Déterminants génétiques et génomiques de la réponse au déficit hydrique chez la tomate (Solanum lycopersicum) et impact sur la qualité des fruits

Elise Albert

Genetic and genomic determinants of response to water deficit in tomato (Solanum lycopersicum) and impact on fruit quality

Elise Albert

ACADEMIE D'AIX-MARSEILLE
UNIVERSITE D'AVIGNON ET DES PAYS DE VAUCLUSE

THESE

présentée pour obtenir le grade de **Docteur en Sciences**
de l'Université d'Avignon et des Pays de Vaucluse

Spécialité : Sciences Agronomiques

Déterminants génétiques et génomiques de la réponse au déficit hydrique chez la tomate (*Solanum lycopersicum*) et impact sur la qualité des fruits

par **Elise ALBERT**

soutenue le 04/01/2017 devant un jury composé de :

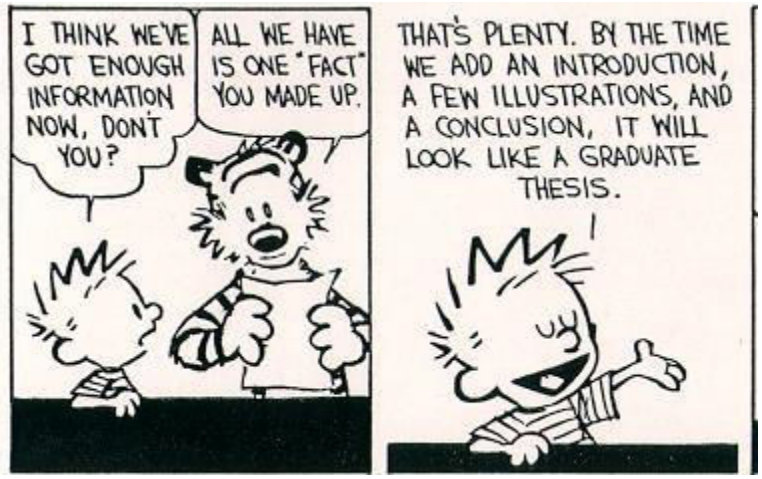
Olivier LOUDET	Directeur de recherche INRA, Versailles (France) (Président du jury)
Antonio MONFORTE	Chercheur CSIC, Valence (Espagne) (Rapporteur)
Christophe ROTHAN	Directeur de recherche INRA, Bordeaux (France) (Rapporteur)
Jean Pierre RENOU	Directeur de recherche INRA, Angers (France) (Examinateur)
Mathilde CAUSSE	Directrice de recherche INRA, Avignon (France) (Directrice de thèse)

Ecole Doctorale :
Sciences et Agrosciences (ED536), Avignon

Laboratoires d'accueil :
INRA-UR1052 «Génétique et Amélioration des Fruits et Légumes», Avignon

Financement : Bourse de thèse INRA & contrat ANR

A Philippe,


I THINK WE'VE GOT ENOUGH INFORMATION NOW, DON'T YOU?
ALL WE HAVE IS ONE "FACT" YOU MADE UP.
THAT'S PLENTY. BY THE TIME WE ADD AN INTRODUCTION, A FEW ILLUSTRATIONS, AND A CONCLUSION, IT WILL LOOK LIKE A GRADUATE THESIS.

Ils ont contribué à la réalisation de cette thèse…

Le département INRA 'Biologie et Amélioration des Plantes' a financé ce projet de thèse et les entreprises **Gautier Semences** et **Vilmorin et Compagnie** ont participé à la mise en place des expérimentations réalisées au cours de ces trois années dans le cadre des projets CTPS TOMSEC et ANR ADAPTOM.

Mathilde Causse, ma directrice de thèse. Elle m'a accueillie au sein de son équipe et soutenue tout au long de cette thèse. Elle m'a donnée la chance de participer à de nombreux workshops et congrès afin de rencontrer les membres de la communauté scientifique *Solanaceae* en France et à l'international. Merci Mathilde pour l'autonomie de travail que tu m'as laissée, la confiance que tu m'as accordée, ta disponibilité (même à l'autre bout du monde et malgré tes nombreuses responsabilités) et ta sérénité à toute épreuve dans les derniers mois de cette thèse ! Merci aussi pour les romans de Craig Johnson que j'ai dévorés !

Olivier Loudet, **Antonio Monforte**, **Jean Pierre Renou** et **Christophe Rothan** m'ont fait l'honneur de constituer mon jury de thèse.

Nadia Bertin, **José Jimenez-Gomez**, **Nicolas Langlade** et **Vincent Segura** ont accepté de faire partie de mon comité de suivi de thèse. Leurs remarques m'ont permis d'envisager mon travail sous de nouveaux angles. Je les remercie pour le temps qu'ils m'ont accordé.

Marie-Laure Martin-Magnette, **Julie Aubert**, **Guillem Rigaill**, **Catherine Bastien**, **Vincent Segura** et **Joël Chadoeuf** m'ont soutenue tout au long de cette thèse pour les aspects statistiques. Merci tout particulièrement à Guillem et Vincent pour leur disponibilité et leur réactivité face à l'avalanche de mes questions.

Jean-Paul Bouchet et **Frédérique Bitton**, m'ont très largement accompagnée pour les aspects bioinformatiques de cette thèse. Un grand MERCI à Jean-Paul que j'ai particulièrement sollicité ces derniers mois. Toujours rigoureux et perfectionniste, ses remarques et ses conseils ont grandement contribué à cette thèse ! J'ai beaucoup apprécié de travailler avec toi !

Christopher Sauvage, co-directeur officieux de cette thèse, il a relu chacun des articles rédigés au cours de ces trois années et ses conseils ont été précieux. Les discussions que j'ai pu avoir avec lui aux pauses café ou autour d'une bière à la *Centrale* m'ont beaucoup enrichie.

Justine Gricourt et **Esther Pelpoir**, deux collaboratrices hors pair pour toutes les expériences menées en serre et au laboratoire aux cours de cette thèse ! Récolter, broyer, peser, extraire, doser et plus encore… des milliers d'échantillons ne leur font pas peur (même durant les jours fériés du mois de Mai…)! Toujours bienveillantes et réconfortantes dans les moments difficiles. J'ai eu énormément de plaisir à travailler avec elles. Je vous remercie pour votre aide et votre amitié aux cours de ces trois années, les fous rires au labo et tous les bons moments passés ensemble… et ceux à venir!

Renaud Duboscq, colocataire de bureau, éleveur de bonzaïs et d'orchidées, fan de football (les Girondins bien sûr), reggae-man, constructeur de maison, papa poule à ses heures et surtout… roi de la biologie moléculaire. Il a contribué à me rendre un peu moins ignare sur les mécanismes et la mise en pratique des aspects moléculaires de cette thèse, m'a initié aux extractions d'ARN et à l'art de la préparation de libraires pour le séquençage. Merci à toi pour tout cela et pour ta patience face à mes nombreuses boulettes.

Sylvain Santoni, **Muriel Latreille**, **Charles Poncet** et **Véronique Gautier.** Merci à eux pour leurs apports techniques dans la création des banques ARN, la réalisation des qPCR microfluidigm et pour leur accueil à Montpellier et à Clermont-Ferrand.

Luisa Bermudez e**t Fernando Carrari** m'ont accueillie à Buenos Aires, fait découvrir leurs activités de recherche à l'INTA (Instituto Nacional de Tecnologia Agropecuaria) et leur ville.

Alain Goujon et les **équipes expérimentales** (**Mara Grumic** et **Christelle Roigt** surtout) pour leurs savoir-faire et compétences dans le suivi des plantes. Merci à **Yolande Carretero** pour la gestion des graines et de l'implantation des expérimentations (et aussi pour ses délicieuses fougasses en début de saison !). Merci à vous sans qui la recherche à l'INRA n'avancerait pas beaucoup. Une petite pensée aussi pour les équipes du **Domaine Margau** et **Julien Bonnefoi** à Agadir, au Maroc !

Les stagiaires, Claire, Romain, Rémi, Margaux, Matthieu et Isidore, pour m'avoir donné du fil à retordre dans la correction de leurs rapports de stage et le suivi de leurs analyses statistiques, mais surtout pour leur aide immense dans la conduite des expérimentations et tous les bons moments passés ensemble.

L'équipe administrative du GAFL, Sébastien, Evelyne JM, Evelyne J, Astrid, Annick et Claudie, pour répondre présent à toutes mes déboires administratifs (inscription à l'université, ordre de mission, réservation d'hôtels, organisation de mes voyages à l'étranger, la liste est longue…) et aussi pour la réception de mes nombreux colis (perso ou non !).

Les thésard(e)s, anciens ou actuel(le)s, **Guillaume B**, **Vincent**, **Elsa**, **Léandro**, **Lucie**, **Gaëtan**, **Mariem, Stéphanie, Anna et Zoé** merci pour les vendredi soir au *Gambrinus* et O'Neills, les matchs de l'euro, les Escapes Games et les randonnées autour d'Avignon.

Les collègues et amis du GAFL, Lilian, René, Hélène, Karine, Rémi, Nasradin, Elodie, Laure, Caroline, Marie-Noëlle, Rebecca, Laura, Carole (Miss Bon Plans), Gisèle, Bénédicte, Véronique S., Henry, Jean-Luc P., Joan, David, Christophe, Thierry, Bruno, Jean, Guillaume R, Martin, Maurice pour les bons moments autour d'un gâteau à la pause-café et lors des repas d'unité..

Les amis sur Avignon et ailleurs, Estanislao, Janejira, Charlotte, Ombeline (Linette), Alice, Aude-Line, Eirini, Pauline M, Steph Chamot, Clément, François, Jean-Charles, Rachel, Eric et la petite Emy… pour m'avoir fait découvrir la vie à Avignon, son festival et ses randonnées, pour les super vacances en Grèce, en Suisse, en Corse et en Argentine … et surtout pour m'avoir permis de couper un peu avec la thèse quand il le fallait.

Enfin un grand MERCI à **ma famille (mum, papa, Rom', Andrée, Mariette, Vinc', Marie-Béatrice, Odile, Corine et tous les autres)** pour leur affection et leur soutien… et à **Benoit** aussi pour m'avoir encouragé et d'avoir été là pour moi…

Merci et toutes mes excuses à ceux que j'aurais pu oublier…

THESE

Déterminants génétiques et génomiques de la réponse au déficit hydrique chez la tomate (*Solanum lycopersicum*) et impact sur la qualité des fruits

PHD THESIS

Genetic and genomic determinants of response to water deficit in tomato (*Solanum lycopersicum*) and impact on fruit quality

Elise ALBERT, 2017

RESUME

A l'échelle du globe, la diminution des ressources en eau est devenue un des principaux facteurs limitants pour les productions agricoles. Jusqu'à présent, les approches génomiques à haut débit conduites chez les espèces modèles ont permis d'identifier des centaines de gènes potentiellement impliqués dans la survie des plantes en conditions de sécheresse, mais très peu ont des effets bénéfiques sur la qualité et le rendement des cultures. Néanmoins, l'application d'un déficit hydrique bien contrôlé peut permettre d'améliorer la qualité des fruits charnus par dilution et/ou accumulation de composés gustatifs majeurs. Dans ce contexte, la première partie du travail de thèse avait pour but de déchiffrer les déterminants génétiques de la réponse au déficit hydrique chez la tomate en explorant les interactions 'génotype x niveau d'irrigation' (G x I) et 'QTL x niveau d'irrigation' (QTL x I) dans deux populations. La première population consistait en un ensemble de lignées recombinantes (RIL) issues du croisement entre deux accessions cultivées, tandis que la seconde était composée de diverses accessions à petits fruits principalement originaires d'Amérique du Sud. Les plantes ont été phénotypées pour un ensemble de caractères agronomiques (vigueur des plantes et qualité des fruits) et génotypées pour des milliers de SNP. Les données ont été analysées en utilisant les méthodologies de la cartographie de liaison et d'association, permettant l'identification de QTL et gènes candidats putatifs pour la réponse de la tomate au déficit hydrique. La deuxième partie du travail de thèse avait pour objectif d'explorer la régulation des gènes dans les fruits et les feuilles de tomates en condition de déficit hydrique. Dans ce but, des données de séquençage du transcriptome ont été recueillies sur les deux génotypes parentaux de la population RIL et leur hybride F1. Les données ont été analysées pour identifier les gènes et les allèles exprimés de manière différentielle. Puis, l'expression de 200 gènes a été mesurée dans les fruits et les feuilles de l'ensemble des lignées de la population RIL par qPCR micro-fluidique à haut débit. Des eQTL et des interactions 'eQTL x niveau d'irrigation' ont été identifiés pour ces gènes par cartographie de liaison. Les colocalisations entre les QTL phénotypiques et les QTL d'expression ont été analysées. Les connaissances produites au cours de cette thèse contribuent à une meilleure compréhension des interactions des plantes de tomate avec leur environnement et fournissent des bases pour l'amélioration de la qualité des fruits en conditions d'irrigation limitée.

ABSTRACT

Water scarcity will constitute a crucial constraint for agricultural productivity in a near future. High throughput approaches in model species have identified hundreds of genes potentially involved in survival under drought conditions, but very few having beneficial effects on quality and yield in crops plants. Nonetheless, controlled water deficits may improve fleshy fruit quality through weaker dilution and/or accumulation of nutritional compounds. In this context, the first part of the PhD was aimed at deciphering the genetic determinants of the phenotypic response to water deficit in tomato by exploring the genotype by watering regime (G x W) and QTL by watering regime (QTL x W) interactions in two populations. The first population consisted in recombinant inbreed lines (RIL) from a cross between two cultivated accessions and the second was composed of diverse small fruit tomato accessions mostly native from South America. Plants were phenotyped for major plant and fruit quality traits and genotyped for thousands of SNP. Data were analyzed within the linkage and association mapping frameworks allowing the identification of QTLs and putative candidate genes for response to water deficit in tomato. The second part of the PhD had the objective to explore gene regulation in green fruit and leaves of tomato plants stressed by water deficit. For this purpose, RNA-Seq data were collected on the two parental genotypes of the RIL population and their F1 hybrid. Data were analyzed to identify differentially expressed genes and allele specific expression (ASE). Then, the expression of 200 genes was measured in leaves and fruits of the whole RIL population by high throughput microfluidic qPCR. eQTLs and eQTL by watering regime interactions were mapped for those genes using linkage mapping. Colocalisations with the phenotypic QTLs were analyzed. The knowledge produced during this PhD will contribute to a better understanding of the tomato plant interaction with their environment and provide bases for improvement of fruit quality under limited water supply.

Communications scientifiques

Communications dans des revues scientifiques

Pascual*, L., **Albert*, E.**, Sauvage, C., Duangjit, J., Bouchet, J. P., Bitton, F., ... & Bruguier, L. (2016). Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. Plant Science, 242, 120-130. [* Co-first author] **(Displayed in Appendix 3)**

Constantinescu, D., Memmah, M-M, Vercambre, G., Genard, G., Baldazzi, V., Causse, M., **Albert, E.,** Brunel, B., Valsesia, P., Bertin, N. (2016) Model-based analysis of the genetic variability in tomato fruit growth under contrasted water conditions. Frontiers in Plant Science, 7, 1841. **(Displayed in Appendix 1)**

Albert, E., Gricourt, J., Bertin, N., Bonnefoi, J., Pateyron, S., Tamby, J. P., ... & Causse, M. (2016). Genotype by watering regime interaction in cultivated tomato: lessons from linkage mapping and gene expression. Theoretical and Applied Genetics, 129(2), 395 - 418. **(Displayed in Chapter 3)**

Albert, E., Segura, V., Gricourt, J., Bonnefoi, J., Derivot, L., Causse, M. (2016). Association mapping reveals the genetic architecture of tomato response to water deficit: focus on major fruit quality traits. Journal of Experimental Botany, 67(22), 6413 - 6430. **(Displayed in Chapter 4)**

Chapitre d'ouvrage

Causse, M., **Albert, E.**, Sauvage, C. (2016) Developing tomato varieties with improved flavor. Book chapter XIII: Achieving sustainable tomato cultivation. Ed A Mattoo and AK Handa. **(Accepted for publication, displayed in Appendix 2)**

Communications dans des congrès internationaux

Albert E., Segura V., Gricourt J., Novaretti R., Carretero Y., Pelpoir E., Bonnefoi J., Derivot L., Causse M. (2016) GWAS of tomato response to water deficit: focus on major fruit quality traits. The 13th Solanaceae Conference, Davis, California, USA. **(Poster)**

Albert, E., Carretero, Y., Gricourt, J., Pelpoir, E., Novaretti, R., Duffes, C., Bonnefoi, J., Bertin, N., Pateyron, S., Tamby, J.P., Causse, M., (2015) Genotype by watering regime interaction in cultivated tomato: from phenotypes to candidate genes. The 12th Solanaceae Conference, Bordeaux, France. **(Oral communication)**

Albert, E., Carretero, Y., Gricourt, J., Duboscq, R., Pelpoir, E., Novaretti, R., Duffes, C., Bonnefoi, J., Pateyron, S., Tamby, J.P., Bitton, F., Sauvage, C., Causse M. (2015) Genotype by watering regime interaction in cultivated tomato. Innovation in Integrated & Organic Horticulture Symposium, Avignon, France. **(Oral communication)**

Albert, E., Carretero, Y., Gricourt, J., Pelpoir, E., Novaretti, R., Duffes, C., Bonnefoi, J., Causse M. (2015) Genetic and genomic control of response to water deficit in cultivated tomato. COST Action FA1106, "QualityFruit" Workshop, Verona, Italy. **(Oral communication)**

Albert, E., Duffes, C., Bonnefoi, J., Gricourt, J., Carretero, Y., Pelpoir, E., Causse, M. (2015) Genetic determinants of response to water deficit in cultivated tomato fruits – QTL x E analysis. TGAC training, Statistics for Ecology, Genetics and Genomics using R, Norwich, United Kingdom. **(Poster)**

Table of contents

Preamble

In a context of global climate change and natural resources scarcity, agricultural systems must evolve towards more sustainable and less intensive forms. In particular, water scarcity will constitute crucial threat in the coming years. Crop productions which are currently consuming up to 80% of the worldwide water resource through irrigation have to limit their water consumption, while maintaining a reasonable productivity to feed the growing world population. Increasing crop water use efficiency is among the solutions to solve this dilemma.

Tomato (*Solanum lycopersicum*) is the second vegetable consumed in the world after potato and its production consistently increased through the world in the last decade. Over the XXth century, tomato breeders have focused on improving the species for yield and yield stability, adaptation to various growth conditions, disease resistances, conservation and appearance of the fruit (diversification of color and shape). The sensory quality was mostly ignored, causing consumer dissatisfaction in the early 90's. In this species, limited water supply can have a favorable impact on fruit quality, on condition to find the right balance to minimize yield losses. Thus, deficit irrigation strategies offer the opportunity to address market expectations of tastier fruits and simultaneously reduce non-beneficial water consumption in tomato production. The characterization of the natural phenotypic variability of tomato response to water deficit and the underlying genetic determinants is greatly needed to lay the necessary basis for fruit quality improvement under water limitation.

This thesis aims at characterizing the genetic and genomic variability of tomato response to water deficit and its impact on fruit quality, through the integration of phenotypic, genomic and transcriptomic data. The outcomes provide basis for understanding tomato plant response to water limitation and for fleshy fruit quality improvement under deficit irrigation while limiting yield loss.

The manuscript consists of six chapters:

Chapter 1 presents a brief literature review, on the impact of water deficit on plant at the physiological and molecular levels. Then, in a second part, the methodologies and tools available to explore the 'genotype by watering regime' and 'QTL/gene by watering regime' interactions in quantitative genetics are presented. Finally, a particular attention is given to

tomato, the impact of water deficit on tomato fruit quality and the state of the art regarding the identification of the underlying genetic determinants.

Chapter 2 summarizes the plant materials and the methods used in the thesis, which will be more detailed in each of the following chapters.

Chapter 3 is in the form of an article published in *Theoretical and Applied Genetics*. The article describes the 'genotype by watering regime' interactions observed in a population of tomato recombinant inbreed lines grown under two irrigation conditions, at the phenotypic and genotypic levels.

Chapter 4 is in the form of an article published in *Journal of Experimental Botany*. The article describes association mapping for fruit quality traits in a collection of small fruit tomatoes grown under contrasted watering conditions. Major QTLs are dissected using expression data, exonic variants and analysis of the gene functions.

Chapter 5 is in the form of an article draft describing the identification of differentially expressed genes, allele specific expression and eQTLs in two contrasted genotypes, a F1 hybrid and a recombinant inbreed line population, grown under two contrasted watering conditions.

Chapter 6 presents a synthesis of the main results, conclusions and perspectives of this thesis.

CHAPTER 1

CHAPTER 1: Bibliographic synthesis: effects of water deficit on plant and fruit quality, identification of the genetic determinants

This chapter corresponds to a short literature review on the effects of water deficit on plant at the physiological and molecular levels. Then, methodologies and tools to explore the 'genotype by watering regime' and 'QTL/gene by watering regime' interactions in quantitative genetics are presented. Finally, a particular attention is given to tomato, the impact of water deficit on tomato fruit quality and the state of the art regarding the identification of the underlying genetic determinants.

---

1. Plant responses to water deficit

In agronomy, water deficit corresponds to the inadequacy of water availability (precipitation and soil storage), in quantity and distribution over time, responsible for a limitation of the expression of the full genetic potential of a plant (Mitra, 2001). It constitutes a major constraint for crop productivity, responsible for large yield losses through the world (Jury and Vaux, 2005; Rost *et al.* , 2008). To maintain sustainable production in the global water scarcity context, improvement of major crops is necessary and requires the knowledge of the plant response mechanisms and their genetic control and variability.

1.1 Plant strategies in response to water deficit

Plants are sessile organisms which cannot migrate when challenged by environmental fluctuations like water limitation. Thus, they have developed mechanisms to adapt to their environmental conditions over centuries. Three evolutionary strategies are commonly distinguished when studying plant response to low water conditions, although they are not exhaustive and plants have developed often more than one strategy at once (Levitt, 1972; Ludlow, 1989):

The **drought escape** strategy gathers mechanisms allowing plants to complete their cycles before soil and plant water deficit happens. Mechanisms promoting developmental plasticity, rapid phenological development and remobilization of pre-anthesis assimilates to reproductive organs allow plant to avoid drought.

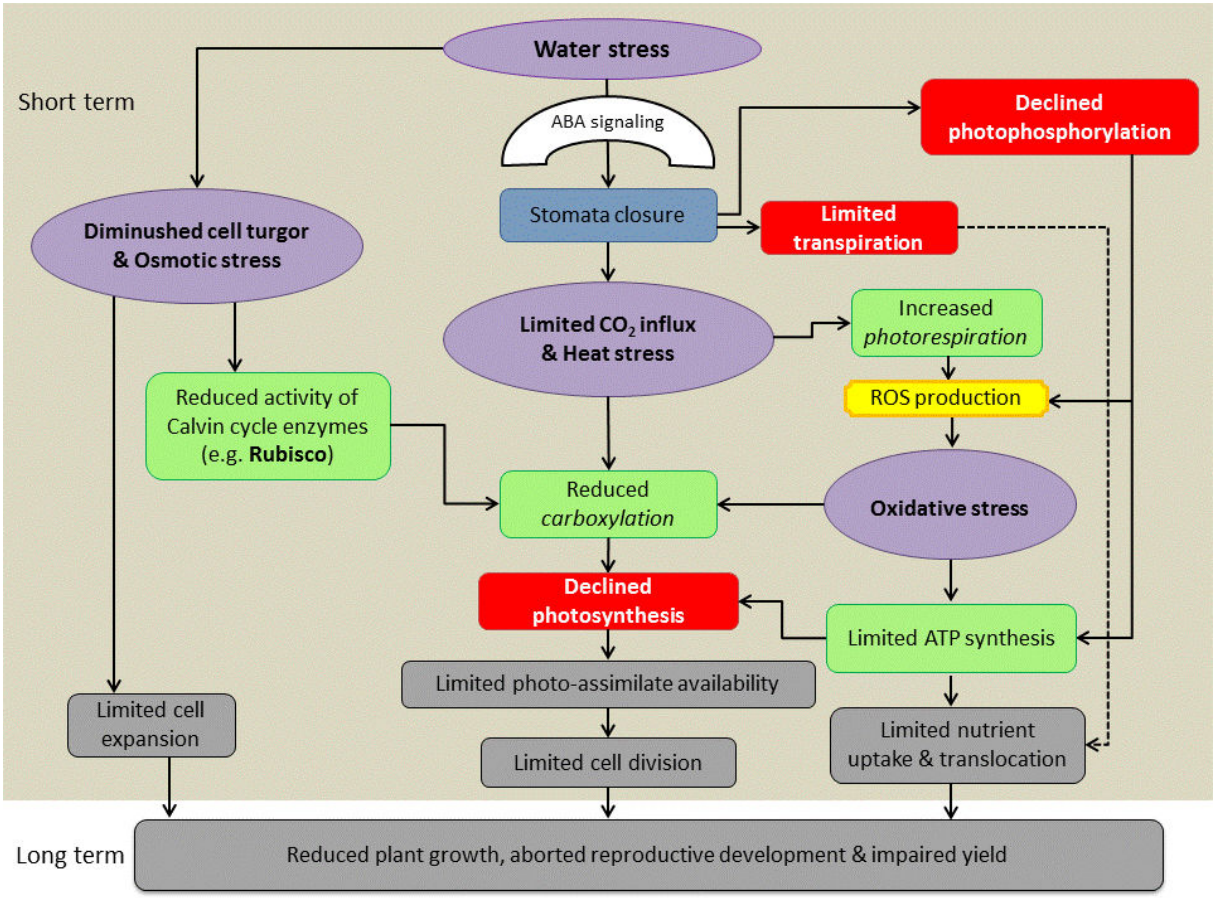


Figure 1: Main effects of water deficit on plant physiology and consequences on plant growth and development. The stomata closure and the secondary stress associated to water deficit (namely oxidative, osmotic and heat stress) affect the whole plant physiology, including key processes such as photosynthesis, photophosphorylation and transpiration. The short term consequences are the limitation of cell division and expansion, the modification of carbohydrate partitioning and the impairment of nutrient uptake and translocation. The long term consequences are a reduction in plant growth, the abortion of the reproductive development and yield impairment. *(Adapted from Farooq et al. , 2009)*

The **drought avoidance** strategy gathers mechanisms allowing plants to maintain relatively high tissue water potential despite water limitation. Mechanisms for improving water uptake, water storage in cell and limit water loss confer drought avoidance. This includes, among other mechanisms, increased investment in roots, limited leaf area through shedding of older leaves, leaf rolling and step leaf angles, development of a dense fruit and leaf cuticles and stomata closing.

The **drought tolerance** strategy gathers mechanisms allowing plants to face water deficit with low tissue water potential. The responses of plants to tissue water deficit determine their level of drought tolerance. Development of an efficient antioxidant system and maintenance of turgor through osmotic adjustment are among the mechanisms allowing drought tolerance.

The various physiological and molecular mechanisms of plant response to water deficit involve numerous interrelated biological pathways and transitions in gene expression (Farooq *et al.* , 2009). The setup of these mechanisms by the plant depends upon the plant species, the genotype sensitivity, the severity, duration and timing of the stress, as well as the interaction with other biotic and abiotic constraints (Plaut, 2003; Tardieu, 2012).

1.2 Physiological mechanisms involved in response to water deficit & consequences on plant growth and development

The main plant physiological mechanisms involved in response to water deficit and consequences on plant growth and development are synthesized below and illustrated in **Figure 1**.

1.2.1 Stomatal closure, maintained water status and impaired photosynthesis

Stomatal closure takes place few minutes after plant perception of a water limitation. Stomata are microscopic pores in the epidermis of the aerial parts of plants, composed of two guard cells, the turgor state of which controls the pore size **(Figure 2)**. They allow plants to optimize gas exchanges (CO_2, O_2 and water vapor) with the atmosphere depending on the environmental conditions (reviewed in Casson and Hetherington, 2010).

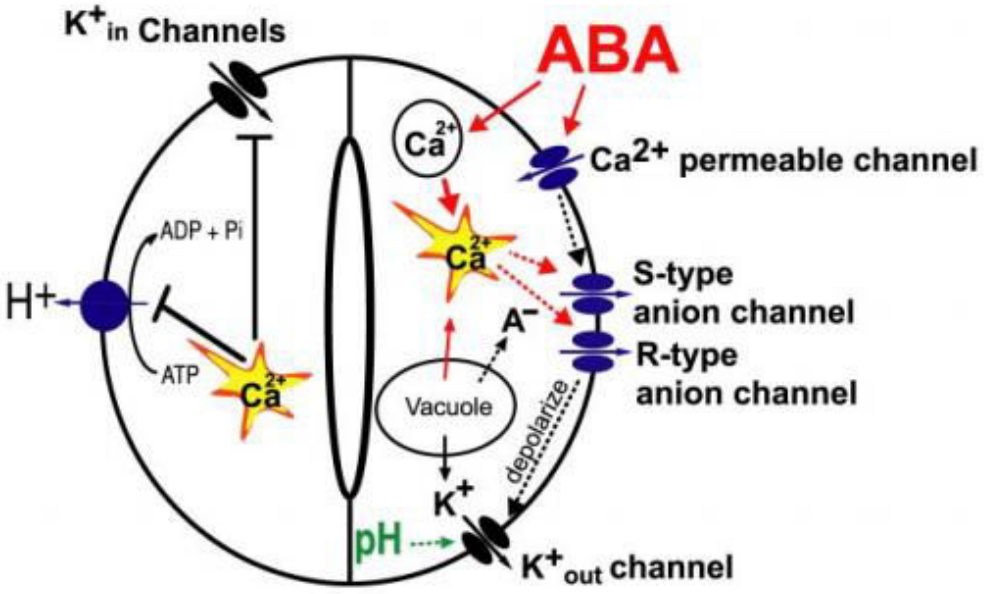


Figure 2: Overview of the ion channel functions in ABA signaling pathway and stomatal closure. The right cell of the stomata shows ion channels and regulators that mediate ABA-induced stomatal closure. The left cell shows the parallel effects of ABA-induced cytoplasmic Ca^{2+} increase that inhibit stomatal opening mechanisms. S-type anion channel; slow channel, R-type anion channel; rapid channel. *(Adapted from Schroeder et al. , 2001)*

In water deficit condition, the reduction in water loss by transpiration associated to stomata closure enables plants to preserve their water balance. However, this positive effect comes at the price of a reduction in the inflow of CO_2 into the leaves and an increase of leaf temperature because the excess of solar radiation evacuated through open stomata under optimal watering conditions is no longer dissipated. In condition of low CO_2/O_2 ratio and leaf heating, photo-respiration (oxygenase activity of the Rubsico enzyme, *ribulose-1, 5-bisphosphate oxygenase*) is favored at the expense of photosynthesis (carboxylation activity of the Rubsico enzyme, *ribulose-1, 5-bisphosphate carboxylase*) resulting in a decrease of carbohydrate synthesis (Zhou *et al.* , 2007, Bota *et al.* , 2004).

Physiological processes that regulate stomatal function have been intensively studied. They involve complex interactions of extrinsic and intrinsic factors. The plant hormone *abscisic acid* (ABA) is intimately involved in regulating the opening and closing of stomata in response to changes in cellular water status, although it is still not clear how cellular water deficit induces ABA biosynthesis. The signal could be an impaired cellular pressure, membrane modifications, solute concentration or cell wall tension (Wasilewska *et al.* , 2008).

In the stomata guard cell, ABA is responsible for the increase in the cytosolic Ca^{2+} content, which activates membrane anion channels and vacuolar K^+ channels. The anion channels allow the release of anions out of the cells (mainly chloride ion Cl^- and nitrate NO_3^-) whereas the vacuolar K^+ channels release the vacuolar K^+ into the cytosol. This causes a plasma membrane depolarization which deactivates inward-rectifying K^+ (K^+ *in*) channels and activates outward-rectifying K^+ (K^+ *out*) channels, resulting in K^+ efflux out of the guard cells. The efflux of anions and K^+ from guard cells contributes to loss of guard cell turgor, which leads to stomatal closure. In addition, ABA directly inhibits the H^+*ATPases*, the K^+ *in* channels and the import of Cl^-, malate and NO_3^- into the cytoplasm, which are the mechanisms involved in the stomata opening (reviewed in Schroeder *et al.* , 2001; Kim *et al.* , 2010) **(Figure 2)**.

Although ABA is the best known hormone regulating stomata closure, other phytohormones such as *jasmonic acid*, *brassinosteroids*, *cytokinins*, *auxin* or *ethylene* are recognized to be involved in this mechanism through a complex interplay with ABA signaling pathway that still

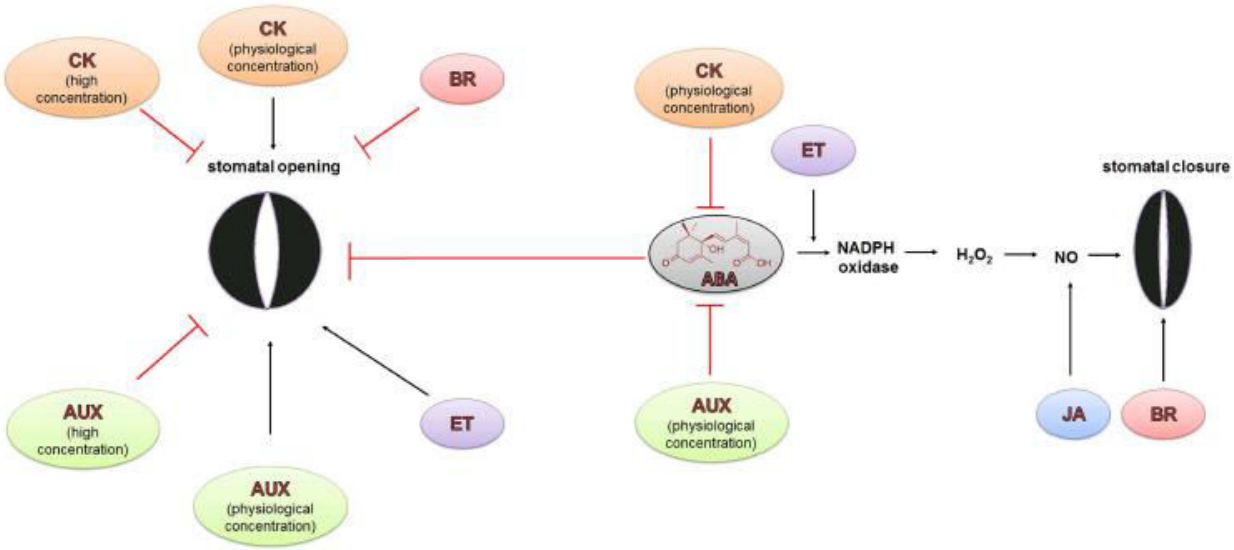


Figure 3: Hormonal cross talk in the regulation of stomatal closure and opening in response to water deficit. The regulation of stomatal opening and closure is not only regulated by ABA, but also by other phytohormones. *Jasmonate* (JA) and *brassinosteroids* (BR) induce stomatal closure and inhibit stomatal opening under water deficit conditions, whereas the role of other hormones is ambiguous. *Cytokinins* (CK) and *auxin* (AUX) in low physiological concentrations promote stomatal opening while in high concentrations, they are able to inhibit this process. The role of *ethylene* (ET) is ambiguous. It can stimulate both the closing and opening of the stomata. *(Adapted from Daszkowska-Golec and Szarejko, 2013)*

needs to be deciphered. *Jasmonate* and *brassinosteroids* were shown to induce stomata closure and inhibit stomata opening under water deficit conditions, whereas the roles of *auxin*, *ethylene* and *cytokinins* remain more ambiguous (Daszkowska-Golec and Szarejko, 2013) **(Figure 3)**.

Non-stomatal limitations may also be involved in the decreased photosynthesis rate observed under severe water limitations (Flexas and Medrano, 2002). They mainly correspond to the inhibition and/or decreased synthesis of the enzymes involved in carbon assimilation (e.i. Calvin cycle) and photophosphorylation in low moisture conditions (Du *et al.* , 1996; Bota *et al.* , 2004) and to lower diffusion of CO_2 across the leaf mesophyll (Flexas and Medrano, 2002). There is a controversy in the scientific community whether water deficit limits photosynthesis mainly through stomatal closure or by metabolic impairments (Cornic, 2000; Bota *et al.* , 2004). First attempts to answer this question showed that the impairment of the Calvin cycle does not limit photosynthesis until drought is very severe. Moreover, the relative cellular water content at which the photosynthetic metabolism is impaired would be strongly species-dependent (Bota *et al.* , 2004). In particular, CAM (*crassulacean acid metabolism*) and C4 plants have developed mechanisms to optimize their photosynthetic machinery under low water availability compared with the C3 plants (reviewed in Szarek and Ting, 1975).

1.2.2 Set up of an efficient defense against oxidative damages

Under water deficit, plant exposure to light intensities that exceed their capacity of CO_2 assimilation under reduced stomatal conductance results in an excessive production of reactive oxygen species (ROS). Indeed, in conditions of low carboxylation, the photosystem II (PSII) activity is reduced to match the available carbon substrate. The excess of electrons from the photophosphorylation electron chain is transferred to dioxygen (O_2) at photosystem I (PSI) in the Mehler reaction, leading to the production of *superoxide anions* ($O2^{\bullet -}$), *hydrogen peroxide* (H_2O_2) and other ROS within the chloroplasts. Besides, the enhanced oxygenase activity of the Rubisco enzyme under water deficit results in H_2O_2 and other ROS production within the peroxisomes **(Figure 4)** (Blokhina *et al.* , 2003; Apel and Hirt, 2004; Unal, 2013).

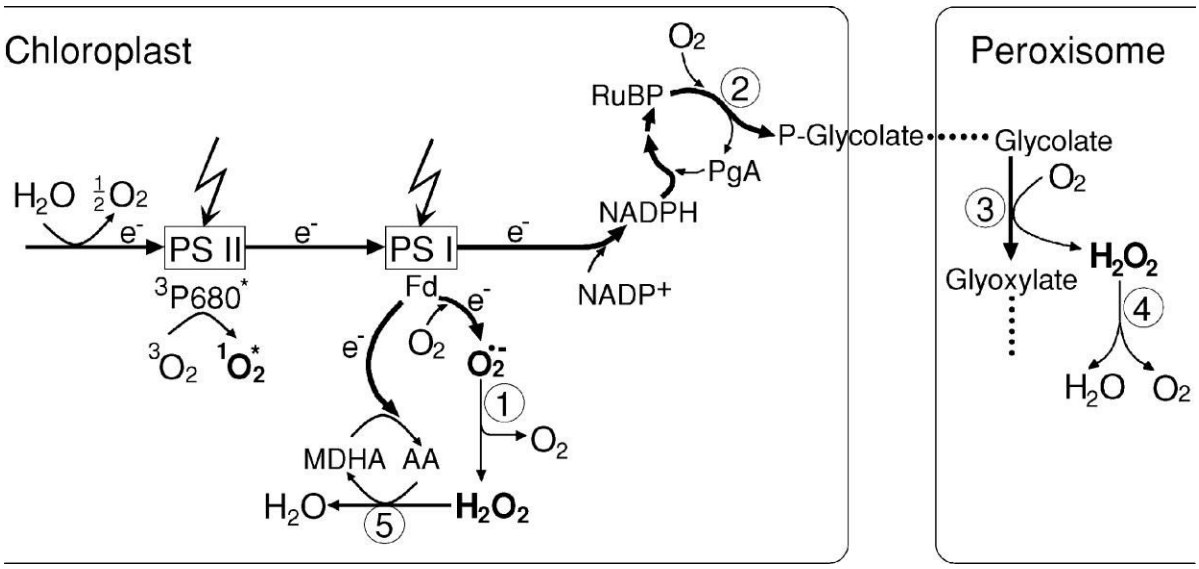


Figure 4: The principal features of photosynthetic electron transport under excess of light energy that lead to the production of reactive oxygen species (ROS) in chloroplasts and peroxisomes. Two electron sinks can be used to alleviate the negative consequences of over-reduction of the photosynthetic electron chain: (a) the reduction of oxygen by PSI that generates superoxide and H_2O_2, and (b) the Rubisco oxygenase reaction and the photorespiratory pathway that lead to H_2O_2 generation within the peroxisome. Under light stress, increasing amounts of singlet oxygen are produced within PSII. Bold arrows show the main routes of electron transport. Key enzymes are shown in encircled numbers: 1) *superoxide dismutase*, 2) Rubisco, 3) *glycolate oxidase*, 4) *catalase*, and 5) ascorbate *peroxidase*. Abbreviation: RuBP; *ribulose-1, 5-bisphosphate.*
(Adapted from Apel and Hirt, 2004)

ROS can generate major oxidative damages in plants and their deleterious effects include DNA degradation, amino acid and protein oxidation and lipid peroxidation. They destabilize cell membrane and enzymes, perturbing cellular functions. Thus, plant survival under water deficit depends on the development of an efficient antioxidant system, with both enzymatic and non-enzymatic components. Enzymatic components are mainly *superoxide dismutases* (SOD), *catalases* (CAT), *peroxidases* and *ascorbate peroxidases* (APX), *glutathione peroxidase* (GPX), *peroxide dismutases*, *polyphenol oxidases*, *laccases*, *anthocyanidin reductase*, *anthocyanidin synthase*, *mono–dehydroascorbate reductases* (MDAR), *dehydroascorbate reductases* (DHAR) and *glutathione reductases* (GR). Non-enzymatic plant antioxidants are mainly amino acid like scavengers, such as *cysteine*, *glutathione* (GSH), *ascorbic acid* (AsA), pigments and polyphenols, such as *carotenoids* and *anthocyanins* (reviewed in Mittler, 2002, 2004; Blokhina *et al.* , 2003; Apel and Hirt, 2004).

Figure 5 displays the water-water cycle, the ascorbate-glutathione cycle, the glutathione peroxidase cycle and the catalase action, which are the main pathways involved in ROS scavenging in plants in response to water deficit. The pathways involving polyphenols are not presented because the antioxidant function of flavonoids is still a matter of debate and need more investigation (Hernández *et al.* , 2009).

1.2.3 Osmotic adjustment to maintain cell turgor and protect the cellular machinery

To limit water loss, plants are able to actively accumulate organic compounds called osmoprotectants. This phenomenon corresponds to the osmotic adjustment. Osmoprotectant compounds are highly soluble molecules, usually neutral at physiological pH, with low molecular weight and non-toxic even in high cytosolic concentrations. They include a wide range of molecules, which depends on the nature of the water deficit, the plant species and the genotypes. Proteins and amino-acids (e.g. *proline*, *aspartic acid* and *glutamic acid*), quaternary amines (e.g. *glycine-betaine* and *alanine–betaine*), polyols and sugars (e.g. *D-pinitol*, *mannitol*, *sorbitol, fructans, sucrose* and *trehalose*), organic acids (e.g. *malic* and *citric acids*), hydrophilic proteins (e.g. *late embryogenesis abundant* LEA, *heat shock proteins* and others *chaperon proteins*) and ions (e.g. calcium, potassium, and chloride ions) are among the plant osmoprotectants.

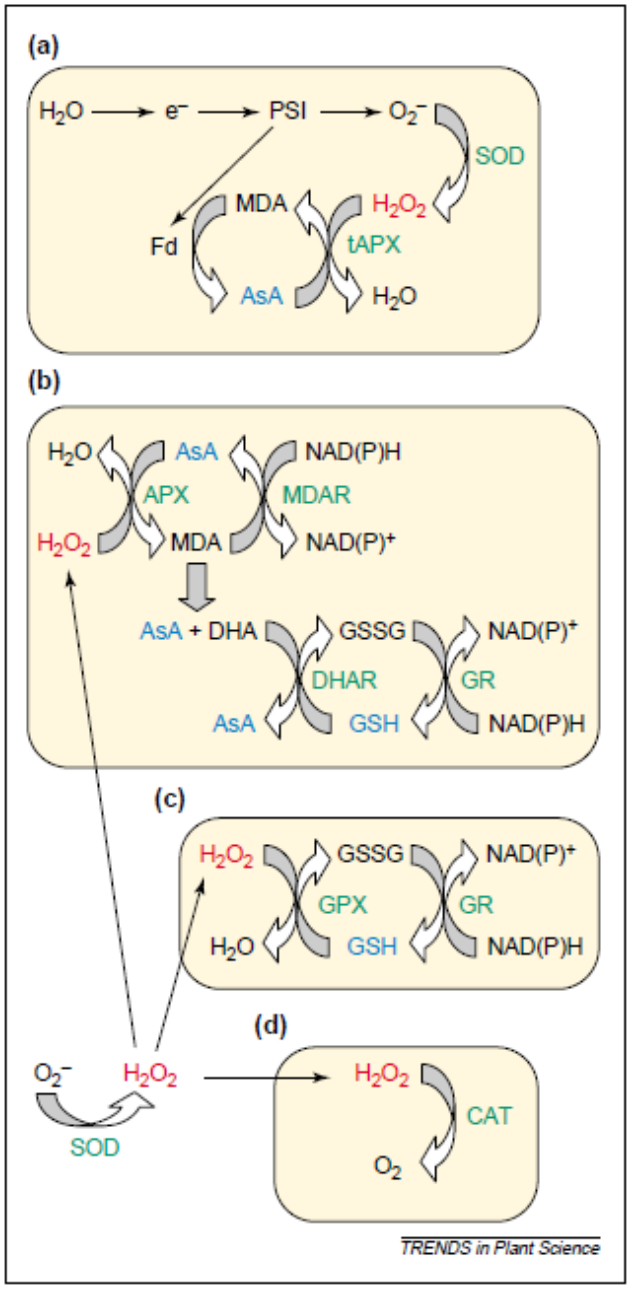


Figure 5. Pathways for reactive oxygen species (ROS) scavenging in plants. (a) The water–water cycle. (b) The *ascorbate–glutathione* cycle. (c) The *glutathione peroxidas*e (GPX) cycle. (d) Catalase (CAT). *Superoxide dismutase* (SOD) acts as the first line of defense converting O_2 into H_2O_2. *Ascorbate peroxidases* (APX), GPX and CAT then detoxify H_2O_2. In contrast to CAT (d), APX and GPX require an ascorbate (AsA) and/or a *glutathione* (GSH) regenerating cycle (a–c). This cycle uses electrons directly from the photosynthetic apparatus (a) or NAD(P)H (b,c) as reducing power. ROS are indicated in red, antioxidants in blue and ROS-scavenging enzymes in green. Abbreviations: DHA, *dehydroascorbate*; DHAR, *DHA reductase*; Fd, *ferredoxin*; GR, *glutathione reductase*; GSSG, *oxidized glutathion*e; MDA, *monodehydroascorbate*; MDAR, *MDA reductase*; PSI, photosystem I; tAPX, thylakoid-bound APX. *(Adapted from Mittler, 2002)*

Osmoprotectant compounds allow plant cell to lower their osmotic potential, thus maintaining water absorption and cell turgor under water deficit conditions. They also protect the proteins, the cell membranes and the metabolic machinery against dehydration, by interacting with water molecules and preventing adverse molecular interactions (reviewed in Zhang *et al.* , 1999; Sanders and Arndt, 2012).

1.2.4 Consequences of water deficit on plant growth and development

Because water deficit affects the main plant physiological processes, namely photosynthesis, transpiration and photophosphorylation **(Figure 1)**, it has major consequences on plant growth and development. The most frequent consequences are presented below.

- **Decreased plant growth through impaired cell elongation and cell division**

In most of the species, water deficit induces a slowdown of the aerial growth because of the impairment of cell elongation and cell division. Cell growth is inhibited mainly because of the interruption of the water flow from the xylem to the elongating cells and the modification of the physicochemical properties of the cell walls which become more rigid, whereas cell division impairments results mainly from the decreased photo-assimilation and carbohydrates availability for cell mitosis (Nonami, 1998; Proseus *et al.* , 1999; Tardieu *et al.* , 2000). When the water deficit is short, cell growth can recover to a large extent upon rehydration. In contrast, the reduction of cell division in meristems may have irreversible consequences and/or result in developmental delays according to the period of the plant development affected by the water limitation, flowering being one of the most sensitive phases (Alves and Setter, 2004).

More than a simple passive consequence of water deficit, the reduced plant growth observed under water limitation constitutes a controlled plant mechanism to limit its evaporative surface when facing long periods of water limitation (Muller *et al.* , 2011).

- **Disturbed nutrient uptake, translocation and metabolism**

Water deficit reduces the availability, uptake, translocation and metabolism of the inorganic nutrients essential for plant nutrition. The reduced uptake results mainly from the reduction of the transpiration water flow through the plant and from interference of nutrient uptake and unloading mechanisms. The reduced translocation and metabolism results mainly from

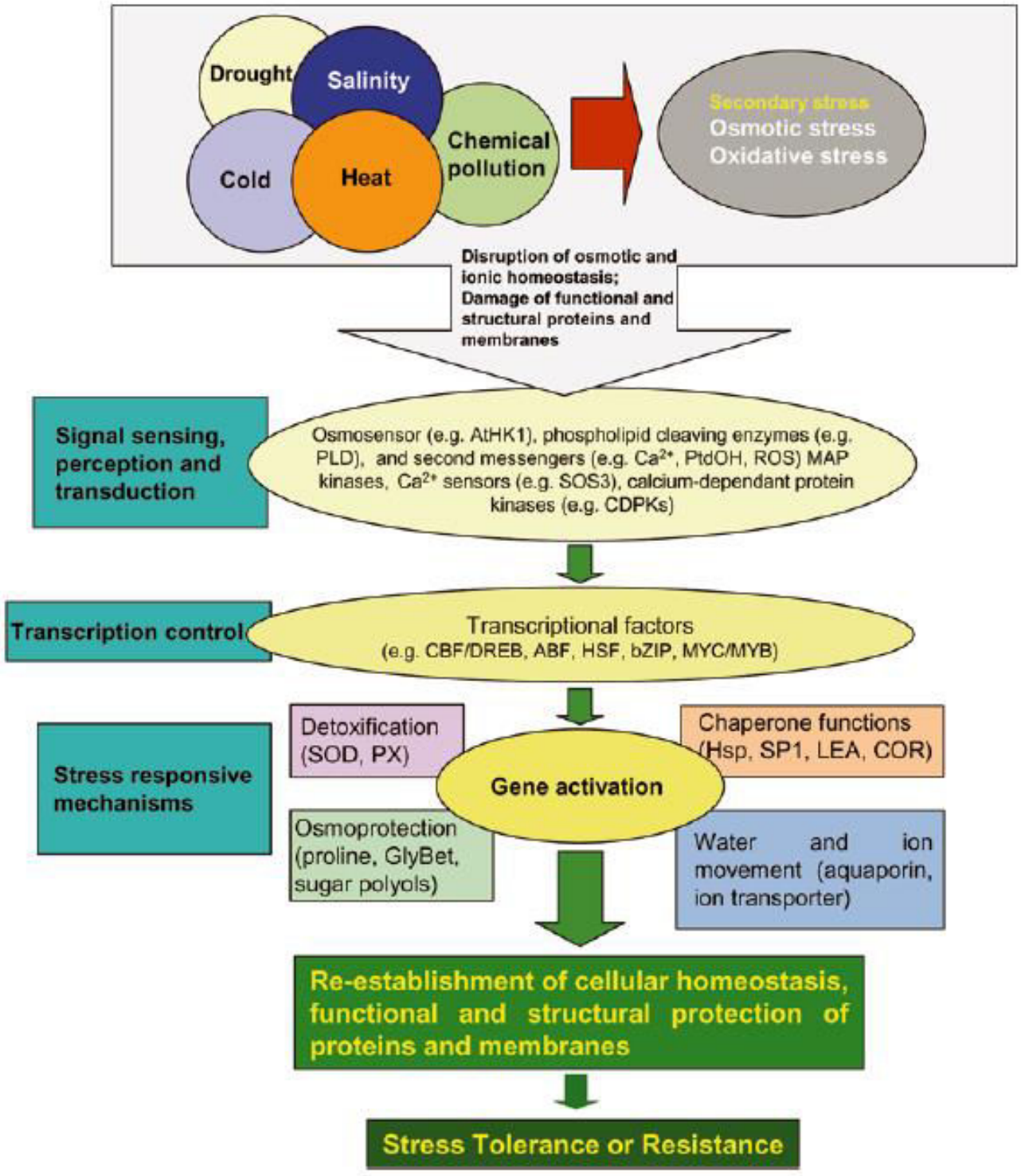


Figure 6: Plant molecular responses to abiotic stress. Primary stresses, such as water deficit, salinity, cold, heat and chemical pollution are often interconnected, and cause cellular damage and secondary stresses, such as osmotic and oxidative stress. The initial stress signals trigger the downstream signaling process and transcription controls which activate stress-responsive mechanisms to re-establish homeostasis and protect and repair damaged proteins and membranes. *(Adapted from Wang et al. , 2003)*

a limited availability of energy for inorganic nutrients assimilation and conversion into organic nutrients to be used for plant growth and development. The assimilation of the different inorganic nutrients is not affected in the same way (Garg, 2003; Peuke and Rennenberg, 2004; Hu and Schmidhalter, 2005).

1.3 Molecular mechanisms involved in response to water deficit

The main molecular plant responses to water deficit are summarized in **Figure 6** and described below. Three major categories of molecular mechanisms are distinguished: (1) those that are involved in signal perception and transduction, (2) those that are involved in transcriptional control and (3) those that are involved in defense mechanisms, including osmotic adjustment, detoxification, protein and membrane protection and water and ion uptake.

1.3.1 Signal perception

First, to develop the appropriate physiological and molecular responses, plants must perceive the water limitation through specific receptors. The nature of these receptors is still largely under debate.

- **AHK1/ATHK1 osmosensor**

Reverse genetic studies have suggested the *Arabidopsis thaliana* plasma membrane *AHK1/AtHK1* as an osmosensor that detects changes in osmotic potential inside the cell and initiates downstream responses in yeast and presumably in plant (Urao *et al.* , 1999; Tran *et al.* , 2007; Wohlbach *et al.* , 2008). Two *high-affinity* K^+ *transporter* (*HKT1* homologues) recently cloned in *Eucalyptus calmaldulensis* may play a similar role to the *Arabidopsis AHK1/ATHK1* (Liu *et al.* , 2001).

- **ABA receptors**

As presented above, ABA plays a crucial role in the initiation of water deficit signaling through the regulating of stomata closure. The ABA biosynthesis and de/conjugation pathways in the vascular parenchyma cell involve several enzymes among which the *zeaxantine epoxidase* (ZEP), the *9-cis epoxycarotenoid dioxygenase* (NCED), the *short-chain alcohol dehydrogenase* (ABA2) the *ABA-aldehyde oxidase* (AAO) and the *ABA*

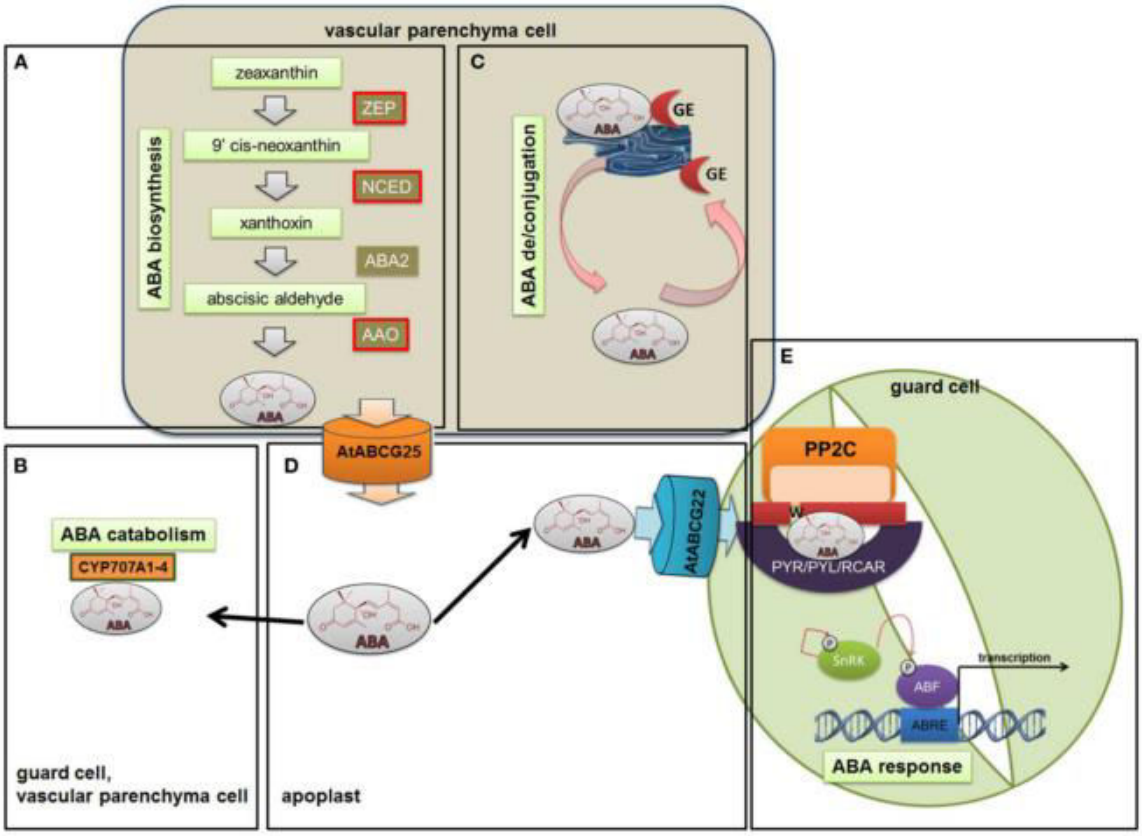


Figure 7: Abscisic acid biosynthesis, catabolism, deconjugation, transport, and signaling. ABA biosynthesis (A) is mainly induced by upregulating NCED3, ZEP, and AAO genes. At the same time as the biosynthesis of ABA is induced, the catabolism (B) that is performed by CYP707A1-4 is inhibited. The balance between active and inactive ABA in the cell is achieved not only by the regulation of biosynthesis and catabolism but also by ABA conjugation and deconjugation. The most widespread conjugate is the ABA glucosyl ester (ABA-GE), which is catalyzed by ABA glucosyltransferase (C). ABA delivery to the guard cells via ABCG transporters such as AGCG22 (D) promotes a cascade of reactions. The early ABA signaling involves ABA receptors (PYR/PYL/RCAR proteins), PP2Cs, and SnRKs (E). After binding ABA to the receptor, the negative regulatory action of PP2Cs is inhibited and SnRKs are able to phosphorylate and activate downstream targets in order to transduce the ABA signal. *(Adapted from Daszkowska-Golec and Szarejko, 2013)*

glucosyltransferase **(Figure 7 A and C)**. The genes coding for these enzymes were shown to be significantly upregulated in response to water deficit (Qin and Zeevaart, 1999, 2002; Thompson *et al.* , 2000*a*; Iuchi *et al.* , 2001; Ye *et al.* , 2011). On the opposite the catabolism of ABA, which is performed by *ABA 8'-hydroxylases* (cytochrome P450 CYP707A family) was shown to be inhibited under water deficit (Kushiro *et al.* , 2004; Umezawa *et al.* , 2006*b*) **(Figure 7 B)**.

Recently the core signalosome of ABA signaling, including ABA transporters, ABA receptors, PP2C phosphatases and protein kinases, was established (reviewed in Umezawa *et al.* , 2010 and Weiner *et al.* , 2010). *ABCG transporters,* such as the *ATP-binding cassettes* G22, G25 and G40 identified in *Arabidopsis thaliana* (Kang *et al.* , 2010; Kuromori and Shinozaki, 2010; Kuromori *et al.* , 2011), are responsible for delivering the ABA produced in vascular parenchyma cells to the stomata guard cells **(Figure 7 D).** Then, the perception of the ABA by ABA receptors induces the inactivation of the PP2C phosphatases (such as ABI1 and ABI2 in *Arabidopsis thaliana*, see Merlot *et al.* , 2001) which are negative regulators of ABA signaling and triggers the phosphorylation/activation of downstream secondary messengers as the SnrK proteins **(Figure 7 E).**

Among the ABA membrane receptors, the PYR/PYL/RCAR (*pyrabactin-resistance 1/ pyrabactin-resistance like/ regulatory component of ABA receptor*) are the most described in literature and their interaction with the PP2C phosphatases was validated and supported by several reverse genetic and crystallographic studies. The G-*protein-coupled receptors* (GPCR) and *Mg- chelatase H subunit* (CHLH or ABAR) are considered as putative ABA membrane receptors because of their affinity for the hormones, but their roles remain to be validated (Klingler *et al.* , 2010).

1.3.2 Signal transduction through secondary messengers

After perception of the water deficit signal by membrane receptors, secondary messengers are involved in the signal transduction to the cytoplasm and nucleus.

- **Calcium signatures**

Ca^{2+}, one of the most important second messengers in response to extracellular stimuli in plants, plays a crucial role in the water deficit signaling cascade. The cellular concentration of

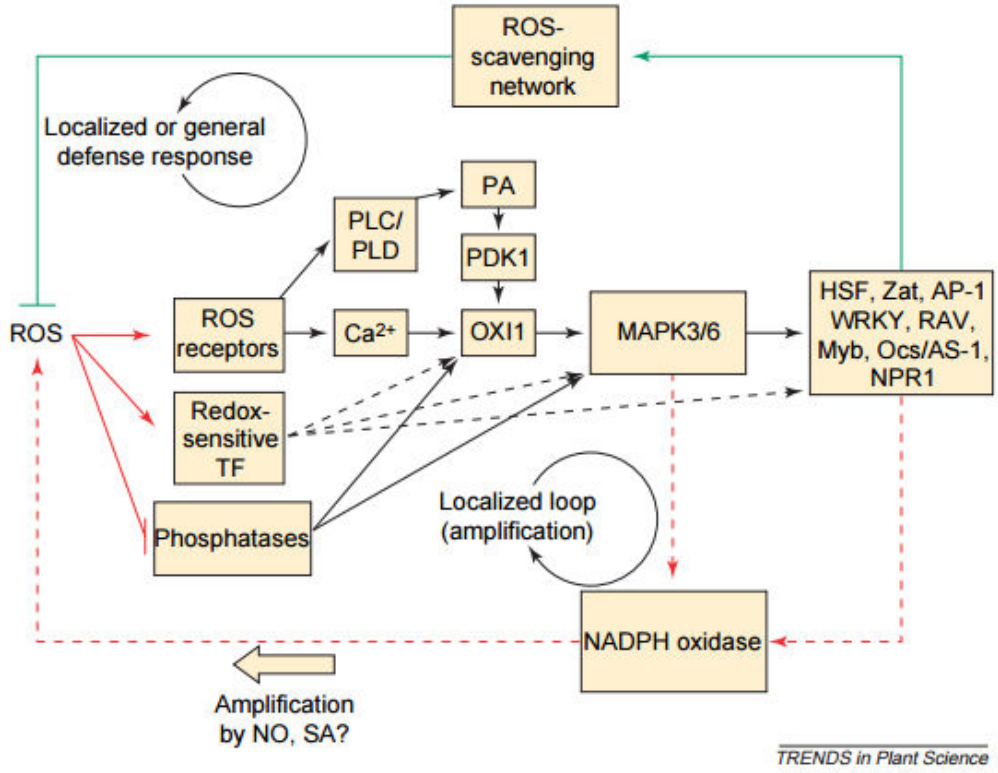


Figure 8. Overview of current knowledge on reactive oxygen species (ROS) signal transduction pathway. ROS can be detected by at least three mechanisms (ROS receptors, redox-sensitive transcription factors and *phosphatases*). Detection of ROS by receptors results in the generation of Ca^{2+} signals and the activation of a *phospholipase* (PLC/PLD) activity that generates *phosphatidic acid* (PA). PA and Ca^{2+} are thought to activate the *protein kinase* OXI1. Activation of OXI1 results in the activation of a *mitogen-activated-protein kinase* cascade (MAPK3/6) and the induction or activation of different transcription factors that regulate the ROS-scavenging and ROS-producing pathways. Two different loops are shown to be involved in the ROS signal transduction pathway. A localized or general defense response (a negative feedback loop; solid green line) can be activated to suppress ROS, whereas a localized amplification loop (positive feedback loop; red dashed line) can be activated to enhance ROS signals via the activity of NADPH *oxidases*. *Salicylic acid* (SA) and *nitric oxide* (NO) might be involved in this amplification. *(Adapted from Mittler et al. , 2004)*

Ca^{2+} is balanced by the presence of 'Ca^{2+} stores' like vacuoles, endoplasmic reticulum, mitochondria and cell walls. In response to water deficit, the Ca^{2+} channels, which are located into the vacuole and endoplasmic reticulum, are activated by ABA through the intermediate of the secondary messenger *inositol-1,4,5-triphosphate* (IP3), leading to the release of Ca^{2+} into the cytosol. When cellular conditions become optimal again, the calcium can be stored back to the vacuoles and endoplasmic reticulum through the activation of the *Ca^{2+} ATPase* and *H^+/Ca^{2+} antiporters*.

The variations of the cytosolic Ca^{2+} concentration (termed as calcium signatures) are recognized by calcium sensors or Ca^{2+} binding proteins, which are mainly *calmodulins* (CaM), *calmodulin like proteins, calcineurin B-like proteins* (CBL) and *Ca^{2+} dependent protein kinases* (CDPK). These calcium sensors are able to induce the expression of water deficit responsive genes (reviewed in Tuteja and Mahajan, 2007).

- **Reactive oxygen species**

The ROS are unequivocally involved as secondary messengers between abiotic stress receptors and the downstream signaling cascade. Whereas Ca^{2+} signaling is predominantly controlled in plants by storage and release, ROS signaling is controlled by production and scavenging. Two different loops are reported to be involved in the ROS signal transduction pathway. A general defense response can be activated to suppress ROS (see **1.2.2** and **Figure 5** for the main ROS scavenging pathways), whereas an amplification response can be activated to enhance ROS signals via the activity of *NADPH oxidases* able to actively produce hydrogen peroxide (H_2O_2) **(Figure 8)**.

H_2O_2 is the most likely ROS to be involved in the stress-response signal transduction pathways because of its relative stable structure and its ease to diffuse from one cellular compartment to another. Up to now, the known downstream events modulated by H_2O_2 are calcium mobilization through activation of Ca^{2+} permeable channels in the plasma membrane, protein phosphorylation through MAPK/CDPK cascades (see below) and gene regulation. There is a need for further analysis in order to elucidate ROS signaling function, for example the role of Ca^{2+} on ROS action in guard cells should be clarified (reviewed in Mittler, 2002; de Carvalho and Cruz de Carvalho, 2008).

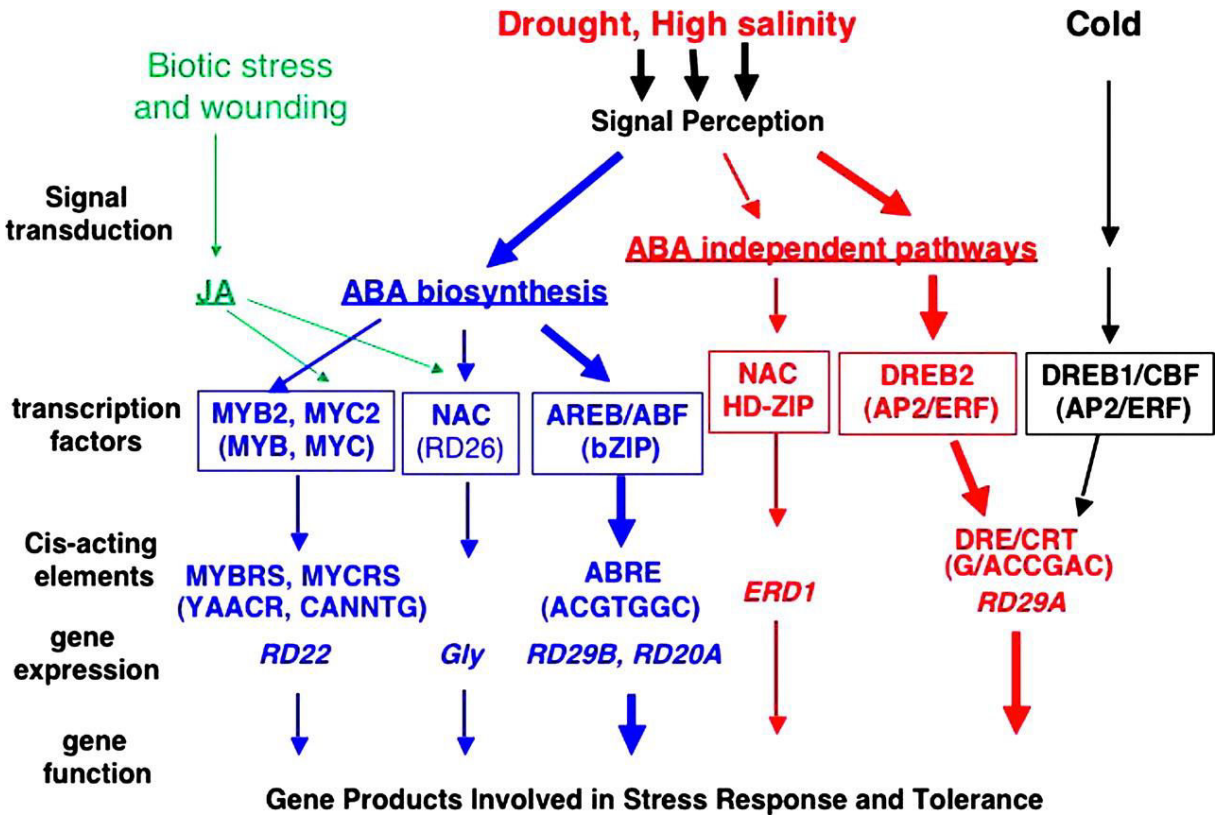


Figure 9. Transcriptional regulatory networks of abiotic stress signals.

(Adapted from Shinozaki and Yamaguchi-Shinozaki, 2007)

- **Protein kinases and protein phosphatases**

Protein kinases are involved in catalyzing the phosphorylation of others proteins, whichcorrespond to the addition of a phosphate ion. On the opposite, *protein phosphatases* catalyze the dephosphorylation of proteins, which means the removal of a phosphate group by hydrolysis. Phosphorylation/dephosphorylation mechanisms are known to be involved in the activation and inactivation of enzyme activity and in modulating protein-protein interactions within signaling networks.

Among kinase proteins, *calcium dependent proteins kinases* (CDPK) (reviewed in Schulz *et al.* , 2013), *mitogen activated protein kinases* (MAPK) (reviewed in Sinha *et al.* , 2011) and *sucrose non-fermentation 1 (SNF1)-related kinases* (SnrK) (reviewed in Kulik *et al.* , 2011) were reported to be involved in transducing abiotic stress signals from the plasma membrane to the nucleus. The MAPK proteins are involved in a regulatory cascade (MAPK cascade) composed of essentially three components, a MAPK kinase kinase (MAPKKK), a MAPK kinase (MAPKK) and a MAPK, connected to each other by events of phosphorylation (Sinha *et al.* , 2011). SnrK proteins from groups 2 and 3 play a crucial role in the ABA signalization pathway through their involvement in the inhibition of the H^+ *ATPase*s and in the activation of anion and K^+ *in* channels involved in stomata closure mechanism (Kulik *et al.* , 2011).

Among phosphatase proteins, two groups are largely described in literature and named according to their ability to bind specific protein residues: the *serine/threonine protein phosphatases* (PPases) and the *tyrosine-specific protein phosphatases* (PTPases). Numerous phosphatases, among which the *tyrosine-specific protein phosphatase PP2C,* are known to be involved in the ABA signaling pathway (Schweighofer *et al.* , 2004).

1.3.3 Transcriptional control

Plant genomes contain thousands of transcription factors, most of them belonging to few large multi-gene families. Individual members of a same family often respond differently to various abiotic stress stimuli, but several stress responsive genes may share the same transcription factors because of considerable overlaps between the different abiotic stress signaling pathways, including salt, heat, cold and drought stress. Classically, the transcription factors involved in response to abiotic stress are distinguished according to their sensitivity

to ABA, defining two types of transcriptional pathways, the ABA-mediated pathways and the non-ABA-mediated pathways, although there are evidences for complex cross talk between both types of pathways **(Figure 9)** (Wang *et al.* , 2003; Fujita *et al.* , 2006; Golldack *et al.* , 2014)**.**

- **ABA-mediated transcriptional responses**

Several ABA-inducible genes contain a conserved cis acting element ACGTGG/TC, named *ABRE*, in their promoter regions. B*asic leucine zipper* (bZIP *ABRE/ABF/TRAB*) transcription factors can recognize this conserved element and activate the expression of these genes.

Recently, the *MYC/MYB (Myeloblastis Oncogen - Myelocytomatosis Oncogen),* NAC and WRKY transcription factors were identified as transcriptional activators in ABA-inducible gene expression, suggesting another regulatory system for gene expression in response to ABA than the *ABRE-bZIP* regulatory system. WRKY transcription factors were reported to bind to the *W box* motif, (T)TGAC(C/T), located in the promoters of several ABA-inducible genes, including others ABA dependant transcription factors (Singh *et al.* , 2002; Kuromori and Shinozaki, 2010; Rushton *et al.* , 2012).

- **Non ABA mediated transcriptional responses**

Numerous ABA independent water deficit induced genes have the conserved *cis* acting element A/GCCGAC in their promoters, named *dehydration responsive element/ C-repeat element* (DRE-CRT). Two groups of *ethylene responsive transcription factors* (ERF) that bind to the *cis* acting DRE-CRT elements have been cloned and characterized: the *dehydration responsive transcription factors* (DREB) and the *C-repeat binding factors* (CBF). They are induced under different abiotic stress, including low temperature, water deficit and osmotic stress. They are encoded by genes of the AP2/EREBP multigene family (*APETALA2/ethylene-responsive-element-binding protein*).

Recently, ERD1 genes which encode *Clp protease regulatory subunit* (ClpD) were identified as involved in controlling the expression of drought inducible genes. They contain in their promoter ABA independent *cis* regulating elements and senescence activation *cis* acting element (ACGTATERD1, ACGT). NAC transcription factors interact with these *cis* acting elements (Singh *et al.* , 2002; Kuromori and Shinozaki, 2010).

Table 1. Classification of differentially expressed genes in response to water deficit in *Arabidopsis thaliana* proposed by Bray (2002) on the basis of microarray analysis and differential expression data available in literature.

Functional process	Genbank ID	Protein entry code*
Amino acid biosynthesis and degradation		
Anthranilate synthase (α-subunit)	M92353	At5g05730
Anthranilate synthase (β-subunit)	L22585	At5g57890
Tryptophan synthase (α-subunit)	U18993	At3g54640
Tryptophan synthase (β-subunit)	M23872	At5g54810
Δ^1-Pyrroline-5-carboxylate synthetase	AB050546	At2g39800
S-Adenosylmethionine synthetase	M33217	At4g01850
Lactoylglutathione lyase-like	AB050576	At1g11840
Chorismate mutase	Z26519	At3g29200
Aromatic metabolism		
4-Coumarate : CoA ligase	U18675	At1g51680
Cinnamyl alcohol dehydrogenase	L37883	At4g39330
Cinnamyl alcohol dehydrogenase	X67816	At4g37980
Chalcone synthase	M20308	At5g13930
Phenylalanine ammonia-lyase	L33677	At2g37040
Cinnamoyl–CoA reductase	T41765	At1g15950
Dihydroflavonol-4-reductase	T20927	At2g33590
O-Methyltransferase	U70424	At5g54160
Fatty acid multifunctional protein		
Hydroperoxide lyase	AF087932	At4g15440
Acyl-CoA oxidase	AF057043	At5g65110
Acyl-CoA oxidase	AF057044	At4g16760
Epoxide hydrolase	D16628	At2g26740
Omega-3 fatty acid desaturase	D14007	At3g11170
Lipoxygenase	L23968	At3g45140
Allene oxide synthase	X92510	At5g42650
Energy		
Oxygen-evolving complex	X52428	At5g66570
PSI, reaction centre sub II	AB050572	At4g02770
Transcription		
14–3-3 like protein, GF14	U60445	At3g02520
Ethylene response element binding protein 4	AB008106	At1g53170
AREB1	AB017160	Not annotated
DREB2A	AB007790	At5g05410
ATMYB2	D14712	At2g47190
ATHB-6	AF104900	At2g22430
ATHB-7	X67032	At2g46680
ATHB-12	AF001949	At3g61890
His1-3	U72241	At2g18050
RNase RNS1	U05206	At2g02990
Cell growth, cell division and DNA synthesis		
Nitrilase (indole-3-aceto-nitrile hydrolysis)	U09958	At3g44300

1.3.4 Genes induced in response to water deficit

In the last decades, thanks to the development of the cDNA microarrays and the RNA sequencing technologies, large-scale parallel analyses of gene expression were conducted in plants, allowing the identification of thousands of genes differentially expressed in response to abiotic and biotic stressors (see for review Cushman and Bohnert, 2000 and Umezawa *et al.* , 2006*a*, for major results in *Arabidopsis thaliana:* Seki *et al.* , 2001, 2002, for major results in rice: Rabbani *et al.* , 2003). Some of these genes were studied in reverse genetic studies to elucidate their biological function, but still a large number of genes have unknown functions. Three major genes family described in literature for their involvement in response to water deficit are presented below. Nevertheless, as water deficit affects cell growth and primary and secondary metabolites, many other genes were found to have their expression modified under water deficit conditions as illustrated in **Table 1.**

- **Heat-shock and molecular chaperon proteins**

Dysfunction of proteins and enzymes is usually observed in plants submitted to water deficit. Maintaining proteins in their functional conformations, preventing the aggregation of denatured proteins and eliminating non-functional peptides are crucial. This is the role of the molecular chaperones, in particular the *heat-shock proteins* (HSPs), which are up-regulated under water deficit and others abiotic stresses. Among the five major families of HSPs, the *small heat-shock proteins* (sHSPs) are the most prevalent in plants.

In addition to their involvement in membrane and protein protection, HSPs were also found to interact with several signaling molecules (including hormone receptors, tyrosine and serine/threonine kinases and cell-cycle/cell death regulators) and with other stress-response mechanisms such as osmotic adjustment and detoxification mechanisms**.** Cross-talk mechanisms between HSPs/chaperon proteins and others stress-responsive mechanisms still need to be explored to provide a further understanding of plant response to water deficit and others abiotic stress (reviewed in Wang *et al.* , 2004).

- **Late embryogenesis abundant proteins and dehydrins**

Late embryogenesis abundant proteins (LEA) and dehydrins are proteins considerably synthetized in response to water deficit and during the last stage of embryogenesis in

Table 1 continued

Protein synthesis		
Ubiquitin (UBQ4)	X12853	At5g20620
Ubiquitin (UBQ1)	J05507	At3g52590
AtHsp81–2	AB011476	At5g56030
AtHSP70–1	M23105	At5g02500
rReg ATP subunit of CLP protease	AB000615	At5g51070
DNAJ homologue	AB050562	At3g62600
Cysteine protease	D13043	At3g19390
Cysteine protease	D13042	At4g39090
Cysteine protease	AB050573	At4g16190
Cysteine protease	X74359	At2g21430
Cysteine protease inhibitor	AB044405	At2g40880
Metallopeptidase	Y13577	At1g51760
Transport		
Aquaporin 2C	D13254	At2g37180
Aquaporin	AB050549	At2g39010
γ-TIP2	AB050557	At3g26520
Sugar transporter (ERD6)	D89051	At1g08930
Intracellular transport protein	AB050567	At2g24420
HVA22-like (YIP2-like)	AB015098	At4g24960
Cell communication/signal transduction		
Ca-binding EF-hand protein	AB039924	At2g33380
CDPK1	D21805	At1g18890
CDPK2	D21806	At1g35670
CDPK	D28582	At2g17290
AtPIP5K1 (phosphatidylinositol-4-phosphate-5-kinase)	AB005902	At1g77740
AtPLC1	D38544	At5g58670
ATMEKK1 (MAPKKK)	D50468	At4g08500
ATMPK3	D21839	At3g45640
ATPK19	D42061	At3g08720
Oxidative stress		
Glutathione-*S*-transferase	D17672	At1g02930
Glutathione-*S*-transferases	D17673	At2g30870
Glutathione-*S*-transferase	D44465	At2g29450
Glutathione-*S*-transferase	AJ012571	At1g78380
L-Ascorbate peroxidase	AB050564	At1g07890
GSH-dep dehydroascorbate reductase	AB050550	At1g19570
Catalase 3	AB050551	At1g20620
Glutathione peroxidase GPX2	AJ000470	At2g31570
Cu, Zn Superoxide dismutase	X60935	At1g08830
Thioredoxin, FL3-2C6	AB050571	At5g42980
Peroxiredoxin TPX1	AB050556	At1g65980
Pathogenesis-related		
Basic chitinase, PR3B1	M38240	At3g12500
β-1–3-Glucanase, PR2	M90509	At3g57260
Extensin-like	T41880	At2g43150
Proline-rich	R64825	At2g14890
PR1 (antifungal protein)	M90508	At2g14610
PR5 (thaumatin-like, acidic)	M90510	At1g75040
Thionin	L41244	At1g72260
Putative lectin	N38164	At3g16460
AIG2-like (phosphate acetyltransferase)	U40857	At3g28930
Metallothionein-like	L15389	At1g07600
Antifungal protein-like	T04323	At5g44420

desiccated seeds. Their functions are still largely unknown. It was suggested that they could be involved in binding water, in ion sequestration and in macromolecule and membrane stabilization (Tunnacliffe and Wise, 2007).

- **Aquaporins**

Aquaporins are proteins allowing water and small neutral solutes to move across cellular membranes. According to amino acid sequence similarity, they are classified in four subfamilies: the plasma membrane intrinsic proteins (PIPs), the tonoplast intrinsic proteins (TIPs), the noduline26-like intrinsic proteins (NIPs) and the small and basic intrinsic proteins (SIPs) (Chaumont *et al.* , 2005).

Aquaporins have been shown to control the extensive water transports from the roots to the leaves during transpiration flux, the transport of assimilates into the phloem, the closure or aperture of the stomata in the leaves, the movement of the leaves and the cytoplasmic homeostasis. They are activated through phosphorylation/dephosphorylation mechanisms induced by CDPK. On the other hand, acid pH and free Ca^{2+} reduce the water permeability of the aquaporins. In maize, Wan *et al.* , (2004) reported that aquaporin gating could be under the control of ABA.

1.3.5 Histone modifications and DNA methylation

Next to transcriptional regulations through specific transcription factors, epigenetic processes can strongly influence the efficiency of gene expression in response to water limitation (see Suji and Joel, 2010 and Wang *et al.* , 2010 in rice or Labra *et al.* , 2002 in pea). Epigenetic processes include histone modifications, DNA de/methylation and non-coding RNA-associated gene and transposable element silencing mechanisms.

These last years, key enzymes involved in de/methylation mechanisms were characterized, namely, *methyltransferase 1* (MET1), *chromomethylase 3* (CMT3) and *domain rearranged methylase* (DRM1 &DRM2) (Chinnusamy and Zhu, 2009; Sahu *et al.* , 2013). Small RNAs were also reported to play an important role in epigenetic regulation in response to abiotic stress via transcriptional gene silencing through RNA directed DNA methylation (Khraiwesh *et al.* , 2010). Besides, among others phytohormones, ABA could be involved in regulating epigenetic modifications (Chinnusamy and Zhu, 2009). However, detailed knowledge on the specific mechanisms that underlay epigenetic regulation under environmental exposure is

Table 1 continued

Cellular organization		
Ferritin	AB050569	At5g01600
β-glucosidase-like	AB050566	At1g52400
Polygalacturonase-like (*rd22*)	D10703	At5g25610
Xyloglucan endo-transglycosylase	AB050552	At1g14720
Ripening-related protein	AB046991	At5g62350
Polygalacturonase-inhibiting	AB010697	At5g06870
Reversibly glycosylated polypeptide-2	AB050560	At5g15650
Unclassified hydrophilic proteins		
rd29B (Lti65)	D13044	At5g52300
rd29A (Lti78)	D13044	At5g52310
Cor15a	U01377	At2g42540
Kin1	X51474	At5g15960
Kin2 COR6·6	X55053	At5g15970
LEA14	Y10085	At1g01470
LEA 76 type 1	AB050548	At1g52690
Group II LEA (Erd10)	D17714	At1g20450
Group II LEA (rd17, cor47)	AB004872	At1g20440
Group II LEA (Erd14)	D17715	At1g76180
Group II LEA (XERO2)	U19536	At3g50970
Classification not clear or unclassified		
Drought-induced protein like	AB050563	At4g02380
Erd7	AB039929	
Erd3	AB039927	
Erd4	AB039928	
Steroid dehydrogenase-like	X99793	At4g24220
Lectin-like (JIP)	N37581	At3g16420
ERD15	D30719	At2g41430
REM3	R90622	At2g41870
Remorin-like	M25268	At2g45820
Glutamate-rich protein	AB050570	At2g05380
Major latex protein-like	AB050543	At4g23670
Ozone and pathogen induced	U20347	At4g00860
Non-specific lipid transfer protein	AB050558	At2g38540
FL5-2H15	AB050559	At5g61820
Non-specific LTP	AB050544	At5g59320
ENOD20-like	AB050542	At4g27520
Cold acclimation protein	AB044404	At2g15970

only slowly emerging. Little is known about the mitotic and meiotic heritability of histone modification and DNA methylation.

2. Deciphering the genotype by environment interactions in quantitative genetic

Water deficit responses presented in the previous section may vary not only across species, but also across individuals of a species. To breed genotypes adapted to various growing environments, there is a need to know the extent of these genetic variations by evaluating genotypes in multi-environment trials. Different tools are available in the quantitative genetic framework to explore the genetic variations in plant response to variations in their environmental conditions, at the phenotype and genome levels.

2.1 Concepts of *phenotypic plasticity* and *genotype by environment interaction* in quantitative genetic

The ability of a genotype to produce distinct phenotypes in different environments is known as *'phenotypic plasticity'* (Via and Lande, 1985; Schlichting, 1986). Phenotypic plasticity is very often genotype dependent and its variations are defined as *genotype by environment interaction* (G x E). G x E interactions can be described in so-called reaction norms, depicting the phenotypic values of different genotypes over several environments **(Figure 10).** In the absence of an interaction between genotypes and environments, the reaction norms will run parallel to each other **(Figure 10 A & B)**. When phenotypic plasticity differs between genotypes, making the reaction norms to potentially cross, this is considered as genotype by environment interaction (G x E) **(Figure 10 C & D)**. A distinction should be made between the most extreme case of a '*cross-over interaction'*, in which reaction norms cross with each other **(Figure 10 D),** and '*scale-effect interaction'*, where no intersection between reactions norms is observed within the environments **(Figure 10 C)** (El-Soda *et al.* , 2014). In practice, a cross-over interaction means that a genotype performing superior in one environment may perform worse in another environment. On the opposite, scale-effect interaction does not change the genotype ranking across environments.

Although the importance of the differential effect of the environment on different plant genotypes has been known for a long time and has been considered in crop breeding

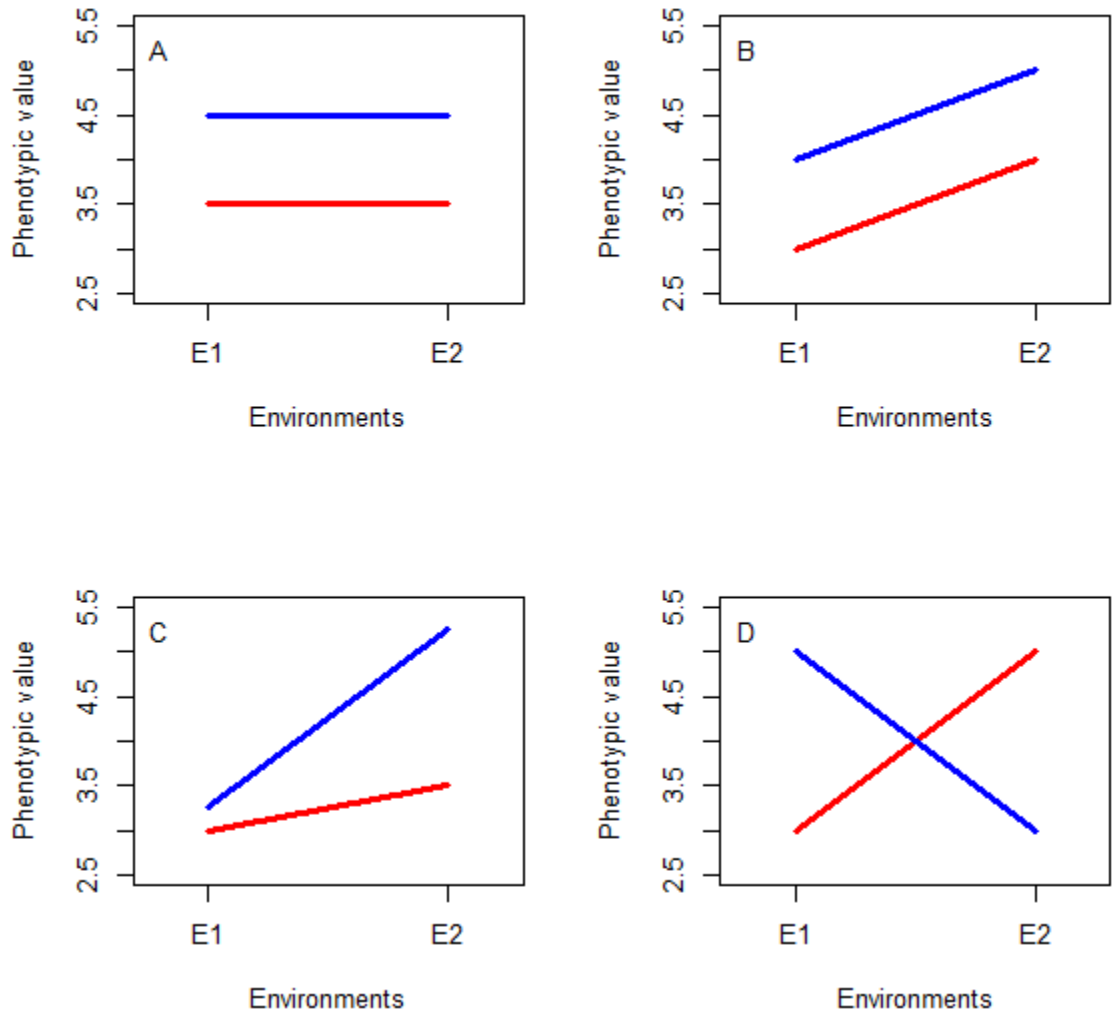


Figure 10. Illustration of the effects of genotype by environment (G x E) interactions on reaction norms, representing two genotypes in two environments. (A) No phenotypic plasticity and no G x E. (B) Phenotypic plasticity but no G x E. (C) Phenotypic plasticity and G x E with changing scale effect. (D) Phenotypic plasticity and G x E with genotype re-ranking across environments. Blue and red colors indicate two different genotypes. E1 and E2 indicate two different environments.

programs, it is generally viewed as a challenging issue. If there is no G x E, a trial in a breeding program conducted at only one location would indicate genetic classification whatever the environment. If there is G x E, discarding genotypes evaluated in only one environment in early stages of a breeding program, bears the risk to discard genotypes with a high potential in another environment. Evaluating genotypes in diverse environments (multi environment trials, MET) will help to identify superior and/or stable genotypes across different environments, providing more robust breeding results and offering insight into the genetic basis of G x E interactions. There is a need to explore genetic variation under both, controlled and natural conditions (at least 'production conditions'), as several studies showed very poor correlations between variations scored in field experiments and variation observed under greenhouse and/or climatic chamber conditions (see for instance Brachi *et al.* , 2010; Mishra *et al.* , 2012; Bac-Molenaar *et al.* , 2016).

In quantitative genetics, phenotypes can be modeled as the combined effects of genetic (G) and environmental (E) factors and their interaction (G x E) using different statistical approaches. The more classical method is the use of a *full interaction linear model* corresponding to the following equation: $P = G + E + G \times E$, where G x E allows to estimate the extent of the genotype by environment interaction. This approach does not allow prediction to be made of phenotypic response to environment variations that were not in the set of tested environments, as the different factors are considered as fixed factors. Setting some of the factors as random, transforming the previous model in a *mixed linear model*, will allow rough prediction as long as the new environment can be ranked in between two environments that were part of the multi-environment trial (van Eeuwijk *et al.* , 2005). Besides, in order to characterize the nature of the interaction, G x E term can be partitioned according to Muir *et al.* (1992) into heterogeneous variances, based on heterogeneity among environments in the scaling of differences among genotypes (method 1) or imperfect correlations, based on heterogeneity among genotypes in the scaling of differences among environments (method 2).

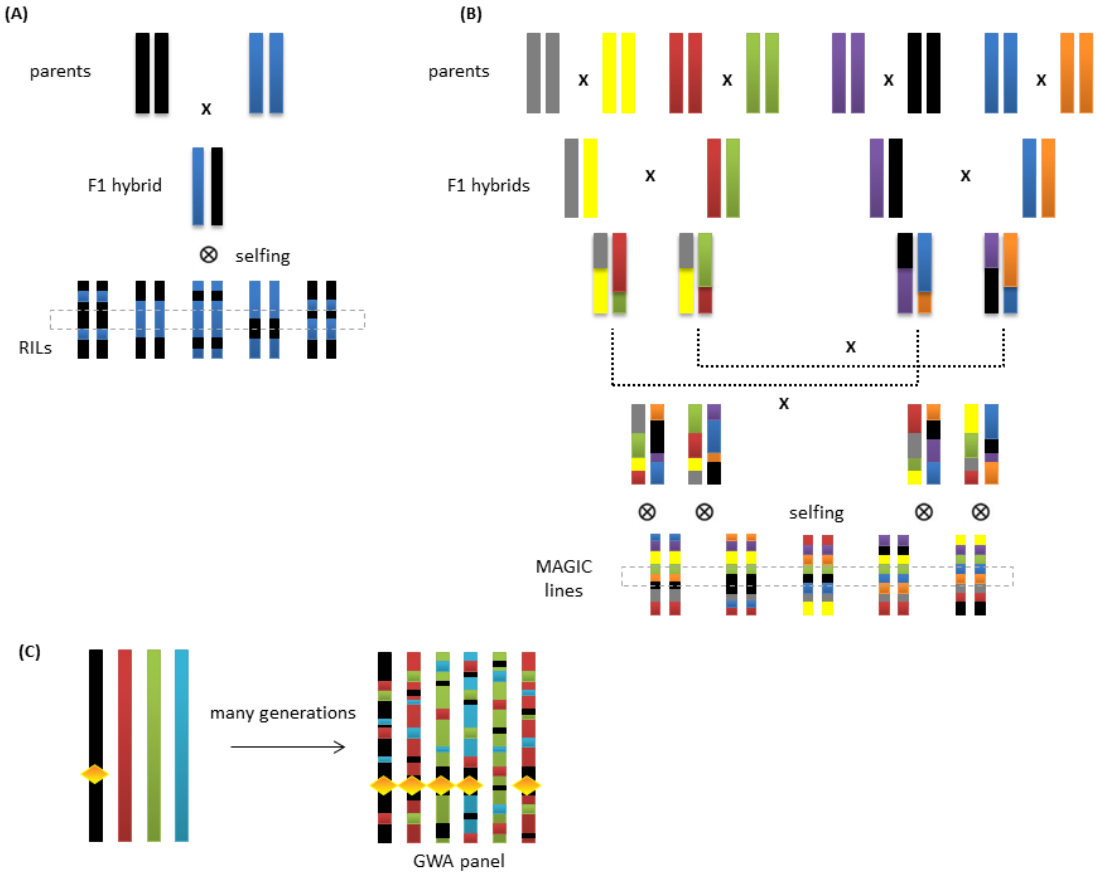
Alternatively, Finlay and Wilkinson (1963) proposed the *regression on the mean model*. The idea behind this model is that in the absence of explicit characterization of an environment, a good approximation to the biological quality of an environment is given by the average

phenotypic value across the genotypes. Thus, the responses of individual genotypes are regressed on the average performance and the G x E is expressed by difference in the slopes between genotypes. The *linear-bilinear models* are an extension of the *regression on the mean model* and were developed to allow more flexible characterization of the environments (see for instance the *AMMI models* in Gauch, 1988; Sabaghnia *et al.* , 2008 and *GGE biplots* in Yan *et al.* , 2007).

All the statistical approaches presented above do not integrate characterization of the environments and are rather difficult to interpret in terms of plant physiology. The *factorial regression models* (Hébert *et al.* , 1995; Vargas *et al.* , 1999; Malosetti *et al.* , 2007) which allow to regress G x E term onto environmental variables (such as light intensity, water availability or temperature) and the *ecophysiological models* constitute appropriate alternatives. Ecophysiological models are adequate tools for analyzing genotype by environment interactions since they integrate environmental and genetic effects on individual processes and are able to predict interactions among processes during plant growth and organ development (Tardieu, 2003; Bertin *et al.* , 2010). Ecophysiological models represent certain features of the plant and their interactions, through a series of mathematical algorithms, and test quantitatively their response to environmental or internal factors of the plant. Each plant trait (e.g. fruit growth or composition) is related to individual processes responsible for its variation and supposed to be less influenced by the environment than integrated traits. To assess model parameters, several environmental conditions have to be analyzed, like fruit load variations which will modulate carbon allocation or water stress which will modulate water intake (Prudent *et al.* , 2011). Some of the model parameters, called genotypic parameters, are genotype dependent while others are generic and independent from the genotype. The genotypic parameters defining a particular genotype represent a phenotypic fingerprint of this genotype and are amenable to genetic analysis (Boote *et al.* , 2001).

2.2 QTLs and mapping populations in quantitative genetic

Many agronomical traits, such as yield or quality traits, are controlled by several genes under the influence of the environment and are known as *quantitative traits*. The genomic regions that contain genes associated to a particular quantitative trait are depicted as *quantitative*



Comparison of advantages and limitations of the RIL, MAGIC and GWA populations.

	RIL	MAGIC	GWA
Time to develop	Intermediate	Long	Short
Precision in mapping common alleles	+	++	+++
Precision in mapping rare alleles	+++	+	–
Access to recombination	+	++	+++
Nb of markers needed	+	++	+++
Population structure	No	No	+
Main advantages	Rare allele mapping Easy analysis	Several alleles segregating Founder allele effect for MAS and QTL identification	Precision due to historical recombination
Main limitations	Large QTL confidence intervals	Time to set up Large population needed	High LD limits the precision Pop structure responsible for false positive

Figure 11. Comparison of the advantages and limitations of the RIL, MAGIC and GWA populations. *(Adapted from Zhu et al. , 2008 and Pascual, Albert et al. , 2016)*

trait loci (QTLs). In the late 1980's, the mapping of QTLs on genomes was made possible thanks to the development of genetic markers allowing to visualize existing polymorphisms between individuals at the DNA level. The first genetic markers were the phenotypic markers (phenotypic traits easily screened) and the biochemical markers (also known as isozymes, limited number), rapidly bypassed by the DNA markers (including *RFLPs Restriction Fragment Length Polymorphisms*, RAPD *Random Amplified Polymorphic DNA*, SSRs *Simple Sequence Repeats*, AFLP *Amplified Fragment Length Polymorphisms* and SNPs *Single Nucleotide Polymorphisms*) (Collard *et al.* , 2005). In parallel to the rise of the DNA markers, different methods for QTL mapping have been developed to link phenotypes to genetic markers, with increasing levels of statistical complexity (Staub *et al.* , 1996; Andersen and Lübberstedt, 2003).

2.2.1 Linkage mapping in progenies

Briefly, linkage QTL mapping methods are based on the principle that genes and markers segregate via chromosome recombination during meiosis (i.e. cross-overs) and thus allowing their analysis in progenies. Genes and markers that are close together are transmitted together to the next generation more frequently than genes and markers that are located faraway. There are four main steps in linkage QTL mapping experiments: (1) obtain a *mapping population* derived from contrasted genotypes for the phenotype of interest; (2) *characterize the phenotype* of a relatively large number of individuals from the population; (3) build *a genetic linkage map* using *genetic markers* genotyped in the whole population and recombination rates between them; (4) perform statistical analysis to identify the loci underlying the genetic architecture of the trait.

Most commonly, mapping populations are derived from the cross of two parental lines, such as a population of F2 or F3 plants, back-cross plants, doubled haploid (DH) lines, or Recombinant Inbred Lines (RILs) **(Figure 11a)** (see for review Collard *et al.* , 2005). The advantage of DH or RIL population is that each line is nearly homozygous and can be easily propagated by selfing.

Different statistical tools can be used to measure the link between phenotypes and markers. The simplest method for QTL mapping is ANOVA (ANalysis Of VAriance) that assesses the relationship of a phenotype with a marker genotype, and thus indicates which markers are

associated with the quantitative trait of interest (Tanksley, 1993). This method is simple, but disadvantages are that individuals with missing genotype data are excluded, QTL location is not precise in low density scan and it only considers one QTL at a time. *Simple interval mapping* (SIM) is another method for QTL mapping based on the estimation of a *genetic linkage map* (Lander and Botstein, 1989). It statistically tests for a single QTL at each location incremented along the ordered markers in the genome. The results of the tests are expressed as LOD (logarithm of the odd ratio) scores, which compare the likelihood function under the null hypothesis (no QTL) with the alternative hypothesis (QTL at the testing position) for the purpose of locating probable QTL. The advantages of this method are (1) it takes into account missing data, (2) it allows higher power in low-density scans and (3) it improves the precision of QTL location. The disadvantages are (1) greater computational effort and (2) it only considers one QTL per chromosome.

The drawbacks of SIM mapping were overcome by *composite interval mapping* (CIM) (Zeng, 1994) and *multiple-QTL models* (MQM) (Jansen and Stam, 1994). Both methods combine interval mapping for a single QTL in a given interval with multiple regression analysis on markers associated with other QTLs. It considers a marker interval plus a few other well-chosen single markers as covariates in each analysis. The advantages of these methods are as follows: (1) mapping of multiple QTLs can be accomplished by the search in one dimension; (2) by using linked markers as cofactors, the test is not affected by QTL outside the region, thereby increasing the precision of QTL mapping; (3) by eliminating much of the genetic variance controlled by other QTL, the residual variance is reduced, thereby increasing the power of detection of QTL. The limitations are: (1) the use of tightly linked markers as cofactors can reduce the statistical power to detect QTL; (2) the test statistic in a marker rich region may not be compared to that in a marker poor region; (3) estimation of the joint contributions of multiple linked QTL and epistasis is difficult (Zeng *et al.* , 1999).

Multiple interval mapping (MIM) is the extension of interval mapping of multiple QTLs, just as multiple regression extends analysis of variance. MLM allows inferring the location of QTLs to positions between markers and enables interactions between QTLs to be tested.

There are many factors that influence the detection of QTLs in linkage mapping experiments. Crucial questions are the definition of the appropriate number of individuals and markers to ensure an acceptable mapping resolution. The number of markers needed

depends on the size of the genome of the species and the level of recombination inside the mapping population, which is strongly dependent upon the size of the population. Usually, larger the population, the more accurate the mapping study and the more likely it is to allow the detection of QTLs with smaller effects (Tanksley, 1993).

2.2.2 Recent advances in the linkage mapping area: multiparental populations

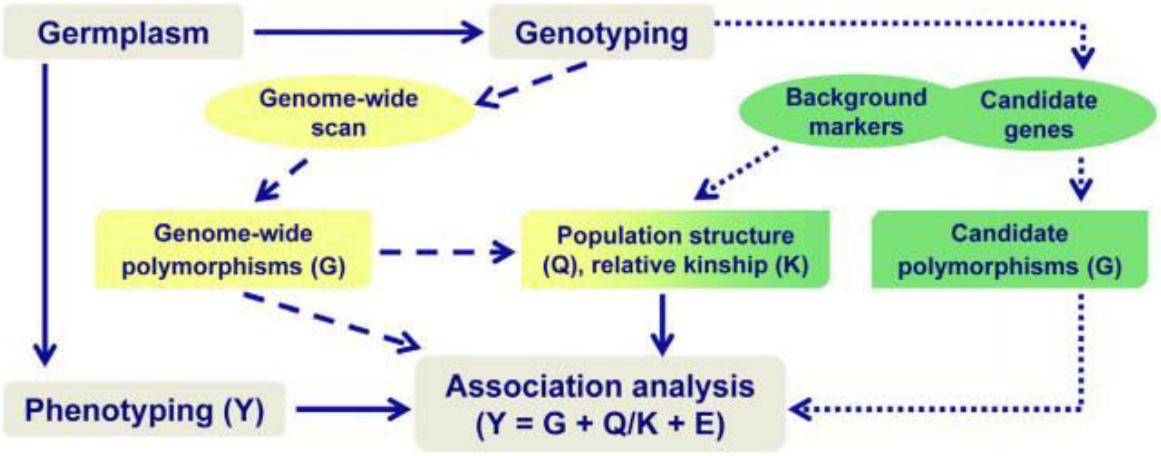
The precision of the classical linkage mapping populations described above is hampered by the limited allelic diversity present in the genitors (usually two) and a poor mapping resolution due to limited recombination events during the creation of the population. To overcome the limited number of recombinations found in the RIL progeny of a biparental cross, advanced intercross RIL populations (AIC-RIL, Balasubramanian *et al.* , 2009) have been developed, in which the number of recombinations is increased by inter-mating F2 plants and later generations before inbred lines are derived.

To even further increase the number of alleles and number of recombinations, the multiple parent populations such as the *Multiple Advanced Generation Inter-Cross* (MAGIC) populations **(Figure 11b)** (Kover *et al.* , 2009; Bandillo *et al.* , 2013; Pascual *et al.* , 2015) and the *multiparent RIL* (AMPRIL) populations (Huang *et al.* , 2011) were proposed. In maize *Nested Association Mapping* panels (NAM) were developed to study the segregation of several alleles in many RIL progenies connected by one parent (McMullen *et al.* , 2009). In trees, where the genealogies of varieties are well known and often related, specific designs were proposed using connected progenies (Allard *et al.* , 2016).

All the approaches listed above need a sufficiently high density of intermediate frequency markers to infer the most likely local founder genotypes. They became thus interesting with the SNP discovery and their possible analysis through SNP chips. Compared with GWAS (see below) these populations are not structured and offer more equilibrated allelic frequencies.

2.2.3 QTL mapping in unrelated populations: genome wide association

To identify genes underlying natural variation, genome wide association studies (GWAS) constitutes an alternative approach to the QTL analysis in progenies (Korte and Farlow, 2013). The advantage of GWAS over progeny analysis is that genotypes from naturally evolved and adapted populations can be used, which make elegant use of historical recombination accumulated over thousands of generations in random mating



Genome-wide association mapping	Candidate-gene association mapping
It is a comprehensive approach to systematically search the genome for causal genetic variation. A large number of markers are tested for association with various complex traits, and prior information regarding candidate genes is not required. It works best for a research consortium with complementary expertise and adequate funding.	Candidate genes are selected based on prior knowledge from mutational analysis, biochemical pathway, or linkage analysis of the trait of interest. An independent set of random markers needs to be scored to infer genetic relationships. It is a low cost, hypothesis-driven, and trait-specific approach but will miss other unknown loci.

Figure 12. The scheme of association mapping for tagging a gene of interest using germplasm accessions *(Adapted from Zhu et al, 2008)*

populations **(Figure 11c and Figure 12).** However, a potentially serious obstacle to association mapping is population structure and relatedness which may cause false positive (Nordborg and Weigel, 2008; Zhu *et al.* , 2008; Ogura and Busch, 2015).

Presently, the method is routinely applied in many crops thanks to large SNP chip availability and statistical improvements enabling to take properly into account both population structure and relatedness into the models to control for false-positive effects (use of mixed models, see Pritchard *et al.* , 2000; Yu *et al.* , 2006; Zhou and Stephens, 2012; Yang *et al.* , 2014). However, when the SNP effects are small and the causal variants have low frequency in the population and/or when the phenotypic trait is strongly correlated with the population structure, GWAS performs poorly (Dickson *et al.* , 2010). Besides, the resolution with which an association can be mapped is a function of how quickly the linkage disequilibrium (LD, non-random association between markers) decays over physical distances along the genome. When a large LD exists in the population, association mapping will not be much more precise than linkage mapping (often the case in autogamous species). On the opposite, when the LD is limited, the genetic resolution offered by GWA is potentially sufficiently high to narrow down the associated region to one or a few genes without the need for additional fine-mapping (Ravel *et al.* , 2007; Myles *et al.* , 2009).

The selection of the most suitable genome wide significant threshold is still a source of discussion in the literature. The GWAS threshold should account for the multiplicity of comparisons that are performed as part of the massive testing in a GWA study. The burden of multiple testing constitutes a major challenge for GWA studies. A variety of statistical approaches accounting for multiple testing in the genome-wide setting have been developed. The most suitable GWAS threshold depends upon the population, the linkage disequilibrium (independency is a major uncertainty concerning these methods), the minor allele frequency in the population (MAF) and type of genetic data (SNP from array, sequencing data…) (Panagiotou and Ioannidis, 2012).

Strategies combining linkage and association mapping have proven their value to account for false positives while limiting false negatives (Brachi *et al.* , 2010; El-Soda *et al.* , 2015) and were applied successfully in crop plants such as rice (Famoso *et al.* , 2011), soybean (Sonah *et al.* , 2015), sunflower (Cadic *et al.* , 2013) or wheat (Mir *et al.* , 2012). GWA statistical methods can be applied in the multiparental population described in **2.2.2**, without the

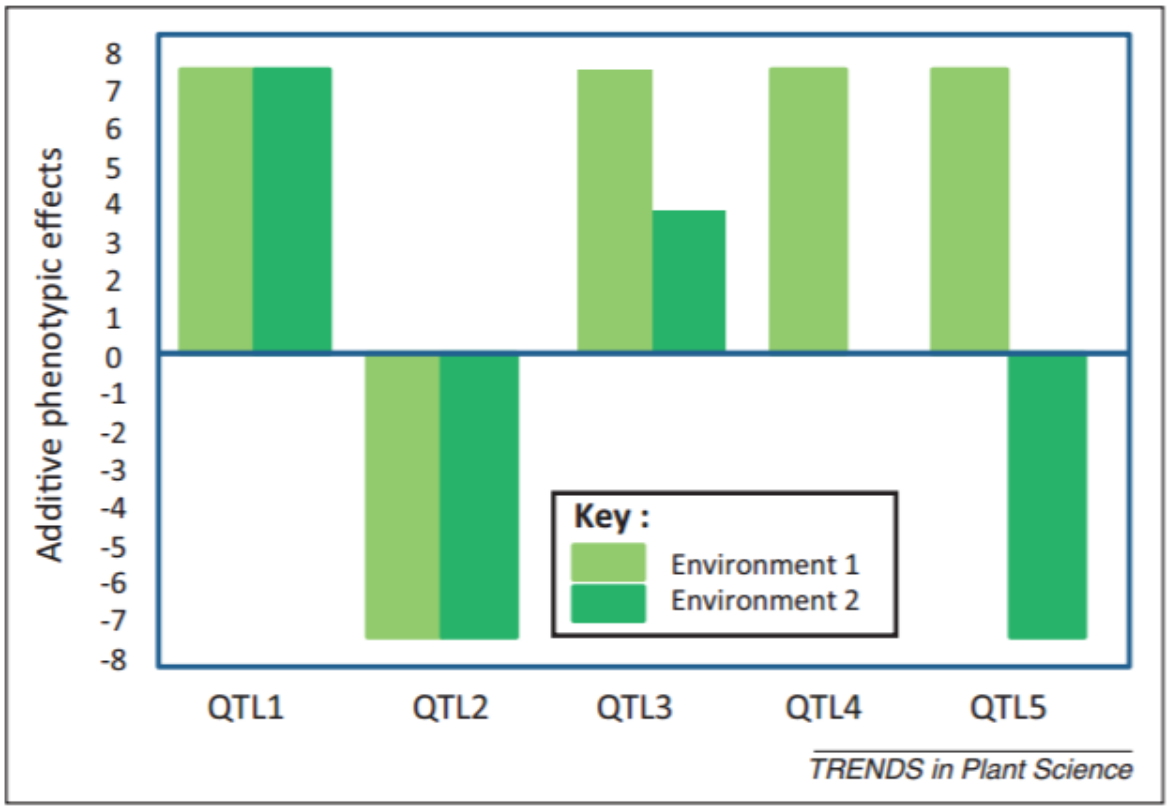


Figure 13. QTL by environnement interactions (QTL x E). QTL1 and QTL2 are constitutive QTL with positive or negative additive phenotypic effects, respectively, which do not show a QTL environment interaction (Q x E). QTL3, QTL4 and QTL5 shows QTL x E because their effects on the phenotype is changing according to the environment. QTL5 shows the strongest Q x E because it has opposite phenotypic effects when comparing both environments. *(Adapted from El-Soda et al. , 2014)*

difficulty of structured populations.

2.3 From genotype by environment to QTL by environment interaction

To dissect G x E into its individual genetic components, the genetic complexity of the phenotypic responses to the environment should be understood in terms of underlying QTLs and their allelic composition. As plant response to the environments is a complex trait, G x E often correspond to several QTLs.

2.3.1 Adding an environmental co-variable in QTL mapping models

Statistical improvements allowing to explicitly take into consideration the effect of environmental variables in QTL mapping models are available in the linkage (van Eeuwijk *et al.* , 2010; Verbyla *et al.* , 2014; El-Soda *et al.* , 2014; Li *et al.* , 2015) and association (Korte *et al.* , 2012; Saïdou *et al.* , 2014) framework and permit to analyze different population designs (RILs, GWA collections, MAGIC populations). These approaches are based on mixed models or multivariate regression models and offer the possibility to properly test the QTL by environment interactions and to identify QTLs whose effects are changing according to the environment. Into the analysis, a trait measured in two environments is regarded as two different traits implying that the physiological mechanisms underlying the same trait might be different across environments, and consequently, the loci underlying that differential performance are also different.

Based on the effect of each QTL in all tested environments, QTLs can be classified into: (1) *constitutive QTLs*, also called environmentally stable QTLs, with a small QTL x E effect but a large main effect in different environments **(Figure 13 QTL1 & QTL2)**; (2) *environment specific QTLs* showing an effect in one environment but no effect or lower effect in another **(Figure 13 QTL3 & QTL4)**; **(**3) *antagonist QTLs* which have effects with opposite directions according to the environment **(Figure 13 QTL5)** (El-Soda *et al.* , 2014). Main target of breeding programs are the constitutive QTLs as their effect is similar in all environments. The interest of specific and antagonist QTLs is more limited in breeding programs, except in situations in which the environment in finely controlled (horticultural productions in greenhouses for instance).

Mixed QTL mapping models are not the only alternative. A classical approach is to map the difference or ratio between a same phenotypic trait measured in two contrasting conditions

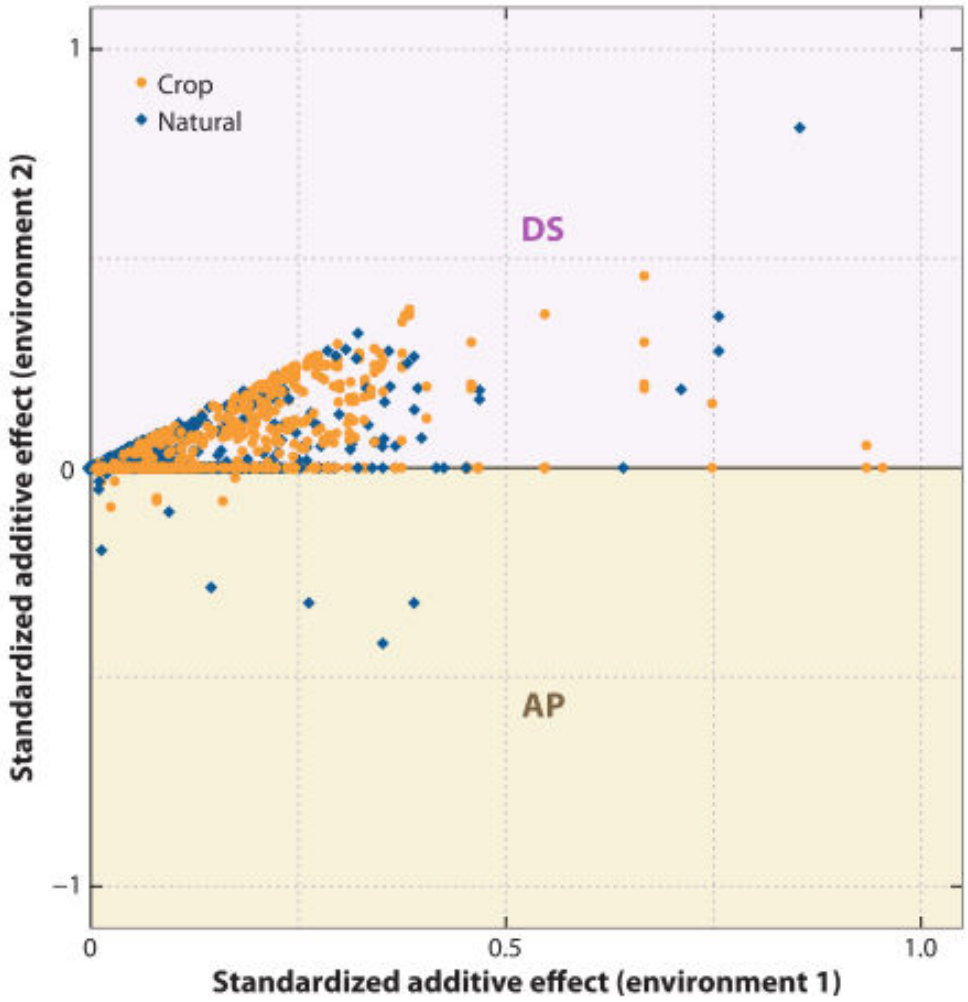


Figure 14. Biplot of quantitative trait locus additive effects from published data and standardized by the trait mean. Each x, y coordinate represents a comparison between an additive effect estimated in two environmental conditions plotted with the largest absolute additive effect on the x-axis. Above the x-axis shows differential sensitivity (DS), below the x-axis shows antagonistic pleiotropy (AP), and along the x-axis are environment- specific effects (censored or true conditional neutrality). Orange circles represent crop species, and blue diamonds represent natural species. (Des Marais *et al.* , 2012)

(thus measuring the level of *phenotypic plasticity*), which effectively translate a multivariate problem into a simpler univariate set-up (Tétard-Jones *et al.* , 2011). Alternatively, when more than two environments are compared, parameters from reaction norms can be used as plasticity traits.

Des Marais *et al.* (2013) performed a meta-analysis of 37 genetic mapping studies reporting QTL by environment interactions, including 11 plant species, multiple designs (Double Haploids, Recombinant Inbreed Lines, F2, backcrosses, Single Seed Descents), a large diversity of phenotypic traits (mainly germination, growth, metabolism, phenology, yield) and six categories of environmental conditions (including water, light and temperature stress). They revealed that only eight studies reported occurrences of QTLs with antagonist effects according to the environment (22% of the studies) and these QTLs represented only 1.4% of the 1,525 QTLs analyzed by the authors (22 QTLs). In contrast, most of the studies reported QTLs with environment specific effects (92% of the studies), with 57% of the QTLs lacking significance in at least one of the environment in which the experiments were carried out **(Figure 14).**

2.3.2 QTL mapping of models parameters

During the last ten years, approaches combining ecophysiological modelling and QTL analyses have been developed to understand the key processes involved in the control of complex traits under the impact of environment **(Figure 15)**. Such an approach has been applied to study specific traits, such as leaf area in barley (Yin *et al.* , 1999), leaf elongation in maize (Reymond *et al.* , 2003) and fruit quality in peach (Quilot *et al.* , 2005) and tomato (Prudent *et al.* , 2011). The method consists in simultaneously studying the genotypic variation of a given trait, and the genotypic variation of ecophysiological model parameters linked to key processes involved in the development of this trait.

Then, co-localizations of QTL for the trait and QTL for parameters give new insights into the processes involved in the trait at the QTL level, and then may help in understanding the physiological processes underlying QTLs, help in choosing candidate genes for characterization and give clues to design new ideotypes. This approach is particularly well-adapted to study interrelated processes linked to complex traits, and appeared to be an essential tool in the context of sugar accumulation (Prudent *et al.* , 2011, 2014), fruit growth

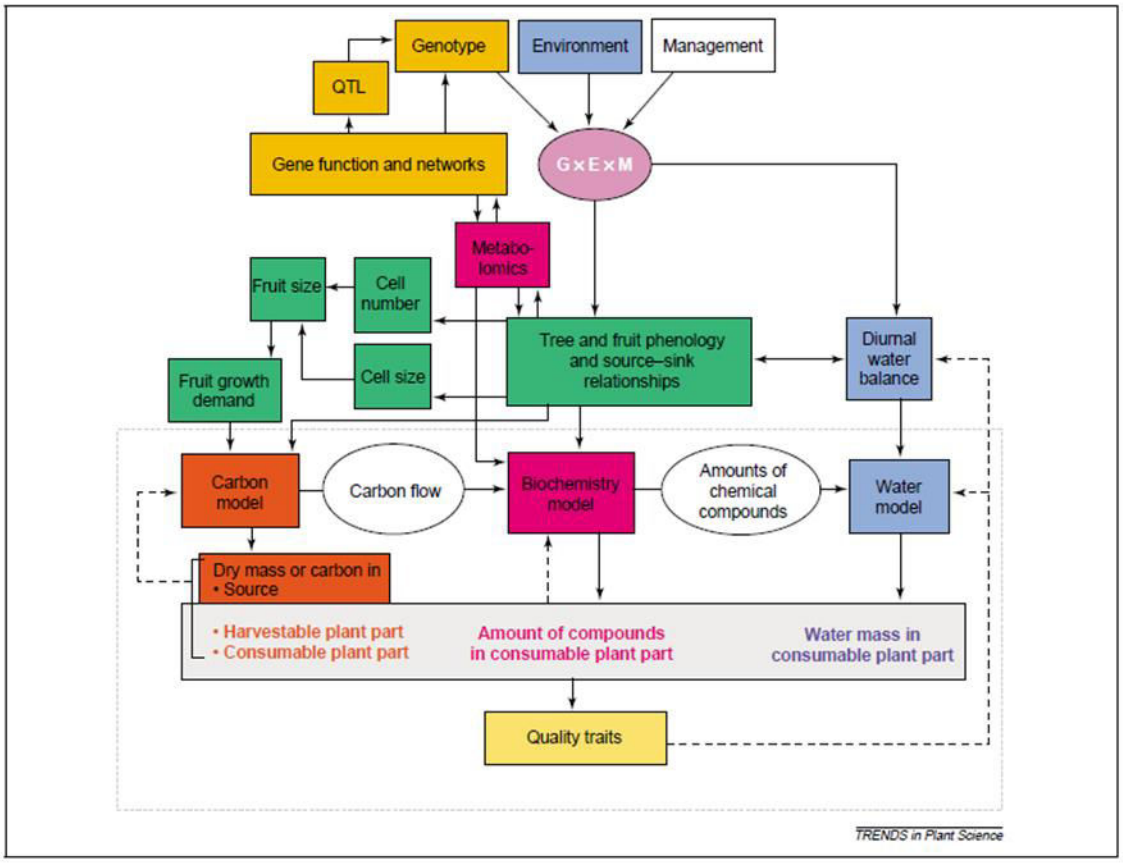


Figure 15. Example of integration of genetic information into an ecophysiological model. The ecophysiological model represented in the grey dashed area is the one developed by Lescourret and Génard (2005). This model combines three previously published models. The first model is a carbon model that quantifies the production of dry mass or carbon in the source, in the harvestable part (e.g. the fruit) and in the consumable part (e.g. the fruit flesh). The second model is a biochemistry model that quantifies the amounts of different quality-determining compounds. The third model is a water model that quantifies the flow of water into and out of the consumable part. The carbon flow from the carbon model is the most important constituent for the biochemistry, which again influences the water flow. The processes described in the carbon and water models are modulated by environmental factors, whereas the processes in the biochemistry model are also modulated by the phenology of the plant. Quality characteristics feedback into the water balance because they affect the osmotic potential of the fruit. Some extension of the model are proposed to take into consideration the effects of phenology (both at tree and fruit level) on the early phases of fruit development and to consider cell number as a function of rate and duration of cell division. With the rapid development of functional genomics, a further extension would be to feed the biochemistry model with information from metabolomics that can be linked to gene function and gene networks and, thereby, the model can be made genotype-specific by including the effect of QTL or genes and their regulatory networks underlying these *QTL. Genetic components are depicted in orange; environmental factors and the water model are depicted in blue: genotype x environment x management interactions are depicted in violet; intrinsic factors, phenology, source–sink relationships and fruit growth demand are depicted in green; dry mass and the carbon part of the model are depicted in red; the biochemistry model and metabolomics are depicted in magenta; resulting quality traits are depicted in yellow. Unbroken arrows indicate information flow; broken arrows indicate a feedback mechanism*.
(Adapted from Struik et al. , 2005)

(Liu *et al.* , 2007; Bertin *et al.* , 2009, 2010) and response to abiotic stresses (Reymond *et al.* , 2003, 2004; Laperche *et al.* , 2006). Nevertheless, up to now, all the studies reporting QTL analysis combined to an ecophysiological model were conducted in biparental populations. The question arises whether the results can be extrapolated to other crosses with other genetic backgrounds (Kraakman, 2004).

In parallel to the main work of the thesis, an article presenting the detection of QTLs of parameters from an ecophysiological model, modelling the genotype variability of tomato fruit growth under contrasted watering conditions, was achieved in collaboration with ecophysiologists (UR PSH INRA PACA Avignon) and accepted for publication in *Frontiers in Plant Sciences* in 2016. The article is entitled "Model-assisted estimation of the genetic variability in physiological parameters related to tomato fruit growth under contrasted water conditions", by Dario Constantinescu, Mohamed-Mahmoud Memmah, Gilles Vercambre, Valentina Baldazzi, Mathilde Causse, Elise Albert, Nadia Bertin and co-authors **(provided in Appendix 1).**

2.4 From QTL by environment interactions to candidate genes: combining gene expression and QTL analysis

Characterizing QTL phenotypic effects and examining if and how the expression of the underlying genes differs across environments is an important step in explaining at the molecular level how G x E determines phenotypes. Following the development of the expression arrays and sequencing technologies, genome-wide assays of gene expression were rapidly adopted by plant molecular biologists. Now publicly available databases contain results of hundreds of experiments measuring differential gene expression in response to biotic and abiotic stressors. However, only recently, researchers started to incorporate experimental designs that explore natural genetic variation in transcriptomic response to the environment using large number of genotypes. Such kinds of data sets constitute treasures for exploring the patterns of functional genetic variations. For example, enrichments tests can be used to ask whether gene lists associated with G x E are related to particular biological functions of genes, transcription factors binding sites in gene promoters or patterns of nucleotide diversity (Des Marais *et al.* , 2012, 2013). Besides, different strategies can be envisaged to link expression and polymorphism data as presented below.

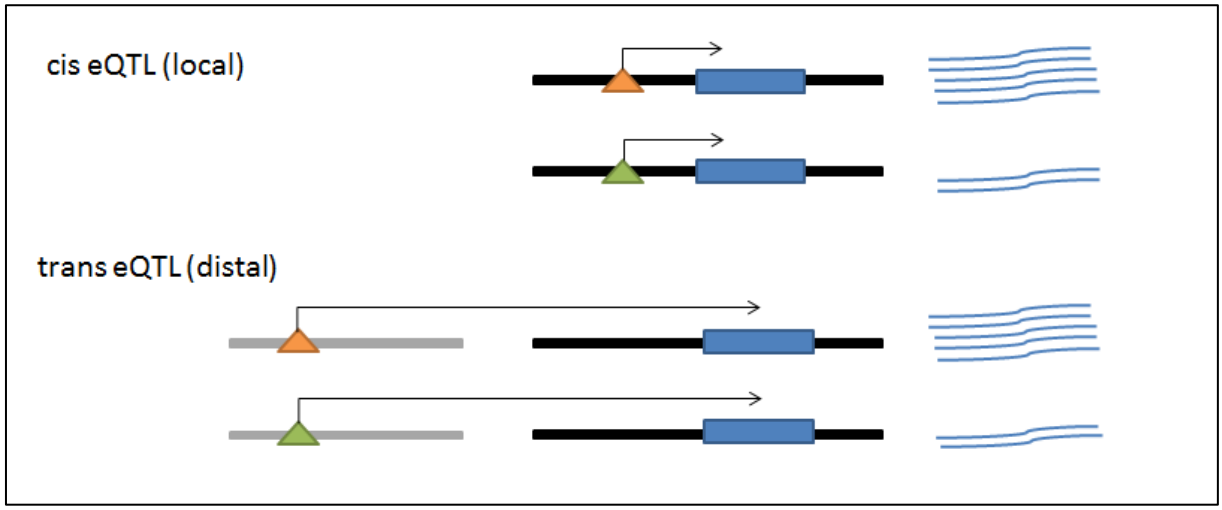


Figure 16. *Cis* and *trans* acting eQTLs identification in mapping populations. Triangles represent polymorphisms responsible for differential gene expression, either upstream of the regulated gene (*cis* eQTL, local) or on a different chromosome from the regulated genes (*trans* eQTL, distal).

2.4.1 Gene expression as phenotypic variable in QTL mapping models: eQTL mapping

A promising method is to consider transcript abundance as a trait in conventional mapping studies, either through linkage or association. Such approach, referred as *expression QTL mapping* (or eQTL mapping), allows inferring the effect and location of genomic regions responsible for the variation in gene expression. Up to now, studies reported in plants identified many transcripts whose heritability was beyond 0.50, both influenced by additive and non-additive effects (West *et al.* , 2007; Druka *et al.* , 2010; Cubillos *et al.* , 2012; Des Marais *et al.* , 2013). These results encourage the generalization of eQTL mapping in model and crop plants.

Classically, eQTLs have been characterized in literature as either *cis* or *trans* acting, depending on the physical distance from the gene they regulate. *Trans* acting eQTLs can be *distant,* when located on a different chromosome compared with the gene they regulate, or *local*, when located on the same chromosome but at a critical distance from the regulated gene **(Figure 16)** (Nica and Dermitzakis, 2013)**.** The definition of the critical distance allowing differentiating between *cis* and *trans* acting is still matter of debate in literature. Cubillos *et al.* (2012) suggested that as much as 70% of the eQTLs detected in maize, rice and *Brassica rapa* are trans-acting loci and that many of them cluster non randomly into genomic hot spots. Such hot spots correspond to master regulators. Nevertheless, the identification of *trans* acting eQTL is still contentious and depends strongly on the statistical power as *trans acting* eQTLs tend to have smaller effect than *cis* acting eQTLs.

Today, just a few plant studies have explored eQTL patterns in response to abiotic or biotic constraints (Hammond *et al.* , 2011; Cubillos *et al.* , 2012, 2014). It will be crucial in future experiments to formally test for *eQTL x E interactions* as it seems that differential expression plays a major role in plant adaptation to their environment **(see 1.3.3).** A pioneer publication on *Arabidopsis thaliana* provided an advanced strategy to use gene expression and *cis* eQTL data as clues for the identification of candidate genes involved in plant response to drought (Lovell *et al.* , 2015). The authors proposed to use gene expression as covariate in a QTL model to link markers, RNA expressions and phenotypes. They selected genes with significant cis-eQTLs and tested the effect of their transcript abundance on the effect of QTLs for phenotypic plasticity traits. Efficiency of the method was proved by

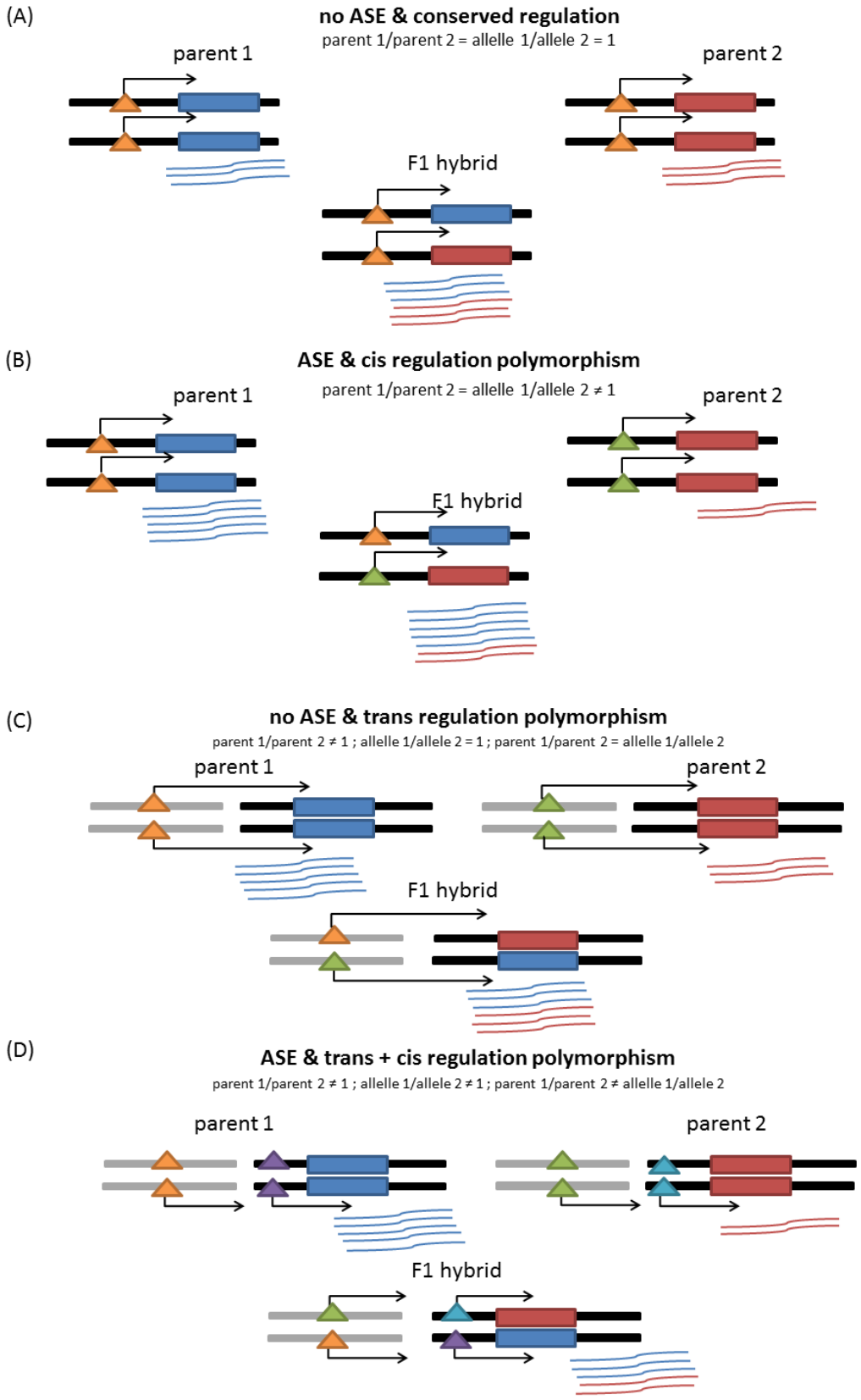


Figure 17. Allele specific expression (ASE), *cis* acting and *trans* acting regulation identification using RNA sequencing data in F1 hybrid compared with the parental lines. Triangles represent polymorphisms responsible for differential gene expression, either upstream of the regulated gene or away from the regulated genes.

recovering the causal locus FRIGIDA (previously cloned by Lovell *et al.* , 2013) among 92 cis-regulated genes in the confidence interval of a QTL for water use efficiency (WUE). This approach will be worth extending in several species to facilitate the identification of causal genes underlying QTLs and QTL x E interactions.

2.4.2 Allele specific expression and RNA sequencing data

Monitoring allele specific expression (ASE) in F1 hybrid is an alternative method to identify regulatory polymorphism at the whole genome scale. This requires allele specific expression data in an F1 hybrid to be compared to expression data in its parental accessions. The recent development of the RNA sequencing technologies makes it possible to extend the approach in a whole genome fashion, combining RNA sequencing and exonic SNP carried out by parental accessions to assign the reads to a specific allele (Knight, 2004; Castel *et al.* , 2015).

In F1 hybrid, both parental alleles are in the same cellular context and exposed to a common set of regulatory factors. In such context, allelic expression can be altered due to *cis*-regulatory divergence between parental species. Besides, allele specific expression in a hybrid compared with allelic expression in its parents makes it possible to identify *trans*-regulatory divergences by comparing the ratio of expression of the two parental alleles in the hybrid with the relative expression of the same alleles in its parents. A similar and balanced allelic expression between parents and hybrid indicates a conserved regulation **(Figure 17 A)**, whereas a conserved unbalanced allelic expression between parents and hybrid is the signature of parental *cis*-regulatory divergences only **(Figure 17 B)**, and a balanced allelic expression only in hybrid revealed parental *trans*-regulatory divergences **(Figure 17 C)**. Different unbalanced allelic expression between parents and hybrid may reflect a combination of both *cis* and *trans* acting effects **(Figure 17 D)**.

