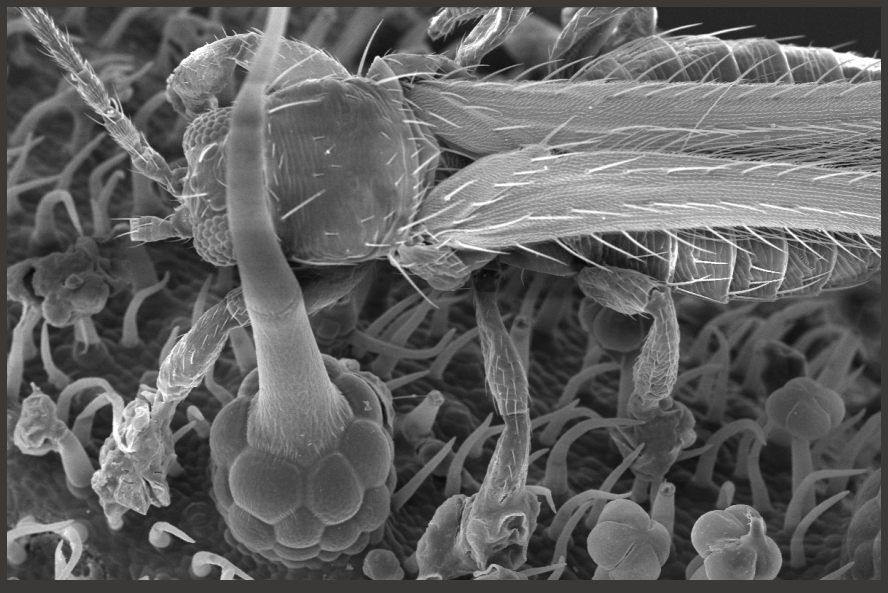
UNIVERSIDAD DE ALMERÍA

Facultad de Ciencias Experimentales

Departamento de Biología y Geología

Área de Genética

La mutagénesis como herramienta de genómica funcional en tomate: caracterización de los mutantes *succulent stamens2* y *hairplus*



Tesis Doctoral

Rocío Fonseca Rodríguez

Almería, 2020

UNIVERSIDAD DE ALMERÍA

FACULTAD DE CIENCIAS EXPERIMENTALES

DEPARTAMENTO DE BIOLOGÍA Y GEOLOGÍA

La mutagénesis como herramienta de genómica funcional en tomate: caracterización de los mutantes *succulent stamens2* y *hairplus*

Rocío Fonseca Rodríguez

Almería, 2020

TESIS DOCTORAL

La mutagénesis como herramienta de genómica funcional en tomate: caracterización de los mutantes *succulent stamens2* y *hairplus*

Mutagenesis as a tool for tomato functional genomics: characterisation of the *succulent stamens2* and *hairplus* mutants

Trabajo realizado en el Área de Genética, Departamento de Biología y Geología, Facultad de Ciencias Experimentales de la Universidad de Almería, por Rocío Fonseca Rodríguez, Licenciada en Biología por la universidad de Granada, para optar al Grado de Doctor por la Universidad de Almería.

Fdo. Rocío Fonseca Rodríguez

Almería, 23 de abril de 2020

UNIVERSIDAD DE ALMERÍA
DPTO. DE BIOLOGÍA Y GEOLOGÍA

El Dr. Juan Capel Salinas, Profesor Titular de Genética de la Universidad de Almería y la Dra. Carmen Capel Salinas Investigadora Postdoctoral Contratada, ambos adscritos al Departamento de Biología y Geología y al Centro de Investigación en Agrosistemas Intensivos Mediterráneos y Biotecnología Agroalimentaria de dicha institución académica,

HACEN CONSTAR

Que el presente trabajo ha sido realizado bajo nuestra dirección y recoge fielmente la labor de Rocío Fonseca Rodríguez, Licenciada en Biología, para optar al Grado de Doctor por la Universidad de Almería. Las investigaciones reflejadas en esta memoria se han llevado a cabo en el Área de Genética, Departamento de Biología y Geología, Facultad de Ciencias Experimentales de la Universidad de Almería.

Y para que así conste, expiden y firman el presente en Almería, a 23 de abril de 2020.

Fdo. Juan Capel Salinas

Fdo. Carmen Capel Salinas

Ctra.Sacramento, s/n. La Cañada de San Urbano 04120 Almería (España) Telf.: 950 015476 FAX: 950 015476 www.ual.es

Publicaciones

Los resultados y conclusiones descritas en la presente memoria han sido objeto de las siguientes publicaciones en revistas científicas incluidas en el *Journal of CItation Reports* (*JCR*):

Pérez-Martín, F., Yuste-Lisbona, F. J., Pineda, B., Angarita-Díaz, M. P., García-Sogo, B., Antón, T., Sánchez, S., Giménez, E., Atarés, A., Fernández-Lozano, A., Ortíz-Atienza, A., García-Alcázar, M., Castañeda, L., Fonseca, R., Capel, C., Goergen, G., Sánchez, J., Quispe, J. L., Capel, J., Angosto, T., Moreno, V. y Lozano, R. (2017). A collection of enhancer trap insertional mutants for functional genomics in tomato. *Plant Biotechnology Journal*, *15*(11), 1439–1452.

Financiación

Este trabajo ha sido financiado por los siguientes proyectos de investigación:

- Genómica funcional y mejora genética de la productividad de tomate: importancia agronómica del balance desarrollo-estrés abiótico (DESTRÉS), (AGL2015-64991-C3-1-R), financiado por el Ministerio de Economía y Competitividad.

- Análisis genómico del desarrollo de tricomas como estrategia para la mejora genética de la resistencia a plagas en tomate (AGL2017-88702-C2-1-R), financiado por el Ministerio de Economía y Competitividad

Asimismo, las actividades de I+D descritas en la presente memoria han recibido el apoyo de los fondos UE-FEDER y del Campus de Excelencia Internacional Agroalimentario (CeiA3).

Resumen

A lo largo de las pasadas décadas se han realizado numerosos avances en el estudio genético de especies hortícolas principalmente para tratar dar respuesta a los retos que enfrenta la agricultura en la actualidad. Entre estos desafíos destacan la presión que ejerce el cambio climático sobre los modelos de cultivo establecidos, así como el incremento de la población mundial. Con este fin, el desarrollo de nuevas variedades comerciales capaces de hacer frente a un entorno cambiante se revela como la alternativa más eficaz. Sin embargo, la variabilidad natural en especies cultivadas parece no ser suficiente para superar todos esos retos. Es por ello que la ONU, además de declara la década 2011-2020 “Década de la ONU sobre Diversidad" ha encargado a la FAO que se implique en impulsar programas de creación de diversidad mediante distintas estrategias de mutagénesis. Solo así podremos continuar desarrollando variedades más productivas a la par que resistentes a estrés bióticos y abióticos.

Entre los cultivos hortícolas de mayor relevancia a nivel mundial destaca el tomate (*Solanum lycopersicum* L.), cuya demanda aumenta continuamente debido a sus propiedades nutricionales y a sus numerosas formas de consumo. Esta especie es además un excelente modelo génico para comprender el desarrollo de plantas de fruto carnoso, ya que presenta enormes ventajas como una elevada tasa de productividad, un ciclo de vida relativamente corto y caracteres agronómicos muy favorables como una elevada tolerancia a estreses abióticos. Todo ello propició que en 2012 un consorcio internacional completara la secuenciación de su genoma, lo que ha permitido realizar enormes progresos en la comprensión de la organización genómica de esta especie. Sin embargo, del total de genes identificados, sólo el 70% aproximadamente tiene una función génica asignada, basada en homologías de secuencias y predicciones *in silico*. De ahí que los esfuerzos actuales se centren en completar la asignación de nuevas funciones génicas mediante estrategias de genómica funcional.

Una de las estrategias de genómica reversa más ampliamente utilizada para la identificación de nuevos genes y su caracterización funcional es el estudio de mutantes, que constituyen un valioso recurso genético y una enorme fuente de variabilidad. La mutagénesis ha sido ampliamente utilizada en tomate y otras hortícolas y el uso de agentes mutagénicos químicos en particular, constituye una de las estrategias más

comúnmente utilizada. Entre las principales ventajas de estas sustancias destaca que son fáciles de manipular y relativamente inocuas. Entre estos agentes se encuentran las sustancias alquilantes como el etil metil sulfonato (EMS), que ha demostrado ser altamente efectivo en la inducción de mutaciones en plantas.

Habida cuenta de la enorme utilidad del EMS, en el presente trabajo se ha obtenido y caracterizado una colección de mutantes empleando este agente mutagénico en el cultivar Moneymaker de tomate. Se han caracterizado alrededor de 8988 plantas mutantes pertenecientes a esta colección, entre las que se han observado alteraciones fenotípicas relevantes de la morfología de la hoja, el tamaño y hábito de crecimiento de la planta, así como alteraciones de la morfología y color de flores y frutos. La combinación de esta estrategia de mutagénesis con técnicas de secuenciación masiva nos ha permitido caracterizar nuevas funciones génicas de enorme interés para comprender en profundidad el desarrollo vegetativo y reproductivo de tomate.

Entre los mutantes identificados como parte de esta colección se encuentra el mutante *succulent stamens2* (*sus2*), que se caracteriza por una morfología floral alteradas, en lo referente a los pétalos y en especial los estambres que están mal formado. Como consecuencia, las plantas mutantes no producen polen y sus frutos son partenocárpicos (sin semillas) y de menor tamaño. Además, durante el desarrollo y la maduración del fruto los estambres se convierten en carpelos carnosos que permanecen adheridos al fruto de las plantas mutantes. El análisis de la morfología de los estambres de plantas mutantes mediante microscopía electrónica de barrido reveló que estos experimentan una conversión homeótica completa en carpelos, así como alteraciones homeóticas parciales en la morfología de las células epidérmicas de los pétalos. La identificación de la localización cromosómica de la mutación *sus2* se realizó mediante mapeo genético utilizando marcadores codominantes en combinación con una estrategia de mapeo por secuenciación. Ello nos permitió identificar dicha mutación en un exón del factor de transcripción *TOMATO MADS-BOX 6* (*TM6*), un gen de clase B implicado en el control de la identidad de los órganos florales del segundo y tercer verticilo. Como resultado de la pérdida de función de TM6, la expresión de otros genes MADS box implicados en el desarrollo floral está alterada en el mutante *sus2*, lo que demuestra que *TM6* es crucial

Resumen

para el mantenimiento de la función de clase B, así como para el correcto funcionamiento de otras proteínas implicadas en la morfogénesis floral, función previamente hipotetizada por algunos autores, pero puesta en tela de juicio por otros.

Por otra parte, durante la caracterización de la colección se identificó también un mutante al que hemos denominado *hairplus* (*hap*), ya que muestra una elevada densidad de tricomas. Este fenotipo es particularmente evidente en los tallos vegetativo y de las inflorescencias, y resulta de enorme interés ya que se ha demostrado que existe una correlación positiva entre el incremento de la densidad de tricomas y la resistencia a plagas de herbívoros. Además, la mutación *hap* incrementa específicamente la densidad de los tricomas glandulares de tipo I, como se deduce del análisis realizado en plantas de fenotipo silvestre y plantas mutantes. La localización cromosómica de la mutación se realizó mediante una estrategia combinada de cartografía genética y mapeo por secuenciación. De esta forma, se localizó la mutación en una Histona N-Lisina Metil Transferasa, una proteína relacionada con el control epigenético de la expresión génica mediante la metilación de histonas. De este modo, la regulación de la densidad de tricomas constituye una nueva función génica identificada para este grupo de proteínas. Por otra parte, el análisis de la densidad de tricomas en líneas de pérdida (RNAi y CRISPR) y de ganancia de función (35S) confirmaron que la mutación identificada es la responsable del fenotipo mutante *hap.* Finalmente, como resultado de la mutación se producen numerosas modificaciones del transcriptoma y del epigenoma de plantas mutantes, que incluyen la expresión diferencial de 92 genes, así como 2113 citosinas diferencialmente metiladas. Todo ello demuestra que *HAP* es un regulador clave de la densidad de tricomas de tomate mediada por la regulación epigenética de la expresión génica, lo cual supone la primera demostración de este tipo de control en un carácter del mayor interés agronómico.

Abstract

Over the past decades, huge advance has been made in the genetic study of crop species in order to overcome the numerous challenges faced by agriculture nowadays. Some of these challenges include modifications of traditional agricultural methods posed by climate change, as well as the increase in world population. With this aim, the development of new commercial varieties capable of coping with a changing environment has revealed as the most effective strategy. However, natural variability among cultivated species is low and as a result, genetic improvement rises as the main approach to develop more productive as well as resistant varieties.

Tomato (*Solanum lycopersicum* L.) is one of the major horticultural crops worldwide, whose demand continues to increase due to its nutritional properties as well as numerous forms of consumption. Tomato is also an excellent model species to understand the development of fleshy fruit plants, since it holds numerous favourable traits such as a high productivity rate, a relatively short life cycle and a high tolerance to abiotic stresses. All these features led in 2012 to the completion of tomato genome sequencing by an international consortium, which has allowed to make an enormous progress in understanding the genomic organization of this species. However, from the total number of identified genes, only 70% approximately have an assigned gene function, based on sequence homologies and *in silico* predictions. Hence, current efforts are focused on the identification of new gene functions by means of functional genomics strategies.

One of the most widely used reverse genetics strategies for the identification of new genes and their functional characterisation is the study of mutants, since they represent a valuable genetic resource as well as an enormous source of variability. Mutagenesis has been extensively used in tomato and other horticultural crops, being chemical mutagenesis one of the most commonly used strategies. Chemical mutagenic agents pose great advantages, since they are easy to handle and relatively harmless. Among them, alkylating substances such as ethyl methanesulfonate (EMS) have been widely employed, since they have proved to be highly effective in inducing plant mutations.

Given the enormous usefulness of EMS, we have used this mutagenic agent in order to obtain a mutant collection in the tomato Moneymaker cultivar. A total number of 8988

Abstract

EMS mutant plants have been characterised, among which we have identified relevant phenotypic alterations related to leaf morphology, size and growth habit, as well as alterations of the morphology and colour of flowers and fruits. The use of this mutagenesis strategy combined with next generation sequencing techniques has allowed us to characterise new gene functions of great interest to deeply understand the vegetative and reproductive development of tomato.

Among the identified mutants is *succulent stamens2* (*sus2*), which exhibits an aberrant stamen morphology. As a result, mutant plants are unable to produce pollen and mutant fruits are parthenocarpic (seedless) and smaller. Furthermore, stamens remain attached to fruits during development and ripening as fleshy structures, whose morphology resembles that of carpels. The analysis of mutant plants stamen morphology was carried out by scanning electron microscopy, which showed that stamens undergo a complete homeotic change into carpels, as well as alterations produced in petals epidermal cells morphology. Mapping strategies allowed us to identify that *sus2* mutation is located in an exon of the *TOMATO MADS BOX 6* (*TM6*) transcription factor, a B class gene involved in the development of the second and third floral whorls. As a result of TM6 loss of function, the expression levels of other MADS box genes involved in flower development are altered in *sus2* mutant plants, revealing a crucial role for *TM6* in the maintenance of B class function as well as in the correct balance of other proteins involved in floral morphogenesis.

On the other hand, a mutant which we have called *hairplus* (*hap*) since it shows a higher trichome density was also identified. This phenotype is particularly obvious in inflorescence stems as well as in vegetative stems of mutant plants, and it is of great interest given that a positive correlation between the increase in trichomes density and pest resistance has been widely described. Furthermore, the *hap* mutation specifically increases type I glandular trichomes density, as deduced from the analysis performed in wild type and mutant plants tissues. Genetic mapping allowed us to identify the *hap* mutation in a Histone N-Lysine Methyl Transferase, a group of proteins related with the epigenetic control of gene expression by means of histone methylation. Therefore, the regulation of trichome density can be considered as a new function of these proteins. On the other hand, the analysis of trichome density in loss of function lines (RNAi and

Abstract

CRISPR) as well as in gain of function lines (35S) confirmed that the identified mutation is the responsible for the *hap* mutant phenotype. Finally, as a result of the mutation, large number of modifications are accounted for the transcriptome and the epigenome of mutant plants, including the differential expression of 92 genes, as well as 2.113 differentially methylated cytosines. Taken together, all these results prove that *HAP* is a key regulator of tomato trichome density by means of epigenetic regulation of gene expression.

Índice

Índice

Apéndice de figuras, tablas y anexos.

Table 4. Sequence of primers used for qRT-PCR and SolGenomics (SGN) gene id

Table 5. Primers used for expression analysis and functional characterisation of *HAP*

ANEXO I. Supplementary figures.

Supplementary Fig. 1. Pollen viability assays performed with tetrazolium chloride.

Supplementary Fig. 2. Fine mapping performed by whole genome sequencing.

Supplementary Fig. 3. Relative Expression of floral identity MADS box genes assessed by qRT-PCR.

ANEXO II. Supplementary figures and tables.

Supplementary Fig. 1. Trichome and stomata density in inflorescence stems, stems and leaves of WT and *hap* mutant plants.

Supplementary Fig. 2. Mapping by sequencing identification of *HAP*.

Supplementary Fig. 3. Expression analysis of *HAP* in RNAi silencing lines.

Supplementary Fig. 4. Allele characterisation in a CRISPR-Cas9 knock-down line.

Supplementary Fig. 5. Spatial expression pattern of *HAP.*

Supplementary Fig. 6. Phylogenetic relationships between *Arabidopsis* and tomato SET domain containing proteins.

Supplementary Fig. 7. Multiple alignment of HAP and other Arabidopsis and tomato SET domain containing proteins.

Supplementary Fig. 8. Distribution in the 12 tomato chromosomes of the DMC caused by the loss of function of *HAP*.

Supplementary Fig. 9. Location of DMCs caused by the loss of function of *HAP*.

Supplementary Fig. 10. Damage caused by the tomato fruit borer *Helicoverpa armigera* (Hübner) after 15 days feeding in WT and *hairplus* mutant plants.

Supplementary Table 1. Trichome density accounts in leaves, stem and inflorescence stem of wild type and *hap* mutant plants.

Supplementary Table 2. Trichome density accounts in leaves, stem and inflorescence stem of wild type, *hap* mutant plants and transgenic lines.

Supplementary Table 3. Summary of up regulated genes detected in a RNAseq transcriptomic analyses. RPKM values are listed for each of three replicates of wild type and hap mutant plants.

Supplementary Table 4. Summary of down regulated genes detected in a RNAseq transcriptomic analyses. RPKM values are listed for each of three replicates of wild type and hap mutant plants.

Introducción

1. El tomate

1.1. Origen y domesticación

El tomate cultivado (*Solanum lycopersicum* L.) tiene su origen en la región andina que comprende desde el sur de Colombia hasta el norte de Chile e incluye parte de Ecuador, Perú y Bolivia (Díez y Nuez, 2008). Se considera que su ancestro más probable es la especie silvestre *S. lycopersicum* var *cerasiforme* (Dun.), que crece de forma silvestre en las zonas tropicales y subtropicales de América (Rick, 1978). Su centro de diversificación es un tópico que genera controversia ya que algunas hipótesis basadas en evidencias botánicas y lingüísticas sostienen que se produjo en Perú (de Candolle, 1882; Luckwill, 1943). Otra teoría sostiene que ocurrió en México (Jenkins, 1948) ya que no existen evidencias del cultivo de tomate en Sudamérica en la época precolombina. Sin embargo, el hecho de que las accesiones sudamericanas y mexicanas de tomate compartan marcadores moleculares sugieren que esta hipótesis es errónea (Rick y Fobes, 1975). Estudios recientes integran ambas hipótesis, concluyendo que el proceso de domesticación de esta especie ocurrió en dos etapas: una primera pre-domesticación en América del Sur y posteriormente una etapa de domesticación en América Central (Blanca et al., 2015). Recientes estudios de genética de poblaciones basados en datos de secuencias genómicas han sugerido que la evolución de tomate podría ser más compleja de lo inicialmente hipotetizado, y *S. lycopersicum* var *cerasiform* sería una especie aparecida de forma natural, no resultado de la domesticación, de la cual habrían llegado a Mesoamérica dos poblaciones a partir de las que se habría domesticado la especie cultivada (Razifard et al. 2020). Desde la región de América Central el tomate fue introducido en Europa alrededor del siglo XVI, donde su uso inicial fue estrictamente ornamental por considerarla tóxica, dada su similitud con otras solanáceas como la belladona (*Atropa belladona* L.) (Rick, 1978). Fue a mediados del siglo XVIII cuando su consumo se popularizó en Europa (Nuez, 1995), desde donde se extendió al resto del mundo hasta alcanzar una distribución global.

Las variedades silvestres de tomate exhiben una gran diversidad genética que les confiere una enorme capacidad de adaptación a las presiones selectivas impuestas por

el ambiente. Sin embargo, como resultado del proceso de domesticación, la especie cultivada ha sufrido un proceso de cuello de botella que seleccionó positivamente caracteres como el tamaño del fruto y la productividad (Tanksley, 2004) en detrimento de caracteres relacionados con la resistencia a plagas y patógenos, así como la resistencia a estrés abiótico (Bai y Lindhout, 2007). De hecho, se ha llegado a estimar que el tomate cultivado contiene menos del 5% de la variación genética encontrada en las especies silvestres (Miller y Tanksley, 1990). Recuperar esa diversidad genética consiste aún hoy un reto para la mejora genética de esta especie, que ha recurrido a numerosas estrategias como la introgresión de caracteres de interés identificados en especies silvestres o la obtención de variedades híbridas (Bauchet y Causse, 2012; Blanca et al., 2015).

1.2. Taxonomía

La clasificación taxonómica del género *Solanum* y del tomate en particular ha variado a lo largo del tiempo. Una de las primeras clasificaciones se atribuye a Miller (1731), que lo incluyó en el género *Lycopersicum*. Posteriormente, en 1753 el naturalista Carolus Linnaeus lo incluyó en el género *Solanum* como parte de su obra *Species Plantarum*. Sin embargo, la clasificación de Miller predominó atendiendo a caracteres fenotípicos como la morfología de la hoja y, de hecho, era la clasificación empleada a mediados del siglo XX (Luckwill, 1943). Finalmente, atendiendo a estudios moleculares el tomate se ha incluido en el género *Solanum* que a su vez incluye el género *Lycopersicum* como una sección con cierta unidad taxonómica (Peralta et al., 2005). Según la base de datos de información taxonómica ITIS (https://www.itis.gov/), la jerarquía taxonómica actual del tomate es la siguiente:

Reino:	Plantae
Subreino:	Viridiplantae
Infrarreino:	Streptophyta
Superdivisión:	Embryophyta
División:	Tracheophyta
Subdivisión:	Spermatophyta

Clase:	Magnoliopsida
Superorden:	Asteranae
Orden:	Solanales
Familia:	Solanaceae
Género:	*Solanum*
Especie:	*Solanum lycopersicum*

1.3. Descripción botánica

El tomate es una planta de porte arbustivo y perenne que se cultiva como anual. Puede desarrollarse de forma rastrera, aunque también es común el porte semierecto o erecto. El tallo principal tiene un grosor de 2 a 4 cm y a partir de él se desarrollan las hojas y en su axila los tallos secundarios. Las hojas son de tipo compuestas e imparipinnadas, con un foliolo terminal de mayor tamaño y hasta 8 foliolos laterales todos peciolados y de borde aserrado. Tanto las hojas como el tallo principal y secundarios están recubiertos de tricomas de tipo glandular y no glandular. El sistema radicular está compuesto por una raíz pivotante profunda y raíces adventicias secundarias.

El crecimiento de la planta es de tipo simpodial o indeterminado, es decir, el tallo está dividido en segmentos denominados simpodios. El primer simpodio está compuesto por entre 6 a 12 hojas, momento en el cual se produce la transición floral y el meristemo apical se determina y adquiere identidad reproductiva, lo que da lugar al desarrollo de la primera inflorescencia. El crecimiento de la planta continúa gracias al mantenimiento de la identidad meristemática del meristemo simpodial (MS), que se desarrolla en la axila de la última hoja. De esta forma, se producen continuamente segmentos simpodiales, en lo adelante compuestos de 2 a 3 hojas y una inflorescencia.

No obstante, algunas variedades de tomate muestran crecimiento determinado y después del desarrollo del primer simpodio el meristemo apical se determina y origina un meristemo inflorescente (MI). Este patrón de crecimiento es resultado de mutaciones en el gen *SELF-PRUNING* (*SP*), ortólogo de los genes *CENTRORADIALIS* y *TERMINAL FLOWER1* de *Antirrhinum majus* L. y *Arabidopsis thaliana* L. respectivamente y da como resultado un crecimiento más compacto que facilita la recolección de los

frutos de forma mecánica (Pnueli et al., 1998). De ahí la enorme utilidad de esta mutación en el procesamiento de las variedades de tomate destinadas a industria.

