Genetic analysis of fruit quality in tomato

Raana Roohanitaziani

Thesis committee

Promotor

Prof. Dr R. G. F. Visser

Professor of Plant Breeding

Wageningen University & Research

Co-Promotor

Dr A.G. Bovy

Senior scientist, Plant Breeding

Wageningen University & Research

Other members

Dr R. Mumm, Wageningen University & Research

Prof. Dr E.J. Woltering, Wageningen University & Research

Prof. Dr J.J.B. Keurentjes, Wageningen University & Research

Dr W. Verkerke, Wageningen University & Research

Genetic analysis of fruit quality in tomato

Raana Roohanitaziani

Thesis

submitted in fulfilment of the requirements for the degree of doctor at Wageningen University by the authority of the Rector Magnificus,

prof. Dr A.P.J. Mol,

in the present of the

Thesis committee appointed by the Academic Board

to be defended in public

on Thursday 31 October 2019

at 11 a.m. in the Aula.

Raana Roohanitaziani

Genetic analysis of fruit quality in tomato

Table of content

Chapter 1

General Introduction

1) Tomato taxonomy and botany

The cultivated tomato, *Solanum lycopersicum* L., belongs to the Solanaceae family, containing more than 3000 species including many plants of economic importance such as potato, eggplant, petunia, tobacco and pepper. Solanum is the largest genus in the Solanaceae family, encompassing 1250 to 1700 species (Bergougnoux, 2014). The phylogenetic classification of the Solanaceae has been recently revised and the genus Lycopersicon re-integrated into the Solanum genus with its new nomenclature. Solanum section Lycopersicon consists of the cultivated tomato, *Solanum lycopersicum*, and 12 wild species: *S. arcanum, S. cheesmaniae, S. chilense, S. chmielewskii, S. corneliomulleri, S. galapagense, S. habrochaites, S. huaylasense, S. neorickii, S. pennellii, S. peruvianum* and *S. pimpinellifolium* (Peralta et al., 2008; Knapp & Peralta, 2016;). All members of sect. Lycopersicon are diploid (2n = 24) (Kimura & Sinha, 2008).

Botanically, tomatoes are fruits. As a true fruit, it develops from the ovary of the plant after fertilization, and contains seeds but they are commonly used as a vegetable ingredient and they are served as part of a salad or main course of a meal, rather than as a dessert. The *S. lycopersicum* is self-compatible but some wild relatives (*S. chilense* and *S. peruvianum*) have a self-incompatibility mechanism (Peralta et al., 2008). The flowers contain functional male (anthers) and female (pistil) parts. They are small, yellow and star shaped although there are some morphological differences regarding the flower parts (sepal, petal, stamen and pistil) between different species of tomatoes (Figure 1). They get together on the same peduncle to form trusses of fruits that can vary in number. Grown under optimal conditions, tomatoes have a 95 to 115 day lifecycle. The first flowers will appear and open 7 to 8 weeks after sowing and mature fruits will follow 6 to 8 weeks later. Normally at least 4 to 8 flowers are borne on each inflorescence, and a single tomato plant may produce as many as 20 or more inflorescences over a season (Fentik, 2017).

Figure 1. Diversity in tomato flowers. Pictures are derived from the plants used in this thesis.

The fruit develops from the ovary after fertilization of the ovules. The walls of the ovary become the pericarp which comprises the fruit flesh. The pericarp surrounds hollow spaces full of seeds and moisture,

called locular cavities. These vary from two to up to 10 or more among fruits and affect fruit shape and size (Muños et al., 2011) (Figure 2). Cultivated varieties produce fruit with more locules, resulting in larger, wider fruits. Tomato fruits are highly diverse in shape and they can be classified into 8 fruit shape categories: flat, rectangular, ellipsoid, obovoid, round, oxheart, long and heart (Figure 3). Tomato fruit color varies from green and yellow to orange, red and purple. In most cases, the color of the fruit is determined by the quantities of yellow to red carotenoids, such as lycopene and β-carotene, the green chlorophyll, the yellow flavonoid naringenin chalcone and purple anthocyanins.

Figure 2. Locule number variation in tomato fruit. The number of locules in the picture varies from 2 to 10. Pictures are derived from the plants used in this thesis.
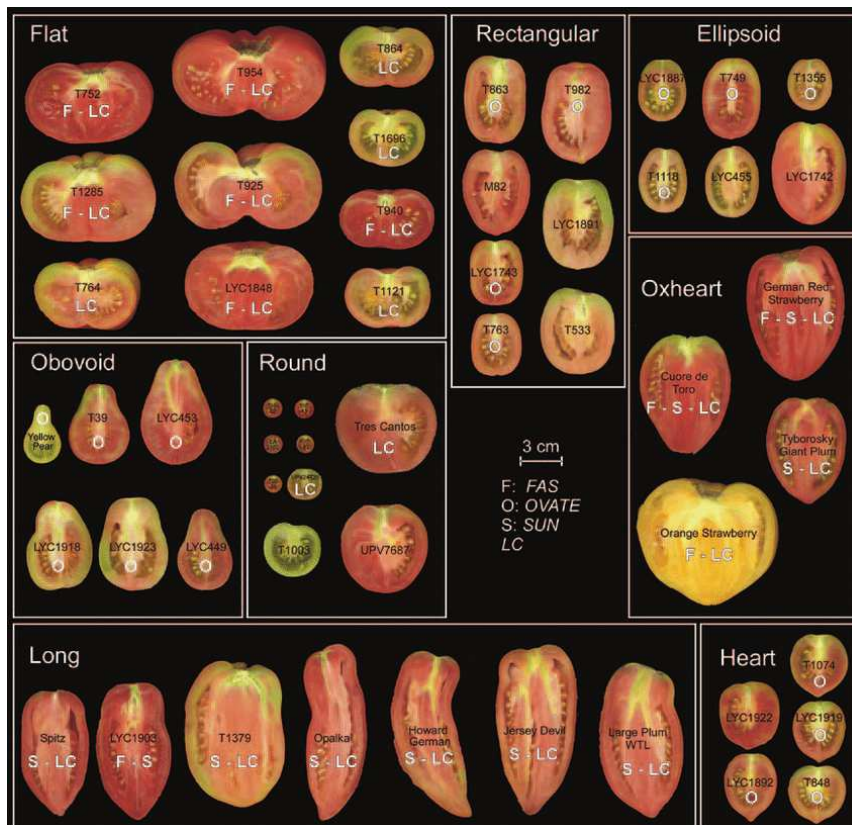

Figure 3. Diversity in tomato fruit shapes (Rodriguez et al., 2011).

The tomato plant has compound leaves. A compound leaf is made up of leaflets which are distributed along the leaf rachis. Some of the leaflets on this leaf are compound as well. Generally there are two basic tomato

leaf types, Regular (typical tomato leaf type) and Potato Leaf. There are also variations of both in terms of width/length of leaf, leaf color and various shades of green (Figure 4). Tomato has sympodial shoot development. The primary vegetative shoot terminates in a flower after the development of 8 to 12 leaves. Subsequently, new vegetative shoots arise from the axillary bud just below the terminating inflorescence. This new shoot, in turn, terminates again after making three leaves, and the next shoot arises from the newly formed axillary bud. This cycle is repeated continuously to form sympodial shoots. By definition, tomato shoots are considered to be "determinate" because each shoot terminates in a flower. However, the wild-type growth habit of tomatoes is classified as "indeterminate" because they continuously produce sympodial units.

Figure 4. Variation in tomato leaf type, as observed with the genotypes studied in this thesis.

2) History and domestication

The wild tomato species are native to the Andean region now encompassed by parts of Chile, Bolivia, Ecuador, Colombia and Peru. The time and place of domestication of tomato is not known with certainty (Peralta & Spooner, 2007). However, tomato had reached a fairly advanced stage of domestication before being taken to Europe in the 15th century and further domestication on a much more intense level occurred throughout Europe in the 18th and 19th centuries (Bai & Lindhout, 2007). Domestication in tomato has triggered a range of morphological and physiological traits that distinguish domesticated crops from their wild ancestors. These traits include growth habits (a more compact growth habit, increased earliness,

reduction/loss of seed dispersal and dormancy) and fruit traits (fruit size, shape and color) (Frary & Doganlar, 2003; Bai & Lindhout, 2007). One of the most dramatic changes in tomato because of domestication is in fruit size. Wild tomato species have tiny fruits that weigh only a few grams each with the purpose to propagate the species and not to feed humans, whereas the modern cultivated tomatoes offer a large variation in fruit size, ranging from cherry to beef tomatoes (Bergougnoux, 2014). Wild tomatoes can grow in many environments and they exhibit a great difference in morphological characters, mating systems, disease susceptibility and stress resistance (Peralta et al., 2008). This is why wild tomatoes are considered to be a valuable source of desirable traits, such as resistance to pathogens and abiotic stresses, to incorporate in modern cultivars. Compared with the rich reservoir in wild species, the cultivated tomato is genetically poor and lack the diversity. It is estimated that the genomes of tomato cultivars contain five percent of the genetic variation of their wild relatives (Miller & Tanksley, 1990). Despite their relatively narrow genetic base, phenotypic variation is large in cultivated tomato, especially in old traditional cultivated tomatoes called heirloom or vintage tomatoes (Barry & Pandey, 2009). These heirloom tomato varieties exhibit a wide range of morphological variation and also exhibit wide diversity in their composition of aroma volatiles (Bennett, 2012). Today, there is an interest to use wild and old accessions maintained in germplasm collections to improve modern varieties of tomato. Most efforts have so far been focussed on the introgression of disease resistance genes and genes involved in abiotic stress tolerance, but this gene pool can be used for many other traits as well (Rick, 1976; Tanksley & McCouch, 1997; Bai & Lindhout, 2007; Aflitos et al., 2014).

3)Breeding for fruit quality traits

The breeding objective for tomato fruit quality includes external characteristics such as size, shape and color, as well as internal qualities like flavour (taste and aroma), texture (firmness, mealiness, juiciness), and nutritional values. These factors have a large influence on the consumer liking and acceptance of the fruits. Visual appearance (size, shape, and color) influences the initial consumer's choice, but after that the eating quality becomes the major influencing factor. Some of the fruit quality attributes such as shape, size and firmness have been the goals of breeding for a long time. But flavour and nutritional compounds have gained attention only recently. In general, breeding goals in tomato have gone through four phases: breeding for yield in the 1970s, for shelf-life in the 1980s, for taste in the 1990s and recently for nutritional value (Bai & Lindhout, 2007).

Today, there is a high demand for improvement of tomato fruit flavour and dissatisfaction of consumers regarding tomato flavour is the main consumer complaint. Several studies suggest that the taste of modern tomato cultivars is not as good as the taste of traditional "vintage" or "heirloom" tomatoes. For instance, two recent GWAS studies with tomato collections including wild relatives, old varieties and

modern cultivars revealed that modern cultivars contain significantly lower amounts of important fruit flavour metabolites, including several volatiles and sugars (Bauchet et al., 2017; D. Tieman et al., 2017). Although in one of these studies (Tieman et al., 2017) the focus has mainly been on the round-type modern cultivars grown in the field and modern cherry tomatoes have not been included in the trial. The decline in flavour of modern cultivars is mainly assigned to the intensive breeding for yield and underlying traits such as disease resistance and fruit size and the secondary importance of selection for flavour and nutritional traits in the existing conventional breeding programmes (Klee & Tieman, 2013). Other reasons for poor flavour of modern cultivars is related to what has been done for improvement of fruit shelf-life via both breeding (selection of long shelf-life genotypes, use of ripening mutants and the introgression of genes that delay ripening) and harvest and postharvest handling (harvesting of fruits at an immature stage, post-harvest physical and chemical treatments) (Boukobza & Taylor, 2002).

To satisfy consumer expectations, tomato breeders are now pursuing flavour improvement as one of their major breeding objectives, although the complexity of this trait has made its improvement a difficult task. Many metabolites, derived from different metabolic pathways, contribute to fruit flavour. Each of these metabolites is controlled by one or more QTLs and changing one of these QTLs may not have a major effect on the perception of taste and aroma (Klee & Tieman, 2013). The second issue complicating tomato flavour improvement is the wide range of consumer preferences, which is influenced by personal, regional and cultural differences (Rambla et al., 2014). Another challenge for improving flavour is putting high flavour in the context of high yield and postharvest shelf life, properties that are essential to the growers and usually have negative correlation with fruit flavour (Klee & Tieman, 2013). Nevertheless, the availability of a tomato genome sequence and recent breakthroughs in the sensory quality analysis, have led to a major progress in our understanding of flavour metabolite pathways and their underlying genetic regulation. To improve tomato flavour several aspects first need to be defined such as the most important metabolites which contribute to flavour and consumer liking, the pathway of the synthesis of these metabolites and finally the genomic region and genes controlling their synthesis. When all of this information is available improvement can be implemented in a series of breeding programs.

3.1) What determines tomato flavour?

The flavour of tomato fruit, is determined by a combination of taste, aroma and texture. Fruit taste is determined by non-volatile primary metabolites, such as sugars, organic acids and free amino acids, as well as salts. Sugar: acid ratio is pointed out as a major determinant of tomato flavour and a desired tomato should have a balanced level of sugars and organic acids. Both, the sugar and acid contents are important traits for breeding and they have been a frequent target for tomato breeders. Fruits of cultivated

tomato contain mainly glucose and fructose and only trace amounts of sucrose (Jones & Scott, 1983; Petro-Turza, 1986). Citric and malic acids are the major organic acids in tomato fruit (Yilmaz, 2001). At all stages citric acid is the dominant organic acid but unripe green tomatoes may contain significant amounts of malic acid while its content in ripe fruits is fairly low (Agius et al., 2018).

Free amino acids form about 2-2.5% of the total dry matter of tomatoes. Glutamic acid, γ-aminobutyric acid, glutamine, and aspartic acid comprise about 80% of the total free amino acids in tomatoes (Yilmaz, 2001).

3.1.1) Glutamate

Glutamic acid is the major free amino acid found in tomato juice which comprises up to 45% of the total weight of free amino acids in fresh tomato juice. Tomato fruits contain a higher concentration of free glutamate compared to many other fruits and vegetables (Jinap & Hajeb, 2010). This high level of free glutamate in tomato fruits provides the characteristic "umami" taste. Umami, in addition to sweet, salt, bitter and sour, is the fifth basic taste quality that humans can detect. Free amino acids have been found as the main umami ingredients in many vegetables, however glutamate is supposed to be the main contributor to umami taste in tomato (Sorrequieta et al., 1998). Umami taste gives meaty and savoury flavours to foods. Monosodium glutamate (MSG) the sodium salt of glutamic acid, is used as a flavour enhancer in both industrial made foods and home cooking, in Western as well as Eastern countries, although the consumption of MSG in eastern countries still prevails over that in western countries (Bellisle, 1999; Bellisle, 2008).

3.1.2) Volatile compounds: Phenylalanine derived volatiles

Volatile compounds are responsible for the aroma and flavour perception of the fruits and, although over 400 aroma volatiles have been identified in tomato fruit, only about 30 of them are considered to be important for flavour based on their odour thresholds. The majority of flavour-related volatiles in tomato are derived from amino acids, such as phenylalanine, leucine and isoleucine, lipids and carotenoids (Buttery & Ling, 1993; Yilmaz, 2001; Rambla et al., 2014).

Several phenylalanine-derived volatiles, also called phenolic volatiles, have been shown to affect tomato fruit flavour either positively or negatively. Among these phenolic compounds, 2-phenylacetaldehyde and 2-phenylethanol are considered the most important volatiles for tomato fruit aroma (Baldwin et al., 2000; Tieman et al., 2007; Tzin et al., 2013; Rambla et al., 2014). Both volatiles have fruity/floral properties and they are major constituents of scent in many flowers. 2-phenylethanol is the major aroma volatile contributing to the scent of roses and this compound is widely used as fragrance

in the cosmetic and food industry (Knudsen et al., 1993; Tieman et al., 2007). Despite their positive association with consumer liking and flavour intensity in tomato fruit, 2-phenylethanol and 2-phenylacetaldehyde can also be perceived as an undesirable flavour at extremely elevated concentrations, as was shown for the *S. pennellii* introgression line IL8.2 in the cultivated M82 tomato background (Tadmor et al., 2002; D Tieman et al., 2017).

3.2) Fruit Shelf-life

Fruit shelf-life is determined by a series of physiological, biochemical and organoleptic changes that happen during ripening to make fruit edible and desirable for consumers. These changes generally include softening of the flesh by modification of the cell wall structure, alterations in pigment biosynthesis, increase in the levels of aromatic volatiles and nutrition, alteration in cuticle architecture and composition and increase in susceptibility to post-harvest pathogens. Decrease in the fruit firmness, resulting in fruit softening, is the result of dissolution of the middle lamella and structural changes in cell wall polysaccharides in an ordered series of modifications by cell wall degrading enzymes (Giovannoni, 2001; Brummell, 2006). The tomato cell wall consists of rigid, inextensible cellulose micro fibrils held together by interpenetrating coextensive networks of matrix glycans, pectins and structural glycoproteins. Many studies have been carried out so far to understand the molecular basis of fruit softening and texture deterioration. In tomato several cell wall degrading enzymes have already been identified and reverse genetics strategies to suppress their expression or create functional knock outs using CRISPR-CAS9 gene editing has been considered as an approach for improving the fruit shelf-life. Despite all of these efforts the molecular basis of fruit texture or in particular fruit cell wall metabolism is not well understood and the precise mechanism of softening remains unknown.

Post-harvest shelf life of tomato has been since long an important trait for breeders and growers, but the importance of shelf-life improvement in modern agriculture is more perceivable because of the globalization of markets and long distances between consumers and producers. Picking fruits at early stages of ripening and use of ripening mutants have been the early solutions for enhancement of fruit shelf-life. Today, various physical and chemical postharvest treatments are used to increase the fruit shelf-life via slowing down the physiological processes of senescence and maturation. However, each of these treatments has its own limitations and disadvantages (Mahajan et al., 2014; Sandarani et al., 2018). Looking for varieties with extended shelf life, which are not ripening mutants but remain firm for an extended period of time after ripening and allow harvesting at a more advanced colour stage is another breeding strategy for shelf-life improvement which recently gained more attention.

4) Molecular breeding

The basis of plant breeding is selection. In classical plant breeding selection of specific plants with desirable traits is based on measurable or visible traits, called phenotype. Selection based on phenotype is slow and influenced by the interaction of genotype with the environment. Discovery of DNA allowed the development of a new type of markers based exclusively on the intrinsic characteristics of an individual genome. Each DNA region (locus) characterised by multiple forms within a population and inherited in a Mendelian fashion can be considered as a molecular marker, which allows the grouping of a set of individuals into as many subgroups as the number of forms (alleles) identified at that locus (Marwal et al., 2014). The first plant DNA markers were based on restriction fragment length polymorphisms (RFLPs). RFLP techniques were inherently challenging and time consuming, and were eventually replaced by less complex, more cost-effective PCR-based markers such as RAPD, CAPS, AFLP and SSR. The improvement of Sanger sequencing in the 1990's, in combination with the start of genome and expressed sequence tag (EST) sequencing programs in model plant species, led to the acceleration in the identification of variation at the single base pair resolution (Deschamps et al., 2012). Today, the most widely used molecular marker technique is based on the analysis of single nucleotide polymorphisms (SNPs) naturally occurring in each organism, which represent the most abundant source of variation within each genome. Markers based on SNPs have rapidly gained the centre stage of molecular genetics during the recent years due to their abundance in the genome and the relative ease in determining their frequency in a cost-effective and high-throughput detection format and platform.

The advent of next-generation DNA sequencing (NGS) technologies has led to the development of rapid genome-wide SNP detection applications in various plant species. Recent improvements in sequencing methods combined with an overall decrease in sequencing costs has increased the information about potentially millions of genome-wide SNPs or small insertion-deletions and their surrounding sequence context which has set the foundation of high-throughput genotyping known as genotyping-by-sequencing (Deschamps et al., 2012). The Availability of this large amount of SNPs and the decrease in sequencing costs has encouraged breeders to re-consider germplasm collections as a valuable trait reservoir and to use biodiversity-based breeding for trait and crop improvement (Lin et al., 2014; Tomato Genome, 2012). For instance, recently hundreds of tomato accessions have been (re) sequenced (Causse et al., 2013; Aflitos et al., 2014; Lin et al., 2014), and they provide an untapped resource of promising genetic variation for mapping and cloning of important domestication and several other agronomic traits which suitably can be deployed for the development of improved cultivars.

Today selection in breeding programs is based on molecular markers through marker assisted selection (MAS). MAS provides prediction of phenotype from genotype, one of the goals in plant breeding,

by use of molecular markers closely linked to a trait of interest (Mammadov et al., 2012). One of the most important functions of MAS in plant sciences is in marker trait association studies. A capability which enabled moving from traditional breeding into the well-known molecular breeding era.

4.1) Marker-trait association study

As mentioned above molecular markers can reveal polymorphisms in the nucleotide sequence at a given locus and this allows to confirm identity between parents and progeny, to determine genetic distances and to construct genetic and physical maps and eventually to localise genes or genomic regions responsible for the expression of a trait of interest. The phenotypic diversity which we observe for quantitative traits such as quality components, is mainly due to the underlying genetic complexity derived from multiple loci interacting together. A genomic region containing one or more genes affecting a quantitative trait is called quantitative trait locus (QTL). QTL mapping relies on finding an association between a genetic marker and the phenotype of interest. This requires the availability of either a genetic linkage map, in which the molecular markers are grouped and ordered according to their recombination frequencies within the studied population, or a physical map in which the order of the markers is based on their physical distances. To find the molecular markers associated with a locus controlling important phenotypic traits we need to develop large mapping populations for applications such as linkage mapping, introgression mapping or association mapping (Figure 5). The segregating populations that are used for linkage mapping can be F2, recombinant inbred lines (RIL), backcross inbred lines (BILs) or doubled haploid lines (DHLs) (Borrelli et al., 2009). These segregating populations are used to develop linkage maps in which the relative position of hundreds of molecular markers is indicated on each chromosome. Then, the gene(s) and QTL(s) position can be determined through the combination of phenotypic characteristics and segregation data of molecular markers. Figure 5 shows a summary of possible strategies for the development of molecular markers linked to gene/s controlling traits of interest. Isolating the gene underlying a QTL is the final goal for the understanding of the molecular basis of the trait and for development of precise molecular markers to be used in breeding programmes.
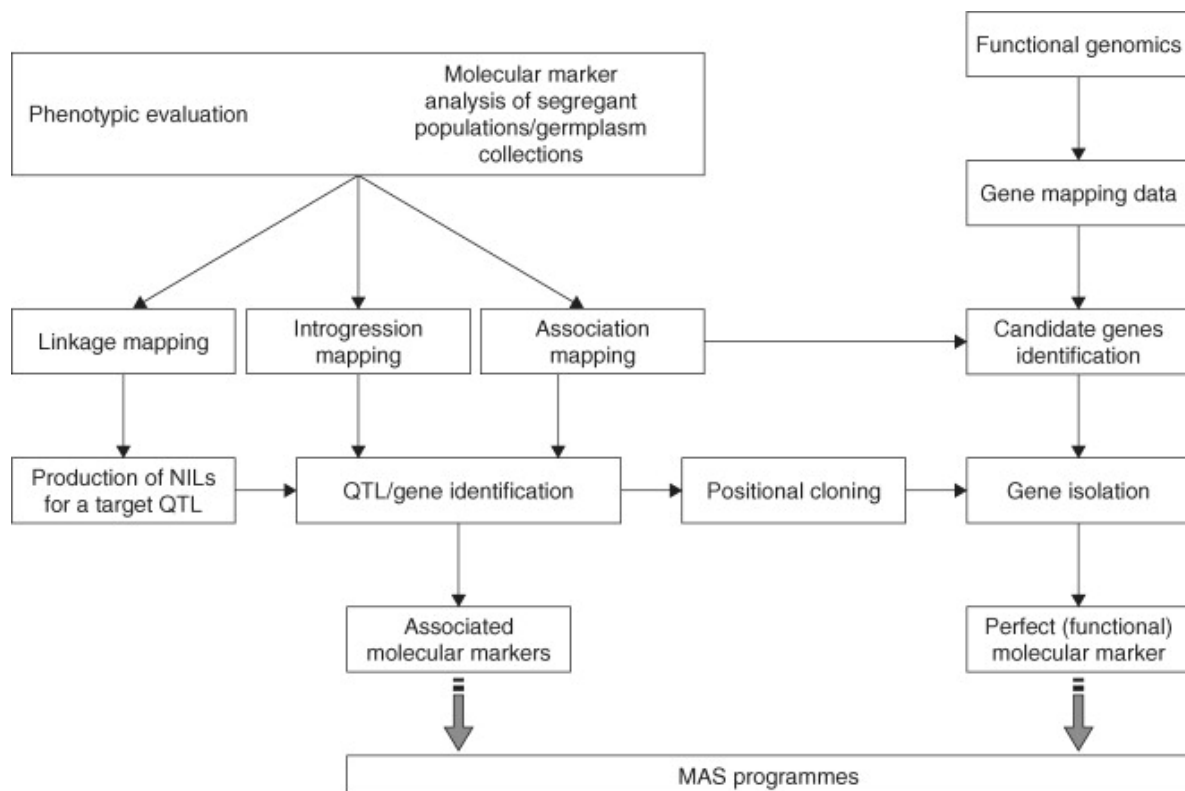
Figure 5. Summary of possible strategies for the development of molecular markers linked to gene controlling trait of interest in marker-trait association studies (Borrelli et al., 2009).

5) Metabolomics: phenotyping approach for quality attributes

The success of genetic analysis of quality traits is largely dependent on reliable and affordable phenotyping methods. Metabolomics approaches based on chromatographic separation techniques connected with mass spectrometry (MS) have been widely used for identification of tomato metabolites. The compounds present in the metabolome of tomato can be targeted by different analytical approaches. These approaches can concentrate on a single class of metabolites like sugars or amino acids, or can be used for a broad untargeted profiling aimed at detecting as many as possible compounds in a given extract. In this study we used three different platforms to perform metabolic profiling of different samples: 1) Gas chromatography(GC) coupled to electron impact time of flight (TOF)-MS. GC-TOF-MS can be used to detect primary metabolites, including organic acids, amino acids and sugars. 2) Solid phase micro extraction (SPME)-GC-MS for detection of volatile compounds. 3) Dionex-HPLC for measurement of cell wall monosaccharide content.

Aim and outline of the thesis project

In this project, we utilised two sources of genetic variation: (1) segregating populations based on cultivated tomato germplasm and (2) a tomato core collection of 122 tomato accessions that consist of wild relatives, land races and old cultivars. The first population was used to fine map and isolate candidate key genes underlying flavour QTL's and the second population was phenotyped and genotyped for several fruit quality and yield related traits with the main focus on fruit shelf-life in order to introduce these novel sources into future breeding programs.

In chapter 2 a tomato core collection consisting of 122 accessions was explored for phenotypic and genotypic variation in several plant growth, yield and fruit quality traits. Plants were phenotyped for the crop growth-related traits abscission zone, inflorescence branching, vegetative outgrowth of the inflorescence, growth rate and earliness of flowering and the fruit characteristics number of ripe fruits per genotype, fruit weight, firmness, total soluble solid content (Brix), colour and shape. Genotyping was also performed for known mutations or variants affecting plant architecture (*self-pruning* (*sp*), *compound inflorescence* (*s*), *jointless* (*j*)-and *potato leaf* (*c*)), fruit size and shape (*fruit weight 2.2* (*fw2.2*), *fruit weight 3.2* (*fw3.2*), *fruit weight 11.3* (*fw11.3*), *locule number* (*lc*), *fasciated* (*fas*), *ovate* and *sun*) and fruit colour (yellow flesh mutation (*r*), *tangerine* (t^{3183}), *old-gold-crimson* (*og*), *green flesh* allele (gf^4) and yellow mutation (*y*), which leads to a transparent epidermis and a pink fruit colour).

In chapter 3 we aimed to evaluate a tomato collection consisting of 93 re-sequenced land races and old cultivars for their fruit post-harvest shelf-life, to identify and introduce novel sources of post-harvest shelf-life into future breeding programs. Shelf-life attributes, such as firmness, water loss and colour pigments were monitored during a 42 days storage period. We also characterized and compared changes in cell wall sugar composition, primary metabolites and volatile compounds of some contrasting lines with long and short post-harvest shelf-life.

In chapter 4 a QTL for 2-phenylethanol at the bottom of CH4 was fine mapped. This QTL was previously found in a RIL population derived from a cross of a cherry tomato line, which exhibited a high overall flavour intensity, and a round fruited line with a common taste. The QTL fine mapping was performed using an F2 population consisting of 5000 individual plants derived from a cross between two RILs with contrasting flavour characteristics including the 2-phenylethanol QTL. Using SNP markers and metabolic analysis by GC-MS this QTL was narrowed down to a region of 110 kb from 54.52 to 54.63 Mb harbouring 11 candidate genes. We identified Solyc04g063350 (*Dcx1*) which is annotated as "3-methyl-2-oxobutanoate dehydrogenase" as a major candidate gene for the variation of 2-phenylethanol content in

our mapped interval. Gene expression analysis and VIGs was used to demonstrate that *Dcx1* is the key gene underlying the QTL for this phenolic volatile.

In chapter 5 we conducted metabolic quantitative trait loci (mQTL) analysis for the identification and mapping of a major glutamate QTL on tomato chromosome 4, using the same F2 population mentioned above. This QTL was fine-mapped to a region of approximately 300-kb from 57 to 57.32 Mb which is harbouring ~30 candidate genes.

Chapter 6 is a general discussion and I integrated the results of the four experimental chapters. The thesis concludes with a summary of the work and key messages.

References

Aflitos, S., Schijlen, E., de Jong, H., de Ridder, D., Smit, S., Finkers, R., . . . Peters, S. (2014). **Exploring genetic variation in the tomato (Solanum section Lycopersicon) clade by whole-genome sequencing.** *Plant J, 80*(1), 136-148.

Agius, C., von Tucher, S., Poppenberger, B., & Rozhon, W. (2018). **Quantification of sugars and organic acids in tomato fruits.** *MethodsX, 5*, 537-550.

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

Baldwin, E., Scott, J. W., & Shewinaker, C. K. (2000). **Flavor Trivia and Tomato Aroma: Biochemistry and Possible Mechanisms for Control of Important Aroma Components.** *HORTSCIENCE, 35*(6).

Barry, C. S., & Pandey, P. (2009). **A survey of cultivated heirloom tomato varieties identifies four new mutant alleles at the green-flesh locus.** *Mol Breed, 24*(3), 269-276.

Bauchet, G., Grenier, S., Samson, N., Segura, V., Kende, A., Beekwilder, J., . . . Causse, M. (2017). **Identification of major loci and genomic regions controlling acid and volatile content in tomato fruit: implications for flavor improvement.** *New Phytol, 215*(2), 624-641.

Bellisle, F. (1999). **Glutamate and the UMAMI taste: sensory, metabolic, nutritional and behavioural considerations. A review of the literature published in the last 10 years.** *Neurosci. Biobehav. Rev, 23*(3), 423-438.

Bellisle, F. (2008). **Experimental studies of food choices and palatability responses in European subjects exposed to the Umami taste.** *Asia Pac J Clin Nutr, 17*(1), 376-379.

Bennett, A B. (2012). **Taste: Unraveling Tomato Flavor.** *Curr. Biol, 22*(11), R443-R444.

Bergougnoux, V. (2014). **The history of tomato: from domestication to biopharming.** *Biotechnol Adv, 32*(1), 170-189.

Borrelli, G. M., De Vita, P., Mastrangelo, A. M., Cattivelli, L., Sadras, V., & Calderini, D. (2009). **Chapter 14 - Integrated Views in Plant Breeding: Modern Approaches for an Old Topic.** In *Crop Physiology* (pp. 327-354). San Diego: Academic Press.

Boukobza, F., & Taylor, A. J. (2002). **Effect of postharvest treatment on flavour volatiles of tomatoes.** *Postharvest Biol Technol, 25*(3), 321-331.

Brummell, D. A. (2006). **Cell wall disassembly in ripening fruit.** *Funct. Plant Biol, 33*(2), 103-119.

Buttery, R. G., & Ling, L. C. (1993). **Volatile Components of Tomato Fruit and Plant Parts.** In *Bioactive Volatile Compounds from Plants* (Vol. 525, pp. 23-34): J. Am. Chem. Soc.

Causse, M., Desplat, N., Pascual, L., Le Paslier, M.-C., Sauvage, C., Bauchet, G., . . . Bouchet, J.-P. (2013). **Whole genome resequencing in tomato reveals variation associated with introgression and breeding events.** *BMC Genom, 14*(1), 791.

Fentik, D. A. (2017). **Review on Genetics and Breeding of Tomato (Lycopersicon esculentum Mill).** *Adv Crop Sci Tech, 5: 306*.

Deschamps, S., Llaca, V., & May, G. D. (2012). **Genotyping-by-Sequencing in Plants.** *Biology (Basel), 1*(3), 460-483.

Frary, A., & Doganlar, S. (2003). **Comparative Genetics of Crop Plant Domestication and Evolution.** *Turk J Agric For, 27*, 59-69.

Giovannoni, J. (2001). **Molecular biology of fruit maturation and ripening.** *Annu Rev Plant Biol,* 52, 725-749.

Jinap, S., & Hajeb, P. (2010). **Glutamate. Its applications in food and contribution to health.** *Appetite, 55*(1), 1-10.

Jones, R. A., & Scott, S. J. (1983). **Improvements of tomato flavor by genetically increasing sugar and acid contents.** *Euphytica, 32*, 845-853.

Kimura, S., & Sinha, N. (2008). **Tomato (Solanum lycopersicum): A Model Fruit-Bearing Crop.** *CSH Protoc, 2008*, pdb.emo105.

Klee, H. J., & Tieman, D. M. (2013). **Genetic challenges of flavor improvement in tomato.** *Trends Genet, 29*(4), 257-262.

Knapp, S., & Peralta, I. E. (2016). ***The Tomato (Solanum lycopersicum L., Solanaceae) and Its Botanical Relatives.***

Knudsen, J. T., Tollsten, L., & Bergström, L. G. (1993). **Floral scents—a checklist of volatile compounds isolated by head-space techniques.** *Phytochemistry, 33*(2), 253-280.

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., & Zheng, Z. (2014). **Genomic analyses provide insights into the history of tomato breeding.** *Nat Genet, 46*, 1220-1226.

Mahajan, P. V., Caleb, O. J., Singh, Z., Watkins, C. B., & Geyer, M. (2014). **Postharvest treatments of fresh produce.** *Philos Trans A Math Phys Eng Sci, 372*(2017), 20130309.

Mammadov, J., Aggarwal, R., Buyyarapu, R., & Kumpatla, S. (2012). **SNP Markers and Their Impact on Plant Breeding.** *Int J Plant Genomics, 2012*, 1-11.

Marwal, A., Sahu, A. K., & Gaur, R. K. (2014). **Chapter 16 - Molecular Markers: Tool for Genetic Analysis.** In A. S. Verma & A. Singh (Eds.), *Anim Biotechnol,* 289-305. San Diego: Academic Press.

Miller, J. C., & Tanksley, S. D. (1990). **RFLP analysis of phylogenetic relationships and genetic variation in the genus Lycopersicon.** *Theor Appl Genet, 80*(4), 437-448.

Peralta, I. E., & Spooner, D. M. (2007). **History, origin and early cultivation of tomato (Solanaceae).** *In: Razdan MK, Mattoo AK (eds) Genetic improvement of Solanaceous crops, vol 2, tomato. Science, Enfield,*, 1–27.

Peralta, I. E., Spooner, D. M., Knapp, S., & Anderson, C. (2008). **Taxonomy of wild tomatoes and their relatives (Solanum sect. Lycopersicoides, sect. Juglandifolia, sect. Lycopersicon; Solanaceae).** *Syst Bot Monogr, 84*.

Petro-Turza, M. (1986). **Flavor of tomato and tomato products.** *FOOD REV INT, 2*(3), 309-351.

Rambla, J. L., Tikunov, Y. M., Monforte, A. J., Bovy, A. G., & Granell, A. (2014). **The expanded tomato fruit volatile landscape.** *J. Exp. Bot., 65*(16), 4613-4623.

Rick, C. M. (1976). **Natural variability in wild species of Lycopersicon and its bearing on tomato breeding.** *Genet Agrar, 30*.

Rodriguez, G. R., Munos, S., Anderson, C., Sim, S. C., Michel, A., Causse, M., . . . van der Knaap, E. (2011). **Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity.** *Plant Physiol, 156*(1), 275-285.

Sandarani, M., Dasanayaka, D., & Jayasinghe, C. (2018). **Strategies Used to Prolong the Shelf Life of Fresh Commodities.** *J AGR FOOD RES, 9*(1), 1-6.

Sorrequieta, A., Ferraro, G., Boggio, S. B., & Valle, E. M. (2010). **Free amino acid production during tomato fruit ripening: a focus on L-glutamate.** *Amino Acids, 38*(5), 1523-1532.

Tadmor, Y., Fridman, E., Gur, A., Larkov, O., Lastochkin, E., Ravid, U., . . . Lewinsohn, E. (2002). **Identification of malodorous, a wild species allele affecting tomato aroma that was aelected against during domestication.** *J Agric Food Chem, 50*(7), 2005-2009.

Tanksley, S. D., & McCouch, S. R. (1997). **Seed banks and molecular maps: unlocking genetic potential from the wild.** *Science (80-), 277*.

Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., . . . Klee, H. (2017). **A chemical genetic roadmap to improved tomato flavor.** *Science, 355*(6323), 391-394.

Tieman, D. M., Loucas, H. M., Kim, J. Y., Clark, D. G., & Klee, H. J. (2007). **Tomato phenylacetaldehyde reductases catalyze the last step in the synthesis of the aroma volatile 2-phenylethanol.** *Phytochemistry, 68*(21), 2660-2669.

Tomato Genome, C. (2012). **The tomato genome sequence provides insights into fleshy fruit evolution.** *Nature, 485*(7400), 635-641.

Tzin, V., Rogachev, I., Meir, S., Moyal Ben Zvi, M., Masci, T., Vainstein, A., . . . Galili, G. (2013). **Tomato fruits expressing a bacterial feedback-insensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase of the shikimate pathway possess enhanced levels of multiple specialized metabolites and upgraded aroma.** *J Exp Bot, 64*(14), 4441-4452.

Valle, E. M., Boggio, S. B., & Heldt, H. W. (1998). **Free Amino Acid Composition of Phloem Sap and Growing Fruit of Lycopersicon esculentum.** *Plant and Cell Physiol, 39*(4), 458-461.

Yilmaz, E. (2001). **The Chemistry of Fresh Tomato Flavor.** *Turk J Agric For, 25* 149-155.

Chapter 2

Exploration of a re-sequenced tomato core collection for phenotypic and genotypic variation in plant growth and fruit quality traits

Raana Roohanitaziani[1,2], Ruud A. de Maagd[3], Michiel Lammers[3], Jos Molthoff[1], Fien Meijer-Dekens[1], Martijn P.V. van Kaauwen[1], Richard Finkers[1], Yury Tikunov[1], Richard G.F. Visser[1] and Arnaud G. Bovy[1]

[1] Plant Breeding, Wageningen University and Research, PO Box 386, 6700 AJ Wageningen, The Netherlands

[2] Graduate School Experimental Plant Sciences, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[3] Bioscience, Wageningen University & Research, PO Box 16, 6700 AA Wageningen, The Netherland

Abstract

A tomato core collection consisting of 122 accessions was explored for variation in several plant growth, yield and fruit quality traits, in order to introduce novel traits from land races, old cultivars and wild relatives into future tomato breeding programs. The re-sequenced accessions were also genotyped with respect to a number of mutations or variations in key genes known to underlie these traits. Plants were phenotyped for the crop growth-related traits abscission zone, inflorescence branching, vegetative outgrowth of the inflorescence, growth rate and earliness of flowering and the fruit characteristics number of ripe fruits per genotype, fruit weight, firmness, total soluble solid content (Brix), colour and shape. The core collection was highly variable for all the investigated traits. The yield-related traits fruit number and fruit weight were much higher in cultivated varieties compared to wild accessions. On the other hand, in wild tomato accessions Brix was higher than in cultivated varieties. Most cultivated accessions contained the large fruit allele of *fruit weight 2.2*, while the "wild" and the "cultivated" allele for *fruit weight 3.2* were more equally distributed. The large fruit alleles of *lc* and *fas* were predominantly present in the cultivated accessions with the highest fruit weights. The occurrence of the *ovate* allele was previously reported for most accessions and newly reported here for Nagcarlan and RZ26. The *sun* locus was found in 9 accessions in our collection. Known mutations in key genes of the carotenoid and flavonoid pathway could well explain the fruit colour variation. The presence and phenotype of several plant architecture affecting mutations, such as *self-pruning* (*sp*), *compound inflorescence* (*s*), *jointless-2* (*j-2*) and *potato leaf* (*c*) was also confirmed. This study provides valuable phenotypic information on important plant growth- and quality-related traits in this collection. The allelic distribution of known genes underlying these traits provides valuable insight in the role and importance of these genes in tomato domestication and breeding. This resource can be used to support (precision) breeding strategies for tomato crop improvement.

Keywords: Tomato**,** *S. lycopersicum*, Tomato Germplasm, Genotyping, Phenotyping, domestication, allele mining

Introduction

Wild relatives, old accessions and landraces held in germplasm collections of crop species represent an underexploited wealth of genetic variation and will therefore offer a valuable gene pool to cope with existing and new breeding challenges (www.tomatogenome.net; Miller & Tanksley, 1990). Among cultivated plants tomato is in a favourable position, due to the availability of related wild species that can be crossed with cultivated varieties and this has been used in recent years by breeders to diversify their genetic material through trait introgression (Rick & Chetelat, 1995; Bai & Lindhout, 2007; Blanca et al., 2012; Blanca et al., 2015). Most efforts have focussed on the introgression of disease resistance genes and genes involved in abiotic stress tolerance, but this gene pool can be used for many other traits as well (Tanksley & McCouch, 1997; Bai & Lindhout, 2007; Aflitos et al., 2014).

The most prominent changes observed during domestication and breeding of tomato concern fruit morphological traits such as fruit size, shape and colour. The molecular basis of these domestication traits has been studied in recent years and several genes affecting these traits have been identified. According to these studies, variation in six loci has been shown to play a role in transforming the small berries of wild tomatoes to the extremely large fruits we observe now in modern cultivars (Grandillo et al., 1999; Lippman & Tanksley, 2001; Van der Knaap & Tanksley, 2003; Tanksley, 2004). Two loci, *fasciated* (*lcn11.1*, on chromosome 11) and *locule number* (*lcn2.1*, on chromosome 2) have been identified as affecting fruit size by determining the number of carpels in the flower (Lippman & Tanksley, 2001; Van der Knaap & Tanksley, 2003; Tanksley, 2004). According to these studies, *fasciated* and *locule number* affect both the final size and the shape of the fruit. Although *fw11.3* and *lcn11.1* were found to be closely linked and originally thought to represent the same underlying gene or locus (Lippman & Tanksley, 2001), cloning of the underlying genes has since then shown that they are distinct loci, with *lcn11.1* renamed *fasciated* (Cong et al., 2008). Other major fruit size loci. Other major fruit size loci, *fw1.1, fw2.1, fw2.2, fw3.1, fw3.2,* and *fw11.3*, exert their effects largely on fruit growth and are able to explain about 67 % of total phenotypic variation, resulting in changes in size with little change in shape (Tanksley, 2004; Muños et al., 2011). Similarly, there are three major loci that modulate fruit shape but with a minimal effect on fruit size. These loci are *ovate* (chromosome 2) (Ku et al., 1990), *sun* (chromosome 7) (Xiao et al., 2008), and *fs8.1* on chromosome 8 (Ku et al., 2000). However, there is no complete separation between the loci controlling fruit size and those controlling fruit shape. Both *ovate* and *sun* lead to the formation of elongated or pear-shaped fruits, while *fs8.1* leads to increased fruit length by increasing the cell number in the proximal-distal direction (Sun et al., 2015). The diversity in fruit colour in tomato is the result of different mutations found during domestication and crop improvement, such as *yellow-flesh (r)* (Fray & Grierson, 1993),

tangerine (t) (Tomes, 1956)*, green-flesh (gf)* (Kerr, 1956)*, old gold (og)*(Thompson, 1956), and *y (yellow)*(Lindstorm, 1925). These mutations have been characterised and reside in genes involved in the biosynthesis of carotenoids and flavonoids or the degradation of chlorophyll (*gf*).

Up to 500 different tomato accessions have been (re)sequenced (Causse et al., 2013; Aflitos et al., 2014; Lin et al., 2014) providing an excellent and untapped resource of promising genetic variation. The availability of such a large number of sequenced tomato genomes facilitates the mapping and cloning of important agronomic or domestication traits, by means of association mapping and/or using different types of mapping populations. In this study, we explored a tomato core collection consisting of 122 tomato accessions for variation in several plant growth and fruit quality-related traits, in order to evaluate the potential of this collection for forward genetics studies. In addition, we evaluated 66 sequenced cultivated accessions of the collection for the presence of known mutations or sequence variations in key genes underlying important domestication and agronomic traits, including inflorescence architecture, fruit pedicel abscission, fruit number, size and shape, fruit colour and soluble solid content. This information is not only valuable for the selection of genotypes for further forward genetics studies, but also demonstrates how sequenced genomes can be used to efficiently mine for allelic variation in candidate genes of interest.

Materials and methods

Plant materials

The core collection for this project consisted of 122 tomato accessions (Table 1). 84 of these accessions were selected from the 150-genome (re)sequencing project (Aflitos et al., 2014) and consisted of 52 cultivated accessions, including tomato landraces and heirloom varieties of *S. lycopersicum* and *S. lycopersicum* var. *cerasiforme* which have been selected from the EU–SOL tomato core collection (https://www.eu-sol.wur.nl) and by participation of companies involved in this project, and 32 accessions comprising wild relatives of tomato. Also, 38 additional *S. lycopersicum* accessions from the EU–SOL tomato core collection (including Heinz 1706, the origin of the reference genome) were selected to be included in this panel with the aim to increase the genetic diversity present in this natural population. After phenotyping these 38 accessions, 14 were selected for further analysis and have been re-sequenced using Illumina Hi-seq.

For greenhouse trials and phenotyping, all self-compatible accessions from the original re-sequencing collection were selected to be included in this project, including related wild species. Of those, 50 cultivated and 20 wild accessions could be grown, in addition to the 38 newly added accessions mentioned above.

Three accessions, cv. Ponderosa (RF_006), RF_017 (*L. esculentum* Mill) and cv. Snowstorm (RF_203) segregated for fruit colour. Fruits of each colour were phenotyped separately.

For allele mining, 52 cultivated accessions from the original re-sequencing collection were used, plus the above-mentioned 14 newly sequenced accessions.

Description of the greenhouse trials

Seeds from all accessions were sown in January 2013. Five weeks later the plants were transplanted to the greenhouse. Plants were grown in two greenhouse compartments as two fully replicated randomised blocks with plots of three plants per accession as experimental units. Each greenhouse contained 8 gutters and on each gutter 14 accessions of each 3 plants were planted. All plants were self-pollinated by vibration or hand-pollination (in case of wild accessions) and seeds were collected. For the phenotyping trial in spring 2014, seeds collected from the tomatoes grown in 2013 were grown in the same setup as described for 2013.

Phenotyping

The collection was phenotyped for 11 traits related to architecture, yield and fruit quality. Some traits were analysed in both years, while others were evaluated in one year only. The fruit characteristics fruit number, fruit weight, Brix, fruit firmness, fruit colour and fruit shape were measured in both 2013 and 2014. The five crop growth-related traits abscission zones of fruit pedicels (AZ category), plant growth rate, the extent of vegetative outgrowth of the inflorescence (VOI), earliness of flowering and inflorescence architecture were measured in 2014 only.

Abscission zones were visually observed and divided in three categories according to their visibility and function as breaking point for the pedicel at harvest: 1. visible and functional; 2: present and visible but less functional; 3: no visible abscission zone (Figure 2). Inflorescence architecture was visually assessed and classified to 5 categories 1. simple/fishbone; 2: simple and forked; 3: forked; 4: forked and compound; 5: compound (Figure 3). Vegetative outgrowth of the inflorescence of each plant was scored in one of 3 categories (1: no outgrowth, 3: outgrowth of leaves; 5: outgrowth of shoots and leaves). Plant growth rate was measured by the number of days from sowing to reaching the attachment wire (3m). Earliness of flowering (EF) was measured by counting the number of nodes up to the first inflorescence.

For fruit number per genotype, the total number of ripe fruits harvested from trusses 1 to 4 of the three plants per accession in each compartment was counted. Fruit trusses were not pruned. The total weight of all fruits was measured and divided by the fruit number to calculate the value for 'fruit weight'. Fruit shape and internal structure characteristics were scored by making flatbed scans of 4 fruits cut

longitudinally as well as transversely. Fruit colour was determined by taking a picture of the fruits and by scoring by eye.

Firmness was measured as the average of 4 measurements per fruit around the mid-height circumference of the fruit using a handheld Fruit Hardness Tester (53215, Turoni Italy) and expressed as average score. Average Firmness of at least 4 (ripe) fruits per genotype at time of harvest was measured and averaged per genotype for each compartment and each season separately. Total soluble solids (Brix) was measured and averaged for at least 4 fruits (one measurement per fruit) per genotype, using an Atago PR-32α brix meter.

Genotyping

Genotyping of the initial set of 232 tomato accessions from the EU-SOL core collection was performed using the SOLCAP infinium SNP array (Finkers et al., 2018). Data were deposited at https://doi.org/10.5281/zenodo.2385441.

Sequencing

Alignment and Variant Call Format files of previously re-sequenced accessions (Aflitos et al., 2014) were available locally and are identical to those deposited at the European Nucleotide Archive (study number: PRJEB5235). Genome resequencing of novel accessions was performed as described previously (Aflitos et al., 2014). Data are deposited at the European Nucleotide Archive (study number PRJEB29506). Genotyping of accessions with respect to mutations or variants that have previously been characterized (Table S 3) was performed using two approaches. For mutations or variants known to be caused by or correlated with small INDEL's or SNP's, VCF files were analysed by SnpEff v3.4 (Cingolani et al., 2012), using the iTAG2.3 annotation on the SL2.40 tomato reference genome version to detect or predict sequence variation affecting protein sequence. For detection of mutations caused by larger deletions, (retrotransposon) insertions, or chromosomal rearrangements, read alignments to the SL2.40 reference genome sequences were inspected using the Integrative Genome Viewer (IGV) software (Robinson et al., 2011). Towards this purpose, reads showing a significantly larger than average distance between pairs were taken as proof for the presence of a deletion at the previously reported genome position underlying the mutation. Nearby accession reads that were not locally paired but paired with reads from various other genomic locations, as well as the presence of truncated mapped reads at the site of the previously characterized retrotransposon insertion were taken as evidence for the presence of the insertion in that accession.

Results

Description of the core collection

A set of 232 genetically diverse tomato accessions were selected from the EU-SOL collection (Aflitos et al., 2014) and genotyped with the SOLCAP infinium SNP array (7414 SNPs) (Sim et al., 2012). A dissimilarity matrix was calculated using the statistical software R (https://www.r-project.org) and was used to construct a neigbour joining tree (Figure S1). The entire collection of 122 accessions used in this study consisted of 84 re-sequenced accessions described by Aflitos et al (Aflitos et al., 2014) and 38 accessions which were additionally selected from the phylogenetic tree. Of the latter, 14 accessions were re-sequenced. Sequence information of in total 66 cultivated accessions, and in addition the reference genome of Heinz 1706, was used to genotype for mutations and known variants of interest. Wild accessions were not included in the genotyping activities, since they generally contain many more polymorphisms in the target genes and it is unclear whether and how these influence the traits under study. In total 107 accessions (88 cultivated and 19 self-pollinating wild accessions) were grown to maturity and phenotyped (Table 1).

Table 1. Origin and analysis of the accessions used in this study

Origin	Type	Number of accessions	Phenotyped	Sequenced genomes for allele mining
150 genome project	Cultivated	52	50	52
	Wild	32	19	
This study	Cultivated	38	38	14
Heinz reference	Cultivated			1
	Total	**122**	**107**	**67**

Of the 122 accessions in this core collection, 107 accessions were phenotyped, including 50 cultivated and 19 self-pollinating wild accessions. Two cultivated accessions failed to grow and hence could not be phenotyped. Sequence data used for genotyping were publicly available (52 cultivated accessions from the 150 genome project plus the Heinz 1706 reference sequence) or newly generated as part of this study (14 cultivated accessions), producing a total of 67.

Phenotyping of the core collection

The set of 108 tomato accessions selected for phenotyping was grown in 2013 and 2014. Plants were phenotyped for architecture, yield, and fruit quality related traits. The core collection was highly diverse for all the investigated traits, as illustrated for fruit morphology, size and colour in Figure 1. The results of the observations are shown in Table S1-S2. Table S1 shows the results obtained for the crop growth related traits abscission zone (AZ), inflorescence branching, vegetative outgrowth of the inflorescence (VOI), growth rate (days required to reach the crop wire at 3 m) and earliness of flowering (EF; number of nodes up to the first inflorescence). Table S2 shows the results for fruit characteristics, such as the number of ripe fruits harvested per genotype, fruit weight, Brix, firmness, colour and shape.

Figure1. Diversity in the core collection. A diverse set of tomato accessions was selected for phenotypic and genotypic evaluation.

Plant Architecture traits

The abscission zones (AZ) were scored in three categories, according to their visibility and function as breaking point for the pedicel at harvest: 1. visible and functional; 2: present and visible but less functional; 3: no visible abscission zone (Figure 2). All three AZ categories were found among cultivated accessions: 23 had a clear and functional abscission zone, 59 had a visible, but less functional abscission zone and only 3 accessions had no abscission zone at all. All six wild accessions analysed for AZ category had a clearly visible and functional abscission zone, except T495 (RF_043) in which the abscission zone was visible but not functional (Table S1).

Figure 2. Categories of fruit pedicel abscission zones. Categories were assigned according to their visibility and function as breaking point for the pedicel at harvest: 1. Visible and functional; 2: Present and visible but less functional; 3: no visible abscission zone.

Three types of Inflorescence architecture were observed in our collection: simple/fishbone, forked and compound (Figure 3). Most of the genotypes had only one type of inflorescence architecture, while 21 accessions had two types. Based on these observations, variation in the branching of the inflorescences in our collection was classified as: 1. simple/fishbone; 2: simple and forked; 3: forked; 4: forked and compound; 5: compound. Three cultivated accessions, cv. Katinka Cherry (RF_007), cv. Lidi (RF_014) and DL/67/248 (RF_226) had a compound inflorescence and this resulted in significantly higher fruit numbers in these accessions (Table S1-S2, Figure S2). For accessions belonging to categories 1-4, no relationship was found between fruit number and inflorescence type. All the wild accessions belonged to categories 1, 2 and 3 (Table S1).

The extent of vegetative outgrowth of the inflorescence (VOI) was scored as categories 1 to 5 (no outgrowth to severe outgrowth) by visual observation of inflorescences per genotype. Cultivars Lidi (RF_014), Dana (RF_018) and accession RF_237 showed severe outgrowth. All wild accessions scored from 1-3, while among cultivated accessions all categories were found (Table S1).

Figure 3. Inflorescence branching categories. 1) simple or fishbone, 3) Forked, 5) compound. Based on these three architecture types our genotypes were classified as : 1. Simple/fishbone; 2: simple and forked; 3: forked; 4: forked and compound; 5: compound.

Crop yield-related traits

The highest growth rate was observed for accession RF_017(*Lycopersicum esculentum* Mill.) with 97 days to reach the crop attachment wire at 3 meter and the lowest growth rate was registered for cv. Tessaloniki (RF_096), cv. Rutgers (RF_004) and cv. Jaune Flamme (RF_230) with 185 days, depending on the compartment (Table S1). A significant correlation was found between the plant growth rate in the two compartments (Figure S3). The variation in plant growth rate observed for wild and cultivated accessions is shown in Figure 4a. In wild accessions, plant growth rate is significantly higher than in cultivated tomato, since they need on average less days to reach the crop attachment wire (142 days in wild accessions compared to 129 days in cultivated accessions).

Earliness of flowering was measured by counting the number of nodes up to the first inflorescence (Table S1). This varied from 6 nodes for cv. John's big orange (RF_008, compartment 1) to 16 for RF_237 ("var. cerasiforme", compartment 2). On average , flowering started 2 nodes earlier in cultivated (9 nodes) compared to wild (11 nodes) accessions (Figure 4b). Since the wild accessions grow faster than cultivated accessions (Figure 4a) we can't exclude that both groups of accessions start flowering at more or less the same time after planting.

Fruit yield related traits

Fruit number varied from less than 10 fruits/plant to more than 500 fruits/plant. Based on the average of both seasons, the highest fruit number was observed for cv. DL/67/248 (RF_226) and cv. Lidi (RF_014), two accessions with compound inflorescences and the lowest fruit number was observed for *S. neorickii* (RF_056) and *S. pennellii* (RF_074) with simple and forked inflorescences, respectively. The variation for fruit number among the collection, based on the average of two seasons, is shown in Figure 4c. Despite the outliers, there was no significant difference in the number of fruits produced by cultivated and wild accessions.
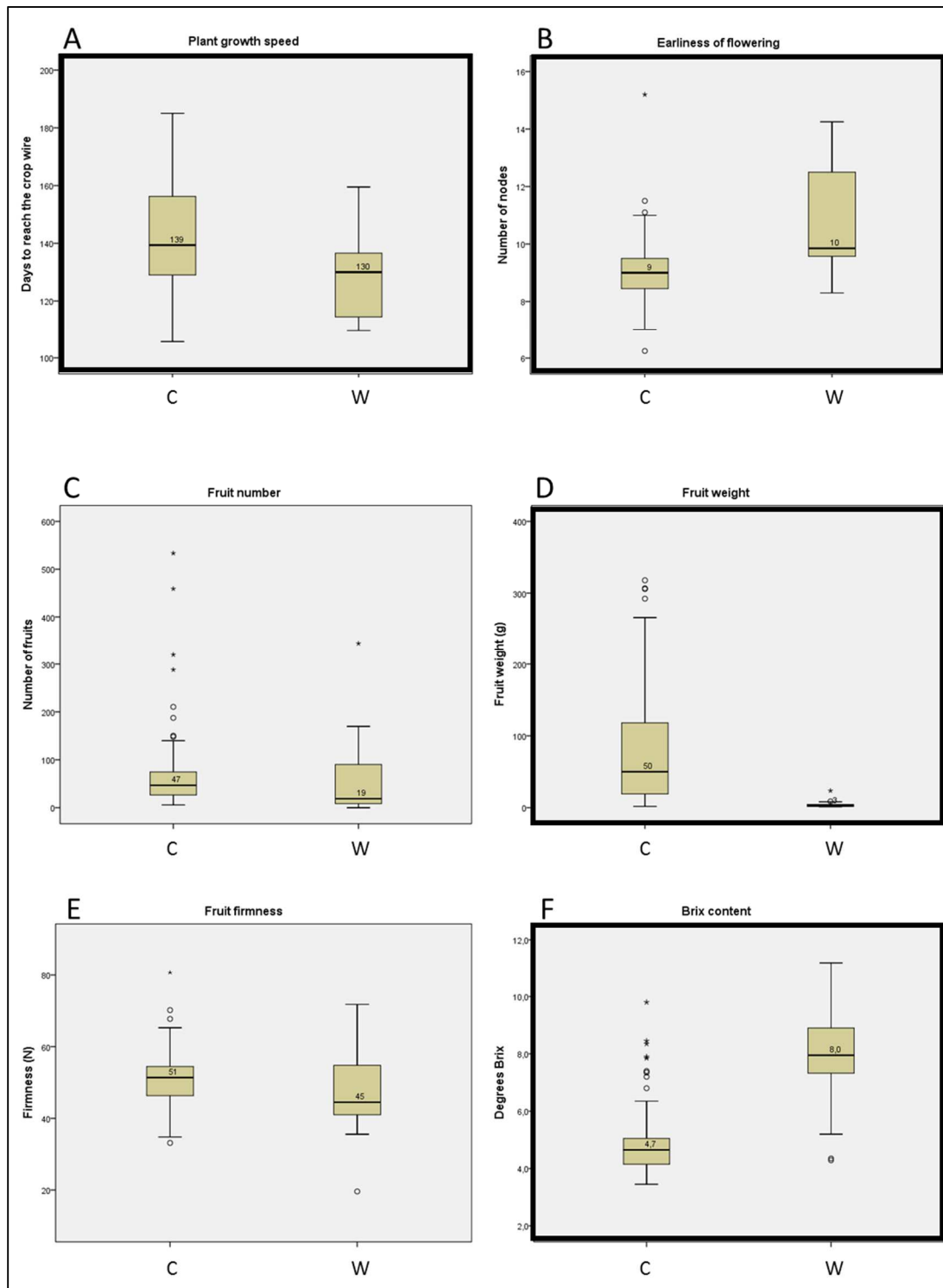
Figure 4. Distribution of fruit yield and fruit quality related traits among the core collection based on the two classes of genotypes, cultivated (C) and wild (W). (A) plant growth speed; (B) earliness of flowering; (C) Fruit number; (D) Fruit weight. (E) Fruit firmness; (F) Brix content. Box plots with thick borders show a significant difference (T-test, P<0.05) between cultivated and wild accessions. Numbers in the plots represent the median value. Outliers are marked with an open circle and far outliers with an asterisk, according to SPSS criteria.

