BMC Plant Biology

RESEARCH Open Access

SABER: a multiparental tomato population leveraging wild relative diversity for high-resolution QTL mapping

Giovanni Gabelli[1], Leonardo Caproni[2], Fabio Palumbo[1], Anna Giulia Boni[3], Giulia Ferrari[3], Maria Lucia Prazzoli[3], Marina Malatrasi[3], Sara Sestili[4], Matteo dell'Acqua[2], Massimiliano Beretta[3] and Gianni Barcaccia[1*]

Abstract

Background The narrow genetic base of cultivated tomato (*Solanum lycopersicum* L.) represents a major constraint on crop improvement, necessitating the exploitation of wild relatives to broaden allelic diversity. Here we present SABER (*Solanum lycopersicum* Allele Biodiversity Enriched Resources), a novel eight-founder Multiparent Advanced Generation Intercross (MAGIC) population that, for the first time, incorporates the Galápagos wild relative *Solanum cheesmaniae* as a founder alongside seven elite *S. lycopersicum* lines.

Results Following a structured crossing scheme and Single Seed Descent (SSD) advancement, F6 recombinant inbred lines were genotyped at 5,850 high-confidence SNP markers using Single Primer Enrichment Technology (SPET). Population structure analyses confirmed low residual heterozygosity, limited substructure among offspring, and successful introgression of *S. cheesmaniae* alleles across all twelve chromosomes. Mapping performance was validated through three Mendelian traits with known genetic determinants, all of which resolved to genomic positions consistent with the literature. Quantitative Trait Loci (QTL) mapping identified known loci for fruit epicarp and flesh color, and two novel QTLs for days to flowering, number of leaves before flowering, and soluble solids content.

Conclusions Together, these results demonstrate that SABER is a powerful and reliable platform for high-resolution QTL mapping and candidate gene discovery, and establish a replicable framework for integrating wild germplasm into multiparental tomato breeding resources.

Keywords Tomato, *Solanum*, *Lycopersycum*, *Cheesmaniae*, MAGIC, QTL, SPET

*Correspondence:
Gianni Barcaccia
gianni.barcaccia@unipd.it
[1]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Legnaro, PD, Italy
[2]Institute of Plant Sciences, Scuola Superiore Sant'Anna, Pisa, Italy
[3]Panora SpA, Fidenza, PR, Italy
[4]Council for Agricultural Research and Economics (CREA) - Research Centre for Vegetable and Ornamental Crops, Monsampolo del Tronto, AP, Italy

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/.

Background

Tomato (*Solanum lycopersicum L.*) stands as one of the most economically significant vegetable crops globally, with annual production figures exceeding 182 million tons [1]. Beyond its commercial dominance, the tomato is the premier model species for studying fleshy fruit development due to its well-documented physiology and the vast array of genetic and genomic tools available to researchers [2, 3]. Despite this status, the genetic diversity of modern cultivars is remarkably narrow, having undergone severe bottlenecks during domestication and subsequent intensive breeding [3–5]. Historically, restricted genetic variation has hindered crop improvement, driving the need to explore wild relatives as essential reservoirs for diversifying the *S. lycopersicum* genome. While previous decades focused on utilizing wild species as sources of resilience through single-trait introgressions [1], the paradigm has recently shifted. Current research prioritizes deciphering the complex regulatory networks underlying agronomic traits and mapping the vast allelic diversity within wild populations. The recent production of pangenomes and super-pangenomes comprising wild and cultivated tomatoes exemplifies this transition toward a more holistic genomic approach [5–10].

Nevertheless, a thorough investigation into the actual agronomic impact of wild traits and alleles is still necessary. Consequently, the development of mapping populations remains an essential tool for identifying and validating valuable traits in the field. While traditional mapping approaches like biparental crosses or Genome-Wide Association Studies (GWAS) have identified major genes, they are often limited by low mapping resolution or population substructure [2]. Multiparent Advanced Generation Intercross (MAGIC) populations have emerged as powerful advanced resources to address both constraints, providing increased genetic diversity and high recombination rates that facilitate high-precision QTL mapping and candidate gene identification [3, 11, 12]. To date, tomato MAGIC resources have included elite lines and cherry types belonging to *S. lycopersicum* [2] or generated with interspecific crosses including the well-known *S. pimpinellifolium* and *lycopersicum* var. *cerasiforme* [3]. These studies facilitated the identification of valuable novel QTLs. It has been demonstrated how an interspecific MAGIC population enabled the identification of key agronomic traits, such as fruit weight, locule number, anthocyanin content, leaf lobing, and the number of leaves below the first inflorescence [3]. The extensive genetic diversity observed among wild tomato species [8, 10] highlights the potential of similar studies utilizing other wild relatives.

The newly developed MAGIC population, called SABER (*Solanum lycopersicum* Allele Biodiversity Enriched Resources), represents a significant innovation in tomato genetics by specifically targeting the untapped potential of the wild relative *Solanum cheesmaniae*. This species is regarded as a valuable source of exotic allelic diversity and has been associated with tolerance to biotic and abiotic stresses, making it potentially useful for broadening the genetic base of cultivated tomato and improving plant resilience [3, 13]. The SABER design introgresses *S. cheesmaniae* as a founder together with seven *S. lycopersicum* elite lines, developed for commercial, breeding and research purposes, placing this unique allelic diversity within a multi-parental framework aimed at producing innovative materials. The SABER population offers a platform for high-precision mapping of complex quantitative traits of agronomic relevance like flowering earliness and sugar and organic acid levels in the fruit (°Brix), untapping allelic diversity that was previously inaccessible. Indeed, SABER also serves as a tool for integrating exotic diversity into modern tomato breeding pipelines, potentially contributing to cultivars better suited to challenges posed by the changing climate.

Methods

Plant material and founder lines

The SABER was developed using an eight-way design. The parental lines were selected to maximize genetic and phenotypic diversity. Seven founders belong to *S. lycopersicum* and include both elite proprietary breeding lines of Panora SpA (formerly ISI Sementi) as well as lines derived from publicly available accessions, while one founder derives from wild relative (*Solanum cheesmaniae)*.

The seven *S. lycopersicum* lines are PAR209, PAR273, PAR480, FLINE, EB5, PAR166, LA2460, and the *S. cheesmaniae* line is called LA1402. A detailed description of the pedigree and the fetaure of the eight founder lines is in Table S1 of Additional file 1.

Crossing scheme and population development

The SABER MAGIC population was developed following a structured eight-founder cross design aimed at maximizing recombination among parental genomes (Fig. 1).

The 8 founder lines were initially paired to generate 4 independent two-way F_1 hybrids. All crosses involving the wild accession *Solanum cheesmaniae* LA1402, were consistently carried out using it as the pollen donor (male parent) to minimize potential fertilization barriers and improve hybrid seed set efficiency. The resulting hybrids (G1) were subsequently intercrossed as follows: the pollen of two groups of G1 was pooled and used to pollinate the individuals of the plants of the other two G1 groups, resulting in two independent four-way hybrid pools (G2). To generate the final eight-way population, 100 individual plants from each of the two four-way pools were randomly selected and grown. The same plants were crossed

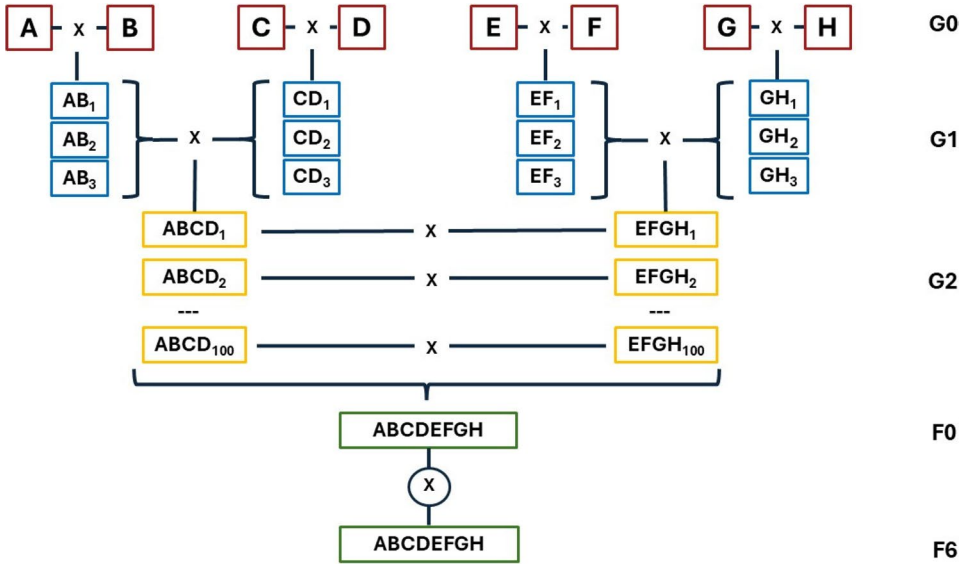


Fig. 1 Crossing scheme of the SABER population. The 'X' symbol between two samples indicates a cross, the circled 'x' (⊗) indicates self-fertilization, and braces group genotypes where pollen was pooled for the subsequent crossing step

in a pairwise (i.e. one-to-one) scheme between pools, resulting in 100 independent eight-way crosses.

Seeds derived from each of the 100 independent eight-way crosses were sown, and 10 plants per cross were planted to the field, resulting in 1,000 plants representing the F0 generation of the MAGIC population. Within each progeny group (cross), 5 plants were randomly selected without phenotypic selection to maintain genetic diversity and avoid early bottleneck effects.

Selected plants were advanced through successive generations using a Single Seed Descent (SSD) approach. At each generation, one seed per plant was harvested and used to establish the subsequent generation, without intentional phenotypic selection, to progressively increase homozygosity while preserving the recombination structure established during the early intercross phases. However, in cases where a progeny row displayed pronounced phenotypic variability suggestive of residual segregation, a modified SSD procedure was applied. Specifically, two individual plants from the same progeny row were independently advanced, thereby generating parallel sub-lines ("forks"). This strategy allowed the preservation of segregating genetic variation that might otherwise have been lost under strict single-line advancement, while maintaining the overall SSD framework.

Plant growth conditions and phenotyping material

Plants were grown under open-field conditions at Panora experimental facilities (Northern Italy, lat.44.8485154, long.10.0295428), following standard agronomic practices for determinate tomato cultivation typical of the Po Valley production system. Seeds were sown in 260-cell trays filled with sterile peat-based substrate and maintained under nursery conditions until seedlings reached the 3°–4° true leaf stage. Planting to open field was carried out approximately 40–45 days after sowing. Field planting was performed mechanically, with an in-row plant spacing of 25 cm and an inter-row distance of 140 cm. Consecutive experimental plots were separated by 100 cm to facilitate management operations. Irrigation was provided with drip lines using well water. Fertigation was scheduled according to crop developmental stage and agronomic requirements. Standard crop protection practices were applied, including calendar-based fungicide treatments to prevent major fungal diseases, in accordance with integrated pest management guidelines for the region.

