



Newly Developed MAGIC Population Allows Identification of Strong Associations and Candidate Genes for Anthocyanin Pigmentation in Eggplant

Giulio Mangino^{1†}, Andrea Arrones^{1†}, Mariola Plazas², Torsten Pook³, Jaime Prohens¹, Pietro Gramazio^{2*} and Santiago Vilanova^{1*}

¹ Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, Valencia, Spain,

² Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-Universitat Politècnica de València, Valencia, Spain, ³ Animal Breeding and Genetics Group, Department of Animal Sciences, Center for Integrated Breeding Research, University of Göttingen, Göttingen, Germany

OPEN ACCESS

Edited by:

Nunzio D'Agostino,
University of Naples Federico II, Italy

Reviewed by:

Monica Rodriguez,
University of Sassari, Italy
Mark Chapman,
University of Southampton,
United Kingdom

*Correspondence:

Pietro Gramazio
piegra@upv.es
Santiago Vilanova
sanvina@upvnet.upv.es

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 03 January 2022

Accepted: 20 January 2022

Published: 07 March 2022

Citation:

Mangino G, Arrones A, Plazas M, Pook T, Prohens J, Gramazio P and Vilanova S (2022) Newly Developed MAGIC Population Allows Identification of Strong Associations and Candidate Genes for Anthocyanin Pigmentation in Eggplant. *Front. Plant Sci.* 13:847789. doi: 10.3389/fpls.2022.847789

Multi-parent advanced generation inter-cross (MAGIC) populations facilitate the genetic dissection of complex quantitative traits in plants and are valuable breeding materials. We report the development of the first eggplant MAGIC population (S3 Magic EGGplant InCanum, S3MEGGIC; 8-way), constituted by the 420 S3 individuals developed from the intercrossing of seven cultivated eggplant (*Solanum melongena*) and one wild relative (*S. incanum*) parents. The S3MEGGIC recombinant population was genotyped with the eggplant 5k probes SPET platform and phenotyped for anthocyanin presence in vegetative plant tissues (PA) and fruit epidermis (FA), and for the light-insensitive anthocyanic pigmentation under the calyx (PUC). The 7,724 filtered high-confidence single-nucleotide polymorphisms (SNPs) confirmed a low residual heterozygosity (6.87%), a lack of genetic structure in the S3MEGGIC population, and no differentiation among subpopulations carrying a cultivated or wild cytoplasm. Inference of haplotype blocks of the nuclear genome revealed an unbalanced representation of the founder genomes, suggesting a cryptic selection in favour or against specific parental genomes. Genome-wide association study (GWAS) analysis for PA, FA, and PUC detected strong associations with two myeloblastosis (MYB) genes similar to *MYB113* involved in the anthocyanin biosynthesis pathway, and with a *COP1* gene which encodes for a photo-regulatory protein and may be responsible for the PUC trait. Evidence was found of a duplication of an ancestral *MYB113* gene with a translocation from chromosome 10 to chromosome 1 compared with the tomato genome. Parental genotypes for the three genes were in agreement with the identification of the candidate genes performed in the S3MEGGIC population. Our new eggplant MAGIC population is the largest recombinant population in eggplant and is a powerful tool for eggplant genetics and breeding studies.

Keywords: multi-parent advanced generation inter crosses (MAGIC), eggplant (*Solanum melongena* L.), *S. incanum*, anthocyanins, pigmentation under calyx (PUC), genome wide association study (GWAS), candidate genes, SPET (single primer enrichment technology)

INTRODUCTION

Multi-parent experimental populations are of great interest for the genetic dissection of quantitative traits and for the development of new recombinant materials for plant breeding (Huang et al., 2015). Despite their complex management and resources requirement, multi-parent advanced generation intercross (MAGIC) populations represent powerful next-generation mapping tools by combining a high genetic diversity and recombination with a low population structure (Arrones et al., 2020; Scott et al., 2020). MAGIC populations are already available in model species, such as *Arabidopsis thaliana* and in several crops, such as cereals, pulses, and vegetables (Kover et al., 2009; Bandillo et al., 2013; Pascual et al., 2015; Huynh et al., 2018), and have demonstrated their power to dissect the structure of complex traits (Dell'Acqua et al., 2015; Stadlmeier et al., 2018).

Although the available MAGIC populations have become a useful resource for genetic studies and breeding, most of them have only exploited intraspecific variation. The incorporation of crop wild relatives (CWRs) as founders could be a way of including multiple wild genomic fragments or introgressions into cultivated background genomes (Arrones et al., 2020). Apart from being of great interest for genetic analysis, interspecific MAGIC populations can be useful for broadening the genetic base of crops and provide new variation for breeding multiple traits, including those related to adaptation to climate change (Gramazio et al., 2020a). However, so far, the potential of interspecific MAGIC populations for plant breeding has largely remained unexploited (Arrones et al., 2020).

Eggplant (*Solanum melongena* L.) is a major vegetable crop of increasing importance, ranking fifth in global production among vegetables (FAOSTAT, 2019). Despite its economic importance, eggplant has lagged behind other major crops and little effort has been made to develop immortal experimental populations and genetic and genomic tools (Gramazio et al., 2018, 2019). So far, only one population of recombinant inbred lines (RIL) and one set of introgression lines (ILs) are publicly available (Lebeau et al., 2013; Gramazio et al., 2017a), while no multiparent population has been developed, so far. Conversely, in other related Solanaceae crops, such as tomato, several experimental populations have been developed, including MAGIC populations, which have allowed great advances in the genetic dissection of traits of interest (Pascual et al., 2015; Campanelli et al., 2019). For this reason, the development of this type of population would represent a landmark in eggplant breeding.

Anthocyanins are responsible for a set of specific and relevant traits in eggplant which can be used as a model plant for other crops (Moglia et al., 2020). Anthocyanins play a key role in plant defence mechanisms, and their synthesis and accumulation may vary in response to specific biotic and abiotic stresses (D'Amelia et al., 2018; Lv et al., 2019; Zhou et al., 2019). In addition, anthocyanins may prevent and alleviate human chronic diseases and provide health benefits (Toppino et al., 2020). In eggplant, anthocyanins are

responsible for the purple colour of the peel, one of the traits of greatest interest for eggplant breeding (Daunay and Hazra, 2012). Purple coloured eggplant fruits are the most demanded in various markets (Li et al., 2018), and developing dark purple-coloured eggplants, which results from the combination of anthocyanins with chlorophylls, is a major objective in eggplant breeding programmes. Eggplants are variable for the presence of anthocyanins not only in fruits, but also in other plant organs and tissues such as hypocotyl, stem, leaves, leaf veins, prickles, flower calyx, or corolla (Toppino et al., 2020). Anthocyanin biosynthesis has been widely studied in Solanaceae species (Van Eck et al., 1994; Borovsky et al., 2004; Gonzali et al., 2009), but its genetic control in eggplant has not been fully clarified. The anthocyanin biosynthetic pathway is a very conserved network in many plant species, where many enzymes and regulatory transcription factors (TFs) are involved (Albert et al., 2014).

The major anthocyanins in the eggplant fruit epidermis are delphinidin-3-p-coumaroylrutinoside-5-glucoside and delphinidin-3-rutinoside (Mennella et al., 2012; Moglia et al., 2020). Fruit anthocyanins (FA) are highly dependent on light (Jiang et al., 2016), however, the genotypes that carry the pigmentation under the calyx (PUC) mutation are able to independently synthesise the anthocyanins in the fruit epidermis of the incidence of light (Tigchelaar et al., 1968). Quantitative trait locus (QTL)-related studies using family based or genome-wide association (GWA) mapping approaches evidenced that chromosome 10 harbours most of the QTL/genes involved in anthocyanin formation, distribution, and accumulation (Doganlar et al., 2002; Barchi et al., 2012; Cericola et al., 2014; Frary et al., 2014; Toppino et al., 2016, 2020; Wei et al., 2020). The availability of high-quality eggplant genome sequences and transcriptomic data allowed the identification of putative candidate genes belonging to the myeloblastosis (MYB) family which controls the variation in anthocyanin content and fruit colour in eggplant (Docimo et al., 2016; Li et al., 2017, 2018, 2021; Xiao et al., 2018; Moglia et al., 2020; Shi et al., 2021), highlighting their synteny with other Solanaceae. However, the genes underlying the PUC phenotype have not been identified, so far.

Here, we report on the first eggplant MAGIC population derived from an interspecific cross of seven accessions of *S. melongena* and its wild relative *S. incanum* (Gramazio et al., 2019). It represents the largest experimental population described, so far, in eggplant, with a similar population size to MAGIC populations in other solanaceous crops. The population has been genotyped by applying the Single Primer Enrichment Technology (SPET) to explore its genetic architecture and the contribution of founders to the final population, and phenotyped for the presence of anthocyanins in the fruit epidermis and other plant organs and for the PUC trait. These traits were chosen due to their physiological, agronomic, and morphological relevance, high stability, and heritability. An association analysis has been performed to locate the genomic regions and to identify the candidate genes involved in the traits under study.

MATERIALS AND METHODS

Multi-Parent Advanced Generation Inter-Cross Population Construction

The eggplant MAGIC population has been developed by intermating seven cultivated eggplants, i.e., MM1597 (A), DH ECAVI (B), AN-S-26 (D), H15 (E), A0416 (F), IVIA-371 (G) and ASI-S-1 (H), and the *S. incanum* accession MM577 (C) (Figure 1A). The wild relative founder was chosen for its tolerance to some biotic and abiotic stresses, mainly drought (Knapp et al., 2013), and for showing a high phenolic content (Prohens et al., 2013). The performance of the founders was comprehensively characterised in previous morphoagronomic and genetic diversity studies (Hurtado et al., 2014; Gramazio et al., 2017b; Kaushik et al., 2018). In addition, their genomes had been resequenced (Gramazio et al., 2019). The latter study highlighted that, in the founder parents, the residual heterozygosity was less than 0.06%.

In order to develop the eggplant S3 Magic EGGplant InCanum (S3MEGGIC) population, founder lines had been inter-crossed by following a simple “funnel” approach (Wang et al., 2017; Arrones et al., 2020; Figure 1A). The eight founders (A–H) had been pairwise inter-crossed to produce two-way or simple F1 hybrids (AB, CD, EF, and GH), which were subsequently inter-crossed in pairs (AB × CD and EF × GH) to obtain two four-way or double hybrids (ABCD and EFGH). In order to achieve a complete admixture of all founder genomes and to avoid assortative mating, the double hybrids were intercrossed following a chain pollination scheme, with each individual being used as female and male parents (Díez et al., 2002; Supplementary Figure 1A).

