MDPI

Review

The Genetic Basis of Tomato Aroma

Matteo Martina [1] , Yury Tikunov [2] , Ezio Portis [1],* and Arnaud G. Bovy [2],*

[1] DISAFA, Plant Genetics and Breeding, University of Turin, 10095 Grugliasco, Italy; matteo.martina@unito.it
[2] Plant Breeding, Wageningen University & Research, P.O. Box 386, 6700 AJ Wageningen, The Netherlands; yury.tikunov@wur.nl
* Correspondence: ezio.portis@unito.it (E.P.); arnaud.bovy@wur.nl (A.G.B.);
Tel.: +39-011-6708807 (E.P.); +31-317-480762 (A.G.B.)

Abstract: Tomato (*Solanum lycopersicum* L.) aroma is determined by the interaction of volatile compounds (VOCs) released by the tomato fruits with receptors in the nose, leading to a sensorial impression, such as "sweet", "smoky", or "fruity" aroma. Of the more than 400 VOCs released by tomato fruits, 21 have been reported as main contributors to the perceived tomato aroma. These VOCs can be grouped in five clusters, according to their biosynthetic origins. In the last decades, a vast array of scientific studies has investigated the genetic component of tomato aroma in modern tomato cultivars and their relatives. In this paper we aim to collect, compare, integrate and summarize the available literature on flavour-related QTLs in tomato. Three hundred and fifty nine (359) QTLs associated with tomato fruit VOCs were physically mapped on the genome and investigated for the presence of potential candidate genes. This review makes it possible to (i) pinpoint potential donors described in literature for specific traits, (ii) highlight important QTL regions by combining information from different populations, and (iii) pinpoint potential candidate genes. This overview aims to be a valuable resource for researchers aiming to elucidate the genetics underlying tomato flavour and for breeders who aim to improve tomato aroma.

Keywords: tomato; *S. lycopersicum*; aroma; volatiles; QTLs

1. Introduction

Tomato (*Solanum lycopersicum* L.) is one of the most important crops on the market, used worldwide as basis in many national traditional dishes [1]. Conventional wisdom suggests that breeding tends to reduce the genetic basis of a cultivated species, but tomato genetic diversity appears to have actually been enhanced in the last fifty years. A recent study [2] investigated the genetic diversity of cultivated tomato varieties in The Netherlands (NL) from 1950s to 2010s, observing that tomato commercial varieties from 1950s and 1960s were mainly homozygous, with narrow genetic variation among them. From 1970s onwards, genetic diversity in tomato has increased, thanks to the application of introgression breeding programs using wild relatives of tomato. The first genetic diversity boost appeared to take place with the introgression of tomato mosaic virus (ToMV), southern root-knot nematode (*Meloidogyne incognita*) and leaf mold disease (*Cladosporium fulvum*) resistance from *Solanum peruvianum* and *Solanum pimpinellifolium* [3–5]. These introgressions varied in size from ~5% of the chromosome (introgression of Cf-4 and Cf-9 on top of Chr.01 for resistance to *Cladosporium fulvum*) up to half a chromosome 9 (ToMV), and together led to a significant increase in overall genetic diversity [2]. A second diversity boost, starting in the late 1980s affected both fruit size and quality traits: in particular the introgression of parts of Chr. 4, 5 and 12 from *S. pimpinellifolium* led to fruit size variation among cultivated varieties and the introduction of cherry, cocktail and large fruited varieties to the NL market [3].

Not only disease resistance and fruit size, but also flavour has been targeted by breeding programs in the last thirty years [2]. The main components of tomato fruit

flavour are: (i) sweetness, mainly determined by sugars; (ii) acidity, determined by the presence of organic acids; (iii) textural attributes, such as firmness and juiciness, and (v) aroma [6]. Tomato aroma is determined by the interaction of volatile compounds released by the tomato fruits with receptors in our nose, leading to a sensorial impression, such as "sweet", "smoky", or "fruity" aroma. More than 400 volatile compounds [7] have been detected in ripe tomato fruit and they are derived from primary metabolites, such as fatty acids and amino acids or secondary metabolites, such as phenylpropanoids and carotenoids. Only a fraction of these volatiles has been associated with specific human sensorial attributes or were shown to be present at concentrations above the so-called odor threshold. From an evolutionary point, these molecules evolved to attract seed dispersers, including animals and men. For humans they have a great economic impact, since they are important determinants of food quality and consumer preference [8,9]. Among them, about 20 compounds have been identified as the most contributing ones, based on their concentrations in fruits and their individual odor thresholds [10]. However, it has to be mentioned that the perception of aroma is determined by both the odor activity of individual VOCs and by interactions between them or with other non-volatile chemicals [11]. For instance, the presence of sugars or organic acids alters the perception of aromatic descriptors of samples with the same concentration of volatiles [12,13], while the perception of basic tastes, e.g., sourness or sweetness, can be modified by variation in VOCs accumulation [8,14,15]. Tomato aroma is therefore a hard-to-define trait and efforts have been made to develop prediction models for its different components [8,12,16,17]. This led to the identification of the main 21 VOCs impacting consumer liking (Table 1). Each of these compounds has its own characteristic odor, as retrieved from the Good Scents Company database [18]. The compounds can be grouped in five clusters, according to their biosynthetic origins [11]: (1) fatty acids-derived VOCs; (2) sulphur-containing and branched chain amino acids-derived VOCs; (3) carotenoid catabolism by-product phenolic VOCs; (4) phenolic VOCs; and (5) phenylpropanoid VOCs [19–25].

Table 1. Overview of the 21 main components of tomato aroma (adapted from [11]). * Odor descriptors were retrieved from the Good Scents Company database [18].

Class	Compound	Odor Descriptors *
Apocarotenoid	6-Methyl-5-hepten-2-one	Citrus, green, musty, lemongrass, apple
Apocarotenoid	Geranial	Sharp, lemon, sweet
Apocarotenoid	β-Damascenone	Apple, rose, honey, tobacco, sweet
Apocarotenoid	Geranylacetone	Fresh, green, fruity, waxy, rose, woody, magnolia, tropical
Apocarotenoid	β-Ionone	Floral, woody, sweet, fruity, berry, tropical, beeswax
BCAA	3-Methylbutanal	Ethereal, aldehydic, chocolate, peach, fatty
BCAA	2-Methylbutanal	Musty, cocoa, coffee, nutty
BCAA	3-Methylbutanol	Fusel, oil, alcoholic, whiskey, fruity, banana
BCAA	2-Isobuthylthiazole	Green, wasabi, privet, tomato, leaf, earthy, vegetable, metallic
Lipid	1-Penten-3-one	Pungent, peppery, mustard, garlic, onion
Lipid	(Z)-3-Hexenal	Green, fatty, grassy, weedy, fruity, apple
Lipid	Hexenal	Sweet, almond, fruity, green, leafy, apple, plum, vegetable
Lipid	(E)-2-Hexenal	Sharp, fresh, leafy, green, clean, fruity, spicy, herbal
Lipid	(E)-2-Heptenal	Pungent, green, vegetable, fresh, fatty
Lipid	(E,E)-2.4-Decadienal	Oily, cucumber, melon, citrus, pumpkin, nut, meat
Phenolic	Phenylacetaldehyde	Green, sweet, floral, hyacinth, clover, honey, cocoa
Phenolic	2-Phenylethanol	Sweet, floral, fresh, bready, rose, honey
Phenolic	1-Nitro-2-phenylethane	Flower, spice
Phenylpropanoid	Guaiacol	Phenolic, smoke, spice, vanilla, woody
Phenylpropanoid	Methyl salicylate	Wintergreen, mint
Phenylpropanoid	Eugenol	Sweet, spicy, clove, woody

Breeders and breeding researchers aim to elucidate the genetic basis underlying important agronomic traits. This requires three essential elements: (1) genetic variation, (2) phenotypic variation and (3) methods to find associations between the genetic and

phenotypic data. As outlined in more detail below, advances in the development of high-throughput (HTP) molecular marker platforms, the availability of genomic information and progress in phenotyping methodologies, such as metabolomics, have led to an increasing importance and use of marker-assisted breeding strategies, such as QTL analysis, association mapping and genomic prediction in current plant breeding practice and transformed plant breeding into a hight-tech industry.

Marker-assisted introgression breeding: Tomato is highly autogamous, a characteristic that, together with the loss of many genes and alleles during domestication and crop improvement, led to a narrow genetic basis of cultivated tomato, compared with its 12 wild relative species [26]. For this reason, wild relatives have been used as potential sources of lost alleles in the development of new cultivated tomato varieties [27,28]. Marker-assisted selection in plant breeding programs relies on genetic linkage analyses, which are based on the principle of genetic recombination during meiosis. This allows the construction of linkage maps, composed of genetic markers linked to genes or Quantitative Trait Loci (QTL's) affecting traits of interest for a specific population. QTL analysis is mostly done using biparental segregating populations based on a cross of two contrasting genotypes. In order to discover and elucidate the genetic basis of agricultural traits, segregating populations have not only been made from intraspecific crosses, but also from interspecific crosses with various tomato wild relatives (*S. pimpinellifolium, S. pennellii, S. lycopersicoides* and *S. habrochaites*) in which genomic regions of the wild donor have been introgressed in the cultivated tomato genetic background, allowing the identification of potential new alleles for traits of interest [29–33]. Interestingly, researchers not only investigated *Solanum lycopersicum* wild relatives, but also *Solanum lycopersicum* var. cerasiforme L, the expected ancestor of the domesticated tomato [34].

The impact of genomics: In the last decades, the advances in genomics have provided new tools for discovering and tagging novel alleles and genes. The advent of next generation sequencing (NGS) techniques has considerably accelerated and simplified the genome-wide detection of single-nucleotide polymorphisms (SNPs), which have become the most popular molecular markers. The development of the reference tomato genome from the inbred cultivar "Heinz 1706" [35] represented a milestone in the genomic era. The comparison of the cultivated tomato genome with the genome of a wild relative *S. pimpinellifolium* revealed the potential of high-throughput sequencing in comparative genetics, confirming the previously reported introgression of *S. pimpinellifolium* in the "Heinz 1706" genome [36] and the identification of thousands of SNPs between the two relatives. Further genome and transcriptome resequencing aimed at detecting genetic variation in tomato paved the way for the development of relatively universal genotyping platforms (i.e., SNP arrays) that can be applied for the genetic analyses of different populations—the SolCap SNParray [37,38] and the CBSG array [3]. A further progress in the application of NGS is represented by the genotyping-by-sequencing (GBS) approaches, based on the use of restriction enzymes to decrease genome complexity before sequencing. These techniques include over a dozen of reduced-representation sequencing (RRS) approaches [39] and have been recently applied for high-resolution QTL mapping in tomato, especially in interspecific crosses [40–43]. These protocols have been applied in the development of high-density genetic maps in many different species [44–55], making it possible to perform comparative and quantitative genetics in virtually any genomic background. The current advances in genome sequencing technologies, such as a revolutionary increase of sequencing throughput and concomitant reduction in costs per sequenced nucleotide, allowed to unravel the genetic variation in tomato to its full extent [25,56,57] and has helped to realize that the concept of one or a few reference genomes is not sufficient to fully understand the genetic control of traits. For this reason, the pangenome concept has been introduced in plant genomics, investigating the sum of genes that can be found in a specific species [58–63]. Specifically, the pangenome is defined as "the full complement of genes of a species, which can be partitioned into a set of core genes that are shared by all individuals and a set of dispensable genes that are partially shared or individual specific" [64]. In the Solanaceae family, the pangenomes of tomato and

pepper (*Capsicum* spp.) have been recently released, identifying missing genes involved in resistance mechanisms and quality traits [26,65,66], but eggplant (*Solanum melongena* L.) and potato (*Solanum tuberosum* L.) are still pangenome-orphan species. The availability of the tomato pangenome allowed the identification of presence/absence variations (PAVs) and the identification of structural variants (SVs) of functionally important genes [26,65]. Interestingly, [26] identified a rare promoter allele for the TomLoxC gene, a lipoxygenase that has been reported to be crucial in C5 and C6 lipid-derived volatiles biosynthesis and apocarotenoid production [67,68]. Moreover, the newly developed tomato pangenome [65] resolved the genomic region flanking *NSGT1*, a functional gene associated with the production of guaiacol, methylsalicylate and eugenol, three phenylalanine-derived volatiles. Five different haplotypes were identified in the analyzed tomato germplasm, providing new insights in the understanding of the genetic variation of *NSGT1*. Pangenome appears to be the most novel tool that breeders and researchers have at their disposal: it may facilitate the mining of natural genetic variation and could contribute to crop improvement by supporting molecular breeding programs and gene function studies.

Linking genetic markers to traits: In the genomics era, novel genotyping techniques and the availability of multitudes of molecular markers, in combination with new high-throughput phenotyping technologies (i.e., phenotyping platforms), supported the development of new methodologies to link genetic markers to phenotypic traits. Not only the above-mentioned QTL mapping has become more precise, also association mapping and genomic prediction are now widely used in breeding [69–72]. Genome-Wide Association Studies (GWASs) are based on genotyping of a set of accessions representing the variability in a given species and rely on the linkage disequilibrium (LD) between a marker and its associated trait [73]. This technique has been applied in tomato, identifying interesting associations for many fruit quality traits [74–79]. Zhao et al. performed a meta-GWAS analysis, by combining datasets of several GWAS panels. This analysis not only led to confirmation of existing, but also to the discovery of novel QTLs and candidate genes for several flavour-related traits [77]. Genomic Prediction (GP) is a selection tool that makes use of genetic markers to predict the genetic potential of untested lines in breeding [80]. While QTL mapping and GWAS rely on the (statistically-significant) association between phenotypic variation and specific molecular markers, GP calculates the genetic potential of breeding candidates by the application of Bayesian or mixed statistical models that take all the genome-wide marker information into account to predict the phenotype. Genomic prediction is a selection tool rather than a research tool and performs better with traits that are controlled by a large number of small-effect QTLs which are hard to detect by QTL analysis of mapping populations or GWAS [81,82]. Genomic prediction is particularly useful for traits for which phenotyping is expensive, difficult or time consuming, since no phenotyping is needed for selection, once a good prediction model based on data of a representative training population is available. This technique has been widely applied in animal selection [83–85], while its practical application in plant breeding is still limited to major crops, such as maize and wheat [82,86,87], in which QTLs for important traits, such as yield, have already been fixed in the elite germplasm [88], or to tree crops, where early selection is very useful and cost-effective [89].

The combination of all the above-mentioned approaches, except GP, led to identification of a multitude of QTLs for many agronomic and quality traits. In this review we aim to collect, compare, integrate and summarize the available literature on flavour-related QTLs in tomato. We selected 16 scientific papers and supplemental data focusing on QTLs for tomato aroma and fruit quality. This not only provides an overview of the known flavour QTLs, but the combined and integrated information also makes it possible to (i) pinpoint potential donors described in literature for specific traits, (ii) highlight important QTL regions by combining information from different populations, and (iii) pinpoint potential candidate genes. This overview aims to be a valuable resource for researchers aiming to elucidate the genetics underlying tomato flavour and for breeders who aim to improve tomato aroma.

2. Construction of a Unified QTL Map of Tomato Aroma

A literature search was performed with the aim of collecting articles reporting QTLs for tomato aroma. In order to compare and integrate QTLs from different studies, the availability of marker information was an essential requirement for inclusion in this review. The identified QTLs have been organized and are available as Supplemental Material (Table S5), reporting the biosynthetic pathway, the QTLs Genomic Regions (QGR), the QTL's original name, the related compound, the chromosome, the correlated markers, their position in cM and bp, their *p*-value, their LOD score, the SolycID of the gene in which the markers have been found, the percentage of explained variation by the QTL, the effect, the donor parent, the crossing population or association panel used and the reference of the primary resource.

We identified 16 articles reporting QTLs for tomato aroma including marker information (Table 2). In the pre-genomics era only genetic linkage positions (cM) of QTLs could be reported, since a reference genome sequence was not available the. This made it difficult to align QTLs of that period with the more recent studies utilizing modern genomics technologies. To circumvent this, the physical position of genetic markers was retrieved from the tomato genome (https://solgenomics.net/search/markers (accessed on 20 November 2020)) whenever marker sequence information was available, using the SL2.50 genome version.

Table 2. List of the articles reporting QTLs for tomato aroma included in the review. RILs: Recombinant Inbred Lines; BCs: Backcross population; ILs: Introgression Lines; GWAS: Genome-Wide Association Studies; DP: Diversity Panel.

