

Discovery of a major QTL controlling trichome IV density in tomato using K-seq genotyping

Estefanía Mata-Nicolás^{1‡}, Javier Montero-Pau^{2‡}, Esther Gimeno-Paez¹, Ana García-Pérez¹, Peio Ziarsolo¹, José Blanca¹, Esther van der Knapp^{3,4}, María José Díez¹ and Joaquín Cañizares^{1#}

¹Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana. COMAV. Universitat Politècnica de València. Spain.

²Instituto Cavanilles de Biodiversidad y Biología Evolutiva. Universitat de València. Spain

³Institute of Plant Breeding, Genetics and Genomics. University of Georgia. GA. USA

⁴Department of Horticulture University of Georgia. GA. USA

‡ Both authors contributed equally

Author for correspondence

Abstract

Trichomes are a common morphological defense against pests, in particular, type IV glandular trichomes have been associated with resistance against different invertebrates. Cultivated tomatoes usually lack or have a very low density of type IV trichomes. Thus, specific breeding programs to incorporate these natural defences, that are common within the *Solanum* genus, might improve a more sustainable management. We have identified a *S. pimpinellifolium* accession with very high density of this type of trichomes. Two F2 mapping populations using two different parents have been developed, characterized and genotyped using a new genotype methodology, K-seq. We have been able to build an ultra-dense genetic map with 147,326 markers with an average distance between markers of 0.2 cM that has allowed us to perform a detailed mapping. We have used two different families and two different approaches, QTL mapping and QTL-seq, to identify several QTLs implicated in the control of trichome type IV developed in this accession on the chromosomes 5, 6, 9 and 11. The QTL located on chromosome 9 is a major QTL that has not been previously reported in *S. pimpinellifolium* that increases by a factor of 9 the density of trichomes.

Keywords

Trichomes type IV, Trichomes, Tomato, QTL, K-seq, *Solanum pimpinellifolium*

Introduction

Agricultural production is greatly affected by pests and diseases in terms of both yield and quality and therefore it is highly dependent on pesticide application. Moreover, the use of pesticides imposes an additional production cost, and can negatively impact human health and the environment. Additionally, pesticides can have negative effects in arthropod biodiversity, which in turn can result in loss of crop production [3]. As a consequence, new approaches are being sought in agricultural management programs to provide more sustainable alternatives [4]. Recent efforts have been mainly focused on the development of crops with genetic resistance.

Presence of trichomes has been shown to be an important morphological defense against insect pests [5,6]. Specifically, glandular trichomes, which accumulate and exude secondary metabolites, are known to interfere with oviposition, larval fixation and development and also can trap or poison insects [4,7,8]. The type of trichomes and stored compound depends on the species but they might include acylsugars, methyl ketones, terpenoids, phenylpropenes or flavonoids [9]. Breeding for resistant cultivars with an increased density of glandular trichomes could be a valuable option to reduce the impact of pesticides. However, this strategy requires a detailed knowledge of the genetic control and molecular basis of trichome development.

Trichomes are a common morphological defense in the genus *Solanum* [1,2]. In particular, presence and density of type IV glandular trichomes --trichomes with a short multicellular stalk and small gland at the tip [10]-- has been associated with pest resistance such as spider mite [11,12], whitefly [13–16] or potato moth [17]. Moreover, resistance against whiteflies and aphids has been associated with the production of acylsugars in type IV trichomes [18]. Cultivated tomato (*Solanum lycopersicum* var. *lycopersicum* L.) usually lacks or has a very low density of type IV trichomes. Thus, specific breeding programs to incorporate these natural defences, that are common within the *Solanum* genus, might improve this crop management.

Several tomato wild relatives have proven to be a rich source of genetic diversity for type IV trichomes; particularly *Solanum pennellii* Correll, *Solanum habrochaites* S.Knapp & D.M Spooner and *Solanum galapagense* S.C.Darwin & M.I. Peralta. Some accessions of these species exhibit trichome densities higher than 90 trichomes/mm² [16,19,20]. Inheritance of

type IV trichome density has been studied using interspecific crosses among these related species and tomato. Presence of trichome type IV on *S. pennelli* has been proposed to be under the control of two dominant unlinked genes [21]. Andrade et al. [16] found that inheritance of trichome type IV on *S. galapagense* is relatively simple and could be controlled by incompletely recessive alleles. Alternatively, genetic factors underlying type IV trichome density have been identified by quantitative trait locus (QTL) mapping studies in other studies. Some QTLs have been confirmed in chromosome 2 of *S. pennelli* [20], in chromosome 2 and 9 of *S. habrochaites* [22] and in chromosome 2 and 9 of *S. galapagense* [23]. Synthesis of acylsugars has also been associated with QTLs in populations derived from *S. pennelli*. These QTLs were located in chromosome 2, 3, 4, 5, 6 and 11 of *S. pennelli* [20,24,25].

Despite the fact that these wild relatives species have been described as a potential source of type IV trichomes, their use in breeding programs is limited because the sexual incompatibility hinders their introgression and their undesirable agronomic traits requires long breeding programs to incorporate desired traits. As a consequence, a more closely related wild species like *Solanum pimpinellifolium* L., a red-fruited, self-compatible species and without incompatibility barriers with the cultivated tomato, is generally preferred in tomato breeding programs as it can avoid an excessive linkage drag in backcrossing programmes. Unfortunately, the number of accessions of *S. pimpinellifolium* with a relatively high density of type IV trichome reported to date is scarce. Rahka et al. [15] described some accessions which were characterized by densities lower than 4.5 trichomes/mm². Fernández-Muñoz et al. [26] found an accession of *S. pimpinellifolium* (TO-937) characterized by a mean of 16.3 type IV trichomes/mm² in the abaxial surface, and Firdaus et al. [27] found an accession of *S. pimpinellifolium* (LA1584) with 21.0 trichomes/mm². Interestingly, a population derived from the accession TO-937 confirmed two QTLs in chromosome 2 involved in resistance to spider mite [28].

Trichome presence and density is also known to depend on developmental factors such as plant and leaf age. Vendemiatti et al. [29] found a decrease in type IV trichomes density from cotyledons (40 trichomes/mm²) up to sixth leaves (< 5 trichomes/mm²) in commercial cultivars of tomato. *S. habrochaites* showed higher pubescence in 9-week old plants compared to 6-week old plants (18 vs. 6 type IV trichomes/mm²) [30]. F2 and F1 populations between *Solanum lycopersicum* and *S. pennelli* also exhibited higher densities in plants at 10 and 13 weeks after transplanting, compared to 7 weeks [31]. However, Wilkens et al. [32] described a reduction of density of type IV trichomes with plant age in an older Ecuadorian tomato cultivar. An effect of light was described in *S. habrochaites*, trichome densities were higher under a photoperiod of 8 hours of light photoperiod, compared to 14-15 hours [33].

The main objective of this study was to map the quantitative trait loci controlling type IV trichome density in a *S. pimpinellifolium* accession that, as far as we are aware, presents the highest density of type IV trichome (> 100 trichomes/mm²) ever described in the literature in this species. This accession has been used to study the genetic basis of the presence and density of type IV trichomes by crossing the accession with two different genetic backgrounds of *S. lycopersicum*. We also show for the first time the practical use of a new K-seq genotyping technology [34].

Materials and methods:

Plant material

During the characterization of the Varitome collection of *S. pimpinellifolium*, *S. lycopersicum* var. *cerasiforme* and *S. lycopersicum* var. *lycopersicum* from Perú, Ecuador and México [35], an accession of *S. pimpinellifolium* with a high density of glandular type IV trichomes was found (BGV016047).

In order to study the genetic basis of trichome IV density, we have used the F1 and F2 collections developed by the Varitome project [35]. BGV016047 was crossed with two different genetic backgrounds with low trichome density: *Solanum lycopersicum* var. *cerasiforme* LA2278 (SLC family) and *Solanum lycopersicum* var. *lycopersicum* cv. Moneymaker (SLL family). F1 populations were developed using BGV016047 as male parent in the SLC family and as a male parent in the SLL family, and the F2 generation was obtained by self-pollination. Additionally, for this study the F1 of SLL family was also backcrossed with Moneymaker (BC1).

All accessions used in this paper were provided and are accessible at the genebank of COMAV at the Polytechnic University of Valencia and at the TGRC (USA).

Glandular type IV trichome density characterization

To evaluate the effect of plant age on glandular type IV trichome density on the BGV016047 accession an experiment was conducted during spring/summer 2018. The effect of plant age was evaluated by measuring trichome density on leaflets on the same developmental stage. Ten plants of this accession were grown in the greenhouse and fertigated under standard doses for tomato. After transplanting, the second leaflet of the second fully expanded leaf from the apex of the plant was sampled weekly from day 9 to 89 after transplanting (13 sample

points). These leaflets were removed from the plant and placed in a Petri dish with wet filter paper to prevent desiccation. Leaflets were cut, perpendicularly to the central vein, into two approximately equal sections to facilitate the observations. These sections were viewed under a 40x magnification microscope and trichome density was estimated in four random areas of 1 mm² per leaflet. Counts of type IV trichomes were made on the abaxial surface of each leaflet, avoiding the interveinal areas and the main vein where trichome density was higher.

The influence of leaf age was explored by following the same leaves over time. Three leaves on 10 different plants were followed during three consecutive weeks, corresponding to the time when the leaves were in the second position, third and fourth position from the apex (at 49, 53 and 60 days after transplanting, respectively). Type IV thricome density was measured as described above.

The SLL family assay was done with 24 BGV016047 plants, 10 of MoneyMaker, 16 of F1, 195 of F2 and 50 of BC1. The phenotypic characterization of the SLC family included 10 plants from each parent and 151 F2 plants. Seeds were sown in trays with commercial potting mix and transplanted to the greenhouse in a completely randomized design five weeks (SLC family) or four weeks (SLL family) after sowing in Paiporta (Valencia, Spain). The phenotyping started once the leaves from BGV016047 exhibited a density higher than 10 trichomes/mm². As a consequence, the SLC family was evaluated the sixth, seventh and eighth weeks after transplanting, whereas the SLL family was evaluated the third, fourth and sixth weeks after transplanting. Type IV trichome density was measured on leaflets from leaves in the second position below apex, following the procedure described above. To reduce the environmental, plant and leaf age effects, measurements of the whole family were made during the same week. SLC family was phenotyped during spring/summer 2018 and SLL family during autumn/winter 2017.

DNA extraction and plant genotyping by K-seq

For the SLC family, a total of 137 individuals from the F2 generation plus 15 parental controls were genotyped. In the case of the SLL family, a Bulk Segregant Approach was used: F2 plants with the highest and lowest trichome density were chosen and individually genotyped 15 plants with an average density higher than 10 trichomes/mm² and 15 plants with a density lower than 0.5 trichomes/mm². Genomic DNA was extracted from leaves using a modified CTAB protocol [36].