Today, a few studies have reported allele specific expression in plant (see Cubillos *et al.* , 2014 and He *et al.* , 2016 in *Arabidopsis*, Combes *et al.* , 2015 in coffee, Verta *et al.* , 2016 in white spruce and Stupar and Springer, 2006; Springer and Stupar, 2007 in maize). Both approaches, eQTL mapping and monitoring allele specific expression in hybrids, are complementary in revealing the genetic basis of differential expression. The first approach can have limited power to detect low effect polymorphisms depending on the population

Table 2. Species list for Solanum section Lycopersicon. *Adapted from Peralta et al. , 2005.*

***Solanum* name**	**Distribution and habitat**
S. arcanum Peralta	North Peru, coastal and inland Andean valleys; lomas, dry valleys and dry rocky slopes; 100 to 2800 m.
S. cheesmaniae (L. Riley) Fosberg	Endemic to the Galapagos Islands, Ecuador; wide variety of habitats; sea level to 500 m.
S. chilense (Dunal) Reiche	South Peru (Tacna) to north Chili (région II); in hyper-arid rocky plains and coastal deserts; sea level to 3250 m.
S. chmielewskii (C.M. Rick, Kesicki, Fobes & M. Holle) D.M. Spooner, G.J. Anderson & R.K. Jansen	South Peru (Apurímac) to north Bolivia (La Paz); high dry Andean valleys; 1600 – 3200 m.
S. corneliomuelleri J.F. Macbr.	Central to south Peru, west slops of the Andes; landslides and rocky slopes; 200 – 3300 m.
S. galapagense S. Darwin & Peralta	Endemic to the Galápagos Islands; mostly occurring on coastal lava to within 1 m of high tide mark within range of sea spray, but occasionally inland; sea level to 50 m.
S. habrochaites S. Knapp & D.M. Spooner	Central Ecuador to central Peru, on the western slopes of the Andes; in a variety of forest types from premontane forests to dry forests; 200 – 3300 m.
S. huaylasense Peralta	North Peru (Ancash); rocky slopes of the Callejón de Huaylas along the Río Santa and in the adjacent Río Fortaleza drainage; 1700 – 3000 m.
S. lycopersicum L.	Known only from cultivation or escapes; worldwide in a variety of habitats, many escaped plants have smaller fruits ("*cerasiforme*"); sea level to 4000 m.
S. neorickii (C.M. Rick, Kesicki, Fobes & M. Holle) D.M. Spooner, G.J. Anderson & R.K. Jansen	South Ecuador (Azuay) to south Peru (Apurímac); dry inter-Andean valleys, often found trailing over rocky banks and roadsides; 1950 – 2600 m.
S. pennellii Correll	North Peru (Piura) to north Chile (Tarapacá); dry rock hillsides and sandy areas; sea level to 2300 m.
S. peruvianum L.	Central Peru (Ancash) to north Chile (région II); coastal lomas formations and occasionally in coastal deserts, occasionally as a weed at field edges in coastal river valleys; sea level to 600 m.
S. pimpinellifolium L.	Central Ecuador to central Chile; dry coastal habitats; 0 – 500 m, but exceptionally up to 1400 m.

design used **(see 2.2).** Furthermore, eQTL mapping requires the collection of gene expression data on a whole mapping population, which can be limiting to extend the approach to the whole genome level. The second approach restricts the identification of regulation pattern to genes presenting molecular polymorphisms in coding sequences between parental accessions, which could be limiting in intraspecific crosses, depending on the level of genetic diversity of the species. Besides, allele specific expression framework does not allow localizing *trans*-acting regulatory factors on the contrary to eQTL mapping and needs adapted processing of the sequencing data to avoid any biases (Castel *et al.* , 2015).

3. Genetics and breeding of tomato sensory quality

Part of this section has been accepted for publication by Autar Mattoo and Avtar Krishan Handa in their book "*Achieving sustainable tomato cultivation",* as chapter XIII, entitled "*Developing tomato varieties with improved flavor*", by Mathilde Causse, Elise Albert and Christopher Sauvage **(provided in Appendix 2)**.

3.1 A brief overview of tomato plant biology, origin and economic importance

In the *Solanaceae* family, the *Solanum* genus gathers the cultivated tomato (*Solanum lycopersicum*) and its 12 wild relative species originated from South America, in the Andes Mountains of Peru, Ecuador and Chile, in the section *Lycopersicon* **(Table2)** (Peralta *et al.* , 2005). All the species of this genus are diploid and have a similar chromosomal structure (2n=2x=24), but vary in their mating system, from autogamous and self-compatible (e.g. cultivated tomato), to facultative allogamous, to allogamous and self-incompatible (Rick, 1979). The domestication of the cultivated tomato is considered to result from a recent divergence from the red fruited species *S. pimpinellifolium,* first in Peru and then in Mexico, leading to a significant increase in fruit size **(Figure 18)** (Jenkins, 1948; Nesbitt and Tanksley, 2002). Several historical and genetic evidences designate *S. lycopersicum* var. *cerasiforme* as a transitional form between the *S. pimpinellifolium* wild species and the *S. lycopersicum* cultivated one (Ranc *et al.* , 2008; Blanca *et al.* , 2012, 2015). In the early XVIth century, the cultivated tomato was introduced to Europe where the species experienced a considerable boom (Robertson and Labate, 2007; Bauchet and Causse, 2010).

Nowadays, tomato is one of the most consumed vegetable through the world, accounting for 14% of the world vegetable production (FAOSTAT, 2012). Its production consistently

Figure 18. Large phenotypic variations among cultivars and wild relative species of tomato for fruit size, shape, and color. Variations are present both for immature fruit color ranging from pale to dark green and for mature fruit color ranging from yellow-green in small-fruited species such as wild tomato (*S. peruvianum*) to red, pink, orange and yellow (*S. pimpinellifolium* and *S. lycopersicum*). Domestication of tomato was accompanied by a dramatic increase in fruit size. *(Adapted from Koornneef and Stam, 2001)*

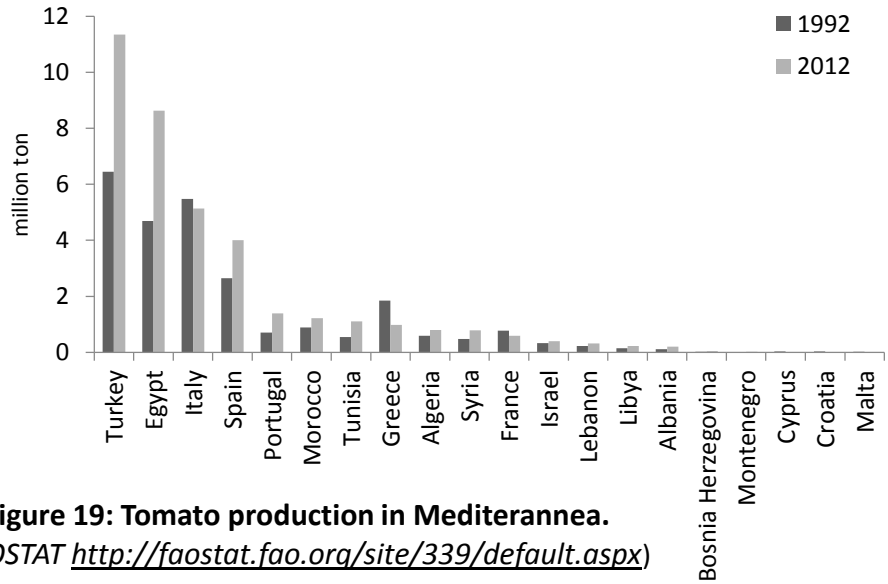


Figure 19: Tomato production in Mediterannea.
(FAOSTAT http://faostat.fao.org/site/339/default.aspx)

increased in the last decade and was estimated at 162 million tons in 2012 (FAOSTAT, 2012). The world production is leaded by China (~31%), India (~11%) and the USA (~8%). In Europe, the countries of the Mediterranean region produced a total of 37 million tons of tomatoes in 2012 and Turkey concentrated 30% of the production **(Figure 19).** Two thirds of the European productions are grown in open field, mechanically harvested and intended for processing industries (puree, concentrate, juice and canned tomato), while the remaining third corresponds to fresh tomato, grown in greenhouses, handpicked and sold on the fresh market (EUROSTAT, 2015).

3.2 Tomato genome sequence and genomic resources

Due to its relatively short generation time and simple growing requirements, tomato has been used in scientific research for a hundred years, serving as model for fleshy fruit development study, quantitative trait locus mapping and development of breeding technologies (Paterson *et al.* , 1988; Tanksley *et al.* , 1992). The tomato genome was completely sequenced in 2012 (~9 Mb, Tomato Genome Consortium, 2012), followed by the release of genotyping arrays allowing to genotype simultaneously thousands of SNP at a genome-wide scale (Hamilton *et al.* , 2012; Sim *et al.* , 2012). Then, resequencing projects were achieved, not only in the cultivated tomato (Aflitos *et al.* , 2014; Lin *et al.* , 2014) but also in its wild relatives (i.e. *Solanum pennellii*, Bolger *et al.* , 2014). A new step in the knowledge of tomato genetics is currently running with the so called 'omics' technologies offering the possibility to access to different levels of variability (phenome, genome, transcriptome, proteome and metabolome) in a high throughput way. Most of the data derived from these technologies are made publicly available through databases and websites (e. g. SGN: https://solgenomics.net/, Mueller *et al.* , 2005), supporting the breeding efforts and the research on the genetic basis of complex traits and their interaction with the environment.

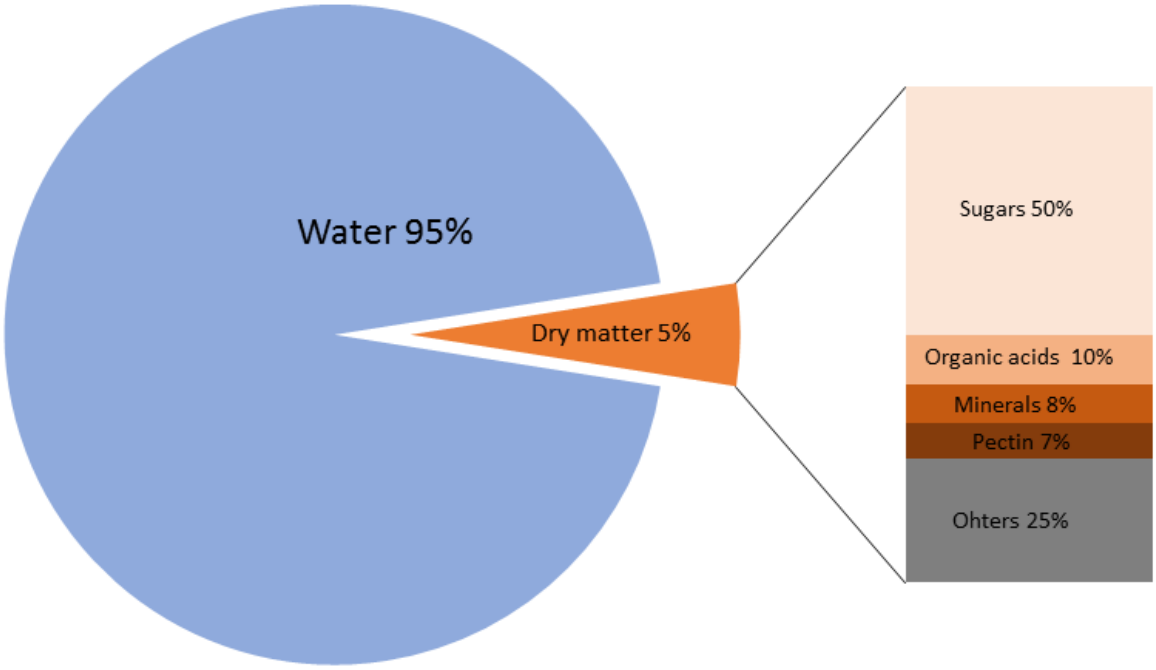


Figure 20. Composition of mature tomato fruit.
(Davies *et al.* , 1981)

3.3 Tomato fruit sensory quality

Over the XXth century, tomato breeders have focused on improving the species for yield and yield stability, adaptation to various growth conditions, disease resistances, conservation and appearance of the fruit (diversification of color and shape). The sensory quality was mostly ignored, causing consumer dissatisfaction in the early 90's (Bruhn *et al.* , 1991). Tomato fruit is mainly composed of water (95%), and 5 % of dry matter, which comprises around 50% sugars (fructose and glucose), 10% organic acids (citric and malic acids, 8% minerals, 7% pectin and 25% of other secondary metabolites (Davies *et al.* , 1981) **(Figure 20)**.

Sensory quality for fresh market tomato is a composite trait determined by both external (size, color, and firmness) and internal (flavor, aroma, texture) characteristics. Besides, sensory quality has a subjective component based on every consumer preferences. The relationships between fruit characteristics and composition and tomato taste have been widely studied (Causse *et al.* , 2003, 2010; Chaïb *et al.* , 2007). Sugar content, acids and their ratio play an important role in determining fruit flavor (Stevens *et al.* , 1977; Stevens, 1979; Bucheli *et al.* , 1999). Sugar and acid contents are related to sweetness and sourness (Stevens *et al.* , 1977) and contribute to sweetness and the overall aroma intensity (Baldwin *et al.* , 2000, 2008). Hundreds of volatile compounds have been identified but the list of those important in tomato aroma is limited to about twenty (Buttery *et al.* , 1989; Baldwin *et al.* , 2000; Klee and Tieman, 2013). Some of the most abundant volatiles do not contribute to consumer liking, whereas other less abundant ones do (Tieman *et al.* , 2012). Besides, certain aroma volatiles make contributions to perceived sweetness independent of sugar concentration. Texture traits are more difficult to relate to instrumental measurements, although firmness perceived when eating is partly related to compression tests (Causse *et al.* , 2001; Chaïb *et al.* , 2007). Because some important components of sensory quality are negatively correlated, like yield and sugar content or fruit shelf life and meltiness (due to physiological and genetic origin), improving simultaneously multiple tomato quality traits is challenging.

Consumer preferences facing genetic diversity have been subject to a few studies (Sinesio *et al.* , 2009; Causse *et al.* , 2010). In the framework of a large European project, Eusol, 806

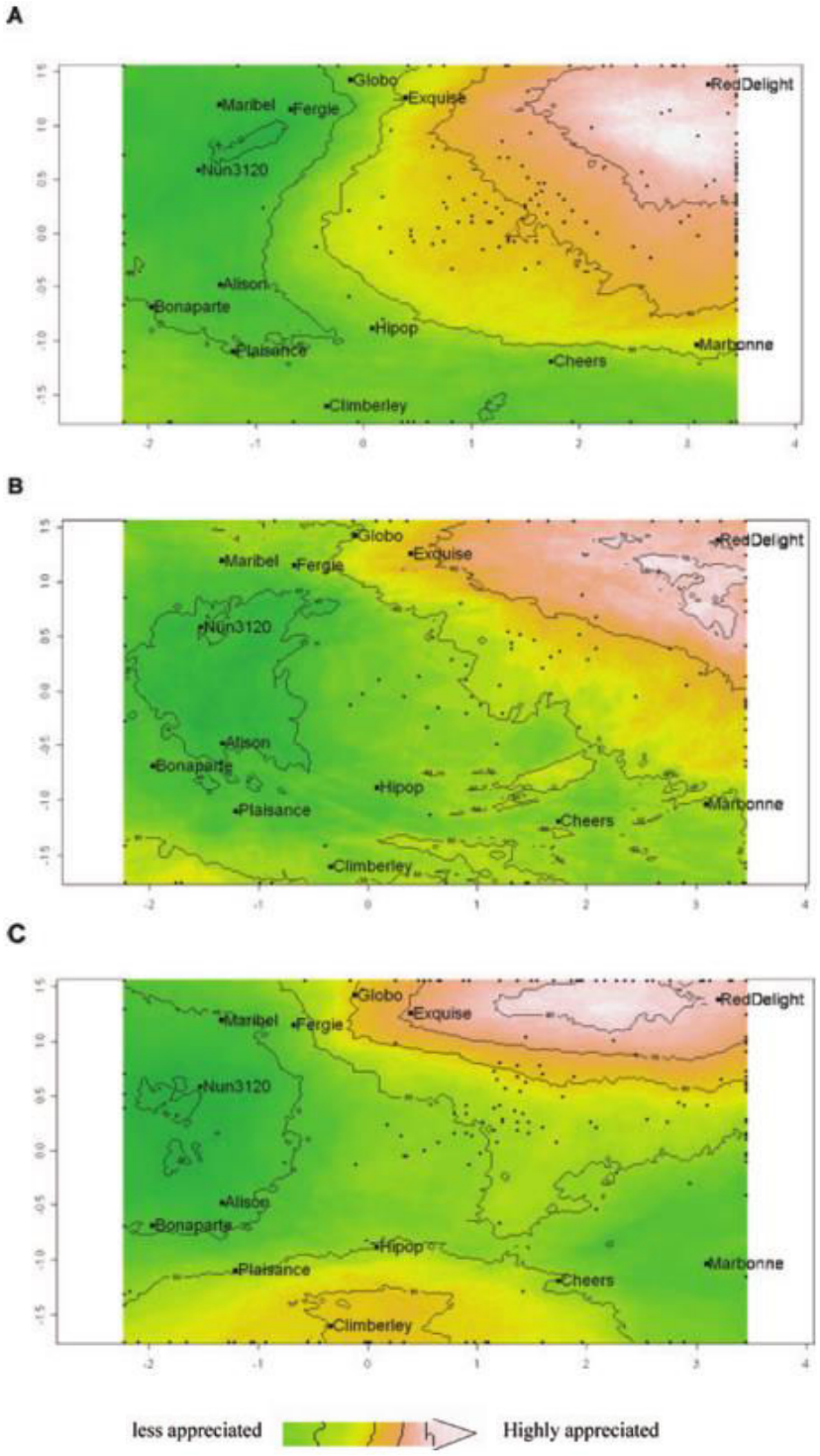


Figure 21. Contour plot of the external preference map of consumers based on a multiple factorial analysis of the sensory profiles from three countries (France, Italy and Netherlands). A total of 16 varieties were evaluated by three consumer panels using various sensory descriptors relative to fruit flavor and basic tastes, texture, odor, appearance and aftertaste. The isolines correspond to the percentage of consumers who gave a score higher than average. Small dots correspond to consumers. Variety names are indicated on each plot. A: French consumers, B: Italian consumers, and C: Dutch consumers.

(Adapted from Causse et al. , 2010)

consumers from three countries (The Netherlands, France, and Italy) were presented with a set of 16 varieties representing the diversity of fresh tomato offer in order to evaluate their preferences. In parallel, expert panels in each country built sensory profiles of the varieties. Preference maps were then constructed in each country revealing the structure of consumer preferences and allowing identification of the most important characteristics. Then a global analysis revealed that preferences were quite homogeneous across countries **(Figure 21).** This study identified the overall flavour and firmness as the most important traits for improving tomato fruit quality. It showed that consumer preferences from different European countries, with different cultures and food practices, are segmented following similar patterns when projected onto a common referential plan. Moreover, the results clearly showed that diversification of taste and texture is required to satisfy all consumers' expectations, as some consumers preferred firm tomatoes, while others preferred melting ones and were more or less demanding in terms of sweetness and flavour intensity. Detailed comparisons also showed the importance of the fruit appearance in consumer preference.

3.4 Genetic determinants of tomato fruit sensory quality

Many mutations involved in fruit development and composition have been discovered and used for fruit quality breeding. **Table 3** lists the major mutations identified, which are directly or indirectly impacting fruit quality. They may induce variation in fruit color or aspect. Some mutations impacting plant architecture, like *sp* which is controlling the determinate/indeterminate growth, are also known to impact fruit quality (Pnueli *et al.* , 1998). Causse *et al.* (2003) have produced and compared seven pairs of nearly isogenic hybrids, with or without the *rin* mutation at the heterozygous level. The presence of the *rin* mutation reduced the consumer preference. Differences were detected by sensory profiles, *rin* hybrids having fruits on average 17% less sweet, with a lower tomato aroma, a higher ‘strange’ aroma and more mealy fruits, although instrumental firmness and sugar content were not different. These results enlighten the negative influence of the *rin* mutation on consumer preference, but also indicated that when transferred into a hybrid with high flavour, the negative influence of the mutation is reduced. Selection could thus be carried out to obtain much sweeter and perfumed lines combined with shelf life in *rin* hybrids.

Table 3. Cloned genes with a phenotype related to fruit quality, plant, leaf or truss architecture; location on the tomato genome assembly.

ITAG gene model	Gene Symbol	Locus name	Chr.	Start position	Phenotypic descriptors	References
Solyc01g079620	*y*	*colorless epidermis*	1	71 255 600	pink epidermis	(Ballester *et al.* , 2010)
Solyc10g081470	*L-2*	*Lutescent-2*	10	61 858 478	altered chloroplast development and delayed ripening	(Barry *et al.* , 2012)
Solyc08g080090	*Gr*	*green flesh*	8	60 582 066	green fruit flesh	(Barry *et al.* , 2008)
Solyc06g074910	*C*	*potato leaf*	6	42 804 036	simple leaves	(Busch *et al.* , 2011)
Solyc03g031860	*r*	*Phytoene synthase 1*	3	8 606 749	yellow fruit	(Fray and Grierson, 1993)
Solyc04g082520	*cwp1*	*cuticular water permeability 1*	4	63 765 366	microfissure/dehydration of fruits	(Hovav *et al.* , 2007)
Solyc10g081650	*t*	*carotenoid isomerase*	10	62 006 972	orange fruit flesh	(Isaacson *et al.* , 2002)
Solyc02g077390	*s*	*compound inflorescence*	2	36 913 957	Inflorescence branching	(Lippman *et al.* , 2008)
Solyc02g077920	*Cnr*	*Colourless non-ripening*	2	37 323 107	Inhibition of ripening	(Manning *et al.* , 2006)
Solyc11g010570	*j*	*jointless*	11	3 640 857	no pedicel abscission zone	(Mao *et al.* , 2000)
Solyc03g118160	*fa*	*falsiflora*	3	61 162 449	leafy inflorescence	(Molinero-Rosales *et al.* , 1999)
Solyc03g063100	*sft*	*single flower truss*	3	30 564 833	single flower truss	(Molinero-Rosales *et al.* , 2004)
Solyc01g056340	*hp-2*	*de-etiolated 1*	1	46 495 644	high pigment	(Mustilli *et al.* , 1999)
Solyc06g074350	*sp*	*self-pruning*	6	42 361 623	determinate plant habit	(Pnueli *et al.* , 1998)
Solyc10g008160	*u*	*uniform ripening*	10	2 293 088	increased chlorophyl content	(Powell *et al.* , 2012)
Solyc12g008980	*Del*	*Delta*	12	2 285 372	orange fruit	(Ronen *et al.* , 1999)
Solyc06g074240	*B*	*Beta-carotene*	6	42 288 127	increased fruit beta-carotene	(Ronen *et al.* , 2000)
Solyc03g083910	*sucr*	*sucrose accumulator*	3	47 401 871	Accumulates predominantly sucrose in mature fruit, rather than glucose and fructose	(Sato *et al.* , 1993)
Solyc07g066250	*ls*	*lateral suppresser*	7	64 958 148	Few or no axillary branches; corolla suppressed; partially male sterile	(Schumacher *et al.* , 1999)
Solyc02g090890	*hp-3*	*zeaxanthin epoxidase*	2	46 947 557	high pigment in fruits	(Thompson *et al.* , 2000*a*)
Solyc05g012020	*rin*	*ripening inhibitor*	5	5 217 073	never ripening	(Vrebalov *et al.* , 2002)
Solyc05g012020	*mc*	*macrocalyx*	5	5 217 073	large sepals	(Vrebalov *et al.* , 2002)
Solyc05g053410	*phyB2*	*apophytochrome B2*	5	62 648 223	red light reception	(Weller *et al.* , 2001)
Solyc10g044670	*phyA*	*apophytochrome A*	10	22 854 459	far red light insensitive	(Weller *et al.* , 2001)
Solyc09g075440	*Nr*	*Never ripe*	9	62 631 866	not ripening	(Wilkinson *et al.* , 1995)
Solyc04g076850	*e*	*Entire leaf*	4	59 354 677	reduced leaf complexity	(Zhang *et al.* , 2007)

Most tomato fruit quality traits are quantitatively inherited. For this reason, tomato was among the first crop for which molecular markers were used to dissect the genetic basis of quantitative traits into QTLs (*Quantitative Trait Loci)* with the pioneer work of Tanksley, (1993). Then, many QTLs controlling tomato yield and fruit quality related traits have been mapped in the linkage framework (Paterson *et al.* , 1988, 1991; DeVicente and Tanksley, 1993; Azanza *et al.* , 1994; Eshed and Zamir, 1995; Goldman *et al.* , 1995; Tanksley *et al.* , 1996; Grandillo and Tanksley, 1996; Fulton *et al.* , 1997, 2000, 2002, Bernacchi *et al.* , 1998*a*,*b*; Chen *et al.* , 1999; Grandillo *et al.* , 1999; Saliba-Colombani *et al.* , 2001; Causse *et al.* , 2001, 2007; Doganlar *et al.* , 2002; Frary *et al.* , 2004; Tieman *et al.* , 2006; Semel *et al.* , 2006; Zanor *et al.* , 2009*b*; Capel *et al.* , 2015), benefiting from the rapid progress of the genetic markers from isozymes, to RFLPs (Restriction Fragment Length Polymorphisms), to RAPDs (Random Amplification of Polymorphic DNA), to microsatellites, to SNPs (Single-Nucleotide Polymorphisms). Due to the very low polymorphism revealed at the within species level, most of these studies were performed on inter-specific progenies derived from crosses between wild tomato species and the cultivated tomato (except the work from Causse *et al.* , 2001; Saliba-Colombani *et al.* , 2001). In most of the studies a few QTLs explained a large fraction (20 to 50%) of the phenotypic variation, acting in concert with minor QTLs that could not be detected. Most of the QTLs acted in an additive manner, but dominant and overdominant QTL have been detected (Paterson *et al.* , 1988, 1991; DeVicente and Tanksley, 1993; Semel *et al.* , 2006). Epistasis (interaction among QTLs) was rarely detected unless a specific experimental design was used (Tanksley *et al.* , 1996; Causse *et al.* , 2007).

The recent sequencing of tomato genome (**see 3.2**), followed by the release of genotyping arrays allowing to genotype simultaneously thousands of SNP at a genome-wide scale, have paved the way to use more complex QTL mapping designs, at the interspecific and intraspecific levels, to close the gap between phenotype and genotype. Using a set of 192 SNP markers genotyped in 188 accessions, Xu *et al.* (2013) identified 2, 16 and 17 loci associated to titrable acidity, soluble solids and sugar contents respectively, demonstrating the feasibility of genome wide association mapping (GWA) for tomato fruit quality traits. Favorable allelic combination between loci associated to fruit quality, such as pH, titrable acidity, SSC or fruit shape were identified using linear mixed models in collections of tomato

Table 4. Overview of the genetic architecture of major tomato fruit quality traits.

Traits	QTLs identified	Identified genes
Fresh weight	**Up to 28 QTLs**, with **6 QTLs explaining more than 20% of the phenotypic variations** (reviewed in Grandillo *et al.* , 1999)	*fw2.2* on chr. 2 (Frary, 2000) *fw3.2* on chr. 3 (Chakrabarti *et al.* , 2013)
Locule number	**Multiple QTLs**, with **2 majors loci** (Lippman and Tanksley, 2001; van der Knaap and Tanksley, 2003; Barrero and Tanksley, 2004)	*lc* on chr. 2 (Muños *et al.* , 2011) *fas* on chr. 11 (Cong *et al.* , 2008; Xu *et al.* , 2015)
Fruit shape	**Up to 11 QTLs**, with **3 major QTLs** explaining most of the variations (reviewed in Grandillo *et al.* , 1999; Brewer *et al.* , 2007)	*ovate* on chr. 2 (Liu *et al.* , 2002) *sun* on chr. 7 (Xiao *et al.* , 2008) *fs8.1* on chr. 8 (Sun *et al.* , 2015)
Sugar content	**Up to 95 QTLs** in 56 chromosomal regions, often **pleiotropic effects on fruit size and sugar content** (reviewed in Labate *et al.* , 2007)	*Lin5* on chr. 9 (Fridman *et al.* , 2000; Zanor *et al.* , 2009*a*)
Fruit firmness	**Up to 56 QTLs** grouped in clusters on chr. 1, 2 , 4, 5, 9, 10 & 11 (reviewed in Labate *et al.* , 2007)	*rin* on chr. 5 (Vrebalov *et al.* , 2002; Ito *et al.* , 2008) *Fir* on chr. 2 (Chapman *et al.* , 2012)
Volatile compounds	**More than 30 QTLs, with few QTLs common between experiments** (Saliba-Colombani *et al.* , 2001; Tieman *et al.* , 2006; Mathieu *et al.* , 2009)	*ADH* (Speirs *et al.* , 1998) *AADC* (Tieman *et al.* , 2006) *PAR* (Tieman *et al.* , 2007) *LoxC* (Chen *et al.* , 2004) *SAMT* (Tieman *et al.* , 2010) *CTOMT* (Mageroy *et al.* , 2012) *CXE1* (Goulet *et al.* , 2012) *CCD1* (Simkin *et al.* , 2004) *GT1* (Tikunov *et al.* , 2013)

cultivars and landraces genotyped for thousands of SNPs (Ruggieri *et al.* , 2014; Sacco *et al.* , 2015). By using a multi-locus mixed model (Segura *et al.* , 2012), Sauvage *et al.* , (2014) provided an extended list of loci associated to important metabolic compounds for flavour (e.g. fructose, SSC and malic and citric acids), demonstrating the interest of multi locus model for revealing the genetic determinants of the highly polygenic fruit quality traits. New types of populations involving several parental lines like MAGIC (Multi-allelic Genetic Intercross) have also been demonstrated useful in tomato to map QTLs for quality traits into small confidence intervals. Combined with the re-sequencing of the parental lines, a direct access to putative polymorphisms under the QTLs could be proposed (Pascual *et al.* , 2015). Recently, Ofner *et al.* (2016) developed a population of 446 backcross inbred lines (BILs) derived after a few generations of backcrosses of the wild species *S. pennellii* with the cultivated tomato, followed by more than seven generations of self-pollination. This genetic material was genotyped for ten thousand SNP markers using a genome wide SNP array and should constitute a fantastic tool to confirm and fine-map QTLs for tomato quality traits in the near future.

An article comparing three population designs for quantitative trait locus (QTL) mapping in tomato, namely recombinant inbreed lines (RILs), genome wide association collection (GWA) and multiple advanced generation inter-cross (MAGIC) using thousands of SNP markers, was achieved in parallel to the main work of this thesis and published in *Plant Sciences* in 2016. The article was entitled “Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels”, by Laura Pascual, Elise Albert, Christopher Sauvage, Janejira Duangiit, Mathilde Causse and co-authors **(provided in Appendix 3).**

Table4 summarizes the genetic architecture of some major tomato fruit quality traits. More details concerning the different QTLs and genes are provided in **Appendix 2**.

3.5 Deficit irrigation strategies as a tool to manage tomato sensory quality

Tomato plants are highly water demanding, mostly produced under irrigation and consuming up to 360 liters of water to produce one kilogram of harvested fresh matter (FAO WATER, 2016) **(Table 5).** They are produced year-round under contrasting environmental conditions, triggering seasonal variations in their sensory quality. Over the tomato growing cycle,

Table 5: Water needs to produce one kilogram of harvested fresh and dry matter for major Mediterranean crop products. *(FAO WATER http://www.fao.org/nr/water/cropinfo.html)*

Species	Water needs for 1kg of harvested **fresh matter** (l)	Water needs for 1kg of harvested **dry matter** (l)	Dry matter content of the harvested product (%)
Tomato	360	1,800	5
Pepper	400	4,000	10
Potato	160	533	30
Grape	330	1,650	20
Olive	660	943	70
Citrus	330	2,200	15
Watermelon	154	1,540	10
Wheat	625	710	88

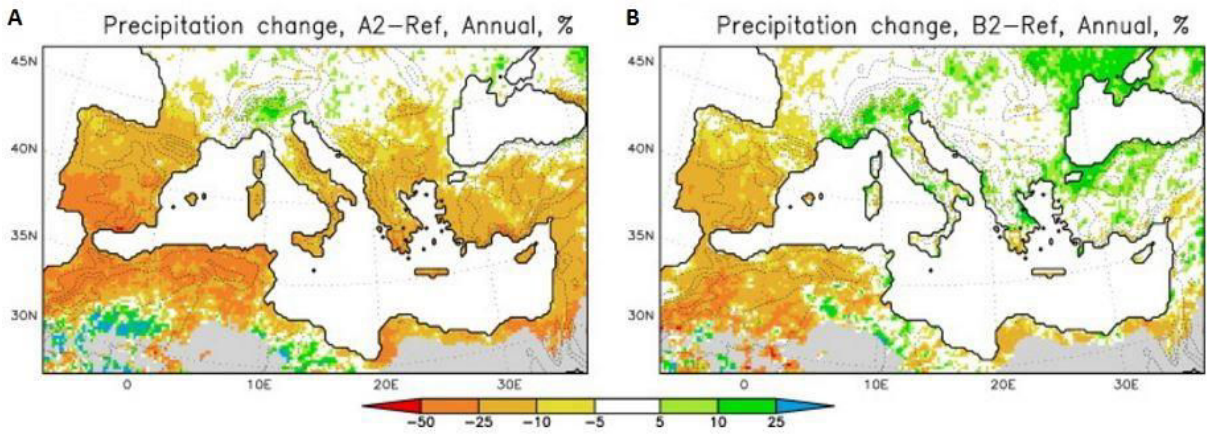


Figure 22. Mean precipitation change in the Mediterranean region (2071–2100 minus 1961–1990). (A) Under the A2 scenario: maintained greenhouse gas production. (B) Under the B2 scenario: reduced greenhouse gas production. *(Adapted from Gao and Giorgi, 2008)*

different factors such as light intensity, air and soil temperatures, plant fruit load, plant mineral nutrition or water availability influence the final fruit quality (reviewed in Davies *et al.* , 1981; Poiroux-Gonord *et al.* , 2010). Water limitation and irrigation with saline water may impact positively tomato fruit quality, mainly through an increase in sugar content in fruit (either by concentration or accumulation effect) and contrasted effects on the secondary metabolite contents (Paterson *et al.* , 1988; Mitchell *et al.* , 1991; Pascale *et al.* , 2001; Nuruddin *et al.* , 2003; Gautier *et al.* , 2008; Ripoll *et al.* , 2016*a*,*b*). The effects reported on fruit composition are associated or not to large yield loss depending upon the intensity and duration of the treatment and the development stage of the plant (see Ripoll *et al.* , 2014 for review) and result from modifications of the water and carbon fluxes imported by the fruit during its growth (Guichard *et al.* , 2001; Albacete *et al.* , 2014; Osorio *et al.* , 2014).

In the context of the global warming and limited water availability **(Figure 22)**, the optimization of water management practices is considered in horticultural production as a tool to manage fruit quality while limiting yield losses, offering the opportunity to address simultaneously environmental issues and consumer expectations of tastier fruits (Stikic *et al.* , 2003; Fereres and Soriano, 2006; Costa *et al.* , 2007). Deficit irrigation consists in water supply below the evapotranspiration demand (Fereres and Soriano, 2006). This irrigation strategies rests on the physiological, biochemical and physical changes undergone by plants under water limitation (Sultan, 2000).

3.6 Genetic determinants of tomato response to water deficit

Large phenotypic variation in response to a wide range of climate and nutrition conditions exists in the genus *Solanum* at both inter and intra species levels (reviewed in Labate *et al.* , 2007). The TGRC (Tomato Genetics Resource Center, UC Davis) maintains wild and cultivated accessions with known or inferred tolerances to various abiotic stresses, including drought, flooding, high temperature, chilling injury, aluminum toxicity, salinity and/or alkalinity, providing useful starting material for breeding, genetic mapping, and other uses.

Several authors attempted to measure genotype by environment interactions on tomato fruit quality by repeating a same experiment in different locations or/and under several growing facilities (Auerswald *et al.* , 1999; Johansson *et al.* , 1999; Causse *et al.* , 2003) or by

Table 6. Examples of genes used to transform tomato with functions related to response to abiotic stresses.

Gene	Annotation	Function	Reference
Trans genesis: genes from another species into tomato			
atnhx1	vacuolar Na^+/H^+ antiporter, from *Arabidopsis thaliana*	Salt tolerance, growth, fruit yield	(Apse *et al.* , 1999; Zhang and Blumwald, 2001)
badh-1	betaine aldehyde dehydrogenase, from sorghum	Osmotic adjustment	(Moghaieb *et al.* , 2000)
CBF1	DREB transcription factor, from *Arabidopsis thaliana*	Elevated tolerance to chilling and oxidative stress	(Hsieh *et al.* , 2002)
AVP1	vacuolar H^+ pyrophosphatase, from *Arabidopsis thaliana*	More robust root systems and improved resistance to water deficit	(Park *et al.* , 2005)
TPS1	trehalose-6-phosphate synthase, from yeast	Water, salt and oxidative stress tolerance	(Cortina and Culiáñez-Macià, 2005)
Tbosm	osmotin gene, from tobacco	Enhanced salt and water stress tolerance, higher chlorophyll and proline content	(Goel *et al.* , 2010)
ATHB-7	homeodomain-leucine zipper transcription factor, from *Arabidopsis thaliana*	Reduced stomatal density and pore size, tolerance to water deficit	(Mishra *et al.* , 2012)
Cis genesis: gene from tomato species into tomato			
tos1	increased ABA sensitivity	Hypersensitive to osmotic stress and exogenous ABA	(Borsani *et al.* , 2002)
SlAREB1 & SlAREB2	leucine zipper transcription factor	Water, salt stress tolerance	(Orellana *et al.* , 2010; Hsieh *et al.* , 2010)
H1-S	linker histone	Water stress tolerance	(Scippa *et al.* , 2004)
ASR1	ABSCISIC ACID RIPENING 1	Enhanced survival under water stress	(Golan *et al.* , 2014)
LeNCED	9- cis-epoxycarotenoid dioxygenase	Higher ABA content and improved water use efficiency	(Thompson *et al.* , 2000*b*; Tung *et al.* , 2008)

building experimental design to isolate the effect of particular environmental factors on large number of genotypes (Semel *et al.* , 2007; Gur *et al.* , 2011 for water availability and Monforte *et al.* , 1996, 1997*a*,*b* for salt stress). In the different experiments, the G x E interaction was significant for the fruit quality traits measured (including fruit fresh weight, secondary and primary metabolism contents and fruit firmness), but generally accounted for a low part of the total variation in comparison to the genotype main effect.

Few QTL studies considering the interaction with environmental variables at the fruit level were reported. They mainly concerned response to salt stress (Monforte *et al.* , 1996, 1997*a*,*b*; Uozumi *et al.* , 2012; Asins *et al.* , 2015) and drought stress (Gur *et al.* , 2011). All these studies identified numerous loci with low to medium effect suggesting a strongly polygenic architecture of tomato fruit response to environmental constraints. Gur *et al.* , (2011) described drought responsive QTLs for fruit fresh weight and sugar content mainly expressed by the shoot in a reciprocal-grafting experiment whereas (Monforte *et al.* , 1997*b*) identified QTLs with changing additive and epistatic effects according to the salinity level of the watering solution. Nevertheless, the authors mostly compared QTLs at different map positions and with different effects across experiments and conditions, which may be questionable as these comparisons depend upon the mapping significance threshold. Besides, populations used were mainly introgression lines involving wild relative species (*Solanum habrochaites*, *Solanum pennellii* and *Solanum pimpinellifolium*) and the confidence intervals obtained remain large and difficult to transpose into the cultivated tomato.

In parallel to QTL mapping, many genes whose expression is increased in response to drought or salt stresses in tomato and/or others species have been cloned and characterized. Transgenic modification of these genes has been used for two different objectives, to determine the function of the protein or gene product in the stress response and to attempt to confer stress tolerance on the transgenic plants. **Table 6** gives examples of transformation experiments achieved in tomato and resulting in improving tolerance to water deficit and others abiotic stresses. Nevertheless, as response to water deficit is complex trait, transgenic plants engineered to express a single gene do not seem likely to result in a robust stress tolerant phenotype (Sultan, 2000; Flowers, 2004). Besides, constitutive overexpression of water deficit responsive genes was often shown to cause deleterious effects. One solution to overcome this problem could be the use of stress

inducible and/or organ specific promoters to restrict the expression of these genes to stress conditions and/or to specific organs.

4. Context and objectives of the thesis

This thesis is included in the TOMSEC (CASDAR CTPS, 2013 - 2015) and AdapTom (ANR, 2014 - 2017) projects which aimed at identifying phenotypes, QTLs, genes and alleles that will enable to maintain tomato yield and improve fruit quality under conditions of limited water availability. These projects rely on the complementary expertises of three laboratories:

- UR1052 '*Génétique et Amélioration des Fruits et Légumes*' (INRA PACA Avignon): quantitative genetic & genomics (project coordination : Mathilde CAUSSE),
- UR1115 '*Plantes et Systèmes de Culture Horticoles*' (INRA PACA Avignon): physiology & ecophysiology,
- UMR1332 '*Biologie du Fruit et Pathologie*' (INRA Bordeaux Aquitaine) : functional genetic & genomics.

The seed companies '*Gautier semences'* and '*Vilmorin & cie'* were partners in both projects and contributed in collecting the genotypic and phenotypic data.

The scientific objectives of this thesis were:

(1) To characterize, **at the phenotypic level**, the patterns of **genotype by watering regime interaction** in diverse tomato accessions phenotyped for various agronomic traits.

(2) To draw an accurate picture of the **QTL by watering regime interactions** and localize the **QTL and genes** involved in tomato phenotypic response to water deficit, using a combination of **linkage and association mapping**.

(3) To identify **differentially expressed genes** and **allele specific expression** in response to water limitation **at whole genome scale** through **RNA sequencing**.

(4) To localize **eQTL and genes** involved in controlling tomato **gene regulation** under water limitation using **microfluidic qPCR** in a biparental mapping progeny.

References

A

Aflitos S, Schijlen E, De Jong H, *et al.* 2014. Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. Plant Journal 80, 136–148.

Albacete A, Martínez-Andújar C, Pérez-Alfocea F. 2014. Hormonal and metabolic regulation of source-sink relations under salinity and drought: From plant survival to crop yield stability. Biotechnology Advances 32, 12–30.

Allard A, Bink MCAM, Martinez S, Kelner JJ, Legave JM, Di Guardo M, Di Pierro EA, Laurens F, Van De Weg EW, Costes E. 2016. Detecting QTLs and putative candidate genes involved in budbreak and flowering time in an apple multiparental population. Journal of Experimental Botany 67, 2875–2888.

Alves AAC, Setter TL. 2004. Response of cassava leaf area expansion to water deficit: Cell proliferation, cell expansion and delayed development. Annals of Botany 94, 605–613.

Andersen JR, Lübberstedt T. 2003. Functional markers in plants. Trends in Plant Science 8, 554–560.

Apel K, Hirt H. 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annual Review of Plant Biology 55, 373–399.

Apse MP, Aharon GS, Snedden W, Blumwald E. 1999. Salt tolerance conferred by overexpression of a vacuolar Na+/H+ antiport in *Arabidopsis*. Science (New York, N.Y.) 285, 1256–1258.

Asins MJ, Raga V, Roca D, Belver A, Carbonell EA. 2015. Genetic dissection of tomato rootstock effects on scion traits under moderate salinity. Theoretical and Applied Genetics 128, 667–79.

Auerswald H, Peters P, Brückner B, Krumbein A, Kuchenbuch R. 1999. Sensory analysis and instrumental measurements of short-term stored tomatoes (*Lycopersicon esculentum* Mill.). Postharvest Biology and Technology 15, 323–334.

Azanza F, Young TE, Kim D, Tanksley SD, Juvik JA. 1994. Characterization of the effect of introgressed segments of chromosome 7 and 10 from *Lycopersion chmielewskii* on tomato soluble solids, pH, and yield. Theoretical and Applied Genetics: International Journal of Plant Breeding Research 87, 965–972.

B

Bac-Molenaar JA, Granier C, Keurentjes JJB, Vreugdenhil D. 2016. Genome-wide association mapping of time-dependent growth responses to moderate drought stress in *Arabidopsis*. Plant, Cell & Environment 39, 88–102.

Balasubramanian S, Schwartz C, Singh A, *et al.* 2009. QTL mapping in new *Arabidopsis thaliana* advanced intercross-recombinant inbred lines. PLoS ONE 4, 1–8.

Baldwin EA, Goodner K, Plotto A. 2008. Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. Journal of Food Science 73.

Baldwin EA, Scott JW, Shewmaker CK, Schuch W. 2000. Flavor trivia and tomato aroma: biochemistry and possible mechanisms for control of important aroma components. HortScience 35, 1013–1022.

Ballester A-R, Molthoff J, de Vos R, *et al.* 2010. Biochemical and molecular analysis of pink tomatoes: deregulated expression of the gene encoding transcription factor SlMYB12 leads to pink tomato fruit color. PLANT PHYSIOLOGY 152, 71–84.

Bandillo N, Raghavan C, Muyco PA, *et al.* 2013. Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. Rice (New York, N.Y.) 6, 11.

Barrero LS, Tanksley SD. 2004. Evaluating the genetic basis of multiple-locule fruit in a broad cross section of tomato cultivars. TAG. Theoretical and applied genetics 109, 669–679.

Barry CS, Aldridge GM, Herzog G, Ma Q, McQuinn RP, Hirschberg J, Giovannoni JJ. 2012. Altered chloroplast development and delayed fruit ripening caused by mutations in a zinc metalloprotease at the lutescent2 locus of tomato. Plant Physiology 159, 1086–1098.

Barry CS, McQuinn RP, Chung MY, Besuden A, Giovannoni JJ. 2008. Amino acid substitutions in homologs of the STAY-GREEN protein are responsible for the green-flesh and chlorophyll retainer mutations of tomato and pepper. Plant Physiology 147, 179–187.

Bauchet G, Causse M. 2010. Genetic diversity in tomato (*Solanum lycopersicum)* and its wild relatives. In: Mahmut Caliskan (Editeur), *Genetic diversity in plants* (p. 133-162). Vienne, AUT : IN-TECH Education and Publishing.

Bernacchi D, Beck-Bunn T, Emmatty D, *et al.* 1998a. Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. Theoretical and Applied Genetics 97, 170–180.

Bernacchi D, Beck-Bunn T, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley S. 1998b. Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum.* Theoretical and Applied Genetics 97, 381–397.

Bertin N, Causse M, Brunel B, Tricon D, Génard M. 2009. Identification of growth processes involved in QTLs for tomato fruit size and composition. Journal of experimental botany 60, 237–48.

Bertin N, Martre P, Genard M, Quilot B, Salon C. 2010. Under what circumstances can process-based simulation models link genotype to phenotype for complex traits? Case-study of fruit and grain quality traits. Journal of Experimental Botany 61, 955–967.

Blanca J, Cañizares J, Cordero L, Pascual L, Diez MJ, Nuez F. 2012. Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato (W Yan, Ed.). PLoS ONE 7, e48198.

Blanca J, Montero-Pau J, Sauvage C, Bauchet G, Illa E, Díez MJ, Francis D, Causse M, van der Knaap E, Cañizares J. 2015. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. BMC Genomics 16, 257.

Blokhina O, Virolainen E, Fagerstedt K V. 2003. Antioxidants, oxidative damage and oxygen deprivation stress: A review. Annals of Botany 91, 179–194.

Bolger A, Scossa F, Bolger ME, *et al.* 2014. The genome of the stress-tolerant wild tomato species *Solanum pennellii.* Nature Genetics 46, 1034–1038.

Boote KJ, Kropff MJ, Bindraban PS. 2001. Physiology and modelling of traits in crop plants: implications for genetic improvement. Agricultural Systems 70, 395–420.

Borsani O, Cuartero J, Valpuesta V, Botella MA. 2002. Tomato tos1 mutation identifies a gene essential for osmotic tolerance and abscisic acid sensitivity. Plant Journal 32, 905–914.

Bota J, Medrano H, Flexas J. 2004. Is photosynthesis limited by decreased Rubisco activity and RuBP content under progressive water stress? New Phytologist 162, 671–681.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F. 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS genetics 6, e1000940.

Bray EA. 2002. Classification of genes differentially expressed during water-deficit stress in *Arabidopsis thaliana:* An analysis using microarray and differential expression data. Annals of Botany 89, 803–811.

Brewer MT, Moyseenko JB, Monforte AJ, Van Der Knaap E. 2007. Morphological variation in tomato: A comprehensive study of quantitative trait loci controlling fruit shape and development. Journal of Experimental Botany 58, 1339–1349.

Bruhn CM, Feldman N, Garlitz C, Harwood J, Ivans E, Marshall M, Riley A, Thurber D, Williamson E. 1991. Consumer perceptions of quality: apricots, cantaloupes, peaches, pears, strawberries and tomatoes. Journal of Food Quality 14, 187–195.

Bucheli P, Voirol E, de la Torre R, López J, Rytz A, Tanksley SD, Pétiard V. 1999. Definition of nonvolatile markers for flavor of tomato (*Lycopersicon esculentum* Mill.) as Tools in Selection and Breeding. Journal of Agricultural and Food Chemistry 47, 659–664.

Busch BL, Schmitz G, Rossmann S, Piron F, Ding J, Bendahmane A, Theres K. 2011. Shoot branching and leaf dissection in tomato are regulated by homologous gene modules. The Plant Cell 23, 3595–3609.

Buttery RG, Teranishi R, Flath RA, Ling LC. 1989. Fresh tomato volatiles: composition and sensory studies. ACS Symposium series-American Chemical Society.

C

Cadic E, Coque M, Vear F, *et al.* 2013. Combined linkage and association mapping of flowering time in Sunflower (*Helianthus annuus* L.). Theoretical and Applied Genetics 126, 1337–56.

Capel C, Fernández del Carmen A, Alba JM, *et al.* 2015. Wide-genome QTL mapping of fruit quality traits in a tomato RIL population derived from the wild-relative species *Solanum pimpinellifolium* L. Theoretical and Applied Genetics 128, 2019–2035.

de Carvalho HMC, Cruz de Carvalho MH. 2008. Drought stress and reactive oxygen species. Plant Signaling & Behavior 3, 156–165.

Casson SA, Hetherington AM. 2010. Environmental regulation of stomatal development. Current Opinion in Plant Biology 13, 90–95.

Castel SE, Levy-Moonshine A, Mohammadi P, Banks E, Lappalainen T. 2015. Tools and best practices for data processing in allelic expression analysis. Genome Biology 16, 195.

Causse M, Buret M, Robini K, Verschave P. 2003. Inheritance of nutritional and sensory quality traits in fresh market tomato and relation to consumer preferences. Journal of Food Science 68, 2342–2350.

Causse M, Chaïb J, Lecomte L, Buret M, Hospital F. 2007. Both additivity and epistasis control the genetic variation for fruit quality traits in tomato. Theoretical and Applied Genetics 115, 429–442.

Causse M, Friguet C, Coiret C, LePicier M, Navez B, Lee M, Holthuysen N, Sinesio F, Moneta E, Grandillo S. 2010. Consumer preferences for fresh tomato at the european scale: a common segmentation on taste and firmness. Journal of Food Science 75.

Causse M, Saliba-Colombani V, Lesschaeve I, Buret M. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 2. Mapping QTLs for sensory attributes. Theoretical and Applied Genetics 102, 273–283.

Chaïb J, Devaux MF, Grotte MG, Robini K, Causse M, Lahaye M, Marty I. 2007. Physiological relationships among physical, sensory, and morphological attributes of texture in tomato fruits. Journal of Experimental Botany 58, 1915–1925.

Chakrabarti M, Zhang N, Sauvage C, *et al.* 2013. A cytochrome P450 regulates a domestication trait in cultivated tomato. Proceedings of the National Academy of Sciences 110, 17125–17130.

Chapman NH, Bonnet J, Grivet L, *et al.* 2012. High-resolution mapping of a fruit firmness-related quantitative trait locus in tomato reveals epistatic interactions associated with a complex combinatorial locus. Plant Physiology 159, 1644–1657.

Chaumont F, Moshelion M, Daniels MJ. 2005. Regulation of plant aquaporin activity. Biology of the cell / under the auspices of the European Cell Biology Organization 97, 749–764.

Chen FQ, Foolad MR, Hyman J, St. Clair DA, Beelaman RB. 1999. Mapping of QTLs for lycopene and other fruit traits in a *Lycopersicon esculent*um × *L. pimpinellifolium* cross and comparison of QTLs across tomato species. Molecular Breeding 5, 283–299.

Chen G, Hackett R, Walker D, Taylor A, Lin Z, Grierson D. 2004. Identification of a specific isoform of tomato lipoxygenase (TomloxC) involved in the generation of fatty acid-derived flavor compounds. Plant physiology 136, 2641–2651.

Chinnusamy V, Zhu J. 2009. Epigenetic regulation of stress responses in plants. Current opinion in plant biology, 12(2), 133-139.

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142, 169–196.

Combes MC, Hueber Y, Dereeper A, Rialle S, Herrera JC, Lashermes P. 2015. Regulatory divergence between parental alleles determines gene expression patterns in hybrids. Genome Biology and Evolution 7, 1110–1121.

Cong B, Barrero LS, Tanksley SD. 2008. Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nature genetics 40, 800–804.

Cornic G. 2000. Drought stress inhibits photosynthesis by decreasing stomatal aperture–not by affecting ATP synthesis. Trends in plant science 5, 187–188.

Cortina C, Culiáñez-Macià FA. 2005. Tomato abiotic stress enhanced tolerance by trehalose biosynthesis. Plant Science 169, 75–82.

Costa JM, Ortuño MF, Chaves MM. 2007. Deficit irrigation as a strategy to save water: physiology and potential application to horticulture. Journal of integrative plant biology 49, 1421–1434.

Cubillos F, Stegle O, Grondin C, Canut M, Tisné S, Gy I, Loudet O. 2014. Extensive cis-regulatory variation robust to environmental perturbation in *Arabidopsis*. The Plant cell 26, 4298–310.

Cubillos F, Yansouni J, Khalili H, *et al.* 2012. Expression variation in connected recombinant populations of *Arabidopsis thaliana* highlights distinct transcriptome architectures. BMC genomics 13, 117.

Cushman JC, Bohnert HJ. 2000. Genomic approaches to plant stress tolerance. Current opinion in plant biology 3, 117–124.

D

Daszkowska-Golec A, Szarejko I. 2013. Open or close the gate - stomata action under the control of phytohormones in drought stress conditions. Frontiers in plant science 4, 138.

Davies JN, Hobson GE, McGlasson WB. 1981. The constituents of tomato fruit — the influence of environment, nutrition, and genotype. Critical Reviews in Food Science and Nutrition 15, 205–280.

DeVicente MC, Tanksley SD. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134, 585–596.

Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. 2010. Rare variants create synthetic genome-wide associations. PLoS biology 8, e1000294.

Doganlar S, Frary A, Ku H-M, Tanksley SD. 2002. Mapping quantitative trait loci in inbred backcross lines of *Lycopersicon pimpinellifolium* (LA1589). Genome 45, 1189–1202.

Druka A, Potokina E, Luo Z, Jiang N, Chen X, Kearsey M, Waugh R. 2010. Expression quantitative trait loci analysis in plants. Plant Biotechnology Journal 8, 10–27.

Du Y, Kawamitsu Y, Nose A, Hiyane S, Murayama S, Wasano K, Uchida Y. 1996. Effects of water stress on carbon exchange rate and activities of photosynthetic enzymes in leaves of sugarcane (*Saccharum* Sp.). Australian Journal of Plant Physiology 23, 719.

E

van Eeuwijk F, Bink MC, Chenu K, Chapman SC. 2010. Detection and use of QTL for complex traits in multiple environments. Current Opinion in Plant Biology 13, 193–205.

van Eeuwijk F, Malosetti M, Yin X, Struik PC, Stam P. 2005. Statistical models for genotype by environment data: from conventional ANOVA models to eco-physiological QTL models. Crop and Pasture Science 56, 883–894.

El-Soda M, Kruijer W, Malosetti M, Koornneef M, Aarts MGM. 2015. Quantitative trait loci and candidate genes underlying genotype by environment interaction in the response of *Arabidopsis thaliana* to drought. Plant, Cell & Environment 38, 585–599.

El-Soda M, Malosetti M, Zwaan BJ, Koornneef M, Aarts MGM. 2014. Genotype × environment interaction QTL mapping in plants: lessons from *Arabidopsis*. Trends in plant science 19, 390–8.

Eshed Y, Zamir D. 1995. An Introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141, 1147–1162.

EUROSTAT. 2015. Agricultural production: crops, http://ec.europa.eu/eurostat/statistics-explained/.

F

Famoso AN, Zhao K, Clark RT, Tung C-W, Wright MH, Bustamante C, Kochian L V, McCouch SR. 2011. Genetic Architecture of Aluminum Tolerance in Rice (*Oryza sativa*) Determined through Genome-Wide Association Analysis and QTL Mapping. PLoS genetics 7, e1002221.

FAO WATER. 2016. Crop Water Information, http://www.fao.org/nr/water/cropinfo.html.

FAOSTAT. 2012. Food and agricultural commodities production, http://faostat.fao.org/site/339/default.aspx.

Farooq M, A.Wahid, Kobayashi N, Fujita D, Basra SMA. 2009. Plant drought stress : effects , mechanisms and management. Agronomy for Sustainable Development 29, 185–212.

Fereres E, Soriano MA. 2006. Deficit irrigation for reducing agricultural water use. Journal of Experimental Botany 58, 147–159.

Finlay K, Wilkinson G. 1963. The analysis of adaptation in a plant-breeding programme. Australian Journal of Agricultural Research 14, 742.

Flexas J, Medrano H. 2002. Drought-inhibition of photosynthesis in C3 plants: Stomatal and non-stomatal limitations revisited. Annals of Botany 89, 183–189.

Flowers TJ. 2004. Improving crop salt tolerance. Journal of Experimental Botany 55, 307–319.

Frary A. 2000. fw2.2: A quantitative trait locus key to the evolution of tomato fruit size. Science 289, 85–88.

Frary A, Fulton TM, Zamir D, Tanksley SD. 2004. Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *L. pennellii* cross and identification of possible orthologs in the Solanaceae. Theoretical and Applied Genetics 108, 485–496.

Fray RG, Grierson D. 1993. Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing, complementation and co-suppression. Plant Molecular Biology 22, 589–602.

Fridman E, Pleban T, Zamir D. 2000. A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proceedings of the National Academy of Sciences of the United States of America 97, 4718–4723.

Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, Shinozaki K. 2006. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. Current Opinion in Plant Biology 9, 436–442.

Fulton TM, Beck-Bunn T, Emmatty D, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD. 1997. QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. TAG Theoretical and Applied Genetics 95, 881–894.

Fulton TM, Bucheli P, Voirol E, López J, Pétiard V, Tanksley SD. 2002. Quantitative trait loci (QTL) affecting sugars, organic acids and other biochemical properties possibly contributing to flavor, identified in four advanced backcross populations of tomato. Euphytica 127, 163–177.

Fulton TM, Grandillo S, Beck-Bunn T, Fridman E, Frampton A, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD. 2000. Advanced backcross QTL analysis of a Lycopersicon esculentum x Lycopersicon parviflorum cross. Theoretical and Applied Genetics 100, 1025–1042.

G

Gao X, Giorgi F. 2008. Increased aridity in the Mediterranean region under greenhouse gas forcing estimated from high resolution simulations with a regional climate model. Global and Planetary Change 62, 195–209.

Garg BK. 2003. Nutrient uptake and management under drought: nutrient-moisture interaction. Curr Agric 27, 1–8.

Gauch JHG. 1988. Model selection and validation for yield trials with interaction. Biometrics 44, 705–715.

Gautier H, Diakou-Verdin V, Bénard C, Reich M, Buret M, Bourgaud F, Poëssel JL, Caris-Veyrat C, Génard M. 2008. How does tomato quality (sugar, acid, and nutritional quality) vary with ripening stage, temperature, and irradiance? Journal of Agricultural and Food Chemistry 56, 1241–1250.

Goel D, Singh AK, Yadav V, Babbar SB, Bansal KC. 2010. Overexpression of osmotin gene confers tolerance to salt and drought stresses in transgenic tomato (*Solanum lycopersicum* L.). Protoplasma 245, 133–141.

Golan I, Guadalupe Dominguez P, Konrad Z, Shkolnik-Inbar D, Carrari F, Bar-Zvi D. 2014. Tomato ABSCISIC ACID STRESS RIPENING (ASR) gene family revisited. PLoS ONE 9, 1–8.

Goldman IL, Paran I, Zamir D. 1995. Quantitative trait locus analysis of a recombinant inbred line population derived from a *Lycopersicon esculent*um x *Lycopersicon cheesmanii* cross. Theoretical and Applied Genetics 90.

Golldack D, Li C, Mohan H, Probst N. 2014. Tolerance to drought and salt stress in plants: Unraveling the signaling networks. Frontiers in Plant Science 5.

Goulet C, Mageroy MH, Lam NB, Floystad A, Tieman DM, Klee HJ. 2012. Role of an esterase in flavor volatile variation within the tomato clade. Proceedings of the National Academy of Sciences of the United States of America 109, 19009–14.

Grandillo S, Ku HM, Tanksley SD. 1999. Identifying the loci responsible for natural variation in fruit size and shape in tomato. TAG Theoretical and Applied Genetics 99, 978–987.

Grandillo S, Tanksley SD. 1996. QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersicon pimpinellifolium.* Theoretical and Applied Genetics 92, 935–951.

Guichard S, Bertin N, Leonardi C, Gary C. 2001. Tomato fruit quality in relation to water and carbon fluxes. Agronomie 21, 385–392.

Gur A, Semel Y, Osorio S, *et al.* 2011. Yield quantitative trait loci from wild tomato are predominately expressed by the shoot. Theoretical and Applied Genetics 122, 405–20.

H

Hamilton JP, Sim S-C, Stoffel K, Van Deynze A, Buell CR, Francis DM. 2012. Single nucleotide polymorphism discovery in cultivated tomato via sequencing by synthesis. The Plant Genome Journal 5, 17.

Hammond JP, Mayes S, Bowen HC, *et al.* 2011. Regulatory hotspots are associated with plant gene expression under varying soil phosphorus supply in *Brassica rapa.* Plant physiology 156, 1230–41.

He F, Arce AL, Schmitz G, Koornneef M, Novikova P, Beyer A, de Meaux J. 2016. The footprint of polygenic adaptation on stress-responsive Cis -regulatory divergence in the *Arabidopsis* genus. Molecular Biology and Evolution 33, 2088–2101.

Hébert Y, Plomion C, Harzic N. 1995. Genotype x environment interaction for root traits in maize, as analysed with factorial regression models. Euphytica 81, 85–92.

Hernández I, Alegre L, Van Breusegem F, Munné-Bosch S. 2009. How relevant are flavonoids as antioxidants in plants? Trends in Plant Science 14, 125–132.

Hovav R, Chehanovsky N, Moy M, Jetter R, Schaffer AA. 2007. The identification of a gene (Cwp1), silenced during Solanum evolution, which causes cuticle microfissuring and dehydration when expressed in tomato fruit. The Plant Journal 52, 627–639.

Hsieh T, Lee J, Charng Y, Chan M. 2002. Tomato plants ectopically expressing Arabidopsis CBF1 show enhanced resistance to water deficit stress 1. Plant Physiology 130, 618–626.

Hsieh T-H, Li C-W, Su R-C, Cheng C-P, Sanjaya, Tsai Y-C, Chan M-T. 2010. A tomato bZIP transcription factor, SlAREB, is involved in water deficit and salt stress response. Planta 231, 1459–1473.

Huang X, Paulo M-J, Boer M, Effgen S, Keizer P, Koornneef M, van Eeuwijk F. 2011. Analysis of natural allelic variation in *Arabidopsis* using a multiparent recombinant inbred line population. Proceedings of the National Academy of Sciences of the United States of America 108, 4488–4493.

Hu Y, Schmidhalter U. 2005. Drought and salinity: A comparison of their effects on mineral nutrition of plants. Journal of Plant Nutrition and Soil Science 168, 541–549.

I

Isaacson T, Ronen G, Zamir D, Hirschberg J. 2002. Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of beta-carotene and xanthophylls in plants. The Plant cell 14, 333–42.

Ito Y, Kitagawa M, Ihashi N, Yabe K, Kimbara J, Yasuda J, Ito H, Inakuma T, Hiroi S, Kasumi T. 2008. DNA-binding specificity, transcriptional activation potential, and the rin mutation effect for the tomato fruit-ripening regulator RIN. Plant Journal 55, 212–223.

Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K. 2001. Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in *Arabidopsis*. Plant J 27, 325–333.

J

Jansen RC, Stam P. 1994. High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136, 1447–1455.

Jenkins JA. 1948. The origin of the cultivated tomato. Economic Botany 2, 379–392.

Johansson L, Haglund Å, Berglund L, Lea P, Risvik E. 1999. Preference for tomatoes, affected by sensory attributes and information about growth conditions. Food Quality and Preference 10, 289–298.

Jury W a, Vaux H. 2005. The role of science in solving the world's emerging water problems. Proceedings of the National Academy of Sciences of the United States of America 102, 15715–15720.

K

Kang J, Hwang J-U, Lee M, Kim Y-Y, Assmann SM, Martinoia E, Lee Y. 2010. PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. Proceedings of the National Academy of Sciences 107, 2355–2360.

Khraiwesh B, Arif MA, Seumel GI, Ossowski S, Weigel D, Reski R, Frank W. 2010. Transcriptional control of gene expression by microRNAs. Cell 140, 111–122.

Kim T-H, Bohmer M, Hu H, Nishimura N, Schroeder JI. 2010. Guard cells signal transduction network: advances in understanding abscisic acid CO2, and Ca2+ signalling. Annual Review of Plant Biology 61, 561–591.

Klee HJ, Tieman DM. 2013. Genetic challenges of flavor improvement in tomato. Trends in Genetics 29, 257–262.

Klingler JP, Batelli G, Zhu JK. 2010. ABA receptors: The START of a new paradigm in phytohormone signalling. Journal of Experimental Botany 61, 3199–3210.

van der Knaap E, Tanksley SD. 2003. The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in Yellow Stuffer tomato. Theoretical and applied genetics 107, 139–147.

Knight JC. 2004. Allele-specific gene expression uncovered. Trends in Genetics 20, 113–116.

Koornneef M, Stam P. 2001. Changing paradigms in plant breeding. Plant Physiology 125, 156–159.

Korte A, Farlow A. 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant methods 9, 29.

Korte A, Vilhjálmsson BJ, Segura V, Platt A, Long Q, Nordborg M. 2012. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. Nature genetics 44, 1066–71.

Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R. 2009. A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. PLoS Genetics 5.

Kraakman ATW. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. Genetics 168, 435–446.

Kulik A, Wawer I, Krzywińska E, Bucholc M, Dobrowolska G. 2011. SnRK2 protein kinases—key regulators of plant response to abiotic stresses. OMICS: A Journal of Integrative Biology 15, 859–872.

Kuromori T, Shinozaki K. 2010. ABA transport factors found in *Arabidopsis* ABC transporters. Plant signaling & behavior 5, 1124–1126.

Kuromori T, Sugimoto E, Shinozaki K. 2011. *Arabidopsis* mutants of AtABCG22, an ABC transporter gene, increase water transpiration and drought susceptibility. Plant Journal 67, 885–894.

Kushiro T, Okamoto M, Nakabayashi K, Yamagishi K, Kitamura S, Asami T, Hirai N, Koshiba T, Kamiya Y, Nambara E. 2004. The *Arabidopsis* cytochrome P450 CYP707A encodes ABA 8'-hydroxylases: key enzymes in ABA catabolism. The EMBO Journal 23, 1647–1656.

L

Labate JA, Grandillo S, Fulton T, *et al.* 2007. Tomato. In *Vegetables* (pp. 1-125). Springer Berlin Heidelberg.

Labra M, Ghiani A, Citterio S, Sgorbati S, Sala F, Vannini C, Bracale M. 2002. Analysis of cytosine methylation pattern in response to water deficit in pea root tips. , 694–699.

Lander ES, Botstein D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121, 185–199.

Laperche A, Devienne-Barret F, Maury O, Le Gouis J, Ney B. 2006. A simplified conceptual model of carbon/nitrogen functioning for QTL analysis of winter wheat adaptation to nitrogen deficiency. Theoretical and Applied Genetics 113, 1131–1146.

Lescourret F, Génard M. 2005. A virtual peach fruit model simulating changes in fruit quality during the final stage of fruit growth. Tree physiology 25, 1303–1315.

Levitt J. 1972. Responses of plants to environmental stresses. New York: Academic press.

Li S, Wang J, Zhang L. 2015. Inclusive composite interval mapping of QTL by environment interactions in biparental populations. PLOS ONE 10, e0132414.

Lin T, Zhu G, Zhang J, *et al.* 2014. Genomic analyses provide insights into the history of tomato breeding. Nature genetics 46(11), 1220-1226.

Lippman ZB, Cohen O, Alvarez JP, Abu-Abied M, Pekker I, Paran I, Eshed Y, Zamir D. 2008. The making of a compound inflorescence in tomato and related nightshades. PLoS Biology 6, 2424–2435.

Lippman Z, Tanksley SD. 2001. Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *Lycopersicon pimpinellifolium* and *L. esculentum* var. Giant Heirloom. Genetics 158, 413–422.

Liu J, Van Eck J, Cong B, Tanksley SD. 2002. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proceedings of the National Academy of Sciences of the United States of America 99, 13302–6.

Liu W, Fairbairn DJ, Reid RJ, Schachtman DP, Plant C, Horticulture I, Box GPO, Osmond G. 2001. Characterization of two HKT1 homologues from *Eucalyptus camaldulensis* that display intrinsic osmosensing capability. 127, 283–294.

Liu HH-F, Génard M, Guichard S, Bertin N. 2007. Model-assisted analysis of tomato fruit growth in relation to carbon and water fluxes. Journal of Experimental Botany 58(13), 3567–80.

Lovell JT, Juenger TE, Michaels SD, Lasky JR, Platt A, Richards JH, Yu X, Easlon HM, Sen S, McKay JK. 2013. Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. Proceedings of the Royal Society B: Biological Sciences 280, 20131043–20131043.

Lovell JT, Mullen JL, Lowry DB, Awole K, Richards JH, Sen S, Verslues PE, Juenger TE, McKay JK. 2015. Exploiting differential gene expression and epistasis to discover candidate genes for drought-associated QTLs in *Arabidopsis thaliana.* The Plant cell 27, 969–983.

Ludlow MM. 1989. Strategies of response to water stress. In: Kreeb HK, Ritcher H, Hinckley TM (Eds) Structural and functional responses to environmental stresses: water shortage. Academic press, The Netherlands 269–281.

M

Mageroy MH, Tieman DM, Floystad A, Taylor MG, Klee HJ. 2012. A *Solanum lycopersicum* catechol-O-methyltransferase involved in synthesis of the flavor molecule guaiacol. Plant Journal 69, 1043–1051.

Malosetti M, Ribaut JM, Vargas M, Crossa J, van Eeuwijk F. 2007. A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). Euphytica 161, 241–257.

Manning K, Tor M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB. 2006. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat Genet 38, 948–952.

Mao L, Begum D, Chuang H, Budiman, Muhammad A. Szymkowiak EJ, Irish EE, Wing RA. 2000. JOINTLESS is a MADS-box gene controlling tomato flower abscission zone development. Nature 406, 910–913.

Des Marais DL, Hernandez KM, Juenger TE. 2013. Genotype-by-environment interaction and plasticity: Exploring genomic responses of plants to the abiotic environment. Annual Review of Ecology, Evolution, and Systematics 44, 5–29.

Des Marais DL, McKay JK, Richards JH, Sen S, Wayne T, Juenger TE. 2012. Physiological genomics of response to soil drying in diverse Arabidopsis accessions. The Plant cell 24, 893–914.

Mathieu S, Cin VD, Fei Z, Li H, Bliss P, Taylor MG, Klee HJ, Tieman DM. 2009. Flavour compounds in tomato fruits: Identification of loci and potential pathways affecting volatile composition. Journal of Experimental Botany 60, 325–337.

McMullen MD, Kresovich S, Villeda HS, *et al.* 2009. Genetic properties of the maize nested association mapping population. Science 325, 737–741.

Merlot S, Gosti F, Guerrier D, Vavasseur A, Giraudat J. 2001. The ABI1 and ABI2 protein phosphatases 2C act in a negative feedback regulatory loop of the abscisic acid signalling pathway. The Plant Journal 25, 295–303.

Mir RR, Kumar N, Jaiswal V, Girdharwal N, Prasad M, Balyan HS, Gupta PK. 2012. Genetic dissection of grain weight in bread wheat through quantitative trait locus interval and association mapping. Molecular Breeding 29, 963–972.

Mishra KB, Iannacone R, Petrozza A, Mishra A, Armentano N, La Vecchia G, Trtílek M, Cellini F, Nedbal L. 2012. Engineered drought tolerance in tomato plants is reflected in chlorophyll fluorescence emission. Plant Science 182, 79–86.

Mitchell JP, Shennann C, Grattan SR, May DM. 1991. Tomato fruit yields and quality under water deficit and salinity. Journal of the American Society for Horticultural 116, 215–221.

Mitra J. 2001. Genetics and genetic improvement of drought resistance in crop plants. Current Science 80, 758–763.

Mittler R. 2002. Oxidative stress, antioxidants and stress tolerance. Trends in Plant Science 7, 405–410.

Mittler R, Vanderauwera S, Gollery M, Van Breusegem F. 2004. Reactive oxygen gene network of plants. Trends in Plant Science 9, 490–498.

Moghaieb REA, Tanaka N, Saneoka H, Hussein HA, Yousef SS, Ewada MA-F, Aly MAM, Fujita K. 2000. Expression of betaine aldehyde dehydrogenase gene in transgenic tomato hairy roots leads to the accumulation of glycine betaine and contributes to the maintenance of the osmotic potential under salt stress. Soil Science and Plant Nutrition 46, 873–883.

Molinero-Rosales N, Jamilena M, Zurita S, Gómez P, Capel J, Lozano R. 1999. FALSIFLORA, the tomato orthologue of FLORICAULA and LEAFY, controls flowering time and floral meristem identity. Plant Journal 20, 685–693.

Molinero-Rosales N, Latorre A, Jamilena M, Lozano R. 2004. SINGLE FLOWER TRUSS regulates the transition and maintenance of flowering in tomato. Planta 218, 427–434.

Monforte AJ, As MJ, Carbonell EA. 1997a. Salt tolerance in *Lycopersicon* species: VI Genotype-by-salinity interaction in quantitative trait loci detection : constitutive and response QTLs. Theoretical and Applied Genetics 95(4), 706–713.

Monforte AJ, As MJ, Carbonell EA. 1997b. Salt tolerance in *Lycopersicon* species: V . Does genetic variability at quantitative trait loci affect their analysis ? Theoretical and Applied Genetics, 95(1-2), 284–293.

Monforte AJ, Asins MJ, Carbonell EA. 1996. Salt tolerance in Lycopersicon species . IV . Efficiency of marker-assisted selection for salt tolerance improvement. Theoretical and applied genetics, 93(5-6), 765–772.

Mueller L, Solow TH, Taylor N, *et al.* 2005. The SOL Genomics Network: a comparative resource for Solanaceae biology and beyond. Plant Physiology 138, 1310–1317.

Muir W, Nyquist WE, Xu S. 1992. Alternative partitioning of the genotype-by-environment interaction. Theoretical and Applied Genetics 84, 193–200.

Muller B, Pantin F, Génard M, Turc O, Freixes S, Piques M, Gibon Y. 2011. Water deficits uncouple growth from photosynthesis, increase C content, and modify the relationships between C and growth in sink organs. Journal of Experimental Botany 62, 1715–1729.

Muños S, Ranc N, Botton E, *et al.* 2011. Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiology 156, 2244–2254.

Mustilli AC, Fenzi F, Ciliento R, Alfano F, Bowler C. 1999. Phenotype of the tomato high pigment-2 mutant is caused by a mutation in the tomato homolog of DEETIOLATED1. The Plant cell 11, 145–157.

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES. 2009. Association mapping: critical considerations shift from genotyping to experimental design. The Plant cell 21, 2194–2202.

N

Nesbitt T, Tanksley SD. 2002. Comparative sequencing in the genus *lycopersicon*: Implications for the evolution of fruit size in the domestication of cultivated tomatoes. Genetics 162, 365–379.

Nica AC, Dermitzakis ET. 2013. Expression quantitative trait loci: present and future. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 368, 20120362.

Nonami H. 1998. Plant water relations and control of cell elongation at low water potentials. Journal of Plant Research 111, 373–382.

Nordborg M, Weigel D. 2008. Next-generation genetics in plants. Nature 456, 720–723.

Nuruddin M, Madramootoo C, Dodds GT. 2003. Effects of water stress at different growth stages on greenhouse tomato yield and quality. 38, 1389–1393.

O

Ofner I, Lashbrooke J, Pleban T, Aharoni A, Zamir D. 2016. *Solanum pennellii* backcross inbred lines (BILs) link small genomic bins with tomato traits. Plant Journal 87, 151–160.

Ogura T, Busch W. 2015. From phenotypes to causal sequences: Using genome wide association studies to dissect the sequence basis for variation of plant development. Current Opinion in Plant Biology 23, 98–108.

Orellana S, Yañez M, Espinoza A, Verdugo I, González E, Ruiz-Lara S, Casaretto JA. 2010. The transcription factor SlAREB1 confers drought, salt stress tolerance and regulates biotic and abiotic stress-related genes in tomato. Plant, Cell and Environment 33, 2191–2208.

Osorio S, Ruan Y-L, Fernie AR. 2014. An update on source-to-sink carbon partitioning in tomato. Frontiers in Plant Science 5, 516.

P

Panagiotou OA, Ioannidis JPA. 2012. What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. International Journal of Epidemiology 41, 273–286.

Park S, Li J, Pittman JK, Berkowitz GA, Yang H, Undurraga S, Morris J, Hirschi KD, Gaxiola RA. 2005. Up-regulation of a H+-pyrophosphatase (H+-PPase) as a strategy to engineer drought-resistant crop plants. Proceedings of the National Academy of Sciences of the United States of America 102, 18830–5.

Pascale SD, Maggio A, Fogliano V, Ambrosino P, Ritieni A. 2001. Irrigation with saline water improves carotenoids content and antioxidant activity of tomato. The Journal of Horticultural Science and Biotechnology 76, 447–453.

Pascual L, Albert E, Sauvage C, *et al.* 2016. Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. Plant Science 242, 120–130.

Pascual L, Desplat N, Huang BE, Desgroux A, Bruguier L, Bouchet JP, Le QH, Chauchard B, Verschave P, Causse M. 2015. Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. Plant Biotechnology Journal 13, 565–577.

Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES, S.D. T. 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generation and environments. Genetics 127, 181–197.

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD. 1988. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335, 721–726.

Peralta IE, Knapp S, Spooner DM. 2005. New species of wild tomatoes (*Solanum* section *Lycopersicon: Solanaceae*) from northern Peru. Systematic Botany 30, 424–434.

Peuke AD, Rennenberg H. 2004. Carbon, nitrogen, phosphorus, and sulphur concentration and partitioning in beech ecotypes (*Fagus sylvatica* L.): Phosphorus most affected by drought. Trees - Structure and Function 18, 639–648.

Plaut Z. 2003. Plant exposure to water stress during specific growth stage. In: Trimble SW, ed. Encyclopedia of water science. London, 673–675.

Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez JP, Ganal M, Zamir D, Lifschitz E. 1998. The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. Development (Cambridge, England) 125, 1979–1989.

Poiroux-Gonord F, Bidel LPR, Fanciullino A-L, Gautier H, Lauri-Lopez F, Urban L. 2010. Health benefits of vitamins and secondary metabolites of fruits and vegetables and prospects to increase their concentrations by agronomic approaches. Journal of Agricultural and Food Chemistry 58, 12065–12082.

Powell A, Nguyen G, Hill T, *et al.* 2012. Uniform ripening encodes a golden 2-like transcription factor regulating tomato fruit chloroplast development. Science 336, 1711–1715.

Pritchard JK, Stephens M, Rosenberg NA, Donnelley P. 2000. Association mapping in structured populations. American journal of human genetics 67, 170–181.

Proseus TE, Ortega JKE, Boyer JS. 1999. Separating growth from elastic deformation during cell enlargement. Plant physiology 119, 775–784.

Prudent M, Dai ZW, Génard M, Bertin N, Causse M, Vivin P. 2014. Resource competition modulates the seed number - fruit size relationship in a genotype-dependent manner: A modeling approach in grape and tomato. Ecological Modelling 290, 54–64.

Prudent M, Lecomte A, Bouchet J-P, Bertin N, Causse M, Genard M. 2011. Combining ecophysiological modelling and quantitative trait locus analysis to identify key elementary processes underlying tomato fruit sugar concentration. Journal of Experimental Botany 62, 907–919.

Q

Qin X, Zeevaart J. 1999. The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory step of abscisic acid biosynthesis in water-stressed bean. Proceedings of the National Academy of Sciences of the United States of America 96, 15354–15361.

Qin X, Zeevaart J. 2002. Overexpression of a 9-cis-epoxycarotenoid dioxygenase gene in *Nicotiana plumbaginifolia* increases abscisic acid and phaseic acid levels and enhances drought tolerance. American Society of Plant Physiologists 128, 544–551.

Quilot B, Kervella J, Génard M, Lescourret F. 2005. Analysing the genetic control of peach fruit quality through an ecophysiological model combined with a QTL approach. Journal of Experimental Botany 56, 3083–3092.

R

Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Shinozaki KY. 2003. Monitoring expression profiles of rice genes under Ccold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNAGel-Blot analyses. Plant Physiol. 133, 1755–1767.

Ranc N, Muños S, Santoni S, Causse M. 2008. A clarified position for *Solanum lycopersicum* var. *cerasiforme* in the evolutionary history of tomatoes (*solanaceae*). BMC Plant Biology 8, 130.

Ravel C, Praud S, Canaguier A, *et al.* 2007. DNA sequence polymorphisms and their application to bread wheat quality. Euphytica 158, 331–336.

Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F. 2003. Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. Plant Physiol. 131, 664–675.

Reymond M, Muller B, Tardieu F. 2004. Dealing with the genotype x environment interaction via a modelling approach: A comparison of QTLs of maize leaf length or width with QTLs of model parameters. Journal of Experimental Botany 55, 2461–2472.

Rick CM. 1979. The biology and taxonomy of the *Solanaceae*. In: Hawkes, JG, Lester, RN, Skelding, AD, Eds. Biosystematic studies in *Lycopersicon* and closely related species of *Solanum*. London, 667–678.

Ripoll J, Urban L, Bertin N. 2016a. The potential of the MAGIC TOM parental accessions to explore the genetic variability in tomato acclimation to repeated cycles of water deficit and recovery. Frontiers in Plant Science 6.

Ripoll J, Urban L, Brunel B, Bertin N. 2016b. Water deficit effects on tomato quality depend on fruit developmental stage and genotype. Journal of Plant Physiology 190, 26–35.

Ripoll J, Urban L, Staudt M, Lopez-Lauri F, Bidel LPR, Bertin N. 2014. Water shortage and quality of fleshy fruits, making the most of the unavoidable. Journal of experimental botany 65, 4097–117.

Robertson LD, Labate J. 2007. Genetic resources of tomato (*Lycopersicon esculentum* Mill.) and wild relatives. Genetic Improvement of *Solanaceous* Crops. Tomato, 2, 25-75.

Ronen G, Carmel-Goren L, Zamir D, Hirschberg J. 2000. An alternative pathway to β-carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. Pnas 97, 11102–11107.

Ronen G, Cohen M, Zamir D, Hirschberg J. 1999. Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant Delta. The Plant Journal 17, 341–351.

Rost S, Gerten D, Bondeau A, Lucht W, Rohwer J, Schaphoff S. 2008. Agricultural green and blue water consumption and its influence on the global water system. Water Resources Research 44.

Ruggieri V, Francese G, Sacco A, Alessandro AD, Rigano MM, Parisi M, Milone M, Cardi T, Mennella G, Barone A. 2014. An association mapping approach to identify favourable alleles for tomato fruit quality breeding. BMC plant biology 14, 1–15.

Rushton DL, Tripathi P, Rabara RC, *et al.* 2012. WRKY transcription factors : key components in abscisic acid signalling. 10(1), 2–11.

S

Sabaghnia N, Sabaghpour SH, Dehghani H. 2008. The use of an AMMI model and its parameters to analyse yield stability in multi-environment trials. The Journal of Agricultural Science 146, 571–581.

Sacco A, Ruggieri V, Parisi M, Festa G, Rigano MM, Picarella ME, Mazzucato A, Barone A. 2015. Exploring a tomato landraces collection for fruit-related traits by the aid of a high-throughput genomic platform. PLoS ONE 10, 1–20.

Sahu PP, Pandey G, Sharma N, Puranik S, Muthamilarasan M, Prasad M. 2013. Epigenetic mechanisms of plant stress responses and adaptation. Plant Cell Reports 32, 1151–1159.

Saïdou A-A, Thuillet A-C, Couderc M, Mariac C, Vigouroux Y. 2014. Association studies including genotype by environment interactions: prospects and limits. BMC genetics 15, 3.

Saliba-Colombani V, Causse M, Langlois D, Philouze J, Buret M. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. Theoretical and Applied Genetics 102, 259–272.

Sanders GJ, Arndt SK. 2012. Osmotic adjustment under drought conditions. Plant Responses to Drought Stress. Berlin, Heidelberg: Springer Berlin Heidelberg, 199–229.

Sato T, Iwatsubo T, Takahashi M, Nakagawa H, Ogura N, Mori H. 1993. Intercellular localization of acid invertase in tomato fruit and molecular cloning of a cDNA for the enzyme. Plant & cell physiology 34, 263–9.

Sauvage C, Segura V, Bauchet G, Stevens R, Do PT, Nikoloski Z, Fernie AR, Causse M. 2014. Genome-Wide Association in tomato reveals 44 candidate loci for fruit metabolic traits. Plant physiology 165, 1120–1132.

Schlichting CD. 1986. The evolution of phenotypic plasticity in plants. Annual Review of Ecology and Systematics 17, 667–693.

Schroeder JI, Allen GJ, Hugouvieux V, Kwak JM, Waner D. 2001. Guard cell signal transduction. Annual Review of Plant Physiology and Plant Molecular Biology 52, 627–658.

Schulz P, Herde M, Romeis T. 2013. Calcium-dependent protein kinases: hubs in plant stress signaling and development. Plant Physiology 163, 523–530.

Schumacher K, Schmitt T, Rossberg M, Schmitz G, Theres K. 1999. The Lateral suppressor (Ls) gene of tomato encodes a new member of the VHIID protein family. Proceedings of the National Academy of Sciences of the United States of America 96, 290–295.

Schweighofer A, Hirt H, Meskiene I. 2004. Plant PP2C phosphatases: Emerging functions in stress signaling. Trends in Plant Science 9, 236–243.

Scippa GS, Di Michele M, Onelli E, Patrignani G, Chiatante D, Bray EA. 2004. The histone-like protein H1-S and the response of tomato leaves to water deficit. Journal of Experimental Botany 55, 99–109.

Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nature genetics 44, 825–30.

Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K. 2001. Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell 13, 61–72.

Seki M, Narusaka M, Ishida J, *et al.* 2002. Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. The Plant Journal 31, 279–292.

Semel Y, Nissenbaum J, Menda N, Zinder M, Krieger U, Issman N, Pleban T, Lippman Z, Gur A, Zamir D. 2006. Overdominant quantitative trait loci for yield and fitness in tomato. Proceedings of the National Academy of Sciences 103, 12981–12986.

Semel Y, Schauer N, Roessner U, Zamir D, Fernie AR. 2007. Metabolite analysis for the comparison of irrigated and non-irrigated field grown tomato of varying genotype. Metabolomics 3, 289–295.

Shinozaki K, Yamaguchi-Shinozaki K. 2007. Gene networks involved in drought stress response and tolerance. Journal of experimental botany 58, 221–7.

Sim S-C, Durstewitz G, Plieske J, *et al.* 2012. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. PloS one 7, e40563.

Simkin AJ, Schwartz SH, Auldridge M, Taylor MG, Klee HJ. 2004. The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles β-ionone, pseudoionone, and geranylacetone. Plant Journal 40, 882–892.

Sinesio F, Cammareri M, Moneta E, Navez B, Peparaio M, Causse M, Grandillo S. 2009. Sensory quality of fresh French and Dutch market tomatoes: a preference mapping study with Italian consumers. Journal of food science 75, S55-67.

Singh KB, Foley RC, Oñate-Sánchez L. 2002. Transcription factors in plant defense and stress responses. Current Opinion in Plant Biology 5, 430–436.

Sinha AK, Jaggi M, Raghuram B, Tuteja N. 2011. Mitogen-activated protein kinase signaling in plants under abiotic stress. Plant signaling & behavior 6, 196–203.

Sonah H, O'Donoughue L, Cober E, Rajcan I, Belzile F. 2015. Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. Plant Biotechnology Journal 13, 211–221.

Speirs, J., Lee, E., Holt, K., Yong-Duk, K., Scott, N. S., Loveys, B., Schuch, W. 1998. Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavor aldehydes and alcohols. Plant physiology 117, 1047–1058.

Springer NM, Stupar RM. 2007. Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. The Plant Cell Online 19, 2391–2402.

Stupar RM, Springer NM. 2006. Cis-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. Genetics 173, 2199–2210.

Staub JE, Serquen FC, Gupta M. 1996. Genetic markers, map construction, and their application in plant breeding. HortScience 31, 729–741.

Stevens MA. 1979. Tomato quality: potential for developing cultivars with improved flavor. Acta Horticulturae 93, 317–329.

Stevens MA, Kader AA, Albright Holton M, Algazi M. 1977. Genotypic variation for flavor and composition in fresh market tomatoes. Journal American Society for Horticultural Science.

Stikic R, Popovic S, Srdic M, Savic D, Jovanovic Z, Zdravkovic J. 2003. Partial Root Drying (Prd): a New technique for growing plants that saves water and improves the quality of fruit. Plant Cell, 164–171.

Struik PC, Yin X, De Visser P. 2005. Complex quality traits: Now time to model. Trends in Plant Science 10, 513–516.

Suji KK, Joel AJ. 2010. An epigenetic change in rice cultivars under water stress conditions. 1, 1142–1143.

Sultan SE. 2000. Phenotypic plasticity for plant development, function and life history. Trends in Plant Science 5, 537–542.

Sun L, Rodriguez GR, Clevenger JP, *et al.* 2015. Candidate gene selection and detailed morphological evaluations of fs8.1, a quantitative trait locus controlling tomato fruit shape. Journal of Experimental Botany 66, 6471–6482.

Szarek SR, Ting IP. 1975. Photosynthetic efficiency of CAM plants in relation to C3 and C4 plants. Environmental and Biological Control of Photosynthesis. Dordrecht: Springer Netherlands, 289–297.

T

Tanksley SD. 1993. Mapping polygenes. Annual Review of Genetics 27, 205–233.

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB. 1992. High density molecular linkage maps of the tomato and potato genomes. Genetics 132, 1141–60.

Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T. 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theoretical and Applied Genetics 92, 213–224.

Tardieu F. 2003. Virtual plants : modelling as a tool for the genomics of tolerance to water deficit. Trends in plant science 8, 9–14.

Tardieu F. 2012. Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. Journal of experimental botany 63, 25–31.

Tardieu F, Reymond M, Hamard P, Granier C, Muller B. 2000. Spatial distributions of expansion rate, cell division rate and cell size in maize leaves: a synthesis of the effects of soil water status, evaporative demand and temperature. Journal of experimental botany 51, 1505–1514.

Tétard-Jones C, Kertesz M a, Preziosi RF. 2011. Quantitative trait loci mapping of phenotypic plasticity and genotype-environment interactions in plant and insect performance. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 366, 1368–79.

Thompson AJ, Jackson AC, Parker RA, Morpeth DR, Burbidge A, Taylor IB. 2000a. Abscisic acid biosynthesis in tomato: Regulation of zeaxanthin epoxidase and 9-cis-epoxycarotenoid dioxygenase mRNAs by light/dark cycles, water stress and abscisic acid. Plant Molecular Biology 42, 833–845.

Thompson AJ, Jackson AC, Symonds RC, Mulholland BJ, Dadswell AR, Blake PS, Burbidge A, Taylor IB. 2000b. Ectopic expression of a tomato 9- cis -epoxycarotenoid dioxygenase gene causes over-production of abscisic acid. The Plant Journal 23, 363–374.

Tieman D, Bliss P, McIntyre LM, *et al.* 2012. The chemical interactions underlying tomato flavor preferences. Current Biology 22, 1035–1039.

Tieman DM, Loucas HM, Kim JY, Clark DG, Klee HJ. 2007. Tomato phenylacetaldehyde reductases catalyze the last step in the synthesis of the aroma volatile 2-phenylethanol. Phytochemistry 68, 2660–2669.

Tieman DM, Zeigler M, Schmelz EA, Taylor MG, Bliss P, Kirst M, Klee HJ. 2006. Identification of loci affecting flavour volatile emissions in tomato fruits. Journal of Experimental Botany 57, 887–896.

Tieman D, Zeigler M, Schmelz E, Taylor MG, Rushing S, Jones JB, Klee HJ. 2010. Functional analysis of a tomato salicylic acid methyl transferase and its role in synthesis of the flavor volatile methyl salicylate. Plant Journal 62, 113–123.

Tikunov YM, Molthoff J, de Vos RCH, *et al.* 2013. NON-SMOKY GLYCOSYLTRANSFERASE1 prevents the release of smoky aroma from tomato fruit. The Plant Cell 25, 3067–3078.

Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485, 635–41.

Tran L-SP, Urao T, Qin F, Maruyama K, Kakimoto T, Shinozaki K, Yamaguchi-Shinozaki K. 2007. Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America 104, 20623–20628.

Tung SA, Smeeton R, White CA, Black CR, Taylor IB, Hilton HW, Thompson AJ. 2008. Over-expression of LeNCED1 in tomato (*Solanum lycopersicum* L.) with the rbcS3C promoter allows recovery of lines that accumulate very high levels of abscisic acid and exhibit severe phenotypes. Plant, Cell & Environment 31, 968–981.

Tunnacliffe A, Wise MJ. 2007. The continuing conundrum of the LEA proteins. Naturwissenschaften 94, 791–812.

Tuteja N, Mahajan S. 2007. Calcium signaling network in plants: an overview. Plant Signaling & Behavior 2, 79–85.

U

Umezawa T, Fujita M, Fujita Y, Yamaguchi-Shinozaki K, Shinozaki K. 2006a. Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. Current Opinion in Biotechnology 17, 113–122.

Umezawa T, Nakashima K, Miyakawa T, Kuromori T, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K. 2010. Molecular basis of the core regulatory network in ABA responses: Sensing, signaling and transport. Plant and Cell Physiology 51, 1821–1839.

Umezawa T, Okamoto M, Kushiro T, Nambara E, Oono Y, Seki M, Kobayashi M, Koshiba T, Kamiya Y, Shinozaki K. 2006b. CYP707A3, a major ABA 8′-hydroxylase involved in dehydration and rehydration response in Arabidopsis thaliana. Plant Journal 46, 171–182.

Unal D. 2013. Effect of abiotic stress on photosystem I-related gene transcription in photosynthetic organisms. Photosynthesis. InTech, .

Uozumi A, Ikeda H, Hiraga M, Kanno H, Nanzyo M, Nishiyama M, Kanahama K, Kanayama Y. 2012. Tolerance to salt stress and blossom-end rot in an introgression line, IL8-3, of tomato. Scientia Horticulturae 138, 1–6.

Urao T, Yakubov B, Satoh R, Yamaguchi-Shinozaki K, Seki M, Hirayama T, Shinozaki K. 1999. A transmembrane hybrid-type histidine kinase in *Arabidopsis* functions as an osmosensor. The Plant cell 11, 1743–54.

V

Vargas M, Crossa J, Van Eeuwijk FA, Ramírez ME, Sayre K. 1999. Using partial least squares regression, factorial regression, and AMMI models for interpreting genotype x environment interaction. Crop Science 39, 955–967.

Verbyla AP, Cavanagh CR, Verbyla KL. 2014. Whole-genome analysis of multienvironment or multitrait QTL in MAGIC. G3 Genes|Genomes|Genetics 4, 1569–1584.

Verta J, Landry CR, Mackay J. 2016. Dissection of expression-quantitative trait locus and allele specificity using a haploid / diploid plant system – insights into compensatory evolution of transcriptional regulation within populations. , 159–171.

Via S, Lande R. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution 39, 505–522.

Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W, Giovannoni J. 2002. A MADS-Box gene necessary for ruit ripening at the tomato ripening-inhibitor (Rin) locus. Science 296, 343–346.

W

Wan X, Steudle E, Hartung W. 2004. Gating of water channels (aquaporins) in cortical cells of young corn roots by mechanical stimuli (pressure pulses): Effects of ABA and of HgCl 2. Journal of Experimental Botany 55, 411–422.

Wang W, Pan Y, Zhao X, Dwivedi D, Zhu L, Ali J, Fu B. 2010. Drought-induced site-specific DNA methylation and its association with drought tolerance in rice (*Oryza sativa* L .). , 1–10.

Wang W, Vinocur B, Altman A. 2003. Plant responses to drought, salinity and extreme temperatures: Towards genetic engineering for stress tolerance. Planta 218, 1–14.

Wang W, Vinocur B, Shoseyov O, Altman A. 2004. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends in Plant Science 9, 244–252.

Wasilewska A, Vlad F, Sirichandra C, Redko Y, Jammes F, Valon C, Frei Dit Frey N, Leung J. 2008. An update on abscisic acid signaling in plants and more ??? Molecular Plant 1, 198–217.

Weiner JJ, Peterson FC, Volkman BF, Cutler SR. 2010. Structural and functional insights into core ABA signaling. Current Opinion in Plant Biology 13, 495–502.

Weller JL, Perrotta G, Schreuder MEL, Van Tuinen A, Koornneef M, Giuliano G, Kendrick RE. 2001. Genetic dissection of blue-light sensing in tomato using mutants deficient in cryptochrome 1 and phytochromes A, B1 and B2. The Plant Journal 25, 427–440.

West MAL, Kim K, Kliebenstein DJ, Van Leeuwen H, Michelmore RW, Doerge RW, St. Clair DA. 2007. Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in *Arabidopsis*. Genetics 175, 1441–1450.

Wilkinson JQ, Lanahan MB, Yen H-C, Giovannoni JJ, Klee HJ. 1995. An ethylene-inducible component of signal transduction encoded by never-ripe. Science 270, 1807–1809.

Wohlbach DJ, Quirino BF, Sussman MR. 2008. Analysis of the *Arabidopsis* histidine kinase ATHK1 reveals a connection between vegetative osmotic stress sensing and seed maturation. The Plant cell 20, 1101–1117.

X

Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E. 2008. A retrotransposon-mediated gene duplication underlies morphological variation of tomato Fruit. Science 319, 1527–1530.

Xu C, Liberatore KL, MacAlister CA, *et al.* 2015. A cascade of arabinosyltransferases controls shoot meristem size in tomato. Nature Genetics 47, 784–792.

Xu J, Ranc N, Muños S, Rolland S, Bouchet JP, Desplat N, Le Paslier MC, Liang Y, Brunel D, Causse M. 2013. Phenotypic diversity and association mapping for fruit quality traits in cultivated tomato and related species. Theoretical and Applied Genetics 126, 567–581.

Y

Yan, W., Kang, M. S., Ma, B., *et al.* 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop science, 47(2), 643-653.

Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. 2014. Advantages and pitfalls in the application of mixed-model association methods. Nature genetics 46, 100–6.

Ye N, Zhu G, Liu Y, Li Y, Zhang J. 2011. ABA controls H2O2 accumulation through the induction of OsCATB in rice leaves under water stress. Plant and Cell Physiology 52, 689–698.

Yin X, Kropff MJ, Stam P. 1999. The role of ecophysiological models in QTL analysis: the example of specific leaf area in barley. Heredity 82, 415–421.

Yu J, Pressoir G, Briggs WH, *et al.* 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature genetics 38, 203–8.

Z

Zanor MI, Osorio S, Nunes-Nesi A, *et al.* 2009a. RNA Interference of LIN5 in tomato confirms its role in controlling brix content, uncovers the influence of sugars on the levels of fruit hormones, and demonstrates the importance of sucrose cleavage for normal fruit development and fertility. Plant physiology 150(3), 1204–1218.

Zanor MI, Rambla JL, Chaïb J, Steppa A, Medina A, Granell A, Fernie AR, Causse M. 2009b. Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. Journal of Experimental Botany 60, 2139–2154.

Zeng ZB. 1994. Precision mapping of quantitative trait loci. Genetics 136, 1457–1468.

Zeng ZB, Kao CH, Basten CJ. 1999. Estimating the genetic architecture of quantitative traits. Genetical research 74, 279–89.

Zhang HX, Blumwald E. 2001. Transgenic salt-tolerant tomato plants accumulate salt in foliage but not in fruit. Nature biotechnology 19, 765–8.

Zhang J, Chen R, Xiao J, Qian C, Wang T, Li H, Ouyang B, Ye Z. 2007. A single-base deletion mutation in SlIAA9 gene causes tomato (*Solanum lycopersicum*) entire mutant. Journal of Plant Research 120, 671–678.

Zhang JX, Nguyen HT, Blum A. 1999. Genetic analysis of osmotic adjustment in crop plants. Journal of Experimental Botany 50, 291–302.

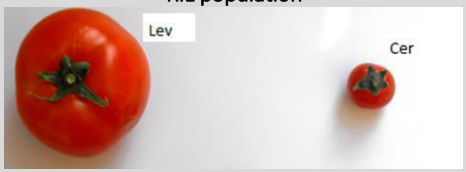
Zhou, Y., Lam, H. M., Zhang, J. 2007. Inhibition of photosynthesis and energy dissipation induced by water and high light stresses in rice. Journal of Experimental Botany, 58(5), 1207-1217.

Zhou X, Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. Nature genetics 44, 821–4.

Zhu C, Gore M, Buckler ES, Yu J. 2008. Status and prospects of association mapping in plants. The Plant Genome Journal 1, 5–20.

CHAPTER 2

Table 1. Overview of the plant material, genomic, phenomic and transcriptomic data analyzed in the successive chapters.

Scientific objectives	**(1)** Characterization of the 'genotype by watering regime' interactions. **(2)** Mapping of the QTLs/genes and 'QTL/gene by watering regime' interactions.		**(3)** Identification of differentially expressed genes and alleles between genotypes and watering regimes. **(4)** Identification of eQTLs and genes involved in controlling differential gene expression.	
Plant material	**GWA population**  141 small fruit accessions	**RIL population**  122 recombinant inbred lines		**Cervil, Levovil and F1 hybrid** 
Genotypic data	**6,100 SNP (Solcap array)** (Sim *et al.* 2012)	**501 SNP gained from parental line re-sequencing & genetic map** (Causse *et al.* 2013 & Pascual *et al.* 2015)		**Genome sequences of the parental lines** (Causse *et al.* 2013)
Trial years and locations	Avignon 2014 Agadir 2014 ← - - - - - - - - - - - - →	Avignon 2013 Agadir 2014	Avignon 2015 ← - - - - - - - - - - - - →	Avignon 2015
Watering regimes	**Control & Drought (-60%)**		**Control & Drought (-40%)**	
Phenotypic measurements *(underlined when measured in both locations, Avignon and Agadir)* *(at plant level in green color, at fruit level in red color)*	flowering date plant height stem diameter leaf length fruit number fresh weight fruit firmness DMW SSC pH total Vitamin C content glucose & fructose content citric & malic acid content	flowering date plant height stem diameter leaf length fruit number fresh weight fruit firmness DMW SSC pH total Vitamin C content / /	flowering date plant height stem diameter / / fresh weight / DMW SSC pH / glucose & fructose content malic acid content	
Transcriptomic measurements & technics	/	**Agilent four-plex array** whole genome leaves of both parents	**Microfluidigm qPCR** ~ 200 genes leaves and green fruits of the RIL	**RNA sequencing, Hiseq3000** whole genome leaves and green fruits of parents and F1
Data analysis	Genome wide association mapping	Linkage mapping Differential expression analysis	Linkage mapping	Differential expression analysis
Thesis chapters	**Chapter 4**	**Chapter 3**	**Chapter 5**	

CHAPTER 2: Material and Methods

This chapter briefly summarizes the plant materials and the methods used to answer the scientific objectives of this thesis. The data analysis is not developed here but provided in the following chapters.

---

1. Plant materials

The plant materials studied consisted in three sets as described in **Table 1.**

Cervil, Levovil and their F1 hybrid

Two undetermined inbred lines and their F1 hybrid were selected and analyzed for their response to water deficit at the phenotypic and transcriptomic levels. Cervil is cherry type tomato (*S. lycopersicum cerasiforme*) with small fruits (6–10 g) selected for its high aroma intensity. Levovil (*S. lycopersicum*) is a large fruited accession (90–160 g) with common taste. The F1 hybrid was obtained using Cervil as male parent and Levovil as female parent.

RIL population

A RIL population composed of 122 F7 recombinant inbred lines derived from the cross between Cervil and Levovil was used for the linkage mapping analysis. Previous QTL mapping studies were conducted using this same mapping population and reported hundreds of QTLs for various sensory, physical and chemical tomato fruit quality traits measured under optimal watering conditions only (Causse *et al.* , 2001; Saliba-Colombani *et al.* , 2001; Lecomte *et al.* , 2004).

GWA population

A GWA population constituted of 141 accessions (fresh weight from 2 to 46 g) encompassing the genetic diversity of the cultivated small fruited tomato was assembled for the association mapping analysis. Among the accessions, 105 accessions were previously investigated in a large scale genetic diversity analysis (Blanca *et al.* , 2015). Ten accessions were *S. pimpinellifolium* (closest wild ancestor from the tomato), among which five originated from Peru and one from Ecuador. A total of 110 accessions were *Solanum lycopersicum* var. *cerasiforme*. Among them, 20 originated from northern Peru and Bolivia, 9

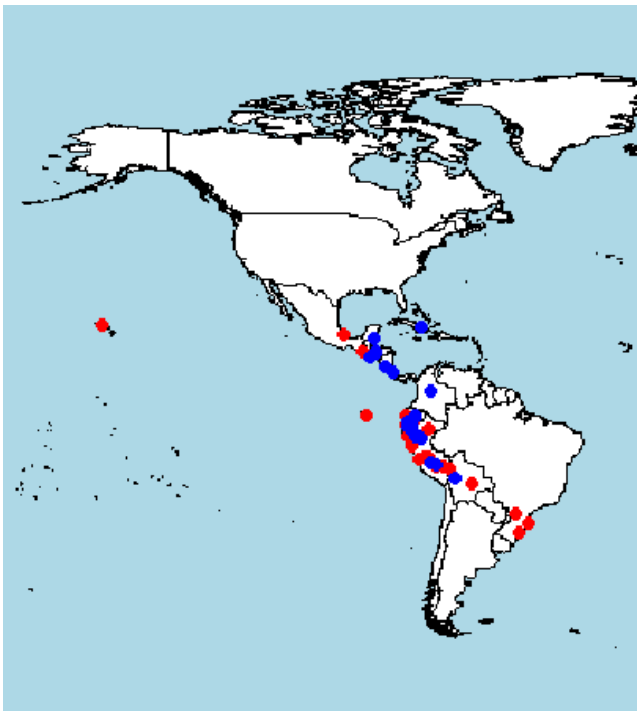


Figure 1: Geographical positions of the GWA accessions for which the Global Positioning System (GPS) coordinates of the collection site were available. Red dots: only GPS data available (23 accessions). Blue dots: GPS and climatic data available (28 accessions).

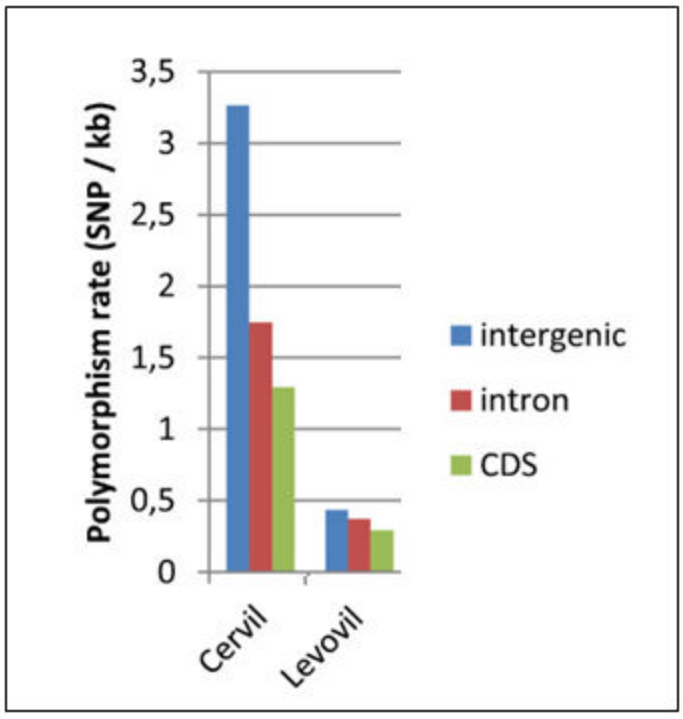


Figure 2: Distribution of the polymorphism rate in intergenic regions, introns and coding sequences (CDS) in Cervil and Levovil when compared with the reference genome sequence Heinz1706 (tomato genome annotation ITAG 2.4). *Adapted from Causse et al. 2013.*

from Ecuador and 81 from all over the subtropical areas of the world. Finally, 21 accessions belonged to a mixture genetic group mainly including commercial cherry tomatoes from the INRA collection and admixed genotypes between *S. pimpinellifolium*, *Solanum lycopersicum* var. *cerasiforme* and *S. lycopersicum* var. *lycopersicum*.

The geographical origin of the collection site was available in the form of Global Positioning System (GPS) coordinates for 51 accessions. Among them, 28 had also information concerning the annual precipitation and mean temperature from their collection site. Some accessions originated from the dry Peruvian coasts and valleys where others originated from the wetter Ecuadorian coasts. They were particularly interesting to include in our GWA population for analyzing the natural genetic diversity of response to water deficit in tomato **(Figure 1)**.

The seeds for the different plant material were kindly provided by the center of biological resources of INRA Avignon (CRB-Leg, France), the tomato genetics resource center of Davis university (TGRC, USA), the department of molecular biology and biochemistry of the university of Malaga (Spain), the Vavilov research institute of plant industry of St. Petersburg (VIR, Russia), the institute for conservation and improvement of Valencian agrodiversity of Valencia (COMAV, Spain) and the center of genetic resources of Wageningen (CGN, the Netherlands).

2. Genotypic data and genetic map

All genotypic data were available before the beginning of the PhD. Both parental lines of the RIL population were re-sequenced with a high depth: 19.6× for Cervil (covering 88.8 % of the genome with a minimum depth of 4×) and 9.2× for Levovil (72.7 % of the genome with a minimal depth of 4×, completed later to 91.5% at minimal depth of 4x through a 55x re-sequencing using the newly realized Illumina sequencing technology) (Causse *et al.* , 2013). These re-sequencing allowed the identification of more than 2 million SNPs and almost 128,000 InDels between both inbred lines and the tomato reference genome **(Figure 2).** Among the variants between Cervil and Levovil, 501 SNP were chosen spread over the genome, genotyped in the 122 RILs and used during the PhD to produce a genetic map. The map covers 1,090 cM corresponding to 98 % of the assembled tomato genome and is published in Plant Science with a re-analysis of the phenotypic data collected by Saliba-Colombani *et al.* (2000) and a comparison with others mapping designs. The article is available in Appendix 2 and includes a detailed description of the genetic map construction.

Figure 3: Picture of the trial conducted in INRA Avignon in 2015. Right and left rows corresponded to the same genotypes, grown under control and drought (-40% water), respectively. *(June)*

Figure 4: Picture of the trial conducted in the experimental site of the company '*GAUTIER Semences*' in 2014. Right and left rows corresponded to the same genotypes, grown under control and drought (-60% water), respectively. *(February)*

The GWA population was genotyped using the Tomato Infinium Array developed within the SolCAP project (http://solcap.msu.edu/). This array contains probes for 8,784 SNP spread over the entire tomato genome, among which 7,663 SNP passed the preliminary quality tests (Hamilton *et al.* , 2012; Sim *et al.* , 2012). After filtering for the missing data and the minor allele frequencies (MAF), the final genotype matrix for the association analysis was constituted of 6,100 SNP (see details in **Chapter 4**).

3. Experimental design and watering conditions

Four different experiments were conducted over three successive years, between 2013 and 2015, in France (Avignon) and Morocco (Agadir). Cervil, Levovil, their F1 hybrid and the RIL population were grown in Avignon in 2013 and 2015 and in Agadir in 2014, whereas the GWA population was grown in Avignon and Agadir in 2014. The three Avignon experiments were conducted in heated glasshouse in the experimental site of INRA Avignon from March to July **(Figure 3).** The Agadir experiment was conducted in the experimental site of the company '*GAUTIER Semences*' from December to May **(Figure 4).**

In each experiment, two watering regimes were applied to the plants: drought and control. Control treatment was applied according to ETp climatic data and the cultural coefficient for tomato crop under greenhouse with a maximal drainage of 25 % and a relative humidity of the peat substrate of 65 %. The drought treatment was progressively applied after flowering of the earliest genotype.

- **In Agadir in 2014 and in Avignon in 2013 and 2014**, water supply was reduced by 25% compared with control for one week, then decreased by 60 % until the end of the experiment, aiming to exacerbate the contrast with the control irrigation.

- **In Avignon in 2015**, water supply was reduced by 25 % compared with control for one week, then decreased by 40 % until the end of the experiment, aiming to achieve a moderate water deficit. Besides, shading screens were installed in the greenhouses to maintain the light intensity below 700 W/m² and subsequently reduce heat peaks during the warmest days.

In each experiment, plants were grown in 4 litters plastic pots filled with peat and watered with nutritive solution (2, 4, 6 mmol l−1, N, P, and K, respectively). Relative humidity of the peat substrate was controlled with GRODAN® moisture probes. Two plants per watering regime per accession were randomized in the greenhouses.

4. Phenotypic measurements

Traits related to phenology, plant vigor and fruit quality were assessed in the different experiments, under both watering conditions (control and drought). Traits related to phenology and plant vigor included flowering date, plant height, stem diameter, leaf length and fruit number. Traits related to fruit quality included fresh weight, firmness, dry mater content (DMW), soluble solid content (SSC), pH, total Vitamin C content, glucose, fructose, citric acid and malic acid content.

Details concerning the methods used to perform the measurements are given in each of the following chapter and **Table 1** indicates which traits were measured in each of the experiments.

5. Transcriptomic measurements

Microarray gene expression experiments

Microarray analyses were performed on samples of young leaves of Cervil and Levovil, grown under control and drought conditions in Avignon in 2014. RNA was extracted using the RNeasy Plant Mini Kit from Qiagen, with DNase I treatment. Hybridization experiments were achieved on Agilent fourplex arrays at IPS2 Transcriptomic Platform (INRA, France). The array contained 33,913 forward and 33,913 reverse probes representing 98 % of the known tomato genes. The array design is available through the GEO at NCBI (GPL20224) and on the CATdb database (array '4PLEX_TOMATO').

Two independent biological replicates per genotype were produced. In total, eight hybridizations were carried out comparing Cervil Control vs Cervil Drought and Levovil Control vs Levovil Drought. For each comparison, one technical replicate with dye swap was performed for each biological replicate (i.e., four hybridizations per comparison). More details concerning this experimentation and the subsequent analysis are given in **Chapter 3**.

RNA sequencing

RNA sequencing was performed on samples of young leaves and green fruits (at cell expansion) of Cervil, Levovil and their F1 hybrid, grown under control and drought conditions in Avignon in 2015 (3 replicates per accession per organ per condition). RNA was extracted using the 'Spectrum Plant Total RNA' kit from Sigma-Aldrich, with DNase I treatment. A total of 32 messenger RNA paired-end strand specific libraries were constructed (collaboration with Dr. Sylvain Santoni and Muriel Latreille at UMR AGAP, Montpellier SUPAGRO). The indexed libraries were combined in two lanes (one for the leaf samples and one for the fruit samples) and were subjected to 150-bp paired-end Illumina next generation sequencing at the GenoTool platform (Toulouse, France). More details concerning this experimentation and the subsequent analyses are given in **Chapter 5**.

Microfluidigm qPCR

Expression changes of 200 target transcripts were quantified in samples of young leaves and greens fruit (at cell expansion) of Cervil, Levovil, their F1 hybrid and the RIL progeny, grown under control and drought conditions in Avignon in 2015 (1 replicate per accession per organ per condition). RNA was extracted using the 'Spectrum Plant Total RNA' kit from Sigma-Aldrich, with DNase I treatment. Transcript expressions were measured using quantitative real time microfluidigm PCR at the Gentiane platform (Clermont-Ferrand, France). Different set of transcripts were quantified in the leaf and fruit samples. The target transcripts were chosen among the differentially expressed genes identified from the parental line RNA-seq libraries and according to their annotation and potential involvement in response to water deficit according to literature. More details concerning this experimentation and the subsequent analysis are given in **Chapter 5**.

Table 1 gives an overview of the plant material and the related genotypic, phenotypic and transcriptomic data collected and analyzed in each of the successive thesis chapters.

References

Blanca J, Montero-Pau J, Sauvage C, Bauchet G, Illa E, Díez MJ, Francis D, Causse M, van der Knaap E, Cañizares J. 2015. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. BMC Genomics 16, 257.

Causse M, Desplat N, Pascual L, *et al.* 2013. Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. BMC genomics 14, 791.

Causse M, Saliba-Colombani V, Lesschaeve I, Buret M. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 2. Mapping QTLs for sensory attributes. Theoretical and Applied Genetics 102, 273–283.

Hamilton JP, Sim S-C, Stoffel K, Van Deynze A, Buell CR, Francis DM. 2012. Single nucleotide polymorphism discovery in cultivated tomato via sequencing by synthesis. The Plant Genome Journal 5, 17.

Lecomte L, Saliba-Colombani V, Gautier A, Gomez-Jimenez MC, Duffé P, Buret M, Causse M. 2004. Fine mapping of QTLs of chromosome 2 affecting the fruit architecture and composition of tomato. Molecular Breeding 13, 1–14.

Saliba-Colombani V, Causse M, Gervais L, Philouze J. 2000. Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. Genome 43, 29–40.

Saliba-Colombani V, Causse M, Langlois D, Philouze J, Buret M. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. Theoretical and Applied Genetics 102, 259–272.

Sim S, Deynze A Van, Stoffel K, *et al.* 2012. High-density SNP genotyping of tomato (*Solanum lycopersicum* L .) reveals patterns of genetic variation due to breeding. 7, 1–18.

CHAPTER 3

CHAPTER 3: QTL mapping in a biparental progeny grown under two watering conditions and study of gene expression in both parental accessions

This chapter is in the form of an article published in *Theoretical and Applied Genetics*. The article describes the interactions 'genotype by watering regime' observed in a population of *recombinant Inbreed lines* (RIL) at the phenotypic (G x W) and genotypic (QTL x W) levels. Then, the study of differentially expressed genes in young leaves of the two parental accessions under two irrigation conditions is connected to some detected QTLs in order to provide candidate genes for response to water deficit in the cultivated tomato.

Genotype by watering regime interaction in cultivated tomato: lessons from linkage mapping and gene expression

doi:10.1007/s00122-015-2635-5 (Accepted: 04/08/2015 / Published online: 18/08/2015)

Authors

Elise Albert[a], Justine Gricourt[a], Nadia Bertin[b], Julien Bonnefoi[c], Stéphanie Pateyron[d], Jean-Philippe Tamby[d], Frédérique Bitton[a], Mathilde Causse[a], [§]

Affiliations

[a] INRA, UR1052, Génétique et Amélioration des Fruits et Légumes, 67 Allée des chênes, Centre de Recherche PACA, Domaine Saint Maurice, CS60094, Montfavet, 84143, France

[b] INRA, UR 1115, Plante et Système de cultures Horticoles, 228 Route de l'aérodrome, Centre de Recherche PACA, Domaine Saint Paul, CS40509, Avignon Cedex9, 84914, France

[c] GAUTIER Semences, route d'Avignon, Eyragues, 13630, France

[d] INRA, Institut of Plant Sciences Paris-Saclay (IPS2), UMR 9213/UMR1403, CNRS, INRA, Université Paris-Sud, Université d'Evry, Université Paris-Diderot, Sorbonne Paris-Cité, Rue de Noetzlin, Plateau du Moulon, 0rsay, 91405, France

[§] Corresponding author

Mathilde Causse

Mathilde.Causse@paca.inra.fr

Tel: +33 (0)4 32 72 28 03

Supplemental materials referred in this chapter are available in **Appendix 4.**

Abstract

As a result of climate change, drought will increasingly limit crop production in the future. Studying genotype by watering regime interactions is necessary to improve plant adaptation to low water availability. In cultivated tomato (*Solanum lycopersicum* L.), extensively grown in dry areas, well-mastered water deficits can stimulate metabolite production, increasing plant defenses and concentration of compounds involved in fruit quality, at the same time. However, few tomato QTLs (Quantitative Trait Loci) and genes involved in response to drought are identified or only in wild species. In this study, we phenotyped a population of 119 recombinant inbred lines derived from a cross between a cherry tomato (*S. l. cerasiforme*) and a large fruit tomato (*S. lycopersicum*), grown in greenhouse under two watering regimes, in two locations. A large genetic variability was measured for 19 plant and fruit traits, under the two watering treatments. Highly significant genotype by watering regime interactions were detected and resulted from re-ranking more than scale changes. The population was genotyped for 679 SNP markers to develop a genetic map. In total, 56 QTLs were identified among which 11 were interactive between watering regimes. These later mainly exhibited antagonist effects according to watering treatment. Variation in gene expression in leaves of parental accessions revealed 2,259 differentially expressed genes, among which candidate genes presenting sequence polymorphisms were identified under two main interactive QTLs. Our results provide knowledge about the genetic control of genotype by watering regime interactions in cultivated tomato and the possible use of deficit irrigation to improve tomato quality.

Key words

Genotype by environment interaction, QTL, Linkage mapping, Water deficit, Tomato, Gene expression

Key message

In tomato, genotype by watering interaction resulted from genotype re-ranking more than scale changes. Interactive QTLs according to watering regime were detected. Differentially expressed genes were identified in some intervals.

Introduction

Today, agriculture is one of the primary water users in many regions of the world, but global warming and drought risks are threatening plant growth and productivity. In particular, the Mediterranean region should experience more frequent drought episodes in the next decades (Gao and Giorgi 2008; Dai 2011). In this area, economic losses due to water limitation could be critical for the fruit and vegetable productions (Katerji *et al*. 2008). Thus, a better management of water resource for crop production is needed. A commonly accepted solution is to improve plant adaptation to low water availability.

Many studies have assessed plant response to different watering regimes in several species and shown the negative impact of water shortage on plant growth and yield. Reviews of the different morphological, physiological and molecular changes induced by water limitation are available (Chaves *et al*. 2003; Hirayama and Shinozaki 2010; Blum 2011; Farooq *et al*. 2012; Silva *et al*. 2013). In particular, these studies highlight the role of secondary metabolites and carbohydrates in plant protection against photo-oxidative stress induced by stomata closure and cell dehydration (Gershenzon 1984; Chaves *et al*. 2009; Shaar-Moshe *et al*. 2015). These drought induced secondary metabolites are also essential compounds for quality of plant food products. For instance, ascorbic acid (Vitamin C), an important antioxidant for human diet, is well known for its role in scavenging reactive oxygen species (ROS) in plant under water stress (Jiang and Zhang 2002; Stevens *et al*. 2008). Evidence of the crucial role of sugars in osmotic adjustment induced by drought has been obtained in several species such as tomato (Bertin *et al*. 2000), *Arabidopsis thaliana* (Anderson and Kohorn 2001) or white lupine (Chaves *et al*. 2002). Well-mastered water deficit can thus help to achieve a tradeoff between crop yield and quality, reducing non-beneficial water consumption in crop production at the same time. Such deficit irrigation strategies are particularly under consideration in fleshy fruits for which consumers are expecting healthier and tastier products (Chaves and Oliveira 2004; Nora *et al*. 2012; Ripoll *et al*. 2014). However, knowledge about the QTLs (Quantitative Trait Loci) and genes involved in plant response to water deficit and their interactions is still lacking (Shinozaki and Yamaguchi-Shinozaki 2007; Ashraf 2010; Tardieu *et al*. 2011). Despite the identification of hundreds of genes involved in response to drought by gene expression analysis associated or not to linkage mapping, their roles and modes of action are still poorly understood (Lovell et al.

2015; Shaar-Moshe *et al*. 2015). Besides, these genes were mainly identified in *Arabidopsis thaliana* under laboratory conditions (Seki *et al*. 2002) or in cereals (Langridge 2006; Barnabas *et al*. 2007). Not all of them are involved in adaptation process (Chaves *et al*. 2003).

Understanding the genetic determinism of genotype by watering regime interactions will constitute a basis for crop improvement, allowing the identification of favorable alleles under drought conditions (Collins *et al*. 2008; Tardieu and Tuberosa 2010). The emergence of high-throughput genomic tools and the availability of genome sequences for many crops facilitate the decomposition of genotype by environment interactions into underlying QTLs and/or genes (Des Marais *et al*. 2013; El-Soda *et al*. 2014b). These approaches will provide a better understanding of the ability of an individual genotype to adapt its phenotype in response to environmental constraints, a phenomenon termed as 'phenotypic plasticity' (Via and Lande 1985; Schlichting 1986). In the context of multiple environments, two main approaches are applied to map QTL by environment (QTL x E) interactions. The first one, looking at the effects of a given QTL in each environment, identifies different interactive QTL types (Malosetti *et al*. 2007; Yang *et al*. 2008; van Eeuwijk *et al*. 2010; Korte *et al*. 2012; Li *et al*. 2015). In most cases, QTLs have a strong effect in one environment, but lower effect in another (differential sensitivity effect). More rarely and mainly in wild species, QTLs can show opposite effects for a same trait in different environments (antagonist effect). The second strategy consists in constructing composite variables measuring phenotypic plasticity to deal with univariate QTL mapping models. These variables can be ratio or difference between the values of a trait measured in two environments or parameters from reaction norms (Tétard-Jones *et al*. 2011; El-Soda *et al*. 2014a; Coupel-Ledru *et al*. 2014). The two methods substantially overlap but the second one gives additional statistical power with more QTLs exceeding the threshold (Tétard-Jones *et al*. 2011; El-Soda *et al*. 2014a).

In cultivated tomato (*Solanum lycopersicum* L.), a water demanding crop extensively grown in Mediterranean region, QTLs for chemical and physical fruit quality were previously mapped, but no attention was paid to the interaction with abiotic factors (Causse *et al*. 2001; Saliba-Colombani *et al*. 2001; Pascual *et al*. 2015). In this species, genes involved in response to abiotic stress were mainly characterized by translational genetics and genetic engineering with genes identified in *Arabidopsis thaliana* (Hsieh *et al*. 2002; Rai *et al*. 2013;

Zhu *et al*. 2014). Studies of natural variability of the interactions with environmental constraints have focused on salt stress (Foolad *et al*. 2003; Foolad 2004; Uozumi *et al*. 2012; Kissoudis *et al*. 2015; Asins *et al*. 2015). In few accessions, authors have reported a positive effect of mild to moderate water deficit on tomato fruit quality, with an increased fruit soluble solids levels and an increased concentration of hexoses (Mitchell *et al*. 1991; Bertin *et al*. 2000; Patanè and Cosentino 2010; Zheng *et al*. 2013). Besides, Foolad *et al*. (2003) and Semel *et al*. (2007) have shown some genetic variability in response to water deficit at the seed and plant levels. However, to date, no QTL by watering regime interaction mapping studies were conducted in the cultivated tomato. More precisely, introgression line populations involving wild relative species (*Solanum habrochaites* and *Solanum pennellii*) were used to map QTLs and the large confidence intervals obtained made the transposition difficult into the cultivated tomato (Gur *et al*. 2011; Easlon *et al*. 2014).

In this context, the aims of the present study are to: (1) describe genotype by watering regime interactions for plant and fruit traits in cultivated tomato genotypes, (2) decipher the inheritance patterns of these interactions and (3) identify candidate genes as putative targets for breeding. We addressed these aims by phenotyping a population of recombinant inbreed lines (RILs), grown in greenhouse under two watering regimes (Drought and Control), in two locations (Morocco and France). Linkage mapping was conducted to identify QTLs controlling genotype by watering regime interactions. Microarray analysis of gene expression in young leaves from the parental genotypes grown under the two watering regimes was performed to identify differentially expressed genes between the watering conditions. Finally, gene expression data were used to identify candidate genes underlying two interactive QTLs. The genetic determinism of genotype by watering regime interactions in cultivated tomato and the possible use of water deficit to improve tomato fruit quality in future breeding programs are discussed.

Material and Methods

Plant material and experimental design

The RIL population consisted in 119 F7 recombinant inbred lines. This population was developed from an intraspecific cross between two inbred lines, Cervil and Levovil (described in Saliba-Colombani *et al*. 2000). Cervil is a cherry type tomato (*S. lycopersicum cerasiforme*) with small fruits (6 – 10 g), whereas Levovil (*S. lycopersicum*) is a large fruited accession (90 – 160 g). In 2013, the plants, including the 119 RILs and the two parents, were grown in a heated glasshouse in INRA Avignon (Avi, France) from March to July. Besides, from December 2013 to May 2014, plants were grown in an unheated plastic greenhouse in the experimental site of the company GAUTIER Semences in Agadir (Aga, Morocco). In the greenhouses, the mean air temperature was 23°C and 26°C during day, 16° and 18°C during night, in France and Morocco, respectively. In each experiment, plants were grown in 4 litters (L) plastic pots filled with peat (Klasmann 165) and watered with nutritive solution (2, 4, 6 mmol l^{-1}, N, P, and K, respectively).

In both locations, two watering regimes were applied to the plants in each trial: drought (D) and control (C). Control treatment was applied according to ETp climatic data and the cultural coefficient for tomato crop under greenhouse with a maximal drainage of 25% and a relative humidity of the peat substrate of 65%. The drought treatment was progressively applied after flowering of the second truss of Cervil (considered as a reference early genotype): water supply was reduced by 25% compared with control for one week, then decreased by 60% until the end of the experiment, aiming to exacerbate the contrast with the control irrigation. Throughout the experiment, relative humidity of the peat substrate was controlled with a GRODAN® moisture probe and monitored in drought pots between 25% and 30%. Genotypes were randomized within rows and watering regime was applied by row. For each experiment, two plants per watering regime per genotype were placed side by side. To insure relatively homogenous environment in the greenhouses, trials were surrounded with one row of border tomato plants.

Plant and fruit phenotyping

In the two trials, under the two watering regimes, RIL plants were phenotyped for traits describing plant performance and fruit characteristics. Vegetative vigor and phenology were measured daily on every plant. Flowering date of the first flower from the 5^{th} truss in

Avignon (Flw.Avi) and 4th in Agadir (Flw.Aga) were assessed in number of days after sowing. su implantation height (Ht.Avi and Ht.Aga, in cm), stem diameter (Diam.Avi and Diam.Aga, in mm) and leaf length (Leaf.Avi and Leaf.Aga, in cm) under truss were measured on the 4th truss in Avignon and the 5th truss in Agadir. The number of fruits per plant (Nbfruits.Avi) was assessed in Avignon only by counting all the fruits from the second truss to the sixth truss.

Fruit measurements were conducted on tomatoes harvested daily on the basis of their red color to ensure a homogeneous ripening stage. At least, ten fruits per genotype per watering regime were harvested in the two trials on 3rd to 6th truss. For each fruit, fresh weight (FW.Avi and FW.Aga, in g) and firmness (FIR.Avi and FIR.Aga, in Durofel index) were measured. Besides, in Avignon only, harvested fruits were pooled in three groups of three to four fruits per watering regime. These pools constituted the three replicates for chemical analysis. In each pool, a quarter of fruit pericarp was sampled and dried in an oven at 60°C for four days to measure dry matter content (DMW.Avi, in %). Then, half of each fruit pool was mixed in juice to measure pH (pH.Avi) and soluble solid content (SSC.Avi, with a refractometer, in °Brix). Pericarps were sampled from the remaining fruit of each pool, frozen with liquid nitrogen and ground into fine powder with an IKA® mill for total Vitamin C (VitCFM.Avi, in mg per 100g of fresh matter) assessment according to Stevens *et al*. (2006). Average total Vitamin C per genotype per watering regime was also expressed in mg per 100g of dry matter (VitCDM.Avi) using DMW.Avi.

The average yield per genotype in Avignon (Yield.Avi, in g fresh weight per plant from truss 2 to 6) was estimated in each watering regime as the product of the average fruit fresh weight (FW.Avi) by the average number of fruits (Nbfruits.Avi). Finally, a total of 19 traits were assessed, under two watering conditions each, considering as two separate traits a same phenotypic measurement carried out in the two locations. The phenotypic means in the RIL population are available in **Supplemental Table 1**.

Statistical analyses on phenotypic data

Statistical analyses were performed on RIL raw data of each trial separately (Avignon and Agadir) using R 3.2.0 (R Development Core Team 2012). Prior to any analysis of variance (ANOVA), data were corrected for normality deviation using Box and Cox transformations (Box and Cox 1964). Effect of watering regime and interaction with genotype were tested by the ANOVA model: $Y_{ij} = \mu + G_i + W_j + G_i * W_j + e_{ij}$, where Y_{ij} was the phenotypic value of

genotype *i* in watering regime *j*, μ the overall mean, *Gi* the fixed effect of genotype *i*, W_j the fixed effect of watering regime *j*, and e_{ij} the residual error effect. Residuals were spatially plotted to control for a potential microenvironment effect due to side by side position of the two replicates of a given genotype in each watering regime in the experiments. No significant pattern was identified and we chose to not include a spatial effect in the ANOVA model. To further describe the genotype by watering regime interaction, the *G x W* sum of squares was partitioned into part associated with heterogeneous variance (scale change) and part due to imperfect correlation between genotypes (rank change) using the method 1 of Muir *et al*. (1992). For FW, Nbfruits, VitCFM and VitCDM, ecovalences were calculated according to Wricke *et al*. (1964) to measure participation of independent genotypes in interaction.

Then, genetic variability expressed at a given watering regime was assessed using the following ANOVA model: $Y_{ij} = \mu + G_i + e_{ij.}$ (G_i and e_{ij} as random). Restricted maximum likelihood estimates (REML) of variances of the random factors (σ^2_G and σ^2_e) were computed. Broad sense heritability was calculated in each watering regime as: $H^2 = \sigma^2_{G\ /\ } \sigma^2_{Total}$, with $\sigma^2_{Total} = \sigma^2_G + \sigma^2_e$. For the different traits in the two trials, correlations between H^2 and σ^2_G measured under drought and under control conditions were estimated by Spearman coefficient and declared significant when *P-value* were below 0.05.

For subsequent analyses, for each watering regime and each trial, the average genotypic values over replicates were computed. For each phenotypic trait *k* in each trial, plasticity (Δk) was calculated on the mean of the trait under each watering regime (drought D_k; control C_k) as: $\Delta k = (D_k - C_k)/C_k$. In the different watering regimes and in the different trials, Pearson correlations between means of traits and between means and plasticity data were calculated. A Mantel test was performed to measure changes in correlation between traits, according to the watering regime, in the two trials. *P-value* was calculated after 9999 permutations and an alpha threshold of 0.05 was considered to declare significance.

Plant genotyping and genetic map building

Genotyping and map construction are described in Pascual *et al*. (2015). Briefly, a set of 754 polymorphic markers between the two parents were genotyped in the RIL population: 679 are SNP markers derived from the re-sequencing of the parent genomes (see Causse *et al*. 2013), two are RAPD markers (random amplified polymorphic DNA) and 73 are RFLP markers

(restriction fragment length polymorphism) present in a previous genetic map from this progeny (Saliba-Colombani *et al*., 2000). The Chi-square test (α = 0.0001%) revealed that 98% of the markers (739/754) did not show any segregation distortions and were used in genetic mapping as described in Pascual *et al*. (2015). When several markers colocalized, only the one with the lowest percentage of missing data was conserved. The final genetic map obtained included 501 loci (501/754) and was covering 1090 cM corresponding to 98% of the assembled tomato genome (Tomato Genome Consortium, 2012). Markers were named according to their positions on the tomato genome (assembly v2.5), as Y01_56000045 at position 56,000,045 pb on chromosome 1. The genotypic data of the RIL population are available in the **Supplemental Table 2.** The genetic map is available in Pascual *et al*. (2015).

QTL and QTL x watering regime mapping

In each watering regime and each trial, the plasticity data and average phenotypic values were used for QTL detection. When distributions were skewed, corrections for normality were applied: Log_{10}(Ht.Avi); Log_{10}(Nbfruits.Avi); Log_{10}(FW.Avi); √(Diam.Aga); Log_{10}(Leaf.Aga); Log_{10}(Ht.Aga); Log_{10}(FW.Aga) and Log_{10}(Nbfruits.Avi). The QTL detection was performed by simple interval mapping (Lander and Botstein 1989) using the EM algorithm method implemented in R/QTL package (Broman *et al*. 2003). A 1000-permutation test was performed to estimate significant threshold. LOD threshold was 3.08, corresponding to a genome-wide significance level of α = 0.05. For each detected QTL, position, LOD score, marker at the LOD score peak, confidence interval (genetic-CI, LOD decrease of one unit), average phenotypic values of the two parental alleles and percentage of phenotypic variation explained (PVE) were displayed. QTL effects were calculated as: (Cervil mean allele - Levovil mean allele)/2. The genetic-CI were translated into physical intervals (Physical-CI in Mbp) onto the tomato genome (assembly v2.5). When a QTL was detected in one watering regime, the effect and PVE were also calculated in the second watering regime. Then, to test for watering regime and interaction with marker, two different ANOVA tests were developed in R 3.2.0 (R Development Core Team 2012):

(1) a “watering regime effect test” (W test) that compares a model with marker genotype and watering regime effect, to a model without watering regime effect ;

(2) an “interaction effect test” (G x W test) that compares the full model, including the effect of the marker genotype and its interaction with the watering regime, to the one that doesn’t include interaction.

This testing method is inspired from the multi-trait mixed model (MTMM) developed by Korte *et al*. (2012) for association analysis, considering only fixed effects. To correct for multiple testing, significance thresholds *P-value* corresponding to a genome wide significance level of α = 0.05 were computed by a 1000-permutation test (*P-value* $_{W\ test}$ = 2.21 x 10^{-4}; *P- value* $_{G\ x\ W\ test}$ = 1.93 x 10^{-4}). This procedure allowed displaying *P-value* for watering regime and interaction effect for marker at the QTL LOD score peak and to identify interactive markers not identified in the QTL mapping step.

Microarray experiment on parental accessions

Microarray analyses were performed on an Agilent four-plex arrays at IPS2 Transcriptomic Platform (INRA, France). For each of the 34,727 tomato genes (assembly v2.4, Tomato Genome Consortium 2012), a set of ten 60-mer probes were designed using the eArray Agilent software. Considering melting temperature and specificity criteria, the best probe for each gene was chosen and synthesized in forward and reverse sense. The array contained 33,913 forward and 33,913 reverse probes representing 98% of the known tomato genes, each printed in technical duplicate and 18 controls in triplicate. The array design is available through the GEO at NCBI (GPL20224) and on the CATdb database (Gagnot *et al*. 2008): array ‘4PLEX_TOMATO’.

