Genetic basis of the lobedness degree in tomato fruit morphology

Dana V. Vazquez, Javier H. Pereira da Costa, Federico.N.I. Godoy, Vladimir Cambiaso, Gustavo R. Rodríguez

PII: S0168-9452(22)00082-6

DOI: https://doi.org/10.1016/j.plantsci.2022.111258

Reference: PSL111258

To appear in: *Plant Science*

Received date: 26 January 2022
Revised date: 8 March 2022
Accepted date: 14 March 2022

Please cite this article as: Dana V. Vazquez, Javier H. Pereira da Costa, Federico.N.I. Godoy, Vladimir Cambiaso and Gustavo R. Rodríguez, Genetic basis of the lobedness degree in tomato fruit morphology, *Plant Science,* (2022) doi:https://doi.org/10.1016/j.plantsci.2022.111258

Title: Genetic basis of the lobedness degree in tomato fruit morphology

Authors: Vazquez, Dana V.[a,b]; Pereira da Costa, Javier H. [a,b]; Godoy, Federico.N.I. [a], Cambiaso, Vladimir[b]; Rodríguez, Gustavo R.[a, b]

ORCID: Vazquez DV (0000-0002-8424-1228), Pereira da Costa JH (0000-0002-9406-6988), Godoy FNI (0000-0001-6902-5511), Cambiaso V (0000-0002-9326-1570), Rodriguez GR (0000-0002-4171-4099)

Author afiliation:

[a]Instituto de Investigaciones en Ciencias Agrarias de Rosario (IICAR-CONICET-UNR). Campo Experimental Villarino, S2125ZAA Zavalla, Santa Fe, Argentina.

[b]Cátedra de Genética, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Campo Experimental Villarino, S2125ZAA Zavalla, Santa Fe, Argentina.

***Corresponding author:** Gustavo Rubén Rodríguez.

Instituto de Investigaciones en Ciencias Agrarias de Rosario (IICAR-CONICET-UNR). Campo Experimental Villarino, S2125ZAA Zavalla, Santa Fe, Argentina. Tel.: +54 - 341 – 4970080 ext. 1162. Fax: +54 - 341 – 4970199. E-mail: rodriguez@iicar-conicet.gob.ar

Abstract

Fruit shape is a key trait in tomato (*Solanum lycopersicum* L.). Since the most studies focused on proximo-distal fruit morphology, we hypothesized that unknown QTLs for medio-lateral direction ones could be found analysing segregating populations where major shape genes are fixed. We examined the diversity of fruit morphology in medio-lateral direction; defined divergent traits in cultivars carrying identical genetic constitution at *LC* and *FAS* genes; and identified QTLs for lobedness degree (LD) by a QTL-seq approach. We found that *LC* and *FAS*

genes were not enough to explain LD variability in a large tomato collection. Then, we derived F_2 populations crossing cultivars divergent for LD where *LC* and *FAS* were fixed (Yellow Stuffer x Heinz 1439 [F_2YSxH] and Voyage x Old Brooks [F_2VxOB]). By QTL-seq we identified a QTL for LD on chromosome 8 in both F_2, which was validated in F_2YSxH by interval mapping accounting for ~ 17% of the variability. Other two QTLs located on chromosomes 6 and 11 with epistasis explained ~ 61% of the variability in the F_2VxOB. In conclusion, three novel QTLs with major effect for LD (*ld6*, *ld8*, and *ld11*) were identified through the study of diversity and genetic segregation in intraspecific tomato crosses.

Keywords: medio-lateral fruit shape; QTL-seq; *Solanum lycopersicum*; bumpiness

1. Introduction

Tomato (*Solanum lycopersicum* L.) displays a great diversity for weight, shapes, and fruit colors compared to its wild ancestor *S. pimpinellifolium* L., which bears red and round fruits with two locules, and weighing only a few grams [1]. Tomato fruit size and shape are important traits, determining yield, quality, and consumer acceptability [2,3]. Medio-lateral features like fruit bumpiness or lobedness degree, locule number, pericarp area, and maximum width are highly associated with shape and fruit weight [4]. Locules are the cavities containing seeds in fruits, and an increase in number determines a more flattened shape leading up to a 50% increase in fruit size [5,6]. Also, positive phenotypic correlation and overlapping of loci underlying lobedness degree and fruit size or mass were previously observed.[7], and it was more frequently observed in flattened shapes [8].

It is well-recognized that fruit shape is quantitatively inherited; however, the great diversity existing in tomato is mainly explained by mutations in a relatively small number of genes [9,10]. The main genes controlling this trait are: *FASCIATED* (*FAS*), *LOCULE NUMBER* (*LC*), *SUN*, *OVATE*, and *SOV1*. The *SUN*, *OVATE*, and the suppressor of *OVATE* (*SOV1*) genes control the occurrence of elongated and pear-shaped fruit [10–13]. On the other hand, *FAS* and *LC* genes act on locule number, fruit size, and flattened shape. The *fas* mutation, located on chromosome 11 is caused by a 294-kb inversion with breakpoints in the first intron of a YABBY transcription factor gene and 1 kb upstream of the tomato *CLAVATA3* (*SlCLV3*) start codon. The inversion disrupts the *SlCLV3* promoter, which results in an increased fruit size and locule number [14,15]. The *lc* mutation, situated on chromosome 2 is related to two single-nucleotide polymorphisms (SNPs) located 1,080 bp downstream of the stop codon of the tomato ortholog of *WUSCHEL* (*WUS*) in a putative CArG box regulatory element. This causes the maintenance of a larger stem cell population resulting in increased locule number and often leads to a flat fruit of a larger size [16,17]. Both *FAS* and *LC* show a synergistic effect on fruit size and weight [8].

Genetic background modulates the effect of major genes underlying fruit shape [9,17], suggesting the presence of unknown genetic modifiers. Furthermore, the effect of these genes on morphological fruit attributes at medio-lateral direction has not been studied so far. Hence, the study of the genetic basis underlying fruit shape and fruit development in the medio-lateral direction represents a novel and important and research area.

Currently, it is possible to identify novel minor and modifier genes that interact with major regulators of tomato fruit shape through characterization of segregating intraspecific

populations using high-throughput SNP technologies [10,13]. Populations that do not segregate for the known fruit shape genes but still vary in morphology are critically important. To map the modifiers, the QTL-seq methodology [18] has been useful to discover new QTLs for important plant traits e.g. seedling vigor, blast resistance, seedling cold resistance, salt tolerance, and grain weight in rice [18–21]; early flowering in cucumber [22]; seed weight, and dry root weight/dry plant weight in chickpea [23]; and fruit weight, elongated shape, locule number, heat tolerance in tomato [13,24,25]. Combining high-throughput SNP genotyping and QTL-seq approaches allowed the rapid and efficient improvement of trait mapping and genomics-assisted breeding.

Based on this background we hypothesized that unknown QTLs for fruit morphology trait at medio-lateral direction could be found on intraspecific populations derived from crosses between cultivars that differ for fruit shape traits and where alleles for *LC* and *FAS* genes were fixed. The aim of this study is to 1) analyse the diversity for medio-lateral fruit morphology traits in the tomato germplasm, 2) define divergent morphological attributes in cultivars carrying identical genetic allelic composition at the *lc* and *fas* loci, 3) identify new QTLs for morphological fruit traits such as lobedness degree by the QTL-seq approach.

2. Materials and methods

2.1. Fruit image dataset from a diverse tomato panel

2.1.1 Tomato fruit images dataset evaluation and medio-lateral direction morphology traits analysis

Tomato fruit images were downloaded from the Solanaceae Genomics Network website (https://solgenomics.net/). The subset image collection was composed of 183 accessions

with at least four fruits each (Supplementary Table 1). The collection has diversity for fruit shape categories, fruit size, germplasm type, and geographic origin. General information and characterization for major shape genes were obtained from Rodríguez et al.[9] . A total of 1,145 medio-lateral sectioned images were analysed for shape attributes with the Tomato Analyzer 3.0 software [26,27]. The analysed traits were: perimeter, area, pericarp area, pericarp thickness and lobedness degree. The pericarp area was calculated as the area within the pericarp boundary and the perimeter whereas the pericarp thickness represents the average length of the pericarp along horizontal and vertical lines through the center of the fruit. Lobedness degree represents the standard deviation of the distances from the center of fruit to the boundary, multiplied by 100. Regular fruits have more similar distances, smaller deviations, and smaller values for this trait, while the opposite situation occurs with irregular fruits. The locule number was counted by the computer vision-based tool LocAnalyzer [28].

2.1.2 Statistical analysis of morphometric fruit traits

Descriptive statistical parameters were estimated for all traits. The normal distribution of the data was verified by graphical methods including histograms, and Q-Q plots. The distribution was analysed with the Shapiro-Wilk test. Traits with non-normal distribution were transformed using the natural logarithm. Broad-sense heritability was calculated by ANOVA for all normal traits [29]. The phenotypic correlations among traits were determined using the Pearson test. The analysis was performed with the R computer program, version 3.6.3 [30] using the basic stats functions and the ggcorrplot [31], agricolae [32], and heritability [33] packages.

A Principal Component Analysis (PCA) was performed to summarize and visualize the multiple inter-correlated quantitative variables. The PCA and plots were performed with the R computer program, version 3.6.3 [30] using basics stats functions and the corrplot [34], factoextra [35], FactoMineR [36], tidyr [37], ggplot2 [38], and Scatterplot3d [39] packages.

The relative frequency distribution was analysed for different allelic compositions at *fas* and *lc* loci, considering locule number and lobedness degree ranges. Five accessions (T1078, T1116, T907, T1697, and T985) with genes at heterozygous state were removed. Ranges for both attributes were created considering: cero to mean value, mean ± one standard deviation, mean ± two standard deviations, mean ± three standard deviations, and higher values.

2.2 Genetic analysis of segregating populations

2.2.1 Plant Material and statistical analysis data

Tomato cultivars (S. *lycopersicum*) named “Voyage” (V), “Old Brooks” (OB), “Yellow Stuffer” (YS), and “Heinz 1439” (H) were selected as progenitors to develop two F_2 populations. All cultivars present semi-indeterminate growth habits, fruits with large size, high locule number, and flattened shape. “Yellow Stuffer” cultivar has yellow mature fruits, while the rest of the cultivars have red fruits. Regarding the known alleles affecting fruit morphology, cultivars, “Voyage” and “Old Brooks” carry the mutant alleles of *LC* and *FAS* ($lc^{-/-}$:$fas^{-/-}$), while the cultivar “Yellow Stuffer” and “Heinz 1439” carry the mutant alleles of *LC* and the wild-type allele of *FAS* ($lc^{-/-}$:$fas^{+/+}$). The four cultivars carry the wild-type alleles for *SUN, OVATE,* and *SOV1* (Supplementary Table 1) genes.