Author, Year	Population	References
Saliba-Colombani et al., 2001	RILs	[90]
Fulton et al., 2002	BCs	[91]
Tadmor et al., 2002	ILs	[92]
Tieman et al., 2006	ILs	[93]
Mathieu et al., 2009	ILs	[6]
Kochevenko & Fernie, 2011	ILs	[94]
Ruggieri et al., 2014	GWAS	[95]
Sauvage et al., 2014	GWAS	[74]
Capel et al., 2015	RILs	[96]
Zhao et al., 2016	GWAS	[97]
Baldina et al., 2016	GWAS	[98]
Bauchet et al., 2017	GWAS	[99]
Tieman et al., 2017	GWAS	[25]
Garbowicz et al., 2018	ILs	[100]
Kimbara et al., 2018	RILs	[101]
Tikunov et al., 2020	F2, F6, DP	[102]

Two criteria were used to cluster QTL information from the different studies into QTLs Genomic Regions (QGRs) in a unified physical map (Figure 1): (i) a biochemical relationship between aroma volatiles and their possible precursors, such as similar chemical structure or common biosynthetic pathway and (ii) overlapping of the QTLs. Since most of the studies only reported the most significant marker(s) of an identified QTL, while information on genetic confidence intervals or (average) linkage disequilibrium (LD) decay was lacking, it was virtually impossible to determine the size of QTL regions. For this reason, we standardized the potential positional error across the reported studies by setting an empirically defined window of ±2.5 Mb around each identified QTL. This window was derived from the average inter-study standard deviation of QTL positions of the three most functionally explored tomato aroma loci—floral aroma on chromosome 4 [25,90,99,102], smoky aroma on chromosome 9 [25,90,102] and the malodorous locus on chromosome 8 [92,93,99]. The major genetic factors underlying these QTLs were identified and, therefore, the dispersion of QTL positions in these different studies could most likely be attributed

to non-genetic sources of variation. This window as well as resulting QGRs only serve as means to classify QTLs and indicate the possibility that the individual QTLs they harbor may be affected by one or a few co-localized genetic factors.
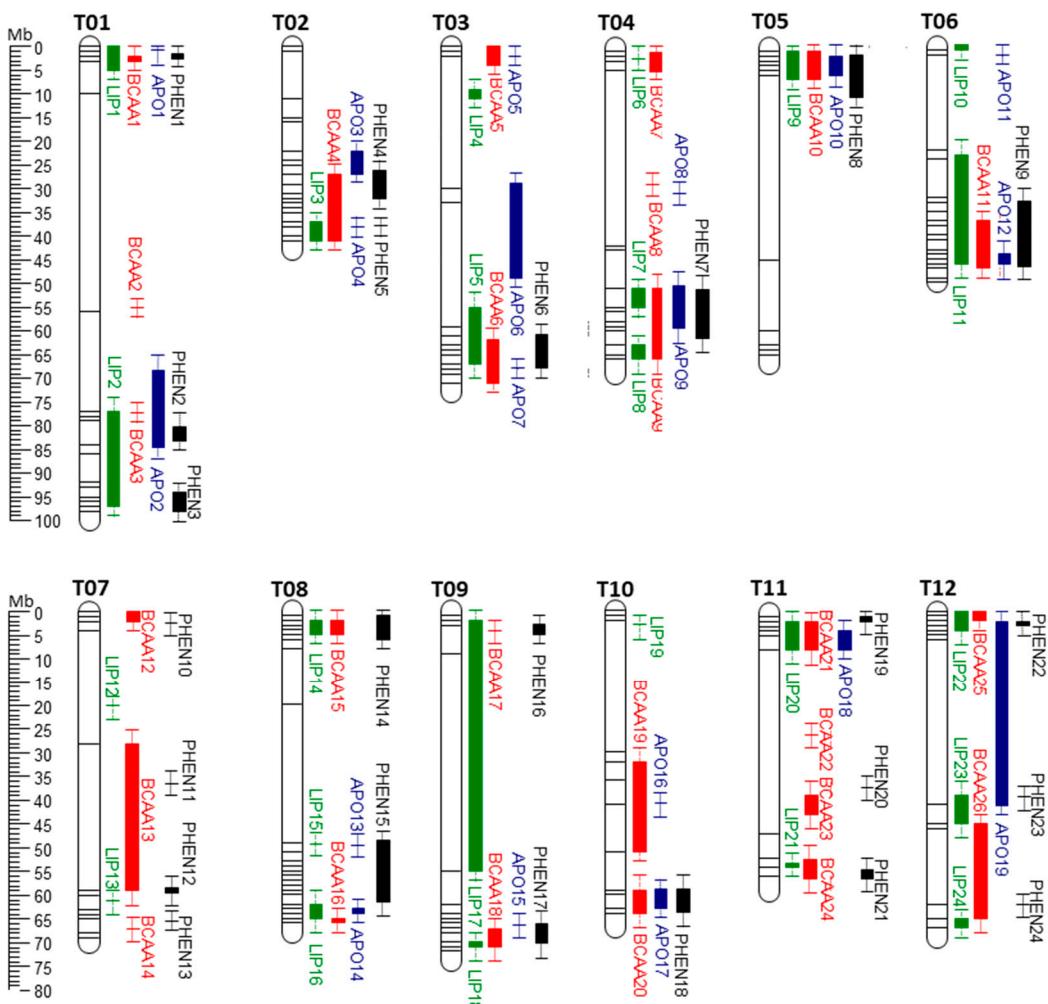
Figure 1. VOCs QGRs chromosome map (T01–T12) obtained from the reported literature. Green: Lipid VOCs QGRs; Red: BCAA VOCs' QGRs; Blue: Apocarotenoid VOCs QGRs; Black: phenolic VOCs QGRs. The estimated confidence interval (±2.5 Mb), has been reported in the figure.

The identified QGRs were mined for the presence of candidate genes based on their annotation (ITAG2.40) and on their expression in tomato fruit, using publicly available dataset and tools [103]. Candidate genes were defined by two criteria: (1) genes belonging to families known from literature to be involved in VOCs biosynthesis and expressed in tomato fruit and (2) genes reported in literature with a demonstrated function in VOCs biosynthesis, irrespective of their expression in tomato fruit. A complete list of the potential candidate genes (with and without expression in the fruit tissues) can be found as Supplemental Material (Tables S1–S4). The genes which have been demonstrated to functionally underlie aroma QTLs in tomato were highlighted in bold.

3. Fatty Acids Derived Volatiles (FA VOCs)

The volatile compounds originating from the degradation of linolenic and linoleic acid accumulate during tomato ripening and can also markedly increase their emission upon fruit tissue disruption. They provide a note of freshly cut grass to the aroma bouquet [104]. These compounds are the most abundant volatiles in tomato fruit and are

mainly represented by the C_5 volatile 1-penten-3-one, a few C_6 volatiles, such as 1-hexanol, (*Z*)-3-hexenal, (*E*)-2-hexenal and hexanal, the C_7 volatile (*E*)-2-heptenal and the C_{10} VOC (*E*,*E*)-2,4-decadienal [8,11,16,17]. Although their high accumulation in ripe fruits may suggest that these compounds are very important determinants of tomato flavour, some studies provide evidence that impact of their quantitative variation on consumer liking may be limited [8,67], likely due to their in general high abundance and low odor thresholds, e.g., for (*Z*)-3-hexenal these values are 12,000 nLL^{-1} and 0.25 nLL^{-1}, respectively [21].

3.1. Biosynthesis of FA VOCs

During tomato fruit ripening, free fatty acids, mainly linolenic and linoleic acid, are derived from the catabolism of acylglycerides from disintegrating cellular membranes, by the action of lipases [87,88]. Linolenic and linoleic acid can be further catabolized by means of β-oxidation, α-oxidation, or the lipoxygenase pathway [68,105–107]. In tomato fruit the latter is the most important for the production of volatiles, which occurs through two steps: (i) fatty acids are deoxygenated by means of lipoxygenases (LOX), which are classified as 13-LOX and 9-LOX and are leading respectively to 13-hydroperoxides and 9-hydroperoxides [67,108]; (ii) hydroperoxides are catabolized by means of hydroperoxide lyases (HPL), also classified as 13-HPL and 9-HPL, leading to an oxoacid and a volatile aldehyde. Volatile aldehydes can be converted into alcohols by means of alcohol dehydrogenases (ADH; [109–114]). According to the literature ([68]), 13-LOX enzymes are mainly involved in the synthesis of (*Z*)-3-hexenal from linolenic acid and hexanal from linoleic acid. Among these enzymes, TomloxC has shown significant correlation with the production of hexanal, together with LeHPL, a 13-HPL [67,68,115]. Another gene, *ADH2*, has been reported as positively related to the production of hexanol and (*Z*)-3-hexenol [116] while ADH1 showed in vitro activity in the conversion of hexanal into hexanol [117]. Figure 2 summarizes the complete biosynthetic pathway of the lipid-derived VOCs, according to the available literature.
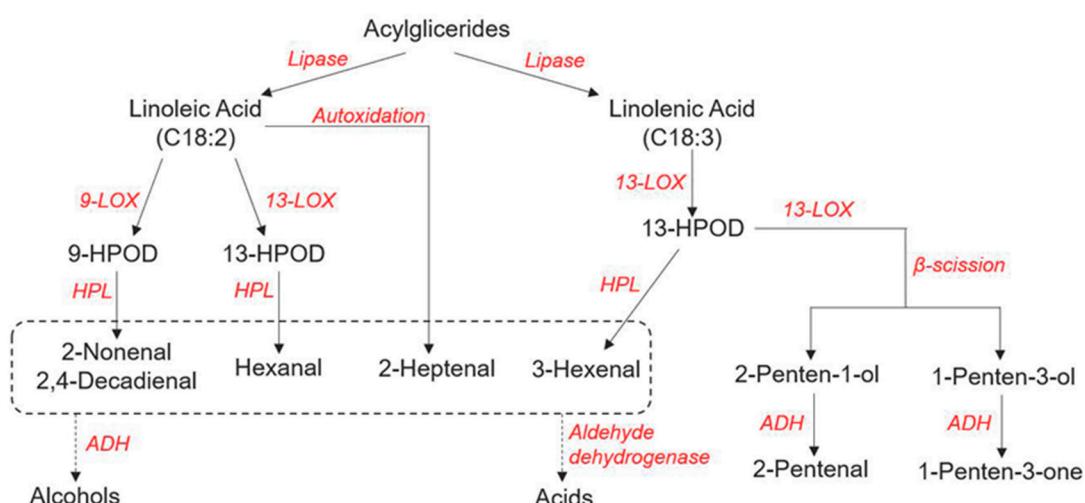

Figure 2. Overview of the lipid VOCs pathway, adapted from [118–120].

3.2. QTLs for FA VOCs

Data collection identified a total number of 108 QTLs reported in 8 different studies and correlated with lipid volatiles biosynthesis (Table S1). Comparing these regions (see "Data Acquisition and Classification"), we identified 24 distinct QGRs (Figure 1, Table S1).

3.3. FA VOCs' Candidate Genes

The collected data allowed the identification of 112 genes potentially involved in lipid VOCs biosynthesis (Table S1). Among them, 28 genes have been reported to be expressed in

at least one fruit tissue (Table 3; [103]), or have been functionally characterized for their role in fatty acids metabolism. Eight lipoxygenases (LOXs) were identified as candidate genes expressed in tomato fruit. Among the identified LOXs, three have been reported for their association with lipid VOCs: *LoxC* (Solyc01g006540) has a major role in the biosynthesis of the most quantitatively prominent C5 and C6 lipid-derived VOCs, *LoxF* (Solyc01g006560) is involved in the production of fatty acid VOCs derived from 13-hydroperoxides [3,121].

Table 3. List of the genes identified as potentially responsible for the lipid VOCs QTLs reported in literature. Genes that have been functionally characterized are highlighted in bold in the table. The last column reports whether the gene has been reported to be expressed in fruit (Y/N—[103]), for a complete overview of the gene expression see Table S1.

Solyc ID	Name	QGR	Start (bp)	Annotation	Reference	Expressed (Y/N)
Solyc01g006540	*loxC*	**LIP1**	**1,113,718**	**Lipoxygenase**	[68]	**Y**
Solyc01g006560	*loxF*	**LIP1**	**1,128,815**	**Lipoxygenase**	[121]	**N**
Solyc01g099210	-	**LIP1**	89,509,072	Lipoxygenase		Y
Solyc01g108150	-	LIP2	95,505,678	Oxidoreductase zinc-binding dehydrogenase family protein		Y
Solyc02g090930	-	LIP2	52,398,542	Lipase		Y
Solyc03g093360	-	LIP3	54,632,041	Lipoxygenase		Y
Solyc03g095360	-	LIP5	56,412,064	Alcohol dehydrogenase zinc-binding		Y
Solyc03g111550	-	LIP5	62,149,409	GDSL esterase/lipase		Y
Solyc04g010250	-	LIP6	3,577,656	Lipase-like protein		Y
Solyc04g054980	-	LIP7	53,517,056	Lipoxygenase homology domain-containing protein 1		Y
Solyc04g054990	-	LIP7	53,522,447	Lipoxygenase homology domain-containing protein 1		Y
Solyc04g064710	-	LIP7	55,858,388	Alcohol dehydrogenase 2		Y
Solyc05g005480	*EO QR*	LIP9	352,211	Oxidoreductase zinc-binding dehydrogenase		Y
Solyc05g009390	-	LIP9	3,544,526	Lipase-like protein		Y
Solyc06g059740	*ADH2*	**LIP11**	**37,606,747**	**Alcohol dehydrogenase 2**		**Y**
Solyc07g045090	-	LIP13	58,206,449	Alcohol dehydrogenase zinc-binding domain protein	[116]	Y
Solyc08g014000	*LOXA*	LIP14	3,516,113	Lipoxygenase		Y
Solyc09g059030	-	LIP17	53,010,842	Alcohol dehydrogenase zinc-containing		Y
Solyc09g059040	-	LIP17	53,050,984	Alcohol dehydrogenase zinc-containing		Y
Solyc09g091050	-	LIP18	70,389,731	Lipase		Y
Solyc11g010960	-	LIP20	4,006,687	Alcohol dehydrogenase		Y
Solyc11g011330	-	LIP20	4,375,432	Cinnamyl alcohol dehydrogenase		Y
Solyc11g065530	-	LIP21	50,971,492	Lipase (Fragment)		Y
Solyc11g071290	-	LIP21	54,819,572	Alcohol dehydrogenase		Y
Solyc12g010950	-	LIP22	3,827,848	Alcohol dehydrogenase zinc-containing		Y
Solyc12g011040	-	LIP22	3,894,074	Lipoxygenase		Y
Solyc12g096760	-	LIP24	65,539,079	Alcohol dehydrogenase zinc-containing		Y
Solyc12g096780	-	LIP24	65,557,520	Mitochondrial trans-2-enoyl-CoA reductase		Y

Eleven alcohol dehydrogenases (ADH) were identified. ADHs are a family of enzymes associated with the interconversion of the aldehyde and alcohol forms of lipid volatiles in tomato and they have been reported to accumulate in the fruit during ripening [6,116]. Among them, Solyc06g059740 has been characterized in tomato fruit as ADH2 [116,122,123]. Solyc11g071290 was identified in the LIP21 QGR, a QTL for the earthy/mushroom odor type volatile 1-octen-3-one—which, unlike the major C6 VOCs, has rarely been suggested as an important fresh tomato fruit odorant, has a much lower concentration than C6 VOCs, but has an extremely low odor threshold of 0.005 nLL^{-1} [124]. A structural variation in the promoter of this gene was reported in [26]. The wild allele was present in *S. pimpinellifolium* and *S. cerasiforme*, but was not found in heirloom tomatoes, suggesting selection against the wild allele during domestication. Although this gene showed a significant expression level in one fruit sample only (*S. pimpinellifolium* fruit at 4 DPA), we cannot exclude that allelic differences in this gene may be responsible for the 1-octen-3-one QTL identified by GWAS [99].

4. Branched-Chain Amino Acids Derivatives (BCAA VOCs)

Branched-chain amino acid (BCAA) derived compounds are highly volatile compounds with a low molecular weight, some of which are considered important in the development of tomato aroma, e.g., 3- and 2-methylbutanal (aldehydic/chocolate/musty odor types), 3- and 2-methylbutanol (roasted/fermented odor types), and 2-isobutyl-thiazole (green/wasabi odor), which is partly derived from BCAA [10,12,16].