Las inflorescencias son racimos que contienen un número variable de flores, generalmente de 5 a 12. Las flores son actinomorfas, hermafroditas y constan de 4 verticilos. El primero es el cáliz, formado por 5-6 sépalos, le sigue la corola con entre 5-6 pétalos de color amarillo lanceolados y al igual que los sépalos fusionados en la base. A continuación, se encuentran los estambres, normalmente 5 y fusionados formando un cono estaminal que envuelve al cuarto verticilo o gineceo, compuesto a su vez por un ovario bi o pluriloculado y un estilo que termina en el estigma. El estilo, a diferencia de algunas especies silvestres emparentadas con tomate que muestran exerción estigmática, es de menor tamaño que los estambres, lo que hace de tomate una especie autógama, si bien la polinización cruzada es igual de exitosa (Lozano et al., 2009).

El fruto es una baya carnosa y se compone del pericarpo, el tejido locular, la placenta y las semillas. Su forma es globular u ovoide de manera general, aunque existen una gran variedad de formas que dependen del cultivar y las condiciones de cultivo (Chamarro, 1995).

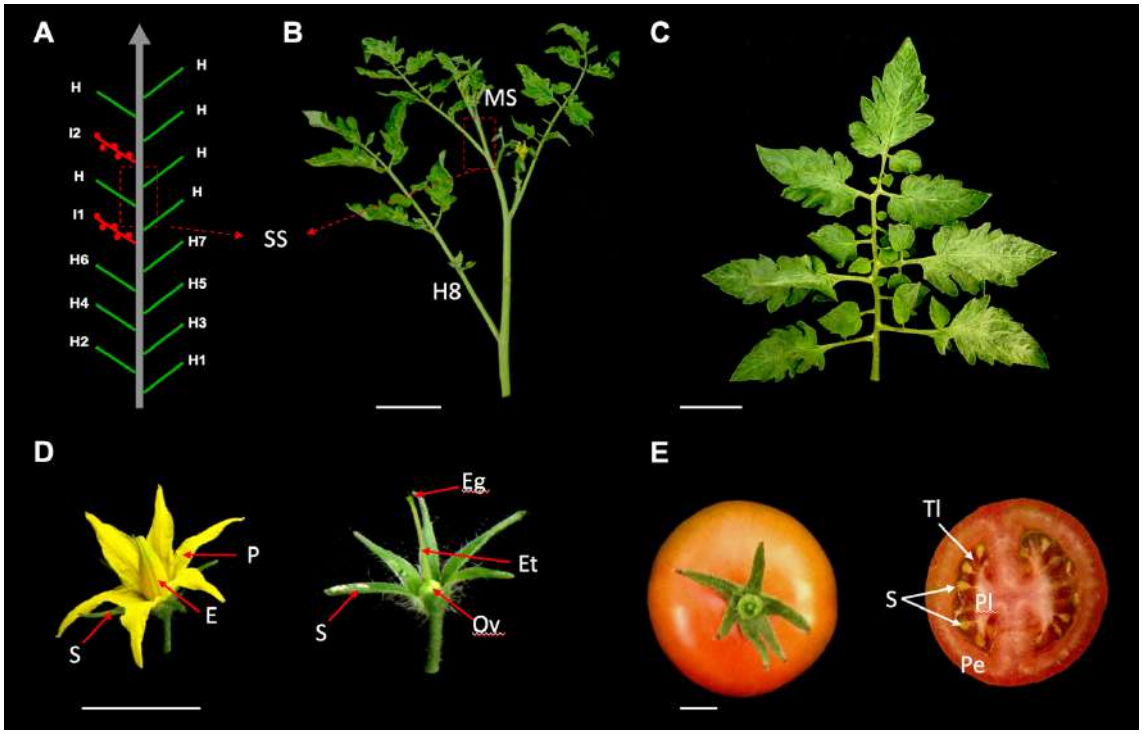


Figura 1: Aspectos del desarrollo vegetativo y reproductivo de tomate *(S. lycopersicum* L.). A. Representación esquemática del desarrollo simpodial de una planta de tomate. **B.** Se muestran las hojas (H), inflorescencias (I), la estructura de un segmento simpodial y la ubicación del meristemo simpodial (MS). **C.** Morfología de la hoja. **D**. Verticilos florales: S: sépalos; P: pétalos; E: estambres; Ov: ovario; Et: estilo y Eg: Estigma. **E.** Fruto de tomate. Izquierda: vista cenital. Derecha: sección longitudinal donde se aprecian el pericarpo (P), la placenta (Pl), el tejido locular (Tl) y las semillas (S). Escala: A, B y C = 5 cm. D y E = 1 cm.

1.4. Importancia económica y nutricional

Con alrededor de 3000 especies, la familia de las solanáceas es la tercera en importancia económica a nivel mundial. Algunas de sus especies, como el propio tomate, patata (*S. tuberosum* L.), pimiento (*Capsicum annuum* L.) y berenjena (*S. melongena* L.) concentran la mayor parte del volumen de producción de hortalizas de todo el planeta. Entre todas ellas, destaca el cultivo de tomate, cuya producción mundial se ha incrementado exponencialmente llegando a un máximo de 182 millones de toneladas en 2017 según datos de la FAO (http://www.fao.org/faostat/es/#data/QC). En cuanto a la producción europea, España con un total de 5.163.466 millones de toneladas en 2017 ocupa el segundo puesto entre los países productores, precedido únicamente por Italia (**Tabla 1**).

Tabla 1: Principales países productores de tomate para consumo en fresco en el año 2017 según datos de la FAO.

Países	Producción (tn)	Superficie (ha)	Rendimiento (hg/ha)
Italia	6.015.868	99.750	603.095
España	5.163.466	60.852	848.529
Rusia	3.230.718	114.300	282.653
Ucrania	2.267.460	74.400	304.766
Portugal	1.747.634	20.873	837.270
Países Bajos	910.000	1.790	5.083.799
Polonia	898.012	11.442	784.838
Grecia	879.000	13.300	660.902
Rumania	726.643	46.815	155.216
Francia	656.408	3.504	1.873.352

Introducción

El consumo de tomate está en constante ascenso ya sea para consumir en fresco, en salsas, sopas u otro tipo de conservas. Ello ha motivado un intenso proceso de mejora genética encaminada a desarrollar variedades de mayor rendimiento. Prueba de ello es que, en España la tendencia desde el año 2000 muestra un incremento de la producción y una disminución de la superficie cultivada, dando como resultado un mayor rendimiento (**Figura 2**).

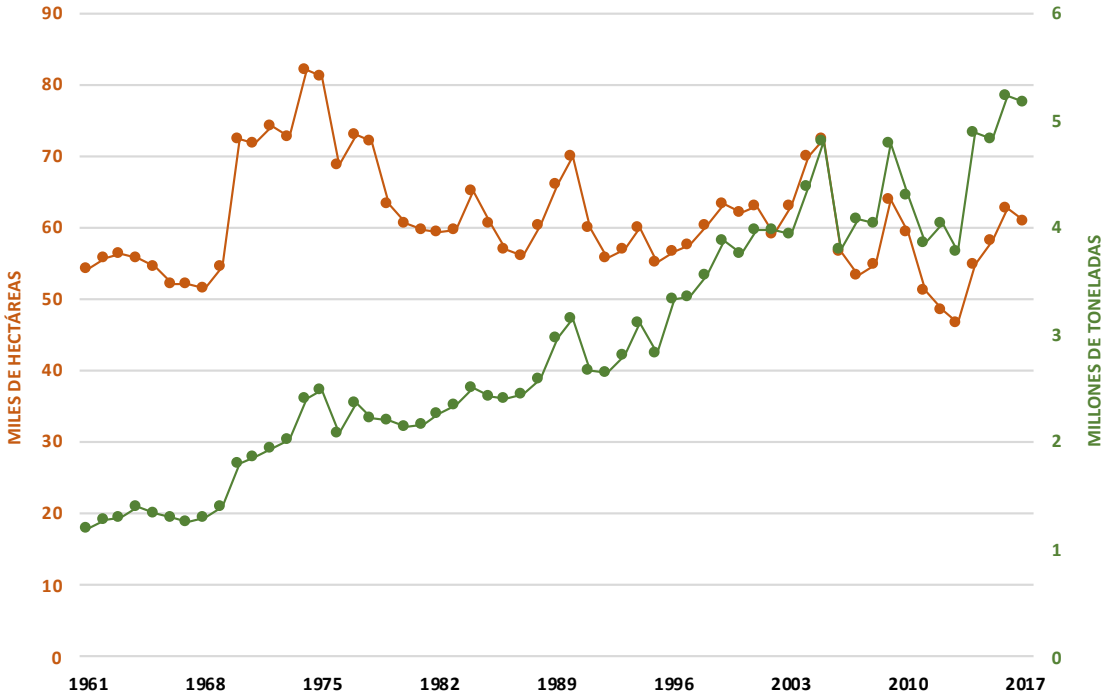


Figura 2: Evolución de la superficie cosechada y la producción de tomate para consumo en fresco en España entre los años 2000-2017.

Gran parte del éxito del cultivo de tomate radica en la gran diversidad de variedades que se han obtenido con diferentes formas, color, sabor, textura o tamaño, que dependen del uso y del mercado al que están destinadas (**Figura 3**). El valor de la producción bruta de esta hortícola en el mercado español se ha cifrado en 1453.163.763 millones de dólares en 2016. Otro valor añadido de su consumo es su elevado valor nutricional. Sus frutos constituyen una fuente de antioxidantes como la vitamina C, así como de β-carotenos precursores de la síntesis de vitamina A, fundamentalmente

licopeno (Moco et al., 2006), que es el responsable del color rojo de los frutos maduros de tomate. Algunos estudios consideran que el licopeno está directamente relacionado con una disminución de la incidencia de enfermedades cardiovasculares (Arab y Steck, 2000), de ahí la importancia de la mejora genética dirigida a incrementar el contenido de este compuesto. De hecho, gran parte de los cultivares modernos son portadores en el cromosoma 6 de la mutación *old-gold-crimson* (*ogc*) que incrementa en un 25% el contenido de licopeno de los frutos (Park et al., 2009).

Figura 3: Variedades cultivadas de tomate en las que se observa una gran diversidad de formas y tamaños. La barra de escala se corresponde con 3 cm.

1.5. Tomate como especie modelo

Entre los vegetales, la especie modelo por excelencia ha sido siempre *Arabidopsis thaliana*, ya que su porte herbáceo y su ciclo de crecimiento permiten obtener numerosas generaciones de estudio en muy corto tiempo. Sin embargo, es complicado establecer paralelismos entre los estudios realizados en esta especie y las hortícolas de fruto carnoso. Para ello, el tomate constituye un excelente modelo de estudio, ya que se trata de una especie autógama y con un ciclo de vida relativamente corto (Rick, 1978). Además, tiene un enorme potencial de adaptación a diferentes condiciones ambientales, lo que facilita la mejora de la resistencia a condiciones extremas, es

insensible al fotoperiodo y presenta aptitud para la transformación genética (Foolad, 2007; Lozano et al., 2009). Se dispone también de un gran número de recursos fitogenéticos, que incluyen alrededor de 62.800 accesiones conservadas en bancos de germoplasma, así como colecciones de mutantes monogénicos y accesiones silvestres entre las que destaca la del *Tomato Genomics Research Centre* (https://tgrc.ucdavis.edu/), que contiene la mayor colección de accesiones de tomate silvestre del mundo.

Otra de las ventajas de tomate es que su genoma es diploide y de tamaño relativamente pequeño. Además, se encuentra secuenciado desde el año 2012 (The Tomato Genome Consortium, 2012) y su secuencia se puede consultar en numerosas bases de datos entre las que destaca Sol Genomics Network (https://solgenomics.net/), dedicada exclusivamente a genomas de solanáceas. Esta base de datos pone a disposición de los investigadores un gran número de recursos bioinformáticos, como el diseño de marcadores moleculares (CAPS - *Cleaved Amplified Polymorphic Sequences*; dCAPS – *derived* CAPS; SNP – *Single Nucleotide Polymorphism*; SSR – *Single Sequence Repeats* o AFLP – *Amplified Fragment Lenght Polymorphism*), alineamiento de secuencias mediante BLAST, predicción de VIGS (*Viral Induced Gene Silencing*), así como numerosas herramientas de cartografía genética.

1.6. Genómica estructural y funcional de tomate

La secuenciación del genoma de tomate se llevó a cabo por un consorcio internacional que reunió a más 300 investigadores de 14 países y que permitió en 2012 obtener la secuencia del cultivar Heinz 1706 así como de la especie silvestre *S. pimpinellifolium* (The Tomato Genome Consortium, 2012). El ensamblaje de las secuencias por el *International Tomato Annotation Group* (ITAG) determinó que el genoma de tomate tiene un tamaño aproximado de 950 Mb que se organizan en 12 cromosomas, cada uno formado por una región heterocromática pericentromérica y dos regiones eucromáticas en las porciones distales. Las regiones eucromáticas representan alrededor del 25% del genoma, pero concentran el 90% de los genes descritos (Wang et al., 2006). El número de genes predichos es de 34.727 genes y de ellos, 30.855 han sido respaldados mediante datos

de secuenciación de ARN mientras que 31.741 comparten homología con genes de *A. thaliana*. Además, el genoma de tomate es altamente sinténico con el de otras solanáceas como la patata y se han identificado 18.320 genes ortólogos entre ambas especies (The Tomato Genome Consortium, 2012).

Toda la información derivada de la secuenciación del genoma de tomate se une a numerosos estudios previos de cartografía y mapeo genético que han contribuido también a la anotación del genoma de esta especie (Causse y Grandillo, 2016). No obstante, del total de genes identificados, solamente alrededor del 70% tiene asignada una función génica, de los que prácticamente la mayoría son asignaciones en base a homología de las secuencias. De ahí que cobren especial interés las estrategias de genómica funcional que persiguen caracterizar la función de cada uno de los genes predichos en el genoma.

1.7. Mutagénesis y genómica funcional

La identificación de nuevos genes y su caracterización funcional tiene un enorme interés no sólo científico sino también aplicado a programas de mejora de caracteres de interés agronómico (Parry et al., 2009). Una de las estrategias más exitosas empleadas a tal fin es el estudio de mutantes, que ha demostrado ser de enorme utilidad en la caracterización de las rutas y genes que controlan procesos clave como el cuajado de los frutos y su desarrollo (Gorguet et al., 2005). El término mutación engloba tanto cambios puntuales producidos en un único gen por sustitución o deleción de un nucleótido, como aquellos que afectan a grandes segmentos cromosómicos o cromosomas completos (Cubero, 2003). Las mutaciones constituyen la fuente primaria de variabilidad en las poblaciones, ya que generan nuevas variantes alélicas con diferentes capacidades de adaptación al ambiente.

Las mutaciones pueden ser de origen espontáneo o inducidas. Las primeras ocurren de forma natural en las poblaciones generalmente debido a errores en la actividad de la ADN polimerasa durante la replicación, mientras que las segundas se producen de forma artificial mediante el empleo de agentes mutagénicos. En el caso de tomate, aunque también podría extrapolarse a otras hortícolas, los alelos mutantes de origen

espontáneo conocidos hasta la fecha representan sólo una pequeña parte del total de genes predichos (Pino-Nunes et al., 2008). Los programas de mutagénesis artificial a gran escala son por tanto imprescindibles para obtener nuevos alelos que no se encuentran de forma natural en el germoplasma.

Los programas de mutagénesis pueden clasificarse según el sitio donde se generan las mutaciones en mutagénesis dirigida o al azar. La mutagénesis dirigida permite obtener nuevos alelos mutantes directamente en genes de interés, previo conocimiento de su secuencia mientras que la mutagénesis al azar origina mutaciones en cualquier punto del genoma. Sin embargo, la clasificación más comúnmente empleada hace referencia al tipo de agente mutagénico empleado y distingue entre biológica, física y química.

1.7.1. Mutágenos físicos

Este grupo se divide fundamentalmente en radiaciones de tipo ionizantes y no ionizantes. Entre las primeras se encuentran los rayos gamma y los rayos X que son los más utilizados (Mba et al., 2012) y sus efectos suelen ser generalmente pequeñas deleciones, aunque en el caso de los rayos gamma también son frecuentes las deleciones de gran tamaño (Yuan et al., 2014). Otro de los agentes físicos empleados son los neutrones rápidos, aunque su uso conlleva la aparición de grandes deleciones, translocaciones e incluso la pérdida de cromosomas enteros, lo que reduce considerablemente la viabilidad de las semillas mutantes (Sikora et al., 2011). A pesar de este inconveniente, los neutrones rápidos han sido utilizados con éxito en Arabidopsis para generar 51.840 líneas mutantes (Li y Zhang, 2002) y en arroz donde se obtuvieron alrededor de 10.000 líneas mutantes junto a otras 20.000 por mutagénesis con rayos X (Wu et al., 2005).

Otro mutágeno físico muy empleado es la luz ultravioleta (UV), que causa daños celulares de origen fotoquímico a las células cuando se exponen durante largos períodos de tiempo a este agente. Los daños causados por la luz UV incluyen modificaciones de las bases pirimidínicas, reducción del crecimiento vegetativo de la planta, así como reducción de la actividad fotosintética y de la tasa de floración (Ries et al., 2000).

1.7.2. Mutágenos biológicos

Entre los primeros mutágenos biológicos empleados destacan los transposones. Entre ellos son de especial interés *Supressor-mutator* (*Spm*) y *Mutator* (*Mu*) del maíz (*Zea mays* L.) y *Tag1* de Arabidopsis, que han permitido clonar y caracterizar más de 60 genes relacionados con el desarrollo (May y Martienssen, 2010). Se trata de elementos móviles que pueden desplazarse de una región del cromosoma a otra con una elevada tasa de transposición. Su uso ha permitido obtener colecciones mutantes en maíz (Brutnell, 2002), Arabidopsis (Kuromori et al., 2004) e incluso en especies ornamentales donde han permitido identificar genes relacionados con la pigmentación en flores de *Petunia hybrida* (van Houwelingen et al., 1998). También destacan los mutantes generados por los 5 transposones activos identificados en *Antirrhinum majus* L., claves para la identificación de genes homeóticos (Li et al., 2019). En tomate no se han descrito elementos transponibles activos, pero si se han empleado transposones exógenos para generar pequeñas colecciones de mutantes (Meissner et al., 2001).

El agente mutagénico de origen biológico más comúnmente empleado es el T-DNA. Se trata de una porción del plásmido Ti (*tumor inducing*) o Ri (*root inducing*) de *Agrobacterium tumefasciens* que se transfiere y se integra en las células vegetales durante la infección con esta bacteria y que constituye un excelente vector de ADN foráneo (Gelvin, 2003). El T-DNA también puede ser transferido al genoma vegetal mediante diferentes metodologías que incluyen desde la electroporación de protoplastos y la biolística entre los métodos físicos, como a través de vectores biológicos (Kikkert et al., 2005). Todo ello ha sido posible gracias a herramientas biotecnológicas que han permitido la manipulación del plásmido Ti de forma tal que este incluya genes reporteros o de resistencia que actúen como marcadores selectivos. Son muy numerosas las colecciones mutantes obtenidas empleando esta metodología en especies modelo como Arabidopsis (Feldmann et al., 1991; Szabados et al., 2002), en gramíneas como arroz (Jeon et al., 2000; Chen et al., 2003) y en hortícolas como la especie cultivada o especies silvestres de tomate (Atarés et al., 2011; Pérez-Martín et al., 2017).

Si bien su uso está muy extendido, cabe destacar que existen algunos inconvenientes relacionados con el uso del T-DNA, principalmente debido al considerable tamaño de

los vectores empleados. Resultados obtenidos del análisis de colecciones mutantes de Arabidopsis demuestran que, con mucha frecuencia la inserción del T-DNA lleva aparejados eventos de translocación y reordenamientos cromosómicos que dificultan la identificación de las secuencias genómicas flanqueantes y por tanto su sitio de inserción (Tax y Vernon, 2000; Jupe et al., 2019).

1.7.3. Mutágenos químicos

Los mutágenos químicos son muy numerosos y de muy diversa naturaleza, por lo que suelen clasificarse según su efecto sobre el ADN. De este modo, existen agentes intercalantes, hidroxilantes, desaminantes, análogos de bases y alquilantes por citar algunos. Generalmente, estas sustancias causan mutaciones puntuales en un único nucleótido, ya sean transiciones o cambios de una base por otra de la misma naturaleza, o transversiones cuando el cambio es entre bases de distinta naturaleza, es decir purinas por pirimidinas o viceversa (Sikora et al., 2011). Su principal ventaja frente a otros agentes mutagénicos como los físicos radica en la posibilidad de regular con mayor precisión la dosis del agente, de forma tal que se reduzca la letalidad embrionaria y se obtengan un mayor número de líneas M_1.

Entre los agentes análogos de bases se encuentran el 5-bromouracilo y la 2-aminopurina, sustancias que durante la replicación se incorporan al ADN como una base más y provocan por tanto transiciones entre bases de la misma naturaleza (Mba, 2013). El ácido nitroso constituye un ejemplo de agente desaminante, cuya acción ocurre preferentemente sobre la guanina seguida de la adenina y la citosina. La reacción de desaminación genera formas aldehído fosfodiéster que se entrelazan con los grupos amino de las bases de la hebra complementaria y provocan transiciones de bases (Zimmerman, 1977). Este compuesto ha demostrado ser altamente mutagénico en bacterias (Kaudewitz, 1959) y hongos como *Neurospora crassa* (Barnett y De Busk, 1960), mientras que en plantas su uso ha estado más limitado ya que en *A. thaliana* induce aberraciones cromosómicas que modifican drásticamente la capacidad de germinación y supervivencia de las semillas (Mabuchi y Arnason, 1969). Se han empleado también agentes intercalantes como el bromuro de etidio, cuyo mecanismo

de acción consiste en insertarse entre las bases nitrogenadas, lo que provoca un estrechamiento de la doble hélice. Dicho estrechamiento es reconocido por la ADN polimerasa como una base adicional de forma tal que inserta una base opuesta con la que emparejar la molécula intercalante, lo que da como resultado cambios en el marco de lectura de los genes (Mba, 2013).

Sin embargo, el mutágeno químico más ampliamente empleado es el agente alquilante etil metano sulfonato o EMS (Van Harten, 1998), ya que su uso es relativamente sencillo y produce una elevada tasa de mutagénesis (Sikora et al., 2011). Esta sustancia induce fundamentalmente transversiones G/C a A/T (Greene et al., 2003). El EMS es un líquido incoloro, poco soluble en agua y que presenta un único grupo alquilo reactivo, lo que hace que su toxicidad sea menor que la de los agentes que contienen dos o varios grupos alquilo reactivos (Kodym y Afza, 2003). Su mecanismo de acción sobre el ADN ha sido descrito por Sega (1984) y puede observarse en la **Figura 4**. Comienza con la formación de triésteres inestables que permiten la liberación de los grupos alquilo que pueden llegar a interferir con la replicación del ADN. En ocasiones estos triésteres sufren reacciones de hidrólisis que resultan en la rotura de la doble cadena de ADN. Una vez libres, los grupos alquilo reaccionan con mayor frecuencia con la guanina, seguida de la adenina y la citosina. La guanina alquilada se ioniza de forma diferente y es capaz de emparejarse con timina en vez de con citosina durante la replicación, lo que introduce una mutación que se hace patente durante la replicación de una de las hebras del ADN. Debido a su mecanismo de acción, las mutaciones provocadas por el EMS son generalmente de naturaleza recesiva, por lo que se requiere cultivar las plantas M_1 para obtener progenies segregantes M_2 donde realizar la caracterización fenotípica.

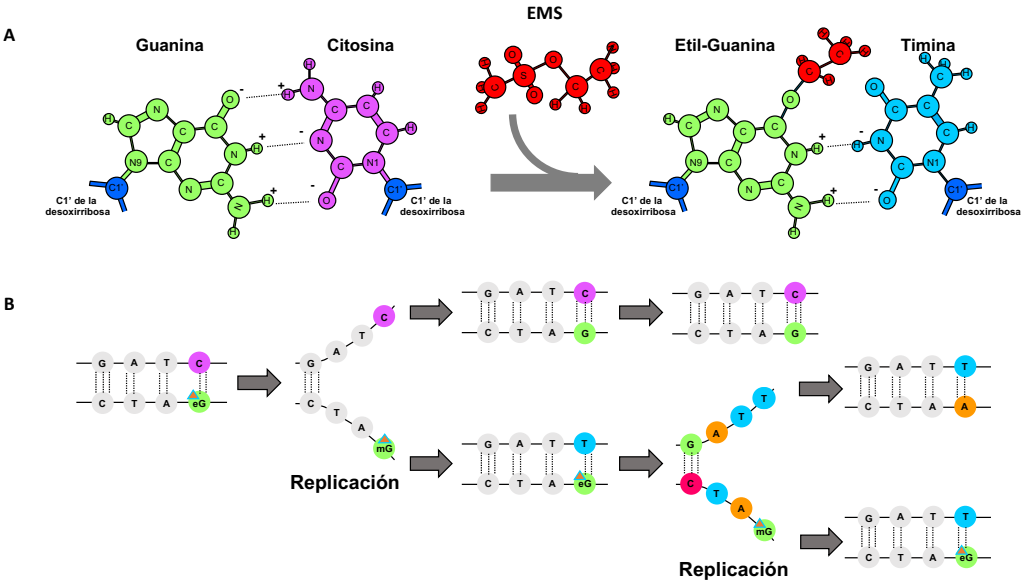


Figura 4: Mecanismo de acción del agente alquilante etil metano sulfonato (EMS) sobre el ADN. A. Alquilación de la guanina por el grupo alquilo reactivo del EMS y formación de etil-guanina. **B.** Después del tratamiento con el mutágeno y durante los siguientes ciclos de replicación del ADN, la etil-guanina se empareja con timina originando el cambio G/C por A/T.