Two independent and intraspecific crosses were carried out: “Voyage” x “Old Brooks”, and “Yellow Stuffer” x “Heinz 1439”. Hybrids were obtained by manual emasculation and pollination. Segregating F_2 populations were obtained by self-fertilization of the hybrids. Five plants of each uniform genotype (parental and hybrids) were grown in a greenhouse with 85 F_2 plants of VxOB population, and 122 plants of YSxH population. Plants were arranged in a completely randomized design, with a distance between plants of 0.35 m and a distance between rows of 1 m. Irrigation, fertilization, and cultural management were carried out according to the standard recommendations for the area, and crop requirements depending on physiological stage, and environmental conditions. The assays were performed at Campo Experimental Villarino, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario (33º S, 61º W).

Seven or eight fruits per plant were harvested at the maximum size stage. Fruits were cut longitudinally at the equatorial plane, placed cut-side down on a scanner, and digitalized at 300 dots per inch as previously described by Brewer et al. [40]. Fruit images were analysed for lobedness degree using the Tomato Analyzer software, version 3.0 [7].

Broad-sense heritability for lobedness degree was calculated by ANOVA in both populations using R, version 3.6.3 [30]. The experimental error was obtained following Mariotti and Collavino [29] procedure.

2.3 DNA isolation and sequencing of DNA-bulks

QTL-seq approach was applied to both segregating populations. High (HLD) and low (LLD) lobedness degree DNA-bulks were generated by mixing an equal amount of DNA from plants with contrasting values for lobedness degree. DNA-bulks included 10 plants from F_2 VxOB

population and 14 plants from F_2 YSxH population. A young leaf was extracted from each F_2 plant at four green leaves-size. Genomic DNA was extracted following the methodologies described by Bernatzky and Tanksley [41] and Fulton et al. [42], with minor modifications. The concentration and integrity of the DNA were evaluated on 1% w/v agarose gels with SYBR® Safe staining (Thermo Fisher Scientific®, Waltham, MA, USA), by comparison against a standard of Phage lambda DNA with a known concentration equal to 25 ng/µl. Also, DNA concentration was quantitatively measured with the QUBIT fluorometer (Thermo Fisher Scientific®, Waltham, MA, USA). Finally, the samples were equalized to a 50 ng/µl final concentration. DNA-bulk quality was also evaluated on 1% w/v agarose gels stained with SYBR® Safe. DNA-bulks samples were sequenced at Macrogen sequencing service (Macrogen, South Korea) in NovaSeq 6000 equipment to obtain 151 bp paired-end reads.

2.4 Data processing and SNPs identification

Raw sequence data were filtered using the Trim Galore program, version 0.4.5 (https://github.com/FelixKrueger/TrimGalore) to remove the Illumina adapter sequences, and reads with poor quality. The trimmed reads were assessed for quality using FastQC, version 0.11.4 [43]. Short reads from the two DNA-bulks were aligned against the tomato consensus genome sequence (*S. lycopersicum* L. cultivar Heinz 1706, reference assembly version SL4.0) [44] using Bowtie 2,version 2.3.2 [45], with "--very-sensitive-local" option. The output SAM files were sorted by coordinates, classified, labeled and converted into a BAM file format using Picard tools, version 1.119 (http://broadinstitute.github.io/picard/). Duplicate readings were identified and labeled with the same program. The resulting BAM files with the aligned sequence information were analysed using the Qualimap program, version 2.2.1 [46]. SNPs and InDels (Insertion–Deletion) calling were performed for each bulk

using the HaplotypeCaller tool from GATK, version 4.0.9.0 [47,48]. The polymorphism information derived from each DNA-bulk was combined into a single file using the CombineGVCFs tool. Genomic variants data (InDels and SNPs) between bulks sequences were obtained with the GenotypeGVCFs tool and exported as VCF files. SNPs were filtered using "--select-type-to-include" option in SelectVariants tool. Low-reliable polymorphisms were filtered out with the "--filter-expression" option included in VariantFiltration tool from the GATK program, version 4.0.9.0 [47,48]. The "--remove-filter-all" option from VCFtools program [49], version 0.1.15, was used to select by quality the variants for further analysis. This information was exported in TABLE format using the VariantsToTable tool from the GATK program, version 4.0.9.0 [47,48]. The alignment and comparison of the genomic sequences were carried out using the High-Performance Computing Center facility at Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)-Rosario.

2.5 Detection of genomic regions associated with lobedness degree by QTL-*seq* analysis

The identification of QTLs was performed with the QTLseqr software [50] package version 0.7.4, developed in R computing language [30]. The reference allele frequency, SNP-index, and the Δ (SNP-index) were obtained for each SNP using the importFromGATK function. Low confidence SNPs were filtered out according to Takagi et al. [18] using the filterSNP option. SNPs with read depth between first to third quartile values, the absolute difference between bulks below 50x, and quality equal to 99 were analysed.

Statistical significance of QTLs was evaluated according to Magwene et al. [51]. A tricube-smoothed of G parameter (G') was calculated for each SNP. Also, the p-value [51] and Benjamini-Hochberg adjusted p-values or Q-value [52] were computed. G' values higher than

the significance threshold suggest a QTLs presence. The primary analysis steps were made using the runGprimeanalysis function. A 2 Mb sliding window was considered. The false discovery rate (FDR) was defined considering a threshold Q-value of 0.05 and 0.01, to determine those sites that deviated significantly from the null distribution of G'. Results were plotted with the plotQTLStat function. Putative QTLs information was exported as a comma-delimited file (CSV) considering the options: method = "Gprime", alpha = 0.05 and export = TRUE of GetQTLTable function.

2.6 Lobedness degree QTLs validation with molecular markers

InDels markers were developed based on InDels-calling data for both populations. The online interface Primer3, version 0.4.0 [53] was used for designed primers flanking InDels ranging from 15 bp up to 50 bp. Polymorphisms with read depths (DP) greater than 10x in both bulks, genotype quality (GQ) equal to 99, and differential genotypes between the segregating groups were considered. According to Takagi et al. [18], polymorphisms with high values of Δ (SNP-index) in the region of interest and near to zero values outside the region were selected.

The F_2 VxOB population was characterized with a total of 22 specific InDels markers designed along the different loci: two markers on chromosome 4, two markers on chromosome 6, 14 on chromosome 8, two on chromosome 9, and two on chromosome 11. The locus *ld8* in the F_2 YSxH population was characterized with 19 InDels markers designed, approximately every 5 Mb, covering the entire chromosome 8. Details on the molecular markers used are summarized in Supplementary Table 3. Specific polymorphic fragments were amplified by

PCR. Electrophoresis was conducted on agarose gels at 3% w/v stained with SYBR® Safe (Thermo Fisher Scientific®, Waltham, MA, USA) for visualization.

Lobedness degree QTLs validation in F_2 VxOB population was performed by single-point one-way ANOVA [54]. Two markers were analysed for each locus (Supplementary Table 3). We verified that the independence, normality, and homoscedasticity of residuals assumptions for ANOVA test were fulfilled. The free software environment R, version 3.6.3 [30] was used to perform a linkage analysis of markers.

Due to a shape locus was found on chromosome 8 in both populations, a genetic linkage map was constructed and a simple interval mapping (SIM) [55] was conducted with R/QTL package [56]. Plants with missing data for half of the total markers or higher were not considered. Markers with distorted segregation or missing data greater than 10% were removed. The minimum logarithm of the odds (LOD) score and recombination frequency were defined as 5 and 0.35, respectively. The “orderMarkers” and “ripple” functions were used to determine the order of the markers in the linkage groups and the distance between markers was calculated with the Kosambi function [57]. QTL detection was carried out based on genotypic and phenotypic data from 72 F_2 individuals in F_2 VxOB population, and 120 F_2 individuals in F_2 YSxH population. Genotypic data were simulated at a maximum distance of two centiMorgan (cM) with the function "sim.geno". The LOD threshold for each character was calculated by performing 1000 permutations with the “scanone” function. Significant peaks or QTL were defined with the “define.peak” function when LOD was greater than 3, which correspond with a p-value between 0.01 and 0.05. The R square value and significance were calculated for all QTL with the “calc.Rsq” function by ANOVA.

Epistatic interactions among genomic regions associated with lobedness degree on chromosomes 6 and 11 were evaluated by performing a two-way ANOVA. The most significant molecular marker per region was chosen. Statistical analyses were carried out with software environment R, version 3.6.3 [30]. Genetic interactions intralocus and interloci were assed applying orthogonal contrast following the Jana [58] procedure with InfoStat software, version 2017 [59].

3. Results

3.1 Diversity of medio-lateral direction fruit shape traits present in a fruit image dataset The tomato germplasm dataset was diverse and representative for fruit shape categories, fruit size, germplasm class, and geographic origin (Supplementary Table 4). The diversity for fruit morphology in the medio-lateral direction is represented in Fig. 1A.

The relatively high values of standard deviation and the interquartile range for some traits suggest a high level of shape diversity (Table 1). Remarkably, 75% of varieties featured a locule number between two and three, and lobedness degree less than or equal to 2.2. Nevertheless, some varieties showed a high locule number and highly irregular shapes (Fig. 1A, Supplementary Table 1) with as much as 17.00 locules (Coure di Bue), and maximum lobedness degree of 9.12 (Zapotec Pink Ribbed) (Table 1, Supplementary Table 2). Pericarp thickness and perimeter had a normal distribution, while locule number, area, lobedness degree, and pericarp area showed a non-normal distribution (Fig. 1B). The broad-sense heritability values were high and significant for all traits ranging from 0.72 for lobedness degree up to 0.94 for the area (Table 1).

Significant phenotypic correlations were found for most of the analysed traits, considering a p-value of 0.05 (Fig. 1C). As expected, there was a very high and positive correlation between perimeter and area. Also, perimeter, area presented a positive intermediate correlation considering the locule number and pericarp thickness. Pericarp thickness showed a positive correlation with the pericarp area, and to a lesser grade with lobedness degree. Perimeter, area, and locule number presented a positive correlation with lobedness degree. A negative correlation was found between pericarp area and locule number. Non-significative correlations were found between the locule number and pericarp thickness, and among the pericarp area and the lobedness degree (Fig. 1C).