4.1. Biosynthesis of BCAA VOCs

Even though the relationship between BCAA VOCs and the tomato aroma is clear, the exact molecular mechanism underlying their quantitative variation in tomato fruit is not fully understood. In general, BCAA VOCs originate from the branched-chain amino acid (BCAA) pathway in many organisms including plants. BCAA biosynthesis has been studied well in plants and it takes place in the chloroplast where leucine and valine are synthesized from pyruvate and isoleucine from threonine via several enzymatic reactions (Figure 3). Catabolism of BCAAs, which occurs in mitochondria is believed to be the source of BCAA VOCs in tomato fruit. It has been suggested that the first step in the catabolic pathway leading to BCAA VOCs is the reversible conversion of branched-chain amino acids (leucine, isoleucine) into their corresponding α-ketoacids by means of branched-chain amino acid aminotransferases (BCATs; [125]). The importance of BCATs in the degradation of BCAAs has been demonstrated in *Arabidopsis* [126]. In tomato different members of the BCAT family have been shown to mediate either synthetic or catabolic reactions of BCAAs [94,127–129]. The products of the reversible BCAT-mediated BCAA deamination—α-ketoacids—have been suggested to be the likely precursors for BCAA VOCs [128], which then could be produced through the combined action of various classes of candidate enzymes (Figure 3), such as α-ketoacid dehydrogenases, decarboxylases and alcohol/aldehyde dehydrogenases [130].
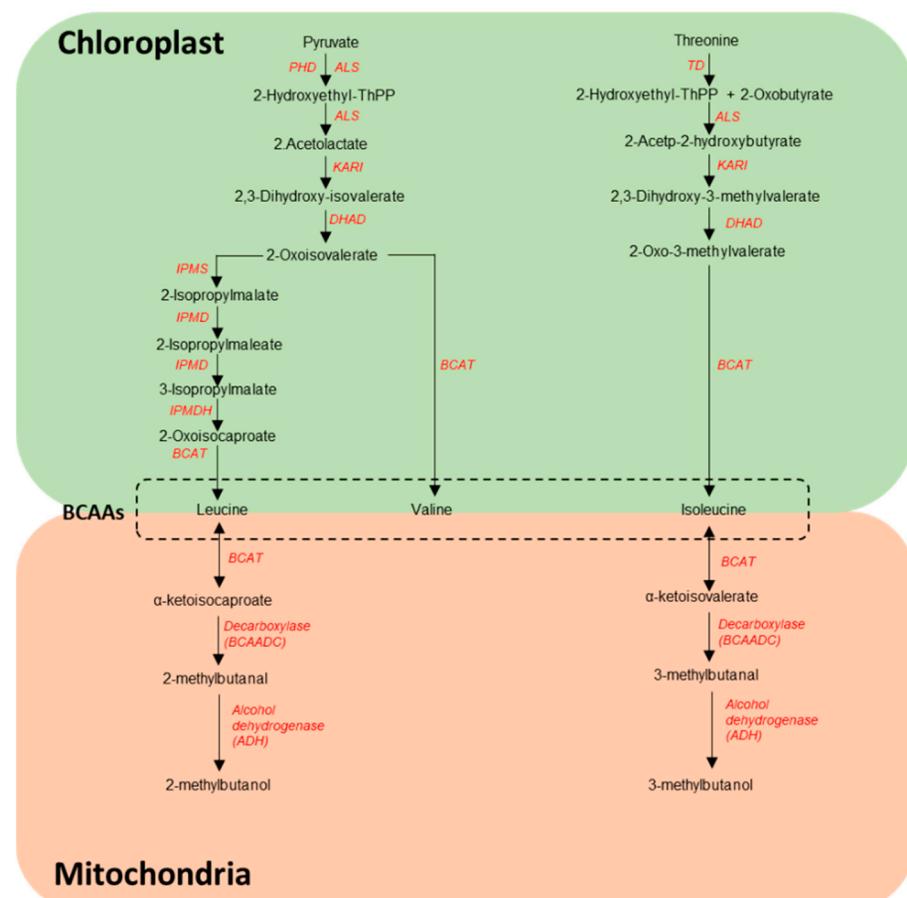
Figure 3. Overview of the proposed BCAA VOCs pathway, adapted from [94,130]. TD: threonine dehydratase; ALS: acetolactate synthase; KARI: ketol-acid reductoisomerase; DHAD: dihydroxy-acid dehydratase; BCAT: branched-chain aminotransferase; IPMS: 2-isopropylmalate synthase; IPMD: 3-isopropylmalate dehydratase; IPMDH: 3-isopropylmalate dehydrogenase; PDH: pyruvate dehydrogenase.

4.2. QTLs for BCAA VOCs

Data collection identified a total number of 129 QTLs reported by seven different authors and correlated with BCAA volatiles biosynthesis (Figure 1; Table S2). These QTLs were classified into 26 distinct QGRs (Table S2).

4.3. BCAA VOC Candidate Genes

Table S2 reveals 75 genes that were identified as potentially involved in the BCAA VOCs metabolism, 28 of which have been reported to be expressed in at least one fruit tissue [103]. An overview of the 30 selected candidate genes is shown in Table 4, including genes expressed in fruits plus two functionally characterized genes (*BCAT2* and *BCAT4*; see below).

Among the genes identified as expressed in tomato fruit, ten belong to enzymatic families that have been associated with BCAA biosynthesis: four genes were annotated as pyruvate dehydrogenases (PDH), three as 3-isopropylmalate dehydratases (IPMD), two as ketol-acid reductoisomerase (KARI) and one as 2-isopropylmalate synthase (IPMS). Reverse genetics analysis of Solyc06g060790 (*IPMD*) revealed that this gene influences the BCAA content in tomato fruit, while a similar analysis failed to support such a conclusion in case of Solyc07g053280 (*KARI*) [94]. Branched chain amino acid aminotransferases (BCATs) can be involved in the last step of BCAA anabolism and/or in the first step of BCAA catabolism depending on their subcellular localization (chloroplast or mitochondrion, respectively). Six BCAT genes were reported as candidate genes in Table 4, of which five (*BCAT1, 2, 3, 4* and *7*) have been functionally characterized [112,114]. Although *BCAT2, 4* and *7* were located outside the QTL intervals, they were included in this review for completeness. *BCAT1* and *BCAT2* were shown to be located in mitochondria and involved in the catabolism of BCAAs, while *BCAT3* and *BCAT4* were shown to be located in chloroplasts and involved in BCAA biosynthesis. In turning and ripe fruits of cv. M82 *BCAT1* expression was up to 10-fold higher compared to *BCAT2, 3* and *4*. Although *BCAT2, 3* and *4* showed low, but detectable expression in fruits, their expression was much higher in leaves (*BCAT2* and *3*) or inflorescence (*BCAT4*) in this tomato background [127]. According to the RNAseq data present in the TomExpress database [103], these three genes are hardly (*BCAT3*) or not at all expressed in tomato fruits (*BCAT2* and *4*). Although its subcellular location is unclear, *BCAT7* was proposed to play a role in BCAA degradation [128]. Finally, we identified eleven alcohol dehydrogenases (ADH) expressed in tomato fruit in the BCAA QGRs. These may play a role in BCAA catabolism, although their functional characteristics remain to be demonstrated.

The available tomato pangenome [26] was investigated for the presence of non-reference (non cv. Heinz) promoter regions for the abovementioned genes. Interestingly, a non-reference allele was reported for Solyc04g063350, annotated as 3-methyl-2-oxobutanoate dehydrogenase, an enzymatic class that has been described as involved in the BCAA catabolism, which involves the decarboxylation of branch chain amino acids [94,129,131]. This gene has recently been named *FLORAL4* after both genetic and functional studies revealed that its involvement was not restricted to BCAA catabolism, but this gene also controlled the quantitative variation of floral phenolic-derived VOCs derived from catabolism of the aromatic amino acid phenylalanine (see below; [102]). Among the investigated tomato varieties, the cultivars harboring the "wild" allele showed significantly lower gene expression than the ones presenting the domesticated promoter, suggesting a positive selection for *FLORAL4* expression during tomato domestication [26].

Table 4. List of the genes identified as potentially responsible for the BCAA VOCs QTL reported in literature. Genes that have been functionally characterized are highlighted in bold in the table. The last column reports whether the gene has been reported to be expressed in fruit (Y/N—[103]); for a complete overview of the gene expression see Table S2.

Solyc ID	Name	QGR	Start (bp)	Annotation	Family	References	Expressed (Y/N)
Solyc01g098700	*SlBCAT7*		**89,114,957**	**Branched-chain-amino-acid aminotransferase 7**	**BCAT**	[128]	**Y**
Solyc02g091970	*SlBCAT3*	**BCAA4**	**53,178,240**	**Branched-chain-amino-acid aminotransferase 3**	**BCAT**	[127]	**Y**
Solyc03g005730		BCAA5	509,260	3-isopropylmalate dehydratase large subunit 2	IPMD		Y
Solyc03g007200		BCAA5	1,775,153	Oxidoreductase zinc-containing alcohol dehydrogenase family	ADH		Y
Solyc03g043880	*SlBCAT4*		**7,542,329**	**Branched-chain amino acid aminotransferase 4**	**BCAT**	[127]	**N**
Solyc03g097680		BCAA6	60,009,332	Pyruvate dehydrogenase E1 component subunit β	PDH		Y
Solyc04g008590		BCAA7	2,195,990	Pyruvate dehydrogenase E1 component subunit β	PDH		Y
Solyc04g063350	*FLORAL4*	**BCAA9**	**55,462,543**	**3-methyl-2-oxobutanoate dehydrogenase**		[102]	**Y**
Solyc04g064710		BCAA9	55,858,388	Alcohol dehydrogenase 2	ADH		Y
Solyc05g005480	*EO QR*	BCAA10	352,211	Oxidoreductase zinc-binding dehydrogenase	ADH		Y
Solyc06g059880		BCAA11	37,781,458	Acetolactate synthase	ALS		Y
Solyc06g060790		**BCAA11**	**38,814,018**	**3-isopropylmalate dehydratase small subunit**	**IPMD**	[94]	**N**
Solyc07g021630	*SlBCAT2*		**20,381,557**	**Branched-chain amino acid aminotransferase 2**	**BCAT**	[127]	**Y**
Solyc07g045090		BCAA13	58,206,449	Alcohol dehydrogenase zinc-binding domain protein	ADH		Y
Solyc07g053280		BCAA14	61,749,215	Ketol-acid reductoisomerase	KARI		Y
Solyc07g061940		BCAA14	64,814,098	Acetolactate synthase	ALS		Y
Solyc08g014130		BCAA15	3,972,785	2-isopropylmalate synthase 1	IPMS		Y
Solyc09g008670	*TD*	BCAA17	2,123,840	Threonine ammonia-lyase biosynthetic	TD		Y
Solyc11g010960		BCAA21	4,006,687	Alcohol dehydrogenase	ADH		Y
Solyc11g011330		BCAA21	4,375,432	Cinnamyl alcohol dehydrogenase	ADH		Y
Solyc11g071280	*LOC778238*	BCAA24	54,809,918	Branched-chain amino acid aminotransferase like protein	BCAT		Y
Solyc11g071290		BCAA24	54,819,572	Alcohol dehydrogenase	ADH		Y
Solyc12g005860		BCAA25	490,745	3-isopropylmalate dehydratase large subunit	IPMD		Y
Solyc12g009400		BCAA25	2,682,210	Pyruvate dehydrogenase E1 component α subunit	PDH		Y
Solyc12g009410		BCAA25	2,687,666	Pyruvate dehydrogenase E1 component α subunit	PDH		Y
Solyc12g010840		BCAA25	3,773,065	Ketol-acid reductoisomerase	KARI		Y
Solyc12g010950		BCAA25	3,827,848	Alcohol dehydrogenase zinc-containing	ADH		Y
Solyc12g088220	*SlBCAT1*	BCAA26	63,663,328	Branched-chain-amino-acid aminotransferase 1	BCAT	[127]	Y
Solyc12g096760		BCAA26	65,539,079	Alcohol dehydrogenase zinc-containing	ADH		Y
Solyc12g096780		BCAA26	65,557,520	Mitochondrial trans-2-enoyl-CoA reductase	ADH		Y

5. Carotenoid-Derived VOCs

Carotenoid-derived volatiles (apocarotenoid VOCs) are perceived as "floral" and accumulate in low amounts in tomato fruit [11]. However, due to their low odor thresholds [10], they have been reported to positively correlate with tomato flavour acceptability [15]. There are two groups of carotenoid VOCs that are relevant for tomato aroma [21,132]: (i) cyclic carotenoid volatiles, such as β-ionone and β-damascenone (both floral odor type); and (ii) open-chain carotenoid-derived volatiles, such as 6-methyl-5-hepten-2-one, geranial (both citrus odor type) and geranylacetone (floral odor type).

5.1. Biosynthesis

Apocarotenoid volatiles are produced in plastids [104] through the cleavage of carotenoids by carotenoid cleavage dioxygenases, like LeCCD1A and LeCCD1B [133], which are particularly expressed during fruit ripening [132]. This family of genes has been reported in other species to act both on cyclic carotenoids (at the 9′,10′ position) and open-chain carotenoids—at the (5′,6′), (7′,8′), or (9′,10′) positions-, leading to the production of the carotenoid VOCs (Figure 4; [133–139]). Although many structural genes in the carotenoid biosynthetic and cleavage pathway are known, the regulation of the carotenoid pathway is still unclear. In this respect it has been proposed that the loss of membrane integrity during the ripening-dependent conversion of chloroplasts into chromoplasts may be a key mechanism in their regulation, as this process may lead to the release of the carotenoids in the cytoplasm, were they will react with cytoplasmatic cleavage enzymes [15,140]. The available knowledge of the carotenoid biosynthetic and cleavage pathway and the underlying biosynthetic genes (Figure 4) is very helpful for the identification of candidate genes within QTLs for carotenoid-derived VOCs in tomato fruit.
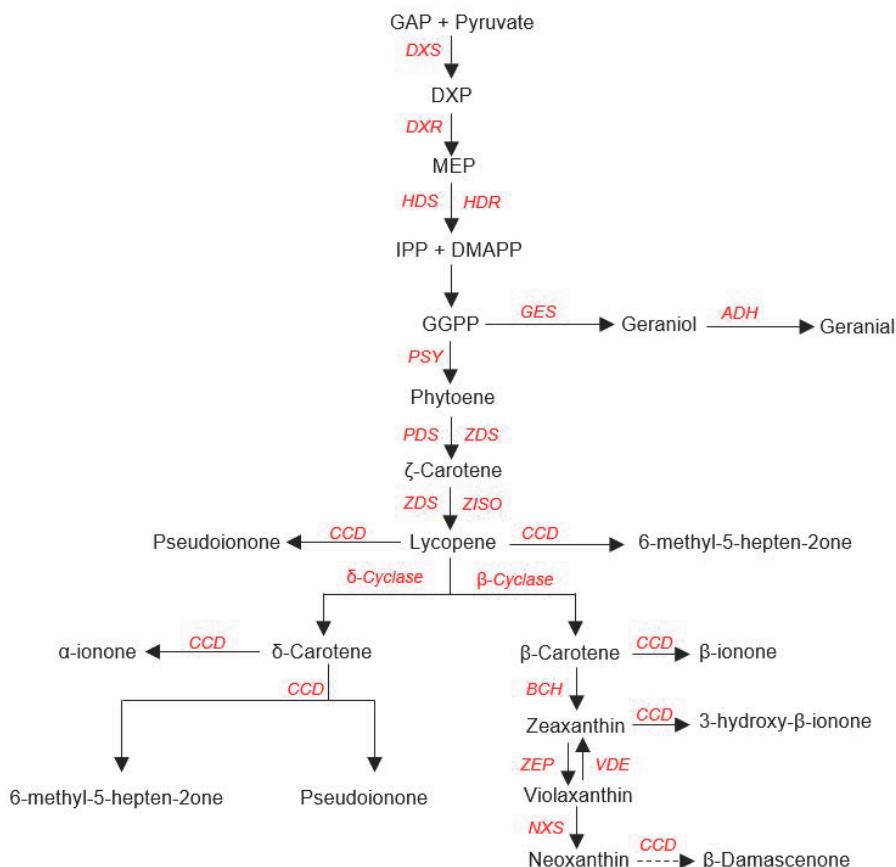
Figure 4. Overview of the carotenoid VOCs pathway, adapted from [135,141,142].

5.2. QTLs for Apocarotenoid VOCs

Data collection identified a total number of 42 QTLs reported in 8 different articles and correlated with apocarotenoid volatiles biosynthesis (Figure 1; Table S3). Comparing these regions (see "Data Acquisition" paragraph for methodology), we identified 19 distinct QGRs (Table S3).

