Genetic diversity in tomato (Solanum lycopersicum) and its wild relatives

Guillaume Bauchet, Mathilde M. Causse

8

Genetic Diversity in Tomato (*Solanum lycopersicum*) and Its Wild Relatives

Guillaume Bauchet and Mathilde Causse
Institut National de la Recherche Agronomique (INRA),
Unité de Génétique et Amélioration des Fruits et Légumes (GAFL),
France

1. Introduction

Tomato, ranking 1st in the world for vegetables, accounts for 14% of world vegetable production (over 100 million metric tons/year $ 1.6 billion market; (Food and Agriculture Organisation [FAO] 2010). Tomato is a rich source of micronutrients for human diet. It is also an acknowledged model species for research on fruit development and metabolite accumulation. The major goals of tomato breeders (higher productivity, better tolerance to biotic and abiotic stresses and increased sensory and health value of the fruit) require a good comprehension and management of tomato genetic resources diversity.

Due to its Latin American origin and related domestication history, cultivated tomato has faced several bottlenecks over ages. This led to a drastic reduction of its genetic diversity. Explorations of tomato center of origin permitted major advances in the characterization of its diversity. In parallel, *ex situ* plant conservation initiatives bloomed, ensuring the collection and conservation of landraces and wild species through development of seed banks. Thus, unraveling the genetic potential of tomato's wild relatives for breeding purpose emerged. In parallel, the ecological and taxonomic diversity of tomato turned it into a model species for evolutionary studies. Since the mid-20th century, new methods such as controlled hybridization allowed crossing between wild and cultivated tomato. Modern genetics and breeding methods contributed to understand the genetic control of agronomical traits but also accentuated the progress. If successful, the accuracy to introgress agronomic traits of interest from wild relatives into cultivated tomato was not always straightforward. This was notably due to inherent linkage between "favorable" and "unfavorable" effects of introgressed fragments.

The advent of molecular biology in the 80's raised great hopes in terms of characterization of the genetic diversity present in both wild and cultivated compartments. Also, great expectations emerged since the development of molecular techniques to "pinpoint" genomic regions involved in targeted traits. Dissection of the genetic control of complex traits, using ad hoc techniques from quantitative genetics, was possible, leading to the identification of key alleles involved in diverse agronomic traits, originating from several wild relatives.

Today the tomato genome is fully sequenced. A new step in the knowledge on tomato diversity with the so called "-omics" and next generation sequencing techniques is coming.

These technologies and related data analysis allow a complete and combined reading of genomes and related levels of expression (transcriptome, proteome, metabolome) in a high throughput way. Among the new approaches, QTL mapping techniques in natural populations or genome wide association studies will facilitate the genetic characterization of complex traits and germplasm management of both wild and cultivated tomatoes.

In this chapter we will first show how tomato diversity evolved from its early domestication until today. We will discuss how valuable tomato genetic resources are, and that investigating natural variation not only highlights existing diversity -which is of critical use for cultivated tomato improvement- but can also provide insights into the evolution and genetic bases of complex traits. In the last part, we will present how molecular markers have completed our view.

2. Diversity of the tomato clade species

Tomato belongs to the large and diverse *Solanaceae* family also called Nightshades which includes more than three thousand species. Among them, major crops arose from Old world (Eggplant from Asia) and New world (pepper, potato, tobacco, tomato from South America). The *Lycopersicon* clade contains the domesticated tomato (*Solanum lycopersicum*) and its 12 closest wild relatives (Peralta and Spooner 2005). The radiation of tomato clade has been estimated to 7.8 (Nesbitt and Tanksley 2002) and to 2.7 Million years between *S. lycopersicum* and *S. pennellii*(Kamenetzky et al. 2010). First detailed studies on this group of wild relatives were made by Charles Rick and colleagues since the 40's. Tomato clade species are originated from the Andean region, including Peru, Bolivia, Ecuador, Colombia and Chile (Figure 1). On Figure 1, *lycopersicon* species distributions are defined according to geographic data from the Tomato Genetics Resource Center, UC Davis http://tgrc. ucdavis.edu/Data/Acc/dataframe.aspx? start=GIS_ dataoption. aspx&navstart=nav.html. Their growing environments range from near sea level to 3,300 m altitude, from arid to rainy climate and from Andeans Highlands to the coast of Galapagos Islands (*S. cheesmaniae; S. galapagense*). Their habitats are often narrow and isolated valleys where they were adapted to particular microclimates and various soil types. Their very large range of ecological conditions contributed to the diversity of the wild species. This broad variation is also expressed at the morphological, physiological, sexual and molecular levels (Peralta and Spooner 2005). Over times, several phylogenetic classifications have been proposed and several adjustments occurred. Being first classified in the *Solanum* genus, the group turned to a specific genus, *Lycopersicum* (Miller, 1731). It recently got renamed *Solanum* within an updated classification (Peralta and Spooner 2001). Taxonomic, ecological, reproductive, breeding specificities for each member of the clade are listed in Table 1 and reviewed by Peralta and colleagues (Peralta, Spooner et al. 2007). The first classification was morphology based (Luckwill 1943). Later molecular data confirmed tomato membership of Linnaeus classification, but also improved subtaxa classification (Spooner 2008). The tomato clade is an interesting example for research on plant biodiversity, notably, on evolution, adaptation, human domestication and nutrition perspectives (Peralta and Spooner 2007). Nowadays, across South America, populations of wild tomatoes are being severely reduced. Their natural habitats are shrinking due to urban development and intensive agriculture as well as goat herding in the highlands, as recently documented by a botanical expedition in Peru. (Grandillo, Chetelat et al. 2011).

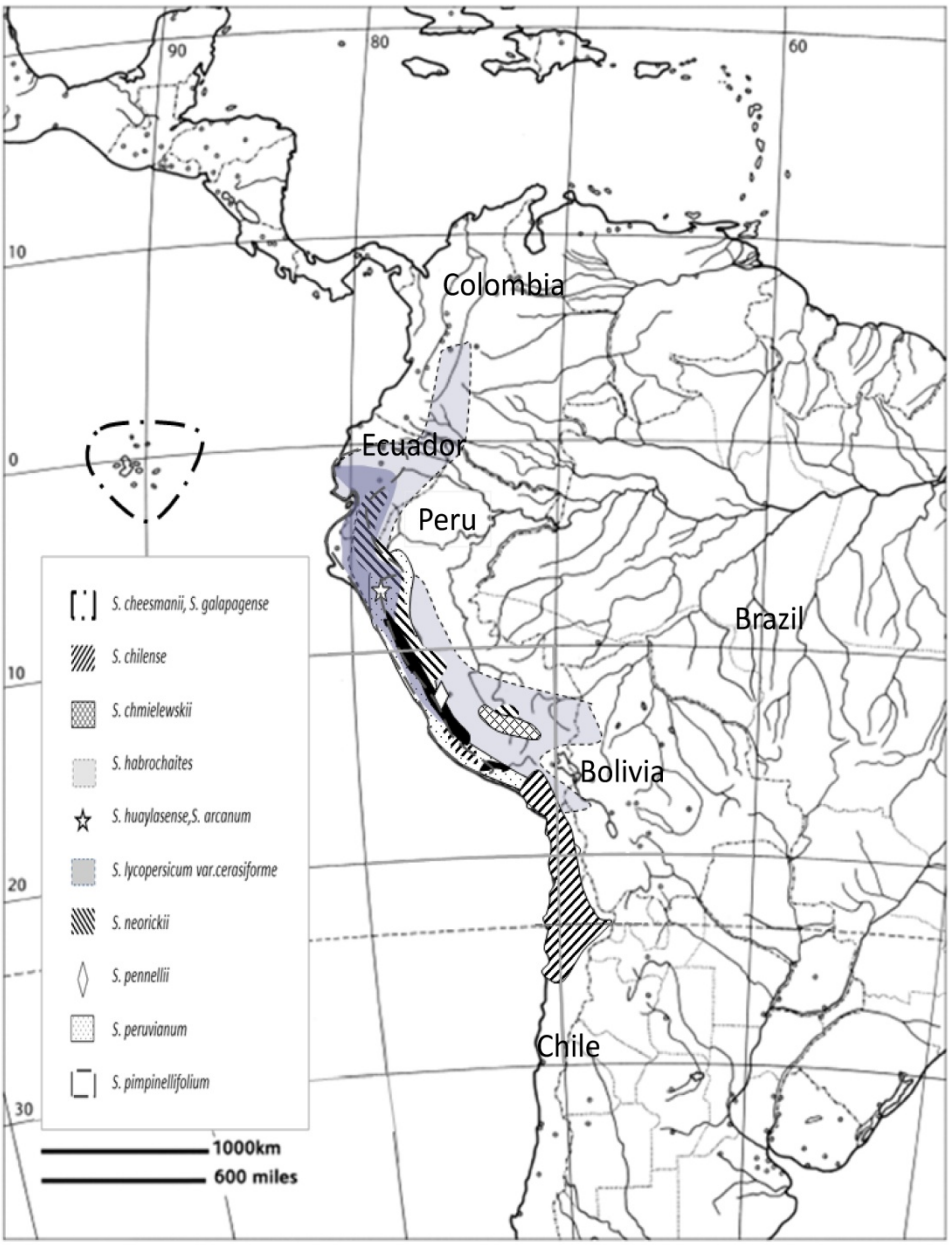

Fig. 1. Geographic distribution of wild species in *Solanum* section *lycopersicon*.

Many studies were conducted on evolutionary aspects of the lycopersicon clade. The mating system was extensively studied, using the clade as a model to study its effects on species variation (Bedinger, Chetelat et al. 2011). Mating system has played a key role in evolution of wild tomatoes, varying from allogamous self-incompatible, to facultative allogamous, to autogamous and self-compatible (Table 1). Flower stigma exertion and gametophytic incompatibility system contribute in greater outcrossing and genetic diversity. All the species of the clade are intercrossable (Table 1), but with a variable success rate (Rick, Fobes et al. 1977a; Rick, Fobes et al. 1979). Fruit color discriminate the wild relative species. Most of the latter carry green fruits, with the exception of the two species from the Galapagos (with yellow and orange fruits) and *S. pimpinellifolium*, which is the only wild relative species with red fruits. *S. pimpinellifolium* fruits are round, small, weighing only few grams. These fruits are edible and the species referred as the currant tomato. The plant presents a reduced apical dominance and prostrate growth habit resulting in a large shrub with inflorescence carrying many flowers and fruits (Paran and van der Knaap 2007). *S. pimpinellifolium* undergone bottleneck only recently with a drastic reduction of its natural habitats and is now an endangered species (Biodiversity-International 2006). *S. lycopersicum* var *cerasiforme* fruit is larger than *S. pimpinellifolium* and is commonly round and red. This subspecies of tomato is referred to as the "cherry tomato". It has been proposed as the direct ancestor of cultivated tomato because of its diversity, its wide spread occurrence in central America and its close genetic relationship with cultivated tomato (Rick and Chetelat 1995). The modern cultivated tomato, *S. lycopersicum*, is cosmopolite. It has spread all around the world and is now cultivated under a broad range of environments and conditions.

3. Tomato domestication in South America

Domestication is a special type of species diversification, distinct from species divergence through natural selection in the wild (Darwin and Wallace 1858). Domesticated species differ from wild and relative species for a set of traits known as the domestication syndrome (Doebley, Gaut et al. 2006). Domestication is often controlled by a limited number of chromosomal regions with major phenotypic effect (Purugganan and Fuller 2009). In tomato, edible fruits, attractive red color and fruit size increase are characterizing this process.

