

REVIEW

Open Access



Combining metabolomic and transcriptomic approaches to assess and improve crop quality traits

Delphine M. Pott, Sara Durán-Soria, Sonia Osorio and José G. Vallarino* 

Abstract

Plant quality trait improvement has become a global necessity due to the world overpopulation. In particular, producing crop species with enhanced nutrients and health-promoting compounds is one of the main aims of current breeding programs. However, breeders traditionally focused on characteristics such as yield or pest resistance, while breeding for crop quality, which largely depends on the presence and accumulation of highly valuable metabolites in the plant edible parts, was left out due to the complexity of plant metabolome and the impossibility to properly phenotype it. Recent technical advances in high throughput metabolomic, transcriptomic and genomic platforms have provided efficient approaches to identify new genes and pathways responsible for the extremely diverse plant metabolome. In addition, they allow to establish correlation between genotype and metabolite composition, and to clarify the genetic architecture of complex biochemical pathways, such as the accumulation of secondary metabolites in plants, many of them being highly valuable for the human diet. In this review, we focus on how the combination of metabolomic, transcriptomic and genomic approaches is a useful tool for the selection of crop varieties with improved nutritional value and quality traits.

Keywords: Crop, Quality, Domestication, Metabolomics, Transcriptomics, Genomics, GWAS

Background

Understanding the genes involved in metabolism and dissection of the metabolic pathways are essential to improve plant adaptation to biotic and abiotic stresses, to improve food quality, and to increase crop yield, all being crucial factors to secure global human nutrition (Gong et al. 2013). Crop domestication and intensive breeding along history contributed to decrease genetic diversity though successive bottlenecks (Bai and Lindhout 2007; Sauvage et al. 2014). While breeders have traditionally focused on improving agronomical important traits, such as yield and pest resistance, crop quality traits, such as

flavour or nutritional value, have long been omitted. In fact, deterioration in flavour quality of important commercial crops, such as tomato, is a major cause of consumer complaint (Klee and Tieman 2018). To reverse this decline, identification and quantification of the compounds responsible for consumer liking must be first undergone, followed by the understanding of what has been lost during domestication and intensive breeding (Tieman et al. 2017). In addition, improving seed quality of main crops is a major target of breeding programs, to ensure human nutrition, as they provide a significant percentage (around 20%) of the consumed energy, protein and dietary fiber (Peng et al. 2018).

Most crop agronomic and quality characteristics are being determined by multiple quantitative traits (QTL) and influenced by the environmental and culture conditions. Loci underlying these economically-important

*Correspondence: vallarino@uma.es

Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora" – Consejo Superior de Investigaciones Científicas, Universidad de Málaga – (IHSM-CSIC-UMA), Málaga, Spain



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

characteristics can be located by linkage mapping and genome-wide association (GWAS). Association between genetic variation and traits of interest can further be used for marker-assisted selection in the breeding programs, in order to obtain high quality crops.

As a result of plant metabolism, fruits and cereals accumulate a myriad of metabolites, which main functions are to protect them against pathogens and herbivores and to attract seed dispersers (Pott et al. 2019). Primary metabolism leads to the synthesis of sugars, amino and organic acids, which are used as building components by the cells and serve as precursors for secondary metabolism. This specialized metabolism, crucial for the plant interaction with its entourage, is extremely diverse and results in the accumulation of a plethora of compounds, *i.e.* polyphenols, terpenoids or nitrogen/sulfur-containing metabolites (Aharoni and Galili 2011; Pott et al. 2019). Many of these secondary metabolites have been associated with health-promoting characteristics (Pott et al. 2019; Sun et al. 2018).

In this sense, recent advances in metabolite profiling technologies, which now allow the simultaneous detection and quantification of thousands of metabolites, combined with genomic and transcriptomic platforms, are fundamental in dissecting crop composition and in identifying the genetic variants underlying metabolic content (de Abreu e Lima et al. 2018; Ballester et al. 2016; Garbowicz et al. 2018; Labadie et al. 2020; Li et al. 2019; Osorio et al. 2011, 2019; Rambla et al. 2016; Vallarino et al. 2019). While the integration of genetic and metabolic information is a powerful strategy to dissect the bases of plant metabolism and to associate complex traits with genotype, it is also a key strategy for the breeding of high-yielding and nutritionally rich crops (Luo 2015; Wen et al. 2018). In this review, we focus on recent studies, published during the last 6 years, which combined metabolomic approaches with genomic or transcriptomic advances to dissect plant biochemistry for crop quality improvement. Web of Science and Scopus databases were screened using the keywords metabolomics, transcriptomics, genomics, proteomics, GWAS, fruit OR crop, quality. Recent studies (2014–2020) were selected when combining metabolomic approach with any of the other above-mentioned omic strategies. As a result of this initial screen, we decided to focus our review on three important and diverse groups of metabolites in different crops associated with fruit quality, *i.e.* (i) essential amino acid content in cereals, (ii) health-promoting and colourful secondary metabolites (anthocyanins and carotenoids) in fruits and (iii) taste- and aroma-related metabolites and the recovery of the associated alleles as a result of tomato domestication and intensive breeding. In this sense, the review will allow us to cover a wide

range of major crop species, including cereals, fruits and tubers, focusing mainly on strategies to improve plant nutritional value. In parallel, the unprecedented example of tomato genetic dissection for fruit quality improvement will enable us to present omic strategies which could be translated to other crop species. In addition, methods and techniques used for metabolomic, genomic and transcriptomic analyses in the studies reviewed here are summarized in Table 1.

Main text

Combining metabolomic and genomic tools to increase amino acid content in cereals and soybean

Primary metabolites, mainly sugars, acids and their derivatives, serve both as cell building blocks and as essential nutrients for animal and human consumption, as they accumulate in sink organs, such as fruits, seeds or tubers (Wen et al. 2018).

The rapid development of next generation sequencing platforms has allowed high-density genotyping with SNPs, favouring the use of GWAS to fine-map loci involved in plant metabolic complexity. In particular, combining GWAS with metabolomic platforms (mGWAS) has permitted the simultaneous screening of an enormous number of accessions for primary metabolic content first in the model plant *Arabidopsis* (Chan et al. 2010) and then in a series of economically important crops, such as maize, rice or tomato (Chen et al. 2014; Sauvage et al. 2014; Wen et al. 2014). Levels of most primary metabolites are found to be normally controlled by multiple loci, each of them explaining a low to moderate percentage of the metabolic variability (Chen et al. 2014; Deng et al. 2017; Wen et al. 2015). Identified QTLs are normally not randomly distributed on the chromosomes, but QTL ‘hotspots’ are commonly observed, suggesting the combined action of closely linked genes or pleiotropic effect (Chen et al. 2014, 2016; Deng et al. 2017; Gong et al. 2013; Peng et al. 2018; Vallarino et al. 2019). It is also important to outline that the overlapping between different populations or varieties in the identified loci is quite limited, as shown by Chen et al. (2014) or Deng et al. (2017). Indeed, combining metabolic profiling of rice (*Oryza sativa*) leaves from *indica* and *japonica* subspecies with GWAS, only 155 out of 514 identified loci were common in both groups, indicating heterogeneous genetic control (Chen et al. 2014). Furthermore, most of the loci identified by GWAS appear to be tissue-specific, implying separate genetic and biochemical regulation (Chen et al. 2016; Gong et al. 2013; Wen et al. 2015, 2018). This tissue-specific metabolic variation is possibly determined by differential allelic expression (Chen et al. 2016). However, common strategies for

Table 1 Summary of the different metabolomic, transcriptomic and genomic methods and softwares (when information available) used in the cited studies, to improve crop quality

Species	Techniques	Platforms	Reference
Maize	Amino acid analysis	Amino acid analyzer	Deng et al. (2017)
	GWAS	GEC software	
	QTL mapping	Windows QTL mapping Cartographer V2.5	
Maize	Primary metabolite profiling	GC-TOF-MS	Wen et al. (2015)
	QTL mapping	Windows QTL mapping Cartographer V2.5	
Maize	Primary metabolite profiling	GC-TOF-MS	Wen et al. (2018)
	GWAS	TASSEL3.0 software	
Rice	Metabolite profiling	LC-ESI-MS/MS	Chen et al. (2016)
	GWAS	FaST-LMM program	
Bread wheat	Metabolite profiling	HPLC-MS	Peng et al. (2018)
	GWAS	FaST-LMM program	
Rice	BCAA analysis	LC-ESI-MS/MS	Sun et al. (2020)
	GWAS	FaST-LMM program	
Soybean	Amino acid analysis	NIR spectroscopy	Zhang et al. (2018)
	GWAS	GAPIT, R	
Soybean	amino acid analysis	Amino acid analyzer	Qin et al. (2019)
	GWAS	TASSEL 5 software	
Potato	Metabolite profiling	UPLC-Triple-TOF-MS	Cho et al. (2016)
	Transcriptomics	RNA-Seq (Bowtie, DESeq package in R)	
Plum	Anthocyanin analysis	Spectrophotometric determination	Fang et al. (2016)
	Transcriptomics	RNA-Seq (in-house Perl scripts), Trinity for de novo assembly	
Pepper	Flavonoid profiling	UPLC-ESI-MS/MS	Liu et al. (2020)
	Carotenoid profiling	UPLC-APCI-MS/MS	
	Transcriptomics	RNASEq (HISAT2, DESeq2 softwares)	
Fig	Metabolite profiling	HPLC-ESI-MS/MS	Wang et al. (2017a)
	Transcriptomics	RNASEq (RSEM software)	
Pear	Anthocyanin analysis	Spectrophotometric determination	Wang et al. (2017b)
	Transcriptomics	RNA-Seq	
Pear	Metabolite profiling	HPLC	Zhang et al. (2020)
	Transcriptomics	RNA-Seq (NGS QC Toolkit, DESeq R package)	
Kiwifruit	Metabolite profiling	UPLC-Triple-TOF-MS	Li et al. (2018b)
	Transcriptomics	RNASEq (iTAK, DESeq R package)	
Wild peach	Metabolite profiling	HPLC-ESI-MS/MS	Ying et al. (2019)
	Transcriptomics	RNA-Seq (TopHat2, cufflinks, cuffcompare packages)	
Kiwifruit	Anthocyanin and phenolic compounds analysis	RP-HPLC	Nardoza et al. (2020)
	Metabolite profiling	HPLC-MS	
	Transcriptomics	RNA-Seq (STAR, DESeq2 R package)	
Maize	Carotenoid profiling	HPLC	Azmach et al. (2018)
	GWAS	TASSEL 4 software	
Cassava	Total carotenoid content	iCheck analytical kit	Esuma et al. (2016)
	GWAS	TASSEL v5.2.9	
Maize	Carotenoid profiling	HPLC	Owens et al. (2014)
	GWAS	Proc MIXED SAS 9.3 software	
Chickpea	Carotenoid profiling	HPLC	Rezaei et al. (2019)
	QTL mapping	IciMapping software	
Maize	Carotenoid profiling	HPLC	Xu et al. (2019)
	QTL mapping	Windows QTL mapping Cartographer 2.5 (Wang	

Table 1 (continued)

Species	Techniques	Platforms	Reference	
Melon	Carotenoid and chlorophyll quantification	HPLC	Oren et al. (2019)	
	GWAS	TASSEL		
Watermelon	QTL mapping	QTL mappingbim R package	Li et al. (2020)	
Watermelon	Lycopene content	HPLC	Wang et al. (2019)	
	QTL mapping	IciMapping V3.3 software		
Sweetpotato	QTL mapping	QTL mappingpoly software	Gemenet et al. (2020)	
Carrot	Carotenoid profiling	HPLC	Ellison et al. (2018)	
	GWAS	TASSEL-GBS version 5.2.26		
Banana	Carotenoid profiling	HPLC	Heng et al. (2019)	
	Proteomics	nanolc-MS/MS		
Tomato	Metabolite profiling	LC-ESI-MS/MS	Zhu et al. (2018)	
	Transcriptomics	RNA-Seq (RNA-SeQC)		
	GWAS	FaST-LMM program	Bauchet et al. (2017)	
	Sugar profiling	High pH ion-exchange chromatography		
	Organic and amino acid profiling	LC-ESI-MS/MS		
	Volatiles	GC-MS		
	GWAS	EMMA algorithm		
	Carotenoid profiling	RP-HPLC		Ruggieri et al. (2014)
	Ascorbic acid content	spectrophotometric determination		Sauvage et al. (2014)
	GWAS	TASSEL v4.0, STRUCTURE 2.3.3		
	Metabolite profiling	GC-TOF-MS		Tiemann et al. (2017)
	GWAS	EMMA algorithm		
	Volatile profiling	GC-MS/SPME-GC-MS	Ye et al. (2019)	
	GWAS	EMMAX, GEC softwares		
	QTL mapping	EM algorithm	D'Angelo et al. (2019)	
	Primary metabolite profiling	GC-MS		
	GWAS	TASSEL v4.0	D'Esposito et al. (2017)	
	Metabolite profiling	³ H-NMR spectroscopy		
	Transcriptomics	GeneChip [®] Tomato arrays (Robin application)	Tikunov et al. (2020)	
	Semi-polar metabolite Profiling	HPLC-ESI-MS		
	Transcriptomics	RNA-Seq (CASAVA, RseqQC softwares. TopHat 2.0.6, Cufflinks 2.1.1)		
	Primary metabolite profiling	GC-MS		
	GWAS	TASSEL v4.0		
	Genomics	Trimmomatic, Megahit		
	Transcriptomics	RNA-Seq (HISAT2)		
	Volatile profiling	SPME-GC-TOF-MS		
	Carotenoid profiling	HPLC		
QTL mapping	QTL mapping R package			
Volatile profiling	SPME-GC-MS	Alonge et al. (2020)		
Primary metabolite profiling	GC-TOF-MS			
Semi-polar secondary metabolite profiling	UPLC-MS	Alonge et al. (2020)		
QTL mapping mapping	GENSTAT software			
Volatile profiling	GC-MS			
Transcriptomics	RNA-Seq (STAR, HTSeq-count, DESeq2 R package)			
Genomics	NGMLR, Sniffles, RaGOO			