Genotyping was performed using K-seq, a novel genotyping method based on the reduction of genome complexity using Klenow polymerase and short oligonucleotides [34]. The genotyping was done as described in Ziarsolo et al [34] but using 25°C as annealing temperature for the Klenow polymerase amplification. A total of 200 ng of DNA was subjected to two cycles of amplification using Klenow polymerase and short primers. After each cycle, an Exonuclease I digestion step was performed to remove uncopied genomic DNA and free primers. Finally, the double strand fragments were amplified during 13 cycles using a standard PCR reaction with indexed primers. Short oligonucleotides and PCR primers are shown in Supplementary File 1. PCR samples were pooled in two pools using the same volume PCR sample, size selected and sequenced in two HiSeq2500 lanes 2x150pb (Illumina, CA, USA) by the CNAG-CRG (Barcelona, Spain). The reads have been deposited under Bioproject PRJNA649673

Raw reads quality was inspected using FastQC v. 0.11.5 [37] and low quality reads and adapters were removed using Trimmomatic v. 0.36 [38]. Clean reads were mapped against *Solanum lycopersicum* genome build SL2.5 using BWA-MEM v. 0.7.17 [39]. SNP calling was performed using freebayes v 1.3.1 [40] and the SNPs with more than 50% of missing data and a minor frequency allele lower than 0.05 were filtered out those using custom Python scripts (<https://github.com/bioinformav>).

Genetic mapping, QTL and QTL-seq analysis

QTL analysis was performed using F2 individuals of SLC family using two different approaches: 1) construction of a genetic linkage map followed by QTL mapping, and 2) QTL-seq analysis [41] between F2 individuals with extreme trichome density values for each mapping population.

For the first approach, the filtered SNPs were transformed into ABH genotype coding using custom scripts for the 107 F2 individuals of the SLC family that passed the quality filters. Transformed genotype data was corrected using GenotypeCorrector [42] using a sliding window size of 41 SNPs and after binning consecutive homozygous markers. Construction of the genetic map and QTL location was performed with *R/qtl* [43] and *ASMap* [44] packages from R software. For the construction of the genetic map, adjacent SNPs with the same genotype were also binned and SNPs in chromosome 0 were discarded. QTL mapping was done by single interval mapping based on a non-parametric test (Kruskal-Wallis test). Statistical significance thresholds were calculated based on 1000 bootstrap. QTL location interval was established based on 1.5 LOD units drop. Additionally, the effect of each QTL

was determined by fitting to a generalized linear model (logistic model) assuming a negative binomial distribution and a log link function using R package *MASS* [45]. Best fitting distribution for trichome density was found using the R package *fitdistrplus* [46]. Goodness of fit was checked for each model. Proportion of deviance (D^2) explained by the model and Nagelkerke R^2 were calculated using functions “Dsquared” and “RsqGLM” from the *modEvA* package [47].

The QTL-seq was done by bulking the genotype information of F2 individuals with the lowest and the highest type IV trichome density. For the SLC family, 15 individuals with an average density of 0.05 trichomes/mm² (range from 0 to 0.25 trichomes/mm²) and 15 with a density of 94.52 trichomes/mm² (range from 53.00 to 212.50 trichomes/mm²) were selected. For the SLL family, 15 individuals with an average density of 0.03 trichomes/mm² (range from 0 to 0.25 trichomes/mm²) and 13 with a density of 28.35 trichomes/mm² (range from 11.25 to 64.25 trichomes/mm²). Allele counts for each group were summed, SNPs with a read depth higher than 150 for SLC family and 250 for SLL family were filtered out. Additionally, SNPs with a reference allele frequency lower than 20% or higher than 80% were also removed. The delta (Δ) SNP-index was calculated by subtracting the SNP-indices of the bulks at each loci, and candidate QTL regions were identified by using a 1 Mb sliding window. Confidence intervals for the Δ SNP were determined using 1,000 simulations. QTL-seq analyses were performed with the QTLseqr R package [48].

Additionally, we also performed an association analysis along the genome for type IV trichome density. We explored the association between genotype for each marker and trichome density by fitting a logistic regression assuming a negative binomial distribution as above. Resulting p-values for each marker were adjusted for multiple comparison by using Benjamini-Yekutieli procedure [49] in R software v.3.6 [50].

Results

Type IV trichome density characterization of S. pimpinellifolium accession BGV016047 with a high density of type IV trichomes

Density of type IV trichomes varied with the age of the plant (Fig. 1A, D). The density of trichomes in the second leaf from the apex ranged from a median of 0 to 9.5 trichomes/mm² (range from 0 to 43 trichomes/mm²) until day 48 after transplanting, afterwards trichome density started to increase to a maximum median of 118 trichomes/mm² (range 8 to 240 trichomes/mm²). Trichome density also showed an increased variability as the plant grew older. In the last week studied (day 89 after transplanting), a drop in the trichome density was observed (median of 57.5 trichomes/mm²). This drop coincided with an increase of minimal

daily temperature and with the death of two analyzed plants; so could be due to environmental factors. The effect of leaf age on density was measured by following three leaves during three consecutive weeks which corresponded with their location as second, third and fourth leaf from the apex (days 49, 53 and 60 after transplanting). This result (Fig. 1B) showed that trichome density on leaflets from young leaves decreased as the leaflet was expanding in the 2nd and 3rd leaves. Trichome density was evaluated in two different seasons (spring/autumn). As it can be seen (Fig. 1C), trichome density during the first week of phenotyping, when parent plants have reached a density of 10 trichomes/mm² in at least one measure, was similar between seasons (median of 35 vs. 34 trichomes/mm²), although with a higher variability in autumn (interquartile range 27 in spring vs 46 in autumn).

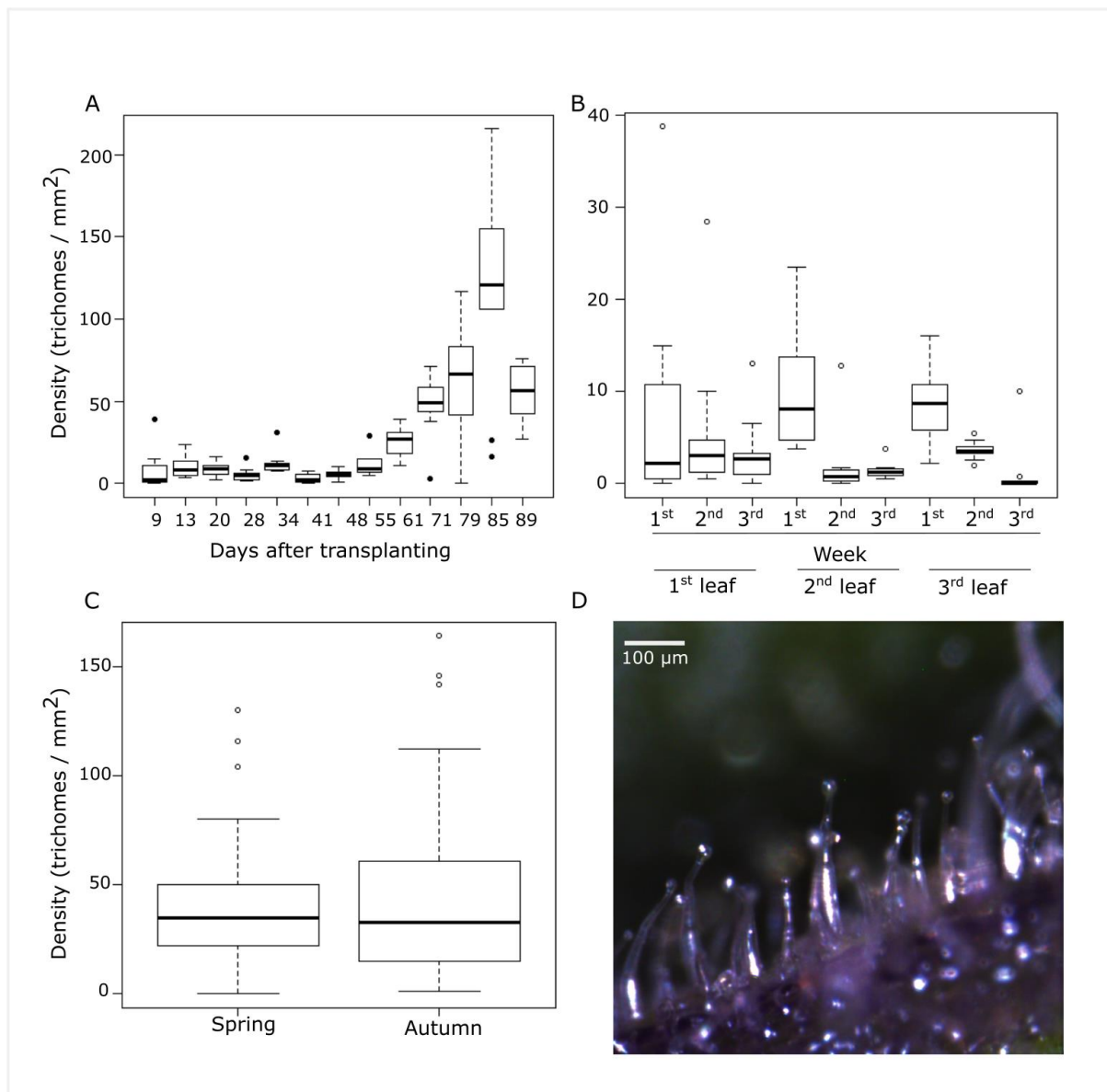


Figure 1. Effect of plant age (A), leaf age (B) and growing season (C) on type IV trichome density.

Trichome type IV density on segregating populations

The *S. pimpinellifolium* accession BGV016047 with a high density of type IV trichomes was crossed against two different genetic backgrounds with low trichome density (*S. lycopersicum* var. *cerasiforme* accession LA2278 and *S. lycopersicum* var. *lycopersicum* 'MoneyMaker' (Suppl. File 2). Figure 2 shows the density distributions of the different crosses. The low trichome density parents exhibit either none or a very low trichome density, whereas the F2 for both families had values between both parents, but showing a clear tendency to low density. However, in some cases some F2 descendants reached or even exceeded the densities of BGV016047. F1 individuals of the SLL family (Fig. 2B) showed, in general, lower values than the corresponding F2, and the backcross generation (BC) obtained from the backcross F1 x MoneyMaker exhibited a very low number of trichomes, recovering the MoneyMaker phenotype. Altogether, this pattern of segregation suggests that the character could be under control of a few QTLs. This pattern did not change with plant age as the same relationship was observed on each of the three weeks that were phenotyped (Suppl. Fig. 1).

