Allele-specific expression and genetic determinants of transcriptomic variations in response to mild water deficit in tomato

Elise Albert, Renaud Duboscq, Muriel Latreille, Sylvain Santoni, Matthieu Beukers, Jean-Paul Bouchet, Frederique Bitton, Justine Gricourt, Charles Poncet, Véronique Gautier, et al.

HAL Id: hal-02620940
https://hal.inrae.fr/hal-02620940

Submitted on 26 May 2020

Allele specific expression and genetic determinants of transcriptomic variations in response to mild water deficit in tomato

Authors

Elise Albert[a,f], Renaud Duboscq[a], Muriel Latreille[b], Sylvain Santoni[b], Matthieu Beukers[a], Jean-Paul Bouchet[a], Fréderique Bitton[a], Justine Gricourt[a], Charles Poncet[c], Véronique Gautier[c], José M Jiménez-Gómez[d], Guillem Rigaill[e], 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

[e] INRA, UMR8071, Laboratoire de Mathématiques et Modélisation d'Evry, Université d'Evry Val d'Essonne, ENSIIE-INRA-CNRS, Évry, 91037, France

[f] Present address : Biochemistry & Molecular Biology Department Michigan State University 603 Wilson Road East Lansing, MI 48823 USA

§ Corresponding author

Mathilde Causse; Mathilde.Causse@inra.fr

Running title: Transcriptome variation in water stressed tomato

Key Words: Transcriptome; Allele specific expression; eQTL; Genotype by environment interaction; Drought; *Solanum lycopersicum*

Summary

Characterizing the natural diversity of gene expression across environments is an important step in understanding how genotype by environment interactions shape phenotypes. Here, we analyzed the impact of water deficit onto gene expression levels in tomato at the genome-wide scale. We sequenced the transcriptome of growing leaves and fruit pericarps at cell expansion stage in a cherry and a large fruited accession and their F_1 hybrid grown under two watering regimes. Gene expression levels were steadily affected by the genotype and the watering regime. Whereas phenotypes showed mostly additive inheritance, ~80% of the genes displayed non-additive inheritance. By comparing allele specific expression in the F_1 hybrid to the allelic expression in both parental lines, respectively, 3,005 genes in leaf and 2,857 genes in fruit deviated from 1:1 ratio independently of the watering regime. Among these genes, ~55% were controlled by *cis factors*, ~25% by *trans* factors and ~20% by a combination of both types. A total of 328 genes in leaf and 113 in fruit exhibited significant allele specific expression by watering regime interaction, among which ~80% presented *trans* by watering regime interaction suggesting a response to water deficit mediated through a majority of *trans* acting loci in tomato. We cross-validated the expression levels of 274 transcripts in fruit and leaves of 124 RILs and identified 163 eQTLs mostly confirming the divergences identified by allele specific expression. Combining phenotypic and expression data, we observed a complex network of variation between genes encoding enzymes involved in the sugar metabolism.

Introduction

Variation in gene expression levels is central to evolution and a key process in plant adaptation (Wray *et al.*, 2003; Fraser, 2011; Wittkopp and Kalay, 2012). Following the rapid development of gene expression arrays and sequencing technologies, genome-wide assays of gene expression were rapidly adopted to explore differential gene expression along plant

development or in response to various biotic and abiotic stresses, including water deficit (Seki *et al.*, 2002; Rabbani *et al.*, 2003; Des Marais *et al.*, 2012). However, exploration of the natural genetic variations in transcriptomic responses to environmental constraints and their genetic determinants remains limited (Druka *et al.*, 2010; Emerson and Li, 2010).

Two complementary strategies have been established to uncover the genetic architecture of gene expression variation. The first one, expression quantitative trait locus (*eQTL*) mapping, takes advantages of methodologies developed for mapping quantitative trait loci (*QTL*), either in the linkage mapping or statistical association framework, considering transcript expression 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 the variation in gene expression levels and distinguishing between *distant* and *local* eQTLs according to their distance to the regulated gene. *Distant* eQTLs are steadily inferable to *trans*-acting regulatory elements, *i.e.* genes that modify or regulate the expression of distant genes. However, eQTL mapping often lacks resolution to determine whether a *local* eQTL corresponds to a polymorphism in a *trans*-acting regulatory element adjacent to the gene it regulates or in 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 comparison of allele specific expression (ASE) in an F_1 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 the hybrid 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 permit to map *trans*-acting elements on the genome but can efficiently determine whether gene expression levels variation result from *cis*-acting elements or *trans*-acting elements.

ASE and eQTL mapping studies were first 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; Morley *et al.*, 2004; Monks *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; Cubillos *et al.*, 2012*b*, 2014; He *et al.*, 2016; Zan *et al.*, 2016), maize (Stupar and Springer, 2006; Springer and Stupar, 2007*a*; Wang *et al.*, 2018) and tree species (Kirst, 2004 in eucalyptus; Drost *et al.*, 2010 in poplar; Combes et al., 2015 in coffee; Verta et al., 2016 in white spruce). These studies successfully identified loci and polymorphisms involved in gene expression levels regulation. However, genotype by environment interactions at the gene expression level remain poorly understood. To the best of our knowledge, eQTL and/or ASE by environment interactions have only been addressed in a few plant studies. Cubillos *et al.* (2014), Lovell *et al.* (2016) and Waters *et al.* (2016) recorded genotype-dependent transcriptomic responses to water, cold and heat stresses mediated through a large majority of *cis* acting by environment interactions in *Arabidopsis thaliana*, *Panicum halli* and *Zea mays*, respectively. On the other hand, Lowry *et al.* (2013), Snoek *et al.* (2013) and Clauw *et al.* (2016) reported a preponderance of *trans* acting eQTL by environment interactions in *Arabidopsis thaliana* subjected to drought and shade conditions.

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. Several transcriptomic analyses revealed the evolution of gene expression levels along tomato fruit development (Tomato Genome Consortium, 2012) or in different fruit tissues (Carrari *et al.*, 2006; Mounet *et al.*, 2009; Matas *et al.*, 2011; Pattison *et al.*, 2015; Zhang *et al.*, 2016). The genetic basis of these variations have been addressed for the first time by Ranjan *et al.* (2016), revealing 7,225 loci regulating the steady-state transcript levels of 5,289 genes using an introgression line population derived from the wild relative species *Solanum pennellii.* More recently, Zhu *et al.* (2018) performed transcriptome based association mapping in a collection of 399 diverse tomato genotypes and identified 2,566 *cis* eQTLs and 93,587 *trans* eQTLs involved in the expression level variations of 3,465 genes. QTL by watering regime interactions were recently reported for tomato fruit quality traits emphasizing their importance for its improvement (Gur *et al.*, 2011; Albert *et al.*, 2016a, b). However, genetic variation in gene expression levels under water deficit has only been considered so far on a limited number of contrasted tomato accessions (Albert *et al.*, 2016*a*; Liu *et al.*, 2017) and the underlying genetic architecture remains to be explored.

In the present study, we sequenced the transcriptomes of leaves and fruit pericarps of a cherry type tomato (Cervil), a large fruited accession (Levovil) and their F_1 hybrid, grown under two contrasted watering regimes. First, we identified differentially expressed genes between genotypes and watering regimes. Then, we characterized the extent of ASE and its interaction with the watering regime in the F_1 hybrid, at the whole transcriptome scale. Finally, we 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. We assessed their expression levels in leaves and fruit pericarps sampled in a recombinant inbred line (RILs) population derived from the same cross. By performing eQTL and eQTL by watering regime interaction mapping, we cross validated the regulatory divergences characterized using ASE and notably report a complex network of variation between genes encoding enzymes involved in the sugar metabolism.

Results

Phenotypic variation and inheritance of plant and fruit traits

Eleven phenotypic traits, three related to plant vigor and phenology and eight related to mature fruit quality, were assessed on the two parental lines and their F_1 hybrid, in control and drought conditions. The three genotypes significantly differed for all the phenotypic traits except for malic acid content **(Table S9)**. The watering regime and genotype by watering regime interaction factors were significant for four (Ht, Dia, FW, Glucose) and three (Ht, Dia and SSC) of the eleven traits, respectively. When comparing the phenotypic means, we observed mostly additive inheritance under both watering regimes, except for Dia (C) and VitC (C & D), which exhibited a recessive inheritance pattern and for Ht (C & D), Fructose (C & D) and Glucose (C) which exhibited a dominant pattern **(Table S9 and Figure S1).**

In the 124 RILs derived from the cross between Cervil and Levovil, phenotypic variations were observed under both watering regimes for every trait (coefficients of variation ranking from 2.7 to 48.9%). Every trait except FW showed transgressions beyond parental values in both directions, under both watering regimes **(Figure S2).** The genotype by watering regime interaction was significant for seven of the eleven traits (except Flw, SSC, pH and VitC), 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%, all *P-values* < 0.001) **(Table S9).** Broad-sense heritabilities ranged from 0.14 (Malic under drought) to 0.93 (FW under control condition) and were conserved between watering regime (r^2 = 0.93). The means of the phenotypic traits in the RILs were significantly correlated between watering regimes **(Table S10)** and with data previously described (Albert *et al.*, 2016*a*) where the RILs were grown under a more severe drought treatment **(Table S11).**

Gene expression variations between genotypes and inheritance of gene expression

RNA sequencing libraries of Cervil, Levovil and their F1 hybrid yielded a total 18,843 and 19,097 genes with expression levels above the background noise, representing 54% and 55% of the total tomato gene models, in leaf and fruit, respectively. Principal coordinate analysis was performed using normalized RNAseq read counts from parents and hybrid samples separating leaf and fruit data. The first principal coordinates distinguished samples according to their genotypes in both tissues, explaining 56% and 71% of the total variation, respectively. The second principal coordinate separated samples according to genotypes and watering regimes, explaining 26% and 12% of the total variation in leaf and fruit, respectively **(Figure S3)**.

We first characterized differential gene expression between the F1 hybrid, Cervil and Levovil, under each watering regime, separately. Overall the proportion of differential expressed genes between genotypes was estimated to 84% (15,822 genes) and 81% (15,509 genes) in leaf and fruit, respectively. Absolute log2 fold changes ranged from 0.11 to 9.15 and adjusted *P-values* ranged from 5.00×10^{-2} to 2.23×10^{-308}. The pairwise comparison between Cervil and the F_1 hybrid fruit transcriptomes under drought condition yielded the lowest number of differentially expressed genes (2,458 genes, 13%), whereas the comparison between Levovil and the F_1 hybrid leaf transcriptomes under drought condition yielded the highest number of differentially expressed genes (11,351 genes, 60%) **(Figure S4)**.

