



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Instituto de Conservación y Mejora
de la Agrodiversidad Valenciana

Utilización de una colección de germoplasma de tomate para la identificación de genes de interés

Tesis doctoral realizada por
Estefanía Mata Nicolás

Directores

María José Díez Niclós

Javier Montero Pau

Tutor

Joaquín Cañizares Sales

Valencia, junio 2021

*“Nuestro conocimiento es
necesariamente finito,
mientras que nuestra
ignorancia es
necesariamente infinita.”*

Karl Popper

*A todas las personas
que han hecho posible este trabajo.*

*Y a todas las que han
aparecido en el camino.*

RESUMEN

La mejora de las especies cultivadas es un área dinámica, sujeta a las necesidades y requerimientos de los diferentes cultivos en cada momento y de las exigencias de productores y consumidores. Una limitación en este proceso de mejora es la diversidad genética existente en los cultivos. Muchas de las especies cultivadas se han visto sometidas a diversos cuellos de botella poblacionales durante su domesticación, ocasionando una notable pérdida de diversidad genética. En estos casos, las especies silvestres emparentadas revisten un especial interés, pudiéndose utilizar como fuente de genes en la mejora genética. Esta circunstancia se da en el caso del tomate cultivado (*Solanum lycopersicum* var. *lycopersicum*, SLL), una de las hortalizas con mayor producción a nivel mundial. Esta especie surgió a partir de *Solanum lycopersicum* var. *cerasiforme* (SLC) que a su vez fue pre-domesticada a partir de *Solanum pimpinellifolium* (SP). Pese al potencial para la mejora de estas especies debido a su mayor diversidad genética y proximidad al tomate, su uso está limitado en gran medida por la falta de información de las colecciones mantenidas en los bancos de germoplasma, principalmente en SLC. Es por ello necesario abordar estudios exhaustivos que hagan posible disponer de una información amplia y detallada de la diversidad de estas especies presente en el centro de origen y difusión del tomate cultivado, con el fin de ser utilizadas de una forma eficiente en mejora y éste es, por tanto, el objetivo general de esta tesis.

Para lograr el objetivo anteriormente expuesto se ha empleado una colección de germoplasma compuesta por 15 accesiones de SLL, procedentes de México; 27 de SP, procedentes de Perú y Ecuador y 121 de SLC, procedentes de Perú, Ecuador, México y Mesoamérica. Esta colección ha sido sometida a una extensa caracterización fenotípica y genética, lo que ha permitido identificar las regiones de genoma asociadas a caracteres fenotípicos mediante estudios de asociación del genoma completo (GWAS). Además, se han creado poblaciones segregantes F2 para cada una de las entradas de la colección a partir de los híbridos obtenidos con tres parentales diferentes, y estos cruces se han puesto a disposición de la comunidad científica. Este recurso permitirá el estudio del control genético de mutantes de interés en mejora. Esta utilidad se explora en el último capítulo de la tesis, donde se estudia el control genético de la densidad de tricomas en una accesión de la especie SP, característica de interés por conferir resistencia a artrópodos.

El estudio realizado con los datos morfológicos ha demostrado la presencia de una gradación morfológica continua entre los tres grupos taxonómicos, y ha permitido la identificación de una serie de caracteres que permiten la diferenciación entre grupos y también entre las distintas procedencias geográficas dentro de cada grupo taxonómico. Los datos analizados indican que, de entre todos los caracteres analizados, el aumento en el tamaño del fruto y la variación en su forma, así como la posición del estigma de las flores, serían los principales cambios asociados a la domesticación de la especie y, por lo tanto, diferencian a la especie silvestre de SLC y SLL. El análisis ha revelado también la existencia de variabilidad dentro de los grupos geográficos de la especie silvestre, conforme se fue expandiendo desde Perú a Ecuador, como consecuencia de su adaptación a diferentes ambientes. Por otro lado, SLC se caracteriza por una enorme variabilidad morfológica en caracteres de planta, inflorescencia y fruto, lo que la convierte en una variedad con un gran potencial para la mejora.

Tras esta caracterización morfológica, se exploró la diversidad genética que presentaba la colección. Para ello, se emplearon las secuencias genómicas públicas de estos materiales obtenidas dentro del marco del proyecto VARITOME, dentro del que se enmarca

esta tesis. Los diferentes polimorfismos de nucleótido único (SNPs) fueron anotados con base a la predicción de sus efectos en la secuencia codificantes mediante el programa SnpEff y se llevaron a cabo estudios de GWAS. El análisis genético reveló a Perú y Ecuador como las regiones con mayores niveles de diversidad, tanto en la especie SP como en SLC. Se observó una reducción en el número de variantes SNP en SLC México, que concuerda con la hipótesis de que Mesoamérica sea centro de domesticación y difusión del tomate cultivado. Por último, se constató el proceso de domesticación como la causa de la menor diversidad presente en la especie cultivada. Los estudios GWAS permitieron la identificación de correlaciones genotipo-fenotipo, revelando la asociación entre 107 SNPs y ocho caracteres cuantitativos y un total de 30 SNPs asociados a 7 caracteres cualitativos. Una parte de los SNPs detectados fueron localizados cerca de regiones genómicas ya asociadas con genes y QTLs relacionados con el carácter, como las asociaciones entre el peso del fruto con los QTLs ya descritos como *fw2.2* ó *fw9.2*. Otra parte se localizaron en regiones con posibles genes candidatos anotados, como fue el caso del color amarillo del fruto, lo que refuerza la validez de la colección para la detección de caracteres de interés agronómico. Por último, el resto de SNPs detectados se corresponden a regiones no descritas previamente, lo que abre el camino a estudios para la detección de nuevos genes candidatos.

La disponibilidad de datos de caracterización morfológica y genética revalorizan el conjunto de materiales estudiados. Con el objeto de aumentar su utilidad, se llevó a cabo el desarrollo de familias segregantes que facilitan el estudio de caracteres de interés por parte de la comunidad científica. Para ello, se cruzó toda la colección con una entrada representativa de cada uno de los tres grupos taxonómicos, SP, SLC y SLL; autofecundando las F1 conseguidas para obtener las poblaciones F2. Estas familias segregantes constituyen una poderosa herramienta como fuente de variantes alélicas naturales, al tiempo que permiten la caracterización de mutantes y la identificación de los genes implicados de una manera rápida y eficiente al disponer tanto de la caracterización genética como de las poblaciones F2.

La utilidad de dichas familias se demuestra en el último capítulo de esta tesis mediante el estudio de la base genética de la densidad de tricomas tipo IV. Durante la caracterización de la colección, se detectó una entrada de la especie SP (BGV016047) con una alta densidad de tricomas glandulares de tipo IV, en los que se ha descrito la presencia de distintos tipos de sustancias relacionadas con resistencia a algunas plagas. En esta tesis se muestra la caracterización fenotípica de la densidad de tricomas en función de la edad de la planta y de la hoja. Por otro lado, se estudia el control genético de la presencia y alta densidad de tricomas tipo IV en dos fondos genéticos diferentes, SLL x SP y SP x SLC. En ambos fondos se han detectado dos QTLs principales en los cromosomas 9 y 11. Además, se han detectado señales en los cromosomas 2, 5, 6, 7 y 8, en función de la familia segregante estudiada. Las regiones detectadas en los cromosomas 9 y 11 habían sido previamente descritas en otras especies por regular la densidad de este tipo de tricomas o los niveles de acilazúcares. En ambas regiones hay genes anotados que intervienen en el desarrollo de los tricomas o en la formación de acilazúcares, como aciltransferasas, glicosiltransferasas o los factores de transcripción MYB.

RESUM

La millora de les espècies cultivades és una àrea dinàmica, subjecta a les necessitats i requeriments dels diferents cultius a cada moment i de les exigències de productors i consumidors. Una limitació en aquest procés de millora és la diversitat genètica existent en els cultius. Moltes de les espècies cultivades s'han vist sotmeses a diversos colls de botella poblacionals durant la seua domesticació, ocasionant una notable pèrdua de diversitat genètica. En aquests casos, les espècies silvestres emparentades revisten un especial interès, podent-se utilitzar com a font de al·lels en la millora genètica. Aquesta circumstància es dona en el cas de la tomaca cultivada (*Solanum lycopersicum* var. *lycopersicum*, SLL), una de les hortalisses amb major producció a nivell mundial. Aquesta espècie va sorgir a partir de *Solanum lycopersicum* var. *cerasiforme* (SLC), que al seu torn va ser pre-domesticada a partir de *Solanum pimpinellifolium* (SP). Pese al potencial per a la millora d'aquestes espècies a causa de la seua major diversitat genètica i proximitat a la tomaca, el seu ús està limitat en gran manera per la falta d'informació de les col·leccions mantingudes en els bancs de germoplasma, principalment SLC. És per això necessari abordar estudis exhaustius que facen possible disposar d'una informació àmplia i detallada de la diversitat d'aquestes espècies present en el centre d'origen i difusió de la tomaca cultivada, amb la finalitat de ser utilitzades d'una forma eficient en millora i aquest és, per tant, l'objectiu general d'aquesta tesi.

Per a aconseguir l'objectiu anteriorment exposat s'ha emprat una col·lecció de germoplasma composta per 15 accessions de SLL, procedents de Mèxic; 27 de SP, procedents del Perú i l'Equador i 121 de SLC, procedents del Perú, l'Equador, Mèxic i Mesoamèrica. Aquesta col·lecció ha sigut sotmesa a una extensa caracterització fenotípica i genètica, la qual cosa ha permès identificar les regions de genoma associades a caràcters fenotípics mitjançant estudis d'associació del genoma complet (GWAS). A més, s'han creat poblacions segregants F2 per a cadascuna de les entrades de la col·lecció creuant-les amb tres parentals diferents, i aquests creus s'han posat a la disposició de la comunitat científica. Aquest recurs permetrà l'estudi del control genètic de mutants d'interès en millora. Aquesta utilitat s'explora en l'últim capítol de la tesi, on s'estudia el control genètic de la densitat de tricomas en una accessió de l'espècie SP, característica d'interès per conferir resistència a artròpodes.

L'estudi realitzat amb les dades morfològiques ha demostrat la presència d'una gradació morfològica contínua entre els tres grups taxonòmics, i ha permès la identificació d'una sèrie de caràcters que permeten la diferenciació entre grups i també entre les diferents procedències geogràfiques dins de cada grup taxonòmic. Les dades analitzades indiquen que l'augment en la grandària del fruit i la variació en la seua forma, així com la posició de l'estigma de les flors, serien els principals canvis associats a la domesticació de l'espècie i, per tant, diferencien a l'espècie silvestre de SLC i SLL. L'anàlisi va revelar també l'existència de variabilitat dins dels grups geogràfics de l'espècie silvestre, conforme es va anar expandint des del Perú a l'Equador, com a conseqüència de la seua adaptació a diferents ambients. D'altra banda, SLC es caracteritza per una enorme variabilitat morfològica en caràcters de planta, inflorescència i fruit, la qual cosa la converteix en una espècie amb un gran potencial per a la millora.

Després d'aquesta caracterització morfològica, es va explorar la diversitat genètica que presentava la col·lecció. Per a això, es van emprar les seqüències genòmiques públiques d'aquests materials obtingudes dins del marc del projecte VARITOME. Els diferents polimorfismes de nucleòtid únic (SNPs) van ser anotats en base a la predicció del seu efecte en la seqüència codificant mitjançant el programa SnpEff i es van dur a terme estudis de GWAS. L'anàlisi genètica va revelar al Perú i l'Equador com les regions amb majors nivells

de diversitat, tant en l'espècie SP com en SLC. Es va observar una reducció en el nombre de variants SNP en SLC Mèxic, que concorda amb la hipòtesi que Mesoamèrica siga centre de domesticació i difusió de la tomaca cultivada. Finalment, es va constatar el procés de domesticació com la causa de la menor diversitat present en l'espècie cultivada. Els estudis GWAS van permetre la identificació de correlacions genotipe-fenotip, revelant l'associació entre 107 SNPs i huit caràcters quantitius i un total de 30 SNPs associats a 7 caràcters qualitius. Una part dels SNPs detectats van ser localitzats prop de regions genòmiques associades amb gens i QTLs relacionat amb el pes del fruit com *fw2.2* o *fw9.2*, o en regions amb possibles gens candidats anotats, com va ser el cas del color groc del fruit, la qual cosa reforça la validesa de la col·lecció per a la detecció de caràcters d'interés agronòmic. La resta de SNPs detectats es corresponen amb regions no descrites prèviament el que obri el camí a estudis per a la detecció de nous gens candidats.

La disponibilitat de dades de caracterització morfològica i genètica revaloritzen el conjunt de materials estudiats. A fi d'augmentar la seua utilitat, es va dur a terme el desenvolupament de famílies segregantes que faciliten l'estudi de caràcters d'interés per part de la comunitat científica. Per a això, es va creuar tota la col·lecció amb una entrada representativa de SLC i una altra de SLL, autofecundando les F1 aconseguides per a obtenir les poblacions F2. Aquestes famílies segregantes constitueixen una poderosa eina com a font de variants al·lèliques naturals, al mateix temps que permeten la caracterització de mutants i la identificació dels gens implicats d'una manera ràpida i eficient en disposar tant de la caracterització genètica com de les poblacions F2.

La utilitat d'aquestes famílies es demostra en l'últim capítol d'aquesta tesi mitjançant l'estudi de la base genètica de la densitat de tricomas tipus IV. Durant la caracterització de la col·lecció, es va detectar una entrada de l'espècie SP (BGV016047) amb una alta densitat de tricomas glandulars de tipus IV, en els quals s'ha descrit la presència de diferents tipus de substàncies relacionades amb resistència a algunes plagues. En aquesta tesi es mostra la caracterització fenotípica de la densitat de tricomas en funció de l'edat de la planta i de la fulla. D'altra banda, s'estudia el control genètic de la presència i alta densitat de tricomas tipus IV en dos fons genètics diferents, SLL x SP i SP x SLC. En tots dos fons s'han detectat dos QTLs principals en els cromosomes 9 i 11. A més, s'han detectat senyals en els cromosomes 2, 5, 6, 7 i 8, en funció de la família segregant estudiada. Les regions detectades en els cromosomes 9 i 11 havien sigut prèviament descrites en altres espècies per regular la densitat d'aquests tricomas o els nivells de acilazúcares. En totes dues regions hi ha gens anotats que intervenen en el desenvolupament dels tricomas o en la formació de acilsucre, com aciltransferases, glicosiltransferases o els factors de transcripció MYB.

ABSTRACT

The improvement of cultivated species is a dynamic field, subject to the needs and requirements of the different crops at different moments as well as to the demands of producers and consumers. A limitation in this improvement process is the amount of genetic diversity available for breeding purposes. Many of the cultivated species have been subjected to population bottlenecks during their domestication, causing a significant loss of genetic diversity. In these cases, wild related species are of special interest, and can be used as a source of alleles in genetic improvement. This circumstance occurs in the case of cultivated tomatoes (*Solanum lycopersicum* var. *lycopersicum*, SLL), reduced genetic diversity in cultivated species. Subsequently, plant breeding led to an additional reduction of genetic diversity inbreeding populations. Due to this narrow genetic diversity, the improvement of cultivated species strongly depends on their wild ancestors. Despite the advances, the improvement of cultivated species faces new challenges related to the requirements from each time or the demands that producers and consumers claim, so the divert of wild relatives are still required. Cultivated tomato (*Solanum lycopersicum* var. *lycopersicum*, SLL) is one of the vegetables with the highest worldwide production. The expansion and the increase of tomato consumption have introduced new breeding goals related to tomato appearance due to consumer preferences. The wild relatives of cultivated tomato are native to the Andean regions of America. SLL evolved from *Solanum lycopersicum* var. *cerasiforme* (SLC), which was pre-domesticated from *Solanum pimpinellifolium* (SP). Despite these species have been used for tomato breeding, their use is still limited due to the lack of detailed information about the collections that are kept in genebanks. Thus, understanding the genetic and phenotypic variation available within these resources is critical to achieve new breeding goals. It is necessary to undertake an exhaustive characterization of these collections to have more detailed information about the species which are native to the center of origin and diffusion of cultivated tomato.

With the aim of achieving this objective, a germplasm collection has undergone an extensive morphological and genetic characterization. This collection comprises a wide range of geographical origins: includes 15 SLL accessions from Mexico; 27 SP accession from Peru and Ecuador and 121 SLC accessions from Peru, Ecuador, Mesoamerica, and Mexico. Genome-wide association studies (GWAS) were performed to detect genomic candidate regions associated with each agronomic trait. Furthermore, a collection of segregating populations has been developed by crossing the whole collection with a representative accession for each of the three species (SLL, SLC and SP). These populations are available to the scientific community and could be very useful to speed up the validation of candidate genes or to study the genetic control of mutants. This last approach will be explored in the last chapter of this thesis, where the genetic control of density of type IV trichomes is studied.

The study carried out with morphological data has demonstrated the presence of a phenotypic variation between our three species. The analysis has allowed the identification of characters which differentiate between species and between different geographical origin within each species. The analysis showed that the main character associated with the domestication process are related to fruit shape and size, as well as the style exertion. The

analysis also revealed morphological variability within wild species SP as a consequence of its adaptation to the different environments. Moreover, SLC was characterized by a wide range of morphological variability for plant architecture, inflorescences and fruit characteristics which makes this species a potential source in tomato breeding programmes.

After this phenotypic characterization, the genetic diversity was explored by using the public genomic sequence of our collections which are available in the framework of VARITOME project. The identified SNPs were annotated, and their putative impacts were predicted by using SnpEff. The lowest number of variants was detected in SLL groups and the highest number of variants was detected in SP. SLC had a variation between SLL and SP, with lower levels in SLC Mexico. This fact agrees with the hypothesis for the domestication and diffusion of cultivated tomato in Mesoamerica. Additionally, GWAS analysis was carried out with this genetic data and revealed significant associations between 107 SNPs and 8 quantitative traits and a total of 30 SNPs associated with 7 qualitative traits. This analysis has allowed the identification of known genomic regions such as the associations between fruit weight and *fw2.2* or *fw9.2*. However, regions with possible novel genes have also been detected such as the case of yellow fruit. These results reinforce the potential of our collection for breeding programmes.

With the aim of using the genetic and morphological potential, the whole collection was crossed with a representative accession of each species. F1 plants were self-pollinated to obtain F2 generations. These different segregating families are a powerful tool as a source of natural allelic variants, they allow the identification of mutants and possible candidate genes or allow to analyze a mutation effect in different genetic backgrounds. The usefulness of these families has been demonstrated in the last chapter where F2 populations have been used to determine the genetic control of density of type IV trichomes. During the characterization of this collection, a SP accession (BGV016047) with a high density of type IV trichomes was detected. These trichomes have been described to accumulate different types of chemical substances related to pest resistance. This thesis shows that this character is influenced by environmental factors such as plant and leave age. On the other hand, the genetic control of this character has been studied in two different genetic backgrounds, SLL x SP and SP x SLC. Two main QTLs have been detected in both families, on chromosome 9 and 11. In addition, the different F2 populations have revealed that this character could be under the control of other QTLs on chromosomes 2, 5, 6, 7 and 8. The associations between detected regions on chromosome 9 and 11 and density of these trichomes or acylsugar levels have been previously described. In both regions, some genes have been annotated to be involved in trichome development, acylsugars synthesis or MYB transcription factor.

ÍNDICES

ÍNDICE GENERAL

CAPÍTULO 1:INTRODUCCIÓN GENERAL

1. Importancia de la mejora genética vegetal -----	2
2. Fuentes de variabilidad genética como recurso en mejora -----	4
3. Diversidad genética y mejora del tomate cultivado -----	5
4. Mejora genética del tomate cultivado -----	14
5. Uso de marcadores genéticos para el aprovechamiento de los recursos fitogenéticos ----	22

CAPÍTULO 2: VARIABILIDAD MORFOLÓGICA DE LA COLECCIÓN ----- 44

CAPÍTULO 3:EXPLOITING THE DIVERSITY OF TOMATO: THE DEVELOPMENT OF PHENOTYPICALLY AND GENETICALLY DETAILED COLLECTION ----- 61

CAPÍTULO 4: DISCOVERY OF A MAJOR QTL CONTROLLING TRICHOME IV DENSITY IN TOMATO USING K-SEQ GENOTYPING ----- 87

DISCUSIÓN GENERAL-----115

ANEXOS -----143

ÍNDICE FIGURAS

CAPÍTULO 1:INTRODUCCIÓN GENERAL

Figura 1. Hortalizas con mayor producción mundial (FAO, 2017) -----	3
Figura 2. Distribución de la producción de tomate por continentes, teniendo en cuenta los datos de producción desde 2007 a 2017 (FAO, 2017) -----	3
Figura 3. Posibles rutas de expansión del tomate cultivado (Modificado de Díez y Nuez, 2008) ----	8
Figura 4. Reducción de la diversidad genética del tomate cultivado y uso de las especies silvestres en mejora. Modificado de Tanksley y McCouch (1997) y Zhu et al. (2018)-----	10
Figura 5. Tipos de tricomas en tomate y especies silvestres relacionadas (Modificado de Peralta et al., 2008)-----	17
Figura 6. Variabilidad en la forma del fruto de tomate cultivado y especies silvestres relacionadas	21
Figura 7. Esquema GWAS para la identificación de genes/QTLs de interés y utilización de estas variantes alélicas para los genes identificados en programas de mejora -----	24

CAPÍTULO 2: VARIABILIDAD MORFOLÓGICA DE LA COLECCIÓN

Figura 8. Distribución geográfica de las entradas empleadas en el ensayo -----	31
Figura 9. Descripción de los parámetros del <i>Tomato Analyzer</i> seleccionados -----	34
Figura 10. Caracteres con una correlación superior a 0,17 con las CP1 y CP2 del PCA realizado con todas las especies -----	37
Figura 11. Análisis de componentes principales efectuado con los caracteres que presentaron diferencias significativas entre las especies-----	38
Figura 12. Caracteres vegetativos que presentan diferencias entre las tres especies -----	40
Figura 13. Caracteres relacionados con flores e inflorescencias que presentan diferencias entre las tres especies-----	41
Figura 14. Caracteres de fruto que presentan diferencias entre las tres especies-----	42
Figura 15. Distribución geográfica de las entradas de la especie <i>S. pimpinellifolium</i> -----	43
Figura 16. Caracteres con una correlación superior a 0,17 con las CP1 y CP2 del PCA realizado con <i>S. pimpinellifolium</i> de Perú y Ecuador -----	44
Figura 17. Análisis de coordenadas principales efectuado con caracteres significativamente diferentes entre los grupos de entradas de <i>S. pimpinellifolium</i> establecidos según su procedencia geográfica - -----	45

Figura 18. Caracteres de tipo vegetativo que presentan diferencias entre las entradas de <i>S. pimpinellifolium</i> procedentes de Ecuador y Perú-----	46
Figura 19. Caracteres relacionados inflorescencias que caracterizan las entradas <i>S. pimpinellifolium</i> procedentes de Ecuador y Perú-----	46
Figura 20. Caracteres relacionados con flores que caracterizan las entradas de <i>S. pimpinellifolium</i> procedentes de Ecuador y Perú-----	47
Figura 21. Caracteres de fruto que presentan diferencias entre <i>S. pimpinellifolium</i> procedentes de Ecuador y Perú-----	47
Figura 22. Distribución geográfica de las entradas de la especie <i>S. lycopersicum</i> var. <i>cerasiforme</i> ---	49
Figura 23. Caracteres con una contribución superior a 0.17 con las CP1 y CP2 del PCA realizado con <i>S. lycopersicum</i> var. <i>cerasiforme</i> de Perú, Ecuador, Mesoamérica y México -----	50
Figura 24. Análisis de coordenadas principales efectuado con los caracteres que presentaron diferencias significativas entre los grupos geográficos de <i>S. lycopersicum</i> var. <i>cerasiforme</i>	54
Figura 25. Caracteres relacionados con hojas que diferencian a <i>S. lycopersicum</i> var. <i>cerasiforme</i> procedentes de Ecuador, Perú, Mesoamérica y México -----	55
Figura 26. Caracteres relacionados con flores que diferencian a <i>S. lycopersicum</i> var. <i>cerasiforme</i> procedentes de Ecuador, Perú, Mesoamérica y México -----	56
Figura 27. Caracteres de fruto que describen los frutos de <i>S. lycopersicum</i> var. <i>cerasiforme</i> procedentes de Ecuador, Perú, Mesoamérica y México -----	57
Figura 28. Principales diferencias en tamaño y forma del fruto en SLC -----	58

CAPÍTULO 3: EXPLOITING THE DIVERSITY OF TOMATO: THE DEVELOPMENT OF PHENOTYPICALLY AND GENETICALLY DETAILED COLLECTION

Figure 1. Morphological variation -----	69
Figure 2. Diversity in leaf, fruit, flower, and inflorescence traits-----	71
Figure 3. Genome-wide association results for some traits that showed significant association	72

CAPÍTULO 4: DISCOVERY OF A MAJOR QTL CONTROLLING TRICHOME IV DENSITY IN TOMATO USING K-SEQ GENOTYPING

Figure 1. Effect of plant age (A), leaf age (B) and growing season (C) on type IV trichome density. Photography of trichomes type IV (D)-----	95
Figure 2. Density of type IV trichomes in the segregating populations-----	96
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 week-----	98
Figure 4. Effect of QTLs TIVd9, TIVd11 and the additive model on trichome density-----	100
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 Money Maker) -----	103

ÍNDICE TABLAS

CAPÍTULO 1: INTRODUCCIÓN GENERAL

Tabla 1. Lista de especies del género <i>Solanum</i> y su correspondencia con el género <i>Lycopersicon</i> . (peralta et al., 2008)-----	6
Tabla 2. Genes de resistencia a enfermedades descritos en <i>S. pimpinellifolium</i> e introgresados en tomate cultivado -----	12
Tabla 3. Especies silvestres del género <i>Solanum</i> , sección <i>lycopersicon</i> usadas en mejora -----	13
Tabla 4. Principales enfermedades y plagas que afectan al cultivo de tomate -----	15
Tabla 5. Principales especies en las que se ha descrito la presencia de tricomas glandulares tipo IV-----	18
Tabla 6. QTLs descritos para la presencia de tricomas tipo iv, resistencia conferida por estos tricomas o síntesis de acilazúcares-----	19
Tabla 7. Principales mutantes de color del fruto descritos en tomate (Stommel, 2014)-----	21

CAPÍTULO 2: VARIABILIDAD MORFOLÓGICA DE LA COLECCIÓN

Tabla 8. Número de entradas de cada especie y origen geográfico-----	31
Tabla 9. Caracteres cuantitativos y cualitativos evaluados-----	32
Tabla 10. Caracteres seleccionados de <i>Tomato analyzer</i> y su descripción -----	33
Tabla 11. Criterios empleados en la conversión de caracteres cuantitativos a caracteres cualitativos -----	35

CAPÍTULO 3: VARIABILIDAD GENÉTICA EN LA COLECCIÓN

Table 1. Summary of significant associations detected for quantitative and qualitative traits -----	73
Table 2. Result of the number of variants per type and the number of effects by impact for each geographical group from SnpEff -----	76

CAPÍTULO 4: POBLACIONES SEGREGANTES Y EL ANÁLISIS DEL CONTROL GENÉTICO DE LA DENSIDAD DE TRICOMAS GLANDULARES TIPO IV

Table 1. Summary statistics for the genetic map built using the F2 generation of the cross BGV016047 x LA2278-----	97
--	----

Table 2. QTL location based on different methods using SLC family and SLL family mapping populations. ----- 99

Table 3. Results of the negative binomial logistic regression for the QTLs found by QTL mapping and association analysis----- 101

Capítulo 1:
INTRODUCCIÓN GENERAL

1. Importancia de la mejora genética vegetal

La población mundial se encuentra en continuo crecimiento, estimándose que en el año 2050 llegue a superar los 9 billones de personas (Naciones Unidas, 2017). Este incremento poblacional requiere un aumento equivalente de la producción de alimentos para poder satisfacer sus necesidades alimenticias. Sin embargo, para asegurar esta producción es necesario el desarrollo de variedades más productivas, adaptadas a condiciones ambientales específicas, con una reducción de insumos y adecuadas a una producción a gran escala. El desarrollo de variedades adaptadas a nuevas condiciones ambientales, además, es clave para poder garantizar la seguridad alimentaria en el mundo. Así se podría evitar una nueva crisis alimentaria como la del año 2008 (FAO, 2008), en la que se produjeron caídas en la producción de cereales debido a fenómenos climáticos extremos, que supusieron un incremento de los precios de hasta el 40%.

El aumento de la producción no es el único reto al que se enfrenta la mejora genética vegetal, ya que también se busca una reducción del coste. Esto se consigue mediante el desarrollo de variedades mejoradas genéticamente para diversos aspectos como la adaptación a la recogida mecánica de los frutos, mayor vida postcosecha y resistencia al transporte, así como el desarrollo de variedades resistentes. Este último aspecto es de especial importancia ya que la globalización del mercado ha traído consigo la diseminación de virus y enfermedades fuera de su hábitat, ocasionando epidemias y, consecuentemente, grandes pérdidas económicas. Otro de los eslabones en esta cadena de retos son las preferencias del consumidor. Los consumidores cada vez hacen una mayor demanda de productos de mejor calidad, lo que ha dado lugar a que la importancia de este objetivo de mejora se haya visto incrementada en los últimos años.

La mejora de los cultivos debe ir adaptándose para satisfacer las necesidades y demandas de cada momento. Pese a la mejora que han ido sufriendo los cultivos en las últimas décadas, es un hecho patente que todavía queda mucho camino por recorrer y que para poder alcanzar los objetivos de mejora es necesario disponer de una gran cantidad de información actualizada y de calidad sobre los cultivos. Un elemento fundamental para la mejora es tener una descripción detallada de la diversidad genética de la especie y de las especies silvestres relacionadas que sirven de fuente de innovación genética.

1.1. Importancia del cultivo del tomate

La producción y el consumo de hortalizas se han incrementado drásticamente en los últimos años. El tomate es una de las hortalizas con mayor producción mundial, solamente superada por la patata, con una producción que supera los 180 millones de toneladas anuales (FAO, 2017) (Figura 1) y que se ha visto incrementada en 40 millones de toneladas en la última década. Estos hechos la convierten en una especie de gran interés en mejora.

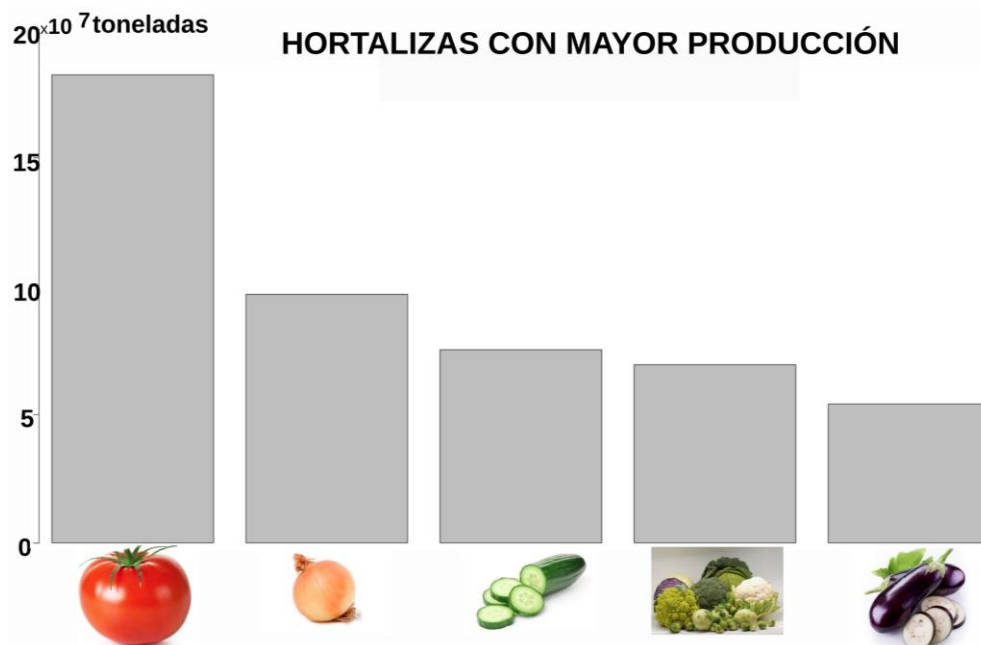


Figura 1. Hortalizas con mayor producción mundial (FAO, 2017).

Más de la mitad de esta producción se genera en el continente asiático (58 %), concretamente en China, que concentra un tercio de la producción mundial de tomate. Le siguen América (16 %), con EEUU como mayor productor del continente, y Europa (14 %), donde los mayores productores son Turquía, Italia y España (Figura 2).

Distribución de la producción (2007-2017)

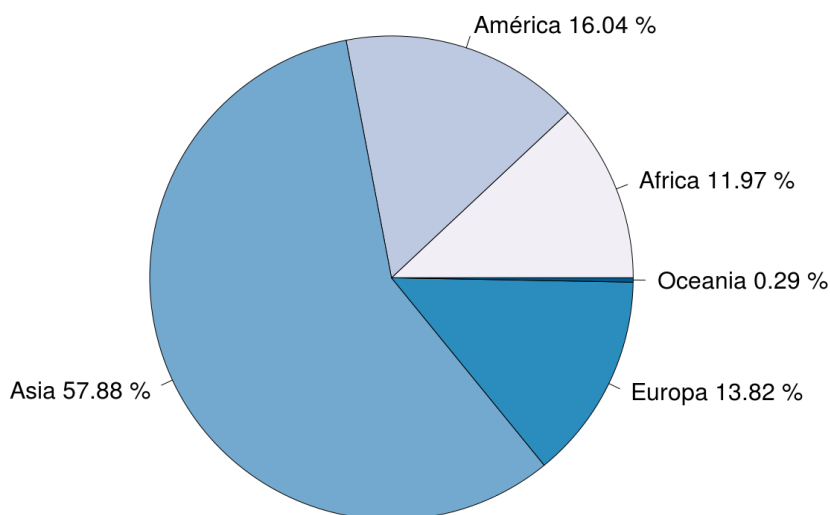


Figura 2. Distribución de la producción de tomate por continentes, teniendo en cuenta los datos de producción desde 2007 a 2017 (FAO, 2017).

El consumo mundial de tomate supone un promedio de 20,81 Kg/persona/año (FAO, 2017), siendo Europa el continente de mayor consumo con una media de 27,29 Kg/persona/año. Estos volúmenes de consumo se deben a que es un alimento muy versátil, ya que puede ser consumido en fresco o bien empleado en la elaboración de alimentos procesados como salsas. Y pese a que no es un alimento que destaque por su elevado valor

nutricional, ha llegado a convertirse en una de las principales fuentes de vitaminas y minerales de la dieta mediterránea debido a las grandes cantidades en las que es consumido (Foolad, 2007). Además, es un alimento beneficioso para la salud por ser fuente de compuestos antioxidantes como el licopeno, que contribuye a reducir el riesgo de sufrir enfermedades cardiovasculares o determinados tipos de cáncer (Mazidi et al., 2020).

2. Fuentes de variabilidad genética como recurso en mejora

Las especies cultivadas suelen presentar una diversidad genética baja. Esto es debido a que durante el proceso de domesticación de muchos cultivos se produjeron cuellos de botella que redujeron su diversidad genética (Meyer & Purugganan, 2013). Ya los primeros agricultores comenzaron a realizar una selección de caracteres de interés para su aprovechamiento a partir de subpoblaciones de especies silvestres (Olsen & Wendel, 2013; Purugganan, 2019). Esto dio lugar a un incremento de la prevalencia de determinados alelos específicos en las nuevas poblaciones, incluyendo la acumulación de alelos con efectos deletéreos como se ha visto en arroz, maíz o girasol (Hufford et al., 2019).

Tras esta primera domesticación, las poblaciones sufrieron un proceso de difusión desde su centro de origen (Meyer & Purugganan, 2013). Durante este proceso, las nuevas poblaciones se fueron seleccionando para su adaptación local a distintos ambientes y prácticas de cultivo. Esta segunda adaptación se produjo o bien mediante la aparición de mutaciones de *novo* o bien por intercambio de material genético con las especies silvestre que vivían en estos nuevos ambientes (Hufford, Teran, & Gepts 2019; Purugganan 2019). Por ejemplo, la patata presenta una alta diversidad genética debido al intercambio de material genético con diferentes especies silvestres durante la expansión del cultivo (Hardigan et al., 2017) y el maíz debe su fenotipo no ramificado a una introgresión procedente de teosinte (Studer et al., 2011). Por el contrario, la judía presenta poca diversidad debido al escaso flujo genético con las especies silvestres durante los procesos de expansión a nuevos ambientes (Papa et al., 2007). Todos estos procesos dieron lugar al desarrollo de variedades tradicionales que mantenían solo una pequeña fracción de la variación genética presente en las especies silvestres y presentaban una gran divergencia fenotípica con éstas (Meyer & Purugganan, 2013). Estas variedades tradicionales posteriormente se emplearon en el desarrollo de las variedades modernas, lo que redujo aún más la diversidad genética de la especie cultivada (Zamir, 2001).

Una consecuencia de la reducción de la diversidad genética es que la mejora genética de las especies cultivadas se ha vuelto dependiente de la variabilidad presente en las especies silvestres. Por ello, y a pesar de que transferir genes desde especies silvestres es un proceso largo y laborioso, éstas ya han sido empleadas con éxito en muchas especies como trigo, tomate, arroz y soja. Estos recursos, generalmente mantenidos en los bancos de germoplasma, se han empleado sobre todo para desarrollar variedades resistentes y de mayor productividad (Zamir, 2001). Por lo tanto, debido al valor de las especies silvestres como posible fuente de genes de interés en mejora es muy importante su conservación y caracterización en los bancos de germoplasma (Tanksley & McCouch, 1997).

3. Diversidad genética y mejora del tomate cultivado

3.1. Clasificación taxonómica del tomate

El tomate cultivado, *Solanum lycopersicum* var. *lycopersicum* L., pertenece a la familia de las solanáceas. Esta familia engloba a más de 2.800 especies, organizadas en 96 géneros, que se caracterizan por ser extremadamente diversas en hábito de crecimiento, hábitat o morfología (Peralta et al., 2008). Además, es una de las familias más importantes desde el punto de vista económico ya que incluye cultivos de interés agronómico como la patata (*Solanum tuberosum* L.), el pimiento (*Capsicum annuum* L.), el tabaco (*Nicotiana tabacum* L.) o la berenjena (*Solanum melongena* L.).

La clasificación taxonómica del tomate cultivado ha sido objeto de múltiples controversias durante años, debido a la gran cantidad de especies que constituyen el género *Solanum* (Miller 1754; Luckwill 1943; Rick 1960; Lester 1991; Peralta & Spooner 2006). En 1694, Joseph Pitton de Tournefort, basándose en los frutos multiloculares del tomate cultivado, lo asignó a un nuevo género, que denominó género *Lycopersicon*, separado del género *Solanum*. Años más tarde, en 1753, Carlos Linneo, clasificó el tomate dentro del género *Solanum*, incluyendo las especies *S. lycopersicum* y *S. peruvianum*. Sin embargo, en 1754, Philip Miller propuso una nueva clasificación del tomate, separándolo del género *Solanum* de nuevo e incluyéndolo en el género *Lycopersicon* como *L. esculentum*. Michel Félix Dunal (1813) continuó con esta clasificación y describió las especies *L. cerasiforme* y *L. pyriforme*, teniendo en cuenta la forma de los frutos. Atendiendo a las características morfológicas de las especies silvestres, el género *Lycopersicon* fue dividido en dos subgéneros, *Eulycopersicon* y *Eriopersicon* (Müller 1940; Luckwill 1943). Posteriormente, Rick (1960, 1979) reconoció nueve especies de tomate silvestre y las agrupó en dos complejos, según las relaciones de cruzabilidad con el tomate cultivado: el complejo *Esculentum* y el complejo *Peruvianum*. El primero estaba constituido por las especies *L. cheesmaniae*, *L. chmielewskii*, *L. parviflorum*, *L. pimpinellifolium*, *L. hirsutum* y *L. pennelli*. Estas especies son capaces de cruzarse con el tomate cultivado, autocompatibles y algunas presentan frutos pigmentados. Por su parte, el complejo *Peruvianum* incluía a las especies *L. peruvianum* y *L. chilense*, especies autoincompatibles, de frutos verdes y pilosos e incapaces de hibridar con el tomate cultivado sin recurrir a técnicas biotecnológicas como el rescate de embriones. Actualmente, la taxonomía aceptada para el tomate cultivado es la siguiente (Peralta et al., 2008):

División: Magnoliophyta

Clase: Magnoliopsida

Subclase: Asteridae

Orden: Solanales

Familia: Solanaceae

Subfamilia: Solanoideae

Género: *Solanum* L.

Especie: *Solanum lycopersicum* var. *lycopersicum*

En la Tabla 1 se recogen los nombres actuales de las especies silvestres del género *Solanum*, sección *Lycopersicoides*, sección *Juglandifolia* y sección *Lycopersicon*, así como su correspondencia con la nomenclatura anterior.

Tabla 1. Lista de especies del género *Solanum* y su correspondencia con el género *Lycopersicon* (Peralta et al., 2008).

Nombre dentro del género <i>Solanum</i>	Equivalente del género <i>Lycopersicon</i>
Sección <i>Lycopersicoides</i>	
<i>S. lycopersicoides</i> Dunal	<i>L. lycopersicoides</i> Dunal A. Child ex J.M.h.Shaw
<i>S. sitiens</i> I.M. Johnst	<i>L. sitiens</i> I.M. Johnst
Sección <i>Juglandifolia</i>	
<i>S. juglandifolium</i> Dunal	<i>L. juglandifolium</i> Dunal
<i>S. ochranthum</i> Dunal	<i>L. ochranthum</i> Dunal
Sección <i>Lycopersicon</i>	
<i>S. pennellii</i> Correll	<i>L. pennellii</i> Correll D'Arcy
<i>S. habrochaites</i> S.Knapp D.M Spooner	<i>L. hirsutum</i> Dunal
<i>S. chilense</i> Dunal Reiche	<i>L. chilense</i> Dunal
<i>S. huaylasense</i> Peralta	Parte de <i>L. peruvianum</i> L. Miller
<i>S. peruvianum</i> L.	Parte de <i>L. peruvianum</i> L. Miller
<i>S. corneliomuelleri</i> J.F.Macbr.L.	Parte de <i>L. peruvianum</i> L. Miller
<i>S. arcanum</i> Peralta	Parte de <i>L. peruvianum</i> L. Miller
<i>S. chmielewskii</i> C.M. Rick, Kesicki, Fobes & M. Holle D.M. Spooner, Anderson & R.K. Jansen	<i>L. chmielewskii</i> C.M. Rick, Kesicki, Fobes & M. Holle
<i>S. neorickii</i> C.M. Rick, Kesicki, Fobes & M. Holle D.M. Spooner, Anderson & R.K. Jansen	<i>L. parviflorum</i> C.M. Rick, Kesicki, Fobes & M. Holle
<i>S. pimpinellifolium</i> L.	<i>L. pimpinellifolium</i> L. Miller
<i>S. lycopersicum</i> L.	<i>L. esculentum</i> Miller
<i>S. cheesmaniae</i> L. Riley Fosberg	<i>L. cheesmaniae</i> L. Riley
<i>S. galapagense</i> S. Darwin & Peralta	Parte de <i>L. cheesmaniae</i> L. Riley

3.2. Origen, domesticación y expansión del tomate cultivado

El origen y el proceso de domesticación del tomate también ha sido objeto de múltiples controversias. De Candolle (1886) fue el primer autor que postuló el origen peruano de la domesticación del tomate, basándose en datos botánicos, arqueológicos, históricos y filológicos. Por un lado, este autor consideraba que el proceso de domesticación del tomate cultivado se produjo antes de la llegada de los conquistadores europeos. Su hipótesis se apoyaba en el hecho de que no había datos que demostraran la existencia del tomate fuera de América antes de la llegada a América de estos conquistadores. Además, este cultivo se introdujo en Italia con los nombres “mala peruviana” y “pomi del Perú”. Por lo tanto, este autor consideraba que esto también podría hacer referencia a su origen peruano. Por último, también postulaba la posibilidad de que el tomate cultivado procediera de los tomates *cherry* que ya se conocían por estar ampliamente distribuidos desde la costa de Perú hasta México.

Sin embargo, la falta de restos de tomate en yacimientos arqueológicos de la región andina junto con las evidencias de su cultivo en México, hicieron que Jenkins (1948) propusiera un origen mexicano del cultivo. Además, también argumentó que el nombre del tomate podría proceder del término azteca “tomatl”.

La incorporación de técnicas moleculares al estudio del proceso de domesticación aportó nuevas evidencias. Rick & Fobes (1975), usando marcadores isoenzimáticos, propusieron que el tomate cultivado se habría originado a partir de la especie *Solanum lycopersicum* var. *cerasiforme*, que a su vez estaría relacionada genéticamente con la especie silvestre *Solanum pimpinellifolium* y sugirieron que el tomate cultivado se introdujo en Europa y Norteamérica desde México. Además, encontraron que los cultivares europeos presentaban mayor semejanza con los cultivares mexicanos que con los de la región andina.

Estudios posteriores han intentado aclarar la posición filogenética y el papel que ha jugado la variedad *S. lycopersicum* var. *cerasiforme* en el origen tomate cultivado. Algunos autores consideraron que esta especie no sería el ancestro del tomate cultivado, sino un híbrido entre *S. lycopersicum* var. *lycopersicum* y *S. pimpinellifolium* (Nesbitt & Tanksley, 2002; Ranc et al., 2008). Sin embargo, estudios posteriores demostraron que *S. lycopersicum* var. *cerasiforme* sí era una entidad taxonómica propia (Blanca et al., 2012; 2015) y que la domesticación de *S. lycopersicum* var. *lycopersicum* se produjo a partir de *S. lycopersicum* var. *cerasiforme*, aunque existen hipótesis diferentes sobre cómo se produjo este proceso de domesticación/especiación. Blanca et al (2012, 2015) sugieren una domesticación en dos pasos con una primera pre-domesticación de *S. lycopersicum* var. *cerasiforme* a partir de *S. pimpinellifolium*, que habría tenido lugar en Perú. Posteriormente, esta especie migró a Mesoamérica donde fue domesticada a *S. lycopersicum* var. *lycopersicum*. Alternativamente a este proceso de pre-domesticación, Razifard et al. (2020) proponen una posible divergencia natural entre *S. pimpinellifolium* y *S. lycopersicum* var. *cerasiforme*. Además, estas especies habrían intercambiado material genético entre ambas tras la divergencia. Posteriormente, *S. lycopersicum* var. *cerasiforme* sufrió un proceso de difusión hacia el norte de América, que fue acompañado de múltiples cambios morfológicos como consecuencia de la domesticación y selección no intencionada por parte del hombre. Por último, *S. lycopersicum* var. *lycopersicum* se habría originado a partir de *S. lycopersicum* var. *cerasiforme* presente en México.

La llegada del tomate a Europa se produjo en el s. XVI, tal y como demuestran las primeras descripciones hechas por botánicos de la época. Estas plantas se trajeron y se mantuvieron como plantas ornamentales, ya que se consideraron venenosas debido a su gran parecido con otras solanáceas europeas como la mandrágora y la belladona (Bauchet & Causse, 2012).

Poco tiempo después de su llegada a Europa, hubo una rápida incorporación del tomate en la alimentación, como demuestran las referencias en libros de cocina del s. XVII. Sin embargo, en el norte de Europa su incorporación fue más tardía ya que en Inglaterra su consumo no fue habitual hasta mediados del siglo XVIII (Bergougnoux, 2014). De la mano de españoles y portugueses, se produjo una expansión del cultivo del tomate conforme iba colonizando el Nuevo Mundo. Por último, también se tienen datos de su presencia en África y Asia a partir del s. XVII (Peralta et al., 2008). La Figura 3 muestra las posibles rutas de expansión que sufrió el tomate cultivado.

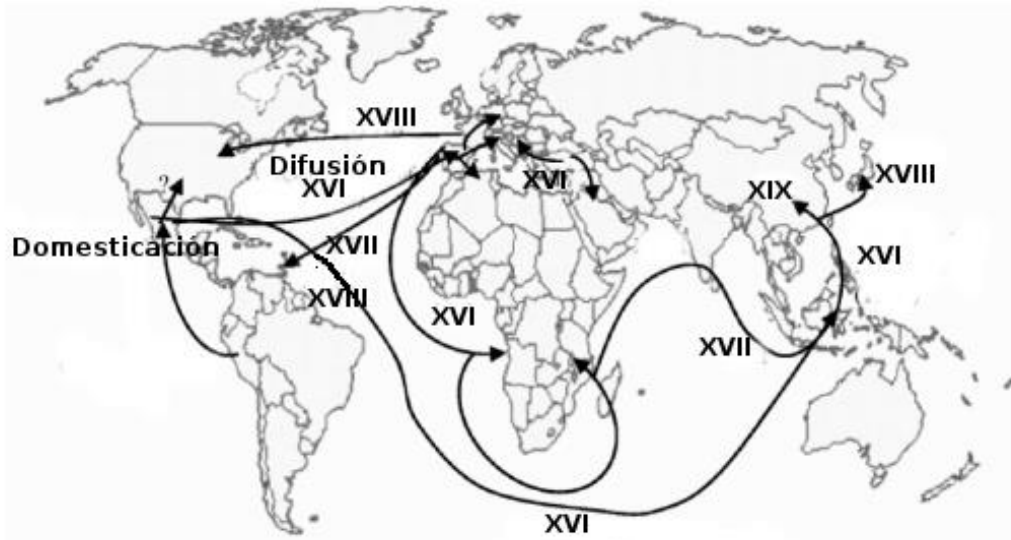


Figura 3. Posibles rutas de domesticación y expansión del tomate cultivado (Modificado de Díez y Nuez, 2008).

El tomate cultivado comenzó a producirse a pequeña escala en los Estados Unidos y Europa durante el s. XIX, generando cultivares mediante polinización abierta. Por lo tanto, el desarrollo de nuevos cultivares se debía a mutaciones espontáneas o a fijación de la variabilidad ya existente (Bergougnoux, 2014).

Como consecuencia del aumento de su consumo, a lo largo del s. XX se incrementó exponencialmente su producción ayudada por el surgimiento del envasado en cadena. También se crearon variedades más productivas, resultado de los programas de mejora que se iniciaron en los años 30 (Labate et al., 2007) y surgieron las primeras casas de semillas, que comenzaron a comercializar los primeros híbridos de tomate (Bergougnoux, 2014). Los híbridos supusieron un aumento del rendimiento comercial del cultivo, ya que el agricultor se veía obligado a comprar semilla cada año puesto que los caracteres segregaban en la descendencia. El primer híbrido de tomate fue “*Single Cross*”, y apareció en Holanda en el año 1946 (Dorst, 1946). En los años 40, también comenzó el intercambio de información genética con las especies silvestres para crear variedades más productivas y adaptadas a las necesidades y demandas del momento.

3.3. Diversidad genética del tomate cultivado

Durante los procesos anteriormente descritos de domesticación y difusión del tomate cultivado, éste sufrió una drástica reducción de su diversidad genética. Esta reducción se debió a diversos cuellos de botella poblacionales que tuvieron lugar debido principalmente a su reproducción dentro de pequeñas poblaciones; acompañado de una selección natural y artificial y al cambio en su sistema reproductivo hacia la autogamia (Sims 1980; Rick & Yoder 1988; Rick 1991).

Los estudios de diversidad genética no solo han puesto de manifiesto su reducción, sino que han revelado la existencia de una mayor variabilidad dentro de las especies próximas *Solanum pimpinellifolium* y *Solanum lycopersicum* var. *cerasiforme*. Así, la especie silvestre *S. pimpinellifolium* es la que presenta una mayor diversidad de las tres (Rick & Fobes 1975; Blanca et al. 2012; Sim et al. 2012b; Razifard et al. 2020) aunque los niveles de diversidad

genética dependen de su origen geográfico. Las poblaciones del noroeste de Perú son las que mayor variabilidad presentan, reduciéndose progresivamente hacia el sur de Perú y hacia Ecuador (Rick, Fobes, & Holle 1977; Caicedo & Schaal 2004; Zuriaga et al. 2009; Blanca et al. 2012; Blanca et al. 2015; Razifard et al. 2020). Rick (1984) estableció que el norte de Perú es el centro de origen de esta especie. Por lo tanto, se produjo una pérdida progresiva de diversidad genética como consecuencia de la migración y colonización de nuevos hábitats desde su centro de origen (Rick 1984; Caicedo & Schaal 2004). Estas nuevas poblaciones se establecieron a partir de un número pequeño de individuos, fijándose unos alelos dentro de la población y perdiéndose otros. Además, se han relacionado las diferencias entre grupos geográficos a las diferencias climáticas entre las distintas regiones geográficas (Zuriaga et al. 2009; Blanca et al. 2012). Esto indica un proceso de adaptación local de la especie a las condiciones ambientales de cada territorio (Blanca et al. 2012).

Con respecto a *S. lycopersicum* var. *cerasiforme*, los niveles de diversidad genética entre ésta y *S. pimpinellifolium* son similares en la región andina (Blanca et al. 2012; Razifard et al. 2020). Sin embargo, esta variabilidad se va reduciendo conforme *S. lycopersicum* var. *cerasiforme* migra desde esta región hacia Mesoamérica y México (Rick & Fobes 1975; Blanca et al. 2015; Razifard et al. 2020). Esta reducción en su diversidad sería una consecuencia de la migración de la especie desde su lugar de origen y la fijación de genotipos exitosos en las nuevas poblaciones (Rick & Holle 1990).

S. lycopersicum var. *lycopersicum* presenta una diversidad genética muy reducida con respecto a *S. lycopersicum* var. *cerasiforme* y *S. pimpinellifolium* (Blanca et al. 2012; Razifard et al. 2020). De hecho, la mayor reducción de la diversidad genética de la especie habría sucedido durante la domesticación y mejora (Gao et al. 2019), ya que los cultivares modernos son aún menos diversos que las variedades tradicionales procedentes de Sudamérica (Blanca et al. 2015). Por lo tanto, para poder aumentar la diversidad genética de la especie queda patente que es necesario recurrir a las especies silvestres relacionadas. Esto hace que sea especialmente importante explorar la diversidad genética existente y conservada en los bancos de germoplasma.

3.4. Uso de la diversidad en mejora

La primera fuente de variabilidad empleada en la mejora del tomate fue la variabilidad intraespecífica. Hasta principios del s. XX, la mejora del tomate se basaba en seleccionar nuevos genotipos que aparecían en los cultivares heterogéneos o en la observación de nueva variabilidad como consecuencia de mutaciones espontáneas o hibridaciones naturales. Por ejemplo, el cultivar ‘Marglobe’ se obtuvo mediante la selección en familias segregantes del cruce entre los cultivares ‘Globe’ y ‘Marvel’. El éxito de esta nueva variedad radicó en su resistencia a fusariosis y alternariosis del tomate (Stevens & Rick 1986).

Sin embargo, la variabilidad existente no fue suficiente para hacer frente a los retos que iban apareciendo en el cultivo del tomate y como consecuencia, fue necesario recurrir a hibridaciones con las especies silvestres (Figura 4).

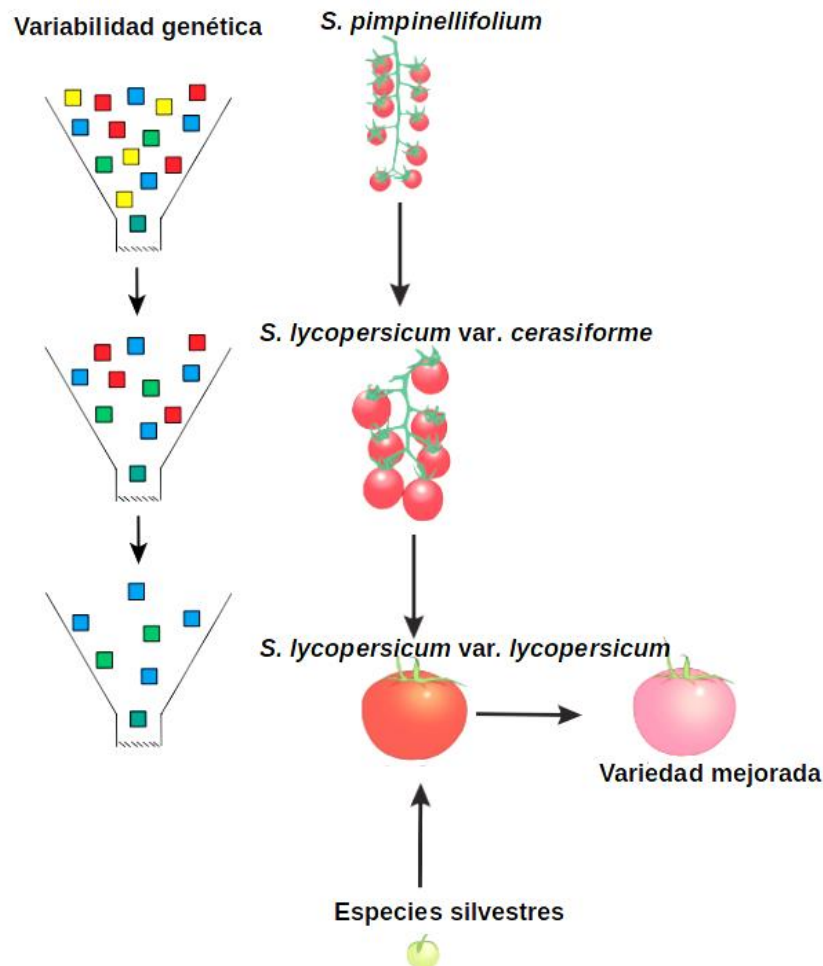


Figura 4. Reducción de la diversidad genética del tomate cultivado y uso de las especies silvestres en mejora. Modificado de Tanksley & McCouch (1997) y Zhu *et al.* (2018).

3.4.1. Colecciones de germoplasma

Como ya se ha comentado anteriormente, las colecciones de germoplasma son una fuente de diversidad con un gran potencial para la mejora genética. Entre las colecciones de germoplasma más importantes de solanáceas a nivel internacional actualmente, podemos destacar la del World Vegetable Center (AVRDC) de Taiwán, el United States Department of Agriculture - Agricultural Research Service (USDA-ARS), Plant Genetic Resources Unit (PGRU) o C.M. Rick Tomato Genetics Resources Center (TGRC) de los Estados Unidos. Dentro de la Unión Europea, cabe destacar la del Institute for Horticultural Plant Breeding (IVT) de Holanda, N.I. Vavilov All-Russian Scientific Research Institute of Plant Industry (VIR) de Rusia, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) de Alemania o la del Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV) y Banco de Germoplasma de Especies Hortícolas del Centro de Investigación y Tecnología Agroalimentaria (BGHZ) de España. Es de destacar la colección que se mantiene en el TGRC que incluye más de 1.000 accesiones de especies silvestres de tomate, una colección de más de 1.000 mutantes monogénicos y 210 cultivares modernos y tradicionales, además de otros recursos fitogenéticos.

Pese a la importancia de los bancos de germoplasma como fuentes de diversidad genética, es importante señalar que, debido a la escasez de recursos, se carece de un buen conocimiento y una buena caracterización de las colecciones disponibles. Por eso, a parte de

la identificación de algunos genes de interés en especies silvestres para mejora, su uso de manera rutinaria es todavía limitado (Ranjan et al., 2012).

3.4.2. Uso de *S. pimpinellifolium*

Solanum pimpinellifolium es una especie autocompatible, de fruto pequeño y rojo, muy inferior al del tomate cultivado. Se encuentra ampliamente distribuida desde el norte de Ecuador al sur de Perú, encontrándose en márgenes y valles de ríos o en zonas costeras (Warnock, 1991).

Atendiendo a su origen, esta especie silvestre presenta diversos grados de diversidad genética (Rick, Fobes, & Holle 1977; Caicedo & Schaal 2004; Sifres et al. 2006; Zuriaga et al. 2009), así como de variabilidad morfológica (Widrechner 1987; Rick, Holle, & Thorp 1978). Esta variabilidad ha sido descrita tanto para el color del fruto como para caracteres morfológicos florales, que incluyen el tamaño de flores e inflorescencias o ejerción estigmática y vegetativos.

Su capacidad para hibridar con el tomate cultivado ha facilitado su uso en los programas de mejora, principalmente aquellos dirigidos a desarrollar variedades resistentes. Dentro de los genes de resistencia descritos en la especie (Tabla 2, basado en (Barrantes et al., 2014; Foolad et al., 2008; Nuez & Díez, 2008; Rivas & Thomas, 2005), cabe destacar el éxito del gen *Pto* que ha demostrado ser una resistencia efectiva frente a *Pseudomonas syringae* (Kraus et al., 2016). Además, esta especie también se ha descrito como posible fuente de resistencia a múltiples enfermedades como enfermedades víricas causadas por TYLCV (Picó et al., 2000) o TLCV (Banerjee & Kalloo, 1987); hongos como *Alternaria solani* ((Alexander & Hoover, 1955; Foolad et al., 2008), *Cladosporium fulvum* (Joosten & de Wit, 1999) u *Oidium lycopersici* (Mieslerová et al., 2000) y bacterias como *Corynebacterium michiganense* (De Jong & Honma, 1976). Por último, *S. pimpinellifolium* también ha sido empleada como fuente de resistencia frente a *Fusarium oxysporum*, *Verticillium dahliae* o *Stemphylium solani* (Soumpourou et al. 2007; Stevens & Rick 1986; Kalloo 1991).

Otra de las características interesantes de esta especie, en lo que a resistencia a insectos se refiere, es que en *S. pimpinellifolium* se ha descrito la presencia de tricomas glandulares de tipo IV, relacionados con la resistencia a diferentes plagas (Fernández-Muñoz et al. 2003; Firdaus et al. 2012; Rakha et al. 2016; Rakha, Hanson, & Ramasamy 2017), entre ellas a mosca blanca, que es el vector del *Tomato yellow leaf curl virus* (TYLCV) y provoca grandes pérdidas económicas en el cultivo del tomate (García-Cano et al., 2008). Pese a que la presencia de tricomas glandulares tipo IV en *S. pimpinellifolium* se desarrollará ampliamente más adelante, es importante destacar que estos tricomas también se han descrito en especies silvestres de fruto verde como *S. habrochaites* y *S. pennellii* (Guo et al., 1993; Gurr & McGrath, 2001a; Oriani et al., 2010). Sin embargo, la ventaja del uso de *S. pimpinellifolium* radica en que es una especie silvestre mucho más cercana al tomate cultivado. Esto permite la disminución de la transmisión de caracteres no deseables, más abundantes en el fondo genético de especies silvestres más alejadas genéticamente.

Por último, cabe destacar la importancia de *S. pimpinellifolium* en mejora como fuente de caracteres de interés relacionados con la calidad del fruto. Dentro de éstos destaca el alto contenido en licopeno (Chen et al. 1999), color del fruto (Grandillo & Tanksley 1996), contenido en azúcares (Grandillo & Tanksley 1996) o compuestos volátiles (Capel et al., 2015; Rambla et al., 2017). También se han descrito QTLs relacionados con la forma del fruto (*ovate*, *fs8.1* o *sun*) en poblaciones procedentes de cruces entre tomate cultivado y *S. pimpinellifolium* (Grandillo, Termolino, & van der Knaap 2013).

Tabla 2. Genes de resistencia a enfermedades descritos en *S. pimpinellifolium* e introgresados en tomate cultivado.

Enfermedad	Agente causal	Gen descrito
Bacterias		
Peca bacteriana	<i>Pseudomonas syringae</i> pv. <i>tomato</i>	<i>Pto</i>
Hongos		
Verticilosis	<i>Verticillium dahliae</i>	<i>Ve</i>
Fusariosis	<i>Fusarium oxysporium</i> f. sp. <i>lycopersici</i>	
	patotipo 0	<i>I</i>
	patotipo 1	<i>I-2</i>
Mancha gris de la hoja	<i>Stemphylium</i> spp.	<i>Sm</i>
Cladosporiasis del tomate	<i>Cladosporium fulvum</i>	<i>Cf 1, 2, 3, 4, 5 y 9</i>
Mildiu del tomate	<i>Phytophthora infestans</i>	<i>Ph-1, Ph-2, Ph-3</i>
Nematodos		
Nematodos del quiste	<i>Globodera restochiensis</i>	<i>Hero</i>
Virus		
Virus del rizado amarillo	TYLCV	<i>Tylc</i>
Virus del rizado de la hoja	TLCV	<i>Tlc</i>

3.4.3. Uso de *S. lycopersicum* var. *cerasiforme*

S. lycopersicum var. *cerasiforme* es una especie autocompatible, que se caracteriza por poseer generalmente frutos rojos y redondos de pequeño tamaño. Sin embargo, Rick & Holle (1990) describieron la presencia de variabilidad morfológica para caracteres relacionados con la morfología del fruto como el tamaño, la forma o el color. Estos mismos autores describieron también la presencia de variabilidad en la forma y tamaño de las hojas, hábito de crecimiento de la planta y pilosidad.

Esta especie es nativa de la región andina de Ecuador y del norte de Perú (Rick & Fobes 1975), pero se encuentra ampliamente distribuida por América (Peralta & Spooner, 2006). Los estudios realizados sobre la variabilidad genética de la especie indican que las accesiones procedentes de esta región de origen presentan una mayor diversidad que las originarias de Mesoamérica (Rick & Fobes 1975; Blanca et al. 2012; Razifard et al. 2020).

Esta especie se distribuye en una gran diversidad de hábitats y en diversos nichos ecológicos que abarcan desde regiones áridas a zonas húmedas y desde el nivel del mar a los 2.400 m de altitud (Cuartero et al., 1985). Debido a su adaptabilidad a una gran diversidad de ambientes, tiene especial interés como fuente de genes de adaptación a estreses bióticos y abióticos (Nuez & Díez, 2008). Por otro lado, también ha sido descrita como fuente de resistencia frente a *Verticillium dahliae* (Kalloo, 1991), *Phytophthora* (Arellano Rodríguez et al., 2013), *Xanthomonas* (Hutton et al., 2010), *Alternaria solani* (Foolad et al., 2008) y *Cladosporium fulvum* (Rivas & Thomas, 2005). En este último caso, esta especie se ha identificado como portadora del gen *Cf-5* que pertenece a la familia de genes de resistencia *Cf*.

Por último, esta especie constituye una fuente interesante en los programas de mejora dirigidos a la mejora de la calidad del fruto. Diferentes líneas de introgresión con tomate cultivado han permitido identificar QTLs relacionados con compuestos volátiles (Saliba-

Colombani et al., 2001; Sauvage et al., 2014), contenido en ácido ascórbico (Stevens et al. 2007), sólidos solubles, dulzor o acidez y carotenos (Saliba-Colombani et al., 2001). Así mismo, se han encontrado mutaciones en los genes que regulan el tamaño y forma del fruto. Esta especie ha sido descrita como portadora de mutaciones en los genes *lc*, *fas* u *ovate* (Rodríguez et al., 2011), cuya función aparece descrita más adelante.

3.4.4. Uso de otras especies

Las especies *S. pimpinellifolium* y *S. lycopersicum* var. *cerasiforme* serían, en principio, las especies de mayor interés para la mejora del tomate cultivado puesto que permiten solventar uno de los mayores inconvenientes en los programas de mejora, el desequilibrio de ligamiento. El empleo de estas especies permite reducir el desequilibrio de ligamiento al ser especies más próximas filogenéticamente al tomate. Pese a esto, durante mucho tiempo, también han sido ampliamente empleadas en mejora otras especies como *S. habrochaites*, *S. peruvianum* y *S. chilense* para el desarrollo de cultivares élite a partir de los que introgresar los genes de interés mediante programas de retrocruzamiento (Bai & Lindhout 2007; Rick 1978). El problema del uso de estas especies más alejadas filogenéticamente es que requieren muchas generaciones para poder eliminar los caracteres indeseables ligados al carácter de interés, haciendo que el proceso sea largo y tedioso (Zamir, 2001). En la Tabla 3 se describen algunos caracteres de interés identificados en las especies pertenecientes al género *Solanum*, sección *Lycopersicon*.

Tabla 3. Especies silvestres del género *Solanum*, sección *Lycopersicon* usadas en mejora. Se recoge el hábitat, sistema reproductivo (SR), color del fruto y uso en mejora. AC: autocompatible, AI:autoincompatible, AL: alógama, ALF: alógama facultativa, AU: autógama. V: Verde. VP: Verde y piloso. VPU: verde con rayas púrpuras. BPU: blanco con rayas púrpuras. VV: verde con rayas verde oscuro. RP: rojo y pequeño. RV: rojo y de tamaño variable. A: Amarillo. N: Naranja. (Spooner et al., 2005; Peralta & Spooner, 2014).

Espece	Hábitat	SR	Fruto	Uso en mejora
<i>S. pennellii</i>	Del norte al sur de Perú. Laderas rocosas secas y zonas arenosas. 2.300m.	AI. AC.	V.	Tolerancia a la salinidad y sequía. Resistencia a bacterias. Calidad.
<i>S. habrochaites</i>	Centro de Ecuador al centro de Perú. 3.300m.	AI. AC.	VP.	Tolerancia a frío. Resistencia a insectos y enfermedades. Calidad.
<i>S. chilense</i>	Desiertos de Chile y Perú. 3.250m.	AI. AL.	VPU. BPU.	Resistencia a sequía y a enfermedades.
<i>S. huaylasense</i>	Laderas rocosas del norte de Perú. 1.700-3.000m.	AI. AL.	VV.	Resistencia a virus, hongos, bacterias e insectos. Calidad*.
<i>S. peruvianum</i>	Lomas peruvianas y desiertos costeros. 2.900 m.	AI. AL.	V.	
<i>S. arcanum</i>	Norte de Perú. Valles costeros y laderas rocosas. 2.800 m.	AI. AL.	V.	
<i>S. chmielewskii</i>	Centro de Perú al norte de Bolivia. Medios rocosos y bien drenados. 1.600-3.200m.	AI. ALF.	VV.	Calidad.
<i>S. neorickii</i>	Centro de Perú al norte de Bolivia. Medios rocosos y bien drenados. 1.600-3.200m.	AC. AU.	VV.	Resistencia a enfermedades. Calidad.
<i>S. cheesmaniae</i>	Endémica de las Islas Galápagos. Desde el nivel del mar a zonas volcánicas.	AC. AU.	A. N.	Tolerancia a la salinidad. Resistencia a insectos y enfermedades.