Phenotyping

The phenotyping of qualitative traits (Table S2 of Additional file 1) was carried out by observation. Obscuravenosa was assessed by observing 3 leaflets per plant, 3 plants/plot against solar light to ensure consistency in scoring. The phenotypes of the fruits (green shoulder, epicarp color, flesh color), were phenotyped by observing 3 fruits present in the first truss of each plant, 3 plants per plot to increase the reliability of the scoring. The hypocotyl color of each accession was classified according to 3 categories: "green" (absence of anthocyanins), "purple" (presence of anthocyanins), or "segregating", observing 13 plants in the nursery trays. Days to Flowering (DTF) and leaves number before flowering (LN) were determined by observing the opening date of the first inflorescence for each plant and counting their leaves. The measurements of Soluble Solids Content (°Brix) were performed using a digital refractometer

(Atago PAL-1). For each accession, the 5 fully mature fruits were weighted and then their juice was squeezed and the average value was calculated. A total of 240 samples were phenotyped for the illustrated traits.

The qualitative phenotype data were formatted for the QTL mapping as following: green shoulder, obscuravenosa and hypocotyl color were transformed in 0/1 character to be suitable for a linear model, with segregating phenotypes assigned to NA; epicarp and flesh fruit color in 0/1/2 (yellow / orange / red) to account for an additive model, or 0/1 (yellow – orange / red). The two transformed data sets gave identical results.

Genotyping of parent lines and SPET probes definition

Seedlings were grown at CREA-OF (Monsampolo del Tronto, Italy) in plateau trays filled with a sterilized peat-based soil mixture, under controlled environmental conditions (25 ± 2 °C; 80–90% relative humidity). At the 4–5 true leaves stage, young leaf tissue was sampled from each plant and stored at -80 °C until processing. Genomic DNA was isolated using the Magnetic Bead System of the Norgen's Plant DNA Isolation Kit (Norgen Biotek Corp, Thorold, ON, Canada) following the manufacturer's protocol. DNA quantity and quality were verified at 260/280 and 260/230, with the NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). Genomic libraries were prepared using the Celero™ DNA-Seq Library Preparation Kit (Tecan Genomics, Redwood City, CA, USA) according to the manufacturer's protocol. DNA input material and the resulting libraries were quantified using a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). Library quality and fragment size distribution were assessed with the Agilent 2100 Bioanalyzer employing the High Sensitivity DNA assay (Agilent Technologies, Santa Clara, CA, USA). Sequencing-ready libraries were subsequently pooled and sequenced on an Illumina NovaSeq 6000 platform in paired-end mode with 150 bp read length. Raw sequencing data were processed for base calling, format conversion, and sample demultiplexing using Bcl Convert v3.9.3 (Illumina). Demultiplexed reads were aligned to the tomato reference genome (SL4.0) using BWA-MEM v0.7.17-r1188 [14], implemented in Clara Parabricks v4.4.0 with the parameters --optical-duplicate-pixel-distance 2500 and --K 1,000,000.

Variant discovery was performed using GATK v4.1.0.0 [15, 16], following the GenomicsDBImport, ReblockGVCF, and GenotypeGVCFs workflow to identify SNPs and small insertions/deletions (indels). To minimize false positives, variants overlapping annotated repetitive regions (ITAG4.0 RepeatModeler annotation [17] were removed using bedtools subtract v2.30.0 [18]. Subsequent quality control was performed using bcftools v1.13 and GATK [16, 19]. After filtering out indels, initial hard filtering for SNPs was applied using the criteria: QD < 2.0 || MQ < 40.0 || FS > 60.0 || SOR > 3.0 || MQRankSum < – 12.5 || ReadPosRankSum < – 8.0, Total site depth INFO/DP < 50. At the genotype level, high-confidence alternative homozygous sites were identified requiring a minimum allele depth ≥ 5, while maintaining sites with ≤ 2 missing calls and ≥ 5 reference or missing genotypes. For heterozygous sites, allelic balance was estimated by calculating variant allele frequencies at positions with a minimum coverage ≥ 10x. Finally, selected variants were converted to BED format, merged across samples using bedtools merge v2.30.0 [18] (maximum distance 20 bp), and filtered based on inter-marker distance thresholds to define the final set of 10,413 high-confidence genomic regions.

F6 samples genotyping

The DNA extracted from the samples was quantified using the Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA). Libraries were prepared using the "Allegro Targeted Genotyping" protocol from NuGEN Technologies (San Carlos, CA), using 2.5 ng/μL of DNA as input and following the manufacturer's instructions. Libraries were quantified using the Qubit 2.0 Fluorometer, and their size was checked using the High Sensitivity DNA assay from Bioanalyzer (Agilent technologies, Santa Clara, CA) or the High Sensitivity DNA assay from Caliper LabChip GX (Caliper Life Sciences, Alameda CA). Libraries were quantified through qPCR using the CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA) and run on the Illumina NovaSeq 6000 (Illumina, San Carlos, CA). Base calling and demultiplexing were performed with Illumina bcl2fastq v2.20. The reads of the F6 samples were processed for adapter trimming and quality filtering with fastp v0.20.1 [20], using default setting, and the resulting files checked with multiqc v1.12 [21].

The resulting reads were then aligned against the fourth release of the *Solanum lycopersicum* (SL4.0 [17]) reference genome produced by the International Tomato Genome Sequencing Project using bwa-mem2 [14] with default setting.

The SAM files of all samples were filtered for alignment quality (-q 5), sorted, and indexed with the samtools suite v1.13 [19]. A quality check was performed with Qualimap bamqc v2.3 [22]. Variant calling for all samples was performed with the GATK suite v4.1.0.0 [16] in three steps: first, variants were called for each sample using HaplotypeCaller in GVCF mode, setting the minimum base quality score to 5; second, the individual GVCF files were combined with CombineGVCFs; finally, variants were jointly genotyped for all samples using GenotypeGVCFs. Due to the size of the data, the last step (GenotypeGVCFs) was performed one chromosome at a

time, and all chromosomes were merged at the end of the process with bcftools concat.

The vcf file with the variants of the founder lines was filtered with bcftools v1.13 [19] to retain only the SNPs that were homozygous in all samples, missing in none, and biallelic across the eight lines. For the F6 lines, the VCF files were filtered using custom awk commands and bcftools and vcftools v0.1.16 programs [19, 23]. We retained only variants satisfying all of the following criteria: being a SNP (indels were excluded), having a mean Phred score of the conditional genotype quality across samples (GQ field in the VCF) higher than 30, having a mean of the read depth across the samples (DP fields) higher than 30, being located on an assembled chromosome (excluding SNPs on chromosome SL4.0ch00, which contains unassembled contigs), being included in the filtered variant list of the eight parent lines, having less than the 20% of the genotypes heterozygous and less of the 20% of the genotypes missing. We then identified samples with >20% missing genotypes or >10% heterozygous genotypes and removed them from the analysis. Finally, the list of variants of the parent lines absent from the list of the F6 lines were removed and vice versa, so that the two lists contain the same variants.

LD decay estimates and sample clustering

The filtered variant file was used in HapMap obtained by the transformation of the vcf file with plink v2.00a3 [24] to estimate the LD decay along all the genome, with a custom R script derived from Caproni et al. [25]. In brief, pairwise LD (expressed as r^2) was calculated for each chromosome independently using the *LDheatmap* package [26]. LD decay as function of physical distance was estimated for each chromosome by fitting a nonlinear model based on the Hill and Weir equation. The expected r^2 values were computed using nonlinear least squares, assuming a population size of 254 individuals. For each chromosomethe LD half-decay distance was calculated using an overall r^2 value corresponding to the maximum interpolated half-decay r^2 (i.e. 0.235).

Genetic structure and diversity within the SABER were assessed using a set of SNP markers pruned for linkage disequilibrium (LD) at an r^2 threshold of 0.3. LD pruning was performed in plink2 using the *indep-pairwise* function with a 150-variant window and a step size of 15 variants. A version of this script is available in Macharia et al. [27]

The pruned dataset was used to estimate genetic pairwise distances among RILs using Euclidean distance with the *adegenet* package [28]. Genetic diversity, was also represented in a Neighbor-Joining (NJ) tree was constructed using the *ape* package [29]. To further investigate genetic structure, a Principal Component Analysis (PCA) was performed using the *SNPRelate* package [30].

Linkage map elaboration and synteny analysis

The genetic map of the SABER was obtained by anchoring the tomato MAGIC genetic positions derived from Pascual et al. [2] to the full SNP dataset on the tomato reference genome V4 (SL4.0). Missing genetic distances were subsequently interpolated using values proportional to the physical distances. In brief, the tomato MAGIC genetic map was lifted over from SL2.4 to SL4.0. The liftover was performed with CrossMap [31], which requires as input the map to be lifted over and a chain file containing the coordinate correspondences. Since a chain file linking SL2.4 and SL4.0 was not publicly available, we generated one using the Python pipeline pyOverChain [32], providing the FASTA files of both reference genome versions, downloaded from the International Tomato Genome Sequencing Project FTP repository [33]. Twenty-one markers showing positional inconsistencies between physical and genetic distances were removed. Using the curated linkage map, we derived genetic positions on the SABER based on marker physical position and interpolated the missing genetic distances proportional to marker physical distances, following a script deposited by Ferguson et al. [12]. The resulting linkage map with the physical and genetic positions, as well as the chain file, are provided in the supplemental material (Table S3 in Additional file 1, Additional file 2).

Genotype probabilities computation and QTL mapping

The QTL mapping was entirely carried out using the R/qtl2 package v0.38 [34]. Genome reconstruction of the RILs were estimated using a Hidden Markov Model (HMM) using the function *calc_genoprob()*, using the previously described genetic map and accounting for a potential genotyping error probability of 0.002. To account for population structure and individual relatedness within the RILs, a kinship matrix was computed with the *calc_kinship* function, using the Leave One Chromosome Out (LOCO) method.

QTL mapping was performed using a linear mixed model (LMM) via the *scan1* function. The Logarithm of Odds (LOD) scores were calculated across the genome for all traits. Significance thresholds were determined through permutation testing (10,000 permutation) to control the genome-wide false discovery rate. Significance levels were set at alpha = 0.05. Significant QTL peaks were identified using the *find_peaks* function. For each identified peak, the 95% Bayesian credible interval was calculated to define the QTL boundaries, and the intervals were expanded to the nearest flanking markers. In the case of BRIX QTL, a second analysis was performed adding the weight of the fruits as covariate in the model (*addcovar* variable in the *scan1* function).