The principal component analysis (PCA) showed that the three PCs represented 55.11%, 28.43%, and 12.33% of the total variance, respectively. The first component was explained primarily by perimeter, area, and in a lesser grade locule number; the second was explained most by pericarp area, and pericarp thickness; meanwhile, the third component was explained almost exclusively by lobedness degree (Fig. 2A). Variation in the medio-lateral traits was correlated with the *LC* and *FAS* alleles, according to the PCA results (Fig. 2A). Accessions carrying the wild-type alleles, $lc^{+/+}$: $fas^{+/+}$, were mainly characterized by small fruit with fewer locules, and a more regular shape (low values for first and third components), and intermediate to high pericarp area values (second component). Most accessions carrying the mutant alleles $lc^{-/-}$: $fas^{-/-}$ carried larger fruits with more locules and bumpy shape (higher values for first and third components) and lowest for pericarp area (second component). However, these accessions also showed a greater spread in the third dimension, indicating more variability for lobedness degree.

We found that the combination of *lc*$^{-/-}$: *fas*$^{-/-}$ showed the highest locule number. On the other hand, lobedness degree was less influenced by the mutant alleles *lc*$^{-/-}$: *fas*$^{-/-}$(Fig. 2B). Thus all genotypes carrying *lc*$^{+/+}$: *fas*$^{+/+}$ composition had a low locule number (≤3), and genotypes carrying *lc*$^{-/-}$: *fas*$^{-/-}$ alleles had fruit with many locules (≥5). For lobedness degree, the *lc*$^{+/+}$: *fas*$^{+/+}$ combination showed a lobedness degree of less than 4.5; while for most of the accession with *lc*$^{-/-}$: *fas*$^{-/-}$ the lobedness degree was higher than 3.6. Note, however, the *lc*$^{+/+}$: *fas*$^{-/-}$ composition was present only in two accessions in the set (Supplementary Table 1).

The accessions distribution considering the locule number and lobedness degree are represented in Fig. 2C. Most genotypes showed values in the range of the mean ± 1 standard deviation for both traits (between 0.91 and 5.35 for locule number and 0.60 and 2.92 for lobedness degree). Some accessions with different compositions of alleles showed high values for lobedness degree and locule number (e.g. cultivars “Cour di Bue”, “Yellow Stuffer”, “Zapotec Pink Ribbed”, “Voyage”, and “LA2845”). Other accessions with different combinations of alleles showed low extreme values of lobedness degree and high locule number (e.g. cultivars “Heinz 1439”, “T546”, “LA1216”, “Person”, and “Old Brooks”). In summary, the characterization of the medio-lateral morphological diversity suggests that populations derived from crosses between accessions with high locule number that carried the same alleles at *LC* and *FAS* can be used to map lobedness degree. The underlying genes can be considered likely modifiers of these genes.

3.2 Genomic regions underlying lobedness degree in tomato

3.2.1 Mapping population for lobedness degree

Two independent F_2 populations were developed from crosses between accessions carrying fruits with high locule number, flat shape, and large size, but with contrasting lobedness degree values. In the first population, “Voyage” showed a bumpy shape whereas “Old Brooks” had a regular external shape. In the other population “Yellow Stuffer” showed a more irregular shape than “Heinz 1439” (Table 2).

The frequency histogram for lobedness degree showed a continuous variation (p-value <0.0001) in both F_2 populations (Fig. 3). The average lobedness degree was equal to 2.62, ranging between 1.10 and 6.77 for VxOB population. In YSxH population, the average value was 1.60, with a range between 0.74 and 3.11 (Supplementary Table 5). Mean values of F_2 plants were skewed towards the smoother shaped parent in both populations. Also, lobedness degree values were more extreme in F_2 VxOB. Broad-sense heritability was significant in both populations with values of 0.32 in VxOB population and 0.20 for YSxH population (Supplementary Table 5).

3.2.2 Genomic regions controlling lobedness degree detected by QTL-seq

A total of 215,875 and 234,153 SNPs were identified between the bulks along 12 tomato chromosomes in populations VxOB and YSxH, respectively. After variant calling, genomic comparison, and quality filtering we kept a total of 76,727 SNPs in population VxOB (35.54%), and 84,364 SNPs in population YSxH (36.03%). The graphs in Fig. 4A (VxOB population) and Fig. 4B (YSxH population) represent the average G’ value across the tomato genome assembly. The QTLs data detail is summarized in Table 3.

According to G’ approach, we identified five genomic regions with G’ values higher than 5 in the VxOB population, located on chromosomes 4, 6, 8, 9, and 11 (Fig. 4A, Table 3). Note, the

region on chromosome 9 was the largest, comprising almost the entire chromosome, meanwhile the more significant G' peak was found in chromosome 4. Considering the YSxH population, five genomic regions associated with lobedness degree were identified across chromosome 8 (Fig. 4B, Table 3). The regions were mapped in the centromeric area, so it is likely that some SNPs were false positives and some regions were part of the same locus. Further marker analysis is necessary to validate these regions. It was remarkable that more than half of SNPs (69.6%) were detected in the last locus, since the first region only contained 3 SNPs, and the maximum value of G' was reached in the third locus.

Since we found a region on chromosome 8 associated with lobedness degree in both populations, further we focus on developing PCR-based markers in this region to validate and define a conserved genomic region underlying this trait across the tomato germplasm.

3.2.3 Lobedness degree QTLs validation with molecular markers

Eight markers on chromosome 8 were genotyped in VxOB population. These markers adjusted to codominant segregation (Fig. 5A). The genetic and physical order of markers was correlated. The length of the map was ~98 cM, with an average spacing between markers of 13.9 cM and maximum spacing of 34.5 cM. Considering the SIM approach, a maximum LOD value equal to 2.2 was found for lobedness degree at PTZ-51 marker (0.96 Mb) (Fig. 5B). The LOD value was significant at 5% but below the defined significance threshold. Therefore, the QTL on chromosome 8 was not validated in this population, which was likely due to few plants in the F_2 population (85 individuals). Moreover, the lobedness degree values for alleles of PTZ-51 marker (Fig. 5C) showed the opposite effect of the expected phenotype according to the parental fruit morphology.

For YSxH population, seven markers amplified polymorphic fragments according to expected molecular weight. These segregated as codominant markers. It was possible to obtain a genetic linkage map of chromosome 8 (Fig. 5D). The total length of the map was ~79 cM, with an average distance between markers of 11.4 cM, and a maximum spacing of 28.7 cM, in the centromeric region, between PTZ-52 and PTZ-64 (Fig. 5D). A single QTL was significantly associated with lobedness degree, *ld8* (Fig. 5E). This presented a maximum LOD value equal to 5.71 located at 22 cM. The significance interval of *ld8* was 24 cM length, from 10 cM up to 34 cM. However, the QTL was located in the centromeric region of the chromosome. The locus *ld8* explained 17% of the phenotypic variation observed for the lobedness degree in F_2 YSxH population. Also, the linked markers showed an additive gene action and the Yellow Stuffer alleles at this QTL increased lobedness degree, in a recessive manner (Fig. 5F).

A total of eight DNA markers were developed for single-point analysis in the VxOB population, two for each one of four chromosomes (4, 6, 8, 9 and 11) (Table 4, Supplementary Table 2). Three major QTLs (%R^2 equal or higher than 15%) located on chromosomes 6 (PTZ-12 or *ld6*, and PTZ-17), and 11 (PTZ-103 or ld11), were validated for lobedness degree (Table 4). However, it was confirmed by simple interval mapping that both QTLs on chromosome 6 were part of the same loci (unpublished results). The *ld11* locus explained the 46% of the observed phenotypic variance for lobedness degree, while *ld6* represented by the PTZ-12 and PTZ-17 markers, explained the 23% and 15% of the variance, respectively. Also, PTZ-103 was associated with fruit weight, locule number, perimeter, and area. PTZ-12 has associated additionally with the locule number.

Epistasis or interaction between *ld6* and *ld11* was significant ($p<0.05$). Together, both markers accounted for 61% of the phenotypic variability for lobedness degree. These two loci have a synergistic effect over fruit lobedness degree, and the presence of "Voyage" alleles leads to an increase in bumpiness (Fig. 6). Due to the highest significance of *ld11* with respect to *ld6* the change of an "Old Brooks" allele for a "Voyage" allele at *ld6* do not generate a great change in the mean value of that trait, and only individuals with the "Voyage" alleles in both loci ($ld6^{-/-}$:$ld11^{-/-}$) showed significant differences (Fig. 6B).

Dominant intralocus interactions were detected for both, *ld6* (p-value = 0.05) and *ld6* (p-value = 0.04), and a highly significant additive effect was found only for *ld11* (p-value<0.0001). Although no significant interloci interaction was detected applying orthogonal contrast, the F value of additive by additive interaction was 2.80 (p = 0.09). These results support that a stronger effect is caused by the QTL detected on chromosome 11, while the QTL on chromosome 6 acts as a modifier with a slight effect.

Altogether these results indicate that different alleles were present in both populations. Thus, a QTL was found by QTL-seq approach, and validated by SIM, in the centromeric region of chromosome 8 (*ld8*) in F_2 YSxH population; whereas two other QTLs, located in chromosomes 6 (*ld6*) and 11 (*ld11*), were identified and validated by single-point analysis in F_2 VxOB population. Despite significant interaction being found between *ld6* and *ld11* by ANOVA, no significant interloci interactions were detected by orthogonal contrast analysis. This suggests the genetic determinants of lobedness degree are found at chromosome 11, while *ld6* acts as a modifier with a slight effect.

4. Discussion

Fruit shape in the medio-lateral direction is a key trait influencing the quality and consumer acceptability in tomato (*Solanum lycopersicum* L.), highly related to the size and fruit weight [4]. In tomato, most of the fruit morphological studies have focused on attributes at the proximal-distal direction, therefore the effect of major genes on features like lobedness degree remains unclear. We found significant H^2 values for all fruit shape attributes in the medio-lateral direction, demonstrating a broad genetic diversity is present in tomato germplasm. Fruit shape diversity is explained by a small number of genes [9] including *LC* and *FAS* that are controlling shape features in medio-lateral direction, as fasciated fruits and locule number [8]. However, our results showed the effect of these genes was not enough to explain the variation for lobedness degree. This suggests the presence of unknown QTLs or modifier genes for this trait.