Son muy numerosas las colecciones mutantes que se han obtenido empleando EMS. Algunos ejemplos incluyen colecciones de maíz (Till et al., 2004; Candela y Hake, 2008), arroz (Hirochika et al., 2004; Till et al., 2007) y por supuesto *A. thaliana* (McCallum et al., 2000; Martín et al., 2009). En tomate su uso está también muy extendido, especialmente en el cultivar Micro-Tom. Este cultivar es particularmente interesante porque es de pequeño tamaño (15 cm) lo que permite una elevada densidad de plantación de hasta 1.357 plántulas por m^2 (Meissner et al., 1997). Además, el crecimiento determinado de esta variedad reduce el ciclo de cultivo y permite recoger los frutos maduros en un periodo de entre 70 y 90 días (Emmanuel y Levy, 2002). Saito et al., (2011) llevaron a cabo la caracterización de 8.598 líneas obtenidas en este cultivar, lo que permitió establecer 1.819 categorías fenotípicas entre todos los mutantes obtenidos e implementar una base de datos denominada TOMATOMA, posteriormente actualizada por Shikata et al., (2016) donde se pueden consultar los fenotipos mutantes de esta colección (http://tomatoma.nbrp.jp/indexAction.do).

Conviene destacar que el empleo de EMS como agente mutagénico presenta ventajas frente a otros sistemas como la mutagénesis insercional con T-DNA. La principal radica en la ausencia de vectores binarios complejos que requieren manipulaciones biotecnológicas al diseñar el experimento de mutagénesis. Además, el potencial para la mejora genética de las plantas mutantes obtenidas es muy elevado ya que no se trata de plantas transgénicas. No obstante, uno de los principales inconvenientes radica en la identificación del sitio del genoma donde se producen las mutaciones, lo que requiere generar poblaciones de mapeo polimórficas y recurrir a técnicas de mapeo por secuenciación.

2. Morfogénesis vegetal

El crecimiento de una planta es un proceso continuo estrechamente regulado por la implicación de diferentes rutas génicas y de una estructura en particular: los meristemos (Evans y Barton, 1997). Una vez que se forma el zigoto, se suceden numerosas divisiones celulares que permiten la polarización del embrión estableciendo dos regiones: el meristemo apical del tallo o SAM (*Shoot Apical Meristem*) que da lugar a la parte aérea de la planta y el meristemo apical de la raíz que da lugar a este órgano. Los meristemos se dividen de forma controlada y manteniendo un equilibrio constante entre las células meristemáticas indiferenciadas y las células que van adquiriendo identidad.

2.1. El meristemo apical del tallo

El SAM es el encargado de la organogénesis de toda la parte aérea de las plantas (tallos, hojas y flores), lo que requiere mantener el equilibro entre la población de células que se mantienen indiferenciadas y las poblaciones celulares que adquieren identidad (Ha et al., 2010). Durante la fase vegetativa, el principal producto del SAM son las hojas, mientras que durante la fase reproductiva produce flores derivadas de meristemos florales (MF) que contienen su propia población de células madre.

Desde el punto de vista anatómico, el SAM se divide en diferentes regiones (**Figura 5A**): la túnica, formada a su vez por la protodermis o capa mas externa (L1) y la capa subdérmica (L2) y el corpus (L3). Las células de las capas L1 y L2 experimentan

exclusivamente divisiones de tipo anticlinal, es decir en un plano perpendicular, mientras que las células que componen el corpus (L3) se dividen en todos los planos (Meyerowitz, 1997). De este modo, las células de la capa L1 originarán la epidermis, a partir de la capa L2 se diferenciarán el tejido mesodérmico y las células germinales y la capa L3 generará los tejidos vasculares (Fletcher y Meyerowitz, 2000).

Desde el punto de vista funcional, existen tres dominios principales en el SAM (Fletcher, 2002) (**Figura 5B**). La zona central, que se caracteriza por divisiones celulares muy lentas y donde se localizan las células meristemáticas indiferencias que migran al resto de zonas. La zona periférica o zona de transición, donde la tasa de divisiones celulares es mayor, es la región donde las células que proceden de la zona central adquieren identidad para formar los primordios de los órganos. Finalmente, encontramos la zona medular implicada en la formación de los tejidos conductores e internos del tallo. Todos estos dominios están delimitados por acción del denominado centro organizador, una región de la zona central que se encarga de controlar la tasa de división celular y diferenciación en el resto de las zonas mediante mecanismos de comunicación célula a célula (Adibi et al., 2016).

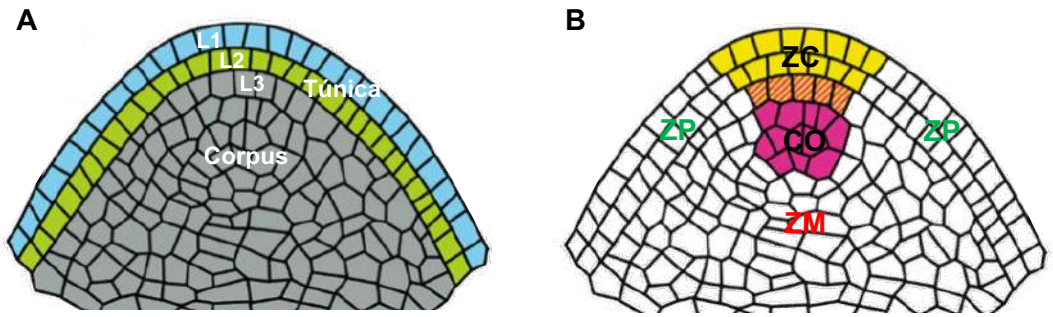


Figura 5: Organización del meristemo apical del tallo (SAM). A. Organización anatómica del SAM. La túnica se compone de la capa L1 da lugar a la epidermis, y de la capa L2 que originará la capa subdérmica. El corpus (L3) da lugar a los tejidos internos del tallo. **B.** Organización funcional del SAM. Se distinguen la zona central (ZC) donde se localizan las células meristemáticas, el centro organizador (CO), así como las zonas periférica (ZP) y media (ZM). Adaptado de Uchida y Torii, 2019.

Entre los principales reguladores de la actividad del SAM se encuentra la ruta compuesta por los genes *WUSCHEL* (*WUS*) y *CLAVATA3* (*CLV3*). *WUS* codifica para un factor de transcripción de la familia *WUSCHEL*-related homeobox (WOX) y fue identificado como parte de la caracterización de mutantes de *A. thaliana* afectados en el desarrollo del SAM. Mutaciones en este gen alteran la identidad de las células madre del meristemo, que se determina prematuramente después de producir un reducido número de órganos (Laux et al., 1996). La expression de *WUS* comienza en el estadio embrionario de 16 células, incluso antes de la formación del SAM. A medida que este se desarrolla, la expresión de *WUS* se restringe a un grupo de células ubicadas por debajo de las células madre, que es donde se requiere su función para el mantenimiento de la indiferenciación. Ello ha llevado a considerar a este gen como el organizador del SAM y el responsable de trasmitir a las células adyacentes una señal que las especifica como células madre pluripotentes (Mayer et al., 1998).

En cuanto a *CLV3*, su función es contraria a la de *WUS*, ya que promueve la progresión de las células meristemáticas hacia la organogénesis (Schoof et al., 2000). Los mutantes *clv3* exhiben un retraso en la formación de primordios de órganos y una acumulación de células meristemáticas indiferenciadas que dan como resultado un incremento de tamaño del SAM (Clark et al., 1995). *CLV3* codifica un péptido señal (Clark et al., 1995) secretado por las células de la zona central, región a través de la cual difunde hasta ser percibido por receptores quinasa en el centro organizador, donde inhibe la expresión de *WUS* (Schoof et al., 2000). La señal de CLV3 es percibida a través de dos rutas independientes. La primera es mediante la unión directa al receptor quinasa CLV1 (Clark et al., 1997) y en la segunda la señal es trasmitida por el dímero CLV2/CRN (CORINE) (Somssich et al., 2016). La ruta *WUS*/*CLV3* se retroalimenta ya que se ha demostrado que WUS migra de las células del centro organizador a las adyacentes, donde se une al promotor de *CLV3* activando su expresión y por lo tanto reprimiendo la suya propia (Yadav et al., 2011).

Más allá de la ruta *WUS*/*CLV3*, existen otros reguladores esenciales de la actividad del meristemo. Es el caso de *SHOOT MERISTEMLESS* (*STM*) un factor de transcripción KNOX1 (*knotted-like homeobox*) que se ha demostrado que se expresa durante el estadio

globular de desarrollo del embrión (Long y Barton, 1998). Los mutantes *stm* presentan un SAM reducido o completamente ausente en los embriones maduros, así como un fusión parcial de los cotiledones (Endrizzi et al., 1996). Ello sugiere su relación con otro gen implicado en el desarrollo del SAM denominado *CUP-SHAPED COTYLEDON* (*CUC*), un factor de transcripción de tipo NAC. Los dobles mutantes de los genes *CUC* muestran un fenotipo de ausencia de meristemo y fusión de los cotiledones. Además, los genes *CUC* regulan la expresión de *STM* (Aida et al., 1999).

2.2. Transición floral

El cambio de la fase vegetativa a la reproductiva esta marcado por la conversión del SAM en un meristemo inflorescente (MI), que dará lugar a los meristemos florales (MF). Este proceso de denomina transición floral y esta regulado por factores ambientales, hormonales y génicos estrechamente vinculados. El principal inductor de la floración en la especie modelo *A. thaliana* es el gen *FLOWERING LOCUS T* (*FT*), un factor de transcripción que contiene un dominio bZIP que se expresa en cotiledones y hojas en condiciones de día largo. FT actúa sobre genes diana del meristemo apical induciendo la morfogenésis floral (Notaguchi et al., 2008). La expresión de *FT* está a su vez condicionada por el gen *CONSTANS* (*CO*), que se une a su promotor y forma complejos con otros factores de transcripción para finalmente activar en el SAM la expresión de los genes MADS *APETALA1* (*AP1*) y *SUPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*) (Khan et al., 2014). *SOC1* a su vez se asocia con *AGAMOUS-LIKE 24* (*AGL24*) e induce la expresión del integrador floral *LEAFY* (*LFY*) (Yu et al., 2004) encargado de establecer la identidad del meristemo floral (Kobayashi et al., 1999). Se han descrito también reguladores negativos de la morfogénesis floral, como los genes *TERMINAL FLOWER 1* (*TFL1*) y *AP2*. *TFL1* presenta gran homología con *FT*, aunque su función es antagónica, ya que previene la formación del meristemo floral mediante la represión de la transcripción de los genes diana de *FT* (Hanano y Goto, 2011). *AP2* tiene una función similar a la de *TFL1*, ya que reprime la expresión de *SOC1* (Yant et al., 2010).

El proceso de morfogénesis floral descrito en Arabidopsis difiere del de tomate en dos puntos esenciales. El primero es que mientras que en Arabidopsis su patrón de

crecimiento monopodial determina que la transición floral sea un evento único, en tomate se produce en cada uno de los simpodios. En segundo lugar, la floración en tomate se inicia independientemente de la duración de los días ya que es una especie insensible al fotoperiodo. El primer meristemo floral se produce después del desarrollo de entre 6 a 12 hojas, momento en que el SAM se determina y a partir de aquí, el proceso se repetirá cada 3 hojas (Molinero-Rosales et al., 2004).

Los principales reguladores de la iniciación de la morfogénesis floral en tomate son los genes *SINGLE FLOWER TRUSS* (*SFT*) y *FALSIFLORA* (*FA*), ortólogos de *FT* y *LFY* respectivamente (Molinero-Rosales et al., 1999). Los mutantes *sft* y *fa* muestran una floración tardía, mientras que en el doble mutante la floración no se produce, lo que indica que ambos genes tienen efecto aditivo y que, en ausencia de *FA*, *SFT* es indispensable para iniciar la floración (Molinero-Rosales et al., 2004). Por otra parte, *SFT* también previene la reversión del meristemo inflorescente a un crecimiento vegetativo, función que comparte con otros genes como *JOINTLESS* (*J*) (Mao et al., 2000) y *MACROCALIX* (*MC*) (Yuste-Lisbona et al., 2016). El gen *COMPOUND INFLORESCENCE* (*S*) promueve también la maduración del meristemo de inflorescencia y su terminación, de ahí que los mutantes de este gen exhiban inflorescencias ramificadas y un elevado número de flores (Quinet et al., 2006). Otro de los reguladores implicados en la determinación de la identidad del meristemos de inflorescencia en tomate es *ANANTHA* (*AN*), que actúa aguas debajo de *FA*. La pérdida de función de este gen origina inflorescencias ramificadas que no producen flores y que generan constantemente meristemos inflorescentes (Lippman et al., 2008).

2.3. El meristemo floral: Modelo ABC(DE)

Como se ha mencionado anteriormente, las flores de tomate están compuestas por verticilos concéntricos de sépalos, pétalos estambres y carpelos. El análisis de mutantes afectados en la morfología y la disposición de estos verticilos ha permitido caracterizar los genes implicados en su identidad y formular un modelo génico que los agrupa denominado modelo ABC (Schwarz-Sommer et al., 1990). Este modelo reúne todos los genes que por su estrecha implicación en el control de la identidad de los órganos

florales se denominan homeóticos. Fue descrito inicialmente a partir de mutantes homeóticos de *A. majus* y *A. thaliana* (Coen y Meyerowitz, 1991) aunque se ha demostrado que el modelo puede ser extrapolado a otras especies como tomate. El modelo ABC establece que las combinaciones de las funciones A, B y C determinan la identidad de cada verticilo. De este modo, la expresión de genes de función A determina la formación de los sépalos, la expresión combinada de los genes de clase A y B controla la formación de los pétalos, la expresión de genes de función B y C regula la formación de los estambres y los genes de clase C controlan la formación de los carpelos. Además, los genes de clase A y C se reprimen mutuamente, por lo que en ausencia de alguno de ellos el otro estaría presente en todos los verticilos de la flor (Weigel y Meyerowitz, 1994). La mayoría de los genes que integran este modelo son factores de transcripción de tipo MADS box que se encuentran ampliamente distribuidos entre las mono y dicotiledóneas y cuya implicación en procesos clave del desarrollo ha sido ampliamente descrita (Castelán-Muñoz et al., 2019).

En Arabidopsis los genes de función A están representados por *APETALA1* (*AP1*) y *AP2*, los de función B por *AP3* y *PISTILLATA* (*PI*) y los de función C por *AGAMOUS* (*AG*). Los mutantes deficientes en cada una de estas funciones génicas han sido ampliamente caracterizados en Arabidopsis y se han añadido dos nuevas funciones génicas que se suman a este modelo, en lo adelante denominado ABC(DE). La función D la confieren los genes *SHATERPROOF* (*SHP1* y *SHP2*) y *SEEDSTICK* (*STK*) que regulan de forma redundante la identidad de los óvulos (Colombo et al., 2010; Mizzotti et al., 2012). La función E la confieren los genes de la familia *SEPALLATA* (*SEP*), que se requieren para la especificación de la identidad de todos los verticilos y para la determinación del meristemo (Pelaz et al., 2000) (**Figura 6)**.

Por otra parte, la función de estos genes homeóticos no se limita estrictamente a la determinación de los verticilos florales. Por ejemplo, *AP1* no sólo determina la identidad de los sépalos, sino que también participa en la determinación del meristemo en estadios tempranos (Bowman et al., 1993). Por otra parte, *AP2* promueve la identidad de los pétalos, a la par que contrarresta la actividad de *AG* mediante un mecanismo

antagonista, basado en que mientras *AG* reprime la expresión del gen de mantenimiento del meristemo *WUS*, *AP2* promueve su expresión (Huang et al., 2017).

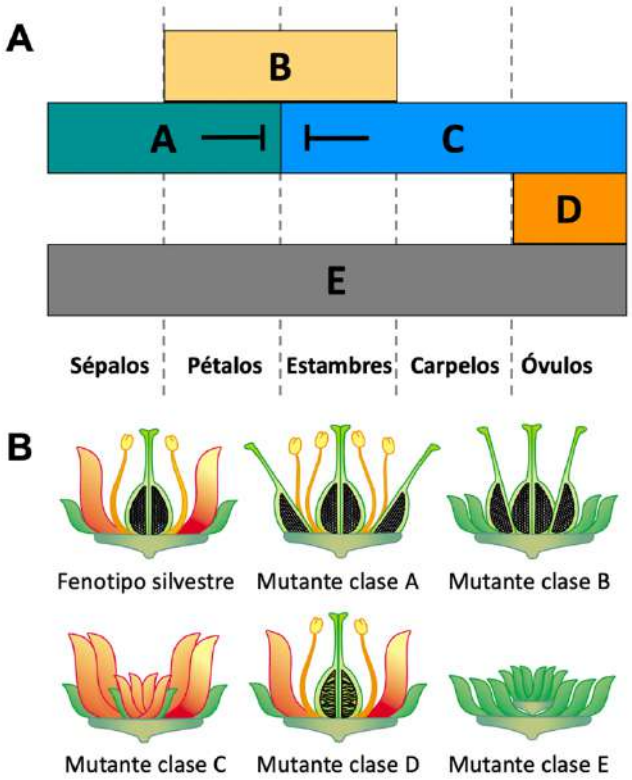


Figura 6: Modelo ABC(DE) de desarrollo floral. A. Representación esquemática de las funciones génicas que componen el modelo ABC(DE) así como de los verticilos controlados por cada una de ellas**. B.** Representación del fenotipo normal o silvestre, así como de las conversiones homeóticas originadas como resultado de mutaciones en cada uno de estos genes. Adaptado de Ferrario et al., 2004.

Como ya se ha comentado, el modelo ABC(DE) es también extrapolable a tomate, donde la función A la asume el gen *MACROCALYX* (*MC*) ortólogo de *AP1* de Arabidopsis. El mutante *mc* presenta inflorescencias de desarrollo aberrante y que revierten a un crecimiento vegetativo después de la formación de unas cuantas flores fértiles además de sépalos de gran tamaño (Vrebalov et al., 2002; Yuste-Lisbona et al., 2016). Entre los homólogos de *AP2* destacan los genes *SlAP2a*, *SlAP2b* y *SlAP2c*. De ellos el más

homólogo es *SlAP2b*, cuya mutación provoca un incremento en la producción de etileno y como consecuencia una maduración temprana (Wang et al., 2019). Sin embargo, no se ha descrito que este gen participe en la determinación de la identidad de los verticilos florales, lo que indica un cierto grado de divergencia funcional dentro del clado al que pertenecen estas proteínas (Chung et al., 2010, Karlova et al., 2011).

Los genes de clase B de tomate también han sido previamente descritos. El gen *STAMENLESS* (*SL*) o *TOMATO APETALLA3* (*TAP3*) se ha revelado esencial para esta función (Quinet et al., 2014). Los mutantes *sl* muestran una severa conversión homeótica de sépalos en el segundo verticilo y estambres convertidos en carpelos en el tercero (Gómez et al., 1999). Entre los homólogos de *AP3* se encuentran *Tomato MADS box gene 6* (*TM6*) (syn. *TDR6*) (Busi et al., 2003) y *Tomato APETALA3* (*TAP3*) (Kramer et al., 1998) y se ha descrito que las mutaciones en *TAP3* y el silenciamiento de *TM6* resultan en la conversión de estambres en carpelos y en conversiones más o menos severas de pétalos en sépalos (de Martino et al., 2006). Los genes *TOMATO PISTILLATA* (*TPI*) también denominados *SlGLOBOSA* (*SlGLO*) son también de clase B y sus mutaciones dan lugar a cambios homeóticos similares a los observados en el mutante *tap3* (de Martino et al., 2006).

En cuanto a la función C, en tomate está representada por los homólogos de *AG*. Los genes *TOMATO AGAMOUS1* (*TAG1*) y TOMATO *AGAMOUS LIKE1* (*TAGL1*) son un claro ejemplo. Ambos muestran conversiones homeóticas en estambres y carpelos e indeterminación en estos últimos (Giménez et al., 2016). La función D en tomate parecen realizarla los genes *TAGL11* y *SlMBP*, homólogos del gen *STK* de Arabidopsis cuya función es establecer el contexto de desarrollo de los óvulos y semillas (Busi et al., 2003). Finalmente, los genes *TOMATO MADS BOX 5* (*TM5*) y *TOMATO MADS BOX 29* (*TM29*) se encuentran descritos como genes de clase E en tomate. *TM5* es homólogo de *SEP3* de Arabidopsis y su expresión se localiza en los tres verticilos internos de la flor (Pnueli et al., 1994). Por su parte, *TM29* (syn. *TAGL2*) por su parte es considerado homólogo de *SEP1* y *SEP2* y su expresión se detecta en los cuatro verticilos (Ampomah-Dwamena et al., 2002).

El estudio de los genes implicados en la morfogénesis floral, así como de sus homólogos en diferentes especies arroja una valiosa información acerca del origen de la flor de las angiospermas. De hecho, estudios comparativos basados en el modelo ABC han demostrado que la mayoría de los genes que lo integran tienen una compleja historia evolutiva caracterizada por duplicaciones frecuentes e incluso la pérdida de algunas familias génicas (Litt y Krammer, 2010). La morfología floral de las angiospermas como la conocemos hoy es por tanto resultado de la interacción entre estos eventos de mutaciones espontáneas y posterior selección y evolución.

3. Tricomas

3.1. Diferenciación celular en la epidermis

La epidermis constituye el tejido que regula la interacción de las plantas con el exterior y como tal modula la respuesta de estos organismos a los cambios ambientales. Como el resto de los tejidos que conforman la parte aérea de las plantas, la epidermis se origina a partir del SAM, concretamente a partir de la protodermis o capa L1.

Dado que la epidermis constituye la principal barrera defensiva de las plantas frente a su entorno, existen un gran número de estructuras diferenciadas a partir de las células epidérmicas que sirven a tal fin. Tal es el caso de la cutícula, que impregna las paredes celulares de la epidermis de prácticamente todas las plantas. La cutícula está compuesta por el poliéster denominado cutina constituido por una mezcla de ácidos grasos y glicerol impregnadas por una capa cerosa, compuesta a su vez por ácidos grasos de cadena muy larga. La cutícula actúa como barrera difusora que evita la pérdida de agua y limita el intercambio de solutos entre las células y el apoplasto (Ingram y Nawrath, 2017).

Sin embargo, sus funciones no se restringen solo a evitar la desecación, ya que muchos mutantes afectados en la permeabilidad de la cutícula presentan a su vez alteraciones en el desarrollo. Tal es el caso del mutante *wax2* de Arabidopsis, en el cual mutaciones en un gen que codifica para una proteína transmembrana de la familia esterol desaturasa dan lugar al desarrollo de una cutícula más delgada y de estructura

desorganizada, además de la fusión de los órganos florales y una marcada esterilidad (Chen et al., 2003).

Otra de las principales modificaciones de la epidermis son los estomas, estructuras cuya principal función es el intercambio gaseoso. Se trata de un poro rodeado de dos células denominadas células guarda u oclusivas que en dicotiledóneas presentan forma arriñonada. Estas estructuras son las encargadas de permitir que el CO_2 llegue a las células del mesófilo y por tanto son críticas para la eficiencia fotosintética y la supervivencia de las plantas.

Se han descrito mutantes alterados en la morfogénesis de los estomas, como el mutante *stomata carpenter1* (*scap1*) o el mutante *too many mouths* (*tmm*). En el primer caso, mutaciones en un factor de transcripción de tipo DOF dan lugar a estomas cuyas células oclusivas son de forma irregular e incapaces de cerrar los estomas como respuesta a la concentración de CO_2 o de abrirlos en respuesta a la luz (Negi et al., 2013). Por otra parte, en el mutante *tmm* se observan alteraciones en el patrón de espaciamiento entre estomas de forma tal que crecen agrupados sin células intermedias que separen unos de otros, como resultado de divisiones progresivas que se producen en diferentes ángulos (Nadeau y Sack, 2002). La información obtenida del estudio de estos mutantes es clave para comprender a un nivel más profundo los factores que regulan la diferenciación de las células epidérmicas.

3.2. Tricomas: función biológica e importancia económica

Las estructuras diferenciadas a partir de la epidermis más conocidas y más extensamente estudiadas son los tricomas, que se originan a partir de las células protodérmicas de la capa L1 del SAM. El término procede del griego *trichos* que significa pelo y que es justamente la forma que adquieren estas estructuras. Los tricomas constituyen una de las primeras estructuras vegetales identificadas al microscopio por los botánicos, para quienes han constituido un importante carácter taxonómico (Behnke, 1984).