Results and discussion

Genotyping of the population

The design of the sequencing loci using SPET (Single Primer Enrichment Technology) [35] resulted in 10,413 primer-targeted sites for genotyping 365 F6 samples. Sequencing and alignment metrics for these samples are reported in Table S4 of Additional file 1. After quality filtering, 5,850 SNPs and 245 samples were retained for downstream analyses. The retained samples showed low residual heterozygosity (mean: 2.47%) and low levels of missing data (mean: 2.69%). Markers exhibited a mean density of 7.53 markers per Mbp and were distributed across the entire genome without major gaps (Fig. 2), including centromeric regions, which in *S. lycopersicum* are characterized by low gene density and low recombination rates [2]. As shown in Fig. 2, specific regions of increased marker density were observed, notably a sms2all segment on chromosome 2 (approx. 8–10 Mbp) and two broader regions on chromosomes 6 and 9 (spanning 1–29 Mbp and 10–60 Mbp, respectively). To investigate the origin of this variability, we examined the distribution of SNPs with alternative genotypes among the eight founder lines (Additional file 3). Across most chromosomes, the wild founder unsurprisingly exhibited the highest level of polymorphism relative to the reference genome. However, in the high-density regions on chromosomes 6 and 9, additional variability originated from other founders: PAR480 and PAR166 on chromosome 6, and LA2460 on chromosome 9. In both regions, the high number of private SNPs in these cultivars resulted in a scarcity of loci capable of distinguishing between the other lines. In other words, the set of variants has reduced discriminatory power between certain parental donors because the majority of SNPs share the same genotype across multiple lines (e.g., PAR206, PAR273, FLINE, E5, and LA2460 in the high-variant region of chromosome 6; PAR206, PAR273, PAR166, and EB5 in the central region of chromosome 9). A similar trend was partially observed in the centromeric region of chromosome 4, where two groups of founder lines showed high similarity: (PAR273, PAR166, and LA2460; PAR480, FLINE, and EB5). These trends may have a technical explanation: when one sample is genetically much more distant than the others, it can "capture" the majority of the polymorphisms during the SPET probe design process, thereby diminishing the observable differences between the remaining samples. However, the biological characteristics of cultivated tomatoes could also contribute; the genetic variability of this highly bred species is low [5], and it is common to observe conserved genomic regions across different varieties due to the selection of specific favorable traits or unknowingly shared ancestors. The high diversity spot on chromosome 2 appears to be a different phenomenon, since almost all the alternative alleles that compose the peak are private SNPs of LA1402 (*S. cheesmaniae*).

Population structure

Principal component analysis (PCA) performed on the LD-pruned SNP data from all lines, founder and offsprings (515 sites) confirmed the high genetic divergence of the wild relative compared to all other samples (Fig. 3a). This pattern is particularly evident along PC1, which explains 3.03% of the variance and captures most of the diversity associated with *S. cheesmaniae*. A similar effect is observed along PC2 (2.64% of the variance),



Fig. 2 SNP density across *Solanum lycopersicum* genome. SNP density plot chromosome wise representing number of SNPs within 1 Mb window size. The horizontal axis shows the chromosome length (Mb); the different color depicts SNP density

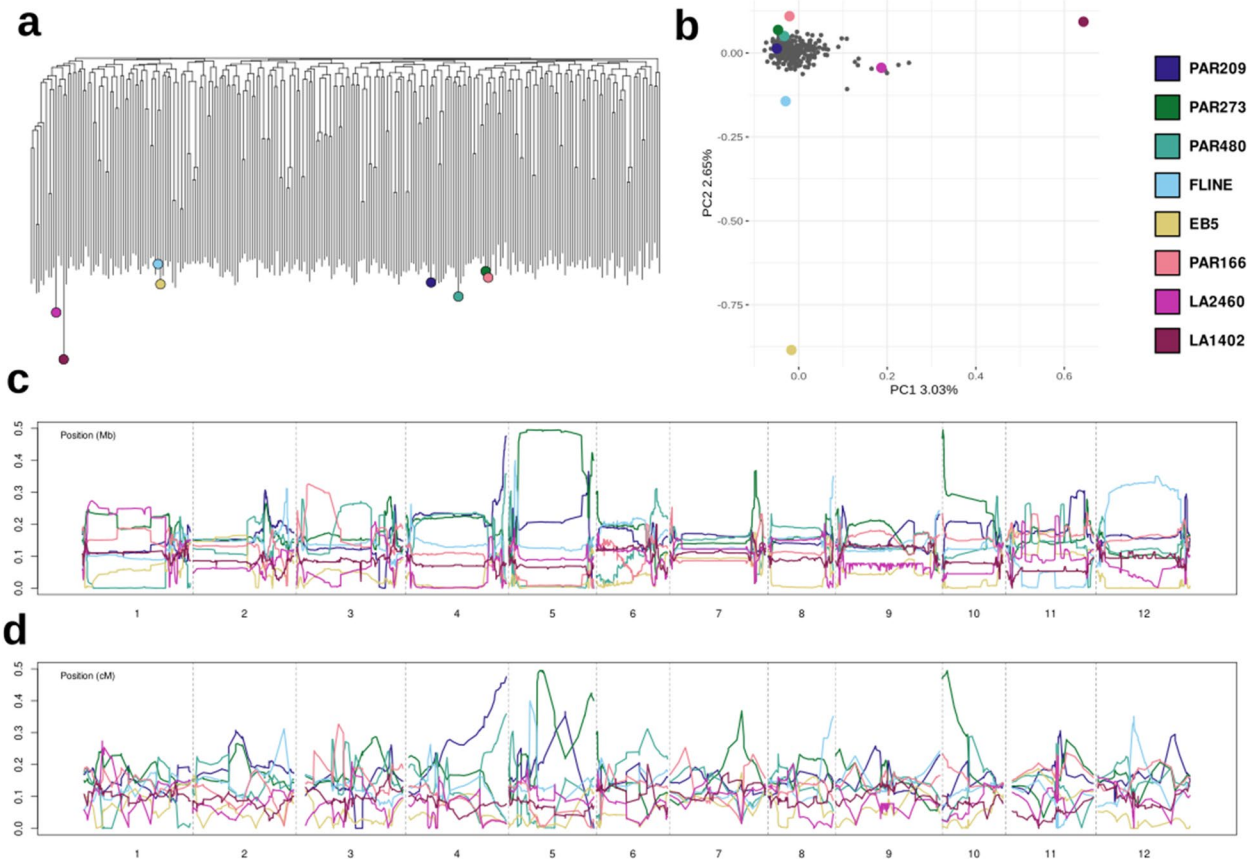


Fig. 3 Population structure analyses of SABER population. The legend at top right corner represents the color assigned to the eight founfer lines in all panels. **a** Hierarchal clustering of the SABER population. The colored dots represent the eight founder lines, **b**) First two principal components of a PCA of the SABER population, colored dots represent the eight founder lines **c**-**d**) Genome-wide founder haplotype nlocks assignment across the 12 chromosomes according to physical (**c**) or genetic position (**d**) of the markers

where EB5 emerges as an outlier, likely due to introgressions from *S. lycopersicum* var. *cerasiforme*. Notably, the loci differentiating LA1402 and EB5 from the other founders appear to be largely mutually exclusive. The remaining founders cluster closely together: LA2460 is genetically closest to LA1402, and FLINE to EB5, while PAR166 occupies the opposite end of the spectrum. PAR209, PAR273, and PAR480 cluster near the center of the PCA space. Offspring samples show tight clustering, indicating low population structure, with only 12 samples deviating towards LA1402 and LA2460, likely reflecting a higher genomic contribution of these founders.

Hierarchical clustering (Fig. 3b) largely corroborates the PCA results, with long branches for all the samples indicating low overall population structure. Nevertheless, some pairs of samples show higher genetic similarity. This can be explained by the breeding scheme of the SABER population: the 500 RILs were selected among the 100 crosses between G3 plants (i.e. the 8-way crosses), therefore inevitably some of the remaining RILs could be descendant from full siblings.

The hierarchical clustering further confirms the wild founder LA1402 as the most genetically distinct sample, as indicated by its long branch length While LA1402 connects adjacent to the main RIL cluster rather than forming a completely isolated outgroup, this topology is a direct consequence of the MAGIC population structure, wherein each RIL retains a significant genetic contribution from this wild founder. PAR166 and PAR273 cluster together, consistent with their breeding history, as PAR273 served as the pollen donor in the development of PAR166. FLINE and EB5 also have a high degree of similarity in this clustering. It is unknown if these lines, both developed by public entities, share some ancestors.

The estimated average contribution of each founder to the overall population (Fig. 3c, d) confirms that regional imbalance is still present, but overall, the mean contribution of each parent is around the theoretical 12.5%. A noticeable exception is represented by founder EB5, that contributes to the genome of the overall population less than expected in all chromosomes but chromosome 3 and chromosome 11 (Table 1). The difficulties in identifying the correct donor parent in some regions due to

Table 1 Mean contribution of each founder lines to the population genome for each chromosome. The measure is an average of the probability assigned to each marker in each sample to descend from the parent line

Parent lines	PAR209	PAR273	PAR480	FLINE	EB5	PAR166	LA2460	LA1402
Chr1	12.8	19.5	5.9	11.4	6.6	15.6	17.5	10.8
Chr2	17.3	16.0	13.2	15.5	10.8	12.2	6.9	8.0
Chr3	12.6	16.5	19.3	12.9	4.3	19.1	6.0	9.3
Chr4	23.7	20.9	21.6	13.3	1.5	10.1	1.6	7.3
Chr5	20.8	38.9	3.6	14.6	1.7	3.4	10.3	6.8
Chr6	17.4	19.2	8.0	19.2	2.1	10.8	11.3	12.1
Chr7	15.9	15.1	14.8	11.9	9.3	9.7	12.1	11.3
Chr8	14.1	16.1	17.5	15.6	3.1	11.6	13.1	8.9
Chr9	15.0	17.6	13.2	12.9	5.4	15.7	7.4	12.8
Chr10	18.7	26.3	12.3	12.4	1.6	15.1	5.6	7.9
Chr11	20.0	15.4	15.0	3.2	12.6	17.9	10.9	5.1
Chr12	16.1	11.4	13.1	27.1	0.7	15.6	6.2	9.8
average	17.0	19.4	13.1	14.2	5.0	13.1	9.1	9.2

shared genotype could help explain smaller imbalances, since they are regional phenomenon, mainly located in the large centromeric region, where few SNPs are genotyped. An example of an imbalance likely produced by technical reason is in chromosome 5 the overperformance of PAR273 and the underperformance of PAR480 and LA2460, where all three share the genotypes in the centromere (Additional file 3, Fig. 3c). The low level of alleles of EB5 in the population, however, appears to have more structural causes. Since no selection has been carried out during the crossing and SSD stages, a random fluctuation in the first crosses could have happened. The *Solanum cheesmaniae* line, LA1402, is present in the population with levels lower than the mean, but stable across all chromosomes. This is an important result, since the interspecific hybrid could have a lower vitality. Instead, we register no regions where the contribution of the wild ancestor to the population genome is missing.