Since lobedness degree is an understudied trait that influences consumer preference, we decided to focus on this characteristic. So, two intraspecific populations were developed from independent crosses between tomato cultivars with contrasting values for lobedness degree and the same allelic composition for main shape genes (Voyage x Old Brooks [F_2VxOB], and Yellow Stuffer x Heinz 1439 [F_2YSxH]). By QTL-seq approach, a region on chromosome 8 was associated with lobedness degree in both populations. This region was validated in the F_2 YSxH population, where *ld8* acts like a major QTL for lobedness degree. This result is consistent with van der Knaap and Tanksley [60], who found two QTLs on top of chromosome 8 linked to fruit bumpiness, and postulated that these QTLs and *fs8.1* are the outcomes of pleiotropic actions of the same gene. In spite of this, *ld8* is a new QTL, since both parents are carrying the wild-type alleles at *fs8.1*, so the locus is not segregating in this population (van der Knaap, personal communication). On the other hand, no marker on

chromosome 8 was significantly associated with lobedness degree in the F_2 VxOB population. Only the PTZ-51 marker reached a LOD value significant at 5%, but it did not surpass the LOD threshold, and the increase in lobedness degree controlled by PTZ-51 was attributable to the Heinz 1439 allele, suggesting the presence of genes acting as negative regulators in the region. Altogether, these results indicate that different alleles were present in both populations.

Additional QTLs with major effect for lobedness degree were found in the F_2 VxOB population by QTL-seq, and they were validated by single-point analysis. The *ld6* (PTZ-12) and *ld11* (PTZ-103) loci showed epistasis, and together explained more than 60% of the phenotypic variance. Significant epistatic interactions between shape traits have been found in many previous studies. Among them, we could name epistatic interaction between *sun*, *ovate*, and *fs8.1* [61], *lc* and *fas* [5,6], or *fs2.1* and *fs8.1* [62]. The *ld11* had a higher significance than *ld6,* and only individuals with the "Voyage" alleles in both loci ($ld6^{-/-}$:$ld11^{-/-}$) presented significant differences. This suggests that the genetic determinants of lobedness degree are present at the bottom of chromosome 11, and the stronger effect of *ld11* could be masking the effect *ld6*, explaining that no significant interloci interactions were found by orthogonal contrasting.

FAS is a partial loss of function caused by an inversion that disrupts the promoter of tomato *CLV3* (*SlCLV3*) [14,63], resulting in increase of locule number, and thus fruit size. *FAS* maps close to *ld11*, which was located at 53.3 Mb. Here we demonstrate that locule number and lobedness degree are correlated traits, which could suggest a pleiotropic effect of *ld11* allele on both traits. This agrees with publications that described the *FAS* gene increase the locule number and fasciated shape [8]. A more recent study demonstrated that the genome edition

of the cis-regulatory element of *SlCLV3* by CRISPR/Cas9 created new alleles displaying a continuum of variation for locule number in tomato fruits [64]. Previous reviews also have reported the existence of mutant alleles of *FAS* has been associated with unfused carpels, a phenotype exclusive of "Voyage" cultivar [65]. In light of these considerations, we postulated *ld11* as a new allele of the *FAS* gene. However, all these studies focus mainly on locule number, inflorescence branching, and fruit size, so the characterization of their impact on lobedness degree constitutes a novel perspective. Additionally, *CELL SIZE REGULATOR* (*CSR*, *Solyc11g071940*) gene, which underlies the *fw11.3* locus, was fine mapped in this region and increase the fruit weight through enlargement of the pericarp areas by increasing cell size [66]. However, there is no evidence that this gene affects fruit shape [67]. Two minor QTLs have been mapped in the centromeric region of chromosome 6, one for elongated fruit shape and the other for fruit size [60]. Moreover, a minor QTL for locule number was identified in the bottom of chromosome 6, and it was postulated as a modifier of mutation *FAS* [24]. However, these regions are not known to have major effects on lobedness degree or fruit shape. Additional experiments will be necessary to confirm the effect of *ld11* and *ld6*.

The QTLs identification accuracy largely depends on population size, and choosing a population size is a compromise between theoretically desirable and feasibly practiced. The size population will depend on the type of inheritance of trait, genetic background, type of mapping population, QTL mapping approach, etc. In tomato, some QTLs with major effects were mapped in small populations and for most of them the gene controlling the studied trait was identified by positional cloning [11,13,44,68–70]. In this context, although both F_2 populations consisted of around 100 individuals, this could be considered populations of

small size, as we expected lobedness degree had a monogenic or oligogenic inheritance, this size is valid for QTLs mapping.

Another restriction for QTL detection is given when an allele is recessive. Thus, regular shape alleles are dominant over irregular shape alleles, and heterozygotes individuals have been included in wild type-like bulk. In this study, the alleles of Heinz 1439 in *ld8* (Fig. 5) and the alleles of Old Brooks in both *ld6* and *ld11* (Fig. 6) are dominant over the alleles of "Yellow Stuffer" and "Voyage", respectively. So, allelic frequency in the LLD-bulk was lower than the average, and the difference between allelic variant frequencies in the two bulks was lower. Although, this could reduce the feasibility to detect variants underlying lobedness degree by QTL-seq approach [71] the major effect of these QTLs can solve this restriction.

5. Conclusion

We examine the fruit morphology diversity in the medio-lateral direction in a tomato germplasm collection. The great variability present for these traits, and particularly for lobedness degree, was not explained by the genetic constitution at *LC* and *FAS*, which are the two known genes involved in fruit shape in the medio-lateral direction. Despite lobedness degree being an economical importance trait related to locule number and fruit size, it has never been mapped yet. By QTL-seq approach applied in two intraspecific populations where alleles for *LC* and *FAS* genes were fixed, we provide valuable information about novel QTLs for lobedness degree, on chromosomes 6, 8, and 11 (*ld6*, *ld8*, and *ld11*, respectively). All QTLs showed a major effect on traits, but different QTLs were present in the analysed populations. The *ld6* and *ld11* loci are epistatic, and *ld11* has a stronger effect while *ld6* seems to act as a modifier on lobedness degree. Here we bring light to the genetic bases of shape in medio-lateral direction and presented an original study that mapped the

lobedness degree trait. Further investigations about these QTLs especially in terms of QTL introgression into elite background may be useful to assist future breeding targeted for improvement of fruit shape and obtain more uniform fruits or fruit with atypical shapes destined to special market niches.

Acknowledgements

We would like to thanks Esther van der Knaap of University of Georgia for comments and valuable suggestions on the manuscript.

Funding

Research was supported by the Agencia Nacional de Promoción Científica y Tecnológica (FONCyT PICT 2018-00824), Consejo Nacional de Investigaciones Científicas y Técnicas (PUE0043), and Universidad Nacional de Rosario (80020190300004UR) to GR Rodríguez.

CRediT authorship contribution statement

GRR conceived the study and supervised the research. DVV, FNIG, and GRR performed experiments. DVV, JPdC, and VC contributed to data analysis and interpretation. DVV prepared the Fig.s and tables. DVV and GRR wrote the manuscript. All authors reviewed and agreed to the published version of the manuscript.

Declarations

Declaration of Competing Interest

The authors declare they have no conflict of interest

Human and animal rights

This study does not include human or animal subjects.

Data Availability

The sequence data generated in this study have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the accession number SRP354680.

References

[1] I. Paran, E. van der Knaap, Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper, J. Exp. Bot. 58 (2007) 3841–3852. https://doi.org/10.1093/jxb/erm257.

[2] S. Grandillo, H.M. Ku, S.D. Tanksley, Identifying the loci responsible for natural variation in fruit size and shape in tomato, Theor. Appl. Genet. 99 (1999) 978–987. https://doi.org/10.1007/s001220051405.

[3] J. Casals, A. Rivera, J. Sabaté, R.R. del Castillo, J. Simó, Cherry and fresh market tomatoes: Differences in chemical, morphological, and sensory traits and their implications for consumer acceptance, Agronomy. 9 (2019). https://doi.org/10.3390/agronomy9010009.

[4] A.N. Nankar, D. Kostova, I. Tringovska, S. Grozeva, D. Ganeva, Tomato phenotypic diversity determined by combined approaches of conventional and high-throughput tomato analyzer phenotyping, Plants. 9 (2020). https://doi.org/10.3390/plants9020197.

[5] Z. Lippman, S.D. Tanksley, Dissecting the Genetic Pathway to Extreme Fruit Size in Tomato Using a Cross Between the Small-Fruited Wild Species Lycopersicon pimpinellifolium and, (2001).

[6] L.S. Barrero, S.D. Tanksley, Evaluating the genetic basis of multiple-locule fruit in a broad cross section of tomato cultivars, Theor. Appl. Genet. 109 (2004) 669–679. https://doi.org/10.1007/s00122-004-1676-y.

[7] M.J. Gonzalo, M.T. Brewer, C. Anderson, D. Sullivan, S. Gray, E. Van Der Knaap, Tomato fruit shape analysis using morphometric and morphology attributes implemented in tomato analyzer software program, J. Am. Soc. Hortic. Sci. 134 (2009) 77–87. https://doi.org/10.21273/jashs.134.1.77.

[8] Y.H. Chu, J.C. Jang, Z. Huang, E. van der Knaap, Tomato locule number and fruit size controlled by natural alleles of lc and fas, Plant Direct. 3 (2019) 1–20. https://doi.org/10.1002/pld3.142.

[9] G.R. Rodríguez, S. Muños, C. Anderson, S.-C. Sim, A. Michel, M. Causse, B.B.M. Gardener, D. Francis, E. van der Knaap, Distribution of SUN, OVATE, LC, and FAS in the Tomato Germplasm and the Relationship to Fruit Shape Diversity, Plant Physiol. 156 (2011) 275–285. https://doi.org/10.1104/pp.110.167577.

[10] S. Wu, B. Zhang, N. Keyhaninejad, G.R. Rodríguez, H.J. Kim, M. Chakrabarti, E. Illa-Berenguer, N.K. Taitano, M.J. Gonzalo, A. Díaz, Y. Pan, C.P. Leisner, D. Halterman, C.R. Buell, Y. Weng, S.H. Jansky, H. van Eck, J. Willemsen, A.J. Monforte, T. Meulia, E. van der Knaap, A common genetic mechanism underlies morphological diversity in fruits and other plant organs, Nat. Commun. 9 (2018) 1–12. https://doi.org/10.1038/s41467-018-07216-8.

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

13306. https://doi.org/10.1073/pnas.162485999.

[12] H. Xiao, N. Jiang, E. Schaffner, E.J. Stockinger, E. Van Der Knaap, A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit, Science (80-.). 319 (2008) 1527–1530. https://doi.org/10.1126/science.1153040.

[13] G.R. Rodríguez, H.J. Kim, E. Van Der Knaap, Mapping of two suppressors of OVATE (sov) loci in tomato, Heredity (Edinb). 111 (2013) 256–264. https://doi.org/10.1038/hdy.2013.45.