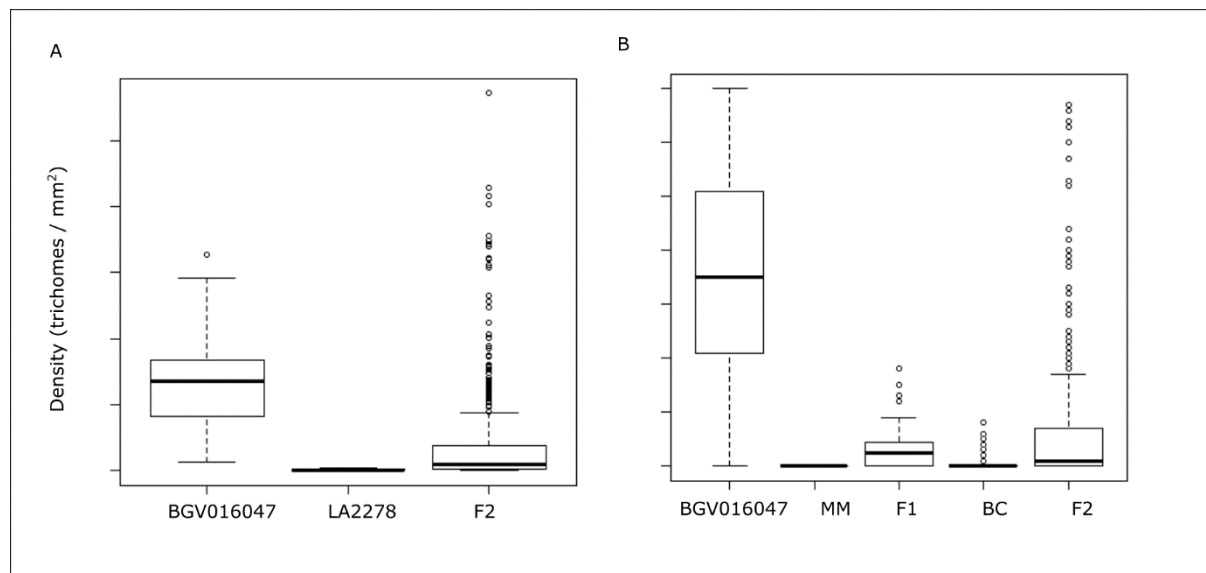


Figure 2. Density of type IV trichomes in the segregating populations for the last week of phenotyping for the SLC family (BGV016047 x LA2278) (A) and SLL family (BGV016047 x MoneyMaker) (B)

QTL and QTLseq analysis in BGV016047 x LA2278 population

The number of raw reads per sample obtained in the K-seq genotyping varied from 399,182 to 7,287,426. This variability was likely due to the fact that the pool of samples was made using equal volumes of PCR product and not quantifying them. In Supplementary Figures 2, 3 and 4 a summary of K-seq read number, mapping and SNP calling results is shown. The average number of raw reads was 2.4 million reads per sample, 1.8 million reads per sample could be mapped against the reference genome with a mapping quality higher than 57 (75.68% of raw reads), which resulted in an average genome coverage of 6,97 % at read depth 1 or higher (1.28% at depth 5 or higher and 0,60 % at depth 10 or higher). After including the parental genome sequences as references into the SNP calling analysis, a total of 4,993,950 SNPs were obtained. The number of SNPs found in the K-seq genotyped samples varied from 121,019 to 729,686. After quality checking and ABH coding 147,326 SNPs and 107 F2 samples were finally used to build the genetic map for the F2 population (Suppl. File 3). Adjacent SNPs with the same parental genotype were binned in 8,285 bin markers. Markers were grouped in 12 linkage groups (Table 1 and Suppl. Fig. 5) that resulted in a total map length of 1,600.4 cM with an average of 12,277.2 SNPs and 690.4 bin markers per chromosome, and an average length of 209,7 cM per chromosome with a distance between markers of 0.2 cM.

Table 1. Summary statistics for the genetic map built using the F2 generation of the cross BGV016047 x LA2278.

	#SNPs	#Bin markers	Length	Avg. spacing	Max. spacing
SL2.50ch01	23,537	1,248	153.4	0.1	3.7
SL2.50ch02	10,368	662	130.3	0.2	1.8
SL2.50ch03	8,300	479	165.2	0.3	8.5
SL2.50ch04	9,482	534	123.9	0.2	1.3
SL2.50ch05	15,374	856	121.4	0.1	7.9
SL2.50ch06	8,375	552	123.7	0.2	11.2
SL2.50ch07	17,919	770	110.9	0.1	1.9
SL2.50ch08	7,555	563	164.7	0.3	13.3
SL2.50ch09	11,342	725	125.9	0.2	3.7
SL2.50ch10	14,855	788	119.9	0.2	5.2
SL2.50ch11	9,925	567	101.8	0.2	2.8
SL2.50ch12	10,294	541	159.2	0.3	6.2
Overall	147,326	8,285	1,600.4	0.2	13.3

The non parametric QTL mapping of BGV016047 x LA2278 population showed two significant QTLs in chromosomes 9 and 11 (Fig. 3 and Table 2) named as TIVd9 and TIVd11. As trichome density varied with plant age, the variance and significance of both QTLs also varied through time, however TIVd9 was significant in all three sets of phenotyping data. Explained variance by the negative binomial logistic regression model exclusively by TIVd9 ranged from 0.18 to 0.26 based on Nalgerke's R^2 , and from 0.30 to 0.37 if both QTLs were considered in an additive model (Fig. 4 and Table 3). The incidence rate ratio (IRR) of the logistic model shows that in the TIVd9 locus each additional allele from the high density parent increased the expected trichome IV density by 3 (IRR varied from 2.61 to 3.24 depending on the analysed week), whereas in TIVd11, each allele of the BGV016047 parent resulted in a decrease of the expected trichome density of about 50% (IRR varied from 0.48 to 0.54).

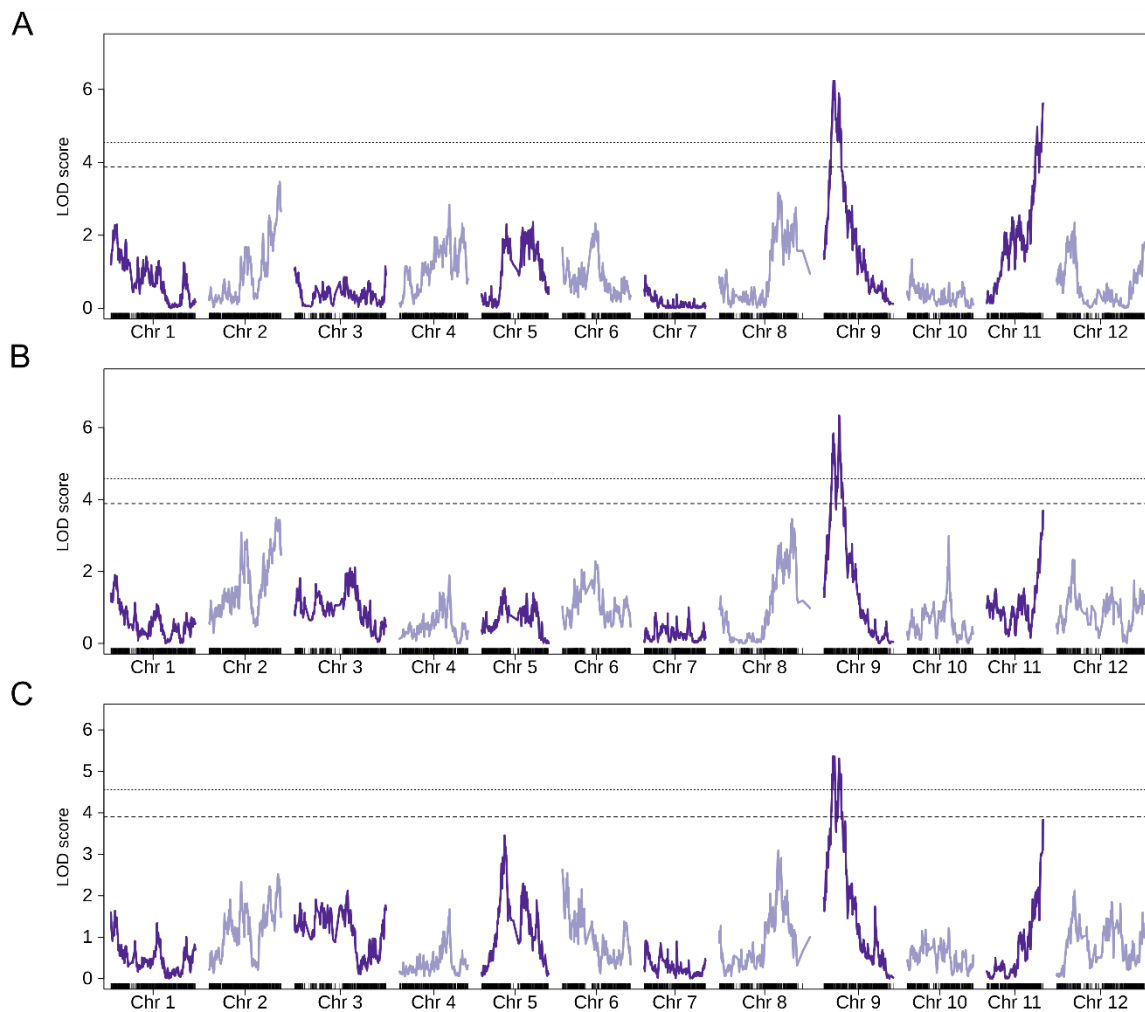


Figure 3. Results of non-parametric single interval mapping QTL analysis for each of the three weeks phenotyped. A) 6th week after transplanting, B) 7th week and C) 8th weeks.