The profiles of linkage disequilibrium (LD) levels along the chromosomes (Fig. 4a) in most of the chromosomes follow the pattern of recombination frequency observed in the genetic map, with large centromeric regions with low recombination and distal region where the majority of genes are located [2, 3]. Chromosomes 2 and 4 have different profiles from the others, since they are acrocentric chromosomes. Chromosomes 6, 9 and, at lesser extent, 4, exhibit high LD levels in the regions already described in this paragraph with reduced discriminating power. Indeed, this is expected since if the genotypes of two or more founder lines in a region are the same or not distinguishable, we will not detect all the recombination events that involve those lines. Considering this, and the Considering this the differences among the LD-decay levels computed along the chromosomes (Fig. 4b) are not surprising, with chromosomes 4, 6, and 9 having the highest levels.

Test of mapping performances of SABER

Since the total number of SABER samples with phenotype and genotype data available is of 240, lower than the optimum recommended for similar mapping populations [36, 37], three Mendelian traits with well-known genetic determinants in the literature (obscuravenosa, green shoulder, and hypocotyl color) were used to test its mapping potential. The comparison between the QTLs mapped in the SABER population and those previously reported in the literature is summarized in Table 2.

In the majority of fresh-market indeterminate tomatoes, chloroplasts are not present in the subepidermal and mesophyll tissues adjacent to the leaf veins, causing the veins to appear translucent under transmitted light. In wild phenotypes and some tomato line breed for industrial processing, however, those cells contain chloroplast and veins appear of the same color for the leaf if not darker, a phenotype called obscuravenosa (*obv*) [41]. In the SABER population, a single QTL was found for this character on chromosome 5, at 62.5 Mb (LOD peak 25.1), a result that corresponds to the known genetic determinant of this character found first by Jones et al. [41], that on SL4.0 assembly is 823 Kb distant [38].

The "green shoulder" trait in tomato is a qualitative character observed during the pre-veraison stage of fruit development. In wild-type genotypes, a distinct area of intense green pigmentation develops in the pericarp tissue surrounding the peduncular insertion. The absence of this pigmentation is governed by the *uniform ripening* (*u*) mutation. While the *u* phenotype is favored in modern cultivars to ensure synchronized maturation and aesthetic uniformity for consumers, it carries significant metabolic trade-offs [39]. The QTL identified in SABER corresponds with the known location of this mutation (chromosome 10, 2.0 Mb, LOD peak 8.9, 159 Kb of distance from the gene).

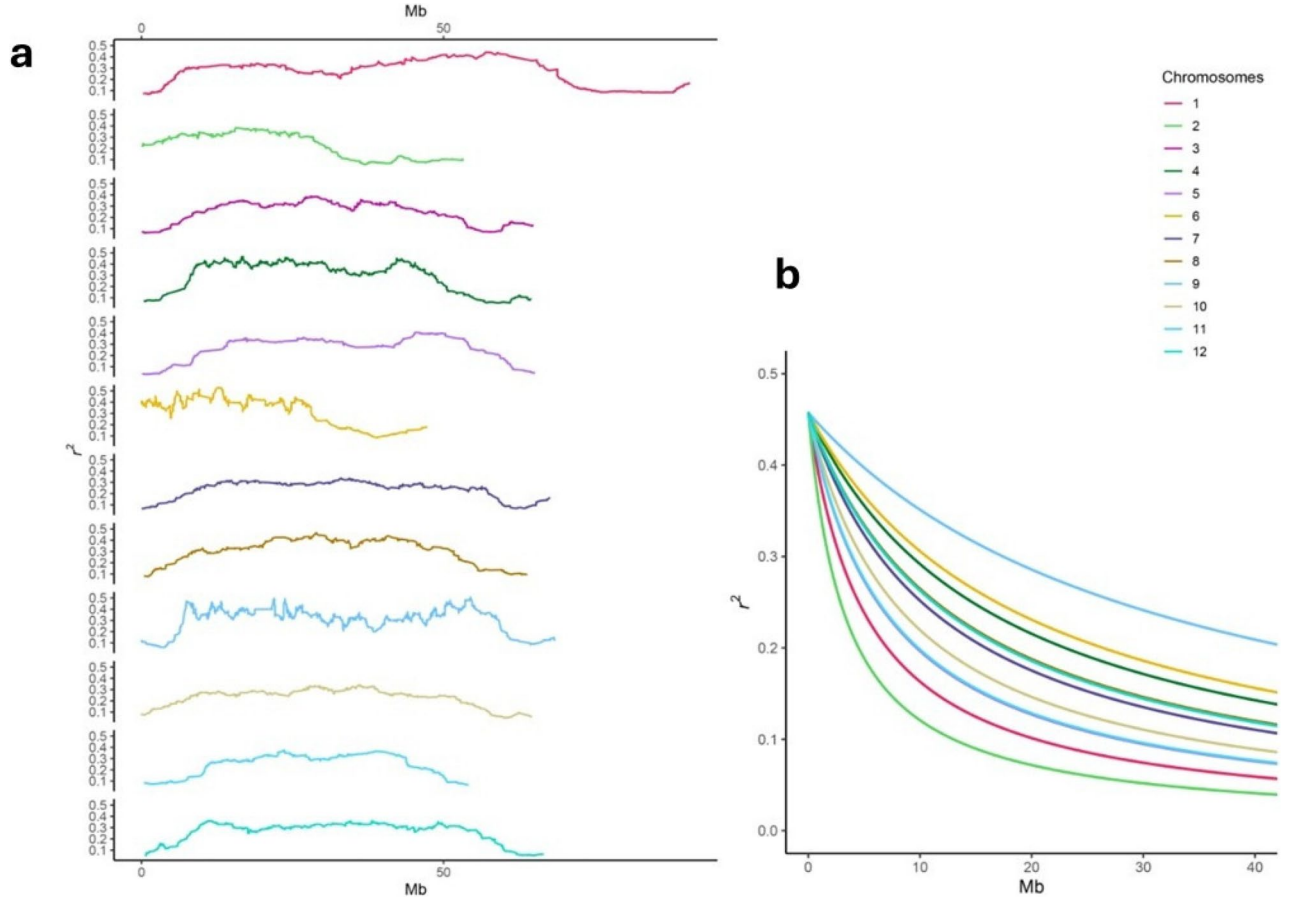


Fig. 4 LD analyses on SABER population. **a** LD levels in the twelve chromosomes computed across sliding windows, **b** LD decay in the twelve chromosomes

Table 2 Comparison between mendelian QTLs and literature data about the genetic determinant of the mendelian traits. The distance between QTL peaks and the gene is calculated considering the nearest boundary

Trait	Candidate gene (SL 4.0)				SABER QTL				
	Name	Chr	Coordinates (bp)	Reference	Phenotype	Chr	coordinates (bp)	LOD	Distance (bp)
obv/obv+	*Solyc05g054030*	Chr5	633,95,462–633,98,588	Lu et al. [38]	Obscuravenosa	Chr5	62,575,099	25.1	823,489
u/ug	*Solyc10g008160*	Chr10	2,169,302-2,174,039	Powell et al. [39]	Green shoulder	Chr10	2,010,429	8.9	158,873
AH/ah	*Solyc09g065100*	Chr9	59,048,134 – 59,055,605	Qiu et al. [40]	Hypocotyl color	Chr9	59,052,650	29.8	0

The third Mendelian trait used to test the model is the color of the hypocotyl, that in tomato is often purple due to the accumulation of anthocyanin pigments [40]. Multiple mutants with low pigmentation of the hypocotyl are known, and this character is used in same cases used as morphological marker [42]. The QTL identified in the SABER population (LOD score peak of 29.8) is positioned at 59.0 Mb in the chromosome 9, that is inside the gene coding for a bHLH transcription factor identified as the most common responsible for absence of pigmentation in the hypocotyl [40] .

Overall, we can say that all three QTL relative to mendelian traits confirm the potential of SABER as mapping population, both in accuracy and power.

Color of the fruit

The color of the tomato fruit at ripeness is a well-studied trait not only to address market preferences, since the bright red of the fruit or tomato-based product is one of the first traits that guide consumer choice [43], but also to unveil the complex metabolic pathways that regulate pigment accumulation in the tomato fruit. The red color of tomatoes is mainly due to carotenoids (mainly lycopene) and naringenin chalcone, which accumulate in the ripening of fruits while the chlorophyll is degraded [44]. The metabolic pathway of carotenoids has been extensively studied, as well as the genes that regulate it. *PSY1* emerged as one of the most important among them, catalyzing the first step committed in carotenoid biosynthesis. Mutants with no transcription of *PSY1*

have yellow flesh [44, 45]. Orange fruits are observed in a number of different mutations, resulting in intermediate lycopene synthesis and accumulation of other carotenoids (*CRTISO*, *Crtl-b*, *Crtl-e*), or limited availability of precursor of carotenoids (*IDI1*) [44, 45]. The founder lines of SABER population are all with red fruits, except for the wild ancestor, *S. cheesmanie*, that has orange fruits. Among the offspring, there is a little diversification between the epicarp color (ECO: 1 yellow, 27 orange, 212 red) and the flesh color (ICO: 36 orange, 204 red). The singular yellow fruit could be a phenotypization error, since the evaluation of the color has been done with the naked eye. However, in literature it is reported a case of cross of a double mutant yellow x orange (PSY1 x CRTISO) that resulted in an orange phenotype even if the *PSY1* mutation should be dominant since the enzyme is earlier in the pathway. Kachanovsky et al. [46] found that the *CRTISO* gene could have epistatic effect on *PSY1* partially restoring its transcription. Theoretically, the yellow phenotype could emerge by segregation of the two loci from an orange individual.

The cases of different colors between flesh and epicarp were rare in the population (3 with orange epicarp and red flesh, 12 with red epicarp and orange flesh, 1 with yellow exocarp and red flesh), so it does not surprise that the QTL identified for the two phenotypes are practically identical; here they will be discussed together. The QTL is positioned at chromosome 6, with a peak at position 43,541,536 and a confidence interval that span a physical region of 0.86 Mb (Table 3; Fig. 5). The genetic determinant identified with high probability is *CYC-B* (*Solyc06g074240*, coordinates according to annotation ITAG4.1: chromosome 6 43,562,526–43,564,022 bp), a lycopene-beta cyclase characterized in the tomato clade. Mohan et al. [47] demonstrated how *CYC-B* underwent purifying selection in cultivated *S. lycopersicum*, while in wild tomatoes present greater diversity in the promotorial region. Higher expression levels of *CYC-B* are associated with beta-carotene accumulation and lower levels of lycopene, resulting in orange fruits. Since *CYC-B* is positioned just outside the confidence interval of the QTL, we analyzed the 122 genes inside that span (Table S5 in Additional file 1), searching for alternative candidates. Among them, we searched for predicted genes showing homology with genes possibly involved in the external and internal coloration of the fruit. Among the most compelling candidate genes within the region, *ACO1* (1-aminocyclopropane-1-carboxylate oxidase 1; *Solyc06g073580*) clearly stands out. *ACO1* is the major *ACO* transcript accumulating during fruit ripening and catalyzes the oxidation of ACC (1-aminocyclopropane-1-carboxylic acid) to ethylene, the final step of the ethylene biosynthesis pathway [48]. Ethylene is essential for climacteric fruit ripening and plays a central role in both internal and external fruit coloration by promoting lycopene accumulation (red pigmentation) and chlorophyll degradation (green-to-red transition) [49, 50]. Transcriptomic analyses suggested that ethylene production via *ACO1* influences carotenoid biosynthesis through *PSY1* expression. Conversely, reduced *ACO1* expression - such as after treatment with 1-methylcyclopropene (1-MCP) - inhibits lycopene accumulation and associated color changes, further confirming its key role in fruit pigmentation [51]. Within the same chromosomal interval, two additional and adjacent genes encode *ETHYLENE INSENSITIVE* 3 (EIN3)-like transcription factors, namely *EIL1* (*Solyc06g073720*) and *EIL4* (*Solyc06g073730*). These transcription factors are key components of the ethylene signaling cascade and may therefore also contribute to fruit pigmentation processes. Although a direct regulation of carotenoid biosynthetic genes by *EIL1* and *EIL4* has not yet been clearly demonstrated in tomatoes, evidence from other climacteric species supports their potential involvement in ripening-related pathways. In melon (*Cucumis melo*), for instance, *EIL1* has been characterized as a ripening-associated gene and shown to function as a transcriptional activator of *ACO1* [52].