Tabla 3. Continuación.

Especie	Hábitat	SR	Fruto	Uso en mejora
<i>S. galapagense</i>	Endémica de las Islas Galápagos. Rocas volcánicas costeras y ocasionalmente tierra adentro. 50 m.	AC. AU.	A. N.	Tolerancia a la salinidad.

* La valoración del uso en mejora de *S. peruvianum*, *S. huaylasense* y *S. arcanum* se hace de forma conjunta, ya que estas especies son el resultado de la escisión de *Lycopersicon peruvianum*.

4. Mejora genética del tomate cultivado

A medida que las necesidades de cultivo y producción del tomate han ido cambiando, el objetivo de los programas de mejora genética se ha ido modificando consecuentemente. El aumento de la producción mediante la reducción de los daños ocasionados por plagas y enfermedades fue uno de los objetivos de mejora iniciales, que se mantiene todavía en la actualidad. A partir de los años 70, comenzó a darse una globalización del mercado que supuso el transporte del tomate hacia regiones lejanas a su lugar de producción. Por lo tanto, los programas de mejora en los años 80 se centraron en alargar la vida postcosecha de los frutos y hacerlos más firmes para soportar la manipulación y el transporte (Bai & Lindhout, 2007). Como resultado, se produjo un detrimento de la calidad del tomate ya que, por ejemplo, el retraso de la maduración conlleva una reducción en la acumulación de carotenoides o compuestos volátiles (Baldwin et al., 2000). Por ello, la preferencia del consumidor por tomates de calidad entró en juego en los años 90. A partir de este momento, los programas de mejora se redirigieron hacia la mejora del sabor y la calidad del fruto (Bai & Lindhout, 2007).

Actualmente, los objetivos de mejora varían ligeramente según se trate de tomate para consumo en fresco o procesado. En el caso del consumo en fresco se incide en la forma y color del fruto, uniformidad en la apariencia, larga vida postcosecha, sabor o textura. Por otro lado, en el tomate procesado se busca un crecimiento compacto y ausencia de zona de abscisión que faciliten la recogida de los frutos a máquina, hábito de crecimiento determinado y caracteres relacionados con la calidad del fruto como son la viscosidad, el pH, acidez o el contenido en sólidos solubles. A parte de estos aspectos relacionados con la calidad, la resistencia a enfermedades y la adaptabilidad de los cultivos a los diferentes ambientes son objetivos comunes a ambos mercados (Foolad, 2007).

4.1. Resistencia a insectos

El tomate cultivado es susceptible a unas 200 enfermedades diferentes que están causadas por hongos, bacterias, virus y nematodos. Adicionalmente, también es susceptible a diversidad de plagas ocasionadas por más de 200 especies distintas de insectos y ácaros, que a su vez además pueden actuar como vectores de transmisión de virus. Tanto las plagas como las enfermedades a las que es susceptible el tomate cultivado pueden ocasionar grandes pérdidas económicas, puesto que se ven afectadas características como la calidad del fruto o la producción de la planta (Stout et al., 2017). A continuación, en la Tabla 4, se recogen algunas de las principales plagas y enfermedades que afectan al cultivo del tomate, así como el agente causante de las mismas.

Tabla 4. Principales enfermedades y plagas que afectan al cultivo de tomate.

Enfermedades bacterianas	
Chancro bacteriano	<i>Clavibacter michiganensis subsp. michiganensis</i>
Marchitamiento bacteriano	<i>Ralstonia solanacearum</i>
Pudrición bacteriana del tallo	<i>Erwinia carotovora subsp. carotovora</i>
Peca bacteriana	<i>Pseudomonas syringae</i> pv. <i>tomato</i> y <i>P. syringae</i> pv. <i>syringae</i>
Mancha bacteriana	<i>Xanthomonas campestris</i> pv. <i>versicatoria</i>
Enfermedades fúngicas	
Mildiu	<i>Phytophthora infestans</i>
Alternariosis	<i>Alternaria solani</i>
Septoriasis	<i>Septoria lycopersici</i>
Cladosporiasis	<i>Cladosporium fulvum</i>
Fusariosis	<i>Fusarium oxysporum</i>
Verticiliosis	<i>Verticillium spp.</i>
Oidio	<i>Oidium lycopersici</i> , <i>O. neolycopersici</i> , <i>Leveillula taurica</i>
Mancha gris de la hoja	<i>Stemphylium spp.</i>
Moho de la hoja	<i>Fulvia fulva</i>
Moho gris	<i>Botrytis cinerea</i>
Pythium	<i>Pythium spp.</i>
Moho blanco	<i>Sclerotinia sclerotiorum</i>
Pudredumbre de la raíz	<i>Pyrenochaeta lycopersici</i>
Nemátodos	
Agallas en la raíz	<i>Meloidogyne spp.</i>
Enfermedades víricas	
Virus del mosaico del tomate (ToMV)	
Virus del mosaico del pepino (PepMV)	
Virus del bronceado del tomate (TSWV)	
Virus del rizado amarillo (TYLCV)	
Virus del rizado de la hoja (TLCV)	
Virus Y de la patata (PVY)	
Mosaico de la alfalfa (AMV)	
Virus del mosaico del pepino dulce (PepMV)	
Insectos	
Mosca blanca	<i>Bemisia tabaci</i>
	<i>Bemisia argentifolii</i>
Trips	<i>Frankliniella occidentalis</i>
Pulgón verde	<i>Myzus persicae</i>
	<i>Macrosiphum euphorbiae</i>
Escarabajo de la patata	<i>Leptinotarsa decemlineata</i>
Polilla del tomate	<i>Tuta absoluta</i>
Minador	<i>Liriomyza huidobrensis</i>
Oruga del tabaco	<i>Helicoverpa zea</i>
Ácaros	
Araña roja	<i>Tetranychus urticae</i>

Debido a su elevada susceptibilidad, el cultivo del tomate es altamente dependiente del uso de pesticidas y plaguicidas. Sin embargo, estas sustancias llevan asociadas grandes desventajas como el aumento de los costes de producción, el impacto en la salud o el medioambiente, el daño a poblaciones de artrópodos beneficiosos o el desarrollo de

resistencias (Stout et al., 2017). Por tanto, el uso de variedades resistentes se presenta como la mejor alternativa para solventar estos problemas. Las principales fuentes de resistencia han sido descritas en las especies silvestres *S. pimpinellifolium*, *S. habrochaites* y *S. peruvianum* (Scott, 2007) que se han empleado en el desarrollo eficiente de cultivares resistentes (Nuez & Díez, 2008).

4.1.1. Mecanismos de resistencia a insectos

En cuanto a la resistencia a insectos, uno de los mecanismos más destacados son los metabolitos secundarios que las especies silvestres sintetizan y almacenan en los tricomas glandulares (Stout et al., 2017). Los tricomas son estructuras uni o pluricelulares que recubren la superficie de las hojas y del tallo y, en ocasiones, de los frutos. Luckwill (1943) hizo la primera clasificación de tricomas, posteriormente actualizada por Channarayappa (1992), que divide a los tricomas en dos tipos según la presencia o ausencia de glándula. De este modo, los tricomas glandulares engloban los tipos I, IV, VI y VII; mientras que el grupo de los tricomas no glandulares está constituido por los tipos II, III, V y VIII (Figura 5).

Los tricomas glandulares se pueden diferenciar por el número de células que los conforman y por el tipo de sustancias químicas que almacenan en el interior de las células secretoras, aunque la diversidad de estas sustancias depende de la especie (McDowell et al., 2011). Los tricomas de los tipos I y IV almacenan y secretan acil azúcares mientras que los tricomas de tipo VI secretan monoterpenos y sesquiterpenos, principalmente. Los tricomas de tipo VII principalmente almacenan alcaloides por lo que son menos implicados en la resistencia a insectos (Glas et al., 2012).

En el caso de los tricomas no glandulares parece que su presencia puede afectar e interferir con el movimiento de los insectos; mientras que los glandulares, al entrar en contacto con el insecto, liberan el contenido almacenado en su interior. Estos compuestos almacenados, al liberarse, afectan al insecto de diversas maneras: a la ovoposición, envenenando al insecto, haciendo que quede pegado a la hoja o actuando como repelente (Stout et al., 2017).

Las resistencias más efectivas han sido encontradas en los tricomas glandulares de tipo IV y VI. Por un lado, los tricomas tipo VI han demostrado su eficiencia frente a araña roja (Guo et al., 1993) o mosca blanca (Channarayappa et al., 1992). Por otro lado, se ha descrito la efectividad de los tricomas tipo IV frente a mosca blanca, araña roja y áfidos (Carter & Snyder 1985; Guo, Weston, & Snyder 1993; Fernández-Muñoz et al. 2003; Oriani et al. 2010; Firdaus et al. 2012; Rakha et al. 2016; Andrade et al. 2017). La presencia de tricomas tipo IV ha sido descrita solo en algunas especies silvestres del género *Solanum* (Tabla 5), mientras que el tipo VI se ha descrito en entradas de todas las especies salvo *S. cheesmaniae* y *S. galapagense* (Glas et al., 2012). Sin embargo, el principal problema de su uso en mejora es que tienen una fuerte influencia ambiental. Se ha descrito el efecto de la edad de la planta y de la hoja (Wilkins et al. 1996; Gurr & McGrath 2001b; Saeidi, Mallik, & Kulkarni 2007; Vendemiatti et al. 2017), de la luz (Weston & Snyder, 1990) o de la disponibilidad de agua (Wilkins et al., 1996) en la presencia y densidad de tricomas glandulares.

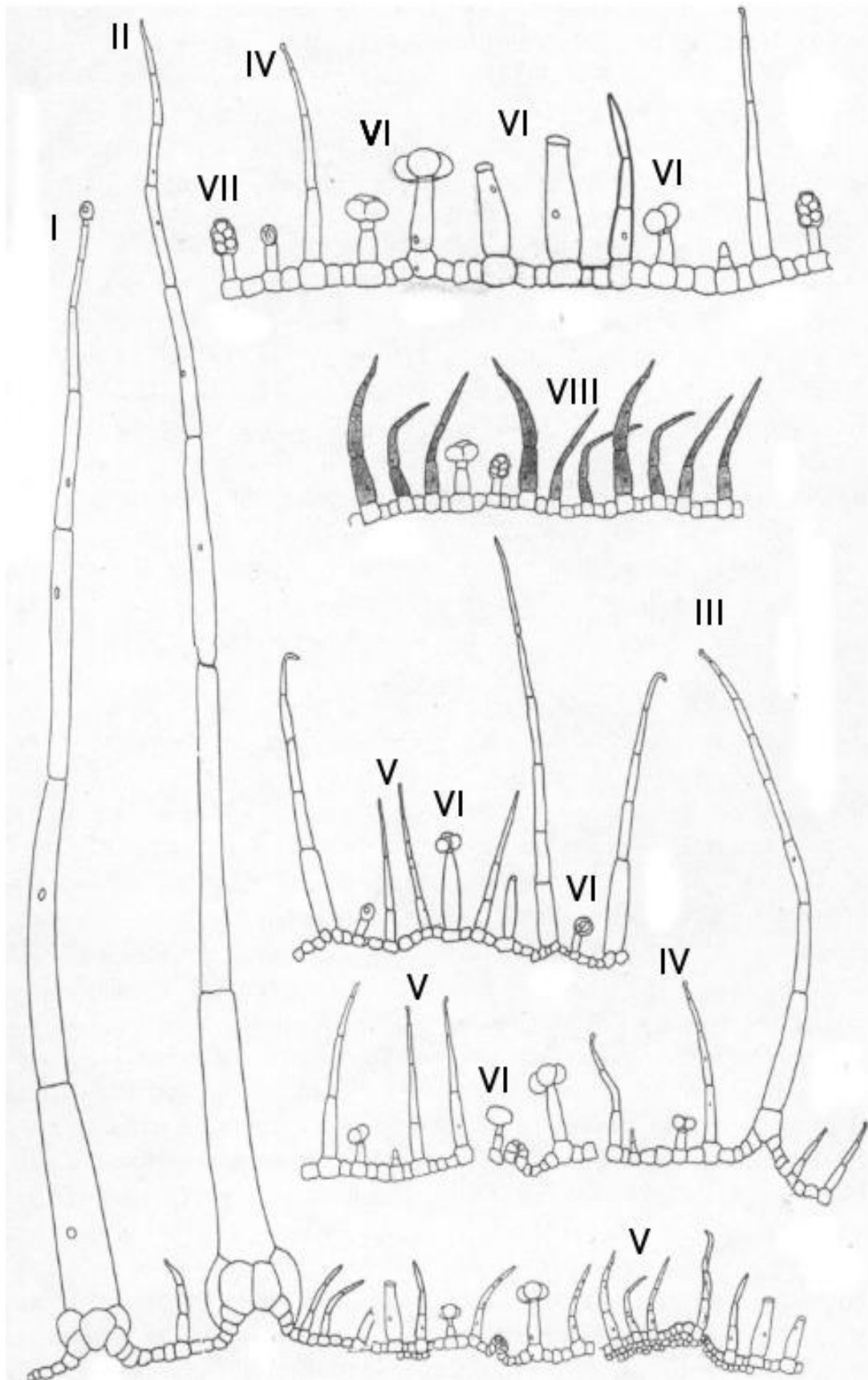


Figura 5. Tipos de tricomas en tomate y especies silvestres relacionadas (Modificado de Peralta, Spooner & Knapp, 2008).

Tabla 5. Principales especies en las que se ha descrito la presencia de tricomas glandulares tipo IV.

Especie	Referencia
<i>S. pennellii</i>	Goffreda & Mutschler (1989); Lemke & Mutschler (1984); Blauth, Churchill, & Mutschler (1998); Mutschler et al. (1996) ; Leckie, de Jong, & Mutschler (2012)
<i>S. galapagense</i>	Andrade et al. (2017; Firdaus et al. (2013)
<i>S. habrochaites</i>	Carter & Snyder (1985); Maliepaard et al. (1995); de Oliveira et al. (2018)
<i>S. pimpinellifolium</i>	Fernández-muñoz & Salinas (2003); Firdaus et al. (2012); Rakha, Hanson y Ramasamy (2017)

4.1.2. QTLs asociados a la resistencia a insectos

Los tricomas glandulares de tipo IV son uno de los principales mecanismos de resistencia a insectos, gracias a su capacidad de almacenar y secretar acil azúcares y otros metabolitos secundarios.

La accesión ‘LA0716’ de *S. pennelli*, es una de las más empleadas en las poblaciones de mapeo de QTLs para resistencia a insectos. Esta accesión ha permitido la identificación de QTLs asociados a niveles de acil azúcares (Mutschler et al. 1996; Blauth, Churchill, & Mutschler 1998; Leckie, de Jong, & Mutschler 2012) y densidad de tricomas tipo IV (Blauth et al., 1998). Sin embargo, el gran desequilibrio de ligamiento junto con la complejidad del control genético del carácter, han hecho que sea muy compleja la transmisión de los QTLs al tomate cultivado (Lawson et al., 1997).

También se han identificado QTLs en poblaciones creadas a partir de *S. galapagense* como parental resistente. Firdaus *et al.* (2013) describieron un QTL responsable de la resistencia a la mosca blanca y de la densidad de tricomas tipo IV en el cromosoma 2. Este mismo QTL fue posteriormente asociado con la resistencia a dos biotipos de mosca blanca (Vosman et al., 2019) y se encuentra próximo a la región detectada en los estudios llevados a cabo por Andrade *et al.* (2018) en esta especie. Este último grupo describe un QTL en el cromosoma 2 y otro en el 3 que estarían asociados a la densidad de tricomas tipo IV en esta especie.

En la especie *S. habrochaites* también se han descrito QTLs en los cromosomas 5 y 9. Sin embargo, no se pudo verificar la asociación entre este tipo de tricomas y la resistencia a mosca blanca (Maliepaard et al., 1995).

Finalmente, la accesión ‘TO-937’ de la especie *S. pimpinellifolium* también ha sido descrita como portadora de tricomas glandulares tipo IV (Fernández-Muñoz et al., 2003), cuya presencia se ha relacionado con la resistencia a araña roja. Posteriormente, se identificaron dos QTLs en el cromosoma 2 responsables de la resistencia en dicha entrada (Salinas et al., 2013).

A continuación, en la Tabla 6, aparecen recogidos todos los QTLs que han sido descritos en el tomate cultivado y especies silvestres relacionadas por estar asociados a la presencia de tricomas glandulares tipo IV o bien, por presentar una resistencia a insectos que está mediada por la presencia de este tipo de tricomas glandulares.

Tabla 6. QTLs descritos para la presencia de tricomas tipo IV, resistencia conferida por estos tricomas o síntesis de acil azúcares.

Cruces	Familia	Cromosoma	Referencia
<i>S. lycopersicum</i> var. <i>lycopersicum</i> x <i>S. pennellii</i> (New Yorkwe x LA716)	F2	2, 3, 4 y 11	Mutschler <i>et al.</i> , 1996
<i>S. pennellii</i> x <i>S. pennellii</i> (LA716 x LA1912)	F2 y BC	2, 4, 5, 6, 7, 10 y 11	Blauth, Churchill & Mutschler, 1998
<i>S. lycopersicum</i> var. <i>lycopersicum</i> x <i>S. pennellii</i> (New Yorwe x LA716)	BC	4, 5, 6, 10 y 11	Leckie, de Jong & Mutschler, 2012
<i>S. lycopersicum</i> var. <i>lycopersicum</i> x <i>S. galapagense</i> (Moneymaker x PRI95004)	F2	2 y 9	Firdaus <i>et al.</i> , 2013
<i>S. lycopersicum</i> var. <i>lycopersicum</i> x <i>S. galapagense</i> (TOM-684 x LA1401)	F2 y BC	2 y 3	Andrade <i>et al.</i> , 2018
<i>S. lycopersicum</i> var. <i>lycopersicum</i> x <i>S. galapagense</i> (Moneymaker x PRI950074)	F7	2	Vosman <i>et al.</i> , 2019
<i>S. lycopersicum</i> var. <i>lycopersicum</i> x <i>S. habrochaites</i> (Moneymaker x CGN1.1561)	F2	5 y 9	Maliepaard <i>et al.</i> , 1995
<i>S. lycopersicum</i> var. <i>lycopersicum</i> x <i>S. pimpinellifolium</i> (Money Maker x TO-937)	F4	2	Salinas <i>et al.</i> , 2013

4.2. Mejora de la calidad

La calidad es un concepto muy amplio que varía según se tengan en cuenta las preferencias del comercializador o del consumidor. El comercializador busca un producto que sea uniforme, que tenga una larga vida postcosecha y que sea resistente a la manipulación desde la recogida hasta su posterior venta. Sin embargo, el consumidor busca un producto de calidad donde son importantes la apariencia del fruto, su sabor, textura y calidad nutricional.

Por lo tanto, el concepto de mejora de la calidad engloba aspectos relacionados tanto con la apariencia externa del fruto, como con su firmeza y sabor. Por un lado, se buscan frutos con determinadas características de tamaño, forma, color y firmeza. Mientras que, por otro lado, la calidad organoléptica engloba aspectos como el sabor del fruto, la textura y el aroma. Además, el sabor del fruto es un carácter complejo que está definido por su contenido en ácidos, azúcares y compuestos volátiles.

4.2.1. Vida postcosecha

Uno de los objetivos de mejora relacionados con la calidad ha sido el desarrollo de cultivares con larga vida postcosecha. Esta mejora ha ido asociada al uso de mutantes naturales que presentan alterada la maduración del fruto. La principal desventaja de estos mutantes es que el aumento de la vida postcosecha ha ido en detrimento de la calidad. Sin embargo, esta desventaja ha sido asumida para reducir al máximo las pérdidas (Nuez & Díez, 2008).

El mutante *Never ripe (Nr)* impide la maduración del fruto mediante su acción en receptores de etileno, volviendo al fruto insensible al mismo (Lanahan et al., 1994). Los mutantes *ripening inhibitor (rin)* y *non-ripening (nor)* también actúan en la ruta del etileno, impidiendo tanto su síntesis como la reacción a la aplicación de etileno de forma exógena (Paran & van der Knaap, 2007). El mutante *Alcobaca (alc)* es un mutante de maduración lenta del fruto. Debido a una disminución de los niveles de poligalacturonasa y etileno, es capaz de retrasar la velocidad de ablandamiento del fruto maduro (Kopeliovitch et al., 1980; Mutschler et al., 1992).

4.2.2. Peso, forma y tamaño

El incremento en el tamaño del fruto fue uno de los cambios más significativos que tuvieron lugar durante el proceso de domesticación del tomate. Este tamaño depende tanto del número y volumen de células como del periodo de crecimiento y expansión celular (Bohner & Bangerth, 1988).

Se han identificado QTLs relacionados con el tamaño del fruto en todos los cromosomas en poblaciones procedentes del cruce de especies silvestres de pequeño tamaño con tomate cultivado de fruto grande (Grandillo, Ku, & Tanksley 1999). Algunos de los QTLs más importantes son *fw1.1*, *fw2.2*, *fw3.1* y *fw4.1*. Todos ellos han sido identificados tanto en varias poblaciones de mapeo como en líneas casi isogénicas y explican un alto porcentaje del incremento del tamaño del fruto (Tanksley 2004). Salvo *fw3.1*, solo afectan al tamaño del fruto y no a otros caracteres como la forma (Tanksley 2004). De todos ellos, *fw2.2* es uno de los más importantes ya que explica un incremento en el tamaño de hasta el 30% en las variedades de tomate cultivado, respecto de las especies silvestres (Frary et al., 2000). Este QTL ha sido clonado y estudiado a nivel molecular (Frary et al., 2000), donde se ha descrito que actúa como represor negativo de la división celular. El QTL *fw3.2* también ha sido clonado y se ha descrito que afecta al tamaño del fruto mediante el incremento del número de capas celulares en el pericarpio (Monforte et al., 2014). Por el contrario, el QTL *fw11.3* actúa mediante el incremento en el tamaño de las células del pericarpio (Mu et al., 2017).

La diversidad en la forma del fruto (Figura 6) está explicada principalmente por mutaciones en los genes *sun*, *ovate*, *locule number (lc)* y *fasciated (fas)* (Monforte et al., 2014). *Sun* y *ovate* controlan la elongación del fruto; siendo los responsables de las formas alargadas o pera, respectivamente. Por el contrario, *fas* y *lc* afectan al tamaño del fruto mediante cambios en el número de carpelos de la flor (Tanksley 2004). De esta forma también se ve afectado el número de lóculos del fruto, que a su vez afecta indirectamente a la forma y tamaño del fruto (Rodríguez et al., 2011). Por último, el locus *fs8.1* es responsable de la forma alargada y cuadrada de los tomates para procesado industrial (Paran & van der Knaap, 2007).



Figura 6. Variabilidad en la forma del fruto de tomate cultivado y especies silvestres relacionadas.

4.2.3. Color del fruto

El color del fruto depende principalmente de su contenido en carotenoides y de la proporción de los distintos tipos de carotenoides. El licopeno es el carotenoide que se encuentra en mayor proporción en el tomate y es responsable de su color rojo. Debido a los beneficios del licopeno para la salud, el interés por incrementar sus niveles en el tomate ha ido en incremento en los últimos años (Mazidi et al., 2020). Mutaciones espontáneas relacionadas con un alto contenido del mismo, *hp1* y *hp2*, han sido introgresadas en cultivares comerciales. Sin embargo, su uso en mejora está limitado por los efectos pleiotrópicos de los genes *hp* (Jarret et al., 1984). Como alternativa, se ha empleado el gen *og^c*, que incrementa el contenido en licopeno a expensas del betacaroteno (Butler, 1962). Por otro lado, también se han identificado mutaciones responsables del color amarillo o naranja de los frutos de tomate como el mutante *yellow flesh*, *sherry*, *tangerine*, *Beta*, *apricot* o *Delta* (Ilan Levin et al. 2006; Foolad 2007; Bergougnoux 2014). En la siguiente tabla (Tabla 7) se recogen los principales mutantes de color en tomate y sus características.

Tabla 7. Principales mutantes de color del fruto descritos en tomate (Stommel, 2014).

Locus	Cromosoma	Descripción	Referencia
<i>hp2</i>	1	Aumenta los niveles de licopeno.	Yen et al. 1997
<i>Colorless fruit epidermis (y)</i>	1	Epidermis no pigmentada que da lugar a color rosa.	Rick & Butler 1956
<i>hp1</i>	2	Aumenta los niveles de licopeno.	Lieberman et al., 2004
<i>yellow (r)</i>	3	Color amarillo. Bajos niveles de carotenos. Epistático con <i>t</i> .	Kachanovsky et al., 2012
<i>Apricot (at)</i>	5	Pulpa de color amarillo con manchas rosadas.	Jenkins & Mackinney 1955
<i>Beta-carotene (B)</i>	6	Color anaranjado. Incremento en β -carotenos y reducción de licopenos.	Ronen et al., 2000
<i>Old gold crimson (og^c)</i>	6	Color rojo por aumento de licopeno y reducción de β -carotenos.	Ronen et al., 2000

4.2.4. Sólidos solubles, acidez y pH

Uno de los factores a tener en cuenta por su efecto en el sabor del tomate es la proporción de sólidos solubles. Los sólidos solubles engloban diversos compuestos como ácidos orgánicos, lípidos, minerales y azúcares. De todos ellos, son los ácidos orgánicos y los azúcares reducidos los que representan la parte principal del contenido de sólidos solubles. Dentro de los ácidos orgánicos destacan el ácido málico y cítrico, mientras que glucosa y fructosa serían los más importantes dentro de azúcares reducidos.

S. habrochaites y *S. chmielewskii* han sido las especies descritas como principales fuentes para la mejora del contenido en azúcares. Por ejemplo, en *S. chmielewskii* se ha descrito la mutación *sucr* como responsable de un alto contenido en sacarosa en vez de glucosa y fructosa (Chetelat et al., 1995). En cuanto a *S. habrochaites*, el gen *Fgr* se ha descrito como responsable del aumento en el dulzor por incremento en la concentración de fructosa del fruto (Levin et al. 2000).

Los ácidos orgánicos determinan la acidez y el pH del producto final. El valor de este último es muy importante en tomate procesado, ya que para evitar el crecimiento de microorganismos patógenos debe ser inferior a 4,5 (Foolad, 2007). Debido a que el ácido cítrico es el principal responsable de la acidez del fruto, la mejora va enfocada a aumentar el contenido de este por encima de 0,35 g / 100 g de peso fresco con un pH inferior a 4,5 (Foolad, 2007). Se han identificado QTLs para acidez y pH mediante el uso de líneas isogénicas de *S. pennellii* (Causse et al. 2003) y mediante análisis de asociación genómica en colecciones de *S. lycopersicum* var. *lycopersicum*, *S. lycopersicum* var. *cerasiforme* y *S. pimpinellifolium* (Causse et al. 2003; Bauchet, Grenier, Samson, Segura, et al. 2017; Tieman et al. 2017; Ye et al. 2019).

5. Uso de marcadores genéticos para el aprovechamiento de los recursos fitogenéticos

El establecimiento y mantenimiento de las colecciones de germoplasma debe ir acompañado de una utilización activa de las mismas, lo que requiere que tanto la diversidad fenotípica como la diversidad genética de dichas colecciones se encuentre bien caracterizada.

El desarrollo de las técnicas de secuenciación masiva ha permitido la secuenciación de genomas completos a un bajo coste y la obtención de millones de marcadores (Sim et al., 2012a). Estas técnicas han supuesto una revolución en la mejora genética asistida por marcadores y en la capacidad para descubrir nuevos genes, validar y evaluar marcadores genéticos en las poblaciones de interés agronómico. Además, también permiten el estudio de grandes poblaciones, facilitando así la conservación de la diversidad genética (Bauchet & Causse, 2012). Por lo tanto, estas técnicas han aumentado enormemente la cantidad de recursos genómicos disponibles.

Así pues, el advenimiento de los secuenciadores de nueva generación y la posibilidad de generar miles de marcadores a un bajo costo ha impulsado, entre otras aplicaciones, el mapeo de rasgos cuantitativos (QTLs) en mapas ultradensos en diversas especies (Gong et al., 2019; S. Li et al., 2019; Liu et al., 2019) con un grado de resolución impensable hasta hace relativamente poco. También ha abierto el camino a estudios de asociación genómica GWAS (Genome-wide association studies) (Huang et al. 2010; Xiao et al. 2017; Zhu et al. 2018a; Hanson et al. 2018). Ambos métodos de análisis permiten explotar la diversidad genética presente en los bancos de germoplasma y acelerar el descubrimiento de genes de interés.

5.1. Análisis de QTLs en mapas ultradensos

El objetivo del mapeo de QTLs es determinar aquellas regiones del genoma que controlan un carácter fenotípico de interés. Esto se hace mediante una búsqueda sistemática de asociaciones entre el genotipo para cada marcador y el valor fenotípico del carácter. Para ello, es necesario disponer de una población de mapeo adecuada que permita obtener la posición y efecto de los QTLs y generar un mapa genético que integre la variación en los marcadores genéticos. Como se ha comentado anteriormente, la diversidad genética del tomate cultivado es muy reducida. Como consecuencia, la alta variación presente en las especies silvestres es necesaria a la hora de desarrollar poblaciones de mapeo, y así poder disponer de un número suficiente de polimorfismos (Grandillo & Cammareri 2016). Las primeras poblaciones de mapeo empleadas en la búsqueda de QTLs fueron poblaciones F2 y retrocruces (BC), aunque posteriormente también se han empleado líneas recombinantes consanguíneas (RIL) o líneas de introgresión (IL) usando generalmente *S. pennellii*, *S. pimpinellifolium* o *S. habrochaites* como parentales masculinos en cruces con el tomate cultivado (Foolad, 2007).

Tras la publicación de la primera versión del genoma completo de tomate en el año 2012 por ‘The Tomato Genome Consortium’ (Sato et al., 2012) y el desarrollo de plataformas de genotipado como SolCAP (Sim et al., 2012a), el desarrollo de mapas ultradensos en tomate ha experimentado un rápido desarrollo, lo que ha permitido un aumento del poder y la precisión en la detección de genes y QTLs (Rothan et al., 2019). Este rápido desarrollo queda reflejado en la evolución desde el primer mapa genético de alta densidad, elaborado con 1.030 marcadores de diferente naturaleza (Tanksley et al. 1992), hasta los mapas actuales con miles de marcadores (Sim et al., 2012a; Wen et al., 2019). Varios de estos mapas genéticos de tomate están disponibles en la página web Solgenomics (<https://solgenomics.net/>).

5.2. Análisis de asociación genómica o GWAS

El objetivo de los estudios de asociación genómica es identificar aquellos polimorfismos de ADN cuyas frecuencias alélicas varían sistemáticamente en función de los valores fenotípicos para el carácter. Se basan en el empleo de muchos individuos no relacionados, aprovechando los eventos naturales de recombinación que han tenido lugar a lo largo del tiempo, y aprovechando así la diversidad genética natural que hay dentro de las poblaciones (Rothan et al., 2019).

Por lo tanto, para llevar a cabo un análisis GWAS (Klein et al., 2005) se requiere disponer de una alta cantidad de marcadores distribuidos a lo largo de todo el genoma y del análisis fenotípico de individuos no relacionados. Al disponer de una suficiente densidad de marcadores genéticos, se consigue capturar una gran proporción de las variantes genéticas presentes en la población de estudio (Mammadov et al., 2012). Por ello, la incorporación de especies silvestres al panel es interesante porque hace que la diversidad genética acumulada sea especialmente alta.

La importancia de estos estudios radica también en su posterior aplicabilidad en los programas de mejora. Al disponer de una población de gran cantidad de individuos diversos y cientos de miles de marcadores distribuidos a lo largo del genoma, se pueden identificar aquellas variantes alélicas que subyacen al fenotipo de interés (Huang & Han 2014). Una vez identificadas las variantes genéticas que están asociadas al carácter, se desarrollan poblaciones de mapeo específicas para validar los posibles genes o QTLs candidatos. Una vez validados, las entradas portadoras de estas variantes se pueden emplear en el desarrollo de

variedades comerciales que incorporen el carácter de interés. Además, se puede explotar la diversidad existente en las colecciones de germoplasma y se pueden buscar nuevas variantes alélicas en las mismas (Figura 7).

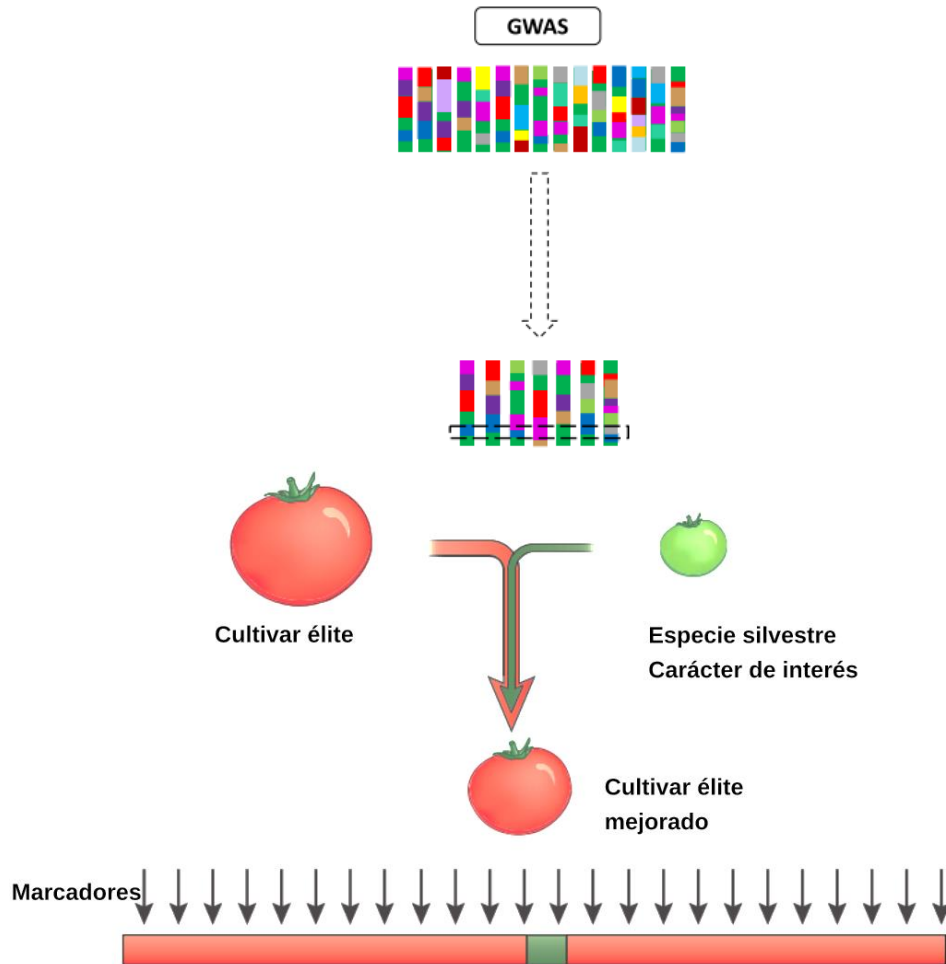


Figura 7. Esquema GWAS para la identificación de genes/QTLs de interés y utilización de estas variantes alélicas para los genes identificados en programas de mejora.

El primer panel estudiado en tomate estaba compuesto por accesiones de tomate cultivado, de *S. lycopersicum* var. *cerasiforme* y de *S. pimpinellifolium* (Sauvage et al., 2014). Este estudio encontró 44 posibles *loci* candidatos relacionados con la composición metabólica del fruto. Desde entonces, este enfoque ha permitido la identificación de diversos genes relacionados con la composición del fruto (Bauchet, Grenier, Samson, Segura, et al., 2017; Ye et al., 2019; Zhu et al., 2018a).

La forma, tamaño o apariencia del fruto también han sido objeto de estudio mediante análisis GWAS. Lin *et al.* (2014) determinaron la variación alélica presente en el cromosoma 1 subyacente al color rosa de los frutos de tomate. Sacco *et al.* (2015) encontraron posibles regiones candidatas del genoma que controlan caracteres morfológicos de forma y tamaño del fruto. Además del peso del fruto, Bauchet *et al.* (2017b) identificaron asociaciones con el desarrollo del fruto y el tiempo de floración.

Razifard *et al.* (2020) en un estudio incluyendo accesiones de las especies *S. pimpinellifolium*, *S. lycopersicum* var. *cerasiforme* y *S. lycopersicum* var. *lycopersicum* determinaron aquellas variantes alélicas responsables de los cambios fenotípicos que tuvieron

lugar durante la domesticación del tomate. Entre estos caracteres, se incluye la cantidad de sólidos solubles, el tamaño y la forma del fruto y el contenido en ácido málico.

Actualmente, el metaanálisis de estudios de asociación genómica se está planteando como una alternativa para mejorar el poder de detección de los estudios de asociación. Estos estudios permiten investigar la consistencia de las asociaciones detectadas en diversos conjuntos de datos y poblaciones de estudios (Zeggini & Ioannidis, 2009). Este enfoque ya ha sido empleado con éxito en tomate, mediante el análisis de los resultados de tres paneles diferentes de estudio que en total incluían 775 accesiones de tomate y más de dos millones de SNPs. Esto ha permitido detectar 305 asociaciones con caracteres relacionados con el sabor del fruto como pueden ser el contenido en azúcares o compuestos volátiles (Zhao et al., 2019).

BIBLIOGRAFÍA

- Alexander, L., & Hoover, M. (1955). Disease resistance in wild species of tomato: Report of the national screening Committee. *Agricultural Experimental Station Research Bulletin*, April, 7.
- Andrade, M. C., da Silva, A. A., Carvalho, R. de C., de Andrade Santiago, J., Souza de Oliveira, A. M., Merrill Francis, D., & Maluf, W. R. (2018). *Quantitative trait loci associated with trichomes in the Solanum galapagense accession LA1401*. 1671–1685. <https://doi.org/10.1007/s10722-018-0644-3>
- Andrade, M. C., da Silva, A. A., Neiva, I. P., Oliveira, I. R. C., De Castro, E. M., Francis, D. M., & Maluf, W. R. (2017). Inheritance of type IV glandular trichome density and its association with whitefly resistance from *Solanum galapagense* accession LA1401. *Euphytica*, 213(2). <https://doi.org/10.1007/s10681-016-1792-1>
- Arellano Rodríguez, L. J., Rodríguez Guzmán, E., Ron Parra, J., Martínez Ramírez, J. L., Lozoya Saldaña, H., Sánchez Martínez, J., & Lépiz Ildefonso, R. (2013). Evaluation of the resistance against *Phytophthora infestans* of wild populations of *Solanum lycopersicum* var. *cerasiforme*. *Revista Mexicana de Ciencias Agrícolas*, 4(5), 753–766. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-09342012000100010
- Bai, Y., & Lindhout, P. (2007). *Domestication and Breeding of Tomatoes: What have We Gained and What Can We Gain in the Future?* 1085–1094. <https://doi.org/10.1093/aob/mcm150>
- Baldwin, E. A., Scott, J. W., Shewmaker, C. K., & Schuch, W. (2000). Flavor Trivia and Tomato Aroma: Biochemistry and Possible Mechanisms for Control of Important Aroma Components. In *1013 HORTSCIENCE* (Vol. 35, Issue 6).
- Banerjee, M. K., & Kalloo, M. K. (1987). Sources and inheritance of resistance to leaf curl virus in *Lycopersicon*. *Theoretical and Applied Genetics*, 73(5), 707–710. <https://doi.org/10.1007/BF00260780>
- Barrantes, W., Fernández-del-Carmen, A., López-Casado, G., González-Sánchez, M. Á., Fernández-Muñoz, R., Granell, A., & Monforte, A. J. (2014). Highly efficient genomics-assisted development of a library of introgression lines of *Solanum pimpinellifolium*. *Molecular Breeding*, 34(4), 1817–1831. <https://doi.org/10.1007/s11032-014-0141-0>
- Bauchet, G., & Causse, M. (2012). Genetic Diversity in Tomato (*Solanum lycopersicum*) and Its Wild Relatives. *Genetic Diversity in Plants*, May. <https://doi.org/10.5772/33073>
- Bauchet, G., Grenier, S., Samson, N., Bonnet, J., Grivet, L., & Causse, M. (2017). Use of modern tomato breeding germplasm for deciphering the genetic control of agronomical traits by Genome Wide Association study. *Theoretical and Applied Genetics*, 130(5), 875–889. <https://doi.org/10.1007/s00122-017-2857-9>
- Bauchet, G., Grenier, S., Samson, N., Segura, V., Kende, A., Beekwilder, J., Cankar, K., Gallois, J. L., Gricourt, J., Bonnet, J., Baxter, C., Grivet, L., & Causse, M. (2017). Identification of major loci and genomic regions controlling acid and volatile content in

- tomato fruit: implications for flavor improvement. *New Phytologist*, 215(2), 624–641. <https://doi.org/10.1111/nph.14615>
- Bergougnoux, V. (2014). The history of tomato: From domestication to biopharming. *Biotechnology Advances*, 32(1), 170–189. <https://doi.org/10.1016/j.biotechadv.2013.11.003>
- Blanca, J., Cañizares, J., Cordero, L., Pascual, L., Díez, M. J., & Nuez, F. (2012). Variation Revealed by SNP Genotyping and Morphology Provides Insight into the Origin of the Tomato. *PLoS ONE*, 7(10). <https://doi.org/10.1371/journal.pone.0048198>
- Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., Francis, D., Causse, M., Knaap, E. Van Der, & Cañizares, J. (2015). Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics*, 257. <https://doi.org/10.1186/s12864-015-1444-1>
- Blauth, S. L., Churchill, G. A., & Mutschler, M. A. (1998). Identification of quantitative trait loci associated with acylsugar accumulation using intraspecific populations of the wild tomato, *Lycopersicon pennellii*. *Theoretical and Applied Genetics*, 96(3–4), 458–467. <https://doi.org/10.1007/s001220050762>
- Bohner, J., & Bangerth, F. (1988). Cell number, cell size and hormone levels in semi-isogenic mutants of *Lycopersicon pimpinellifolium* differing in fruit size. *Physiologia Plantarum*, 72, 316–320.
- Caicedo, A. L., & Schaal, B. A. (2004). Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. *Molecular Ecology*, 13(7), 1871–1882. <https://doi.org/10.1111/j.1365-294X.2004.02191.x>
- Capel, C., del Carmen, A.F., Alba, J. M., Lima-Silva, V., Hernández-Gras, F., Salinas, M., Boronat, A., Angosto, T., Botella, M. A., Fernández-Muñoz, R., Granell, A., Capel, J., & Lozano, R. (2015). Wide-genome QTL mapping of fruit quality traits in a tomato RIL population derived from the wild-relative species *Solanum pimpinellifolium* L. *Theoretical and Applied Genetics*, 128(10), 2019–2035. <https://doi.org/10.1007/s00122-015-2563-4>
- Carter, C. D., & Snyder, J. C. (1985). Mite responses in relation to trichomes of *Lycopersicon esculentum* x *L. hirsutum* F2 hybrids. *Euphytica*, 34(1), 177–185. <https://doi.org/10.1007/BF00022877>
- Causse, M., Buret, M., Robini, K., & Verschave, P. (2003). Inheritance of Nutritional and Sensory Quality Traits in Fresh Market Tomato and Relation to Consumer Preferences. *Journal of Food Science*, 68(7), 2342–2350.
- Channarayappa, C., Shivashankar, G., Muniyappa, V., & Frist, R. H. (1992). Resistance of *Lycopersicon* species to *Bemisia tabaci*, a tomato leaf curl virus vector. *Can. J. Bot.*, 70(Stoner 1970), 2184–2192. <https://doi.org/10.1139/b92-270>
- Chen, F. Q., Foolad, M. R., Hyman, J., St Clair, D. A., & Beelaman, R. B. (1999). Mapping of QTLs for lycopene and other fruit traits in a *Lycopersicon esculentum* x *L. pimpinellifolium* cross and comparison of QTLs across tomato species. In *Molecular*

Breeding (Vol. 5).

- Chetelat, R. T., DeVerna -A B Bennett, J. W., Bennett, A. B., & DeVerna, J. W. (1995). Effects of the *Lycopersicon chmielewskii* sucrose accumulator gene (sucr) on fruit yield and quality parameters following introgression into tomato. In *Theor Appl Genet* (Vol. 91). Springer-Verlag.
- Cuartero, J., Gomez-Guillamon, M. L., & Diaz, A. (1985). Catalog of collections of *Lycopersicon* from Peruvian Central Areas. *TGC Report*, 35, 32–35.
- De Jong, J., & Honma, S. (1976). EVALUATION OF SCREENING TECHNIQUES AND DETERMINATION OF CRITERIA FOR ASSESSING RESISTANCE TO *CORYNEBACTERIUM MICHIGANENSE* IN TOMATO 1. In *Euphytica* (Vol. 25).
- De Candolle, A. (1886). Plants Cultivated for Their Seeds. In *Origin of Cultivated Plants: Vol. XLIX* (2nd ed., pp. 376–384). Kegan Paul Trench & Co.
- Dorst, J. C. E. A. (1946). *Een en twintigste beschrijvende rassenlijst voor landbouwgewassen*.
- Dunal, M. F. (1813). *Histoire naturelle, medicinale et économique des Solanum*. Paris: A. Koenig.
- Fernández-Muñoz, R., Salinas, M., Álvarez, M., & Cuartero, J. (2003). *Inheritance of Resistance to Two-spotted Spider Mite and Glandular Leaf Trichomes in Wild Tomato Lycopersicon pimpinellifolium (Jusl.) Mill.* 128(2), 188–195.
- 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. (2013). Identification and QTL mapping of whitefly resistance components in *Solanum galapagense*. *Theoretical and Applied Genetics*, 126(6), 1487–1501. <https://doi.org/10.1007/s00122-013-2067-z>
- Firdaus, S., van Heusden, A. W., Hidayati, N., Supena, E. D. J., Visser, R. G. F., & Vosman, B. (2012). Resistance to *Bemisia tabaci* in tomato wild relatives. *Euphytica*, 187(1), 31–45. <https://doi.org/10.1007/s10681-012-0704-2>
- Foolad, M. R. (2007). Genome mapping and molecular breeding of tomato. *International Journal of Plant Genomics*, 2007. <https://doi.org/10.1155/2007/64358>
- Foolad, M. R., Merk, H. L., & Ashrafi, H. (2008). Genetics, genomics and breeding of late blight and early blight resistance in tomato. *Critical Reviews in Plant Sciences*, 27(2), 75–107. <https://doi.org/10.1080/07352680802147353>
- Frery, A., Nesbitt, T. C., Frery, A., Grandillo, S., Cong, B., Liu, J., Meller, J., Elber, R., Alpert, K. B., & Tanksley, S. D. (2000). *fw2.2: A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size.* 289(July), 85–88. <https://doi.org/10.1126/science.289.5476.85>
- Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., Burzynski-chang, E. A., Fish, T. L., Stromberg, K. A., Sacks, G. L., Thannhauser, T. W., Foolad, M. R., Diez, M. J., Blanca, J., Canizares, J., Xu, Y., Knaap, E. van der, Huang, S., Klee, H. J., ... Fei, Z. (2019). The tomato pan-genome uncovers new genes and a Glas, J. J., Schimmel, B. C. J., Alba, J. M., Escobar-Bravo, R., Schuurink, R. C., & Kant, M. R. (2012). Plant

- glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences*, 13(12), 17077–17103. <https://doi.org/10.3390/ijms131217077>
- Goffreda, J. C., & Mutschler, M. A. (1989). Inheritance of potato aphid resistance in hybrids between *Lycopersicon esculentum* and *L. pennellii*. *Theoretical and Applied Genetics*, 78, 210–216.
- Gong, H., Rehman, F., Yang, T., Li, Z., Zeng, S., Pan, L., Li, Y., & Wang, Y. (2019). Construction of the first high-density genetic map and QTL mapping for photosynthetic traits in *Lycium barbarum* L. *Molecular Breeding*, 39(7). <https://doi.org/10.1007/s11032-019-1000-9>
- Grandillo, S., Ku, H. M., & Tanksley, S. D. (1999). Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theoretical and Applied Genetics*, 99(6), 978–987. <https://doi.org/10.1007/s001220051405>
- Grandillo, S., Termolino, P., & van der Knaap, E. (2013). Molecular mapping of Complex Traits in Tomato. In *Genetics, Genomic and Breeding of Tomato* (pp. 150–228). CRC Press.
- Grandillo, S., and Steven D Tanksley. 1996. “QTL Analysis of Horticultural Traits Differentiating the Cultivated Tomato from the Closely Related Species *Lycopersicon Pimpinellifolium*.” *TAG Theoretical and Applied Genetics* 92: 935–51.
- Grandillo, S., & Cammareri, M. (2016). Molecular Mapping of Quantitative Trait Loci in Tomato. In Causse M., Giovannoni J., Bouzayen M., & Zouine M. (Eds.), *The Tomato Genome. Compendium of Plant Genomes*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-53389-5_4
- Guo, Z., Weston, P. A., & Snyder, J. C. (1993). Repellency to two-spotted spider mite, *Tetranychus urticae* Koch, as related to leaf surface chemistry of *Lycopersicon hirsutum* accessions. *Journal of Chemical Ecology*, 19(12), 2965–2979. <https://doi.org/10.1007/BF00980596>
- Gurr, G. M., & McGrath, D. (2001). Effect of plant variety, plant age and photoperiod on glandular pubescence and host-plant resistance to potato moth (*Phthorimaea operculella*) in *Lycopersicon* spp. *Annals of Applied Biology*, 138(2), 221–230. <https://doi.org/10.1111/j.1744-7348.2001.tb00106.x>
- Hanson, A. A., Lorenz, A. J., Hesler, L. S., Bhusal, S. J., Bansal, R., Michel, A. P., Jiang, G., & Koch, R. L. (2018). Genome-Wide Association Mapping of Host-Plant Resistance to Soybean Aphid. *The Plant Genome*, 11(3), 180011. <https://doi.org/10.3835/plantgenome2018.02.0011>
- Hardigan, M. A., Laimbeer, F. P. E., Newton, L., Crisovan, E., Hamilton, J. P., Vaillancourt, B., Wiegert-Rininger, K., Wood, J. C., Douches, D. S., Farré, E. M., Veilleux, R. E., & Buell, C. R. (2017). Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proceedings of the National Academy of Sciences of the United States of America*, 114(46), E9999–E10008. <https://doi.org/10.1073/pnas.1714380114>

- Huang, X., & Han, B. (2014). Natural variations and genome-wide association studies in crop plants. In *Annual Review of Plant Biology* (Vol. 65, pp. 531–551). Annual Reviews Inc. <https://doi.org/10.1146/annurev-arplant-050213-035715>
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z., Li, M., Fan, D., Guo, Y., Wang, A., Wang, L., Deng, L., Li, W., Lu, Y., Weng, Q., Han, B. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics*, 42(11), 961–967. <https://doi.org/10.1038/ng.695>
- Hufford, M. B., Berny Mier Y Teran, J. C., & Gepts, P. (2019). *Annual Review of Plant Biology Crop Biodiversity: An Unfinished Magnum Opus of Nature*. <https://doi.org/10.1146/annurev-arplant-042817>
- Hutton, S. F., Scott, J. W., Yang, W., Sim, S. C., Francis, D. M., & Jones, J. B. (2010). Identification of QTL associated with resistance to bacterial spot race T4 in tomato. *Theoretical and Applied Genetics*, 121(7), 1275–1287. <https://doi.org/10.1007/s00122-010-1387-5>
- Jarret, R. L., Sayama, H., & Tigchelaar, E. C. (1984). Pleiotropic effects associated with the chlorophyll intensifier mutations high pigment and dark green in tomatoes. *Journal of the American Society of Horticultural Science*, 109, 873–878.
- Jenkins, J. A. (1948). The Origin of the Cultivated Tomato. *Economic Botany*, 379–392.
- Jenkins, J. A., & Mackinney, G. (1955). CAROTENOIDS OF THE APRICOT TOMATO AND ITS HYBRIDS WITH YELLOW AND TANGERINE. *Genetics*, 40(5), 715–720.
- Joosten, M. H. A. J., & de Wit, P. J. G. M. (1999). The tomato-Cladosporium fulvum interaction: A versatile experimental system to study plant-pathogen interactions. *Annual Review of Phytopathology*, 37(1). <https://doi.org/10.1146/annurev.phyto.37.1.335>
- Kachanovsky, D. E., Filler, S., Isaacson, T., & Hirschberg, J. (2012). Epistasis in tomato color mutations involves regulation of phytoene synthase 1 expression by cis-carotenoids. *Proceedings of the National Academy of Sciences of the United States of America*, 109(46), 19021–19026. <https://doi.org/10.1073/pnas.1214808109>
- Kaloo, G. (Ed.). (1991). *Genetic Improvement of Tomato*. Springer-Verlag Berlin Heidelberg. <https://doi.org/10.1111/j.1526-100X.2009.00595.x>
- Klein, R. J., Zeiss, C., Chew, E. Y., Tsai, J. Y., Sackler, R. S., Haynes, C., Henning, A. K., SanGiovanni, J. P., Mane, S. M., Mayne, S. T., Bracken, M. B., Ferris, F. L., Ott, J., Barnstable, C., & Hoh, J. (2005). Complement factor H polymorphism in age-related macular degeneration. *Science*, 308(5720), 385–389. <https://doi.org/10.1126/science.1109557>
- Kopeliovitch, E., Mizrahi, Y., Rabinowitch, H. D., & Kedar, N. (1980). Physiology of the tomato mutant alcobaca. *Physiologia Plantarum*, 48(2), 307–311. <https://doi.org/10.1111/j.1399-3054.1980.tb03260.x>
- Kraus, C. M., Munkvold, K. R., & Martin, G. B. (2016). Natural Variation in Tomato Reveals Differences in the Recognition of AvrPto and AvrPtoB Effectors from Pseudomonas

- syringae. *Molecular Plant*, 9(5), 639–649. <https://doi.org/10.1016/j.molp.2016.03.001>
- Labate, J. A., Grandillo, S., Fulton, T., Muños, S., Caicedo, A. L., Peralta, I., Yuanfu, J., Chetelat, R. T., J. W. Scott, Gonzalo, M. J., Francis, D., Wencai, Y., van der Knaap, E., Baldo, A. M., Smith-White, B., Mueller, L. A., Prince, J. P., Blanchard, N. E., Storey, D. B., ... Causse, M. (2007). Tomato. In C. Kole (Ed.), *GenomeMapping andMolecular Breeding in Plants - Vegetables* (pp. 1–95). Springer Berlin Heidelberg New York.
- Lanahan, M. B., Yen Hsiao Ching, Giovannoni, J. J., & Klee, H. J. (1994). The Never ripe mutation blocks ethylene perception in tomato. *Plant Cell*, 6(4), 521–530. <https://doi.org/10.1105/tpc.6.4.521>
- Lawson, D. M., Lunde, C. F., & Mutschler, M. A. (1997). Marker-assisted transfer of acylsugar-mediated pest resistance from the wild tomato, *Lycopersicon pennellii*, to the cultivated tomato, *Lycopersicon esculentum*. In *Molecular Breeding* (Vol. 3). Kluwer Academic Publishers.
- Leckie, B. M., de Jong, D. M., & Mutschler, M. A. (2012). Quantitative trait loci increasing acylsugars in tomato breeding lines and their impacts on silverleaf whiteflies. *Molecular Breeding*, 30(4), 1621–1634. <https://doi.org/10.1007/s11032-012-9746-3>
- Lemke, C. A., & Mutschler, M. (1984). Inheritance of glandular trichomes in crosses between *Lycopersicon esculentum* and *Lycopersicon pennellii*. *Journal of the American Society for Horticultural Science*, 109, 592–596.
- Lester, R. N. (1991). Evolutionary relationships of tomato, potato and pepino, and wild species of *Lycopersicon* and *Solanum*. In R. N. Hawkes, M. N. Lester, & N. Estrada (Eds.), *Solanaceae III: taxonomy, chemistry, evolution* (pp. 283–301). Royal Botanic Gardens.
- Levin, I., Gilboa, N., Yeselson, E., Shen, S., & Schaffer, A. A. (2000). Fgr, a major locus that modulates the fructose to glucose ratio in mature tomato fruits. *Theoretical and Applied Genetics Volume*, 100, 256–262.
- Levin, Ilan, de Vos, C., Tadmor, Y., Bovy, A., Lieberman, M., Oren-Shamir, M., Segev, O., Kolotilin, I., Keller, M., Ovadia, R., Meir, A., & Bino, R. (2006). High pigment tomato mutants—more than just lycopene (a review). *Israel Journal of Plant Sciences*, 54(3). https://doi.org/10.1560/IJPS_54_3_179
- Li, S., Yang, G., Yang, S., Just, J., Yan, H., Zhou, N., Jian, H., Wang, Q., Chen, M., Qiu, X., Zhang, H., Dong, X., Jiang, X., Sun, Y., Zhong, M., Bendahmane, M., Ning, G., Ge, H., Hu, J. Y., & Tang, K. (2019). The development of a high-density genetic map significantly improves the quality of reference genome assemblies for rose. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-019-42428-y>
- Lieberman, M., Segev, O., Gilboa, N., Lalazar, A., & Levin, I. (2004). The tomato homolog of the gene encoding UV-damaged DNA binding protein 1 (DDB1) underlined as the gene that causes the high pigment-1 mutant phenotype. *Theoretical and Applied Genetics*, 108(8), 1574–1581. <https://doi.org/10.1007/s00122-004-1584-1>
- Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., Zhang, Z., Lun, Y., Li, S., Wang, X.,

- Huang, Z., Li, J., Zhang, C., Wang, T., Zhang, Y., Wang, A., Zhang, Y., Lin, K., Li, C., ... Huang, S. (2014). Genomic analyses provide insights into the history of tomato breeding. In *Nature Genetics* (Vol. 46, Issue 11, pp. 1220–1226). <https://doi.org/10.1038/ng.3117>
- Linneo, C. (1753). *Species plantarum*. Stockholm: L. Salvius.
- Lippman, Z. B., Cohen, O., Alvarez, J. P., Abu-Abied, M., Pekker, I., Paran, I., Eshed, Y., & Zamir, D. (2008). The making of a compound inflorescence in tomato and related nightshades. *PLoS Biology*, 6(11), 2424–2435. <https://doi.org/10.1371/journal.pbio.0060288>
- Liu, T. J., Zhang, Y. J., Agerbirk, N., Wang, H. P., Wei, X. C., Song, J. P., He, H. J., Zhao, X. Z., Zhang, X. H., & Li, X. X. (2019). A high-density genetic map and QTL mapping of leaf traits and glucosinolates in *Barbarea vulgaris*. *BMC Genomics*, 20(1). <https://doi.org/10.1186/s12864-019-5769-z>
- Luckwill, L. C. (1943). The Genus *Lycopersicon*: An Historical, Biological and Taxonomic Survey of the Wild and Cultivated Tomatoes. In U. Press (Ed.), *Aberdeen University Studies*.
- Maliepaard, C., Bas, N., Van Heusden, S., Kos, J., Pet, G., Verkerk, R., Vrielink, R., Zabel, P., & Lindhout, P. (1995). Mapping of QTLs for glandular trichome densities and trialeurodes vaporariorum (Greenhouse whitefly) resistance in an F2 from *lycopersicon esculentum* × *lycopersicon hirsutum* f. *glabratum*. *Heredity*, 75(4), 425–433. <https://doi.org/10.1038/hdy.1995.155>
- Mammadov, J., Aggarwal, R., Buyyarapu, R., & Kumpatla, S. (2012). SNP markers and their impact on plant breeding. *International Journal of Plant Genomics*, 2012. <https://doi.org/10.1155/2012/728398>
- Mazidi, M., Ferns, G. A., & Banach, M. (2020). A high consumption of tomato and lycopene is associated with a lower risk of cancer mortality: results from a multi-ethnic cohort. *Public Health Nutrition*. <https://doi.org/10.1017/S1368980019003227>
- McDowell, E. T., Kapteyn, J., Schmidt, A., Li, C., Kang, J. H., Descour, A., Shi, F., Larson, M., Schillmiller, A., An, L., Daniel Jones, A., Pichersky, E., Soderlund, C. A., & Gang, D. R. (2011). Comparative Functional Genomic Analysis of *Solanum* Glandular Trichome Types. *Plant Physiology*, 155(1), 524–539. <https://doi.org/10.1104/pp.110.167114>
- Meyer, R. S., & Purugganan, M. D. (2013). Evolution of crop species: Genetics of domestication and diversification. In *Nature Reviews Genetics* (Vol. 14, Issue 12). <https://doi.org/10.1038/nrg3605>
- Mieslerová, B., Lebeda, A., & Chetelat, T. (2000). Variation in Response of Wild *Lycopersicon* and *Solanum* spp. against Tomato Powdery Mildew (*Oidium lycopersici*) *Journal of Phytopathology*, 148, 303–311.
- Miller, P. (1754). The gardeners dictionary abridged. In *London: Printed for the author*.

- Monforte, A. J., Diaz, A., Caño-Delgado, A., & van der Knaap, E. (2014). The genetic basis of fruit morphology in horticultural crops: Lessons from tomato and melon. In *Journal of Experimental Botany* (Vol. 65, Issue 16, pp. 4625–4637). <https://doi.org/10.1093/jxb/eru017>
- Mu, Q., Huang, Z., Chakrabarti, M., Illa-Berenguer, E., Liu, X., Wang, Y., Ramos, A., & van der Knaap, E. (2017). Fruit weight is controlled by Cell Size Regulator encoding a novel protein that is expressed in maturing tomato fruits. *PLoS Genetics*, *13*(8), 1–26. <https://doi.org/10.1371/journal.pgen.1006930>
- Müller, C. H. (1940). *A revision of the genus Lycopersicon*. Washington, D.C.: U.S. Dept. of Agriculture.
- Mutschler, M. A., Doerge, R. W., Liu, S. C., Kuai, J. P., Liedl, B. E., & Shapiro, J. A. (1996). QTL analysis of pest resistance in the wild tomato *Lycopersicon pennellii*: QTLs controlling acylsugar level and composition. *Theoretical and Applied Genetics*, *92*(6), 709–718. <https://doi.org/10.1007/BF00226093>
- Mutschler, M.A., Wolfe, D. W., Cobb, E. D., & Yourstone, K. S. (1992). Tomato Fruit Quality and Shelf Life in Hybrids Heterozygous for the alc Ripening Mutant. *HortScience*, *27*, 1985–1988.
- Nesbitt, T. C., & Tanksley, S. D. (2002). Comparative Sequencing in the Genus *Lycopersicon*: Implications for the Evolution of Fruit Size in the Domestication of Cultivated Tomatoes. *Genetics Society of America*, *162*, 365–379.
- Nuez, F., & Díez, M. J. (2008). Tomato. In J. Prohens & F. Nuez (Eds.), *Vegetables II. Handbook of Plant Breeding* (pp. 249–323). Springer. https://doi.org/10.1007/978-0-387-74110-9_7
- Oliveira, J.R.F., de Resende J.T.V, Maluf, W.R., Lucini, T., de Lima Filho, R.B., de Lima, I.P., and Nardi, C. (2018). Trichomes and Allelochemicals in Tomato Genotypes Have Antagonistic Effects upon Behavior and Biology of *Tetranychus Urticae*. *Frontiers in Plant Science* (9 (August): 1–9). <https://doi.org/10.3389/fpls.2018.01132>.
- Olsen, K. M., & Wendel, J. F. (2013). A bountiful harvest: Genomic insights into crop domestication phenotypes. In *Annual Review of Plant Biology* (Vol. 64, pp. 47–70). <https://doi.org/10.1146/annurev-arplant-050312-120048>
- Organización de las Naciones Unidas para la Agricultura y la Alimentación. (2008). *El estado de la inseguridad alimentaria en el mundo 2008: los precios elevados de los alimentos y la seguridad alimentaria; amenazas y oportunidades*. 56.
- Oriani, M. A., Vendramim, J. D., Oriani, M. A. de G., & Vendramim, J. D. (2010). Influence of trichomes on attractiveness and ovipositional preference of *Bemisia tabaci* (Genn.) B biotype (Hemiptera: Aleyrodidae) on tomato genotypes. *Neotropical Entomology*, *39*(6), 1002–1007. <https://doi.org/10.1590/S0103-90162011000100006>
- Papa, R., Bellucci, E., Rossi, M., Leonardi, S., Rau, D., Gepts, P., Nanni, L., & Attene, G. (2007). Tagging the signatures of domestication in common bean (*Phaseolus vulgaris*) by means of pooled DNA samples. *Annals of Botany*, *100*(5), 1039–1051.

<https://doi.org/10.1093/aob/mcm151>

- Paran, I., & van der Knaap, E. (2007). Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *Journal of Experimental Botany*, 58(14), 3841–3852. <https://doi.org/10.1093/jxb/erm257>
- Peralta, I. E., & Spooner, D. M. (2006). History, origin and early cultivation of tomato (Solanaceae). In M. K. Razdan & A. K. Mattoo (Eds.), *Genetic improvement of Solanaceous crops, vol. 2: Tomato*. Enfield: Science Publishers.
- Peralta, I. E., Spooner, D. M., & Knapp, S. (2008). Taxonomy of Wild Tomatoes and Their Relatives (Solanum sect. Lycopersicoides, sect. Juglandifolia, sect. Lycopersicon; Solanaceae). *Systematic Botany Monographs*, 84, 1–186.
- Picó, B., Sifres, A., Mónica, E., Díez, M. J., & Nuez, F. (2000). Searching for new resistance sources to tomato yellow leaf curl virus within a highly variable wild Lycopersicon genetic pool. *BIOTIC STRESSES*, 344–350.
- Purugganan, M. D. (2019). Evolutionary Insights into the Nature of Plant Domestication. In *Current Biology* (Vol. 29, Issue 14, pp. R705–R714). Cell Press. <https://doi.org/10.1016/j.cub.2019.05.053>
- Rakha, M., Bouba, N., Ramasamy, S., Regnard, J. L., & Hanson, P. (2016). Evaluation of wild tomato accessions (Solanum spp.) for resistance to two-spotted spider mite (Tetranychus urticae Koch) based on trichome type and acylsugar content. *Genetic Resources and Crop Evolution*, 64(5), 1011–1022. <https://doi.org/10.1007/s10722-016-0421-0>
- Rakha, M., Hanson, P., & Ramasamy, S. (2017). 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*, 64(2), 247–260. <https://doi.org/10.1007/s10722-015-0347-y>
- Rambla, J. L., Medina, A., Fernández-Del-Carmen, A., Barrantes, W., Grandillo, S., Cammareri, M., López-Casado, G., Rodrigo, G., Alonso, A., García-Martínez, S., Primo, J., Ruiz, J. J., Fernández-Muñoz, R., Monforte, A. J., & Granell, A. (2017). Identification, introgression, and validation of fruit volatile QTLs from a red-fruited wild tomato species. *Journal of Experimental Botany*, 68(3), 429–442. <https://doi.org/10.1093/jxb/erw455>
- Ranc, N., Mũos, S., Santoni, S., & Causse, M. (2008). A clarified position for solanum lycopersicum var. cerasiforme in the evolutionary history of tomatoes (solanaceae). *BMC Plant Biology*, 8, 1–18. <https://doi.org/10.1186/1471-2229-8-130>
- Ranjan, A., Ichihashi, Y., & Sinha, N. R. (2012). *The tomato genome: implications for plant breeding, genomics and evolution. Box 1*, 1–8.
- Razifard, H., Ramos, A., Della Valle, A. L., Bodary, C., Goetz, E., Manser, E. J., Li, X., Zhang, L., Visa, S., Tieman, D., van der Knaap, E., & Caicedo, A. L. (2020). Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution*, 1–15.