Table 2. QTL location based on different methods using SLC family and SLL family mapping populations. For each QTL, chromosome, start and end position, location of the most significant value of the region (Max) and LOD, adjusted p-value or Δ SNP value for that location are reported,

QTL mapping (SLC family)					
Phenotype	Chr	Start	Max	End	LOD
Trichome density 6th week	SL2.50ch09	2,055,839	2,542,786	4,250,325	6.24
	SL2.50ch11	54,298,873	56,297,460	56,297,460	5.62
Trichome density 7th week	SL2.50ch09	2,078,180	3,756,370	4,250,325	6.33
	SL2.50ch11	55,093,420	56,297,460	56,297,460	3.69
Trichome density 8th week	SL2.50ch09	1,966,640	2,309,105	4,371,756	5.36
	SL2.50ch11	55,548,327	56,297,460	56,297,460	3.83
Association (SLC family)					
Phenotype	Chromosome	Start	Max	End	Adj. pval.
Trichome density 6th week	SL2.50ch09	2,309,105	2,542,786	2,545,171	0.0090
Trichome density 7th week	SL2.50ch09	2,078,180	3,727,914	4,969,831	$7.7 \cdot 10^{-6}$
Trichome density 8th week	SL2.50ch06	8,008	154,414	615,838	0.0002
	SL2.50ch09	2,069,769	2,492,691	5,693,896	$4.6 \cdot 10^{-5}$
QTLseq					
Mapping population	Chromosome	Start	Max	End	ΔSNP
SLC family	SL2.50ch09	3,042,382	3,966,023	4,560,795	0.57

SLL family	SL2.50ch05	9,189,958	33,460,771	59,293,229	0.52
	SL2.50ch06	23,628	23,628	943,353	0.46
	SL2.50ch09	2,105,569	2,865,099	4,800,975	0.56

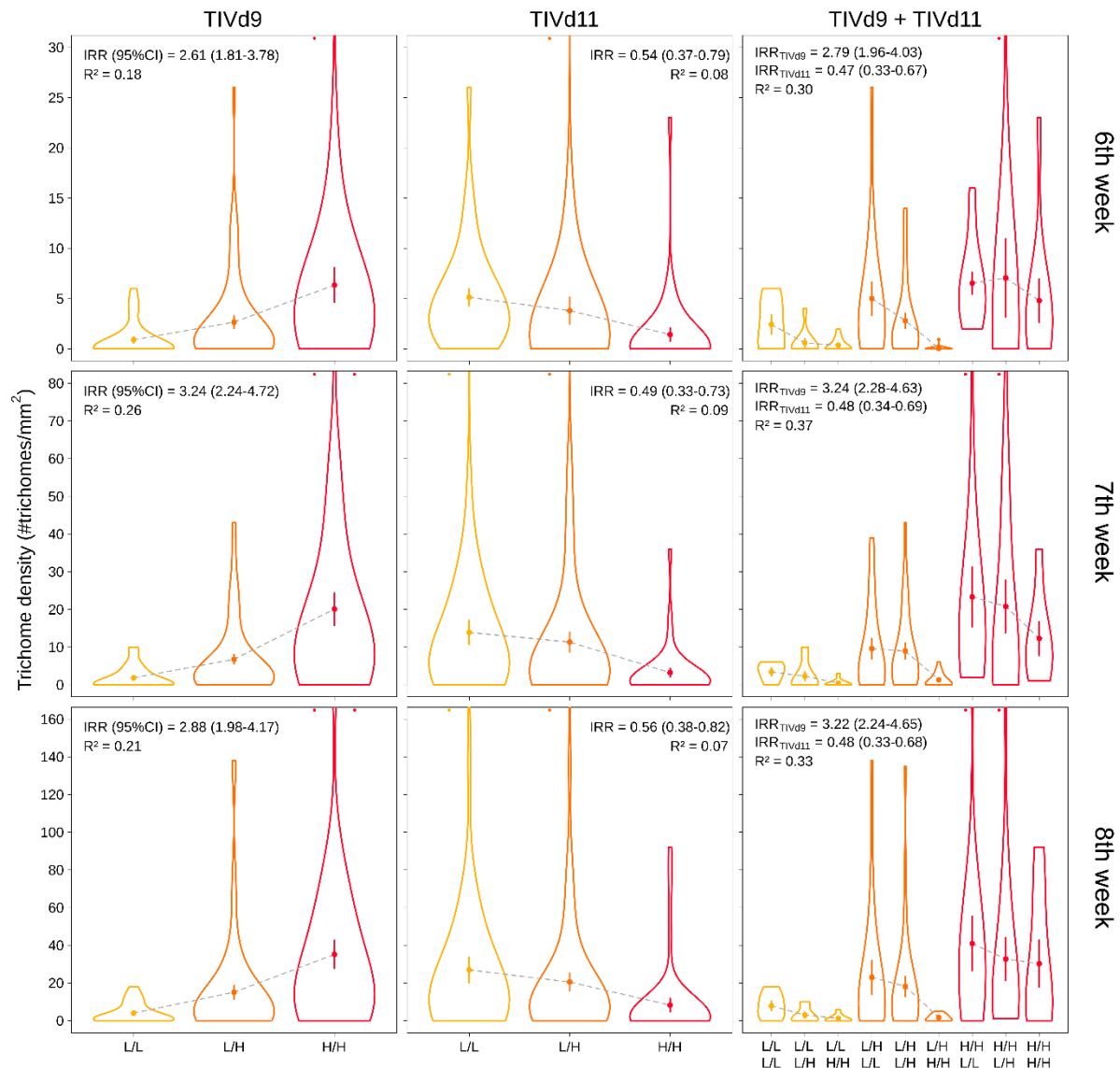


Figure 4. Effect of QTLs TIVd9, TIVd11 and the additive model on trichome density along the three phenotyped weeks (6th, 7th, 8th weeks after transplanting). Incidence rate ratio (IRR) and 95% confidence intervals (values between brackets) and Nagelkerke R² of the negative binomial logistic regression are shown for each combination. Violin plots for the two homozygotes (L/L and H/H) and the heterozygote (L/H) are represented. L refers to an allele

of the low trichome density parental (LA2278) and H to an allele of the high trichome density parental (BGV016047). Mean trichome density and standard error for each genotype is shown inside each violin plot. Notice that dots on top of a genotype violin plot represent plants with a trichome density higher that exceeds the maximum density of the axis

Table 3. Results of the negative binomial logistic regression for the QTLs found by QTL mapping and association analysis. Different statistical models using trichome density (TD) after 6,7 or 8 weeks after transplanting and either a single predictor or an additive model are shown. Independent variables used correspond to the number of alleles of the high density parental at the marker with the highest statistical support. For each model, χ^2 goodness of fit p-value of the model (GOF), Nagelkerke R^2 , explained deviance (D^2) and point estimated of incidence rate ratio and 95% confidence interval are reported. *** p-value < 0.001

QTL mapping							
Phenotype	Model	GOF	R^2	D^2	Variable	IRR (95% CI)	Sig.
TD 6th w.	TIVd9	0.361	0.18	0.16	SL2.50ch09_2542786	2.61 (1.81-3.78)	***
	TIVd11	0.348	0.08	0.07	SL2.50ch11_56297460	0.54 (0.37-0.78)	***
	TIVd9 + TIVd11	0.386	0.30	0.26	SL2.50ch09_2542786	2.79 (1.96-4.03)	***
					SL2.50ch11_56297460	0.47 (0.33-0.67)	***
TD 7th w.	TIVd9	0.145	0.26	0.21	SL2.50ch09_3756370	3.24 (2.24-4.72)	***
	TIVd11	0.144	0.09	0.08	SL2.50ch11_56297460	0.49 (0.33-0.73)	***
	TIVd9 + TIVd11	0.149	0.37	0.29	SL2.50ch09_3756370	3.24 (2.28-4.63)	***
					SL2.50ch11_56297460	0.48 (0.34-0.69)	***
TD 8th w.	TIVd9	0.084	0.21	0.17	SL2.50ch09_2309105	2.88 (1.98-4.17)	***
	TIVd11	0.076	0.07	0.06	SL2.50ch11_56297460	0.56 (0.38-0.82)	***
	TIVd9 + TIVd11	0.091	0.33	0.26	SL2.50ch09_2309105	3.22 (2.24-4.65)	***
					SL2.50ch11_56297460	0.48 (0.33-0.68)	***

Association							
Phenotype	Model	GOF	R ²	D ²	Variable	IRR (95% CI)	Sig.
TD 6th w.	TIVd9	0.361	0.18	0.16	SL2.50ch09_2542786	2.61 (1.81-3.78)	***
TD 7th w.	TIVd9	0.147	0.25	0.20	SL2.50ch09_3727914	3.19 (2.21-4.64)	***
TD 8th w.	TIVd6	0.090	0.22	0.17	SL2.50ch06_154414	2.47 (1.79-3.41)	***
	TIVd9	0.085	0.22	0.18	SL2.50ch09_2492691	2.94 (2.03-4.26)	***
	TIVd6 + TIVd9	0.085	0.32	0.25	SL2.50ch06_154414	1.86 (1.33-2.60)	***
					SL2.50ch09_2492691	2.16 (1.44-3.22)	***

As the QTL mapping was performed using a single interval mapping based on a non parametric method, we decided also to test the association by adjusting to a negative binomial distribution, thus we fitted the density of type IV trichomes with a logistic regression using genotype data as a explanatory factor (Table 2 and Suppl. Fig. 6). After adjusting for multiple comparisons, the QTL in chromosome 9 appeared in all three weeks of phenotyping, coinciding with the TIVd9 region. Interestingly, no region in the chromosome 11 reached a significant association in any case (Suppl. Fig. 6). However, an additional QTL was found when using the phenotypic data of the third week at the beginning of chromosome 6 (TIVd6). This QTL was able to explain a 22% of variance according to Nalgerke's R^2 when considered in a model with a single factor or 32% when considered in an additive model together with TIVd9 (Table 3). IRR for TIVd6 and TIVd9 was 1.86 and 2.16 respectively in the additive model.

A QTLseq analysis comparing F2 individuals with high and low trichome density was also performed. A total of 124,514 SNPs were retained after filtering by depth and MAF. Results show a significant QTL at 99% CI in chromosome 9 overlapping with the location of TIVd9 (Table 2 and Fig. 5A). Additionally, most part of chromosome 5 seems to present a significant deviation of the Δ SNP index above the 95% confidence interval.