GC: gas chromatography; TOF: time-of-flight; MS: mass spectrometry; LC: liquid chromatography; ESI: electrospray ionization; MS/MS: tandem mass spectrometry; HPLC: high-performance LC; NIR: near-infrared reflectance; UPLC: ultra-performance LC; APCI: atmospheric-pressure chemical ionization; RP-HPLC: reverse-phase HPLC; ³H-NMR: proton nuclear magnetic resonance; SPME-GC-MS: solid phase microextraction GC-MS

the genetic control of certain metabolites may be conserved between plant organs (Wen et al. 2018), subspecies (Chen et al. 2014) or even species, as demonstrated by a comparative GWAS between the two major crops maize and rice (Chen et al. 2016). In this sense, integrative approaches, combining metabolomics with functional genomics, may help crop rapid improvement. In the next paragraphs, we will focus on loci and gene identification for amino acid content in major crops, an essential nutritional trait to secure global food.

Indeed, there is an increasing demand for major crops with enhanced nutritional quality, to both sustain a growing population and improve human health (Galili et al. 2002). In particular, improving the content of minerals, vitamins and proteins of major crops, which are the main nutrient sources in developing countries, is a key strategy to solve severe widespread health problems associated with nutritional deficiencies (Deng et al. 2017; Galili et al. 2002; Galili and Amir 2013). For example, the seed protein content of most consumed cereals, such as maize or wheat, is directly associated with their nutritional quality, and, as animal proteins are more expensive, part of the humanity depends on it to fulfil this dietary requirement (Mandal and Mandal 2000). While protein content depends on the availability of free amino acids, breeding improvement has been hindered due to the tight regulation of their synthesis and to a highly efficient catabolic rate (Galili et al. 2016). An additional difficulty is the limited availability of some essential amino acids, which cannot be synthesized by humans, in most cereals and legumes. (Galili and Amir 2013). In particular, the limited availability of lysine, methionine, threonine or tryptophan in most crops made necessary the identification of the genes regulating these metabolites, and is a key step for breeding improvement and crop biofortification as a critical strategy to struggle against global health nutritional deficiencies (Deng et al. 2017; Galili et al. 2016; Galili and Amir 2013; Yang et al. 2020).

GWAS and linkage mapping studies were combined to examine the genetic architecture of amino acids in the maize (*Zea mays*) kernel, with a special focus on essential amino acids, using a panel of diverse inbred lines and different recombinant inbred line populations (Deng et al. 2017; Wen et al. 2015, 2018). While Deng et al. (2017) were able to outline 308 candidate genes, that could be associated to 528 identified loci, Wen et al. (2018) found 153 significant loci related to the content of 61 primary metabolites, including 23 amino acids. Furthermore, it appeared that epistatic interactions may play an important role in controlling free amino acid content, as shown by Loudet et al. (2003) and Wen et al. (2015). Epistasis can be explained by physical or functional interactions between genes or gene products, such as protein–protein

interactions or transcription factors (TF) controlling the expression of structural genes (Wen et al. 2015).

Interestingly, only 27% of the candidate genes identified by Deng et al. (2017) encodes enzymes or proteins which affect amino acid metabolism, and the function of up to 35% was unknown. Similarly, only a small amount of the 153 loci were known or well characterized previously to the study published by Wen et al. (2018), confirming the necessity of mGWAS and association studies for the identification of natural variation for metabolic traits. Once loci are identified, with underlying putative causal genes, further validation of these candidates is fundamental, before marker development for breeding programs. Due to the usual high number of identified loci, this downstream analysis can be complex, for which it is highly important to prioritise the selection of the putative candidates, according to strong evidence.

These evidences may include expression QTL analysis (eQTL), as undergone by Deng et al. (2017), which took advantage of previously published RNA-sequencing data of young maize kernels (Fu et al. 2013). Interestingly, 16.2% of the 308 candidate genes were found to possibly affect amino acid variation via transcriptional regulation, as significant correlations between gene expression, identified eQTL and metabolite content were established (Deng et al. 2017). Further evidences may consist of gene functional annotation matching corresponding metabolite of the direct association between gene expression levels and metabolite content (Wen et al. 2018). Wen et al. (2015) noticed that most candidate genes they found were only putatively annotated; however, their genetic mapping fell together with gene predicted functions in many cases, including candidate for essential amino acid content. In particular, six genes involved in the aspartate-derived amino acid pathway were proposed as candidate genes, as they co-localized with QTL for lysine, methionine and homoserine.

Among the promising candidate genes found to be associated with essential amino acid content by GWAS studies in maize kernels were an *arogenate dehydratase*, responsible for phenylalanine accumulation, and *Opaque2 (O2)*, a basine leucine zipper (bZIP) TF involved in the regulation of the expression of several genes during kernel development, and associated with lysine/total amino acid ratio (Table 2) (Deng et al. 2017; Wen et al. 2015, 2018). A 811-bp insertion was found in the 5' untranslated region of the *arogenate dehydratase* with an allelic frequency of 46% and explaining both the decrease in gene expression and in phenylalanine content (Wen et al. 2018). The Lys/total amino acid ratio was significantly correlated with *O2* expression, for which a very strong *cis*-eQTL was also found. Furthermore, a duplication present in chromosome 7 was also seen to affect

Table 2 Candidate genes involved in nutritional and organoleptic crop quality traits

Candidate gene	Gen function	Specie	Technique	Associated metabolite(s)	Candidate validation	Reference
<i>Arogenate dehydratase</i>	Aromatic amino acid synthesis	Maize	GWAS	Phenylalanine	InDel in 5'UTR correlated with lower gene expression and Phe content	Wen et al. (2015, 2018)
<i>Opaque2</i>	bZIP TF	Maize	GWAS / QTL mapping	Lysine/total amino acid ratio	<i>Cis</i> -eQTL identification	Deng et al. (2017); Liu et al. (2017)
<i>Acetolactate synthase 1</i>	Branched-chain amino acid metabolism	Maize	GWAS / QTL mapping	Leucine/total amino acid ratio	<i>Cis</i> - and <i>trans</i> -eQTL identification	Deng et al. (2017)
<i>Tryptophan decarboxylase</i>	Tryptamine synthesis from tryptophan	Bread wheat	GWAS	Tryptamine	In vitro assay	Peng et al. (2018)
<i>Lin5</i>	Cell wall invertase	Tomato	GWAS / QTL mapping	Sugar content (glucose, fructose, SSC)	Amino acid change in the protein sequence / Stable silencing and overexpression of the two different alleles	Bauchet et al. (2017); Sauvage et al. (2014); Tieman et al. (2017)
<i>Unknown gene</i>	Unkown	Tomato	GWAS	Malate	Polymorphism in gene sequence	Sauvage et al. (2014); Bauchet et al. (2017)
<i>Aluminium-activated malate transporter-like</i>	Malate accumulation/fruit acidity	Tomato	GWAS	Malate	Nonsynonymous mutation in protein transmembrane domain	Sauvage et al. (2014); Bauchet et al. (2017)
<i>Aluminium-activated malate transporter9</i>	Malate accumulation	Tomato	GWAS	Malate	3pb deletion in promoter sequence resulting in higher gene expression / gene editing by CRISPR	Sauvage et al. (2014); Ye et al. (2019)
<i>SibHLH59</i>	Regulation of D-man/L-Gal ascorbate biosynthetic pathway	Tomato	GWAS / QTL mapping	Ascorbate	8-pb-insertion in the promoter resulting in higher downstream gene expression / Stable RNAi silencing	Sauvage et al. (2014); Ye et al. (2019)
<i>TomLoxC</i>	Lipid-derived C5 and C6 volatiles	Tomato	QTL mapping	Lipid-derived and apocarotenoid volatiles	4-kb substitution in promoter region / Stable silencing	D'Angelo et al. (2018, 2019); Gao et al. 2019
<i>FLORAL4</i>	Phenylalanine decarboxylation	Tomato	QTL mapping	Phenylpropanoid-derived volatiles	Gene editing by CRISPR	Tikunov et al. (2020)
<i>E8 and NSGT1</i>	Non-smoky glycosyltransferase	Tomato	GWAS / omic approaches	Phenylpropanoid-derived volatiles	Haplotype analysis confirmed multiple mutant alleles for <i>nsgt1</i> / F ₂ population segregating for the 23-kb deletion removing <i>E8</i> and <i>NSGT1</i>	Alonge et al. (2020); Tieman et al. (2017); Tikunov et al. (2013, 2020)

Table 2 (continued)

Candidate gen	Gen function	Specie	Technique	Associated metabolite(s)	Candidate validation	Reference
<i>APPR2</i>	Ripening-related TF	Melon, watermelon, cucumber, tomato, pepper	Linkage analysis	Chlorophyll and carotenoid pigments	Haplotype analysis confirmed multiple independent polymorphisms. Stable overexpression in tomato	Liu et al. (2016); Oren et al. (2019); Pan et al. (2013)
<i>Orange</i>	Chromoplast biogenesis	Sweet potato, melon, carrot	QTL mapping / GWAS	Carotenoid pigments	Haplotype analysis with polymorphisms leading to amino acid changes / Site-directed mutagenesis of <i>CmOr</i> in <i>Arabidopsis</i> calli	Ellison et al. (2018); Gemenet et al. (2020); Tzuri et al. (2015)

lysine content and its ratio to total amino acids and was identified both by GWAS in the diverse inbred lines and by QTL mapping (Deng et al. 2017; Liu et al. 2017). The expression of five genes, located within the duplicated region, were found to be significantly correlated with lysine to total amino acid ratio. Additionally, haplotype analysis for one of the genes showed significant differences both for Lys/total amino acid and expression. Combination of duplication and haplotype variants had greater effect on phenotype than single variation (Deng et al. 2017). Another candidate, an acetolactate synthase 1, involved in branched-chain amino acid metabolism, was found to be associated with leucine/total amino acid (Table 2). Interestingly, two eQTL were detected for this gene, including a *cis*- and *trans*-eQTL, which results to be O_2 (Deng et al. 2017). Taken together, the natural variation underlying the above mentioned loci and genes can be used for the improvement of essential amino acid composition of cereal crops.