<https://doi.org/10.1093/molbev/msz297>

- Rick, C. M. (1960). HYBRIDIZATION BETWEEN LYCOPERSICON ESCULENTUM AND SOLANUM PENNELLII: PHYLOGENETIC AND CYTOGENETIC SIGNIFICANCE*. In *Univ. of Texas Publ* (Vol. 34, Issue 5721).
- Rick, C. M. (1979). Biosystematic studies in Lycopersicon and closely related species of Solanum. In L. S. symposium Series (Ed.), *The biology and taxonomy of Solanaceae* (pp. 667–677). Academic Press. London.
- Rick, C. M., & Holle, M. (1990). Andean Lycopersicon esculentum var. cerasiforme: genetic variation and its evolutionary significance. *Economic Botany*, 44(3 Supplement), 69–78. <https://doi.org/10.1007/BF02860476>
- Rick, C. M., & Yoder, J. I. (1988). Classical And Molecular Genetics Of Tomato: Highlights And Perspectives. *Annual Review of Genetics*, 22(1), 281–300. <https://doi.org/10.1146/annurev.genet.22.1.281>
- Rick, C.M. (1991). Tomato Paste: A Concentrated Review of Genetic Highlights From the Beginnings to the Advent of Molecular Genetics. *Genetics Society of America*, 128, 1–5
- Rick, C. M. (1984). Evolution of mating systems: evidence from allozyme variation. *Genetics: New Frontiers*, 216–221.
- Rick, C. M., & Butler, L. (1956). *Cytogenetics of the Tomato* (Vol. 8). [https://doi.org/10.1016/S0065-2660\(08\)60504-0](https://doi.org/10.1016/S0065-2660(08)60504-0)
- Rick, C. M., Holle, M., & Thorp, R. W. (1978). Rates of cross-pollination in Lycopersicon pimpinellifolium: Impact of genetic variation in floral characters. *Plant Systematics and Evolution*, 129(1–2), 31–44. <https://doi.org/10.1007/BF00988982>
- Rick, C. M. (1978). The Tomato. *Scientific American*.
- Rick, C. M., & Fobes, J. F. (1975). Allozyme Variation in the Cultivated Tomato and Closely Related Species. *Bulletin of the Torrey Botanical Club*, 102(6), 376–384.
- Rick, C. M., Fobes, J. F., & Holle, M. (1977). Genetic Variation in Lycopersicon pimpinellifolium: Evidence of Evolutionary Change in Mating Systems *. *Plant Systematics and Evolution*, 127, 139–170.
- Rivas, S., & Thomas, C. M. (2005). Molecular Interactions Between Tomato and the Leaf Mold Pathogen Cladosporium fulvum . *Annual Review of Phytopathology*, 43(1), 395–436. <https://doi.org/10.1146/annurev.phyto.43.040204.140224>
- Rodríguez, G. R., Muños, S., Anderson, C., Sim, S.-C., Michel, A., Causse, M., Gardener, B. B. M., Francis, D., & van der Knaap, E. (2011). Distribution of SUN, OVATE, LC, and FAS in the Tomato Germplasm and the Relationship to Fruit Shape Diversity. *Plant Physiology*, 156(1), 275–285. <https://doi.org/10.1104/pp.110.167577>
- Ronen, G., Carmel-Goren, L., Zamir, D., & Hirschberg, J. (2000). An alternative pathway to β -carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. *Proceedings of the National Academy of*

Sciences of the United States of America, 97(20), 11102–11107.
<https://doi.org/10.1073/pnas.190177497>

- Rothan, C., Diouf, I., & Causse, M. (2019). Trait discovery and editing in tomato. *Plant Journal*, 97(1), 73–90. <https://doi.org/10.1111/tpj.14152>
- Sacco, A., Ruggieri, V., Parisi, M., Festa, G., Rigano, M., Picarella, M. E., Mazzucato, A., & Barone, A. (2015). Exploring a Tomato Landraces Collection for Fruit-Related Traits by the Aid of a High- Throughput Genomic Platform. *PLoS ONE*, 10. <https://doi.org/10.1371/journal.pone.0137139>
- Saeidi, Z., Mallik, B., & Kulkarni, R. S. (2007). Inheritance of glandular trichomes and two-spotted spider mite resistance in cross *Lycopersicon esculentum* “Nandi” and *L. pennellii* “LA2963.” *Euphytica*, 154(1–2), 231–238. <https://doi.org/10.1007/s10681-006-9291-4>
- Saliba-Colombani, V., Causse, M., Langlois, D., Philouze, J., & Buret, M. (2001). Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theoretical and Applied Genetics*, 102, 259–272.
- Salinas, M., Capel, C., Alba, J. M., Mora, B., Cuartero, J., Fernández-Muñoz, R., Lozano, R., & Capel, J. (2013). Genetic mapping of two QTL from the wild tomato *Solanum pimpinellifolium* L. controlling resistance against two-spotted spider mite (*Tetranychus urticae* Koch). *Theoretical and Applied Genetics*, 126(1), 83–92. <https://doi.org/10.1007/s00122-012-1961-0>
- Sato, S., Tabata, S., Hirakawa, H., Asamizu, E., Shirasawa, K., Isobe, S., Kaneko, T., Nakamura, Y., Shibata, D., Aoki, K., Egholm, M., Knight, J., Bogden, R., Li, C., Shuang, Y., Xu, X., Pan, S., Cheng, S., Liu, X., ... Gianese, G. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 485(7400), 635–641. <https://doi.org/10.1038/nature11119>
- Sauvage, C., Segura, V., Bauchet, G., Stevens, R., Do, P. T., Nikoloski, Z., Fernie, A. R., Causse, M., Molekulare, P., & Germany, P. T. D. (2014). *Genome-Wide Association in Tomato Reveals 44 Candidate Loci for Fruit Metabolic Traits I [W]*. 165(July), 1120–1132. <https://doi.org/10.1104/pp.114.241521>
- Scott, J. W. (2007). Breeding for resistance to viral pathogens. In M. K. Razdan & A. K. Mattoo (Eds.), *Genetic Improvement of Solanaceous Crops: Vol. Vol 2: Tomato*. (pp. 447–474). Enfield, New Hampshire, USA: Science Publishers, US.
- Sifres, A., Picó, B., Blanca, J. M., De Frutos, R., & Nuez, F. (2006). Genetic structure of *Lycopersicon pimpinellifolium* (Solanaceae) populations collected after the ENSO event of 1997-1998. *Genetic Resources and Crop Evolution*, 54(2), 359–377. <https://doi.org/10.1007/s10722-005-5725-4>
- Sim, S. C., Durstewitz, G., Plieske, J., Wieseke, R., Ganal, M. W., van Deynze, A., Hamilton, J. P., Buell, C. R., Causse, M., Wijeratne, S., & Francis, D. M. (2012a). Development of a large snp genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE*, 7(7). <https://doi.org/10.1371/journal.pone.0040563>
- Sim, S. C., Durstewitz, G., Plieske, J., Wieseke, R., Ganal, M. W., van Deynze, A., Hamilton,

- J. P., Buell, C. R., Causse, M., Wijeratne, S., & Francis, D. M. (2012b). Development of a large snp genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE*, 7(7). <https://doi.org/10.1371/journal.pone.0040563>
- Sims, W. L. (1980). History of Tomato Production for Industry Around the World. In *Acta Horticulturae* (Issue 100, pp. 25–26). <https://doi.org/10.17660/actahortic.1980.100.1>
- Soumpourou, E., Iakovidis, M., Chartrain, L., Lyall, V., & Thomas, C. M. (2007). The *Solanum pimpinellifolium* Cf-ECP1 and Cf-ECP4 genes for resistance to *Cladosporium fulvum* are located at the Milky Way locus on the short arm of chromosome 1. *Theoretical and Applied Genetics*, 115(8), 1127–1136. <https://doi.org/10.1007/s00122-007-0638-6>
- Stevens, M. A., & Rick, C. M. (1986). Genetics and Breeding. In J. Atherton & J. Rudich (Eds.), *The Tomato Crop: a scientific basis for improvement* (p. 661). Chapman and Hall Ltd. <https://doi.org/10.1007/978-94-009-3137-4>
- Stevens, R., Buret, M., Duffé, P., Garchery, C., Baldet, P., Rothan, C., & Causse, M. (2007). Candidate genes and quantitative trait loci affecting fruit ascorbic acid content in three tomato populations. *Plant Physiology*, 143(4), 1943–1953. <https://doi.org/10.1104/pp.106.091413>
- Stout, M. J., Kurabchew, H., & Leite, G. L. D. (2017). Host-Plant Resistance in Tomato. In *Sustainable Management of Arthropod Pests of Tomato*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-802441-6.00009-7>
- Studer, A., Zhao, Q., Ross-Ibarra, J., & Doebley, J. (2011). Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nature Genetics*, 43(11), 1160–1163. <https://doi.org/10.1038/ng.942>
- Tanksley, S. D., Ganai, M. W., Prince, J. P., de Vicente, M. C., Bonierbale, M. W., Broun, P., Fulton, T. M., Giovannoni, J. J., Grandillo, S., Martin, G. B., Messeguer, R., Miller, J. C., Miller, L., Paterson, A. H., Riider, M. S., Wing, R. A., Wu, W., & Young, N. D. (1992). High Density Molecular Linkage Maps of the Tomato and Potato Genomes. *Genetics Society of America*, 132, 1141–1160.
- Tanksley, Steven D. (2004). The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell*, 16, 181–190. <https://doi.org/10.1105/tpc.018119>
- Tanksley, Steven D., & McCouch, S. R. (1997). Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild The Narrow Genetic Base of Crop Plants. *Science*, 277(August), 1063–1066. <https://doi.org/10.1126/science.277.5329.1063>
- Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., Rambla, J. L., Beltran, K. S. O., Taylor, M., Zhang, B., Ikeda, H., Liu, Z., Fisher, J., Zemach, I., Monforte, A., Zamir, D., Granell, A., Kirst, M., Huang, S., & Klee, H. (2017). A chemical genetic roadmap to improved tomato flavor. *Science (New York, N.Y.)*, 355(6323), 391–394. <https://doi.org/10.1126/science.aal1556>
- Tournefort, J. P. (1694). *Éléments de botanique*. Paris I'Imprimerie royale.

- 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. (2017). Loss of type-IV glandular trichomes is a heterochronic trait in tomato and can be reverted by promoting juvenility. *Plant Science*, 259, 35–47. <https://doi.org/10.1016/j.plantsci.2017.03.006>
- 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. (2019). QTL mapping of insect resistance components of *Solanum galapagense*. *Theoretical and Applied Genetics*, 132(2), 531–541. <https://doi.org/10.1007/s00122-018-3239-7>
- Warnock, S. J. (1991). Natural Habitats of *Lycopersicon* Species. *HortScience*, 26(5), 466–471. <https://doi.org/10.21273/hortsci.26.5.466>
- Wen, J., Jiang, F., Weng, Y., Sun, M., Shi, X., Zhou, Y., Yu, L., & Wu, Z. (2019). Identification of heat-tolerance QTLs and high-temperature stress-responsive genes through conventional QTL mapping, QTL-seq and RNA-seq in tomato. *BMC Plant Biology*, 19(1). <https://doi.org/10.1186/s12870-019-2008-3>
- Weston, P. A., & Snyder, J. C. (1990). Thumbtack bioassay: a quick method for measuring plant resistance to twospotted spider mites (Acari: Tetranychidae). *Journal of Economic Entomology*, 83(2), 500–504. <https://doi.org/10.1093/jee/83.2.500>
- Widrechner, M. P. (1987). Variation in Breeding System of *Lycopersicon pimpinellifolium*: Implications for Germplasm Maintenance. *Plant Genetic Resources Newsletter*, 70, 38–43.
- Wilkens, R. T., Shea, G. O., Halbreich, S., & Stamp, N. E. (1996). Resource availability and the trichome defenses of tomato plants. *Oecologia*, 106(2), 181–191. <https://doi.org/10.1007/BF00328597>
- Xiao, Y., Liu, H., Wu, L., Warburton, M., & Yan, J. (2017). Genome-wide Association Studies in Maize: Praise and Stargaze. In *Molecular Plant* (Vol. 10, Issue 3, pp. 359–374). Cell Press. <https://doi.org/10.1016/j.molp.2016.12.008>
- Ye, J., Li, W., Ai, G., Li, C., Liu, G., Chen, W., Wang, B., Wang, W., Lu, Y., Zhang, J., Li, H., Ouyang, B., Zhang, H., Fei, Z., Giovannoni, J. J., Ye, Z., & Zhang, Y. (2019). Genome-wide association analysis identifies a natural variation in basic helix-loop-helix transcription factor regulating ascorbate biosynthesis via D-mannose/L-galactose pathway in tomato. *PLoS Genetics*, 15(5), e1008149. <https://doi.org/10.1371/journal.pgen.1008149>
- Yen, H. C., Shelton, B. A., Howard, L. R., Vrebalov, S. L. J., & Giovannoni, J. J. (1997). The tomato high-pigment (hp) locus maps to chromosome 2 and influences plastome copy number and fruit quality. *Theoretical and Applied Genetics*, 95, 1069–1079.
- Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. *Nature Reviews Genetics*, 2(12), 983–989. <https://doi.org/10.1038/35103590>
- Zeggini, E., & Ioannidis, J. P. A. (2009). Meta-analysis in genome-wide association studies. In *Pharmacogenomics* (Vol. 10, Issue 2, pp. 191–201). <https://doi.org/10.2217/14622416.10.2.191>

- Zhao, J., Sauvage, C., Zhao, J., Bitton, F., Bauchet, G., Liu, D., Huang, S., Tieman, D. M., Klee, H. J., & Causse, M. (2019). Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. *Nature Communications*, *10*(1), 1–12. <https://doi.org/10.1038/s41467-019-09462-w>
- Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., Lin, T., Qin, M., Peng, M., Yang, C., Cao, X., Han, X., Wang, X., van der Knaap, E., Zhang, Z., Cui, X., Klee, H., Fernie, A. R., Luo, J., & Huang, S. (2018). Rewiring of the Fruit Metabolome in Tomato Breeding. *Cell*, *172*(1–2), 249–261. <https://doi.org/10.1016/j.cell.2017.12.019>
- Zuriaga, E., Blanca, J. M., Cordero, L., Sifres, A., Blas-Cerdán, W. G., Morales, R., & Nuez, F. (2009). Genetic and bioclimatic variation in *Solanum pimpinellifolium*. *Genetic Resources and Crop Evolution*, *56*(1), 39–51. <https://doi.org/10.1007/s10722-008-9340-z>

OBJETIVOS

Para asegurar la futura productividad y calidad de los cultivos no es suficiente con el mantenimiento de la variabilidad presente en los mismos, sino que es necesario disponer de fuentes de variabilidad genética y fenotípica para su uso en mejora. En este sentido, las colecciones de germoplasma suponen un recurso fundamental y muy valioso, pero para que puedan resultar útiles y se pueda hacer un uso activo y exitoso de los materiales conservados, es necesaria una buena caracterización a nivel morfológico y genético de los mismos. Además, es necesario que tanto estos materiales, como la información asociada a ellos esté disponible y accesible a la comunidad científica y a los mejoradores. Es por ello por lo que el **objetivo principal** de la presente tesis se centra en la caracterización morfológica y genética de la diversidad presente en una colección de entradas procedentes del centro de domesticación y difusión del tomate cultivado que ha sido seleccionada con el fin de recoger la diversidad presente en esa área y que incluye entradas de *S. pimpinellifolium* (SP), *S. lycopersicum* var. *cerasiforme* (SLC) y *S. lycopersicum* var. *lycopersicum* (SLL), así como la creación de recursos que permitan su uso en mejora. Para poder abordar este trabajo, este objetivo principal se concretó en los siguientes **objetivos parciales**:

1) Caracterizar la diversidad morfológica de una colección de germoplasma de las especies SP, SLC y SLL procedentes del centro de origen y de domesticación del tomate.

2) Generar una colección de SNPs de alta calidad, y anotada en función del efecto en la proteína codificante, que sirva como recurso para el análisis genético de las entradas fenotipadas.

3) Desarrollar una colección pública y accesible de familias segregantes mediante el cruce de todas las entradas de la colección con una entrada representativa de cada una de las especies para facilitar y acelerar el análisis genético de caracteres de interés.

4) Demostrar la utilidad y aplicabilidad de los recursos elaborados en los objetivos anteriores mediante la realización de:

a) un estudio de asociación genómica (GWAS) para la identificación de regiones genómicas implicadas en los caracteres morfológicos evaluados.

b) un estudio del control genético de la presencia y densidad de tricomas tipo IV en una entrada de la especie SP con alta densidad de este tipo de tricomas.

La descripción morfológica de la colección para caracteres de tipo vegetativo, de floración y de fruto se presenta en el Capítulo 2 de esta tesis. La finalidad de este estudio es determinar los caracteres que identifican a cada una de las distintas especies y tener información acerca de la variabilidad morfológica existente en el centro de origen y domesticación del tomate cultivado. Ello facilitará el empleo de esta colección en mejora.

En el Capítulo 3 se muestra el desarrollo de la colección anotada de SNPs para el análisis genético de las entradas, así como el desarrollo de una colección de familias segregantes que pueden ser de utilidad en estudios genéticos. Esta colección de familias se ha desarrollado cruzando toda la colección de entradas original con una entrada representativa de cada especie, SP, SLC y SLL. Por último, se ejemplifica el potencial morfológico y genético de la población para la identificación de regiones genómicas implicadas en el control de algunos caracteres morfológicos evaluados mediante un análisis GWAS.

Finalmente, en el Capítulo 4 se muestra la utilidad de las familias segregantes para llevar a cabo estudios genéticos. La caracterización morfológica de la colección permitió identificar una entrada de SP con alta densidad de tricomas glandulares tipo IV, y estas familias segregantes se han empleado para determinar el control genético de este carácter y estudiar su efecto en dos fondos genéticos distintos.

Capítulo 2:

VARIABILIDAD MORFOLÓGICA DE LA COLECCIÓN

1. Introducción

El tomate cultivado se domesticó a partir de la especie *Solanum lycopersicum* var. *cerasiforme* que, a su vez, fue pre-domesticada a partir de *Solanum pimpinellifolium* (Blanca et al. 2012; Blanca et al. 2015). Durante este proceso de domesticación, sufrió diversos cuellos de botella poblacionales que hicieron que su diversidad genética se viera drásticamente reducida y que ha dado como resultado la baja diversidad de los tomates modernos (Lin et al. 2014; Blanca et al. 2015; Razifard et al. 2020). Esto hace que la mejora genética del tomate cultivado dependa de la variabilidad existente en las especies silvestres y en la variedad *cerasiforme*. Por ello, explorar el centro de origen y la difusión de la especie es un punto crucial para tener información acerca de la diversidad existente en las especies más próximas al cultivo (Bauchet & Causse, 2012). En paralelo a la exploración de la diversidad existente, es importante que ésta sea conservada en los bancos de germoplasma. Sin embargo, la simple recolección y preservación de los recursos genéticos no es suficiente para poder asegurar una futura productividad del cultivo (Zamir, 2001). El establecimiento de las colecciones de germoplasma debe ir acompañado de una buena caracterización de los materiales, tanto a nivel fenotípico como molecular (Tanksley & McCouch 1997). Además, para hacer un uso activo de los recursos de forma eficiente, esta información debe estar disponible para la comunidad científica.

La especie silvestre *S. pimpinellifolium* se encuentra ampliamente distribuida desde el norte de Ecuador al sur de Perú (Warnock, 1991), existiendo diferencias en cuanto al nivel de diversidad genética de las poblaciones en función de su origen geográfico (Rick & Fobes 1975; Rick, Fobes, & Holle 1977; Caicedo & Schaal 2004; Sifres et al. 2006; Zuriaga et al. 2009; Blanca et al. 2012; Razifard et al. 2020)). Así pues, el mayor grado de diversidad de la especie se da en las entradas procedentes del noroeste de Perú. Las poblaciones de Perú y Ecuador no solo constituyen dos grupos diferenciados a nivel genético, sino que también existen diferencias morfológicas entre ellos en caracteres florales, como la ejerción estigmática, inflorescencias, forma de la hoja, contenido en antocianinas, longitud de entrenudos o tamaño y color del fruto (Rick, Fobes, & Holle 1977; Widrlechner 1987; Sifres et al. 2006; Blanca et al. 2012).

S. lycopersicum var. *cerasiforme* es una variedad mundialmente distribuida en zonas de clima cálido (Peralta & Spooner, 2006), aunque es nativa de la región andina de Ecuador y del norte de Perú (Rick & Fobes 1975). Los estudios realizados sobre su variabilidad genética indican que la migración desde esta región hacia México supuso una reducción del tamaño poblacional, con la consiguiente reducción de su diversidad genética (Rick & Fobes 1975; Blanca et al. 2012; Razifard et al. 2020). Esto hace que la diversidad presente en Perú y Ecuador sea comparable a la descrita en la especie silvestre *S. pimpinellifolium*, mientras que la de México es similar a la variabilidad presente en el tomate cultivado. Pese a que generalmente se ha considerado a esta especie como de fruto pequeño o tomate *cherry*, hay diversos estudios que revelan la presencia de una gran diversidad morfológica. De hecho, estudios genéticos realizados por Blanca et al. (2015) revelaron que tomates procedentes de mercados locales en Ecuador, considerados tomates tradicionales, eran en realidad de la especie *S. lycopersicum* var. *cerasiforme*. Previamente, Rick & Holle (1990), habían descrito también una notable diversidad en el tamaño, la forma y el color de los frutos, forma y tamaño de las hojas, hábito de crecimiento de la planta y pilosidad.

La variabilidad descrita en ambas especies y su capacidad para hibridar con el tomate cultivado las convierten en una fuente de genes de interés en mejora. Pese a ello, solo la especie silvestre *S. pimpinellifolium* ha sido ampliamente estudiada y utilizada en mejora, mientras que el potencial de la especie *S. lycopersicum* var. *cerasiforme* todavía no ha sido muy explotado. Sin embargo, esta especie es una potencial fuente de diversidad molecular para los programas de mejora, ya que es más diversa genéticamente que el tomate cultivado y más variable morfológicamente que la especie silvestre *S. pimpinellifolium* (Ranc et al., 2008).

El objetivo de este capítulo es caracterizar la variabilidad morfológica para caracteres vegetativos, de flor y de fruto presente en un conjunto de accesiones de las especies *S. pimpinellifolium*, *S. lycopersicum* var. *cerasiforme* y *S. lycopersicum* var. *lycopersicum*, seleccionadas para representar la variabilidad existente en el centro de origen y domesticación del tomate cultivado.

2. Material y métodos

2.1. Material vegetal

Se ha analizado una colección de germoplasma compuesta por un total de 163 accesiones, 15 pertenecientes a *S. lycopersicum* var. *lycopersicum* (SLL), 27 a *S. pimpinellifolium* (SP) y 121 a *S. lycopersicum* var. *cerasiforme* (SLC) provenientes de los bancos de germoplasma del Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV, códigos BGV), del Tomato Genetics Resource Center (TGRC, códigos LA) y del United States Department of Agriculture (USDA, códigos PI).

La colección fue seleccionada como parte de un estudio financiado por la National Science Foundation (NSF- IOS 1564366) cuyo objetivo es la identificación de genes de interés en la mejora de calidad del fruto de tomate y la identificación de alelos relacionados con su domesticación. Por ello, esta colección representa gran parte de la diversidad genética, morfológica y geográfica de la especie SLC, así como del tomate cultivado tradicional mexicano y su especie silvestre más próxima, SP.

El rango geográfico de procedencia en estas accesiones engloba las entradas de SLL procedentes de México, las entradas de SP procedentes de Ecuador y Perú y las entradas de SLC procedentes tanto de estos tres países como de otras regiones de Mesoamérica. Las entradas se han agrupado atendiendo a criterios de procedencia geográfica para su análisis. En la Figura 8 se representa el origen geográfico de las entradas y en la Tabla 8 se indica el número de entradas por grupo geográfico según los datos de pasaporte. La información adicional de pasaporte y bancos de germoplasma proveedores está recopilada en el Anexo I. Las entradas se nombraron con un código de 2-3 letras y números, donde la primera letra hace referencia a la especie (P: *S. pimpinellifolium*, C: *S. lycopersicum* var. *cerasiforme*, L: *S. lycopersicum* var. *lycopersicum*), la segunda y tercera al origen geográfico (p: Perú, e: Ecuador, m: México, ma: Mesoamérica) y los números corresponden al orden de las entradas ordenadas por especie y origen geográfico de norte a sur, de manera que números más bajos para entradas de una misma especie indican posiciones más septentrionales.

Tabla 8. Número de entradas de cada especie y origen geográfico.

	Perú	Ecuador	México	Mesoamérica
SP	12	15	0	0
SLC	35	42	24	21
SLL	0	0	15	0



Figura 8. Distribución geográfica de las entradas empleadas en el ensayo. SP: *Solanum pimpinellifolium*. SLC: *Solanum lycopersicum* var. *cerasiforme*. SLL: *Solanum lycopersicum* var. *lycopersicum*.

2.2. Caracterización morfológica

El ensayo de caracterización se llevó a cabo en los invernaderos pertenecientes al Instituto de Conservación y Mejora de Agrobiodiversidad Valenciana (COMAV) de la Universitat Politècnica de València (39.482836, -0.337882), durante la campaña de primavera-verano de 2016, según un diseño aleatorizado con dos plantas de cada accesión, cada una de ellas cultivada en un invernadero diferente. Las plantas se dispusieron al azar en cada invernadero, en macetas de 12L con fibra de coco y con las dosis de fertirrigación adecuadas para ese cultivo.

Para la caracterización morfológica, se incluyeron caracteres de importancia agronómica, así como caracteres que permitieran identificar cada una de las especies o grupos realizados en función de su origen geográfico. La selección de estos caracteres se hizo atendiendo a la experiencia previa de los autores en el estudio de estas especies. Se seleccionaron una serie de descriptores de tomate recogidos en el IPGRI (1996) que describen caracteres vegetativos, de flor y de fruto. Sin embargo, algunos descriptores fueron modificados para una mejor representación de la variabilidad de la colección (Tabla 9 y Anexo II).

Tabla 9. Caracteres cuantitativos y cualitativos evaluados.

Cuantitativos	Cualitativos
Vegetativos	
Longitud de la primera a la décima hoja	Diseción de la hoja
Número de hojas entre la segunda y la tercera inflorescencia	Borde de la hoja
Número de hojas entre la tercera y la cuarta inflorescencias	Posición de la hoja
Altura hasta la primera inflorescencia	Contenido de antocianinas en el tallo
Altura hasta la última inflorescencia	Tipo de hoja
Número total de inflorescencias	Hábito de crecimiento
Anchura del tallo	Color de la hoja
Número de folíolos primarios	
Número de folíolos secundarios	
Longitud y anchura de la hoja	
Flores e inflorescencias	
Longitud de la inflorescencia	Exerción estigmática
Número de flores por inflorescencia	Curvatura del pétalo
Número de pétalos	Color estambres
Número de sépalos	Color del pétalo
Longitud pétalos	Inflorescencia bifurcada
Longitud sépalos	Inflorescencia irregular
Anchura pétalos	Inflorescencia unípara
Anchura sépalos	Inflorescencia espina de pescado
	Inflorescencia con hojas
Fruto	
Color LAB	Protrusión distal
Peso	Obovoide
Número de lóculos	Asimetría
Perímetro	Asimetría ovoide
Índice externo de forma de fruto	Posición máxima anchura
Forma proximal	Hombro verde
Forma distal fruto	Cicatriz pistilar
Altura del hombro	Fasciación del fruto
Ángulo proximal micro	Color del fruto inmaduro
Ángulo proximal macro	Color de la piel
Ángulo distal micro	Anchura de la cicatriz del pedicelo
Ángulo distal macro	Acostillado
Área indentación	Contenido vascular

Para las medidas relacionadas con la forma y tamaño de fruto se usó el programa *Tomato Analyzer* versión 4 (Ramos et al., 2018). Este programa mide diversos caracteres a partir de una serie de parámetros calculados en la sección longitudinal del fruto. Puesto que diversos parámetros dan información sobre el mismo carácter, y para evitar información redundante, se determinaron los coeficientes de correlación entre parámetros que describen un mismo tipo de carácter. Se seleccionaron aquellos con menos de un 0,70 de coeficiente de correlación (todos los coeficientes de correlación entre parámetros que miden un mismo carácter se recogen en el Anexo III). De esta manera, se seleccionaron 16 caracteres poco correlacionados entre sí y representativos de las características del fruto (Tabla 10 y Figura 9).

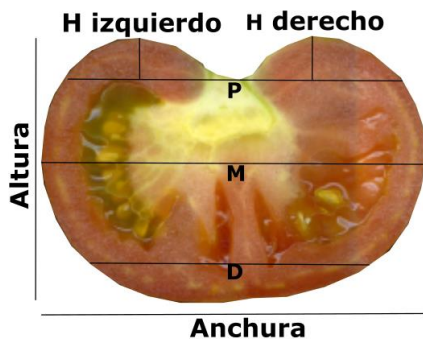
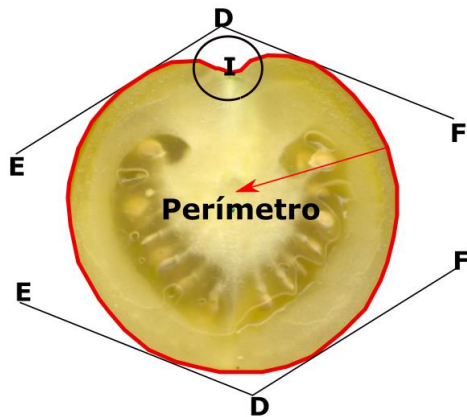
Tabla 10. Caracteres seleccionados de *Tomato Analyzer* y su descripción.

Carácter	Descripción
Perímetro	Longitud del contorno del fruto. Medido en la sección longitudinal.
Índice externo de forma de fruto	Relación entre altura y anchura máximas.
Forma proximal	Relación de la anchura en la parte superior del fruto con la anchura del fruto a la mitad de su altura.
Forma distal	Relación de la anchura en la parte inferior del fruto con la anchura del fruto a la mitad de su altura.
Forma fruto triangular	Relación entre la anchura en la parte superior e inferior del fruto.
Altura del hombro	Relación entre la altura del hombro respecto de la altura máxima.
Ángulo proximal micro	Ángulo entre las líneas dibujadas a través del perímetro en la zona proximal del fruto. Los puntos para el ajuste comprenden el 1% del perímetro a cada lado del punto central.
Ángulo proximal macro	Ángulo entre las líneas dibujadas a través del perímetro en la zona proximal del fruto. Los puntos para el ajuste comprenden el 5% del perímetro a cada lado del punto central.
Ángulo distal micro	Ángulo entre las líneas dibujadas a través del perímetro en la zona distal del fruto. Los puntos para el ajuste comprenden el 1% del perímetro a cada lado del punto central.
Ángulo distal macro	Ángulo entre las líneas dibujadas a través del perímetro en la zona distal del fruto. Los puntos para el ajuste comprenden el 5% del perímetro a cada lado del punto central.
Área de indentación	Área de indentación respecto al área total del fruto y multiplicada por 10.
Protrusión distal	Relación entre el área de la protrusión en la zona distal y el área total del fruto.
Asimetría	Distancia promedio entre una línea vertical a la mitad de la anchura del fruto y el punto medio del ancho del fruto a cada altura.
Obovoide	Si el área es mayor por debajo de la altura media que por encima, este parámetro se calcula teniendo en cuenta la anchura máxima (W), la altura a la máxima anchura (y), el ancho promedio por encima de esa altura (w1) y el ancho promedio por debajo de esa altura (w2). Si no se cumple la condición, el parámetro es 0.
Asimetría ovoide	Si el área es mayor a la altura media del fruto, se calcula como la distancia promedio entre una línea horizontal en el fruto y el punto medio de la altura de la fruta en cada anchura. Si no se cumple la condición, el parámetro es 0.
Posición máxima anchura	Relación de la altura donde el fruto alcanza su máxima anchura, con respecto a la altura total del fruto.

Ángulo distal = ángulo EDF

Ángulo proximal = ángulo EDF

$$\text{Indentación} = \frac{\text{Área I}}{\text{Área total}} \times 10$$

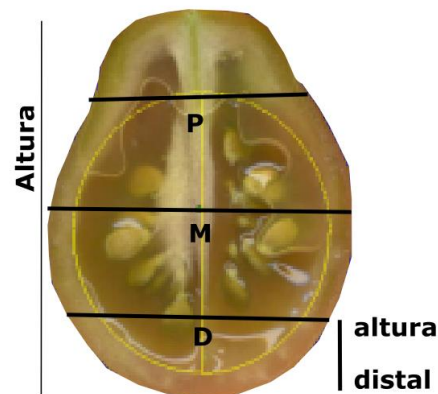


$$\text{Índice externo forma fruto} = \frac{\text{Altura}}{\text{Anchura}}$$

$$\text{Forma distal} = \frac{\text{Anchura distal (D)}}{\text{Anchura mital altura (M)}}$$



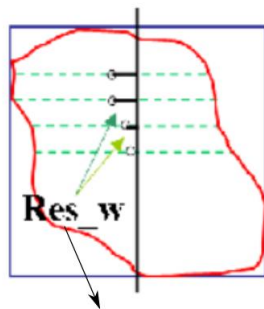
$$\text{Protusión distal} = \frac{\text{Área protusión distal}}{\text{Área del fruto}} \times 100$$



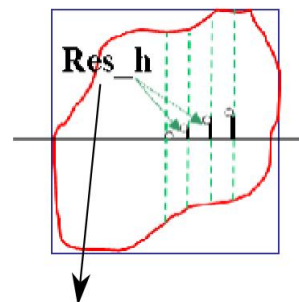
$$\text{Forma proximal} = \frac{\text{Anchura proximal (P)}}{\text{Anchura mital altura (M)}}$$

$$\text{Altura hombro} = \frac{\text{Media(H derecho, H izquierdo)}}{\text{Altura}}$$

$$\text{Obovoide} = 1/2 * \text{scale_ob}(\text{altura a la máxima anchura}) * (1 - P/M + D/M)$$



$$\text{Asimetría} = \text{Media Res}_w$$



$$\text{Asimetría ovoide} = \text{Media Res}_h$$

Figura 9. Descripción de los parámetros del *Tomato Analyzer* seleccionados. Imagen modificada de *Tomato Analyzer user manual*, versión 4.

2.3. Análisis de los datos

Con el fin de descartar un posible efecto de bloque debido al factor invernadero y poder combinar los datos de ambos invernaderos, se analizaron las diferencias debidas a este factor para cada carácter. Para ello, en el caso de caracteres cuantitativos se realizó una prueba t de Student o de Mann-Withney-Wilcoxon en función de si presentaban o no una distribución normal. Para los datos cualitativos se realizó un test exacto de Fisher. Puesto que no se detectó efecto de invernadero, se combinaron los datos de ambos usando la media en el caso de los caracteres cuantitativos. En el caso de los cualitativos, se emplearon los datos obtenidos en cada invernadero, ya que, en ocasiones, éstos fueron distintos para un mismo carácter. Todos los análisis estadísticos y representaciones gráficas se realizaron usando el software estadístico R versión 3.6.2 (R Core Team, 2020).

Para determinar qué caracteres morfológicos presentaban diferencias estadísticamente significativas entre especies o entre los distintos grupos geográficos establecidos dentro de cada especie, se utilizó el test exacto de Fisher en el caso de los caracteres cualitativos. Los p-valores se ajustaron para corregir por comparaciones múltiples usando el método FDR. En el caso de los datos cuantitativos, primero se comprobó la normalidad de los datos para cada carácter usando el test de Shapiro-Wilk. Si el carácter no seguía una distribución normal, se realizó la transformación de los datos usando el paquete de R *bestNormalize* (Peterson, 2019) y se volvió a comprobar la normalidad una vez transformados. En aquellos casos en los que los datos transformados tampoco se ajustaron a la normalidad, se transformaron en caracteres cualitativos mediante la creación de escalas para agrupar los valores numéricos (Tabla 11). Estas agrupaciones se establecieron con base en los valores máximo y mínimo que se observaron en los datos fenotípicos, tanto de toda la colección como de cada grupo geográfico. Con aquellos datos que finalmente cumplían el criterio de normalidad se realizó un MANOVA, con la función “manova” del paquete *stats* de R (Wang et al., 2019).

Tabla 11. Criterios empleados en la conversión de caracteres cuantitativos a caracteres cualitativos.

Caracteres vegetativos	Categorías del carácter cualitativo		
	0	1	2
Número de hojas entre inflorescencias	<2	>2	
Altura 1ª inflorescencia	<55 cm	56 a 91 cm	>91 cm
Altura última inflorescencia	<180 cm	>180 cm	
Número inflorescencias	<6	>6	
Anchura del tallo	<12 mm	>12 mm	
Número foliolos	<7	>7	
Número de foliolillos	<15	>15	
Longitud de la hoja	20 a 40 cm	>40 cm	
Anchura hoja	<40 cm	40 a 55 cm	>55 cm
Caracteres flores e inflorescencias			
Número de pétalos	<5	>5	
Longitud del pétalo	<1,2 cm	>1,2cm	
Anchura del pétalo	<0,45cm	>0,45 cm	
Número de sépalos	<5	>5	
Longitud de sépalos	<1cm	>1cm	
Anchura del sépalo	<0,15 cm	>0,15cm	
Longitud de anteras	<0,8 cm	>0,8cm	

Longitud inflorescencia	<30 cm	30-60cm	>60 cm
Número de flores en la inflorescencia	<10	>10	
Caracteres fruto			
Peso	<3 g	3 a 20 g	>20 g
Perímetro	<10cm	>10 cm	
Índice externo	<1	>1	
Forma distal del fruto	<0,62	>0,62	
Forma proximal	<0,68	>0,68	
Altura del hombro	<0,01	>0,01	
Posición máxima anchura	<0,5	>0,5	
Ángulo proximal	<130	>130	
Ángulo distal	<120	>120	
Protrusión	<0,05	>0,05	
Asimetría	<0,01	>0,01	
Asimetría ovoide	0	>0	
Obovoide	0	>0	
Área de indentación	<0,01 cm ²	>0,01 cm ²	
Número de lóculos	<3	3 a 10	>10
Color L	<41	>41	
Color a	<0	>0	

Con el objetivo de explorar la variabilidad existente entre especies y dentro de especie, se realizó un análisis de componentes principales (PCA). Este análisis se realizó combinando datos cuantitativos y cualitativos, incluyendo solamente aquellos caracteres significativamente diferentes entre grupos. Para ello, los caracteres cuantitativos se transformaron en variables cualitativas, siguiendo los criterios establecidos en la Tabla 11. Una vez transformados todos los datos en caracteres cualitativos, éstos se convirtieron en variables binarias o dummy con el paquete *fastDummies* de R (Kaplan y Schlege, 2020). Es decir, el carácter fue escindido en tantos niveles como presentaba, asignando ceros y unos dependiendo de la ausencia o presencia de ese nivel para un carácter determinado en cada entrada. Los PCAs se realizaron con la función “prcomp” de la librería *stats* de R, tras estandarización de los datos. Para facilitar la visualización de los resultados, solamente se mostrarán en las gráficas de los siguientes apartados aquellos caracteres con una contribución a las componentes principales primera o segunda superior a 0,17.

3. Resultados

3.1. Análisis de la variabilidad morfológica entre grupos taxonómicos.

Al analizar las diferencias entre los taxones en estudio, SP, SLC y SLL se encontró que gran parte de los caracteres vegetativos, de inflorescencia y flor y del fruto mostraron diferencias estadísticamente significativas según los resultados de los análisis MANOVA y test de Fisher (Anexo IV).

Los caracteres evaluados que no pudieron ser normalizados fueron el número de hojas entre inflorescencias, el número total de inflorescencias, la anchura de la hoja y el número de folíolos primarios, el número de pétalos y sépalos, el número de lóculos del fruto y los parámetros L, a y b del color del fruto. Además, tampoco se pudieron normalizar algunos caracteres medidos con *Tomato Analyzer*, como el índice de la forma externa del fruto, forma distal del fruto, ángulo proximal macro, ángulo distal macro, protrusión distal, forma obovoide, altura del hombro, altura a la que el fruto presenta su máxima anchura y asimetría ovoide.

En el análisis de componentes principales realizado con todos los caracteres convertidos en variables dummy, la componente principal 1 (CP1) y 2 (CP2) explicaron un 14,6% y 6,9% de la varianza observada, respectivamente. En la Figura 10 se representa la contribución de los caracteres a las dos primeras componentes principales y en la Figura 11 la distribución de las entradas de acuerdo con estas componentes. Tanto los caracteres como la contribución de cada uno de ellos a las tres primeras componentes aparecen recogidos en el Anexo V. Los caracteres con una mayor contribución negativa a la CP1 fueron el tipo de hoja estándar y de mayor longitud, peso elevado del fruto, perímetro y número de lóculos, mientras que la hoja tipo pimpinellifolium, sin disección y borde entero, y el peso del fruto pequeño tuvieron una contribución positiva. Estas contribuciones dieron lugar a la separación de las entradas de SLL y SP a lo largo de la CP1 (Figura 11). Por su parte, las entradas de SLC, con características intermedias en muchos casos a los dos taxones anteriores, quedaron distribuidas en la parte central del gráfico.

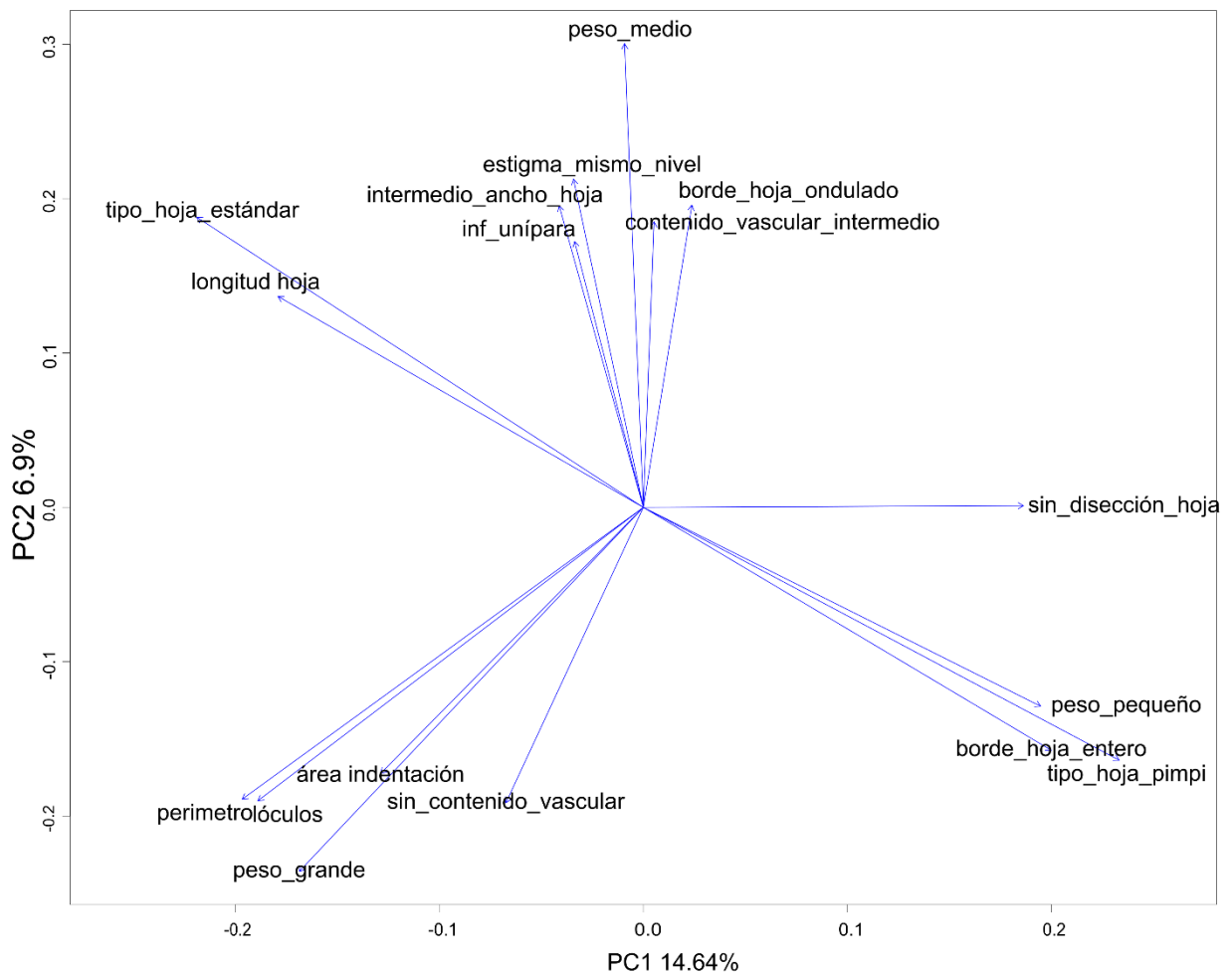


Figura 10. Caracteres con una correlación superior a 0,17 con las CP1 y CP2 del PCA realizado con todas las especies.

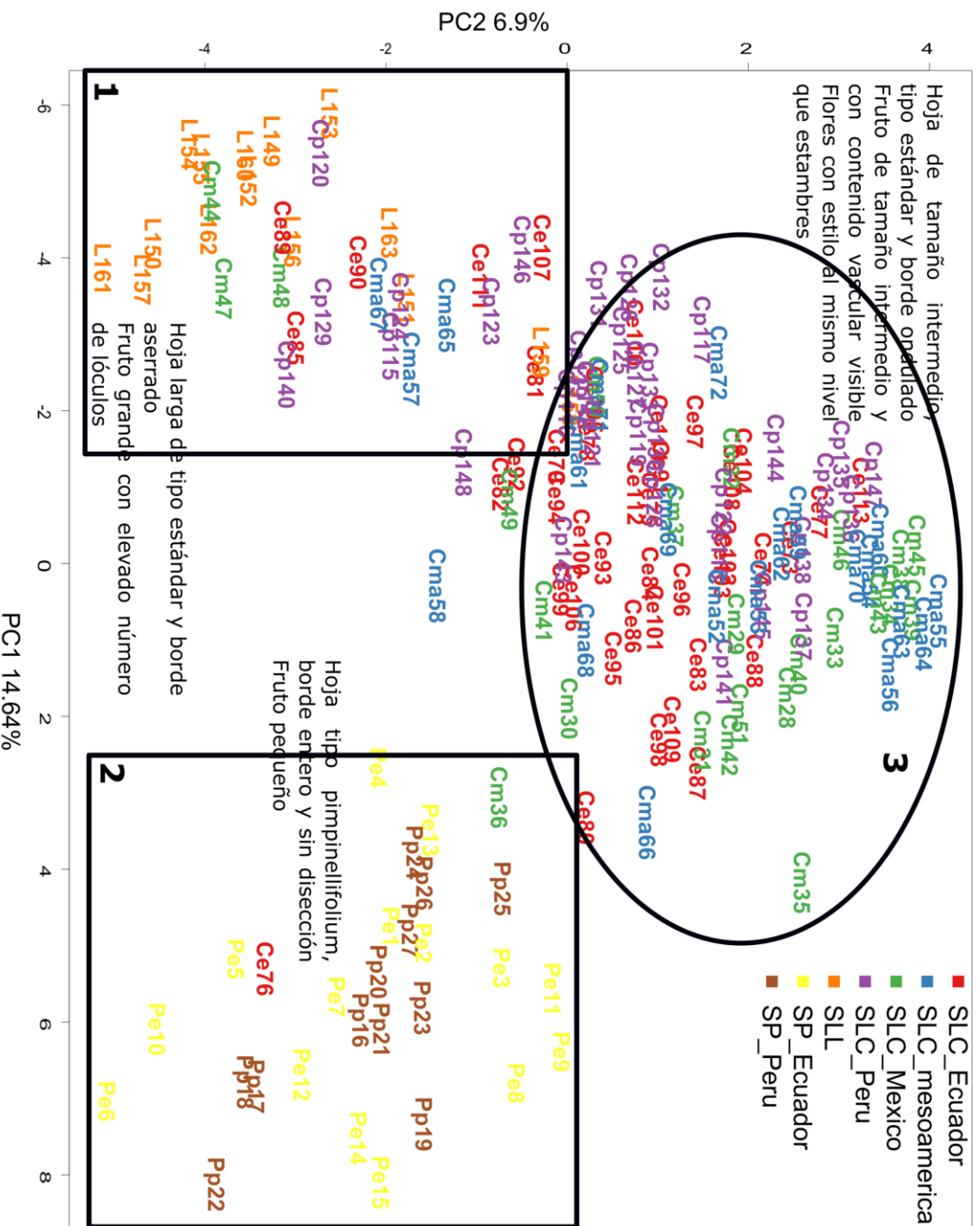


Figura 11. Análisis de componentes principales efectuado con los caracteres que presentaron diferencias significativas entre especies. El grupo 1 representa a las entradas de la especie SLL, junto con algunas SLC de fruto grande. El grupo 2 engloba principalmente a la especie SP. El grupo 3 está constituido por la mayoría de las entradas SLC.

La CP2 correlacionó positivamente con el tipo de hoja estándar, de anchura intermedia y borde ondulado, peso intermedio del fruto, frutos con contenido vascular intermedio, inflorescencias de tipo uníparo y estigma al mismo nivel que los estambres. Estos caracteres son los que mejor definen al grupo de entradas de SLC y determinan el desplazamiento de la mayor parte de entradas de este grupo taxonómico hacia la parte superior de la gráfica (Figura 11). Por otro lado, esta componente estuvo negativamente correlacionada con los caracteres anteriormente mencionados y que definían a los grupos SP y SLL. Estos caracteres incluyen el peso pequeño y grande del fruto, perímetro y número de lóculos y hojas de tipo *pimpinellifolium*, entre otros.

Cabe destacar algunas excepciones concretas a las agrupaciones anteriormente descritas. Las entradas Cm36 y Ce76 de la especie SLC se agrupan junto con la especie silvestre SP debido principalmente a la similitud de sus hojas a las de tipo *pimpinellifolium*. También hay una serie de entradas del grupo SLC que se agrupan con la especie SLL, principalmente debido a su gran tamaño de fruto, similar a los de la especie cultivada. Éstas presentaron frutos superiores a los 9,2 g de peso medio que tiene el grupo SLC, situándose entre 13,5 g (Ce107) y 81,8 g (Cm44). Además, algunas presentaron un valor elevado de área de indentación, superior al valor medio del grupo SLC de 0,09 cm², como las entradas Ce85 (0,11), Cm44 (0,18), Cm47 y Cm 48 (0,16), Ce90 (0,14) y Cp129 (0,21). Por último, salvo la entrada Cma57 que presentó frutos de 2 ó 3 lóculos, el resto de las entradas se caracterizaron por presentar frutos con un número de lóculos superior a los 2 lóculos que tiene de moda el grupo SLC.

Con base en los resultados anteriores, se han identificado una serie de caracteres que diferencian principalmente a la especie silvestre del tomate cultivado, teniendo *S. lycopersicum* var. *cerasiforme* características intermedias entre ellas y conteniendo una gran diversidad morfológica. A continuación, se da una resumida descripción de cada taxón, incluyendo imágenes para visualizar las diferencias morfológicas entre ellos (Figuras 12-14).

La especie *S. pimpinellifolium* se caracteriza por presentar hojas sencillas de tipo *pimpinellifolium*, de menor tamaño, sin disección y con borde entero. Además, esta especie suele presentar de forma habitual inflorescencia uníparas. Sus flores tienden a presentar el estigma exerto y pétalos muy curvados. Por último, sus frutos son de pequeño tamaño, de forma redondeada y generalmente de color rojo intenso. La especie *S. lycopersicum* var. *lycopersicum* se caracteriza principalmente por presentar hojas de mayor tamaño y complejidad. Además, sus frutos son de mayor tamaño, forma aplastada, multiloculares y, en ocasiones, fasciados. La especie *S. lycopersicum* var. *cerasiforme* presenta alta variabilidad morfológica, generalmente con frutos y flores de características intermedias entre tomate cultivado y la especie silvestre. Sin embargo, la mayoría de las accesiones de SLC se diferenciarían de los otros dos taxones principalmente por presentar el estigma inserto e inflorescencias de tipo uníparo, hojas de anchura intermedia y con el borde ondulado.

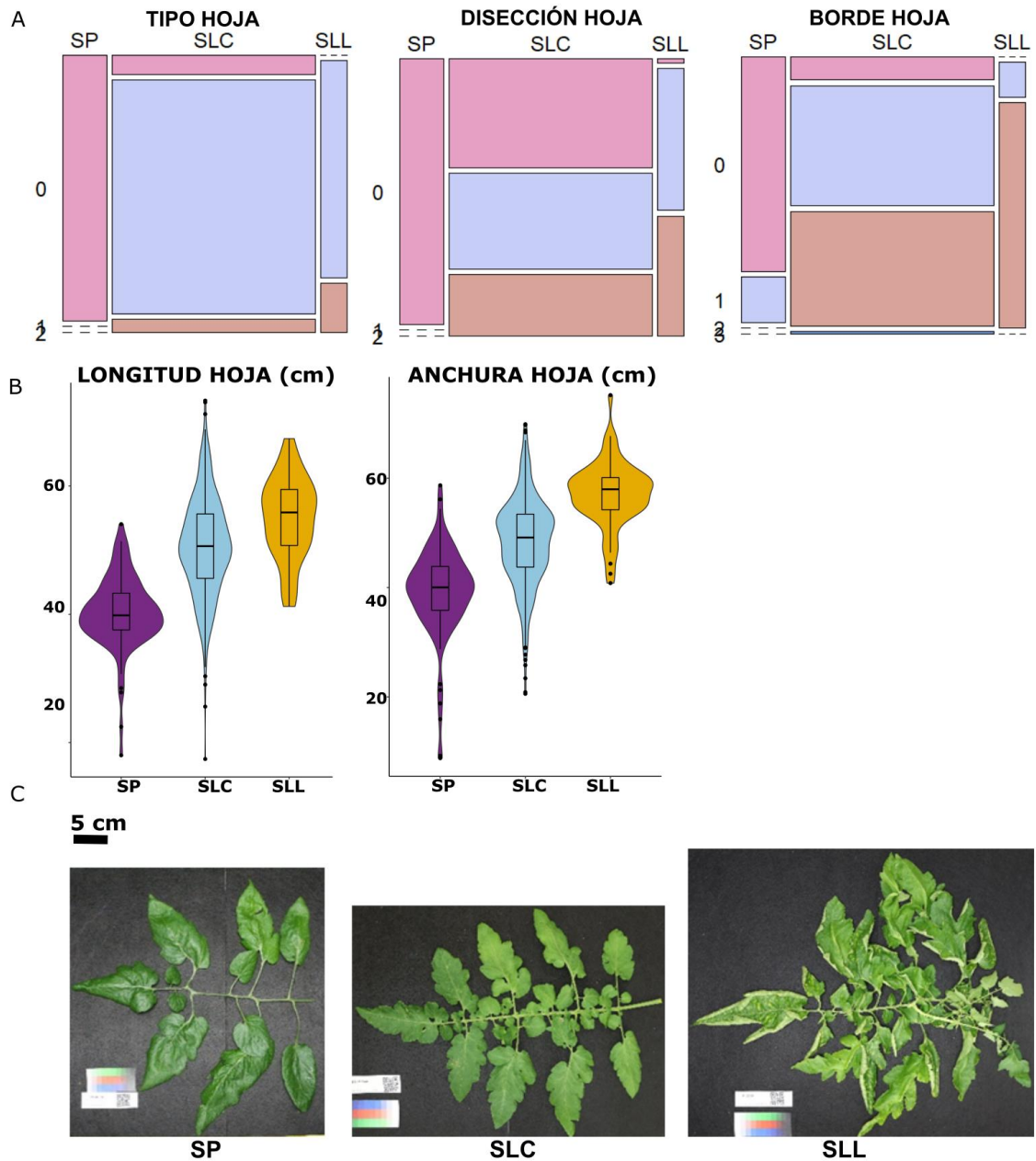


Figura 12. A. Gráficos mosaico de los caracteres vegetativos de y que permite diferenciar a *S. pimpinellifolium* (SP), *S. lycopersicum* var. *cerasiforme* (SLC) y *S. lycopersicum* var. *lycopersicum* (SLL) relacionados con hojas con mayor contribución a las CP1 y CP2 del PCA. Tipo hoja: 0: pimpinellifolium, 1: estándar, 2: doblemente dividida. Disección hoja: 0: baja, 1: media, 2: alta. Borde hoja: 0: entero, 1: ondulado, 2: aserrado, 3: fuertemente aserrado. Nótese que la anchura de las barras es proporcional al número de entradas de cada grupo. B. Gráfico violín con los datos cuantitativos originales para anchura y longitud de las hojas. C. Principales diferencias entre las hojas de SP, SLC y SLL.

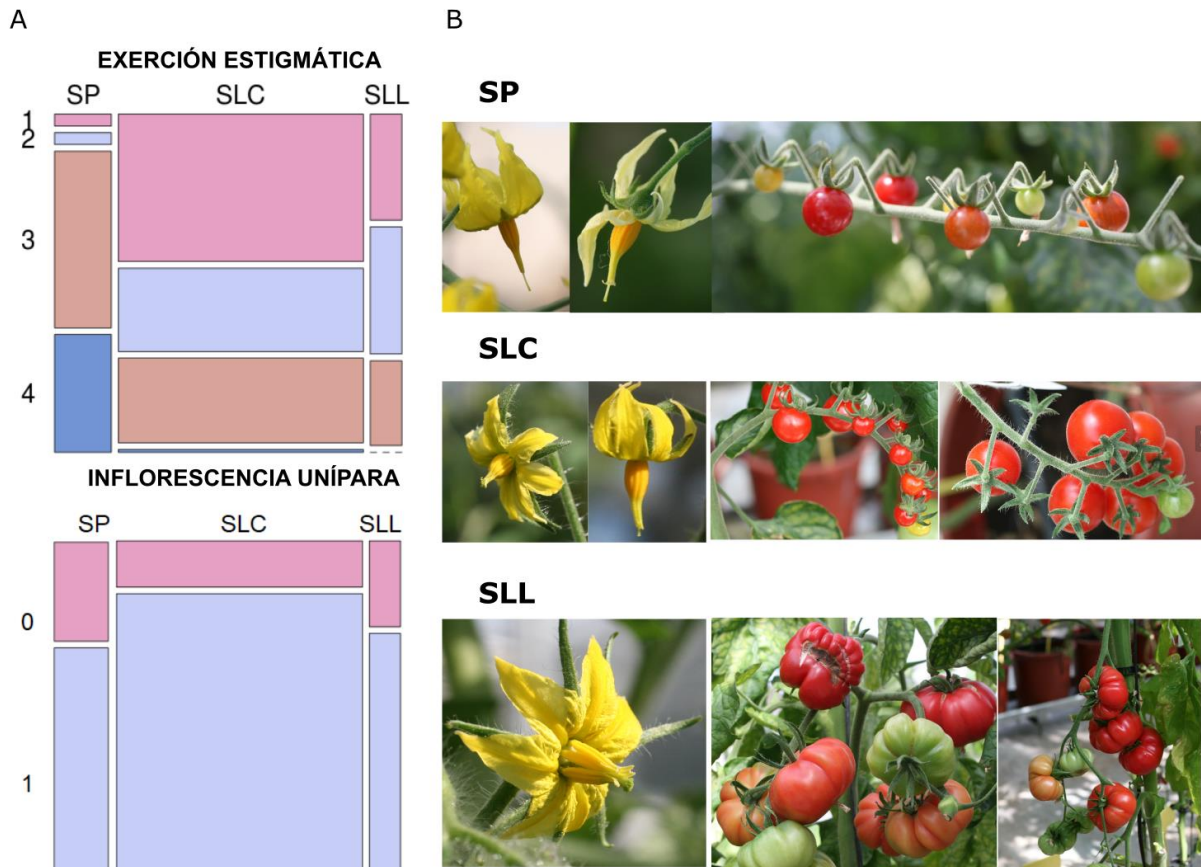


Figura 13. A. Gráficos mosaico para los caracteres de flores e inflorescencias en las tres especies con mayor correlación en el PCA. Exerción estigmática: 1: inserto, 2: mismo nivel, 3: ligeramente exerto, 4: altamente exerto. Inflorescencia unípara: 0: ausente; 1: presente. Nótese que la anchura de las barras es proporcional al número de entradas de cada grupo. B. Principales diferencias en flores e inflorescencias de SP, SLC y SLL.

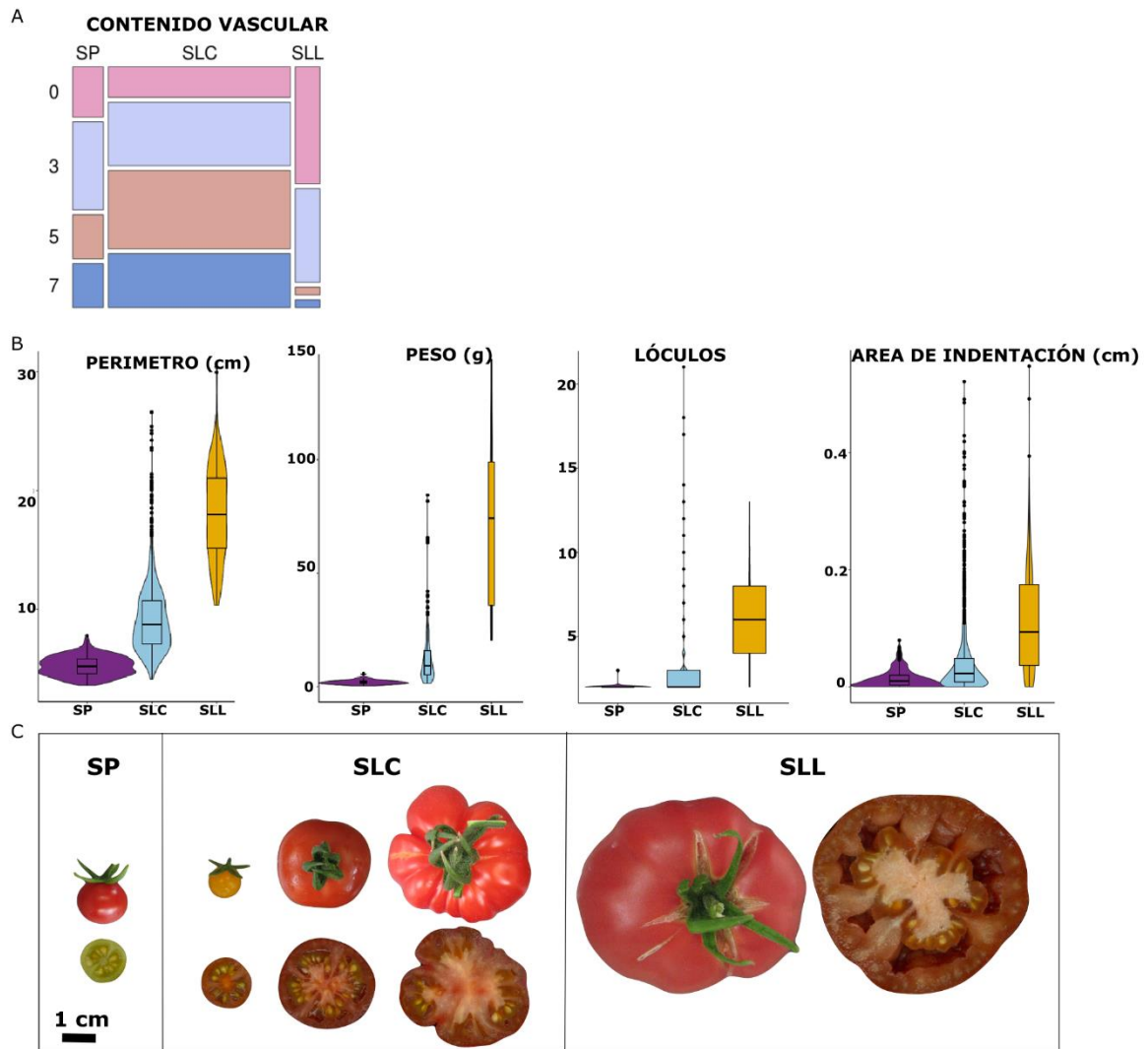


Figura 14. A. Gráficos mosaico para los caracteres vegetativos relacionados con hojas con mayor correlación en el PCA para las tres especies. Contenido vascular: 0: ausente, 3: bajo, 5: intermedio, 7: severo. B. Gráfico violín con los datos cuantitativos originales para perímetro, peso, número de lóculos y área de indentación del fruto. C. Principales diferencias en frutos de SP, SLC y SLL.

3.2. Variabilidad morfológica en *S. pimpinellifolium*

3.2.1. Origen geográfico

Las entradas de *S. pimpinellifolium* se dividieron en dos grupos atendiendo a su origen geográfico: Ecuador y Perú (Figura 15, Tabla Suplementaria 1). Las de Ecuador fueron colectadas en regiones con climas húmedos y templados o calurosos como las dos entradas de Esmeraldas (Pe1 y Pe2) y las 10 de Manabí (Pe3 a Pe12). Las entradas Pe13 (Azúay) y Pe14 y Pe15 (El Oro) provinieron del sur de Ecuador, de una región con clima tropical árido. Las entradas de Perú proceden mayoritariamente de la región costera, donde predomina el clima tropical o subtropical árido. Estas entradas proceden de las provincias de Piura (Pp16 a Pp21) y Lambayeque (Pp22, Pp26 y Pp27). Además, se incluyeron tres entradas procedentes de la región andina (Pp23 a Pp25), donde el clima es templado húmedo.



Figura 15. Distribución geográfica de las entradas de la especie *S. pimpinellifolium*.
3.2.2. Análisis de la variabilidad morfológica

Varios caracteres relacionados con flores e inflorescencias y tamaño y apariencia del fruto presentaron diferencias significativas entre los dos grupos geográficos (Anexo VI). Sin embargo, otros caracteres como el número de pétalos y sépalos de la flor (cinco en todos los casos), frutos biloculares y hojas con siete folíolos principales fueron invariantes dentro de la especie. El parámetro “a” del color del fruto, la longitud de la hoja y el número de hojas entre inflorescencias no siguieron una distribución normal tras la transformación de los datos, por lo que se transformaron en caracteres cualitativos para su análisis.

En el análisis de componentes principales realizado con todos los caracteres convertidos en variables dummy, las CP1 y CP2 explicaron un 22,95 % y 12,95 % de la varianza observada, respectivamente. Los caracteres que presentaron una contribución superior a 0,17 a alguna de estas componentes se representan en la Figura 16, incluyéndose en el Anexo VII la contribución de todos los caracteres.

La CP1 estuvo positivamente correlacionada con caracteres relativos a las inflorescencias, flores y frutos. Éstos fueron, inflorescencias largas, con muchas flores y terminadas en hojas, flores con el estigma muy exerto y frutos con hombros verdes. Los caracteres que contribuyeron negativamente a esta componente fueron inflorescencias cortas, sin hojas ni brotes, flores con menor ejerción estigmática y pétalos de color amarillo claro, y frutos sin hombros verdes. Esto determinó el posicionamiento de las entradas de Perú hacia la derecha y las de Ecuador hacia la izquierda del gráfico (Figura 17). Por su parte, la CP2 tuvo una mayor contribución positiva de los siguientes caracteres, ausencia de antocianina en el tallo, inflorescencias de longitud intermedia, ejerción estigmática intermedia y flores con pétalos amarillos y una contribución negativa de presencia de antocianina, elevada ejerción estigmática y longitud de las anteras (Figura 16). Esto contribuyó a separar las entradas de ambas procedencias a lo largo de esta componente, como consecuencia de la variación existente para estos caracteres en las entradas de ambos grupos geográficos (Figura 17).

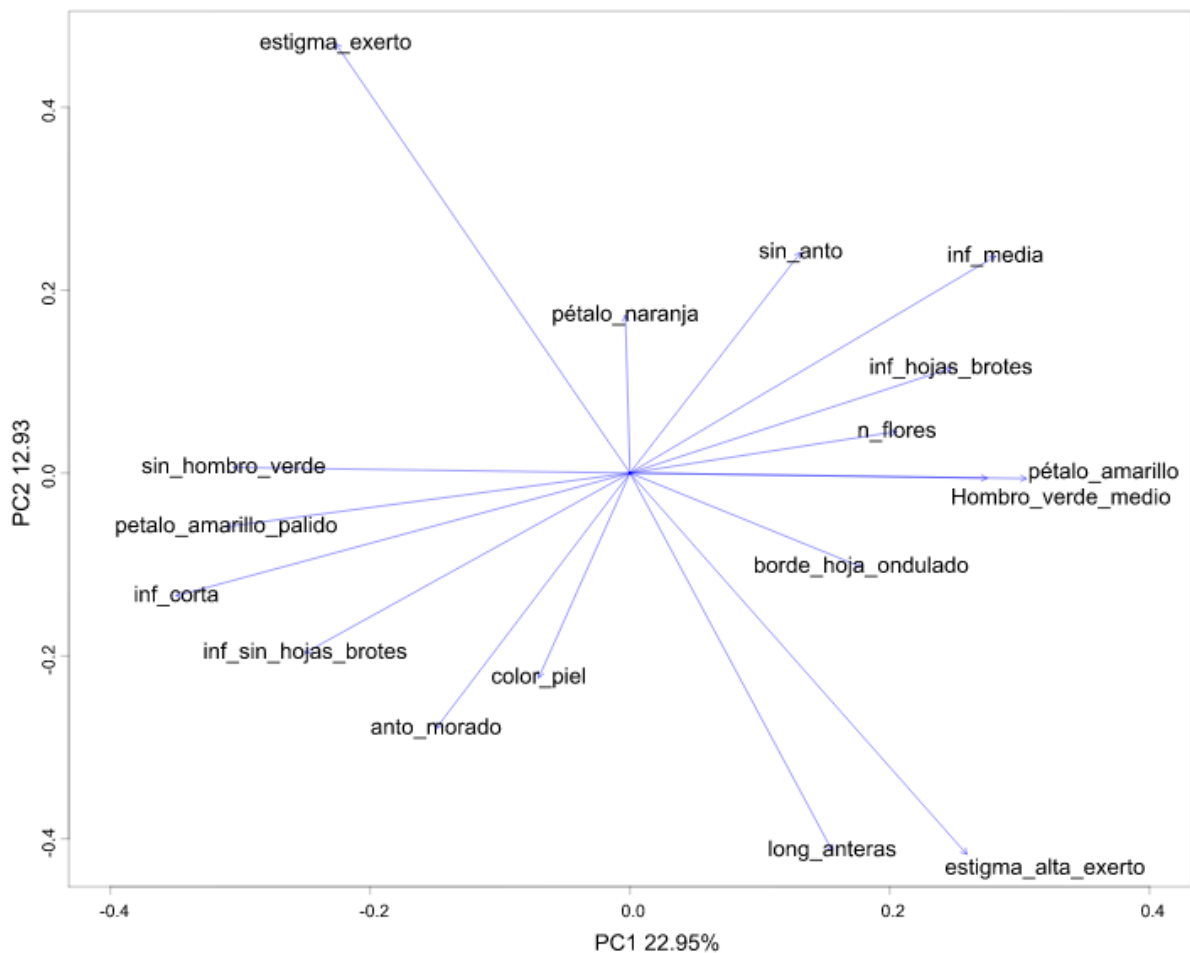


Figura 16. Caracteres con una correlación superior a 0,17 con las CP1 y CP2 del PCA realizado con *S. pimpinellifolium* de Perú y Ecuador. inf : inflorescencia, n: número, anto: antocianinas.