No obstante, existen varios tipos de células que por su origen epidérmico y su forma de diferenciación pueden clasificarse como tricomas. Por ejemplo, las papilas son

estructuras poco pronunciadas que apenas sobresalen de la epidermis y que en ocasiones tienen función sensitiva. Tal es el caso de las papilas que se desarrollan en el estigma del gineceo de las flores y que controlan de forma directa el proceso de polinización. Cuando estas papilas detectan que el polen es compatible favorecen su correcta hidratación, mientras que la restringen en presencia de granos de polen incompatibles (Dickinson, 1995). Las escamas o pelos peltados son otro tipo de tricomas pluricelulares cuyas células se desarrollan todas en un único plano formando un escudo y que en ocasiones pueden ser pedunculados. Son frecuentes en especies del género *Tillandsia* y otras Bromeliáceas, así como en la epidermis de las hojas de olivo (*Olea europea* L.), donde evitan la pérdida de agua y favorecen su absorción a nivel foliar además de ofrecer protección frente a la desecación.

Sin embargo, el prototipo más comúnmente asociado a los tricomas es el de estructuras de forma filiforme que crecen perpendiculares a la epidermis. Estudios realizados en la especie modelo *A. thaliana* han demostrado que una vez que adquieren identidad, los incipientes tricomas detienen las divisiones mitóticas e inician ciclos consecutivos de endoreduplicación, es decir, el genoma nuclear se replica en ausencia de divisiones celulares. A continuación, las células aumentan de tamaño y la dirección de su crecimiento pasa a ser perpendicular a la epidermis (Hülskamp, 2004).

La estructura de los tricomas puede sufrir ramificaciones, como en el caso de los tricomas unicelulares de Arabidopsis, puede contener varias células como en el caso de los tricomas multicelulares de tomate e incluso pueden ser de tipo glandular cuando en su punta se diferencian células secretoras. La función biológica de los tricomas depende por tanto de su morfología. Los tricomas no glandulares contribuyen a la disipación del calor en la superficie de la hoja, incrementan la tolerancia al frío, favorecen la absorción de agua, actúan como barrera para los rayos UV y no menos importante, constituyen una barrera mecánica que frena la diseminación de las plagas de herbívoros (Serna y Martin, 2006; Simmons y Gurr, 2005). Entre los tricomas no glandulares más conocidos destacan las fibras de algodón (*Gossypium hirsutum* L.) de tan extensivo uso en la industria textil (Wilkins et al., 2000). Por otra parte, los tricomas glandulares constituyen auténticas biofactorías productoras de una gran diversidad de compuestos químicos

que varían desde aceites esenciales como en el caso de algunas Lamiáceas, sustancias psicotrópicas como en caso del Cannabis o más frecuentemente sustancias que resultan tóxicas para plagas de herbívoros (Glas et al., 2012; Huchelmann et al., 2017).

3.3. Diferenciación y morfogénesis de tricomas en *Arabidopsis thaliana* L.

Los tricomas constituyen un excelente modelo para el estudio de los procesos de diferenciación celular en plantas, ya que son de fácil acceso y visualización bajo condiciones de laboratorio. La iniciación del desarrollo de los tricomas en la epidermis sigue un patrón predeterminado, ya que rara vez se observan tricomas que se desarrollen en células adyacentes o en grupos. Existen evidencias de este patrón incluso en el caso de las fibras de algodón que se desarrollan en gran número a partir del óvulo (Martin y Glover, 2007).

Hasta la fecha, la mayoría de los estudios sobre tricomas se han realizado en la especie modelo *A. thaliana*, que presenta tricomas en toda la superficie aérea incluyendo las hojas de la roseta, las hojas caulinares, sépalos y tallos. Los tricomas de Arabidopsis son de tipo unicelular, si bien en las hojas de la roseta presentan tres ramificaciones mientras que en el resto de los tejidos son menos ramificados o no presentan ramificación alguna (Hülskamp et al., 1999).

Una vez que una célula protodérmica adquiere la identidad de tricoma se detienen las divisiones celulares y la célula inicia alrededor de 4 ciclos de endoreduplicación que multiplican su contenido en ADN hasta 32 C (Schwab et al., 2000; Hülskamp, 2004). Tomando como modelo los tricomas de las hojas de la roseta, el primer ciclo de endoreduplicación se completa antes de que la célula cambie de polaridad y el incipiente tricoma comience a emerger de la epidermis. El segundo ciclo ocurre en paralelo a la primera ramificación, así como el tercero se acompasa con la segunda ramificación. Durante el último ciclo, la célula se expande dando lugar a un tricoma maduro con tres ramificaciones que mide alrededor de 0,5 mm (Schnittger y Hülskamp, 2002). Se ha descrito también el patrón de ramificación que experimentan los tricomas en desarrollo: el primero se produce alineado con el eje distal-proximal de la hoja, el segundo ocurre en la base de la primera ramificación en un ángulo de 90º (Folkers et al., 1997). Cuando

los procesos de ramificación se completan, se inicia una nueva fase de crecimiento que se caracteriza por una rápida vacuolización acompañada de la expansión de las ramificaciones (Schwab et al., 2000).

El análisis genético de mutantes de Arabidopsis afectados en la morfogénesis de los tricomas ha permitido esclarecer las rutas génicas implicadas en los diferentes estadios de diferenciación de estas estructuras. Se han descrito mutaciones que afectan al patrón de diferenciación celular en la epidermis, también mutaciones que regulan el cambio de los ciclos de división mitótica a los ciclos de endoreduplicación y que por consiguiente alteran el contenido de ADN de los tricomas, mutaciones que afectan al patrón de ramificación y mutaciones que afectan a la dirección del crecimiento y que originan tricomas aberrantes (Hülskamp, 2004).

En cuanto al patrón de iniciación, se cree que este se rige por un mecanismo de inhibición celular mutua. Las células destinadas a diferenciarse en tricomas producen un factor promotor del desarrollo de tricomas, a la par que producen factores de inhibición que suprimen el desarrollo de tricomas en las células vecinas. El análisis genético y molecular de mutantes apoya este modelo y se han caracterizado genes que pueden considerarse como reguladores positivos y negativos. Entre los primeros destacan *GLABRA1* (*GL1*) y *TRANSPARENT TESTA GLABRA1* (*TTG1*), cuyas mutaciones originan fenotipos de ausencia de tricomas (Koornneef, 1981). También es el caso de los genes *GLABRA3* (*GL3*) y *ENHANCER OF GLABRA3* (*EGL3*), que actúan de forma redundante (Zhang et al., 2003). Los mutantes *gl3* presentan menos tricomas que las plantas de fenotipo silvestre, mientras que los dobles mutantes *gl3 egl3* muestran una ausencia total de tricomas (Zhang et al., 2003). En cuanto a los represores de la formación de tricomas están representados por tres genes también de función redundante. Las mutaciones en el gen *TRYPTYCHON* (*TRY*) dan lugar a la formación de agrupaciones de tricomas, mientras que mutaciones en el gen *CAPRICE* (*CPC*) causan un incremento de la densidad de tricomas en la hoja (Schellmann et al., 2002). Finalmente, mutaciones en el gen *ENHANCER OF TRYPTYCHON AND CAPRICE* (*ETC1*), un inductor de *TRY* y *CPC* dan lugar a un fenotipo indistinguible del fenotipo silvestre (Kirik et al., 2004).

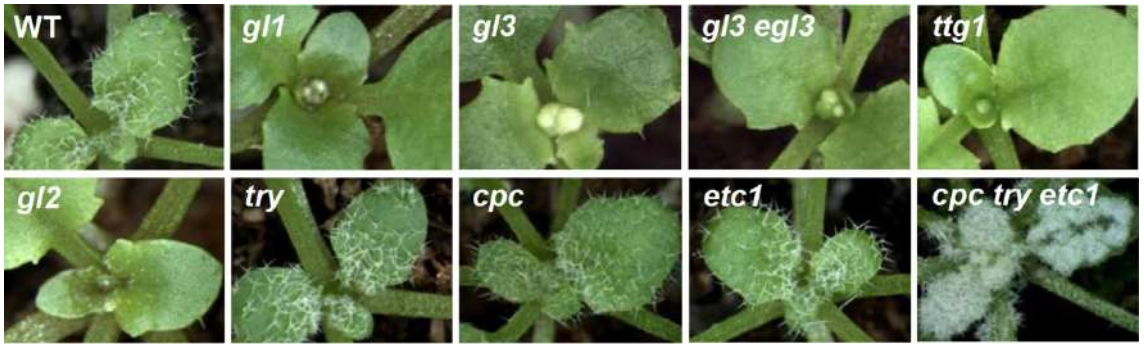


Figura 7: Mutantes de genes inductores y represores de la morfogénesis de tricomas unicelulares en *Arabidopsis thaliana*. *gl1*: mutante *glabra1*; *gl3*: mutante *glabra3*; *gl3 egl3*: doble mutante *glabra3-enhancer of glabra3*; *ttg1*: mutante *transparent testa glabra1*; *gl2*: mutante *glabra2*; *try*: mutante *tryptychon*; *cpc*: mutante *caprice*; *etc1*: mutante *enhancer of tryptychon and caprice1*; *cpc try etc1*: triple mutante *caprice-tryptychon-enhancer of tryptychon and caprice1*. Adaptado de Wester, (2009).

Estudios posteriores han demostrado que existen interacciones físicas entre las proteínas codificadas por los genes antes citados y que el complejo proteico que forman es el responsable directo de la iniciación del proceso de morfogénesis de los tricomas. *GL1* codifica un factor de transcripción de tipo MYB R2R3 (Oppenheimer et al., 1991) que se ha demostrado que interactúa tanto con el dominio N-terminal de GL3 como con EGL3 (Payne et al., 2000), siendo estas últimas proteínas factores de transcripción de tipo bHLH (*basic helix-loop-helix*) (Zhang et al, 2003). También se ha demostrado que la proteína TTG1, que contiene cuatro dominios conservados de naturaleza WD interacciona tanto con GL3 como con EGL3, pero a través de una región diferente de la empleada en la interacción con GL1 (Zimmermann et al., 2004). La diana aguas abajo de este complejo inductor parece ser el gen que codifica para la proteína GLABROUS2, un factor de transcripción de tipo *homeodomain leucine zipper* (Serna y Martin, 2006).

Por otra parte, las proteínas represoras de la iniciación de los tricomas también forman complejos de interacción. *TRY* y *CPC* codifican para factores de transcripción de tipo MYB que interaccionan con el dominio N-terminal de GL3 y EGL3 y que por tanto compiten activamente con GL1 por el sitio de unión a GL3 y EGL3, previniendo la formación del

complejo de activación (Payne et al., 2000; Larkin et al., 2003). Dos modelos que representan la interacción entre las proteínas que componen los complejos inductor y represor de la formación de tricomas descritos en *A. thaliana* se muestran en la **Figura 8**.

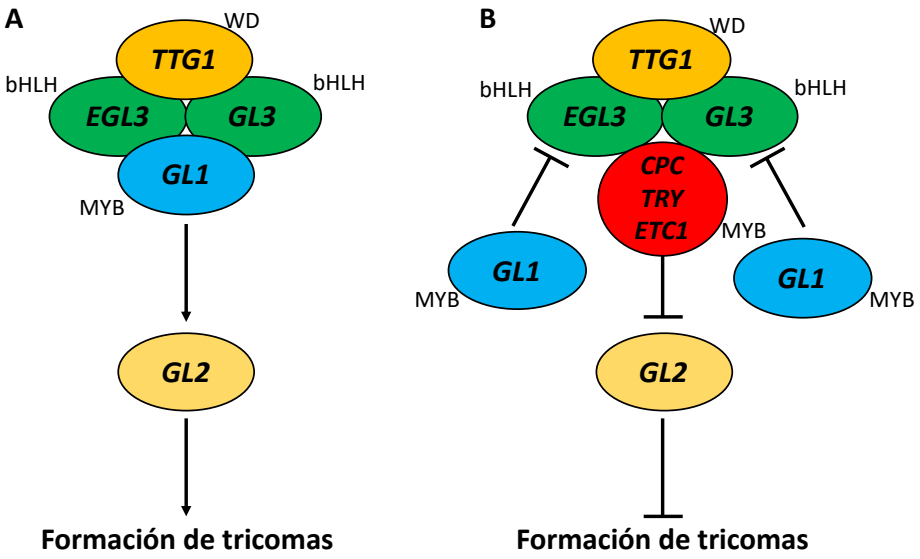


Figura 8: Complejos proteicos implicados en la morfogénesis de tricomas unicelulares de *Arabidopsis thaliana*. A, representación del complejo formado por la interacción de las proteínas GL1, TTG1 y GL3-EGL3 y de su diana aguas abajo, la proteína GL2 que promueve la iniciación de tricomas en la epidermis. **B,** la unión de las proteínas TRY y CPC a los dominios bHLH de las proteínas Gl3 y EGL3 previene la unión de estas a GL1 y en consecuencia reprime la formación de tricomas.

3.4. Tricomas multicelulares en tomate

Las rutas génicas que controlan la morfogénesis de tricomas en Arabidopsis a través del complejo WD-bHLH-MYB parecen estar restringidas a las rósidas y se encuentran muy bien caracterizadas, pero este no es el caso de las astéridas, que engloban a familias de enorme interés agronómico como las Solanáceas.

Los tricomas del género *Solanum* fueron descritos primero por Luckwill (1943) y esta clasificación ha sido posteriormente revisada por Channarayappa (1992). Todos son de tipo multicelular y pueden diferenciarse en glandulares cuando las células distales

producen y secretan sustancias de diferente naturaleza y en no glandulares. Los tricomas no glandulares constituyen una barrera física que restringe la movilidad de las plagas de herbívoros mientras que los de tipo glandular secretan sustancias tóxicas que atrapan o incluso matan a los insectos (Simmons y Gurr, 2005). La clasificación detallada de los tricomas presentes en el género *Solanum* se recoge en la **Tabla 2**.

Tabla 2: Tricomas presentes en el género *Solanum* según la descripción de Luckwill (1943) revisada por Channarayappa (1992).

Tipo	Descripción
I	Glandular formado por entre 6–10 células y 2–3 mm de largo. De base multicelular globular y con una célula glandular pequeña en el extremo distal.
II	Similar al tipo I, pero no glandular y más pequeño (0,2-1mm). De base globular y multicelular.
III	Delgado, no glandular formado por entre 4-8 células y de 0,4-1 mm de largo. Base unicelular y plana. Las paredes externas carecen de secciones intercelulares.
IV	Glandular de pequeño tamaño (0,2-04 mm) y base unicelular y plana.
V	Similar al tipo IV, pero no glandular.
VI	Tricomas glandulares y cortos formados por un pedúnculo de dos células y cuatro células glandulares.
VII	Tricomas glandulares muy pequeños (0,05 mm) con 4-8 células secretoras.
VIII	No glandulares compuestos por una célula basal gruesa

Existen diferencias en los tipos de tricomas que se encuentran en la especie cultivada de tomate y los que pueden aparecer en especies silvestres cercanas. De este modo, la especie cultivada tiene 5 de los tipos de tricomas antes descritos: de tipo I, III, V, VI y VII, mientras que en las especies silvestres *S. pennelli* y *S. habrochaites* no se encuentran los de tipo II y V (Simmons y Gurr, 2005). Numerosos estudios han caracterizado los perfiles de exudados producidos por cada uno de los tipos de tricomas, así como su impacto en las plagas.

Los tricomas predominantemente asociados con efectos negativos sobre las plagas son los de tipo IV y VI. Los tricomas de tipo IV producen sustancias denominadas acil azúcares mientras que los de tipo VI producen una gran variedad de terpenos y metil quetonas (Glas et al., 2012). En el caso de los tricomas de tipo VI, se ha demostrado que

las metil quetonas 2-undecanona y 2-tridecanona se asocian a diferentes efectos negativos sobre plagas de lepidópteros como *Tuta absoluta* (Vercosa de Magalhaes et al., 2001) o de hemípteros como *Aphis gossypii* (Williams et al., 1980). Estos tricomas producen también terpenos como los denominados sesquiterpenos. Entre los efectos negativos de estos compuestos sobre las plagas de herbívoros se ha demostrado un efecto repelente sobre la mosca blanca *Bemisia tabaci* (Bleeker et al., 2009), una reducción en la oviposición y la alimentación de *T. absoluta* (Eigenbrode et al., 1994) y que su eliminación de la superficie de la hoja mediante tratamientos con metanol incrementa la supervivencia de las larvas de *Spodoptera exigua* en un 65% (De Azevedo et al., 2003).

En cuanto a los acil azúcares o ésteres de azúcares, son exudados no volátiles producidos por los tricomas de tipo IV y sus efectos negativos sobre las plagas están ampliamente documentados. Estas sustancias son conjugados de azúcares y ácidos grasos alifáticos que en el caso de la especie silvestre S*. pennelli* representan hasta el 20% del peso seco de la hoja (Fobes et al., 1985). Estos compuestos son tóxicos para las plagas, pero también actúan como emulsificadores y surfactantes que las atrapan (Glas et al., 2012). Wagner et al (2004) demostraron que los áfidos en contacto con los tricomas producidos por plantas de tabaco (*Nicotiana tabacum*) son rápidamente atrapados por los acil azúcares. En tomate se ha demostrado que el elevado contenido en tricomas de tipo IV y el elevado contenido de acil azúcares incrementan la mortalidad de la araña roja *Tetranychus urticae* así como disminuyen su tasa de oviposición (Fernández-Muñoz et al., 2003; Alba et al., 2009).

De todo lo anteriormente expuesto se deduce la enorme importancia de los tricomas como primera línea de defensa en la resistencia de las especies cultivadas a plagas, para lo cual constituyen una excelente alternativa al uso extensivo de pesticidas. La capacidad de producción de estas estructuras de un gran número de sustancias favorece su uso potencial como biofactorías. De ahí que se hayan convertido en los últimos años en la diana de numerosos estudios con el fin de dilucidar los mecanismos moleculares implicados en su morfogénesis (Huchelmann et al., 2017).

3.5. Regulación génica del desarrollo de tricomas multicelulares en solanáceas: tomate

Como se ha mencionado anteriormente, las rutas e interacciones génicas que controlan la morfogénesis de tricomas multicelulares en solanáceas no están descritas al nivel de detalle de *A. thaliana*. Ello ha motivado en los últimos años numerosos estudios con el objetivo de determinar los genes implicados en este proceso. La herramienta más útil y comúnmente empleada para ello ha sido el análisis genético de mutantes.

Entre los primeros genes caracterizados se encuentra un factor de transcripción de tipo MYB denominado *MIXTA* identificado en *A. majus*, que regula la formación de células cónicas, consideradas estadios iniciales de la diferenciación de tricomas (Payne et al., 1999). La sobreexpresión de *MIXTA* bajo el control del promotor 35S del virus del mosaico de la coliflor promueve la formación ectópica de tricomas multicelulares en las hojas en *Antirrhinum* (Martin y Paz-Ares, 1997) y su sobreexpresión en tabaco (*Nicotiana tabacum*) tiene el mismo efecto, lo que sugiere que en Solanáceas existen proteínas de tipo MYB implicadas en la formación de tricomas (Glover et al., 2004). El papel de MIXTA es por tanto similar al desempeñado por otras proteínas de tipo MYB implicadas en la formación de tricomas en Arabidopsis, como GL1 y podría parecer que este proceso está regulado de forma similar en Rósidas y Astéridas. Sin embargo, este argumento ha quedado refutado, toda vez que la sobreexpresión de *GL1* en tabaco no tiene ningún efecto en la formación de tricomas (Payne et al., 1999). Tampoco tienen efecto alguno la sobreexpresión en tabaco de un factor de transcripción de tipo bHLH identificado en maíz y que induce la formación de tricomas en hojas de Arabidopsis (Lloyd et al., 1992).

Estudios recientes encaminados a diseccionar las redes génicas que controlan la morfogénesis de tricomas en solanáceas se han centrado en tomate como especie modelo. Hasta el momento se han identificado genes clave para la formación y morfología de los tricomas, particularmente los de tipo I. Tal es el caso del mutante *woolly* (*wo*). Se trata de una mutación espontánea detectada en la especie cultivada que produce un incremento en la densidad de tricomas del tipo I acompañada de letalidad embrionaria (Huang y Paddock., 1962). *WO* codifica para un factor de transcripción de tipo HD-Zip que se ha demostrado que interacciona físicamente con una ciclina de tipo

B denominada *SlCYCB2*. Dado que las ciclinas regulan la transición del estadio G2 a mitosis, la hipótesis más probable para el mecanismo de acción de estos genes es que *Wo* promueve la formación de tricomas mediante la regulación de *SlCYCB2*, que a su vez promueve el cambio de endoreduplicación a mitosis (Yang et al., 2011).

Otro de los mutantes caracterizados es el mutante *hair absent* (*h*), que muestra un fenotipo de ausencia de tricomas de tipo I en todos los tejidos. La mutación responsable de este fenotipo se localiza en el gen *HAIR* (*H*), que codifica para un factor de transcripción de tipo C2H2 dedos de zinc (Chang et al., 2018). Se ha demostrado que la proteína H interacciona físicamente con WO mediante ensayos de dobles híbridos y de coprecipitación d*e* ambas proteínas. Ello unido al hecho de que los dobles mutantes *h* y *wo* presentan un fenotipo de ausencia de tricomas de tipo I y un incremento en la densidad de tricomas de pequeño tamaño, sugiere que estos dos genes constituyen una ruta de control de la formación de tricomas multicelulares en tomate (Chang et al., 2018) (**Figura 9**).

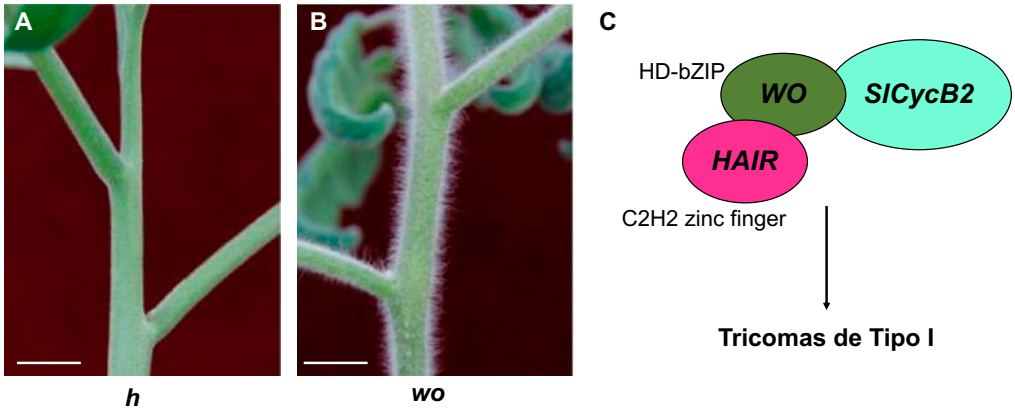


Figura 9: Mutantes de tricomas caracterizados en tomate (*Solanum lycopersicum*). A. Fenotipo de una planta mutante *hair absent* (*h*). **B.** Mutante *woolly* (*wo*). **C.** Ruta de interacción génica propuesta para los genes *HAIR* (*H*), *WOOLLY* (*WO*) y *SlCYCB2* que regula la morfogénesis de tricomas multicelulares de tipo I en tomate. Las barras de escala en **A** y **B** se corresponden con 1 cm. Adaptado de Chang et al., 2018.

Las mutaciones *wo* y *h* ocurren en genes represores e inductores respectivamente de la síntesis de tricomas. Sin embargo, se han descrito mutaciones de distinta naturaleza que

afectan a este proceso, como las implicadas en la organización del citoesqueleto y los filamentos de actina. El ejemplo más evidente lo constituye el mutante *hairless* (*hl*), que se caracteriza por presentar tricomas de morfología aberrante (Rick y Butler, 1956). *HL* codifica la proteína altamente conservada SRA1, que constituye una de las subunidades del complejo proteico WAVE, conocido por su implicación en la regulación de la ramificación de los filamentos de actina (Kang et al., 2016). Los mutantes *hl* también presentan deficiencias en la acumulación de sesquiterpenos y polifenoles en los tricomas glandulares de tipo VI, fenotipo asociado con una mayor susceptibilidad a herbívoros (Kang et al., 2010).

El análisis genético de los mutantes de tomate hasta ahora descritos ha permitido esclarecer las rutas que controlan la morfogénesis de tricomas multicelulares de tipo I en esta especie. Sin embargo, este conocimiento es aún muy limitado y se requieren estudios que en el futuro permitan caracterizar nuevas rutas e identificar nuevas funciones génicas relacionadas con este proceso, de una importancia vital para la resistencia de esta especie cultivada a plagas que en última instancia afectan su productividad.

Objetivos

El tomate (*Solanum lycopersicum* L.) se ha convertido en los últimos años en el sistema modelo por excelencia entre las hortícolas de fruto carnoso. Su enorme importancia económica, así como algunas de sus características agronómicas como un ciclo de vida relativamente corto, han colocado a esta especie en el foco de la investigación científica reciente. Todo ello impulsó el proyecto de secuenciación de su genoma y el desarrollo de un gran número de herramientas moleculares encaminadas al estudio de nuevas funciones génicas. A pesar de todos estos avances, la mayoría de los genes anotados en el genoma de tomate continúan siendo predicciones bioinformáticas que deben ser complementadas con una caracterización funcional. Con este fin, el estudio de mutantes constituye una herramienta de genética directa y de genética inversa de enorme utilidad, tanto para la identificación de genes clave del desarrollo vegetal, como para la selección de nuevos fenotipos de caracteres de interés agronómico. Para contribuir a este fin, el grupo de investigación Genética y Fisiología del Desarrollo Vegetal (AGR-176) ha generado colecciones de mutantes de tomate mediante programas de mutagénesis insercional (T-DNA) y química (EMS). Como parte de la caracterización agronómica y molecular de dichas colecciones se han identificado mutantes de enorme interés agronómico. El objetivo principal de esta tesis doctoral es profundizar en el estudio de algunos de estos mutantes, para lo que nos hemos propuesto los siguientes objetivos específicos:

PRIMERO. - Caracterizar fenotípicamente una colección de mutantes de tomate generados mediante mutagénesis química con etil metano sulfonato (EMS), prestando especial atención a los caracteres relacionados con el desarrollo vegetativo y reproductivo.