[14] C. Xu, K.L. Liberatore, C.A. Macalister, Z. Huang, Y.H. Chu, K. Jiang, C. Brooks, M. Ogawa-Ohnishi, G. Xiong, M. Pauly, J. Van Eck, Y. Matsubayashi, E. Van Der Knaap, Z.B. Lippman, A cascade of arabinosyltransferases controls shoot meristem size in tomato, Nat. Genet. 47 (2015) 784–792. https://doi.org/10.1038/ng.3309.

[15] Z. Huang, J. Van Houten, G. Gonzalez, H. Xiao, E. van der Knaap, Genome-wide identification , phylogeny and expression analysis of SUN , OFP and YABBY gene family in tomato, (2013) 111–129. https://doi.org/10.1007/s00438-013-0733-0.

[16] S. Muños, N. Ranc, E. Botton, A. Berard, S. Rolland, P. Duffe, Y. Carretero, M.-C. Le Paslier, C. Delalande, M. Bouzayen, D. Brunel, M. Causse, Increase in Tomato Locule Number Is Controlled by Two Single-Nucleotide Polymorphisms Located Near WUSCHEL, Plant Physiol. 156 (2011) 2244–2254. https://doi.org/10.1104/pp.111.173997.

[17] E. van der Knaap, M. Chakrabarti, Y.H. Chu, J.P. Clevenger, E. Illa-Berenguer, Z. Huang, N. Keyhaninejad, Q. Mu, L. Sun, Y. Wang, S. Wu, What lies beyond the eye: The molecular mechanisms regulating tomato fruit weight and shape, Front. Plant Sci. 5 (2014) 1–14. https://doi.org/10.3389/fpls.2014.00227.

[18] H. Takagi, A. Abe, K. Yoshida, S. Kosugi, S. Natsume, C. Mitsuoka, A. Uemura, H. Utsushi, M. Tamiru, S. Takuno, H. Innan, L.M. Cano, S. Kamoun, R. Terauchi, QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations, Plant J. 74 (2013) 174–183. https://doi.org/10.1111/tpj.12105.

[19] Z. Yang, D. Huang, W. Tang, Y. Zheng, K. Liang, A.J. Cutler, W. Wu, Mapping of Quantitative Trait Loci Underlying Cold Tolerance in Rice Seedlings via High-Throughput Sequencing of Pooled Extremes, PLoS One. 8 (2013). https://doi.org/10.1371/journal.pone.0068433.

[20] R. Bommisetty, N. Chakravartty, R. Bodanapu, J.B. Naik, S.K. Panda, S.P. Lekkala, K. Lalam, G. Thomas, S.J. Mallikarjuna, G.R. Eswar, G.M. Kadambari, S.N. Bollineni, K. Issa, S. Akkareddy, C. Srilakshmi, K. Hariprasadreddy, P. Rameshbabu, P. Sudhakar, S. Gupta, V.B.R. Lachagari, L.R. Vemireddy, Discovery of genomic regions and candidate genes for grain weight employing next generation sequencing based QTL-seq approach in rice (Oryza sativa L.), Mol. Biol. Rep. 47 (2020) 8615–8627. https://doi.org/10.1007/s11033-020-05904-7.

[21] L. Lei, H. Zheng, Y. Bi, L. Yang, H. Liu, J. Wang, J. Sun, H. Zhao, X. Li, J. Li, Y. Lai, D. Zou, Identification of a Major QTL and Candidate Gene Analysis of Salt Tolerance at the Bud Burst Stage in Rice (Oryza sativa L.) Using QTL-Seq and RNA-Seq, Rice. 13 (2020) 1–14. https://doi.org/10.1186/s12284-020-00416-1.

[22] H. Lu, T. Lin, J. Klein, S. Wang, J. Qi, Q. Zhou, J. Sun, Z. Zhang, Y. Weng, S. Huang, QTL-seq identifies an early flowering QTL located near Flowering Locus T in cucumber, Theor. Appl. Genet. 127 (2014) 1491–1499. https://doi.org/10.1007/s00122-014-2313-z.

[23] V.K. Singh, A.W. Khan, D. Jaganathan, M. Thudi, M. Roorkiwal, H. Takagi, V. Garg, V. Kumar, A. Chitikineni, P.M. Gaur, T. Sutton, R. Terauchi, R.K. Varshney, QTL-seq for rapid identification of candidate genes for 100-seed weight and root/total plant dry weight ratio under rainfed conditions in chickpea, Plant Biotechnol. J. 14 (2016) 2110–2119. https://doi.org/10.1111/pbi.12567.

[24] E. Illa-Berenguer, J. Van Houten, Z. Huang, E. van der Knaap, Rapid and reliable identification of tomato fruit weight and locule number loci by QTL-seq, Theor. Appl. Genet. 128 (2015) 1329–1342. https://doi.org/10.1007/s00122-015-2509-x.

[25] J. Wen, F. Jiang, Y. Weng, M. Sun, X. Shi, Y. Zhou, L. Yu, Z. Wu, Identification of heat-tolerance QTLs and high-temperature stress-responsive genes through conventional QTL mapping , QTL- seq and RNA-seq in tomato, BMC Plant Biol. 19 (2019) 398.

[26] G.R. Rodríguez, J.B. Moyseenko, M.D. Robbins, N. Huarachi Morejón, D.M. Francis, E. van der Knaap, Tomato Analyzer: A Useful Software Application to Collect Accurate and Detailed Morphological and Colorimetric Data from Two-dimensional Objects, J. Vis. Exp. (2010) 1–12. https://doi.org/10.3791/1856.

[27] G.R. Rodríguez, D.M. Francis, E. van der Knaap, J. Strecker, I. Njanji, J. Thomas, A. Jack, New features and many Improvements to analyze morphology and color of digitalized plant organs are available in Tomato Analyzer 3.0, Proc. Twenty-Second Midwest Artif. Intell. Cogn. Sci. Conf. Vol-710 (2011) 160–163. http://ceur-ws.org/.

[28] F.E. Spetale, J. Murillo, D. V. Vazquez, P. Cacchiarelli, G.R. Rodríguez, E. Tapia, LocAnalyzer: A computer vision method to count locules in tomato fruits, Comput. Electron. Agric. 173 (2020) 105382. https://doi.org/10.1016/j.compag.2020.105382.

[29] J.A. Mariotti, N.G. Collavino, Los caracteres cuantitativos en la mejora genética de los cultivos, Orientación Gráfica Editora, Buenos Aires, 2014.

[30] R Core Team, R: A language and environment for statistical computing, (2014). http://www.r-project.org/.

[31] A. Kassambara, ggcorrplot: Visualization of a correlation matrix using “ggplot2”. R package version 0.1. 3, (2019).

[32] F. de Mendiburu, Statistical procedures for agricultural research. Package ‘Agricolae,’’ version 1.4-4.,’ (2013).

[33] W. Kruijer, M.P. Boer, M. Malosetti, P.J. Flood, B. Engel, R. Kooke, J.J.B. Keurentjes, F.A. Van Eeuwijk, Marker-based estimation of heritability in immortal populations, Genetics. 199 (2014) 379–398. https://doi.org/10.1534/genetics.114.167916.

[34] T. Wei, V. Simko, R package "corrplot": Visualization of a Correlation Matrix. Version 0.84., (2017).

[35] A. Kassambara, F. Mundt, Package 'factoextra'. Extract and visualize the results of multivariate data analyses, (2017).

[36] S. Lê, J. Josse, F. Husson, FactoMineR: An R Package for Multivariate Analysis, J. Stat. Softw. 25 (2008) 1–18. https://doi.org/10.18637/jss.v025.i01.

[37] H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the Tidyverse, J. Open Source Softw. 4 (2019) 1686. https://doi.org/10.21105/joss.01686.

[38] H. Wickham, ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009. http://link.springer.com/10.1007/978-0-387-98141-3.

[39] U. Ligges, M. Maechler, S. Schnackenberg, The scatterplot3d Package, Recover. from Https//Cran. Rproject. Org/Web/Packages/Scatterplot3d/Scatterplot3d. Pdf. (2018).

[40] M.T. Brewer, L. Lang, K. Fujimura, N. Dujmovic, S. Gray, E. van der Knaap, Development of a Controlled Vocabulary and Software Application to Analyze Fruit Shape Variation in Tomato and Other Plant Species, Plant Physiol. 141 (2006) 15–25. https://doi.org/10.1104/pp.106.077867.

[41] R. Bernatzky, S.D. Tanksley, Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences, Genetics. 112 (1986) 887–898.

[42] T.M. Fulton, J. Chunwongse, S.D. Tanksley, Microprep protocol for extraction of DNA from tomato and other herbaceous plants, Plant Mol. Biol. Report. 13 (1995) 207–209.

[43] S. Andrews, F. Krueger, A. Segonds-Pichon, L. Biggins, C. Krueger, S. Wingett, FastQC: a Quality Control Tool for High Throughput Sequence Data, (2010) https://www.bioinformatics.babraham.ac.uk/projects.

[44] The Tomato Genome Consortium, The tomato genome sequence provides insights into fleshy fruit evolution, Nature. 485 (2012) 635–641. https://doi.org/10.1038/nature11119.

[45] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2., Nat. Methods. 9 (2012) 357–9. https://doi.org/10.1038/nmeth.1923.

[46] F. García-Alcalde, K. Okonechnikov, J. Carbonell, L.M. Cruz, S. Götz, S. Tarazona, J.

Dopazo, T.F. Meyer, A. Conesa, Qualimap: evaluating next-generation sequencing alignment data., Bioinformatics. 28 (2012) 2678–2679. https://doi.org/10.1093/bioinformatics/bts503.

[47] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M.A. DePristo, The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data, Genome Res. 20 (2010) 1297–1303. https://doi.org/10.1101/gr.107524.110.

[48] M.A. DePristo, M.A. Rivas, A. McKenna, C. Hartl, G. del Angel, A.Y. Sivachenko, A.A. Philippakis, M. Hanna, M.J. Daly, D. Altshuler, S.B. Gabriel, T.J. Fennell, R. Poplin, K. V Garimella, A.M. Kernytsky, K. Cibulskis, J.R. Maguire, E. Banks, A framework for variation discovery and genotyping using next-generation DNA sequencing data, Nat. Genet. 43 (2011) 491–498. https://doi.org/10.1038/ng.806.

[49] P. Danecek, A. Auton, G. Abecasis, C.A. Albers, E. Banks, M.A. DePristo, R.E. Handsaker, G. Lunter, G.T. Marth, S.T. Sherry, G. McVean, R. Durbin, The variant call format and VCFtools, Bioinformatics. 27 (2011) 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

[50] B.N. Mansfeld, R. Grumet, QTLseqr: An R Package for Bulk Segregant Analysis with Next-Generation Sequencing, Plant Genome. 11 (2018) 0. https://doi.org/10.3835/plantgenome2018.01.0006.