Como se ha indicado, la proyección de las entradas en el plano definido por las dos primeras componentes permitió separar las entradas en dos grupos de acuerdo con su procedencia, Perú y Ecuador (Figura 17), según la PC1. Sin embargo, no son dos grupos estancos, encontrándose una distribución continua a lo largo de esta CP y con algunas excepciones. El grupo 1 engloba a parte de las entradas de Perú, junto con las entradas Pe9 y Pe12 (Manabí, Ecuador), por poseer un número de flores por inflorescencia y una longitud de

las anteras superiores a las medias en Ecuador para dichos caracteres. El grupo 2 lo constituyen las accesiones de Ecuador, y también engloba a las entradas Pp23 y Pp24 de Cajamarca, Perú. En conjunto, las entradas de este grupo se caracterizan por presentar el estigma ligeramente exerto, inflorescencias cortas y sin hojas ni brotes, color morado del tallo por la presencia de antocianina y ausencia de hombros verdes en los frutos. Las entradas Pp23 y Pp24 de Perú se agrupan con las entradas ecuatorianas debido a la longitud de la inflorescencia (18 cm y 20 cm, respectivamente), más corta que la media de los 36 cm de Perú. Además, presentan estigmas ligeramente exertos y frutos sin hombros verdes, a diferencia de las características generales del grupo 1. Cabe destacar la presencia en la colección de tres entradas con frutos de color amarillo (Pe13, de Ecuador y Pp26 y Pp27, de Perú). Así mismo, es remarcable la posición que ocupan en el gráfico las entradas Pe13, Pe14 y Pe15, todas ellas procedentes del sur de Ecuador, en ambientes más similares a los de las zonas costeras áridas de Perú donde fueron colectadas la mayor parte de las entradas de este país.

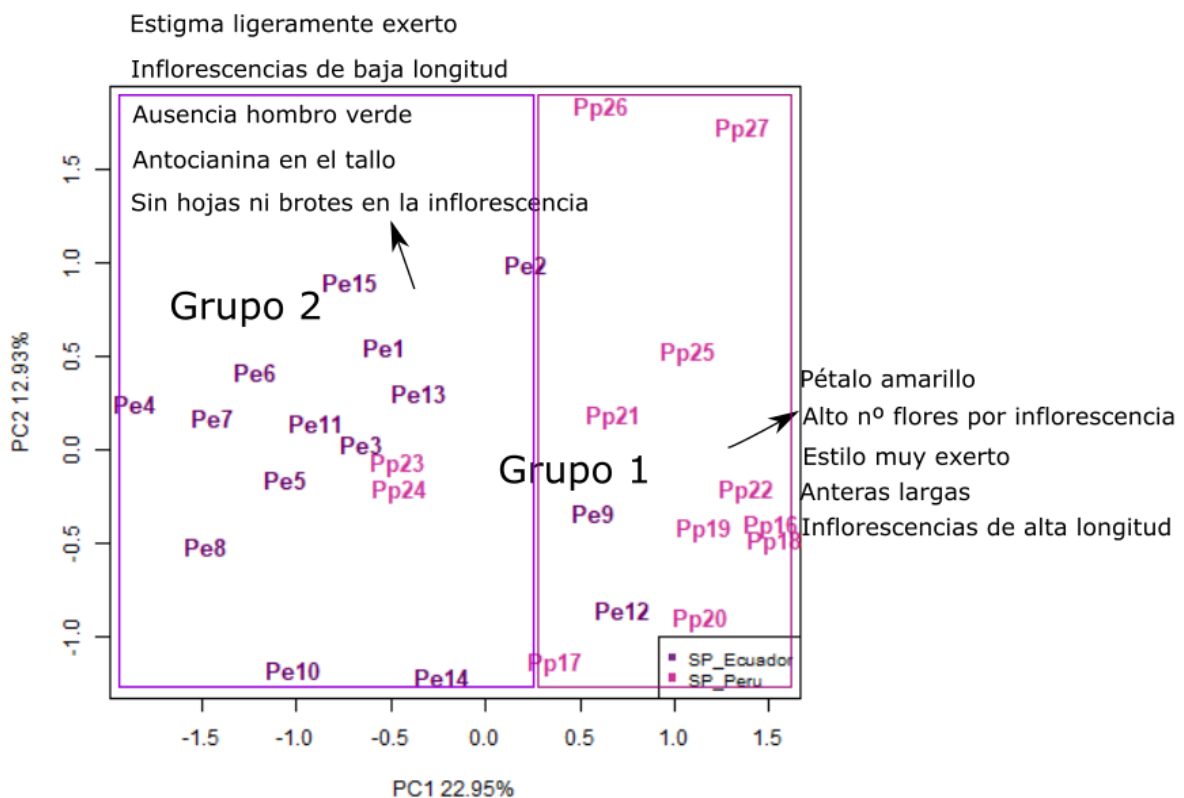


Figura 17. Análisis de componentes principales efectuado con caracteres significativamente diferentes entre los grupos de entradas de *S. pimpinellifolium* establecidos según su procedencia geográfica.

Con base en los resultados anteriores, se pueden establecer algunas diferencias generales entre ambos grupos en cuanto a caracteres vegetativos, de inflorescencia, flor y fruto. Se incluyen a continuación una descripción resumida de los caracteres más discriminantes de cada grupo, así como gráficos mosaico, violín e imágenes representativas de dichos caracteres para cada uno de ellos (Figura 18-21).

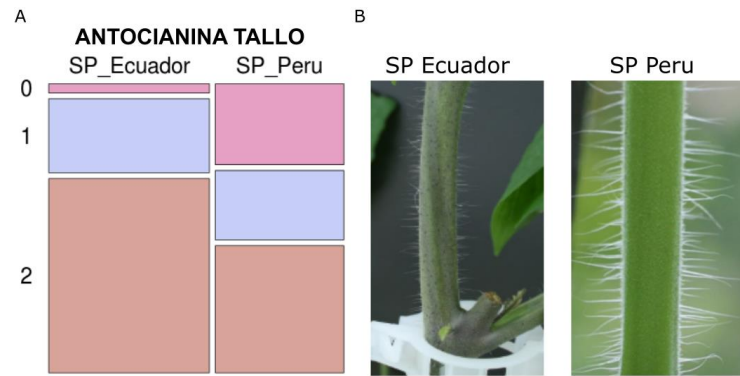


Figura 18. A. Gráfico mosaico de los caracteres de tipo vegetativo que presentaron diferencias significativas entre las entradas de *S. pimpinellifolium* (SP) procedentes de Ecuador y Perú. Antocianina tallo: 0: Ausente, 1: Morado claro, 2: Morado. **B.** Principales diferencias en el contenido en antocianinas entre *S. pimpinellifolium* de Ecuador y Perú.

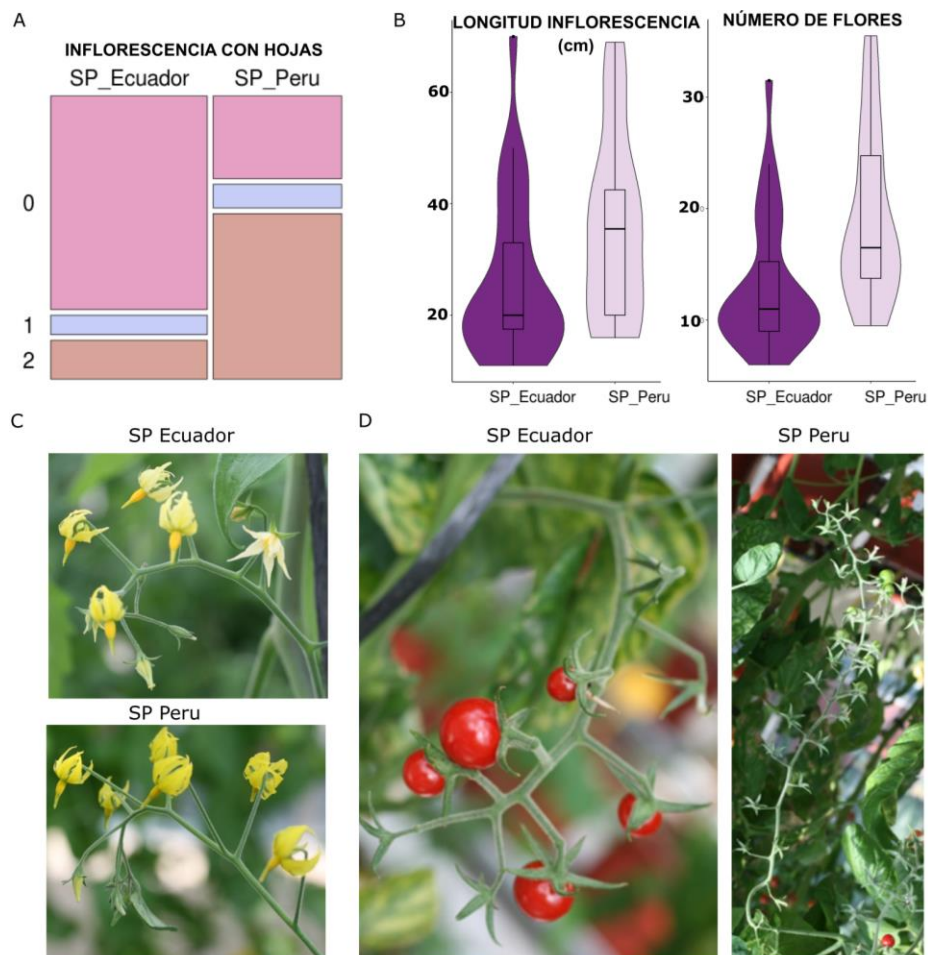


Figura 19. A. Gráfico mosaico de los caracteres de inflorescencia que caracterizan las entradas de *S. pimpinellifolium* procedentes de Ecuador y Perú. Inflorescencia con hojas: 0, ausente; 1, presente; 2, brotes y hojas. Longitud de la inflorescencia: 0: menor de 30 cm, 1: entre 30 y 60cm, 2: mayor a 60cm. Número de flores en la inflorescencia: 0: menor 10, 1: mayor 10. **B.** Gráfico violín con los datos cuantitativos originales para longitud y número de flores en la inflorescencia. **C.** Diferencias en la presencia y ausencia de hojas y brotes entre *S. pimpinellifolium* (SP) de Ecuador y Perú. **D.** Diferencias en la longitud y número de flores de las inflorescencias de *S. pimpinellifolium* (SP) de Ecuador y Perú.

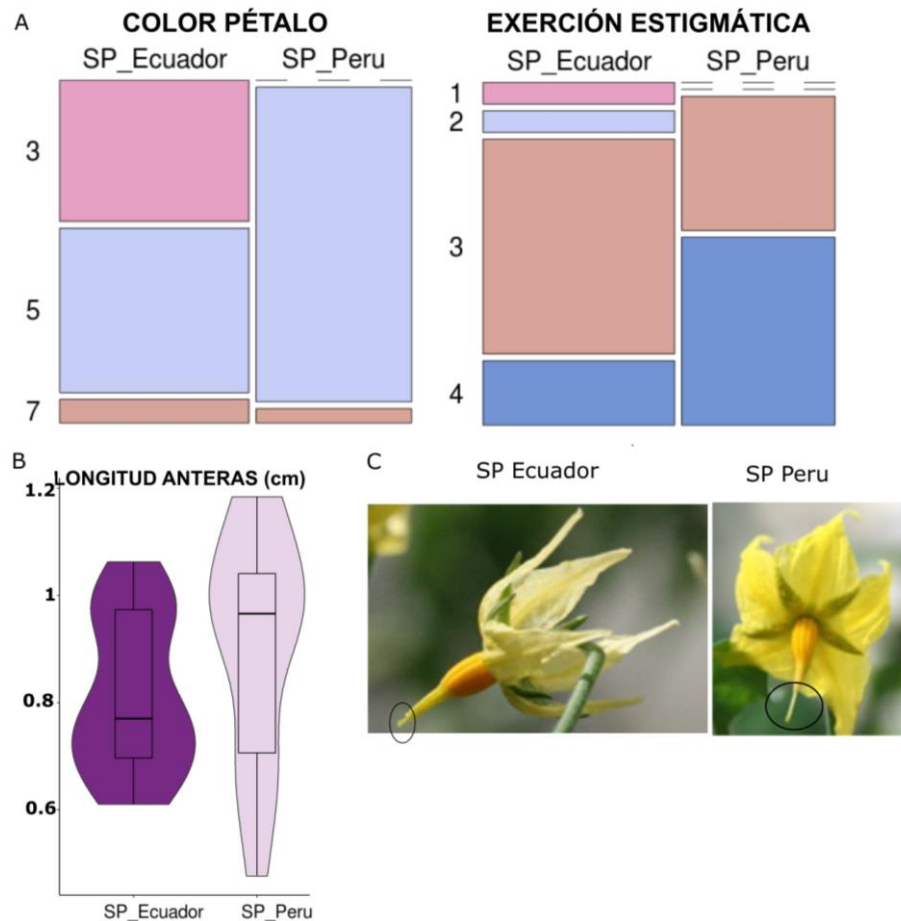


Figura 20. A. Gráfico mosaico de los caracteres de flor que caracterizan las entradas de *S. pimpinellifolium* procedentes de Ecuador y Perú. Color del pétalo: 3, amarillo pálido; 5, amarillo; 7, naranja. Exerción estigmática: 1, inserto; 2, mismo nivel; 3, ligeramente exerto, 4: muy exerto. **B. Gráfico violín con los datos cuantitativos originales para longitud de anteras. **C. Diferencias en el color de los pétalos y la exerción estigmática entre *S. pimpinellifolium* (SP) de Ecuador y Perú.****

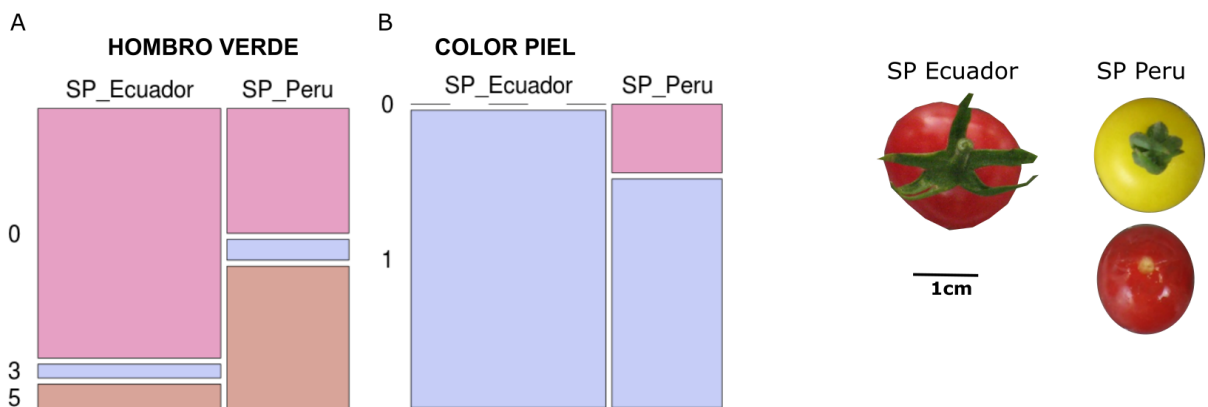


Figura 21. A. Gráfico mosaico que describen los caracteres de fruto que presentan diferencias significativas entre *S. pimpinellifolium* procedentes de Ecuador y Perú. Hombro verde: 0: uniforme, 3: verde claro, 5: verde medio. Color piel: 0: sin color, 1: amarillo. **B. Diferencias en el color de la piel entre frutos rojos y frutos amarillos de *S. pimpinellifolium* (SP).**

Las entradas de SP Ecuador se caracterizaron por presentar un mayor contenido de antocianina en el tallo (Figura 18), confiriéndole un color morado más intenso, inflorescencias cortas, sin hojas ni brotes (Figura 19). Sus flores fueron generalmente de color amarillo pálido o amarillo, con estigmas ligeramente exertos (Figura 20) y sus frutos fueron de coloración uniforme sin hombros verdes (Figura 21).

Las entradas SP Perú presentaron ausencia o baja concentración de antocianina en tallo, hojas y brotes en la inflorescencia, mayor longitud de las mismas y mayor número de flores. Sus flores se caracterizaron por una mayor longitud de las anteras y una alta ejerción estigmática y por el color amarillo de los pétalos. Los frutos presentaron hombros verdes en algunos casos.

3.3. Variabilidad morfológica en *S. lycopersicum* var. *cerasiforme*

3.3.1. Origen geográfico

Las entradas de la especie *S. lycopersicum* var. *cerasiforme* se dividieron en cuatro grupos, atendiendo a su origen geográfico, Perú, Ecuador, México y Mesoamérica. Este último grupo engloba a las entradas procedentes de Honduras, El Salvador, Nicaragua, Colombia y Costa Rica.

Las entradas colectadas en México (Figura 22A) proceden de diversos lugares generalmente cercanos a la costa, de climas subhúmedos (desde la entrada Cm31 a Cm51), a excepción de las entradas Cm28, Cm29 y Cm30 que proceden de regiones de clima seco como Sinaloa y Nuevo León. Las entradas de Mesoamérica (de la 52 a la 72) proceden de los países con clima tropical citados anteriormente (Figura 22B).

La mayoría de las entradas de Ecuador (Figura 22C) fueron recolectadas en la región amazónica, caracterizada por un clima tropical muy húmedo, como las procedentes de Napo (Ce73 a Ce80), Pastaza (Ce81, Ce82, Ce85 a Ce88), Morona Santiago (Ce91 a Ce99, Ce101 a Ce105 y Ce114), Guayas (Ce100) y Zamora Chinchipe (Ce107, Ce108, Ce110 a Ce113). También forman parte de este grupo dos entradas de la región de Tungurahua (Ce83 y Ce84), donde predomina el clima tropical muy húmedo, y entradas procedentes de regiones de clima seco-húmedo, próximas a la costa como Guayaquil (Ce89, Ce90), El Oro (Ce106) y Loja (Ce109).

Las entradas que conforman el grupo de Perú tienen una amplia distribución geográfica (Figura 22D). Las entradas del norte procedentes de Amazonas (Cp115) y San Martín (Cp116 a Cp131) se encuentran en zonas de clima tropical, al igual que la entrada Cp133 de Ucayali, situado hacia el sur. En cuanto al resto de entradas, procedentes del sur, se diferencian según su procedencia de regiones de clima frío o cálido. De clima frío son las de las provincias de Junín (Cp135), Cuzco (Cp137, Cp139 a Cp142), Ayacucho (Cp138 y Cp143) y Apurímac (Cp144 y Cp145), mientras que en Huánuco (Cp134), Madre de Dios (Cp136) y Puno (Cp147), el clima es cálido. En cuanto a las entradas procedentes de la costa (Cp132 de Santa Cruz y Cp148 de Arequipa), son de climas templados.

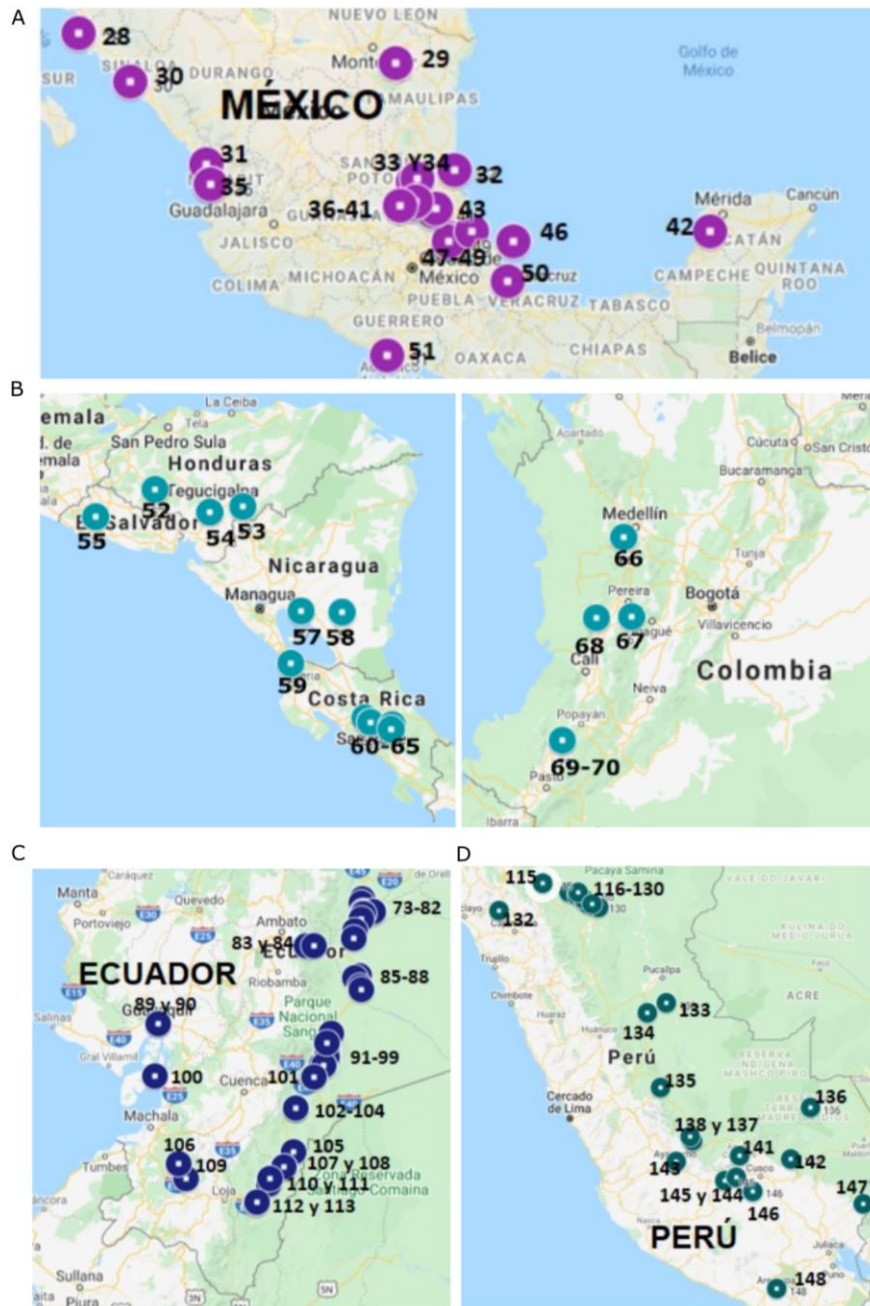


Figura 22. Distribución geográfica de las entradas de A. México, B. Mesoamérica, C. Ecuador y D. Perú de la especie *S. lycopersicum* var. *cerasiforme*.

3.3.2. Análisis de la variabilidad morfológica

Los análisis MANOVA y test de Fisher revelaron diferencias significativas entre los cuatro grupos geográficos establecidos para el taxón *Solanum lycopersicum* var. *cerasiforme*, principalmente en caracteres relacionados con flores y frutos (Anexo VIII). Algunos caracteres cuantitativos como son la anchura y el número de folíolos de la hoja, la altura hasta la última inflorescencia, el número de hojas entre inflorescencias, el número total de inflorescencias, el número de pétalos y sépalos de la flor, el número de lóculos del fruto, los parámetros L y b del color del fruto, el índice externo de la forma del fruto, la forma distal, la altura del hombro, el ángulo distal macro, el ángulo proximal macro, la forma obovoide y la protrusión del fruto no fue posible normalizarlos y se trataron como caracteres de tipo cualitativo.

En el análisis de componentes principales realizado con todos los caracteres convertidos en variables dummy, las CP1 y CP2 explicaron un 14,83 % y 8,52 % de la varianza observada, respectivamente. Los caracteres cuya contribución a la CP1 o CP2 fue superior a 0,17 se representan en la Figura 23 y en la Figura 24 se muestra la proyección de las entradas según estas componentes. El Anexo IX incluye la contribución de todos los caracteres a las tres primeras componentes principales.

La CP1 estuvo negativamente correlacionada con hojas de borde ondulado, flores con estigmas al mismo nivel que los estambres y frutos con un bajo número de lóculos, todos ellos caracteres que describen a las entradas procedentes mayoritariamente de México y Mesoamérica. Por el contrario, esta componente tuvo una mayor correlación positiva con mayor número de foliolos en las hojas, mayor número de pétalos de las flores, sépalos de mayor tamaño y mayor longitud de las anteras y frutos de mayor perímetro, peso y forma aplastada (perímetro, forma proximal y ángulos distal y proximal del fruto). Todos estos caracteres identifican en mayor medida a las accesiones de Perú y parte de Ecuador, distribuidas hacia la parte derecha del gráfico. Respecto a la CP2, los caracteres con una mayor contribución negativa fueron principalmente de flores y frutos: flores con pétalos largos, pétalos y sépalos anchos y pétalos muy curvados y frutos con bajo número de lóculos y alto ángulo proximal. Éstos son los caracteres más frecuentes en gran parte de las entradas de Ecuador, por lo que éstas se distribuyeron en la parte baja del gráfico. La CP2 estuvo positivamente correlacionada principalmente con caracteres de hoja: hojas altamente diseccionadas, con bordes aserrados y muchos foliolos.

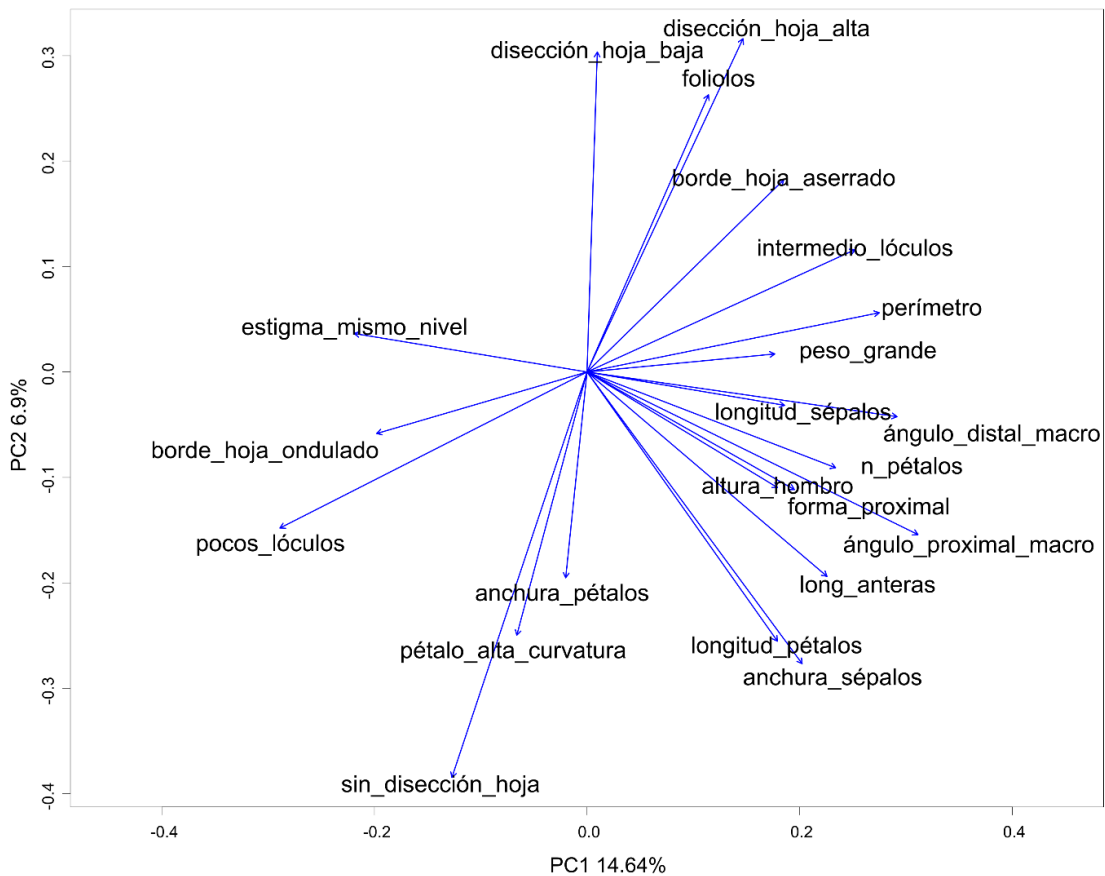


Figura 23. Caracteres con una contribución superior a 0,17 con las CP1 y CP2 del PCA realizado con *S. lycopersicum* var. *cerasiforme* de Perú, Ecuador, Mesoamérica y México.

La proyección de las entradas sobre las dos primeras componentes principales permitió visualizar su agrupación en tres grandes grupos según su origen geográfico (Figura 24). Sin embargo, la separación de los tres grupos no llegó a ser completa como consecuencia de la gran variabilidad morfológica que existe en esta especie a lo largo de toda su distribución geográfica.

El grupo 1 (grupo MEX_MESO) engloba la mayor parte de accesiones procedentes tanto de México como de Mesoamérica, junto con algunas entradas concretas de Ecuador y Perú. En el caso concreto de México, SLC se encuentra creciendo silvestre en ambientes que no han sido modificados por el hombre, además de en espacios próximos a asentamientos humanos o incluso se llega a comercializar, vendiendo sus frutos en mercados locales. En función de la procedencia y de el uso que se ha dado a las distintas entradas, existen una serie de diferencias notables en sus características morfológicas. Por ejemplo, existe una clara distinción entre las entradas de fruto pequeño o las entradas comercializadas: las entradas silvestres pertenecen al grupo 1 ya que son de frutos pequeños y redondos, mientras que las que se comercializan, tanto de fruto pequeño como más grandes y en ocasiones acostillados, se agrupan con las entradas de Perú en el grupo 3 (entradas Cm47 y Cm48).

El grupo 2 (grupo ECU) constituye la mayor parte de las accesiones de Ecuador. En este país se encuentra una enorme variabilidad morfológica para este taxón. Por un lado, nos encontramos con entradas que presentan una morfología intermedia entre SP y SCL (Napo, Ce74 y Ce77; Guayas, Ce100 y Zamora Chinchipe, Ce113), con frutos muy pequeños y que se agrupan con el grupo 1, debido a esta característica. Sin embargo, por otro lado, también existen entradas que se caracterizan por poseer frutos que son de mayor tamaño, con formas más aplastadas y en ocasiones con acostillado. Las accesiones con frutos de mayor tamaño corresponden a las entradas Ce89 y Ce90 que fueron colectadas en el mercado de Guayaquil. Estas entradas se agrupan, al igual que las de mercados de México, junto con las entradas de Perú en el grupo 3.

El grupo 3 (grupo PER) está integrado por la mayor parte de las entradas de Perú y las excepciones comentadas anteriormente acerca de las entradas procedentes de México y Ecuador. También las entradas de Perú albergan una notable variabilidad en cuanto a sus características morfológicas, en función de su origen geográfico. Por un lado, la mayor parte de entradas incluidas en el grupo 3 proviene de la provincia de San Martín y son las de frutos de mayor tamaño y acostillados, con un número de lóculos entre 3 y 10. Por otro lado, las entradas de otras procedencias son de frutos más pequeños y redondeados, localizándose la mayor parte de ellas en el grupo 1.

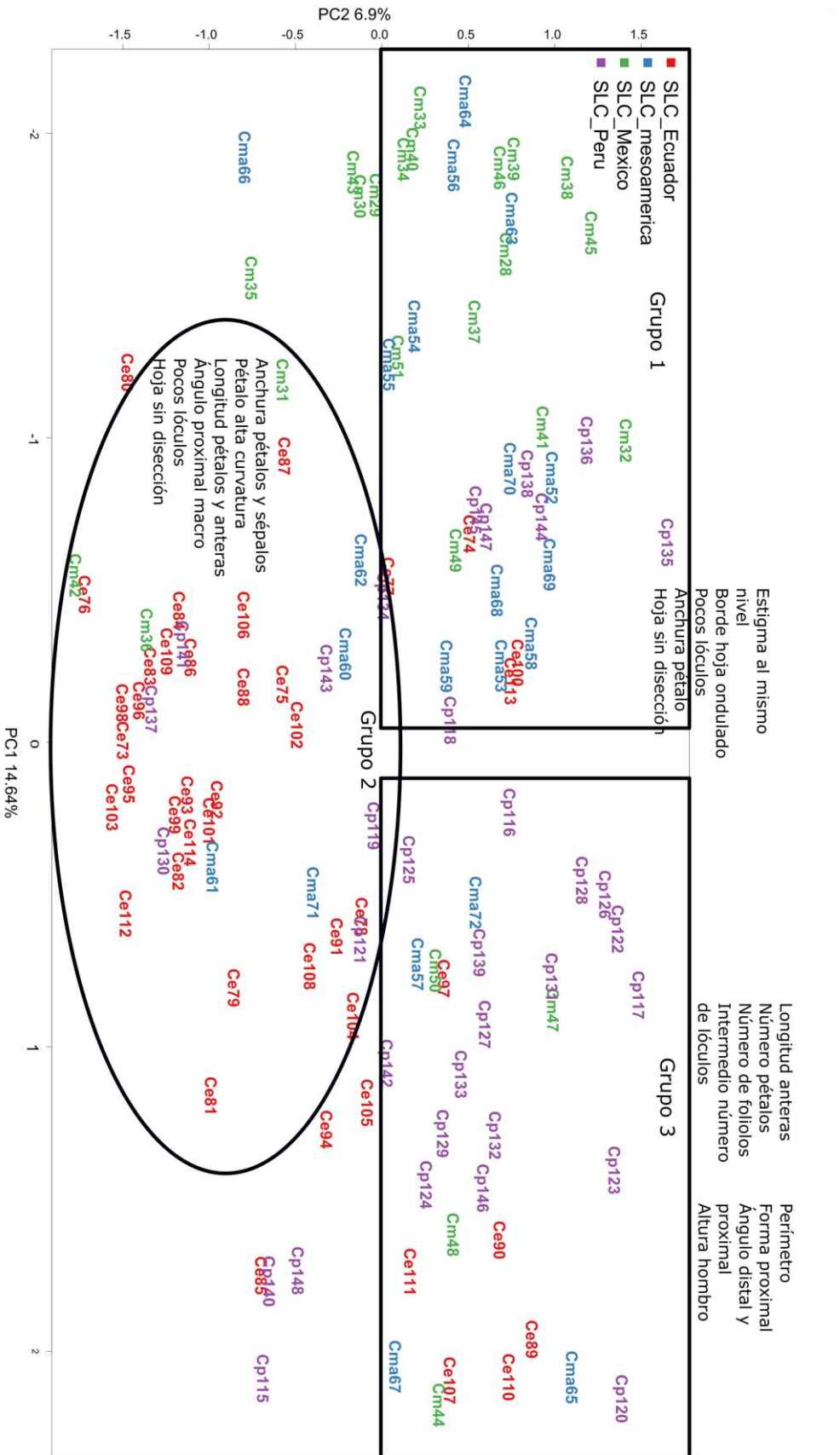


Figura 12. Análisis de componentes principales efectuado con los caracteres que presentaron diferencias significativas entre los grupos geográficos de *S. lycopersicum* var. *cerasiforme*.

Con base en los resultados anteriores, una clasificación más precisa hubiera sido agrupar las entradas procedentes de México y Mesoamérica, debido a las similitudes morfológicas entre ellas. Sin embargo, en los gráficos posteriores se mantendrá la agrupación de las entradas en los cuatro grupos establecidos según su procedencia geográfica, para representar los caracteres vegetativos, de inflorescencia, de flor y de fruto que mejor identifican cada grupo. Además, se incluyen imágenes representativas de dichos caracteres (Figura 25-28).

Los caracteres vegetativos que diferencian los grupos geográficos están relacionados con las hojas (Figura 25). El grupo SLC Ecuador y México presentaron menor número de folíolos en sus hojas y el mayor porcentaje de hojas con poca disección. Por el contrario, el grupo SLC Perú y Mesoamérica, presentaron altas proporciones de hojas con disección y la mayor proporción de borde de la hoja aserrado. Además, el tipo de borde fuertemente aserrado estuvo presente únicamente en algunas accesiones de Mesoamérica.

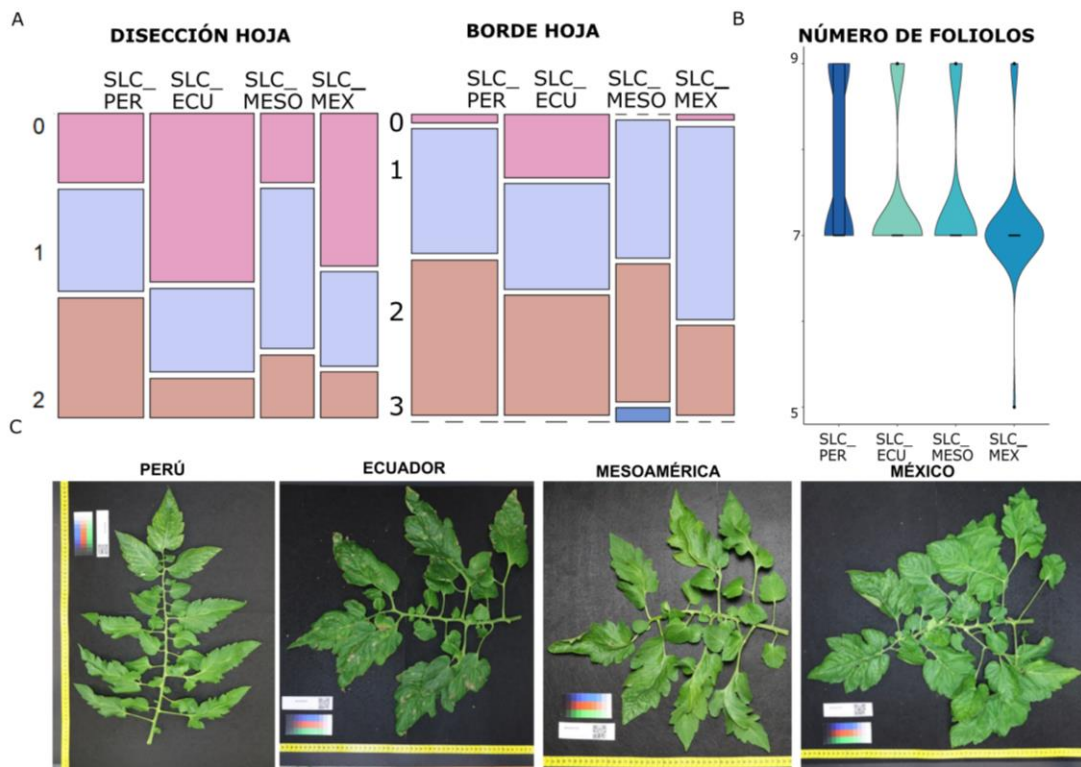


Figura 25. A. Gráfico mosaico que describen las hojas las entradas de *S. lycopersicum* var. *cerasiforme* procedentes de Ecuador, Perú, Mesoamérica y México. Disección de la hoja: 0: baja, 1: intermedia, 2: alta. Borde de la hoja: 0: entero, 1: ondulado, 2: aserrado, 3: fuertemente aserrado. **B. Gráfico violín con los datos cuantitativos originales para número de folíolos. SLC_ECU: Ecuador. SLC_PER: Perú. SLC_MA: Mesoamérica. SLC_MEX: México. **C. Hojas representativas de los cuatro grupos geográficos.****

Múltiples caracteres relacionados con las flores presentaron diferencias entre los cuatro grupos geográficos, caracterizando principalmente a las accesiones de Ecuador (Figura 26). La longitud de las anteras y el tamaño de pétalos y sépalos disminuyó progresivamente desde Ecuador hasta México. En Ecuador se encuentran las flores más pequeñas ya que, como se ha dicho anteriormente, se encuentran accesiones semejantes a la especie silvestre. Pese a que el estigma fue generalmente inserto o al mismo nivel, Perú y México presentaron menores proporciones de estigma exerto que Ecuador y Mesoamérica. Además, Ecuador es el único

grupo donde aparece el estigma fuertemente exerto, por la presencia de entradas en el grupo intermedias entre las especies SP y SLC. Además, estas entradas también serían las responsables de que en Ecuador aparezca el mayor porcentaje de pétalos altamente curvados. Por último, cabe destacar la mayor complejidad de las flores en las entradas de fruto grande de Perú y Ecuador.

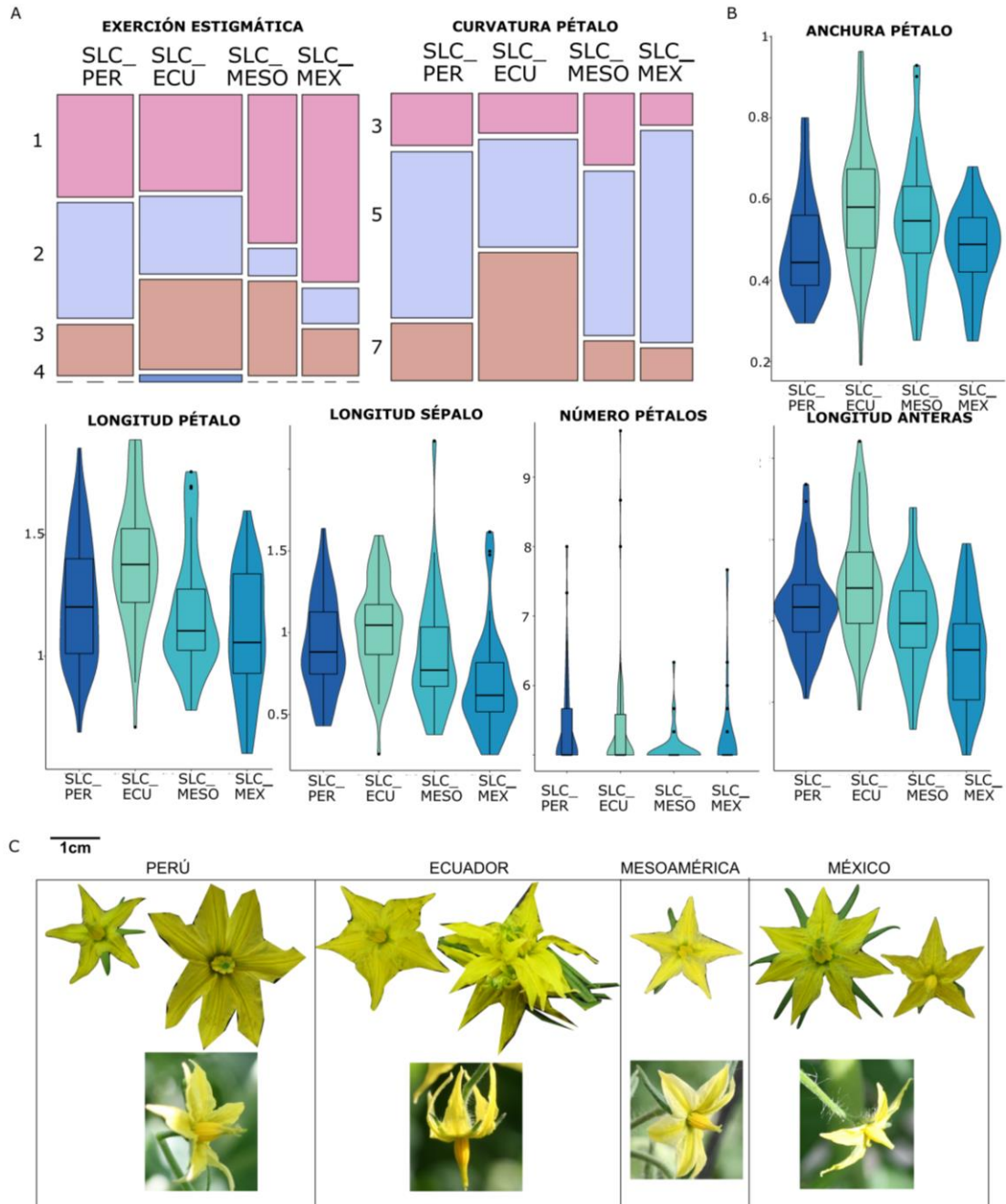


Figura 26. A. Gráfico mosaico que describe las principales diferencias en las flores entre *S. lycopersicum* var. *cerasiforme* procedentes de Ecuador, Perú, Mesoamérica y México. Exerción estigmática: 1: inserto, 2: mismo nivel, 3: ligeramente exerto, 4: altamente exerto. Curvatura del pétalo: 3: baja, 5: media, 7: alta. **B. Gráfico violín con los datos cuantitativos originales para número anchura y longitud del pétalo, número de pétalos y longitud de las anteras. SLC_ECU: Ecuador. SLC_PER: Perú. SLC_MA: Mesoamérica. SLC_MEX: México. **C. Características de las flores de los cuatro grupos geográficos. La referencia de tamaño solamente es aplicable a la fila superior de flores.****

Otra característica importante a la hora de diferenciar a los grupos geográficos ha sido el tamaño y la forma de los frutos (Figura 27 y 28). En cuanto a la forma de los frutos, los procedentes de México y Mesoamérica son los que presentaron una forma más redondeada, en comparación con los otros grupos geográficos. Esto queda representado por sus menores valores para forma distal, ángulos distal y proximal y altura del hombro. Por el contrario, las entradas de Perú y Ecuador presentaron una forma de fruto más aplastada, lo que es indicativo de un mayor acostillamiento del fruto, tal y como se observa en la altura del hombro. En cuanto al tamaño del fruto, representado por el perímetro y el peso, se pudo observar una reducción progresiva del mismo desde Perú a México. Por último, se observó la misma tendencia en el número de lóculos.

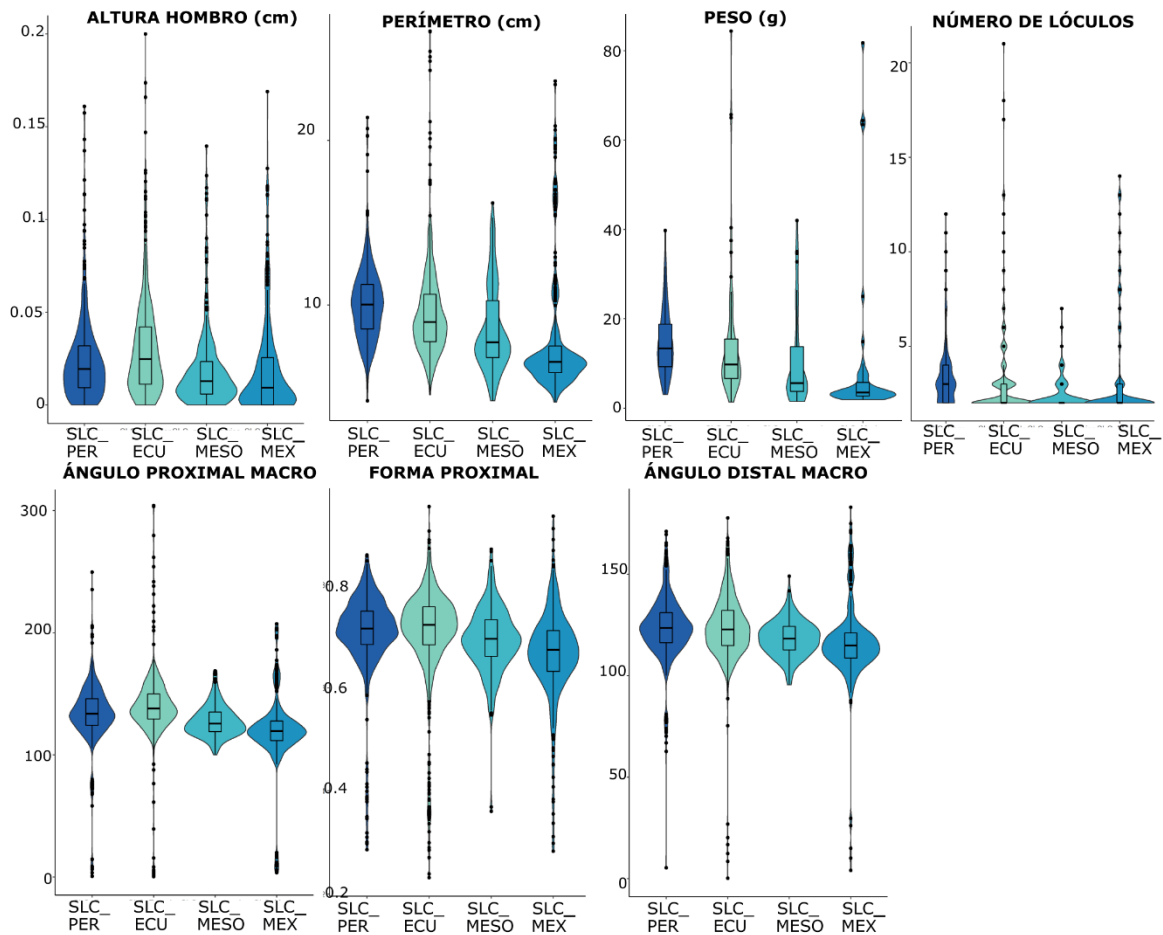


Figura 27. Gráfico violín con los datos cuantitativos originales que describen los frutos entre *S. lycopersicum* var. *cerasiforme* procedentes de Ecuador, Perú, Mesoamérica y México. SLC_ECU: Ecuador. SLC_PER: Perú. SLC_MA: Mesoamérica. SLC_MEX: México.

1 cm

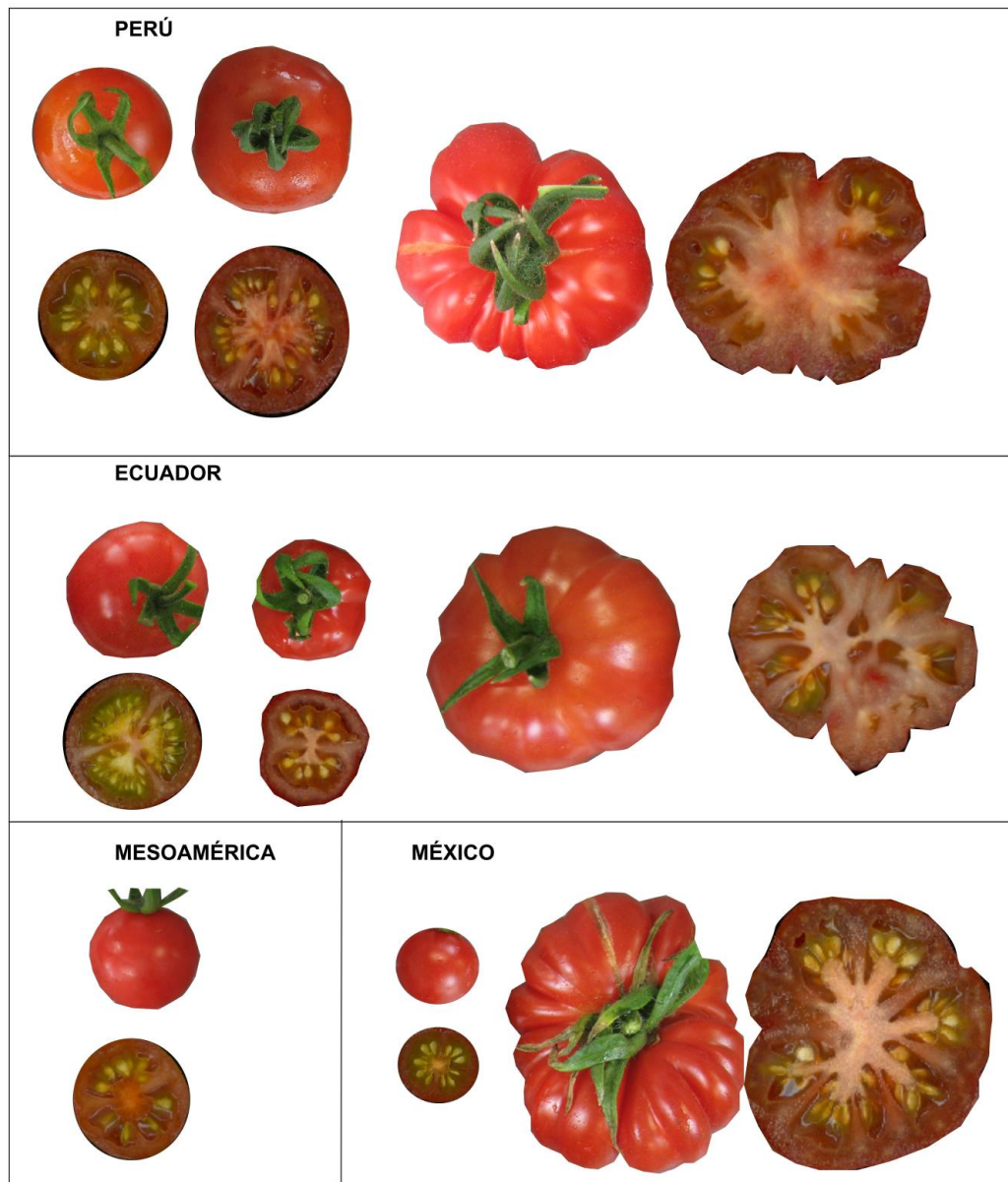


Figura 28. Principales diferencias en tamaño y forma del fruto en SLC.

4. Discusión

Explorar y caracterizar la diversidad existente en el centro de origen y difusión de los cultivos es crucial para su utilización en los programas de mejora. La colección de germoplasma estudiada en este trabajo ha sido seleccionada para representar la variabilidad existente tanto en el centro de origen como en el de domesticación del tomate cultivado. En ella se incluye una extensa representación de *Solanum lycopersicum* var *cerasiforme*, el ancestro del tomate cultivado, además de una selección de entradas de *Solanum pimpinellifolium*, abarcando gran parte del área de distribución de esta especie, y un conjunto de entradas de la especie cultivada procedentes de México. En paralelo al conocimiento de la variabilidad morfológica, es importante su conservación y una documentación exhaustiva y fácilmente accesible que posibilite su utilización.

S. pimpinellifolium, *S. l.* var *cerasiforme* y *S. l.* var *lycopersicum*

Considerando la relación entre los tres taxones estudiados, SP (especie silvestre más próxima al tomate cultivado), SLC (ancestro) y SLL (tomate cultivado), las diferencias morfológicas encontradas entre ellos deben deberse, en parte, a los procesos implicados en la domesticación a partir de SP (Meyer & Purugganan, 2013). En este proceso hay que considerar distintas causas de la variación, una de tipo de adaptación ecológica, al migrar las plantas a hábitats distantes, y otra debida al factor humano: la selección ejercida por el hombre. Los descriptores usados para llevar a cabo la caracterización morfológica han permitido recoger estos aspectos, identificando los descriptores más discriminantes entre taxones y separando el conjunto de entradas estudiadas en función de su taxón y, dentro de éste, en función de su origen geográfico.

Teniendo en cuenta los resultados del PCA efectuado, se ha podido constatar una gradación morfológica continua entre los tres taxones que coincide con estudios previos (Rick 1958; Rick 1958; Blanca et al. 2012). Uno de los cambios más destacables durante el proceso de domesticación fue el del tamaño y forma de los frutos. Los frutos de la especie silvestre SP son homogéneos, caracterizándose por su pequeño tamaño y forma redondeada. Tanto el tamaño como la diversidad de formas fue aumentando al pasar de SP a SLC y se incrementó todavía más al pasar de SLC a SLL. Lin et al. (2014) han determinado que, efectivamente, este cambio se produjo en dos fases, la primera durante el paso de SP a SLC y la segunda en el paso de SLC al tomate cultivado. De la gran cantidad de QTLs que determinan el tamaño del fruto (Grandillo, Ku, & Tanksley 1999; Huang & van der Knaap 2011; Chakrabarti et al. 2013; Barrantes et al. 2014; Munos et al. 2011), los QTLs implicados en el paso de SP a SLC fueron *fw1.1*, *fw5.2*, *fw7.2*, *fw12.1* y *lcn12.1* y los que determinaron el aumento todavía mayor en el tamaño de fruto desde SLC a SLL fueron *fw2.1*, *fw2.2*, *fw2.3*, *lcn2.1*, *lcn2.2*, *fw3.2*, *lcn3.1*, *fw9.3*, *lcn10.1*, *fw11.1*, *fw11.2* y *fw11.3* (Lin et al., 2014). Además del peso del fruto, se sucedieron cambios drásticos en su morfología durante el proceso de domesticación. Estos cambios han sido descritos como consecuencia de alelos derivados de los genes *fasciated* (*fas*), *sun*, *ovate* y *locule* (*lc*) (Rodríguez et al., 2011). Alelos alternativos para el gen *lc* ya han sido descritos en baja frecuencia en algunas poblaciones de SP, y más frecuentemente en poblaciones de SLC. En esta última especie también se han descrito variantes para los genes *fas* y *ovate*, mientras que para *sun* solamente han sido descritas en SLL (Blanca et al. 2015). De estos genes, *sun* y *ovate* regulan el alargamiento del fruto mientras que *fas* y *lc* regulan la forma aplastada del fruto y la fasciación y el número de lóculos, respectivamente, y determinan también un aumento del peso del mismo (Monforte et al., 2014). La aparición de estas mutaciones dio lugar a nuevos tipos de frutos que fueron seleccionados por el hombre dando lugar a la extensa variabilidad existente en SLC y SLL.

Otra clara diferencia entre SP y SLL es la producción, carácter de extrema importancia en mejora. Dos de los factores implicados en el aumento de la producción son el tipo de inflorescencia y el número de flores, y por tanto número de frutos, por inflorescencia. Se conocen varios genes implicados en el tipo de inflorescencia, como el gen *multiflora* y el *compound inflorescence* (*s*) (Lippman et al., 2008) y el *bifurcate flower truss* (*bif*) (Silva Ferreira et al., 2018). Lippman et al. (2008) propusieron que mutaciones en el gen *s* habrían dado lugar a la variación en la ramificación y abundancia de flores en las inflorescencias existente en el tomate domesticado. Por su parte, Zsögön et al. (2018) acometieron la “domesticación *de novo*” de *S. pimpinellifolium* mediante la edición genómica de varios genes, siendo el gen *multiflora* uno de los editados. Esta modificación originó inflorescencias

ramificadas a partir de las simples propias de la especie silvestre, con el consiguiente incremento en el número de frutos y la producción.

Una planta con un aumento considerable en el número y tamaño de los frutos requiere de otros cambios morfológicos en la estructura de la misma que le permitan soportar el peso de los frutos y llevar a buen término el desarrollo de los mismos, llegando a las producciones propias del tomate cultivado. Estos cambios morfológicos pasan por un aumento del vigor, que se traduce en un tallo más grueso y unas hojas más grandes capaces de aumentar la capacidad fotosintética y la producción de fotoasimilados. Los datos obtenidos en nuestra colección son consistentes con esta afirmación. El tallo de las plantas de la especie silvestre son de menor diámetro que los de SLC y SLL. De la misma manera, la especie SP es la que presenta las hojas más sencillas en cuanto a tipo, disección, borde de los folíolos y tamaño. La descripción de hojas simples dentro de la especie SP concuerda con la descripción previa hecha por Rick (1958) y Blanca et al. (2012), principalmente para las entradas de SP procedentes del norte de Perú. El tomate cultivado se caracteriza por una mayor complejidad y tamaño de sus hojas (Sifres et al., 2006) y la especie SLC por una mayor diversidad morfológica, concordante con la variabilidad observada en SLC de regiones andinas por Rick y Holle (1990).

Otro carácter para el que existe variación en el conjunto de entradas estudiado es la ejerción estigmática, que determina el paso de la alogamia propia de su ancestro silvestre a la autogamia de la especie cultivada. Este cambio se produjo ya en SP al desplazarse desde su centro de origen, en el norte de Perú, hacia el sur de Perú y el norte de Ecuador, ya que las máximas ejerciones se encuentran en las provincias del norte Piura, Lambayeque y La Libertad (Rick, Fobes, & Holle 1977; Rick, Holle, & Thorp 1978). Este patrón es común en otras especies y puede deberse a la ausencia de insectos polinizadores en áreas muy distantes al centro de origen o a las distintas condiciones ambientales, por ejemplo, una mayor humedad ambiental que dificulte el desprendimiento y liberación del polen disminuyendo la polinización cruzada. Posteriormente, durante el proceso de domesticación se produjo una modificación de la estructura floral y de la posición del estilo, que dio lugar a un cambio aún mayor en el sistema reproductivo de las plantas. Las plantas con flores conteniendo estigmas insertos habrían sido seleccionadas por el hombre para fijar los genotipos adaptados a nuevas condiciones o con caracteres de interés. Este hecho es concordante con nuestros resultados que describen a SP como la especie con la mayor ejerción estigmática, que fue disminuyendo progresivamente hacia SLC y SLL (Rick & Fobes 1975; Chen et al. 2007; Blanca et al. 2012).

Considerando el conjunto de entradas estudiado, podemos resaltar la gran variabilidad morfológica encontrada para la mayoría de los caracteres de planta, inflorescencia, flor y fruto surgida por la historia evolutiva del tomate y por la acción del hombre. Esta variabilidad puede ser aprovechada por los mejoradores, ya que el total de entradas descritas se ha secuenciado y son públicamente accesibles (Mata-Nicolás et al., 2020). De hecho, algunas de ellas han sido utilizadas en estudios relativos a la domesticación del tomate (Razifard et al., 2020), interacciones entre los genes que determinan la longitud y arquitectura de las inflorescencias (Soyk et al., 2017) y para el descubrimiento de nuevos genes (Gao et al. 2019).

S. pimpinellifolium

La especie SP se extiende por las regiones costeras del oeste de los Andes desde el norte de Ecuador hasta el sur de Perú. Las entradas incluidas en este estudio fueron colectadas

en ambientes muy distintos, desde las regiones con clima húmedo y templado o caluroso del norte y noroeste de Ecuador, hasta la región costera del norte de Perú, donde predomina el clima seco y árido. Durante la colonización de nuevos espacios hasta llegar a ocupar un rango de distribución tan amplio, debieron ocurrir procesos de diferenciación en los caracteres vegetativos y reproductivos de las plantas con el objeto de adaptarse a los distintos ambientes, dándose también una divergencia genética. En esta especie, la relación entre el clima y la diferenciación genética ha sido demostrada por diversos autores (Nakazato, Bogonovich, & Moyle 2008; Zuriaga et al. 2009; Blanca et al. 2012; Gibson & Moyle 2020). Nakazato, Bogonovich y Moyle (2008) llevaron a cabo un estudio con entradas de Perú de diferentes procedencias de las laderas oeste de los Andes y de las regiones adyacentes costeras de Perú, demostrando que los factores ambientales son los causantes de la diferenciación fenotípica que lleva a la adaptación a los diferentes ambientes y estreses abióticos. El rango de distribución del conjunto de entradas estudiadas en este trabajo fue bastante mayor que el incluido en el trabajo de Nakazato, Bogonovich, & Moyle (2008), abarcando Ecuador y Perú. La caracterización fenotípica realizada nos ha permitido constatar la existencia de una amplia variabilidad fenotípica en esta especie para caracteres vegetativos, de inflorescencias, flores y frutos. Las diferencias morfológicas más notables se dieron entre las entradas colectadas en la costa oeste de Ecuador, con clima húmedo y caluroso, y las del norte de Perú, con clima marcadamente seco.

Rick, Fobes y Holle pusieron ya de manifiesto en 1977 la baja uniformidad de esta especie, que entonces se creía más uniforme. Estos autores basaron su estudio en una muestra de 43 entradas, tres de las cuales eran de Ecuador y el resto de Perú. En nuestro estudio hemos incrementado el número de entradas de Ecuador, así como sus lugares de procedencia, y se han caracterizado para un conjunto de descriptores mucho más amplio, lo que ha permitido descubrir variabilidad en caracteres antes no considerados. Así, existe una serie de caracteres vegetativos que han demostrado una elevada variabilidad y, además, han sido caracteres discriminantes entre las poblaciones de Perú y Ecuador. Estos caracteres incluyen el contenido en antocianinas del tallo, el número de foliolillos de las hojas y el color y borde los folíolos. Las hojas de las plantas de Ecuador fueron de mayor tamaño, con mayor número de foliolillos y bordes más ondulados, más similares a las hojas estándar del tomate cultivado, en contraposición a las de Perú que fueron de menor tamaño, más simples (menor número de foliolillos) y de borde entero, la típica hoja “pimpinellifolium” indicada en los descriptores de tomate del IPGRI (1996).

El tamaño de las inflorescencias fue otro carácter que varió marcadamente en la colección estudiada y entre las entradas de Ecuador y Perú. Las entradas de Ecuador presentaron inflorescencias cortas, sin brotes ni hojas, mientras que en Perú tuvieron un mayor número de flores y mayor longitud, terminando con cierta frecuencia con hojas o brotes. El número de flores en la inflorescencia es un carácter de interés en el desarrollo de cultivares de tomate tipo *cherry*, donde tanto el tamaño pequeño del fruto como la longitud de la inflorescencia son caracteres clave. Disponer de entradas con una gran diversidad para estos caracteres (se encontraron entradas con más de 100 flores por inflorescencia) permite su utilización en los programas de mejora para poder aumentar el rendimiento de estas variedades. Se constató también la gran influencia del ambiente en la expresión de este carácter, ya que este elevado número de flores se observó principalmente en cultivo bajo invernadero, disminuyendo sensiblemente cuando las mismas entradas se cultivaron al aire libre en otros experimentos realizados por nuestro grupo (datos no publicados). Aunque se conoce que el gen *lin* es el responsable de la longitud de la inflorescencia (Lin et al., 2014), sería conveniente profundizar en el estudio de su expresión en diferentes ambientes. Además,

también sería de interés describir su posible interacción con otros genes implicados en el desarrollo de la inflorescencia y en la longitud de los sépalos, carácter variable en la colección, como son los genes *s*, *s2*, *j2* y *ej2w*. La importancia de la influencia ambiental es un factor a tener en cuenta ya que la mayor parte del tomate tipo *cherry* se produce en invernadero, por lo tanto, permitiría maximizar esta característica. La elección de las entradas de SP que presenten inflorescencias tan largas podría tener una repercusión deseable en los programas de mejora encaminados a la obtención de cultivares de tomate tipo *cherry*.

El tamaño de los pétalos y la tonalidad amarilla de los mismos se redujo en Ecuador, y también la ejerción estigmática y longitud de las anteras. Esta variabilidad en función del origen geográfico para el grado de ejerción estigmática ya fue puesta de manifiesto por Rick y Fobes (1975); Rick, Fobes y Holle (1977); Rick (1978) o Rick, Holle y Thorp (1978) quienes describieron las mayores ejerciones en las entradas de las regiones de Piura y Lambayeque, coincidiendo este hecho con otras observaciones (Widrlchner, 1987). Es sabido que la posición del estigma está relacionada con el grado de alogamia de la especie, determinando de esta manera su sistema reproductivo. Rick, Fobes y Holle (1977) y Georgiady, Whitkus y Lord (2002) describieron un mayor grado de alogamia de la especie cuanto mayor es su ejerción estigmática. Estos patrones coinciden con las diferencias observadas en nuestros grupos geográficos y sugieren que SP sufrió una inserción progresiva del estigma conforme migraba fuera de su centro de origen, favoreciendo así su autopolinización. Este proceso concuerda con la hipótesis de Stebbins (1957) quien propuso que las especies autógamias habrían surgido a partir de sus ancestros alógamos. La autogamia supondría una ventaja en ausencia de insectos polinizadores en los nuevos hábitats conquistados y favorecería la fijación de caracteres que permitirían la adaptación de la especie a las nuevas condiciones ambientales.

Las poblaciones de Ecuador y Perú difirieron también en las características de sus frutos. En Perú fueron de menor tamaño, más redondeados y en ocasiones de una tonalidad roja más intensa debida al contenido en licopeno. La especie SP ha sido empleada en mejora de la calidad debido al elevado contenido en licopeno que presentan sus frutos, ya que es uno de los compuestos que determinan el valor nutricional del tomate (Bergougnoux, 2014). Este compuesto se sintetiza de forma natural, en mayor cantidad en frutos de color rojo intenso, y es beneficioso para la salud, ya que presenta propiedades antioxidantes (Mazidi et al., 2020). Un alelo de la mutación *Beta*, *og^c*, ha sido empleado para el desarrollo de variedades comerciales con un alto contenido en licopeno (Baranski et al., 2016; Kinkade & Foolad, 2013). Además, el color naranja de las anteras ha sido descrito como indicativo de mayores acumulaciones de licopeno en el fruto (Zsögön et al., 2018). En nuestra colección, se ha descrito variabilidad dentro de esta especie para el color de las anteras, tanto en las entradas procedentes de Ecuador como en las de Perú. Sería de interés realizar una cuantificación más precisa del contenido en licopeno en las entradas que presentaron esta característica.

En la colección estudiada también se identificaron tres entradas con frutos amarillos, una de ellas procedente de Ecuador (Azuay) y dos de Perú (Lambayeque). Se han descrito mutantes de color amarillo que presentan diferencias en la composición de carotenoides con respecto a los frutos de color rojo. Estos mutantes han sido empleados para el desarrollo de cultivares comerciales con frutos de color amarillo o anaranjado. Así, el mutante *yellow fruit flesh* o *r* presenta color amarillo por ausencia de la síntesis de carotenos, los mutantes *Beta* y *Delta* incrementan los niveles de β -caroteno y δ -caroteno, respectivamente, a expensas del licopeno y los tomates portadores de la mutación *tangerine* (*t*) contienen bajos contenidos de β -caroteno y licopeno (Yoo et al., 2017). Además de la presencia de frutos de color amarillo,

descrita en Perú (Rick, Fobes, & Holle 1977), también se ha descrito la presencia de piel incolora en los frutos (Blanca et al. 2012). La asociación entre la presencia de la mutación *yellow fruit flesh*, responsable de frutos amarillos o anaranjados, y la mutación *colorless fruit epidermis*, responsable de la ausencia del color de la piel del fruto, ya ha sido descrita previamente en entradas de esta especie procedentes de Lambayeque (Sifres et al., 2006), dando lugar a tonalidades diversas dentro del color amarillo. Por último, tal y como expuso Rick, Fobes y Holle (1977), la especie SP es variable en el contenido de antocianinas del tallo y también en los frutos en función de su origen geográfico. El interés en este compuesto radica en su función antioxidante, anticancerígena o antiinflamatoria (Al-Sane, Povero y Perata 2011) y se han descrito una serie de mutantes con un alto contenido en el mismo como *Abg* o *Aft*.