A total of 6,665 (10,519) genes and 11,918 (6,171) genes presented an expression level in the F_1 hybrid significantly different from the expression in at least one parent in leaf (fruit) under control and drought **(Table S12).** Among these genes all the different forms of alteration in the gene expression levels were observed, i.e. additive, dominant, recessive,

over-dominant and over-recessive. Noteworthy, in both organs and under both watering regimes, less than 15% of the genes presented additive inheritance patterns. The proportion of genes showing transgressive inheritance pattern (over-dominant or over-recessive) varied greatly between watering regimes and organs. A total of 1,188 genes and 7,043 genes presented such inheritance pattern in leaf under control and drought conditions, respectively. In fruit, 4,140 genes presented transgressive inheritance under the control condition and 295 under drought condition. Inheritance patterns were conserved between watering regimes for 42% of the genes in leaf and 51% of the genes in fruit and conserved between organs for 38% of the genes in control and 30% of the genes in drought conditions.

Variations in gene expression levels between watering regimes

To assess to which extent the watering regime affected gene expression levels, we estimated the number of differentially expressed genes between drought and control conditions in each genotype. A total of 12,679 genes in leaves and 8,397 genes in fruit were differentially expressed across comparisons, representing 67% and 44% of the analyzed genes, respectively **(Figure 1)**. A total 9,999 genes were assessed for differential expression in both organs, among which 6,293 (63%) were differentially expressed between drought and control conditions in at least one genotype in both leaf and fruit. Absolute log2 fold changes ranged from 0.15 to 6.98 and adjusted *P-values* ranged from 5.00×10^{-2} to 5.53×10^{-303}. While almost as many genes were differentially expressed between watering regimes in leaf of Cervil (6,601 genes) and Levovil (5,853 genes), eight fold less genes were differentially expressed between watering regimes in Cervil (389 genes) compared to Levovil (3,073 genes) in fruit. The F_1 hybrid displayed two to 15 fold more differentially expressed genes in response to water deficit in comparison to parental genotypes in both organs.

Enrichment tests revealed several significantly over-represented gene ontology terms (GO) associated to biological process **(Table S13)**. Among the 12,679 genes differentially expressed between watering regimes in leaf, GO related to "cellular process", "translation and signal transduction", "response to abiotic stimulus, "response to stress", "generation of precursor metabolites and energy" and "cellular homeostasis" were significantly over-represented. Among the 8,397 genes differentially expressed between watering regimes in fruit, GO related to "translation", "response to abiotic stimulus", "response to stress" and "photosynthesis" were significantly over-represented.

Allele specific expression and constitutive regulatory divergences

Exonic SNPs between Cervil and Levovil permitted quantification of allele specific expression (ASE) in the F1 hybrid for 8,434 genes (24% of the known tomato CDS) in the leaf and 7,459 genes (21% of the know tomato CDS) in the fruit. At least two SNPs were required to retain a gene. Initially, we compared the ASE in the F1 hybrid to the allelic expression in both parental lines independently of the watering regime in order to distinguish genes presenting *cis*, *trans* and combination of both types of regulatory divergence (*cis + trans*) maintained across watering regimes. Among the genes assessed, only 3,005 genes in leaf (36%) and 2,857 genes in fruit (38%) showed significant constitutive regulatory divergences. In both organs, a majority of *cis* regulated genes were identified (1,878 in leaf and 1,428 in fruit), and fewer *trans* regulated genes (610 in leaf and 906 in fruit). A total of 517 and 523 genes in leaf and in fruit, respectively, combined both *cis* and *trans*-regulatory divergences **(Figure 2A).** Among the genes presenting significant *cis* regulatory divergences, absolute allelic log2 fold change in the F1 hybrid ranged from 0.20 to 11.00. Among *trans* regulated genes, absolute log2 fold change between parental allelic expressions ranged from 0.26 to 6.00 **(Figure 3A and B).** A set of 6,845 genes was assessed for allele specific expression in both organs among which 1,083 genes were *cis*-regulated and 266 genes *trans*-regulated both in fruit and leaf.

Interestingly, 70% and 67% of the *cis* regulated genes exhibited greater expression level from Levovil allele than Cervil allele in the F1 hybrid, in leaf and fruit, respectively. In fruit, 56% of the *trans* regulated genes presented significantly higher expression in Levovil compared to Cervil. The opposite was observed in leaf where 66% of the *trans* regulated genes showed higher expression in Cervil compared to Levovil. In the sets of genes presenting both types of regulatory mechanisms, we categorized the direction of the *cis* and *trans* effect as "convergent" when allelic effects were in the same direction in parents and F1 or "compensatory" when allelic effects were in opposite direction. For 58 % and 63% of the genes in the fruit and in the leaf, allelic effects converged between *cis* and *trans* effects.

GO enrichment tests for terms related to biological process among the genes presenting constitutive *cis, trans* and *cis+trans* regulatory divergences in leaf showed a significant enrichment in GO terms related to "metabolic process", "cellular process",

"photosynthesis", "translation", "signal transduction" and "response to stress". In fruit terms, related to "cellular process", "response to abiotic stimulus", "translation" and "signal transduction were significantly over-represented among the *cis*, *trans* and *cis+trans* regulated genes. **(Table S14)**.

Allele specific expression by watering regime interactions

When including the genotype by watering regime interaction in the modeling procedure, we identified a total of 328 genes (4% of the genes assessed) and 113 genes (1.51 % of the genes assessed) presenting significant ASE by watering regime interaction, in leaf and fruit, respectively **(Figure 2B)**. Contrary to the pattern observed for constitutive regulatory divergences, a larger proportion of genes regulated through *trans* by watering interaction were identified (251 in leaf and 106 in fruit), and fewer *genes* regulated through *cis* by watering regime interaction (32 in leaf and 1 in fruit). A total of 77 genes in leaf and 7 in fruit presented a combination of both categories of ASE by watering regime interaction.

The most common pattern of *cis*-by-watering regime effect was found in the case where the alleles conserved their relative ranking in the F1 hybrid across watering regimes but the difference between the allelic means was shifted (differential effect, 60% of the *cis*-by-watering regime interactions). Among the genes presenting *trans*-by-watering regulatory divergences, we observed 47% of differential effects where ranking of parental allelic expression was conserved across watering regimes but difference between allelic means was shifted and 53% of antagonist effects where the direction of allelic mean difference was opposed between control and drought conditions.

Four GO terms related to "response to biotic stimulus", "response to stress", "biological_process" and "response to external stimulus" were over-represented among the 328 genes presenting significant ASE by watering regime interaction in leaf. One GO term related to "catabolic process" was over-represented in the 113 genes associated to significant ASE by watering regime interaction in fruit **(Table S14)**.

QTL mapping in the RIL progeny for plant vigor and fruit quality traits

A total of 39 QTLs were mapped for 10 of the 11 phenotypic traits assessed in the RILs (no QTL detected for VitC content in fruit) **(Tables S15 and S16, Figure 4).** The 39 QTLs explained more than 11% of the total phenotypic variation (PVE) of the traits, with a maximum of 37% for FW under drought condition. LOD scores varied from 3.15 to 12.74. The confidence intervals were smaller than 6 Mbp for 87% of the QTLs (encompassing between 31 and 715 genes), whereas five QTLs mapped across centromeric regions had larger confidence intervals (from 19 to 59 Mbp, encompassing between 518 and 1,448 genes). Twenty-eight QTLs co-localized within five clusters on chromosomes 1, 2, 3, 9 and 12. Thirteen QTLs were detected under control condition only and 9 QTLs under drought condition only, while 11 QTLs were constitutive (detected under both watering regimes with consistent effect, **Table S15)**. Six additional QTLs were detected using the plasticity traits **(Table S16)**. Twenty QTLs colocalized with previously identified in Albert et al. (2016a) for the same traits, suggesting an overall good repeatability of the measurements and conservation of the QTLs despite lower drought intensity in the present experiment.

eQTL mapping in the RIL progeny for 274 transcripts

We then selected 274 genes among the differentially expressed genes between watering regimes and/or between parental genotypes. They were selected according to their annotation and potential involvement in the response to water deficit, and/or their positional overlap with the main QTLs and QTL by watering regime interaction regions identified in Albert et al. (2016a) and confirmed in the present study. Gene expression levels of the 274 selected genes assessed by microfluidic qPCR was consistent with RNA sequencing data in both leaf and fruit in Cervil, Levovil and their F_1 hybrid, under both watering regimes (Pearson r^2 > 0.65, *P-values* < 1.36 x 10^{-12}) **(Figures S5 and S6).** Large expression variations were observed in the RILs under both conditions, with transgression beyond parental values for all transcripts assessed. A total of 163 eQTLS controlling the expression of 115 genes were detected (125 in fruit and 38 in leaf) **(Tables S17 to S20, Figure 4)**. At least one significant eQTL was associated to 49% of the genes in the fruit (89 genes) and to 30% of the genes in leaf (27 genes), with one to five eQTLs per gene. LOD scores varied from 3.13 to 32.61 and percentage of variation explained (PVE) from 4.45 to

71.57 %. A total of 127 eQTLs had confidence intervals smaller than 5 Mbp (covering between 2 and 633 genes).

Among the 163 eQTLs mapped, 41 were significant under both watering regimes ('constitutive'), whereas 122 were either specific to one watering condition (48 drought and 60 control specific) or detected using expression ratio between watering regimes (14) ('interactive'). Considering the chromosomic location of eQTL relative to their target gene, 47 were *local* eQTLs (21 'constitutive', 26 'interactive') whereas 116 were *trans* acting eQTLs (20 'constitutive', 96 'interactive'). The proportion of *local* vs *trans* eQTLs was not uniform across the LOD score categories. Up to 83% of the eQTLs mapped with LOD scores above 8 were local eQTLs, whereas 78% of the eQTLs detected with LOD scores below 8 were *trans* acting **(Figure 5)**. Moreover, local eQTLs explained on average a higher proportion of the phenotypic variation per transcript compared to *trans* acting eQTLs (*local* eQTLs: average PVE = 25.85 % , sd = 15.94; *trans* eQTLs: average PVE = 14.13 %, sd = 3.69). Contrary to constitutive eQTLs which were equally represented among *local* and *trans* acting eQTLs, up to 79 % of the interactive eQTLs presented a distant location, including all eQTLs identified using expression ratio between watering regimes (**Figure 5**).