5.3. Apocarotenoid VOCs' Candidate Genes

Twenty-three genes were identified as potentially associated with apocarotenoid VOCs metabolism (Table S3). For nineteen genes there is functional evidence for their involvement in the carotenoid pathway and five of those (Solyc03g007960, Solyc03g031860, Solyc03g123760, Solyc06g036260, Solyc10g081650) have also been shown to be expressed in at least one fruit tissue based on [103]. Solyc06g036260 (Crtz1), Solyc03g114340 (DXR), Solyc03g123760 (PDS), Solyc04g040190 (LCY-B1) are located outside the QTL regions, but were included for completeness (Table 5; [137,143–154]).

Two genes have been annotated as 1-deoxy-D-xylulose 5-phosphate synthases (DXS), reported as first and key regulatory step of the MEP pathway required for the production of carotenoids [141,143,155]. Most of the structural genes of the carotenoid pathway are found within the QTL regions, including three phytoene synthases required for the production of phytoene, the first carotenoid of the pathway [144,149,156–160], followed by phytoene desaturase [152] and one carotenoid isomerase [161], required for the production of lycopene, two lycopene cyclases, involved in the production of α- and β-carotene [153,154,162,163], one carotene hydroxylase to form zeaxanthin [159,164], and finally one zeaxanthin epoxidase, an enzymatic class that converts zeaxanthin, precursor of 3-hydroxy-β-ionone, into violaxanthin [132,135,151,165]. As mentioned above, lycopene β-cyclase 1 (Solyc04g040190) was not associated with any of the apocarotenoid QGRs. In total four carotenoid cleavage dioxygenases (CCDs) were present in the QGRs. Two CCD's were annotated as carotenoid cleavage dioxygenase 1A (CCD1A—Solyc01g087250) and 1b (CCD1B—Solyc01g087260) and have been shown to be directly involved in carotenoid VOCs accumulation in tomato [132,138], one was annotated as carotenoid cleavage dioxygenase 2 (CCD2—Solyc01g087270), reported to be involved in the formation of carotenoid VOCs in Crocus sativus [137], and one as carotenoid cleavage dioxygenase 7 (CCD7—Solyc01g090660), associated with β-carotene cleavage to carlactone [166]. Two CDDs (CCD4A—Solyc08g075480, CCD4B—Solyc08g075490) have been reported to be associated with tomato fruit color, flavour, and aroma [167–169], but were not comprised in any identified QGR.

Phytoene synthase 2 (PSY2—Solyc02g081330) represents an exception to our candidate gene mining approach. This gene has been annotated at 45Mb on chr 2 and is located between QGRs APO3-4. Although this gene is not comprised among the regions identified by this review, its activity has been associated with carotenoid production in fruit in other Solanaceae [170]. For this reason, it has been proposed here as potential candidate gene for tomato apocarotenoid VOCs.

Table 5. List of the genes identified as potentially responsible for the carotenoid VOCs QTL reported in literature. Genes that have been functionally characterized are highlighted in bold in the table. The last column reports whether the gene has been reported to be expressed in fruit (Y/N—[103]); for a complete overview of the gene expression see Table S3.

Solyc ID	Name	QGR	Start (bp)	Annotation	References	Expressed (Y/N)
Solyc01g005940	*PSY3*	APO1	613,955	**Phytoene synthase 3**	[171]	N
Solyc01g067890	*DXS1*	APO2	76,868,469	**1-Deoxy-D-xylulose 5-phosphate synthase 1**	[145,155]	N
Solyc01g087250	*CCD1A*	APO2	82,184,585	**Carotenoid cleavage dioxygenase 1A**	[132,138]	N
Solyc01g087260	*CCD1B*	APO2	82,196,996	**Carotenoid cleavage dioxygenase 1B**	[132,138]	N
Solyc01g087270	*CCD2*	APO2	82,209,237	**Carotenoid cleavage dioxygenase 2**	[137]	N
Solyc01g090660	*CCD7*	APO2	84,307,951	**Carotenoid cleavage dioxygenase 7**	[166]	N
Solyc02g081330	*PSY2*	APO3/4	45,335,358	Phytoene synthase 2		N
Solyc02g090890	*ZEP*	APO4	52,369,708	**Zeaxanthin epoxidase, 2C chloroplastic**	[151]	N
Solyc03g007960	*CrtZ-2*	APO5	2,447,872	**β-Carotene hydroxylase 2**	[164]	Y
Solyc03g031860	*PSY1*	APO5	4,326,134	**Phytoene synthase 1**	[171]	Y
Solyc03g114340	*DXR*		64,347,674	**1-Deoxy-D-xylulose-5-phosphate reductoisomerase**	[143]	N
Solyc03g123760	*PDS*		70,501,011	**Phytoene desaturase**	[152]	Y
Solyc04g040190	*LCY-B1*		11,947,053	**Lycopene β-cyclase 1**	[153]	N
Solyc06g036260	*CrtZ-1*		25,742,578	**β-Carotene hydroxylase 1**	[162]	Y
Solyc06g074240	*LCY-B2*	APO12	45,898,227	**Lycopene β-cyclase 2**	[153]	N
Solyc08g075480	*CCD4A*		59,627,153	**Carotenoid cleavage dioxygenase 4A**	[167]	N
Solyc08g075490	*CCD4B*		59,643,898	**Carotenoid cleavage dioxygenase 4B**	[167]	N
Solyc10g079480	*LCY-b*	APO17	61,024,821	**Lycopene β-cyclase**	[153]	N
Solyc10g081650	*CRTISO*	APO17	62,682,440	**Carotenoid isomerase, 2C chloroplastic**	[161]	Y
Solyc11g010850	*DXS2*	APO18	3,870,455	**1-Deoxy-D-xylulose 5-phosphate synthase 2**	[145,155]	N

A search for reported wild promoter regions [26] was performed for the above-mentioned candidate genes, identifying Solyc03g114340 (*DXR*—1-deoxy-D-xylulose-5-phosphate reductoisomerase) as the only candidate gene showing a wild allele for its promoter region in the available tomato pangenome. This non-reference allele consists of a 645 bp promoter located at 582 bp upstream of *DXR*, and has been associated with an occurrence frequency of 0.58 in the *Solanum pimpinellifolium* L. accessions investigated by [26]. On the other hand, its presence in the *Solanum lycopersicum* var. *cerasiforme* and *S. lycopersicum* heirlooms has been reported to be rare, with an occurrence frequency of 0.05 and 0.02 respectively. The cultivars presenting the wild allele showed a significantly higher expression of Solyc03g11434 compared to the ones harboring the common allele, suggesting a selection against a higher expression of *DXR* during tomato domestication.

6. Phenylalanine-Derived Volatiles (Phe VOCs)

Phenolic and phenylpropanoid volatiles originate from the catabolism of phenylalanine. Phenolic VOCs include compounds such as phenylacetaldehyde, 2-phenylethanol, 1-nitro-2-phenylethane and 2-phenylacetonitrile (benzylnitrile; all floral odor type) which have been reported to affect consumer liking of tomato fruit, [8,12,16,17]. However, the effect of these phenolic compounds on flavour and consumer preference seems not easy to predict, since some studies show positive while others show negative effects of these compounds on consumer liking [13,92,172]. This apparent inconsistency may be caused by differences in the concentrations of these compounds in the tomato materials studied. The main phenylpropanoid VOCs in tomato are guaiacol, methylsalicylate and eugenol. They are associated with a smoky, pharmaceutical aroma and are generally considered as off-flavours.

6.1. Biosynthesis of Phe VOCs

Phenolic volatiles (C_6–C_2) have been reported to have a high impact on tomato aroma [11]. Their biochemical pathway starts with decarboxylation of phenylalanine, leading to the production of phenylethylamine [93]. According to the proposed phenolic volatile biosynthetic pathway in tomato, phenylethylamine is the precursor for the synthesis of the two nitrogen-containing volatiles nitrophenylethane and benzylnitrile, as well as the production of phenylacetaldehyde. Extremely high phenylacetaldehyde levels were found in tomato introgression lines carrying the malodorous locus from S. pennellii on chromosome 8 [92]. The high phenylacetaldehyde production in this line was associated with the expression of AADC1A, AADC1B, and AADC2 located in the malodorous QTL

region [93,173,174]. Transgenic approaches revealed that this family of genes was capable of decarboxylating phenylalanine, leading to phenylethylamine, the direct precursor of phenylacetaldehyde. The subsequent deamination of phenethylamine to produce phenylacetaldehyde has been reported to be related to an amine oxidase. Finally, 2-phenyethanol is produced from phenylacetaldehyde by means of two reductases, PAR1 and PAR2 [175–177]. These enzymes are not only reducing phenylacetaldehyde, but they are also able to catalyze the reduction of benzaldehyde and cinnamaldehyde to their respective alcohols as well [175]. More recently, FLORAL4 (Solyc04g063350—3-methyl-2-oxobutanoate dehydrogenase- has been fine mapped in a diversity panel of cultivated contemporary tomato varieties and tomato RIL populations, and associated with the floral phenolic volatiles accumulation in tomato fruit [102]. Based on the protein sequence FLORAL4 belongs to the mitochondrial 2-oxoisovalerate dehydrogenase/decarboxylase enzyme family which is involved in the catabolism of BCAAs and constitutes the E1 subunit of the BCKDC complex, catalyzing the decarboxylation of the BCAA deamination products in plants [131]. A complete knock-out of the FLORAL4 gene by CRISPR-Cas9-mediated gene editing in tomato plants led to a major depletion of the phenylalanine-derived volatiles as well as a notable depletion of BCAA VOCs. This suggests involvement of FLORAL4 in both BCAA and PHE VOC metabolism, possibly via decarboxylation of the corresponding amino or keto acids.

As mentioned above, phenylpropanoid volatiles (C_6–C_3) are the second group of phenylalanine derived VOCs. Although their biosynthesis in tomato has not been fully characterized, results from other species suggest that phenylpropanoid volatiles are derived from intermediates of the lignin pathway [178–181]. For example, eugenol has been reported in other species to be produced from coniferyl acetate by means of a eugenol synthase [182]. The mechanisms of methyl salicylate and guaiacol biosynthesis have been investigated in tomato [11]. Methyl salicylate can be produced from salicylic acid by means of SlSAMT1, an O-methyltransferase [183], while guaiacol may be produced from catechol, by means of the catechol-O-methyltransferase CTOMT1 [184]. Although CTOMT1 could not be connected to any of the QTL regions found, it has been included for completeness. Furthermore, the conjugation of these three compounds after their biosynthesis has been reported to be linked to the activity of two classes of enzymes: glycosyltransferases and glycosyl hydrolases. Among these classes of enzymes, NSGT1 was shown to prevent the wound-induced release of the smoky aroma associated with phenylpropanoid VOCs in ripening tomato fruit [185]. Figure 5 summarizes the metabolic reactions described above.
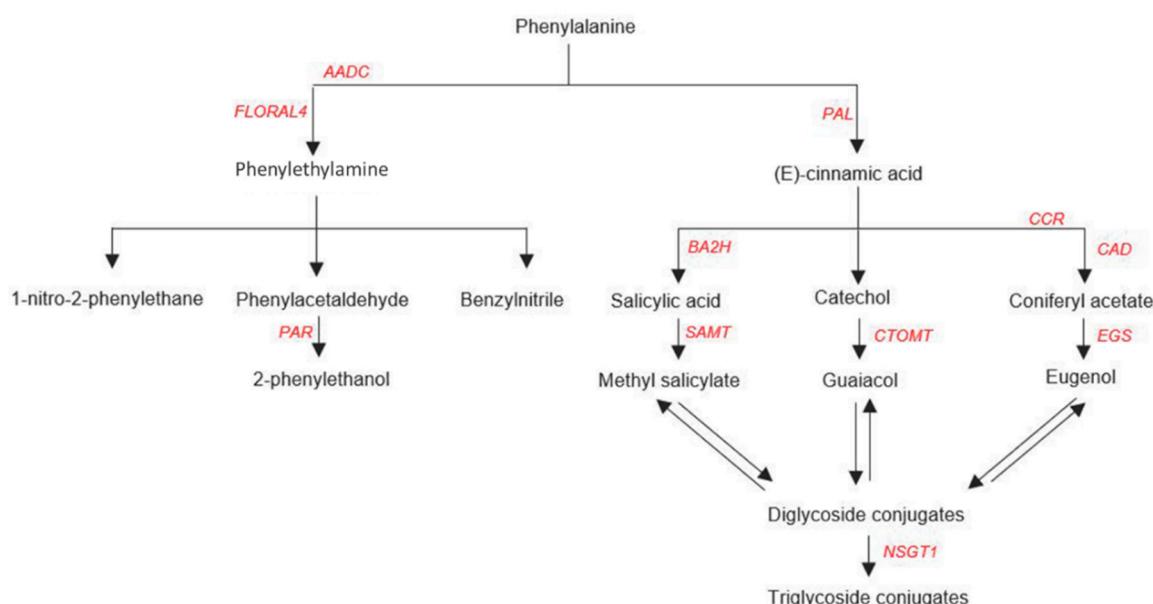

Figure 5. Overview of the phenolic and phenylpropanoid VOCs pathways, adapted from [11].

6.2. QTLs for Phe VOCs

Data collection identified a total number of 81 QTLs reported by eight different authors and correlated with phenylalanine-derived volatiles biosynthesis (Figure 2; Table S4). Comparing these regions (see "Data Acquisition" paragraph for methodology), we identified 24 distinct QGRs (Table S3).

6.3. Phe VOCs' Candidate Genes

Seventy-five genes were identified as potentially associated with phenylalanine derived VOCs, including all candidate genes present in the QGRs plus the above-mentioned CTOMT1 (Solyc10g005060—Table S4). Among them, twenty-one have been reported to be expressed in at least one fruit tissue or were characterized for their role in VOCs metabolism (Table 6 and Table S4; [103]). In addition, the above-mentioned NSGT1, present in QGR PHEN17, was also included in Table 6. NSGT1 expression in tomato fruit was reported by [185], but this gene has not been predicted by the tomato genome and its genomic organization has only recently been resolved [65], and hence gene expression data for this gene could not be retrieved from the TomExpress database.

Table 6. List of the genes identified as potentially responsible for the phenylalanine-derived VOCs QTL reported in literature. Genes that have been functionally characterized are highlighted in bold in the table. The last column reports whether the gene has been reported to be expressed in fruit (Y/N—[103]); for a complete overview of the gene expression see Table S4.

Solyc ID	Name	QGR	Start (bp)	Annotation	References	Expressed (Y/N)
Solyc01g107910		PHEN3	95,296,421	Caffeoyl CoA 3-*O*-methyltransferase		Y
Solyc02g093270		PHEN5	54,178,272	Caffeoyl CoA *O*-methyltransferase		Y
Solyc03g097700		PHEN6	60,029,697	*O*-methyltransferase		Y
Solyc03g111830		PHEN6	62,444,512	*O*-methyltransferase		Y
Solyc04g063350	*FLORAL4*	**PHEN7**	**55,462,543**	**3-methyl-2-oxobutanoate dehydrogenase**	[102]	**Y**
Solyc04g071140		PHEN7	58,070,304	Decarboxylase family protein		Y
Solyc05g013440		PHEN8	6,501,962	Primary amine oxidase		Y
Solyc06g059840	LOC778303	PHEN9	37,729,682	3-methyl-2-oxobutanoate dehydrogenase		Y
Solyc08g006740	*AADC2*	**PHEN14**	**1,295,712**	**Decarboxylase family protein**	[175]	**N**
Solyc08g066240		PHEN15	54,722,572	Decarboxylase family protein		Y
Solyc08g066250	*Hdc*	**PHEN15**	**54,745,923**	**Decarboxylase family protein**		**Y**
Solyc08g068600		PHEN15	57,730,921	Decarboxylase family protein		Y
Solyc08g068610	*AADC1B*	**PHEN15**	**57,740,004**	**Decarboxylase family protein**	[175]	**Y**
Solyc08g068680	*AADC1A*	**PHEN15**	**57,812,621**	**Decarboxylase family protein**	[175]	**N**
Solyc08g079430		PHEN15	62,954,528	Primary amine oxidase		Y
Solyc09g007890		PHEN16	1,413,536	Phenylalanine ammonia-lyase		Y
Solyc09g007900		PHEN16	1,419,041	Phenylalanine ammonia-lyase		Y
Solyc09g007910		PHEN16	1,429,132	Phenylalanine ammonia-lyase		Y
Solyc09g007920		PHEN16	1,435,451	Phenylalanine ammonia-lyase		Y
Solyc09g091550	**SAMT1**	**PHEN17**	**70,802,564**	**Salicylate methyltransferase 1**	[183]	**N**
	NSGT1	PHEN17	64,653,692	Glycosyltransferase	[185]	Y
Solyc10g005060	*CTOMT1*		**53,330**	**Catechol-*O*-methyltransferase 1**	[184]	**Y**
Solyc10g085830		PHEN18	64,899,145	*O*-methyltransferase 1		Y
Solyc10g086180		PHEN18	65,098,335	Phenylalanine ammonia-lyase		Y
Solyc12g013690	*SlFMO1*	**PHEN22**	**4,532,510**	**Monooxygenase FAD-binding protein**	[175]	**Y**

Among the genes identified as expressed in tomato fruit, one was annotated as 3-methyl-2-oxobutanoate dehydrogenase (Solyc04g063350), the above-mentioned FLORAL4 gene and was demonstrated to be the causal gene for the variation in phenolic VOCs in the PHEN7 QGR on chromosome 4.