Samples of young leaves of Cervil and Levovil, grown under the two watering regimes in Avignon, were harvested, immediately frozen and ground in liquid nitrogen with an IKA® mill. RNA was extracted using the RNeasy Plant Mini Kit (Qiagen) following the manufacturer's protocol, with DNase I treatment. RNA quality was assessed on Agilent 2100 Bioanalyser using Nano 6000 kit. Two independent biological replicates per genotype were produced. The labeling of cRNAs with Cy3-dUTP or Cy5-dUTP was randomly performed as described in Two-Color Microarray-Based Gene Expression Analysis Low Input Quick Amp Labeling manual (© Agilent Technologies, Inc.). Hybridization and washing were performed according to Agilent Microarray Hybridization Chamber User Guide instructions (© Agilent Technologies, Inc.). A two microns resolution scanning was performed using InnoScan900

scanner (InnopsysR, Carbonne, France) and raw data were extracted using the MapixR software (InnopsysR, Carbonne, France). In total, eight hybridizations were carried out comparing Cervil Control *vs* Cervil Drought and Levovil Control *vs* Levovil Drought. For each comparison, one technical replicate with dye swap was performed for each biological replicate (i.e. four hybridizations per comparison).

Microarray statistical analyses

Analysis was conducted with the R 3.2.0 software (R Development Core Team 2012). For each array, the raw data comprised the logarithm of the median feature pixel intensity at wavelengths 635 nm (red) and 532 nm (green). A global intensity-dependent normalization using the loess procedure (Yang *et al*. 2002) was performed to correct the dye bias. The differential analysis was based on the $\log_{10}$ of the fold changes between watering regimes averaging over the duplicate probes and over the technical replicates. Hence, the numbers of available data for each gene equals the number of biological replicates. Empirical Bayes posterior means were computed to smooth the specific variances and used to calculate the moderated t-test (function SqueezeVar of the library limma, Smyth 2005). Under the null hypothesis, no evidence that the specific variances changed between probes was highlighted and consequently the moderated t-statistic was assumed to follow a standard normal distribution. To control the false discovery rate, *P-values* were adjusted with the Bonferroni approach (Storey 2007) using the R library *kerfdr* (Guedj *et al*. 2009). We considered as being differentially expressed the probes with a Bonferroni adjusted *P-values* below 0.05. A Venn diagram was drawn to indicate genes differentially expressed in Cervil and/or in Levovil. Gene Ontology (GO) terms were associated with the differentially expressed genes using genome annotation v2.4 (Tomato Genome Consortium 2012). A maximum of seven GO terms were associated to each gene. A total of 33% of the differentially expressed genes (DEG) were not associated to any GO term due to a lack in the genome annotation. Identification of GO terms related to biological process that were significantly enriched within the differentially expressed genes compared with the tomato genome was achieved using the 'GO term enrichment analysis' tool (http://bioinfo.bti.cornell.edu/tool/GO/GO_enrich.html) based on the 'GO::TermFinder' program described in Boyle *et al*. (2004). GO terms were declared significantly enriched when Bonferroni corrected *P-value* was below 0.05. Separate analyses for up and down

regulated genes in Cervil only, in Levovil only and common to the two accessions were conducted. All raw and normalized data are available through the GEO at NCBI (GSE69898) and through the CATdb database (Gagnot *et al*. 2008): project '4PLEX_TOMATO_2013_03'. The list of genes differentially expressed and the GO data are available in the **Supplemental Table 3**.

Candidate gene selection under interactive QTLs

Microarray data on young leaves were used to identify candidate genes under interactive QTLs for plant traits potentially controlled at leaf level. We focused on two QTLs with short physical intervals: a QTL for flowering time on chromosome 2 (3.23 Mbp) and a QTL for stem diameter on chromosome 4 (2.55 Mbp). Within the QTL confidence intervals, differentially expressed genes between watering regimes in Cervil and/or in Levovil were selected (adjusted *P-values* below 0.05). Among these genes, polymorphism data obtained through the re-sequencing of parental accessions (Causse *et al*. 2013) was screened to identify nucleotide variants between Cervil and Levovil (SNPs and Indels). Re-sequencing depth was 19.6 x for Cervil (covering 88.8 % of the genome with a minimum depth of 4x) and 9.2 x for Levovil (72.7% of the genome with a minimal depth of 4x). Variants were classified in four categories as specified in Causse *et al*. (2013): 'High' for polymorphisms which modified splice sites or start/stop codons (loss or gain); 'Moderate' for non-synonymous polymorphisms in coding regions, 'Low' for variants in coding regions which do not change the amino acid sequence and 'Modifier' for polymorphisms located in upstream and downstream regions or in UTR or intergenic regions.

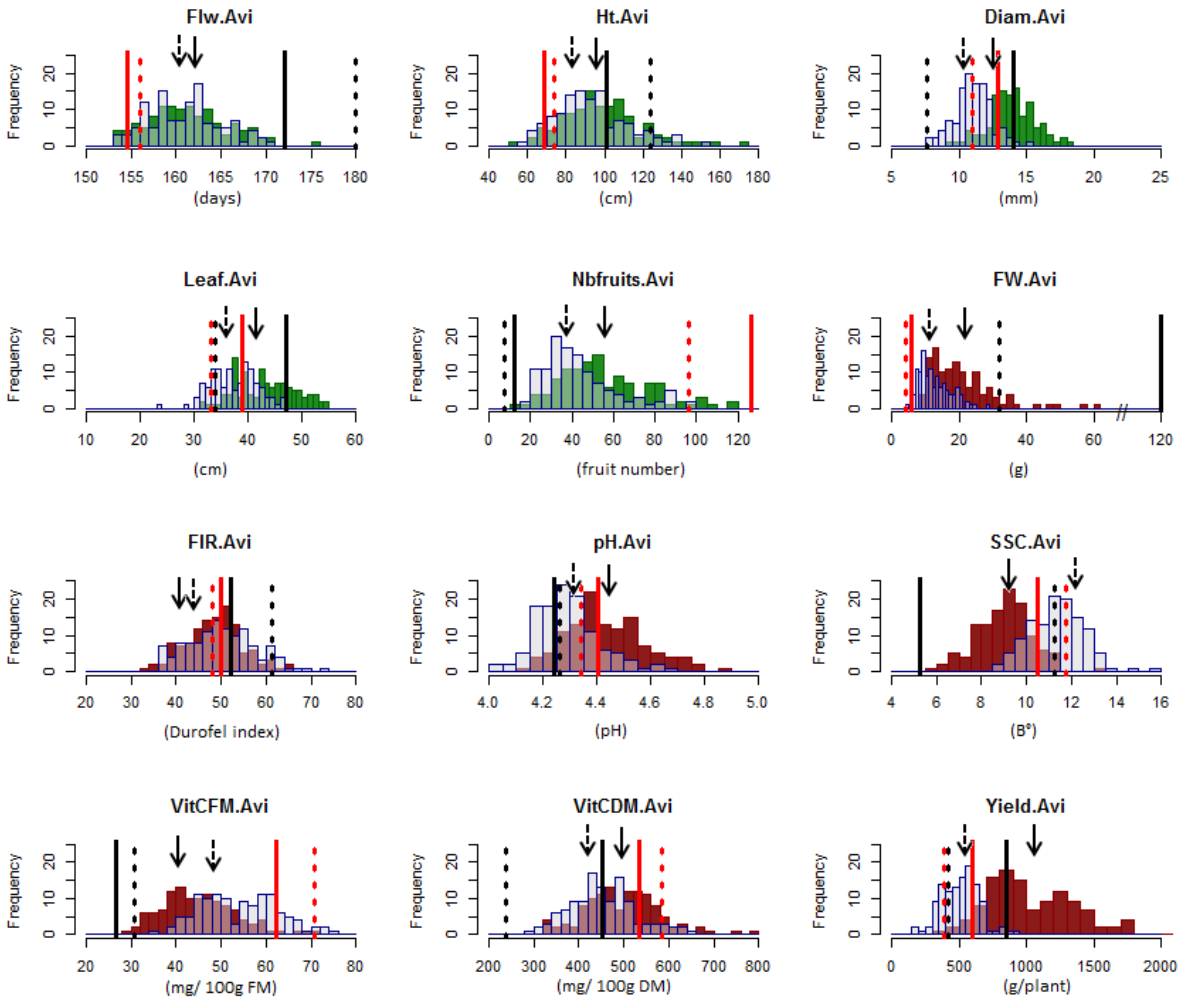


Fig 1 Distribution of the average plant and fruit traits in the recombinant inbred lines (RILs) grown under two watering regimes in Avignon. Opaque color indicates trait values under control treatment and transparent color trait values under drought treatment (green: plant traits; red: fruit traits). The parental mean values are indicated: full red line for Cervil in control treatment, dashed red line for Cervil in drought treatment, full black line for Levovil in control treatment and dashed black line for Levovil in drought treatment. The black arrows represent the RIL population means: dashed arrow for drought and full arrow for control treatment.

Results

To study the genetic variability of tomato plant and fruit response to water deficit, we mapped QTLs in a population derived from a cross between a cherry tomato and a large fruit accession, grown under two watering regimes (control and drought), in two locations (Agadir and Avignon). Phenotypic data from the two locations were analyzed separately because fewer and/or different phenotypic traits were measured in Morocco. Thus, a total of 19 traits were assessed in each watering condition, considering as two separate traits a same phenotypic measurement carried out in the two locations. However, for the six common traits, correlations between the two experiments under the two watering regimes were highly significant (*P-value* < 0.001), suggesting a good repeatability of the measurements through the experiments **(Supplemental Table 4)**. Significant watering regime by location interactions were detected for these six traits (data not shown), which may reflect the consequences of differences in temperature and day length between locations.

1. Phenotypic variability and genotype by watering regime interactions

Both parental accessions were impacted by the drought treatment. From 1.1 to 33.1 times significantly higher percentages of phenotypic changes due to drought were observed in Levovil than in Cervil for 15 of the 19 measured traits (excepted for Diam.Aga, Leaf.Aga, FW.Aga and pH.Avi) **(Fig 1 and Supplemental Fig 1)**. It suggested a higher susceptibility to water deficit in the large fruit accession. In particular, under drought, SSC was increased by 114.6% and FW.Avi decreased by 71.8% in Levovil, whereas the SSC gain was only 11.7% and the FW loss 33.3% in Cervil.

In the RILs, we surveyed ample phenotypic variation for the plant and fruit traits, under the two watering regimes, in the two locations (coefficient of variation ranking from 2.48 to 46.27%; average CV = 17.68%) **(Fig 1 and Supplemental Fig 1)**. Transgressions beyond parental values in the two directions were observed for all traits in the two watering regimes, except for Flw.Aga, FW.Aga, FW.Avi and Nbfruits.Avi for which RIL phenotypic means were comprised between Cervil and Levovil means. At plant level, in average in the RILs and in the two locations, the drought treatment tended to reduce stem diameter (Avi : -20.7%; Aga: -30.3%), leaf length (Avi: -13.4%; Aga: -25.8%) and fruit number (Avi: -21.7%). At fruit level, drought treatment reduced FW (Avi: -37.7%; Aga: -25.4%) and

Table 1 Effect of genotype (G), watering regime (W) and the interaction (G x W) on the plant and fruit traits. G, W and G x W indicate the significance of the ANOVA test for genotype, watering regime and interaction effects, respectively. SS G, SS W and SS G x W display the proportion of each effect in the total sum of squares, respectively. Scale and Rank are the proportion of interaction associated with heterogeneous variance and imperfect correlation between genotypes, respectively, using method 1 according to Muir *et al*. (1992). 'H^2 Control' and 'H^2 drought' indicate the broad sense heritabilities in control and drought treatment, respectively.

Trait	G	SS G (%)	W	SS W (%)	G x W	SS G x W (%)	Scale (%)	Rank (%)	H^2 Control	H^2 Drought
Plant traits										
Flw.Avi	***	75.90	ns	0.14	**	10.67	1.44	98.56	0.74	0.65
Flw.Aga	***	85.87	ns	0.37	***	6.47	15.06	84.94	0.88	0.83
Diam.Avi	***	28.72	***	39.72	***	16.86	3.84	96.16	0.52	0.47
Diam.Aga [a]	***	27.07	***	49.62	***	14.15	1.69	98.31	0.64	0.62
Leaf.Avi	***	50.71	***	21.87	***	13.63	6.54	93.46	0.63	0.65
Leaf.Aga [a]	***	34.19	***	41.57	***	14.93	3.02	96.98	0.72	0.60
Ht.Avi [a]	***	80.92	***	1.92	***	9.76	0.62	99.38	0.87	0.82
Ht.Aga [a]	***	79.45	***	0.68	***	11.35	1.32	98.68	0.81	0.83
Nbfruits.Avi [a]	***	68.82	***	7.57	***	12.51	0.30	99.70	0.76	0.74
Fruit traits										
FW.Avi [a]	***	55.65	***	21.13	***	5.70	5.83	94.17	0.83	0.69
FW.Aga [a]	***	59.88	ns	9.80	***	6.95	3.98	96.02	0.71	0.71
FIR.Avi	***	26.74	***	0.26	***	7.97	3.01	96.99	0.32	0.30
FIR.Aga	***	37.89	*	0.03	***	11.00	0.19	99.81	0.41	0.46
DMW.Avi	***	23.96	***	29.82	***	14.87	0.00	100.00	0.32	0.43
pH.Avi	***	38.06	***	14.66	ns	10.65	4.75	95.25	0.89	0.90
SSC.Avi	***	30.10	***	39.88	***	12.14	0.15	99.85	0.66	0.44
VitCFM.Avi	***	48.73	***	16.33	***	10.75	0.00	100.00	0.59	0.53

[a] Data transformed for skewed distribution

*** *P-value* below 0.001, ** between 0.001 and 0.01; and * between 0.01 and 0.05.

ns indicates non-significant *P-value*.

yield (Avi: -50.3%), but increased SSC (Avi: +26.3%) and DMW (Avi: +30.7%). The average plant height, flowering time, fruit firmness and pH were poorly affected by watering deficit whatever the location (Ht.Avi = -5.6%; Ht.Aga = 2.4%; Flw.Avi = -0.2%; Flw.Aga = +0.6%; FIR.Avi = +3.4%; FIR.Aga = +0.8%; pH.Avi = -3.2%). Vitamin C was differently impacted by drought depending on the unit in which it was expressed: in average increased when expressed relatively to fresh matter (+26.3%) and reduced when expressed relatively to dry matter (-8.9%).

For all the traits measured with replicates, in the two locations, genotype by watering regime interaction was significant (*P-value* < 0.01), except for pH, which was also poorly variable in the population ($CV_{drought}$ = 2.80%; $CV_{control}$ = 3.34%) **(Table 1)**. These interactions represented between 5.70 (FW.Avi) and 16.86% (Diam.Avi) of the total sum of square, a proportion lower than the one due to the genotype (between 23.96 and 85.87%) for all traits and lower than the one due to the watering regime (between 14.66 and 49.62%) for ten of the seventeen traits. Interaction partitioning according to method 1 from Muir *et al*. (1992) showed that the observed interactions were mainly due to genotype re-ranking across watering regimes (84.94 to 100%) and poorly to scale changes (0 to 15.6%, **Table 1**).

The broad-sense heritabilities were comprised between 0.30 (FIR.Avi under drought) and 0.90 (DMW.Avi under drought), with consistence between experiments for the six common traits in the two watering regimes (*P-value* $_{drought}$ = 0.03 and r $_{drought}$ = 0.85; *P-value* $_{control}$ = 0.03 and r $_{control}$ = 0.86) **(Table 1, Supplemental Fig 2).** Correlations between heritability and genetic variance were significant in both conditions (for H^2: *P-value* < 2.2×10^{-16} and r = 0.92; for varG: *P-value* = 8.56×10^{-11} and r = 0.97). Thus, genetic variability was conserved across watering treatments, interactions being associated with re-ranking among genotypes more than heterogeneous variance between watering regimes **(Table 1)**.

2. Changes in correlations according to fruit weight and watering regime

We observed a negative linear relationships between FW in control condition and FW plasticity (ΔFW) in the two experiments (r $_{Avi}$ = -0.51 and *P-value* $_{Avi}$ = 2.08×10^{-09}; r $_{Aga}$ = -0.44 and *P-value* $_{Aga}$ = 9.69×10^{-07}) **(Fig 2a and Supplemental Fig 3).** As ΔFW measures the percentage of FW gain or loss due to the drought treatment, these negative relationships indicated higher fresh weight loss for lines with large fruits. On another side, correlation between FW in control condition and vitamin C plasticity when expressed relatively to fresh

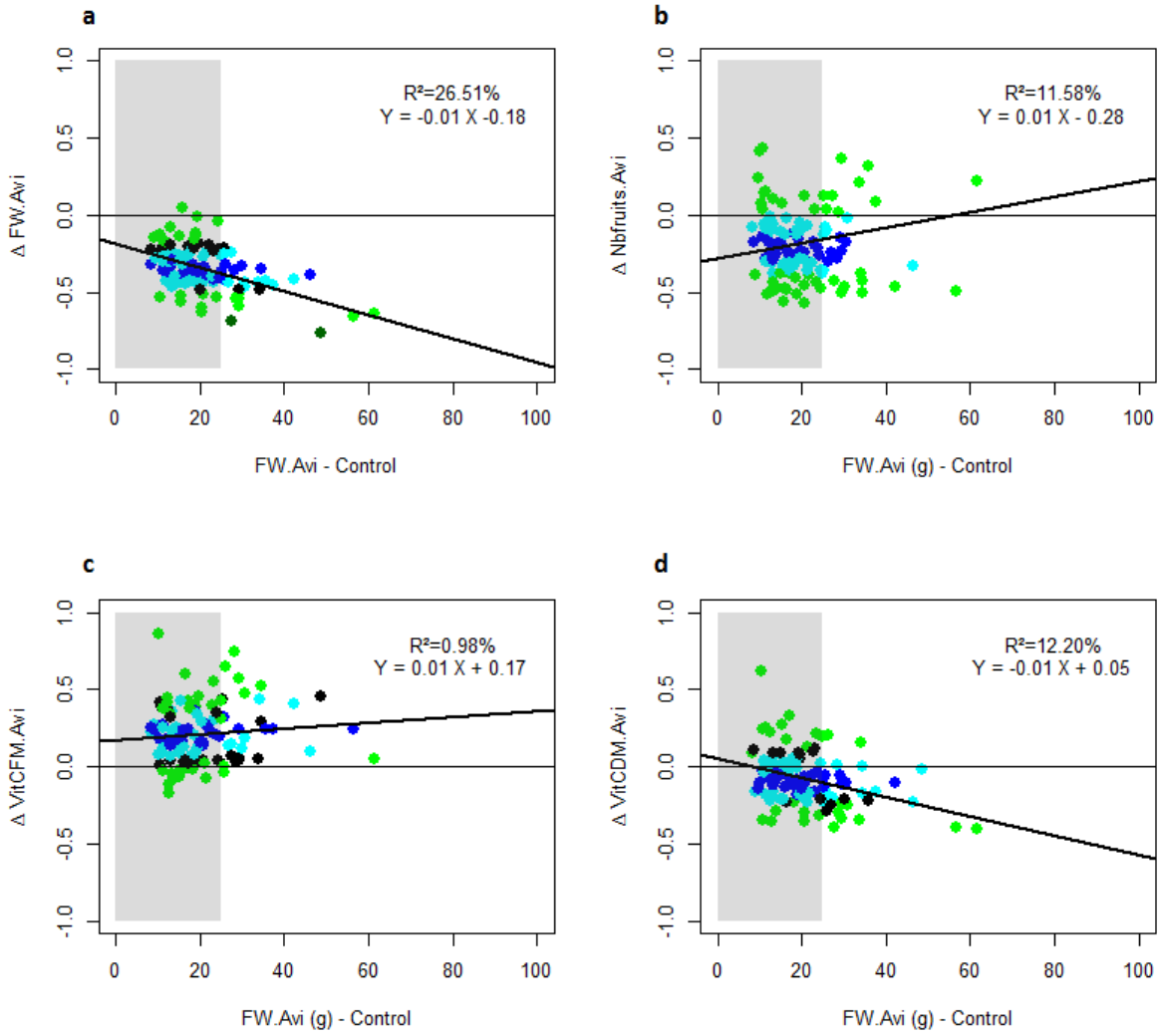


Fig 2 Linear relationships between FW in control condition and plasticity for (a) fresh weight, (b) fruit number, (c) vitamin C content relatively to fresh weight and (d) vitamin C content relatively to dry weight in Avignon. Equation and R^2 of the linear regression lines are displayed. Color indicates relative ecovalence classes: blue < 0.1; 0.1 < cyan < 0.5; 0 .5 < black < 1; green > 1. The gray areas indicate small fruit accessions (FW < 25 g).

matter (VitCFM) was not significant (*P-value* = 0.29 and r = 0.09), whereas this correlation was significantly negative (*P-value* = 1.05 x 10^{-03} and r = -0.35) when vitamin C plasticity was expressed relatively to dry matter (VitCDM) **(Fig 2c and d)**. Together, these results suggest important water losses in large fruits under drought, responsible for a FW decrease and concentration of vitamin C in fruit, without clear increase of vitamin C synthesis or accumulation in fruit.

Among small fruit accessions (FW below 25 g), ΔFW was comprised between -0.6 and 0.1 in Avignon and between -0.6 and 0.5 in Agadir **(Fig 2a and Supplemental Fig 3).** Plasticity for fruit number was comprised between -0.6 and 0.4 **(Fig 2b).** Thus, part of the small fruit accessions had positive delta value for FW.Avi (1), FW.Aga (8) and Nbfruits.Avi (12), indicating improved yield components under drought. Besides, in Agadir, two accessions showed, at the same time, a positive delta value for FW and fruit number (SSD168 and SSD172) **(Supplemental Table 1)**. Ecovalence measurement is a method developed by Wricke *et al*. (1964) to partition the sum of square of the interaction term and measure participation to individual genotype or environment to the genotype by environment interaction. In the small fruit accessions, the median values of the ecovalence distributions were 0.30, 0.18 and 0.25 for FW.Avi, FW.Aga and Nbfruits.Avi, respectively (blue and cyan colors in the gray area on Fig 2a and b; Supplemental Fig 3). These low values observed for 50% of the small fruit accessions indicated a relatively stable FW and fruit number for these genotypes, whatever the watering regime.

For all phenotypic traits, correlations between mean values under control and drought were highly significant (*P-value* < 0.01), correlation coefficients ranking from 0.26 (Yield.Avi) to 0.88 (Flw.Aga) **(Supplemental Table 5).** Nonetheless, we exhibited a significant change in correlation network between phenotypic traits according to the watering treatment, in the two experiments (*P-value* $_{Avi}$ = 4.0 x 10^{-4} and *P-value* $_{Aga}$ = 9.7 x10^{-3}) **(Fig 3, Supplemental Fig 4).** Part of the correlations observed under control treatment was reinforced under water deficit, as for the positive correlation between yield and fruit number ($r_{control}$ = 0.31 and $r_{drought}$ = 0.49) **(Fig 3)**. Others correlations were reduced under drought treatment. FW was slightly less positively correlated with yield under drought (0.28) than under control watering (0.47) **(Fig 3).** Together, these changes in correlation between FW and yield and between fruit number and yield suggested that fruit number was a major yield component under

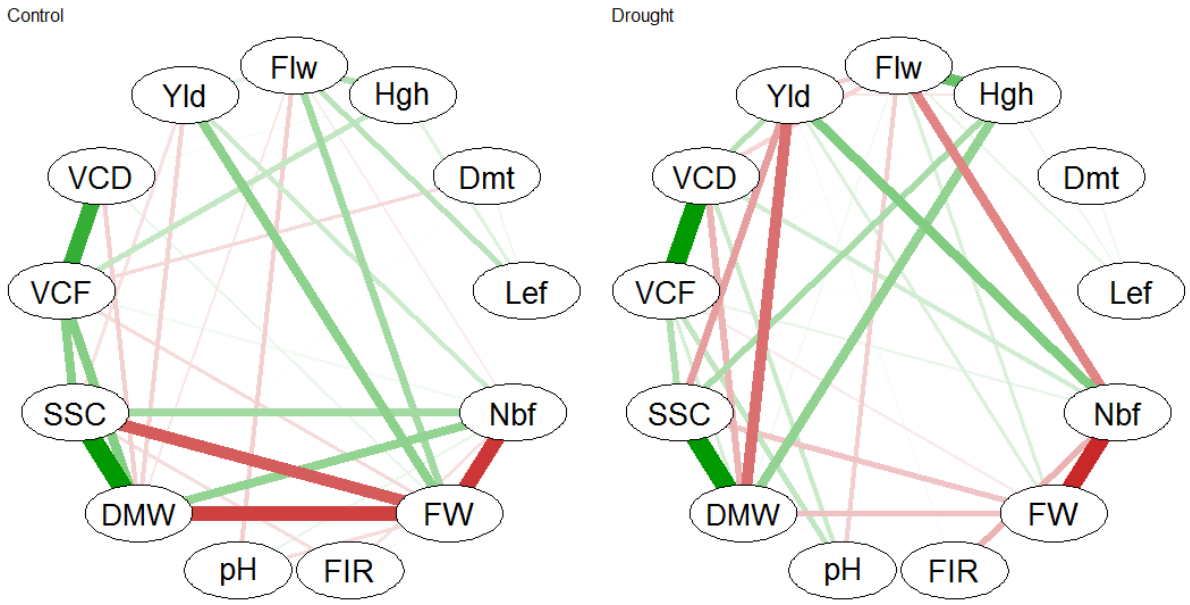


Fig 3 Changes in phenotypic correlation network between the two watering regimes. The figure displays Pearson correlation coefficients between average phenotypic values measured in control and drought treatment, in Avignon. Only coefficients higher than 0.2 are shown (*P-value* < 0.05). The line width is proportional to correlation coefficient value. The line color indicates direction of the correlation: green for positive correlations and red for negative correlations. Abbreviations meanings: 'Flw' for Flw.Avi ; 'Hgh' for Ht.Avi ; 'Dmt' for Diam.Avi ; 'Lef' for Leaf.Avi ; 'Nbf' for Nbfruits.Avi ; 'FW' for FW.Avi ; 'FIR' for FIR.Avi ; 'pH' for pH.Avi ; 'DMW' for DMW.Avi ; 'SSC' for SSC.Avi ; 'VCF' for VitCFM.Avi ; 'VCD' for VitCDM.Avi and 'Yld' for Yield.Avi.

drought. Only one situation of correlation reversal between watering regimes was observed, between flowering date and DMW **(Fig 3)**.

3. QTLs and QTL by watering regime interactions

A total of 56 QTLs were mapped and 44 of them colocalized within five clusters on chromosomes 2, 3, 4, 6 and 11 **(Table 2 and Fig 4).** The 56 QTLs explained more than 5% of the total phenotypic variance (PVE), with a median value of 14% and a maximum of 41% for FW.Avi in control treatment **(Supplemental Table 6, Supplemental Table 7)**. Eight QTLs were detected both in Avignon and Agadir experiments (33% of the QTLs detected on the six common traits between the two locations). The confidence intervals were smaller than 11 Mbp for 88% of the QTLs. Seven QTLs mapped around the centromeres encompassing more than 30 Mbp. Besides, the size of the confidence intervals in cM and in Mbp were poorly correlated (*P-value* = 0.01 and r = 0.38), due to differences in recombination rates along the genome. Pascual *et al*. (2015) and Sim *et al*. (2012) reported similar results in tomato and explained these findings by large genomic regions around the centromeres with roughly no recombination.

Twenty QTLs were detected only under the control conditions and 12 QTLs only under drought. Thirteen QTLs were constitutive as they were detected under both watering regimes. Distinguishing between specific and constitutive QTLs is not straightforward as it depends on the magnitude of the effect and the chosen detection threshold. Thus, we calculated the effects and PVE for all QTLs in the two watering regimes **(Supplemental Table 6, Supplemental Table 7)**. On chromosomes 2 and 11, constitutive QTLs for FW colocalized with FW QTLs previously fine mapped or cloned (Frary 2000; Lecomte *et al*. 2004; Huang and van der Knaap 2011; Illa-Berenguer *et al*. 2015) **(Fig 4).**

Eleven QTLs were significantly interactive between watering treatments, with two of them mapped both with the ANOVA testing procedure and with the plasticity data. The plasticity data gave more power to detect QTL by watering regime interaction, mapping ten interactive QTLs against three for the ANOVA testing procedure. One more interactive QTL, for yield on chromosome 8, was just below the threshold according to the ANOVA procedure (*P-value* = 0.005) and was not detected with the plasticity data. Among interactive QTLs detected, four were associated to plant traits and seven to fruit quality traits. Seven antagonist QTLs had opposite allelic effects when comparing both watering treatments and

Table 2 Characteristics of QTLs detected for plant and fruit traits. QTLs significant under both watering regimes are referred as 'constitutive'. QTLs significant under one watering regime only ('control' or 'drought') are designated as 'specific'. QTLs detected with the plasticity data and/or with a significant interaction in the ANOVA procedure are designated as 'interactive'. For each phenotypic trait and each QTL type, number of QTL, minimum and maximum confidence interval (CI in Mbp on genome assembly v2.5) and percentage of phenotypic variation (PVE) explained are displayed. For interactive QTL detected only with the ANOVA procedure, we were not able to asses CI.

Trait	Nb QTL total [a]	Constitutive			Specific						Interactive			
					Control			Drought						
		Nb QTL	Min - Max CI (Mbp)	Min - Max PVE	Nb QTL	Min - Max CI (Mbp)	Min - Max PVE	Nb QTL	Min - Max CI (Mbp)	Min - Max PVE	Nb Antagonist	Nb Differential	Min – Max CI (Mbp)	Min – Max PVE
Flw	6 (1)	1 (0)	34.28	0.13 - 0.19	1 (0)	5.98	0.12	3 (1)	2.83 - 33.49	0.12 - 0.13	0	1 (0)	3.23	0.20
Diam	7 (3)	1 (0)	0.76	0.14 - 0.22	1 (0)	2.56	0.17	3 (2)	0.79 - 37.00	0.13 - 0.29	2 (1)*	0	1.02 - 2.55	0.14 - 0.15
Leaf	4 (0)	0	0	0	4 (0)	0.84 - 2.14	0.11 - 0.19	0	0	0	0	0	0	0
Ht	2 (0)	0	0	0	1 (0)	8.45	0.13	1 (0)	2.81	0.14	0	0	0	0
Nbfruits	2	1	4.95	0.19 - 0.33	0	0	0	0	0	0	1	0	3.77	0.09
FW	10 (3)	4 (1)	0.33 - 6.93	0.11 - 0.41	3 (1)	3.28 - 5.14	0.11 - 0.15	1 (1)	1.34	0.12	0	2 (0)	11.09 - 45.71	0.12 - 0.19
FIR	4 (1)	2 (1)	2.54 - 2.90	0.15 - 0.21	1 (0)	0.87	0.14	1 (0)	1.48	0.15	0	0	0	0
pH	3	1	1.45	0.13 – 0.14	0	0	0	1	53.47	0.16	1	0	2.75	0.14
DMW	4	0	0	0	3	0.68 - 5.93	0.11 - 0 .18	0	0	0	1	0	NA	0.05 – 0.08
SSC	6	1	2.27	0.13 - 0.14	2	0.78 - 3.47	0.17 - 0.22	1	0.99	0.18	1*	1	2.24 - 6.75	0.11 - 0.24
VitCFM	2	1	1.95	0.17 - 0.20	0	0	0	1	53.47	0.15	0	0	0	0
VitCDM	3	1	1.25	0.13 – 0.14	1	4.18	0.17	0	0	0	1	0	4.18	0.17
Yield	3	0	0	0	3	0.68 - 6.35	0.13 - 0.18	0	0	0	0	0	0	0
Total	**56 (8)**	**13 (2)**			**20 (1)**			**12 (4)**			**7 (1)**	**4 (0)**		

[a] For traits measured in the two experiments, number of QTLs common to Avignon and Agadir experiments (same trait in the same watering regime) are indicated between brackets.

* Indication of interactive QTL confirmed with both methods, plasticity data and ANOVA procedure.

NA Not applicable.

four differential QTLs had effect intensity changed according to treatment **(Table 2 and Fig 4)**. In average, the interactive QTLs explained 14% of the phenotypic variance (sd = 5%). One of them, controlling variation in FW, was in the centromeric region of chromosome 11 and covered 46 Mbp. The ten others encompassed in average 4 Mpb (sd = 3 Mbp) and genomic regions carrying between 221 and 1009 genes **(Supplemental Table 8)**.

Among the interactive QTLs, the differential QTL for SSC on chromosome 2 mapped in the same genomic region as *ssc2.2*, a SSC QTL fine-mapped between two FW QTLs by Lecomte *et al*. (2004). The differential QTL for FW on chromosome 3 overlapped with *fw3.2* which was recently cloned (Chakrabarti *et al*. 2013) **(Fig 4).** This QTL was shown to control the increase in cell layers, the delay of fruit ripening and the decrease in fruit number as well. Besides, this QTL was reported to have a minor effect on fruit shape and to be sensitive to the growing environment (Zhang *et al*. 2012). In this same region at the extreme end of chromosome 3, we mapped also three antagonist QTLs for SSC, DMW and VitCDM. An antagonist QTL for pH mapped at 38 Mbp on chromosome 6 could be related to detected associations for organic acid content in an unrelated tomato population. These QTLs/associations are close from two putative malate transporters (Solyc06g072910 and Solyc06g072920) identified in a previous study (Sauvage *et al*. 2014). The differential QTL for FW on chromosome 11 mapped 2 Mbp ahead of the fine-mapped QTLs *fw11.2* and *fw11.3* (Huang and van der Knaap 2011; Illa-Berenguer *et al*. 2015). No interaction with the environmantal condition was demonstrated for these two QTLs until now.

Three examples of interactive QTLs are displayed in **Fig 5**. On chromosome 3, an interactive QTL had antagonist effect on DMW.Avi: in control treatment Cervil allele increased the trait value of 0.4 units whereas under drought Cervil allele reduced DMW of 0.3 units **(Fig 5a)**. On chromosome 4, an interactive QTL with antagonist effect on stem diameter was mapped: in control treatment Cervil allele decreased stem diameter of 0.7 units whereas under drought Cervil allele increased Diam.Avi of 0.1 units **(Fig 5b)**. On chromosome 2, an interactive QTL was detected with changes in effect intensity according to watering regime for flowering time **(Fig 5c)**. Cervil allele effect was increased (+1.2 units) under drought (meaning an earlier flowering).

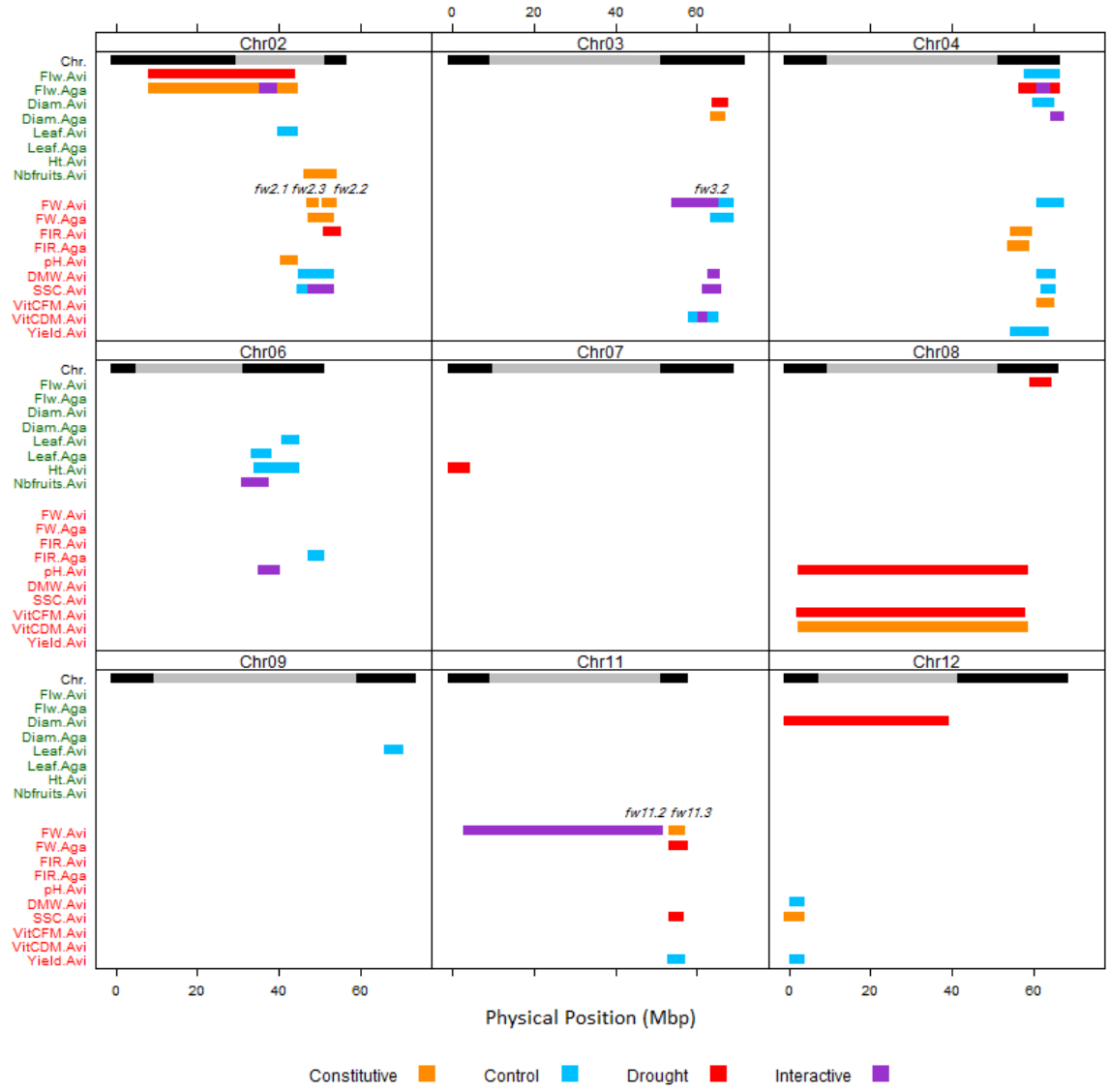


Fig 4 Overview of plant and fruit QTL identified on the tomato genome by QTL analysis in RILs. At the top of the panels, lines are representing tomato chromosomes where the lengths are proportional to chromosome physical sizes in million base pairs (Mbp). Centromeric regions with low recombination frequency are indicated in grey and peripheral parts in black (according to Sim *et al*., 2012). QTL are represented by square. Color codes correspond to the QTL types: constitutive (commons to control and drought treatment) in orange; detected only in control treatment in blue; detected only in drought treatment in red; interactive between the two watering regimes in purple. When an interactive QTL is colocalized with a non-interactive one, the interactive QTL is represented in the first plan in purple, surrounded by square with the color of the colocalized QTL. Positions of five major FW QTLs (: *fw2.1*, *fw2.3*, *fw2.2*, *fw3.2*, *fw11.2* and *fw11.3)* are indicated.

4. Genes differentially expressed and candidate gene identification under interactive QTLs

To go further in the understanding of the genetic control of tomato response to water deficit, gene expression was measured in young leaves of Cervil and Levovil, grown under the two watering regimes. Among the 33,913 tomato genes carried on the microarrays, 2,259 were differentially expressed between watering treatments in young leaves of Cervil and/or Levovil **(Fig 6 and Supplemental Table 3)**. More genes were differentially expressed in Levovil (1911), than in Cervil (786). Roughly identical counts of up (Cer: 43% and Lev: 44%) and down (Cer: 57% and Lev: 57%) regulated genes were observed in the two accessions. A total of 438 genes were differentially expressed in both parental accessions and 405 of them showed regulation in the same direction in Levovil and Cervil.

The enrichment analysis of GO terms related to biological process was achieved on the differentially expressed genes. The 405 genes differentially expressed in the same direction in Cervil and Levovil contained more genes associated to microtubule process, to lipid metabolism and response to wound stress than the proportions observed in the whole tomato genome. Among the genes differentially expressed only in Cervil, processes associated with the defense against biotic stress and cell-wall process were significantly overrepresented compared with the tomato genome **(Supplemental Fig 5a).** On the other hand, the list of genes differentially expressed only in Levovil was significantly enriched in genes related to cellular homeostasis, oxidation-reduction and metabolic process **(Supplemental Fig 5b)**. The differences between Levovil and Cervil in the enriched functions of the differentially expressed genes supported the differences observed between small (tolerance behavior) and large (avoidance behavior) fruit accessions at the phenotypic level. Finally, among the 33 genes regulated in different direction between Cervil and Levovil, three genes were associated to response to stress stimulus **(Supplemental Fig 5d)**: Solyc11g028060 (‘defensin-like protein’), Solyc06g009140 (‘Late embryogenesis abundant protein 3’) and Solyc07g006380 (‘defensin-like protein’). Interestingly, two of these genes were located in close vicinity of interactive QTLs **(Fig 4)**: Solyc06g009140 was located 2 Mbp above a differential QTL for fruit number on chromosome 6, whereas Solyc11g028060 was located in the interval of the differential QTL for FW on chromosome 11. They could be related to the phenotypic difference observed between Levovil and Cervil under drought and represent candidate genes for future studies,

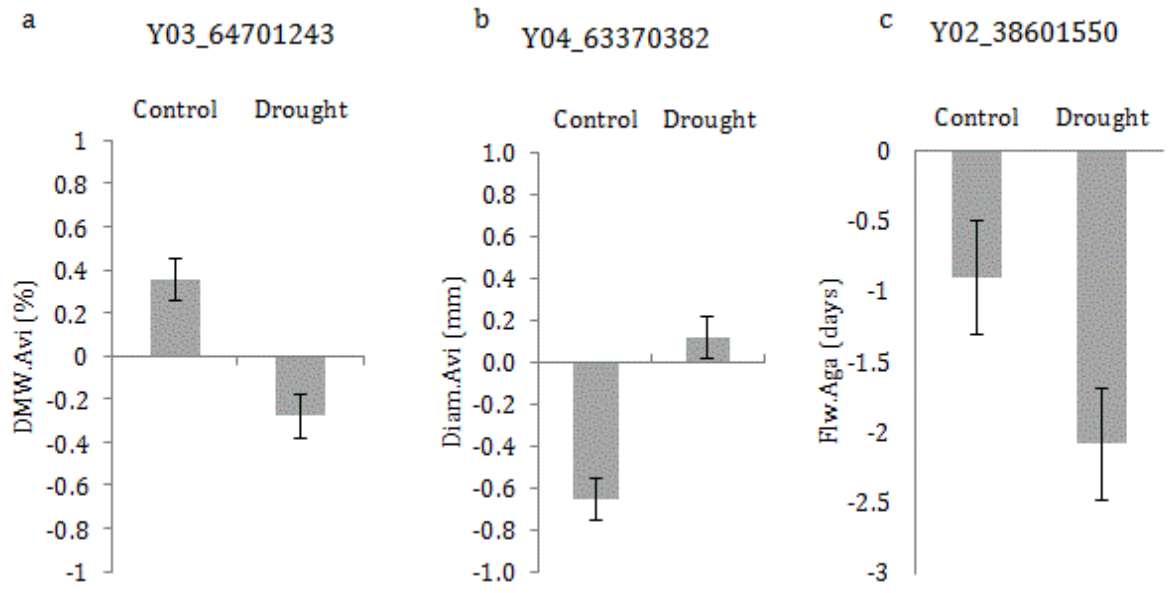


Fig 5 Examples of interactive QTL effects (a) 'antagonist' interactive QTL on chromosome 3 for fruit DMW measured in Avignon (marker at the LOD peak: Y03_64701243). (b) 'antagonist' interactive QTL on chromosome 4 for stem diameter measured in Avignon (marker at the LOD peak: Y04_63370382). (c) 'differential' interactive QTL on chromosome 2 for flowering time measured in Avignon (marker at the LOD peak: Y02_38601550).

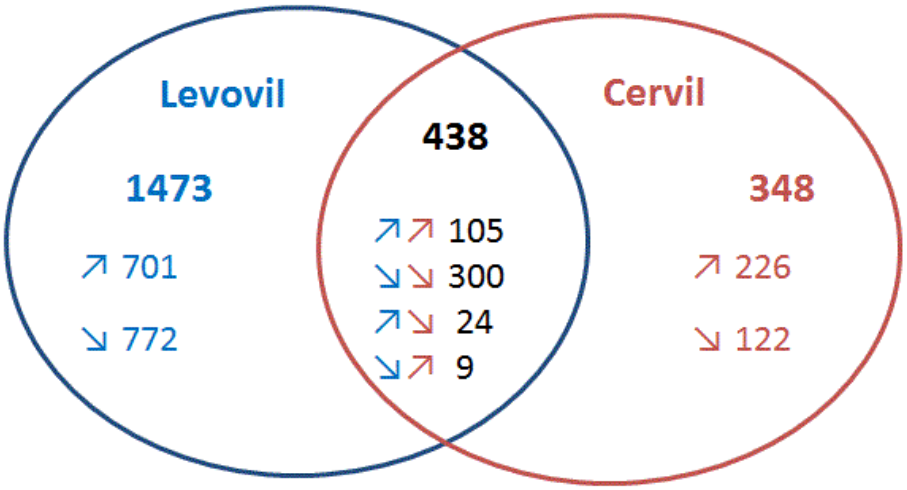


Fig 6 Venn diagram of differentially expressed genes between watering regimes in Cervil and/or Levovil. Genes were considered as being differentially expressed when the Bonferroni adjusted *P-value* was below 0.05. Blue and red colors indicate gene differentially expressed in Levovil and Cervil, respectively. Up and down arrows show genes up and down regulated under water deficit, respectively.

although these results should be taken cautiously as their expression was studied in leaves and related to fruit traits.

Then, we focused on two short genomic regions where interactive QTLs for stem diameter (chromosome 4, antagonist QTL, detected in Agadir and Avignon) **(Fig 5b)** and for flowering time (chromosome 2, differential QTL, Agadir) **(Fig 5c)** were mapped to look deeper at the differentially expressed genes **(Supplemental Table 6, Supplemental Table 7)**. The interactive QTL for Flw.Aga corresponded to a genomic region carrying 357 genes, whereas the QTL for stem diameter encompassed 289 genes (Tomato Genome Consortium 2012) **(Supplemental Table 8)**. Selecting the differentially expressed genes in Cervil and/or Levovil in these intervals reduced the candidate gene list to 24 and 29 genes, for Flw QTL and Diam QTL, respectively **(Table 3)**.

The re-sequencing of Cervil and Levovil genomes identified polymorphisms between these accessions and constituted a powerful tool to further reduce the differentially expressed gene lists (Causse *et al*. 2013). Under the interactive QTL for stem diameter, 24 genes differentially expressed were polymorphic between Cervil and Levovil. Among them, seventeen genes exhibited moderate effect polymorphism and two genes had a polymorphism with a high impact on the protein sequence: Solyc04g077640 coding for a '*serine carboxypeptidase 1*' (splice site donor) and Solyc04g079080 coding for a '*calmodulin*' (frame shift) **(Table 3)**. Among the 24 differentially expressed genes under the interactive QTL for Flw, eleven were polymorphic between Cervil and Levovil. Five genes presented moderate effect polymorphisms (non-synonymous variants in coding region) and only one had a polymorphism with a high impact on the protein sequence: Solyc02g069060 coding for a '*phloem lectin*' (loss of a stop codon) **(Table 3)**. These moderate to high effect polymorphisms and differentially expressed genes constitute putative candidates for the genetic control of tomato response to water deficit and have to be further investigated. Nevertheless, others polymorphic genes in these QTL intervals represent others putative candidates. Apart from the polymorphisms described above and in **Table 3**, we identified 17 genes with high effect variants under the Diam QTL (Solyc04g076410, Solyc04g076840, Solyc04g076940, Solyc04g077050, Solyc04g077330, Solyc04g077630, Solyc04g077700, Solyc04g077710, Solyc04g077920, Solyc04g078080, Solyc04g078180, Solyc04g078230, Solyc04g078260, Solyc04g078350, Solyc04g078360, Solyc04g078660 and Solyc04g078910)

Table 3 Genes underlying interactive QTL for flowering date and stem diameter differentially expressed between watering regimes in young leaves of parental genotypes. Only genes with t-test Bonferonni corrected *P-value* below 0.05 for Cervil and/or for Levovil are displayed. 'Gene ID', 'Start' and "End' indicate the gene identification code and its position on tomato genome assembly v2.5. 'ratio Cer' and 'ratio Lev' indicate the LOG ratio of the gene expression under drought relatively to control treatment, for Cervil and Levovil, respectively. The stars show the significance of the *P-value* for the LOG ratio test. The interaction type (Inter.) is specified: 'Ant.' when gene expression changes are in opposite directions for the two parents and 'Diff.' when expression changes are in the same direction. 'Gene polymorphism effects' shows the variants between Cervil and Levovil genetic sequence according to Causse *et al*. (2013). 'Modifier' indicates polymorphisms located in upstream and downstream regions or in UTR regions. 'Low' indicates synonymous polymorphisms. 'Moderate' indicates the non-synonymous variants in coding regions. 'High' indicates polymorphisms which modified splice sites or start/stop codons (loss or gain).

Gene				Gene expression			Gene polymorphism effects				
GENE_ID	start (Mbp)	end (Mbp)	Annotation	ratio Cer	ratio Lev	Inter.	Modifier	Low	Moderate	High	Total
Solyc02g065210	36.384	36.386	Cytochrome P450	0.41	1.37 **	-	1	0	0	0	1
Solyc02g065220	36.388	36.391	Cytochrome P450	0.64	1.47 ***	-	0	0	0	0	0
Solyc02g065240	36.416	36.419	Hydrolase alpha/beta fold family protein	-1.22	-1.24 *	-	3	0	0	0	3
Solyc02g065380	36.562	36.570	Cold acclimation protein COR413-like	-1.33	-1.56 ***	-	0	0	0	0	0
Solyc02g065470	36.621	36.621	Pathogenesis-related protein	-1.30	-1.25 *	-	0	0	0	0	0
Solyc02g067050	37.292	37.294	Uncharacterized ACR COG1678 family protein	0.05	1.85 ***	-	0	0	0	0	0
Solyc02g067100	37.330	37.334	Targeting protein for Xklp2	-1.62 **	-1.89 ***	Diff.	0	0	0	0	0
Solyc02g067310	37.494	37.496	Zinc finger-homeodomain protein 1	2.06 ***	0.95	-	0	0	0	0	0
Solyc02g067380	37.582	37.585	Transcription factor style2.1	-0.98	-2.03 ***	-	0	0	0	0	0
Solyc02g067750	37.862	37.865	Carbonic anhydrase	-6.46 ***	-1.53 ***	Diff.	0	0	0	0	0
Solyc02g067970	38.037	38.040	Tripartite motif-containing 25	-1.49 *	-1.15	-	0	0	0	0	0
Solyc02g068080	38.125	38.131	Voltage-gated chloride channel	0.68	1.86 ***	-	0	0	0	0	0
Solyc02g068170	38.196	38.197	Unknown Protein	0.63	1.73 ***	-	0	0	0	0	0
Solyc02g068200	38.218	38.219	TCP family transcription factor	0.04	1.24 *	-	1	0	0	0	1
Solyc02g068400	38.353	38.355	AT4G20050-like protein	-0.49	-1.29 **	-	0	0	0	0	0
Solyc02g068510	38.464	38.468	Zinc-binding protein	-0.42	1.30 **	-	4	0	0	0	4
Solyc02g068610	38.581	38.582	Genomic DNA chromosome 5 P1 clone MIK22	-0.23	-1.70 ***	-	0	0	0	0	0
Solyc02g068840	38.746	38.747	Unknown Protein	-0.89	-1.21 *	-	9	0	1	0	10
Solyc02g069060	39.033	39.035	Phloem lectin	2.77 ***	-0.45	-	22	0	0	1	23
Solyc02g069110	39.070	39.073	Cathepsin B	-0.04	-2.09 ***	-	44	0	2	0	46
Solyc02g069260	39.203	39.212	ARGONAUTE 1	-1.96 ***	-0.03	-	64	11	10	0	85
Solyc02g069490	39.357	39.360	FAD linked oxidase domain protein	-0.98	-2.83 ***	-	31	4	0	0	35
Solyc02g069560	39.407	39.412	Chloroplast unusual positioning 1A	-1.53 *	-0.94	-	21	3	1	0	25
Solyc02g069680	39.518	39.524	Charged multivesicular body protein 2a	-0.77	-1.56 ***	-	78	9	1	0	88

and 4 genes (Solyc02g065250, Solyc02g068970, Solyc02g069140 and Solyc02g069270) under the Flw QTL (polymorphism details in Causse *et al*. 2013).

Discussion

The aims of the study were (1) to outline genotype by watering regime interactions for 19 plant and fruit traits in 119 RILs from a cross between a small fruit accession and a large fruit accession, grown under two watering regimes, (2) to elucidate the inheritance patterns of these interactions and (3) to identify candidate genes as putative targets for tomato breeding under deficit irrigation. The results provided a basis for improving the use of deficit irrigation strategies for tomato production.

1. Genotype by watering regime interaction at the phenotypic level

At the phenotypic level, we identified significant genotype by watering regime interactions for most of the traits evaluated (except pH). The importance of the interaction with respect to watering and genotype factors may depend on the phenotypic traits and the genotypes studied, as well as the plant developmental stage and the intensity/duration of the water deficit suffered by the plants. In our study, the interactions, although marginal in regard to the magnitude of the effect of the genotype factor (24 to 86%), represented up to 17% of the total sum of squared deviations. They were essentially due to genotype re-ranking (85 to 100%), and poorly to heterogeneous variance between watering conditions. In a previous publication on leaf water content in six cultivated tomato accessions grown under two watering regimes, the authors reported a significant genotype by watering treatment interaction representing 16% of the total sum of square, against 5% and 72% for the genotype and watering factor, respectively (Jureková *et al*. 2011). On the other hand, in thirty wheat lines phenotyped for traits measuring vigor of seeds and seedlings under two watering conditions, genotype by watering interactions represented between 8 and 39% of the total sum of square and their weight was always equal or lower than the weight of the genotype factor (Dhanda *et al*. 2004).

Heritabilities and genetic variance were highly correlated between the two watering treatments, in the two locations. These results are contrasting with the only detailed study of genotype by watering regime interaction in tomato reported by Gur *et al*. (2011) in a set of introgression lines (ILs) derived from a cross between *S. pennellii* and *S. lycopersicum.* The

Following Table 3.

Diam 4	Solyc04g076310	61.227	61.241	Kinesin-like protein	-1.84 ***	-1.56 ***	Diff.	60	4	2	0	66
	Solyc04g076870	61.796	61.801	Glutamyl-tRNA reductase	-1.56 **	0.12	-	0	0	0	0	0
Detected on	Solyc04g076880	61.815	61.823	Phosphoenolpyruvate carboxykinase	-0.46	1.53 ***	-	66	5	0	0	71
Δ and	Solyc04g076950	61.892	61.895	Multidrug resistance protein mdtK	0.43	1.24 *	-	24	1	1	0	26
average data												
	Solyc04g076990	61.947	61.950	Receptor like kinase, RLK	-0.51	1.49 ***	-	21	4	1	0	26
	Solyc04g077010	61.969	61.974	Receptor like kinase, RLK	0.82	1.29 **	-	29	2	2	0	33
	Solyc04g077030	61.993	61.999	Xylulose kinase	0.08	1.41 ***	-	0	0	0	0	0
	Solyc04g077150	62.086	62.090	Erg28 like protein expressed	-0.95	-1.58 ***	-	49	0	0	0	49
	Solyc04g077210	62.134	62.144	Knotted-like homeobox protein	0.03	1.79 ***	-	70	1	1	0	72
	Solyc04g077270	62.242	62.247	Serine/threonine kinase receptor	-0.78	-1.29 **	-	0	0	0	0	0
	Solyc04g077430	62.339	62.355	Squalene monooxygenase	-0.42	-1.22 *	-	57	3	2	0	62
	Solyc04g077440	62.351	62.357	Squalene monooxygenase	0.10	-1.87 ***	-	23	1	1	0	25
	Solyc04g077530	62.490	62.490	Unknown Protein	0.75	2.66 ***	-	37	0	1	0	38
	Solyc04g077540	62.496	62.507	Kinesin-like protein	-1.76 ***	-1.53 ***	Diff.	0	0	0	0	0
	Solyc04g077640	62.582	62.585	Serine carboxypeptidase 1	1.49 *	0.46	-	48	1	1	1	51
	Solyc04g077650	62.588	62.594	Serine carboxypeptidase 1	-0.05	1.57 ***	-	29	2	1	0	32
	Solyc04g077690	62.625	62.627	Unknown Protein	0.48	-1.44 ***	-	28	0	2	0	30
	Solyc04g077810	62.701	62.702	Unknown Protein	-0.89	1.25 *	-	29	0	0	0	29
	Solyc04g077990	62.859	62.860	LOB domain protein 38	-0.92	1.76 ***	-	35	1	0	0	36
	Solyc04g078010	62.881	62.883	Unknown Protein	-0.78	-1.49 ***	-	53	1	1	0	55
	Solyc04g078090	62.935	62.942	Acyl-CoA-binding domain-containing protein 6	0.49	1.24 *	-	93	4	6	0	103
	Solyc04g078310	63.067	63.069	Cyclin A-like protein	-0.78	-1.82 ***	-	22	0	3	0	25
	Solyc04g078460	63.184	63.186	N(4)-(Beta-N-acetylglucosaminyl)-L-asparaginase	-0.90	-4.02 ***	-	20	3	0	0	23
	Solyc04g078470	63.202	63.205	Cyclin D3-1	-0.76	-2.27 ***	-	39	1	6	0	46
	Solyc04g078540	63.242	63.247	Cathepsin B-like cysteine proteinase	-0.68	-1.22 *	-	0	0	0	0	0
	Solyc04g078590	63.274	63.278	Receptor like kinase, RLK	-1.23	-2.37 ***	-	22	2	5	0	29
	Solyc04g078840	63.500	63.505	BZIP transcription factor	-1.53 *	0.66	-	25	0	1	0	26
	Solyc04g078900	63.574	63.578	Cytochrome P450	1.90 ***	-0.07	-	44	0	0	0	44
	Solyc04g079080	63.670	63.675	Calmodulin	-0.04	-1.56 ***	-	38	1	0	1	40

*** shows *P-value* below 0.001, ** between 0.001 and 0.01; and * between 0.01 and 0.05.

authors described decreased genetic variances and heritabilities under drought. The discrepancy with our results could be caused by the different genetic basis of the populations, a lower number of tested lines and/or a more drastic drought treatment (no water supply) in Gur *et al*. (2011). In rice, intermediate results on 151 lines were reported, with a conserved heritability under arid conditions for some traits (grain and biomass yield), decreased (plant height, 1000-grain-weight) or increased (harvest index) for others, with differences according to water deficit intensity (Babu *et al*. 2003).

The pattern of genotype by environment interaction and the level of heritability under stress conditions are important features to consider when choosing a breeding strategy. When correlations are imperfect between environmental conditions (revealed through re-ranking of genotypes) and if there is genetic variability under stress, selecting one genotype for a specific environmental condition seems to be the best strategy. However, such breeding strategy is limited by the variability across years in the intensity and frequency of the drought episodes. Alternative approaches could be to improve drought adaptation in elite varieties by incorporating morphological and physiological mechanisms maintaining genotype performances under drought or to improve yield potential in already drought adapted accessions (Mitra 2001).

2. Interaction between genotype and the watering regime at the genotypic level

For the first time, in a cross between two cultivated tomato accessions, we identified QTL by watering regime interaction, deciphering the genetic architecture of tomato response to water deficit. Such an approach for dissecting G x E interaction into underlying genetic loci is not new and was already performed in numerous plant species. In the first QTL by environment studies, authors performed independent QTL mapping in each environment and compared the QTLs obtained through the experiments (Paterson *et al*. 1991 in tomato; Jansen *et al*. 1995 in *Arabidopsis thaliana*; Lu *et al*. 1997 in rice). Today, more complex mapping strategies are undertaken and allow a more refined understanding of QTL x E interactions. Models can test for the presence of QTL whose effect vary between environments or plasticity QTLs can be mapped using composite traits measuring genotypic response to environmental constraints (van Eeuwijk *et al*. 2010; Des Marais *et al*. 2013; Li *et al*. 2015).

Applying these complementary strategies, we identified 56 QTLs with moderate (10%) to high (40%) percentage of phenotypic variance explained. Among them, 13 were constitutive, 20 were control-specific, 12 were drought-specific and 11 were interactive between watering regimes. Excepted the seven QTLs encompassing centromeric regions with low recombination frequency (Sim *et al*. 2012; Pascual *et al*. 2015), we mapped QTLs with relatively small confidence intervals covering 0.33 to 11 Mbp. The others reported QTL x watering regime studies in tomato used introgression lines and reported QTLs with large confidence intervals covering up to an entire arm of chromosome (Gur *et al*. 2011; Easlon *et al*. 2014). The plasticity data gave more power to detect QTL by watering regime interaction, mapping ten interactive QTLs against three for the ANOVA testing procedure (with two common between methods). Tétard-Jones *et al*. (2011) and El-Soda *et al*. (2014a) obtained such proportion when comparing both mapping methods and proposed that mapping QTL using directly phenotypic difference or ratio could give additional statistical power by exacerbating contrasts between two environmental conditions. Alternatively, when more than two environments are compared, parameters from reaction norms can be used as plasticity variable.

Although it is difficult to make a precise comparison because the authors did not exactly test QTL x E interaction, the relative proportion of different QTL types was relatively similar in the study of Gur *et al*. (2011) on tomato introgression lines. In this latter study on tomato yield and quality traits, a majority of the QTLs was constitutive (45%) and control-specific (39%) and few drought-specific QTLs (16%) were mapped. Part of our constitutive QTLs confirmed the constitutive loci identified by these authors, in particular for the FW QTLs located on chromosomes 2 and 11 **(Fig 4)**. These QTLs were colocalized with cloned or fine-mapped genes controlling tomato FW: *fw2.1* and *fw2.3* (fine-mapped in Lecomte *et al*. 2004), fw2.2 (cloned in Frary 2000), *fw11.2* and *fw11.3* (fine-mapped in Huang and van der Knaap 2011; Illa-Berenguer *et al*. 2015). Their constitutive feature is promising in regard to tomato yield improvement for diverse environments. From these results, Gur *et al*. (2011) proposed that tomato yield under drought conditions would be mostly controlled by QTLs determining the productivity of the plant, rather than QTLs providing a physiological improvement for drought tolerance. However, we moderate this hypothesis arguing that distinguishing between specific and constitutive QTLs is not straightforward as it depends on the

magnitude of the effect and the chosen detection threshold. We agree with Des Marais *et al*. (2013) who encourage authors to display effects and PVE for all QTLs in the different environmental conditions to make fair comparison between studies. Besides, quantification of QTL × E interactions appears necessary to be able to detect QTLs with varying effects depending on the environmental conditions and to identify QTLs/genes potentially involved in tolerance mechanisms against abiotic stress.

On average, the eleven interactive QTLs detected in our study explained 14% of the phenotypic variance for the respective traits. Among them, a majority had antagonistic effects according to watering regime (7 QTLs) and four QTLs showed a decreased effect under drought. To date, only four studies have reported QTLs with changing direction of allelic effects according to the environmental constraints (Des Marais *et al*. 2012). Such antagonist QTLs were described in rice in response to planting density (Liu *et al*. 2012) and between different water regimes in *Arabidopsis thaliana* (Hausmann *et al*. 2005; El-Soda *et al*. 2014a) and sorghum (Sabadin *et al*. 2012). In our study, the detection of loci with antagonist effect could result from a drastic drought treatment and testing different levels of watering in a next QTL study may give different results. Negative correlation between water deficit intensity and fruit fresh weight decrease was already observed in tomato (Durán Zuazo *et al*. 2011), but the effect of different water reduction intensity on the genetic determinants of tomato plant response to drought has been poorly investigated. In *Arabidopsis thaliana*, the expression patterns of 6,180 genes that were differentially expressed under severe drought was not significantly changed under moderate water deficit (Harb *et al*. 2010). Besides, 18 *Arabidopsis thaliana* mutants that behaved better under severe water stress did not present any superiority under moderate stress (Skirycz *et al*. 2011). Such results are in line with a different genetic determinism depending on the level of water deficit suffered by the plant. Skirycz *et al*. (2011) proposed that severe drought may be favorable for detecting QTLs relative to limitation of water depletion in plant tissues (referred as avoidance strategy) whereas mild stress would allow to identify loci responsible for maintening growth, photosynthesis and metabolism during water deficit (tolerance strategy).

Knowledge of the interactive QTLs is crucial in breeding programs because the presence of such QTLs can limit breeding efficiency if the favorable alleles do not have the same effect

under different environmental conditions. We identified antagonist QTLs for DMW, SSC, stem diameter, fruit number and pH on chromosomes 3, 4 and 6. Luckily, tomato and others horticultural crops are widely grown under irrigation and the level of water deficit imposed to the plants can be controlled and used to manage fruit quality, yield and water consumption. Furthermore, the antagonist QTLs identified could be used in marker-assisted selection (MAS) to build either genotypes for growth specifically under deficit irrigation or genotypes to cultivate under well-watered conditions. However, specific consideration should be given to the overlap of interactive QTLs at the terminal end of chromosome 3. In this region, we mapped one differential QTL for FW and three antagonist QTLs for DMW, SSC and VitCDM. Under control, Levovil alleles increased FW and VitCDM, whereas it decreased SSC and DMW. In drought conditions, the Levovil alleles increased FW in a lower extend, but decreased VitCDM and increased SSC and DMW. This cluster of QTL colocalized with the cloned QTL *fw3.2* which was shown to be environmentally sensitive and to have pleiotropic effects on fruit cell numbers, fruit shape, ripening date and fruit number (Zhang *et al*. 2012; Chakrabarti *et al*. 2013). Besides, previous mapping studies in tomato grown under well watered condition identified QTLs for sugar content and titrable acidity in the same genomic region (Saliba-Colombani *et al*. 2001; Pascual *et al*. 2015). It is difficult to determine if the interactive QTLs correspond to an unique QTL with pleiotropic effect on the different traits whatever the watering regime or to different QTLs controlling the different traits under the two watering treatments. If there are several QTLs, recombination events could help to gather the most favorable alleles (or less unfavorable) to build a genotype suitable for growing under deficit irrigation.

3. Could we stand on water management to improve tomato fruit quality?

Deficit irrigation and partial root drying are watering strategies under consideration in fruit crops, aiming to reduce non-beneficial water consumption in horticultural production while maintaining the economic feasibility of the cropping systems (Kirda *et al*. 2004; Cui *et al*. 2008; Zheng *et al*. 2013). Major fruit species are highly water demanding crops often cultivated with abundant irrigation, which can lead to overuse of groundwater and environmental degradations. Beyond the concerns of gaining in water productivity and controlling yield losses, these practices may also contribute to improve fruit flavour and nutritional quality. In tomato, flavour perception is an important criteria for genetic

improvement since consumers started complaining about lack of taste and aroma in the new long shelf life varieties (Kader *et al*. 1977; Bruhn *et al*. 1991; Ratanachinakorn *et al*. 1997). Tomato flavour results from complex interactions between sugars, organic acids and tens of volatile aromas (Stevens 1972; Yilmaz 2001). Besides, an abundant scientific literature have reported the favorable effects of tomato consumption on human health (Giovannucci 1999; Khachik *et al*. 2002; Giovannucci 2002), in particular through its content in ascorbic acid (vitamin C) and carotenoids which are among the most effective antioxidant in plants (Smirnoff 1996). Sugars, acids and antioxidants are also involved in plant response against stressing factors and their contents in fruits may be improved by the application of abiotic stress (in particular drought; Lester 2006; Dorais *et al*. 2008).

Due to the application of a 60% water deficit, we observed reduced plant vigor (stem diameter, leaf length and plant height) and productivity (fruit number and FW). On average, in the RIL population, the yield decrease reached the substantial value of 50%, hardly compatible with a sustainable production system. However, at the same time, tomato fruit soluble solid and dry matter content were increased by 26% and 31% on average, respectively. These results confirmed decreased yield and improved fruit quality previously reported in many fruit crop species cultivated under water limitation, among which peach tree (Mirás-Avalos *et al*. 2013), grape wine (Santesteban and Royo 2006) and tomato (Guichard *et al*. 2005; Zheng *et al*. 2013; Ripoll *et al*. 2014). In accordance with others studies (Zairi *et al*. 2003; Kirda *et al*. 2004), we support that a limited water deficit could be the best compromise between crop yield and fruit quality. Response of large genotypic sets to low and mild water stress should be studied to assess the optimal stress level.

Concerning vitamin C content in fruit, results were different depending if concentration were reported on the basis of fresh weight or dry weight. When expressed relatively to fresh weight, vitamin C content was increased by 26% in average in the RIL population, whereas when expressed relatively to dry weight it was decreased by 9%. In their review, Ripoll *et al*. (2014) pointed out such discrepancy for acid and sugar content in many fruits. Regarding how metabolite contents are measured, it is thus difficult to define if an increase in concentration results from a dilution/concentration effect, solute accumulation or synthesis in the fruit. Recent studies in tomato showed increased concentration of vitamin C with various extend depending on genotypes and water stress level, but the contents were

expressed relatively to fruit fresh weight only (Favati *et al*. 2009; Murshed *et al*. 2013). More refined study of various metabolite contents in tomato fruits grown under drought should be achieved to clearly state on the effect of water limitation on the nutritional value of tomato.

The highlight of our study is that the response to water stress depends on fruit size leading to different water management strategies. Large fruits suffered from a greater loss of FW due to the water deficit, but they were also those whose dry matter and soluble solid contents most increased. Growing large fruit tomatoes under a slight water deficit could improve tomato flavour and limit water consumption. To compensate the yield loss for growers, the fruits of plants grown with a slight water deficit could be marketed as 'tasty and environment friendly food' (Dorais *et al*. 2008). On the other hand, in small fruit accessions, we observed a large genotypic variability for fresh weight and fruit number plasticity under drought. In particular, a dozen lines showed stable or increased FW and/or fruit number under drought. They represent good candidates for tomato yield improvement under drought. Genotypic variability in small fruit lines could be further explored to limit water consumption in cherry tomato greenhouses. In the long term, interesting 'plasticity alleles' identified in small fruit genotypes could be introgressed in large fruit plants provided they have no pleiotropic effect on fruit size.

4. Studying G x E interactions in the omic era

Genetic and genomic methods are now available to accelerate the identification of candidate genes and polymorphisms, through fine-mapping and functional genomic studies. It gives opportunity to further decipher the genetic basis of phenotypic plasticity. Here, we combined QTLs and genomic approaches to gain knowledge on the genetic architecture of tomato response to water deficit. The re-sequencing of the RIL parents identified thousands of SNPs to build a new genetic map covering fully the tomato genome (Causse *et al*. 2013; Pascual *et al*. 2015). The availability of the reference genome sequence and its functional annotation allowed the projection of the QTL confidence intervals onto the tomato physical map, identifying hundreds of genes located in these intervals (Tomato Genome Consortium, 2012). Then, we focused on two interactive QTLs potentially controlled at leaf level and with short confidence intervals: an antagonist QTL for stem diameter on chromosome 4 and a differential QTL for flowering time on chromosome 2. Combining the results of QTL mapping

to the analysis of gene expression of the RIL parents drastically reduced the list of candidate genes under these two interactive QTLs by targeting differentially expressed genes between watering regimes. The catalog of polymorphisms between the parental accessions gained from their re-sequencing reduced further the putative gene list. Under the two QTLs, three promising candidate genes with differential expression and high impact polymorphisms were identified: Solyc02g069060 was coding for a '*phloem lectin*' (loss of a stop codon), Solyc04g077640 for a '*serine carboxypeptidase 1*' (splice site donor) and Solyc04g079080 for a '*calmodulin*' (frame shift) **(Table 3)**. The involvement of genes of these three protein families in plant protection and/or signaling in response to biotic and abiotic stress was reported in many studies and strengthened the interest paid to these three loci. Lectins constitute a class of carbohydrate-binding proteins with a known role in plant protection against cold, drought, salinity and biotic stress. They seem to be involved in cellular regulation and signaling in many plants (Van Damme *et al*. 2004; Jiang *et al*. 2010). Calmodulin is involved in plant response to abiotic stress through the involvement in osmotic adjustment and stress signaling in interaction with cellular calcium (Gong *et al*. 1997; Perruc *et al*. 2004; Reddy *et al*. 2011). A serine carboxypetidase type 1 was identified for its role in response to wound stress in tomato (Moura *et al*. 2001) and a rice serine carboxypeptidase-like gene was shown to be involved in response to biotic and oxidative stress (Liu *et al*. 2008). However, numerous serine carboxypeptidases genes are present in tomato genome and more precise study of these specific genes should be done.

On the other hand, others polymorphic genes in the QTL intervals remain equally good candidates as differential expression is not always a requisite and differences in stability and activity of proteins may not be related to differences in mRNA production. Besides, polymorphisms in non-coding regions may also affect gene expression or protein stability. Furthermore, we identified 17 and 4 genes not differentially expressed but with high effect variants under the stem diameter and the flowering time QTL, respectively. In particular, one of the polymorphic genes under the interactive QTL for flowering time (Solyc02g069270 coding for '*SlAGO2b*' – frame shift) belonged to the argonaute (AGO) gene family known to be involved in RNA silencing pathways and interaction with microRNAs in plants (Vaucheret 2008). This gene is located in a region of chromosome 2 including two others AGO genes (SIAGO2a and SIAGO3), the first (Solyc02g069260 in **Table 3**) presenting moderate effect

polymorphisms and being differentially expressed depending on the watering regime in our experiments. In a previous study on tomato, this set of genes was shown to be up-regulated under tomato yellow leaf curl virus (TYLCV) infection and under several abiotic stress, including drought stress (Bai *et al*. 2012). The argonaute genes may play an important role in reproductive development of tomato plant subjected to biotic and abiotic stressors by involving miRNA, in line with several results obtained in *Arabidopsis thaliana* (Vaucheret 2004; Lee *et al*. 2010; Westwood *et al*. 2013).

A pioneer publication on *Arabidopsis thaliana* provided another more advanced strategy to use gene expression data as clues for the identification of candidate genes involved in plant response to drought (Lovell *et al*. 2015). The authors proposed to use gene expression as covariate in a QTL model to link markers, RNA expression and phenotypes. They selected genes with significant cis-eQTLs and tested the effect of their transcript abundance on the effect of QTLs for phenotypic plasticity traits. Efficiency of the method was proved by recovering the causal locus FRIGIDA (previously cloned by Lovell *et al*. 2013) among 92 cis-regulated genes in the confidence interval of a QTL for water use efficiency (WUE). However, a limited number of studies have explored the genetic variation in transcriptome response to environmental constraints in large populations (Des Marais *et al*. 2013). The development of new sequencing technologies at a reduced cost may help to improve the quality of genome sequences, produce marker and gene expression datasets and allow the automatic functional annotation in many crops, which is an essential condition to implement strategies combining gene expression and QTL analysis in future research.

Genome wide association studies (GWAs) are another possible framework for the dissection of the genetic basis of phenotypic plasticity in diverse plant populations. GWA approaches benefit from the many recombination events experienced in natural populations to achieve QTL mapping leading to a few candidate genes or even to the identification of the causal polymorphisms. Besides, GWAs access to a larger genetic diversity, with more alleles than in a bi-parental cross. GWA models testing G x E interactions have been developed by plant biologists (Korte *et al*. 2012; Saïdou *et al*. 2014), but the combination of GWAs and gene expression was essentially reported in human genetics until now (Cheung *et al*. 2005; Cookson *et al*. 2009) and in few studies on *Arabidopsis thaliana* (Nicolae *et al*. 2010; Chan *et al*. 2011). It is probably because GWA models may be less powerful than bi-parental

population to identify QTLs when they have small effects, which is the case for most of the QTLs involved in phenotypic plasticity (Des Marais *et al*. 2012). This limitation can be bypassed using large populations designed to balance the allele frequencies. Combining QTL mapping, GWAs in large population and gene expression will constitute a complete framework to obtain a fine picture of the genetic control of genotype by watering regime interactions in plant. The advantage of a dual mapping strategy would be to reduce the rate of false positives and detect false negatives suffered by GWA due to structuration of the mapping panel, while taking advantage of the great allelic diversity in diverse populations (Brachi *et al*. 2010; Sterken *et al*. 2012).

Conclusion

This work is the first QTL study of response to water deficit in cultivated tomato (*S. lycopersicum*). At the phenotypic levels, significant genotype by watering regime interactions were reported. Large fruit tomatoes were more sensitive to drought and will require specific breeding considerations for growing under deficit irrigation, to achieve a trade-off between fruit quality improvement and yield. At the genotypic level, we identified interactive QTLs, many exhibiting effects changing direction depending on the watering regime. In the scope of plant breeding program, these QTLs could be used in marker-assisted selection (MAS) to develop tomato genotypes adapted to water limitation with intent to limit overuse of groundwater. In regard to genetic developments, we demonstrated a convenient way of combining QTL data, gene expression analysis and polymorphism data to identify candidate genes for plant adaptation to drought in the high-throughput area. Further studies need to confirm their roles. Besides, organ specific transcriptome analysis will be of main interest to reveal the regulation network of tomato response to water deficit in a more refined scale.

Author contributions

E.A. conducted experiments in France, analyzed data and wrote the manuscript. J.G. and N.B sampled and collected phenotypic data in France. J.B. sampled and collected phenotypic data in Morocco. S.P., J.P.T. and F.B. performed microarray experiments and gene expression quantification. M.C. supervised the project, built the experimental design and revised the manuscript. All authors discussed the results and commented the manuscript. Authors declared no conflict of interest in the authorship and publication of this document.

Acknowledgements

We acknowledge the experimental teams of UR GAFL and Gautier SEMENCES for their help in experimentation. We thank especially Yolande Carretero, Alain Goujon, Esther Pelpoir, Renaud Duboscq, Claire Duffes and the employees of “Domaine Margau” (Agadir) for their help in phenotyping plants. Thanks to Christopher Sauvage for his proofreading. The CTPS project TOMSEC supported this work. E.A. was supported by an INRA PhD fellowship.

Supplemental Material

Supplemental figures 1 to 5 and supplemental tables 4 to 8 are available at the end of the manuscript in Appendix 4. Supplemental table 1 to 3 display phenotypic and genotypic data and are available online on the editor website.

References

Anderson CM, Kohorn BD (2001) Inactivation of *Arabidopsis* SIP1 leads to reduced levels of sugars and drought tolerance. J Plant Physiol 158:1215–1219. doi: 10.1078/S0176-1617(04)70149-2

Ashraf M (2010) Inducing drought tolerance in plants: Recent advances. Biotechnol Adv 28:169–183. doi: 10.1016/j.biotechadv.2009.11.005

Asins MJ, Raga V, Roca D, *et al* (2015) Genetic dissection of tomato rootstock effects on scion traits under moderate salinity. Theor Appl Genet 128:667–79. doi: 10.1007/s00122-015-2462-8

Babu RC, Nguyen BD, Chamarerk V, *et al* (2003) Genetic analysis of drought resistance in rice by molecular markers. Crop Sci 43:1457 – 1469. doi: 10.2135/cropsci2003.1457

Bai M, Yang G-S, Chen W-T, *et al* (2012) Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum.* Gene 501:52–62. doi: 10.1016/j.gene.2012.02.009

Barnabas B, Jager K, Feher A (2007) The effect of drought and heat stress on reproductive processes in cereals. Plant Cell Environ 31:11 – 38. doi: 10.1111/j.1365-3040.2007.01727.x

Bertin N, Guichard S, Leonard C, *et al* (2000) Seasonal evolution of the quality of fresh glasshouse tomatoes under Mediterranean conditions, as affected by air vapour pressure deficit and plant fruit load. Ann Bot 85:741–750. doi: 10.1006/anbo.2000.1123

Blum A (2011) Plant water relations, plant stress and plant production. Plant Breeding for Water-Limited Environments. Springer New York, New York, pp 11 – 52

Box G, Cox D (1964) An analysis of transformations. J R Stat Soc Ser B 26:211 – 252.

B**oyle EI, Weng S, Gollub J, *et al* (2004)** GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics 20:3710–3715. doi: 10.1093/bioinformatics/bth456

Brachi B, Faure N, Horton M, *et al* (2010) Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS Genet 6:e1000940. doi: 10.1371/journal.pgen.1000940

Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890. doi: 10.1093/bioinformatics/btg112

Bruhn CM, Feldman N, Garlitz C, *et al* (1991) Consumer perceptions of quality: apricots, cantaloupes, peaches, pears,strawberries and tomatoes. J Food Qual 14:187–195. doi: 10.1111/j.1745-4557.1991.tb00060.x

Causse M, Desplat N, Pascual L, *et al* (2013) Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. BMC Genomics 14:791. doi: 10.1186/1471-2164-14-791

Causse M, Saliba-Colombani V, Lesschaeve I, Buret M (2001) Genetic analysis of organoleptic quality in fresh market tomato. 2. Mapping QTLs for sensory attributes. TAG Theor Appl Genet 102:273–283. doi: 10.1007/s001220051644

Chakrabarti M, Zhang N, Sauvage C, *et al* (2013) A cytochrome P450 regulates a domestication trait in cultivated tomato. Proc Natl Acad Sci 110:17125–17130. doi: 10.1073/pnas.1307313110

Chan EKF, Rowe HC, Corwin JA, *et al* (2011) Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in *Arabidopsis thaliana.* PLoS Biol 9:e1001125. doi: 10.1371/journal.pbio.1001125

Chaves MM, Flexas J, Pinheiro C (2009) Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. Ann Bot 103:551–60. doi: 10.1093/aob/mcn125

Chaves MM, Maroco JP, Pereira JS (2003) Understanding plant responses to drought — from genes to the whole plant. Funct Plant Biol 30:239. doi: 10.1071/FP02076

Chaves MM, Oliveira MM (2004) Mechanisms underlying plant resilience to water deficits: prospects for water-saving agriculture. J Exp Bot 55:2365–84. doi: 10.1093/jxb/erh269

Chaves MM, Pereira JS, Maroco J, *et al* (2002) How plants cope with water stress in the field? Photosynthesis and Growth. Ann Bot 89:907–916. doi: 10.1093/aob/mcf105

Cheung VG, Spielman RS, Ewens KG, *et al* (2005) Mapping determinants of human gene expression by regional and genome-wide association. Nature 437:1365–1369. doi: 10.1038/nature04244

Collins NC, Tardieu F, Tuberosa R (2008) Quantitative trait loci and crop performance under abiotic stress: where do we stand? Plant Physiol 147:469–86. doi: 10.1104/pp.108.118117

Cookson W, Liang L, Abecasis G, *et al* (2009) Mapping complex disease traits with global gene expression. Nat Rev Genet 10:184–194. doi: 10.1038/nrg2537

Coupel-Ledru A, Lebon E, Christophe A, *et al* (2014) Genetic variation in a grapevine progeny (*Vitis vinifera* L. cvs Grenache×Syrah) reveals inconsistencies between maintenance of daytime leaf water potential and response of transpiration rate under drought. J Exp Bot 65:6205–18. doi: 10.1093/jxb/eru228

Cui N, Du T, Kang S, *et al* (2008) Regulated deficit irrigation improved fruit quality and water use efficiency of pear-jujube trees. Agric Water Manag 95:489–497. doi: 10.1016/j.agwat.2007.11.007

Dai A (2011) Drought under global warming: a review. Wiley Interdiscip Rev Clim Chang 2:45–65. doi: 10.1002/wcc.81

Des Marais DL, Hernandez KM, Juenger TE (2013) Genotype-by-environment interaction and plasticity: Exploring genomic responses of plants to the abiotic environment. Annu Rev Ecol Evol Syst 44:5–29. doi: 10.1146/annurev-ecolsys-110512-135806

Des Marais DL, McKay JK, Richards JH, *et al* (2012) Physiological genomics of response to soil drying in diverse *Arabidopsis* accessions. Plant Cell 24:893–914. doi: 10.1105/tpc.112.096180

Dhanda SS, Sethi GS, Behl RK (2004) Indices of drought tolerance in wheat genotypes at early stages of plant growth. J Agron Crop Sci 190:6–12. doi: 10.1111/j.1439-037X.2004.00592.x

Dorais M, Ehret DL, Papadopoulos AP (2008) Tomato (*Solanum lycopersicum*) health components: from the seed to the consumer. Phytochem Rev 7:231–250. doi: 10.1007/s11101-007-9085-x

Durán Zuazo VH, Pleguezuelo CRR, Tarifa DF (2011) Impact of sustained-deficit irrigation on tree growth, mineral nutrition, fruit yield and quality of mango in Spain. Fruits 66:257–268. doi: 10.1051/fruits/2011038

Easlon HM, St. Clair D, Bloom AJ (2014) An introgression from wild tomato (*Solanum habrochaites*) affects tomato photosynthesis and water relations. Crop Sci 54:779–784. doi: 10.2135/cropsci2013.06.0401

El-Soda M, Boer MP, Bagheri H, *et al* (2014a) Genotype-environment interactions affecting preflowering physiological and morphological traits of *Brassica rapa* grown in two watering regimes. J Exp Bot 65:697–708. doi: 10.1093/jxb/ert434

El-Soda M, Malosetti M, Zwaan BJ, *et al* (2014b) Genotype × environment interaction QTL mapping in plants: lessons from *Arabidopsis*. Trends Plant Sci 19:390–8. doi: 10.1016/j.tplants.2014.01.001

Farooq M, Hussain M, Wahid A, Siddique KHM (2012) Plant responses to drought stress. Springer Berlin Heidelberg, Berlin, Heidelberg

Favati F, Lovelli S, Galgano F, *et al* (2009) Processing tomato quality as affected by irrigation scheduling. Sci Hortic (Amsterdam) 122:562–571. doi: 10.1016/j.scienta.2009.06.026

Foolad MR (2004) Recent advances in genetics of salt tolerance in tomato. Plant Cell Tissue Organ Cult 76:101–119. doi: 10.1023/B:TICU.0000007308.47608.88

Foolad MR, Zhang LP, Subbiah P (2003) Genetics of drought tolerance during seed germination in tomato: inheritance and QTL mapping. Genome 46:536–45. doi: 10.1139/g03-035

Frary A (2000) fw2.2: A quantitative trait locus key to the evolution of tomato fruit size. Science (80-) 289:85–88. doi: 10.1126/science.289.5476.85

Gagnot S, Tamby J-P, Martin-Magniette M-L, *et al* (2008) CATdb: a public access to *Arabidopsis* transcriptome data from the URGV-CATMA platform. Nucleic Acids Res 36:D986–D990. doi: 10.1093/nar/gkm757

Gao X, Giorgi F (2008) Increased aridity in the Mediterranean region under greenhouse gas forcing estimated from high resolution simulations with a regional climate model. Glob Planet Change 62:195–209. doi: 10.1016/j.gloplacha.2008.02.002

Gershenzon J (1984) Changes in the levels of plant secondary metabolites under water and nutrient stress. In: Timmermann BN, Steelink C, Loewus FA (eds) Phytochemical Adaptations to Stress. Springer US, Boston, MA, pp 273–320

Giovannucci E (1999) Tomatoes, tomato-based products, lycopene and cancer: Review of the epidemiologic literature. JNCI J Natl Cancer Inst 91:317–331. doi: 10.1093/jnci/91.4.317

Giovannucci E (2002) A prospective study of tomato products, lycopene and prostate cancer risk. CancerSpectrum Knowl Environ 94:391–398. doi: 10.1093/jnci/94.5.391

Gong M, Chen S-N, Song Y-Q, Li Z-G (1997) Effect of calcium and calmodulin on intrinsic heat tolerance in relation to antioxidant systems in maize seedlings. Aust J Plant Physiol 24:371. doi: 10.1071/PP96118

Guedj M, Robin S, Celisse A, Nuel G (2009) Kerfdr: a semi-parametric kernel-based approach to local false discovery rate estimation. BMC Bioinformatics 10:84. doi: 10.1186/1471-2105-10-84

Guichard S, Gary C, Leonardi C, Bertin N (2005) Analysis of growth and water relations of tomato fruits in relation to air vapor pressure deficit and plant fruit load. J Plant Growth Regul 24:201–213. doi: 10.1007/s00344-005-0040-z

Gur A, Semel Y, Osorio S, *et al* (2011) Yield quantitative trait loci from wild tomato are predominately expressed by the shoot. Theor Appl Genet 122:405–20. doi: 10.1007/s00122-010-1456-9

Harb A, Krishnan A, Ambavaram MMR, Pereira A (2010) Molecular and physiological analysis of drought stress in *Arabidopsis* reveals early responses leading to acclimation in plant growth. Plant Physiol 154:1254–1271. doi: 10.1104/pp.110.161752

Hausmann NJ, Juenger TE, Sen S, *et al* (2005) Quantitative trait loci affecting ɕ13 C and response to differential water availibility in *Arabidopsis thaliana.* Evolution (N Y) 59:81–96. doi: 10.1111/j.0014-3820.2005.tb00896.x

Hirayama T, Shinozaki K (2010) Research on plant abiotic stress responses in the post-genome era: past, present and future. Plant J 61:1041–52. doi: 10.1111/j.1365-313X.2010.04124.x

Hsieh T, Lee J, Charng Y, Chan M (2002) Tomato plants ectopically expressing *Arabidopsis* CBF1 show enhanced resistance to water deficit stress 1. Plant Physiol 130:618–626. doi: 10.1104/pp.006783.KIN1

Huang Z, van der Knaap E (2011) Tomato fruit weight 11.3 maps close to fasciated on the bottom of chromosome 11. Theor Appl Genet 123:465–474. doi: 10.1007/s00122-011-1599-3

Illa-Berenguer E, Van Houten J, Huang Z, van der Knaap E (2015) Rapid and reliable identification of tomato fruit weight and locule number loci by QTL-seq. Theor Appl Genet 1 – 14. doi: 10.1007/s00122-015-2509-x

Jansen RC, Van Ooijen JW, Stam P, *et al* (1995) Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. Theor Appl Genet 91:33–37. doi: 10.1007/BF00220855

Jiang M, Zhang J (2002) Water stress-induced abscisic acid accumulation triggers the increased generation of reactive oxygen species and up-regulates the activities of antioxidant enzymes in maize leaves. J Exp Bot 53:2401–2410. doi: 10.1093/jxb/erf090

Jiang S-Y, Ma Z, Ramachandran S (2010) Evolutionary history and stress regulation of the lectin superfamily in higher plants. BMC Evol Biol 10:79. doi: 10.1186/1471-2148-10-79

Jureková Z, Németh-molnár K, Paganová V (2011) Physiological responses of six tomato (*Lycopersicon esculentum* Mill .) cultivars to water stress. J Hortic For 3:294–300.

Kader A a, Stevens MA, Albright-Holton M, *et al* (1977) Effect of fruit ripeness when picked on flavor and composition in fresh market tomatoes. J. Am. Soc. Hortic. Sci. 102:724–731.

Katerji N, Mastrorilli M, Rana G (2008) Water use efficiency of crops cultivated in the Mediterranean region: Review and analysis. Eur J Agron 28:493–507. doi: 10.1016/j.eja.2007.12.003

Khachik F, Carvalho L, Bernstein PS, *et al* (2002) Chemistry, distribution, and metabolism of tomato carotenoids and their impact on human health. Exp Biol Med 227:845–851.

Kirda C, Cetin M, Dasgan Y, *et al* (2004) Yield response of greenhouse grown tomato to partial root drying and conventional deficit irrigation. Agric Water Manag 69:191–201. doi: 10.1016/j.agwat.2004.04.008

Kissoudis C, Chowdhury R, van Heusden S, *et al* (2015) Combined biotic and abiotic stress resistance in tomato. Euphytica 202:317–332. doi: 10.1007/s10681-015-1363-x

Korte A, Vilhjálmsson BJ, Segura V, *et al* (2012) A mixed-model approach for genome-wide association studies of correlated traits in structured populations. Nat Genet 44:1066–71. doi: 10.1038/ng.2376

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199.

Langridge P (2006) Functional genomics of abiotic stress tolerance in cereals. Briefings Funct Genomics Proteomics 4:343–354. doi: 10.1093/bfgp/eli005

Lecomte L, Saliba-Colombani V, Gautier A, *et al* (2004) Fine mapping of QTLs of chromosome 2 affecting the fruit architecture and composition of tomato. Mol Breed 13:1–14. doi: 10.1023/B:MOLB.0000012325.77844.0c

Lee H, Yoo SJ, Lee JH, *et al* (2010) Genetic framework for flowering-time regulation by ambient temperature-responsive miRNAs in *Arabidopsis*. Nucleic Acids Res 38:3081–3093. doi: 10.1093/nar/gkp1240

Lester GE (2006) Environmental regulation of human health nutrients (ascorbic acid, carotene and folic acid) in fruits and vegetables. HortScience 41:1.

Li S, Wang J, Zhang L (2015) Inclusive composite interval mapping of QTL by environment interactions in biparental populations. PLoS One 10:e0132414. doi: 10.1371/journal.pone.0132414

Liu G, Zhu H, Zhang G, *et al* (2012) Dynamic analysis of QTLs on tiller number in rice (*Oryza sativa* L.) with single segment substitution lines. Theor Appl Genet 125:143–153. doi: 10.1007/s00122-012-1822-x

Liu H, Wang X, Zhang H, *et al* (2008) A rice serine carboxypeptidase-like gene OsBISCPL1 is involved in regulation of defense responses against biotic and oxidative stress. Gene 420:57–65. doi: 10.1016/j.gene.2008.05.006

Lovell JT, Juenger TE, Michaels SD, *et al* (2013) Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. Proc R Soc B Biol Sci 280:20131043–20131043. doi: 10.1098/rspb.2013.1043

Lovell JT, Mullen JL, Lowry DB, *et al* (2015) Exploiting differential gene expression and epistasis to discover candidate genes for drought-associated QTLs in *Arabidopsis thaliana*. Plant Cell 27:969–983. doi: 10.1105/tpc.15.00122

Lu C, Shen L, He P, *et al* (1997) Comparative mapping of QTLs for agronomic traits of rice across environments by using a doubled-haploid population. Theor Appl Genet 94:145–150. doi: 10.1007/s001220050393

Malosetti M, Ribaut JM, Vargas M, *et al* (2007) A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). Euphytica 161:241–257. doi: 10.1007/s10681-007-9594-0

Mirás-Avalos JM, Alcobendas R, Alarcón JJ, *et al* (2013) Assessment of the water stress effects on peach fruit quality and size using a fruit tree model, QualiTree. Agric Water Manag 128:1–12. doi: 10.1016/j.agwat.2013.06.008

Mitchell JP, Shennann C, Grattan SR, May DM (1991) Tomato fruit yields and quality under water deficit and salinity. J Am Soc Hortic 116:215–221.

Mitra J (2001) Genetics and genetic improvement of drought resistance in crop plants. Curr Sci 80:758–763.

Moura DS, Bergey DR, Ryan CA (2001) Characterization and localization of a wound-inducible type I serine-carboxypeptidase from leaves of tomato plants (*Lycopersicon esculentum* Mill.). Planta 212:222–230. doi: 10.1007/s004250000380

Muir W, Nyquist WE, Xu S (1992) Alternative partitioning of the genotype-by-environment interaction. Theor Appl Genet 84:193–200. doi: 10.1007/BF00224000

Murshed R, Lopez-Lauri F, Sallanon H (2013) Effect of water stress on antioxidant systems and oxidative parameters in fruits of tomato (*Solanum lycopersicon* L, cv. Micro-tom). Physiol Mol Biol Plants 19:363–378. doi: 10.1007/s12298-013-0173-7

Nicolae DL, Gamazon E, Zhang W, *et al* (2010) Trait-associated SNPs are more likely to be eQTLs: Annotation to enhance discovery from GWAS. PLoS Genet 6:e1000888. doi: 10.1371/journal.pgen.1000888

Nora L, Dalmazo GO, Nora F, Rombaldi CV (2012) Controlled water stress to improve fruit and vegetable postharvest quality. In: Ismal M, Mofizur R, Hiroshi H (eds) Water Stress. Tech Open Science, Rijeka, pp 59 – 72

Pascual L, Albert E, Sauvage C, *et al* (2015) Dissecting quantitative trait variation in the resequencing era: Complementarity of bi-parental, multi-parental and association panels. Plant Sci. doi: 10.1016/j.plantsci.2015.06.017

Patanè C, Cosentino SL (2010) Effects of soil water deficit on yield and quality of processing tomato under a Mediterranean climate. Agric Water Manag 97:131–138. doi: 10.1016/j.agwat.2009.08.021

Paterson H, Damon S, Hewitt JD, *et al* (1991) Mendelian factors underlyiing quantitative traits in tomato: comparison across species, generation and environments. Genetics 127:181–197.

Perruc E, Charpenteau M, Ramirez BC, *et al* (2004) A novel calmodulin-binding protein functions as a negative regulator of osmotic stress tolerance in *Arabidopsis thaliana* seedlings. Plant J 38:410–420. doi: 10.1111/j.1365-313X.2004.02062.x

R Development Core Team (2012) R: a language and environment for statistical computing.

Rai AC, Singh M, Shah K (2013) Engineering drought tolerant tomato plants over-expressing BcZAT12 gene encoding a C_2H_2 zinc finger transcription factor. Phytochemistry 85:44–50. doi: 10.1016/j.phytochem.2012.09.007

Ratanachinakorn B, Klieber A, Simons DH (1997) Effect of short-term controlled atmospheres and maturity on ripening and eating quality of tomatoes. Postharvest Biol Technol 11:149–154. doi: 10.1016/S0925-5214(97)00021-5

Reddy ASN, Ali GS, Celesnik H, Day IS (2011) Coping with stresses: roles of calcium- and calcium/calmodulin-regulated gene expression. Plant Cell 23:2010–2032. doi: 10.1105/tpc.111.084988

Ripoll J, Urban L, Staudt M, *et al* (2014) Water shortage and quality of fleshy fruits, making the most of the unavoidable. J Exp Bot 65:4097–117. doi: 10.1093/jxb/eru197

Sabadin PK, Malosetti M, Boer MP, *et al* (2012) Studying the genetic basis of drought tolerance in sorghum by managed stress trials and adjustments for phenological and plant height differences. Theor Appl Genet 124:1389–1402. doi: 10.1007/s00122-012-1795-9

Saïdou A-A, Thuillet A-C, Couderc M, *et al* (2014) Association studies including genotype by environment interactions: prospects and limits. BMC Genet 15:3. doi: 10.1186/1471-2156-15-3

Saliba-Colombani V, Causse M, Gervais L, Philouze J (2000) Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. Genome 43:29–40. doi: 10.1139/gen-43-1-29

Saliba-Colombani V, Causse M, Langlois D, *et al* (2001) Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. TAG Theor Appl Genet 102:259–272. doi: 10.1007/s001220051643

Santesteban LG, Royo JB (2006) Water status, leaf area and fruit load influence on berry weight and sugar accumulation of cv. “Tempranillo” under semiarid conditions. Sci Hortic (Amsterdam) 109:60–65. doi: 10.1016/j.scienta.2006.03.003

Sauvage C, Segura V, Bauchet G, *et al* (2014) Genome-Wide Association in tomato reveals 44 candidate loci for fruit metabolic traits. Plant Physiol 165:1120–1132. doi: 10.1104/pp.114.241521

Schlichting CD (1986) The evolution of phenotypic plasticity in plants. Annu Rev Ecol Syst 17:667–693.

Seki M, Narusaka M, Ishida J, *et al* (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J 31:279–292. doi: 10.1046/j.1365-313X.2002.01359.x

Semel Y, Schauer N, Roessner U, *et al* (2007) Metabolite analysis for the comparison of irrigated and non-irrigated field grown tomato of varying genotype. Metabolomics 3:289–295. doi: 10.1007/s11306-007-0055-5

Shaar-Moshe L, Hübner S, Peleg Z (2015) Identification of conserved drought-adaptive genes using a cross-species meta-analysis approach. BMC Plant Biol 15:111. doi: 10.1186/s12870-015-0493-6

Shinozaki K, Yamaguchi-Shinozaki K (2007) Gene networks involved in drought stress response and tolerance. J Exp Bot 58:221–7. doi: 10.1093/jxb/erl164

Silva D, Albuquerque D, de Azevedo (2013) Drought and its consequences to plants – From individual to ecosystem. In: Akinci S (ed) Responses of Organisms to Water Stress. InTech, Turkey, p 17

Sim S-C, Durstewitz G, Plieske J, *et al* (2012) Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. PLoS One 7:e40563. doi: 10.1371/journal.pone.0040563

Skirycz A, Vandenbroucke K, Clauw P, *et al* (2011) Survival and growth of *Arabidopsis* plants given limited water are not equal. Nat Biotechnol 29:212–214. doi: 10.1038/nbt.1800

Smirnoff N (1996) The function and metabolism of ascorbic acid in plant. Ann Bot 78:661–669.

Smyth GK (2005) Limma: linear models for microarray data. Bioinformatics and computational biology solutions using R and Bioconductor, Springer N. pp 397 – 420

Sterken R, Kiekens R, Boruc J, *et al* (2012) Combined linkage and association mapping reveals CYCD5;1 as a quantitative trait gene for endoreduplication in *Arabidopsis*. Proc Natl Acad Sci 109:4678–4683. doi: 10.1073/pnas.1120811109

Stevens MA (1972) Relationships between components contributing to quality variation among tomato lines. J Am Soc Hortic Sci 70 – 76.

Stevens R, Buret M, Garchery C, *et al* (2006) Technique for rapid, small-scale analysis of vitamin C levels in fruit and application to a tomato mutant collection. J Agric Food Chem 54:6159–65. doi: 10.1021/jf061241e

Stevens R, Page D, Gouble B, *et al* (2008) Tomato fruit ascorbic acid content is linked with monodehydroascorbate reductase activity and tolerance to chilling stress. Plant Cell Environ 31:1086–96. doi: 10.1111/j.1365-3040.2008.01824.x

Storey JD (2007) The optimal discovery procedure: a new approach to simultaneous significance testing. J R Stat Soc Ser B (Statistical Methodol 69:347–368. doi: 10.1111/j.1467-9868.2007.005592.x

Tardieu F, Granier C, Muller B (2011) Water deficit and growth. Co-ordinating processes without an orchestrator? Curr Opin Plant Biol 14:283–9. doi: 10.1016/j.pbi.2011.02.002

Tardieu F, Tuberosa R (2010) Dissection and modelling of abiotic stress tolerance in plants. Curr Opin Plant Biol 13:206–12. doi: 10.1016/j.pbi.2009.12.012

Tétard-Jones C, Kertesz M a, Preziosi RF (2011) Quantitative trait loci mapping of phenotypic plasticity and genotype-environment interactions in plant and insect performance. Philos Trans R Soc Lond B Biol Sci 366:1368–79. doi: 10.1098/rstb.2010.0356

Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635–41. doi: 10.1038/nature11119

Uozumi A, Ikeda H, Hiraga M, *et al* (2012) Tolerance to salt stress and blossom-end rot in an introgression line, IL8-3, of tomato. Sci Hortic (Amsterdam) 138:1–6. doi: 10.1016/j.scienta.2012.01.036

Van Damme EJM, Barre A, Rougé P, Peumans WJ (2004) Cytoplasmic/nuclear plant lectins: a new story. Trends Plant Sci 9:484–489. doi: 10.1016/j.tplants.2004.08.003

Van Eeuwijk F a, Bink MC, Chenu K, Chapman SC (2010) Detection and use of QTL for complex traits in multiple environments. Curr Opin Plant Biol 13:193–205. doi: 10.1016/j.pbi.2010.01.001

Vaucheret H (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev 18:1187–1197. doi: 10.1101/gad.1201404

Vaucheret H (2008) Plant ARGONAUTES. Trends Plant Sci 13:350–358. doi: 10.1016/j.tplants.2008.04.007

Via S, Lande R (1985) Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution (N Y) 39:505–522.

Westwood JH, Mccann L, Naish M, *et al* (2013) A viral RNA silencing suppressor interferes with abscisic acid-mediated signalling and induces drought tolerance in *Arabidopsis thaliana*. Mol Plant Pathol 14:158–170. doi: 10.1111/j.1364-3703.2012.00840.x

Wricke G (1964) The calculation of ecovalence in summer wheat and oat. Z Pflanzenzuecht 52:127 – 138.

Yang J, Hu C, Hu H, *et al* (2008) QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. Bioinformatics 24:721–3. doi: 10.1093/bioinformatics/btm494

Yang YH, Dudoit S, Lin DM, *et al* (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res 30:15e–15. doi: 10.1093/nar/30.4.e15

Yilmaz E (2001) The Chemistry of Fresh Tomato Flavor. Turk J Agric 25:149–155.

Zairi A, El Amami H, Slatni A, *et al* (2003) Coping with drought: Deficit irrigation strategies for cereals and field horticultural crops in central Tunisia. Tools for Drought Mitigation in Mediterranean Regions, Springer. The Netherlands, pp 181–201

Zhang N, Brewer MT, van der Knaap E (2012) Fine mapping of fw3.2 controlling fruit weight in tomato. Theor Appl Genet 125:273–84. doi: 10.1007/s00122-012-1832-8

Zheng J, Huang G, Jia D, *et al* (2013) Responses of drip irrigated tomato (*Solanum lycopersicum* L.) yield, quality and water productivity to various soil matric potential thresholds in an arid region of Northwest China. Agric Water Manag 129:181–193. doi: 10.1016/j.agwat.2013.08.001

Zhu M, Chen G, Zhang J, *et al* (2014) The abiotic stress-responsive NAC-type transcription factor SlNAC4 regulates salt and drought tolerance and stress-related genes in tomato (*Solanum lycopersicum*). Plant Cell Rep 33:1851–1863. doi: 10.1007/s00299-014-1662-z

CHAPTER 4

CHAPTER 4: Association mapping in an unrelated collection of small fruit tomatoes grown under two watering conditions and identification of the genetic determinants of major fruit quality traits

This chapter is in the form of an article published in *Journal of Experimental Botany*. The article describes association mapping for fruit quality traits in a collection of small fruit tomatoes grown under two watering conditions in two locations. Results are combined with those reported in the biparental population (**Chapter 3**) to draw a detailed characterization of the genetic variations and genomic determinants of response to water deficit in tomato. Then, QTLs for major fruit quality traits are dissected using publicly available expression data, exonic variants gained from re-sequencing of four accessions of the collection and functional analysis of the gene annotations in the confidence intervals.

Association mapping reveals the genetic architecture of tomato response to water deficit: focus on major fruit quality traits.

doi:10.1093/jxb/erw411 (Accepted: 13/10/2016, Published: 18/11/2016)

Authors

Elise Albert[a], Vincent Segura[b], Justine Gricourt[a], Julien Bonnefoi[c], Laurent Derivot[c], Mathilde Causse[a], §

Affiliations

[a] INRA, UR1052, Génétique et Amélioration des Fruits et Légumes, 67 Allée des chênes, Centre de Recherche PACA, Domaine Saint Maurice, CS60094, Montfavet, 84143, France

[b] INRA, UR0588, Amélioration, Génétique et Physiologie Forestières, 2163 Avenue de la Pomme de Pin, Centre de Recherche Val de Loire, CS 40001, Orléans, 45075, France

[c] GAUTIER Semences, route d'Avignon, Eyragues, 13630, France

§ Corresponding author

Mathilde Causse

Mathilde.Causse@paca.inra.fr

Tel: +33 (0)4 32 72 28 03

Supplemental materials referred in this chapter are available in **Appendix 5.**

Abstract

Water scarcity constitutes a crucial constraint for agriculture productivity. High-throughput approaches in model plant species identified hundreds of genes potentially involved in survival under drought, but few having beneficial effects on quality and yield. Nonetheless, controlled water deficit may improve fruit quality through higher concentration of flavour compounds. The underlying genetic determinants are still poorly known. In this study, we phenotyped 141 highly diverse small fruit tomato accessions for 27 traits under two contrasted watering conditions. A subset of 55 accessions exhibited increased metabolite contents and maintained yield under water deficit. Using 6,100 SNP, association mapping revealed 31, 41 and 44 QTLs under drought, control and both conditions, respectively. Twenty five additional QTLs were interactive between conditions, emphasizing the interest of accounting for QTL by watering regime interactions in fruit quality improvement. Combining our results with the loci previously identified in a biparental progeny resulted in eleven common QTLs and contribute to a first detailed characterization of the genetic determinants of response to water deficit in tomato. Major QTLs for fruit quality traits were dissected and candidate genes were proposed using expression and polymorphism data. The outcomes provide basis for fruit quality improvement under deficit irrigation while limiting yield losses.

Key words

Drought, Fleshy fruit quality, Genotype by environment interaction, GWA, QTL, *Solanum* lycopersicum

Highlight

Tomato quality could be improved under deficit irrigation while maintaining yield. The underlying genetic architecture is polygenic and varies with water availability. Candidate genes related to primary metabolism were identified.

Introduction

Global water scarcity will constitute a crucial challenge in the coming years (Jury and Vaux, 2005). Agriculture which is consuming up to 80% of the worldwide water resource through irrigation has to move towards a more sustainable use of water (Rost *et al*., 2008). Utilization of advanced irrigation strategies and development of drought-adapted crops are among the solutions to solve this dilemma (Fereres and Soriano, 2006; Costa *et al*., 2007).

Beyond these concerns, deficit irrigation practices constitute a way to manage fruit flavor by exploiting the morphological, physiological and molecular changes (referred as 'phenotypic plasticity') occurring in water stressed plants (Ripoll *et al*., 2014). Under water deficit, plants close their stomata to limit transpiration, impacting resource availability from photosynthetic sources, which may result in a decrease in number and/or size of the fruits. On the other hand, a mild water deficit tends to shift photo-assimilates partitioning towards synthesis of antioxidants compounds (in particular vitamin C) involved in defense against stress induced reactive oxygen species and compatible solutes (including sugars and acids) involved in osmotic adjustment (Lemoine *et al*., 2013; Albacete *et al*., 2014; Osorio *et al*., 2014). Evidence of the efficiency of deficit irrigation to concentrate the major flavor and nutritional components in fleshy fruits (mainly sugars, acids and antioxidants), either by concentration or accumulation effect, was obtained in many species such as tomato (Kirda *et al*., 2004; Zheng *et al*., 2013), grapevine (Chaves *et al*., 2007), apple (Leib *et al*., 2006) and mango (Durán Zuazo *et al*., 2011). However, these studies focused on small number of genotypes while responses to deficit irrigation seem to be highly genotype dependent (Ripoll *et al*., 2016a,b).

Gene expression studies have revealed hundreds of genes involved in plant survival under severe water limitation, but usually associated to detrimental effects on yield under realistic drought scenario (Tardieu, 2012; Bac-Molenaar *et al*., 2016). These studies focused on model species, mainly *Arabidopsis thaliana* (Seki *et al*., 2002; Des Marais *et al*., 2012) and cereals (Langridge, 2006; Barnabas *et al*., 2007). Up to now, the identification of the genetic determinants of drought response from the natural diversity of fleshy fruit crops remains limited. Quantitative trait locus (QTL) mapping might be particularly valuable to address this question (Des Marais *et al*., 2013).

Two complementary approaches are commonly applied to dissect genotype by environment interactions into their underlying QTLs (QTL by environment interactions). The first one consists in computing the effects of a given QTL across the environmental conditions using multivariate QTL mapping models (van Eeuwijk *et al*., 2010; El-Soda *et al*., 2014b). The second one uses the construction of composite variables measuring phenotypic plasticity and univariate mapping models (El-Soda *et al*., 2014a). With both approaches, QTLs can be classified according to the prevalence of their effect under the different conditions. A QTL is considered 'constitutive' when its effect is conserved whatever the environment. QTLs whose effect is not significant in every environment are called 'specific', while 'interactive' QTLs have their effect changing direction ('antagonist') or intensity ('differential') according to the environment. With the availability of high throughput genotyping assay, this classification can be considered in crop species via conventional linkage mapping (Malosetti *et al*., 2007; Verbyla *et al*., 2014) as well as genome-wide association study (GWAS) (Korte *et al*., 2012; Saïdou *et al*., 2014). GWAS has the advantage over linkage mapping to allow exploration of the genetic diversity and the numerous recombination events present in germplasm collections and may lead to higher resolution mapping if the LD (linkage disequilibrium) is low enough in the population (Brachi *et al*., 2010; Korte and Farlow, 2013; El-Soda *et al*., 2015; Pascual *et al*., 2016).

In tomato (*Solanum lycopersicum* L.), QTLs were mapped for fruit quality traits measured under optimal watering conditions using linkage (Causse *et al*., 2001; Saliba-Colombani *et al*., 2001; Tieman *et al*., 2006; Zanor *et al*., 2009b; Capel *et al*., 2015) and association mapping (Xu *et al*., 2013; Ruggieri *et al*., 2014; Sauvage *et al*., 2014; Sacco *et al*., 2015). The studies of QTL by water regime interactions focused on introgression lines between the cultivated tomato and its wild relatives (mainly *S. habrochaites* and *S. pennellii*) leading to low mapping resolution (Semel *et al*., 2007; Gur *et al*., 2011; Arms *et al*., 2015). Recently, we analyzed QTL by watering regime interaction in a segregating population derived from a cross between a small and a big fruited *S. lycopersicum* accessions (Albert *et al*., 2016). A total of 56 QTLs were identified for 19 traits, among which 20% were interactive between the control and deficit watering regimes. Nevertheless, these QTLs were limited to the allelic diversity present in the two parental accessions and the confidence intervals were broad.

The aims of the present study were (1) to explore the pattern of genotype by watering regime interaction in a GWAS panel with broad genetic basis (including *S. pimpinellifolium, S. lycopersicum* var *cerasiforme* and admixture genotypes) grown under two different watering regimes in two locations and phenotyped for 27 traits, (2) to identify with a high resolution QTLs and QTL by watering regime interactions in this collection, (3) to combine the results to those obtained in the bi-parental progeny to draw an accurate picture of the genetic variability and the genetic determinants of tomato response to water deficit and (4) to identify candidate genes related to the variation of major fruit quality traits under water deficit by dissecting some of the QTLs.

Material and Methods

1. Plant material

The population consisted in 141 accessions (fresh weight from 2 to 46 g) encompassing the genetic diversity of the cultivated small fruit tomato. Among these, 105 accessions were previously investigated in Blanca *et al*. (2015). Preliminary genetic analysis of our collection confirmed the genetic structure described by these authors, with clusters reflecting the species and the geographic origin of the accessions (Supplemental Figure 1 A to D). Ten accessions were *S. pimpinellifolium* (SP, closest wild ancestor from the tomato) originated from Peru and Ecuador. A total of 110 accessions were *Solanum lycopersicum* var. *cerasiforme* (SLC) originated mainly from South America. Finally, 21 accessions belonged to a mixture genetic group mainly including commercial cherry tomatoes and admixed genotypes between *SP*, *SLC* and *S. lycopersicum* var. *lycopersicum*. Description of the accessions and their origin are available in **Supplemental Table 1**. The genetic groups (SLC, SP and mixture) are used below in the statistical analysis.

2. Experimental design

The plants were cultivated with the same experimental design as in Albert *et al*. (2016). Plants were grown in heated glasshouse in INRA Avignon (Avi, France) from March to July 2014 and in unheated plastic greenhouse on the experimental site of the seed company GAUTIER Semences in Agadir (Aga, Morocco) from December 2013 to March 2014. Two watering regimes were applied to the plants: control (C) and drought (D). The control

treatment was set according to ETp and the cultural coefficient for tomato under greenhouse (FAO Water, 2015). A maximal drainage of 25% and a relative humidity of the substrate of 65% were established in the control pots. Drought treatment was applied progressively after flowering of the second truss of the earliest accession. Watering was first reduced by 25% compared with control for one week and then reduced by 60% until the end of the experiments. Relative humidity of the peat substrate was controlled with GRODAN® moisture probes and monitored between 25% and 30% in drought pots. In both experiments, two plants per watering regime per accession were randomized in the greenhouse.

3. Plant and fruit phenotyping

A total of 27 traits were assessed in the GWA population as described in Albert *et al*. (2016). Flowering date (Flw, days after sowing), stem diameter (Diam, mm), leaf length (Leaf, cm) and truss implantation height (Ht, cm) were measured on each plant both in Avignon (6th truss) and Agadir (5th truss). Plant fruit number (Nbfruits, all fruits from 3rd to 6th truss) was measured in Avignon solely.

Fruit quality measurements were carried out on a minimum of twenty mature fruits per accession per watering regime harvested daily on the 3rd to the 6th truss. All the fruits were weighted (FW, g) and their firmness was measured with a Durofel device (FIR). In Avignon solely, fruits were pooled in three groups in each watering regime. Half of the fruits of each pool were used to assess dry matter weight (DMW, %), pH and soluble solid content (SSC, °Brix). From the second half of the fruit replicates, pericarps were crushed in liquid nitrogen and assayed for total vitamin C content (VitCFM) according to the microplate method described in Stevens *et al*. (2006), for sugar content (Glucose and Fructose) according to the enzymatic method described in Gomez *et al*. (2007) and for organic acid content (Malic and Citric) according to the HPLC method reported in Wu *et al*. (2002). The different metabolite concentrations were expressed relatively to fresh matter (g per 100g of FM) and relatively to dry matter (g per 100g of DM). Yield (Yield, g/plant) was computed by multiplying average fruit fresh weight by average fruit number per plant.

4. Plant genotyping and SNP filtering

The GWA population was genotyped using the Tomato Infinium Array developed within the SolCAP project (http://solcap.msu.edu/) (Hamilton *et al*., 2012; Sim *et al*., 2012). The maximum rates of missing data were fixed at 25% per accessions and 10% per SNP. A MAF threshold of 0.04 was applied to discard markers with very rare alleles according to Aulchenko *et al*. (2007). After filtering, the set of markers was constituted of 6,100 SNP. Prior to any genetic analysis, the remaining missing genotypes were replaced by the allele frequency of the major allele. The SNP were renamed according to their positions on the tomato genome (SL2.50), as S01_58000085 at position 58,000,085 pb on chromosome 1 **(Supplemental Table 2).**

5. Statistical analysis of the phenotypic data

All statistical analyses were performed using R (R Development Core Team, 2012). Because fewer and different traits were measured in Agadir experiments, data from both locations were analyzed separately (Pearson correlations for the common trait means available in **Supplemental Table 4** – all significant). Prior to the analyses of variance (ANOVA) and when distributions were skewed, phenotypic data were normalized using Box and Cox transformations. The ANOVA were performed according to the following model:

$Y_{ijkl} = \mu + Gr_i + Gr_i(G_j) + W_k + Gr_i * W_k + Gr_i(G_j) * W_k + e_{ijkl}$

Y_{ijkl} was the phenotypic value of accession *j* from genetic group *i* in watering regime *k*, μ the overall mean, Gr_i the fixed effect of genetic group *i*, $Gr_i(G_j)$ the fixed effect of accession *j* nested in genetic group *i*, W_k the fixed effect of watering regime *k* and e_{ijkl} the residual error effect. No significant microenvironment pattern was identified and we chose to not include any spatial effect in the model. When the interaction *Gr x W* was significant, we computed a Tukey's post-hoc test to compare the means.

Then, in both watering regimes, restricted maximum likelihood estimates of the genetic and residual variances (σ^2_G and σ^2_e) were computed with a second linear model: $Y_{ijk} = \mu + Gr_i + Gr_i(G_j) + e_{ijk}$ (Gr_j fixed, G_i and e_{ijk} random). Broad-sense heritabilities (H^2) were calculated under both watering regimes as the ratio between the genetic variance and the total phenotypic variance: $H^2 = \sigma^2_G / \sigma^2_{Total}$, with $\sigma^2_{Total} = \sigma^2_G + 1/n*\sigma^2e$ (with n the number of replicates per accession). Spearman coefficients estimated the correlations between H^2 and σ^2_G under drought and control conditions for a same trait.

Average values per accession in each watering regime and location were used for subsequent analyses. Plasticity was computed on the accession means as: $\Delta ki = (D_{ki} - C_{ki})/C_{ki}$, with Δki the plasticity value for trait k and accession i, D_{ki} the mean of trait k under drought condition for accession i and C_{ki} the mean of trait k under control condition for accession i.

6. Construction of kinship and structure matrices

We performed a Principal Coordinate Analysis (PCoA) on the genotype matrix. The coordinates of the accessions on the first three components are available in **Supplemental Table 3** and displayed graphically in **Supplemental Figure 1**. A kinship matrix (K) based on identity by state among the 6,100 SNP was estimated.

7. GWA mapping

Average values for each trait following the transformation giving the least-skewed distribution were used in the mapping models. GWAS were performed using correction for population structure (PCoA) and modeling genetic variance with the kinship matrix (K). Two mixed models were implemented.

First, the bivariate multi-trait mixed model (MTMM) developed by Korte *et al.* (2012) to take into account the correlation structure of multi-environment datasets and increase the detection power was implemented. The MTMM approach includes two different tests: (1) The 'global test' compared a model including only the genotype effect to a null model to identify markers with common effect between watering regimes ('constitutive QTL'). (2) The 'G x W test' compared a full model to a model including only the genotype effect to identify markers with interactive effect between the watering conditions ('interactive QTL'). SNP with *P-value* below 10^{-4} were considered as significant. From each test, the percentage of variation explained by the marker (individual PVE for each significant marker) was computed.

Secondly, the univariate multi-locus mixed model (MLMM) developed by Segura *et al.* (2012) to increase the detection power for polygenic characters was used to identify associations for each trait under each watering regime ('specific QTL') and for the Δ values ('interactive QTL'). We implemented a new model selection criterion in the MLMM framework to allow

for a more permissive detection threshold to compromise between type I (false positive) and type II (false negative) errors, while limiting the number of cofactors selected to avoid overestimation of the *P-values* due to the relatively small size of the population. Models with a maximum of five cofactors having all a raw *P-value* below 10^{-4} were retained. From the optimal model selected, the percentages of variation explained by the selected markers (global PVE for all the significant markers) were computed for each trait.

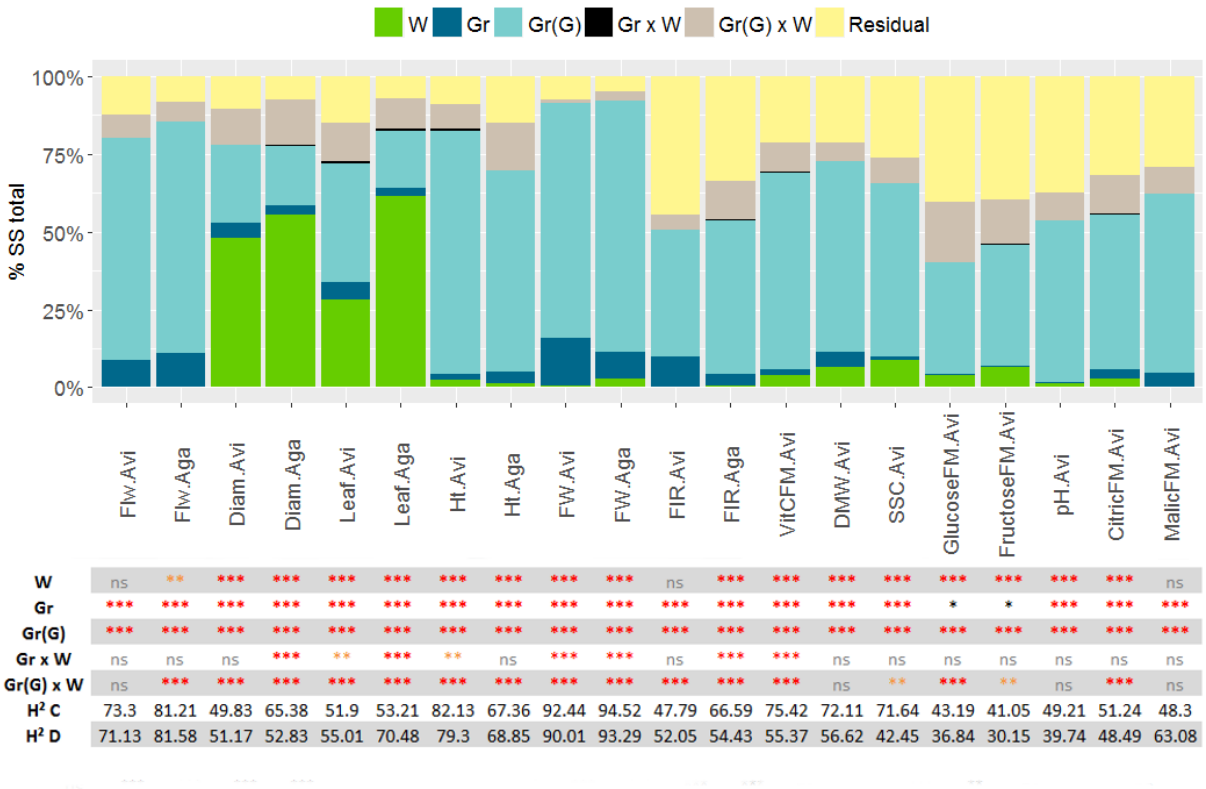
For all the QTLs identified, we computed phenotypic effects under both watering conditions as: (Minor allele mean − Major allele mean) / 2. Among the interactive QTLs, we distinguished between 'antagonist QTLs' (effect changing direction according to the watering regime) and 'differential QTLs' (effect changing intensity according to the watering regime).

8. Linkage disequilibrium estimation and confidence interval definition

To define intervals around QTLs, we used a strategy based on LD between pair of markers inspired from Cormier *et al.* (2014). We used the r^2 estimator implemented in the package *genetics* (Warnes and Leisch, 2012) to assess LD between marker pairs. First, we performed LD calculation between 100,000 randomly chosen pairs of unlinked loci (on different chromosomes). The 95^{th} percentile of the unlinked-r^2 distribution equal to 0.28 was considered as critical LD threshold. Then, for each significant marker, we computed LD with all the markers upstream and downstream on the same chromosome. We defined the lower (upper) boundary of the interval as the last marker downstream (upstream) on the chromosome that presented a LD with the significant marker above the 'critical LD' threshold. For the QTLs detected with the MTMM procedure, when two markers presented a LD higher than the LD threshold, we considered them as a unique QTL. The number of genes within each interval was identified from the tomato genome (ITAG2.4).

9. Comparison between linkage and association QTLs and identification of candidate genes

For the comparison with the QTLs detected in the RILs grown under the same conditions and phenotyped for the same traits (Albert *et al.*, 2016), we projected the QTLs detected in both populations onto the tomato genome (SL2.50). In the comparison, we considered related traits as a single one: pH, malic and citric acid contents were grouped as 'acids', as well as SSC, Glucose and Fructose contents as 'sugars'. Besides, whatever the QTL type ('interactive',



	FIw.Avi	FIw.Aga	Diam.Avi	Diam.Aga	Leaf.Avi	Leaf.Aga	Ht.Avi	Ht.Aga	FW.Avi	FW.Aga	FIR.Avi	FIR.Aga	VitCFM.Avi	DMW.Avi	SSC.Avi	GlucoseFM.Avi	FructoseFM.Avi	pH.Avi	CitricFM.Avi	MalicFM.Avi
W	ns	**	***	***	***	***	***	***	***	***	ns	***	***	***	***	***	***	***	***	ns
Gr	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	*	*	***	***	***
Gr(G)	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Gr x W	ns	ns	ns	***	**	***	**	ns	***	***	ns	***	***	ns	ns	ns	ns	ns	ns	ns
Gr(G) x W	ns	***	***	***	***	***	***	***	***	***	***	***	***	ns	**	***	**	ns	***	ns
H^2 C	73.3	81.21	49.83	65.38	51.9	53.21	82.13	67.36	92.44	94.52	47.79	66.59	75.42	72.11	71.64	43.19	41.05	49.21	51.24	48.3
H^2 D	71.13	81.58	51.17	52.83	55.01	70.48	79.3	68.85	90.01	93.29	52.05	54.43	55.37	56.62	42.45	36.84	30.15	39.74	48.49	63.08

Figure 1. Dissection of the total phenotypic variation. For each phenotypic trait, the top figure displays the proportion of each effect in the total sum of squares: green for watering regime (W); dark blue for genetic group (Gr); light blue for genotype nested in genetic group (Gr(G)); black for the interaction genetic group by watering regime (Gr x W); grey for the interaction genotype by watering regime (Gr(G) x W) and yellow for the residual. The table shows the significance of the *P-value* for the different effects: *** *P-value* below 0.001, ** between 0.001 and 0.01, * between 0.01 and 0.05 and ns above 0.05. 'H^2 C' and 'H^2 D' indicate the broad sense heritabilities in control and drought conditions, respectively.

'constitutive' or 'specific') and the location of the trial, we considered that a single QTL was present when the intervals overlapped between RIL and GWA QTLs. We then focused on the QTLs for Vitamin C, sugars and acids content including less than 100 genes to identify putative candidate genes with a reasonable confidence. Under those QTLs, we refined the set of candidates by selecting the genes expressed in tomato fruits according to gene expression data published by the Tomato Genome Consortium (2012). Then, we examined their functional annotations and focused on genes with annotations corresponding to related functions. Finally, we screened the polymorphism data obtained through the whole-genome resequencing of four accessions of our GWA population chosen to represent a large range of the molecular variability present in small fruit tomato (Causse *et al.*, 2013): Cervil (13.3x sequence depth), Criollo (8.1x), LA1420 (12.5x) and Plovdiv (12.2x). First, we considered the nucleotide variants with moderate (non-synonymous polymorphisms in coding regions) to high (modification of splice sites or start/stop codons) effect on the protein sequence (detected using SnpEff, Cingolani *et al.*, 2012). Then, the predicted impacts of the variants on the protein function were assessed using the web interfaces of *PROVEAN* (*http://provean.jcvi.org/seq_submit.php*) (Choi and Chan, 2015).

Results

1. Dissection of the phenotypic variations in the GWA population

In the variance analysis, the part of the total variation attributed to the genotype effect was predominant (35 to 80%, all *P-values* < 0.001) compared with the one attributed to the genetic group (0 to 15%, all *P-values* < 0.05) and the watering regime (0 to 28%, significant for 17 traits), except for leaf length in Agadir and stem diameter in Avignon and Agadir **(Figure 1 and Supplemental Table 5)**. For those vigor traits, the watering regime represented 48 to 61% of the total variation.

The genetic group by watering regime interactions represented less than 2% of the total sum of square for all traits and was non-significant for 12 traits. The eight significant traits were Diam.Aga, Leaf.Avi, Leaf.Aga, Ht.Avi, FW.Avi, FW.Aga, FIR.Aga and VitCFM.Avi. Tuckey's post-hoc test indicated that these interactions were mainly driven by a singular behavior of the SP group in response to water deficit **(Supplemental Figure 2)**. In contrast, the genotype by watering regime interaction represented 1 to 19% of the total variation and was significant for all traits, except Flw.Avi, DMW.Avi, pH.Avi and MalicFM.Avi. Interaction partitioning

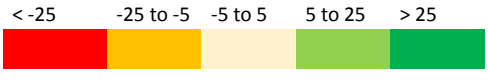
Table 1. Average relative difference between control and drought conditions for the fruit and plant traits measured in the GWA and RIL populations (%). The average relative differences were computed as: (Mean $_{Drought}$ – Mean $_{Control}$)/Mean $_{Control}$.

Plant traits	GWA	RIL[a]
Flw.Avi	0.0	-0.2
Flw.Aga	-0.6	+0.6
Diam.Avi	-27.5	-20.7
Diam.Aga	-37.4	-30.3
Leaf.Avi	-19.7	-13.4
Leaf.Aga	-31.8	-25.8
Ht.Avi	-5.1	-5.6
Ht.Aga	-4.2	+2.4
Nbfruits.Avi	-2.5	-21.7
Fruit traits		
FW.Avi	-19.0	-37.7
FW.Aga	-30.5	-25.4
FIR.Avi	-1.0	+3.4
FIR.Aga	+3.5	+0.8
VitCFM.Avi	+12.7	+26.3
VitCDM.Avi	+0.6	-8.9
DMW.Avi	+11.4	+30.7
SSC.Avi	+12.6	+26.3
GlucoseFM.Avi	+13.8	NA
FructoseFM.Avi	+17.7	NA
GlucoseDM.Avi	+0.5	NA
FructoseDM.Avi	+4.3	NA
pH.Avi	-1.3	-3.2
CitricFM.Avi	+10.7	NA
MalicFM.Avi	-3.6	NA
CitricDM.Avi	-1.2	NA
MalicDM.Avi	-14.8	NA
Yield.Avi	-18.8	-50.3

[a] : Data for the RIL population were reported in Albert *et al*. (2016)

NA: traits non-measured in the RIL population.

Color scale:



FM: metabolite concentrations expressed relatively to fresh matter

DM: metabolite concentrations expressed relatively to dry matter

according to method 1 from Muir *et al.* (1992) indicated that the genotype by watering regime interactions were mainly due to accessions re-ranking across watering regimes (80 – 100 %) and poorly to scale changes (0 – 20%, data not shown). The broad-sense heritabilities ranged from 30% for FructoseFM.Avi.D to 92% for FW.Avi.C. These values were correlated across watering regimes (r_{H^2} = 0.80), as well as the genetic variances ($r_{\sigma^2 G}$ = 0.99), confirming genotype re-ranking across watering regimes **(Figure 1** and **Supplemental Table 5).**

2. Impact of the water deficit on fruit quality and yield components

The RIL and GWA populations were grown in Avignon and Agadir in separate greenhouse trials over the years 2013 and 2014, while ensuring similar watering conditions (control and drought) (Albert *et al.* (2016) for details concerning the RILs). On average, in both locations, water deficit impacted plant and fruit traits in the same direction in the GWA and RIL populations with a decline in plant vigor, a decrease in yield and a higher concentration of the metabolites in fruits (in percentage of fresh matter) **(Table 1).** However, when applying the drought treatment, FW.Avi was decreased twofold and Nbfruits.Avi ninefold in the RILs compared with the GWA accessions. It resulted in a yield decrease reaching the level of -50% in the RILs against -20% in the GWA accessions. On the other hand, SSC, DMW and VitCFM were more strongly enhanced in the RILs (SSC: +26.3%, DMW: +30.7% and VitCFM: +26.3%) than in the GWA accessions (SSC: +12.6%, DMW: +11.4% and VitCFM: +12.7%).

The correlation between fruit fresh weight in control condition (indicator of fruit size) and ΔFW was strongly negative in the GWA accessions (Avi: r = -0.55, *P-value* = 2.70 x 10^{-12}, Aga: r = -0.52, *P-value* = 2.65 x 10^{-10}) as it was previously noticed in the RILs. This advocated greater FW loss in larger fruited accessions under drought and increased metabolite contents resulting mainly from reduced amount of water in the fruits. Thus, the differences observed between the populations may mostly reflect differences in fruit size, with larger fruits among the RILs (8 to 61 g, mean = 20 g, sd = 9 g) compared with the GWA accessions (2 to 46 g, mean = 13 g, sd = 10 g). Nevertheless, a larger range of variation was observed among the GWA accessions for ΔYield.Avi and ΔNbfruits.Avi compared with the RILs **(Figure 2, Supplemental Figures 3 to 4).** In particular, 55 accessions exhibited an increased yield under drought in the GWA population against only two among the RILs. No noticeable geographic origin or genetic group was obvious among these 55 accessions of the GWA population (10 mixture, 43 SLC and 2 SP).

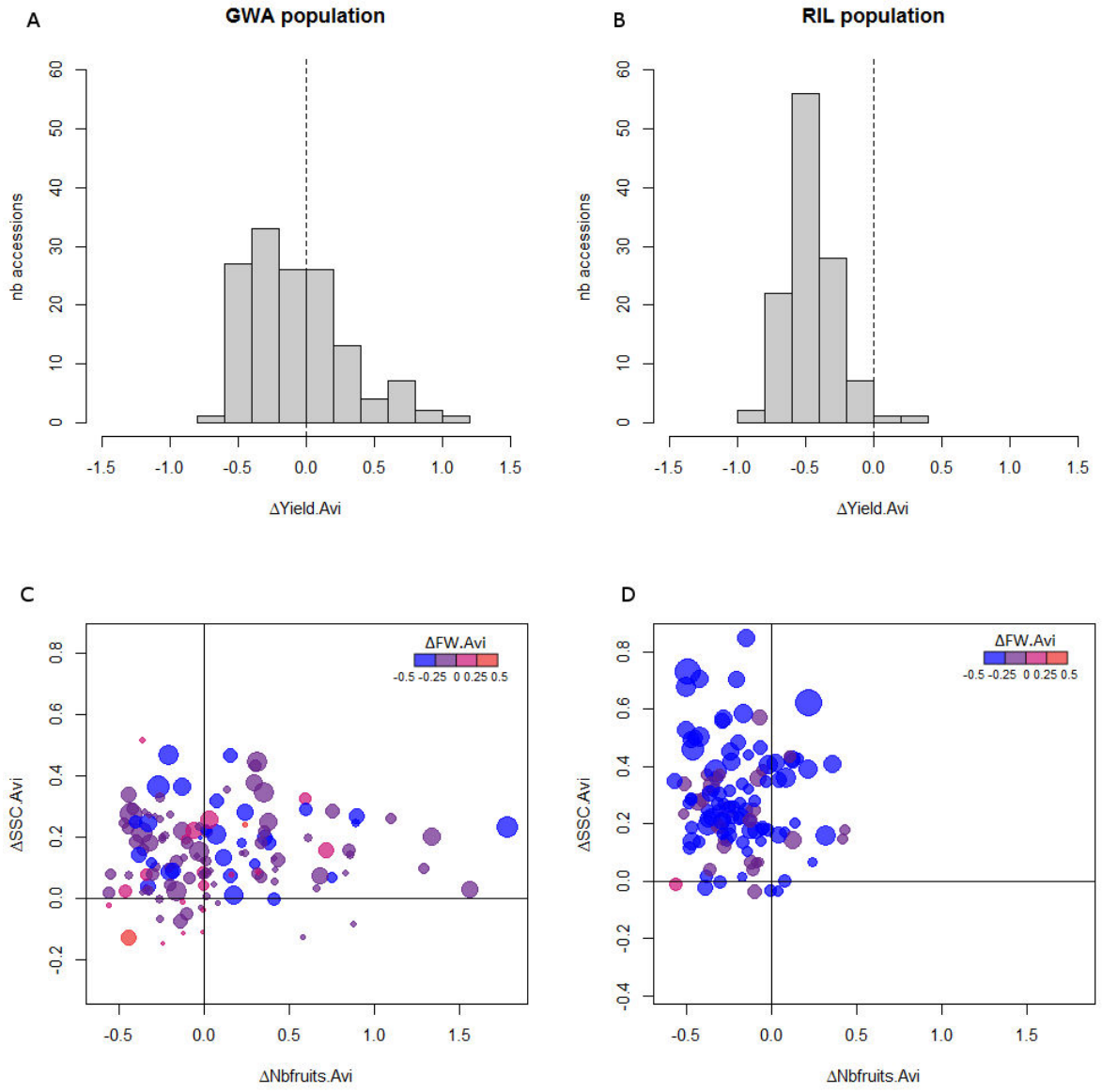


Figure 2. Impact of water deficit on yield, fruit number, fruit fresh weight (FW) and soluble solid content (SSC) in fruit. (A) and (B) Histograms of yield plasticity (ΔYield) in the GWA and RIL populations, respectively. (C) and (D) Relationship between plasticity of fruit number (ΔNbfruits) and plasticity of SSC (ΔSSC), in view of FW plasticity (ΔFW), in the GWA and RIL populations, respectively. In the bottom figures, the color scale indicates the variation in FW plasticity: blue for values below -0.5, purple for values between -0.25 and 0, magenta for values between 0 and 0.25 and red for values beyond 0.5. The size of the points is proportional to FW in control watering condition.