The domestication time of tomato is unclear. It is supposed to be due to a recent divergence from S. *pimpinellifolium.* The first hypothesis supports Peru as the center of origin and domestication (de Candolle 1882). This hypothesis gives emphasis on botanical evidences and has been complemented by botanical, linguistic and historical aspects. It was further supported by other colleagues (Müller 1940a; Müller 1940b; Luckwill 1943) and recent molecular studies (Nesbitt and Tanksley 2002). Nevertheless, very little and unclear archeological evidences are available to clearly support this hypothesis (McMeekin 1992). The second hypothesis supports that domestication occurred primarily in Mexico in the Vera Cruz Puebla area (Jenkins 1948), as there is no evidence for pre-Colombian cultivation of tomato in South America but good evidences in Mexico. Referring to Guilandini (1572), Jenkins also argued that tomato name comes most probably from the Mexican Nahua people word "Tomatl" that described "plants bearing globous and juicy fruit" (Sahagún 1988). Based on its

Subsection	Species New Nomenclature [previous nomenclature]	Geographic distribution and habitat	Mating system / Cross compatibility *S.lycopersicum*	Fruit color	Genetic polymorphism
Arcanum	*S. arcanum* [*L. peruvianum*]	Northern Peru, 100-2500m, Coastal and inland Andean valleys, in lomas, dry valleys, and on dry rocky slopes	SI[a], facultative AL[2] / UI[1], EL[2]	green with dark green stripes	Intermediate
	S. chmielewskii [*L. chmielewskii*]	South Peru to North Bolivia native, 1500-3000 m, dry and drained areas	SC[b], facultative AL / reciprocal	green with dark green stripes	Intermediate
	S. neorickii [*L. parviflorum*]	South Equador-south center Peru native, 1500-3000m, rocky, humid and well drained areas	SC, highly AT[d] / reciprocal	green with dark green stripes	Low
Neolycopersicon	*S. pennellii* [*L. pennellii*]	Peruvian cost native, 0-2000 m, dry and rocky hillsides.	SI usually, SC populations in southern parts /reciprocal	green	High
Eriopersicon	*S. habrochaites* [*L. hirsutum*]	South west Equador to south center Peru native , 500-3300, forest regions	SI, SC populations in southern parts / UI	green with darker green stripes	High
	S. chilense [*L.chilense*]	South Peru to North Chili, 0-3000 m, dry river bed	SI, AL / UI, EL	green to whitish	High
	S. huaylasense [*partly L. peruvianum*]	Peru, 1700-3000m, rocky slopes around Callejón de Huaylas	SI, AL / UI, EL	green with purple green with dark green stripes	High
	S. peruvianum [*L. peruvianum*]	Central Peru to northern Chile,0-600 m, lomas formations and occasionally in coastal deserts	SI, AL / UI, EL	green to greenish white,sometimes	High
	S. corneliomulleri [*partly L. peruvianum, known as L. glandulosum*]	Southern Peru,1000-3000 m, Middle to high elevations on the western slope of the Andes, lower slopes on the edges of landslides	SI, AL / UI, EL	green with dark green stripes	High
Lycopersicon	*S. chessmaniae* [*L. chessmaniae*]	Galápagos islands endemic species,0-1300 m, From sea shore to volcanic area .	SC, AT / reciprocal	yellow, orange	Low
	S. galapagense [*partly L. chessmaniae*]	Galápagos islands endemic species, sea shore.	SC, AT / reciprocal	yellow, orange	Low
	S. pimpinellifolium [*L. pimpinellifolium*]	South Equador-North Peru native, under 1000 m, south valleys of the pacific coast.	SC, AT, facultative AL / reciprocal	red	Intermediate
	S. lycopersicum var. *cerasiforme** [*L. esculentum* var. *cerasiforme*]	Adventive worldwide in tropics and subtropics, probably native from Andeans region	SC, AT, facultative AL / reciprocal	red	Low
	S. lycopersicum [*L. esculentum*]	Probably Equador-Peru, nowadays widely spread, various range of habitats	SC, AT	red	Very Low

(a) Self-incompatible (b) Self compatible (c) Alloganous (d) Autoganous * cults group
(1) Unilateral incompatibility (2) Embryo lethality (embryo-rescue technique required)

Table 1. Principal features of the *lycopersicon* subsection (*Solanum* sect. *Lycopersicon*) Data are compiled from Peralta *et al*. 2007, Moyle *et al*. 2008, Grandillo *et al*. 2011

center theory, Harlan suggested that biloculed domesticated forms found in south Mexico and Guatemala are the oldest cultivated types (Harlan 1971). Quoting Sahagun, Diez argued that tomato was totally "integrated" in the Aztec civilization food consumption in XVI century, contrary to South American Incas (Diez and Nuez 2008). Nevertheless, two authors identified Quechua names possibly referring to tomato: "pirca" (Horkheimer 1973) and "pesco-tomate" (Yakovleff 1935). However, botanists consider the origins of tomato domestication as unsolved (Peralta and Spooner 2007b). These authors mention recent evidences showing that the Mexican hypothesis is not supported by comparative data, as South American and Mexican tomato accessions share similar isozymes (Rick and Fobes 1975) as well as molecular markers (Villand, Skroch et al. 1998). So far, no evidence appears to be enough conclusive and tomatoes may have been domesticated independently in both areas. To go further a more extensive analysis of molecular polymorphism in the wild and cultivated tomatoes is needed. This would allow investigating demographic scenarios and estimating the parameters of these scenarios (bottleneck intensity, ancestral population size, migration rates) using Markovian model implemented in tools such as IM* program (Hey and Nielsen 2004) or ABC[1] methodology (Beaumont, Zhang et al. 2002; Lopes and Beaumont 2010). Very recently, this approach has been implemented to infer past demography and ecological parameters of two tomato wild relatives, *S. chilense* and *S. peruvianum* (Tellier, Laurent et al. 2011).

Many authors consider *S. lycopersicum* var. *cerasiforme* as ancestral form of the cultivated tomato. It is present in both Mexico and Peru, on the contrary to *S. pimpinellifolium* which is absent from Mexico. If we assume that *S.l. cerasiforme* results from direct domestication from *S. pimpinellifolium*, a consequence of this domestication is that *S.l. cerasiforme* suffered a decrease of its population effective size during domestication (Bai and Lindhout 2007). Subsequent changes occurred for domestication traits such as growth habit, mating system, gigantism and fruit morphological diversity. Notably a change from exerted to inserted stigmas is responsible for the change from partial allogamy to strict autogamy. Selection for self-pollinating as well as shortening of the stigma compared to close wild relatives such as *S. pimpinellifolium* has allowed a yield increase (Rick 1977b). This "selfing syndrome" (Sicard and Lenhard 2011) is striking in tomato where a mutation in gene controlling stigma length has been identified in cultivated germplasm (Chen, Cong et al. 2007).

4. Early cultivation in Europe and in the world

Probably only a few tomato seeds were brought back from Mexico to Europe, leading, after domestication, to a new genetic bottleneck. George McCue has extensively reviewed the history of tomato diverse uses, tracking back the first references by country upon the time (McCue 1952). Most remote reference available comes from Petrus Matthiolus, an Italian Physician (1544). Due to its botanical closeness with toxic *Solanum* species common in Europe (Mandrake, Belladonna), tomato was for long mostly used as an ornamental. Two centuries later, it was referred as a cultivated plant in Italy by Saccardo (1769). Southern Europe was precursor in use of tomato for human consumption. In France, Bois at first mentions it as ornamental (1760). The same author reported it as vegetable seeds sold in the catalogue of the seeds of the "Maison grainière Andrieux Vilmorin" (1778). Lamarck mentioned it in 1798. Extensive consumption in Spain is described by Quer (1784).

[1]Approximate Bayesian Computation

Progressively, following South-North gradient, tomato consumption reached Northern Europe (Sabine 1819). Similarly in USA, Bartram (1766) reported tomatoes being used as food plants. Boyd (1784) mentioned that The David Landreth Seed Co. started to sell tomato seeds for vegetable consumption. Selection for diverse fruit shapes and local adaptation probably rapidly occurred through bulk selection. The crop gained in economic importance by the end of XIXth century with the establishment of tomato breeding programs. Most of the plant material at that time can be considered as landraces: selected for subsistence agriculture environments, producing low but relatively stable yield. At the end of the XIXth century, tomato cultivars (nowadays called landraces or heirlooms) were open pollinated from which seeds were saved by the farmers from a year to the other. Selection of new genotypes within heterogeneous cultivars (or selection of chance variance) resulted from spontaneous mutations, natural outcrossing or recombination of pre-existing genetic variation. Thus, *S. lycopersicum* found in Europe a secondary centre for diversification (García-Martínez, Andreani et al. 2006). In the XIXth century, establishment of commercial routes and colonies contributed to spread the species worldwide (Diez and Nuez 2008). In United States, prior to 1850 and "Trophy" the first commercially successful variety, no breeding programs were effective (Smith 2000). On an evolutionary perspective, domestication and implementation of breeding programs induced physiological changes. Artificial selection has reduced the genetic diversity of the crop which suffered a new bottleneck.

5. Tomato breeding in the XXth century: Seeking for diversity and intensive production

After domestication and adaptation to North hemisphere growing conditions for two centuries, the crop started the XXth century with benefits of two major scientific discoveries: The rediscovery of Mendel pioneering work to set up the basis of experimental methods on the use of plant hybridization (Mendel 1866). Second are established domestication concept (Darwin and Wallace 1858) and selection theory (Darwin 1859).

This context has seen the emergence in public institutes of plant germplasm banks, starting point for collecting existing genetic diversity, preserving and valorizing it, following the pioneer work of Nikolai Vavilov (1887-1943) (Kurlovich, Rep'ev et al. 2000). Later on, he was followed by Charles Rick (1915-2002) who dedicated his life to discover, collect and characterize exotic tomato germplasm (Tanksley and Khush 2002). Today, more than 83,000 tomato accessions are stored in seed banks worldwide, ranking 1st among vegetable species collected (FAO 2010). The main collections in the world are: In USA, the Tomato Genetic Resources Center in California (TGRC), (www.tgrc.ucdavis.edu) and the USDA[2] collection (www.ars.usda.gov), the World Vegetable Center in Taiwan (www.avrdc.org) and several Europeans collections. The establishment of tomato resource collections made major contributions to understand the distribution of its diversity around the world. Nevertheless the lack of coordination and conflicting passport data is a pitfall for an efficient tomato germplasm management. Efforts are now made to coordinate national initiatives in global or regional approaches. Since 2007, The European Cooperative Programme for Plant Genetic Resources (www.ecpgr.cgiar.org) is a collaborative project between most European countries for long-term conservation and utilization of plant genetic resources in Europe.

[2] United States Department of Agriculture, Geneva

This project is based on large network of national centers for tomato genetic resources including COMAV[3] (Spain), CGN[4] (Netherlands), INRA (France), IPK[5] (Germany), Vavilov Institute (Russia) and others. These institutions share their germplasm informations through a database (http://documents.plant.wur.nl/cgn/pgr/tomato/). More recently, in the context of a European Solanaceae project (EU-SOL, www.eu-sol.wur.nl), a collection of more than 6,000 domesticated tomato accessions was established and phenotyped, accompanied by an *ad hoc* database (Finkers, de Weerd et al. 2011). Finally, since 2008, a world initiative, is conducted under the Plant Biodiversity Inventories (www.nhm.ac.uk/research-curation/projects/solanaceaesource/). Project aim is to produce a worldwide taxonomic monograph of the species occurring within the plant genus *Solanum.* As well, tomato is part of long term collection of plant species project, launched by the Svalbard Global Seed Vault initiative (Food 2008).

Due to its broad use for food consumption and adaptation to many environmental conditions (from Alaska summers to tropical conditions) and different crop systems, tomato experienced an important phenotypic diversification. Hundreds of past and present cultivated varieties are now available. Cultivars are dedicated to two main markets, processing and fresh market. Processing tomatoes are cultivated as a field crop, whereas fresh market tomatoes are grown outdoor or indoor (heated and non-heated greenhouses). Breeding objectives have evolved over time, with the evolution of production systems. Nevertheless, three main objectives remain: adaptation to growth constraints, disease and pest resistances and fruit productivity and quality. Wild species were first used as source of adaptation to biotic stress. Disease resistance selection started in United States early XXth century. The first *fusarium* wilt –resistant cultivar "Tennessee red" was released in 1912. Early 1920's breeders used hybridization with selection in segregating generations. By the mid-30's, plant breeders developed technical procedures to improve selection, such as pedigree selection. Later, existing or emerging private companies enhanced their development with the release of F1 hybrid varieties. Selection for disease resistance was successful as dominant resistance genes were found in the wild relatives for most of the diseases and pests. Modern cultivars can cumulate up to 12 different disease resistance genes which all derive from wild species. Wild germplasm has been primarily used as a source of major resistances.