Among the 46 genes for which *local* eQTLs were identified (35 in fruit, 11 in leaf), 30 were investigated in ASE tests and 63% of them presented a significant *cis* regulatory pattern in the F1 hybrid. Moreover, we identified significant correlations between allelic ratio in the RILs and allelic ratio in the F1 hybrid for these 30 genes (r^2 = 0.65, *P-value* = 1.84 x 10^{-8}), suggesting consistent direction and intensity of the regulatory divergences between the two independent experiments **(Figure S7).** Among the 87 genes for which *trans* eQTLs were identified (67 in fruit, 19 in leaf), 42 were investigated in ASE tests and 60% of them were identified with a *trans* regulatory pattern when comparing ASE in F1 to allelic expression in parental genotypes.

Focus on a complex network of variation between sugar metabolism related genes

Sixteen of the genes whose expression level was assessed in fruit in the RILs were related to sugar metabolism. These genes corresponded to two cell wall invertases (LIN5: Solyc09g010080 and LIN9: Solyc08g079080), one vacuolar invertase (Solyc03g083910), three neutral invertases (Solyc01g111100, Solyc06g065210 and Solyc11g020610), two invertase inhibitors (Solyc01g088590 and Solyc12g099200), two sucrose synthases (Solyc02g081300 and Solyc12g009300), two sugar transporters (Solyc01g098560 and Solyc04g076960), one fructokinase (Frk3: Solyc02g091490) and three fructose phosphatases (Solyc04g071340, Solyc05g052600 and Solyc07g065900). For six of these genes, we observed significant correlations (Pearson $r^2 > 0.18$ and *P-values* < 0.05) between their expression levels in fruit pericarp at cell expansion stage and sugar-related traits (DMW, SSC, Glucose, Fructose) measured in the mature fruit of the RILs, under control and drought conditions **(Figure S8).** The strongest correlations were observed between LIN9 and DMW under drought condition (Pearson $r^2 = 0.33$, *P-value* = 0.0004) and between LIN5 and SSC under control condition (Pearson $r^2 = 0.26$, *P-value* = 0.004).

Four of these 16 genes were associated to *local* eQTLs with medium to high effects (PVE ranking from 13.93% for Frk3 to 44.34% for Solyc03g083910), whereas five of them were associated with at least one significant *trans* acting eQTL **(Figure 6).** When lowering the detection threshold to an alpha risk of 10% (LOD = 2.81), one additional *local* eQTL for Solyc12g009300 and four additional *trans* eQTLs for LIN9, Solyc01g088590, Solyc03g083910 and Solyc12g009300 were detected. Excepted one *trans* eQTL regulating LIN5 located on chromosome 2 and one *local* eQTL regulating Solyc03g083910, all these eQTLs were either watering condition specific or detected using expression ratio between watering conditions. Remarkably, we observed what seemed to be a complex network of interactions between these genes/loci. The genomic region of chromosome 2 including Frk3, concentrated four *trans* eQTLs for three genes (LIN5, LIN9 and Solyc03g083910) and could constitute a master regulatory region. A second genomic region on chromosome 3, including a vacuolar invertase, concentrated three *trans* eQTLs for an invertase inhibitor, a fructose phosphatase and LIN9 and could represent a second regulatory node. Further functional studies are required to characterize and confirm the interplay among these genes.

Discussion

A large variation in gene expression levels between organs, genotypes and watering regimes and non-additive inheritance of gene expression level in the hybrid

In this work, we first showed an important impact of the genotype, water regime and organ on the transcription levels of a large number of genes. This is consistent with previous studies on the impact of water stress on transcriptomic variations (Seki *et al.*, 2002 in rice; Rabbani *et al.*, 2003 in *Arabidopsis*; Des Marais *et al.*, 2012 in *Arabidopsis*; Albert *et al.*, 2016*a* in *tomato*) and transcriptome specificities in various organs and tissues (Libault *et al.*, 2010 in soybean; Slane *et al.*, 2014 in Arabidopsis; Matas *et al.*, 2011, Koenig *et al.*, 2013 and Pattison *et al.*, 2015 in tomato). The impact of the genetic background on gene expression variations in response to biotic and abiotic stressors is more rarely studied. Notably, in our study, the fruit of the cherry type genotype 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 variation in response to water deficit comparable to the two others genotypes. This could reflect a different source to sink relationship and a probable buffering effect in small fruited tomato genotypes as proposed in Ripoll *et al.* (2016).

Transcript abundance can be considered as a quantitative trait and its heritability and inheritance vary from one gene to the other. In tomato, until now, inheritance studies were performed on phenotypic, metabolic and proteomic traits (Steinhauser *et al.*, 2011; Pascual *et al.*, 2013). These studies showed that heterosis was rare, with few plant and fruit traits showing over dominance. Our results on plant vigor and fruit quality phenotypes confirmed this pattern, with a majority of additive inheritance whatever the watering regime. On the contrary, at the gene expression levels, we identified a majority of genes showing non-additive inheritance (> 80%) and a limited conservation of the inheritance patterns between organs and watering regimes. Previous plant studies have investigated gene expression inheritance patterns in inter and intra-specific hybrids. The fraction of non-additive gene expression level inheritance varied greatly between species, tissues, experimental procedures and statistical approaches. Nevertheless, non-additive gene expression in hybrids appeared to be a common occurrence in plants (Swanson-Wagner *et al.*, 2006, Jahnke *et al.*, 2010 and Paschold *et al.*, 2012 in maize; Bell *et al.*, 2013 in *Cirsium arvense*; Combes *et al.*, 2015 in coffee tree; Lovell *et al.*, 2016 in *Panicum hallii*). Lippman and Zamir

(2007) suggested the absence of direct link between expression changes caused by heterozygosity and hybrid vigor.

Genetic control of expression variation and interaction with the water regime

We combined two approaches in order to assess the genetic control of variation in gene expression levels in interaction with the watering regime. First, using 48,909 exonic SNPs identified between parental genotypes and comparing ASE in the F_1 to the parental allelic expression, we assessed the extent of regulatory divergence at the whole transcriptome scale. Although almost two million SNPs could be identified between Cervil and Levovil (Causse *et al.*, 2013), most of them are in non-coding regions and only one third of the genes had informative SNPs to assess ASE in the F_1. The approach was nevertheless powerful enough to identify 4,424 genes showing constitutive regulatory divergences (3,005 in leaf and 2,857 in fruit) and 425 genes showing significant ASE by watering regime interaction (328 in leaf and 113 in fruit).

A total of 70% of the *cis* regulated genes exhibited greater expression level from Levovil allele than Cervil allele in the F1 hybrid both in fruit and in leaf. The non-symmetrical divergence of *cis*-regulatory polymorphisms between parental alleles could have, in part, resulted from a mapping bias, as Levovil genome was shown to be closer to the tomato reference genome than Cervil (Causse *et al.*, 2013). However, the read mapping success rates at the gene and allele level were equivalent for Cervil, Levovil and F1 libraries, suggesting that our results are unlikely to be severely affected by mapping bias. Hybrids are widely used in modern agriculture, either for their heterosis (the advantage of a hybrid compared to both parents) or for the combination of dominant traits. This is particularly the case in tomato where six to eight dominant disease resistance genes are cumulated in modern F1 varieties (Scott, 2005). ASE could be at the origin of heterosis. In maize, attempts were made to investigate the heterosis phenomenon in regard 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, 2007*a,b*). Such

directional bias has also been described in tomato for the *SINGLE FLOWER TRUSS* locus (Jiang et al, 2013).

Swinnen et al. (2016) reported that in different crop species 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 (Cubillos *et al.*, 2012*a*; Signor and Nuzhdin, 2018). 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* mutations 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). Similar observations was made for the FW2.2 and FW3.2 QTLs controlling tomato fruit size (Frary et al. 2000; Chakrabarti et al. 2013). The ultimate identification of the mutations responsible for *cis* regulatory divergences is feasible using the ASE approach but requires the comparison of several F_1 as proposed by Kang et al. (2016).

ASE analysis merges all non-*cis* regulatory divergences as *trans* effects and does not permit estimation of the number, the genomic location and the specific effect of *trans* acting polymorphisms. We thus completed the ASE approach by studying the variation in gene expression levels for 274 transcripts in RILs. We identified 163 eQTLs (47 *local* and 116 *trans*), mostly confirming the regulatory divergences observed in the ASE approach. Most of local eQTLs were constitutive over watering conditions with moderate to high effects, probably corresponding to *cis* regulated genes. On the contrary, *trans* acting eQTLs mostly presented low to moderate effects varying with watering conditions. Ranjan et al. (2016) also identified in an interspecific progeny of tomato a majority of trans acting eQTL.

To the best of our knowledge the genetic determinants of natural transcriptomic variations in response to environmental constraints have not yet been studied in tomato. Cubillos *et al.* (2014), Lovell *et al.* (2016) and Waters *et al.* (2016) recorded genotype-dependent transcriptomic responses to water, cold and heat stresses mediated through a large majority of *cis* acting by environment interactions in *Arabidopsis thaliana*, *Panicum halli* and *Zea*

mays, respectively. On the other hand, Lowry *et al.* (2013), Snoek *et al.* (2013) and Clauw *et al.* (2016) reported a preponderance of *trans* acting eQTL by environment interactions in *Arabidopsis thaliana* subjected to drought and shade conditions. In our study, with both approaches (eQTL and ASE), we identified a majority of trans by watering regime interactions, suggesting that water deficit resulted in the activation of *trans* acting regulatory elements specific to adaptation to the abiotic constraint.

A complex regulatory network of invertase genes and a few candidates for sugar content

There is increasing evidence that sucrose and hexoses may play a non-nutritive role as a regulator of cellular metabolism, possibly by altering gene expression and playing the role of signal molecules in source-sink regulation (Eveland and Jackson, 2012). Eleven invertase genes are identified in the tomato genome. Cell wall invertases play a crucial role in sucrose partitioning, plant development and cell differentiation, whereas vacuolar invertases are involved in cell expansion, sugar storage and regulation of cold induced sweetening. The role of cytoplasmic neutral invertases remains unclear. They seem to play a role in the antioxidant system involved in cellular reactive oxygen species homeostasis and they have been shown to be involved in responses to abiotic stresses (Albacete *et al.*, 2014, 2015; Chen *et al.*, 2016; Liu *et al.*, 2016; Xu *et al.*, 2017). We showed a strong connectivity between several genes related to sugar metabolism. Coordination of their expression could play a central role in source-sink regulation. Further functional studies are required to characterize the interplay among these genes and their polymorphisms.