When plotting ΔNbfruits against ΔSSC in regard to fruit size and ΔFW.Avi, the RIL and GWA plants presented different patterns **(Figure 2).** Among the RILs, only 18 accessions were present in the top right quarter of the plot corresponding to accessions with increased SSC and Nbfruits under water deficit. Besides, all the top right quarter RILs had a negative ΔFW.Avi (blue and purple color) meaning a decreased fresh weight under drought compared with the control condition for these accessions. On the other hand, 40% of the GWA accessions were present in the top right quarter of the plot and six of them had a positive ΔFW.Avi (magenta and red color) and small to medium fruit size (FW in control from 2 to 28 g). Similar figures were obtained when considering fruit ascorbate **(Supplemental Figure 5),** malic and citric acid contents **(Supplemental Figure 6).**

3. QTL and QTL by watering regime interactions identified by association mapping

The multi-trait mixed model (MTMM) mapping approach detected 53 unique associations for 15 out of 27 phenotyped traits in the GWA population with *P-values* below 10^{-4} and percentages of variation explained varying from 5.45% to 18.22% (individual PVE per marker) **(Supplemental Table 6)**. A total of 49 associations were 'constitutive' irrespectively of the watering regime. Among these associations, the most significant were observed for malic acid content with *P-values* comprised between 2.40×10^{-6} and 1.33×10^{-13} in the global test (chromosomes 6 and 7) **(Supplemental Figure 7).** Four associations were declared 'interactive' between the watering regimes, two for Flw.Avi (chromosomes 9 and 11) and two for GlucoseDM.Avi (chromosomes 4 and 5), with *P-values* ranging from 1.48×10^{-5} to 7.04×10^{-5} (**Figure 3**).

The multi-locus mixed model (MLMM) approach identified a total of 124 associations (*P-values* < 1×10^{-4}) for the 27 studied phenotypic traits. Among them, 94 associations were '*specific*' (39 and 55 to drought and control conditions, respectively), 23 '*interactive*' (detected on Δ values) and 7 '*constitutive*' (detected under both conditions**; Supplemental Tables 7 and 8).** The explained percentages of phenotypic variation ranged from 8.16% (1 SNP for Leaf.Aga.C) to 63.85% (6 SNP for SSC.Avi.D) (global PVE for all the significant markers for a trait). Constitutive and/or specific associations were observed for all the traits. The most significant *P-values* were associated to MalicFM.Avi.D (S06_44955568: 1.88×10^{-19}), MalicDM.Avi.D (S06_44955568: 1.27×10^{-17}), pH.Avi (S04_66307772: 9.95×10^{-11}, **Figure 3**) and SSC.Avi.C (S10_64149793: 5.96×10^{-10}). The 23 interactive SNP were associated

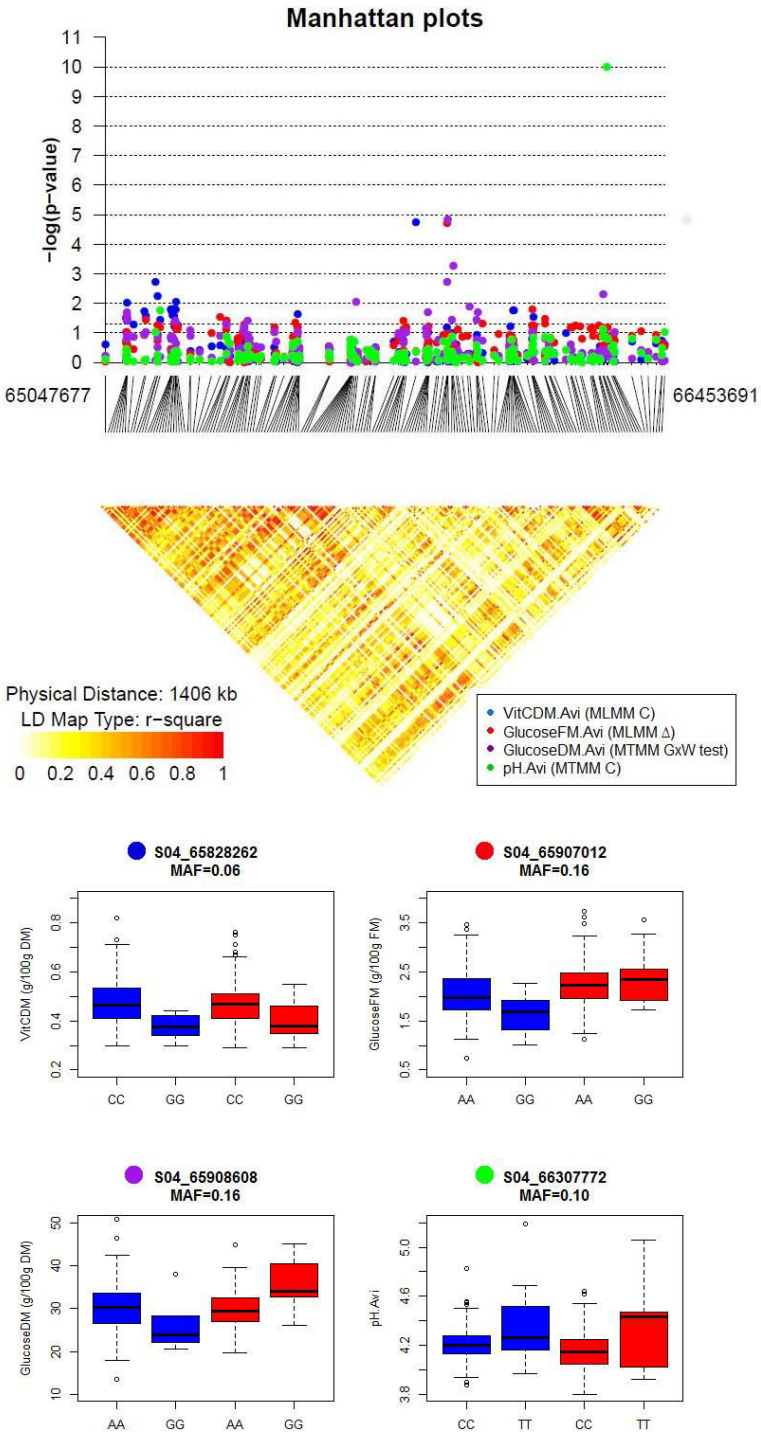


Figure 3. Focus on QTLs detected for fruit quality traits at the bottom of chromosome 4. (A) Manhattan plot displaying the –log10(P-values) (Y-axis) over genomic positions (X-axis) in a window of 1.46 Mbp corresponding to the common confidence interval of QTLs detected for VitCDM.Avi (MLMM control condition, blue), GlucoseDM.Avi (MTMM GxW test, purple), GlucoseFM.Avi (MLMM Δ, red) and pH.Avi (MLMM control, green) on chromosome 4 in the GWA population. P-values below 10-4 were considered as significant (4 in logit values). The pairwise LD heatmap was drawn using the R package 'snp.plotter' (Luna and Nicodemus, 2007). (B) Box-plot of the allelic effects for the four associated markers: S04_65828262 (VitCDM, 'control specific'), S04_65907012 (GlucoseFM, 'antagonist'), S04_65908608 (GlucoseDM, 'antagonsit') and S04_6630772 (pH, 'control specific'). Blue: Allelic effects under control condition. Red: Allelic effects under drought condition.

to 11 out of 27 traits. Their *P-values* ranged from 7.59 x 10^{-5} (ΔFlw.Avi: S06_36868039) to 2.75 x 10^{-11} (ΔFW.Aga: S11_50391249, **Supplemental Figure 8)**.

When gathering the associations obtained with MLMM and MTMM, 20 associations were detected in common (same trait and same QTL type) resulting in a total of 157 associations for the 27 traits **(Supplemental Tables 6 to 8).** Sixteen associations were detected between two and three times with related traits ('acid' and 'sugar' traits) and/or for the same trait in the two locations. Thus, a total of 141 different associations were identified, spread unevenly over the genome **(Table 2).** Chromosomes carried out six (chromosomes 7 and 8) to 23 associations (chromosome 2**; Supplemental Figure 7).** Thirty percent of the associations were '*constitutive*' (44/141), 30% were '*control specific*' (41/141), 22% were '*drought specific*' (31/141) and 17% were '*interactive*' (25/141). Among the interactive associations, 16 showed '*differential*' effects (effect intensity changing according to watering regime) whereas nine presented '*antagonist*' effects (effect direction changing according to watering regime). Up to 14, 24 and 28 different associations were mapped for vitamin C, 'acid' and 'sugar' content in fruit, respectively.

4. Confidence intervals and candidate gene selection under QTLs for fruit quality traits

We observed large differences in size and number of underlying genes when drawing confidence intervals around the association peaks. Eighteen QTLs mapped around the weakly recombinant centromeres covered more than 10 Mbp and included between 410 and 2,573 genes, whereas 84 QTLs covered less than 5.5 Mbp and encompassed between 1 and 97 genes **(Supplemental Figure 9)**. In the RILs grown in the same conditions (Albert *et al.*, 2016), only four QTLs covered less than 100 genes on a total of 56 QTLs. The comparison of the QTL positions between the RIL and GWA populations resulted in a total of 11 common QTLs to both populations **(Table 2)**, whereas 45 were specific to the RILs and 130 to the GWA population **(Supplemental Figure 10)**.

To propose putative candidate genes, we focused on QTLs for vitamin C, sugars and acids content in fruit including less than 100 genes (42 among 66 QTLs) and selected in their intervals genes showing expression in the fruits according to the Tomato Genome Consortium data (2012). This reduced the gene list to screen for between one and 87 genes depending on the QTL intervals. Annotations were analyzed to identify genes with functions related to vitamin C, sugar or acid metabolism under 'constitutive' QTLs and functions

related to primary metabolism and/or defense against abiotic stress under 'specific' and 'interactive' QTLs. A total of 41 putative candidates were proposed for three 'constitutive' QTLs **(Table 3)** and 15 'interactive' or 'specific' QTLs **(Table 4).** Among them, 22 presented polymorphisms inducing changes at the gene sequence level when comparing the four accessions of our GWA population that were re-sequenced in Causse *et al.* (2013). For four genes, these polymorphisms were predicted to have an impact on the biological function of a protein at the amino-acid level.

From the 18 dissected QTLs, '*SSC.Avi_9.1'* (control specific) likely corresponded to the cloned QTL '*Brix9.2.5*' controlling SSC in fruit and associated to a polymorphism in a cell wall invertase gene (Solyc09g010080: Lin5) (Fridman *et al.*, 2000) **(Table 4)**. A second QTL ('*Malic.Avi_6.3')* colocalized with a previously mapped QTL for acid content in fruit in different tomato populations and for which two '*aluminum-activated malate transporter-like*' genes (Solyc06g072910 and Solyc06g072920) were pointed out as putative candidate genes by Sauvage *et al.* (2014) **(Table 3)**. Although these two genes presented promising polymorphisms between our four re-sequenced accessions, they displayed a very low expression in fruit (Tomato Genome Consortium, 2012 and personal data) and will need further validation to be clearly associated to the phenotypes.

Ten QTLs colocalized with loci identified in the RILs (Albert *et al.*, 2016, control and drought conditions) and/or in the three tomato population analyzed by Pascual *et al.* (2016: RIL, GWA and MAGIC, control conditions) but for which no candidate gene was proposed until now, while six were present in genomic regions where to the best of our knowledge no QTLs for related traits were mapped thus far. In the intervals of four of them, controlling vitamin C and fructose content in a drought specific manner ('*VitCDM.Avi_1.1'*, '*FructoseDM.Avi_4.1*' and '*FructoseDM.Avi_10.1*'), three genes coding for '*chaperone proteins dnaJ*' were identified (Solyc01g105340, Solyc04g009770 and Solyc10g078560; **Table 4).** Five more genes coding for '*heat/cold shock proteins*' (Solyc01g111280, Solyc01g111300, Solyc01g111750, Solyc04g011440 and Solyc04g011450) were identified under antagonist and drought specific QTLs for fructose and malic acid content ('*FructoseDM.Avi_1.1*' and '*MalicDM.Avi_4.1*', **Table 4).**

Table 2. Description of QTLs detected for plant and fruit traits in the GWA population through association mapping and comparison with the one detected in the RIL population through linkage analysis. QTLs detected in the GWA population were classified according to their type. QTLs significant under both watering regimes are referred as 'constitutive'. QTLs significant under one watering regime only ('control' or 'drought') are designated as 'specific'. QTLs detected with the plasticity data and/or with the interaction test are designated as 'interactive'. For each phenotypic trait and each QTL type, number of QTL, minimum and maximum confidence interval (CI in Mbp on genome assembly v2.5) and minimum and maximum number of genes in the interval are displayed. We considered related traits as a single one: pH, acid malic (DM and FM) and acid citric (DM and FM) were grouped in 'acids', as well as SSC, Glucose (DM and FM) and Fructose (DM and FM) in 'sugars'. We gathered QTLs detected in the both trial locations (Agadir and Avignon) for a same trait. For the comparison with the RIL population (Albert *et al*. 2015), whatever the QTL type, we considered that a single QTL was present when the CI overlapped between RIL and GWA QTLs.

Trait		Constitutive QTL					Specific QTL										Interactive QTL					
							Control					Drought										
	Nb QTL total	Nb	Chr.	Min - Max CI (Mpb)	Min - Max Nb genes	Com. RIL	Nb	LG	Min - Max CI (Mpb)	Min - Max Nb genes	Com. RIL	Nb	LG	Min - Max CI (Mpb)	Min - Max Nb genes	Com. RIL	Nb ant.	Nb diff.	LG	Min - Max CI (Mpb)	Min - Max Nb genes	Com. RIL
Plant traits																						
Flw	10	2	1 ; 12	0.08 - 0.94	17 - 117	0	1	3	0.33	30	0	3	4 ; 5 ; 10	0.00 - 59.94	1 - 1653	1	0	4*	1 ; 6 ; 9 ; 11	0.06 - 16.69	7 - 794	0
Diam	14	1	10	2.56	336	0	4	2 ; 5 ; 6 ; 11	0.14 - 3.64	16 - 500	0	4	2 ; 4 ; 9 ; 12	0.02 - 36.64	2 - 600	1	3	2	2 ; 5 ; 6	0.02 - 12.80	2 - 516	0
Leaf	12	6	1 ; 2 ; 3 ; 11	0.03 – 45.30	1 - 1147	1	2	2 ; 4	3.50 - 4.86	232 - 463	1	3	1 ; 2	0.08 - 9.03	6 - 284	0	0	1	8	1.60	96	0
Ht	8	3	1 ; 2 ; 3	0.22 - 32.22	10 - 720	0	4§	2 ; 3 ; 7 ; 9	0.10 - 5.34	14 - 140	0	1	12	0.63	31	0	0	0	-	-	-	0
Nbfruits	7	0	-	-	-	0	4	4 ; 7 ; 11 ; 12	0.34 - 50.43	40 - 741	0	3	9 ; 10	1.28 - 7.54	97 - 819	0	0	0	-	-	-	0
Fruit traits																						
FW	6	2§	2 ; 3	0.07 - 1.87	6 - 250	0	0	-	-	-	0	1	2	31.33	677	0	1	2	1 ; 10 ; 11	0.23 - 56.14	17 - 1287	1
FIR	15	6§	1 ; 2 ; 5 ; 6 ; 11	0.02 - 32.56	2 - 858	0	7	1 ; 2 ; 3 ; 5 ; 9 ; 12	0.04 - 48.91	5 - 928	0	2	4 ; 10	0.04 - 65.29	4 - 2573	1	0	0	-	-	-	0
VitC	14	4	8 ; 9 ; 10 ; 11	0.15 - 8.50	18 - 899	0	5	4 ; 7 ; 9 ; 12	0.27 - 41.12	18 - 796	0	4	1 ; 2 ; 4; 11	0.08 - 4.30	7 - 494	1	0	1	10	0.26	39	0
DMW	2	0	-	-	-	0	2	4 ; 9	0.01 - 0.93	2 - 137	0	0	-	-	-	0	0	0	-	-	-	0
Sugars	28	9¤	4 ; 5 ; 7 ; 8 ; 9 ; 10 ; 11	0.07 - 59.37	18 - 1602	0	5	3 ; 9 ; 11	0.01 - 3.34	2 - 417	2	7	1 ; 4; 6 ; 10 ; 11	0.03 - 2.62	5 - 327	0	2	5	1 ; 2 ; 4 ; 5 ; 11	0.04 - 59.37	8 - 1602	1
Acids	24	11 ¤	5 ; 6 ; 7 ; 8 ; 9	0.00 - 2.09	1 - 289	1	6	2 ; 3 ; 4 ; 6 ; 11	0.00 - 0.93	1 - 137	0	3	1 ; 4 ; 6	0.04 - 0.47	6 - 31	0	3	1	2 ; 4 ; 10 ; 11	0.09 - 64.21	10 - 2427	0
Yield	1	0	-	-	-	0	1	1	0.35	49	0	0	-	-	-	0	0	0	-	-	-	0
Total	**141**	**44**				2	**41**				3	**31**				4	**9**	**16**				2
Com. RIL	11																					

* Indication of interactive QTLs confirmed with both methods plasticity data and interaction test.

§ Indication of QTLs confirmed in both locations, Agadir and Avignon (with same type: 'constitutive', 'specific' or 'interactive').

¤ Indication of QTLs for acids and sugars confirmed with several measurement methods (pH and acid content, SSC and sugar content).

Three constitutive QTLs, the first two on chromosome 7 controlling glucose and malic acid content and the third on chromosome 10 controlling fructose content, seemed particularly promising. The first two ('GlucoseDM.Avi_7.2' and 'Malic.Avi.7_2' in **Table 3**) shared a common interval including a gene coding for a '*phosphoenolpyruvate carboxylase*' (Solyc07g062530: PEPC) and a gene coding for a '*malate deshydrogenase*' (Solyc07g062650). The PEPC gene presented a non-synonymous polymorphism with a predicted impact on the protein function when comparing the four re-sequenced accessions. The third one ('FructoseDM.Avi_10.2' in **Table 3**) contained two genes coding for '*cell wall invertases*', Lin6 (Solyc10g083290) and Lin8 (Solyc10g083300), presenting three non-synonymous polymorphisms between the re-sequenced accessions.

Discussion

To assess the extent of natural variation in tomato responses to water deficit, we phenotyped a collection of 141 small fruit accessions for plant and fruit traits, under control and drought conditions. Using 6,100 SNP genotyped over the genome, we achieved association mapping using univariate and bivariate mixed models. QTLs, QTL by watering regime interactions and putative candidate genes were identified. This study, in combination with the results reported in RILs grown under the same watering conditions, contributed to a first detailed characterization of the genetic variations and genomic determinants of response to water deficit in tomato.

1. Improving fruit quality while maintaining yield in tomato under water limitation

Deficit irrigation strategies aiming to reduce non-beneficial water consumption while maximizing fruit quality and minimizing yield losses are studied in horticultural production to simultaneously address environmental issues and market expectations. It seems particularly relevant for tomato since consumers complain about lack of taste in the new varieties (Bruhn *et al.*, 1991; Causse *et al.*, 2010). In our trials, after a decrease in 60 % of the water supply throughout the plant growth, we observed on average reduced plant vigor and yield, while fruit quality was improved or stable depending whether metabolite concentrations were expressed relatively to fresh or dry matter. This antagonistic relationship between quality and yield performances confirmed the results obtained in RILs (Albert *et al.*, 2016)

Table 3. Putative candidate genes in the confidence interval around constitutive GWA QTLs for vitamin C, sugar and acid content in fruit. We focused on QTLs encompassing less than 100 genes. Comparisons with the QTLs detected in Albert *et al*. (2016) (RIL under control and drought conditions) and Pascual *et al*. (2016) (MAGIC, RIL and GWA populations under control conditions) for related traits are indicated. For each QTL, significant marker(s), confidence interval (CI), number of genes in the interval and among them number of genes which are expressed in the tomato fruits according to gene expression data published by the Tomato Genome Consortium (2012) are indicated. Putative candidate genes are proposed on the basis of their expression in the fruit, their functional annotation and the scientific literature. 'Variants' displays the number of moderate (non-synonymous polymorphisms in coding regions) to high (modification of splice sites or start/stop codons) effect polymorphisms identified from the resequencing of four accessions of the GWA population (Causse *et al*., 2013). Variants which have a deleterious impact on the protein structure according to PROVEAN are indicated by '#'.

QTL(s)*	QTL type	Coloc. Albert *et al*. (2016) and Pascual *et al*. (2015)	Marker(s)	CI (Mbp)	Nb genes	Nb genes expressed in fruit	Putative candidate genes and annotations	Related functions	Non-syn. variants
MalicDM.Avi_6.3; MalicFM.Avi_6.3	C & D	MAGIC + GWA	S06_44955568	44.92 – 44.96	8	5	*Solyc06g072910: Aluminum-activated malate transporter-like* **	Carbon metabolism and malate compartmentation (Martinoia and Rentsch, 1995; Sauvage *et al*., 2014)	1
							Solyc06g072920: Aluminum-activated malate transporter-like **		2
GlucoseDM.Avi_7.1; MalicDM.Avi_7.2; MalicFM.Avi_7.2	C & D	MAGIC + GWA	S07_64878195; S07_65079667	64.86 –65.60	97	87	Solyc07g062530: Phosphoenolpyruvate carboxylase 2	Malic and citric acid accumulation (Guillet *et al.*, 2002)	**1#**
							Solyc07g062650: Malate dehydrogenase	Carbon metabolism and malate compartmentation (Martinoia and	0
FructoseDM.Avi_10.2	C & D	NO	S10_63163119	63.10 – 63.24	18	16	Solyc10g083290: Beta-fructofuranosidase insoluble isoenzyme 2 (Lin6)	Sugar metabolism (Fridman *et al.*, 2004*a*; Proels and Roitsch, 2009; Ruan *et al.*, 2010*a*; Li *et al.*, 2012)	1
							Solyc10g083300: Beta-fructofuranosidase insoluble isoenzyme 2 (Lin8)		2

* QTL names make reference to the map representation in Supplemental Figure 7. They are highlighted when they were identified with *P-values* below 10^{-5}.

*** Genes poorly expressed in the fruit.*

and the tendencies reported by others authors in tomato (Guichard *et al.*, 2001; Kirda *et al.*, 2004; Zheng *et al.*, 2013), peach (Mirás-Avalos *et al.*, 2013) or grapevine (Santesteban and Royo, 2006).

Nevertheless, fifty accessions (with small to medium fruit size) had both improved fruit quality and maintained yield (or even improved) under water deficit compared with the control watering regime, although their vigor (measured through leaf length and stem diameter) was decreased. These accessions emphasized the opportunity to increase metabolite content in tomato fruits using deficit irrigation without achieving parallel limitation of the yield. On the opposite, not any RIL presented such response to the water deficit treatment and the increased sugar and acid contents observed reflected mainly concentration effects due to a decreased amount of water in fruit (Albert *et al.*, 2016).

The large phenotypic variations observed mainly resulted from genotype effects (35 to 80%) and less from genotype by watering regime interactions (1 to 19%). The watering regime effect represented a significant part of the total phenotypic variability (up to 40%) only for stem diameter and leaf length. This suggests that tomato plants buffer the negative effect of water limitation by limiting their vegetative growth and reallocating the photo-assimilates to the fruits (Lemoine *et al.*, 2013; Osorio *et al.*, 2014).

2. Benefits and limits of GWA to dissect the genetic architecture of response to water deficit in tomato

Association studies aiming to identify alleles whose effects are modulated by environmental conditions are still few in plants. To date, such studies were only reported in *Arabidopsis thaliana* (Li *et al.*, 2010; Morrison and Linder, 2014; Sasaki *et al.*, 2015; El-Soda *et al.*, 2015) and maize (Saïdou *et al.*, 2014). Explicitly accounting for 'QTL by environment interactions' in QTL studies can help to discover novel genes that act synergistically with the environment, potentially leading to the identification of superior genotypes according to the environments (Des Marais *et al.*, 2013).

Table 4. Putative candidate genes in the confidence interval around specific and interactive GWA QTLs for vitamin C, sugar and acid content in fruit. We focused on QTLs encompassing less than 100 genes. Comparisons with the QTLs detected in Albert *et al*. (2016) (RIL under control and drought conditions) and Pascual *et al*. (2016) (MAGIC, RIL and GWA populations under control conditions) for related traits are indicated. For each QTL, significant marker(s), confidence interval (CI), number of genes in the interval and among them number of genes which are expressed in the tomato fruits according to gene expression data published by the Tomato Genome Consortium (2012) are indicated. Putative candidate genes are proposed on the basis of their expression in the fruits, their functional annotation and the scientific literature. 'Variants' displays the number of moderate (non-synonymous polymorphisms in coding regions) to high (modification of splice sites or start/stop codons) effect polymorphisms identified from the resequencing of four accessions of the GWA population (Causse *et al*., 2013). Variants which have a deleterious impact on the protein structure according to PROVEAN are indicated by '#'.

QTL(s)*	QTL type	Coloc. Albert *et al*. (2016) and Pascual *et*	Marker(s)	CI (Mbp)	Nb genes	Nb genes expressed in fruit	Putative candidate genes and annotations	Related functions	Non-syn. variants
CitricDM.Avi_1.1	D	RIL	S01_86174739	86.15 – 86.20	6	6	Solyc01g094720: Vesicular glutamate transporter	Nitrogen transporter (Rentsch *et al.*, 2007)	1
VitCDM.Avi_1.1	D	NO	S01_93702068	93.47 – 93.76	42	36	Solyc01g105340: Chaperone protein dnaJ	Protein protection (Wang *et al.*, 2014)	0
							Solyc01g105540: 2-oxoglutarate/malate translocator	Carbon metabolism and malate compartmentation (Martinoia and Rentsch, 1995)	0
							Solyc01g105630: Calmodulin	Osmotic adjustment and stress signaling in interaction with cellular calcium (Perruc *et al.*, 2004; Reddy *et al.*, 2011)	1
SSC_Avi_1.2	D	MAGIC	S01_96226845	96.22 – 96.25	7	5	Solyc01g109220: Mitochondrial import receptor	Oxidative stress (Frank *et al.*, 2007)	**1#**
FructoseDM.Avi_1.1	ant.	MAGIC	S01_97877551	97.43 – 97.99	79	61	Solyc01g111280: Cold shock protein-1	Protein protection under salt and drought stress (Kim *et al.*, 2013)	2
							Solyc01g111300: Cold shock protein-1		0
							Solyc01g111320: Thaumatin-like protein	Sweet-tasting protein, sugar accumulation and plant defense (Kim *et al.*, 2002; Petre *et al.*, 2011)	0
							Solyc01g111330: Thaumatin-like protein		0
							Solyc01g111510: Ascorbate peroxidase	Oxidative stress (Pignocchi *et al.*, 2006)	0
							Solyc01g111630: Glyoxylate/hydroxypyruvate reductase B	Recycling fatty acids into glucose (Cornah *et al.*, 2004)	0
							Solyc01g111680: Ubiquitin-conjugating enzyme 22	Osmotic adjustment and oxidative stress response (Zhou *et al.*, 2010)	5
							Solyc01g111660: Aquaporin-like protein	Water and solute transport, osmotic adjustment (Reuscher *et al.*, 2013; Ricardi *et al.*, 2014*a*)	0
							Solyc01g111750: Heat shock protein dnaJ	Oxidative stress, fruit maturation (Banzet *et al.*, 1998; Neta-Sharir *et al.*, 2005)	0

We identified a total of 141 QTLs with low to medium effects. The phenotyped traits were strongly polygenic and justified the use of a multi-locus GWA mapping model (MLMM: Segura *et al.*, 2012). In particular, up to 14, 24 and 28 different QTLs were identified for vitamin C, acid and sugar content, respectively. Among the loci identified, 51% were specific to one watering condition, 31% were constitutive and detected whatever the condition and 18% were interactive between the watering conditions. These proportions of QTL types are relatively similar to those reported in the RILs grown in the same conditions (Albert *et al.*, 2016) and in the study of Gur *et al.* (2011) on tomato introgression lines. However, while most of the interactive QTLs identified in the RILs presented antagonist effects, a majority of differential effects was observed in the GWA study. These discrepancies between both populations may reflect their different genetic basis: the RILs segregate between a small and a large fruited accession, whereas the GWA collection focuses on the polymorphisms between several diverse small-fruited accessions.

Because of the large number of markers to be used in GWA analysis, it is not straightforward to choose an appropriate significance threshold controlling for false positives while maintaining the statistical power. We thus opted for a lowered threshold of 10^{-4}. If we used Bonferroni correction usually applied to exclude false positives, we should have used a significance threshold of 10^{-5}. This would reduce the number of associations detected to 69 (9 'interactive', 44 'specific' and 16 'constitutive'). With this stringent threshold, we would not have recovered some well described tomato QTLs, as for example FW11.2 and FW11.3 on chromosome 11 (fruit fresh weight QTLs: Huang & van der Knaap, 2011; Illa-Berenguer *et al.*, 2015). The necessity of more permissive thresholds in GWAS is often claimed. Strategies based on enrichment tests using known candidate genes from the literature to evaluate the false positive rate and choose the appropriate threshold values are proposed (Atwell *et al.*, 2010; Sasaki *et al.*, 2015). However, these approaches are limited to well annotate model genomes and simple traits with already well described genetic architecture. Another solution to solve the multiple testing issues could be to use haplotypes instead of individual markers to minimize the number of tests, especially in species where the LD spans large genomic regions (Bader, 2001; McClurg *et al.*, 2006). This has already been successfully applied in crops (Gawenda *et al.*, 2015) and would be worth to be tested in tomato, but may need more markers to identify haplotypes correctly.

Following Table 4.

SSC.Avi_2.2	dif.	MAGIC + RIL + GWA	S02_40059311	40.02 – 40.11	8	7	Solyc02g070270: Amino acid transporter	Transport	6
							Solyc02g070280: Amino acid transporter	Transport	0
							Solyc02g070290: Potassium/chloride transporter	Transport	0
pH.Avi_2.3	dif.	MAGIC + RIL + GWA	S02_49491595	49.40 – 49.50	10	9	Solyc02g086820: Carbonic anhydrase	Enhanced photosynthesis under drought (Gu *et al.*, 2013)	0
FructoseDM.Avi_4.1	D	NO	S04_03214865	3.05 – 3.22	20	18	Solyc04g009770: DNAJ chaperone	Protein protection (Wang *et al.*, 2014)	1
							Solyc04g009830: Stress responsive gene	Gene regulation under abiotic stress (Chen *et al.*,	0
MalicDM.Avi_4.1	D	NO	S04_03821452	3.64 – 4.10	31	26	Solyc04g010330: Auxin-regulated protein	Abiotic stress signaling (Bianchi *et al.*, 2002*a*; Gong *et al.*, 2010)	**3#**
							Solyc04g011440: Heat shock protein	Oxidative stress, fruit maturation (Banzet *et al.*, 1998; Neta-Sharir *et al.*, 2005)	0
							Solyc04g011450: Heat shock cognate		0
GlucoseFM.Avi_6.1	D	NO	S06_38712034	38.34 – 38.73	36	29	Solyc06g060360: Universal stress protein	Gene regulation under abiotic stress (Chen *et al.*, 2011)	2
							Solyc06g060370: Organic anion transporter	Metabolism	0
							Solyc06g060620: Nitrate transporter	Nitrogen transport (Rentsch *et al.*, 2007)	1
FructoseFM.Avi_6.1	D	NO	S06_42161946	42.00 – 42.20	28	19	Solyc06g066820: Gibberellin 3-beta-hydroxylase	Water status and reduced transpiration (Nir *et al.*, 2014)	2
VitCFM.Avi_7.1	C	NO	S07_02439123	2.29 – 2.56	25	24	Solyc07g007790: Sucrose phosphate synthase	Sugar compartmentation, sink strength (Nguyen-Quoc and Foyer, 2001)	1
SSC.Avi_9.1	C	MAGIC + RIL + GWA	S09_03477979	0.34 – 0.35	2	2	Solyc09g010080: Beta-fructofuranosidase, insoluble isoenzyme 1 (Lin5)	Sugar metabolism, heat and drought tolerance (Fridman *et al.*, 2004*b*; Zanor *et al.*, 2009*a*; Ruan *et al.*, 2010*b*; Li *et al.*, 2012)	4
							Solyc09g010090: Beta-fructofuranosidase insoluble isoenzyme 2 (Lin7)		1
VitCFM.Avi_10.1	dif.	NO	S10_00934508	0.08 – 0.11	39	38	Solyc10g006130: Ethylene responsive TrF	Abiotic stress signaling (Pan *et al.*, 2012)	1
FructoseDM.Avi_10.1	D	NO	S10_60291460	60.12 – 60.37	31	30	Solyc10g078370: Auxin efflux carrier	Abiotic stress signaling (Bianchi *et al.*, 2002*b*; Gong *et al.*, 2010)	0
							Solyc10g078490: Aquaporin	Water and solute transport, osmotic adjustment (Reuscher *et al.*, 2013; Ricardi *et al.*, 2014*b*)	0
							Solyc10g078560: Chaperone protein dnaJ	Protein protection (Wang *et al.*, 2014)	**1#**
FructoseFM.Avi_11.3	D	MAGIC	S11_52838456	52.80 – 52.84	5	5	Solyc11g067050: Neutral invertase	Sugar metabolism, heat and drought tolerance (Ruan *et al.*, 2010*b*; Li *et al.*, 2012)	1

* QTL names make reference to the map representation in Supplemental Figure 7. They are highlighted when they were identified with *P-values* below 10^{-5}.

The projection of the QTL intervals onto the physical map of tomato allowed the comparison of QTL positions between the RIL and GWA population even though they were genotyped with different markers. This projection resulted in a total of 11 QTLs conserved between both populations. On the other hand, 45 were specific to the RIL population and 130 to the GWA population. This may seem a relatively small number of common QTLs between the populations, but the RIL parental accessions reflected only a limited fraction of the genetic variation present in the GWA population.

3. Searching for candidate genes under QTLs for fruit quality traits

Our approach, combining linkage and association mapping was powerful in recovering previously identified loci associated to fruit quality. As an example, we mapped a QTL associated with fruit fructose content on chromosome 9 which included in its interval the gene Lin5 (Solyc09g010080) known to encode a cell wall invertase affecting tomato fruit sugar content (Fridman *et al.*, 2000). Apart from recovering previously described genes, we identified QTLs in genomic regions where QTLs associated to related trait were previously identified in other populations but for which no candidate gene was proposed until now (probably because of too large confidence intervals) or in genomic regions where to the best of our knowledge no QTL was reported for related traits thus far. The confidence intervals around the association peaks obtained using a LD based approach were mostly shorter (1 to 97 genes for 84 intervals) compared with the intervals obtained using the RILs or introgression lines (Semel *et al.*, 2007; Gur *et al.*, 2011; Arms *et al.*, 2015).

Combining publicly available expression data (Tomato Genome Consortium, 2012), exonic variants gained from re-sequencing of four accessions of the GWA collection (Causse *et al.*, 2013) and functional analysis of the gene annotations in the confidence intervals, we proposed 41 putative candidate genes under three constitutive QTLs and 15 interactive or specific QTLs. Under the interactive and specific QTLs, genes related to protein protection (chaperone and heat/cold shock proteins), water and solute transport (aquaporins and others transporters), sugar metabolism (sucrose phosphate synthase and invertases) and hormonal signaling (auxin, gibberellin and ethylene) were identified and may play a crucial role in responses to water deficit (Wang *et al.*, 2003; Shinozaki and Yamaguchi-Shinozaki, 2007). Some of them presented polymorphisms with predicted impacts on the protein

function when comparing the re-sequenced accessions and constitute promising targets for future functional validations.

On bottom of chromosome 7, two QTLs, controlling glucose and malic acid content, shared a common interval including a gene coding for a 'phosphoenolpyruvate carboxylase' (PEPC) and a gene coding for a 'malate dehydrogenase'. The PEPC gene presented a non-synonymous polymorphism with a predicted impact on the protein function in the four re-sequenced accessions. As the PEPC is catalyzing the carboxylation of the phosphoenolpyruvate arising from the glycolysis into oxaloacetate which is then converted into malate by the malate dehydrogenase or enters the Krebs cycle (Guillet *et al.*, 2002), this gene constitutes a likely candidate. Nevertheless, although if the 'malate dehydrogenase' gene did not present any exonic SNP in our data, it remains an interesting candidate as our four re-sequenced accessions probably did not represent the full genetic diversity present in the GWA population and the phenotypic variations observed may result from regulation change more than modifications of the protein. On bottom of chromosome 10, a QTL interval controlling fructose content contained two genes coding for 'cell wall invertases' (Lin6 and Lin8). Both genes presented non-synonymous polymorphisms between the re-sequenced accessions. Contrary to Lin5 on chromosome 9, Lin6 and Lin8 have not yet been associated to variation in sugar content in fruit. Cell wall invertases are extracellular hydrolases which cleave sucrose to glucose and fructose, which are then transported into the cell. They play a central role in regulating, amplifying and integrating different signals that lead to the source-sink transition in plants.

Subsequent analyses based on either fine mapping of the candidate genes using target re-sequencing approach or functional validation, for example by genome editing, could clarify the involvement of these genes in the phenotypic variations observed.

Acknowledgements

We acknowledge the experimental teams of UR-GAFL and Gautier SEMENCES for their collaboration in implementing the experimentations. We especially thank Yolande Carretero, Esther Pelpoir, Romain Novaretti, Doriane Bancel and the employees of “Domaine Margau” (Agadir) for their help in phenotyping. Thanks to Christopher Sauvage for proof reading and script sharing. The CTPS project TOMSEC supported this work. E.A. was supported by an INRA PhD fellowship.

Author contributions

E.A. conducted experiments in France, analyzed data and wrote the manuscript. V.S. developed scripts for the GWA mapping. J.G. sampled and collected phenotypic data in France. J.B. and L.D. supervised sample collection and phenotypic measurements in Morocco. M.C. supervised the project, built the experimental design and revised the manuscript. All authors discussed the results and commented the manuscript. Authors declared no conflict of interest in the authorship and publication of this document.

Supplemental Material

Supplemental figures 1 to 10 and supplemental table 4 are available at the end of the manuscript in Appendix 5. Supplemental tables 1 to 3 and 5 to 8 are available online on the editor website.

References

Albacete A, Martínez-Andújar C, Pérez-Alfocea F. 2014. Hormonal and metabolic regulation of source-sink relations under salinity and drought: From plant survival to crop yield stability. Biotechnology Advances **32**, 12–30.

Albert E, Gricourt J, Bertin N, Bonnefoi J, Pateyron S, Tamby J-P, Bitton F, Causse M. 2016. Genotype by watering regime interaction in cultivated tomato: lessons from linkage mapping and gene expression. Theoretical and Applied Genetics **129**, 395–418.

Arms EM, Bloom AJ, St Clair DA. 2015. High-resolution mapping of a major effect QTL from wild tomato Solanum habrochaites that influences water relations under root chilling. Theoretical and Applied Genetics **128**, 1713–1724.

Atwell S, Huang YS, Vilhjálmsson BJ, *et al.* 2010. Genome-wide association study of 107 phenotypes in a common set of *Arabidopsis thaliana* inbred lines. Nature **465**, 627–631.

Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics **23**, 1294–1296.

Bac-Molenaar JA, Granier C, Keurentjes JJB, Vreugdenhil D. 2016. Genome-wide association mapping of time-dependent growth responses to moderate drought stress in *Arabidopsis*. Plant, Cell & Environment **39**, 88–102.

Bader JS. 2001. The relative power of SNPs and haplotype as genetic markers for association tests. Pharmacogenomics **2**, 11–24.

Banzet N, Richaud C, Deveaux Y, Kazmaier M, Gagnon J, Triantaphylidès C. 1998. Accumulation of small heat shock proteins, including mitochondrial HSP22, induced by oxidative stress and adaptive response in tomato cells. Plant Journal **13**, 519–527.

Barnabas B, Jager K, Feher A. 2007. The effect of drought and heat stress on reproductive processes in cereals. Plant, Cell & Environment **31**, 11 – 38.

Bianchi MW, Damerval C, Vartanian N. 2002*a*. Identification of proteins regulated by cross-talk between drought and hormone pathways in *Arabidopsis* wild-type and auxin-insensitive mutants, axr1 and axr2. Functional Plant Biology **29**, 55.

Bianchi MW, Damerval C, Vartanian N. 2002*b*. Identification of proteins regulated by cross-talk between drought and hormone pathways in *Arabidopsis* wild-type and auxin-insensitive mutants, axr1 and axr2. Functional Plant Biology **29**, 55.

Blanca J, Montero-Pau J, Sauvage C, Bauchet G, Illa E, Díez MJ, Francis D, Causse M, van der Knaap E, Cañizares J. 2015. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. BMC Genomics **16**, 257.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F. 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS genetics **6**, e1000940.

Bruhn CM, Feldman N, Garlitz C, Harwood J, Ivans E, Marshall M, Riley A, Thurber D, Williamson E. 1991. Consumer perceptions of quality: apricots, cantaloupes, peaches, pears, strawberries and tomatoes. Journal of Food Quality **14**, 187–195.

Capel C, Fernández del Carmen A, Alba JM, *et al.* 2015. Wide-genome QTL mapping of fruit quality traits in a tomato RIL population derived from the wild-relative species *Solanum pimpinellifolium* L. Theoretical and Applied Genetics **128**, 2019–2035.

Causse M, Desplat N, Pascual L, *et al.* 2013. Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. BMC genomics **14**, 791.

Causse M, Friguet C, Coiret C, LePicier M, Navez B, Lee M, Holthuysen N, Sinesio F, Moneta E, Grandillo S. 2010. Consumer preferences for fresh tomato at the european scale: a common segmentation on taste and firmness. Journal of Food Science **75**.

Causse M, Saliba-Colombani V, Lesschaeve I, Buret M. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 2. Mapping QTLs for sensory attributes. Theoretical and Applied Genetics **102**, 273–283.

Chaves MM, Santos TP, Souza CR, Ortuño MF, Rodrigues ML, Lopes CM, Maroco JP, Pereira JS. 2007. Deficit irrigation in grapevine improves water-use efficiency while controlling vigour and production quality. Annals of Applied Biology **150**, 237–252.

Chen H, Zhang B, Hicks LM, Xiong L. 2011. A nucleotide metabolite controls stress-responsive gene expression and plant development. PLoS ONE **6**.

Choi Y, Chan AP. 2015. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics **31**, btv195.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w 1118; iso-2; iso-3. Fly **6**, 80–92.

Cormier F, Le Gouis J, Dubreuil P, Lafarge S, Praud S. 2014. A genome-wide identification of chromosomal regions determining nitrogen use efficiency components in wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics **127**, 2679–2693.

Cornah JE, Germain V, Ward JL, Beale MH, Smith SM. 2004. Lipid utilization, gluconeogenesis, and seedling growth in *Arabidopsis* mutants lacking the glyoxylate cycle enzyme malate synthase. Journal of Biological Chemistry **279**, 42916–42923.

Costa JM, Ortuño MF, Chaves MM. 2007. Deficit irrigation as a strategy to save water: physiology and potential application to horticulture. Journal of integrative plant biology **49**, 1421–1434.

Durán Zuazo VH, Pleguezuelo CRR, Tarifa DF. 2011. Impact of sustained-deficit irrigation on tree growth, mineral nutrition, fruit yield and quality of mango in Spain. Fruits **66**, 257–268.

Van Eeuwijk F a, Bink MC, Chenu K, Chapman SC. 2010. Detection and use of QTL for complex traits in multiple environments. Current Opinion in Plant Biology **13**, 193–205.

El-Soda M, Boer MP, Bagheri H, Hanhart CJ, Koornneef M, Aarts MGM. 2014*a*. Genotype-environment interactions affecting preflowering physiological and morphological traits of *Brassica rapa* grown in two watering regimes. Journal of experimental botany **65**, 697–708.

El-Soda M, Kruijer W, Malosetti M, Koornneef M, Aarts MGM. 2015. Quantitative trait loci and candidate genes underlying genotype by environment interaction in the response of *Arabidopsis thaliana* to drought. Plant, Cell & Environment **38**, 585–599.

El-Soda M, Malosetti M, Zwaan BJ, Koornneef M, Aarts MGM. 2014*b*. Genotype×environment interaction QTL mapping in plants: lessons from *Arabidopsis*. Trends in plant science **19**, 390–8.

FAO Water. 2015. Crop Water Information: Soybean. http://www.fao.org/nr/water/cropinfo_tomato.html.

Fereres E, Soriano MA. 2006. Deficit irrigation for reducing agricultural water use. Journal of Experimental Botany **58**, 147–159.

Frank W, Baar KM, Qudeimat E, Woriedh M, Alawady A, Ratnadewi D, Gremillon L, Grimm B, Reski R. 2007. A mitochondrial protein homologous to the mammalian peripheral-type benzodiazepine receptor is essential for stress adaptation in plants. Plant Journal **51**, 1004–1018.

Fridman E, Carrari F, Liu Y-S, Fernie AR, Zamir D. 2004*a*. Zooming in on a quantitative trait for tomato yield using interspecific introgressions. Science (New York, N.Y.) **305**, 1786–9.

Fridman E, Carrari F, Liu Y-S, Fernie AR, Zamir D. 2004*b*. Zooming in on a quantitative trait for tomato yield using interspecific introgressions. Science (New York, N.Y.) **305**, 1786–9.

Fridman E, Pleban T, Zamir D. 2000. A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proceedings of the National Academy of Sciences of the United States of America **97**, 4718–4723.

Gawenda I, Thorwarth P, Gunther T, Ordon F, Schmid KJ. 2015. Genome-wide association studies in elite varieties of German winter barley using single-marker and haplotype-based methods. Plant Breeding **134**, 28–39.

Gomez L, Bancel D, Rubio E, Vercambre G. 2007. The microplate reader: an efficient tool for the separate enzymatic analysis of sugars in plant tissues—validation of a micro-method. Journal of the Science of Food and Agriculture **87**, 1893–1905.

Gong P, Zhang J, Li H, Yang C, Zhang C, Zhang X. 2010. Transcriptional profiles of drought-responsive genes in modulating transcription signal transduction, and biochemical pathways in tomato. **61**, 3563–3575.

Gu J-F, Qiu M, Yang J-C. 2013. Enhanced tolerance to drought in transgenic rice plants overexpressing C4 photosynthesis enzymes. The Crop Journal **1**, 105–114.

Guichard S, Bertin N, Leonardi C, Gary C. 2001. Tomato fruit quality in relation to water and carbon fluxes. Agronomie **21**, 385–392.

Guillet C, Just D, Bernard N, Destrac-Irvine A, Baldet P, Hernould M, Causse M, Raymond P, Rothan C. 2002. A fruit-specific phosphoenolpyruvate carboxylase is related to rapid growth of tomato fruit. Planta **214**, 717–726.

Gur A, Semel Y, Osorio S, *et al.* 2011. Yield quantitative trait loci from wild tomato are predominately expressed by the shoot. Theoretical and Applied Genetics **122**, 405–20.

Hamilton JP, Sim S-C, Stoffel K, Van Deynze A, Buell CR, Francis DM. 2012. Single nucleotide polymorphism discovery in cultivated tomato via sequencing by synthesis. The Plant Genome Journal **5**, 17.

Huang Z, van der Knaap E. 2011. Tomato fruit weight 11.3 maps close to fasciated on the bottom of chromosome 11. Theoretical and Applied Genetics **123**, 465–474.

Illa-Berenguer E, Van Houten J, Huang Z, van der Knaap E. 2015. Rapid and reliable identification of tomato fruit weight and locule number loci by QTL-seq. Theoretical and Applied Genetics **128**, 1329–1342.

Jury W a, Vaux H. 2005. The role of science in solving the world's emerging water problems. Proceedings of the National Academy of Sciences of the United States of America **102**, 15715–15720.

Kim YS, Park JY, Kim KS, Ko MK, Cheong SJ, Oh BJ. 2002. A thaumatin-like gene in nonclimacteric pepper fruits used as molecular marker in probing disease resistance, ripening, and sugar accumulation. Plant Molecular Biology **49**, 125–135.

Kim M-H, Sato S, Sasaki K, Saburi W, Matsui H, Imai R. 2013. COLD SHOCK DOMAIN PROTEIN 3 is involved in salt and drought stress tolerance in *Arabidopsis*. FEBS open bio **3**, 438–42.

Kirda C, Cetin M, Dasgan Y, Topcu S, Kaman H, Ekici B, Derici MR, Ozguven AI. 2004. Yield response of greenhouse grown tomato to partial root drying and conventional deficit irrigation. Agricultural Water Management **69**, 191–201.

Korte A, Farlow A. 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant methods **9**, 29.

Korte A, Vilhjálmsson BJ, Segura V, Platt A, Long Q, Nordborg M. 2012. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. Nature genetics **44**, 1066–71.

Langridge P. 2006. Functional genomics of abiotic stress tolerance in cereals. Briefings in Functional Genomics and Proteomics **4**, 343–354.

Leib BG, Caspari HW, Redulla CA, Andrews PK, Jabro JJ. 2006. Partial rootzone drying and deficit irrigation of 'Fuji' apples in a semi-arid climate. Irrigation Science **24**, 85–99.

Lemoine R, Camera S La, Atanassova R, *et al.* 2013. Source-to-sink transport of sugar and regulation by environmental factors. Frontiers in Plant Science **4**.

Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO. 2010. Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the USA **107**, 21199–204.

Li Z, Palmer WM, Martin AP, Wang R, Rainsford F, Jin Y, Patrick JW, Yang Y, Ruan YL. 2012. High invertase activity in tomato reproductive organs correlates with enhanced sucrose import into, and heat tolerance of, young fruit. Journal of Experimental Botany **63**, 1155–1166.

Luna A, Nicodemus KK. 2007. snp.plotter: An R-based SNP/haplotype association and linkage disequilibrium plotting package. Bioinformatics **23**, 774–776.

Malosetti M, Ribaut JM, Vargas M, Crossa J, van Eeuwijk F a. 2007. A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). Euphytica **161**, 241–257.

Des Marais DL, Hernandez KM, Juenger TE. 2013. Genotype-by-environment interaction and plasticity: Exploring genomic responses of plants to the abiotic environment. Annual Review of Ecology, Evolution, and Systematics **44**, 5–29.

Des Marais DL, McKay JK, Richards JH, Sen S, Wayne T, Juenger TE. 2012. Physiological genomics of response to soil drying in diverse *Arabidopsis* accessions. The Plant cell **24**, 893–914.

Martinoia E, Rentsch D. 1995. Malate comprtmentation-responses to a complex metabolisme. Annual review of plant biology **45**, 447 – 467.

McClurg P, Pletcher MT, Wiltshire T, Su AI. 2006. Comparative analysis of haplotype association mapping algorithms. BMC bioinformatics **7**, 61.

Mirás-Avalos JM, Alcobendas R, Alarcón JJ, Valsesia P, Génard M, Nicolás E. 2013. Assessment of the water stress effects on peach fruit quality and size using a fruit tree model, QualiTree. Agricultural Water Management **128**, 1–12.

Morrison GD, Linder CR. 2014. Association mapping of germination traits in *Arabidopsis thaliana* under light and nutrient treatments: searching for G x E effects. G3 (Bethesda, Md.) **4**, 1465–1478.

Muir W, Nyquist WE, Xu S. 1992. Alternative partitioning of the genotype-by-environment interaction. Theoretical and Applied Genetics **84**, 193–200.

Neta-Sharir I, Isaacson T, Lurie S, Weiss D. 2005. Dual role for tomato heat shock protein 21: protecting photosystem II from oxidative stress and promoting color changes during fruit maturation. The Plant cell **17**, 1829–38.

Nguyen-Quoc B, Foyer CH. 2001. A role for 'futile cycles' involving invertase and sucrose synthase in sucrose metabolism of tomato fruit. Journal of experimental botany **52**, 881–889.

Nir I, Moshelion M, Weiss D. 2014. The *Arabidopsis* GIBBERELLIN METHYL TRANSFERASE 1 suppresses gibberellin activity, reduces whole-plant transpiration and promotes drought tolerance in transgenic tomato. Plant, Cell and Environment **37**, 113–123.

Osorio S, Ruan Y-L, Fernie AR. 2014. An update on source-to-sink carbon partitioning in tomato. Frontiers in Plant Science **5**.

Pan Y, Seymour GB, Lu C, Hu Z, Chen X, Chen G. 2012. An ethylene response factor (ERF5) promoting adaptation to drought and salt tolerance in tomato. Plant Cell Reports **31**, 349–360.

Pascual L, Albert E, Sauvage C, *et al.* 2016. Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. Plant Science **242**, 120–130.

Perruc E, Charpenteau M, Ramirez BC, Jauneau A, Galaud J-P, Ranjeva R, Ranty B. 2004. A novel calmodulin-binding protein functions as a negative regulator of osmotic stress tolerance in *Arabidopsis thaliana* seedlings. The Plant Journal **38**, 410–420.

Petre B, Major I, Rouhier N, Duplessis S. 2011. Genome-wide analysis of eukaryote thaumatin-like proteins (TLPs) with an emphasis on poplar. BMC plant biology **11**, 33.

Pignocchi C, Kiddle G, Hernández I, Foster SJ, Asensi A, Taybi T, Barnes J, Foyer CH. 2006. Ascorbate oxidase-dependent changes in the redox state of the apoplast modulate gene transcript accumulation leading to modified hormone signaling and orchestration of defense processes in tobacco. Plant physiology **141**, 423–435.

Proels RK, Roitsch T. 2009. Extracellular invertase LIN6 of tomato: A pivotal enzyme for integration of metabolic, hormonal, and stress signals is regulated by a diurnal rhythm. Journal of Experimental Botany **60**, 1555–1567.

R Development Core Team. 2012. R: a language and environment for statistical computing.

Reddy ASN, Ali GS, Celesnik H, Day IS. 2011. Coping with stresses: roles of calcium- and calcium/calmodulin-regulated gene expression. The Plant cell **23**, 2010–2032.

Rentsch D, Schmidt S, Tegeder M. 2007. Transporters for uptake and allocation of organic nitrogen compounds in plants. FEBS Letters **581**, 2281–2289.

Reuscher S, Akiyama M, Mori C, Aoki K, Shibata D, Shiratake K. 2013. Genome-wide identification and expression analysis of aquaporins in tomato. PLoS ONE **8**.

Ricardi MM, González RM, Zhong S, *et al.* 2014*a*. Genome-wide data (ChIP-seq) enabled identification of cell wall-related and aquaporin genes as targets of tomato ASR1, a drought stress-responsive transcription factor. BMC plant biology **14**, 29.

Ricardi MM, González RM, Zhong S, *et al.* 2014*b*. Genome-wide data (ChIP-seq) enabled identification of cell wall-related and aquaporin genes as targets of tomato ASR1, a drought stress-responsive transcription factor. BMC plant biology **14**, 29.

Ripoll J, Urban L, Bertin N. 2016*a*. The potential of the MAGIC TOM parental accessions to explore the genetic variability in tomato acclimation to repeated cycles of water deficit and recovery. Frontiers in Plant Science **6**.

Ripoll J, Urban L, Brunel B, Bertin N. 2016*b*. Water deficit effects on tomato quality depend on fruit developmental stage and genotype. Journal of Plant Physiology **190**, 26–35.

Ripoll J, Urban L, Staudt M, Lopez-Lauri F, Bidel LPR, Bertin N. 2014. Water shortage and quality of fleshy fruits, making the most of the unavoidable. Journal of experimental botany **65**, 4097–117.

Rost S, Gerten D, Bondeau A, Lucht W, Rohwer J, Schaphoff S. 2008. Agricultural green and blue water consumption and its influence on the global water system. Water Resources Research **44**.

Ruan YL, Jin Y, Yang YJ, Li GJ, Boyer JS. 2010*a*. Sugar input, metabolism, and signaling mediated by invertase: Roles in development, yield potential, and response to drought and heat. Molecular Plant **3**, 942–955.

Ruan YL, Jin Y, Yang YJ, Li GJ, Boyer JS. 2010*b*. Sugar input, metabolism, and signaling mediated by invertase: Roles in development, yield potential, and response to drought and heat. Molecular Plant **3**, 942–955.

Ruggieri V, Francese G, Sacco A, Alessandro AD, Rigano MM, Parisi M, Milone M, Cardi T, Mennella G, Barone A. 2014. An association mapping approach to identify favourable alleles for tomato fruit quality breeding. BMC plant biology **14**, 1–15.

Sacco A, Ruggieri V, Parisi M, Festa G, Rigano MM, Picarella ME, Mazzucato A, Barone A. 2015. Exploring a tomato landraces collection for fruit-related traits by the aid of a high-throughput genomic platform. PLoS ONE **10**, 1–20.

Saïdou A-A, Thuillet A-C, Couderc M, Mariac C, Vigouroux Y. 2014. Association studies including genotype by environment interactions: prospects and limits. BMC genetics **15**, 3.

Saliba-Colombani V, Causse M, Langlois D, Philouze J, Buret M. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. Theoretical and Applied Genetics **102**, 259–272.

Santesteban LG, Royo JB. 2006. Water status, leaf area and fruit load influence on berry weight and sugar accumulation of cv. 'Tempranillo' under semiarid conditions. Scientia Horticulturae **109**, 60–65.

Sasaki E, Zhang P, Atwell S, Meng D, Nordborg M. 2015. 'Missing' G x E Variation Controls Flowering Time in Arabidopsis thaliana. PLoS Genetics **11**, 1–18.

Sauvage C, Segura V, Bauchet G, Stevens R, Do PT, Nikoloski Z, Fernie AR, Causse M. 2014. Genome-Wide Association in tomato reveals 44 candidate loci for fruit metabolic traits. Plant physiology **165**, 1120–1132.

Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nature genetics **44**, 825–30.

Seki M, Narusaka M, Ishida J, *et al.* 2002. Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. The Plant Journal **31**, 279–292.

Semel Y, Schauer N, Roessner U, Zamir D, Fernie AR. 2007. Metabolite analysis for the comparison of irrigated and non-irrigated field grown tomato of varying genotype. Metabolomics **3**, 289–295.

Shinozaki K, Yamaguchi-Shinozaki K. 2007. Gene networks involved in drought stress response and tolerance. Journal of experimental botany **58**, 221–7.

Sim S-C, Durstewitz G, Plieske J, *et al.* 2012. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. PloS one **7**, e40563.

Stevens R, Buret M, Garchery C, Carretero Y, Causse M. 2006. Technique for rapid, small-scale analysis of vitamin C levels in fruit and application to a tomato mutant collection. Journal of agricultural and food chemistry **54**, 6159–65.

Tardieu F. 2012. Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. Journal of experimental botany **63**, 25–31.

Tieman DM, Zeigler M, Schmelz EA, Taylor MG, Bliss P, Kirst M, Klee HJ. 2006. Identification of loci affecting flavour volatile emissions in tomato fruits. Journal of Experimental Botany **57**, 887–896.

Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature **485**, 635–41.

Verbyla AP, Cavanagh CR, Verbyla KL. 2014. Whole-genome analysis of multienvironment or multitrait QTL in MAGIC. G3 Genes|Genomes|Genetics **4**, 1569–1584.

Wang G, Cai G, Kong F, Deng Y, Ma N, Meng Q. 2014. Overexpression of tomato chloroplast-targeted DnaJ protein enhances tolerance to drought stress and resistance to Pseudomonas solanacearum in transgenic tobacco. Plant Physiology and Biochemistry **82**, 95–104.

Wang W, Vinocur B, Altman A. 2003. Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta **218**, 1–14.

Warnes G, Leisch F. 2012. Package 'genetics'.

Wu B-H, Genard M, Lescourret F, Gomez L, Li S-H. 2002. Influence of assimilate and water supply on seasonal variation of acids in peach (cv Suncrest). Journal of the Science of Food and Agriculture **82**, 1829–1836.

Xu J, Ranc N, Muños S, Rolland S, Bouchet JP, Desplat N, Le Paslier MC, Liang Y, Brunel D, Causse M. 2013. Phenotypic diversity and association mapping for fruit quality traits in cultivated tomato and related species. Theoretical and Applied Genetics **126**, 567–581.

Zanor MI, Osorio S, Nunes-Nesi A, *et al.* 2009*a*. RNA Interference of LIN5 in tomato confirms its role in controlling brix content, uncovers the influence of sugars on the levels of fruit hormones, and demonstrates the importance of sucrose cleavage for normal fruit development and fertility. Plant physiology **150**, 1204–1218.

Zanor MI, Rambla JL, Chaïb J, Steppa A, Medina A, Granell A, Fernie AR, Causse M. 2009*b*. Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. Journal of Experimental Botany **60**, 2139–2154.

Zheng J, Huang G, Jia D, Wang J, Mota M, Pereira LS, Huang Q, Xu X, Liu H. 2013. Responses of drip irrigated tomato (*Solanum lycopersicum* L.) yield, quality and water productivity to various soil matric potential thresholds in an arid region of Northwest China. Agricultural Water Management **129**, 181–193.

Zhou GA, Chang RZ, Qiu LJ. 2010. Overexpression of soybean ubiquitin-conjugating enzyme gene GmUBC2 confers enhanced drought and salt tolerance through modulating abiotic stress-responsive gene expression in *Arabidopsis*. Plant Molecular Biology **72**, 357–367.

CHAPTER 5

CHAPTER 5: Expression inheritance patterns, allele specific expression and regulatory divergences in water stressed tomato

This chapter is in the form of an article draft describing the identification of differentially expressed genes, expression inheritance patterns, allele specific expression and eQTLs in two contrasted genotypes, an F1 hybrid and a recombinant inbreed line population, grown under two contrasted watering conditions.

Genetic architecture of transcriptome changes in water stressed tomato

Authors

Elise Albert[a], Renaud Duboscq[a], Murielle Latreille[b], Sylvain Santoni[b], Matthieu Beukers[a], Jean-Paul Bouchet[a], Fréderique Bitton[a], Charles Poncet[c], Véronique Gautier[c], José Jimenez-Gomez[d], Mathilde Causse[a,§]

Affiliations

[a] INRA, UR1052, Génétique et Amélioration des Fruits et Légumes, 67 Allée des Chênes, Centre de Recherche PACA, Domaine Saint Maurice, CS60094, Montfavet, 84143, France

[b] INRA, UMR1334, Amélioration génétique et Adaptation des Plantes, Montpellier SupAgro-INRA-IRD-UMII, 2 Place Pierre Viala, Montpellier, 34060, France

[c] INRA, UMR1095, Génétique Diversité et Ecophysiologie des Céréales, 5 Chemin de Beaulieu, Clermont Ferrand, 63039, France

[d] INRA, UMR1318, Institut Jean-Pierre Bourgin, AgroParisTech-INRA-CNRS, Route de Saint Cyr, Versailles, 78026, France

[§] Corresponding author

Mathilde Causse

Mathilde.Causse@paca.inra.fr

Tel: +33 (0)4 32 72 28 03

Supplemental materials referred in this chapter are available in **Appendix 6.**

Summary

Characterizing the natural diversity of gene expression variations across environments is an important step in explaining how genotype by environment interaction shapes phenotypes. In tomato, limited water supply can have a favorable impact on fruit quality but may reduce yield. QTL by watering regime interactions were recently reported for quality traits and candidate genes were proposed. Here, we analyzed the impact of water deficit at the genome expression level. For this purpose, we first sequenced the transcriptomes of growing leaves and fruit pericarps at cell expansion stage of a cherry type tomato, a large fruited tomato accession and their F1 hybrid, grown under two contrasted watering regimes. Gene expression was steadily affected by the watering regimes, notably in the F1 hybrid with 5,034 and 4,044 differentially expressed genes in the fruits and leaves, despite a low effect of the watering treatment on plant and fruit phenotypes. Whereas phenotypes showed mostly additive inheritance whatever the watering regime, gene expression inheritance varied according to the watering regime and the organ. The expression of roughly half of the genes presented over dominant or over recessive inheritance. Then, using 61,304 exonic SNPs identified in 11,719 genes polymorphic between parental genotypes and comparing allele specific expression in the hybrid with allelic expression in both parents, we identified 4,093 genes (35%) showing deviation from a 1 to 1 ratio, among which 1,982 *cis* regulated genes, 1,450 *trans* regulated genes and 715 genes presenting a combination of both types of regulatory divergence. Contrary to *trans* regulatory divergences, *cis* regulatory divergences were strongly conserved between organs and between watering regimes. Measuring the expression of a subset of 274 transcripts in fruits and leaves of 124 RILs, we identified 246 eQTLs (107 *local* and 139 *trans*), mostly confirming the regulatory divergences identified with the allele specific expression approach. Finally, combining phenotypic data and expression data, we characterized a complex cross talk between several genes coding for enzymes involved in the sugar metabolisms.

Key words

Allele specific expression, eQTL, Interaction genotype by environment, Fruit quality, Drought, *Solanum lycopersicum*

Introduction

Differences in gene expression are central to evolution and a key process of plant adaption (Anderson et al., 2012). Following the rapid development of expression arrays and sequencing technologies, genome-wide assays of gene expression were rapidly adopted by plant biologists to explore differential gene expression along plant development or in response to various biotic and abiotic stressors, including water deficit (Seki et al., 2002; Rabbani et al., 2003; Des Marais et al., 2012). However, researchers started only recently to explore the natural genetic variation in transcriptomic responses to environmental constraints.

Characterizing the genetic diversity of gene expression variations across environments is an important step in explaining at the molecular level how genotype by environment interaction shapes phenotypes. Currently, two complementary strategies are used for uncovering the genetic basis of variation in gene expression (Druka et al., 2010; Emerson and Li, 2010). The first strategy, expression quantitative trait locus (*eQTL*) mapping, takes advantages of methodologies developed for mapping quantitative trait loci (*QTL*), either in the linkage or association framework, considering transcript levels as quantitative traits (Nica and Dermitzakis, 2013). With the availability of genome sequences, this approach allows inferring the location of genomic regions responsible for variations in gene expression and distinguishing between *distant* and *local* eQTLs according to the distance to the regulated gene. *Distant* eQTLs are steadily inferable to *trans*-acting regulatory elements, *i.e.* variations in the coding region of a gene that affects the expression of other genes. However, eQTL mapping often lacks resolution to determine whether a *local* eQTL corresponds to a *trans*-acting regulatory element immediately adjacent to the gene it regulates or to a *cis*-acting element in the regulatory region of a gene that affects its own expression (Rockman and Kruglyak, 2006). The second strategy is based on the comparisons of allele specific expression (ASE) in an F1 hybrid with allelic expression in parental lines. In the hybrid both parental alleles are in the same cellular context and exposed to a common set of regulatory factors. Therefore, a conserved unbalanced allelic expression between parents and hybrid is the signature of parental *cis*-regulatory divergences whereas a balanced allelic expression only in hybrids reveals parental *trans*-regulatory divergences. Different unbalanced allelic expression between parents and hybrid may reflect a

combination of both *cis* and *trans*-regulatory divergences (Liu et al., 2014; Castel et al., 2015). In contrast to eQTL mapping, monitoring ASE does not allow mapping *trans*-acting elements on the genome, but can efficiently determine whether gene expression variations result from *cis-acting* elements or *trans*-acting elements closely linked to the gene they regulate.

The recent technological advance has encouraged the examination of variations, heritability and inheritance patterns of gene expression in model biological systems at a whole genome scale. ASE and eQTL mapping studies have been reported in yeast (Brem et al., 2002; Yvert et al., 2003; Emerson et al., 2010), mouse (Hubner et al., 2005; Wang et al., 2006) and human (Schadt et al., 2003; Monks et al., 2004; Morley et al., 2004; Bystrykh et al., 2005; Chesler et al., 2005; Cheung et al., 2005). In plants, most of the studies focused on *Arabidopsis thaliana* (DeCook, 2005; Kiekens et al., 2006; Keurentjes et al., 2007; West et al., 2007; Jiménez-Gómez et al., 2010; Cubillos et al., 2012; Cubillos et al., 2014; He et al., 2016), maize (Stupar and Springer, 2006; Springer and Stupar, 2007a) and tree species (Kirst, 2004 in Eucalyptus; Combes *et al.*, 2015 in coffee; Verta *et al.*, 2016 in white spruce), leading to the successful identification of loci and polymorphisms involved in gene regulation. In *Arabidopsis thaliana*, approximatively one-third of the eQTLs identified were *local* whereas a majority of *distant* eQTLs mapped to hotspots considered as 'master regulator' loci (Keurentjes et al., 2007; West et al., 2007). On the other hand, variations between different maize genotypes seem to be largely controlled by *cis*-acting polymorphisms (Stupar and Springer, 2006; Springer and Stupar, 2007a). However, to the best of our knowledge the eQTL by environment interaction issues have been addressed in only three reports in plant (Hammond *et al.*, 2011 in *Brassica rapa* under various soil phosphorus content; Cubillos *et al.*, 2014 in *Arabidopsis* under water deficit; van Muijen *et al.*, 2016 in diploid potato under water deficit).

Tomato (*Solanum lycopersicum*) is a crop of particular interest, as the fruit is an important source of nutrients for the human diet and a model for the study of fleshy fruit development (Giovannoni, 2001; Willcox et al., 2003). Several transcriptome analyses reveal the evolution of gene expression along fruit development (Tomato Genome Consortium, 2012) or in different parts of the fruit (Mounet et al., 2009; Matas et al., 2011; Koenig et al., 2013; Pattison et al., 2015). In this species, limited water supply can have a favorable impact on fruit quality, but requires finding the right balance to minimize yield loss

(Ripoll et al., 2014). QTL by watering regime interactions were recently reported for tomato major fruit quality traits and various candidate genes were proposed (Gur et al., 2011; Albert et al., 2016a; Albert et al., 2016b). However, tomato genetic variation in gene expression under water deficit has only been considered so far on a limited number of contrasted genotypes. First studies have reported up to 2,250 genes differently expressed in leaves in response to water limitation in two cultivars (Albert et al., 2016a). Characterizing the pattern of inheritance and the genetic architecture of gene expression variations and their interaction with water deficit in tomato will bring crucial knowledge for understanding the mechanisms of fleshy fruit crops adaptation to their environment. Besides, colocalizations between QTLs and eQTLs may speed up the identification of the causal genes responsible for tomato fruit quality variations in response to environmental constraints.

In the present study, we sequenced the transcriptomes of fruit pericarps at the cell expansion stage and growing leaves of a cherry type tomato line, a large fruited line and their F1 hybrid, grown under two contrasted watering regimes. We first identified differentially expressed genes between genotypes and according to watering regimes and characterized the patterns of inheritance of gene expression variations in the F1 hybrid. Then using 61,304 exonic SNPs identified in 11,719 genes from parental line whole genome re-sequencing (Causse et al., 2013), we characterized ASE and regulatory divergences in fruits and leaves of the F1 hybrid under both watering regimes at a whole transcriptome scale. We thus selected a subset of 274 differentially expressed genes located in the main QTL and QTL by watering regime interaction regions identified in Albert et al. (2016a) and confirmed in the present study. Their expression was measured in fruits and leaves of 124 recombinant inbreed lines. Performing eQTL and eQTL by watering regime interaction mapping, we cross validated the regulatory divergences characterized using ASE and proposed candidate genes for fruit quality variations in response to water deficit in tomato.

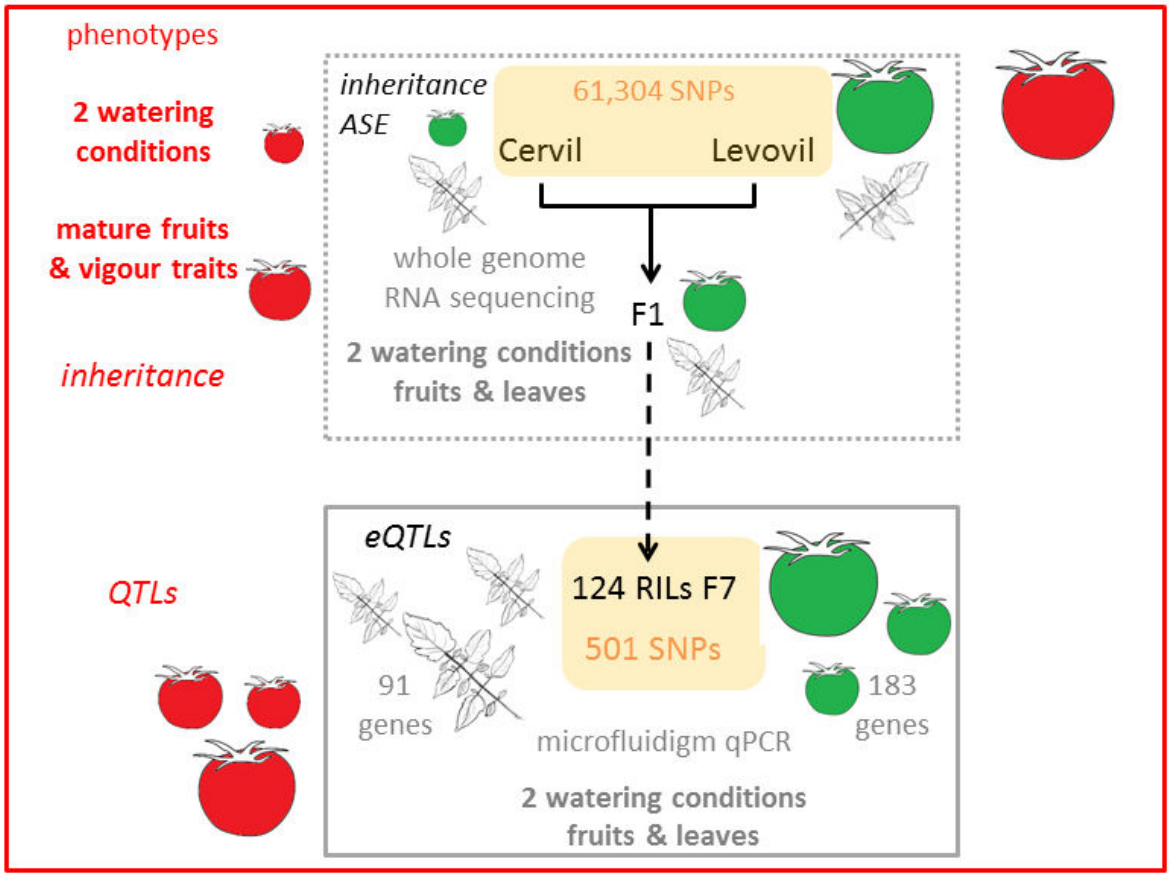


Figure 1. Experimental design. (1) Phenotypic data (related to plant vigor and mature fruit quality) and whole genome RNA sequencing data in two organs (cell expansion fruit pericarps and growing leaves) were collected on two tomato lines (Cervil: *S. lycopersicum* var *cerasiforme* and Levovil: *S. lycopersicum*) and their F1 hybrid, grown under two watering regimes (Control and Drought). RNA sequencing data were analyzed in order to identify allele specific expression (ASE) in the hybrid compared to the parental lines, using 61,304 high quality exonic SNPs carried by 11,719 genes, obtained from whole genome re-sequencing of both parental lines (Causse et al., 2013). Fruit quality phenotypic data and gene expression data were analyzed to characterize their patterns of inheritance. (2) Plant vigor and fruit quality phenotypic data were collected on 124 F7 RILs resulting from the cross between both parental lines, grown under control and drought watering regimes. Using 501 SNPs genotyped across the genome (Pascual et al., 2016), QTL and QTL by watering regime were mapped. In the main QTL regions, on the basis of the differentially expressed genes between genotypes and between watering regimes in the RNA sequencing data, considering genes with annotations related to response to abiotic stresses, 274 genes were selected and their expressions were measured by microfluidigm qPCR in 124 RILs, in cell expansion fruit pericarps (183 genes) and growing leaves (91 genes), under control and drought watering regimes. The expression data were analyzed in the linkage mapping framework to identify *distant* and *local* eQTLs.

Material and Methods

The experimental design supporting this study is summarized in **Figure 1.**

Plant material and experimental design

The plant material was composed of two inbred lines, their F1 hybrid and 124 F7 recombinant inbred lines (RILs) developed from the cross between the two lines. The male parent (Cervil) was a cherry type tomato (*S. lycopersicum cerasiforme*) with high aroma intensity, whereas the female parent (Levovil) was a large fruited accession (*S. lycopersicum,* 90–160 g) with common taste (described in Saliba-Colombani et al., 2000). In 2015, plants were grown in a glasshouse in INRA Avignon, from March to July, with shade screens to maintain light intensity during day below 700 w/m² and heater to maintain temperature during night beyond 18°C. Plants were grown in 4 liter plastic pots filled with peat (Klasmann 165) and watered with nutritive solution (2, 4, 6 mmol l^{-1}, N, P, and K, respectively). Two watering regimes were applied to the plants: drought (D) and control (C). Control treatment was applied according to ETP (evapotranspiration) climatic data and the cultural coefficient for tomato crop with a maximal drainage of 25 % and a relative humidity of the peat substrate of 65 %. The drought treatment was progressively applied after flowering of the second truss of Cervil (earliest genotype): water supply was reduced by 25 % compared to control for one week, then decreased by 40 % until the end of the experiment, aiming to apply a mild water deficit. Through the experiment, relative humidity of the peat substrate was controlled with a GRODAN® moisture probe and monitored in drought pots between 25 and 30 %. Two plants per watering regime for the RILs and three plants per watering regime for the parental lines and the F1 hybrid were randomized within the greenhouse.

Plant and fruit phenotyping

Plants were phenotyped for traits describing plant performance and fruit characteristics as described in Albert et al. (2016a). Flowering date (Flw) was assessed in number of days after sowing. The implantation height (Ht, in cm), stem diameter (Diam, in mm) and leaf length (Leaf, in cm) were measured on each plant. Fruit measurements were conducted on mature tomatoes harvested daily on the basis of their red color to ensure homogeneous ripening stage. At least, ten fruits per genotype per watering regime were harvested on 3rd to 6th truss. For each fruit, fresh weight (FW, in g) and firmness (FIR, in Durofel index) were

measured. Besides, harvested fruits were pooled in three groups of three to four fruits per watering regime. These pools constituted three replicates for the biochemical analysis. In each pool, a quarter of fruit pericarp was sampled and dried in an oven at 60 °C for 4 days to measure dry matter content (DMW, in %). Then, half of each fruit pool was mixed in juice to measure pH and soluble solid content (SSC, with a refractometer, in °Brix). Pericarps were sampled from the remaining fruit of each pool, frozen with liquid nitrogen and ground into powder for sugar (Glucose and Fructose) and acid malic (Malic) content assessment according to enzymatic protocols described in Garcia and Renard (2014) with minor adaptations. The different metabolic content were expressed both relative to fresh matter (g 100 g^{-1} of FM) and relative to dry matter (g 100 g^{-1} of DM) using DMW. Phenotypic data are available in **Supplemental Table 1.**

Statistical analysis on phenotypic data

All statistical analyses were performed using R (R Development Core Team, 2012) and all *P-values* were considered to be statistically significant when below 0.05. Prior to the analyses of variance (ANOVA) and when distributions were skewed, phenotypic data were normalized using Box and Cox transformations. ANOVAs were performed first on the parent and F1 hybrid individual data, then on the RIL individual data, according to the following model:

$$Y_{ijk} = \mu + G_i + W_j + G_i * W_j + e_{ijk}$$

Y_{ijk} was the phenotypic value of accession i in watering regime j, μ the overall mean, G_i the fixed effect of accession i, W_j the fixed effect of watering regime j and e_{ijk} the residual error effect. For the parent and F1 hybrid, when the interaction factor G x W and/or the factor G were significant, we computed a Tukey's post-hoc test to compare the means. Besides, we estimated additivity (A) and dominance (D) components of genetic variation. A was computed as the absolute value of half of the difference between two parental line means. D was computed as the difference between the F1 hybrid mean and the parental mean. The inheritance patterns were assessed by the dominance/additivity (D/A) ratio and classified as over-recessive ($D/A < -1.2$), recessive ($-1.2 \leq D/A \leq -0.8$), additive ($-0.8 < D/A < 0.8$), dominant ($0.8 \leq D/A \leq 1.2$), or overdominant ($D/A > 1.2$) (according to Pascual et al., 2013).

For the RILs, under both watering regimes, restricted maximum likelihood estimates of the

genetic and residual variances (σ^2_G and σ^2_e) were computed with a second linear model: $Y_{ijk} = \mu + Gr_i + e_{ijk}$ (G_i and e_{ijk} random). Broad-sense heritabilities (H^2) were computed under both watering regimes as the ratio between the genetic variance and the total phenotypic variance: $H^2 = \sigma^2_G / \sigma^2_{Total}$, with $\sigma^2_{Total} = \sigma^2_G + 1/n*\sigma^2e$ (with n the number of replicates per accession). Spearman coefficients estimated the correlations between H^2 and σ^2_G under drought and control conditions for a same trait.

RIL average values in each watering regime were used for subsequent analyses. Pearson coefficients estimated the correlations between means under both watering regimes and with data collected in Albert et al. (2016a) for the same genotypes within an equivalent experimental design. Plasticity was computed on the means as: $\Delta_{ki} = (D_{ki} - C_{ki})/C_{ki}$, with Δ_{ki} the plasticity value for trait k and accession *i*, D_{ki} the mean of trait k under drought condition for accession *i* and C_{ki} the mean of trait k under control condition for accession *i*.

Fruit and leaf sampling and RNA extraction

Samples of growing young leaves and pericarps from cell expansion (CE) fruits were collected on each plant and immediately frozen in liquid nitrogen. CE fruits were collected 21 DAA (days after anthesis) for large fruited accessions and 14 DAA for small fruited accessions (*according to Nadia Bertin personal communication*, **Supplemental Table 1**). Frozen leaves and fruit pericarps were pooled and ground to constitute the biological replicates per watering regime per organ: three for the parental genotypes, two for the F1 hybrid and one for each RIL. For each replicate, RNA was extracted using the '*Spectrum Plant Total RNA*' kit (Sigma-Aldrich, adapted to plant tissue samples) following the manufacturer's protocol, with '*On-Column DNase I Digestion Set*' (Sigma-Aldrich) treatment to remove genomic DNA traces. RNA purity was assessed on Nanodrop 1000 (ThermosFisher). All 260/280 nm ratios were comprised between 1.8 and 2.2. RNA integrity was assessed on Bioanalyser 2100 (Agilent) using *'RNA 6000 Nano*' kits. All RNA integrity numbers (RIN) were beyond 6. Samples were assayed on Qubit 3.0 Fluorometer (ThermoFisher) to determine the concentration for each RNA sample using '*Qubit RNA Broad Range Assay*' kits.

RNA sequencing for parental lines and F1 hybrid

A total of 32 messenger RNA paired-end strand specific libraries were constructed from 1 µg of total RNA for the parental lines and the F1 hybrid (one library per biological replicate per

organ per watering regime). All libraries were constructed using '*TrueSeq Stranded mRNA Sample Preparation' kits* (Illumina, San Diego, CA) following the manufacturer instructions. Briefly, poly-A mRNA were selected using poly-T oligo-attached magnetic beads, subjected to enzymatic fragmentation and purified. Then, DNA first-strands were synthetized using random-hexamer primers. For the second strand synthesis, dUTP were incorporated in place of dTTP to build strand specific libraries. An end repair process was applied to cDNA fragments, including dA-tailing to the 3' end and index ligation. Finally, libraries were purified and enriched through PCR using Illumina primers. Fragment sizes were controlled using Bioanalyser 2100 (Agilent) with '*DNA 7500*' kits aiming to design libraries with insert sizes ranging from 100 to 400 bp. The indexed parental and F1 hybrid libraries were combined in two lanes (one for the leaf samples and one for the fruit samples) and were subjected to 150-bp paired-end Illumina next generation sequencing at the GenoToul platform (INRA Toulouse, France).

RNA-sequencing data processing and read count generation

The data concerning the parents and the F1 hybrid were treated using the same method. Read quality was assessed with *FASTQC v.0.11.5* (Babraham, 2011). Sequencing adapters and size markers were trimmed out using *Cutadapt 1.9.1* (Martin, 2011). We allowed 10% error with at least 10 bases overlap in a first *Cutadapt* round, then no error and at least four bases overlap in a second round. Reads shorter than 30 bp after trimming and having an average quality of their three first bases lower than 25 were discarded from the analysis.

Then, remaining reads were filtered using *Selqual* and *Selpairs* (homemade Python programs, available upon demand). *Selqual* was used to select for each read the longest segment fulfilling the chosen quality criteria whatever its position, using a sliding window of length 6 bp and prohibiting any missing base in the selected segment. The average quality inside the sliding windows was calculated as a Phred score deduced from the average probability of error deduced from the Phred score at each position. When the average quality of the bases within the window felt bellow 20, the read was cut at that point and the window moved over the rest of sequence with the same criteria. Of the resulting fragment, the longest for each read was kept in the analysis only when its size was equal or above 30 bp. Then, *Selpairs* was used in order to properly reconstruct read pairs, keeping only paired-end reads. After quality filtering, a total of 654 million paired-end reads were retained (93%

of the initial total read count) for the 32 libraries (in average of 20 million read pairs per library). Finally, read pairs from each individual sample were aligned to the tomato reference genome (Heinz 1706, v2.5) using *tophat2 2.1.1* (Kim et al., 2013) setting the mate inner distance to 300 bp. The tomato gene model (annotation 2.4) was provided to help the mapping process. All other mapping parameters were kept on default values. Alignments were sorted on the leftmost coordinates and filtered to keep only concordantly mapped reads using *SAMtools 1.3.1* (Li et al., 2009). For each gene and each library, read counts were generated using *HTSeq-Count 0.6.1* (Anders et al., 2014), providing a modified gene model file with a gene id field. In average, 18 million reads were mapped per library (**Supplemental Figures 1 and 2**).

Differential gene expression analysis

Mapped reads of each sample were analyzed using the Bioconductor package *DESeq2* (Love et al., 2014) in *R version 3.3.2* (R Development Core Team, 2012). Fruit and leaf samples were analyzed separately. To remove the negative effect of background expression noise on differential expression analysis, we restricted the analysis to genes with a minimum cumulated read counts of 15 across replicates, resulting in 10,960 genes (32%) discarded among leaf samples and 11,627 genes (34%) discarded among fruit samples.

Using *DESeq2*, count data were first normalized on the total number of counts using the Trimmed Mean of M-values (TMM) method (Robinson and Oshlack, 2010) in order to correct for differences among library sizes. Then, the per-gene dispersions were estimated by incorporating data-driven prior distributions and negative binomial generalized linear models were fitted for each gene to estimate moderated LOG2 fold changes between genotypes (F1, Cervil or Levovil) and between watering conditions (Control or Drought). Per-gene Wald test statistics were computed to identify significantly differentially expressed genes between genotypes in each watering regime and between watering regimes for a given genotype. Per-gene likelihood ratio tests were computed to identify genes for which the watering regime effect significantly differed across watering regimes, comparing a full model to a reduced model without the interaction term. In all tests, a FDR threshold of 5% was fixed to call significantly differentially expressed genes (Benjamini and Hochberg, 1995). Normalized count data are available in **Supplemental Tables 2 and 3.** A variance stabilizing transformation was applied to normalized gene expression data prior to performing principal

components analyses (PCA) and plotting the first two components.

Identification of GO terms related to biological process that were significantly enriched within the differentially expressed genes between watering regimes compared to the tomato genome was achieved using the 'GO term enrichment analysis' tool (http://bioinfo.bti.cornell.edu/tool/GO/GO_enrich.html) based on the 'GO::TermFinder' program described in (Boyle et al., 2004). To restrict the gene lists to the most significant genes, we focused on genes differentially expressed with a FDR *P-values* below 0.01 and GO terms were declared significantly enriched when their Bonferroni corrected *P-values* were below 0.01.

Expression inheritance patterns

Expression inheritance was determined for differentially expressed genes in one or more genotypic comparisons, under both watering regimes, in fruits and leaves. Transformed *DESeq2* normalized expression values were used to estimate additivity (A) and dominance (D) components as presented above for phenotypic data. Pearson chi-squared tests were computed to compare the frequency of the different inheritance patterns between organs (fruits and leaves) and between watering regimes (Control and Drought), estimating *P-values* with a 1,000 permutations Monte Carlo simulation and considering significance when *P-values* were below 0.05.

Allele specific quantification and allele specific expression test

For each library, allele-specific gene expression (ASE) levels were estimated using 61,304 high quality exonic SNPs between both parental accessions from previously published genome re-sequencing data (Causse et al., 2013) and a homemade PERL program (José Jimenez, available upon demand). The exonic SNP were distributed over 11,719 tomato genes (annotation 2.4, ~1/3 of the total number of tomato genes). Individual RNA-seq reads tended to span multiple segregating alleles. Thus, to avoid double-counting, allele-specific counts were created on the basis of individual reads summed across the gene, ensuring consistency of the reads originating either from the Cervil or Levovil haplotype. Reads that could not be consistently assigned to either haplotype were discarded (on average 0.09 reads per gene, sd= 1.64). An allele-specific detection PERL program was applied both, on the parental and hybrid libraries, to ensure fair comparisons. A total of 289 genes presented

unexpected allele assignations in the parental libraries and were discarded from the analysis. Genes with no count or less than 15 allele specific counts over the replicates were discarded from the analysis (4,112 genes in fruit data and 3,777 genes in leaves).

Under both watering regime separately, allele specific expression in the F1 hybrid was compared with allelic expression in parental lines using *DESeq2* (Love et al., 2014) in order to identify *cis*, *trans* and combinations of *cis* and *trans* regulatory divergences. The *calcNormFactors* function was used to scale both maternal and paternal counts for the F1 hybrid by the same normalization factor to avoid absorbing the allele specific expression when normalizing read counts. The per-gene dispersions were estimated by incorporating data-driven prior distributions and negative binomial generalized linear models were fitted for each gene to estimate log fold changes between generation (F1 or parent) and between allele (Cervil or Levovil). Then, per gene Wald tests were computed to identify genes for which alleles where significantly differentially expressed in the F1 hybrid (*cis* regulatory divergences) and to identify genes for which the allelic ratio differed between the F1 hybrid and the parental genotypes (*trans* regulatory divergences). As for differential gene expression, FDR threshold of 5% was fixed to call significant differences. Pearson chi-squared tests were computed to compare the frequency of the different regulatory patterns between organs (fruits and leaves) and between watering regimes (Control and Drought), computing *P-values* with 1,000 permutations Monte Carlo simulation and considering significance when *P-values* were below 0.05.

Identification of GO terms related to biological process that were significantly enriched within the different regulation classes (*cis, trans, cis + trans*) compared to the tomato genome was achieved using the 'GO term enrichment analysis' tool as for differentially expressed genes. GO terms were declared significantly enriched when their Bonferroni corrected *P-values* were below 0.01.

Primer design and quantitative gene expression analysis by microfluidigm qPCR in the RILs

Expression variation of 274 target genes (183 in fruits and 91 in leaves) were quantified under both watering regimes in the RIL RNA samples by quantitative real time microfluidigm PCR on the Gentyane platform (INRA Clermont-Ferrand, France). The target genes were chosen among the differentially expressed genes between watering regimes and/or

between parental accessions identified in the parental line RNA-sequencing data, according to (1) their annotation and potential involvement in response to water deficit according to literature, and/or (2) their location in the main QTL and QTL by watering regime interaction regions identified in Albert et al. (2016a) using the same design and confirmed in the present study (see below). Genes selected mainly belonged to 12 functional categories **(Supplemental Figures 3 and 4).**

A custom Python program based on Primer3 (Rozen and Skaletsky, 1999) was developed to design primers for each target transcript checking for amplification specificity (blast on the tomato CDS, annotation v2.4) and positioning at least one of the primers on an exon-exon junction to avoid bias of quantification due to potential genomic DNA contaminations (program available upon demand). Gene-specific primer lists are available in **Supplemental Tables 4 to 5**. Primer sizes were comprised between 17 and 25 pb, with Tm varying between 58 and 62°C, GC% comprised between 40 and 60% and amplicon sizes varying from 200 to 300 pb.

First strand cDNAs were synthetized from the RIL RNA samples using oligod(T) and Superscript III, followed by a Ribonuclease H treatment to delete any RNA traces in the samples. Then, a PCR preamplification was achieved pooling all primers in order to increase the amount of the initial cDNA molecules several fold, while preserving the relationships between the transcripts. Thermal cycling conditions consisted of 10 min at 95°C and 14 cycles of 15 s at 95°C and 4 min at 60°C. Primer traces eventually present in the samples were deleted using an exonuclease I treatment to avoid any interference in the qPCR reaction. Preamplicons were deposed on the BioMark HD system (Fluidigm). On each BioMark HD cheap, a negative control, four dilution points and reference genes were included.

RT-Q-PCR results were captured and analyzed using the BioMark HD software (Fluidigm). The EvaGreen fluorescent signal was standardized to a passive reference dye (ROX) included in the EvaGreen PCR master mix. The BioMark HD software allowed computing the cycle number at which the fluorescence passed the cycle threshold (CT) for each reaction. Relative expression levels were obtained by normalization to three reference genes for fruit samples and two reference genes for leaves samples **(Supplemental Table 6),** using the ΔCT method and considering reaction efficiency equal to two in accordance with the analysis of

Table 1. Effect of genotype (G), watering regime (W) and the interaction (G x W) on the phenotypic traits in Cervil, Levovil, their F1 hybrid and the 124 RILs. G, W and G × W indicate the significance of the ANOVA test for the genotype, watering regime and interaction effects, respectively. 'Inheritance control' and 'inheritance drought' were determined according to the dominance/additivity (D/A) ratio between both parental lines and the F1 hybrid (see Material and Methods). For the RILs, SS G, SS W and SS G × W display the proportion of each effect in the total sum of squares, respectively. 'H^2 control' and 'H^2 drought' indicate the broad sense heritabilities measured in the RILs in control and drought conditions, respectively.

	Cervil, Levovil and F1 hybrid					**124 RILs**							
	G	**W**	**G x W**	**Inheritance Control (D/A)**	**Inheritance Drought (D/A)**	**G**	**SS G (%)**	**W**	**SS W (%)**	**G x W**	**SS G x W (%)**	**H^2 Control**	**H^2 Drought**
Traits													
Plant traits													
Flw	***	NS	NS	Additive (-0.24)	Additive (-0.33)	***	84.05	NS	0.1	NS	6.25	80.14	79.54
Ht	***	***	*	Additive (0.79)	Dominant Lev (1.06)	***	88.96	**	0.82	***	6.61	92.65	92.3
Dia	*	***	***	NS	Recessive Lev (-0.89)	***	46.42	***	23.46	***	15	56.13	62.01
Fruit traits													
FW[a]	***	*	NS	Additive (-0.63)	Additive (-0.71)	***	88.07	***	3.19	***	3.99	93.36	92.1
DMW	***	NS	NS	Additive (0.26)	Additive (0.11)	***	71.11	***	1.16	***	8.82	72.18	70.54
SSC	***	NS	*	Additive (0.40)	Additive (0.07)	***	69.89	NS	0.14	NS	6.52	68.24	61.67
FructoseFM	***	NS	NS	Dominant Cer (1.02)	Additive (0.32)	***	44.22	***	0.43	***	15.42	35.57	43.76
GlucoseFM	***	*	NS	Additive (0.71)	Additive (-0.01)	***	56.67	NS	0.19	***	13.05	45.25	63.03
pH	***	NS	NS	NS	NS	***	44.34	NS	3.17	NS	11.15	38.25	33.8
MalicFM	NS	NS	NS	NS	NS	***	28.72	*	0.51	***	20.12	31.33	13.91

[a] Transformed to ensure a normal distribution (LOG10).

*** shows *P-value* below 0.001, ** between 0.001 and 0.01 and * between, 0.01 and 0.05. NS non-significant.

the dilution curves. Geometric means were computed over the reference genes in order to compute a reference CT for each sample in the normalization procedure. Expression data are available in **Supplemental Tables 4 and 5.**

Relative expression levels under both watering regimes and LOG10 ratio of the expression levels under drought relative to the expression levels under control condition for each genotype were used in the subsequent eQTL mapping analysis. Pearson correlation between expression means and phenotypic means were performed, considering significance when *P-values* were below 0.05.

(e)QTL and (e)QTL by watering regime interaction mapping

Phenotypic (mean values under both watering regimes and plasticity values) and expression (individual value under both watering regimes and LOG10 ratio) traits were used for QTL detection. When distributions were skewed, corrections for normality were applied. The QTL detection was performed by simple interval mapping (Lander and Botstein 1989) using the EM algorithm method implemented in R/QTL package (Broman et al. 2003) following the procedure indicated in Albert et al. (2016a) and using a genetic map constituted of 501 SNP markers genotyped in the whole population and covering 98% of the assembled tomato genome (assembly 2.5) (available in Pascual et al. 2016).

Briefly, a 1000-permutation test was performed to estimate significant threshold. LOD thresholds were 3.13 and 2.81, corresponding to genome wide significance levels of $\alpha = 0.05$ and $\alpha = 0.10$, respectively. When an (e)QTL was detected in one watering regime, the effect and PVE were also calculated in the second watering regime. Then, ANOVA tests were computed to test for genotype (G test) and interaction (G x W test) effects at each marker. This testing method was inspired from the multi-trait mixed model (MTMM) developed by Korte et al. (2012) for association analysis, considering only fixed effects. To correct for multiple testing in the ANOVA, significance thresholds corresponding to a genome wide significance level of $\alpha = 0.05$ were computed by a 1000-permutation test ($P\text{-value}_{\text{G test}} = 2.10 \times 10^{-4}$; $P\text{-value}_{\text{G} \times \text{W test}} = 2.00 \times 10^{-4}$).

For each detected (e)QTL, position, LOD score, marker at the LOD score peak, confidence interval (genetic CI, LOD decrease of two units), average phenotypic values of the two parental alleles and percentage of phenotypic variation explained (PVE) were displayed.

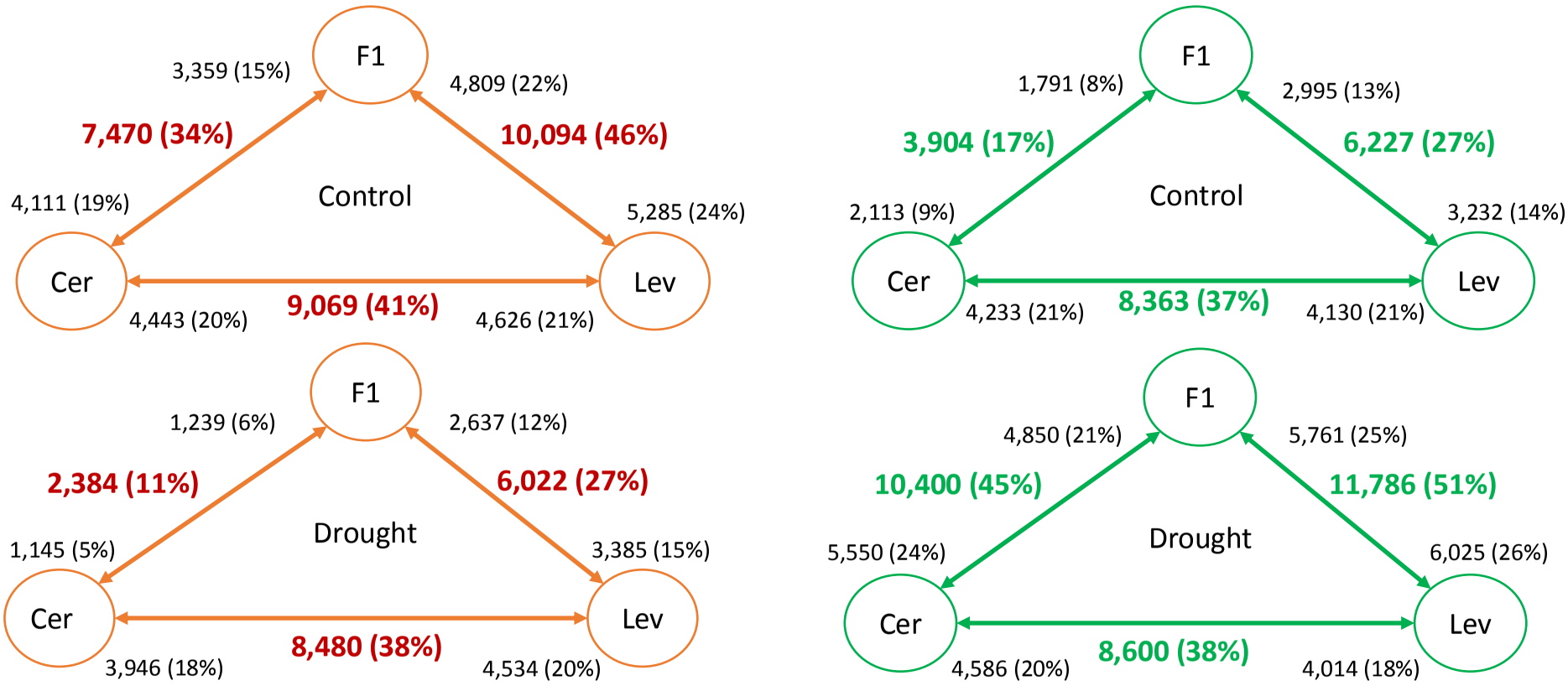


Figure 2: Differentially expressed genes in each possible comparison between parents and F1 hybrid in fruits (left) and leaves (right), under both watering regimes. Bold and colored numbers indicate the number and fraction of genes differentially expressed in each comparison among the 22,204 and 22,871 genes analyzed in fruits and leaves, respectively. For each comparison, the number and fraction of upregulated genes are indicated at the end of the arrows.

(e)QTL effects were calculated as: (Cervil mean allele − Levovil mean allele)/2. The genetic CI were translated into physical intervals (physical CI in Mbp) onto the tomato genome (assembly v2.5). We distinguished *constitutive* (e)QTLs (detected under both watering regimes), from *specific* (e)QTLs (detected under control or drought only) and *interactive* (e)QTLs (detected in the G x W test or using expression LOG10 ratio or plasticity values). Interactive (e)QTLs with effect direction changing between watering regimes were considered as *antagonist*, whereas (e)QTLs with intensity of effect changed intensity between watering conditions were considered as *differential*. Besides, for eQTLs, we distinguished between *trans* eQTLs when the regulated gene was not comprised in the eQTL CI, from *local* one when the regulated gene was inside the eQTL CI.

Results

1. Phenotypic variation and inheritance of phenotypic traits under both watering regimes

A total of thirteen phenotypic traits, three related to plant vigor and phenology and ten related to mature fruit quality, were assessed on Cervil, Levovil, their F1 hybrid and the 124 RILs derived from Cervil and Levovil cross, grown under control and drought conditions. Cervil, Levovil and their F1 hybrid differed for all the phenotypic traits except malic acid content **(Table 1).** However, the watering regime and/or genotype by watering regime interaction factors were only significant for Ht, Dia, FW, SSC and GlucoseFM. This limited phenotypic response to the watering treatment reflects our intention to perform physiologically relevant mild water deficit treatment (-40% water supply) compared to Albert et al. (2016a) (-60% water supply). The mode of inheritance of the traits that were significantly different between the parental genotypes and the F1 hybrid was assessed. We observed mostly additive inheritance for the different traits under both watering regimes, except for Ht and Dia for which we identified dominant and recessive inheritance with higher values of the Levovil parent and for FructoseFM for which we identified dominant inheritance patterns with higher value of the Cervil parent **(Table 1 and Supplemental Figures 4 to 6).**

Among the 124 F7 RILs, we observed large phenotypic variations under both watering regimes. For all traits except FW, transgressions beyond parental values were showed in both directions, under both watering regimes **(Supplemental Figures 7 and 6).** In the

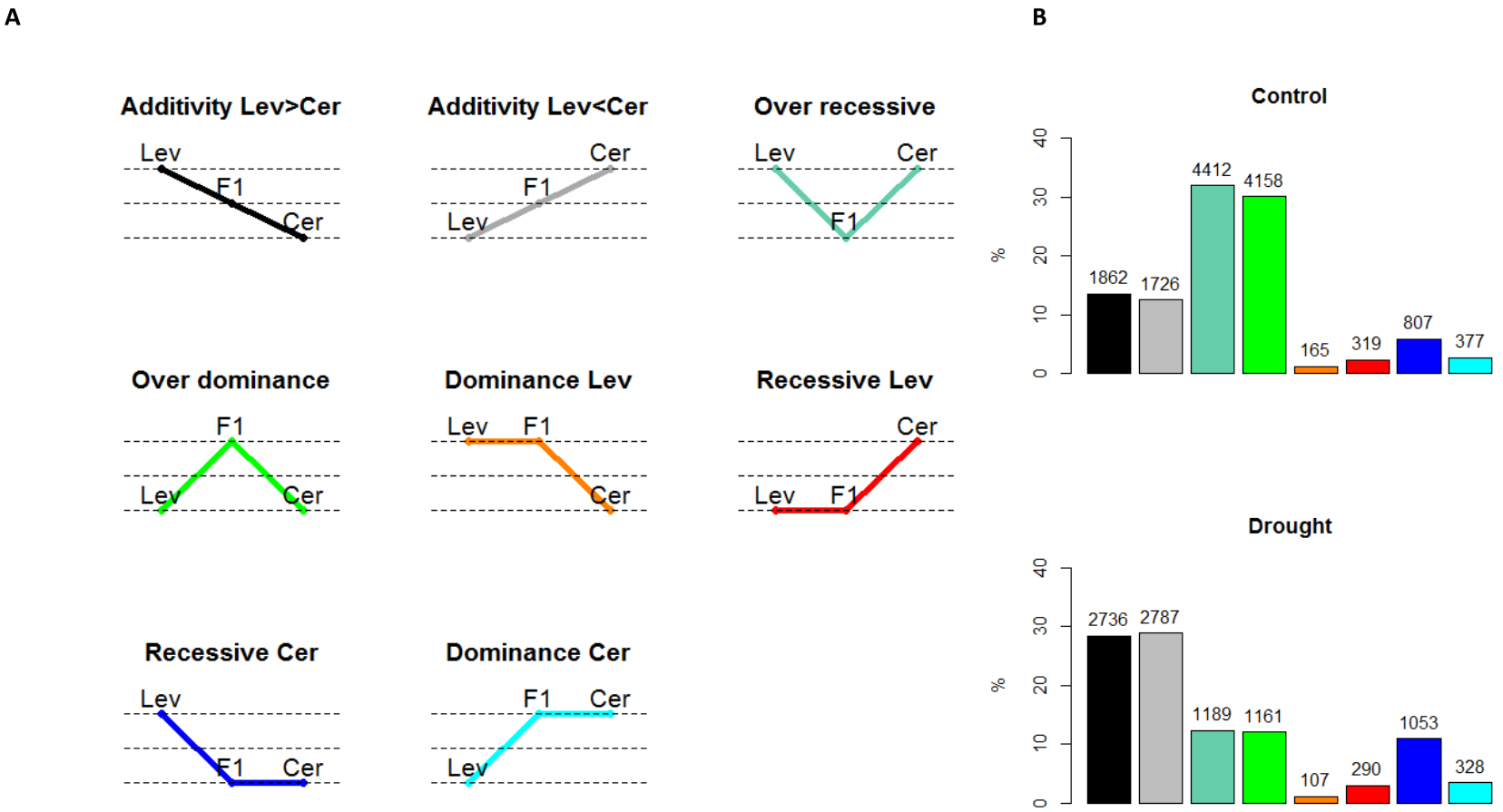


Figure 3: Inheritance of gene expression levels in fruits of the Cer x Lev F1 hybrid under both watering conditions. (A) The eight hypothetical patterns of gene expression in the two parental accessions and their F1 hybrid. (B) Bar plots showing the number and frequency of genes in each expression inheritance category (color defined in A), under control and drought conditions. A total of 13,826 and 9,651 genes, differentially expressed in one or more genotypic comparisons, were considered for the inheritance assessments, in control and drought conditions, respectively.

ANOVAs, the genotype by watering regime interaction was significant for most of the traits (except Flw, SSC and pH), but accounted for a low to medium part of the total phenotypic variation (4% to 20%) in comparison to the genotype main effect (29% to 89%) **(Table 1).** The broad sense heritabilities were comprised between 0.14 (MalicFM under drought) and 0.93 (for FW under control) and were conserved between watering regimes (Spearman correlation, r^2=0.93, *P-value* = 0.0001). For the different phenotypic traits, RIL means were significantly correlated between watering regimes **(Supplemental Table 7)** and with data previously described (Albert et al., 2016a) despite a less severe drought treatment and impact in the present experiment **(Supplemental Table 8).**

2. Differential expression between parental genotypes and F1 hybrid in each of the watering regime and inheritance patterns of gene expression

On the principal components analysis performed on read counts, the first axes distinguished samples according to their genotypes both in the fruit and leaf data, explaining 73% and 56% of the total variation, respectively**.** On the other hand, the second axes separated samples according to watering conditions, explaining 11% and 26% of the total variation in the fruits and leaves, respectively **(Supplemental Figures 9 to 10)**. Also, before studying gene expression patterns between watering regimes, we first characterized differential gene expression between the F1 hybrid, Cervil and Levovil, under each of the watering regimes separately.

Between 11% and 46% of the ~22,000 genes analyzed in the fruits under both watering regimes were differentially expressed in the different genotypic comparisons. Whatever the watering conditions, the comparison between Cervil and the F1 hybrid displayed always at least 1.20 fold lower numbers of differentially expressed genes (7,470 genes under control and 2,384 under drought). Besides, in the transcriptome comparisons between F1 hybrid and parental lines, in fruits, at least twice more genes were differentially expressed in the control than in drought condition **(Figure 2)**.

In the leaves, between 17% and 51% of the ~23,000 genes analyzed under control and drought conditions were differentially expressed in the different comparisons between genotypes. Contrary to the results observed in the fruits, always at least twice more genes were differentially expressed in drought than in control condition in the comparisons

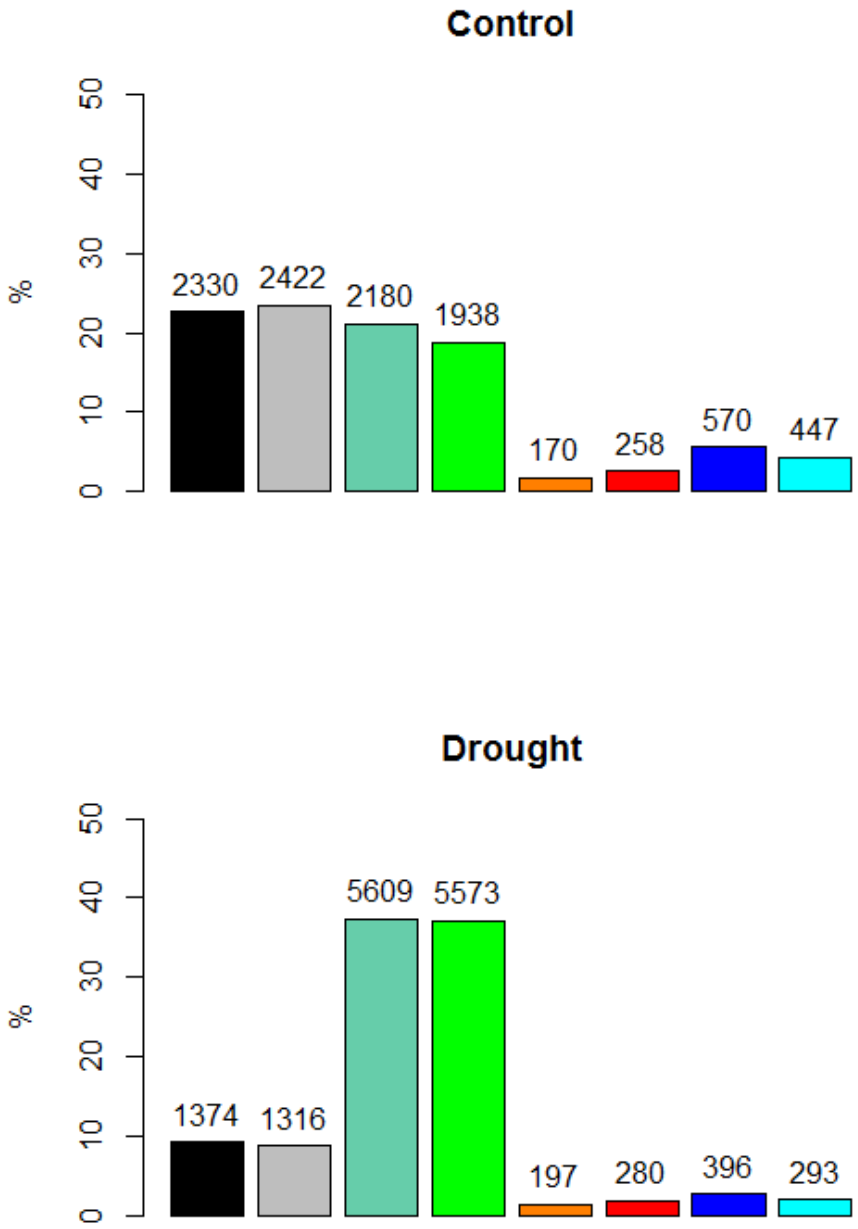


Figure 4: Inheritance of gene expression levels in leaves of the Cer x Lev F1 hybrid under both watering regimes. Bar plots show the number and frequency of genes in each expression inheritance category, under control and drought conditions. A total of 10,315 and 15,038 genes, differentially expressed in one or more genotypic comparisons, were considered for the inheritance assessments, in control and drought conditions, respectively. For color legend see **Figure 3**.

between leaf transcriptomes of the F1 hybrid and both parents **(Figure 2)**.

Classifying gene expression inheritance patterns also evidenced changes in gene expression in F1 hybrid. We focused on genes differentially expressed in one or more genotypic comparisons, i.e. 13,826 and 9,651 genes in the fruits under control and drought conditions, respectively **(Figure 3)**, and 10,315 and 15,038 genes in the leaves under control and drought conditions, respectively **(Figure 4).** We observed all the different forms of altered gene expression, i.e. additive, dominant and recessive, over-dominant and over-recessive. The difference in the proportion of the different inheritance patterns between control and drought conditions in each organ, or between fruits and leaves in each watering condition, were significant when computing Pearson chi-squared tests (all *Pvalues* < 1.0×10^{-3}). Noteworthy, in fruits, only 26% of the assessed genes presented additive inheritance patterns under control condition, whereas up to 57% of the genes presented such pattern under drought condition. The contrary was observed in the leaves, with 46% of the genes showing additive inheritance under control condition, whereas only 18% presented such inheritance pattern under drought condition. Inheritance patterns were conserved between watering regimes for 45% (3,493 over 7,753) and 44% (3,889 over 8,749) of the genes in the fruits and leaves respectively, and conserved between organs for 30% (2,103 over 7,061) and 18% (1,316 over 7,118) of the genes in control and drought conditions, respectively.

3. Differential expression between watering regimes in the parents and F1 hybrid

To assess the extent to which the watering regimes affected gene expression levels, we estimated the number of genes in F1 hybrid and parental accessions differentially expressed between control and drought conditions. In the fruits, we identified 5,034, 125 and 2,049 genes exhibiting significant differences in expression between both watering regimes in the F1 hybrid, Cervil and Levovil, respectively (over a total of 22,204 genes, **Figure 5**). On the other hand, in the leaves, 4,044, 1,323 and 1,271 genes were significantly differentially expressed between watering regimes in the F1 hybrid, Cervil and Levovil, respectively (over a total of 22,871 genes, **Figure 6**). Surprisingly, the F1 hybrid displayed between 3 and 40 fold more differentially expressed genes in response to water deficit in the fruits and leaves in comparison with its parents. Besides, while almost as many genes were differentially expressed between watering regimes in the leaves of Cervil (1,323 genes) and Levovil (1,271 genes), 16 fold less genes were differentially expressed between watering regimes in the

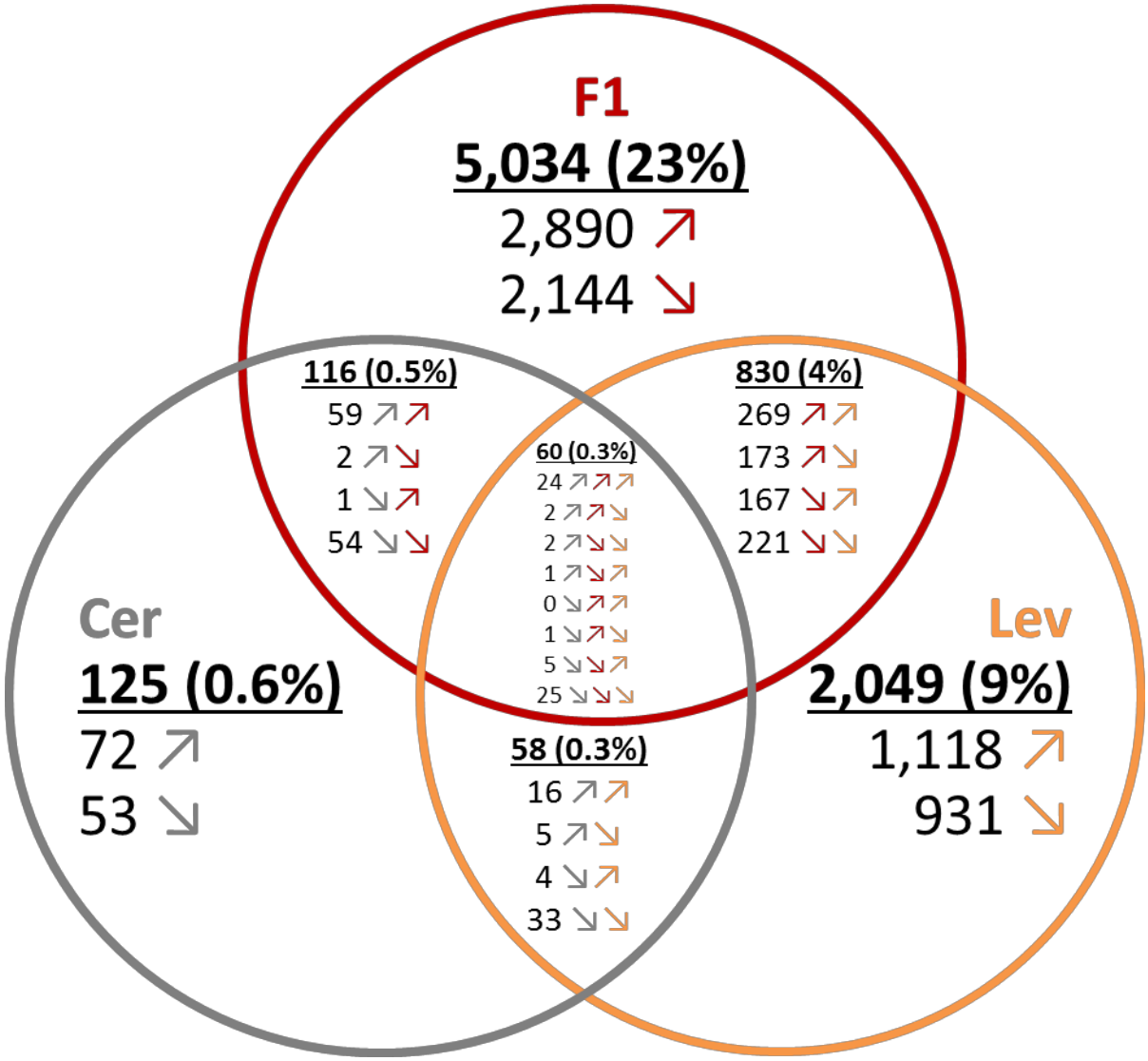


Figure 5: Vein diagram showing the differentially expressed genes between drought and control conditions in fruits for each genotype and comparisons between genotypes. Bold and highlighted numbers indicate the total number and fraction of genes differentially expressed in each comparison among the 22,204 genes analyzed. '↗' indicates the up regulated genes in stress compared to control. '↘' indicates the down regulated genes in stress compared to control.

fruits of Cervil (125 genes) compared to the fruits of Levovil (2,049 genes), suggesting a buffering effect in the Cervil genotype. A total of 4,602 (21% of tested genes) and 7,696 (34% of tested genes) genes presented a significantly different watering regime effect on their expression across the genotypes, in the fruits and leaves, respectively.

We performed gene ontology (GO) enrichment tests for terms related to biological process among the genes differentially expressed between watering regimes in each of the three genotypes and each of the organs. In the fruits of Levovil and the F1 hybrid, GO terms related to "secondary and primary metabolism", "oxidation reduction", "protein modification & transport" and "phosphorylation" were over represented compared to the proportion observed in the whole tomato genome. In Cervil fruits, only the GO terms related to "carbon fixation" were significantly over represented **(Supplemental Figures 11 A to C)**. In the leaves of the three genotypes, GO terms related to "cell redox homeostasis", "cellular regulation & signal transduction", "secondary and primary metabolisms", "protein modification", "catabolism", "ion transport", "phosphorylation" and "response to stimuli" were over represented **(Supplemental Figures 11 D to E)**.

4. Allele specific expression in the F1 hybrid

We used 61,304 exonic SNPs between Cervil and Levovil to identified allele specific expression (ASE) in the F1 hybrid. We detected 780 and 1,475 genes exhibiting ASE over 7,318 genes tested in the fruits of the F1 hybrid under control and drought conditions **(Figure 7A)**, representing 11% and 20% of the genes assessed. On the other hand, in leaves, over 7,653 genes tested, 1,240 (16%) and 1,203 (16%) genes exhibited allele-specific imbalance in the F1 hybrid under control and drought conditions, respectively **(Figure 7C)**.

The ASE ratios were highly correlated between control and drought conditions both in fruits and leaves (Spearman correlation test: r^2_{fruits} =0.87 and *P-value* $_{fruits} < 2.2 \times 10^{-16}$ and r^2_{leaves} = 0.77 and *P-value* $_{leaves} < 2.2 \times 10^{-16}$). A total of 663 and 770 genes exhibited significant ASE in the same direction under both conditions, in the fruits and leaves, respectively, suggesting *cis*-regulatory divergence sustained across watering regimes in both organs **(Figure 7A & D).** Besides, 72% (564 over 780 in control; 1,064 over 1,475 in drought) and 73% (901 over 1,240 in control; 895 over 1,203 in drought) of the significant ASE genes exhibited greater expression levels from the Levovil allele both under the control and drought conditions, in

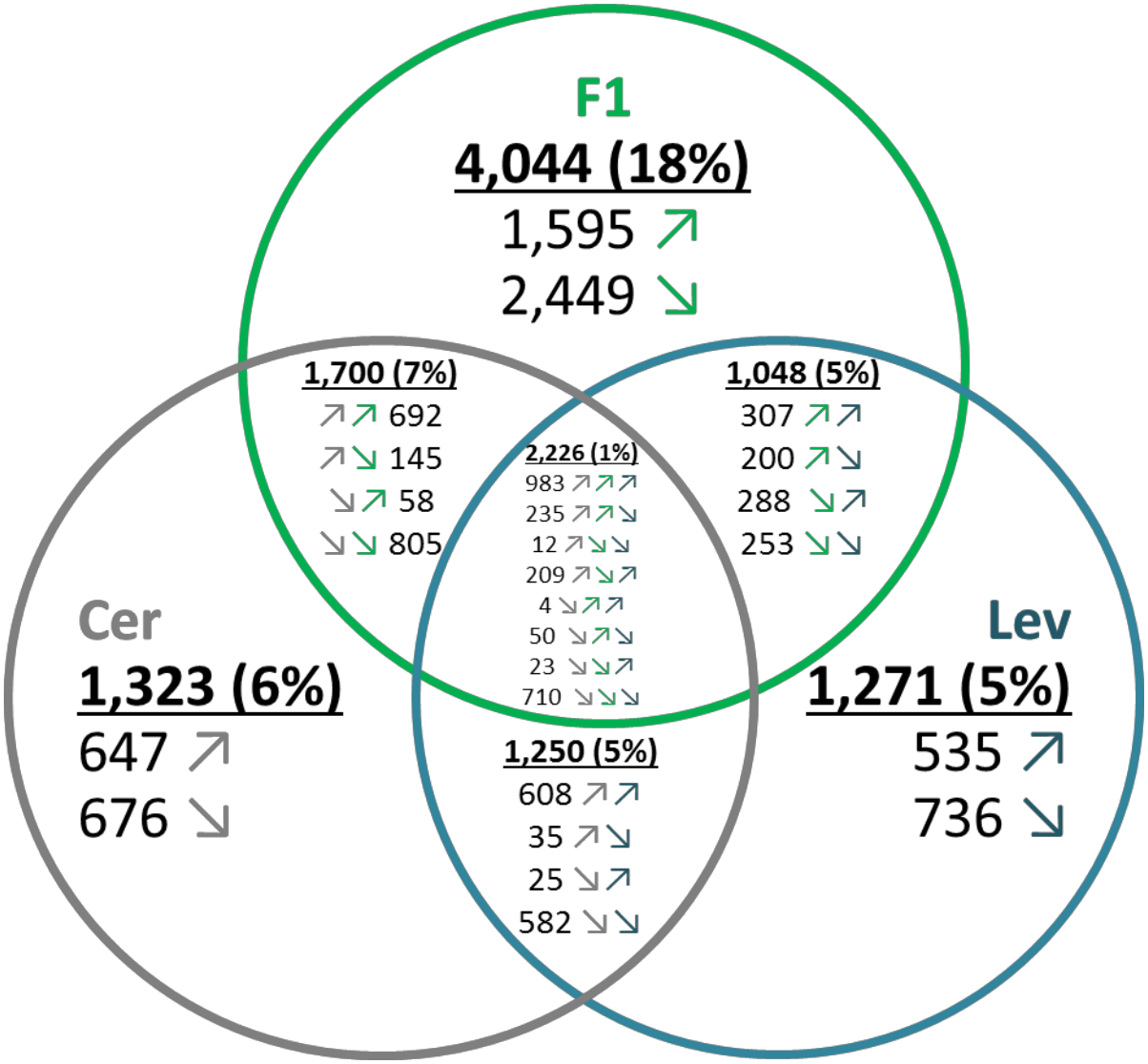


Figure 6: Vein diagram showing the differentially expressed genes between drought and control conditions in leaves for each genotype and comparisons between genotypes. Bold and highlighted numbers indicate the total number and fraction of genes differentially expressed in each comparison among the 22,871 genes analyzed. '↗' indicates the up regulated genes in stress compared to control. '↘' indicates the down regulated genes in stress compared to control.

the fruits and leaves, respectively **(Figure 7A & C).** The non-symmetrical divergence of *cis*-regulatory polymorphisms between both parental alleles (Cervil vs Levovil) could have, in part, resulted from a mapping bias, as Levovil genome was shown to be closer to the tomato reference genome than that of Cervil (Causse et al., 2013). However, the read mapping success rates equivalent between Cervil and Levovil libraries did not support this hypothesis and suggested that our results are unlikely to be severely affected by reference bias **(Supplemental Figures 1 and 2)**.

Among the 5,514 genes evaluated both in the fruits and leaves, 339 and 531 genes exhibited ASE in both organs under control and drought conditions, respectively, among which 98% (323/331) and 97% (517/531) with an allelic imbalance in the same direction. On the other hand, 892 genes and 757 genes presented ASE in only one of the organs assessed, under control and drought conditions, respectively.

5. Regulatory divergences in the F1 hybrid compared to the parents

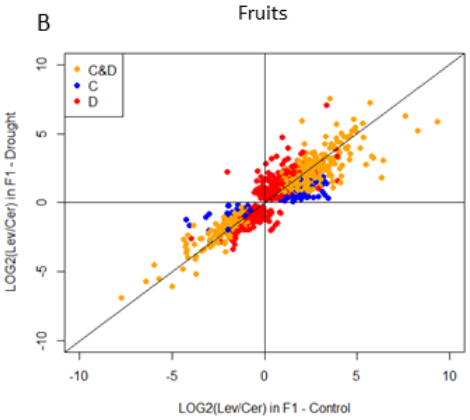
We compared the ASE in the F1 hybrid to the allelic expression in both parental lines to distinguish between *cis*, *trans* and combination of both regulatory divergences (*cis + trans*) at the whole genome scale, in the fruits and the leaves, under both watering regimes **(Figure 8 and Supplemental Figure 12)**. We considered that genes presented *cis* regulatory divergence when the imbalanced allelic expression was conserved between both parents and the F1 hybrid, whereas we considered that genes presented *trans* regulatory divergence when the allelic expression was imbalanced only between parental genotypes. When different imbalanced allelic expressions were observed between parents and hybrid, the genes were considered as presenting a combination of both types of regulatory divergences.

Between 69% and 84% of the ~7,000 tomato genes assessed (7,318 in fruit and 7,653 in leaves) showed no expression divergence between parental lines and no allelic expression changes in the F1 hybrid. Among the remaining genes in the fruits, we identified 658 genes and 1,111 genes (9% and 15%) with a significant *cis* regulatory divergence, 369 genes and 806 genes with a significant *trans* regulatory divergence (5% and 11%) and 122 genes and 364 genes (2% and 5%) presenting a combination of *cis* and *trans* regulatory divergences, under control and drought conditions, respectively **(Figure 8)**. In the leaves, 1,040 genes and 1,022 genes (14% and 13%) presented a significant *cis* regulatory divergence, 359 genes and

A

Fruits		Control			
		No ASE	Lev > Cer	Lev < Cer	Total
Drought	No ASE	5,726	79	38	5,843
	Lev > Cer	579	485	0	1,064
	Lev < Cer	233	0	178	411
	Total	6,538	564	216	7,318

B



C

Leaves		Control			
		No ASE	Lev > Cer	Lev < Cer	Total
Drought	No ASE	5,982	329	139	6,450
	Lev > Cer	322	572	1	895
	Lev < Cer	109	0	199	308
	Total	6,413	901	339	7,653

D

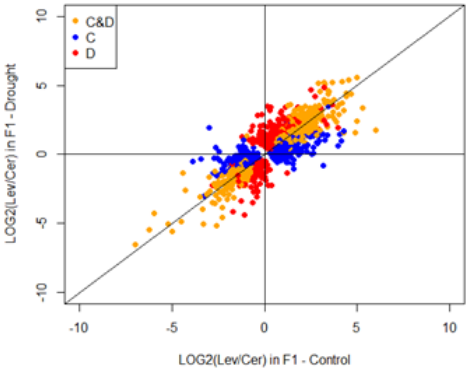


Figure 7. Overview of the numbers of genes showing allele specific expression under control and drought conditions, in fruits and leaves. In tables A and C, dominant allele in the F1 hybrid is indicated (either Lev: Levovil allele, or Cer: Cervil allele), in the fruits and leaves, respectively. Plots B and C display correlation of the allelic ratios in the F1 between watering conditions, in fruits and leaves, respectively (ratio computed after summing the replicates).

317 genes a significant *trans* regulatory divergence (5% and 4%) and 200 genes and 181 genes (3% and 2%) a combination of *cis* and *trans* regulatory divergences, under control and drought conditions respectively **(Supplemental Figure 12)**.

The difference in the proportion of each of the different regulatory classes between control and drought conditions was significant in the fruits (*P-value* < 1.0 X 10^{-3}) but not in the leaves (*P-value = 0.22*). The difference in the proportion of each of the different regulatory classes between leaves and fruits was significant both under control and drought condition (*P-values* < 1.0 X 10^{-3}). Regulatory divergences were conserved between watering regimes for 36% (894 over 2,511) and 42% (906 over 2,179) of the genes in the fruits and leaves, respectively, and conserved between organs for 54% (614 over 1,178) and 67% (790 over 1,178) of the genes in control and drought conditions, respectively. When comparing the frequencies of the inheritance patterns in the different regulation categories, we observed significant differences both in fruits and leaves, under both watering regimes (*P-values* < 1.0 x 10^{-2}) **(Supplemental Figure 13).**

Performing gene ontology (GO) enrichment tests for terms related to biological process among the genes presenting *trans* and *cis+trans* regulatory divergence in the fruits and leaves under both watering regime conditions, we identified a significant enrichment in GO terms related to “regulation of transcription”, “signal transduction” , “cellular regulation”. Among the genes presenting *cis* regulatory divergence, GO terms related to “secondary and primary metabolisms”, “oxidation-reduction process” and “phosphorylation” were significantly enriched. Remarkably, among the genes presenting *trans* regulatory divergence in fruits and leaves under drought condition, we observed an enrichment in terms related to “response to stimulus” and “response to fungus and other organisms”, supporting the hypothesis that *trans* regulation may contribute to the fine-tuning of gene expression related to defense mechanisms under stress conditions **(Supplemental Figure 14).**

6. QTL mapping in the RIL progeny for plant vigor and fruit quality traits

Using 501 SNP genotypes in 124 F7 RILs developed from the cross between the Cervil and Levovil, a total of 46 QTLs were mapped for the thirteen phenotypic traits related to plant vigor and fruit quality (α = 0.10, **Supplemental Tables 9 and 10; Supplemental Figure 15).** The 46 QTLs explained more than 8% of the total phenotypic variance (PVE), with a median

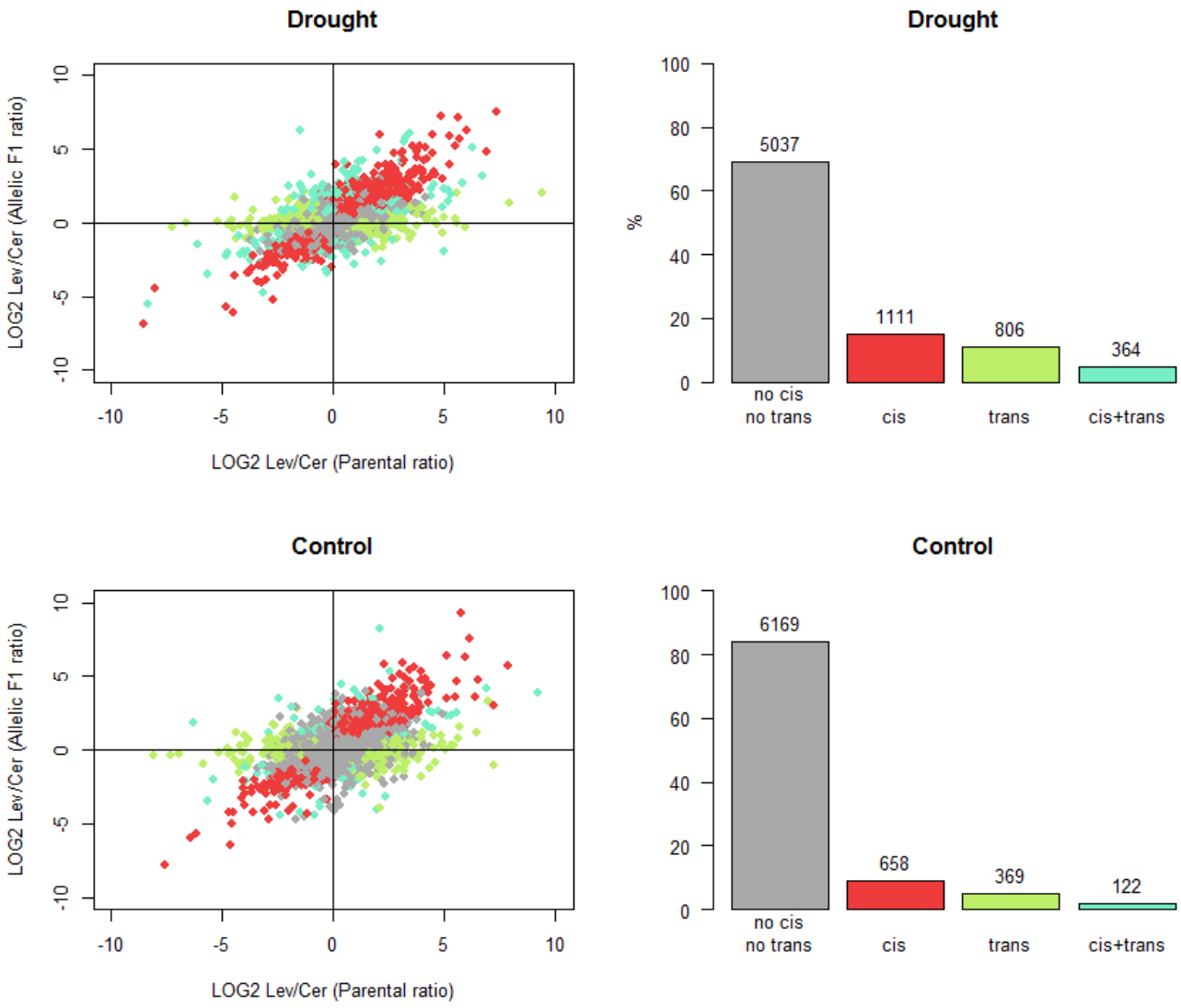


Figure 8: *Cis* and *trans* regulatory divergence between parental genotypes for 7,318 genes the expression of which was measured in fruits through RNA sequencing, under control and drought conditions. The left bar plots display numbers and frequencies of genes in each regulatory category in control and drought conditions. The right plots summarize the relative allelic-specific expression levels in parents and F1 hybrid (parental ratio on x axis and F1 ratio on y axis), in control and drought conditions.

value of 14% and a maximum of 37% for FW.Avi in control condition. The confidence intervals were smaller than 9 Mbp for 70 % of the QTLs, whereas 14 QTLs mapped around the centromeres and their confidence intervals encompassed more than 19 Mbp. Thirty four QTLs colocalized within five clusters on chromosomes 2, 3, 4, 9 and 12.

Thirteen QTLs were detected under control condition only and 13 QTLs under drought condition only. Fourteen QTLs were constitutive as they were detected under both watering regimes **(Supplemental Table 9)**. Six QTLs were significantly interactive between watering treatments, with one of them mapped both with the ANOVA testing procedure and with the plasticity data (MalicFM on chromosome 11, PVE = 17%) **(Supplemental Table 10)**. Among the interactive QTLs, three presented 'antagonist' effects (for Dia, GlucoseDM, MalicFM) and three presented 'differential' effects (one for SSC and two for GlucoseFM) according to the watering regime. Twenty two QTLs confirmed previously identified QTLs in Albert et al. (2016a), suggesting an overall good repeatability of the measurements and conservation of the QTLs despite slightly different drought treatments between both experiments.