GWAS for amino acid content was also performed in bread wheat, rice or soybean (Chen et al. 2016; Lee et al. 2019; Peng et al. 2018; Qin et al. 2019; Sun et al. 2020). A highly diverse panel of 182 accessions of *Triticum aestivum* was used to map 328 significant quantitative trait nucleotides with six different multi-locus models. Based on chemical structure and previous knowledge of the synthetic pathways, 15 candidate genes could be tentatively associated to free amino acids levels, in particular six genes annotated as amino acid transporters and permeases co-localized with identified loci. Furthermore, the function of a tryptophan decarboxylase, found in a locus associated with tryptamine levels, was validated in vitro as a proof of concept (Table 2) (Peng et al. 2018). Two genes annotated as tryptophan decarboxylases have also been associated with tryptophan content in maize kernel (Wen et al. 2014). Very recently, a combination of GWAS and functional analysis in rice leaves of 520 accessions validated a bZIP TF, *OsZIP18*, as being the main genetic determinant for branched-chain amino acid (BCAA) content in rice (Sun et al. 2020). As humans and animals are not able to synthesize BCAA, *OsZIP18* may be a promising candidate for increasing rice nutritional value; however, *OsZIP18* is mainly expressed in leaf tissues, for which its impact on grain BCAA content needs further validation. In soybean (*Glycine max*), GWAS for seed composition in 321 accessions, including amino acid content, pinpointed a major-effect QTL on chromosome 8, harbouring an aspartokinase-homoserine dehydrogenase, a key enzyme involved in the synthesis of aspartate-family amino acids (Zhang et al. 2018). Interestingly, the same study suggests that free amino acid profiles in soybean seeds may be under a different genetic control from total protein levels, implying that the former may

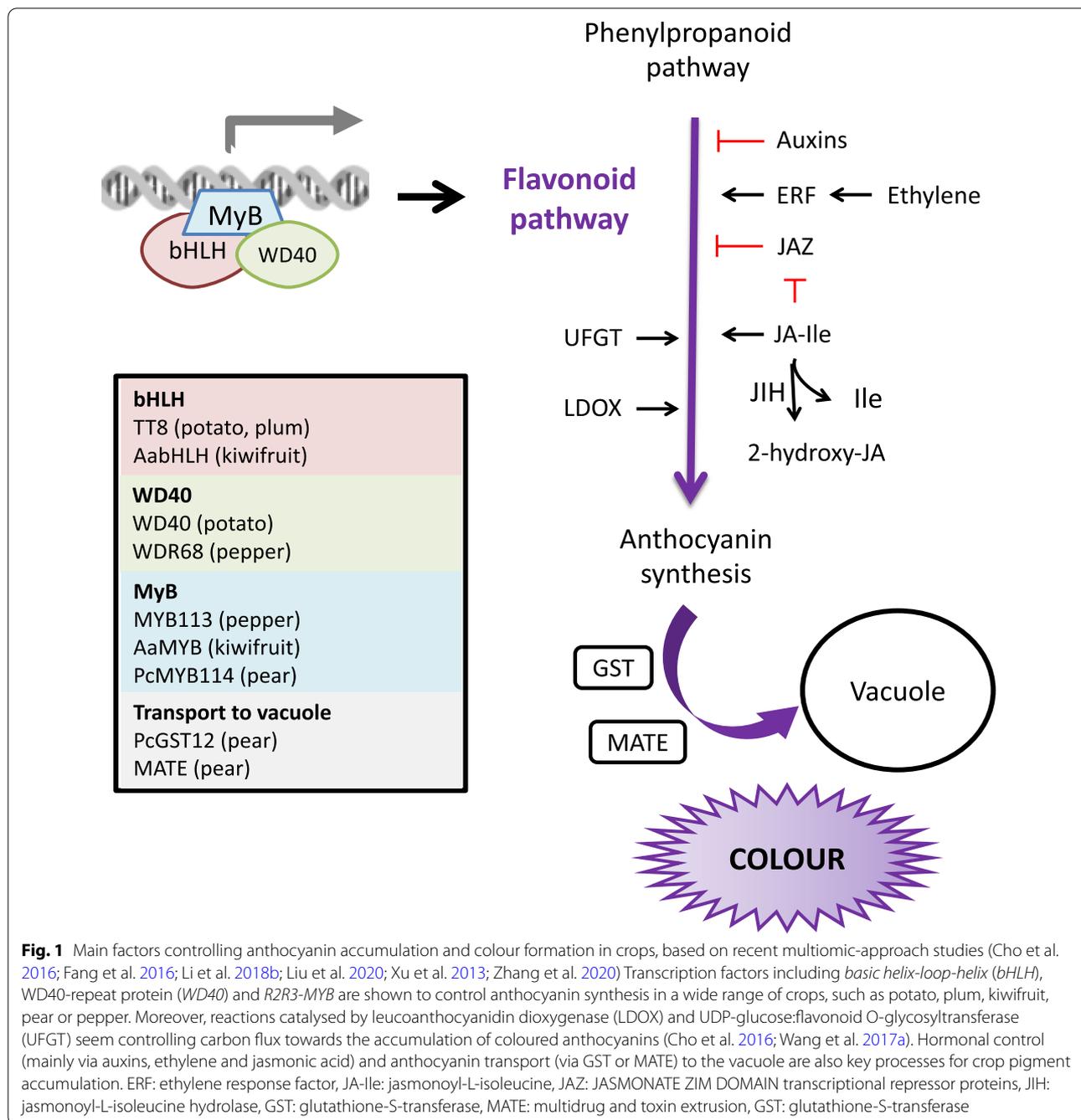
be improved without affecting protein levels. Indeed, as 92 amino acid-associated QTL were detected, only four were common with both dry weight- and protein-based amino acid content (Zhang et al. 2018). (Qin et al. (2019) detected 15 amino acid-associated SNPs located near 14 candidate genes in a GWAS analysis comprising 249 soybean accessions, which will be further validated with the objective to develop molecular markers for breeding purposes.

Transcriptomic and metabolomic platforms for improved crop colour

Anthocyanin pigments

Another group of plant metabolites which are particularly attractive for breeding programs are pigments, responsible for the appealing aspect of many ripe crops. In particular, anthocyanins have widely attracted research interest, for conferring red, blue, pink or purple colours to a large number of fruits and vegetables (Pott et al. 2019). Also, they are important for their health-promoting role due to their high antioxidant capacity and their ability to modulate mammalian cell signalling pathway (Butelli et al. 2008; Giampieri et al. 2015; Petrusa et al. 2013; Pott et al. 2019). In fact, engineering methodologies for increased anthocyanin content have been proposed in crops with suboptimal concentration of these metabolites, such as tomato fruits (Butelli et al. 2008; Zhang et al. 2013, 2015). Anthocyanin synthesis through the phenylpropanoid and flavonoid pathways has been widely studied (Vogt 2010; Fraser and Chapple 2011; Hassan and Mathesius 2012). However, due to the complexity and diversity of plant secondary metabolism, i.e. more than 630 anthocyanins have been identified (He and Giusti 2010), OMIC data set integration may be an optimal strategy to identify potential genes involved in pigment formation regulation. In particular, metabolomic platforms, using ultrahigh-performance liquid chromatography coupled with mass spectrometry is the technique of choice for anthocyanin and other polyphenol metabolite detection in plant tissues (Cho et al. 2016; Liu et al. 2020; Pott et al. 2020; Vallarino et al. 2018). Metabolite profiling combined with transcriptomic and genomic data and multivariate statistical models can be key in revealing the molecular mechanisms associated with crop pigment accumulation, as highlighted in the next section and in Fig. 1.

Combining metabolite profiling with high-throughput RNA-sequencing analysis, Cho et al. (2016) established a correlation network between metabolites and genes involved in the pigmentation of three potato (*Solanum tuberosum*) cultivars, 'Hongyoung' (light-red skin and flesh), 'Jayoung' (dark-purple skin and flesh) and 'Atlantic' (white-coloured). In particular, they focused



on compounds belonging to the flavonoid class, including anthocyanins, *i.e.* delphinidin, petunidin and malvidin-derived metabolites, responsible for purple and dark hues, and derivatives of cyanidin and pelargonidin, mainly associated with bright-red-coloured tones (Jaakola 2013), and on differentially-expressed transcripts between cultivars, belonging to flavonoid metabolism, hormone metabolism, cell signalling or regulation of transcription (Cho et al. 2016). Interaction networks

divided genes and metabolites into five clusters (I to V) and four groups (A to D), respectively, and correlation between cluster and group may indicate functional connection between a particular set of genes and flavonoid metabolites, shedding light on the regulation controlling the specific synthesis of anthocyanins in potato. Common mechanisms leading to the synthesis of anthocyanin may encompass the upregulation of cluster III genes, which expression was enhanced in both coloured

cultivars, and which include *TT8* and *WD40-repeat protein (WD40)*. Interestingly, *TT8* is a bHLH-type regulation factor, which forms a complex with *WD40* and *R2R3-MYB* TF, responsible for the regulation of the flavonoid pathway, and precisely anthocyanin biosynthesis in *Arabidopsis* and other crops (Fig. 1) (Cho et al. 2016; Schaart et al. 2013; Xu et al. 2013). This observation suggests that *TT8*-mediated pathway also regulates pigmentation in coloured potato cultivars. Similarly, an RNA-Seq-based comparative transcriptome analysis allowed to identify TF which were coexpressed with anthocyanin biosynthetic genes during the ripening of the red-fleshed plum 'Furongli' cultivar (*Prunus salicina*). A bHLH gene, homologous to *AtTT8*, was found to accumulate to higher levels in late ripening stages and to show significant correlation with anthocyanin biosynthetic genes (Fang et al. 2016). In pepper (*Capsicum spp.*), some varieties show purple or black tones during the immature period due to the accumulation of anthocyanins (Lightbourn et al. 2008). By combining metabolite profiling with RNA-seq analysis, Liu et al. (2020) built a weighted gene co-expression network, highlighting the strongly connected genes. This way, a *forin-like protein 11*, an unknown gene (*Capana02g003118*), *WDR68* (a *WD40* gene), a *solute carrier family 40 member 1* and *MYB113* may be key genes in flavonoid synthesis in pepper, suggesting a key role of TF and transport in anthocyanin accumulation (Liu et al. 2020).

Hormonal roles in anthocyanin accumulation were also outlined by several transcriptomic and metabolomic studies (Cho et al. 2016; Wang et al. 2017b, a; Zhang et al. 2020). Indeed, most auxin-related genes were negatively correlated to the anthocyanins found in the both coloured potato cultivars, suggesting that their reduced expression may enhance pigment synthesis by removing the inhibition upon the flavonoid pathway. On the opposite, as ethylene-related genes were upregulated in the two pigmented potatoes in comparison with the colourless cultivar, it can be hypothesized that this phytohormone is associated with anthocyanin synthesis (Cho et al. 2016). However, *NAC2* was down-regulated in both coloured cultivars. *NAC* genes are plant-specific TF, and are involved in a myriad of physiological processes, inclusive of fruit ripening and leaf formation and senescence, and stress responses (Shan et al. 2012; Zhang and Gan 2012). Some of them have been described to take part in hormone cascade signalling, including ethylene (He et al. 2005). Indeed, while *NAC2* promoter was shown to be activated by ethylene in banana, *NAC1* and *NAC2* was shown to interact with ethylene insensitive (EIN) protein, a downstream component of ethylene signalling (Shan et al. 2012). *EIN2-NAC2* pathway induces the expression of genes involved in senescence, for which anthocyanin

accumulation in potato tubers may not be senescence-dependent and *EIN2-NAC2* ethylene transduction pathway should be discarded (Cho et al. 2016; Woo et al. 2013). The positive role of ethylene-responsive genes (ERF) upon anthocyanins accumulation has also been outlined in pear and fig (Wang et al. 2017a, b; Zhang et al. 2020).

Specific to light-red potato cultivar compared to the dark-purple one, up-regulation of jasmonic acid (JA) signalling genes was observed by Cho et al. (2016), including *JIH*, *JAZ8* and *JAZ10* genes. JA is known to increase anthocyanin synthesis by favouring the degradation of negative regulators of JA (*JAZ* genes). Up-regulation of *JIH* in 'Hongyoung' cultivar, which catalyses the cleavage of JA-isoleucine, the JA conjugate which promotes *JAZ* degradation, combined with enhanced transcripts of *JAZ8* and *JAZ10* suggests that *JAZ* degradation-mediated anthocyanin biosynthesis is possibly inactive in the light-red cultivar compared to 'Jayoung', and may explain the lower colouration in the first one (Cho et al. 2016; Thines et al. 2007). Furthermore, the increased expression of several enzymes involved in flavonoid first-step reactions, a *UDP-glucose:flavonoid O-glycosyltransferase (UGFT)* and a *leucoanthocyanidin dioxygenase (LDOX)* in 'Jayoung' may explain how metabolic flux is favoured towards the synthesis of delphinidin- and petunidin-derived anthocyanins, conferring its purple tone (Cho et al. 2016). *LDOX* and *UGFT* genes have also been pointed out as being predominant structural genes for red-colour formation of kiwifruit (*Actinidia arguta*) flesh, together with *AaF3H*, a flavanone 3-hydroxylase (Li et al. 2018b). A model was proposed in which their expression was controlled mainly by *AaMYB*, *AabHLH* and *AaHB2* TF (Li et al. 2018b).

In pears (*Pyrus communis*), bud mutation is an important method for the selection of new varieties, including the natural occurrence of red-skin fruits, accumulating anthocyanins and proanthocyanidins (condensed tannins) in the peel (Zhang et al. 2020). A comparative metabolic and transcriptomic analysis of the young fruit of 'CF' cultivar and its red mutant 'RCF', whose colouration appears as early as five days after full bloom, may facilitate the understanding of the mechanisms underlying pigment accumulation due to the high similarity between the two fruits. Twenty candidate genes associated with anthocyanins and proanthocyanidins, which shared common flavonoid precursors, were outlined, including six structural flavonoid genes and four *R2R3-MYB* genes. Interestingly, while one the *MYB* candidate, *PcMYB114*, has been previously known to be involved in anthocyanin synthesis in pears (Yao et al. 2017), this approach allowed to outline upon which specific anthocyanins it acts and to newly identify three *MYB* genes associated

with this process (Zhang et al. 2020). Furthermore, other TF found to be associated with anthocyanin and proanthocyanidins accumulation were *PcKNAT1*, *PcbZIP1* and *WRKY28* (negatively correlated). The latter two have been previously described to be involved in anthocyanin accumulation in other species, though *WRKY28* putative function as a repressor was found to be the opposite in apple (An et al. 2019). *PcKNAT1* belongs to the class I KNOX HD gene family, known to be involved in meristem development and leaf morphogenesis (Lincoln et al. 1994). Although a role in anthocyanin synthesis has been described for TF belonging to other HD subfamilies, the identification of *PcKNAT1* as a candidate gene for this process is novel (Zhang et al. 2020). Finally, a glutathione-S-transferase, *PcGST12*, was found to be correlated with seven anthocyanins and seven proanthocyanidins and is functionally characterized as an anthocyanin and proanthocyanidin transporter from cytosol to vacuole. Zhang et al. (2020) found that *PcGST12* expression is positively regulated by *PcMYB114* and stable overexpression of *PcGST12* in *Arabidopsis* pointed out that *PcGST12* may be involved not only in transport, which is also a key process in determining anthocyanin levels, but in other steps of anthocyanin and proanthocyanidin biosynthesis.