[51] P.M. Magwene, J.H. Willis, J.K. Kelly, The statistics of bulk segregant analysis using next generation sequencing, PLoS Comput. Biol. 7 (2011) 1–9. https://doi.org/10.1371/journal.pcbi.1002255.

[52] Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, J. R. Stat. Soc. Ser. B. 57 (1995) 289–300.

[53] A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm, S.G. Rozen, Primer3-new capabilities and interfaces, Nucleic Acids Res. 40 (2012) 1–12. https://doi.org/10.1093/nar/gks596.

[54] S.D. Tanksley, MAPPING POL YGENES, Annu. Rev. Genet. 27 (1993) 205–233. https://doi.org/10.1146/annurev.ge.27.120193.001225.

[55] E.S. Lander, D. Botstein, Mapping mendelian factors underlying quantitative traits using RFLP linkage maps, Genetics. 121 (1989) 185–199.

[56] K.W. Broman, H. Wu, Ś. Sen, G.A. Churchill, R/qtl: QTL mapping in experimental crosses, Bioinformatics. 19 (2003) 889–890. https://doi.org/10.1093/bioinformatics/btg112.

[57] D.D. Kosambi, The estimation of map distances from recombination values, Ann. Eugen. 12 (1943) 172–175. https://doi.org/doi:10.1111/j.1469-1809.1943.tb02321.x.

[58] S. Jana, Simulation of quantitative characters from qualitatively acting genes - II.

Orthogonal subdivision of hereditary variance in two-locus genetic systems, Theor. Appl. Genet. 42 (1972) 119–124. https://doi.org/10.1007/BF00583413.

[59] J.A. Di Rienzo, J.A. Di Rienzo, F. Casanoves, M.G. Balzarini, L. Gonzalez, M. Tablada, C.W. Robledo, InfoStat: software estadístico., (2001).

[60] E. van der Knaap, S.D. Tanksley, The making of a bell pepper-shaped tomato fruit: Identification of loci controlling fruit morphology in Yellow Stuffer tomato, Theor. Appl. Genet. 107 (2003) 139–147. https://doi.org/10.1007/s00122-003-1224-1.

[61] S. Wu, J.P. Clevenger, L. Sun, S. Visa, Y. Kamiya, Y. Jikumaru, J. Blakeslee, E. van der Knaap, The control of tomato fruit elongation orchestrated by sun, ovate and fs8.1 in a wild relative of tomato, Plant Sci. 238 (2015) 95–104. https://doi.org/10.1016/j.plantsci.2015.05.019.

[62] G.Y. Green, J.H. Pereira da Costa, V. Cambiaso, G.R. Pratta, R. Zorzoli, G.R. Rodríguez, Single and joint effect of the basal region of chromosome 2 and centromeric region of chromosome 8 on morphological and fruit quality traits in tomato, Euphytica. 210 (2016) 327–339. https://doi.org/10.1007/s10681-016-1689-z.

[63] Z. Huang, E. van der Knaap, Tomato fruit weight 11.3 maps close to fasciated on the bottom of chromosome 11, Theor. Appl. Genet. 123 (2011) 465–474. https://doi.org/10.1007/s00122-011-1599-3.

[64] D. Rodríguez-Leal, Z.H. Lemmon, J. Man, M.E. Bartlett, Z.B. Lippman, Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing, Cell. 171 (2017) 470-480.e8. https://doi.org/10.1016/j.cell.2017.08.030.

[65] S.D. Tanksley, The Genetic , developmental, and molecular bases of fruit size in tomato and shape variation, Plant Cell. 16 (2004). https://doi.org/10.1105/tpc.018119.S182.

[66] Q. Mu, Z. Huang, M. Chakrabarti, E. Illa-Berenguer, X. Liu, Y. Wang, A. Ramos, E. van der Knaap, Fruit weight is controlled by Cell Size Regulator encoding a novel protein that is expressed in maturing tomato fruits, PLoS Genet. 13 (2017). https://doi.org/10.1371/journal.pgen.1006930.

[67] R. Roohanitaziani, R.A. de Maagd, M. Lammers, J. Molthoff, F. Meijer-Dekens, M.P.W. van Kaauwen, R. Finkers, Y. Tikunov, R.G.F. Visser, A.G. Bovy, Exploration of a resequenced tomato core collection for phenotypic and genotypic variation in plant growth and fruit quality traits, Genes (Basel). 11 (2020) 1–22. https://doi.org/10.3390/genes11111278.

[68] E. van der Knaap, A. Sanyal, S.A. Jackson, S.D. Tanksley, High-resolution fine mapping and fluorescence in Situ hybridization analysis of sun, a locus controlling tomato fruit shape, reveals a region of the tomato genome prone to DNA rearrangements, Genetics. 168 (2004) 2127–2140. https://doi.org/10.1534/genetics.104.031013.

[69] I. Celik, N. Gurbuz, A.T. Uncu, A. Frary, S. Doganlar, Genome-wide SNP discovery and

QTL mapping for fruit quality traits in inbred backcross lines (IBLs) of solanum pimpinellifolium using genotyping by sequencing, BMC Genomics. 18 (2017) 1–10. https://doi.org/10.1186/s12864-016-3406-7.

[70] P. Adhikari, J. McNellie, D.R. Panthee, Detection of quantitative trait loci (Qtl) associated with the fruit morphology of tomato, Genes (Basel). 11 (2020) 1–15. https://doi.org/10.3390/genes11101117.

[71] V. Garcia, C. Bres, D. Just, L. Fernandez, F.W.J. Tai, J.P. Mauxion, M.C. Le Paslier, A. Bérard, D. Brunel, K. Aoki, S. Alseekh, A.R. Fernie, P.D. Fraser, C. Rothan, Rapid identification of causal mutations in tomato EMS populations via mapping-by-sequencing, Nat. Protoc. 11 (2016) 2401–2418. https://doi.org/10.1038/nprot.2016.143.

Figure captions

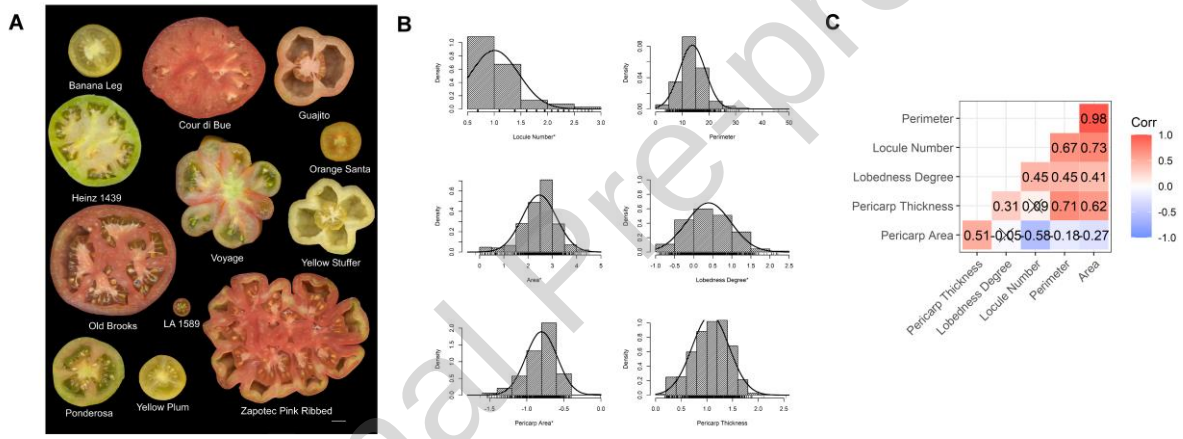

Figure 1: Diversity for medio-lateral direction fruit shape traits.

(a) Representative images for fruit cut at medio-lateral direction. Scale: 1cm. (b) Density histograms for analysed shape traits in a subset of the images. *Variables were transformed using natural logarithm (c) Correlation matrix showing the phenotypic association between traits. Cell colours indicate the strength of Spearman's rank correlations (positive or negative) between pairs of traits. Positive strong correlations are highlighted in dark red, and lighter colours indicate weaker relations. Crosses indicate non-significative correlations

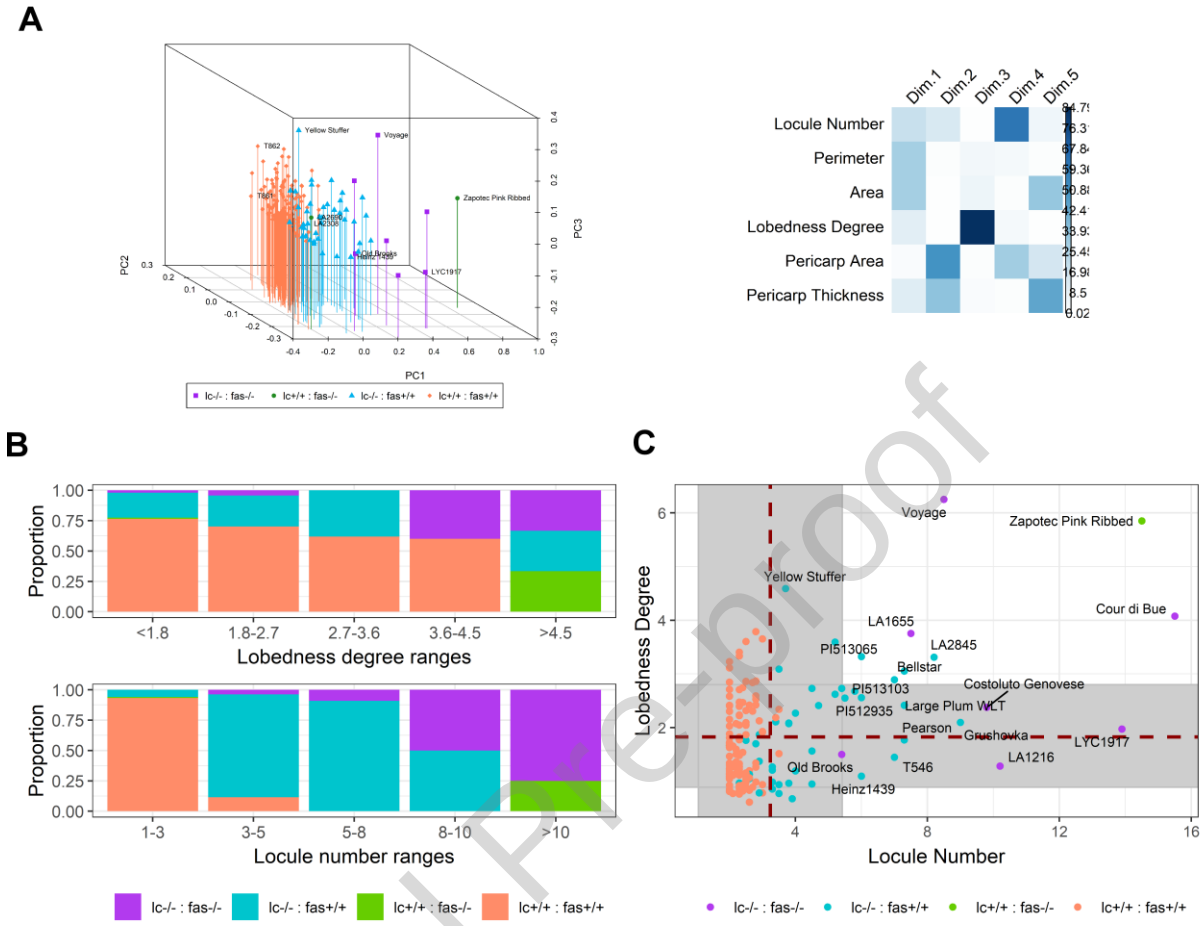

Figure 2: Impact of alleles at *LC* and *FAS* genes on main shape traits at medio-lateral direction