Fruit weight ranged from 1 gram per fruit (N 481 -*S. pimpinellifolium* RF-044, 2013) to up to 360 gram per fruit (cv*.* The Dutchman, RF_028, 2014). Fruit weight showed a high correlation between the two years (Figure S4). Based on the average of the two seasons, cultivated varieties showed a much larger range of variation in fruit weight compared to wild accessions and their fruit weight was much higher (a median fruit weigh of 50 versus 3 gram)(Figure 4d). There was a clear negative correlation between fruit number and fruit weight, especially in the range from 10-100 gram, representing cherry and round type tomatoes (Figure 5). This inverse correlation was less obvious in very small-fruited, wild accessions in which fruit number varied strongly with little impact on fruit weight, and in large-fruited accessions, where fruit weight varied strongly with little impact on fruit number. The harvestable yield (total fruit production per plant) was predominantly influenced by fruit weight and much less so by fruit number (Figure 6 A,B). An exception is formed by the two accessions DL/67/248 (RF_226) and cv. Lidi (RF_014) which had a relatively high yield compared to other genotypes with similar fruit weight, due to their very high fruit number resulting from their compound inflorescences.
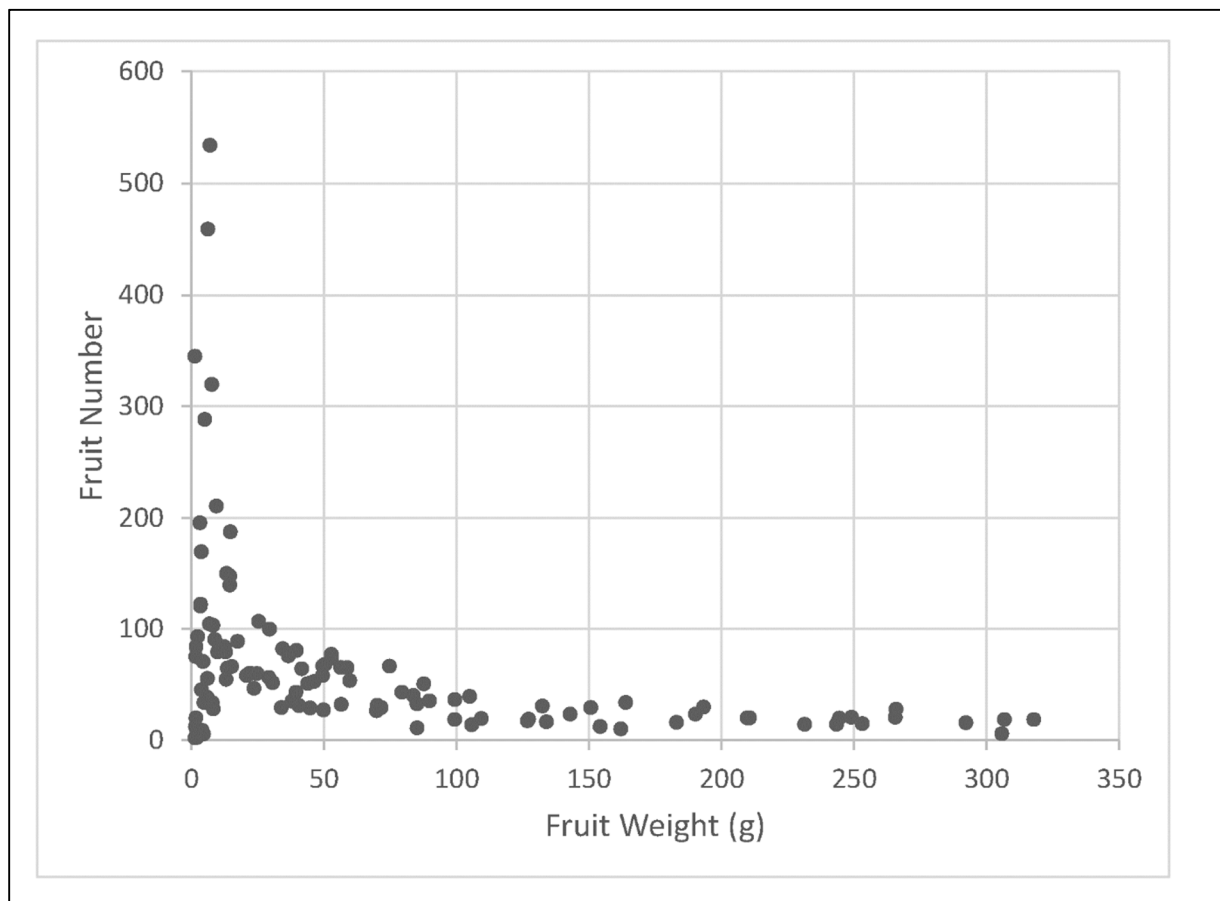
Figure 5. Scatter plot of fruit number versus fruit weight.
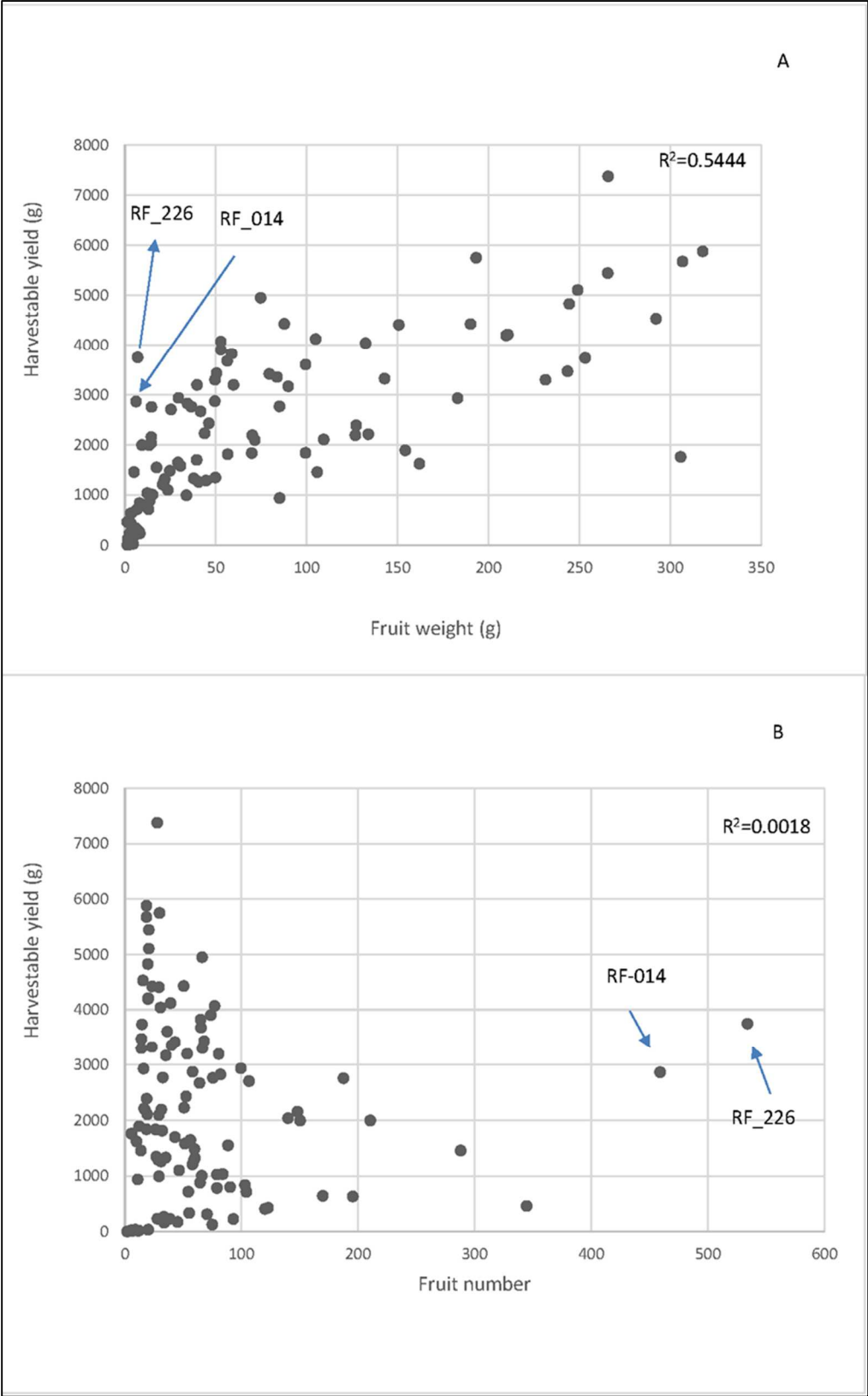
Figure 6. Scatter plot of fruit number (A) and fruit weight (B) versus harvestable yield.

Fruit quality related traits

The collection harboured accessions with different fruit shapes, such as round, ellipsoid, ovate, rectangular, flat, heart and ox-heart, and varying colours, ranging from pink, yellow, orange, light and dark red, to purple and striped (Figure 1, Table S2).

There was extensive variation in firmness among the genotypes in our collection (Table S2). The highest firmness at harvest in the 2013 season was found for fruits of OH88119 (RF_ 232) and RZ26 (RF_238) with firmness values of 75 and 74.5 Newton (N), respectively and the lowest firmness was found for the accession var. *cerasiforme* (RF_102) with a firmness of 24.4 N. In the 2014 season the highest firmness observed was 80.8 N for the genotypes EZ 033 (RF_231, a reported *rin* mutant), and the lowest firmness was found for *S. neorickii* (RF_057) with a firmness of 19.6 N. As shown in Figure 4e, there was no significant difference in firmness between the groups of cultivated and wild accessions. Both groups harbour several accessions with a very high fruit firmness that might potentially be used as novel donors for fruit firmness. Total soluble solids content (Brix) varied from 3.5 (var. *cerasiforme* RF_103) to 9.8 (*Lycopersicum esculentum* Mill, RF_017) degrees in 2013 and from 3.2 (ES 58 Heinz, RF_040) to 11.2 (*S. chmielewskii*, RF_051) degrees in 2014 . Brix values correlated well (R^2=0.75) between the two seasons (Figure S5). Based on the average of the two seasons the distribution of brix in wild and cultivated accessions is shown in Figure 4f. The box plot reveals that the brix content in most of the wild accessions is significantly higher than in cultivated ones.

We observed an inverse correlation between fruit weight and soluble solids content in accessions with an average fruit weight less than 15 gram – cherry type *S. lycopersicum* accessions as well as wild species (Figure 7). They showed a very wide Brix range (from 3 to 10 degrees) and the highest soluble solids levels were found in the smallest fruits. In contrast, fruits with an average weight above 15 gram never showed BRIX levels higher than 5.6 degrees in this experiment, but, there was no decrease in soluble solids content when fruit weight increased further in the entire range from 15 to 300 gram. A similar, but less pronounced trend was observed when Brix and harvestable yield were compared (Figure 8): genotypes with high Brix content (higher than 7 degrees) had the lowest yield, but in the lower Brix range, yield increases seemed to be possible without a large penalty on Brix content. No relationship was found between the Brix content and fruit firmness (Figure S6).
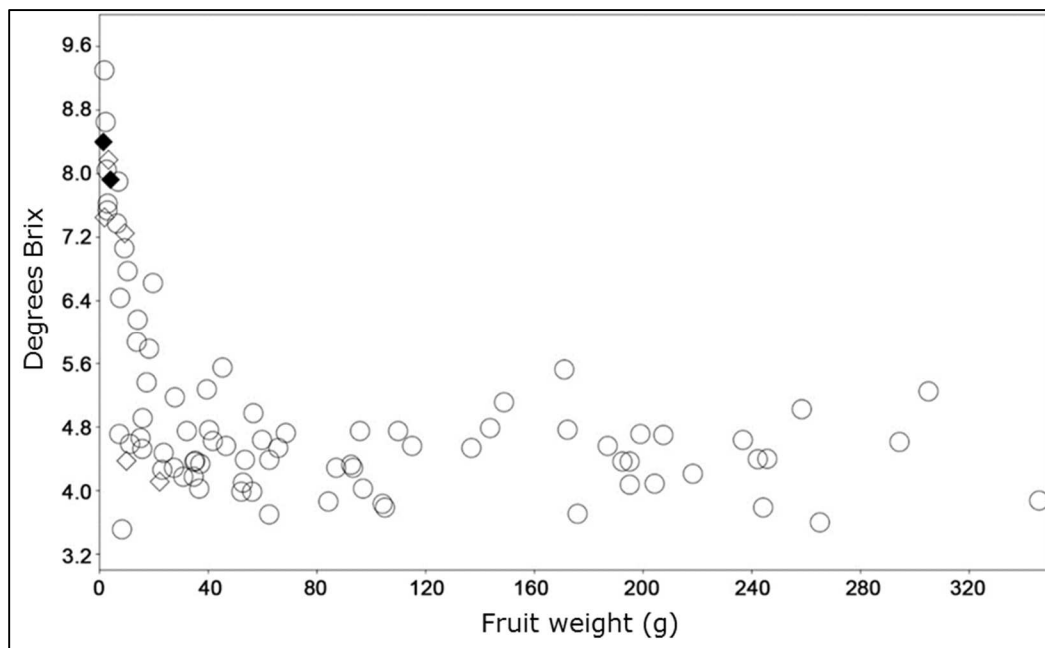
Figure 7. Relationship between Brix (soluble solids content) and fruit weight. Circles represent *S. lycopersicum* accessions, diamonds – *S. pimpinellifolium and S. cheesmaniae*, filled diamonds – *S. chilense* and *S. habrochaites*
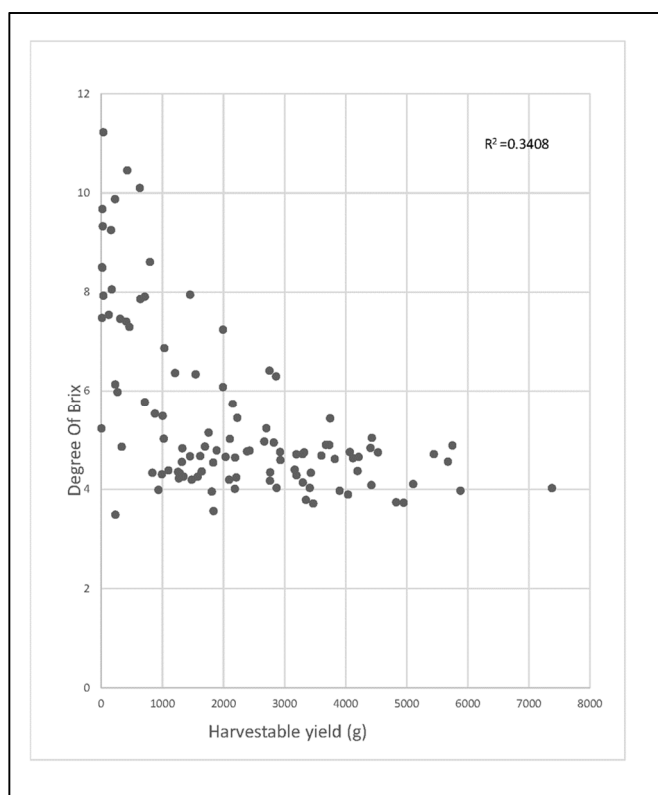
Figure 8. The relationship between Brix content and harvestable yield.

Genotyping of accessions for known mutations or variants affecting plant architecture, fruit size and shape, fruit colour

All re-sequenced accessions were genotyped with respect to a number of mutations or variants that are caused by, or strongly linked to known (combinations of) SNP's, small INDEL's, or larger deletions, insertions, or rearrangements (Table S3). SNP's and INDEL's were extracted from VCF files, while larger deletions, insertions, or rearrangements were identified by unusual or absent read pairing of accession reads mapped to the reference genome. All detected mutations or variants in the re-sequenced accessions are listed in Table 2. For a small number of mutations the presence in particular accessions was already listed in the TGRC database or characterized in literature (as indicated in Table S3), and their presence was confirmed by genotyping in this study. The accession from which the reference genome was derived, Heinz 1706, itself contains a number of mutations such as *self-pruning* (*sp)*(Pnueli et al., 1998), *uniform* (*u)*(Powell et al., 2012), and *ovate* (*o)*(Liu et al., 2002), so that an apparent SNP or INDEL in the majority of other accessions often actually indicates the presence of the wild type or ancestral allele. The classic *sp* allele, leading to a determinate growth habit, caused by a missense mutation leading to substitution of a Proline residue at position 76 by a Leucine occurs in eleven cultivated accessions (of which three as heterozygous). We identified a likely novel *sp* allele in the determinate accession "Nagcarlan" (RF_227), with a missense mutation leading to substitution of Glutamine 128 by a Lysine residue. Prediction of the substitution's effect on protein function using Provean revealed that this substitution with an effect score of -3.7 is likely to be deleterious and thus may well explain the determinate phenotype. Other mutations known to affect plant architecture are rare in our accessions. Two accessions with highly branched inflorescences cv*.* Lidi (RF_014) and accession DL/67/248 (RF_226) contain the *compound inflorescence* (*s*) (Lippman et al., 2008) mutation affecting the function of a *WUSCHEL* homolog (Table S1). Four accessions contain the previously identified *c* or "potato leaf" mutation as the *c-1* allele, caused by a retrotransposon *Rider* insertion in the third exon of a MYB transcription factor encoding gene (Figure S7A.) (Busch et al., 2008). A novel mutant allele of *C*, characterized by an approximately 400 base pair deletion in the second exon, was found in cv. Galina (RF_005) (Figure S7A and S8). The recently characterized *jointless-2* mutation caused by a *Rider* transposon insertion in the first intron of MADS-box protein encoding gene *MBP21* (Soyk et al., 2017; Roldan et al., 2017) (Figure S7B) was found in the only four accessions having no visible pedicel abscission zone (AZ score 3): Cal J TM VF (RF_027), OH88119 (RF_232), NCEBR2 (RF_233) and 981136 (RF_234) (Table S1).

Table 2. Detected mutations or variants in the re-sequenced accessions.

Accession	EU-SOL ID	Variety or Cultivar Name	fas	fw2	lc	SU	fw3	ova	fwt	c	gf	nor	og	ry	s	sp	t31	u	ug	y	j-2
RF_001	EA05097	Moneymaker	-	+	-	-	d	-	d	-	-	-	-	-	-	-	-	+	-	-	-
RF_002	EA05096	Ailsa Craig	-	+	-	-	d	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF_003	EA06086	Garderners Delight	-	+	p	-	d	-	a	-	-	-	-	-	-	-	-	-	-	-	-
RF_004	EA00465	Rutgers	-	+	+	-	d	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF_005	EA00325	Galina (Galina's yellow; Galina's cherry)	-	+	-	-	a	-	d	+	-	-	-	+	-	-	-	-	-	-	-
RF_006	EA00448	Ponderosa	+	+	+	-	d	-	d	-	-	-	-	-	-	-	+	-	-	-	-
RF_007	EA00375	Katinka Cherry	-	+	p	-	d	-	a	-	-	-	-	-	-	-	+	+	-	-	-
RF_008	EA00371	John's big orange	-	+	+	-	d	-	d	-	-	-	-	-	-	+	-	+	-	-	-
RF_011	EA02617	All Round	-	+	-	-	d	-	d	-	-	-	-	-	-	-	-	+	-	-	-
RF_012	EA02724	Sonato	-	+	-	-	d	-	d	-	-	-	-	-	-	-	-	+	-	-	-
RF_013	EA03701	Cross Country	-	+	-	-	a	+	d	-	-	-	-	-	-	+	-	+	-	-	-
RF_014	EA03362	Lidi	-	+	-	-	d	+	d'	-	-	-	-	+	+	-	-	-	-	-	-
RF_015	TR00003	Momatero	-	+	+	-	d	-	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_016	EA01965	Rote Beere	-	p/+	p	-	a	-	d'	-	-	-	-	-	-	-	-	-	-	-	-
RF_017	EA03306	"Lycopersicum esculentum Mill."	-	+	p	-	a	-	d'	-	-	-	-	+	-	-	-	-	-	+	-
RF_018	EA01155	DANA	+	+	+/-	-	d	-	d	-	-	-	-	-	-	+	-	+	-	-	-
RF_019	EA01049	Large Pink	+	+	+	+	d	-	d	+	-	-	-	-	-	-	-	-	-	+	-
RF_020	EA03221	"Lycopersicon esculentum Mill."	+	+	+	-	d	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF_021	EA03222	Bolivar' Lycopersicon esculentum Mill.	+	+	+	-	d	-	d	-	-	-	-	-	-	+	-	+	-	-	-
RF_022	EA04710	"Lycopersicum esculentum "	-	+	x/-	+	a	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF_023	EA05170	"Lycopersicum esculentum "	-	x	-	-	a	-	d'	-	-	-	-	-	-	-	-	-	-	+	-
RF_024	EA00990	Jersey Devil	-	+	+	+	d	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF-025	LA0118	S. corneliomulleri	-	+	-	a	-	-	?	-	-	-	-	-	-	-	-	-	-	-	-
RF_026	EA00157	Polish Joe	+	+	+	+	d	-	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_027	EA02054	Cal J TM VF	-	+	-	-	d	-	d	-	-	-	-	-	-	+	-	+	-	-	-
RF_028	EA05581	The Dutchman	+	+	+	-	d	-	d	-	-	-	-	-	-	-	-	+	-	+	-
RF_029	EA00027	Black Cherry	-	p	p	-	d	-	a	-	+	-	+	-	-	-	-	-	-	+	-
RF_030	EA01835	ANTO	+	+	-	-	d	+	d	-	-	-	-	-	-	-	-	-	+	-	-
RF_031	EA01854	Winter Tipe (nor)	-	+	-	-	a	-	d	-	-	+	-	-	-	-	-	-	-	-	-
RF_032	EA04243	Chang Li' Lycopersicon esculentum	-	+	-	-	a	+	d'	-	-	-	-	+	-	-	-	-	-	-	-
RF_033	EA00892	Belmonte	+	+	+	+	d	-	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_034	EA01088	Tiffen mennonite	+	+	+	+	d	-	d	+	-	-	-	-	-	-	-	-	-	+	-
RF_035	EA04939	Wheatley's Frost Resistant	-	+	-	-	a	+	d'	-	-	-	-	-	-	-	-	-	-	+	-
RF_036	EA05701	"Lycopersicum esculentum "	-	x	-	+	a	+	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_037	EA05891	"Lycopersicum esculentum "	-	p	p	-	a	-	a/d	-	-	-	-	-	-	-	-	+	-	-	-
RF_038	EA04828	Chih-Mu-Tao-Se	-	+	+	+	d	-	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_039	EA00526	"Lycoperscium esculentum "	-	p-	+	-	a	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF_040	EA02655	ES 58 Heinz	-	+	+	-	d	-	d	-	-	-	-	-	-	+	-	+	-	-	-
RF_041	EA04861	Dolmalik	+	+	+	-	a	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF_042	LYC2962	Ventura	-	p	p	-	a	-	a	-	-	-	-	-	-	-	-	-	-	-	-
RF_077	TR00018	Large Red Cherry	-	+	-	-	a	-	d'	-	-	-	-	-	-	-	-	-	-	-	-
RF_078	EA00940	Porter	-	+	-	+	d	+	d'	-	-	-	-	-	-	-	-	-	-	+	-
RF_088	TR00019	Bloody Butcher	+	+	+	-	d	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF_089	EA01019	Brandywine	+	+	+	-	d	-	d	+	-	-	-	-	-	-	-	-	-	+	-
RF_090	TR00020	Dixy Golden Giant	+	+	+	-	d	-	d	+	-	-	-	-	-	-	+	-	-	+	-
RF_091	EA01037	Giant Belgium	+	+	+	-	d	-	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_093	TR00021	Kentucky Beefsteak	+	+	+	-	d	-	d	-	-	-	-	-	-	+	-	-	-	+	-
RF_094	TR00022	Marmande VFA	+	+	+	-	d	-	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_096	TR00023	Thessaloniki	-	+	+	-	d	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF_097	EA01640	Watermelon Beefsteak	+	+	+	-	d	-	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_102	TR00026	"var. cerasiforme "	-	+	-	-	a	-	d'	-	-	-	-	-	-	-	-	-	-	-	-
RF_103	TR00027	"var. cerasiforme "	-	p/-	x+/-	?	a	-	a/d'	-	-	-	-	-	-	-	-	-	-	+	-
RF_105	TR00028	"var. cerasiforme "	-*	p	p	-	a	-	d'	-	-	-	-	-	-	-	-	-	-	-	-
RF_201	EA00282	Blondokee	+	+	+	+	d	-	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_202	EA00450	Purple Russian	-	-	-	-		-		-	-	-	-	-	-	-	-	-	-	-	-
RF_203	EA00470	Snowstorm	-	+	+/-	-	d/a	-	d	-	-	-	-	-	-	-	-	-	-	-	-
RF_204	EA00488	Taxi	-	-	-	-		-		-	-	-	-	-	-	-	-	-	-	-	-
RF_205	EA00907	Indian Stripe(d)	-	-	-	-		-		-	-	-	-	-	-	-	-	-	-	-	-

Table 2 continued

Accession	EU-SOL ID	Variety or Cultivar Name	fas	fw2.2	lc	SUN	fw3.2	ovate	fw11.3	c	gf	nor	ogc	ry	s	sp	t31	u	ug	y	j-2
RF_206	EA00915	ABC Potato Leaf	+	+	+/-	-	a	-	d	+	-	-	-	-	-	-	-	-	-	-	-
RF_214	EA01579	Heinz 1706 (reference genome)	-	-	-	-	+		-	-	-	-	-	-	-	+	-	-	-	-	-
RF_215	EA01709	"Italian cherry tomato "	-	-	-	-		-		-	-	-	-	-	-	-	-	-	-	-	-
RF_216	EA01985	Sonora	-	+	+	-	d	-	d	-	-	-	-	-	-	+	-	+	-	-	-
RF_226	EA05721	DL/67/248	-	p	-	-	d	-	d'	-	-	-	-	+	-	-	-	-	-	-	-
RF_227	EA05732	Nagcarlan	+	x	-	-	d	+	d'	-	-	-	-	-	-	+	-	-	-	-	-
RF_228	EA05789	V 145-E	-	-	-	-		-		-	-	-	-	-	-	-	-	-	-	-	-
RF_229	EA05979	Morne a L'Eau	-	+	+	-	d	-	d'	-	-	-	-	-	-	+	-	-	-	-	-
RF_230	EA06631	Jaune Flammee	-	-	-	-		-		-	-	-	-	-	-	-	-	-	-	-	-
RF_231	EA06746	EZ 033	-	-	-	-		-		-	-	-	-	-	-	-	-	-	-	-	-
RF_232	EA06902	OH88119	-	+	-	-	d	-	d	-	-	-	-	-	-	**+**	-	+	+	-	**+**
RF_233	EA06929	NCEBR2	-	+	+	-	d	-	d	-	-	-	-	-	-	+	-	+	-	-	**+**
RF_234	EA06932	981136	-	+	-	-	d	-	d	-	-	-	-	-	-	+	-	+	-	-	-
RF_235	EA03075	Kecskemeti Koria Bibor	-	+	+	-	d	-	d	-	-	-	-	-	-	+	-	+	-	-	-
RF_236	EA01042	Grosse Cotelee	+	+	+	-	d	-	d	-	-	-	-	-	-	-	-	-	-	+	-
RF_237	EA06002	"var. *cerasiforme* "	new	p	-	-	a	-	a	-	-	-	-	-	-	-	-	-	-	-	-
RF_238	EA06699	RZ26	new	p	-	-	d	+	a	-	-	-	-	-	-	-	-	-	-	-	-

For fas (+=present). For fw2.2 (+: modern allele for large fruit; p: pimpinellifolium-like allele; x: allele of unknown origin). For lc (+: high number allele; p: pimpinellifolium allele; x: allele of unknown origin). For sun ("+" indicates the presence of the duplicated /translocated copy; "-"= wild ancestral allele). For fw3.2 ("+"=modern large fruit size allele; "-"= wild ancestral allele). For ovate ("+"=presence of the allele). For fw11.3 ("+"= modern allele, as in reference; "-"= wild ancestral allele as in S. pimpinellifolium (a 22nt deletion and large insertion compared to reference; "-*" appears to have only the derived 22nt insertion). For r ("+"=ry allele; "r"=alternative r allele). For all other genes, "+" represents the presence of the mutant allele.

Allelic variation for seven quantitative trait loci (QTLs) for fruit size and weight was investigated by comparing the resequencing data with the published underlying genotype: *fruit weight 2.2* (*fw2.2*) (Frary et al., 2000) *fruit weight 3.2* (*fw3.2*) (Chakrabarti et al., 2013), *fruit weight 11.3* (*fw11.3*) (Mu et al., 2017), *locule number* (*lc*) (Muños et al., 2011), *fasciated* (*fas*) (Xu et al., 2015), *ovate*(Liu et al., 2002) and *sun* (Xiao et al., 2008). Most cultivated accessions contained the modern, cultivated (large fruit) allele of *fw2.2*, while a minority consisting of more primitive *S. lycopersicum* accessions contained alleles that were highly similar to those found in our re-sequenced *S. pimpinellifolium* accessions (Table 2). Two additional accessions contained novel alleles that could not be matched to either the modern cultivated allele or the "S. *pimpinellifolium* allele". The relationship between size locus haplotype and fruit weight derived from our phenotyping effort is shown in Figure 9. The cultivated accessions with the *S. pimpinellifolium* allele of *fw2.2* have fruits smaller than 10 gram, whereas the modern, cultivated allele of *fw2.2* is present in all large-fruited accessions in the collection, suggesting that this allele is required for a large fruit size. For *fw3.2*, cultivated accessions were more equally distributed between having the "wild" or the "modern" (larger fruit) allele. The results of our phenotyping show that all the fruits with weight higher than 50 gr, with one exception, contain the modern allele of *fw3.2* (Figure 9). Most modern cultivars, with the exception of a few cherry-type accessions, contain the modern large fruit allele of *fw11.3*.
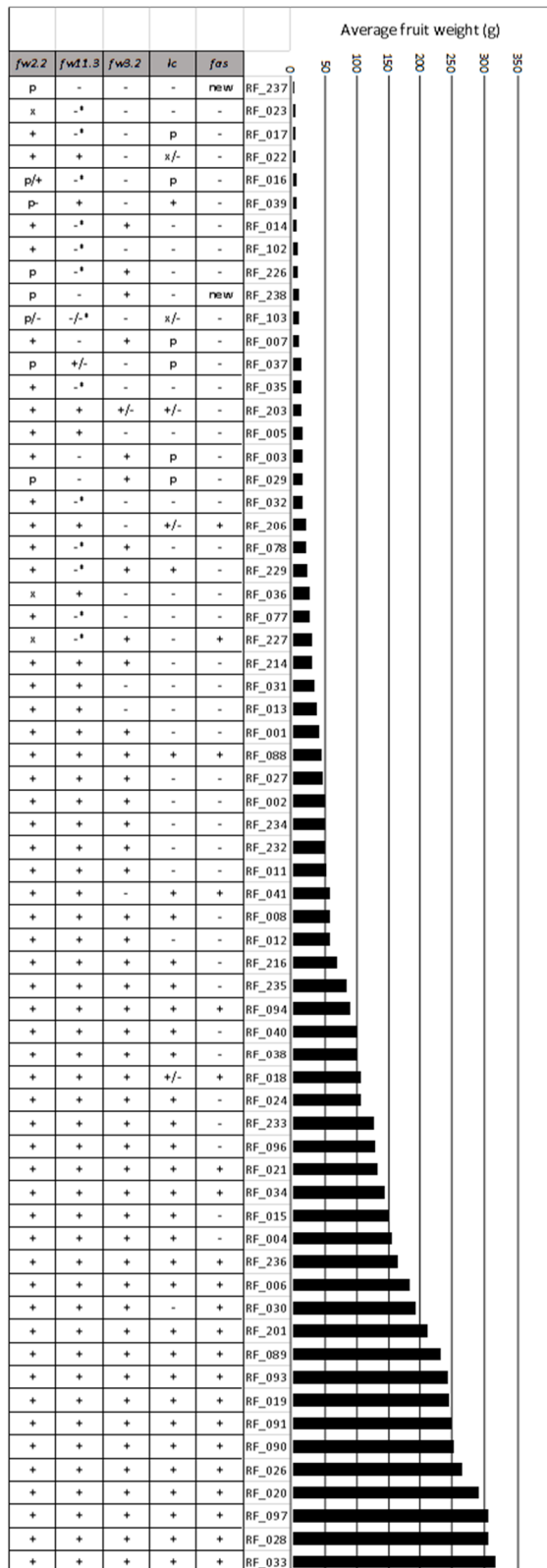
Figure 9. The relationship between size locus haplotype and fruit weight. The different allelic forms are described in the caption of Table 2.

The *locule number* (*lc*) QTL is determined by two SNP's near the gene encoding tomato WUSCHEL (Muños et al., 2011) but their effect on *WUS* expression or function has yet to be determined. The large and small fruit alleles of *lc*, respectively, occur throughout the re-sequenced collection, as does also an allele resembling that of *S. pimpinellifolium* accessions and two unique alleles (Table S3). Most of the big-fruited cultivated accessions have the modern, large fruit *lc* allele, although this allele is neither a requirement for a big fruit size (e.g. ANTO; RF_030), nor a guarantee (e.g. Morne a L'Eau; RF_229) (Figure 9). The *fasciated* (*fas*) QTL has been shown to have a greater effect on locule number and fruit size than *lc* (Muños et al., 2011). Indeed, the modern, large fruit allele of *fas* is present in the cultivated accessions that have the highest average fruit weights (Figure 9, Figure S7C). Two mutations affecting fruit shape, *ovate* (Liu et al., 2002) and *sun* (Xiao et al., 2008), were also investigated. The occurrence of the *ovate* allele, which is also present in the reference accession, was previously reported for most accessions (Tanksley & McCouch, 1997) and is newly reported here only for Nagcarlan (RF_227) and RZ26 (RF_238). The *sun* locus, which causes elongated fruit types, was found in 9 accessions in our collection (Figure S7D). Figure 10 shows examples of the effect of fruit shape mutations in our collection.

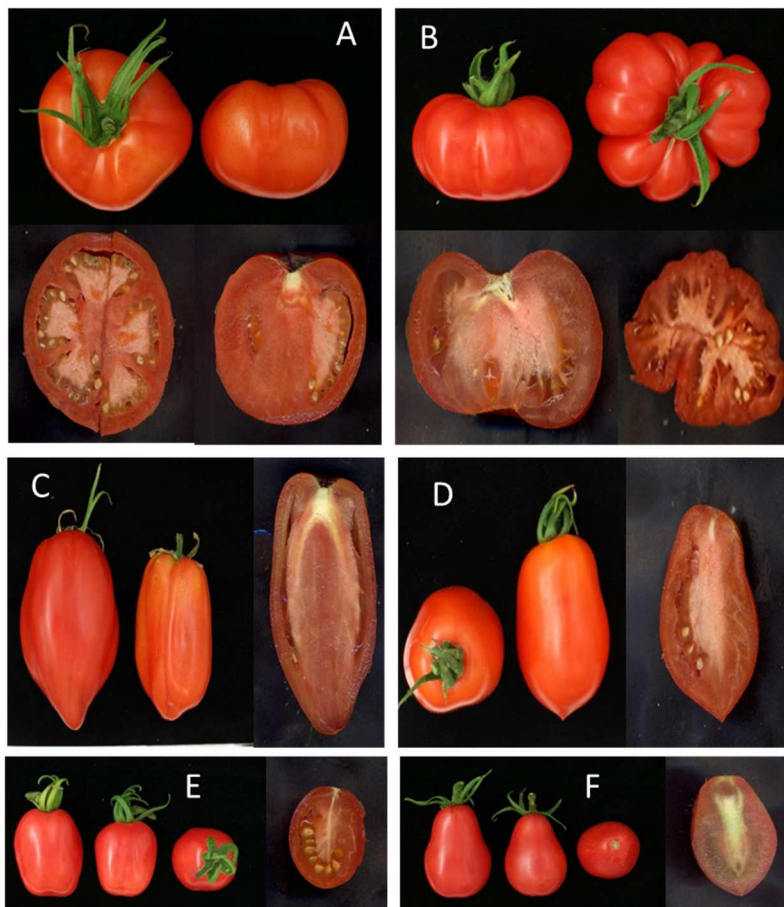

Figure 10. Examples of fruit shape mutations in our collection. (A) lc mutation in RF_096; (B) fas mutation in RF_041; (C) sun mutation in RF_024; (D) ovate mutation RF_214; (E) ovate mutation in RF_035; (F) sun and ovate mutation in RF_043.

Fruit colours in our collection of cultivated tomato accessions ranged from red through pink to orange or yellow. Moreover, a number of more modern accessions, including Heinz 1706, contain the *uniform* (*u*) mutation, having a frameshift- causing deletion in the open reading frame of the transcription factor gene *GOLDEN LIKE 2* (*GLK2*). The wild type allele is responsible for the "green shoulder" phenotype of more ancient accessions, such as cv. Ailsa Craig (Powell et al., 2012). All ripe yellow fruited accessions appeared to contain the r^y allele of the yellow flesh mutation *r*, which is detectable as an approximately 6 kB deletion running from the last exon of *PHYTOENE SYNTHASE 1* (*PSY1*) to the first exon of the neighbouring gene (Fray & Grierson, 1993) (Figure 11a, Figure S7E). Only one example of the alternative *r* allele was found:

cv. Snowstorm (RF_203) was heterozygous for *r* and hence this cultivar segregated for red and yellow fruit colour (results not shown). Another mutation affecting lycopene synthesis, *tangerine*, leads to orange fruits in cv. Ponderosa (RF_006; segregating), cv. Katinka Cherry (RF_007), cv. Dixy Golden Giant (RF_090) and cv. Kentucky Beefsteak (RF_093) and is caused by disruption of *CAROTENE ISOMERASE* activity, in this case from the t^{3183} allele having a deletion located 5' to the open reading frame (Isaacson et al., 2002)(Figure 11b, Figure S7F).
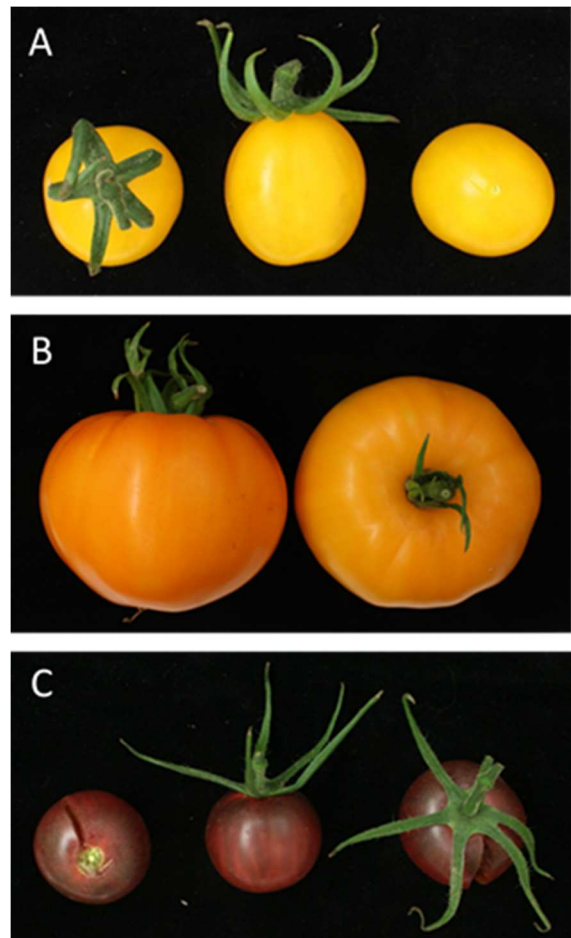
Figure 11. (A) yellow flesh mutation in RF_032; (B) Tangerine mutation in RF_006; (C) The *old-gold-crimson* and *green flesh* mutations in RF_ 029.

The *old-gold-crimson* mutation (Ronen et al., 200) affecting *LYCOPENE β-CYCLASE* in cv. Black Cherry (RF_029) was reported earlier (Aflitos et al., 2014). The same accession holds the *green flesh* allele gf^4 leading to retention of chlorophyll during ripening, as has been reported earlier(Barry et al., 2009). This combination of alleles causes the fruit of this accession to have a deep red, almost black color at the ripe stage (Figure 11c). Finally, the pink mutation *y* causes pink fruits through the lack of the yellow-colored naringenin chalcone in the skin of mutant accessions, caused by deregulated expression of transcription factor gene *MYB12* (Ballester e al., 2010). Of all sequenced accessions, twenty were homozygous and two were heterozygous for the previously identified *y* allele,

having a 603 base pair deletion upstream of the *MYB12* open reading frame as seen earlier in other cultivated accessions (Lin et al., 2014).

There is a single accession, cv. Winter Tipe (RF_031), containing the *non-ripening* (*nor*) mutation (Tigchelaar et al., 1973) This variety contains a two nucleotide deletion in the third exon of the transcription factor encoding gene *NAC-NOR*, causing a frameshift mutation and a premature stop codon, which is highly likely to affect its function.

Discussion

In this study, a collection of 122 tomato accessions, including wild relatives, old cultivars and landraces, was characterised. The collection presented a wide range of phenotypic and genotypic diversity for yield and quality-related traits. This information is expected to be valuable for subsequent tomato breeding programs and crop improvement. In the phenotyping part we evaluated several traits and compared these traits between wild and cultivated accessions (Figure 4). We observed that there is a clear difference between wild and cultivated accessions for several domestication traits. Cultivated accessions have been selected for a higher growth rate, earlier flowering, production of more and larger fruits and a higher firmness compared to wild accessions. On the other hand, total soluble solids content was on average higher in wild accessions compared to cultivated accessions. This may be partly due to a dilution of fruit soluble solids with increasing fruit weight or increasing total yield, which might partially explain the inverse correlation between fruit size and Brix in small fruited accessions (<30 grams), but this correlation is lacking in bigger fruits (Figure 8). Similar trends, though less clear, were observed when total harvestable yield was compared with fruit soluble solids content (Figure 9). This suggests that soluble solids content can be influenced by genetic factors that do not affect fruit weight or total yield. This was also shown for the Brix 9-2-5 QTL, caused by variation in the *LIN5* gene, that was introgressed from *S. pennellii* in a cultivated tomato background and led to a significant increase in Brix without an adverse effect on fruit size and yield (Fridman et al., 2001). Discovery and stacking of such QTL's and their underlying alleles offers opportunities for the development of large-fruited tomato genotypes with increased Brix.

Part of the collection, in as far as it was re-sequenced by (Aflitos et al., 2014) was investigated earlier by (Sahu et al., 2017) who focussed on the analysis of genome-wide sequence variations between wild and cultivated species to identify genomic regions and genes that were selected for during domestication (Frary.et al., 2000). In our study, allele mining was performed for known traits related to fruit morphology (shape, colour and size) and plant architecture. We identified new alleles of the plant

architecture mutations *potato leaf* (*c)* and *self pruning* (*sp*). Although the accessions carrying these new alleles show the expected phenotype (potato leaves and a determinate growth habit), genetic and/or functional studies are needed to demonstrate that these alternative alleles are indeed causal to the observed phenotypes. We confirmed the effect of previously identified mutations or alleles for fruit size and shape genes such as *ORFX* (*fw2.2*)*, KLUH* (*fw3.2*)*, CSR (fw11.3), WUS* (*lc*)*, CLV3* (*FAS*)*, SUN,* and *OVATE* in the cultivated germplasm. The genes controlling fruit size and shape can be divided in three categories:1) the genes that influence only the mass (weight) of the fruit, 2) the genes that influence fruit mass by influencing the locule number and 3) the genes that influence only the shape of the fruit with no effect on fruit weight . The loci belonging to the first category are *fw1.1, fw2.1, fw2.2, fw3.1, fw3.2* and *fw11.3.* Lin et al. postulated a two-step model of tomato domestication or improvement from *S. pimpinellifolium* via *S. lycopersicum* cv. *cerasiforme* to modern large *S. lycopersicum* (Lin et al. 2014). All five fruit size loci that were cloned and sequenced, and could thus be used for genotyping our collection (Figure 9) are involved in the second of those steps. Thus the overall increasing presence of modern alleles of these genes correlated with increasing fruit size can be seen as to represent the evolution and selection of more primitive cv. *cerasiforme* to more modern cultivated tomato. The *fw2.2* locus is considered as the most important fruit weight locus so far, and mutation of the underlying gene is considered to be the first step in domestication with regard to fruit size (Tanksley, 2004). In our collection most of the cultivated accessions contain the modern (large fruit) allele of *fw2.2*, while a minority consisting of more primitive *S. lycopersicum* accessions contained alleles that were highly similar to those found in our re-sequenced *S. pimpinellifolium* accessions (Table 2) and these accessions have small fruits. *fw3.2* is another important tomato fruit weight QTL explaining 19% of the fruit mass v ariance (van der Knaap & Tanksley, 2003). It has been reported that the *fw3.2* locus also controls fruit shape in some tomato varieties and has pleiotropic effects on fruit development (Tanksley , 2004; Brewer et al., 2007; Gonzalo & Van Der Knaap, 2008). For *fw3.2*, the "wild" and the "modern" (larger fruit) allele (Chakrabarti et al., 2013). were more or less equally distributed among cultivated accessions. The "modern" large fruit allele of *fw11.3* was found in most of the cultivated accessions in our collection and is caused by a 1.4 kb 3'deletion in the *Cell Size Regulator* (Solyc11g071940) gene, leading to a 194 amino acid truncation of the predicted wild-type protein. This resulted in a partially-dominant gain of function protein that affects fruit size by increasing the size of mesocarp cells in the fruit pericarp (Mu et al., 2013). *Fw11.3* and *fas* are tightly linked and therefore were long thought to be identical. With the cloning and further characterization of the underlying loci or genes in the last few years it has become apparent that although *CSR* (Solyc11g071940) is only approximately 10 kB downstream from the start of the inversion underlying the *fas* allele (Table S3), the two are distinct.

Fasciated (*fas*, chromosome 11) (Lippman &Tanksley., 2001) and *locule-number* (*lc*, chromosome 2) (Muños et al., 2011) belong to the second category and influence the fruit size by influencing the number of carpels in the flower. Wild tomato species, and many cultivated varieties, produce flowers with a gynoecium containing two to four carpels. After fertilization, each carpel develops into a locule in the fruit. Varieties that produce fruit with more locules generally have larger, wider, sometimes ribbed fruits (van der Knaap & Tanksley., 2003). The large and small fruit alleles of *lc*, respectively, occur throughout the re-sequenced collection, as does an allele resembling that of *S. pimpinellifolium* accessions and two unique alleles (Table 1). The *fas* QTL has been shown to have a greater effect on locule number and fruit size than *lc* (Muños et al., 2011). Although it was originally shown to be linked to two SNP's nearby a YABBY transcription factor-encoding gene (Cong et al., 2008) it was more recently shown that the high locule number accessions contain an inversion on chromosome 11 that not only affects *YABBY*, but more importantly also affects expression of *CLAVATA3* (Xu et al., 2015). In our collection, the large fruit alleles of *lc* and *fas* are present in the cultivated accessions with the largest average fruit weights. The genes of category 3 include two mutations that cause variation in fruit shape with little effect on fruit size. These loci are *ovate* and *sun*. The *ovate* mutation is associated with a change from round to elongated or pear-shaped fruit. OVATE encodes a protein belonging to the Ovate Family Protein (OFP) and is thought to negatively regulate transcription of target genes (Liu et al., 2002). Recent genetic analyses have further identified *ovate* as a major quantitative trait (QTL) controlling pear-shaped fruit development in both tomato and eggplant (Liu et al., 2002). It has been reported that the mutation is not associated with a single phenotype (van der Knaap & Tanksley, 2003; Tanksley, 2004) which we also observed in our collection. In some accessions the *ovate* mutation led to elongated fruit with highly constricted neck growth indicative of pears, but in some backgrounds, neck constriction was not noticeable and the degree of fruit elongation was not so prominent (Figure 10). Some reports suggested that the *ovate* locus interacts with the *sun* locus on chromosome 7 (Tanksley, 2004; Monforte et al., 2014). The *sun* locus was found in 9 accessions in our collection. This locus causes elongated fruit types and arose from a *Rider-mediated* transposition event placing an IDQ10 domain protein encoding gene from the ancestral locus on chromosome 10 under the expression control of the *DEFL1* gene on chromosome 7 (Xiao et al.,2008). Although allelic variation at both *ovate* and *sun* can cause elongated fruit shape, the two loci differ in some important genetic, morphological, and developmental aspects (Tanksley, 2004).

In this study we also examined fruit colour and linked our phenotypic observations with known mutations in the carotenoid and flavonoid pigment pathways. Four yellow accessions in our collection contain the recessive *yellow-flesh* mutation, which is linked to a single locus, *r (red),* on chromosome 3 (Table 2). Locus *r* encodes a fruit-specific phytoene synthase (PSY1), which catalyses a rate-limiting step in the carotenoid pathway (Kachanovsky et al., 2012). The first step in carotenoid biosynthesis in plants is

the formation of phytoene from two molecules of geranylgeranyl diphosphate by the enzyme phytoene synthase (PSY). Four double bonds are then introduced by two enzymes, phytoene desaturase (PDS) and ζ-carotene desaturase (ZDS), each catalysing two symmetric desaturation steps to yield ζ-carotene and lycopene, respectively. Intermediates in this part of the pathway are *cis*-configurations. A specific enzyme, ζ-carotene isomerase (ZISO), is required for *cis*-to-*trans* conversion of the 15—15′ *cis* double bond in tri-*cis*-ζ-carotene, and another isomerase, carotenoid *cis–trans* isomerase (CRTISO), produces *all-trans*-lycopene from tetra-*cis*-lycopene (pro lycopene) (Giuliano et al., 1993; Bramley, 2013). Fruits of tomato with the recessive mutation *tangerine* (*t*) lack a functional form of the enzyme CRTISO (due to deletion of 348 bp in the promoter), which results in accumulation of tetra-*cis*-lycopene and its precursors upstream in the carotenoid pathway, in particular phytofluene Tetra-cis-lycopene and phytofluene in tangerine tomatoes have a light absorption peak approximately 35 nm below that of all-trans-lycopene, resulting in tomatoes with an orange colour Twenty-two accessions in our collection contained the previously identified, recessive *y* allele, having a 603 base pair deletion upstream of the *MYB12* transcription factor gene, which regulates the accumulation of flavonoids in tomato fruit (Ballester et al., 2010). Most pink accessions characterized to date harbour this promoter mutation (Lin et al., 2014), which leads a ripening-dependent suppression of *MYB12* gene expression, resulting in a lack of accumulation of the yellow flavonoid naringenin chalcone in the fruit peel and, consequently, a transparent fruit peel and pink appearance of the fruit (Ballester et al., 2010; Fernandez-Moreno et al., 2016). This phenotype could indeed be observed in all 20 accessions with a homozygous *y* allele.

Among all the accessions in our collection, 3 accessions (cv. Lidi, DL/67/248 and cv. Katinka Cherry) had a highly branched inflorescence and 2 of them (cv. Lidi, DL/67/248) had the *s* (compound inflorescence) allele. The *s* allele transforms the simple inflorescence into a highly branched structure with many flowers. The *S* gene encodes a homolog of the WUSCHEL HOMEOBOX 9 transcription factor and is involved in regulating inflorescence architecture (Lippman et al., 2008). Wild tomatoes have a simple inflorescence and branched inflorescences occurred during domestication. Although in some crops (e.g. cereals), selection for a branched inflorescence seems to have been common, in tomato breeding it is rare, presumably due to poor fruit set. Yet our results showed that two accessions with a branched inflorescence produced the largest number of fruits and had a significantly higher harvestable yield compared to other accessions with the same fruit size. Recent work by others has shown that mutations leading to mild branching can boost yield (Soyk et al., 2017). All wild accessions in our collection had inflorescence architectures belonging to categories 1 (simple), 2 (simple and forked) and 3 (forked), while in the cultivated accessions all 5 categories were found (Table S 1).

Conclusion

Current efforts in tomato breeding are focused on broadening the genetic basis of modern tomato, by introgression of favourable traits from old cultivars, land races and wild materials. We characterised a diverse collection of sequenced tomato accessions, by phenotypic analysis of important plant growth, yield and fruit quality traits and by genotyping the collection with respect to a number of mutations or variations in key genes underlying these traits. The results of this study can be used to select the most optimal sources for genetic studies of important agronomic and fruit quality traits and crop improvement.

Supplementary data

Table S1. Variation among the genotypes for crop growth related traits. Abscission zone (AZ) of fruit pedicels was analysed in 2014 in two greenhouse compartments (1 and 2) and categorized in three classes. Inflorescence architecture was evaluated in 2014 in one greenhouse compartment and genotypes were categorized in 5 classes. Vegetative outgrowth of the inflorescence (VOI) was evaluated in one greenhouse compartment in 2014 and scored from 1 to 5. Plant growth speed was evaluated based on the days from sowing to reach the wire in two compartments in 2014. Earliness of Flowering was measured by counting the number of nodes up to the first inflorescence in 2014 in two greenhouse compartments. Values for plant growth speed and flowering time represent the average of three plants. W:Wild. C:Cultivated

RF	Accession ID	Genotype name	Type	AZ category		Inflorescence branching	VOI	Plant growth speed		Earliness of flowering	
				2014 (1)	2014 (2)	2014 (1)	2014 (1)	2014 (1)	2014 (2)	2014 (1)	2014 (2)
RF_001	EA05097	Moneymaker	C	1.0	1.0	2	3	133.0	139.0	8.3	9.3
RF_002	EA05096	Ailsa Craig	C	1.0	1.0	2	1	145.0	146.0	9.3	8.3
RF_003	EA06086	Garderners Delight	C	1.0	2.0	2	3	155.0	160.0	-	9.7
RF_004	EA00465	Rutgers	C	1.0	1.0	1	1	185.0	135.0	8.3	8.3
RF_005	EA00325	Galina	C	1.0	1.0	1	1	129.7	136.3	8.7	9.7
RF_006R	EA00448	Ponderosa	C	2.0	-	3	1	136.3	143.7	8.3	8.5
RF_006Y	EA00448	Ponderosa	C	-	2.0	3	2	131.0	139.0	8.7	9.0
RF_007	EA00375	Katinka Cherry	C	1.0	1.0	5	1	139.0	131.0	8.3	11.3
RF_008	EA00371	John's big orange	C	1.0	1.0	3	1	-	-	6.0	6.5
RF_011	EA02617	All Round	C	1.0	1.0	3	3	128.3	131.0	9.0	9.0
RF_012	EA02724	Sonato	C	1.0	1.0	3	1	131.0	133.7	10.0	10.0
RF_013	EA03701	Cross Country	C	1.0	1.0	1	1	-	-	8.3	10.0
RF_014	EA03362	Lidi	C	2.0	1.0	5	5	117.0	117.0	9.7	10.5
RF_015	TR00003	Momatero	C	1.0	1.0	2	4	155.0	155.0	9.3	10.7
RF_016	EA01965	Rote Beere	C	1.0	1.0	4	4	159.7	143.7	8.7	8.0
RF_017R	EA03306	"*L. esculentum* Mill."	C	1.0	1.0	1	3	145.0	164.3	9.3	11.0
RF_017Y	EA03306	"*L. esculentum* Mill."	C	2.0	2.0	1	1	97.0	110.0	8.7	8.7
RF_018	EA01155	DANA	C	1.5	1.5	3	5	138.5	155.0	9.7	10.0
RF_019	EA01049	Large Pink	C	1.0	-	3	1	129.7	133.7	7.3	8.0
RF_020	EA03221	"*L. esculentum* Mill."	C	2.0	2.0	3	4	136.3	138.3	8.7	9.5

RF	Accession ID	Genotype name	Type	AZ category		Inflorescence branching	VOI	Plant growth speed		Earliness of flowering	
RF_021	EA03222	Bolivar *"L. esculentum"*	C	2.0	1.0	3	3	169.0	169.0	9.0	9.0
RF_022	EA04710	"*L. esculentum*"	C	1.0	1.0	1	3	139.0	137.0	10.7	11.3
RF_023	EA05170	"*L. esculentum*"	C	1.0	1.0	1	4	110.0	111.0	8.0	9.7
RF_024	EA00990	Jersey Devil	C	1.0	1.0	3	4	164.3	155.0	11.3	10.7
RF_026	EA00157	Polish Joe	C	2.0	2.0	3	3	133.0	120.3	7.7	8.7
RF_027	EA02054	Cal J TM VF	C	2.0	3.0	1	1	180.0	-	8.0	8.0
RF_028	EA05581	The Dutchman	C	1.0	2.0	3	3	111.0	112.3	10.3	9.3
RF_029	EA00027	Black Cherry	C	1.0	1.0	3	1	110.0	101.3	9.0	8.7
RF_030	EA01835	ANTO	C	2.0	1.5	3	3	110.0	114.3	8.0	8.0
RF_031	EA01854	Winter Tipe (nor)	C	-	-	2	3	146.0	139.0	10.0	11.3
RF_032	EA04243	Chang Li *"L. esculentum"*	C	1.0	-	-	-	131.0	131.0	9.0	8.3
RF_033	EA00892	Belmonte	C	2.0	1.5	3	3	169.0	159.7	8.0	8.7
RF_034	EA01088	Tiffen mennonite	C	1.0	1.0	3	3	139.0	160.0	7.7	9.3
RF_035	EA04939	Wheatley's Frost Resistant	C	1.0	-	-	-	141.0	155.0	10.3	11.7
RF_036	EA05701	"*L. esculentum*"	C	1.0	1.0	3	3	172.0	-	9.0	9.5
RF_037	EA05891	"*L. esculentum*"	C	2.0	1.0	1	1	133.7	139.0	9.0	10.0
RF_038	EA04828	Chih-Mu-Tao-Se	C	1.0	1.0	2	1	146.0	155.0	10.3	11.0
RF_039	EA00526	"*L. esculentum*"	C		1.0	1	1	169.0	169.0	11.7	10.5
RF_040	EA02655	ES 58 Heinz' *L. esculentum*	C	1.0	1.0	1	1	-	-	9.0	10.0
RF_041	EA04861	Dolmalik	C	1.0	1.0	3	4	136.3	141.0	9.0	9.5
RF_043	EA03058	*S. pimpinellifolium* (Jusl.) Mill.	W	2.0	2.0	1	1	119.0	127.0	10.0	10.5
RF_044	EA02994	N 481 -*S. pimpinellifolium* (Jusl.) Mill.	W	1.0	-	-	-	110.0	110.0	8.3	10.7
RF_045	EA00676	*S. pimpinellifolium*	W	1.0	1.0	1	1	139.0	180.0	9.7	10.0
RF_046	EA00674	*S. pimpinellifolium*	W	1.0	1.0	1	3	112.0	119.0	8.7	11.0
RF_047	TR00005	*S. pimpinellifolium*	W	1.0	1.0	1	3	106.0	117.0	9.0	10.3
RF_051	TR00007	*S. chmielewskii*	W	-	-	0	3	133.7	131.7	8.7	11.0

RF	Accession ID	Genotype name	Type	AZ category		Inflorescence branching	VOI	Plant growth speed		Earliness of flowering	
RF_052	EA00759	*S. chmielewskii*	W	-	-	0	0	131.0	131.0	-	-
RF_054	TR00024	*S. cheesmaniae*	W	1.0	1.0	2	1	146.0	155.0	10.7	11.3
RF_056	EA00729	*S. neorickii (prev: L. parviflorum)*	W	-	-	1	1	159.0	159.0	9.0	
RF_057	TR00025	*S. neorickii (prev: L. parviflorum)*	W	-	-	2	1	131.0	129.0	12.3	12.0
RF_058	TR00008	*S. arcanum (prev: L. peruvianum)*	W	-	-	1	3	105.0	117.0	8.7	
RF_064	TR00012	*S. chilense*	W	-	-	3	1	130.7	131.0	12.7	13.0
RF_065	TR00013	*S. chilense (prev: L. peruvianum)*	W	-	-	3	3	-	145.0	13.7	14.0
RF_066	TR00014	*S. habrochaites f. glabratum*	W	-	-	-	0	127.0	135.0	-	-
RF_067	TR00015	*S. habrochaites f. glabratum*	W	-	-	3	1	122.3	120.3	-	14.0
RF_068	TR00016	*S. habrochaites f. glabratum*	W	-	-	3	1	110.0	117.0	-	-
RF_071	EA00703	*S. habrochaites (prev: L. hirsutum f. glabratum)*	W	-	-	-	0	104.7	114.3	13.0	15.5
RF_072	EA00558	*S. habrochaites*	W	-	-	-	0	117.0	135.0	-	-
RF_074	EA00585	*S. pennellii*	W	-	-	3	1	139.0	142.0	8.3	8.3
RF_077	TR00018	Large Red Cherry	C	1.0	1.0	2	4	119.0	127.0	9.0	9.0
RF_078	EA00940	Porter	C	1.0	1.0	2	1	169.0	169.0	10.0	10.0
RF_088	TR00019	Bloody Butcher	C	1.0	1.5	3	4	141.3	139.0	8.7	10.3
RF_089	EA01019	Brandywine	C	1.0	-	1	1	139.0	142.0	9.3	9.7
RF_090	TR00020	Dixy Golden Giant	C	1.0	1.0	3	1	131.0	136.3	9.7	8.5
RF_091	EA01037	Giant Belgium	C	2.0	2.0	3	2	139.0	142.0	7.3	7.7
RF_093	TR00021	Kentucky Beefsteak	C	1.5	1.0	4	3	133.7	127.0	6.7	9.3
RF_094	TR00022	Marmande VFA	C	1.0	1.5	3	1	152.0	180.0	9.0	8.0
RF_096	TR00023	Thessaloniki	C	1.0	1.0	3	1	185.0	169.0	10.0	8.5
RF_097	EA01640	Watermelon Beefsteak	C	2.0		3	4	131.0	133.3	8.3	8.5
RF_102	TR00026	"var. *cerasiforme*"	C	1.0	2.0	1	3	121.0	117.0	9.3	9.0

RF	Accession ID	Genotype name	Type	AZ category		Inflorescence branching	VOI	Plant growth speed		Earliness of flowering	
RF_103	TR00027	"var. *cerasiforme*"	C	1.0	2.0	1	1	146.0	164.3	8.3	9.3
RF_201	EA00282	Blondokee	C	1.0	1.5	3	4	135.0	132.0	8.7	9.3
RF_202	EA00450	Purple Russian	C	2.0	1.5	3	3	119.7	127.0	8.3	9.0
RF_203R	EA00470	Snowstorm	C	2.0	2.0	1	1	126.0	129.7	7.7	7.0
RF_203Y	EA00470	Snowstorm	C	1.0	1.0	2	3	118.7	117.0	8.7	9.3
RF_204	EA00488	Taxi	C	1.0	1.0	1	1	-	-	7.0	7.0
RF_205	EA00907	Indian Striped	C	1.0	1.0	3	3	155.0	155.0	11.7	9.0
RF_206	EA00915	ABC Potato Leaf	C	1.5	1.0	3	4	169.0	149.7	7.7	10.0
RF_207	EA00999	Napoli	C	1.0	1.0	3	3	178.0	173.0	9.7	9.3
RF_208	EA01031	Eckert Polish	C	2.0	2.0	3	4	137.7	147.0	10.0	9.7
RF_209	EA01120	Brandywine OTV	C	1.0	1.0	3	1	152.0	155.0	8.7	8.3
RF_210	EA01165	Earlibell	C	1.0	1.0	3	4	152.0	169.0	8.0	9.3
RF_211	EA01284	Super Beafsteak	C	1.0	1.0	2	3	146.0	155.0	10.0	9.0
RF_212	EA01345	Glory of Moldova	C	2.0	2.0	1	1	-	-	9.0	9.5
RF_213	EA01349	Golden Queen	C	1.0	1.0	1	4	145.7	144.3	9.0	9.7
RF_214	EA01579	Heinz 1706	C	1.0	1.0	1	1		181.7	8.7	7.0
RF_215	EA01709	"Italian cherry tomato "	C	-	-	2	1	116.0	127.0	9.0	9.3
RF_216	EA01985	Sonora	C	1.0	1.0	1	3	-	-	9.3	9.0
RF_217	EA02015	Felsozsolcai	C	1.0	1.0	1	1	-	-	7.0	8.0
RF_218	EA02021	Mao Tao Shi Zi	C	1.0	1.0	1	3	119.0	127.0	8.0	8.7
RF_219	EA02550	Dippes Nz 802	C	1.0	1.0	3	1	136.3	144.3	9.0	9.0
RF_220	EA02660	Madara	C	2.0	2.0	2	1	127.0	131.0	8.3	9.3
RF_221	EA02734	Stam Bovoy	C	2.0	1.0	3	3	115.0	112.0	11.0	9.0
RF_222	EA03444	Agrappoli d'Inverno	C	1.0	1.0	1	1	169.0	169.0	10.7	10.3
RF_223	EA05480	Early Dwarf	C	1.0	1.0	1	1	141.0	151.3	8.3	8.0
RF_224	EA05520	San-Marzano	C	1.5	2.0	2	4	169.0		8.3	8.7

RF	Accession ID	Genotype name	Type	AZ category		Inflorescence branching	VOI	Plant growth speed		Earliness of flowering	
RF_225	EA05677	Tomate Ciruela	C	1.0	2.0	2	3	119.7	127.0	8.7	9.7
RF_226	EA05721	DL/67/248	C	1.0	1.0	5	1	113.0	119.0	11.7	11.3
RF_227	EA05732	Nagcarlan	C	1.5	1.0	2	1	118.7	132.3	8.3	8.3
RF_228	EA05789	V 145-E	C	1.0	1.0	2	3	-	-	8.0	8.0
RF_229	EA05979	Morne a L'Eau	C	2.0	2.0	3	4	155.0	159.7	8.3	9.0
RF_230	EA06631	Jaune Flamme	C	1.0	1.0	3	1	-	185.0	8.3	9.0
RF_231	EA06746	EZ 033	C	-	-	1	1	136.3	136.3	8.7	8.5
RF_232	EA06902	OH88119	C	3.0	3.0	2	1	-	-	9.3	7.5
RF_233	EA06929	NCEBR2	C	3.0	-	1	1	-	-	8.0	9.3
RF_234	EA06932	981136	C	3.0	3.0	0	1	-	-	7.3	7.7
RF_235	EA03075	Kecskemeti Koria Bibor	C	1.0	1.0	2	1	-	-	7.3	8.3
RF_236	EA01042	Grosse Cotelee	C	-	2.0	4	3	129.0	120.3	9.0	8.3
RF_237	EA06002	"var. cerasiforme"	C	1.0	-	3	5	135.0	133.0	14.7	15.7
RF_238	EA06699	RZ26	C	1.0	1.0	3	1	110.0	103.7	8.3	8.7

Table S2. Variation among genotypes for fruit characteristics. For fruit number, fruit weight and firmness, the number presented in each year is the average per plant, based on observations in two greenhouse compartments. For Brix the number in each year is the average of four fruits. Accession RF_071 failed to produce any fruits and therefore is omitted from this table.