SEGUNDO. - Realizar una caracterización molecular mediante mapeo por secuenciación de algunas de las líneas mutantes seleccionadas, con el fin de determinar la utilidad de la mutagénesis química para la identificación de los genes responsables de caracteres de interés agronómico.

TERCERO. - Llevar a cabo la caracterización fenotípica complementada con técnicas de microscopía electrónica y genética molecular del mutante de tomate *succulent stamens2* (*sus2*), que hagan posible la identificación del gen responsable de dicha mutación.

CUARTO. - Realizar una caracterización agronómica y molecular del mutante *hairplus* (*hap*), clonando el gen *HAP* y demostrando que es el responsable del fenotipo mediante la obtención de líneas transgénicas de complementación, silenciamiento génico mediante RNAi y la obtención de nuevos alelos mediante edición genética (CRISPR).

QUINTO. – Proponer el mecanismo molecular a través del cual el gen *HAP* controla la formación de tricomas glandulares en tomate.

Capítulo I: Characterisation of a tomato mutant collection as a valuable resource for breeding and tomato functional genomics

Abstract

Tomato (*Solanum lycopersicum* L.) is a major crop that is gaining importance as a model species among the eudicots, especially for trait related to fleshy fruit development. Although a considerable progress has been achieved since the tomato genome sequence project was completed, most of the genes identified remain as predictions with unknow or with a hypothetical function. This lack of functional characterization hampered the use of the huge amount of information contained in the genome to increase the quality and productivity of this crop. Reverse genetics strategies such as artificial mutagenesis, as well as next generation sequencing approaches form the perfect tandem for increasing knowledge in tomato gene functional annotation. This study reports the phenotypic and molecular characterization a tomato mutant collection generated from a chemical mutagenesis program carried out with EMS. Causal mutations of the selected mutant lines have been isolated by a mapping by sequencing approach and most of them account for novel gene functions. All these results support that tomato mutagenesis is an essential tool for functional genomics in this fleshy-fruited model species and a highly valuable resource for future breeding programs of this crop species.

Capítulo I

I.1. Background

Global agriculture systems and food production currently face important issues. Mainly, a global population likely to grow to nearly 10 billion people for 2050 makes necessary a major crop productivity capable to face this demographic grow as well as it minimises environmental impact. The development of new crop varieties with desirable agronomical traits such as high productivity rates becomes crucial for this purpose. Since natural variation is limited, genetic improvement programs become a suitable alternative and particularly induced mutagenesis programs (Gulfishan et al., 2015). A successful breeding program should be ideally based on a combination of induced mutagenesis and advanced molecular biology techniques (Jain, 2010).

Tomato (*Solanum lycopersicum* L.) is one of the major crop species whose global demand has increased over the last decades (Laskar et al., 2018), mostly due to its diverse forms of consumption (fresh or processed) and to its highly nutritive values. Tomato is considered to be beneficial since it is a rich source of antioxidants and beta-carotene as well as vitamins (Hobson & Grierson, 1993; Kalloo, 1991) while it may prevent cardiovascular disease as well (Arab y Steck, 2000). This species is a good example of the successful use of mutations in breeding programs. Moreover, mutations as *self-pruning* (*sp*), *ovate* (*o*) or *jointless2* (*j2*) are widely used in the processing industry to determinate compact growing plants, with elongated fruits and more suitable for mechanical harvesting (Rick et al., 1976). Tomato is considered as a model species among fleshy fruit plants due to its favourable agronomic traits such as short life cycle or self-pollination, as well as biotic and abiotic stress resistance (Kissoudis et al., 2015; Bai et al., 2018). Furthermore, tomato genome is relatively small (950 Mb) and its full sequence is available, which eases the genetic improvement of this species (The Tomato Genome Consortium, 2012). Nevertheless, only a small percentage of the about 35.000 genes annotated in the tomato genome have an associated function and given that natural mutant alleles are also a few, mutational analysis is revealed as a suitable strategy for gene functional characterisation and tomato genetic improvement.

Induced mutagenesis techniques available include insertional mutagenesis as well as the employment of physical and chemical agents that have very diverse effects on target

genomes. Insertional mutagenesis methods such as T-DNA have been also employed in tomato using methods like enhancer trapping (Atarés et al., 2011; Pérez-Martín et al., 2017) but the need for *in planta* transformation using *Agrobacterium tumefaciens* remains a time-consuming task. Among the physical agents, fast neutrons have been employed successfully in *Arabidopsis thaliana* and rice (*Oryza sativa*), where a total of 51.840 and 10.000 mutant lines respectively have been obtained (Li et al., 2001; Wu et al., 2005). Nevertheless, fast neutrons cause huge deletions or even the complete loss of entire chromosomes, which diminishes mutant seeds viability (Vrebalov et al., 2002; Sikora et al., 2011). Other physical agents include gamma and X rays as well as UV light, whose effects range from small modifications to deletions (Yuan et al., 2014; Ries et al., 2000).

Alkylating agents such as EMS (ethyl methanesulfonate) are highly effective and relatively easy to handle. EMS commonly alkylates guanine (G) rendering O6-etylguanine, which pairs with thymine (T) instead of cytosine (C) and as a result, C/G to T/A transitions are produced. Nevertheless, G-C to C-G or C/G-C to T-A transversions by 7-methylguanine hydrolysis are also produced, although at a lower frequency (Krieg, 1963). EMS mutagenesis combined with high-throughput methods for detecting point mutations such as TILLING (targeting induced local lesions in genomes) are a powerful tool for reverse genetics in tomato (Okabe et al., 2011). With the aim to characterise new gene functions and mutant alleles in loci of agronomical interest, we have carried out a mutagenesis program using EMS in seeds of the tomato cultivar Moneymaker (MM) which has allowed us to obtain a mutant collection that comprises almost 9.000 mutant lines. Part of this mutant collection has been screened in this work under greenhouse conditions and a large number of new mutant alleles have been identified. This way, EMS induced mutagenesis reveals once more as a suitable approach for identification of key regulators of plant growth and reproductive development.

I.2. Methods

EMS mutagenesis

Mutant collection development began with M_0 seeds mutagenesis. Briefly, seeds of the tomato cultivar Moneymaker (MM) were incubated in 250 ml of a solution containing 0.7% of EMS (Sigma Aldrich) (w/v). Incubation proceeded for 16 hours with gentle shake at 30º. Afterwards, seeds were washed twice in distilled water and sown for M_1 plant cultivation. The used EMS and wash solutions were treated with NaOH for mutagen neutralization and afterwards disposed as a regular waste. M_1 plants were cultivated under greenhouse conditions and once reached fruit ripening phase, seeds from single M_1 plants were collected. M_2 seeds were incubated for 12 hours in a solution containing 1% of hydrochloric acid (v/v) and then washed with distilled water for 10 minutes in individual metal mesh. After room temperature dry, seeds were kept for 24 hours in a drying oven at 80º for virus elimination. Phenotypic characterisation of M_2 plants was performed along with WT plants under greenhouse conditions and trough four cultivation cycles distributed between spring 2014 and autumn 2018.

Phenotypic characterisation

Mutants were grouped in six main categories depicted in **Table 3,** which included plant habit, plant vegetative development, plant reproductive development, parthenocarpic fruit production, seed production and fruit quality traits. When it comes to fruit quality traits, Brix degrees were measured using a digital refractometer Atago PR-101 and a PH 25+ (Crison) was used for pH determination. Fruit firmness was assessed with a digital firmness tester (Durofel DFT 100) using a 5.64 mm diameter tip. Regarding pollen viability assays, these were carried out by *in vitro* staining of pollen grains in a 0.5 % solution of 2,3,5-triphenil tetrazolium chloride (TTC) in a 0.5 M solution of sucrose and were incubated two hours at 50ºC in darkness in a humid box. Finally, an OPTIPHOT-2 (Nikon) optical microscopy was used for results visualization.

Genetic characterisation of mutants of interest

Following mutant identification, genetic characterisation was carried out by complementation tests in those mutant lines suspected to be allelic of a previously

described mutation. In the rest of cases, a mapping by sequencing strategy was performed. First, F_2 population in obtained by crossing mutant plants with wild relative *S. pimpinellifo*lium accession LA1589 as pollen donor. This wild relative accession was obtained from Tomato Genetics Resource Center (http://tgrc.ucdavis.edu/). Following F_2 characterisation, plant material of mutant and WT plants is collected for DNA isolation. Genomic DNA was isolated with DNAzol® Reagent Kit following manufacturer´s instructions and equimolar amounts of DNA from each WT and mutant F_2 plant were used to form two independent pools. Afterwards, mapping of these two pools was performed and allele frequencies analysis was performed.

Mapping by sequencing bioinformatic analysis

Illumina TruSeq protocol was followed for library generation and sequencing was carried out in an Illumina HiSeq2000 platform (Illumina, Inc., USA). Alignment of paired-end-reads was performed against the tomato genome reference sequence version 2.5 (ITAG2.4) using Bowtie2 (Langmead & Salzberg, 2012). Duplicated reads were removed using Picard 1.65 and for indel realignment GATK 2.2-8 was employed. VCFtools was used for variant calling (Danecek et al., 2011) and SAMtools 1.2 for obtaining reference and non-reference allele counts for each position of the genome. Finally, determination of the chromosomic region where the mutations are located was carried out with a custom script in the R environment for statistical computing (R Development Core Team, 2011).

I.3. Results

Development of EMS mutant collection

EMS treatment was performed in seeds of the tomato cultivar Moneymaker (MM), which is characterised by an indeterminate growth and medium size fruits. With the aim to optimise mutagenesis, the mutagen doses required to obtain a large number of mutant lines without compromising seed viability and sterility was established as a preliminary step. Three treatments with different EMS concentrations were then assessed i.e. 0.5%, 0.7% and 1%. The number of seedlings was lower the higher EMS concentration and a delay in seed germination was observed as well. Afterwards, LD_{50}

was assessed for each of the treatments employed. This parameter measures not only the doses required for 50% of individuals to die, but also the amount of survival plants leaving descendance. In our conditions, 0.7% can be considered as the LD_{50}, since at this concentration germination rate decreases to a 76.65%, whereas a 49.28% of these M_1 individuals yielded M_2 plants (**Figure 10**). Afterwards, a large-scale mutagenesis experiment was performed at 0.7% EMS concentration in MM seeds.

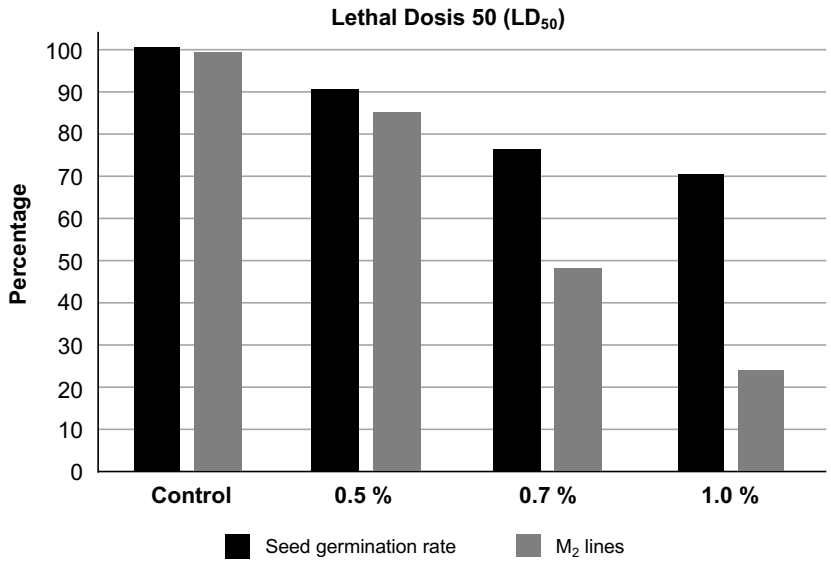


Figure 10: Lethal dosis 50 (LD50%) assessed in three different EMS mutagenesis experiments. EMS 0.7% concentration is considered as the LD_{50}, since 49.28% of M_1 plants leave descendance. At 0.5% of EMS greater survival rate is observed, but mutant number was significantly low, which is the opposite of the results obtained using 1% concentration of EMS, where M_2 seeds obtained drops to 23.85%.

Phenotypic screening of EMS mutant lines

A total of 1417 M_1 plants were screened under greenhouse conditions and three of them (0.21%) showed phenotypic alterations due to dominant mutations. 177 M_1 (12.06%) plants were detected to be parthenocarpic and 24 plants died during cultivation making impossible to obtain M_2 seeds. For the rest, mature fruits were collected in order to obtain M_2 segregating families for the study of recessive mutations. Screening of 749 out of the 1216 M_2 families obtained was performed throughout four cultivation cycles.

In order to detect recessive mutants 12 plants of each M_2 family were transplanted meaning that around 8988 M_2 were screened for alterations in visible phenotypes. All mutant phenotypes detected were grouped into six main categories: plant habit, plant vegetative development, plant reproductive development, parthenocarpic fruit production, seed production and fruit quality traits (**Table 3**).

Table 3: Phenotypic categories used for M_2 mutant plants screening.

Main classes	Subclasses
Plant habit	Larger or smaller size due to internode length Branching Plant architecture
Plant Vegetative Development	Leaf morphology: Size, colour and number of leaflets Trichome morphology and density
Plant Reproductive Development	Flowering time Flower Development: homeotic mutations, whorls size and number and flower colour Inflorescence structure Fruit: morphology, shape, size and colour Ripening: early or late ripening
Parthenocarpy	
Seed production	Number of seeds per plant
Fruit Quality traits	Fruits ºBrix, pH and diameter

Part of the mutants that exhibited vegetative development and plant growth alterations are represented in **Figure 11**. The most abundant category in this group was plant size (25.30%), followed by leaf morphology (11.27%). We have characterised mutants altered in plant architecture (Fig. 11A), branching (Fig. 11B), as well as dwarfism and compact growth (Fig. 11C y Fig. 11D). Also, extreme albinism was detected in mutant line UAL-7632 (Fig. 11E), whereas chlorosis and leaf colour phenotypes were detected in lines UAL-7308 (Fig. 11F) and UAL-7368 (Fig. 11G) respectively.

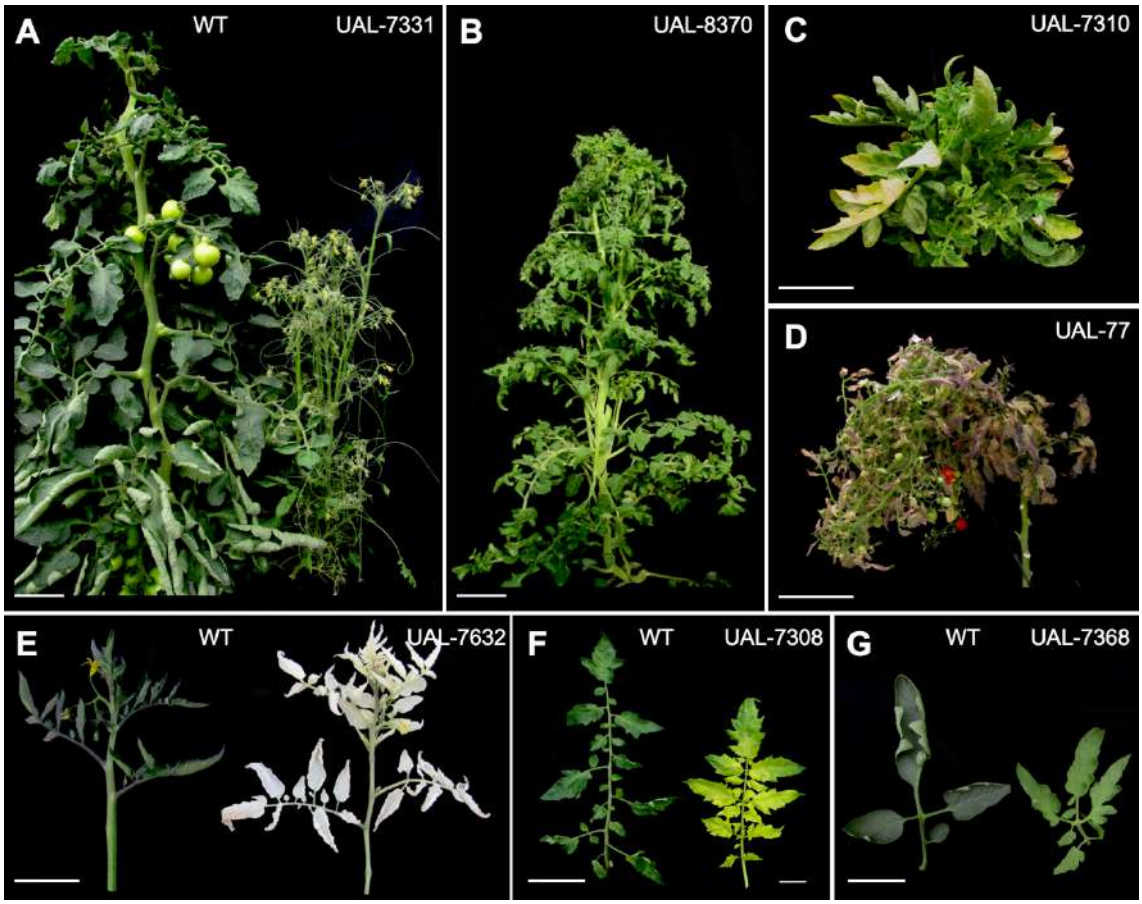


Figure 11: Plant growth and vegetative development mutants. A. Left Moneymaker wild type (WT) plant and right UAL-7331 mutant line, which exhibits severe plant architecture alterations and a wiry like phenotype. **B.** Mutant line UAL-8370, characterised by an extreme branching, rounded leaflets and late flowering time. **C.** Dwarf mutant line UAL-7310. **D.** Mutant line UAL-77 which shows compact growth. **E.** Albino mutant line UAL-7632 (right). **F.** Leaf chlorosis in UAL-7308 mutant line (right). **G.** Leaf morphology and colour in mutant line UAL-7368 (right). Scale bars apply for 2 cm in **A** and **B** and to 5 cm from **C** to **G**.

Regarding reproductive development, the most common phenotype characterised was parthenocarpy (33.76%), followed by fruit morphology and colour (17.83%). Some of these mutant phenotypes are displayed in **Figure 12**.

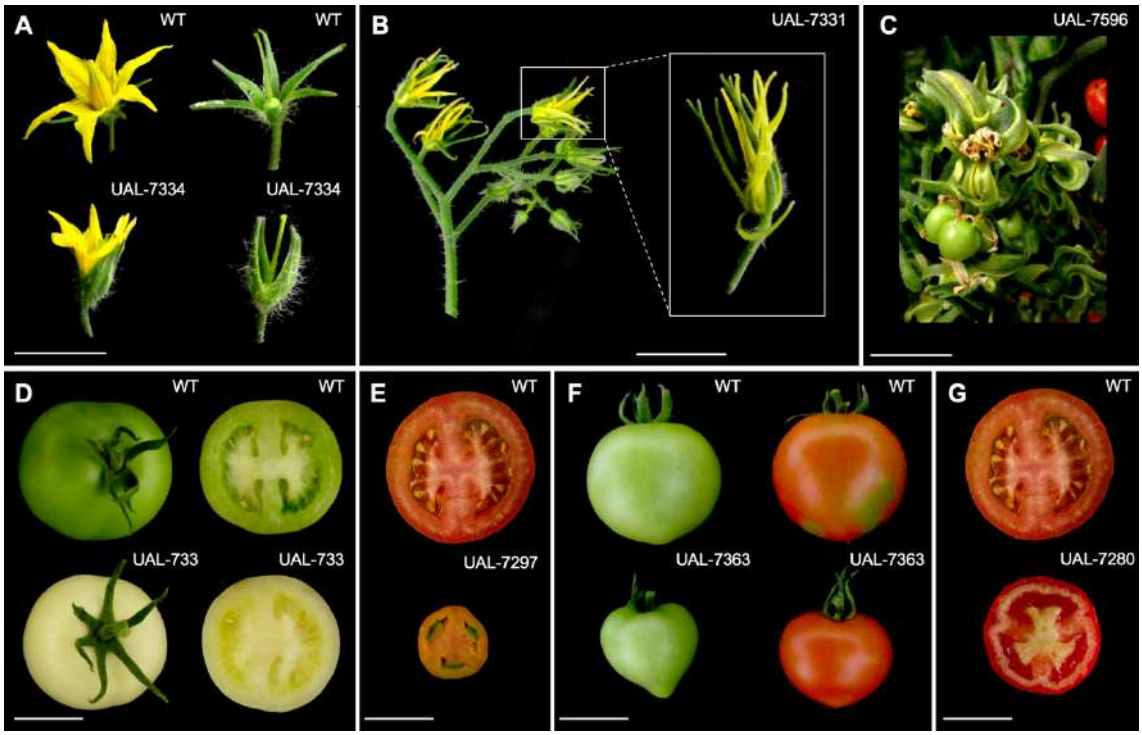


Figure 12: Reproductive development mutants. A. Upper: Moneymaker wild type (WT) flowers and UAL-7334 mutant line (down), which exhibits fusion of the first whorl of sepals. **B.** Mutant line UAL-7331, whose flowers have wiry phenotype and yield parthenocarpic fruits. **C.** Mutant line UAL-7596, characterised by extreme floral homeotic changes, as wells as aberrant inflorescence architecture and fruit morphology. **D.** Mutant line UAL-733, whose fruits at mature green stage exhibit a pale-yellow colour when compared to WT plants. **E.** Parthenocarpic and smaller fruits of the mutant line UAL-7297 (down). **F.** Fruits of UAL-7363 mutant line (down) with heart-shaped fruits. **G.** Mutant phenotype of UAL-7280 line whose fruits exhibit a wrinkled aspect and a pericarp of lignified aspect in a longitudinal cut. Scale bars apply for 1 cm from **A** to **C** and to 5 cm from **D** to **G**.

Another important phenotypic category was that of fruit quality traits. Along all the cultivation cycles, Brix degrees, pH and diameter of 10 fruits of each M_2 plants were measured, as well as fruit firmness. A total number of 95 mutant lines were detected (12.68%) to exhibit increased values in any of these three categories.

As previously described by Menda et al., 2004, reproducibility of phenotyping must always be assessed as a verification step, and M_2 should only be treated as likely

mutants. Thus, a total of 217 M_3 seeds were re-planted. Out of them, 187 showed the same phenotype described in M_2, while 11 lines showed a new phenotype.

Genetic and molecular characterisation of EMS induced mutations

The genetic characterisation of some of the mutant lines here described was carried out following two main strategies. First, complementation tests were performed when a mutation was suspected to be allelic of other previously described. This is the case of UAL-733 mutant line, whose fruits at mature green stage are pale-yellow coloured (**Figure 12D**). Crosses performed with the mutants *lutescent1* (*l1*) and *lutescent2* (*l2*) previously described by Barry et al., 2012 proved that the mutation detected in line UAL-733 was allelic with *l1* (**Figure 13**).

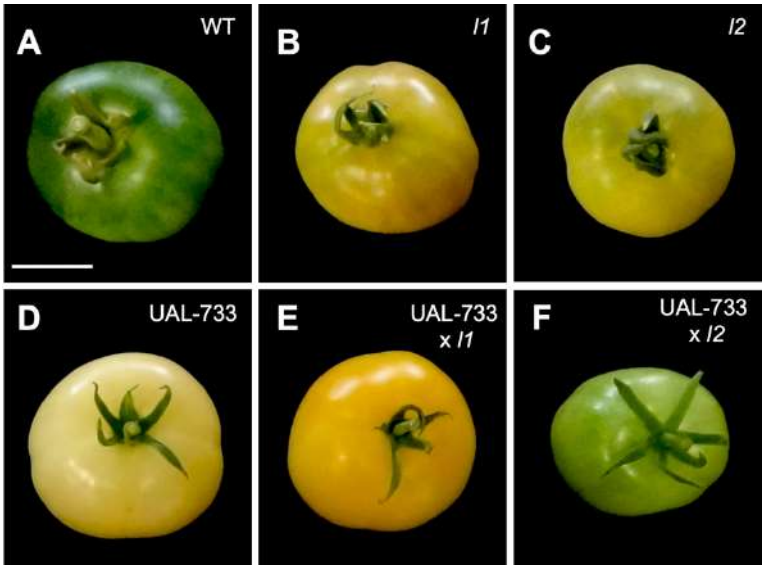


Figure 13: Complementation test performed in line UAL-33. A. Wild type MM plants. **B.** *lutescent1* (*l1*) mutant phenotype. **C.** *lutescent2* (*l2*) mutant phenotype. **D.** UAL-733 mutant phenotype. **E.** Phenotype observed in lines derived from the cross of UAL-733 mutant line and *l1* mutant prove that both mutations are allelic. **F.** Phenotype observed in lines derived from the cross of UAL-733 mutant line and *l2* mutant. Scale bars apply for 2.5 cm.