Flowering time

Days to flowering (DTF) and leaves number before flowering (LN) are two measurements used for determining the length of the vegetative phase and are essential for the management of the plants in an industrial context. In the domestication and varieties breeding process short and uniform vegetative cycles were favored, with the goal

Table 3 QTLs mapped in SABER population. Start and end position of the QTL represent the 95% Bayesian credible interval

Phenotype	Chromosome	Peak position (pb)	LOD score	Start (pb)	End (pb)
Hypocotyl color	Chr9	59,052,650	29.83	57,908,608	59,165,600
Obscura venosa	Chr5	62,575,099	25.15	62,575,099	63,569,772
Green shoulder	Chr10	2,010,429	8.94	1,208,771	19,083,107
Epicarp color	Chr6	43,541,490	24.05	42,683,303	43,541,554
Flesh color	Chr6	43,541,536	16.38	42,683,303	43,541,554
BRIX°	Chr6	5,023,437	6.89	20,652	27,019,138
Leaves before flowering	Chr10	1,738,579	7.13	1,413,363	1,829,496
Days before flowering	Chr3	57,494,257	6.80	56,893,414	59,699,304

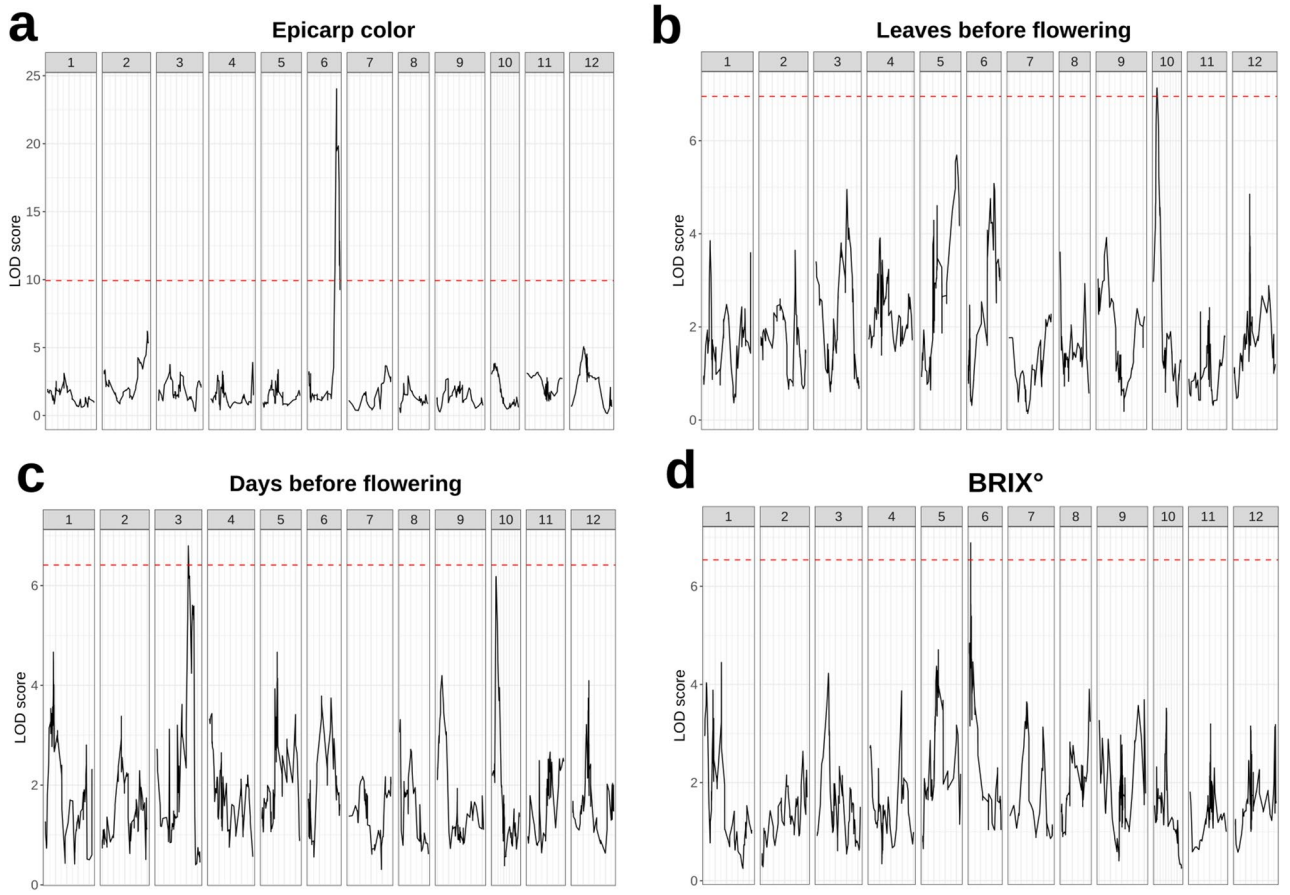


Fig. 5 LOD scores of QTLs across the genome, divided in the 12 chromosomes, indicated by their numbers. The LOD peaks are relative to QTLs of: epicarp color (**a**), number of leaves before flowering (**b**), days before flowering (**c**) and BRIX value of the ripe fruit (**d**). Markers are positioned according to the linkage map. The dotted red line indicates the genome-wide significance threshold (alpha = 0.05) determined individually for each trait via 10,000 permutations

to accelerate production and, most of all, obtain synchronous ripening of the fruits, that could be harvested at the same moment [53]. Given the importance, time flowering has been extensively documented in the literature. QTLs that influence DTF or LN have been mapped in chromosomes 1, 2, 3, 5, 6, 7, 9, and 11 (53, 54, 55, 56, 57] and in many cases functional characterization of the candidate genes confirmed their role (58, 59, 60]. Since the cultivated tomato shows a phenotype divergent from the wild tomatoes, in many cases the association studies have been performed on hybrid population with a wild ancestor: *S. pimpinellifolium* [55], *S. Chmielewski* [53] or *S. pennelli* [56]. This is the first study of the character *S. cheesmaniae* as a parent in order to obtain variability of the phenotype. The two measurements of the trait, the time from transplant to first flowering and the number of leaves produced are positively correlated in SABER population (Pearson correlation: 0.57, p-value < 2.2 e -16), similarly to what can be observed in literature (*ibidem*). However, the genes controlling DTF and LN are only partially overlapping, as confirmed by both QTL mapping [53] and functional studies [60]. The QTLs mapped in SABER population are consistent with the literature. The LOD peaks of the two phenotypes are positioned in two different chromosomes (Table 3) but is evident how comparing the LOD profiles there is a spot on chromosome 10 around 1.7 Mb with influx in both characters, even if it passes the significance threshold only for LN (Fig. 5).

The QTL confidence interval in this region encompassed a 0.42 Mb region and included 59 annotated genes in the mapping performed with the LN phenotype (Table S7 in Additional file 1). Among them, we focused our attention on two consecutive genes codifying for as many Gibberellin 2-oxidases (GA2ox, *Solyc10g007570* and *Solyc10g007560*). The GA2ox family, a class of 2-oxoglutarate-dependent dioxygenases, regulates the levels of bioactive gibberellins by catalyzing their deactivation, which influences various aspects of plant growth and development [61]. While direct evidence linking GA2ox specifically to the number of leaves between cotyledons and the first truss is limited, this family broadly affects shoot architecture, leaf development, and branching

patterns through modulation of gibberellin levels. For example, overexpression of GA2ox genes can lead to dwarfism, smaller leaves, and modifications in the primary and secondary growth of the stem, indicating their role in controlling vegetative growth [62–64]. Additionally, in the frame of soybean domestication, copy number variations in certain GA2ox genes were found to correlate with traits related to overall plant architecture, such as reduced shoot length and trailing growth [65]. Although proposing the expression regulation of the two GA2ox genes as the underlying cause of the observed QTL remains speculative at this early stage, their role would logically align with explaining the broad-spectrum effects observed for both LN and DTF traits.

The highest LOD peak for DTF is located on chromosome 3 (56.8–59.7 Mb), a region lacking noticeable peaks in the LN mapping. This novel locus has not been previously reported in the literature to contribute to tomato flowering time. The LOD peak spans a 2.8 Mb region on chromosome 3 and includes 376 annotated genes (Table S6 in Additional file 1), among which two resulted of particular interest. Foremost among these is FRUITFULL-like MADS-box 2 (*FUL2*; *Solyc03g114830*), given its well-established role in tomato in promoting the transition from vegetative to reproductive development by inducing floral meristem maturation and repressing inflorescence branching [66]. Moreover, it has been demonstrated that the *ful2 mbp20* (*Solyc02g089210*) double mutant exhibits delayed flowering [67]. Notably, the gene immediately adjacent to *FUL2*, *MADS1* (also known as ENHANCER OF JOINTLESS 2, *EJ2*; *Solyc03g114840*), has likewise been reported to influence inflorescence branching [68]. Another strong candidate within the QTL interval is SQUAMOSA PROMOTER BINDING PROTEIN 6a (*SlSBP6a*, *Solyc03g114850*), located immediately downstream of the two aforementioned genes (*FUL2* and *MADS1*), further highlighting this genomic cluster as a key regulatory hotspot for flowering-related traits [69]. *SlSBP6a* in tomato is part of the miR156-targeted SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL/SBP) family, which plays a crucial role in regulating flowering time by promoting floral meristem maturation and transition to flowering. In particular, these miR156-targeted SBPs interact with gibberellin (GA) signaling pathways, synergistically regulating floral meristem determinacy and influencing flowering time [69].

As previously noted, the flowering time QTL identified in the SABER population does not appear to co-localize with any major loci previously described in the literature, including well-characterized genetic determinants such as SP5G and SFT.