Aunque en este trabajo no hemos abordado el estudio de la adaptación a estreses bióticos y abióticos de SP, numerosos autores han abordado estos trabajos, encontrando esta especie como una de las más utilizadas en mejora por su resistencia a diversos patógenos y su adaptación a estreses abióticos como déficit hídrico y salinidad (Rao et al., 2013, 2015). Su proximidad al tomate cultivado, la adaptación a diferentes tipos de estreses y la gran variabilidad morfológica para caracteres de interés agronómico han ocasionado que algunos autores hayan planteado su “domesticación *de novo*”, mediante la edición genética de genes que controlan aspectos clave en mejora. Zsögön et al. (2018), mediante la modificación de seis *loci* implicados en caracteres de planta, inflorescencia y fruto (hábito de crecimiento, *self pruning sp*; forma del fruto, *ovate*; tamaño del fruto, *fasciated*, *fruit weight 2.2*; número de frutos, *multiflora*; y calidad nutricional, *lycopeno beta ciclasa*), han conseguido modificar el desarrollo vegetativo de la planta, aumentar en número y tamaño de los frutos, modificar su forma y aumentar el nivel de licopeno, lo que les llevó a conseguir aumentos significativos en la producción. Las técnicas de edición genética permitirán aprovechar de maneras distintas a las habituales esta especie como un recurso para la mejora. Para ello, el conocimiento detallado de la variación en la morfología de la planta resulta de especial interés, a fin de seleccionar las entradas con un mayor número de características favorables.

La variabilidad encontrada para algunos caracteres ha permitido separar las entradas incluidas en este estudio según su origen geográfico, encontrando una clara tendencia para ciertos alelos en determinados lugares (es el caso de las hojas más complejas y grandes y frutos de mayor tamaño en Ecuador, entre otros) pero también hay que destacar la existencia de variabilidad para muchos caracteres a lo largo de todo el rango de distribución, no encontrando una relación clara entre las distintas opciones para cada descriptor y su procedencia geográfica. Este hecho es de importancia para los mejoradores y enfatiza la necesidad de realizar extensivos trabajos de caracterización, dentro y entre poblaciones, para identificar caracteres específicos de interés.

S. lycopersicum var. cerasiforme

A diferencia de otras especies relacionadas con el tomate, SLC no ha sido extensivamente utilizada en mejora. La similitud de sus frutos con los del tipo *cherry* comercial en ocasiones no ha contribuido de forma positiva a profundizar en el conocimiento de esta variedad. Esta falta de conocimiento ha dado lugar a la inclusión en estudios de variedades tipo *cherry* comerciales en lugar de la verdadera variedad *cerasiforme*, dando lugar a malinterpretaciones sobre su relación con el tomate cultivado (Ranc et al., 2008). En este trabajo hemos abordado la caracterización de 121 entradas de zonas donde se produjo la domesticación del tomate como son Perú, Ecuador y Mesoamérica.

La variabilidad para todos los caracteres evaluados ha sido muy elevada, existiendo un continuo entre SP y SLL para muchos de ellos, lo que hace difícil establecer los límites que la diferencian de SP y de la morfológicamente diversa especie cultivada. Para entender cuáles son las causas de esta variación es preciso considerar la historia evolutiva y la situación actual de esta variedad, descritas ya en otros trabajos previos (Blanca et al. 2012; 2015). SLC se halla extendida en la actualidad principalmente por la vertiente este de los Andes en Ecuador y Perú y por las zonas tropicales y subtropicales de gran parte del mundo. Aunque puede encontrarse de forma silvestre, lo habitual es encontrarla asociada a los asentamientos humanos. Muchas de las colectas se han llevado a cabo en los patios traseros de las casas de los agricultores, donde se deja crecer sin realizar un cultivo consciente, beneficiándose de la recolección y consumo de sus frutos. Además, en ocasiones, se ha realizado un cultivo consciente, llegando a comercializar sus frutos en mercados locales. Se unen pues, en esta variedad, varias causas que pueden dar lugar a la variabilidad encontrada en este estudio: las características desarrolladas cuando crece silvestre para adaptarse a los factores ambientales de su entorno, la variabilidad introgresada desde SP en los cruzamientos naturales en áreas simpátricas con esta especie y la acción del hombre que, de forma más o menos deliberada, ha ido ejerciendo para determinados caracteres de interés. Estos factores pueden no darse en la misma medida en el amplio rango de distribución de SLC, lo que ocasiona diferencias más o menos claras entre orígenes, aunque sin una separación total.

En el PCA realizado con las entradas de SLC los caracteres que más contribuyeron negativamente a la primera componente principal fueron los relacionados con hojas de borde ondulado, flores con estigmas al mismo nivel que los estambres y frutos con un bajo número de lóculos, mientras que los que más influyeron positivamente fueron hojas con mayor número de folíolos, flores con mayor número de pétalos y mayor longitud de las anteras, sépalos de mayor tamaño y frutos de mayor perímetro y peso y de forma aplastada (forma proximal y ángulos distal y proximal del fruto). De acuerdo con estos caracteres, las entradas quedaron separadas, empezando por las de México y Mesoamérica a la izquierda del gráfico, seguidas de las de Ecuador y finalmente las de Perú, aunque esta separación no fue completa.

Diversos autores han demostrado (Jenkins 1948; Rick & Holle 1990; Blanca et al. 2012; Blanca et al. 2015; Razifard et al. 2020) el papel fundamental de México en la domesticación del tomate. Por ello, varios investigadores de este país han dedicado numerosos estudios al conocimiento de la distribución y usos de esta variedad (Rodríguez Guzmán et al. 2009; Álvarez-Hernández, Cortez-Madrigal, & García-Ruiz 2009; Rodríguez-Guzmán et al. 2018; Délices et al. 2019). Los rangos climáticos de distribución de SLC en la región del occidente de México muestran que esta especie se distribuye entre 7 y 2000 m de altitud, de 495 a 1591 mm de lluvia anual, de 7,1 a 21,6 °C de temperatura mínima media anual, 22,6 a 38,4 °C de temperatura máxima media anual, y entre 15,8 y 28,1 °C de temperatura media anual. Presenta amplia variabilidad en su capacidad de adaptación climática, al crecer en amplitud de climas tropicales y subtropicales, con preferencia por tipos subhúmedos cálidos (Rodríguez Guzmán et al., 2009). En diversas regiones se encuentra creciendo de forma silvestre, siendo en estos casos los frutos de tamaño muy pequeño, comparable a los de SP, aunque el resto de las características son más similares a las de SLC. Además, como se ha comentado, SLC se cultiva en algunas ocasiones y esto conlleva una selección por parte del agricultor hacia frutos de mayor tamaño y variaciones en la forma. Dos de las entradas con estas características fueron la Cm47 y Cm48, vendidas en mercados locales y que, debido al mayor tamaño de sus frutos y cierto acostillado, se unieron a las de Perú en la parte derecha del PCA, explicando este hecho la diversidad de formas y tamaños encontradas en las entradas de este país.

Las entradas de Ecuador analizadas en este trabajo proceden de varias provincias situadas principalmente en la ladera este de los Andes. Morfológicamente, gran parte de las entradas colectadas en estas provincias presentan características intermedias entre SP y SCL. Así, son frecuentes las plantas sin pilosidad en el tallo y con flores presentando exéresis estigmática y disposición de los pétalos curvados hacia atrás, características típicas de SP, pero con frutos de tamaño variable, desde los muy pequeños que son comparables a los de SP hasta los más grandes y acostillados. Estas formas intermedias fueron encontradas en las provincias de Napo, Pastaza, Morona Santiago y Zamora Chinchipe, es decir, a lo largo de toda la vertiente oriental ecuatoriana de los Andes. La existencia de cruzamientos entre SP y tomate fue demostrada ya por Rick (1958) y Blanca et al. (2012), demostrando estos últimos autores que compartían parte de su genoma con SP en un experimento en el que se genotiparon con la plataforma SolCAP (7.414 marcadores SNPs). Además de estas formas intermedias, existen otras de características variadas para caracteres de fruto, principalmente en cuanto a tamaño, forma, acostillado y color, apareciendo frutos de color rosado debido a la ausencia de flavonoides en la piel del fruto (Adato et al., 2009; Ballester et al., 2010) y una alta frecuencia de frutos con el contenido de los haces vasculares visibles. Finalmente, las entradas PI 129028 y PI 129033 provienen del mercado de Guayaquil, donde fueron adquiridas en el año 1938. La selección realizada por el hombre en estos materiales dio lugar a un aumento considerable del tamaño de los frutos y aparición de acostillado en los mismos, siendo similares a los del tomate cultivado. De hecho, fueron considerados como SLL, sin embargo, el análisis molecular reveló que su genoma es similar al de SLC (Blanca et al. 2012). Aunque SP se distribuye en Ecuador por la vertiente oeste de los Andes, existen corredores que atraviesan la cordillera y pueden favorecer el acercamiento entre las dos especies. Por ello, SLC en Ecuador es una variedad sometida a introgresiones desde SP debido a la proximidad de su área de distribución y sometida además a la acción del hombre en las milpas, donde es tolerada y consumida. El cultivo local que se realiza en determinadas ocasiones lleva a una selección todavía mayor, aumentando de esta manera su variabilidad morfológica hasta completar el rango de variación desde las características de SP hasta las del tomate cultivado. Estos cambios en el tamaño y forma del fruto han sido descritos como uno de los principales cambios que tuvieron lugar en el proceso de domesticación y Blanca et al. (2015) han demostrado la existencia de las mutaciones *fas*, *sun*, *ovate* y *lc* en la colección de SLC estudiada, demostrando que, efectivamente, contribuyeron en la diversificación del tomate.

Al igual que en Ecuador, SCL se encuentra distribuida en Perú a lo largo de toda la vertiente este de los Andes. Existe también una marcada diversidad para caracteres de planta y fruto. Sin embargo, la abundancia de formas intermedias detectada en Ecuador no se da en Perú. Únicamente en la provincia de Puno, en el valle del río Urubamba, se detectaron hibridaciones naturales que dieron lugar a formas intermedias, con plantas de características vegetativas más similares a SLC, pero con frutos de muy pequeño tamaño similares a los de SP (Cuartero et al., 1985). A pesar de tratarse de formas intermedias, presentan marcadas diferencias con las de Ecuador. Esto puede ser debido a las diferencias entre las entradas de SP que hibridaron con SLC de estas procedencias (Blanca et al. 2012) y a los procesos que siguieron tras la hibridación. Según las características morfológicas apreciables, las formas intermedias de Ecuador son más similares a SP, mientras que las de Puno lo son a SLC. Además, las diferencias ambientales de estas áreas en Ecuador y Perú son marcadamente distintas, habiendo en Puno un clima frío (6 °C – 12 °C) y subhúmedo (500-1000 ml), mientras que en las provincias de Ecuador el clima es templado y húmedo. Como se ha comentado, Nakazato, Bogonovich y Moyle (2008) demostró que los órganos de la planta se

adaptan a las condiciones ambientales, exacerbando por tanto las diferencias morfológicas entre las entradas de ambos orígenes.

Además de las entradas de Puno, más al norte, en la provincia de San Martín, se encuentra una diversidad para características de fruto muy apreciable. Este hecho fue observado ya por Rick y Holle en 1990. Estos autores estudiaron la variabilidad morfológica e isoenzimática para un amplio conjunto de entradas colectadas en esta área, determinando la existencia de variabilidad para pilosidad de la planta, amargor del fruto, forma, tamaño, acostillado y color de la piel, muchos de estos caracteres de control monogénico. Lo mismo fue observado en posteriores expediciones realizadas por nuestro grupo. (Rick y Holle (1990) atribuyeron esta variabilidad a la evolución de este taxón durante un largo periodo de tiempo, durante el cual ha estado sometido a diferentes fuerzas ejercidas por el ambiente, la acción del hombre y las posibles hibridaciones con tomate cultivado.

La colección presentada en esta tesis recoge una alta variabilidad morfológica y podría actuar como una valiosa fuente de genes de interés para los programas de mejora del tomate dirigidos a la mejora de la producción o calidad. En cuanto a la mejora de la producción, hemos encontrado en la colección entradas con alto número de flores por inflorescencia, llegando a alcanzar más de 100 flores, así como una tremenda variabilidad para la posición del estigma, carácter floral que afecta al cuajado del fruto y, por tanto, a la producción final de la planta (Srinivas, 2001). La calidad es un concepto que engloba tanto la apariencia externa del fruto (tamaño, forma o color) como la calidad organoléptica. La amplia gama de color de fruto que ha presentado nuestra colección es una característica de interés para el desarrollo de cultivares que satisfagan las demandas del consumidor, ya que se ha demostrado en diversos estudios que el color del fruto influye en su aceptabilidad y preferencia por parte de los mismos (Francis 1995; Causse et al. 2003; Causse et al. 2010). El color del fruto se debe principalmente a su contenido en carotenoides y la proporción de los distintos tipos de carotenoides. Tanto el color del fruto como el color de las anteras han sido descritos como parámetros de interés ya que existen mutantes de color amarillo que presentan diferencias en la composición de carotenoides con respecto a los frutos de color rojo (Yoo et al., 2017) y el color naranja de las anteras puede actuar como indicativo de mayores acumulaciones de licopeno en el fruto (Zsögön et al., 2018). Por último, también es importante considerar el enorme potencial que podría tener esta colección para el desarrollo de cultivares resistentes. Esta colección engloba entradas que cubren un amplio rango de ambientes, desde las zonas desérticas costeras de Perú y de Sinaloa en México, hasta las selvas tropicales húmedas de las laderas este de los Andes en Ecuador, con pluviometrías que pueden alcanzar los 4500 mm/año. Esto supone un gran abanico de adaptaciones a diferentes condiciones y posiblemente patógenos como describió Rick (1973), quien describió cómo SLC es capaz de sobrevivir en estos ambientes, donde el tomate cultivado es víctima de enfermedades fúngicas y bacterianas prevalentes en estas condiciones. Pese a que tampoco se ha llevado a cabo un cribado específico de la colección para encontrar resistencia a plagas o enfermedades, sí que se ha observado una enorme variabilidad en la presencia de diversos tipos de tricomas en las accesiones (véase Capítulo 4). La presencia de tricomas en la planta es un mecanismo que puede conferir resistencia a insectos de diversas maneras puesto que afectan a su movimiento, facilitan su envenenamiento o dificultan la ovoposición (Stout et al., 2017). En definitiva, nos encontramos ante una colección de germoplasma de enorme potencial para emplear en programas de mejora de tomate, tanto por la variabilidad presente en los caracteres evaluados como la posible existencia de caracteres de interés aún por evaluar.

BIBLIOGRAFÍA

- Adato, A., Mandel, T., Mintz-Oron, S., Venger, I., Levy, D., Yativ, M., Domínguez, E., Wang, Z., De Vos, R. C. H., Jetter, R., Schreiber, L., Heredia, A., Rogachev, I., & Aharoni, A. (2009). Fruit-surface flavonoid accumulation in tomato is controlled by a SLMYB12-regulated transcriptional network. *PLoS Genetics*, 5(12). <https://doi.org/10.1371/journal.pgen.1000777>
- Al-Sane, K. O., Povero, G., & Perata, P. (2011). Anthocyanin tomato mutants: Overview and characterization of an anthocyanin-less somaclonal mutant. *Plant Biosystems*, 145(2), 436–444. <https://doi.org/10.1080/11263504.2010.548914>
- Álvarez-Hernández, J. C., Cortez-Madriral, H., & García-Ruiz, I. (2009). EXPLORACIÓN Y CARACTERIZACIÓN DE POBLACIONES SILVESTRES DE JITOMATE (SOLANACEAE) EN TRES REGIONES DE MICHOACÁN, MÉXICO. *Polibotánica*, 28, 139–159.
- Ballester, A. R., Molthoff, J., de Vos, R., Hekkert, B. te L., Orzaez, D., Fernández-Moreno, J. P., Tripodi, P., Grandillo, S., Martin, C., Heldens, J., Ykema, M., Granell, A., & Bovy, A. (2010). Biochemical and molecular analysis of pink tomatoes: Deregulated expression of the gene encoding transcription factor SLMYB12 leads to pink tomato fruit color. *Plant Physiology*, 152(1), 71–84. <https://doi.org/10.1104/pp.109.147322>
- Baranski, R., Goldman, I., Nothnagel, T., & Scott, J. W. (2016). Improving Color Sources by Plant Breeding and Cultivation. In *Handbook on Natural Pigments in Food and Beverages: Industrial Applications for Improving Food Color* (pp. 429–472). Elsevier Inc. <https://doi.org/10.1016/B978-0-08-100371-8.00019-1>
- Barrantes, W., Fernández-del-Carmen, A., López-Casado, G., González-Sánchez, M. Á., Fernández-Muñoz, R., Granell, A., & Monforte, A. J. (2014). Highly efficient genomics-assisted development of a library of introgression lines of *Solanum pimpinellifolium*. *Molecular Breeding*, 34(4), 1817–1831. <https://doi.org/10.1007/s11032-014-0141-0>
- Bauchet, G., & Causse, M. (2012). Genetic Diversity in Tomato (*Solanum lycopersicum*) and Its Wild Relatives. *Genetic Diversity in Plants*, May. <https://doi.org/10.5772/33073>
- Bergougnoux, V. (2014). The history of tomato: From domestication to biopharming. *Biotechnology Advances*, 32(1), 170–189. <https://doi.org/10.1016/j.biotechadv.2013.11.003>
- Blanca, J., Cañizares, J., Cordero, L., Pascual, L., Díez, M. J., & Nuez, F. (2012). Variation Revealed by SNP Genotyping and Morphology Provides Insight into the Origin of the Tomato. *PLoS ONE*, 7(10). <https://doi.org/10.1371/journal.pone.0048198>
- Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., Francis, D., Causse, M., Knaap, E. Van Der, & Cañizares, J. (2015). Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics*, 257. <https://doi.org/10.1186/s12864-015-1444-1>
- Caicedo, A. L., & Schaal, B. A. (2004). Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. *Molecular Ecology*, 13(7), 1871–1882.

<https://doi.org/10.1111/j.1365-294X.2004.02191.x>

- Causee, M., Buret, M., Robini, K., & Verschave, P. (2003). Inheritance of Nutritional and Sensory Quality Traits in Fresh Market Tomato and Relation to Consumer Preferences. *Journal of Food Science*, 68(7), 2342–2350.
- Causee, M., Friguier, C., Coiret, C., Lépiciier, M., Navez, B., Lee, M., Holthuysen, N., Sinesio, F., Moneta, E., & Grandillo, S. (2010). Consumer Preferences for Fresh Tomato at the European Scale: A Common Segmentation on Taste and Firmness. *Journal of Food Science*, 75(9). <https://doi.org/10.1111/j.1750-3841.2010.01841.xi>
- Chakrabarti, M., Zhang, N., Sauvage, C., Muños, S., Blanca, J., Cañizares, J., Diez, M. J., Schneider, R., Mazourek, M., McClead, J., Causee, M., & van der Knaap, E. (2013). A cytochrome P450 regulates a domestication trait in cultivated tomato. *Proceedings of the National Academy of Sciences of the United States of America*, 110(42), 17125–17130. <https://doi.org/10.1073/pnas.1307313110>
- Chen, K. Y., Cong, B., Wing, R., Vrebalov, J., & Tanksley, S. D. (2007). Changes in regulation of a transcription factor lead to autogamy in cultivated tomatoes. *Science*, 318(5850), 643–645. <https://doi.org/10.1126/science.1148428>
- Cuartero, J., Gomez-Guillamon, M. L., & Diaz, A. (1985). Catalog of collections of *Lycopersicon* from Peruvian Central Areas. *TGC Report*, 35, 32–35.
- Délices, G., Leyva Ovalle, O. R., Mota-Vargas, C., Núñez Pastrana, R., Gámez Pastrana, R., Meza, P. A., & Serna-Lagunes, R. (2019). Biogeografía del tomate *Solanum lycopersicum* var. *cerasiforme* (Solanaceae) en su centro de origen (sur de América) y de domesticación (México). *Revista de Biología Tropical*, 67(4), 1023–1036.
- Francis, F. J. (1995). QUALITY AS INFLUENCED BY COLOR. *Food Quality and Preference*, 6, 149–155.
- Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., Burzynski-chang, E. A., Fish, T. L., Stromberg, K. A., Sacks, G. L., Thannhauser, T. W., Foolad, M. R., Diez, M. J., Blanca, J., Canizares, J., Xu, Y., Knaap, E. van der, Huang, S., Klee, H. J., ... Fei, Z. (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nature Genetics*, 51, 1044–1051. <https://doi.org/10.1038/s41588-019-0410-2>
- Georgiady, M. S., Whitkus, R. W., & Lord, E. M. (2002). Genetic Analysis of Traits Distinguishing Outcrossing and Self-Pollinating Forms of Currant Tomato, *Lycopersicon pimpinellifolium* (Jusl.) Mill. *Genetics*, 161(1), 333–344.
- Gibson, M. J. S., & Moyle, L. C. (2020). Regional differences in the abiotic environment contribute to genomic divergence within a wild tomato species. *Molecular Ecology*, 29(12), 2204–2217. <https://doi.org/10.1111/mec.15477>
- Grandillo, S., Ku, H. M., & Tanksley, S. D. (1999). Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theoretical and Applied Genetics*, 99(6), 978–987. <https://doi.org/10.1007/s001220051405>
- Huang, Z., & van der Knaap, E. (2011). Tomato fruit weight 11.3 maps close to fasciated on

the bottom of chromosome 11. *Theoretical and Applied Genetics*, 123(3), 465–474. <https://doi.org/10.1007/s00122-011-1599-3>

IPGRI. (1996). *Descriptors for Tomato (Lycopersicon spp.)*.

Jenkins, J. A. (1948). The Origin of the Cultivated Tomato. *Economic Botany*, 379–392.

Kinkade, M. P., & Foolad, M. R. (2013). Validation and fine mapping of lyc12.1, a QTL for increased tomato fruit lycopene content. *Theoretical and Applied Genetics*, 126(8), 2163–2175. <https://doi.org/10.1007/s00122-013-2126-5>

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., Zhang, Z., Lun, Y., Li, S., Wang, X., Huang, Z., Li, J., Zhang, C., Wang, T., Zhang, Y., Wang, A., Zhang, Y., Lin, K., Li, C., Huang, S. (2014). Genomic analyses provide insights into the history of tomato breeding. In *Nature Genetics* (Vol. 46, Issue 11, pp. 1220–1226). <https://doi.org/10.1038/ng.3117>

Lippman, Z. B., Cohen, O., Alvarez, J. P., Abu-Abied, M., Pekker, I., Paran, I., Eshed, Y., & Zamir, D. (2008). The making of a compound inflorescence in tomato and related nightshades. *PLoS Biology*, 6(11), 2424–2435. <https://doi.org/10.1371/journal.pbio.0060288>

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., & Díez, M. J. (2020). Exploiting the diversity of tomato: the development of a phenotypically and genetically detailed germplasm collection. *Horticulture Research*, 7(1). <https://doi.org/10.1038/s41438-020-0291-7>

Mazidi, M., Ferns, G. A., & Banach, M. (2020). A high consumption of tomato and lycopene is associated with a lower risk of cancer mortality: results from a multi-ethnic cohort. *Public Health Nutrition*. <https://doi.org/10.1017/S1368980019003227>

Meyer, R. S., & Purugganan, M. D. (2013). Evolution of crop species: Genetics of domestication and diversification. In *Nature Reviews Genetics* (Vol. 14, Issue 12). <https://doi.org/10.1038/nrg3605>

Monforte, A. J., Diaz, A., Caño-Delgado, A., & van der Knaap, E. (2014). The genetic basis of fruit morphology in horticultural crops: Lessons from tomato and melon. In *Journal of Experimental Botany* (Vol. 65, Issue 16, pp. 4625–4637). <https://doi.org/10.1093/jxb/eru017>

Munos, S., Ranc, N., Botton, E., Berard, A., Rolland, S., Duffe, P., Carretero, Y., Le Paslier, M.-C., Delalande, C., Bouzayen, M., Brunel, D., & Causse, M. (2011). Increase in Tomato Locule Number Is Controlled by Two Single-Nucleotide Polymorphisms Located Near WUSCHEL. *Plant Physiology*, 156(4), 2244–2254. <https://doi.org/10.1104/pp.111.173997>

Nakazato, T., Bogonovich, M., & Moyle, L. C. (2008). Environmental factors predict adaptive phenotypic differentiation within and between two wild Andean tomatoes. *Evolution*, 62(4), 774–792. <https://doi.org/10.1111/j.1558-5646.2008.00332.x>

Peralta, I. E., & Spooner, D. M. (2006). History, origin and early cultivation of tomato

- (Solanaceae). In M. K. Razdan & A. K. Mattoo (Eds.), *Genetic improvement of Solanaceous crops, vol. 2: Tomato*. Enfield: Science Publishers.
- Ranc, N., Mũos, S., Santoni, S., & Causse, M. (2008). A clarified position for *Solanum lycopersicum* var. *cerasiforme* in the evolutionary history of tomatoes (solanaceae). *BMC Plant Biology*, 8, 1–18. <https://doi.org/10.1186/1471-2229-8-130>
- Rao, E. S., Kadirvel, P., Symonds, R. C., & Ebert, A. W. (2013). Relationship between survival and yield related traits in *Solanum pimpinellifolium* under salt stress. *Euphytica*, 190(2), 215–228. <https://doi.org/10.1007/s10681-012-0801-2>
- Rao, E. S., Kadirvel, P., Symonds, R. C., Geethanjali, S., Thontadarya, R. N., & Ebert, A. W. (2015). Variations in DREB1A and VP1.1 genes show association with salt tolerance traits in wild tomato (*Solanum pimpinellifolium*). *PLoS ONE*, 10(7). <https://doi.org/10.1371/journal.pone.0132535>
- Razifard, H., Ramos, A., Della Valle, A. L., Bodary, C., Goetz, E., Manser, E. J., Li, X., Zhang, L., Visa, S., Tieman, D., van der Knaap, E., & Caicedo, A. L. (2020). Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution*, 1–15. <https://doi.org/10.1093/molbev/msz297>
- Rick, C. M. (1958). The role of natural hybridization in the derivation of cultivated tomatoes of western South America. *Economic Botany*, 12, 346–367.
- Rick, C. M., & Holle, M. (1990). Andean *Lycopersicon esculentum* var. *cerasiforme*: genetic variation and its evolutionary significance. *Economic Botany*, 44(3 Supplement), 69–78. <https://doi.org/10.1007/BF02860476>
- Rick, C. M. (1973). Potential Genetic Resources in Tomato Species: Clues from Observations in Native Habitats. In *Genes, Enzymes, and Populations*. Springer US. https://doi.org/10.1007/978-1-4684-2880-3_17
- Rick, C. M., Holle, M., & Thorp, R. W. (1978). Rates of cross-pollination in *Lycopersicon pimpinellifolium*: Impact of genetic variation in floral characters. *Plant Systematics and Evolution*, 129(1–2), 31–44. <https://doi.org/10.1007/BF00988982>
- Rick, Charles M. (1978). The Tomato. *Scientific American*.
- Rick, C. M., & Fobes, J. F. (1975). Allozyme Variation in the Cultivated Tomato and Closely Related Species. *Bulletin of the Torrey Botanical Club*, 102(6), 376–384.
- Rick, C. M., Fobes, J. F., & Holle, M. (1977). Genetic Variation in *Lycopersicon pimpinellifolium*: Evidence of Evolutionary Change in Mating Systems *. *Plant Systematics and Evolution*, 127, 139–170.
- Rodríguez-Guzmán, E., Ruiz-Corral, J. A., Lépiz-Ildefonso, R., & Pimienta-Barrios, E. (2018). Adaptación climática y edáfica de *Solanum lycopersicum* var. *cerasiforme* en el Occidente de México. *ECORFAN Revista de Ciencias Ambientales y Recursos Naturales*, 4(12), 1–9. www.ecorfan.org/spain,revista@ecorfan.

- Rodríguez, G. R., Muños, S., Anderson, C., Sim, S.-C., Michel, A., Causse, M., Gardener, B. B. M., Francis, D., & van der Knaap, E. (2011). Distribution of SUN, OVATE, LC, and FAS in the Tomato Germplasm and the Relationship to Fruit Shape Diversity. *Plant Physiology*, 156(1), 275–285. <https://doi.org/10.1104/pp.110.167577>
- Rodríguez Guzmán, E., Vargas Canela, D., Sánchez González, J. de J., Lépiz Ildelfonso, R., Rodríguez Contreras, A., Ruiz Corral, J. A., Puente Ovalle, P., & Miranda Medrano, R. (2009). Etnobotánica de *Solanum lycopersicum* var. cerasiforme en el occidente de México. *Naturaleza y Desarrollo*, 7(2), 45–57.
- Sifres, A., Picó, B., Blanca, J. M., De Frutos, R., & Nuez, F. (2006). Genetic structure of *Lycopersicon pimpinellifolium* (Solanaceae) populations collected after the ENSO event of 1997-1998. *Genetic Resources and Crop Evolution*, 54(2), 359–377. <https://doi.org/10.1007/s10722-005-5725-4>
- Silva Ferreira, D., Kevei, Z., Kurowski, T., De Noronha Fonseca, M. E., Mohareb, F., Boiteux, L. S., & Thompson, A. J. (2018). BIFURCATE FLOWER TRUSS: A novel locus controlling inflorescence branching in tomato contains a defective MAP kinase gene. *Journal of Experimental Botany*, 69(10), 2581–2593. <https://doi.org/10.1093/jxb/ery076>
- Soyk, S., Lemmon, Z. H., Oved, M., Fisher, J., Liberatore, K. L., Park, S. J., Goren, A., Jiang, K., Ramos, A., van der Knaap, E., Van Eck, J., Zamir, D., Eshed, Y., & Lippman, Z. B. (2017). Bypassing Negative Epistasis on Yield in Tomato Imposed by a Domestication Gene. *Cell*, 169(6), 1142–1155. <https://doi.org/10.1016/j.cell.2017.04.032>
- Srinivas, T. R. (2001). Floral descriptors of field evaluated tomato germplasm. *Journal of Applied Horticulture*, 3(1), 53–55.
- Stebbins, G. L. (1957). Self Fertilization and Population Variability in the Higher Plants. *American Society of Naturalists*, 91(861), 337–354.
- Stout, M. J., Kurabchew, H., & Leite, G. L. D. (2017). Host-Plant Resistance in Tomato. In *Sustainable Management of Arthropod Pests of Tomato*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-802441-6.00009-7>
- Tanksley, S. D., & McCouch, S. R. (1997). Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild The Narrow Genetic Base of Crop Plants. *Science*, 277(August), 1063–1066. <https://doi.org/10.1126/science.277.5329.1063>
- Warnock, S. J. (1991). Natural Habitats of *Lycopersicon* Species. *HortScience*, 26(5), 466–471. <https://doi.org/10.21273/hortsci.26.5.466>
- Widrechner, M. P. (1987). Variation in Breeding System of *Lycopersicon pimpinellifolium*: Implications for Germplasm Maintenance. *Plant Genetic Resources Newsletter*, 70, 38–43.
- Yoo, H. J., Park, W. J., Lee, G. M., Oh, C. S., Yeam, I., Won, D. C., Kim, C. K., & Lee, J. M. (2017). Inferring the Genetic Determinants of Fruit Colors in Tomato by Carotenoid Profiling. *Molecules*, 22(5), 1–14. <https://doi.org/10.3390/molecules22050764>

- Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. *Nature Reviews Genetics*, 2(12), 983–989. <https://doi.org/10.1038/35103590>
- Zsögön, A., Čermák, T., Naves, E. R., Notini, M. M., Edel, K. H., Weinl, S., Freschi, L., Voytas, D. F., Kudla, J., & Peres, L. E. P. (2018). De novo domestication of wild tomato using genome editing. *Nature Biotechnology*, 36(12), 1211–1216. <https://doi.org/10.1038/nbt.4272>
- Zuriaga, E., Blanca, J. M., Cordero, L., Sifres, A., Blas-Cerdán, W. G., Morales, R., & Nuez, F. (2009). Genetic and bioclimatic variation in *Solanum pimpinellifolium*. *Genetic Resources and Crop Evolution*, 56(1), 39–51. <https://doi.org/10.1007/s10722-008-9340-z>

Capítulo 3:

EXPLOITING THE DIVERSITY OF TOMATO: THE DEVELOPMENT OF PHENOTYPICALLY AND GENETICALLY DETAILED COLLECTION

Exploiting the diversity of tomato: The development of phenotypically and genetically detailed germplasm collection

Estefanía Mata-Nicolás¹, Javier Montero-Pau², Esther Gimeno-Paez¹, Víctor Garcia-Carpintero¹, Peio Ziarsolo¹, Naama Menda³, Lukas A. Mueller³, José Blanca¹, Joaquín Cañizares^{1*}, Esther van der Knaap^{4,5*} and Maria José Díez^{1*}

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

²Department of Biochemistry and Molecular Biology. Universitat de València. Spain

³Boyce Thompson Institute, Ithaca, NY. USA.

⁴Institute of Plant Breeding, Genetics and Genomics. University of Georgia. GA. USA

⁵Department of Horticulture University of Georgia. GA. USA

* Authors for correspondence

Referencia completa:

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., & Díez, M. J. (2020). Exploiting the diversity of tomato: the development of a phenotypically and genetically detailed germplasm collection. *Horticulture Research*, 7(1). <https://doi.org/10.1038/s41438-020-0291-7>

ABSTRACT

A collection of 163 accessions, including *Solanum pimpinellifolium*, *Solanum lycopersicum* var. *cerasiforme* and *Solanum lycopersicum* var. *lycopersicum*, has been selected to represent the genetic and morphological variability of tomato in its center of origin and domestication: Andean regions of Peru and Ecuador and Mesoamerica. The collection is enriched with *S. lycopersicum* var. *cerasiforme* from the Amazonian region that has not been previously analyzed and used extensively. The collection has been morphologically characterized, showing diversity for fruit, flower and vegetative characters. Their genomes were sequenced in the Varitome project and are publicly available (solgenomics.net/projects/varitome). The identified SNPs have been annotated with respect to their impact and a total number of 37,974 of 19,364,146 SNPs have been described as high impact by SnpEff analysis. GWAS has shown associations for different traits, proving the potential of this collection for this kind of analysis. We have identified known QTLs and genes, but also new regions associated with characters such as fruit color, number of flowers or inflorescence architecture. To speed up and facilitate the use of this information, F2 populations were constructed by crossing the whole collection with three different parents. This F2 collection is useful for testing immediately SNPs identified by GWAs, selection sweeps or any other candidate gene. All data is available on Solanaceae Genomics Network and the accession and F2 seeds are freely accessible at COMAV and TGRC genebank. All together make this collection a good candidate for other genetic studies.

INTRODUCTION

Tomato, *Solanum lycopersicum* var. *lycopersicum* L. (SLL), is one of the most consumed vegetables all over the world with a production that exceeds 180 million tons (FAO, 2017). Its cultivation has become highly efficient thanks to the introduction of technological advances and the development of modern varieties. These modern varieties are the result of intensive plant breeding programs since the beginning of the 20th century, and the natural biodiversity of tomato wild species has been key in this success.

The cultivated tomato and its wild relatives came from the Peruvian and Ecuadorian regions of South America. According to allozyme variation, Rick & Fobes (1975) proposed that SLL evolved from *S. lycopersicum* var. *cerasiforme* (Dunal) Spooner, Anderson & Jansen (SLC). Recently, Blanca et al. (2012; 2015) proposed a two-step domestication process from SLC to SLL based on molecular and morphological evidence. The first step involves the pre-domestication of SLC in the Amazonian region of Southern Ecuador and Northern Peru. Subsequently, SLC would have migrated to Mesoamerica where it would be domesticated to SLL. Razifard et al. (2020) proposed that many traits considered typical of cultivated tomatoes arose in South America. However, these domestication traits were lost or diminished once these partially domesticated forms spread to Mesoamerica, where it was finally morphed into the SLL (Bai & Lindhout, 2007; Peralta et al., 2008). This domestication and diffusion process was accompanied by a selection of alleles related to fruit color, size and shape and also changes in plant architecture (Paran & Van Der Knaap 2007; Rodríguez et al. 2011; Ichihashi & Sinha 2014; Monforte et al. 2014). This process also included various genetic bottlenecks that progressively narrowed the genetic diversity of modern tomato, compared to its wild species (García-Martínez et al. 2006; Blanca et al. 2015). The main loss of variability occurred during the migration to Mesoamerica from the Peruvian and Ecuadorian Amazon region. Most of the allelic variants present in European vintage tomato are already present in these Amazonian SLC populations (Blanca et al. 2015).

Solanum pimpinellifolium L. (SP) is the closest wild relative to SLC and SLL. It is also a red-fruited species and native to coastal areas from Ecuador to Southern Peru. According to its distribution, this species presents varying degrees of genetic variation (Rick, Fobes, and Holle 1977; Caicedo & Schaal 2004; Sifres et al. 2006; Zuriaga et al. 2009) and morphological differences such as flower and inflorescence size, style exertion, or fruit color (Rick, Fobes, & Holle 1977). This fact and its capacity to hybridize with tomato, make this species a valuable source of desired traits in tomato breeding. For instance, SP has been used as a genetic source for quality improvement related to solid content, firmness, fruit color (Rick 1978; Stevens and Rick 1986), volatile compounds (Capel et al., 2015; Rambla et al., 2017), or resistance against fungi or viruses such as Tomato leaf curl virus (Banerjee & Kalloo, 1987), *Alternaria solani*, *Fusarium oxysporum*, and *Phytophthora infestans* (Alexander & Hoover, 1955) or *Cladosporium fulvum* (Walter, 1967). SLC has a worldwide distribution in tropical regions, but it is native to the Andean region of Ecuador and North of Peru (Rick and Fobes 1975). This species is found over a vast range of environmental conditions such as tropical or arid regions, sea level or high altitudes (Warnock, 1991), and it has also been collected at native markets (Rick 1958). It usually bears red and small fruits, but (Rick & Holle 1990) described a remarkable morphological variability in fruits, plant habit, or leaf size and shape. A higher genetic variability has been described in Ecuadorian and Peruvian accessions (Rick & Fobes 1975; Blanca et al. 2012) due to the development of morphological diversity during a predomestication phase. In fact, tomatoes collected in local markets of Ecuador were morphologically classified as vintage tomato; but they have been

genetically classified as SLC (Blanca et al. 2015). These studies and data show that SLC from Northern Peru are very close to Mexican and vintage tomatoes. Despite that, Amazonian SLC has not been used frequently for tomato improvement as opposed to SP. Furthermore, SLC has been characterized as a valuable genetic source for abiotic and biotic stresses, such as moisture-tolerance (Nuez & Díez, 2008) or resistance to root rot caused by *Phytophthora* (Arellano Rodríguez et al., 2013); traits related with the global climate change and sustainability challenges currently facing agriculture.

Most modern breeding programs have usually focused on resistance, yield, and quality traits, such as firmness, color, or texture (Foolad, 2007), plant habit and adaptation to machine harvesting in processing cultivars or traits related to fruit appearance for fresh market (Foolad, 2007). However, nowadays, the new objectives of tomato breeding focus on sustainable production or adaptation to unfavorable environmental conditions due to climate change and nutritional quality. The genetic variation of exotic germplasm collections has been used in tomato breeding to bypass the limited genetic diversity of SLL. These germplasm collections have mainly included *S. pimpinellifolium*, *S. chilense* (Dunal) Reiche, *S. peruvianum* L. s. str., *S. habrochaites* S. Knapp & D.M. Spooner and *S. pennellii* Correl. Thus, the maintenance and characterization of germplasm collections are essential to achieve these breeding goals. Germplasm is a good source of natural allelic variants, useful for genetic analyses and subsequent breeding applications. Consequently, the creation of genebank collections characterized at genetic and phenotypic level is a primary objective for a sustainable breeding. In addition, it is crucial that these data and genetic resources are easily available to the scientific community to exploit this extensive amount of information.

The advent of NGS technologies has created a huge amount of available genetic information about germplasm held in genebanks (Robertson & Labate, 2014) that can be useful for improving breeding cultivars (Rothan, Diouf, & Causse 2019). For instance, the availability of its genomes in association with its characterization at phenotypic and molecular level allows the development of genome-wide association studies (GWAS). GWAS studies have already identified regions of the genome related to morphological and metabolic diversity (Bauchet et al. 2017a; Tieman et al. 2017). For example, Bauchet et al. (2017a; 2017b) detected associations for traits such as fruit weight, flowering time, early fruit development, malate, and phenylacetaldehyde/phenylethanol content. Finally, the first meta-analysis of GWAS has revealed numerous candidate genes involved in tomato flavor (Zhao et al., 2019). Full genome sequences have been published in several studies and more than 725 genome sequences of tomato accessions are available (Aflitos et al. 2014; Gao et al. 2019; Soyk et al. 2017; Zhu et al. 2018b). A pan-genome analysis of tomato including SLL, SLC, and SP has discovered 4873 genes that are not present in the reference genome (Gao et al. 2019) thus increasing the interest of these populations for tomato breeding. Once a candidate region of the genome, gene or SNP has been characterized as significantly associated with a trait, it is necessary to validate its role in the control of the trait by using segregating families or mutants. However, this latter step sometimes becomes limiting as the development of such populations is time consuming and costly.

In the present study, we have morphologically characterized the variability of fruit, flower, and vegetative characters from a collection of 163 tomato accessions of the Varitome project, for which the full genome is available (Gao et al. 2019). These accessions include SP, SLC, and SLL and represent the diversity at the center of origin and domestication of tomato. We have annotated the identified SNPs within our collection using SnpEff. We have performed GWAS analysis for all our morphological descriptors with the aim of detecting

candidate regions. In addition, a collection of segregating families has been developed by crossing the complete set of accessions with a representative accession for each of the three species. These populations could help to speed up the validation of candidate genes and SNPs. The combination of passport, phenotypic, genetic information, and germplasm with easy accessibility converts this collection into a powerful instrument for genetic studies and breeding.

MATERIAL AND METHODS

Plant material

A germplasm collection of 163 accessions was selected with the aim of representing a broad range of geographical, morphological, and genetic diversity. These plant materials consisted of 15 accessions of SLL from Mexico; 121 accessions of SLC coming from Ecuador, Peru, Mexico, and different countries of Mesoamerica and 27 accessions of *Solanum pimpinelifolium* (SP), from Ecuador and Peru. Accessions were grouped according to their geographical origin and previous genetic results (Blanca et al. 2012; 2015) as is shown in Fig. S1 and Table S1-ANEXO B. These accessions were provided by different germplasm banks such as Tomato Genetics Resource Center (TGRC), United States Department of Agriculture (USDA), and Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV) of Universitat Politècnica de València. Passport data of these accessions are available in Supplementary Table S1-ANEXO B and COMAV, Solanaceae Genomics Network web pages. For each accession, seeds obtained after a double round of self-pollination from a single plant of each original accession were collected and they were used both for the morphological and genetic characterization and for the creation of segregating populations.

Morphological characterization

Plants used for the morphological characterization were cultivated in a greenhouse at Universitat Politècnica de València (Valencia, Spain) during the spring-summer seasons of 2016. Plants were grown in 12-l pots with coconut fiber and fertirrigated under standard dosages for tomato in our area. A completely randomized experimental design was conducted with two plants per accession, each replicate in a different greenhouse.

Twenty-six quantitative and 27 qualitative traits based on the descriptors developed by IPGRI (1996), mainly related to plant architecture, inflorescences and flowers, leaves and fruit size were evaluated. Some descriptors were modified for a better representation of the morphological variability exhibited in the collection. The descriptors and their definitions are listed in Supplementary Table S2-ANEXO B. All traits were added to the Solanacea Phenotype Ontology, available at SGN (<https://solgenomics.net/search/traits>). Prior to any analysis, all traits were manually curated to detect possible errors. Differences due to a greenhouse effect were assessed using Student's t test and Mann-Whitney-Wilcoxon test for quantitative traits depending on whether the data was normally distributed. Fisher's exact test was conducted on qualitative data. As no differences between the two greenhouses were found, data from both greenhouses were joined, and the mean value was calculated for quantitative traits. For qualitative data, a new level for each qualitative trait (named as 10) was created to include the accessions which presented different scale values for this qualitative trait. Robust ANOVA and Fisher test were conducted to detect significance differences between species, depending on whether the data was quantitative or qualitative. A

Bonferroni correction of p values was conducted.

Genetic analysis

Genome sequences of the accessions of this collection have been published previously (Aflitos et al., 2014) and the SNPs identified in these accessions are publicly available in Solanaceae Genomics Network (<https://solgenomics.net/projects/varitome>). Using these data, the collection of SNPs has been annotated to detect the localization and possible impact of changes using SnpEff (Cingolani et al., 2012), and statistics were calculated for each geographical group.

GWAS between genotypes and phenotypes were calculated for all quantitative and qualitative traits using R package GENESIS v.2.14.1 (Gogarten et al., 2019). A total number of 1.479.141 high quality SNPs were used for the GWAS analysis. A PCoA has been done to visualize genetic structure (Supplemental figure S3-ANEXO B). To test the association, a generalized linear mixed model (GLMM) using the genetic relationship matrix (GRM) as random effects was used in order to account for population stratification. GRM was computed using GCTA v.1.92.1 (Cingolani et al., 2012). For count data, a Poisson distribution of residuals was assumed, while for the rest of the quantitative data a Gaussian distribution was applied. Normality was checked using a Shapiro-Wilk normality test and a Box-Cox power transformation was used when necessary. For qualitative traits, a binomial distribution was assumed. For multinomial qualitative traits, each category level was treated as a dummy binary variable. Quantile-quantile plots were used to assess the GWAS model (Supplemental figure S4-ANEXO B). Significant level of association was estimated using GEC (Genetic type 1 Error Calculator) v.0.2 (Li et al. 2012).

Development of breeding population

In order to help to exploit the variability detected in our collection and facilitate its use to the research community, F1 and F2 generations were constructed for the 163 accessions by crossing each accession with one accession representative of each species (SP, SLC and SLL). The accessions, BGV007109 of SP, LA2278 of SLC and Money Maker of SLL were selected as parents. Each fruit from each individual cross was maintained separately in order to facilitate the detection of possible mistakes. Two different F1 plants of each combination were self-pollinated to obtain the set of two independent F2 breeding populations. The culture for the F2 family generation was done during the years 2017, 2018 and 2019 in the Centro de experiencias Cajamar de Paiporta (Valencia, Spain). Plants were grown in greenhouses in soil and fertirrigated under standard dosages for tomato in our area.

The complete list of these available materials is listed in Supplementary Table S4-ANEXO B.

RESULTS

Morphological analysis

A germplasm collection of 163 accessions was selected with the aim of representing the geographical, morphological, and genetic diversity of tomato and its closest wild relatives in their region of origin (Supplemental Figure S1 and Supplemental Table S1- ANEXO B). These materials consisted of 15 accessions of SLL from Mexico; 121 accessions of SLC coming from Ecuador, Peru, Mexico and different countries of Mesoamerica and 27

accessions of SP from Ecuador and Peru. The accessions have been grouped based on their geographical origin and on previous genetic studies (Blanca et al. 2012; 2015). Plants were evaluated for a total of 54 morphological traits (supplemental Table S2-*ANEXO B*) describing the variability of this collection for plant architecture, leaves, inflorescences, flowers, and fruits (Figure 1, 2, Supplementary Table S3-*ANEXO B*). The lowest morphological variability was found in quantitative traits related to plant architecture such as height until first or last inflorescence (Figure 1a) or stem width (Figure 1b). However, SP can be differentiated from the rest of species by this last trait. Qualitative traits related to plant architecture showed that most accessions had an indeterminate growth habit and that exists a wide range of variation related to the way that leaves were held naturally (Figure 1c).

Quantitative traits related to leaves were the leaf size, the number of primary leaflets and small leaflets, whereas the qualitative ones described the leaf morphology, complexity and leaflet dissection and shape. The collection exhibited a low variability for number of primary leaflets while differences were remarkably greater for leaf size and the number of small leaflets, as it is shown in Figure 2d. Figures 1d and 1e show that SP group was characterized by smaller leaves whereas the maximum values were found in SLC group. Observations related to the type of leaf revealed that SP group was generally characterized by *pimpinellifolium* type leaf, SLC group exhibited all types but generally leaves were classified as standard ones and SLL exhibited standard and double feathered types. SP leaf was generally characterized by a lack of dissection (Figure 1f) and entire or undulating borders. However, SLC and SLL groups exhibited more variability in leaflet dissection (Figure 1f) and border.

Traits related to inflorescences included inflorescence length, number of flowers per inflorescence or type of inflorescence whereas flowers were evaluated for number of petals and sepals and their length, width, and style exertion, among others. The values observed for inflorescence length and the number of flowers per inflorescence demonstrated a wide variability (Figures 1g and 1h). In addition, the complexity of the inflorescence exhibited a remarkably diversity (Figure 2e-2g). For flower traits, most accessions had between 5 and 6 petals and sepals per flower, but several accessions were much more complex (Figure 2b). SP Ecuador and Peru and SLC Mesoamerica exhibited the simplest flowers and low variability, opposite to the complexity observed in SLC Ecuador, SLC Peru, SLC Mexico and SLL. Also, the observed variability related to the position of style is represented in Figures 2c and 1i.

The high variability for fruit weight and locule number is shown in Figures 1j and 1i, respectively. SP was characterized by the smallest fruits whereas SLL group presented the biggest ones. However, the highest variability appeared in SLC group which reached smaller values than SP or bigger than SLL. Finally, qualitative traits related to fruit appearance revealed that most accessions produced red fruits, although other colors were also present. Some accessions belonging to SP species presented an intense red fruit, and others belonging to SP Peru and SLC Mexico groups exhibited colors ranging from yellow to orange. Other qualitative fruit traits presented a high variability, such as the presence and intensity of green shoulders (Figure 1l and 2h). This variability in fruit size, color and shape is shown in Figure 2a.

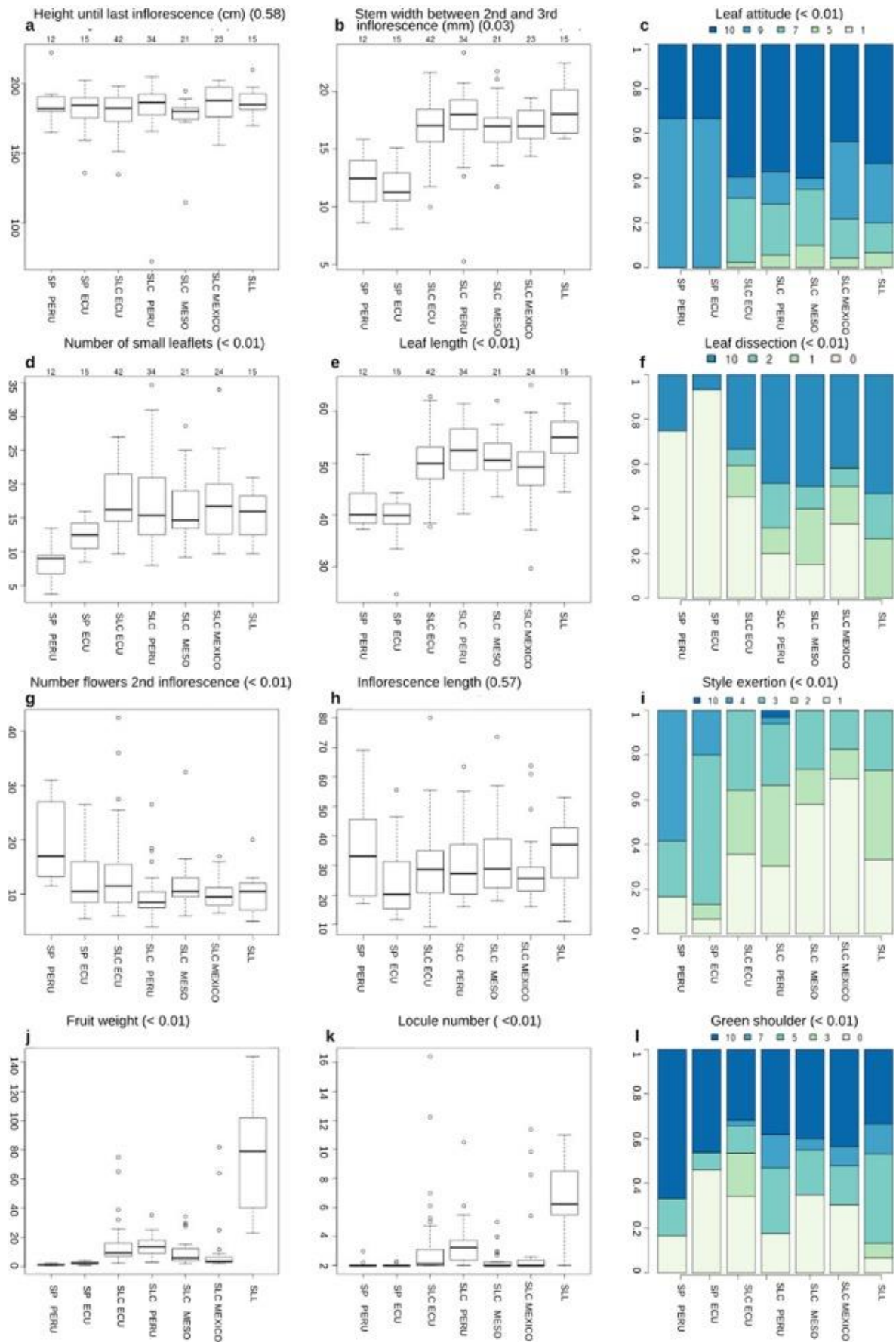


Figure 1. Morphological variation.

Distribution for 8 quantitative and 4 qualitative morphological traits related to vegetative (a-c), leaf (d-f), flower (g-i) and fruit (j-l) descriptors for each geographical group. P-values (in brackets) of the differences between species are shown. Morphological traits were measured as: a. Plant height until last inflorescence, measured in cm. b. Stem width between 2nd and 3rd inflorescence, measured in mm. c. The way that leaves are held naturally (1:semi-erect, 3:semi-horizontal, 5:horizontal, 7:horizontal-drooping, 9:drooping, 10:accessions that exhibited variability for their measures). d. Number of small leaflets. e. Leaf length, measured in cm. f. Leave dissection (0:low, 1:intermediate, 2:high, 10:accessions that exhibited variability for their measures). g. Number of flowers in the second inflorescence. h. Distance from the stem to the last flower of the inflorescence. i. Position of the style in relation to stamens (1:inserted, 2:same level as stamen, 3:slightly exerted, 4:highly exerted, 10: accessions that exhibited variability for their measures). j. fruit weight, measured in grams. k. Number of locules in the transversal section of the fruit. l. Presence and color of green shoulder (0:uniform, 3:light green, 5:medium green, 7:dark green, 10: accessions that exhibited variability for their measures).

Genome-wide association analysis

GWAS analysis revealed significant associations with a total of 15 traits. We found SNPs associated with eight quantitative traits (Figure 3 and in Table S5-ANEXO B). For the total number of inflorescences and petal length traits, each was associated with a single SNP located on chromosomes 1 and 9 respectively (Table 1). The number of flowers in the second inflorescence revealed associations with two SNPs located in chromosome 7 and 11. The result of leaf length analysis revealed two associated regions on chromosome 2 and 8. Associations with locule number were detected on chromosome 1, 2 and 11 and associations with fruit weight were detected on chromosomes 2, 7, 9, 12. The most remarkable associations occurred on chromosome 2, since associated SNP were located in the genomic region where *locule number* and *fw2.2* QTLs have been described. On chromosome 11, the association with the trait number of locules is located on the *fas* gene. On chromosome 9, the previous QTL *fw9.2* was detected for fruit weight. Interestingly, there is not a close described QTL for chromosomes 1 and 12 related to locule number or weight, respectively. Several associations for fruit color have been detected, listed in Table 1 and Table S5-ANEXO B. For instance, GWAS for LAB color space's b value revealed associations on chromosome 1 that were located in a genomic region with an annotated gene as carotenoid cleavage dioxygenase 1B. Regions on chromosome 3 and 10 were close to annotated genes involved in yellow and orange fruit flesh. Finally, the analysis detected also a region on chromosome 5 that has not been previously described for this trait. For LAB color space's L value, the association detected on chromosome 3 lacks of annotated genes. GWAS analysis also showed associations between SNPs and qualitative traits, as it is shown in Figure 3 and Table 1. The genomic region on chromosome 9 associated to dark-green leaves lacked annotated genes and only one significant SNP for low petal curvature was detected on chromosome 7. For the type of inflorescences, one genomic region on chromosome 9 could be involved in forked inflorescence and chromosome 11 could carry another region that could be involved in uniparous inflorescence. For fruit traits, associations with the presence of longitudinal stripes, fasciated fruit, ribbing at calix end and fruit scar were detected. The most remarkable result was the association for irregular pistil scar, covering a region of 355 kb on chromosome 11 that included several genes and three of them were also associated to ribbing at calix end. Finally, two genomic regions on chromosome 1 were associated with pink fruits (175 kb) and fasciated fruits (200 Kb).

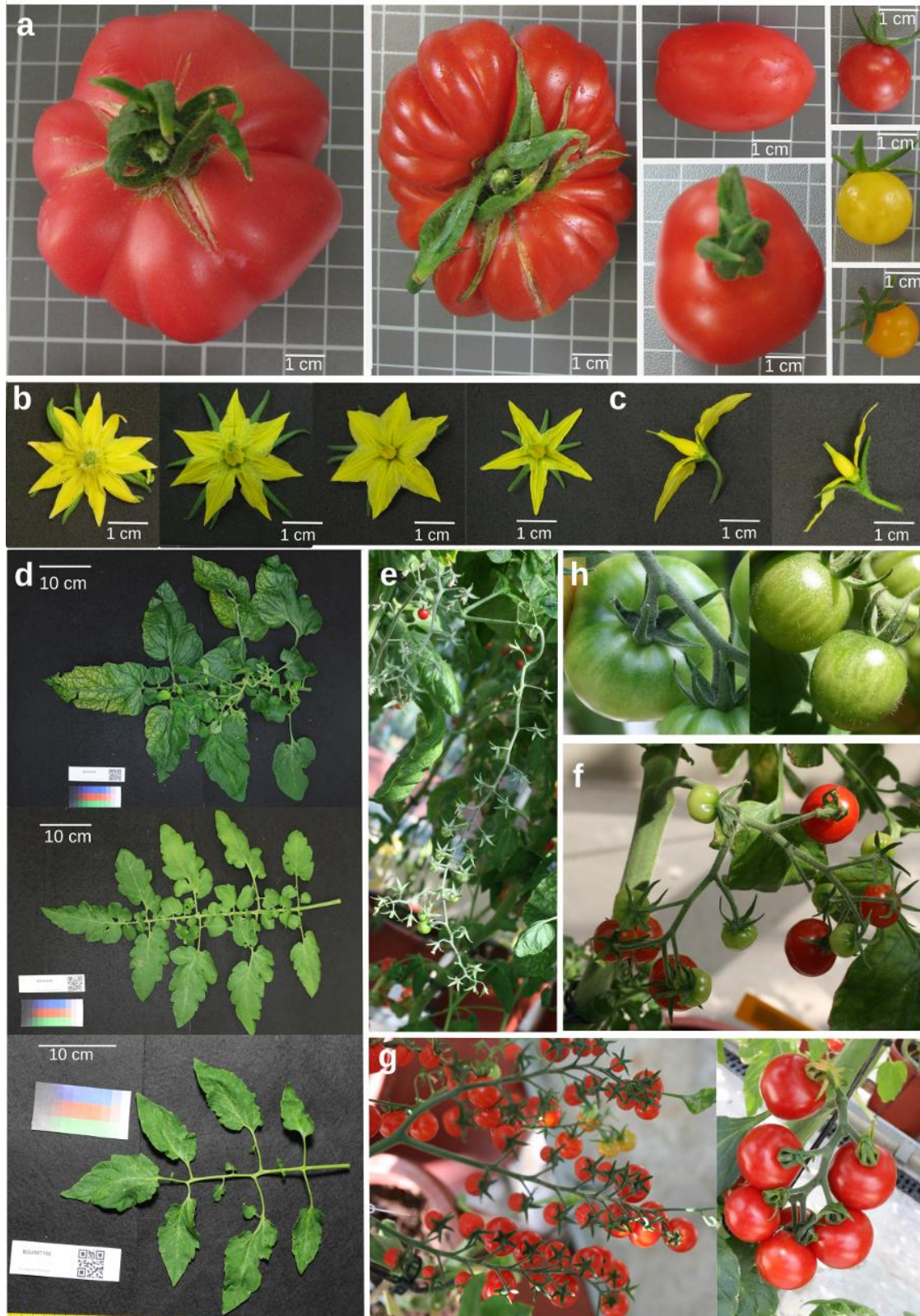


Figure 2. Diversity in leaf, fruit, flower, and inflorescence traits. a. Tomato fruit size, shape and color. b. Variability for flower complexity, related to the number of petals and sepals and their sizes. c. Differences between exerted and inserted styles. d. Diversity in leaf size, number of small leaflets and border or dissection of small leaflets. e. Uniparous inflorescence. f. Forked inflorescence. g. Irregular inflorescence. h. Differences between presence and absent of green shoulder.

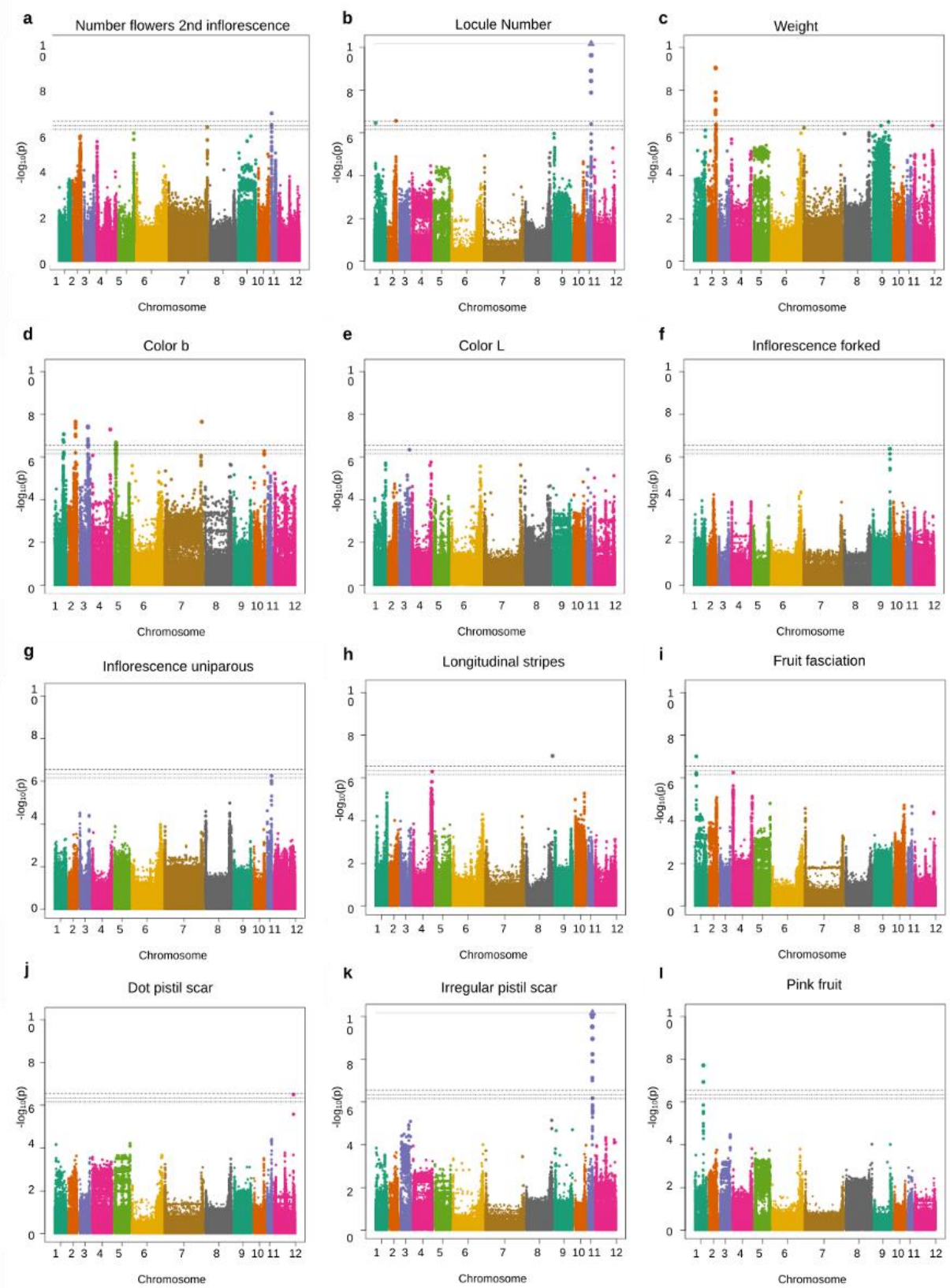


Figure 3. Genome-wide association results for some traits that showed significant association.

Table 1. Summary of significant associations detected for quantitative and qualitative traits. For each trait, the position in bp on the chromosome, the corresponding annotated gene, or close annotated gene, known genes related to the trait and number of SNPs with a high putative effect.

Trait	Chromosome	Position	Locus name	Close annotated locus name	Know QTL	Number of high impact SNPs
color b	1	78911282-82203699	<i>Solyc01g079760</i> , <i>Solyc01g087260</i>			<i>Solyc01g087260</i> : 2
color b	2	44794084	<i>Solyc02g080610</i>			
color b	3	52153309-52406077	<i>Solyc03g081260</i> , <i>Solyc03g081300</i> , <i>Solyc03g082480</i>			<i>Solyc03g081260</i> : 14
color b	5	9596743-9779918	<i>Solyc05g015030</i> , <i>Solyc05g015070</i>			<i>Solyc05g015030</i> : 2
color b	10	60308826	<i>Solyc10g078510</i>			
color L	3	65751759				
dark green leaf	9	65565848-65566230	-	<i>Solyc09g072880</i>		
fruit fasciation	1	3690368-3918681	<i>Solyc01g009510</i>			<i>Solyc01g009510</i> : 1
fruit fasciation	4	3935728	-			
fruit weight	2	49587330-52662216	<i>Solyc02g087050</i> , <i>Solyc02g087780</i> , <i>Solyc02g087810</i> , <i>Solyc02g091330</i>		<i>fw2.2</i>	<i>Solyc02g087050</i> : 1, <i>Solyc02g091330</i> : 1
fruit weight	7	1318982	<i>Solyc07g006520</i>			<i>Solyc07g006520</i> : 1
fruit weight	9	31453421	-		<i>fw9.2</i>	
fruit weight	9	60775035	-		<i>fw9.2</i>	
fruit weight	12	61805734	<i>Solyc12g055810</i>			

Table 1. Continuation.

Trait	Chromosome	Position	Locus name	Close annotated locus name	Know QTL	Number of high impact SNPs
inflorescence forked	9	68492821-68496154	<i>Solyc09g082830</i>			
inflorescence uniparous	11	55020323-55194697	<i>Solyc11g071600</i> , <i>Solyc11g071830</i> , <i>Solyc11g071840</i>			<i>Solyc11g071600</i> : 1
leaf length	2	39577220-39581181	-	<i>Solyc02g069760</i>		
leaf length	8	60002113-60010971	-			
locule number	1	3717866	-			
locule number	2	49617538	<i>Solyc02g087100</i>		<i>locule number (lc)</i>	<i>Solyc02g087100</i> : 1
locule number	11	54838887-55194697	<i>Solyc11g071310</i> , <i>Solyc11g071600</i> , <i>Solyc11g071820</i> , <i>Solyc11g071830</i> , <i>Solyc11g071840</i>		<i>fas and fw11.3</i>	
longitudinal stripes	4	60283297	-	<i>Solyc04g074290</i>		
longitudinal stripes	8	65493583	<i>Solyc08g082820</i>			<i>Solyc08g082820</i> : 2
Number flowers in 2nd inflorescence	7	65679033	-			

Table 1. Continuation.