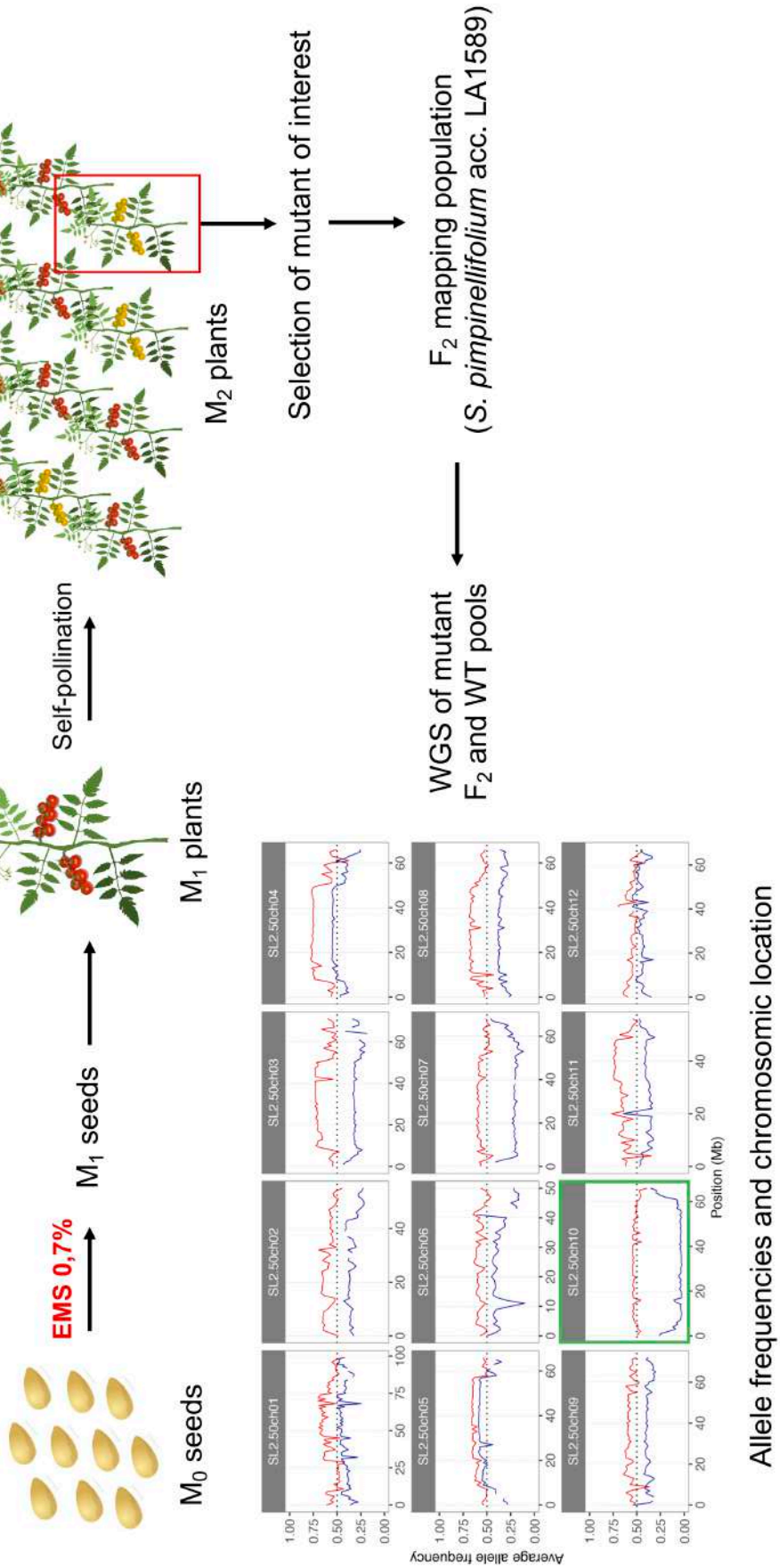

M0 seeds
EMS 0,7%
M1 seeds
M1 plants
Self-pollination
M2 plants
Selection of mutant of interest
F2 mapping population
(S. pimpinellifolium acc. LA1589)
WGS of mutant
F2 and WT pools
Allele frequencies and chromosomic location
Average allele frequency
Position (Mb)
SL2.50ch01
SL2.50ch02
SL2.50ch03
SL2.50ch04
SL2.50ch05
SL2.50ch06
SL2.50ch07
SL2.50ch08
SL2.50ch09
SL2.50ch10
SL2.50ch11
SL2.50ch12

Figure 14: An overview of mutagenesis and mapping by sequencing. First, M_0 seeds are exposed to EMS at 0.7% for 16 hours and M_1 seeds are sowed. M_1 plants are phenotypically characterised in order to detect dominant mutations (scarce) and M_2 seeds are obtained for further characterisation. After detection of mutant phenotypes of interest, F_2 interspecific population derived from the cross with *S. pimpinellifolium* accession LA1589 as pollen donor are obtained. Sequencing of mutant and WT pools of DNA allow to obtain allele frequencies along the 12 tomato chromosomes. Causal mutation is located in the region where allele frequency of mutant pool (red line) falls to 0 when compared to WT (blue line), which in this example is outlined by a green box.

Another approach was the development of F_2 mapping populations derived from the cross of mutant lines and *Solanum pimpinellifolium* accession LA1589 as pollen donor. This way, a mapping by sequencing strategy was carried out by sequencing DNA pools formed by equimolar amounts of DNA from mutant F_2 and WT plants. Analysis of allele distributions in both pools allows the identification of the causal mutation in the genome position were the allele frequency of mutant pool reaches cero (Garcia et al., 2016). Following this approach, we have been able to determine the exact chromosomic location of some of the mutant lines. An overview of the whole process including mutagenesis and identification of causal mutations as well as allele frequency detection is showed in **Figure 14**.

Among the identified mutations following the mapping by sequencing strategy is the one responsible for the mutant phenotype observed in the line UAL-7334 (**Figure 15**). Mutant plants exhibit a fusion in sepals, that develop in a bilateral axis and not in the pentameric form characteristic of tomato flowers (Fig. 15A). Also, mutant plants develop smaller fruits which are parthenocarpic or have a lower number of seeds when compared to WT plants (Fig. 15B). Given the phenotype observed in sepals of mutant plants, we have called this mutant *sepal indehiscent* (*sin*).

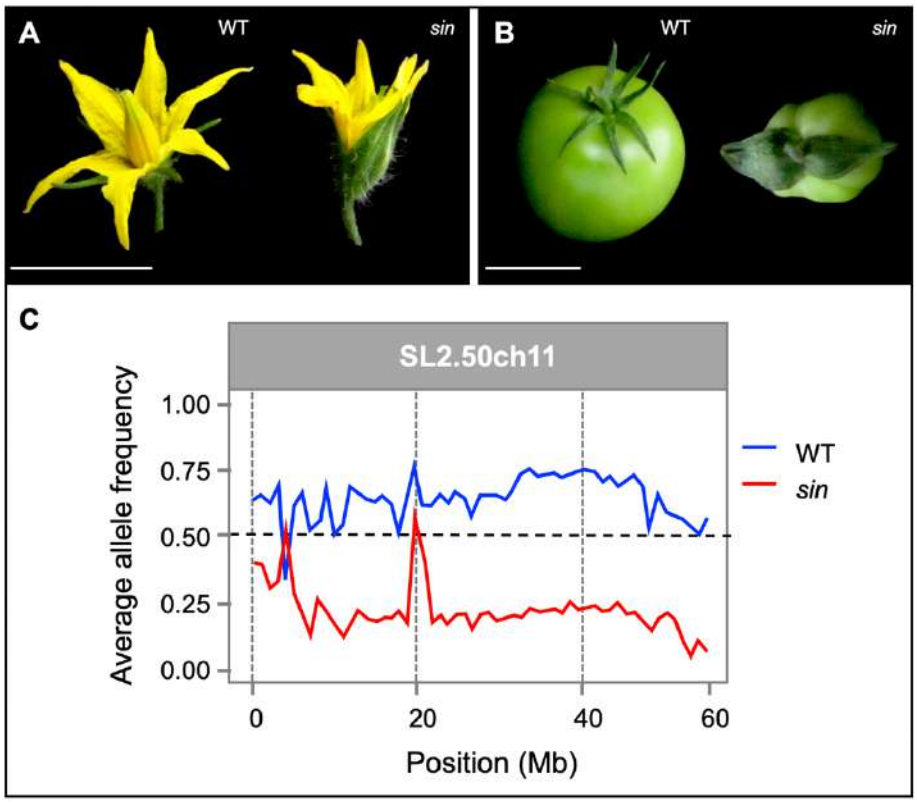


Figure 15: Mapping by sequencing performed in the mutant line UAL-7334 named *sepal indehiscent* (*sin*). A. Flowers of *sin* mutant plants show fused sepals. **B**. Fruits of mutant plants are smaller than those of WT plants. **C.** Mapping by sequencing performed in a pool of 15 F2 mutant plants allowed to identify the causal mutation in chromosome 11. Scale bars apply for 1 cm in **A.** and for 5 cm in **B**.

An F_2 mapping population derived from the cross of a *sin* mutant plant and *S. pimpinellifolium* accession LA 1589 was obtained. Sequencing of a pool of equimolar amounts of DNA of 15 F_2 mutant plants allowed us to detect that causal mutation was located in chromosome 11. Afterwards, variant analysis was carried out and a single nucleotide variant in an Auxin Response Factor (ARF) was detected in this candidate region. ARF´s have been related to a broad spectrum of biological processes, and their implication as mediators of the auxin perception in tomato in response to abiotic and biotic stress has been described (Bouzroud et al., 2018), as well as control of developmental issues such as leaf shape (Wu et al., 2018), fruit chlorophyll and sugar accumulation (Yuan et al., 2018). Thus, a novel function of this gene as a regulator of floral organ morphology and particularly of sepals is presented here. Also, identification

of this gene is a proof of concept of the methodology carried out for genetic characterisation of EMS mutant lines.

I.4. Discussion

Since full tomato genome sequence is available (Tomato Genome Consortium, 2012), huge progress has been made in comprehension of genome organisation and gene functions. As a result, 34.727 protein coding genes have been annotated by the international tomato annotation group (ITAG), although their functions remain unknown for most of them. Functional genomics approaches become particularly interesting when it comes to gene functional characterisation and among them, induced mutagenesis rises as one of the most relevant genomic tools. Tomato mutants have been and still are directly used in classical breeding programs (Pnueli et al, 1998; Mao et al., 2000), but natural variation of agronomically interesting alleles is scarce in natural germplasm. Given that, mutagenesis allows to create new allelic variants in a short period of time with the aim to shed light on gene functions through a reverse genetics strategy (Emmanuel & Levi, 2002). Moreover, tomato is a major crop whose favourable agronomic traits combined with all the genomic information available, made it a model species for fleshy fruit biology.

Chemical mutagenesis and particularly alkylating agents such as EMS have been widely employed in tomato and other crop species. In rice (*O. sativa*), this strategy has allowed to obtain mutants affected in agronomical traits of interest such as photosynthetic rate of leaves (Feldman et al., 2014), abiotic stress tolerance such as drought resistance related to root length and volume (Mohapatra et al., 2014) as well as a heat tolerant mutants that exhibits higher photosystem II efficiency (Poli et al., 2013). In maize (*Z. mays*), mutations induced by EMS affecting embryo morphogenesis have been described, whose main effects range from reduced germination, delayed development and aberrant morphology to lethality (Brunelle et al., 2017). Furthermore, identification of two maize EMS induced mutations related to dwarf and pale-green phenotypes respectively have been recently carried out by Heuermann et al., 2019. All these studies demonstrate the potential use of EMS mutagenesis for higher yield and stress resistance in grasses.

When it comes to tomato, chemical EMS mutagenesis collections have been mainly obtained in the cultivar Micro-Tom. This cultivar was developed for garden purposes and it is characterised by a dwarf phenotype determined by a combination of mutations affecting hormone perception as well as a determinate growth habit (Carvalho et al., 2011). Small size of Micro-Tom makes it suitable for cultivation of a large number of plants in a reduced space and a relatively short life cycle allows to obtain M_2 seeds faster. In contrast, Micro-Tom may not be the best option when it comes to identification of genes related to plant size or vigour as well as fruit size (Pérez-Martín et al., 2017). Nevertheless, Watanabe et al (2007) have carried out a characterisation of 3.839 mutant M_2 lines obtained in the Micro-Tom genetic background allowing them to identify a broad range of mutants. Furthermore, 8.598 Micro-Tom mutant lines have been characterised and added to a database called TOMATOMA (Shikata et al., 2016), which integrates mutant phenotypes as well as associated data making this information available for consult (Saito et al., 2011). On the other hand, another genetic background such as M82 have been used, with the advantage that can be grown as a determinate plant in an open field when it carries the *self-pruning* mutation as well as in an indeterminate form. This cultivar has been employed by Menda et al (2004) in the generation of 3.417 mutant lines, which provide of new isogenic alleles of the monogenic mutants already catalogued in The Tomato Genetics Resource Center (https://tgrc.ucdavis.edu/), although it also includes new ones.

In this study, the tomato commercial cultivar Moneymaker (MM) has been employed for EMS mutagenesis, as previously performed by other authors (Pérez-Martín et al., 2017). This background shows an indeterminate growth and a 4 to 6 months life cycle like most of the commercial varieties. Given that, time and a considerable greenhouse space has been consumed in characterisation of this mutant collection. Most relevant vegetative and reproductive phenotypes detected are depicted in **Figure 11** and **Figure 12** and account mainly for parthenocarpy (33.76%) and plant size (25.30%), as well as a large amount of fruit quality traits mutants (11.84%). Mutation identification performed by mapping by sequencing has allowed us to identify new alleles of previously described mutations such as *wiry*, whose novel allele has been detected in line UAL-7331.

Nevertheless, most mutations identified account for novel gene functions. Such is the case of the line UAL-7334 (*sin)*, where this approach has allowed us to identify novel roles for Auxin Response Factors (ARF) in tomato sepal development. Thus, the striking phenotypes detected in our mutant collection, as well as the functional characterisation of causal mutations provide new evidence of the key role of mutagenesis and mutant analysis as a valuable tool for tomato functional genomics.

Capítulo II: A homeotic mutant with impaired petal and stamen development confirms that *TM6* is a tomato B-class gene

Abstract

Flower development is a crucial step towards the completion of plants life cycle. Regulatory gene networks underlying their formation have been extensively characterised and implication of MADS-box transcription factors as primary regulators of flower morphology has been widely described mainly due to the analysis of loss of function mutants in model species. Nevertheless, detailed characterisation of mutant alleles of some of the homologous genes in crops species have not been described yet. Here, we have characterised a tomato mutant with aberrant petal and stamen development. Mutant plants exhibit changes in petal epidermal cells identity, as well as a homeotic change of stamens into carpels that in most cases results in succulent organs. Molecular analysis proved that a loss of function mutation in the proposed B-class gene *TOMATO MADS-BOX 6* (*TM6*) accounted for this mutant phenotype. Furthermore, misregulation of the expression of other MADS-box genes implicated in flower development due to the loss of function of *TM6* has been assessed. Our findings demonstrate that *TM6* is a B-class gene and a key player of the complex regulatory network of MADS-box proteins related with flower whorl identity.

Capítulo II

II.1. Background

With nearly 260.000 species classified into 453 families (Soltis et al., 2018), Angiosperms are without a doubt the more successful group among terrestrial plants mainly due to the development of a unique feature: the flower. Although there is a huge diversity of colours, forms and morphology among angiosperms flowers, the basic developmental program is highly conserved and flowers are usually formed by four whorls: sepals, petals, stamens and carpels. The organization of this structure has been in the centre of a large amount of investigations with the aim to dissect the molecular mechanisms controlling their development. The completion of these works has been the ABC model of flower development, that establishes the gene functions that control the four whorls identity (Coen & Meyerowitz, 1991).

According to the ABC model described in the model species *Arabidopsis thaliana* L., the A-class gene activity specifies the development of sepals in the outermost whorl, a role assumed by the *Arabidopsis* genes *APETALA* 1 and *APETALA* 2 (*AP1-2*) (Irish & Sussex, 1990). The A- and B-class genes coordinated activity specifies petals in the second whorl. The B-class function is represented by *AP3* and *PISTILATA* (*PI*), whose downregulation gives rise to sepal-like petals and carpels instead of stamens in the third whorl (Krizek & Meyerowitz, 1996). B- and C-class genes together are responsible for stamen development and finally C-class genes alone specify carpels in the fourth whorl (Irish, 2017). The C-class function is represented by *AGAMOUS* (*AG*), whose loss of function causes flowers to develop petals instead of stamens in the third whorl and instead of the gynoecium in the fourth whorl, another *ag* flower is developed (Yanofski et al., 1990). Furthermore, new gene functions have contributed to the enrichment of this model. Thus, D-class genes *SHATERPROOF* (*SHP1-2*) redundantly specify ovule identity (Colombo et al., 2010), whereas *SEPALLATA* (*SEP*) E-class gene function is crucial for meristem determination and all the four whorls identity (Pelaz et al., 2000).

Most ABC genes are MADS-box transcription factors, a widely distributed group of proteins with many important genetic, molecular and biochemical roles in plant life cycle (Riechmann & Meyerowitz, 1997). Major part of these genes had undergone

frequent duplication events and losses. In some cases, there is no evidence that these duplications have led to functional divergence and in others these paralogs adopt new functions or functions partially inherited from ancestors (Litt & Kramer, 2010). In spite of this complex evolutionary history, strong evidence supports that the ABC model as well as the basic developmental program that underlies flower ontogeny are widely conserved among angiosperms.

In tomato (*Solanum lycopersicum* L.), a major vegetable crop, ABC gene functions have been extensively characterised although not all the functions are supported by well characterised mutants. Mutations linked with stamen development and male sterility results particularly interesting, since they have potential uses in hybrid seeds production (Kim & Zhang, 2018). In this work, we have characterised a tomato male sterile mutant that shows clear alterations in flower morphogenesis which we have named *succulent stamens2* (*sus2*). Stamens of *sus2* mutant plants are aberrant and show carpel-like identity. Fine mapping and genetic complementation analysis showed that an *AP3* orthologue named *TOMATO MADS BOX 6* (*TM6*), a proposed B-class function gene was responsible for this phenotype. Our findings not only identify a new allele of this gene, whose functional characterisation has been performed only trough RNAi silencing approaches (de Martino et al., 2006), but also shed light on implications of B class genes loss of function in tomato fruit development.

II.2. Methods

Plant material and phenotypic characterisation

The *sus2* mutant was identified as part of the screening of a chemically mutagenised population obtained in the tomato cultivar Moneymaker (MM) by using ethylmethanesulfonate (EMS) as previously described (Capítulo 1). Briefly, seed were incubated with 0.7 % EMS at 30ºC with gentle shaking and after extensive wash were sown in seedling beds and germination was assessed prior to transplant. M_1 plants were cultivated under greenhouse conditions and self-pollinated to obtain M_2 segregating populations. The *sus2* mutant was selected by the conspicuous homeotic changes observed in the flowers of one M_2 segregating family. Phenotypic characterisation was

performed in M_2 populations grew along with control MM plants. Wild relative *S. pimpinellifolium* accession LA1589 retrieved from Tomato Genetics Resource Center (http://tgrc.ucdavis.edu/) was employed for generation of F_2 mapping population. All experiments were conducted under greenhouse conditions and following standard management practices, including regular fertilization.

Pollen viability assays

Pollen viability was assessed in vitro by staining pollen grains in a 0.5 % solution of 2,3,5-triphenil tetrazolium chloride (TTC) in a 0.5 M solution of sucrose. A total number of 20 MM wild type and 20 *sus2* mutant flowers were employed. Incubation took place for two hours at 50ºC in darkness in a humid box. Results were visualized in an OPTIPHOT-2 (Nikon) optical microscopy.

Scanning Electron Microscopy

Epidermal cell morphology of the four floral whorls of wild type MM and *sus2* mutant plants was assessed by Scanning Electron Microscopy (SEM), following the methodology previously described by Lozano et al., 1998. Briefly, plant material was fixed in a FAEG solution (3.7 % formaldehyde, 5.0 % acetic acid, 50% absolute ethanol and 0.5 % glutaraldehyde) and after 72 hours of incubation was stored in 70 % ethanol. Samples were dehydrated in ethanol increasing concentrations and then dried with liquid CO_2 in a Bal-Tech CPD 030 critical drier. Afterwards, samples were gold coated in a Bal-Tec SCD005 sputter coater and visualised in a Hitachi S-3500N scanning electron microscope.

Genetic mapping of the *sus2* mutation

In order to determine the chromosomic location of the *sus2* mutation, an F_2 interspecific population composed of 129 plants descendant of an F1 plant obtained from a cross between a plant from *S. pimpinellifolium* accession LA1589 and *sus2* mutant plant was characterised. Leaves from F_2 plants were frozen in liquid nitrogen and grounded using a Retsch MM301 mixer mill shaker. Genomic DNA of individual F_2 plants was isolated using DNAzol® Reagent Kit (Invitrogen Life Technologies, USA) and following manufacturer´s instructions. DNA concentration was estimated using a Nanodrop 2000

spectrophotometer (ThermoFisher Scientific, USA) and by comparison with DNA standard markers after electrophoresis. All 129 F_2 plants were individually genotyped and mapping was carried out using codominant markers distributed along the genome (Capel et al., 2015). Genetic linkage and distances were determined using JoinMap® 4 software (Van Ooijen, 2006). A 600 Kb candidate interval was identified in chromosome 2 and fine mapping was completed by a whole genome sequencing approach. Whit this aim, a pool formed by equimolar amounts of DNA of 19 F_2 mutant plants as well as a pool of WT plants was constructed.

Whole genome sequencing analysis

Library generation was carried out using Illumina TruSeq DNA protocol and sequencing was performed in an Illumina HiSeq2000 platform (Illumina, Inc., USA) with 150 pb paired ends. The obtained reads were aligned against the tomato genome reference sequence version 2.5 (ITAG2.4) using Bowtie2 version 2-2.0.0-b5 with default parameters (Danecek et al., 2011). Duplicated reads were eliminated using Picard 1.65 and indel realignment was performed using GATK v2.2-8 under default settings (Li et al., 2011). GATK v2.2-8 was also used along with VCFtools for variant calling (R Development Core Team, 2011). With the aim to determine the allele frequency ratio (i.e. non-reference allele counts / total allele counts) for bi-allelic variants the reference and non-reference allele counts for each position were obtained from SAMtools 1.2 (Li et al., 2009). Finally, the chromosomic candidate region to harbor the *sus2* mutation was detected by plotting the average allele frequencies determined for each chromosome, using a custom script in the R environment for statistical computing (Yuste-Lisbona et al., 2020) and the results obtained were compared to those sequences retrieved from the WT pool.

RNA isolation and gene expression analysis

Total RNA was isolated from flowers harvested at five different developmental stages following the description of Mazzucato et al., 1998, i.e.: flower bud 0 (FB0), flower bud 1 (FB1), flower bud 2 (FB2), flower at pre-anthesis (PA) and flower at anthesis day (AD). A total number of three biological replicates per stage of WT and *sus2* mutant flowers

were collected. RNA isolation was performed using TRIzol® Reagent (Invitrogen Life Technologies, USA) and following manufacturer´s protocol. DNA contamination was eliminated by treatment of the samples with DNA-free™ DNA removal kit (Invitrogen Life Technologies, USA). RNA quantitation was assessed using a Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, USA) and RNA integrity was determined by denaturing agarose gel electrophoresis. First-strand cDNA synthesis was carried out with M-MuLV reverse transcriptase (ThermoFisher Scientific, USA), using a mixture of random hexamer and oligo(dT)$_{18}$ primers. M-MuLV reverse transcriptase was used for first-strand cDNA synthesis (ThermoFisher Scientific, USA), using a mixture of random hexamer and oligo(dT)$_{18}$ primers. Gene expression analysis was conducted by qRT-PCR using the 7300 Real-Time PCR System (Applied Biosystems, ThermoFisher Scientific, USA) and SYBR Green PCR Master Mix (Applied Biosystems, ThermoFisher Scientific, USA). In all qRT-PCR experiments three biological replicates and two technical replicates were used with the primers pairs listed in **Table 4**. Expression data was analysed using the 7300 System Sequence Detection Software v1.2 (Applied Biosystems, ThermoFisher Scientific, USA). The housekeeping gene *UBIQUITIN3* was employed for sample normalization and Ct calculation method was conducted to quantitation of relative gene expression (Livak & Schmittgen, 2001). Finally, differences in gene expression levels were statistically analysed by the least significant difference (LSD) test and a $P < 0.05$ was considered statistically significant.