(a) Tri-plot from principal components analysis for shape traits at medio-lateral direction. The contribution of variables to principal components (dim.), as percentages, is displayed as a colour scale. Darker blue colours mean a greater contribution of the variable on the principal component. (b) Relative frequency distribution for alleles at *LC* and *FAS* genes according to lobedness degree and locule number ranges. (c) Accessions characterization for lobedness degree/locule number relation and LC and FAS allelism. Grey area stands out the traits values among mean ± one standard deviation. Dots colours represent tomato accessions with specific *LC* and *FAS* allelic constitution. Groups are: $lc^{+/+}$: $fas^{+/+}$ (wild alleles for both genes), $lc^{-/-}$: $fas^{-/-}$ (mutated alleles for both genes), and the combined options, $lc^{+/+}$: $fas^{-/-}$, and $lc^{-/-}$: $fas^{+/+}$. Examples of accession with extremes values for fruit shape traits are named.

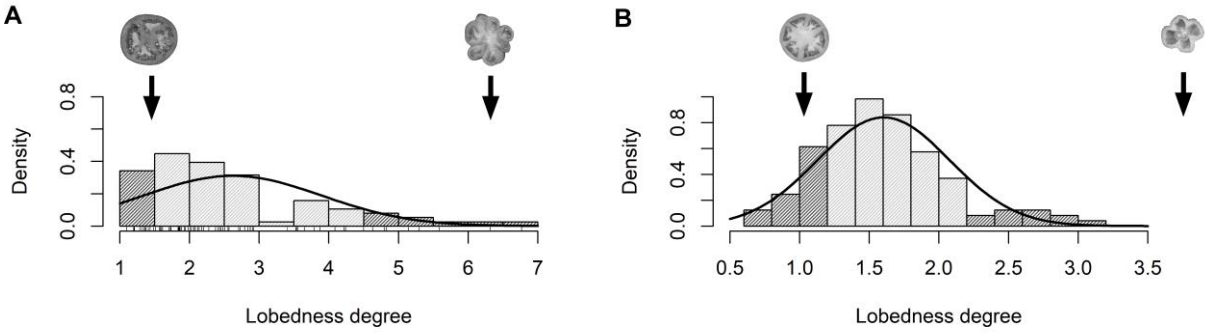

Figure 3: Phenotype distribution for lobedness degree in the F_2 analysed populations.

(a) F_2 Voyage x Old Brooks population (b) F_2 Yellow Stuffer x Heinz 1439 population. An image fruit representative of each parental genotype is shown. The mean values of the parental genotypes are indicated by arrows. Normal distribution trends are indicated with a curve. Parallel bars at the bottom of the graph show data concentration: the thicker line the greater concentration.

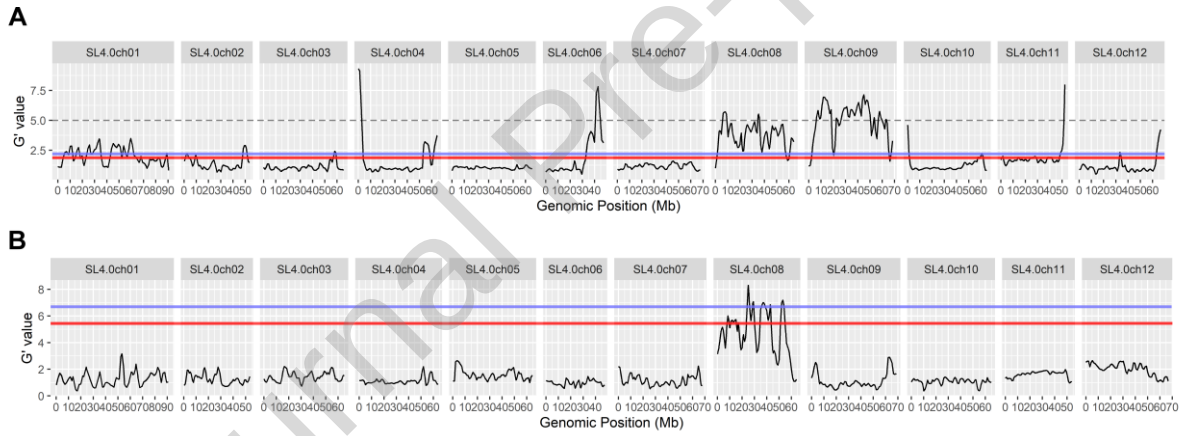

Figure 4: Genomic regions associated with lobedness degree along the 12 tomato chromosomes in two F_2 analysed populations.

(a) F_2 Voyage x Old Brooks population (b) F_2 Yellow Stuffer x Heinz 1439 population. Genome-wide tricube-smoothed G' value. Dash line indicate G' value of 5. Benjamini-Hochberg false discovery rate (FDR) threshold of 0.05 is indicated in red and 0.01 is indicated in blue. SL4.0ch: chromosome. Sliding window = 2Mb.

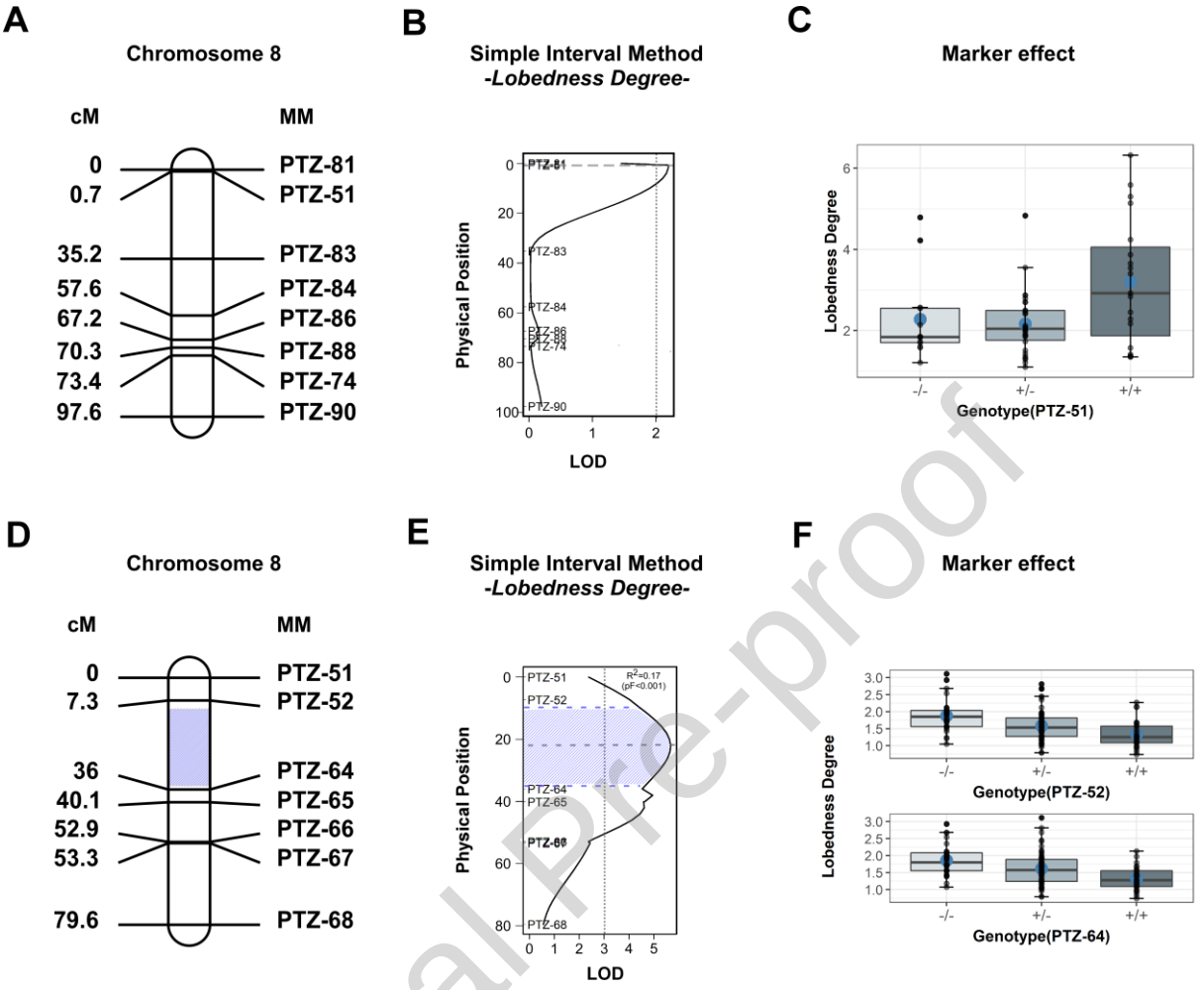

Figure 5: Lobedness degree QTLs validation in chromosome 8 for both F_2 populations.