All the obtained eight-way or quadruple hybrids (S0 generation) presented all the eight randomly shuffled genomes and only differed on the cytoplasm inherited from the maternal parent. The S0 progenies obtained using the double hybrid ABCD, as a female parent carried the cytoplasm of the wild *S. incanum* MM577, while those derived using the double hybrid EFGH as a female parent carried the cytoplasm of the cultivated *S. melongena* ASI-S-1. Subsequently, the S0 progenies were selfed for three generations by a single seed descent (SSD) to obtain the S3 segregating individuals that were phenotyped and genotyped in this study. To ensure the continuity of the S0 progenies and to accelerate the self-fertilisation process, four plants of each S0 progeny were germinated, and only the first two that set a viable seed were selected for the next generation (S1; Supplementary Figure 1B). From each of the two S0 selected plants, two S1 plants were germinated, and only the first setting fruit was selected for the S2 generation. The same was done for the S3 generation so that for each progeny, two plants were germinated and phenotyped. Despite this, only one was used for originating the next generation. On the other hand, in the S3 progenies, if two individuals displayed some phenotypic differences, both of them were included in the S3MEGGIC population.

Cultivation Conditions

Seeds were germinated in Petri dishes, following the protocol developed by Ranil et al. (2015) and, subsequently, transferred to seedling trays in a climatic chamber under a photoperiod and temperature regime of 16 h light (25°C) and 8 h dark (18°C). After acclimatisation, plantlets were transplanted to 15 L pots and grown in a pollinator-free benched glasshouse of the Universitat Politècnica de València (UPV), Valencia, Spain (GPS coordinates: latitude, 39° 28′ 55″ N; longitude, 0° 20′ 11″ W; 7 m above sea level). Plants were spaced 1.2 m between rows and 1 m within the row, and fertirrigated using a drip irrigation system and trained with vertical strings. Pruning was done manually to regulate the vegetative growth and flowering. Phytosanitary treatments were performed when necessary. In order to shorten generation time of subsequent generations (S0–S3), plantlets were transplanted to individual thermoformed pots (1.3 L capacity) in a pollinator-free glasshouse, and selfings were stimulated by a mechanical vibration.

High-Throughput Genotyping

Young leaf tissue was sampled from 420 S3 individuals, the eight founders, and the four two-way hybrids. Genomic DNA was extracted using the silica matrix extraction (SILEX) extraction method (Vilanova et al., 2020) and checked for quality and integrity by agarose electrophoresis and NanoDrop ratios (260/280 and 260/230), while its concentration was estimated with Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, United States). After dilution, the samples were sent to Identity Governance and Administration (IGA) Technology Services (IGATech, Udine, Italy) for library preparation and sequencing with NextSeq500 sequencer (150 paired-end) for a high-throughput genotyping using the SPET technology using the 5k probes eggplant SPET platform (Barchi et al., 2019a). The latter comprises 5,093 probes, and was developed by filtering out the most informative and reliable polymorphisms (3,372 of them in coding sequences or CDS and 1,721 in introns and untranslated regions or UTR regions) from the set of over 12 million single-nucleotide polymorphisms (SNPs) identified among the MAGIC founders (Gramazio et al., 2019).

Raw reads were demultiplexed and the adapters were removed using a standard Illumina pipeline and Cutadapt (Martin, 2011), while trimming was performed by ERNE (Del Fabbro et al., 2013). Clean reads were mapped onto the eggplant reference genome “67/3” (Barchi et al., 2019b) using BWA-MEM (Li, 2013) with default parameters, and only the uniquely aligned reads were selected for the variant calling performed with GATK 4.0 (DePristo et al., 2011), following the best practice recommended by the Broad Institute.¹

The SNPs identified by SPET were filtered using the Trait Analysis by Association, Evolution, and Linkage (TASSEL) software (ver. 5.0, Bradbury et al., 2007) in order to retain the most reliable ones (minor allele frequency > 0.01, missing data < 10% and maximum marker heterozygosity < 70%). In addition, a linkage disequilibrium (LD) k-nearest neighbour

¹<http://www.broadinstitute.org>

genotype imputation method (LD KNNi) was performed to fill the missing calls or genotyping gaps.

Population Structure, Heterozygosity, and Haplotype Blocks Inferring

A principal component analysis (PCA) was performed to assess the population structure of S3MEGGIC using the R package *vcfR* (Knaus and Grünwald, 2017) and the function *gPCA* of the *Adegenet* package (Jombart, 2008). Finally, the PCA was graphically plotted with *ggplot2* (Wickham, 2016). An Analysis of Molecular Variance (AMOVA) was performed to estimate the population differentiation according to the cytoplasm (cultivated vs. wild) of the individuals of the S3MEGGIC population by using the function *poppr.amova* of the *poppr* R package (Kamvar et al., 2014). The residual heterozygosity and its distribution were evaluated with *TASSEL* software (ver. 5.0, Bradbury et al., 2007). Parental contribution to S3MEGGIC individuals and haplotype blocks were estimated by using R-package *HaploBlocker* (Pook et al., 2019).

Phenotyping and Genome-Wide Association Study

Phenotypic data were collected from the 420 S3 individuals grown during the 2019/2020 season. Three traits were screened using a binary classification (presence/absence), namely, anthocyanins presence in vegetative plant tissues (PA) and fruit epidermis (FA), and anthocyanic PUC (**Supplementary Figure 4**). The presence of PA was phenotyped when the purple coloration was observed in any vegetative plant parts such as in stem, branches, leaf veins, or prickles. For FA, anthocyanins were considered as present when the purple coloration was observed in the fruit epidermis regardless of their distribution (uniform, listed, etc.) or intensity. The PUC trait could only be phenotyped in anthocyanic fruits by removing the calyx and observing the presence of anthocyanins under it. FA and PUC traits were screened at the stage of commercial maturity (e.g., when the fruit was physiologically immature) which is the best stage for phenotyping these traits. Phenotypic characteristics of the eight founders and two-way hybrids are described in **Figure 1A**.

Using the phenotypic and genotypic data collected from the S3MEGGIC individuals, a Genome-Wide Association Study (GWAS) was performed for the selected traits using the *TASSEL* software (ver. 5.0, Bradbury et al., 2007). For the association study, mixed linear model (MLM) analyses were conducted. The MLM analysis uses both fixed and random effects, which incorporates the kinship among the individuals. The multiple testing was corrected with the Bonferroni and the false discovery rate (FDR) methods (Holm, 1979; Benjamini and Hochberg, 1995) to identify candidate-associated regions at the significance level of 0.05 (Thissen et al., 2002). SNPs with limit of detection (LOD) [$-\log_{10}(p\text{-value})$] over these thresholds or cut off values were declared to be significantly associated with the anthocyanin presence. The R *qqman* package (Turner, 2018) was used to visualise the Manhattan plots and *LDBlockShow* (Dong et al., 2020) to determine the LD and the plot haplotype block structure. The pattern of pairwise LD between SNPs was measured by

LD correlation coefficient (r^2) by considering haplotype blocks with default r^2 values greater than 0.5 supported by the solid spine of LD method (Gabriel et al., 2002; Barrett et al., 2005). The LD was used to narrow down the genomic regions with significant associations. The genes underlying the associated regions were retrieved from the “67/3” eggplant reference genome (ver. 3) (Barchi et al., 2019b). Genes were considered as potential candidates in controlling the assessed traits when carrying homozygous allelic variants classified as “high impact” according to *SnPEff* software v 4.2 prediction (Cingolani et al., 2012) of the eight MAGIC founders (Gramazio et al., 2019). The Integrative Genomics Viewer (IGV) tool was used for visual exploration of founder genome sequences to validate *SnPEff* results and to confirm the presence of the so-called “high impact” variants (Robinson et al., 2020). In addition, eggplant and tomato syntenies were assessed by a BLASTx search of candidate genes sequences against the tomato genome (version SL4.0) in the Sol Genomics Network database.²

RESULTS

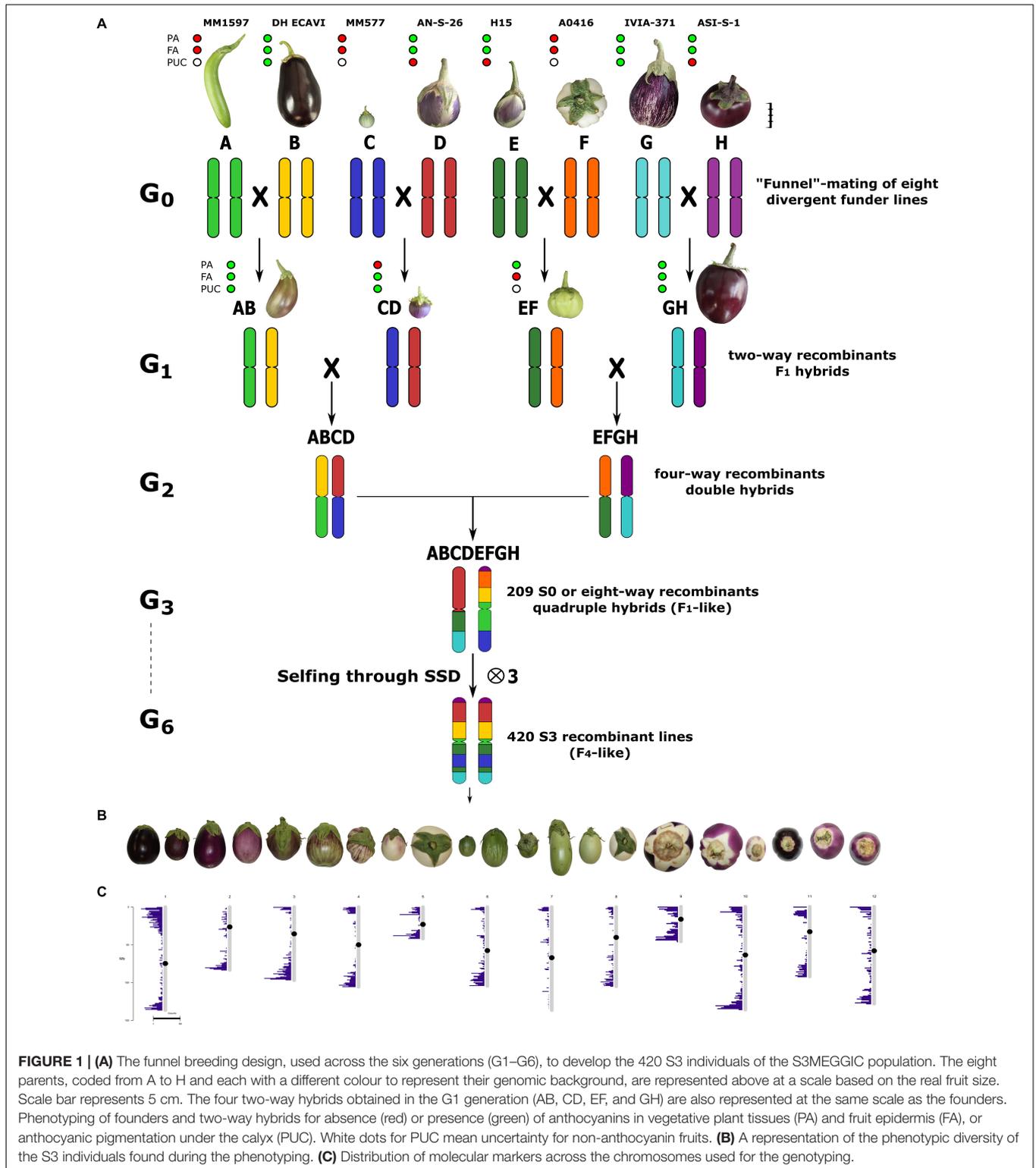
Multi-Parent Advanced Generation Inter-Cross Population Construction

Seven accessions of eggplant and one of the wild relative *S. incanum* were selected as founder parents (A-H) for the construction of the eggplant MAGIC population (S3MEGGIC). Following a funnel breeding scheme (**Figure 1A** and **Supplementary Figure 1**), a total of 420 individuals of the MAGIC populations were obtained. First, founders were pairwise inter-crossed to produce four two-way hybrids (AB, CD, EF, and GH), which were subsequently inter-crossed in pairs to obtain four-way hybrids (ABCD and EFGH). One-hundred and forty-nine individuals of each of the two four-way hybrids were inter-crossed using a chain pollination scheme (**Supplementary Figure 1**). Out of the theoretical maximum of 298 eight-way hybrid progenies (S0), seeds were obtained for 209 of them, of which 116 carried the *S. melongena* ASI-S-1 cytoplasm and 93 carried the *S. incanum* MM577 cytoplasm. Two plants per S0 progeny were used to advance the population, reaching 402 S1 progenies. These S1 progenies were advanced through single seed descend (SSD) to obtain the 391 S2 and 305 S3 MAGIC progenies. The final S3MEGGIC population was constituted by 420 S3 individuals, of which 348 individuals carried the cultivated cytoplasm and 72 the wild cytoplasm.