Table 4: Sequence of primers used for qRT-PCR and SolGenomics (SGN) gene id

Gene Name	Forward and Reverse sequences	SGN gene id
MC	TGAATGGCACCAGCAAACTA CCTCCCATATTTAGGCATTGA	*Solyc05g056620*
TM6	TGCAAGAAAACTTGAAGAGATTG CCTGCAAGTGACATAGTTCCTG	*Solyc02g084630*
TAP3	ACTTACGCCTTCAACCCAAC ATCAGAGCCACCTCCACTGT	*Solyc04g081000*
TPI	AATGGCTCCCATCTCCTTTT GAGGAGGAAAATAAGCAACTTCAA	*Solyc06g059970*
TPIB	GGAAGATGCACTTGAAAATGG TCCCTTGTCTGCTCAAACACT	*Solyc08g067230*
TAG1	CTTGATGCCAGGGAGTTCAT TCGAATTGCTGAGGTGGAG	*Solyc02g071730*
TAGL1	AAAAGAGGGAGATTGAGCTGC CTCTACCTCTGCTATCTTTGCG	*Solyc07g055920*
TAGL11	TGATGCTGAGATTGCTCTCATT CGCACTAGACGTTTCTGCTG	*Solyc11g028020*
SlMPB3	AGAAGGCAACTGCTGAAACC ATCCTTCACCAACCAGATGC	*Solyc06g064840*
TM5	CAACACGAGAAGCACTGGAA CAAGAAGATTCCTCTGTGATCG	*Solyc05g015750*
TM29	GTCAGCAGCAACATCCTCAA CATTCACAGCATCCAACCAG	*Solyc02g089200*

II.3. Results

sus2 mutation impairs flower development

The *sus2* mutant was identified as part of the screening of the EMS mutant collection obtained in *S. lycopersicum* cv. Moneymaker (MM). Mutant plants show no significant alterations of vegetative development (**Figure 16A**), but reproductive structures were severely altered. Floral organs in the second and mainly the third whorl of mutant flowers are affected, since petals are smaller than those of WT plants and stamens are not fused and appear curled (Fig. 16B). Moreover, pollen viability assays performed proved that mutant stamens were unable to produce pollen (Supplementary Fig. 1). Also, all mutant plants produce smaller and parthenocarpic fruits probably due to the absence of viable pollen (Fig. 16C). Genetic analyses of the *sus2* mutation was

performed in M_3 segregating populations. Among them, the mutant phenotype was observed in 64 out of 267 plants (23.97%). The Chi-square statistic test confirmed that the segregation ratios observed were consistent with a monogenic recessive inheritance for the mutant phenotype (χ^2 = 0.69; *P* = 0.15).

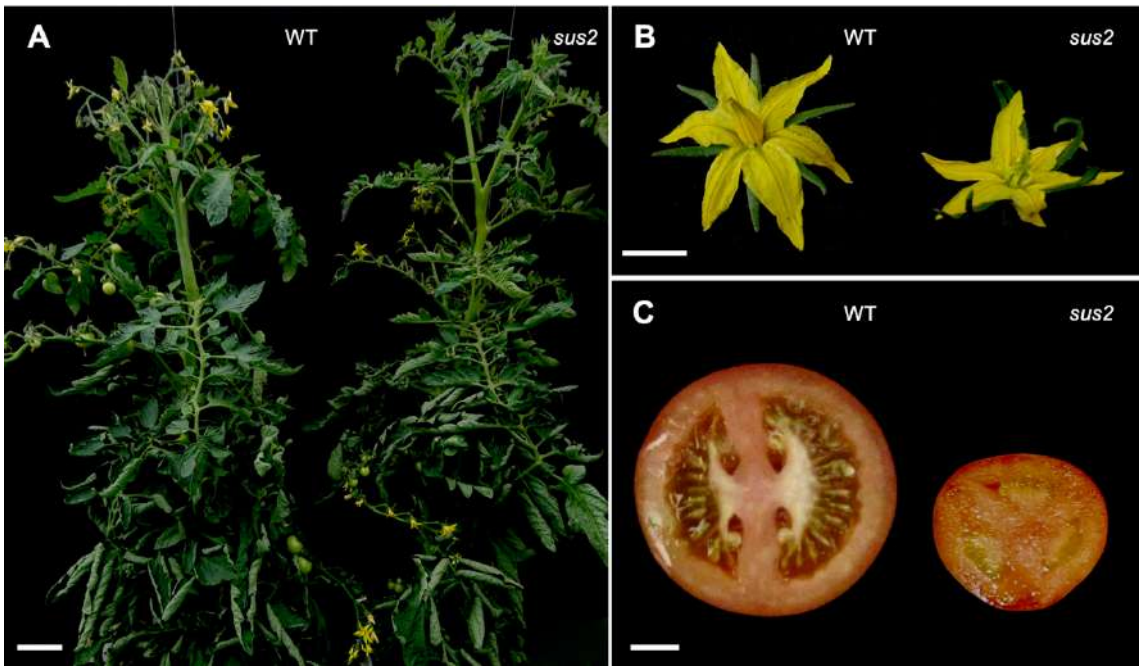


Figure 16: The *sus2* mutant phenotype. A. *sus2* mutant plants show no differences in vegetative growth when compared to WT ones. **B.** *sus2* flowers are smaller and stamens exhibit aberrant morphology. **C.** Mutant fruits are smaller and parthenocarpic (seedless) given to a complete absence of pollen. Scale bar in **A** accounts for 2 cm. The remaining scale bars apply for 1 cm.

On the other hand, stamens of mutant plants are transformed into succulent organs that remain in the fruits during growth and ripening (Figure 17A, B), exhibiting a similar structure to that of carpels when observed in a longitudinal cut (Fig. 17C). These succulent stamens were observed in all mutant plants characterised, although in a variable number of fruits per plant. Given that, a total of 124 M_3 homozygous mutant plants were analysed with the aim to determine the number of succulent stamens developed in 10 fruits of each of these plants. Among these fruits, 54.03 % developed no succulent stamens, 33.06 % developed between 1 to 3 succulent stamens and 12.90

% of the fruits developed 4 to 6 succulent stamens (Fig. 17D). Given that, complete penetrance but variable expressivity can be assumed for this aspect of the *sus2* mutant phenotype.

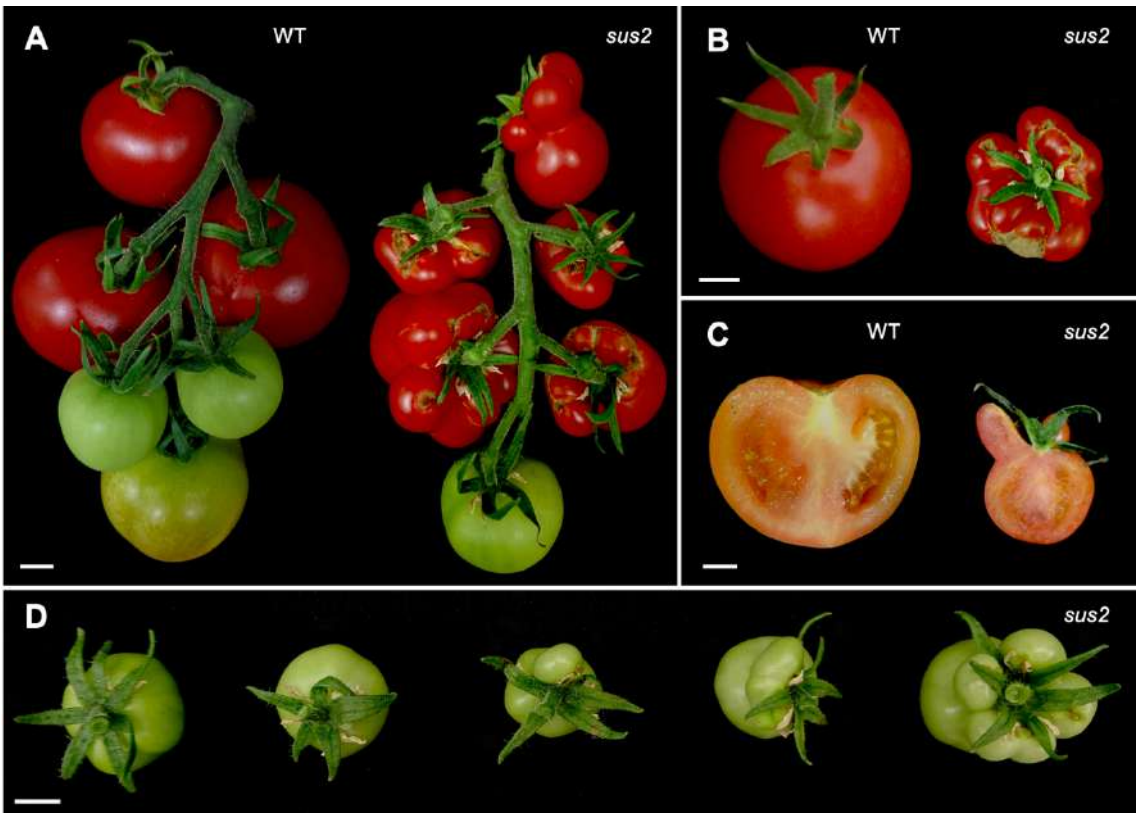


Figure 17: *sus2* mutant fruits develop succulent stamens. A. All mutant plants develop different number of succulent stamens that remain fused to fruits during growth and ripening. **B.** Upper view of a WT and a mutant fruit were all stamens appear attached as succulent organs. **C.** Longitudinal cut of a mutant fruit shows the resemblance between succulent stamens and carpel morphology. **D.** Succulent stamens appear in a variable number, ranging from 0 to all stamens transformed into succulent structures. Scale bars apply for 1 cm.

sus2 is a B class homeotic mutant

In order to elucidate the morphological changes observed in *sus2* mutant plants we performed Scanning Electron Microscopy (SEM) analysis of identity of the epidermal cells present in the four whorls of anthesis flowers from WT and mutant plants. No changes in cell identity were observed in cells of the first and fourth whorls. However,

epidermal cells of *sus2* petals showed changes in cell identity reflected in differences in size and shape (**Figure 18**).

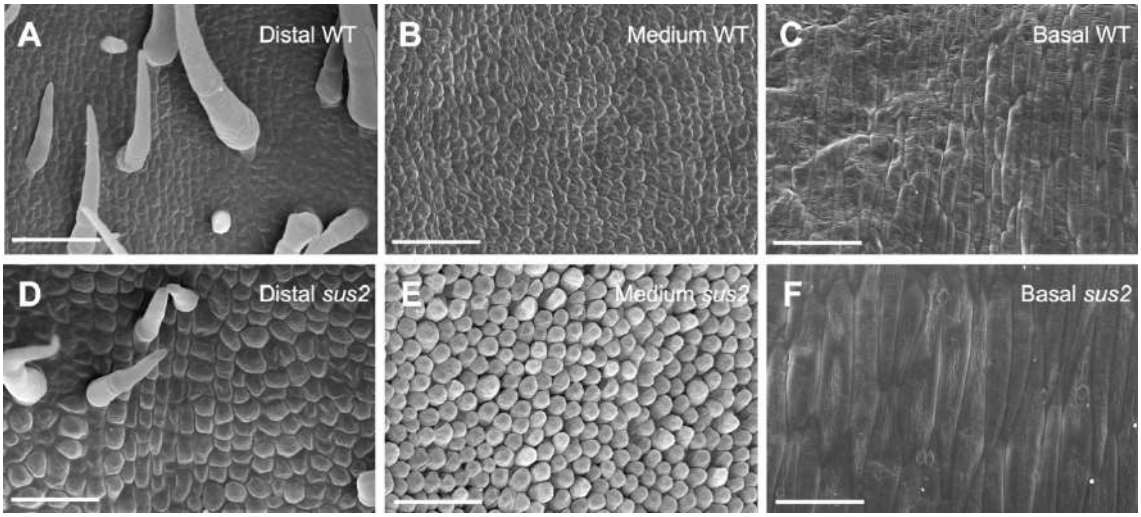


Figure 18: Cell morphology observed in the ventral side of petals of WT and *sus2* plants by SEM. Distal (**A**), medium (**B**) and basal (**C**) section of ventral side of WT petals. Cell morphology and identity is altered in all the studied sections (**D-F**) of ventral side of *sus2* petals. All scale bars apply for 50 µm.

Even more striking is the morphology observed in epidermal cells of stamens, where SEM revealed a complete homeotic change to carpeloid cells in all the analysed portions (**Figure 19**). Cells in the distal portion of stamens have a papillae morphology which is similar to WT stigma (Fig. 19A-C). Middle section cells are elongated instead of puzzled rounded and are similar to those of WT style (Fig. 19D-F). Finally, cells in the basal section of mutant stamen are smaller and irregular when compared to WT ones and resembled those of WT ovaries (Fig.19G-I). Altogether, the homeotic changes observed in *sus2* suggest that it can be considered as a B-class mutant.

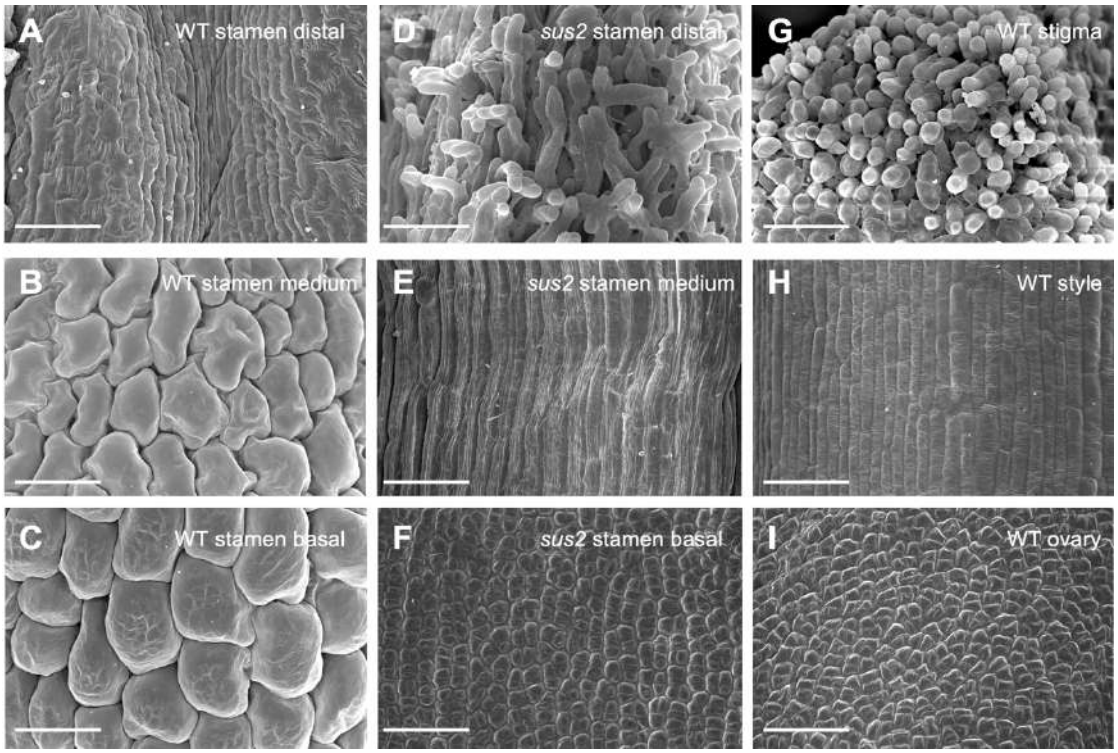


Figure 19: Scanning electron microscopy (SEM) of anthesis flowers of WT and *sus2* mutant plants. A-C. Cell morphology of WT stamens assessed in distal, medium and basal sections. **D-F.** Stamens of *sus2* mutant plants. **G-I.** Cell morphology of WT carpels, with details of stigma, style and ovary. *sus2* stamens morphology resembles that of WT carpels in all studied sections. All scale bars apply for 50 µm.

Cloning and molecular characterisation of *SUS2*

With the aim to identify the gene that underlies the *sus2* mutation, mapping strategies were carried out. An F_2 segregating population derived from the cross of a mutant plant with a plant from the wild relative *S. pimpinellifolium* accession LA1589 was obtained. F_1 plant was self-pollinated and F_2 plants were characterised. Among these, 27 mutant plants were detected out of 129 F_2 plants, which confirmed the monogenic recessive inheritance (χ^2 = 1.13; *P* = 0.28) of the *sus2* mutant phenotype in this interspecific segregating population. Mapping was performed by genotyping all F_2 plants with codominant markers distributed along the genome (Capel et al., 2015). This strategy allowed us to identify a genomic region of 600 Kb that harbor the mutation located in chromosome 2 (**Figure 20A**). Fine mapping was completed by a whole genome

sequencing approach carried out by sequencing a pool formed by equimolar amounts of DNA of 19 F_2 mutant plants as well as sequencing of a WT DNA pool. Allele frequencies confirmed the location of *sus2* mutation in chromosome 2 (Supplementary Fig. 2) Afterwards, variant analysis of this interval encompassing the candidate region allowed us to identify a unique mutation, a single base deletion in an exon of *Solyc02g084630* syn. *Tomato MADS box gene 6* (*TM6*) (Fig. 20B). *TM6* is gene that belongs to a paralogous lineage of Arabidopsis *AP3* that originated as a result of a duplication event in *AP3* lineage and proposed to be implicated in stamen development (de Martino et al., 2006). The mutation detected in *sus2* causes a frameshift in *TM6* and as a result, mutant protein lacks all functional domains, including the MADS domain (Fig. 20C). Since *sus2* is a B-class homeotic mutant, *TM6* is a strong candidate to be responsible for this mutant phenotype. Recently, *TM6* tomato mutants with malformed stamens and male sterility have been described (Cao et al., 2019). However, given that *sus2* alterations are extended to the second and third whorls, *TM6* can be considered a true homeotic B-class gene.

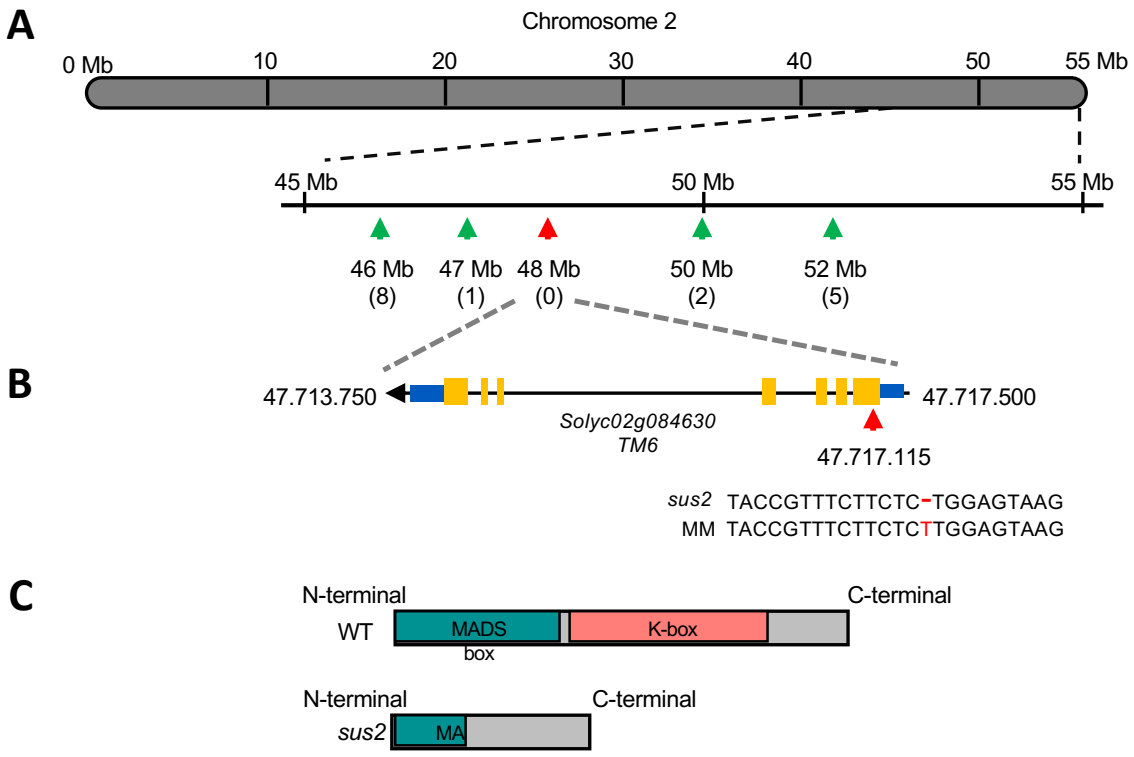


Figure 20: Cloning and molecular characterisation of *SUS2.* A. Mapping of an F_2 segregating population using codominant markers identified a candidate region in chromosome 2. Numbers in parenthesis indicate the number of recombinant chromosomes identified between the *SUS2* gene and each genetic marker analysed. **B.** Fine mapping of this interval was carried out by a mapping sequencing approach and variant analysis identified a single-base deletion in the first exon of *Solyc02g084630* also named *TM6.* **C.** As a result of the mutation detected, TM6 mutant protein lacks functional domains including the MADS-box.

ABC genes expression analysis

In order to assess whether if mutations detected in *TM6* also induced changes in the expression of other genes involved in flower development and whorl identity, a comparative qRT-PCR analysis was carried out at five stages of flower development, i.e. FB0 (flower bud 0), FB1 (flower bud 1), FB2 (flower bud 2), PA (flower at pre-anthesis stage) and AD (flower at anthesis day stage). Expression pattern of *TM6* as well as from 11 additional genes previously described as key regulators of tomato flower development was assessed.

The analysed genes are orthologues of those described by Coen & Meyerowitz (1991) in *A. majus* and *A. thaliana* as part of the ABC model. Thus, we determined the relative expression level of the A-class gene *MACROCALYX* (*MC*) (Yuste-Lisbona et al., 2016), the B-class MADS-box genes orthologues of *PISTILLATA* (*PI*), i.e. *TPI* (*TOMATO PISTILLATA*) and *TPIB* (Mazzucato et al., 1998) as well as the *APETALA3* orthologues *TOMATO APETAL3* (*TAP3* syn. *STAMENLESS*) (Kramer et al., 1998) and the own *TM6* gene (Pnueli et al., 1991). Relative expression of C-class gene *TOMATO AGAMOUS1* (*TAG1*) (Pnueli et al., 1994) and *TOMATO AGAMOUS LIKE1* (*TAGL1*) (Gimenez et al., 2016), as well as E-class genes *TOMATO MADS BOX 5* (*TM5*) and *TOMATO MADS BOX 29* (*TM29*) (Pnueli et al., 1994; Ampomah-Dwamena et al., 2002) and the of D-class genes *TOMATO AGAMOUS LIKE11* (*TAGL11*) and *SLMBP3* (Daminato et al., 2014) was also analysed.

Changes detected in gene relative expression trough all developmental stages analysed are represented in **Figure 21**. The A-class gene *MC* is induced in *sus2* flower bud at stage 0 (FB0) and pre-anthesis (PA), but other differences in gene expression were observed. Meanwhile, relative expression of other B-class genes such as *TPI* and the *TM6*

paralogous *SL-TAP3* remains unaltered between WT and *sus2* mutant plants. The only exception is *TPI* paralogous *TPIB*, which is induced in FB1 (flower bud 1), PA and AD (anthesis day) in *sus2* mutant plants. When it comes to the rest of gene functions, C-class gene *TAGL1* is also induced in all analysed stages except for AD while *TAG1* is induced in FB2. Finally, for the E-class genes *TM5* and *TM29*, significant differences were accounted only in the earliest developmental stage of *sus2* flowers (FB0), where *TM29* is induced. Gene expression of D-class genes *TOMATO AGAMOUS LIKE11* (*TAGL11*) and *SLMBP3* could not be assessed at these early flower developmental stages, since these genes seems to be induced during early stages of fruit development (Busi et al., 2003; Zhang et al., 2019). Regarding the own *TM6* gene, results obtained demonstrated that is downregulated in *sus2* flowers in almost all the analysed stages, except for pre-anthesis (PA), suggesting that *TM6* regulates its own expression as well as the other MADS-box genes in different stage of flower development.