7. eQTL mapping in the RIL progeny for 274 transcripts

We selected 274 genes (183 in fruits and 91 in leaves) according to (1) their annotation and potential involvement in response to water deficit according to literature **(Supplemental Figures 3 and 4)**, and/or (2) their location in the main QTL and QTL by watering regime interaction regions identified in Albert et al. (2016a) and confirmed in the present study. Expression profiles measured in the RILs were used to investigate the genetic basis of mRNA expression variation in the fruits and leaves, under control and drought conditions. Using an alpha threshold equal to 0.10, we identified at least one significant eQTL for 117 of the 183 transcripts investigated (64%) in fruits **(Supplemental Tables 11 and 12)** and for 36 of the 91 transcripts investigated (40%) in the leaves **(Supplemental Tables 13 and 14).**

A total of 246 eQTLs (190 in fruits and 56 in leaves) were detected and between 1 and 6 eQTLs were associated to the different transcripts (average of 1.60). The percentage of variation explained (PVE) by the different eQTLs varied from 4.45 to 67.40%. A total of 121 eQTLs had confidence intervals smaller than 10 Mbp (including between 1 and 3,300 genes), whereas 123 eQTLs mapped in centromeric regions encompassed more than 10 Mbp (including between 500 and 4,300 genes). Among the detected eQTLs, 47 were identified

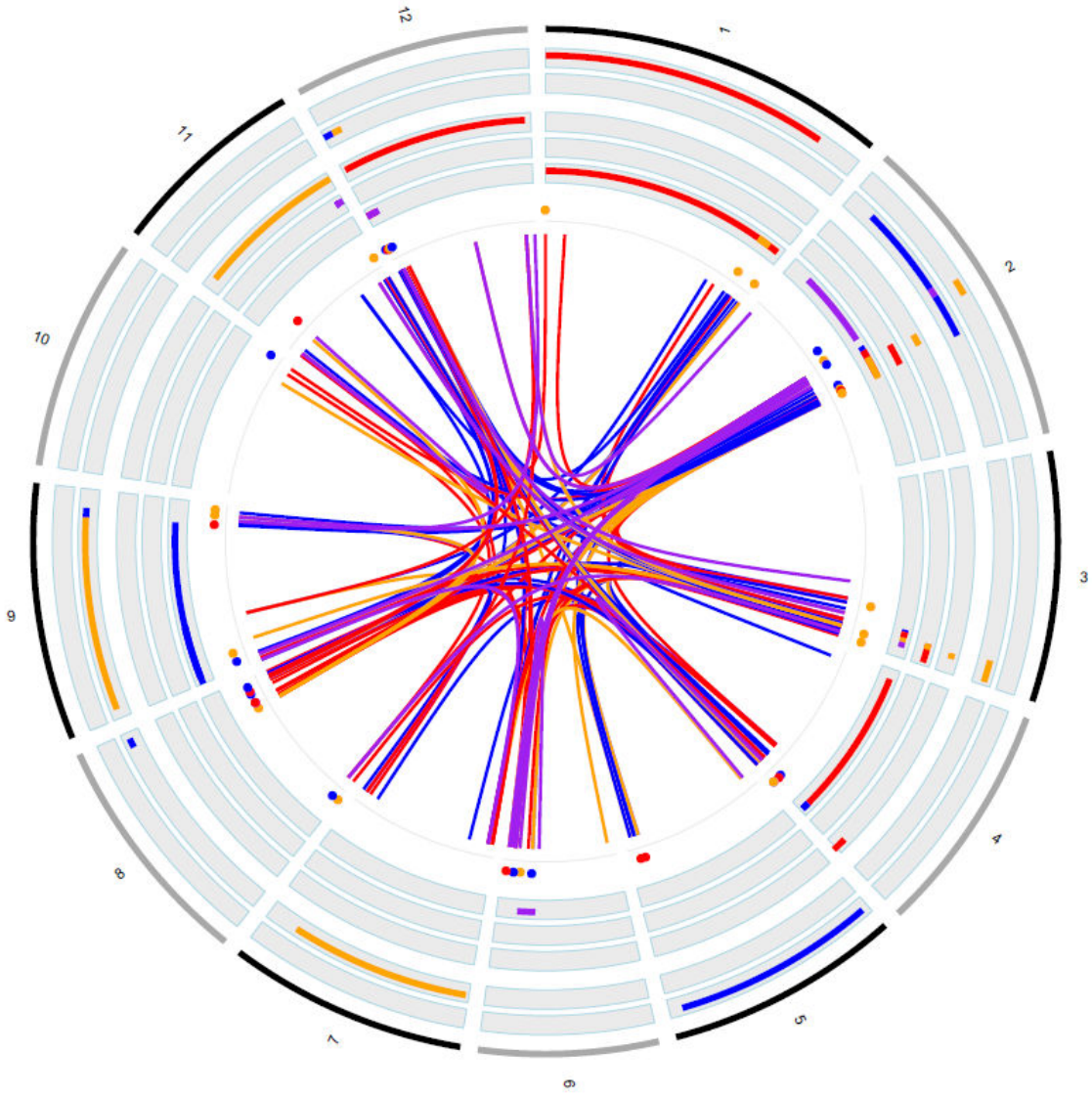


Figure 9: Overview of the QTLs and fruit eQTLs identified on the twelve chromosomes of the tomato genome by linkage analysis in the RILs. The diagram displays the twelve chromosomes of the tomato genome proportionally to their physical size (assembly 2.5). In five first layers, lines represent QTLs detected for phenotypic traits. From outside to inwards: (1) flw, (2) vigor traits (Ht and Dia), (3) FW, (4) acid traits (pH, MalicFM, MalicDM, CirticFM and CitricDM) and (5) sugar traits (DMW, SSC, GlucoseFM, GlucoseDM, FructoseFM and FructoseDM). In the inner part, links represent *trans acting eQTLs* and dots *local eQTLs*, detected on expression data measured in fruits. Colors indicate QTL and eQTL types. Orange: *constitutive*. Blue: *control specific*. Red: *drought specific*. Purple: *interactive*.

under both watering conditions ('constitutive'), 81 under control condition only ('control specific') and 88 under drought condition only ('drought specific'). Thirty additional eQTLs were interactive between watering regimes, among which 11 with 'antagonist' effects and 19 with 'differential' effects according to the watering regimes.

Considering the distance between each eQTL and their regulated target, 107 corresponded to *local* eQTLs (25 constitutive, 62 control specific, 19 drought specific and 1 interactive) whereas 139 corresponded to *trans* acting eQTLs (22 constitutive, 19 control specific, 69 drought specific and 29 interactive) **(Figure 9 and Supplemental Figure 16).** The average PVE of the *local* eQTLs was 22.30% (sd = 9.13), against 13% (sd = 3.26) for the *trans* acting eQTLs.

Among the 153 genes for which we mapped eQTLs in the fruits and/or in the leaves, 54 were also assessed for allele specific expression. For these 54 genes, 89 eQTLS (34 local and 55 trans) were mapped in the RILs and 75 of them (83%, 28 local and 35 trans) confirmed the regulation patterns observed in the allele specific expression assay, 41 of which (55%, 18 local and 23 trans) presented consistent effects between qPCR and RNA sequencing data **(Supplemental Tables 11 and 14)**.

8. Connecting gene expression to fruit quality phenotypes

Nineteen of the genes selected for eQTL mapping were related to sugar metabolism **(Supplemental Tables 4 and 5)** and eleven of them were significantly associated to one to six eQTLs. These eleven genes corresponded to three 'apoplastic cell wall invertases' (Lin5 Solyc09g010080, Lin6 Solyc10g083290 and Lin9 Solyc08g079080), three 'neutral invertases' (Solyc01g111100, Solyc10g083290 and Solyc11g020610), one 'fructokinase' (FRK3 Solyc02g091490), one 'acid invertase' (Solyc03g083910), one 'invertase inhibitor' (Solyc01g088590), one 'fructose-1-6-bisphosphate' (Solyc04g071340) and one 'sucrose synthase' (Solyc12g009300) **(Table 2)**. Remarkably, we observed what seemed to be a complex cross talk between them **(Figure 10).** Seven of the eleven genes were associated to *local* eQTLs with medium to high effects (Solyc03g083910 PVE = 40%; Lin6 PVE = 37%), whereas eight of them were associated with at least one *trans* eQTL. A genomic region of 14.41 Mbp at the bottom of chromosome 2 (from 40Mbp to 55Mbp; ~2,000 genes), including the 'frutokinase 3' gene, concentrated four *trans* eQTLs for three of the eleven genes (Lin5, Lin9 and Solyc03g083910) and could constitute a master regulatory region.

Table 2. Legend: blue for control specific, red for drought specific, orange for constitutive, purple for interactive. Grey: QTLs identified in others experiments).

Genes	Annotation	Organe	eQTL	Positive allele	ASE	Correlation to phenotypes in RILs	Colocalisation with phenotypic QTLs
Solyc01g088590	Invertase inhibitor	Fruit	**1 TRANS (LG3)**	LEV	NA	-0.20 SSC -0.20 DMW	DMW
Solyc01g111100	Neutral invertase	Fruit	1 TRANS (LG8)	LEV	NA (in leaves: NA)	+0.27 FructoseDM	**FructoseFM**
Solyc02g091490	Frutokinase3	Fruit	1 LOCAL (PVE=10%)	CER	No CIS, No TRANS	-0.24 pH	FW, DMW, SSC, **FructoseFM,** **MalicDM**
Solyc03g083910	Acid invertase	Fruit	1 LOCAL (PVE=40%) + **1 TRANS (LG2)**	CER	TRANS	+0.19 GlucoseDM +0.17 FW	*FW control* *(RIL, Albert et al. 2016a)*
Solyc04g071340	Fructose-1 6-bisphosphatase	Fruit	1 TRANS (LG3)	CER	NA (in leaves: CIS+TRANS)	**-0.19 GlucoseFM** **+0.20 FW** **+0.22 MalicDM**	**FructoseDM**
Solyc06g065210	Neutral invertase	Leaf	1 TRANS (LG11)	LEV	No CIS, No TRANS	-0.28 MalicFM -0.22 MalicDM	**GlucoseDM**
Solyc08g079080	Acid invertase	Fruit	1 LOCAL (PVE=18%) + 5 TRANS (LG2, 3, 9, 10, 11)	LEV	NA	-0.23 SSC -0.30 DMW +0.23 FructoseDM **+0.32 FW** **-0.27 GlucoseFM** **+0.29 MalicDM** **+0.30 pH**	*GlucoseDM* *control & drought* *(GWA, Albert et al. 2016b)*
Solyc09g010080	Lin5 Appoplastic cell wall invertase	Fruit	1 LOCAL (PVE=10%) + 1 TRANS (LG2) + **1 TRANS (LG2)**	LEV	TRANS	+0.26 SSC +0.24 DMW +0.23 GlucoseFM +0.19 FructoseFM +0.20 pH -0.20 MalicDM	DMW SSC FructoseFM GlucoseFM
Solyc10g083290	Lin6 Appoplastic cell wall invertase	Leaf	1 LOCAL (PVE=37%)	LEV	NA	+0.25 GlucoseFM +0.22 FructoseFM +0.20 GlucoseFM	*SSC, FructoseFM, GlucoseFM* *control & drought* *(GWA, Albert et al. 2016b)*
Solyc11g020610	Neutral invertase	Fruit	**1 LOCAL (PVE=10%)** + **1 TRANS (LG11)**	LEV	NA	**-0.22 MalicDM** **-0.19 FW**	FW
Solyc12g009300	Sucrose synthase	Fruit	**1 LOCAL** + 1 TRANS (LG4)	LEV	**CIS+TRANS** & TRANS	+0.20 FructoseDM -0.22 DMW	**GlucoseFM,** **FructoseFM,** **FW, GlucoseFM**

A second genomic region at the bottom of chromosome 3 (from 67Mbp to 71Mbp; ~400 genes, including the acid invertase Solyc03g083910) concentrated three *trans* eQTL for Solyc01g088590, FRK3 and Lin9 and could represent a second regulatory node.

We observed significant correlation ($r^2 > 0.20$ and *P-values* < 0.05) between the expressions of the eleven sugar related genes in cell expansion fruits or young leaves and phenotypes related to fruit quality (DMW, SSC, Malic, Glucose, Fructose) assessed on mature fruits in the RILs, under control and/or drought conditions. The strongest correlations were observed for the apoplastic cell wall invertases (Lin5, Lin6 and Lin9) and SSC, DMW, fructoseFM and glucoseFM content in the fruits, in accordance with Fridman et al. (2000) who described Lin5 as a major QTL controlling sugar content in tomato fruits. Furthermore, the genomic locations of eight of the eleven genes colocalized with sugar content phenotypic QTLs identified in the present study. Solyc03g083910, Solyc08g079080 and Solyc10g083290 did not colocalize with any phenotypic QTL identified in our study, but colocalized with sugar QTLs identified in Albert et al. (2016a) using the same population or in Albert et al. (2016b) using genome wide association mapping in a diverse population composed of small fruited tomato accessions.

Three more genes seemed promising candidates regarding the strong correlations of their expressions in fruit and fruit quality phenotypes in mature fruits, in the RILs. The first one, Solyc03g115920 is coding for a '**Zinc finger protein-like protein'** at the bottom of chromosome 3. Its expression was strongly correlated to FW (Pearson r^2; -0.22 in C; -0.38 in D), pH (-0.28 in D), DMW (+0.17 in C; +0.40 in D) and GlucoseDM (-0.31 in C) under control and/or drought conditions. This gene presented a *local* eQTL with a constitutive and strong effect (PVE > 57%) in the eQTL mapping analysis, which was confirmed in the ASE assay. This *local* eQTL colocalized with phenotypic QTLs detected for FW (C&D), DMW (C&D), SSC (D), GlucoseFM (D), GlucoseDM (C) and MalicDM (D). The second promising candidate gene was Solyc04g077050 coding for an '**amino acid permease 6'** gene at the bottom of chromosome 4, the expression of which was strongly correlated to FW (-0.38 in C; -0.37 in D), SSC (+0.41 in C; +0.39 in D), DMW (+0.40 in C; +0.46 in D), FructoseFM (+0.18 in C; +0.29 in D), MalicDM (-0.29 in C; -0.38 in D) and Glucose FM (+0.30 in C; +0.37 in D). The eQTL analysis revealed that Solyc04g077050 was controlled by a *local* eQTL (PVE=11%) and a *trans* acting eQTL (on chromosome 3, PVE=12%), both specific to control condition. The local eQTL colocalized with

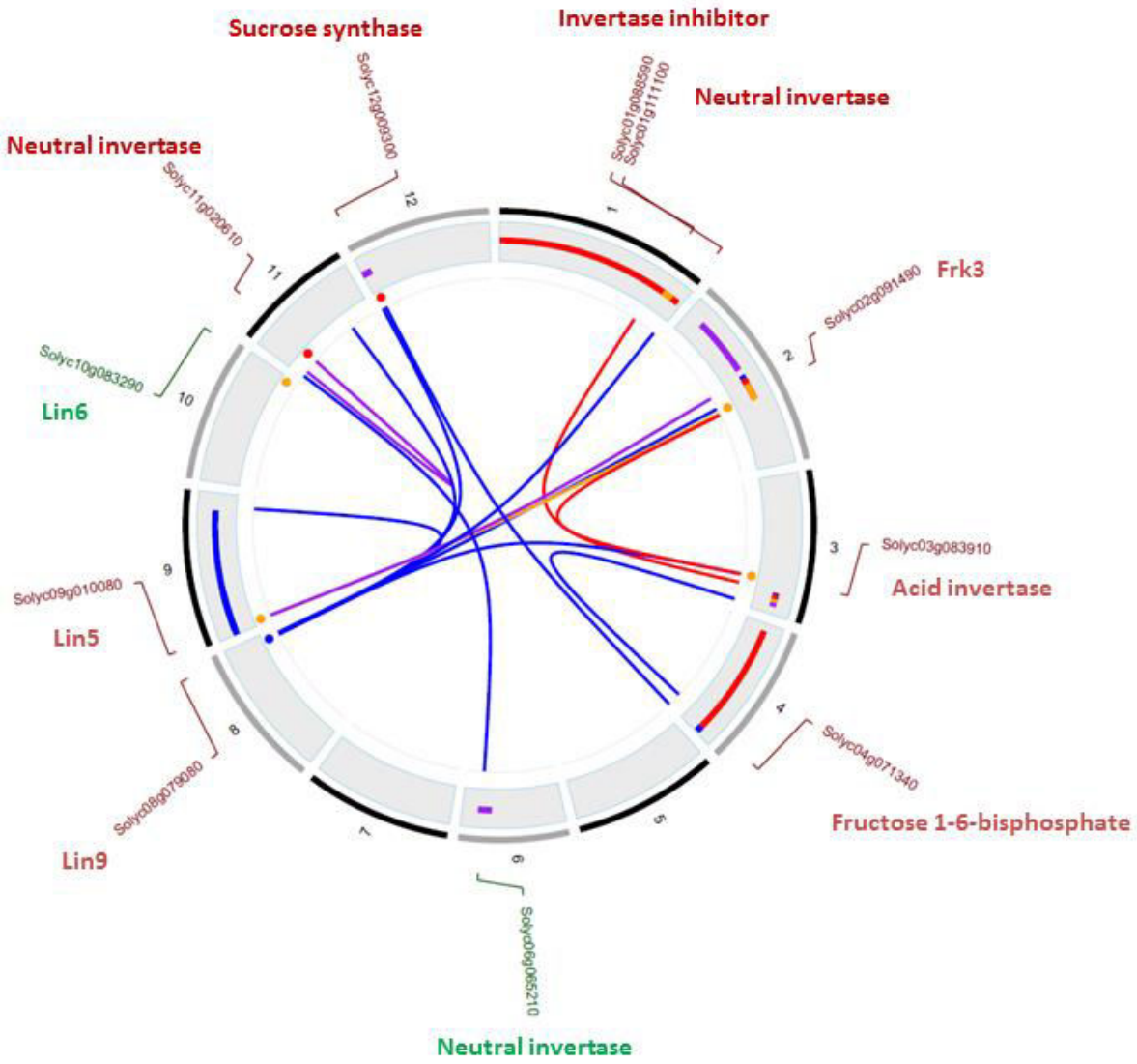


Figure 10: Overview of the cross talk between eleven sugar related genes. The diagram displays the twelve chromosomes of the tomato genome proportionally to their physical size (assembly 2.5). In the first layer, lines represent QTLs detected for sugar traits (DMW, SSC, GlucoseFM, GlucoseDM, FructoseFM and FructoseDM). In the inner part, links represent *trans acting eQTLs* and dots *local eQTLs*, detected on the expression of eleven sugar related gene measured in fruits (gene names in reds) and leaves (gene names in green). Colors indicate QTL and eQTL types. Orange: *constitutive*. Blue: *control specific*. Red: *drought specific*. Purple: *interactive*.

phenotypic QTLs detected for FW (D), DMW (C), FructoseFM (C) and FructoseDM (D) in the present study. Finally, the expression of Solyc04g082500, coding for an **'ATP binding / serine-threonine kinase'** on chromosome 4, was correlated to FW (-0.39 in C; -0.35 in D), SSC (+0.32 in C; + 0.29 in D), DMW (+0.42 in C; +0.34 in D), pH (-0.23 in C; -0.29 in D), MalicDM (-0.19 in C; -0.31 in D) and FructoseFM (+0.18 in C; +0.29 in D). A local eQTL with a constitutive strong effect (PVE > 39%) and a *trans* eQTL with a constitutive moderate effect (PVE=13%) were identified as controlling the expression of Solyc04g082500. The local eQTL was colocalized with phenotypic QTLs for FW (D), DMW (C&D) and FructoseFM (C).

Discussion

1. Variations in gene expression in two organs of three genotypes grown under two contrasted watering regimes

First, we showed a large number of genes and gene functions impacted at the transcript level by water stress, organ and genotypes. This is consistent with previous studies on the impact of water stress on transcriptomic variations (Seki et al., 2002; Rabbani et al., 2003; Des Marais et al., 2012; Albert et al., 2016a) and transcriptome specificities in various organs (Libault et al., 2010; Matas et al., 2011; Koenig et al., 2013; Slane et al., 2014; Pattison et al., 2015). The impact of genetic background was more rarely studied, but the large number of variations identified in our study underlined how much it deserves to be taken into account. Notably, the fruit of the small fruited accession (Cervil) appeared much less affected by the watering regime at the transcript level than the two other genotypes, whereas the leaf transcriptome of the same genotype showed a range of variations in response to water deficit comparable to the two others genotypes. This reflects a different source to sink relationship in the cherry tomato. Finally, the F1 hybrid displayed between 3 and 40 fold more differentially expressed genes in response to water deficit in the fruits and leaves in comparison with its parents. This can be related to the large amount of non-additive inheritance detected among the three genotypes.

2. Inheritance of gene expression and allele specific expression

Transcript abundance can be considered as any quantitative trait. Their heritability and inheritance vary from one gene to the other, as well as the impact of environmental conditions. In tomato, until now only inheritance studies of phenotypic, metabolic and

proteomic traits were performed by (Pascual et al., 2013) and (Steinhauser et al., 2011). They showed that heterosis was rare, with few plant and fruit traits showing over dominance. Our results concerning plant vigor and fruit quality phenotypes confirmed this pattern, with a conservation of trait inheritance through watering regimes. Concerning gene expression inheritance, we observed different inheritance patterns according to the watering regime and the organ. The expression of roughly half of the genes presented over dominant or over recessive inheritance.

Hybrids are widely used in modern agriculture, either for heterosis (the advantage of a hybrid compared to both parents) of for combination of dominant traits. This is particularly the case for tomato. In our study ASE concerned 11% to 20% of the genes assessed, depending on the organ and/or watering regime. This phenomenon could be at the origin of heterosis. In maize, attempts were made to investigate the heterosis phenomenon largely observed in this species in regards to the dominance variations found in ASE studies. However, until now, this led to inconclusive results and provided only speculations on the precise relationship between dominance of ASE patterns and hybrid vigour (Stupar and Springer, 2006; Springer and Stupar, 2007a; Springer and Stupar, 2007b).

3. Genetic control of regulatory divergence in response to water deficit

Using 61,304 exonic SNPs identified in 11,719 genes polymorphic between parental genotypes and comparing allele specific expression in the hybrid with allelic expression in both parents, we assessed regulatory divergence at a whole transcriptome scale. Although almost two million SNP could be identified between Cervil and Levovil (Causse et al., 2013), most of them were in non-coding regions and only one third of the genes had informative SNPs to measure ASE in the F1. The approach was nevertheless powerful enough to identify 1,982 *cis* regulated genes, 1,450 *trans* regulated genes and 715 genes presenting a combination of both types of regulatory divergence. A special analysis should be performed on the regulating regions of the genes showing *cis* regulatory divergences. The ultimate identification of the mutations responsible for *cis* regulatory divergences may require a larger number of F1 to be tested, as proposed by Kang et al. (2016).

Swinnen et al. (2016) reported that in different crop species many phenotypic changes reside in *cis*-regulatory elements that control the expression of an unmodified coding

sequence. Sequence variation in *cis* regulatory elements could impact gene expression levels, but also developmental timing and tissue specificity of expression. Besides, mutations in *cis* regulatory elements may be favored by domestication in contrast to mutations in coding sequences due to less detrimental pleiotropic effects. In tomato, for example the locule number (lc) and fasciated (fas) mutations responsible for phenotype variation are both located downstream of a gene. Selection of lc and fas thus fine-tuned the expression of regulators in a network controlling floral meristem size, which resulted in supernumerary locules (Muños et al., 2011; Sánchez-Rodríguez et al., 2011). The same was observed for FW2.2 and FW3.2, two major cloned QTLs controlling tomato fruit size (Frary, 2000; Zhang et al., 2012).

On the contrary to eQTL mapping, ASE analysis does not allow localizing regulatory variant acting in *trans*. We thus completed the ASE approach by studying the variation of 274 transcripts in RILs. We could identify 246 eQTL, among which 107 were local and 139 were *trans* acting. As the choice of the genes was skewed for either genes involved in water stress or fruit quality or genes located in phenotypic QTL regions (based on results from Albert et al. 2016a, Albert et al. 2016b), it is difficult to assess the distribution of eQTL. Nevertheless, several *trans* eQTL were located on chromosomes 4 and 8 and could correspond to regulatory hubs. The distinction between local and distant regulatory element is somewhat arbitrary if a regulatory element maps to the same chromosome as the gene it regulates. It depends strongly on the resolution of the method used to map the variation (Emerson and Li, 2010). In the present experiment, the population size was limited, so QTL confidence intervals were often very large. Nevertheless, many local eQTLs with very strong effect (LOD > 6) were identified and probably corresponded to *cis* regulated genes. The *local* eQTLs were usually constitutive over watering conditions with moderate to high effects, whereas *trans* acting eQTLs mostly presented low to moderate effects varying with watering conditions. Remarkably, 97% of the interactive eQTLs corresponded to *trans* acting eQTLs, i.e. eQTLs distant from their regulated targets. Besides, 79% of the drought specific eQTLs were *trans* acting whereas 76% of the control specific eQTLs were local eQTLs (either *trans* acting elements close to the gene they regulate or *cis* acting elements). These different proportions may indicate that the drought treatment resulted in the activation of *trans* acting regulatory elements specific to plant adaptation to abiotic constraint.

4. A complex regulatory network of invertase genes and a few candidates

Focusing on genes involved in sugar metabolisms, we showed a strong interrelation between invertase genes, whether the apoplastic (cell wall), cytoplasmic (neutral) and vacuolar (acid). All these genes seem to be related to the variation of source to sink relationship. A putative role could be also played by FRK3 on chromosome 2 which is located in the confidence intervals of *trans* eQTLs for several of these genes. FRK3 was detected as a candidate gene for sugar content at the proteome level (Pascual et al., 2013) and fine mapping experiments confirmed its role (unpublished data). Remarkably, the expression of the neutral invertase of chromosome 10 (Lin6) in leaves was strongly correlated with sugar content in mature fruits. Albert et al. (2016b) proposed this gene as a possible candidate for sugar content variation in their association mapping study.

The strong involvement of other genes was underlined based on the strong correlations with phenotype data. Zinc finger proteins are transcription factors regulating transcription by DNA binding. They are numerous in the genome and uncover a large range of functions, but some have been shown to be involved in drought stress response (Li et al., 2013). The strong local eQTL (PVE> 57%) suggests a major mutation effect controlling its variation. The amino peptidase 6 expression is also correlated with several fruit traits. Kohl et al. (2012) showed that this gene is related to the source to sink relationship through nitrogen transport. The strong relation between nitrogen and carbon metabolism has been already underlined (Prudent et al., 2011). A polymorphism in this gene could thus significantly impact the phenotypes studied.

Acknowledgments

We acknowledge the experimental teams of UR GAFL for their help in phenotyping plants. We thank especially Yolande Carretero, Margaux Duberos, Justine Gricourt and Gisèle Riqueau for their involvement in the experimentations. Thanks to Christopher Sauvage, Marie-Laure Martin-Magnette, Guillem Rigaill and Joel Chadoeuf for scripts sharing and for their advices concerning RNA sequencing data analysis and statistical analysis. The ANR project ADAPTOM supported this work. E.A. was supported by an INRA PhD fellowship.

Author contributions

E.A. supervised sample collection and phenotypic measurements, analyzed data and wrote the manuscript. R.D. performed RNA extraction and RNA-seq library preparation. M.L. and S.S. supervised library preparation and sequencing. M.B, J.P.B. and F.B. developed the program for qPCR primers design and analyzed RNA-seq reads. C.P. and V.G. performed microfluidigm qPCR experiment. J.J.G developed script for ASE analysis. M.C. supervised the project, built the experimental design and revised the manuscript. All authors discussed the results and commented the manuscript. Authors declared no conflict of interest in the authorship and publication of this document.

References

Albert E., Gricourt J., Bertin N., Bonnefoi J., Pateyron S., Tamby J.-P., Bitton F., Causse M. (2016a) Genotype by watering regime interaction in cultivated tomato: lessons from linkage mapping and gene expression. Theor Appl Genet **129**: 395–418

Albert E., Segura V., Gricourt J., Bonnefoi J., Derivot L., Causse M. (2016b) Association mapping reveals the genetic architecture of tomato response to water deficit: focus on major fruit quality traits. J Exp Bot **67**: 6413–6430

Anders S., Pyl P.T., Huber W. (2014) 10D-HTSeq – A Python framework to work with high-throughput sequencing data. Bioinformatics **31**: 0–5

Anderson J.T., Willis J.H., Mitchell-olds T. (2012) Evolutionary genetics of plant adaptation. Trends Genet **27**: 258–266

Babraham B. (2011) FastQC A quality control tool for high throughput sequence data.

Benjamini Y., Hochberg Y. (1995) Controlling the false discovery rate : a practical and powerful approach to multiple testing. J R Stat Soc **57**: 289–300

Boyle E.I., Weng S., Gollub J., Jin H., Botstein D., Cherry J.M., Sherlock G. (2004) GO:TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics **20**: 3710–3715

Brem R.B., Yvert G., Clinton R., Kruglyak L. (2002) Genetic dissection of transcriptional regulation in budding yeast. Science (80-) **296**: 752–755

Bystrykh L., Weersing E., Dontje B., Sutton S., Pletcher M.T., Wiltshire T., Su A.I., Vellenga E., et al (2005) Uncovering regulatory pathways that affect hematopoietic stem cell function using “genetical genomics.” Nat Genet **37**: 225–232

Castel S.E., Levy-Moonshine A., Mohammadi P., Banks E., Lappalainen T. (2015) Tools and best practices for data processing in allelic expression analysis. Genome Biol **16**: 195

Causse M., Desplat N., Pascual L., Le Paslier M.-C., Sauvage C., Bauchet G., Bérard A., Bounon R., et al (2013) Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. BMC Genomics **14**: 791

Chesler E.J., Lu L., Shou S., Qu Y., Gu J., Wang J., Hsu H.C., Mountz J.D., et al (2005) Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. Nat Genet **37**: 233–242

Cheung V.G., Spielman R.S., Ewens K.G., Weber T.M., Morley M., Burdick J.T. (2005) Mapping determinants of human gene expression by regional and genome-wide association. Nature **437**: 1365–1369

Combes M.C., Hueber Y., Dereeper A., Rialle S., Herrera J.C., Lashermes P. (2015) Regulatory divergence between parental alleles determines gene expression patterns in hybrids. Genome Biol Evol **7**: 1110–1121

Cubillos F., Stegle O., Grondin C., Canut M., Tisné S., Gy I., Loudet O. (2014) Extensive cis-regulatory variation robust to environmental perturbation in *Arabidopsis*. Plant Cell **26**: 4298–310

Cubillos F., Yansouni J., Khalili H., Balzergue S., Elftieh S., Martin-Magniette M.-L., Serrand Y., Lepiniec L., et al (2012) Expression variation in connected recombinant populations of *Arabidopsis thaliana* highlights distinct transcriptome architectures. BMC Genomics **13**: 117

DeCook R. (2005) Genetic regulation of gene expression during shoot development in Arabidopsis. Genetics **172**: 1155–1164

Druka A., Potokina E., Luo Z., Jiang N., Chen X., Kearsey M., Waugh R. (2010) Expression quantitative trait loci analysis in plants. Plant Biotechnol J **8**: 10–27

Emerson J., Li W.-H.H. (2010) The genetic basis of evolutionary change in gene expression levels. Philos Trans R Soc Lond B Biol Sci **365**: 2581–2590

Emerson J.J., Hsieh L.-C., Sung H.-M., Wang T.-Y., Huang C.-J., Lu H.H.-S., Lu M.-Y.J., Wu S.-H., et al (2010) Natural selection on *cis* and *trans* regulation in yeasts. Genome Res **20**: 826–836

Frary A. (2000) fw2.2: A quantitative trait locus key to the evolution of tomato fruit size. Science (80-) **289**: 85–88

Fridman E., Pleban T., Zamir D. (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proc Natl Acad Sci U S A **97**: 4718–4723

Garcia C., Renard C. (2014) Validation des dosages enzymatiques des sucres (glucose, fructose, saccharose) et acides (acide citrique et malique) par un spectrophotomètre avec lecteur de microplaques. Le Cah des Tech l'INRA **81**: 1–18

Giovannoni J. (2001) Molecular biology of fruit maturation and ripening. Annu Rev Plant Physiol Plant Mol Biol **52**: 725–749

Gur A., Semel Y., Osorio S., Friedmann M., Seekh S., Ghareeb B., Mohammad A., Pleban T., et al (2011) Yield quantitative trait loci from wild tomato are predominately expressed by the shoot. Theor Appl Genet **122**: 405–20

Hammond J.P., Mayes S., Bowen H.C., Graham N.S., Hayden R.M., Love C.G., Spracklen W.P., Wang J., et al (2011) Regulatory hotspots are associated with plant gene expression under varying soil phosphorus supply in *Brassica rapa*. Plant Physiol **156**: 1230–41

He F., Arce A.L., Schmitz G., Koornneef M., Novikova P., Beyer A., de Meaux J. (2016) The footprint of polygenic adaptation on stress-responsive Cis -regulatory divergence in the *Arabidopsis* genus. Mol Biol Evol **33**: 2088–2101

Hubner N., Wallace C.A., Zimdahl H., Petretto E., Schulz H., Maciver F., Mueller M., Hummel O., et al (2005) Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. Nat Genet **37**: 243–253

Jiménez-Gómez J.M., Wallace A.D., Maloof J.N. (2010) Network analysis identifies ELF3 as a QTL for the shade avoidance response in *Arabidopsis*. PLoS Genet. doi: 10.1371/journal.pgen.1001100

Kang E.Y., Martin L., Mangul S., Isvilanonda W., Zou J., Ben-David E., Han B., Lusis A.J., et al (2016) Discovering SNPs regulating human gene expression using allele specific expression from RNA-Seq Data. Genetics **204**: 1057–1064

Keurentjes J.J.B., Fu J., Terpstra I.R., Garcia J.M., van den Ackerveken G., Snoek L.B., Peeters A.J.M., Vreugdenhil D., et al (2007) Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. Proc Natl Acad Sci U S A **104**: 1708–13

Kiekens R., Vercauteren A., Moerkerke B., Goetghebeur E., Van Den Daele H., Sterken R., Kuiper M., van Eeuwijk F., et al (2006) Genome-wide screening for cis-regulatory variation using a classical diallel crossing scheme. Nucleic Acids Res **34**: 3677–3686

Kim D., Pertea G., Trapnell C., Pimentel H., Kelley R., Salzberg S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol **14**: R36

Kirst M. (2004) Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of *Eucalyptus*. PLANT Physiol **135**: 2368–2378

Koenig D., Jiménez-Gómez J.M., Kimura S., Fulop D., Chitwood D.H., Headland L.R., Kumar R., Covington M.F., et al (2013) Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. Proc Natl Acad Sci U S A **110**: E2655-62

Kohl S., Hollmann J., Blattner F.R., Radchuk V., Andersch F., Steuernagel B., Schmutzer T., Scholz U., et al (2012) A putative role for amino acid permeases in sink-source communication of barley tissues uncovered by RNA-seq. BMC Plant Biol **12**: 154

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., et al (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**: 2078–2079

Li W.T., He M., Wang J., Wang Y.P. (2013) Zinc finger protein (ZFP) in plants-A review. Plant Omics **6**: 474–480

Libault M., Farmer A., Joshi T., Takahashi K., Langley R.J., Franklin L.D., He J., Xu D., et al (2010) An integrated transcriptome atlas of the crop model Glycine max, and its use in comparative analyses in plants. Plant J **63**: 86–99

Liu Z., Yang J., Xu H., Li C., Wang Z., Li Y., Dong X., Li Y. (2014) Comparing computational methods for identification of allele-specific expression based on next generation sequencing data. Genet Epidemiol **38**: 591–598

Love M.I., Huber W., Anders S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**: 550

Des Marais D.L., McKay J.K., Richards J.H., Sen S., Wayne T., Juenger T.E. (2012) Physiological genomics of response to soil drying in diverse *Arabidopsis* accessions. Plant Cell **24**: 893–914

Martin M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal **17**: 10

Matas a. J., Yeats T.H., Buda G.J., Zheng Y., Chatterjee S., Tohge T., Ponnala L., Adato A., et al (2011) Tissue- and Cell-Type Specific Transcriptome Profiling of expanding tomato fruit provides insights into metabolic and regulatory specialization and cuticle formation. Plant Cell **23**: 3893–3910

Monks S.A., Leonardson A., Zhu H., Cundiff P., Pietrusiak P., Edwards S., Phillips J.W., Sachs A., et al (2004) Genetic inheritance of gene expression in human cell lines. Am J Hum Genet **75**: 1094–1105

Morley M., Molony C.M., Weber T.M., Devlin J.L., Ewens K.G., Spielman R.S., Cheung V.G. (2004) Genetic analysis of genome-wide variation in human gene expression. Nature **430**: 743–747

Mounet F., Moing A., Garcia V., Petit J., Maucourt M., Deborde C., Fruit I.B., Bordeaux C. De, et al (2009) Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit. **149**: 1505–1528

van Muijen D., Anithakumari A.M., Maliepaard C., Visser R.G.F., van der Linden C.G. (2016) Systems genetics reveals key genetic elements of drought induced gene regulation in diploid potato. Plant Cell Environ **39**: 1895–1908

Muños S., Ranc N., Botton E., Bérard A., Rolland S., Duffé P., Carretero Y., Paslier M.-C. Le, et al (2011) Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiol **156**: 2244–2254

Nica A.C., Dermitzakis E.T. (2013) Expression quantitative trait loci: present and future. Philos Trans R Soc Lond B Biol Sci **368**: 20120362

Pascual L., Albert E., Sauvage C., Duangjit J., Bouchet J.-P., Bitton F., Desplat N., Brunel D., et al (2016) Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. Plant Sci **242**: 120–130

Pascual L., Xu J., Biais B., Maucourt M., Ballias P., Bernillon S., Deborde C., Jacob D., et al (2013) Deciphering genetic diversity and inheritance of tomato fruit weight and composition through a systems biology approach. J Exp Bot **64**: 5737–5752

Pattison R.J., Csukasi F., Zheng Y., Fei Z., van der Knaap E., Catalá C. (2015) Comprehensive tissue-specific transcriptome analysis reveals distinct regulatory programs during early tomato fruit development. Plant Physiol **168**: 1684–1701

Prudent M., Lecomte A., Bouchet J.-P., Bertin N., Causse M., Genard M. (2011) Combining ecophysiological modelling and quantitative trait locus analysis to identify key elementary processes underlying tomato fruit sugar concentration. J Exp Bot **62**: 907–919

Rabbani M.A., Maruyama K., Abe H., Khan M.A., Katsura K., Ito Y., Yoshiwara K., Seki M., et al (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNAGel-Blot Analyses. Plant Physiol **133**: 1755–1767

R Development Core Team (2012) R: a language and environment for statistical computing.

Ripoll J., Urban L., Staudt M., Lopez-Lauri F., Bidel L.P.R., Bertin N. (2014) Water shortage and quality of fleshy fruits, making the most of the unavoidable. J Exp Bot **65**: 4097–117

Robinson M., Oshlack A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol **11**: R25

Rockman M. V, Kruglyak L. (2006) Genetics of global gene expression. Nat Rev Genet **7**: 862–872

Rozen S., Skaletsky H. (1999) Primer3 on the WWW for general users and for biologist programmers. Bioinforma methods Protoc 365–386

Saliba-Colombani V., Causse M., Gervais L., Philouze J. (2000) Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. Genome **43**: 29–40

Sánchez-Rodríguez E., Moreno D. a, Ferreres F., Rubio-Wilhelmi M.D.M., Ruiz J.M. (2011) Differential responses of five cherry tomato varieties to water stress: changes on phenolic metabolites and related enzymes. Phytochemistry **72**: 723–9

Schadt E.E., Monks S.A., Drake T.A., Lusis A.J., Che N., Colinayo V., Ruff T.G., Milligan S.B., et al (2003) Genetics of gene expression surveyed in maize, mouse and man. Nature **422**: 297–302

Seki M., Narusaka M., Ishida J., Nanjo T., Fujita M., Oono Y., Kamiya A., Nakajima M., et al (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J **31**: 279–292

Slane D., Kong J., Berendzen K.W., Kilian J., Henschen A., Kolb M., Schmid M., Harter K., et al (2014) Cell type-specific transcriptome analysis in the early *Arabidopsis thaliana* embryo. Development **141**: 4831–40

Springer N.M., Stupar R.M. (2007a) Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. Plant Cell Online **19**: 2391–2402

Springer N.M., Stupar R.M. (2007b) Allelic variation and heterosis in maize: How do two halves make more than a whole? Genome Res **17**: 264–275

Steinhauser M.-C., Steinhauser D., Gibon Y., Bolger M., Arrivault S., Usadel B., Zamir D., Fernie A.R., et al (2011) Identification of enzyme activity quantitative trait loci in a *Solanum lycopersicum x Solanum pennellii* introgression line population. Plant Physiol **157**: 998–1014

Stupar R.M., Springer N.M. (2006) Cis-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. Genetics **173**: 2199–2210

Swinnen G., Goossens A., Pauwels L. (2016) Lessons from domestication: targeting cis-regulatory elements for crop improvement. Trends Plant Sci **21**: 506–515

Tomato Genome Consortium ' (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature **485**: 635–41

Verta J., Landry C.R., Mackay J. (2016) Dissection of expression-quantitative trait locus and allele specificity using a haploid / diploid plant system – insights into compensatory evolution of transcriptional regulation within populations. 159–171

Wang S., Yehya N., Schadt E.E., Wang H., Drake T.A., Lusis A.J. (2006) Genetic and genomic analysis of a fat mass trait with complex inheritance reveals marked sex specificity. PLoS Genet **2**: e15

West M.A.L., Kim K., Kliebenstein D.J., Van Leeuwen H., Michelmore R.W., Doerge R.W., St. Clair D.A. (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in *Arabidopsis*. Genetics **175**: 1441–1450

Willcox J.K., Catignani G.L., Lazarus S. (2003) Tomatoes and cardiovascular health. Crit Rev Food Sci Nutr **43**: 1–18

Yvert G., Brem R.B., Whittle J., Akey J.M., Foss E., Smith E.N., Mackelprang R., Kruglyak L. (2003) Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. Nat Genet **35**: 57–64

Zhang N., Brewer M.T., van der Knaap E. (2012) Fine mapping of fw3.2 controlling fruit weight in tomato. Theor Appl Genet **125**: 273–84

CHAPTER 6

CHAPTER 6: Discussion and conclusion

This thesis aimed at characterizing the genetic and genomic variability of tomato response to water deficit and its impact on fruit quality, through the integration of phenotypic, genomic and transcriptomic data. The outcomes provide basis for understanding tomato plant response to water limitation and for fleshy fruit quality improvement under deficit irrigation while limiting yield loss. The individual results were already discussed at the end of each of the three previous chapters. We will thus here first identify some limiting factors of our approach and propose a few possible improvements. We will then discuss our results in terms of impact of water stress and genetic architecture of the traits studied. We will finally discuss the consequences of these results for tomato quality breeding.

---

1. Limiting factors of our approach and a few possible improvements

- Our results rely on the analysis on two different panels, a RIL population and a GWAS panel constituted of cherry tomatoes. The **size of each panel** was limited by the capacity of the greenhouse, as we wanted to simultaneously grow the populations in both conditions side by side. The size of each population were thus limited to less than 140 individuals which allowed us to detect only QTLs with large effects and without a great precision. Nevertheless the high heritability of most of the traits allowed the identification of a large number of QTLs.

- For the RILs we studied a population whose parents were chosen according to their contrasted fruit quality but **not for a specific adaptation to water deficit**. This may explain that the variability in the response to water deficit was much higher in the GWAS panel than in the RIL population.

- The **way we applied water stress** can be criticized as actually all the plants received an identical volume of water whatever their growth and water need. The limited effect observed on small-fruited line could thus, in part, result from a difference in plant development and water requirement. Several phenotyping platforms are now available

to better control the environmental conditions (http://www.plant-phenotyping-network.eu/), but are usually limited to young or small plants like *Arabidopsis* and would not allow the simultaneous analysis of 600 plants of more than 2m high.

- Furthermore, we had **different temperature** and relative **air humidity** conditions in Morocco and Avignon. This may explain the differences observed in terms of impacts of water deficit on the plants. Nevertheless these differences reflect the variability of production conditions. It could even have been interesting to study the populations in winter conditions with a more limited light and lower temperatures.

- In parallel the **phenotypic traits** assessed were limited to flowering time, plant vigour, and fruit quality parameters. We did not measure any trait related to water limitation response such as osmotic adjustment, solute accumulation, transpiration, photosynthesis, which would have given us some clues about the parameters affected by water deprivation in our material. Measuring such traits in large number of genotypes is nevertheless not easy.

- Although the recent availability of the tomato reference genome largely improved our ability to screen for candidate gene, a **higher number of SNP markers** could have been interesting to increase the precision in the GWAS panel and for the identification of haplotypes around the QTLs. This could have allowed us to relate our results with the haplotypes of the large sets of re-sequenced lines available (Aflitos et al., 2014; Lin et al., 2014) to identify possible non-synonymous or deleterious polymorphisms.

- In the transcriptome analysis, two main limitation can be mentioned :
    - The analysis of Allele Specific Expression was limited by the **number of polymorphic genes** (although 11,000 genes is already quite high at the intraspecific level, and the closest the parents the lowest the polymorphism rate), and the **number of F1** studied should be increased in order to be able to study the polymorphisms present in the non-coding regions around the genes to

try to identify some of the polymorphism that could be linked to cis ASE, as proposed by (Kang et al., 2016). A further analysis of 4 F1 hybrids including the Cervil x Levovil one is planned and will be highly informative.

- The **eQTL analysis is limited to 250 genes**, chosen for their location or their variation to be able to quantify the expression by microfluidigm technology. It is thus difficult to discuss the dispersion of the QTLs and identify key regulators corresponding to hot spots of eQTLs, as shown in *Arabidopsis* (Keurentjes et al., 2007; West et al., 2007) following a whole genome eQTL analysis. A whole genome GWAS analysis could be highly informative and becomes realistic according to the very high throughput of new sequencing machines.

- **Annotation** of the genome is also a limiting factor for the interpretation of the results. We have identified a few genes with a wrong annotation and automatic annotation sometimes leads to unprecise function. This should be improved with the new genome version coming soon.

- Finally, our approach could have been much enriched by following recent approaches such as the study of **small ARN**, which are known to regulate gene expression, particularly during stress conditions (Khraiwesh et al., 2010). **Alternative splicing** or **epigenetic marks** could also be interesting to study, in order to give clues about the candidate genes (Chinnusamy and Zhu, 2009; Sahu et al., 2013; Sun and Xiao, 2015).

2. Impact of water stress and genetic variability

We have shown that deficit irrigation increased fruit quality while reducing plant growth and yield but not in the same range according to genotype fruit size in the RILs and to the genotypes in the association panel, where 50 accessions showed an absence of relation between water deficit and yield reduction, although we are not able to indicate the physiological mechanism of their response to water deprivation. This underlines the importance of genetic resources. We could not link the origin of the best performing accessions to specific origin of prospection in

Latin America. Nevertheless, for further studies it could be interesting to enrich the collection with Mediterranean accessions traditionally grown with limited water amount, like the one studied in Balearic Islands which present a strong adaptation (Galmés et al., 2011). Several wild species which can be crossed with cultivated tomatoes are known to be adapted to very dry conditions (Labate et al., 2007). Nevertheless their use for breeding for drought tolerance has not yet proven its efficiency probably because favorable alleles may be linked to unfavorable ones for other traits. It thus seems preferable to continue to study the variability across cultivated accessions.

3. Genetic architecture of the traits studied

We have identified a large number of QTLs, in both populations for every trait. Several methods are available to explore the G x E interaction at the QTL level. We chose to dissect the interactions or study the differential effects and obtained complementary results with both approaches. Other approaches integrate the environmental factors a covariate, but are more adapted for trials performed in many conditions, or use ecophysiological modeling as illustrated on our data in **Appendix 1**. Using the RIL population data of plant water status, fruit growth and composition our ecophysiologist colleagues calibrated a process-based model describing water and carbon fluxes in growing fruit as a function of plant and environmental conditions. Eight genotype-dependent parameters of the model were estimated in order to minimize model prediction errors of fruit dry and fresh mass increases during fruit development. The variability in model parameters allowed them to explore diverse genetic strategies in response to water deficit. We performed a QTL analysis of model parameters and detected three main QTLs related to xylem and phloem conductivity, on chromosome 2, 4 and 8.

A synthesis of QTLs detected in our populations and others would deserve to be performed and made available to the tomato geneticists, now that physical positions of the markers allow the comparison of maps obtained with different set of markers. A tool like Biomercator combined with a QTL database could be highly useful for such study (Sosnowski et al., 2012).

A few genomic regions seem particularly rich in QTLs, but these regions differ in RILs and GWA panel, which is probably due to the different range of variability explored. To precisely identify if a master QTL is responsible for the variation of several traits or if clusters of linked QTLs are due to highly polymorphic regions, fine mapping should be performed, as it was done previously on chromosome 2 (Lecomte et al., 2004) and 4 and 9 (Chaïb et al., 2007). Such analyses often conclude to the presence of several linked QTLs. The near isogenic lines produced at that time could be studied in contrasted environment to further explore these QTL clusters.

The comparison between the RIL population, the GWAS panel and a Multi-allelic MAGIC population was discussed in (Pascual et al., 2016, **Appendix 3**). According to the number of QTLs detected and the lower confidence intervals in the RIL population, the GWAS panel seems to be the most efficient for QTL discovery. Nevertheless the panel does not include large fruited lines to avoid the strong structure linked to fruit size. Thus to study fruit size and the link between quality traits and fruit size, the MAGIC population derived from the intercross of four cherry tomato accessions and four large fruited lines is more adapted. Furthermore all these panels are composed of homozygous lines and do not allow to assess dominance effect, which may be interesting for breeding purpose as all the varieties developed today are F1 hybrids. For such purpose it would be necessary to study genotype performance at the hybrid level, when crossed with a tester.

The genome sequence availability allowed us to propose a set of candidate genes in the regions of interest. Apart from their location in the confidence intervals, these genes were chosen based on a cluster of arguments (polymorphism effect, function related to the trait, expression in the relevant organ and variation of expression, haplotype structure). The screen for these arguments is still complex and an integrated database for all these parameters would be useful to systemize it.

Unless to reduce the interval around a QTL to a unique gene by fine mapping the QTL in isogenic lines or very large GWAS panel and low LD, the final proof that a candidate gene is responsible for the variation of a trait can only be obtained by functional validation. Until recently this was not straightforward as the use of knock out or overexpressed genes by plant transformation

with 35S promoter (or even an organ specific promoter) could induce large phenotypic effect without corresponding to the effect of allele replacement. Today genome editing and allele replacement become available in plant and should allow more rapid and efficient functional validation of the role of candidate genes (Bortesi and Fischer, 2015).

4. Genome expression

Analyzing the transcriptome of the three genotypes Cervil, Levovil and their F1, we identified a large number of genes differentially expressed, corresponding to a wide range of functions, whatever the factor studied, genotype or water condition. The question of the development stage of harvested organs is always questionable. Here, we intended more to have a picture of long term response to stress as we sampled fruit and leaves after several weeks of stress, than to identify the genes involved in early response. We also studied fruit at cell expansion, as it is an important stage in the source to sink relationships. We decided to benefit from the ecophysiological study of the RILs to identify for each line the best time corresponding to the cell expansion stage and thus not harvest all the fruits on the same date and after the same number of days post anthesis. Nevertheless this may induce a bias, as for instance the environmental conditions (other than irrigation) were not the same.

Interesting results concerning sugar metabolisms, but also several candidate genes whose expression was related to phenotypes were identified. Due to time limitation, we could not yet fully analyze these data, and many other questions remain to be explored, notably the variation and role of chaperon genes and transcription factors or the polymorphisms around the *cis* eQTLS. ASE and eQTLs provided complementary results and ASE could be useful for the validation of candidate genes for phenotypic QTLs. Using specific transcripts as covariable in the QTL mapping model of phenotypic traits may be also a way to determine the role of the gene in the trait variation. This still has to be assessed.

Finally, we are conscious that transcription is still far from the phenotypes, and post-transcriptional modifications, as well as proteome variations would deserve to be analyzed but are not yet as easy and high throughput as transcriptome analysis.

5. Consequences for tomato quality breeding

Our results together with previous ones showed that a moderate water stress could increase the concentration in metabolites favorable to fruit quality. The problem for growers is then to manage the stress to avoid a yield loss and the optimal genotypes have to be identified. In the ecophysiological study of the RIL population, a group of genotypes could be discriminated for their low loss of fresh mass under stress associated with an increase active uptake of sugars and low value of the maximum cell wall extensibility, and for their high dry matter content in control condition, associated with mass flow. In the GWAS study we also identified a set of interesting accessions. These individuals constitute an interesting starting point for breeding.

The next question is the ideotype design to follow. The ecophysiological modeling approach was applied to design ideotypes with high dry matter content in control conditions and low loss of water in water deficit condition. The ideotypes outperformed the RILs especially for large and medium fruit-size genotypes, by combining high pedicel conductance and high active uptake of sugars. Interestingly, five small fruit-size RILs were close to the selected ideotypes, and likely bear interesting traits and alleles for adaptation to water deficit. Such approach may allow a better definition of the final ideotypes required.

If selection was based on integrated traits such as yield, fruit weight or soluble solid content, the colocalisation of QTLs with opposite effects is to be considered. QTLs for fruit size and number did not all co-localize with sugar and acid QTLs and the panel of QTLs may be screened to identify the best targets for molecular breeding. Then Marker-Assisted Selection or genomic selection can be planned. Duangjit et al. (2016) showed that the prediction of genomic value derived from a GWAS panel for tomato fruit quality was strongly related to the trait heritability but quite good for many metabolic traits. The way to integrate the environmental impact in genomic value prediction has still to be considered.

Another question is the prediction of F1 hybrid value as all the tomato varieties today are hybrids. The knowledge of trait inheritance as well as gene expression inheritance may help in this prediction and in the construction of the most interesting combinations. Furthermore,

F1 hybrids have been proposed as a way to benefit of a higher stability than in homozygous lines (Blum, 2013). In a longer term, the identification of genes whose expression or polymorphisms are related to the response to environmental conditions may also help in the construction of the right ideotypes.

Finally, in the frame of global change, in order to breed new sustainable varieties, it will be necessary to take into account other stresses, notably high temperature or salty conditions, which are frequently associated to water stress. The genetic and genomic approaches we followed could be used for such purpose. Geneticists will thus need to identify the right targets for future breeding and then to integrate results from ecophysiological and agronomy studies in order to optimize the Genotype by environment by management interaction.

References

Aflitos S., Schijlen E., De Jong H., De Ridder D., Smit S., Finkers R., Wang J., Zhang G., et al (2014) Exploring genetic variation in the tomato (Solanum section Lycopersicon) clade by whole-genome sequencing. Plant J **80**: 136–148

Blum A. (2013) Heterosis, stress, and the environment: A possible road map towards the general improvement of crop yield. J Exp Bot **64**: 4829–4837

Bortesi L., Fischer R. (2015) The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol Adv **33**: 41–52

Chaïb J., Devaux M.F., Grotte M.G., Robini K., Causse M., Lahaye M., Marty I. (2007) Physiological relationships among physical, sensory, and morphological attributes of texture in tomato fruits. J Exp Bot **58**: 1915–1925

Chinnusamy V., Zhu J. (2009) Epigenetic regulation of stress responses in plants. 133–139

Duangjit J., Causse M., Sauvage C. (2016) Efficiency of genomic selection for tomato fruit quality. Mol Breed. doi: 10.1007/s11032-016-0453-3

Galmés J., Conesa M.A., Ochogavía J.M., Perdomo J.A., Francis D.M., Ribas-Carbó M., Savé R., Flexas J., et al (2011) Physiological and morphological adaptations in relation to water use efficiency in Mediterranean accessions of Solanum lycopersicum. Plant, Cell Environ **34**: 245–260

Kang E.Y., Martin L., Mangul S., Isvilanonda W., Zou J., Ben-David E., Han B., Lusis A.J., et al (2016) Discovering SNPs regulating human gene expression using allele specific expression from RNA-Seq Data. Genetics **204**: 1057–1064

Keurentjes J.J.B., Fu J., Terpstra I.R., Garcia J.M., van den Ackerveken G., Snoek L.B., Peeters A.J.M., Vreugdenhil D., et al (2007) Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. Proc Natl Acad Sci U S A **104**: 1708–13

Khraiwesh B., Arif M.A., Seumel G.I., Ossowski S., Weigel D., Reski R., Frank W. (2010) Transcriptional control of gene expression by microRNAs. Cell **140**: 111–122

Labate J.A., Grandillo S., Fulton T., Muños S., Caicedo A.L., Peralta I., Ji Y., Chetelat R.T., et al (2007) 1 Tomato. 5:

Lecomte L., Saliba-Colombani V., Gautier A., Gomez-Jimenez M.C., Duffé P., Buret M., Causse M. (2004) Fine mapping of QTLs of chromosome 2 affecting the fruit architecture and composition of tomato. Mol Breed **13**: 1–14

Lin T., Zhu G., Zhang J., Xu X., Yu Q., Zheng Z., Zhang Z., Lun Y., et al (2014) Genomic analyses provide insights into the history of tomato breeding. Nat Genet. doi: 10.1038/ng.3117

Pascual L., Albert E., Sauvage C., Duangjit J., Bouchet J.-P., Bitton F., Desplat N., Brunel D., et al (2016) Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. Plant Sci **242**: 120–130

Sahu P.P., Pandey G., Sharma N., Puranik S., Muthamilarasan M., Prasad M. (2013) Epigenetic mechanisms of plant stress responses and adaptation. Plant Cell Rep **32**: 1151–1159

Sosnowski O., Charcosset A., Joets J. (2012) Biomercator V3: An upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. Bioinformatics **28**: 2082–2083

Sun Y., Xiao H. (2015) Identification of alternative splicing events by RNA sequencing in early growth tomato fruits. BMC Genomics **16**: 948

West M.A.L., Kim K., Kliebenstein D.J., Van Leeuwen H., Michelmore R.W., Doerge R.W., St. Clair D.A. (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. Genetics **175**: 1441–1450

Appendices

Appendix 1: 'Model-assisted estimation of the genetic variability in physiological parameters related to tomato fruit growth under contrasted water conditions'. (Article accepted for publication in Frontiers in Plant Sciences, 2016)

Appendix 2: 'Developing tomato varieties with improved flavor' (Book chapter XIII: Achieving sustainable tomato cultivation. Ed A. Matto and A. Handa. Accepted for publication, 2016)

Appendix 3: 'Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels.' (Article published in Plant Science, 2015)

Appendix 4: Supplemental figures and tables from chapter 3.

Appendix 5: Supplemental figures and tables from chapter 4.

Appendix 6: Supplemental figures and tables from chapter 5.

Appendix 1: Model-assisted estimation of the genetic variability in physiological parameters related to tomato fruit growth under contrasted water conditions

Article accepted for publication in Frontiers in Plant Sciences, 2016

Appendix 1

frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 December 2016
doi: 10.3389/fpls.2016.01841

Model-Assisted Estimation of the Genetic Variability in Physiological Parameters Related to Tomato Fruit Growth under Contrasted Water Conditions

*Dario Constantinescu[1], Mohamed-Mahmoud Memmah[1], Gilles Vercambre[1], Michel Génard[1], Valentina Baldazzi[1], Mathilde Causse[2], Elise Albert[2], Béatrice Brunel[1], Pierre Valsesia[1] and Nadia Bertin[1]**

[1] Plantes et Systèmes de Culture Horticoles, Institut National de la Recherche Agronomique - Centre PACA, Avignon, France, [2] Unité Génétique et Amélioration des Fruits et Légumes, Institut National de la Recherche Agronomique – Centre PACA, Montfavet, France

OPEN ACCESS

Edited by:
Lifeng Xu,
Zhejiang University of Technology, China

Reviewed by:
Teemu Hölttä,
University of Helsinki, Finland
Tsu-Wei Chen,
Leibniz University of Hanover, Germany

***Correspondence:**
Nadia Bertin
nadia.bertin@inra.fr

Specialty section:
This article was submitted to Plant Biophysics and Modeling, a section of the journal Frontiers in Plant Science

Received: *26 August 2016*
Accepted: *22 November 2016*
Published: *09 December 2016*

Citation:
Constantinescu D, Memmah M-M, Vercambre G, Génard M, Baldazzi V, Causse M, Albert E, Brunel B, Valsesia P and Bertin N (2016) Model-Assisted Estimation of the Genetic Variability in Physiological Parameters Related to Tomato Fruit Growth under Contrasted Water Conditions. Front. Plant Sci. 7:1841. doi: 10.3389/fpls.2016.01841

Drought stress is a major abiotic stress threatening plant and crop productivity. In case of fleshy fruits, understanding mechanisms governing water and carbon accumulations and identifying genes, QTLs and phenotypes, that will enable trade-offs between fruit growth and quality under Water Deficit (WD) condition is a crucial challenge for breeders and growers. In the present work, 117 recombinant inbred lines of a population of *Solanum lycopersicum* were phenotyped under control and WD conditions. Plant water status, fruit growth and composition were measured and data were used to calibrate a process-based model describing water and carbon fluxes in a growing fruit as a function of plant and environment. Eight genotype-dependent model parameters were estimated using a multiobjective evolutionary algorithm in order to minimize the prediction errors of fruit dry and fresh mass throughout fruit development. WD increased the fruit dry matter content (up to 85%) and decreased its fresh weight (up to 60%), big fruit size genotypes being the most sensitive. The mean normalized root mean squared errors of the predictions ranged between 16–18% in the population. Variability in model genotypic parameters allowed us to explore diverse genetic strategies in response to WD. An interesting group of genotypes could be discriminated in which (i) the low loss of fresh mass under WD was associated with high active uptake of sugars and low value of the maximum cell wall extensibility, and (ii) the high dry matter content in control treatment (C) was associated with a slow decrease of mass flow. Using 501 SNP markers genotyped across the genome, a QTL analysis of model parameters allowed to detect three main QTLs related to xylem and phloem conductivities, on chromosomes 2, 4, and 8. The model was then applied to design ideotypes with high dry matter content in C condition and low fresh mass loss in WD condition. The ideotypes outperformed the RILs especially for large and medium fruit-size genotypes, by combining high pedicel conductance and high active uptake of sugars. Interestingly, five small fruit-size RILs were close to the selected ideotypes, and likely bear interesting traits and alleles for adaptation to WD.

Keywords: fleshy fruit, quality, ideotype, *Solanum lycopersicum*, virtual fruit model, water stress, multiobjective optimization

INTRODUCTION

Drought stress is one of the major abiotic stresses, which represents the primary cause of crop loss worldwide, and the development of more water efficient cropping systems is becoming critical (Bodner et al., 2015). Nonetheless, in the case of fleshy fruits, moderate drought has been suggested to improve both organoleptic quality and nutritive value (Ripoll et al., 2014). Trade-offs between quality and yield seem realistic, but depend strongly on stress intensity and genotypes (Ripoll et al., 2016). Indeed, recent studies on tomato revealed a strong genetic variability in the response to drought from negative to nil to positive impact on fruit size and quality (Ripoll et al., 2015). A large number of genes and molecular mechanisms involved in survival under drought have been identified, in particular in *Arabidopsis thaliana* (L.) Heynh. (Blum, 2011). These genes are involved in the control of many physiological processes, but they do not necessarily confer a stress resistance and they may entail detrimental effects on yield and quality in crop plants facing long periods of drought combined with high temperature (Gong et al., 2010; Tardieu, 2012). In tomato, only a few QTLs/genes involved in the response to water deficit are known (Labate et al., 2009). In a recent study (Albert et al., 2016), a RIL population of 117 F7 recombinant inbred tomato lines has been genotyped for 501 SNP markers and phenotyped under control (C) and water deficit (WD). This study revealed a total of 56 QTLs of plant and fruit traits, among which 11 depended on watering regime. Interestingly, these authors observed a large genetic diversity in plant and fruit responses to WD and significant genotype by watering regime interactions, suggesting the possibility to develop tomato genotypes adapted to grow under water limitation. The diversity present in genetic resources of tomato species is a vital source of traits and alleles for crops, many of which may have been inadvertently lost during selection. Thus, identifying main mechanisms governing fruit adaptation to water deficit and pinpointing genes, QTLs and phenotypes that will enable a fruit to maintain growth and improve quality under conditions of limited water supply is a crucial challenge for breeders and growers in the light of current issues related to climate change.

Crop models are adequate tools for analyzing genotype by environment interactions, since they integrate environmental and genetic effects on individual physiological processes and are able to predict interactions among processes during fruit development (Bertin et al., 2010). The Virtual Fruit Model (Fishman and Génard, 1998), an eco-physiological process-based model which describes both water and dry matter accumulation rates in fleshy fruits, has already proven its robustness and genericity under contrasted environmental conditions and for different fruit species: peach (Quilot et al., 2005), mango (Lechaudel et al., 2007), kiwifruit (Hall et al., 2013), and tomato (Liu et al., 2007). Notably, this model has been used to assess water deficit impacts on fruit growth (Lescourret and Génard, 2005; Baldazzi et al., 2013). In such mechanistic models, the parameters are linked to physiological traits or processes which can be linked to loci or genes. Each parameter is in fact related to a set of interconnected processes controlled by a group of genes, which was defined by Tardieu (2003) as "meta-mechanism." Though plant traits generally depend on genotype, environment and cultural practices, model parameters should be, ideally, independent of the environment and management. Some of these parameters,—called genotypic parameters,—are genotype dependent while others are generic and independent of the genotype (Boote et al., 2001). The set of genotypic parameters related to a particular genotype represents a phenotypic fingerprint of this genotype and it is amenable to QTL analysis (Bertin et al., 2010). Several attempts have been made to include genetic information into process-based models and to link model parameters to genes or QTLs (White and Hoogenboom, 1996; Chapman et al., 2003; Reymond et al., 2003; Quilot et al., 2005; Xu et al., 2012; Rebolledo et al., 2015). The main difficulty is that the model should capture sufficient physiological functionalities, to simulate the expression of single genes or a gene network.

An ultimate goal is then to use these enriched process-based models for the design of ideotypes adapted to biotic and abiotic stress environments. Here, an ideotype designates a "plant model which is expected to perform or behave in a predictable manner within a defined environment." However, the fitness landscape (objectives space) to be explored to design ideotypes is often very complex and a large number of parameter combinations must be evaluated in order to identify the best-adapted genotypes. This difficulty comes from the nonlinear and non-convex nature of antagonist criteria and the complex nature of the process-based models, such as the "Virtual Fruit." Consequently, the model-based design of ideotypes is a difficult nonlinear multi-objective optimization problem that resists to the classical simulation and optimization methods. To deal with such multi-objective optimization problems, nature-inspired optimization algorithms (e.g., genetic algorithms and particle swarm optimization algorithms) are suitable and increasingly used. Multi-Objective Evolutionary Algorithms (MOEAs) are amongst the best-known and most effective nature inspired optimization algorithms. They allow exploring high dimensional solution spaces and they do not require any derivative information. MOEAs generate many feasible and non-dominated solutions, i.e., elements of the Pareto optimal set (best tradeoffs between conflicting objectives). Many papers have been published on the use of evolutionary algorithms for ideotype model-based design. For the sake of conciseness, we mention only the works of Letort et al. (2008) on beech trees, of Qi et al. (2010) on maize, of Lu et al. (2012) on wheat, of Quilot-Turion et al. (2012) and Sidi et al. (2014) on peach, and of Ding et al. (2016) on rice.

In the present study our objectives were to use the Virtual Fruit Model (i) to phenotype a RIL population of tomatoes at the process level; (ii) to better understand the fruit growth mechanisms (water and dry matter accumulation) involved in the response to water deficit; (iii) to look for optimized sets of genotypic parameters/genotypes which could reduce the loss of fruit fresh weight under WD and at the same time maintain/improve high fruit dry matter content. The RIL population is the one previously genotyped by Albert et al. (2016). The genetic variability in fruit traits and model parameters was analyzed by Principal Component Analysis (PCA) and through QTL analysis. This step helped to explore diverse genetic

strategies in response to water deficit and to discuss potential processes/genes involved in this response. Then the model was applied to design ideotypes in terms of fruit size and quality.

MATERIALS AND METHODS

RIL Population and Experimental Design

The RIL population, including 117 F7 recombinant inbred lines, was developed from an intraspecific cross between two inbred lines, Cervil and Levovil (described in Saliba-Colombani et al., 2001). Cervil is a cherry type tomato (*S. lycopersicum* var. *cerasiforme*, 6–10 g), whereas Levovil (*S. lycopersicum*) is a large fruited accession (90–160 g). The 117 RILs, the F1 hybrid and the two parents were grown in a heated glasshouse in INRA Avignon (France) from March to July 2013. Based on previous data, eight genotypes were selected in the population in order to have a good representation of the ranges in fruit size and dry matter content. These eight genotypes included the two parents and the F1 hybrid (CxL). Some input parameters of the Virtual Fruit Model (initial fresh and dry weights, fruit surface conductance to water, stem water potential) were accurately measured on these eight representative genotypes and then the same values were applied to all genotypes of the group (see below and **Figure 1**).

Plants were grown in 4 l plastic pots filled with peat (Klasmann 165) and watered with nutritive solution (2, 4, 6 mmol l−1, N, P, and K, respectively). All trusses were pollinated with an electrical bee. The number of flowers per truss was regulated to get homogeneous fruit load and comparable source:sink ratios among plants of a given genotype. The first two trusses of the small fruit genotypes (final fruit size < 30g) were pruned to 8 fruits and the following trusses to 12 fruits. Regarding the medium and large fruit genotypes (final fruit size > 30g), the first two trusses were pruned to 4 fruits and the following trusses to 6 fruits. Climate conditions (temperature, humidity and light intensity) in the glasshouse were recorded every minute and data were averaged hourly throughout the experiment.

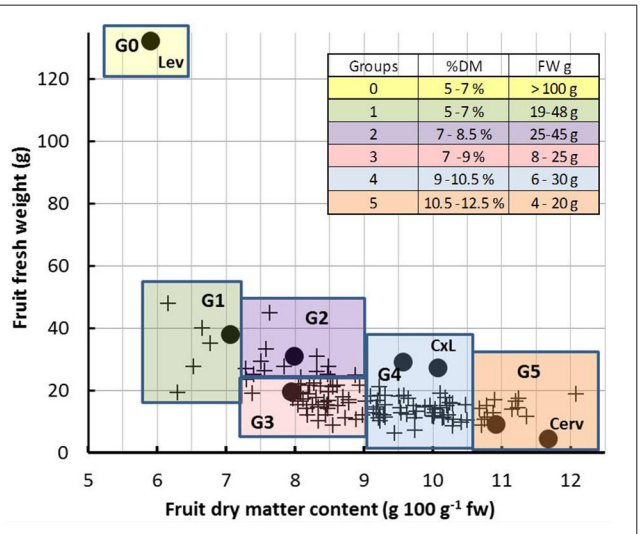


FIGURE 1 | Relationship between fruit fresh weight and dry matter content of ripe fruits under control condition. Each symbol represents one genotype (means of 15 to 20 fruits). Black dots indicate the five representative genotypes, the two parents (Lev and Cerv) and the F1 hybrid (CxL). The colored squares represent the six groups of genotypes (G1 to G6) which were considered for model inputs. The insert gives the ranges of fresh weight (FW) and dry matter content (%dm) of ripe fruits in each group.

From anthesis of the second truss of Cervil (considered as a reference early genotype), two irrigation treatments were applied: control (C) and water deficit (WD). Control plants were irrigated in order to get drainage around 25%. In the WD treatment, water supply was reduced by 64% compared to the control, corresponding to 49% of the potential evapotranspiration on average over the experimental period. The peat substrate humidity was assessed continuously with 12 small soil moisture sensors (EC-5 Decagon devices, USA) inserted in the substrate and randomly distributed in the glasshouse, and twice a week with a water content sensor (WCM-control, Grodan, Roermond, The Netherlands). Peat substrate humidity averaged 60–65% in control plants and 25–30% in WD plants (no drainage). Within the glasshouse, irrigation treatments were applied by row, and the genotypes were randomized within rows. Two plants of each genotype (10 for the parent lines and for the six representative genotypes) were grown under each treatment. The trial plants were surrounded with one row of border tomato plants.

Phenotypic Measurements

Stem water potential was measured using a pressure chamber (SAM Précis 2000 Gradignan, France) at predawn and at solar noon. Measurements were performed twice during the stress period on five plants of the eight representative genotypes under both conditions. The fruit conductance was measured on three ripe fruits of the eight genotypes in both treatments, according to the weight loss method described in Lescourret et al. (2001). Flower anthesis was recorded on four successive trusses on all plants (excluding the first two trusses). The fruit fresh and dry masses were measured from 8 days after anthesis (daa) until fruit ripening (from beginning of June to beginning of July) on the whole population. About 4–5 fruits were sampled every 7 days for the 8 representative genotypes. For all other genotypes, three fruits were sampled at 8–10 daa, 12–15 daa, and 20–25 daa. At ripening about 15 to 20 fruits were sampled on all genotypes. For a given developmental stage, fruits were sampled on trusses which developed during the same time window. Within a truss, the first proximal and last distal fruits were not sampled, in order to avoid fruit position effects. The fruit fresh mass was measured after harvest and the fruit dry mass was measured after drying in a ventilated oven for 72 h. All sampled fruits from the WD treatment were grown after water deficit onset, which means that cell division, cell expansion and ripening processes were all affected by WD.

Virtual Fruit Model Description

Fishman and Génard (1998) developed a biophysical model which simulates water and dry matter accumulation rates in the fruit, using as inputs two climatic variables (fruit temperature and air humidity) and two variables describing the plant status (stem water potential, and phloem sap concentration in sugars). This model describes the biophysical processes involved in fruit

growth, with appropriate equations computing uptakes from the xylem and phloem across composite membranes, and losses of dry matter and water due to respiration and transpiration, respectively. Hall et al. (2013) extended the model formulation by adding a pedicel, which contributes to the major hydraulic resistance of the pathway to the fruit (Mazzeo, 2008). The extended version of Hall et al. (2013) was used in our study. Water and sugar flow from the stem through the pedicel into an intermediate compartment, that we called the fruit vasculature, and then through composite membranes into the fruit. The equations describing flows from the fruit vasculature into the fruit (U_p = mass flow from phloem, U_x = mass flow from xylem, U_s = sugar flow) and those describing fruit respiration (R_f) and transpiration (T_f) are the same as those given by Fishman and Génard (1998). The model simulates two state variables (w = mass of water in fruit, s = dry mass of fruit), whose rates of change are:

$$\frac{dw}{dt} = U_x + U_p + r_w R_f - T_f \quad (1)$$

$$\frac{ds}{dt} = U_s - R_f \quad (2)$$

where r_w is the proportion of dry mass converted to water during respiration ($r_w = 9/16$ according to Hall et al., 2013).

Three parallel mechanisms involved in sugar uptake (U_s) from the phloem were considered: active uptake (using Michaelis-Menten kinetics), mass flow, and diffusion (equations are described in Liu et al., 2007).

The rate of fruit volume (V) increase is given by:

$$\frac{dV}{dt} = \begin{cases} V\phi\left(P_f - Y\right) & P_f > Y \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where ϕ and Y are respectively, the cell wall extensibility and yield threshold parameters of the Lockhart equation. When this is equated to the rate of volume increase calculated from the mass balance, we get an algebraic equation for P_f (fruit turgor pressure).

The water and carbon fluxes through the pedicel xylem (which primarily carries water) and phloem (which carries water and sugar) were considered, as in the model developed by Hall et al. (2013).

To identify the main genotypic parameters that affect the model outputs, we performed a sensitivity analysis of the "Virtual Fruit" model, which includes 30 parameters. Three sensitivity analysis methods were used for this purpose: one elementary effects method, i.e., Morris method, and two methods based on the variance decomposition, i.e., the Fourier Amplitude Sensitivity Test (FAST) and the Sobol's methods (Saltelli et al., 2008). Based on the conclusions of those methods, a cross selection of the most important parameters was performed. Accordingly, six genotypic parameters involved in different processes had significant impacts on model outputs (**Table 1**). Two additional parameters were chosen because of their impact on carbon and water transports, which are main processes on which this study focusses. The first one (tauS) drives the mass flow, whereas the second one (lp1) is related to the pedicel conductivity which is strongly involved in water uptake from the phloem. These eight genotypic parameters are described in **Table 1**.

Model Calibration

As mentioned above, the model genotypic parameters were assumed to be genotype dependent and environment independent, i.e., they do not depend on the irrigation conditions. Thus, each set of parameters is a footprint of one of the 117 tomato RILs. To account for the different plant and fruit status under C and WD conditions, some of the model inputs were measured experimentally under each treatment: the stem water potential, the fruit surface conductivity to water vapor, the initial dry and fresh masses, and the fruit osmotic pressure related to soluble compounds other than sugars.

The model calibration aims at estimating the values of the eight selected genotypic parameters in order to minimize the fitting errors (observed vs. simulated fruit fresh and dry weights) for each genotype. The performance index used in the model calibration was the Normalized Root Mean Squared Error (NRMSE), a dimensionless indicator that takes into account the time steps in which more observations were available along with fewer observations at other time steps. This index is suggested in Wallach et al. (2013):

$$NRMSE\ [\%] = 100^* \frac{\sqrt{\frac{1}{n}^* \sum_{i=1}^{n} (O_i - S_i)^2}}{\frac{1}{n}^* \sum_{i=1}^{n} O_i} \quad (4)$$

where Oi and Si are respectively, the observed and simulated values of fruit fresh and dry masses, and n is the number of observations.

The four objectives corresponding to the four NRMSE values, related to the fruit dry and fresh masses under C and WD conditions, were aggregated into two objectives. For this purpose the mean NRMSE value calculated under each irrigation condition was considered in order to have a balanced fitting error between the fruit weight components:

$$f_1(\boldsymbol{X}) = NRMSE_{aggr_C} = \frac{NRMSE_{f_C} + NRMSE_{dC}}{2} \quad (5)$$

$$f_2(\boldsymbol{X}) = NRMSE_{aggr_{WD}} = \frac{NRMSE_{f_{WD}} + NRMSE_{dWD}}{2} \quad (6)$$

where $\boldsymbol{X} = (x_1, x_2, x_3, \ldots, x_8)^T$ is the vector of parameters generating the (f_1,f_2) objective values. $NRMSE_{fC}$ and $NRMSE_{dC}$ are related to respectively, the fruit fresh weight and fruit dry weight predictions in the control (C) condition. $NRMSE_{fWD}$ and $NRMSE_{dWD}$ are related to respectively, the fruit fresh weight and fruit dry weight predictions in the water deficit (WD) condition.

The model calibration was therefore formulated as a multi-objective problem as follows:

$$\min_{X \in D} \left\{ f_1(\boldsymbol{X}), f_2(\boldsymbol{X}) \right\} \quad (7)$$

where D is the search space defined by boundaries of the considered parameters. The problem solutions $\boldsymbol{X}^*$ are all the

Appendix 1

TABLE 1 | Description of the eight genotypic parameters used in the calibration step and of the three additional parameters used for designing ideotypes.

Parameter name	Description	Boundaries			
		Calibration		Ideotypes design	
		Lower	Upper	Lower	Upper
phiMax [bar^{-1} h^{-1}]	Maximum cell wall extensibility. Involved in cell expansion rate	1.0E-04	0.01	0.002	0.02
Lp [g cm^{-2} bar^{-1} h^{-1}]	Conductivity of the composite membrane for water transport from phloem to fruit cells	5.0E-04	0.4	0.02	0.6
nuM [gs h^{-1}]	Maximum sugar active uptake rate. Involved in the sugar active uptake calculus (Ua)	0.002	0.15	0.002	0.2
tstar [h]	Involved in the sugar active uptake calculus. The higher is tstar, later the active uptake begins to decrease	10	900	10	900
tauA [h]	Involved in the sugar active uptake calculus. The higher is tauA, the slower is the active uptake decreasing rate in the growth stage	5	900	72	900
tauS [h^{-2}]	Involved in the calculus of the reflection coefficient of the composite membrane (sigmaP) which increases with tauS. sigmaP is involved in phloem mass flow	5.0E-06	1.5E-05	1.5E-06	2E-05
lp1 [g bar^{-1} h^{-1}]	Pedicel conductivity for the water transport in phloem	5.0E-05	0.1	0.002	0.2
rxp [dimensionless]	Rxp=Lx/Lp=Lx1/Lp1. Lx and Lx1 have the same meaning as Lp and Lp1 but they refer to the xylem	0.1	0.6	0.1	0.8
s0 [g][1]	Initial fruit dry weight			0.019	0.086
w0 [g][1]	Initial fruit water weight			0.126	1.0
bssrat [dimensionless][1]	Involved in the soluble sugar concentration calculus. Ssrat = assrat*t/24 + bssrat Ssrat is the ratio between soluble sugars mass and the total dry mass			0.043	0.22

1 = irrigation dependent parameter, optimized only for the ideotype design.

The parameter ranges used in the calibration step are based on literature data. The lower boundaries of phiMax, lp, and lp1 are set near to 0 for computational stability reasons; tauA, tstar, and tauS are based on experimental information on fruit development.

parameter sets belonging to the Pareto front, i.e., the set of solutions that consists in the best tradeoffs between the two conflicting objectives.

Design of Ideotypes

In this step, we aimed to design ideotypes of tomato adapted to WD conditions. The term ideotype designates a combination of genotypic parameters that represent virtual tomato genotypes with optimized tolerance to water deficit. For this purpose, we considered a set of 11 genotypic parameters, adding three new genotype dependent parameters to the search space (**Table 1**). In this study, the ideotype design aimed at (i) maximizing the ratio between dry weight and fresh weight at the ripe stage (dry matter content *dm*) until a maximal value of 10% under C condition, and (ii) minimizing the fresh weight loss associated with water deficit. Thus the ideotype design was formulated as a multi-objective problem as follows:

$$f_1(\boldsymbol{X}_{id}) = dm_C\,[\%] = 100^* \frac{dry_{weight_C}}{fresh_{weight_C}} \quad (8)$$

$$f_2(\boldsymbol{X}_{id}) = loss\,[\%] = 100^* \sqrt{\left(\frac{\left(fresh_{weight_C} - fresh_{weight_{WD}}\right)}{fresh_{weight_C}}\right)^2} \quad (9)$$

The problem formulation becomes:

$$\min_{X_{id} \in D_{id}} \left\{-f_1(\boldsymbol{X}_{id}), f_2(\boldsymbol{X}_{id})\right\}, \text{ Subject to } dm_{C,WD}\,[\%] < 10\% \quad (10)$$

where $\boldsymbol{X}_{id}$ is the parameter vector belonging to the set $D_{\mathbf{id}}$, which represents the ideotypes search space. The negative sign of $f_1(\boldsymbol{X})$ objective is introduced to transform the minimization into maximization.

Because the sensitivity to WD depends on fruit size, we considered three groups of tomatoes differing by their final fresh weight in control conditions: large size (100–300 g), medium size (20–80 g) and small size (5–15 g).

Optimization Algorithm for Model Calibration and Design of Ideotypes

The NSGA-II developed by Deb et al. (2002) has proven to be one of the most efficient algorithms for solving multi-objective problems. Therefore, we used this algorithm both for the "Virtual Fruit" model calibration and for the tomato ideotype design. For sake of simplicity, we do not give a full description of this algorithm. The interested readers can refer to the above cited. The NSGA-II algorithm was applied through the Java package *jMetal*. As the NSGA-II algorithm depends on random variables, the optimization process was repeated 10 times in the calibration phase and 20 times for the ideotype design. At the end of the process, we could have high number of similar solutions. Therefore, the choice of the best compromise solution for the

calibration step was based on the *min-max* decision criterion, to avoid high mean fitting errors in each condition. Therefore, among the solutions $\boldsymbol{X}_i^*$ belonging to the Pareto-optimal solution set P, we chose the solution $\overline{\boldsymbol{X}}$ that satisfied the following condition:

$$\min\left\{\max_{\boldsymbol{X}^*\in P}\left\{f_1\left(\boldsymbol{X}^*\right), f_2\left(\boldsymbol{X}^*\right)\right\}\right\} \quad (11)$$

where P is the set of Pareto-optimal solutions. We also checked that among the best sets of parameters estimated for one genotype (solutions that all have similar objective values), parameters were not correlated (data not shown).

For the design of ideotypes, we performed a Principal Component Analysis on the parameter sets, whose corresponding objective values matched the following decision criteria:

$$dm_C\,[\%] \geq 8\%\ and\ loss\ [\%] \leq 15\% \quad (12)$$

Principal Component Analysis and Hierarchical Clustering on PCA Individuals Score

A Principal Component Analysis (*ade4* package developed for the *R* software) (Dray and Dufour, 2007) was performed on the parameter values estimated for each recombinant line. This analysis was also applied to study the ideotype features. Genotypic parameters obtained for both calibration and ideotype design, were set as active variables. The dry and fresh weights under C and WD conditions and the dry matter content and fresh mass reduction under WD conditions were added as supplementary variables for the first PCA (model calibration step), while for the second (ideotype design) the initial dry weight, the initial fresh weight, and the *bssrat* parameter (contributing to the soluble sugar concentration calculus) were added. Data were previously normalized and centered (subtracting the mean value and dividing by the standard deviation). Among the 120 calibrated individuals, we excluded one outlier individual.

The PCA individual scores of the model calibration were subjected to hierarchical cluster analysis, using the *complete-linkage* clustering method with the *hclust* R function. The cluster number was chosen according to a visual criterion based on the cluster dendogram. For the ideotype analysis, the three groups of fruit size (large, medium and small size) were used to group the individual scores.

QTL Analysis of Model Genotypic Parameters

The best estimations of the eight genotypic parameters for the 120 genotypes and the coordinates of the RILs on the three first axis of the PCA were used as phenotypic traits in the QTL detection. When distributions were skewed, the best corrections for normality were applied: LOG10(nuM); $\sqrt{tstar}$; LOG10(lp1); LOG10(lx); LOG10(lx1); 1/rxp. The QTL detection was performed as presented in Albert et al. (2016) using the genetic map developed by Pascual et al. (2016) which included 501 SNP markers covering 80% of the tomato genome. Briefly, the simple interval parametric mapping model (Lander and Botstein, 1989) based on the EM algorithm method implemented in the R/QTL package (Broman et al., 2003) was used. A 1000-permutation test was performed to estimate the significant thresholds. Firstly, a LOD threshold equal to 3.13 and corresponding to a genome wide significance level of $\alpha = 0.05$ was considered. Then, we also considered lower significance levels to detect more QTLs: $\alpha = 0.10$ (LOD threshold = 2.76), $\alpha = 0.20$ (LOD threshold = 2.42) and $\alpha = 0.30$ (LOD threshold = 2.20). For each detected QTL, position, LOD score, marker at the LOD score peak, confidence interval (CI, LOD decrease of one unit), average phenotypic values of the two parental alleles and percentage of phenotypic variation explained (PVE) were displayed. The CIs were expressed both in cM Haldane (genetic distance) and in Mbp onto the tomato genome (assembly v2.5) (physical distance). The number of genes within each interval was identified from the tomato genome annotation (2.4). We reported the locations between the detected QTLs and the QTLs identified on phenotypic traits (plant and fruit traits) measured on the same plants (see Albert et al., 2016).

RESULTS

Water Deficit Effects on the Observed Dry Matter Content and Fresh Weight

The observed values of fresh weight and dry matter content measured at the ripe stage under control and WD conditions are shown on **Figures 1**, **2**. WD generally increased the dry matter content (up to 85%) and decreased the fruit fresh weight (up to 60%). This was directly connected to the lower influx of water to the fruit under WD conditions. Cervil—characterized by a low fresh weight—was the less sensitive to WD. The dry matter content of the F1 hybrid (CxL) increased substantially, while its fresh weight decreased slightly. On the contrary, Levovil was the most sensitive to WD, since it lost more than half of its fresh weight and it doubled its dry matter content under WD. In the population, the relative decrease in fruit fresh weight under WD was negatively correlated to the fresh weight under control conditions, indicating that large fruit genotypes were the most sensitive to WD, as mentioned in Albert et al. (2016). On the contrary, the increase in dry matter content under WD was rather independent of the dry matter content observed under control condition. Interestingly a few genotypes were close to the bisector and thus, get comparable fresh weight or dry matter content under both conditions. For these genotypes (Cervil, SSD12, SSD17, SSD49, SSD61, SSD65, SSD140, and SSD154), the differences between C and WD conditions was less than 5 g fresh mass and 1% dry matter content (**Figure 2**).

Model Calibration and Genetic Variability in Model Genotypic Parameters in the RIL Population

Eight genotypic parameters of the model (**Table 1**) were estimated for the RILs and for the two parent lines, in order to predict the dry and fresh masses (output variables) during fruit growth. The fittings were fairly good. **Table 2** shows the

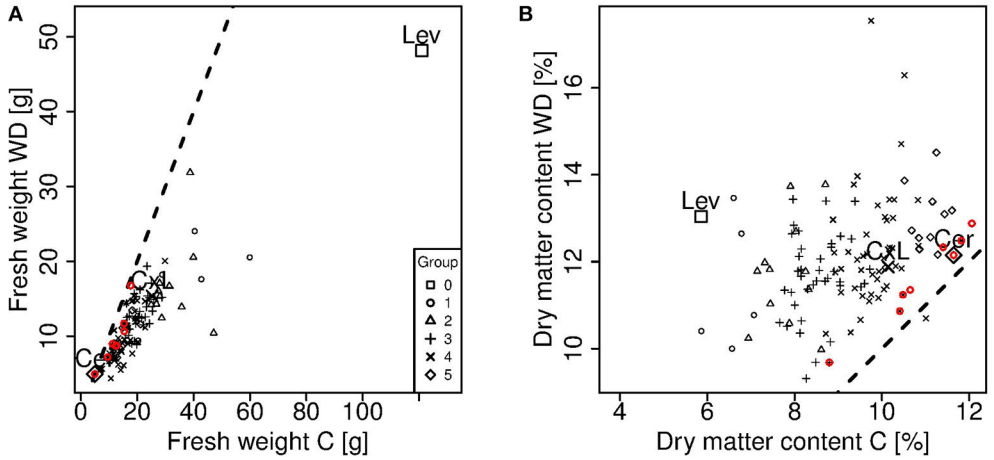


FIGURE 2 | Comparison of fresh weight (A) and dry matter content **(B)** measured on ripe fruits under control (C) and water deficit (WD) conditions. Each point represents one genotype (means of 15 to 20 fruits) and crosses indicate the parental lines (Cervil and Levovil) and the F1 hybrid (CxL). The different symbols represent the six groups of genotypes shown in **Figure 1**. The dashed lines represent the condition in which the plotted variables are equal. The red points represent the genotypes that are near to both dashed lines. These genotypes are the same in **A** and **B**.

TABLE 2 | Statistical summary of the Normalized Relative Mean Squared Errors (NRMSE) obtained with the model calibration under control (C) and water deficit (WD) conditions.

NRMSE	Mean [%]	Standard deviation [%]	Minimum [%]	Maximum [%]	Parents and F1 [%]	
Fresh weight in C condition	17.41	5.35	7.88	34.00	Cer	8.65
					CxL	16.36
					Lev	34.00
Dry weight in C condition	16.48	4.03	8.61	28.18	Cer	9.08
					CxL	8.72
					Lev	15.78
Fresh weight in WD condition	17.76	4.10	9.34	34.20	Cer	9.42
					CxL	12.88
					Lev	24.11
Dry weight in WD condition	17.65	4.19	8.29	27.46	Cer	8.29
					CxL	12.70
					Lev	25.19

The dry and fresh weight increases were fitted from 8 daa until fruit maturation and NRMSE were calculated over this developmental period for each genotype and condition. Mean and standard deviations were calculated for the whole RIL population (including the parent lines). Minimum and maximum refer to the lower and upper values of NRMSE obtained in the population. On the last column, the parents and the F1 hybrid values are shown.

NRMSE values obtained under C and WD conditions that were obtained for the whole population, for the parental lines and for the F1 hybrid. Considering the dry and fresh mass increases over the developmental period (from 8 daa to maturity), the mean NRMSE of the population ranged between 16 and 18% (standard deviation ∼ 4–5%) whatever the condition and output variables (**Table 2**). The total variation of NRMSE values in the population was in the range of 5–34%. NRMSE values obtained for Cervil were close to the minimum for all objectives. The dry mass increase of Levovil fruits was better simulated in C than in WD condition, while the prediction of their fresh mass increase was the worst under C condition.

Considering the ripe stage, the final fruit dry mass was more accurately predicted by the model than the final fresh mass (**Figure 3**), and predictions were better under WD than C condition. Indeed, the model underestimated the fresh mass of the largest fruit-size genotypes in C condition, in particular for Levovil (−34 g). For this genotype, the prediction errors of fresh mass in C conditions were high over the whole development period, as indicated by the NRMSE value in **Table 2**. On the contrary, the model predictions were more accurate for Cervil and the CxL hybrid under both conditions. As a consequence, the final dry matter content was mostly overestimated (differences from −1.73 to 6.17%) and underestimated (differences from −3.16 to 1.37%) in C and WD conditions, respectively (**Figure 4**).

The frequency distributions of the eight genotypic parameters were widely spread over the parameter search spaces, except for lp1 (pedicel conductivity for water transport in the phloem), which varied in a narrow range in the population (**Figure 5**). A principal component analysis was performed on the estimated parameter values. The first three components explained 24, 22, and 16% of the variance, respectively (62% in total). On the first principal component (**Figure 6B**), we observed negative loadings of lx and lp—which are the parameters related to the membrane permeability—and lx1 and lp1—related to the pedicel hydraulic conductance (**Table 1**). So the first axis was mainly associated with parameters controlling water inflow to the fruit from the xylem and phloem tissues. PhiMax which impacts the cell wall plasticity, as well as tauA, nuM, and tauS, which tunes the active sugar uptake and the sugar mass flow intensities, respectively, had a high loading on the second principal component. Parameter phiMax had the highest impact with a negative value, opposite

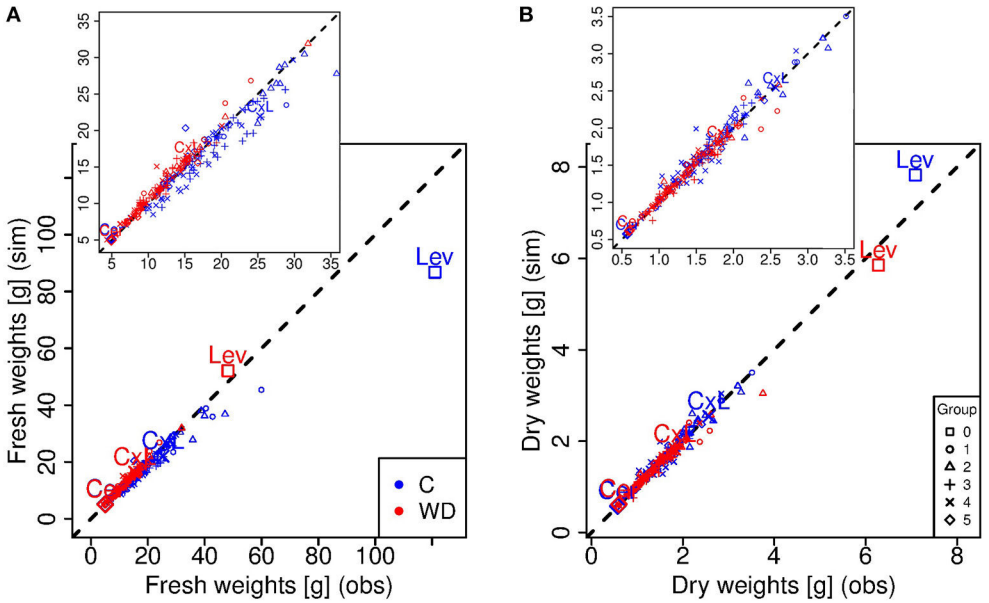


FIGURE 3 | Simulated (y-axis) *versus* observed (x-axis) fresh (A) and dry **(B)** weight values at the red ripe stage for each genotype in control (C, blue) and water deficit (WD, red) conditions. Each point represents one genotype (means of 15 to 20 fruits) and crosses indicate the parental lines (Cervil and Levovil) and the F1 hybrid (CxL). The different symbols represent the six groups of genotypes shown in **Figure 1**. The dashed line represents the points in which the simulated values are equal to the observed one.

with respect to nuM. So the second axis was mainly associated to turgor-driven cell expansion and active sugar uptake.

The fruit dry and fresh masses, the dry matter content under C and WD conditions, and the fresh mass reduction under WD were projected as inactive variables on the correlation circle (**Figure 6C**). Dry matter content in C condition and fresh mass reduction in WD condition correlated negatively. High dry matter content in C condition was associated with high value of tauS (parameter referring to mass flow process). The fresh mass reduction was associated with low values of nuM (maximum active uptake of sugar) and tauS, while it was associated with high values of phiMax (cell wall extensibility) and tauA (whose high values mean a slower decrease of active uptake rate of sugar during fruit development). Five clusters were selected through hierarchical cluster analysis (**Figure 6D**). The two parental lines and the F1 hybrid belonged to different groups. Levovil constituted a single cluster characterized by high fresh weight (fw) and dry weight (dw) under both conditions (first axis) and high reduction in fresh mass (second axis), with high values of lp (on the first component), phiMax and tauA (on the second component). Cluster 1 (including Cervil) and cluster 5 were characterized by high dm content under C conditions and low loss of fresh mass under WD, despite the fact they were associated with different active variables. Cluster 1 was associated with high membrane-permeability and high pedicel-conductance, whereas cluster 5 was associated with high values of nuM and low values of phiMax and tauA. Clusters 2 and 4 overlapped near to the origin of the first two components plan and were characterized by a high loss of fresh mass under WD and low dry matter content under control conditions. The CxL F1 hybrid belonged to cluster 2 and was positioned far from the cluster center.

QTL of Model Genotypic Parameters in the Population

A QTL analysis of the eight model parameters (**Table 1**, best estimation for each RIL line) and the coordinates of the RIL on the PCA axes, was performed independently of the irrigation level. Results are presented in **Table 3**. Two QTLs were detected with a genome wide significance level of $\alpha = 0.05$, on chromosome 2 and 8. These QTLs were associated with lp1 (pedicel conductivity) and lp (composite membrane conductivity) and explained 14 and 13% of the trait variations, respectively (**Table 3**). When considering less stringent significance levels, six supplementary QTLs became significant which explained between 9 and 11% of the parameter variations. Three of them were related to conductivity (lx1, rxp and lx, α between 0.10 and 0.30), one was related to sugar active uptake (nuM, α between 0.20 and 0.30), and two QTLs were associated with the second and third axes of the PCA (α between 0.10 and 0.20). No significant QTL was detected for phiMax, a parameter associated with cell wall extensibility, even when considering lowered significance thresholds.

Among the eight identified QTLs, two colocalized on top of chromosome 8 (for lp and lx), two in the centromeric region of chromosome 4 (for rxp and lx1) and two on top of chromosome 7 (for nuM and axis 2), which may correspond to three unique QTLs. Except for the QTL for rxp on chromosome 4 which included 192 genes over 0.18 Mbp, the QTL intervals were rather large (from 3.6 to 54.95 Mbp, including between 480 and 1180 genes). QTLs for fruit and plant traits detected in the same regions in the same population grown under control and WD conditions (Albert et al., 2016) are indicated in the last column of **Table 3**.

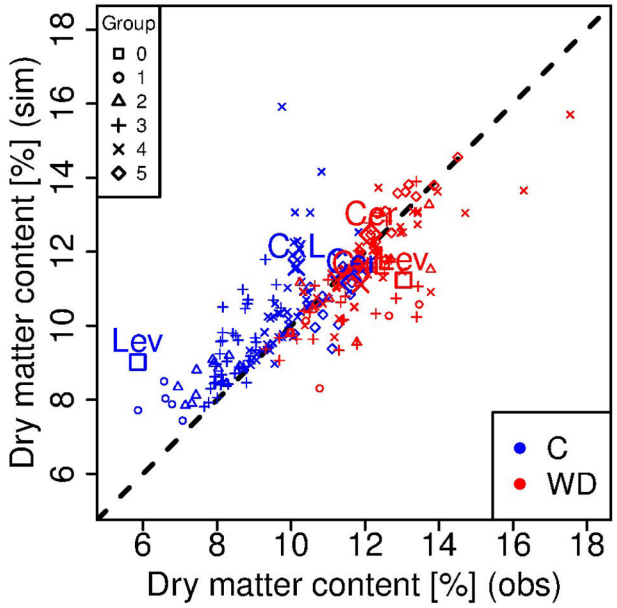


FIGURE 4 | Observed *vs.* simulated dry matter contents at the red ripe stage for each genotype in control (C, blue) and water deficit (WD, red) conditions. Each point represents one genotype (means of 15 to 20 fruits) and crosses indicate the parental lines (Cervil and Levovil) and the F1 hybrid (CxL). The different symbols represent the six groups of genotypes shown in **Figure 1**. The dashed line represents the points in which the simulated values are equal to the observed one.

Design of Tomato Ideotypes to Minimize the Reduction of Fruit Fresh Mass under WD Conditions and Optimize Fruit Dry Matter Content under C Conditions

The design of ideotypes consisted in finding sets of model parameters to match one or more objectives under a given condition (C or WD). Because we observed significant decrease in fruit fw under WD conditions and because fruit dm content is associated with high sugar and acid contents, our objectives were to maximize the dry matter content in C conditions and to minimize the fresh weight loss under the water deficit modality. Moreover, since the sensitivity to WD depends on the fruit fw (**Figure 2**), three classes of fruit grades were considered: large size (100–300 g), medium size (20–80 g) and small size (5–15 g).

In this step, 11 genotypic parameters were considered (**Table 1**), the eight parameters estimated in the calibration step, the initial fresh and dry mass of fruit and one parameter related to sugar content. The threshold between the two objectives is highlighted on **Figure 7**. For the group of large fruits (**Figure 7A**), the objectives were largely improved (dm content around 9% under C conditions and fresh mass loss around 15% under WD) with respect to Levovil (5.5% dm content under C condition and 60% fresh mass loss under WD), the only genotype belonging to the 100–300 g interval. The median fresh weight of the ideotypes in this group was 113 g in C conditions. In the medium-size group (**Figure 7B**) and in the small-size group (**Figure 7C**) of fruits, we obtained a large improvement too, since our selected ideotypes contained between 8 and 10% dm, which was comparable to the RIL population; however, they lost less than 15% fresh mass under WD conditions, which was two to three times less than the RIL population. In these two groups, the median fresh weight in C condition was 21 and 7 g, respectively. Interestingly in the small-size group of fruits, five RILs were in the scatter plot of selected ideotypes, and likely bear interesting traits and alleles for adaptation to WD. These are Cervil, SSD84, SSD107, SSD121, and SSD154.

A PCA was performed on the ideotype parameters obtained through the optimization problem resolution in order to understand the mechanisms of water stress resistance that could be combined in "ideal" fruits. In order to compare the ideotypes and the RILs, the eight genotypic parameters calibrated on the RILs were used as active variables (**Figure 8B**). The three additional parameters estimated for the ideotypes (bssrat, s0, w0) as well as the calculated fresh mass loss under WD, the fresh and dry mass and the dm content under non limiting water supply were projected as inactive variables (**Figure 8C**). The first two principal components explained 49.4 and 14.9% of the variation, respectively (**Figure 8A**). On the first component, lx and lp (composite membrane conductivity) had a positive loading and lx1 and lp1 (pedicel conductivities) had a negative loading. nuM (maximum rate of active sugar uptake) and tauA (negatively linked to the decrease in active sugar uptake rate during the fruit development) showed a negative and a positive loading, respectively, on this first component. The three groups of fruit grades were well-separated in the PC space especially for the big-size fruits (**Figure 8D**, cluster 1 blue colored). The big fruit-size ideotypes (100–300 g) were associated with high value of nuM, high pedicel conductivity and low fruit composite membrane conductivity, suggesting that sugar transport and pedicel conductivity may be interesting issues for improving adaptation of large-fruited genotypes to WD. They were also associated with high fresh mass loss (considered as inactive variable; **Figure 8C**). On the contrary, the cell wall extensibility (phiMax) and the mass flow characteristics (tauS) did not strongly discriminate the ideotype population.

The active variables were correlated in a different way with respect to the calibration situation (**Figure 6A**). In the ideotype principal component space, tstar and tauS correlated to each other, whereas in the calibration parameters space they were uncorrelated. Parameters regarding conductivities were highly correlated in both spaces; however, in the ideotype situation, the pedicel conductivities lp1 and lx1 were negatively correlated with the composite membrane ones (lp and lx). nuM did not show any positive correlation in the calibration case, whereas it correlated positively with the pedicel conductivities in the ideotype case. Most of the RILs, when projected as inactive individuals (**Figure 8A**), lied for in the positive values of the first principal component and in the negative values of the second principal component. Most of the RILs have medium to small-sized fruits (**Figure 2**), but, in the ideotype principal component plan, their positions did not completely overlap the corresponding group of ideotypes (n°2). Levovil is the only large fruited line and it did not belong to ideotype group 1. Cervil

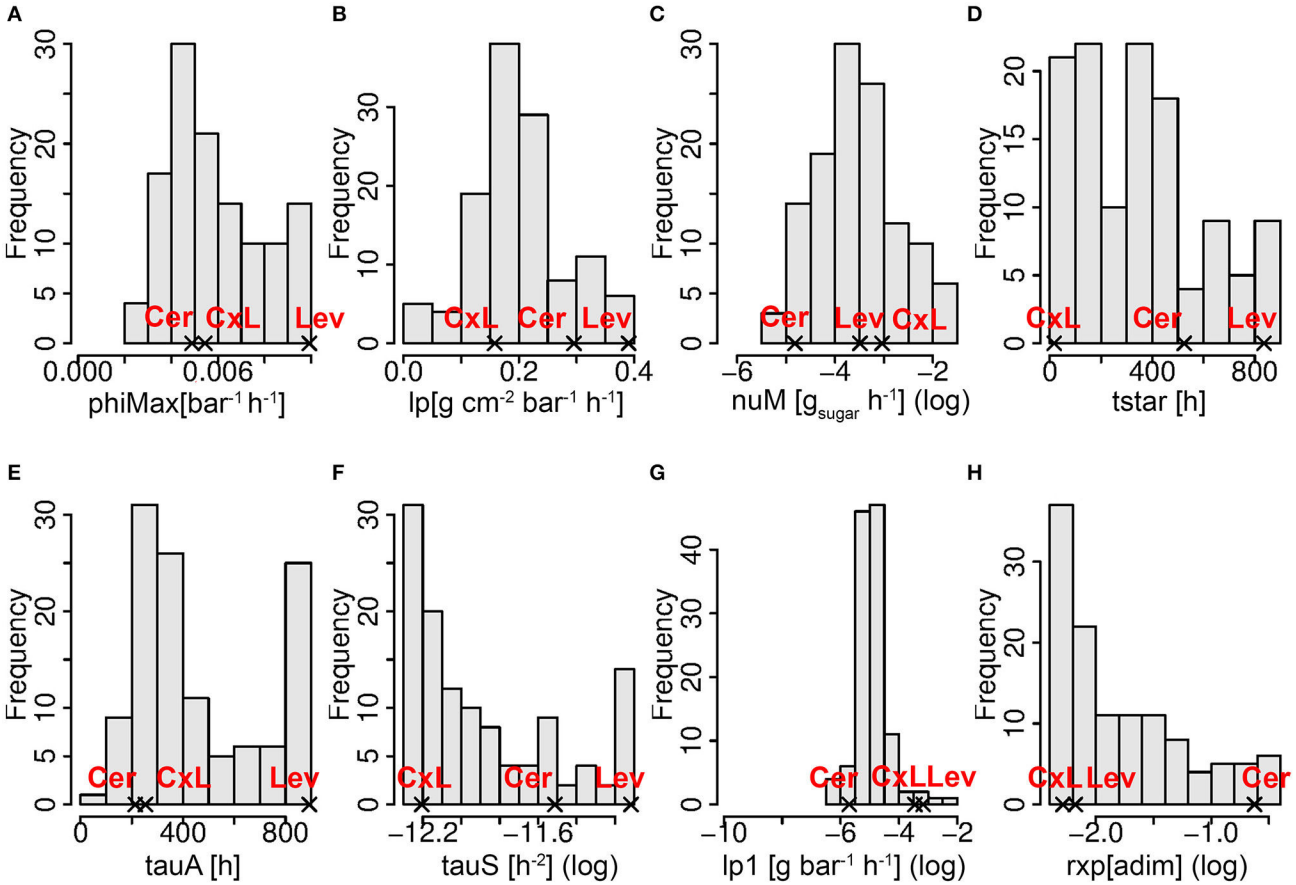


FIGURE 5 | Frequency distributions of the best estimated parameter sets in the RIL population including the two parental lines (Cer: cervil and Lev: levovil), the F1 hybrid (CxL), and the 117 RILs. Panels **(A–H)** represent the frequency distributions of the eight genotypic parameters as indicated in the x-axes title. The x-axes range corresponds to the parameter search space (see **Table 1**). The log in brackets indicates that a natural logarithmic transformation was applied on the variable, for a better legibility.

belonged to its size group. CxL belonged to a region that is the farthest one to group 2.

DISCUSSION

Model-Based Analysis of the Processes Involved in Genetic Variability of Fruit Response to Water Deficit

In this study, we applied a long and severe WD, which caused significant decrease in fruit fresh mass and increase in dm content for most of the RILs. Large fruits were the most sensitive in terms of fresh mass reduction, in agreement with previous studies (Ripoll et al., 2015; Albert et al., 2016). The model was able to reproduce fairly well the genetic variabilities observed in the population and the WD effects, after the calibration of eight genotypic parameters, which are related to cell expansion, water transport, and sugar uptake. These three processes were strongly discriminant in the RIL population (**Figure 6**), and appeared as main traits to be considered in the future for breeding tomato adapted to WD conditions. Interestingly in the population, a few genotypes (among which Cervil) reached comparable fresh and dry masses under C and WD conditions. All these genotypes are in the range 5–20 g FW and 9–12% dm (**Figure 2**), but they clustered in groups 2, 4, and 5, whereas Cervil was in group 1 (**Figure 6**). Thus, the low sensitivity to WD could not be related to one single parameter/process, and a more detailed phenotyping of these RILs at the plant and fruit levels could be useful to identify the main discriminating traits. According to the PCA on the calibrated genotypic parameters, conductivities merit special attention since high conductivities were associated with high dry matter content and heavy fruit weights. The sugar active uptake seemed to play an important role as well: the higher the maximum uptake rate was (nuM), the lower the decrease in fresh mass, which was probably associated with osmotic regulations.

The mean NRMSE value of the population was around 17%, which is quite performing (**Table 2**). The worst value was obtained for Levovil, but the fitting were largely improved when predictions were done independently under C or WD conditions (not shown), suggesting that some WD effects were poorly taken into account in this case. During the calibration step, the WD treatment was taken into account through several

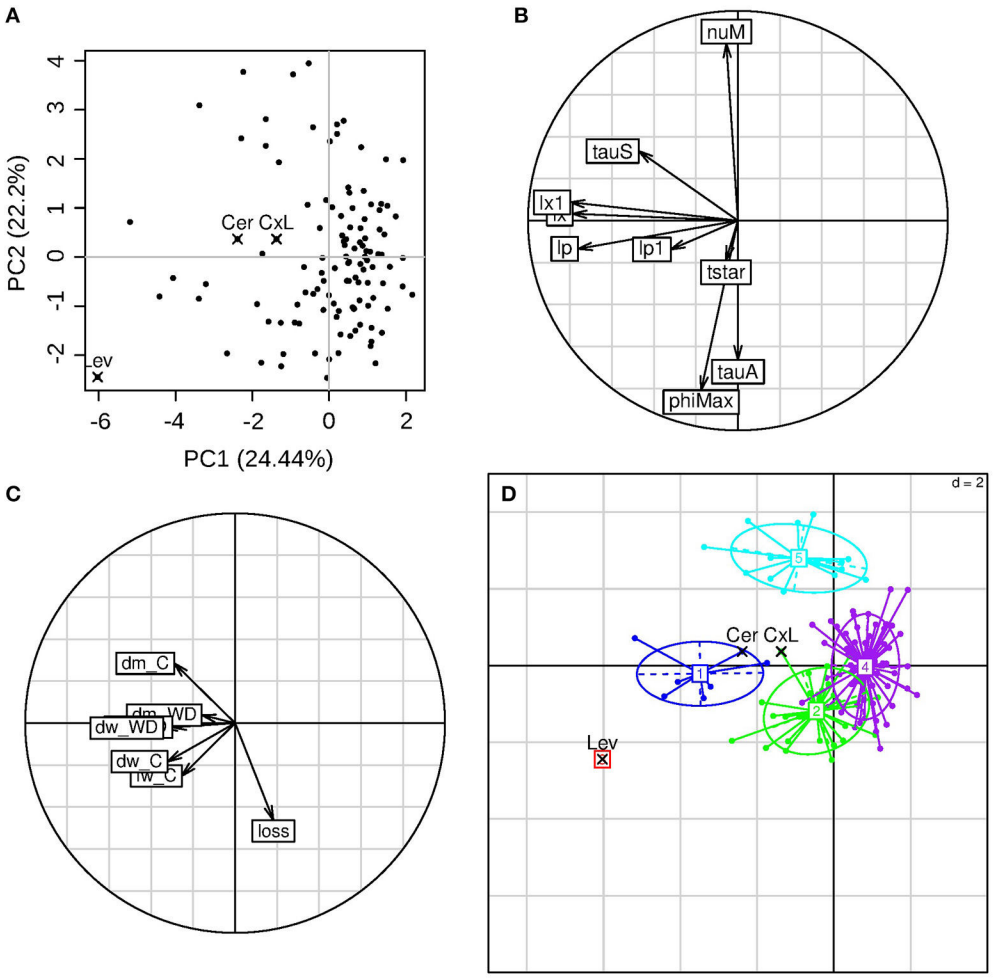


FIGURE 6 | PCA of the calibration parameter results. (A) Individual's positions in the first two principal components plane, with highlighted values for the parental lines (Cervil and Levovil) and F1 hybrid (CxL); each component explained variance is showed in parenthesis. **(B)** Variable coordinates normed to the square root of the eigenvalue represented in the correlation circle. **(C)** Inactive variables coordinates in the correlation circle: *dm* indicates the dry matter content, *dw* and *fw* state for dry and fresh weights of ripe fruits, respectively, *loss* indicates the fresh mass loss index as computed in Equation (11); **(D)** Output clusters from the hierarchical cluster analysis.

plant and fruit variables, i.e., the stem water potential, the initial dry and fresh masses, the sugar concentration of the dry matter, the fruit conductance involved in transpiration and the fruit osmotic pressure related to soluble components other than sugars. These variables were measured experimentally. On the contrary, other parameters were fixed since they cannot be easily measured. In the future, these parameters could be more deeply investigated. For instance, the impact of drought on phloem transport has been nicely illustrated through current model hypotheses (Sevanto, 2014). Accordingly, in the Virtual fruit model, the assumptions of impermeable conduit walls in the fruit pedicel and semipermeable walls in the fruit cells, implicitly involve that phloem transport in the pedicel is vulnerable to the increase of viscosity and to the geometry (number and size) of the conducting vessels. Sevanto (2014) demonstrated that wider or more numerous conduits are required to compensate for the increase in sap viscosity in order to maintain phloem transport under drought. In the present study, WD effects on these two parameters were overlooked. Indeed, in the absence of experimental value, the sugar concentration in phloem sap (Cp) was supposed to be constant over fruit development and the surface (cm2) of exchange of the vascular networks entering the fruit, was assumed to be proportional to the fruit surface area (Af), according to a non-dimensional constant coefficient, leading to smaller exchange area in case of the WD treatment. Thus, the conductivity of the phloem and xylem in the pedicel (lp1 and lx1) or in the fruit (lp and lx), which were estimated, likely integrated several properties of the conducting tissues. Both in the calibrating step (**Figure 6**) and the ideotype design (**Figure 8**), these parameters were highly discriminant and undoubtedly involved in the reduction of fresh weight loss under WD. So sugar concentration in

TABLE 3 | QTLs detected on eight genotypic parameters of the model and on the first three axes of the PCA estimated on the RIL population.

Trait	Sign.	LOD	Chr	Pos	Marker	CI cM (Mpb)	Nb genes	Mean Cer (sd)	Mean Lev (sd)	PVE	Coloc. (Albert et al., 2016)**
lp1*	0.05	3.85	2	95.60	TG167_Y02_52393366	89.73–107.19 (51.19 –54.79)	480	0.01 (0.00)	0.01 (0.00)	13.75	Nbfruits.C&WD fw.C&WD FIR.WD FIR.WD dw.C SSC.Int
rxp*	0.20	2.68	4	36.73	Y04_03230589	33.63–52.44 (0.30–0.48)	192	0.17 (0.12) (0.12)	0.22 (0.13)	10.22	Ø
lx1*	0.10	2.81	4	61.27	Y04_53862540	2.06–63.34 (0.42–55.37)	1604	0.001 (0.00)	0.002 (0.00)	10.7	FIR.C and WD
Axis3	0.20	2.46	4	86.96	Y04_61146494	61.27–95.70 (53.86–62.08)	589 589	−0.36 (0.12)	0.26 (0.13)	9.37	Flw.C Flw.WD Diam.C fw.C FIR.C&WD dw.C VitCFM.C&WD Yield.C
nuM	0.30	2.30	7	93.31	Y07_67908188	82.11–93.31 (65.13–67.90)	408	0.05 (0.01)	0.03 (0.00)	8.80	Ø
Axis2	0.20	2.49	7	88.00	Y07_64327204	73.63–93.31 (63.64–67.90)	575	0.51 (0.20)	−0.44 (0.19)	9.15	Ø
lp	0.05	3.31	8	42.12	Y08_57208257	31.67–58.97 (54.32–59.92)	479	0.22(0.08)	0.16 (0.06)	12.64	pH.WD VitCFM.WD VitCDM.C&WD
lx*	0.30	2.24	8	42.12	Y08_57208257	31.67–101.95 (54.32–65.60)	1180	0.51(0.05)	0.03 (0.01)	8.71	Flw.WD pH.WD VitCFM.WD VitCDM.C&WD Yield.Int

**Traits transformed to ensure a normal distribution, LOG10(lp1); 1/rxp; LOG10(lx1); LOG10(lx).*

***Nbfruits, plant fruit number; fw, fruit fresh weight; FIR, fruit firmness; dw, fruit dry matter weight; SSC, solid soluble content; Flw, flowering time; Diam, stem diameter; pH, fruit pH; VitCFM, vitamin C content in fruit on a fresh weight basis; VitCDM, vitamin C content in fruit on a dry weight basis; Yield, fruit fresh weight per plant; C, control; WD, water deficit; Int, interactive between watering regimes.*

"Sign." indicates the significance threshold at which the QTL was detected. LOD is the log-likelihood at that marker. The chromosome is indicated under "Chr" and the position of the QTL is expressed in Haldane cM under "Pos." The most closely associated marker is indicated. CI indicates the genetic confidence interval in Haldane cM calculated by LOD decrease of one unit, and its physical equivalent (Mpb) on genome assembly 2.5 between brackets. The number of genes in the QTL interval (genome annotation 2.4) is indicated. The average value of the two parental alleles (Cer and Lev, with the standard deviation between brackets) and the percentage of phenotypic variation explained by the QTL (PVE) are displayed. Colocalizations with phenotypic QTLs detected in Albert et al. (2016) in the same RIL population are indicated. C, QTL specific to the control condition; D, QTL specific to the water deficit condition; C&WD, QTL detected under both condition; Int, QTL with effect changing intensity or direction according to the watering conditions.

the phloem and geometry of conducts appeared as important components of the water deficit adaptation strategies, which have to be more deeply investigated as well as their genetic variability.

The maximum cell wall extensibility (phiMax) also appeared as a discriminant parameter in the population. Despite the growing number of studies and methods to investigate cell wall extensibility and elasticity (Cosgrove, 2016), data are currently missing to parameterize fruit models. In the present model, cell wall extensibility decreases exponentially with time from phiMax, which was considered to be genotype dependent, whereas the rate of decrease was constant and taken from Liu et al. (2007). In the RIL population, phiMax was negatively correlated with the maximum rate of sugar active uptake (nuM) and was associated with high loss of fresh mass under WD, likely because the large fruit-size genotypes (mainly Lev) combined both traits in this population.

QTL Analysis of Model Genotypic Parameters

The added value of QTLs for model parameters lies in their expected stability, on the contrary to other QTLs for phenotypic

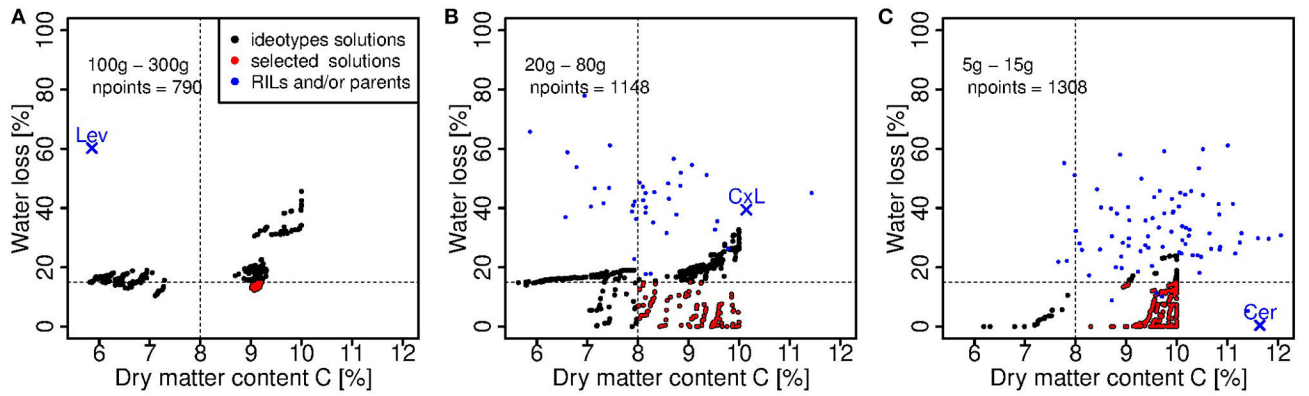


FIGURE 7 | Ideotype solutions considering a fruit fresh weight interval of (A): 100–300 g, **(B)**: 20–80 g, **(C)**: 5–15 g. The black and red points represent the dry matter content of ripe fruits in C conditions and the fresh mass loss value of all the ideotype solutions obtained solving the multi-objective problem for each interval of fruit fresh weight. The red points are the ideotype solutions satisfying simultaneously a high dry matter content in C condition (>8%) and a low loss of fresh mass under WD (<15%), according to Equation (13). npoints is the number of solutions. The blue points represent the dry matter content of ripe fruits in C conditions and the fresh mass loss values computed for the observed individuals in the RIL population, in the respective weight class. The parental lines (Cervil and Levovil) and F1 hybrid (CxL) are highlighted.

traits, which fluctuate with environmental conditions. We detected QTLs for six of the 10 genotypic parameters and two PCA axes, four of them (for Lp1, Lx1, and Lp) with a significance level below 0.1. The QTLs were located in four chromosomic regions. The QTL for lp1 on chromosome 2 colocalized with QTLs detected in Albert et al. (2016) for plant fruit number (constitutive effect under control and WD conditions), fruit firmness (constitutive effect under control and WD conditions), dm (specific to control condition) and soluble solid content (SSC, with changing effect according to the watering regime). Besides, this QTL was present in the genomic region carrying the cloned tomato fresh weight QTL FW2.2 (Frary et al., 2000). Close to this QTL, several QTLs for sugar content were fine mapped (Lecomte et al., 2004) The QTLs for rxp, lx1 and axis3 in the centromeric region of chromosome 4 colocalized with a QTL for fruit firmness (FIR, specific to the WD condition) detected in Albert et al. (2016); this connection between conductivity and firmness under WD may result from the effect of turgor regulation on fruit firmness (Shackel et al., 1991). QTLs detected for nuM and axis 2 (related to sugar active uptake) did not colocalize with any QTL identified in Albert et al. (2016). However, QTL for soluble solid content were identified in this genomic region in other tomato populations (Pascual et al., 2016). Finally, the QTLs for lp and lx on top of chromosome 8 colocalized with QTLs for flowering time (Flw, specific to WD condition), fruit pH (specific to WD condition), vitamin C content in fruit on a fresh weight and on a dry weight basis (WD specific and constitutive, respectively) and yield (with changing effect according to the watering regime). Unfortunately the confidence intervals were too large to check for candidate genes, but future studies should more deeply investigate these regions, in particular regarding the pedicel and fruit conductivities.

Interestingly, the seven genotypes belonging to group 1 in the PCA (**Figure 6**) all carried the Cervil allele for the lp and lx QTLs on chromosome 8. Besides, these genotypes also carried the Cervil allele for a yield QTL detected by (Albert et al., 2016) on chromosome 4, close to the QTLs identified for rxp and lx1 (3.16 Mbp upper on the chromosome). No specific allele at the QTL was identified for the other groups of the PCA. We detected only one QTL for one of the four genotypic parameters related to sugar uptake, hypothetically due to their skewed distribution and to the estimation error. In the Fruit model, sugars may be allocated to fruit through passive diffusion, mass flow or active transport, the last being the most discriminant in our population. Different transporters are required for efficient phloem unloading in fruit pericarp at the rapid expansion phase (Ruan and Patrick, 1995), operating with different energetic and kinetic constraints (reviewed by Osorio et al., 2014). In the model, all these transporters are compiled into one single Michaelis-Menten function, which might explain that no specific QTL was detected.

Ideotypes of Tomato Adapted to WD Depending on Trade-Offs between Fresh Weight and Dry Matter Content

In the light of experimental data, one challenge for producing tomatoes under WD conditions will be to avoid the reduction of fresh mass of large fruited genotypes, while maintaining or increasing the fruit dm content, which correlates with the accumulation of sugars and acids, both involved in fruit taste. In confront to previous works (Semenov et al., 2014; Rötter et al., 2015), the problem was complex, first, because the process-based model used in this study was relatively sophisticated, second, because we aimed at maintaining quality and increasing yield under WD conditions. We were able to design large-fruited ideotypes rich in dm (9% dm content) and with reasonable fresh mass loss under WD (<20%) which outperformed Levovil

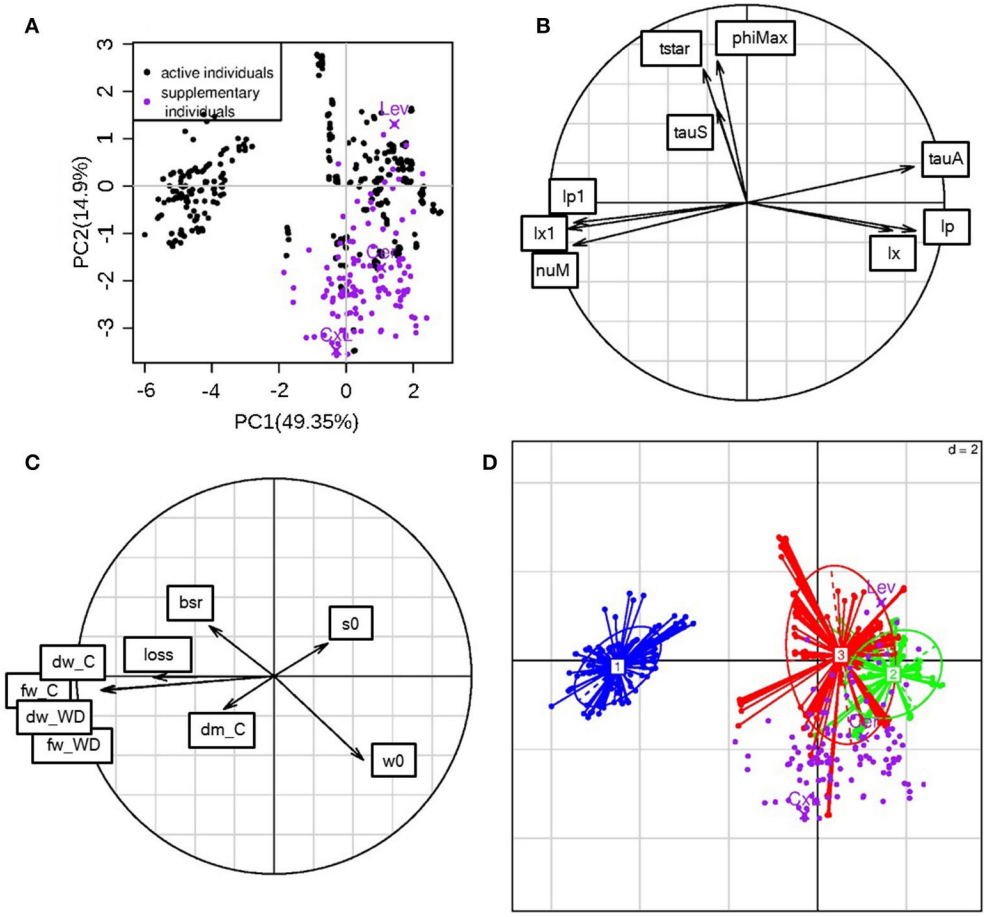


FIGURE 8 | PCA performed on the ideotype parameters. (A): ideotype positions in the first two principal components plane (black dots), and projection as inactive variables of the parameter values calibrated for the RIL population (purple dots); The parental lines (Cervil and Levovil) and F1 hybrid (CxL) are highlighted; the % of variance explained on each component is given in parenthesis. **(B)**: variable coordinates normed to the square root of the eigenvalue represented in the correlation circle. **(C)** Inactive variable coordinates in the correlation circle. *bsr* is the *bssrat* parameter, *s0* and *w0* are the estimated initial dry and fresh weights; *dm_C* states for dry matter content in C condition and *dw_C* for dry weight in C condition. Three more variables were hidden by this last one: *fw_C, fw_WD,* and *dw_WD* representing fresh weight under C and WD, and dry weight under WD, respectively **(D)**: ideotypes divided into fruit grade groups: (1) 100–300 g (blue), (2) 20–80 g (green), (3) 5–15 g (red). Purple dots represent the projections of the RILs.

(60% fresh mass loss and 6% dm content). Pedicel conductivity and fruit composite membrane conductivity were opposed in the ideotype population (**Figure 8**). The model considers three pathways for water and carbon flows: the plant-to-pedicel, the pedicel-to-fruit and the fruit apoplasm-to-cell, which all differ in carbon concentration and water potential. Conductance is mainly axial in the first two pathways (plant-pedicel and pedicel-fruit), whereas it is radial within the fruit. Thus high conductance in the pedicel which promotes water and sugar inflows in combination with high active uptake of sugars could be a successful strategy to produce large fruit-size ideotypes able to maintain, under WD conditions, a fresh weight above 100 g and dm content above 6 % (group 1 on **Figure 8**). These ideotypes also exhibit a low tauA value, indicating that the active uptake decreases slowly during the growth stage. On the contrary, the medium fruit-size ideotypes (group 2) were associated with low pedicel conductance and sugar uptake rate, but high fruit composite membrane conductivity. Such interactions between pedicel and fruit conductivities in relation to the demand for water and carbon, is intriguing and should deserve further attention. As mentioned above, the genetic variability of the conducting tissue geometry in the pedicel and fruit has been hardly described. In an anatomical descriptive study, Rančić et al. (2010) suggested that the low phloem efficiency (defined as the ratio between fruit dry weight and phloem pedicel area) of tomato flacca mutants was responsible for low fruit growth, whereas the phloem area and the functional xylem area were not affected compared to the wild type. In agreement, a modeling approach showed that, under a wide range of conditions, water import in young tomato fruit would be limited by the pedicel resistance (Bussières, 2002) and by phloem conductivity in relation to sap viscosity (Bussiéres et al., 2011). In this respect, QTL observed on chromosome 2, 4, and 8 may be really interesting. Regarding the small fruit-size genotypes, the four RILs, which scattered among

the ideotypes (**Figure 7C**) could probably bring new interesting source of genetic variations for breeding programs, as their fruit fresh mass at maturity is two to three times higher than the fresh mass of Cervil, which is already known to be WD resistant (Albert et al., 2016; Ripoll et al., 2016).

Comparing the RIL population (**Figure 6**) with the ideotype population (**Figure 8**), a main difference states in the orthogonality of nuM and lx/lp or lx1/lp1 in the RIL population, suggesting that fruit growth was limited either by the incoming fluxes (group 5 on **Figure 6**) or by the active transport of sugars (Lev or group 1 on **Figure 6**). Thus, the uptake of carbon was likely the limiting step for fruit growth of large fruit-size genotypes such as Lev, which bears large fruits with low dry matter content in the C condition.

CONCLUSIONS

The fruit model was able to reproduce contrasting behaviors in the RIL population, regarding fresh weight loss and/or dm content increase under WD. Cell expansion, water transport and sugar uptake were all involved in the genetic variability of the fruit response to WD, but pedicel conductivity and active uptake of sugars seemed to be the key-mechanisms. In the future, model improvements should account for WD effects on cell wall extensibility, sugar uptake and exchanges of water between phloem and xylem tissues. Such advances will boost our understanding of the complex interactions between osmotic adjustments, changes in cell wall extensibility and maintenance of cell turgor under WD. The present study also outlined three interesting QTLs that deserve attention in breeding program for adaptation to WD and 4 RILs, which could bring new interesting traits in this regard. Finally, we are aware of the fact that Levovil is the only big fruit size genotype in the studied population; all other RILs ranged between 5 and 60 g FW. Thus, applying our approach to other tomato populations will be valuable. In a longer-term perspective, using a plant-fruit model (Baldazzi et al., 2013) would allow a better assessment of the respective contribution of source and sink capacities to the genetic variability. From a methodological point of view, other algorithms e.g., the Reference-point-based Non-dominated Sorting Genetic Algorithm R-NSGA-II could be used to calibrate the process-based model considering individual errors for each output variables, which could open new perspectives regarding the accurate integration of genetic information into the process-based model. Indeed, here ideotypes were discussed without considering the genetic constraints (epistasis or linkage). For this purpose, we should first integrate the genetic information into the process-based model through the QTL analysis. Then, we shall consider the allelic combinations of the loci involved in the genetic models allowing the computation of the parameter values. In this step, taking into account the genetic constraints (probabilities of two loci to be identical) shall be achieved either through their direct integration into the genetic model or through the optimization algorithm (mathematical formulation of the problem). In the future, such genetic models could be used to test virtual scenarios of fruit adaptation to water stress and identify key-regions on the tomato genome.

AUTHOR CONTRIBUTIONS

NB, MG, GV, and MM conceived and designed the work; BB and NB collected the experimental data; DC, PV, and MM devised the algorithms and performed the simulations; DC, NB, MG, GV, EA, MC, and MM analyzed and interpreted the data; DC, NB, MM, GV, and EA wrote the paper; MG, MC, and VB revised it critically; all authors approved the final version.

FUNDING

INRA provided all experimental and modeling supports as well as permanent manpower.

ACKNOWLEDGMENTS

This project has been supported by the CTPS Project TOMSEC, by the ANR project ADAPTOM (ANR-13-ADAP-0013-01) and by Agropolis Foundation under the reference ID 1202-039 through the "Investissements d'avenir" Programme (Labex Agro: ANR-10-LABX-0001-01). We acknowledge the experimental teams of UR GAFL (Yolande Carretero, Alain Goujon) and PSH.

REFERENCES

Albert, E., Gricourt, J., Bertin, N., Bonnefoi, J., Pateyron, S., Tamby, J. P., et al. (2016). Genotype by watering regime interaction in cultivated tomato: lessons from linkage mapping and gene expression. *Theor. Appl. Genet.* 129, 395–418. doi: 10.1007/s00122-015-2635-5

Baldazzi, V., Pinet, A., Vercambre, G., Bénard, C., Biais, B., and Génard, M. (2013). *In-silico* analysis of water and carbon relations under stress conditions. A multi-scale perspective centered on fruit. *Front. Plant Sci.* 4:495. doi: 10.3389/fpls.2013.00495

Bertin, N., Martre, P., Génard, M., Quilot, B., and Salon, C. (2010). Under what circumstances can process-based simulation models link genotype to phenotype for complex traits? Case-study of fruit and grain quality traits. *J. Exp. Bot.* 61, 955–967. doi: 10.1093/jxb/erp377

Blum, A. (2011). *Plant Breeding for Water-Limited Environments*. New York, NY: Springer.

Bodner, G., Nakhforoosh, A., and Kaul, H.-P. (2015). Management of crop water under drought: a review. *Agron. Sustain. Dev.* 35, 401–442. doi: 10.1007/s13593-015-0283-4

Boote, K. J., Kropff, M. J., and Bindraban, P. S. (2001). Physiology and modelling of traits in crop plants: implications for genetic improvement. *Agric. Syst.* 70, 395–420. doi: 10.1016/S0308-521X(01)00053-1

Broman, K. W., Wu, H., Sen, Ś., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, 19, 889–890. doi: 10.1093/bioinformatics/btg112

Bussières, P. (2002). Water import in the young tomato fruit limited by pedicel resistance and calyx transpiration. *Funct. Plant Biol.* 29, 631–641. doi: 10.1071/PP00144

Bussiéres, P., Bertin, N., Morris, C. E., Vigne, C., Orlando, P., Glaux, C., et al. (2011). High external sucrose concentration inhibits the expansion of detached tomato fruits grown in a novel semi-open device. *In Vitro Cell. Dev. Biol. Plant* 47, 743–751. doi: 10.1007/s11627-011-9378-z

Appendix 1

Chapman, S., Cooper, M., Podlich, D., and Hammer, G. (2003). Evaluating plant breeding strategies by simulating gene action and dryland environment effects. *Agron. J.* 95, 99–113. doi: 10.2134/agronj2003.0099

Cosgrove, D. J. (2016). Plant cell wall extensibility: connecting plant cell growth with cell wall structure, mechanics, and the action of wall-modifying enzymes. *J. Exp. Bot.* 67, 463–476. doi: 10.1093/jxb/erv511

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6, 182–197. doi: 10.1109/4235.996017

Ding, W., Xu, L., Wei, Y., Wu, F., Zhu, D., Zhang, Y., et al. (2016). Genetic algorithm based approach to optimize phenotypical traits of virtual rice. *J. Theor. Biol.* 403, 59–67. doi: 10.1016/j.jtbi.2016.05.006

Dray, S., and Dufour, A.-B. (2007). The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* 22, 1–20. doi: 10.18637/jss.v022.i04

Fishman, S., and Génard, M. (1998). A biophysical model of fruit growth: simulation of seasonal and diurnal dynamics of mass. *Plant Cell Environ.* 21, 739–752. doi: 10.1046/j.1365-3040.1998.00322.x

Frary, A., Nesbitt, T. C., Frary, A., Grandillo, S., Van Der Knaap, E., Cong, B., et al. (2000). Fw2. 2: a quantitative trait locus key to the evolution of tomato fruit size. *Science* 289, 85–88. doi: 10.1126/science.289.5476.85

Gong, P., Zhang, J., Li, H., Yang, C., Zhang, C., Zhang, X., et al. (2010). Transcriptional profiles of drought-responsive genes in modulating transcription signal transduction, and biochemical pathways in tomato. *J. Exp. Bot.* 61, 3563–3575. doi: 10.1093/jxb/erq167

Hall, A. J., Minchin, P. E. H., Clearwater, M. J., and Génard, M. (2013). A biophysical model of kiwifruit (*Actinidia deliciosa*) berry development. *J. Exp. Bot.* 64, 5473–5483. doi: 10.1093/jxb/ert317

Labate, J. A., Robertson, L. D., and Baldo, A. M. (2009). Multilocus sequence data reveal extensive departures from equilibrium in domesticated tomato (*Solanum lycopersicum* L.). *Heredity (Edinb).* 103, 257–267. doi: 10.1038/hdy.2009.58

Lander, E. S., and Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121, 185–199.