RF	EU SOL ID	Genotype name	Geno type Type	Number of fruits	Number of fruits	Fruit Weight	Fruit weight	Brix	Brix	Firmness	Firmness	colour	shape
				2013	*2014*	*2013*	*2014*	*2013*	*2014*	*2013*	*2014*		
RF_001	EA05097	Moneymaker	C	48	80	51.8	31.6	4.6	5.2	49.3	56.5	red	round
RF_002	EA05096	Ailsa Craig	C	84.5	48.5	56.3	42.9	4.1	4.1	56.6	54.1	red	round
RF_003	EA06086	Gardeners Delight	C	133.5	162.5	15.9	13.2	4.9	6.5	39.4	54.8	red	round
RF_004	EA00465	Rutgers	C	7.5	17	157.4	151.1	4.7	4.7	50.6	44.8	red	Ox-heart
RF_005	EA00325	Galina yellow cherry	C	166.5	134	14.4	12.1	6.1	5.9	53.3	53.6	orange	round
RF_006	EA00448	Ponderosa	C	6	26	183.3	140.8	-	4.7	-	54.5	red/orange	round
RF_007	EA00375	Katinka Cherry	C	209.5	211.5	10.6	8.4	6.7	7.7	50.6	53.2	orange	round
RF_008	EA00371	John's big orange	C	34.5	29.5	59.7	53.5	3.7	4.2	55.2	64	orange	flat round
RF_011	EA02617	All Round	C	53	101	55.9	49.8	4.5	4.9	51.2	51.7	red	round
RF_012	EA02724	Sonato	C	56.5	73.5	60.4	57.3	4.3	4.8	53	50.2	red	round
RF_013	EA03701	Cross Country	C	66	85	35.6	37.6	4	4.3	57.3	48.3	red	round
RF_014	EA03362	Lidi	C	354	564	6.7	5.7	6.4	6.1	35.2	39.6	yellow	ovate
RF_015	TR00003	Momatero	C	23.5	35	130.7	170.7	5.1	4.55	51	59.8	pink	flat
RF_016	EA01965	Rote Beere	C	234.5	341.5	5.3	4.8	8.2	7.6	42.1	49.4	red	round
RF_017	EA03306	*"L. esculentum"*	C	93	-	2.4	-	9.8	-	36.9	-	red/yellow	round
RF_018	EA01155	DANA	C	28	50.5	126.9	83	4.5	4.7	42.5	33.2	red	flat
RF_019	EA01049	Large Pink	C	13	26.5	311.8	177.1	3.7	3.6	59.1	51	pink	flat
RF_020	EA03221	*"L. esculentum"*	C	10	21	299.3	285.2	5	4.4	50.8	55.8	light red	flat
RF_021	EA03222	Bolivar' Lycopersicon esculentum	C	26	35	146.2	118.7	4	3.7	55.2	65.4	red	round
RF_022	EA04710	*"L. esculentum"*	C	98	43	3.5	5.4	8	6.8	32.7	49.5	light red	round
RF_023	EA05170	*"L. esculentum"*	C	114	126.5	3.1	3.7	7.6	7.1	32.1	39	pink	round

RF	EU SOL ID	Genotype name	Geno type Type	Number of fruits	Number of fruits	Fruit Weight	Fruit weight	Brix	Brix	Firmness	Firmness	colour	shape
RF_024	EA00990	Jersey Devil	C	9	18.5	115.7	95.9	4.5	4.8	49.2	48.7	red	oxheart
RF_026	EA00157	Polish Joe	C	22.5	33	262.9	268.8	4.3	3.5	53.2	53.5	pink	heart
RF_027	EA02054	Cal J TM VF	C	65	40	46.9	45.6	4.5	4.9	57.7	61.8	red	round
RF_028	EA05581	The Dutchman	C	16	21	250.7	362.7	4.7	4.4	48.3	47.5	pink	flat
RF_029	EA00027	Black Cherry	C	141.5	233.5	13.8	15.6	5.8	6.9	48.4	55.3	dark pink	round
RF_030	EA01835	ANTO	C	27.5	32	148	238.4	4.5	5.2	58.9	55.8	red	flat
RF_031	EA01854	Winter Tipe (*nor*)	C	19.5	39	30.8	37.1	4.1	4.4	59.5	71.1	Green	round
RF_032	EA04243	Chang Li " *L.esculentum*"	C	69	63	15.6	14.9	5.3	5.6	49	48	yellow	round
RF_033	EA00892	Belmonte	C	15.5	21.5	332.6	303.1	3.8	4	52.1	47.1	dark pink	flat
RF_034	EA01088	Tiffen mennonite	C	10	36.5	181.8	104	5.5	3.9	52.8	55.2	pink	flat
RF_035	EA04939	Wheatley's Frost Resistant	C	70	88	15.8	10.1	4.6	5.3	41.4	42.1	pink	rectangular
RF_036	EA05701	"L. esculentum"	C	49	70.5	24.5	25	4.2	4.1	59.4	45.3	pink	round
RF_037	EA05891	"L. esculentum"	C	70	98	20	4.6	7.3	6.3	40	43.4	red	round
RF_038	EA04828	Chih-Mu-Tao-Se	C	25	47.5	89.4	109.5	4.7	4.6	43.3	49.2	pink	flat
RF_039	EA00526	"L. esculentum"	C	97	13.5	8.7	3.4	4.5	5.1	39.6	-	red	round
RF_040	EA02655	ES 58 Heinz " *L. esculentum*"	C	10.5	26.5	105.5	93.2	3.8	3.2	55	58.3	red	round
RF_041	EA04861	Dolmalik	C	74.5	56	64	48.7	4.9	4.8	46.6	39.6	red	flat
RF_043	EA03058	*S. pimpinellifolium (Jusl.)*	W	28	65	24.8	22.5	4.1	4.6	59.3	48.7	dark pink	ovate
RF_044	EA02994	*N 481 -S. pimpinellifolium (Jusl.)*	W	518	171	1.1	1.6	7.4	7.1	33.6	48.4	red	round
RF_045	EA00676	*S. pimpinellifolium*	W	99	107	9.9	6.3	4.375	4.2	47.1	51.9	red	round
RF_046	EA00674	*S. pimpinellifolium*	W	36	114	2	1.2	7.2	7.8	28.9	43.9	red	round
RF_047	TR00005	*S. pimpinellifolium*	W	71.5	109	16.5	1.1	-	8.6	31.6	39.5	red	round
RF_051	TR00007	*S. chmielewskii*	W	16	24	1.62	1.8	-	11.2	37.6	51.4	green	round
RF_052	EA00759	*S. chmielewskii*	W	0	0	-	-	-	-	-	-		
RF_054	TR00024	*S. cheesmaniae*	W	189	150.5	3.3	4.2	8.1	7.5	40.6	47.7	red	round

RF	EU SOL ID	Genotype name	Geno type Type	Number of fruits	Number of fruits	Fruit Weight	Fruit weight	Brix	Brix	Firmness	Firmness	colour	shape
RF_056	EA00729	*S. neorickii (prev: L. parviflorum)*	W	4	0	1.375	-	-	-	-	-	green	round
RF_057	TR00025	*S. neorickii (prev: L. parviflorum)*	W	59	8.5	8.4	1.1	-	9.2	-	19.6	green	round
RF_058	TR00008	*S. arcanum (prev: L. peruvianum)*	W	6.5	14	1.8	1.7	-	9.6	-	-	green	round
RF_064	TR00012	*S. chilense*	W	-	90.5	-	3.8	-	8	-	54.8	purple	round
RF_065	TR00013	*S. chilense (prev: L. peruvianum)*	W	17	7	1.3	1.7	8.4	8.5	-	-	green	round
RF_066	TR00014	*S. habrochaites f. glabratum*	W	8.5	-	2.7	-	7.4		43	-	green	round
RF_067	TR00015	*S. habrochaites f. glabratum*	W	-	11	-	4.5	-	9.3	-	69.5	green	round
RF_068	TR00016	*S. habrochaites f. glabratum*	W	7.5	6	1.8	2.1	-	8.5	56.7	-	-	-
RF_072	TR00017	*S. habrochaites*	W	17.5	-	4	-	7.9	-	71.8	-	green	round
RF_074	EA00585	*S. pennellii*	W	-	4.5	-	2.1	-	5.2	-	-	green	round
RF_077	TR00018	Large Red Cherry	C	122.5	90.5	26.2	24.4	5.1	5.2	47.9	50.6	red	round
RF_078	EA00940	Porter	C	42.5	75.5	25.7	17.3	4.2	4.1	49.1	47.5	pink	round
RF_088	TR00019	Bloody Butcher	C	53	48.5	47.3	40.3	5.5	5.3	44.05	39.6	red	flat
RF_089	EA01019	Brandywine	C	18	10.5	228.5	234.2	4	5.3	51.3	54.3	dark pink	round
RF_090	TR00020	Dixy Golden Giant	C	15	14.5	261.8	244.4	4.6	5.1	43.1	43.5	orange	flat
RF_091	EA01037	Giant Belgium	C	16	25	252.1	245.9	4	4.1	50.7	47.4	dark pink	flat
RF_093	TR00021	Kentucky Beefsteak	C	8	20.5	218.4	268.4	3.6	3.8	51.8	47.2	orange	flat
RF_094	TR00022	Marmande VFA	C	24.5	46	117.06	62.6	4.5	4.2	51.7	50.3	red	flat
RF_096	TR00023	Thessaloniki	C	18.5	19	125.8	128.7	4.7	4.7	47.4	32	red	flat
RF_097	EA01640	Watermelon Beefsteak	C	5	6.5	303.3	308.1	5.2	5	43.1	47.9	pink	ovate
RF_102	TR00026	"var. *cerasiforme*"	C	119	89.5	2.3	11.3	8.6	7.1	24.4	45.2	red	round
RF_103	TR00027	"var. *cerasiforme*"	C	14	42	7.7	8.7	3.5	3.4	43.8	56.1	red	round
RF_201	EA00282	Blondokee	C	16	24	172.6	248.5	4.4	4.9	53.2	42.2	orange stripe	round flat

RF	EU SOL ID	Genotype name	Geno type Type	Number of fruits	Number of fruits	Fruit Weight	Fruit weight	Brix	Brix	Firmness	Firmness	colour	shape
RF_202	EA00450	Purple Russian	C	38	63	80.4	94.9	4.7	5.2	49.4	47	dark red	ovate
RF_203	EA00470	Snowstorm	C	28.5	48.5	4.9	6.9	7	5.2	26.9	39.4	red/yellow	round
RF_204	EA00488	Taxi	C	51.5	34.5	89.5	69.2	4	4	50.6	46.3	yellow	round
RF_205	EA00907	Indian Striped	C	26.5	20	182.2	198.2	4.3	3.8	43	48.5	dark red	round
RF_206	EA00915	ABC Potato Leaf	C	61.5	54.5	22.3	19.3	6.6	6.1	46.7	46.6	pink	flat
RF_207	EA00999	Napoli	C	55	48	34.4	26.8	4.1	4.3	54	51.1	red	Ox-heart
RF_208	EA01031	Eckert Polish	C	23	18	225.7	305.3	4.6	4.7	54.7	48.2	red	flat
RF_209	EA01120	Brandywine OTV	C	56	51	81	38.4	4.3	4.2	60.8	62.7	red	heart
RF_210	EA01165	Earlibell	C	73	59.5	71.5	77.9	3.8	3.5	45.8	47.4	red	flat
RF_211	EA01284	Super Beafsteak	C	29.5	35.5	80.3	89.8	4.7	3.9	53	59.07	red	flat
RF_212	EA01345	Glory of Moldova	C	90	71	32.4	46.8	4.7	4.6	52.5	54.5	orange	round flat
RF_213	EA01349	Golden Queen	C	10	28.5	97.3	121.4	-	5	-	50.3	orange	round
RF_214	EA01579	Heinz 1706	C	69	43.5	33.02	25.4	4.3	4.3	57.1	54.8	red	rectangular
RF_215	EA01709	"Italian cherry tomato "	C	-	67	-	7.9	-	5.9	-	53.3	green	round
RF_216	EA01985	Sonora	C	27.5	25	63.7	75.9	4.7	4.3	61.8	42.4	red	flat
RF_217	EA02015	Felsozsolcai	C	14	43.5	42.8	46.6	4.3	4.2	53.2	53	red	round
RF_218	EA02021	Mao Tao Shi Zi	C	21	41	39.02	42.1	4.3	4.3	42.9	41.2	pink	flat
RF_219	EA02550	Dippes Nz 802	C	41	75	58.08	40.8	3.9	4	51.5	52.1	red	round
RF_220	EA02660	Madara	C	88.5	88.5	19.1	15.7	5.7	6.8	62.4	61.9	red	ovate
RF_221	EA02734	Stam Bovoy	C	95	52.5	49.03	56.8	3.9	3.9	49.2	57.9	red	flat
RF_222	EA03444	Agrappoli d'Inverno	C	140.5	139.5	15.8	13.1	4.5	4.7	53.5	49.1	red	round
RF_223	EA05480	Early Dwarf	C	26	36.5	80.5	59.6	4.2	4.9	54.4	51.3	red	flat
RF_224	EA05520	San-Marzano	C	30	40	42.5	33.3	4.6	5	62.7	49.2	red	rectangular
RF_225	EA05677	Tomate Ciruela	C	29.5	56.5	40.7	38.2	4.7	4.9	49.3	43.5	red	oxheart
RF_226	EA05721	DL/67/248	C	-	533.5	6.6	7.3	4.5	6.3	61.1	53.5	red	round

RF	EU SOL ID	Genotype name	Geno type Type	Number of fruits	Number of fruits	Fruit Weight	Fruit weight	Brix	Brix	Firmness	Firmness	colour	shape
RF_227	EA05732	Nagcarlan	C	85.5	113.5	32.1	26.8	4.3	4.8	56.6	45.9	pink	round
RF_228	EA05789	V 145-E	C	31.5	27	78.1	64.9	4.2	4	61.1	57.8	red	heart
RF_229	EA05979	Morne a L'Eau	C	73.5	46.5	23.6	20.3	4.4	4.6	37.9	38.7	dark pink	flat
RF_230	EA06631	Jaune Flamme	C	89.5	74.5	36.9	31.9	5.2	4.6	40.9	52.7	orange	round
RF_231	EA06746	EZ 033	C	-	22	-	85.1	-	3.9	-	80.8	green	flat
RF_232	EA06902	OH88119	C	52	84	60.1	40.7	4.1	4.5	75.1	65.3	red	ellipsoid
RF_233	EA06929	NCEBR2	C	11	23.5	171.9	81.7	3.7	4.3	69.7	53.8	dark pink	round
RF_234	EA06932	981136	C	33	21	45.8	53.8	-	4.2	-	64.1	-	-
RF_235	EA03075	Kecskemeti Koria Bibor	C	34.5	45.5	95.2	72.3	3.7	3.7	55.9	52.9	red	flat
RF_236	EA01042	Grosse Cotelee	C	26.5	41	172.4	155.3	4.2	3.8	45.4	46.5	dark pink	flat
RF_237	EA06002	"var. *cerasiforme*"	C	81	86.5	1.8	1.71	9.3	7.6	30.6	39.7	red	round
RF_238	EA06699	RZ26	C	412	227	6.5	8.8	7.9	8.8	74.5	61	red	ellipsoid

Table S3. List of mutations and variants screened.

Abbreviation	Name	Genome location (SL2.40)	Type/effect	(nearest) gene	Gene name	Reference
c	*potato leaf*	ch06:42805810	Rider insertion in open reading frame	Solyc06g074910	*Bli-2*	Busch et al., 2011
fas	*fasciated*	ch11:51960800-52254400	294 kB inversion	Solyc11g071380/Solyc11g071810	*CLV3/YABBY*	Xu et al., 2015
fw2.2	*fruit weight 2.2*		promoter SNP	Solyc02g090730	*ORFX/CNR*	Frary et al., 2008
fw3.2	*fruit weight 3.2*		promoter SNP	Solyc03g114940	*KLUH*	Chakrabarty et al., 2014
fw11.3	*Fruit weight 11.3*	ch11: 51960800-52254400	1.4 kB deletion and 22 nt insertion	Solyc11g071940	*CSR*	Mu et al., 2017
gf	*green flesh*	ch08:60581508	nonsense mutation	Solyc08g080090	*SGR1*	Barry et al., 2009
lc	*locule number*	ch02:41766348,41766354	intergenic SNPs	Solyc02g083950	*WUS*	Muños et al., 2011
j-2	*jointless-2*	ch12:35900973	Rider insertion in 1st intron	Solyc12g038510	*MBP21*	Soyk et al., 2017
nor	*non-ripening*	ch10:1306295,1306296	2 nt deletion/frameshift	Solc10g006880	*NAC-NOR*	Eriksson et al., 2004
o	*ovate*	ch02:42946735	nonsense mutation	Solyc02g085500	*O*	Liu et al., 2002
og^c	*old-gold crimson*	ch06:42289512	1 nt deletion/frameshift	Solyc06g074240	*CYC-B*	Ronen et al., 2000
r^y	*yellow flesh*	ch03: 8610000-8616000	6 kB deletion from 3' of orf	Solyc03g031860	*PSY1*	Aflitos et al., 2014
s	*compound inflorescence*	ch02:36915588	Non-synonymous substitution	Solyc02g077390	*WOX9*	Lippman et al., 2008
sp	*self-pruning*	ch06:42362163	Non-synonymous substitution	Solyc06g074350	*SP*	Pnueli et al., 1998
sun		ch07:2395260	duplication/translocation	Solyc07g007760/Solyc10g079240	*DEFL1/SUN*	Xiao et al., 2008
t^3183	*tangerine*	ch10:62011800-62012200	384 nt deletion in promoter	Solyc10g081650	*CRTISO*	Isaacson et al. 2002
u	*uniform*	ch10:2295926	1 nt deletion/frameshift	Solyc10g008160	*GLK2*	Powell et al., 2012
ug	*uniform gray-green*	ch01:82279043	Non-synonomous substitution	Solyc01g100510	*tKN4*	Nadakuduti et al., 2014
y	*yellow*	ch01:71250000	603 nt promoter deletion	Solyc01g079620	*MYB12*	Lin et al., 2014
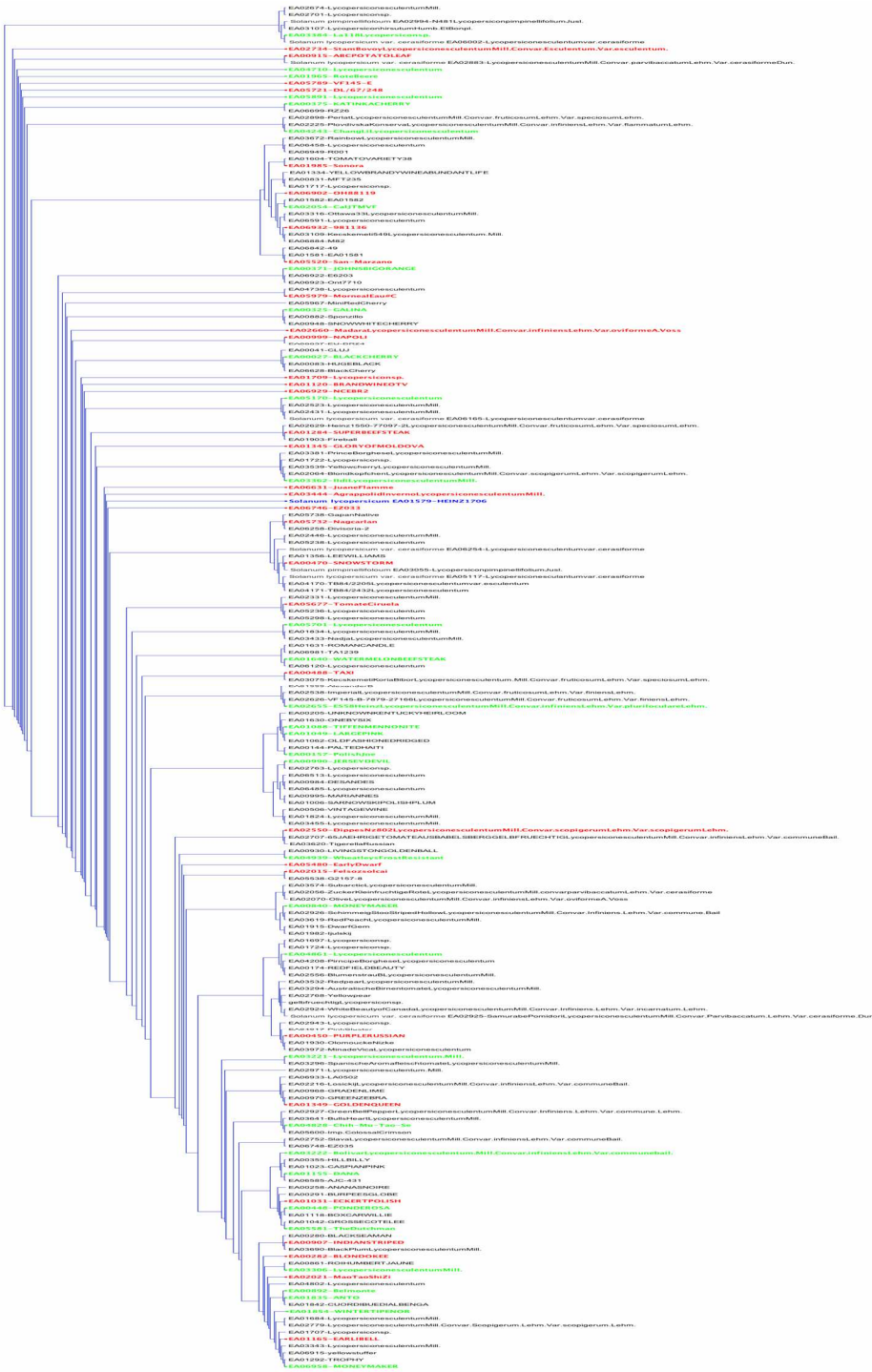
Figure S 1. Neighbour Joining tree, on the basis of approximately 7500 markers. In green, 33 accessionss sequenced in the 150+ tomato genome (re)seq are indicated (Aflitos 2014) while the 34 accessions additionally selected from this tree for this project are shown in red. The positon of the tomato genome reference cultivar (Heinz 1706) is highlighhted in blue.
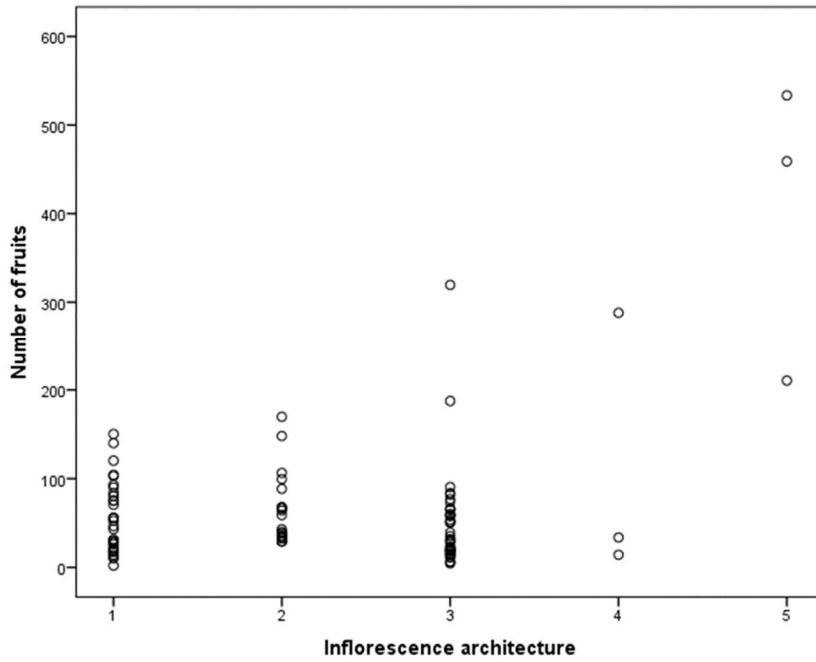
Figure S2. Scatter plot of the relationship between inflorescence architecture (categories 1 to 5) and number of fruits. One-way ANOVA with Post Hoc Tuckey analysis revealed that there is no statistical difference in fruit number among Inflorescence architecture categories 1 to 4, while plants with inflorescence type 5 have significantly increased fruit numbers.
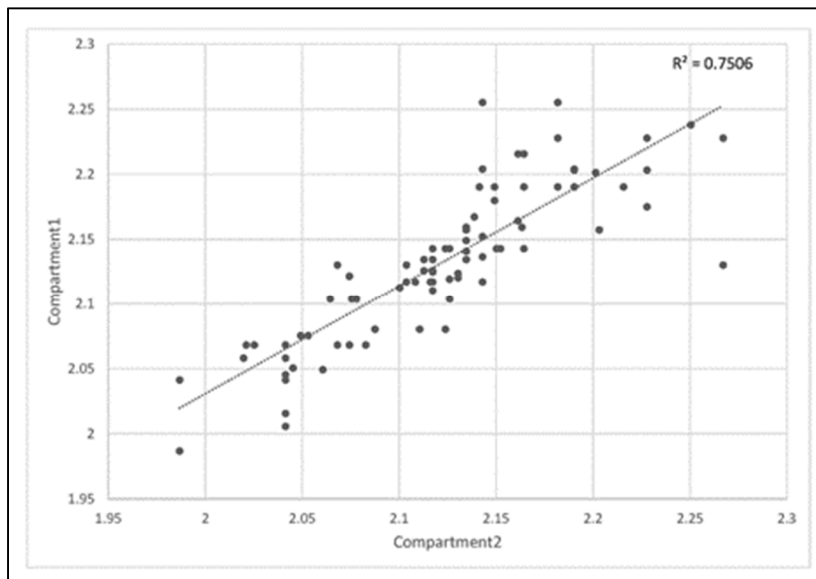
Figure S 3. Scatter plot of plant growth speed between two compartments. Numbers indicate the days until the plants reach the crop attachment wire. Values have been transformed to Log scale (base 10).
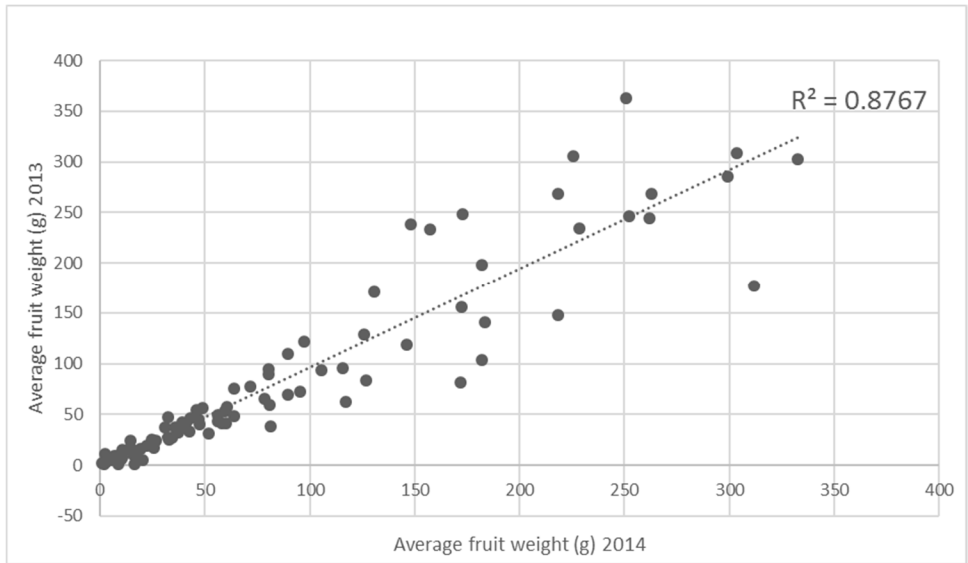
Figure S 4. Scatter plot of fruit weight (g) between 2013 and 2014 seasons.
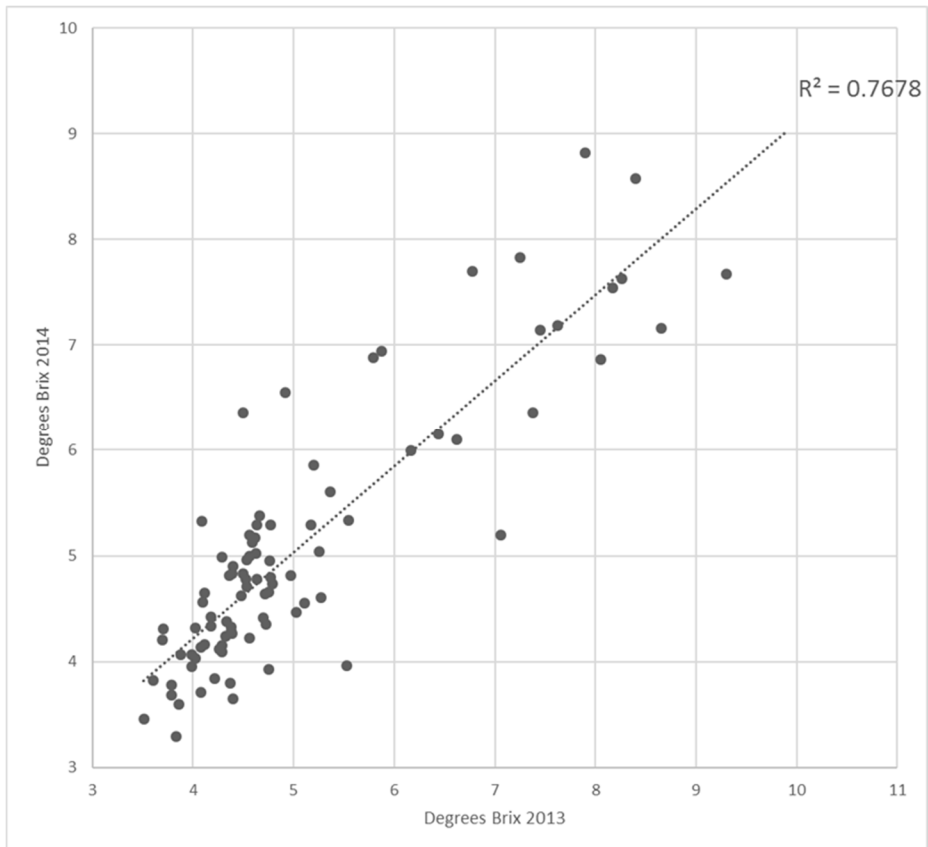
Figure S 5. Scatter plot of Brix content between 2013 and 2014 seasons.
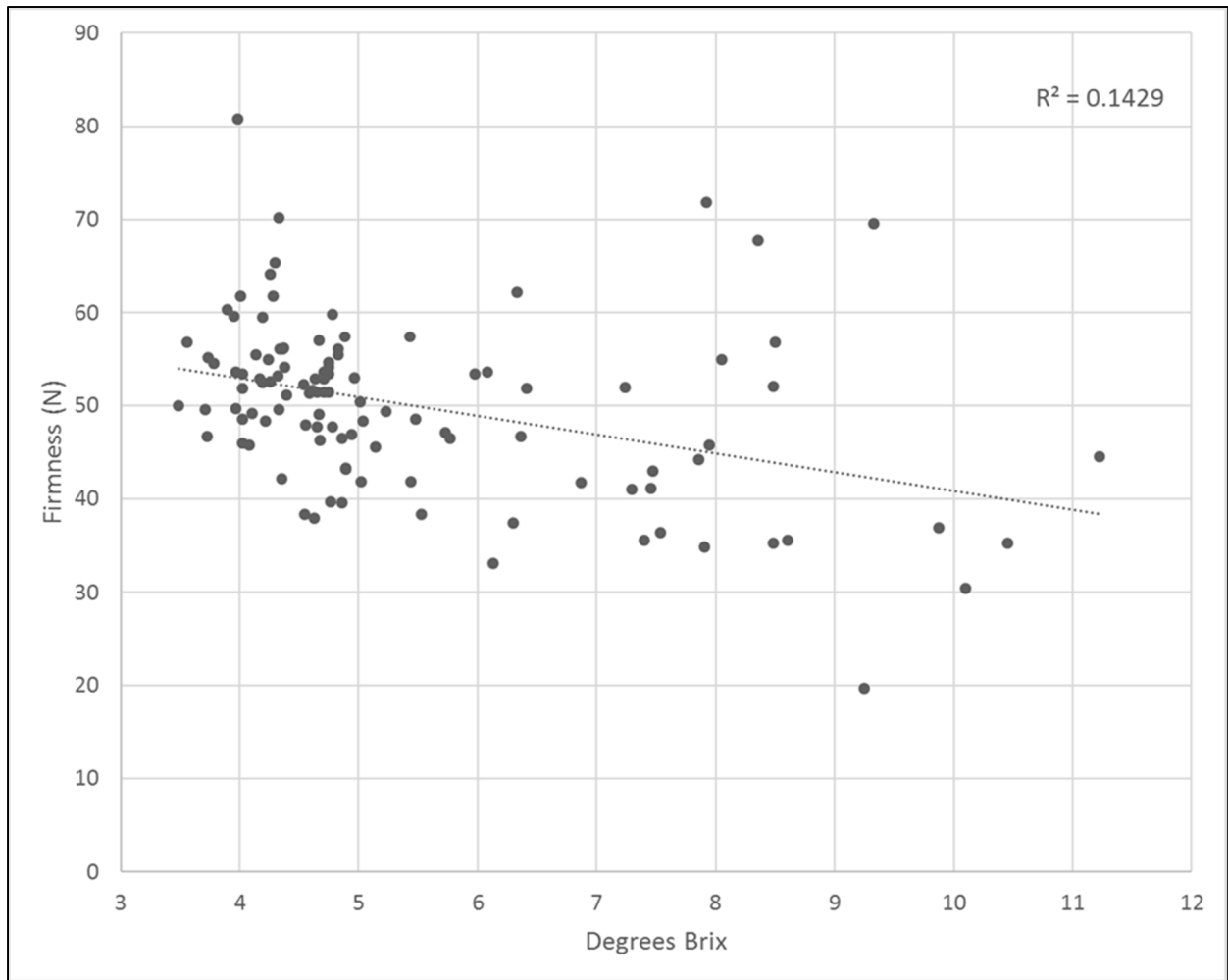
Figure S 6. Relationship between fruit Brix and fruit Firmness.
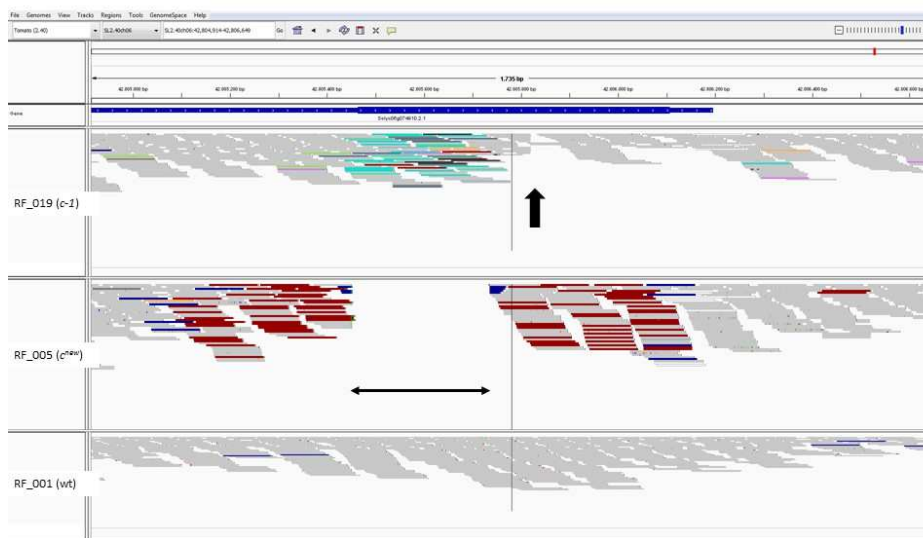
A
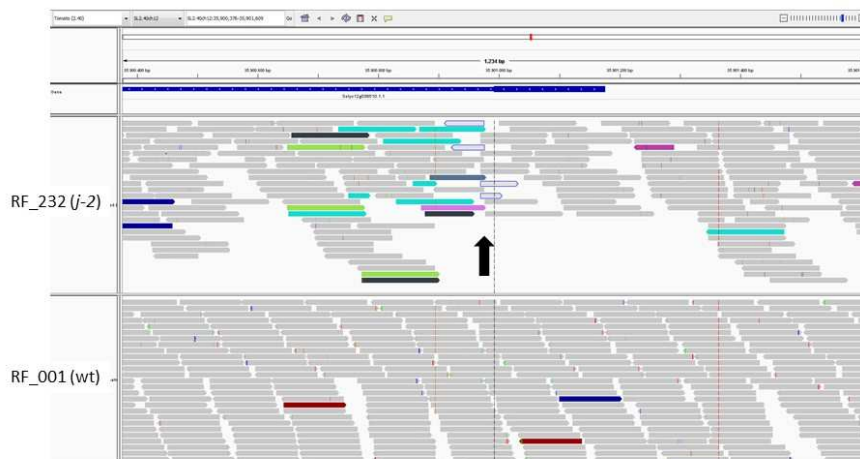
B
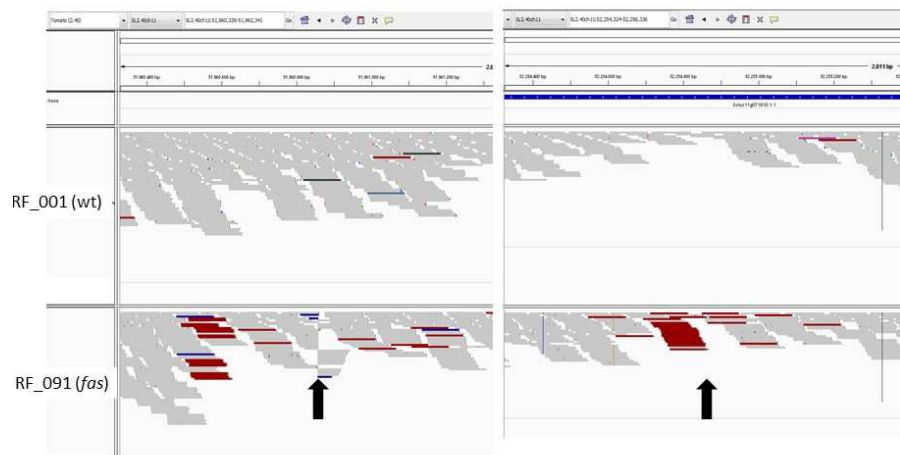
C
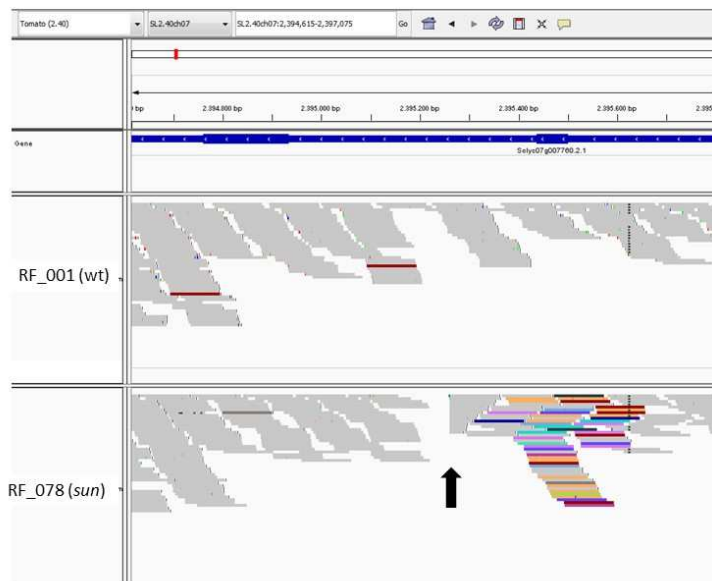
D
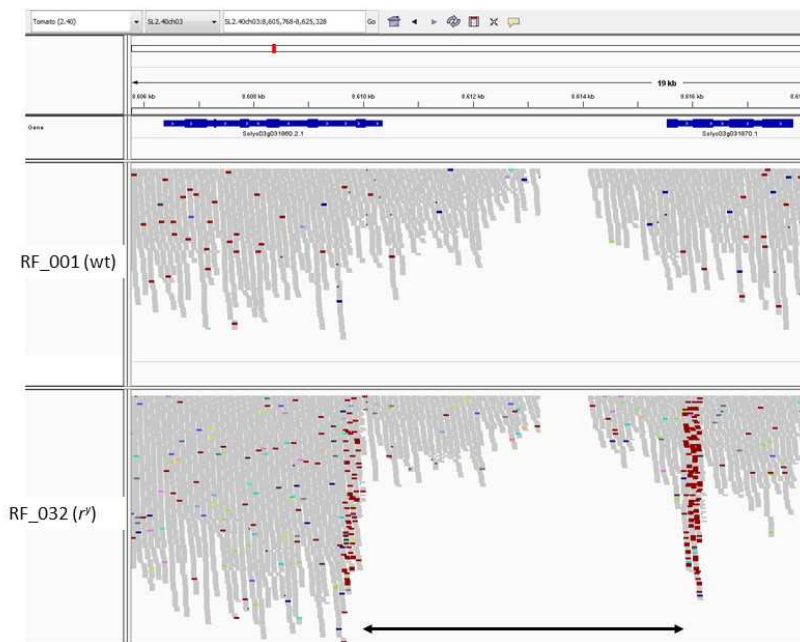
E
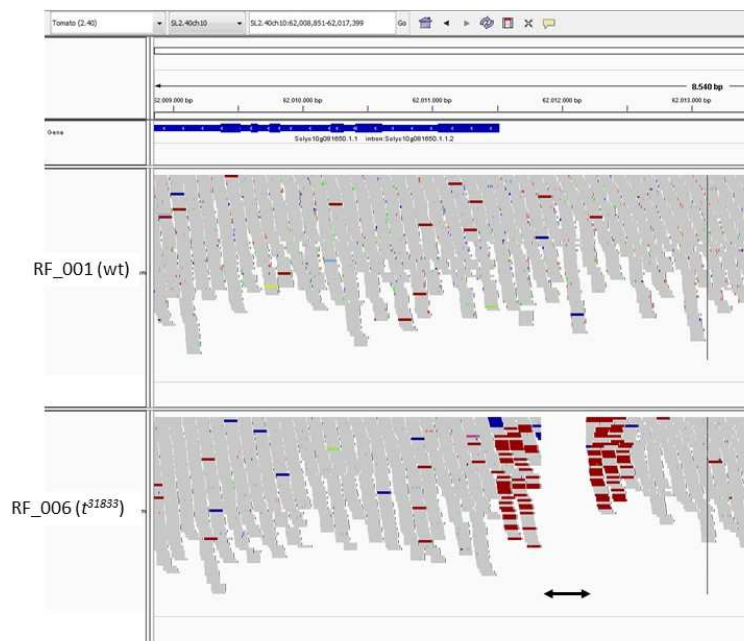
F

Figure S 7. IGV output of read alignments demonstrating the presence and nature of structural variations leading to mutations: **A.** *potato leaf* (*c*). **B.** *jointless-2* (*j-2*). **C.** *fasciated (fas).* **D.** *sun*. **E.** *yellow flesh* (*r^y*). **F.** *tangerine* (*t^{3183}*). Vertical arrows indicate insertion positions in various mutants or the break sites of the inversions in *fas* Horizontal double-headed arrows indicate the extent of deletions in mutations.

Figure S 8. Potato leaf mutant allele (c) in cv. Galina (RF_005).

References

Adato, A., Mandel, T., Mintz-Oron, S., Venger, I., Levy, D., Yativ, M., . . . Aharoni, A. (2009). **Fruit-surface flavonoid accumulation in tomato is controlled by a SlMYB12-regulated transcriptional network.** *PLoS Genet, 5*(12), e1000777.

Aflitos, S., Schijlen, E., de Jong, H., de Ridder, D., Smit, S., Finkers, R., . . . Peters, S. (2014). **Exploring genetic variation in the tomato (Solanum section Lycopersicon) clade by whole-genome sequencing.** *Plant J, 80*(1), 136-148.

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

Ballester, A. R., Molthoff, J., de Vos, R., Hekkert, B., Orzaez, D., Fernandez-Moreno, J. P., . . . Bovy, A. (2010). **Biochemical and molecular analysis of pink tomatoes: deregulated expression of the gene encoding transcription factor SlMYB12 leads to pink tomato fruit color.** *Plant Physiol, 152*(1), 71-84.

Barry, C. S., & Pandey, P. (2009). **A survey of cultivated heirloom tomato varieties identifies four new mutant alleles at the green-flesh locus.** *Mol Breed, 24*(3), 269-276.

Bartley, G. E., Scolnik, P. A., & Giuliano, G. (1994). **MOLECULAR-BIOLOGY OF CAROTENOID BIOSYNTHESIS PLANTS.** *Annu Rev Plant Physiol Plant Mol Biol., 45*, 287-301.

Blanca, J., Cañizares, J., Cordero, L., Pascual, L., Diez, M. J., & Nuez, F. (2012). **Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato.** *PLoS One, 7*.

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., . . . Cañizares, J. (2015). **Genomic variation in tomato, from wild ancestors to contemporary breeding accessions.** *BMC Genom, 16*(1), 257.

Bramley, P. M.(2013). **Carotenoid Biosynthesis and Chlorophyll Degradation.** In: Seymour G, Mervin Poole M, Giovannoni J, Tucker GA (eds). *The Molecular Biology and Biochemistry of Fruit Ripening*. Blackwell Publishing Ltd.: Oxford, 2013, pp 75–116.

Brewer, M. T., Moyseenko, J. B., Monforte, A. J., & van der Knaap, E. (2007). **Morphological variation in tomato: a comprehensive study of quantitative trait loci controlling fruit shape and development.** *J Exp Bot, 58*(6), 1339-1349.

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

Causse, M., Desplat, N., Pascual, L., Le Paslier, M.-C., Sauvage, C., Bauchet, G., . . . Bouchet, J.-P. (2013). **Whole genome resequencing in tomato reveals variation associated with introgression and breeding events.** *BMC Genom, 14*(1), 791.

Causse, M., Duffe, P., Gomez, M. C., Buret, M., Damidaux, R., Zamir, D., . . . Rothan, C. (2004). **A genetic map of candidate genes and QTLs involved in tomato fruit size and composition.** *J Exp Bot, 55*(403), 1671-1685.

Chakrabarti, M., Zhang, N., Sauvage, C. Muños, S. Blanca, J. Cañizares, J. (2013). **A cytochrome P450 regulates a domestication trait in cultivated tomato.** *Proc Natl Acad Sci U S A*, 110, 17125–30.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., . . . Ruden, D. M. (2012). **A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3.** *Fly (Austin), 6*(2), 80-92.

Cong, B., Barrero, L. S., & Tanksley, S. D. (2008). **Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication.** *Nat Genet, 40*, 800-804.

Cooperstone, J. L. (2014). **Tangerine Tomato Caretenoids: Processing, Structure, Bioavailability and Biological Implications of Consumption.** *(The Ohio State Unive. PhD dissertation)*.

Fernandez-Moreno, J. P., Tzfadia, O., Forment, J., Presa, S., Rogachev, I., Meir, S., . . . Granell, A. (2016). **Characterization of a New Pink-Fruited Tomato Mutant Results in the Identification of a Null Allele of the SlMYB12 Transcription Factor.** *Plant Physiol, 171*(3), 1821-1836.

Finkers, R., Causse, M., Giuliano , G., Klein Lankhorst, R., & Zamir, D. (2018). **SolCap Genotyping of approx. 400 tomato and tomato crop wild relative genotypes.** *Zenodo, Open acess*.

Fray, R. G., & Grierson, D. (1993). **Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing, complementation and co-suppression.** *Plant Mol Biol, 22*(4), 589-602.

Fridman, E., Liu, Y. S., Carmel-Goren, L., Gur, A., Shoresh, M., Pleban, T., . . . Zamir, D. (2002). **Two tightly linked QTLs modify tomato sugar content via different physiological pathways.** *Mol Genet Genomics, 266*(5), 821-826.

Gonzalo, M. J., & van der Knaap, E. (2008). **A comparative analysis into the genetic bases of morphology in tomato varieties exhibiting elongated fruit shape.** *Theor Appl Genet, 116*(5), 647-656.

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

Giuliano, G., Bartley, G. E., Scolnik, P. A. (1993). **Regulation of carotenoid biosynthesis during tomato development.** *Plant Cell* 1993, 5, 379–387.

Isaacson, T. (2002). **Cloning of tangerine from Tomato Reveals a Carotenoid Isomerase Essential for the Production of beta-Carotene and Xanthophylls in Plants.** *Plant Cell 14*(2), 333-342.

Kachanovsky, D. E., Filler, S., Isaacson, T., & Hirschberg, J. (2012). **Epistasis in tomato color mutations involves regulation of phytoene synthase 1 expression by cis-carotenoids.** *Proc Natl Acad Sci U S A, 109*(46), 19021-19026.

Kerr, E. **Green flesh, gf. (1956).** *Rep Tomato Genet Coop*, 6, 17.

Ku, H. M., Doganlar, S., Chen K. Y., & Tanksley, S. D. (1999). **The genetic basis of pear-shaped tomato fruit.** *Theor Appl Genet*, 99,844–850.

Ku, H. M., Grandillo, S., & Tanksley S. D. (2000). *fs8.1***, a major QTL, sets the pattern of tomato carpel shape well before anthesis.** *Theor Appl Genet*, 101, 873–878.

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., & Zheng, Z. (2014). **Genomic analyses provide insights into the history of tomato breeding.** *Nat Genet, 46*, 1220-1226.

Lindstrom, W. **Inheritance in tomatoes.** (1925). *Genetics*, 10, 305–317.

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

Lippman, Z. B., Cohen, O., Alvarez, J. P., Abu-Abied, M., Pekker, I., Paran, I., . . . Zamir, D. (2008). **The making of a compound inflorescence in tomato and related nightshades.** *PLoS Biol, 6*(11), e288.

Liu, J., Eck, J., Cong, B., & Tanksley, S. D. (2002). **A new class of regulatory genes underlying the cause of pear-shaped tomato fruit.** *Proc Natl Acad Sci U S A, 99*, 13302-13306.

Miller, J. C., & Tanksley, S. D. (1990). **RFLP analysis of phylogenetic relationships and genetic variation in the genus Lycopersicon.** *Theor Appl Genet, 80*(4), 437-448.

Monforte, A. J., Diaz, A., Cano-Delgado, A., & van der Knaap, E. (2014). **The genetic basis of fruit morphology in horticultural crops: lessons from tomato and melon.** *J Exp Bot, 65*(16), 4625-4637.

Mu, Q., Huang, Z., Chakrabarti, M., Illa-Berenguer, E., Liu, X., Wang, Y., . . . van der Knaap, E. (2017). **Fruit weight is controlled by Cell Size Regulator encoding a novel protein that is expressed in maturing tomato fruits**. *PLoS Genet, 13*(8), e1006930.

Muños, S., Ranc, N., Botton, E., Bérard, A., Rolland, S., & Duffé, P. (2011). **Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL.** *Plant Physiol, 156*, 2244-2254.

Roldan, M. V. G., Périlleux, C., Morin, H. (2017). **Natural and induced loss of function mutations in *SlMBP21* MADS-box gene led to *jointless-2* phenotype in tomato.** *Sci Rep* 2017, 7, 4402.

Ronen, G., Carmel-Goren, L., Zamir, D., Hirschberg, J. (2000). **An alternative pathway to ß-carotene formation in plant chromoplasts discovered by map-based cloning of *Beta* and *old-gold* color mutations in tomato.** *Proc Natl Acad Sci U S A* , 97,11102–11107.

Paran, I., & van der Knaap, E. (2007). **Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper.** *J Exp Bot, 58*(14), 3841-3852.

Park, Y. H., West, M. A. L., & St Clair, D. A. (2004). **Evaluation of AFLPs for germplasm fingerprinting and assessment of genetic diversity in cultivars of tomato (Lycopersicon esculentum L).** *Genome, 47*, 647-656.

Pnueli, L., Carmel-Goren, L., Hareven, D. (1998). **The *SELF-PRUNING* gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of *CEN* and *TFL1*.** *Development*, 125, 1979–89.

Porrini, M., Riso, P., & Testolin, G. (1998). **Absorption of lycopene from single or daily portions of raw and processed tomato.** *Br J Nutr, 80*(4), 353-361.

Powell, A. L., Nguyen, C. V., Hill, T., Cheng, K. L., Figueroa-Balderas, R., Aktas, H., . . . Bennett, A. B. (2012). **Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development.** *Science, 336*(6089), 1711-1715.

Rick, C. M., & Chetelat, R. T. (1995). **utilization of related wild species for tomato improvement.** *Acta Horticuture, 412*, 21-35.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). **Integrative Genomics Viewer.** *Nat biotechnol, 29*(1), 24-26.

Rodriguez, G. R., Munos, S., Anderson, C., Sim, S. C., Michel, A., Causse, M., . . . van der Knaap, E. (2011). **Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity.** *Plant Physiol, 156*(1), 275-285.

Sahu, K. K., & Chattopadhyay, D. (2017). **Genome-wide sequence variations between wild and cultivated tomato species revisited by whole genome sequence mapping.** *BMC Genom, 18*(1), 430

Sim, S.-C., Van Deynze, A., Stoffel, K., Douches, D. S., Zarka, D., Ganal, M. W., . . . Francis, D. M. (2012**). High-Density SNP Genotyping of Tomato (Solanum lycopersicum L.) Reveals Patterns of Genetic Variation Due to Breeding.** *PLOS ONE, 7*(9), e45520.

Soyk, S., Lemmon, Z. H., Oved, M., Fisher, J., Liberatore, K. L., Park, S. J., . . . Lippman, Z. B. (2017). **Bypassing Negative Epistasis on Yield in Tomato Imposed by a Domestication Gene.** *Cell, 169*(6), 1142-1155.e1112.

Sun, L., Rodriguez, G. R., Clevenger, J. P., Illa-Berenguer, E., Lin, J., Blakeslee, J. J., . . . van der Knaap, E. (2015). **Candidate gene selection and detailed morphological evaluations of fs8.1, a quantitative trait locus controlling tomato fruit shape.** *J Exp Bot, 66*(20), 6471-6482.

Tanksley, S. D. (2004). **The genetic, developmental, and molecular bases of fruit size and shape variation in tomato.** *Plant Cell, 16 Suppl*, S181-189.

Tanksley, S. D., & McCouch, S. R. (1997). **Seed banks and molecular maps: unlocking genetic potential from the wild.** *Science (80-), 277*.

Thompson, A.E., Tomes ML, Wann EV, McCollum J, Stoner AK. (1965). **Characterization of crimson tomato fruit color.** *Proc Amer Soc Hort Sci*, 86, 610–616.

Tigchelaar, E. C., & Barman, R. J. (1985). **Allelism of the alcobaca ripening mutant and nor.** Rep Tomato Genet Coop, 35, 20.

Tomes, M. L. (1952). **Flower color modification associated with the gene t.** Rep Tomato Genet Coop, 2, 12.

Van der Knaap, E., & Tanksley, S. D. (2003). **The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in Yellow Stuffer tomato.** Theor Appl Genet, 107(1), 139-147.

Xiao, H., Jiang, N., Schaffner, E., Stockinger, E. J., & Knaap, E. (2008). **A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit.** Science, 319, 1527-1530.

Xu, C., Liberatore, K. L., MacAlister, C. A., Huang, Z., Chu, Y. H., Jiang, K., . . . Lippman, Z. B. (2015). **A cascade of arabinosyltransferases controls shoot meristem size in tomato.** Nat Genet, 47(7), 784-792.

Chapter 3

Evaluation of a diverse tomato collection for post-harvest shelf-life

Raana Roohanitaziani[1,2], Michiel Lammers[3], Jos Molthoff[1], Yury Tikunov[1], Richard G.F. Visser[1] , Jeroen van Arkel[3] and Arnaud G. Bovy[1]

[1] Plant Breeding, Wageningen University and Research, PO Box 386, 6700 AJ Wageningen, The Netherlands

[2] Graduate School Experimental Plant Sciences, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[3] Bioscience, Wageningen University & Research, PO Box 16, 6700 AA Wageningen, The Netherland

Abstract

In all fruit and vegetable crops, reduction in quality during postharvest storage leads to substantial losses of primary produce with enormous economic consequences. Also in tomato, fruit shelf-life is an important quality trait. In this study, a tomato collection consisting of 90 heirlooms and landraces was phenotyped for several shelf-life parameters and biochemical characterisation was performed during the postharvest shelf-life of fruits from selected accessions. The core collection was grown in a controlled greenhouse, fruits were harvested at breaker-turning stage and stored in a controlled climate chamber for 42 days at 18^0C. The shelf-life related parameters fruit firmness, weight loss and pigment production were measured once a week and evaluated over time. All three shelf-life related parameters varied markedly among genotypes, resulting in fruits with different shelf life. The most promising genotypes of the first screen were re-grown and analysed to validate the initial results and 6 genotypes with contrasting shelf-life were selected for metabolite analysis. Fruits were harvested at Breaker stage and stored for 35 days at $18\ ^0$C. Samples were taken at weekly intervals and analysed for volatile compounds, primary metabolites and cell wall polysaccharide monomers. During storage long and short shelf-life genotypes showed considerable differences in their content of sugars, such as galactose and polyamines, such as putrescine. Interpretation of cell wall composition data based on the ratios of pectin monosaccharides revealed that the pectin of firm genotypes is more linear and has a lower degree of branching compared to soft genotypes. This knowledge provides a better insight in the biochemical basis of long shelf life tomatoes.

Key words: Tomato, Shelf-life, Firmness, post-harvest, Cell wall, Metabolite

1 Introduction

Managing postharvest losses of fruits in modern agriculture is of great importance due to globalization of the markets and often very long distances between producers and consumers. In tomato the early solution to manage post-harvest losses has been picking green or unripe fruits instead of picking ripe fruits which are more prone to damage during the delivery period. Picking unripe fruits on the one hand can increase the fruit shelf life, but on the other hand comes with sacrifice of quality (Beckles, 2012). Today, various physical and chemical postharvest treatments are used to enhance the fruit shelf-life via slowing down physiological processes of senescence and maturation. However each of these treatments has its own limitations (Mahajan et al., 2014; Sandarani et al., 2018). The first breeding strategy to improve the tomato fruit shelf-life was use of ripening mutants and deploying their alleles in the development of hybrids with good quality and extended shelf-life. F1 hybrids such as rin/+, nor/+ and alc (alc/+) have been widely commercialized and they have played an important role in increasing the availability of "vine ripened" tomatoes in the market (Kopeliovitch et al., 1979; McGlasson et al., 1983; Nguyen et al., 1991; Markovic et al., 2012). The second breeding strategy which recently has gained more attention, is looking for varieties with extended shelf life, which are not ripening mutants but remain firm for an extended period of time after ripening and allow harvesting at a more advanced colour and thus riper stage. Tomato germplasm collections, fortunately, offer a valuable gene pool for both fruit and crop related traits and these sources of novel genes recently have been utilized for improvement of several traits and they have the potential to cope with this post-harvest quality challenge as well (Rick & Chetelat, 1995; Blanca et al., 2015).

Fruit shelf-life is determined by a series of physiological, biochemical and organoleptic changes that happen during ripening to make fruit edible and desirable for consumers. These changes generally include softening of the flesh by modification of the cell wall structure, alterations in pigment biosynthesis, increase in the levels of aromatic volatiles and nutrition, alteration in cuticle architecture and composition and increase in susceptibility to post-harvest pathogens (Giovannoni, 2001). Decrease in fruit firmness, resulting in fruit softening, is the result of dissolution of the middle lamella and structural changes in cell wall polysaccharides (mainly in pectin but also to a minor extent in hemicellulose and cellulose), in an ordered series of modifications by cell wall degrading enzymes (Brummell & Harpster, 2001; Brummell, 2005, 2006; Lahaye et al., 2012). Although many details about the cell wall modification and changes in the structure of polysaccharides during ripening of tomato fruit are still unclear, several earlier studies

have shown that loss of Galactorunic acid (Gal A), Galactose (Gal) and Arabinose (Ara) residues from the cell wall and concomitant increase in the solubility of Gal and Gal A in the pericarp occur upon ripening (Gross & Wallner, 1979; Labavitch, 1981; Gross, 1983; Gross & Sams, 1984; Prasanna et al., 2007). Gal A, Gal and Ara are major components of pectic polysaccharides and changes in their content during ripening have shown to be different from fruit to fruit and especially between normal ripening tomatoes and ripening tomato mutants (Gross & Wallner, 1979; Gross, 1983; Gross & Sams, 1984; Lunn et al., 2013). Pectin in the fruit cell wall comprises of acidic polysaccharides made of variable proportions of three structural domains: homogalacturonan (HG), rhamnogalacturonan I (RG-I) and rhamnogalacturonan II (RG-II). HG is built of linearly linked 1,4-α-d-galacturonic acids. RG-I is the major component of the primary cell wall and middle lamella in dicot plants and it is responsible for the structural diversity of pectin (Prasanna et al., 2007). The backbone of RG-I consists of repeating disaccharides of Rha and Gal A. Almost 50 percent of Rha residues in the backbone of RG-I are branched with side chains consisting of D-galactose and L-arabinose residues. RG-II is a HG substituted by complex side chains, some of which containing rhamnose. RG-II is invariably present as a minor component of the cell wall and has an extremely complex structure (Prasanna et al., 2007; Houben et al., 2011).