In conclusion, our study underlined the importance of the genotype background in gene expression variation. It also revealed the impact of water stress and its interaction with genotypes on phenotypes and gene expression. Finally we showed how ASE and eQTL are complementary approaches to reveal *cis* and *trans* regulated effects responsible for the variation of gene expression and to identify networks of co-regulated genes.

Experimental Procedures

Plant material and experimental design

The plant material was composed of two inbred genotypes, their F_1 hybrid and 124 recombinant inbred lines (RILs) developed from the same cross. The male parent (Cervil) is a cherry type tomato (*S. lycopersicum cerasiforme, 4 – 8 g*), whereas the female parent (Levovil) is a large fruited accession (*S. lycopersicum,* 90 – 160 g) (Saliba-Colombani et al., 2000). Plants were grown in a glasshouse at INRA Avignon, from March to July 2015, following the same experimental design as in Albert *et al.* (2016a). Briefly, two watering regimes were applied to the plants: drought (D) and control (C). The drought treatment was set progressively 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, in order to apply a mild water deficit. Two plants per watering regime for the RILs and three plants per watering regime for the parental lines and the F_1 hybrid were randomized within the greenhouse.

Plant growth and fruit quality phenotyping

Plants were phenotyped for traits describing plant growth and fruit quality. Briefly, flowering date (Flw, days after sowing), height (Ht, cm) and stem diameter (Diam, mm) were measured on each plant. Ten mature fruits per genotype and watering regime were harvested on 3rd to 6th truss and weighted (FW, g). Fruits were pooled in three groups of three to four fruits per watering regime to constitute 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 blended to measure pH and soluble solid content (SSC, °brix). Fresh pericarps were sampled from the remaining fruits of each pool, frozen in liquid nitrogen and ground into powder for sugar (Glucose and Fructose, g 100 g^{-1} FM), malic acid (Malic, g 100 g^{-1} FM) and total vitamin C (VitC, mg 100 g^{-1} FM) content assessment according to protocols described in Wu *et al.* (2002) and Stevens *et al.* (2006), respectively.

Statistical analysis on plant growth and fruit quality phenotypes

All statistical analyses were performed using R version 3.1.2 (R Development Core Team, 2012) considering significance with an alpha risk of 5%. 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 parents and F_1 hybrid individual data, then on the data collected in RILs, according to the following model:

$Y_{ijk} = \mu + G_i + W_j + G_i * W_j + e_{ijk}$ where

Y_{ijk} was the phenotypic value of genotype i in watering regime j, μ the overall mean of the trait, G_i the fixed effect of genotype i, W_j the fixed effect of watering regime j and e_{ijk} the residual error effect. Then, for each phenotypic trait, we compared the low parent (LP), the high parent (HP) and the F_1 hybrid means using Tukey's post-hoc tests. The comparisons distinguished between six inheritance classes. Inheritance was considered as '*additive*' when (LP < F1 < HP); '*recessive*' when (LP = F1 < HP); '*dominant*' when (LP < F1 = HP); '*over recessive*' when (F1 < LP ≤ HP); '*over dominant*' when (LP ≤ HP < F1) and '*non-significantly different*' (NS) when (LP ≤ F1 ≤ HP, LP < HP) or (F1 ≤ LP ≤ HP, F1 < HP) or (LP ≤ HP ≤ F1, LP < F1) or (LP = HP = F1).

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: $Y_{ij} = \mu + G_i + e_{ij}$ (G_i and e_{ij} random). Broad-sense heritability estimates (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 $\sigma^2_{Total} = \sigma^2_G + 1/n*\sigma^2e$ (n, number of replicates per genotype). Correlation between H^2 measured under control and drought conditions were estimated using the Spearman coefficient. Pearson coefficients were computed to estimate correlations between trait means under both watering regimes and with phenotypic data collected in Albert et al. (2016a) for the same genotypes under a similar experimental design. Phenotypic plasticity was computed for each trait as : $\Delta_{ki} = (D_{ki} - C_{ki})/C_{ki}$, with Δ_{ki} the plasticity value for trait k and genotype i, D_{ki} the mean of trait k under drought condition for genotype i and C_{ki} the mean of trait k under control condition for genotype i. The RIL mean trait values in each watering regime and the phenotypic plasticity traits were used in subsequent QTL analyses **(Table S1).**

Fruit and leaf tissue sampling and RNA extraction

Samples of growing leaves and pericarps from fruits at the cell expansion stage were collected on each plant and immediately frozen in liquid nitrogen. Fruits were collected at 21 DAA (days after anthesis) for large fruited genotypes and 14 DAA for small fruited genotypes to normalize for the fruit developmental stage **(Table S1).** For each organ, frozen samples were separately ground to get biological replicates for each watering regime: three for the parental genotypes, two (control) to three (drought) for the F_1 hybrid and one for each RIL. RNA was extracted using the Spectrum Plant Total RNA kit (Sigma-Aldrich) following the manufacturer's protocol and treated with On-Column DNase I Digestion Set (Sigma-Aldrich) to remove any remaining genomic DNA. RNA purity and integrity were assessed on Nanodrop 1000 (ThermosFisher) and Bioanalyser 2100 (Agilent), respectively.

RNA sequencing of parental lines and F1 hybrid

A total of 33 messenger RNA paired-end strand specific libraries were constructed from parental lines and F1 hybrid RNA (one library per biological replicate per organ per watering regime). The library preparation was conducted using '*TruSeq Stranded mRNA Sample Preparation' kits* (Illumina, San Diego, CA) following the manufacturer instructions, aiming to obtain insert sizes ranging from 100 to 400 base pairs (bp). Indexed parental and F1 hybrid libraries were then combined in two lanes (one for the leaf and one for the fruit samples) and were subjected to 150-bp paired-end Illumina high-throughput sequencing on a Hiseq 3000 at the GeT-PlaGe platform (INRA Toulouse, France). The minimal, maximal and average amount of raw sequencing data per sample were estimated to 1.62×10^9 bp, 4.94×10^9 bp and 3.21×10^9 bp, respectively.

RNA-sequencing data processing and read count generation

Raw sequencing data quality was assessed with *FASTQC v.0.11.5* (Babraham, 2011). Sequencing adapters were trimmed out using *Cutadapt 1.9.1* (Martin, 2011). Then, data were cleaned and filtered as described in Causse et al. (2013). On average, cleaning steps removed 7.38% of the data (min = 2.70%, max = 9.37%). Remaining data were aligned to the tomato reference genome (Heinz 1706, v2.5) using *TopHat2 v2.1.1* (Kim *et al.*, 2013) with

default parameters and providing the tomato gene model (annotation v2.4) to support the mapping process. Alignments were filtered to keep only concordantly mapped reads using *SAMtools 1.3.1* (Li *et al.*, 2009) and read counts per gene were generated for each library using *HTSeq-Count 0.6.1* (Anders *et al.*, 2014). On average, 18 million pairs of reads were successfully mapped per library (min = 9×10^{6} read pairs, max = 27×10^{6} read pairs), representing 84 to 94% of the cleaned data. No significant difference was observed in the percentage of mapped reads onto the reference genome between genotypes (ANOVA, *mapped reads ~ genotype*, *P-value* > 0.05). Non-zero read counts were obtained for 25,998 genes (77% of the known tomato CDS) across leaf samples and 25,572 genes (76% of the known tomato CDS) across fruit samples.

Differential gene expression analysis

Counts from mapped reads of leaf and fruit samples were analyzed separately using the Bioconductor package *DESeq2* version 1.6.3 (Love *et al.*, 2014). Count data were first normalized using the Relative Log Expression (RLE) method (Anders and Huber, 2010) to correct for differences among library sizes. Then, weakly expressed genes were filtered out using the *HTSFilter* function as recommended in Rau et al. (2013) to avoid the negative effect of background expression noise. A total of 18,843 genes among leaf samples and 19,097 genes among fruit samples were retained in the analysis, corresponding to genes with maximum normalized read count across samples above 25. Leaf and fruit normalized read counts are available in **Tables S2 and S3,** respectively.

A regularized log transformation was applied to normalize gene expression levels prior to perform a principal coordinate analysis (PCA). Per-gene dispersions were estimated by incorporating data-driven prior distributions and negative binomial generalized linear models were fitted to estimate moderated $\log_2$ fold changes in gene expression levels between genotypes (F1, Cervil or Levovil) and between watering regimes (control or drought). Per-gene Wald test statistics were computed in order to identify significantly differentially expressed genes between genotypes in each watering regime and between watering regimes for a given genotype. A FDR threshold of 5% was applied to call significantly differentially expressed genes (Benjamini and Hochberg, 1995). *P-values* from

the Wald tests between pairs of genotypes under each watering regime were used to determine the gene expression level inheritance patterns following the classification defined in the phenotypic analysis.

Allele specific quantification and allele specific expression test

Allele-specific gene expression (ASE) levels were estimated using 48,909 high quality exonic SNPs between both parental genotypes identified from high depth genome re-sequencing (Cervil: 13.3x, Levovil: 6.7x completed to 50x, Causse *et al.*, 2013). The exonic SNPs were distributed over 12,391 tomato genes (36% of the known tomato CDS). Individual RNA-seq reads tended to span multiple SNPs. 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 from either parental haplotype, using a homemade Perl script (available upon request). The allele-specific detection was applied on the parental and F_1 libraries. Reads that could not be consistently assigned to either haplotype were discarded (in average 4.5 reads per gene across the 33 libraries). On average 25% of the reads mapped at the gene level were assigned to one allele per gene and per library.

Leaf and fruit allele specific counts were analyzed separately using *DESeq2* version 1.6.3 (Love *et al.*, 2014). First, the *calcNormFactors* function was used to scale both maternal and paternal counts from each F_1 hybrid library by the same normalization factor to avoid biasing the allele specific expression when normalizing read counts. Normalized allele specific counts were filtered using the *HTSfilter* function. A total of 8,434 genes among leaf samples and 7,459 genes among fruit samples with maximum normalized allele specific read count across samples above five were retained. Leaf and fruit normalized allele specific read counts are available in **Tables S4 and S5**, respectively.

Per-gene dispersion of expression levels were estimated by incorporating data-driven prior distributions and negative binomial generalized linear models were fitted for each gene considering *generation* (G: F1 or parent), *allele* (A: Cervil or Levovil) and *watering regime* (W: control or drought) factors and their two ways and three ways interactions (G x A; G x W; A x W and G x A x W). Per gene Wald tests were computed to identify *cis* regulatory divergences (A main effect in F1), *trans* regulatory divergences (G x A interactions), *cis by*

watering interactions (A x W in F1) and *trans* by *watering* interactions (G x A x W) as in Lovell et al. (2016). A FDR threshold of 5% called significant differences.