The role of anthocyanin transport has also been highlighted as a key factor controlling pigment accumulation in other crops, such as carrots (Meng et al. 2020), or in two different red-flesh pears varieties, 'Red Bartlett', which shows a decrease in anthocyanin content along ripening concomitant with colour fading, and 'Starkrimson', which maintain purple-red colour throughout maturation (Wang et al. 2017b). Indeed, six *GST* and two multidrug and toxic compound extrusion transporters, also involved as flavonoid carrier, were downregulated at the later stage of 'Red Bartlett' ripening and could explain at least partially colour fading mechanism. In addition, Wang et al. (2017b) highlighted the possible role of two *LDOX* genes and a flavonol synthase (*FLS*) in this differential behaviour between the two varieties. While *LDOX* genes were downregulated during 'Red Bartlett' ripening, *FLS* was upregulated and its expression was negatively associated with anthocyanin content. In addition, several TF were also outlined as possible candidates in the differential colouration pattern between the two varieties, including *MYB*, *bHLH*, *NAC*, *WRKY* and *ERF* TF, however, no clear candidate could be highlighted. Finally, two peroxidases and an intracellular lacase, with a possible function in anthocyanin degradation, were differentially expressed between the two cultivars. Interestingly, comparing the transcriptome and metabolome of three varieties of a wild peach species (*Prunus mira*), with different flesh pigmentations, Ying et al.

(2019) also highlighted a peroxidase, which expression was strongly enhanced in the white-fleshed pear. This enzyme has been associated with anthocyanin degradation in several fruits, and this result suggests that besides TF and structural gene downregulation, anthocyanin enhanced degradation favoured white-flesh colouration in wild peach (Ying et al. 2019).

Apart from pears, colour mutation is also common in other important crop trees, such as fig (*Ficus carica*). Four cyanidin glucosides, identified in 'Purple Peel' fig cultivar, were responsible for the colour difference between 'Green Peel' cultivar and its mutant. Interestingly, accumulation of other colourless flavonoids was also observed in the mutant, suggesting increased nutraceutical properties (Wang et al. 2017a). The authors of the study also outlined that even 'Purple Peel' is considered the mutant of 'Green Peel', it should probably be considered the first one as the wild type condition, based on the fact that the seeds inside fig syconia are bird-dispersed, and fruits with this characteristic are generally black or red (Willson and Whelan 1990). In this sense, green-fruit phenotypes can be understood by a functional loss of key enzymes in the flavonoid or anthocyanin pathway or by mutations, such as retrotransposon insertion, in key regulatory genes, such as *MYB* TF, as it has been described for white grapes (Kobayashi et al. 2004; Wang et al. 2017a). Interestingly, differential expression of transposons and retrotransposons were observed between 'Green Peel' and 'Purple Peel' cultivars, and in the first one a general upregulation of reverse transcriptases, integrases and gag sequences could be noticed (Wang et al. 2017a).

Another regulating mechanism underlying anthocyanin accumulation during fruit ripening may involve sugar signalling cascade (Dai et al. 2014; Jia et al. 2016; Luo et al. 2019). In particular, carbon starvation may affect pigment accumulation in red-fleshed kiwifruits (Nardoza et al. 2020). Indeed, metabolite profiling combined with RNA-Seq analysis in fruits grown under low (one leaf per fruit) or high (four leaves per fruit) carbon supply showed a general downregulation of genes and metabolites from both carbohydrate and anthocyanin biosynthetic pathways. In particular, the sugar trehalose-6-phosphate, a sucrose sensor and regulator, was proposed as a signalling molecule reflecting the levels of imported sugars to heterotrophic tissues (Figuroa and Lunn 2016). Trehalose-6-phosphate may act upon *MYB27*, a repressor of anthocyanin synthesis, negatively controlling its expression via a *cis*-acting sugar-repressive element present in *MYB27* promoter (Nardoza et al. 2020). In this sense, it was proposed that under carbon starvation, decreased levels of trehalose-6-phosphate mirroring carbon status favoured *MYB27*-mediated repression upon anthocyanin synthesis (Nardoza et al. 2020).

Carotenoid pigments

Carotenoid are the second largest group of pigments, after anthocyanins, contributing to the red, orange or yellow tones of many fruits and crops. They play important roles *in planta*, such as photoprotection, and serves as precursors of a series of physiological-important molecules, both for plants (strigolactones, abscisic acid or apocarotenoid volatiles, among others) and humans (provitamin A) (Ashokkumar et al. 2020; Pott et al. 2019; Sun et al. 2018). While QTL mapping for carotenoid genes and underlying candidate genes has been reviewed in cereal crops such as durum wheat (Colasuonno et al. 2019) and will not be discussed here, GWAS and linkage analysis have recently shed light on the genetic control of carotenoid metabolism in important fruit and crop species, and will be summarized in this section. In addition, current QTL mapping and GWAS in important crops, such as maize or chickpea, remains a valid approach to unravel carotenogenesis genetic architecture (Azmach et al. 2018; Esuma et al. 2016; Owens et al. 2014; Rezaei et al. 2019; Xu et al. 2019; Zhai et al. 2018). However, most genes underlying pinpointed QTL are still unknown, and further studies are needed.

Rind and flesh melon (*Cucumis melo*) colours are important quality traits, linked to consumer liking, and which depends on chlorophyll/carotenoid synthesis and accumulation. Surprisingly, GWAS performed on 120 melon accessions with contrasting rind colours yielded no significant association between pigment content and genetic markers (Oren et al. 2019). However, whole-genome linkage analysis in two segregating bi-parental populations led to the identification of unique highly significant locus on chromosome 4 (Oren et al. 2019). Sequencing of an *Arabidopsis pseudo-response regulator2-like* (*CmAPPR2*) gene, found within the locus confidence interval, outlined five major polymorphisms, including two independent base substitutions leading to premature stop codons in the two light-rind parents (Table 2). Of particular interest, carotenoid profiles in the melon flesh were also associated with *CmAPPR2* haplotypes. Furthermore, *APPR2* overexpression in tomato resulted in higher levels of both chlorophyll in unripe fruits and carotenoid pigments in ripe fruits (Pan et al. 2013); in addition, white immature fruit colour in cucumber (*Cucumis sativus*) was consequence of a frameshift mutation leading to a premature stop codon in *APPR2* open reading frame (Liu et al. 2016). These multiple and independent polymorphisms found in *CmAPPR2*, and their low frequency, could explain the lack of GWAS success in pinpointing any associated locus. While this example outlines the need to improve GWAS resolution for low-frequency causative variants, alternative layers of information obtained from OMIC data integration

can be used as a valuable alternative approach to identify important loci (Lee and Lee 2018). Interestingly, the watermelon (*Citrullus lanatus*) homolog *ClAPRR2* was also mapped by whole-genome linkage analysis and a putative causative SNP, leading to a 16-pb deletion in the mRNA of the light rind parental line, was identified (Oren et al. 2019). Watermelon red flesh intensity, which depends on the accumulation of the red carotenoid lycopene, is also considered an important quality trait. Whole genome variation analysis in populations derived from scarlet- and coral-red fleshed accessions mapped a 40 kbp region on chromosome 6 responsible for lycopene variation. Strikingly, three unknown and four glycine-rich cell wall structural genes were found in the confidence interval, and showed different expression along fruit ripening (Li et al. 2020). While two InDel markers were able to explain the observed phenotypic variance and could be useful for watermelon breeding programs, the underlying gene and the molecular mechanism responsible for differential lycopene accumulation in the flesh still remains elusive (Li et al. 2020). In another QTL mapping study, a major locus was identified on chromosome 4, with a lycopene β -cyclase being the most probable candidate gene underlying the QTL (Wang et al. 2019). A SNP variation leading to a nonsynonymous substitution in the candidate amino acid sequence may explain differences in protein levels or functionality between red- and yellow-flesh lines (Wang et al. 2019). Out of curiosity, some watermelon accessions accumulate mainly the orange carotenoid β -carotene, the main source of provitamin A, instead of lycopene (Tadmor et al. 2005). Genome-wide mapping in a population segregating for flesh colour led to the identification of a single major QTL on chromosome 1 (Branham et al. 2017). No clear candidate gene could be identified, however, several TF were located within the confidence interval. Furthermore, a second still unidentified locus acting epistatically with the major QTL may explain flesh colour segregation in watermelon (Branham et al. 2017).

Sweet potato (*Ipomoea batatas*) orange-flesh varieties are able to accumulate high levels of β -carotene. However, a negative correlation between this important source of provitamin A and starch is observed, due to the linkage between phytoene and sucrose synthase enzymes, which catalyse the rate-limiting reactions of carotenoid and starch biosynthetic pathways, respectively. Indeed, QTL mapping in a biparental F_1 progeny identified two loci on linkage groups 3 and 12 affecting both β -carotene and starch content (Gemenet et al. 2020). Interestingly, both phytoene and sucrose synthases were found within the locus interval on chromosome 3, suggesting that physical linkage, instead of pleiotropic effects, was responsible for the negative correlation between starch and carotenoid

accumulation. In addition, fine mapping on chromosome 12 allowed to pinpoint the *Orange (IbOr)* gene as the candidate underlying the identified QTL (Gemenet et al. 2020). *Or* gene is involved in chromoplast biogenesis and, additionally, positively affects phytoene synthase protein levels (Park et al. 2016; Sun et al. 2018; Zhou et al. 2015). As no correlation between *IbOr* expression and β -carotene content was observed in sweet potato, it can be hypothesized that nucleotide polymorphisms leading to protein changes may impact chromoplast biogenesis and phytoene synthase levels (Gemenet et al. 2020). In fact, melon *Or* gene, which co-segregate with flesh colour in F_2 and backcross populations, present two major haplotypes, with a single SNP inducing an amino acid change from a highly conserve arginine to a histidine and responsible for β -carotene accumulation (Table 2) (Tzuri et al. 2015). Carrot (*Daucus carota*) *Or* gene was also identified by GWAS analysis in 674 domesticated and wild carrot accessions. A nonsynonymous mutation was found in *Or* exon 5, causing a serine to leucine change in the amino acid sequence. Interestingly, *Or* sequence fell within a domestication sweeps, suggesting that alleles favouring carotene accumulation have been fixed by artificial selection (Ellison et al. 2018).

Another omic approach, combining metabolite profiling with proteomic analysis to decipher carotenoid accumulation in banana (*Musa* spp.) pulp, is worth mentioning (Heng et al. 2019). Both proteome and metabolome analyses outlined a relation between enhanced glycolysis activity and carotenoid accumulation, suggesting that carotenogenesis may be supported by carbohydrate metabolism. Together, these results imply that multi-target approach may be indispensable to improve carotenoid content in crops (Heng et al. 2019).

Transcriptomic, GWAS and metabolomic approaches to understand flavour quality changes during domestication and intensive breeding: tomato as an example

Tomato (*Solanum lycopersicum*) is an important crop for the human diet, being a source of fiber and micronutrients, in addition to be a study model for fruit development and ripening (Giovannoni 2001; Sauvage et al. 2014). Due to intensive breeding in the last centuries, a reduction of the genetic diversity through two successive bottlenecks has led to the loss of important fruit quality parameters, such as flavour and nutritional value (Bai and Lindhout 2007; Bauchet et al. 2017). As a consequence, consumer complaint in the last decades about flavourless tomatoes has become a major issue for breeding programs, and the need to understand the genetic mechanisms underlying this quality loss is shared by both academic research and food industry. In addition,

the combination of metabolomics, transcriptomics and mGWAS is a key approach to understand the fruit metabolic and quality changes which have occurred along domestication and improvement (Zhu et al. 2018).