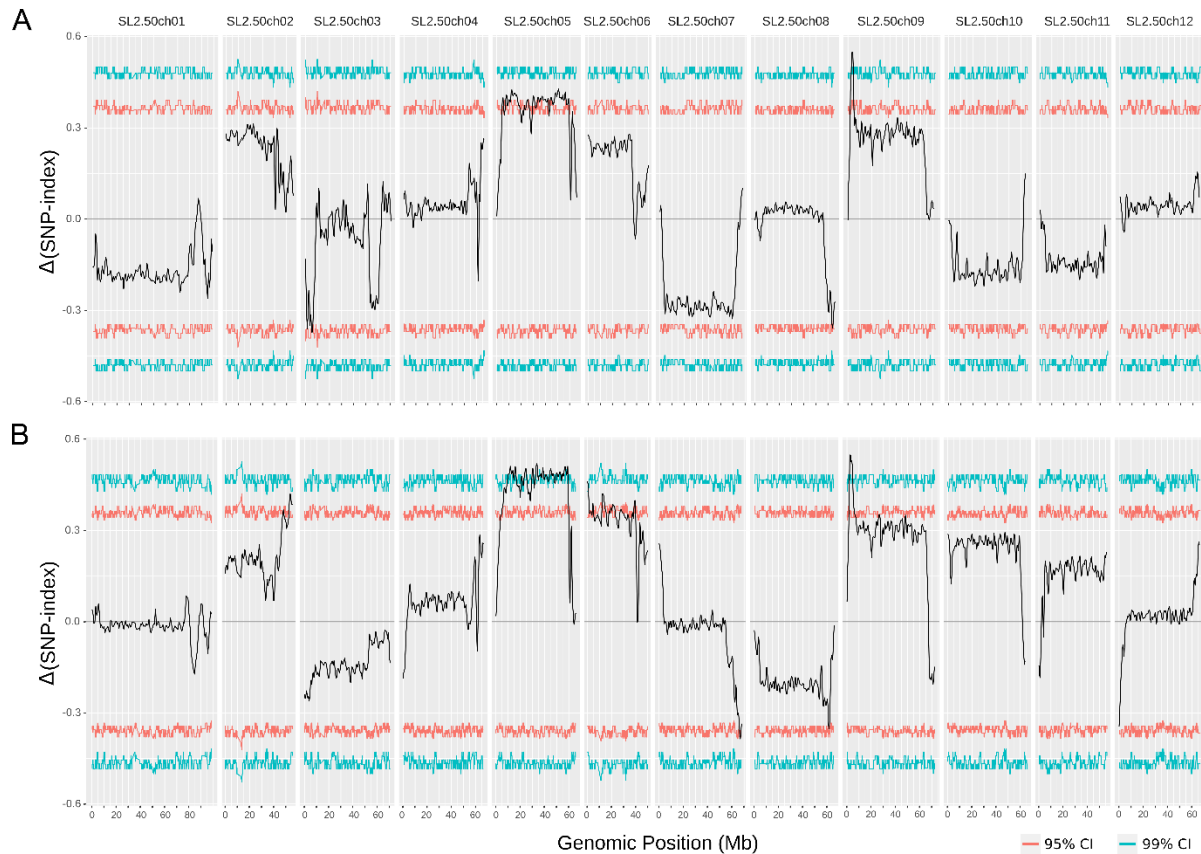


Figure 5. Distribution of $\Delta(\text{SNP-index})$ along the chromosomes based on 1Mb windows for the A) SLC family (BGV016047 x LA2278) and B) SLL family (BGV016047 x MoneyMaker). The 95% and 99% confidence thresholds are shown.

QTL validation using *S. lycopersicum* var. *lycopersicum* as genetic background

The accession BGV016047 was also crossed with MoneyMaker, a *S. lycopersicum* var. *lycopersicum* accession, in order to test if the same genetic control was maintained with a different genetic background. A QTLseq analysis was performed on the F2 offspring of SLL family by bulking individuals with the highest and lowest trichome IV density. A total of 107,112 SNPs were used after filtering by genotype depth and MAF. The results show that the QTL in chromosome 9 was also detected and it overlaps the region detected in SLC family (Table 2 and Fig. 5). Chromosome 5 also presented the same pattern that in the cross with *S. lycopersicum* var. *cerasiforme*, although the signal was stronger. Additionally, a region at the beginning of chromosome 6 also crossed the 99% CI threshold. This region overlaps with the signal found for SLC family in the association analysis when using phenotype data of the 8th week after transplanting. Other regions in chromosome 2, 7 and 8 showed some signal above the 95% confidence interval.

Candidate genes

The region of TIVd9 (SL2.50ch09: 1,966,640-4,371,756 bp) comprises 248 genes (Suppl. Table 1), which include two MYB transcription factors Solyc09g010820, a MYB R3 transcription, and Solyc09g009450, a Myb/SANT-like protein paralogous of *Arabidopsis* AT2G24960. The MYB family is involved in development and cell division. Additionally, two trichome birefringence-like proteins are located (Solyc09g010260 and Solyc09g010270) close to the location where the maximum significance is found (about 3,7 Mb). This gene family has been shown to regulate the density of trichomes [51,52].

Besides these genes related to regulation of expression and cell division, other genes in this region of chromosome 9 are linked to genes related to the production of secondary metabolites that accumulate in glandular trichomes [53]: genes related to acylsugar production like acyltransferases (Solyc09g008520) and glycosyltransferases (Solyc09g008510, Solyc09g009010, Solyc09g010760), fatty acid synthesis like fatty acyl-CoA reductase (Solyc09g009570, Solyc09g009580) and phosphopantetheinyl transferase family protein (Solyc09g009960) or transporters like ABC transporters (Solyc09g009910) that are involved in lipid and secondary metabolites transport. Additionally, other genes related to other plant-defenses mechanisms that also are known to be upregulated in glandular trichomes like subtilisin-like protease (Solyc09g009750) are also found.

In the region of TIVd11 (SL2.50ch11:54,298,873-56,297,460 bp), two out of 265 genes are transcription factors of the family MYB R2R3, Solyc11g072060 paralogous of *Arabidopsis* MYB104 and Solyc11g073120 paralogous of MYB58, and an Agamous-like MADS-box protein (Solyc11g069770). Additionally, the gene for gibberellin 20-oxidase-3 (Solyc11g072310) is also located in this region. This enzyme plays a central role on the synthesis of gibberellins which have a positive effect on trichome formation [54]. Other genes related to acylsugars metabolism that are found in TIVd11 region are Solyc11g072980 and Solyc11g072990 that code for a 3-ketoacyl-CoA synthase, a long fatty acid dehydratase (Solyc11g073130), ABC transporters (Solyc11g069710 and Solyc11g069820), a glycosyltransferase (Solyc11g071230) and an acyltransferase (Solyc11g069680). Interestingly, this last gene presents an SNPs that was classified as a high impact SNPs by SNPEff.

In the detected region of chromosome 6 (from 8,008 to 943,353 bp and 95 genes) three more MYB transcription factors are found, two paralogs of *Arabipodis* MYB48 (Solyc06g005310 and Solyc06g005320) which is related to the synthesis of flavonoids that also accumulate in type IV trichomes and a paralog of MYB59 (Solyc06g005330) which is involved in the regulation of

cell cycle progression and root growth. The region of chromosome 5 covers a high proportion of this chromosome and includes 843 genes (from 9,189,958 bp to 59,293,229) (Suppl. Table 1). Several genes related to acylsugars are located in this region like several acyltransferases (Solyc05g025890, Solyc05g039950, Solyc05g047610, Solyc05g047640, Solyc05g016030), long chain acyl-CoA synthetase 2 (Solyc05g041520) or ABC transporters (Solyc05g018510, Solyc05g023940). Also defense related proteins like glutathione S-transferase (Solyc05g026210), kirola-like (Solyc05g046140, Solyc05g046150, Solyc05g046160, Solyc05g046170, Solyc05g046210, Solyc05g046220) or FLOWERING LOCUS D (Solyc05g016300) that are upregulated in trichomes [53]. Besides that, transcription factors including a Myb/SANT-like (Solyc05g018830) and two MADS-box transcription factor (Solyc05g015720 and Solyc05g015730) are also located in this region as well as a protein trichome birefringence-like (Solyc05g019980).

Discussion

Presence of glandular trichomes, especially type IV trichomes, is involved in pest resistances, which is a desirable characteristic for more sustainable culture management. Glandular type IV trichomes have been described in wild *Solanum* species such as *S. habrochaites*, *S. pennellii*, *S. galapagense*, *S. cheesmaniae* and *S. neorickii*, with accessions reaching densities higher than 70 trichomes/mm² [16,19,20,23]. Unfortunately, these species are distant relatives from cultivated tomatoes and their use as sources for tomato improvement is limited due to the incorporation of undesirable characters and incompatibility problems. *S. pimpinellifolium* is the closest wild relative of cultivated tomato and a common source of alleles for breeders. However, until now, presence of type IV trichomes has only been described at low densities (< 20 trichomes/mm²) [27,55,56]. In the present study, we have characterized a *S. pimpinellifolium* accession (BGV016047) with a high density of type IV trichomes (from 9.5 trichomes/mm² at 48 days after transplanting, to a maximum median of 118 trichomes/mm²) that could be of great interest for tomato genetic improvement. Density of this type of trichomes is known to be dependent on developmental conditions, such as leaf and plant age [30]. Accordingly, we have observed a density increase with plant age, reaching maximum trichome densities of 240 trichomes/mm² in young leaves (118 trichomes/mm² maximum median density) several weeks after transplanting, while trichome density tends to decrease as leaves continue growing and the foliar area increases. Besides plant and leaf age, some degree of variability has been observed between measures and seasons, which is likely to be the result of environmental conditions or resources availability [30,32].

K-seq is a reduced representation library sequencing method that provides a high-throughput and cost-effective genotyping alternative to existing methods. K-seq genotyping achieves enough genome coverage and depth for many genotyping needs and produces a high number of SNPs. We have used the K-seq genotyping to develop here an ultra-dense genetic map with almost 150 thousand SNPs. K-seq could be easily adapted to different necessities varying the primer number and the K-mer sequences [34]. In this study, we have genotyped 96 samples per HiSeq2500 lane, resulting in a low cost by sample. However, sample cost could be further reduced using more multiplexing indexes and reads by lane as the number of SNPs obtained exceeds the needs for the development of a detailed F2 map. So, K-seq is a cheap and easy methodology to genotype that can be used by any laboratory.

Using K-seq we have been able to build an ultra-dense genetic map with 147,326 markers with an average distance between markers of 0.2 cM that has allowed us to perform a detailed mapping of the phenotype density of type IV trichomes in two F2 populations and to identify genes candidates for this trait. The inheritance of type IV trichomes has been previously addressed in other accessions and species, but the genes are still unknown. More information is available regarding QTLs related to insect resistance and metabolism of acylsugars, the most common chemical accumulated in type IV glandular trichomes, that also affects density and development of trichomes [20,24,25,57–59]. In *S. galapagense* and *S. habrochaites* the density of type IV trichomes was considered to be controlled by an incompletely recessive allele in a major locus [16,60]. In *S. pennelli* and *S. pimpinellifolium*, results supported two dominant unlinked genes [21,26,31]. The pattern of inheritance found in the two segregating families derived for this study would also support the involvement of a limited number of genes.