Gene ontology enrichment analyses

Enrichments in gene ontologies (GO) associated to biological process was achieved using the Wallenius non-central hypergeometric distribution in the Bioconductor package *GOseq*, an extension of the standard Fisher's exact test able to account for selection bias due to gene length (Young *et al.*, 2010). Gene lengths (based on the ORF, in bp) and GO associated to each gene were retrieved from the ITAG v2.4 (Slim GO terms, International Tomato Annotation Genome). The gene space was composed of the 24,968 annotated tomato genes (72% of the known CDS). A FDR threshold of 5% called significantly enriched GO.

Primer design and quantitative gene expression analysis in the RILs

Gene expression variation of 274 target genes (183 in fruit and 91 in leaves; **Tables S6 and S7**) were quantified under both watering regimes in Cervil, Levovil, their F_1 hybrid and the 124 RILs by quantitative real time microfluidigm PCR (performed on the Gentyane platform, INRA Clermont-Ferrand, France).

Primer3 (Rozen and Skaletsky, 1999) designed primers pairs for each targeted transcript. Gene-specific primer lists are available in **Tables S6 and S7**. First strand cDNAs were synthetized from the RIL RNA samples using oligod(T) and Superscript III enzyme, followed by a Ribonuclease H treatment to remove any RNA in the samples. 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. In each sample, remaining primer pairs were removed using an exonuclease I treatment avoiding any interference in the qPCR reaction. Preamplicons were loaded into the BioMark HD system (Fluidigm). On each of the BioMark HD chip, a negative control, four dilution points and three 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 from two reference genes for leaf samples and three reference genes for fruit samples using the $2^{-\Delta CT}$ method **(Table S8)**. Leaf and fruit normalized expression data are available in **Tables S6 and S7**, respectively.

Normalized expression levels under both watering regimes and ratio of normalized expression under drought relative to normalized expression under control for each genotype were used in the subsequent eQTL mapping analysis. Pearson coefficients were computed to estimate correlations between gene expression measured through RNA sequencing and expression measured though microfluidigm qPCR in leaf and fruit samples of Cervil, Levovil and the F1 hybrid, under control and drought conditions. Pearson coefficients were computed as well to estimate correlation between the expression levels of 16 genes related to sugar metabolism measured in fruit and phenotypic traits related to sugar metabolism measured in the RILs.

(e)QTL and (e)QTL by watering regime interaction mapping

Plant vigour, fruit quality and gene expression levels recorded in the RILs were used in mapping analysis. Log10 transformations were applied when distribution were skewed. Analysis were performed using a genetic map constituted of 501 SNP markers (Albert et al. 2016a). (e)QTLs were identified by simple interval using the EM algorithm method implemented in R/QTL package (Broman et al. 2003). A 1000-permutation test was computed to identify the LOD threshold at a genome wide significance level α = 0.05 (LOD=3.13). For each detected QTL, position, LOD score, marker at the peak, confidence interval (LOD decrease one unit, genetic CI in cM), average values of the two parental alleles and percentage of phenotypic variation explained (PVE) were reported. Genetic CI were translated into physical intervals (in Mbp) onto the tomato genome (assembly v2.5). When a QTL was detected under one watering regime, the average values of parental alleles and PVE was also calculated in the second watering regime. We distinguished *constitutive*

(e)QTLs (detected under both watering regimes with consistent effects), from *interactive* (e)QTLs detected under control or drought only or using plasticity traits. 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.

Data availability

Illumina sequencing data can be found in the NCBI SRA repository under project number PRJNA472702. The allelic value and location of SNP markers can be found in supplemental data of Albert et al, 2016. Resequencing data of the parental lines can be found in the European Nucleotide Archive (ENA) with the project number PRJEB4395 (http://www.ebi.ac.uk/ena/data/view/PRJEB4395). BAM files and SNP characteristics of the genomic sequences are available on the SolGenomics ftp site ftp.solgenomics.net/projects/causse_tomato_snp8lines. Other data (read counts, expression data from fluidigm) are available as supplemental data.

Author contributions

E.A. supervised sample collection and phenotypic measurements, analyzed data and wrote the manuscript. R.D. performed RNA extractions and RNA-seq library preparation. M.L. and S.S. supervised library preparation and sequencing. M.B, J.P.B. and F.B. analyzed RNA-seq reads. J.G and M.C.D collected samples, performed phenotypic measurements and RNA extractions. C.P. and V.G. performed microfluidigm qPCR experiment. J.J.G developed and shared allele specific expression mapping scripts. G.R. supported statistical analyses. M.C. supervised the project, built the experimental design and revised the manuscript. All authors discussed the results and commented the manuscript.

Albert, E., Duboscq, R., Latreille, M., Santoni, S., Beukers, M., Bouchet, J.-P., Bitton, F., Gricourt, J., Poncet, C., Gautier, V., Jimenez-Gomez, J., Rigaill, G., Causse, M. (2018). Allele-specific expression and genetic determinants of transcriptomic variations in response to mild water deficit in tomato. Plant Journal, 96 (3), 635-650. , DOI : 10.1111/tpj.14057

Acknowledgments

Supporting Legends

Figure S1. Mean comparisons for eleven phenotypic traits measured in Cervil, Levovil and their F1 under two watering regimes.

Figure S2. Distribution of the average values for eleven phenotypic traits measured in the RILs under two watering regimes.

Figure S3. Principal Coordinate Analysis (PCoA) performed on gene expression measured in leaf (A) and fruit (B) through RNA sequencing. Normalized counts were transformed using the regularized log transformation in *DESeq2* package.

Figure S4. Differentially expressed genes in each possible comparison between parents and F1 hybrid in leaf (left, green) and fruit (right, red), under both watering regimes.

Figure S5. Pearson correlation between gene expression measured by RNA sequencing and microfluidigm qPCR in leaf samples of Cervil, Levovil and their F1 hybrid under control (C) and drought (D) conditions.

Figure S6. Pearson correlation between gene expression measured by RNA sequencing and microfluidigm qPCR in fruit samples of Cervil, Levovil and their F1 hybrid under control (C) and drought (D) conditions.

Figure S7. Correlation between allelic ratio estimated in the RILs using microfluidigm qPCR expression data and allelic ratio estimated in the F1 hybrid using normalized allele specific counts for a total of 30 genes with significant local eQTLs assessed in ASE tests.

Figure S8. Correlation between sugar related traits measured in mature fruit pericarps and the expression of 16 genes related to sugar metabolism measured in fruit pericarps at cell expansion stage in the RILs under control and drought conditions.

Table S1. Effect of the genotype (G), the watering regime (W) and the interaction (G x W) on the phenotypic traits in Cervil, Levovil, their F1 hybrid and the 124 RILs.

Table S2. Pearson correlations between phenotypic traits measured in the RILs in 2013 (Albert *et al.* 2016) and 2015.

Table S3. Pearson correlations between phenotypic traits measured in control and drought in the RILs in 2015.

Table S4. Inheritance of gene expression levels in the F1 hybrid in leaf and fruit, under control and drought conditions. P-values from the Wald tests between pairs of genotypes under each watering regime were used to determine the gene expression level inheritance patterns.

Table S5. Complete lists of GO terms associated to biological process over-represented within the set of genes differentially expressed between watering regimes across comparisons in leaf and fruit.

Table S6. Complete list of GO terms associated to biological process over-represented within the set genes presenting significant *Cis*, *Trans*, *Cis* x W and *Trans* x W regulatory divergences in leaf and fruit.

Table S7. QTLs detected for plant vigor and fruit quality traits average values measured in the 124 RILs.

Table S8. QTLs detected for phenotypic plasticity traits measured in the 124 RILs.

Table S9. eQTLs detected on log10 transformed normalized gene expression values measured by microfluidgm qPCR in leaf samples of the 124 RILs.

Table S10. eQTLs detected on log10 transformed ratio of expression under drought relative to expression under control condition measured by microfluidgm qPCR in leaf samples of the 124 RILs.

Table S11. eQTLs detected on log10 transformed normalized gene expression values measured by microfluidgm qPCR in fruit samples of the 124 RILs.

Table S12. eQTLs detected on log10 transformed ratio of expression under drought relative to expression under control condition measured by microfluidgm qPCR in fruit samples of the 124 RILs.

Table S13. Untransformed phenotypic means for the 11 plant vigour and fruit quality traits measured on Cervil, Levovil, their F1 and 124 RILs derived from the same cross.

Table S14. Normalized gene level RNAseq counts for leaf libraries.

Table S15. Normalized gene level RNAseq counts for fruit libraries.

Table S16. Normalized allele level RNAseq counts for leaf libraries.

Table S17. Normalized allele level RNAseq counts for fruit libraries.

Table S18. Gene specific primers and untransformed normalized microfruidigm qPCR expression data for the 91 genes assessed in leaf.

Table S19. Gene specific primers and untransformed normalized microfruidigm qPCR expression data for the 183 genes assessed in fruit.

Table S20. Reference genes used for normalization of microfluidigm qPCR expression data.

References

Albacete, A., Cantero-Navarro, E., Großkinsky, D.K., et al. (2015) Ectopic overexpression of the cell wall invertase gene CIN1 leads to dehydration avoidance in tomato. *J. Exp. Bot.*, **66**, 863–78. Available at: https://academic.oup.com/jxb/article-lookup/doi/10.1093/jxb/eru448.

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. *Biotechnol. Adv.*, **32**, 12–30. Available at: http://dx.doi.org/10.1016/j.biotechadv.2013.10.005.

Albert, E., Gricourt, J., Bertin, N., Bonnefoi, J., Pateyron, S., Tamby, J.-P., Bitton, F. and Causse, M. (2016) Genotype by watering regime interaction in cultivated tomato: lessons from linkage mapping and gene expression. *Theor. Appl. Genet.*, **129**, 395–418. Available at: http://link.springer.com/10.1007/s00122-015-2635-5.

Albert, E., Segura, V., Gricourt, J., Bonnefoi, J., Derivot, L. and Causse, M. (2016) Association mapping reveals the genetic architecture of tomato response to water deficit: focus on major fruit quality traits. *J. Exp. Bot.*, **67**, 6413–6430. Available at: https://academic.oup.com/jxb/article-lookup/doi/10.1093/jxb/erw411.

Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106. Available at: http://genomebiology.biomedcentral.com/articles/10.1186/gb-2010-11-10-r106.

Anders, S., Pyl, P.T. and Huber, W. (2015) HTSeq: a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–9. Available at: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btu638.