Processing tomato industry was developed to provide North American and European households canned tomato, tomato paste and ketchup. Processing tomato varieties differ from fresh market ones in their pulp volume. Their growing conditions are dramatically different from fresh tomatoes (open field, mechanical harvesting). Thus the main criteria for processing tomatoes are fruit firmness, plant type with short fruit set period to produce a high percentage of ripe fruits simultaneously. Compact fruit set was obtained from a natural mutation discovered at the beginning of the XXth century, named *sp* (for self-pruning), conferring to the plant a determinate growth. This mutation was introduced in the well known "Roma" variety, whose long fruit type became a specific trait in processing tomatoes. In the 60's, VF145 was the first variety mechanically harvested. This cultivar has been the major cultivar for tomato ketchup industry for more than a decade in California. Apart from *sp*, several other mutations (detected in cultivated tomato or through interspecific

[3] Centro de Conservación y Mejora de la Agrodiversidad Valenciana, Valencia
[4] Center for Genetic resources, Wageningen
[5] Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben

hybridization) were used in tomato breeding. For example, the jointless (Szymkowiak and Irish 2005; Quinet, Kinet et al. 2011) *j2* allele was introgressed from *S. Cheesmanii*. In combination with *sp*, *j2* has been bred into many processing varieties, allowing a large scale mechanical harvest of tomato fruits. Major genes used in tomato breeding are listed in Table2.

Today, after a rapid evolution towards very firm fruits and long shelf life varieties (with the major success of the variety *Daniela*, which carried the spontaneous *rin* mutation), consumers request more diverse texture and tastes (Causse, Friguet et al. 2010). The fresh tomato market faces rapid developments and diversification (Navez 2011). New products and varieties are emerging always faster and their life cycle gets shorter, 5 years in average (Bai and Lindhout 2007). Consumption trend is for broader and diverse choice of fruit types. After the development of truss and cherry type tomatoes, new cultivars resembling to old heirloom varieties are developed. If access to allelic diversity is a must to improve fruit quality, choices in breeding objectives are critical to maintain organoleptic fruit values. Strong associations are often made by consumers between morphology and sensory values. Association of "Oxheart" or "Marmande" fruit shape with a pleasant texture is a good example. This link can be lost through modern breeding (Casals, Pascual et al. 2011). Improved content in potential health beneficial components such as anti-oxidants (lycopene, vitamin C) is also promoted. This can be obtained thanks to specific mutations like *hp* (Lieberman, Segev et al. 2004) or, again, by the introgression of genes from wild relatives.

Farmers and breeders have shaped diversity over years in an ever-evolving process that is hard to track and to record. Intensive breeding of crop varieties by modern science has increased the genetic erosion which started with domestication. Nevertheless the introgression from wild relatives allowed major progress and introduced a new source of diversity. Charles Rick observed that crosses between wild and cultivated species generated a large diversity of novel phenotypic diversity. Rick's work represents milestones for the modern use of genetic diversity in tomato. It led to uncover positive transgressive variation within interspecific progenies. This encouraged a greater use of exotic germplasm and thus larger gene pools to unlock causal polymorphism.

6. Biotechnology as a source of new diversity

Many natural mutants discovered in cultivated tomato have been extensively studied and characterized (http://tgrc.ucdavis.edu/), but their amount is limited. Thus reverse genetic techniques were developed aiming to discover gene function by analyzing the phenotypic effects of specific variants of targeted gene sequence. This approach is complementary to classical ("forwards") approaches, as they allow silencing or promoting the expression of targeted gene. They can also be used to generate genetic diversity within DNA sequences.

A delayed ripening tomato, named *Flavr-savr*® tomato, with a reduced expression of a cell wall protein (a polygalacturonase), obtained by an antisense construction, was the first transgenic plant released on the fresh market (Kramer and Redenbaugh 1994; Sanders and Hiatt 2005). This transformation significantly improved fruit shelf life and storage quality. Nevertheless, it was a commercial failure. Few years later, consumer's concerns about GMO[6],

[6] Genetically Modified Organism

Gene Name	Plant trait	Gene product	Phenotype	Chromosome(s)	Germplasm source	Reference
fer*	abiotic stress	bHLH protein	Iron-uptake response in roots	6	*S. lycopersicum*	Ling et al. (2002)
chlorosis*	abiotic stress	Nicotianamine synthase	Iron uptake	1	*S. pennellii*	Ling et al. (1999)
asc *	biotic stress	Novel protein of Longevity Assurance Gene (LAG1) family	*Alternaria alternata* (stem canker) resistance	3	*S. lycopersicum*	Brandwagt et al. (2000)
bs4 *	biotic stress	Toll/interleukin-1-receptor-nucleotide-binding-leucine-rich-repeat protein (TIR-NB	*Xanthomonas campestris* pv. vesicatoria resistanc	5	*S. pennellii*	Balhova et al. (2001); Schornack et al. (2004)
I-3 *	biotic stress	Toll/interleukin-1-receptor-nucleotide- binding site-leucine-rich-repeat (TIR-NBS-	*Fusarium oxysporum f. sp. lycopersici* resistance	7	*S. pennellii*	Hemming et al. (2004)
sw-5 *	biotic stress	Nucleotide-binding site-leucine-rich-repeat protein (NBS-LRR)	GNBV, GRSV, TCSV	9	*S. peruvianum*	Brommonschenkel et al. (2000)
mi-1*	biotic stress	Coiled-coil-nucleotide-binding site- leucine-rich repeat protein (CC-NBS-LRR)	Root-knot nematode resistance	6	*S. peruvianum*	Zhong et al. (1999)
tmv-1, tmv-2*	biotic stress	Unknown	Tobacco mosaic virus resistance	2, 9	*S. peruvianum*	Lanfermeijer et al. (2005)
pto*	biotic stress	Serine threonine protein kinase	*Pseudomonas syringae* pv. tomato resistance	5	*S. pimpinellifolium*	Martin et al. (1993)
hero *	biotic stress	Protein (NBS-LRR) Nucleotide-binding-site-leucine-rich-repeat	Nematode resistance	4	*S. pimpinellifolium*	Ernst et al.(2002)
cf-2 *	biotic stress	Leucine-rich repeat protein	Cladosporium fulvum resistance	6	*S. pimpinellifolium*	Dixon et al. (1996)
I-2*	biotic stress	Coiled-coil-nucleotide-binding site-leucine-rich repeat protein (CC-NBS-LRR)	*Fusarium oxysporum f. sp. lycopersici* resistance	11	*S. pimpinellifolium*	Ori et al. (1997)
ph-2, ph-3	biotic stress	Unknown	Phytophthora infestance resistance	9, 10	*S. pimpinellifolium*	Moreau et al. (1998); Chunwongse et al. (1998)
by-1	biotic stress	Unknown	Tomato yellow leaf curl virus resistance	6	*S.chilense, S. habrochaites*	Zamir (1994)
style2.1	flower development	Transcription factor regulating cell elongation	Flower style lenght	2	*S. lycopersicum*	Chen et al. (2007)
S	flower development	S-ribonucleases (pistil) , S-locus F-box (SLF) proteins (pollen)	Control unilateral interspecific incompatibility	2	*S. pennellii*	Li et al. (2010)
hp1*, hp2	fruit color	Damaged DNA binding protein 1 (DDB1)	High pigment, imunature fruit dark green	1,2	*S. lycopersicum*	Yen et al. (1997); Lieberman et al. (2004)
tangerine*	fruit color	CRTSO, Carotenoid isomerase	Carotenoid desaturation ,orange fruit	10	*S. lycopersicum*	Isaacson et al. (2002)
cr	fruit color	Chromoplast-specific lycopene beta cyclase Cyc-B	Crimson, increase lycopene	6	*S. lycopersicum*	Ronen et al. (2000)
beta oldgold*	fruit color	Lycopene -cyclase	Carotene synthesis lycopene increase, orange fru	6	*S. pennellii*	Ronen et al. (2000)
fgr	fruit content	Unknown	Increase fructose level	4	*S. habrochaites*	Levin et al. (2000)
brt-9-2-5*	fruit content	Apoplastic invertase (LIN5)	Increases sugar content and tomato yield	9	*S. pennellii*	Fridman et al. (2004)
cnr*	fruit development	SQUAMOSA promoter binding protein (SBP) box transcription factor	Colorless non-ripening mutant	2	*S. chmielewskii*	Manning et al. (2006)
Gr/Nr-2 *	fruit development	Novel protein, block ethylene perception	Green ripe/never ripe mutant, center turns red	1	*S. lycopersicum*	Barry and Giovannoni (2006)
tagl1	fruit development	PLENA subfamily of MADS-BOX genes	Yellow-orange fruit, reduced carotenoids	7	*S. lycopersicum*	Itkin et al. (2009)
rin *	fruit maturity	MADS-box transcription factor	Ripening inhibitor	5	*S. chmielewskii, S. pennellii*	Vrebalov et al. (2002)
nor	fruit maturity	MADS-box transcription factor	Ripening inhibitor	10	*S. lycopersicum*	Lanahan et al. (1994)
nr	fruit maturity	Unknown	Ripening inhibitor	9	*S. lycopersicum*	Yen et al. (1995)
ovate *	fruit shape	Novel protein with bipartite nuclear localization signal	Growth supressor, Pear shaped fruit	2	*S. lycopersicum*	Ku et al. (1999); Liu, J. et al. (2002)
sun*	fruit shape	IQ67 domain-containing protein	Elongated fruit shape	7	*S. lycopersicum*	van der Knaap et al. (2004); Xiao et al. (2008)
lcn*	fruit size	3' end of WUSCHEL (homeodomain protein)	Fruit size and locule number increase	2	*S. lycopersicum*	Munos et al. (2011)
fas*	fruit size	YABBY like transcription factor	Fruit size and locule number increase	11	*S. lycopersicum*	Cong et al. (2008)
fw2.2*	fruit size	ORFX similar to human oncogene c-H-ras p21	Major fruit weight locus	2	*S. pennellii*	Alpert and Tanksley (1996); Frary (2000)
fs8.1*	fruit size	Unknown	Fruit size, Imparts blocky, elongated shape	8	*S. pimpinellifolium*	Ku et al. (2000)
Obv*	leaf development	Unknown	Opaque leaf veins	5	*S. pennellii*	Jones et al. (2007)
ls*	plant development	VHIID protein	Lateral suppressor	7	*S. lycopersicum*	Schumacher et al. (1999)
sft*	plant development	Florigen precursor	Regulates transition and maintenance of flowers	3	*S. lycopersicum*	Molinero-Rosales et al. (2004)
an.s	plant development	F-box protein involved in transcriptional co-activation with the transcription factor	Inflorescence branching and floral identity	2	*S. lycopersicum*	Lippman et al. (2008)
j, j2*	plant development	MADS-box transcription factor	Jointless, flower abscission zone development	11,12	*S. lycopersicum, S. chmielewskii*	Zhang et al. (1994); Mao et al. (2000)
sp *	plant development	Ortholog of CENTRORADIALIS and TERMINAL FLOWER1	Self pruning	6	*S. pennellii*	Pnueli et al. (1998)
sw4.1*	seed development	ABC transporter	Seed weight	4	*S. lycopersicum S. pimpinellifolium*	Orsi et al. (2009)

* cloned genes

Table 2. List of genes characterized through molecular techniques with their related function and germplasm origin

as well as the high engineering cost, stopped further commercial developments. In the research field, transformation with *Agrobacterium thumefaciens* is still widely used for the functional characterization of specific genes. For instance, transformed tomato plants were produced to enable study of endotoxins genes (Zhang, Buehner et al. 2006) plant disease resistance genes, abiotic stress genes or to produce molecules useful in human medicine (Sharma, Singh et al. 2008).

TILLING (Target Induced Local Lesion In Genomes), a mutagenesis technique, has experienced important development (Comai and Henikoff 2006). Early days of this technique were in the 50's (Rick 1991). It is now widely used for reverse genetics to generate and identify induced point mutations in genomes. A chemical reagent (Ethylmethane Sulphonate) is used to induce genetic mutations. Collections of tomatoes carrying artificially induced genetic variants, called mutant libraries are currently available (Menda, Semel et al. 2004; Minoia et al. 2010) or under development (Okabe, Asamizu et al. 2011). In contrast to transgenic methods, mutagenesis is random, cost effective and is not submitted to GMO regulation. TILLING allows generating variants in cultivated genetic background (Piron, Nicolaï et al. 2010) and thus transfer rapidly interesting mutations into cultivars (Gady, Hermans et al. 2009). Application of TILLING technique to screen for natural variation within tomato germplasm collection is now performed (Rigola, van Oeveren et al. 2009).