Trait	Chromosome	Position	Locus name	Close annotated locus name	Know QTL	Number of high impact SNPs
Number flowers in 2nd inflorescence	11	101719-612480	<i>Solyc11g005110</i>	<i>Solyc11g005510</i> , <i>Solyc11g005570</i>		
petal curvature low	7	1760949	-	<i>Solyc07g006910</i>		
petal length	9	69065327	-	<i>Solyc09g083410</i>		
pink mature fruit	1	78736178-78911282	<i>Solyc01g079760</i>	<i>Solyc01g079620</i>		<i>Solyc01g079620</i> : 5
pistil scar dot	12	61805734	<i>Solyc12g055810</i>			<i>Solyc12g055810</i> : 2
pistil scar irregular	11	54838887-55194697	<i>Solyc11g071310</i> , <i>Solyc11g071580</i> , <i>Solyc11g071600</i> , <i>Solyc11g071820</i> , <i>Solyc11g071830</i> , <i>Solyc11g071840</i>			<i>Solyc11g071580</i> : 9
Ribbing calix end	11	55020323-55183870	<i>Solyc11g071600</i> , <i>Solyc11g071820</i> , <i>Solyc11g071830</i>			
total number of inflorescences	1	3717866	-			

Annotation and prediction of the SNP effects

The SNPs identified in this collection are available at Solanaceae Genomics Network (<https://solgenomics.net/>). The SNPs were annotated, and their putative impacts were predicted by using SnpEff. The effects were classified by the impact of these variants, type of effect and region. The lowest number of variants was detected in our SLL group. SLC groups had a variation between SLL and SP groups with lower levels of variants in SLC Mexico. A total of 37,974 of 19,364,146 SNPs detected in this collection have been designated as high impact in the SnpEff analysis. The number of variants per type and the number of effects by impact for each group are summarized in Table 2. However, it is important to take into account that the number of SNPs is influenced by the different number of accessions in each of our groups. Among other mutations, the generation or the loss of stop codon could be one of the most interesting changes because the synthesis of an essential protein could be affected, and its function would have changed. As it is shown in Figure S2, the same pattern of this SNP distribution was observed for the number of these mutations. Finally, genomic regions of candidate genes from GWAS analysis were used to find out allelic variants in our collection labelled as high impact. These 37,974 SNPs with a high putative impact were related to 12 candidate genes (Table 1) and are summarized in Supplemental Table S6-ANEXO B.

Table 2. Result of the number of variants per type and the number of effects by impact for each geographical group from SnpEff. SNP- Single-Nucleotide Polymorphism, Ins-Insertion and MIXED-Multiple-nucleotide and an InDel.

	Number of variants per type			Number of effects by impact			
	SNP	INS	MIXED	HIGH	LOW	MODERATE	MODIFIER
all samples	15,700,927	2,736,310	926,909	45,111	143,153	196,099	27,035,721
SP Ecuador	7,705,076	1,976,689	738,373	28,070	77,621	97,472	17,935,822
SP Peru	7,752,552	1,995,127	747,838	28,673	81,778	102,291	15,074,861
SLC Peru	6,476,228	1,834,571	699,351	25,197	65,290	83,083	12,905,503
SLC Ecuador	6,358,145	1,726,621	615,355	24,137	63,308	81,036	12,515,017
SLC mesoamerica	5,963,635	1,628,278	619,202	21,682	55,103	72,443	11,580,385
SLC Mexico	3,781,905	1,112,989	371,858	13,919	31,684	45,342	7,342,190
SLL	658,721	387,752	87,763	4,178	9,168	11,567	1,793,184

Development of segregating families

The whole collection was crossed with BGV007109 (SP), LA2278 (SLC) and Money Maker (SLL). The 163 accessions were used as the female parent to obtain the F1 generations, except for some accessions, mainly SP, where flowers were too difficult to emasculate due to the small size. F1 plants were self-pollinated to obtain the F2 generations. A collection of 485 F1 populations and 457 F2 populations were achieved (Supplemental Table S4-ANEXO B). Considering that most of the cross collections from each accession can have various independent F2 populations, created from different F1s, the total number of different F1 and F2 populations are 1430 and 672, respectively. The seeds of these segregating families are

available at COMAV.

DISCUSSION

Morphological variability

The results of our study revealed a wide range of diversity in our collection for most of the evaluated traits, mainly traits related to leaves, fruit shape and size, and color or flower morphology. Regarding SLC group, it generally exhibited the highest grade of morphological diversity since the group comprises a wide range of geographical origins. For traits related to leaf shape and size, this group has much higher diversity in comparison to the simple leaves of SP (Rick 1958; Rick & Holle 1990).

The high variation related to fruit color and shape was an interesting observation for breeding and genetic purposes. Some SP and SLC accessions, collected in Peru and Mexico respectively, exhibited colors ranging from yellow to orange. In fact, yellow fruits have been previously described in these localizations (Rick 1958; Rick, Fobes, & Holle 1977; Rick & Holle 1990). The fruit shape of SLC group exhibited a considerable variation ranging from round to flattened, fasciated or elongated fruits. The transition from small and uniform fruits of SP to diversity in fruit size, shape and locule number was a consequence of variations in flower complexity and an increase in ovary size. For example, the appearance of fasciated phenotype (*fas*) has been suggested to have arrived in Europe from Mexico in the 16th century (Rodríguez et al., 2011). These changes in fruit size and shape have been described to be a consequence of derived alleles of *fas*, *sun*, *ovate* and *lc* genes (Rodríguez et al., 2011). According to the study of Blanca et al. (2015), some accessions from our SLC group carry *fas* and *ovate* and some of our SLL also carries derived alleles. These results support our observations, and these derived alleles could be an explanation of the diversity that we have detected.

Changes in flower complexity and style exertion could be other interesting changes related to domestication and further selection processes. For example, the number of petals or sepals tended to increase in SLC and SLL groups, although not in SLC Mesoamerica. The style position has also been altered from highly exerted in SP Peru to slightly exerted or even inserted in SP Ecuador. In SLC, the degree of style exertion tended to decrease from Amazonian SLC to SLC Mexico, whereas in the SLL group it tended to be inserted. This correlation between stigma exertion and SP geographical origin had been previously described (Rick, Fobes, & Holle 1977; Widrlechner 1987), and variation for SLC from South America has also been described (Rick 1958; Rick & Holle 1990). This insertion process is related with the migration from the center of origin and results in the development and increasing of the autogamy levels. Interestingly, two different subgroups can be discerned in SLC Mexico. Accessions collected as wild that exhibited inserted styles, and accessions with fruit sizes similar to cultivated tomato that were characterized by exerted ones. This presence of exertion is also detected in SLL and probably related to fasciated and big fruit morphology.

Genetic variability

As expected, the highest level of diversity was found in Peru and Ecuador for both SP and SLC groups (Blanca et al. 2012; 2015). Rick & Fobes (1975) described that variation in SLC depended on its geographical origin, being SLC from other countries less variable than SLC from these regions. The analysis within the SLC group also revealed a considerable decrease in the number of SNP variants detected in SLC Mexico. This is in agreement with

the loss of variability that took place during the migration to Mesoamérica (Blanca et al. 2012; 2015; Williams & Clair 1993). The detected 37,974 SNPs labelled as high impact show that this collection may be an interesting source of new alleles. This is supported by the SNPs with a high putative effect that were detected for some of our candidate genes from GWAS analysis (Supplementary Table S6-ANEXO B). For example, candidate genes related to yellow fruit color had a total of 18 allelic variants with high effect. Two of them correspond to a carotenoid cleavage dioxygenase 1B (*Solyc01g087260*), 14 of them correspond to a protease-like protein (*Solyc03g081260*) and the last 2 correspond to a homeobox leucine-zipper protein (*Solyc05g015030*). Also, a high impact SNP was detected in the genomic region where *lc* gene is located. Besides, the genomic region associated to *fw2.2* presented 2 allelic variants with high effect. These variants are related with a nodulin MtN21 family protein (*Solyc02g087050*) and with an uncharacterized protein (*Solyc02g091330*).

GWAS analysis

GWAS analysis has revealed a total number of 107 SNPs associated to eight quantitative traits and 30 SNPs associated to seven qualitative traits. This analysis has allowed the identification of known and novel genes for these traits. Despite of QTLs for flowers per inflorescence had been previously described on chromosomes 2, 3 and 5 (Georgiady et al., 2002), our analysis identified a possible novel genomic region on chromosome 11, which carries genes encoding for Agenet and cellulose synthase proteins. Agenet has been described to be involved in flower development (Brasil et al., 2015) and the expression of cellulose synthase has also been detected in flowers of *Arabidopsis thaliana* (Holland et al., 2000). The association identified for forked inflorescence on chromosome 9 corresponds to a SNP in ARGONAUTE 1 gene (*Solyc09g082830*). This gene is a member of AGO gene family, which is known to regulate vegetative and reproductive development and stress response (Zhang et al., 2015). The expression of these genes has been detected in flower and fruit of tomato (Meskauskiene et al., 2001). A significant association with uniparous inflorescence was identified on chromosome 11 and was located approximately 5 Mb away from a mapped region which is considered to be involved in branched inflorescences of *fin* mutants (Xu et al., 2015). Six associations for dark green leaves were detected on chromosome 9. An annotated gene as chloroplast FLU-like protein was located 2 Kb away from this region. FLU is a nuclear-encoded plastid protein that interacts with enzymes involved in chlorophyll synthesis (Meskauskiene et al., 2001). Besides that, some new identified SNPs lacked in functional annotation, for example SNP associated to the total number of inflorescences. Finally, another several traits were associated with SNPs located on genes that were not apparently related to the trait they are associated with, such as the association between leaf length and a RING-finger protein-like which could regulate ubiquitination processes (Joazeiro & Weissman, 2000) or the association between longitudinal stripes and heat shock proteins. All these novel detected regions would require further experiments for validation and identification of candidate genes suitable for tomato breeding.

As expected, our analysis has identified several SNPs located close to genes or candidate regions previously characterized. GWAS analysis has allowed the identification of previously described *loci* associated to fruit size such as *fw2.2* (Frery et al., 2000), *fw9.2* (Grandillo, Ku, & Tanksley 1999), *locule number (lc)* and *fas* (Rodríguez et al., 2011). For instance, we detected an association between fruit weight and SNPs close to *fw2.2* and also close to SNPs that have already been identified in other GWAS analyses (Frery et al., 2000; Rothan et al., 2019). In the case of *lc* and *fas* genes, our study revealed associations between the trait number of locules and SNPs located genetically close to both QTLs, which are

located on chromosome 2 and 11, respectively. Sacco et al. (2015) detected *lc* gene and also one of our annotated candidate genes on chromosome 11, *Solyc11g071840*. This detected region on chromosome 11 was located in a region previously described as *fas* and really close to the annotated *fw11.3*. The *fw11.3* is a QTL controlled by Cell Size Regulator, which regulates weight by the control of cell size in the pericarp (Mu et al., 2017). Strikingly, no association has been found between the qualitative trait fruit-fasciation and chromosome 11 (where *fas* gene is located). However, two regions previously undescribed that could be associated to fasciated phenotype on chromosomes 1 and 4 were revealed. Despite that *fas* gene is not located in these detected regions, this result suggests the involvement of new genome regions. In fact, the SNP located on chromosome 1 at 3,717,866 bp was also associated with the number of locules in our analysis.

An association signal for fruit color was identified on chromosome 1 and located in a region with a candidate gene described as carotenoid cleavage dioxygenase 1B. Carotenoids are an important factor implied in fruit color and modifications or absence of their syntheses are the reason for the orange color of mutants such as tangerine (*t*), delta (*Del*) and beta (*B*) or the yellow flesh (*r*) mutant (Yoo et al., 2017). The genome region involved in this *r* mutant is located on chromosome 3 at 9 Mbs from our associated region. For the pink color, associations were located in a genomic region with an annotated gene as colorless fruit epidermis (*y* gene). This association between the pink fruit color and this gene had already been detected by GWAS analysis and a deletion in this region has been hypothesized to control this trait (Sacco et al., 2015).

The utility of this germplasm collection

The present work has revealed a wide range of variability in our collection. The novelty of our study is the inclusion of a wide range of geographical origins of SLC accessions and SP from North Ecuador, which has not been widely studied. Moreover, the potential of Andean SLC is still not widely explored and it could be a novel source of interesting agronomic traits for tomato breeding. The genetic variability present in SLC from the Amazonian region is huge in comparison with the variability of the traditional tomato, although it has notably increased recently due to introgressions from wild species. The close phylogenetic relatedness of SLC makes this species especially useful for being exploited in tomato breeding, much more than other more phylogenetically distant species.

The high morphological variability found in our study may be an evidence of the potential variability in other traits not evaluated in this work. This collection is being analyzed for biochemical composition of fruit and deeper approaches for specific morphology analyses of fruit in the context of the Varitome project. The genome sequences of all these accessions are published and available, together with the identified annotated SNPs. The number of allelic variants present in this collection is huge and many of them may have interesting effects, such as lost or gained stop codons, frameshift variants or splice variants. A pan-genome analyses that includes the set of accessions we have used in our work, has described 4873 genes not present in the reference genome (Gao et al. 2019). Part of this gene variability is present in our collection and easily accessible. This increases its usefulness and makes our collection in one of the most characterized of tomato and related species.

The GWAS study has shown that the size and population structure of this collection make it feasible for this type of analysis. Further studies with other traits probably will increase the identification of candidate genes and alleles. Seeds from these sequenced

accessions are available from two genebanks, one in Europe (COMAV) and the other in America (TGRC). The characterized and sequenced plants came from a double round of self-pollination of a single plant, so they are quite homozygous, and it is possible to use the genotype data to do other GWAS analyses with other traits. Different segregating families have been developed and have led to the creation of a powerful tool to speed up genetic studies based on this collection. The use of three different parents in crosses with all accessions, allows the testing of the same alleles in different genetic backgrounds. The F2 populations will facilitate the analysis of the segregation of any variant in this collection. Studies can be conducted starting from SNP alleles, presence or absence of a determinate gene or phenotypic variants. By choosing an accession carrying a selected allele and one of the three parental accession which carries the alternative allele, the F1 and F2 segregation families are available for the genetic study of this variant. The availability of these segregating families allows speeding up research to confirm possible candidate genes. Besides sparing the effort of developing segregating families, researchers could analyze different natural mutants of the same gene or study the mutation effect in different genetic backgrounds.

The usefulness of this collection is based on the fact that all these resources are freely and easily available. Seeds of the original and self-pollinating accessions and F2 families are available at COMAV and TGRC genebanks. Passport and characterization data, pedigree information, genome sequences, SNPs and GWAs results are available and integrated at Solanaceae Genomics Network (solgenomics.net). All these resources build up a powerful platform for tomato genetics and breeding that could be reinforced with new studies performed on it.

Data availability

Sequences, SNPs, passport data, characterization data and images of the original collection are available in Solanaceae Genomics Network (<https://solgenomics.net/>). Seeds of the original germplasm collection are available by request to COMAV genebank (mdiezni@btc.upv.es) and to the TGRC (<https://tgrc.ucdavis.edu/>). Requests for the available seeds of F1 and F2 families should be addressed to COMAV genebank.

Acknowledgements: This research was supported by the National Natural Science Foundation of USA Varitome project (NSF IOS 1564366). We would like to thank the Centro de experiencias Cajamar de Paiporta (Valencia, Spain) for their excellent work done in growing the tomato plants in their greenhouses. We thank TGRC, ARS-GRIN and COMAV genebanks for the provided seeds and to all genebanks for their titan efforts to preserve biodiversity.

Authors' contributions: JC, MJD, JB and EVK conceived the experiment. EM, EG, VGC performed the research. JB, PZ, NM, LM did the data management. EM, JMP, PZ, VGC, JB, JC analyzed the data. JC, EM, MJD, EVK wrote the manuscript. All authors reviewed and approved this submission.

Ethics declarations

Conflict of interest: The authors declare that they have no conflict of interest.

REFERENCES

- Aflitos, S., Schijlen, E., De Jong, H., De Ridder, D., Smit, S., Finkers, R., Wang, J., Zhang, G., Li, N., Mao, L., Bakker, F., Dirks, R., Breit, T., Gravendeel, B., Huits, H., Struss, D., Swanson-Wagner, R., Van Leeuwen, H., Van Ham, R. C. H. J., Peters, S. (2014). Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *Plant Journal*, 80(1), 136–148. <https://doi.org/10.1111/tpj.12616>
- Alexander, L., & Hoover, M. (1955). Disease resistance in wild species of tomato: Report of the national screening Committee. *Agricultural Experimental Station Research Bulletin*, April, 7.
- Arellano Rodríguez, L. J., Rodríguez Guzmán, E., Ron Parra, J., Martínez Ramírez, J. L., Lozoya Saldaña, H., Sánchez Martínez, J., & Lépiz Ildelfonso, R. (2013). Evaluation of the resistance against *Phytophthora infestans* of wild populations of *Solanum lycopersicum* var. *cerasiforme*. *Revista Mexicana de Ciencias Agrícolas*, 4(5), 753–766. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-09342012000100010
- Bai, Y., & Lindhout, P. (2007). *Domestication and Breeding of Tomatoes: What have We Gained and What Can We Gain in the Future?* 1085–1094. <https://doi.org/10.1093/aob/mcm150>
- Banerjee, M. K., & Kalloo, M. K. (1987). Sources and inheritance of resistance to leaf curl virus in *Lycopersicon*. *Theoretical and Applied Genetics*, 73(5), 707–710. <https://doi.org/10.1007/BF00260780>
- Bauchet, G., Grenier, S., Samson, N., Bonnet, J., Grivet, L., & Causse, M. (2017). Use of modern tomato breeding germplasm for deciphering the genetic control of agronomical traits by Genome Wide Association study. *Theoretical and Applied Genetics*, 130(5), 875–889. <https://doi.org/10.1007/s00122-017-2857-9>
- Bauchet, G., Grenier, S., Samson, N., Segura, V., Kende, A., Beekwilder, J., Cankar, K., Gallois, J. L., Gricourt, J., Bonnet, J., Baxter, C., Grivet, L., & Causse, M. (2017). Identification of major loci and genomic regions controlling acid and volatile content in tomato fruit: implications for flavor improvement. *New Phytologist*, 215(2), 624–641. <https://doi.org/10.1111/nph.14615>
- Blanca, J., Cañizares, J., Cordero, L., Pascual, L., Díez, M. J., & Nuez, F. (2012). Variation Revealed by SNP Genotyping and Morphology Provides Insight into the Origin of the Tomato. *PLoS ONE*, 7(10). <https://doi.org/10.1371/journal.pone.0048198>
- Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., Francis, D., Causse, M., van der Knaap, E., & Cañizares, J. (2015). Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics*, 257. <https://doi.org/10.1186/s12864-015-1444-1>
- Brasil, J. N., Cabral, L. M., Eloy, N. B., Primo, L. M. F., Barroso-Neto, I. L., Grangeiro, L. P. P., Gonzalez, N., Inzé, D., Ferreira, P. C. G., & Hemerly, A. S. (2015). AIP1 is a novel Aget/Tudor domain protein from *Arabidopsis* that interacts with regulators of DNA

- replication, transcription and chromatin remodeling. *BMC Plant Biology*, 15(270), 270. <https://doi.org/10.1186/s12870-015-0641-z>
- Caicedo, A. L., & Schaal, B. A. (2004). Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. *Molecular Ecology*, 13(7), 1871–1882. <https://doi.org/10.1111/j.1365-294X.2004.02191.x>
- Capel, C., Fernández del Carmen, A., Alba, J. M., Lima-Silva, V., Hernández-Gras, F., Salinas, M., Boronat, A., Angosto, T., Botella, M. A., Fernández-Muñoz, R., Granell, A., Capel, J., & Lozano, R. (2015). Wide-genome QTL mapping of fruit quality traits in a tomato RIL population derived from the wild-relative species *Solanum pimpinellifolium* L. *Theoretical and Applied Genetics*, 128(10), 2019–2035. <https://doi.org/10.1007/s00122-015-2563-4>
- Cingolani, P., Platts, A., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, 6(2), 80–92.
- Foolad, M. R. (2007). Genome mapping and molecular breeding of tomato. *International Journal of Plant Genomics*, 2007. <https://doi.org/10.1155/2007/64358>
- Frary, A., Nesbitt, T. C., Frary, A., Grandillo, S., Cong, B., Liu, J., Meller, J., Elber, R., Alpert, K. B., & Tanksley, S. D. (2000). *fw2.2*: A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size. 289(July), 85–88. <https://doi.org/10.1126/science.289.5476.85>
- Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., Burzynski-chang, E. A., Fish, T. L., Stromberg, K. A., Sacks, G. L., Thannhauser, T. W., Foolad, M. R., Diez, M. J., Blanca, J., Canizares, J., Xu, Y., van der Knaap, E., Huang, S., Klee, H. J., Fei, Z. (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nature Genetics*, 51, 1044–1051. <https://doi.org/10.1038/s41588-019-0410-2>
- García-Martínez, S., Andreani, L., Garcia-Gusano, M., Geuna, F., & Ruiz, J. J. (2006). Evaluation of amplified fragment length polymorphism and simple sequence repeats for tomato germplasm fingerprinting: utility for grouping closely related traditional cultivars. *Genome*, 49(6), 648–656. <https://doi.org/10.1139/g06-016>
- Georgiady, M. S., Whitkus, R. W., & Lord, E. M. (2002). Genetic Analysis of Traits Distinguishing Outcrossing and Self-Pollinating Forms of Currant Tomato, *Lycopersicon pimpinellifolium* (Jusl.) Mill. *Genetics*, 161(1), 333–344.
- Gogarten, S. M., Sofer, T., Chen, H., Yu, C., Brody, J. A., Thornton, T. A., Rice, K. M., & Conomos, M. P. (2019). Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics*, July, 1–3. <https://doi.org/10.1093/bioinformatics/btz567>
- Grandillo, S., Ku, H. M., & Tanksley, S. D. (1999). Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theoretical and Applied Genetics*, 99(6), 978–987. <https://doi.org/10.1007/s001220051405>
- Holland, N., Holland, D., Helentjaris, T., Dhugga, K. S., Xoconostle-Cazares, B., & Delmer,

- D. P. (2000). A Comparative Analysis of the Plant Cellulose Synthase (CesA) Gene Family. *Plant Physiology*, 123(4), 1313–1323.
- Ichihashi, Y., & Sinha, N. R. (2014). From genome to phenome and back in tomato. In *Current Opinion in Plant Biology* (Vol. 18, Issue 1, pp. 9–15). Elsevier Ltd. <https://doi.org/10.1016/j.pbi.2013.12.004>
- IPGRI. (1996). *Descriptors for Tomato (Lycopersicon spp.)*.
- Joazeiro, C. A. P., & Weissman, A. M. (2000). RING Finger Proteins: Mediators of Ubiquitin Ligase Activity. *Cell*, 102(5), 549–552. [https://doi.org/10.1016/S0092-8674\(00\)00077-5](https://doi.org/10.1016/S0092-8674(00)00077-5)
- Li, M. X., Yeung, J. M. Y., Cherny, S. S., & Sham, P. C. (2012). Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human Genetics*, 131(5), 747–756. <https://doi.org/10.1007/s00439-011-1118-2>
- Meskauskiene, R., Nater, M., Goslings, D., Kessler, F., Op den Camp, R., & Apel, K. (2001). FLU: A negative regulator of chlorophyll biosynthesis in arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, 98(22), 12826–12831. <https://doi.org/10.1073/pnas.221252798>
- Monforte, A. J., Diaz, A., Caño-Delgado, A., & Van Der Knaap, E. (2014). The genetic basis of fruit morphology in horticultural crops: Lessons from tomato and melon. In *Journal of Experimental Botany* (Vol. 65, Issue 16, pp. 4625–4637). <https://doi.org/10.1093/jxb/eru017>
- Mu, Q., Huang, Z., Chakrabarti, M., Illa-Berenguer, E., Liu, X., Wang, Y., Ramos, A., & van der Knaap, E. (2017). Fruit weight is controlled by Cell Size Regulator encoding a novel protein that is expressed in maturing tomato fruits. *PLoS Genetics*, 13(8), 1–26. <https://doi.org/10.1371/journal.pgen.1006930>
- Nuez, F., & Díez, M. J. (2008). Tomato. In J. Prohens & F. Nuez (Eds.), *Vegetables II. Handbook of Plant Breeding* (pp. 249–323). Springer. https://doi.org/10.1007/978-0-387-74110-9_7
- Paran, I., & van der Knaap, E. (2007). Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *Journal of Experimental Botany*, 58(14), 3841–3852. <https://doi.org/10.1093/jxb/erm257>
- Peralta, I. E., Spooner, D. M., & Knapp, S. (2008). Taxonomy of Wild Tomatoes and Their Relatives (Solanum sect. Lycopersicoides, sect. Juglandifolia, sect. Lycopersicon; Solanaceae). *Systematic Botany Monographs*, 84, 1–186.
- Rambla, J. L., Medina, A., Fernández-Del-Carmen, A., Barrantes, W., Grandillo, S., Cammareri, M., López-Casado, G., Rodrigo, G., Alonso, A., García-Martínez, S., Primo, J., Ruiz, J. J., Fernández-Muñoz, R., Monforte, A. J., & Granell, A. (2017). Identification, introgression, and validation of fruit volatile QTLs from a red-fruited wild tomato species. *Journal of Experimental Botany*, 68(3), 429–442. <https://doi.org/10.1093/jxb/erw455>

- Razifard, H., Ramos, A., Della Valle, A. L., Bodary, C., Goetz, E., Manser, E. J., Li, X., Zhang, L., Visa, S., Tieman, D., van der Knaap, E., & Caicedo, A. L. (2020). Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution*, 1–15. <https://doi.org/10.1093/molbev/msz297>
- Rick, C. M. (1958). The role of natural hybridization in the derivation of cultivated tomatoes of western South America. *Economic Botany*, 12, 346–367.
- Rick, C. M. (1978). Potential improvement of tomatoes by controlled introgression of genes from wild species. *Proceedings of the Conference Broadening the Genetic Base of Crops*, 167–176.
- Rick, C. M., & Holle, M. (1990). Andean *Lycopersicon esculentum* var. *cerasiforme*: genetic variation and its evolutionary significance. *Economic Botany*, 44(3 Supplement), 69–78. <https://doi.org/10.1007/BF02860476>
- Rick, C. M., & Fobes, J. F. (1975). Allozyme Variation in the Cultivated Tomato and Closely Related Species. *Bulletin of the Torrey Botanical Club*, 102(6), 376–384.
- Rick, C. M., Fobes, J. F., & Holle, M. (1977). Genetic Variation in *Lycopersicon pimpinellifolium*: Evidence of Evolutionary Change in Mating Systems *. *Plant Systematics and Evolution*, 127, 139–170.
- Robertson, L., & Labate, J. (2014). *Genetic Improvement of Solanaceous Crops Volume 2: Tomato* (M. K. Razdan & A. K. Mattoo (eds.); Vol. 2). Science Publishers. <https://doi.org/10.1201/b10744-3>
- Rodríguez, G. R., Muños, S., Anderson, C., Sim, S. C., Michel, A., Causse, M., Gardener, B. B. M., Francis, D., & van der Knaap, E. (2011). Distribution of SUN, OVATE, LC, and FAS in the Tomato Germplasm and the Relationship to Fruit Shape Diversity. *Plant Physiology*, 156(1), 275–285. <https://doi.org/10.1104/pp.110.167577>
- Rothan, C., Diouf, I., & Causse, M. (2019). Trait discovery and editing in tomato. *Plant Journal*, 97(1), 73–90. <https://doi.org/10.1111/tpj.14152>
- Sacco, A., Ruggieri, V., Parisi, M., Festa, G., Rigano, M., Picarella, M. E., Mazzucato, A., & Barone, A. (2015). Exploring a Tomato Landraces Collection for Fruit-Related Traits by the Aid of a High- Throughput Genomic Platform. *PLoS ONE*, 10. <https://doi.org/10.1371/journal.pone.0137139>
- Sifres, A., Picó, B., Blanca, J. M., De Frutos, R., & Nuez, F. (2006). Genetic structure of *Lycopersicon pimpinellifolium* (Solanaceae) populations collected after the ENSO event of 1997-1998. *Genetic Resources and Crop Evolution*, 54(2), 359–377. <https://doi.org/10.1007/s10722-005-5725-4>
- Soyk, S., Lemmon, Z. H., Oved, M., Fisher, J., Liberatore, K. L., Park, S. J., Goren, A., Jiang, K., Ramos, A., van der Knaap, E., Van Eck, J., Zamir, D., Eshed, Y., & Lippman, Z. B. (2017). Bypassing Negative Epistasis on Yield in Tomato Imposed by a Domestication Gene. *Cell*, 169(6), 1142–1155. <https://doi.org/10.1016/j.cell.2017.04.032>

- Stevens, M. A., & Rick, C. M. (1986). Genetics and Breeding. In J. Atherton & J. Rudich (Eds.), *The Tomato Crop: a scientific basis for improvement* (p. 661). Chapman and Hall Ltf. <https://doi.org/10.1007/978-94-009-3137-4>
- Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., Rambla, J. L., Beltran, K. S. O., Taylor, M., Zhang, B., Ikeda, H., Liu, Z., Fisher, J., Zemach, I., Monforte, A., Zamir, D., Granell, A., Kirst, M., Huang, S., & Klee, H. (2017). A chemical genetic roadmap to improved tomato flavor. *Science (New York, N.Y.)*, 355(6323), 391–394. <https://doi.org/10.1126/science.aal1556>
- Walter, J. M. (1967). Hereditary Resistance to Disease in Tomato. *Annual Review of Phytopathology*, 5(1), 131–160.
- Warnock, S. J. (1991). Natural Habitats of Lycopersicon Species. *HortScience*, 26(5), 466–471. <https://doi.org/10.21273/hortsci.26.5.466>
- Widrechner, M. P. (1987). Variation in Breeding System of Lycopersicon pimpinellifolium: Implications for Germplasm Maintenance. *Plant Genetic Resources Newsletter*, 70, 38–43.
- Williams, C. E., & ST. Clair, D. A. (1993). Phenetic relationships and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accessions of Lycopersicon esculentum. *Genome*, 36(3), 619–630. <https://doi.org/10.1139/g93-083>
- Xu, C., Liberatore, K. L., Macalister, C. A., Huang, Z., Chu, Y. H., Jiang, K., Brooks, C., Ogawa-Ohnishi, M., Xiong, G., Pauly, M., Van Eck, J., Matsubayashi, Y., Van Der Knaap, E., & Lippman, Z. B. (2015). A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nature Genetics*, 47(7), 784–792. <https://doi.org/10.1038/ng.3309>
- Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. (2011). GCTA: a tool for genomewide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82.
- Yoo, H. J., Park, W. J., Lee, G. M., Oh, C. S., Yeam, I., Won, D. C., Kim, C. K., & Lee, J. M. (2017). Inferring the Genetic Determinants of Fruit Colors in Tomato by Carotenoid Profiling. *Molecules*, 22(5), 1–14. <https://doi.org/10.3390/molecules22050764>
- Zhang, H., Xia, R., Meyers, B. C., & Walbot, V. (2015). Evolution, functions, and mysteries of plant ARGONAUTE proteins. *Current Opinion in Plant Biology*, 27, 84–90. <https://doi.org/10.1016/j.pbi.2015.06.011>
- Zhao, J., Sauvage, C., Zhao, J., Bitton, F., Bauchet, G., Liu, D., Huang, S., Tieman, D. M., Klee, H. J., & Causse, M. (2019). Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. *Nature Communications*, 10(1), 1–12. <https://doi.org/10.1038/s41467-019-09462-w>
- Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., Lin, T., Qin, M., Peng, M., Yang, C., Cao, X., Han, X., Wang, X., van der Knaap, E., Zhang, Z., Cui, X., Klee, H., Fernie, A. R., Luo, J., & Huang, S. (2018). Rewiring of the Fruit Metabolome in Tomato Breeding. *Cell*, 172(1–2), 249–261.e12. <https://doi.org/10.1016/j.cell.2017.12.019>

Zuriaga, E., Blanca, J. M., Cordero, L., Sifres, A., Blas-Cerdán, W. G., Morales, R., & Nuez, F. (2009). Genetic and bioclimatic variation in *Solanum pimpinellifolium*. *Genetic Resources and Crop Evolution*, 56(1), 39–51. <https://doi.org/10.1007/s10722-008-9340-z>

Capítulo 4:

DISCOVERY OF A MAJOR TLS CONTROLLING TRICHOME IV DENSITY IN TOMATO USING K-SEQ GENOTYPING

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-Paez ¹, Ana García-Pérez ¹, Peio Ziarsolo ¹, José Blanca ¹, Esther van der Knapp ^{3,4}, María José Díez ¹ and Joaquín Cañizares ^{1,*}

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

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

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

⁴ Department of Horticulture University of Georgia. GA. USA

‡ Both authors contributed equally

* Correspondence: jcanizares@upv.es

Referencia completa:

Mata-Nicolás, E., Montero-Pau, J., Gimeno-Paez, E., García-Pérez, A., Ziarsolo, P., Blanca, J., Van Der Knaap, E., Díez, M. J., & Cañizares, J. (2021). Discovery of a Major QTL Controlling Trichome IV Density in Tomato Using K-Seq Genotyping. *Genes* 2021, 12, 243. <https://doi.org/10.3390/genes12020243>

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. Therefore, for sustainable management of this crop, breeding programs could incorporate some natural defense mechanisms, such as those afforded by trichomes, present in certain *Solanum* species. We have identified a *S. pimpinellifolium* accession with very high density of this type of trichomes. This accession was crossed with a *S. lycopersicum* var. *cerasiforme* and a *S. lycopersicum* var. *lycopersicum* accessions, and the two resulting F2 populations have been characterized and genotyped using a new genotyping methodology, K-seq. We have been able to build an ultra-dense genetic map with 147,326 SNP 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*. This QTL could be easily introgressed in cultivated tomato due to the close genetic relationship between both species.

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

INTRODUCTION

Agricultural production is greatly affected by pests and diseases related to yield and quality. Successful production is, therefore, 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 (Woodcock et al., 2019). As a consequence, new approaches are being sought in agricultural management programs to provide sustainable alternatives to pesticide applications (Glas et al., 2012). Recent efforts to enhance sustainability have focused on the development of crops with genetic resistance to various pests and diseases.

The presence of trichomes is an important morphological defence against insect pests (Peter, Shanower, and Romeis 1995; Huchelmann, Boutry, and Hachez 2017). Specifically, glandular trichomes that accumulate and exude secondary metabolites, are known to interfere with oviposition, larval fixation and development, while also trapping or poisoning insects (Liedl et al. 1995; Peiffer et al. 2009; Glas et al. 2012). The type of trichomes and stored compound depends on the species and could include acyl sugars, methyl ketones, terpenoids, phenylpropenes or flavonoids (Leong et al., 2019). Breeding for resistant cultivars with an increased density of glandular trichomes is a valuable option to reduce the impact of arthropod pests. However, this strategy requires a detailed knowledge of the genetic control and molecular basis of trichome development.

Trichomes are a common morphological defence in the genus *Solanum* (Cho et al., 2017; McDowell et al., 2011). Eight types of trichomes have been described on the stems of plants of this genus (Luckwill 1943) type I, IV, VI and VII are glandular trichomes. These glandular trichomes accumulate different compounds -Type I and IV, acyl sugars; type VI, monoterpenes and sesquiterpenes and type VII, alkaloids--and provide different levels of protection against pests (Glas et al., 2012). In particular, the presence and density of type IV glandular trichomes --trichomes with a short multicellular stalk and small gland at the tip (Channarayappa et al., 1992) -- has been associated with pest resistance such as spider mite (Maluf, Campos, & Das Graças Cardoso 2001; Alba, Montserrat, & Fernández-Muñoz 2009), whitefly (Andrade et al., 2017; S Firdaus, 2012; Oriani et al., 2010; Rakha et al., 2017) and potato moth (Gurr & McGrath, 2001b). Moreover, resistance against whiteflies and aphids has been associated with the production of acyl sugars in this type of trichome (Fan et al., 2019). Cultivated tomato (*Solanum lycopersicon* var. *lycopersicon* L.) typically lacks or has a very low density of type IV trichomes (Glas et al., 2012). Thus, specific breeding programs that will incorporate these natural defences that are common within the *Solanum* genus, may improve crop management against a host of arthropod pests.

Several wild relatives of tomato have shown 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 display trichome densities higher than 90 trichomes/mm² (Guo, Weston, and Snyder 1993; Blauth, Churchill, and Mutschler 1998; Andrade et al. 2017). Inheritance of type IV trichome density has been studied using interspecific crosses among these related species and tomato. The presence of trichome type IV on *S. pennellii* has been proposed to be under the control of two dominant unlinked genes (Lemke & Mutschler, 1984). Andrade et al. (2017) found that inheritance of trichome type IV on *S. galapagense* is relatively simple and could be controlled by incomplete 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 on chromosome 2 of *S. pennellii* (Blauth et al., 1998), in chromosome 2 and 9 of *S. habrochaites* (Momotaz et al., 2010) and *S. galapagense* (Firdaus et al. 2013). Synthesis of acyl sugars has also been associated with QTLs in populations derived from *S. pennellii*. These QTLs were located on chromosomes 2, 3, 4, 5, 6 and 11 of *S. pennellii* (Mutschler et al. 1996; Blauth, Churchill, and Mutschler 1998; Leckie, de Jong, and Mutschler 2012).

Despite the fact that these wild species are described as a potential source of type IV trichomes density, their use in breeding programs is limited because the sexual incompatibility hinders their introgression and due to linkage drag resulting in undesirable agronomic traits such lower fruit size or higher morphological heterogeneity in addition to increased type IV trichome density. As a consequence, a more closely related wild species like *Solanum pimpinellifolium* L., a red-fruited, facultative self-compatible species is generally preferred in tomato breeding programs as excessive linkage drag can be avoided 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. Rakha, Hanson, and Ramasamy (2017) described accessions that were characterized by densities lower than 4.5 trichomes/mm². Fernández-Muñoz et al. (2003) found a *S. pimpinellifolium* (TO-937) accession characterized by a mean of 16.3 type IV trichomes/mm² in the abaxial surface, and Firdaus et al. (2012) found a *S. pimpinellifolium* (LA1584) accession with 21.0 trichomes/mm². Interestingly, a population derived from the accession TO-937 confirmed two QTLs in chromosome 2 conferring resistance to spider mites (Salinas et al., 2013).

Trichome presence and density is also known to depend on developmental factors such as plant and leaf age. Vendemiatti et al. (2017) 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) (Gurr & McGrath, 2001a). F2 and F1 populations between *Solanum lycopersicum* and *S. pennellii* also exhibited higher densities in plants at 10 and 13 weeks after transplanting, compared to 7 weeks aged plants (Saeidi et al., 2007). However, Wilkens et al. (1996) 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, compared to 14-15 hours (Weston et al., 1989).

The main objective of this study was to map the quantitative trait loci controlling type IV trichome density in a *S. pimpinellifolium* accession (BGV016047) that appears to present the highest density of type IV trichome (> 100 trichomes/mm²) to be described for this species. Populations derived from this accession using two different genetic backgrounds have been used to map the loci controlling the presence and density of type IV trichomes on chromosomes 9 and 11. We demonstrate the practical use of K-seq, a new genotyping technology based on the amplification with the Klenow polymerase of genomic regions with short oligonucleotides, followed by standard PCR and Illumina sequencing (Ziarsolo et al., 2020).

MATERIALS AND METHODS

Plant material

During the characterization of the Varitome collection (<https://solgenomics.net/projects/varitome>) of *S. pimpinellifolium* (27 accessions), *S. lycopersicum* var. *cerasiforme* (121) and *S. lycopersicum* var. *lycopersicum* (15) from Perú, Ecuador and México (Mata-Nicolás et al., 2020), an accession of *S. pimpinellifolium* with a high density of glandular type IV trichomes was found (BGV016047).

To study the genetic basis of trichome IV density, we have used the F1 and F2 collections developed by the Varitome project (Mata-Nicolás et al., 2020). BGV016047 (also coded as PAS014479 in Varitome) was crossed with two different genetic backgrounds with low trichome density: *Solanum lycopersicum* var. *cerasiforme* LA2278 (LA2278 x BGV016047, referred hereafter as SLC family) and *Solanum lycopersicum* var. *lycopersicum* cv. MoneyMaker (MoneyMaker x BGV016047, referred hereafter as SLL family). F1 populations were developed using BGV016047 as the male parent in both families, and the F2 generation was obtained by self-pollination. BGV016047 was chosen as the male parent because flowers of *S. pimpinellifolium* are smaller and more difficult to emasculate. 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 available at the genebank of Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV) at the Polytechnic University of Valencia and at the Tomato Genetics Resource Center (TGRC, USA).

Glandular type IV trichome density characterization

To evaluate the effect of plant age on glandular type IV trichome density in 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 were grown in the greenhouse and fertigated with a nutrient solution containing 14 mM NO₃⁻, 1 mM H₂PO₄⁻, 2 mM SO₄²⁻, 1 mM NH₄⁺, 16 mM K⁺, 4 mM Ca²⁺, 2mM Mg²⁺, 15 µM Fe²⁺, 10 µM Mn²⁺, 5 µM Zn²⁺, 30 µM B³⁺, 0.75 µM Cu²⁺ and 0.6 Mo⁶⁺ (Maynard & Hochmuth, 2006). 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). 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 (Alphasot YS-2, Nikkon) 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 (Supplementary Fig. 1- ANEXO C)

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, starting when the leaves were in the second position from the apex (at 49, 53 and 60 days after transplanting, respectively). Type IV trichome density was measured as described above.

Trichome density was evaluated in 195 F2, 50 BC1 and controls (24 BGV016047, 10 MM, 16 F1) for the SLL family and in 151 F2 and controls (10 BGV016047 and 10 LA2278) for the SLC family. 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 F2 individuals were genotyped out of the 151 that were phenotyped. Additionally, 15 parental controls of this family were also 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 CTAB protocol (Doyle & Doyle, 1990).

Genotyping was performed using K-seq, a novel genotyping method based on the reduction of genome complexity using Klenow polymerase and short oligonucleotides (Ziarsolo et al., 2020). The genotyping was done as described in Ziarsolo et al. (2020) but using 25°C as annealing temperature for the Klenow polymerase amplification. A total of 200 ng of DNA of each individual 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 products of each individual 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 (Andrews et al., 2012) and low quality reads and adapters were removed using Trimmomatic v. 0.36 (Bolger et al., 2014). Clean reads were mapped against *Solanum lycopersicum* genome build SL2.5 (Shearer et al., 2014) using BWA-MEM v. 0.7.17 (Li 2013). SNP calling was performed using freebayes v 1.3.1 (Garrison & Marth, 2012) and the SNPs with more than 50% of missing data and a minor allele frequency lower than 0.05 were filtered out those using custom Python scripts that are publicly available at <https://github.com/bioinfcomav>.

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 (Takagi et al., 2013) 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 (Miao et al., 2018)

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* (Broman et al., 2003) and *ASMap* (Taylor & Butler, 2014) 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* (Venables & Ripley, 2002). Best fitting distribution for trichome density was found using the R package *fitdistrplus* (Delignette-Muller & Dutang, 2015). 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 “RsquaredGLM” from the *modEVA* package (Barbosa et al., 2016).

The QTL-seq was done by bulking the genotype information of F₂ individuals with the lowest and the highest type IV trichome average 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 an average 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. For each bulk, a SNP-index per loci was calculated as the percentage of reads containing a different allele from the reference genome. The delta (Δ) SNP-index was calculated by subtracting the SNP-indices of the bulks at each loci [44], 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 package (Mansfeld & Grumet, 2018).

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 (Yekutieli & Benjamini, 2001) in R software v.3.6 (R Core Team, 2015).

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 of the experiment (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) (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).

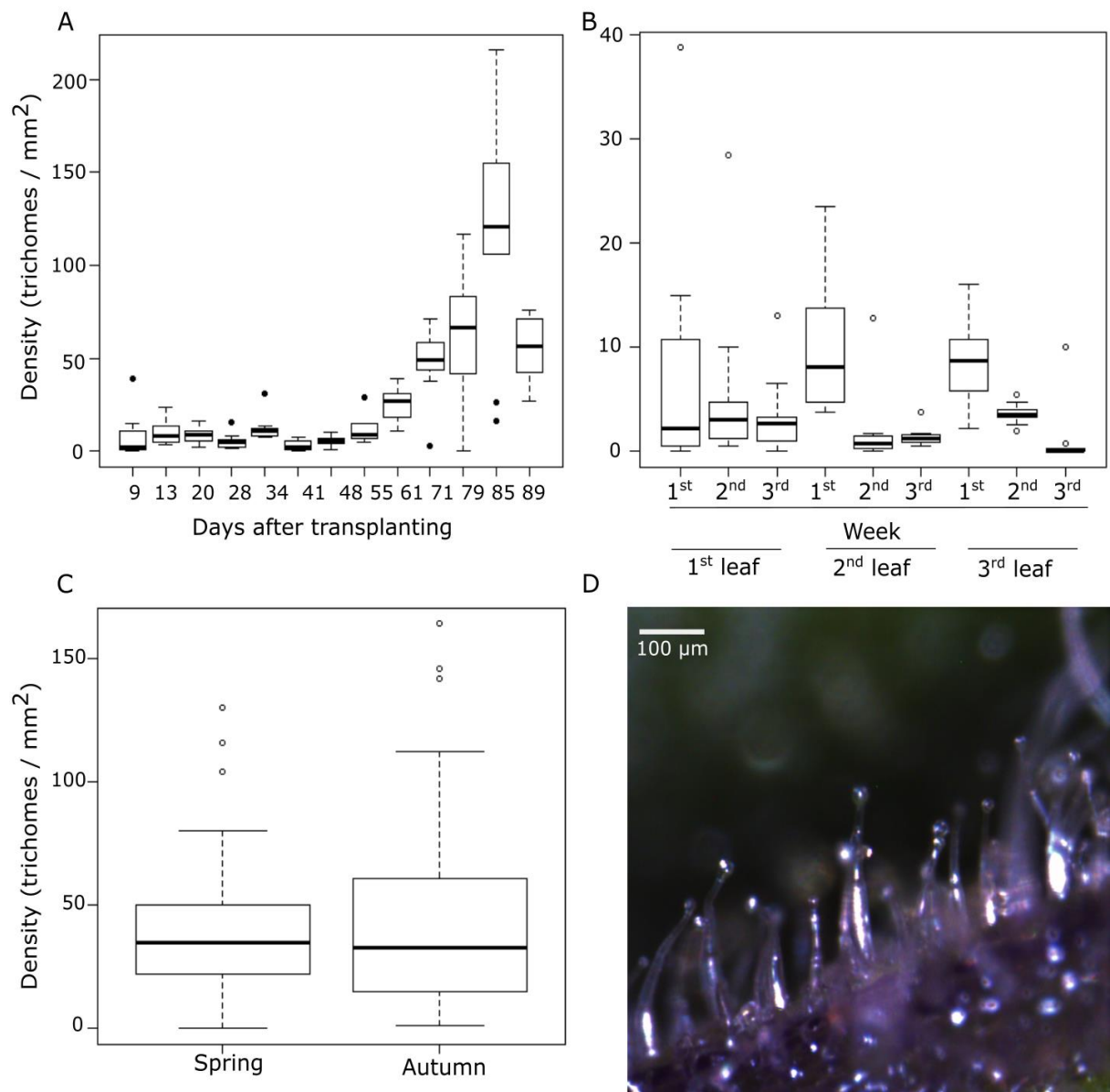


Figure 1. Effect of plant age (A), leaf age (B) and growing season (C) on type IV trichome density. Photography of trichomes type IV (D).

Trichome type IV density on segregating populations

The *S. pimpinellifolium* accession BGV016047 was crossed with low trichome density *S. lycopersicum* var. *cerasiforme* accession LA2278 and *S. lycopersicum* var. *lycopersicum* ‘MoneyMaker’ (Supplementary File 2- ANEXO C). Figure 2 shows the density distributions of the different populations derived from these crosses. The low trichome density parents exhibit either none or a very low trichome density, whereas the F2 for both families showed values that were skewed towards the low density parent. However, in some cases some F2 reached or even exceeded the densities of BGV016047. F2 individuals of the SLL family (Fig. 2B) showed a higher average density than the corresponding F1 (5.5 vs. 3.6) but a lower median (1 vs. 2.5), indicating a distribution shifted towards lower or higher values than the F1. 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 (Supplementary Fig. 2-ANEXO C).

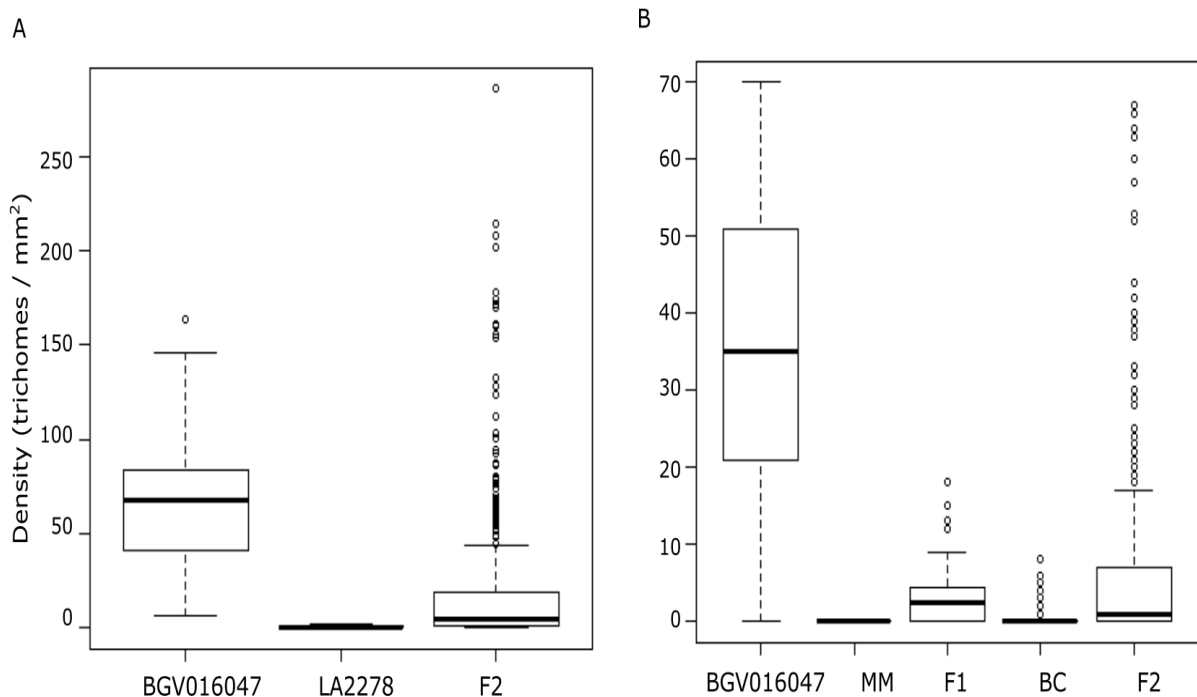


Figure 2. Density of type IV trichomes in the segregating populations for the last week of phenotyping for the SLC family (BGV016047 x LA2278) (A) and SLL family (BGV016047 x MoneyMaker) (B).

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 of each sample and not quantifying them. In Supplementary Figures 3, 4 and 5 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 (Shearer et al., 2014) 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 removing SNPs with a minor allele frequency lower than 5%, 4,136,452 SNPs were kept and 308,400 were finally retained after removing those with 50% of missing data. Out of those, 147,326 SNPs could be transformed to ABH coding. 107 F2 samples were finally used to build the genetic map for the F2 population (Supplementary File 3-ANEXO C). Adjacent SNPs with the same parental genotype were binned in 8,285 bin markers. Markers were grouped in 12 linkage groups (Table 1 and Supplementary Fig. 6-ANEXO C) 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 133.4 cM per chromosome with an average distance 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), so a homozygote bearing the two SP alleles has a 9 fold expected density. 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).

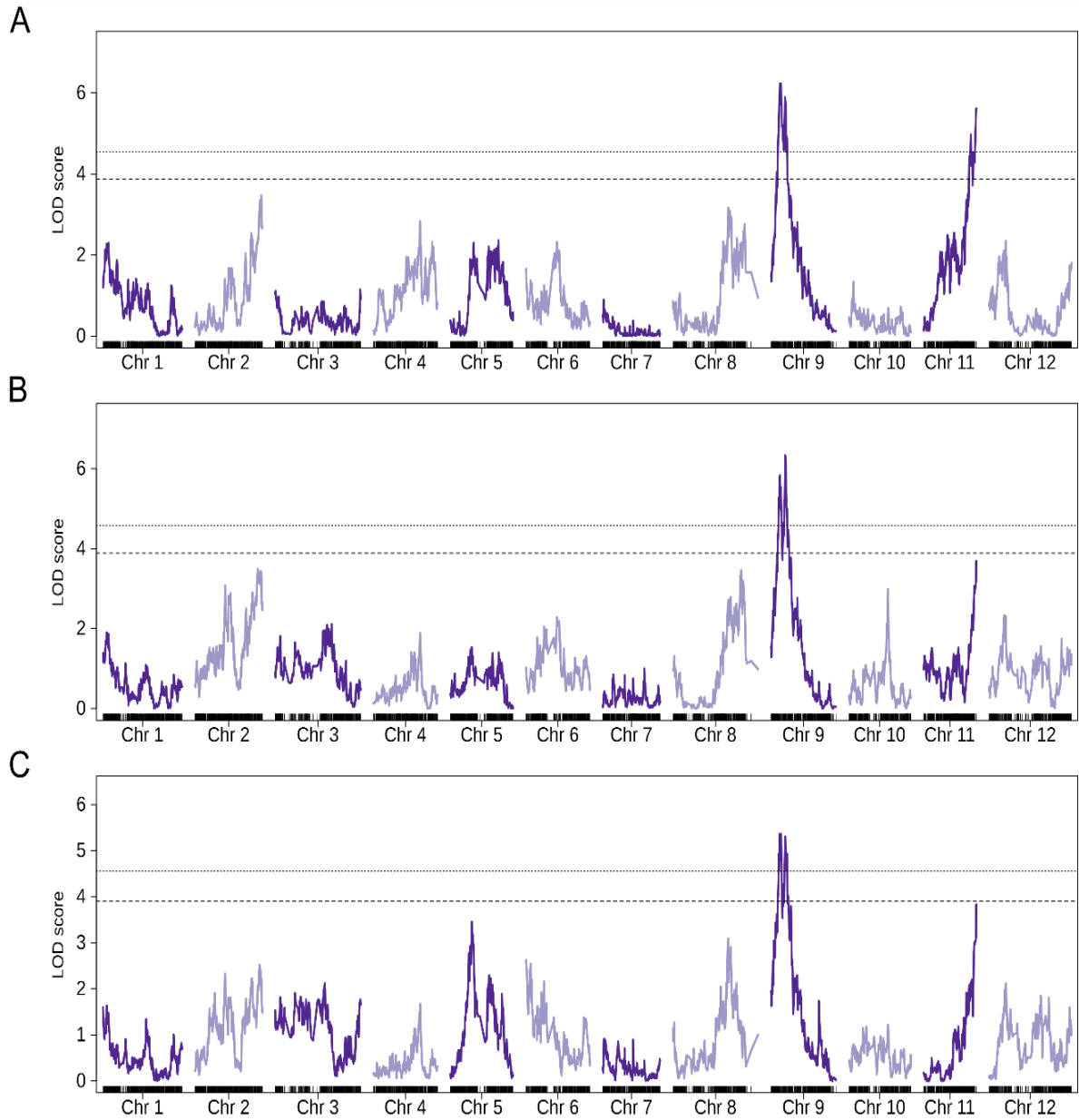


Figure 3. Results of non-parametric single interval mapping QTL analysis for each of the three weeks phenotyped. A) 6th week after transplanting, B) 7th week and C) 8th week.

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

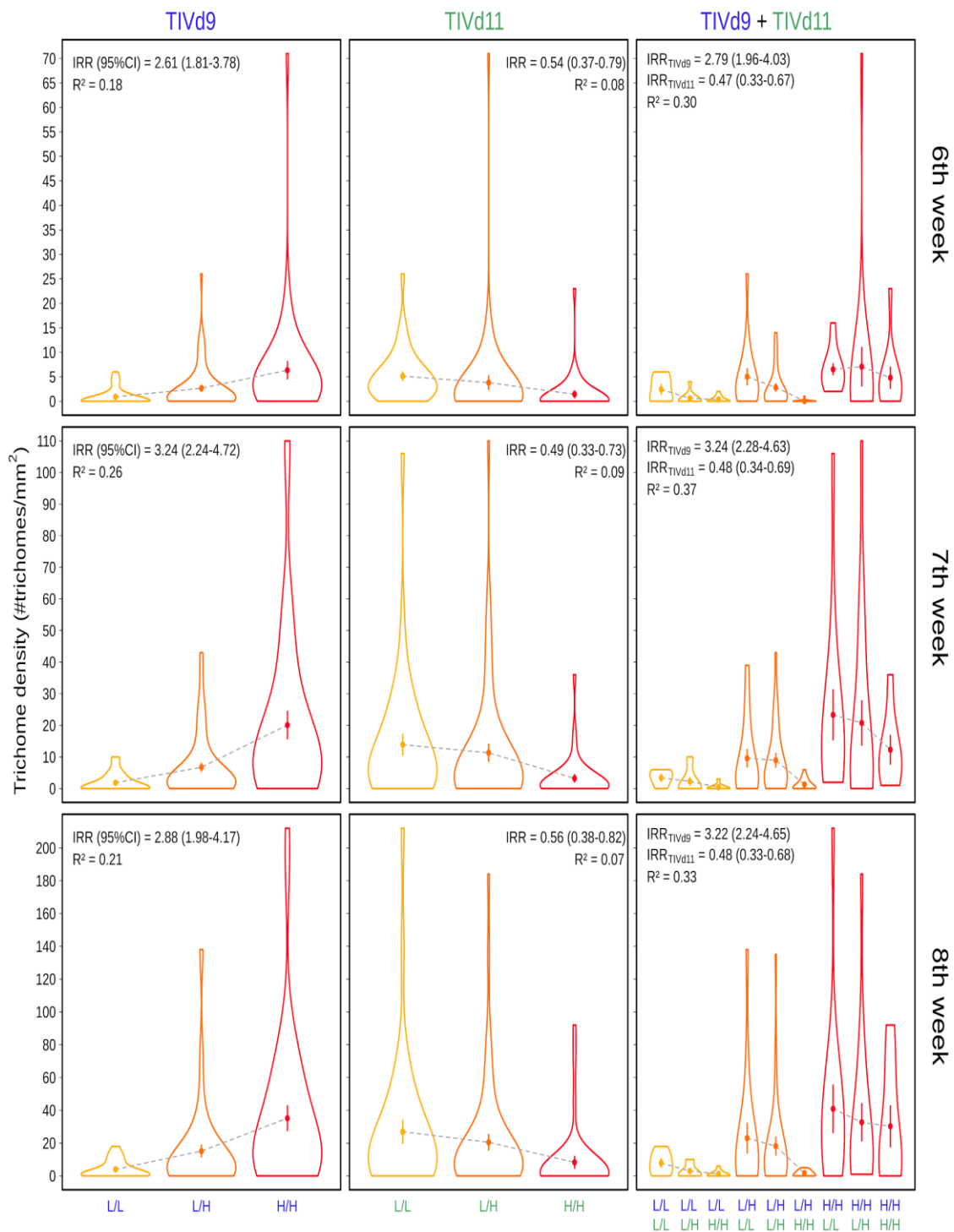


Figure 4. Effect of QTLs TIVd9, TIVd11 and the additive model on trichome density along the three phenotyped weeks (6th, 7th, 8th weeks after transplanting. Incidence rate ratio (IRR) and 95% confidence intervals (values between brackets) and Nagelkerke R² of the negative binomial logistic regression 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.

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 parent at the marker with the highest statistical support. For each model, χ^2 goodness of fit 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)	***

Table 3. Continuation.

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 an explanatory factor (Table 2 and Supplementary Fig. 7-ANEXO C). 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 (Supplementary Fig. 7-ANEXO C). 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.

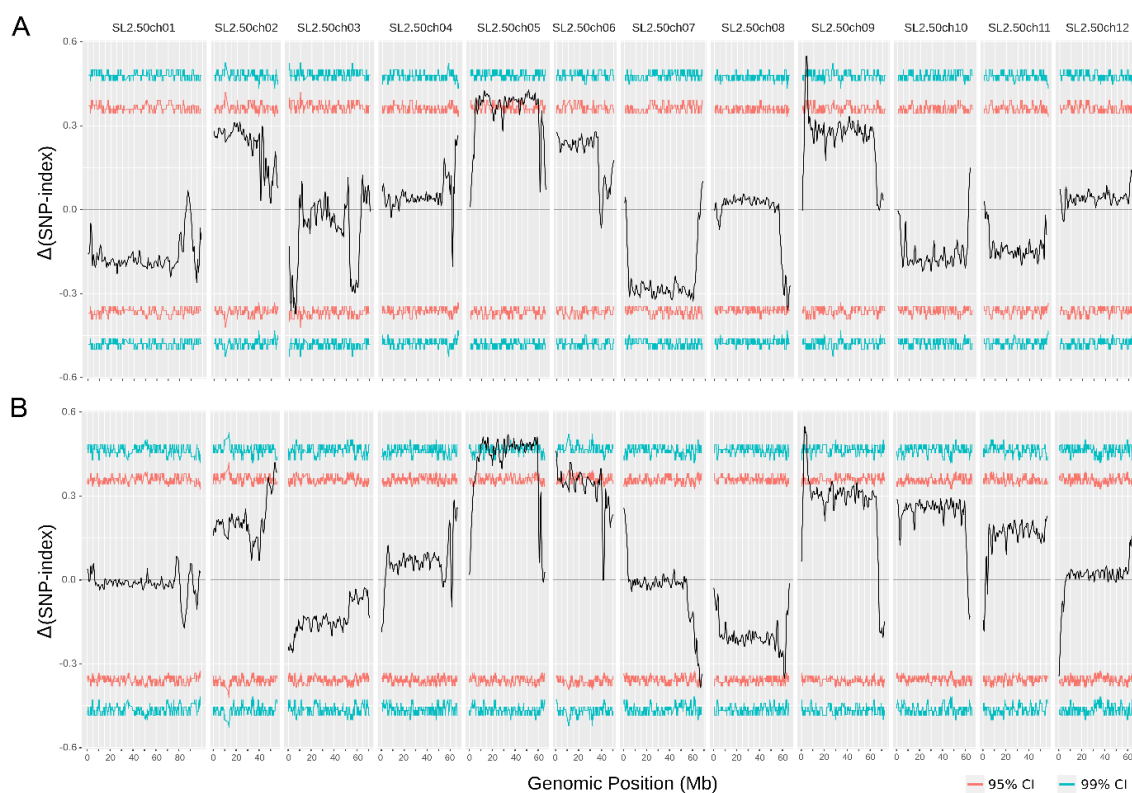


Figure 5. Distribution of Δ (SNP-index) along the chromosomes based on 1Mb windows for the A) SLC family (BGV016047 x LA2278) and B) SLL family (BGV016047 x MoneyMaker). The 95% and 99% confidence thresholds are shown.

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 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 (Supplementary Table 1- *ANEXO C*), 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 (Potikha and Delmer 1995; Gao et al. 2017).

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 (Mandal et al., 2020): 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 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 (Chien & Sussex, 1996). 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) (Supplementary Table 1-ANEXO C). 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 (Mandal et al., 2020). Besides that, transcriptions factors including a Myb/SANT-like (Solyc05g018830) and two MADS-box transcription factors (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 arthropod pest resistance, which is a desirable trait for sustainable agriculture. 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 (Guo, Weston, & Snyder 1993; Blauth, Churchill, & Mutschler 1998; Firdaus et al. 2013; Andrade et al. 2017). 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. *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) (Firdaus et al. 2012; Rakha, Hanson, and Ramasamy 2017; Rodríguez-López, Moriones, and Fernández-Muñoz 2020). 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 (Gurr & McGrath, 2001a). 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) 12 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 (Gurr & McGrath, 2001a; Wilkens et al., 1996).

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 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 (Ziarsolo et al., 2020). 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 built an ultra-dense genetic map with 147,326 markers at an average distance between markers of 0.2 cM that 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 (Blauth, Churchill, & Mutschler 1998; Fernández-Muñoz et al. 2003; Leckie, de Jong, & Mutschler 2012; Smeda et al. 2016; 2018). 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 (Freitas et al. 2002; Andrade et al. 2017). In *S. pennelli* and *S. pimpinellifolium*, results supported two dominant unlinked genes (Fernández-Muñoz et al., 2003; Lemke & Mutschler, 1984; Saeidi et al., 2007). 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 making feasible their management in a breeding programme.

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 increased 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 (Salinas et al., 2013). 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%) (Vosman et al., 2019). Another region in chromosome 9, adjacent to TIVd9 and with minor effects on trichome density, has also been detected in *S. galapagense* (Firdaus et al. 2013). 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 (Momotaz et al., 2010). The QTL on chromosome 9 overlaps with our QTL TIVd9, although explains a lower percentage of the variance (11%) than in our study. Interestingly, this QTL is also associated with a reduced oviposition by *Bemisia tabaci*. The QTL on chromosome 11 does not overlap with our QTL as it was 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 is on chromosome 9 in *S. galapagense* explaining 8% of variance (Firdaus et al. 2013) and is located 61 Mb away from TIVd9. Therefore, it is unlikely that both QTLs overlap. Moreover, in a posterior study with RIL populations this QTL was not found (Vosman et al., 2019). However, the chromosomal region of TIVd9 does overlap with a QTL for adult survival to white fly (Firdaus et al. 2013). TIVd5 and TIVd6 overlap with previously detected QTLs for trichome density in *S. pennelli* (Maliepaard et al. 1995; Blauth, Churchill, & Mutschler 1998). Additionally, TIVd5 was also previously found to be associated with acylsugar accumulation levels (Blauth et al., 1998).

Several types of genes have been found to control the development of trichomes (Doroshkov et al., 2019; Fambrini & Pugliesi, 2019), including MYB transcription factors

that are known to regulate positively and negatively trichome development (Gan et al., 2011; Kirik et al., 2005), and genes related to the production of acylsugars (Mandal et al., 2020) 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 (Potikha and Delmer 1995; Gao et al. 2017). The different gene members are likely to provide a wide range of activities and tissue specificities (Gao et al. 2017). 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* (Ó'Maoiléidigh et al., 2013), which is the same effect found in our population.

The QTLs identified in this work, which increase the density of type IV trichomes and, probably, the synthesis of acylsugars, represent a valuable resource for the development of new tomato varieties. One of the main advantages is that these QTLs have been described in a *S. pimpinellifolium* accession, a species that has been widely and successfully used in tomato breeding and in introgression of alleles. Incorporation of insect resistance into commercial tomatoes is a key point in breeding programs to facilitate the cultivation of crops reducing the use of pesticides. Type IV trichomes could provide broad-spectrum resistance to pests with low impact on tomato fruit as trichomes are not present in fruits. The identification of these QTLs could facilitate the rapid development of new varieties through genetic marker-assisted selection and contribute to the identification of key genes in the control of trichome development.