The aim of the present study was to evaluate a collection of tomato genotypes consisting of 90 land races and old cultivars for their fruit post-harvest shelf-life, to be able to introduce novel sources of improved post-harvest shelf-life into future breeding programs. Shelf-life attributes such as firmness, water loss and colour pigments were monitored during a 42 days storage period. The other objective of this study was to characterize and compare changes in cell wall sugar composition, primary metabolites and volatile compounds of fruits with long and short post-harvest shelf-life.

2 Materials & Methods

1.2 Plant materials and growth conditions

The collection for this project consisted of 90 tomato genotypes. 52 of these genotypes were selected from the 150-genome (re)sequencing project (Aflitos et al., 2014) and consisted of tomato landraces and heirloom varieties of *S. lycopersicum* and *S.lycopersicum* var. *cerasiforme.* The other additional 38 genotypes are *S. lycopersicum* and were selected from the EU–SOL tomato core collection (https://www.eu-sol.wur.nl) and by participation of companies involved in this project.

For the first shelf-life trial seeds were sown in spring 2014, mid-November, after a temperature treatment of 48 hours at 75˚C to prevent virus infection. Five weeks later the plants were transplanted to the

greenhouse. Three plants per genotype for round and cherry tomatoes and 6 plants per genotype for beef tomatoes were grown.

For the repetition of the experiment with promising accessions of 2014, seeds of 17 accessions were sown in mid-January 2015 and fruits were harvested from mid-May onwards. All genotypes were grown in three blocks, using a complete randomised block design. Each plot contained 2-6 plants of one accession grown alongside each other. In total, 15 plants per genotype were grown for beef tomatoes (five plants per plot in each block), nine plants per genotype for round (three plants per plot in each block) and six plants per genotype for cherry tomatoes (two plants per plot in each block) . Five fruits per genotype were harvested from each block at breaker-turning stage (in total 3 x 5=15 fruits per genotype) for the shelf-life analysis.

For biochemical characterization of the six selected genotypes, 30 plants per genotype were grown in the greenhouse as described for the trial in 2015. In total 45 fruits per genotype were harvested, six fruits in green stage and 39 fruits in breaker stage. From these 39 breaker fruits, 15 fruits per genotype were stored in a climate chamber (18 degrees °C, 85% Humidity) for 35 days and they were subjected to a shelf life trial (firmness, weight loss and pigment measurements) once a week. The other 30 fruits were used for metabolite analysis at different time points: Green, Breaker+1 (one day after storage), Breaker+14 (14 days after storage), Breaker+21 (21 days after storage) and Breaker +35 (35 days after storage).

2.2 Postharvest Shelf-life trial

Harvested fruits were stored in a controlled climate chamber (18 degrees °C, 85% Humidity) for 42 days. Firmness, fruit weight and chlorophyll and lycopene were measured once a week and evaluated over time. Firmness was measured as the average of four measurements per fruit around the mid-height circumference of the fruit using a handheld Fruit Hardness Tester (53215, Turoni Italy) and expressed as average score.

For weight loss measurement, each fruit was weighed with a precision balance (at least +/-10 mg).

A hand-held photodiode array spectrophotometer (Pigment Analyzer PA1101, CP, Germany) was applied to measure the chlorophyll and lycopene levels non-destructively. Remittance was assessed between 350 and 1100 nm to calculate the normalised difference vegetation index (NDVI) and the normalised anthocyanin index (NAI) which are indicators for chlorophyll and lycopene respectively and produces a normalized value between -1 and $+1$ (Schouten et al., 2014)**.** Measurements were carried out on two positions on the equator of each tomato and reported as an average NDVI and NAI value per tomato.

Measurements were performed at the day of harvest and subsequently once a week until the end of the storage period.

During the storage period, appearance of wrinkles or loss of shape and infections with pathogens were monitored every week. Before each measurement a picture was taken from each analysed fruit.

3.2 Measurements of volatile compounds

Six fruits from each genotype at each post-harvest time point were cut in small pieces and samples pooled together and immediately frozen in liquid nitrogen, ground in an analytical electric mill (IKA A11 basic, Germany), and stored at -80°C before analyses. Volatile compounds were quantified and identified using Gas Chromatography/Mass Spectrometry (GC/MS) as described in (Tikunov et al., 2005). Frozen fruit powder (1 g fresh weight) was weighed in a 5-mL screw-cap vial, closed, and incubated at 30 °C for 10 min. An EDTA-NaOH water solution was prepared by adjusting of 100mM EDTA to a pH of 7.5 with NaOH. Then, 1 mL of the EDTA-NaOH solution was added to the sample. 1.2 gram of solid CaCl2 was then immediately added to give a final concentration of 5 M. The closed vials were vortexed by hand for 1 min to allow CaCl2 to dissolve. A 1-mL aliquot was transferred into a 10-mL crimp cap vial (Waters), capped, and used for SPME-GC-MS analysis. In addition, a number of identical analytical quality control samples (QC) were made by mixing fruit powders from several genotypes. The QCs were routinely analysed every 10 samples during the entire analysis period as an external control in order to monitor the stability of the analytical system. The samples were automatically extracted and injected into the GC-MS via a Combi PAL auto sampler (CTC Analytics AG). Headspace volatiles were extracted by exposing a 65-mm polydimethylsiloxane-divinylbenzene SPME fiber (Supelco) to the vial headspace for 20 min under continuous agitation and heating at 50 °C. The fiber was inserted into a GC 8000 (Fisons Instruments) injection port and volatiles were desorbed for 1 min at 250 °C. Chromatography was performed on an HP-5 (50 m30.32 mm3 .05 mm) column with helium as carrier gas (37 kPa). The GC interface and MS 1 source temperatures were 260 °C and 250 °C, respectively. The GC temperature program began at 45 °C (2min), was then raised to 250 °C at a rate of 10 C/min, and finally held at 250 °C for 5min. The total run time, including oven cooling, was 30 min. Mass spectra in the 35 to 400 m/z range were recorded by an MD800 electron impact MS (Fisons Instruments) at a scanning speed of 2.8 scans/s and an ionization energy of 70 eV. The chromatography and spectral data were evaluated using Xcalibur software (http://www.thermo.com).

4.2 Primary metabolite analyses

The detection of amino acids, sugars and organic acids was performed by GC-TOF-MS. For extraction of samples 700 µl methanol containing 0.5 mg/ml ribitol (as an internal standard) was added to 300 mg

frozen fruit powder in a 2ml Eppendorf vial. Samples were vortexed for 20 minutes and centrifuged at max speed for 8 min. 500 µl of the methanol extract was transferred to a new vial to which 450 µl water and 250 µl of chloroform were added. The samples were shortly vortexed and centrifuged at max speed for 8 minutes. Supernatant was diluted 8 times with pure methanol. 40 µl of supernatant was transferred in to a crimp cap with insert and dried overnight in a speed vac and capped under argon. Samples were analysed by GC-TOF-MS as described in (Carreno-Quintero et al. 2012) using a detector voltage of 1700 V. Leco Chroma TOF software 2.0 was used for pre-processing of the raw data.

5.2 Preparation of cell wall (alcohol insoluble residue; AIR)

The preparation of cell wall material (alcohol insoluble residue; AIR) was performed based on the procedure described in (Uluisik et al., 2016). Briefly, fresh tomato pericarp (100 g) was peeled, seeds and jelly materials were separated, cubed and boiled in 95% EtOH (100 mL) at 80 °C for 30 min. The samples were cooled to room temperature, homogenised using a coffee grinder, then filtered through Miracloth and washed successively with hot 85% EtOH (200 mL), chloroform/methanol (1:1 v/v) (200 mL) and 100% acetone. The samples were then air dried overnight.

6.2 Preparation of monosaccharide fraction from cell walls

25-30 mg of each cell wall sample was weighed in a 15 ml pressure glass tube. Each sample was weighed in triplicate and the weight of each sample was recorded. The glass tubes were labelled with permanent marker. 3 ml of 72% sulfuric acid was added to the tube and vortexed until the sample was thoroughly mixed. The tubes were placed in thermomixer set at 30 °C for 60 minutes. After removing the tubes from the thermomixer, acid was diluted to 4% concentration by adding 11.00±0.04 ml MQ water (to a final volume of 12 ml). The samples were mixed by inverting the tubes several times to eliminate phase separation between high and low concentration acid layers. The samples were heated for 1 hour at 121 °C . After completion of autoclave cycle, samples were allowed to cool to near room temperature. 1 ml of each sample was filtered through a 0.2 µm syringe and transferred to vials with inserts for monosaccharide analysis. A set of sugar recovery standards (including glucose, galactose, arabinose, mannose, xylose, rhamnose, glucuronic acid and galacturonic acid) were prepared and analysed along with the samples.

7.2 Determination of monosaccharide composition of cell wall fractions

Monosaccharide composition was determined by high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) on a CarboPac PA1, 4- × 250-mm column preceded by a guard column (CarboPac PA 1, 4 × 50 mm) mounted in a DX-500 system, Dionex®. Saccharides were separated by elution in an increasing concentration of NaOH (5–20 mM) and sodium acetate (0-100mM)

with a flow rate of 1 ml per minute as described by (van Arkel et al., 2012) , with the described modifications in the eluents concentrations. Peaks were identified by co-elution of standards.

8.2 Statistical and data analyses

Principal component analysis (PCA), was performed in GeneMaths XT created by applied Maths NV (http://www.appliedmaths.com). Averages, standard deviations and LSD, were calculated by SPSS.

3 Results

1.3 Shelf-life trial

1.1.3 Firmness measurements

In the 2014 trial, the firmness scores measured by fruit hardness tester ranged from 36 (cv. Lidi, RF_014; cv. Rote Beere, RF_016) to 85 (cv. The Dutchman, RF_028; cv. Flora Dade, RF_ 140013; cv. S. esc. A0011-6-3, RF_140014) at the day of harvest and from 15 (cv. Kecskemeti Koria Bibor, RF_235) to 81 (cv. EZ 033-rin, RF-231) after 21 days (Table 1). Scores were calculated based on the average of five fruits for each accession. The firmness loss (FL) of each accession was evaluated during storage and is expressed as the average of the % of FL and was calculated as $FL=100-(F_t*100/F_0)$. Where F_t is the measured firmness at 21 days of storage and F_0 the measured firmness at the harvesting time. Based on their FL rate, accessions were scored as; 1) fast FL (FL > 50%), 2) medium FL (FL ranging between 30-50%), 3) slow FL (FL ranging between 10-30%), 4) very slow FL (FL < 10%) in relation to the reference time . The lowest FL after 3 weeks storage was found for cv. EZ 033-rin, RF-231 (1%) and the highest FL was found for. S. pimpinellifolium, RF_046 (72%) (Table 1, Figure 1).

Table 1. Measurements of shelf-life parameters; weight, firmness and fruit decay in 90 tomato genotypes.

RF	EU SOL Accession ID	Genotype Name	F (0)	F (21)	FL	FL score	%WL	(D)
RF_001	EA05097	Moneymaker	78	40	49	Medium	5	Medium
RF_002	EA05096	Ailsa Craig	83	37	56	Fast	4	Medium
RF_003	EA06086	Garderners Delight	71	26	63	Fast	18	High
RF_004	EA00465	Rutgers	72	42	41	Medium	5	Medium
RF_005	EA00325	Galina	73	33	54	Fast	10	Medium
RF_006	EA00448	Ponderosa	68	33	52	Fast	15	medium
RF_007	EA00375	Katinka Cherry	77	30	61	Fast	7	High
RF_008	EA00371	John's big orange	84	59	30	Slow	5	Low
RF_011	EA02617	All Round	82	40	51	Fast	8	High
RF_012	EA02724	Sonato	48	21	57	Fast	4	Medium
RF_013	EA03701	Cross Country	54	31	42	Medium	7	Medium
RF_014	EA03362	Lidi	36	24	34	Medium	6	Medium
RF_015	TR00003	Momatero	83	60	28	Slow	6	Low
RF_016	EA01965	Rote Beere	36	30	16	Slow	7	Medium
RF_017	EA03306	"Lycopersicum esculentum Mill."	59	21	65	Fast	7	Medium
RF_018	EA01155	DANA	46	37	21	Slow	4	Medium
RF_019	EA01049	Large Pink	79	36	55	Fast	5	Medium
RF_020	EA03221	"Lycopersicon esculentum Mill."	80	44	44	Medium	5	Medium
RF_021	EA03222	Bolivar' Lycopersicon esculentum	79	30	61	Fast	8	Medium
RF_022	EA04710	"Lycopersicum esculentum"	66	28	57	Fast	11	Low
RF_023	EA05170	"Lycopersicum esculentum"	very small				9	Low
RF_024	EA00990	Jersey Devil	78	27	65	Fast	17	High
RF_026	EA00157	Polish Joe	83	44	47	Medium	5	medium
RF_027	EA02054	Cal J TM VF	64	56	11	Slow	4	Low
RF_028	EA05581	The Dutchman	85	50	42	Medium	6	Medium
RF_029	EA00027	Black Cherry	75	43	43	Medium	8	Medium
RF_030	EA01835	ANTO	82	48	42	Medium	5	Low
RF_032	EA01854	Chang Li' Lycopersicon esculentum	72	35	52	Fast	9	High
RF_033	EA04243	Belmonte	80	51	36	Medium	4	Low
RF_034	EA00892	Tiffen mennonite	49	34	30	Slow	5	low
RF_035	EA01088	Wheatley's Frost Resistant	70	25	65	Fast	7	High
RF_036	EA04939	"Lycopersicum esculentum"	79	34	57	Fast	6	High
RF_037	EA05701	"Lycopersicum esculentum"	56	24	57	Fast	8	High
RF_038	EA05891	Chih-Mu-Tao-Se	68	29	58	Fast	11	very high
RF_039	EA04828	"Lycoperscium esculentum"	41	28	31	Medium	6	Low
RF_040	EA00526	ES 58 Heinz' Lycopersicon esculentum	77	54	29	Slow	4	Low
RF_041	EA04861	Dolmalik	72	30	58	Fast	6	High
RF_043	EA03058	S. pimpinellifolium (Jusl.) Mill.	81	29	64	Fast	15	High
RF_045	EA02960	S. pimpinellifolium	76	47	38	Medium	7	Low
RF_046	EA00676	S. pimpinellifolium	68	19	72	Fast	22	High

Table 1 continued.

RF	EU SOL Accession ID	Genotype Name	F (0)	F (21)	FL	FL score	%WL	(D)
RF_077	TR00018	Large Red Cherry	69	28	59	Fast	16	High
RF_078	EA00940	Porter	80	27	66	Fast	5	Medium
RF_088	TR00019	Bloody Butcher	65	28	57	Fast	6	Medium
RF_089	EA01019	Brandywine	73	22	71	Fast	10	Medium
RF_090	TR00020	Dixy Golden Giant	72	34	52	Fast	10	Medium
RF_091	EA01037	Giant Belgium	82	47	43	Medium	4	Medium
RF_093	TR00021	Kentucky Beefsteak	68	36	48	Medium	5	Medium
RF_094	TR00022	Marmande VFA	70	30	58	Fast	5	Medium
RF_096	TR00023	Thessaloniki	79	33	58	Fast	5	High
RF_097	EA01640	Watermelon Beefsteak	78	51	34	Medium	5	Low
RF_102	TR00026	"var. cerasiforme"	very small				10	Low
RF_103	TR00027	"var. cerasiforme"	80	37	54	Fast	8	Low
RF_201	EA00282	Blondokee	71	35	51	Fast	7	Medium
RF_202	EA00450	Purple Russian	47	26	51	Fast	9	Medium
RF_203	EA00470	Snowstorm	64	28	55	Fast	12	Medium
RF_204	EA00488	Taxi	50	29	42	Medium	4	Medium
RF_205	EA00907	Indian Striped	74	36	52	Fast	3	Medium
RF_206	EA00915	ABC Potato Leaf	81	49	40	Medium	7	Low
RF_207	EA00999	Napoli	82	46	44	Medium	4	Low
RF_208	EA01031	Eckert Polish	77	41	47	Medium	6	Low
RF_209	EA01120	Brandywine OTV	76	37	52	Fast	6	Low
RF_210	EA01165	Earlibell	73	29	60	Fast	5	High
RF_211	EA01284	Super Beafsteak	70	41	41	Medium	24	very high
RF_212	EA01345	Glory of Moldova	54	35	34	Medium	4	Low
RF_213	EA01349	Golden Queen	72	35	51	Fast	4	Medium
RF_214	EA01579	Heinz 1706	83	39	53	Fast	9	Low
RF_215	EA01709	"Italian cherry tomato "	75	53	30	Slow	3	very low
RF_216	EA01985	Sonora	55	49	11	Slow	5	Low
RF_217	EA02015	Felsozsolcai	80	52	35	Medium	5	Low
RF_218	EA02021	Mao Tao Shi Zi	80	29	64	Fast	10	High
RF_219	EA02550	Dippes Nz 802	78	35	55	Fast	4	Low
RF_220	EA02660	Madara	82	29	65	Fast	9	High
RF_221	EA02734	Stam Bovoy	72	27	62	Fast	9	High
RF_222	EA03444	Agrappoli d'Inverno	79	29	63	Fast	6	High
RF_223	EA05480	Early Dwarf	57	27	53	Fast	7	Medium
RF_224	EA05520	San-Marzano	81	41	49	Medium	9	Medium
RF_225	EA05677	Tomate Ciruela	72	31	57	Fast	5	Medium
RF_226	EA05721	DL/67/248	79	50	38	Medium	9	Very low

Table 1 continued.

RF	EU SOL Accession ID	Genotype Name	F (0)	F (21)	FL	FL score	%WL	(D)
RF_227	EA05732	Nagcarlan	77	39	49	Medium	4	Low
RF_229	EA05979	Morne a L'Eau	52	38	29	Slow	6	Low
RF_230	EA06631	Jaune Flamme	62	34	45	Medium	4	very high
RF_231	EA06746	EZ 033-rin	82	81	1	Very Slow	4	very low
RF_232	EA06902	OH88119	79	47	41	Medium	6	Medium
RF_233	EA06929	NCEBR2	71	36	49	Medium	12	HIGH
RF_235	EA03075	Kecskemeti Koria Bibor	44	15	67	Fast	10	High
RF_236	EA01042	Grosse Cotelee	55	33	40	Medium	5	Low
RF_237	EA06002	"var. cerasiforme"	59	25	58	Fast	6	High
RF_238	EA06699	RZ26	74	47	37	Medium	18	Medium
RF_140013	-	Flora Dade	85	60	29	Slow	6	very low
RF_140014	-	S. esc. A0011-6-3	85	64	25	Slow	5	Low

F(0): Firmness at the harvesting time. F(21): firmness after 21 days. FL: % Firmness loss. FL score: Firmness loss score. % WL: % water(weight) loss. (D): Fruit decay.
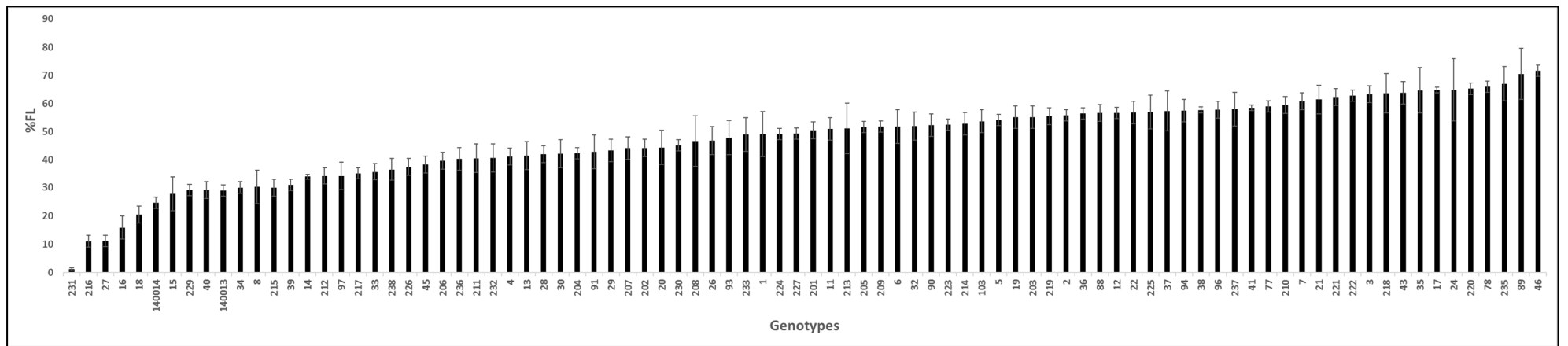
Figure 1. Percentage of fruit firmness loss after 21 days of storage in 88 tomato genotypes (two of the original 90 genotypes were too small to measure the firmness by firmness meter). Genotypes are presented based on their RF number. Error bars represent the variation in firmness between fruits.

2.1.3 Water loss

Water loss (WL) was expressed as the % of water loss and calculated by subtracting the fruit weight after 21 days from the initial fruit weight, divided by the initial fruit weight. The performance of genotypes for water loss is shown in Table 1 and Figure 2. The best performing genotype was cv. Italian cherry, RF_215 (3% water loss) which seems to be a ripening mutant, since its fruits did not get completely ripe until the end of storage (data not shown). Genotypes cv. Super Beafsteak (RF_211) and S. pimpinellifolium (RF_046) lost most water during three weeks storage (24 and 22% respectively). We did not find a direct correlation between the firmness loss and water loss (Figure S 1). Neither we found any correlation between the initial fruit weigh and percentage water loss (Figure S 2): in both big and small fruits varying water losses were observed.

3.1.3 Fruit decay (D)

Fruit Decay (D) was calculated as the % of fruits with an acceptable consumption quality after the end of each storage period in relation to the number of fruits at harvesting time and the results are presented in Table 1. The genotypes were classified based on their shelf life behaviour as follows; 1) Very high decay (VHD): genotypes for which >50% of the fruits were discarded (because they did not meet the minimal consumption quality) within 2 weeks storage, 2) High decay (HD): genotypes for which >50% of fruits were discarded within 3 weeks storage, 3) Medium decay (MD): genotypes for which >50% of fruits were discarded within 4 weeks storage, 4) Low decay (LD): genotypes for which > 50% were discarded within 5 weeks storage 5) very low decay (VLD): genotypes for which 50% or more of the fruits remained of acceptable consumer quality without apparent signs of decay until the end of the experiment. For genotypes cv. Italian cherry tomato (RF_215), cv. DL/67/248 (RF_226), cv. EZ 033-rin, (RF_231) and cv. Floradade (RF_140013) at least 50 percent of the fruits remained of acceptable consumption quality during the 6 weeks storage period.
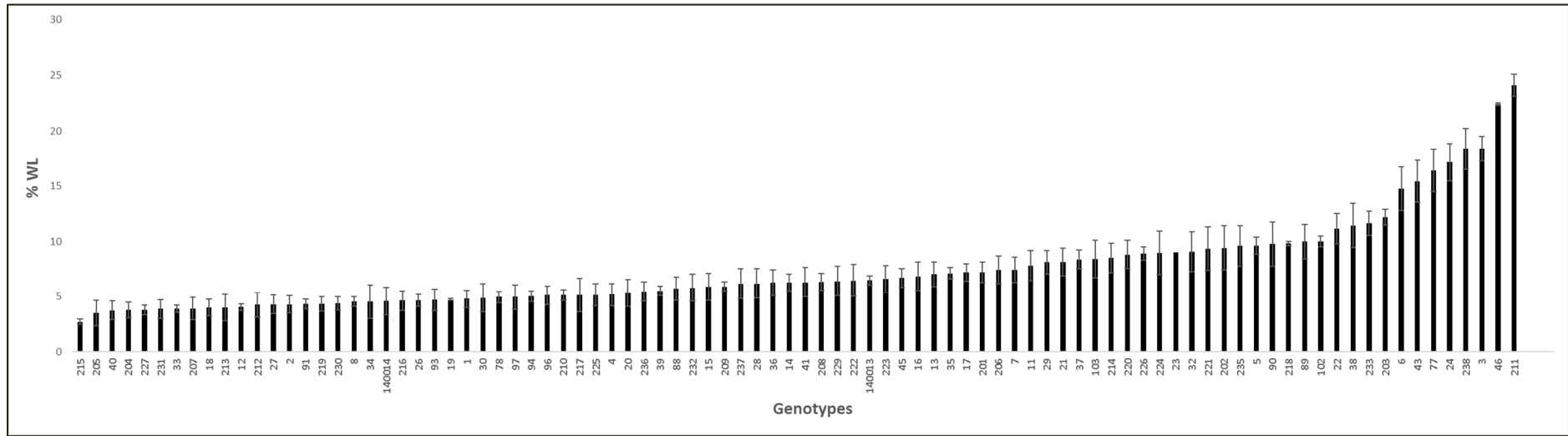
Figure 2. Comparison of genotypes based on their water loss (weight loss). Genotypes are presented by their RF number.

4.1.3 Chlorophyll and lycopene contents

The level of chlorophyll and lycopene based on Normalised Difference Vegetation Index (NDVI) and Normalised Anthocyanin Index (NAI) obtained from remittance VIS spectroscopy are given in Table S 1. Previously it has been reported that NDVI and NAI obtained from remittance VIS spectroscopy is closely related to the lycopene and chlorophyll content in pericarp tissue as measured by HPLC and other biochemical methods (Kuckenberg et al., 2008; Farneti et al., 2012). The level of lycopene reached its peak 14 to 21 days after harvest and after that remained constant. The higher lycopene content was found for cv. Black cherry (RF_029), cv. Purple Russian (RF_202) and cv. Indian Striped (RF_205). Chlorophyll had disappeared after 7 to 14 days of storage (reached its lowest peak) for most of the genotypes. In some genotypes decrease in chlorophyll content was very slow and they did not lose all of the chlorophyll content until the end of the storage period such as: cv. Black cherry (RF_29); cv. Indian Striped (RF_205); cv. Purple Russian (RF_202); cv. *Lycopersicum esculentum* (RF_037) and cv. DL/67/24 (RF_226). One of these genotypes (RF_029, cv. Black cherry) is a known colour mutant and holds two mutations; the *old-gold-crimson* mutation affecting *LYCOPENE β-CYCLASE* leading to an increase lycopene content and the *green flesh* allele *gf⁴* leading to retention of chlorophyll during ripening (Aflitos et al., 2014), but the other genotypes in this class are not among known colour and/or ripening mutants. In yellow and orange tomatoes the level of lycopene was much lower than in red tomatoes, but chlorophyll break down was the same as in normal red tomatoes (Figure 3). No correlation between lycopene content and firmness loss was found (Figure S 3).
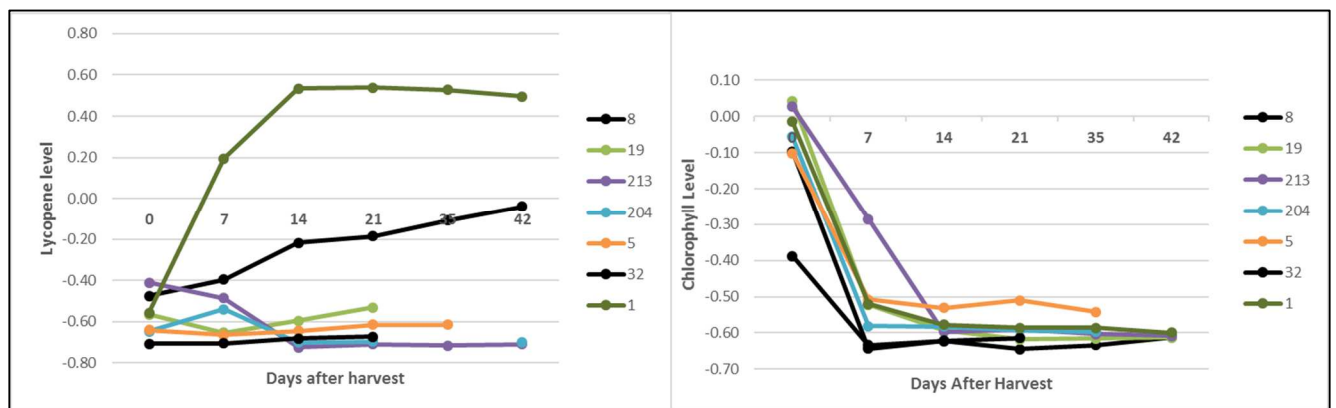

Figure 3. Lycopene accumulation and chlorophyll breakdown in eight yellow and orange fruit accessions in comparison to the performance of cv. Moneymaker (RF001). A) Lycopene accumulation, B) Chlorophyll breakdown. Genotypes are presented by their RF number.

5.1.3 Repetition of the shelf-life experiment with selected promising accessions

Based on the results of the shelf life experiment in the first season, 13 promising genotypes were selected and grown together with two standard genotypes cv. Moneymaker (RF_001) and cv. Gardeners delight (RF_003) and two negative control(soft) genotypes cv. Mao Tao Shi Zi (RF_218) and cv. Madara (RF_220) which showed a short shelf life performance in the first season. The result of firmness measurement for these 17 genotypes is shown in Figure 4. In Figure 4 A the black bars show firmness at the day of harvest and the grey bars show firmness after 21 days of storage. As is clear from this figure all of the 13 selected promising genotypes had acceptable firmness after 21 days (we empirically determined with a commercial cultivar that the lower threshold for an acceptable firmness was 40 N). The two negative controls and the two standard genotypes had the lowest firmness after 21 days. These four genotypes also showed higher firmness loss (Fig 4 B). Of all genotypes cv. Floradade (RF_140013) showed the lowest firmness loss. Performance of genotypes for water loss is shown in Figure 4C. The highest water loss was found for RF_218 and the lowest water loss for RF_140014. For four genotypes (RF_140013, 140014, 226, 040) fruit decay after 42 days was zero and all of the fruits harvested kept an acceptable consumer quality until the end of the experiment (see Figure 4D). Fruit decay for the two negative control genotypes RF_218 and 220 was 100%. Finally, considering all shelf life parameters, we selected four genotypes (RF_226, RF_140013, RF_027, RF_40) as long shelf-life (LSL), RF_218 as short shelf-life (SSL) and RF_001 (cv. Moneymaker) as standard genotype to study the mechanisms underlying shelf-life in more detail through biochemical characterisation.
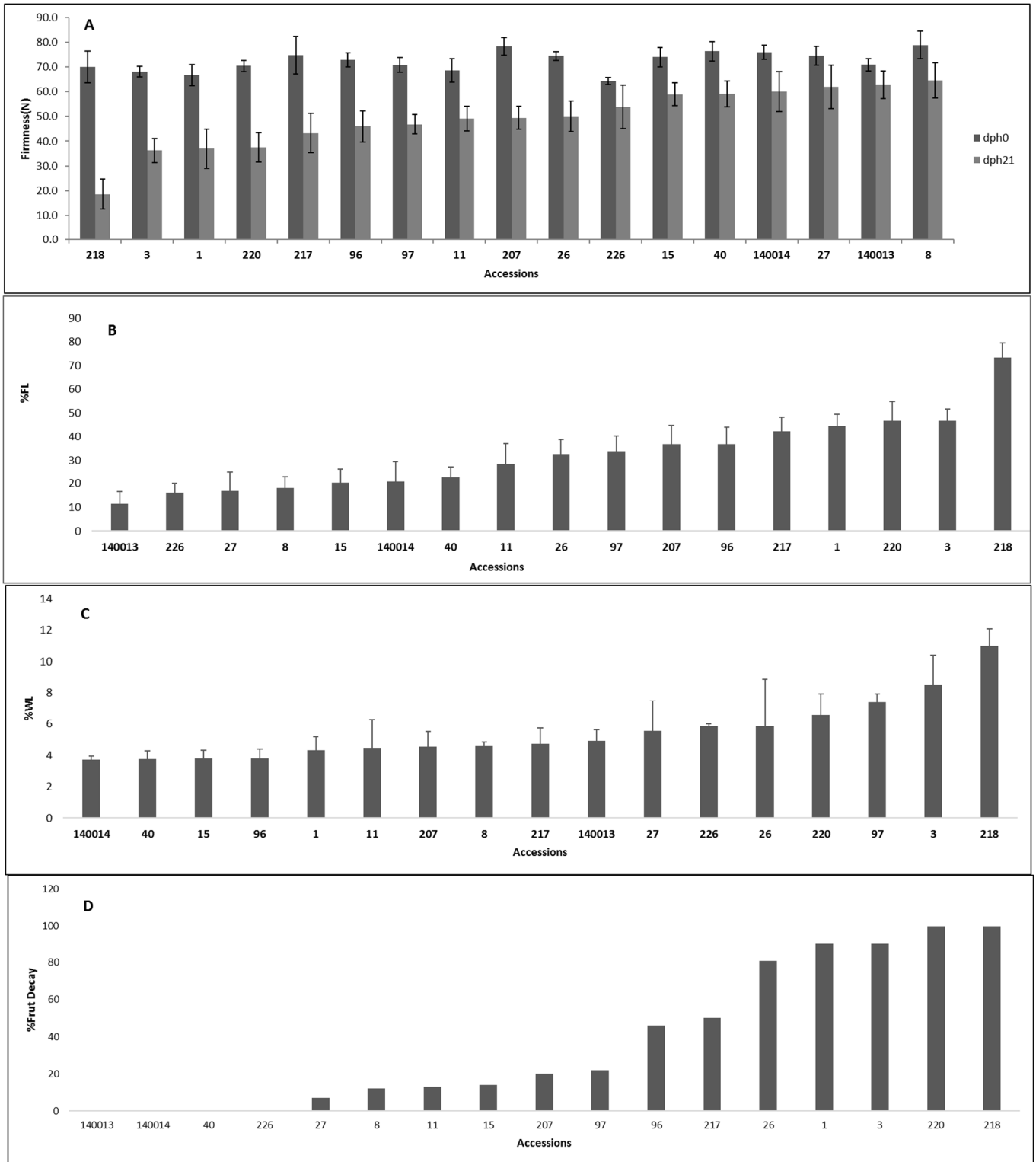
Figure 4. Performance of 17 genotypes for water loss, firmness and fruit decay during 42 days of storage. A: percentage of firmness loss. B: Firmness at the day of harvest and after 21 days of storage. C: Percentage of water loss(weight loss). D: Percentage of fruit decay.

2.3 Cell wall analyses

1.2.3 Changes in sugar composition

The sugar composition of the six genotypes at the five postharvest time points, expressed in mg/100mg AIR is shown in Table 2. Components of the pectic polysaccharides (arabinose, galactose, rhamnose, galacturonic acid) comprised approximately 60% of the pericarp cell wall of mature green tomatoes.

RF_218, the SSL genotype, remained firm enough until 21 days after storage but after that the quality of the fruits was not good enough anymore for cell wall extraction. Therefore the amount of sugars for this genotype has only been reported until B+21. From the sugar composition data it can be observed that the level of rhamnose (Rha) for all the genotypes, except RF_226 and RF-001, was constant during 35 days of storage (Figure 5 and Table 2). For RF_226 the amount of Rha decreased significantly after 21 days of storage, while in RF_001 Rha levels decreased at 14 d storage and increased again at 35 days of storage (Figure 5).

Arabinose (Ara) levels decreased 2 to 3-fold in RF_ 140013, 40, 27, 226 and 001 from the first day of storage until 14 days and after that remained constant. Interestingly, for RF_218, the SSL genotype, no change in cell wall Ara was observed during storage (Table 2 and Figure 6), and its level after 21 days storage is up to 2-fold higher than in the LSL genotypes.

Galactose (Gal) showed the most drastic decline (3 to 5-fold) during storage compared to the other sugars although this decline was different among the different genotypes. The decline occurred from B+1 until B+14 and after that there was no further change (Table 2 and Figure 7). The cell wall Gal content in RF_218, the SSL genotype, was significantly higher than in the other genotypes. RF_226 had the lowest Gal content in green stage and B+1.

The amount of Galacturonic acid (Gal A) increased for all the genotypes, except RF_001, until 14 days storage and after that it decreased (Table 2 and Figure 8). For RF_001 the amount of Gal A did not change until B+14 and after that it decreased. After 35 days of storage the GalA level in RF_001 was lower than those found in the four LSL genotypes. Unfortunately, it was not possible to harvest RF_218 at B+35 stage, since all fruits were already spoiled and discarded. Therefore we cannot make any comparison regarding the loss of Gal A for this genotype at B+35 days.

We were not able to separate the mannose and xylose by Dionex HPLC and the amount of these two sugars is reported together. The content of xylose+mannose (Xyl+Man) remained the same or showed an apparent increase during storage (Table 2). No remarkable difference between LSL and SSL genotypes was observed for the Xyl+Man content.

The amount of Glucose (Glc) remained constant during storage for all the genotypes (Table 2). The reported amount is for total glucose and we did not separate the cellulosic and non-cellulosic glucose.

The high amount of Glc in the mature green fruits was unexpected and may originate from starch and may be a remnant of soluble sugars that were not completely removed during the AIR isolation procedure.

Table 2. Changes in cell wall neutral sugars and galacturonic acid content of six contrasting shelf life tomato genotypes.

line/stage	Sugar content (mg/100mg AIR)					
RF_001	Rha	Ara	Gal	Gal A	Xyl+Man	Glc
GS	1.0 ± 0.06[a]	3.5 ± 0.20	8.9 ± 0.10	13.0 ± 0.5	3.7 ± 0.10	31.0 ± 1.30
B+1	1.1 ± 0.10	2.0 ± 0.05	7.0 ± 0.00	12.4 ± 0.3	3.2 ± 0.10	21.4 ± 0.40
B+14	0.7 ± 0.00	1.2 ± 0.08	1.8 ± 0. 04	12.7 ± 0.3	3.6 ± 0.10	21.5 ± 0.10
B+21	0.8 ± 0.01	1.5 ± 0.01	2.3 ± 0.07	8.8 ± 0.00	4.0 ± 0.00	22.8 ± 1.00
B+35	0.9 ± 0.05	1.4 ± 0.20	2.1 ± 0.07	9.2 ± 0.10	3.6 ± 0.03	19.8 ± 0.50
RF_140013						
GS	1.0 ± 0.06	3.5 ± 0.07	11.8 ±0.00	10.2 ± 0.30	2.5 ± 0.10	31.3 ± 0.10
B+1	0.8 ± 0.01	1.5 ± 0.04	8.8 ±0.00	13.7 ± 0.09	3.6 ± 0.08	22.7 ± 0.30
B+14	0.8 ± 0.05	0.7 ± 0.01	2.2 ±0.00	20.4 ± 0.70	3.4 ± 0.05	23.9 ± 0.70
B+21	0.8 ±0.01	0.8 ± 0.02	2.1 ±0.05	16.6 ± 0.30	3.4 ± 0.09	23.9 ± 0.01
B+35	0.9 ± 0.05	0.9 ± 0.00	2.4 ±0.04	14.3 ± 0.02	3.8 ± 0.02	22.1 ± 0.10
RF_226						
GS	1.5 ± 0.06	3.0 ± 0.04	5.6 ± 0.10	12.5 ± 0.40	2.9 ± 0.09	27.4 ± 0.40
B+1	1.2 ± 0.01	1.8 ± 0.06	3.2 ± 0.00	14.6 ± 0.00	4. ± 0.01	23.3 ± 0.30
B+14	1.1 ± 0.02	1.3 ± 0.00	1.8 ± 0.10	19.6 ± 0.07	3.4 ± 0.09	24.1 ± 0.04
B+21	1.0 ± 0.02	1.3 ± 0.00	2.2 ± 0.05	15.1 ± 0.08	3.5± 0.00	22.7 ± 0.10
B+35	0.9 ± 0.05	1.1 ± 0.04	2.4 ± 0.05	11.3 ± 0.07	3.8 ± 0.06	22.2 ± 0.02
RF_40						
GS	1.2 ±0.00	3.8 ± 0.01	10.1 ± 0.10	12.3 ± 0.10	3.0 ± 0.09	28.5 ± 0.20
B+1	0.9 ±0.02	1.4 ± 0.01	6.8 ± 0.10	13.7 ± 0.01	2.9 ± 0.05	22.5 ± 0.30
B+14	0.9 ±0.01	0.9 ± 0.00	2.3 ± 0.01	17.4 ± 0.20	3.0 ± 0.00	24.4 ± 0.10
B+21	1.0 ±0.00	0.8 ± 0.00	1.9 ± 0.00	15.0 ± 0.00	3.2 ± 0.07	24.2 ± 0.30
B+35	1.0 ±0.02	0.9 ± 0.01	2.4 ± 0.10	10.9 ± 0.40	3.9 ± 0.20	21.0 ± 0.70
RF_218						
GS	0.9 ± 0.03	2.0 ± 0.00	9.3 ± 0.10	7.8 ± 0.04	1.0 ± 0.02	37.5 ± 0.20
B+1	1.1 ± 0.01	1.8 ± 0.02	10.1 ± 0.10	14.9 ± 0.40	3.4 ± 0.02	21.4 ± 0.50
B+14	1.0 ± 0.00	1.8 ± 0.00	5.6 ± 0.10	17.5 ± 0.05	3.4 ± 0.03	21.0 ± 0.20
B+21	1.0 ± 0.00	1.9 ± 0.00	5.4 ± 0.05	15.0 ± 0.00	3.3 ± 0.05	21.2 ± 0.20
B+35	-		-	-	-	-
RF_27						
GS	0.8 ± 0.00	3.0 ± 0.00	6.8 ± 0.05	8.4 ± 0.01	1.5 ± 0.04	37.6 ± 0.60
B+1	1.1 ± 0.05	1.8 ± 0.00	4.0 ± 0.01	14.3 ± 0.02	3.4 ± 0.08	23.3 ± 0.10
B+14	1.0 ± 0.02	1.0 ± 0.00	1.3 ± 0.10	17.6 ± 0.03	2.8 ± 0.05	22.0 ± 0.80
B+21	1.0 ± 0.04	0.9 ± 0.00	1.3 ± 0.03	13.7 ± 0.30	3.0 ± 0.01	21.3 ± 1.00
B+35	1.0 ± 0.06	0.9 ± 0.03	2.1 ± 0.05	11.8 ± 0.08	3.5 ± 0.60	21.0 ± 0.40

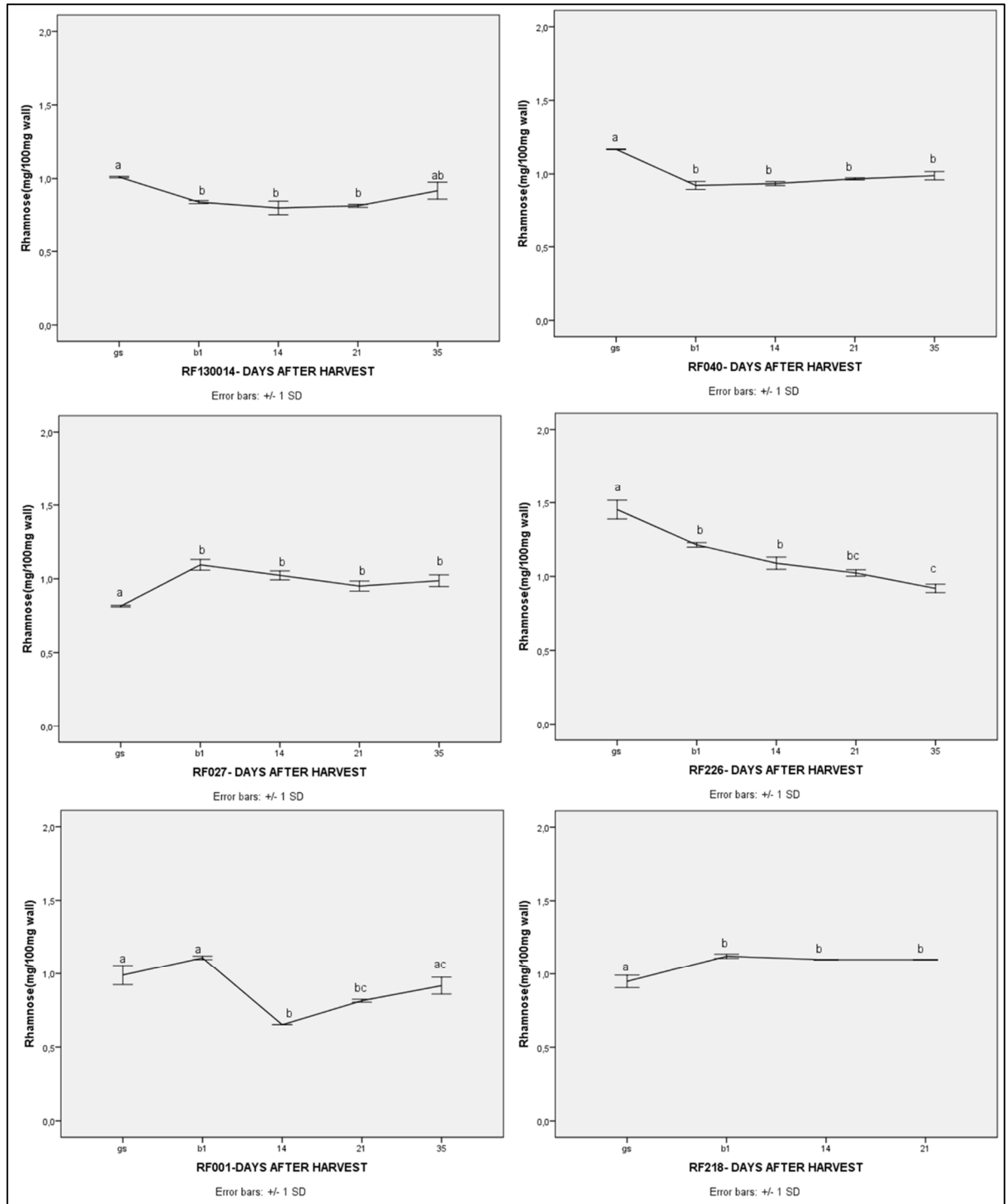Data are the means of three replicates. a: standard deviation.

Figure 5. Changes in cell wall rhamnose content during 35 days of postharvest storage in six tomato genotypes. Data shown are means of three replicates. Points (on a curve) which are labelled with the same letter are not significantly different from each other at the .05 level (LSD).
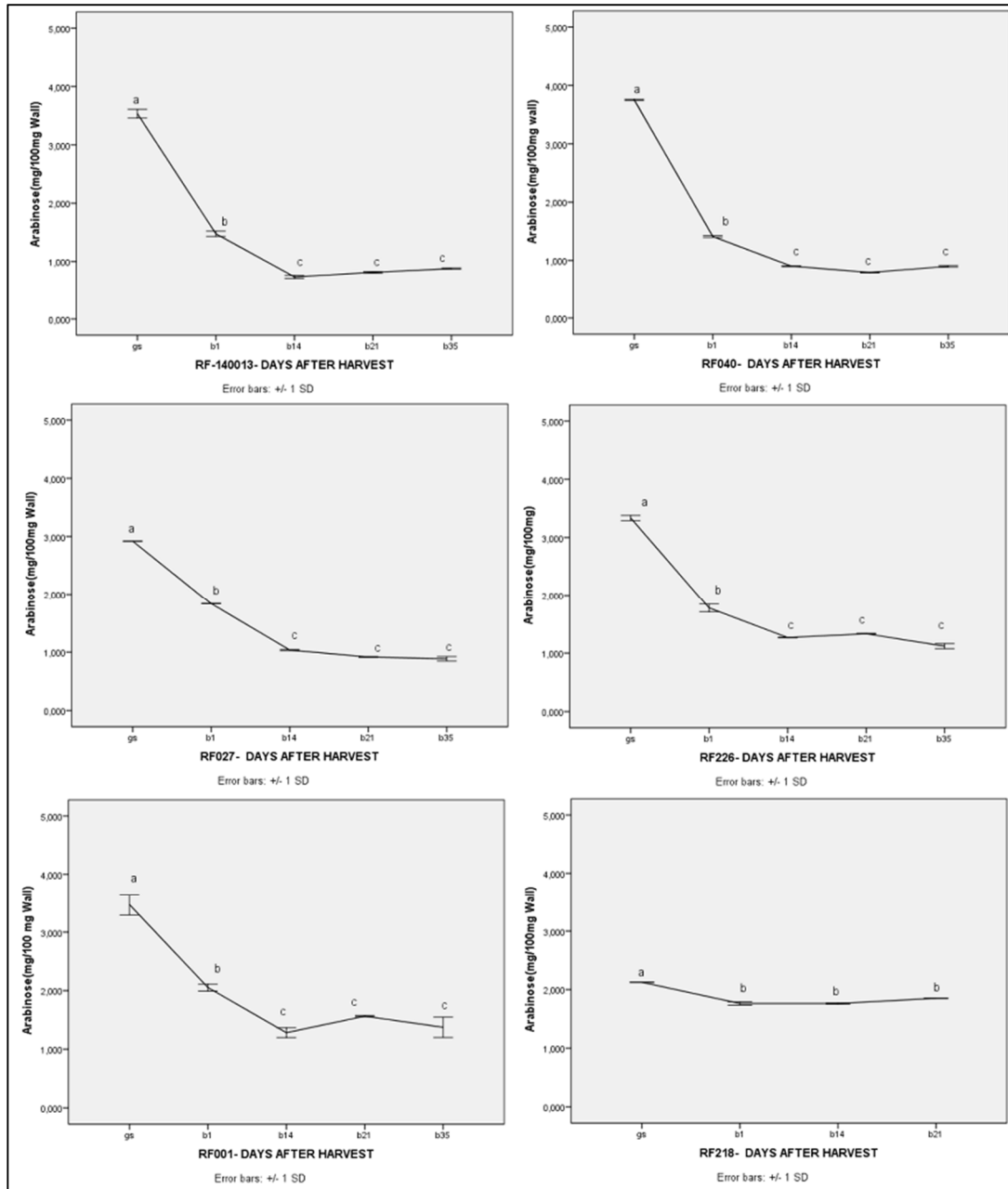
Figure 6. Changes in cell wall arabinose content during 35 days of postharvest storage in six tomato genotypes. Data shown are means of three replicates. Points (on a curve) which are labelled with the same letter are not significantly different from each other at the .05 level (LSD).

Figure 7. Changes in cell wall galactose content during 35 days of postharvest storage in six tomato genotypes. Data shown are means of three replicates. Points (on a curve) which are labelled with the same letter are not significantly different from each other at the .05 level (LSD).
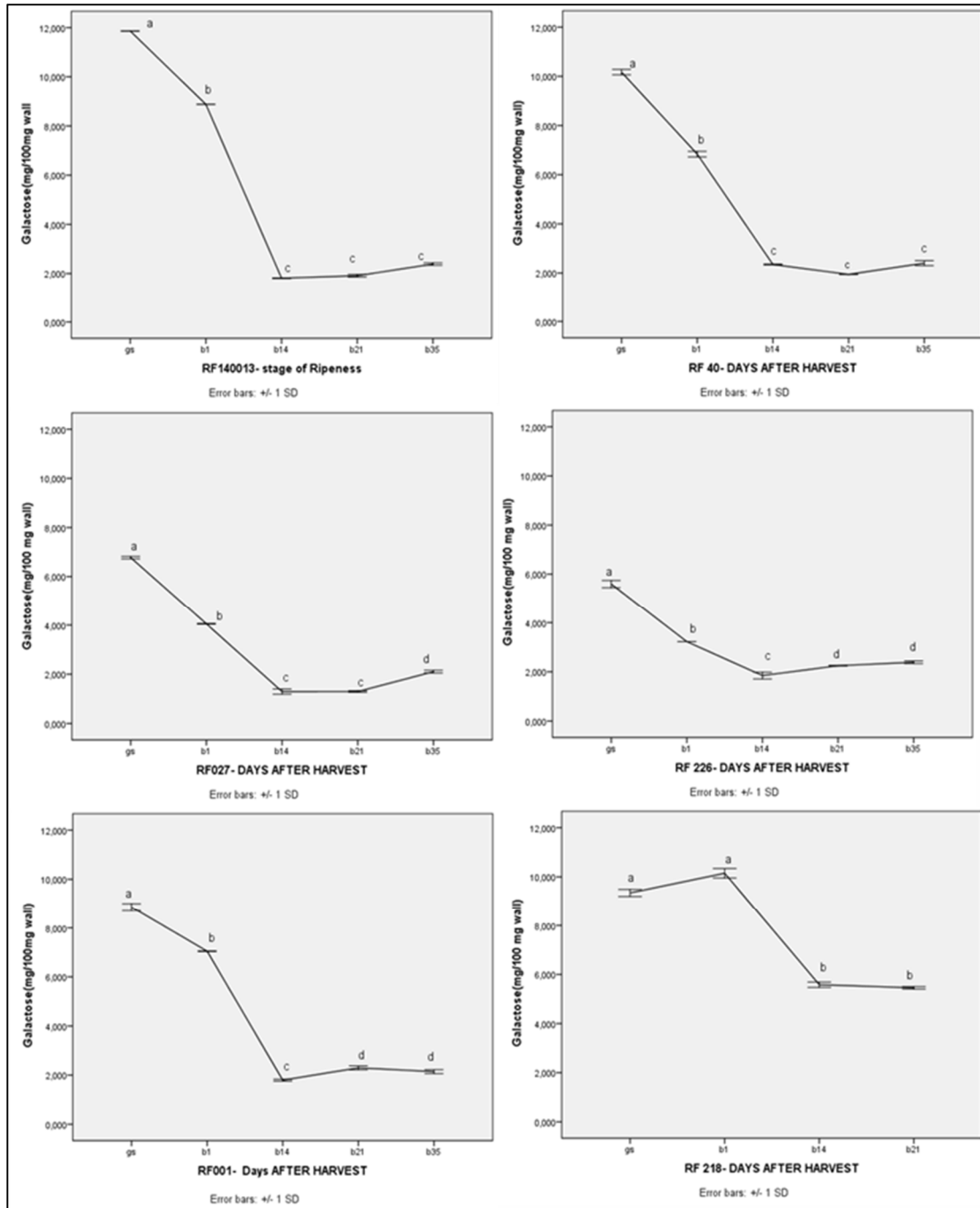
Figure 8. Changes in cell wall galacturonic acid content during 35 days of postharvest storage in six tomato genotypes. Data shown are means of three replicates. Points (on a curve) which are labelled with the same letter are not significantly different from each other at the .05 level (LSD).
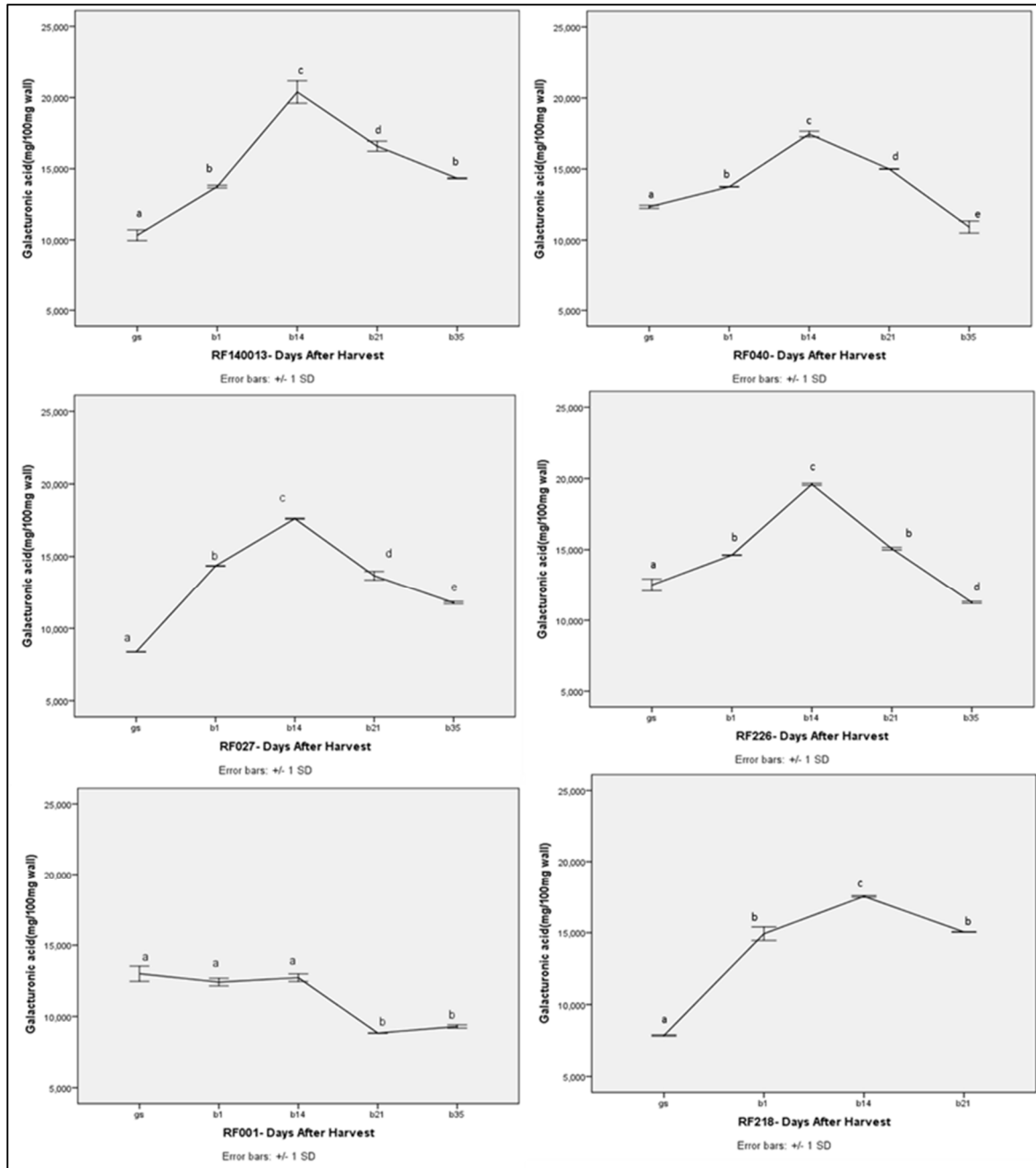
2.2.3 Sugar Ratios

For interpretation of sugar composition data, the use of ratios between the composing sugars can help to obtain information on the polymerisation level. Therefore, we defined three 'sugar ratios' based on (Houben et.al., 2011) that are used as an expression for the occurrence and properties of certain pectin structures from sugar composition data (Table 3). The first sugar ratio is embodied by the ratio of the

pectic backbone sugar Gal A to the neutral pectic sugars involved in side chains, thus being a measure for the linearity of pectin. The second ratio is defined as the proportion of Rha to Gal A, indicative for the contribution of RG (rhamnogalacturonan) to the entire pectin population. Ratio 3 compares the amount of RG-I side-chain sugars to Rha, as a measure for the extent of branching of RG-I.

Table 3. Sugar ratios based on sugar composition data.

	Sugar ratio	property
1	$\dfrac{Gal\ A}{Rha + Ara + Gal}$	Linearity of pectin
2	$\dfrac{Rha}{Gal\ A}$	Contribution of RG to pectin population
3	$\dfrac{Ara + Gal}{Rha}$	Branching of RG-I

The sugar ratios for the different genotypes, calculated based on the sugar content are displayed in Table 4. Summarizing the sugar composition data as sugar ratios reveals that the linearity of pectin (sugar ratio 1) goes up and then down in all genotypes, but the SSL genotype RF_218 and RF_001(MM) exhibit the lowest linearity of pectin when fruits are ripe (from B+14 onwards)(Table 4 and Figure 9). The contribution of RG to the pectin population (sugar ratio 2) goes down slightly during ripening in all genotypes and then up again, no difference in RG contribution was observed between LSL and SSL genotypes (Table 4). RG-I branching (sugar ratio 3) goes down in all genotypes, but remains highest in RF_218 and RF_001 in the last stages of ripeness.

Table 4. Sugar ratios for six examined tomato genotypes at five post-harvest time points. The definitions of the three ratios are given in Table 3.

| genotypes	Sugar ratios	GS	B+1	B+14	B+21	B+35
RF_140013	1	0.60	1.20	5.50	4.50	3.40
	2	0.10	0.06	0.04	0.05	0.06
	3	15.30	12.90	3.60	3.60	3.70
RF_226	1	1.20	2.40	4.70	3.40	2.60
	2	0.12	0.08	0.06	0.07	0.08
	3	5.70	4.20	2.80	3.50	3.90
RF_40	1	0.80	1.50	4.20	4.10	2.50
	2	0.10	0.07	0.05	0.07	0.09
	3	11.60	9.00	3.60	2.70	3.30
RF_27	1	0.80	2.00	5.30	4.30	3.00
	2	0.10	0.08	0.06	0.07	0.10
	3	12.30	5.30	2.30	2.20	3.00
RF_001	1	1.00	1.20	3.40	1.90	2.10
	2	0.08	0.09	0.06	0.09	0.10
	3	12.40	8.20	4.30	4.80	3.90
RF_218	1	0.60	1.10	2.10	1.80	-
	2	0.12	0.07	0.06	0.07	-
	3	12.60	10.80	7.40	7.30	-
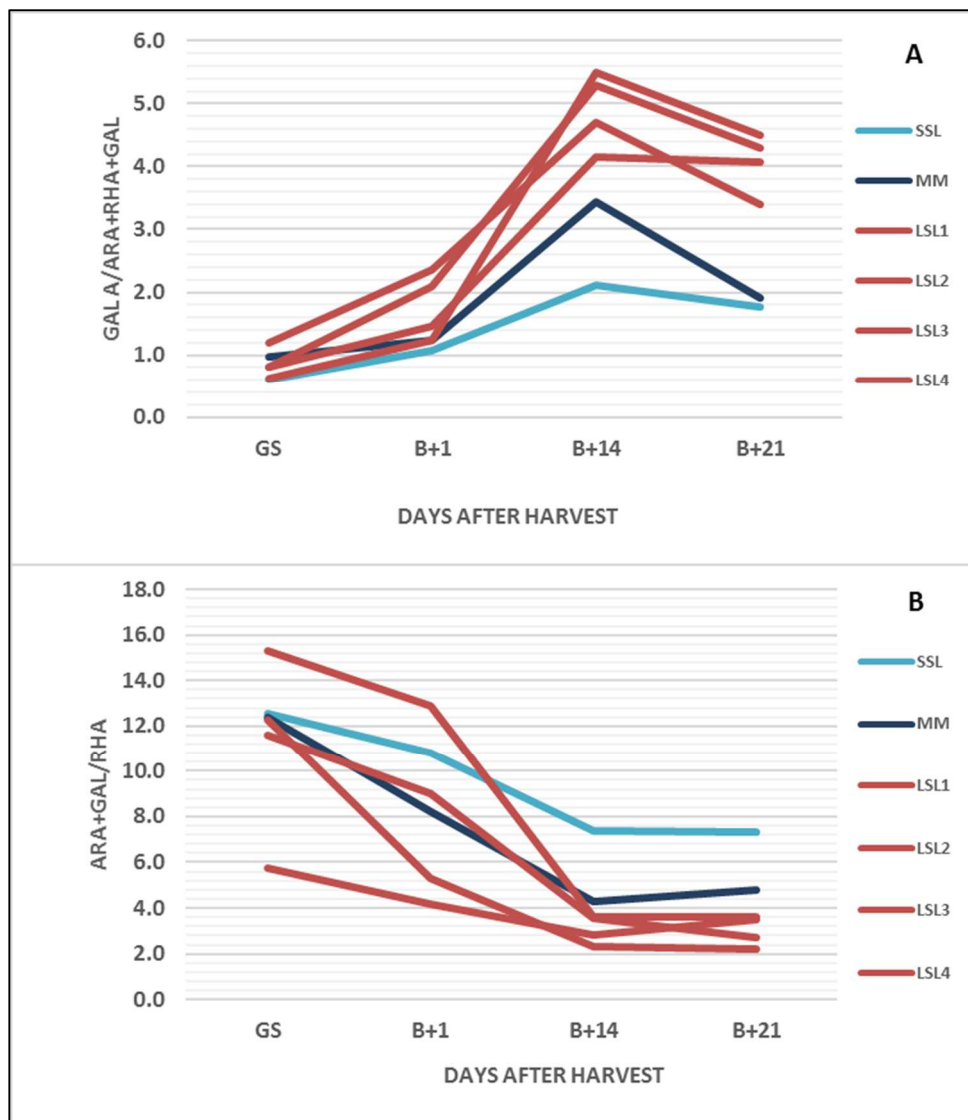
Figure 9. Change in two pectin properties; linearity and branching of RG1 which were derived from sugar ratios. Figure 9A shows the change in linearity of pectin during ripening and figure 9B shows the change in branching of RG-I.