Single Primer Enrichment Technology Genotyping

The genotyping of the 420 S3 MAGIC individuals, the eight founders, and the four two-way hybrids by the eggplant SPET platform yielded 22,146 SNPs. After filtering, 7,724 high-confidence SNPs were retained for the subsequent analysis, and the low percentage of missing calls (0.53%) was imputed. Filtered SNPs were distributed across the entire

²<http://www.solgenomics.net>

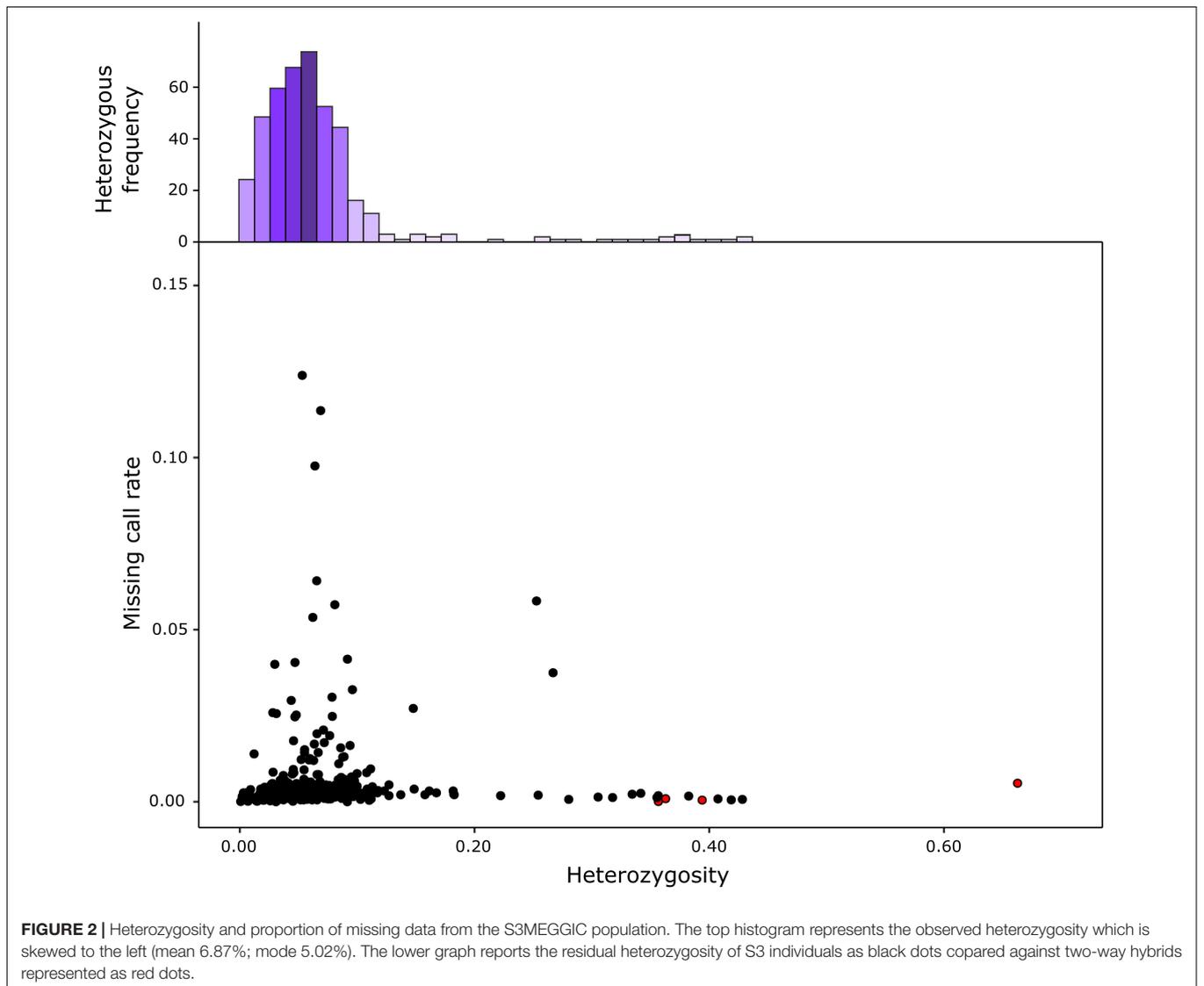


eggplant genome, although the distribution of SNPs varied within and among chromosomes (Table 1 and Figure 1C). Chromosome 9 had the highest average marker density after the SNP filtering, with 173.07 SNPs per Mb, while

chromosome 7 had the lowest with an average of 23.03 SNPs per Mb. Generally, most of the SNPs were located in regions with a high gene density and decayed around the centromere (Figure 1C). The S₃ MAGIC individuals exhibited

TABLE 1 | Statistics of the genotyping using the eggplant 5k probes Single Primer Enrichment Technology (SPET) platform of the 420 S3MEGGIC population individuals using the “67/3” eggplant reference genome (Barchi et al., 2019b).

Chr	Markers	Filtered markers	% Markers	% Filtered markers	Chr length (Mb)	Marker density (Mb)	Filtered marker density
1	3,234	1,205	14.60	15.60	13.66	236.82	88.24
2	1,119	424	5.05	5.49	8.34	134.21	50.85
3	2,004	758	9.05	9.81	9.71	206.41	78.07
4	1,567	588	7.08	7.61	10.57	148.23	55.62
5	1,438	571	6.49	7.39	4.39	327.80	130.16
6	2,108	794	9.52	10.28	10.90	193.38	72.84
7	2,483	328	11.21	4.25	14.24	174.34	23.03
8	1,392	550	6.29	7.12	10.96	126.98	50.17
9	1,841	625	8.31	8.09	3.61	509.80	173.07
10	2,069	798	9.34	10.33	10.67	193.86	74.77
11	1,260	453	5.69	5.86	7.23	174.23	62.64
12	1,631	630	7.36	8.16	10.05	162.33	62.70
Total	22,146	7,724	100.00	100.00	114.33		
Average						215.70	76.85



a heterozygosity average of 6.87%, with only 15 individuals (3.57%) with a proportion of residual heterozygosity higher than 20% (Figure 2).

Population Structure

Population stratification, performed by PCA, indicated the absence of subgroups in the S3 individuals since no clear clustering was observed (Figure 3). The first two principal components (PCs) account for 5.15% (PC1) and 3.30% of the total variation, respectively, while the first 10 PCs explain together only 25.14% of the total variation, revealing the absence of genetic structure in the S3MEGGIC population. No differentiated clusters among individuals carrying the wild (*S. incanum* MM577) or the cultivated (*S. melongena* ASI-S-1) cytoplasm were observed. An Analysis of Molecular Variance (AMOVA) was also performed, revealing that only 0.29% of the total sums of squares is accounted for the molecular variation among the *S. melongena* and *S. incanum* cytoplasm groups, thereby resulting in a very small phi-value of 0.0019, which indicates a low level of differentiation and supporting that no population structure exists.

The genome mosaics reconstruction of the S3 MAGIC individuals, in terms of the eight founder haplotypes, showed different haplotype block proportions depending on the genomic position for all chromosomes (Figure 4). The estimated contribution of some founders to the overall S3MEGGIC population differed from the expected value of 12.5%. Two of the founder genomes (A0416 and IVIA-371) had a high representation in the genome of the S3 individuals (32.6% and 23.6%, respectively), while two others (AN-S-26 and H15) had a small representation (0.3% in both cases). The wild founder, *S. incanum*, had an average haplotype representation of 5.8%.

Phenotypic Variation Among Multi-Parent Advanced Generation Inter-Cross Individuals and Association Analysis

The screening for absence or presence of PA, FA, and PUC in purple fruits of the 420 S3 MAGIC individuals revealed a considerable variation (Figure 1B). Out of the 420 S3 individuals, 57.6% displayed PA and 37.5% FA. Among the individuals displaying FA, 64.3% had the PUC phenotype.

Given that no population structure was observed for the S3MEGGIC population, the phenotypic data, together with the genotypic information, were used for GWAS analysis (Figure 5). The GWAS was performed taking into account the kinship in an MLM leading to the identification of significant associations for the evaluated traits.

Plant Anthocyanins

The Manhattan plot for PA revealed two major peaks on chromosome 9 and 10 (PA9 and PA10), with seven significant SNPs over the FDR threshold (LOD > 4.09), and three of them over the Bonferroni threshold (LOD > 5.16) (Figure 5A). One SNP (LOD = 4.25) was mapped on PA9 region (between 17.0 and 17.4 Mb; Supplementary Figure 2A) and six SNPs, with an LOD

between 4.31 and 6.48, were mapped on a PA10 region (between 91.08 and 94.81 Mb; Figure 5A).