QTL approach in biparental populations has shed light on the genetic architecture of a series of valuable traits for fruit quality improvement (Causse 2004; Goulet et al. 2012; Mageroy et al. 2012; Tieman et al. 2006; Tikunov et al. 2013; Zanon et al. 2009b); however, as QTL mapping does not take into account the entire genetic diversity present in germplasm collections, GWAS strategy, which allows the screening of a wide range of accessions, has been more recently used to decipher the inheritance of important fruit metabolic traits (Bauchet et al. 2017; Ruggieri et al. 2014; Sauvage et al. 2014; Tieman et al. 2017; Ye et al. 2019; Zhu et al. 2018). Additionally, combination of transcriptomic and metabolomic profiling also depicted some important genetic regulation of quality traits (D'Angelo et al. 2019; D'Esposito et al. 2017).

GWAS for primary metabolites, in particular, sugars and organic acids, have been performed in several hundreds of tomato genotypes, including domesticated *S. lycopersicum*, wild *S. lycopersicum* variant *cerasiforme* and the most closely related wild species *S. pimpinellifolium* accessions (Bauchet et al. 2017; Sauvage et al. 2014; Tieman et al. 2017; Ye et al. 2019). The identification of candidate genes underlying the associated loci can be crucial for modern tomato flavour improvement, the main consumer complaint, as tomato flavour is defined by sugars, acids and a specific set of volatile compounds (Baldwin et al. 2008; Tieman et al. 2012).

Sauvage et al. (2014) identified 44 loci for 19 primary metabolic traits, some of the peak SNPs targeting previously described or putative candidate genes. Indeed, in one of the detected loci associated to soluble solid content was found *lin5* gene, which encodes a cell wall invertase, a key enzyme for fruit sugar uptake and a major locus controlling sugar content in this organ (Table 2) (Fridman et al. 2004). *Lin5* was previously identified in a QTL study for fruit sugar content and further validated by stable silencing in tomato plants (Baxter et al. 2005; Fridman et al. 2004; Vallarino et al. 2017; Zanon et al. 2009a). Bauchet et al. (2017), which also implemented GWAS for both primary and secondary metabolites in 300 tomato accessions and detected 79 loci for 33 metabolic traits, were able to validate *lin5* gene as being responsible for fructose, glucose and soluble solid content. Furthermore, Tieman et al. (2017) identified by GWAS performed in 398 tomato accessions two loci, on chromosomes 9 and 11, associated with glucose and fructose levels. Interestingly, they detected the same locus on chromosome 9, and which corresponds to *lin5* gene by QTL mapping in a F_2 population, derived from

a cross between a variety producing small but flavourful fruits and a modern large-fruit inbred line (Tieman et al. 2017). This strong candidate for tomato sugar content was validated at molecular level, as the *lin5* SNP resulted in an amino acid change (Asn to Asp) in the protein sequence. Stable transgenic tomato plants overexpressing one or the other *lin5* variants allowed to confirm that significant differences in sugar content were observed based on the amino acid change in *lin5* sequence (Tieman et al. 2017). Furthermore, an explanation about flavour loss along domestication could be provided, as the two hexose-associated loci on chromosomes 9 and 11 were located within domestication and improvement sweeps. By continuously selecting for fruit size, breeders have set the combination of alleles at the two loci which result in the lowest sugar content, and which are found in almost all the modern cultivars included in the GWAS analysis (Tieman et al. 2017).

Apart from main sugars present in ripe tomatoes, the predominant organic acids, malate and citrate, are essential for fruit taste perception (Etienne et al. 2013). Interestingly, Sauvage et al. (2014) identified two significant SNPs located on chromosomes 2 and 6 linked to malate content, explaining 39% of the metabolic variation. Interestingly, the SNP on chromosome 6 was found within the coding sequence of an unknown protein, but two aluminium-activated malate transporters-like (likely a unique gene) were detected in the vicinity (Bauchet et al. 2017; Sauvage et al. 2014). Deeper analysis of chromosome 6 candidate region indicated that the aluminium-activated malate transporter-like expression correlated with malate content in eight tested accessions, and that a nonsynonymous mutation in a transmembrane domain may explain this correlation (Table 2). While a polymorphism in the gene encoding the unknown protein cannot be discarded in impacting malate levels in the assessed accessions (Bauchet et al. 2017), another GWAS analysis in 272 tomato accessions, combined with linkage mapping and transgenic complementation approaches, allowed to outline that expression of *aluminium-activated malate transporter9* on chromosome 6 is the main determinant for tomato fruit malate levels in cultivated accessions (Ye et al. 2017). Indeed, a 3pb deletion in the promoter sequence resulted in higher *aluminium-activated malate transporter9* expression, increasing fruit malate content (Ye et al. 2017, Table 2).

Another important acid, for its essential function in human health and which has to be obtained through fruit and vegetable ingestion is ascorbate (vitamin C). Ascorbate synthesis takes place mainly through the D-Mannose/L-Galactose pathway, also known as the Smirnoff-Wheeler (SW) pathway, although alternative routes have also been proposed, through *myo*-inositol

and D-galacturonate. In fruits, both the SW and the galacturonate pathways have been described to be functional, depending on the ripening stage (Fenech et al. 2019). Interestingly, in tomato, ascorbate levels showed the most significant differences between *S. lycopersicum*, *S. lycopersicum* variant *cerasiforme* and *S. pimpinellifolium* accessions, being higher in the wild species genotypes (Sauvage et al. 2014; Ye et al. 2019) and suggesting the possibility of lost alleles during domestication, which could be used by breeding programs to improve this important trait. In fact, a SNP associated with fruit ascorbate content was found 423 kb upstream of a monodehydroascorbate reductase (NADH)-like protein on chromosome 9 (Sauvage et al. 2014). More recently, Ye et al. (2019) identified a SNP associated with ascorbate content on chromosome 9 by GWAS on 302 tomato accessions, and explaining 15.9% of the variation in this metabolic trait. By measuring the expression of possible candidate genes in accessions showing contrasting ascorbate levels, a bHLH TF, *SlbHLH59*, could be highlighted as the likely candidate underlying this locus. Furthermore, a deeper analysis of the gene sequence outlined four different haplotypes, with haplotype 4, found in most *S. pimpinellifolium* tested accessions, harbouring nine consensus polymorphisms. In particular, an 8 pb-insertion in the promoter region resulted in the formation of a 5'UTR Py-rich stretch motif associated with elevated expression of downstream genes. The candidate was then functionally validated by stable RNAi silencing in tomato plants (Table 2) (Ye et al. 2019). In addition, expression analysis in *SlbHLH59*-silenced and overexpressed lines and yeast two-hybrid allowed to establish that *SlbHLH59* protein can directly bind to the promoter region of enzymes of the D-Man/L-Gal ascorbate biosynthetic pathway and modulate their expression during tomato fruit development (Ye et al. 2019). Decreased ascorbate fruit content was selected during domestication, as suggested by the analysis of the InDel present in *SlbHLH59* promoter sequence in more than 500 accessions. Indeed, the insertion frequency was high in *S. pimpinellifolium* genotypes (83.3%), and sharply decreased in *S. lycopersicum* *cerasiforme* variants (16.3%) and modern varieties (1.8%). The low nucleotide diversity in the latter ones further confirmed that *SlbHLH59* was located in a domestication and improvement sweep (Lin et al. 2014).

Ruggieri et al. (2014) performing a GWAS on landraces, vintage and modern *S. lycopersicum* varieties, found three associations for ascorbate content, different from the ones identified by Sauvage et al. (2014) and Ye et al. (2019). In the genomic region in linkage disequilibrium with one of the SNP (on chromosome 3) a cluster of *pectinesterases*, *pectate lyase* and *polygalacturonase*

genes were identified. Furthermore, another SNP on chromosome 5 was found to be associated with ascorbate content and was located in the sequence of an *ethylene-responsive transcription factor 1* (Ruggieri et al. 2014). Interestingly, a previous study using introgression lines from a cross between *S. lycopersicum* 'M82' and *S. pennellii*, outlined a *pectinesterase* and two *polygalacturonases* to be upregulated in a line with higher levels of ascorbate (Di Matteo et al. 2010). Furthermore, several ethylene-related transcripts showed increased expression in this line, suggesting that ascorbate accumulation in tomato fruit may occur through pectin degradation and is triggered by ethylene signalling, which corroborate the GWAS analysis by (Ruggieri et al. 2014). Taken together, these data imply that both the predominant SW pathway and the alternative D-galacturonate pathway are important for the regulation of tomato fruit ascorbate content.

A combined metabolomic and transcriptomic study in Andean landrace accessions and the wild specie *S. pimpinellifolium* shed light upon the domestication of important quality traits, such as fruit aroma. In particular, significant differences between the wild ancestor and the Andean varieties were observed in the content of branched-chain amino acids and threonine, important precursors of tomato aroma volatiles. In particular, isoleucine and its substrate, threonine, was increased in five Andean accessions, and correlated with a lower expression of threonine aldolase gene, favouring both essential amino acid content and aroma formation (D'Angelo et al. 2019). Furthermore, one of the accessions showed a remarkable flavour when compared to the other genotypes; interestingly, a gene involved in volatile synthesis, *TomLoxC* was highly upregulated in this cultivar (D'Angelo et al. 2019, 2018). In fact, *TomLoxC* was outlined in a QTL study in a recombinant inbred line population, as controlling lipid- and carotenoid-derived volatile content (Table 2) (Gao et al. 2019). A nonreference allele for *TomLoxC* was identified in a pan-genome analysis, present in 91.2% of *S. pimpinellifolium* accessions and only in 2.2% of *S. lycopersicum* heirlooms, implying a strong negative selection during domestication. Interestingly, modern tomato varieties showed increased frequency of the rare allele (7.2%) compared to heirlooms, possibly as a consequence of recent introgressions from wild species into breeding cultivars (Gao et al. 2019).

Among volatiles impacting tomato aroma, the phenylpropanoid phenylethanol and its precursor phenylacetaldehyde have a positive effect on the fruit flavour, conferring sweet and fruity aromas (Baldwin et al. 2008). A highly significant association was found for phenylethanol and phenylacetaldehyde on chromosome 4. In this region, a glucosyltransferase was identified as the most

probable candidate gene (Bauchet et al. 2017). Indeed, an allelic variation on the third exon of the gene appeared to be nonsynonymous, and significant variation of glycosides were observed according to the different alleles. Taken together, this result suggests that the release of volatiles may be regulated by glycosylation. Interestingly, glycosyltransferases have been previously described to play an important role in volatile conjugation (Bauchet et al. 2017; Tikunov et al. 2013; Yauk et al. 2014). However, further validation of the candidate is needed to confirm its impact on the two volatiles.

A recent study also mapped a QTL for phenylpropanoid-derived volatiles (2-phenylethanol, phenylacetaldehyde and 1-nitro-2-phenylethane) on chromosome 4, using a diversity panel of 94 cultivars and several F₂ and F₆ populations of a half diallel cross (Tikunov et al. 2020). Near isogenic lines, harbouring the marker alleles associated to the genomic region on chromosome 4 for either low or high phenylpropanoid volatiles, allowed to confirm the correlation of this locus with fruity and rose-hip aroma (Tikunov et al. 2020). Fine-mapping the locus on chromosome 4 narrowed down to 11 genes in the associated region, and one of them showed high similarity to a 3-methyl-2-oxobutanoate dehydrogenase, an enzyme with decarboxylase activity, and which may be involved in the decarboxylation of phenylalanine, to produce the common precursor of the three volatiles (Table 2). Gene editing of the candidate gene, named *FLORAL4*, by CRISPR led to an important reduction in phenylalanine-derived volatiles and to a less extent in 3-methylbutanol levels. On the other hand, 3-methylbutanol precursor, the branched-chain amino acid leucine, was significantly increased, while no differences were observed in phenylalanine content. Even if gene editing validated the effect of *FLORAL4* deleterious mutations *in planta*, the decarboxylase activity of the gene could not be verified *in vivo* (Tikunov et al. 2020).

Other phenylpropanoid-derived volatiles, guaiacol and methylsalicylate, have a negative impact on consumer liking, being associated with 'smoky' or 'medicinal' aroma (Zanor et al. 2009b). Tieman et al. (2017) detected a locus on chromosome 9 linked to the two metabolically-related volatiles, and suggested that *E8*, a specific highly-expressed ripening-related gene, could be the candidate based on the analysis of silenced tomato lines, even if the molecular mechanism remained unclear. On the other hand, using a combination of omics approaches, Tikunov et al. (2013, 2020) highlighted the role of a *non-smoky glycosyltransferase* (*NSGT1*), also located on chromosome 9, and responsible for the volatile release. Very recently, and taking advantage of the new long-read technology sequencing, a collection of 100 tomato accessions was sequenced, shedding light upon structural variants

underlying important domestication traits, including the *smoky* locus (Alonge et al. 2020). Indeed, this new assembly allowed a better resolution of the locus sequence and to fill the gaps flanking *E8* with *NSGT1* present in the previous reference genome. Furthermore, haplotype analysis seemed to confirm that multiple mutant alleles of *nsgt1* underlie the natural variation in guaiacol and methylsalicylate volatile content (Alonge et al. 2020). GWAS analysis and the generation of an F₂ population between two accessions segregating for haplotype V (23-kpb deletion removing both *E8* and *NSGT1*) and functional *NSGT1* analysis validated that the deletion of both *E8* and *NSGT1* results in higher guaiacol and methylsalicylate content (Table 2) (Alonge et al. 2020).