7. Molecular markers offer a new vision of tomato diversity

Natural genetic diversity is the fuel of evolution. No evolutive forces or adaptation to environment changes can apply without it (Alonso-Blanco, Aarts et al. 2009). Consequently it is a vital characteristic for species adaptation in general and for crop breeding in particular. Genetic variation occurs both within cultivated tomato (intraspecific) and between wild species (interspecific). Tomato breeding for adaptation to specific growing areas is in progress for more than two centuries now (Stevens and Rick 1986). Since the early days of quantitative genetics, initiatives were developed to improve the understanding of trait inheritance. Attempts to construct genetic maps based on interspecific crosses (*S. pimpinellifolium* x *S. lycoperpersicum*) and to map disease resistance genes are performed for years (Langford 1937). A linkage map showing the distribution of agronomic trait with Mendelian inheritance, based on linkage between two or three mutations, was proposed (Butler 1952). Nevertheless, the lack of polymorphic and neutral markers was strongly limitant. Development of isozymes allowed a first evaluation of wild germplasm (Rick and Fobes 1975) and introgression diagnostic (Tanksley, Medina-Filho et al. 1981), but isozyme marker scarcity and their low polymorphism was still limitant. This limitation was progressively overcome since the 80's thanks to the discovery of several molecular marker types.

7.1 Ecological and evolution in wild tomato related species

Molecular studies provide important clues into ecological and evolutionary questions in wild tomatoes species. In speciation process, hybrid sterility is frequently due to dysfunctional interactions between loci that accumulate between different lineages. A "snowballing effect" characterizes loci controlling such reproductive barrier and hybrid sterility that should accumulate faster than linearly with time. Such "snowballing" effect has been recently described within distinct populations derived from crosses of *S. lycopersicum* with *S. pennellii*,

S. habrochaites and *S. lycopersicoides* (Moyle and Nakazato 2010). However, further investigations are suggested to confirm these results (Stadler, Florez-Rueda et al. 2011).

Tellier and colleagues quantified the number of adaptive and deleterious mutations and the distribution of fitness effects of new mutations within housekeeping genes in 4 species, *S. arcanum, S. chilense, S. habrochaites* and *S. peruvianum*. Little evidence for adaptive mutations was shown but strong purifying selection in coding regions was detected (Tellier, Fischer et al. 2011). This suggests that closely related species with similar genetic backgrounds but contrasted environments differ in the frequency of deleterious fitness effects.

The west coastal area between the Andes and the ocean, from Ecuador to Chile is widely recognized as the center of origin of the species from the Solanum sect. *lycopersicon*. This area covers a wide range of geographical conditions. Complex geography and ecology of Andes had a major impact in species divergence and hybridization between *S. pimpinellifolium* and *S. lycopersicum* (Nakazato and Housworth 2011). The two species present a distinct lineage, separated by the Andes. They hybridize extensively in north and central Ecuador. Nakazato and colleagues demonstrated using molecular markers and geographic information system (GIS) data that *S. lycopersicum* has likely experienced a severe population bottleneck during the colonization of the eastern Andes followed by a rapid population expansion. In plant, resistance genes and homologs (RGA) tend to be highly variable. Caicedo et al (2004) studied the geographic distribution of a RGA family Cf-2 (see Table 2) within and among plant populations of *S. pimpinellifolium*. They underlined that the geographical distribution of RGA diversity has been primarily shaped by demographic factors and selective pressure (Caicedo and Schaal 2004; Caicedo 2008). The authors underlined the reduction of natural habitat. This phenomenon is also observed on Galapagos Islands. The endemic species *S. cheesmanii* shows a reduction of its population due to human activity. Differentiation within *S. cheesmanii* was also observed (Nuez, Prohens et al. 2004) as well as hybridization with the two introduced species *S. lycopersicum* and *S. pimpinellifolium* (Darwin, Knapp et al. 2003).

7.2 Diversity analysis among wild and cultivated germplasm

Allelic richness (number of different alleles segregating in the population) is used to measure the genetic diversity and is considered as a key parameter for genetic resources management. It reveals past fluctuations in population size (Nei, Maruyama et al. 1975). Molecular differences between more than 200 Peruvian and Ecuadorian *S. pimpinellifolium* accessions were highlighted by Zuriaga and colleagues. Climate and genetic data were highly correlated. Thus the non-uniform nature of climates between the two countries is shown to be an important factor. Highest diversity was found in North Peru, lowest on Galapagos Islands. Authors stressed the fact that interspecific variation between *S. pimpinellifolium* and *S. lycopersicum* was indicating a very close relatedness between the two species (Zuriaga, Blanca et al. 2009).

Cherry tomato accessions show typically a large genetic diversity and an intermediate fruit size *between S. pimpinellifolium* and large cultivated ones. Botanists postulate that cherry tomato accessions are feral plants (also called revertant) or a possible genetic admixture of wild and cultivated germplasm (Rick and Holle 1990; Peralta et al. 2007a). Recently molecular analysis of the structure of a large set of accessions of wild *S. pimpinellifolium*, cherry tomato and cultivated accessions showed that domesticated and wild tomatoes have

evolved as a species complex with intensive hybridization. This highlighted the admixture position of *S. lycopersicum* var. *cerasiforme* (Ranc, Muños et al. 2008) which is illustrated on figure 2 using a data from Ranc et al (2010) and analyzed using Structure 2.0 (Pritchard et al. 2000) output data. Accessions display clustering patterns (circled) following two phenotypic traits: fruit size and stigma insertion. Structuration effect of those domestication traits can be observed. The emergence of molecular markers has allowed quantifying with accuracy the diversity within germplasm material. The first molecular diversity studies on cultivated tomato revealed the very low polymorphism compared to wild species, whether it was based on RFLP[7] (Miller and Tanksley 1990), SSR[8] (Bredemeijer, Cooke et al. 2002; He, Poysa et al. 2003) AFLP[9] (Park, West et al. 2004; Berloo, Zhu et al. 2008), SSAP[10] (Tam, Mhiri et al. 2005) or SNP[11] (Yang, Bai et al. 2004; Labate and Baldo 2005). However, Bredemeijer et al (2002) characterized 500 cultivated lines from European lines and showed that it was possible to distinguish them all from each other using a set of 20 SSR markers. When comparing old varieties (or landraces) to modern hybrids, a higher level of molecular diversity in landraces is usually observed (Mazzucato, Papa et al. 2008; van Berloo, Zhu et al. 2008).

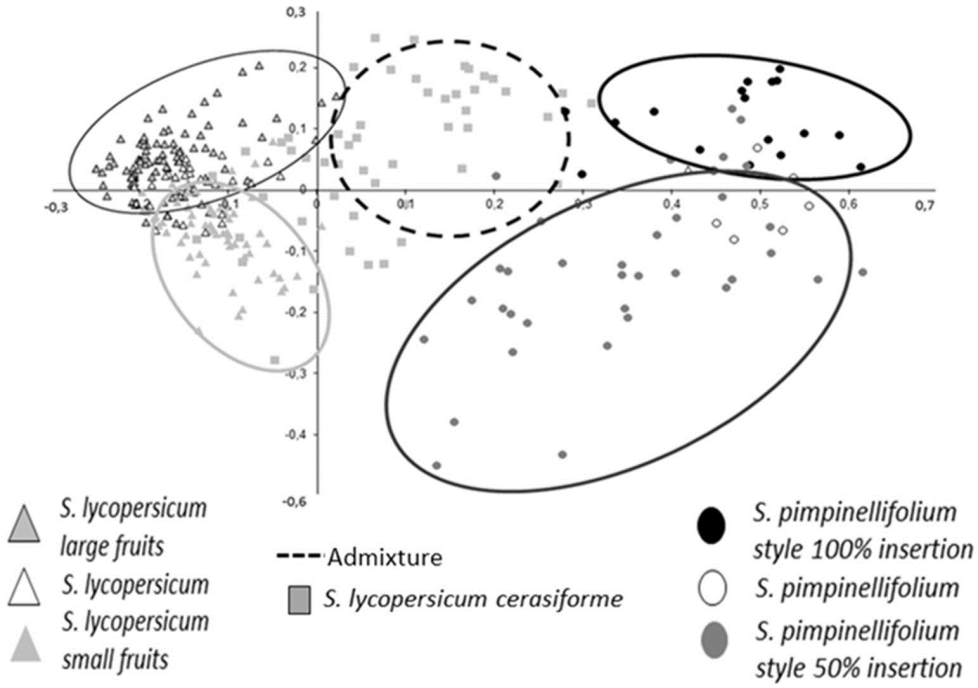

Fig. 2. Principal Coordinate Analysis of 318 accessions tomato core collection.

[7] Restriction Fragment Lenght Polymorphism
[8] Simple Sequence Repeats
[9] Amplified Fragment Lenght Polymorphism
[10] Sequence-Specific Amplification Polymorphism
[11] Single Nucleotide Polymorphism

If interspecific populations for genetic analyses and diversity studies answered to many questions, it has left a void in the understanding of genotypic variation within tomato breeding programs which focus on intra-specific populations (Van Deynze, Stoffel et al. 2007). The recent discovery of SNP markers, first detected in EST (expressed sequenced tag) sequences (Van Deynze et al. 2007; Jimenez-Gomez and Maloof 2009) then in non-coding sequences (Labate et al. 2009) provided access to a higher level of polymorphism. Labate and colleagues estimated parameters of diversity among *S. lycopersium* accessions, first using the SNP detected in 50 loci that were resequenced in a diversity panel of 31 accessions. In a second investigation, multilocus estimates of polymorphism were obtained and led to rejection of the neutral equilibrium model of evolution within the studied collection (Labate, Robertson et al. 2009). Public germplasm are potential allele mining sources for crop improvement as illustrated by previous authors who sampled among US seed banks 30 accessions from the five continents. The study confirmed that history of crossing with wild tomato species and distribution among different environments across the world has spread allelic variation (Labate, Sheffer et al. 2011).

Molecular markers have proven their efficiency in sampling and maximizing allelic richness (Schoen and Brown 1993) through the development of nested core collections (McKhann, Camilleri et al. 2004). Such nested core collections (from 8 to 96 accessions) were constructed in tomato, capturing most of the molecular and phenotypic variation present in a set of 360 constituted of wild, feral and cultivated accessions (Ranc, Muños et al. 2008).

7.3 Use of molecular diversity to dissect phenotypes

Molecular markers allowed the construction of high density genetic maps of the tomato genome (Tanksley, Ganal et al. 1992). This permitted the dissection of quantitative traits into Mendelian factors or QTL (Quantitative Trait Loci) (Paterson, Lander et al. 1988; Tanksley et al. 1992). This strategy also opened the way to investigate physical mapping and molecular cloning of genetic factors underlying quantitative traits (Paterson, Damon et al. 1991). Moreover, *Lycopersicon* varieties and related species are all diploid and chromosomally collinear, making genetic dissection straightforward. The first gene cloned by positional cloning was the *Pto* gene, confering resistance to *Pseudomonas syringae* (Martin, Brommonschenkel et al. 1993). Since then, interspecific crosses with each wild species were performed. Due to the low genetic diversity within the cultivated compartment (Miller and Tanksley 1990), most of the mapping populations are based on interspecific crosses between a cultivar and related wild species from the lycopersicon group (as reviewed by Foolad (2007) ;Labate, Grandillo et al.(2007); Grandillo et al (2011)) or from lycopersicoides (Pertuzé, Ji et al. 2002) and juglandifolia group (Albrecht, Escobar et al. 2010). However, maps based on intraspecific crosses have proved their interest notably on fruit quality aspects (Saliba-Colombani, Causse et al. 2001). All those populations allowed discovering and/or characterizing a myriad of major genes (Table 2) and QTLs involved in various traits.