Fruit Anthocyanins

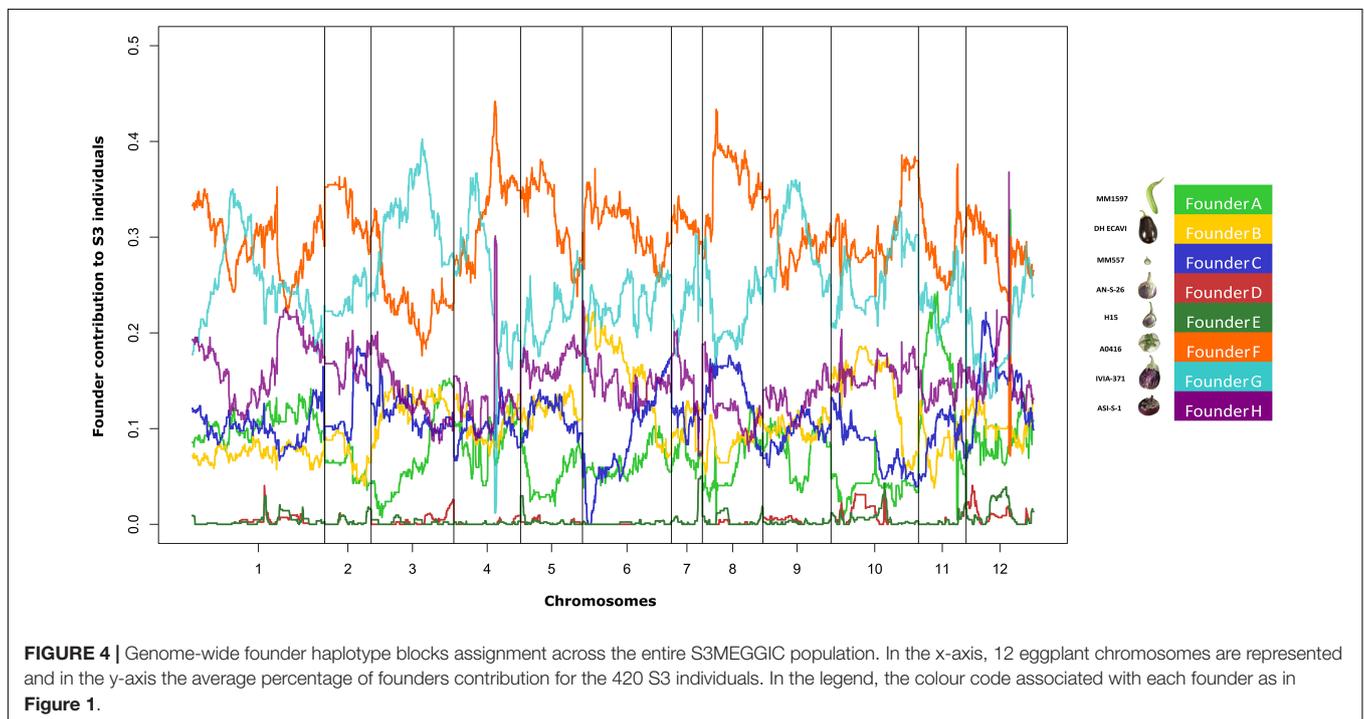
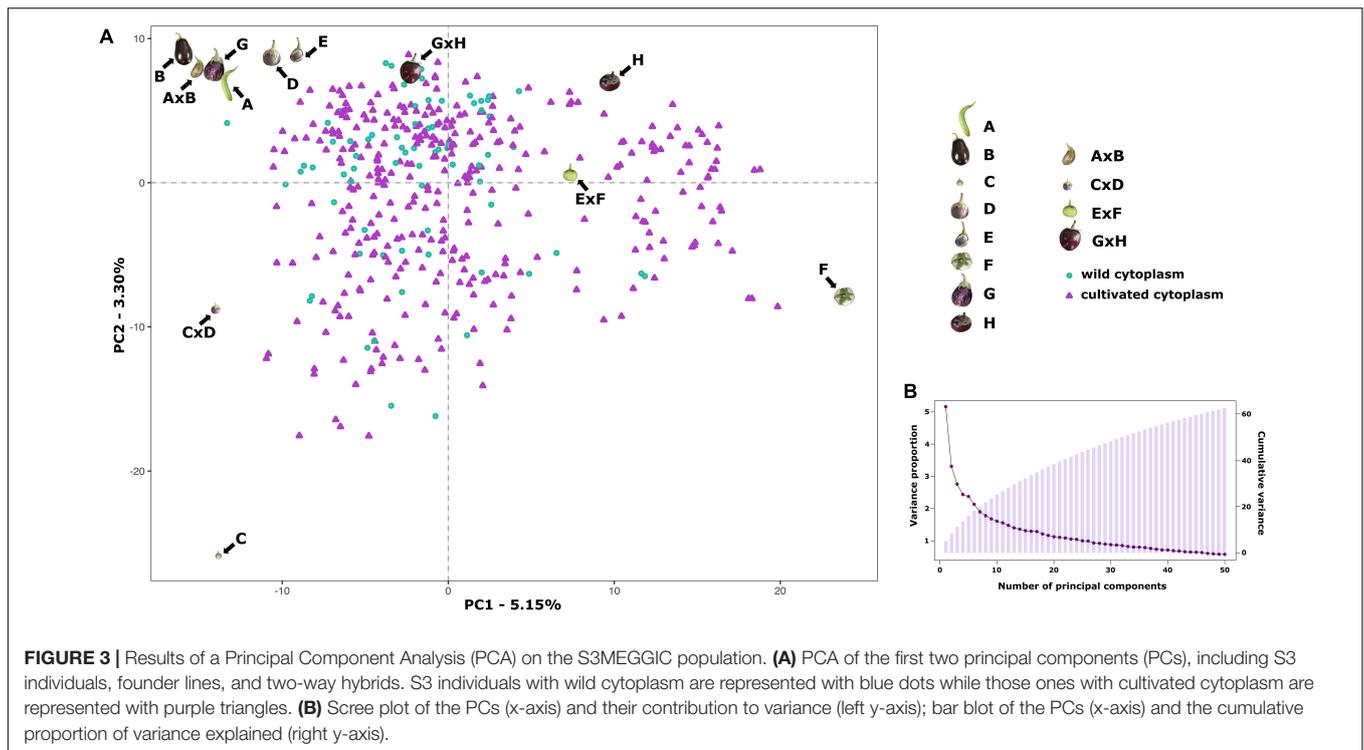
For FA, 22 significant SNPs above FDR threshold (LOD > 3.6), which included 11 SNPs above the Bonferroni threshold (LOD > 5.16), were plotted on three major peaks located on chromosomes 1, 9, and 10 (FA1, FA9, and FA10, respectively) (Figure 5C). One SNP (LOD = 4.07) was detected on FA1 region (between 5.11 and 5.88 Mb) in position 5,346,977 (Figure 5D). Five SNPs, with an LOD between 3.86 and 4.22, were located on FA9 region (between 16.2 and 17.4 Mb; Supplementary Figure 2B), which overlapped with the PA9 region (Supplementary Figure 2A). The other sixteen SNPs (LOD between 4.52 and 9.06) were detected on FA10 region between 91.08 and 94.81 Mb (Figure 5E), corresponding to the PA10 region, where significant associations were detected for the PA (Figure 5B).

Pigmentation Under the Calyx

Among the three traits evaluated, PUC was the one with the highest number of significant associations. The Manhattan plot for PUC revealed seven main peaks on chromosomes 1 (PUC1), 3 (PUC3.1 and PUC3.2), and 10 (PUC10.1–10.4), with 36 SNPs over the FDR threshold (LOD > 3.7), and 23 of them are above the Bonferroni threshold (LOD > 5.16) (Figure 5F). One SNP with an LOD of 4.21 was located on the PUC1 region (between 5.11 and 5.88 Mb) in position 5,480,282 (Figure 5G), 133,305 bp away of the significant SNP on the FA1 region (Figure 5D). On the chromosome 3, four SNPs (LOD between 4.41 and 6.18) were detected on the PUC3.1 region (between 7.22 and 8.65 Mb; Supplementary Figure 2C) and one SNP (LOD = 4.58) on the PUC3.2 region (between 31.54 and 40.09 Mb; Supplementary Figure 2D). Thirty SNPs were mapped on chromosome 10 in four genomic regions. One of the SNPs (LOD = 4.2) was identified on the PUC10.1 region (between 2.08 and 2.67 Mb) in position 2,388,375 (Supplementary Figure 2E), and other one (LOD = 4.2) on the PUC10.2 region (between 3.94 and 4.34 Mb; Figure 5H). Another twelve SNPs were located on the PUC10.3 region, between 91.08 and 94.81 Mb (LOD between 3.86 and 12.44; Figure 5I), as observed for PA (Figure 5B) and FA (Figure 5D). The 16 remaining SNPs (LOD between 3.71 and 7.93) were found in the PUC10.4 region, between 98.25 and 100.55 Mb (Supplementary Figure 2F).

Candidate Genes for Anthocyanin Biosynthesis

Based on the results of the GWAS analysis, putative candidate genes were identified close to or within the LD blocks defined in the genomic regions with significant associations (Supplementary Table 1). On chromosome 10, in the genomic region (91.08–94.81 Mb) associated with all the evaluated traits (PA10, FA10, and PUC10.3), a candidate gene was identified as similar to *MYB113* (SMEL_010g351850.1), a well-known regulatory transcription factor controlling anthocyanin synthesis in eggplant (Zhou et al., 2019; Shi et al., 2021). In addition, in the



genomic region associated with FA and PUC (FA1 and PUC1) on chromosome 1 (5.11–5.88 Mb), another candidate gene was identified similar to *MYB113* (SMEL_001g120500.1). Variants that predicted high impact effects on protein function were annotated by SnpEff for both *MYB113* genes in the population founders (A, C, and F) that do not present anthocyanins in plants

and fruits, as confirmed by aligning the founder gene sequences (Figures 5J,K). Specifically, for the founders A and C, single frameshift variants were identified in two different positions while a disruptive inframe deletion and a slice region variant were identified for the founder F in a third region of the eggplant gene SMEL_010g351850.1 (Figure 5J). For SMEL_001g120500.1,