SP5G, located on chromosome 5, is a thoroughly characterized flowering regulator [70] whose expression in domesticated tomato differs from wild species due to a cis-regulatory mutation. Several hypotheses may explain why the effect of this gene was not identified as statistically significant by our mapping model. First, according to Soyk et al., *S. cheesmaniae* exhibits the least divergent flowering behavior relative to *S. lycopersicum* among wild tomato species. Furthermore, we observed a severe segregation distortion on chromosome 5 regarding parental allele contributions to the progeny. Although the LA2460 genomic segment is present, its frequency drops to just below 10% on the long arm of chromosome 5, where SP5G is physically located. It is possible that this combination of low parental divergence and distorted allele frequencies reduced the statistical power, ultimately preventing the detection of a significant QTL at this locus.

As regards SFT (SINGLE FLOWER TRUSS), It is worth noting that while the founder line LA2460 is known to carry the sft allele, a major regulator of tomato flowering time and architecture [71], the characteristic sft mutant phenotype was not observed in the SABER population. It is well established that the phenotypic effects of this gene vary depending on the genomic background into which it is introduced, due to epistatic interactions with other loci [72]. It is highly probable that the introduction of alleles from the wild relative contributed to masking the expected mutant phenotype.

Sugars and organic acids concentration in the fruits

Brix Degrees measure the levels of dissolved solids in a solution and are the first estimate of the sugar and acids contained in tomato fruit. It is an important assessment of the quality of the product for tomatoes for every destination, industrial or fresh market. Obviously, the character has a substantial body of literature dedicated to it, and since it is a measurement of a metabolic state of the plant, it is considered a highly polygenic character, with a high number of QTLs identified. Already in 2000 Fridman et al. [73] characterized a QTL on chromosome 9, since then many more were identified in chromosomes 1, 2, 7, 8, 11 and 12 [74–78] .It has been extensively reported that Brix values can be negatively correlated with fruit size. However, in the SABER population, Brix degree and fruit weight showed little to no correlation (Pearson's correlation: -0.069, p-value: 0.28). To confirm that fruit weight did not significantly influence the Brix evaluation, we performed a corrected QTL mapping using fruit weight as a covariate. This analysis yielded nearly identical results compared to the one without covariates (Additional file 4)."

The QTL identified in SABER population is located in the first arm of chromosome 6 (0–27 Mb, Table 3), a position where no QTL for BRIX values was identified, with the exception of a QTL linked to the exposition of the plant to a LED light, located at 44.6 Mb [78]. The absence of several historically well-described QTLs

is not surprising, given that many were originally identified using introgression lines from wild relatives distinct from *S. cheesmaniae*, most notably *S. pennellii* (e.g., the Brix9-2-5 and Malodorous QTLs [73, 79] and *S. pimpinellifolium* (e.g., fw2.2 and fw3.2 [80, 81]). Because *S. cheesmaniae* is a biologically distinct species from these wild relatives, the genetic diversity within these specific loci is, in all probability, insufficient to drive such a pronounced variation in Brix values in our population. The confidence interval for the position of the QTL is very broad and encompasses nearly 700 genes (Table S8 in Additional file 1); nevertheless, it is worth discussing some of the genes comprised in this region, considering the importance of the trait and the novelty of the identified QTL. *Solyc06g005070* is a member of the Major Facilitator Superfamily protein (MFS), a superfamily of sugar transporters characterized by 12 transmembrane. Many members of the superfamily have been already characterized for their involvement in Brix determination and constitute the probable genetic determinant of identified QTLs [82]. Isoamlylase 3 (*ISA3*, *Solc06g009220*) is a chloroplast enzyme deputed to the debranching of starch, identified as one of the key enzymes in starch catabolism in apple and play a role also in grape and sweet orange [83]. Starch is found only in early stages of tomato fruit [83] and its degradation could be a source for the production of sugars and organic acids. Finally, Li et al. [84] in their analysis of BRIX-related genes using the correlation of expression level and Brix values, identified S*olyc06g005970* positively correlated with sugar levels (fructose and glucose) negatively correlated with organic acid levels (malic and citric acid). The gene, then uncharacterized, is now annotated as a protein of the glycosyl hydrolase family.

Conclusions

Expanding the genetic base of cultivated tomatoes beyond the narrow diversity retained after domestication is a central objective of contemporary breeding and genomic research. The SABER population contributes directly to this effort, being the first MAGIC population to include as a founder line an accession of *Solanum cheesmaniae*, an extremophile wild relative from the Galápagos. The successful introgression of *S. cheesmaniae* alleles across all chromosomes, with no genomic regions showing a complete absence of wild-founder contribution, demonstrates that the SABER design effectively captured exotic genetic variation within a usable quantitative genetic framework. The relatively uniform mean founder contributions and the little population structure among the offspring further supports the population's suitability for unbiased QTL discovery. The testing of the potential of the population through three well-characterized Mendelian traits (obscuravenosa, uniform ripening, and hypocotyl color) confirmed the reliability of the analytical pipeline and the quality of the genotypic data obtained through SPET-based genotyping. Beyond these controls, the QTL of quantitative traits yielded both results consistent with the existing literature and novel genomic regions not previously associated with these phenotypes, suggesting that the allelic diversity introduced by the wild founder is actively contributing to phenotypic variation within the population. From a broader perspective, SABER establishes a replicable model for integrating wild germplasm into multiparental breeding resources. As climate change increasingly threatens the productivity and resilience of major crops, the ability to exploit the adaptive potential encoded in wild relatives becomes strategically essential. Future work should focus on expanding phenotyping efforts — particularly for abiotic stress-related traits — and on fine-mapping the novel QTL regions identified here to accelerate candidate gene characterization and marker-assisted selection.

Abbreviations

ACC	1-aminocyclopropane-1-carboxylic acid
DTF	Days To Flowering
GWAS	Genome-Wide Association Study
LD	Linkage Disequilibrium
LED	Light Emitting Diod
Kb	Kilo base pairs (thousands of base pairs)
LMM	Linear Mixed Model
LN	Leaves Number (before flowering)
LOCO	Leaves One Choromosome Out
LOD	Logarithm of Odds
MAGIC	Multi-parent Advanced Generation Inter-Cross
Mb	Mega base pairs (million of base pairs)
MFS	Major Facilitator Superfamily
PCA	Principal Component Analysis
PC	Principal Component
QTL	Quantitative Trait Locus
RIL	Recombinant Inbred Lines
SABER	*Solanum lycopersicum* Allele Biodiversity Enriched Resources
SL2.4	*Solanum lycopersicum* reference genome V2.4
SL4.0	*Solanum lycopersicum* reference genome V4
SNP	Single Nucleotide Polymorphism
SPET	Single Primer Enrichment Technology
SSD	Single Seed Descent

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-026-09726-w.

Additional file 1: Supplementary tables. Supplementary tables, one each sheet. The name of the sheet corresponds to the name of the table as cited in the main text. Table S1 – Pedigree and proprieties of the eight founder lines. Table S2 – Phenotypes of parental lines and F6 samples. Hypocotyl color (ANT = purple, AH = green), Obscuravenosa (NT = not transparent, T = transparent), Green shoulder (G = present, UG = not present), Epicarp and Flesh color (R = red, Beta = orange, Y = yellow). The quantitative traits (°Brix values, leaves before flowering LN, Days to flowering DTF, Weight) are measured as descripted in chapter 2.4. Table S3 – Linkage map of SABER with the physical (bp) and genetic (cM) position of each marker. Table S4 – Sequencing, alignment and SNP calling metrics. The columns contain: sample ID, number of pairs of total unfiltered reads (total sequenced reads), number of pairs of filtered reads (filtered reads), number of successfully aligned reads (aligned reads), number of reads aligned with a quality equal or higher than 5 (aligned reads q5), average

depth of reads on SNP called (Average depth), number and percentage of: loci with homozygous genotype equal to the reference genome SL4.0 (Reference homozygous), loci with homozygous genotype alternative to the reference genome SL4.0 (Alternative Homozygous), heterozygous loci (Heterozygous) and loci with missing genotypes (Missing). Table S5 – List of genes and their functional annotation comprised in LOD curve of QTL for epicarp and flesh color, according to annotation ITAG4.1. Table S6 - List of genes and their functional annotation comprised in LOD curve of QTL for DTF, according to annotation ITAG4.1. Table S7 - List of genes and their functional annotation comprised in LOD curve of QTL LN, according to annotation ITAG4.1. Table S8 - List of genes and their functional annotation comprised in LOD curve of QTL for °Brix, according to annotation ITAG4.1

Additional file 2: Chain file from SL2.4 to SL4.0. chain files to lift over the coordinates from S. lycopersicum annotation SL2.4 to SL4.0 with CrossMap (https://crossmap.readthedocs.io/).

Additional file 3: Supplementary figure S1. Density of SNPs with genotypes alternative to the reference SL4.0 of the eight founder lines, plotted binning the genomes in 200Kb windows. The twelve vertical bars represent the twelve chromosomes, the eight founder lines are represented by the colored box under each chromosome. The coordinates of the chromosomes are in the y-axis.

Additional file 4: Supplementary figure S2. Comparison of LOD score across the genome using as phenotype in the model Brix degree (black line) and Brix degree corrected for the weight of the fruit (red line), inserting the values as covariate in the model.

Acknowledgements
The authors gratefully acknowledge Daniela Palma for her contribution to the laboratory work, specifically regarding DNA extraction processes.

Authors' contributions
M.B., S.S., and G.B. conceived and designed the study. M.B., S.S., and M.D.A. coordinated the research group and secured funding for the project. A.G.B., G.F., L.P., M.M., S.S., and M.B. were responsible for population planning and maintenance, as well as all phenotyping and genotyping processes. G.G. and L.C. performed data management, designed and executed the formal analysis, and prepared the figures. The manuscript was written by G.G., L.C., M.B., F.P., and S.S., and revised by all authors. All authors have read and agreed to the published version of the manuscript.

Funding
This research was supported by funding from Panora seeds srl and the following research projects: Agritech National Research Center, funded by the European Union – NextGenerationEU (National Recovery and Resilience Plan, PNRR, Mission 4, Component 2, Investment 1.4 – Project CN2, Spoke 4); LIVESEED, European Union's Horizon 2020 research and innovation programme [Grant Agreement No. 727230]; National Recovery and Resilience Plan (NRRP) – Mission 4, Component 2, Investment 1.4 – Strengthening research infrastructures and creating "National Champions in R&D" in selected Key Enabling Technologies. National Research Centre for Agricultural Technologies (AGRITECH) – CN_00000022 funded by the European Union – NextGenerationEU. CUP: J53C22001610007.