Lechaudel, M., Vercambre, G., Lescourret, F., Normand, F., and Génard, M. (2007). An analysis of elastic and plastic fruit growth of mango in response to various assimilate supplies. *Tree Physiol.* 27, 219–230. doi: 10.1093/treephys/27.2.219

Lecomte, L., Saliba-Colombani, V., Gautier, A., Gomez-Jimenez, M., Duffé, P., Buret, M., et al. (2004). Fine mapping of QTLs for the fruit architecture and composition in fresh market tomato, on the distal region of the long arm of chromosome 2. *Mol. Breed.* 13, 1–14. doi: 10.1023/B:MOLB.0000012325.77844.0c

Lescourret, F., and Génard, M. (2005). A virtual peach fruit model simulating changes in fruit quality during the final stage of fruit growth. *Tree Physiol.* 25, 1303–1315. doi: 10.1093/treephys/25.10.1303

Lescourret, F., Génard, M., Habib, R., and Fishman, S. (2001). Variation in surface conductance to water vapor diffusion in peach fruit and its effects on fruit growth assessed by a simulation model. *Tree Physiol.* 21, 735–741. doi: 10.1093/treephys/21.11.735

Letort, V., Mahe, P., Cournède, P. H., de Reffye, P., and Courtois, B. (2008). Quantitative genetics and functional-structural plant growth models: simulation of quantitative trait loci detection for model parameters and application to potential yield optimization. *Ann. Bot.* 101, 1243–1254. doi: 10.1093/aob/mcm197

Liu, H. F., Génard, M., Guichard, S., and Bertin, N. (2007). Model-assisted analysis of tomato fruit growth in relation to carbon and water fluxes. *J. Exp. Bot.* 58, 3567–3580. doi: 10.1093/jxb/erm202

Lu, C., Han, J. L., Hu, F. J., and Qin, T. G. (2012). Mathematical model of wheat stalk lodging-resistance during the later growth period. *Math. Pract. Theory* 42, 46–53. Available online at: http://en.cnki.com.cn/Journal_en/A-A002-SSJS-2012-15.htm

Mazzeo, M. (2008). *Xylem Transport Efficiency and Calcium Accumulation in Fruit of Actinidia deliciosa: Implications for Fruit Quality*. PhD. thesis, University of Basilicata, Potenza.

Osorio, S., Ruan, Y.-L., and Fernie, A. R. (2014). An update on source-to-sink carbon partitioning in tomato. *Front. Plant Sci.* 5:516. doi: 10.3389/fpls.2014.00516

Pascual, L., Albert, E., Sauvage, C., Duangjit, J., Bouchet, J. P., Bitton, F., et al. (2016). Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. *Plant Sci.* 242, 120–130. doi: 10.1016/j.plantsci.2015.06.017

Qi, R., Ma, Y. T., Hu, B. G., De Reffye, P., and Cournede, P. H. (2010). Optimization of source-sink dynamics in plant growth for ideotype breeding: a case study on maize. *Comput. Electron. Agric.* 71, 96–105. doi: 10.1016/j.compag.2009.12.008

Quilot, B., Kervella, J., Génard, M., and Lescourret, F. (2005). Analysing the genetic control of peach fruit quality through an ecophysiological model combined with a QTL approach. *J. Exp. Bot.* 56, 3083–3092. doi: 10.1093/jxb/eri305

Quilot-Turion, B., Ould-Sidi, M.-M., Kadrani, A., Hilgert, N., Génard, M., and Lescourret, F. (2012). Optimization of parameters of the 'Virtual Fruit' model to design peach genotype for sustainable production systems. *Eur. J. Agron.* 42, 34–48. doi: 10.1016/j.eja.2011.11.008

Rančić, D., Quarrie, S. P., and Pećinar, I. (2010). "Anatomy of tomato fruit and fruit pedicel during fruit development," in *Microscopy: Science, Technology, Applications and Education,* eds A. Méndez-Vilas and J. Díaz (Badajoz: Formatex), 851–861.

Rebolledo, M. C., Dingkuhn, M., Courtois, B., Gibon, Y., Clement-Vidal, A., Cruz, D. F., et al. (2015). Phenotypic and genetic dissection of component traits for early vigour in rice using plant growth modelling, sugar content analyses and association mapping. *J. Exp. Bot.* 66, 5555–5566. doi: 10.1093/jxb/erv258

Reymond, M., Muller, B., Leonardi, A., Charcosset, A., and Tardieu, F. (2003). Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiol.* 131, 664–675. doi: 10.1104/pp.013839

Ripoll, J., Urban, L., and Bertin, N. (2015). The potential of the MAGIC TOM parental accessions to explore the genetic variability in tomato acclimation to repeated cycles of water deficit and recovery. *Front. Plant Sci.* 6:1172. doi: 10.3389/fpls.2015.01172

Ripoll, J., Urban, L., Brunel, B., and Bertin, N. (2016). Water deficit effects on tomato quality depend on fruit developmental stage and genotype. *J. Plant Physiol.* 190, 26–35. doi: 10.1016/j.jplph.2015.10.006

Ripoll, J., Urban, L., Staudt, M., Lopez-Lauri, F., Bidel, L. P. R., and Bertin, N. (2014). Water shortage and quality of fleshy fruits–making the most of the unavoidable. *J. Exp. Bot.* 65, 4097–4117. doi: 10.1093/jxb/eru197

Rötter, R. P., Tao, F., Höhn, J. G., and Palosuo, T. (2015). Use of crop simulation modelling to aid ideotype design of future cereal cultivars. *J. Exp. Bot.* 66, 3463–3476. doi: 10.1093/jxb/erv098

Ruan, Y.-L., and Patrick, J. W. (1995). The cellular pathway of postphloem sugar transport in developing tomato fruit. *Planta*, 196, 434–444. doi: 10.1007/BF00203641

Saliba-Colombani, V., Causse, M., Langlois, D., Philouze, J., and Buret, M. (2001). Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theor. Appl. Genet.* 102, 259–272. doi: 10.1007/s001220051643

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Galtelli, D., et al. (2008). *Global Sensitivity Analysis: The Primer.* Chichester: John Wiley and sons. doi: 10.1002/9780470725184

Semenov, M. A., Stratonovitch, P., Alghabari, F., and Gooding, M. J. (2014). Adapting wheat in Europe for climate change. *J. Cereal Sci.* 59, 245–256. doi: 10.1016/j.jcs.2014.01.006

Sevanto, S. (2014). Phloem transport and drought. *J. Exp. Bot.* 65, 1751–1759. doi: 10.1093/jxb/ert467

Shackel, K. A., Greve, C., Labavitch, J. M., and Ahmadi, H. (1991). Cell turgor changes associated with ripening in tomato pericarp tissue. *Plant Physiol.* 97, 814–816. doi: 10.1104/pp.97.2.814

Sidi, M.-M. O., Quilot-Turion, B., Kadrani, A., Génard, M., and Lescourret, F. (2014). The Relationship between metaheuristics stopping criteria and performances: cases of NSGA-II and MOPSO-CD for sustainable peach fruit design. *Int. J. Appl. Metaheuristic Comput. (IJAMC)* 5, 44–70. doi: 10.4018/ijamc.2014070104

Tardieu, F. (2003). Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. *Trends Plant Sci.* 8, 9–14. doi: 10.1016/S1360-1385(02)00008-0

Tardieu, F. (2012). Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. *J. Exp. Bot.* 63, 25–31. doi: 10.1093/jxb/err269

Appendix 1

Wallach, D., Makowski, D., Jones, J. W., and Brun, F. (2013). *Working with Dynamic Crop Models: Methods, Tools and Examples for Agriculture and Environment*. London: Academic Press; Elsevier Ltd. doi: 10.1016/B978-0-12-397008-4.00011-3

White, J. W., and Hoogenboom, G. (1996). Simulating effects of genes for physiological traits in a process-oriented crop model. *Agron. J.* 88, 416–422. doi: 10.2134/agronj1996.00021962008800030009x

Xu, L., Zhu, J., Henke, M., Kurth, W., Ding, W., and Buck-Sorlin, G. H. (2012). "Simulating superior genotypes for plant height based on QTLs: towards virtual breeding of rice," in *The 4th International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications (PMA 12),* eds M. Kang, Y. Dumont, and Y. Guo Shanghai, 31.10.-03.11, 2012, IEEE, 447–454.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix 2: 'Developing tomato varieties with improved flavor'.

Book chapter XIII: Achieving sustainable tomato cultivation.

Ed A Matto and A Handa.

Accepted for publication, 2016

Appendix 2

Developing tomato varieties with improved flavour

M. Causse, E. Albert and C. Sauvage, INRA, France

Abstract:

Tomato fruit quality is a complex trait involving a number of components, including appearance, flavour, aroma and texture. There is a large range of genetic diversity in tomato for fruit quality components. Although a few major mutations may have a huge effect on fruit quality (notably the rin mutation), most of the components have a quantitative inheritance. Several Quantitative Trait Loci (QTL) mapping experiments have been performed, mostly on interspecific progeny. Many loci and QTL have thus been detected, revealing some QTL cluster regions. Tomato is a model plant for fruit development and composition, and knowledge about its physiology is rapidly increasing. This chapter examines the use of QTL to identify and determine favourable sensory characteristics, exploring current technologies and suggesting future trends for research in this area.

Key words: Tomato genetics; Tomato genome; Tomato sensory characteristics; Tomato flavour; Tomato texture; Quantitative Trait Loci

Appendix 2

1. Introduction

Today, tomato flavour is a key issue for tomato breeders. Over the last century, tomato breeders have improved tomato yield and yield stability and have adapted to diverse growth conditions. They have also introgressed many disease resistance genes from wild tomato relatives. Fruit quality has been improved mainly in the areas of shelf life, fruit homogeneity, shape and colour. However, consumers have frequently complained about tomato flavour over many years.

Improving taste by breeding is complex for several reasons. To begin with, sensory quality is a composite trait involving many components. Sugars and acids (responsible for sweet and sour flavours), as well as aroma (involving several volatiles) and texture (linked to firmness, meltiness, mealiness), contribute to flavour perception. Furthermore, the measurement of these compounds may be difficult; some are measured by only sensory analyses, but most of the components can be related to the chemical composition of the fruit. Tens of volatile compounds have been identified, but the list of those that are important for tomato aroma is very limited (Baldwin et al., 2000; Klee and Tieman, 2013).

In addition, most taste components are strongly influenced by the environment during plant and fruit growth and development (Causse et al., 2003), by the harvest stage (often immature) but also by post-harvest conditions (Kader et al., 1978; Whitaker, 2008). Many actors affect and can damage the flavour of a variety. Some important quality traits are also negatively correlated, like yield and sugar content or fruit shelf life and meltiness (due to physiological and genetic origin). This limits the options available for improving one trait without reducing others. Finally, quality has a subjective component based on individual consumer preferences. Economic factors can be a brake to optimizing taste which is not always sufficiently valued to attract a premium price for better varieties (Bellec-Gauche et al., 2015).

Nevertheless, it is well recognized that genetics has a fundamental impact on tomato flavour. Advances in molecular markers, and more recently the availability of the tomato genome sequence (Tomato Genome Consortium, 2012), have paved the way towards a better understanding of the genetic factors involved in fruit quality.

To improve flavour, several questions have to be addressed: What are consumer expectations? What is the genetic diversity available for breeding? What is the genetic control of tomato flavour? How to efficiently breed flavour? How can recent advances in genetics and genomics allow more efficient breeding? How environment influences flavour components? Emerging results to these questions will be presented in this section.

2. Genetic diversity of tomato flavour and consumer expectations

The first requirement to improve a trait is availability of genetic variability. Characterization of the genetic diversity tomato accessions for fruit quality components has revealed a large variation among traditional varieties and among wild related species for many traits as reviewed by a number of authors (Davies and Hobson, 1981; Stevens, 1986; Dorais et al., 2001; Causse et al., 2011). Metabolome profiling assessed several panels of tomato varieties and identified a large range of variations for primary and secondary metabolites (Schauer et al., 2006; Sauvage et al., 2014) as well as for volatile compounds (Tikunov et al., 2005; Bartoshuk and Klee, 2013; Rambla et al., 2014). All the metabolites appear to be influenced by growth conditions, varieties, ripening stages and storage conditions (Klee and Tieman, 2013).

Consumer preferences facing genetic diversity have been subject to a few studies (Sinesio et al., 2009; Causse et al., 2010). In the framework of a large European project, Eusol, 806 consumers from three countries (The Netherlands, France, and Italy) were presented with a set of 16 varieties representing the diversity of fresh tomato offer in order to evaluate their preferences. In parallel, expert panels in each country built sensory profiles of the varieties. Preference maps were then constructed in each country, revealing the structure of consumer preferences and allowing identification of the most important characteristics. Then, a global analysis revealed that preferences were quite homogeneous across countries. This study identified the overall flavour and firmness as the most important traits for improving tomato fruit quality. It showed that consumer preferences from different European countries, with different cultures and food practices, are segmented following similar patterns when projected onto a common referential plan. Moreover, the results clearly showed that diversification of taste and texture is required to satisfy all consumers' expectations as some consumers preferred firm tomatoes, while others preferred melting ones and were more or less demanding in terms of sweetness and flavour intensity. Detailed comparisons also showed the importance of the fruit appearance in consumer preference.

To study the inheritance of taste-related traits and the influence of growth conditions, Causse et al. (2003) analysed the genetic variation of quality attributes in 35 hybrids and their 13 parental lines, grown in two contrasted environments. The 13 parental lines had various origins (old traditional inbred cultivars, experimental lines bred in the 1980s and lines used as parents of modern hybrid varieties). Each experiment was grown in spring in soil-less glasshouse conditions and in summer in the open field or under unheated plastic tunnels, in order to estimate the overall influence of environmental conditions on quality traits. As fruit size influences the judgement of taste panels, two experiments were set up, one involving large fruits, the other small fruits from hybrids between

cherry tomato lines and large fruited lines. Among the main results on the genetic control of quality traits one can retain:

- Differences for sensory traits among genotypes may be related to differences in fruit composition (sweetness and sourness with sugars and acids), but texture traits are more difficult to relate to instrumental measurements;
- Consumers particularly liked hybrids between old and modern lines with intermediate firmness. The preference for hybrids between large and cherry tomatoes confirmed the major role of sweetness and acidity in preference, which appeared more important than texture traits. The results also showed the importance of texture in consumer preference; thus if a good flavour is obtained, a good texture is the second criteria needed, at least in large-fruit hybrids.
- Correlations between sensory profiles and fruit composition allowed identification of the major components to be selected.
- Most of the physico-chemical traits, flavour attributes and firm texture showed a simple additive inheritance, in contrast to the aroma and other texture traits.

Several mutations affecting fruit ripening and shelf life are described. The most widely used in tomato breeding is *rin* (ripening inhibitor), which, in the heterozygous state, enables fruits to be kept for a few weeks (Davies and Hobson, 1981). Long shelf life cultivars have invaded the tomato market, but in the 1990's, their quality, particularly their colour and flavour, had been criticized by consumers (Jones, 1986; McGlasson et al., 1987). In the previous experiment, Causse et al. (2003) had produced and compared seven pairs of nearly isogenic hybrids, with or without the *rin* mutation at the heterozygous level. The presence of the *rin* mutation reduced consumer preference. Differences were detected by sensory profiles, *rin* hybrids having fruits on average 17% less sweet, with a lower tomato aroma, a higher 'strange' aroma and more mealy fruits. Instrumental firmness and sugar content were no different. These results confirmed the negative influence of the *rin* mutation on consumer preference, but also indicated that when transferred into a hybrid with high flavour, the negative influence of the mutation is reduced. Selection could thus be carried out to obtain much sweeter and perfumed lines combined with shelf life in *rin* hybrids.

3. Genes and Quantitative Trait Loci (QTLs) affecting flavour

Many mutations involved in fruit development and composition have been discovered and used for fruit quality breeding. Table 1 lists the major mutations identified, which directly or indirectly impact fruit quality. They may induce variation in fruit colour or aspect. Some mutations impacting plant architecture, like *sp*, which controls the determinate/indeterminate growth, are also known to

impact fruit quality. Today, several populations of EMS mutants have been produced in a few genetic backgrounds (Okabe et al., 2011). They enlarge the range of variations available and allow the rapid discovery of the responsible genes (Austin et al., 2011).

Table 1 Cloned genes with a phenotype related to fruit quality, plant, leaf or truss architecture, location on the tomato genome assembly

ITAG gene model	Gene Symbol	locus_name	Chromosome	Start position	Phenotypic descriptors	References
Solyc01g079620	*y*	colourless epidermis*	1	71 255 600	pink epidermis	Ballester et al. (2010)
Solyc10g081470	*L-2*	Lutescent-2	10	61858478	altered chloroplast development and delayed ripening	Barry et al. (2012)
Solyc08g080090	*Gr*	green flesh	8	60 582 066	green fruit flesh	Barry et al. (2008)
Solyc06g074910	*C*	potato leaf	6	42 804 036	simple leaves	Busch et al. (2011)
Solyc03g031860	*r*	Phytoene synthase 1	3	8 606 749	yellow fruit	Fray and Grieson (1993)
Solyc04g082520	*cwp1*	cuticular water permeability 1	4	63 765 366	microfissure/dehydration of fruits	Hovav et al. (2007)
Solyc10g081650	*t*	carotenoid isomerase	10	62 006 972	orange fruit flesh	Isaacson et al. (2002)
Solyc02g077390	*s*	compound inflorescence	2	36 913 957	Inflorescence branching	Lippman et al. (2008)
Solyc02g077920	*Cnr*	Colourless non-ripening	2	37 323 107	Inhibition of ripening	Manning et al. (2006)
Solyc11g010570	*j*	jointless	11	3 640 857	no pedicel abscission zone	Mao et al. (2000)
Solyc03g118160	*fa*	falsiflora	3	61 162 449	leafy inflorescence	Molinero-Rosales et al. (1999)
Solyc03g063100	*sft*	single flower truss	3	30 564 833	single flower truss	Molinero-Rosales et al. (2004)

Appendix 2

Solyc01g056340	*hp-2*	de-etiolated 1	1	46 495 644	high pigment	Mustilli et al. (1999)
Solyc06g074350	*sp*	self-pruning	6	42 361 623	determinate plant habit	Pnueli et al. (1998)
Solyc10g008160	*u*	*uniform ripening*	10	2 293 088	increased chlorophyl content	Powell et al. (2012)
Solyc12g008980	*Del*	Delta	12	2 285 372	orange fruit	Ronen et al. (1999)
Solyc06g074240	*B*	Beta-carotene	6	42 288 127	increased fruit beta-carotene	Ronen et al. (2000)
Solyc03g083910	*sucr*	sucrose accumulator	3	47 401 871	Accumulates predominantly sucrose in mature fruit, rather than glucose and fructose	Sato et al. (1993)
Solyc07g066250	*ls*	lateral suppresser	7	64 958 148	Few or no axillary branches; corolla suppressed; partially male sterile	Schumacher et al. (1999)
Solyc02g090890	*hp-3*	zeaxanthin epoxidase	2	46 947 557	high pigment in fruits	Thompson et al. (2000)
Solyc05g012020	*rin*	ripening inhibitor	5	5 217 073	never ripening	Vrebalov et al. (2002)
Solyc05g012020	*mc*	macrocalyx	5	5 217 073	large sepals	Vrebalov et al. (2002)
Solyc05g053410	*phyB2*	apophytochrome B2	5	62 648 223	red light reception	Weller et al. (2001)
Solyc10g044670	*phyA*	apophytochrome A	10	22 854 459	far red light insensitive	Weller et al. (2001)
Solyc09g075440	*Nr*	Never ripe	9	62 631 866	not ripening	Wilkinson et al. (1995)
Solyc04g076850	*e*	Entire leaf	4	59 354 677	reduced leaf complexity	Zhang et al. (2007)

Most tomato fruit quality traits are quantitatively inherited. Tomato was among the first crop for which molecular markers were used to dissect the genetic basis of quantitative traits into QTL (Quantitative Trait Loci, Tanksley, 1993). Since then, many QTL controlling yield and fruit quality–related traits have been mapped (Paterson et al., 1988, 1990, 1991; Azanza et al., 1994; Goldman et al., 1995; Grandillo and Tanksley, 1996; Tanksley et al., 1996; Fulton et al., 1997, 2000, 2002; Bernacchi et al., 1998a,b; Chen et al., 1999; Doganlar et al., 2002; Frary et al., 2004; Eshed and Zamir, 1995; see Labate et al. (2007) for review). Due to the very low polymorphism revealed at the within species level, these studies were performed on interspecific progenies derived from crosses between wild tomato species and tomato inbreds. In most of the studies a few QTL explained a large fraction (20–50%) of the phenotypic variation, acting in concert with minor QTL that could not be detected. Most of the QTL act in an additive manner, but dominant and overdominant QTL have been detected (Paterson et al., 1988, 1991; de Vicente et al., 1993; Semel et al., 2006). Epistasis (interaction among QTL) was rarely detected unless a specific experimental design is used (Eshed and Zamir, 1996; Causse et al., 2007).

3.1. QTL for fruit size and shape

Grandillo et al. (1999) summarized the results of QTL mapping for fruit weight obtained in 17 studies. Six QTL explained more than 20% of the phenotypic variation. A common set of 28 QTL could be identified that frequently segregated in at least two populations. Nevertheless, only QTL cloning and complementation permits determination of whether each consensus QTL location corresponds to a single gene. Nowadays, only two fruit weight QTLs have been cloned by a map-based cloning approach. The first fruit size QTL to be cloned, *fw2.2* (Frary et al., 2000), controls up to 30% of the fruit size variation. It corresponds to an unknown function gene, ORFX, which acts on cell number in carpels before anthesis where it is differentially expressed between large and small fruits. However, its precise function is still unclear. The wild-type allele of ORFX negatively regulates cell division. The second QTL cloned for fruit weight (*fw3.2*), corresponds to a cytochrome P450 (Chakrabarti et al., 2013).

Locule number is another major component of fruit size and shape. Several QTL have been mapped (Lippman and Tanksley, 2001; van der Knaap and Tanksley, 2003; Barrero and Tanksley, 2004) for this trait. The two major QTL correspond to the mutations fasciated on chromosome 11 and *lc* on chromosome 2, with a strong epistatic interaction between these two genes (Lippman and Tanksley, 2001). Both mutations have been identified using a map-based cloning approach. The *lc* mutation is located near *Wuschel*, a gene that is responsible for stem cell fate in apical meristems, but 1500 bp upstream (Muños et al., 2011). Compared to *lc*, the fasciated mutation has a strong effect on locule

number, increasing the trait from 3 to more than 6 locules. The QTL is located close to a Yabby-like transcription factor (Cong et al., 2008). Analysis of molecular diversity of the locus finally showed that the phenotype resulted from a large invertion (several kilobases) between the YABBY gene and the *CLV3* gene (Xu et al., 2015). Furthermore, Xu et al. (2015) underlined the role of Clavata pathway in interaction with arabinosyltransferase genes in meristem size and subsequently in fruit size.

For fruit shape, Grandillo et al. (1999) identified a set of 11 QTL from the six studies where the fruit length:diameter ratio was segregating. Three major QTL were identified, *ovate* on chromosome 2, *sun* on chromosome 7 and *fs8.1* on chromosome 8 (van der Knaap et al., 2002). The gene *ovate* encodes a predicted 40.7 kDa protein with an unknown function Liu et al. (2002). A mutation in the second exon of the ORF leads to a premature stop codon in the protein sequence. Plants containing this truncated protein exhibit the ovate phenotype. The gene is expressed at the early developmental stages in flowers and fruits. Another mutation responsible for fruit length, *sun*, has been cloned (Xiao et al., 2008). The gene responsible for *sun* encodes an IQD protein. IQD proteins are found in plants and contain an IQ67 motif, which corresponds to a 67 amino acid motif. The function of this protein family is unknown, except for AtIQD1, which plays a role in the regulation of cytochrome P450 genes. The *sun* locus results from a retrotransposon duplication event. Functionally, the *sun* phenotype is due to a difference in the IQD gene expression. In wild-type plants, the gene is less expressed. The differential expression pattern could result from the new genomic context of the gene after duplication.

Rodriguez et al. (2011) showed that the combination of *lc*, *fas*, *sun* and *ovate* allows the classification of most shapes of the tomato fruit. Nevertheless, some QTLs modifying the effect of these genes remain to be detected. Fruit shape and size phenotypes are well described in thousands of natural accessions. The challenge is now to identify the molecular nature of QTLs with a weaker effect. Fruit shape and size are directly linked to developmental processes. To understand heterochrony, it is essential to characterize the natural diversity at the cellular level. For this purpose, new high-throughput tools have to be developed. Combining histological and molecular genetic regulation studies will help to clarify the precise mechanisms leading from stem cells to developed tomato fruits (Xu et al., 2015). Cell division and cell growth are two important mechanisms in fruit size; the two phases are distinct during fruit development. Consequently, genes involved in flower meristem development can be used as candidate genes for fruit size or shape (Barrero et al., 2006; Bauchet et al., 2014).

3.2. QTL for sugar and acid content

The review of Labate et al. (2007) summarizes the chromosome regions carrying QTL for sugar content or related traits (Soluble Solid Content-SSC, fructose, glucose or sucrose content) on the basis of 14 populations involving 8 different species (Paterson et al., 1988, 1990, 1991; Goldman et al., 1995; Azanza et al., 1994; Bernacchi et al., 1998a; Fulton et al., 1997, 2000, 2002a; Tanksley et al., 1996; Doganlar et al., 2002; Grandillo and Tanksley, 1996; Chen et al., 1999; Eshed and Zamir, 1995; Causse et al., 2004; Frary et al., 2004; Saliba-Colombani et al., 2001). A total of 95 QTL were detected in 56 chromosomal regions. For the majority of QTL, the wild species alleles increased the sugar content. In 28 regions, QTL were detected in more than one population, and may possibly correspond to the same QTL. The large number of regions involved suggests that many mechanisms are responsible in increasing fruit sugar content. The same results were obtained for acid content (Fulton et al., 2002a; Causse et al., 2002, 2004), with only a few regions common to acid and sugar content. In contrast, frequent colocations between QTL for sugar content and fruit weight (Grandillo et al., 1999) with opposite allelic effects were detected, suggesting pleiotropic effects of some common QTL. Few studies have reported QTLs for SSC with no apparent effect on fruit size (Yousef and Juvik, 2001; Fridman et al., 2004) and such antagonism may be responsible for the difficulty in simultaneously increasing fruit size and sugar content (Prudent et al., 2009). Part of this relationship is due to a dilution effect, but another part may be due to gene linkage as shown by fine mapping results (Lecomte et al., 2004).

The first QTL controlling SSC variation has been identified in a series of introgression lines derived from *S. pennellii* in an *S*. *lycopersicum* background (Eshed and Zamir, 1995). The QTL has been delimited to a region encompassing *Lin5* (Fridman et al., 2000), a gene encoding an apoplastic invertase expressed exclusively in fruits and flowers (Godt and Roitsch, 1997; Fridman and Zamir, 2003). Fridman et al. (2004) revealed that the wild species allele of *Lin5* was more efficient than the cultivated allele, due to a single nucleotide substitution that coded for an amino acid residue close to the fructosyl-binding site of the enzyme. *In planta*, proof of the importance of *Lin5* in the control of the total soluble solids content in tomato has been confirmed by RNAi approach (Zanor et al., 2009a). The sucrose accumulation in *S. chmielewskii* and *S. habrochaites* fruits is associated with low-level acid invertase activity (Yelle et al., 1988, 1991; Chetelat et al., 1993).

Starch accumulates at the early stages of tomato fruit development, contributing approximately 20% of the dry weight of the fruit tissue at peak concentration, prior to the mature green stage. This starch is completely degraded in the ripe fruit, serving as a reservoir contributing to the soluble solids content (Dinar and Stevens, 1981; Ho, 1996). ADP-Glc pyrophosphorylase (AGPase) catalyses the

synthesis of ADP-Glc, and is considered the first committed step in starch synthesis. Tomato plants (*S. lycopersicum*) harbouring the allele for the AGPase large subunit (AgpL1) derived from the wild species *S. habrochaites* (AgpL1 (H)) are characterized by higher AGPase activity and increased starch content in the immature fruit, as well as higher soluble solids in the mature fruit following the breakdown of the transient starch, as compared to fruits from plants harbouring the cultivated tomato allele (AgpL1(E), Schaffer et al., 2000). The increased activity of the *AgpL1H* in tomato fruit is due to an extended period of *AgpL1H* gene expression and subsequent stability of the S1–L1 heterotetramer (Petreikov et al., 2006). Similarly, the small subunit of ADP-glucose pyrophosphorylase (AGPase SS on chromosome 7) colocalized with QTLs for reducing sugars and fructose content (Causse et al., 2004).

3.3. QTL for volatile compounds

Volatiles are derived from the degradation of amino acids, fatty acids, carotenoids or phenolic compounds. A large range of variations for individual volatiles have been shown in panels of accessions (Klee and Tieman, 2013). QTL for volatile compounds have been mapped in three populations. Saliba-Colombani et al. (2001) detected QTL for 12 volatile compounds among 18 that were quantified in the progeny of a cross involving a cherry tomato. Tieman et al. (2006) identified QTL for 23 volatiles in the population of introgression lines derived from *S. pennellii*. Twenty-five QTL were identified. Although ten volatiles were analysed in both studies, only three QTL were detected in the same regions, for phenylacetaldehyde on chromosome 8 (confirming the effect of the QTL *Malodorous*, named by Tadmor et al., 2002), on chromosome 9 for 2-methylbutanal and on chromosome 12 for pentanal. The content in some volatile compounds appeared strongly variable over years or environments (Tieman et al., 2006). This could partly explain the small number of QTL common to the two studies. In both studies, QTL for several volatiles were frequently in clusters. In a few cases these clusters corresponded to volatiles derived from the same metabolic pathway (related to fatty acid, carotenoid or amino acid degradation), suggesting the action of a gene within a single pathway. More frequently, colocalizations of QTL for volatiles derived from various metabolic pathways were shown, suggesting the presence of regulatory gene acting on several pathways. In *S. habrochaites* introgression lines, 30 QTL affecting the emission of one or more volatiles were mapped (Mathieu et al., 2009).

A few genes responsible for volatile accumulation have been identified (Table 2). The *ADH* gene coding for an alcohol dehydrogenase is involved in the ratio of hexanal to hexanol in the fruit (Speirs et al., 1998). *TomloxC*, a gene coding for a fruit-specific lipoxygenase has been shown to be related to the generation of volatile C6 aldehyde and alcohol compounds including hexanal, hexenal and

hexenol (Chen et al., 2004). Two genes *LeAADC1* and *LEAADC2* are responsible for the decarboxylation of phenylalanine and subsequent synthesis of phenylethanol and related compounds (Tieman et al., 2006). The gene coding for the carotenoid cleavage dioxygenase 1 enzyme (*CCD1*) is involved in the synthesis of several aroma volatiles derived from carotenoid cleavage (Vogel et al., 2008). Tieman et al. (2007) showed that phenylacetaldehyde reductases (PAR) catalyse the last step in the synthesis of the aroma volatile 2-phenylethanol. A salicylic acid methyl transferase has been shown to be involved in the synthesis of methyl salicylate (Tieman et al., 2010). Tikunov et al. (2013) identified a mutation in a glycosyltransferase which is responsible for the release of smoky aroma related to phenylpropanoid compounds.

Table 2 Genes associated with volatile production in tomato

Gene	Associated volatile	Identification method	Reference
ADH	Hexanal:heanol ratio	BP	Speirs et al. (1998)
AADC	Phenylacetaldehyde, 2-phenylethanol, 1-nitro-2-phenethane, 2-phenylacetonitrile	BP/QTL	Tieman et al. (2006)
PAR	2-Phenylethanol	BP	Tieman et al. (2007)
LoxC	*Z*-3-Hexenal, *Z*-3-hexenol, hexanal, hexanol	CG	Chen et al. (2004)
SAMT	Methylsalicylate	BP	Tieman et al. (2010)
CTOMT	2-Methoxyphenol	BP	Mageroy et al. (2012)
CXE1	Multiple alcohols	QTL	Goulet et al. (2012)
CCD1	Multiple apocarotenoids	CG	Simkin et al. (2004)
GT1	Smoky aroma (phenylpropanoids)	QTL	Tikunov et al. (2013)

Abbreviations: AADC, aromatic amino acid decarboxylase; PAR, phenylacetaldehyde reductase; LoxC, 13-lipoxygenase; SAMT, salicylic acid methyltransferase; CTOMT, catechol-*O*-methyltransferase; CCD1, carotenoid cleavage dioxygenase; CXE1, carboxylesterase. BP: Biochemical pathway; CG: candidate gene; QTL: positional cloning; Adapted from Klee et al. (2013)

4. Tomato texture

Fruit texture is a complex breeding objective, as it involves the fruit firmness and shelf life, but also refers to a wider range of sensory attributes such as crispiness, juiciness, meltiness or mealiness. Texture is dependent on the overall fruit structure and spatial organization, the cellular morphology of main tissues, the cell turgor in addition to the biochemical and mechanical properties of the cell walls (Shackel et al., 1991; Harker et al., 1997; Chaib et al., 2007). In fleshy fruits, texture does not only influence the purchasing power of the consumer and consumer acceptance, but it also has a significant impact on overall organoleptic quality, shelf life, and transportability (Seymour et al., 2002) and it strongly interferes with the perception of flavour and aroma (Causse et al., 2003, 2011). After harvest, texture evolves rapidly, while membrane and cell wall breakdown occurs in relation to turgor loss and to enzyme-orchestrated cell wall loosening. Internal hormonal stimulation, as well as environmental factors such as light, temperature, water and nutrient supply, regulates ripening. Fruit texture is, thus, essentially, an unstable characteristic closely related to shelf life (Seymour et al., 2013).

Fruit firmness has been studied in several quantitative genetic studies. Labate et al. (2007) present a summary of QTL controlling fruit firmness in nine populations (Bernacchi et al., 1998a; Causse et al., 2002; Doganlar et al., 2002; Frary et al., 2003, 2004; Fulton et al., 1997, 2000; Tanksley et al., 1996). Forty-six QTL controlling firmness were mapped using seven different populations. More than half of the QTL were grouped in clusters of three to four QTL. These clusters were localized on chromosomes 1, 2, 4, 5, 9, 10 and 11. Chapman et al. (2012) dissected a fruit firmness QTL on chromosome 2 and revealed a complex locus with epistatic interactions.

Our current understanding of ripening mechanisms and the molecular basis of fruit texture in fleshy fruit mainly relies on transgenic or mutant plant analysis (CF CHAPTER X; Giovannoni, 2007; Seymour et al., 2002; Vicente *et al.,* 2007).

The pleiotropic *rin* (ripening inhibitor) recessive mutation blocks the ripening process. Mutant fruits fail to produce ethylene and are unable to ripen under ethylene treatment, although they are responsive to ethylene. Breeders have extensively used the *rin* mutation and hybrids (*rin/Rin*) form the basis for most present-day production of slow ripening, long shelf life, fresh-market tomatoes. The gene underlying the mutant was cloned; it encodes a partially deleted MADS-box protein of the SEPALATTA clade (Vrebalov *et al.,* 2002; Ito *et al.,* 2008; Hileman *et al.,* 2006). Another mutation, *Cnr* (colourless non-ripening) corresponds to an epigenetic mutation in a member of the same gene family (Manning *et al.,* 2006).

The decrease in tomato firmness coincides with the dissolution of the cell wall middle lamella, resulting in lower intercellular adhesion, depolymerization and solubilization of pectic and hemicellulosic cell wall polysaccharides (Brummell and Harpster, 2001; Rose *et al.,* 2004). Although many genes have been identified, their role in the natural variation of fruit texture has been rarely demonstrated. Polygalacturonase and pectin methylesterase were long considered as major enzymes for pectin depolymerization and de-esterification, but antisense mRNA-mediated suppression had only a minor effect on cell wall loosening and failed to reduce fruit softening (Tieman and Handa, 2004; Brummell and Harpster, 2001). Similarly, weak effects on fruit texture were obtained with several other ripening-related cell wall–modifying enzymes (reviewed by Rose *et al.,* 2003). A multigene family of 7 members encodes β-galactosidases, whose activity increases during ripening. Tomato fruits with suppressed β-galactosidase expression soften more slowly during ripening (Smith *et al.,* 2002), demonstrating that pectic side chains contribute to fruit texture. The possible involvement of pectin lyase and acetylesterases in pectin breakdown is proposed by Vicente et al., (2007).

During ripening xyloglucan depolymerization occurs within the hemicellulose fraction without clear identification of the responsible proteins. Some experimental clues indicate that endoglucanases and endo transglucosylases might catalyse xyloglucan degradation, but this possibility has to be further explored (Vicente *et al.,* 2007; Saladie *et al.,* 2006).

Expansins are proteins that contribute to cell expansion by a non-enzymatic cell wall–loosening biomechanical process (Cosgrove, 1998). They are present and active in ripening tomato fruit (Rose *et al.,* 2000) where they participate in cell wall disassembly and may enhance cellulose degradation by cellulases. Their exact contribution to fruit softening is yet to be demonstrated.

The difficulty encountered in identifying one or a few key determinants of fruit softening is due to the complexity of the process probably involving many different cell wall actors in a fine orchestrated manner. Meli et al., (2009) suggested that N-glycoprotein-modifying enzymes such as α-mannosidase and β-D-acetylhexosaminidase may play a role in tomato ripening–associated fruit softening. RNAi downregulation of these two ethylene-regulated genes led to subsequent downregulation of many genes that are associated with tomato ripening and cell wall degradation. Moreover, identification of new cell wall–modifying enzymes in order to gain new insight into the biochemical processes underlying fruit ripening is the present-day challenge; proteomic studies represent a promising perspective in this area because of the high level of post-transcriptional regulation of cell wall proteins (Minic *et al.,* 2009).

Finally, tomato fruit texture and shelf life may rely on other physiological mechanisms unrelated to cell wall loosening (Matas et al., 2009). In particular, fruit water status and cuticle structure may be important factors related to shelf life. The Delayed Fruit Deterioration (DFD) tomato cultivar which is able to remain firm for several months, exhibits normal ripening and cell wall loosening but very low fruit transpiration, high cellular turgor and a different cutin composition as compared to a control cultivar (Saladie et al., 2007). Such results, therefore, emphasize the possibility of a disconnection between pericarp firmness and fruit shelf life. Moreover, fruit firmness at harvest may be disconnected from its ability to remain firm after a period of storage especially at cold temperature. Recently, Page et al., (2010) compared the behaviour of two isogenic lines for a firmness QTL at harvest and during cold storage and found that the line possessing the favourable allele for firmness had the lowest storage ability. The lack of ability to remain firm was correlated with the lower expression of genotype-specific protective proteins, among others, heat shock proteins. Ascorbic acid redox state has also been shown to be involved in fruit shelf life and tolerance to cold storage (Stevens et al., 2008).

Fruit texture and shelf life capacity, and potentially, the interaction with susceptibility to pathogens (Cantu et al., 2008) represent a highly challenging research area which is currently benefiting from several genomic approaches.

5. New approaches to tomato flavour diversity and genetic control

During the last decade, the advent of the high-throughput sequencing and genotyping technologies enabled the collection of data at the genome-wide scale for hundreds of thousands of single nucleotide polymorphisms (SNP) at a reasonable cost. This task was facilitated by the release of the reference genome of major crops, among which was the tomato genome in 2012 (TGC, 2012). Then, tools such as the SolCAP SNP genotyping array were derived in tomato (Hamilton et al., 2012; Sim et al., 2012). On the basis of 7720 SNP markers, this array was largely used for different purposes, including the investigation of the tomato worldwide germplasm nucleotide diversity (Blanca et al., 2012; 2015), the study of the linkage disequilibrium decay along chromosomes (Sim et al., 2012a) or the establishment of reference linkage maps (Sim et al., 2012b) to pave the way for the mapping of quantitative traits linked to agronomical traits. In parallel, the phenotyping of multiple traits related to fruit quality in multiple environments for large populations obtained from bi-parental or multi-parental crosses was achieved to decipher the genetic basis (i.e. broad sense of heritability, number of loci) of these traits and their interaction with the environmental conditions.

To overcome the main limitation of the QTL experimental design (essentially the lack of recombination), benefit was taken from the ancestral polymorphism found in natural population or

germplasm core collections to identify the underlying molecular determinants of agronomical traits. In 2006, a linear mixed model (MLM), based on the statistical model described in Henderson (1975), was proposed to test the statistical link between the genotypic and the phenotypic data in a collection of maize varieties (Yu et al., 2006), while taking into account the confounding effect of the pairwise genetic relatedness (the so-called K-matrix) between accessions and population structure (the Q-matrix). The genome-wide association (GWA) approach was adopted in many crops, including the tomato. Ranc et al., (2008) initiated the building of a reference core collection of 360 accessions on the basis of the genetic diversity revealed at 20 microsatellite markers. Despite being restricted to a single chromosome, a proof of concept was presented with the mapping of genotype–phenotype associations for a flavour trait (solid soluble content) in this core collection. Subsequent studies applied the approach at the genome-wide scale while detecting more and more loci. Using a set of 192 SNP markers genotyped in 188 accessions, Xu et al., (2012) identified 2, 16 and 17 loci associated to titrable acidity, soluble solids and sugar contents of the fruit, while the phenotypic heritabilities were estimated to 0.75, 0.73 and 0.63, respectively. A similar study, in terms of experimental design, with 174 accessions (both *S. lycopersicum* and *S. l.* var *cerasiforme*) and 182 SSR, identified 17 and 22 associated loci for fruit weight and ascorbic acid content, supporting the polygenic genetic architecture of these traits (Zhang et al., 2016). Favourable allelic combination between loci associated to fruit quality, such as pH, titrable acidity or SSC, were defined from a classical (K+Q) linear mixed model in the core collection of 96 fresh market and processing tomatoes (Ruggieri et al., 2014).

To deepen the framework of GWA, extensive work towards developing SNP was also achieved through resequencing projects, especially for larger collections of tomato accessions. For example, Yamamoto et al., (2016) identified over 50,000 markers that were implemented in a GWA approach for traits of agronomical interest (i.e. plant height, fruit size), including traits related to flavour such as soluble solid content. By using a multi-locus mixed model (see Segura et al., 2012), Sauvage et al., (2014) provided an extended list of loci notably associated to important metabolic compounds for flavour such as fructose, SSC and malic and citric acids. In this experimental design, the broad sense of heritability was estimated to 0.56, 0.6, 0.64 and 0.42 respectively, demonstrating that not all the genetic variabilities of these traits have been captured by the molecular markers. In addition, the genotype by environment interaction may certainly bias these heritability estimates as flavour traits are under the influence of growing conditions.

In recent years, the high-throughput genomic produced large amounts of SNP genotyping data that are now overcome by large resequencing projects not only in the cultivated tomato (see Aflitos et al., 2014 and Lin et al., 2014) but also in its wild relatives (i.e. *Solanum pennellii*, see Bolger et al., 2014),

from which fragments carrying genes of interest were introgressed. The statistical approaches were refined, notably to take into account confounding effects such as genetic relatedness or population stratification to improve the power of the GWA approach (see Tucker et al., 2014). One of the main conclusions of these GWA studies is the polygenic architecture of the traits related to flavour in tomato that may explain why breeding for improving such traits remains a challenging endeavour. Another conclusion is the intraspecific genetic variability that might be still exploited for enhancing tomato fruit quality, especially in the *S.l.* var *cerasiforme* group. While being genetically diverse, this group has the advantages to be in admixture between the closest wild relative tomato (*S. pimpinellifolium*) and the big-fruited cultivated tomato (*S. lycopersicum*) and having a large phenotypic diversity for many quality components. Thus, focusing on this group might improve the power of GWA and remove the population structure confounding effect, especially when the quality traits are correlated to the population structure (i.e. fruit weight or sugar content). In addition to cheery-type tomato, landraces remain underexploited and might also provide a viable indirect selection tool in future practical breeding programmes (Sacco et al., 2015).

New types of populations involving several parental lines like MAGIC (Multi-allelic Genetic Intercross) have also been found useful to map QTL into small confidence intervals. Combined with the resequencing of the parental lines, a direct access to putative polymorphisms under the QTL could be proposed (Pascual et al., 2015). The comparison of biparental, MAGIC and GWAS panels confirmed the complementarity of the three kinds of populations (Pascual et al., 2016).

On the basis of these sets of associated molecular markers, two different steps forward have to be achieved. The first one relies on the dissection of the molecular mechanisms underlying flavour traits. Here again, to reach this objective, the broad panel of biotechnologies (the so-called 'new breeding technologies') offered is plethora: RNAi, TALEN or CRISPR\Cas9 can be implemented to validate functionally the most promising loci identified by the GWA. The second step forward aims at implementing the molecular information gathered into a marker-assisted selection (MAS) breeding scheme and sustain the varietal innovation. For traits governed by a low to moderate number of genes (<10), this approach is feasible. However, for complex and polygenic traits, such as the ones determining flavour in tomato, the MAS approach is limiting in this case.

6. From marker-assisted selection to genomic selection for flavour breeding

Candidate loci affecting traits of agronomical interest are plentiful, but very few markers have been exploited so far, compared to the numerous linkage or association studies published (Jonas and de Koning, 2013). One of the main reasons is the variation of the marker effects between environments and populations leading to non-consistent results (Bernardo et al., 2008). MAS is still successful in

assisting breeding but remains limited to a moderate number of markers. Thus, MAS is now being extended by the latest selective breeding approach, the genomic selection, a multiple markers and genome-wide scale approach.

In the early 2000s, marker-assisted selection for quality traits was initiated for five QTLs controlling fruit quality traits in tomato. This investigation revealed epistatic interactions between the QTLs and the genetic background, limiting the breeding efficiency according to the recipient parent (Lecomte et al., 2004). While being efficient in improving quality traits, the introgression of large chromosomic regions favoured the linkage drag of undesired alleles. This study demonstrated how challenging is the MAS for complex traits and how many generations of crosses will be required to clean the genetic background from the linkage drag, especially for multiple traits.

Facing this limitation, plant breeders are now evaluating the potential of the latest selective breeding strategy, the genomic selection (GS). The funder paper published by Meuwissen et al., (2001) describes this approach that is aimed at estimating the breeding value of an individual from the genome-wide genetic information. Precisely, an effect is attributed to all the markers identified in the genome of individuals, from which genotype and phenotype are known. These individuals compose the so-called training population (TP). Then, the sum of all these effects is the genomic estimate of the breeding value (GEBV) of each individual, for the considered trait: the larger, the more interesting is the individual to reproduce. Within the cross-validation step of the GS process, statistical models are tested for their accuracy to predict the phenotype of the individuals of the TP, on the only basis of the genotypic information. One or several may be accurate when the coefficient of correlation between the observed phenotype and the predicted one is high to very high (0.5–0.9). Then, the accurate prediction model is implemented in a real breeding scheme. GS has been largely investigated and implemented in animals, especially for dairy cattle, for which the evaluation of the real genetic gain has been published (Patri et al., 2011), supporting the success of the approach. However, transferring the methods and knowledge obtained from animal to plant breeding is not trivial as the animal model does not take into account biological parameters like genotype by environment interactions.

In plant breeding, the GS potential is now being tested. More advanced experiments have been conducted in the annual crop wheat and maize (see Poland et al., 2012 and Bernardo and Yu, 2007). The potential of GS is also being tested in perennial species such as Spruce (Beaulieu et al., 2014), wine grape (Fodor et al., 2014) or apple (Kumar et al., 2012; Muranty et al., 2015). In tomato, the amount of data produced for the GWA experiments was further used to test the potential of GS to improve quality traits. The published studies rely on either the combination of GWA, GS cross

validation and recurrent GS simulation (Yamamoto et al., 2016) and on the estimation of the effects of various parameters (i.e. marker density, size and composition of the TP) on the accuracy of the statistical models (Duangjit et al., 2016). In both studies, the targeted traits are mainly related to quality traits, especially flavour with sugar content or acidity. The accuracy of the GEBV in the first study was evaluated from a set of ~16k SNP in a TP of 96 accessions, while for the second study, the accuracy was evaluated from ~7k SNP in a TP of 122 accessions. For the common phenotypes to both studies, the estimated accuracy was similar with 0.807 and 0.714 for SCC, for example. This demonstrated the reliability of the approach. However, these results have to be carefully interpreted, as in both studies, phenotype values were estimated across several years and growing environment, controlling for the G×E interactions. Briefly, both studies demonstrated that reliable phenotype predictions could be obtained in tomato for highly and moderately heritable traits, stimulating interest to implement this approach in a large-scale breeding scheme. However, GS is still in its infancy in tomato and other crops. Priority has first been given to cross validation, but GS has to take the next step by delivering more of its promises.

7. Interactions genotype by environment: a tool for breeding good tomato

Tomatoes are produced year-round under contrasting environmental conditions, triggering seasonal variations in their sensory quality. Over the tomato growing cycle, different factors such as light intensity, air and soil temperatures, plant fruit load, plant mineral nutrition or water availability influence the final fruit quality (reviewed in Davies and Hobson, 1981 and Poiroux-Gonord *et al.,* 2010). Variations in temperature and irradiance during ripening affect carotene, ascorbic acid and phenolic compound content in the fruit, although acid and sugar content are not modified considerably by these two factors (Venter *et al.,* 1977; Rosales *et al.,* 2006 and Gautier *et al.,* 2008). Changes in plant fruit load through trust pruning modify fruit dry matter content and final fruit fresh weight by disrupting the carbon flux entering to the fruit (Bertin *et al.,* 2000; Guichard *et al.,* 2005). Water limitation and irrigation with saline water may impact positively tomato fruit quality, mainly through an increase in sugar content in fruit (either by concentration or by accumulation effect) and through contrasted effects on the secondary metabolite contents (Mitchell *et al.,* 1991; De Pascale *et al.,* 2001; Nuruddin *et al.,* 2003; Johnstone *et al.,* 2005; Gautier *et al.,* 2009; Ripoll 2016a; Ripoll 2016b). The effects reported on fruit composition are associated or not to large yield loss depending upon the intensity and duration of the treatment and the development stage of the plant (see Ripoll *et al.,* 2014 for review). They result from modifications of the water and carbon fluxes imported by the fruit during its growth (Guichard *et al.,* 2001; Albacete *et al.,* 2013; Osorio *et al.,* 2014).

Thus, the optimization of the growing practice, in particular water management, is considered in horticultural production as a tool to manage fruit quality while limiting yield losses, offering the

opportunity to address simultaneously environmental issues and consumer expectations of tastier fruits (Stikic *et al.,* 2003; Fereres *et al.,* 2006; Costa *et al.,* 2007). The genetic variability of tomato response to water limitations and other abiotic constraints and their combination still need to be deciphered to develop genotypes adapted to these practices (Poiroux-Gonord *et al.,* 2010; Ripoll *et al.,* 2014). Large phenotypic variation in response to a wide range of climate and nutrition conditions exists in the genus *Solannum* at both inter- and intra-species levels (reviewed in Labate, 2007). The TGRC (Tomato Genetics Resource Center, UC Davis) maintains wild and cultivated accessions with known or inferred tolerances to various abiotic stresses, including drought, flooding, high temperature, chilling injury, aluminium toxicity, salinity and/or alkalinity, providing useful starting material for breeding, genetic mapping and other uses.

Several authors attempted to measure genotype by environment (G x E) interactions on tomato fruit quality by repeating a same experiment in different locations or/and under several growing facilities (Auerswald *et al.,* 1999; Johansson *et al.,* 1999; Causse *et al.,* 2003) or by building experimental design to isolate the effect of particular environmental factors on a large number of genotypes (see Semel *et al.,* 2007; Albert *et al.,* 2016; Gur *et al.,* 2011 for water availability and Monforte *et al.,* 1996; Monforte *et al.,* 1997a, Monforte *et al.,* 1997b for salt stress). In the different experiments, the G x E interaction was significant for the fruit quality traits measured (including fruit fresh weight, secondary and primary metabolism contents and fruit firmness), but generally accounted for a low part of the total variation in comparison with the genotype main effect. Albert *et al.,* (2016) dissected further the genotype by watering regime interaction in an intraspecific *S. lycopersicum* recombinant inbred line population grown under two contrasting watering regimes in two locations. In their studies, the interaction resulted from genotype re-ranking across the watering regime rather than scale changes. Besides, they identified large genetic variation and genetic heritabilities under both watering regimes, encouraging the possibility of developing tomato genotypes with an improved fruit quality under deficit irrigation.

The emergence of high-throughput genomic tools and the availability of genome sequences facilitate the decomposition of the genotype by environment interactions into underlying QTLs and/or genes. For this purpose, the first approach consists in modelling the effect of QTLs across different environmental conditions. In tomato, a few QTL studies considering the interaction with environmental variables at the fruit level were reported. They mainly pertained to response to salt stress (Monforte *et al.,* 1996; Monforte *et al.,* 1997a; Monforte *et al.,* 1997b; Uozumi *et al.,* 2012; Asins *et al.,* 2015) and drought stress (Gur *et al.,* 2011; Albert *et al.,* 2016). All these studies identified numerous loci with low to medium effect, suggesting a strongly polygenic architecture of tomato fruit response to environmental constraints. Gur *et al.,* (2011) described drought-responsive QTLs for

fruit fresh weight and sugar content mainly expressed by the shoot in a reciprocal-grafting experiment, whereas Monforte *et al.,* (1997b) identified QTLs with changing additive and epistatic effects according to the salinity level of the watering solution. Nevertheless, the authors mostly compared QTLs at different map positions and with different effects across experiments and conditions, which may be questionable as these comparisons depend upon the mapping significance threshold. Besides, the populations used were mainly introgression lines involving wild relative species (*Solanum habrochaites, Solanum pennellii* and *Solanum pimpinellifolium*) and the confidence intervals obtained remain large and difficult to transpose into the cultivated tomato. Statistical improvements allowing to explicitly take into consideration the effect of environmental variables in mapping models are available in the linkage (Van Eeuwijk *et al.,* 2010; El Soda *et al.,* 2014; Li *et al.,* 2015; Verbyla *et al.,* 2014) and association (Korte *et al.,* 2012; Saïdou *et al.,* 2014) framework and permit to analyse more complex population designs (RILs, GWA collections, multi-allelic MAGIC populations). These models offer the possibility to test properly the QTL by environment interactions and to identify QTLs whose effects are changing according to the environment. Recently, by applying such models in their *S. lycopersicum* RIL population grown under two contrasted watering regimes, Albert *et al.,* (2016) mapped a total of 56 QTLs for plant and fruit quality traits, among which 20% presented effects changing direction or intensity according to the irrigation treatment. This proportion of interactive QTL is roughly identical to the results obtained in other crop species (Tinker *et al.,* (1996) in barley; Sari-Gorla *et al.,* (1997) and Melchinger *et al.,* (1998) in maize) and should be considered for crop improvement.

The second strategy to decipher the G x E interaction into QTLs consists in constructing composite variables measuring phenotypic plasticity to deal with univariate QTL mapping models. These variables can be simple ratio or difference between the values of a trait measured under two contrasted conditions or parameters derived from more or less complex crop models integrating multiple environments. The ecophysiological models constitute adequate tools for analysing the genotype by environment interactions since they integrate environmental and genetic effects on individual processes and are able to predict interactions among processes during fruit development (Bertin *et al.,* 2010). The Virtual Fruit Model was developed by Fischman and Génard (1998) to describe both the water and dry matter accumulation rates in fleshy fruits. This model was powerful in assessing the impacts of fruit load in tomato of fruit load (Prudent et al., 2011) or of water deficit on fruit growth in different species, among which are peach (Quilot et al., 2005), mango (Lechaudel *et al.,* 2006), kiwifruit (Hall *et al.,* 2013) and tomato (Liu *et al.,* 2007). When plant traits are generally dependent on genotype, environment and cultural practices, model parameters are, ideally, independent of the environment and management and are amenable to QTL analysis within

univariate mapping models. Then, plants carrying different combinations of QTL might be simulated and tested under different environmental constraints, helping to select the best ideotypes as illustrated by Reymond et al., (2003) in maize and Quilot et al., (2005) in peach.

The major lock to decipher the genetic determinants of fruit quality and its response to environmental constraints remains in the ability to phenotype large number of accessions (to maintain a satisfying power in the genetic analysis) under contrasted environmental conditions. This limitation tends to be bypassed by the development of high-throughput phenotyping platforms, allowing to phenotype a large number of plants in finely controlled environments. Nevertheless, these platforms are generally difficult to adapt to grow tomato up to the production of mature fruits and may differ strongly from the conditions suffered by the plants in real production conditions. The second type of high-throughput phenotypic platforms will characterize the 'invisible' phenotypes, such as secondary metabolites, that are major determinants of flavour (see Tikunov et al., 2013 for an example for aroma in tomato). Thus, a second wave of large phenotyping data may submerge the discipline of plant breeding. This will require new statistical models that handle these large datasets and the mixing with the other 'omics' data (i.e. transcriptomic profiling, RNAseq) to push the plant breeding into the era of system biology. Tomato has proven to be one of the best models for the integration of multiple levels of information to sustain breeding (Pascual et al., 2013). Finally, changes in gene expression were shown to be the key process of tomato response to environmental constraints (Chen and Tabaeizadeh 1992; Thompson *et al.,* 1995; Zhou *et al.,* 2007). In the near future, measures of gene expression at whole genome scale (using RNAseq technologies or microfluidic qPCR or microarrays) in a large set of accessions under contrasted conditions associated to eQTL (expression QTL) mapping may help to decipher the molecular basis of tomato response to environmental constraints.

8. Future trends

At the genetic level, many studies focused on the genetic variation and the genes regulating the fruit quality components, but only a few QTL were finely characterized. Furthermore, the integration of these results into breeding process is still incomplete.

Today, many new genomic and genetic resources are available. Several tilling mutant collections were developed and constitute novel sources of variation (Okabe et al., 2011). A high-quality tomato genome sequence is publicly available together with a large set of transcriptome data in several accessions and for different organs, stages and conditions. The high-throughput sequencing technology allows rapid mapping and cloning of new genes. Thus in the coming years, polymorphisms and genes involved in tomato flavour should be identified and used for breeding

better tomatoes. Combined with adapted growth conditions they should better answer consumer expectations.

9. Conclusion

Tomato fruit quality is a complex trait involving a number of components, including fruit aspect, flavour, aroma and texture. A large range of genetic diversities have been shown in tomato for fruit quality components. Although a few major mutations may have a huge effect on fruit quality (notably the rin mutation), most of the components have a quantitative inheritance. Several QTL mapping experiments have been performed, mostly on interspecific progeny. Many loci and QTL have thus been detected, revealing some QTL cluster regions. Tomato is a model plant for fruit development and composition, and knowledge about its physiology rapidly increases and several genes affecting fruit quality are discovered. New approaches such as genome-wide association studies or MAGIC populations using the genome information allow a higher precision of QTL location.

Environment and post-harvest conditions may also strongly affect fruit quality and interact with the genotype limiting the genetic progress. More research in this field is thus necessary to identify the processes affected by the environment and assay whether light stress can improve sensory quality.

Although many QTL studies have been performed, marker-assisted selection has been rarely set up and results are mitigated. Particularly the negative correlation between fruit size and sugar content has limited genetic progress. Today, new hopes arise from genomic selection, although the impact of such method still needs to be demonstrated.

10. Where to look for further information

- *Introductions to the subject for non-specialists*

Book chapters:

- Quilot-Turion B, Causse M (2014) Natural Diversity and Genetic Control of Fruit Sensory Quality. In: Pravendra N, Bouzayen M, Mattoo AK, Pech JC, editors. Fruit ripening Physiology, signalling and genomics: CABI. pp. 228–45.
- Causse, M. 2008 Genetic background of flavour: the case of the tomato. In Brückner, B., Wyllie, S.G. (Ed). Fruit and vegetable flavour. Recent advances and future prospects. CRC Press, Boca Raton (USA), 229–53.
- Labate JA, S Grandillo, T Fulton, S Muños, AL Caicedo, I Peralta, Y Ji, RT Chetelat, JW Scott, MJ Gonzalo, D Francis, W Yang, E van der Knaap, AM Baldo, B Smith-White, LA Mueller, JP Prince, NE Blanchard, DB Storey, MR Stevens, MD Robbins, J Fen Wang, BE Liedl, MA O'Connell, JR Stommel, K Aoki, Y Iijima, AJ Slade, SR Hurst, D Loeffler, MN Steine, D Vafeados, C McGuire, C

Freeman, A Amen, J Goodstal, D Facciotti, J Van Eck, M Causse (2007) 1 Tomato. In "Genome Mapping and Molecular Breeding in Plants", Volume 5, Vegetables, C. Kole (Ed.), Springer-Verlag Berlin Heidelberg, 11–135.

- Any seminal articles or books which have shaped the subject:
    - Causse M, R Damidaux, P Rousselle (2007) Traditional and enhanced breeding for fruit quality traits in tomato. In Genetic Improvement of Solanaceous Crops, Vol.2: Tomato. Eds: M.K.Razdan and A. K. Mattoo, Science Publishers, Enfield, USA, 637 pp.
    - Klee HJ, Tieman DM (2013) Genetic challenges of flavor improvement in tomato. Trends Genet 29:257–62.
    - Bartoshuk LM, Klee HJ. 2013. Better Fruits and Vegetables through Sensory Analysis. Current Biology 23: R374–R378.
    - Causse M., Saliba-Colombani V.,Lecomte L., Duffé P, Rousselle P., Buret M. (2002) Genetic analysis of fruit quality attributes in fresh market tomato. J Exp Bot 53/377: 2089–98.

- *Any key websites worth visiting to keep up to date with trends:*
    - Solgenomics

- *Any key journals or conferences:*
    - Solanaceae genome congress, Tomato Eucarpia meetings, Tomato roundtable meetings; Solanaceae workshop of the Plant and Animal Genomes meeting

- *Any major international research projects:*
    - Tomato Genome cooperative for sequencing the tomato genome (https://solgenomics.net/); FP6 EUSOL (https://www.eu-sol.wur.nl/), H2020 Traditom (http://traditom.eu/)

- *top five or more research centres that readers can investigate for possible collaboration as well as to keep up with research trends*:
    - USA (Jim Giovanonni, Harry Klee), France-INRA (Christophe Rothan, Mondher Bouzayen), Netherlands-WUR (Arnaud Bovy), Spain-CSIC-Valencia (Antonio Granell).

11. References

Aflitos, S., Schijlen, E., de Jong, H., et al. (2014). Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *The Plant Journal* 80(1): 136–48.

Albacete, A. A., Martínez-Andújar, C. and Pérez-Alfocea, F. (2014). Hormonal and metabolic regulation of source–sink relations under salinity and drought: From plant survival to crop yield stability. *Biotechnology Advances*, 32(1), 12–30.

Albert, E., Gricourt, J., Bertin, N., Bonnefoi, J., Pateyron, S., Tamby, J. P. and Causse, M. (2016). Genotype by watering regime interaction in cultivated tomato: lessons from linkage mapping and gene expression. *Theoretical and Applied Genetics* 129(2), 395–418.

Asins, M. J., Raga, V., Roca, D., Belver, A., and Carbonell, E. A. (2015). Genetic dissection of tomato rootstock effects on scion traits under moderate salinity. *Theoretical and Applied Genetics* 128(4), 667–79.

Auerswald H, Peters P, Bruckner B, et al. (1999). Sensory analysis and instrumental measurements of short-term stored tomatoes (*Lycopersicon esculentum* Mill.). *Postharvest Biology and Technology* 15, 323–34.

Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang J, Fung P, Gong Y, Wang PW, McCourt P and Guttman DS. (2011). Next-generation mapping of Arabidopsis genes. *The Plant Journal* 67, 715–25.

Azanza F, Young TE, Kim D, Tanksley SD and Juvik JA (1994). Characterization of the effects of introgressed segments of chromosome 7 and 10 from *Lycopersicon chmielewskii* on tomato soluble solids, pH and yield. *Theoretical and Applied Genetics* 87, 965–72.

Baldwin E, Scott J, Shewmaker C and Schuch W (2000). Flavor trivia and tomato aroma: biochemistry and possible mechanisms for control of important aroma components. *Hortscience* 35, 1013–22.

Ballester AR, Molthoff J, de Vos R., et al. (2010). Biochemical and molecular analysis of pink tomatoes: deregulated expression of the gene encoding transcription factor SlMYB12 leads to pink tomato fruit color. Plant Phys 152: 71–84

Barrero LS, Cong B, Wu F., et al. (2006). Developmental characterization of the fasciated locus and mapping of Arabidopsis candidate genes involved in the control of floral meristem size and carpel number in tomato. Genome 49, 991–1006.

Barrero LS, Tanksley SD (2004). Evaluating the genetic basis of multiple-locule fruit in a broad cross section of tomato cultivars. *Theoretical and Applied Genetics* 109:669–79

Barry CS, Aldridge GM, Herzog G., et al. (2012). Altered chloroplast development and delayed fruit ripening caused by mutations in a zinc metalloprotease at the lutescent2 locus of tomato. *Plant Physiology* . 159(3): 1086–98

Barry CS, McQuinn RP, Chung MY., et al. (2008). Amino acid substitutions in homologs of the STAY-GREEN protein are responsible for the green-flesh and chlorophyll retainer mutations of tomato and pepper. *Plant Physiology* 147(1), 179–87.

Bartoshuk LM and Klee HJ (2013). Better fruits and vegetables through sensory analysis. *Current Biology* 23, R374–8.

Bassel GW, Mullen RT and Bewley JD (2008). Procera is a putative DELLA mutant in tomato (*Solanum lycopersicum*): effects on the seed and vegetative plant. *Journal of Experimental Botany* 59(3), 585–93.

Bauchet G, Munos, S., Sauvage, C., Bonnet, J., Grivet, L and Causse, M. (2014). Genes involved in floral meristem in tomato exhibit drastically reduced genetic diversity and signature of selection. *BMC Plant Biology* 14, 279.

Beaulieu, J., Doerksen, T. K., MacKay, J., et al. (2014). Genomic selection accuracies within and between environments and small breeding groups in white spruce. *BMC Genomics* 15(1), 1–16.

Bellec-Gauche, et al. (2015). Case Study: multidimensional comparison of local and global fresh tomato supply chains. *GLAMUR Project Report*, p. 56.

Bernacchi D, Beck-Bunn T, Emmatty D, Eshed Y, Inai S, Lopez J, Petiard V, Sayama H, Uhlig J, Zamir D and Tanksley S (1998b). Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. *Theoretical and Applied Genetics* 97, 170–80.

Bernacchi D, Beck-Bunn T, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D and Tanksley S (1998a). Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theoretical and Applied Genetics* 97, 381–97.

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Science* 48(5), 1649–64.

Bernardo, R. and J. Yu (2007). Prospects for Genomewide Selection for Quantitative Traits in Maize All rights reserved. *Crop Science* 47(3), 1082–90.

Bertin, N., Guichard, S., Leonardi, C., Longuenesse, J. J., Langlois, D. and Navez, B. (2000). Seasonal evolution of the quality of fresh glasshouse tomatoes under Mediterranean conditions, as affected by air vapour pressure deficit and plant fruit load. *Annals of Botany* 85(6), 741–50.

Bertin, N., Martre, P., Génard, M., Quilot, B. and Salon, C. (2009). Under what circumstances can process-based simulation models link genotype to phenotype for complex traits? Case-study of fruit and grain quality traits. *Journal of Experimental Botany,* 377.

Blanca, J., Canizares, J., Cordero, L., et al. (2012). Variation Revealed by SNP Genotyping and Morphology Provides Insight into the Origin of the Tomato. *PLoS ONE* 7(10), e48198.

Blanca, J., Montero-Pau, J., Sauvage, C., et al. (2015). Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics* 16(1), 257.

Bolger, A., Scossa, F., Bolger, M. E., et al. (2014). The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nature Genetics* 46(9), 1034–8.

Brummell DA and Harpster MH. (2001). Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Molecular Biology* 47, 311–39.

Busch BL, Schmitz G, Rossmann S, Piron F, Ding J, Bendahmane A and Theres K (2011). Shoot branching and leaf dissection in tomato are regulated by homologous gene modules. *The Plant Cell* 23(10), 3595–609.

Cantu D, Vicente AR, Greve LC., et al. (2008). The intersection between cell wall disassembly, ripening, and fruit susceptibility to Botrytis cinerea. *Proceedings of the National Academy of Sciences of the USA* 105, 859–64.

Causse M, Buret M, Robini K and Verschave P (2003). Inheritance of nutritional and sensory quality traits in fresh market tomato and relation to consumer preferences. *Journal of Food Science* 68, 2342–50.

Causse M, Chaïb J, Lecomte L., et al. (2007). Both additivity and epistasis control the genetic variation for fruit quality traits in tomato. *Theoretical and Applied Genetics* 115, 429–42.

Causse M, Duffe P, Gomez MC, Buret M, Damidaux R, Zamir D, Gur A, Chevalier C, Lemaire-Chamley M and Rothan C (2004). A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. *Journal of Experimental Botany* 55, 1671–85.

Causse M, Friguet C, Coiret C., et al. (2010). Consumer preferences for fresh tomato at the European scale: a

common segmentation on taste and firmness. *Journal of Food Science* 75, S531–41.

Causse M, Stevens R, Ben Amor B., et al. (2011). Breeding for fruit quality. In Jenks M and Bebelli PJ (eds), Breeding for Fruit Quality, Wiley Online, pp. 279–305.

Chaïb J, Devaux, MF., Grotte, M., Robini, K., Causse, M., Lahaye, M. and Marty, I. (2007). Physiological relationships among physical, sensory, and morphological attributes of texture in tomato fruits. *Journal of Experimental Botany* 58, 1915–25.

Chakrabarti, M., Zhang, N., Sauvage, C., Munos, S., Blanca, J., Canizares, J., Diez, MJ., Schneider, R., Mazurek, M., McClead, J., Causse, M. and van der Knaap, E. (2013). A cytochrome P450 CYP78A regulates a domestication trait in tomato (*Solanum lycopersicum*). *Proceedings of the National Academy of Sciences of the USA* 110(42), 17125–30.

Chapman NH, Bonnet, J., Grivet, L., Lynn, J., Graham, N., Smith, R., Sun, G., Walley, P. G., Poole, M., Causse, M., King, G. J., Baxter, C. and Seymour, G. B. (2012). High-resolution mapping of a fruit firmness-related quantitative trait locus in tomato reveals epistatic interactions associated with a complex combinatorial locus. *Plant Physiology* 159, 1644–57.

Chen FQ, Foolad MR, Hyman J, St. Clair DA and Beelman RB (1999). Mapping of QTLs for lycopene and other fruit traits in a *Lycopersicon esculentum × L. pimpinellifolium* cross and comparison of QTLs across tomato species. *Molecular Breeding* 5, 283–99.

Chen GP, Hackett R, Walker D., et al. (2004). Identification of a specific isoform of tomato lipoxygenase (TomloxC) involved in the generation of fatty acid-derived flavor compounds. *Plant Physiology* 136, 2641–51.

Chen R.-D. and Tabaeizadeh, Z. (1992b). Expression and molecular cloning of drought-induced genes in the wild tomato *Lycopersicon chilense*. *Biochemistry and Cell Biology* 70, 199–206.

Chetelat RT, Klann E, DeVerna JW, Yelle S and Bennett AB (1993). Inheritance and genetic mapping of fruit sucrose accumulation in *Lycopersicon chmielewskii. The Plant Journal* 4, 643–50.

Cong B, Barrero LS and Tanksley SD (2008). Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication

Cosgrove DJ. (1998). Cell wall loosening by expansins. *Plant Physiology* 118, 333–9.

Costa, J. M., Ortuño, M. F., and Chaves, M. M. (2007). Deficit irrigation as a strategy to save water: physiology and potential application to horticulture. *Journal of Integrative Plant Biology* 49(10), 1421–34.

Davies JN and Hobson GE (1981). The constituents of tomato fruit - The influence of environment, nutrition and genotype. *Critical Review of Food Science and Nutrition* 15, 205–80.

de Vicente MC and Tanksley SD (1993). QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics* 134, 585–96.

Dinar M and Stevens MA (1981). The relationship between starch accumulation and soluble solids content of tomato fruits. *Journal of the American Society for Horticultural Science* 106, 415–18.

Doganlar S, Frary A, Ku H-K and Tanksley SD (2002). Mapping quantitative trait loci in inbred backcross lines of *Lycopersicon pimpinellifolium* (LA1589). *Genome* 45, 1189–202.

Dorais M, Papadopoulos AP and Gosselin A (2001). Greenhouse tomato fruit quality. *Horticulture Review* 26, 239–319.

Duangjit, J., Causse, M. and Sauvage, C. (2016). Efficiency of genomic selection for tomato fruit quality. *Molecular Breeding* 36(3), 1–16.

El-Soda, M., Malosetti, M., Zwaan, B. J., Koornneef, M. and Aarts, M. G. (2014). Genotype× environment interaction QTL mapping in plants: lessons from Arabidopsis. *Trends in Plant Science* 19(6), 390–8.

Eshed Y and Zamir D (1995). An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield associated QTL. *Genetics* 141, 1147–62.

Eshed Y and Zamir D (1996). Less-than-additive epistatic interactions of quantitative trait loci in tomato. *Genetics* 143, 1807–17.

Fereres, E. and Soriano, M. A. (2007). Deficit irrigation for reducing agricultural water use. *Journal of Experimental Botany* 58(2), 147–59.

Fishman, S. and Génard, M. (1998). A biophysical model of fruit growth: simulation of seasonal and diurnal dynamics of mass. *Plant, Cell & Environment* 21(8), 739–52.

Fodor, A., Segura, V., Denis, M., et al. (2014). Genome-Wide Prediction Methods in Highly Diverse and Heterozygous Species: Proof-of-Concept through Simulation in Grapevine. *PLoS ONE* 9(11), e110436.

Frary A, Doganlar S, Frampton A, Fulton T, Uhlig J, Yates H and Tanksley S (2003). Fine mapping of quantitative trait loci for improved fruit characteristics from *Lycopersicon chmielewskii* chromosome 1. *Genome* 46, 235–43.

Frary A, Fulton TM, Zamir D and Tanksley SD (2004). Advance backcross QTL analysis of a *Lycopersicon esculentum* x L. pennellii cross and identification of possible orthologs in the Solanaceae. *Theoretical and Applied Genetics* 108, 485–96.

Frary A, Nesbitt TC, Frary A, Grandillo S, Van der Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB and Tanksley SD (2000). fw-2.2: a quantitative trait locus key to the evolution of tomato fruit size. *Science* 289:85–8

Fray RG, Grierson D (1993). Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing, complementation and co-suppression. *Plant Molecular Biology* 22(4):589–602.

Fridman E, Carrari F, Liu YS, Fernie AR, Zamir D (2004). Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* 305, 1786–9.

Fridman E, Pleban T and Zamir D (2000). A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proceedings of the National Academy of Sciences of the USA* 97, 4718–23.

Fridman E and Zamir D (2003). Functional divergence of a syntenic invertase gene family in tomato, potato, and Arabidopsis. *Plant Physiology* 131, 603–9.

Fulton TM, Beck-Bunn T, Emmatty D, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D and Tanksley SD (1997). QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. *Theoretical and Applied Genetics* 95, 881–94.

Fulton TM, Bucheli P, Voirol E, Lopez J, Petiard V and Tanksley SD (2002). Quantitative trait loci (QTL) affecting sugars, organic acids and other biochemical properties possibly contributing to flavor, identified in four advanced backcross populations of tomato. *Euphytica* 127, 163–77.

Fulton TM, Grandillo S, Beck-Bunn T, Fridman E, Frampton A, Lopez J, Petiard V, Uhlig J, Zamir D and Tanksley SD (2000). Advanced backcross QTL analysis of a *Lycopersicon esculentum x Lycopersicon parviflorum* cross. *Theoretical and Applied Genetics* 100, 1025–42.

Gautier, H., Diakou-Verdin, V., Bénard, C., Reich, M., Buret, M., Bourgaud, F. and Génard, M. (2008). How does tomato quality (sugar, acid, and nutritional quality) vary with ripening stage, temperature, and irradiance?. *Journal of Agricultural and Food Chemistry*, 56(4), 1241–50.

Gautier, H., Lopez-Lauri, F., Massot, C., Murshed, R., Marty, I., Grasselly, D. and Genard, M. (2010). Impact of ripening and salinity on tomato fruit ascorbate content and enzymatic activities related to ascorbate

recycling. *Functional Plant Science and Biotechnology*, 4(1), 66–75.

Giovannoni J.J. (2004). Genetic regulation of fruit development and ripening. *The Plant Cell* 16, 170–80.

Godt DE and Roitsch T (1997). Regulation and tissue-specific distribution of mRNAs for three extracellular invertase isoenzymes of tomato suggests an important function in establishing and maintaining sink metabolism. *Plant Physiology* 115, 273–82.

Goldman IL, Paran I and Zamir D (1995). Quantitative trait locus analysis of a recombinant inbred line population derived from a *Lycopersicon esculentum × L. cheesmanii* cross. *Theoretical and Applied Genetics* 90, 925–32.

Goulet, C., et al. (2012). Role of an esterase in flavor volatile variation within the tomato clade. *Proceedings of the National Academy of Sciences of the USA* 109, 19009–14.

Grandillo S, Ku HM and Tanksley SD (1999). Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theoretical and Applied Genetics* 99, 978–87.

Grandillo S and Tanksley SD (1996). QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersicon pimpinellifolium*. *Theoretical and Applied Genetics* 92, 935–51.

Guichard, S., Bertin, N., Leonardi, C. and Gary, C. (2001). Tomato fruit quality in relation to water and carbon fluxes. *Agronomie* 21(4), 385–92.

Guichard, S., Gary, C., Leonardi, C. and Bertin, N. (2005). Analysis of growth and water relations of tomato fruits in relation to air vapor pressure deficit and plant fruit load. *Journal of Plant Growth Regulation* 24(3), 201–13.

Gur, A., Semel, Y., Osorio, S., Friedmann, M., Seekh, S., Ghareeb, B. and Zamir, D. (2011). Yield quantitative trait loci from wild tomato are predominately expressed by the shoot. *Theoretical and Applied Genetics* 122(2), 405–20.

Hall, A. J., Minchin, P. E., Clearwater, M. J. and Génard, M. (2013). A biophysical model of kiwifruit (Actinidia deliciosa) berry development. *Journal of Experimental Botany*, 317.

Hamilton, J. P., Sim, S.-C., Stoffel, K., et al. (2012). Single Nucleotide Polymorphism Discovery In Cultivated Tomato Via Sequencing By Synthesis. *Plant Gene* 5(1), 17–29.

Harker FR, Redgwell RJ, Hallett IC., et al. (1997). Texture of fresh fruit. *Horticultural Reviews* 20, 121–224.

Henderson, C. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31(423), 423.

Hileman LC, Sundstrom JF, Litt A, et al. (2006). Molecular and phylogenetic analyses of the MADS-Box gene family in tomato. *Molecular Biology and Evolution* 23, 2245–58.

Ho L.C. (1996). The mechanism of assimilate partitioning and carbohydrate compartmentation in fruit in relation to the quality and yield of tomato. *Journal of Experimental Botany* 47, 1239–43.

Hobson GE and Bedford L (1989). The composition of cherry tomatoes and its relation to consumer acceptability. *Journal of Horticulture Science* 64, 321–9.

Hovav R, Chehanovsky N, Moy M., et al. (2007). The identification of a gene (Cwp1), silenced during Solanum evolution, which causes cuticle microfissuring and dehydration when expressed in tomato fruit. *The Plant Journal* 52(4), 627–39.

Isaacson T, Ronen G, Zamir D., et al. (2002). Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of beta-carotene and xanthophylls in plants. *The Plant Cell* 14(2), 333–42.

Ito Y, Kitagawa M, Ihashi N, et al.. (2008). DNA-binding specificity, transcriptional activation potential, and the

rin mutation effect for the tomato fruit-ripening regulator RIN. *The Plant Journal* 55, 212–23.

Jansen, R. C., Van Ooijen, J. W., Stam, P., Lister, C. and Dean, C. (1995). Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. *Theoretical and Applied Genetics*, 91(1), 33–7.

Jin S, Chen CCS and Plant AL (2000). Regulation byABAof osmoticstress-induced changes in protein synthesis in tomato roots. *Plant, Cell & Environment* 23, 51–60.

Johansson L, Haglund A, Berglund L, et al. (1999). Preference for tomatoes, affected by sensory attributes and information about growth conditions. *Food Quality and Preference* 10, 289–98.

Johansson L, Haglund A, Berglund L, et al. (1999). Preference for tomatoes, affected by sensory attributes and information about growth conditions. *Food Quality and Preference* 10, 289–98.

Johnstone, P. R., Hartz, T. K., LeStrange, M., Nunez, J. J. and Miyao, E. M. (2005). Managing fruit soluble solids with late-season deficit irrigation in drip-irrigated processing tomato production. *Hortscience* 40(6), 1857–61.

Jonas, E. and de Koning, D.-J. (2013). Does genomic selection have a future in plant breeding? *Trends in Biotechnology* 31(9), 497–504.

Jones RA.. (1986). Breeding for improved post-harvest tomato quality: genetical aspects. *Acta Horticulturae* 190, 77–87.

Kader A, Morris L, Stevens M and Albrightholton M (1978). Composition and flavor quality of fresh market tomatoes as influenced by some post-harvest handling procedures. *Journal of the American Society for Horticultural Science* 103, 6–13.

Klee H.J. (2010). Improving the flavor of fresh fruits: genomics, biochemistry, and biotechnology. *New Phytologist* 187, 44–56.

Klee HJ and Tieman DM (2013). Genetic challenges of flavor improvement in tomato. *Trends in Genetics* 29, 257–62.

Korte, A., Vilhjálmsson, B. J., Segura, V., Platt, A., Long, Q. and Nordborg, M. (2012). A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature Genetics* 44(9), 1066–71.

Kumar, S., ChagnÄ, D., Bink, M. C. A. M., et al. (2012). Genomic selection for fruit quality traits in apple (Malus domestica Borkh.). *PLoS ONE* 7.

Labate JA, Grandillo, S., Fulton, T., Muños, S., Caicedo, AL., Peralta, I., Ji, Y., Chetelat, RT., Scott, JW., Gonzalo, MJ., Francis, D., Yang, W., van der Knaap, E., Baldo, AM., Smith-White, B., Mueller, LA., Prince, JP., Blanchard, NE., Storey, DB., Stevens, MR., Robbins, MD., Fen Wang, J., Liedl, BE., O'Connell, MA., Stommel, JR., Aoki, K., Iijima, Y., Slade, AJ., Hurst, SR., Loeffler, D., Steine, MN., Vafeados, D., McGuire, C., Freeman, C., Amen, A., Goodstal, J., Facciotti, D., Van Eck, J. and Causse, M. (2007). 1 Tomato. In C. Kole (Ed.), *Genome Mapping and Molecular Breeding in Plants*, Vol. 5, Springer-Verlag, Berlin Heidelberg, 11–135.

Léchaudel, M. and Joas, J. (2006). Quality and maturation of mango fruits of cv. Cogshall in relation to harvest date and carbon supply. *Crop and Pasture Science* 57(4), 419–26.

Lecomte L, Saliba-Colombani V, Gautier A, Gomez-Jimenez MC, Duffé P, Buret M and Causse M (2004). Fine mapping of QTLs of chromosome 2 affecting the fruit architecture and composition of tomato. *Molecular Breeding* 13, 1–14.

Lecomte, L., Duffé, P., Buret, M., et al. (2004). Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. TAG *Theoretical and Applied Genetics* V109(3), 658–68.

Li, S., Wang, J. and Zhang, L. (2015). Inclusive composite interval mapping of QTL by environment interactions

in biparental populations. *PloS ONE*, 10(7), e0132414.

Lin, T., Zhu, G., Zhang, J., et al. (2014). Genomic analyses provide insights into the history of tomato breeding. *Nature Genetics* 46(11), 1220–6.

Lippman Z and Tanksley SD (2001). Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *Lycopersicon pimpinellifolium* and *L. esculentum* var. giant heirloom. *Genetics* 158, 413–22.

Lippman ZB, Cohen O, Alvarez JP., et al. (2008). The making of a compound inflorescence in tomato and related nightshades. *PLoS Biol* (11), e288.

Liu JP, Van Eck J, Cong B and Tanksley SD (2002). A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proceedings of the National Academy of Sciences of the USA* 99, 13302–6

Liu, H. F., Génard, M., Guichard, S. and Bertin, N. (2007). Model-assisted analysis of tomato fruit growth in relation to carbon and water fluxes. *Journal of Experimental Botany*, 58(13), 3567–80.

Mageroy, M.H., et al. (2012). A *Solanum lycopersicum* catechol-Omethyltransferase involved in synthesis of the flavor molecule guaiacol. *The Plant Journal* 69, 1043–51

Manning K, Tör M, Poole M., et al. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics* 38(8), 948–52.

Manning K, Tor M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ and Seymour GB (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics* 38, 948–52.

Mao L, Begum D, Chuang HW., et al. (2000). JOINTLESS is a MADS-box gene controlling tomato flower abscission zone development. *Nature* 406(6798), 910–13.

Matas AJ, Gapper NE, Chung MY., et al. (2009). Biology and genetic engineering of fruit maturation for enhanced quality and shelf-life. *Current Opinion in Biotechnology* 20, 197–203.

Mathieu S, Cin VD, Fei ZJ., et al. 2009. Flavour compounds in tomato fruits: identification of loci and potential pathways affecting volatile composition. *Journal of Experimental Botany* 60, 325–37.

McGlasson WB, Last JH, Shaw KJ., et al. (1987). Influence of the non-ripening mutants rin and nor on the aroma of tomato fruit. *Hortscience* 22, 632–4.

Melchinger AE, Utz HF and Scho¨n CC. (1998). QTL mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* 149, 383–403.

Meli VS, Ghosh S, Prabha TN., et al. (2009). Enhancement of fruit shelf life by suppressing N-glycan processing enzymes. *Proceedings of the National Academy of Sciences of the USA* 107, 2413–18.

Meuwissen, T. H. E., Hayes, B. J. and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157(4), 1819–29.

Minic Z, Jamet E, San-Clemente H, et al. (2009). Transcriptomic analysis of Arabidopsis developing stems: a close-up on cell wall genes. *BMC Plant Biology* 9, 17.

Mitchell, J. P., Shennan, C., Grattan, S. R. and May, D. M. (1991). Tomato fruit yields and quality under water deficit and salinity. *Journal of the American Society for Horticultural Science*, 116(2), 215–21.

Molinero-Rosales N, Jamilena M, Zurita S, et al. (1999). FALSIFLORA, the tomato orthologue of FLORICAULA and LEAFY, controls flowering time and floral meristem identity. *The Plant Journal* 20(6), 685–93.

Molinero-Rosales N, Latorre A, Jamilena M., et al. (2004). SINGLE FLOWER TRUSS regulates the transition and

maintenance of flowering in tomato. *Planta* 218(3), 427–34.

Monforte, A. J., Asins, M. J. and Carbonell, E. A. (1996). Salt tolerance in *Lycopersicon* species. IV. Efficiency of marker-assisted selection for salt tolerance improvement. *Theoretical and Applied Genetics* 93(5–6), 765–72.

Monforte, A. J., Asins, M. J. and Carbonell, E. A. (1997a). Salt tolerance in *Lycopersicon* species. V. Does genetic variability at quantitative trait loci affect their analysis?. *Theoretical and Applied Genetics* 95(1–2), 284–93.

Monforte, A. J., Asins, M. J. and Carbonell, E. A. (1997b). Salt tolerance in *Lycopersicon* species VI. Genotype-by-salinity interaction in quantitative trait loci detection: constitutive and response QTLs. *Theoretical and Applied Genetics* 95(4), 706–13.

Muños S, Ranc N, Botton E, Bérard A, Rolland S, Duffé P, Carretero Y, Le Paslier MC, Delalande C, Bouzayen M, Brunel D and Causse M. (2011). Increase in tomato locule number is controlled by two key SNP located near Wuschel. *Plant Physiology* 4, 2244–54.

Muranty, H. l. N., Troggio, M., Sadok, I. S. B., et al. (2015). Accuracy and responses of genomic selection on key traits in apple breeding. *Horticulture Research* 2, 15060.

Mustilli AC, Fenzi F, Ciliento R., et al. (1999). Phenotype of the tomato high pigment-2 mutant is caused by a mutation in the tomato homolog of DEETIOLATED1. *The Plant Cell* 11, 145–57.

Nuruddin, M. M., Madramootoo, C. A. and Dodds, G. T. (2003). Effects of water stress at different growth stages on greenhouse tomato yield and quality. *Hortscience*, 38(7), 1389–93.

Okabe Y, Asamizu E, Saito T, Matsukura C, Ariizumi T, Brès C, Rothan C, Mizoguchi T and Ezura H. (2011). Tomato TILLING technology: development of a reverse genetics tool for the efficient isolation of mutants from Micro-Tom mutant libraries. Plant and Cell Physiology 52(11), 1994–2005.

Osorio, S., Ruan, Y. L. and Fernie, A. R. (2014). An update on source-to-sink carbon partitioning in tomato. *Frontiers in Plant Science*, 5, 516.

Page D, Gouble B, Valot B., et al. (2010). Down-regulated protective proteins in tomato correlating with decreased tolerance to low-temperature storage. *Planta* (in press).

Pascale, S. D., Maggio, A., Fogliano, V., Ambrosino, P. and Ritieni, A. (2001). Irrigation with saline water improves carotenoids content and antioxidant activity of tomato. *The Journal of Horticultural Science and Biotechnology* 76(4), 447–53.

Pascual-Banuls, L., Xu, J., Biais, B., et al. (2013). Deciphering genetic diversity and inheritance of tomato fruit weight and composition through a systems biology approach. *Journal of Experimental Botany.* doi: 10.1093/jxb/ert349.

Pascual, L., Desplat, N., Huang, B.E., Desgroux, A., Bruguier, L., Bouchet, J.-P., Le, Q.H., Chauchard, B., Verschave, P. and Causse, M. (2015). Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. *Plant Biotechnology Journal* 13, 565–77.

Pascual L, Albert, E., Sauvage, C., Duangjit, J., Bouchet, JP., Bitton, F., Desplat, N., Brunel, D., Le Paslier, MC., Ranc, N., Bruguier, L., Chauchard, B., Verschave, P. and Causse, M. (2016). Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels. *Plant Science* 242, 120–30.

Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES and Tanksley SD (1991). Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. *Genetics* 127, 181–97.

Paterson AH, de Verna JW, Lanini B and Tanksley SD (1990). Fine mapping of quantitative trait loci using

selected overlapping recombinant chromosomes, in an interspecies cross of tomato. *Genetics* 124, 735–42.

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE and Tanksley SD (1988). Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335, 721–6.

Pnueli L, Carmel-Goren L, Hareven D., et al. (1998). The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. *Development* 125, 1979–89.

Poiroux-Gonord, F., Bidel, L. P., Fanciullino, A. L., Gautier, H., Lauri-Lopez, F. and Urban, L. (2010). Health benefits of vitamins and secondary metabolites of fruits and vegetables and prospects to increase their concentrations by agronomic approaches. *Journal of Agricultural and Food Chemistry* 58(23), 12065–82.

Poland, J., Endelman, J., Dawson, J., et al. (2012). Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *The Plant Genome* 5(3), 103–13.

Powell AL, Nguyen CV, Hill T., et al. (2012). Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. *Science*. 336(6089), 1711–15.

Prudent M, Lecomte, A., Bouchet, JP., Bertin, N., Causse, M. and Génard, M. (2011). Combining ecophysiological modelling and quantitative trait loci analysis to identify key elementary processes underlying tomato fruit sugar concentration. *Journal of Experimental Botany* 62, 907–11.

Prudent M, Causse M, Génard M, Tripodi P, Grandillo S and Bertin N (2009). Genetic and ecophysiological analysis of tomato fruit weight and composition – Influence of carbon availability on QTL detection. *Journal of Experimental Botany* 60(3), 923–37.

Quilot, B., Kervella, J., Génard, M. and Lescourret, F. (2005). Analysing the genetic control of peach fruit quality through an ecophysiological model combined with a QTL approach. *Journal of Experimental Botany* 56(422), 3083–92.

Rambla JL, Tikunov YM, Monforte AJ, et al. (2014). The expanded tomato fruit volatile landscape. *Journal of Experimental Botany* 65, 4613–23.

Ranc, N., Munos, S., Santoni, S., et al. (2008). A clarified position for *Solanum lycopersicum* var. *cerasiforme* in the evolutionary history of tomatoes (solanaceae). *BMC Plant Biology* 8, 130.

Reymond, M., Muller, B., Leonardi, A., Charcosset, A. and Tardieu, F. (2003). Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiology* 131(2), 664–75.

Ripoll, J., Urban, L. and Bertin, N. (2016b). The potential of the MAGIC TOM parental accessions to explore the genetic variability in tomato acclimation to repeated cycles of water deficit and recovery. *Frontiers in Plant Science*, 6.

Ripoll, J., Urban, L., Brunel, B. and Bertin, N. (2016a). Water deficit effects on tomato quality depend on fruit developmental stage and genotype. *Journal of Plant Physiology* 190, 26–35.

Ripoll, J., Urban, L., Staudt, M., Lopez-Lauri, F., Bidel, L. P. and Bertin, N. (2014). Water shortage and quality of fleshy fruits—making the most of the unavoidable. *Journal of Experimental Botany* 65(15), 4097–117.

Rodríguez GR, Muños, S., Anderson, C., Sim, SC., Michel, A., Causse, M., McSpadden Gardener, BB., Francis, D. and van der Knaap, E. (2011). Distribution of SUN, OVATE, LC, and FAS Alleles in Tomato Germplasm and their Effect on Fruit Morphology. *Plant Physiology* 156, 275–85.

Ronen G, Carmel-Goren L, Zamir D., et al. (2000). An alternative pathway to β-carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. *Proceedings of the National Academy of Sciences of the USA* 97, 11102–7.

Ronen GL, Cohen M, Zamir D., et al. (1999). Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant Delta. *The Plant Journal* 17, 341–51.

Rosales, M. A., Ruiz, J. M., Hernández, J., Soriano, T., Castilla, N. and Romero, L. (2006). Antioxidant content and ascorbate metabolism in cherry tomato exocarp in relation to temperature and solar radiation. *Journal of the Science of Food and Agriculture* 86(10), 1545–51.

Rose J, Catalá C, Gonzalez-Carranza Z, et al. (2003). Plant cell wall disassembly. In JKC Rose (Ed.), *The Plant Cell Wall*. Oxford, Blackwell, 264–324.

Rose JKC, Bashir S, Giovannoni JJ, Jahn MM and Saravanan RS (2004). Tackling the plant proteome: practical approaches, hurdles and experimental tools. *The Plant Journal* 39, 715–33.

Rose JKC, Cosgrove DJ, Albersheim P, et al. (2000). Detection of Expansin Proteins and Activity during Tomato Fruit Ontogeny. *Plant Physiology* 123, 1583–92.

Ruggieri, V., Francese, G., Sacco, A., et al. (2014). An association mapping approach to identify favourable alleles for tomato fruit quality breeding. *BMC Plant Biology* 14, 337.

Rutkoski, J. E., Heffner, E. L. and Sorrells, M. E. (2011). Genomic selection for durable stem rust resistance in wheat. *Euphytica* 179(1), 161–73.

Sacco, A., Ruggieri, V., Parisi, M., et al. (2015). Exploring a Tomato Landraces Collection for Fruit-Related Traits by the Aid of a High-Throughput Genomic Platform. *PLoS ONE* 10(9), e0137139.

Saïdou, A. A., Thuillet, A. C., Couderc, M., Mariac, C. and Vigouroux, Y. (2014). Association studies including genotype by environment interactions: prospects and limits. *BMC Genetics* 15(1), 1.

Saladie M, Rose JKC, Cosgrove DJ, et al. (2006). Characterization of a new xyloglucan endotransglucosylase/hydrolase (XTH) from ripening tomato fruit and implications for the diverse modes of enzymic action. *The Plant Journal* 47, 282–95.

Saliba-Colombani V, Causse M, Langlois D, Philouze J and Buret M (2001). Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theoretical and Applied Genetics* 102, 259–72.

Sari-Gorla M, Calinski T, Kaczmarek Z and Krajewski P. 1997. Detection of QTL3environment interaction in maize by a least squares interval mapping method. *Heredity* 78, 146–57.

Sato T, Iwatsubo T, Takahashi M., et al. (1993). Intercellular localization of acid invertase in tomato fruit and molecular cloning of a cDNA for the enzyme. *Plant Cell Physiol*ogy 34(2), 263–9.

Sauvage C, Segura V, Bauchet G, et al. (2014). Genome-Wide Association in Tomato Reveals 44 Candidate Loci for Fruit Metabolic Traits. *Plant Physiology* 165, 1120–32.

Schauer N, Semel Y, Roessner U, et al. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature Biotechnology* 24, 447–54.

Schumacher K, Schmitt T, Rossberg M., et al. (1999). The Lateral suppressor (Ls) gene of tomato encodes a new member of the VHIID protein family. *Proceedings of the National Academy of Sciences of the USA* 96(1), 290–5.

Segura, V., Vilhjalmsson, B. J., Platt, A., et al. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature Genetics* 44(7), 825–30.

Semel Y, Nissenbaum J, Menda N, Zinder M, Krieger U, Issman N, Pleban T, Lippman Z, Gur A and Zamir D (2006). Overdominant quantitative trait loci for yield and fitness in tomato. *Proceedings of the National Academy of Sciences of the USA* 103, 12981–6.

Semel, Y., Schauer, N., Roessner, U., Zamir, D. and Fernie, A. R. (2007). Metabolite analysis for the comparison of irrigated and non-irrigated field grown tomato of varying genotype. *Metabolomics* 3(3), 289–95.

Seymour GB, Manning K, Eriksson EM, Popovich AH and King GJ (2002). Genetic identification and genomic organization of factors affecting fruit texture. *Journal of Experimental Botany* 53, 2065–71.

Seymour GB, Ostergaard L, Chapman NH, Knapp S and Martin C (2013). Fruit Development and Ripening. *Annual Review of Plant Biology* 64, 219–41.

Shackel KA, Greve C, Labavitch JM, et al. (1991). Cell turgor changes associated with ripening in tomato pericarp tissue. *Plant Physiology* 97, 814–16.

Sim, S.-C., Van Deynze, A., Stoffel, K., et al. (2012b). High-Density SNP Genotyping of Tomato (*Solanum lycopersicum* L.) Reveals Patterns of Genetic Variation Due to Breeding. *PLoS ONE* 7(9), e45520.

Sim, S.-C., Durstewitz, G., Plieske, J. R. et al. (2012a). Development of a Large SNP Genotyping Array and Generation of High-Density Genetic Maps in Tomato. *PLoS ONE* 7(7), e40563.

Simkin AJ., et al. (2004). The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles b-ionone, pseudoionone, and geranylacetone. *The Plant Journal* 40, 882–92.

Sinesio F, Cammareri M, Moneta E, et al. (2009). Sensory quality of fresh French and Dutch market tomatoes: a preference mapping study with Italian consumers. *Journal of Food Science* 75, S55–67.

Smith DL, Abbott JA and Gross KC. (2002). Down-regulation of tomato beta-galactosidase 4 results in decreased fruit softening. *Plant Physiology* 129, 1755–62.

Speirs J, Lee E, Holt K, et al. (1998). Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavor aldehydes and alcohols. *Plant Physiology* 117, 1047–58.

Stevens MA (1986). Inheritance of tomato fruit quality components. *Plant Breeding Reviews* 4, 273–311.

Stevens R, Page, D., Gouble, B., Garchery, C., Zamir, D. and Causse, M. (2008). Tomato fruit ascorbic acid content is linked with monodehydroascorbate reductase activity and tolerance to chilling stress. *Plant Cell and Environment* 31 (8), 1086–96.

Stikic, R., Popovic, S., Srdic, M., Savic, D., Jovanovic, Z., Prokic, L. J. and Zdravkovic, J. (2003). Partial root drying (PRD): a new technique for growing plants that saves water and improves the quality of fruit. Bulg. J. *Plant Physiology* 29(3–4), 164–71.

Tadmor Y, Fridman E, Gur A, Larkov O, Lastochkin E, Ravid U, Zamir D and Lewinsohn E (2002). Identification of malodorous, a wild species allele affecting tomato aroma that was selected against during domestication. *Journal of Agricultural and Food Chemistry* 50, 2005–9.

Tanksley SD. (1993). Mapping polygenes. *Annual Review of Genetics* 27, 205–33.

Tanksley SD (2004). The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *The Plant Cell* 16, S181–S189

Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed T, Pétiard V, Lopez J and Beck-Bunn T (1996). Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinnellifolium*. *Theoretical and Applied Genetics* 92, 213–24.

The Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485(7400), 635–41.

Thompson AJ and Corlett JE. (1995). mRNA levels of four tomato (*Lycopersicon esculentum* Mill. L.) genes are related to fluctuating plant and soil water status. *Plant, Cell & Environment* 18, 773–80.

Thompson AJ, Jackson AC, Parker RA., et al. (2000). Abscisic acid biosynthesis in tomato: regulation of zeaxanthin epoxidase and 9-cis-epoxycarotenoid dioxygenase mRNAs by light/dark cycles, water stress and abscisic acid. *Plant Molecular Biology* 42(6), 833–45.

Tieman, D.M. and Handa, A.K. (1994). Regulation in pectin methylesterase activity modifies tissue integrity and cation levels in ripening tomato (Lycopersicon esculentum Mill.) fruits. *Plant Physiology* 106, 429–36.

Tieman D, Zeigler M, Schmelz E., et al. (2010). Functional analysis of a tomato salicylic acid methyl transferase and its role in synthesis of the flavor volatile methyl salicylate. *The Plant Journal*ournal 62, 113–23.

Tieman DM, Zeigler M, Schmelz EA, Taylor MG, Bliss P, Kirst M and Klee HJ (2006). Identification of loci affecting flavour volatile emissions in tomato fruits. *Journal of Experimental Botany* 57, 887–96.

Tieman, D.M., et al. (2006). Aromatic amino acid decarboxylases participate in the synthesis of the flavor and aroma volatiles 2-phenylethanol and 2-phenylacetaldehyde in tomato fruits. *Proceedings of the National Academy of Sciences of the USA* 103, 8287–92.

Tieman, D.M., et al. (2007). Tomato phenylacetaldehyde reductases catalyze the last step in the synthesis of the aroma volatile 2-phenylethanol. *Phytochemistry* 68, 2660–9.

Tikunov Y, Molthoff, J., de Vos, R C., Beekwilder, J., van Houwelingen, A., van der Hooft, J.J., Nijenhuis-de Vries, M., Labrie, C W., Verkerke, W., van de Geest, H., Viquez Zamora, M., Presa, S., Rambla, JL., Granell, A., Hall, RD. and Bovy, A. G. (2013). NON-SMOKY GLYCOSYLTRANSFERASE1 Prevents the Release of Smoky Aroma from Tomato Fruit. *The Plant Cell* 25, 8 3067–78.

Tikunov Y, Lommen A, de Vos CHR, et al. (2005). A Novel Approach for Nontargeted Data Analysis for Metabolomics. Large-Scale Profiling of Tomato Fruit Volatiles. *Plant Physiology* 139, 1125–37.

Tikunov, Y. M., Molthoff, J., de Vos, R. C. H., et al. (2013). Non-smoky glycosyl transferase1 prevents the release of smoky aroma from tomato fruit. *The Plant Cell* 25(8), 3067–78.

Tinker NA, Mather DE, Rossnagel BG, Kasha KJ, Kleinhofs A and Hayes PM. (1996). Regions of the genome that affect agronomic performance in two-row barley. *Crop Science* 36, 1053–62.

Tomato-Genome-Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–41.

Tucker, G., Price, A. L. and Berger, B. (2014). Improving the power of GWAS and avoiding confounding from population stratification with PC-Select. *Genetics* 197(3), 1045–9.

Uozumi, A., Ikeda, H., Hiraga, M., Kanno, H., Nanzyo, M., Nishiyama, M. and Kanayama, Y. (2012). Tolerance to salt stress and blossom-end rot in an introgression line, IL8–3, of tomato. *Scientia Horticulturae* 138, 1–6.

van der Knaap E, Lippman ZB and Tanksley SD (2002). Extremely elongated tomato fruit controlled by four quantitative trait loci with epistatic interactions. *Theoretical and Applied Genetics* 104 (2–3), 241–7.

van der Knaap E and Tanksley SD (2003). The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in Yellow Stuffer tomato. *Theoretical and Applied Genetics* 107, 139–47.

van Eeuwijk, F. A., Bink, M. C., Chenu, K. and Chapman, S. C. (2010). Detection and use of QTL for complex traits in multiple environments. *Current Opinion In Plant Biology* 13(2), 193–205.

Venter, F. (1977). Solar radiation and vitamin C content of tomato fruits. *Acta Horticulturae*, 58, 121–7.

Verbyla, A. P., Cavanagh, C. R. and Verbyla, K. L. (2014). Whole-Genome Analysis of Multienvironment or Multitrait QTL in MAGIC. *G3: Genes| Genomes| Genetics* 4(9), 1569–84.

Vicente AR, Saladie M, Rose JKC, et al. (2007). The linkage between cell wall metabolism and fruit softening: looking to the future. *Journal of the Science of Food and Agriculture* 87, 1435–48.

Vogel JT, Tan BC, McCarty DR., et al. (2008). The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. *Journal of Biological Chemistry* 283, 11364–73.

Vrebalov J, Ruezinsky D, Padmanabhan V., et al. (2002). A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. *Science* 296(5566), 343–6.

Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W and Giovannoni J (2002). A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (Rin) locus. *Science* 296, 343–6.

Weller JL, Perrotta G, Schreuder ME., et al. (2001). Genetic dissection of blue-light sensing in tomato using mutants deficient in cryptochrome 1 and phytochromes A, B1 and B2. *The Plant Journal* 25(4), 427–40.

Whitaker BD (2008). Postharvest flavor deployment and degradation in fruits and vegetables. In Bruckner B and Grant Willie S (Eds), *Fruit Vegetable Flavour*. CRC Press, Cambridge, UK, 103–31.

Wilkinson JQ, Lanahan MB, Yen HC., et al. (1995). An ethylene-inducible component of signal transduction encoded by never-ripe. *Science* 270(5243), 1807–9.

Xiao H, Jiang N, Schaffner E., et al. (2008). A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science* 319, 1527–30.

Xu C, Liberatore, K L., MacAlister, C A., Huang, Z., Chu, YH., Jiang, K., Brooks, C., Ogawa-Ohnishi, M., Xiong, G., Pauly, M., Van Eck, J., Matsubayashi, Y., van der Knaap, E. and Lippman, ZB. (2015). A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nature Genetics* 47, 784–95.

Xu, J., Ranc, N., Munos, S., et al. (2013). Phenotypic diversity and association mapping for fruit quality traits in cultivated tomato and related species. *Theoretical and Applied Genetics* 126(3), 567–81.

Yamamoto, E., Matsunaga, H., Onogi, A., et al. (2016). A simulation-based breeding design that uses whole-genome prediction in tomato. *Scientific Reports* 6, 19454.

Yelle S, Chetelat RT, Dorais M, DeVerna JW and Bennett AB (1991). Sink metabolism in tomato fruit. IV. Genetic and biochemical analysis of sucrose accumulation. *Plant Physiology* 95, 1026–35.

Yelle S, Hewitt JD, Robinson NL, et al. (1988). Sink metabolism in tomato fruit .3. Analysis of carbohydrate assimilation in a wild-species. *Plant Physiology* 87, 737–40.

Yousef GG and Juvik JA (2001). Evaluation of breeding utility of a chromosomal segment from *Lycopersicon chmielewskii* that enhances cultivated tomato soluble solids. *Theoretical and Applied Genetics* 103, 1022–7.

Yu, J., Pressoir, G., Briggs, W. H., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38(2), 203–8.

Zanor MI, Rambla, JL., Chaïb, J., Steppa, A., Medina, A., Granell, A., Fernie, A. and Causse, M. (2009). Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. *Journal of Experimental Botany* 60, 2139–54.

Zhang J, Chen R, Xiao J., et al. (2007). A single-base deletion mutation in SlIAA9 gene causes tomato (*Solanum lycopersicum*) entire mutant. *Journal of Plant Research* 120(6), 671–8.

Zhang, J., Zhao, J., Liang, Y., et al. (2016). Genome-wide association-mapping for fruit quality traits in tomato. *Euphytica* 207(2), 439–51.

Zhou, S., Wei, S., Boone, B. and Levy, S. (2007). Microarray analysis of genes affected by salt stress in tomato. *African Journal of Environmental Science and Technology* 1(2), 14–26.

Appendix 3: Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels.

Article published in Plant Science, 2015

Appendix 3

Plant Science 242 (2016) 120–130

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Dissecting quantitative trait variation in the resequencing era: complementarity of bi-parental, multi-parental and association panels

Laura Pascual[a,1,5], Elise Albert[a,5], Christopher Sauvage[a], Janejira Duangjit[a,2], Jean-Paul Bouchet[a], Frédérique Bitton[a], Nelly Desplat[a,3], Dominique Brunel[b], Marie-Christine Le Paslier[b], Nicolas Ranc[a,4], Laure Bruguier[c], Betty Chauchard[c], Philippe Verschave[c], Mathilde Causse[a,*]

[a] INRA, UR1052, Centre de Recherche PACA, 67 Allée des Chênes CS60094, 84143 Montfavet Cedex, France
[b] INRA, US1279, Etude du Polymorphisme des Génomes végétaux (EPGV), CEA-IG/CNG, 2 rue Gaston Crémieux, 91057 Evry, France
[c] Vilmorin S.A. – Groupe Limagrain, Centre de Recherche de La Costière, Route de Meynes, 30210 Ledenon, France

ARTICLE INFO

Article history:
Received 13 April 2015
Received in revised form 12 June 2015
Accepted 16 June 2015
Available online 23 June 2015

Keywords:
Tomato
QTL mapping
Genome-wide association
Fruit quality
Resequencing

ABSTRACT

Quantitative trait loci (QTL) have been identified using traditional linkage mapping and positional cloning identified several QTLs. However linkage mapping is limited to the analysis of traits differing between two lines and the impact of the genetic background on QTL effect has been underlined. Genome-wide association studies (GWAs) were proposed to circumvent these limitations. In tomato, we have shown that GWAs is possible, using the admixed nature of cherry tomato genomes that reduces the impact of population structure. Nevertheless, GWAs success might be limited due to the low decay of linkage disequilibrium, which varies along the genome in this species.

Multi-parent advanced generation intercross (MAGIC) populations offer an alternative to traditional linkage and GWAs by increasing the precision of QTL mapping. We have developed a MAGIC population by crossing eight tomato lines whose genomes were resequenced. We showed the potential of the MAGIC population when coupled with whole genome sequencing to detect candidate single nucleotide polymorphisms (SNPs) underlying the QTLs. QTLs for fruit quality traits were mapped and related to the variations detected at the genome sequence and expression levels. The advantages and limitations of the three types of population, in the context of the available genome sequence and resequencing facilities, are discussed.

1. Introduction

Agronomic traits are usually under the control of several genes with variable effects modulated by the environment. Since the pioneer work of Paterson and colleagues [1], deciphering the genetic control of quantitative traits into quantitative trait loci (QTLs) has been studied through QTL mapping [2,3]. Quantitative trait loci have been mapped in many crops in biparental populations segregating after one (F2 populations) or a few selfing generations (in recombinant inbred lines, RIL), when selfing is possible, or on advanced backcross progenies. Populations of introgression lines covering the whole genome are also helpful to identify QTLs from wild species in a cultivated genetic background [4]. Among hundreds of QTLs mapped, only a few were identified following positional cloning [5]. Nevertheless such populations allow the identification of the QTLs differing only between the two parental lines. The confidence intervals around QTLs are usually large as they only rely on one or two efficient recombination generations. Until the recent advent in genome sequencing, the number of available molecular markers was also limiting the power of this approach, particularly to fine map genes and QTLs. Since the discovery of SNP markers, thousands of markers are available, drastically changing the paradigm of QTL mapping. In the early 2000s, it was proposed

* Corresponding author. Tel.: +33 432722803.
E-mail address: mathilde.causse@avignon.inra.fr (M. Causse).
[1] Present address: Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UA, Universidad de Barcelona, Barcelona 08193, Spain.
[2] Present address: Department of Horticulture, Faculty of Agriculture, Kasetsart University, 10900 Bangkok, Thailand.
[3] Present address: Institut Bergonié, 229, Cours de l'Argonne, 33000 Bordeaux, France.
[4] Present address: Syngenta Seeds 12, 31790 St Sauveur, France.
[5] These authors contributed equally to the article.

http://dx.doi.org/10.1016/j.plantsci.2015.06.017

Appendix 3

to extend the QTL mapping approach to panels of unrelated lines through Genome- Wide Association Studies (GWAs) as first used in human genetics. The GWAs allow the discovery of QTLs in broad panels. It is particularly efficient in species with low linkage disequilibrium (LD) [6,7]. The population structure of the studied panel must be taken into account as it can lead to false positive association discovery [8,9]. If LD is sufficiently low and the number of markers is high, GWAs can land on the causal polymorphism [10].

Multi-parental populations represent intermediate populations, with more equilibrated allelic frequencies than GWAs panels and higher efficient recombination than biparental populations. Two main types of populations were proposed, Nested Association Mapping, mainly used in maize [11] and Multi-allelic Genetic Intercross (MAGIC), which have been developed in *Arabidopsis* [12], rice [13], wheat [14], barley [15] and tomato [16]. Multi-parental populations constitute a unique resource that can overcome the main limitations of GWAs and RIL studies and provide complementary information [17]. Generating new phenotypes by mixing different gene alleles permits the exploitation of QTL effects on the different founders of the population and quickly identifies causal variants [16]. Additionally, these new phenotypes constitute a highly valuable pre-breeding resource and a potential tool to develop genomic selection models. Evaluating GWAs offers unique information by allowing the analysis of a wider range of diversity, and usually provide greater precision, as they are based on recombination that has taken place during a greater number of generations. Other connected population designs were proposed [18,19] with related interests. We recently developed a tomato MAGIC population based on eight cultivated lines and showed its potential to map QTLs for fruit weight [16]. Furthermore, the genomes of the eight parental lines were sequenced [20] and the list of candidate genes was reduced by combining the predicted allelic effect at the QTLs with SNP haplotypes.

To illustrate the pros and cons of each of the three strategies, QTL mapping (in RIL and MAGIC populations) and GWAs (in a panel of accessions), we used the cultivated tomato (*Solanum lycopersicum* L.) as a model. Tomato is commonly cultivated vegetables worldwide and a model species for fruit quality and development [21]. For years QTL mapping among cultivated accessions of tomato was hampered by the low polymorphisms in the species [22], but many progenies involving distant related species were characterized [23]. Several QTL controlling fruit weight or fruit composition were mapped and characterized [24,25]. A high quality tomato genome sequence is now available [26] allowing the resequencing of several accessions [27–29] and the detection of several million of SNPs, which aids in the development of a SNP chip for diversity analyses [30]. In cultivated tomato, the molecular polymorphism is low and LD is high, although varied along the chromosomes [31]. Using a panel of highly variable cherry tomato accessions, we showed that GWAs were possible in tomato for fruit metabolite traits [32]. It was also particularly helpful to identify causative SNPs for a QTL identified by map based cloning [33].

In the present article, we compare original results of QTL and association mapping experiments using three populations: (1) a RIL population that was first mapped using RFLP markers [34]. The resequencing of the parental lines allowed the construction of a saturated map and QTL mapping using this new map; (2) a MAGIC population derived from eight lines whose genomes were resequenced and (3) a GWAs experiment based on a core collection. QTL were mapped for fruit quality and agronomic traits and their locations and effects were compared. Finally we discuss and compare these populations for QTL mapping and characterization in the new genome era.

2. Materials and methods

2.1. RIL mapping population

A population of 124 F7 recombinant inbred lines (RIL) was developed from the intraspecific cross of two inbred lines Cervil and Levovil as described in [34]. Cervil is a cherry tomato (*Solanum lycopersicum* var. *cerasiforme*) with small fruits (6–10 g) and high aroma intensity, whereas Levovil (*S. lycopersicum*) has much larger fruits (90–160 g) with common taste. In 1996, the RIL were phenotyped for plant and fruit quality traits in a fully randomized trial in a greenhouse at Chateaurenard in Southern France. Plant traits were flowering date of the first flower on the third truss (FLW) and height of the 6th truss on plant stem (HT). The quality traits measured on red fruits were: fresh weight (FW), firmness (FIR), external color (COB, corresponding to the b parameter – blue to yellow – of L, a*, b* parameters), soluble solids content (SSC), pH and titratable acidity (TA), as detailed in [34].

2.2. Genetic data and mapping in RIL

Following the resequencing of the parental lines [20], 754 polymorphic markers were genotyped on the progeny: 679 SNP from parent re-sequencing, 2 RAPD (random amplified polymorphic DNA) and 73 RFLP (restriction fragment length polymorphism) mapped in the previous genetic map from this progeny [22]. The average rate of missing data per marker was estimated at 3% while 98% of the markers passed the Chi-square test ($\alpha = 0.0001\%$). Markers with significant segregation distortion were excluded. Linkage analysis was performed using JoinMap 4.1 [35]. The 12 linkage groups (LG) corresponding to the 12 chromosomes of the tomato genome were built with a grouping logarithm of odds (LOD)-threshold of 4.0, except LG05 for which the grouping threshold was lowered to 3.0. The regression-mapping algorithm was used to order markers within each LG. Genetic distances between markers were calculated using the Haldane mapping function. When several markers colocalized, only the one with the lower rate of missing data was conserved.

2.3. QTL detection in RIL

Quantitative trait loci detection was performed by simple interval mapping [36] using the expectation maximization (EM) algorithm method implemented in R/QTL package [37]. A $\log_{10}$ transformation was applied to FW, FIR and COB as trait distributions deviated from normality. A 1000-permutation test was performed to estimate significant threshold. The LOD threshold was 2.76, corresponding to a genome-wise significance level of $\alpha = 0.10$. For each detected QTL, position, LOD score, confidence interval (CI – for a decrease in the LOD score of one unit), average phenotypic values of the two parental alleles and percentage of phenotypic variation explained (PVE) were displayed. The genetic-CIs were translated into physical intervals (Physical-CI) onto the tomato genome (assembly 2.4).

2.4. MAGIC population

The MAGIC population (397 lines) was obtained by crossing eight tomato lines (including the two parents of the RIL population), selected to include a wide range of genetic diversity of the species as described in [16]. The population was grown in two locations in the South of France in Avignon (location INRA) and La Costière (location VCo). In each location, the 397 lines (one plant per line) and five replicates of each founder were grown in greenhouses during spring-summer 2012, as described in [16]. The traits measured were truss height at second truss (HT), flowering date at third truss

(FLW) and fruit quality traits. Red-fruit quality traits were fresh weight (FW), firmness (FIR), external color (COB), soluble solids content (SSC), pH and titratable acidity (TA) as in RILs. Traits were evaluated from a minimum of five ripe fruits per genotype, collected during two different harvests from truss two to six. A $\log_{10}$ transformation was carried out to normalize the data for FW and HT.

2.5. Genetic data, mapping and QTL detection in MAGIC population

The whole genomes of the eight founder lines were re-sequenced allowing the identification of more than 4 million SNPs [20]. Polymorphism information was used to select a subset of 1486 markers specially designed to analyze the MAGIC population. The selected markers were employed to develop a saturated map [16]. Briefly, genetic distances were estimated with the '*computemap*' *mpMap* function, using a 15-marker window and Haldane distances were computed. This map and the genotype of the parental lines were used to predict the haplotype origin of each locus along the MAGIC population lines genome (using '*mpprob*' function from *mpMap*) [16]. Based on this information QTLs were detected by Simple Interval Mapping using the *R/MpMap* package [38]. QTLs were called when *p*-values were smaller than the empirical threshold *p*-value (1.31×10^{-4}) derived after computing 1000 permutations, to reflect a genome-wide significance threshold of 0.05. When the QTL profile showed more than one QTL peak per chromosome, multiple QTLs were considered significant when peaks were separated by more than 20 cM and the LOD score dropped by more than one. In order to compare QTLs detected in several populations, a second less stringent QTL detection was performed, where QTLs were called when *p*-values were smaller than 10^{-3} corresponding to a LOD 3 value. As in the RIL population, QTL support intervals (SI) were determined with a 1-LOD drop support and translated in physical intervals (physical-SI) on the tomato genome (assembly 2.4).

Recombination events were imputed at locations where the parent/founder allele/haplotype changed along the chromosome in the RIL/MAGIC lines. To calculate haplotype size, recombination locations were translated to physical positions according to their position on the tomato genome (assembly 2.4). Then, haplotypes were defined as the blocks between the beginning/end of the chromosome and the closer recombination, or two consecutive recombinations. Using the haplotype predictions along the MAGIC line genomes, we performed a joint Wald-test for the significance of all founder effects at putative QTL positions.

2.6. GWAs panel

The tomato diversity panel consisted of 163 accessions composed of 28 *S. lycopersicum* (S.L.), 119 *S. lycopersicum* cv. *cerasiforme* (S.C.) and 16 *S. pimpinellifolium* (S.P.) as described in [32], and [39]. Plants (four replicates) were grown in plastic tunnel in Avignon, France during summers 2007 and 2008. At least 10 fruits per plot were measured for the same traits, as in the RIL and MAGIC populations, as described in [34,39], except for flowering time and plant height. Phenotypic data collected in 2007 and 2008 were averaged over the two years and $\log_{10}$ transformed when the Shapiro–Wilk test evidenced a non-normal distribution.

2.7. Genetic data and GWAs analysis

Genotyping was performed using the Infinium assay (Illumina Inc.), developed by the Solanaceae Coordinated Agricultural Project (SolCAP) [40]. After filtering for quality, missing data and low allele frequency, a set of 5995 SNP markers remained. Briefly, pairwise kinship coefficients (*K* matrix) were estimated using the Ritland formula implemented in SPAGeDI [41]. For population stratification, the most likely number of clusters K in all simulations were assumed to be in the range of $K = 1$ to $K = 10$. Ten replicates were conducted in the structure software [42] for each *K* with a burn-in period of 1×10^6, followed by 5×10^6 Markov chain Monte Carlo (MCMC) steps. Then, the Evanno correction was applied [43]. GWA analyses were performed with correction for population structure (Q for FIR, pH and SSC or PCoA for FW, COB and TA) and modeling phenotypic covariance with the kinship (*K*) matrix. These matrices were implemented into a modified version of the multi-locus mixed model (MLMM) described in [44]. The analysis followed the same steps as in [32]. Levels of significance were assessed according to [44].

2.8. Comparison of QTLs and screen for candidate genes and polymorphisms

We projected all markers on the SL2.40 genome sequence and thus mapped all the QTLs/associations on the same framework. We compared QTL SI and decided that a single QTL was present when the SI overlapped or when an association lay in the SI.

To screen for candidate genes and polymorphisms, we selected QTLs from the MAGIC population with SI lower than 1 Mb and listed all the polymorphisms detected among the parental lines in the interval. Then based on founder allelic effects at the QTLs, we identified two successive conditions (pairs of lines with identical or different alleles at the QTL) and listed polymorphisms corresponding to the conditions.

2.9. Data availability

Input RIL data (genotypes and phenotypes) are provided as Supplemental data S1. MAGIC map details and genotype data are available in [16]. MAGIC phenotypes are provided in Supplemental data S2. MAGIC SNPs, polymorphisms and QTLs are deposited on the GNPis repository hosted at https://urgi.versailles.inra.fr/gnpis [45]. Genome-wide association input data (K and Q matrices as well as genotypes and phenotypic data) and results are deposited on the GNPis repository hosted at https://urgi.versailles.inra.fr/association

3. Results

3.1. QTLs in RIL population

3.1.1. A saturated map of the intraspecific RIL population

After resequencing the two parental lines of the RIL population, the new genetic map constructed with SNP markers included 501 distinct loci covering 1090 cM. The average number of markers per chromosome was 42, with an average distance between markers of 2.60 cM. The map covered 98% of the assembled tomato genome, against 70% for the genetic map obtained earlier with the same progeny [22]. In particular, coverage of chromosomes 1, 7 and 8 was improved from 9% to 99%, 15% to 99% and 30% to 99%, respectively (Supplemental data S3). This map was then used to map QTLs with the phenotype data earlier described [34] and unpublished data for FLW and HT.

3.1.2. QTLs in the RIL population

In the RIL population, 25 QTLs were detected for eight traits, explaining 8 (*flw5.1*) to 36% (*fir4.1*) of the phenotypic variation (Table 1 and Supplemental data S4). The percentage of variation explained per trait ranged from 20 (for pH) to 67% (for FW). Several clusters of QTLs were identified, particularly on chromosomes 2, 4 and 9. Most of the QTLs detected earlier [34], using the same phenotypic data and a genetic map with a lower coverage rate, were confirmed. The TA and pH QTLs, on chromosome 12, were no longer

Appendix 3

Table 1
Characteristics of QTLs and associations detected in the 3 populations. The number of individuals and markers per population are indicated. The total number of QTLs and the number of QTLs per trait are indicated, together with the percentage of variation explained (PVE) and the minimum and maximum support intervals (CI or SI, and distance of LD decay for GWAs). The numbers of QTL common to two or three populations are indicated.

	RIL			MAGIC			GWAs			Common QTL RIL/MAGIC	Common QTL MAGIC/GWAs; RIL/GWAS	Common to 3 pops
Nb genotypes	140			397			163					
Nb markers	500			1500			5500					
Map size (cM)	1200			2100			–					
QTL number	25			63			28			16	8; 5	4
	nb QTL	PVE	Min–max CI (Mb)	nb QTL[a,b]	PVE[c,d]	Min–max SI (Mb)	nb QTL	PVE	Min–max distance of LD decay (kb)			
Fruit weight	5	0.67	1.38–48.41	11 (2, 3)	0.51–0.34	0.95–57.33	9	0.80	1.1–4097	4	3; 2	2 (*fw3.2*; *fw12.1*)
Soluble Solid content	2	0.36	1.62–4.74	7 (1, 0)	0.41–0.13	1.10–3.39	9	0.55	1.3–269.7	2	3; 2	2 (*ssc2.1*; *ssc9.1*)
Titratable acidity	4 + 1	0.59	0.89–46.14	4 (0, 0)	0.36–NA	1.89–4.16	6	0.47	3.7–152.9	2	0; 1	0
pH	1 + 1	0.20	1.09–1.86	9 (5, 0)	0.48–0.32	0.90–9.29	2	0.24	11.4–181.1	1	2; 0	0
Firmness	3	0.51	1.03–11.46	8 (2, 2)	0.48–0.42	0.50–53.19	1	0.43	96.3	2	0; 0	0
Color-b	3	0.59	1.24–8.92	8 (2, 2)	0.43–NA	1.04–59.37	1	0.43	116.8	2	0; 0	0
Flowering T3	4	0.40	2.24–62.21	9 (2, 0)	0.45–0.33	0.77–51.75				2		
Plant height	1	0.35	10.63	7 (2, 1)	0.33–0.16	0.34–51.72				1		

[a] Number of QTLs just below the threshold.
[b] Number of QTLs detected as 2nd peak.
[c] PVE for INRA location.
[d] PVE for VCo location.

detected, because of a large gap without any marker. Two new QTLs were detected on chromosomes 1 and 4, for fruit firmness (*fir1.1*) and fruit fresh weight (*fw4.1*), respectively. They were located in genomic regions weakly covered by the previous map.

3.2. QTLs in the MAGIC population

3.2.1. QTL mapping in the MAGIC population

The MAGIC linkage map is composed of 1345 SNP markers, covering 758 Mb (84% of the 900 Mb tomato genome size, and almost all the 760 Mb assembled genome [26]) and 2156 cM. It thus more than doubled the map size compared to the RIL population. We could predict the haplotype origin for an average of 89% of the MAGIC line genomes [16].

A total of 63 QTLs (corresponding to 78 QTLs over the two locations) were detected for the eight traits (Table 1 and Supplemental data S5), with four to eleven QTLs per trait. The PVE per trait ranged from 13% (for SSC in VCo) to 51% (for FW in INRA location). Lower PVE in VCo compared to INRA location was due to more homogeneous growing conditions in INRA trial. For the six traits assessed in the two locations, 15 QTLs were detected in both locations, while nine were specific of VCo and 27 of INRA location. Support intervals ranged from 5.5 (*ta6.1*) to 86 cM (*cob9.2*) and from 340 kb (*ht4.1*) to 64 Mb (*flw1.1*). Segregation of different QTLs according to the founders lead to variation in allelic effects according to the QTLs. Allele effects according to the parental line are detailed in Supplemental data S5 and illustrated for FLW in Fig. 1.

If we combine the support intervals of QTLs detected at two locations to limit the interval boundaries, nine QTLs were mapped in an interval close to or less than 1 Mb. Table 2 illustrates for these QTLs the number of genes and polymorphisms detected in the regions and the number of mutations with an effect on the protein sequence. Several thousands of polymorphisms were frequently detected, but the number with an effect on the protein was much lower. By assessing the allele effect of the eight founder lines at the QTLs, it was possible to determine combinations of parental alleles that should be similar or different. This strongly reduced the number of candidate polymorphisms. Polymorphisms with an effect on the protein sequence were detected for four QTLs (for FIR and SSC) providing a short list of candidate genes to be further studied. In some cases, we could not find any polymorphism corresponding to the condition. This could be due to missing or ambiguous sequence data or the causal variant may be due to a long Indel (not detected) or a copy number variants (CNV). The analysis of the founder genome sequences revealed several regions with CNV [20] covering 35 Mb (around 4.4% of the genome). Copy number variants were detected in at least five of the nine regions scanned in Table 2. Epigenetic modifications could also account for the QTL variation as shown in the case of the CNR gene variant [46].

3.2.2. Haplotypes, recombination and linkage disequilibrium

The LD, recombination rate and haplotype sizes will determine the power to detect genetic associations. In the MAGIC population LD decayed quickly from an average of 0.47 at 1 kb to less than 0.2 at 2 Mb, reaching a minimum of 0.08 at 20 Mb. However, for more distant markers (40 Mb), LD increased again (higher than 0.13) to fall again to previous values at distances around 50 Mb [16]. This is caused by the large centromeric regions with low recombination rate in the tomato genome that comprise around 70% of the chromosomes [30]. In natural populations used for GWAs studies LD is lower especially in the centromeric regions and baseline is reached before 50 cM [31].

Higher apparent recombination rates in the MAGIC population reduced haplotype sizes and conferred greater precision to QTL detection when compared with RIL population. The MAGIC genetic map (2156 cM) is 97.8% longer than the RIL map (1090 cM). This

Appendix 3

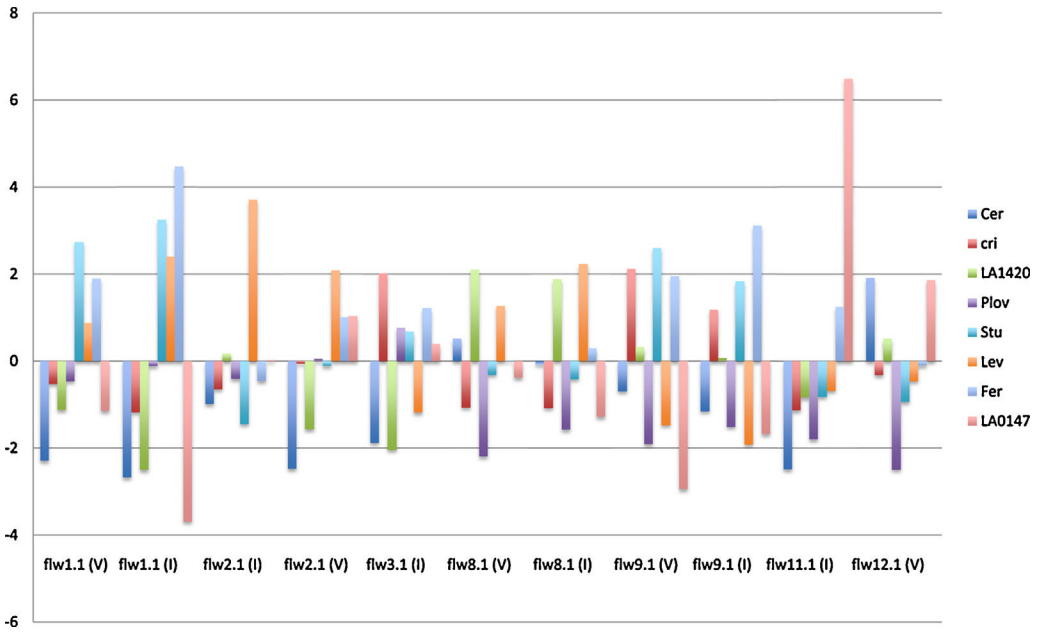


Fig. 1. Founder allelic effects at the flowering time QTLs in the MAGIC population. Centered effects for the eight parental alleles (Cervil, Levovil, Criollo, Stupicke PR, Plovdiv24A, LA1420, Ferum and LA0147, from left to right). The QTL name and location of the trial (Inra, I or VCo, V) is indicated below the QTL.

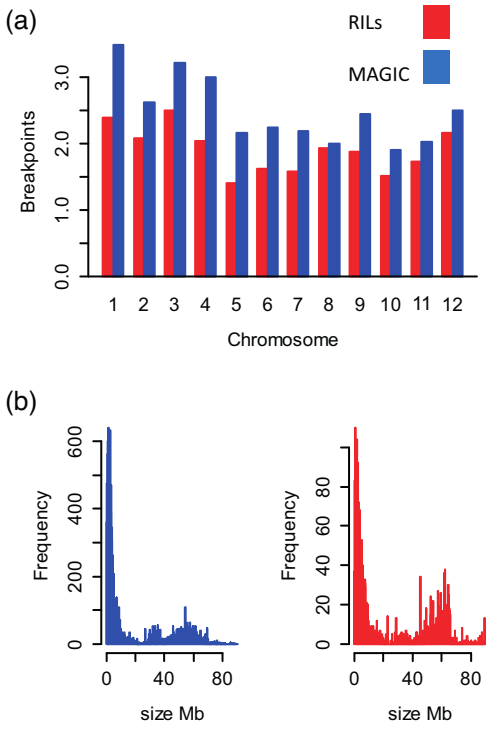


Fig. 2. Comparison of recombination in RIL and MAGIC populations. (a) Distribution of average number of break points per chromosome in RILs and MAGIC lines; (b) distribution of the haplotype block size in RILs and MAGIC lines.

is consistent with the average number of break points per line and chromosome, 2.49 and 1.46 in MAGIC and RIL population, respectively (Fig. 2a). This increase was not that obvious when we compared the haplotype block physical size, which was just 33% higher in the RIL (22.33 Mb) than in the MAGIC population (16.77 Mb). However the higher recombination rates in the MAGIC population clearly reduced the frequency of long haplotype blocks, corresponding to centromeric regions (Fig. 2b).

3.3. GWAs

3.3.1. Linkage disequilibrium, kinship and population stratification

Briefly, within the GWA experiment, we took benefit from the most recent development achieved in the estimation of LD decay by using the *LDcorSV* measurement that takes into account kinship and population stratification in the studied population [47]. Based on this measurement, the average intra-chromosomal LD estimations ranged from 0.337 in *S. pimpinellifolium* to 0.567 in *S. lycopersicum*. This demonstrated that selection tends to increase LD level especially in cultivated accessions. As earlier reported [32], the average degree of relatedness was low with an average value of 0.074 while the number of ancestral populations was estimated to be two ($K = 2$).

3.3.2. Phenotype–genotype associations

A total of 41 associations were detected for the six traits also measured in the RIL and MAGIC populations (Table 1). The number of associated loci ranged from one (for COB and FIR) to nine (for FW and SSC). In terms of genomic location, chromosomes carried varying numbers of associated SNP with chromosome 7 carrying only one association (i.e. SSC) while up to five associations were detected onto chromosome 2 for FW, SSC and TA. The estimated heritability (estimated at step 0 of the model, based on the variance component σ^2_g computed for all markers and g representing the estimated genetic variance of the trait) and PVE ranged from 0.42 (FIR) to 0.88 (FW) and from 0.24 (pH) to 0.80 (FW), respectively. Detailed information regarding these results, such as peak SNP annotation, is reported in Supplemental data S6 and S7.