CONCLUSIONS

The results presented in this study increase the knowledge of the genetic factors involved in trichomes density 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 creation of tomato lines with a higher density of type IV trichomes, and likely higher insect resistance (Rakha et al., 2017). Several efforts have been made to develop introgressed lines with type IV trichomes enriched in acylsugars (Leckie et al., 2012; Rodríguez-López et al., 2020; Smeda et al., 2016). We found a major QTL that has not been previously reported in *S. pimpinellifolium* that increases the density of trichomes by a factor of 9 if the alleles are in homozygosis. This accession contains probably different alleles of the QTLs detected in the other species. The main advantage of this accession is that this QTL could be easily introgressed 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 perform 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.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

- Alba, J. M., Montserrat, M., & Fernández-Muñoz, R. (2009). Resistance to the two-spotted spider mite (*Tetranychus urticae*) by acylsucroses of wild tomato (*Solanum pimpinellifolium*) trichomes studied in a recombinant inbred line population. *Experimental and Applied Acarology*, *47*(1), 35–47. <https://doi.org/10.1007/s10493-008-9192-4>
- Andrade, M. C., da Silva, A. A., Neiva, I. P., Oliveira, I. R. C., De Castro, E. M., Francis, D. M., & Maluf, W. R. (2017). Inheritance of type IV glandular trichome density and its association with whitefly resistance from *Solanum galapagense* accession LA1401. *Euphytica*, *213*(2). <https://doi.org/10.1007/s10681-016-1792-1>
- Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., & Wingett, S. (2012). *FastQC: a quality control tool for high throughput sequence data*.
- Barbosa, A. M., Brown, J. A., Jimenez-Valverde, A., & Real, R. (2016). *modEvA: Model Evaluation and Analysis*.
- Blauth, S. L., Churchill, G. A., & Mutschler, M. A. (1998). Identification of quantitative trait loci associated with acylsugar accumulation using intraspecific populations of the wild tomato, *Lycopersicon pennellii*. *Theoretical and Applied Genetics*, *96*(3–4), 458–467. <https://doi.org/10.1007/s001220050762>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15). <https://doi.org/10.1093/bioinformatics/btu170>
- Broman, K. W., Wu, H., Sen, S., & Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, *19*(7). <https://doi.org/10.1093/bioinformatics/btg112>
- Channarayappa, C., Shivashankar, G., Muniyappa, V., & Frist, R. H. (1992). Resistance of *Lycopersicon* species to *Bemisia tabaci*, a tomato leaf curl virus vector. *Can. J. Bot.*, *70*(Stoner 1970), 2184–2192. <https://doi.org/10.1139/b92-270>
- Chien, J. C., & Sussex, L. M. (1996). Differential Regulation of Trichome Formation on the Adaxial and Abaxial Leaf Surfaces by Gibberellins and Photoperiod in *Arabidopsis thaliana* (L.) Heynh.?. *Plant Physiol*, *111*, 1321–1328. <https://plantphysiol.org>
- Cho, K. S., Kwon, M., Cho, J. H., Im, J. S., Park, Y. E., Hong, S. Y., Hwang, I. T., & Kang, J. H. (2017). Characterization of trichome morphology and aphid resistance in cultivated and wild species of potato. *Horticulture Environment and Biotechnology*, *58*(5), 450–457. <https://doi.org/10.1007/s13580-017-0078-4>
- Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software*, *64*, 1–34.
- Doroshkov, A. V., Konstantinov, D. K., Afonnikov, D. A., & Gunbin, K. V. (2019). The evolution of gene regulatory networks controlling *Arabidopsis thaliana* L. trichome development. *BMC Plant Biology*, *19*(S1). <https://doi.org/10.1186/s12870-019-1640-2>

- Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, *12*, 13–15.
- Fambrini, M., & Pugliesi, C. (2019). The Dynamic Genetic-Hormonal Regulatory Network Controlling the Trichome Development in Leaves. *Plants*, *8*(8). <https://doi.org/10.3390/plants8080253>
- Fan, P., Leong, B. J., & Last, R. L. (2019). Tip of the trichome: evolution of acylsugar metabolic diversity in Solanaceae. In *Current Opinion in Plant Biology* (Vol. 49, pp. 8–16). Elsevier Ltd. <https://doi.org/10.1016/j.pbi.2019.03.005>
- Fernández-Muñoz, R., Salinas, M., Álvarez, M., & Cuartero, J. (2003). *Inheritance of Resistance to Two-spotted Spider Mite and Glandular Leaf Trichomes in Wild Tomato Lycopersicon pimpinellifolium (Jusl.) Mill.* *128*(2), 188–195.
- Firdaus, S. (2012). *Identification of whitefly resistance in tomato and hot pepper*. Wageningen University.
- Firdaus, Syarifin, van Heusden, A. W., Hidayati, N., Supena, E. D. J., Mumm, R., de Vos, R. C. H., Visser, R. G. F., & Vosman, B. (2013). Identification and QTL mapping of whitefly resistance components in *Solanum galapagense*. *Theoretical and Applied Genetics*, *126*(6), 1487–1501. <https://doi.org/10.1007/s00122-013-2067-z>
- Firdaus, Syarifin, van Heusden, A. W., Hidayati, N., Supena, E. D. J., Visser, R. G. F., & Vosman, B. (2012). Resistance to *Bemisia tabaci* in tomato wild relatives. *Euphytica*, *187*(1), 31–45. <https://doi.org/10.1007/s10681-012-0704-2>
- Freitas, J. A., Maluf, W. R., Das Graças Cardoso, M., Gomes, L. A. A., & Bearzotti, E. (2002). Inheritance of foliar zingiberene contents and their relationship to trichome densities and whitefly resistance in tomatoes. *Euphytica*, *127*(2), 275–287. <https://doi.org/10.1023/A:1020239512598>
- Gan, L., Xia, K., Chen, J.-G., & Wang, S. (2011). Functional characterization of TRICHOMELESS2, a new single-repeat R3 MYB transcription factor in the regulation of trichome patterning in *Arabidopsis*. *BMC Plant Biology*, *11*(1). <https://doi.org/10.1186/1471-2229-11-176>
- Gao, Y., He, C., Zhang, D., Liu, X., Xu, Z., Tian, Y., Liu, X.-H., Zang, S., Pauly, M., Zhou, Y., & Zhang, B. (2017). Two Trichome Birefringence-Like Proteins Mediate Xylan Acetylation, Which Is Essential for Leaf Blight Resistance in Rice. *Plant Physiology*, *173*(1). <https://doi.org/10.1104/pp.16.01618>
- Garrison, E., & Marth, G. (2012). *Haplotype-based variant detection from short-read sequencing*. arXiv [q-bio.GN].
- Glas, J. J., Schimmel, B. C. J., Alba, J. M., Escobar-Bravo, R., Schuurink, R. C., & Kant, M. R. (2012). Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences*, *13*(12), 17077–17103. <https://doi.org/10.3390/ijms131217077>
- Guo, Z., Weston, P. A., & Snyder, J. C. (1993). Repellency to two-spotted spider mite,

- Tetranychus urticae Koch, as related to leaf surface chemistry of Lycopersicon hirsutum accessions. *Journal of Chemical Ecology*, 19(12), 2965–2979. <https://doi.org/10.1007/BF00980596>
- Gurr, G. M., & McGrath, D. (2001a). Effect of plant variety, plant age and photoperiod on glandular pubescence and host-plant resistance to potato moth (*Phthorimaea operculella*) in *Lycopersicon* spp. *Annals of Applied Biology*, 138(2), 221–230. <https://doi.org/10.1111/j.1744-7348.2001.tb00106.x>
- Gurr, G. M., & McGrath, D. (2001b). Foliar pubescence and resistance to potato moth, *Phthorimaea operculella*, in *Lycopersicon hirsutum*. *Entomologia Experimentalis et Applicata*, 103(1), 35–41. <https://doi.org/10.1023/A:1019819722203>
- Huchelmann, A., Boutry, M., & Hachez, C. (2017). Plant glandular trichomes: Natural cell factories of high biotechnological interest. In *Plant Physiology* (Vol. 175, Issue 1, pp. 6–22). American Society of Plant Biologists. <https://doi.org/10.1104/pp.17.00727>
- Kirik, V., Lee, M. M., Wester, K., Herrmann, U., Zheng, Z., Oppenheimer, D., Schiefelbein, J., & Hulskamp, M. (2005). Functional diversification of MYB23 and GL1 genes in trichome morphogenesis and initiation. *Development*, 132(7), 1477–1485. <https://doi.org/10.1242/dev.01708>
- Leckie, B. M., de Jong, D. M., & Mutschler, M. A. (2012). Quantitative trait loci increasing acylsugars in tomato breeding lines and their impacts on silverleaf whiteflies. *Molecular Breeding*, 30(4), 1621–1634. <https://doi.org/10.1007/s11032-012-9746-3>
- Lemke, C. A., & Mutschler, M. (1984). Inheritance of glandular trichomes in crosses between *Lycopersicon esculentum* and *Lycopersicon pennellii*. *Journal of the American Society for Horticultural Science*, 109, 592–596.
- Leong, B. J., Lybrand, D. B., Lou, Y.-R., Fan, P., Schillmiller, A. L., & Last, R. L. (2019). Evolution of metabolic novelty: A trichome-expressed invertase creates specialized metabolic diversity in wild tomato. In *Sci. Adv* (Vol. 5). <http://advances.sciencemag.org/>
- Li, H. (2013). *Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM*. arXiv [q-bio.GN].
- Liedl, B. E., Lawson, D. M., White, K. K., Shapiro, J. A., Cohen, D. E., Carson, W. G., Trumble, J. T., & Mutschler, M. A. (1995). Acylsugars of wild tomato *Lycopersicon pennellii* alters settling and reduces oviposition of *Bemisia argentifolii* (Homoptera: Aleyrodidae). *Journal of Economic Entomology*, 88(3), 742–748. <https://doi.org/10.1093/jee/88.3.742>
- Luckwill, L. (1943). The genus *Lycopersicon*: historical, biological, and taxonomic survey of the wild and cultivated tomatoes. *Aberdeen University Press*.
- Maliepaard, C., Bas, N., Van Heusden, S., Kos, J., Pet, G., Verkerk, R., Vrielink, R., Zabel, P., & Lindhout, P. (1995). Mapping of QTLs for glandular trichome densities and *trialeurodes vaporariorum* (Greenhouse whitefly) resistance in an F2 from *lycopersicon esculentum* × *lycopersicon hirsutum* f. *glabratum*. *Heredity*, 75(4), 425–433. <https://doi.org/10.1038/hdy.1995.155>

- Maluf, W. R., Campos, G. A., & Das Graças Cardoso, M. (2001). Relationships between trichome types and spider mite (*Tetranychus evansi*) repellence in tomatoes with respect to foliar zingiberene contents. In *Euphytica* (Vol. 121).
- Mandal, S., Ji, W., & McKnight, T. D. (2020). Candidate Gene Networks for Acylsugar Metabolism and Plant Defense in Wild Tomato *Solanum pennellii*. *The Plant Cell*, 32(1), 81–99. <https://doi.org/10.1105/tpc.19.00552>
- Mansfeld, B. N., & Grumet, R. (2018). QTLseqr: An R Package for Bulk Segregant Analysis with Next-Generation Sequencing. *The Plant Genome*, 11. <https://doi.org/10.3835/plantgenome2018.01.0006>
- 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., & Díez, M. J. (2020). Exploiting the diversity of tomato: the development of a phenotypically and genetically detailed germplasm collection. *Horticulture Research*, 7(1). <https://doi.org/10.1038/s41438-020-0291-7>
- Maynard, D. N., & Hochmuth, G. J. (2006). *Knott's Handbook for Vegetable Growers*. John Wiley & Sons.
- McDowell, E. T., Kapteyn, J., Schmidt, A., Li, C., Kang, J. H., Descour, A., Shi, F., Larson, M., Schillmiller, A., An, L., Daniel Jones, A., Pichersky, E., Soderlund, C. A., & Gang, D. R. (2011). Comparative Functional Genomic Analysis of *Solanum* Glandular Trichome Types. *Plant Physiology*, 155(1), 524–539. <https://doi.org/10.1104/pp.110.167114>
- Miao, C., Fang, J., Li, D., Liang, P., Zhang, X., Yang, J., Schnable, J. C., & Tang, H. (2018). Genotype-Corrector: improved genotype calls for genetic mapping in F2 and RIL populations. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-018-28294-0>
- Momotaz, A., Scott, J. W., & Schuster, D. J. (2010). Identification of Quantitative Trait Loci Conferring Resistance to *Bemisia tabaci* in an F2 Population of *Solanum lycopersicum* × *Solanum habrochaites* Accession LA1777. *Journal of the American Society for Horticultural Science*, 135(2). <https://doi.org/10.21273/JASHS.135.2.134>
- Mutschler, M. A., Doerge, R. W., Liu, S. C., Kuai, J. P., Liedl, B. E., & Shapiro, J. A. (1996). QTL analysis of pest resistance in the wild tomato *Lycopersicon pennellii*: QTLs controlling acylsugar level and composition. *Theoretical and Applied Genetics*, 92(6), 709–718. <https://doi.org/10.1007/BF00226093>
- 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., & Wellmer, F. (2013). Control of reproductive floral organ identity specification in arabidopsis by the C function regulator AGAMOUS. *Plant Cell*, 25(7), 2482–2503. <https://doi.org/10.1105/tpc.113.113209>
- Oriani, M. A., Vendramim, J. D., Oriani, M. A. de G., & Vendramim, J. D. (2010). Influence of trichomes on attractiveness and ovipositional preference of *Bemisia tabaci* (Genn.) B biotype (Hemiptera: Aleyrodidae) on tomato genotypes. *Neotropical Entomology*, 39(6), 1002–1007. <https://doi.org/10.1590/S0103-90162011000100006>

- Peiffer, M., Tooker, J. F., Luthe, D. S., & Felton, G. W. (2009). Plants on early alert: Glandular trichomes as sensors for insect herbivores. *New Phytologist*, *184*(3), 644–656. <https://doi.org/10.1111/j.1469-8137.2009.03002.x>
- Peter, A. J., Shanower, T. G., & Romeis, J. (1995). The role of plant trichomes in insect resistance: a review. *Phytophaga*, *7*, 41–63.
- Potikha, T., & Delmer, D. P. (1995). A mutant of *Arabidopsis thaliana* displaying altered patterns of cellulose deposition. *The Plant Journal*, *7*(3). <https://doi.org/10.1046/j.1365-313X.1995.7030453.x>
- R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Rakha, M., Hanson, P., & Ramasamy, S. (2017). 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*, *64*(2), 247–260. <https://doi.org/10.1007/s10722-015-0347-y>
- Rodríguez-López, M. J., Moriones, E., & Fernández-Muñoz, R. (2020). An Acylsucrose-Producing Tomato Line Derived from the Wild Species *Solanum pimpinellifolium* Decreases Fitness of the Whitefly *Trialeurodes vaporariorum*. *Insects*, *11*(9). <https://doi.org/10.3390/insects11090616>
- Saeidi, Z., Mallik, B., & Kulkarni, R. S. (2007). Inheritance of glandular trichomes and two-spotted spider mite resistance in cross *Lycopersicon esculentum* “Nandi” and *L. pennellii* “LA2963.” *Euphytica*, *154*(1–2), 231–238. <https://doi.org/10.1007/s10681-006-9291-4>
- Salinas, M., Capel, C., Alba, J. M., Mora, B., Cuartero, J., Fernández-Muñoz, R., Lozano, R., & Capel, J. (2013). Genetic mapping of two QTL from the wild tomato *Solanum pimpinellifolium* L. controlling resistance against two-spotted spider mite (*Tetranychus urticae* Koch). *Theoretical and Applied Genetics*, *126*(1), 83–92. <https://doi.org/10.1007/s00122-012-1961-0>
- Shearer, L. A., Anderson, L. K., de Jong, H., Smit, S., Goicoechea, J. L., Roe, B. A., Hua, A., Giovannoni, J. J., & Stack, S. M. (2014). Fluorescence In Situ Hybridization and Optical Mapping to Correct Scaffold Arrangement in the Tomato Genome. *G3*, *4*(8). <https://doi.org/10.1534/g3.114.011197>
- Smeda, J. R., Schillmiller, A. L., Anderson, T., Ben-Mahmoud, S., Ullman, D. E., Chappell, T. M., Kessler, A., & Mutschler, M. A. (2018). Combination of Acylglucose QTL reveals additive and epistatic genetic interactions and impacts insect oviposition and virus infection. *Molecular Breeding*, *38*(1). <https://doi.org/10.1007/s11032-017-0756-z>
- Smeda, J. R., Schillmiller, A. L., Last, R. L., & Mutschler, M. A. (2016). Introgression of acylsugar chemistry QTL modifies the composition and structure of acylsugars produced by high-accumulating tomato lines. *Molecular Breeding*, *36*(12). <https://doi.org/10.1007/s11032-016-0584-6>
- Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., Uemura, A., Utsushi, H., Tamiru, M., Takuno, S., Innan, H., Cano, L. M., Kamoun, S., & Terauchi,

- R. (2013). QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant Journal*, *74*(1), 174–183. <https://doi.org/10.1111/tpj.12105>
- Taylor, J. D., & Butler, D. (2014). *ASMap: Linkage map construction using the MSTmap algorithm* (R package version 0.3--3).
- Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth). Fourth.; Springer: New York.
- 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. (2017). Loss of type-IV glandular trichomes is a heterochronic trait in tomato and can be reverted by promoting juvenility. *Plant Science*, *259*, 35–47. <https://doi.org/10.1016/j.plantsci.2017.03.006>
- 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. (2019). QTL mapping of insect resistance components of *Solanum galapagense*. *Theoretical and Applied Genetics*, *132*(2), 531–541. <https://doi.org/10.1007/s00122-018-3239-7>
- Weston, P. A., Johnson, D. A., Burton, H. T., & Snyder, J. C. (1989). Trichome secretion composition, trichome densities, and spider mite resistance of ten accessions of *Lycopersicon hirsutum*. *Journal of the American Society for Horticultural Science*.
- Wilkins, R. T., Shea, G. O., Halbreich, S., & Stamp, N. E. (1996). Resource availability and the trichome defenses of tomato plants. *Oecologia*, *106*(2), 181–191. <https://doi.org/10.1007/BF00328597>
- 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., Jauker, F., McCracken, M. E., Zou, Y., Potts, S. G., Rundlöf, M., Noriega, J. A., Greenop, A., Smith, H. G., Bommarco, R., ... Pywell, R. F. (2019). Meta-analysis reveals that pollinator functional diversity and abundance enhance crop pollination and yield. *Nature Communications*, *10*(1). <https://doi.org/10.1038/s41467-019-09393-6>
- Yekutieli, D., & Benjamini, Y. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, *29*(4). <https://doi.org/10.1214/aos/1013699998>
- Ziarsolo, P., Hasing, T., Tech, V., Hilario, R., & Canizares, J. (2020). K-Seq, An Affordable, Reliable, and Open Klenow NGS-Based Genotyping Technology. *Cold Spring Harbor Laboratory*, 1–15. <https://doi.org/10.21203/rs.3.rs-127367/v1>

DISCUSIÓN GENERAL

DISCUSIÓN GENERAL

El continuo crecimiento de la población mundial (Naciones Unidas, 2017), inevitablemente conduce hacia un escenario donde será necesaria una producción de alimentos cada vez superior que permita satisfacer sus necesidades alimenticias. Para poder afrontar esta demanda de alimentos es necesario, entre otras cosas, el desarrollo de nuevas variedades más productivas, capaces de adaptarse a las condiciones ambientales que irán surgiendo como consecuencia del cambio climático y resistentes a plagas y enfermedades. Sin embargo, uno de los principales problemas que existe para la mejora de las especies cultivadas es la reducida diversidad genética de estas especies como consecuencia de su proceso de domesticación (Meyer and Purugganan, 2013). Esto da lugar a que en muchos casos la mejora de las especies cultivadas sea altamente dependiente del uso de especies silvestres próximas como fuentes de variabilidad genética. En este sentido, es crucial explorar y caracterizar la diversidad existente en los centros de origen y domesticación de las especies de cultivo, zonas que, por lo general, presentan los mayores índices de diversidad genética. Gracias a los esfuerzos de recolección y conservación, gran parte de esta diversidad se encuentra actualmente recogida en múltiples colecciones de germoplasma. Sin embargo, la falta de recursos hace que todavía se carezca de una buena caracterización tanto de la diversidad genética como morfológica presente en estas colecciones (Ranjan et al., 2012).

La mejora del tomate cultivado (*Solanum lycopersicum* var. *lycopersicum*, SLL) ejemplifica perfectamente esta problemática. Se trata de una especie de gran interés desde un punto de vista alimentario y económico, por ser una de las hortalizas con mayor consumo y producción a nivel mundial (FAO, 2017), al tiempo que presenta una muy baja diversidad genética debido a su proceso de domesticación. Esta baja diversidad, además, contrasta con la alta diversidad genética y morfológica descrita en las especies próximas *S. pimpinellifolium* (SP) y *S. lycopersicum* var. *cerasiforme* (SLC), especialmente en las regiones de Perú y Ecuador (Blanca et al., 2012; Blanca et al., 2015; Razifard, 2020). Esta alta diversidad, junto con su capacidad para hibridar con el tomate cultivado, hacen de estas especies fuentes potenciales de caracteres de interés en mejora. Sin embargo, dichas especies se encuentran infrutilizadas en parte porque sigo siendo necesario llevar a cabo una caracterización más extensa de los recursos disponibles.

Así pues, el objetivo principal de esta tesis ha sido llevar a cabo una detallada caracterización morfológica y genética de la diversidad presente en una colección de germoplasma integrada por SP, SLC y SLL representativa de los centros de origen y domesticación del tomate cultivado. Una de las principales ventajas de esta colección es que ha incluido un gran número de entradas provenientes de un amplio rango de orígenes geográficos, ambientes y climas previamente no estudiados. Este hecho es de especial relevancia en el caso de SLC, ya que al ser una especie que se encuentra tanto de forma silvestre como asociada a la actividad humana y en climas muy diversos, un muestreo amplio favorece la captura de la diversidad de esta especie. Además, en la colección se ha incrementado el número de entradas de SP de Ecuador, un grupo diferenciado de SP de Perú y que generalmente ha estado infrarrepresentado en los estudios previos (Rick et al., 1977; Rick et al., 1978; Caicedo y Schaal, 2004; Tieman et al., 2017).

Diversidad morfológica

En la caracterización morfológica se han evaluado 63 descriptores morfológicos (Rick & Fobes, 1975; Rick, Fobes & Holle, 1977; Cuartero et al., 1985; Sifres et al., 2006; Blanca.

et al. 2012), entre los cuales se encontraban algunos descriptores nuevos que no habían sido evaluados previamente en SP o SLC como el ángulo de caída de las hojas. Además, también se incluyeron nuevos niveles de caracterización para descriptores previamente evaluados en estas especies debido a que se había observado una mayor variabilidad en nuestra colección, con respecto a la variabilidad descrita previamente, como por ejemplo colores amarillos claro en los pétalos. Por ejemplo, cabe destacar el caso de SP donde se ha podido detectar variabilidad incluso para caracteres que se habían considerado invariantes en esta especie (Rick et al., 1976; Rick et al., 1978) como el tipo de foliolo de las hojas o el número de flores por inflorescencia. Adicionalmente, este análisis ha revelado una alta diversidad en toda la colección principalmente en caracteres relacionados con la morfología de las flores, longitud y tipo de inflorescencia, forma, tamaño y color del fruto. Además, SLC se ha presentado como la especie con mayor grado de diversidad morfológica.

Los cambios en la morfología floral y la ejerción estigmática, junto con el incremento del tamaño y la variabilidad en la forma del fruto, son los cambios más característicos asociados al proceso de domesticación y selección. Durante este proceso, se pasó de flores sencillas con 5 pétalos y sépalos que son características de SP, a otras más complejas. En SLC, las flores más simples de Mesoamérica contrastan con la variabilidad observada en el resto de los grupos geográficos. Esta variabilidad en la región andina ya había sido descrita anteriormente por Rick and Fobes (1975). El proceso de domesticación también conllevó cambios en la posición del estigma que pasó de altamente exerto, más propio de SP Perú, a inserto o al mismo nivel de los estambres, en SLL. Dentro de la especie SP, este carácter también presenta variabilidad entre los dos grupos geográficos ya que es superior en Perú con respecto a Ecuador (Rick, 1976; Rick et al., 1978; Widrlechner, 1987). La posición del estigma es un carácter relacionado con el sistema reproductivo de la especie, siendo más exerto en poblaciones alógamas y más inserto en autógamias. En el proceso de difusión y dispersión desde el centro de origen, se habría seleccionado la autogamia y por tanto una reducción de la ejerción estigmática (Stebbins, 1957). En la especie SLC se ha observado una inserción progresiva del estigma desde SLC Perú hasta SLC México, como ya había sido previamente descrito (Rick, 1958; Rick y Holle, 1990). Cabe destacar el hecho de que las entradas de Ecuador, procedentes de varias provincias situadas principalmente en la ladera este de los Andes, presentan características intermedias entre las especies SP y SCL que se reflejan en las flores. Por ejemplo, algunas de ellas se han caracterizado por un estigma exerto y pétalos curvados hacia atrás, principales características asociadas a las flores de SP. La existencia de estas formas intermedias puede deberse a cruzamientos esporádicos con SP, hecho que ya ha sido demostrado por Rick et al. (1958) y Blanca et al. (2012). Por el contrario, las entradas procedentes de Mesoamérica se caracterizan principalmente por sus flores de menores dimensiones, similares a SP en tamaño, pero con el estigma inserto o al mismo nivel que los estambres, más propio de SLL. Estos resultados son consistentes con la variación ya observada para estas especies y grupos geográficos (Blanca et al., 2012; Rick and Fobes, 1975) y están relacionados con la migración desde el centro de origen, que resultó en la fijación de alelos dentro de las poblaciones.

Estos cambios en morfología floral, unidos a un aumento del tamaño del ovario, implicaron una transición desde los frutos pequeños, redondeados y uniformes de SP a la variabilidad presente en frutos del tomate cultivado. Tanto la forma como el tamaño del fruto fueron caracteres especialmente variables dentro de la especie SLC, lo que la convierte en una potencial fuente de alelos de interés en programas de mejora. Las entradas con frutos de tamaño muy pequeño y comparables a los de SP son propias de entradas que se encuentran en regiones donde crece de forma silvestre. Por otro lado, las entradas procedentes de mercados

locales tuvieron un mayor tamaño de sus frutos y cierto acostillado. En cuanto a las entradas de Ecuador, encontramos frutos pequeños tipo SP junto con frutos grandes y acostillados, más propias de SLL. El cambio en la forma y el tamaño del fruto desde la especie silvestre a la especie cultivada habría sucedido en dos fases: primero fue aumentando al pasar de SP a SLC y, posteriormente, se diversificó todavía más al pasar de SLC a SLL (Lin et al., 2014, Blanca et al., 2012; Blanca et al., 2015). El tamaño del fruto está asociado a múltiples QTLs como *fw1.1*, *fw5.2*, *fw7.2*, *fw12.1* y *lcn12.1*, *fw2.1*, *fw2.2*, *fw2.3*, *fw3.2*, *lcn3.1*, *fw9.3*, *lcn10.1*, *fw11.1*, *fw11.2* y *fw11.3* (Grandillo et al., 1999; Frary et al., 2000; Tanksley et al., 2004; Huang and Van der Knaap, 2011; Muños et al., 2011; Chakrabarti et al., 2013; van der Knaap et al., 2014) mientras que la forma está asociada a los genes descritos anteriormente *fas*, *sun*, *ovate* y *lc* (Rodríguez et al., 2011). De estos genes, *sun* y *ovate* regulan el alargamiento del fruto, mientras que *fas* y *lc* regulan la fasciación y el número de lóculos, respectivamente (Monforte et al., 2014). La aparición de mutaciones en estos genes dio lugar a nuevos tipos de frutos que fueron seleccionados por el hombre dando lugar a la extensa variabilidad existente en SLC y SLL. La variabilidad observada en nuestra colección podría ser debida a la presencia de alguno de estos alelos, ya que en algunas de las entradas incluidas en este trabajo de la especie SLC y SLL, ya se ha descrito la presencia de alelos alternativos para los genes *fas* y *ovate* (Blanca et al., 2015).

Aparte de los cambios característicos asociados al proceso de domesticación, también se han detectado caracteres que pueden ser de interés en distintos programas de mejora del tomate cultivado. En nuestra colección de SP se han llegado a detectar inflorescencias de más de 100 flores que son interesantes desde un punto de mejora, ya que tanto la longitud de la inflorescencia como el número de flores por inflorescencia tienen especial interés en el desarrollo de variedades tipo *cherry*. Otro carácter de interés es la enorme variabilidad en el color del fruto que se ha detectado. Tanto en la especie SP como SLC se han encontrado variedades de color amarillo pálido hasta anaranjado, además de distintas tonalidades de rojo. El color rojo intenso está asociado a un mayor contenido en licopeno (Bergougnoux, 2014; Mazidi et al., 2019), mientras que los frutos de color amarillo o anaranjado son debidos a variaciones en el contenido en carotenoides (Yoo et al., 2018). Por tanto, este carácter podría ser de interés en la mejora de la calidad nutricional.

Diversidad genética

Las entradas que se han empleado en esta tesis habían sido secuenciadas previamente, dentro del proyecto VARITOME, y sus SNPs están disponibles en *Solanaceae Genomics Network*. El programa SnpEff ha permitido anotar todos estos SNPs, determinar su posible impacto y efecto. Esto ha contribuido a poner en valor el potencial genético de la colección, ya que se han identificado una gran cantidad de variantes de alto impacto, algunas de las cuales se encuentran localizadas en las regiones genómicas candidatas del análisis GWAS. Adicionalmente, de entre todos los efectos detectados, se han encontrado variantes que pueden producir la generación o pérdida de codones de parada, afectando gravemente a la función del gen. Estos resultados son de interés en la búsqueda de genes candidatos y variantes alélicas naturales, que son de notable interés en el desarrollo de los programas de mejora.

En cuanto a la diversidad genética, los resultados indican que el mayor nivel de diversidad se encuentra en la especie silvestre SP y el menor, en la especie cultivada. SLC, por su parte, mostró distintos grados de diversidad en función del origen geográfico, siendo semejante a SP en Perú y Ecuador, mientras que existe una reducción desde Mesoamérica

hasta México, donde los niveles son equiparables a SLL. El resultado de nuestro estudio es consistente con las descripciones previas de una mayor diversidad genética en SP, que se va reduciendo progresivamente conforme migra su centro de origen en el norte de Perú hacia el sur de Perú y Ecuador (Rick et al., 1976; Zuriaga et al., 2009; Caicedo and Schaal, 2004; Blanca et al., 2012; Blanca et al., 2015; Razifard et al., 2020). En SLC la mayor diversidad en la región andina y su progresiva reducción hacia Mesoamérica y México también se había observado anteriormente (Williams and Clair, 2008, Rick and Holle, 1990). Esta pérdida de diversidad es consecuencia de la migración desde su lugar de origen, acompañada de la fijación de genotipos exitosos en las nuevas poblaciones (Rick and Fobes, 1975; Rick and Holle, 1990; Blanca, 2012; Blanca et al., 2015; Razifard, 2020). Además, concuerda con la hipótesis de que Mesoamérica sea el centro de domesticación y difusión del tomate cultivado (Blanca et al., 2012; Williams and Clair, 2008). Por último, y como consecuencia del proceso de domesticación, se habría producido una gran pérdida de diversidad genética en SLL (Blanca et al., 2012; Razifard et al., 2020).

Asociaciones fenotipo-genotipo

Para demostrar el potencial genético de la colección para la mejora, se llevó a cabo un análisis de asociación genómica (GWAS) con el objetivo de identificar regiones asociadas a los caracteres fenotípicos evaluados. Las asociaciones más relevantes detectadas son aquellas relacionadas con el tamaño, color y forma del fruto o la longitud y el tipo de inflorescencia. El análisis ha permitido la identificación de genes ya descritos, confirmando la utilidad de la colección, así como posibles nuevas regiones genómicas implicadas en el control de caracteres de interés. El GWAS ha revelado asociaciones entre el número de lóculos y las regiones donde se localizan los genes *lc* y *fas* y entre el peso del fruto y los QTLs *fw2.2* (Frary et al., 2000) y *fw9.2* (Grandillo et al., 1999). El QTL *fw2.2* ya había sido detectado en otros estudios de GWAS (Bauchet et al., 2017) y, además, se habían detectado SNPs muy cercanos a las posiciones asociadas en nuestro análisis (Sacco et al., 2018). Además, también se detectó una región en el cromosoma 11 que estaba cercana al QTL anotado como *fw11.3*, controlado por un gen que actúa en el tamaño de las células del pericarpio (Mu et al., 2017). Adicionalmente a estos genes conocidos, también se han descrito otras dos nuevas regiones en los cromosomas 1 y 4 que podrían estar asociadas a la presencia de frutos fasciados.

El potencial de la colección en mejora de la calidad nutricional queda de manifiesto en las asociaciones detectadas con el color del fruto: en el cromosoma 1, en una región portadora de un gen anotado como una enzima que participa en la síntesis de carotenoides (*carotenoid cleavage dioxygenase 1B*), en el cromosoma 3, a 8Mb de un gen involucrado en la biosíntesis de carotenoides en fruto amarillo, y en el cromosoma 10, cercana a una región con un gen anotado por participar en el color de frutos naranjas. Además, el color rosa del fruto que fue medido cualitativamente reveló una asociación en una región donde se encuentra anotado el gen *y*, relacionado con la ausencia de color en la epidermis del fruto. En cuanto a los caracteres relacionados con inflorescencias, cabe destacar la asociación entre la inflorescencia de tipo unípara y el cromosoma 11, en una región a 5 Mb del gen responsable de la inflorescencia ramificada de los mutantes *fin* (Xu et al., 2015) o la asociación entre inflorescencia bifurcada y una región del cromosoma 9, con un gen asociado a la regulación del desarrollo vegetativo y reproductivo en respuesta a estrés. Por último, cabe destacar la asociación entre el número de flores por inflorescencia y una región genómica del cromosoma 11 que podría ser una nueva posible región candidata, distinta a QTLs que ya han sido descritos en los cromosomas 2, 3 y 5 (Georgiady et al., 2002).

Desarrollo de poblaciones segregantes

Se ha desarrollado una colección de familias segregantes F2 con el fin de que puedan estar disponibles para la comunidad científica y poder explotar la variabilidad morfológica y genética presente en la colección, especialmente dentro de la especie SLC. Estas familias permiten acelerar la identificación de genes de interés y su transferencia a variedades comerciales dentro de los programas de mejora. En total se han producido 457 poblaciones F2 cruzando todas las entradas de la colección con una entrada representativa de cada especie, que se corresponden con BGV007109 (SP), LA2278 (SLC) y Money Maker (SLL). De todas estas familias, dos con fondos genéticos diferentes se han empleado en la identificación de los factores genéticos que controlan la presencia y densidad de tricomas glandulares tipo IV. Este tipo de tricomas se detectaron en altas densidades en una entrada de SP de la colección (BGV016047), durante la caracterización morfológica de la colección. La presencia de tricomas glandulares de tipo IV ha sido descrita en altas densidades en especies silvestres lejanas al tomate cultivado como *S. habrochaites*, *S. pennellii*, *S. galapaguense*, *S. cheesmaniae* and *S. neorikii* (Guo et al., 1993; Blauth et al., 1998; Firdaus et al., 2013; Andrade et al., 2017). Sin embargo, el uso de estas especies en los programas de mejora está altamente limitado debido a la incorporación de caracteres que no son deseables o a problemas de incompatibilidad entre estas especies y la especie cultivada. La entrada BGV016047 es especialmente interesante ya que se trata de una entrada de SP, la especie silvestre más próxima al tomate cultivado, lo que permite evitar los problemas anteriormente descritos asociados al uso de especies más alejadas filogenéticamente. Por otro lado, las densidades descritas en esta entrada (superiores a 100 tricomas/mm²) son similares a los niveles descritos en las especies filogenéticamente más alejadas y se encuentran muy por encima de los niveles descritos para SP en estudios previos (Firdaus et al., 2012; Fernández-Muñoz, 2013; Alba et al., 2009; Rahka et al., 2016; Rahka et al., 2017).

La caracterización detallada de la densidad de tricomas en la entrada BGV016047 mostró que tanto la edad de la hoja como la edad de la planta afectan a la presencia de tricomas glandulares. Este efecto ha sido observado en otras accesiones (Gurr y Mc Grath, 2001). También se ha detectado una amplia variabilidad en la densidad entre las diferentes campañas que puede ser debida a las diferencias en las condiciones ambientales entre las distintas épocas del año. Las dos familias se han genotipado con el nuevo método de genotipado K-seq (Ziarsolo et al., 2020), que ha demostrado ser una técnica eficiente y de bajo coste para la elaboración de un mapa genético ultradenso y el mapeo de QTLs. El análisis ha mostrado que la densidad de los tricomas tipo IV está controlado principalmente por un QTL mayor en el inicio del cromosoma 9 y un QTL menor en el cromosoma 11, así como otros posibles QTLs menores en los cromosomas 6 y 5. Estos QTLs no habían sido previamente descritos en *S. pimpinellifolium*. La región detectada en el cromosoma 9 coincide con la detectada en *S. galapaguense* (Vosman et al., 2019) y en *S. habrochaites* (Maliepard et al., 1995; Momotaz, 2010), asociadas con una reducción en la ovoposición de *Bemisia tabaci*, aunque con un menor efecto que el encontrado en nuestro estudio. En el cromosoma 11 se han descrito también QTLs relacionados con la presencia de tricomas glandulares y resistencia a *Bemisia tabaci* en *S. habrochaites* (Momotaz, 2010) o niveles de acil azúcares (Leckie et al., 2012) que no solapan con los QTLs descritos para nuestras familias segregantes. Finalmente, se han descrito QTLs en el cromosoma 5 relacionados con resistencia a mosca blanca y *Bemisia tabaci* en *S. galapaguense* (Vosman et al., 2019), presencia de tricomas glandulares en *S. habrochaites* (Maliepard et al., 1995) y en *S. pennellii* (Blauth et al., 1998) y niveles de acilazúcares en *S. pennellii* (Blauth et al., 1998), tanto en el cromosoma 5 como en el 6 (Leckie et al., 2012). En las regiones genómicas descritas se encuentran anotados genes de

interés como dos “trichome birefringence-like proteins” que se ha visto que están implicadas en la regulación de la densidad de tricomas (Potikha et al., 1995; Gao et al., 2017), varios factores de transcripción MYB, que participa en la regulación del desarrollo de tricomas (Gan et al., 2011; Kirik et al., 2005) y genes relacionados con la producción de acilazúcares (Mandal et al., 2020).

En su conjunto, esta tesis muestra el potencial que tiene el disponer de una colección de germoplasma extensamente caracterizada tanto morfológica como genéticamente para la mejora vegetal. Como se ha mostrado, el conjunto de entradas recoge una gran variabilidad de las dos especies más próximas al tomate cultivado. Se ha encontrado una gran diversidad morfológica para un gran número de los caracteres analizados y, dados los criterios usados para establecer la colección, es muy probable que exista variabilidad incluso para caracteres de importancia en mejora que no se han considerado en el ámbito de desarrollo de la tesis como caracteres relacionados con la calidad del fruto, resistencia a plagas y enfermedades y adaptación a estreses abióticos. Además de una buena caracterización morfológica, el disponer de un buen genotipado y una buena anotación de las variantes permite una detección sencilla de regiones y genes candidatos, orientando la búsqueda de alelos. Para que se pueda producir una rápida transferencia de los resultados de investigación, especialmente de aquella investigación financiada con fondos públicos, a otros grupos de investigación o mejoradores es necesario que dichos resultados se hagan disponibles a la comunidad. Es por ello que tanto la colección de entradas, como los datos de caracterización morfológica, así como los de genotipado y anotación se han puesto a disposición de toda la comunidad. Además, con el objetivo de hacer más accesibles los resultados y acelerar los procesos de mejora se ha desarrollado una colección de familias F2 que también se encuentra accesible para toda la comunidad científica.

CONCLUSIONES GENERALES

CONCLUSIONES GENERALES

- El análisis de la variabilidad existente en el centro de origen y domesticación del tomate cultivado a través de la caracterización morfológica y genética de una colección de germoplasma compuesta por las especies *Solanum pimpinellifolium*, *Solanum lycopersicum* var. *cerasiforme* y *Solanum lycopersicum* var. *lycopersicum* ha puesto de manifiesto el enorme potencial de la colección de cara a programas de mejora del tomate cultivado.
- La caracterización morfológica ha mostrado la existencia de diferencias entre los tres taxones, debidas principalmente al proceso de domesticación, para los caracteres de planta, inflorescencia, flor y fruto. También se han encontrado caracteres discriminantes de tipo vegetativo y reproductivo entre los distintos grupos geográficos de la especie *S. pimpinellifolium*. La especie *S. lycopersicum* var. *cerasiforme* es la que mayor variabilidad morfológica ha presentado, fruto de los procesos de selección ejercida por el hombre junto con la adaptación a nuevos ambientes. El rango de variación de este taxón abarca desde *S. pimpinellifolium* hasta *S. lycopersicum*.
- El análisis genético de la colección indica una reducción de la diversidad genética de la especie cultivada en comparación con *S. pimpinellifolium* y *S. lycopersicum* var. *cerasiforme*. Dentro de estas especies, las regiones de Perú y Ecuador se presentan como las regiones geográficas donde se pueden encontrar los mayores niveles de diversidad genética.
- Los estudios de asociación genómica o GWAS han descrito la presencia de asociaciones entre diversas regiones genómicas y ocho caracteres cuantitativos y siete caracteres cualitativos. En parte de estas regiones, existen o bien genes involucrados en el control del carácter con el que presentaron asociación o bien posibles genes candidatos.
- El estudio de la densidad de tricomas glandulares tipo IV en la entrada BGV016047 ha permitido determinar que la densidad de tricomas aumenta con la edad de la planta, mientras que la edad de la hoja tienen un efecto negativo en la densidad. Adicionalmente, se ha llevado a cabo un estudio del control genético de este carácter en dos fondos genéticos diferentes, SLL x SP y SLC x SP. En ambos fondos se han descrito dos QTLs localizados en los cromosomas 9 y 11 previamente desconocidos en SP. Las regiones asociadas presentan genes que intervienen en el desarrollo de los tricomas o en la formación de acil azúcares.
- La entrada BGV016047 de la especie *Solanum pimpinellifolium* presenta una densidad de tricomas glandulares tipo IV muy superior a otras entradas de esta especie. Algunos de los QTLs asociados al carácter ya han sido descritos por conferir resistencia a diferentes plagas o controlar el desarrollo de tricomas tipo IV en otras especies silvestres, más alejadas filogenéticamente del tomate cultivado. Por lo tanto, esta entrada se presenta como una potencial fuente de genes de interés en programas de mejora dirigidos a resistencia a plagas, ya que es la especie silvestre más cercana al tomate cultivado y se pueden cruzar sin que aparezcan problemas de incompatibilidad.

REFERENCIAS

REFERENCIAS

- Adato, A., Mandel, T., Mintz-Oron, S., Venger, I., Levy, D., Yativ, M., Domínguez, E., Wang, Z., De Vos, R. C. H., Jetter, R., Schreiber, L., Heredia, A., Rogachev, I., & Aharoni, A. (2009). Fruit-surface flavonoid accumulation in tomato is controlled by a SLMYB12-regulated transcriptional network. *PLoS Genetics*, 5(12). <https://doi.org/10.1371/journal.pgen.1000777>
- Al-Sane, K. O., Povero, G., & Perata, P. (2011). Anthocyanin tomato mutants: Overview and characterization of an anthocyanin-less somaclonal mutant. *Plant Biosystems*, 145(2), 436–444. <https://doi.org/10.1080/11263504.2010.548914>
- Alexander, L., & Hoover, M. (1955). Disease resistance in wild species of tomato: Report of the national screening Committee. *Agricultural Experimental Station Research Bulletin*, April, 7.
- Álvarez-Hernández, J. C., Cortez-Madrugal, H., & García-Ruiz, I. (2009). EXPLORACIÓN Y CARACTERIZACIÓN DE POBLACIONES SILVESTRES DE JITOMATE (SOLANACEAE) EN TRES REGIONES DE MICHOACÁN, MÉXICO. *Polibotánica*, 28, 139–159.
- Andrade, M. C., da Silva, A. A., Carvalho, R. de C., de Andrade Santiago, J., Souza de Oliveira, A. M., Merrill Francis, D., & Maluf, W. R. (2018). *Quantitative trait loci associated with trichomes in the Solanum galapagense accession LA1401*. 1671–1685. <https://doi.org/10.1007/s10722-018-0644-3>
- Andrade, M. C., da Silva, A. A., Neiva, I. P., Oliveira, I. R. C., De Castro, E. M., Francis, D. M., & Maluf, W. R. (2017). Inheritance of type IV glandular trichome density and its association with whitefly resistance from *Solanum galapagense* accession LA1401. *Euphytica*, 213(2). <https://doi.org/10.1007/s10681-016-1792-1>
- Arellano Rodríguez, L. J., Rodríguez Guzmán, E., Ron Parra, J., Martínez Ramírez, J. L., Lozoya Saldaña, H., Sánchez Martínez, J., & Lépiz Ildefonso, R. (2013). Evaluation of the resistance against *Phytophthora infestans* of wild populations of *Solanum lycopersicum* var. *cerasiforme*. *Revista Mexicana de Ciencias Agrícolas*, 4(5), 753–766. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-09342012000100010
- Bai, Y., & Lindhout, P. (2007). *Domestication and Breeding of Tomatoes: What have We Gained and What Can We Gain in the Future?* 1085–1094. <https://doi.org/10.1093/aob/mcm150>
- Baldwin, E. A., Scott, J. W., Shewmaker, C. K., & Schuch, W. (2000). Flavor Trivia and Tomato Aroma: Biochemistry and Possible Mechanisms for Control of Important Aroma Components. In *1013 HORTSCIENCE* (Vol. 35, Issue 6).
- Ballester, A. R., Molthoff, J., de Vos, R., Hekkert, B. te L., Orzaez, D., Fernández-Moreno, J. P., Tripodi, P., Grandillo, S., Martín, C., Heldens, J., Ykema, M., Granell, A., & Bovy, A. (2010). Biochemical and molecular analysis of pink tomatoes: Deregulated expression of the gene encoding transcription factor SLMYB12 leads to pink tomato fruit color. *Plant Physiology*, 152(1), 71–84. <https://doi.org/10.1104/pp.109.147322>

- Banerjee, M. K., & Kalloo, M. K. (1987). Sources and inheritance of resistance to leaf curl virus in *Lycopersicon*. *Theoretical and Applied Genetics*, 73(5), 707–710. <https://doi.org/10.1007/BF00260780>
- Baranski, R., Goldman, I., Nothnagel, T., & Scott, J. W. (2016). Improving Color Sources by Plant Breeding and Cultivation. In *Handbook on Natural Pigments in Food and Beverages: Industrial Applications for Improving Food Color* (pp. 429–472). Elsevier Inc. <https://doi.org/10.1016/B978-0-08-100371-8.00019-1>
- Barrantes, W., Fernández-del-Carmen, A., López-Casado, G., González-Sánchez, M. Á., Fernández-Muñoz, R., Granell, A., & Monforte, A. J. (2014). Highly efficient genomics-assisted development of a library of introgression lines of *Solanum pimpinellifolium*. *Molecular Breeding*, 34(4), 1817–1831. <https://doi.org/10.1007/s11032-014-0141-0>
- Bauchet, G., & Causse, M. (2012). Genetic Diversity in Tomato (*Solanum lycopersicum*) and Its Wild Relatives. *Genetic Diversity in Plants*, May. <https://doi.org/10.5772/33073>
- Bauchet, G., Grenier, S., Samson, N., Bonnet, J., Grivet, L., & Causse, M. (2017). Use of modern tomato breeding germplasm for deciphering the genetic control of agronomical traits by Genome Wide Association study. *Theoretical and Applied Genetics*, 130(5), 875–889. <https://doi.org/10.1007/s00122-017-2857-9>
- Bauchet, G., Grenier, S., Samson, N., Segura, V., Kende, A., Beekwilder, J., Cankar, K., Gallois, J. L., Gricourt, J., Bonnet, J., Baxter, C., Grivet, L., & Causse, M. (2017). Identification of major loci and genomic regions controlling acid and volatile content in tomato fruit: implications for flavor improvement. *New Phytologist*, 215(2), 624–641. <https://doi.org/10.1111/nph.14615>
- Bergougnoux, V. (2014). The history of tomato: From domestication to biopharming. *Biotechnology Advances*, 32(1), 170–189. <https://doi.org/10.1016/j.biotechadv.2013.11.003>
- Blanca, J., Cañizares, J., Cordero, L., Pascual, L., Díez, M. J., & Nuez, F. (2012). Variation Revealed by SNP Genotyping and Morphology Provides Insight into the Origin of the Tomato. *PLoS ONE*, 7(10). <https://doi.org/10.1371/journal.pone.0048198>
- Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., Francis, D., Causse, M., Knaap, E. Van Der, & Cañizares, J. (2015). Genomic variation in tomato , from wild ancestors to contemporary breeding accessions. *BMC Genomics*, 257. <https://doi.org/10.1186/s12864-015-1444-1>
- Blauth, S. L., Churchill, G. A., & Mutschler, M. A. (1998). Identification of quantitative trait loci associated with acylsugar accumulation using intraspecific populations of the wild tomato, *Lycopersicon pennellii*. *Theoretical and Applied Genetics*, 96(3–4), 458–467. <https://doi.org/10.1007/s001220050762>
- Bohner, J., & Bangerth, F. (1988). Cell number, cell size and hormone levels in semi-isogenic mutants of *Lycopersicon pimpinellifolium* differing in fruit size. *Physiologia Plantarum*, 72, 316–320.
- Caicedo, A. L., & Schaal, B. A. (2004). Population structure and phylogeography of *Solanum*

- pimpinellifolium inferred from a nuclear gene. *Molecular Ecology*, 13(7), 1871–1882. <https://doi.org/10.1111/j.1365-294X.2004.02191.x>
- Capel, C., del Carmen, A.F., Alba, J. M., Lima-Silva, V., Hernández-Gras, F., Salinas, M., Boronat, A., Angosto, T., Botella, M. A., Fernández-Muñoz, R., Granell, A., Capel, J., & Lozano, R. (2015). Wide-genome QTL mapping of fruit quality traits in a tomato RIL population derived from the wild-relative species *Solanum pimpinellifolium* L. *Theoretical and Applied Genetics*, 128(10), 2019–2035. <https://doi.org/10.1007/s00122-015-2563-4>
- Carter, C. D., & Snyder, J. C. (1985). Mite responses in relation to trichomes of *Lycopersicon esculentum* x *L. hirsutum* F2 hybrids. *Euphytica*, 34(1), 177–185. <https://doi.org/10.1007/BF00022877>
- Causse, M., Buret, M., Robini, K., & Verschave, P. (2003). Inheritance of Nutritional and Sensory Quality Traits in Fresh Market Tomato and Relation to Consumer Preferences. *Journal of Food Science*, 68(7), 2342–2350.
- Causse, M., Friguet, C., Coiret, C., Lépicier, M., Navez, B., Lee, M., Holthuysen, N., Sinesio, F., Moneta, E., & Grandillo, S. (2010). Consumer Preferences for Fresh Tomato at the European Scale: A Common Segmentation on Taste and Firmness. *Journal of Food Science*, 75(9). <https://doi.org/10.1111/j.1750-3841.2010.01841.xi>
- Chakrabarti, M., Zhang, N., Sauvage, C., Muños, S., Blanca, J., Cañizares, J., Diez, M. J., Schneider, R., Mazourek, M., McClead, J., Causse, M., & van der Knaap, E. (2013). A cytochrome P450 regulates a domestication trait in cultivated tomato. *Proceedings of the National Academy of Sciences of the United States of America*, 110(42), 17125–17130. <https://doi.org/10.1073/pnas.1307313110>
- Channarayappa, C., Shivashankar, G., Muniyappa, V., & Frist, R. H. (1992). Resistance of *Lycopersicon* species to *Bemisia tabaci*, a tomato leaf curl virus vector. *Can. J. Bot.*, 70(Stoner 1970), 2184–2192. <https://doi.org/10.1139/b92-270>
- Chen, F. Q., Foolad, M. R., Hyman, J., St Clair, D. A., & Beelaman, R. B. (1999). Mapping of QTLs for lycopene and other fruit traits in a *Lycopersicon esculentum* x *L. pimpinellifolium* cross and comparison of QTLs across tomato species. In *Molecular Breeding* (Vol. 5).
- Chen, K. Y., Cong, B., Wing, R., Vrebalov, J., & Tanksley, S. D. (2007). Changes in regulation of a transcription factor lead to autogamy in cultivated tomatoes. *Science*, 318(5850), 643–645. <https://doi.org/10.1126/science.1148428>
- Chetelat, R. T., DeVerna -A B Bennett, J. W., Bennett, A. B., & DeVerna, J. W. (1995). Effects of the *Lycopersicon chmielewskii* sucrose accumulator gene (sucr) on fruit yield and quality parameters following introgression into tomato. In *Theor Appl Genet* (Vol. 91). Springer-Verlag.
- Cuartero, J., Gomez-Guillamon, M. L., & Diaz, A. (1985). Catalog of collections of *Lycopersicon* from Peruvian Central Areas. *TGC Report*, 35, 32–35.
- De Candolle, A. (1886). Plants Cultivated for Their Seeds. In *Origin of Cultivated Plants:*

Vol. XLIX (2nd ed., pp. 376–384). Kegan Paul Trench & Co.

- De Jong, J., & Honma, S. (1976). EVALUATION OF SCREENING TECHNIQUES AND DETERMINATION OF CRITERIA FOR ASSESSING RESISTANCE TO CORYNEBACTERIUM MICHIGANENSE IN TOMATO 1. In *Euphytica* (Vol. 25).
- de Oliveira, J. R. F., de Resende, J. T. V., Maluf, W. R., Lucini, T., de Lima Filho, R. B., de Lima, I. P., & Nardi, C. (2018). Trichomes and allelochemicals in tomato genotypes have antagonistic effects upon behavior and biology of tetranychus urticae. *Frontiers in Plant Science*, 9(August), 1–9. <https://doi.org/10.3389/fpls.2018.01132>
- Délices, G., Leyva Ovalle, O. R., Mota-Vargas, C., Núñez Pastrana, R., Gámez Pastrana, R., Meza, P. A., & Serna-Lagunes, R. (2019). Biogeografía del tomate *Solanum lycopersicum* var. *cerasiforme* (Solanaceae) en su centro de origen (sur de América) y de domesticación (México). *Revista de Biología Tropical*, 67(4), 1023–1036.
- Dorst, J. C. E. A. (1946). *Een en twintigste beschrijvende rassenlijst voor landbouwgewassen*.
- Dunal, M. F. (1813). *Histoire naturelle, medicinale et économique des Solanum*. Paris: A. Koenig.
- Fernández-Muñoz, R., Salinas, M., Álvarez, M., & Cuartero, J. (2003). *Inheritance of Resistance to Two-spotted Spider Mite and Glandular Leaf Trichomes in Wild Tomato Lycopersicon pimpinellifolium (Jusl.) Mill.* 128(2), 188–195.
- 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. (2013). Identification and QTL mapping of whitefly resistance components in *Solanum galapagense*. *Theoretical and Applied Genetics*, 126(6), 1487–1501. <https://doi.org/10.1007/s00122-013-2067-z>
- Firdaus, S., van Heusden, A. W., Hidayati, N., Supena, E. D. J., Visser, R. G. F., & Vosman, B. (2012). Resistance to *Bemisia tabaci* in tomato wild relatives. *Euphytica*, 187(1), 31–45. <https://doi.org/10.1007/s10681-012-0704-2>
- Foolad, M. R. (2007). Genome mapping and molecular breeding of tomato. *International Journal of Plant Genomics*, 2007. <https://doi.org/10.1155/2007/64358>
- Foolad, M. R., Merk, H. L., & Ashrafi, H. (2008). Genetics, genomics and breeding of late blight and early blight resistance in tomato. *Critical Reviews in Plant Sciences*, 27(2), 75–107. <https://doi.org/10.1080/07352680802147353>
- Francis, F. J. (1995). QUALITY AS INFLUENCED BY COLOR. *Food Quality and Preference*, 6, 149–155.
- Frary, A., Nesbitt, T. C., Frary, A., Grandillo, S., Cong, B., Liu, J., Meller, J., Elber, R., Alpert, K. B., & Tanksley, S. D. (2000). *fw2.2: A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size.* 289(July), 85–88. <https://doi.org/10.1126/science.289.5476.85>
- Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., Burzynski-chang, E. A., Fish, T. L., Stromberg, K. A., Sacks, G. L., Thannhauser, T. W., Foolad, M. R., Diez, M. J., Blanca, J., Canizares, J., Xu, Y., Knaap, E. van der, Huang, S., Klee, H. J., ... Fei, Z.

- (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nature Genetics*, *51*, 1044–1051. <https://doi.org/10.1038/s41588-019-0410-2>
- García-Cano, E., Resende, R. O., Boiteux, L. S., Giordano, L. B., Fernández-Muñoz, R., & Moriones, E. (2008). Phenotypic Expression, Stability, and Inheritance of a Recessive Resistance to Monopartite Begomoviruses Associated with Tomato Yellow Leaf Curl Disease in Tomato. *Phytopathology*, *98*(5). <https://doi.org/10.1094/PHYTO-98-5-0618>
- Georgiady, M. S., Whitkus, R. W., & Lord, E. M. (2002). Genetic Analysis of Traits Distinguishing Outcrossing and Self-Pollinating Forms of Currant Tomato, *Lycopersicon pimpinellifolium* (Jusl.) Mill. *Genetics*, *161*(1), 333–344.
- Gibson, M. J. S., & Moyle, L. C. (2020). Regional differences in the abiotic environment contribute to genomic divergence within a wild tomato species. *Molecular Ecology*, *29*(12), 2204–2217. <https://doi.org/10.1111/mec.15477>
- Glas, J. J., Schimmel, B. C. J., Alba, J. M., Escobar-Bravo, R., Schuurink, R. C., & Kant, M. R. (2012). Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences*, *13*(12), 17077–17103. <https://doi.org/10.3390/ijms131217077>
- Goffreda, J. C., & Mutschler, M. A. (1989). Inheritance of potato aphid resistance in hybrids between *Lycopersicon esculentum* and *L. pennellii*. *Theoretical and Applied Genetics*, *78*, 210–216.
- Gong, H., Rehman, F., Yang, T., Li, Z., Zeng, S., Pan, L., Li, Y., & Wang, Y. (2019). Construction of the first high-density genetic map and QTL mapping for photosynthetic traits in *Lycium barbarum* L. *Molecular Breeding*, *39*(7). <https://doi.org/10.1007/s11032-019-1000-9>
- Grandillo, S., Ku, H. M., & Tanksley, S. D. (1999). Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theoretical and Applied Genetics*, *99*(6), 978–987. <https://doi.org/10.1007/s001220051405>
- Grandillo, S., Termolino, P., & van der Knaap, E. (2013). Molecular mapping of Complex Traits in Tomato. In *Genetics, Genomic and Breeding of Tomato* (pp. 150–228). CRC Press.
- Grandillo, S., & Tanksley, S. D. (1996). QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersicon pimpinellifolium*. *TAG Theoretical and Applied Genetics*, *92*, 935–951.
- Grandillo, S., & Cammareri, M. (2016). Molecular Mapping of Quantitative Trait Loci in Tomato. In Causse M., Giovannoni J., Bouzayen M., & Zouine M. (Eds.), *The Tomato Genome. Compendium of Plant Genomes*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-53389-5_4
- Guo, Z., Weston, P. A., & Snyder, J. C. (1993). Repellency to two-spotted spider mite, *Tetranychus urticae* Koch, as related to leaf surface chemistry of *Lycopersicon hirsutum* accessions. *Journal of Chemical Ecology*, *19*(12), 2965–2979. <https://doi.org/10.1007/BF00980596>

- Gurr, G. M., & McGrath, D. (2001). Effect of plant variety, plant age and photoperiod on glandular pubescence and host-plant resistance to potato moth (*Phthorimaea operculella*) in *Lycopersicon* spp. *Annals of Applied Biology*, 138(2), 221–230. <https://doi.org/10.1111/j.1744-7348.2001.tb00106.x>
- Hanson, A. A., Lorenz, A. J., Hesler, L. S., Bhusal, S. J., Bansal, R., Michel, A. P., Jiang, G., & Koch, R. L. (2018). Genome-Wide Association Mapping of Host-Plant Resistance to Soybean Aphid. *The Plant Genome*, 11(3), 180011. <https://doi.org/10.3835/plantgenome2018.02.0011>
- Hardigan, M. A., Laimbeer, F. P. E., Newton, L., Crisovan, E., Hamilton, J. P., Vaillancourt, B., Wiegert-Rininger, K., Wood, J. C., Douches, D. S., Farré, E. M., Veilleux, R. E., & Buell, C. R. (2017). Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proceedings of the National Academy of Sciences of the United States of America*, 114(46), E9999–E10008. <https://doi.org/10.1073/pnas.1714380114>
- Huang, X., & Han, B. (2014). Natural variations and genome-wide association studies in crop plants. In *Annual Review of Plant Biology* (Vol. 65, pp. 531–551). Annual Reviews Inc. <https://doi.org/10.1146/annurev-arplant-050213-035715>
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z., Li, M., Fan, D., Guo, Y., Wang, A., Wang, L., Deng, L., Li, W., Lu, Y., Weng, Q., Han, B. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics*, 42(11), 961–967. <https://doi.org/10.1038/ng.695>
- Huang, Z., & van der Knaap, E. (2011). Tomato fruit weight 11.3 maps close to fasciated on the bottom of chromosome 11. *Theoretical and Applied Genetics*, 123(3), 465–474. <https://doi.org/10.1007/s00122-011-1599-3>
- Hufford, M. B., Berny Mier Y Teran, J. C., & Gepts, P. (2019). *Annual Review of Plant Biology Crop Biodiversity: An Unfinished Magnum Opus of Nature*. <https://doi.org/10.1146/annurev-arplant-042817>
- Hutton, S. F., Scott, J. W., Yang, W., Sim, S. C., Francis, D. M., & Jones, J. B. (2010). Identification of QTL associated with resistance to bacterial spot race T4 in tomato. *Theoretical and Applied Genetics*, 121(7), 1275–1287. <https://doi.org/10.1007/s00122-010-1387-5>
- IPGRI. (1996). *Descriptors for Tomato (Lycopersicon spp.)*.
- Jarret, R. L., Sayama, H., & Tigchelaar, E. C. (1984). Pleiotropic effects associated with the chlorophyll intensifier mutations high pigment and dark green in tomatoes. *Journal of the American Society of Horticultural Science*, 109, 873–878.
- Jenkins, J. A. (1948). The Origin of the Cultivated Tomato. *Economic Botany*, 379–392.
- Jenkins, J. A., & Mackinney, G. (1955). CAROTENOIDS OF THE APRICOT TOMATO AND ITS HYBRIDS WITH YELLOW AND TANGERINE. *Genetics*, 40(5), 715–720.
- Joosten, M. H. A. J., & de Wit, P. J. G. M. (1999). The tomato-*Cladosporium fulvum*

- interaction: A versatile experimental system to study plant-pathogen interactions. *Annual Review of Phytopathology*, 37(1). <https://doi.org/10.1146/annurev.phyto.37.1.335>
- Kachanovsky, D. E., Filler, S., Isaacson, T., & Hirschberg, J. (2012). Epistasis in tomato color mutations involves regulation of phytoene synthase 1 expression by cis-carotenoids. *Proceedings of the National Academy of Sciences of the United States of America*, 109(46), 19021–19026. <https://doi.org/10.1073/pnas.1214808109>
- Kaloo, G. (Ed.). (1991). *Genetic Improvement of Tomato*. Springer-Verlag Berlin Heidelberg. <https://doi.org/10.1111/j.1526-100X.2009.00595.x>
- Kinkade, M. P., & Foolad, M. R. (2013). Validation and fine mapping of lyc12.1, a QTL for increased tomato fruit lycopene content. *Theoretical and Applied Genetics*, 126(8), 2163–2175. <https://doi.org/10.1007/s00122-013-2126-5>
- Klein, R. J., Zeiss, C., Chew, E. Y., Tsai, J. Y., Sackler, R. S., Haynes, C., Henning, A. K., SanGiovanni, J. P., Mane, S. M., Mayne, S. T., Bracken, M. B., Ferris, F. L., Ott, J., Barnstable, C., & Hoh, J. (2005). Complement factor H polymorphism in age-related macular degeneration. *Science*, 308(5720), 385–389. <https://doi.org/10.1126/science.1109557>
- Kopeliovitch, E., Mizrahi, Y., Rabinowitch, H. D., & Kedar, N. (1980). Physiology of the tomato mutant alcobaca. *Physiologia Plantarum*, 48(2), 307–311. <https://doi.org/10.1111/j.1399-3054.1980.tb03260.x>
- Kraus, C. M., Munkvold, K. R., & Martin, G. B. (2016). Natural Variation in Tomato Reveals Differences in the Recognition of AvrPto and AvrPtoB Effectors from *Pseudomonas syringae*. *Molecular Plant*, 9(5), 639–649. <https://doi.org/10.1016/j.molp.2016.03.001>
- Labate, J. A., Grandillo, S., Fulton, T., Muñoz, S., Caicedo, A. L., Peralta, I., Yuanfu, J., Chetelat, R. T., J. W. Scott, Gonzalo, M. J., Francis, D., Wencai, Y., van der Knaap, E., Baldo, A. M., Smith-White, B., Mueller, L. A., Prince, J. P., Blanchard, N. E., Storey, D. B., ... Causse, M. (2007). Tomato. In C. Kole (Ed.), *GenomeMapping andMolecular Breeding in Plants - Vegetables* (pp. 1–95). Springer Berlin Heidelberg New York.
- Lanahan, M. B., Yen Hsiao Ching, Giovannoni, J. J., & Klee, H. J. (1994). The Never ripe mutation blocks ethylene perception in tomato. *Plant Cell*, 6(4), 521–530. <https://doi.org/10.1105/tpc.6.4.521>
- Lawson, D. M., Lunde, C. F., & Mutschler, M. A. (1997). Marker-assisted transfer of acylsugar-mediated pest resistance from the wild tomato, *Lycopersicon pennellii*, to the cultivated tomato, *Lycopersicon esculentum*. In *Molecular Breeding* (Vol. 3). Kluwer Academic Publishers.
- Leckie, B. M., de Jong, D. M., & Mutschler, M. A. (2012). Quantitative trait loci increasing acylsugars in tomato breeding lines and their impacts on silverleaf whiteflies. *Molecular Breeding*, 30(4), 1621–1634. <https://doi.org/10.1007/s11032-012-9746-3>
- Lemke, C. A., & Mutschler, M. (1984). Inheritance of glandular trichomes in crosses between *Lycopersicon esculentum* and *Lycopersicon pennellii*. *Journal of the American Society for Horticultural Science*, 109, 592–596.

- Lester, R. N. (1991). Evolutionary relationships of tomato, potato and pepino, and wild species of *Lycopersicon* and *Solanum*. In R. N. Hawkes, M. N. Lester, & N. Estrada (Eds.), *Solanaceae III: taxonomy, chemistry, evolution* (pp. 283–301). Royal Botanic Gardens.
- Levin, I., Gilboa, N., Yeselson, E., Shen, S., & Schaffer, A. A. (2000). Fgr, a major locus that modulates the fructose to glucose ratio in mature tomato fruits. *Theoretical and Applied Genetics Volume*, *100*, 256–262.
- Levin, I., de Vos, C., Tadmor, Y., Bovy, A., Lieberman, M., Oren-Shamir, M., Segev, O., Kolotilin, I., Keller, M., Ovadia, R., Meir, A., & Bino, R. (2006). High pigment tomato mutants—more than just lycopene (a review). *Israel Journal of Plant Sciences*, *54*(3). https://doi.org/10.1560/IJPS_54_3_179
- Li, S., Yang, G., Yang, S., Just, J., Yan, H., Zhou, N., Jian, H., Wang, Q., Chen, M., Qiu, X., Zhang, H., Dong, X., Jiang, X., Sun, Y., Zhong, M., Bendahmane, M., Ning, G., Ge, H., Hu, J. Y., & Tang, K. (2019). The development of a high-density genetic map significantly improves the quality of reference genome assemblies for rose. *Scientific Reports*, *9*(1). <https://doi.org/10.1038/s41598-019-42428-y>
- Lieberman, M., Segev, O., Gilboa, N., Lalazar, A., & Levin, I. (2004). The tomato homolog of the gene encoding UV-damaged DNA binding protein 1 (DDB1) underlined as the gene that causes the high pigment-1 mutant phenotype. *Theoretical and Applied Genetics*, *108*(8), 1574–1581. <https://doi.org/10.1007/s00122-004-1584-1>
- Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., Zhang, Z., Lun, Y., Li, S., Wang, X., Huang, Z., Li, J., Zhang, C., Wang, T., Zhang, Y., Wang, A., Zhang, Y., Lin, K., Li, C., ... Huang, S. (2014). Genomic analyses provide insights into the history of tomato breeding. In *Nature Genetics* (Vol. 46, Issue 11, pp. 1220–1226). <https://doi.org/10.1038/ng.3117>
- Linneo, C. (1753). *Species plantarum*. Stockholm: L. Salvius.
- Lippman, Z. B., Cohen, O., Alvarez, J. P., Abu-Abied, M., Pekker, I., Paran, I., Eshed, Y., & Zamir, D. (2008). The making of a compound inflorescence in tomato and related nightshades. *PLoS Biology*, *6*(11), 2424–2435. <https://doi.org/10.1371/journal.pbio.0060288>
- Liu, T. J., Zhang, Y. J., Agerbirk, N., Wang, H. P., Wei, X. C., Song, J. P., He, H. J., Zhao, X. Z., Zhang, X. H., & Li, X. X. (2019). A high-density genetic map and QTL mapping of leaf traits and glucosinolates in *Barbarea vulgaris*. *BMC Genomics*, *20*(1). <https://doi.org/10.1186/s12864-019-5769-z>
- Luckwill, L. C. (1943). The Genus *Lycopersicon*: An Historical, Biological and Taxonomic Survey of the Wild and Cultivated Tomatoes. In U. Press (Ed.), *Aberdeen University Studies*.
- Maliepaard, C., Bas, N., Van Heusden, S., Kos, J., Pet, G., Verkerk, R., Vrieling, R., Zabel, P., & Lindhout, P. (1995). Mapping of QTLs for glandular trichome densities and trialeurodes vaporariorum (Greenhouse whitefly) resistance in an F2 from *lycopersicon esculentum* × *lycopersicon hirsutum* f. *glabratum*. *Heredity*, *75*(4), 425–433.