3.3 Determination of primary metabolites

In total, 37 compounds were successfully identified (Table 5, Table S 2) using the GC/MS Metabolomics RTL library software. They were mostly represented by sugars, organic acids and amino acids.

Based on principal components analysis (Figure 10) 2 groups of metabolites can be recognised. The first group of metabolites (PC1) shows variation as a function of ripening stage which is the major variation in the data set (32%). The second group (PC2) shows variation between LSL and SSL genotypes (12%).

Table 5. Metabolites identified by GC-TOF-MS as components of a methanol extract from tomato fruit powder.

Amino acids	Organic acids	Sugars	Others
Alanine	Nicotinic acid	Glucopyranose	Sorbitol-6-p
Threonine	Lactate	Fructose	Phosphate
Serine	Maleate	Galactose	Suberate
Aspartate	Malate	Glucose	Hexadecanoic acid
n-acetylglutamic acid	Citrate	Mannose	5-Aminocarboxy-4,6-dihydroxypyrimidine
Lysine	Dehydroascorbate	Myo-Inositol	
Glutamate	Mevalonolactone	Sucrose	
phenylalanine	Galacturonate	3-O-Methylglucose	
Asparagine	Quinate		
Putrescine			
Glutamine			
Pyroglutamate			
Asparagine			
Gaba			
Glycine			

PC1 gave an overall view of the metabolite changes during post-harvest storage, and revealed discriminatory metabolites. In Figure 10, ripe tomato samples (B+14, B+21 and B+35 post-harvest time points) are located to the positive side of the X axis, which is characterized by high levels of, amongst others, Lysine, Asparagine, Alanine, Galacturonate, Aspartate, Glutamate, Galactose, Methyl glucose and Putrescine, whereas mature green and breaker samples are located towards the negative side of the X axis which is characterized by high levels of, amongst others, Malate, Citrate, Sucrose, Mannose, Glucose, Fructose and Myo-inositol. The further a variable is located from the axis origin the more influential the variable is on the principle component. Progress of ripening from the early pr-eharvest stages (green and breaker) to the postharvest ripe and over-ripe stages (B+14, B+21, B+ 35) seems to be characterized by a decrease in organic acids (Citrate and Malate), several sugars (sucrose, fructose, glucose, mannose) and an increase in amino acids and sugars such as galactose and methyl Glucose.

The key metabolites to distinguish LSL and SSL genotypes are putrescine (Ptc), galactose, myo-inositol, fructose, glucose, mannose and methyl glucose. In Figure 10, firm fruits are located to the positive side of the Y axis, while SSL fruits are located towards the negative side. Fruits of LSL genotypes contain the much more putrescine than SSL genotypes. In three of our LSL genotypes (RF_130014, 040, 226) the putrescine level went up during ripening and ripe fruits of these genotypes contain 2-3 times higher putrescine compared to the green and breaker stages (Figure 10, Figure 11). One of the LSL genotypes (RF-027) did not accumulate these high levels of putrescine in the ripe fruits. The amount of putrescine in

RF-001 did not change during the storage period but in RF_218 putrescine levels showed a steady decrease and green fruits of this genotype contained more putrescine than ripe fruits. Galactose was another metabolite which showed a high quantitative difference between LSL and SSL genotypes. Galactose levels were similar in all genotypes at G and B+1 stage. In SSL genotypes, however, galactose levels increased dramatically upon postharvest storage, while this increase was only minor in LSL genotypes (Table S2, Figure 12). The amount of mannose in the early stage of ripening(pre-harvest) was higher than at the later stages and this amount was up to 2-fold higher in the two SSL genotypes (Figure 10, Table S2).

The amount of glucose was highest at the early ripening stages (G and B+1) and decreased in most genotypes when ripening progressed. A difference between the glucose content of SSL and LSL genotypes could be observed at later ripening stages: SSL genotypes contained higher glucose levels than LSL genotypes at B+21 and B+35 days (Figure 10, Table S 2). The change in Fructose levels during storage and its variation between SSL and LSL genotypes was the same as for Glucose. Myo-inositol in SSL genotypes was higher than in the LSL genotypes and the biggest difference was observed at B+21 and B+35 (Table S 2).
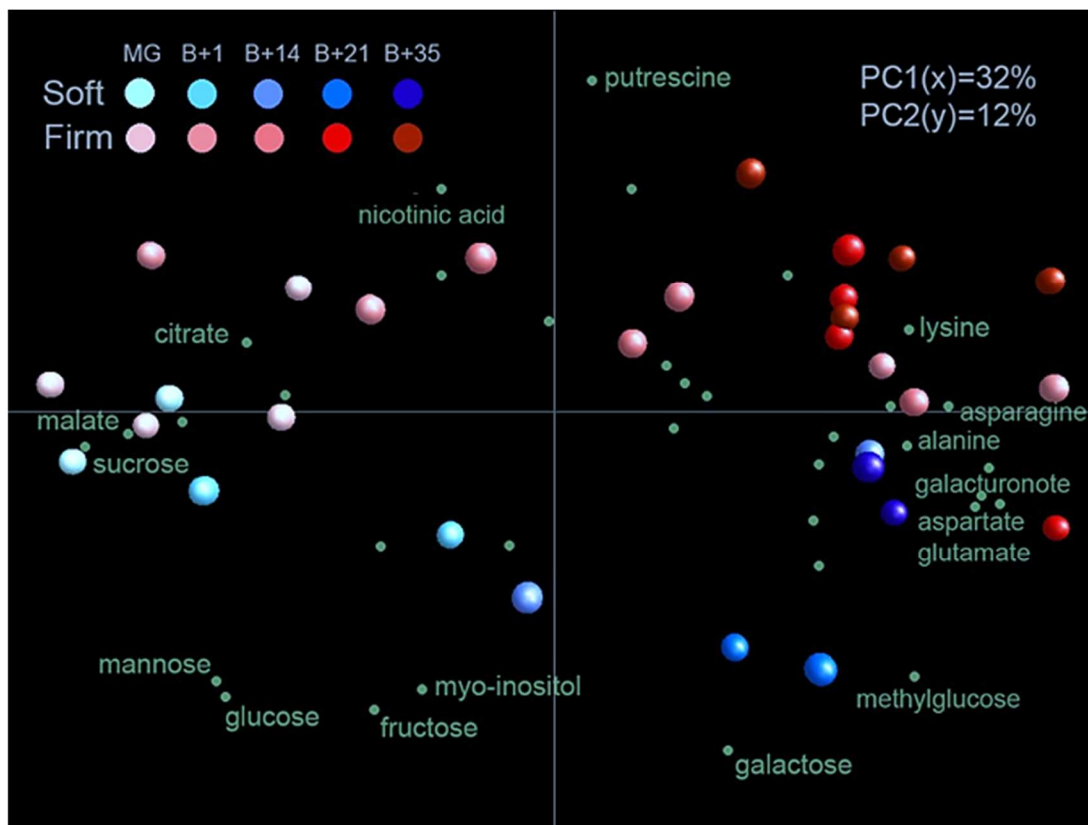

Figure 10. Principal Components Analysis of the variation of primary metabolite accumulation between SSL and LSL fruits and between post-harvest time points. LSL genotypes (cv. Moneymaker and 218) are coloured from light to dark blue depending on their ripening stages. LSL genotypes are depicted in light to dark red.
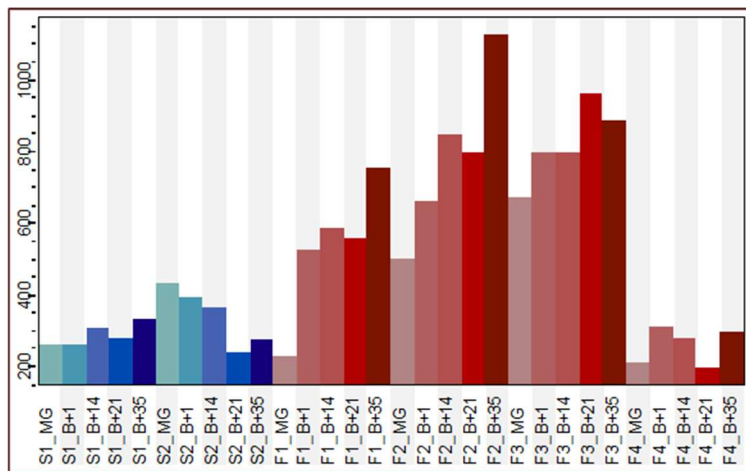
Figure 11. Change in the levels of putrescine during storage. S1 and S 2 represent the two soft genotypes RF_001 and RF_218. F1 to F4 represent the four firm genotypes RF_140013, 226, 040 and 027 respectively.
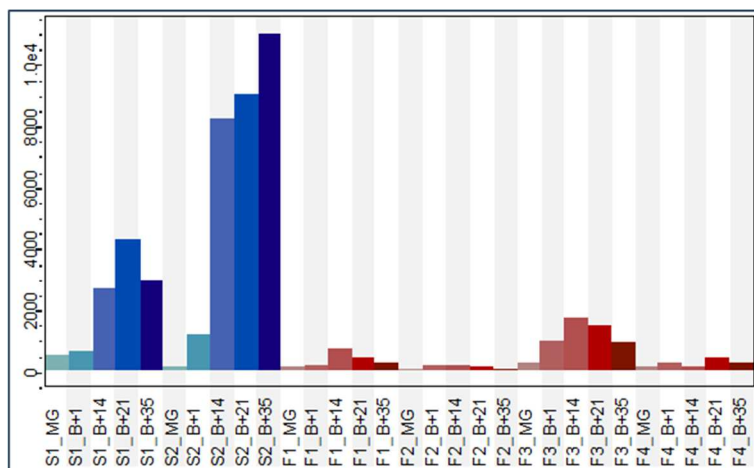

Figure 12. Changes in soluble galactose content during post-harvest storage. S1 and S 2 is indicative for the two soft genotypes RF_001 and RF_218 respectively. F1 to F4 represent the four firm genotypes RF_140013, 226, 040 and 027 respectively.

4.3 Determination of Volatile compounds

In total 174 volatile compounds were identified and semi-quantified in fruits of six genotypes at five post-harvest time points. Six major biochemical classes of volatiles were detected: 1) lipid-derived, 2) phenolic volatiles derived from phenylalanine, 3) phenylpropanoid volatiles 4) terpenoids, 5) volatiles derived from the amino acids Leu and Ile, 6) open chain carotenoid derived volatiles. PCA shows no variation for volatile compounds between SSL and LSL fruits and fruits of both types of genotypes were distributed as a function of ripening stages (Figure 13, A). In Figure 13 we observe that green fruits from both class of genotypes, are characterised by phenylpropanoid "smoky" VOCs whereas the breaker fruits accumulated more lipid

100

derived volatiles and terpenoids. On the other hand, carotenoid derived, and amino acid derived VOCs increased during ripening and were typical for fully ripe fruits.
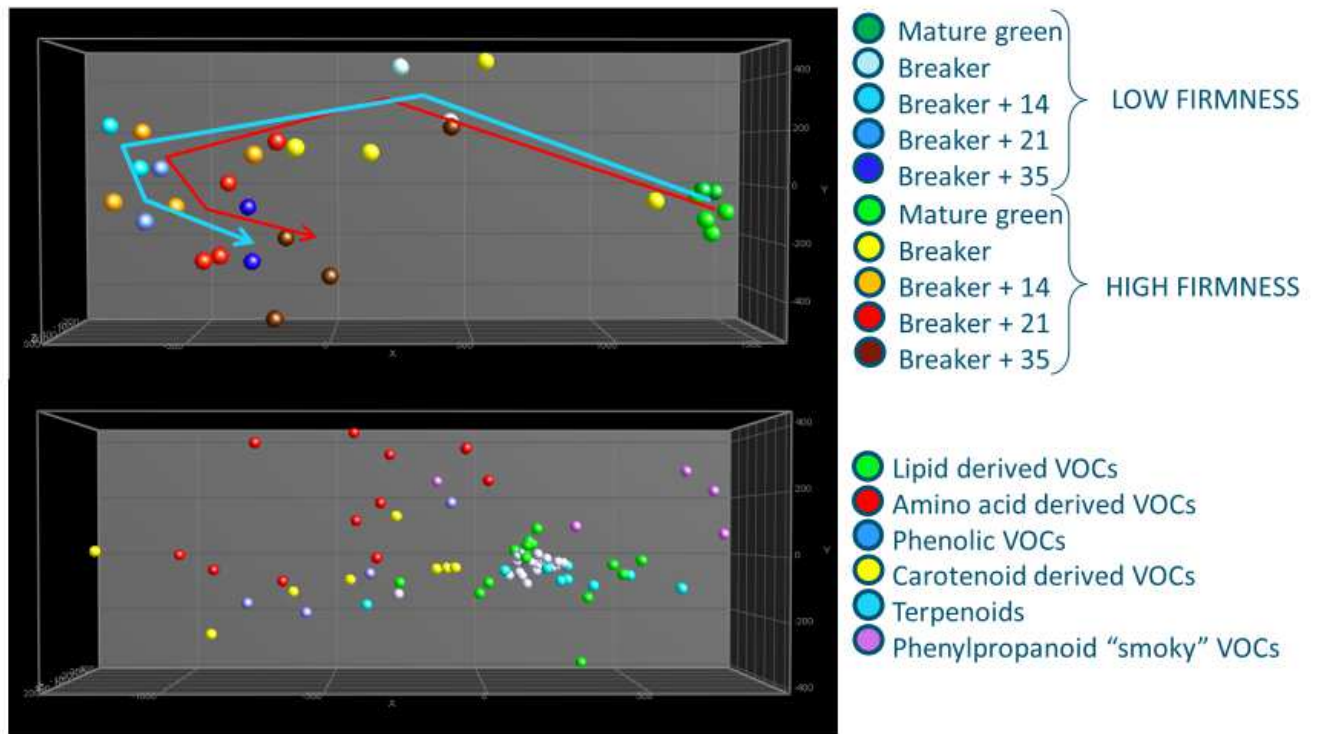

Figure 13. Principal Components Analysis of the variation of volatile compounds between post-harvest time points and soft/firm fruits. A: Tomato samples from both soft and firm genotypes at five post- harvest time-points. Soft genotypes are coloured from light to dark blue depending on their ripening stages. A: six different classes of Volatile compounds. B: Volatile compounds content distribution.

4 Discussion

1.4 Shelf-life trial

In this study, a collection of 90 tomato genotypes, including old cultivars and landraces was characterised for fruit post-harvest shelf-life. The collection presented a wide range of diversity for shelf-life attributes such as firmness, water (weight) loss and colour pigments. This information is expected to be valuable for subsequent tomato breeding programs and tomato crop improvement. We observed that there are significant differences among genotypes in firmness at the day of harvest and that firmness at harvest does not guarantee the maintenance of fruit firmness during storage. In some genotypes the firmness at harvesting time was not high, but loss of firmness was slow and after 21 days they still exhibited an acceptable firmness, which was empirically set at >40 N. On the other hand there were genotypes that at

the day of harvest showed very high firmness but they lost considerable firmness during storage. Neither firmness at harvest, nor firmness after 21 days was related to the size of the fruits (fruit weight). We observed a diverse range of firmness scores in both small and big fruits. Water loss during storage was also not correlated to the initial fruit weight and water loss and firmness loss were not well correlated to each other (Figurer S 1 & S 2).

Chlorophyll breakdown and lycopene accumulation was monitored during post-harvest storage and we observed that some genotypes such as cv. Black cherry, RF_029, cv. Indian Striped, RF_205 and cv. Purple Russian (RF_202) accumulated more lycopene compared to others after 21 days of storage when lycopene reached its highest peak. Notably, in these three genotypes chlorophyll breakdown was also very slow and chlorophyll content in the ripe stages (14-21 days after storage) was much higher compared to other genotypes. Among the 13 genotypes selected as exhibiting a long shelf-life, RF_226 exhibited slow breakdown of chlorophyll, but the others showed normal (compared to standard genotype cv. Moneymaker) breakdown in chlorophyll and accumulation of lycopene. In general there was no correlation between the accumulation of lycopene and firmness loss (also shelf-life) (Figure S 3).

2.4 Cell wall analyses

During storage there was a decrease in the cell wall galactose (Gal) content in all six genotypes and concomitantly, there was a steady increase in soluble galactose in the short shelf-life (SSL) genotype RF_218 and the standard tomato cv. Moneymaker. The long shelf-life (LSL) genotypes exhibited no noticeable change in soluble galactose level during postharvest storage (Figure 12, Table S 2). In previous studies in which the cell wall sugars of *rin* and *nor* mutants had been analysed in comparison to normal ripening tomato cv. Heinz and cv. Rutgers, the decrease in cell wall galactose content was noticed for both ripening mutants and normal ripening tomatoes but the increase in soluble galactose was only noticed in normal ripening tomatoes (Gross & Wallner, 1979; Gross, 1983; Gross & Sams, 1984; Gross, 1985; Seymour et al., 1990; Kim et al., 1991). The mechanism of net loss in galactosyl residues from the cell wall has been suggested to be related to a reduced rate of de novo galactan synthesis or hydrolysis of galactan by the enzyme β-galactosidase II (Gross & Wallner, 1979; Lackey et al., 1980; Gross, 1983). The exact source(s) of soluble galactose remains as yet unknown. However, the cell wall is by far the major galactose-containing organelle and soluble galactose in the fruit pericarp is most likely due to galactosyl solubilisation from the cell wall (Kim et al., 1991; Prasanna et al., 2007). The current idea is that most of the cell wall galactose is present in the side chains of RG-I pectin which is very firmly integrated into the wall, via its side-chains and / or its backbone (Brummell, 2006; Zykwinska et al., 2007). De-branching of RG-I side chains has been considered to be an important component of the changes that alter fruit firmness and textural properties and in this regard galactan and arabinan, the side chains of RG-I, play important

roles in cell wall structure and function (Brummell, 2006). In pea cotyledon, potato tuber and tomato fruit has already been shown that loss of pectic galactan side chains has correlation with decrease in wall firmness (McCartney et al. 2000; Smith et al., 2002; Ulvskov et al., 2005; Brummell, 2006). Several observations in our study indicate that cell wall galactose content in our SSL genotype is higher than in the other genotypes. The reason for the high amount of soluble galactose in SSL genotypes may be derived from the high amount of galactose in their cell wall and hydrolysis of cell wall galactan chains during ripening contributes to the free galactose in their pericarp. Results obtained in the present study (Figure 6 and 8) show that besides cell wall galactose, the content of two other pectin sugars arabinose and galacturonic acid underwent considerable changes during postharvest storage, although these changes were different among the different genotypes. The decline in arabinose and galactose started from green stage until 14 days after storage when fruits are completely ripe. After that time point significant changes in the amounts of these two sugars were no longer observed. The amount of cell wall galacturonic acid increased from green stage until 14 days after storage and after 14 days it decreased. At the same time, 14 days after storage, the increase in the amount of free galacturonic acid in the fruit pericarp was observed in all six genotypes, but higher levels of free galacturonic acid solubilisation was observed for the SSL genotypes RF_218 (Figure 10, Figure S 4, Table S 2).

Mitcham et al. demonstrated that cell wall synthesis continued throughout ripening of tomato fruit, even though there was a net loss of cell wall material (Mitcham et al., 1989). Other studies also revealed that the cell wall undergoes compositional changes during ripening, i.e. degradation of cell wall polysaccharides coincident with synthesis and insertion of replacement polymers. But they also pointed out that turnover occurs with degradation exceeding synthesis, resulting in a net decrease in total dry weight (Mitcham et al., 1989; Labavitch, 1981; Mitcham et al., 1991; Goulao & Oliveira, 2008). The change in the composition of cell wall sugars in our study can be interpreted as the result of "cell wall polysaccharide turnover", as described above. Based on sugar composition data at the examined postharvest time points three sugar ratios were defined to obtain information on the occurrence and properties of certain pectin structures (Table 3). Based on these sugar ratios we observed that the amount of linear pectin increased in all genotypes until B+14 and then decreased upon further storage. RF_218 and cv. Moneymaker had the lowest levels of linear pectin at all developmental stages (Figure 9 A and B, Table 4). The extent of branching of RG-I decreased in all genotypes during ripening and postharvest storage, but remained highest in the SSL genotype RF_218 and cv. Moneymaker. Based on these results we suggest that the decline in the amount of pectin sugars in the fruit cell wall of LSL genotypes is most likely due to reduced rates of *de novo* synthesis of these sugars rather than to hydrolysis and removal of them from the cell

wall. The lack of an increase in the levels of free galactose and the low amount of free galacturonic acid in the pericarp of ripening LSL genotypes

(Table S2) confirm this suggestion. In contrast, the free galactose and galacturonic acid levels clearly increase upon ripening in LSL genotypes RF_218 and MM (Table S2), suggesting that polysaccharide hydrolysis may play a significant role in these two SSL genotypes. The decline in the amount of pectin sugares does not seem to be directly related to fruit firmness, since firm fruits in our study and ripening mutants in other studies soften only slightly despite a considerable decrease in their cell wall galactose and arabinose content during ripening. Considering this information, we think that pectin structure and modification during fruit ripening has a large effect on fruit softening, since changes to pectin structure and bonding will affect the movement of enzymes within the wall and the existence of a more branched pectin structure in the cell wall may lead to more solubilisation of pectin and therefore more fruit softening. Arabinan and galactan side chains are known to be much more flexible than the HG (homogalacturonan) components that form the backbones of pectic polysaccharides and they may therefore increase the porosity changes necessary for the access of cell wall modifying enzymes such as polygalacturonase and expansin to their site of action during ripening (Cosgrove, 2000; Smith et al. 2002). As has previously been shown the accumulation of polygalacturonase activity is not the only factor responsible for the solubilisation of pectin, but also an increase in the accessibility of the enzyme to its homogalacturonan substrate during ripening is another important component of cell wall changes (Giovannoni et al. 1989; Brummell, 2006). A study of the fractionation of polysaccharides and oligosaccharide phenotyping at each post-harvest time point will provide more information and details about the presence of certain pectin or polysaccharide structures in these two classes of genotypes (long shelf-life and short shelf-life).

3.4 Primary metabolite variation

The fruits of the LSL genotypes showed a clear increase in putrescine level during ripening as compared to the soft genotypes (Figure 10 and 11). Earlier (Dibble et al., 1988) it has been reported that fruits of the *Alcobaca (Alc)* landrace of tomato, which ripen slowly and have prolonged keeping qualities, contain three times as much putrescine as the normal variety at the ripe stage. It was suggested that the enhanced putrescine levels in this line may be responsible for the ripening and storage features. As polyamines have been shown to delay senescence in some plant tissues they may also act in this way in tomato fruit. The long keeping quality of *Alc* has been shown to be related to its ability to reduce the over-ripening effect caused by ethylene, which regulates the senescence process. Other studies have shown that the functions of ethylene and polyamines are antagonistic in senescence not only in tomato fruit of the Alcobaca landrace, but also in apple fruit and tobacco leaves (Apelbaum et.al., 1982; Smith, 1985; Dibble et al., 1988). The two pathways are closely linked as they share a common intermediate, i.e. S-adenosyl-methionine for their

biosynthesis. The elevated putrescine levels in Alc fruit have been reported not to be due to changes in putrescine conjugation or metabolism, but rather to an increase and rise in free putrescine levels (Rastogi, 1991; Rastogi, 1990). The exogenous application of polyamines including putrescin either at pre-harvest (during fruit growth and ripening on tree) or after harvest (postharvest treatments) to delay the postharvest ripening process is a common method of postharvest handling in several fruits such as pomegranate, strawberry, plum, apricot, mango, tomato and cucumber (Barman, et al., 2011; Koushesh saba et al., 2012; Jia et al., 2018; Wannabussapawich & Seraypheap, 2018). Exogenous application of putrescine will result in maintaining fruit quality attributes such as colour, firmness, acidity, and total soluble solids and reduced mechanical damage and chilling injury susceptibility during postharvest storage (Serrano et al., 2003; Serrano & Valero, 2018). Polyamines are involved in many aspects of plant development and considered as important molecules associated with both abiotic and biotic stresses. Exogenous application of polyamines has been shown to increase the levels of antioxidant compounds and the activity of antioxidant enzymes, which results in a decrease in accumulation of ROS (reactive oxygen species) and in this way senescence processes and post-harvest over ripening in fruits will be delayed (Sharma et al., 2017; Serrano & Valero, 2018). Nevertheless, the precise physiological and molecular mechanisms by which putrescine increases fruit shelf-life are not clear yet and need further research.

Acknowledgement

Supplementary Data
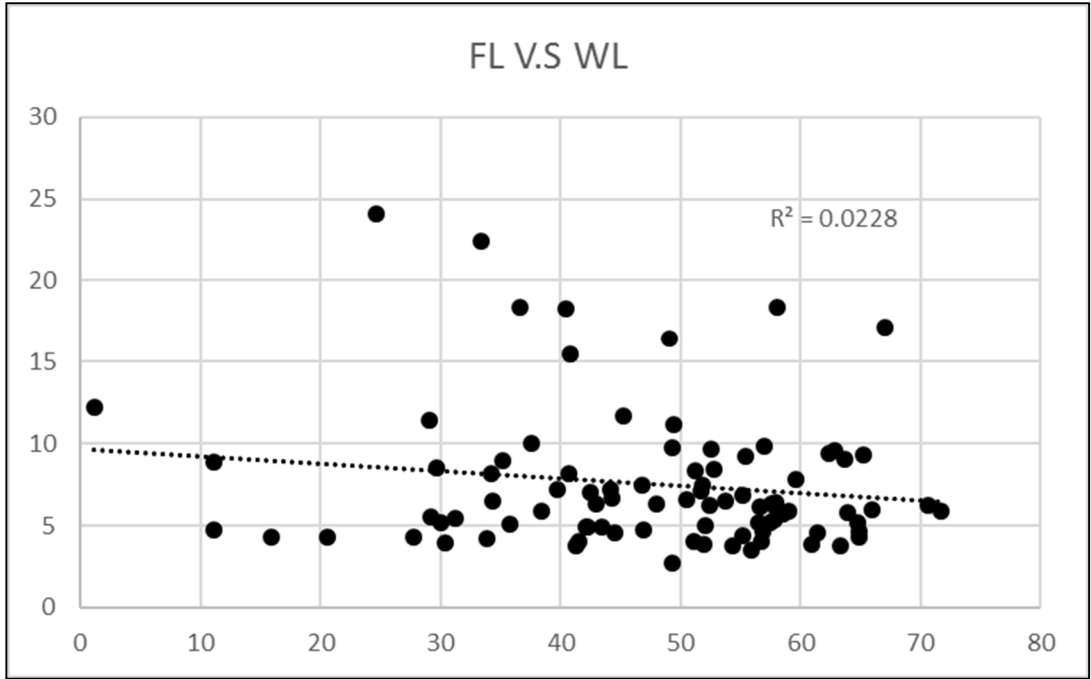
Figure S 1. The correlation between firmness loss(FL) and Water loss(WL).
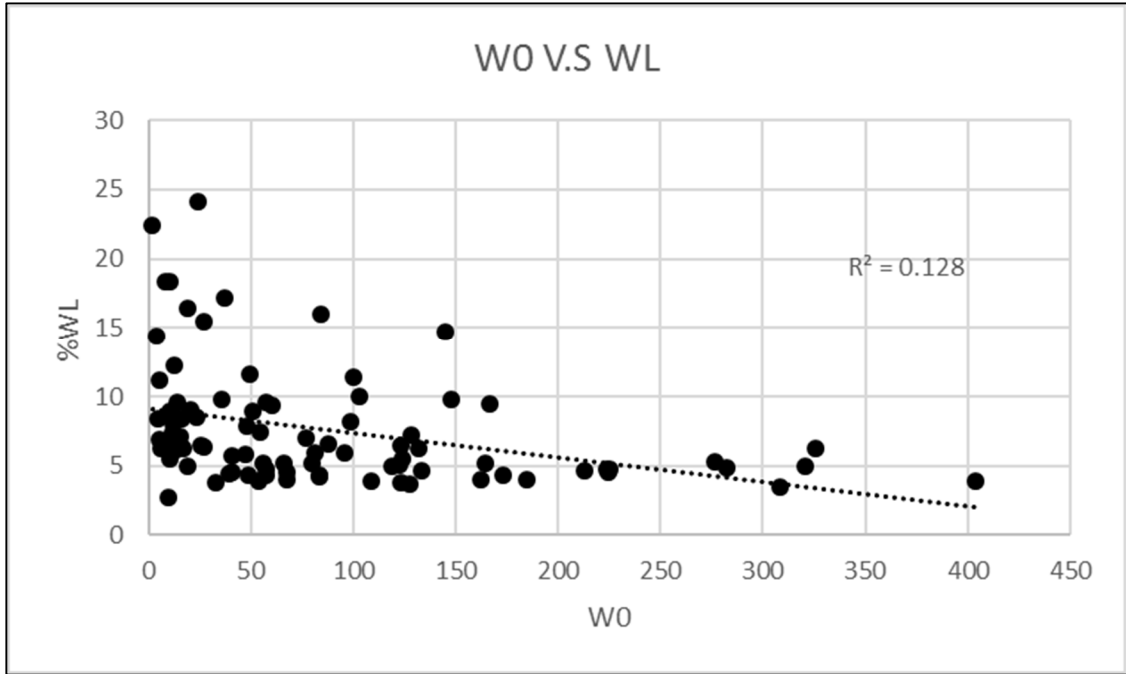
Figure S 2. The correlation between initial weight (W0) and fruit weight loss loss (WL).
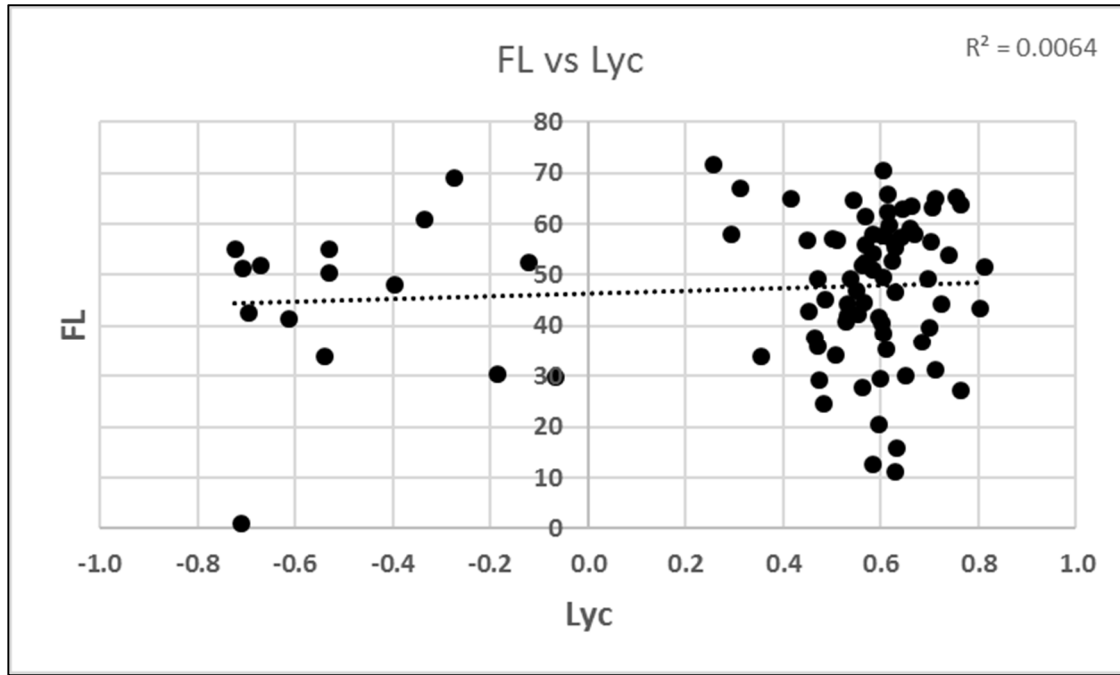
Figure S 3. The correlation between lycopene content and firmness loss.

	B+1	B+14	B+21	B+35	
1	1940.3	165	692	2897	4729
140013	5790	179	613	1290	1560
226	2235	116	152	3400	4022
40	133	2635	2972	3359	3845
218	167	575	3368	5541	9910
27	170	1834	1310	1200	2610

Figure S 4. Changes in free galactorunic acid (Gal A) content during post-harvest storage among six genotypes.

Table S 1. Chlorophyll and lycopene content measured by pigment analyser during 42 days of post-harvest storage.

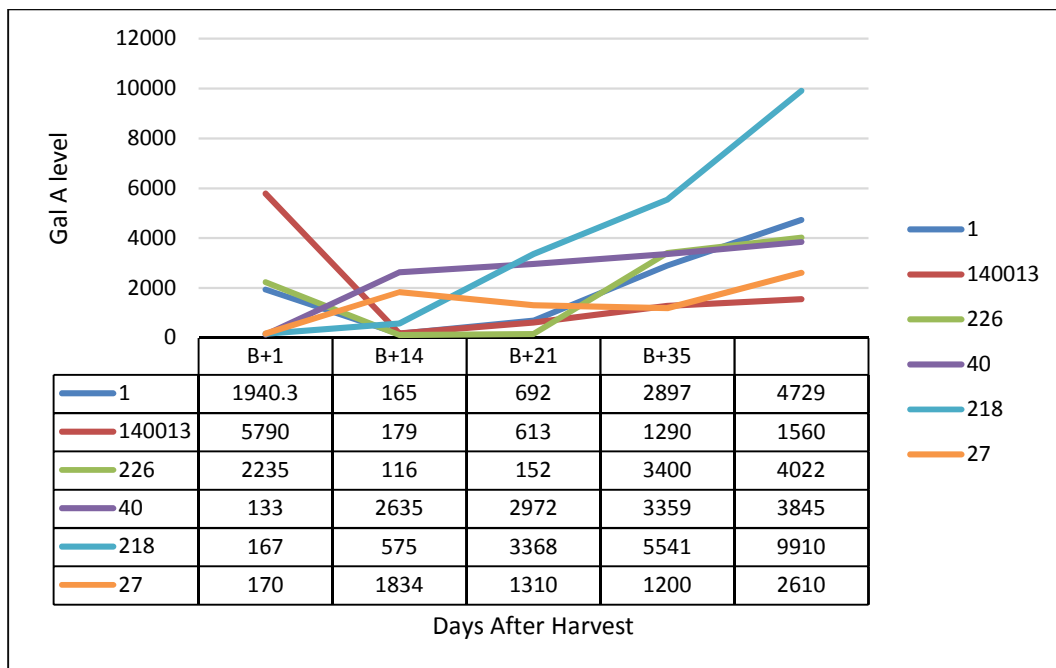| Days After Harvest	1		7		14		21		28		35		42	
RF_number	chl	lyc	chl	lyc	chl	lyc	chl	lyc	chl	lyc	chl	lyc	chl	lyc
RF_001	-0.02	-0.56	-0.52	0.20	-0.58	0.53	-0.59	0.54	-0.59	0.53	-0.60	0.50		
RF_002	0.09	-0.31	-0.57	0.44	-0.58	0.56	-0.60	0.57	-0.60	0.53	-0.59	0.54		
RF_003	-0.06	-0.13	-0.35	0.56	-0.36	0.70	-0.39	0.71	-0.45	0.72				
RF_004	-0.10	-0.64	-0.51	-0.67	-0.53	-0.64	-0.51	-0.61	-0.54	-0.61				
RF_005	0.03	-0.41	-0.55	0.43	-0.58	0.57	-0.61	0.59	-0.59	0.54				
RF_006	-0.01	-0.54	-0.60	-0.53	-0.62	-0.35	-0.62	-0.27						
RF_007	0.04	-0.48	-0.33	-0.47	-0.32	-0.34	-0.40	-0.34	-0.40	-0.30	-0.42	-0.29		
RF_008	-0.10	-0.47	-0.64	-0.40	-0.62	-0.22	-0.65	-0.19	-0.63	-0.11	-0.61	-0.04		
RF_011	-0.11	-0.31	-0.50	0.49	-0.53	0.66	-0.55	0.58	-0.50	0.56	-0.52	0.55	-0.53	0.55
RF_012	-0.16	-0.53	-0.58	0.34	-0.59	0.49	-0.58	0.51	-0.59	0.49				
RF_013	0.25	-0.48	-0.30	0.14	-0.56	0.52	-0.57	0.60	-0.56	0.58	-0.55	0.58		
RF_014	-0.09	-0.59	-0.45	-0.49	-0.47	-0.58	-0.45	-0.54	-0.49	-0.22				
RF_015	-0.17	-0.41	-0.54	0.43	-0.58	0.56	-0.59	0.56	-0.59	0.57	-0.59	0.57		
RF_016	0.41	-0.16	-0.20	0.57	-0.30	0.67	-0.30	0.63	-0.27	0.64				
RF_017	-0.03	0.05	-0.44	0.60	-0.45	0.66	-0.48	0.71	-0.46	0.71				
RF_018	0.08	-0.19	-0.44	0.29	-0.58	0.57	-0.58	0.60	-0.58	0.61	-0.59	0.57		
RF_019	0.04	-0.57	-0.52	-0.65	-0.59	-0.60	-0.62	-0.53	-0.62	-0.44	-0.61	-0.40	-0.56	
RF_020	-0.15	-0.26	-0.62	0.47	-0.63	0.52	-0.63	0.56						
RF_021	0.10	-0.16	-0.53	0.44	-0.59	0.56	-0.59	0.57						
RF_022	0.20	-0.01	-0.06	0.30	-0.14	0.38	-0.11	0.45	-0.17	0.43	-0.18	0.46		
RF_023	0.45	0.28	-0.01	0.63	-0.03	0.67	-0.14	0.67	-0.21	0.69	-0.23	0.46		
RF_024	0.22	-0.30	0.00	0.01	-0.32	0.34	-0.45	0.42						
RF_026	0.42	-0.46	-0.16	0.08	-0.57	0.46	-0.58	0.55	-0.59	0.55	-0.60	0.55	-0.60	0.57
RF_027	-0.25	0.01	-0.59	0.58	-0.57	0.57	-0.57	0.58	-0.58	0.59	-0.57	0.66		
RF_028	0.11	-0.23	-0.60	0.47	-0.63	0.54	-0.62	0.53						
RF_029	0.41	-0.01	0.44	0.75	0.39	0.80	0.27	0.80	0.21	0.79				
RF_030	-0.14	-0.33	-0.60	0.49	-0.61	0.55	-0.63	0.55	-0.66	0.58	-0.64	0.55		
RF_032	-0.39	-0.71	-0.63	-0.70	-0.62	-0.68	-0.61	-0.67						
RF_033	-0.29	-0.24	-0.36	-0.02	-0.64	0.46	-0.63	0.47						
RF_034	0.24	-0.10	-0.51	0.53	-0.55	0.61	-0.55	0.65	-0.63	0.48				
RF_035	0.00	-0.07	-0.22	0.43	-0.38	0.52	-0.39	0.55	-0.56	0.65	-0.56			
RF_036	-0.18	0.17	-0.52	0.70	-0.55	0.70	-0.57	0.70	-0.59	0.73	-0.58	0.70		
RF_037	0.54	0.05	0.33	0.55	0.27	0.64	0.25	0.64						
RF_038	0.06	-0.20	-0.57	0.47	-0.62	0.58	-0.60	0.61	-0.62	0.61	-0.66	0.61		
RF_039	0.26	-0.17	-0.05	0.58	-0.10	0.72	-0.15	0.71	-0.61	0.57	-0.60	0.59	-0.60	0.56
RF_040	-0.32	-0.53	-0.65	0.33	-0.65	0.45	-0.63	0.48	-0.68	0.70	-0.66	0.48		
RF_041	-0.04	-0.44	-0.47	0.44	-0.52	0.67	-0.55	0.58	-0.68	0.47	-0.66	0.46		
RF_043	0.24	-0.27	-0.08	0.36	-0.31	0.70	-0.35	0.77						
RF_045	0.11	-0.44	-0.51	0.30	-0.52	0.55	-0.51	0.61	-0.53	0.67	-0.51	0.59	-0.53	0.70
RF_046	0.47	-0.02	-0.09	0.38	-0.27	0.36	-0.27	0.48	-0.28	0.53				
RF_077	-0.17	-0.13	-0.43	0.35	-0.47	0.55	-0.55	0.66	-0.59	0.59	-0.59	0.59		
RF_078	0.19	-0.34	-0.46	0.49	-0.53	0.60	-0.54	0.61						
RF_088	-0.02	0.04	-0.50	0.63	-0.52	0.63	-0.54	0.63	-0.54	0.64	-0.55	0.62		
RF_089	-0.02	-0.15	-0.48	0.52	-0.58	0.63	-0.58	0.61						
RF_090	-0.47	-0.52	-0.60	-0.37	-0.61	-0.24	-0.61	-0.12	-0.62	-0.09	-0.60	0.00	-0.61	-0.04

Continuation of table S 1.

RF_091	0.06	-0.48	-0.63	0.43	-0.64	0.51	-0.64	0.55	-0.66	0.58	-0.66	0.57		
RF_093	-0.02	-0.61	-0.59	-0.64	-0.61	-0.52	-0.60	-0.40	-0.63	-0.36	-0.60	-0.29		
RF_094	0.19	-0.41	-0.41	0.43	-0.53	0.59	-0.56	0.61	-0.55	0.61	-0.57	0.60		
RF_096	-0.44	0.16	-0.63	0.57	-0.62	0.60	-0.65	0.67	-0.32	0.68				
RF_097	-0.09	-0.61	-0.64	-0.26	-0.65	0.17	-0.64	0.36	-0.64	0.45	-0.64	0.40		
RF_102	0.06	0.10	-0.19	0.50	-0.23	0.55	-0.28	0.50	-0.36	0.49	-0.46	0.38		
RF_103	-0.05	-0.37	-0.42	0.64	-0.41	0.71	-0.45	0.74	-0.47	0.76	-0.48	0.76		
RF_201	0.04	-0.57	-0.52	-0.65	-0.59	-0.60	-0.62	-0.53	-0.62	-0.44	-0.61	-0.40	-0.56	0.10
RF_202	0.57	0.10	0.52	0.66	0.51	0.70	0.52	0.73	0.36	0.84				
RF_203	-0.53	-0.74	-0.57	-0.74	-0.57	-0.73	-0.57	-0.72						
RF_204	-0.06	-0.65	-0.58	-0.54	-0.58	-0.70	-0.59	-0.70	-0.60	-0.70				
RF_205	0.50	0.21	0.46	0.73	0.36	0.83	0.23	0.81	-0.22	0.78				
RF_206	0.02	-0.18	-0.54	0.67	-0.55	0.71	-0.57	0.70	-0.57	0.67				
RF_207	0.04	-0.42	-0.61	0.54	-0.61	0.55	-0.61	0.53	-0.53	0.53	-0.63	0.68		
RF_208	-0.07	-0.17	-0.58	0.58	-0.60	0.65	-0.60	0.63	-0.60	0.57	-0.60	0.68		
RF_209	-0.35	-0.35	-0.59	0.44	-0.61	0.56	-0.61	0.56	-0.61	0.53	-0.61	0.52	-0.60	0.51
RF_210	-0.08	0.01	-0.59	0.61	-0.60	0.61	-0.60	0.62						
RF_211	-0.35	0.10	-0.50	0.55	-0.51	0.55								
RF_212	-0.03	-0.46	-0.55	0.39	-0.59	0.46	-0.59	0.51	-0.59	0.51	-0.59	0.52		
RF_213	0.03	-0.41	-0.29	-0.49	-0.60	-0.72	-0.59	-0.71	-0.60	-0.72	-0.61	-0.71		
RF_214	0.16	-0.39	-0.51	0.49	-0.57	0.64	-0.56	0.62						
RF_215	0.00	-0.09	-0.03	-0.10	0.00	-0.09	-0.06	-0.07						
RF_216	-0.27	0.24	-0.58	0.63	-0.58	0.61	-0.58	0.63	-0.59	0.64	-0.58	0.62	-0.59	0.63
RF_217	-0.14	-0.36	-0.62	0.57	-0.62	0.63	-0.62	0.61	-0.57	0.68	-0.63	0.67		
RF_218	0.19	-0.33	-0.60	0.53	-0.62	0.62	-0.61	0.66						
RF_219	0.30	-0.20	-0.45	0.59	-0.51	0.62	-0.51	0.63	-0.50	0.62	-0.37	0.60		
RF_220	-0.05	0.03	-0.43	0.69	-0.40	0.75	-0.45	0.75						
RF_221	-0.44	-0.06	-0.61	0.53	-0.59	0.67	-0.61	0.61						
RF_222	0.11	-0.27	-0.57	0.70	-0.58	0.69	-0.58	0.65						
RF_223	-0.17	-0.49	-0.59	0.42	-0.59	0.56	-0.59	0.57	-0.60	0.49				
RF_224	0.26	-0.46	-0.60	0.57	-0.62	0.66	-0.63	0.70	-0.54	0.70	-0.60	0.71		
RF_225	-0.15	-0.21	-0.61	0.48	-0.64	0.56	-0.62	0.50	-0.57	0.62				
RF_226	0.36	-0.10	-0.02	0.26	-0.23	0.43	-0.29	0.47	-0.38	0.49	-0.33	0.56	-0.44	0.52
RF_227	0.16	-0.32	-0.39	0.46	-0.53	0.59	-0.55	0.61	-0.55	0.61	-0.57	0.59	-0.59	0.55
RF_229	-0.03	0.01	-0.41	0.71	-0.39	0.78	-0.43	0.77	-0.46	0.76	-0.44	0.76		
RF_230	0.12	-0.56	-0.25	-0.05	-0.50	0.42	-0.54	0.49	-0.57	0.42	-0.55	0.42		
RF_231	-0.31	-0.33	-0.29	-0.69	-0.31	-0.70	-0.41	-0.71	-0.45	-0.71	-0.49	-0.73		
RF_232	-0.32	-0.36	-0.62	0.44	-0.61	0.54	-0.62	0.53	-0.61	0.52				
RF_233	-0.50	0.19	-0.61	0.35	-0.61	0.38	-0.58	0.47	-0.61	0.47				
RF_235	-0.39	0.41	-0.40	0.40	-0.39	0.37	-0.29	0.31						
RF_236	-0.37	-0.11	-0.54	0.43	-0.60	0.59	-0.61	0.60	-0.61	0.58	-0.60	0.57		
RF_237	-0.06	-0.13	-0.39	0.21	-0.36	0.61	-0.39	0.30	-0.60	0.55	-0.58	0.55		
RF_238	0.10	-0.24	-0.46	0.61	-0.48	0.69	-0.48	0.69	-0.50	0.69				
RF_140013	-0.04	-0.49	-0.60	0.30	-0.59	0.44	-0.61	0.60	-0.62	0.63	-0.60	0.58		
RF_140014	-0.29	-0.55	-0.65	0.40	-0.64	0.48	-0.64	0.48	-0.63	0.48	-0.65	0.58		

Table S 2. Variation of primary metabolites in six selected genotypes at five postharvest time points.

Annotated	Retention time	RF-001 gs	RF-001 b1	RF-001 b14	RF-001 b21	RF-001 b35	RF-140013gs	RF-140013b1	RF-140013b14	RF-140013b21	RF-140013b35	RF-226gs	RF-226b1	RF-226b14	RF-226b21	RF-226b35
Nicotinic acid	6154300	1329	1067	898	891	926	955	1123	1090	1052	1003	965	993	1088	885	1478
Suberate	6249290	759	755	762	769	774	805	800	786	796	814	739	830	778	798	794
lactate	6402620	469	480	418	753	581	466	834	504	1116	834	357	477	455	544	461
alanine	7041760	298	227	253	552	1524	320	183	692	1062	4237	74	123	204	163	491
Threonine	7230920	418	389	367	337	310	539	413	403	370	348	364	416	406	336	316
maleate	7610070	4013	4850	4901	3511	5952	4499	4908	5208	3318	6045	4680	4752	3603	4511	5163
Phosphate	9451660	6047	7977	6708	7244	7521	8558	10467	8700	8120	10225	4729	5153	6123	6157	6648
Glycine	9957480	163	185	127	138	220	189	235	132	202	435	47.2883	47.9724	62	45.4741	170
Mevalonolactone	10650000	5680	5917	5990	5272	5324	5674	6104	5794	6193	5781	5168	5980	5745	5473	5813
serine	10658300	642	1212	654	644	813	474	1345	756	943	1411	197	316	258	259	337
threonine	10994900	318	512	688	535	774	209	757	416	480	839	77	88	90	94	314
5-Aminocarboxy-4,6-dihydroxypyrimidine	12213200	558	727	457	181	317	635	1062	893	692	711	215	290	215	128	321
malate	12332400	63503	55311	22296	20097	17624	26276	20624	12634	12603	11518	46650	53513	16635	17121	15233
aspartate	12734000	1373	3187	7700	6893	9209	987	4559	7077	5591	9757	391	707	1721	1710	4856
pyroglutamate	12844000	1166	2337	3550	2051	3327	1340	3842	3212	3649	4559	672	970	1086	1054	2885
gaba	12890700	8308	9617	4169	2901	6476	6910	5765	2712	1851	3035	1785	1745	270	143	958
n-acetylglutamic acid	12952400	41.5194	413	1032	659	1007	49.1539	567	1189	866	1217	52	94	381	376	1102
glutamate	13939000	649	5156	22099	16493	21623	506	7689	20144	12992	21336	395	871	5415	5618	16559
phenylalanine	13946500	303	264	319	253	385	96	287	149	298	430	52	48	120	102	300

Continuation of table S 2.

Annotated	Retention time	RF-001 gs	RF-001 b1	RF-001 b14	RF-001 b21	RF-001 b35	RF-140013gs	RF-140013b1	RF-140013b14	RF-140013b21	RF-140013b35	RF-226gs	RF-226b1	RF-226b14	RF-226b21	RF-226b35
Putrescine	15296400	250	251	299	272	324	222	518	580	548	750	494	652	844	790	1119
glutamine	15693100	1859	4599	6639	3150	5224	1949	6979	4428	4464	7830	710	951	963	1288	2525
Citrate	16124700	42224	57192	43773	41531	51786	45512	67762	38883	28788	32854	57263	72690	69721	51246	84338
dehydroascorbate	16479700	48.8139	930	675	969	944	640	1099	946	756	1074	46.8009	807	920	749	1034
Fructose	16628900	177343	329187	123340	232482	179420	160740	238106	159373	116640	139423	116452	132895	152730	90232	134017
Galactose	16827200	471	574	2682	4245	2927	131	187	712	437	245	75	156	187	97	69
Glucose	16892200	288129	409239	105109	286064	146702	275336	351339	158766	122872	115287	147268	132048	130815	61625	82006
Mannose	17097200	44656	70466	19655	43941	26587	44958	55578	29156	23182	21608	28260	25288	23810	11886	16380
Lysine	17217200	222	227	362	264	448	179	413	345	293	656	151	141	194	190	413
galacturonate	17368000	165	692	2897	4729	5790	179	613	1290	1560	2235	116	152	3400	4022	3667
Hexadecanoic acid	18512100	792	680	610	682	744	822	776	739	687	868	698	756	701	618	642
inositol	18813000	3030	4224	1929	4286	2767	2616	3086	1800	1320	1360	968	796	344	223	1491
Glucopyranose	21620300	172	101	918	4150	3218	88	162	203	264	455	248	206	312	2874	371
3-O-Methylglucose	21630300	172	101	918	4150	3218	88	162	203	264	455	248	206	312	2874	371
Sucrose	23305300	11163	14907	4430	4230	4022	10542	12330	8618	3860	5086	10872	10334	6018	3423	6028

Continuation of Table S 2.

Annotated	Retention time	RF-040gs	RF-040b1	RF-040b14	RF-040b21	RF-040b35	RF-218gs	RF-218b1	RF-218b14	RF-218b21	RF-218b35	RF-027gs	RF-027b1	RF-027b14	RF-027b21	RF-027b35
Nicotinic acid	6154300	1035	938	1085	1048	1152	956	907	1025	927	1055	938	970	968	793	1080
Suberate	6249290	801	798	766	793	740	790	772	749	784	808	729	821	804	718	834
lactate	6402620	445	499	534	490	490	792	408	447	519	449	445	404	1020	380	590
alanine	7041760	159	759	269	1005	1530	479	271	415	776	2916	840	1264	582	941	2483
Threonine	7230920	387	321	406	317	329	447	362	406	383	388	444	338	349	333	412
maleate	7610070	4486	5649	3466	5748	5725	4817	5418	4901	4225	5353	4650	5519	3702	5599	5872
Phosphate	9451660	5150	5371	7118	7112	6748	7224	6903	6960	7569	6568	9902	9569	8722	6428	8880
Glycine	9957480	148	162	152	199	253	220	81	57	107	256	430	184	153	128	273
Mevalonolactone	10650000	5629	5628	5893	6105	5871	5921	5987	5835	5962	5897	6021	5838	5524	5731	5772
serine	10658300	617	676	808	858	919	721	438	280	361	486	1168	915	816	713	1025
threonine	10994900	229	404	505	563	525	497	361	285	409	343	583	680	748	498	964
5-Aminocarboxy-4,6-dihydroxypyrimidine	12213200	423	294	673	470	456	658	268	381	149	339	1165	374	500	271	732
malate	12332400	16415	7133	9950	9268	7614	54654	22744	8489	6670	3324	17565	8425	8797	5815	5219
aspartate	12734000	501	3484	3710	4838	4891	1936	3085	4208	5005	4308	2520	9998	10224	6164	14530
pyroglutamate	12844000	993	2369	3143	2770	2674	1820	2076	2380	1815	2055	2891	3916	3854	2539	4810
gaba	12890700	5173	4948	4092	6417	4612	10423	5844	2967	1207	2311	6708	2100	2186	1352	3573
n-acetylglutamic acid	12952400	44.984	460	512	604	588	73	459	1308	738	1128	92	830	923	565	1119
glutamate	13939000	287	7629	9989	11260	10132	720	8131	20158	16897	15223	890	17174	19816	12270	25593
phenylalanine	13946500	155	234	438	311	432	570	377	389	626	526	767	612	923	448	983

Continuation of table S 2.

Annotated	Retention time	RF-040gs	RF-040b1	RF-040b14	RF-040b21	RF-040b35	RF-218gs	RF-218b1	RF-218b14	RF-218b21	RF-218b35	RF-027gs	RF-027b1	RF-027b14	RF-027b21	RF-027b35
Putrescine	15296400	663	792	788	956	878	423	390	357	233	268	206	307	275	192	291
glutamine	15693100	1304	3154	4194	4747	3623	3281	3371	2526	2339	2102	5747	5879	6472	3989	6802
Citrate	16124700	38735	35413	57395	49585	41333	48208	53662	51152	39670	30888	56985	28832	28703	18160	27218
dehydroascorbate	16479700	848	1032	1152	1381	1488	602	612	43.0265	1051	790	613	959	741	549	924
Fructose	16628900	109759	105713	149303	153929	134964	188284	214098	265298	291038	219111	193781	129111	139070	87426	125693
Galactose	16827200	221	985	1725	1437	913	78	1160	8165	9012	10952	114	242	128	382	221
Glucose	16892200	147504	107224	167991	157542	126194	321919	318876	277254	289259	199551	338740	126502	152968	82027	89506
Mannose	17097200	28317	19982	29785	28360	23834	50682	47743	43016	43521	34033	52746	24017	29137	16508	16886
Lysine	17217200	260	290	356	413	420	198	185	275	270	351	360	468	372	285	863
galacturonate	17368000	133	2635	2972	3359	3845	167	575	3368	5541	9910	170	1834	1310	1200	2610
Hexadecanoic acid	18512100	715	712	693	653	680	757	744	710	649	723	831	659	667	683	688
inositol	18813000	1138	840	1850	1121	1184	3934	1521	2155	2866	1452	2862	1652	1618	1196	1617
Glucopyranose	21620300	129	1632	204	682	292	232	1130	420	4512	397	104	1868	637	6419	440
3-O-Methylglucose	21630300	129	1632	204	682	292	232	1130	420	4512	397	104	1868	637	6419	440
Sucrose	23305300	5711	2274	4414	3179	3418	13835	10289	8317	4840	2837	7397	2393	3778	2110	3272

gs: green stage; b1: brekaer+1; b14: breaker+14; b21: breaker+21; b35: breaker+35

References

Aflitos, S., Schijlen, E., de Jong, H., de Ridder, D., Smit, S., Finkers, R., . . . Peters, S. (2014). **Exploring genetic variation in the tomato (Solanum section Lycopersicon) clade by whole-genome sequencing.** *Plant J, 80*(1), 136-148.

Apelbaum, A., Icekson, I., Burgoon, A. C., & Lieberman, M. (1982). **Inhibition by polyamines of macromolecular synthesis and its implication for ethylene production and senescence processes.** Plant Physiol, 70(4), 1221-1223.

Barman, K., Asrey, R., & Pal, R. K. (2011). **Putrescine and carnauba wax pretreatments alleviate chilling injury, enhance shelf life and preserve pomegranate fruit quality during cold storage.** *Sci Hortic, 130,* 795–800.

Beckles, D. M. (2012). **Factors affecting the postharvest soluble solids and sugar content of tomato (Solanum lycopersicum L.) fruit.** *Postharvest Biol Technol, 63*, 129-140.

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., . . . Cañizares, J. (2015). **Genomic variation in tomato, from wild ancestors to contemporary breeding accessions.** *BMC Genom, 16*(1), 257.

Brummell, D. A., & Harpster, M. H. (2001). **Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants.** *Plant Mol Biol, 47*(1-2), 311-340.

Brummell, D. A. (2005). **Regulation and genetic manipulation of ripening in climacteric fruit.** *Stewart Postharvest Review, 1*(3), 1.

Brummell, D. A. (2006). **Cell wall disassembly in ripening fruit.** *Funct Plant Biol, 33*(2), 103-119.

Carreno-Quintero, N., Acharjee, A., Maliepaard, C., Bachem, C., Mumm, R., Bouwmeester, H., . . . Keurentjes, J. (2012). **Untargeted metabolic quantitative trait loci (mQTL) analyses reveal a relationship between primary metabolism and potato tuber quality.** *Plant Physiol*, 158(3), 1306-1318.

Cosgrove, D. J. (2000). **Loosening of plant cell walls by expansins.** *Nature, 407*(6802), 321-326.

Dibble, A. R. G., Davies, P. J., & Mutschler, M. A. (1988). **Polyamine Content of long-keeping Alcobaca Tomato Fruit.** *Plant Physiol, 86*(2), 338-340.

Farneti, B., Schouten, R. E., & Woltering, E. J. (2012). **Low temperature-induced lycopene degradation in red ripe tomato evaluated by remittance spectroscopy.** *Postharvest Biol Technol, 73*, 22-27.

Giovannoni, J. J., DellaPenna, D., Bennett, A. B., & Fischer, R. L. (1989**). Expression of a chimeric polygalacturonase gene in transgenic rin (ripening inhibitor) tomato fruit results in polyuronide degradation but not fruit softening.** Plant Cell, 1(1), 53-63.

Giovannoni, J. (2001). **Molecular biology of fruit maturation and ripening.** *Annu Rev Plant Biol,* 52, 725-749.

Goulao, L. F., & Oliveira, C. M. (2008). **Cell wall modifications during fruit ripening: when a fruit is not the fruit.** *Trends Food Sci Technol, 19*(1), 4-25.

Gross, K. C. (1983). **Changes in free galactose, myo-inositol and other monosaccharides in normal and non-ripening mutant tomatoes.** *Phytochemistry, 22*(5), 1137-1139.

Gross, K. C. (1985). **Promotion of ethylene evolution and ripening of tomato fruit by galactose.** *Plant Physiol, 79*(1), 306-307.

Gross, K. C., & Sams, C. E. (1984). **Changes in cell wall neutral sugar composition during fruit ripening: a species survey.** *Phytochemistry, 23*(11), 2457-2461.

Gross, K. C., & Wallner, S. J. (1979). **Degradation of cell wall polysaccharides during tomato fruit ripening.** *Plant Physiol, 63*(1), 117-120.

Jia, B., Zheng, Q., Zuo, J., Gao, L., Wang, Q., Guan, W., & Shi, J. (2018). **Application of postharvest putrescine treatment to maintain the quality and increase the activity of antioxidative enzyme of cucumber.** *Scientia Horticulturae, 239*, 210-215.

Houben, K., Jolie, R. P., Fraeye, I., Van Loey, A. M., & Hendrickx, M. E. (2011). **Comparative study of the cell wall composition of broccoli, carrot, and tomato: structural characterization of the extractable pectins and hemicelluloses.** *Carbohydr Res, 346*(9), 1105-1111.

Kim, J., Gross, K. C., & Solomos, T. (1991). **Galactose metabolism and ethylene production during development and ripening of tomato fruit.** *Postharvest Biol Technol, 1*(1), 67-80.

Kopeliovitch, E., Rabinowitch, H. D., Mizrahi, Y., & Kedar, N. (1979). **The potential of ripening mutants for extending the storage life of the tomato fruit.** *Euphytica, 28*(1), 99-104.

Koushesh saba, M., Arzani, K., & Barzegar, M. (2012). **Postharvest polyamine application alleviates chilling injury and affects apricot storage ability.** *J Agric Food Chem, 60*(36), 8947-8953.

Kuckenberg, J., Tartachnyk, I., & Noga, G. (2008). **Evaluation of fluorescence and remission techniques for monitoring changes in peel chlorophyll and internal fruit characteristics in sunlit and shaded sides of apple fruit during shelf-life.** *Postharvest Biol Technol, 48*(2), 231-241.