3.4. Comparison of QTLs across populations and candidate gene identification

Table 1 presents the total number of QTLs per population and QTLs detected in overlapping intervals across populations. For the six common traits, 17 of 95 QTLs were detected in at least two populations. Fig. 3 summarizes the QTLs detected in the three populations. A few chromosome regions present clusters of QTLs,

Appendix 3

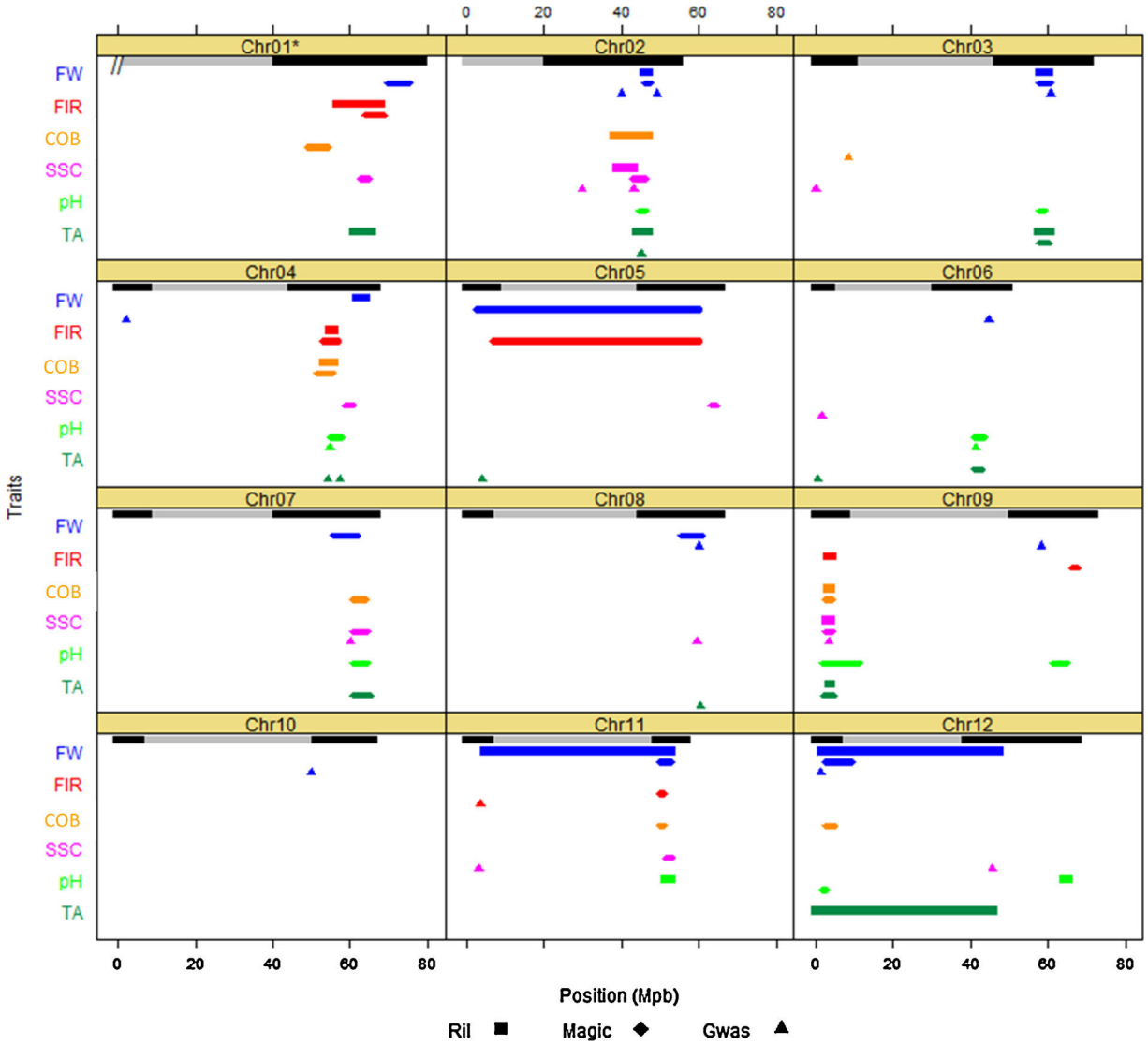


Fig. 3. Overview of fruit quality QTLs identified on the tomato genome by mapping analysis in RIL, MAGIC and GWA populations. At the top of the twelve panels, lines proportional to chromosome physical size in million base pairs (Mb) represent tomato chromosomes. Chromosome 1 is truncated of the first 20 Mbp for representation comfort (marked by // and *). Centromeric parts with low recombination frequency are indicated in gray and peripheral parts in black (according to [30]). QTLs are represented by square, diamond and triangle symbols in the RIL, MAGIC and GWAs populations, respectively. Color codes correspond to the six fruit quality traits: fresh weight (FW) in blue, firmness (FIR) in red, b color parameter (COB) in orange, soluble solids content (SSC) in pink, pH in light green and titrable acidity (TA) in dark green. For the MAGIC population, only QTLs identified by simple interval mapping are represented. Besides, when a QTL was found in two locations, only the one with the shorter confidence interval is represented.

particularly on chromosomes 2, 4, 7, 9 and 11. We then reviewed the traits considering a single QTL when SI overlapped. For FW, a total of 20 different QTLs were detected. The two previously cloned QTLs, *fw2.2* [24] and *fw3.2* [33] were detected in the three populations, confirming their major role in the difference between cherry and large-fruited tomato accessions. The QTL *fw11.2* was detected in both RIL and MAGIC populations and probably corresponds to a QTL close to the fasciated (*fas*) locus, which has been fine mapped to 149 kb [48]. On chromosome 2, several linked QTLs seem to be present in a small region as already showed [16] and this region should be precisely dissected as it contains many QTLs and major genes for fruit size, shape and sugar composition [49].

For SSC, 13 QTLs were detected. The locus *ssc9.1* was detected in the three populations. It likely corresponds to a previously cloned QTL (*Brix9.2.5*) exhibiting a polymorphism in a cell wall invertase gene (*lin5* [25]). Another QTL was detected on chromosome 2 in RIL and GWA panel around 42 Mb, and it seems linked to another QTL (*ssc2.2*) detected in the MAGIC population with a peak around 45 Mb. The smallest support interval in the MAGIC population concerned *ssc2.2*, which covered 5.8 cM and 430 kb. When looking at the allelic effect of the founders, we poorly reduced the list of candidate polymorphisms (1368 in 46 genes) as the allelic effects of the parents corresponded to the major haplotype in the region. A total of 24 polymorphisms had an effect on the coding sequence in 12 genes but we could not identify any specific candidate gene based on their annotation.

Six QTLs revealed colocations for TA and pH, which are assumed to be related traits. Thus, we considered the two traits as a single one. A total of 14 QTLs were detected, with five in at least two out of the three populations analyzed. The strongest effects concerned *ta9.1, ta3.1* (in RIL and MAGIC), *ph6.1* (in MAGIC and GWA populations) and the association on chromosome 2 (position around 45 Mb which colocalized with a QTL for TA in RIL and pH in MAGIC populations). The association on chromosome 6 could be related with that detected on the same panel for citrate (position 41,345,468) in the close vicinity of two malate transporters previously identified [32].

Table 2
Screening for candidate SNP in the MAGIC population. Number of possible causal polymorphisms underlying the MAGIC QTLs with SI smaller or around 1 Mb. QTL start and end are the SI limits translated to physical bp, calculated by LOD decrease of one. Nb genes and nb pol correspond to the number of genes and polymorphisms in the SI without selection, after selection based on founder effects (as described by the different contrasts according to allelic effects), in the SI and nb pol high/moderate the number of polymorphisms with an effect on the protein sequence.

QTL	Intervals			Total polymorphisms			Selection on allelic effect						
	Start	End	Size (bp)	nb genes	nb pol	nb pol high/moderate	cond1	Nb cond1	cond2	Nb cond2	nb both conditions	in nb genes	Nb pol high/moderate
fir1.1	86,982,244	87,980,000	997,756	126	3998	224 (in 91 genes)	Cer ≠ Cri	2615	Cri = Plo	2190	1400	61 genes	71 (in 30 genes)
fir4.1	54,840,000	55,870,000	1,030,000	70	3173	171 (in 60 genes)	Fer ≠ Lev	1333	Plo = Lev	952	109	31 genes	4 (in 4 genes)
fir11.1	49,998,845	50,507,387	508,542	64	2374	133 (in 41 genes)	Fer = La0	2350	Cer = Cri	168	157	54 genes	5 (in 3 genes)
ssc2.2	43,049,064	43,480,000	430,936	46	1763	63 (in 31 genes)	Cri = Lev	1707	Cer ≠ Lev	1404	1368	46 genes	24 (in 12 genes)
ph6.1	41,150,000	42,150,000	1,000,000	129	3803	153 (in 65 genes)	Lev ≠ La0	5	La0 = Plo	3711	1	1 gene	0
ph12.1	1,921,369	2,821,857	900,488	177	4236	197 (in 105 genes)	Fer ≠ La01	0	La0 = Plo	841	0	0	0
flw9.1	64,036,532	65,053,302	1,016,770	108	4623	206 (in 76 genes)	Stu ≠ La0	82	Cri = La0	1814	27	2 genes	0
flw11.1	50,290,034	51,066,236	776,202	87	3276	107 (in 48 genes)	La0 ≠ Plo	17	Plo = Cer	240	3	2 genes	0
ht4.1	62,567,262	62,907,824	340,562	28	584	40 (in 16 genes)	LA0 ≠ Stu	0	–				

For fruit firmness, seven QTLs were detected, among which two were detected in RIL and MAGIC populations on chromosomes 1 and 4. The QTL *fir11.1*, which colocalized with *cob11.1*, had the smallest support interval (0.5 Mb). The selection based on allelic effect of founders allowed reducing the list of candidate polymorphisms from 2374 in 64 genes to 157 in 54 genes. Three genes presented polymorphisms with coding effects (Table 2). Among them a vacuolar sorting protein (Solyc11g067230) also showed a strong correlation ($r = -0.81$; $p = 0.002$) between its expression in growing fruits of parental lines and their allele effect at the QTL (unpublished data). Again, for fruit firmness, GWA signal and QTL overlapped especially on chromosome 11 supporting our results.

For COB, eight QTLs were detected in total, among which two were detected both in RIL and MAGIC populations on chromosome 4 and 9, respectively, but with large confidence intervals, while GWAs detected other associations on chromosome 3.

FLW and HT were assessed only in the RIL and MAGIC populations.

For FLW, 11 QTLs were detected, with two common to both populations on chromosomes 2 and 12. In the MAGIC population, taking advantage of the allelic effects of the QTLs with the smallest support interval allowed identifying 3 and 27 candidate polymorphisms for *flw11.1* and *flw9.1*, respectively (Table 2). For HT, seven QTLs were detected with one common to both populations on chromosome 6. The QTL *ht4.1* had a small support interval, carrying only 584 polymorphisms. None of them corresponded to the allelic pattern of the founders at the QTL, suggesting either low coverage, presence of an undetected large Indel, missing data or epigenetic effect. The QTL on chromosome 3 for these two traits could be related to a pleiotropic effect of *fw3.2* as this QTL was shown to affect also earliness and plant vigor [33].

4. Discussion

4.1. Common QTLs in the three populations

The results describe QTLs and associations detected for fruit and plant traits in three panels of (or derived from crosses between) large fruited and cherry tomato accessions. Using three different panels, we detected 71 QTLs for the six traits evaluated in the three populations, among which 17 were at least detected in two populations (Table 1). The large proportion of QTLs detected in RILs also detected in the MAGIC population was expected as the two parents of the first population were among the parents of the second. On the contrary in the GWAs panel a larger set of QTLs may segregate explaining the number of differences (32% of associations mapped in the support interval of a QTL). We must underline that we supposed that overlapping support interval corresponded to a single QTL although only fine mapping experiments could prove that two linked QTLs do not segregate as sometimes shown after fine mapping experiments [49,50].

4.2. The benefit from the genome sequence

For the first time, the availability of the reference tomato genome sequence [26] allowed the projection of the QTLs and their support intervals onto the physical map of the tomato. This allows comparing QTL positions even in populations with maps constructed with different marker sets. Several clusters and most of the QTLs fall in regions where a recent diversity study based on 360 resequenced accessions [28] identified selective sweeps due to the rise in frequency of favorable haplotypes and leading to a drastic reduction of the nucleotide diversity when comparing cherry accessions to large fruited lines. Although these 133 regions only cover 7% (54.5 Mb) of the assembled genome, 52% (33 of 63) of the

QTLs detected in the MAGIC population have their support interval in one of these regions. In the future, a special attention should be brought to these regions as they contain important genes for breeding.

The uneven distribution of crossovers in tomato, with large chunks of chromosome around centromeres which almost do not recombine, leads to a few QTLs encompassing more than 50 Mbp. Luckily such QTLs are not frequent as they represent less than 10% of the QTL. Fewer genes are present in these regions but the low recombination frequency hampers their use in tomato breeding.

The availability of a high quality reference genome sequence and the development of next generation sequencing technologies [51,52] eased the resequencing the genomes of the parental lines of the RIL and MAGIC populations and the discovery of more than 4 million polymorphisms among the eight founders with a very high level of accuracy [20]. Thus, combining founder allelic effects and SNP catalogs reduced the number of candidate polymorphisms and allowed targeting candidate genes or regions more precisely than ever. Furthermore, polymorphisms with an effect on the coding sequences are quite rare, but many QTLs have been discovered in non-coding sequences [53].

4.3. Interests and limitations of the three population types

Table 3 summarizes the main pro and cons of the use of each type of panel.

4.3.1. RIL and MAGIC populations

Briefly, RIL population are easy to set up and to analyze. They are interesting for mapping rare alleles such as disease resistance genes or other specific traits, but they lead to large support intervals due to the low recombination while the genetic background effect (epistasis) may hamper QTL detection. MAGIC populations are more complex to set up and to analyze and population size need to be much larger (at least 50 individuals per founder [17]). However, MAGIC allows better detection as a larger set of QTLs segregate among the population founders (63 QTLs *vs* 25 in RIL in our experiments performed on the same traits). Several methodologies were proposed to analyze such populations. On RILs, Composite Interval Mapping provided results very similar to Simple Interval Mapping (data not shown). For the MAGIC data, we used a regression of phenotypic values on the predicted haplotypes of the lines, as the percentage of prediction was high. We could have used the SNP alleles as in GWAs approaches [14], or intermediate approach grouping the haplotypes [18]. These approaches may lead to small differences as shown for FW in tomato [16]. One of the main interests in detecting QTLs in MAGIC population is the dissection of allelic effects when the founder genomes are sequenced. This is useful for combining positive alleles through marker-assisted selection or genomic selection and for QTL identification. In the MAGIC population, support intervals are smaller in average than in RILs (from 10 to 7 Mb on average, in our experiment). Successive inter-crossing before selfing generations could even increase the apparent growth in recombination [12,50]. The increase in recombination reduces support interval size and subsequently reduces the number of candidate genes or polymorphisms to be studied. If the population is large enough (e.g. about 1000 *Arabidopsis* accessions) then the QTLs can be directly identified [12].

4.3.2. Genome-wide association study

Genome-wide association scans have the potential to detect more precisely loci underlying the variation of traits due to the high density of markers and the rapid decay of LD even though the studied population is stratified. Thus, this approach is complementary to linkage-based approaches (either the population is bi-parental or multi-parental). In the present study, a total of 28 associations were identified for the six studied traits explaining varying part of the variance. This confirmed the polygenic architecture of the traits with a large number of small effect loci that we were not able to detect. Stringent threshold, lack of statistical significance for the control of false negatives caused by small effect sizes [8] may explain these results. The estimations of the missing heritability (Supplemental data S6) tend to support our observations and means that small to medium effect loci remain to be identified. However, for the fruit weight, the associated loci explained 80% of the variation suggesting strong effect QTLs, but the missing heritability remains high. Furthermore, identifying associated loci that have been previously cloned (*lin5*; *fw2.2*) validates our methodological approach.

When examining the annotation of the peak SNPs related to the traits (Supplemental data S6), few of these seemed to be directly related to the traits they are associated to. This means that peak SNPs are in LD with the candidate gene or polymorphism. Defining the shortest physical distance that contains the candidate gene is much more complex in GWAs than in classical linkage populations, where methods such as 1-LOD support interval or bootstrapping are commonly used to assess QTL confidence interval [36] followed by polymorphism examination. As the number of SNPs is limited (5595), we examined LD decay around each peak SNP to define a 'LD bin' in which looking for putative candidate genes provided by the tomato genome annotation. In our case, when considering an arbitrary LD decay of 0.2 around the peak SNP, the estimated length of the LD bins were ranging from 1.1 kb to 4.1 Mb with a median value of 57 kb, reflecting the different degree of LD decay in the tomato genome. Thus, looking for candidate loci in a LD bin may be time consuming or nearly impossible as hundreds of genes may be included within the same LD bin [54]. To circumvent this inherent problem in GWAs, two approaches have been proposed. The first one tends to predict a minimal genomic region around a genetic association signal within a LD bin with a high degree of accuracy by observing around an association signal LD between polymorphic markers that is known to be stronger in cases compared to controls [55]. The second approach tends to recover power in regions of high LD by whether estimating the kinship with all the markers that are not located on the same chromosome as the tested SNP or taking into account the correlation between markers to weight the contribution of each marker to the kinship [56]. Thus, as previously stated [57], the method chosen to define an associated chromosomal region influences GWAs reliability and this issue remains under investigated.

Regarding the design of the GWA study, much improvement could be achieved, especially through (1) optimizing the panel population by choosing the individuals on the basis of their relatedness to maximize its reliability, and (2) increasing the SNP density. In the context of genomic selection, an approach that discriminates which individuals must be included in the calibration set was proposed [58]. Applying this approach to define the optimal association panel would be worth testing. In parallel, increasing the number of markers would definitely help to detect more associations and reduce the missing heritability part. However, this statement largely relies onto the LD patterns of the species. In the cultivated tomato, several studies reported that LD decays over large genomic regions (up to several Mb [31,32,39,59]) limiting the interest of high-density SNP arrays, in addition to the ascertainment bias introduced by the use of SNP arrays. Increasing the SNP density would be of interest in recombining regions of the genome. However, larger set of SNP would imply more stringent threshold due to correction for multiple testing. To overcome this limitation, one way is to test for genotype–phenotype associations using haplotypes (blocks of LD) rather than single markers [60]. Thanks to NGS and imputation methods, the use of haplotypes has already been tested [61] and applied [62] in plants, demonstrating its increasing interest.

Appendix 3

Table 3
Comparison of advantages and limitations of the RIL, MAGIC and GWA populations.

	RIL	MAGIC	GWA
Time to develop	Intermediate	Long	Short
Precision in mapping common alleles	+	++	+++
Precision in mapping rare alleles	+++	+	–
Access to recombination	+	++	+++
Nb of markers needed	+	++	+++
Population structure	No	No	+
Main advantages	Rare allele mapping Easy analysis	Several alleles segregating Founder allele effect for MAS and QTL identification	Precision due to historical recombination
Main limitations	Large QTL confidence intervals	Time to set up Large population needed	High LD limits the precision Pop structure responsible for false positive

4.3.3. Combining populations to close the genotype–phenotype gap

Overall, the interest of this study relies on combining results from linkage mapping experiments in RIL and MAGIC populations with GWA analysis to decipher the genetic architecture of traits related to the fruit quality in tomato. The combination of the populations seems efficient as not all the loci affecting these complex traits are expected to be detected in a single population because of allele specific effects. Such combination may reinforce QTL relevance and restrict the support interval for their characterization [63], as association signal as well as QTL support interval may span over large genomic regions (see above) directly reflecting the patterns of LD decay.

From a wider point of view, closing the gap between the genotype and the phenotype is not a recent idea. The pioneer work of Mitchell-Olds on *Arabidopsis* [64] clearly demonstrate the interest of combining quantitative genetics and population genetics to decipher the genetic architecture of adaptive traits and solve the "genotype-to-phenotype problem". In the same ecological context, Stinchcombe and Hoekstra [65] reviewed the advances of this combination demonstrating its power to identify candidate genes. More recently, Mitchell-Olds [66] dissected two studies in *Arabidopsis* related to flowering traits, respectively based on population [67] and pedigree (MAGIC) [12]. He suggested that an increasing number of small effect loci will be detected but also that the combination of pedigree and natural populations will elucidate the patterns of trait variation. In addition, the combination of quantitative and population genetics makes sense as breeding system, effective population size, selective history and population demography influence the genetic architecture of traits, as illustrated in *Arabidopsis* to detect genes associated to the resistance to the PPV virus [68].

4.4. Prospects: QTL characterization in genome and resequencing era

Nowadays genome sequences and next generation sequencing technologies provide a number of changes in QTL detection strategies: a number of wild tomato relatives and cultivated accessions have been resequenced [27–29]. Polymorphism discovery is no longer limiting and Genotyping-by-Sequencing [69] may allow the rapid discovery of SNPs necessary for the construction of new genetic maps at a reduced cost. It is thus possible to map new QTLs at the intraspecific level for traits differing in cultivated accessions without the large effect of major QTLs, which distinguish wild from cultivated accessions. For example, among the eight founders of the MAGIC population, the cherry tomato accessions differ from the reference genome by 1–2 million SNPs, while the four large fruited lines only differ by 180,000–350,000 SNPs. The availability of the catalog of polymorphisms among parental lines also considerably facilitated the fine mapping of candidate genes. If two lines only differ in a small number of large effect QTLs, then Bulk Segregant Analysis can be combined to NGS to speed up gene discovery. This has been shown efficient to directly identify the polymorphism responsible of major mutations [70,71] but also used for QTL mapping as illustrated in tomato for FW [28].

Genome information is also important to compare and identify QTLs. We have shown that physical positions of markers allow the projection of QTLs on a reference map independent of the progeny. It is thus possible to perform meta-analysis on a large number of studies using dedicated software [72,73] and thus reducing the support interval around the QTLs [45]. Thus, managing all the phenotype data in a common database is highly important today [74]. In tomato, the Sol Genomics Network [75] (http://solgenomics.net/) concentrates Solanaceae genome sequences, polymorphisms as well as a few QTLs and phenotypic data. Genome annotation provides gene catalog under the QTLs. Thus high quality gene annotation is also strongly needed, as a large percentage of genes are still with unknown function. Finally transcriptome (RNAseq) data on several organs, developmental stages and genotypes [76] also provide cues for the identification of candidate genes. The identification of candidate genes underlying a QTL relies on a set of arguments related to gene location, function, expression and polymorphism. When a candidate gene has expression variation linked to the phenotype in parental lines, eQTLs can be mapped and the colocation of a trait QTL, a related gene and its eQTL may confirm that a polymorphism close to the gene is responsible for its variation and putatively responsible for trait variation [77]. The validation of such guilty-by-association candidate polymorphism may not be easy by traditional transgenic approaches as knockout or overexpressing a gene may have an effect on the phenotype even though it is not the QTL. Today, the genome editing technologies make possible to precisely perform genome modifications in plants and thus validate a specific polymorphism [78]. Screening and characterizing mutants in a candidate gene in Tilling populations is another way to validate the effect of a candidate gene [79].

Taken together, all these results illustrate that finding the genes underlying the phenotype of interest is only feasible in species for which genetic information is abundant and even in model organism, such as tomato, the ability to move from QTL to QTN is still not that easy [80]. However, compared to the decade needed to clone the first QTL responsible for fruit weight (*fw2.2*) in tomato [24], biologists gained the power to prove that a variant is responsible for the trait variation with a much larger variety of genomic tools and experimental designs that speed up the process. Thus, even though the statistical approaches used in QTL and GWAs present some caveats, we are still discovering and understanding new molecular determinants underlying traits of economical and agronomic interest in crops. However, major challenges remain especially toward the understanding of the role of non-coding 'junk' DNA and epigenetic marks [81] onto the regulatory landscape of genomes and the adaptation of crops to their environmental conditions notably in their response to biotic and abiotic stresses.

Appendix 3

5. Conclusion

Genome sequences and NGS technologies provided flood of genomic information such as genetic variants responsible of quantitative traits that we have to manage. But the remaining limitation is no longer genotyping or sequencing but is to properly phenotype in a high throughput and reproducible way. Relevant populations are now of high importance together with the phenotype precision. It is urgent to gather all QTL data in databases in order to be able to perform meta-analyses to decipher the genetic determinants of agronomic traits. Thus, we clearly demonstrated that the combination of QTL analysis (in RIL and MAGIC populations) and GWAs precisely mapped and identified the QTLs and avoided false positives. Combined to data of polymorphisms in large populations and expression profiles we should quickly identify new causal variants responsible for the variation of important traits.

Acknowledgements

We acknowledge the experimental team of INRA GAFL for their help in experimentation, Yolande Carretero, Justine Gricourt, Esther Pelpoir and Renaud Duboscq for their help in phenotyping. We thank the experimental and informatic INRA-EPGV team: Aurélie Bérard, Auréle Chauveau, Rémi Bounon, Maria Tchumakov and Elodie Marquand. This work was supported by CEA-IG/CNG, by performing the DNA QC and providing access to INRA-EPGV to their Illumina Sequencing Platform. We acknowledge groups of Anne Boland (DNA and Cell Bank service) and Marie-Thérèse Bihoreau (Illumina HT Sequencing). The ANR MAGIC-Tom SNP project 09-GENM-109G and the European Solanaceae Integrated Project EUSOL (Food-CT-2006-016214) supported this work. LP was supported by a postdoctoral INRA fellowship, EA by an INRA PhD fellowship and JD by a grant from the Embassy of France in Thailand in Junior Research Fellowship Program 2014.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.plantsci.2004.08.011

References

[1] A.H. Paterson, E.S. Lander, J.D. Hewitt, S. Peterson, S.E. Lincoln, S.D. Tanksley, Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms, Nature 335 (1988) 721–726.
[2] M.J. Kearsey, A.G.L. Farquhar, QTL analysis in plants; where are we now? Heredity (Edinb.) 80 (1998) 137–142.
[3] J.B. Holland, Genetic architecture of complex traits in plants, Curr. Opin. Plant Biol. 10 (2007) 156–161.
[4] Z.B. Lippman, Y. Semel, D. Zamir, An integrated view of quantitative trait variation using tomato interspecific introgression lines, Curr. Opin. Genet. Dev. 17 (2007) 545–552.
[5] A. Price, Believe it or not, QTLs are accurate!, Trends Plant Sci. 11 (2006) 213–216.
[6] A. Korte, A. Farlow, The advantages and limitations of trait analysis with GWAS: a review, Plant Methods 9 (2013) 29.
[7] P.K. Gupta, S. Rustgi, P.L. Kulwal, Linkage disequilibrium and association studies in higher plants: present status and future prospects, Plant Mol. Biol. 57 (2005) 461–485.
[8] P.M. Visscher, M.A. Brown, M.I. McCarthy, J. Yang, Five years of GWAS discovery, Am. J. Hum. Genet. 90 (2012) 7–24.
[9] A.L. Price, N.A. Zaitlen, D. Reich, N. Patterson, New approaches to population stratification in genome-wide association studies, Nat. Rev. Genet. 11 (2010) 459–463.
[10] I. Baxter, J.N. Brazelton, D. Yu, Y.S. Huang, B. Lahner, E. Yakubova, et al., A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1;1, PLoS Genet. 6 (2010) e1001193.
[11] J. Yu, J.B. Holland, M.D. McMullen, E.S. Buckler, Genetic design and statistical power of nested association mapping in maize, Genetics 178 (2008) 539–551.
[12] P.X. Kover, W. Valdar, J. Trakalo, N. Scarcelli, I.M. Ehrenreich, M.D. Purugganan, et al., A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*, PLoS Genet. 5 (2009) e1000551.
[13] N. Bandillo, C. Raghavan, P.A. Muyco, M.A.L. Sevilla, I.T. Lobina, C.J. Dilla-Ermita, et al., Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding, Rice (N.Y.) 6 (2013) 11.
[14] B.E. Huang, A.W. George, K.L. Forrest, A. Kilian, M.J. Hayden, M.K. Morell, et al., A multiparent advanced generation inter-cross population for genetic analysis in wheat, Plant Biotechnol. J. 10 (2012) 826–839.
[15] W. Sannemann, B.E. Huang, B. Mathew, J. Léon, Multi-parent advanced generation inter-cross in barley: high-resolution quantitative trait locus mapping for flowering time as a proof of concept, Mol. Breed. 35 (2015), http://dx.doi.org/10.1007/s11032-015-0284-7
[16] L. Pascual, N. Desplat, B.E. Huang, A. Desgroux, L. Bruguier, J.-P. Bouchet, et al., Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era, Plant Biotechnol. J. (2014), http://dx.doi.org/10.1111/pbi.12282
[17] C. Cavanagh, M. Morell, I. Mackay, W. Powell, From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants, Curr. Opin. Plant Biol. 11 (2008) 215–221.
[18] N. Bardol, M. Ventelon, B. Mangin, S. Jasson, V. Loywick, F. Couton, et al., Combined linkage and linkage disequilibrium QTL mapping in multiple families of maize (*Zea mays* L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism, Theor. Appl. Genet. 126 (2013) 2717–2736.
[19] J.R. Klasen, H.-P. Piepho, B. Stich, QTL detection power of multi-parental RIL populations in *Arabidopsis thaliana*, Heredity (Edinb.) 108 (2012) 626–632.
[20] M. Causse, N. Desplat, L. Pascual, M.-C. Le Paslier, C. Sauvage, G. Bauchet, et al., Whole genome resequencing in tomato reveals variation associated with introgression and breeding events, BMC Genomics 14 (2013) 791.
[21] J.J. Giovannoni, Genetic regulation of fruit development and ripening, Plant Cell 16 (2004) 170–180.
[22] V. Saliba-Colombani, M. Causse, L. Gervais, J. Philouze, Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome, Genome 43 (2000) 29–40.
[23] L.M. Labate JA, S. Grandillo, T. Fulton, S. Muños, A.L. Caicedo, I. Peralta, Y. Ji, R.T. Chetelat, J.W. Scott, M.J. Gonzalo, D. Francis, W. Yang, E. van der Knaap, A.M. Baldo, B. Smith-White, M.C. Eller, J.P. Prince, N.E. Blanchard, D.B. Storey, M.R. Stevens, M.D. Robbins, J. Fen Wang, B.E. Liedl, M.A. O'Connell, J.R. Stommel, K. Aoki, Y. Iijima, A.J. Slade, S.R. Hurst, D. Loeffler, M.N. Steine, D. Vafeados, C. McGuire, C. Freeman, A. Amen, J. Goodstal, D. Facciotti, J. Van Eck, Labate_et_al_2007.pdf, in: C. Kole (Ed.), Genome Mapp. Mol. Breed. Plants, vol. 5, Veg., Springer-Verlag, Berlin Heidelberg, 2007, pp. 11–135.
[24] A. Frary, T.C. Nesbitt, S. Grandillo, E. van der Knaap, B. Cong, J.P. Liu, et al., fw2.2: a quantitative trait locus key to the evolution of tomato fruit size, Science 289 (2000) 85–88.
[25] E. Fridman, T. Pleban, D. Zamir, A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 4718–4723.
[26] Tomato-Genome-Consortium, The tomato genome sequence provides insights into fleshy fruit evolution, Nature 485 (2012) 635–641.
[27] S. Aflitos, E. Schijlen, H. de Jong, Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing, Plant J. 80 (2014) 136–148.
[28] T. Lin, G. Zhu, J. Zhang, X. Xu, Q. Yu, Z. Zheng, et al., Genomic analyses provide insights into the history of tomato breeding, Nat. Genet. (2014), http://dx.doi.org/10.1038/ng.3117
[29] A. Bolger, F. Scossa, M.E. Bolger, C. Lanz, F. Maumus, T. Tohge, et al., The genome of the stress-tolerant wild tomato species *Solanum pennellii*, Nat. Genet. 46 (2014) 1034–1038.
[30] S.C. Sim, G. Durstewitz, J. Plieske, R. Wieseke, M.W. Ganal, A. Van Deynze, et al., Development of a large SNP genotyping array and generation of high-density genetic maps in tomato, PLoS ONE 7 (2012) e40563.
[31] S.-C. Sim, A. Van Deynze, K. Stoffel, D.S. Douches, D. Zarka, M.W. Ganal, et al., High-density SNP genotyping of tomato (*Solanum lycopersicum* L.) reveals patterns of genetic variation due to breeding, PLoS ONE 7 (2012) e45520.
[32] C. Sauvage, V. Segura, G. Bauchet, R. Stevens, P.T. Do, Z. Nikoloski, et al., Genome-wide association in tomato reveals 44 candidate loci for fruit metabolic traits, Plant Physiol. 165 (2014) 1120–1132.
[33] M. Chakrabarti, N. Zhang, C. Sauvage, S. Muños, J. Blanca, J. Cañizares, et al., A cytochrome P450 regulates a domestication trait in cultivated tomato, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 17125–17130.
[34] V. Saliba-Colombani, M. Causse, D. Langlois, J. Philouze, M. Buret, Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits, Theor. Appl. Genet. 102 (2001) 259–272.
[35] J.W. van Ooijen, JoinMap® 4, Software for the Calculation of Genetic Linkage Maps in Experimental Populations, Kyazma B.V., Wageningen, Netherlands, 2006.
[36] E. Lander, D. Botstein, Mapping Mendelian factors underlying quantitative traits using Rflp linkage maps, Genetics 121 (1989) 185–199.
[37] K.W. Broman, H. Wu, S. Sen, G.A. Churchill, R/qtl. QTL mapping in experimental crosses, Bioinformatics 19 (2003) 889–890.
[38] B.E. Huang, A.W. George, R/mpMap: a computational platform for the genetic analysis of multiparent recombinant inbred lines, Bioinformatics 27 (2011) 727–729.

Appendix 3

[39] J. Xu, N. Ranc, S. Muños, S. Rolland, J.P. Bouchet, N. Desplat, et al., Phenotypic diversity and association mapping for fruit quality traits in cultivated tomato and related species, Theor. Appl. Genet. 126 (2013) 567–581.

[40] S.-C. Sim, G. Durstewitz, J. Plieske, R. Wieseke, M.W. Ganal, A. Van Deynze, et al., Development of a large SNP genotyping array and generation of high-density genetic maps in tomato, PLoS ONE 7 (2012) e40563.

[41] O.J. Hardy, X. Vekemans, SPAGeDI. A versatile computer program to analyse spatial genetic structure at the individual or population levels, Mol. Ecol. Notes 2 (2002) 618–620.

[42] D. Falush, M. Stephens, J.K. Pritchard, Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies, Genetics 164 (2003) 1567–1587.

[43] G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software structure: a simulation study, Mol. Ecol. 14 (2005) 2611–2620.

[44] V. Segura, B.J. Vilhjálmsson, A. Platt, A. Korte, Ü. Seren, Q. Long, et al., An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations, Nat. Genet. 44 (2012) 825–830.

[45] D. Steinbach, M. Alaux, J. Amselem, N. Choisne, S. Durand, R. Flores, et al., GnpIS: an information system to integrate genetic and genomic data from plants and fungi, Database (Oxford) 2013 (2013) bat058.

[46] K. Manning, M. Tör, M. Poole, Y. Hong, A.J. Thompson, G.J. King, et al., A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening, Nat. Genet. 38 (2006) 948–952.

[47] B. Mangin, A. Siberchicot, S. Nicolas, A. Doligez, P. This, C. Cierco-Ayrolles, Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness, Heredity (Edinb.) 108 (2012) 285–291.

[48] Z. Huang, E. van der Knaap, Tomato fruit weight 11.3 maps close to fasciated on the bottom of chromosome 11, Theor. Appl. Genet. 123 (2011) 465–474.

[49] L. Lecomte, V. Saliba-Colombani, A. Gautier, M.C. Gomez-Jimenez, P. Duffe, M. Buret, et al., Fine mapping of QTLs of chromosome 2 affecting the fruit architecture and composition of tomato, Mol. Breed. 13 (2004) 1–14.

[50] Y.-F. Huang, D. Madur, V. Combes, C.L. Ky, D. Coubriche, P. Jamin, et al., The genetic architecture of grain yield and related traits in *Zea maize* L. revealed by comparing intermated and conventional populations, Genetics 186 (2010) 395–404.

[51] J.W. Davey, P.A. Hohenlohe, P.D. Etter, J.Q. Boone, J.M. Catchen, M.L. Blaxter, Genome-wide genetic marker discovery and genotyping using next-generation sequencing, Nat. Rev. Genet. 12 (2011) 499–510.

[52] E.L. van Dijk, H. Auger, Y. Jaszczyszyn, C. Thermes, Ten years of next-generation sequencing technology, Trends Genet. 30 (2014) 418–426.

[53] A.A. Pai, J.K. Pritchard, Y. Gilad, The genetic and mechanistic basis for variation in gene regulation, PLoS Genet. 11 (2015) e1004857.

[54] M.I. McCarthy, J.N. Hirschhorn, Genome-wide association studies: potential next steps on a genetic journey, Hum. Mol. Genet. 17 (2008) 157–165.

[55] Z. Bochdanovits, J. Simon-Sanchez, M. Jonker, W.J. Hoogendijk, A. van der Vaart, P. Heutink, Accurate prediction of a minimal region around a genetic association signal that contains the causal variant, Eur. J. Hum. Genet. 22 (2013) 238–242.

[56] R. Rincent, L. Moreau, H. Monod, E. Kuhn, A.E. Melchinger, R.A. Malvar, et al., Recovering power in association mapping panels with variable levels of linkage disequilibrium, Genetics 197 (2014) 375–387.

[57] F. Cormier, J. Le Gouis, P. Dubreuil, S. Lafarge, S. Praud, A genome-wide identification of chromosomal regions determining nitrogen use efficiency components in wheat (*Triticum aestivum* L.), Theor. Appl. Genet. 127 (2014) 2679–2693.

[58] R. Rincent, D. Laloë, S. Nicolas, T. Altmann, D. Brunel, P. Revilla, et al., Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.), Genetics 192 (2012) 715–728.

[59] M.D. Robbins, S.C. Sim, W. Yang, A. Van Deynze, E. van der Knaap, T. Joobeur, et al., Mapping and linkage disequilibrium analysis with a genome-wide collection of SNPs that detect polymorphism in cultivated tomato, J. Exp. Bot. 62 (2011) 1831–1845.

[60] J. Ross-Ibarra, P.L. Morrell, B.S. Gaut, Plant domestication, a unique opportunity to identify the genetic basis of adaptation, Proc. Natl. Acad. Sci. U. S. A. 104 (Suppl.) (2007) 8641–8648.

[61] A.J. Lorenz, M.T. Hamblin, J.-L. Jannink, Performance of single nucleotide polymorphisms versus haplotypes for genome-wide association analysis in barley, PLoS ONE 5 (2010) e14079.

[62] I. Gawenda, P. Thorwarth, T. Günther, F. Ordon, K.J. Schmid, Genome-wide association studies in elite varieties of German winter barley using single-marker and haplotype-based methods, Plant Breed. 134 (2015) 28–39.

[63] B. Brachi, N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, et al., Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature, PLoS Genet. 6 (2010) e1000940.

[64] T. Mitchell-Olds, The molecular basis of quantitative genetic variation in natural populations, Trends Ecol. Evol. 10 (1995) 324–328.

[65] J.R. Stinchcombe, H.E. Hoekstra, Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits, Heredity (Edinb.) 100 (Suppl.) (2008) 158–170.

[66] T. Mitchell-Olds, Complex-trait analysis in plants, Genome Biol. 11 (2010) 113.

[67] S. Atwell, Y.S. Huang, B.J. Vilhjálmsson, G. Willems, M. Horton, Y. Li, et al., Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines, Nature 465 (2010) 627–631.

[68] G. Pagny, P.S. Paulstephenraj, S. Poque, O. Sicard, P. Cosson, J.-P. Eyquard, et al., Family-based linkage and association mapping reveals novel genes affecting Plum pox virus infection in *Arabidopsis thaliana*, New Phytol. 196 (2012) 873–886.

[69] J.A. Poland, T.W. Rife, Genotyping-by-sequencing for plant breeding and genetics, Plant Genome J. 5 (2012) 92–102.

[70] R.S. Austin, D. Vidaurre, G. Stamatiou, R. Breit, N.J. Provart, D. Bonetta, et al., Next-generation mapping of *Arabidopsis* genes, Plant J. 67 (2011) 715–725.

[71] H. Candela, R. Casanova-Sáez, J.L. Micol, Getting started in mapping-by-sequencing, J. Integr. Plant Biol. (2014), http://dx.doi.org/10.1111/jipb.12305

[72] O. Sosnowski, A. Charcosset, J. Joets, Biomercator V3: an upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms, Bioinformatics 28 (2012) 2082–2083.

[73] F.S. Khowaja, G.J. Norton, B. Courtois, A.H. Price, Improved resolution in the position of drought-related QTLs in a single mapping population of rice by meta-analysis, BMC Genomics 10 (2009) 276.

[74] D. Zamir, Where have all the crop phenotypes gone? PLoS Biol. 11 (2013) e1001595.

[75] A. Bombarely, N. Menda, I.Y. Tecle, R.M. Buels, S. Strickler, T. Fischer-York, et al., The sol genomics network (solgenomics.net): growing tomatoes using Perl, Nucleic Acids Res. 39 (2011) D1149–D1155.

[76] L.B.B. Martin, Z. Fei, J.J. Giovannoni, J.K.C. Rose, Catalyzing plant science research with RNA-seq, Front. Plant Sci. 4 (2013) 66.

[77] B.G. Hansen, B.A. Halkier, D.J. Kliebenstein, Identifying the molecular basis of QTLs: eQTLs add a new dimension, Trends Plant Sci. 13 (2008) 72–77.

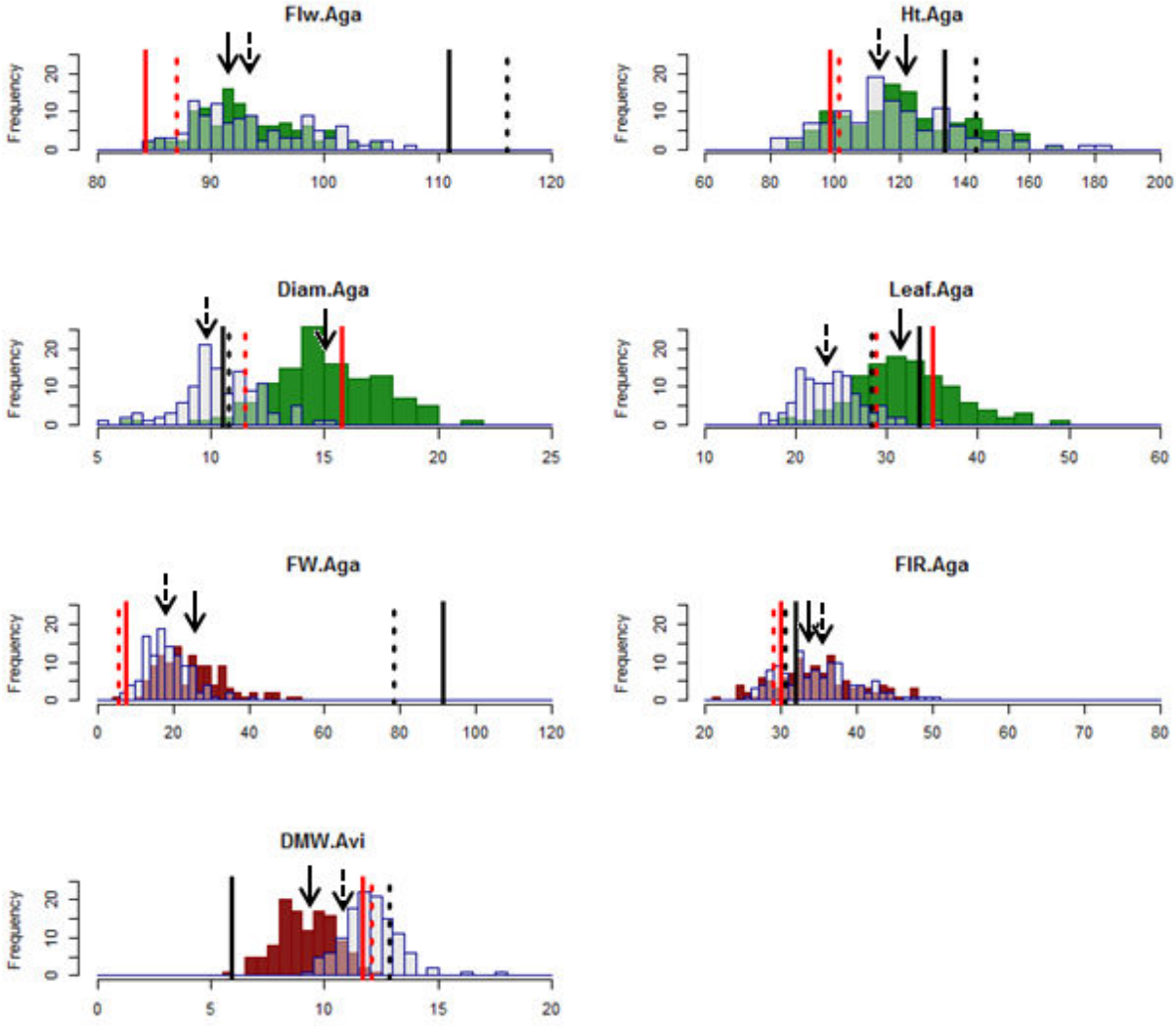
[78] M.G. Palmgren, A.K. Edenbrandt, S.E. Vedel, M.M. Andersen, X. Landes, J.T. Østerberg, et al., Are we ready for back-to-nature crop breeding? Trends Plant Sci. 20 (2014) 155–164.

[79] F. Piron, M. Nicolai, S. Minoia, E. Piednoir, A. Moretti, A. Salgues, et al., An induced mutation in tomato eIF4E leads to immunity to two potyviruses, PLoS ONE 5 (2010).

[80] R. Mauricio, Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology, Nat. Rev. Genet. 2 (2001) 370–381.

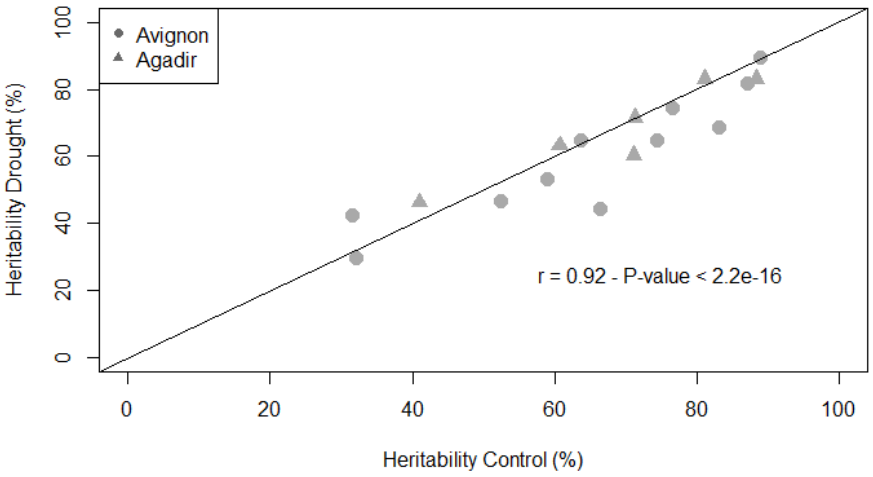
[81] S. Zhong, Z. Fei, Y.-R. Chen, Y. Zheng, M. Huang, J. Vrebalov, et al., Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening, Nat. Biotechnol. 31 (2013) 154–159.

Appendix 4: Supplemental figures and tables from chapter 3.

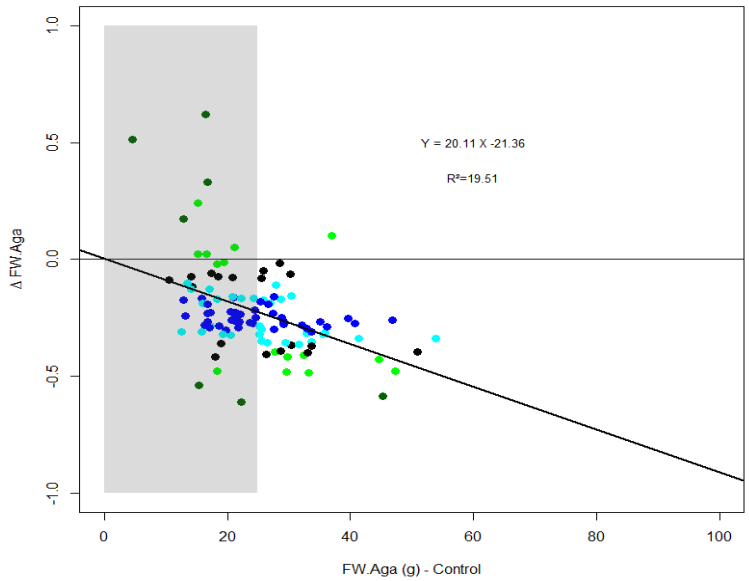


Supplemental Fig 1 Distribution of the average plant and fruit traits in the recombinant inbred lines population (RILs) grown under two watering regimes. Opaque color indicates trait values under control treatment and transparent color trait values under drought treatment. The parental mean values are indicated: full red line for Cervil in control treatment, dashed red line for Cervil in drought treatment, full black line for Levovil in control treatment and dashed black line for Levovil in drought treatment. The black arrows represent the RIL population means: dashed arrow for drought and full arrow for control treatment.

Appendix 4

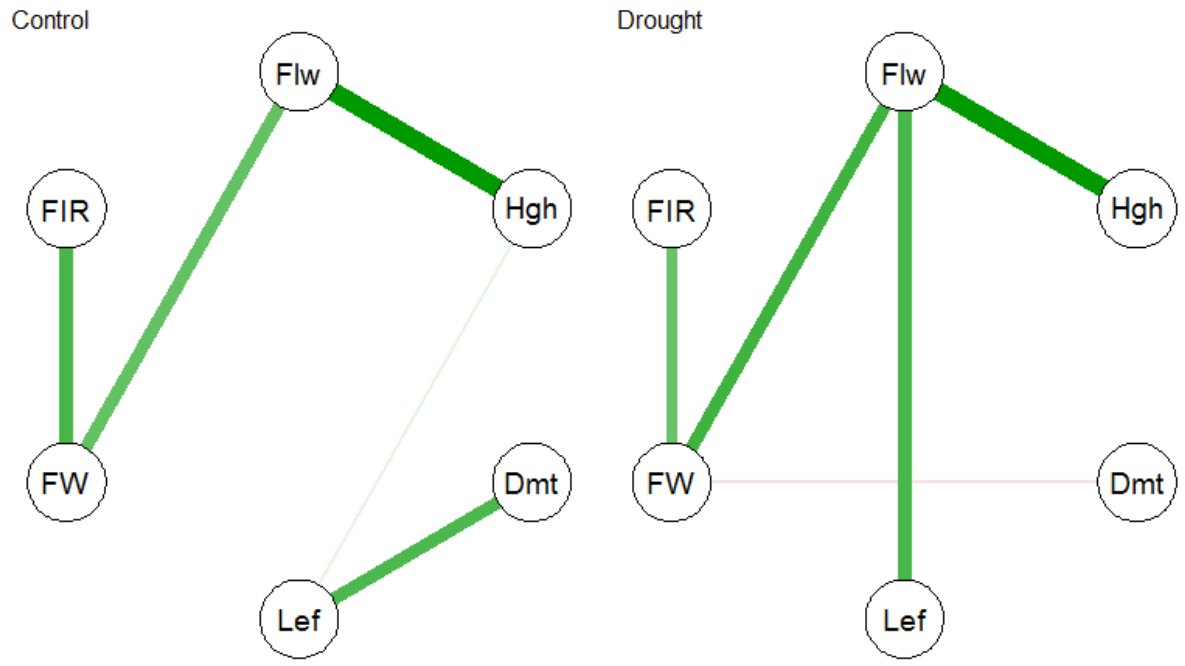


Supplemental Fig 2 Correlation between broad sense heritability measured under drought and control treatments, for plant and fruit traits. Dots and triangles correspond to trait heritabilities measured in Avignon and Agadir, respectively. r indicates the spearman correlation coefficient between heritabilities. The associated *P-value* is displayed.



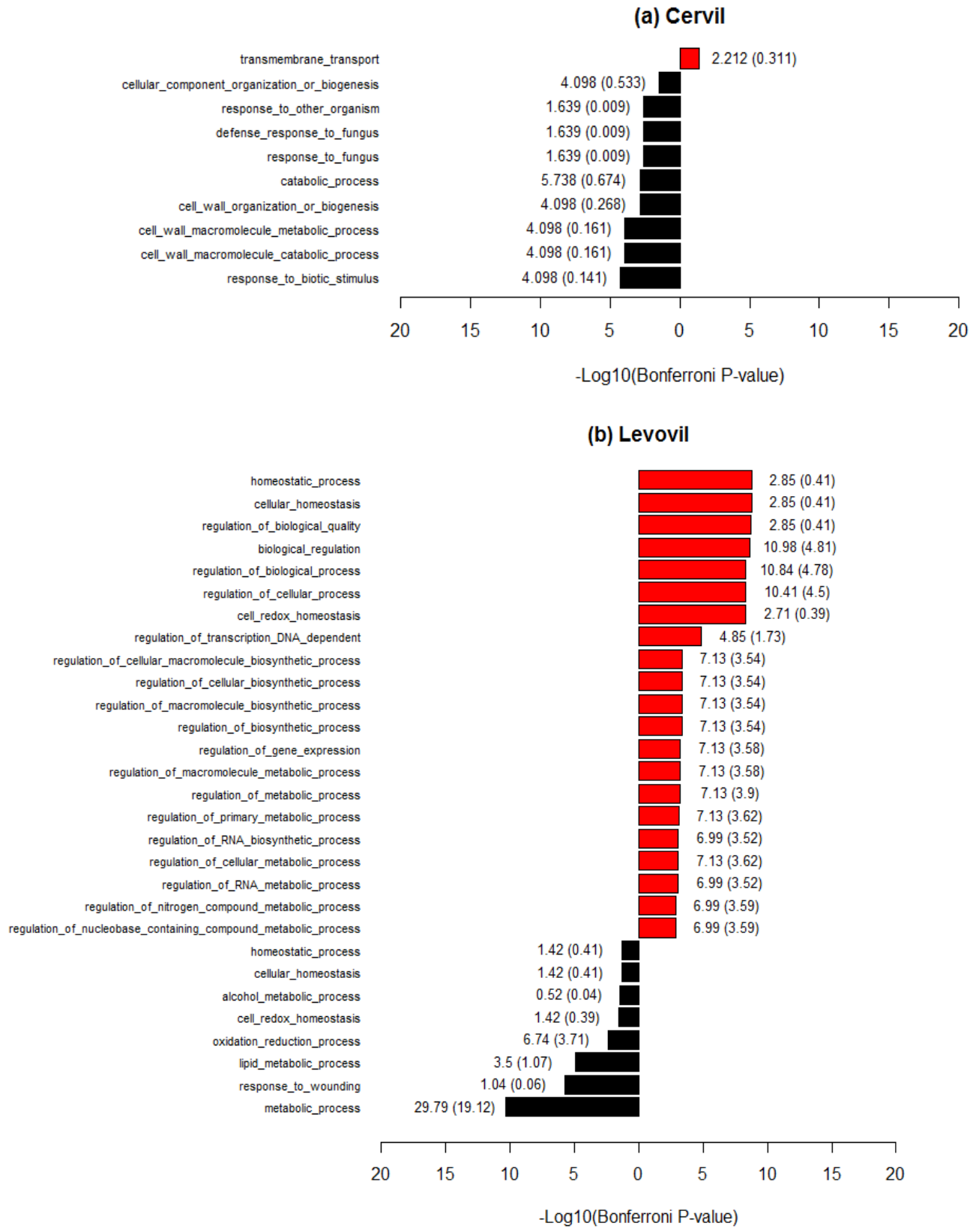
Supplemental Fig 3 Linear relationships between Fresh Weight in control condition and Δ Fresh Weight in the RIL population in Agadir Equation and R^2 of the linear regression lines are displayed. Color indicates relative ecovalence classes: blue < 0.1; 0.1 < cyan < 0.5; 0.5 < black < 1; green > 1. The gray areas indicate small fruit accessions (FW < 25 g).

Appendix 4



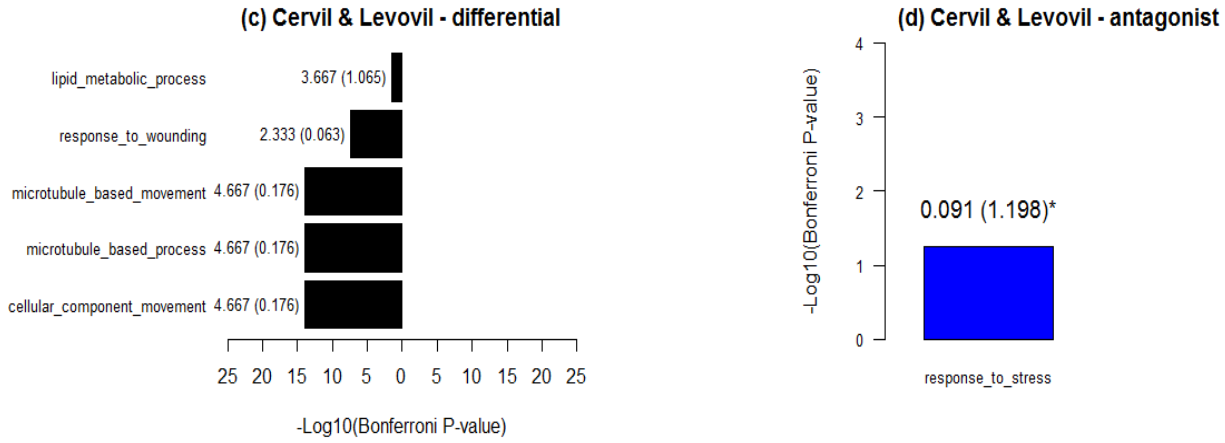
Supplemental Fig 4 Changes in phenotypic correlation network between the two watering regimes. The figure displays Pearson correlation coefficients between average phenotypic values measured in control and drought treatment, in Agadir trial. Only coefficients beyond 0.2 are shown (*P-value* < 0.05). The line width is proportional to correlation coefficient value. The line color indicates direction of the correlation: green for positive correlations and red for negative correlations. *Abbreviations meanings: 'Flw' for Flw.Aga ; 'Hgh' for Ht.Aga ; 'Dmt' for Diam.Aga ; 'Lef' for Leaf.Aga; 'FW' for FW.Aga and 'FIR' for FIR.Aga.*

Appendix 4



Supplemental Fig 5.

Appendix 4



Supplemental Fig 5 Enrichment analysis of Gene Ontology terms (GO) related to biological process for genes differentially expressed between watering regimes in (a) Cervil only, in (b) Levovil only and in both accessions with regulation change in the same direction ('differential') (c) or in opposite direction ('antagonist') (d). Only GO terms with a Bonferonni adjusted *P-value* below 0.05 are shown, excepted on graph (c) marked with a '*' indicating a *P-value* close to significant (0.057). At the top of each bar, frequences of the GO terms in the corresponding differentially expressed gene clusters followed by the frequence in the tomato genome between brackets are displayed in percentage. Black: Down regulated genes. Red: Up regulated genes. Blue: Genes regulated in different direction according to the accession.

Supplemental Table 4 Correlations between the two trials. Pearson correlation coefficients between average trait values observed in Avignon and Agadir trials, under the two watering regimes (control and drought), are displayed.

Trait	Control	Drought
Flw	0.71 ***	0.77 ***
Diam [a]	0.27 ***	0.42 ***
Leaf [a]	0.43 ***	0.30 ***
Ht [a]	0.68 ***	0.66 ***
FW [a]	0.70 ***	0.72 ***
FIR	0.46 ***	0.53 ***

[a] Data transformed for skewed distribution

*** shows P-value below 0.001.

Supplemental Table 5 Correlations between the two watering regimes. Pearson correlation coefficients between average trait value observed in control and drought treatments, in Avignon and Agadir trials, are displayed. We indicated "-" for traits not measured in Agadir.

Trait	Avignon	Agadir
Flw	0.76 ***	0.88 ***
Diam	0.22 **	0.32 ***
Leaf	0.56 ***	0.37 ***
Ht [a]	0.78 ***	0.76 ***
Nbfruits [a]	0.71 ***	-
FW [a]	0.82 ***	0.81 ***
FIR	0.54 ***	0.55 ***
pH	0.57 ***	-
DMW	0.38 ***	-
SSC	0.41 ***	-
VitCFM	0.62 ***	-
VitCDM	0.40 ***	-
Yield	0.26 ***	-

[a] Data transformed for skewed distribution

*** shows P-value below 0.001, ** between 0.001 and 0.01.

Appendix 4

Supplemental Table 6 QTLs detected for plant and fruit average phenotypic values based on SIM analyses and ANOVA procedure. For each QTL, the watering regime is indicated by 'WR': C for 'Control', D for 'Drought' and Int for QTL identified as interactive by ANOVA testing procedure but not significant by SIM analysis. The position of the QTL in cM Haldane and its chromosome are indicated by Pos and Chr, respectively. The most closely associated marker is indicated. LOD is the log-likelihood at that position, when significant (between brackets when just under significance). CI indicates the genetic confidence interval in cM Haldane calculated by LOD decrease of one unit, and its physical equivalent (Mpb) on genome assembly 2.5 between brackets (when significant in the two watering regime, only the shorter interval is displayed). For each watering regime, the average value of the two parental alleles (Cer and Lev, with the standard deviation between brackets) and the percentage of phenotypic variation explained by the QTL (PVE) are displayed 'P-value W' and 'P-value G x W' indicates the marker P-value in the ANOVA procedure, for watering effect and interaction effect, respectively. Significant P-values are in bold underlined letters and P-value just under threshold in bold italic letters. The interactive QTL type is specified: 'antagonist' for antagonist effect and 'differential' for differential sensitivity effect.

Trait	QTL						Control			Drought			Testing		
	WR	Chr	Pos	Marker	CI cM (Mpb)	LOD [b]	Cer	Lev	PVE	Cer	Lev	PVE	P-value W [c]	P- value G x W [c]	Interaction type
Flw.Avi	D	2	0	Y02_09141111	0.00 - 47.58 (9.14 - 42.63)	3.67	160.45 (3.57)	162.85 (4.84)	7.60	159.90 (3.44)	162.83 (4.05)	13.47	5.54E-01	6.07E-01	ns
	C	4	108	Y04_64155108	76.16 - 114.97 (59.33 - 65.31)	3.11	160.12 (3.88)	163.12 (4.37)	11.8	160.07 (3.78)	162.56 (3.88)	9.67	5.53E-01	6.30E-01	ns
	D	8	79	Y08_63002251	63.80 - 86.21 (60.62 - 63.45)	3.37	160.40 (4.42)	163.06 (3.88)	9.27	160.03 (3.76)	162.84 (3.81)	12.27	5.54E-01	8.82E-01	ns
Flw.Aga	C & D	2	43	Y02_41344166	0.00 - 50.94 (9.14 - 43.42)	3.58 & 5.06	92.19 (4.18)	95.49 (4.04)	13.3	92.12 (4.64)	97.00 (5.25)	19.33	3.29E-01	2.07E-01	ns
	D	4	108	Y04_64155108	66.84 - 116.21 (57.90 - 65.45)	3.19	92.07 (4.13)	94.90 (4.21)	10.50	92.18 (4.47)	96.02 (5.63)	12.77	3.42E-01	4.13E-01	ns
Diam.Avi	D	3	105	Y03_65566757	102.80 - 109.91 (65.38 - 66.17)	8.57	14.41 (1.51)	13.56 (1.77)	6.31	11.83 (1.05)	10.38 (1.18)	29.42	**1.13E-37**	1.21E-01	ns
	C & Int	4	102	Y04_63370382	86.96 - 104.98 (61.15 - 63.71)	4.58	13.28 (1.56)	14.67 (1.50)	17.33	11.10 (1.19)	10.98 (1.46)	0.21	**5.03E-35**	**1.00E-04**	antagonist
	D	12	32	Y12_02821857	0.00 - 53.73 (0.09 - 37.09)	3.32	14.17 (1.44)	13.79 (1.88)	1.24	11.55 (1.27)	10.58 (1.24)	13.34	**3.95E-35**	1.32E-01	ns
Diam.Aga [a]	C & D	3	103	Y03_65380813	99.29 - 106.14 (64.90 - 65.66)	6.15 & 3.78	4.08 (0.25)	3.76 (0.31)	22.40	3.38 (0.18)	3.17 (0.31)	13.93	**2.56E-45**	1.70E-01	ns
	Int	4	118	Y04_65935430	Ø	ns	3.84 (0.37)	3.95 (0.26)	3.03	3.30 (0.26)	3.20 (0.29)	3.10	**8.62E-40**	***8.14E-03***	ns (antagonist)
	D	12	22	Y12_01921369	4.84 - 24.30 (0.67 - 2.37)	3.49	3.95 (0.33)	3.85 (0.32)	2.26	3.35 (0.23)	3.15 (0.29)	12.69	**1.90E-41**	1.94E-01	ns
Leaf.Avi	C	2	47	Y02_42277511	43.44 - 50.94 (41.34 - 43.42)	3.35	41.80 (5.50)	45.89 (4.78)	13.06	37.05 (4.46)	38.35 (4.01)	2.19	**1.19E-17**	3.06E-02	ns
	C	6	52	Y06_43009935	50.48 - 54.86 (42.83 - 43.67)	5.24	41.18 (4.81)	45.90 (5.16)	18.55	36.46 (4.13)	38.81 (4.23)	7.50	**1.53E-18**	5.38E-02	ns
	C	9	75	Y09_68297506	60.90 - 82.37 (67.30 - 69.13)	3.08	41.80 (5.09)	45.58 (5.43)	11.51	36.59 (4.22)	38.92 (4.14)	7.21	**5.30E-18**	2.51E-01	ns

Appendix 4

Following Supplemental Table 6

Leaf.Aga [a]	C	6	15	Y06_35446106	13.48 - 20.81 (34.68 - 36.82)	3.11	1.48 (0.08)	1.53 (0.06)	11.28	1.37 (0.06)	1.38 (0.06)	1.37	**2.23E-38**	2.83E-03	ns
Ht.Avi [a]	C	6	50	Y06_42831928	15.39 - 56.28 (35.45 - 43.88)	3.79	1.95 (0.09)	2.02 (0.08)	13.39	1.94 (0.08)	1.98 (0.08)	6.77	2.41E-02	3.39E-01	ns
	D	7	16	Y07_01872441	0.00 - 26.58 (0.14 - 2.95)	3.70	1.96 (0.09)	2.01 (0.09)	5.77	1.93 (0.07)	2.00 (0.08)	13.92	2.54E-02	4.30E-01	ns
Nbfruits.Avi [a]	C & D	2	79	Y02_49107705	71.08 - 96.66 (47.61 - 52.56)	8.70 & 6.48	1.83 (0.14)	1.62 (0.15)	33.42	1.70 (0.15)	1.53 (0.12)	19.38	**3.42E-08**	4.80E-01	ns
FW.Avi [a]	C & D	2	79	Y02_49107705	79.31 - 82.22 (49.03 - 49.36)	9.02 & 12.66	1.17 (0.14)	1.37 (0.16)	31.18	0.98 (0.12)	1.17 (0.12)	40.97	**1.52E-22**	9.90E-01	ns
	C & D	2	93	Y02_52105460	89.73 - 96.66 (51.18 - 52.56)	11.3 & 10.84	1.16 (0.14)	1.37 (0.15)	34.34	0.99 (0.13)	1.16 (0.12)	33.88	**1.71E-22**	4.25E-01	ns
	C	3	97	Y03_64701243	81.40 - 118.07 (62.66 - 67.80)	3.56	1.20 (0.16)	1.34 (0.18)	14.97	1.03 (0.14)	1.11 (0.16)	6.47	**4.24E-18**	1.43E-01	ns
	C	4	118	Y04_65935430	95.70 - 118.95 (62.08 - 66.47)	3.15	1.21 (0.15)	1.33 (0.19)	11.32	1.04 (0.14)	1.11 (0.16)	5.85	**8.41E-18**	2.67E-01	ns
	C & D	11	91	Y11_55578287	87.19 - 96.73 (54.97 - 55.96)	3.06 & 6.69	1.21 (0.17)	1.33 (0.17)	10.78	1.01 (0.13)	1.15 (0.14)	22.14	**6.67E-19**	5.05E-01	ns
FW.Aga [a]	C & D	2	79	Y02_49034333	77.59 - 95.60 (48.89 - 52.39)	5.21 & 6.91	1.29 (0.16)	1.44 (0.13)	20.26	1.17 (0.14)	1.32 (0.10)	26.24	**2.43E-22**	7.35E-01	ns
	C	3	110	Y03_65831575	95.60 - 118.07 (64.52 – 67.80)	3.22	1.30 (0.16)	1.42 (0.15)	12.30	1.20 (0.13)	1.28 (0.14)	9.20	**9.11E-18**	1.39E-01	ns
	D	11	97	Y11_55958042	84.45 - 97.35 (54.76 – 56.10)	3.34	1.34 (0.17)	1.40 (0.14)	4.24	1.20 (0.15)	1.30 (0.12)	11.84	**2.34E-18**	5.90E-01	ns
FIR.Avi	D	2	99	Y02_53187144	93.19 - 101.20 (52.10 - 53.58)	5.49	47.51 (5.97)	49.74 (6.79)	2.99	47.46 (0.90)	53.2 (0.92)	14.65	4.72E-02	5.65E-02	ns
	C & D	4	67	Y04_57905541	65.92 - 70.51 (55.73 - 58.27)	5.49 & 5.03	51.12 (0.70)	45.38 (0.82)	20.00	52.92 (7.12)	46.40 (6.38)	18.70	4.72E-02	3.89E-01	ns
FIR.Aga	C & D	4	67	Y04_57905541	63.34 - 70.51 (55.37 - 58.27)	3.88 & 5.21	36.77 (5.35)	31.69 (4.39)	20.68	36.64 (5.37)	32.49 (4.17)	15.15	6.85E-01	5.01E-01	ns
	C	6	77	Y06_49506384	74.30 - 77.69 (48.82 - 49.69)	3.45	32.83 (5.34)	37.06 (5.07)	14.25	33.35 (4.95)	36.74 (5.65)	9.47	6.95E-01	5.63E-01	ns
pH.Avi	C & D	2	47	Y02_42277511	33.67 - 50.94 (41.64 - 43.09)	4.07 & 3.22	4.48 (0.15)	4.36 (0.12)	13.58	4.33 (0.12)	4.24 (0.09)	12.91	**2.27E-16**	3.01E-01	ns
	D	8	30	Y08_53246077	25.05 - 42.12 (3.74 - 57.21)	4.58	4.46 (0.14)	4.38 (0.13)	7.47	4.33 (0.12)	4.24 (0.09)	15.87	**9.86E-16**	6.15E-01	ns
DMW.Avi	C	2	79	Y02_49034333	65.08 - 93.19 (46.17 - 52.10)	4.82	9.83 (1.02)	8.76 (1.25)	17.87	12.41 (1.34)	11.81 (1.07)	5.79	**8.73E-47**	1.49E-01	ns
	Int	3	97	Y03_64701243	Ø	ns	9.54 (1.08)	8.83 (1.28)	8.01	11.69 (0.96)	12.25 (1.40)	4.95	**1.02E-43**	**1.48E-04**	antagonist

Appendix 4

Following Supplemental Table 6

	C	4	107	Y04_64091490	95.70 - 109.52 (62.08 - 64.29)	3.12	9.60 (1.09)	8.75 (1.28)	11.37	12.19 (1.25)	11.85 (1.28)	1.84	**2.49E-45**	1.21E-01	ns
	C	12	22	Y12_01865755	19.45 - 24.29 (1.69 - 2.37)	3.78	9.67 (1.14)	8.76 (0.22)	13.16	12.34 (1.05)	11.73 (1.40)	5.70	**1.90E-46**	3.32E-01	ns
SSC.Avi	C	2	75	Y02_48329646	60.84 - 79.47 (45.64 - 49.11)	6.07	9.66 (1.06)	8.49 (1.16)	21.73	11.70 (1.25)	11.16 (1.02)	5.50	**1.67E-39**	4.64E-02	ns
	Int	3	89	Y03_63932680	Ø	ns (1.11 & 2.34)	9.44 (0.99)	8.72 (1.15)	8.30	11.12 (1.10)	11.63 (1.10)	4.59	**2.11E-36**	**1.04E-04**	antagonist
	C	4	105	Y04_63784202	102.03 - 108.43 (63.37 - 64.15)	4.95	9.52 (1.07)	8.49 (1.19)	17.42	11.73 (1.23)	11.04 (1.04)	8.39	**7.35E-40**	2.43E-01	ns
	D	11	90	Y11_55409944	87.19 - 96.73 (54.97 - 55.96)	5.05	9.22 (1.23)	8.76 (1.24)	3.46	11.81 (1.08)	10.81 (1.04)	18.48	**9.15E-39**	7.61E-02	ns
	C & D	12	22	Y12_01865755	0.00 - 24.30 (0.10 - 2.37)	3.57	9.49 (1.03)	8.58 (1.28)	13.13	11.83 (1.09)	10.98 (1.15)	13.00	**5.27E-40**	8.89E-01	ns
VitCFM.Avi	C & D	4	102	Y04_63370382	95.70 - 106.29 (62.08 - 64.03)	4.58 & 5.45	48.58 (9.09)	41.84 (5.68)	16.53	57.69 (8.52)	50.19 (6.67)	19.52	**2.11E-16**	7.00E-01	ns
	D	8	28	Y08_06066102	22.85 - 42.12 (3.04 - 57.21)	4.51	47.06 (8.95)	42.64 (6.59)	7.01	56.88 (8.48)	50.20 (6.83)	15.36	**2.56E-15**	2.88E-01	ns
VitCDM.Avi	C	3	81	Y03_62665611	71.82 - 85.77 (59.44 - 63.62)	4.75	461.40 (73.61)	530.97 (79.15)	17.33	453.93 (78.29)	451.68 (70.36)	0.02	**1.56E-05**	3.21E-04	ns
	C & D	8	30	Y08_53246077	25.05 - 42.12 (3.74 - 57.21)	3.51 & 4.12	523.80 (91.83)	460.90 (58.27)	13.73	477.26 (76.43)	422.19 (57.53)	13.83	**5.97E-06**	6.90E-01	ns
Yield.Avi	C	4	71	Y04_58537245	65.92 - 95.70 (55.73 - 62.08)	3.40	921.89 (313.74)	1155.91 (298.93)	12.69	488.91 (129.04)	539.05 (143.82)	3.31	**5.85E-40**	4.57E-03	ns
	Int	8	102	Y08_65595099	Ø	ns	931.74 (326.80)	1130.15 (310.10)	8.97	498.812 (138.98)	521.41 (134.25)	0.69	**3.76E-39**	***6.65E-03***	ns (differential)
	C	11	77	TG030_Y11	76.36 - 96.73 (54.09 - 55.96)	3.39	912.73 (294.77)	1167.17 (333.00)	14.33	489.97 (149.00)	537.17 (121.19)	2.93	**3.31E-40**	1.39E-03	ns
	C	12	23	Y12_02159240	19.46 - 24.30 (1.69 - 2.37)	4.86	881.79 (260.82)	1163.02 (338.88)	17.97	474.34 (126.70)	548.81 (138.57)	7.40	**3.03E-41**	1.07E-03	ns

[a] Data transformed for skewed distribution

[b] LOD threshold from 1000-permutations ($\alpha = 0.05$) = 3.08

[c] P-value WR threshold from 1000-permutations ($\alpha = 0.05$) = $2.21\ 10^{-4}$; P-value G x WR threshold from 1000-permutations ($\alpha = 0.05$) = $1.93\ 10^{-4}$

Appendix 4

Supplemental Table 7 QTLs mapped by SIM for phenotypic plasticity. For each QTL, the position of the QTL in cM Haldane and its chromosome are indicated by Pos and Chr respectively. The most-closely associated marker is indicated. CI indicates the genetic confident interval in cM Haldane calculated by LOD decrease of one unit, and its physical equivalent (Mpb) on genome assembly 2.5 between bracket. 'Nb genes' indicates the number of genes underlying QTL. LOD is the log-likelihood at that position. Average delta values of the two parental alleles (ΔCer and ΔLev, with the standard deviation between brackets) and percentage of phenotypic variation explained by the QTL (PVE) are indicated. Besides, for each watering regime, the average phenotypic values of the two parental alleles ('Mean Cer' and 'Mean Lev', with the standard deviation between brackets) are displayed. The interactive QTL type is specified: 'antagonist' for antagonist effect and 'differential' for differential sensitivity effect. The QTL common to interactive QTL in Table 4 are indicated.

Trait	QTL									Control		Drought		Interaction type	Common to Table 4
	Chr	Pos	Marker	CI cM (Mpb)	Nb genes	LOD[b]	ΔCer	ΔLev	PVE	Mean Cer	Mean Lev	Mean Cer	Mean Lev		
Δ Flw.Aga	2	21.81	Y02_38601550	10.44 - 28.91 (36.39 - 39.62)	357	5.21	0.00 (0.02)	- 0.02 (0.03)	19.52	92.77 (4.36)	94.57 (4.14)	92.42 (4.74)	96.58 (5.45)	differential	No
Δ Diam.Avi	4	100.81	Y04_62963379	86.96 - 104.98 (61.15 - 63.70)	289	4.39	-0.16 (0.12)	-0.25 (0.11)	14.65	13.33 (1.58)	14.64 (1.52)	11.14 (1.24)	10.91 (1.41)	antagonsit	Yes
Δ Diam.Aga[a]	4	118.95	Y04_66467467	116.21 - 118.94 (65.45 - 66.47)	290	3.17	-0.12 (0.08)	-0.19 (0.08)	14.37	14.69 (2.58)	15.85 (2.19)	11.00 (1.79)	10.28 (1.72)	antagonsit	No *(just under threshold)*
Δ Nbfruits.Avi	6	15.39	Y06_35446106	9.42 - 17.68 (32.35 - 36.12)	221	3.11	-0.02 (0.03)	-0.04 (0.04)	9.50	55.71 (20.43)	57.09 (23.46)	48.08 (19.54)	39.69 (15.66)	antagonsit	No
Δ FW.Avi	3	99.29	Y03_64899693	65.39 - 109.91 (55.08 - 66.17)	1009	3.26	-0.29 (0.12)	-0.39 (0.14)	11.91	17.32 (7.45)	23.01 (10.33)	11.82 (3.88)	13.50 (5.09)	differential	No
	11	46.66	Y11_40877070	34.05 - 54.28 (4.78 - 50.49)	1196	4.52	-0.29 (0.12)	-0.41 (0.14)	18.87	17.39 (6.42)	24.87 (11.92)	12.22 (4.48)	13.64 (4.98)	differential	No
Δ pH.Avi	6	21.82	CT083_Y01_372 59141	17.68 - 28.57 (36.12 - 38.87)	245	3.56	-0.04 (0.02)	-0.02 (0.03)	14.32	4.45 (0.14)	4.41 (0.15)	4.26 (0.10)	4.32 (0.13)	antagonsit	No
Δ SSC.Avi	2	71.08	Y02_47608392	60.84 - 95.60 (45.64 - 52.39)	901	3.08	0.22 (0.14)	0.33 (0.18)	10.90	9.56 (1.07)	8.49 (1.16)	11.56 (1.29)	11.17 (1.05)	differential	No
	3	83.69	Y03_63211070	81.40 - 99.29 (62.66 - 64.90)	289	7.25	0.18	0.36	24.51	9.37 (0.98)	8.69 (1.36)	10.99 (1.02)	11.69 (1.23)	antagonsit	Yes
Δ VitCDM.Avi	3	79.08	Y03_61970262	71.82 - 85.76 (59.44 - 63.62)	430	4.49	-0.00 (0.17)	-0.14 (0.14)	16.73	464.65 (78.12)	529.04 (77.27)	457.83 (80.82)	447.98 (67.10)	antagonsit	No

[a] Data transformed for skewed distribution

[b] LOD threshold from 1000-permutations (α = 0.05) = 3.08

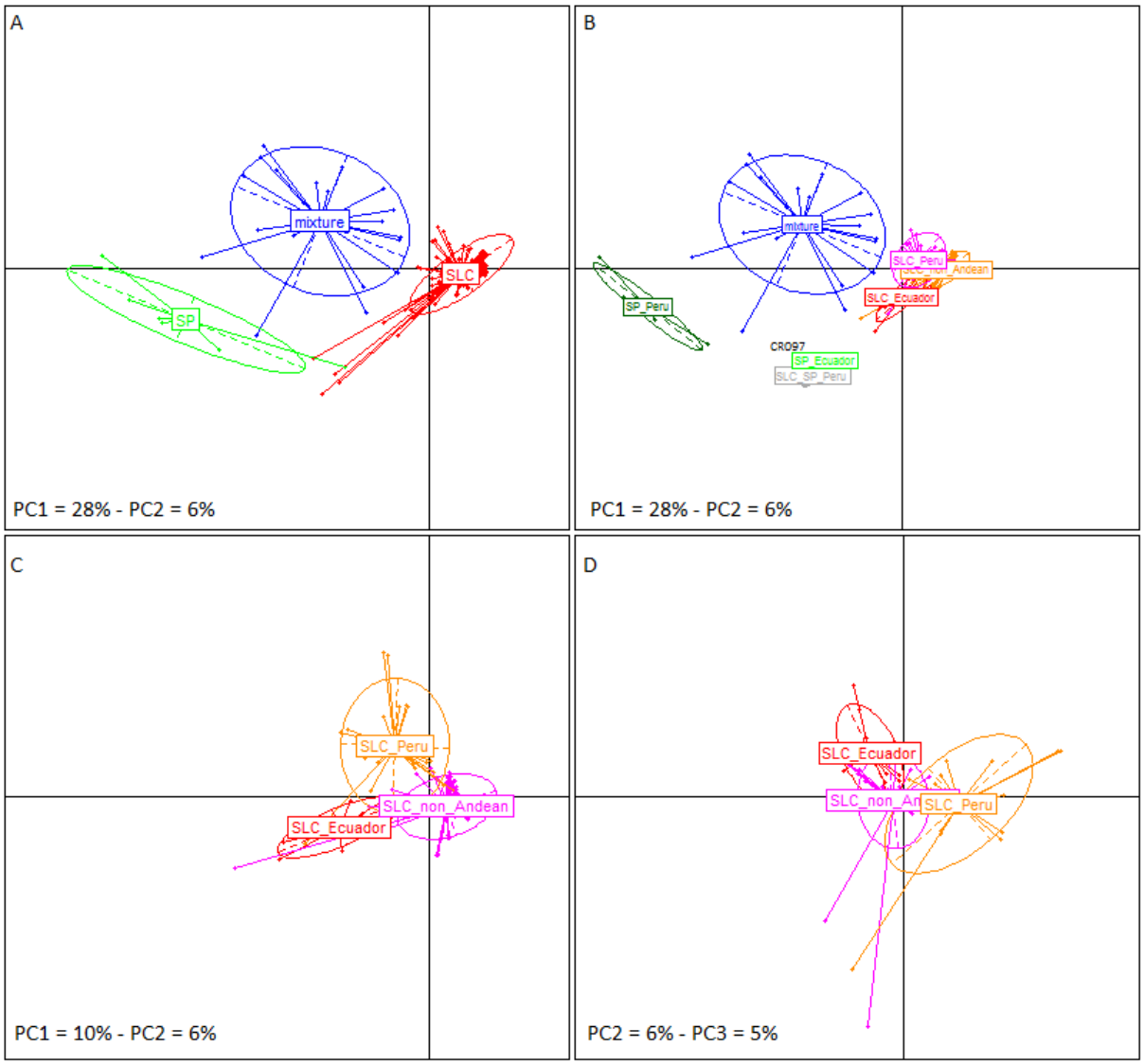
Appendix 4

Supplemental Table 8 Genes underlying interactive QTLs For each interactive QTL, 'Chr', 'CI/Pos' and 'Size' indicate the chromosome, the confident interval/position on tomato genome assembly 2.5 (Mpb) and the genomic size of the interval (Mpb), respectively. When a QTL was detected with both interaction testing methods, we used the shortest confidence interval.

Trait	Chr	CI/Pos	Size	Nb genes
Flw	2	36.39 - 39.62	3.23	357
Diam	4	61.15 - 63.70	2.55	289
Nbfruits	6	32.35 - 36.12	3.77	221
FW	3	55.08 - 66.17	11.09	1009
	11	4.78 - 50.49	45.71	1196
pH	6	36.12 - 38.87	2.75	245
DMW	3	64.70	NA	NA
SSC	2	45.64 - 52.39	6.75	901
	3	62.66 - 64.90	2.24	289
VitCDM	3	59.44 - 63.62	4.18	430

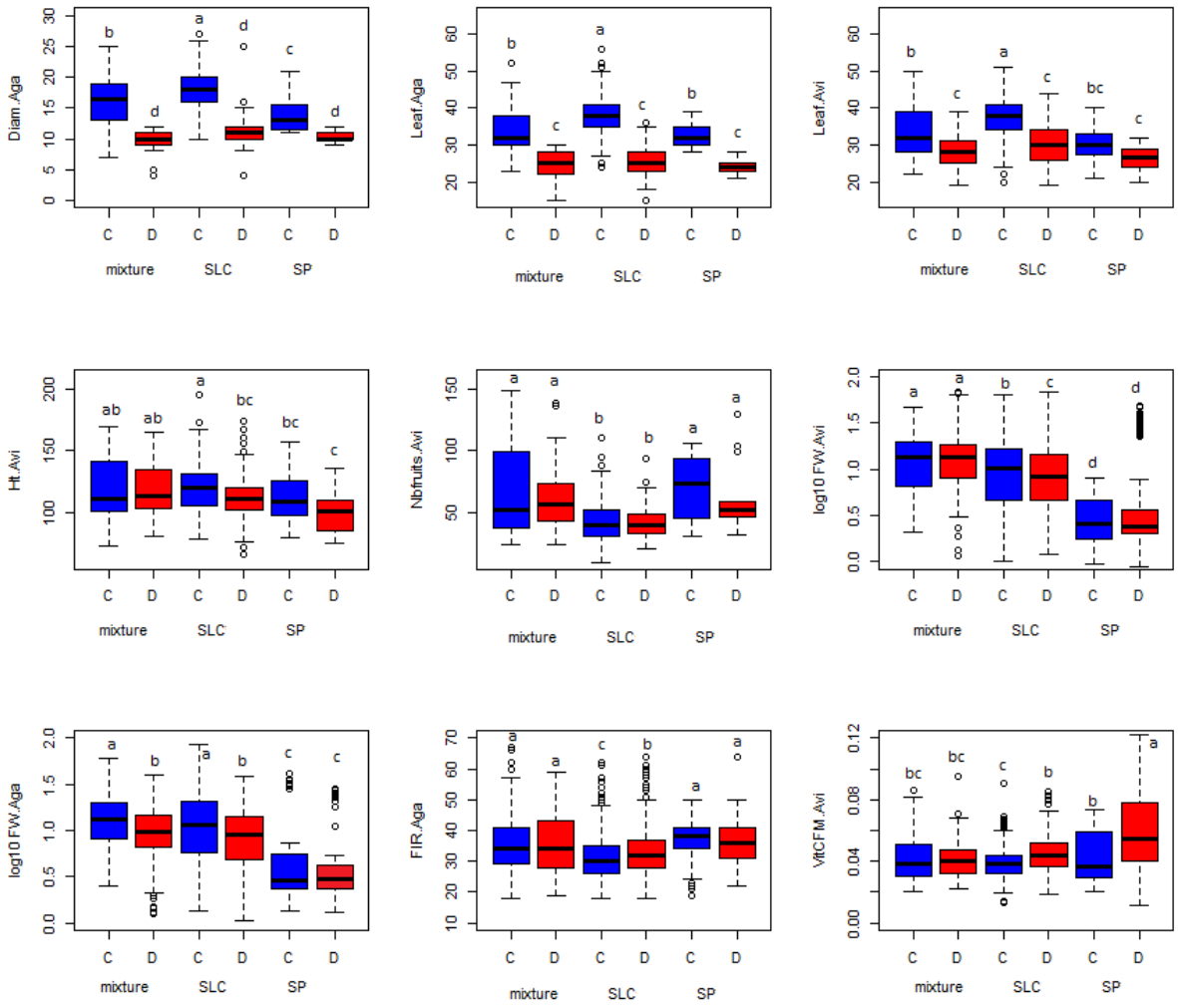
NA For interactive QTL detected only with the ANOVA procedure, we were not able to display CI and PVE.

Appendix 5: Supplemental figures and tables from chapter 4.



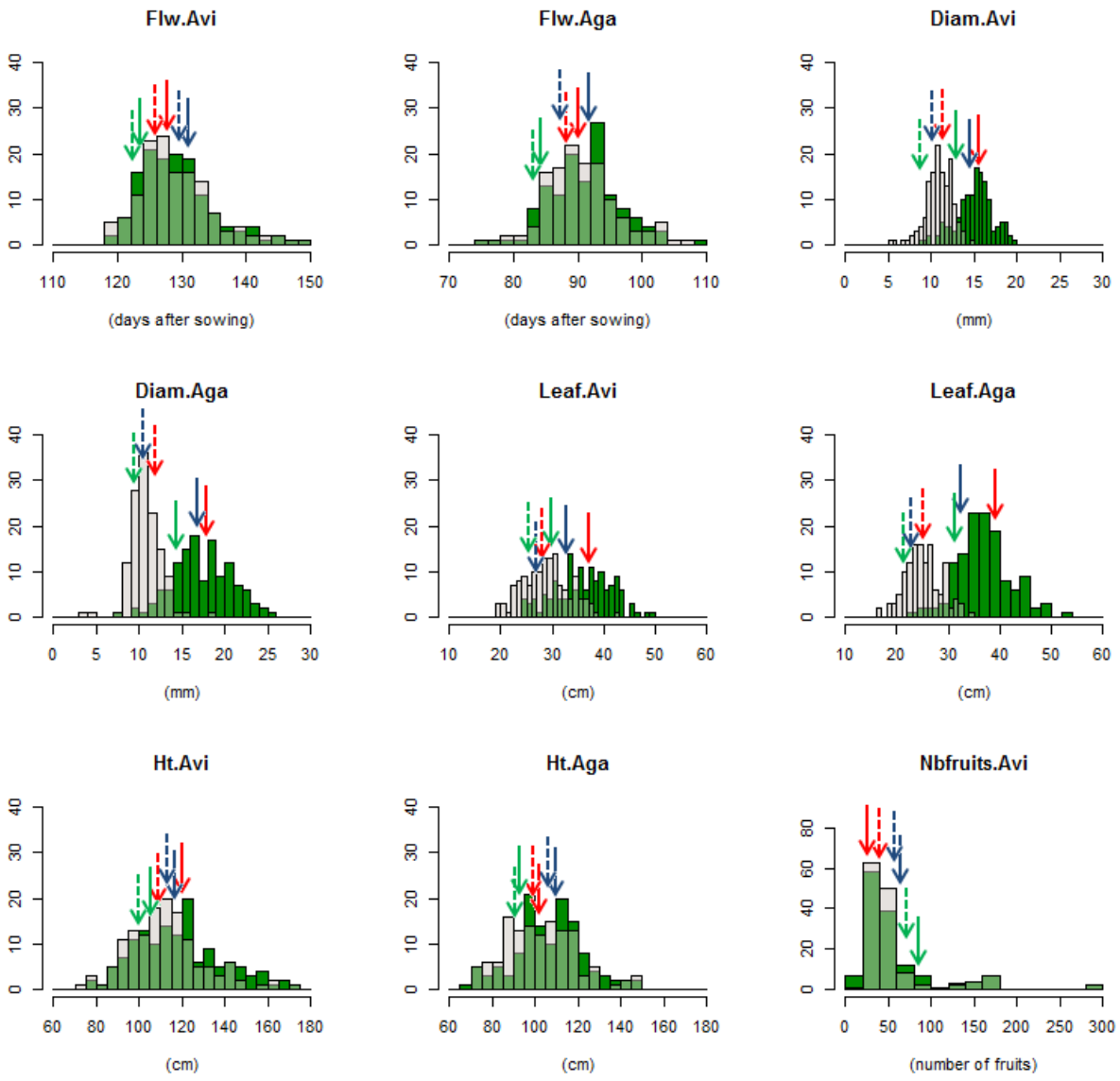
Supplemental Figure 1. Structuration observed in the GWA population based on principal coordinate analysis (PCoA) on 6,100 SNP data. (A) Analysis on the full population with coloration according to genetic specie affiliations reported in passport data ('SP': *Solanum pimpinellifolium*; 'SLC': *S. lycopersicum* var. *cerasiforme* and 'mixture': admixed accessions). (B) Analysis on the full population with coloration according to genetic sub-group affiliations ('non-Andean SLC'; 'SLC Peru'; 'SLC Ecuador'; 'SLC-SP Peru', 'SP Peru', 'SP Ecuador' and the unclassified accession 'CR097') proposed by Blanca *et al*. (2015). (C) and (D) Analysis reduced to the SLC accessions with coloration according to genetic sub-group affiliations. Ellipses of dispersion gather 67% of the individuals for a given grouping factor (dispersion coefficient: k =1.5).

Appendix 5



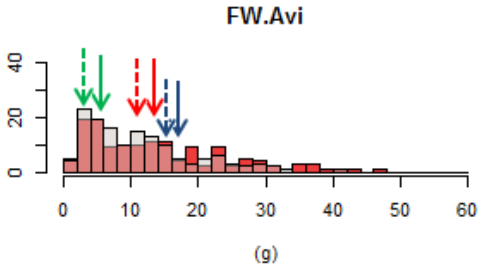
Supplemental Figure 2. Box-plot of the mean distribution for the 9 traits that showed a significant genetic group by watering regime interaction in the ANOVA tests. 'SP' stands for *Solanum pimpinellifolium* and 'SLC' for *Solanum lycopersicum* var. *cerasiforme*. Means values labeled with different letters were significantly different in the Tukey's tests (*P-value* < 0.05). Blue: control (C). Red: drought (D).

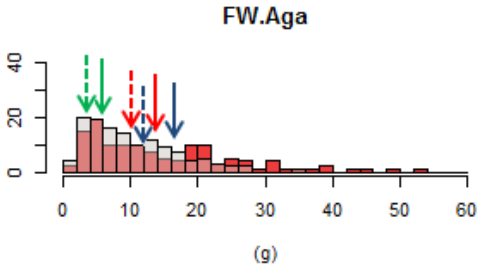
Appendix 5

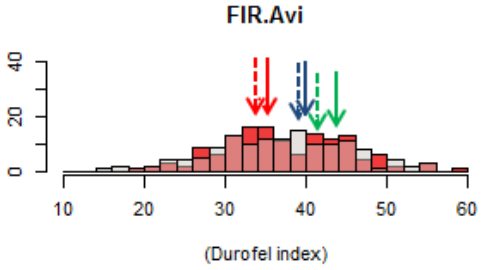


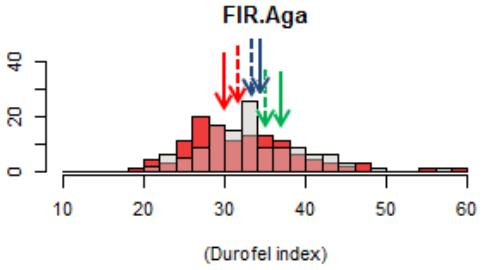
Supplemental Figure 3. Distribution of the accession means for plant traits in the GWA population grown under two watering regimes. Dark color shows trait values under control treatment and transparent color trait values under drought treatment. The full and dashed arrows indicate the average values in the population under control and water deficit treatments, respectively. Arrow color indicates the genetic groups: green for 'SP', red for 'SLC'and blue for 'mixture'.

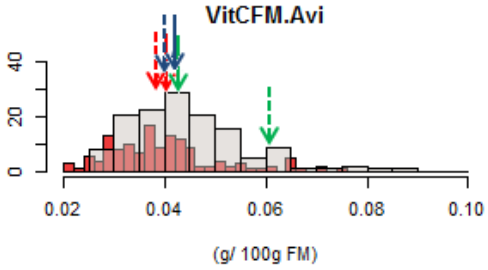
Appendix 5

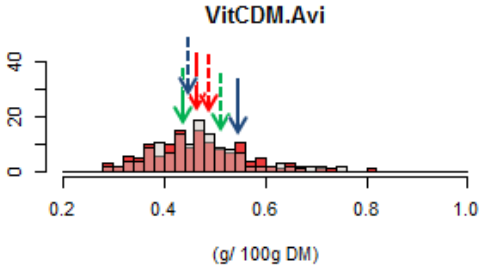

FW.Avi
0
20
40
0
10
20
30
40
50
60
(g)

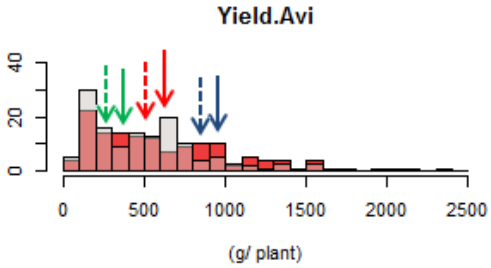

FW.Aga
0
20
40
0
10
20
30
40
50
60
(g)


FIR.Avi
0
20
40
10
20
30
40
50
60
(Durofel index)

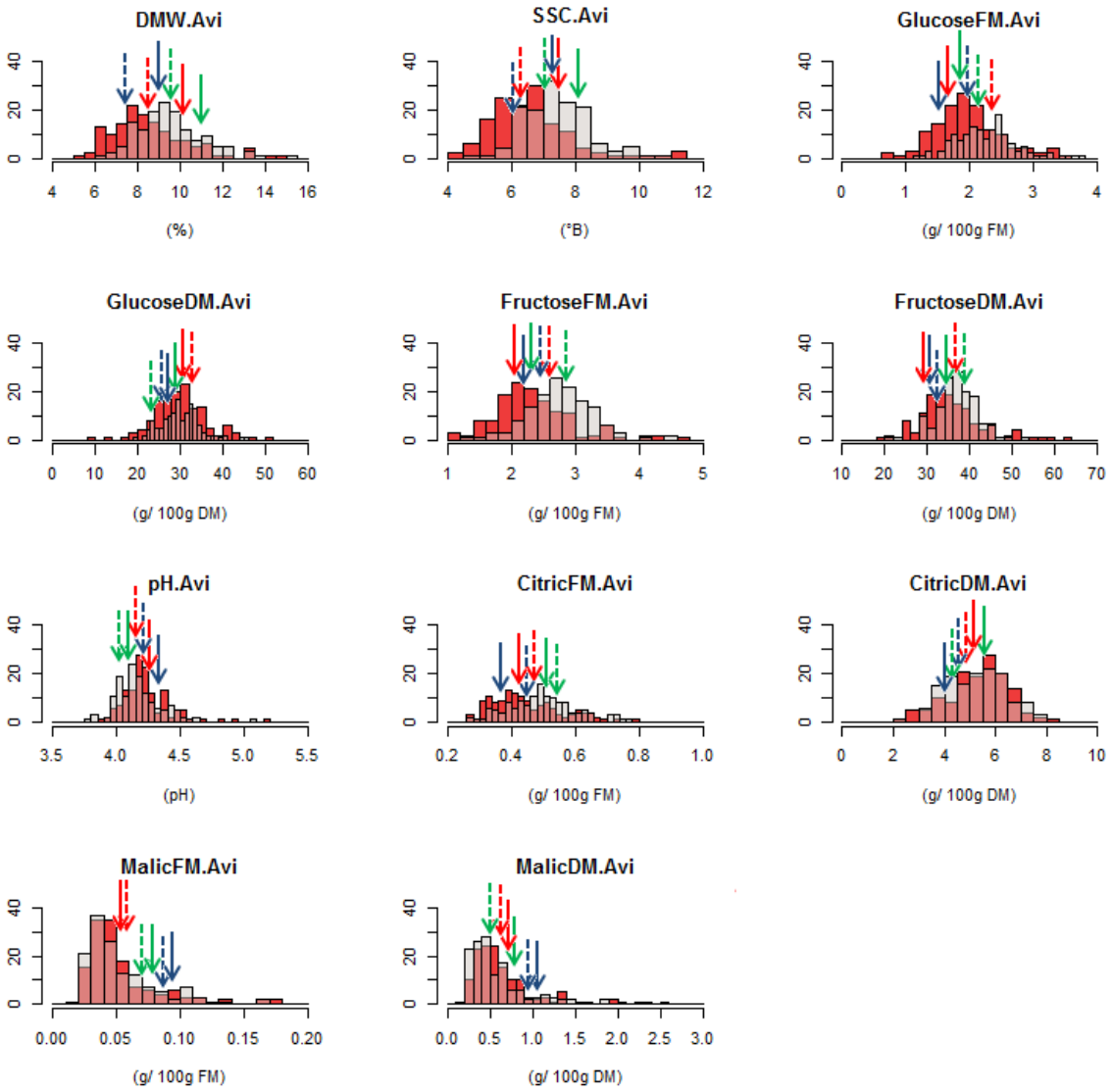

FIR.Aga
0
20
40
10
20
30
40
50
60
(Durofel index)


VitCFM.Avi
0
20
40
0.02
0.04
0.06
0.08
0.10
(g/ 100g FM)


VitCDM.Avi
0
20
40
0.2
0.4
0.6
0.8
1.0
(g/ 100g DM)

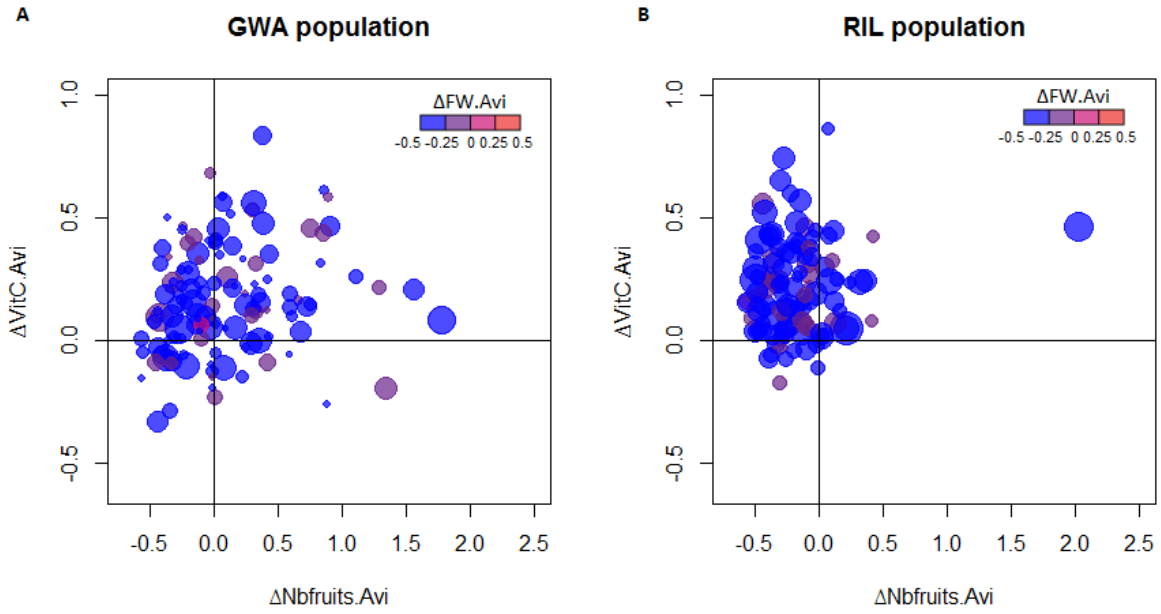

Yield.Avi
0
20
40
0
500
1000
1500
2000
2500
(g/ plant)

Appendix 5



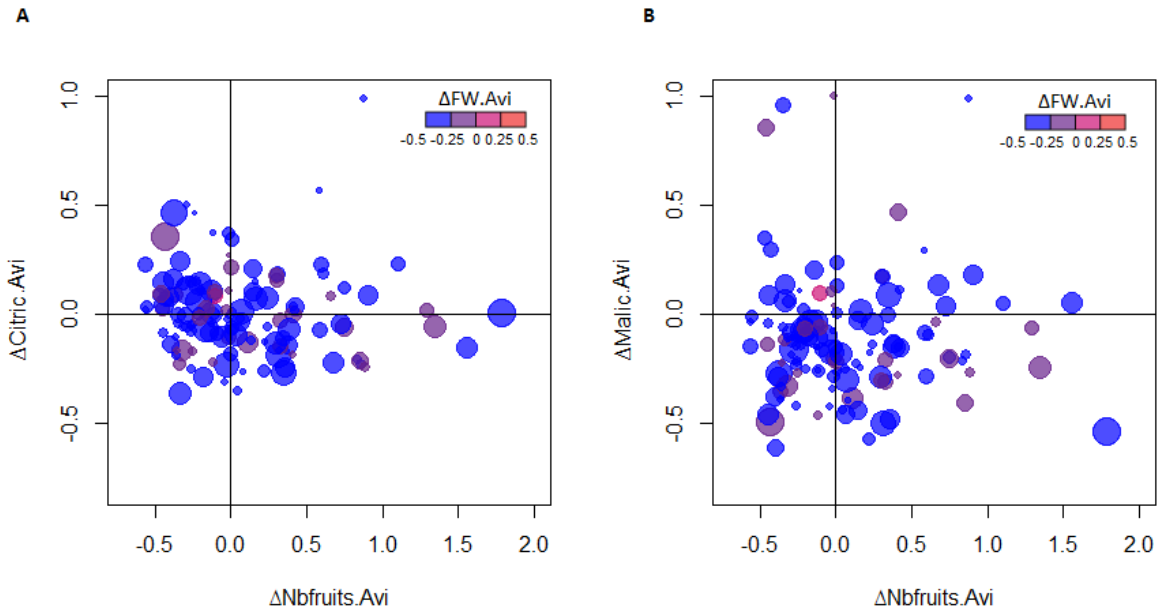
Supplemental Figure 4. Distribution of the accession means for fruit traits in the GWA population grown under two watering regimes. Dark color shows trait values under control treatment and transparent color trait values under drought treatment. The full and dashed arrows indicate the average values in the population under control and water deficit treatments, respectively. Arrow color indicates the genetic groups: green for 'SP', red for 'SLC'and blue for 'mixture'.

Appendix 5



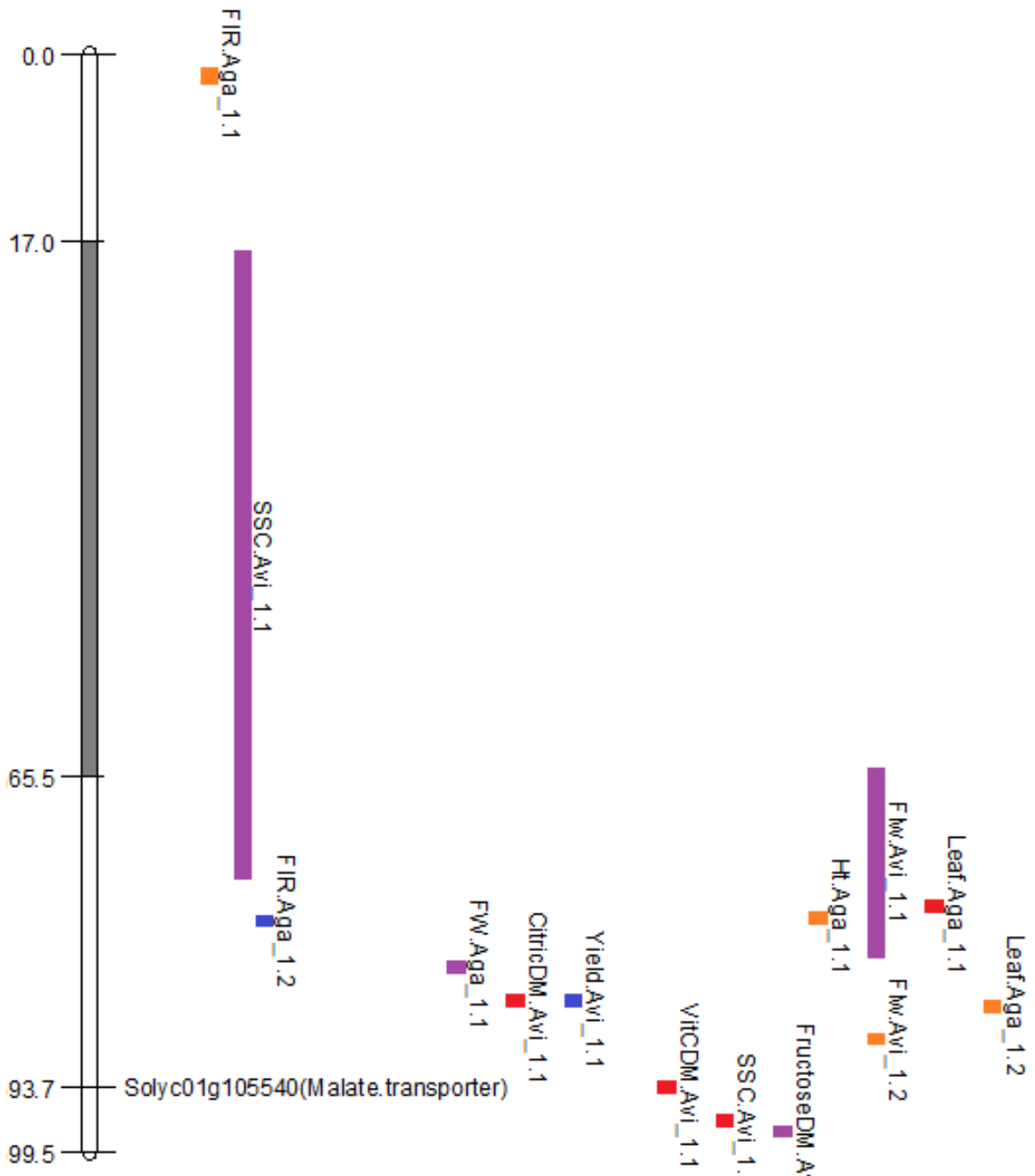
Supplemental Figure 5. Relationship between plasticity of fruit number (ΔNbfruits) and plasticity of Vitamin C (ΔVitC, relatively to fresh weight) content in fruit, in view of the fruit fresh weight plasticity (ΔFW), in the GWA and RIL populations, respectively. The color scale indicates the variation in FW plasticity: blue for values below -0.5, purple for values between -0.25 and 0, magenta for values between 0 and 0.25 and red for values beyond 0.5. The size of the points is proportional to fruit fresh weight in control watering condition.

Appendix 5

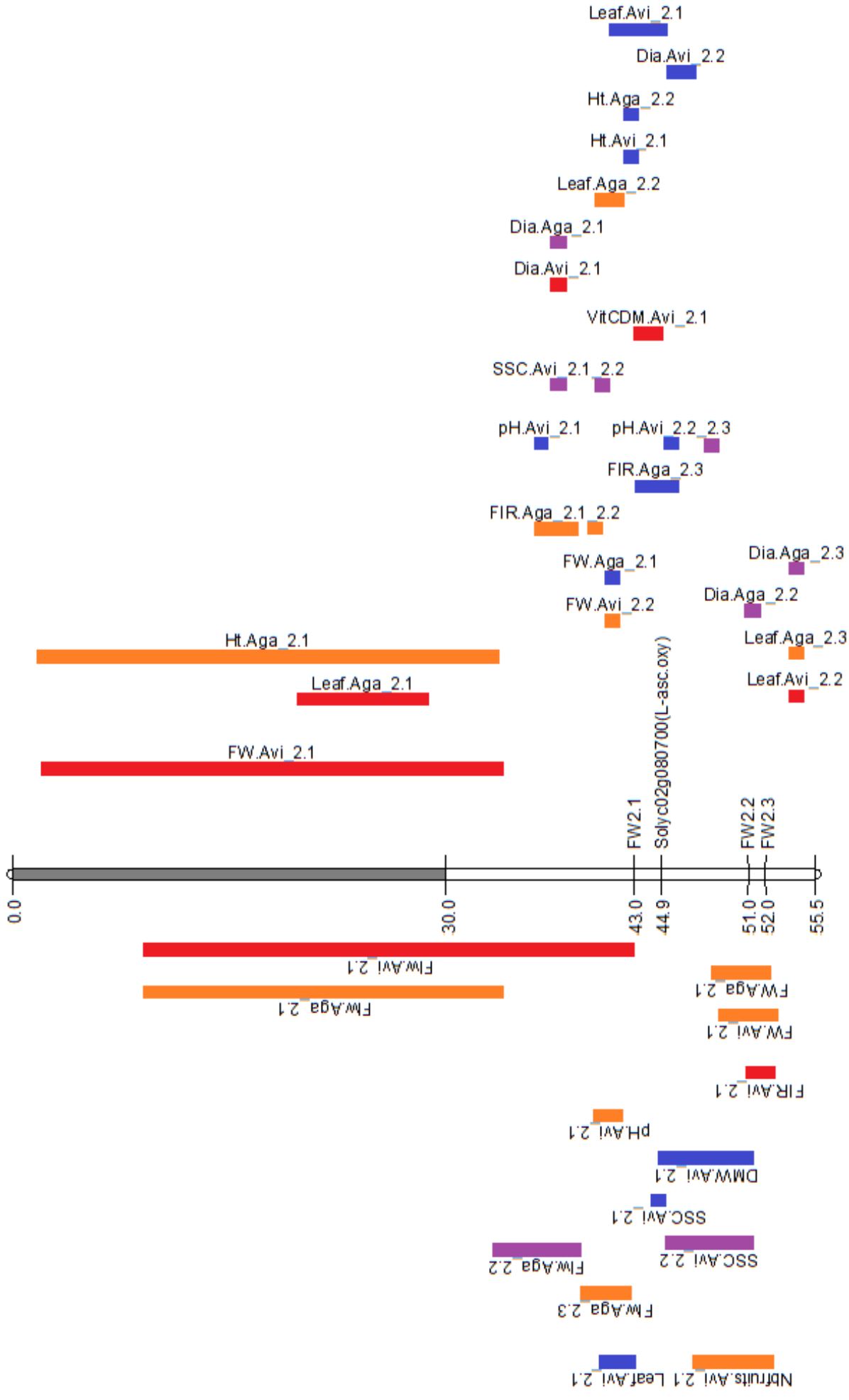


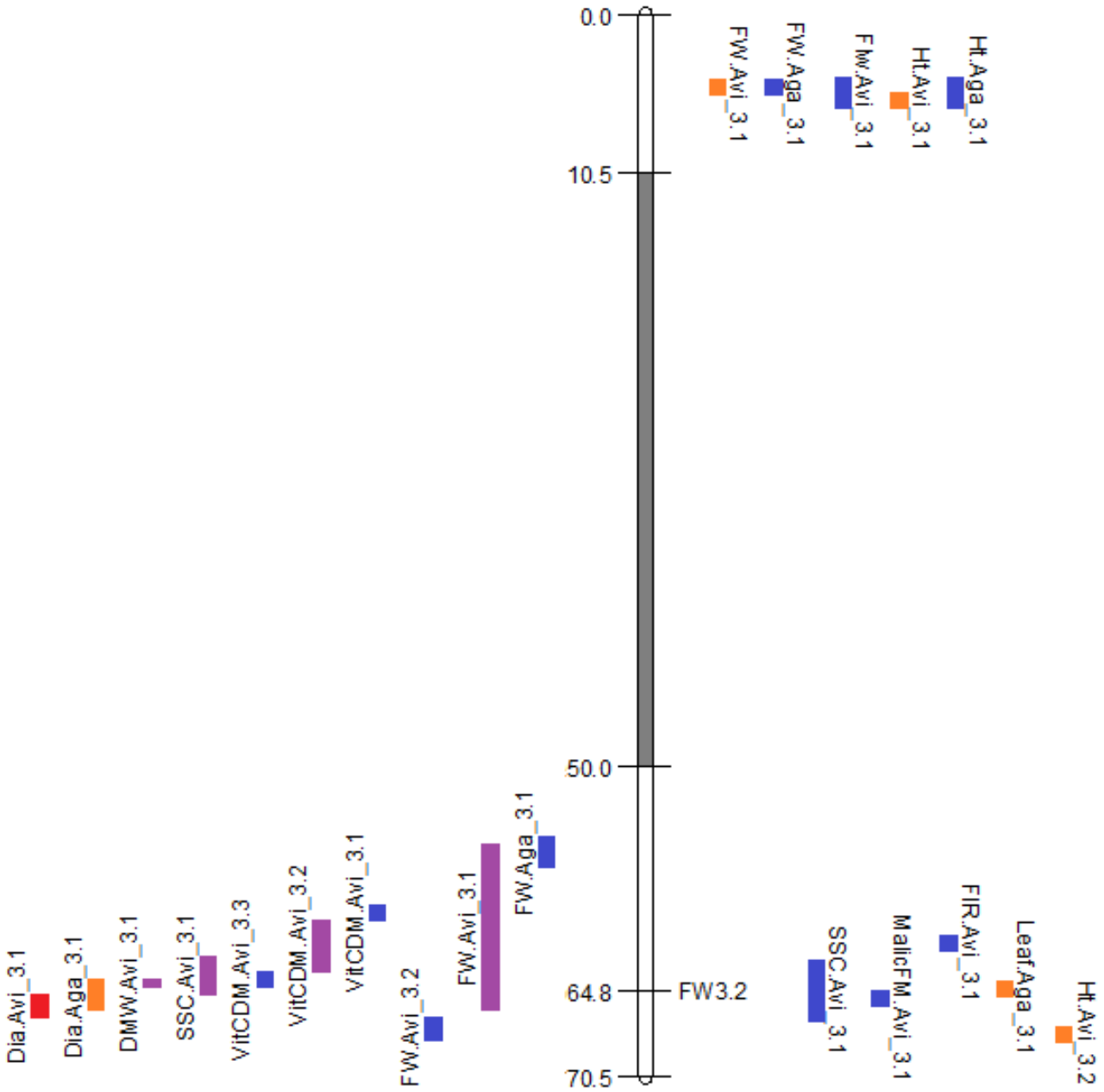
Supplemental Figure 6. Relationship between plasticity of fruit number (ΔNbfruits) and plasticity of Citric (ΔCitric) (A) and Malic (ΔMalic) (B) content in fruit (relatively to fresh weight), in view of the fruit fresh weight plasticity (ΔFW), in the GWA population. The color scale indicates the variation in FW plasticity: blue for values below -0.5, purple for values between -0.25 and 0, magenta for values between 0 and 0.25 and red for values beyond 0.5. The size of the points is proportional to fruit fresh weight in control watering condition.

Supplemental Figure 7. Physical map of the QTLs detected in the GWA and RIL populations. Distances are expressed in million bp on the tomato genome assembly 2.5. For each chromosome, QTLs detected in the GWA population are drawn to the right and QTLs detected in the RIL population to the left. Grey color on the chromosome bars indicates the centromeric regions with low recombination frequency according to Sim *et al*. 2012. Orange: constitutive QTLs. Red: drought specific QTLs. Blue: control specific QTLs. Purple: interactive QTLs. Candidate genes under some QTLs are indicated through their solyc code (tomato genome annotation 2.4).

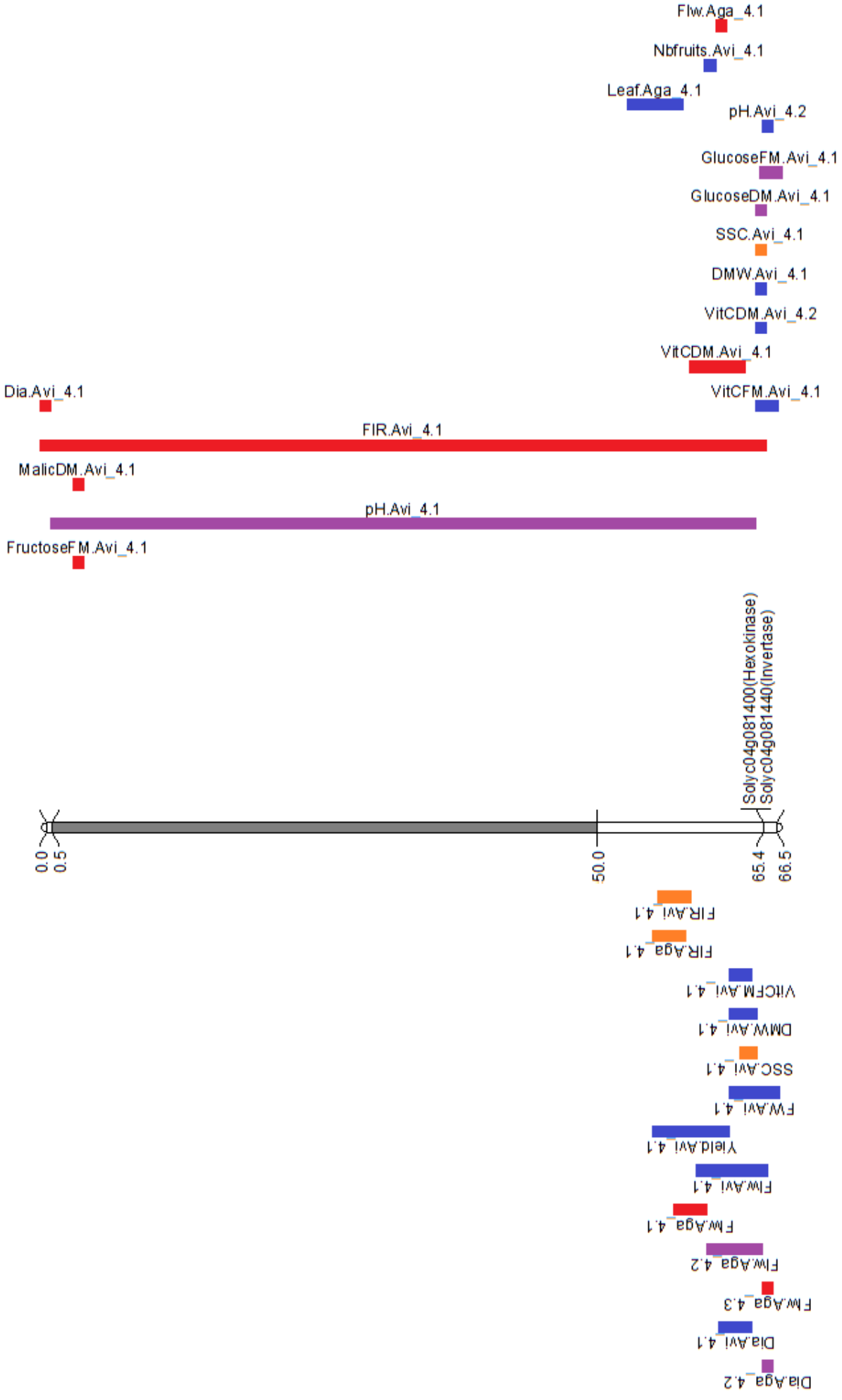

0.0
17.0
65.5
93.7
99.5
FIR.Aga_1.1
SSC.Avi_1.1
FIR.Aga_1.2
FW.Aga_1.1
CitricDM.Avi_1.1
Yield.Avi_1.1
VitCDM.Avi_1.1
SSC.Avi_1.2
FructoseDM.Avi_1.1
Ht.Aga_1.1
Fw.Avi_1.1
Fw.Avi_1.2
Leaf.Aga_1.1
Leaf.Aga_1.2
Solyc01g105540(Malate.transporter)

Appendix 5

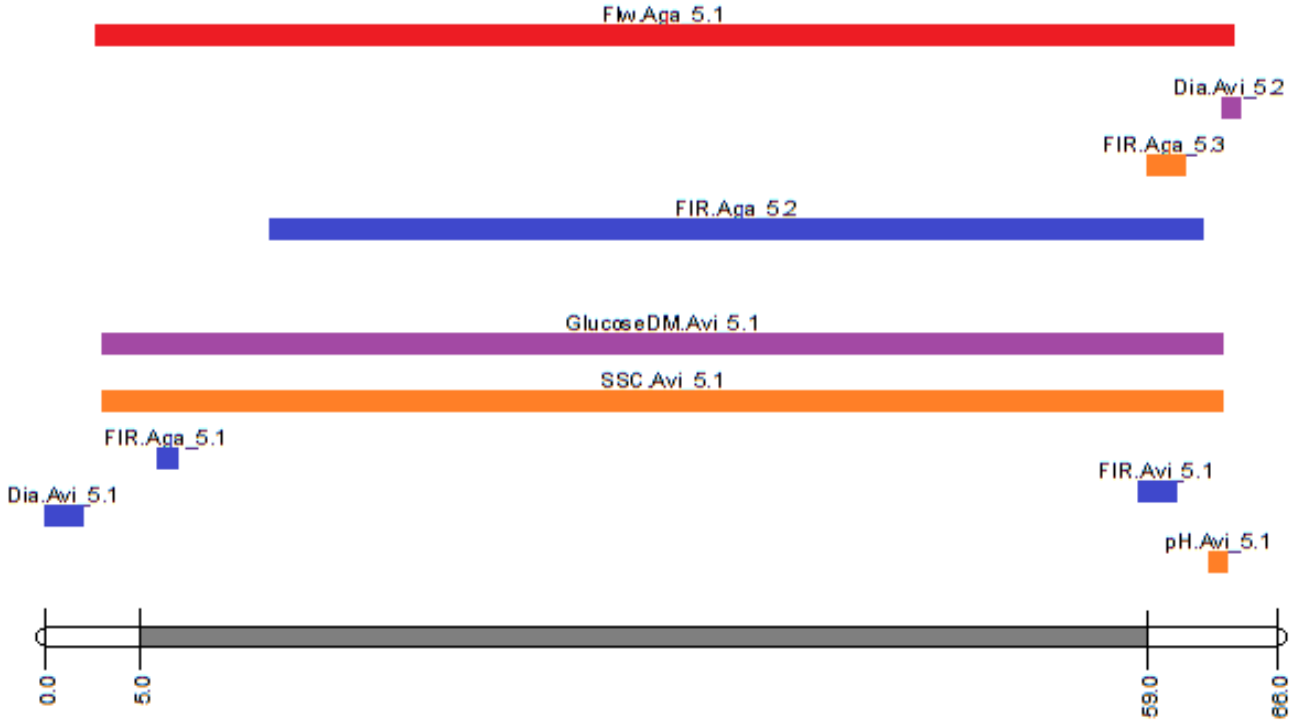

GWA
Leaf.Avi_2.1
Dia.Avi_2.2
Ht.Aga_2.2
Ht.Avi_2.1
Leaf.Aga_2.2
Dia.Aga_2.1
Dia.Avi_2.1
VitCDM.Avi_2.1
SSC.Avi_2.1_2.2
pH.Avi_2.1
pH.Avi_2.2_2.3
FIR.Aga_2.3
FIR.Aga_2.1_2.2
FW.Aga_2.1
Dia.Aga_2.3
FW.Avi_2.2
Dia.Aga_2.2
Leaf.Aga_2.3
Leaf.Avi_2.2
Chr02
Ht.Aga_2.1
Leaf.Aga_2.1
FW.Avi_2.1
FW2.1
Solyc02g080700(L-asc.oxy)
FW2.2
FW2.3
0.0
30.0
43.0
44.9
51.0
52.0
55.5
RIL
FW.Avi_2.1
FW.Aga_2.1
FW.Aga_2.1
FW.Avi_2.1
FIR.Avi_2.1
pH.Avi_2.1
DMW.Avi_2.1
SSC.Avi_2.1
SSC.Avi_2.2
FW.Aga_2.2
FW.Aga_2.3
Nbfruits.Avi_2.1
Leaf.Avi_2.1


0.0
10.5
50.0
64.8
70.5
FW3.2
Dia.Avi_3.1
Dia.Aga_3.1
DMW.Avi_3.1
SSC.Avi_3.1
VitCDM.Avi_3.3
VitCDM.Avi_3.2
VitCDM.Avi_3.1
FW.Avi_3.2
FW.Avi_3.1
FW.Aga_3.1
FW.Avi_3.1
FW.Aga_3.1
Fw.Avi_3.1
Ht.Avi_3.1
Ht.Aga_3.1
SSC.Avi_3.1
MalicFM.Avi_3.1
FIR.Avi_3.1
Leaf.Aga_3.1
Ht.Avi_3.2

Appendix 5


GWA
Flw.Aga_4.1
Nbfruits.Avi_4.1
Leaf.Aga_4.1
pH.Avi_4.2
GlucoseFM.Avi_4.1
GlucoseDM.Avi_4.1
SSC.Avi_4.1
DMW.Avi_4.1
VitCDM.Avi_4.2
VitCDM.Avi_4.1
Dia.Avi_4.1
VitCFM.Avi_4.1
FIR.Avi_4.1
MalicDM.Avi_4.1
pH.Avi_4.1
FructoseFM.Avi_4.1
Chr04
Solyc04g081400(Hexokinase)
Solyc04g081440(Invertase)
0.0
0.5
50.0
65.4
66.5
RIL
FIR.Avi_4.1
FIR.Aga_4.1
VitCFM.Avi_4.1
DMW.Avi_4.1
SSC.Avi_4.1
FW.Avi_4.1
Yield.Avi_4.1
Flw.Avi_4.1
Flw.Aga_4.1
Flw.Aga_4.2
Flw.Aga_4.3
Dia.Avi_4.1
Dia.Aga_4.2

GWA



Chr05

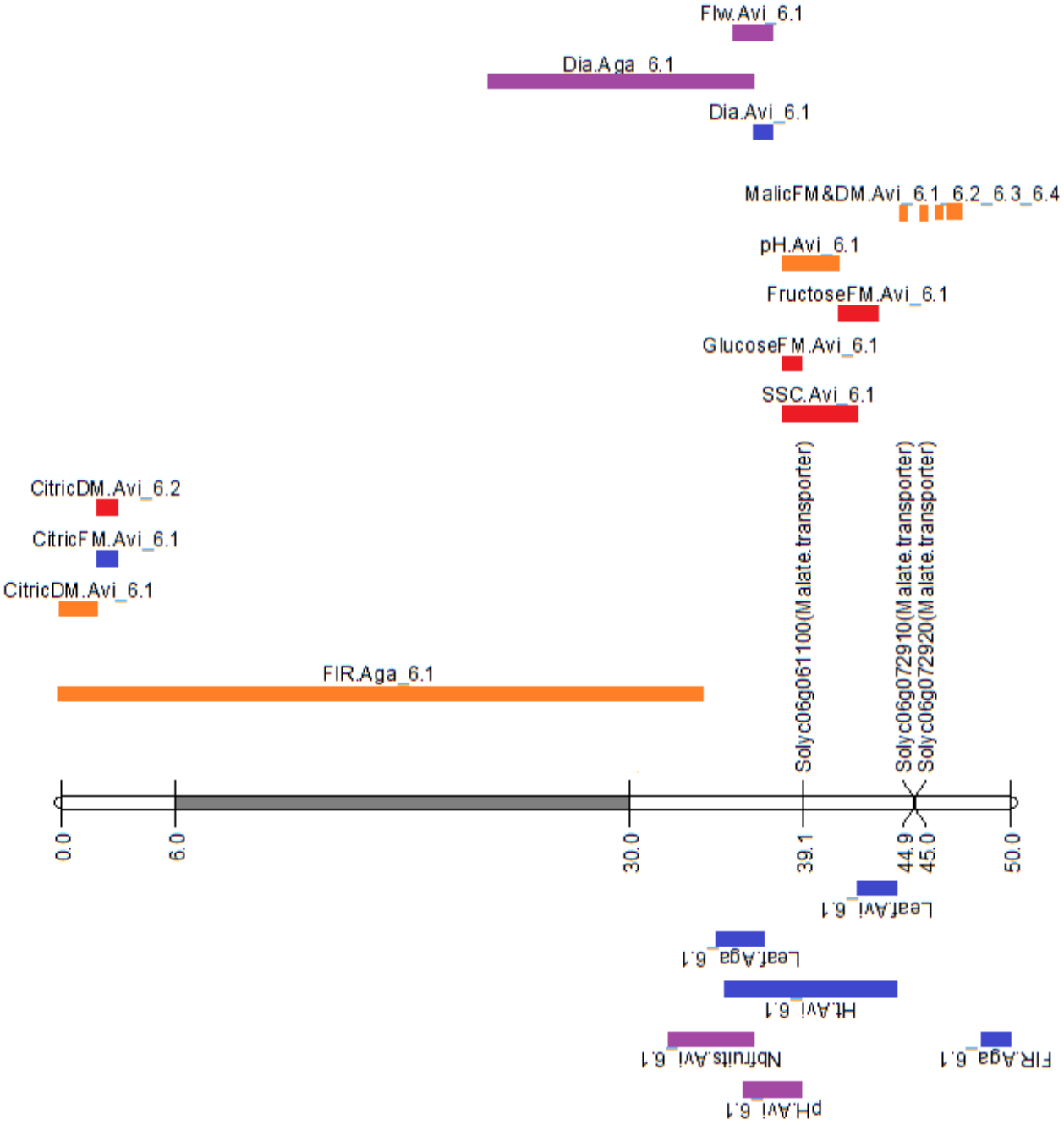
RIL

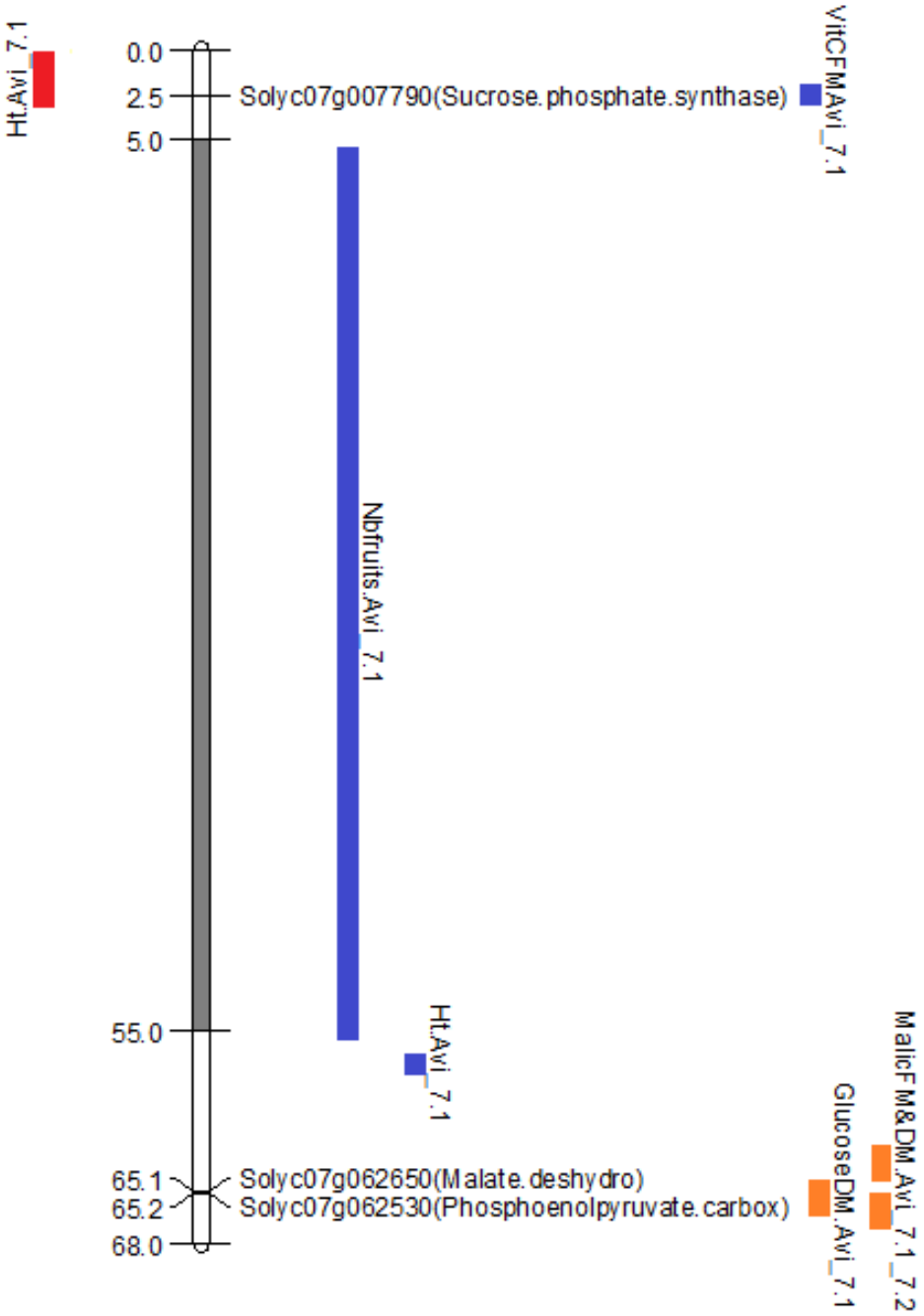
Appendix 5

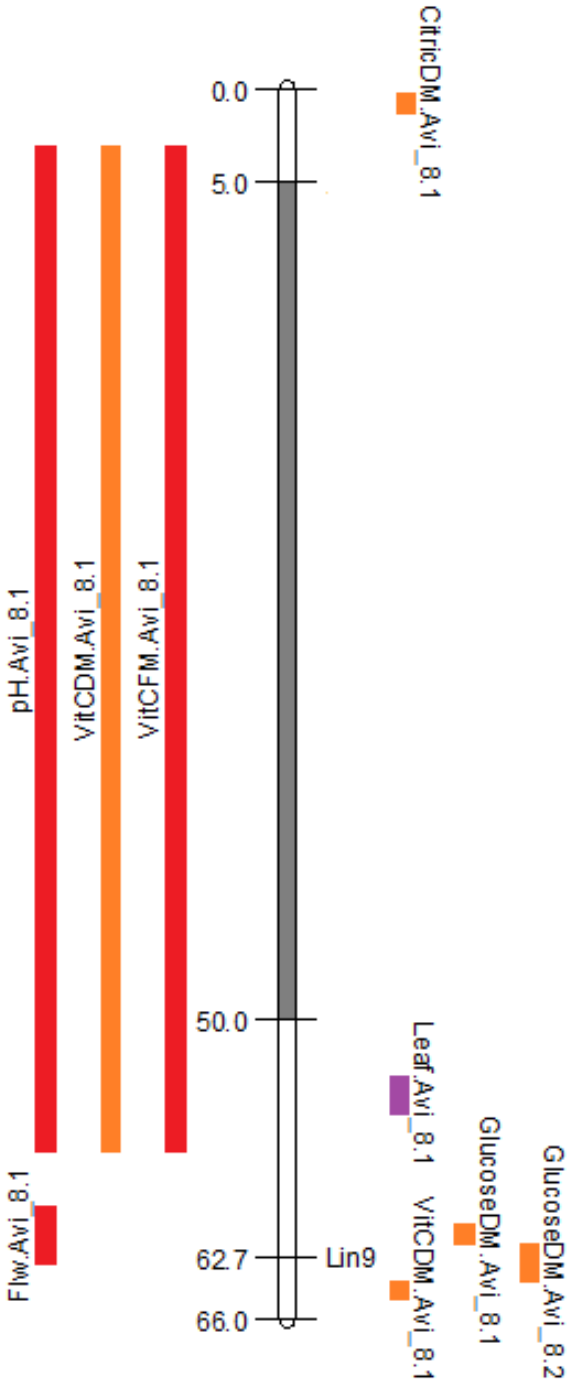
GWA

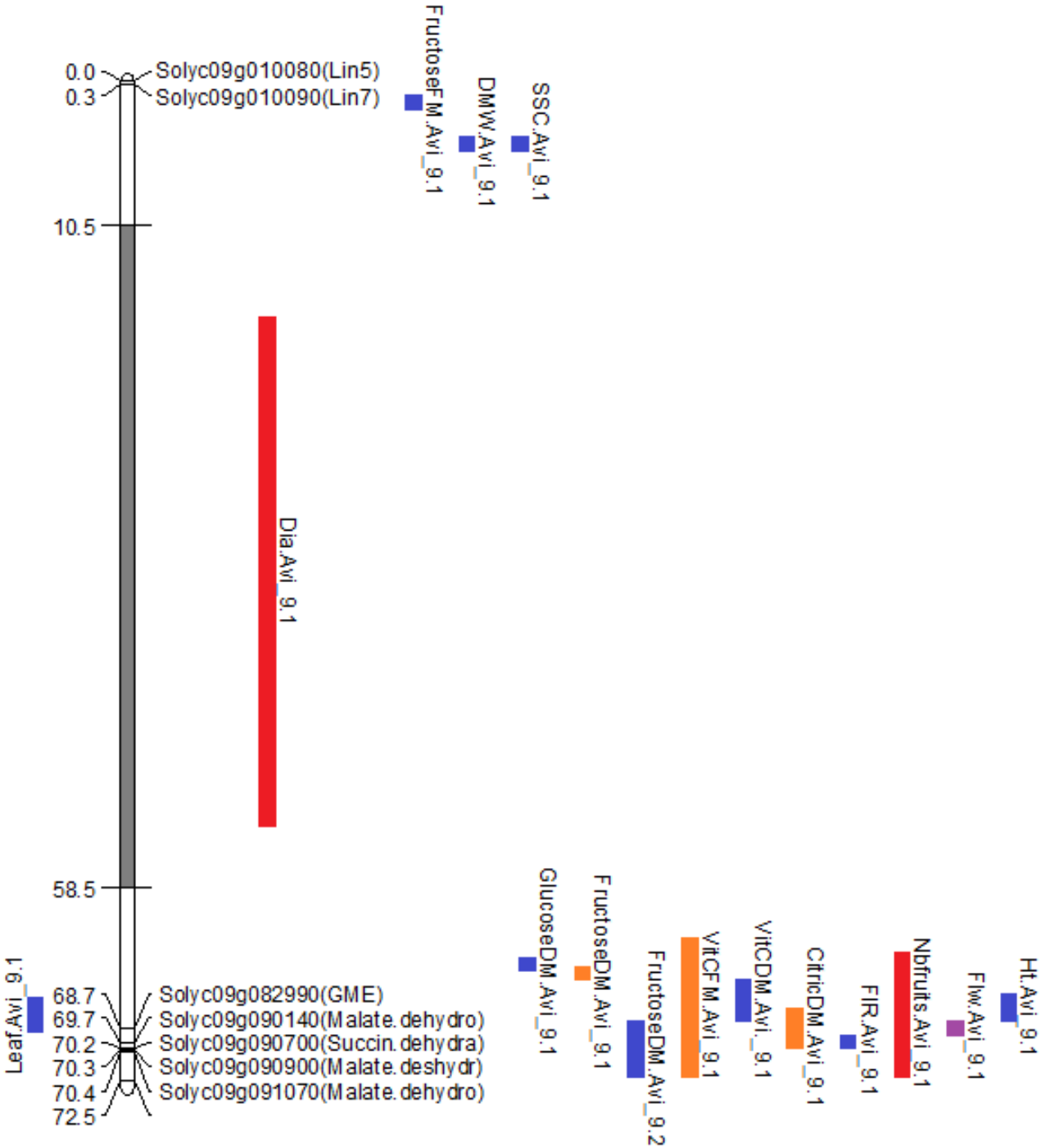
Chr06

RIL



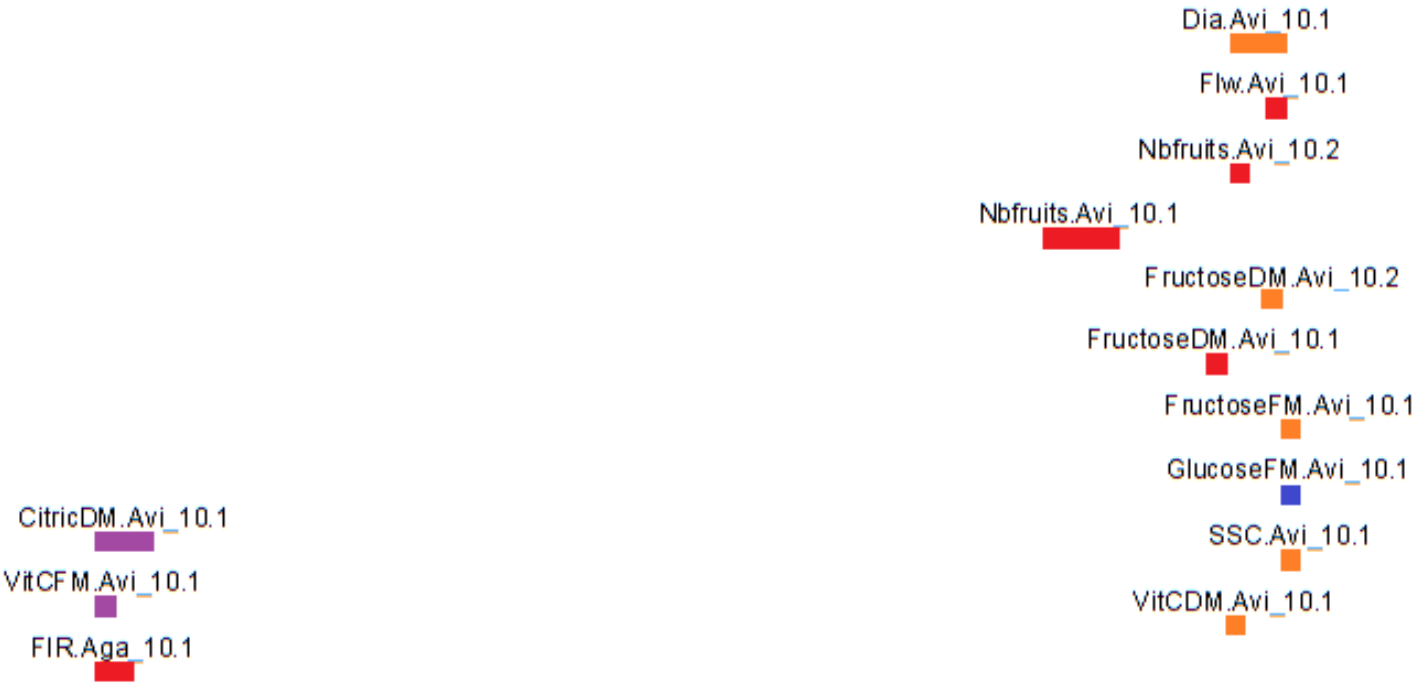

Ht.Avi_7.1
0.0
2.5
5.0
Solyc07g007790(Sucrose.phosphate.synthase)
VitCFM.Avi_7.1
Nbfruits.Avi_7.1
55.0
Ht.Avi_7.1
65.1
65.2
68.0
Solyc07g062650(Malate.deshydro)
Solyc07g062530(Phosphoenolpyruvate.carbox)
GlucoseDM.Avi_7.1
MalicFM&DM.Avi_7.1_7.2