By using the information of the two populations derived from crosses between the accession BGV016047 and two different genetic backgrounds, we have detected a new major QTL at the beginning of chromosome 9 (TIVd9) in *S. pimpinellifolium*, which accounts up to 26% of the explained variance, each allele increasing the expected trichome IV density by 3. Additionally, a minor QTL has been found in chromosome 11 (TIVd11) and likely in chromosomes 6 and 5. The major QTL TIVd9 has not been previously detected in *S. pimpinellifolium*, where only two QTLs likely involved in trichome formation have been described in the telomeric region of chromosome 2 [28]. This region in chromosome 9 has also been linked to type IV trichome density in *S. galapagense* although the percentage of explained variance was much lower (2.8-8.3%) [61]. Another region in chromosome 9, adjacent to TIVd9 and with minor effects on trichome density, has also been detected in *S. galapagense* [23]. In *S. habrochaites*, QTLs for type IV trichome density in chromosome 9 have also been described. In this species, three QTLs for trichome density in chromosome 9, 10

and 11 were detected, all together explaining the 22% of the variance [22]. The QTL in chromosome 9 overlaps our QTL TIVd9, although explains a lower percentage of the variance (11% of the variance). Interestingly, this QTL is also associated with a lower oviposition of *Bemisia tabaci*. The QTL in chromosome 11 does not overlap our QTL but is located 4Mb away from the lower limit of TIVd11, however this QTL was only found to be significant in an additive model with the other two QTLs, but not when considered alone. Another QTL for type IV trichomes was also found in chromosome 9 in *S. galapagense* explaining 8% of variance [23] although it is located 61 Mb away from TIVd9, so it is unlikely that both QTLs are the same. Moreover, in a posterior study with RIL populations this QTL was not found [61]. However, the chromosomal region of TIVd9 does overlap with a QTL for adult survival to white fly [23]. TIVd5 and TIVd6 overlap with previously detected QTLs for trichome density in *S. penelli* [20,62]. Additionally TIVd5 was also previously found to be associated with acylsugar accumulation levels [20].

Several types of genes have been found to control the development of trichomes [63,64], including MYB transcription factor factors that are known to regulate positively and negatively trichome development [65,66], and genes related to the production of acylsugars [53] or plant defenses. Several strong gene candidates for TIVd9 have been found in the QTL region. Besides several enzymes related to acylsugar production, two MYB genes (Solyc09g010820 and Solyc09g009450) are located within the range and especially two trichome birefringence-like proteins (Solyc09g010260 and Solyc09g010270). This gene family influences synthesis of cellulose on the secondary wall, affects resistance and has been shown to regulate the density of trichomes [51,52]. The different gene members are likely to provide a wide range of activities and tissue specificities [52]. In the region of TIVd11 another two MYB R2R3 genes have been found (Solyc11g072060 and Solyc11g073120) and an Agamous-like MADS-box protein (Solyc11g069770). Interestingly, *AGAMOUS* has been shown to negatively interfere with the development of trichomes in *Arabidopsis* [67], which is the same effect found in our population.

Conclusions

These results increase the knowledge of the genetic factors involved in the development of trichomes as several strong candidate regions have been detected, so further studies can address the effect of such genes. The two genetic backgrounds used have detected some different QTL regions, showing the complex mechanism involved in the control of trichome development. Additionally, our results contribute to the development of tomato lines with a higher density of type IV trichomes, and likely insect resistance [15]. Several efforts have been made to develop introgressed lines with type IV trichomes enriched in acylsugars [25,56,58].

We found a major QTL that has not been previously reported in *S. pimpinellifolium* that increases by a factor of 9 the density of trichomes. This accession contains probably different alleles of the QTLs detected on other species. The main advantage of this accession is that this QTL could be easy to introgress due to the close genetic relationship between *S. pimpinellifolium* and cultivated tomato. Besides, this study is the first example of the utility and efficiency of a new genotyping methodology, K-seq, that with a low read number by sample allows the genotype of millions of bases at low cost. As has been shown, the methodology is easy and robust enough to generate high density maps and to do QTL-seq.

Acknowledgments

We would like to thank the Centro de Experiencias Cajamar de Paiporta (Valencia, Spain) for its excellent work done in growing the tomato plants in their greenhouses. We thank TGRC and COMAV genebanks for providing seeds.

Funding

This research was supported by the National Natural Science Foundation of USA Varitome project (NSF IOS 1564366).

Authors contributions

Conceptualization, JC, MJD, JMP; Methodology, JC, JMP; Investigation, EMN, EGP, AGP; Formal Analysis, JMP, PZ, JB, EMN.; Writing – Original Draft Preparation, JC, JMP, EMN; Writing – Review & Editing, JB, EK, MJD; Supervision, MJD, JC; Funding Acquisition, JC, MJD, EK.

Conflict of interest

The authors declare no conflict of interest.

Bibliography

1. McDowell, E.T.; Kapteyn, J.; Schmidt, A.; Li, C.; Kang, J.-H.; Descour, A.; Shi, F.; Larson, M.; Schillmiller, A.; An, L.; et al. Comparative functional genomic analysis of *Solanum glandular* trichome types. *Plant Physiol.* **2011**, *155*, 524–539, doi:10.1104/pp.110.167114.
2. Cho, K.-S.; Kwon, M.; Cho, J.-H.; Im, J.-S.; Park, Y.-E.; Hong, S.-Y.; Hwang, I.-T.; Kang, J.-H. Characterization of trichome morphology and aphid resistance in cultivated and

- wild species of potato. *Horticulture, Environment, and Biotechnology* **2017**, *58*, 450–457, doi:10.1007/s13580-017-0078-4.
3. Woodcock, B.A.; Garratt, M.P.D.; Powney, G.D.; Shaw, R.F.; Osborne, J.L.; Soroka, J.; Lindström, S.A.M.; Stanley, D.; Ouvrard, P.; Edwards, M.E.; et al. Meta-analysis reveals that pollinator functional diversity and abundance enhance crop pollination and yield. *Nat. Commun.* **2019**, *10*, 1481, doi:10.1038/s41467-019-09393-6.
 4. Glas, J.J.; Schimmel, B.C.J.; Alba, J.M.; Escobar-Bravo, R.; Schuurink, R.C.; Kant, M.R. Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *Int. J. Mol. Sci.* **2012**, *13*, 17077–17103, doi:10.3390/ijms131217077.
 5. Huchelmann, A.; Boutry, M.; Hachez, C. Plant Glandular Trichomes: Natural Cell Factories of High Biotechnological Interest. *Plant Physiol.* **2017**, *175*, 6–22, doi:10.1104/pp.17.00727.
 6. Peter, A.J.; Shanower, T.G.; Romeis, J. The role of plant trichomes in insect resistance: a selective review. *Phytophaga* **1995**, *7*, 41–63.
 7. Liedl, B.E.; Lawson, D.M.; White, K.K.; Shapiro, J.A.; Cohen, D.E.; Carson, W.G.; Trumble, J.T.; Mutschler, M.A. Acylsugars of Wild Tomato *Lycopersicon pennellii* Alters Settling and Reduces Oviposition of *Bemisia argentifolii* (Homoptera: Aleyrodidae). *J. Econ. Entomol.* **1995**, *88*, 742–748, doi:10.1093/jee/88.3.742.
 8. Peiffer, M.; Tooker, J.F.; Luthe, D.S.; Felton, G.W. Plants on early alert: glandular trichomes as sensors for insect herbivores. *New Phytol.* **2009**, *184*, 644–656, doi:10.1111/j.1469-8137.2009.03002.x.
 9. Leong, B.J.; Lybrand, D.B.; Lou, Y.-R.; Fan, P.; Schillmiller, A.L.; Last, R.L. Evolution of metabolic novelty: A trichome-expressed invertase creates specialized metabolic diversity in wild tomato. *Sci Adv* **2019**, *5*, eaaw3754, doi:10.1126/sciadv.aaw3754.
 10. Channarayappa, C.; Shivashankar, G.; Muniyappa, V.; Frist, R.H. Resistance of *Lycopersicon* species to *Bemisia tabaci*, a tomato leaf curl virus vector. *Can. J. Bot.* **1992**, *70*, 2184–2192, doi:10.1139/b92-270.
 11. Alba, J.M.; Montserrat, M.; Fernández-Muñoz, R. Resistance to the two-spotted spider mite (*Tetranychus urticae*) by acylsucroses of wild tomato (*Solanum pimpinellifolium*) trichomes studied in a recombinant inbred line population. *Exp. Appl. Acarol.* **2009**, *47*, 35–47.
 12. Maluf, W.R.; Campos, G.A.; das Graças Cardoso, M. Relationships between trichome types and spider mite (*Tetranychus evansi*) repellence in tomatoes with respect to foliar zingiberene contents. *Euphytica* **2001**, *121*, 73–80, doi:10.1023/A:1012067505361.
 13. Oriani, M.A. de G.; Vendramim, J.D. Influence of trichomes on attractiveness and ovipositional preference of *Bemisia tabaci* (Genn.) B biotype (Homoptera: Aleyrodidae) on tomato genotypes. *Neotrop. Entomol.* **2010**, *39*, 1002–1007, doi:10.1590/s1519-566x2010000600024.
 14. Firdaus, S. Identification of whitefly resistance in tomato and hot pepper. **2012**.
 15. Rakha, M.; Hanson, P.; Ramasamy, S. Identification of resistance to *Bemisia tabaci* Genn. in closely related wild relatives of cultivated tomato based on trichome type analysis and choice and no-choice assays. *Genetic Resources and Crop Evolution* **2017**, *64*, 247–260.
 16. Andrade, M.C.; da Silva, A.A.; Neiva, I.P.; Oliveira, I.R.C.; De Castro, E.M.; Francis, D.M.; Maluf, W.R. Inheritance of type IV glandular trichome density and its association with whitefly resistance from *Solanum galapagense* accession LA1401. *Euphytica* **2017**, *213*, 52, doi:10.1007/s10681-016-1792-1.
 17. Gurr, G.M.; McGrath, D. Foliar pubescence and resistance to potato moth, *Phthorimaea operculella*, in *Lycopersicon hirsutum*. *Entomol. Exp. Appl.* **2002**, *103*, 35–41, doi:10.1046/j.1570-7458.2002.00960.x.
 18. Fan, P.; Leong, B.J.; Last, R.L. Tip of the trichome: evolution of acylsugar metabolic diversity in Solanaceae. *Curr. Opin. Plant Biol.* **2019**, *49*, 8–16, doi:10.1016/j.pbi.2019.03.005.
 19. Guo, Z.; Weston, P.A.; Snyder, J.C. Repellency to two-spotted spider mite, *Tetranychus urticae* Koch, as related to leaf surface chemistry of *Lycopersicon hirsutum* accessions.