Rapidly, molecular breeding strategies were set up and implemented to "pyramid" genes of interest for agronomical traits, notably using Advanced Backcross QTL method (AB-QTL) (Tanksley, Grandillo et al. 1996). Using this approach with a *S. lycopersicum x S. pimpinellifolium* progeny, in which agronomical favorable QTL alleles were detected, Grandillo and colleagues showed how a wild species could contribute to improve

cultivated tomato (Tanksley, Grandillo et al., 1996). Introgression Lines (IL) derived from interspecific crosses allowed to dissect the effect of chromosome fragments from a donor (usually from a wild relative) introgressed into a recurrent elite line. IL offer the possibility to evaluate the agronomic performance of a specific set of QTL (Paran, Goldman et al. 1995). IL was used as a base for fine mapping and positional cloning of several genes and QTL of interest. The first IL library was developed between *S. pennellii* and *S. lycopersicum* (Eshed and Zamir 1995; Zamir 2001). QTL mapping power was increased compared to biallelic QTL mapping population, and was again improved by the constitution of sub-IL set with smaller introgressed fragments. This progeny was successful in identifying QTLs for fruit traits (Causse, Duffe et al 2004); anti-oxidants (Rousseaux, Jones et al. 2005), vitamin C (Stevens, Buret et al. 2007) and volatile aromas (Tadmor, Fridman et al. 2002). The introgression of a QTL identified in these IL has allowed plant breeders to boost the level of soluble solids (brix) in commercial varieties and largely increased tomato yield in California (Fridman, Carrari et al. 2004). Such exotic libraries were thus designed with several species, involving *S. pimpinellifolium* (Doganlar, Frary et al. 2002), *S. habrochaites* (Monforte and Tanksley 2000; Finkers, van Heusden et al. 2007) and *S. lycopersicoides* (Canady, Meglic et al. 2005).

Introgression lines were also used to dissect the genetic basis of heterosis (Eshed and Zamir 1995). Heterosis refers to phenomenon where hybrids between distant varieties or crosses between related species exhibit greater biomass, speed of development, and fertility than both parents (Birchler, Yao et al. 2010). Heterosis involves genome–wide dominance complementation and inheritance model such as locus–specific overdominance (Lippman and Zamir 2007). Heterotic QTL for several trait were identified in tomato IL (Semel and Nissenbaum, 2006). A unique QTL was shown to display at the heterozygous level improved harvest index, earliness and metabolite content (sugars and amino acids) in processing tomatoes (Gur, Osorio et al. 2010; Gur, Semel et al. 2011) Furthermore, a natural mutation in the SFT gene, involved in flowering (Shalit, Rozman et al. 2009), was shown to correspond to a single overdominant gene increasing yield in hybrids of processing tomato (Krieger, Lippman et al. 2010).

Metabolite detection is an approach of choice to identify compounds involved in fruits quality traits. Metabolite QTL (mQTL) can be now identified for non-volatile metabolites like sugars, pigments or volatiles compounds (Bovy, Schijlen et al. 2007). This was done on several interspecific populations, notably on *S. lycopersicum x S. Chmielewskii* (Do, Prudent et al. 2010) and intraspecific crosses (Saliba-Colombani et al. 2001; Causse, Saliba-Colombani et al. 2002; Zanor, Rambla et al. 2009). Recent technologies allowed screening for diversity in a wide range of components on whole genomes. This can be done in a targeted way to better characterize known metabolites (Tieman, Taylor et al. 2006) or untargeted manner to identify new metabolites (Tikunov, Lommen et al., 2005). Further than identify and quantify compounds, metabolomics can be of great help to decipher biosynthetic pathways (Keurentjes 2009). Metabolome studies can be combined to transcriptomic data to identify the key factors (Mounet, Moing et al. 2009; Do, Prudent et al. 2010). Metabolomics has an important role to play in characterization of natural diversity in tomato (Schauer, Zamir et al. 2004; Fernie et al. 2011). As well, it can boost the biochemical understanding of fruit content and be an enhancer for quality breeding (Fernie and Schauer 2009; de Vos, Hall et al. 2011).

7.4 Dissection of the molecular bases of domestication and diversification

Product of human domestication and later diversification of fruit types, led to a large morphological diversity in tomato fruit (with small to large, round, blocky, elongated, pear shaped fruits, with color ranging from red to green, white, black, pink, orange or yellow). On the contrary, wild tomato species carry small, round red or green fruits, with a low intraspecific phenotypic diversity. This has drawn scientist attention on the inheritance and development of fruit size and shape in the tomato (Yeager 1937). Influence of chromosome 2 in fruit morphology was noticed (Butler 1964). Thus, using available molecular techniques, fruit traits genetic control has been widely dissected (Grandillo, Ku et al. 1999; Lippman and Tanksley 2001; Barrero and Tanksley 2004). The first QTL, fw2.2, controlling fruit weight variation was cloned (Frary, Nesbitt et al 2000). It has been suggested that diversity of fruit shape in cultivated germplasm can be explained to a large extent by four genes (Rodriguez, Muños et al. 2011). The study established a model for fruit shape evolution in tomato. This model includes four major mutations recently identified: FAS which increases locule number, fruit fasciation and size (Cong, Barrero et al. 2008), LC which increases locule number and fruit size (Muños, Ranc et al. 2011), OVATE which gives ovoid fruit shape (Liu, van Eck et al 2002) and SUN which gives an elongated fruit shape (van der Knaap, Lippman et al. 2002; Xiao, Jiang et al. 2008) or the oxheart shape when associated to LC and FAS. The allelic distribution of the four genes was associated with morphologic, geographical and historical data in a collection of diverse cultivated accessions. This study established that the selection occurred in distinct chronologic and historic periods: LC arose first, followed by OVATE, both in *S.l. cerasiforme* background but in distinct populations. FAS arose later in a LC background. Presence of those three mutations in Latin American germplasm suggests Pre-Columbian mutations. Combined with fw2.2, they must have strongly contributed to the increase in fruit size during tomato domestication. On the contrary, SUN mutation is not carried by any Latin American material tested, suggesting that SUN mutation appeared post domestication in European material (probably in Italy). This study also showed that the selection for fruit shape is strongly responsible for the underlying genetic structure in tomato cultivars. The recent discoveries of the molecular events shaping tomato fruit indicate that the germplasm is frequently more diverse phenotypically than the wild related germplasm but not necessarily showing a similar pattern at the molecular level. *"The irony of all this,"* says Steve Tanksley (geneticist at Cornell University, and precursor of all these studies) *"is all that diversity of heirlooms can be accounted for by a handful of genes. There are probably no more than 10 mutant genes that create the diversity of heirlooms you see"* (Borrell 2009). Tomato selection and spread worldwide has led to the immense diversity of varieties that characterizes many domesticated plant species (Purugganan and Fuller 2009).

8. Association genetics: New valorization of natural diversity

Recent advance in molecular genetics and computation has allowed the emergence of association mapping (Myles, Peiffer et al. 2009). Association mapping takes advantage of historical recombination events and natural genetic diversity. By using large numbers of lines and molecular markers over the whole genome, the resolution of Genome Wide Association studies (GWAS) is much higher than in conventional segregating populations. Such approach requires an accurate estimate of the genetic structure of the sample studied (Price, Zaitlen et al. 2010) and linkage disequilibrium (LD) extend among loci. Yu and

colleagues (2005) proposed a unified mixed model taking into account the genetic structure of the sample, based on single locus analysis. This model is being updated by integrating a multi-locus analysis (Ayers et al. 2010).In autogamous crops, it is expected that large extent of LD will reduce the resolution and risks to lead to false positive associations. Nevertheless, successful results have been obtained in selfing crops (Atwell, Huang et al. 2010; Ramsay, Comadran et al. 2011).

In tomato, several studies revealed contrasted results according to the samples studied. First studies of the linkage disequilibrium revealed large LD in cultivated tomatoes (Mazzucato, Papa et al, 2008; van Berloo, Zhu et al, 2008; Robbins, Sim et al, 2010). Van Berloo and colleagues performed association mapping within a collection of 94 accessions containing both old and elite (hybrids) European germplasm and about 300 markers (AFLP). Structure coinciding with fruit size was identified allowing grouping between cherry tomato and round-beef types, extensive LD was observed (15 cM average). Robbins and colleagues investigated the population structure among 70 tomato cultivars (modern and vintage, from fresh and processing market). The STRUCTURE analysis (Pritchard, Stephens et al. 2000) revealed groups predefined by market niche and age into distinct subpopulations. Furthermore, they detected two subpopulations within the processing varieties, corresponding to historical patterns of breeding conducted for specific production environments. They found no subpopulation within fresh-market varieties. High levels of admixture were shown in several varieties representing a transition in the demarcation between processing and fresh-market. Mapping and LD analysis on a genome wide level was performed (Robbins, Sim et al. 2010). Using a panel of 102 accessions including 95 cultivars (heirloom, fresh and processing cultivars) and 9 wild species), effect of selection on genome variation was studied using 340 markers (SNP, SSR, and INDEL[12]). LD value varied from 6-8 cM (all accessions) up to 3-16 cM (fresh market cultivars). Inter-chromosomal LD appeared to be population dependent, suggesting cautious approach for association mapping. Notably, a genetic divergence between fresh market and processing types was also shown. On the contrary, the use of cherry tomato allowed the construction of core collection with a reduced structure and lower LD (Ranc et al, 2008; 2010). In a pilot study on chromosome 2, using markers distant from several cM to few kb, Ranc (2010) showed that LD varied strongly from one region to the other. A few distant markers remained in strong LD, but could be removed from the analysis.

The first association study was performed by Nesbitt and Tanksley (2002) to identify the SNP responsible for FW2.2 gene they had cloned. They failed to find any association between fruit size and genomic sequence of the *fw2.2* region in a collection of 39 cherry tomato accessions. Ranc and colleagues (2010) identified significant association in the promoter region, thanks to a larger and more representative sample. From a breeding point of view, the admixture mapping between the cultivated tomato and its closest relative is a method of choice for allele mining in wild germplasm. Muños and colleagues (2011) used this approach to identify causal polymorphism of QTL controlling locule number on chromosome 2. New SNP arrays are now available thanks to Next Generation Sequencing technologies (NGS), as the genotyping array developed under the Solanaceae Coordinated Agriculture Project (SolCAP) initiative carrying 7,000 effective SNP (SolCAP 2008). These tools will be very useful to scan the whole genome for associations.

[12] Insertion-Deletion

9. Conclusion: Toward a change in the way to manage and use diversity

Crossing wild and cultivated species can reveal alleles left behind during the domestication process. Molecular markers strongly helped to reinforce the use of wild relatives (Zamir 2008). Interfacing genetic resources management and plant breeding, pre-breeding is now recognized as an important adjunct to plant breeding, as a way to introduce new traits from non-adapted populations and wild relatives, notably for abiotic stress (FAO 2010). Nevertheless, the extensive use of this genetic richness contained in seed banks and germplasm collection faces limits. The difficulty to introgress accurately the targeted allele (with favorable effect) without unfavorable ones, carried on by "linkage drag", remains.

With the emergence of bioinformatics and nanotechnologies -so called "post-genomics" era- the last decade has opened high throughput sequencing era. Now, conducting large intra-specific studies becomes a reality in tomato, allowing a better characterization of its genetic diversity. With the completion of its genome sequence (Mueller, Lankhorst et al. 2009) and rich annotation as well as a large number of tools available via SGN (SOL Genome Network; http://solgenomics.net/organism/solanum_lycopersicum/genome) platform (Bombarely, Menda et al. 2010), tomato and its relatives is the most advanced vegetable crop. A draft of the genome sequence of *S. pimpinellifolium* LA1589 is also released by D. Ware, W. R. McCombie, and Z. B. Lippman at Cold Spring Harbor Laboratory allowing a detailed comparison of both species. The genome sequences of tomato provide clues for understanding the Solanum clade evolutive history and identify genes involved in fleshy fruit development.

Progress in sequencing technologies has reached the point where genotyping by sequencing (GBS) is now possible (Davey, Hohenlohe et al. 2011; Elshire, Glaubitz et al. 2011). This opens new perspectives in terms of genetic diversity management, notably toward conservation and survey of large populations. In a near future, techniques such as GBS may allow breeders and scientists of the tomato community to determine population characteristics prior concretely establishing genome or nucleotide diversity. GBS opens ways to a global and quantitative management of diversity, and let foresee an *a priori* genetic resource management. It also opens perspectives in allele based breeding called genomic selection (Hamblin, Buckler et al. 2011).