Another group of flavour-important volatiles, such as geranylacetone and 6-methyl-5-hepten-2-one (MHO) are carotenoid-derived compounds (apocarotenoids). Tieman et al. (2017) identified one associated locus for MHO, four for geranylacetone and two for both compounds. Interestingly, by comparing the frequency of alleles for both volatile content in wild, heirloom and modern accessions, they found that allelic combinations were progressively lost along domestication. While breeders have selected the two alleles responsible for the highest MHO content, the two most frequent alleles for geranylacetone were not the most advantageous ones for geranylacetone levels (Tieman et al. 2017). Indeed, breeders have indirectly favoured increased MHO levels, by enhancing lycopene content, the main carotenoid pigment responsible for the bright red ripe fruit colour and MHO precursor. On the opposite, they could not easily select for geranylacetone accumulation, as it is derived from the oxidative cleavage of carotenoids present in small amounts and which do not contribute to the fruit colour (Tieman et al. 2017). In addition, Tieman et al. (2017) could identify a phytoene synthase, the enzyme catalysing the first step of carotenoid pathway, in a region where a strong association with both apocarotenoids was detected.

Conclusions

Current challenges in modern breeding include increasing yield and quality under climatic changing conditions to secure global food production. From one hand, there is an urging need to improve cereal and legume nutrient composition, as a key strategy to fight against health issues in developing countries. On the other hand, recent advances in multiomic platforms is allowing to understand better the complexity of crop and fruit metabolism and genetics, in order to increase their organoleptic and nutraceuticals quality, as requested by consumers. While metabolomic and/or proteomic approaches enables to outline genotypes

with quality-related characteristics, we can expect that deep sequencing combined with other omics platforms will increased our knowledge about the genetic architecture of quality traits. In a foreseeable future, this information will be easily translated to breeding programs for crop improvement, after trait (phenotype)-genotype validation, as it is already occurring for tomato fruit (Li et al. 2018a; Rodríguez-Leal et al. 2017; Rothan et al. 2019). In the sense, it is worth mentioning that the rapid progresses made in multiomic technologies are matched with advances in gene validation and editing techniques. Indeed, recent development of gene editing applications, and in particular CRISPR/Cas9 system, will allow base editing or targeted gene replacement to facilitate allele substitution, where conventional marker-assisted selection may not be achievable (Gaston et al. 2020; Rothan et al. 2019).

Furthermore, as breeding targets can be shared between crop species, mining tomato data can be a fundamental approach to help the exploitation of less studied and genetically more complex species, as suggested for the octoploid cultivated strawberry (*F. x ananassa*) by a recent review (Gaston et al. 2020). Other future perspectives may consist of GWAS meta-analysis, as recently performed in tomato by Zhao et al. (2019), which identified 37 promising candidate genes for flavour improvement and in particular sugar content. This strategy, together with marker-assisted selection or gene editing, may be implemented for other crop species, and will allow to substitute undesirable alleles in modern cultivars by favourable alleles for fruit nutrition and flavour traits.

Abbreviations

QTL: Quantitative trait locus; GWAS: Genome wide association study; mGWAS: Metabolic-GWAS; TF: Transcription factor; O2: Opaque2; eQTL: Expression QTL; WD40: WD40-repeat protein; ERF: Ethylene-responsive genes; JA: Jasmonic acid; UFGT: UDP-glucose:flavonoid O-glycosyltransferase; LDOX: Leucoanthocyanidin dioxygenase; SW pathway: Smirnov-Wheeler pathway; NSGT1: Non-smoky glycosyltransferase 1.

Acknowledgements

S.D.-S. and S.O. acknowledges the support by Plan Propio from University of Malaga.

Authors' contributions

All authors did the literature research and drafted the review and helped writing the final manuscript. All authors read and approved the final manuscript.

Funding

This work was supported through grants RTI2018-099797-B-100 (Ministerio de Ciencia, Innovación y Universidades, Spain) and UMA18-FEDERJA-179 (FEDER-Junta Andalucía).

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 27 September 2020 Accepted: 15 December 2020

Published online: 04 January 2021

References

- Aharoni A, Galili G. Metabolic engineering of the plant primary-secondary metabolism interface. *Curr Opin Biotechnol*. 2011. <https://doi.org/10.1016/j.copbio.2010.11.004>.
- Alonge M, Wang X, Benoit M, Van Der Knaap E, Schatz MC, Lippman ZB, et al. Major impacts of widespread structural variation on gene expression and crop improvement in tomato in brief comprehensive structural variant identification in tomato genomes allows insight into the evolution and domestication of tomato and serves as a res. *Cell*. 2020. <https://doi.org/10.1016/j.cell.2020.05.021>.
- An JP, Zhang XW, You CX, Bi SQ, Wang XF, Hao YJ. MdWRKY40 promotes wounding-induced anthocyanin biosynthesis in association with MdMYB1 and undergoes MdBT2-mediated degradation. *New Phytol*. 2019. <https://doi.org/10.1111/nph.16008>.
- Ashokkumar K, Govindaraj M, Karthikeyan A, Shobhana VG, Warkentin TD. Genomics-integrated breeding for carotenoids and folates in staple cereal grains to reduce malnutrition. *Front Genet*. 2020. <https://doi.org/10.3389/fgene.2020.00414>.
- Azmach G, Menkir A, Spillane C, Gedil M. Genetic loci controlling carotenoid biosynthesis in diverse tropical maize lines. *G3 Genes Genomes Genet*. 2018. <https://doi.org/10.1534/g3.117.300511>.
- Bai Y, Lindhout P. Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? *Ann Bot*. 2007. <https://doi.org/10.1093/aob/mcm150>.
- Baldwin EA, Goodner K, Plotto A. Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. *J Food Sci*. 2008. <https://doi.org/10.1111/j.1750-3841.2008.00825.x>.
- Ballester A-R, Tikunov Y, Molthoff J, Grandillo S, Viquez-Zamora M, de Vos R, et al. Identification of loci affecting accumulation of secondary metabolites in tomato fruit of a *Solanum lycopersicum* × *Solanum chmielewskii* introgression line population. *Front Plant Sci*. 2016. <https://doi.org/10.3389/fpls.2016.01428>.
- Bauchet G, Grenier S, Samson N, Segura V, Kende A, Beekwilder 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. <https://doi.org/10.1111/nph.14615>.
- Baxter CJ, Carrari F, Bauke A, Overy S, Hill SA, Quick PW, et al. Fruit carbohydrate metabolism in an introgression line of tomato with increased fruit soluble solids. *Plant Cell Physiol*. 2005. <https://doi.org/10.1093/pcp/pci040>.
- Branham S, Vexler L, Meir A, Tzuri G, Frieman Z, Levi A, et al. Genetic mapping of a major codominant QTL associated with β -carotene accumulation in watermelon. *Mol Breed*. 2017. <https://doi.org/10.1007/s11032-017-0747-0>.
- Butelli E, Titta L, Giorgio M, Mock HP, Matros A, Peterek S, et al. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat Biotechnol*. 2008. <https://doi.org/10.1038/nbt.1506>.
- Causse M. A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. *J Exp Bot*. 2004. <https://doi.org/10.1093/jxb/erv207>.
- Chan EKF, Rowe HC, Hansen BG, Kliebenstein DJ. The complex genetic architecture of the metabolome. *PLoS Genet*. 2010. <https://doi.org/10.1371/journal.pgen.1001198>.
- Chen W, Gao Y, Xie W, Gong L, Lu K, Wang W, et al. Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat Genet*. 2014. <https://doi.org/10.1038/ng.3007>.
- Chen W, Wang W, Peng M, Gong L, Gao Y, Wan J, et al. Comparative and parallel genome-wide association studies for metabolic and agronomic traits in cereals. *Nat Commun*. 2016. <https://doi.org/10.1038/ncomms12767>.
- Cho K, Cho KS, Sohn HB, Ha IJ, Hong SY, Lee H, et al. Network analysis of the metabolome and transcriptome reveals novel regulation of potato pigmentation. *J Exp Bot*. 2016. <https://doi.org/10.1093/jxb/erv549>.
- Colasuonno P, Marcotuli I, Blanco A, Maccaferri M, Condorelli GE, Tuberosa R, et al. Carotenoid Pigment Content in Durum Wheat (*Triticum turgidum* L. var durum): An Overview of Quantitative Trait Loci and Candidate Genes. *Front Plant Sci*. 2019. <https://doi.org/10.3389/fpls.2019.01347>.
- Dai ZW, Meddard M, Renaud C, Merlin I, Hilbert G, Delro S, et al. Long-term in vitro culture of grape berries and its application to assess the effects of sugar supply on anthocyanin accumulation. *J Exp Bot*. 2014. <https://doi.org/10.1093/jxb/ert489>.
- Deng M, Li D, Luo J, Xiao Y, Liu H, Pan Q, et al. The genetic architecture of amino acids dissection by association and linkage analysis in maize. *Plant Biotechnol J*. 2017. <https://doi.org/10.1111/pbi.12712>.
- de Abreu e Lima F, Li K, Wen W, Yan J, Nikoloski Z, Willmitzer L, et al. Unraveling lipid metabolism in maize with time-resolved multi-omics data. *Plant J*. 2018. <https://doi.org/10.1111/tipj.13833>.
- D'Angelo M, Zanor MI, Sance M, Cortina PR, Boggio SB, Asprelli P, et al. Contrasting metabolic profiles of tasty Andean varieties of tomato fruit in comparison with commercial ones. *J Sci Food Agric*. 2018. <https://doi.org/10.1002/jsfa.8930>.
- Angelo M, Zanor MI, Burgos E, Asprelli PD, Boggio SB, Carrari F, et al. Fruit metabolic and transcriptional programs differentiate among Andean tomato (*Solanum lycopersicum* L.) accessions. *Planta*. 2019. <https://doi.org/10.1007/s00425-019-03274-4>.
- D'Esposito D, Ferriello F, Molin AD, Diretto G, Sacco A, Minio A, et al. Unraveling the complexity of transcriptomic, metabolomic and quality environmental response of tomato fruit. *BMC Plant Biol*. 2017. <https://doi.org/10.1186/s12870-017-1008-4>.
- Di Matteo A, Sacco A, Anacleria M, Pezzotti M, Delledonne M, Ferrarini A, et al. The ascorbic acid content of tomato fruits is associated with the expression of genes involved in pectin degradation. *BMC Plant Biol*. 2010. <https://doi.org/10.1186/1471-2229-10-163>.
- Ellison SL, Luby CH, Corak KE, Coe KM, Senalik D, Iorizzo M, et al. Carotenoid presence is associated with the or gene in domesticated carrot. *Genetics*. 2018. <https://doi.org/10.1534/genetics.118.301299>.
- Esuma W, Herselman L, Labuschagne MT, Ramu P, Lu F, Baguma Y, et al. Genome-wide association mapping of provitamin A carotenoid content in cassava. *Euphytica*. 2016. <https://doi.org/10.1007/s10681-016-1772-5>.
- Etienne A, Génard M, Lobit P, Mbéguié-A-Mbéguié D, Bugaud C. What controls fleshy fruit acidity? A review of malate and citrate accumulation in fruit cells. *J Exp Bot*. 2013. <https://doi.org/10.1093/jxb/ert035>.
- Fang ZZ, Zhou DR, Ye XF, Jiang CC, Pan SL. Identification of candidate anthocyanin-related genes by transcriptomic analysis of 'furogl' plum (*Prunus salicina* Lindl.) during fruit ripening using RNA-seq. *Front Plant Sci*. 2016. <https://doi.org/10.3389/fpls.2016.01338>.
- Fenech M, Amaya I, Valpuesta V, Botella MA. Vitamin C content in fruits: biosynthesis and regulation. *Front Plant Sci*. 2019. <https://doi.org/10.3389/fpls.2018.02006>.
- Figuerola CM, Lunn JE. A tale of two sugars: trehalose 6-phosphate and sucrose. *Plant Physiol*. 2016. <https://doi.org/10.1104/pp.16.00417>.
- Fraser CM, Chapple C. The phenylpropanoid pathway in arabidopsis. *Arab B*. 2011. <https://doi.org/10.1199/tab.0152>.
- Fridman E, Carrari F, Sheng Y, Fernie A, Zamir D. Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science*. 2004. <https://doi.org/10.1126/science.1101666>.
- Fu J, Cheng Y, Linghu J, Yang X, Kang L, Zhang Z, et al. RNA sequencing reveals the complex regulatory network in the maize kernel. *Nat Commun*. 2013. <https://doi.org/10.1038/ncomms3832>.
- Galili G, Amir R. Fortifying plants with the essential amino acids lysine and methionine to improve nutritional quality. *Plant Biotechnol J*. 2013. <https://doi.org/10.1111/pbi.12025>.
- Galili G, Amir R, Fernie AR. The regulation of essential amino acid synthesis and accumulation in plants. *Annu Rev Plant Biol*. 2016. <https://doi.org/10.1146/annurev-arplant-043015-112213>.
- Galili G, Galili S, Lewinsohn E, Tadmor Y. Genetic, molecular, and genomic approaches to improve the value of plant foods and feeds. *CRC Crit Rev Plant Sci*. 2002. <https://doi.org/10.1080/0735-260291044232>.