One gene (SlFMO1—Solyc12g013690) was annotated as flavin-dependent monooxygenases, a class of enzymes that was shown to catalyze the hydroxylation of aromatic compounds in prokaryotes and plants [174–176]. This gene was suggested to play a role in the synthesis of nitrogenous phenolic volatiles, like 2-phenylacetonitrile and 1-nitro-2-phenylethane in tomato [186], although this needs experimental confirmation. Furthermore, we identified one primary amine oxidase (Solyc08g079430) that may potentially be involved in the production of 2-phenylacetaldehyde from phenethylamine [93].

We identified five phenylalanine ammonia-lyase (Solyc09g007890, Solyc09g007900, Solyc09g007910, Solyc09g007920, and Solyc10g086180) genes expressed in tomato fruit. These enzymes have been reported to be responsible for the first key step in phenylpropanoid metabolism, catalyzing the conversion of l-phenylalanine into trans-cinnamate [187].

Finally, six decarboxylase genes were reported to be expressed in tomato fruit during ripening [103,188]. Among them, Solyc08g066250 has been reported as SlHDC11 and has been associated with fruit-ripening in tomato [189]. Moreover, LeAADC1A (Solyc08g068680), together with LeAADC1B (Solyc08068610) and LeAADC2 (Solyc08g006740), have been related with the conversion of phenylalanine to phenylethylamine in a S. lycopersicum var. M82 × S. pennellii tomato IL population [93].

The tomato pangenome [26] was investigated for potential non-reference alleles in the promoter region of the identified candidate genes. A presence-absence variant was detected in the promoter of FLORAL4 (Solyc04g063350). This variant has an occurrence frequency of 0.74 in cultivated heirloom tomato and 0.64 in *S. lycopersicum* var. cerasiforme, while it occurs at a frequency of 0.02 only in *S. pimpinellifolium* WL. Gene expression analyses pinpointed a significant difference (*p*-value: 5.48×10^{-6}) between accessions harboring the domesticated versus the wild allele, suggesting that the domesticated allele of FLORAL4 has a lower expression compared to the wild allele.

7. Concluding Remarks

In the last decades, a vast array of scientific studies has investigated the genetic components of tomato aroma in modern tomato cultivars and their relatives. However, the methodological differences between different studies, such as the source materials, the type of mapping population, the number of markers, the reference map used, the influence of environmental factors on trait performance and analytical variation in determination of the aroma phenotype, make it difficult for breeders and researchers to efficiently use the available data in their research. This review summarized the state of the art on the understanding of tomato aroma genetics and provides a tool that can be used by breeders and researchers to collect additional evidence for the robustness of their QTL data. The identified QGRs, especially when defined by the overlapping of QTLs from different sources, represent the regions that most likely contain genetic elements that regulate tomato aroma. The QGRs differ from each other by two major parameters—the size and the number of studies contributing to it. The smaller the size of a resulting QGR and the higher the number of studies referring to it suggests a high robustness of such a QGR in a breeding program. Some of the QGRs, however, appeared to be quite large. This most likely indicates the presence of multiple genetic factors controlling the levels of (related) aromatic compounds in such a region, although we cannot exclude the possibility that, in some cases, a large QGR is caused by inaccurate prediction of a QTL due to experimental and/or methodological differences, as mentioned above. For a breeder such large QGRs could indicate a need for further inspection and dissection of the respective QGR in selected donor material, using high-resolution (fine) mapping approaches.

There are various potential applications of these data in practical breeding and breeding research:

(i) The QTL information presented in this review can be directly used to support marker-assisted breeding programs aimed at introgressing large-effect aroma QTLs into elite germplasm, for example using the donors indicated in Table S5.

(ii) The current development of pangenome projects paves the way for a new step in tomato breeding research. Advances in computational genomics and long-read sequencing allow an easier and more comprehensive investigation of the genetic variation in tomato collections worldwide. This makes it possible to identify genetic elements that are missing in the reference genome and to discover and use novel markers—such as Structural Variants (SVs) and Present Absent Variants (PAVs) [190]. Thanks to the ongoing reduction in sequencing costs, large sets of genotypes can nowadays be re-sequenced, allowing the application of SV markers in GWAS projects [191, 192]. Furthermore, SV identification in large sets of genotypes can lead to downstream breeding approaches. For example, SV-based linkage mapping can be applied by genotyping mapping populations using SV markers that showed polymorphism in the parental lines [193]. Furthermore, SV studies may help to get more insight in

the mechanisms leading to a certain phenotype [26,65,194,195], pinpointing the best donors for a certain allele cross. By providing a comprehensive set of candidate genes for tomato aroma, our review may guide researchers and breeders in the selection of the most interesting genes that can be investigated for structural variation.

(iii) The data in this review can be used to support the identification and use of the key genes underlying these QTLs. In combination with available (pan)genomic and transcriptomic information candidate genes present in the QTL regions can be selected and tested for their effects in vivo using either stable transgenic approaches such as CRISPR-CAS9 mediated gene editing or quicker transient overexpression or silencing in tomato fruit. This may not only lead to the identification of the causal genes controlling a trait, but also to the detection of the causal genetic variants underlying trait variation. Such variants, also called functional markers, are the best possible molecular markers for MAS, since they are functionally linked to the trait rather than genetically linked and their use as marker does not need validation in other populations, which is always required with genetically-linked markers [196].

(iv) The information on large effect aroma QTLs provided in this review can alo be used to improve the performance of genomic prediction models, since both genetically-linked markers and in particular functional markers have been shown to significantly improve the prediction power of GP models compared to the use of random neutral markers [197].

Last but not least we hope that this review may facilitate the development of more tasty tomatoes.

References

1. Peralta, I.; Knapp, S.; Spooner, D.M. New species of wild tomatoes (Solanum section lycopersicon: Solanaceae) from Northern Peru. *Syst. Bot.* **2005**, *30*, 424–434. [CrossRef]
2. Schouten, H.J.; Tikunov, Y.; Verkerke, W.; Finkers, R.; Bovy, A.; Bai, Y.; Visser, R.G. Breeding has increased the diversity of cultivated tomato in The Netherlands. *Front. Plant Sci.* **2019**, *10*, 1606. [CrossRef] [PubMed]
3. Víquez-Zamora, M.; Vosman, B.; Van De Geest, H.; Bovy, A.; Visser, R.G.F.; Finkers, R.; Van Heusden, A.W. Tomato breeding in the genomics era: Insights from a SNP array. *BMC Genom.* **2013**, *14*, 1–13. [CrossRef] [PubMed]
4. Lanfermeijer, F.C.; Dijkhuis, J.; Sturre, M.J.; De Haan, P.; Hille, J. Cloning and characterization of the durable tomato mosaic virus resistance gene Tm-22 from Lycopersicon esculentum. *Plant Mol. Biol.* **2003**, *52*, 1039–1051. [CrossRef] [PubMed]

5.   Lin, T.; Zhu, G.; Zhang, J.; Xu, X.; Yu, Q.; Zheng, Z.; Zhang, Z.; Lun, Y.; Li, S.; Wang, X.; et al. Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* **2014**, *46*, 1220–1226. [CrossRef] [PubMed]

6.   Mathieu, S.; Cin, V.D.; Fei, Z.; Li, H.; Bliss, P.; Taylor, M.G.; Klee, H.J.; Tieman, D.M. Flavour compounds in tomato fruits: Identification of loci and potential pathways affecting volatile composition. *J. Exp. Bot.* **2008**, *60*, 325–337. [CrossRef]

7.   Petro-Turza, M. Flavor of tomato and tomato products. *Food Rev. Int.* **1986**, *2*, 309–351. [CrossRef]

8.   Tieman, D.; Bliss, P.; McIntyre, L.M.; Blandon-Ubeda, A.; Bies, D.; Odabasi, A.Z.; Rodríguez, G.R.; Van Der Knaap, E.; Taylor, M.G.; Goulet, C.; et al. The chemical interactions underlying tomato flavor preferences. *Curr. Biol.* **2012**, *22*, 1035–1039. [CrossRef]

9.   Dudareva, N.; Klempien, A.; Muhlemann, J.K.; Kaplan, I. Biosynthesis, function and metabolic engineering of *Plant* volatile organic compounds. *New Phytol.* **2013**, *198*, 16–32. [CrossRef]

10.   Buttery, R.G.; Teranishi, R.; Flath, R.A.; Ling, L.C. Fresh Tomato Volatiles: Composition and Sensory Studies. In *Flavour Chemistry*; Teranishi, R., Buttery, R.G., Shahidi, F., Eds.; American Chemical Society: Washington, DC, USA, 1989; Volume 388, pp. 213–222.

11.   Rambla, J.L.; Tikunov, Y.M.; Monforte, A.J.; Bovy, A.G.; Granell, A. The expanded tomato fruit volatile landscape. *J. Exp. Bot.* **2013**, *65*, 4613–4623. [CrossRef]

12.   Tandon, K.; Baldwin, E.A.; Scott, J.; Shewfelt, R. Linking sensory descriptors to volatile and nonvolatile components of fresh tomato flavor. *J. Food Sci.* **2003**, *68*, 2366–2371. [CrossRef]

13.   Baldwin, E.A.; Goodner, K.; Plotto, A. Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. *J. Food Sci.* **2008**, *73*, S294–S307. [CrossRef] [PubMed]

14.   Baldwin, E.A.; Goodner, K.L.; Plotto, A.; Pritchett, K.; Einstein, M.A. Effect of volatiles and their concentration on perception of tomato descriptors. *J. Food Sci.* **2004**, *69*, S310–S318. [CrossRef]

15.   Vogel, J.T.; Tieman, D.M.; Sims, C.A.; Odabasi, A.Z.; Clark, D.G.; Klee, H.J. Carotenoid content impacts flavor acceptability in tomato (*Solanum lycopersicum*). *J. Sci. Food Agric.* **2010**, *90*, 2233–2240. [CrossRef]

16.   Abegaz, E.; Tandon, K.; Scott, J.; Baldwin, E.; Shewfelt, R. Partitioning taste from aromatic flavor notes of fresh tomato (*Lycopersicon esculentum*, Mill) to develop predictive models as a function of volatile and nonvolatile components. *Postharvest Biol. Technol.* **2004**, *34*, 227–235. [CrossRef]

17.   Piombino, P.; Sinesio, F.; Moneta, E.; Cammareri, M.; Genovese, A.; Lisanti, M.T.; Mogno, M.R.; Peparaio, M.; Termolino, P.; Moio, L.; et al. Investigating physicochemical, volatile and sensory parameters playing a positive or a negative role on tomato liking. *Food Res. Int.* **2013**, *50*, 409–419. [CrossRef]

18.   The Good Scents Company-Flavour, Fragrance, Food and Cosmetics Ingredients Information. Available online: http://www.thegoodscentscompany.com/ (accessed on 26 November 2020).

19.   Buttery, R.G.; Teranishi, R.; Ling, L.C. Fresh tomato aroma volatiles: A quantitative study. *J. Agric. Food Chem.* **1987**, *35*, 540–544. [CrossRef]

20.   Buttery, R.G.; Teranishi, R.; Ling, L.C.; Flath, R.A.; Stern, D.J. Quantitative studies on origins of fresh tomato aroma volatiles. *J. Agric. Food Chem.* **1988**, *36*, 1247–1250. [CrossRef]

21.   Baldwin, E.A.; Scott, J.W.; Shewmaker, C.K.; Schuch, W. Flavor trivia and tomato aroma: Biochemistry and possible mechanisms for control of important aroma components. *HortScience* **2000**, *35*, 1013–1022. [CrossRef]

22.   Tandon, K.; Baldwin, E.; Shewfelt, R. Aroma perception of individual volatile compounds in fresh tomatoes (*Lycopersicon esculentum*, Mill.) as affected by the medium of evaluation. *Postharvest Biol. Technol.* **2000**, *20*, 261–268. [CrossRef]

23.   Tandon, K.S.; Baldwin, E.A.; Goodner, K.L.; Jordán, M. Characterization of fresh tomato aroma volatiles using GC-Olfatometry. In Proceedings of the Annual Meeting of the Florida State Horticultural Society, Stuart, FL, USA, 10–12 June 2001.

24.   Selli, S.; Kelebek, H.; Ayseli, M.T.; Tokbas, H. Characterization of the most aroma-active compounds in cherry tomato by application of the aroma extract dilution analysis. *Food Chem.* **2014**, *165*, 540–546. [CrossRef] [PubMed]

25.   Tieman, D.; Zhu, G.; Resende, M.F.; Lin, T.; Nguyen, C.; Bies, D.; Rambla, J.L.; Beltran, K.S.O.; Taylor, M.; Zhang, B.; et al. A chemical genetic roadmap to improved tomato flavor. *Science* **2017**, *355*, 391–394. [CrossRef]

26.   Gao, L.; Gonda, I.; Sun, H.; Ma, Q.; Bao, K.; Tieman, D.M.; Burzynski-Chang, E.A.; Fish, T.L.; Stromberg, K.A.; Sacks, G.L.; et al. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* **2019**, *51*, 1044–1051. [CrossRef]

27.   Gur, A.; Semel, Y.; Osorio, S.; Friedmann, M.; Seekh, S.; Ghareeb, B.; Mohammad, A.; Pleban, T.; Gera, G.; Fernie, A.R.; et al. Yield quantitative trait loci from wild tomato are predominately expressed by the shoot. *Theor. Appl. Genet.* **2010**, *122*, 405–420. [CrossRef]

28.   Liu, Z.; Alseekh, S.; Brotman, Y.; Zheng, Y.; Fei, Z.; Tieman, D.M.; Giovannoni, J.J.; Fernie, A.R.; Klee, H.J. Identification of a *Solanum pennellii* chromosome 4 fruit flavor and nutritional quality-associated metabolite QTL. *Front. Plant Sci.* **2016**, *7*, 1671. [CrossRef]

29.   Eshed, Y.; Zamir, D. An introgression line population of lycopersicon pennellii in the cultivated tomato enables the identi-fication and fine mapping of yield-associated QTL. *Genetics* **1995**, *141*, 1147–1162. [CrossRef]

30.   Fridman, E.; Carrari, F.; Liu, Y.-S.; Fernie, A.R.; Zamir, D. Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* **2004**, *305*, 1786–1789. [CrossRef] [PubMed]

31.   Canady, M.A.; Meglic, V.; Chetelat, R.T. A library of *Solanum lycopersicoides* introgression lines in cultivated tomato. *Genome* **2005**, *48*, 685–697. [CrossRef] [PubMed]

32. Barrantes, W.; López-Casado, G.; García-Martínez, S.; Alonso, A.; Rubio, F.; Ruiz, J.J.; Fernández-Muñoz, R.; Granell, A.; Monforte, A. Exploring new alleles involved in tomato fruit quality in an introgression line library of *Solanum pimpinellifolium*. *Front. Plant Sci.* **2016**, *7*, 1172. [CrossRef]

33. Rambla, J.L.; Medina, A.; Fernández-Del-Carmen, A.; Barrantes, W.; Grandillo, S.; Cammareri, M.; López-Casado, G.; Rodrigo, G.; Alonso, A.; García-Martínez, S.; et al. Identification, introgression, and validation of fruit volatile QTLs from a red-fruited wild tomato species. *J. Exp. Bot.* **2016**, *68*, 429–442. [CrossRef]

34. Blanca, J.; Montero-Pau, J.; Sauvage, C.; Bauchet, G.; Illa-Berenguer, E.; Díez, M.J.; Francis, D.M.; Causse, M.; Van Der Knaap, E.; Canizares, J. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genom.* **2015**, *16*, 257. [CrossRef]

35. The Tomato Genome Consortium. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **2012**, *485*, 635–641. [CrossRef]

36. Ozminkowski, R.; Ozminkowski, R. Pedigree of variety heinz 1706. *Rep. Tomato Genet. Coop.* **2004**, *54*, 26.

37. Hamilton, J.P.; Sim, S.-C.; Stoffel, K.; Van Deynze, A.; Buell, C.R.; Francis, D.M. Single nucleotide polymorphism discovery in cultivated tomato via sequencing by synthesis. *Plant Genome* **2012**, *5*, 17–29. [CrossRef]