- J. Chem. Ecol.* **1993**, *19*, 2965–2979, doi:10.1007/BF00980596.
20. Blauth, S.L.; Churchill, G.A.; Mutschler, M.A. Identification of quantitative trait loci associated with acylsugar accumulation using intraspecific populations of the wild tomato, *Lycopersicon pennellii*. *Theor. Appl. Genet.* **1998**, *96*, 458–467, doi:10.1007/s001220050762.
 21. Lemke, C.A.; Mutschler, M.A. Inheritance of glandular trichomes in crosses between *Lycopersicon esculentum* and *Lycopersicon pennellii*. *J. Am. Soc. Hortic. Sci.* **1984**, *109*, 592–596.
 22. Momotaz, A.; Scott, J.W.; Schuster, D.J. Identification of quantitative trait loci conferring resistance to *Bemisia tabaci* in an F2 population of *Solanum lycopersicum* × *Solanum habrochaites* accession LA1777. *J. Am. Soc. Hortic. Sci.* **2010**, *135*, 134–142.
 23. Firdaus, S.; van Heusden, A.W.; Hidayati, N.; Supena, E.D.J.; Mumm, R.; de Vos, R.C.H.; Visser, R.G.F.; Vosman, B. Identification and QTL mapping of whitefly resistance components in *Solanum galapagense*. *Theor. Appl. Genet.* **2013**, *126*, 1487–1501, doi:10.1007/s00122-013-2067-z.
 24. Mutschler, M.A.; Doerge, R.W.; Liu, S.C.; Kuai, J.P.; Liedl, B.E.; Shapiro, J.A. QTL analysis of pest resistance in the wild tomato *Lycopersicon pennellii*: QTLs controlling acylsugar level and composition. *Theor. Appl. Genet.* **1996**, *92*, 709–718, doi:10.1007/BF00226093.
 25. Leckie, B.M.; De Jong, D.M.; Mutschler, M.A. Quantitative trait loci increasing acylsugars in tomato breeding lines and their impacts on silverleaf whiteflies. *Mol. Breed.* **2012**, *30*, 1621–1634, doi:10.1007/s11032-012-9746-3.
 26. Fernández-Muñoz, R.; Salinas, M.; Álvarez, M.; Cuartero, J. Inheritance of resistance to two-spotted spider mite and glandular leaf trichomes in wild tomato *Lycopersicon pimpinellifolium* (Jusl.) Mill. *J. Am. Soc. Hortic. Sci.* **2003**, *128*, 188–195.
 27. Firdaus, S.; van Heusden, A.W.; Hidayati, N.; Supena, E.D.J.; Visser, R.G.F.; Vosman, B. Resistance to *Bemisia tabaci* in tomato wild relatives. *Euphytica* **2012**, *187*, 31–45, doi:10.1007/s10681-012-0704-2.
 28. Salinas, M.; Capel, C.; Alba, J.M.; Mora, B.; Cuartero, J.; Fernández-Muñoz, R.; Lozano, R.; Capel, J. Genetic mapping of two QTL from the wild tomato *Solanum pimpinellifolium* L. controlling resistance against two-spotted spider mite (*Tetranychus urticae* Koch). *Theor. Appl. Genet.* **2013**, *126*, 83–92.
 29. Vendemiatti, E.; Zsögön, A.; Silva, G.F.F.E.; de Jesus, F.A.; Cutri, L.; Figueiredo, C.R.F.; Tanaka, F.A.O.; Nogueira, F.T.S.; Peres, L.E.P. Loss of type-IV glandular trichomes is a heterochronic trait in tomato and can be reverted by promoting juvenility. *Plant Sci.* **2017**, *259*, 35–47, doi:10.1016/j.plantsci.2017.03.006.
 30. Gurr, G.M.; McGrath, D. Effect of plant variety, plant age and photoperiod on glandular pubescence and host-plant resistance to potato moth (*Phthorimaea operculella*) in *Lycopersicon* spp. *Ann. Appl. Biol.* **2001**, *138*, 221–230, doi:10.1111/j.1744-7348.2001.tb00106.x.
 31. Saeidi, Z.; Mallik, B.; Kulkarni, R.S. Inheritance of glandular trichomes and two-spotted spider mite resistance in cross *Lycopersicon esculentum* “Nandi” and *L. pennellii* “LA2963”. *Euphytica* **2007**, *154*, 231–238.
 32. Wilkens, R.T.; Shea, G.O.; Halbreich, S.; Stamp, N.E. Resource availability and the trichome defenses of tomato plants. *Oecologia* **1996**, *106*, 181–191, doi:10.1007/BF00328597.
 33. Weston, P.A.; Johnson, D.A.; Burton, H.T.; Snyder, J.C. Trichome secretion composition, trichome densities, and spider mite resistance of ten accessions of *Lycopersicon hirsutum*. *J. Am. Soc. Hortic. Sci.* **1989**.
 34. Ziarsolo, P.; Hasing, T.; Hilario, R.; Garcia-Carpintero, V.; Blanca, J.; Bombarely, A.; Cañizares, J. K-seq, an affordable, reliable, and open Klenow NGS-based genotyping technology. bioRxiv 2020.11.16.360446; doi: <https://doi.org/10.1101/2020.11.16.360446>
 35. Mata-Nicolás, E.; Montero-Pau, J.; Gimeno-Paez, E.; Garcia-Carpintero, V.; Ziarsolo, P.; Menda, N.; Mueller, L.A.; Blanca, J.; Cañizares, J.; van der Knaap, E.; et al. Exploiting the diversity of tomato: the development of a phenotypically and genetically

- detailed germplasm collection. *Hortic Res* **2020**, *7*, 66, doi:10.1038/s41438-020-0291-7.
36. Doyle, J.J.; Doyle, J.L. Isolation of plant DNA from fresh tissue. *Focus* **1990**, *12*, 13–15.
 37. Andrews, S.; Krueger, F.; Segonds-Pichon, A.; Biggins, L.; Krueger, C.; Wingett, S. FastQC: a quality control tool for high throughput sequence data 2012.
 38. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120, doi:10.1093/bioinformatics/btu170.
 39. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bio.GN]* 2013.
 40. Garrison, E.; Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv [q-bio.GN]* 2012.
 41. Takagi, H.; Abe, A.; Yoshida, K.; Kosugi, S.; Natsume, S.; Mitsuoka, C.; Uemura, A.; Utsushi, H.; Tamiru, M.; Takuno, S.; et al. QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* **2013**, *74*, 174–183, doi:10.1111/tbj.12105.
 42. Miao, C.; Fang, J.; Li, D.; Liang, P.; Zhang, X.; Yang, J.; Schnable, J.C.; Tang, H. Genotype-Corrector: improved genotype calls for genetic mapping in F2 and RIL populations. *Sci. Rep.* **2018**, *8*, 10088, doi:10.1038/s41598-018-28294-0.
 43. Broman, K.W.; Wu, H.; Sen, S.; Churchill, G.A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **2003**, *19*, 889–890, doi:10.1093/bioinformatics/btg112.
 44. Taylor, J.D.; Butler, D. ASMap: Linkage map construction using the MSTmap algorithm. *R package version 0.3--3* **2014**.
 45. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S*; Fourth.; Springer: New York, 2002;.
 46. Delignette-Muller, M.L.; Dutang, C. fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software* 2015, *64*, 1–34.
 47. Barbosa, A.M.; Brown, J.A.; Jimenez-Valverde, A.; Real, R. modEvA: Model Evaluation and Analysis 2016.
 48. Mansfeld, B.N.; Grumet, R. QTLseqr: An R Package for Bulk Segregant Analysis with Next-Generation Sequencing. *Plant Genome* **2018**, *11*, doi:10.3835/plantgenome2018.01.0006.
 49. Yekutieli, D.; Benjamini, Y. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **2001**, *29*, 1165–1188, doi:10.1214/aos/1013699998.
 50. R Core Team *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015;.
 51. Potikha, T.; Delmer, D.P. A mutant of *Arabidopsis thaliana* displaying altered patterns of cellulose deposition. *Plant J.* **1995**, *7*, 453–460, doi:10.1046/j.1365-313X.1995.7030453.x.
 52. Gao, Y.; He, C.; Zhang, D.; Liu, X.; Xu, Z.; Tian, Y.; Liu, X.-H.; Zang, S.; Pauly, M.; Zhou, Y.; et al. Two Trichome Birefringence-Like Proteins Mediate Xylan Acetylation, Which Is Essential for Leaf Blight Resistance in Rice. *Plant Physiol.* **2017**, *173*, 470–481, doi:10.1104/pp.16.01618.
 53. Mandal, S.; Ji, W.; McKnight, T.D. Candidate Gene Networks for Acylsugar Metabolism and Plant Defense in Wild Tomato *Solanum pennellii*. *Plant Cell* **2020**, *32*, 81–99, doi:10.1105/tpc.19.00552.
 54. Chien, J.C.; Sussex, I.M. Differential regulation of trichome formation on the adaxial and abaxial leaf surfaces by gibberellins and photoperiod in *Arabidopsis thaliana* (L.) Heynh. *Plant Physiol.* **1996**, *111*, 1321–1328, doi:10.1104/pp.111.4.1321.
 55. Rakha, M.; Zekeya, N.; Sevgan, S.; Musembi, M.; Ramasamy, S.; Hanson, P. Screening recently identified whitefly/spider mite-resistant wild tomato accessions for resistance to *Tuta absoluta*. *Plant Breed.* **2017**, *136*, 562–568.
 56. Rodríguez-López, M.J.; Moriones, E.; Fernández-Muñoz, R. An Acylsucrose-Producing Tomato Line Derived from the Wild Species *Solanum pimpinellifolium* Decreases Fitness of the Whitefly *Trialeurodes vaporariorum*. *Insects* **2020**, *11*, doi:10.3390/insects11090616.
 57. Leckie, B.M.; De Jong, D.M.; Mutschler, M.A. Quantitative trait loci regulating sugar