If *ex situ* germplasm conservation is well developed and will benefit of these developments, *in situ* conservation of tomato and its wild relatives is becoming critical due to major ecological changes in its area origin. Efforts on *in situ* conservation and participatory approaches as proposed by Jarvis, Brown et al (2008) and (Thomas, Dawson et al. 2011) could be very useful to maintain the adaptive potential of tomato genetic resources. Nuez and colleagues proposed to use *S. cheesmanii* accessions now stored in germplasm banks to reinstate some extinct populations in Galapagos Islands (Nuez et al. 2004). This could help avoiding the present paradox: the more knowledge we gain on tomato diversity and its evolutive history, the less available those genetic resources are available in the wild.

10. References

Albrecht, E., M. Escobar, et al. (2010). "Genetic diversity and population structure in the tomato-like nightshades Solanum lycopersicoides and S. sitiens." *Annals of Botany* 105(4): 535-554.

Alonso-Blanco, C., M. G. M. Aarts, et al. (2009). "What Has Natural Variation Taught Us about Plant Development, Physiology, and Adaptation?" *The Plant Cell Online* 21(7): 1877-1896.

Atwell, S., Y. S. Huang, et al. (2010). "Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines." *Nature* 465(7298): 627-631.

Ayers, K. L. and H. J. Cordell (2010). "SNP Selection in genome-wide and candidate gene studies via penalized logistic regression." Genetic Epidemiology 34(8): 879-891.

Bai, Y. and P. Lindhout (2007). "Domestication and Breeding of Tomatoes: What have We Gained and What Can We Gain in the Future?" *Annals of Botany* 100(5): 1085-1094.

Barrero, L. S. and S. D. Tanksley (2004). "Evaluating the genetic basis of multiple-locule fruit in a broad cross section of tomato cultivars." *Theoretical and Applied Genetics* 109(3)): 669-679.

Barry, C. S. and J. J. Giovannoni (2006). "Ripening in the tomato Green-ripe mutant is inhibited by ectopic expression of a protein that disrupts ethylene signaling." *Proceedings of the National Academy of Sciences* 103(20): 7923-7928.

Beaumont, M. A., W. Zhang, et al. (2002). "Approximate Bayesian Computation in Population Genetics." *Genetics* 162(4): 2025-2035.

Bedinger, P., R. Chetelat, et al. (2011). "Interspecific reproductive barriers in the tomato clade: opportunities to decipher mechanisms of reproductive isolation." *Sexual Plant Reproduction* 24(3): 171-187.

Berloo, R., A. Zhu, et al. (2008). "Diversity and linkage disequilibrium analysis within a selected set of cultivated tomatoes." *Theoretical and Applied Genetics* 117(1): 89-101.

Biodiversity-International (2006) "Crop wild relatives." 25.

Birchler, J. A., H. Yao, et al. (2010). "Heterosis." *The Plant Cell Online* 22(7): 2105-2112.

Bombarely, A., N. Menda, et al. (2010). "The Sol Genomics Network (solgenomics.net): growing tomatoes using Perl." *Nucleic Acids Research* 39(Database): D1149-D1155.

Borrell, B. (2009). "How to Grow a Better Tomato: The Case against Heirloom Tomatoes." *Scientific American*
http://www.scientificamerican.com/article.cfm?id=case-against-heirloom-tomatoes

Bovy, A., E. Schijlen, et al. (2007). "Metabolic engineering of flavonoids in tomato (*Solanum lycopersicum*): the potential for metabolomics." *Metabolomics* 3(3): 399-412.

Brandwagt, B. F., L. A. Mesbah, et al. (2000). "A longevity assurance gene homolog of tomato mediates resistance to *Alternaria alternata* f. sp. *lycopersici* toxins and *fumonisin* B1." *Proceedings of the National Academy of Sciences* 97(9): 4961-4966.

Bredemeijer, M., J. Cooke, et al. (2002). "Construction and testing of a microsatellite database containing more than 500 tomato varieties." *Theoretical and Applied Genetics* 105(6-7): 1019-1026.

Brommonschenkel, S. H., A. Frary, et al. (2000). "The Broad-Spectrum Tospovirus Resistance Gene Sw-5 of Tomato Is a Homolog of the Root-Knot Nematode Resistance Gene Mi." *Molecular Plant-Microbe Interactions* 13(10): 1130-1138.

Butler, L. (1952). "The linkage map of the tomato." *Journal of Heredity* 43: 25-36.

Caicedo, A. L. (2008). "Geographic diversity cline of R gene homologs in wild populations of *Solanum pimpinellifolium* (Solanaceae)." *American Journal of Botany* 95(3): 393-398.

Caicedo, A. L. and B. A. Schaal (2004). "Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene." *Molecular Ecology* 13(7): 1871-1882.

Canady, M. A., V. Meglic, et al. (2005). "A library of *Solanum lycopersicoides* introgression lines in cultivated tomato." *Genome* 48(4): 685-697.

Casals, J., L. Pascual, et al. (2011). "The risks of success in quality vegetable markets: Possible genetic erosion in Marmande tomatoes (*Solanum lycopersicum* L.) and consumer dissatisfaction." *Scientia Horticulturae* 130(1): 78-84.

Causse, M. Duffe, P. et al (2004). "A genetic map of candidate genes and QTLs involved in tomato fruit size and composition." *Journal of Experimental Botany* 55(403): 1671-1685.

Causse, M., C. Friguet, et al. (2010). "Consumer Preferences for Fresh Tomato at the European Scale: A Common Segmentation on Taste and Firmness." *Journal of Food Science* 75(9): S531-S541.

Causse, M., V. Saliba-Colombani, et al. (2002). "QTL analysis of fruit quality in fresh market tomato: a few chromosome regions control the variation of sensory and instrumental traits." *Journal of Experimental Botany* 53(377): 2089-2098.

Chen, K. Y., B. Cong, et al. (2007). "Changes in Regulation of a Transcription Factor Lead to Autogamy in Cultivated Tomatoes." *Science* 318(5850): 643-645.

Chunwongse, J. C., L. Black, et al. (2002). Molecular mapping of the Ph-3 gene for late blight resistance in tomato. *Journal of Horticultural science and Technology* 77(3): 281-286.

Comai, L. and S. Henikoff (2006). "TILLING: practical single-nucleotide mutation discovery." *The Plant Journal* 45(4): 684-694.

Cong, B., L. S. Barrero, et al. (2008). "Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication." *Nature Genetics* 40(6): 800-804.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection*. London, J. Murray.

Darwin, C. and A. Wallace (1858). "On the Tendency of Species to form Varieties; and on the Perpetuation of Varieties and Species by natural means of selection." *Journal of the Proceedings of the Linnean Society of London. Zoology* 3(9): 45-62.

Darwin, S. C., S. Knapp, et al. (2003). "Taxonomy of tomatoes in the Galpagos Islands: native and introduced species of *Solanum* section *Lycopersicon* (Solanaceae) " *Systematics and Biodiversity* 1(1): 29-53.

Davey, J. W., P. A. Hohenlohe, et al. (2011). "Genome-wide genetic marker discovery and genotyping using next-generation sequencing." *Nature Reviews Genetics* 12(7): 499-510.

de Candolle, A. P. (1882). *L'origine des plantes cultivées*. Paris.

de Vos, R. C. H., R. D. Hall, et al. (2011). Metabolomics of a Model Fruit: Tomato. *Annual Plant Reviews* 109-155.

Diez M.-J. and Nuez, F. (2008). Tomato. *Vegetables II Fabaceae, Liliaceae, Solanaceae, and Umbelliferae*. Prohens J. and Nuez, F.: 249-323.

Dixon, M. S., D. A. Jones, et al. (1996). "The Tomato Cf-2 Disease Resistance Locus Comprises Two Functional Genes Encoding Leucine-Rich Repeat Proteins." *Cell* 84(3): 451-459.

Do, P. T., M. Prudent, et al. (2010). "The influence of fruit load on the tomato pericarp metabolome in a *Solanum chmielewskii* introgression line population." *Plant Physiology* 154(3): 1128-1142.

Doebley, J. F., B. S. Gaut, et al. (2006). "The molecular genetics of crop domestication." *Cell* 127(7): 1309-1321.

Doganlar, S., A. Frary, et al. (2002). "Mapping quantitative trait loci in inbred backcross lines of *Lycopersicon pimpinellifolium* (LA1589)." *Genome* 45(6): 1189-1202.

Elshire, R. J., J. C. Glaubitz, et al. (2011). "A robust, simple genotyping-by-sequencing (GBS) Approach for High Diversity Species." *PLoS ONE* 6(5): e19379.

Ernst, K., A. Kumar, et al. (2002). "The broad-spectrum potato cyst nematode resistance gene (Hero) from tomato is the only member of a large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region." *The Plant Journal* 31(2): 127-136.

Eshed, Y. and D. Zamir (1995). "An Introgression Line Population of Lycopersicon pennellii in the Cultivated Tomato Enables the Identification and Fine Mapping of Yield-Associated QTL." *Genetics* 141(3): 1147-1162.

FAO (2010). *Plant genetic ressource for food and agriculture*. Rome, Food and Agriculture Organisation of the United Nations

Fernie, A. and H. Klee (2011). "The use of natural genetic diversity in the understanding of metabolic organization and regulation." Frontiers in Plant Science 2.

Fernie, A. R. and N. Schauer (2009). "Metabolomics-assisted breeding: a viable option for crop improvement?" *Trends in Genetics* 25(1): 39-48.

Finkers, R., H. de Weerd, et al. (january 2011). *BreeDB:* A database suppporting quantitative aspects of plant breeding. Plant and Animal Genomes XIX, San Diego.

Finkers, R., A. van Heusden, et al. (2007). "The construction of a Solanum habrochaites LYC4 introgression line population and the identification of QTLs for resistance to Botrytis cinerea." *Theoretical and Applied Genetics* 114(6): 1071-1080.

Food, Royal Norvegian Ministry of Agriculture (2008) "Svalbard global seedvault - more about the constuction from the directorate of public construction and property." 27.

Foolad, M. R. (2007). "Genome Mapping and Molecular Breeding of Tomato." *International Journal of Plant Genomics* 2007: 1-52.

Frary, A., Nesbitt et al (2000). "fw2.2: A quantitative trait locus key to the evolution of tomato fruit size." *Science* 289(5476): 85-88.

Fridman, E., F. Carrari, et al. (2004). "Zooming on a quantitative trait for tomato yield using interspecific introgressions." *Science* 305(5691): 1786-1789.

Gady, A., F. Hermans, et al. (2009). "Implementation of two high through-put techniques in a novel application: detecting point mutations in large EMS mutated plant populations." *Plant Methods* 5(1): 13.

García-Martínez, S., L. Andreani, et al. (2006). "Evaluation of amplified fragment length polymorphism and simple sequence repeats for tomato germplasm fingerprinting: utility for grouping closely related traditional cultivars." *Genome* 49(6): 648-656.

Grandillo, S., R. Chetelat, et al. (2011). *Solanum* sect. *Lycopersicon Wild Crop Relatives: Genomic and Breeding Resources*. C. Kole, Springer Berlin Heidelberg: 129-215.

Grandillo, S., H. M. Ku, et al. (1999). "Identifying the loci responsible for natural variation in fruit size and shape in tomato." *Theoretical and Applied Genetics* 99(6): 978-987.

Gur, A., S. Osorio, et al. (2010). "*hi2-1*, A QTL which improves harvest index, earliness and alters metabolite accumulation of processing tomatoes." *Theoretical and Applied Genetics* 121(8): 1587-1599.

Gur, A., Y. Semel, et al. (2011). "Yield quantitative trait loci from wild tomato are predominately expressed by the shoot." *Theoretical and Applied Genetics* 122(2): 405-420.

Hamblin, M. T., E. S. Buckler, et al. (2011). "Population genetics of genomics-based crop improvement methods." *Trends in Genetics* 27(3): 98-106.

Harlan, J. R. (1971). "Agricultural Origins: Centers and Noncenters." *Science* 174(4008): 468-474.

He, C., V. Poysa, et al. (2003). "Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars." *Theoretical and Applied Genetics* 106(2): 363-373.