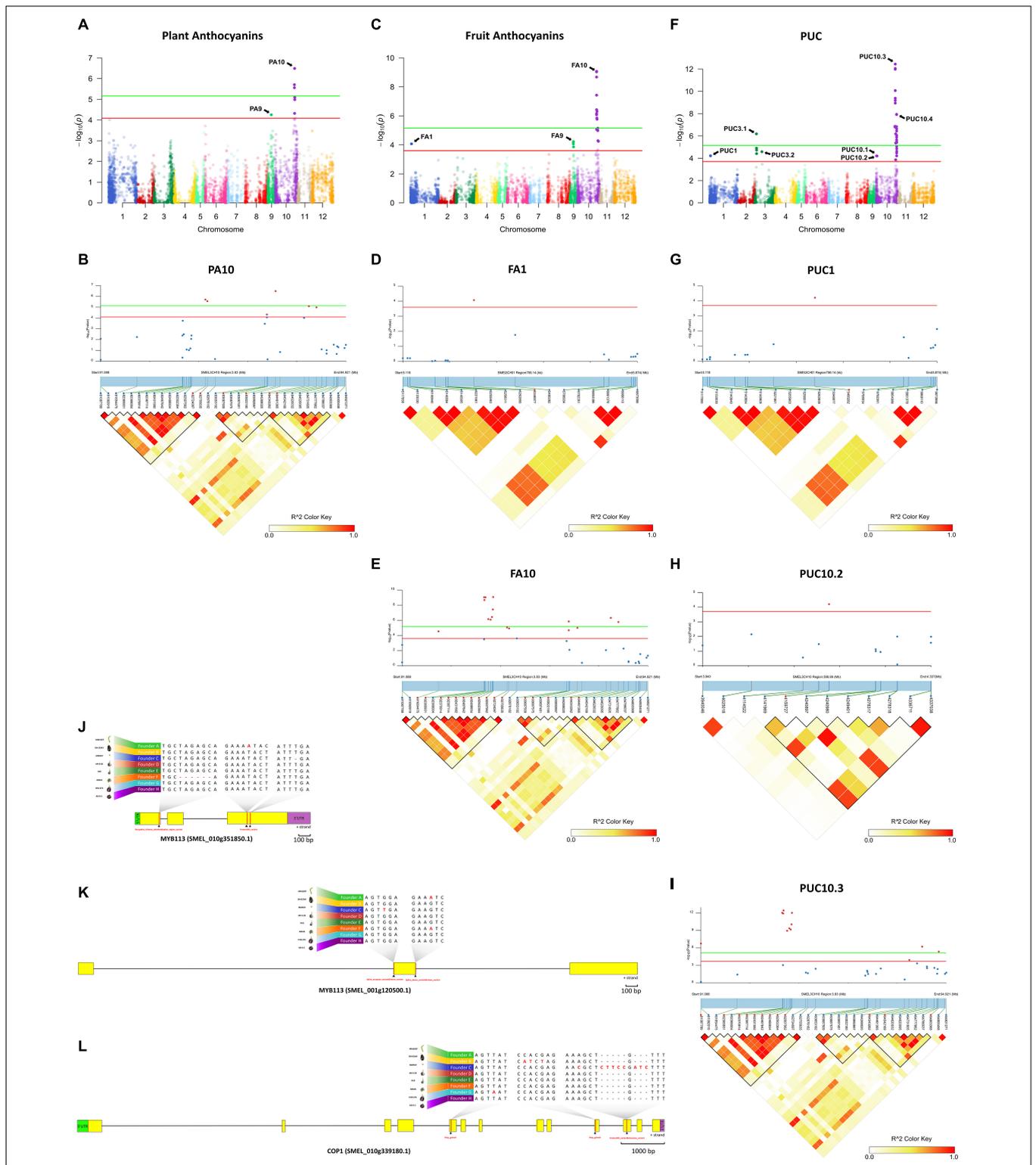
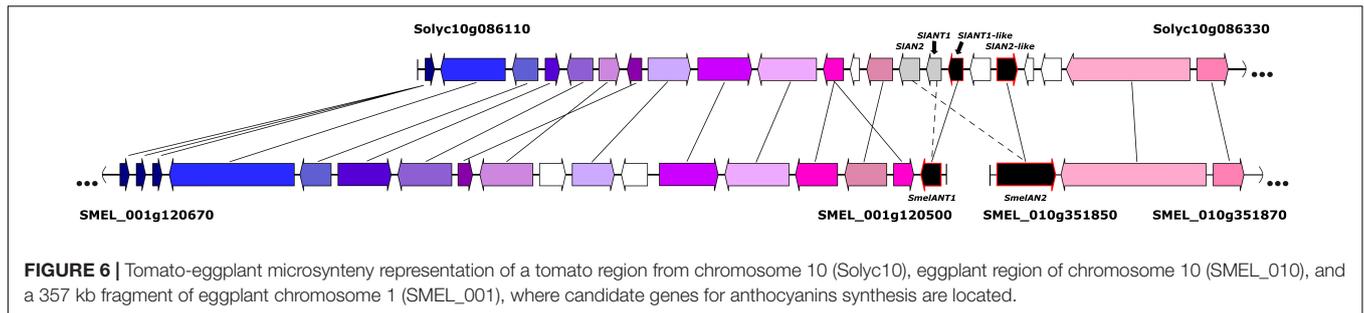


FIGURE 5 | Genome-wide association mapping with linkage disequilibrium (LD) block heatmap of significant regions and potential candidate genes controlling PA, FA, and PUC in eggplant. **(A,C,F)** Manhattan plot for PA **(A)**, FA **(C)**, and PUC **(F)**. Arrows indicate the position of main peaks detected for each trait. The red and green horizontal lines represent, respectively, FDR and Bonferroni significance thresholds at $p = 0.05$. **(B,D,E,G-I)** Local Manhattan plot (top) and LD heatmap (bottom) surrounding the peaks PA10 **(B)**, FA1 **(D)**, FA10 **(E)**, PUC1 **(G)**, PUC10.2 **(H)** and PUC10.3 **(I)**. Pairwise LD between SNPs is indicated as values of R^2 values: red indicates a value of 1 and white indicates 0. **(J-L)** Structure of candidate genes MYB113 (SMEL_010g351850.1) **(J)**, MYB113 (SMEL_001g120500.1) **(K)**, and COP1 (SMEL_010g339180.1) **(L)** and effect of the high-impact variants detected in those of the eight founders.



founders A and F exhibited the exact same splice donor, and the intron variant predicted as high impact, while in the founder C, a splice acceptor and intron variant were identified (Figure 5K). Founder reconstruction of the S3 individuals and haplotype blocks were estimated for both candidate gene regions (chromosome 10: 91.08–94.81 Mb; chromosome 1: 5.11–5.88 Mb) (Supplementary Figures 3A,B). Founders that contribute most to anthocyanin-related traits in these regions are IVIA-371 (111 and 97 S3 individuals, respectively) and ASI-S-1 (74 and 75 S3 individuals, respectively). Furthermore, a reciprocal best hit BLAST analysis of the SMEL_001g120500.1 onto the tomato genome indicated that this gene corresponded to the orthologue *SIANT1*-like with 70.99% identity. The eggplant SMEL_010g351850.1 corresponded to the tomato orthologue *SIAN2*-like with a 73.73% identity, which has been described as the best candidate gene for the anthocyanin fruit biosynthesis in tomato (Yan et al., 2020). These results suggest that a single duplication event and a translocation of a fragment of at least 357 kb from chromosome 10 to 1 occurred during eggplant evolution (Figure 6).

A candidate gene that corresponds to a *COPI* (SMEL_010g339180.1), a gene that encodes for photo-regulatory proteins (Jiang et al., 2016; Li et al., 2018; He et al., 2019; Naeem et al., 2019), was reported in the genomic regions associated with PUC (PUC10.2) on chromosome 10 (3.94–4.34 Mb; Supplementary Table 1). High-effect variants were detected in *COPI* gene in those founders that are able to synthesise anthocyanin under the calyx (B and G) and in the green wild species *S. incanum* (C) (Figure 5L). The C founder mutation was confirmed by the F1 hybrid phenotype from the C × D inter-cross, PUC (anthocyanic pigmentation under the calyx) × *puc* (no anthocyanic pigmentation under the calyx), which showed the anthocyanin fruits with pigmentation under the calyx (Figure 1A). The wild founder presented multiple SNPs compared with the rest that were predicted to cause a frameshift and missed variants, while founders B and G exhibited variants that were predicted to produce stop codons (Figure 5L). Founders with a higher haplotype representation contributing to PUC in this region (chromosome 10: 3.94–4.34 Mb) are IVIA-371 (113 S3 individuals) and DH_ECAVI (71 S3 individuals) (Supplementary Figure 3C).

Furthermore, we found two candidate genes annotated as similar to *BHLH*, basic helix loop helix protein A (SMEL_009g326640.1), and, similar to *SPA3*, protein SPA1-RELATED 3 (SMEL_010g338090.1), respectively (data

not shown), close to LD blocks in the PA9 and FA9 (between 17,862,090 and 17,872,428 Mb), and PUC10.1 (between 3,128,678 and 3,143,068 Mb) regions. Although they were described as genes related to the anthocyanin biosynthetic pathway (Maier et al., 2013; Li et al., 2018; Shi et al., 2021), no high-effects variants were identified in these genes for any of the founders.

DISCUSSION

Multi-parent advanced generation inter-cross (MAGIC) populations are outstanding genetic materials for identifying gene-trait associations with high resolution (Arrones et al., 2020; Scott et al., 2020). The introduction of multiple founders with an increased genetic and phenotypic diversity, together with the multiple rounds of inter-crossing and selfing, increases the number of accumulated recombinant events and, thus, improves the mapping accuracy (Scott et al., 2020). By introducing a wild relative as a founder parent, the genetic variability in the population increases, which is a key point for QTL identification (Gramazio et al., 2020a). Here, we present the first eggplant S3MEGGIC population of which one of the founders was one accession of the close wild relative *S. incanum*.

Large population sizes are essential to increase the power and mapping resolution in MAGIC populations (Collard et al., 2005; Valdar et al., 2006; Jaganathan et al., 2020). Following a simple “funnel” scheme design, the population was kept as large as possible to gather a large number of recombination events. However, a sharp reduction in the number of progenies was observed at the S0 generation, which might be related to the use of the wild species *S. incanum* (C) as a female founder and parent to obtain the simple (CD) and the double (ABCD) hybrids. This interspecific crossing dragged the maternal cytoplasmic background of the wild parent, which might have caused a partial sterility and bias in subsequent generations. Some studies confirmed a strong effect of wild *Solanum* cytoplasm in the reduction in pollen fertility of alloplasmic lines (Khan et al., 2020; Isshiki et al., 2021). However, the PCA highlighted the absence of a population structure, also confirmed by the lack of genetically differentiated cytoplasmic groups.

The genotyping of the S3MEGGIC population was carried out with the 5k probes eggplant SPET platform with a well-distributed marker density along all chromosomes (Barchi et al., 2019a). This genotyping strategy has already been used in the analysis of biparental populations (Herrero et al., 2020).

In this study, we have verified that its use can be extended to multiparent populations. The genotyping revealed a low heterozygosity in the S3MEGGIC individuals, similar to the expected value for an F5-like biparental inter-cross generation (6.25%). The contributions of each of the founder parents to the S3MEGGIC population revealed that some parents had a higher representation than others. Apart from drift effects, several biological reasons could potentially explain cryptic selection processes, causing the unbalanced representation of the genomes (Rockman and Kruglyak, 2008; Thépot et al., 2015), including seed dormancy, delayed germination, precocity, reduced fertility, and parthenocarpy associated to some genomes which have already been reported in eggplant and other crops (Barchi et al., 2010; Khan et al., 2015; Prohens et al., 2017). The rather limited contribution of the wild species *S. incanum* to the final S3 MAGIC individuals may have been caused by a selection pressure, as progenies bred from crosses involving the two different species tend to suffer from a reduced fertility and show a segregation distortion (Lefebvre et al., 2002; Barchi et al., 2010). In addition, *S. incanum* has a recalcitrant germination and a very erratic flowering and fruit set, which strongly depends on environmental conditions (Gisbert et al., 2011; Mangino et al., 2020, 2021). Other reasons for the segregation distortion could be the inability of the current genotyping density to efficiently distinguish between the founders that are genetically closer, like AN-S-26 and H15 genotypes. This phenomenon has already been observed in previous MAGIC populations (Dell'Acqua et al., 2015). A deeper genotyping, resulting in a better haplotype reconstruction, might shed light on the mechanisms that have led to the unbalanced representation of the founder genomes in the S3MEGGIC population. Due to this low contribution, care should be taken in the future when analysing the traits that are present only in these founders.