Andrews, S. (2011) FastQC A quality control tool for high throughput sequence data. Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bell, G.D.M., Kane, N.C., Rieseberg, L.H. and Adams, K.L. (2013) RNA-seq analysis of allele-specific expression, hybrid effects, and regulatory divergence in hybrids compared with their parents from natural populations. *Genome Biol. Evol.*, **5**, 1309–23. Available at: https://academic.oup.com/gbe/article-lookup/doi/10.1093/gbe/evt072.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate : a practical and powerful approach to multiple testing. *J. R. Stat. Soc.*, **57**, 289–300. Available at: https://www.jstor.org/stable/2346101.

Brem, R.B., Yvert, G., Clinton, R. and Kruglyak, L. (2002) Genetic dissection of transcriptional regulation in budding yeast. *Science*, **296**, 752–5. Available at: http://www.sciencemag.org/cgi/content/abstract/296/5568/752.

Bystrykh, L., Weersing, E., Dontje, B., et al. (2005) Uncovering regulatory pathways that affect hematopoietic stem cell function using "genetical genomics". *Nat. Genet.*, **37**, 225–32. Available at: http://www.nature.com/articles/ng1497.

Carrari, F., Baxter, C., Usadel, B., et al. (2006) Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiol.*, **142**, 1380–96. Available at: http://www.plantphysiol.org/cgi/doi/10.1104/pp.106.088534.

Castel, S.E., Levy-Moonshine, A., Mohammadi, P., Banks, E. and Lappalainen, T. (2015) Tools and best practices for data processing in allelic expression analysis. *Genome Biol.*, **16**, 195. Available at: http://genomebiology.com/2015/16/1/195.

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. Available at: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-791.

Chakrabarti, M., Zhang, N., Sauvage, C., et al. (2013) A cytochrome P450 regulates a domestication trait in cultivated tomato. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 17125–30. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1307313110.

Chen, S.-F., Liang, K., Yin, D.-M., Ni, D.-A., Zhang, Z.-G. and Ruan, Y.-L. (2016) Ectopic expression of a tobacco vacuolar invertase inhibitor in guard cells confers drought tolerance in Arabidopsis. *J. Enzyme Inhib. Med. Chem.*, **31**, 1381–1385. Available at: https://www.tandfonline.com/doi/full/10.3109/14756366.2016.1142981.

Chesler, E.J., Lu, L., Shou, S., et al. (2005) Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat. Genet.*, **37**, 233–42. Available at:

http://www.nature.com/doifinder/10.1038/ng1518.

Cheung, V.G., Spielman, R.S., Ewens, K.G., Weber, T.M., Morley, M. and Burdick, J.T. (2005) Mapping determinants of human gene expression by regional and genome-wide association. *Nature*, **437**, 1365–9. Available at: http://www.nature.com/doifinder/10.1038/nature04244.

Clauw, P., Coppens, F., Korte, A., et al. (2016) Leaf growth response to mild drought: natural variation in Arabidopsis sheds light on trait architecture. *Plant Cell*, **28**, 2417–2434. Available at: http://www.plantcell.org/lookup/doi/10.1105/tpc.16.00483.

Combes, M.-C., Hueber, Y., Dereeper, A., Rialle, S., Herrera, J.-C. and Lashermes, P. (2015) Regulatory divergence between parental alleles determines gene expression patterns in hybrids. *Genome Biol. Evol.*, **7**, 1110–21. Available at: https://academic.oup.com/gbe/article-lookup/doi/10.1093/gbe/evv057.

Cubillos, F.A., Coustham, V. and Loudet, O. (2012) Lessons from eQTL mapping studies: non-coding regions and their role behind natural phenotypic variation in plants. *Curr. Opin. Plant Biol.*, **15**, 192–8. Available at: http://dx.doi.org/10.1016/j.pbi.2012.01.005.

Cubillos, F.A., Stegle, O., Grondin, C., Canut, M., Tisné, S., Gy, I. and Loudet, O. (2014) Extensive cis-regulatory variation robust to environmental perturbation in Arabidopsis. *Plant Cell*, **26**, 4298–310. Available at: http://www.plantcell.org/lookup/doi/10.1105/tpc.114.130310.

Cubillos, F.A., Yansouni, J., Khalili, H., et al. (2012) Expression variation in connected recombinant populations of Arabidopsis thaliana highlights distinct transcriptome architectures. *BMC Genomics*, **13**, 117. Available at: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-13-117.

DeCook, R., Lall, S., Nettleton, D. and Howell, S.H. (2006) Genetic regulation of gene expression during shoot development in Arabidopsis. *Genetics*, **172**, 1155–64. Available at: http://www.genetics.org/cgi/doi/10.1534/genetics.105.042275.

Drost, D.R., Benedict, C.I., Berg, A., et al. (2010) Diversification in the genetic architecture of gene expression and transcriptional networks in organ differentiation of Populus. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 8492–7. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0914709107.

Druka, A., Potokina, E., Luo, Z., Jiang, N., Chen, X., Kearsey, M. and Waugh, R. (2010) Expression quantitative trait loci analysis in plants. *Plant Biotechnol. J.*, **8**, 10–27. Available at: http://doi.wiley.com/10.1111/j.1467-7652.2009.00460.x.

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. and Li, W.-H. (2010) Natural selection on cis and trans regulation in yeasts. *Genome Res.*, **20**, 826–836. Available at: http://genome.cshlp.org/cgi/doi/10.1101/gr.101576.109.

Emerson, J.J. and Li, W.-H. (2010) The genetic basis of evolutionary change in gene expression levels. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **365**, 2581–90. Available at: http://rstb.royalsocietypublishing.org/cgi/content/abstract/365/1552/2581.

Eveland, A.L. and Jackson, D.P. (2012) Sugars, signalling, and plant development. *J. Exp. Bot.*, **63**, 3367–3377. Available at: https://academic.oup.com/jxb/article-lookup/doi/10.1093/jxb/err379.

Frary, A., Nesbitt, T.C., Grandillo, S., et al. (2000) fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. *Science*, **289**, 85–8. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.289.5476.85.

Fraser, H.B. (2011) Genome-wide approaches to the study of adaptive gene expression evolution: systematic studies of evolutionary adaptations involving gene expression will allow many fundamental questions in evolutionary biology to be addressed. *Bioessays*, **33**, 469–77. Available at: http://doi.wiley.com/10.1002/bies.201000094.

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. Available at: http://link.springer.com/10.1007/s00122-010-1456-9 [Accessed November 21, 2014].

He, F., Arce, A.L., Schmitz, G., Koornneef, M., Novikova, P., Beyer, A. and Meaux, J. de (2016) The footprint of polygenic adaptation on stress-responsive cis-regulatory divergence in the Arabidopsis genus. *Mol. Biol. Evol.*, **33**, 2088–101. Available at: https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msw096.

Hubner, N., Wallace, C.A., Zimdahl, H., et al. (2005) Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat. Genet.*, **37**, 243–53. Available at: http://www.nature.com/doifinder/10.1038/ng1522.

Jahnke, S., Sarholz, B., Thiemann, A., Kühr, V., Gutiérrez-Marcos, J.F., Geiger, H.H., Piepho, H.-P. and Scholten, S. (2010) Heterosis in early seed development: a comparative study of F1 embryo and endosperm tissues 6 days after fertilization. *Theor. Appl. Genet.*, **120**, 389–400. Available at: http://link.springer.com/10.1007/s00122-009-1207-y.

Kang, E.Y., Martin, L.J., Mangul, S., et al. (2016) Discovering single nucleotide polymorphisms regulating human gene expression using allele specific expression from RNA-seq data. *Genetics*, **204**, 1057–1064. Available at: http://www.genetics.org/cgi/doi/10.1534/genetics.115.177246.

Keurentjes, J.J.B., Fu, J., Terpstra, I.R., 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. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0610429104.

Kiekens, R., Vercauteren, A., Moerkerke, B., Goetghebeur, E., Daele, H. Van Den, Sterken, R., Kuiper, M., Eeuwijk, F. van and Vuylsteke, M. (2006) Genome-wide screening for cis-regulatory variation using a classical diallel crossing scheme. *Nucleic Acids Res.*, **34**, 3677–86. Available at: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkl510.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36. Available at:

http://genomebiology.biomedcentral.com/articles/10.1186/gb-2013-14-4-r36.

Kirst, M., Myburg, A.A., León, J.P.G. De, Kirst, M.E., Scott, J. and Sederoff, R. (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**, 2368–78. Available at: http://www.plantphysiol.org/cgi/doi/10.1104/pp.103.037960.

Koenig, D., Jiménez-Gómez, J.M., Kimura, S., et al. (2013) Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, E2655-62. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1309606110.

Li, H., Handsaker, B., Wysoker, A., et al. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–9. Available at: http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btp352.

Libault, M., Farmer, A., Joshi, T., 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. Available at: http://doi.wiley.com/10.1111/j.1365-313X.2010.04222.x.

Lippman, Z.B. and Zamir, D. (2007) Heterosis: revisiting the magic. *Trends Genet.*, **23**, 60–6. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0168952506004094.

Liu, M., Yu, H., Zhao, G., Huang, Q., Lu, Y. and Ouyang, B. (2017) Profiling of drought-responsive microRNA and mRNA in tomato using high-throughput sequencing. *BMC Genomics*, **18**, 481. Available at: http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-017-3869-1.

Liu, Y.-H., Offler, C.E. and Ruan, Y.-L. (2016) Cell wall invertase promotes fruit set under heat stress by suppressing ROS-independent cell death. *Plant Physiol.*, **172**, 163–80. Available at: http://www.plantphysiol.org/lookup/doi/10.1104/pp.16.00959.

Liu, Z., Yang, J., Xu, H., Li, C., Wang, Z., Li, Y., Dong, X. and Li, Y. (2014) Comparing computational methods for identification of allele-specific expression based on next generation sequencing data. *Genet. Epidemiol.*, **38**, 591–8. Available at: http://doi.wiley.com/10.1002/gepi.21846.

Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550. Available at: http://genomebiology.com/2014/15/12/550.

Lovell, J.T., Schwartz, S., Lowry, D.B., et al. (2016) Drought responsive gene expression regulatory divergence between upland and lowland ecotypes of a perennial C4 grass. *Genome Res.*, **26**, 510–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26953271.