<https://doi.org/10.1038/hdy.1995.155>

- Mammadov, J., Aggarwal, R., Buyyarapu, R., & Kumpatla, S. (2012). SNP markers and their impact on plant breeding. *International Journal of Plant Genomics*, 2012. <https://doi.org/10.1155/2012/728398>
- 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., & Díez, M. J. (2020). Exploiting the diversity of tomato: the development of a phenotypically and genetically detailed germplasm collection. *Horticulture Research*, 7(1). <https://doi.org/10.1038/s41438-020-0291-7>
- Mazidi, M., Ferns, G. A., & Banach, M. (2020). A high consumption of tomato and lycopene is associated with a lower risk of cancer mortality: results from a multi-ethnic cohort. *Public Health Nutrition*. <https://doi.org/10.1017/S1368980019003227>
- McDowell, E. T., Kapteyn, J., Schmidt, A., Li, C., Kang, J. H., Descour, A., Shi, F., Larson, M., Schillmiller, A., An, L., Daniel Jones, A., Pichersky, E., Soderlund, C. A., & Gang, D. R. (2011). Comparative Functional Genomic Analysis of Solanum Glandular Trichome Types. *Plant Physiology*, 155(1), 524–539. <https://doi.org/10.1104/pp.110.167114>
- Meyer, R. S., & Purugganan, M. D. (2013). Evolution of crop species: Genetics of domestication and diversification. In *Nature Reviews Genetics* (Vol. 14, Issue 12). <https://doi.org/10.1038/nrg3605>
- Mieslerová, B., Lebeda, A., & Chetelat, T. (2000). Variation in Response of Wild *Lycopersicon* and *Solanum* spp. against Tomato Powdery Mildew (*Oidium lycopersici*) *Journal of Phytopathology*, 148, 303–311.
- Miller, P. (1754). The gardeners dictionary abridged. In *London: Printed for the author*.
- Monforte, A. J., Diaz, A., Caño-Delgado, A., & van der Knaap, E. (2014). The genetic basis of fruit morphology in horticultural crops: Lessons from tomato and melon. In *Journal of Experimental Botany* (Vol. 65, Issue 16, pp. 4625–4637). <https://doi.org/10.1093/jxb/eru017>
- Mu, Q., Huang, Z., Chakrabarti, M., Illa-Berenguer, E., Liu, X., Wang, Y., Ramos, A., & van der Knaap, E. (2017). Fruit weight is controlled by Cell Size Regulator encoding a novel protein that is expressed in maturing tomato fruits. *PLoS Genetics*, 13(8), 1–26. <https://doi.org/10.1371/journal.pgen.1006930>
- Müller, C. H. (1940). *A revision of the genus Lycopersicon*. Washington, D.C. : U.S. Dept. of Agriculture.
- Munos, S., Ranc, N., Botton, E., Berard, A., Rolland, S., Duffe, P., Carretero, Y., Le Paslier, M.-C., Delalande, C., Bouzayen, M., Brunel, D., & Causse, M. (2011). Increase in Tomato Locule Number Is Controlled by Two Single-Nucleotide Polymorphisms Located Near WUSCHEL. *Plant Physiology*, 156(4), 2244–2254. <https://doi.org/10.1104/pp.111.173997>

- Mutschler, M. A., Doerge, R. W., Liu, S. C., Kuai, J. P., Liedl, B. E., & Shapiro, J. A. (1996). QTL analysis of pest resistance in the wild tomato *Lycopersicon pennellii*: QTLs controlling acylsugar level and composition. *Theoretical and Applied Genetics*, 92(6), 709–718. <https://doi.org/10.1007/BF00226093>
- Mutschler, M.A., Wolfe, D. W., Cobb, E. D., & Yourstone, K. S. (1992). Tomato Fruit Quality and Shelf Life in Hybrids Heterozygous for the alc Ripening Mutant. *HortScience*, 27, 1985–1988.
- Nakazato, T., Bogonovich, M., & Moyle, L. C. (2008). Environmental factors predict adaptive phenotypic differentiation within and between two wild Andean tomatoes. *Evolution*, 62(4), 774–792. <https://doi.org/10.1111/j.1558-5646.2008.00332.x>
- Nesbitt, T. C., & Tanksley, S. D. (2002). Comparative Sequencing in the Genus *Lycopersicon*: Implications for the Evolution of Fruit Size in the Domestication of Cultivated Tomatoes. *Genetics Society of America*, 162, 365–379.
- Nuez, F., & Díez, M. J. (2008). Tomato. In J. Prohens & F. Nuez (Eds.), *Vegetables II. Handbook of Plant Breeding* (pp. 249–323). Springer. https://doi.org/10.1007/978-0-387-74110-9_7
- Oliveira, J.R.F., de Resende J.T.V, Maluf, W.R., Lucini, T., de Lima Filho, R.B., de Lima, I.P., and Nardi, C. (2018). Trichomes and Allelochemicals in Tomato Genotypes Have Antagonistic Effects upon Behavior and Biology of *Tetranychus Urticae*. *Frontiers in Plant Science* (9 (August): 1–9). <https://doi.org/10.3389/fpls.2018.01132>.
- Olsen, K. M., & Wendel, J. F. (2013). A bountiful harvest: Genomic insights into crop domestication phenotypes. In *Annual Review of Plant Biology* (Vol. 64, pp. 47–70). <https://doi.org/10.1146/annurev-arplant-050312-120048>
- Organización de las Naciones Unidas para la Agricultura y la Alimentación. (2008). *El estado de la inseguridad alimentaria en el mundo 2008: los precios elevados de los alimentos y la seguridad alimentaria; amenazas y oportunidades*. 56.
- Oriani, M. A., Vendramim, J. D., Oriani, M. A. de G., & Vendramim, J. D. (2010). Influence of trichomes on attractiveness and ovipositional preference of *Bemisia tabaci* (Genn.) B biotype (Hemiptera: Aleyrodidae) on tomato genotypes. *Neotropical Entomology*, 39(6), 1002–1007. <https://doi.org/10.1590/S0103-90162011000100006>
- Papa, R., Bellucci, E., Rossi, M., Leonardi, S., Rau, D., Gepts, P., Nanni, L., & Attene, G. (2007). Tagging the signatures of domestication in common bean (*Phaseolus vulgaris*) by means of pooled DNA samples. *Annals of Botany*, 100(5), 1039–1051. <https://doi.org/10.1093/aob/mcm151>
- Paran, I., & van der Knaap, E. (2007). Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *Journal of Experimental Botany*, 58(14), 3841–3852. <https://doi.org/10.1093/jxb/erm257>
- Peralta, I. E., & Spooner, D. M. (2006). History, origin and early cultivation of tomato (*Solanaceae*). In M. K. Razdan & A. K. Mattoo (Eds.), *Genetic improvement of Solanaceous crops, vol. 2: Tomato*. Enfield: Science Publishers.

- Peralta, I. E., Spooner, D. M., & Knapp, S. (2008). Taxonomy of Wild Tomatoes and Their Relatives (Solanum sect. Lycopersicoides, sect. Juglandifolia, sect. Lycopersicon; Solanaceae). *Systematic Botany Monographs*, 84, 1–186.
- Picó, B., Sifres, A., Mónica, E., Díez, M. J., & Nuez, F. (2000). Searching for new resistance sources to tomato yellow leaf curl virus within a highly variable wild Lycopersicon genetic pool. *BIOTIC STRESSES*, 344–350.
- Purugganan, M. D. (2019). Evolutionary Insights into the Nature of Plant Domestication. In *Current Biology* (Vol. 29, Issue 14, pp. R705–R714). Cell Press. <https://doi.org/10.1016/j.cub.2019.05.053>
- Rakha, M., Bouba, N., Ramasamy, S., Regnard, J. L., & Hanson, P. (2016). Evaluation of wild tomato accessions (Solanum spp.) for resistance to two-spotted spider mite (Tetranychus urticae Koch) based on trichome type and acylsugar content. *Genetic Resources and Crop Evolution*, 64(5), 1011–1022. <https://doi.org/10.1007/s10722-016-0421-0>
- Rakha, M., Hanson, P., & Ramasamy, S. (2017). 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*, 64(2), 247–260. <https://doi.org/10.1007/s10722-015-0347-y>
- Rambla, J. L., Medina, A., Fernández-Del-Carmen, A., Barrantes, W., Grandillo, S., Cammareri, M., López-Casado, G., Rodrigo, G., Alonso, A., García-Martínez, S., Primo, J., Ruiz, J. J., Fernández-Muñoz, R., Monforte, A. J., & Granell, A. (2017). Identification, introgression, and validation of fruit volatile QTLs from a red-fruited wild tomato species. *Journal of Experimental Botany*, 68(3), 429–442. <https://doi.org/10.1093/jxb/erw455>
- Ranc, N., Mũos, S., Santoni, S., & Causse, M. (2008). A clarified position for solanum lycopersicum var. cerasiforme in the evolutionary history of tomatoes (solanaceae). *BMC Plant Biology*, 8, 1–18. <https://doi.org/10.1186/1471-2229-8-130>
- Ranjan, A., Ichihashi, Y., & Sinha, N. R. (2012). *The tomato genome: implications for plant breeding, genomics and evolution. Box 1*, 1–8.
- Rao, E. S., Kadirvel, P., Symonds, R. C., & Ebert, A. W. (2013). Relationship between survival and yield related traits in Solanum pimpinellifolium under salt stress. *Euphytica*, 190(2), 215–228. <https://doi.org/10.1007/s10681-012-0801-2>
- Rao, E. S., Kadirvel, P., Symonds, R. C., Geethanjali, S., Thontadarya, R. N., & Ebert, A. W. (2015). Variations in DREB1A and VP1.1 genes show association with salt tolerance traits in wild tomato (Solanum pimpinellifolium). *PLoS ONE*, 10(7). <https://doi.org/10.1371/journal.pone.0132535>
- Razifard, H., Ramos, A., Della Valle, A. L., Bodary, C., Goetz, E., Manser, E. J., Li, X., Zhang, L., Visa, S., Tieman, D., van der Knaap, E., & Caicedo, A. L. (2020). Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution*, 1–15. <https://doi.org/10.1093/molbev/msz297>

- Rick, C. M. (1958). The role of natural hybridization in the derivation of cultivated tomatoes of western South America. *Economic Botany*, 12, 346–367.
- Rick, C. M. (1960). HYBRIDIZATION BETWEEN LYCOPERSICON ESCULENTUM AND SOLANUM PENNELLII: PHYLOGENETIC AND CYTOGENETIC SIGNIFICANCE*. In *Univ. of Texas Publ* (Vol. 34, Issue 5721).
- Rick, C. M. (1979). Biosystematic studies in Lycopersicon and closely related species of Solanum. In L. S. symposium Series (Ed.), *The biology and taxonomy of Solanaceae* (pp. 667–677). Academic Press. London.
- Rick, C. M., & Holle, M. (1990). Andean Lycopersicon esculentum var. cerasiforme: genetic variation and its evolutionary significance. *Economic Botany*, 44(3 Supplement), 69–78. <https://doi.org/10.1007/BF02860476>
- Rick, C. M., & Yoder, J. I. (1988). Classical And Molecular Genetics Of Tomato: Highlights And Perspectives. *Annual Review of Genetics*, 22(1), 281–300. <https://doi.org/10.1146/annurev.genet.22.1.281>
- Rick, C.M. (1991). Tomato Paste: A Concentrated Review of Genetic Highlights From the Beginnings to the Advent of Molecular Genetics. *Genetics Society of America*, 128, 1–5.
- Rick, C. M. (1973). Potential Genetic Resources in Tomato Species: Clues from Observations in Native Habitats. In *Genes, Enzymes, and Populations*. Springer US. https://doi.org/10.1007/978-1-4684-2880-3_17
- Rick, C. M. (1984). Evolution of mating systems: evidence from allozyme variation. *Genetics: New Frontiers*, 216–221.
- Rick, C. M., & Butler, L. (1956). *Cytogenetics of the Tomato* (Vol. 8). [https://doi.org/10.1016/S0065-2660\(08\)60504-0](https://doi.org/10.1016/S0065-2660(08)60504-0)
- Rick, C. M., Holle, M., & Thorp, R. W. (1978). Rates of cross-pollination in Lycopersicon pimpinellifolium: Impact of genetic variation in floral characters. *Plant Systematics and Evolution*, 129(1–2), 31–44. <https://doi.org/10.1007/BF00988982>
- Rick, C. M. (1978). The Tomato. *Scientific American*.
- Rick, C. M., & Fobes, J. F. (1975). Allozyme Variation in the Cultivated Tomato and Closely Related Species. *Bulletin of the Torrey Botanical Club*, 102(6), 376–384.
- Rick, C. M., Fobes, J. F., & Holle, M. (1977). Genetic Variation in Lycopersicon pimpinellifolium: Evidence of Evolutionary Change in Mating Systems *. *Plant Systematics and Evolution*, 127, 139–170.
- Rivas, S., & Thomas, C. M. (2005). Molecular Interactions Between Tomato and the Leaf Mold Pathogen Cladosporium fulvum . *Annual Review of Phytopathology*, 43(1), 395–436. <https://doi.org/10.1146/annurev.phyto.43.040204.140224>
- Rodríguez, G. R., Muños, S., Anderson, C., Sim, S.-C., Michel, A., Causse, M., Gardener, B. B. M., Francis, D., & van der Knaap, E. (2011). Distribution of SUN, OVATE, LC, and

FAS in the Tomato Germplasm and the Relationship to Fruit Shape Diversity. *Plant Physiology*, 156(1), 275–285. <https://doi.org/10.1104/pp.110.167577>

- Rodríguez-Guzmán, E., Ruiz-Corral, J. A., Lépiz-Ildefonso, R., & Pimienta-Barrios, E. (2018). Adaptación climática y edáfica de *Solanum lycopersicum* var. cerasiforme en el Occidente de México. *ECORFAN Revista de Ciencias Ambientales y Recursos Naturales*, 4(12), 1–9. www.ecorfan.org/spain,revista@ecorfan.
- Rodríguez Guzmán, E., Vargas Canela, D., Sánchez González, J. de J., Lépiz Ildefonso, R., Rodríguez Contreras, A., Ruiz Corral, J. A., Puente Ovalle, P., & Miranda Medrano, R. (2009). Etnobotánica de *Solanum lycopersicum* var. cerasiforme en el occidente de México. *Naturaleza y Desarrollo*, 7(2), 45–57.
- Ronen, G., Carmel-Goren, L., Zamir, D., & Hirschberg, J. (2000). An alternative pathway to β -carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. *Proceedings of the National Academy of Sciences of the United States of America*, 97(20), 11102–11107. <https://doi.org/10.1073/pnas.190177497>
- Rothan, C., Diouf, I., & Causse, M. (2019). Trait discovery and editing in tomato. *Plant Journal*, 97(1), 73–90. <https://doi.org/10.1111/tpj.14152>
- Sacco, A., Ruggieri, V., Parisi, M., Festa, G., Rigano, M., Picarella, M. E., Mazzucato, A., & Barone, A. (2015). Exploring a Tomato Landraces Collection for Fruit-Related Traits by the Aid of a High- Throughput Genomic Platform. *PLoS ONE*, 10. <https://doi.org/10.1371/journal.pone.0137139>
- Saeidi, Z., Mallik, B., & Kulkarni, R. S. (2007). Inheritance of glandular trichomes and two-spotted spider mite resistance in cross *Lycopersicon esculentum* “Nandi” and *L. pennellii* “LA2963.” *Euphytica*, 154(1–2), 231–238. <https://doi.org/10.1007/s10681-006-9291-4>
- Saliba-Colombani, V., Causse, M., Langlois, D., Philouze, J., & Buret, M. (2001). Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theoretical and Applied Genetics*, 102, 259–272.
- Salinas, M., Capel, C., Alba, J. M., Mora, B., Cuartero, J., Fernández-Muñoz, R., Lozano, R., & Capel, J. (2013). Genetic mapping of two QTL from the wild tomato *Solanum pimpinellifolium* L. controlling resistance against two-spotted spider mite (*Tetranychus urticae* Koch). *Theoretical and Applied Genetics*, 126(1), 83–92. <https://doi.org/10.1007/s00122-012-1961-0>
- Sato, S., Tabata, S., Hirakawa, H., Asamizu, E., Shirasawa, K., Isobe, S., Kaneko, T., Nakamura, Y., Shibata, D., Aoki, K., Egholm, M., Knight, J., Bogden, R., Li, C., Shuang, Y., Xu, X., Pan, S., Cheng, S., Liu, X., ... Gianese, G. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 485(7400), 635–641. <https://doi.org/10.1038/nature11119>
- Sauvage, C., Segura, V., Bauchet, G., Stevens, R., Do, P. T., Nikoloski, Z., Fernie, A. R., Causse, M., Molekulare, P., & Germany, P. T. D. (2014). *Genome-Wide Association in Tomato Reveals 44 Candidate Loci for Fruit Metabolic Traits 1 [W]*. 165(July), 1120–1132. <https://doi.org/10.1104/pp.114.241521>

- Scott, J. W. (2007). Breeding for resistance to viral pathogens. In M. K. Razdan & A. K. Mattoo (Eds.), *Genetic Improvement of Solanaceous Crops: Vol. Vol 2: Tomato*. (pp. 447–474). Enfield, New Hampshire, USA: Science Publishers, US.
- Sifres, A., Picó, B., Blanca, J. M., De Frutos, R., & Nuez, F. (2006). Genetic structure of *Lycopersicon pimpinellifolium* (Solanaceae) populations collected after the ENSO event of 1997-1998. *Genetic Resources and Crop Evolution*, 54(2), 359–377. <https://doi.org/10.1007/s10722-005-5725-4>
- Silva Ferreira, D., Kevei, Z., Kurowski, T., De Noronha Fonseca, M. E., Mohareb, F., Boiteux, L. S., & Thompson, A. J. (2018). BIFURCATE FLOWER TRUSS: A novel locus controlling inflorescence branching in tomato contains a defective MAP kinase gene. *Journal of Experimental Botany*, 69(10), 2581–2593. <https://doi.org/10.1093/jxb/ery076>
- Sim, S. C., Durstewitz, G., Plieske, J., Wieseke, R., Ganal, M. W., van Deynze, A., Hamilton, J. P., Buell, C. R., Causse, M., Wijeratne, S., & Francis, D. M. (2012a). Development of a large snp genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE*, 7(7). <https://doi.org/10.1371/journal.pone.0040563>
- Sim, S. C., Durstewitz, G., Plieske, J., Wieseke, R., Ganal, M. W., van Deynze, A., Hamilton, J. P., Buell, C. R., Causse, M., Wijeratne, S., & Francis, D. M. (2012b). Development of a large snp genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE*, 7(7). <https://doi.org/10.1371/journal.pone.0040563>
- Sims, W. L. (1980). History of Tomato Production for Industry Around the World. In *Acta Horticulturae* (Issue 100, pp. 25–26). <https://doi.org/10.17660/actahortic.1980.100.1>
- Soumpourou, E., Iakovidis, M., Chartrain, L., Lyall, V., & Thomas, C. M. (2007). The *Solanum pimpinellifolium* Cf-ECP1 and Cf-ECP4 genes for resistance to *Cladosporium fulvum* are located at the Milky Way locus on the short arm of chromosome 1. *Theoretical and Applied Genetics*, 115(8), 1127–1136. <https://doi.org/10.1007/s00122-007-0638-6>
- Soyk, S., Lemmon, Z. H., Oved, M., Fisher, J., Liberatore, K. L., Park, S. J., Goren, A., Jiang, K., Ramos, A., van der Knaap, E., Van Eck, J., Zamir, D., Eshed, Y., & Lippman, Z. B. (2017). Bypassing Negative Epistasis on Yield in Tomato Imposed by a Domestication Gene. *Cell*, 169(6), 1142–1155. <https://doi.org/10.1016/j.cell.2017.04.032>
- Srinivas, T. R. (2001). Floral descriptors of field evaluated tomato germplasm. *Journal of Applied Horticulture*, 3(1), 53–55.
- Stebbins, G. L. (1957). Self Fertilization and Population Variability in the Higher Plants. *American Society of Naturalists*, 91(861), 337–354.
- Stevens, M. A., & Rick, C. M. (1986). Genetics and Breeding. In J. Atherton & J. Rudich (Eds.), *The Tomato Crop: a scientific basis for improvement* (p. 661). Chapman and Hall Ltf. <https://doi.org/10.1007/978-94-009-3137-4>
- Stevens, R., Buret, M., Duffé, P., Garchery, C., Baldet, P., Rothan, C., & Causse, M. (2007). Candidate genes and quantitative trait loci affecting fruit ascorbic acid content in three

- tomato populations. *Plant Physiology*, *143*(4), 1943–1953. <https://doi.org/10.1104/pp.106.091413>
- Stout, M. J., Kurabchew, H., & Leite, G. L. D. (2017). Host-Plant Resistance in Tomato. In *Sustainable Management of Arthropod Pests of Tomato*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-802441-6.00009-7>
- Studer, A., Zhao, Q., Ross-Ibarra, J., & Doebley, J. (2011). Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nature Genetics*, *43*(11), 1160–1163. <https://doi.org/10.1038/ng.942>
- Tanksley, S. D., Ganal, M. W., Prince, J. P., de Vicente, M. C., Bonierbale, M. W., Broun, P., Fulton, T. M., Giovannoni, J. J., Grandillo, S., Martin, G. B., Messeguer, R., Miller, J. C., Miller, L., Paterson, A. H., Riider, M. S., Wing, R. A., Wu, W., & Young, N. D. (1992). High Density Molecular Linkage Maps of the Tomato and Potato Genomes. *Genetics Society of America*, *132*, 1141–1160.
- Tanksley, Steven D. (2004). The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell*, *16*, 181–190. <https://doi.org/10.1105/tpc.018119>
- Tanksley, Steven D., & McCouch, S. R. (1997). Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild The Narrow Genetic Base of Crop Plants. *Science*, *277*(August), 1063–1066. <https://doi.org/10.1126/science.277.5329.1063>
- Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., Rambla, J. L., Beltran, K. S. O., Taylor, M., Zhang, B., Ikeda, H., Liu, Z., Fisher, J., Zemach, I., Monforte, A., Zamir, D., Granell, A., Kirst, M., Huang, S., & Klee, H. (2017). A chemical genetic roadmap to improved tomato flavor. *Science (New York, N.Y.)*, *355*(6323), 391–394. <https://doi.org/10.1126/science.aal1556>
- Tournefort, J. P. (1694). *Éléments de botanique*. Paris l'Imprimerie royale.
- 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. (2017). Loss of type-IV glandular trichomes is a heterochronic trait in tomato and can be reverted by promoting juvenility. *Plant Science*, *259*, 35–47. <https://doi.org/10.1016/j.plantsci.2017.03.006>
- 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. (2019). QTL mapping of insect resistance components of *Solanum galapagense*. *Theoretical and Applied Genetics*, *132*(2), 531–541. <https://doi.org/10.1007/s00122-018-3239-7>
- Warnock, S. J. (1991). Natural Habitats of *Lycopersicon* Species. *HortScience*, *26*(5), 466–471. <https://doi.org/10.21273/hortsci.26.5.466>
- Wen, J., Jiang, F., Weng, Y., Sun, M., Shi, X., Zhou, Y., Yu, L., & Wu, Z. (2019). Identification of heat-tolerance QTLs and high-temperature stress-responsive genes through conventional QTL mapping, QTL-seq and RNA-seq in tomato. *BMC Plant Biology*, *19*(1). <https://doi.org/10.1186/s12870-019-2008-3>
- Weston, P. A., & Snyder, J. C. (1990). Thumbtack bioassay: a quick method for measuring

- plant resistance to twospotted spider mites (Acari: Tetranychidae). *Journal of Economic Entomology*, 83(2), 500–504. <https://doi.org/10.1093/jee/83.2.500>
- Widrechner, M. P. (1987). Variation in Breeding System of *Lycopersicon pimpinellifolium*: Implications for Germplasm Maintenance. *Plant Genetic Resources Newsletter*, 70, 38–43.
- Wilkens, R. T., Shea, G. O., Halbreich, S., & Stamp, N. E. (1996). Resource availability and the trichome defenses of tomato plants. *Oecologia*, 106(2), 181–191. <https://doi.org/10.1007/BF00328597>
- Xiao, Y., Liu, H., Wu, L., Warburton, M., & Yan, J. (2017). Genome-wide Association Studies in Maize: Praise and Stargaze. In *Molecular Plant* (Vol. 10, Issue 3, pp. 359–374). Cell Press. <https://doi.org/10.1016/j.molp.2016.12.008>
- Ye, J., Li, W., Ai, G., Li, C., Liu, G., Chen, W., Wang, B., Wang, W., Lu, Y., Zhang, J., Li, H., Ouyang, B., Zhang, H., Fei, Z., Giovannoni, J. J., Ye, Z., & Zhang, Y. (2019). Genome-wide association analysis identifies a natural variation in basic helix-loop-helix transcription factor regulating ascorbate biosynthesis via D-mannose/L-galactose pathway in tomato. *PLoS Genetics*, 15(5), e1008149. <https://doi.org/10.1371/journal.pgen.1008149>
- Yen, H. C., Shelton, B. A., Howard, L. R., Vrebalov, S. L. J., & Giovannoni, J. J. (1997). The tomato high-pigment (hp) locus maps to chromosome 2 and influences plastome copy number and fruit quality. *Theoretical and Applied Genetics*, 95, 1069–1079.
- Yoo, H. J., Park, W. J., Lee, G. M., Oh, C. S., Yeam, I., Won, D. C., Kim, C. K., & Lee, J. M. (2017). Inferring the Genetic Determinants of Fruit Colors in Tomato by Carotenoid Profiling. *Molecules*, 22(5), 1–14. <https://doi.org/10.3390/molecules22050764>
- Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. *Nature Reviews Genetics*, 2(12), 983–989. <https://doi.org/10.1038/35103590>
- Zeggini, E., & Ioannidis, J. P. A. (2009). Meta-analysis in genome-wide association studies. In *Pharmacogenomics* (Vol. 10, Issue 2, pp. 191–201). <https://doi.org/10.2217/14622416.10.2.191>
- Zhao, J., Sauvage, C., Zhao, J., Bitton, F., Bauchet, G., Liu, D., Huang, S., Tieman, D. M., Klee, H. J., & Causse, M. (2019). Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. *Nature Communications*, 10(1), 1–12. <https://doi.org/10.1038/s41467-019-09462-w>
- Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., Lin, T., Qin, M., Peng, M., Yang, C., Cao, X., Han, X., Wang, X., van der Knaap, E., Zhang, Z., Cui, X., Klee, H., Fernie, A. R., Luo, J., & Huang, S. (2018). Rewiring of the Fruit Metabolome in Tomato Breeding. *Cell*, 172(1–2), 249–261. <https://doi.org/10.1016/j.cell.2017.12.019>
- Zsögön, A., Čermák, T., Naves, E. R., Notini, M. M., Edel, K. H., Weinl, S., Freschi, L., Voytas, D. F., Kudla, J., & Peres, L. E. P. (2018). De novo domestication of wild tomato using genome editing. *Nature Biotechnology*, 36(12), 1211–1216. <https://doi.org/10.1038/nbt.4272>

Zuriaga, E., Blanca, J. M., Cordero, L., Sifres, A., Blas-Cerdán, W. G., Morales, R., & Nuez, F. (2009). Genetic and bioclimatic variation in *Solanum pimpinellifolium*. *Genetic Resources and Crop Evolution*, 56(1), 39–51. <https://doi.org/10.1007/s10722-008-9340-z>

ANEXOS

ANEXO A: ANEXO CAPÍTULO 2.

Anexo I. Datos de pasaporte de las entradas. SP: *Solanum pimpinellifolium*. SLC: *Solanum lycopersicum* var. *cerasiforme*. SLL: *Solanum lycopersicum* var. *lycopersicum*. PER: Perú. ECU: Ecuador. MEX: México. MA: Mesoamérica.

Accesión	País	Lugar de recolecta	Latitud	Longitud	Banco	Grupo	Código
BGV007109	ECU	Pampa. Tacusa. Esmeraldas; Esmeraldas	0,99	-79,53	COMAV	SP_ECU	Pe1
BGV007111	ECU	Camarones. Esmeraldas	1,00	-79,56	COMAV	SP_ECU	Pe2
BGV007181	ECU	Pedernales. Manabí	0,07	-80,06	COMAV	SP_ECU	Pe3
BGV007169	ECU	El Humedal. Pedernales. Manabí	0,04	-79,95	COMAV	SP_ECU	Pe4
BGV007161	ECU	El Achote. Pedernales. Manabí	0,01	-79,88	COMAV	SP_ECU	Pe5
BGV007152	ECU	Puerto Nuevo. El Carmen. Manabí	-0,08	-79,68	COMAV	SP_ECU	Pe6
BGV007155	ECU	La Poderosa. Chone. Manabí	-0,09	-79,76	COMAV	SP_ECU	Pe7
BGV007158	ECU	Chibunga. Chone. Manabí	-0,09	-79,76	COMAV	SP_ECU	Pe8
BGV007151	ECU	Chilagualan. El Carmen. Manabí	-0,13	-79,59	COMAV	SP_ECU	Pe9
BGV007194	ECU	Punta Prieta. Jama. Manabí	-0,13	-80,22	COMAV	SP_ECU	Pe10
BGV007149	ECU	El Carmen. Manabí	-0,18	-79,52	COMAV	SP_ECU	Pe11
BGV007198	ECU	Tambuche. Jama. Manabí	-0,30	-80,33	COMAV	SP_ECU	Pe12
BGV006208	ECU	Chaguarpamba. Santa Isabel. Azuay	-3,31	-79,35	COMAV	SP_ECU	Pe13
BGV007339	ECU	Juana de Oro. Pasaje. El Oro	-3,32	-79,83	COMAV	SP_ECU	Pe14
BGV007366	ECU	Uzhcurrumi. Uzhcurrumi. El Oro	-3,33	-79,61	COMAV	SP_ECU	Pe15
BGV006370	PER	Mallares. Sullana. Piura	-4,85	-80,85	COMAV	SP_PER	Pp16
BGV006327	PER	Tambo Grande. Piura	-4,89	-80,38	COMAV	SP_PER	Pp17
BGV006454	PER	Piura	-5,15	-80,27	COMAV	SP_PER	Pp18
BGV006457	PER	Vicus. Morropón. Lambayeque. Piura	-5,17	-80,14	COMAV	SP_PER	Pp19
BGV006347	PER	Piura. Piura	-5,28	-80,11	COMAV	SP_PER	Pp20
BGV006336	PER	Buenos Aires. Morropón. Piura	-5,33	-79,92	COMAV	SP_PER	Pp21
BGV006353	PER	Lambayeque	-5,59	-79,97	COMAV	SP_PER	Pp22
BGV015380	PER	San Agustín. Bellavista. Jaén. Cajamarca	-5,60	-78,78	COMAV	SP_PER	Pp23
BGV016047	PER	Linderos. Jaén. Jaén. Cajamarca	-5,69	-78,80	COMAV	SP_PER	Pp24
BGV015382	PER	Magunchal. Jamalca. Utcubamba. Amazonas	-5,89	-78,18	COMAV	SP_PER	Pp25
BGV006363	PER	Olmos. Lambayeque	-6,00	-79,72	COMAV	SP_PER	Pp26
BGV006478	PER	Olmos. Lambayeque.	-6,03	-79,68	COMAV	SP_PER	Pp27
BGV007931	Mex	La Biznaga. Ahome. Sinaloa	26,06	-109,38	COMAV	SLC_MEX	Cm28

BGV008051	Mex	General Teran. Nuevo León	25,26	-99,68	TGRC (LA1455)	SLC_MEX	Cm29
BGV007927	Mex	El Vergel. Culiacán. Sinaloa	24,73	-107,79	COMAV	SLC_MEX	Cm30
BGV007933	Mex	Tecuala. Nayarit	22,39	-105,46	COMAV	SLC_MEX	Cm31
BGV008070	Mex	Rio Tamesi. Tamaulipas	22,23	-97,86	TGRC (LA1703)	SLC_MEX	Cm32
BGV007911	Mex	Cuesta de los monos. Tamazunchale. San Luis Potosí	21,99	-98,99	COMAV	SLC_MEX	Cm33
BGV007910	Mex	Palo de Arco. Ciudad Valles. San Luis de Potosí	21,91	-99,16	COMAV	SLC_MEX	Cm34
BGV007934	Mex	Sentispac. Nayarit	21,81	-105,34	COMAV	SLC_MEX	Cm35
BGV007935	Mex	Santispac. Nayarit	21,81	-105,34	COMAV	SLC_MEX	Cm36
BGV007920	Mex	El Balcón. Xilitla. San Luis Potosí	21,34	-99,04	COMAV	SLC_MEX	Cm37
BGV007918	Mex	Rioverdito. Querétaro	21,26	-99,13	COMAV	SLC_MEX	Cm38
BGV007921	Mex	Ahuacatlan de Guadalupe. Querétaro	21,21	-99,54	COMAV	SLC_MEX	Cm39
BGV007909	Mex	Huejutla. Hidalgo	21,14	-98,43	COMAV	SLC_MEX	Cm40
BGV007908	Mex	Tehuatlan. San Luis de Potosí	21,06	-98,51	COMAV	SLC_MEX	Cm41
BGV008058	Mex	Becal Ruins. Campeche	20,46	-90,03	TGRC (LA1545)	SLC_MEX	Cm42
BGV012614	Mex	Papantla de Olarte. Veracruz	20,46	-97,33	COMAV	SLC_MEX	Cm43
BGV007895	Mex	Xicotepac. Puebla	20,27	-97,95	COMAV	SLC_MEX	Cm44
BGV007901	Mex	Coahuila. Huauchinango. Puebla	20,19	-98,06	COMAV	SLC_MEX	Cm45
BGV007902	Mex	Huauchinango. Puebla	20,18	-96,06	COMAV	SLC_MEX	Cm46
BGV007899	Mex	Huauchinango. Puebla	20,17	-98,05	COMAV	SLC_MEX	Cm47
BGV007900	Mex	Huauchinango. Puebla	20,17	-98,05	COMAV	SLC_MEX	Cm48
BGV007894	Mex	Huauchinango. Puebla	20,17	-98,06	COMAV	SLC_MEX	Cm49
BGV008061	Mex	Jamapa. Vera Cruz	19,04	-96,24	TGRC (LA1569)	SLC_MEX	Cm50
BGV008067	Mex	Acapulco de Juárez. Guerrero	16,86	-99,91	TGRC (LA1668)	SLC_MEX	Cm51
BGV008347	HON	Intibucá	14,36	-88,26	COMAV	SLC_MA	Cma52
PI406890	HON	El Paraiso, 1 km from Danli	14,04	-86,57	USDA	SLC_MA	Cma53
BGV008223	HON	Morazán. Yoro	13,93	-87,21	COMAV	SLC_MA	Cma54
BGV008221	SAL	Ruinas de Santa Ana. San Salvador. La Libertad	13,83	-89,41	COMAV	SLC_MA	Cma55
BGV008345	SAL				COMAV	SLC_MA	Cma56
BGV008224	NIC	Río El Barco. Boaco	12,06	-85,46	COMAV	SLC_MA	Cma57
BGV008225	NIC	Santo Tomás. Chontales	12,03	-84,66	COMAV	SLC_MA	Cma58
BGV008219	CR	La Cruz. Guanacaste	11,08	-85,64	COMAV	SLC_MA	Cma59
BGV008218	CR	San Pablo. Heredia	10,03	-84,23	COMAV	SLC_MA	Cma60
BGV008354	CR	San José	9,94	-84,11	COMAV	SLC_MA	Cma61
BGV008348	CR	San José	9,94	-84,11	COMAV	SLC_MA	Cma62
BGV013134	CR	Turrialba. Cartago	9,91	-83,69	COMAV	SLC_MA	Cma63
LA1712	CR	Pejibaye	9,81	-83,71	TGRC	SLC_MA	Cma64

PI487625	CR				USDA	SLC_MA	Cma65
BGV012615	COL	Titiribí. Antioquia	6,06	-75,79	TGRC (LA2783)	SLC_MA	Cma66
PI129088	COL	Calarca	4,53	-75,64	USDA	SLC_MA	Cma67
BGV008110	COL	Mata de Cana. El Dovio. Valle del Cauca	4,49	-76,31	TGRC (LA2697)	SLC_MA	Cma68
BGV013175	COL	Villa Hermosa. Cauca	2,12	-76,98	TGRC (LA1425)	SLC_MA	Cma69
BGV012626	COL	Purillo, between El Bordo and Dolores	2,12	-76,98	COMAV	SLC_MA	Cma70
BGV012627	COL				COMAV	SLC_MA	Cma71
BGV004584	COL	Cauca			COMAV	SLC_MA	Cma72
BGV006753	ECU	Archidona. Archidona. Napo	-0,91	-77,81	COMAV	SLC_ECU	Ce73
BGV006767	ECU	Santa Inés. Napo	-0,96	-77,82	COMAV	SLC_ECU	Ce74
BGV006768	ECU	Puerto Napo. Napo	-1,04	-77,79	COMAV	SLC_ECU	Ce75
BGV006775	ECU	El Mirador. Puerto Misahuallí. Napo	-1,04	-77,69	COMAV	SLC_ECU	Ce76
BGV006777	ECU	El Mirador. Puerto Misahuallí. Napo	-1,04	-77,69	COMAV	SLC_ECU	Ce77
BGV006779	ECU	Las Colinas. Puerto Napo. Napo	-1,05	-77,80	COMAV	SLC_ECU	Ce78
BGV006792	ECU	Costa Azul. Tena. Napo	-1,07	-77,79	COMAV	SLC_ECU	Ce79
BGV006806	ECU	Nueva Esperanza. Arosamena Tola. Napo	-1,12	-77,82	COMAV	SLC_ECU	Ce80
BGV006825	ECU	Santa Clara. Pastaza	-1,24	-77,88	COMAV	SLC_ECU	Ce81
BGV006828	ECU	Recinto San Vicente. Pastaza	-1,32	-77,89	COMAV	SLC_ECU	Ce82
BGV007015	ECU	La Ciénaga. Baños de Agua Santa. Tungurahua	-1,40	-78,38	COMAV	SLC_ECU	Ce83
BGV007017	ECU	Río Verde. Tungurahua	-1,40	-78,30	COMAV	SLC_ECU	Ce84
BGV006852	ECU	Simón Bolívar. Pastaza	-1,70	-77,84	COMAV	SLC_ECU	Ce85
BGV006859	ECU	Tashapi. Pastaza	-1,72	-77,88	COMAV	SLC_ECU	Ce86
BGV006865	ECU	Pitirishca. Pastaza	-1,81	-77,83	COMAV	SLC_ECU	Ce87
BGV006867	ECU	Kunkuki. Pastaza	-1,85	-77,82	COMAV	SLC_ECU	Ce88
PI129026	ECU	Market Guayaquil	-2,20	-79,89	USDA	SLC_ECU	Ce89
PI129033	ECU	Market Guayaquil	-2,20	-79,89	USDA	SLC_ECU	Ce90
BGV006881	ECU	Macas. Morona Santiago	-2,30	-78,12	COMAV	SLC_ECU	Ce91
BGV007023	ECU	Arapicos. Morona Santiago	-2,39	-78,17	COMAV	SLC_ECU	Ce92
BGV006896	ECU	Sucúa. Morona Santiago	-2,45	-78,17	COMAV	SLC_ECU	Ce93
BGV006231	ECU	Sucúa. Sucúa. Morona Santiago	-2,48	-78,16	COMAV	SLC_ECU	Ce94
BGV006899	ECU	Huambi. Morona Santiago	-2,54	-78,17	COMAV	SLC_ECU	Ce95
BGV006901	ECU	Huambi. Morona Santiago	-2,54	-78,17	COMAV	SLC_ECU	Ce96
BGV006232	ECU	Logroño. Méndez. Morona Santiago	-2,61	-78,22	COMAV	SLC_ECU	Ce97
BGV006904	ECU	Logroño. Morona Santiago	-2,62	-78,20	COMAV	SLC_ECU	Ce98
BGV006907	ECU	Yurupari. Morona Santiago	-2,62	-78,20	COMAV	SLC_ECU	Ce99
BGV012639	ECU	Puna. Guayas	-2,74	-79,93	COMAV	SLC_ECU	Ce100
BGV006910	ECU	Chinganaza Nuevo. Morona	-2,75	-78,30	COMAV	SLC_ECU	Ce101

		Santiago					
BGV006229	ECU	Indanza. Limón Indanza. Morona Santiago	-3,06	-78,49	COMAV	SLC_ECU	Ce102
BGV006230	ECU	Indanza. Limón Indanza. Morona Santiago	-3,06	-78,49	COMAV	SLC_ECU	Ce103
BGV006235	ECU	Indanza. Limón Indanza. Morona Santiago	-3,06	-78,49	COMAV	SLC_ECU	Ce104
BGV006225	ECU	Comunidad Shuar Sharup. Gualaquiza. Morona Santiago	-3,51	-78,51	COMAV	SLC_ECU	Ce105
BGV006175	ECU	San Jacinto. Ayapamba. Atahualpa. El Oro	-3,63	-79,68	COMAV	SLC_ECU	Ce106
BGV006927	ECU	Pachicutza. Zamora Chinchipe	-3,67	-78,61	COMAV	SLC_ECU	Ce107
BGV006931	ECU	Pachicutza. Zamora Chinchipe	-3,67	-78,61	COMAV	SLC_ECU	Ce108
BGV006148	ECU	El Rosario. Chaguarpamba. Loja	-3,79	-79,61	COMAV	SLC_ECU	Ce109
BGV006934	ECU	Simón Bolívar. Yantzaza. Zamora Chinchipe	-3,79	-78,75	COMAV	SLC_ECU	Ce110
BGV005912	ECU	Yantzaza. Zamora Chinchipe	-3,84	-78,76	COMAV	SLC_ECU	Ce111
BGV006234	ECU	Timbara. Zamora. Zamora Chinchipe	-4,03	-78,88	COMAV	SLC_ECU	Ce112
BGV005895	ECU	Timbara. Zamora Chinchipe	-4,04	-78,89	COMAV	SLC_ECU	Ce113
BGV006906	ECU	Yurupari. Morona Santiago	-2,62	-78,20	COMAV	SLC_ECU	Ce114
BGV008100	PER	Jumbilla. Amazonas	-5,91	-77,79	TGRC (LA2312)	SLC_PER	Cp115
BGV008095	PER	Soritor. San Martín	-6,14	-77,09	TGRC (LA2251)	SLC_PER	Cp116
BGV015730	PER	El Triunfo. Japelacio. Moyobamba. San Martín	-6,15	-76,84	COMAV	SLC_PER	Cp117
BGV008096	PER	Casaría de Pacayzapa. San Martín	-6,19	-76,84	TGRC (LA2265)	SLC_PER	Cp118
BGV013161	PER	Lahuarpia. San Martín	-6,23	-76,86	TGRC (LA2263)	SLC_PER	Cp119
BGV014508	PER	Alta Tangarana-Pueblo Nuevo. Tabalosos. Lamas. San Martín	-6,32	-76,70	COMAV	SLC_PER	Cp120
BGV015727	PER	Ponazapa. Alonso Alvarado. Lamas. San Martín	-6,32	-76,68	COMAV	SLC_PER	Cp121
BGV008098	PER	Tabalosas. San Martín	-6,36	-76,66	TGRC (LA2278)	SLC_PER	Cp122
BGV014514	PER	Lamas. San Martín	-6,42	-76,52	COMAV	SLC_PER	Cp123
BGV014515	PER	Lamas. San Martín	-6,43	-76,51	COMAV	SLC_PER	Cp124
BGV014522	PER	San Rafael del Río Mayo. Cuñumbuqui. Lamas. San Martín	-6,44	-76,58	COMAV	SLC_PER	Cp125
BGV014516	PER	Rumisapa. Lamas. San Martín	-6,45	-76,47	COMAV	SLC_PER	Cp126
BGV014518	PER	Rumisapa. Lamas. San Martín	-6,45	-76,47	COMAV	SLC_PER	Cp127
BGV014519	PER	Rumisapa. Lamas. San Martín	-6,45	-76,47	COMAV	SLC_PER	Cp128
BGV015734	PER	Rumisapa. Lamas. San Martín	-6,45	-76,47	COMAV	SLC_PER	Cp129
BGV008065	PER	Tarapoto. San Martín	-6,51	-76,36	TGRC (LA1654)	SLC_PER	Cp130

BGV015726	PER	Bello Horizonte. Tarapoto. San Martín	-6,53	-76,30	COMAV	SLC_PER	Cp131
LA2309	PER	Puerto Santa Cruz	-6,62	-78,94	TGRC	SLC_PER	Cp132
BGV007981	PER	La Divisoria. Ucayali	-9,04	-74,49	COMAV	SLC_PER	Cp133
BGV008189	PER	Río Huallaga. Tingo María. Huánuco	-9,29	-79,99	COMAV	SLC_PER	Cp134
BGV008036	PER	Granja. Satipo. Junín	-11,25	-74,64	TGRC (LA1291)	SLC_PER	Cp135
BGV008108	PER	Yunguyo. Madre de Dios	-11,76	-70,65	TGRC (LA2686)	SLC_PER	Cp136
BGV008042	PER	Convento de Sivia. Pichari. Cuzco	-12,51	-73,86	TGRC (LA1313)	SLC_PER	Cp137
BGV008037	PER	Hotel Oasis. San Francisco. Ayacucho	-12,62	-73,79	TGRC (LA1307)	SLC_PER	Cp138
BGV008041	PER	Paisanato. Cuzco	-12,62	-73,78	TGRC (LA1312)	SLC_PER	Cp139
BGV012625	PER	Quillabamba. Cuzco	-12,84		COMAV	SLC_PER	Cp140
BGV007992	PER	Chilcachaca. Cuzco	-13,01	-72,54	COMAV	SLC_PER	Cp141
BGV013945	PER	Pilcopata. Cuzco	-13,08	-71,18	TGRC (LA2692)	SLC_PER	Cp142
BGV012613	PER	Ayacucho	-13,15	-74,22	COMAV	SLC_PER	Cp143
BGV012640	PER	Hacienda Carmen. Apurímac	-13,56	-72,63	COMAV	SLC_PER	Cp144
BGV007989	PER	Molinopata. Apurímac	-13,64	-72,93	COMAV	SLC_PER	Cp145
BGV007990	PER	Bellavista. Cotabambas. Apurímac	-13,92	-72,18	COMAV	SLC_PER	Cp146
BGV008106	PER	Yanahuanca. Puno	-14,25	-69,23	TGRC (LA2664)	SLC_PER	Cp147
BGV008077	PER	Arequipa	-16,41	-71,54	TGRC (LA1953)	SLC_PER	Cp148
BGV007872	Mex	Mercado. Oxkutzcab. Yucatán	20,31	-89,43	COMAV	SLL	L149
BGV007878	Mex	Mercado. Dzitbalché. Yucatán	20,33	-90,04	COMAV	SLL	L150
BGV007870	Mex	Mercado. Muna. Yucatán	20,49	-89,71	COMAV	SLL	L151
BGV007871	Mex	Mercado. Muna. Yucatán	20,49	-89,71	COMAV	SLL	L152
BGV007876	Mex	Mercado. Valladolid. Yucatán	20,69	-88,23	COMAV	SLL	L153
BGV007875	Mex	Deaboku. Yucatán	20,81	-89,41	COMAV	SLL	L154
BGV007867	Mex	Mercado. Umán. Yucatán	20,88	-89,74	COMAV	SLL	L155
BGV007936	Mex	Tekax. Mérida. Yucatán	20,98	-89,64	COMAV	SLL	L156
BGV007857	Mex	Xcanatun. Yucatán	21,09	-89,63	COMAV	SLL	L157
BGV007862	Mex	San Eduardo. Yucatán	21,23	-88,81	COMAV	SLL	L158
BGV007863	Mex	San Eduardo. Yucatán	21,23	-89,81	COMAV	SLL	L159
BGV007864	Mex	San Eduardo. Yucatán	21,23	-89,81	COMAV	SLL	L160
BGV007865	Mex	San Eduardo. Yucatán	21,23	-89,81	COMAV	SLL	L161
BGV007860	Mex	Hacienda Tchoilen. Dzidzantun. Yucatán	21,26	-89,04	COMAV	SLL	L162
BGV007854	Mex	Santa Elena. Telchac Puerto. Yucatán	21,29	-89,26	COMAV	SLL	L163

Anexo II. Descriptores empleados en el análisis morfológico.

Nombre	Descripción
Vegetativos	
Longitud de la primera a la décima hoja	Longitud (cm) desde la primera a la décima hoja.
Número de hojas entre segunda y tercera inflorescencia	Número de hojas entre inflorescencias.
Número de hojas entre tercera y cuarta inflorescencia	
Altura hasta la primera inflorescencia	Altura hasta la primera inflorescencia (cm).
Altura hasta última inflorescencia	Altura hasta la última inflorescencia (cm).
Número total de inflorescencias	Inflorescencias totales.
Anchura del tallo	Media de los valores entre la segunda y tercera inflorescencia y entre la tercera y la cuarta. Medido en mm.
Longitud de la inflorescencia	Media entre la longitud (mm) de la segunda y tercera inflorescencia.
Número de flores	Media del número de flores de la segunda y tercera inflorescencia.
Número de foliolos primarios	Foliolillos primarios de una hoja completamente desarrollada.
Número de foliolos secundarios	Foliolillos secundarios de una hoja completamente desarrollada.
Longitud de la hoja	Distancia desde la zona basal hasta el ápice (cm).
Diseción de la hoja	0: baja, 1: intermedia, 2: alta
Borde de la hoja	0: entero, 1: ondulado, 2: aserrado, 3: fuertemente aserrado
Posición de la hoja	Descripción del modo en el que la hoja cae de forma natural. 1: semierecto, 3: semihorizontal, 5: horizontal, 7: horizontal caído, 9: caído
Contenido de antocianinas	Medido en el tallo de plantas adultas. 0: ausente, 1: morado claro, 2: morado.
Tipo de hoja	0: pimpinellifolium, 1: standard, 2: doblemente dividida
Hábito de crecimiento	0: determinado, 1: semideterminado, 2: indeterminado
Color de la hoja	3: verde claro, 5: verde medio, 7: verde oscuro
Inflorescencia cola de pescado	0: ausente, 1: presente
Inflorescencia bifurcada	0: ausente, 1: presente
Inflorescencia irregular	0: ausente, 1: presente
Inflorescencia unípara	0: ausente, 1: presente
Inflorescencia con hojas	0: ausente, 1: hojas, 2: hojas y brotes
Flores	
Número de pétalos	Número de pétalos en flores completamente abiertas.
Número de sépalos	Número de sépalos en flores completamente abiertas.
Longitud pétalo	Media de la longitud (cm) de todos los pétalos de una flor representativa de la planta. La longitud se mide desde el centro de la flor al final del pétalo.
Anchura pétalo	Media de la anchura (cm) de todos los pétalos de una flor representativa de la planta. La longitud se mide desde el centro de la flor al final del pétalo.
Longitud sépalo	Media de la longitud (cm) de todos los sépalos de una

	flor representativa de la planta. La longitud se mide desde el centro de la flor al final del pétalo.
Anchura sépalos	Media de la anchura (cm) de todos los sépalos de una flor representativa de la planta. La longitud se mide desde el centro de la flor al final del pétalo.
Exerción estigmática	Posición del estilo respecto de los estambres. 1: inserto, 2: mismo nivel, 3: ligeramente exerto, 4: altamente exerto.
Curvatura del pétalo	Posición del pétalo en flores completamente abiertas. 3: baja, 5: media, 7: alta.
Curvatura estilo	1: ausente, 2: baja, 3: fuerte.
Color estambres	3: amarillo, 5: naranja pálido, 7: naranja
Color del pétalo	3: amarillo pálido, 5: amarillo, 7: amarillo anaranjado
Fruto	
Color LAB	El color del fruto se midió en un punto cercano a la cicatriz pistilar y en un máximo de 4 frutos por accesión. La escala de colores empleada es LAB donde L representa la luminosidad, a el balance entre rojo y verde, y b el balance entre azul y amarillo.
Peso	El peso se midió en un máximo de 8 frutos por accesión.
Número de lóculos	
Perímetro	Longitud del contorno del fruto. Medido en la sección longitudinal.
Índice externo de forma de fruto	Relación entre altura y anchura máximas.
Forma proximal	Relación de la anchura en la parte superior del fruto con la anchura del fruto a la mitad de su altura.
Forma distal	Relación de la anchura en la parte inferior del fruto con la anchura del fruto a la mitad de su altura.
Altura del hombro	Relación entre la altura del hombro respecto de la altura máxima.
Ángulo proximal micro	Ángulo entre las líneas dibujadas a través del perímetro en la zona proximal del fruto. Los puntos para el ajuste comprenden el 1% del perímetro a cada lado del punto central.
Ángulo proximal macro	Ángulo entre las líneas dibujadas a través del perímetro en la zona proximal del fruto. Los puntos para el ajuste comprenden el 5% del perímetro a cada lado del punto central.
Ángulo distal micro	Ángulo entre las líneas dibujadas a través del perímetro en la zona distal del fruto. Los puntos para el ajuste comprenden el 1% del perímetro a cada lado del punto central.
Ángulo distal macro	Ángulo entre las líneas dibujadas a través del perímetro en la zona distal del fruto. Los puntos para el ajuste comprenden el 5% del perímetro a cada lado del punto central.
Área de indentación	Área de indentación, respecto al total del fruto y multiplicada por 10.
Protusión distal	Relación entre el área de la protusión en la zona distal y

	el área total del fruto.
Asimetría	Distancia promedio entre una línea vertical a la mitad de la anchura del fruto y el punto medio del ancho del fruto a cada altura.
Obovoide	Si el área es mayor por debajo de la altura media que por encima, este parámetro se calcula teniendo en cuenta la anchura máxima (W), la altura a la máxima anchura (y), el ancho promedio por encima de esa altura (w1) y el ancho promedio por debajo de esa altura (w2). Si no se cumple la condición, el parámetro es 0.
Asimetría ovoide	Si el área es mayor a la altura media del fruto que debajo de ella, se calcula como la distancia promedio entre una línea horizontal en el fruto y el punto medio de la altura de la fruta en cada anchura. Si no se cumple la condición, el parámetro es 0.
Posición máxima anchura	Relación de la altura donde el fruto alcanza su máxima anchura, con respecto a la altura total del fruto.
Hombro verde	0: uniforme, 3: verde claro, 5: verde medio, 7: verde oscuro
Cicatriz pistilar	1: punto, 3: línea, 5: estrella, 7: irregular
Fasciación del fruto	0: ausente, 3: baja, 5: intermedia, 7: severa
Color del fruto inmaduro	3: verde blanquecino, 5: verde claro, 7: verde, 9: verde oscuro
Color de la piel	0: sin color, 1: amarillo
Anchura de la cicatriz del pedicelo	3: estrecha, 5: intermedia, 7: ancha
Acostillado	0: ausente, 3: muy débil, 5: débil, 7: intermedio, 9: fuerte
Contenido vascular	Haz vascular presente en frutos verdes maduros. 0: ausente, 3: bajo, 5: intermedio, 7: severo
Líneas longitudinales	Presentes en frutos inmaduros. 0: ausente, 1: presente.

Anexo III. Índices de correlación entre las medidas tomadas por el programa *Tomato Analyzer*.

Medidas básicas	Perímetro	Área	Altura máxima	Anchura máxima	Altura hasta la mitad de la anchura	Anchura hasta la mitad de la altura	Altura curva
Perímetro	1	0,98	0,95	0,99	0,86	0,97	0,95
Área		1	0,93	0,98	0,84	0,97	0,93
Altura máxima			1	1	0,96	0,91	0,97
Anchura máxima				1	0,81	0,91	0,91
Altura hasta la mitad de la anchura					1	0,96	0,91
Anchura hasta la mitad de la altura						1	0,90
Altura curva							1

Índices de forma del fruto	Índice externo I	Índice externo II	Índice externo de forma curva
Índice externo I	1	0,98	0,84
Índice externo II		1	0,83
Índice externo de forma curva			1

Forma	Forma proximal	Forma distal	Forma triangular
Forma proximal	1	0,11	0,75
Forma distal		1	0,66
Forma triangular			1

Forma del extremo proximal	Altura del hombro	Ángulo proximal macro	Ángulo proximal micro	Área de indentación proximal
Altura del hombro	1	0,42	0,14	0,85
Ángulo proximal I		1	0,10	0,27
Ángulo proximal II			1	0,1
Área de indentación proximal				1

Forma del extremo distal	Ángulo distal macro	Ángulo distal micro	Área de indentación	Protrusión
Ángulo distal I	1	0,05	0,11	0
Ángulo distal II		1	0,24	0
Área de indentación			1	0,55
Protrusión				1

Asimetría	Obovoide	Ovoide	Asimetría	Asimetría obovoide	Asimetría ovoide	Posición anchura máxima
Obovoide	1	0,85	0,27	0,65	0,4	0,48
Ovoide		1	0,13	0,54	0,62	0,62
Asimetría			1	0,7	0,34	0,1
Asimetría obovoide				1	0,25	0,2
Asimetría ovoide					1	0,46
Posición anchura máxima						1

Anexo IV. Caracteres significativamente diferentes entre las especies *S. pimpinellifolium*, *S. lycopersicum* var. *cerasiforme* y *S. lycopersicum* var. *lycopersicum*. El p valor procede del análisis MANOVA o test Chi cuadrado y Fisher.

Vegetativos	
Carácter	p valor ANOVA
Anchura del tallo	<0,01
Longitud de la hoja	<0,01
Número de foliolillos	<0,01

Altura hasta la 1ª inflorescencia	0,03
Carácter	p valor FISHER
Número de inflorescencias	0,02
Anchura de la hoja	0,03
Número de foliolos	<0,01
Color de la hoja	<0,01
Tipo de hoja	<0,01
Diseción de la hoja	<0,01
Borde de la hoja	<0,01
Contenido antocianinas del tallo	<0,01
Flores e inflorescencias	
Carácter	p valor ANOVA
Longitud de los sépalos	<0,01
Número de flores por inflorescencia	<0,01
Carácter	p valor FISHER
Número de pétalos	<0,01
Número de sépalos	<0,01
Inflorescencias con hojas	<0,01
Inflorescencia espina de pescado	<0,01
Inflorescencia bifurcada	<0,01
Inflorescencia unípara	0,01
Exerción estigmática	<0,01
Curvatura del estilo	<0,01
Curvatura del pétalo	<0,01
Color de los pétalos	<0,01
Color de los estambres	<0,01
Fruto	
Carácter	p valor ANOVA
Perímetro	<0,01
Peso	<0,01
Asimetría	<0,01
Ángulo proximal micro	<0,01
Ángulo distal micro	0,02
Área indentación	0,03
Carácter	p valor FISHER
Contenido vascular	<0,01
Líneas longitudinales	<0,01
Hombro verde	<0,01
Color piel	<0,01
Ángulo proximal macro	<0,01
Ángulo distal macro	<0,01
Forma distal	<0,01
Altura del hombro	<0,01
Número de lóculos	<0,01

Anexo V. Análisis de componentes principales realizado con todas las especies.

	CP1	CP2	CP3
tipo de hoja estándar	-0,22	0,19	-0,03
tipo de hoja pimpinellifolium	0,23	-0,16	0,04
tipo de hoja doblemente dividida	-0,08	-0,01	-0,06

hoja poca disección	-0,15	0,06	-0,15
hoja mucha disección	-0,15	-0,03	-0,12
hoja sin disección	0,19	0,00	0,15
borde hoja aserrado	-0,17	-0,07	-0,05
borde hoja ondulado	0,02	0,20	-0,04
borde hoja entero	0,20	-0,16	0,11
borde hoja fuertemente aserrado	0,00	-0,02	-0,02
estilo inserto	-0,03	0,21	-0,10
estilo exerto	0,05	-0,10	0,11
estilo mismo nivel	-0,10	-0,04	0,08
estilo altamente exerto	0,14	-0,15	-0,14
sin curvatura del estilo	-0,13	0,00	0,10
baja curvatura del estilo	0,12	0,08	-0,12
alta curvatura del estilo	0,08	-0,09	0,01
sin antocianina	-0,03	-0,06	0,00
antocianina alta	0,10	0,06	-0,02
antocianina intermedia	-0,05	-0,01	0,01
hoja verde medio	-0,04	-0,05	0,04
hoja verde oscuro	0,01	0,08	-0,07
hoja verde claro	0,03	-0,05	0,08
inflorescencia con hojas y brotes	-0,14	0,09	-0,09
inflorescencia sin hojas ni brotes	0,13	-0,12	0,02
inflorescencia con hojas	0,01	0,03	0,17
curvatura media pétalo	-0,10	0,08	-0,15
curvatura baja pétalo	-0,12	-0,03	-0,02
curvatura alta pétalo	0,16	0,03	0,22
pétalo amarillo	-0,02	0,06	0,10
pétalo amarillo pálido	0,00	-0,06	-0,04
pétalo amarillo-anaranjado	0,06	-0,04	0,02
estambre amarillo	-0,12	0,04	0,10
estambre naranja pálido	0,08	0,04	-0,02
estambre naranja	0,12	-0,13	-0,09
sin hombro verde	0,07	0,00	-0,04
hombro verde medio	-0,05	-0,04	-0,06
hombro verde claro	-0,03	0,01	0,22
hombro verde oscuro	-0,06	0,09	-0,06
contenido vascular intermedio	0,01	0,19	0,06
contenido vascular severo	0,03	0,16	0,02
contenido vascular bajo	-0,01	-0,12	-0,03
sin contenido vascular	-0,07	-0,19	-0,06
inflorescencia espina de pescado	0,15	-0,12	-0,06
inflorescencia bifurcada	-0,10	-0,06	0,06
inflorescencia unípara	-0,03	0,17	0,14
color piel	0,11	0,16	0,00
líneas longitudinales	0,05	-0,04	-0,07
número de pétalos	-0,16	-0,15	0,08
número de sépalos	-0,17	-0,14	0,08
número foliolos	-0,14	0,05	-0,12
número de inflorescencias	0,08	0,12	0,22
número de lóculos	-0,19	-0,19	0,02
ángulo distal macro	-0,15	-0,15	0,19

ángulo proximal macro	-0,14	-0,06	0,27
altura hombro	-0,11	-0,08	0,25
forma distal	-0,13	-0,04	0,17
baja anchura hoja	0,16	-0,11	0,11
anchura hoja media	-0,04	0,20	0,01
alta anchura hoja	-0,11	-0,14	-0,12
peso pequeño	0,19	-0,13	-0,11
peso medio	-0,01	0,30	0,14
peso grande	-0,17	-0,24	-0,06
poca altura primera inflorescencia	0,04	0,07	0,30
media altura primera inflorescencia	-0,03	-0,05	-0,30
gran altura primera inflorescencia	-0,02	-0,05	0,04
longitud sépalo	-0,10	-0,07	0,00
perímetro	-0,20	-0,19	0,00
ángulo proximal micro	-0,07	-0,09	0,17
ángulo distal micro	-0,06	-0,05	0,02
indentation	-0,13	-0,17	0,05
asimetría	-0,13	0,08	0,17
número de flores por inflorescencia	0,07	-0,05	0,02
anchura tallo	-0,16	0,12	-0,09
número de foliolillos	-0,12	0,07	-0,02
longitud hoja	-0,18	0,14	-0,03

Anexo VI. Caracteres diferentes significativamente entre SP Ecuador y SP Peru. El p valor procede del análisis MANOVA o test Chi cuadrado y Fisher.

Vegetativos	
Carácter	p valor MANOVA
Número de foliolillos	0,03
Contenido antocianinas tallo	0,03
Carácter	p valor FISHER
Color hoja	0,02
Borde hoja	<0,01
Inflorescencias y flores	
Carácter	p valor MANOVA
Longitud anteras	0,01
Número de flores por inflorescencia	0,01
Longitud inflorescencia	0,04
Carácter	p valor FISHER
Inflorescencias con hojas	<0,01
Color pétalo	<0,01
Exerción estigmática	0,01
Inflorescencia fishbone	0,04
Fruto	
Carácter	p valor MANOVA
Peso	0,03
Carácter	p valor FISHER
Hombro verde	<0,01
Color piel	0,03
Contenido vascular	0,03

Anexo VII. Análisis de componentes principales en la especie *S. pimpinellifolium*.

	CP1	CP2	CP3	CP4
sin antocianina	0,13	0,24	-0,05	-0,08
antocianina alta	-0,15	-0,28	0,07	-0,25
antocianina intermedia	0,08	-0,03	-0,22	0,56
borde hoja aserrado	0	0	0	0
borde hoja ondulado	0,18	-0,1	-0,13	-0,31
borde hoja entero	-0,07	-0,07	0,02	0,18
borde hoja fuertemente aserrado	0	0	0	0
estilo inserto	0,02	-0,02	0,17	0,01
estilo exerto	-0,23	0,47	-0,13	0,21
estilo mismo nivel	-0,05	-0,03	0,04	-0,08
estilo altamente exerto	0,26	-0,42	-0,09	-0,14
	-0,03	0,01	-0,2	0,32
hoja verde oscuro	-0,04	-0,13	0,24	-0,12
hoja verde claro	0,07	0,12	-0,04	-0,19
inflorescencia con hojas y brotes	0,25	0,11	-0,3	-0,05
inflorescencia sin hojas ni brotes	-0,25	-0,2	0,31	0,22
inflorescencia con hojas	0,02	0,11	-0,03	-0,04
pétalo amarillo	0,27	-0,01	-0,17	0,08
pétalo amarillo pálido	-0,31	-0,06	-0,03	-0,02
pétalo amarillo-anaranjado	0	0,17	0,09	-0,04
sin hombro verde	-0,31	0,01	-0,19	-0,19
hombro verde medio	0,31	-0,01	0,19	0,19
hombro verde claro	0,06	-0,02	0,03	0,13
hombro verde oscuro	0	0	0	0
	-0,07	-0,22	-0,02	0,11
peso pequeño	0,1	-0,06	-0,35	-0,1
peso medio	-0,1	0,06	0,35	0,1
hoja corta	-0,35	-0,13	-0,31	0
hoja media	0,28	0,24	0,34	-0,03
hoja larga	0,07	-0,1	-0,03	0,03
longitud anteras	0,16	-0,41	-0,01	0,3
número de flores por inflorescencia	0,21	0,05	-0,02	-0,01
número de foliolillos	-0,05	0,09	-0,14	-0,01

Anexo VIII. Caracteres significativamente diferentes entre SLC Ecuador, SLC Perú, SLC Mesoamérica y SLC México. El p valor procede del análisis MANOVA o test Chi cuadrado y Fisher.

Vegetativos	
Carácter	p valor FISHER
Anchura de la hoja	0,04
Número de foliolos	<0,01
Color de la hoja	<0,01
Tipo de hoja	<0,01

Diseción de la hoja	<0,01
Borde de la hoja	<0,01
Flores e inflorescencias	
Carácter	p valor MANOVA
Anchura del pétalo	<0,01
Longitud del pétalo	<0,01
Anchura del sépalo	<0,01
Longitud sépalos	<0,01
Número de flores	<0,01
Longitud de las anteras	<0,01
Carácter	p valor FISHER
Número de pétalos	<0,01
Número de sépalos	0,02
Inflorescencias con hojas	0,02
Exerción etigmática	<0,01
Curvatura de los pétalos	<0,01
Fruto	
Carácter	p valor MANOVA
Perímetro	<0,01
Ángulo proximal micro	<0,01
Asimetría ovoboide	<0,01
Asimetría	<0,01
Área de indentación	<0,01
Posición máxima anchura	<0,01
Forma proximal	<0,01
Peso	<0,01
Carácter	p valor FISHER
Ángulo distal macro	<0,01
Ángulo proximal macro	<0,01
Número de lóculos	<0,01
Altura del hombro	<0,01
Color L	<0,01
Hombro verde	<0,01
Líneas longitudinales	<0,01

Anexo IX. Análisis de componentes principales en la especie *S. lycopersicum* var. *cerasiforme*.

	CP1	CP2	CP3	CP4
color hoja verde medio	-0,01	0,00	0,03	-0,06
color hoja verde oscuro	0,03	0,01	-0,03	0,10
color hoja verde claro	0,00	0,00	0,00	0,00
tipo hoja estándar	0,03	0,12	-0,04	-0,03
tipo hoja <i>pimpinellifolium</i>	-0,03	-0,12	0,04	0,03
tipo hoja doblemente dividida	0,02	0,09	-0,02	-0,05
baja disección hoja	0,01	0,30	-0,05	-0,02
media disección hoja	0,15	0,32	-0,08	-0,03
sin disección hoja	-0,13	-0,38	0,11	0,12
borde hoja aserrado	0,19	0,18	-0,20	0,10
borde hoja ondulado	-0,20	-0,06	0,16	-0,01
borde hoja entero	0,01	-0,15	0,02	-0,11
borde hoja fuertemente	0,00	0,01	0,01	0,00

aserrado				
inflorescencia con hojas y brotes	0,03	0,09	0,05	-0,01
inflorescencia sin hojas ni brotes	-0,04	-0,05	-0,06	-0,02
inflorescencia con hojas	0,03	-0,04	0,01	-0,03
curvatura media pétalo	0,00	0,08	-0,02	0,26
curvatura baja pétalo	0,11	0,12	-0,08	0,11
curvatura alta pétalo	-0,07	-0,25	0,05	-0,34
estilo inserto	-0,22	0,04	0,14	0,17
estilo exerto	0,08	-0,14	-0,19	-0,04
estilo mismo nivel	0,12	0,11	0,06	-0,14
estilo altamente exerto	0,01	0,01	-0,01	0,00
sin hombro verde	-0,02	0,09	-0,37	-0,28
hombro verde medio	0,00	-0,01	0,22	0,40
hombro verde claro	0,04	-0,13	0,07	-0,19
hombro verde oscuro	-0,02	0,03	0,14	0,10
líneas longitudinales	0,03	0,01	0,12	0,18
número de pétalos	0,23	-0,09	0,09	-0,04
Color L	-0,05	0,06	-0,19	0,06
número de folíolos	0,11	0,26	-0,26	-0,05
pocos lóculos	-0,29	-0,15	-0,19	-0,04
número lóculos intermedio	0,25	0,12	0,21	0,02
alto número lóculos	0,04	0,03	-0,03	0,02
ángulo distal macro	0,29	-0,04	0,23	-0,06
ángulo proximal macro	0,31	-0,15	0,13	-0,23
altura hombro	0,18	-0,11	0,06	-0,12
peso pequeño	-0,08	-0,02	-0,04	0,11
peso medio	-0,10	0,00	-0,04	-0,25
peso grande	0,18	0,02	0,07	0,14
longitud anteras	0,23	-0,19	-0,06	0,15
longitud pétalos	0,18	-0,26	-0,26	0,16
anchura pétalos	-0,02	-0,20	-0,20	0,18
longitud sépalos	0,19	-0,03	-0,19	0,07
anchura sépalos	0,20	-0,28	-0,27	0,16
perímetro	0,27	0,06	0,15	0,13
ángulo proximal micro	0,07	-0,03	0,05	-0,09
forma proximal	0,20	-0,11	0,08	-0,23
área indentación	0,16	-0,04	0,02	0,01
asimetría	0,02	-0,01	0,09	-0,04
asimetría ovoide	0,00	-0,01	0,02	-0,02
número de flores por inflorescencia	0,08	-0,14	-0,27	0,10
posición máxima anchura	-0,06	0,01	0,07	0,03

ANEXO B: ANEXO CAPÍTULO 3.

En la siguiente tabla se recoge la dirección donde se puede consultar la información complementaria relativa al capítulo 3.

Material suplementario capítulo 3	Dirección de enlace
Figure S1	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM1_ESM.png
Figure S2	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM5_ESM.jpg
Figure S3	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM2_ESM.png
Figure S4	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM3_ESM.png
Table S1	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM4_ESM.xlsx
Table S2	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM6_ESM.pdf
Table S3	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM8_ESM.xlsx
Table S4	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM9_ESM.xlsx
Table S5	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM10_ESM.xlsx
Table S6	https://static-content.springer.com/esm/art%3A10.1038%2Fs41438-020-0291-7/MediaObjects/41438_2020_291_MOESM6_ESM.pdf

ANEXO C: ANEXO CAPÍTULO 4 .

Todo el material suplementario relativo al capítulo 4 se encuentra disponible en la dirección <https://www.mdpi.com/2073-4425/12/2/243>.