The anthocyanin biosynthetic pathway is one of the most studied biochemical routes in plants due to its physiological importance. Major structural genes of this pathway are under the control of a regulatory complex, where myeloblastosis (MYB) TFs are recognised as main regulators alone or in complexes with other TFs (Ramsay and Glover, 2005; Kiferle et al., 2015; Liu et al., 2018). Activator or repressor MYB proteins directly and competitively bind the basic-helix-loop-helix (bHLH) *via* the amino terminus domain and can act as positive or negative transcriptional regulators in a tissue-specific mode to modulate the anthocyanin synthesis (Barchi et al., 2019b; Moglia et al., 2020).

In eggplant, anthocyanin-related MYB protein-encoding genes have been reported to be related to fruit peel colouration (Zhang et al., 2014; Docimo et al., 2016; Xiao et al., 2018; Moglia et al., 2020; Toppino et al., 2020). In tomato, a cluster of four different MYB proteins was reported to be involved in the anthocyanin synthesis located on chromosome 10 and encoded by *SIAN2*, *SIANT1*, *SIANT1*-like, and *SIAN2*-like genes (Solyc10g086250, Solyc10g086260, Solyc10g086270, and Solyc10g086290, respectively). However, genetic associations to only two paralogue *MYB113* genes were detected in the S3MEGGIC population on chromosomes 1 and 10 (SMEL_001g120500.1 and SMEL_010g351850.1). The same

occurs in potato, in which, only two anthocyanin genes have been identified, i.e., Sotub10g028550 and Sotub10g028540, both located on chromosome 10. Similarly, in pepper, in addition to CA10g11690 and CA10g11650, a third gene (CA10g11710) has been identified as an orthologue to *SIAN2*-like (Barchi et al., 2019b). In eggplant, these two tomato orthologs were previously described as *SmelANT1* and *SmelAN2* (Docimo et al., 2016; Barchi et al., 2019b), corresponding to *SIANT1* and *SIAN2*. However, the reciprocal best hit BLAST analysis of the eggplant coding proteins showed a stronger homology, respectively, to tomato *SIANT1*-like and *SIAN2*-like. Although previous studies in eggplant indicated that the overexpression of *SmelANT1* accounts for constitutive upregulation of most anthocyanin biosynthetic genes (Zhang et al., 2014; Shi et al., 2021), we considered the *SmelAN2* as the best candidate gene for different reasons. Previous studies in eggplant have located the major anthocyanin-related QTLs on chromosome 10 (Barchi et al., 2012; Cericola et al., 2014; Toppino et al., 2016), which are in agreement with the highest association signals for anthocyanin-related traits in our GWAS results. Furthermore, recent studies in tomato suggested that *SIAN2*-like functions as an activator to regulate biosynthesis genes, including *SIANT1*-like, and controls the accumulation of anthocyanins (Yan et al., 2020). However, in the S3MEGGIC population, high-impact variants on protein function were found in both *MYB113* genes for non-anthocyanin fruits. These results could suggest that a duplication of function occurred during an eggplant evolution, and both genes may be required for anthocyanin synthesis.

Although it has been demonstrated that the activation of the anthocyanin biosynthetic pathway in eggplant is strongly regulated by light (Li et al., 2018; Xiao et al., 2018), the PUC mutation confers to some genotypes the ability to synthesise anthocyanins under the calyx regardless of incidence of light (Tigchelaar et al., 1968). In this study, a candidate gene, related with light-dependent anthocyanin biosynthesis in fruits, was detected at the beginning of chromosome 10. The constitutive photomorphogenic1 (*COP1*) gene has been reported to be a regulatory TF responsible for mediating a light-regulated gene expression and development (Jiang et al., 2016; Li et al., 2018; He et al., 2019; Naeem et al., 2019). The *COP1* gene has been demonstrated to act as a light-inactivable repressor interacting with MYB TFs (Jiang et al., 2016), and is considered a “molecular switch” in metabolic processes, which are stimulated by light. In presence of light, *COP1* expression is inhibited, and its concentration decreases rapidly, promoting anthocyanins synthesis. Under dark conditions, *COP1* expression is induced and promotes the degradation of the photomorphogenesis-promoting TF and MYB inhibition by conforming the *COP1*/Suppressor of *phyA*-105 (*SPA*) ubiquitin ligase complex (Maier et al., 2013; Jiang et al., 2016; Li et al., 2018). Although PUC can only be observed in anthocyanic fruits, high-impact variants were also identified in the wild species *S. incanum* *COP1* gene (founder C, green fruit). The PUC allele is more common in eggplants from western countries, where European markets demand for more homogeneously pigmentation eggplants. On the other hand, in Asian

eggplants, it is more common to find the *puc* phenotype, such as in the ASI-S-1 founder.

CONCLUSION

In conclusion, the S3MEGGIC population represents a landmark breeding material and a tool of great value, which allows the study and fine-mapping of complex traits due to (i) the highly phenotypically diverse founders; (ii) the large population size, being the largest developed eggplant experimental population so far; (iii) the high degree of homozygosity of the final individuals, which constitute a population of fixed “immortal” lines nearly homozygous at each locus; and (iv) the tailored genotyping SPET platform used for the genetic analysis of the population, which has been developed from the whole genome sequencing (WGS) of the founders and allows the comparison with the genotyped materials with the same setup (Gramazio et al., 2020b). In addition, the S3MEGGIC population has demonstrated its potential usefulness for association studies, allowing the establishment of marker associations to anthocyanin-related genes and identification of candidate genes for plant and fruit anthocyanins, including the first identification of a candidate gene for an economically relevant breeding trait in eggplant such as PUC.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/>, PRJNA392603.

AUTHOR CONTRIBUTIONS

SV, PG, and JP conceived the idea and supervised the manuscript. GM, AA, MP, and PG performed the field trials. GM and AA prepared a first draft of the manuscript. All other authors reviewed and edited the manuscript.

REFERENCES

- Albert, N. W., Davies, K. M., Lewis, D. H., Zhang, H., Montefiori, M., Brendolise, C., et al. (2014). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in Eudicots. *Plant Cell* 26, 962–980. doi: 10.1105/tpc.113.122069
- Arrones, A., Vilanova, S., Plazas, M., Mangino, G., Pascual, L., Díez, M. J., et al. (2020). The dawn of the age of multi-parent magic populations in plant breeding: novel powerful next-generation resources for genetic analysis and selection of recombinant elite material. *Biology* 9, 1–25. doi: 10.3390/biology9080229
- Bandillo, N., Raghavan, C., Muyco, P. A., Sevilla, M. A. L., Lobina, I. T., Dilla-Ermita, C. J., et al. (2013). Multi-parent advanced generation inter-cross

FUNDING

This work was supported by the Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación and Fondo Europeo de Desarrollo Regional (grant RTI2018-094592-B-I00 from MCIU/AEI/FEDER, UE) and European Union’s Horizon, 2020 Research and Innovation Programme under grant agreement no. 677379 (G2P-SOL project: Linking genetic resources, genomes and phenotypes of Solanaceous crops). AA is grateful to Spanish Ministerio de Ciencia, Innovación y Universidades for a pre-doctoral (FPU18/01742) contract. MP is grateful to Spanish Ministerio de Ciencia e Innovación for a post-doctoral grant (IJC2019-039091-I/AEI/10.13039/501100011033). PG is grateful to Spanish Ministerio de Ciencia e Innovación for a post-doctoral grant (FJC2019-038921-I/AEI/10.13039/501100011033). Funding for open access charge: Universitat Politècnica de València.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.847789/full#supplementary-material>

Supplementary Figure 1 | (A) Chain pollination scheme of the four-way hybrids followed to obtain the eight-way hybrids. **(B)** For each S0 progeny, four plants were germinated, selecting for the next generation (S1) only the first two that set fruits with viable seed. For subsequent generations, two plants were germinated and only the first one that set fruit was selected for the next generation.

Supplementary Figure 2 | Local Manhattan plot (top) and LD heatmap (bottom) surrounding the peaks PA9 **(A)**, FA9 **(B)**, PUC3.1 **(C)**, PUC3.2 **(D)**, PUC10.1 **(E)**, and PUC10.4 **(F)**. The red and green horizontal lines represent, respectively, FDR and Bonferroni significance thresholds. Pairwise LD between SNPs is indicated as values of R^2 values: red indicates a value of 1 and white indicates 0.

Supplementary Figure 3 | Founder haplotype blocks representation predicted for each of the S3 individuals for the three anthocyanin-related candidate gene regions: **(A)** MYB113 (SMEL_001g120500.1) on chromosome 1 between 5.11 and 5.88 Mb identified by the FA1 and PUC1 associations; **(B)** COP1 (SMEL_010g339180.1.01) on chromosome 10 between 3.94 and 4.34 Mb identified by the PUC10.2 association; and **(C)** MYB113 (SMEL_010g351850.1) on chromosome 10 between 91.08 and 94.81 Mb identified by the PA10, FA10, and PUC10.3 associations.

Supplementary Figure 4 | Phenotyping of the S3MEGGIC population for presence or absence of PA, FA, and PUC.

- (MAGIC) populations in rice: progress and potential for genetics research and breeding. *Rice* 6, 1–15. doi: 10.1186/1939-8433-6-11
- Barchi, L., Acquadro, A., Alonso, D., Aprea, G., Bassolino, L., Demurtas, O., et al. (2019a). Single Primer Enrichment Technology (SPET) for high-throughput genotyping in tomato and eggplant germplasm. *Front. Plant Sci.* 10:1005. doi: 10.3389/fpls.2019.01005
- Barchi, L., Pietrella, M., Venturini, L., Minio, A., Toppino, L., Acquadro, A., et al. (2019b). A chromosome-anchored eggplant genome sequence reveals key events in Solanaceae evolution. *Sci. Rep.* 9, 1–13. doi: 10.1038/s41598-019-47985-w
- Barchi, L., Lanteri, S., Portis, E., Stägel, A., Valè, G., Toppino, L., et al. (2010). Segregation distortion and linkage analysis in eggplant (*Solanum melongena* L.). *Genome* 53, 805–815. doi: 10.1139/g10-073