38. Sim, S.-C.; Durstewitz, G.; Plieske, J.; Wieseke, R.; Ganal, M.W.; Van Deynze, A.; Hamilton, J.P.; Buell, C.R.; Causse, M.; Wijeratne, S.; et al. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE* **2012**, *7*, e40563. [CrossRef] [PubMed]

39. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **2011**, *6*, e19379. [CrossRef]

40. Celik, I.; Gurbuz, N.; Uncu, A.T.; Frary, A.; Doganlar, S. Genome-wide SNP discovery and QTL mapping for fruit quality traits in inbred backcross lines (IBLs) of *Solanum pimpinellifolium* using genotyping by sequencing. *BMC Genom.* **2017**, *18*, 1–10. [CrossRef]

41. Ohlson, E.W.; Ashrafi, H.; Foolad, M.R. Identification and mapping of late blight resistance quantitative trait loci in tomato accession PI 163245. *Plant Genome* **2018**, *11*, 180007. [CrossRef]

42. Gonda, I.; Ashrafi, H.; Lyon, D.A.; Strickler, S.R.; Hulse-Kemp, A.M.; Ma, Q.; Sun, H.; Stoffel, K.; Powell, A.F.; Futrell, S.; et al. Sequencing-based bin map construction of a tomato mapping population, facilitating high-resolution quantitative trait loci detection. *Plant Genome* **2019**, *12*, 1–14. [CrossRef]

43. Lin, Y.-P.; Liu, C.-Y.; Chen, K.-Y. Assessment of genetic differentiation and linkage disequilibrium in *Solanum pimpinellifolium* using genome-wide high-density SNP markers. *G3 Genes Genomes Genet.* **2019**, *9*, 1497–1505. [CrossRef]

44. Acquadro, A.; Barchi, L.; Gramazio, P.; Portis, E.; Vilanova, S.; Comino, C.; Plazas, M.; Prohens, J.; Lanteri, S. Coding SNPs analysis highlights genetic relationships and evolution pattern in eggplant complexes. *PLoS ONE* **2017**, *12*, e0180774. [CrossRef]

45. Edae, E.A.; Olivera, P.D.; Jin, Y.; Rouse, M.N. Genotyping-by-sequencing facilitates a high-density consensus linkage map forAegilops umbellulata, a wild relative of cultivated wheat. *G3 Genes Genomes Genet.* **2017**, *7*, 1551–1561. [CrossRef]

46. Guo, X.; Cericola, F.; Fè, D.; Pedersen, M.G.; Lenk, I.; Jensen, C.S.; Jensen, J.; Janss, L.L. Genomic prediction in tetraploid ryegrass using allele frequencies based on genotyping by sequencing. *Front. Plant Sci.* **2018**, *9*, 1165. [CrossRef]

47. Hussain, W.; Baenziger, P.S.; Belamkar, V.; Guttieri, M.J.; Venegas, J.P.; Easterly, A.; Sallam, A.; Poland, J. Genotyping-by-sequencing derived high-density linkage map and its application to QTL mapping of flag leaf traits in bread wheat. *Sci. Rep.* **2017**, *7*, 16394. [CrossRef] [PubMed]

48. Yao, C.; Chen, G.; Song, C.; Keefe, J.; Mendelson, M.; Huan, T.; Sun, B.B.; Laser, A.; Maranville, J.C.; Wu, H.; et al. Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease. *Nat. Commun.* **2018**, *9*, 1–11. [CrossRef]

49. Jin, Y.; Zhao, W.; Nie, S.; Liu, S.-S.; El-Kassaby, Y.A.; Wang, X.-R.; Mao, J.-F. Genome-wide variant identification and high-density genetic map construction using RADseq for *Platycladus orientalis* (*Cupressaceae*). *G3 Genes Genomes Genet.* **2019**, *9*, 3663–3672. [CrossRef] [PubMed]

50. Qi, P.; Eudy, D.; Schnable, J.C.; Schmutz, J.; Raymer, P.L.; Devos, K.M. High density genetic maps of *Seashore paspalum* using genotyping-by-sequencing and their relationship to the sorghum bicolor genome. *Sci. Rep.* **2019**, *9*, 12183. [CrossRef]

51. Toppino, L.; Barchi, L.; Mercati, F.; Acciarri, N.; Perrone, D.; Martina, M.; Gattolin, S.; Sala, T.; Fadda, S.; Mauceri, A.; et al. A new intra-specific and high-resolution genetic map of eggplant based on a RIL population, and location of QTLs related to *Plant* anthocyanin pigmentation and seed vigour. *Genes* **2020**, *11*, 745. [CrossRef]

52. Jordon-Thaden, I.E.; Beck, J.B.; Rushworth, C.A.; Windham, M.D.; Diaz, N.; Cantley, J.T.; Martine, C.T.; Rothfels, C.J. A basic dd RAD seq two-enzyme protocol performs well with herbarium and silica-dried tissues across four genera. *Appl. Plant Sci.* **2020**, *8*, e11344. [CrossRef]

53. Sunde, J.; Yıldırım, Y.; Tibblin, P.; Forsman, A. Comparing the performance of microsatellites and RADseq in population genetic studies: Analysis of data for pike (*Esox lucius*) and a synthesis of previous studies. *Front. Genet.* **2020**, *11*, 218. [CrossRef]

54. Tsujimoto, M.; Araki, K.S.; Honjo, M.N.; Yasugi, M.; Nagano, A.J.; Akama, S.; Hatakeyama, M.; Shimizu-Inatsugi, R.; Sese, J.; Shimizu, K.K.; et al. Genet assignment and population structure analysis in a clonal forest-floor herb, *Cardamine leucantha*, using RAD-seq. *AoB Plants* **2019**, *12*. [CrossRef]

55. Feng, J.; Zhao, S.; Li, M.; Zhang, C.; Qu, H.; Li, Q.; Li, J.; Lin, Y.; Pu, Z. Genome-wide genetic diversity detection and population structure analysis in sweetpotato (*Ipomoea batatas*) using RAD-seq. *Genomics* **2020**, *112*, 1978–1987. [CrossRef]

56. Aflitos, S.A.; Schijlen, E.; De Jong, H.; De Ridder, D.; Smit, S.; Finkers, R.; Wang, J.; Zhang, G.; Li, N.; Mao, L.; et al. Exploring genetic variation in the tomato (*Solanumsection Lycopersicon*) clade by whole-genome sequencing. *Plant J.* **2014**, *80*, 136–148. [CrossRef] [PubMed]

57. Roohanitaziani, R.; De Maagd, R.A.; Lammers, M.; Molthoff, J.; Meijer-Dekens, F.; Van Kaauwen, M.P.W.; Finkers, R.; Tikunov, Y.; Visser, R.G.F.; Bovy, A. Exploration of a resequenced tomato core collection for phenotypic and genotypic variation in *Plant* growth and fruit quality traits. *Genes* **2020**, *11*, 1278. [CrossRef]

58. Hirsch, C.N.; Foerster, J.M.; Johnson, J.M.; Sekhon, R.S.; Muttoni, G.; Vaillancourt, B.; Peñagaricano, F.; Lindquist, E.; Pedraza, M.A.; Barry, K.; et al. Insights into the maize pan-genome and pan-transcriptome. *Plant Cell* **2014**, *26*, 121–135. [CrossRef] [PubMed]

59. Li, Y.-H.; Zhou, G.; Ma, J.; Jiang, W.; Jin, L.-G.; Zhang, Z.; Guo, Y.; Zhang, J.; Sui, Y.; Zheng, L.; et al. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nat. Biotechnol.* **2014**, *32*, 1045–1052. [CrossRef]

60. Schatz, M.C.; Maron, L.G.; Stein, J.C.; Wences, A.H.; Gurtowski, J.; Biggers, E.; Lee, H.; Kramer, M.; Antoniou, E.; Ghiban, E.; et al. Whole genome de novo assemblies of three divergent strains of rice, *Oryza sativa*, document novel gene space of aus and indica. *Genome Biol.* **2014**, *15*, 1–16. [CrossRef]

61. Golicz, A.A.; Bayer, P.E.; Barker, G.C.; Edger, P.P.; Kim, H.; Martinez, P.A.; Chan, C.-K.K.; Severn-Ellis, A.; McCombie, W.R.; Parkin, I.A.P.; et al. The pangenome of an agronomically important crop *Plant Brassica oleracea*. *Nat. Commun.* **2016**, *7*, 13390. [CrossRef]

62. Song, J.-M.; Guan, Z.; Hu, J.; Guo, C.; Yang, Z.; Wang, S.; Liu, D.; Wang, B.; Lu, S.; Zhou, R.; et al. Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nat. Plants* **2020**, *6*, 34–45. [CrossRef]

63. Khan, A.W.; Garg, V.; Roorkiwal, M.; Golicz, A.A.; Edwards, D.; Varshney, R.K. Super-pangenome by integrating the wild side of a species for accelerated crop improvement. *Trends Plant Sci.* **2020**, *25*, 148–158. [CrossRef] [PubMed]

64. Tettelin, H.; Masignani, V.; Cieslewicz, M.J.; Donati, C.; Medini, D.; Ward, N.L.; Angiuoli, S.V.; Crabtree, J.; Jones, A.L.; Durkin, A.S.; et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13950–13955. [CrossRef] [PubMed]

65. Alonge, M.; Wang, X.; Benoit, M.; Soyk, S.; Pereira, L.; Zhang, L.; Suresh, H.; Ramakrishnan, S.; Maumus, F.; Ciren, D.; et al. Major impacts of widespread structural variation on gene expression and crop improvement in tomato. *Cell* **2020**, *182*, 145–161.e23. [CrossRef]

66. Ou, L.; Li, D.; Lv, J.; Chen, W.; Zhang, Z.; Li, X.; Yang, B.; Zhou, S.; Yang, S.; Li, W.; et al. Pan-genome of cultivated pepper (*Capsicum*) and its use in gene presence-absence variation analyses. *New Phytol.* **2018**, *220*, 360–363. [CrossRef]

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

68. Shen, J.; Tieman, D.; Jones, J.B.; Taylor, M.G.; Schmelz, E.; Huffaker, A.; Bies, D.; Chen, K.; Klee, H.J. A 13-lipoxygenase, TomloxC, is essential for synthesis of C5 flavour volatiles in tomato. *J. Exp. Bot.* **2014**, *65*, 419–428. [CrossRef] [PubMed]

69. Yang, W.; Guo, Z.; Huang, C.; Duan, L.; Chen, G.; Jiang, N.; Fang, W.; Feng, H.; Xie, W.; Lian, X.; et al. Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* **2014**, *5*. [CrossRef] [PubMed]

70. Yang, J.; Jiang, H.; Yeh, C.; Yu, J.; Jeddeloh, J.A.; Nettleton, D.; Schnable, P.S. Extreme-phenotype genome-wide association study (XP-GWAS): A method for identifying trait-associated variants by sequencing pools of individuals selected from a diversity panel. *Plant J.* **2015**, *84*, 587–596. [CrossRef]

71. Minamikawa, M.F.; Takada, N.; Terakami, S.; Saito, T.; Onogi, A.; Kajiya-Kanegae, H.; Hayashi, T.; Yamamoto, T.; Iwata, H. Genome-wide association study and genomic prediction using parental and breeding populations of Japanese pear (*Pyrus pyrifolia Nakai*). *Sci. Rep.* **2018**, *8*, 11994. [CrossRef]

72. Mir, R.R.; Reynolds, M.; Pinto, F.; Khan, M.A.; Bhat, M.A. High-throughput phenotyping for crop improvement in the genomics era. *Plant Sci.* **2019**, *282*, 60–72. [CrossRef]

73. Visscher, P.M.; Wray, N.R.; Zhang, Q.; Sklar, P.; McCarthy, M.I.; Brown, M.A.; Yang, J. 10 Years of GWAS discovery: Biology, function, and translation. *Am. J. Hum. Genet.* **2017**, *101*, 5–22. [CrossRef] [PubMed]

74. Sauvage, C.; Segura, V.; Bauchet, G.; Stevens, R.; Do, P.T.; Nikoloski, Z.; Fernie, A.R.; Causse, M. Genome-wide association in tomato reveals 44 candidate loci for fruit metabolic traits. *Plant Physiol.* **2014**, *165*, 1120–1132. [CrossRef]

75. Zhang, J.; Zhao, J.; Xu, Y.; Liang, J.; Chang, P.; Yan, F.; Li, M.; Liang, Y.; Zou, Z. Genome-wide association mapping for tomato volatiles positively contributing to tomato flavor. *Front. Plant Sci.* **2015**, *6*, 6. [CrossRef]

76. Bauchet, G.; Grenier, S.; Samson, N.; Bonnet, J.; Grivet, L.; Causse, M. Use of modern tomato breeding germplasm for deciphering the genetic control of agronomical traits by genome wide association study. *Theor. Appl. Genet.* **2017**, *130*, 875–889. [CrossRef]

77. Zhao, J.; Sauvage, C.; Zhao, J.; Bitton, F.; Bauchet, G.; Liu, D.; Huang, S.; Tieman, D.M.; Klee, H.J.; Causse, M. Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. *Nat. Commun.* **2019**, *10*, 1–12. [CrossRef] [PubMed]

78. Phan, N.T.; Trinh, L.T.; Rho, M.-Y.; Park, T.-S.; Kim, O.-R.; Zhao, J.; Kim, H.-M.; Sim, S.-C. Identification of loci associated with fruit traits using genome-wide single nucleotide polymorphisms in a core collection of tomato (*Solanum lycopersicum* L.). *Sci. Hortic.* **2019**, *243*, 567–574. [CrossRef]

79. Ye, J.; Li, W.; Ai, G.; Li, C.; Liu, G.; Chen, W.; Wang, B.; Wang, W.; Lu, Y.; Zhang, J.; et al. Genome-wide association analysis identifies a natural variation in basic helix-loop-helix transcription factor regulating ascorbate biosynthesis via D-mannose/L-galactose pathway in tomato. *PLoS Genet.* **2019**, *15*, e1008149. [CrossRef] [PubMed]

80. Gao, N.; Teng, J.; Ye, S.; Yuan, X.; Huang, S.; Zhang, H.; Zhang, X.; Li, J.; Zhang, Z. Genomic prediction of complex phenotypes using genic similarity based relatedness matrix. *Front. Genet.* **2018**, *9*, 364. [CrossRef] [PubMed]

81. Daetwyler, H.D.; Calus, M.P.L.; Pong-Wong, R.; Campos, G.D.L.; Hickey, J.M. Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. *Genetics* **2013**, *193*, 347–365. [CrossRef] [PubMed]

82. Yamamoto, E.; Matsunaga, H.; Onogi, A.; Kajiya-Kanegae, H.; Minamikawa, M.; Suzuki, A.; Shirasawa, K.; Hirakawa, H.; Nunome, T.; Yamaguchi, H.; et al. A simulation-based breeding design that uses whole-genome prediction in tomato. *Sci. Rep.* **2016**, *6*. [CrossRef]

83. Tusell, L.; Pérez-Rodríguez, P.; Forni, S.; Wu, X.-L.; Gianola, D. Genome-enabled methods for predicting litter size in pigs: A comparison. *Animal* **2013**, *7*, 1739–1749. [CrossRef]

84. Ni, G.; Cavero, D.; Fangmann, A.; Erbe, M.; Simianer, H. Whole-genome sequence-based genomic prediction in laying chickens with different genomic relationship matrices to account for genetic architecture. *Genet. Sel. Evol.* **2017**, *49*, 8. [CrossRef] [PubMed]

85. Duenk, P.; Calus, M.P.L.; Wientjes, Y.C.J.; Breen, V.P.; Henshall, J.; Hawken, R.J.; Bijma, P. Validation of genomic predictions for body weight in broilers using crossbred information and considering breed-of-origin of alleles. *Genet. Sel. Evol.* **2019**, *51*, 1–12. [CrossRef]

86. Pandey, M.K.; Chaudhari, S.; Jarquin, D.; Janila, P.; Crossa, J.; Patil, S.C.; Sundravadana, S.; Khare, D.; Bhat, R.S.; Radhakrishnan, T.; et al. Genome-based trait prediction in multi-environment breeding trials in groundnut. *Theor. Appl. Genet.* **2020**, *133*, 3101–3117. [CrossRef]

87. Lenz, P.; Nadeau, S.; Azaiez, A.; Gérardi, S.; DesLauriers, M.; Perron, M.; Isabel, N.; Beaulieu, J.; Bousquet, J. Genomic prediction for hastening and improving efficiency of forward selection in conifer polycross mating designs: An example from white spruce. *Heredity* **2020**, *124*, 562–578. [CrossRef]