CitricDM.Avi_8.1
0.0
5.0
pH.Avi_8.1
VitCDM.Avi_8.1
VitCFM.Avi_8.1
50.0
Leaf.Avi_8.1
GlucoseDM.Avi_8.1
GlucoseDM.Avi_8.2
Flw.Avi_8.1
62.7
Lin9
VitCDM.Avi_8.1
66.0

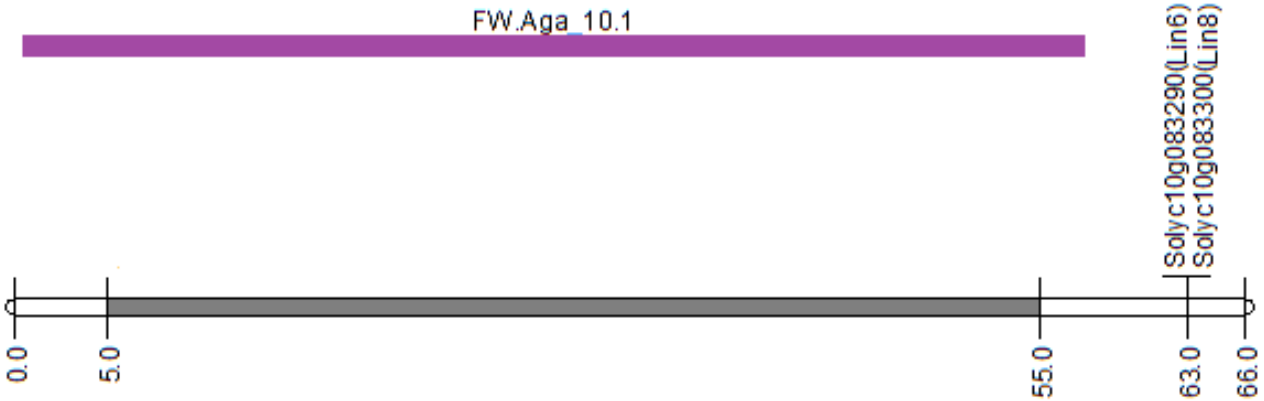

0.0
0.3
10.5
58.5
68.7
69.7
70.2
70.3
70.4
72.5
Solyc09g010080(Lin5)
Solyc09g010090(Lin7)
Solyc09g082990(GME)
Solyc09g090140(Malate.dehydro)
Solyc09g090700(Succin.dehydra)
Solyc09g090900(Malate.deshydr)
Solyc09g091070(Malate.dehydro)
FructoseFM.Avi_9.1
DMW.Avi_9.1
SSC.Avi_9.1
Dia.Avi_9.1
LeafAvi_9.1
GlucoseDM.Avi_9.1
FructoseDM.Avi_9.1
FructoseDM.Avi_9.2
VitCFM.Avi_9.1
VitCDM.Avi._9.1
CitricDM.Avi_9.1
FIR.Avi_9.1
Nbfruits.Avi_9.1
Flw.Avi_9.1
Ht.Avi_9.1

Appendix 5

GWA



Chr10



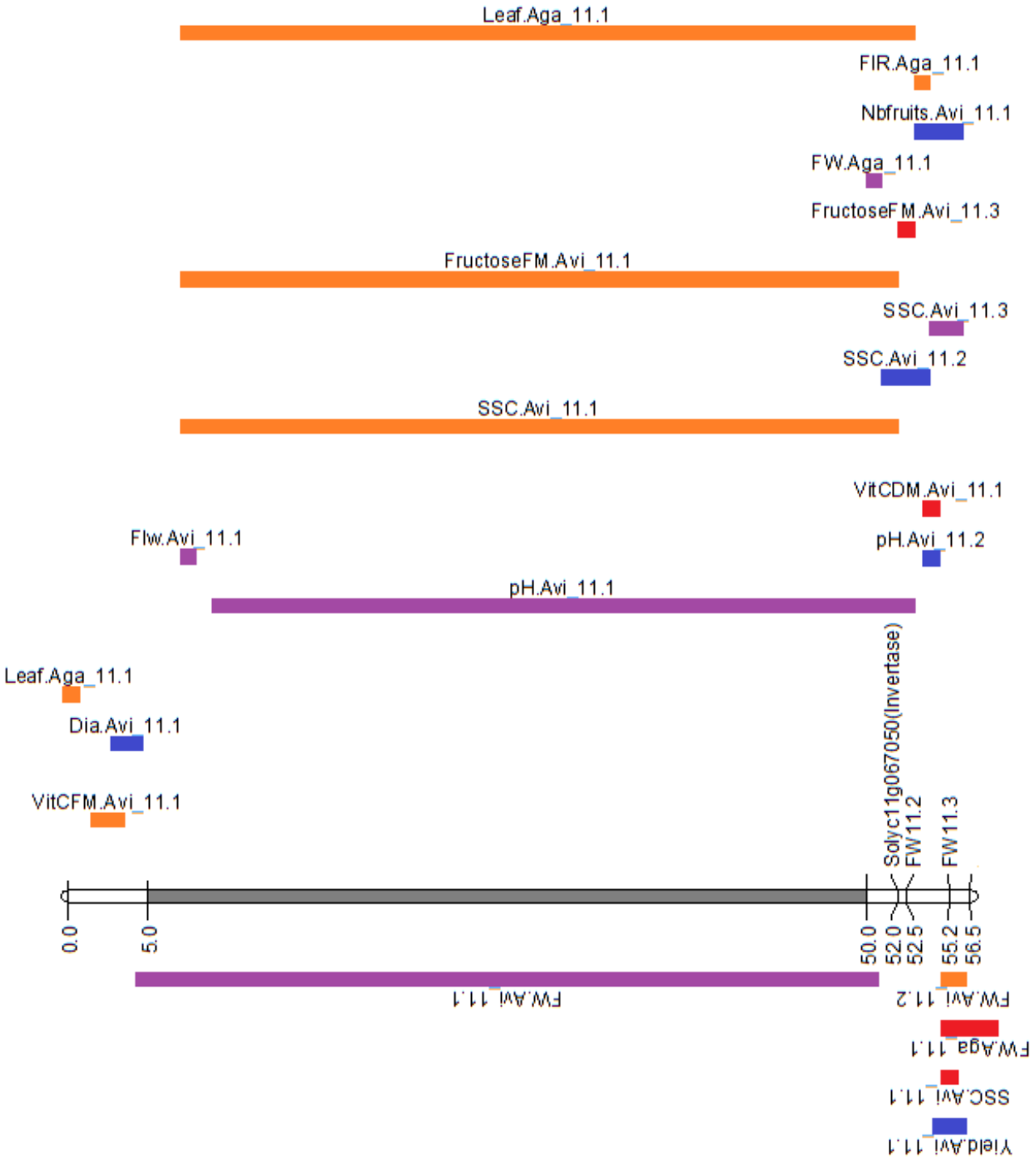
RIL

Appendix 5

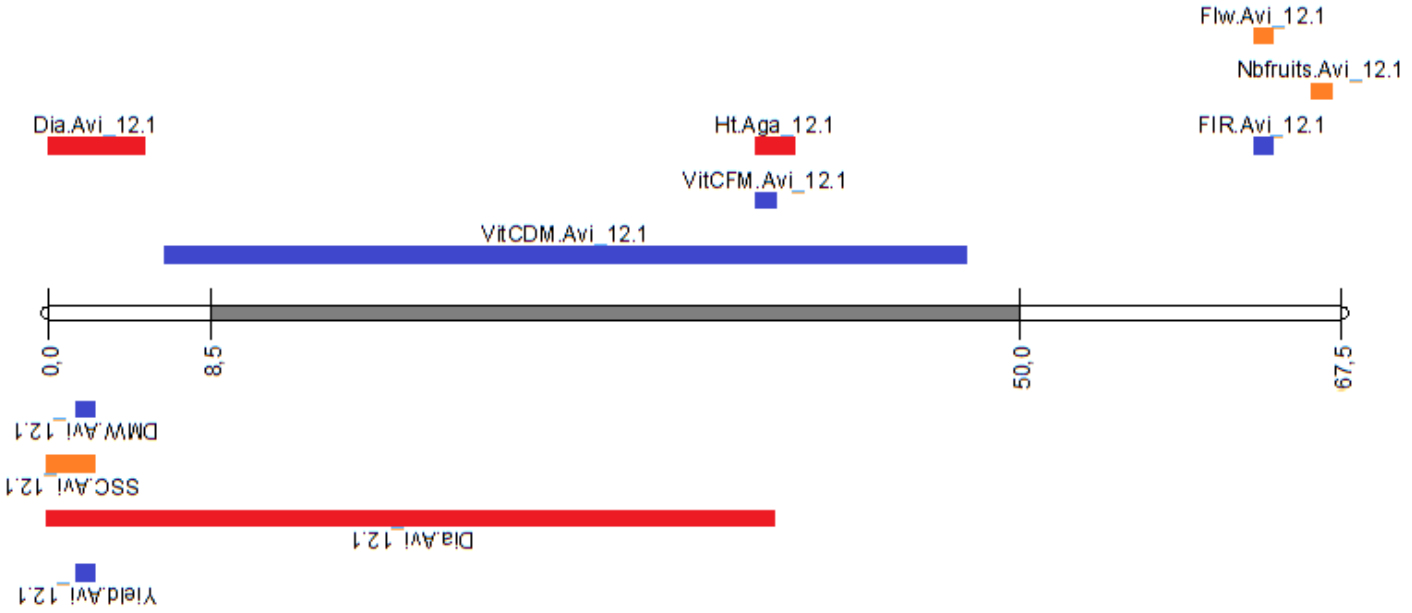
GWA

Chr11

RIL

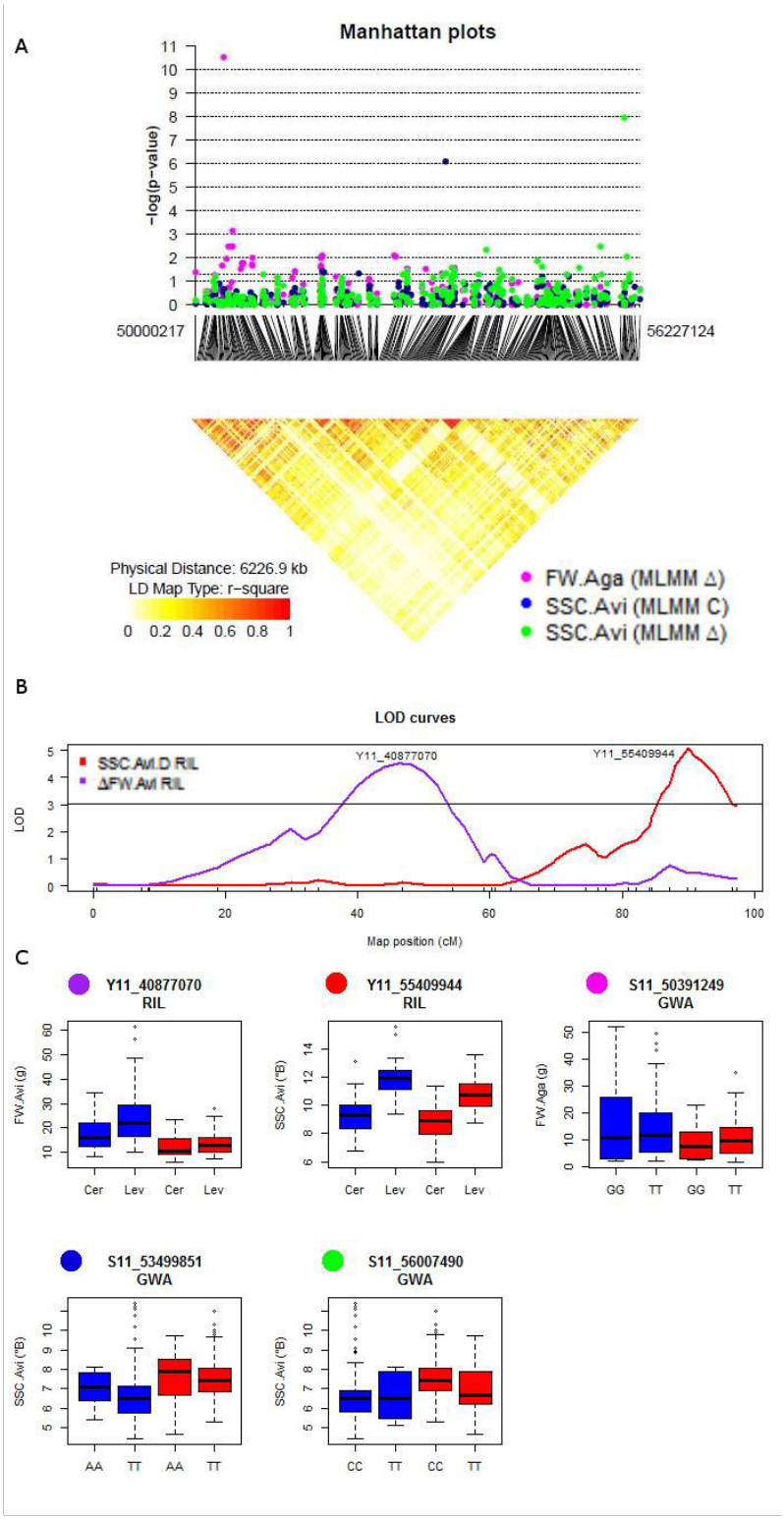


GWA



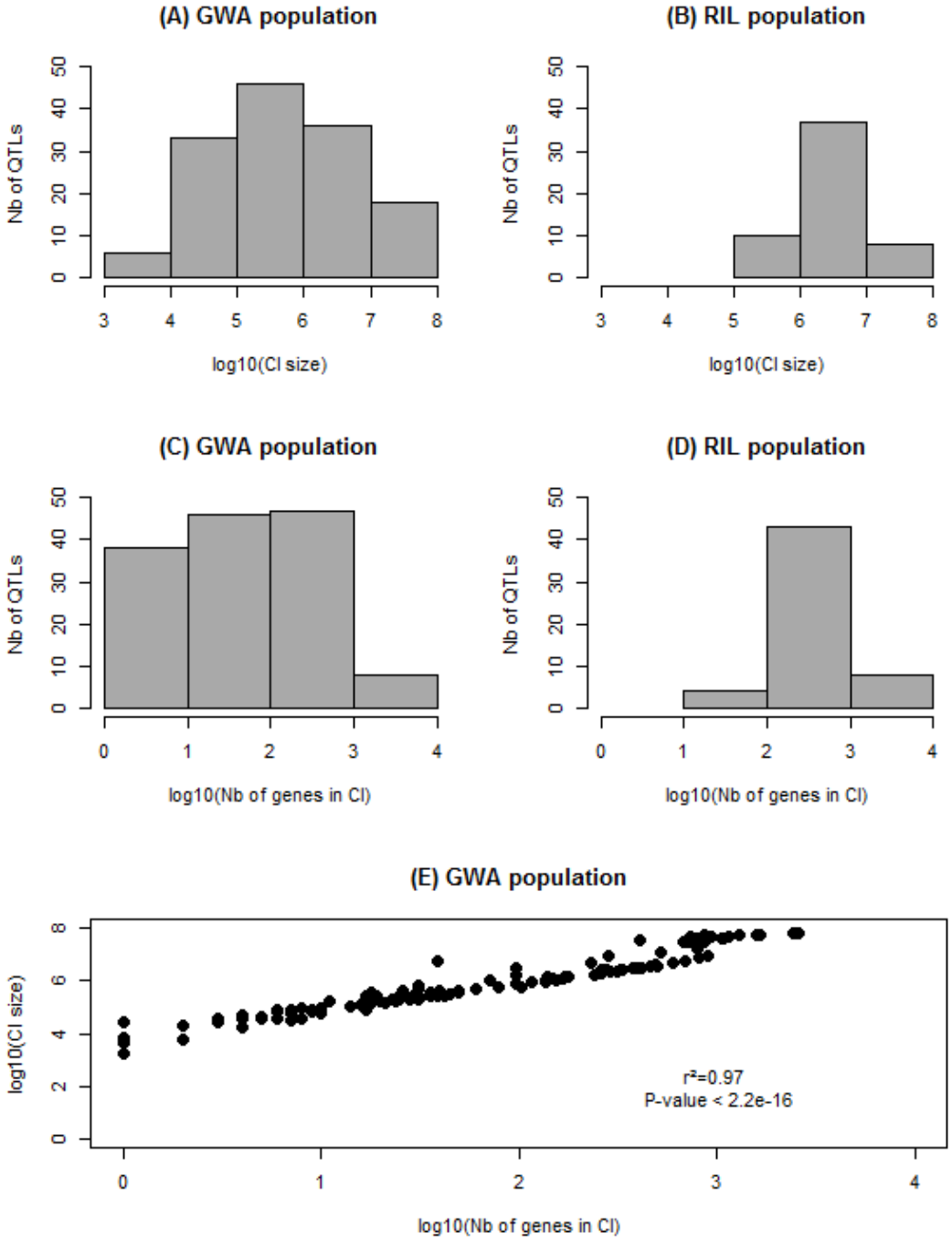
Chr12

RIL

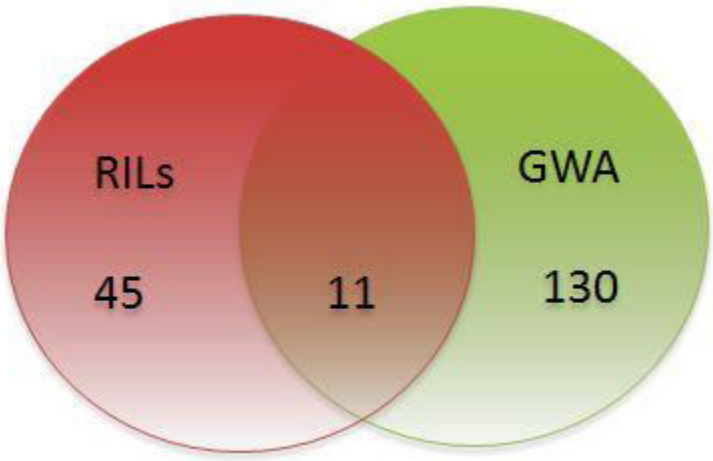


Supplemental Figure 8. Example of colocalisations between GWA and RIL QTLs for soluble solid content (SSC) and fruit fresh weight (FW) on top of chromosome 11.

Legend Supplemental Figure 8. (A) Manhattan plots displaying the –log10(*P-values*) (Y-axis) over genomic positions (X-axis) in a window of 6.23 Mbp corresponding to the genomic region encompassing three QTLs detected for FW.Aga (MLMM Δ, magenta), SSC.Avi (MLMM control condition, dark blue) and SSC.Avi (MLMM Δ, green) on chromosome 11 in the GWA population. *P-values* below 10^{-4} were considered as significant (4 in logit values). The heatmap of the pairwise LD was drawn using *the R package 'snp.plotter'* (Luna and Nicodemus, 2007). (B) Likelihood curves of the LOD score for two QTLs detected for SSC.Avi (Simple Interval Mapping, drought condition, red) and FW.Avi (SIM, Δ, purple) in the RIL population. Marker at the LOD score peak is indicated for each QTL. Distances are expressed in cM (see genetic map in Albert *et al*. (2016)). (C) Allelic effects for the five detected QTLs: Y11_40877070 (RIL, FW.Avi, '*differential*'), Y11_55409944 (RIL, SSC.Avi, '*drought specific*'), S11_50391249 (GWA, FW.Aga, '*antagonist*'), S11_53499851 (GWA, SSC.Avi, '*control specific*') and S11_56007490 (GWA, SSC.Avi, '*differential*'). Blue: Allelic effects under control condition. Red: Allelic effects under drought condition.



Supplemental Figure 9. Confidence interval (CI) sizes and numbers of genes underlying the QTLs in the GWA and RIL populations. Distribution of the CI sizes expressed in bp in the GWA (A) and RIL (B) populations. Distribution of the number of genes underlying QTLs in the GWA (C) and RIL (D) populations. (E) Relation between number of underlying genes and CI sizes in bp for the QTLs detected in the GWA population. The r^2 corresponds to the Spearman correlation between CI sizes and number of underlying genes for the different QTLs.



Supplemental Figure 10. Venn diagram representing common QTLs between the RIL population (linkage mapping) and the GWA population (association mapping). For the comparison, we considered related traits as a single one: pH, acid malic (DM and FM) and acid citric (DM and FM) were grouped, as well as SSC, Glucose (DM and FM) and Fructose (DM and FM). Besides, whatever the QTL type ('specific', 'constitutive' or 'interactive') and the location of the trial (Agadir and Avignon), we considered that a single QTL was present when the CI overlapped between RIL and GWA QTLs.

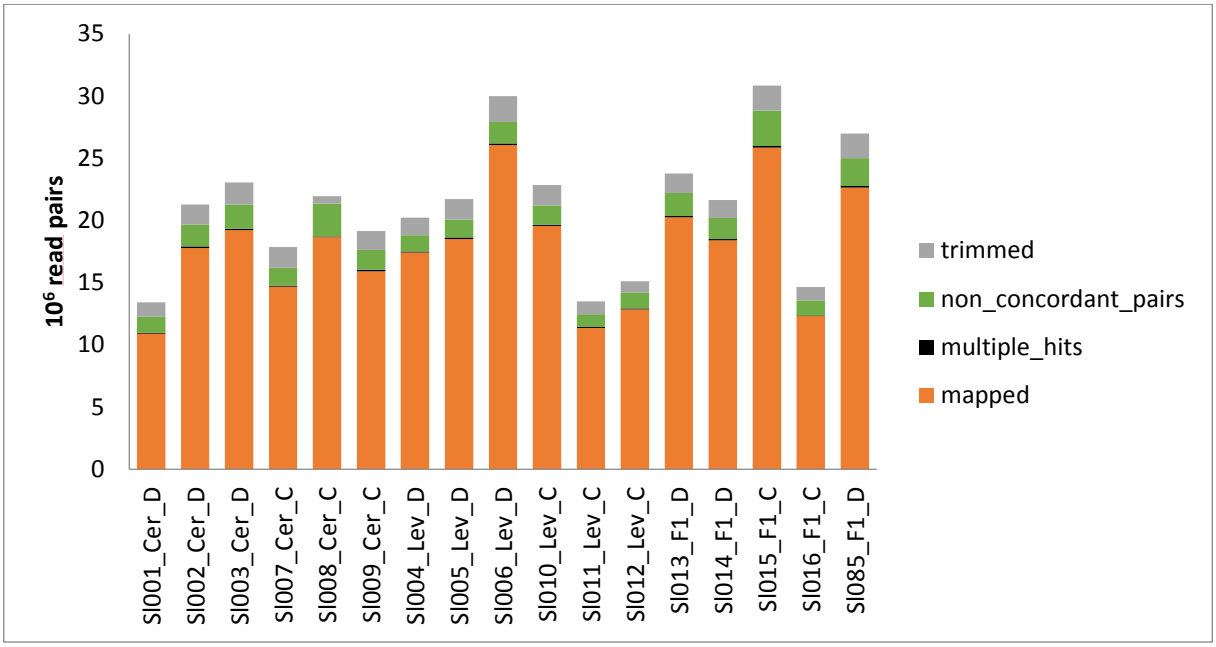
Supplemental Table 4. Correlations between Avignon and Agadir trials. Pearson correlation coefficients between average trait values measured in Avignon and Agadir trials, under the two watering regimes (control and drought), are displayed.

Trait	Control	Drought
Flw	0.70 ***	0.74 ***
Diam	0.51 ***	0.43 ***
Leaf	0.49 ***	0.41 ***
Ht[a]	0.58 ***	0.65 ***
FW[a]	0.96 ***	0.93 ***
FIR	0.57 ***	0.51 ***

[a] Data transformed for skewed distribution

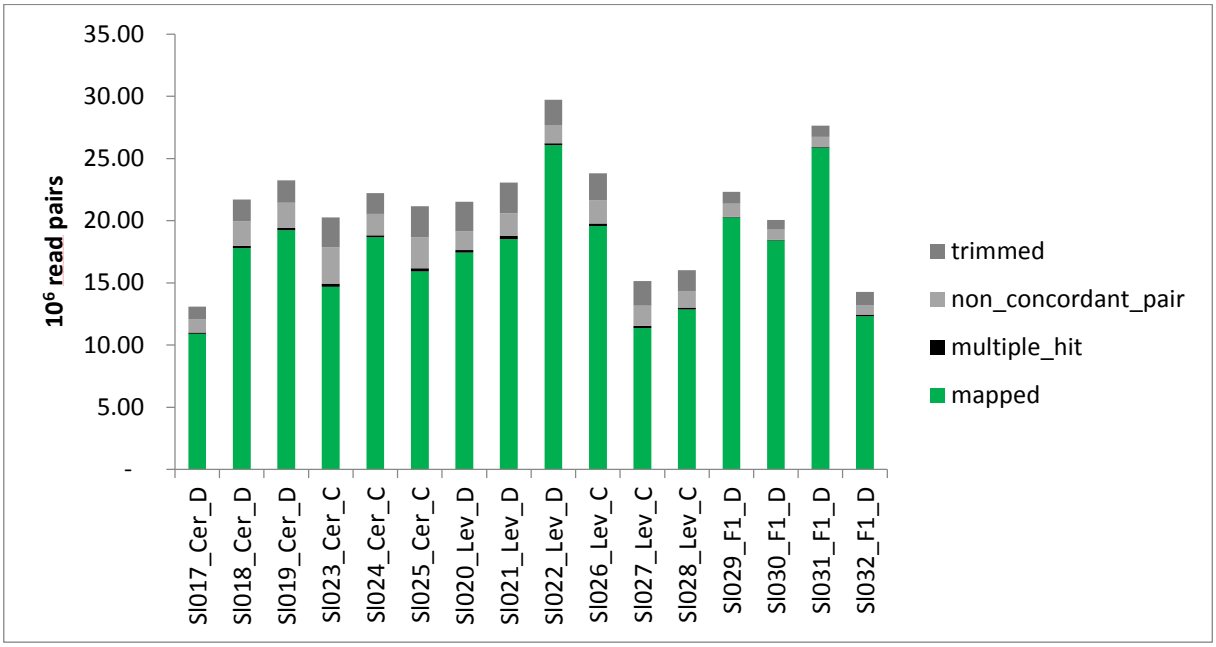
*** shows *P-value* below 0.001, ** between 0.001 and 0.01; and * between 0.01 and 0.05.

Appendix 6: Supplemental figures and tables from chapter 5.



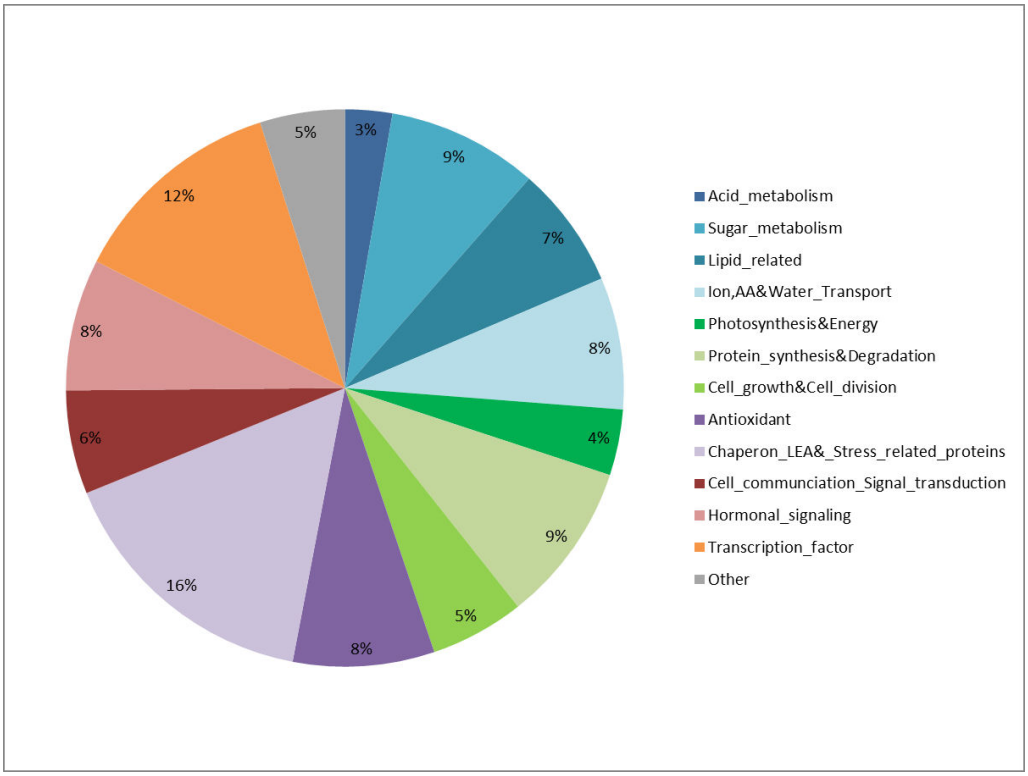
Supplemental Figure 1. Summary of read processing for the fruit libraries. The grey color indicates the reads trimmed out based on the presence of adaptors, poor sequence quality and presence of full read pairs (*Cutadapt* 1.9.1, *Selqual* and *Selpair* programs, see M&M). The green and black colors indicate the non-concordant read pairs and the read pairs with multiple hits that were discarded during read mapping on the tomato genome (tophat2, see M&M). The orange color indicates the read pairs that were considered in the subsequent statistical analysis. In average, 84% of the reads were mapped.

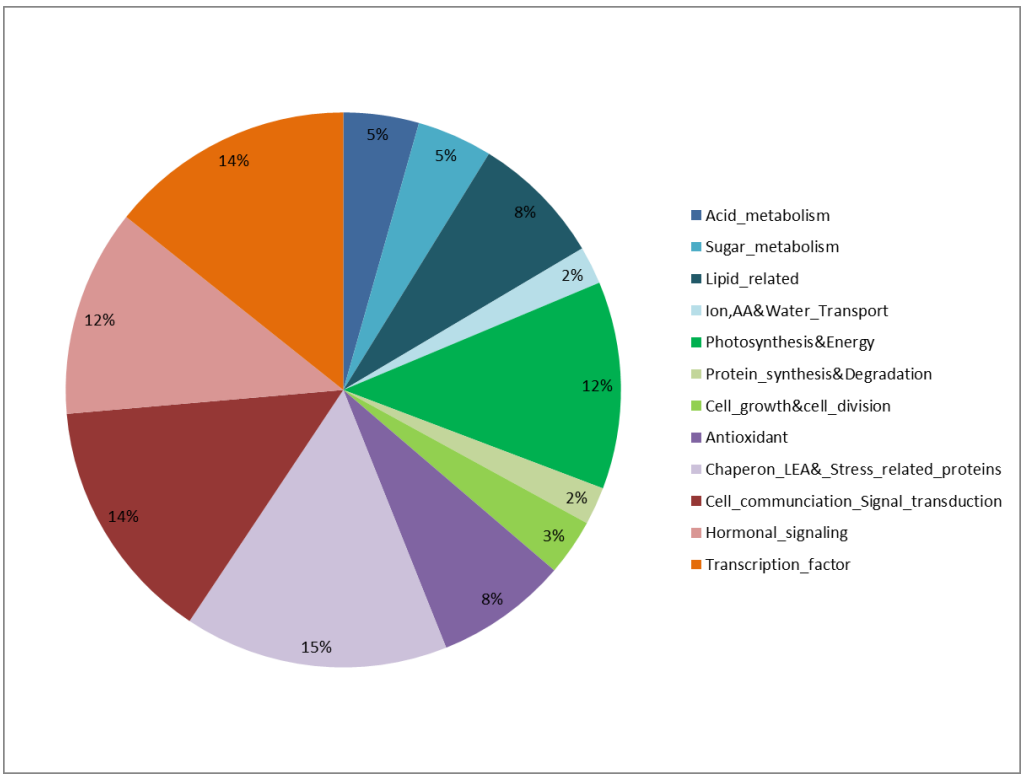
Appendix 6



Supplemental Figure 2. Summary of read processing for the leaf libraries. The dark grey color indicates the reads trimmed out based on the presence of adaptors, poor sequence quality and presence of full read pairs (*Cutadapt* 1.9.1, *Selqual* and *Selpair* programs, see M&M). The light grey and black colors indicate the non-concordant read pairs and the read pairs with multiple hits that were discarded during read mapping on the tomato genome (tophat2, see M&M). The dark green color indicates the read pairs that were considered in the subsequent statistical analysis. In average, 84 % of the reads were mapped.

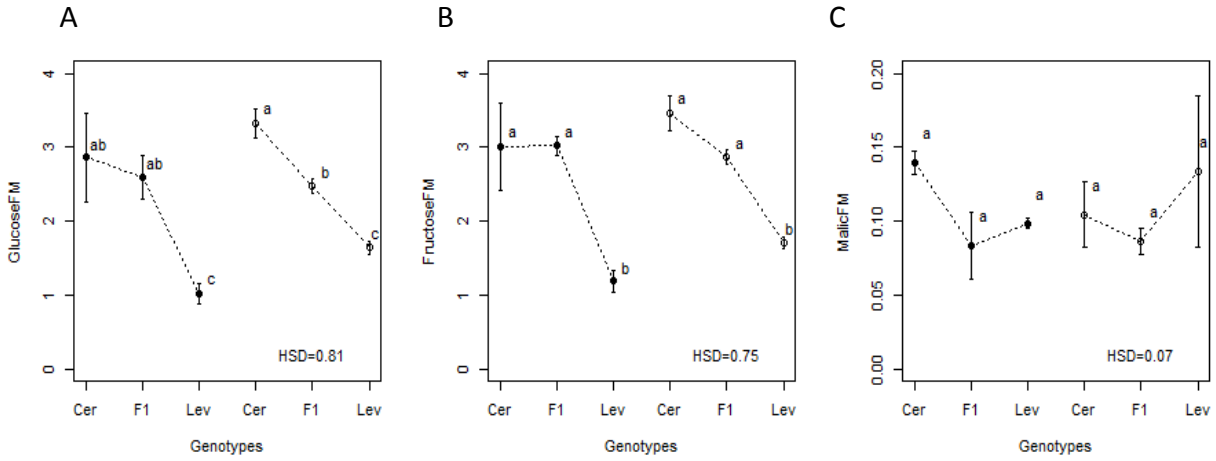
Appendix 6



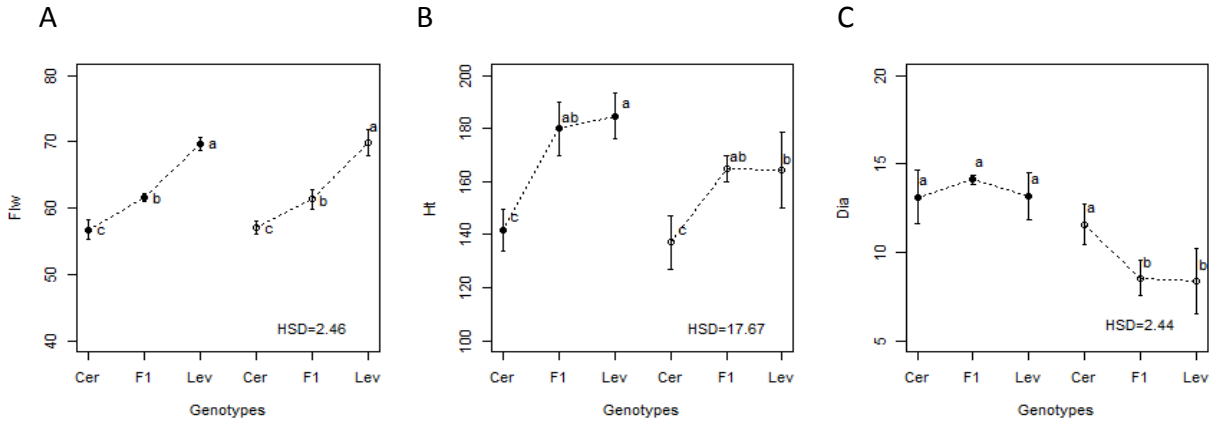


Supplemental Figure 3. Functional categories of the selected genes for quantitative real time microfluidigm PCR assessment in fruits (above, 183 genes) and leaves (below, 91 genes) of the RILs, respectively.

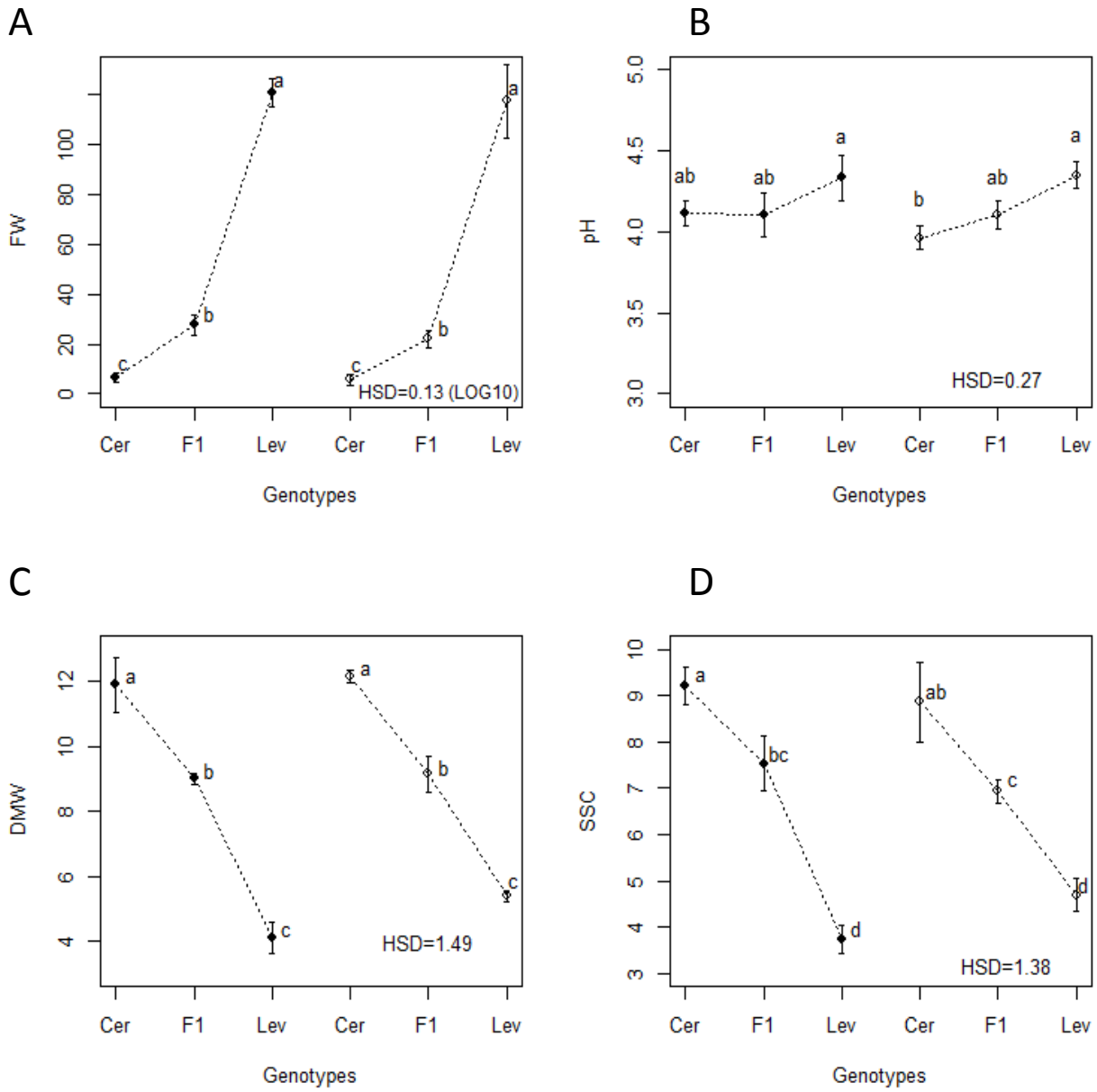
Appendix 6



Supplemental Figure 4: Inheritance of glucose (A), fructose (B) and malic acid (C) contents in mature fruits under both watering conditions. Watering conditions are indicated with different symbols: full symbol for control condition (left), empty symbol for drought condition (right). Letters indicates significantly different mean values according to Tukey's tests (*P-value* < 0.05).

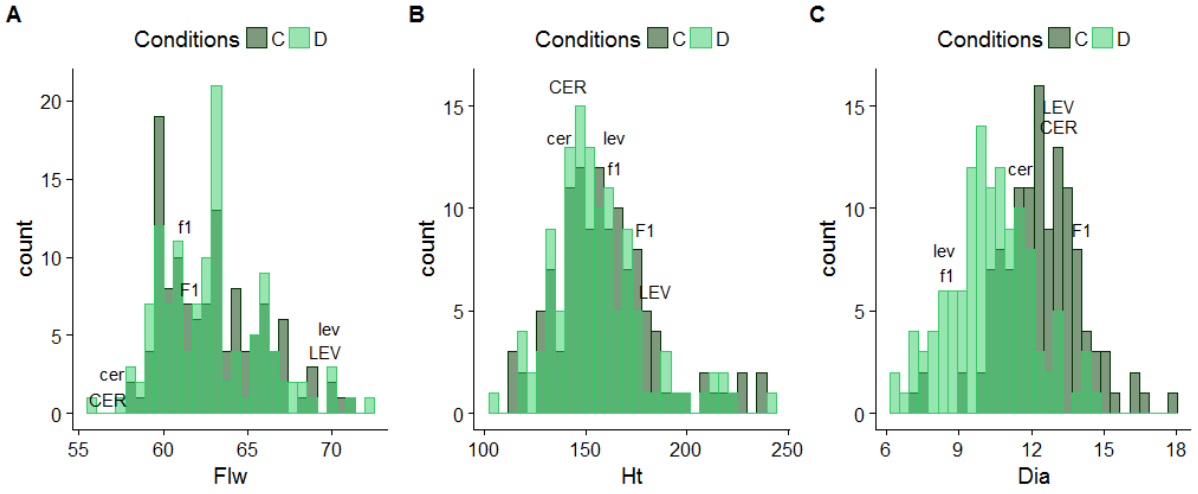


Supplemental Figure 5. Inheritance of Flw (A), Ht (B) and Dia (C) under both watering conditions. Watering conditions are indicated with different symbols: full symbol for control condition (left), empty symbol for drought condition (right). Letters indicates significantly different mean values according to Tukey's tests (*P-value* < 0.05).



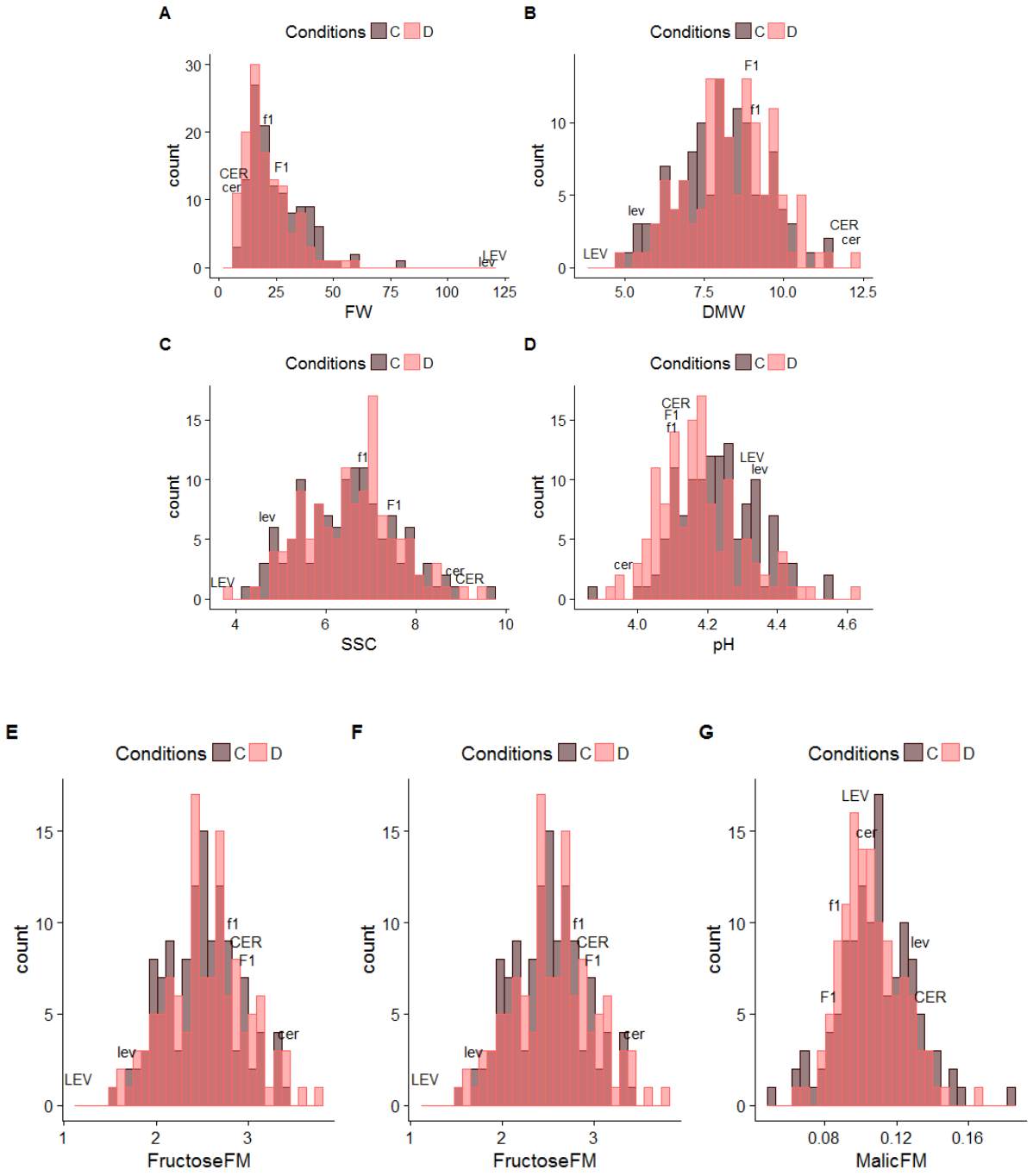
Supplemental Figure 6. Inheritance of FW (A), pH (B), DMW (C) and SSC (D) under both watering conditions. Watering conditions are indicated with different symbols: full symbol for control condition (left), empty symbol for drought condition (right). Letters indicates significantly different mean values according to Tukey's tests (*P-value* < 0.05).

Appendix 6



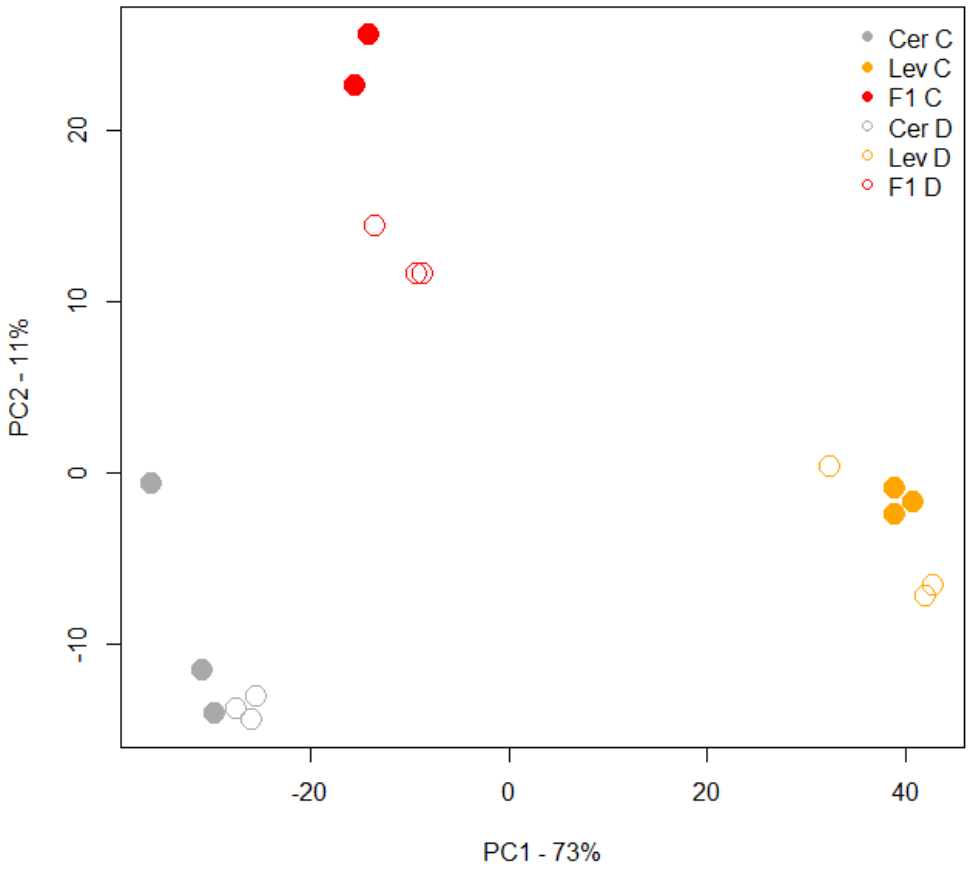
Supplemental Figure 7. Distribution of the average Flw (A), Ht (B) and Dia (C) in the recombinant inbred lines (RILs) grown under two watering regimes. Dark green indicates trait values under control condition and light green trait values under drought condition. The parental and hybrid F1 mean values are indicated: uppercase letters for mean values in control condition and lowercase letters for mean values in drought condition.

Appendix 6



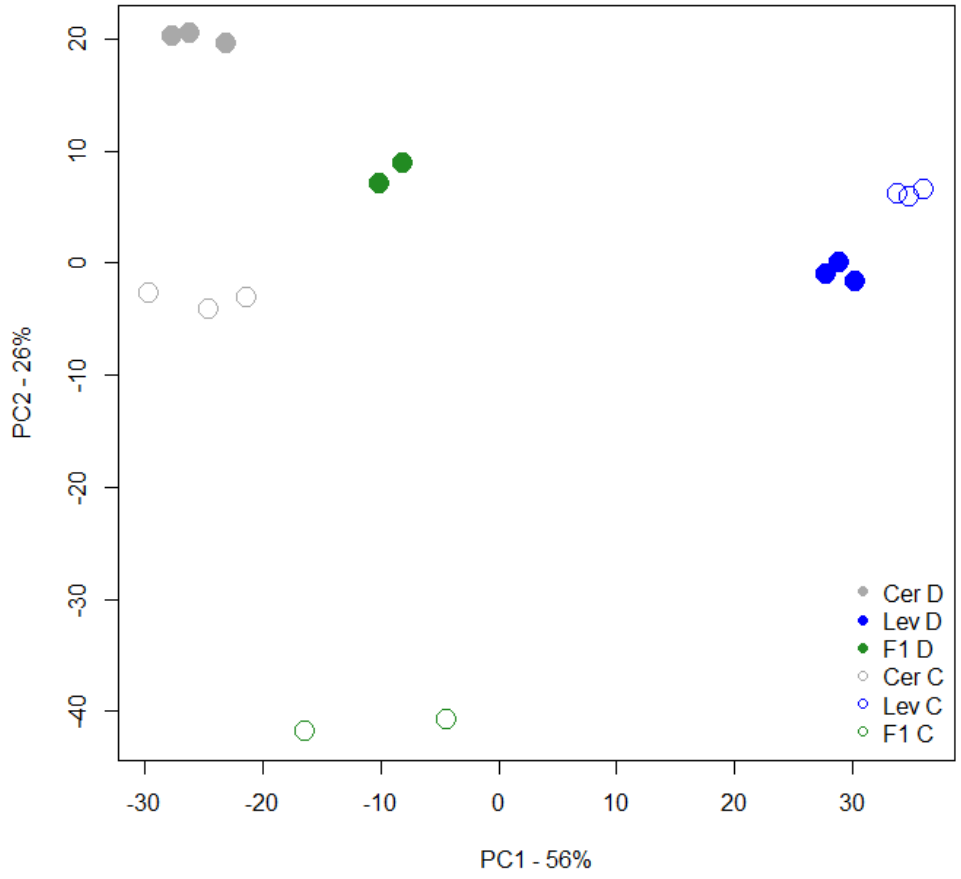
Supplemental Figure 8. Distribution of the average FW (A), DMW (B), SSC (C), pH (D), FructoseFM (E), GlucoseFM (F), and MalicFM (G) in the recombinant inbred lines (RILs) grown under two watering regimes. Dark red indicates trait values under control condition and light red trait values under drought condition. The parental and hybrid F1 mean values are indicated: uppercase letters for mean values in control condition and lowercase letters for mean values in drought condition.

Appendix 6



Supplemental Figure 9. Principal Component Analysis (PCA) of gene expression data in the fruits measured through RNA sequencing. Normalized counts were transformed using the 'regularized log' transformation of the *DESeq2* package.

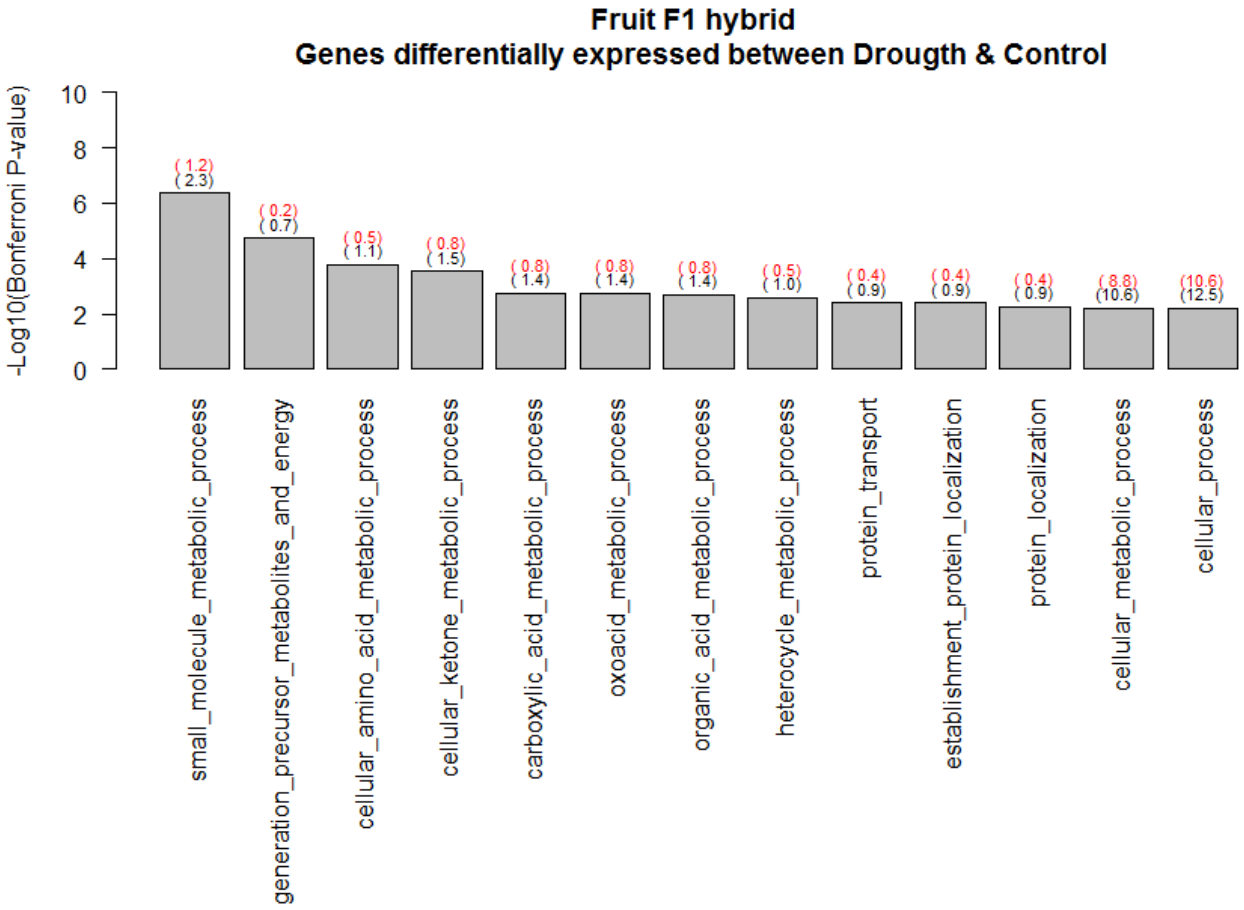
Appendix 6



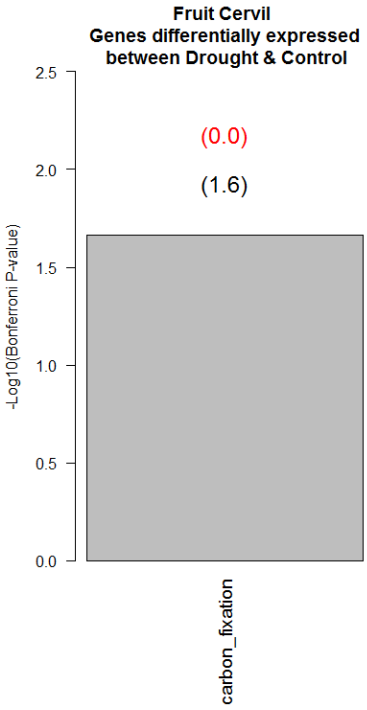
Supplemental Figure 10. Principal Component Analysis (PCA) of gene expression data in the leaves measured through RNA sequencing. Normalized counts were transformed using the 'regularized log' transformation of the *DESeq2* package.

Appendix 6

A

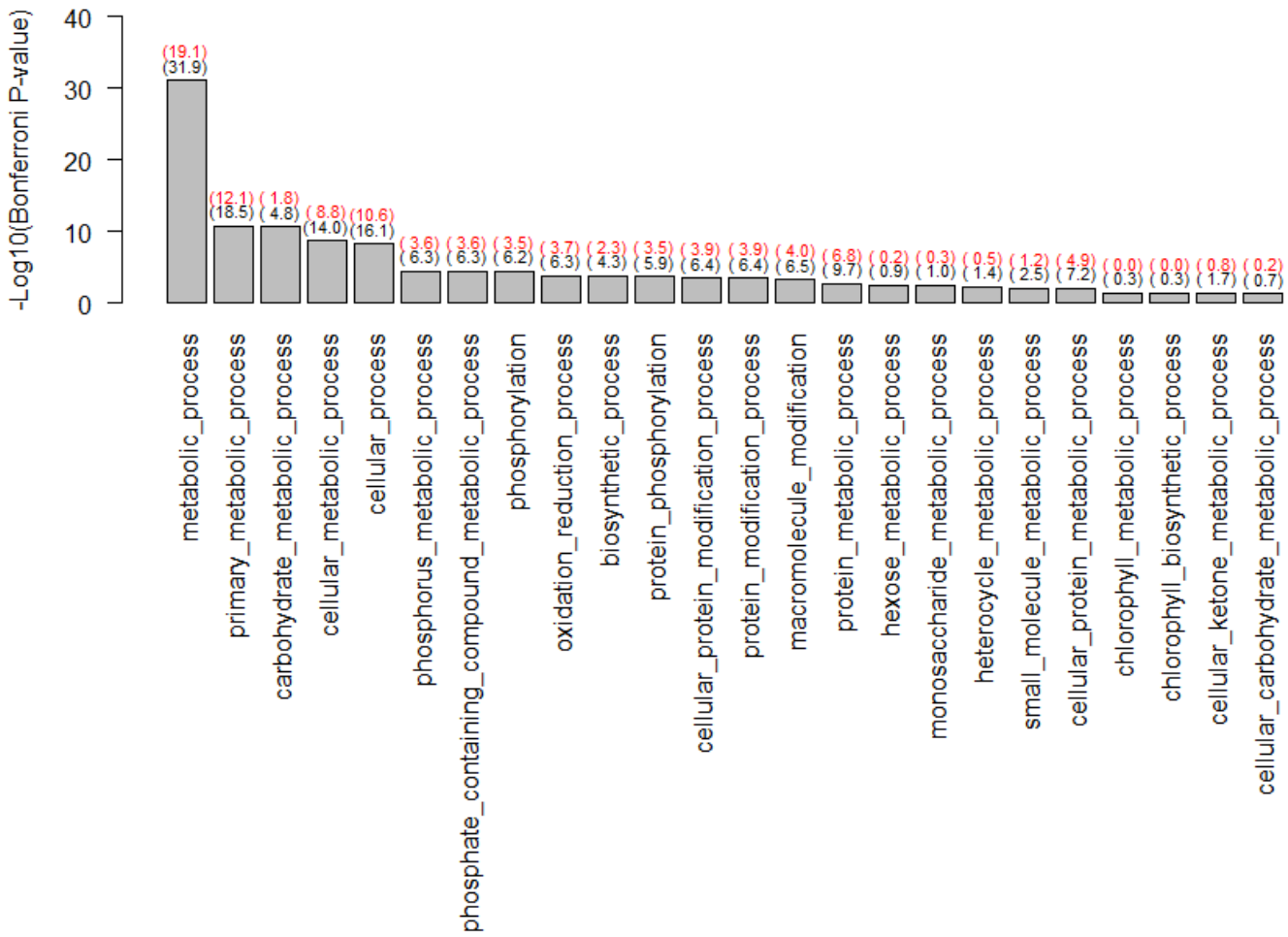


B



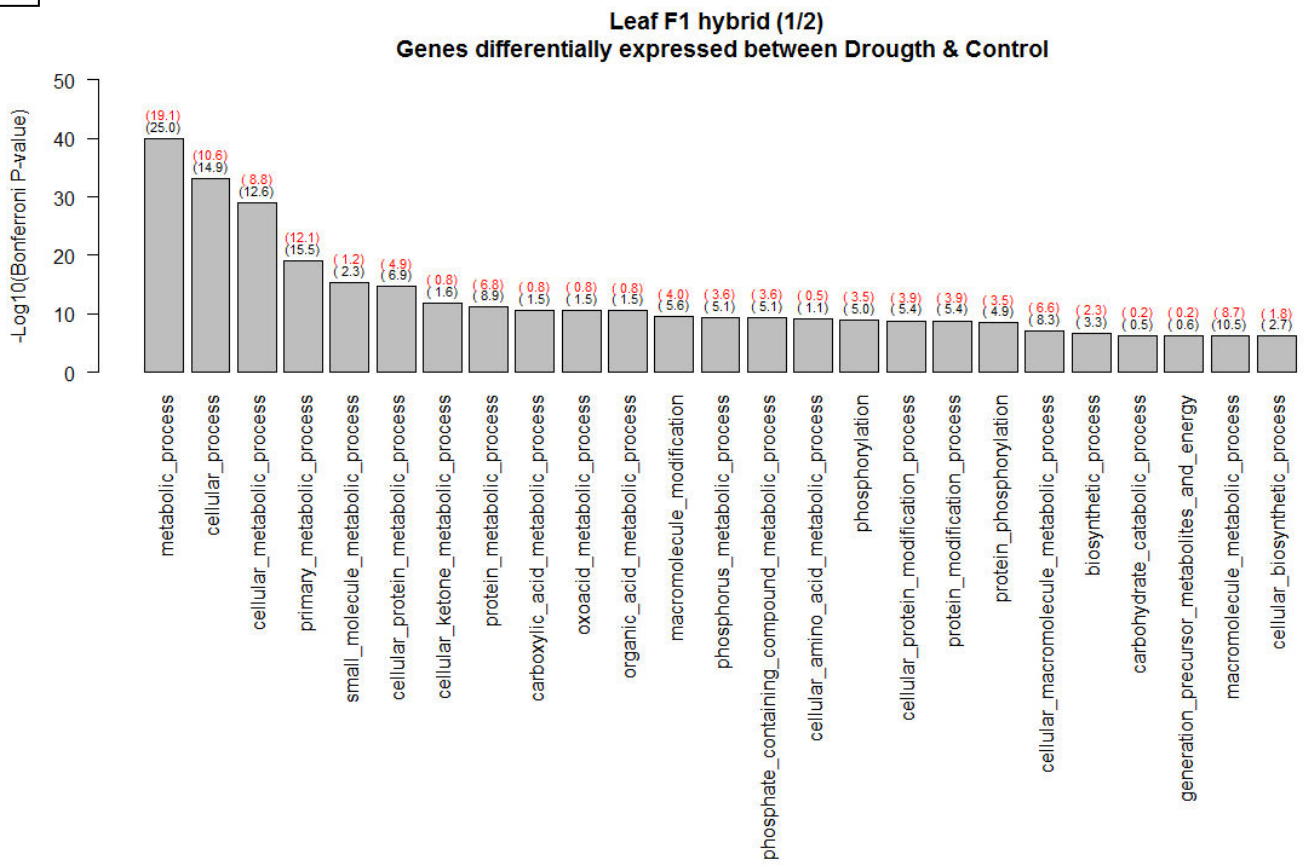
Appendix 6

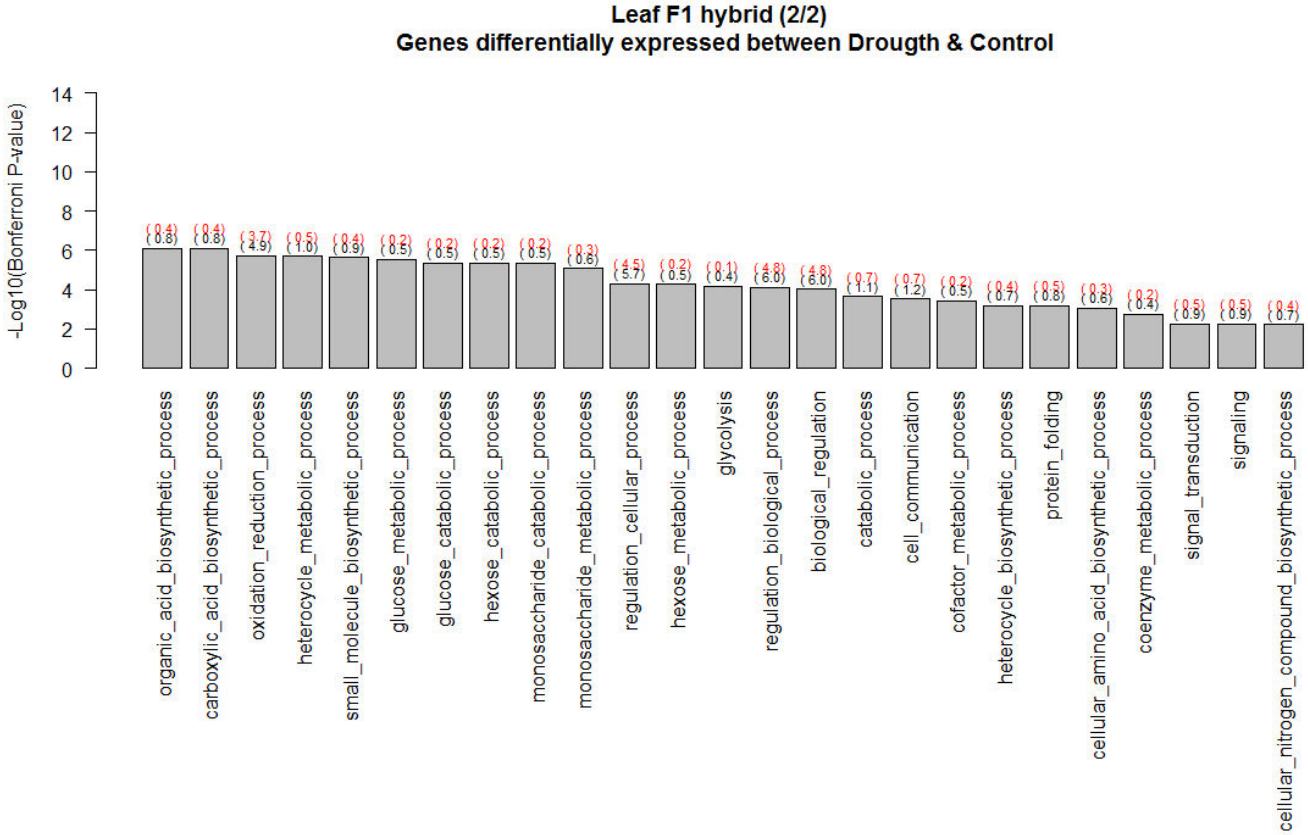
C


Fruit Levovil
Genes differentially expressed between Drought & Control
-Log10(Bonferroni P-value)
0
10
20
30
40
(19.1) (31.9)
(12.1) (18.5)
(1.8) (4.8)
(8.8) (14.0)
(10.6) (16.1)
(3.6) (6.3)
(3.6) (6.3)
(3.5) (6.2)
(3.7) (6.3)
(2.3) (4.3)
(3.5) (5.9)
(3.9) (6.4)
(3.9) (6.4)
(4.0) (6.5)
(6.8) (9.7)
(0.2) (0.9)
(0.3) (1.0)
(0.5) (1.4)
(1.2) (2.5)
(4.9) (7.2)
(0.0) (0.3)
(0.0) (0.3)
(0.8) (1.7)
(0.2) (0.7)
metabolic_process
primary_metabolic_process
carbohydrate_metabolic_process
cellular_metabolic_process
cellular_process
phosphorus_metabolic_process
phosphate_containing_compound_metabolic_process
phosphorylation
oxidation_reduction_process
biosynthetic_process
protein_phosphorylation
cellular_protein_modification_process
protein_modification_process
macromolecule_modification
protein_metabolic_process
hexose_metabolic_process
monosaccharide_metabolic_process
heterocycle_metabolic_process
small_molecule_metabolic_process
cellular_protein_metabolic_process
chlorophyll_metabolic_process
chlorophyll_biosynthetic_process
cellular_ketone_metabolic_process
cellular_carbohydrate_metabolic_process

Appendix 6

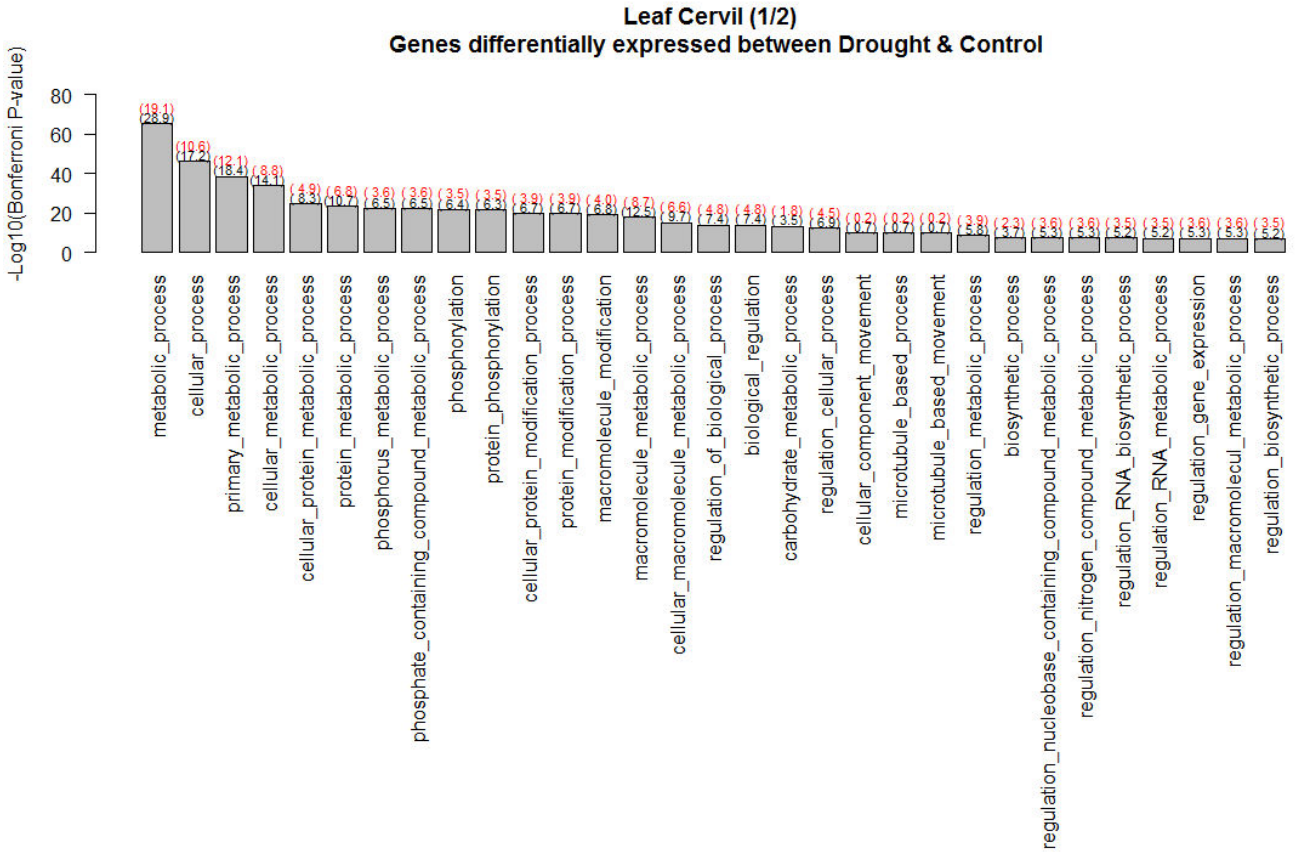
D

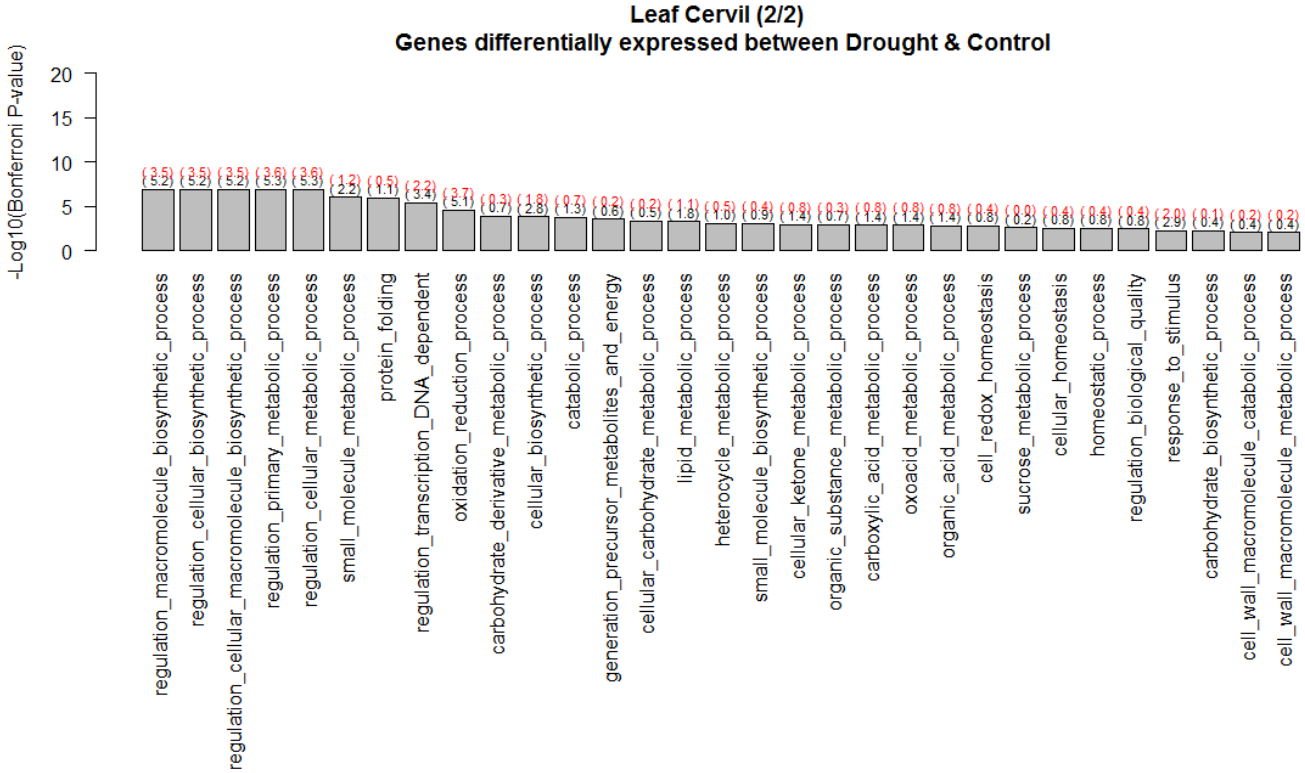

Leaf F1 hybrid (1/2)
Genes differentially expressed between Drougth & Control
-Log10(Bonferroni P-value)
0
10
20
30
40
50
(19.1) (25.0)
(10.6) (14.9)
(8.8) (12.6)
(12.1) (15.5)
(1.2) (2.3)
(4.9) (6.9)
(0.8) (1.6)
(6.8) (8.9)
(0.8) (1.5)
(0.8) (1.5)
(0.8) (1.5)
(4.0) (5.6)
(3.6) (5.1)
(3.6) (5.1)
(0.5) (1.1)
(3.5) (5.0)
(3.9) (5.4)
(3.9) (5.4)
(3.5) (4.9)
(6.6) (8.3)
(2.3) (3.3)
(0.2) (0.5)
(0.2) (0.6)
(8.7) (10.5)
(1.8) (2.7)
metabolic_process
cellular_process
cellular_metabolic_process
primary_metabolic_process
small_molecule_metabolic_process
cellular_protein_metabolic_process
cellular_ketone_metabolic_process
protein_metabolic_process
carboxylic_acid_metabolic_process
oxoacid_metabolic_process
organic_acid_metabolic_process
macromolecule_modification
phosphorus_metabolic_process
phosphate_containing_compound_metabolic_process
cellular_amino_acid_metabolic_process
phosphorylation
cellular_protein_modification_process
protein_modification_process
protein_phosphorylation
cellular_macromolecule_metabolic_process
biosynthetic_process
carbohydrate_catabolic_process
generation_precursor_metabolites_and_energy
macromolecule_metabolic_process
cellular_biosynthetic_process


Leaf F1 hybrid (2/2)
Genes differentially expressed between Drougth & Control
-Log10(Bonferroni P-value)
0
2
4
6
8
10
12
14
(0.4) (0.8)
(0.4) (0.8)
(3.7) (4.9)
(0.5) (1.0)
(0.4) (0.9)
(0.2) (0.5)
(0.2) (0.5)
(0.2) (0.5)
(0.2) (0.5)
(0.3) (0.6)
(4.5) (5.7)
(0.2) (0.5)
(0.1) (0.4)
(4.8) (6.0)
(4.8) (6.0)
(0.7) (1.1)
(0.7) (1.2)
(0.2) (0.5)
(0.4) (0.7)
(0.5) (0.8)
(0.3) (0.6)
(0.2) (0.4)
(0.5) (0.9)
(0.5) (0.9)
(0.4) (0.7)
organic_acid_biosynthetic_process
carboxylic_acid_biosynthetic_process
oxidation_reduction_process
heterocycle_metabolic_process
small_molecule_biosynthetic_process
glucose_metabolic_process
glucose_catabolic_process
hexose_catabolic_process
monosaccharide_catabolic_process
monosaccharide_metabolic_process
regulation_cellular_process
hexose_metabolic_process
glycolysis
regulation_biological_process
biological_regulation
catabolic_process
cell_communication
cofactor_metabolic_process
heterocycle_biosynthetic_process
protein_folding
cellular_amino_acid_biosynthetic_process
coenzyme_metabolic_process
signal_transduction
signaling
cellular_nitrogen_compound_biosynthetic_process

Appendix 6

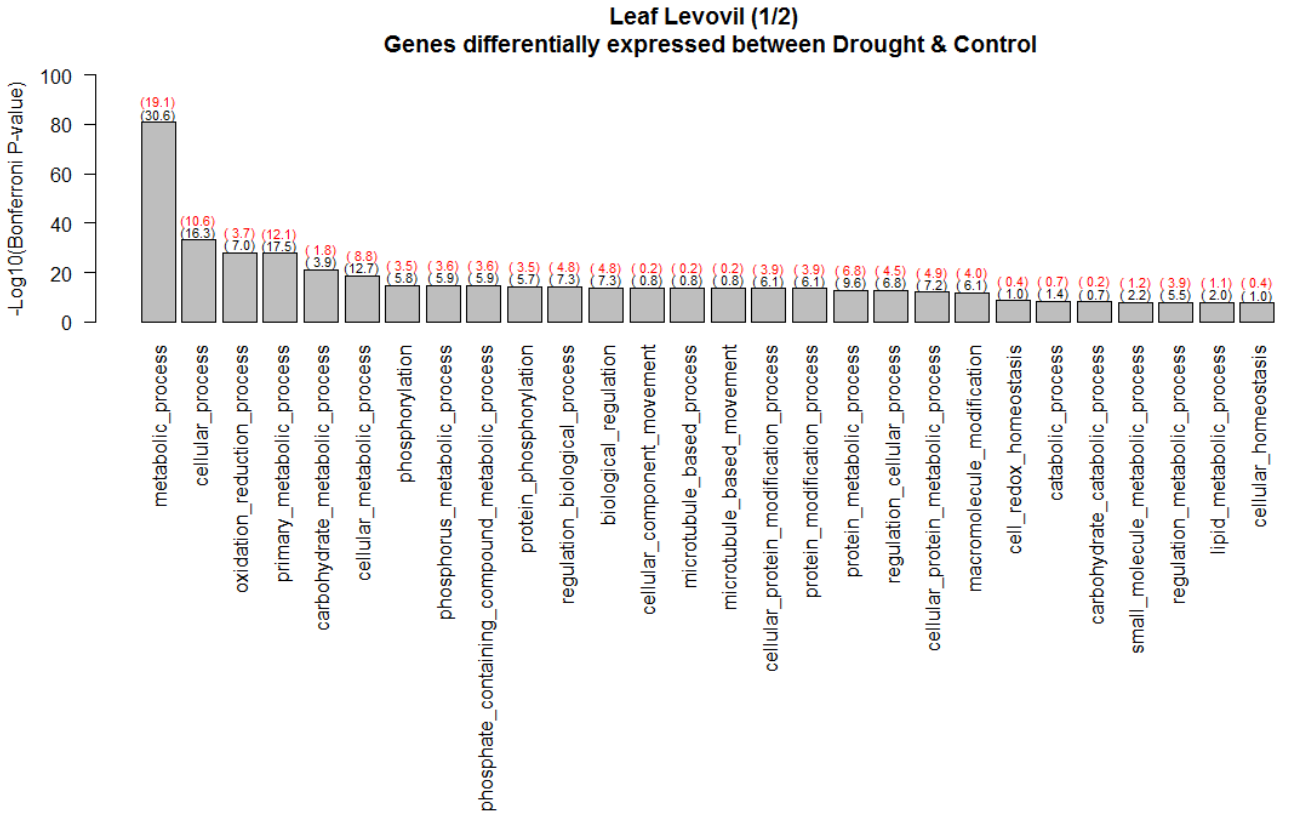
E

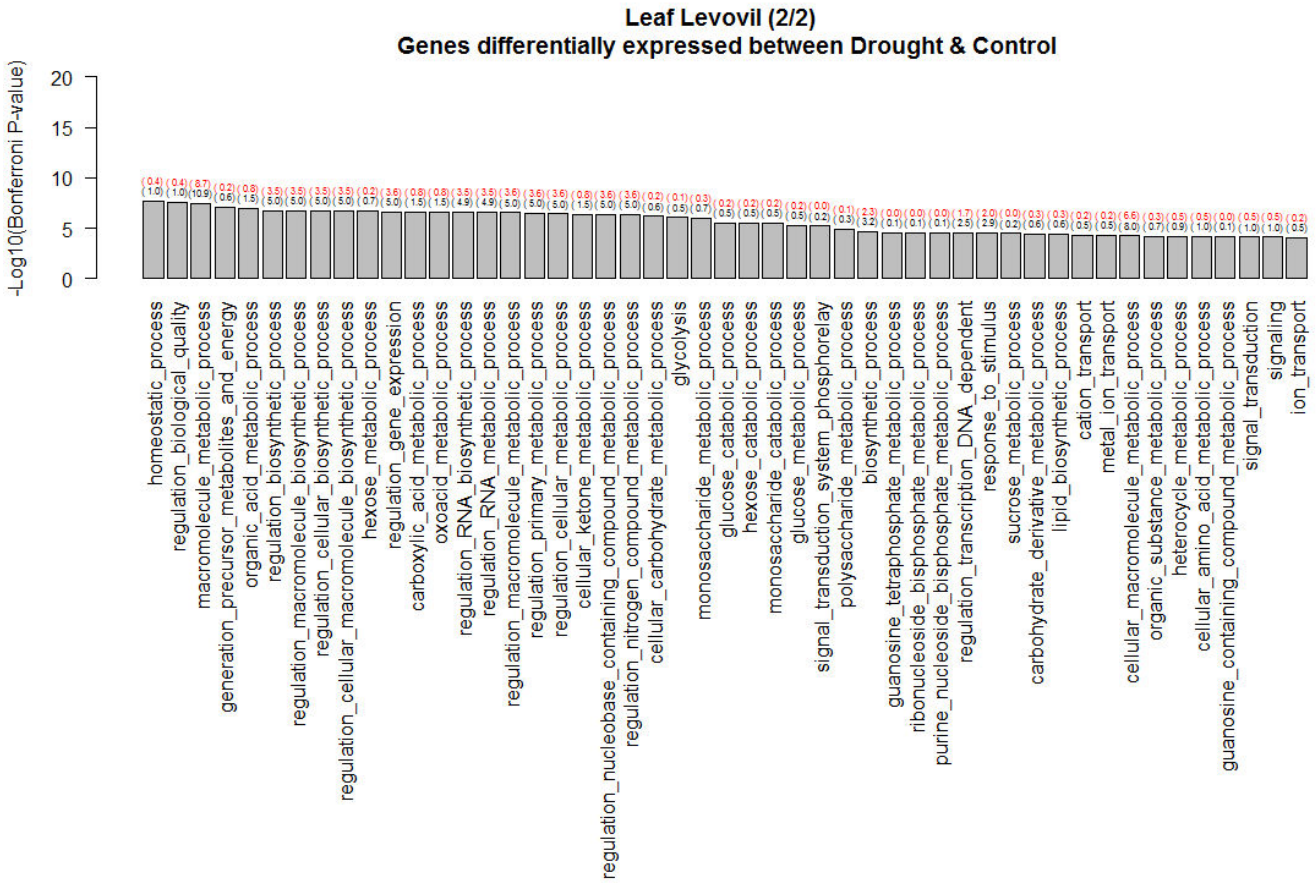

Leaf Cervil (1/2)
Genes differentially expressed between Drought & Control
-Log10(Bonferroni P-value)
0
20
40
60
80
metabolic_process
cellular_process
primary_metabolic_process
cellular_metabolic_process
cellular_protein_metabolic_process
protein_metabolic_process
phosphorus_metabolic_process
phosphate_containing_compound_metabolic_process
phosphorylation
protein_phosphorylation
cellular_protein_modification_process
protein_modification_process
macromolecule_modification
macromolecule_metabolic_process
cellular_macromolecule_metabolic_process
regulation_of_biological_process
biological_regulation
carbohydrate_metabolic_process
regulation_cellular_process
cellular_component_movement
microtubule_based_process
microtubule_based_movement
regulation_metabolic_process
biosynthetic_process
regulation_nucleobase_containing_compound_metabolic_process
regulation_nitrogen_compound_metabolic_process
regulation_RNA_biosynthetic_process
regulation_RNA_metabolic_process
regulation_gene_expression
regulation_macromolecul_metabolic_process
regulation_biosynthetic_process


Leaf Cervil (2/2)
Genes differentially expressed between Drought & Control
-Log10(Bonferroni P-value)
0
5
10
15
20
regulation_macromolecule_biosynthetic_process
regulation_cellular_biosynthetic_process
regulation_cellular_macromolecule_biosynthetic_process
regulation_primary_metabolic_process
regulation_cellular_metabolic_process
small_molecule_metabolic_process
protein_folding
regulation_transcription_DNA_dependent
oxidation_reduction_process
carbohydrate_derivative_metabolic_process
cellular_biosynthetic_process
catabolic_process
generation_precursor_metabolites_and_energy
cellular_carbohydrate_metabolic_process
lipid_metabolic_process
heterocycle_metabolic_process
small_molecule_biosynthetic_process
cellular_ketone_metabolic_process
organic_substance_metabolic_process
carboxylic_acid_metabolic_process
oxoacid_metabolic_process
organic_acid_metabolic_process
cell_redox_homeostasis
sucrose_metabolic_process
cellular_homeostasis
homeostatic_process
regulation_biological_quality
response_to_stimulus
carbohydrate_biosynthetic_process
cell_wall_macromolecule_catabolic_process
cell_wall_macromolecule_metabolic_process

Appendix 6

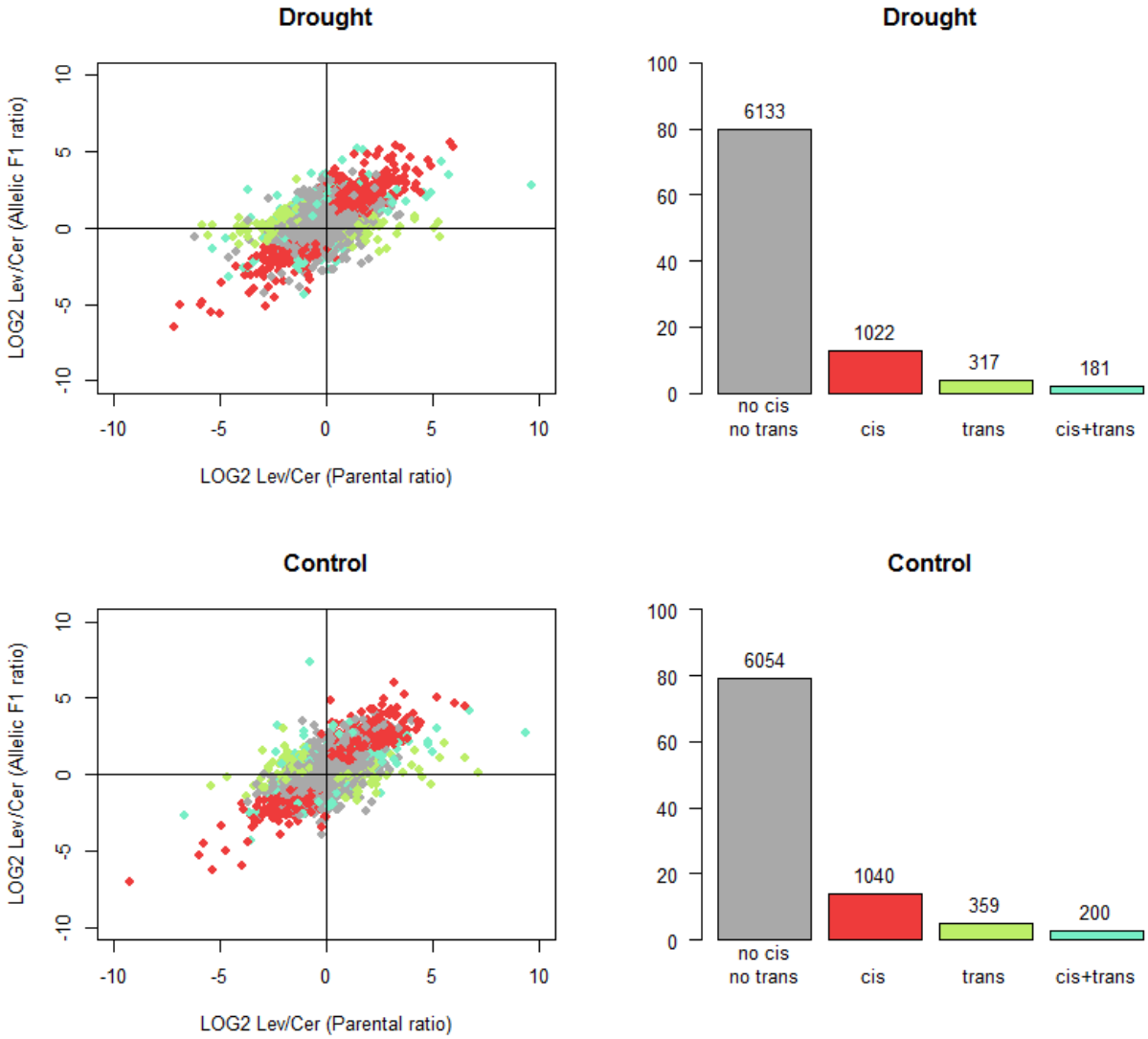
F


Leaf Levovil (1/2)
Genes differentially expressed between Drought & Control
-Log10(Bonferroni P-value)
metabolic_process
cellular_process
oxidation_reduction_process
primary_metabolic_process
carbohydrate_metabolic_process
cellular_metabolic_process
phosphorylation
phosphorus_metabolic_process
phosphate_containing_compound_metabolic_process
protein_phosphorylation
regulation_biological_process
biological_regulation
cellular_component_movement
microtubule_based_movement
cellular_protein_modification_process
protein_modification_process
protein_metabolic_process
regulation_cellular_process
cellular_protein_metabolic_process
macromolecule_modification
cell_redox_homeostasis
catabolic_process
carbohydrate_catabolic_process
small_molecule_metabolic_process
regulation_metabolic_process
lipid_metabolic_process
cellular_homeostasis


Leaf Levovil (2/2)
Genes differentially expressed between Drought & Control
-Log10(Bonferroni P-value)
homeostatic_process
regulation_biological_quality
macromolecule_metabolic_process
generation_precursor_metabolites_and_energy
organic_acid_metabolic_process
regulation_biosynthetic_process
regulation_macromolecule_biosynthetic_process
regulation_cellular_biosynthetic_process
regulation_cellular_macromolecule_biosynthetic_process
hexose_metabolic_process
regulation_gene_expression
carboxylic_acid_metabolic_process
oxoacid_metabolic_process
regulation_RNA_biosynthetic_process
regulation_RNA_metabolic_process
regulation_macromolecule_metabolic_process
regulation_primary_metabolic_process
regulation_cellular_metabolic_process
cellular_ketone_metabolic_process
regulation_nucleobase_containing_compound_metabolic_process
regulation_nitrogen_compound_metabolic_process
cellular_carbohydrate_metabolic_process
glycolysis
monosaccharide_metabolic_process
glucose_catabolic_process
hexose_catabolic_process
monosaccharide_catabolic_process
glucose_metabolic_process
signal_transduction_system_phosphorelay
polysaccharide_metabolic_process
biosynthetic_process
guanosine_tetraphosphate_metabolic_process
ribonucleoside_bisphosphate_metabolic_process
purine_nucleoside_bisphosphate_metabolic_process
regulation_transcription_DNA_dependent
response_to_stimulus
sucrose_metabolic_process
carbohydrate_derivative_metabolic_process
lipid_biosynthetic_process
cation_transport
metal_ion_transport
cellular_macromolecule_metabolic_process
organic_substance_metabolic_process
heterocycle_metabolic_process
cellular_amino_acid_metabolic_process
guanosine_containing_compound_metabolic_process
signal_transduction
signaling
ion_transport

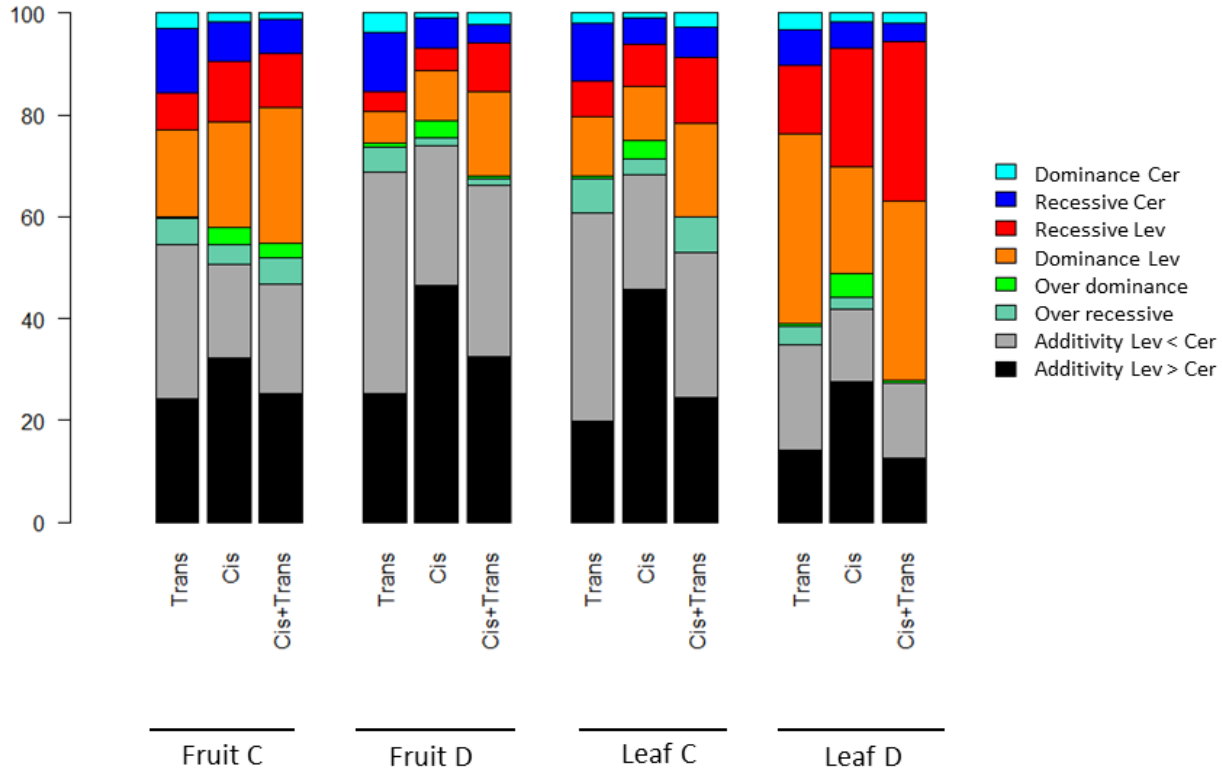
Supplemental Figure 11. Enrichment analysis of Gene Ontology (GO) related to biological process for genes differentially expressed between watering regimes (considering FDR *P-values* < 0.01). (A to C) in the fruits of the F1 hybrid (4,080 genes), Cervil (126 genes) and Levovil (1,519 genes). (D to E) in the leaves of the F1 hybrid (7,028 genes), Cervil (4,646 genes) and Levovil (4,215 genes). Only GO terms with a Bonferroni adjusted *P-value* below 0.01 are shown. At the top of each bar, frequencies of the GO terms in the corresponding differentially expressed gene cluster (black) and in the tomato genome (red) are indicated as percentage.

Appendix 6



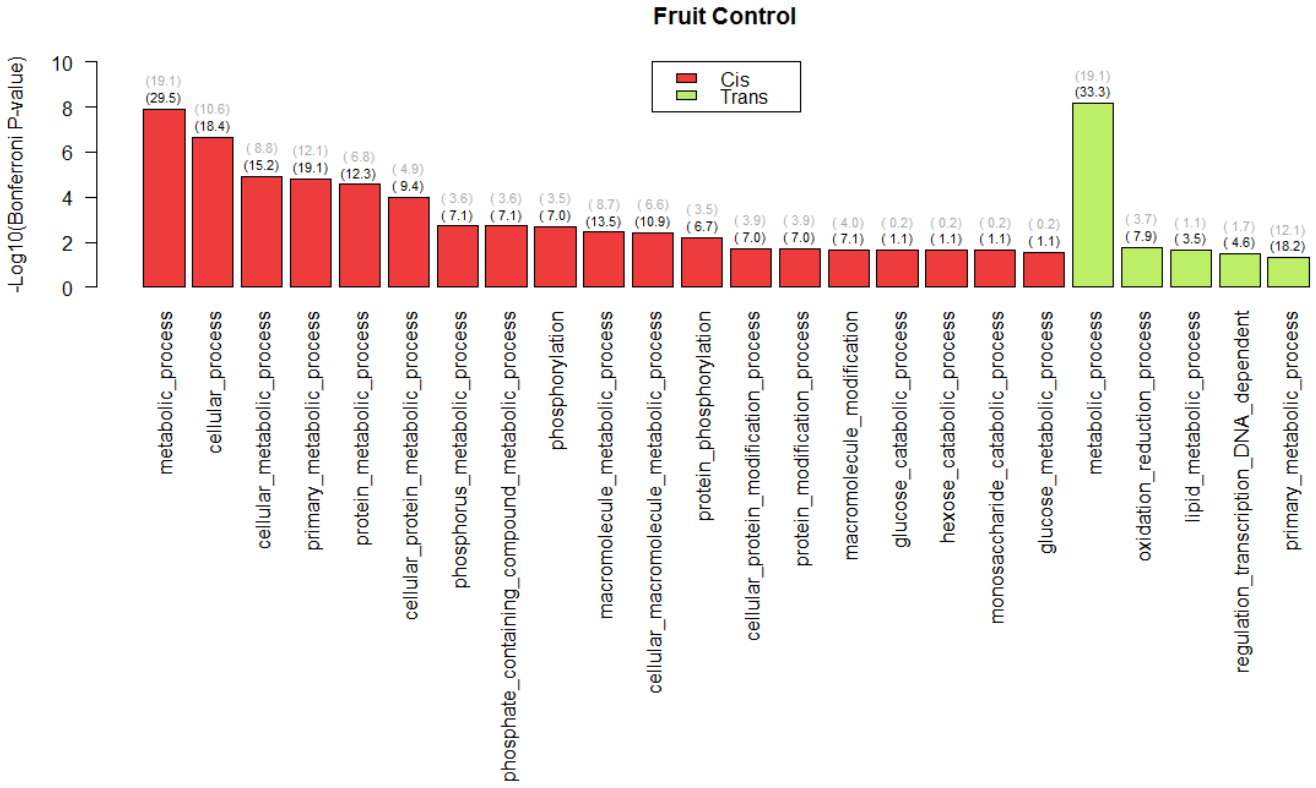
Supplemental Figure 12. *Cis* and *trans* regulatory divergence between parental genotypes for 7,653 genes the expression of which was measured in leaves through RNA sequencing, under control and drought conditions. The left bar plots display numbers and frequencies of genes in each regulatory category, in control and drought conditions. The right plots summarize the relative allelic-specific expression levels in parents and F1 hybrid (parental ratio on x axis and F1 ratio on y axis), in control and drought conditions.

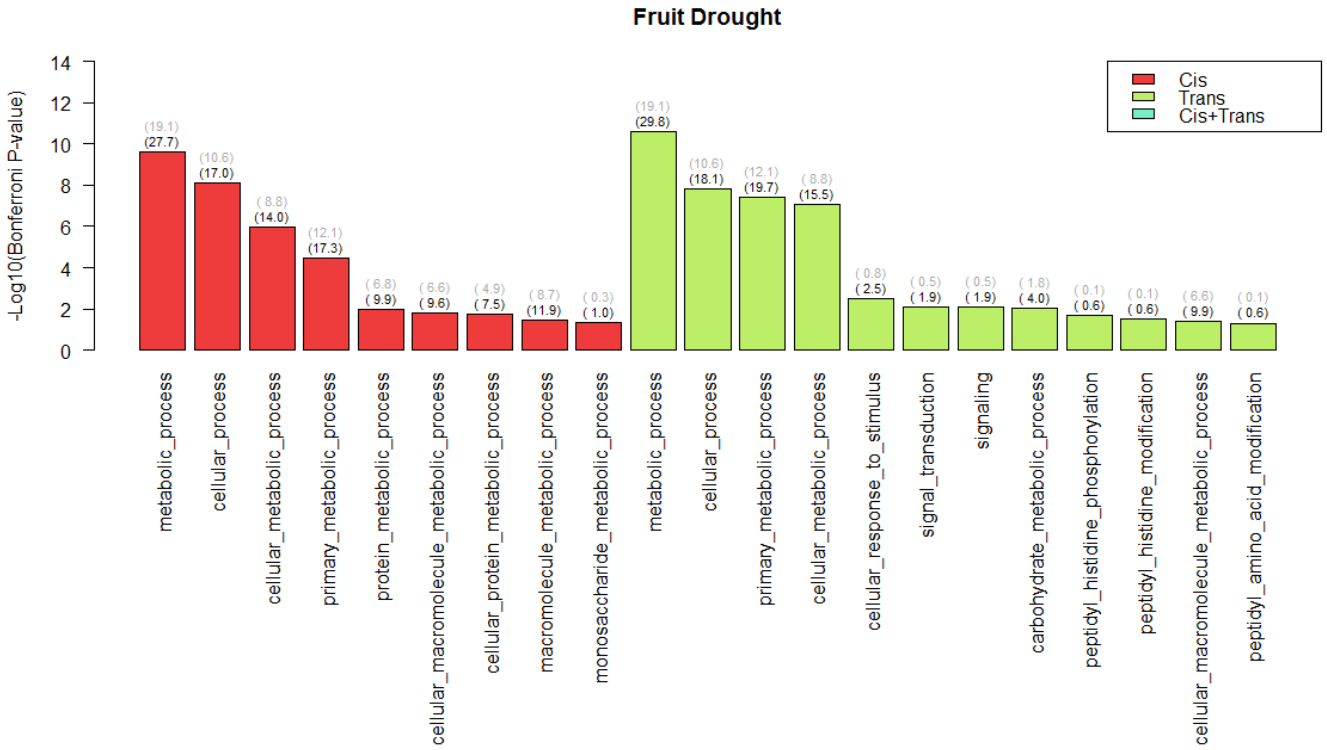
Appendix 6



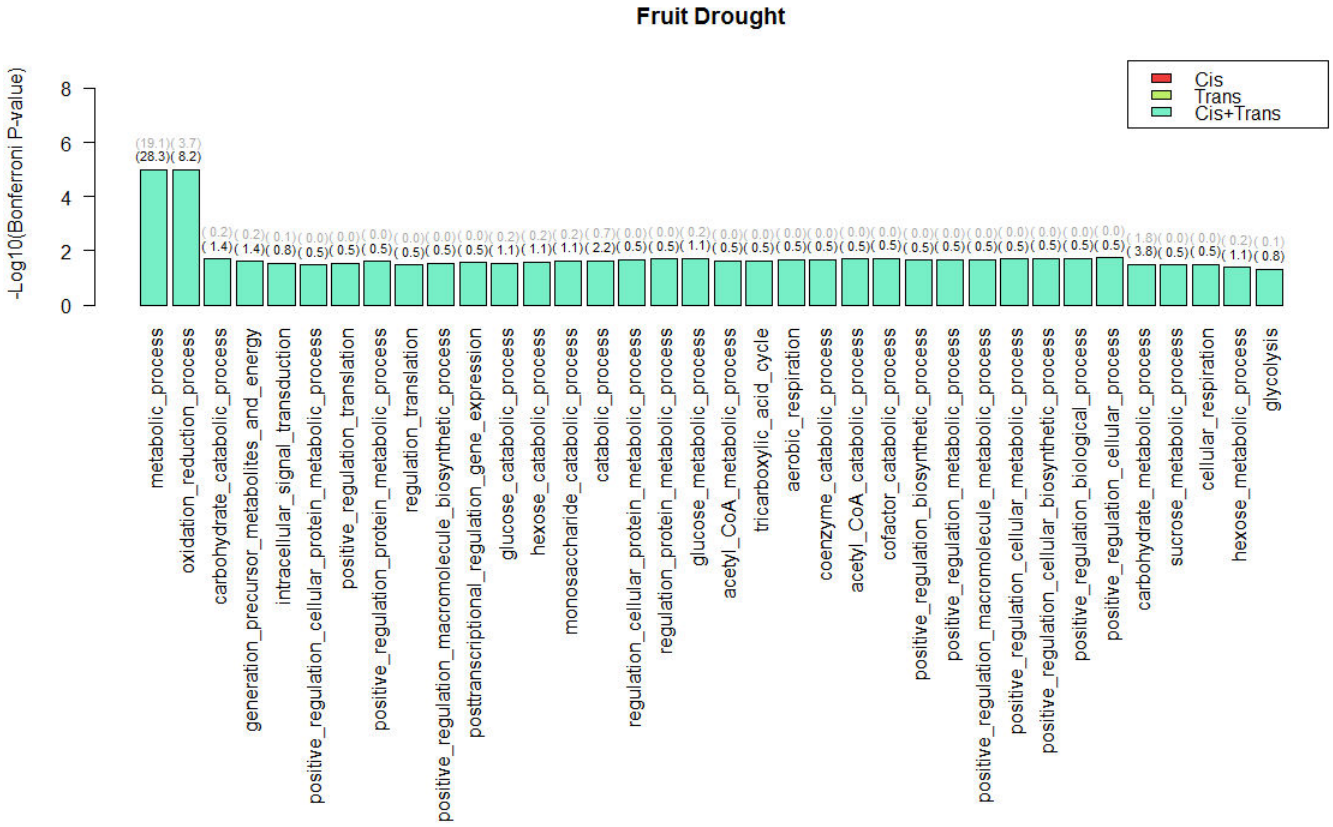
Supplemental Figures 13. Inheritance patterns of gene expression according to regulation categories, in fruit and leaf tissues, under control (C) and drought (D) conditions. 'Trans': *Trans* regulatory divergence. 'Cis': *Cis* regulatory divergence. 'Cis+Trans': Combination of both types of regulatory divergence. Inheritance patterns and regulation categories were defined from the analysis of the RNA sequencing data in Cervil, Levovil and their F1 hybrid. *Fruit Control, Trans: 308 genes, Cis: 457 genes, Cis+Trans: 75 genes. Fruit Stress, Trans: 671 genes, Cis: 674 genes, Cis+Trans: 234 genes. Leaf Control, Trans: 319 genes, Cis: 825 genes, Cis+Trans: 147 genes. Leaf Stress, Trans: 275 genes, Cis: 797 genes, Cis+Trans: 143 genes.*

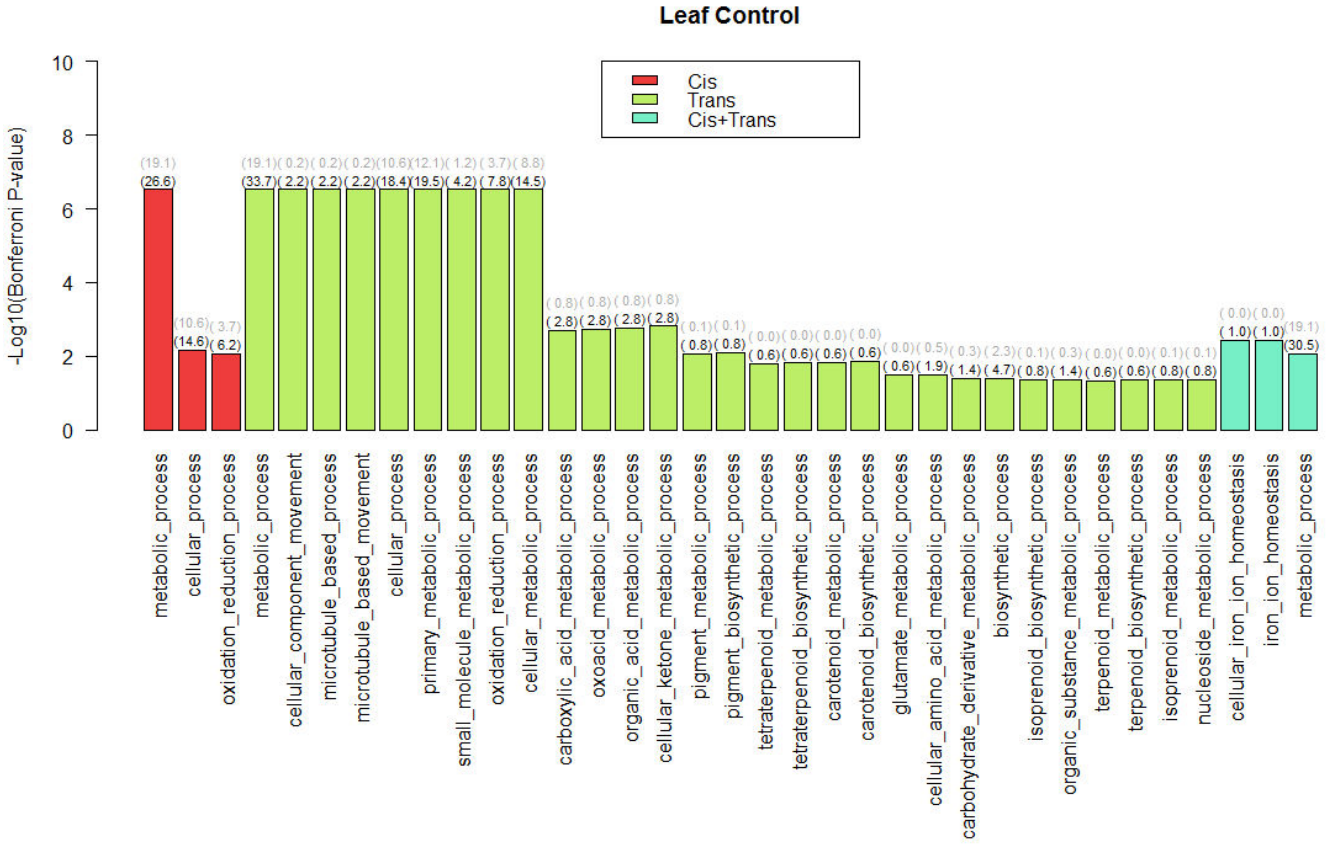
Appendix 6


Fruit Control
Cis
Trans
-Log10(Bonferroni P-value)
0
2
4
6
8
10
(19.1) (29.5)
(10.6) (18.4)
(8.8) (15.2)
(12.1) (19.1)
(6.8) (12.3)
(4.9) (9.4)
(3.6) (7.1)
(3.6) (7.1)
(3.5) (7.0)
(8.7) (13.5)
(6.6) (10.9)
(3.5) (6.7)
(3.9) (7.0)
(3.9) (7.0)
(4.0) (7.1)
(0.2) (1.1)
(0.2) (1.1)
(0.2) (1.1)
(0.2) (1.1)
(19.1) (33.3)
(3.7) (7.9)
(1.1) (3.5)
(1.7) (4.6)
(12.1) (18.2)
metabolic_process
cellular_process
cellular_metabolic_process
primary_metabolic_process
protein_metabolic_process
cellular_protein_metabolic_process
phosphorus_metabolic_process
phosphate_containing_compound_metabolic_process
phosphorylation
macromolecule_metabolic_process
cellular_macromolecule_metabolic_process
protein_phosphorylation
cellular_protein_modification_process
protein_modification_process
macromolecule_modification
glucose_catabolic_process
hexose_catabolic_process
monosaccharide_catabolic_process
glucose_metabolic_process
metabolic_process
oxidation_reduction_process
lipid_metabolic_process
regulation_transcription_DNA_dependent
primary_metabolic_process

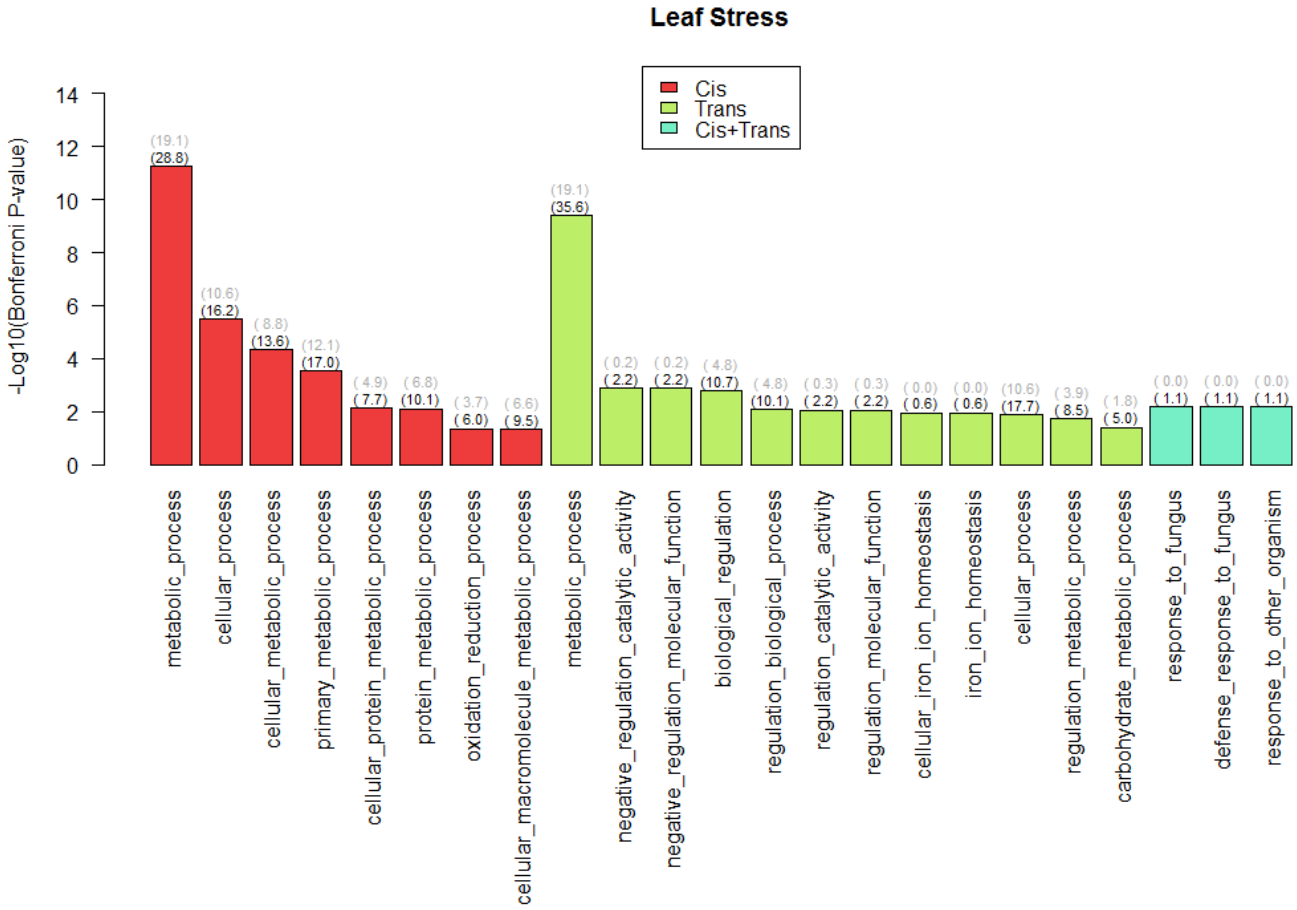

Fruit Drought
Cis
Trans
Cis+Trans
-Log10(Bonferroni P-value)
0
2
4
6
8
10
12
14
(19.1) (27.7)
(10.6) (17.0)
(8.8) (14.0)
(12.1) (17.3)
(6.8) (9.9)
(6.6) (9.6)
(4.9) (7.5)
(8.7) (11.9)
(0.3) (1.0)
(19.1) (29.8)
(10.6) (18.1)
(12.1) (19.7)
(8.8) (15.5)
(0.8) (2.5)
(0.5) (1.9)
(0.5) (1.9)
(1.8) (4.0)
(0.1) (0.6)
(0.1) (0.6)
(6.6) (9.9)
(0.1) (0.6)
metabolic_process
cellular_process
cellular_metabolic_process
primary_metabolic_process
protein_metabolic_process
cellular_macromolecule_metabolic_process
cellular_protein_metabolic_process
macromolecule_metabolic_process
monosaccharide_metabolic_process
metabolic_process
cellular_process
primary_metabolic_process
cellular_metabolic_process
cellular_response_to_stimulus
signal_transduction
signaling
carbohydrate_metabolic_process
peptidyl_histidine_phosphorylation
peptidyl_histidine_modification
cellular_macromolecule_metabolic_process
peptidyl_amino_acid_modification

Appendix 6

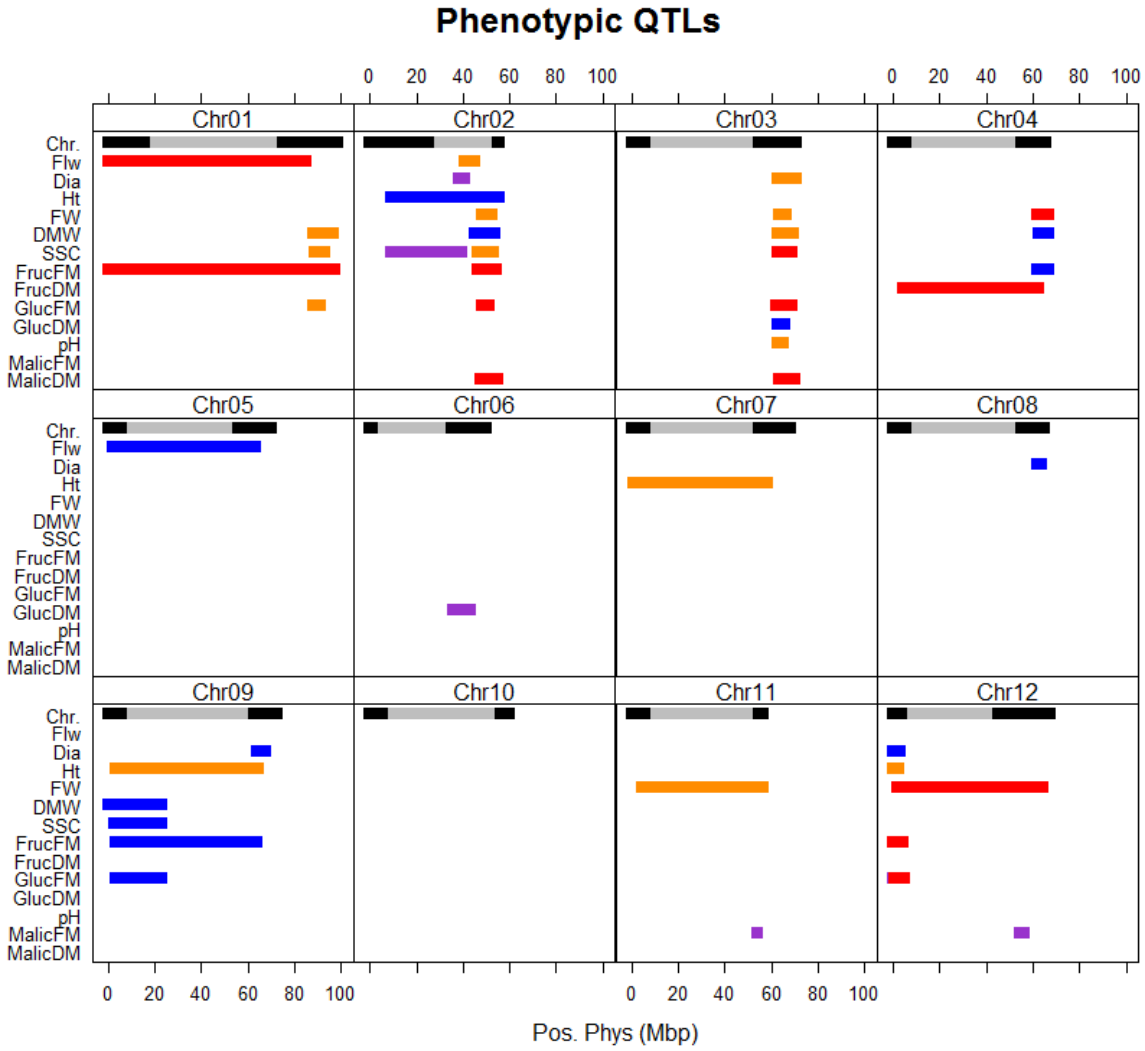

Fruit Drought
-Log10(Bonferroni P-value)
Cis
Trans
Cis+Trans
metabolic_process
oxidation_reduction_process
carbohydrate_catabolic_process
generation_precursor_metabolites_and_energy
intracellular_signal_transduction
positive_regulation_cellular_protein_metabolic_process
positive_regulation_translation
positive_regulation_protein_metabolic_process
regulation_translation
positive_regulation_macromolecule_biosynthetic_process
posttranscriptional_regulation_gene_expression
glucose_catabolic_process
hexose_catabolic_process
monosaccharide_catabolic_process
catabolic_process
regulation_cellular_protein_metabolic_process
regulation_protein_metabolic_process
glucose_metabolic_process
acetyl_CoA_metabolic_process
tricarboxylic_acid_cycle
aerobic_respiration
coenzyme_catabolic_process
acetyl_CoA_catabolic_process
cofactor_catabolic_process
positive_regulation_biosynthetic_process
positive_regulation_metabolic_process
positive_regulation_macromolecule_metabolic_process
positive_regulation_cellular_metabolic_process
positive_regulation_cellular_biosynthetic_process
positive_regulation_biological_process
positive_regulation_cellular_process
carbohydrate_metabolic_process
sucrose_metabolic_process
cellular_respiration
hexose_metabolic_process
glycolysis


Leaf Control
-Log10(Bonferroni P-value)
Cis
Trans
Cis+Trans
metabolic_process
cellular_process
oxidation_reduction_process
metabolic_process
cellular_component_movement
microtubule_based_process
microtubule_based_movement
cellular_process
primary_metabolic_process
small_molecule_metabolic_process
oxidation_reduction_process
cellular_metabolic_process
carboxylic_acid_metabolic_process
oxoacid_metabolic_process
organic_acid_metabolic_process
cellular_ketone_metabolic_process
pigment_metabolic_process
pigment_biosynthetic_process
tetraterpenoid_metabolic_process
tetraterpenoid_biosynthetic_process
carotenoid_metabolic_process
carotenoid_biosynthetic_process
glutamate_metabolic_process
cellular_amino_acid_metabolic_process
carbohydrate_derivative_metabolic_process
biosynthetic_process
isoprenoid_biosynthetic_process
organic_substance_metabolic_process
terpenoid_metabolic_process
terpenoid_biosynthetic_process
isoprenoid_metabolic_process
nucleoside_metabolic_process
cellular_iron_ion_homeostasis
iron_ion_homeostasis
metabolic_process

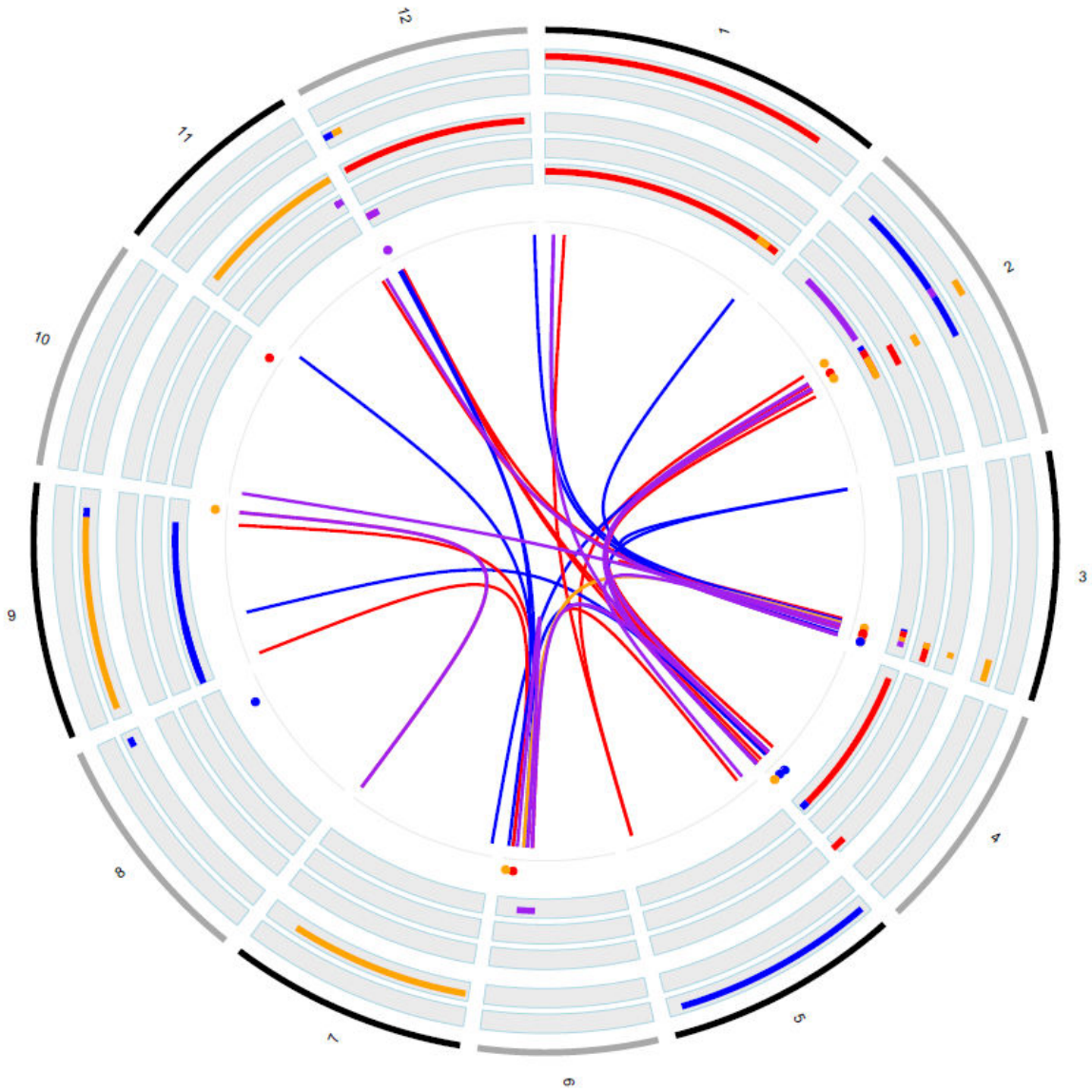
Appendix 6



Supplemental Figures 14. Enrichment analysis of Gene Ontology (GO) related to biological process for genes presenting *Cis*, *Trans* and *Cis+Trans* regulatory divergence in fruit and leaf RNA sequencing data, under both watering regimes. Only GO terms with a Bonferroni adjusted *P-value* below 0.05 are shown. At the top of each bar, frequencies of the GO terms in the corresponding differentially expressed gene cluster (black) and in the tomato genome (grey) are indicated as percentage.



Supplemental Figure 15. Overview of phenotypic QTL identified on the tomato genome by linkage analysis in the RILs. At the top of the panels, lines are representing tomato chromosomes where the lengths are proportional to chromosome physical sizes in million base pairs (Mbp). Centromeric regions with low recombination frequency are indicated in grey and peripheral parts in black (according to Sim et al., 2012). QTL are represented by square. Color codes correspond to the QTL types: constitutive (common to control and drought treatment) in orange; detected only in control treatment in blue; detected only in drought treatment in red; interactive between the two watering regimes in purple.



Supplemental Figure 16: Overview of the QTLs and leaf eQTLs identified on the twelve chromosomes of the tomato genome by linkage analysis in the RILs. The diagram displays the twelve chromosomes of the tomato genome proportionally to their physical size (assembly 2.5). In five first layers, lines represent QTLs detected for phenotypic traits. From outside to inwards: (1) flw, (2) vigor traits (Ht and Dia), (3) FW, (4) acid traits (pH, MalicFM, MalicDM, CirticFM and CitricDM) and (5) sugar traits (DMW, SSC, GlucoseFM, GlucoseDM, FructoseFM and FructoseDM). In the inner part, links represent *trans acting eQTLs* and dots *local eQTLs*, detected on expression data measured in leaves. Colors indicate QTL and eQTL types. Orange: *constitutive*. Blue: *control specific*. Red: *drought specific*. Purple: *interactive*.

Supplemental Table 7. Pearson correlations between phenotypic traits measured in the RIL in 2013 (Albert *et al.* 2016) and 2015.

Traits	Control	Drought
Plant traits		
Flw	0.69 ***	0.73 ***
Ht	0.75 ***	0.79 ***
Dia	0.21 *	0.53 ***
Fruit traits		
FW[a]	0.85 ***	0.82 ***
DMW	0.61 ***	0.23 *
SSC	0.65 ***	0.27 **
pH	0.37 ***	0.36 ***

[a] Transformed to ensure a normal distribution (LOG10).

*** shows P-value below 0.001, ** between 0.001 and 0.01 and * between, 0.01 and 0.05.

Supplemental Table 8. Pearson correlations between phenotypic traits measured in control and drought in the RIL (data 2015).

Traits	
Plant traits	
Flw	0.84 ***
Ht	0.86 ***
Dia	0.49 ***
Fruit traits	
FW[a]	0.91 ***
DMW	0.78 ***
SSC	0.83 ***
pH	0.60 ***
GlucoseFM	0.65 ***
FructoseFM	0.50 ***
MalicFM	0.15 NS

[a] Transformed to ensure a normal distribution (LOG10).

*** shows P-value below 0.001, NS non-significant.