- Gao L, Gonda I, Sun H, Ma Q, Bao K, Tieman DM, et al. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat Genet*. 2019. <https://doi.org/10.1038/s41588-019-0410-2>.
- Garbowicz K, Liu Z, Alseekh S, Tieman D, Taylor M, Kuhalskaya A, 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. <https://doi.org/10.1016/j.molp.2018.06.003>.
- Gaston A, Osorio S, Denoyes B, Rothan C. Applying the solanaceae strategies to strawberry crop improvement. *Trends Plant Sci*. 2020. <https://doi.org/10.1016/j.tplants.2019.10.003>.
- Gemenet DC, da Silva PG, De Boeck B, Wood JC, Mollinari M, Olukolu BA, et al. Quantitative trait loci and differential gene expression analyses reveal the genetic basis for negatively associated β -carotene and starch content in hexaploid sweetpotato [*Ipomoea batatas* (L.) Lam.]. *Theor Appl Genet*. 2020. <https://doi.org/10.1007/s00122-019-03437-7>.
- Giampieri F, Forbes-Hernandez TY, Gasparri M, Alvarez-Suarez JM, Afrin S, Bompadre S, et al. Strawberry as a health promoter: an evidence based review. *Food Funct*. 2015. <https://doi.org/10.1039/c5fo00147a>.
- Giovannoni J. MOLECULAR BIOLOGY OF FRUIT MATURATION AND RIPENING. *Annu Rev Plant Physiol Mol Biol*. 2001. <https://doi.org/10.1146/annurev.arplant.52.1.725>.
- Gong L, Chen W, Gao Y, Liu X, Zhang H, Xu C, et al. Genetic analysis of the metabolome exemplified using a rice population. *Proc Natl Acad Sci USA*. 2013. <https://doi.org/10.1073/pnas.1319681110>.
- Goulet C, Mageroy MH, Lam NB, Floystad A, Tieman DM, Klee HJ. Role of an esterase in flavor volatile variation within the tomato clade. *Proc Natl Acad Sci USA*. 2012. <https://doi.org/10.1073/pnas.1216515109>.
- Hassan S, Mathesius U. The role of flavonoids in root-rhizosphere signalling: Opportunities and challenges for improving plant-microbe interactions. *J Exp Bot*. 2012. <https://doi.org/10.1093/jxb/err430>.
- He J, Giusti MM. Anthocyanins: natural colorants with health-promoting properties. *Annu Rev Food Sci Technol*. 2010. <https://doi.org/10.1146/annurev.food.080708.100754>.
- He XJ, Mu RL, Cao WH, Zhang ZG, Zhang JS, Chen SY. AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. *Plant J*. 2005. <https://doi.org/10.1111/j.1365-3113X.2005.02575.x>.
- Heng Z, Sheng O, Huang W, Zhang S, Fernie AR, Motorykin I, et al. Integrated proteomic and metabolomic analysis suggests high rates of glycolysis are likely required to support high carotenoid accumulation in banana pulp. *Food Chem*. 2019. <https://doi.org/10.1016/j.foodchem.2019.125016>.
- Jaakola L. New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci*. 2013. <https://doi.org/10.1016/j.tplants.2013.06.003>.
- Jia H, Jiu S, Zhang C, Wang C, Tariq P, Liu Z, et al. Abscisic acid and sucrose regulate tomato and strawberry fruit ripening through the abscisic acid-stress-ripening transcription factor. *Plant Biotechnol J*. 2016. <https://doi.org/10.1111/pbi.12563>.
- Klee HJ, Tieman DM. The genetics of fruit flavour preferences. *Nat Rev Genet*. 2018. <https://doi.org/10.1038/s41576-018-0002-5>.
- Kobayashi S, Goto-Yamamoto N, Hirochika H. Retrotransposon-induced mutations in grape skin color. *Science*. 2004. <https://doi.org/10.1126/science.1095011>.
- Labadie M, Vallin G, Petit A, Ring L, Hoffmann T, Potier A, et al. Metabolite Quantitative Trait Loci for flavonoids provide new insights into the genetic architecture of strawberry (*Fragaria x ananassa*) fruit quality. *J Agric Food Chem*. 2020. <https://doi.org/10.1111/1/j.1753-4887.2009.00178.x>.
- Lee T, Lee I. AraGWAB: Network-based boosting of genome-wide association studies in Arabidopsis thaliana. *Sci Rep*. 2018. <https://doi.org/10.1038/s41598-018-21301-4>.
- Lee S, Van K, Sung M, Nelson R, LaMantia J, McHale LK, et al. Genome-wide association study of seed protein, oil and amino acid contents in soybean from maturity groups I to IV. *Theor Appl Genet*. 2019. <https://doi.org/10.1007/s00122-019-03304-5>.
- Li T, Yang X, Yu Y, Si X, Zhai X, Zhang H, et al. Domestication of wild tomato is accelerated by genome editing. *Nat Biotechnol*. 2018a. <https://doi.org/10.1038/nbt.4273>.
- Li Y, Fang J, Qi X, Lin M, Zhong Y, Sun L, et al. Combined analysis of the fruit metabolome and transcriptome reveals candidate genes involved in flavonoid biosynthesis in *Actinidia arguta*. *Int J Mol Sci*. 2018b. <https://doi.org/10.3390/ijms19051471>.
- Li K, Wen W, Alseekh S, Yang X, Guo H, Li W, et al. Large-scale metabolite quantitative trait locus analysis provides new insights for high-quality maize improvement. *Plant J*. 2019. <https://doi.org/10.1111/tpj.14317>.
- Li N, Shang J, Wang J, Zhou D, Li N, Ma S. Discovery of the Genomic Region and Candidate Genes of the Scarlet Red Flesh Color (Yscr) Locus in Watermelon (*Citrullus Lanatus* L.). *Front Plant Sci*. 2020. <https://doi.org/10.3389/fpls.2020.00116>.
- Lightbourn GJ, Griesbach RJ, Novotny JA, Clevidence BA, Rao DD, Stommel JR. Effects of anthocyanin and carotenoid combinations on foliage and immature fruit color of *Capsicum annuum* L. *J Hered*. 2008. <https://doi.org/10.1093/jhered/esm108>.
- Lin T, Zhu G, Zhang J, Xu X, Yu Q, Zheng Z, et al. Genomic analyses provide insights into the history of tomato breeding. *Nat Genet*. 2014. <https://doi.org/10.1038/ng.3117>.
- Lincoln C, Long J, Yamaguchi J, Serikawa K, Hake S. A knotted1-like homeobox gene in Arabidopsis is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. *Plant Cell*. 1994. <https://doi.org/10.1105/tpc.6.12.1859>.
- Liu H, Jiao J, Liang X, Liu J, Meng H, Chen S, et al. Map-based cloning, identification and characterization of the w gene controlling white immature fruit color in cucumber (*Cucumis sativus* L.). *Theor Appl Genet*. 2016. <https://doi.org/10.1007/s00122-016-2700-8>.
- Liu H, Luo X, Niu L, Xiao Y, Chen L, Liu J, et al. Distant eQTLs and non-coding sequences play critical roles in regulating gene expression and quantitative trait variation in maize. *Mol Plant*. 2017. <https://doi.org/10.1016/j.molp.2016.06.016>.
- Liu Y, Lv J, Liu Z, Wang J, Yang B, Chen W, et al. Integrative analysis of metabolome and transcriptome reveals the mechanism of color formation in pepper fruit (*Capsicum annuum* L.). *Food Chem*. 2020. <https://doi.org/10.1016/j.foodchem.2019.125629>.
- Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F. Quantitative trait loci analysis of nitrogen use efficiency in Arabidopsis. *Plant Physiol*. 2003. <https://doi.org/10.1104/pp.102.010785.and>.
- Luo J. Metabolite-based genome-wide association studies in plants. *Curr Opin Plant Biol*. 2015. <https://doi.org/10.1016/j.pbi.2015.01.006>.
- Luo Y, Lin Y, Mo F, Ge C, Jiang L, Zhang Y, et al. Sucrose promotes strawberry fruit ripening and affects ripening-related processes. *Int J Genomics*. 2019. <https://doi.org/10.1155/2019/9203057>.
- Mageroy MH, Tieman DM, Floystad A, Taylor MG, Klee HJ. A Solanum lycopersicum catechol-O-methyltransferase involved in synthesis of the flavor molecule guaiacol. *Plant J*. 2012. <https://doi.org/10.1111/j.1365-3113X.2011.04854.x>.
- Mandal S, Mandal RK. Seed storage proteins and approaches for improvement of their nutritional quality by genetic engineering. *Curr Sci*. 2000. <https://doi.org/10.2174/1389202917666160202215934>.
- Meng G, Clausen SK, Rasmussen SK. Transcriptome analysis reveals candidate genes related to anthocyanin biosynthesis in different carrot genotypes and tissues. *Plants*. 2020. <https://doi.org/10.3390/plants9030344>.
- Nardoza S, Boldingh HL, Kashuba MP, Feil R, Jones D, Thrimawithana AH, et al. Carbon starvation reduces carbohydrate and anthocyanin accumulation in red-fleshed fruit via trehalose 6-phosphate and MYB27. *Plant Cell Environ*. 2020. <https://doi.org/10.1111/pce.13699>.
- Oren E, Tzuri G, Vexler L, Dafna A, Meir A, Faigenboim A, et al. The multi-allelic APRR2 gene is associated with fruit pigment accumulation in melon and watermelon. *J Exp Bot*. 2019. <https://doi.org/10.1093/jxb/erz182>.
- Osorio S, Alba R, Damasceno CMB, Lopez-Casado G, Lohse M, Zanon MI, et al. Systems biology of tomato fruit development: combined transcript, protein, and metabolite analysis of tomato transcription factor (*nor*, *rin*) and ethylene receptor (*Nr*) mutants reveals novel regulatory interactions. *Plant Physiol*. 2011. <https://doi.org/10.1104/pp.111.175463>.
- Osorio S, Carneiro RT, Lytovchenko A, McQuinn R, Sørensen I, Vallarino JG, et al. Genetic and metabolic effects of ripening mutations and vine detachment on tomato fruit quality. *Plant Biotechnol J*. 2019. <https://doi.org/10.1111/pbi.13176>.
- Owens BF, Gore MA, Magallanes-Lundback M, Tiede T, Diepenbrock CH, Kandianis CB, et al. A foundation for provitamin A biofortification of maize: Genome-wide association and genomic prediction models of carotenoid levels. *Genetics*. 2014. <https://doi.org/10.1534/genetics.114.169979>.