88. Robertsen, C.D.; Hjortshøj, R.L.; Janss, L. Genomic selection in cereal breeding. *Agronomy* **2019**, *9*, 95. [CrossRef]

89. Cortés, A.J.; Restrepo-Montoya, M.; Bedoya-Canas, L.E. Modern strategies to assess and breed forest tree adaptation to changing climate. *Front. Plant Sci.* **2020**, *11*. [CrossRef] [PubMed]

90. Saliba-Colombani, V.; Causse, M.; Langlois, D.; Philouze, J.; Buret, M. Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theor. Appl. Genet.* **2001**, *102*, 259–272. [CrossRef]

91. Fulton, T.; Bucheli, P.; Voirol, E.; Lopez, J.; Pétiard, V.; Tanksley, S. 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* **2002**, *127*, 163–177. [CrossRef]

92. Tadmor, Y.; Fridman, E.; Gur, A.; Larkov, O.; Lastochkin, E.; Ravid, U.; Zamir, D.; Lewinsohn, E. Identification of malodorous, a wild species allele affecting tomato aroma that was selected against during domestication. *J. Agric. Food Chem.* **2002**, *50*, 2005–2009. [CrossRef] [PubMed]

93. Tieman, D.M.; Zeigler, M.; Schmelz, E.A.; Taylor, M.G.; Bliss, P.; Kirst, M.; Klee, H.J. Identification of loci affecting flavour volatile emissions in tomato fruits. *J. Exp. Bot.* **2006**, *57*, 887–896. [CrossRef]

94. Kochevenko, A.; Fernie, A.R. The genetic architecture of branched-chain amino acid accumulation in tomato fruits. *J. Exp. Bot.* **2011**, *62*, 3895–3906. [CrossRef]

95. Ruggieri, V.; Francese, G.; Sacco, A.; D'Alessandro, A.; Rigano, M.M.; Parisi, M.; Milone, M.; Cardi, T.; Mennella, G.; Barone, A.; et al. An association mapping approach to identify favourable alleles for tomato fruit quality breeding. *BMC Plant Biol.* **2014**, *14*, 1–15. [CrossRef]

96. Capel, C.; Fernández-Del-Carmen, A.; Alba, J.M.; Lima-Silva, V.; Gras, F.H.; Salinas, M.; Boronat, A.; Angosto, T.; Botella, J.R.; Fernández-Muñoz, R.; et al. Wide-genome QTL mapping of fruit quality traits in a tomato RIL population derived from the wild-relative species *Solanum pimpinellifolium* L. *Theor. Appl. Genet.* **2015**, *128*, 2019–2035. [CrossRef] [PubMed]

97. Zhao, J.; Xu, Y.; Ding, Q.; Huang, X.; Zhang, Y.; Zou, Z.; Li, M.; Cui, L.; Zhang, J. Association mapping of main tomato fruit sugars and organic acids. *Front. Plant Sci.* **2016**, *7*. [CrossRef]

98. Ebaldina, S.; Picarella, M.E.; Troise, A.D.; Epucci, A.; Eruggieri, V.; Eferracane, R.; Ebarone, A.; Efogliano, V.; Mazzucato, A. Metabolite profiling of italian tomato landraces with different fruit types. *Front. Plant Sci.* **2016**, *7*. [CrossRef]

99. Bauchet, G.; Grenier, S.; Samson, N.; Segura, V.; Kende, A.; Beekwilder, J.; Cankar, K.; Gallois, J.; Gricourt, J.; Bonnet, J.; et al. Identification of major loci and genomic regions controlling acid and volatile content in tomato fruit: Implications for flavor improvement. *New Phytol.* **2017**, *215*, 624–641. [CrossRef] [PubMed]

100. Garbowicz, K.; Liu, Z.; Alseekh, S.; Tieman, D.; Taylor, M.; Kuhalskaya, A.; Ofner, I.; Zamir, D.; Klee, H.J.; Fernie, A.R.; et al. Quantitative trait loci analysis identifies a prominent gene involved in the production of fatty acid-derived flavor volatiles in tomato. *Mol. Plant* **2018**, *11*, 1147–1165. [CrossRef]

101. Kimbara, J.; Ohyama, A.; Chikano, H.; Ito, H.; Hosoi, K.; Negoro, S.; Miyatake, K.; Yamaguchi, H.; Nunome, T.; Fukuoka, H.; et al. QTL mapping of fruit nutritional and flavor components in tomato (*Solanum lycopersicum*) using genome-wide SSR markers and recombinant inbred lines (RILs) from an intra-specific cross. *Euphytica* **2018**, *214*, 210. [CrossRef]

102. Tikunov, Y.; Roohanitaziani, R.; Meijer-Dekens, F.; Molthoff, J.; Paulo, J.; Finkers, R.; Capel, I.; Moreno, F.C.; Maliepaard, C.; Vries, M.N.; et al. The genetic and functional analysis of flavor in commercial tomato: The FLORAL4 gene underlies a QTL for floral aroma volatiles in tomato fruit. *Plant J.* **2020**, *103*, 1189–1204. [CrossRef]

103. Zouine, M.; Maza, E.; Djari, A.; Lauvernier, M.; Frasse, P.; Smouni, A.; Pirrello, J.; Bouzayen, M. TomExpress, a unified tomato RNA-Seq platform for visualization of expression data, clustering and correlation networks. *Plant J.* **2017**, *92*, 727–735. [CrossRef] [PubMed]

104. Fall, R.; Karl, T.; Hansel, A.; Jordan, A.; Lindinger, W. Volatile organic compounds emitted after leaf wounding: On-line analysis by proton-transfer-reaction mass spectrometry. *J. Geophys. Res. Space Phys.* **1999**, *104*, 15963–15974. [CrossRef]

105. Ferrie, B. The cloning of two tomato lipoxygenase genes and their differential expression during fruit ripening. *Plant Physiol.* **1994**, *106*, 109–118. [CrossRef] [PubMed]

106. Jansen, G.A.; Wanders, R.J. α-Oxidation. *Biochim. Biophys. Acta BBA Bioenerg.* **2006**, *1763*, 1403–1412. [CrossRef] [PubMed]

107. Li, C.; Schilmiller, A.L.; Liu, G.; Lee, G.I.; Jayanty, S.; Sageman, C.; Vrebalov, J.; Giovannoni, J.J.; Yagi, K.; Kobayashi, Y.; et al. Role of β-Oxidation in *Jasmonate biosynthesis* and systemic wound signaling in tomato. *Plant Cell* **2005**, *17*, 971–986. [CrossRef] [PubMed]

108. Liavonchanka, A.; Feussner, I. Lipoxygenases: Occurrence, functions and catalysis. *J. Plant Physiol.* **2006**, *163*, 348–357. [CrossRef] [PubMed]

109. Prestage, S.; Linforth, R.S.; Taylor, A.; Lee, E.; Speirs, J.; Schuch, W. Volatile production in tomato fruit with modified alcohol dehydrogenase activity. *J. Sci. Food Agric.* **1999**, *79*, 131–136. [CrossRef]

110. Matsui, K. Green leaf volatiles: Hydroperoxide lyase pathway of oxylipin metabolism. *Curr. Opin. Plant Biol.* **2006**, *9*, 274–280. [CrossRef]

111. Chen, H.; Cao, S.; Jin, Y.; Tang, Y.; Qi, H. The relationship between CmADHs and the diversity of volatile organic compounds of three aroma types of melon (*Cucumis melo*). *Front. Physiol.* **2016**, *7*, 254. [CrossRef]

112. Vieira, D.D.S.S.; Emiliani, G.; Michelozzi, M.; Centritto, M.; Luro, F.; Morillon, R.; Loreto, F.; Gesteira, A.S.; Maserti, B. Poly-ploidization alters constitutive content of volatile organic compounds (VOC) and improves membrane stability under water deficit in Volkamer lemon (*Citrus limonia* Osb.) leaves. *Environ. Exp. Bot.* **2016**, *126*, 1–9. [CrossRef]

113. Zhou, Z.; Wu, Q.; Yao, Z.; Deng, H.; Liu, B.; Yue, C.; Deng, T.; Lai, Z.; Sun, Y. Dynamics of ADH and related genes responsible for the transformation of C 6-aldehydes to C 6-alcohols during the postharvest process of oolong tea. *Food Sci. Nutr.* **2019**, *8*, 104–113. [CrossRef]

114. Furuhashi, T.; Ishii, R.; Onishi, H.; Ota, S. Elucidation of biochemical pathways underlying VOCs production in A549 cells. *Front. Mol. Biosci.* **2020**, *7*, 116. [CrossRef]

115. Domínguez, T.; Hernández, M.L.; Pennycooke, J.C.; Jiménez, P.; Martínez-Rivas, J.M.; Sanz, C.; Stockinger, E.J.; Sánchez-Serrano, J.J.; Sanmartín, M. Increasing ω-3 desaturase expression in tomato results in altered aroma profile and enhanced resistance to cold stress. *Plant Physiol.* **2010**, *153*, 655–665. [CrossRef]

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

117. Moummou, H.; Kallberg, Y.; Tonfack, L.B.; Persson, B.; Van Der Rest, B. The *Plant* short-chain dehydrogenase (SDR) superfamily: Genome-wide inventory and diversification patterns. *BMC Plant Biol.* **2012**, *12*, 219. [CrossRef] [PubMed]

118. Gosset, V.; Harmel, N.; Göbel, C.; Francis, F.; Haubruge, E.; Wathelet, J.-P.; Du Jardin, P.; Feussner, I.; Fauconnier, M.-L. Attacks by a piercing-sucking insect (*Myzus persicae* Sultzer) or a chewing insect (*Leptinotarsa decemlineata* Say) on potato plants (*Solanum tuberosum* L.) induce differential changes in volatile compound release and oxylipin synthesis. *J. Exp. Bot.* **2009**, *60*, 1231–1240. [CrossRef]

119. Campestre, C.; Angelini, G.; Gasbarri, C.; Angerosa, F. The compounds responsible for the sensory profile in monovarietal virgin olive oils. *Molecules* **2017**, *22*, 1833. [CrossRef] [PubMed]

120. Vong, W.C.; Liu, S.-Q. Bioconversion of green volatiles in okara (*Soybean residue*) into esters by coupling enzyme catalysis and yeast (*Lindnera saturnus*) fermentation. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 10017–10026. [CrossRef]

121. Mariutto, M.; Duby, F.; Adam, A.; Bureau, C.; Fauconnier, M.-L.; Ongena, M.; Thonart, P.; Dommes, J. The elicitation of a systemic resistance by *Pseudomonas putida* BTP1 in tomato involves the stimulation of two lipoxygenase isoforms. *BMC Plant Biol.* **2011**, *11*, 29. [CrossRef]

122. Longhurst, T.J.; Tung, H.F.; Brady, C.J. Developmental Regulation of the Expression of Alcohol Dehydrogenase in Ripening Tomato Fruits. *J. Food Biochem.* **1990**, *14*, 421–433. [CrossRef]

123. Van Der Straeten, D.; Rodrigues-Pousada, R.A.; Gielen, J.; Van Montagu, M. Tomato alcohol dehydrogenase: Expression during fruit ripening and under hypoxic conditions. *FEBS Lett.* **1991**, *295*, 39–42. [CrossRef]

124. Buttery, R.G.; Teranishi, R.; Flath, R.A.; Ling, L.C. Identification of additional tomato paste volatiles. *J. Agric. Food Chem.* **1990**, *38*, 792–795. [CrossRef]

125. Brosnan, J.T.; Brosnan, M.E. Branched-chain amino acids: Enzyme and substrate regulation. *J. Nutr.* **2006**, *136*, 207S–211S. [CrossRef] [PubMed]

126. Schuster, J.; Binder, S. The mitochondrial branched-chain aminotransferase (AtBCAT-1) is capable to initiate degradation of leucine, isoleucine and valine in almost all tissues in Arabidopsis thaliana. *Plant Mol. Biol.* **2005**, *57*, 241–254. [CrossRef] [PubMed]

127. Maloney, G.S.; Kochevenko, A.; Tieman, D.M.; Tohge, T.; Krieger, U.; Zamir, D.; Taylor, M.G.; Fernie, A.R.; Klee, H.J. Characterization of the branched-chain amino acid aminotransferase enzyme family in tomato. *Plant Physiol.* **2010**, *153*, 925–936. [CrossRef] [PubMed]

128. Kochevenko, A.; Klee, H.J.; Fernie, A.R.; Araújo, W.L. Molecular identification of a further branched-chain aminotransferase 7 (BCAT7) in tomato plants. *J. Plant Physiol.* **2012**, *169*, 437–443. [CrossRef]

129. Tohge, T.; Alseekh, S.; Fernie, A.R. On the regulation and function of secondary metabolism during fruit development and ripening. *J. Exp. Bot.* **2013**, *65*, 4599–4611. [CrossRef]

130. Marilley, L. Flavours of cheese products: Metabolic pathways, analytical tools and identification of producing strains. *Int. J. Food Microbiol.* **2004**, *90*, 139–159. [CrossRef]

131. Mooney, B.P.; Miernyk, J.A.; Randall, D.D. Thecomplexfate ofα-ketoacids. *Annu. Rev. Plant Biol.* **2002**, *53*, 357–375. [CrossRef]

132. Simkin, A.J.; Schwartz, S.H.; Auldridge, M.; Taylor, M.G.; Klee, H.J. The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles β-ionone, pseudoionone, and geranylacetone. *Plant J.* **2004**, *40*, 882–892. [CrossRef]

133. Wei, Y.; Wan, H.; Wu, Z.; Wang, R.; Ruan, M.; Ye, Q.; Li, Z.; Zhou, G.; Yao, Z.; Yang, Y.; et al. A comprehensive analysis of carotenoid cleavage dioxygenases genes in *Solanum lycopersicum*. *Plant Mol. Biol. Rep.* **2015**, *34*, 512–523. [CrossRef]

134. Floss, D.S.; Walter, M.H. Role of carotenoid cleavage dioxygenase 1 (CCD1) in apocarotenoid biogenesis revisited. *Plant Signal. Behav.* **2009**, *4*, 172–175. [CrossRef]

135. Vogel, J.T.; Tan, B.-C.; Mccarty, D.R.; Klee, H.J. The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. *J. Biol. Chem.* **2008**, *283*, 11364–11373. [CrossRef] [PubMed]

136. Lashbrooke, J.G.; Young, P.R.; Dockrall, S.J.; Vasanth, K.; A Vivier, M. Functional characterisation of three members of the *Vitis vinifera* L. carotenoid cleavage dioxygenase gene family. *BMC Plant Biol.* **2013**, *13*, 156. [CrossRef]

137. Frusciante, S.; Diretto, G.; Bruno, M.; Ferrante, P.; Pietrella, M.; Prado-Cabrero, A.; Rubio, A.-M.; Beyer, P.D.; Gomez-Gomez, L.; Al-Babili, S.; et al. Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 12246–12251. [CrossRef]

138. Ilg, A.; Bruno, M.; Beyer, P.; Al-Babili, S. Tomato carotenoid cleavage dioxygenases 1A and 1B: Relaxed double bond specificity leads to a plenitude of dialdehydes, mono-apocarotenoids and isoprenoid volatiles. *FEBS Open Bio* **2014**, *4*, 584–593. [CrossRef]

139. Qin, X.; Fischer, K.; Yu, S.; Dubcovsky, J.; Tian, L. Distinct expression and function of carotenoid metabolic genes and homoeologs in developing wheat grains. *BMC Plant Biol.* **2016**, *16*, 155. [CrossRef]

140. Klee, H.J. Improving the flavor of fresh fruits: Genomics, biochemistry, and biotechnology. *New Phytol.* **2010**, *187*, 44–56. [CrossRef] [PubMed]

141. Simpson, K.; Quiroz, L.F.; Rodriguez-Concepción, M.; Stange, C. Differential contribution of the first two enzymes of the MEP pathway to the supply of metabolic precursors for carotenoid and chlorophyll biosynthesis in carrot (*Daucus carota*). *Front. Plant Sci.* **2016**, *7*. [CrossRef] [PubMed]

142. Manoharan, R.K.; Jung, H.-J.; Hwang, I.; Jeong, N.; Kho, K.H.; Chung, M.-Y.; Nou, I.S. Molecular breeding of a novel orange-brown tomato fruit with enhanced β-carotene and chlorophyll accumulation. *Hereditas* **2017**, *154*, 1. [CrossRef]