Labavitch, J. M. (1981). **Cell wall turnover in plant development**. *Ann. Rev. Plant PhysioL, 31*, 385-406.

Lackey, G. D., Gross, K. C., & Wallner, S. J. (1980). **Loss of tomato cell wall galactan may involve reduced rate of synthesis.** Plant Physiol, 66(3), 532-533.

Lahaye, M., Quemener, B., Causse, M., & Seymour, G. B. (2012). **Hemicellulose fine structure is affected differently during ripening of tomato lines with contrasted texture.** *Int J Biol Macromol, 51*(4), 462-470.

Lunn, D., Phan, T. D., Tucker, G. A., & Lycett, G. W. (2013). **Cell wall composition of tomato fruit changes during development and inhibition of vesicle trafficking is associated with reduced pectin levels and reduced softening.** *Plant Physiol Biochem, 66*, 91-97.

Mahajan, P. V., Caleb, O. J., Singh, Z., Watkins, C. B., & Geyer, M. (2014). **Postharvest treatments of fresh produce.** *Philos Trans A Math Phys Eng Sci, 372*(2017), 20130309.

Markovic, Z., Zdravkovic, J., Cvikic, D., Pavlovic, N., & Zdravkovic, M. (2012). **Breeding new F1 tomato hybrids of rin genotype.** *Acta Hortic, 960, 165-168.*

McCartney, L., Ormerod, A. P., Gidley, M. J., & Knox, J. P. (2000). **Temporal and spatial regulation of pectic (1-->4)-beta-D-galactan in cell walls of developing pea cotyledons: implications for mechanical properties.** *Plant J, 22*(2), 105-113.

McGlasson, W., Sumeghy, J., Morris, L., McBride, R., Best, D., & Tigchelaar, E. (1983). **Yield and evaluation of F1 tomato hybrids incorporating the non-ripening nor gene.** *Aust J Exp Agric, 23*(120), 106-112.

Mitcham, E. J., Gross, K. C., & Ng, T. J. (1989). **Tomato fruit cell wall synthesis during development and senescence.** *In Vivo Radiolabeling of Wall Fractions Using ^{14}C Sucrose. Plant physiol, 89*(2), 477-481.

Mitcham, E. J., Gross, K. C., & Ng, T. J. (1991). **Ripening and cell wall synthesis in normal and mutant tomato fruit.** *Phytochemistry, 30*(6), 1777-1780.

Nguyen, V. Q., Ashcroft, W. J., Jones, K. H., & McGlasson, W. B. (1991). **Evaluation of F1 hybrids incorporating the rin (ripening inhibitor) gene to improve the storage life and fruit quality of fresh market tomatoes (Lycopersicon esculentum Mill.).** *Aust J Exp Agric, 31*, 407-413.

Prasanna, V., Prabha, T. N., & Tharanathan, R. N. (2007). **Fruit ripening phenomena-an overview.** *Crit. Rev. Food Sci. Nutr., 47*(1), 1-19.

Rastogi, R., & Davies, P. J. (1990). **Polyamine metabolism in ripening tomato fruit : I. Identification of metabolites of putrescine and spermidine.** *Plant Physiol, 94*(3), 1449-1455.

Rastogi, R., & Davies, P. J. (1991). **Polyamine metabolism in ripening tomato fruit: II. Polyamine metabolism and synthesis in relation to enhanced putrescine content and storage life of tomato fruit***. Plant Physiol, 95*(1), 41-45.

Rick, C. M., & Chetelat, R. T. (1995). **utilization of related wild species for tomato improvement.** *Acta Hortic, 412*.

Sandarani, M., Dasanayaka, D., & Jayasinghe, C. (2018). **Strategies used to prolong the shelf Life of fresh Commodities.** *J AGR FOOD RES, 9*(1), 1-6.

Schouten, R. E., Farneti, B., Tijskens, L. M. M., Alarcón, A. A., & Woltering, E. J. (2014). **Quantifying lycopene synthesis and chlorophyll breakdown in tomato fruit using remittance VIS spectroscopy.** *Postharvest Biol Technol, 96*, 53-63.

Serrano, M., Martinez-Romero , D., Guillen , F., & Valero , D. (2003). **Effects of exogenous putrescine on improving shelf life of four plum cultivars.** *Postharvest Biol Technol, 30* 259-271.

Serrano, M., & Valero, D. (2018). **Application of Polyamines to Maintain Functional Properties in Stored Fruits.** *Methods Mol Biol, 1694*, 449-458.

Seymour, G. B., Colquhoun, I. J., Dupont, M. S., Parsley, K. R., & Selvendran, R. (1990). **Composition and structural features of cell wall polysaccharides from tomato fruits.** *Phytochemistry, 29*(3), 725-731.

Sharma, S., Pareek, S., Sagar, N. A., Valero, D., & Serrano, M. (2017). **Modulatory Effects of Exogenously Applied Polyamines on Postharvest Physiology, Antioxidant System and Shelf Life of Fruits: A Review.** *Int J Mol Sci, 18*(8).

Smith, T. A. (1985). **Polyamines.** Annu Rev Plant Biol*, 36*(1), 117-143.

Smith, D. L., Abbott, J. A., & Gross, K. C. (2002). **Down-regulation of tomato beta -galactosidase 4 results in decreased fruit softening.** *Plant Physiol, 129*(4), 1755-1762.

Tikunov, Y., Lommen, A., de Vos, C. H. R., Verhoeven, H. A., Bino, R. J., Hall, R. D., & Bovy, A. G. (2005). **A novel approach for nontargeted data analysis for metabolomics.** Large-Scale Profiling of Tomato Fruit Volatiles. *Plant Physiol, 139*(3), 1125-1137.

Ulvskov, P., Wium, H., Bruce, D., Jorgensen, B., Qvist, K. B., Skjot, M., . . . Sorensen, S. O. (2005). **Biophysical consequences of remodeling the neutral side chains of rhamnogalacturonan I in tubers of transgenic potatoes.** *Planta, 220*(4), 609-620.

Uluisik, S., Chapman, N. H., Smith, R., Poole, M., Adams, G., Gillis, R. B., . . . Seymour, G. B. (2016). **Genetic improvement of tomato by targeted control of fruit softening.** *Nat. Biotechnol 34*(9), 950-952.

van Arkel, J., Vergauwen, R., Sévenier, R., Hakkert, J. C, van Laere, A., Bouwmeester, H. J, Koops, A. J, & van der Meer, I. M. (2012). **Sink filling, inulin metabolizing enzymes and carbohydrate status in field grown chicory (Cichorium intybus L.).** *plant physiol, 169*, 1520-1529.

Wannabussapawich, B., & Seraypheap, K. (2018). **Effects of putrescine treatment on the quality attributes and antioxidant activities of 'Nam Dok Mai No.4' mango fruit during storage.** *Scientia Horticulturae, 233*, 22-28.

Zykwinska, A., Thibault, J. F., & Ralet, M. C. (2007). **Organization of pectic arabinan and galactan side chains in association with cellulose microfibrils in primary cell walls and related models envisaged.** *J Exp Bot, 58*(7), 1795-1802.

Chapter 4

Fine mapping of a major QTL controlling phenylalanine-derived flavour volatiles in tomato

Raana Roohanitaziani[1,2], Yury Tikunov [1], Fien Meijer-Dekens [1], Jos Molthoff [1], Richard G.F. Visser[1] and Arnaud G. Bovy[1]

[1] Plant Breeding, Wageningen University and Research, PO Box 386, 6700 AJ Wageningen, The Netherlands

[2] Graduate School Experimental Plant Sciences, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

Abstract

An important class of volatile organic compounds (VOCs) are phenolic volatiles, derived from the amino acid phenylalanine. Among these phenolic compounds, 2-phenylethanol is considered one of the most important volatiles and is often associated with fruity or floral notes of aroma in tomato fruit. We aimed to elucidate the genetic regions controlling this metabolite and identify the encoding genes involved in the biosynthesis of this metabolite in cultivated tomato. A marker-trait association study on a collection of 94 tomato cultivars revealed a genetic region on chromosome 4 strongly associated with quantitative variation of phenolic volatiles. In the present study we have developed a fine mapping population based on RILs contrasting in this region. Genetic and metabolic analyses of this population allowed to narrow down the region associated with the phenolic volatiles from 54.52 to 54.63 Mb harbouring 11 candidate genes. Among the 11 genes we identified Solyc04g063350 (*Dcx1*) gene, encoding a putative 2-oxoacid dehydrogenase enzyme that may possess amino acid decarboxylation activity, as a major candidate gene for the variation of phenylethanol content in our mapped interval. Gene expression analysis of fruits of the parental genotypes of the crossing population and fruits derived from contrasting plants of F3 families showed a higher expression of *Dcx1* in fruits with a higher content of the phenolic volatiles. To test the function of the *Dcx1* gene a VIGS (Virus Induced Gene Silencing) experiment was carried out by injection of a VIGS construct in detached immature green fruits of tomato cultivar "Solarino" which has the positive *Dcx1* allele, a high *Dcx1* expression and a high content of the phenolic volatiles in fruits. Metabolic analysis of silenced fruits revealed a severe reduction in the levels of all three phenolic volatiles 2-phenylethanol, phenylacetaldehyde and 2-nitrophenylethane.

Key words: Tomato, *S. lycopersicum*, fine mapping, 2-phenylethanol, Phenolic VOCs

Introduction

The flavour of fruits, including tomato, is determined by a combination of taste, aroma and texture, each of which is determined by different chemical compounds accumulated in the fruit during ripening. Fruit taste is mainly determined by non-volatile primary metabolites, such as sugars, organic acids and free amino acids, as well as minerals (Stevens et al., 1979; Jones & Scott., 1983; Acree, 1993; Yilmaz, 2000). Volatile compounds are responsible for the aroma of the fruits (Baldwin et al., 2000; Baldwin et al., 2008). A few hundred volatile compounds have already been detected in tomato fruits, but only a small proportion of those have been associated with tomato flavour. The majority of flavour-related volatiles in tomato are derived from amino acids, such as phenylalanine, leucine and isoleucine, lipids and carotenoids (Tikunov et al., 2005; Rambla et al., 2014). Several phenylalanine-derived volatiles, also called phenolic volatiles, have been shown to affect tomato fruit flavour either positively or negatively. Among these phenolic compounds, phenylacetaldehyde and 2-phenylethanol are considered the most important volatiles for tomato fruit aroma (Baldwin et.al., 2000; Tieman et al., 2007; Tzin et al., 2013; Rambla et al., 2014). Both volatiles have fruity/floral properties and they are major constituents of scent in many flowers. 2-phenylethanol is the major aroma volatile contributing to the scent of roses and this compound is widely used as fragrance in the cosmetic and food industry (Knudsen et al., 1993; Tieman et al., 2007). Despite their positive association with consumer liking and flavour intensity in tomato fruit, 2-phenylethanol and phenylacetaldehyde can also be perceived as an undesirable flavour at extremely elevated concentrations, as was shown for the *S. pennelli*i introgression line IL8.2 in the cultivated M82 tomato background (Tadmor et al., 2002).

Several studies suggest that the taste of modern tomato cultivars is not as good as the taste of traditional "vintage" or "heirloom" tomatoes, due to many years of breeding focus on producer traits such as crop yield and disease resistance, which has indirectly led to a loss of fruit flavour. Two recent GWAS studies with tomato collections including wild relatives, old varieties and modern cultivars revealed that modern cultivars contain significantly lower amounts of important fruit flavour metabolites, including several volatiles and sugars (Bauchet et al., 2017; Tieman et al., 2017). In recent years consumer traits, like taste and nutritional value, gained increasingly more attention from breeders and researchers. This is facilitated by the advances made in the detection of flavour–related metabolites by mass spectrometry-based metabolomics technologies and the availability of more than 500 sequenced tomato genomes, which makes identification of candidate QTLs and genes responsible for the production of taste-related compounds, for example through GWAS approaches, more feasible (The Tomato Genome Consortium, 2012; Lin et al., 2014). Nevertheless, despite the availability of these X-omics tools, genetic improvement of tomato flavour is not straightforward. Firstly, flavour is a complex trait (Rambla et al., 2014): many

metabolites, derived from different metabolic pathways, contribute to fruit flavour. Each of these metabolites is controlled by one or more QTLs and changing one of these QTLs may not have a major effect on the perception of taste and aroma (Klee & Tieman, 2013). The second issue complicating tomato flavour improvement is the wide range of consumer preferences, which is influenced by personal, regional and cultural differences. Therefore, unlike yield or pathogen resistance, flavour is not a unique single trait.

We previously characterised a tomato collection consisting of 94 tomato genotypes, mostly cultivars and hybrids, to study the biochemical and genetic basis of fruit flavour (Tikunov et al., 2005, 2013; Ursem et al., 2008; Menéndez et al., 2012). These studies confirmed the presence of a strong QTL for phenolic volatiles at the bottom half of chromosome 4, with a QTL maximum at 54.7 Mb (unpublished results). In the present study we aimed to fine map this QTL using an F2 population based on a cross of two RILs contrasting for the 2-phenylethanol QTL on CH4. This narrowed down the QTL region to 110 kb containing 11 genes. We identified Solyc04g063350 (*Dcx1*) which encodes for a putative 2-oxoacid dehydrogenase gene family as a major candidate gene for the variation of phenylethanol content in our mapped interval. Gene expression analysis and VIGS was used to demonstrate that Solyc04g063350 is indeed the key gene underlying the QTL for these three phenolic volatiles.

Materials and Methods

Plant materials and growth conditions

The empirical basis of this project is based on 94 tomato cultivars from the CBSG (The Center For Biosystem Genomics In Netherlands) tomato quality project for which there was a complete dataset consisting of AFLP marker scores, metabolic profiles, organoleptic trait evaluations, plant and fruit morphology measurements and consumer appreciation assessments. All cultivars were greenhouse cultivars provided by the five international breeding companies involved in the project, and almost all were F1 hybrids. Based on the metabolomics analysis of these 94 cultivars, 4 genotypes contrasting for many flavour traits were selected and 6 RIL populations were developed. In this project we have used one of these RIL populations (named population 4) based on one round (R104) and one cherry parent (C085). Since –for IP reasons- we were not allowed to use the original parents of the RIL population, two lines (named 4-066 and 4-128) which were contrasting for as many as possible QTL regions were selected and used as parents to develop a large F2 population of around 5000 individuals. Young leaves of all F2 plants were harvested and used to isolate genomic DNA for recombinant screening using SNP markers. Plants were grown to maturity, F3 seed was collected from each individual plant and stored for further use within

the project. F3 plants used for volatile analysis of ripe fruits were grown in the greenhouses of Unifarm, Wageningen University & Research while the 5000 F2 population plants were grown in the greenhouses of Rijk Zwaan (Fijnaart/de Lier, NL) and Enza Zaden (Enkhuizen, NL).

Development of contrasting NILs

Four F6 lines from population 4 (C85 x R104) have been selected because they were heterozygous in the chromosome 4 QTL region and almost completely homozygous for the rest of the genome. The purpose was to select F7 plants which are homozygous for either of the contrasting alleles (derived from C85 or R104). Up to 20 F7 plants per NIL were grown in the greenhouse of Unifarm (Wageningen University & Research) and selected by Marker Assisted Selection (MAS). Fruits of contrasting F8 NILs were harvested at fully ripe stage and used for both sensory and metabolic analysis. Sensory analysis was performed at Wageningen UR Greenhouse Horticulture, Bleiswijk. A sensory panel consisting 15 trained judges was used to assess the flavour of the NIL's fruits. Metabolic analysis of the fruits was carried out as described below.

Metabolomic Analyses

Six fruits from each F3 genotype were harvested at fully ripe stage. Fruits were harvested starting from the second fruit truss onwards. The six fruits from a plant were cut into small pieces, pooled together and immediately frozen in liquid nitrogen, subsequently ground in an analytical electric mill (IKA A11 basic, Germany) and stored at -80°C before analysis. Volatile compounds were quantified and identified using Gas Chromatography/Mass Spectrometry (GC/MS) as described in (Tikunov et al., 2005). Frozen fruit powder (1 g fresh weight) was weighed in a 5-mL screw-capvial, closed, and incubated at 30 °C for 10 min. An EDTA-NaOH water solution was prepared by adjusting of 100mM EDTA to pH 7.5 with NaOH. Then, 1 mL of the EDTA-NaOH solution was added to the sample. 1.2 gram solid $CaCl_2$ was then immediately added to give a final concentration of 5 M. The closed vials were vortexed by hand for 1 min to allow $CaCl_2$ to dissolve. A 1-mL aliquot was transferred into a 10-mL crimp cap vial (Waters), capped, and used for SPME-GC-MS analysis. In addition, a number of identical analytical quality control samples (QC) were made by mixing fruit powders from several genotypes. TheQCs were routinely analyzed every 10 samples during the entire analysis period as an external control in order to monitor the stability of the analytical system. The samples were automatically extracted and injected into the GC-MS via a Combi PAL autosampler (CTC Analytics AG). Headspace volatiles were extracted by exposing a 65-mm polydimethylsiloxane-divinylbenzene SPME fiber (Supelco) to the vial headspace for 20 min under continuous agitation and heating at 50 °C. The fiber was inserted into a GC 8000 (Fisons Instruments) injection port and volatiles were desorbed for 1 min at 250 °C. Chromatography was performed on an HP-5 (50 m30.32 mm3 .05 mm) column with helium as carrier gas (37 kPa). The GC interface and MS 1 source temperatures were

260 °C and 250 °C, respectively. The GC temperature program began at 45 °C (2min), was then raised to 250 °C at a rate of 10 °C /min, and finally held at 250 °C for 5min. The total run time, including oven cooling, was 30 min. Mass spectra in the 35 to 400 m/z range were recorded by an MD800 electron impact MS (Fisons Instruments) at a scanning speed of 2.8 scans/s and an ionization energy of 70 eV. The chromatography and spectral data were evaluated using Xcalibur software (http://www.thermo.com).

Genotyping

Genomic DNA was extracted from young leaves of individual plants using the CTAB method (Devi et al., 2013)**.** Genotyping was performed by KASP assay using SNP markers in Dr. van Haeringen Laboratorium, Wageningen, the Netherlands.

QTL analysis

QTL analyses were performed with the MapQTL 6.0 software program (Van Ooijen, 2009). The non-parametric Kruskal-Wallis (KW) statistical test was used to detect the association between markers and VOCs level. Of the 89 detected volatiles, 13 were selected for QTL analysis.

Real Time Quantitative PCR

Quantitative PCR was performed on the Solyc04g063350 transcript in an ABI 7500 Fast Real-Time PCR system (Applied Biosystems). RNA was extracted from tomato fruit of eight F3 tomato lines harvested at three different ripening stages: mature green, turning and ripe. RNA was converted into cDNA using reverse transcriptase according to manufacturer's protocol (Invitrogen). Primers used for amplification of the target gene were designed using online Primer3 Plus software. The RT-PCR mixture (duplo) consisted of: 22 µl 2x iQ SYBR Green super mix (Bio-Rad), 11 µl MQ, 4.4 µl forward primer (3 µM), 4.4 µl reverse primer (3 µM) and finally 2.2 µl cDNA (10 ng/µl) for each sample to get a final volume of 44 µl. The house keeping gene β-Actin was used as reference gene for quantification. The relative gene expression for each sample was calculated using the Ct values of the gene of interest and the reference gene.

Results

Developing a stable resource for QTL fine mapping

A F6 RIL population was developed based on a cross between a round (R104) and a cherry (C85) breeding line. Two RILs, contrasting for many genetic regions on all chromosomes were selected as parents for the development of a large F2 population for fine mapping of flavour-related QTLs (Figure 1). The two RILs, 4-066 and 4-128 contrasted for a large part of chromosome 4 (from 5.8 Mb until the end of the chromosome). Ripe fruits of these two RILs also showed contrasting levels of phenolic volatiles, as did the original cherry and round parental breeding lines (Figure S 1). Five thousand individual F2 plants were grown to maturity in the greenhouse. Leaves of every young F2 plant were harvested and DNA was isolated. To immortalize the fine mapping population, F3 seed was harvested from each individual F2 plant and stored for further analysis.
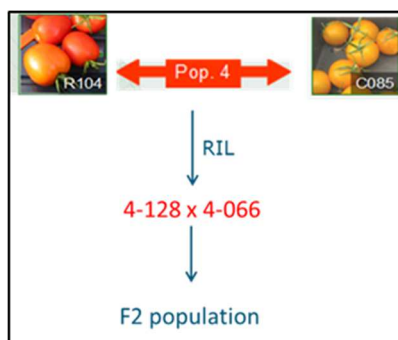
Figure 1. Development of the mapping population based on two contrasting RILs; 4-128 and 4-066.

Selection of recombinants in F2

The first screen of 5000 F2 plants for recombinants in the selected QTL region was carried out with 4 markers spanning the segregating region of chromosome 4 and led to the selection of 384 recombinants (Figure S 2). The 384 F2 recombinants were genotyped with 50 markers covering a 7.7 Mb region from 53.0 to 60.7 Mb which spans the QTL maximum at 54.7 Mb and 50 F2 genotypes with recombinations distributed over the entire QTL region were selected for metabolic analysis in the F3 generation (Figure 2).
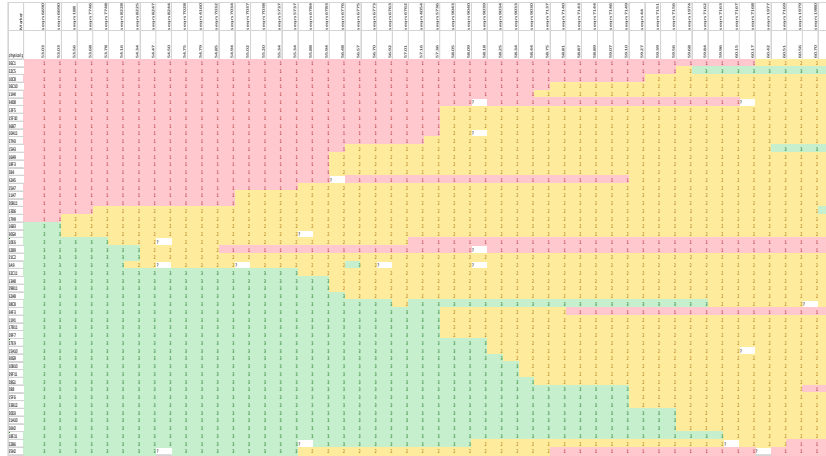
Figure 2. Genotype of 50 F2 recombinants selected for further analysis in the F3 generation. The red colour (A score) represents the homozygous allele of the R104 parent and the green colour (B allele) represents the homozygous allele of parent C085. The yellow colour represents heterozygous regions.

Selection of homozygous recombinants in F3 for phenotyping

For every selected F2 recombinant, 20 F3 seedlings were grown and genotyped with 4 markers in the QTL region to determine the segregation pattern in the F3 generation. An example of this analysis is shown in Figure S 3. From each F3 family, 2 plants with a homozygous recombination in the QTL region were selected for metabolic analysis.

Metabolite profile of F3 fruits

Selected F3 plants were grown to maturity and ripe fruits were harvested for metabolic analysis of flavour-related volatiles, using SPME-GC-MS. Data analysis focussed on 13 important flavour-related VOCs, which show a key role in tomato flavour based on previous studies, including fatty acid-derived VOCs (the C6 volatiles 1-hexanol, (Z)-3-hexenal, (E)-2-hexenal and hexanal, and the C5 volatile 1-penten-3-one), phenylalanine-derived VOCs (2-phenylethanol, phenylacetaldehyde, 2-nitrophenylethane, benzyl alcohol and bezaldehyde), branched-chain amino acid-derived VOCs (2 and 3-methylbutanol and 3-methylbutanoic acid) and isobutylthiazole (Table S1).

QTL discovery

The first round of QTL mapping resulted in identification of 5 significant QTLs, for phenylethanol, phenylacetaldehyde, 2-nitrophenylethane, 2-methyl-1-buthanol and 3-methylbutanoic- acid. Their positions in the region are shown in Table 1. A strong QTL was found for the three phenolic-derived volatiles phenylethanol, phenylacetaldehyde and 2-nitrophenylethane, with a QTL maximum between 54.5 and 54.8 Mb on chromosome 4 (Table 2 and Figure S 4). The phenylethanol values belonging to the different alleles

at one of the most significant markers (seq-rs4160) are shown in Figure 3. The average level of 2-phenylethanol in genotypes with a homozygous C085 allele (b) is 3.5-fold higher than in genotypes with a homozygous R104 allele (a).

Table 1. QTL analysis results for the 5 most significant VOCs in selected F3 recombinants.

Metabolite	Ch	Position (MB)	Locus	K*	Signif.	log P	Mean-a	Mean-b	R square
2-phenylethanol	4	54.5	seq-rs8247	33.785	*******	8.499771	333658	1.16E+06	0.47
phenylacetaldehyde	4	54.5	seq-rs8247	31.792	*******	8.053794	192668	602161	0.457
2-phenylenitroethane1	4	54.5	seq-rs8247	23.155	*******	6.109457	313180	764043	0.317
2-methyl-1-buthanol	4	54.5	seq-rs8247	13.953	******	4.001279	26961.6	54305.4	0.147
3-methylbutanoic- acid	4	58.3	seq-rs9034	16.688	*******	4.634045	40459.5	113331	0.187

Ch: chromosome; Position: physical position of the markers; Locus: name of most significant marker; K*: Kruskal-Wallis K value; LOGP: ^{10}LOG P-value of the Chi-square distribution; Mean a: mean value of "a" allele, mean b: mean value of the "b" allele; R square: Explained variation after regression analysis.

Table 2. QTL analysis of the phenylethanol content using the selected F3 recombinants.

| Position (MB)	Locus	K*	Signif.	log P	Mean-a	Mean-b
53	seq-rs6690	20.979	*******	5.6	421374	1.06E+06
53.6	seq-rs188	22.005	*******	5.8	432519	1.08E+06
53.7	seq-rs7746	22.005	*******	5.8	432519	1.08E+06
53.8	seq-rs7748	26.167	*******	6.8	354902	1.11E+06
54.2	seq-rs8228	27.794	*******	7.2	349454	1.14E+06
54.3	seq-rs8225	30.434	*******	7.7	342305	1.16E+06
54.5	seq-rs8247	33.785	*******	8.5	333658	1.16E+06
54.5	seq-rs8244	33.785	*******	8.5	333658	1.16E+06
54.8	seq-rs7028	33.785	*******	8.5	333658	1.16E+06
54.8	seq-rs4160	33.785	*******	8.5	333658	1.16E+06
54.8	seq-rs7032	33.785	*******	8.5	333658	1.16E+06
54.9	seq-rs7034	30.51	*******	7.8	387489	1.16E+06
55	seq-rs7037	29.747	*******	7.6	387993	1.14E+06
55.2	seq-rs7038	29.747	*******	7.6	387993	1.14E+06
55.3	seq-rs5737	29.747	*******	7.6	387993	1.14E+06
55.9	seq-rs6784	28.575	*******	6.5	409700	1.15E+06
55.9	seq-rs6783	27.645	*******	7.1	409700	1.15E+06
56.5	seq-rs6776	19.734	*******	5.3	458935	1.11E+06
56.6	seq-rs6775	16.83	*******	4.7	471303	1.08E+06
56.7	seq-rs6773	16.83	*******	4.7	471303	1.08E+06
56.9	seq-rs6763	16.83	*******	4.7	471303	1.08E+06
57	seq-rs6762	16.83	*******	4.7	471303	1.08E+06
57.2	seq-rs4054	14.611	******	4.2	500232	1.07E+06
57.4	seq-rs5736	14.5	******	4.1	499451	1.06E+06
58	seq-rs9043	4.183	-	1.2	710869	864551
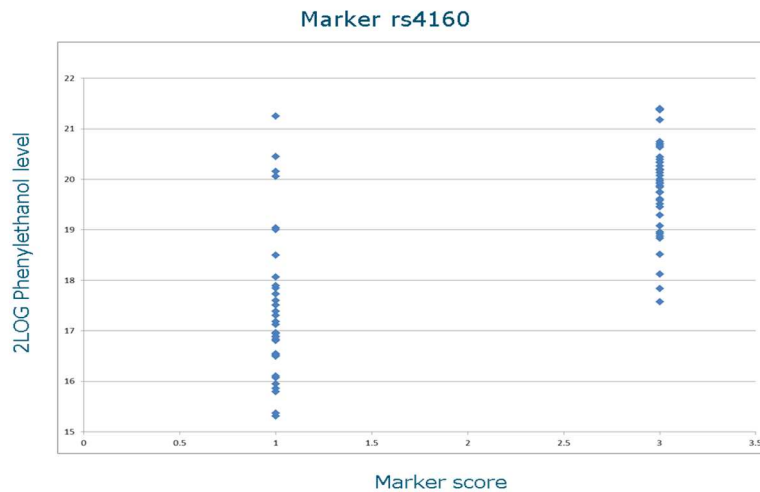
Figure 3. Scatter plot of RS4160marker scores versus 2-phenylethanol level for the two parental alleles.

Fine mapping of the phenylethanol QTL on CH4

In order to further fine-map the 2-phenylethanol QTL interval reached in the first round of fine mapping, we screened the entire F2 population of 5000 genotypes with two markers at 54.34 and 54.94 Mb. This screen resulted in 236 recombinants in this region. These recombinants were screened with an additional set of 39 markers covering this region, in order to determine the position of the recombinations. Based on the distribution of recombinations, 65 F3 families were selected and screened for homozygous recombinants. This led to a selection of 75 F3 plants for phenotypic evaluation of the fruits by GC-MS (Figure S 5). Plants were grown to maturity, fruits were harvested at fully ripe stage and analysed for variation in flavour VOCs. The levels of phenylethanol varied strongly over the samples (Figure S 6). The second round of QTL analysis revealed a strong QTL for 2-phenylethanol with a maximum ^{2}LogP score at 54.64 Mb, spanning an interval of 310 kb, based on 2 LogP units on both sides of the QTL maximum (Table 3, Figure S 7). The presence of the homozygous C085 allele led to an average 10-fold increase in the level of 2-phenylethanol compared to the homozygous R104 allele. Similar QTL results were obtained for the other phenolic VOCs phenylacetaldehyde and 2-nitrophenylethane, which belong to the same metabolic pathway (data not shown).

Table 3. Results of the second QTL analysis for 2-phenylethanol

Position (MB)	Locus	K*	LogP	Mean-a	Mean-b
54.3	ultra04-0001	10.7	2.6	825283	333683
54.31	ultra04-0004	10.4	3.2	825283	333683
54.31	ultra04-0006	10.4	3.2	825283	333683
54.32	ultra04-0009	10.4	3.2	825283	333683
54.34	seq-rs8225	10.4	3.2	825283	333683
54.34	ultra04-0010	10.4	3.2	825283	333683
54.35	ultra04-0011	6.3	2.2	765235	377929
54.36	ultra04-0013	3.0	1	680804	450697
54.37	ultra04-0014	1.9	0.7	595289	507707
54.39	ultra04-0016	0.5	0.4	515946	561777
54.40	ultra04-0019	0.3	0.4	441051	626900
54.41	ultra04-0021	1.6	0.6	397944	682723
54.45	ultra04-0024	1.6	0.6	397944	682723
54.47	ultra04-0027	3.4	1	337004	703436
54.49	ultra04-0030	6.3	1.7	314121	733059
54.50	ultra04-0031	11.5	2.8	236630	813851
54.52	ultra04-0032	19.2	4.5	164601	869678
54.54	ultra04-0034	28.9	6.6	88406.2	934328
54.58	ultra04-0039	32.0	7.2	84002.3	899134
54.60	ultra04-0041	34.4	7.8	93784.6	939643
54.61	ultra04-0042	34.4	7.8	93784.6	939643
54.63	ultra04-0044	36.9	8.3	89053.4	943944
54.64	ultra04-0045	38.3	8.6	89053.4	916317
54.65	ultra04-0046	38.0	8.5	89250.5	916148
54.72	ultra04-0054	38.0	8.5	89250.5	916148
54.75	ultra04-0056	38.0	8.5	89250.5	916148
54.75	ultra04-0057	38.0	8.5	89250.5	916148
54.75	ultra04-0058	37.2	8.4	90276.7	916148
54.75	ultra04-0059	37.2	8.4	90276.7	916148
54.78	ultra04-0061	37.2	8.4	90276.7	916148
54.79	ultra04-0062	37.2	8.4	90276.7	916148
54.81	ultra04-0064	35.2	7.9	100734	950878
54.81	ultra04-0065	33.0	7.5	98592.5	881639
54.85	ultra04-0067	18.9	4.4	256734	800735
54.86	ultra04-0068	18.9	4.4	256734	800735
54.86	ultra04-0070	18.9	4.4	256734	800735
54.87	ultra04-0071	19.5	4.5	250265	843687
54.89	ultra04-0074	15.7	3.7	289637	839670
54.94	seq-rs7034	13.1	3.2	284149	788667

Determining the QTL effects in selected F3 families

In order to confirm the previous results regarding the obtained QTL for phenolic VOCs, six F3 families segregating for specific regions in the QTL interval were selected based on the previous results . F3 plants representing either the homozygous A or B allele were selected using four markers spanning the QTL region and 10 plants of each allelic class were grown to maturity. From each individual plant ripe fruits were harvested and analysed for their content of VOCs. A T-test was used to determine if there was a significant difference between the presence of the A or B allele in each F3 family (Figure 4). As shown in panel A, three out of six F3 families showed a significant difference between the A or B allele in the specific region while three did not. Of these three, two families (1605 and 4146) showed high 2-phenylethanol levels in all lines, indicating that the 2-phenylethanol QTL is outside the segregating region in these families and represented by a homozygous B (cherry) allele. The third F3 family, 3359, had intermediate phenylethanol levels in all lines and defined the right border of the QTL at 54.63Mb. The left border of the QTL was defined by family 2973 at 54.52 Mb. Results for phenylacetaldehyde (panel B) and 2-nitrophenylethane (panel C) are fully in line with those of 2-phenylethanol. Combining these results narrows down the QTL region for the three phenolic VOCs to a region of 110 kb, from 54.52 to 54.63 Mb (marked in red in Figure 4).

Figure 4. Analysis of phenolic VOCs in six segregating F3 families. Results represent GCMS detector response. CV: coefficient of variation. A: Phenylethanol. B: phenylacetaldehyde. C: 2-nitrophenylethane. Red coloured markers are within the 110 kb QTL region.

Accumulation of 2-phenylethanol is ripening-dependent

To evaluate whether the accumulation of phenylethanol is ripening dependent, twelve selected F3 plants representing high and low phenylethanol accumulators were propagated in triplicate through cuttings. At two time points, fruits of three ripening stages were harvested from each individual plant. GC-MS analysis revealed that both low and high phenylethanol accumulators have similar very low phenylethanol levels in mature green fruits. During ripening, however, the increase in phenylethanol is much more pronounced in the high phenylethanol accumulators compared to the low accumulators (Table 4).

Table 4. Phenylethanol levels in cuttings of F3 plants in three ripening stages. Data are semi-quantitative and represent "detector response" of the GC-MS.

	Genotype	Date	green	turning	ripe
High	4006	1	4,622	292,377	644,921
	4006	2	11,994	131,283	396,243
	2358-1	1	3,747	665,617	1,694,926
	2358-1	2	3,401	525,940	1,256,691
	4146	1	11,029	103,005	1,048,340
	4146	2	11,865	194,578	915,344
	1501-4	1		120,694	1,062,960
	1501-4	2	21,969	239,042	1,233,960
	3	1	5,045	167,680	683,062
	3	2	14,691	685,144	1,018,140
Low	2550	1	4,197	10,638	30,455
	2550	2	5,630	10,229	63,278
	2451	1	6,902		
	2451	2	2,253	5,727	36,777
	524	1	2,730	13,137	12,334
	524	2	5,305	10,923	68,125
	610	1	5,440	31,683	20,313
	610	2	3,712	39,928	
	1452	1	69,123	17,391	15,858
	1452	2	23,528	50,366	32,281
	4776	1	8,101	36,444	
	4776	2	9,477	26,401	64,754
	3328-12	1	3,757	10,187	31,397
	3328-12	2	4,329	19,302	15,439

Sensory and metabolic analysis of contrasting NIL's

Based on segregation of a RIL heterozygous in the chromosome 4 QTL region, NIL pairs contrasting for the phenolic volatile QTL region were developed, in order to study whether the difference in content of phenolic volatiles in fruits of these NILs would have a sensorial effect. Plants were grown in the greenhouse and fruits were harvested at ripe stage for analysis by a trained sensory panel. NIL2, which possesses the cherry allele of parent C85, scored significantly higher compared to NIL1 with the round R104 allele for the attributes aroma presence, fruity aroma and rosehip aroma (Figure 5).

Table 5. Genotyping of the CH4 QTL regions in a pair of contrasting NILs.

Chromosome	Position	Name	NIL number: 1 PV165851	2 PV165852
4	4711066	seq-rs95	BB	BB
4	5295394	seq-rs6572	BB	BB
4	6179278	seq-rs6583	BB	BB
4	11747731	seq-rs4381	BB	BB
4	50582780	seq-rs8115	BB	BB
4	51776644	seq-rs6696	BB	BB
4	52664580	seq-rs6691	BB	BB
4	53710784	seq-rs3404	AA	BB
4	54785444	seq-rs4160	AA	BB
4	55209992	seq-rs7040	AA	BB
4	55936272	seq-rs6783	AA	BB
4	56475308	seq-rs6777	AA	BB
4	56475724	seq-rs6776	AA	BB
4	56748992	seq-rs8169	AA	BB
4	57161780	seq-rs4054	BB	BB
4	58174884	seq-rs9039	BB	BB
4	58340636	seq-rs9033	BB	BB
4	59390024	seq-rs7153	BB	BB
4	59883176	seq-rs4651	AA	AA
4	60168744	seq-rs7168	AA	AA
4	61089840	seq-rs7185	AA	AA
4	62821436	seq-rs1000	AA	AA
4	62997364	seq-rs7224	AA	AA

AA: alleles from parent 104; BB: alleles from parent C85.
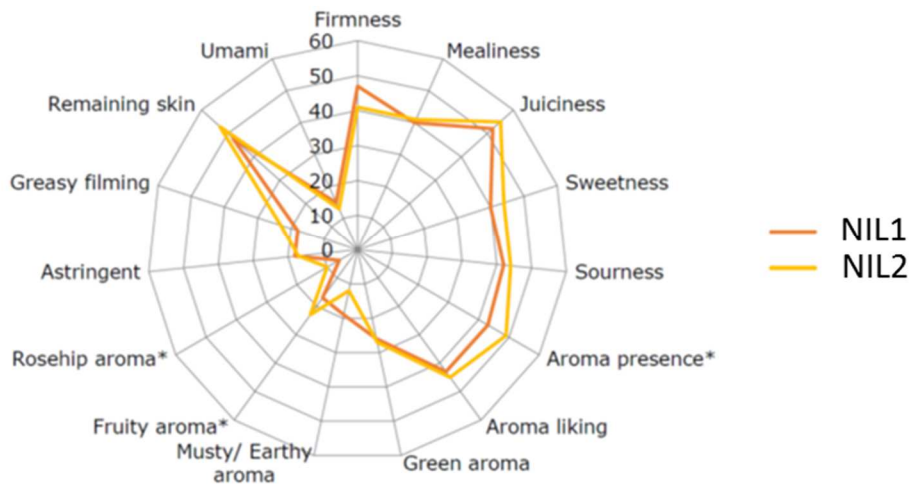
Figure 5. Spider diagram representing the attributes scored for NILs 1 and 2.   *: Significant at P<0.05.

Metabolic analysis of volatiles in the two NILs revealed that the phenolic-derived VOCs 2-phenylethanol, phenylacetaldehyde and nitrophenylethane had strongly increased levels in NIL2 compared to NIL1 (Figure 6 A). In addition, NIL2 also showed elevated levels of the amino acid-derived VOCs 2-methylbutanol and 3-methylbutanol (Figure 6 B).
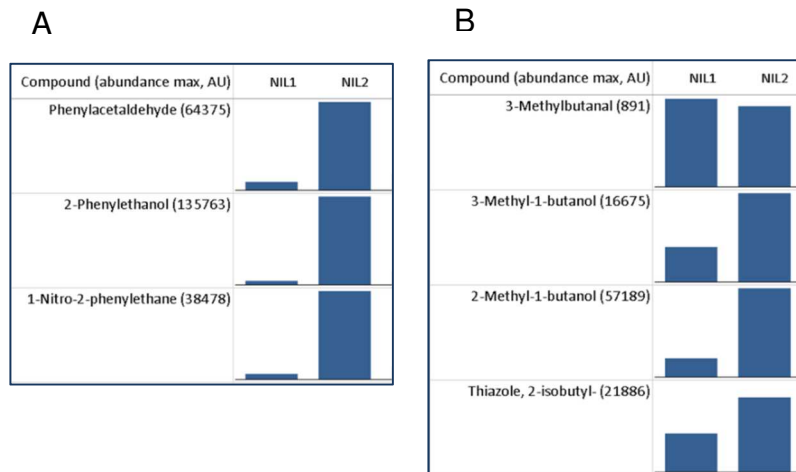
Figure 6. Relative abundance of flavour-related VOCs in the ripe fruits of the four NILs tested. A: phenolic-derived VOCs. B: amino acid-derived VOCs. The maximum abundance value (in brakets on the right of the VOC name) is set at 100% and other value are relative to it.

Of all flavour-active VOCs detected, the three phenolic VOCs showed the largest quantitative difference between NIL 2 and 1 (Figure 7), suggesting they might be the primary cause of the sensorial differences between these two NILs shown in Figure 5.
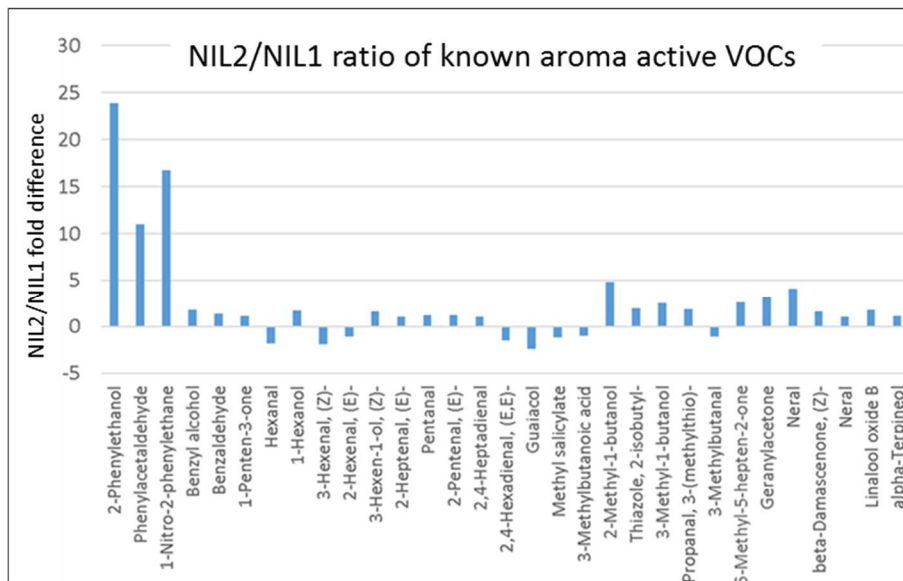
Figure 7. Fold difference in abundance of known tomato aroma active compounds in NIL2 versus NIL1.

Candidate gene identification and characterisation

The QTL region contains 11 genes, of which 3 are annotated as "unknown" (Table 6). Among them, Solyc04g063350 is annotated as a putative3-methyl-2-oxobutanoate dehydrogenase, a member of the family of branched chain amino acid dehydrogenase enzymes. These enzymes form a complex consisting

of three different subunits (E1, E2 and E3). Solyc04g063350 (further denoted as *Dcx1*) shows homology to the E1 subunit, which has decarboxylase activity. The first step in the phenylethanol biosynthetic pathway is the decarboxylation of phenylalanine, leading to the production of phenylethylamine (Tieman et al., 2006, Figure 8). Although this reaction can be carried out by several members of the aromatic amino acid decarboxylase (AADC) family, there are no AADC homologs in the phenylethanol QTL region analysed in this study. Expression of the *Dcx1* gene during fruit ripening was determined in RNA isolated from parental lines C85 (high phenylethanol) and R104 (low phenylethanol). As shown in Figure 9, *Dcx1* expression is low in line R104 and decreases upon ripening, while the expression is 2-fold higher at green stage in C85 and increases upon ripening with a peak at B/T stage. This indicates that the differences in phenylethanol accumulation may at least in part be due to differences in the expression of the *Dcx1* gene.

Table 6. Genes in the mapped interval.

| Name	Chromosome	Start	Stop	Annotation
Solyc04g063280.2	SL2.40ch04	54521549	54535248	Vacuolar protein sorting-associated protein 27 (AHRD V1 *---
Solyc04g063290.2	SL2.40ch04	54541016	54543059	30S ribosomal protein S5 (AHRD V1 ***- C5K1T7_AJEDS); contains Interpro
Solyc04g063300.1	SL2.40ch04	54545986	54548082	Pentatricopeptide repeat-containing protein (AHRD V1 ***- D7LR64_ARALY);
Solyc04g063310.1	SL2.40ch04	54548467	54550041	Unknown Protein (AHRD V1)
Solyc04g063320.1	SL2.40ch04	54558989	54559183	Unknown Protein (AHRD V1)
Solyc04g063330.1	SL2.40ch04	54560901	54561098	Unknown Protein (AHRD V1)
Solyc04g063340.1	SL2.40ch04	54565840	54567843	Pentatricopeptide repeat-containing protein IPR002885 Pentatricopeptide repeat
Solyc04g063350.2	SL2.40ch04	54581013	54588153	3-methyl-2-oxobutanoate dehydrogenase (2-methylpropanoyl-transferring)
Solyc04g063360.2	SL2.40ch04	54590862	54606336	Cell division protease ftsH homolog (AHRD V1 *--- FTSH_CYACA); contains
Solyc04g063370.1	SL2.40ch04	54608074	54614389	Receptor-like kinase (AHRD V1 *-*- A7VM65_9VIRI); contains Interpro
Solyc04g063380.2	SL2.40ch04	54633946	54637333	F-box/LRR-repeat protein 2 (AHRD V1 ***- B6TAC3_MAIZE); contains Interpro
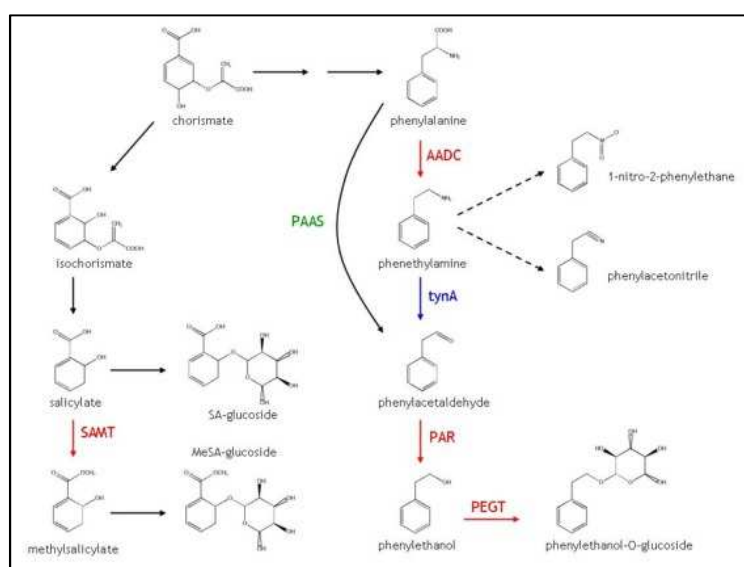
Figure 8. Proposed pathway for production of the phenolic volatiles 2-phenylethanol, phenylacetaldehyde and 2-nitrophenylethane (htpp://hos.ufl.edu/kleeweb/flavorresearch.html).
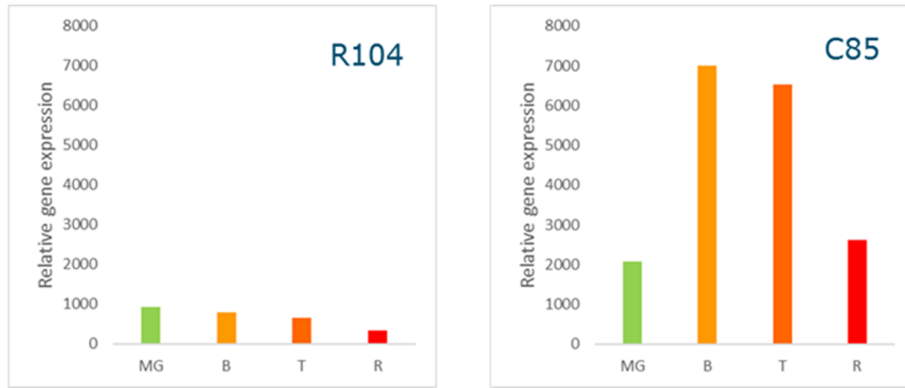
Figure 9. Expression levels of the candidate gene (*Dcx1*) in the QTL region in fruits of the parental lines R104 and C85, harvested at different ripening stages. MG: mature green. B: breaker. T: turning. R: ripe

Expression of the candidate gene in the contrasting F3 lines

The expression of *Dcx1* was also measured in 8 F3 lines which accumulate low (3328-12, 4776, 0610, 0524, 2451) or high levels of phenylethanol (003, 1501, 4146, 2358-1, 4006), based on previous results of this study. The result of real time PCR showed that the expression of the *Dcx1* gene in the fruits of high phenylethanol accumulator lines is higher than in the low accumulators (Figure 10).
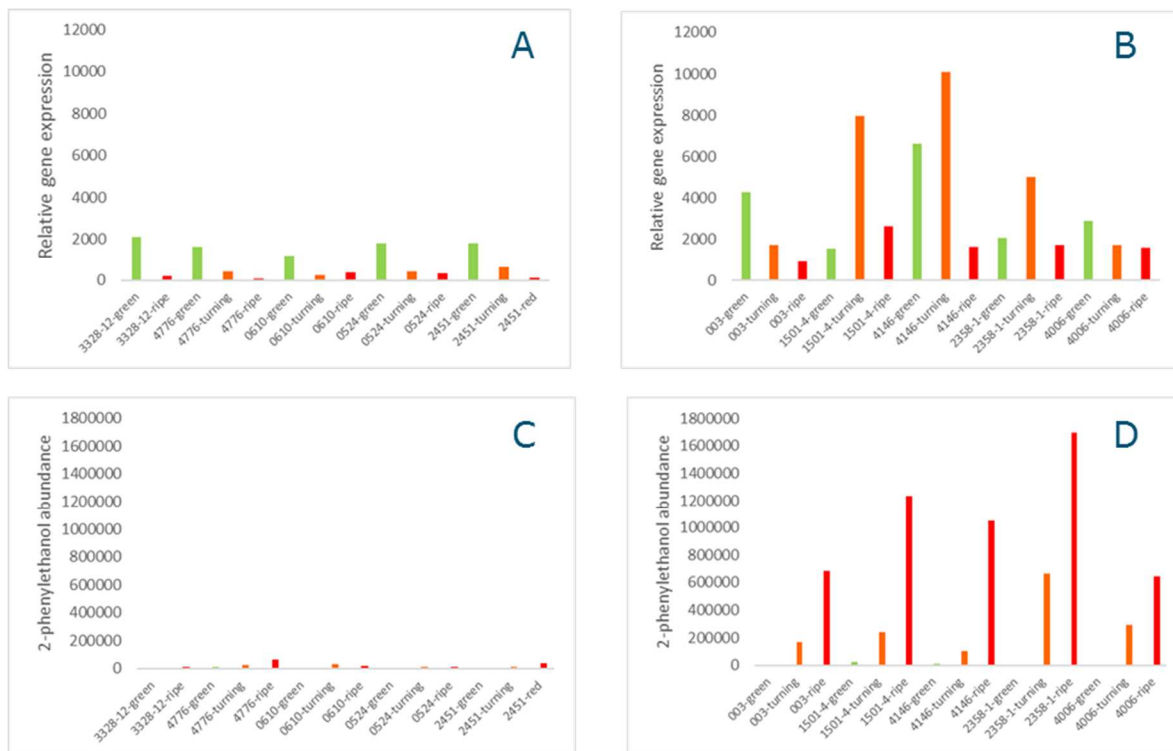
Figure 10. Result of Real time QPCR for the expression of *Dcx1* in 8 F3 lines. Gene expression analysis of *Dcx1* and 2-phenylethanol abundance in ripening fruits (G, T, R) of five F3 genotypes 'low in 2-phenylethanol' and five F3 genotypes 'high in 2-phenylethanol'. Relative gene expression of *Dcx1* in low (panel A) and high (panel B) 2-phenylethanol tomato fruits. Relative abundance of 2-phenylethanol was measured in the same fruit samples, low phenylethanol genotypes (panel C) and high phenylethanol genotypes (Panel D).

136

VIGS analysis of *Dcx1* in tomato fruit

To test the function of the *Dcx1* gene a VIGS (Virus Induced Gene Silencing) experiment was carried out by injection of a VIGS construct in detached immature green fruits of tomato cultivar Solarino, using the Agrobacterium-based pTRV2 system. Solarino has the positive *Dcx1* allele and a high *Dcx1* expression (results not shown). Metabolic analysis of red ripe fruits revealed a severe reduction in the levels of all three phenolic-derived volatiles phenylethanol, phenylacetaldehyde and nitrophenylethane (Figure 11). Initially, the annotation of this gene was predicted to act on another class of amino acids namely leucine and isoleucine. We also observed a severe reduction in leucine and isoleucine-derived volatiles, such as 2,3 methylbutanol/al and isobutylthiazole in the VIGS plants (Figure 12). There was no effect of the VIGS treatment on any other VOCs (data not shown).
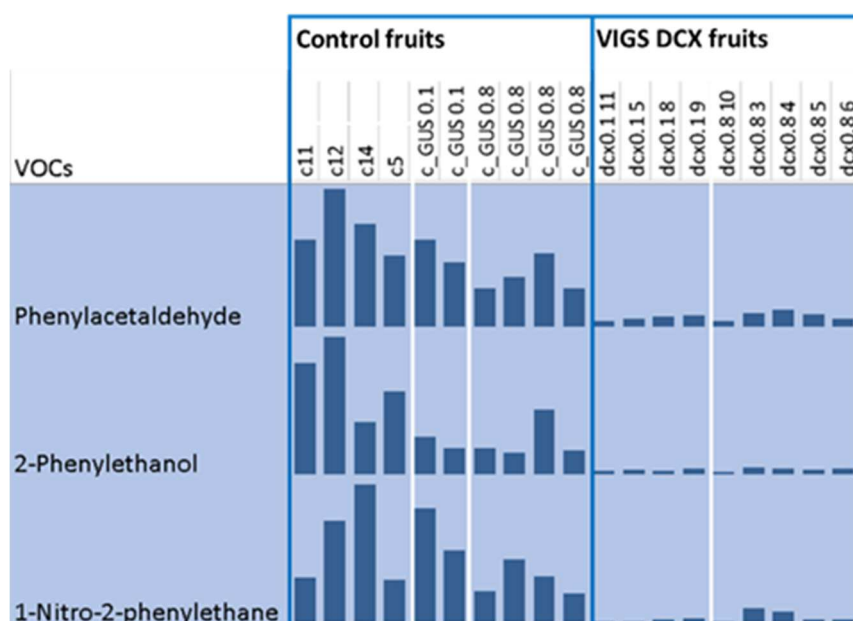
Figure 11. Metabolic analysis of phenolic VOCs in VIGS silenced *Dcx1* fruits. Each column represents a single fruit.
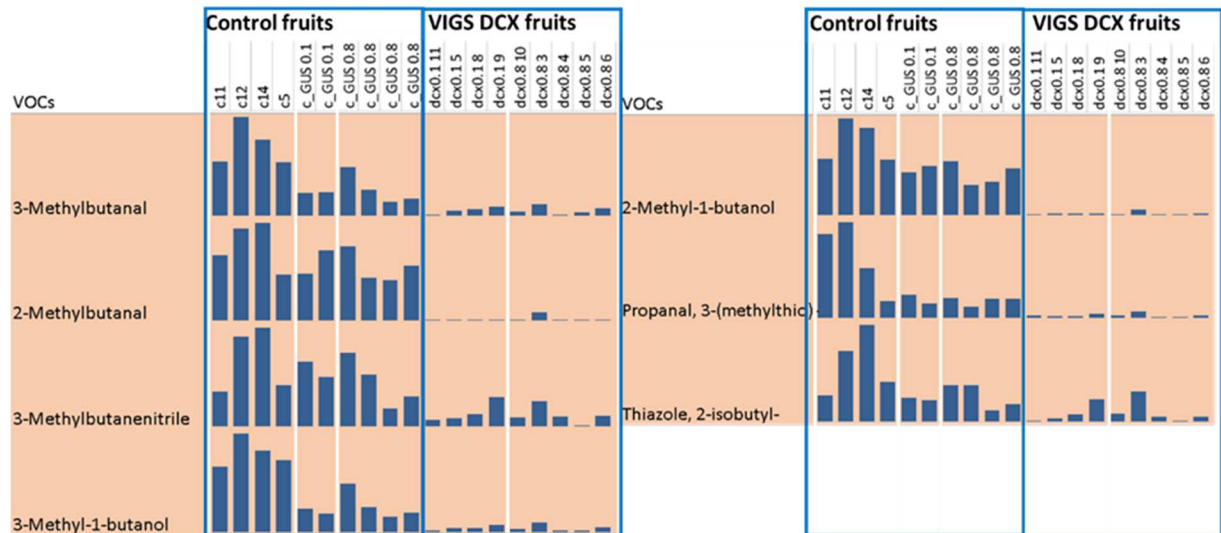
Figure 12. Metabolic analysis of branched-chain amino acid-derived VOCs in VIGS silenced Dcx1 fruits. Each column represents a single fruit.

Discussion

2-phenylethanol is one of the most abundant volatile compounds often present in tomato fruits and its positive effect on tomato flavour has been described in several studies. We implemented bi-parental qtl mapping in order to fine map a major QTL for 2-phenylethanol on the distal part of chromosome 4 and were able to narrow down this QTL to a region of 110 kb. We identified a candidate key gene for the variation of this volatile in the mapped interval.

The biosynthetic pathway of 2-phenyethanol has previously been studied in several plant species in which this volatile was shown to have a significant effect on their aromas, such as rose (Sakai et al., 2007; Hirata et al., 2012, 2016; Sheng et al., 2018; Roccia et al., 2019), petunia (Kaminaga et al., 2006; Oliva et al., 2017), melon (Gonda et al., 2010, 2018) and tomato (Tieman et al., 2006, 2007). In none of these plants the regulatory mechanisms underlying 2-phenylethanol metabolism have been fully elucidated. 2-phenylethanol is produced from the amino acid phenylalanine in two competing pathways. In the first pathway phenyl acetaldehyde is an intermediate and in the second pathway phenylpyruvate. In tomato, 2-phenylethanol is produced from phenylacetaldehyde. The pathway starts with decarboxylation of phenylalanine to phenethylamine, by a set of related aromatic L-amino acid decarboxylases (AADCs) such as AADC1A, AADC1B, and AADC2 (Tieman et al., 2006, Figure 8). In the second step phenethylamine is presumably de-aminated by an, yet uncharacterised, amine oxidase to

produce phenylacetaldehyde. Alternatively, phenethylamine could be converted to 1-nitro-2-phenylethane or 2-phenylacetonitrile by means of other yet unknown enzymes. The final step in this pathway is the conversion of 2-phenylacetaldehyde to 2-phenylethanol by two phenylacetaldehyde reductases PAR1 and PAR2 (Tieman et al., 2007). A similar pathway has been suggested in rose (Sakai et al., 2007). Formation of phenylacetaldehyde from phenylalanine can also be carried out by means of a bifunctional phenylacetaldehyde synthase enzyme catalysing both decarboxylation of phenylalanine and the posterior amine oxidation, as has been reported in rose and petunia (Kaminaga et al., 2006; Roccia et al., 2019). The second pathway for the formation of 2-phenylethanol from phenylalanine starts with transamination of phenylalanine to phenylpyruvate, followed by its decarboxylation. This pathway has first been found in microorganisms and yeast and been validated in plants such as melon (Gonda et al., 2010, 2018), petunia (Oliva et al., 2017) and rose (Hirata et al., 2012, 2016; Sheng et al., 2018). In rose and petunia it was speculated that several pathways are involved in 2-phenylethanol synthesis (Oliva et al., 2017; Sheng et al., 2018).

The existence of a phenylethanol QTL on Chromosome 4 has previously been reported in other studies (Bauchet et al., 2017; Saliba-Colombani et al., 2001; Zanor et al., 2009). The recent study from Bauchet et al. (2017) identified a strong QTL for the phenolic flavour volatiles phenylethanol, phenylacetaldehyde and nitrophenylethane in a GWAS panel including wild tomato relatives, old varieties and modern cultivars. The QTL interval spanned a 115 kb region at the bottom of chromosome 4 from 54.75 to 54.86 where the major candidate gene for the variation of phenylacetaldehyde and phenylethanol content was proposed to be Solyc04g064490.2.1, a glucosyltransferase of the GT8 glycosyltransferase family. Our fine-mapped interval did not overlap with this region, but was located about 230 kb upstream, from 54.52 to 54.63 Mb. We identified Solyc04g063350 (*Dcx1)* which is annotated as 3-methyl-2-oxobutanoate dehydrogenase as the key gene responsible for the variation of 2-phenylethanol in our mapped interval. *Dcx1* is a member of the family of 2-oxoacid dehydrogenases which consist of three related multi-enzyme complexes: the 2-oxoglutarate dehydrogenase, pyruvate dehydrogenase and branched-chain 2-oxoacid dehydrogenase complexes which catalyse the oxidative decarboxylation of their respective substrates (Dickinson & Dawes, 1992; Yeaman, 1986, 1989). Since the first step in the phenylethanol biosynthesis pathway is a decarboxylation reaction (Tieman et al. 2006; Figure 8) we hypothesised that the *Dcx1* gene might carry out the decarboxylation of phenylalanine to phenylethylamine. Co-segregation of the nitrogen-containing compound 2-nitrophenylethane with phenylethanol and phenylacetaldehyde in our study and the severe reduction in the levels of all these three phenolic VOCs in VIGS-silenced *Dcx1* fruits (Figure 11), can be considered as a validation of our hypothesis that this gene acts at a step prior to deamination. In addition to reduction of phenolic VOCs, a severe

reduction in leucine and isoleucine-derived volatiles, such as 2,3 methylbutanol/al and isobutylthiazole was also observed in the VIGs plants (Figure 12). This is not surprising since it has been demonstrated that branched-chain 2-oxoacid dehydrogenases catalyse an irreversible step in the catabolism of the branched-chain amino acids leucine, isoleucine and valine. Catabolism of branched-chain amino acids starts with an aminotransferase reaction to yield 2-oxoacids. These 2-oxoacids then undergo an oxidative decarboxylation by a branched-chain 2-oxoacid dehydrogenase (Jones & Yeaman, 1986; Yeaman, 1986; van der Hijden & Bom, 1996). Our results suggest that the substrate specificity of the Dcx1 enzyme is not restricted to the amino acid phenylalanine, but may also act on branched chain amino acid-derived keto acids. This hypothesis, however, needs further investigation. We did not observe any effect of the VIGS treatment on other VOCs (data not shown). Other reverse genetics approaches are currently in progress to unequivocally demonstrate the function of the *Dcx1* gene.