- Barchi, L., Lanteri, S., Portis, E., Valè, G., Volante, A., Pulcini, L., et al. (2012). A RAD tag derived marker based eggplant linkage map and the location of QTLs determining anthocyanin pigmentation. *PLoS One* 7:e43740. doi: 10.1371/journal.pone.0043740
- Barrett, J. C., Fry, B., Maller, J., and Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21, 263–265. doi: 10.1093/bioinformatics/bth457
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300.
- Borovsky, Y., Oren-Shamir, M., Ovadia, R., De Jong, W., and Paran, I. (2004). The A locus that controls anthocyanin accumulation in pepper encodes a MYB transcription factor homologous to Anthocyanin2 of *Petunia*. *Theor. Appl. Genet.* 109, 23–29. doi: 10.1007/s00122-004-1625-9
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Campanelli, G., Sestili, S., Acciarri, N., Montemurro, F., Palma, D., Leteo, F., et al. (2019). Multi-parental advances generation inter-cross population, to develop organic tomato genotypes by participatory plant breeding. *Agronomy* 9:119.
- Cericola, F., Portis, E., Lanteri, S., Toppino, L., Barchi, L., Acciarri, N., et al. (2014). Linkage disequilibrium and genome-wide association analysis for anthocyanin pigmentation and fruit color in eggplant. *BMC Genomics* 15:896. doi: 10.1186/1471-2164-15-896
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6, 80–92. doi: 10.4161/fly.19695
- Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., and Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* 142, 169–196.
- D'Amelia, V., Aversano, R., Chiaiese, P., and Carputo, D. (2018). The antioxidant properties of plant flavonoids: their exploitation by molecular plant breeding. *Phytochem. Rev.* 17, 611–625. doi: 10.1007/s11101-018-9568-y
- Daunay, M. C., and Hazra, P. (2012). “Eggplant,” in *Handbook of Vegetables*, eds K. V. Peter and P. Hazra (Houston, TX: Studium Press), 257–322.
- Del Fabbro, C., Scalabrin, S., Morgante, M., and Giorgi, F. M. (2013). An extensive evaluation of read trimming effects on illumina NGS data analysis. *PLoS One* 8:e85024. doi: 10.1371/journal.pone.0085024
- Dell'Acqua, M., Gatti, D. M., Pea, G., Cattonaro, F., Coppens, F., Magris, G., et al. (2015). Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in *Zea mays*. *Genome Biol.* 16, 1–23. doi: 10.1186/s13059-015-0716-z
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498. doi: 10.1038/ng.806
- Díez, M. J., Picó, B., and Nuez, F. (2002). *Cucurbit Genetic Resources in Europe: Ad Hoc Meeting held in Adana, Turkey, 19 January 2002*. Rome: International Plant Genetic Resources Institute.
- Docimo, T., Francese, G., Ruggiero, A., Batelli, G., De Palma, M., Bassolino, L., et al. (2016). Phenylpropanoids accumulation in eggplant fruit: characterization of biosynthetic genes and regulation by a MYB transcription factor. *Front. Plant Sci.* 6:1233. doi: 10.3389/fpls.2015.01233
- Doganlar, S., Frary, A., Daunay, M. C., Lester, R. N., and Tanksley, S. D. (2002). Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. *Genetics* 161, 1713–1726. doi: 10.1093/genetics/161.4.1713
- Dong, S. S., He, W. M., Ji, J. J., Zhang, C., Guo, Y., and Yang, T. L. (2020). LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Brief. Bioinform.* 22:bbaa227. doi: 10.1093/bib/bbaa227
- FAOSTAT (2019). *FAOSTAT*. Available online at: <http://www.fao.org> (accessed February 20, 2020).
- Frary, A., Frary, A., Daunay, M. C., Huvenaars, K., Mank, R., and Doganlar, S. (2014). QTL hotspots in eggplant (*Solanum melongena*) detected with a high resolution map and CIM analysis. *Euphytica* 197, 211–228.
- Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., et al. (2002). The structure of haplotype blocks in the human genome. *Science* 296, 2225–2229. doi: 10.1126/science.1069424
- Gisbert, C., Prohens, J., and Nuez, F. (2011). Treatments for improving seed germination in eggplant and related species. *Acta Horticult.* 898, 45–52.
- Gonzali, S., Mazzucato, A., and Perata, P. (2009). Purple as a tomato: towards high anthocyanin tomatoes. *Trends Plant Sci.* 14, 237–241. doi: 10.1016/j.tplants.2009.02.001
- Gramazio, P., Pereira-Dias, L., Vilanova, S., Prohens, J., Soler, S., Esteras, J., et al. (2020a). Morphoagronomic characterization and whole-genome resequencing of eight highly diverse wild and weedy *S. pimpinellifolium* and *S. lycopersicum* var. cerasiforme accessions used for the first interspecific tomato MAGIC population. *Horticult. Res.* 7:174. doi: 10.1038/s41438-020-00395-w
- Gramazio, P., Jaén-Molina, R., Vilanova, S., Prohens, J., Marrero, Á., Caujapé-Castells, J., et al. (2020b). Fostering conservation via an integrated use of conventional approaches and high-throughput SPET genotyping: a case study using the endangered canarian endemics *Solanum lidii* and *S. vespertilio* (Solanaceae). *Front. Plant Sci.* 11:757. doi: 10.3389/fpls.2020.00757
- Gramazio, P., Prohens, J., Borràs, D., Plazas, M., Herraiz, F. J., and Vilanova, S. (2017a). Comparison of transcriptome-derived simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers for genetic fingerprinting, diversity evaluation, and establishment of relationships in eggplants. *Euphytica* 213:264.
- Gramazio, P., Prohens, J., Plazas, M., Mangino, G., Herraiz, F. J., and Vilanova, S. (2017b). Development and genetic characterization of advanced backcross materials and an introgression line population of *Solanum incanum* in a *S. melongena* background. *Front. Plant Sci.* 8:1477. doi: 10.3389/fpls.2017.01477
- Gramazio, P., Prohens, J., Plazas, M., Mangino, G., Herraiz, F. J., García-Fortea, E., et al. (2018). Genomic tools for the enhancement of vegetable crops: a case in eggplant. *Not. Bot. Horti Agrobot. Cluj Napoca* 46, 1–13. doi: 10.15835/nbha46110936
- Gramazio, P., Yan, H., Hasing, T., Vilanova, S., Prohens, J., and Bombarely, A. (2019). Whole-genome resequencing of seven eggplant (*Solanum melongena*) and one wild relative (*S. incanum*) accessions provides new insights and breeding tools for eggplant enhancement. *Front. Plant Sci.* 10:1220. doi: 10.3389/fpls.2019.01220
- He, Y., Chen, H., Zhou, L., Liu, Y., and Chen, H. (2019). Comparative transcription analysis of photosensitive and non-photosensitive eggplants to identify genes involved in dark regulated anthocyanin synthesis. *BMC Genomics* 20:678. doi: 10.1186/s12864-019-6023-4
- Herrero, J., Santika, B., Herrán, A., Erika, P., Sarimana, U., Wendra, F., et al. (2020). Construction of a high density linkage map in Oil Palm using SPET markers. *Sci. Rep.* 10, 1–9. doi: 10.1038/s41598-020-67118-y
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.
- Huang, B. E., Verbyla, K. L., Verbyla, A. P., Raghavan, C., Singh, V. K., Gaur, P., et al. (2015). MAGIC populations in crops: current status and future prospects. *Theor. Appl. Genet.* 128, 999–1017. doi: 10.1007/s00122-015-2506-0
- Hurtado, M., Vilanova, S., Plazas, M., Gramazio, P., Andújar, I., Herraiz, F. J., et al. (2014). Enhancing conservation and use of local vegetable landraces: the Almagro eggplant (*Solanum melongena* L.) case study. *Genet. Resour. Crop Evol.* 61, 787–795.
- Huynh, B. L., Ehlers, J. D., Huang, B. E., Muñoz-Amatriáin, M., Lonardi, S., Santos, J. R. P., et al. (2018). A multi-parent advanced generation inter-cross (MAGIC) population for genetic analysis and improvement of cowpea (*Vigna unguiculata* L. Walp.). *Plant J.* 93, 1129–1142. doi: 10.1111/tpj.13827
- Isshiki, S., Nakamura, I., Ureshino, K., and Khan, M. M. R. (2021). Pollen fertility differences in the progenies obtained from a cross between eggplant (*Solanum melongena* L.) as a seed parent and eggplant cytoplasmic substitution lines as pollen parents. *Austral. J. Crop Sci.* 15, 233–237. doi: 10.21475/ajcs.21.15.02.p2785
- Jaganathan, D., Bohra, A., Thudi, M., and Varshney, R. K. (2020). Fine mapping and gene cloning in the post-NGS era: advances and prospects. *Theor. Appl. Genet.* 133, 1791–1810. doi: 10.1007/s00122-020-03560-w