143. Enfissi, E.M.; Nogueira, M.; Bramley, P.M.; Fraser, P. The regulation of carotenoid formation in tomato fruit. *Plant J.* **2017**, *89*, 774–788. [CrossRef] [PubMed]

144. Liu, L.; Shao, Z.; Zhang, M.; Wang, Q. Regulation of carotenoid metabolism in tomato. *Mol. Plant* **2015**, *8*, 28–39. [CrossRef] [PubMed]

145. García-Alcázar, M.; Giménez, E.; Pineda, B.; Capel, C.; García-Sogo, B.; Sánchez, S.; Yuste-Lisbona, F.J.; Angosto, T.; Capel, J.; Moreno, V.; et al. Albino T-DNA tomato mutant reveals a key function of 1-deoxy-D-xylulose-5-phosphate synthase (DXS1) in *Plant* development and survival. *Sci. Rep.* **2017**, *7*, 45333. [CrossRef] [PubMed]

146. D'Andrea, L.; Simon-Moya, M.; Llorente, B.; Llamas, E.; Marro, M.; Loza-Alvarez, P.; Li, L.; Rodríguez-Concepción, M. Interference with Clp protease impairs carotenoid accumulation during tomato fruit ripening. *J. Exp. Bot.* **2018**, *69*, 1557–1568. [CrossRef] [PubMed]

147. McQuinn, R.P.; Wong, B.; Giovannoni, J.J. AtPDSoverexpression in tomato: Exposing unique patterns of carotenoid self-regulation and an alternative strategy for the enhancement of fruit carotenoid content. *Plant Biotechnol. J.* **2018**, *16*, 482–494. [CrossRef] [PubMed]

148. Chetelat, R.T.; Qin, X.; Tan, M.; Burkart-Waco, D.; Moritama, Y.; Huo, X.; Wills, T.; Pertuzé, R. Introgression lines of *Solanum sitiens* a wild nightshade of the Atacama Desert, in the genome of cultivated tomato. *Plant J.* **2019**, *100*, 836–850. [CrossRef]

149. Xiong, C.; Luo, D.; Lin, A.; Zhang, C.; Shan, L.; He, P.; Li, B.; Zhang, Q.; Hua, B.; Yuan, Z.; et al. A tomato B-box proteinSlBBX20 modulates carotenoid biosynthesis by directly activatingPHYTOENE SYNTHASE 1, and is targeted for 26S proteasome-mediated degradation. *New Phytol.* **2019**, *221*, 279–294. [CrossRef]

150. Britton, G.; Pfander, H.; Liaaen-Jensen, S. *Carotenoids*; Birkhäuser Verlag: Basel, Switzerland, 1998; Volume 3.

151. Galpaz, N.; Wang, Q.; Menda, N.; Zamir, D.; Hirschberg, J. Abscisic acid deficiency in the tomato mutant high-pigment 3 leading to increased plastid number and higher fruit lycopene content. *Plant J.* **2008**, *53*, 717–730. [CrossRef]

152. Mann, V.; Pecker, I.; Hirschberg, J. Cloning and characterization of the gene for phytoene desaturase (Pds) from tomato (*Lycopersicon esculentum*). *Plant Mol. Biol.* **1994**, *24*, 429–434. [CrossRef]

153. Pecker, I.; Gabbay, R.; Cunningham, F.X.; Hirschberg, J. Cloning and characterization of the cDNA for lycopene β-cyclase from tomato reveals decrease in its expression during fruit ripening. *Plant Mol. Biol.* **1996**, *30*, 807–819. [CrossRef]

154. Ronen, G.; Carmel-Goren, L.; Zamir, D.; Hirschberg, J. An alternative pathway to β-carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 11102–11107. [CrossRef] [PubMed]

155. Lois, L.M.; Rodriguez-Concepcion, M.; Gallego, F.; Campos, N.; Boronat, A. Carotenoid biosynthesis during tomato fruit development: Regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant J.* **2000**, *22*, 503–513. [CrossRef] [PubMed]

156. Bartley, G.E.; Scolnik, P.A. cDNA cloning, expression during development, and genome mapping of PSY2, a second tomato gene encoding phytoene synthase. *J. Biol. Chem.* **1993**, *268*, 25718–25721. [CrossRef]

157. Fraser, P.D.; Truesdale, M.R.; Bird, C.R.; Schuch, W.; Bramley, P.M. Carotenoid biosynthesis during tomato fruit development (Evidence for tissue-specific gene expression). *Plant Physiol.* **1994**, *105*, 405–413. [CrossRef]

158. Fraser, P.D.; Enfissi, E.M.; Halket, J.M.; Truesdale, M.R.; Yu, D.; Gerrish, C.; Bramley, P.M. Manipulation of phytoene levels in tomato fruit: Effects on isoprenoids, plastids, and intermediary metabolism. *Plant Cell* **2007**, *19*, 3194–3211. [CrossRef]

159. Giorio, G.; Stigliani, A.L.; D'Ambrosio, C. Phytoene synthase genes in tomato (*Solanumlycopersicum*, L.) -new data on the structures, the deduced amino acid sequences and the expression patterns. *FEBS J.* **2007**, *275*, 527–535. [CrossRef]

160. Rodríguez-Villalón, A.; Gas, E.; Rodríguez-Concepción, M. Phytoene synthase activity controls the biosynthesis of carotenoids and the supply of their metabolic precursors in dark-grown *Arabidopsis seedlings*. *Plant J.* **2009**, *60*, 424–435. [CrossRef]

161. Isaacson, T.; Ronen, G.; Zamir, D.; Hirschberg, J. Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of β-carotene and xanthophylls in plants. *Plant Cell* **2002**, *14*, 333–342. [CrossRef]

162. Ronen, G.; Cohen, M.; Zamir, D.; Hirschberg, J. Regulation of carotenoid biosynthesis during tomato fruit development: Expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant delta. *Plant J.* **1999**, *17*, 341–351. [CrossRef] [PubMed]

163. Apel, W.; Bock, R. Enhancement of carotenoid biosynthesis in transplastomic tomatoes by induced Lycopene-to-Provitamin a conversion. *Plant Physiol.* **2009**, *151*, 59–66. [CrossRef] [PubMed]

164. D'Ambrosio, C.; Stigliani, A.L.; Giorio, G. Overexpression of CrtR-b2 (carotene β hydroxylase 2) from *S. lycopersicum* L. differentially affects xanthophyll synthesis and accumulation in transgenic tomato plants. *Transgenic Res.* **2010**, *20*, 47–60. [CrossRef]

165. Gonzalez-Jorge, S.; Mehrshahi, P.; Magallanes-Lundback, M.; Lipka, A.E.; Angelovici, R.; Gore, M.A.; DellaPenna, D. zeaxanthin epoxidase activity potentiates carotenoid degradation in maturing seed. *Plant Physiol.* **2016**, *171*, 1837–1851. [CrossRef]

166. Alder, A.; Jamil, M.; Marzorati, M.; Bruno, M.; Vermathen, M.; Bigler, P.; Ghisla, S.; Bouwmeester, H.; Beyer, P.; Al-Babili, S.; et al. The path from-carotene to carlactone, a strigolactone-like plant hormone. *Science* **2012**, *335*, 1348–1351. [CrossRef]

167. Huang, F.-C.; Molnár, P.; Schwab, W. Cloning and functional characterization of carotenoid cleavage dioxygenase 4 genes. *J. Exp. Bot.* **2009**, *60*, 3011–3022. [CrossRef]

168. Campbell, R.; Ducreux, L.J.; Morris, W.L.; Morris, J.A.; Suttle, J.C.; Ramsay, G.; Bryan, G.J.; Hedley, P.E.; Taylor, M.A. The metabolic and developmental roles of carotenoid cleavage dioxygenase4 from potato. *Plant Physiol.* **2010**, *154*, 656–664. [CrossRef]

169. Brandi, F.; Bar, E.; Mourgues, F.; Horváth, G.; Turcsi, E.; Giuliano, G.; Liverani, A.; Tartarini, S.; Lewinsohn, E.; Rosati, C. Study of "Redhaven" peach and its white-fleshed mutant suggests a key role of CCD4 carotenoid dioxygenase in carotenoid and norisoprenoid volatile metabolism. *BMC Plant Biol.* **2011**, *11*, 24. [CrossRef]

170. Jang, S.-J.; Jeong, H.-B.; Jung, A.; Kang, M.-Y.; Kim, S.; Ha, S.-H.; Kwon, J.-K.; Kang, B.-C. Phytoene synthase 2 can compensate for the absence of PSY1 in the control of color in *Capsicum* fruit. *J. Exp. Bot.* **2020**, *71*, 3417–3427. [CrossRef]

171. Fantini, E.; Falcone, G.; Frusciante, S.; Giliberto, L.; Giuliano, G. Dissection of tomato *Lycopene biosynthesis* through virus-induced gene silencing. *Plant Physiol.* **2013**, *163*, 986–998. [CrossRef] [PubMed]

172. Tzin, V.; Galili, G. New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. *Mol. Plant* **2010**, *3*, 956–972. [CrossRef] [PubMed]

173. Sakai, M.; Hirata, H.; Sayama, H.; Sekiguchi, K.; Itano, H.; Asai, T.; Dohra, H.; Hara, M.; Watanabe, N. Production of 2-Phenylethanol in roses as the dominant floral scent compound fromL-phenylalanine by two key enzymes, a PLP-dependent decarboxylase and a phenylacetaldehyde reductase. *Biosci. Biotechnol. Biochem.* **2007**, *71*, 2408–2419. [CrossRef]

174. Xie, Q.; Liu, Z.; Meir, S.; Rogachev, I.; Aharoni, A.; Klee, H.J.; Galili, G. Altered metabolite accumulation in tomato fruits by coexpressing a feedback-insensitive AroG and the Ph ODO 1 MYB-type transcription factor. *Plant Biotechnol. J.* **2016**, *14*, 2300–2309. [CrossRef] [PubMed]

175. Tieman, D.M.; Loucas, H.M.; Kim, J.Y.; Clark, D.G.; Klee, H.J. Tomato phenylacetaldehyde reductases catalyze the last step in the synthesis of the aroma volatile 2-phenylethanol. *Phytochemistry* **2007**, *68*, 2660–2669. [CrossRef] [PubMed]

176. Costa, M.A.; Marques, J.V.; Dalisay, D.S.; Herman, B.; Bedgar, D.L.; Davin, L.B.; Lewis, N.G. Transgenic hybrid poplar for sustainable and scalable production of the commodity/specialty chemical, 2-Phenylethanol. *PLoS ONE* **2013**, *8*, e83169. [CrossRef] [PubMed]

177. Liu, P.; Cheng, Y.; Yang, M.; Liu, Y.; Chen, K.; Long, C.-A.; Deng, X. Mechanisms of action for 2-phenylethanol isolated from *Kloeckera apiculata* in control of *Penicillium* molds of citrus fruits. *BMC Microbiol.* **2014**, *14*. [CrossRef] [PubMed]

178. Singh, R.; Rastogi, S.; Dwivedi, U.N. Phenylpropanoid metabolism in ripening fruits. *Compr. Rev. Food Sci. Food Saf.* **2010**, *9*, 398–416. [CrossRef]

179. Vogt, T. Phenylpropanoid biosynthesis. *Mol. Plant* **2010**, *3*, 2–20. [CrossRef] [PubMed]

180. Fraser, C.M.; Chapple, C. The phenylpropanoid pathway in arabidopsis. *Arab. Book* **2011**, *9*, e0152. [CrossRef]
181. Biała, W.; Jasiński, M. The phenylpropanoid case-it is transport that matters. *Front. Plant Sci.* **2018**, *9*, 9. [CrossRef]
182. Koeduka, T.; Fridman, E.; Gang, D.R.; Vassão, D.G.; Jackson, B.L.; Kish, C.M.; Orlova, I.; Spassova, S.M.; Lewis, N.G.; Noel, J.P.; et al. Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10128–10133. [CrossRef]
183. Tieman, D.; Zeigler, M.; Schmelz, E.; Taylor, M.G.; Rushing, S.; Jones, J.B.; Klee, H.J. Functional analysis of a tomato salicylic acid methyl transferase and its role in synthesis of the flavor volatile methyl salicylate. *Plant J.* **2010**, *62*, 113–123. [CrossRef]
184. Mageroy, M.H.; Tieman, D.M.; Floystad, A.; Taylor, M.G.; Klee, H.J. A *Solanum lycopersicum* catechol-O-methyltransferase involved in synthesis of the flavor molecule guaiacol. *Plant J.* **2012**, *69*, 1043–1051. [CrossRef]
185. Tikunov, Y.; Molthoff, J.; De Vos, R.C.; Beekwilder, J.; Van Houwelingen, A.; Van Der Hooft, J.J.; Vries, M.N.-D.; Labrie, C.W.; Verkerke, W.; Van De Geest, H.; et al. Non-smoky glycosyltransferase1 prevents the release of smoky aroma from tomato fruit. *Plant Cell* **2013**, *25*, 3067–3078. [CrossRef]
186. Klee, H.J.; Tieman, D.M. The genetics of fruit flavour preferences. *Nat. Rev. Genet.* **2018**, *19*, 347–356. [CrossRef]
187. Rigano, M.M.; Raiola, A.; Docimo, T.; Ruggieri, V.; Calafiore, R.; Vitaglione, P.; Ferracane, R.; Frusciante, L.; Barone, A. Metabolic and molecular changes of the phenylpropanoid pathway in tomato (*Solanum lycopersicum*) lines carrying different *Solanum pennellii* wild chromosomal regions. *Front. Plant Sci.* **2016**, *7*. [CrossRef] [PubMed]
188. Kim, A.Y.; Kim, H.M.; Ma, S.H.; Park, S.Y.; Dat, M.T.; Jang, G.; Joung, Y.H. The promoter of tomato *Histidine decarboxylase* A is fruit-specific, and its expression is stably maintained in fruits during ripening. *Plant Biotechnol. Rep.* **2019**, *13*, 43–50. [CrossRef]
189. Kumar, R. Evolutionary trails of plant group ii pyridoxal phosphate-dependent decarboxylase genes. *Front. Plant Sci.* **2016**, *7*. [CrossRef] [PubMed]
190. Wellenreuther, M.; Mérot, C.; Berdan, E.; Bernatchez, L. Going beyond SNPs: The role of structural genomic variants in adaptive evolution and species diversification. *Mol. Ecol.* **2019**, *28*, 1203–1209. [CrossRef] [PubMed]
191. Guo, J.; Cao, K.; Deng, C.; Li, Y.; Zhu, G.; Fang, W.; Chen, C.; Wang, X.; Wu, J.; Guan, L.; et al. An integrated peach genome structural variation map uncovers genes associated with fruit traits. *Genome Biol.* **2020**, *21*, 258. [CrossRef] [PubMed]
192. Zhang, Z.; Mao, L.; Chen, H.; Bu, F.; Li, G.; Sun, J.; Li, S.; Sun, H.; Jiao, C.; Blakely, R.; et al. Genome-wide mapping of structural variations reveals a copy number variant that determines reproductive morphology in cucumber. *Plant Cell* **2015**, *27*, 1595–1604. [CrossRef]
193. Shen, X.; Liu, Z.-Q.; Mocoeur, A.; Xia, Y.; Jing, H.-C. PAV markers in *Sorghum bicolour*: Genome pattern, affected genes and pathways, and genetic linkage map construction. *Theor. Appl. Genet.* **2015**, *128*, 623–637. [CrossRef] [PubMed]
194. McHale, L.K.; Haun, W.J.; Xu, W.W.; Bhaskar, P.B.; Anderson, J.E.; Hyten, D.L.; Gerhardt, D.J.; Jeddeloh, J.A.; Stupar, R.M. Structural variants in the soybean genome localize to clusters of biotic stress-response genes. *Plant Physiol.* **2012**, *159*, 1295–1308. [CrossRef] [PubMed]
195. Gabur, I.; Chawla, H.S.; Lopisso, D.T.; Von Tiedemann, A.; Snowdon, R.J.; Obermeier, C. Gene presence-absence variation associates with quantitative *Verticillium longisporum* disease resistance in *Brassica napus*. *Sci. Rep.* **2020**, *10*, 4131. [CrossRef] [PubMed]
196. Salgotra, R.K.; Stewart, C.N. Functional markers for precision *Plant* breeding. *Int. J. Mol. Sci.* **2020**, *21*, 4792. [CrossRef] [PubMed]
197. Fu, Y.-B.; Yang, M.-H.; Zeng, F.; Biligetu, B. Searching for an accurate marker-based prediction of an individual quantitative trait in molecular plant breeding. *Front. Plant Sci.* **2017**, *8*, 1182. [CrossRef] [PubMed]