- moiety of acylsugars in tomato. *Mol. Breed.* **2013**, *31*, 957–970, doi:10.1007/s11032-013-9849-5.
58. Smeda, J.R.; Schillmiller, A.L.; Last, R.L.; Mutschler, M.A. Introgression of acylsugar chemistry QTL modifies the composition and structure of acylsugars produced by high-accumulating tomato lines. *Mol. Breed.* **2016**, *36*, 160, doi:10.1007/s11032-016-0584-6.
 59. Smeda, J.R.; Schillmiller, A.L.; Anderson, T.; Ben-Mahmoud, S.; Ullman, D.E.; Chappell, T.M.; Kessler, A.; Mutschler, M.A. Combination of Acylglucose QTL reveals additive and epistatic genetic interactions and impacts insect oviposition and virus infection. *Mol. Breed.* **2017**, *38*, 3, doi:10.1007/s11032-017-0756-z.
 60. Freitas, J.A.; Maluf, W.R.; das Graças Cardoso, M.; Gomes, L.A.A.; Bearzotti, E. Inheritance of foliar zingiberene contents and their relationship to trichome densities and whitefly resistance in tomatoes. *Euphytica* **2002**, *127*, 275–287, doi:10.1023/A:1020239512598.
 61. Vosman, B.; Kashaninia, A.; Van't Westende, W.; Meijer-Dekens, F.; van Eekelen, H.; Visser, R.G.F.; de Vos, R.C.H.; Voorrips, R.E. QTL mapping of insect resistance components of *Solanum galapagense*. *Theor. Appl. Genet.* **2019**, *132*, 531–541, doi:10.1007/s00122-018-3239-7.
 62. Maliepaard, C.; Bas, N.; van Heusden, S.; Kos, J.; Pet, G.; Verkerk, R.; Vrieunk, R.; Zabel, P.; Lindhout, P. Mapping of QTLs for glandular trichome densities and *Trialeurodes vaporariorum* (greenhouse whitefly) resistance in an F₂ from *Lycopersicon esculentum* × *Lycopersicon hirsutum* f. *glabratum*. *Heredity* **1995**, *75*, 425–433.
 63. Fambrini, M.; Pugliesi, C. The Dynamic Genetic-Hormonal Regulatory Network Controlling the Trichome Development in Leaves. *Plants* **2019**, *8*, doi:10.3390/plants8080253.
 64. Doroshkov, A.V.; Konstantinov, D.K.; Afonnikov, D.A.; Gunbin, K.V. The evolution of gene regulatory networks controlling *Arabidopsis thaliana* L. trichome development. *BMC Plant Biol.* **2019**, *19*, 53, doi:10.1186/s12870-019-1640-2.
 65. Kirik, V.; Lee, M.M.; Wester, K.; Herrmann, U.; Zheng, Z.; Oppenheimer, D.; Schiefelbein, J.; Hulskamp, M. Functional diversification of MYB23 and GL1 genes in trichome morphogenesis and initiation. *Development* **2005**, *132*, 1477–1485, doi:10.1242/dev.01708.
 66. Gan, L.; Xia, K.; Chen, J.-G.; Wang, S. Functional characterization of TRICHOMELESS2, a new single-repeat R3 MYB transcription factor in the regulation of trichome patterning in *Arabidopsis*. *BMC Plant Biol.* **2011**, *11*, 176, doi:10.1186/1471-2229-11-176.
 67. Ó'Maoiléidigh, D.S.; Wuest, S.E.; Rae, L.; Raganelli, A.; Ryan, P.T.; Kwaśniewska, K.; Das, P.; Lohan, A.J.; Loftus, B.; Graciet, E.; et al. Control of reproductive floral organ identity specification in *Arabidopsis* by the C function regulator AGAMOUS. *Plant Cell* **2013**, *25*, 2482–2503.

Tables

Table 1. Summary statistics for the genetic map built using the F₂ generation of the cross BGV016047 × LA2278.

	#SNPs	#Bin markers	Length	Avg. spacing	Max. spacing
SL2.50ch01	23,537	1,248	153.4	0.1	3.7
SL2.50ch02	10,368	662	130.3	0.2	1.8

SL2.50ch03	8,300	479	165.2	0.3	8.5
SL2.50ch04	9,482	534	123.9	0.2	1.3
SL2.50ch05	15,374	856	121.4	0.1	7.9
SL2.50ch06	8,375	552	123.7	0.2	11.2
SL2.50ch07	17,919	770	110.9	0.1	1.9
SL2.50ch08	7,555	563	164.7	0.3	13.3
SL2.50ch09	11,342	725	125.9	0.2	3.7
SL2.50ch10	14,855	788	119.9	0.2	5.2
SL2.50ch11	9,925	567	101.8	0.2	2.8
SL2.50ch12	10,294	541	159.2	0.3	6.2
Overall	147,326	8,285	1,600.4	0.2	13.3

Table 2. QTL location based on different methods using SLC family and SLL family mapping populations. For each QTL, chromosome, start and end position, location of the most significant value of the region (Max) and LOD, adjusted p-value or Δ SNP value for that location are reported,

QTL mapping (SLC family)					
Phenotype	Chr	Start	Max	End	LOD
Trichome density 6th week	SL2.50ch09	2,055,839	2,542,786	4,250,325	6.24
	SL2.50ch11	54,298,873	56,297,460	56,297,460	5.62
Trichome density 7th week	SL2.50ch09	2,078,180	3,756,370	4,250,325	6.33
	SL2.50ch11	55,093,420	56,297,460	56,297,460	3.69
Trichome density 8th week	SL2.50ch09	1,966,640	2,309,105	4,371,756	5.36
	SL2.50ch11	55,548,327	56,297,460	56,297,460	3.83
Association (SLC family)					
Phenotype	Chromosome	Start	Max	End	Adj. pval.

Trichome density 6th week	SL2.50ch09	2,309,105	2,542,786	2,545,171	0.0090
Trichome density 7th week	SL2.50ch09	2,078,180	3,727,914	4,969,831	7.7·10 ⁻⁶
Trichome density 8th week	SL2.50ch06	8,008	154,414	615,838	0.0002
	SL2.50ch09	2,069,769	2,492,691	5,693,896	4.6·10 ⁻⁵
QTLseq					
Mapping population	Chromosome	Start	Max	End	ΔSNP
SLC family	SL2.50ch09	3,042,382	3,966,023	4,560,795	0.57
SLL family	SL2.50ch05	9,189,958	33,460,771	59,293,229	0.52
	SL2.50ch06	23,628	23,628	943,353	0.46
	SL2.50ch09	2,105,569	2,865,099	4,800,975	0.56

Table 3. Results of the negative binomial logistic regression for the QTLs found by QTL mapping and association analysis. Different statistical models using trichome density (TD) after 6,7 or 8 weeks after transplanting and either a single predictor or an additive model are shown. Independent variables used correspond to the number of alleles of the high density parental at the marker with the highest statistical support. For each model, χ^2 goodness of fit p-value of the model (GOF), Nagelkerke R^2 , explained deviance (D^2) and point estimated of incidence rate ratio and 95% confidence interval are reported. *** p-value < 0.001

QTL mapping							
Phenotype	Model	GOF	R^2	D^2	Variable	IRR (95% CI)	Sig.
TD 6th w.	TIVd9	0.361	0.18	0.16	SL2.50ch09_2542786	2.61 (1.81-3.78)	***
	TIVd11	0.348	0.08	0.07	SL2.50ch11_56297460	0.54 (0.37-0.78)	***
	TIVd9 + TIVd11	0.386	0.30	0.26	SL2.50ch09_2542786	2.79 (1.96-4.03)	***

					SL2.50ch11_56297460	0.47 (0.33-0.67)	***
TD 7th w.	TIVd9	0.145	0.26	0.21	SL2.50ch09_3756370	3.24 (2.24-4.72)	***
	TIVd11	0.144	0.09	0.08	SL2.50ch11_56297460	0.49 (0.33-0.73)	***
	TIVd9 + TIVd11	0.149	0.37	0.29	SL2.50ch09_3756370	3.24 (2.28-4.63)	***
					SL2.50ch11_56297460	0.48 (0.34-0.69)	***
TD 8th w.	TIVd9	0.084	0.21	0.17	SL2.50ch09_2309105	2.88 (1.98-4.17)	***
	TIVd11	0.076	0.07	0.06	SL2.50ch11_56297460	0.56 (0.38-0.82)	***
	TIVd9 + TIVd11	0.091	0.33	0.26	SL2.50ch09_2309105	3.22 (2.24-4.65)	***
					SL2.50ch11_56297460	0.48 (0.33-0.68)	***
Association							
Phenotype	Model	GOF	R²	D²	Variable	IRR (95% CI)	Sig.
TD 6th w.	TIVd9	0.361	0.18	0.16	SL2.50ch09_2542786	2.61 (1.81-3.78)	***
TD 7th w.	TIVd9	0.147	0.25	0.20	SL2.50ch09_3727914	3.19 (2.21-4.64)	***
TD 8th w.	TIVd6	0.090	0.22	0.17	SL2.50ch06_154414	2.47 (1.79-3.41)	***
	TIVd9	0.085	0.22	0.18	SL2.50ch09_2492691	2.94 (2.03-4.26)	***
	TIVd6 + TIVd9	0.085	0.32	0.25	SL2.50ch06_154414	1.86 (1.33-2.60)	***
					SL2.50ch09_2492691	2.16 (1.44-3.22)	***

Figure legends

Figure 1. Effect of plant age (A), leaf age (B) and growing season (C) on type IV trichome density.

Figure 2. Density of type IV trichomes in the segregating populations for the last week of phenotyping for the SLC family (BGV016047 x LA2278) (A) and SLL family (BGV016047 x MoneyMaker) (B)

Figure 3. Results of non-parametric single interval mapping QTL analysis for each of the three weeks phenotyped. A) 6th week after transplanting, B) 7th week and C) 8th weeks.

Figure 4. Effect of QTLs TIVd9, TIVd11 and the additive model on trichome density along the three phenotyped weeks (6th, 7th, 8th weeks after transplanting. Incidence rate ratio (IRR) and 95% confidence intervals (values between brackets) and Nagelkerke R^2 of the negative binomial logistic regression are shown for each combination. Violin plots for the two homozygotes (L/L and H/H) and the heterozygote (L/H) are represented. L refers to an allele of the low trichome density parental (LA2278) and H to an allele of the high trichome density parental (BGV016047). Mean trichome density and standard error for each genotype is shown inside each violin plot. Notice that dots on top of a genotype violin plot represent plants with a thricome density higher that exceeds the maximum density of the axis.

Figure 5. Distribution of Δ (SNP-index) along the chromosomes based on 1Mb windows for the A) SLC family (BGV016047 x LA2278) and B) SLL family (BGV016047 x MoneyMaker). The 95% and 99% confidence thresholds are shown.

Supplementary Table 1. Genes in the QTL candidate region of chromosome 5, 6, 9 and 11. Number of changes with high, moderate and low impact on gene function as well as those with an uncertain effect (modifier) inferred by SNPEff are shown. Annotation of ITAG4.1 when available is also provided.

Supplementary Figure 1. Density of type IV trichomes in the matting populations along the three weeks of phenotyping for the SLC family (BGV016047 x LA2278) (A) and SLL family (BGV016047 x MoneyMaker) (B)

Supplementary Figure 2. Read number by sample

Supplementary Figure 3. Bases sequenced at 5X and 20X coverages by sample

Supplementary Figure 4. Genotyped SNPs by sample

Supplementary Figure 5 Genetic map of the F2 population of the SLC family based on 8,285 bin markers.

Supplementary Figure 6. Results of the association analysis of the trichome IV density along the genome A) 6th week after transplanting, B) 7th week and C) 8th weeks. The minus logarithm of corrected p-values obtained after fitting to a negative binomial logistic

regression model are shown. Horizontal dashed line represents the p-value threshold of 0.01.

Supplementary Materials

Supplementary File 1. K-seq primer sequences

Supplementary File 2. Data recorded during trichome density measures. Mean: mean of trichome density from 4 measures per leaf, leafA: first measure, leafB: second measure, leafC: third measure, position: the position of the leaf, counting from the ground to the apex.

Supplementary File 3. VCF file with F2 K-seq genotypes