Hemming, M. N., S. Basuki, et al. (2004). "Fine mapping of the tomato gene for fusarium wilt resistance and elimination of a co-segregating resistance gene analogue as a candidate for *I-3*." *Theoretical and Applied Genetics* 109(2): 409-418.

Hey, J. and R. Nielsen (2004). "Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*." *Genetics* 167(2): 747-760.

Horkheimer, H. (1973). *Alimentación y obtención de alimentos en el Perú prehispánico.* Lima.

Isaacson, T., G. Ronen, et al. (2002). "Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of β-carotene and xanthophylls in plants." *The Plant Cell Online* 14(2): 333-342.

Itkin, M., H. Seybold, et al. (2009). "TOMATO AGAMOUS-LIKE 1 is a component of the fruit ripening regulatory network." *The Plant Journal* 60(6): 1081-1095.

Jarvis, D. I., A. H. D. Brown, et al. (2008). "From the Cover: A global perspective of the richness and evenness of traditional crop-variety diversity maintained by farming communities." *Proceedings of the National Academy of Sciences* 105(14): 5326-5331.

Jenkins, J. (1948). "The origin of the cultivated tomato." *Economic Botany* 2(4): 379-392.

Jimenez-Gomez, J. and J. Maloof (2009). "Sequence diversity in three tomato species: SNPs, markers, and molecular evolution." *BMC Plant Biology* 9(1): 85.

Kamenetzky, L., R. Asís, et al. (2010). "Genomic analysis of wild tomato introgressions determining metabolism- and yield-associated traits." *Plant Physiology* 152(4): 1772-1786.

Keurentjes, J. J. B. (2009). "Genetical metabolomics: closing in on phenotypes." *Current Opinion in Plant Biology* 12(2): 223-230.

Kramer, M. G. and K. Redenbaugh (1994). "Commercialization of a tomato with an antisense polygalacturonase gene: The FLAVR SAVR™ tomato story." *Euphytica* 79(3): 293-297.

Krieger, U., Z. B. Lippman, et al. (2010). "The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato." *Nature Genetics* 42(5): 459-463.

Ku, H. M., S. Doganlar, et al. (1999). "The genetic basis of pear-shaped tomato fruit." *Theoretical and Applied Genetics* 99(5): 844-850.

Kurlovich, B. S., S. I. Rep'ev, et al. (2000). "The significance of Vavilov's scientific expeditions and ideas for development and use of legume genetic resources." *Plant Genetic Resources* 124: 23-32.

Labate, J. A. and A. Baldo (2005). "Tomato SNP Discovery by EST Mining and Resequencing." *Molecular Breeding* 16(4): 343-349.

Labate, J. A., S. Grandillo, et al. (2007). Tomato. *Genome mapping and molecular breeding in plants.* C. Kole. NY, Springer Publishing. 5: 1-125.

Labate, J. A., L. D. Robertson, et al. (2009). "Multilocus sequence data reveal extensive departures from equilibrium in domesticated tomato (*Solanum lycopersicum* L.)." *Heredity* 103(3): 257-267.

Labate, J. A., S. M. Sheffer, et al. (2011). "Diversity and population structure in a geographic sample of tomato accessions." *Crop Science* 51(3): 1068-1079.

Lanahan, M. B., H. C. Yen, et al. (1994). "The never ripe mutation blocks ethylene perception in tomato." *The Plant Cell Online* 6(4): 521-530.

Lanfermeijer, F. C., J. Warmink, et al. (2005). "The products of the broken Tm-2 and the durable Tm-22 resistance genes from tomato differ in four amino acids." *Journal of Experimental Botany* 56(421): 2925-2933.

Langford, A. N. (1937). "The parasitism of *Cladosporium fulvum* cooke and the genetics of resistance to it." *Canadian Journal of Research*(15): 108-128.

Levin, I., N. Gilboa, et al. (2000). "*Fgr* , a major locus that modulates the fructose to glucose ratio in mature tomato fruits." *Theoretical and Applied Genetics* 100(2): 256-262.

Li, W. and R. T. Chetelat (2010). " A pollen factor linking inter- and intraspecific pollen rejection in tomato." *Science* 330(6012): 1827-1830.

Lieberman, M., O. Segev, et al. (2004). "The tomato homolog of the gene encoding UV-damaged DNA binding protein 1 (DDB1) underlined as the gene that causes the high pigment-1 mutant phenotype." *Theoretical and Applied Genetics* 108(8): 1574-1581.

Ling, H.-Q., P. Bauer, et al. (2002). "The tomato fer gene encoding a bHLH protein controls iron-uptake responses in roots." *Proceedings of the National Academy of Sciences* 99(21): 13938-13943.

Ling, H.-Q., G. Koch, et al. (1999). "Map-based cloning of chloronerva, a gene involved in iron uptake of higher plants encoding nicotianamine synthase." *Proceedings of the National Academy of Sciences* 96(12): 7098-7103.

Lippman, Z. and S. D. Tanksley (2001). "Dissecting the Genetic Pathway to Extreme Fruit Size in Tomato Using a Cross Between the Small-Fruited Wild Species *Lycopersicon pimpinellifolium* and *L. esculentum* var. Giant Heirloom." *Genetics* 158(1): 413-422.

Lippman, Z. B. and D. Zamir (2007). "Heterosis: revisiting the magic." *Trends in Genetics* 23(2): 60-66.

Liu, J., van Eck et al (2002). "A new class of regulatory genes underlying the cause of pear-shaped tomato fruit." *Proceedings of the National Academy of Sciences* 99(20): 13302-13306.

Lopes, J. S. and M. A. Beaumont (2010). "ABC: A useful Bayesian tool for the analysis of population data." *Infection, Genetics and Evolution* 10(6): 825-832.

Luckwill, L. C. (1943). "The genus *Lycopersicon*: an historical, biological, and taxonomical survey of the wild and cultivated tomatoes. ." *Aberdeen University Studies*(120): 1–44.

Manning, K., M. Tör, et al. (2006). "A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening." *Nature Genetics* 38(8): 948-952.

Martin, G., S. Brommonschenkel, et al. (1993). "Map-based cloning of a protein kinase gene conferring disease resistance in tomato." *Science* 262(5138): 1432-1436.

Mazzucato, A., R. Papa, et al. (2008). "Genetic diversity, structure and marker-trait associations in a collection of Italian tomato (Solanum lycopersicum L.) landraces." *Theoretical and Applied Genetics* 116(5): 657-669.

McCue, G. A. (1952). "The history of the use of the tomato: an annotated bibliography." *Annals of the Missouri Botanical Garden* 39(4): 289-348.

McKhann, H. I., C. Camilleri, et al. (2004). "Nested core collections maximizing genetic diversity in *Arabidopsis thaliana*." *The Plant Journal* 38(1): 193-202.

McMeekin, D. (1992). "Representations on pre-columbian spindle whorls of the floral and fruit structure of economic plants." *Economic Botany* 46(2): 171-180.

Menda, N., Y. Semel, et al. (2004). "In silico screening of a saturated mutation library of tomato." *The Plant Journal* 38(5): 861-872.

Mendel, G. (1866). Experiments in plant hybridization : original traduction from *Versuche über Plflanzenhybriden* by W. Bateson completed by R. Blumberg, Electronic Scholarly Publishing Project.

Miller, J. C. and S. D. Tanksley (1990). "RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*." *Theoretical and Applied Genetics* 80(4): 437-448.

Minoia, S., A. Petrozza, et al. (2010). "A new mutant genetic resource for tomato crop improvement by TILLING technology." *BMC Research Notes* 3(1): 69.

Monforte, A. J. and S. D. Tanksley (2000). "Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: A tool for gene mapping and gene discovery." *Genome* 43(5): 803-813.

Moreau, P., P. Thoquet, et al. (1998). "Genetic Mapping of Ph-2, a Single Locus Controlling Partial Resistance to Phytophthora infestans in Tomato." *Molecular Plant-Microbe Interactions* 11(4): 259-269.

Mounet, F., A. Moing, et al. (2009). " Gene and metabolite regulatory network of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development." *Plant Physiology* 149(3): 1505-1528.

Mueller, L. A., R. K. Lankhorst, et al. (2009). "A Snapshot of the emerging tomato genome sequence." *Plant Gen.* 2(1): 78-92.

Müller, C. H. (1940a). "A revision of the genus *Lycopersicon*. ." *U.S.D.A. Miscellanous Publications*(382): 1–28.

Müller, C. H. (1940b). "The taxonomy and distribution of the genus *Lycopersicon*." *National Horticulture Magazine*(19): 157–160.

Muños, S., N. Ranc, et al. (2011). "Increase in tomato locule number is controlled by two SNPs located near WUSCHEL." *Plant Physiology*. 156(4): 2244-2254.

Myles, S., J. Peiffer, et al. (2009). " Association mapping: critical considerations shift from genotyping to experimental design.." *The Plant Cell Online* 21(8): 2194-2202.

Nakazato, T. and E. A. Housworth (2011). "Spatial genetics of wild tomato species reveals roles of the Andean geography on demographic history." *American Journal of Botany* 98(1): 88-98.

Navez, B. (2011). *Tomate: Qualité et préférences* Centre Technique Interprofessionnel Fruits Legumes.

Nei, M., T. Maruyama, et al. (1975). " The bottleneck effect and genetic variability in populations." *Evolution* 29(1): 1-10.

Nesbitt, T. C. and S. D. Tanksley (2002). "Comparative sequencing in the genus *Lycopersicon*: implications for the evolution of fruit size in the domestication of cultivated tomatoes." *Genetics* 162(1): 365-379.

Nuez, F., J. Prohens, et al. (2004). "Relationships, origin, and diversity of Galápagos tomatoes: implications for the conservation of natural populations." *American Journal of Botany* 91(1): 86-99.

Okabe, Y., E. Asamizu, et al. (2011). "Tomato TILLING Technology: development of a reverse genetics tool for the efficient isolation of mutants from Micro-Tom mutant libraries." *Plant and Cell Physiology* 52(11): 1994-2005.

Ori, N., Y. Eshed, et al. (1997). "The I2C Family from the wilt disease resistance Locus I2 belongs to the nucleotide binding, Leucine-Rich Repeat superfamily of plant resistance genes." *The Plant Cell Online* 9(4): 521-532.

Paran, I., I. Goldman, et al. (1995). "Recombinant inbred lines for genetic mapping in tomato." *Theoretical and Applied Genetics* 90(3): 542-548.

Paran, I. and E. van der Knaap (2007). "Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper." *Journal of Experimental Botany* 58(14): 3841-3852.

Park, Y.-J., M. West, et al. (2004). "Evaluation of AFLPs for germplasm fingerprinting and assessment of genetic diversity in cultivars tomato (*Lycopersicon esculentum* L.)." *Genome* 47(3): 510-518.

Paterson, A., S. Damon, et al. (1991). "Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments." *Genetics* 127: 181-187.

Paterson, A. H., E. S. Lander, et al. (1988). "Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms." *Nature* 335: 721-726.

Peralta, I. E. and D. Spooner (2007). History, origin and early cultivation of tomato (*Solanaceae*). *Genetic improvement of Solanaceous crops*. M. K. Razdan and A. K. Mattoo. Enfield (NH), Science Publisher. 2: 1-24.

Peralta, I. E., D. Spooner, et al. (2007). "Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *Lycopersicon*; *Solanaceae*)." *Systematic Botany Monographs* 84.

Peralta, I. E. and D. M. Spooner (2001). "Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*)." *American Journal of botany* 88(10): 1888-1902.

Peralta, I. E. and D. M. Spooner (2005). Morphological characterization and relationships of wild tomatoes (*Solanum L.* Section *Lycopersicon*). *A Festschrift for William G. D'Arcy*. T. B. Croat, V. C. Hollowell and R. C. Keating, Missouri Botanical Garden Press. 104: 227-257.

Pertuzé, R. A., Y. Ji, et al. (2002). "Comparative linkage map of the *Solanum lycopersicoides* and *S. sitiens* genomes and their differentiation from tomato." *Genome* 45(6): 1003-1012.

Piron, F., M. Nicolaï, et al. (2010). "An Induced Mutation in Tomato eIF4E Leads to Immunity to Two potyviruses." *PLoS ONE* 5(6): e11313.