Data availability
The sequencing data generated and analyzed during the current study are available in NCBI's SRA repository, under the project ID PRJNA1417763.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 15 May 2026 / Accepted: 7 August 2026
Published online: 15 August 2026

References

1. Gramazio P, Pereira-Dias L, Vilanova S, Prohens J, Soler S, Esteras J, et al. 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. Hortic Res. 2020;7(1):174. https://doi.org/10.1038/s41438-020-00395-w.
2. Pascual L, Desplat N, Huang BE, Desgroux A, Bruguier L, Bouchet JP, et al. 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. 2015;13(4):565–77. https://doi.org/10.1111/pbi.12282.
3. Arrones A, Antar O, Pereira-Dias L, Solana A, Ferrante P, Aprea G, et al. A novel tomato interspecific (*Solanum lycopersicum* var. *cerasiforme* and *Solanum pimpinellifolium*) MAGIC population facilitates trait association and candidate gene discovery in untapped exotic germplasm. Hortic Res. 2024;11(7):uhae154. https://doi.org/10.1093/hr/uhae154.
4. Blanca J, Montero-Pau J, Sauvage C, Bauchet G, Illa E, Díez MJ, et al. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. BMC Genomics. 2015;16(1):257. https://doi.org/10.1186/s12864-015-1444-1.
5. 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;51(6):1044–51. https://doi.org/10.1038/s41588-019-0410-2.
6. Alonge M, Wang X, Benoit M, Soyk S, Pereira L, Zhang L, et al. Major Impacts of Widespread Structural Variation on Gene Expression and Crop Improvement in Tomato. Cell. 2020;182(1):145–e16123. https://doi.org/10.1016/j.cell.2020.05.021.
7. Zhou Y, Zhang Z, Bao Z, Li H, Lyu Y, Zan Y, et al. Graph pangenome captures missing heritability and empowers tomato breeding. Nature. 2022;606(7914):527–34. https://doi.org/10.1038/s41586-022-04808-9.
8. Li N, He Q, Wang J, Wang B, Zhao J, Huang S, et al. Super-pangenome analyses highlight genomic diversity and structural variation across wild and cultivated tomato species. Nat Genet. 2023;55(5):852–60. https://doi.org/10.1038/s41588-023-01340-y.
9. Yu J, Chen Q, Yuan L, Feng S, Huang M, Zheng P, et al. Phylogenomic and super-pangenome analyses unveil the genetic landscape of tomato evolution and domestication. Plant Biotechnol J. 2025;23(9):3783–97. https://doi.org/10.1111/pbi.70199.
10. Shi C, Chen S, Wang J, Chen W, Sun C, Guo Q, et al. A tomato telomere-to-telomere super-pangenome empowers stress resilience breeding. Nat Genet. 2026;1–13. https://doi.org/10.1038/s41588-026-02508-y.
11. Arrones A, Vilanova S, Plazas M, Mangino G, Pascual L, Díez MJ, et al. 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. 2020;9(8). https://doi.org/10.3390/biology9080229.
12. Ferguson JN, Caproni L, Walter J, Shaw K, Arce-Cubas L, Baines A, et al. A deficient CP24 allele defines variation for dynamic nonphotochemical quenching and photosystem II efficiency in maize. Plant Cell. 2025;37(4):koaf063. https://doi.org/10.1093/plcell/koaf063.
13. Campanelli G, Sestili S, Acciarri N, Montemurro F, Palma D, Leteo F, et al. Multi-Parental Advances Generation Inter-Cross Population, to Develop Organic Tomato Genotypes by Participatory Plant Breeding. Agronomy. 2019;9(3). https://doi.org/10.3390/agronomy9030119.
14. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv; 2013. Available from: http://arxiv.org/abs/1303.3997 https://doi.org/10.48550/arXiv.1303.3997. Cited 2026 Mar 12.
15. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, der Auwera GAV et al. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv; 2018. p. 201178. Available from: https://www.biorxiv.org/content/ https://doi.org/10.1101/201178v3 doi:10.1101/201178. Cited 2026 Mar 12.
16. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, der Auwera GAV, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv; 2018. p. 201178. Available from: https://www.biorxiv.org/content/ https://doi.org/10.1101/201178v3 doi:10.1101/201178. Cited 2022 Oct 26.
17. Hosmani PS, Flores-Gonzalez M, Geest H, van de, Maumus F, Bakker LV, Schijlen E, et al. An improved de novo assembly and annotation of the tomato reference genome using single-molecule sequencing, Hi-C

proximity ligation and optical maps. bioRxiv; 2019. p. 767764. Available from: https://www.biorxiv.org/content/ https://doi.org/10.1101/767764v1 doi:10.1101/767764. Cited 2026 Mar 12.
18. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841–2. https://doi.org/10.1093/bioinformatics/btq033.
19. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2):giab008. https://doi.org/10.1093/gigascience/giab008.
20. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90. https://doi.org/10.1093/bioinformatics/bty560.
21. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32(19):3047–8. https://doi.org/10.1093/bioinformatics/btw354.
22. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics. 2016;32(2):292–4. https://doi.org/10.1093/bioinformatics/btv566.
23. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156–8. https://doi.org/10.1093/bioinformatics/btr330.
24. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4(1):s13742. 015-0047-8.
25. Caproni L, Lakew BF, Kassaw SA, Miculan M, Ahmed JS, Grazioli S, et al. The genomic and bioclimatic characterization of Ethiopian barley (Hordeum vulgare L.) unveils challenges and opportunities to adapt to a changing climate. Glob Change Biol. 2023;29(8):2335–50. https://doi.org/10.1111/gcb.16560.
26. Shin JH, Blay S, McNeney B, Graham J. LDheatmap: An R Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms. J Stat Softw. 2006;16:1–9. https://doi.org/10.18637/jss.v016.c03.
27. Macharia MW, Caproni L, Solemanegy M, Takele R, Chiulele RM, Munisse P, et al. Cowpea (Vigna unguiculata L. Walp.) landraces in Mozambique and neighbouring Southern African countries harbour genetic loci with potential for climate adaptation. Commun Biol. 2025;8(1):834. https://doi.org/10.1038/s42003-025-08227-0.
28. Jombart T, Ahmed I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics. 2011;27(21):3070–1. https://doi.org/10.1093/bioinformatics/btr521.
29. Paradis E, Claude J, Strimmer K. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics. 2004;20(2):289–90. https://doi.org/10.1093/bioinformatics/btg412.
30. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics. 2012;28(24):3326–8. https://doi.org/10.1093/bioinformatics/bts606.
31. Zhao H, Sun Z, Wang J, Huang H, Kocher JP, Wang L. CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics. 2014;30(7):1006–7. https://doi.org/10.1093/bioinformatics/btt730.
32. TAO Z. tao-bioinfo/pyOverChain [Python]. 2026. Available from: https://github.com/tao-bioinfo/pyOverChain. Cited 2026 Mar 12.
33. Index of /ftp/tomato_genome. /. Available from: https://solgenomics.net/ftp/tomato_genome/. Cited 2025 Nov 28.
34. Broman KW, Gatti DM, Simecek P, Furlotte NA, Prins P, Sen Ś, et al. R/qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations. Genetics. 2019;211(2):495–502. https://doi.org/10.1534/genetics.118.301595.
35. Barchi L, Acquadro A, Alonso D, Aprea G, Bassolino L, Demurtas O, et al. Single Primer Enrichment Technology (SPET) for High-Throughput Genotyping in Tomato and Eggplant Germplasm. Front Plant Sci. 2019;10. https://doi.org/10.3389/fpls.2019.01005.
36. Valdar W, Flint J, Mott R. Simulating the Collaborative Cross: Power of Quantitative Trait Loci Detection and Mapping Resolution in Large Sets of Recombinant Inbred Strains of Mice. Genetics. 2006;172(3):1783–97. https://doi.org/10.1534/genetics.104.039313.
37. Dell'Acqua M, Gatti DM, Pea G, Cattonaro F, Coppens F, Magris G, et al. Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in Zea mays. Genome Biol. 2015;16(1):167. https://doi.org/10.1186/s13059-015-0716-z.
38. Lu J, Pan C, Li X, Huang Z, Shu J, Wang X, et al. OBV (obscure vein), a C2H2 zinc finger transcription factor, positively regulates chloroplast development and bundle sheath extension formation in tomato (Solanum lycopersicum) leaf veins. Hortic Res. 2021;8:230. https://doi.org/10.1038/s41438-021-00659-.z PubMed PMID: 34719693; PubMed Central PMCID: PMC8558323.
39. Powell ALT, Nguyen CV, Hill T, Cheng KL, Figueroa-Balderas R, Aktas H, et al. Uniform ripening Encodes a Golden 2-like Transcription Factor Regulating Tomato Fruit Chloroplast Development. Science. 2012;336(6089):1711–5. https://doi.org/10.1126/science.1222218.
40. Qiu Z, Wang X, Gao J, Guo Y, Huang Z, Du Y. The Tomato Hoffman's Anthocyaninless Gene Encodes a bHLH Transcription Factor Involved in Anthocyanin Biosynthesis That Is Developmentally Regulated and Induced by Low Temperatures. PLoS ONE. 2016;11(3):e0151067. https://doi.org/10.1371/journal.pone.0151067.
41. Jones CM, Rick CM, Adams D, Jernstedt J, Chetelat RT. Genealogy and fine mapping of obscuravenosa, a gene affecting the distribution of chloroplasts in leaf veins, and evidence of selection during breeding of tomatoes (Lycopersicon esculentum; Solanaceae). Am J Bot. 2007;94(6):935–47. https://doi.org/10.3732/ajb.94.6.935.
42. Zhang L, Huang Z, Wang X, Gao J, Guo Y, Du Y, et al. Fine mapping and molecular marker development of anthocyanin absent, a seedling morphological marker for the selection of male sterile 10 in tomato. Mol Breed. 2016;36(8):107. https://doi.org/10.1007/s11032-016-0531-6.
43. Casals J, Romero del Castillo R, Pons C, Mazzucato A, Tringovska I, Pasev G, et al. European fresh-market tomato sensory ideotypes based on consumer preferences. Sci Hort. 2024;335:113351. https://doi.org/10.1016/j.scienta.2024.113351.
44. Yang T, Ali M, Lin L, Li P, He H, Zhu Q, et al. Recoloring tomato fruit by CRISPR/Cas9-mediated multiplex gene editing. Hortic Res. 2023;10(1):uhac214. https://doi.org/10.1093/hr/uhac214.
45. Duduit JR, Kosentka PZ, Miller MA, Blanco-Ulate B, Lenucci MS, Panthee DR, et al. Coordinated transcriptional regulation of the carotenoid biosynthesis contributes to fruit lycopene content in high-lycopene tomato genotypes. Hortic Res. 2022;9:uhac084. https://doi.org/10.1093/hr/uhac084.
46. Kachanovsky DE, Filler S, Isaacson T, Hirschberg J. Epistasis in tomato color mutations involves regulation of phytoene synthase 1 expression by cis-carotenoids. Proc Natl Acad Sci. 2012;109(46):19021–6. https://doi.org/10.1073/pnas.1214808109.
47. Mohan V, Pandey A, Sreelakshmi Y, Sharma R. Neofunctionalization of Chromoplast Specific Lycopene Beta Cyclase Gene (CYC-B) in Tomato Clade. PLoS ONE. 2016;11(4):e0153333. https://doi.org/10.1371/journal.pone.0153333.
48. Kende H. Ethylene Biosynthesis. Annu Rev Plant Biol. 1993;44(44, 1993):283–307. https://doi.org/10.1146/annurev.pp.44.060193.001435.
49. Alexander L, Grierson D. Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. J Exp Bot. 2002;53(377):2039–55. https://doi.org/10.1093/jxb/erf072.
50. Liu L, Wei J, Zhang M, Zhang L, Li C, Wang Q. Ethylene independent induction of lycopene biosynthesis in tomato fruits by jasmonates. J Exp Bot. 2012;63(16):5751–61. https://doi.org/10.1093/jxb/ers224.
51. Hoeberichts FA, Van Der Plas LHW, Woltering EJ. Ethylene perception is required for the expression of tomato ripening-related genes and associated physiological changes even at advanced stages of ripening. Postharvest Biol Technol. 2002;26(2):125–33. https://doi.org/10.1016/S0925-5214(02)00012-1.
52. Huang S, Sawaki T, Takahashi A, Mizuno S, Takezawa K, Matsumura A, et al. Melon EIN3-like transcription factors (CmEIL1 and CmEIL2) are positive regulators of an ethylene- and ripening-induced 1-aminocyclopropane-1-carboxylic acid oxidase gene (*CM-ACO1*). Plant Sci. 2010;178(3):251–7. https://doi.org/10.1016/j.plantsci.2010.01.005.
53. Jiménez-Gómez JM, Alonso-Blanco C, Borja A, Anastasio G, Angosto T, Lozano R, et al. Quantitative genetic analysis of flowering time in tomato. Genome. 2007;50(3):303–15. https://doi.org/10.1139/G07-009.
54. Cagas CC, Lee ON, Nemoto K, Sugiyama N. Quantitative trait loci controlling flowering time and related traits in a *Solanum lycopersicum* × *S. pimpinellifolium* cross. Sci Hort. 2008;116(2):144–51. https://doi.org/10.1016/j.scienta.2007.12.003.
55. Sumugat MR, Lee ON, Nemoto K, Sugiyama N. Quantitative trait loci analysis of flowering-time-related traits in tomato. Sci Hort. 2010;123(3):343–9. https://doi.org/10.1016/j.scienta.2009.10.002.
56. Fulop D, Ranjan A, Ofner I, Covington MF, Chitwood DH, West D, et al. A New Advanced Backcross Tomato Population Enables High Resolution Leaf QTL Mapping and Gene Identification. (Bethesda). 2016;G3(10):3169–84. https://doi.org/10.1534/g3.116.030536 . PubMed PMID: 27510891; PubMed Central PMCID: PMC5068939.