Lowry, D.B., Logan, T.L., Santuari, L., Hardtke, C.S., Richards, J.H., DeRose-Wilson, L.J., McKay, J.K., Sen, S. and Juenger, T.E. (2013) Expression quantitative trait locus mapping across water availability environments reveals contrasting associations with genomic features in Arabidopsis. *Plant Cell*, **25**, 3266–79. Available at:

http://www.plantcell.org/cgi/doi/10.1105/tpc.113.115352.

Marais, D.L. Des, McKay, J.K., Richards, J.H., Sen, S., Wayne, T. and Juenger, T.E. (2012) Physiological genomics of response to soil drying in diverse Arabidopsis accessions. *Plant Cell*, **24**, 893–914. Available at: http://www.plantcell.org/cgi/doi/10.1105/tpc.112.096180 [Accessed August 1, 2014].

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **17**, 10. Available at: http://journal.embnet.org/index.php/embnetjournal/article/view/200.

Matas, A.J., Yeats, T.H., Buda, G.J., 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**, 3893–910. Available at: http://www.plantcell.org/lookup/doi/10.1105/tpc.111.091173.

Monks, S.A., Leonardson, A., Zhu, H., Cundiff, P., Pietrusiak, P., Edwards, S., Phillips, J.W., Sachs, A. and Schadt, E.E. (2004) Genetic inheritance of gene expression in human cell lines. *Am. J. Hum. Genet.*, **75**, 1094–105. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0002929707600767.

Morley, M., Molony, C.M., Weber, T.M., Devlin, J.L., Ewens, K.G., Spielman, R.S. and Cheung, V.G. (2004) Genetic analysis of genome-wide variation in human gene expression. *Nature*, **430**, 743–7. Available at: http://www.nature.com/doifinder/10.1038/nature02797.

Mounet, F., Moing, A., Garcia, V., et al. (2009) Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development. *Plant Physiol.*, **149**, 1505–28. Available at: http://www.plantphysiol.org/cgi/doi/10.1104/pp.108.133967.

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 Physiol.*, **156**, 2244–54. Available at: http://www.plantphysiol.org/cgi/doi/10.1104/pp.111.173997.

Nica, A.C. and Dermitzakis, E.T. (2013) Expression quantitative trait loci: present and future. *Philos. Trans. R. Soc. B Biol. Sci.*, **368**, 20120362–20120362. Available at: http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2012.0362.

Paschold, A., Jia, Y., Marcon, C., et al. (2012) Complementation contributes to transcriptome complexity in maize (Zea mays L.) hybrids relative to their inbred parents. *Genome Res.*, **22**, 2445–54. Available at: http://genome.cshlp.org/cgi/doi/10.1101/gr.138461.112.

Pascual, L., Xu, J., Biais, B., et al. (2013) Deciphering genetic diversity and inheritance of tomato fruit weight and composition through a systems biology approach. *J. Exp. Bot.*, **64**, 5737–52. Available at: https://academic.oup.com/jxb/article-lookup/doi/10.1093/jxb/ert349.

Pattison, R.J., Csukasi, F., Zheng, Y., Fei, Z., Knaap, E. van der and Catalá, C. (2015) Comprehensive tissue-specific transcriptome analysis reveals distinct regulatory

programs during early tomato fruit development. *Plant Physiol.*, **168**, 1684–701. Available at: http://www.plantphysiol.org/lookup/doi/10.1104/pp.15.00287.

R Development Core Team (2012) R: a language and environment for statistical computing. Available at: http://softlibre.unizar.es/manuales/aplicaciones/r/fullrefman.pdf.

Rabbani, M.A., Maruyama, K., Abe, H., et al. (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiol.*, **133**, 1755–67. Available at: http://www.plantphysiol.org/cgi/doi/10.1104/pp.103.025742.

Ranjan, A., Budke, J.M., Rowland, S.D., et al. (2016) eQTL regulating transcript levels associated with diverse biological processes in tomato. *Plant Physiol.*, **172**, 328–40. Available at: http://www.plantphysiol.org/lookup/doi/10.1104/pp.16.00289.

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. Available at: http://journal.frontiersin.org/Article/10.3389/fpls.2015.01172/abstract.

Rockman, M. V and Kruglyak, L. (2006) Genetics of global gene expression. *Nat. Rev. Genet.*, **7**, 862–72. Available at: http://www.nature.com/doifinder/10.1038/nrg1964.

Saliba-Colombani, V., Causse, M., Gervais, L. and 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. Available at: http://www.nrc.ca/cgi-bin/cisti/journals/rp/rp2_abst_e?gen_g99-096_43_ns_nf_gen43-00.

Sánchez-Rodríguez, E., Moreno, D. a, Ferreres, F., Rubio-Wilhelmi, M.D.M. and 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. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0031942211000975.

Schadt, E.E., Monks, S.A., Drake, T.A., et al. (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature*, **422**, 297–302. Available at: http://www.nature.com/doifinder/10.1038/nature01434.

Scott, J.W. (2005) Perspectives on tomato disease resistance breeding: past, present and future. *Acta Hortic.*, 217–224. Available at: https://www.actahort.org/books/695/695_24.htm.

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**, 279–92. Available at: http://doi.wiley.com/10.1046/j.1365-313X.2002.01359.x.

Signor, S.A. and Nuzhdin, S. V. (2018) The evolution of gene expression in cis and trans. *Trends Genet.*, **34**, 532–544. Available at: https://doi.org/10.1016/j.tig.2018.03.007.

Slane, D., Kong, J., Berendzen, K.W., et al. (2014) Cell type-specific transcriptome analysis in the early Arabidopsis thaliana embryo. *Development*, **141**, 4831–40. Available at:

http://dev.biologists.org/cgi/doi/10.1242/dev.116459.

Snoek, L.B., Terpstra, I.R., Dekter, R., Ackerveken, G. Van den and Peeters, A.J.M. (2012) Genetical genomics reveals large scale genotype-by-environment interactions in Arabidopsis thaliana. *Front. Genet.*, **3**, 317. Available at: http://journal.frontiersin.org/article/10.3389/fgene.2012.00317/abstract.

Springer, N.M. and Stupar, R.M. (2007a) Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. *Plant Cell*, **19**, 2391–402. Available at: http://www.plantcell.org/cgi/doi/10.1105/tpc.107.052258.

Springer, N.M. and Stupar, R.M. (2007b) Allelic variation and heterosis in maize: how do two halves make more than a whole? *Genome Res.*, **17**, 264–75. Available at: http://www.genome.org/cgi/doi/10.1101/gr.5347007.

Steinhauser, M.-C., Steinhauser, D., Gibon, Y., Bolger, M., Arrivault, S., Usadel, B., Zamir, D., Fernie, A.R. and Stitt, M. (2011) Identification of enzyme activity quantitative trait loci in a Solanum lycopersicum x Solanum pennellii introgression line population. *Plant Physiol.*, **157**, 998–1014. Available at: http://www.plantphysiol.org/cgi/doi/10.1104/pp.111.181594.

Stevens, R., Buret, M., Garchery, C., Carretero, Y. and Causse, M. (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. Available at: http://pubs.acs.org/doi/abs/10.1021/jf061241e.

Stupar, R.M. and 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**, 2199–210. Available at: http://www.genetics.org/cgi/doi/10.1534/genetics.106.060699.

Swanson-Wagner, R.A., Jia, Y., DeCook, R., Borsuk, L.A., Nettleton, D. and Schnable, P.S. (2006) All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 6805–10. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0510430103.

Swinnen, G., Goossens, A. and Pauwels, L. (2016) Lessons from domestication: targeting cis-regulatory elements for crop improvement. *Trends Plant Sci.*, **21**, 506–515. Available at: http://dx.doi.org/10.1016/j.tplants.2016.01.014.

Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, **485**, 635–41. Available at: http://www.nature.com/articles/nature11119 [Accessed July 10, 2014].

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. (2012) Primer3--new capabilities and interfaces. *Nucleic Acids Res.*, **40**, e115. Available at: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gks596.

Verta, J., Landry, C.R. and 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. *New Phytol.*,

211, 159–71. Available at: http://doi.wiley.com/10.1111/nph.13888.

Wang, S., Yehya, N., Schadt, E.E., Wang, H., Drake, T.A. and Lusis, A.J. (2006) Genetic and genomic analysis of a fat mass trait with complex inheritance reveals marked sex specificity. *PLoS Genet.*, **2**, e15. Available at: http://dx.plos.org/10.1371/journal.pgen.0020015.

Wang, X., Chen, Q., Wu, Y., et al. (2018) Genome-wide analysis of transcriptional variability in a large maize-teosinte population. *Mol. Plant*, **11**, 443–459. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1674205217303817.

Waters, A.J., Makarevitch, I., Noshay, J., Burghardt, L.T., Hirsch, C.N., Hirsch, C.D. and Springer, N.M. (2017) Natural variation for gene expression responses to abiotic stress in maize. *Plant J.*, **89**, 706–717. Available at: http://doi.wiley.com/10.1111/tpj.13414.

West, M.A.L., Kim, K., Kliebenstein, D.J., Leeuwen, H. van, Michelmore, R.W., Doerge, R.W. and St Clair, D.A. (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. *Genetics*, **175**, 1441–50. Available at: http://www.genetics.org/cgi/doi/10.1534/genetics.106.064972.

Wittkopp, P.J. and Kalay, G. (2011) Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.*, **13**, 59–69. Available at: http://dx.doi.org/10.1038/nrg3095.

Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M. V and Romano, L.A. (2003) The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.*, **20**, 1377–419. Available at: https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msg140.

Wu, B.-H., Génard, M., Lescourret, F., Gomez, L. and Li, S.-H. (2002) Influence of assimilate and water supply on seasonal variation of acids in peach (cv Suncrest). *J. Sci. Food Agric.*, **82**, 1829–1836. Available at: http://doi.wiley.com/10.1002/jsfa.1267.

Xu, X., Hu, Q., Yang, W. and Jin, Y. (2017) The roles of call wall invertase inhibitor in regulating chilling tolerance in tomato. *BMC Plant Biol.*, **17**, 195. Available at: https://bmcplantbiol.biomedcentral.com/articles/10.1186/s12870-017-1145-9.

Young, M.D., Wakefield, M.J., Smyth, G.K. and Oshlack, A. (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.*, **11**, R14. Available at: http://genomebiology.biomedcentral.com/articles/10.1186/gb-2010-11-2-r14.

Yvert, G., Brem, R.B., Whittle, J., Akey, J.M., Foss, E., Smith, E.N., Mackelprang, R. and Kruglyak, L. (2003) Trans-acting regulatory variation in Saccharomyces cerevisiae and the role of transcription factors. *Nat. Genet.*, **35**, 57–64. Available at: http://www.nature.com/doifinder/10.1038/ng1222.