Price, A. L., N. A. Zaitlen, et al. (2010). "New approaches to population stratification in genome-wide association studies." *Nature Reviews Genetics* 11(7): 459-463.

Pritchard, J. K., M. Stephens, et al. (2000). "Inference of population structure ssing multilocus genotype data." *Genetics* 155(2): 945-959.

Purugganan, M. D. and D. Q. Fuller (2009). "The nature of selection during plant domestication." *Nature* 457(7231): 843-848.

Quinet, M., J.-M. Kinet, et al. (2011). "Flowering response of the uniflora:blind:self-pruning and jointless:uniflora:self-pruning tomato (*Solanum lycopersicum*) triple mutants." *Physiologia Plantarum* 141(2): 166-176.

Ramsay, L., J. Comadran, et al. (2011). "INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1." *Nature Genetics* 43(2): 169-172.

Ranc, N. (2010). Analyse du polymorphisme de gènes de composantes de la qualité des fruits dans les ressources génétiques sauvages et cultivées de tomate ; recherche d'association gènes/QTL. Montpellier, SupAgro. Thèse de doctorat: 348.

Ranc, N., S. Muños, et al. (2008). "A clarified position for *Solanum lycopersicum* var. cerasiforme in the evolutionary history of tomatoes (*Solanaceae*)." *BMC Plant Biology* 8(1): 130.

Rick (1977). "Conservation of tomato species germplasm. ." *California Agriculture* 31((9)): 32-33

Rick, C. M. (1991). "Tomato Paste: A concentrated review of genetic highlights from the beginnings to the advent of molecular." *Genetics* 128(1): 1-5.

Rick, C. M. and R. T. Chetelat (1995). "Utilization of related wild species for tomato improvement." *Acta Horticulturae* 412: 21-38.

Rick , C. M. and F. Fobes (1975). "Allozyme variation in the cultivated tomato and closely related species." *Bulletin of the Torrey Botanical Club* 102(6): 376-384.

Rick, C. M., J. F. Fobes, et al. (1977). "Genetic variation in *Lycopersicon pimpinellifolium*: Evidence of evolutionary change in mating systems." *Plant Systematics and Evolution* 127(2): 139-170.

Rick, C. M., J. F. Fobes, et al. (1979). "Evolution of mating systems in *Lycopersicon hirsutum* as deduced from genetic variation in electrophoretic and morphological characters." *Plant Systematics and Evolution* 132(4): 279-298.

Rick , C. M. and M. Holle (1990). "Andean *Lycopersicum esculentum* var. *cerasiforme*: Genetic Variation and Its Evolutionary Significance." *Economic Botany* 44: 69-78.

Rigola, D., J. van Oeveren, et al. (2009). "High-throughput detection of induced mutations and natural variation using KeyPoint™ technology." *PLoS ONE* 4(3): 4761.

Robbins, M. D., S. C. Sim, et al. (2010). "Mapping and linkage disequilibrium analysis with a genome-wide collection of SNPs that detect polymorphism in cultivated tomato." *Journal of Experimental Botany* 62(6): 1831-1845.

Rodriguez, G. R., S. Munos, et al. (2011). "Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity." *Plant Physiology* 156(1): 275-285.

Ronen, G., L. Carmel-Goren, et al. (2000). "An alternative pathway to β-carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato." *Proceedings of the National Academy of Sciences* 97(20): 11102-11107.

Rousseaux, M., C. Jones, et al. (2005). "QTL analysis of fruit antioxidants in tomato using *Lycopersicon pennellii* introgression lines." *Theoretical and Applied Genetics* 111(7): 1396-1408.

Sahagún, B. d. (1988). *Historia General de las Cosas de Nueva España*. Madrid.

Saliba-Colombani, V., M. Causse, et al. (2001). "Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits." *Theoretical and Applied Genetics* V102(2): 259-272.

Sanders, R. A. and W. Hiatt (2005). "Tomato transgene structure and silencing." *Nature Biotechnology* 23(3): 287-289.

Schauer, N., D. Zamir and A.R. Fernie (2004). "Metabolic profiling of leaves and fruit of wild species tomato: a survey of the Solanum lycopersicum complex." Journal of Experimental Botany 56(410): 297-307.

Schoen, D. and A. Brown (1993). " Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers." *Procceedings of the National Academy of Sciences* 90(22): 10623-10627.

Semel, Y., Nissenbaume et al (2006). "From the Cover: Overdominant quantitative trait loci for yield and fitness in tomato." *Proceedings of the National Academy of Sciences* 103(35): 12981-12986.

Shalit, A., A. Rozman, et al. (2009). "The flowering hormone florigen functions as a general systemic regulator of growth and termination." *Proceedings of the National Academy of Sciences* 106(20): 8392-8397.

Sharma, M., N. Singh, et al. (2008). "Expression of toxin co-regulated pilus subunit A (TCPA) of Vibrio cholerae; and its immunogenic epitopes fused to cholera toxin B subunit in transgenic tomato (*Solanum lycopersicum*)." *Plant Cell Reports* 27(2): 307-318.

Sicard, A. and M. Lenhard (2011). "The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants." *Annals of Botany* 107(9): 1433-1443.

Smith, A. F. (2000). *Souper tomatoes: the story of America's favorite food*, Rutgers University Press.

SolCAP (2008). USDA-CSREES awards $5.4 million to SolCAP *SolCAP News* . Michigan State University, East Lansing.

Stadler, T., A. M. Florez-Rueda, et al. (2012). "Testing for "Snowballing" hybrid incompatibilities in *Solanum*: impact of ancestral polymorphism and divergence estimates." *Molecular Biology and Evolution.* 29(1): 31-34

Stevens, M. A. and C. M. Rick. (1986). Genetics and breeding. *The Tomato Crop: A Scientific Basis for Improvement.* J. G. A. a. J. Rudich. London, Chapman and Hall: 35-109.

Stevens, R., M. Buret, et al. (2007). "Candidate genes and quantitative trait loci affecting fruit ascorbic acid content in three tomato populations." *Plant Physiology* 143(4): 1943-1953.

Szymkowiak, E. J. and E. E. Irish (2005). "JOINTLESS suppresses sympodial identity in inflorescence meristems of tomato." *Planta* 223(4): 646-658.

Tadmor, Y., E. Fridman, et al. (2002). "Identification of malodorous, a wild species allele affecting tomato aroma that was selected against during domestication." *Journal of Agricultural and Food Chemistry* 50(7): 2005-2009.

Tam, S. M., C. Mhiri, et al. (2005). "Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR." *Theoretical and Applied Genetics* 110(5): 819-831.

Tanksley, D., H. Medina-Filho, et al. (1981). "The effect of isozyme selection on metric characters in an interspecific backcross of tomato — basis of an early screening procedure." *Theoretical and Applied Genetics* 60(5): 291-296.

Tanksley, S. D., M. W. Ganal, et al. (1992). "High density molecular linkage maps of the tomato and potato genomes." *Genetics* 132(4): 1141-1160.

Tanksley, S. D., S. Grandillo, et al. (1996). "Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*." *Theoretical and Applied Genetics* 92(2): 213-224.

Tanksley, S. D. and G. S. Khush (2002). Charles Madera Rick. *Bibliographical Memoirs* 84: 307-319.

Tellier, A., I. Fischer, et al. (2011). "Fitness effects of derived deleterious mutations in four closely related wild tomato species with spatial structure." *Heredity* 107(3): 189-199.

Tellier, A., S. J. Y. Laurent, et al. (2011). "Inference of seed bank parameters in two wild tomato species using ecological and genetic data." *Proceedings of the National Academy of Sciences* 108(41): 17052-17057.

Thomas, M., J. Dawson, et al. (2011). "Seed exchanges, a key to analyze crop diversity dynamics in farmer-led on-farm conservation." *Genetic Resources and Crop Evolution* 58(3): 321-338.

Tieman, D., M. Taylor, et al. (2006). "Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde." *Proceedings of the National Academy of Sciences* 103(21): 8287-8292.

Tikunov, Y. A. Lommen , et al. (2005). "A Novel Approach for Nontargeted Data Analysis for Metabolomics. Large-Scale Profiling of Tomato Fruit Volatiles." Plant Physiology 139(3): 1125-1137.

van Berloo, R., A. Zhu, et al. (2008). "Diversity and linkage disequilibrium analysis within a selected set of cultivated tomatoes." *Theoretical and Applied Genetics* 117(1): 89-101.

van der Knaap, E., Z. B. Lippman, et al. (2002). "Extremely elongated tomato fruit controlled by four quantitative trait loci with epistatic interactions." *Theoretical and Applied Genetics* 104(2): 241-247.

van der Knaap, E., A. Sanyal, et al. (2004). "High-Resolution Fine Mapping and Fluorescence in Situ Hybridization Analysis of sun, a Locus Controlling Tomato Fruit Shape, Reveals a Region of the Tomato Genome Prone to DNA Rearrangements." *Genetics* 168(4): 2127-2140.

Van Deynze, A., K. Stoffel, et al. (2007). "Diversity in conserved genes in tomato." *BMC Genomics* 8(1): 465.

Villand, J., P. W. Skroch, et al. (1998). "Genetic Variation among tomato accessions from primary and secondary centers of diversity." *Crop Science* 38: 1339-1347.

Vrebalov, J., D. Ruezinsky, et al. (2002). "A MADS-Box gene necessary for fruit ripening at the tomato ripening-inhibitor (Rin) locus." *Science* 296(5566): 343-346.

Xiao, H., N. Jiang, et al. (2008). "A Retrotransposon-Mediated Gene Duplication Underlies Morphological Variation of Tomato Fruit." *Science* 319(5869): 1527-1530.

Yakovleff, E. a. F. L. H. (1935). "El mundo vegetal de los antiguos peruanos. ." *Revista del Museo Nacional Botánica etnológica* 3(3): 3-55.

Yang, W., X. Bai, et al. (2004). "Discovery of single nucleotide polymorphisms in Lycopersicon esculentum by computer aided analysis of expressed sequence tags." *Molecular Breeding* 14(1): 21-34.

Yeager, A. (1937). "Studies on the inheritance and development of fruit size and shape in the tomato." *Journal of Agricultural Research* 55: 141-152.

Yen, H. C., S. Lee, et al. (1995). "The Tomato Never-ripe Locus Regulates Ethylene-Inducible Gene Expression and Is Linked to a Homolog of the Arabidopsis ETR1 Gene." *Plant Physiology* 107(4): 1343-1353.

Yen, H. C., B. A. Shelton, et al. (1997). "The tomato high-pigment (hp) locus maps to chromosome 2 and influences plastome copy number and fruit quality." *Theoretical and Applied Genetics* 95(7): 1069-1079.

Yu, J., G. Pressoir, et al. (2005). "A unified mixed-model method for association mapping that accounts for multiple levels of relatedness." *Nature Genetics* 38(2): 203-208.

Zamir, D. (2001). "Improving plant breeding with exotic genetic libraries." *Nature Review Genetics* 2(12): 983 - 989.

Zamir, D. (2008). "Plant breeders go back to nature." *Nature Genetics* 40(3): 269-270.

Zamir, D., I. Ekstein-Michelson, et al. (1994). "Mapping and introgression of a tomato yellow leaf curl virus tolerance gene, *TY-1*." *Theoretical and Applied Genetics* 88(2): 141-146.

Zanor, M. I., J.-L. Rambla, et al. (2009). "Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents." *Journal of Experimental Botany* 60(7): 2139-2154.

Zhang, X., N. A. Buehner, et al. (2006). "Tomato is a highly effective vehicle for expression and oral immunization with Norwalk virus capsid protein." *Plant Biotechnology Journal* 4(4): 419-432.

Zhong, X. B., J. Bodeau, et al. (1999). "FISH to meiotic pachytene chromosomes of tomato locates the root-knot nematode resistance gene Mi-1 and the acid phosphatase gene Aps-1 near the junction of euchromatin and pericentromeric heterochromatin of chromosome arms 6S and 6L, respectively." *Theoretical and Applied Genetics* 98(3): 365-370.

Zuriaga, E., J. M. Blanca, et al. (2009). "Genetic and bioclimatic variation in *Solanum pimpinellifolium*." *Genetic Resources and Crop Evolution* 56: 39-51.