57. Ruangrak E, Du Y, Htwe NMPS, Pimorat P, Gao J. Identification of Early Tomato Fruit Ripening Loci by QTL-seq. J Agric Sci. 2019;11(2):p51. https://doi.org/10.5539/jas.v11n2p51.
58. Molinero-Rosales N, Jamilena M, Zurita S, Gómez P, Capel J, Lozano R. FALSIFLORA, the tomato orthologue of FLORICAULA and LEAFY, controls flowering time and floral meristem identity. Plant J. 1999;20(6):685–93. https://doi.org/10.1046/j.1365-313X.1999.00641.x.
59. MacAlister CA, Park SJ, Jiang K, Marcel F, Bendahmane A, Izkovich Y, et al. Synchronization of the flowering transition by the tomato TERMINATING FLOWER gene. Nat Genet. 2012;44(12):1393–8. https://doi.org/10.1038/ng.2465.
60. Zhang D, Ai G, Ji K, Huang R, Chen C, Yang Z, et al. EARLY FLOWERING is a dominant gain-of-function allele of FANTASTIC FOUR 1/2c that promotes early flowering in tomato. Plant Biotechnol J. 2024;22(3):698–711. https://doi.org/10.1111/pbi.14217.
61. Li C, Zheng L, Wang X, Hu Z, Zheng Y, Chen Q, et al. Comprehensive expression analysis of Arabidopsis *GA2-oxidase* genes and their functional insights. Plant Sci. 2019;285:1–13. https://doi.org/10.1016/j.plantsci.2019.04.023.
62. Mauriat M, Sandberg LG, Moritz T. Proper gibberellin localization in vascular tissue is required to control auxin-dependent leaf development and bud outgrowth in hybrid aspen. Plant J. 2011;67(5):805–16. https://doi.org/10.1111/j.1365-313X.2011.04635.x.
63. Hu YX, Tao YB, Xu ZF. Overexpression of Jatropha Gibberellin 2-oxidase 6 (JcGA2ox6) Induces Dwarfism and Smaller Leaves, Flowers and Fruits in Arabidopsis and Jatropha. Front Plant Sci. 2017;8. https://doi.org/10.3389/fpls.2017.02103.
64. Cheng J, Ma J, Zheng X, Lv H, Zhang M, Tan B, et al. Functional Analysis of the Gibberellin 2-oxidase Gene Family in Peach. Front Plant Sci. 2021;12. https://doi.org/10.3389/fpls.2021.619158.
65. Wang X, Li MW, Wong FL, Luk CY, Chung CYL, Yung WS, et al. Increased copy number of gibberellin 2-oxidase 8 genes reduced trailing growth and shoot length during soybean domestication. Plant J. 2021;107(6):1739–55. https://doi.org/10.1111/tpj.15414.
66. Jiang X, Lubini G, Hernandes-Lopes J, Rijnsburger K, Veltkamp V, de Maagd RA, et al. FRUITFULL-like genes regulate flowering time and inflorescence architecture in tomato. Plant Cell. 2022;34(3):1002–19. https://doi.org/10.1093/plcell/koab298.
67. Jiang X, Zahn IE, Thoris K, Roelofsen C, Roque E, Gómez-Mena C, et al. Tomato flowering depends on overlapping functions of AP1/FUL-like genes in reproductive meristem specification. New Phytol. 2025;248(2):1002–20. https://doi.org/10.1111/nph.70451.
68. Soyk S, Lemmon ZH, Oved M, Fisher J, Liberatore KL, Park SJ, et al. Bypassing Negative Epistasis on Yield in Tomato Imposed by a Domestication Gene. Cell. 2017;169(6):1142–e115512. https://doi.org/10.1016/j.cell.2017.04.032. PubMed PMID: 28528644.
69. Silva GFF, Silva EM, Correa JPO, Vicente MH, Jiang N, Notini MM, et al. Tomato floral induction and flower development are orchestrated by the interplay between gibberellin and two unrelated microRNA-controlled modules. New Phytol. 2019;221(3):1328–44. https://doi.org/10.1111/nph.15492.
70. Soyk S, Müller NA, Park SJ, Schmalenbach I, Jiang K, Hayama R, et al. Variation in the flowering gene SELF PRUNING 5G promotes day-neutrality and early yield in tomato. Nat Genet. 2017;49(1):162–8. https://doi.org/10.1038/ng.3733.
71. Molinero-Rosales N, Latorre A, Jamilena M, Lozano R. SINGLE FLOWER TRUSS regulates the transition and maintenance of flowering in tomato. Planta. 2004;218(3):427–34. https://doi.org/10.1007/s00425-003-1109-1.
72. Krieger U, Lippman ZB, Zamir D. The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. Nat Genet. 2010;42(5):459–63. https://doi.org/10.1038/ng.550.
73. Fridman E, Pleban T, Zamir D. A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proc Natl Acad Sci. 2000;97(9):4718–23. https://doi.org/10.1073/pnas.97.9.4718.
74. Petreikov M, Yeselson L, Shen S, Levin I, Schaffer AA, Efrati A, et al. Carbohydrate Balance and Accumulation during Development of Near-isogenic Tomato Lines Differing in the AGPase-L1 Allele. J Am Soc Hortic Sci. 134(1):134–40. https://doi.org/10.21273/JASHS.134.1.134.
75. Lippman ZB, Semel Y, Zamir D. An integrated view of quantitative trait variation using tomato interspecific introgression lines. Curr Opin Genet Dev. 2007;17(6):545–52. https://doi.org/10.1016/j.gde.2007.07.007.
76. Ikeda H, Hiraga M, Shirasawa K, Nishiyama M, Kanahama K, Kanayama Y. Analysis of a tomato introgression line, IL8-3, with increased Brix content. Sci Hort. 2013;153:103–8. https://doi.org/10.1016/j.scienta.2013.02.006.
77. Zhang J, Lyu H, Chen J, Cao X, Du R, Ma L, et al. Releasing a sugar brake generates sweeter tomato without yield penalty. Nature. 2024;635(8039):647–56. https://doi.org/10.1038/s41586-024-08186-2.
78. Prinzenberg AE, van der Schoot H, Visser RGF, Marcelis LFM, Heuvelink E, Schouten HJ. Genetic mapping of the tomato quality traits brix and blossom-end rot under supplemental LED and HPS lighting conditions. Euphytica. 2021;217(12):213. https://doi.org/10.1007/s10681-021-02946-1.
79. Tadmor Y, Fridman E, Gur A, Larkov O, Lastochkin E, Ravid U, et al. Identification of malodorous, a Wild Species Allele Affecting Tomato Aroma That Was Selected against during Domestication. J Agric Food Chem. 2002;50(7):2005–9. https://doi.org/10.1021/jf011237x.
80. Frary A, Nesbitt TC, Frary A, Grandillo S, van der Knaap E, Cong B, et al. fw2.2: A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size. Science. 2000;289(5476):85–8. https://doi.org/10.1126/science.289.5476.85.
81. Chakrabarti M, Zhang N, Sauvage C, Muños S, Blanca J, Cañizares J, et al. A cytochrome P450 regulates a domestication trait in cultivated tomato. Proc Natl Acad Sci. 2013;110(42):17125–30. https://doi.org/10.1073/pnas.1307313110.
82. Reuscher S, Akiyama M, Yasuda T, Makino H, Aoki K, Shibata D, et al. The Sugar Transporter Inventory of Tomato: Genome-Wide Identification and Expression Analysis. Plant Cell Physiol. 2014;55(6):1123–41. https://doi.org/10.1093/pcp/pcu052.
83. Shangguan L, Song C, Leng X, Kayesh E, Sun X, Fang J. Mining and comparison of the genes encoding the key enzymes involved in sugar biosynthesis in apple, grape, and sweet orange. Sci Hort. 2014;165:311–8. https://doi.org/10.1016/j.scienta.2013.11.026.
84. Li N, Wang J, Wang B, Huang S, Hu J, Yang T, et al. Identification of the Carbohydrate and Organic Acid Metabolism Genes Responsible for Brix in Tomato Fruit by Transcriptome and Metabolome Analysis. Front Genet. 2021;12. https://doi.org/10.3389/fgene.2021.714942.

Publisher's note

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