- Jiang, M., Ren, L., Lian, H., Liu, Y., and Chen, H. (2016). Novel insight into the mechanism underlying light-controlled anthocyanin accumulation in eggplant (*Solanum melongena* L.). *Plant Sci.* 249, 46–58. doi: 10.1016/j.plantsci.2016.04.001
- Jombart, T. (2008). ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129
- Kamvar, Z. N., Tabima, J. F., and Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2014, 1–14. doi: 10.7717/peerj.281
- Kaushik, P., Plazas, M., Prohens, J., Vilanova, S., and Gramazio, P. (2018). Diallel genetic analysis for multiple traits in eggplant and assessment of genetic distances for predicting hybrids performance. *PLoS One* 13:e0199943. doi: 10.1371/journal.pone.0199943
- Khan, M. M. R., Arita, T., Iwayoshi, M., Ogura-Tsujita, Y., and Isshiki, S. (2020). Development of the functional male sterile line of eggplant utilizing the cytoplasm of *Solanum kurzii* by way of the amphidiploid. *Environ. Control Biol.* 58, 79–83. doi: 10.2525/ecb.58.79
- Khan, M. M. R., Hasnunnahar, M., Iwayoshi, M., Ogura-Tsujita, Y., and Isshiki, S. (2015). Pollen degeneration in three functional male-sterile lines of eggplant with the wild *Solanum cytoplasm*s. *Horticult. Environ. Biotechnol.* 56, 350–357. doi: 10.1007/s13580-015-0015-3
- Kiferle, C., Fantini, E., Bassolino, L., Povero, G., Spelt, C., Buti, S., et al. (2015). Tomato R2R3-MYB proteins SLANT1 and SLAN2: same protein activity, different roles. *PLoS One* 10:e0136365. doi: 10.1371/journal.pone.0136365
- Knapp, S., Vorontsova, M. S., and Prohens, J. (2013). Wild relatives of the eggplant (*Solanum melongena* L.: Solanaceae): new understanding of species names in a complex group. *PLoS One* 8:e57039. doi: 10.1371/journal.pone.0057039
- Knaus, B. J., and Grünwald, N. J. (2017). VCFR: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* 17, 44–53. doi: 10.1111/1755-0998.12549
- Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., et al. (2009). A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* 5:e1000551. doi: 10.1371/journal.pgen.1000551
- Lebeau, A., Gouy, M., Daunay, M. C., Wicker, E., Chiroleu, F., Prior, P., et al. (2013). Genetic mapping of a major dominant gene for resistance to *Ralstonia solanacearum* in eggplant. *Theor. Appl. Genet.* 126, 143–158. doi: 10.1007/s00122-012-1969-5
- Lefebvre, V., Pflieger, S., Thabuis, A., Caranta, C., Blattes, A., Chauvet, J. C., et al. (2002). Towards the saturation of the pepper linkage map by alignment of three intraspecific maps including known-function genes. *Genome* 45, 839–854. doi: 10.1139/g02-053
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [Preprint]*. arXiv, 1303.3997.
- Li, J., He, Y. J., Zhou, L., Liu, Y., Jiang, M., Ren, L., et al. (2018). Transcriptome profiling of genes related to light-induced anthocyanin biosynthesis in eggplant (*Solanum melongena* L.) before purple color becomes evident. *BMC Genomics* 19:201. doi: 10.1186/s12864-018-4587-z
- Li, J., Ren, L., Gao, Z., Jiang, M., Liu, Y., Zhou, L., et al. (2017). Combined transcriptomic and proteomic analysis constructs a new model for light-induced anthocyanin biosynthesis in eggplant (*Solanum melongena* L.). *Plant Cell Environ.* 40, 3069–3087. doi: 10.1111/pce.13074
- Li, L., He, Y., Ge, H., Liu, Y., and Chen, H. (2021). Functional characterization of SmMYB86, a negative regulator of anthocyanin biosynthesis in eggplant (*Solanum melongena* L.). *Plant Sci.* 302:110696. doi: 10.1016/j.plantsci.2020.110696
- Liu, Y., Tikunov, Y., Schouten, R. E., Marcelis, L. F. M., Visser, R. G. F., and Bovy, A. (2018). Anthocyanin biosynthesis and degradation mechanisms in Solanaceous vegetables: a review. *Front. Chem.* 6:52. doi: 10.3389/fchem.2018.00052
- Lv, L. L., Feng, X. F., Li, W., and Li, K. (2019). High temperature reduces peel color in eggplant (*Solanum melongena*) as revealed by RNA-seq analysis. *Genome* 62, 503–512. doi: 10.1139/gen-2019-0021
- Maier, A., Schrader, A., Kokkelink, L., Falke, C., Welter, B., Iniesto, E., et al. (2013). Light and the E3 ubiquitin ligase COP1/SPA control the protein stability of the MYB transcription factors PAP1 and PAP2 involved in anthocyanin accumulation in *Arabidopsis*. *Plant J.* 74, 638–651. doi: 10.1111/tpj.12153
- Mangino, G., Plazas, M., Vilanova, S., Prohens, J., and Gramazio, P. (2020). Performance of a set of eggplant (*Solanum melongena*) lines with introgressions from its wild relative *S. incanum* under open field and greenhouse conditions and detection of QTLs. *Agronomy* 10:467. doi: 10.3390/agronomy10040467
- Mangino, G., Vilanova, S., Plazas, M., Prohens, J., and Gramazio, P. (2021). Fruit shape morphometric analysis and QTL detection in a set of eggplant introgression lines. *Sci. Horticult.* 282:110006. doi: 10.1016/j.scienta.2021.110006
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12. doi: 10.1089/cmb.2017.0096
- Mennella, G., Lo Scalzo, R., Fibiani, M., D'Alessandro, A., Francese, G., Toppino, L., et al. (2012). Chemical and bioactive quality traits during fruit ripening in eggplant (*S. melongena* L.) and allied species. *J. Agric. Food Chem.* 60, 11821–11831. doi: 10.1021/jf3037424
- Moglia, A., Francesco, E. F., Sergio, I., Alessandra, G., Milani, A. M., Cinzia, C., et al. (2020). Identification of a new R3 MYB type repressor and functional characterization of the members of the MBW transcriptional complex involved in anthocyanin biosynthesis in eggplant (*S. melongena* L.). *PLoS One* 15:e0232986. doi: 10.1371/journal.pone.0232986
- Naem, M., Muqarab, R., and Waseem, M. (2019). The *Solanum melongena* COP1 delays fruit ripening and influences ethylene signaling in tomato. *J. Plant Physiol.* 240:152997. doi: 10.1016/j.jplph.2019.152997
- Pascual, L., Desplat, N., Huang, B. E., Desgroux, A., Bruguier, L., Bouchet, J. P., et al. (2015). Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. *Plant Biotechnol. J.* 13, 565–577. doi: 10.1111/pbi.12282
- Pook, T., Schlather, M., De Los Campos, G., Mayer, M., Carolin Schoen, C., and Simianer, H. (2019). HaploBlocker: creation of subgroup-specific haplotype blocks and libraries. *Genetics* 212, 1045–1061. doi: 10.1534/genetics.119.302283
- Prohens, J., Gramazio, P., Plazas, M., Dempewolf, H., Kilian, B., Díez, M. J., et al. (2017). Introgressomics: a new approach for using crop wild relatives in breeding for adaptation to climate change. *Euphytica* 213:158.
- Prohens, J., Whitaker, B. D., Plazas, M., Vilanova, S., Hurtado, M., Blasco, M., et al. (2013). Genetic diversity in morphological characters and phenolic acids content resulting from an interspecific cross between eggplant, *Solanum melongena*, and its wild ancestor (*S. incanum*). *Ann. Appl. Biol.* 162, 242–257. doi: 10.1111/aab.12017
- Ramsay, N. A., and Glover, B. J. (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci.* 10, 63–70. doi: 10.1016/j.tplants.2004.12.011
- Ranil, R. H. G., Niran, H. M. L., Plazas, M., Fonseka, R. M., Fonseka, H. H., Vilanova, S., et al. (2015). Improving seed germination of the eggplant rootstock *Solanum torvum* by testing multiple factors using an orthogonal array design. *Sci. Horticult.* 193, 174–181.
- Robinson, J. T., Thorvaldsdóttir, H., Turner, D., and Mesirov, J. P. (2020). igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). *bioRxiv [Preprint]*. bioRxiv, 075499.
- Rockman, M. V., and Kruglyak, L. (2008). Breeding designs for recombinant inbred advanced intercross lines. *Genetics* 179, 1069–1078. doi: 10.1534/genetics.107.083873
- Scott, M. F., Ladejobi, O., Amer, S., Bentley, A. R., Biernaskie, J., Boden, S. A., et al. (2020). Multi-parent populations in crops: a toolbox integrating genomics and genetic mapping with breeding. *Heredity* 125, 396–416. doi: 10.1038/s41437-020-0336-6
- Shi, S., Liu, Y., He, Y., Li, L., Li, D., and Chen, H. (2021). R2R3-MYB transcription factor SmMYB75 promotes anthocyanin biosynthesis in eggplant (*Solanum melongena* L.). *Sci. Horticult.* 282:110020. doi: 10.1016/j.scienta.2021.110020
- Stadlmeier, M., Hartl, L., and Mohler, V. (2018). Usefulness of a multiparent advanced generation intercross population with a greatly reduced mating design for genetic studies in winter wheat. *Front. Plant Sci.* 9:1825. doi: 10.3389/fpls.2018.01825
- Thépot, S., Restoux, G., Goldringer, I., Hospital, F., Gouache, D., Mackay, I., et al. (2015). Efficiently tracking selection in a multiparental population: the case of earliness in wheat. *Genetics* 199, 609–623.
- Thissen, D., Steinberg, L., and Kuang, D. (2002). Quick and easy implementation of the Benjamini-Hochberg procedure for controlling the false positive rate in multiple comparisons. *J. Educ. Behav. Stat.* 27, 77–83. doi: 10.3102/10769986027001077

- Tigchelaar, E. C., Janick, J., and Erickson, H. T. (1968). The genetics of anthocyanin coloration in eggplant (*Solanum Melongena* L.). *Genetics* 60, 475–491. doi: 10.1093/genetics/60.3.475
- Toppino, L., Barchi, L., Lo Scalzo, R., Palazzolo, E., Francese, G., Fibiani, M., et al. (2016). Mapping quantitative trait loci affecting biochemical and morphological fruit properties in eggplant (*Solanum melongena* L.). *Front. Plant Sci.* 7:256. doi: 10.3389/fpls.2016.00256
- Toppino, L., Barchi, L., Mercati, F., Acciarri, N., Perrone, D., Martina, M., et al. (2020). A new intra-specific and high-resolution genetic map of eggplant based on a RIL population, and location of QTLs related to plant anthocyanin pigmentation and seed vigour. *Genes* 11, 1–29. doi: 10.3390/genes11070745
- Turner, S. D. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* 3:731. doi: 10.21105/joss.00731
- Valdar, W., Flint, J., and Mott, R. (2006). Simulating the collaborative cross: power of quantitative trait loci detection and mapping resolution in large sets of recombinant inbred strains of mice. *Genetics* 172, 1783–1797. doi: 10.1534/genetics.104.039313
- Van Eck, H. J., Jacobs, J. M. E., Van Den Berg, P. M. M., Stiekema, W. J., and Jacobsen, E. (1994). The inheritance of anthocyanin pigmentation in potato (*Solanum tuberosum* L.) and mapping of tuber skin colour loci using RFLPs. *Heredity* 73, 410–421. doi: 10.1038/hdy.1994.189
- Vilanova, S., Alonso, D., Gramazio, P., Plazas, M., García-Forte, E., Ferrante, P., et al. (2020). SILEX: a fast and inexpensive high-quality DNA extraction method suitable for multiple sequencing platforms and recalcitrant plant species. *Plant Methods* 16:110. doi: 10.1186/s13007-020-00652-y
- Wang, H., Guo, X., Pandey, M. K., Ji, X., Varshney, R. K., Nwosu, V., et al. (2017). “History and impact of the international peanut genome initiative: the exciting journey toward peanut whole-genome sequencing,” in *The Peanut Genome*, eds R. Varshney, M. Pandey, and N. Puppala (New York, NY: Springer.), 117–134. doi: 10.1007/978-3-319-63935-2_8
- Wei, Q., Wang, W., Hu, T., Hu, H., Wang, J., and Bao, C. (2020). Construction of a SNP-based genetic map using SLAF-Seq and QTL analysis of morphological traits in eggplant. *Front. Genet.* 11:178. doi: 10.3389/fgene.2020.00178
- Wickham, H. (2016). *ggplot2 Elegant Graphics for Data Analysis (Use R!)*. New York, NY: Springer.
- Xiao, X. O., Li, K., Feng, X. F., and Jin, H. (2018). Transcriptome analyses reveal anthocyanins biosynthesis in eggplant. *PeerJ* 6:e27289v1. doi: 10.1186/s12870-019-1960-2
- Yan, S., Chen, N., Huang, Z., Li, D., Zhi, J., Yu, B., et al. (2020). Anthocyanin Fruit encodes an R2R3-MYB transcription factor, SLAN2-like, activating the transcription of SIMYBATV to fine-tune anthocyanin content in tomato fruit. *New Phytol.* 225, 2048–2063. doi: 10.1111/nph.16272
- Zhang, Y., Hu, Z., Chu, G., Huang, C., Tian, S., Zhao, Z., et al. (2014). Anthocyanin accumulation and molecular analysis of anthocyanin biosynthesis-associated genes in eggplant (*Solanum melongena* L.). *J. Agric. Food Chem.* 62, 2906–2912. doi: 10.1021/jf404574c
- Zhou, L., He, Y., Li, J., Liu, Y., and Chen, H. (2019). CBFs function in anthocyanin biosynthesis by interacting with MYB113 in eggplant (*Solanum melongena* L.). *Plant Cell Physiol.* 61, 416–426. doi: 10.1093/pcp/pcz209

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Mangino, Arrones, Plazas, Pook, Prohens, Gramazio and Vilanova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.