Conclusions

Using a F2 population derived from two contrasting RILs we were able to fine map a QTL for 2-phenylethanol on the distal part of chromosome 4 and identify a putative key gene in its biosynthetic pathway. These results help to better understand the biosynthesis pathway underlying 2-phenylethanol and the other two phenolic VOCs 2-phenylacetaldehyde and 1-nitro-2-phenylethane which have a positive effect on tomato aroma. VIGS analysis of the identified *Dx1* gene revealed that even the manipulation of one gene in the biosynthetic pathway of 2-phenylethanol can dramatically change the level of this volatile in the fruits, although additional reverse genetics approaches are needed to obtain final proof for the exact role of this gene in metabolic pathways affecting tomato aroma. Nevertheless, molecular markers based on this gene can already be used to improve tomato aroma by marker-assisted selection.

Acknowledgement
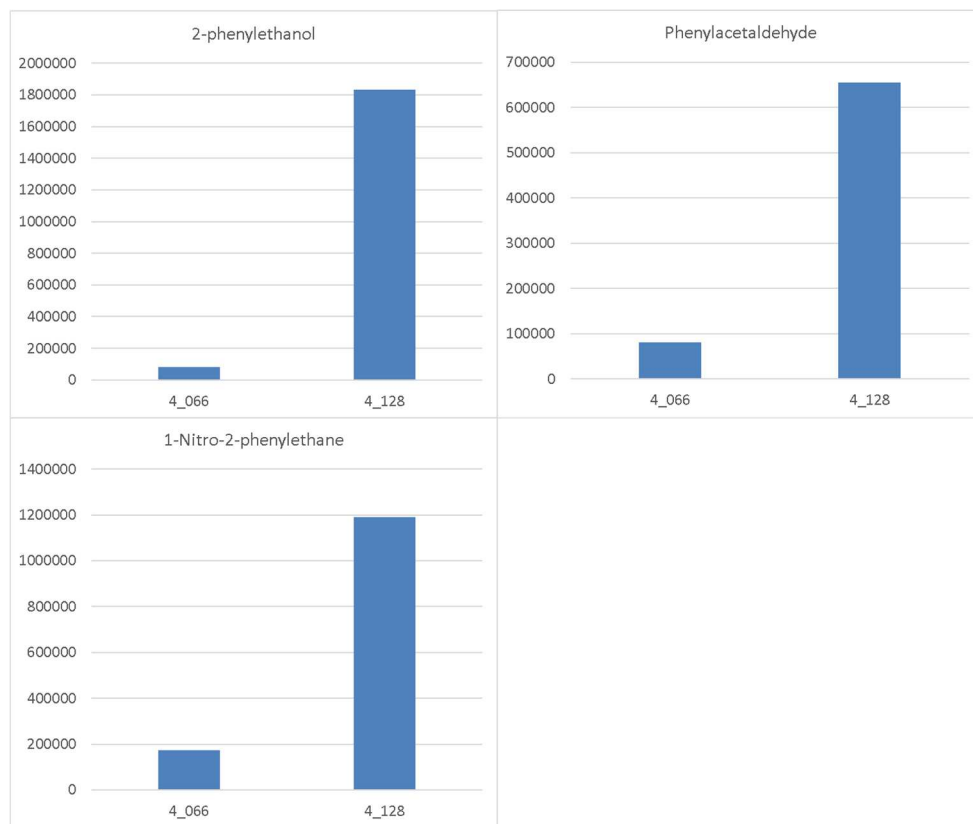
Supplementary Data

Figure S1. Accumulation of phenolic volatiles in fruits of the F6 parents of the fine mapping population. The Y-axis shows thecompound abundance measured by GC-MS and expressed as arbitrary units of the mass spectrometer response.
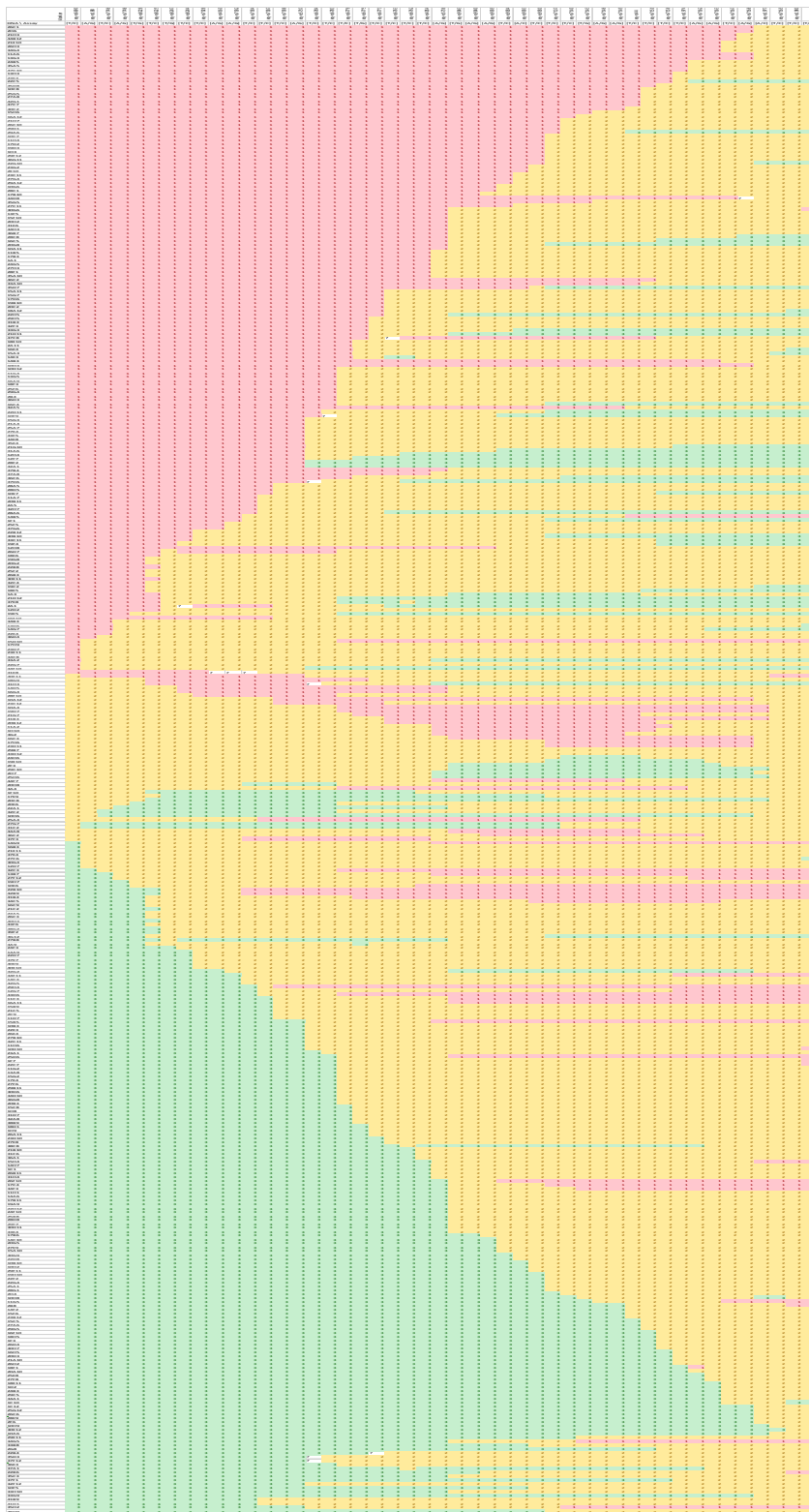
Figure S 2. 384 F2 recombinants in the segregating region of chromosome 4. The red colour (A score) represents the homozygous allele of the R104 parent and the green colour (B allele) represents the homozygous alleles of parent C085. The yellow colour represents heterozygous regions.

physical position	53029140	55883979	58047432	60699722	
DNA \ Assay	seq-rs6690	seq-rs6784	seq-rs9043	seq-rs1980	
10B10-0848-F2	3	3	3	2	
10B10-0848-1	3	3	3	2	
10B10-0848-2	3	3	3	2	
10B10-0848-3	3	3	3	1	X
10B10-0848-4	3	3	3	1	X
10B10-0848-5	3	3	3	2	
10B10-0848-6	3	3	3	2	
10B10-0848-7	3	3	3	1	
10B10-0848-8	3	3	3	3	
10B10-0848-9	3	3	3	1	
10B10-0848-10	3	3	3	1	
10B10-0848-11	3	3	3	2	
10B10-0848-12	3	3	3	2	

Figure S 3. Genotypic analysis of F3 seedlings. Homozygous recombinants, such as those marked "X" were selected for phenotypic analysis.

Table S 1. Result of SPME-GC-MS for 13 important flavour-related VOCs.

Sample	2-Phenylethanol	Phenylacetaldehyde	2-Phenylnitroethane	2-isobutylthiazole	Benzyl_alcohol	Benzaldehyde	Guaiacol	Methyl_salicylate	3-Methyl-1-butanol	2-Methyl-1-butanol	3Methylbutanoic_acid	cis-3-Hexenal	Hexanal
11H1-1001-8	2.71E+06	921870	371807	24574	62207	86749	63795	12460	24105	139819	26732	1308470	2499080
10E6-0878-2	1.43E+06	504988	336402	49742	22304	61686	24243	5209	14563	34835	19311	837371	1108180
20D3-1785-8	497605	527708	2.27E+06	92633	38800	66371	25022	2815	2611	9912	6319	372664	2259710
34D8-3074-7	59462	28393	4770	5023	56300	57241	61960	7463	29606	57022	62165	1909420	2440200
10C8-0858-6	127797	71238	36343	27591	35108	60371	44152	2867	15745	19980	39147	1611570	2540960
15A3-1291-11	63139	34291	18445	173919	54515	81044	9632	2715	116449	36199	88687	1741350	1659550
27H3-2475-12	95477	75935	58291	70871	38980	77622	23904	2480	25121	40620	49604	999806	1709830
10B10-0848-8	2.76E+06	1.10E+06	556262	35772	52377	2432	26951	4798	18962	61987	121075	1445470	2608310
3A9-0193-5	120674	85118	67626	54923	41617	50777	20633	2590	8975	8087	21790	1781490	2721560
22C5-1959-1	149029	75089	11951	50062	63915	83222	40942	6716	18857	41893	20027	1313940	1903310
38C8-3430-10	195375	81275	27941	390846	160313	140099	8861	2591	79427	44891	58805	938445	2181980
11H4-1004-1	116441	69253	31508	96855	21027	52311	113533	7029	28982	39829	73012	707669	1703860
34F3-3091-8	990362	647020	849940	164702	72067	92398	18115	2679	19747	40901	94171	1361540	2212980
10C8-0858-3	70633	56541	37457	36751	80295	133765	28882	2523	36363	22815	70373	860396	1354860
22B6-1948-1	1.10E+06	736189	1.49E+06	299780	40997	63354	9727	2708	23266	23807	38389	843852	2453430

42B9-3787-1	2.16E+06	787819	461511	44746	163397	153143	14546	3824	33085	153717	35601	1989460	2588360
42A5-3773-1	2.49E+06	1.46E+06	2.76E+06	17798	49940	114843	22133	2828	2776	51883	11729	724941	2550360
25A7-2215-14	243692	129459	174801	243306	41846	67836	18012	2627	64963	29759	161059	1390600	2052500
26C10-2332-6	161716	68208	63231	49401	58439	72541	28939	2441	15210	32040	20003	1788890	2451310
23B12-2046-1	799243	477479	636980	187431	207133	177200	8447	4974	13678	15382	33226	679695	2526270
16A9-1389-11	143148	48837	7552	61867	388543	173000	27599	2885	71031	55300	16728	866521	2121680
13F5-1165-5	217715	95341	22675	302301	53330	54808	10591	2839	60345	31336	109943	1424140	2711930
25A7-2215-7	273962	159565	225999	227859	24577	46828	15274	4566	13277	21605	29513	703499	2461040
13A8-1112-7	641427	305117	130352	103445	27418	64530	30710	2851	26296	25028	71346	819676	1611080
10A11-0839-8	171796	111313	36255	326735	49648	100957	4512	2524	32200	18701	53423	789856	1560860
26H2-2382-1	478835	259790	113632	85455	25104	2995	18161	2557	25272	20239	80313	1397960	2093380
32C11-2881-2	468045	349560	316250	146182	61835	135288	16174	2680	32766	22538	414286	1165730	1870430
11H4-1004-4	198155	145375	226450	256440	51233	111763	10256	2520	25156	36103	47431	960066	2056510
11A7-0927-16	114610	78953	46721	252476	63488	64112	41676	9928	41908	22260	40612	1008300	2761880
2B8-0110-1	1.02E+06	446515	185670	24476	60684	70935	355790	5832	41648	158015	51756	1016170	2195770
36B7-3233-1	234865	176190	94408	228614	87012	149342	11818	2485	49275	36806	79807	1334370	2761510
11H4-1004-2	537685	308282	348366	217262	18336	43714	15452	2671	18805	15032	47863	613412	2489440
13A8-1112-2	1.15E+06	471623	843370	54174	30761	63079	14831	2808	5842	23845	18673	660598	2598790
16B3-1393-3	94245	58321	12413	204961	71911	80188	24370	2812	33655	20000	60077	1827790	2761620
23B12-2046-13	747558	476502	1.20E+06	190341	90848	81704	7553	4934	11362	18895	33154	875888	2761460
15F10-1354-2	92517	57448	9906	8868	17750	36794	76450	4735	34563	30327	182056	1154020	1451370
15F6-1350-1	880829	660704	623003	233786	74050	89187	25446	2960	44167	96125	211074	1196980	1873090
10E6-0878-11	186014	181417	157569	116452	59287	119115	24014	5006	17754	17970	69754	1463410	2378540
21C2-1864-9	507912	346294	167936	223043	97258	2663	53828	3487	66527	92476	231997	1262540	1659020
27E9-2445-5	1.19E+06	635701	1.11E+06	123315	101799	132891	14600	2627	32658	59783	209025	1014690	2761430
26H2-2382-3	285110	178264	138719	266118	41528	73598	18736	2816	42442	40917	73804	712428	2388070
11H1-1001-9	2.37E+06	1.65E+06	2.75E+06	137948	63734	100179	9080	2438	14807	40225	37841	852429	2293010
30G4-2736-2	1.09E+06	696275	1.93E+06	200366	29206	40097	7728	2738	6476	9012	26012	1179150	2739860
29F7-2635-9	552460	283521	822513	219950	28277	48991	22530	4089	56845	43788	122068	1093560	2527490
29B11-2593-4	997803	438583	293532	85291	39214	55586	5537	2412	14247	15540	71150	1234200	1408130
10B10-0848-3	1.37E+06	574780	324073	51381	26354	48981	24494	6437	46065	84167	188534	1122290	1866220
20B11-1769-16	1.17E+06	483051	312741	41449	38351	49772	20098	7052	7205	27540	33285	1187080	1840190
10B10-0848-4	1.71E+06	597717	424334	90079	72331	95690	16944	5297	23509	64129	92010	1636610	2761660
28G1-2553-5	233704	72381	2483	2477	64748	93908	202098	43568	10076	53965	21415	2608180	1851410
20D3-1785-10	719639	388609	1.15E+06	123180	164092	2412	12354	2870	11591	14859	23754	1089280	2379820
17B11-1493-5	1.32E+06	733487	360949	58894	38497	95687	26116	3830	7407	62797	12839	1329440	2507000

34F3-3091-4	2.74E+06	1.08E+06	818230	26061	140484	95683	52891	2432	25601	157351	31200	2423060	1859690
35H2-3206-15	1.76E+06	1.01E+06	1.70E+06	65561	36124	2820	71586	2766	27495	79537	183850	2044610	2543160
18F3-1623-8	68805	49129	10847	253280	79361	105712	7420	2605	63251	22784	119325	2531800	2022880
29F11-2639-7	940660	706143	1.24E+06	251967	101922	103041	19802	2789	65721	42681	179895	826012	1937830
17B11-1493-13	877241	405695	478133	93972	157904	139865	33286	6902	18551	44053	41878	1489720	2761620
21C2-1864-10	56953	47738	18349	33056	18283	34480	65757	2532	8956	17395	20936	1556210	2420290
27H3-2475-2	525224	540070	2.41E+06	106220	42839	81291	25444	3358	2537	10121	5494	498650	2700890
21A10-1850-4	1.19E+06	947924	2.13E+06	276973	27975	2558	16140	2487	34747	61140	144475	841463	1947470
13D6-1144-2	2.76E+06	1.09E+06	547836	33522	75232	84935	22757	5147	17709	61166	122000	1697020	2542730
16A9-1389-2	121027	94935	39829	28257	32930	37145	133157	2723	20106	10974	131724	1135450	1857080
21A10-1850-8	784809	373560	322551	56897	62704	99547	53572	2716	17329	26891	265961	1669790	1945100
29F11-2639-6	373947	229565	962603	359258	38493	49013	12949	3736	28338	13749	266772	1801080	2753460
28G1-2553-4	1.05E+06	914526	1.05E+06	90117	29870	71698	10292	2939	9432	36352	16104	729427	2184640
15F10-1354-15	40680	24162	36557	168555	27677	40508	6855	2971	15961	8676	18760	785869	1198270
18E11-1619-2	1.67E+06	813978	728527	82140	53496	67234	32076	3326	16085	72730	24869	1615920	2449410
15A10-1298-5	1.25E+06	655539	547223	144881	35588	68448	7419	2591	28192	67112	56680	1586210	2729330
32C11-2881-13	960028	561466	988190	139175	32629	62806	8552	2969	14462	14433	89287	1000450	1667720
26C10-2332-5	125883	68893	40357	154865	44156	59444	32679	2694	25052	39645	31851	2264290	2552910
30G9-2741-2	1.32E+06	788216	1.33E+06	163933	280090	152360	30702	2833	22300	87881	34962	774512	2058040
30G4-2736-6	1.63E+06	562489	183935	29540	47946	61868	44008	2688	10759	54784	49518	2111530	2570170
27E9-2445-2	1.20E+06	694362	800766	44433	49201	50482	59567	2748	15237	97797	43973	1339910	2252980
34D8-3074-5	1.43E+06	621088	1.03E+06	5701	31552	57454	21371	2603	6469	22756	14910	891694	2558720
20B11-1769-3	370862	128591	47448	67678	55301	71939	198219	3982	28351	21311	90515	1066480	1864670
18F3-1623-1	42256	32695	10927	12942	35408	44050	39845	2577	23967	13952	231054	1853810	1585220

Pos (Mb)	Marker	K value		Significance
54.200	seq-rs8228	27.852	1	*******
54.300	seq-rs8225	30.495	1	*******
54.500	seq-rs8247	33.850	1	*******
54.500	seq-rs8244	33.850	1	*******
54.800	seq-rs7028	33.850	1	*******
54.800	seq-rs4160	33.850	1	*******
54.800	seq-rs7032	33.850	1	*******
54.900	seq-rs7034	30.570	1	*******
55.000	seq-rs7037	29.807	1	*******
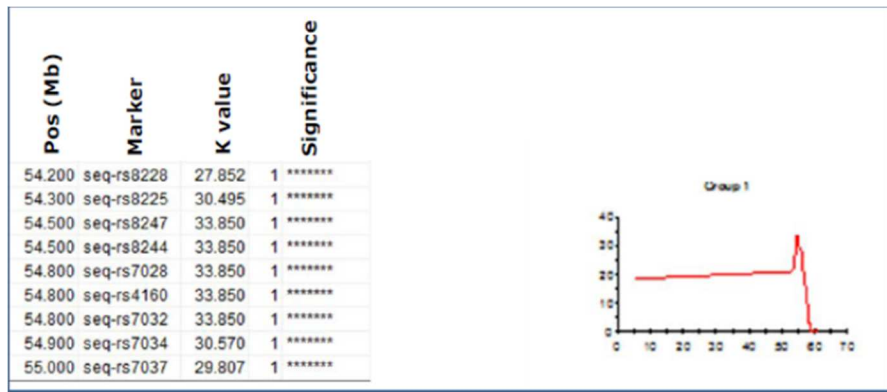
Figure S 4. Kruskal-Wallis analysis of F3 genotypes for phenylethanol. Left panel: scores. Right panel: graphical representation. *******: P<0.000001.

Figuer S 5. 75 F3 plants for phenotypic evaluation of the fruits by GC-MS in the second round of fine mapping.

Figure S 6. Metabolic analysis of F3 genotypes for variation in phenylethanol.
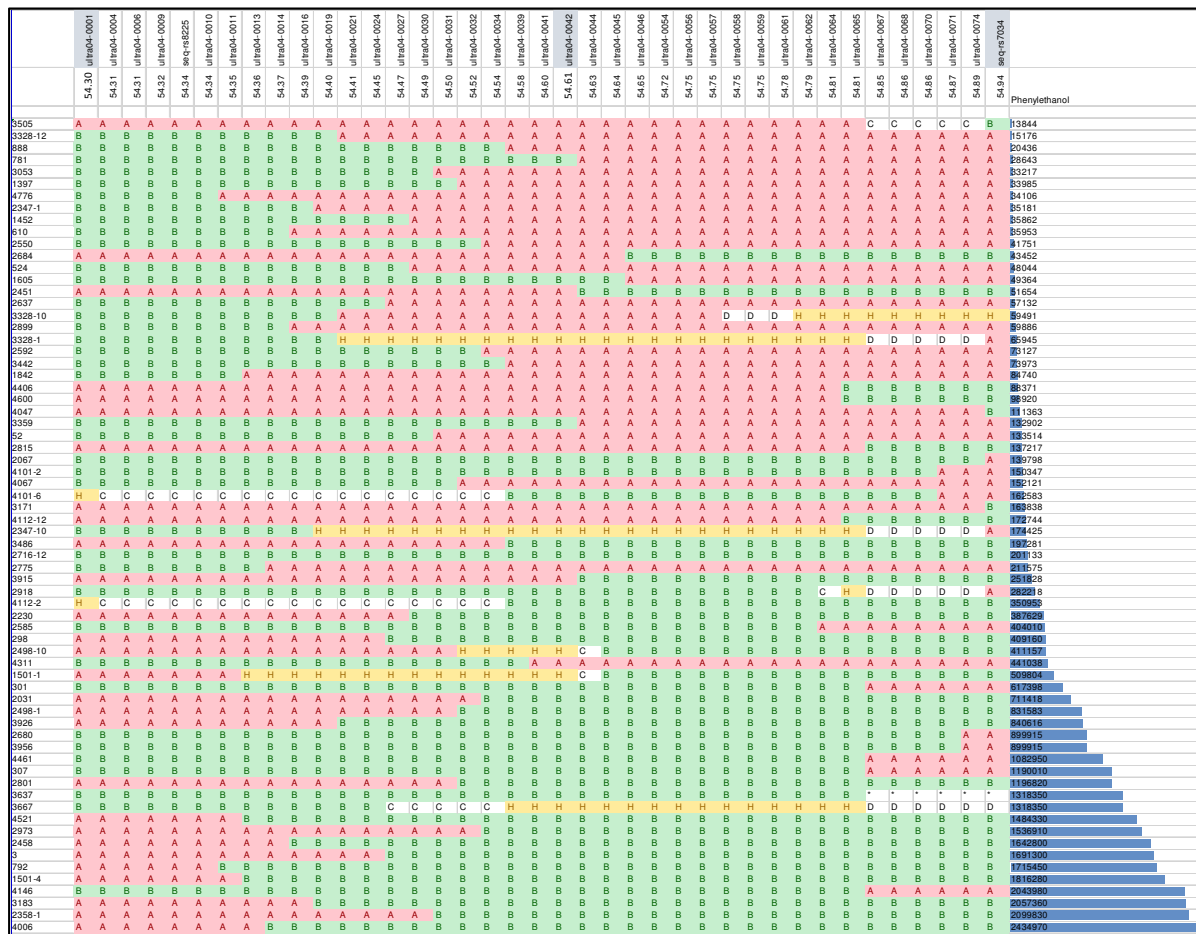
Figure S 7. Graphical representation of the F3 genotypes and their association with the VOC phenylethanol. The red colour (A score) is where the recombinants show the homozygous allele of the R104 parent and the green colour (B allele) is where the recombinants show the homozygous alleles of cherry parent C085. The yellow colour (H) represents heterozygous regions. D: A or H , C: H or B.

References

Acree, T. E. (1993). **Flavor science: Sensible principles and techniques**. *Amer. Chem.Soc., Washington, D.C.*, 1-20.

Baldwin, E., Scott, J. W., & Shewinaker, C. K. (2000). **Flavor Trivia and Tomato Aroma:Biochemistry and Possible Mechanisms for Control of Important Aroma Components.** *HORTSCIENCE, 35*(6).

Baldwin, E. A., Goodner, K., & Plotto, A. (2008). **Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors.** *J Food Sci, 73*(6), S294-307.

Bauchet, G., Grenier, S., Samson, N., Segura, V., Kende, A., Beekwilder, J., . . . Causse, M. (2017). **Identification of major loci and genomic regions controlling acid and volatile content in tomato fruit: implications for flavor improvement.** *New Phytol, 215*(2), 624-641.

Devi, K. D., Punyarani, K., Singh, N. S., & Devi, H. S. (2013). **An efficient protocol for total DNA extraction from the members of order Zingiberales- suitable for diverse PCR based downstream applications.** *SpringerPlus, 2*, 669-669.

Dickinson, J. R., & Dawes, I. W. (1992). **The catabolism of branched-chain amino acids occurs via 2-oxoacid dehydrogenase in S'charomyces cerevisiae.** *J Gen Microbiol, 138*, 2029-2033.

Gonda, I., Bar, E., Portnoy, V., Lev, S., Burger, J., Schaffer, A. A., . . . Lewinsohn, E. (2010). **Branched-chain and aromatic amino acid catabolism into aroma volatiles in Cucumis melo L. fruit.** *J Exp Bot, 61*(4), 1111-1123.

Gonda, I., Davidovich-Rikanati, R., Bar, E., Lev, S., Jhirad, P., Meshulam, Y., . . . Lewinsohn, E. (2018). **Differential metabolism of L-phenylalanine in the formation of aromatic volatiles in melon (Cucumis melo L.) fruit.** *Phytochemistry, 148*, 122-131.

Hirata, H., Ohnishi, T., Ishida, H., Tomida, K., Sakai, M., Hara, M., & Watanabe, N. (2012**). Functional characterization of aromatic amino acid aminotransferase involved in 2-phenylethanol biosynthesis in isolated rose petal protoplasts.** *J Plant Physiol, 169*(5), 444-451.

Hirata, H., Ohnishi, T., & Watanabe, N. (2016). **Biosynthesis of floral scent 2-phenylethanol in rose flowers.** *Biosci Biotechnol Biochem, 80*(10), 1865-1873.

Jones, R. A., & Scott, S. J. (1983). **Improvement of tomato flavor by genetically increasing sugar and acid contents.** *Euphytica, 32*(3), 845-855.

Jones, S. M., & Yeaman, S. J. (1986). **Oxidative decarboxylation of 4-methylthio-2-oxobutyrate by branched-chain 2-oxo acid dehydrogenase complex.** *Biochem J, 237*(2), 621-623.

Kaminaga, Y., Schnepp, J., Peel, G., Kish, C. M., Ben-Nissan, G., Weiss, D., . . . Dudareva, N. (2006). **Plant phenylacetaldehyde synthase is a bifunctional homotetrameric enzyme that catalyzes phenylalanine decarboxylation and oxidation.** *J Biol Chem, 281*(33), 23357-23366.

Klee, H. J., & Tieman, D. M. (2013). **Genetic challenges of flavor improvement in tomato.** *Trends Genet, 29*(4), 257-262.

Knudsen, J. T., Tollsten, L., & Bergström, L. G. (1993). **Floral scents—a checklist of volatile compounds isolated by head-space techniques.** *Phytochemistry, 33*(2), 253-280.

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., . . . Huang, S. (2014). **Genomic analyses provide insights into the history of tomato breeding.** *Nat Genet, 46*(11), 1220-1226.

Menéndez, P., Eilers, P., Tikunov, Y., Bovy, A., & van Eeuwijk, F. (2012). **Penalized regression techniques for modeling relationships between metabolites and tomato taste attributes.** *Euphytica, 183*(3), 379-387.

Oliva, M., Bar, E., Ovadia, R., Perl, A., Galili, G., Lewinsohn, E., & Oren-Shamir, M. (2017). **Phenylpyruvate Contributes to the Synthesis of Fragrant Benzenoid-Phenylpropanoids in Petunia x hybrida Flowers.** *Front Plant Sci, 8*, 769.

Rambla, J. L., Tikunov, Y. M., Monforte, A. J., Bovy, A. G., & Granell, A. (2014). **The expanded tomato fruit volatile landscape.** *J Exp Bot, 65*(16), 4613-4623.

Roccia, A., Hibrand-Saint Oyant, L., Cavel, E., Caissard, J. C., Machenaud, J., Thouroude, T., . . . Baudino, S. (2019). **Biosynthesis of 2-Phenylethanol in Rose Petals Is Linked to the Expression of One Allele of RhPAAS.** *Plant Physiol, 179*(3), 1064-1079.

Saliba-Colombani, V., Causse, M., Langlois, D., Philouze, J., & Buret, M. (2001). **Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits.** *Theor Appl Genet, 102*(2), 259-272.

Sakai, M., Hirata, H., Sayama, H., Sekiguchi, K., Itano, H., Asai, T., . . . Watanabe, N. (2007). **Production of 2-phenylethanol in roses as the dominant floral scent compound from L-phenylalanine by two key enzymes, a PLP-dependent decarboxylase and a phenylacetaldehyde reductase.** *Biosci Biotechnol Biochem, 71*(10), 2408-2419.

Sheng, L., Zeng, Y., Wei, T., Zhu, M., Fang, X., Yuan, X., . . . Feng, L. (2018). **Cloning and Functional Verification of Genes Related to 2-Phenylethanol Biosynthesis in Rosa rugosa.** *Genes (Basel), 9*(12).

Stevens, M. A., Kader, A. A., & Albright, M. (1979). **potential for increasing tomato flavour via increased sugar and acid contents.** *J.Amer.Soc.Hort, 104*(1), 40-42.

Tadmor, Y., Fridman, E., Gur, A., Larkov, O., Lastochkin, E., Ravid, U., . . . Lewinsohn, E. (2002). **Identification of malodorous, a wild species allele affecting tomato aroma that was aelected against during domestication.** *J Agric Food Chem, 50*(7), 2005-2009.

The Tomato Genome Consortium. (2012). **The tomato genome sequence provides insights into fleshy fruit evolution.** *Nature, 485*, 635.

Tieman, D., Taylor, M., Schauer, N., Fernie, A. R., Hanson, A. D., & Klee, H. J. (2006**). Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde.** *Proc Natl Acad Sci U S A, 103*(21), 8287-8292.

Tieman, D. M., Loucas, H. M., Kim, J. Y., Clark, D. G., & Klee, H. J. (2007). **Tomato phenylacetaldehyde reductases catalyze the last step in the synthesis of the aroma volatile 2-phenylethanol.** *Phytochemistry, 68*(21), 2660-2669.

Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., . . . Klee, H. (2017). **A chemical genetic roadmap to improved tomato flavor.** *Science, 355*(6323), 391-394.

Tikunov, Y., Lommen, A., de Vos, C. H. R., Verhoeven, H. A., Bino, R. J., Hall, R. D., & Bovy, A. G. (2005). **A Novel Approach for Nontargeted Data Analysis for Metabolomics. Large-Scale Profiling of Tomato Fruit Volatiles.** *Plant Physiol, 139*(3), 1125-1137.

Tikunov, Y. M., Molthoff, J., de Vos, R. C. H., Beekwilder, J., van Houwelingen, A., van der Hooft, J. J. J., . . . Bovy, A. G. (2013). **NON-SMOKY GLYCOSYLTRANSFERASE1 Prevents the Release of Smoky Aroma from Tomato Fruit.** *Plant Cell, 25*(8), 3067-3078.

Tzin, V., Rogachev, I., Meir, S., Moyal Ben Zvi, M., Masci, T., Vainstein, A., . . . Galili, G. (2013). **Tomato fruits expressing a bacterial feedback-insensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase of the shikimate pathway possess enhanced levels of multiple specialized metabolites and upgraded aroma.** *J Exp Bot, 64*(14), 4441-4452.

Ursem, R., Tikunov, Y., Bovy, A., van Berloo, R., & van Eeuwijk, F. (2008). **A correlation network approach to metabolic data analysis for tomato fruits.** *Euphytica, 161*(1), 181.

van der Hijden, H. T. W. M., & Bom, I. J. (1996). **ENZYMES INVOLVED IN THE METABOLIC PATHWAY LEADING TO 3-METHYLBUTANAL IN TOMATO FRUIT.** In A. J. Taylor & D. S. Mottram (Eds.), *Flavour Science* (pp. 130-133): Woodhead Publishing.

Van Ooijen, J. W. (2009). **Software for the mapping of quantitative trait loci in experimental populations of diploid species.**

Yeaman, S. J. (1986). **The mammalian 2-oxoacid dehydrogenases: a complex family.** *Trends Biochem Sci, 11*(7), 293-296.

Yeaman, S. J. (1989). **The 2-oxo acid dehydrogenase complexes: recent advances.** *Biochem. J, 257*(3), 625-632.

Yilmaz, E. (2001). **The Chemistry of Fresh Tomato Flavor.** *Turk J Agric For, 25 (2001)* 149-155.

Zanor, M. I., Rambla, J. L., Chaib, J., Steppa, A., Medina, A., Granell, A., . . . Causse, M. (2009). **Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents.** *J Exp Bot, 60*(7), 2139-2154.

Chapter 5

Fine mapping of a QTL controlling glutamate content responsible for the umami taste in tomato

Raana Roohanitaziani[1,2], Fien Meijer-Dekens [1], Richard GF Visser[1] and Arnaud G. Bovy[1]

[1] Plant Breeding, Wageningen University and Research, PO Box 386, 6700 AJ Wageningen, The Netherlands

[2] Graduate School Experimental Plant Sciences, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

Abstract

Flavour in tomato is governed by many different components of which the amino acid glutamate is of main importance since it is linked to the so called umami taste in tomato. We conducted metabolic quantitative trait loci (mQTL) analysis to localise the genomic region controlling glutamate content in tomato (*Solanum lycopersicum* L.) and to identify the encoding genes involved in the biosynthesis of this metabolite. A major QTL for glutamate content was identified on chromosome 4. With the objective of fine mapping this QTL, we developed an F2 population consisting of 5000 individual plants derived from a cross between two RILs contrasting for the glutamate QTL region on chromosome 4. F3 seed was harvested from each individual F2 plant to immortalize the fine mapping population. Using SNP markers and metabolic analysis by GC-TOF-MS, this QTL was fine-mapped to a region of approximately 300-kb from 57 to 57.32 Mb which is harbouring ~30 candidate genes. Of these 30 genes, seventeen showed a significantly different expression level in genotypes contrasting for the glutamate QTL. Identification of the responsible gene or genes needs further research.

Key words: Glutamate, Umami, mQTL, fine mapping, Tomato, *S. lycopersicum*

Introduction

Tomato flavour is the result of the interaction of taste components and aromatic volatiles. Sugars, organic acids, free amino acids and minerals are the non-volatile constituents of tomato flavour. Free amino acids form about 2-2.5% of the total dry matter of tomatoes (Kader et al., 1978; Yilmaz, 2001). Glutamic acid (GLU) is the major free amino acid found in tomato juice which comprises up to 45% of the total weight of free amino acids in fresh tomato juice (Yilmaz, 2001). This high level of free glutamate in tomato fruits provides the characteristic "umami" taste. Umami, in addition to sweet, salt, bitter and sour, is the fifth basic taste quality that humans can detect. Several small molecules eliciting umami taste have been identified such as monosodium glutamate (MSG), nucleotides (monophosphates of inosinate or guanylate, inosine 5'-monophosphate, and guanosine-5'monophosphate), free amino acids and organic acids (Kurihara, 2009; Kurihara, 2015; Wu et al., 2017). Among them, free amino acids have been found as the main umami ingredients in many vegetables and glutamate is supposed to be the main contributor of umami taste in tomato (Valle et al., 1998; Sorrequieta et al., 2010). The taste of umami is very familiar in west Asia, especially in Japan where pure umami solutions have been used for cooking for a long time. Perception of umami taste is very difficult for people who are not living in west Asia because umami is not a familiar taste for them and consumers have only recently learned to discriminate it from the other basic tastes, although they have enjoyed the palatability of several traditional dishes with umami taste for centuries (Bellisle, 2008; Jinap & Hajeb, 2010).

The aims of the present study were firstly to identify the chromosomal regions controlling glutamate content on the tomato genome and secondly to fine map the most significant QTL and identify the underlying responsible genes. As described in chapter 4 we previously characterised a tomato collection consisting of 94 tomato genotypes, mostly cultivars and hybrids, to study the biochemical and genetic basis of fruit flavour (Tikunov et al., 2005, 2013; Ursem et al., 2008; Menéndez et al., 2012). Based on the metabolomics analyses of this collection 4 breeding lines contrasting for many flavour traits were selected and half diallel crosses were made between all parents. This resulted in the development of 6 F2 and 3 RIL populations. A marker-trait association study on a F6 population based on the cross of a round (R104) and a cherry (C085) parental line revealed a genetic region on chromosome 4 strongly associated with quantitative variation of several flavour metabolites including glutamate. We aimed to fine map this QTL using a large segregating F2 population derived from a cross between two RILs (4-066 and 4-128) contrasting for the glutamate QTL on CH4. This QTL was fine mapped and narrowed down to a region of approximately 300 kb.

Materials & Methods

Plant materials and growth conditions

The plant materials that were used in this study were the same as the materials used for fine mapping of Phenolic VOCs (This thesis; chapter 4). In summary, based on genetic and metabolic analysis of 94 tomato cultivars, 4 genotypes contrasting for many flavour traits were selected, crossed in a half diallel and 6 F2 populations and 3 RIL populations were developed. In this project we have used one of these RIL populations based on a round parent (R104) and a cherry parent (C085). From this RIL population 2 contrasting lines 4-066 and 4-128 which were contrasting for as many as possible QTL regions were selected and a large F2 population of around 5000 individuals was developed. Plants were grown to maturity, F3 seed was collected from each individual plant and stored for further use within the project.

Development of contrasting NILs

Four F6 lines from population 4 (C85 x R104) have been selected because they were heterozygous in the QTL region on chromosome 4 and almost completely homozygous for the rest of the genome. The purpose was to select F7 plants which are homozygous for the regions of interest, but with contrasting alleles (derived from C85 or R104). Up to 20 F7 plants per NIL were grown in the greenhouse of Unifarm (Wageningen University & Research) and analysed using diagnostic markers in the QTL region. NILs (F8 generation) contrasting for the glutamate QTL region only were selected and grown to maturity. Fruits were harvested at ripe stage and used for both sensory and metabolic analysis. For sensory analysis, freshly harvested fruits were transferred to the taste lab in Bleiswijk (Wageningen UR Greenhouse Horticulture). A taste panel consisting of 15 trained judges was used to assess the taste of the different fruits. For metabolomics analysis, freshly harvested fruits were cut into small pieces, snap-frozen in liquid nitrogen and ground to a powder using an IKA grinder. Samples were stored at -80 ºC until further use.

Metabolomics Analysis

The detection of amino acids, sugars and organic acids was performed by GC-TOF-MS. For extraction of samples 700 µl methanol containing 0.5 mg/ml ribitol (as an internal standard) was added to 300 mg frozen fruit powder. Samples were vortexed for 20 minutes and centrifuged at max speed for 8 min. 500 µl of the methanol extract was transferred to a new vial to which 450 µl water and 250 µl of chloroform were added. The samples were shortly vortexed and centrifuged at max speed for 8 minutes. Supernatant was diluted 8 times with pure Methanol. 40 µl of supernatant was transferred in to a crimp cap with insert and dried overnight in a speed vac and capped. Samples were analysed by GC-TOF as described in

(Carreno-Quintero et al., 2012) using a detector voltage of 1700 V. leco Chroma TOF software 2.0 was used for pre-processing of the raw data.

Genotyping

Genomic DNA was extracted from young leaves of the individual plants using the CTAB method (Devi et al., 2013)**.** Genotyping was performed by analysing SNPs in the QTL region using KASP assays, performed at the Dr. van Haeringen Laboratorium, Wageningen, the Netherlands.

QTL analysis

QTL analyses were performed with the MapQTL 6.0 software program (Van Ooijen, 2009). The non-parametric Kruskal-Wallis(KW) statistical test was used to detect associations between SNP markers and glutamate content.

Results

Developing a stable resource for QTL fine mapping

An F6 RIL population was developed based on a cross between a round (R104) and a cherry (C85) breeding line and used to map QTLs for several flavour-related compounds, including a major QTL for glutamate content on chromosome 4 at approximately 57 Mb (unpublished results). To fine map this QTL, an F2 population was developed based on two RILs, 4-066 and 4-128, which were contrasting for a large part of chromosome 4 (from 53.0 Mb until the end of the chromosome), including the glutamate content QTL region. The development of this RIL population has been explained in more detail in chapter 4. Five thousand individual F2 plants were grown to maturity in a greenhouse. Leaves of every young F2 plant were harvested and DNA was isolated. To immortalize the fine mapping population, F3 seed was harvested from each individual F2 plant and stored for further analysis.

Fine mapping of the glutamate content QTL on chromosome 4

The first step of QTL mapping was performed in the same way as the fine mapping of phenolic VOCs described in chapter 4. All 5000 individuals of the F2 population were screened using 4 markers (seq-rs6690 at 53.03 Mb, seq-rs5737 at 55.34 Mb, seq-rs7144 at 58.89 Mb, seq-rs1979 at 60.56 Mb) in the segregating region of chromosome 4, covering a 7.7 Mb region from 53.0 to 60.7 Mb. This led to the selection of 384 recombinants in this region. These 384 recombinants were screened with an additional set of 50 markers to select 50 F2 genotypes with recombinations distributed over the entire QTL region.

For every selected F2 recombinant, 20 F3 seedlings were grown and genotyped with 4 markers in the QTL region to determine the segregation pattern in the F3 generation. From each F3 family (each set of 20 seedlings), 2 plants with a homozygous recombination in the QTL region were selected for metabolic analysis. Selected F3 plants were grown to maturity and ripe fruits were harvested to determine the level of glutamate, using GC-TOF-MS. Glutamate levels ranged from 150 to 680 mg/100g (Figure S1). The first round of QTL mapping revealed a QTL for glutamate content with a maximum at 57.2 Mb (locus seq-rs4054) with a ^{10}LogP value of 6.8 which explained 35 % of the phenotypic variation. This QTL spans an interval of 2.1 Mb from 55.9 (seq-rs 6783) to 58.0 Mb (locus seq-rs 9043), based on 2 LogP units on both sides of the QTL maximum (Table 1).

Table 1. QTL analysis of the glutamate content using the selected F3 recombinants.

Position	Locus	K*	Signif.	LOGP	Mean-a	Mean-b	Ratio a/b
53.0	seq-rs6690	12.755	******	3.7	6318.5	4013.62	1.6
53.6	seq-rs188	13.789	******	4.0	6234.29	3957.32	1.6
53.7	seq-rs7746	13.789	******	4.0	6234.29	3957.32	1.6
53.8	seq-rs7748	14.281	******	4.1	6301.26	3968.76	1.6
54.2	seq-rs8228	14.281	******	4.1	6301.26	3968.76	1.6
54.3	seq-rs8225	13.882	******	4.0	6235.21	3956.68	1.6
54.5	seq-rs8247	12.916	******	3.8	6141.1	3940.03	1.6
54.5	seq-rs8244	12.916	******	3.8	6141.1	3940.03	1.6
54.8	seq-rs7028	12.916	******	3.8	6141.1	3940.03	1.6
54.8	seq-rs4160	12.916	******	3.8	6141.1	3940.03	1.6
54.8	seq-rs7032	12.916	******	3.8	6141.1	3940.03	1.6
54.9	seq-rs7034	13.186	******	3.8	6104.4	3940.03	1.5
55.0	seq-rs7037	16.092	*******	4.5	6243.55	3892.05	1.6
55.2	seq-rs7038	16.092	*******	4.5	6243.55	3892.05	1.6
55.3	seq-rs5737	16.092	*******	4.5	6243.55	3892.05	1.6
55.9	seq-rs6784	17.842	******	4.2	6154.21	3852.74	1.6
55.9	seq-rs6783	16.092	*******	4.5	6154.21	3852.74	1.6
56.5	seq-rs6776	22.119	*******	5.9	6235.9	3724.3	1.7
56.6	seq-rs6775	24.733	*******	6.5	6341.34	3709.92	1.7
56.7	seq-rs6773	24.733	*******	6.5	6341.34	3709.92	1.7
56.9	seq-rs6763	24.733	*******	6.5	6341.34	3709.92	1.7
57.0	seq-rs6762	24.733	*******	6.5	6341.34	3709.92	1.7
57.2	seq-rs4054	26.077	*******	6.8	6309.27	3664.81	1.7
57.4	seq-rs5736	22.639	*******	6.0	6199.76	3817.97	1.6
58.0	seq-rs9043	14.28	*****	3.4	6070.78	4066.64	1.5
58.1	seq-rs9040	12.458	******	3.7	6103.16	4066.64	1.5
58.2	seq-rs9039	12.827	******	3.7	6054.77	4046.16	1.5
58.3	seq-rs9034	12.959	******	3.8	5952.61	4014.03	1.5
58.3	seq-rs9033	12.959	******	3.8	5952.61	4014.03	1.5
58.4	seq-rs9030	8.695	****	2.8	5680.97	4092.97	1.4
58.7	seq-rs7137	8.662	**	2.2	5528.52	4153.47	1.3
58.8	seq-rs7140	7.113	**	1.8	5428.23	4202.97	1.3
58.9	seq-rs7143	6.049	**	1.6	5352.1	4238.44	1.3
58.9	seq-rs7144	6.049	**	1.6	5352.1	4238.44	1.3
59.1	seq-rs7146	6.049	**	1.6	5352.1	4238.44	1.3
59.1	seq-rs7149	6.049	**	1.6	5352.1	4238.44	1.3

Position: physical position of the markers. Locus; name of marker. K*: Kruskal-Wallis value. LOGP: [10]LOG P-value of the Chi-square distribution, Mean a: mean value of "a" allele, mean b: mean value of the "b" allele. Phenotypic values are in mg/kg FW.

Second round of fine mapping

In order to further fine-map the glutamate content QTL interval, we screened the entire F2 population of 5000 genotypes with two markers flanking the interval at 55.9 and 57.4 Mb. This screen resulted in 800 recombinants. We selected 381 of these F2 recombinants and screened them with 30 markers. Based on the recombination pattern, 63 F2 recombinants (Figure 1) were selected for further phenotypic analysis in the corresponding F3 plants and homozygous recombinants were selected as explained earlier.
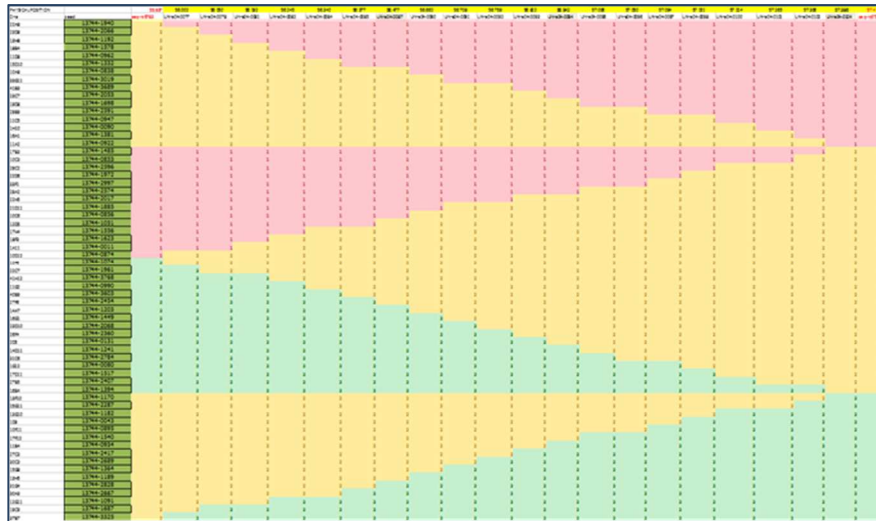
Figure 1. Genotypic scores of 63 selected F2 recombinants. Score 1 (red): homozygous for parent R104 allele; score 2 (yellow): heterozygous; score 3 (green): homozygous for parent C85 allele.

The second QTL analysis revealed an interval for glutamate content from 57.21 Mb up to the last marker used in this experiment at 57.39 Mb (Table 2). The second round of fine mapping appeared to narrow-down the QTL interval on the left side only and the QTL maximum was close to the last marker used to find recombinants, at 57.31 Mb. This indicates that the interval we had selected for the ultra-fine mapping was too narrow, resulting in a lack of recombinants on the right side of the QTL maximum. To get a better estimation of the QTL interval, we developed 15 additional markers in the QTL region to get a more precise mapping of the recombinants analysed in fine mapping round 1, which covered a region up to 60 Mb. QTL analysis of this new set showed the QTL maximum at 57.31 Mb with a LogP score of 6.5 and narrowed down the glutamate content QTL to a region of 220kb at the right side of the interval to 57.53 Mb (Table 3). Combination of all fine mapping results narrowed down the glutamate QTL region to a 320 kb region, from 57.21 to 57.53 Mb with a QTL maximum at 57.31 Mb.

Table 2. QTL analysis of the glutamate content region using the selected F3 recombinants of the second round of fine mapping.

Position	Locus	K*	LOGP	Signif.	Mean-a	Mean-b	Ratio a/b
56.48	Ultra04-0087	0.3	0.2	-	12632400	11953500	1.1
56.53	Ultra04-0088	1.8	0.7	-	12632400	11644000	1.1
56.66	Ultra04-0090	0.5	0.3	-	13153500	11323900	1.2
56.71	Ultra04-0091	2.1	1	-	13440700	10976800	1.2
56.76	Ultra04-0092	5.1	1.9	**	13860400	10605400	1.3
56.81	Ultra04-0093	6.6	2.2	**	14133900	10446600	1.4
56.94	Ultra04-0094	8.7	2.8	****	14104000	10175400	1.4
57.01	Ultra04-0095	13.4	3.9	******	14312000	9395580	1.5
57.05	Ultra04-0096	14.2	4.1	******	14247400	9150830	1.5
57.09	Ultra04-0097	21.8	5.8	*******	14758500	8967930	1.6
57.15	Ultra04-0098	20.4	5.5	*******	14359600	8909330	1.6
57.21	Ultra04-0100	23.6	6.2	*******	14566500	8909330	1.6
57.27	Ultra04-0101	23.6	6.2	*******	14566500	8909330	1.6
57.31	Ultra04-0102	32.4	8.2	*******	15168900	8615590	1.8
57.36	seq-rs5736	35.1	7.9	*******	15530200	8624520	1.8
57.39	Ultra04-0104	31.6	8	*******	15530200	8996540	1.7

Position: physical position of the markers. Locus; name of marker. K*: Kruskal-Wallis value. LOGP: 10LOG P-value of the Chi-square distribution, Mean a: mean value of "a" allele, mean b: mean value of the "b" allele. Phenotypic values represent the GCMS detector response.

Table 3. QTL analysis of the glutamate content QTL region with a set of 65 markers, using the selected F3 recombinants of the fine mapping experiment.

Position	Locus	K*	Signif.	LOGP	Mean-a	Mean-b	Ratio a/b
54.94	seq-rs7034	14.4	******	4.1	5958.97	3876.05	1.5
55.02	seq-rs7037	17.5	*******	4.8	6097.89	3829.72	1.6
55.20	seq-rs7038	17.5	*******	4.8	6097.89	3829.72	1.6
55.34	seq-rs5737	17.5	*******	4.8	6097.89	3829.72	1.6
55.88	seq-rs6784	20.2	*******	4.7	5898.52	3788.76	1.6
55.94	seq-rs6783	20.2	*******	4.7	5898.52	3788.76	1.6
56.00	Ultra04-0077	20.2	*******	4.7	5898.52	3788.76	1.6
56.48	seq-rs6776	26.7	*******	6.1	6070.09	3685.92	1.6
56.48	Ultra04-0087	23.3	*******	5.4	6070.09	3685.92	1.6
56.57	seq-rs6775	28.2	*******	6.4	6207.81	3672.55	1.7
56.70	seq-rs6773	28.2	*******	6.4	6207.81	3672.55	1.7
56.92	seq-rs6763	28.2	*******	6.4	6207.81	3672.55	1.7
56.94	Ultra04-0094	26.8	*******	6.1	6049.91	3672.55	1.6
57.01	seq-rs6762	26.8	*******	6.1	6049.91	3672.55	1.6
57.09	Ultra04-0097	28.2	*******	6.4	6021.96	3626.43	1.7
57.16	seq-rs4054	28.2	*******	6.4	6021.96	3626.43	1.7
57.21	Ultra04-0100	28.2	*******	6.4	6021.96	3626.43	1.7
57.27	Ultra04-0101	28.5	*******	6.5	5992.52	3626.43	1.7
57.31	Ultra04-0102	28.5	*******	6.5	5992.52	3626.43	1.7
57.32	Ultra04-0120	25.5	*******	5.8	5847	3740.92	1.6
57.36	seq-rs5736	25.5	*******	5.8	5847	3740.92	1.6
57.39	Ultra04-0104	25.5	*******	5.8	5847	3740.92	1.6
57.41	Ultra04-0121	22.9	*******	5.3	5799.04	3823.21	1.5
57.42	Ultra04-0122	22.9	*******	5.3	5799.04	3823.21	1.5
57.47	Ultra04-0123	19.7	*******	4.6	5634.58	3875.08	1.5
57.53	Ultra04-0124	18.8	*******	4.4	5764.8	3875.08	1.5
57.67	Ultra04-126	18.3	******	4.3	5764.8	3832.28	1.5
57.92	Ultra04-130	18.3	******	4.3	5764.8	3832.28	1.5
58.05	seq-rs9043	15.1	******	3.7	5826.58	3889.72	1.5

Position: physical position of the markers. Locus; name of marker. K*: Kruskal-Wallis value. LOGP: 10LOG P-value of the Chi-square distribution, Mean a: mean value of "a" allele, mean b: mean value of the "b" allele. Phenotypic values are in mg/kg FW.

The distribution of glutamate levels in genotypes scored at the most associated marker is shown in Figure 2. High glutamate levels are associated with the presence of the homozygous allele derived from the R104 parent.
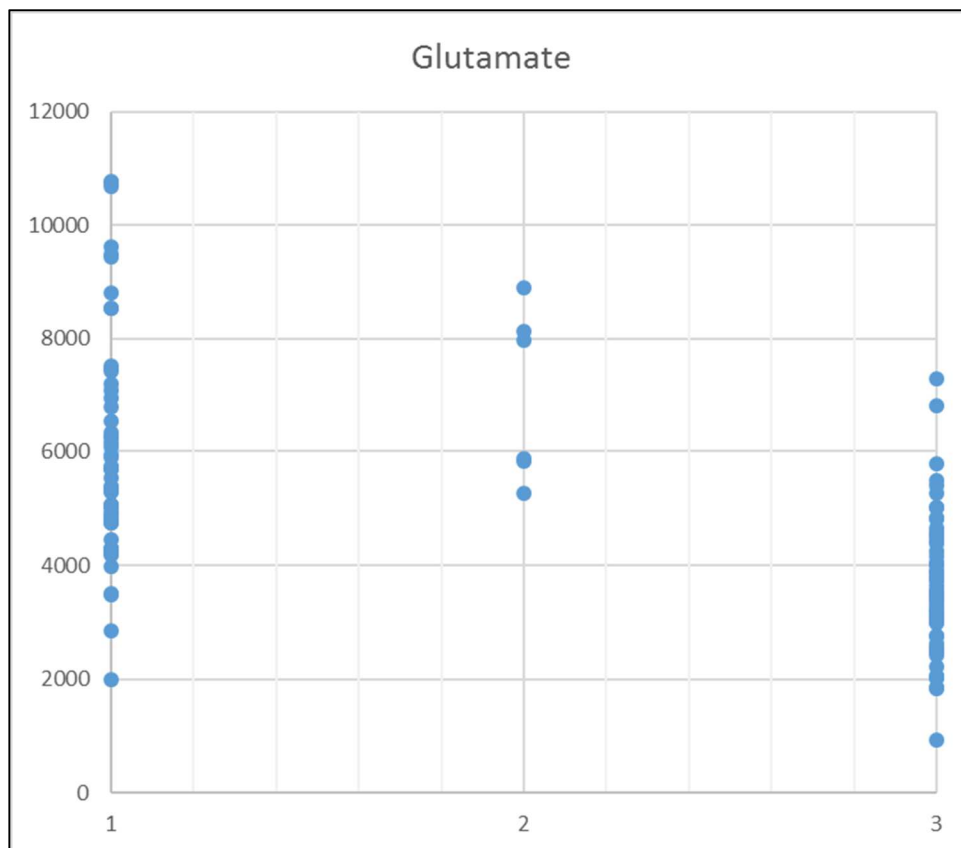
Figure 2. Scatter plot of genotypic scores of marker ultra4-0102 at 57.31 Mb on chromosome 4 and glutamate levels in the F3 recombinants used. Y-axis values represent the detector response. X-axis values are as described in Figure 3.

Sensory and metabolic analysis of contrasting NIL's

A pair of NILs was developed in this study based on the position of the glutamate QTL on chromosome 4 (Table 4), named NIL3 and NIL4. Metabolic analysis of primary metabolites revealed a 2.2-fold higher level of glutamate in NIL3 compared to NIL4 (Figure 3). Despite the fact that the glutamate level in both NILs was above the human detection threshold (1.2 mg/g), taste panel data revealed no significant difference for umami taste in this NIL pair (Figure 4). Nevertheless, NIL4 was perceived as significantly more sweet than NIL3, despite the fact that there was no difference between the NILs in sugar content (results not shown).

Table 4. Genotyping of the CH4 QTL regions in two pairs of contrasting NILS.

Chromosome	Position	Name	NIL number: 3 PV165853	4 PV165854
4	4711066	seq-rs95	BB	BB
4	5295394	seq-rs6572	BB	BB
4	6179278	seq-rs6583	BB	BB
4	11747731	seq-rs4381	BB	BB
4	50582780	seq-rs8115		
4	51776644	seq-rs6696	BB	BB
4	52664580	seq-rs6691	AA	AA
4	53710784	seq-rs3404	AA	AA
4	54785444	seq-rs4160	AA	AA
4	55209992	seq-rs7040	AA	BB
4	55936272	seq-rs6783	AA	BB
4	56475308	seq-rs6777	AA	BB
4	56475724	seq-rs6776	AA	BB
4	56748992	seq-rs8169	AA	BB
4	57161780	seq-rs4054	AA	BB
4	58174884	seq-rs9039	AA	BB
4	58340636	seq-rs9033	AA	BB
4	59390024	seq-rs7153	AA	BB
4	59883176	seq-rs4651	AA	BB
4	60168744	seq-rs7168	AA	BB
4	61089840	seq-rs7185	BB	BB
4	62821436	seq-rs1000	BB	BB
4	62997364	seq-rs7224	BB	BB

AA: homozygous parent 104; BB: homozygous parent C85.
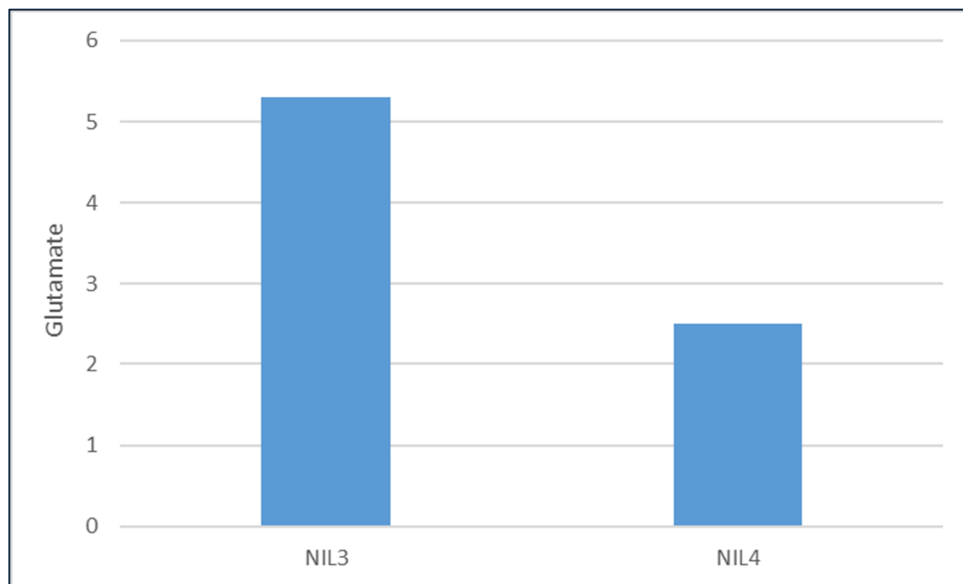
Figure 3. Level of glutamate detected in a pair of NILs analysed. Values are in mg/g FW.
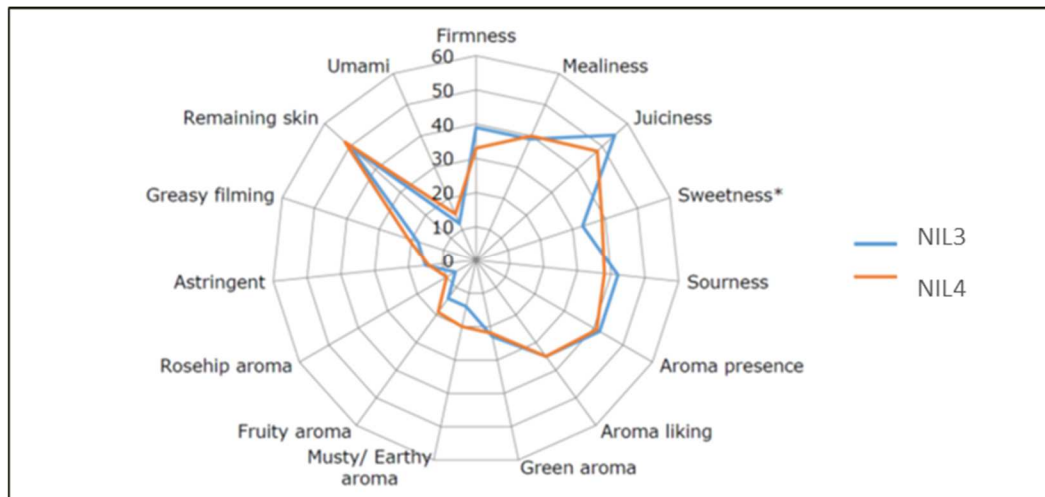
Figure 4. Spider diagram representing the attributes scored for NIL3 (PV165853) and NIL4 (PV165854).
*: Significant at P<0.05.