- Pan Y, Bradley G, Pyke K, Ball G, Lu C, Fray R, et al. Network inference analysis identifies an APRR2-like gene linked to pigment accumulation in tomato and pepper fruits. *Plant Physiol*. 2013. <https://doi.org/10.1104/pp.112.212654>.
- Park S, Kim HS, Jung YJ, Kim SH, Ji CY, Wang Z, et al. Orange protein has a role in phytoene synthase stabilization in sweetpotato. *Sci Rep*. 2016. <https://doi.org/10.1038/srep33563>.
- Peng Y, Liu H, Chen J, Shi T, Zhang C, Sun D, et al. Genome-wide association studies of free amino acid levels by six multi-locus models in bread wheat. *Front Plant Sci*. 2018. <https://doi.org/10.3389/fpls.2018.011196>.
- Petrussa E, Braidot E, Zancani M, Peresson C, Bertolini A, Patui S, et al. Plant flavonoids-biosynthesis, transport and involvement in stress responses. *Int J Mol Sci*. 2013. <https://doi.org/10.3390/ijms140714950>.
- Pott DM, Osorio S, Vallarino JG. From central to specialized metabolism : an overview of some secondary compounds derived from the primary metabolism for their role in conferring nutritional and organoleptic characteristics to fruit. *Front Plant Sci*. 2019. <https://doi.org/10.3389/fpls.2019.00835>.
- Pott DM, de Abreu e Lima F, Soria C, Willmitzer L, Fernie AR, Nikoloski Z, et al. Metabolic reconfiguration of strawberry physiology in response to postharvest practices. *Food Chem*. 2020. <https://doi.org/10.1016/j.foodchem.2020.126747>.
- Qin J, Shi A, Song Q, Li S, Wang F, Cao Y, et al. Genome wide association study and genomic selection of amino acid concentrations in soybean seeds. *Front Plant Sci*. 2019. <https://doi.org/10.3389/fpls.2019.01445>.
- Rambla JL, Trapero-Mozos A, Direccion G, Rubio-Moraga A, Granell A, Gómez-Gómez L, et al. Gene-metabolite networks of volatile metabolism in airen and tempranillo grape cultivars revealed a distinct mechanism of aroma bouquet production. *Front Plant Sci*. 2016. <https://doi.org/10.3389/fpls.2016.01619>.
- Rezaei MK, Deokar AA, Arganosa G, Roorkiwal M, Pandey SK, Warkentin TD, et al. Mapping quantitative trait loci for carotenoid concentration in three F₂ populations of chickpea. *Plant Genome*. 2019. <https://doi.org/10.3835/plantgenome2019.07.0067>.
- Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB. Engineering quantitative trait variation for crop improvement by genome editing. *Cell*. 2017. <https://doi.org/10.1016/j.cell.2017.08.030>.
- Rothan C, Diouf I, Causse M. Trait discovery and editing in tomato. *Plant J*. 2019. <https://doi.org/10.1111/tpj.14152>.
- Ruggieri V, Francese G, Sacco A, D'Alessandro A, Rigano MM, Parisi M, et al. An association mapping approach to identify favourable alleles for tomato fruit quality breeding. *BMC Plant Biol*. 2014. <https://doi.org/10.1186/s12870-014-0337-9>.
- Sauvage C, Segura V, Bauchet G, Stevens R, Do PT, Nikoloski Z, et al. Genome-wide association in tomato reveals 44 candidate loci for fruit metabolic traits. *Plant Physiol*. 2014. <https://doi.org/10.1104/pp.114.241521>.
- Schaart JG, Dubos C, Romero De La Fuente I, van Houwelingen AMML, de Vos RCH, Jonker HH, et al. Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria × ananassa*) fruits. *New Phytol*. 2013. <https://doi.org/10.1111/nph.12017>.
- Shan W, Kuang fei J, Chen L, Peng H, Xiao Y, Li X, et al. Molecular characterization of banana NAC transcription methylation and chromatin patterning factors and their interactions with ethylene signalling component EIL during fruit ripening. *J Exp Bot*. 2012. <https://doi.org/10.1093/jxb/err313>.
- Sun Y, Shi Y, Liu G, Yao F, Zhang Y, Yang C, et al. Natural variation in the OsbZIP18 promoter contributes to branched-chain amino acid levels in rice. *New Phytol*. 2020. <https://doi.org/10.1111/nph.16800>.
- Sun T, Yuan H, Cao H, Yazdani M, Tadmor Y, Li L. Carotenoid metabolism in plants: the role of plastids. *Mol Plant*. 2018. <https://doi.org/10.1016/j.molp.2017.09.010>.
- Tadmor Y, King S, Levi A, Davis A, Meir A, Wasserman B, et al. Comparative fruit colouration in watermelon and tomato. *Food Res Int*. 2005. <https://doi.org/10.1016/j.foodres.2004.07.011>.
- Thines B, Katsir L, Melotto M, Niu Y, Mandaokar A, Liu G, et al. JAZ repressor proteins are targets of the SCFCO11 complex during jasmonate signalling. *Nature*. 2007. <https://doi.org/10.1038/nature05960>.
- Tieman D, Bliss P, McIntyre LM, Blandon-Ubeda A, Bies D, Odabasi AZ, et al. The chemical interactions underlying tomato flavor preferences. *Curr Biol*. 2012. <https://doi.org/10.1016/j.cub.2012.04.016>.
- Tieman DM, Zeigler M, Schmelz EA, Taylor MG, Bliss P, Kirst M, et al. Identification of loci affecting flavour volatile emissions in tomato fruits. *J Exp Bot*. 2006. <https://doi.org/10.1093/jxb/erj074>.
- Tieman D, Zhu G, Resende MFR, Lin T, Nguyen C, Bies D, et al. A chemical genetic roadmap to improved tomato flavor. *Science*. 2017. <https://doi.org/10.1126/science.aal1556>.
- Tikunov YM, Molthoff J, de Vos RCH, Beekwilder J, van Houwelingen A, van der Hoof JJJ, et al. Non-smoky GLYCOSYLTRANSFERASE1 prevents the release of smoky aroma from tomato fruit. *Plant Cell*. 2013. <https://doi.org/10.1105/tpc.113.114231>.
- Tikunov Y, Meijer-Dekens F, Molthoff J, Paulo J, Finkers R, Capel I, 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. <https://doi.org/10.1111/tpj.14795>.
- Tzuri G, Zhou X, Chayat N, Yuan H, Portnoy V, Meir A, et al. A "golden" SNP in CmOr governs the fruit flesh color of melon (*Cucumis melo*). *Plant J*. 2015. <https://doi.org/10.1111/tpj.12814>.
- Vallarino JG, Yeats TH, Maximova E, Rose JK, Fernie AR, Osorio S. Postharvest changes in LIN5-down-regulated plants suggest a role for sugar deficiency in cuticle metabolism during ripening. *Phytochemistry*. 2017. <https://doi.org/10.1016/j.phytochem.2017.06.007>.
- Vallarino JG, de Abreu e Lima F, Soria C, Tong H, Pott DM, Willmitzer L, et al. Genetic diversity of strawberry germplasm using metabolomic biomarkers. *Sci Rep*. 2018. <https://doi.org/10.1038/s41598-018-32212-9>.
- Vallarino JG, Pott DM, Cruz-rus E, Miranda L, Medina-minguez JJ, Valpuesta V, et al. Identification of quantitative trait loci and candidate genes for primary metabolite content in strawberry fruit. *Hortic Res*. 2019. <https://doi.org/10.1038/s41438-018-0077-3>.
- Vogt T. Phenylpropanoid biosynthesis. *Mol Plant*. 2010. <https://doi.org/10.1093/mp/ssp106>.
- Wang Z, Cui Y, Vainstein A, Chen S, Ma H. Regulation of fig (*Ficus carica* L.) fruit color: Metabolomic and transcriptomic analyses of the flavonoid biosynthetic pathway. *Front Plant Sci*. 2017a. <https://doi.org/10.3389/fpls.2017.01990>.
- Wang Z, Du H, Zhai R, Song L, Ma F, Xu L. Transcriptome analysis reveals candidate genes related to color fading of 'red bartlett' (pyrus communis L.). *Front Plant Sci*. 2017b. <https://doi.org/10.3389/fpls.2017.00455>.
- Wang C, Qiao A, Fang X, Sun L, Gao P, Davis AR, et al. Fine mapping of lycopene content and flesh color related gene and development of molecular marker-assisted selection for flesh color in watermelon (*Citrullus lanatus*). *Front Plant Sci*. 2019. <https://doi.org/10.3389/fpls.2019.01240>.
- Wen W, Li D, Li X, Gao Y, Li W, Li H, et al. Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nat Commun*. 2014. <https://doi.org/10.1038/ncomms4438>.
- Wen W, Li K, Alseekh S, Omranian N, Zhao L, Zhou Y, et al. Genetic determinants of the network of primary metabolism and their relationships to plant performance in a maize recombinant inbred line population. *Plant Cell*. 2015. <https://doi.org/10.1105/tpc.15.00208>.
- Wen W, Jin M, Li K, Liu H, Xiao Y, Zhao M, et al. An integrated multi-layered analysis of the metabolic networks of different tissues uncovers key genetic components of primary metabolism in maize. *Plant J*. 2018. <https://doi.org/10.1111/tpj.13835>.
- Willson MF, Whelan CJ. The evolution of fruit color in fleshy-fruited plants. *Am Nat*. 1990. <https://doi.org/10.1086/285132>.
- Woo HR, Kim HJ, Nam HG, Lim PO. Plant leaf senescence and death—regulation by multiple layers of control and implications for aging in general. *J Cell Sci*. 2013. <https://doi.org/10.1242/jcs.109116>.
- Xu J, Fu X, Cai Y, Wang M, Xu S, Li J, et al. Uncovering the genetic basis of carotenoid variations in maize kernels using two segregating populations. *Mol Breed*. 2019. <https://doi.org/10.1007/s11032-019-0982-7>.
- Xu W, Grain D, Le Gourrierc J, Harscoët E, Berger A, Jauvion V, et al. Regulation of flavonoid biosynthesis involves an unexpected complex transcriptional regulation of TT8 expression, in Arabidopsis. *New Phytol*. 2013. <https://doi.org/10.1111/nph.12142>.
- Yang Q, Zhao D, Liu Q. Connections between amino acid metabolisms in plants: lysine as an example. *Front Plant Sci*. 2020. <https://doi.org/10.3389/fpls.2020.00928>.
- Yao G, Ming M, Allan AC, Gu C, Li L, Wu X, et al. Map-based cloning of the pear gene MYB114 identifies an interaction with other transcription factors to coordinately regulate fruit anthocyanin biosynthesis. *Plant J*. 2017. <https://doi.org/10.1111/tpj.13666>.

- Yauk YK, Ged C, Wang MY, Matich AJ, Tessarotto L, Cooney JM, et al. Manipulation of flavour and aroma compound sequestration and release using a glycosyltransferase with specificity for terpene alcohols. *Plant J*. 2014. <https://doi.org/10.1111/tpj.12634>.
- Ye J, Wang X, Hu T, Zhang F, Wang B, Li C, et al. An InDel in the promoter of AI-ACTIVATED MALATE TRANSPORTER9 selected during tomato domestication determines fruit malate contents and aluminum tolerance. *Plant Cell*. 2017. <https://doi.org/10.1105/tpc.17.00211>.
- Ye J, Li W, Ai G, Li C, Liu G, Chen W, 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. <https://doi.org/10.1371/journal.pgen.1008149>.
- Ying H, Shi J, Zhang S, Pingcui G, Wang S, Zhao F, et al. Transcriptomic and metabolomic profiling provide novel insights into fruit development and flesh coloration in *Prunus mira* Koehne, a special wild peach species. *BMC Plant Biol*. 2019. <https://doi.org/10.1186/s12870-019-2074-6>.
- Zanor MI, Osorio S, Nunes-Nesi A, Carrari F, Lohse M, Usadel B, et al. RNA interference of LIN5 in tomato confirms its role in controlling brix content, uncovers the influence of sugars on the levels of fruit hormones, and demonstrates the importance of sucrose cleavage for normal fruit development and fertility. *Plant Physiol*. 2009a. <https://doi.org/10.1104/pp.109.136598>.
- Zanor MI, Rambla JL, Chaïb J, Steppa A, Medina A, Granell A, et al. Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. *J Exp Bot*. 2009b. <https://doi.org/10.1093/jxb/erp086>.
- Zhai S, Liu J, Xu D, Wen W, Yan J, Zhang P, et al. A genome-wide association study reveals a rich genetic architecture of flour color-related traits in bread wheat. *Front Plant Sci*. 2018. <https://doi.org/10.3389/fpls.2018.01136>.
- Zhang K, Gan SS. An abscisic acid-AtNAP transcription factor-SAG113 protein phosphatase 2C regulatory chain for controlling dehydration in senescing *Arabidopsis* leaves. *Plant Physiol*. 2012. <https://doi.org/10.1104/pp.111.190876>.
- Zhang Y, Butelli E, De Stefano R, Schoonbeek HJ, Magusin A, Pagliarani C, et al. Anthocyanins double the shelf life of tomatoes by delaying overripening and reducing susceptibility to gray mold. *Curr Biol*. 2013. <https://doi.org/10.1016/j.cub.2013.04.072>.
- Zhang Y, Butelli E, Alseekh S, Tohge T, Rallapalli G, Luo J, et al. Multi-level engineering facilitates the production of phenylpropanoid compounds in tomato. *Nat Commun*. 2015. <https://doi.org/10.1038/ncomms9635>.
- Zhang J, Wang X, Lu Y, Bhusal SJ, Song Q, Cregan PB, et al. Genome-wide scan for seed composition provides insights into soybean quality improvement and the impacts of domestication and breeding. *Mol Plant*. 2018. <https://doi.org/10.1016/j.molp.2017.12.016>.
- Zhang Z, Tian C, Zhang Y, Li C, Li X, Yu Q, et al. Transcriptomic and metabolomic analysis provides insights into anthocyanin and procyanidin accumulation in pear. *BMC Plant Biol*. 2020. <https://doi.org/10.1111/ppl.12835>.
- Zhao J, Sauvage C, Zhao J, Bitton F, Bauchet G, Liu D, et al. Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. *Nat Commun*. 2019. <https://doi.org/10.1038/s41467-019-09462-w>.
- Zhou X, Welsch R, Yang Y, Álvarez D, Riediger M, Yuan H, et al. *Arabidopsis* OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. *Proc Natl Acad Sci USA*. 2015. <https://doi.org/10.1073/pnas.1420831112>.
- Zhu G, Wang S, Huang Z, Zhang S, Liao Q, Zhang C, et al. Rewiring of the Fruit Metabolome in Tomato Breeding. *Cell*. 2018. <https://doi.org/10.1016/j.cell.2017.12.019>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