Zan, Y., Shen, X., Forsberg, S.K.G. and Carlborg, Ö. (2016) Genetic regulation of transcriptional variation in natural Arabidopsis thaliana accessions. *G3 Genes|Genomes|Genetics*, **6**, 2319–28. Available at: http://g3journal.org/lookup/doi/10.1534/g3.116.030874.

Zhang, S., Xu, M., Qiu, Z., Wang, K., Du, Y., Gu, L. and Cui, X. (2016) Spatiotemporal transcriptome provides insights into early fruit development of tomato (Solanum lycopersicum). *Sci. Rep.*, **6**, 23173. Available at: http://www.nature.com/articles/srep23173.

Zhu, G., Wang, S., Huang, Z., et al. (2018) Rewiring of the fruit metabolome in tomato breeding. *Cell*, **172**, 249–261.e12. Available at: https://linkinghub.elsevier.com/retrieve/pii/S009286741731499X.

Legend of Figures

Figure 1. Venn diagram showing the differentially expressed genes between drought and control conditions in leaf (A) and fruit (B) for each genotype and comparisons between genotypes. Bold and highlighted numbers indicate the total number of genes differentially expressed in each comparison among the 18,843 and 19,097 genes analyzed in leaf and fruit, respectively. '↗' indicates the up-regulated genes in stress compared to control. '↘' indicates the down-regulated genes in stress compared to control.

Figure 2. Cumulative numbers of genes significantly affected by *Cis*, *Trans* or both types of regulation patterns in fruit and leaf based on allele specific expression in F1 hybrid and parental RNAseq libraries. (A) Constitutive effect; (B) interaction with watering regime effect.

Figure 3. Distribution of the *Cis*, *Trans* and *Cis+Trans* constitute regulatory divergences detected using allele specific expression in F1 hybrid and parental RNAseq libraries in leaf (A) and fruit (B). Average log2FC of Levovil vs Cervil allele in parental genotypes (x-axis) is plotted against the average log2FC of Levovil vs Cervil allele in the F1 hybrid (y-axis). A pseudo counts of 1 was added to allelic expression counts to allow log scale transformation. Log2FC were averaged across biological replicates and watering regimes. Green : *Trans* effect only. Blue : *Cis* effect only. Salmon : *Cis* and *Trans* effects.

Figure 4. Distribution of (e)QTLs across genome. On the left panel, locations of eQTLs are plotted against the location of the regulated genes. Circle : *Local* eQTL. Triangle : *Trans* eQTL. Open symbol : Gene expression measured in leaf. Close symbol : Gene expression measured in fruit. On the right panel, QTLs for plant vigor and fruit quality traits are represented according to their position on the genome. QTLs detected for Fructose, Glucose, DMW and SSC traits are gathered under 'Sugar'. QTLs detected for pH and Malic traits are gathered under 'Acid'. In both panels, size of symbol is proportional to LOD score. For representation ease, LOD scores for constitutive (e)QTLs were computed as the average of LOD scores under control and drought conditions. Orange : Constitutive
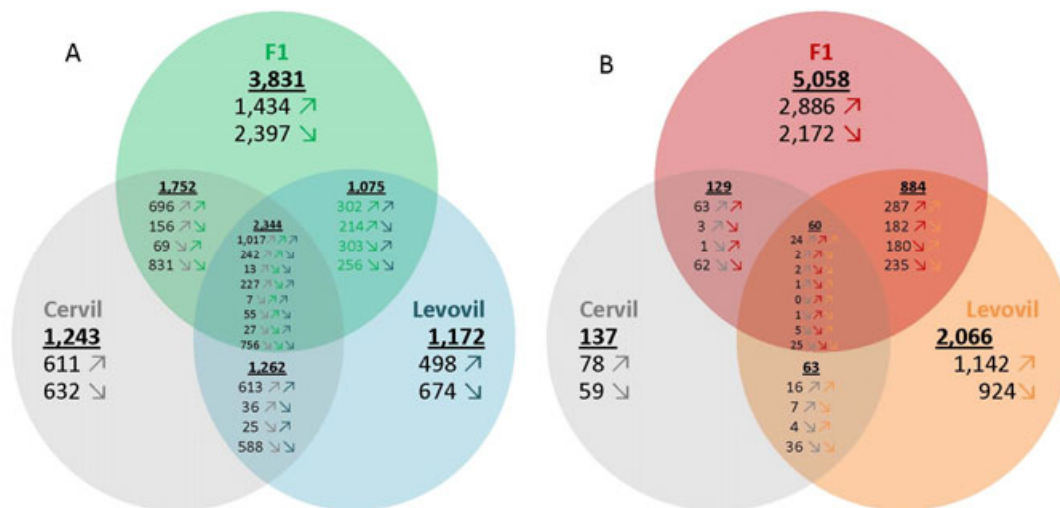
(e)QTLs. Purple : interactive (e)QTLs. Grey dashed lines represent limits between chromosomes. Positions are expressed in Mbp cumulated across chromosomes CH1 to CH12.
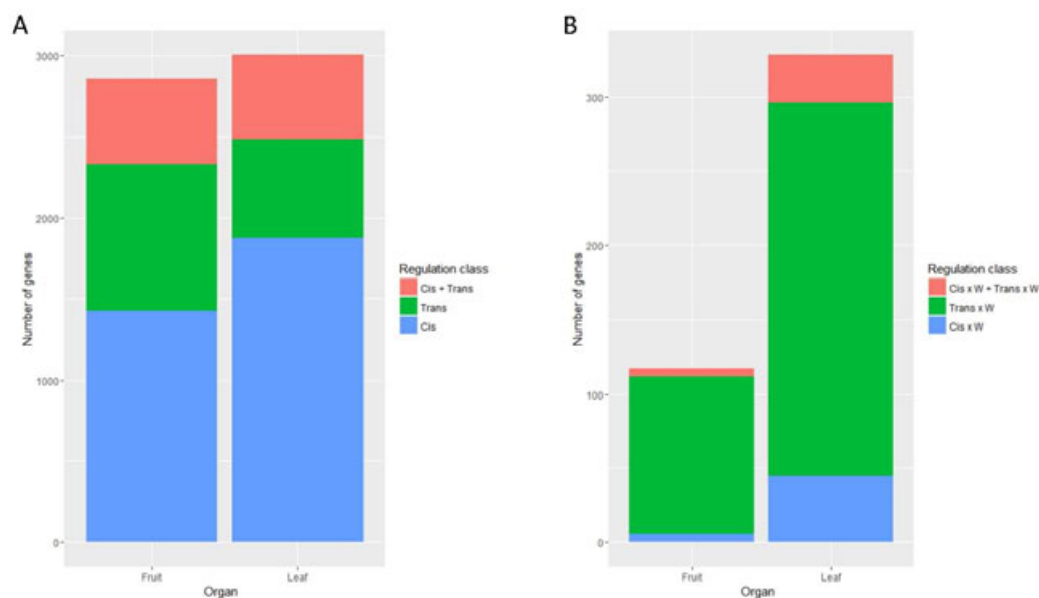
Figure 5. Bar graph partitioning eQTLs according to LOD score (low : LOD < 4, medium : 4 ≤ LOD < 8, high : LOD > 8), regulation class (*Trans* or *Local*, see Material & Methods) and interaction with watering regime (Constitutive or Interactive, see Material & Methods). For representation ease, LOD scores for constitutive eQTLs were computed as the average of LOD scores under control and drought conditions.

Figure 6. Overview of (e)QTLs related to sugar metabolism in tomato fruit identified by linkage analysis in the RILs. On the first layer of the circle, the twelve tomato chromosomes are represented proportionally to their physical size (assembly 2.5). QTLs detected for sugar related traits in tomato fruit (DMW, SSC, Glucose & Fructose) are represented with lines on the second layer of the circle. *Trans* and *local* eQTLs identified for the expression of 8 genes related to sugar metabolism measured through microfluidgm qPCR on fruit samples of 124 RILs are represented in the inner part of the circle using links and large dots, respectively. Orientation of the *trans* eQTLs is shown with small dots to indicate position of the regulated genes and an arrow to indicate the localization of the eQTL. Orange : Constitutive (e)QTLs. Purple : interactive (e)QTLs. Light purple indicates *trans* & *local* interactive eQTLS that were significant only at a lowered threshold (alpha = 0.10). InvInh (Invertase Inhibitor) : Solyc01g088590; nInv1 (Neutral Invertase) : Solyc01g111100; Frk3 (Fructokinase 3) : Solyc02g091490; acInv (Vacuolar Acid Invertase) : Solyc03g083910; FBPase (Fructose 1-6 bisphosphatase) : Solyc04g071340; Lin9 (Cell Wall Acid invertase) : Solyc08g079080; Lin5 (Cell Wall Acid Invertase): Solyc09g010080; nInv11 (Neutral Invertase): Solyc11g020610. Susy (Sucrose synthase) : Solyc12g009300.

A

F1
3,831
1,434 ↗
2,397 ↘

1,752
696 ↗↗
156 ↗↘
69 ↘↗
831 ↘↘

1,075
302 ↗↗
214 ↗↘
303 ↘↗
256 ↘↘

2,344
1,017 ↗↗↗
242 ↗↗↘
13 ↗↘↘
227 ↗↘↗
7 ↘↗↗
55 ↘↗↘
27 ↘↘↗
756 ↘↘↘

Cervil
1,243
611 ↗
632 ↘

Levovil
1,172
498 ↗
674 ↘

1,262
613 ↗↗
36 ↗↘
25 ↘↗
588 ↘↘

B

F1
5,058
2,886 ↗
2,172 ↘

129
63 ↗↗
3 ↗↘
1 ↘↗
62 ↘↘

884
287 ↗
182 ↗
180 ↘
235 ↘

60
24 ↗↗↗
2 ↗↗↘
2 ↗↘↘
1 ↗↘↗
0 ↘↗↗
1 ↘↗↘
5 ↘↘↗
25 ↘↘↘

Cervil
137
78 ↗
59 ↘

Levovil
2,066
1,142 ↗
924 ↘

63
16 ↗↗
7 ↗↘
4 ↘↗
36 ↘↘

Figure 5. Bar graph partitioning eQTLs according to LOD score (low : LOD < 4, medium : 4 ≤ LOD < 8, high : LOD > 8), regulation class (*Trans* or *Local*, see Material & Methods) and interaction with watering regime (Constitutive or Interactive, see Material & Methods). For representation ease, LOD scores for constitutive eQTLs were computed as the average of LOD scores under control and drought conditions.