(a) Chromosome 8 linkage map for lobedness degree in F_2VxOB. (b) Logarithm of the odds (LOD) scores along chromosome 8 in F_2VxOB. cM: Centimorgan. MM: identifier of the molecular marker. Physical position expressed in megabases (Mb). The dotted line signals the 3 LOD significance thresholds. The dashed line, point to the maximum LOD score (2.2) at 0.69 (0.96 Mb). (c) Boxplot of lobedness degree value against the genotype at marker PTZ-51 in F_2VxOB. Blue bigger dot indicates mean lobedness degree value and black dots correspond to imputed genotypes. Genotype: -/- mutant homozygous (alleles as irregular parent), -/+ heterozygous, +/+ wild type homozygous (alleles as regular parent). Error bars indicate ± 1 standard error. (d) Chromosome 8 linkage map for lobedness degree in F_2YSxH. Region significantly associated with lobedness degree is stand out in blue colour. (e) Logarithm of the odds (LOD) scores along chromosome 8 in F_2YSxH. The dotted line signals the 3 LOD significance thresholds. The grey dashed line, point to the maximum LOD score (5.71) at 22 cM position. The blue dashed lines indicate the beginning and end of locus (f) Boxplot of lobedness degree against the genotype at marker PTZ-52 and PTZ-64.

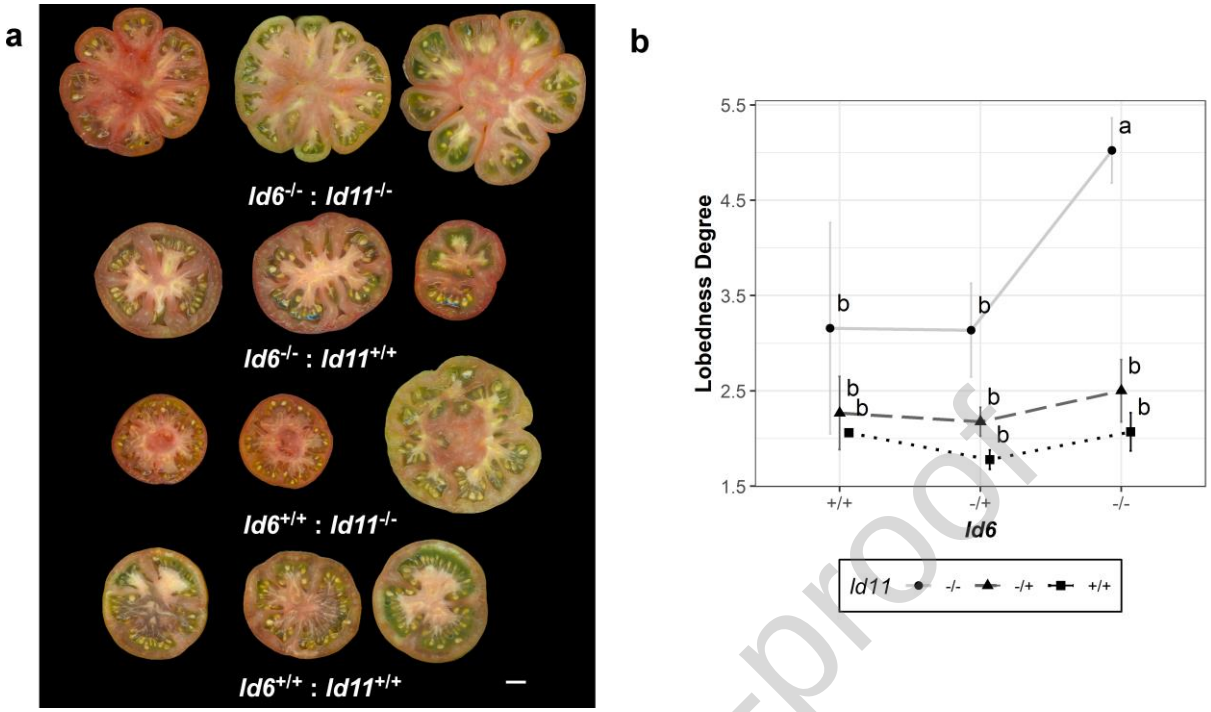

Figure 6: Effects of interaction between *ld6* (PTZ-12 marker) and *ld11* (PTZ-103 marker) loci on lobedness degree.

(a) Representative fruit images for different allele combinations of ld6 and ld11 loci. (b) Phenotype x genotype interaction plot. Genotypes: -/- mutant homozygous genotype, -/+ heterozygous genotype, +/+ wild type homozygous genotype. Dots represent mean genotype value and error bars showed ± 1 standard deviation.

Table 1: Descriptive statistics parameters and heritability for latitudinal shape traits

Trait	Mean ± s.d.	Min	Max	Median	Q1	Q3	IQ range	Heritability ± s.e.
Locule Number*	3.13 ± 2.22	2.00	17.00	2.00	2.00	3.00	1.00	0.79 ± 0.023
Perimeter	13.78 ± 4.89	3.43	49.15	13.39	10.72	16.17	5.45	0.91 ± 0.018
Area*	14.61 ± 10.15	0.83	90.60	12.74	8.06	18.13	10.07	0.94 ± 0.014
Lobedness Degree*	1.76 ± 1.16	0.38	9.12	1.44	0.96	2.23	1.27	0.79 ± 0.023
Pericarp Area*	0.46 ± 0.09	0.20	1.00	0.46	0.40	0.51	0.11	0.72 ± 0.024
Pericarp Thickness	1.07 ± 0.36	0.18	2.55	1.11	0.84	1.32	0.48	0.88 ± 0.020

s.d.: Standard deviation, Min: Minimum value, Max: Maximum value, Q1: first quartile, Q3: third quartile, IQ range: interquartile range, s.e.: standard error. Units: perimeter (cm), area (cm2). *: Variables transformed using natural logarithm.

Table 2: Mean values for fruit locule number and lobedness degree, and molecular description in the parental genotypes

F_2 population	Parental name	Specie	Locule number	Lobedness degree	Shape genes				
					FAS	LC	SUN	OVATE	SOV1
Population 1	Voyage (V)	*Solanum lycopersicum* L.	8.50 ± 0	6.25 ± 4	1	1	3	3	3
	Old Brooks (OB)	*Solanum lycopersicum* L.	5.43 ± 8	1.50 ± 9	1	1	3	3	3
Population 2	Yellow Stuffer (YS)	*Solanum lycopersicum* L.	3.67 ± 1	4.59 ± 6	3	1	3	3	3
	Heinz 1439 (H)	*Solanum lycopersicum* L..	6.00 ± 7	1.10 ± 7	3	1	3	3	3

Shape genes: 1: mutant allele, 3: wild-type allele, a data obtained from Rodríguez et al. (2011) and Wu et al. (2018) as described at the supplementary data. Values are given as the mean ± s.d.

Table 3: Description of fields length, number of SNPs, value, position, and significance of maximum Δ (SNP-index) and G', for the putative QTLs of lobedness degree identified in the different chromosomes (genome version SL 4.0)

Population	Chromosome	Start (Mb)	End (Mb)	Length (Mb)	nSNPs	peak Δ (SNP-index)	pos peak Δ (SNP-index)	Max. G'	Pos max. G'	Mean Qvalue
F_2 VxOB	4	0.11	4.06	3.95	583	-0.44	1.10	9.32	0.11	0.00
	6	32.71	47.25	14.54	2515	0.39	42.83	7.82	42.83	0.00
	6	1.16	58.22	57.05	1652	0.27	52.98	5.71	8.05	0.01
	9	2.62	65.96	63.34	2140	0.37	44.90	7.14	44.90	0.00
	11	47.93	52.29	4.36	679	0.12	51.46	7.98	52.29	0.01
F_2 YSxH	8	8.46	9.07	0.61	3	-0.27	8.46	6.02	8.46	0.03
		10.73	14.94	4.21	27	-0.26	12.04	5.74	14.94	0.04
		23.45	29.74	6.29	49	-0.34	29.16	8.31	24.95	0.02
		34.99	43.11	8.12	51	-0.37	40.00	7.01	37.11	0.02
		51.13	54.84	3.71	256	0.38	52.98	7.20	52.98	0.02

Population: F_2 population derived from crosses between tomato cultivars (*Solanum lycopersicum* L. cv) "Voyage" x "Old Brooks" (F_2 VxOB) and Yellow Stuffer" x "Heinz 1439" (F_2 VxOB). Chromosome: the chromosome on which the region was identified. Start: the position in megabases (Mb) of the first SNP that passed the FDR threshold on that chromosome. End: the end position expressed in Mb. Length: the length in Mb from start to end of the region. nSNPs: the number of SNPs in the region. peak Δ (SNP-index): the tricube-smoothed Δ (SNP-index) value at the peak summit. pos peak Δ(SNP-index): the position of the absolute maximum tricube-smoothed Δ (SNP-index). Max. G': the maximum G' score in the region. Pos max. G': the genomic position of the maximum G' value in the QTL. Mean Qvalue: the average adjusted p-value in the region.

Table 4: QTL detection by single-point analysis for fruit shape traits in F_2 VxOB

Marker	Chr	Pos	Fruit Weight		Locule Number		Perimeter		Area		Pericarp Area		Pericarp Thickness		Lobedness Degree	
			F value	p value	F value	p value	F value	p value	F value	p value	F value	p value	F value	p value	F value	p value
PTZ-	4	0.	0.5	0.56	0.8	0.43	0.1	0.83	0.2	0.80	0.2	0.76	0.9	0.4	1.0	0.35

93		97	8		5		9		3		8		2	0	7	
PTZ-97	4	62.53	0.30	0.74	0.14	0.87	0.82	0.45	0.94	0.40	0.97	0.39	2.74	0.07	0.33	0.72
PTZ-12	6	40.98	1.30	0.28	4.24	0.02	0.92	0.40	0.44	0.65	0.05	0.95	1.39	0.25	10.88	0.0001
PTZ-17	6	0.96	1.75	0.18	1.94	0.15	2.30	0.11	1.14	0.32	0.78	0.46	0.15	0.86	6.35	0.0029
PTZ-99	9	12.78	1.37	0.26	0.07	0.94	1.84	0.17	2.02	0.14	0.98	0.38	1.94	0.15	0.37	0.69
PTZ-101	9	65.08	1.16	0.32	0.62	0.54	0.72	0.49	0.80	0.45	0.43	0.65	2.27	0.11	0.27	0.77
PTZ-103	11	53.26	13.62	<0.0001	139.32	<0.0001	18.15	<0.0001	11.65	<0.0001	2.55	0.09	0.16	0.85	30.39	<0.0001

Chr: chromosome; Pos: Physical position corresponding to the tomato reference genome version SL4.0 in megabases; F value and p value correspond to ANOVA test. Units: perimeter (cm), area (cm^2).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper'

Highlights

- First report about diversity in the medio-lateral direction of tomato fruits
- LC and FAS genes highly explain locule number but not lobedness degree variability
- QTL-seq was applied in two F_2 populations with no segregation for FAS and LC
- Three QTLs with major effect for lobedness degree (*ld6*, *ld8*, and *ld11*) were mapped
- *ld6* and *ld11* showed epistatic interaction and accounted for ~61% of the variability