Candidate genes in the interval

The 320 kb QTL region harbours ~30 genes (Figure 5). We examined the expression of these genes in ripening fruits of parent C85 (minus allele) and R104 (plus allele). Of these 30 genes, seventeen showed a significant expression level but none of those were higher expressed in the plus parent versus the minus parent.

ID	chr start	Max	G	B	T	R	G	B	T	R	ITAG Hit description
				C085				R104			
Solyc04g072510	57108668	7									Mitochondrial transcription termination factor-like (AHRD V1 ***- C
Solyc04g072520	57113111	35									Mitochondrial transcription termination factor family protein (AHRD
Solyc04g072530	57116309	14									Mitochondrial transcription termination factor family protein (AHRD
Solyc04g072540	57120525	51									Mitochondrial transcription termination factor-like (AHRD V1 ***- C
Solyc04g072550	57126270	23									Mitochondrial transcription termination factor-like (AHRD V1 ***- C
Solyc04g072560	57131457	950									IST1 homolog (AHRD V1 *--- C1BSR8_9MAXI); contains Interpro dom
Solyc04g072570	57137313	58									Receptor like kinase, RLK
Solyc04g072580	57151228	4									6-phosphofructokinase 2 (AHRD V1 ***- B6TE72_MAIZE); contains In
Solyc04g072590	57161291	6									F-box/LRR-repeat protein At5g02910 (AHRD V1 ***- FBL79_ARATH);
Solyc04g072610	57169847	15									Unknown Protein (AHRD V1)
Solyc04g072620	57176804	120									Unknown Protein (AHRD V1)
Solyc04g072630	57188492	72									Os10g0363700 protein (Fragment) (AHRD V1 *-*- Q0IY59_ORYSJ)
Solyc04g072640	57199085	15									Membrane-anchored ubiquitin-fold protein 2 (AHRD V1 ***- D7M70
Solyc04g072650	57213467	8									Unknown Protein (AHRD V1)
Solyc04g072660	57219673	221									Ribosomal protein L30 (AHRD V1 ***- B3TLP4_ELAGV); contains Inte
Solyc04g072670	57223017	340									Apoptotic chromatin condensation inducer 1 (Fragment) (AHRD V1 '
Solyc04g072680	57233109	12									SAP domain containing protein (AHRD V1 *-*- C3SA89_BRADI)
Solyc04g072690	57241593	237									Aspartyl protease family protein (AHRD V1 **-- D7M2T2_ARALY); co
Solyc04g072700	57252741	9									Metal ion binding protein (AHRD V1 **-* B6U811_MAIZE); contains I
Solyc04g072740	57302588	10									High affinity sulfate transporter 1 (AHRD V1 **** SUT1_STYHA); cont
Solyc04g072760	57324262	611									High affinity sulfate transporter 2 (AHRD V1 **** SUT2_STYHA); cont
Solyc04g072770	57328456	67									tRNA modification GTPase mnmE (AHRD V1 ***- B1WVF4_CYAA5); c
Solyc04g072780	57350814	360									Genomic DNA chromosome 5 TAC clone K14A3 (AHRD V1 *--- Q9FGC
Solyc04g072790	57357337	76									Cellulose synthase (AHRD V1 **** B8XPP7_9ROSI); contains Interpro
Solyc04g072800	57368435	6									2 3-bisphosphoglycerate-dependent phosphoglycerate mutase (AH
Solyc04g072810	57371024	121									CRAL/TRIO domain containing protein expressed (AHRD V1 **-- Q75
Solyc04g072820	57381248	165									Unknown Protein (AHRD V1)
Solyc04g072830	57381394	126									Helicase sen1 (AHRD V1 *--- B6K5E8_SCHJY)
Solyc04g072840	57386200	542									tRNA-splicing endonuclease positive effector-related protein (AHR
Solyc04g072850	57401704	10									Alpha-L-arabinofuranosidase/beta-D-xylosidase (AHRD V1 **** D9D
Solyc04g072860	57410312	11									Beta-D-xylosidase (AHRD V1 **** Q8W011_HORVU); contains Interp
Solyc04g072870	57420743	10									Beta-D-xylosidase (AHRD V1 **** Q8W011_HORVU); contains Interp
Solyc04g072880	57425740	267									RNA polymerase II holoenzyme cyclin-like subunit (AHRD V1 **** S
Solyc04g072890	57438605	595									WD-40 repeat family protein (AHRD V1 **-- D7KN06_ARALY); contai
Solyc04g072900	57461646	981									Ethylene responsive transcription factor 2b (AHRD V1 *-*- C0J9I6_9F
Solyc04g072910	57470274	181									Endoplasmic reticulum-Golgi intermediate compartment protein 3
Solyc04g072920	57502534	21									Trehalose-6-phosphate phosphatase (AHRD V1 ***- D7KV74_ARALY
Solyc04g072930	57525991	89									F-box family protein (AHRD V1 ***- D7KV76_ARALY); contains Interp
Solyc04g072940	57535397	5									Unknown Protein (AHRD V1)
Solyc04g073940	57543217	6									Tetratricopeptide repeat-containing protein (AHRD V1 **-- D7MT84
Solyc04g073950	57556961	574									VQ motif family protein expressed (AHRD V1 *-*- Q10KE8_ORYSJ); c
Solyc04g073960	57582626	87									Major facilitator superfamily transporter (AHRD V1 *--- Q2NDT2_ERY
Solyc04g073970	57593111	103									cDNA clone J033025P19 full insert sequence (AHRD V1 ***- B7ERC6_

Figure 5. Microarray gene expression analysis of genes in the glutamate content QTL region in the parent lines C85 and R104. Genes with a maximum expression above the background are pink coloured in column 3.

Based on these expression values and the potential identity of the candidate genes more work has to be carried out in order to decrease the number of potential candidate genes and identify the most likely gene responsible for the glutamate content QTL.

Discussion

The first introduction of glutamic acid as a source of umami taste was in 1908 by a Japanese scientist named Ikeda, who extracted glutamic acid from sea foods (Ikeda, 2002). Umami is a Japanese word, which means delicious. Before the formal introduction of umami taste, glutamate rich foods have been used in cooking in several cultures in order to enhance the sensory qualities (Bellisle, 1999). Umami taste gives meaty and savoury flavours to foods. Only the free form of glutamate, in its L-configuration presents flavour enhancing properties (Bellisle, 1999; Populin, 2007). Recently, monosodium glutamate (MSG) the sodium salt of glutamic acid, has been used to increase the palatability of foods in both industrial made foods and home cooking, in Western as well as Eastern countries, although the consumption of MSG in eastern countries still prevails over that in western countries (Bellisle, 1999; Bellisle, 2008). Food additives which provide umami taste are categorized as a flavour enhancer. As glutamate itself has been placed in the category of safe food additives by the World Health Organization, it can be a good alternative for flavour enhancers such as sodium chloride which causes a number of health problems for human (Jinap & Hajeb, 2010).

Tomato fruits contain a high concentration of free glutamate compared to many other fruits and vegetables (Jinap & Hajeb, 2010). The concentration of glutamate increases during fruit ripening and ripe tomatoes contain much higher levels of glutamate than green tomato fruits. It has also been revealed that the inner part of the tomato fruit contains higher levels of glutamate than the outer part (skin) and pulp contains much higher concentrations of glutamate than flesh (Oruna-Concha et al., 2007). Besides de novo synthesis, glutamate present in ripe tomato fruits can be derived from two other sources: phloem sap and degradation of endogenous peptides (Sorrequieta et al., 2010; Valle et al., 1998). Peptidases are very active in ripening fruits, and they are able to release free amino acids from endogenous proteins. However, glutamate levels of ripe fruits seem to be mostly related to glutamate dehydrogenase (GDH) and gamma aminobutyric acid transferase (GABA-T) activities, two enzymes involved in the glutamate production pathway of which the activities increase when the ripening process starts, resulting in increased L-glutamate levels during the ripening transition (Forde & Lea, 2007; Sorrequieta et al., 2010).

In this study we were able to fine map our identified glutamate QTL on Chromosome 4 to an interval of 320 kb using bi-parental QTL fine mapping. One other study reported the presence of a QTL for glutamate content on chromosome 4 of tomato (Fulton et al., 2002) and these authors mentioned that this QTL might have a significant effect on tomato flavour, although this was not further elucidated. The cross between two RILs (4-128 and 4-66) that were contrasting for many flavour metabolites generated a varied F2 population which segregated for glutamate content as well as for several other taste metabolites. In

addition to glutamate we were able to map several other taste-relate metabolite QTLs such as for Brix, sucrose, acidity, aspartate and GABA in this population (data not shown). For glutamate content we also found QTLs on Chromosome 2 and 9 (data not shown), but the QTL on Chromosome 4 was the most significant. This QTL on Chromosome 4 was narrowed down to a 320 kb interval which contains 30 genes based on the current annotation of the tomato genome. Among these genes 17 showed a significant expression level in fruit of the parental lines R104 and C085 (Figure 5). Nevertheless, based on the primary gene expression data none of these genes showed a clear higher expression in the parent with the plus allele R104 (4-066), while 7 out of these 17 genes showed a higher expression level in the parent with the minus allele (C085). This leads to the possibility that the GLU content may be determined by a gene encoding a protein whose action directly or indirectly leads to decreased levels of glutamate in the minus parent compared to the plus parent. Such a gene may, for example, encode a glutamate metabolising enzyme or a repressor of the glutamate biosynthesis pathway. In addition it is possible that the difference between the two alleles of the underlying gene is mediated at the posttranscriptional level rather than at the gene expression level. Clearly, additional research is needed to identify and validate the key gene underlying the GLU QTL.

Sensory and metabolic analysis of NILs was very useful to increase our knowledge regarding the genomic region controlling our trait of interest. Metabolic analysis of the NILs confirmed the location of our primary mapped QTL for glutamate on CH4 between 55 to 60 Mb. It also confirmed that the plus allele is derived from parent R104, since NIL3 showed a higher amount of glutamate compared to NIL4 (Table 4, Figure 3). Despite a 2.2-fold difference in glutamate content, the taste panel was not able to perceive a significant difference between the two NILs in umami taste (Figure 4). This may be due to the difficulty for Western people, which made up the sensory panel, to sense umami flavour. It is noteworthy to mention, however, that NIL4 was perceived as significantly more sweet than NIL3, despite the fact that there was no difference between the NILs in sugar and acid content (results not shown). We can't rule out the possibility that the perception of sweetness is negatively influenced by glutamate, as it was recently shown that high glutamate levels (more than 4 mg/g), as found in NIL3, leads to a significant reduction in the response of the human sweetness receptor to high sugar concentrations (Shim et al., 2015).

CONCLUSION

A QTL for glutamate content was mapped on the distal part of chromosome 4. Using an F2 population derived from two contrasting RILs, this QTL was fine mapped to a region of 300kb spanning from 57 to 57.3 Mb. This region harbours 30 candidate genes. The positive allele at this interval is from parent R104

(4-066). More research is needed to identify the underlying gene responsible for glutamate content but we can already use the knowledge from this study to design molecular markers to introgress the glutamate QTL into commercial cultivars and to efficiently select for glutamate content in future breeding programs.
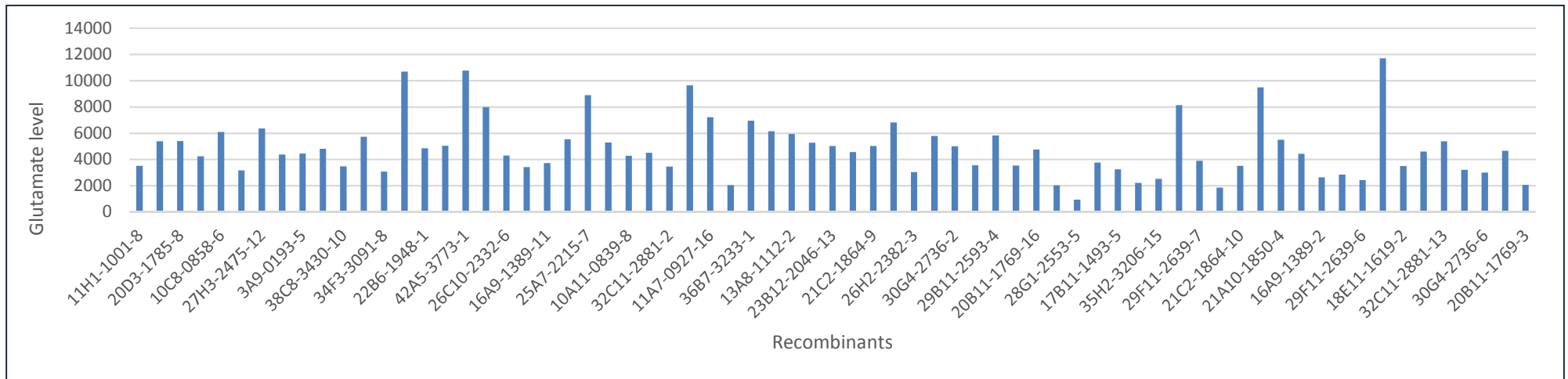
Figure S 1. Variation of Glutamate in the selected F3 recombinants used for QTL mapping.

References

Bellisle, F. (1999). **Glutamate and the UMAMI taste: sensory, metabolic, nutritional and behavioural considerations. A review of the literature published in the last 10 years**. *Neurosci. Biobehav. Rev, 23*(3), 423-438.

Bellisle, F. (2008). **Experimental studies of food choices and palatability responses in European subjects exposed to the Umami taste.** *Asia Pac J Clin Nutr, 17*(*1*), 376-379.

Carreno-Quintero, N., Acharjee, A., Maliepaard, C., Bachem, C. W. B., Mumm, R., Bouwmeester, H., . . . Keurentjes, J. J. B. (2012). **Untargeted Metabolic Quantitative Trait Loci Analyses Reveal a Relationship between Primary Metabolism and Potato Tuber Quality.** *Plant Physiol, 158*(3), 1306-1318.

Devi, K. D., Punyarani, K., Singh, N. S., & Devi, H. S. (2013). **An efficient protocol for total DNA extraction from the members of order Zingiberales- suitable for diverse PCR based downstream applications.** *SpringerPlus, 2*, 669-669.

Forde, B. G., & Lea, P. J. (2007). **Glutamate in plants: metabolism, regulation, and signalling.** *J. Exp. Bot, 58*(9), 2339-2358.

Fulton, T. M., Bucheli, P., Voirol, E., López, J., Pétiard, V., & Tanksley, S. D. (2002). **Quantitative trait loci (QTL) affecting sugars, organic acids and other biochemical properties possibly contributing to flavor, identified in four advanced backcross populations of tomato.** *Euphytica, 127*(2), 163-177.

Ikeda, K. (2002). **New Seasonings.** *Chem Senses, 27*(9), 847-849.

Jinap, S., & Hajeb, P. (2010). **Glutamate. Its applications in food and contribution to health.** *Appetite, 55*(1), 1-10.

Kader, A. A., Stevens, M. A., Marjiorie, A., & Morris, L. L. (1978). **Amino acid composition and flavor of fresh market tomatoes as influenced by fruit ripeness when harvested.** *J.Amer.Soc.Hort.Sci., 103(4)*, 541-544.

Kurihara, K. (2009). **Glutamate: from discovery as a food flavor to role as a basic taste (umami).** *Am J Clin Nutr, 90*(3), 719s-722s.

Kurihara, K. (2015). **Umami the Fifth Basic Taste: History of Studies on Receptor Mechanisms and Role as a Food Flavor.** *BioMed. Res. Int., 2015*, 189402.

Oruna-Concha, M. J., Methven, L., Blumenthal, H., Young, C., & Mottram, D. S. (2007). **Differences in glutamic acid and 5'-ribonucleotide contents between flesh and pulp of tomatoes and the relationship with umami taste.** *J Agric Food Chem, 55*(14), 5776-5780.

Shim, J., Son, H. J., Kim, Y., Kim, K. H., Kim, J. T., Moon, H., . . . Rhyu, M. R. (2015). **Modulation of sweet taste by umami compounds via sweet taste receptor subunit hT1R2.** *PLoS One, 10*(4), e0124030.

Sorrequieta, A., Ferraro, G., Boggio, S. B., & Valle, E. M. (2010). **Free amino acid production during tomato fruit ripening: a focus on l-glutamate.** *Amino Acids, 38*(5), 1523-1532.

Populin, T., Moret, S., Truant, S., & Conte, L. S. (2007). **A survey on the presence of free glutamic acid in foodstuffs, with and without added monosodium glutamate.** *Food Chem, 104*(4), 1712-1717.

Valle, E. M., Boggio, S. B., & Heldt, H. W. (1998). **Free Amino Acid Composition of Phloem Sap and Growing Fruit of Lycopersicon esculentum.** *Plant and Cell Physiol, 39*(4), 458-461.

Van Ooijen, J. W. (2009). **Software for the mapping of quantitative trait loci in experimental populations of diploid species.**

Yilmaz, E. (2001). **The Chemistry of Fresh Tomato Flavor.** *Turk J Agric For, 25 (2001)* 149-155.

Wu, X., Xu, P., Wu, X., Wang, B., Lu, Z., & Li, G. (2017). **Genome-wide association analysis of free glutamate content, a key factor conferring umami taste in the bottle gourd [Lagenaria siceraria (Mol.) Standl.].** Sci. Hortic., 225, 795-801.

Chapter 6

General Discussion

Overall goals and findings regarding genetic analysis of fruit quality in tomato

Tomato fruit quality for fresh tomato consumption is determined by a set of attributes, describing external (size, color, shape) and internal (flavour, nutritional value, texture) properties. Some of these traits have since long been the target of breeding, such as fruit size, shape and firmness, but breeding for flavour and nutritional value have only recently received attention. The aim of this thesis was an in-depth study of two important fruit quality aspects, namely shelf-life and flavour in tomato, by utilising two sources of genetic variation: (1) segregating populations based on cultivated tomato germplasm and (2) a diverse tomato core collection that consisted of wild relatives, land races and old cultivars.

Cultivated tomato germplasm has limited genetic diversity and thus is genetically 'poor' compared to the rich reservoir that exists in wild tomato species (Bai & Lindhout, 2007). Diversity of alleles for traits of interest is the key for its improvement and decrease in diversity will limit the capacity of selection and therefore genetic improvement of the crop. There are a large number of *Solanum* accessions maintained in gene banks but much of these materials remain poorly characterized, making it difficult to utilize them for crop improvement (Bauchet & Causse, 2012; Smýkal et al., 2018). Current efforts in tomato breeding are geared towards discovering and exploiting genes for important traits in tomato, by using wild and old genotypes maintained in gene banks. In chapter 2 and 3 of this thesis we have evaluated a tomato core collection consisting of 122 landraces, heirloom varieties and wild accessions for variation in several plant growth and fruit quality-related traits. This collection provided a good resource of superior alleles to incorporate in future breeding programs. 66 of the 88 cultivated accessions of this collection have been sequenced and explored for the presence of known mutations or sequence variations in key genes underlying important domestication and agronomic traits. In chapter 3, the 88 cultivated accessions of this collection plus two additional modern genotypes have been characterised in more detail for post-harvest shelf-life and thus we have now uncovered in that material novel sources for fruit shelf-life improvement.

The consumption of tomato has regularly increased during recent years, but consumers have become more and more unsatisfied about tomato fruit flavour. Flavour in tomato has not been the primary objective of crop improvement by breeders, but today genetic improvement of tomato fruit flavour is a priority as it is a subject of consumer complaints (Klee & Tieman, 2013). Most attempts for improvement of flavour in tomato have so far been focused on sugars, organic acids and brix content and more recently on fruit aroma (volatile compounds) and specific flavour enhancers such as certain amino acids. To this end, in chapter 4 and 5 of this thesis, we have investigated two metabolite QTLs involved in flavour of tomato by using a segregating population based on cultivated tomato genotypes. In chapter 4 we have fine mapped a major QTL for a volatile organic compound (2-phenyethanol) which is highly associated with

consumer liking and flavour intensity (Knudsen et al., 1993; Tieman et al., 2007; Tieman et al., 2017). In chapter 5 we have identified and fine mapped the genomic region controlling the amino acid glutamate, which is a flavour enhancer and key factor conferring umami taste in tomato (Yilmaz, 2000). These results facilitate the molecular breeding of cultivars with improved flavour and also lead to a better understanding of the biosynthesis pathway underlying these two metabolites.

Breeding for fruit flavour

Identification and narrowing down the genomic regions related to metabolite quantitative trait loci (mQTL) is a prerequisite for molecular breeding for enhanced or improved fruit flavour and represents a valuable tool for setting up novel breeding programs aimed to develop new varieties with improved taste. The majority of QTLs in plants have been identified by two approaches, either bi-parental crosses exploiting recent recombinations or association analysis which exploits historical recombinations (Cavanagh et al., 2008; Ranc et al., 2012; Korte & Farlow, 2013). Bi-parental populations have been widely employed by selecting lines displaying large phenotypic differences for one or more traits, usually with unrelated parents selected to maximize marker polymorphism and creating a single population, such as recombinant inbred lines (RILs) in which all progenies are homozygous at each locus with varying contributions from either parent. In tomato there are several examples of quantitative trait loci (QTL) analysis in bi-parental populations to dissect the genetic architecture of metabolic traits. This helped to identify several candidate and key genes underlying primary and secondary metabolites (Saliba-Colombani et al., 2001; Causse et al., 2004; Fridman et al., 2004; Tieman et al., 2006a; Tieman et al., 2006b; Tieman et al., 2007; Zanor et al., 2009; Goulet et al., 2012; Tikunov et al., 2013) and validated the efficiency of such an approach for elucidating genetic regions underlying flavour traits.

In chapter 4 and 5 of this thesis we have used a RIL population for QTL analysis of our traits of interest. Based on the metabolomics analyses of a collection of 94 tomato cultivars, 4 genotypes contrasting for many flavour traits were selected and 6 segregating F2 and 3 RIL populations were developed. In this project we have used one of these RIL populations (named population 4) based on one round (R104) and one cherry parent (C085) (Figure 1). A marker-trait association study on this F6 population revealed a genetic region on chromosome 4 strongly associated with quantitative variation of several flavour metabolites including phenolic volatiles and glutamate. In this thesis (chapter 4 and 5) we have fine mapped these two QTL regions using an F2 population based on two contrasting RILs from population 4. Since –for IP reasons- we were not allowed to use the original parents of the RIL population, two lines (named 4-066 and 4-128) which were contrasting for as many as possible QTL regions were selected and used as parents to develop a large F2 population of around 5000 individuals. F3 seed was harvested from each individual F2 plant to immortalize the fine mapping population. The two RILs, 4-066

and 4-128 contrasted for a large part of chromosome 4 (from 5.8 Mb until the end of the chromosome). Ripe fruits of these two RILs also showed contrasting levels of both metabolites, as did the original cherry and round parental breeding lines (chapter 4).
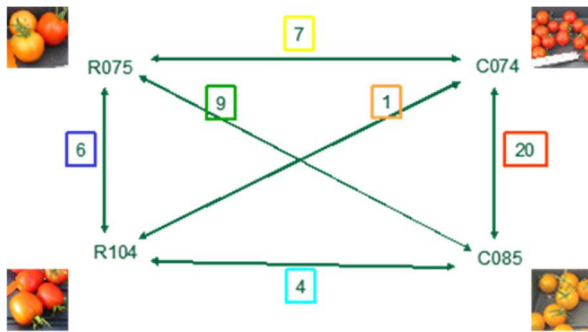

Figure 1. Development of segregating populations based on 4 contrasting genotypes.

Fine mapping of a 2-phenylethanol QTL: An important class of volatile organic compounds (VOCs) are phenolic volatiles, derived from the amino acid phenylalanine. Among these phenolic compounds, 2-phenylethanol is considered one of the most important volatiles and is often associated with fruity or floral notes of aroma in tomato fruit. We were able to fine map a region associated with this volatile on chromosome 4 to 110 kb from 54.52 to 54.63 Mb harbouring 11 candidate genes. Among the 11 genes we identified Solyc04g063350 (*Dcx1*), which is annotated as 3-methyl-2-oxobutanoate dehydrogenase, as candidate gene responsible for the variation of 2-phenylethanol in our mapped interval. *Dcx1* is a member of the family of 2-oxoacid dehydrogenases which catalyse the oxidative decarboxylation of their respective substrates (Yeaman, 1986, 1989; Dickinson & Dawes, 1992). With the previous knowledge of the role of decarboxylases on 2-phenylethanol biosynthesis in tomato (Tieman et al., 2006a) we suggest that the underlying gene (*Dcx1*) might carry out the conversion of phenylalanine to phenylethylamine in a decarboxylation reaction. The observed co-segregation of a nitrogen-containing compound (2-nitrophenylethane) with phenylethanol and phenylacetaldehyde in our study, coupled to a severe reduction in the levels of all these three phenolic VOCs in VIGS treated fruits (chapter 4), can be seen as a validation of our hypothesis that this gene acts at a step prior to deamination. In spite of this circumstantial evidence, a more appropriate reverse genetics approach is needed to unequivocally demonstrate the function of the target gene. Transformation of a high-phenylethanol tomato cultivar with the positive *Dcx1* allele using a CRISP-CAS construct to knocking out the *Dcx1* gene and analysis of the expression of the gene in transgenic plants is currently in progress in our group. Subsequently, the candidate gene can be overexpressed in a low phenylethanol accumulator line to see whether overexpression leads to an increase

in the levels of phenylethanol.The most important application of these results is the use of this gene as a perfectly coupled genetic marker to improve this trait in breeding programs using marker assisted selection.

Fine mapping of a Glutamate QTL: Glutamate is the major free amino acid found in tomato juice and comprises up to 45% of the total weight of free amino acids in fresh tomato juice (Yilmaz, 2001). Glutamate is the key factor for delivering the umami taste in tomato. Only the free form of glutamate, in its L-configuration, presents flavour enhancing properties (Bellisle, 1999; Populin et al., 2007). We identified a major QTL for glutamate on chromosome 4 and this QTL was fine-mapped into a region of approximately 300 kb from 57 to 57.32 Mb (chapter 5). To our knowledge this is the first fine mapping study for glutamate content in tomato. The detected QTL was validated in a pair of near Isogenic Lines (NIL) contrasting for the QTL region in the same genetic background (chapter 5). This current 300 kb region harbours ~30 candidate genes. Of these 30 genes, seventeen showed a significantly different expression level but based on the primary gene expression data none of these genes showed a higher expression in the parent with the plus allele. Based on these expression values and the potential identity of the candidate genes more work has to be carried out in order to decrease the number of potential candidate genes and identify the most likely gene responsible for the glutamate content QTL. The fine mapped region can be further delimited by adding more markers in this region or by performing another round of fine mapping. Alternatively the QTL can be narrowed down by determining the QTL effects in selected F3 families segregating for specific regions in the QTL interval, in a similar way as we have done for the phenylethanol QTL in chapter 4. Although additional research is needed to identify the underlying gene(s) responsible for glutamate content we can already use the knowledge from this study by designing molecular markers to introgress the allele from one of the parents , or other sources of the same QTL, into commercial cultivars by marker assisted selection.

Breeding for fruit shelf-life

Fruit shelf-life represents one of the most investigated quality traits in tomato and today with the presence of global markets for trading of fresh products, minimizing postharvest losses and maintaining very good (export) quality is of increasing interest. Postharvest shelf-life of tomato like other fleshy fruits is under the influence of the ripening process which is a genetically determined program of tissue differentiation involving many physiological processes, including the production of aromatic compounds, colour changes, flesh softening and cellular turgor changes (Prasanna et al., 2007; Mercado et al., 2011). An effective way to minimize fruit postharvest losses is the use of prolonged shelf-life varieties. The currently used long

shelf-life varieties in tomato are ripening mutants, like *rin*, *nor* and *alcobaca*, in which the entire ripening program is halted or slowed down. These mutants not only have a longer shelf-life, but also have less colour and a poor taste. In our study we aimed to find novel genetic resources with a long shelf-life, but with a normal progression of ripening, in order to avoid the negative effects of ripening mutants on flavour. We have characterised a collection of tomato genotypes consisting of 90 heirloom tomatoes and land races for their fruit post-harvest shelf-life during 42 days of storage (chapter 3). In addition to the identification of several promising shelf-life varieties with superior storability properties we have provided valuable material for more detailed studies on the molecular mechanisms governing shelf-life. The collection that we characterised for shelf-life was very variable regarding shelf-life attributes and these results indicate that there is still a high potential in tomato germplasm collections for fruit shelf-life improvement. Six accessions of the collection with contrasting shelf-life were selected for biochemical characterisation. This provided insight in the changes in cell wall polysaccharide monomers and in the levels and changes of primary and secondary metabolites during the post-harvest ripening process. The materials that we have selected as contrasting shelf-life genotypes are suitable for analyses of other physiological processes that affect shelf-life, such as cell wall turgor, cuticle architecture and thickness. Given that 66 genotypes of this collection have already been sequenced, a comprehensive study of the biochemical processes underlying the shelf-life of contrasting genotypes will be very helpful for identification of key genes underlying the variation in shelf-life and for designing breeding programs for improved tomato fruit shelf-life.

In this study we have identified several promising genotypes for post-harvest shelf-life. They exhibited very good performance regarding firmness loss, water loss and fruit decay during the storage period. These genotypes have also been characterised for flavour metabolites (results not shown) and some of them exhibited high levels of important flavour-related metabolites. In most cases, especially in case of ripening mutants, genotypes with good shelf-life performance have a poor flavour. In follow-up experiments we intend to cross the long shelf-life genotypes with established high-flavour tomato varieties to test if we can make tomato hybrids with extended shelf-life and better flavour. In addition we intend to study the genetics of the long shelf-life traits found in this study using population development and QTL analysis.

Analysis of fruit cell wall composition

Fruit softening is the major determining factor of fruit shelf-life. Cell wall metabolism during ripening has been considered the main factor determining fruit softening and textural changes in ripening fruits and is mainly associated with the dissolution of the middle lamella and the modification of the composition and structure of polymers present in the primary cell wall (Brummell, 2006). Cell wall composition and structure and its related degrading enzymes have been the targets of research for more than 40 years but despite

this intensive research only a few key individual determinants of fruit softening have been identified (Mercado et al., 2011). From that research it has become clear that pectin metabolism plays a critical role in softening of most fruits, but the exact structural modifications and the role of enzymes underlying this process are unresolved as yet. Further knowledge of cell wall structure and composition, as well as the interaction of this cellular compartment with the different families of cell wall-degrading enzymes, is fundamental towards gaining insight into the process of fruit softening and therefore the control of it.

We monitored the changes in the content of three pectin sugars during postharvest storage of tomato fruit. The content of the pectin sugars galactose (Gal), arabinose (Ara) and galacturonic acid (Gal A) underwent considerable changes during postharvest storage, although these changes were different among the six selected genotypes (chapter 3). The decline in Ara and Gal started from green stage until 14 days after storage when fruits are completely ripe. After that time point, there was no significant change in the levels of these two sugars. The amount of cell wall Gal A ncreased from green stage until 14 days after storage and after that time point it decreased. At the same time, 14 days after storage, the increase in the amount of soluble Gal A in the fruit pericarp was observed in all six genotypes, but more solubilisation of Gal A was observed for the two short shelf-life genotypes.

The decrease in the amount of cell wall galactose from tomato fruit was reported for the first time in 1979 by Gross and Wallner (Gross & Wallner, 1979) when they observed that both in the *rin* mutant and normal ripening tomato cell walls become lower in their galactose content during ripening despite the fact that in the *rin* mutant the enzyme polygalacturonase is not active. After that report the galactose content of tomato fruit cell walls during ripening has been a major research target in many studies. In several previous studies it has been observed that the decrease in the cell wall galactose content occurs in both ripening mutants and normal ripening tomatoes, but the amount of solubilized Gal in the pericarp of normal ripening tomatoes was always much higher than in the ripening mutants (Gross & Wallner, 1979; Gross, 1983; Gross & Sams, 1984; Gross, 1985; Seymour et al., 1990; Kim et al., 1991). Similar results were observed in our study (chapter 3). The decline in cell wall Gal occurred in all six genotypes, but the amount of solubilized Gal in the two short shelf-life genotypes was much higher. Several hypothesis were brought forward in the past years to account for these results. The high amount of soluble galactose in the pericarp of normal ripening tomatoes (soft fruits), but not in ripening mutants (firm fruits) might be due to its relationship with ethylene or the decrease in the ability to metabolize solubilized galactose in normal ripening tomatoes. Both of these hypothesises were rejected later and eventually it was proposed that increase in solubilized galactose is most likely due to an increase in the rate of galactosyl solubilisation from the cell wall rather than other fruit ripening-related processes (Kim et al., 1991, 1999). The mechanism of net loss in galactosyl residues from the cell wall has been suggested to be related to a

reduced rate of *de novo* galactan synthesis or to hydrolysis of galactan by the enzyme β-galactosidase II (Gross & Wallner, 1979; Lackey et al., 1980; Gross, 1983). During ripening, substantial modification of cell wall integrity has shown to occur and several studies have revealed that the cell walls undergo turnover, i.e. degradation of cell wall polysaccharides coincident with synthesis and insertion of replacement polymers (Labavitch, 1981; Mitcham et al., 1989; Mitcham et al., 1991; Goulao & Oliveira, 2008). Our observations also point in the direction that synthesis and extension of the galactan chains play an important role in cell wall polysaccharide synthesis in the soft genotypes studied here and likewise the cell wall Gal content in these soft genotypes is higher than in the other genotypes. The reason for the high amount of soluble Gal in the pericarp of soft genotypes might be due to hydrolysis of cell wall galactan chains form the highly branched, Gal-rich cell walls during ripening. . Last but not least, the results presented in our study support the suggestion that reduced synthesis accounts for the loss of galactose from cell walls during ripening and the decline in the amount of other pectin sugars in the fruit cell walls may also be related to turnover of these sugars rather than hydrolysis and removal of them from the cell wall. Detailed further studies such as polysaccharide fractionation and oligosaccharide phenotyping in each post-harvest time-point should provide more information about the structure of the cell wall polysaccharides and their turnover during ripening and post-harvest of the present material.

Putrescine to maintain the quality of fruits during post-harvest storage?

During storage, long and short shelf-life genotypes showed considerable differences in their content of the polyamine putrescine (chapter 3). The fruits with long shelf-life showed a clear increase in putrescine (Ptc) levels during storage as compared to the short shelf-life genotypes. A high content of putrescine was reported earlier in the fruits of the *Alcobaca* landrace of tomato. The fruits of this ripening mutant ripen slowly and have prolonged keeping qualities and it was suggested that the enhanced putrescine levels found in this line may be responsible for the superior ripening and storage features (Dibble et al., 1988). Exogenous application of polyamines including putrescine either at pre-harvest (during fruit growth and ripening on tree) or after harvest (postharvest treatments) a common method to delay the postharvest ripening process in several fruits such as pomegranate, strawberry, plum, apricot, mango, tomato and cucumber (Barman, et al., 2011; Koushesh saba et al., 2012; Jia et al., 2018; Wannabussapawich & Seraypheap, 2018). Exogenous application of putrescine will result in maintaining fruit quality attributes such as colour, firmness, acidity, and total soluble solids content during postharvest storage (Serrano & Valero, 2018). One of the proposed mechanism by which endogenous putrescine may increase shelf-life and maintain fruit quality is its inverse relationship with ethylene production, since the ethylene and polyamine biosynthetic pathways share and compete for s-adenosylmethionine (SAM) as common

precursor (Smith, 1985; Serrano & Valero, 2018). In addition, exogenous application of polyamines has been shown to increase the content of antioxidant compounds and the activity of antioxidant enzymes, which results in a decrease in accumulation of ROS (reactive oxygen species) and in this way senescence processes and postharvest over-ripening in fruits will be delayed (Sharma et al., 2017; Serrano & Valero, 2018). Nevertheless, the precise physiological and molecular mechanisms by which putrescine and other polyamines increase the fruit shelf-life are not clear yet and need further research.

Conclusions

The aim of this study was to address two issues that tomato breeding is currently facing: (i) a lack of genetic diversity in cultivated germplasm which limits the capacity for crop improvement and (ii) a lack of breeding tools for the improvement of tomato flavour, in order to cope with the dissatisfaction of consumers regarding tomato fruit flavour. Taking into account that tomato is considered as a model species for fleshy fruit development, ripening and composition, knowledge provided in this thesis will help in understanding the quality attributes in other fleshy fruits as well.

Recent breakthroughs in ~Omics approaches and the availability of a tomato genome sequence have led to a significant progress in providing fundamental knowledge about the chemistry of flavour metabolites and their biosynthesis pathways. The mapping and further identification of QTLs controlling several flavour components including sugars, acids, and volatiles (although less than other metabolites) has progressed significantly in the last decade and now starts to provide a molecular toolbox to assist breeders in improving fruit flavour. The current challenge of tomato fruit improvement may be to achieve an improved flavour in the context of high yield and long postharvest shelf-life. Volatile organic compounds have been identified that show a high correlation with consumer liking, after taste sensation and perception of sweetness, independently of sugar content which was shown to have a negative correlation with yield. In this regard the observation that increasing the levels of volatiles can improve the fruit flavour without a negative influence on yield is very promising (Klee & Tieman, 2013). On the other hand, many volatiles contribute to flavour of tomato and many genes influence their chemistry and synthesis, so altering one or a few of them may not have a major effect on the final flavour. Identification and improvement of the key regulators of the flux through aroma production pathways will likely be more advantageous. The amino acid phenylalanine is the direct precursor of several important flavour volatiles in tomato and thus identification and manipulation of the rate-determining enzymes in this pathway maylead to higher increases in all of the downstream volatiles which previously have been shown to have positive effects on flavour in general and in tomato fruit in particular.

A certain percentage of consumer dissatisfaction with the quality of modern hybrids is associated with the harvest of immature fruits and subsequent postharvest handling to increase the fruit shelf-life. Use of long shelf-life varieties which are not ripening mutants, but remain firm for an extended period of time after harvest and allow harvesting of fruits at riper stages will result in fruits with higher quality and more appreciation in the market. Our results showed that there is still a high potential for shelf-life improvement by the use of genotypes maintained in germplasm collections.

One important aspect we did not address in this study is the environmental influence on fruit quality aspects such as fruit flavour and storability. The chemical composition of a fruit can be affected by changing the climate condition in the greenhouse and fruits harvested in different seasons and under different environmental conditions can have different chemical properties. The environmental factors directly affect the expression of quality traits and thus may affect QTL detection. To reduce the effect of environment and experimental effect of phenotyping on genetic studies of quality traits it is necessary that experiments are conducted on the same sites and in various seasons. In our study we tried to kept the environmental variation as low as possible, by working in controlled greenhouses and growing the plants in different seasons. Our assumption was based on the fact that the greenhouse condition during all the performed experiments was the same in terms of environmental factors such as light intensity, CO_2 concentration, temperature and nutrition. But the best strategy is indeed to check and record these environmental factors during each experiment and make them completely under control.

References

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

Barman, K., Asrey, R., & Pal, R. K. (2011). **Putrescine and carnauba wax pretreatments alleviate chilling injury, enhance shelf life and preserve pomegranate fruit quality during cold storage.** *Scientia Horticulturae, 130* 795–800.

Bauchet, G., Causse, M. (2012). **Genetic diversity in tomato (Solanum lycopersicum) and its wild relatives.** Open access peer-reviewed chapter.

Brummell, D. A. (2006). **Cell wall disassembly in ripening fruit.** *Funct. Plant Biol, 33*(2), 103-119.

Cavanagh, C., Morell, M., Mackay, I., & Powell, W. (2008). **From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants.** Curr. Opin. Plant Biol, 11(2), 215-221.

Dibble, A. R. G., Davies, P. J., & Mutschler, M. A. (1988). **Polyamine Content of Long-Keeping Alcobaca Tomato Fruit.** *Plant Physiol, 86*(2), 338-340.

Dickinson, J. R., & Dawes, I. W. (1992). **The catabolism of branched-chain amino acids occurs via 2-oxoacid dehydrogenase in S'charomycescerevisiae.** *J. Gen. Appl. Microbiol, 138*, 2029-2033.

Fridman, E., Carrari, F., Liu, Y. S., Fernie, A. R., & Zamir, D. (2004). **Zooming in on a quantitative trait for tomato yield using interspecific introgressions.** *Science, 305*(5691), 1786-1789.

Goulao, L. F., & Oliveira, C. M. (2008). **Cell wall modifications during fruit ripening: when a fruit is not the fruit.** *Trends Food Sci Technol, 19*(1), 4-25.

Goulet, C., Mageroy, M. H., Lam, N. B., Floystad, A., Tieman, D. M., & Klee, H. J. (2012). **Role of an esterase in flavor volatile variation within the tomato clade.** *Proceedings of the National Academy of Sciences, 109*(46), 19009.

Gross, K. C., & Wallner, S. J. (1979). **Degradation of Cell Wall Polysaccharides during Tomato Fruit Ripening.** *Plant Physiol, 63*(1), 117-120.

Gross, K. C. (1983). **Changes in free galactose, myo-inositol and other monosaccharides in normal and non-ripening mutant tomatoes.** *Phytochemistry, 22*(5), 1137-1139.

Gross, K., C. (1984). **Fractionation and partial characterization of cell walls from normal and non-ripening mutant tomato fruit.** *PHYSIOL. PLANT., 62*, 25-32.

Jia, B., Zheng, Q., Zuo, J., Gao, L., Wang, Q., Guan, W., & Shi, J. (2018). **Application of postharvest putrescine treatment to maintain the quality and increase the activity of antioxidative enzyme of cucumber.** *Scientia Horticulturae, 239*, 210-215.

Kim, J., Gross, K. C., & Solomos, T. (1991). **Galactose metabolism and ethylene production during development and ripening of tomato fruit.** *Postharvest Biol. Technol, 1*(1), 67-80.

Kim, J., Solomos, T., & Gross, K. C. (1999). **Changes in cell wall galactosyl and soluble galactose content in tomato fruit stored in low oxygen atmospheres.** *Postharvest Biol. Technol, 17*(1), 33-38.

Klee, H. J., & Tieman, D. M. (2013). **Genetic challenges of flavor improvement in tomato.** *Trends Genet, 29*(4), 257-262.

Knudsen, J. T., Tollsten, L., & Bergström, L. G. (1993). **Floral scents—a checklist of volatile compounds isolated by head-space techniques**. *Phytochemistry, 33*(2), 253-280.

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

Koushesh saba, M., Arzani, K., & Barzegar, M. (2012). **Postharvest polyamine application alleviates chilling injury and affects apricot storage ability.** *J Agric Food Chem, 60*(36), 8947-8953.

Labavitch, J. M. (1981). **CELL WALL TURNOVER IN PLANT DEVELOPMENT.** *Ann. Rev. Plant PhysioL, 31*, 385-406.

Lackey, G. D., Gross, K. C., & Wallner, S. J. (1980). **Loss of tomato cell wall galactan may involve reduced rate of synthesis.** *Plant Physiol, 66*(3), 532-533.

Mercado, J. A., Pliego-Alfaro, F., Quesada, M. A (2011). **Fruit Shelf Life and Potential for Its Genetic Improvement.** In *Breeding for Fruit Quality* (First ed.): John Wiley & Sons, Inc.

Mitcham, E. J., Gross, K. C., & Ng, T. J. (1989). **Tomato Fruit Cell Wall Synthesis during Development and Senescence: In Vivo Radiolabeling of Wall Fractions Using [^{14}C] Sucrose.** *Plant physiol, 89*(2), 477-481.

Mitcham, E. J., Gross, K. C., & Ng, T. J. (1991). **Ripening and cell wall synthesis in normal and mutant tomato fruit.** *Phytochemistry, 30*(6), 1777-1780.

Populin, T., Moret, S., Truant, S., & Conte, L. S. (2007). **A survey on the presence of free glutamic acid in foodstuffs, with and without added monosodium glutamate.** *Food Chemistry, 104*(4), 1712-1717.

Prasanna, V., Prabha, T. N., & Tharanathan, R. N. (2007). **Fruit ripening phenomena-an overview.** *Crit. Rev. Food Sci. Nutr, 47*(1), 1-19.

Ranc, N., Muños, S., Xu, J., Le Paslier, M.-C., Chauveau, A., Bounon, R., . . . Causse, M. (2012). **Genome-wide association mapping in tomato (Solanum lycopersicum) is possible using genome admixture of Solanum lycopersicum var. cerasiforme.** G3 (Bethesda, Md.), *2*(8), 853-864.

Saliba-Colombani, V., Causse, M., Langlois, D., Philouze, J., & Buret, M. (2001). G**enetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits.** *Theor Appl Genet, 102*(2), 259-272.

Seymour, G. B., Colquhoun, I. J., Dupont, M. S., Parsley, K. R., & R. Selvendran, R. (1990). **Composition and structural features of cell wall polysaccharides from tomato fruits.** *Phytochemistry, 29*(3), 725-731.

Serrano, M., & Valero, D. (2018). **Application of Polyamines to Maintain Functional Properties in Stored Fruits.** *Methods Mol Biol, 1694*, 449-458.

Sharma, S., Pareek, S., Sagar, N. A., Valero, D., & Serrano, M. (2017). **Modulatory Effects of Exogenously Applied Polyamines on Postharvest Physiology, Antioxidant System and Shelf Life of Fruits: A Review.** *Int J Mol Sci, 18*(8).

Smith, T. A. (1985). **Polyamines.** Annu Rev Plant Biol*, 36*(1), 117-143.

Smýkal, P., Nelson, M. N., Berger, J. D., & Wettberg, E. J. B. v. (2018). **The Impact of Genetic Changes during Crop Domestication**. *agronomy, 8, 119*.

Tieman, D., Taylor, M., Schauer, N., Fernie, A. R., Hanson, A. D., & Klee, H. J. (2006a**). Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde.** *Proc Natl Acad Sci U S A, 103*(21), 8287-8292.

Tieman, D. M., Zeigler, M., Schmelz, E. A., Taylor, M. G., Bliss, P., Kirst, M., & Klee, H. J. (2006b). **Identification of loci affecting flavour volatile emissions in tomato fruits.** *J Exp Bot, 57*(4), 887-896.

Tieman, D. M., Loucas, H. M., Kim, J. Y., Clark, D. G., & Klee, H. J. (2007). **Tomato phenylacetaldehyde reductases catalyze the last step in the synthesis of the aroma volatile 2-phenylethanol.** *Phytochemistry, 68*(21), 2660-2669.

Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., . . . Klee, H. (2017). **A chemical genetic roadmap to improved tomato flavor.** *Science, 355*(6323), 391-394.

Tikunov, Y. M., Molthoff, J., de Vos, R. C. H., Beekwilder, J., van Houwelingen, A., van der Hooft, J. J. J., . . . Bovy, A. G. (2013). **NON-SMOKY GLYCOSYLTRANSFERASE1 Prevents the Release of Smoky Aroma from Tomato Fruit.** *Plant Cell, 25*(8), 3067-3078.

Wannabussapawich, B., & Seraypheap, K. (2018). **Effects of putrescine treatment on the quality attributes and antioxidant activities of 'Nam Dok Mai No.4' mango fruit during storage.** *Scientia Horticulturae, 233*, 22-28.

Yeaman, S. J. (1986). **The mammalian 2-oxoacid dehydrogenases: a complex family.** *Trends Biochem Sci, 11*(7), 293-296.

Yeaman, S. J. (1989). **The 2-oxo acid dehydrogenase complexes: recent advances.** *Biochem. J, 257*(3), 625-632.

Yilmaz, E. (2001). **The Chemistry of Fresh Tomato Flavor.** *Turk J Agric For, 25*, 149-155.

Zanor, M. I., Rambla, J. L., Chaib, J., Steppa, A., Medina, A., Granell, A., . . . Causse, M. (2009). **Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents.** *J Exp Bot, 60*(7), 2139-2154.

Summary

The quality of fresh fruits comprises several properties including external characteristics such as size, shape and colour, as well as internal qualities like flavour (taste and aroma), texture (firmness, mealiness, juiciness), and nutritional values. These factors have a large influence on the consumer liking and acceptance of the fruits. This thesis focused on two important fruit quality aspects, fruit shelf life and fruit flavour by utilisation of two sources of genetic variation: (1) segregating populations based on cultivated tomato germplasm and (2) a diverse tomato core collection that consist of wild relatives, land races and old cultivars.

In **chapter 2** a tomato core collection consisting of 122 accessions (90 cultivated+ 32 wild accessions) was explored for phenotypic and genotypic variation in several plant growth, yield and fruit quality traits in order to introduce novel traits from land races, old cultivars and wild relatives into future tomato breeding programs. The collection was phenotyped for the crop growth-related traits abscission zone, inflorescence branching, vegetative outgrowth of the inflorescence, growth rate and earliness of flowering and the fruit characteristics number of ripe fruits per plant, fruit weight, firmness, total soluble solid content (Brix), colour and shape. 66 accessions of this collection have been re-sequenced, either in this study or in previous studies, and used for allele mining of known mutations or variants affecting plant architecture (*self-pruning* (*sp*), *compound inflorescence* (*s*), *jointless* (*j*)-and *potato leaf* (*c*)), fruit size and shape (*fruit weight 2.2* (*fw2.2*), *fruit weight 3.2* (*fw3.2*), *fruit weight 11.3* (*fw11.3*), *locule number* (*lc*), *fasciated* (*fas*), *ovate* and *sun*) and fruit colour (yellow flesh mutation (*r*), *tangerine* (t^{3183}), *old-gold-crimson* (*og*), *green flesh* allele (*gf⁴*) and pink mutation (*y*). In **chapter 3** 90 cultivated accessions including heirloom and land races were characterised for fruit post-harvest shelf-life during 42 days storage. Genotypes were characterised for shelf-life attributes such as colour pigments, water loss, firmness and fruit decay. Genotypes were categorized based on their performances regarding shelf-life attributes during storage. Several promising genotypes with superior storability performance were identified and are good candidates for incorporation in breeding programs aimed at developing improved shelf-life varieties. Six genotypes of the collection with contrasting shelf-life were selected for biochemical characterisation of shelf-life and changes in polysaccharide monomers, primary and secondary metabolites were monitored during 35 days at 5 post-harvest time-points. During storage, long and short shelf-life genotypes showed considerable differences in their content of sugars, such as galactose, and polyamines, such as putrescine. Interpretation of cell wall composition data based on the ratios of pectin monosaccharides revealed that the pectin of firm genotypes is more linear and has a lower degree of branching compared to soft genotypes. This knowledge provides a better insight in the biochemical basis of long shelf life tomatoes. **Chapter 4** aimed at dissection

of genomic regions controlling the volatile organic compound 2-phenylethanol. A previous marker-trait association study on a collection of 94 tomato cultivars revealed a genetic region on chromosome 4 strongly associated with quantitative variation of phenolic volatiles. This QTL on CH4 was fine mapped in this study using a F2 population based on two RILs contrasting for this associated region. Genetic and metabolic analyses of this population allowed to narrow down the region associated with the phenolic volatiles from 54.52 to 54.63 Mb, harbouring 11 candidate genes and 110kb in size. We identified the gene Solyc04g063350 (*Dcx1*), encoding for a putative 2-oxoacid dehydrogenase enzyme that may possess amino acid decarboxylation activity, as a major candidate gene for the variation of phenylethanol content in our mapped interval. Gene expression analysis of fruits of the parental genotypes of the crossing population and fruits derived from contrasting plants of F3 families showed a higher expression of *Dcx1* in fruits with a higher content of the phenolic volatiles. To test the function of the *Dcx1* gene a VIGS (Virus Induced Gene Silencing) experiment was carried out by injection of a *Dcx1* VIGS construct in detached immature green fruits of tomato cultivar "Solarino" which has the positive *Dcx1* allele, a high *Dcx1* expression and a high content of the phenolic volatiles in fruits. Metabolic analysis of silenced fruits revealed a severe reduction in the levels of all three phenolic volatiles 2-phenylethanol, phenylacetaldehyde and 2-nitrophenylethane. Other reverse genetics approaches are currently in progress to unequivocally demonstrate the function of the target gene. In **chapter 5** we conducted metabolic quantitative trait loci (mQTL) analysis to localise and narrow down the genomic region controlling the levels of the amino acid glutamate, conferring Umami taste in tomato. A major QTL for glutamate content was identified on chromosome 4. This QTL was fine mapped using a F2 population consisting of 5000 individual plants derived from a cross between two RILs contrasting for this region. F3 seed was harvested from each individual F2 plant to immortalize the fine mapping population. Using SNP markers and metabolic analysis by GC-TOF-MS, this QTL was fine-mapped into a region of approximately 300 kb from 57 to 57.32 Mb. The detected QTL was validated in a pair of near Isogenic Lines (NILs), which were genetically identical, but contrasting in the QTL region only. The current 300kb region harbours ~30 candidate genes. Of these 30 genes, seventeen showed a significant expression level, but based on the primary gene expression data none of these genes showed a higher expression in the parent with the plus allele. This suggests that the key gene underlying this QTL may be a factor encoded by the minus allele that reduces the level of glutamate, for example by converting glutamate into its direct product GABA. Clearly, identification of the key gene or genes responsible for the variation in glutamate needs further research.

Acknowledgements

This journey would not have been possible without the support of my family, colleagues and friends. Therefore, I desire to express my gratefulness here.

My first acknowledgements go to my daily supervisor **Dr. Arnaud G. Bovy**. I would like to thank him firstly for giving me the opportunity of doing a PhD in Wageningen University & Research, and secondly for his guidance and all the knowledge he shared with me during these 5 years.

I also would like to thank the other members of the fruit quality group; **Dr. Yury Tikunov, Fien Meijer-Dekens** and **Jos Molthoff** for all their technical and lab support during my PhD.

I would like to present my gratitude to my PhD promoter, **Prof. Richard Visser**, for his contributions and being supportive throughout the project. Richard! I have to specially thank you for helping me to make this thesis ready for submission otherwise I was not able to make my defence at this time.

I would also to express my appreciation to the PBR secretary ladies **Janneke, Nicole, Letty** and **Danielle**, thanks for helping me with the bureaucratic stuff during my PhD. Of course you made the PhD life so easier for me and other PhDs!

During these years in Wageningen University I spend a great amount of my time with my two office mates, **Dr. Rients Niks** and **Atiyeh Kashaninia**. I would like to thank you both for all of the scientific and non-scientific discussions that we shared during these years. You have been of an incredible help and I am deeply grateful for that.

My next gratitude will go to my fellow PhDs and post-docs in the fruit quality group, **Marine Paupiere**, **Ying Liu**, **Yi Wu** and **Fatima Carvajal moreno**. I thank all four of you for the nice moments that we shared together. I would like to extend my acknowledgement to Marine a little further, because of all of her emotional support during the four years that we were in this group. Marine! I am glad that we were in this adventure together.

Dr. Ruud de Maagd, Dr. Richard Finkers and Michiel Lammers, members of the phenotyping and allele mining group, I would like to thank you for all your practical support, scientific discussions, suggestions and comments on my experiments during our meetings.

Many thanks to the technicians in PBR and Bioscience. Especially to Irma, Annemarie Dechesne, Jeroen and Henriette. Thanks for your help in the labs and teaching me lab works and using the instruments.

My PBR fellow PhDs, I enjoyed the exchange of talks and laughs during our coffee breaks, luncheons and events. I've got to know many cultures because of you. I wish you success and happiness wherever you are. Xuan Xu, Cynara, Lorena, Noor, Mathilde, Charlotte, Behzad, Narges, Anne, Carolina, Pauline, Jarst, Jasper, Kim, Katharina, Meng Jin, Yan Zhe, Gurnoor, Ehsan, Ashikin, Mas, Peter, Sara, Sri, Ernest, Jordi, Viviana.

My "Paranymphs", **Carolina** and **Ying** also deserve to be separately thanked. Thanks for accepting this responsibility and helping me with all the submission stuff and ceremony preparation.

دوستان عزیز همزبان و هم وطنم: **رباب، خرارد و تونی**(نیمه هم وطن :))، **نفیسه و آرمان، نرگس و مهدی، ثمین و سیاوش، شقایق و جلال،** به خاطر همه لحظات خوبی که در این چند سال در کنارشما داشتم و باعث شد که غم غربت رو کمتر احساس کنم ازتون ممنونم. جا داره که تشکر کنم از همتون به خاطر همه مشاوره هایی که بهم دادین در زمانی که تو کارم مشکل پیش میومد. امیدوارم هر جا که هستید موفق و شاد و سلامت باشید.

جا داره که همینجا از خانواده عزیزم تشکر کنم به خاطر همه حمایت های مالی و معنویشون توای چند سال. میدونم که غم دوری همینقدر که برای من سخت بود به همین میزان هم به اونها سخت گذشت. بخصوص باید از خواهرم تشکر کنم که تو این مدت به عنوان وکیل مسئولیت کارای اداری و بانکیم برعهدش بود. خوشحالم و خدا رو شکر میکنم که شما رو در کنار خودم دارم.

Last but not least! I would like to dedicate this thesis to my mother for her endless support and encouragement and especially for her key sentence to me whenever I was talking to her during these five years: **"Everything is going to be alright, be strong!"**

Education Statement of the Graduate School

Experimental Plant Sciences
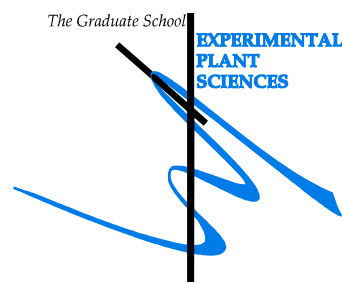
Issued to: **Raana Roohanitaziani**
Date: **31 October 2019**
Group: **Plant Breeding**
University: **Wageningen University & Research**

1) Start-Up Phase

	date	*cp*
► **First presentation of your project**		
Variation of lycopene, chlorophyll and weight loss, during postharvest storage of tomato core collection	30 Sep 2014	1.5
► **Writing or rewriting a project proposal**		
Genetic analysis of fruit quality traits in tomato	01 Dec 2014	6.0
► **Writing a review or book chapter**		
► **MSc courses**		
Subtotal Start-Up Phase		7.5

2) Scientific Exposure

	date	*cp*
► **EPS PhD student days**		
EPS PhD student days 'Get2Gether', Soest, NL	28-29 Jan 2016	0.6
EPS PhD student days 'Get2Gether', Soest, NL	9-10 Feb 2017	0.6
► **EPS theme symposia**		
EPS Theme 3 Symposium 'Metabolism and Adaptation', Wageningen, NL	10 Feb 2015	0.3
EPS Theme 1 Symposium 'Developmental Biology of Plants, Wageningen, NL	21 Jan 2016	0.3
EPS Theme 3 Symposium 'Metabolism and Adaptation, Wageningen, NL	17 Mar 2017	0.3
EPS Theme 1 Symposium 'Developmental Biology of Plants, Wageningen, NL	29 Jan 2018	0.3
EPS Theme 3 Symposium 'Metabolism and Adaptation', Wageningen, NL	13 Mar 2018	0.3
► **Lunteren Days and other national platforms**		
Annual meeting 'Experimental plant sciences', Lunteren, NL	10-11 April 2017	0.6
Annual meeting 'Experimental plant sciences', Lunteren, NL	9-10 April 2018	0.6
► **Seminars (series), workshops and symposia**		
Plant Breeding Day	24 Sep 2014	0.3
EPS flying seminar:PIN-ning down the function of D6 PROTEIN KINASES in auxin transport'	21 Oct 2014	0.1
Symposium Omics advances for academia and industry, Wageningen University	11 Dec 2014	0.3
EPS flying seminar:Seasonal flowering in annual and perennial plants	19 Jan 2015	0.1
Plant Breeding Day	29 Sep 2015	0.3
EPS flying seminar: functional analysis of circadium long non-coding RNAs by dr Rossana Henrigues	11 Nov 2015	0.1
EPS flying seminar: polyploidy in wild relatives of soybean and other legumes: systematics, comparative and functional genomics	12 May 2015	0.1

EPS flying seminar: Membrane traffic and fusion in plant cytokinesis - an evolutionary perspective by Prof Jürgen	11 May 2016	0.1
EPS flying seminar:Plant intracellular immunity: evolutionary and molecular underpinnings	21 Jan 2016	0.1
Wur Omics symposium	15 Dec 2016	0.3
Seminar: Genetic and Genomic approach of fruit quality and adaptation to stress in tomato	12 Sep 2017	0.1
Seminar: Recombination in crop plants	16 Oct 2017	0.1

► **Seminar plus**

► **International symposia and congresses**

International Plant Phenotyping Symposium, Barcelona (Spain)	11-12 Nov 2015	0.6
XIV Solanacea & 3rd cucurbitacea conference, Valencia (Spain)	03-06 Sep 2017	1.1
XIX EUCARPIA Meeting of the Tomato Working Group, Naples (Italy)	02-04 May 2018	0.9

► **Presentations**

Poster: International Plant Phenotyping Symposium, Barcelona, Spain	11-12 Nov 2015	1.0
Oral: XIV Solanacea & 3rd cucurbitacea conference, Valencia, Spain	04 Sep 2017	1.0
Poster : Annual meeting 'Experimental plant sciences', Lunteren, NL	10-11 Apr 2017	1.0
Oral: TKI tomato meeting, Wageningen, NL	07 Dec 2016	1.0
Oral: TKI tomato meeting, Wageningen, NL	09 Nov 2017	1.0
Oral: Annual meeting 'Experimental plant sciences', Lunteren, NL	09 Apr 2018	1.0
Oral: XIX EUCARPIA Meeting of the Tomato Working Group, Naples, Italy	02 May 2018	1.0

► **IAB interview**

► **Excursions**

Company visit Genetwister and In2care in wageningen	19 Sep 2014	0.3
Company visiting ENZA zaden in Enkhuizen	12 Jun 2015	0.2
Subtotal Scientific Exposure		16.0

3) In-Depth Studies

	date	cp

► **Advanced scientific courses & workshops**

Course 'Bioinformatics - A User's Approach', Wageningen, NL	24-28 Aug 2015	1.5
Course 'Genome assembly', Wageningen, NL	28-29 Apr 2015	0.6
Summer School 'Image Analysis for Plant Phenotyping', Wageningen, NL	12-15 Jul 2016	1.0
Course 'Data analyses and visualization in R', Wageningen, NL	11-12 May 2017	0.6

► **Journal club**

Literature discussion group at Plant Breeding	2014-2018	3.0

► **Individual research training**

Subtotal In-Depth Studies		6.7

4) Personal Development

	date	cp

► **General skill training courses**

EPS PhD introduction course	20 Jan 2015	0.2
Course 'Information literacy PhD including EndNote Introduction', Wageningen, NL	09-10 Jun 2015	0.6
Course 'Effective Writing Strategies', Wageningen, NL	Apr - May 2016	1.3
Course 'Project and Time Management', Wageningen, NL	Sep - Nov 2016	1.5
Course 'Effective behaviour in your professional surrondings', Wageningen, NL	May - Jun 2016	1.3

► **Organisation of meetings, PhD courses or outreach activities**

► **Membership of EPS PhD Council**

Subtotal Personal Development	5.2

TOTAL NUMBER OF CREDIT POINTS*	**35.4**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS with a minimum total of 30 ECTS credits.

* A credit represents a normative study load of 28 hours of study.