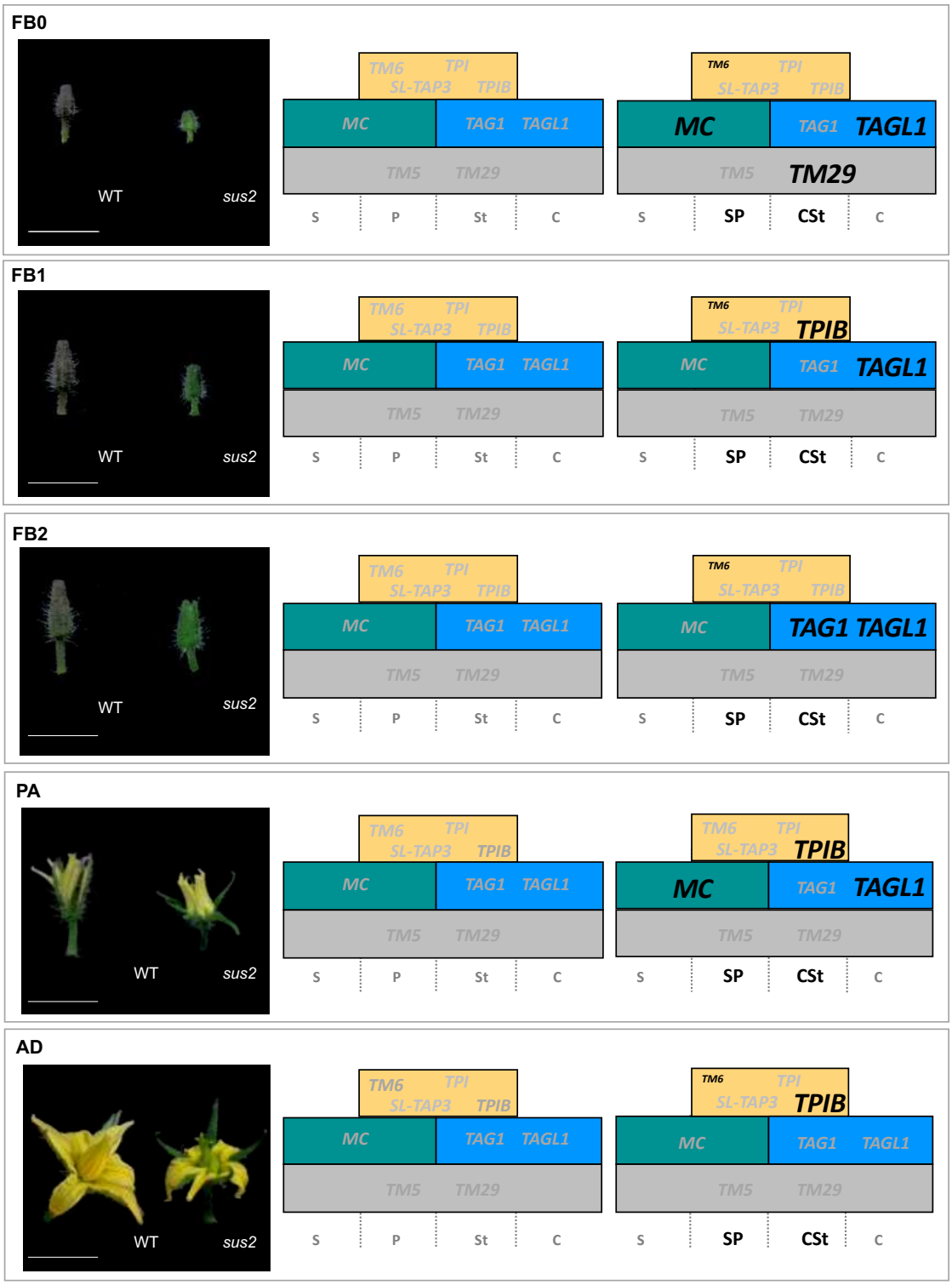


Figure 21: Changes in relative expression of genes related with flower development and whorl identity in tomato trough five developmental stages. Flower of each developmental stage analysed is showed at the left, while relative expression of each gene is schematically

represented for WT (centre) and *sus2* mutant plants (right). Black bold letters account for induced genes, while downregulated ones are in a smaller font size. S, sepals; P, petals; St, stamens; C, carpels; SP, sepaloid petals; CSt, carpeloid stamens. Scale bars apply for 1 cm.

II.4. Discussion

Over the past decades, great progress has been made in understanding the genetic network that controls flower development based on the studies performed in the model species *A. thaliana* and *Antirrhinum majus*. The genes underlaying this crucial developmental process are described in the ABCDE model (Coen & Meyerowitz, 1991; Colombo et al., 2010; Pelaz et al., 2000), which is widely conserved among Angiosperms. Extension of this model to other species of agronomic interest such as tomato has been possible due to the characterisation of homeotic mutants, most of them of spontaneous origin. This approach along with the completion of the tomato genome sequencing have allowed to describe the genes accounting for each of the ABCDE gene functions, proving that the establishment of floral organs is controlled by highly conserved molecular mechanisms. However, for some of the genes predicted in the model there are no described tomato mutants, which demonstrates the need of mutagenesis programs in this crop species.

Here, a tomato mutant identified in a chemically mutagenised population and characterised by aberrant petal and stamen morphology as well as male sterility has been described. Mutations in the predicted B-class gene *TM6*, seems to account for this mutant phenotype as proved by mapping by sequencing. In tomato as in other species of the core eudicots clade, major duplication events have given rise to two *AP3* paralogous lineages, the *SL-TAP3* and the *TM6* lineage, both distinguishable by their divergent C-terminal motifs (Kramer et al., 1998). Furthermore, as part of the functional characterisation of tomato *AP3* paralogous genes, *TM6* downregulation by RNAi silencing proved to result in aberrant stamen development and no petal homeotic changes (de Martino et al., 2006), on the contrary to the observed in the *sl-tap3* mutant, where homeotic changes affect petals as well (Gomez et al., 1999). The *AP3* lineage has also been studied in the Solanaceae species *Petunia hybrida*, where downregulation of *AP3* orthologue *DEFICIENS* (*DEF*) causes homeotic changes from petals to sepals in the

second whorl whereas stamens identity remains unaltered (van der Krol & Chua, 1993). Although Petunia *TM6* has not been related with petal determination, petal defects observed in the *def* mutant can be restored by complementation with an overexpression construct of *PhTM6* (Rijpkema et al., 2006), which confirms the implication of *TM6* in petal determination. A probable explanation for the absence of homeotic changes observed in petals of *TM6* RNAi silencing lines by de Martino et al., (2006) is that these lines may retain some residual expression levels of this gene. Recently, it has been described that *TM6* is a candidate gene to *male sterile-15* locus, since *ms-15* mutants bear mutated alleles of *TM6* (Cao et al., 2019). *ms-15* mutant plants develop flowers with reduced and deformed anthers and, as a consequence, the stigmas are exerted an easily hand pollinated. Anthesis-day *ms-15* flowers also show alterations in the expression pattern of MADS-box genes as well as other genes involved in pollen ontogeny (Cao et al., 2019). Nevertheless, neither changes in petal morphology nor stamens conversion to carpels have been addressed for *ms-15* mutant plants. Since *sus2* mutant plants exhibit a strong downregulation of *TM6* and changes in petal epidermal cell morphology have been accounted, *TM6* can be considered without a doubt as a homeotic B-class gene linked to petal and stamen identity determination. Taken together, all these results provide new evidence to the role of *TM6* lineage in higher eudicots like tomato, questioning the functional divergence hypothesis previously assumed for the two *AP3* paralogous lineages (Kramer et al., 1998; de Martino et al., 2006).

On the other hand, our results suggest that the B-class paralogous genes *TM6* and *SL-TAP3* do not regulate each other, since *SL-TAP3* expression remains unaltered in all flower stages analysed in *sus2* mutant plants. In this way, it has been described that *TM6* and *TPI* relative expression in *sl-tap* loss of function mutants is similar to that of WT plants, and that *TM6* downregulation do not alters *SL-TAP3* and *TPI* expression (de Martino et al., 2006). Moreover, Quinet et al. (2014) demonstrated that *SL-TAP3* downregulation do not alters *TM6* or *TPI* expression. Furthermore, downregulation of the *PI* gene in *P. hybrida* do not affects *TM6* expression (Vandenbussche et al., 2004). All these results provide strong evidence for the assumption that no changes in B-class

gene expression are expected from the downregulation of their paralogous genes. Nevertheless, our findings suggest that while *SL-TAP3* and *TPI* expression do not change in *sus2* mutant plants when compared to WT ones, full downregulation of *TM6* does results in an induction of *TPIB* in at least three out of the five analysed flower developmental stages. Interestingly, expression of *TPI* and *TPIB* has been proved to be absent in second and third whorls of *sl-tap3* knock-down mutants, whereas *SL-TAP3* expression increases in *TPIB* RNAi plants (Geuten & Irish, 2010). Given that our observations suggest that downregulation of *TM6* induces *TPIB* expression, further experiments should be achieved focused on B-class genes interaction in different loss of function mutants, with special attention to the mutual regulation among *TPIB* with *TM6* and *SL-TAP3*.

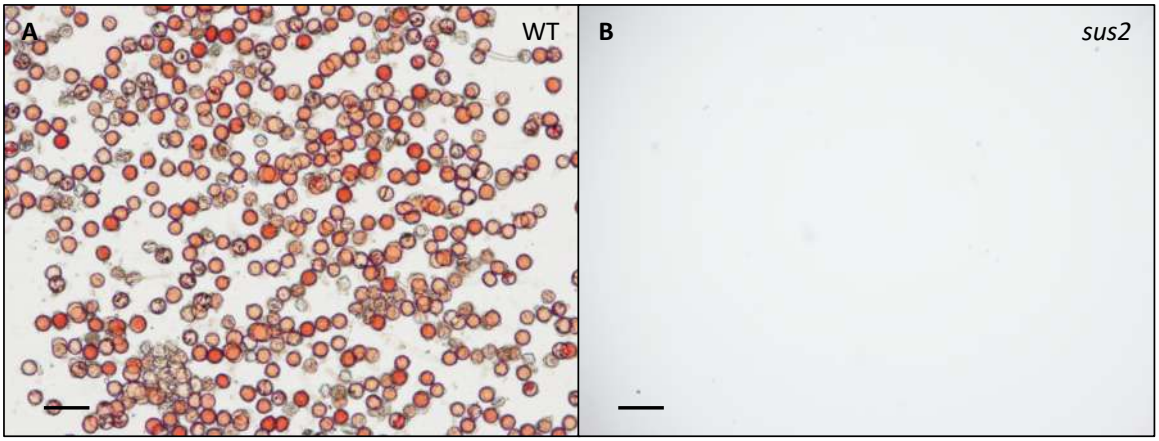
We have also observed that downregulation of *TM6* in *sus2* mutant plants increases the expression of the A-class gene *MC*, which is a similar effect to that observed after downregulation of *SL-TAP3* (Quinet et al., 2014). In *A. thaliana* during the earlier flower developmental stages, *AP1* expression can be found throughout the floral meristem, but it gets soon restricted to the sepal and petal primordia as a result of negative regulation by the complex *PI*/*AP3*. In Arabidopsis *ap3* knock-down mutants this protein complex cannot be formed and expression of *AP1,* the *MC* orthologue, is induced (Sundström et al., 2006). A similar mechanism has been suggested to explain induction of *MC* in *SL-TAP3* mutants (Quinet et al., 2014), but this putative regulation requires further investigation and the possible formation of protein complexes involving *TM6* and *PI* must be deeply assessed.

Finally, C- and E-class genes expression is similar among WT and *sus2* mutant plants except for *TAGL1*, which is induced in four out of the five analysed stages. C-class genes *TAG1* and *TAGL1* belong to the euAG and PLE lineages that arose from duplication of the tomato *AG* clade (Kramer et al., 2004). Both genes are reported to have similar expression patterns during flower development given that their transcripts preferentially accumulate in stamens and carpels (Pnueli et al., 1994; Gimenez et al., 2010). This is not the case of *sus2* mutant plants, where downregulation of *TM6* induces *TAGL1* expression while *TAG1* transcripts levels remains unaltered, indicating functional

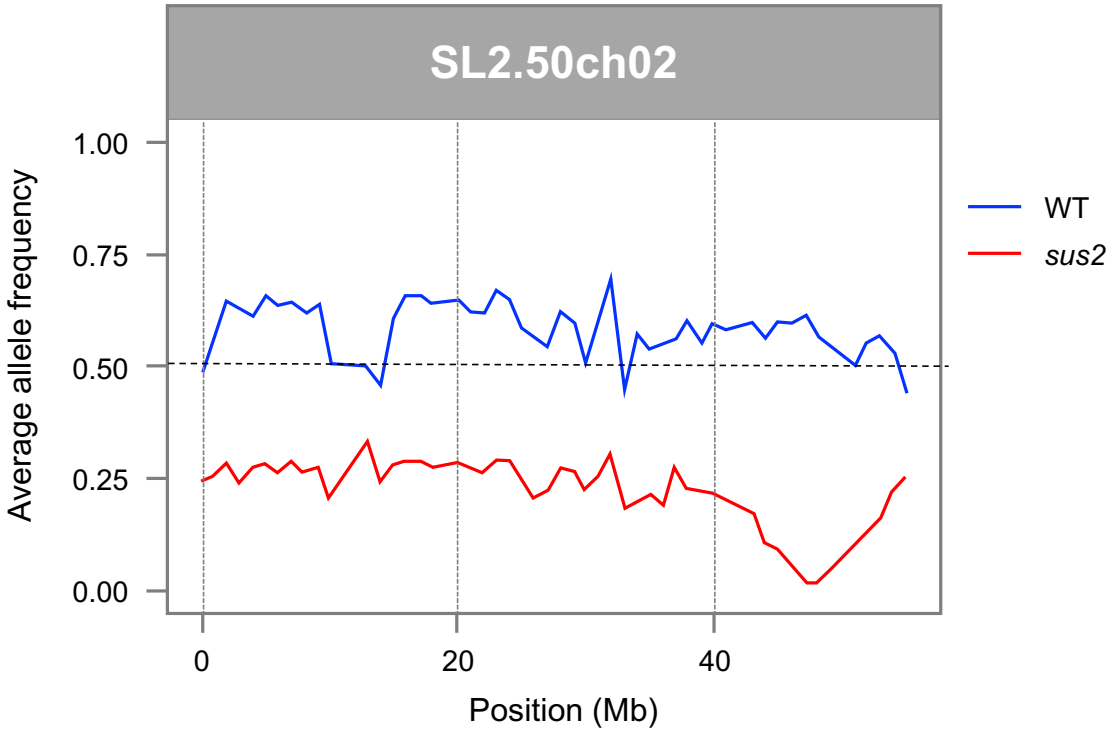
divergence among both genes regarding the implication of TM6 in their regulation. In fact, whereas *TAG1* is involved in stamen and carpel identity specification (Pnueli et al., 1994), *TAGL1* function is not restricted to flower development but it extends to fruit ripening (Gimenez et al., 2016). Nevertheless, some partial redundancy when it comes to floral whorl identity is also evident, since overexpression lines of *TAGL1* exhibit homeotic changes from sepals to carpels and from petals to stamens (Vrebalov et al., 2009; Gimenez et al., 2010) in a similar way to *TAG1* overexpression lines (Pnueli et al., 1994). Given the effect of *TM6* downregulation in expression levels of C-class *TAGL1* as well as in the A-class gene *MC*, it could be assumed that B-class genes and particularly *TM6* are crucial for balance maintenance between the other different homeotic genes function.

Together, the results reported here contribute to shed light on the regulatory interactions of MADS-box genes in flower development as well as to reinforce the evidence of regulatory function of *TM6* during petal and stamen morphogenesis. Although great progress has been achieved in understanding flower development and MADS-box transcription factors implication in this process, further research in protein interactions and CHIP-seq analysis should be carried out in the future with the aim to fully dissect the regulatory interactions of *TM6* with other key players in floral organ identity.

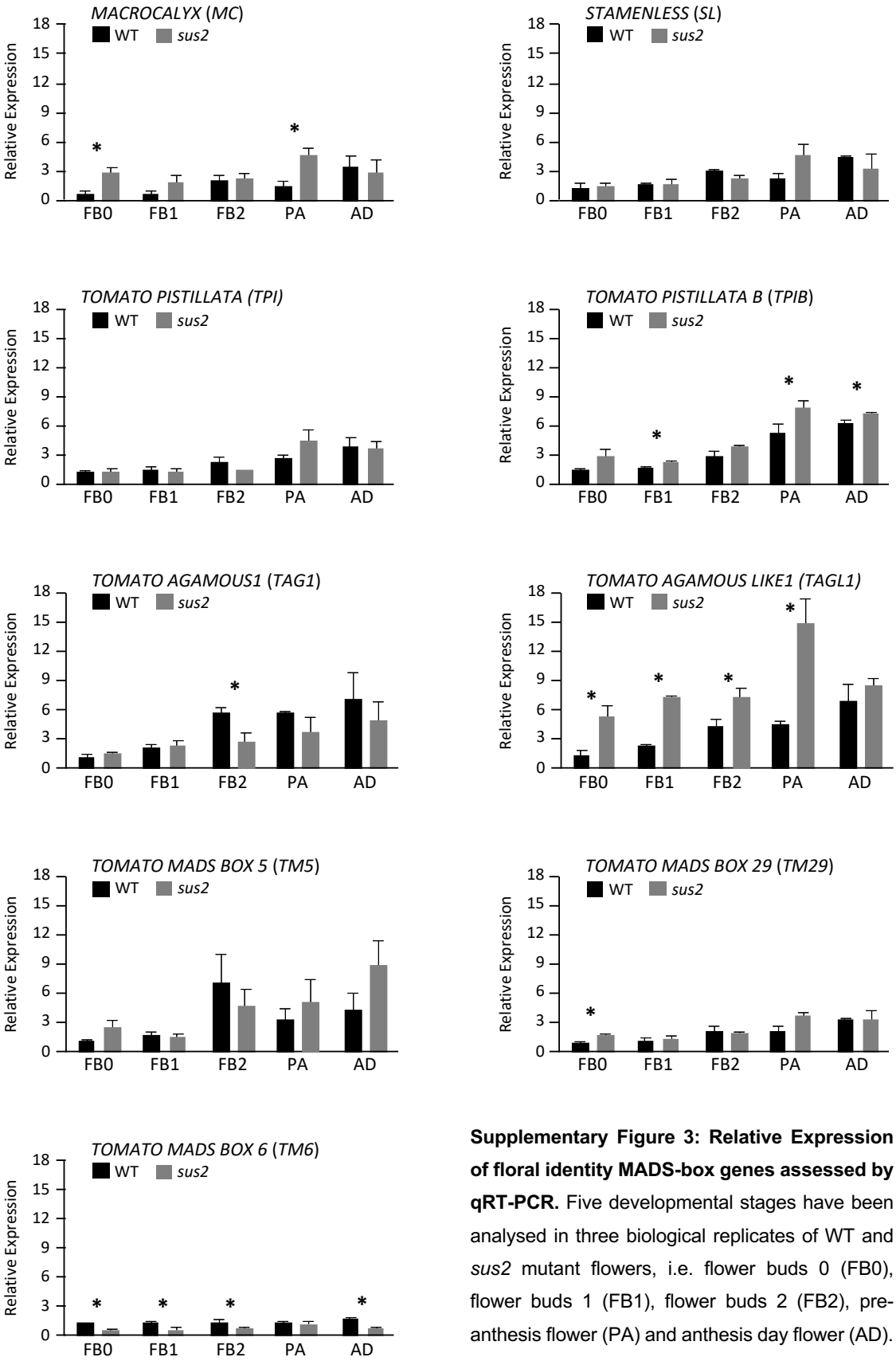
Anexo I. Supplementary figures



Supplementary Figure 1: Pollen viability assays performed with tetrazolium chloride. WT plants produce a large number of viable stained pollen (**A**), on the contrary of *sus2* mutant plants (**B**), where no pollen grains are observed. Scale bars correspond to 50 µm.



Supplementary Figure 2: Fine mapping performed by whole genome sequencing. Mapping was performed in pools of equimolar amounts of DNA of F_2 mutant and WT plants. Allele frequencies analysis confirmed that *sus2* mutation is located in chromosome 2.



Supplementary Figure 3: Relative Expression of floral identity MADS-box genes assessed by qRT-PCR. Five developmental stages have been analysed in three biological replicates of WT and *sus2* mutant flowers, i.e. flower buds 0 (FB0), flower buds 1 (FB1), flower buds 2 (FB2), pre-anthesis flower (PA) and anthesis day flower (AD).

Capítulo III: Functional diversification of HAIRPLUS, a Histone N-Lysine Methyl Transferase, leads to trichome development in tomato

Abstract

Trichomes are specialised epidermal cells that cover the aerial surface of almost all terrestrial plants. These structures form physical barriers that prevent plagues from spreading which combined with their capability of synthesis of complex molecules confer trichomes a key role in defence against herbivores. Here we have identified by a map-based strategy that *HAP* a histone methyl transferase promotes glandular type I trichome initiation in stems of tomato plants. Hap belongs to a group of proteins involved in histone tail modifications and some members also bind methylated DNA. Furthermore, transcriptome and epigenome modifications were accounted in *hap* mutant plants, allowing to detect large amounts of differentially expressed genes (DEG's) and differentially methylated cytosines (DMC's). Taken together, all our findings suggest that *HAP* links transcriptome remodelling and multicellular type I trichome development.

Capítulo III

III.1. Background

Plants exhibit a large number of adaptive traits that allow them to succeed in their environments and among these the formation of epidermal protrusions called trichomes are one of the most recurrent. Trichomes are epidermal structures that can be classified as uni- or multicellular based on their cell number and as glandular or non-glandular depending on their secretory capability (Chang et al., 2016). Trichomes are implicated in abiotic stress response, including extreme temperatures or excessive radiation, and mainly involved in biotic stress response including pest resistance (Martin & Glover, 2007; Kang et al., 2010).

Trichomes in the model species *Arabidopsis thaliana* are unicellular and non-glandular (Hülskamp et al., 1999; Schwab et al., 2000), and the genetic network controlling their initiation is well known. The first gene is this network is *GLABRA1* (*GL1*), which encodes a R2R3 MYB transcription factor whose loss of function mutation results in a complete absence of trichomes (Marks & Feldmann, 1989). The MYB-domain of GL1 physically interacts with the N-terminal domain of the proteins encoded by *GLABRA3* (*GL3*) and *ENHANCER OF GLABRA3* (*EGL3*) (Zhang et al, 2003; Serna & Martin, 2006). GL3 and EGL3 are both bHLH proteins that promote trichome formation in a redundant manner (Zhang et al, 2003). The fourth gene, *TRANSPARENT TESTA GLABRA 1* (*TTG1*), encodes a protein with four conserved WD-domains that interacts with both GL3 and EGL3 and all together with GL1 form a multimeric complex that triggers trichome initiation in *A. thaliana*. *GLABRA* (*GL2*), a homeodomain-leucine zipper protein seems to be the downstream target of this molecular complex, since it contains a MYB binding site to which *GL1*, as part of the multimeric complex, can bind (Serna & Martin, 2006). On the other hand, four MYB proteins encoded by *CAPRICE* (*CPC*), *TRIPTHYCHON* (*TRY*), *ENHACER OF TRY AND CPC1* (*ETC1*) and *ENHACER OF TRY AND CPC2* (*ETC2*), repress trichome initiation in a redundant manner by interacting themselves with GL3 (Esch et al., 2003). As a result of this interaction, GL1 cannot bind to GL3 and the activation complex is not formed (Serna & Martin, 2006).

Whether there is a similar complex controlling trichome formation in Asterid species and particularly within the Solanaceae remains unclear. However, in Solanaceae the

MIXTA gene, that codes for a MYB-related transcription factor, seems to hold a key role in trichome initiation. This gene is well known for inducing ectopic trichome formation on the anthers of the woody nightshade *Solanum dulcamara* (Glover et al., 2004). In *Nicotiana tabacum, MIXTA* promotes multicellular trichome and conical cells formation, but ectopic expression of *MIXTA* in *A. thaliana* has no effect in trichome formation, as well as overexpression of *GL1* in tobacco (Payne et al., 1999). All these findings suggest that trichome initiation in *A. thaliana* and Solanaceae species could follow divergent regulatory pathways (Yang et al., 2011).

Five types of trichomes have been described in tomato (*Solanum lycopersicum* L.) (Luckwill, 1943): type I are 2-3 mm long, have a multicellular base and a glandular cell in the tip; type III are non-glandular and 0.4-1 mm long with a unicellular flat base; type V also have unicellular and flat base like type III but are 0.2-0.4 mm long; type VI are short trichomes with two stalk cells and a glandular head composed of four secretory cells and type VII are very small trichomes (0.05 mm) with glandular heads formed by 4 to 8 cells. All non-glandular trichomes act as a mechanical barrier that prevents pest movement along the plant whereas glandular ones produce specialized metabolites so called exudates that entrap and sometimes are toxic or deterrent to the pest (Simmons & Gurr, 2005).

Despite its importance in pest resistance, very few trichome mutants have been described in tomato that allow to shed light on the genetic regulation of multicellular trichome initiation. One of those mutants is *hairless*, characterised by a distorted morphology of type I trichomes and a deficient accumulation of sesquiterpenes in type VI trichomes (Kang et al., 2016). The *hair absent* mutant, which exhibits a complete absence of type I trichomes on the epidermis, has also been characterised and encodes a C2H2 zinc-finger protein named *Hair* (Chang et al., 2018). On the other hand, the *woolly* mutant is characterised by an increase in type I trichome density and embryo lethality (Huang & Paddock, 1962; Yang et al., 2011). Fine mapping of the *woolly* mutation located in chromosome 2 showed that *Wo* encodes an HD-Zip protein containing a START-domain that physically interacts with a B type cyclin named *SlCycB2* (Yang et al., 2011). These findings demonstrate the existence in tomato of protein

complexes of very different nature to the protein complex that regulates trichome initiation in *Arabidopsis*.

In this work we have characterised a tomato mutant that exhibits a high type I trichome density which we have named *hairplus* (*hap*). Fine mapping and genetic complementation analysis showed that *HAP* encodes a SuvH3 Histone N-Lysine Methyl Transferase. This protein family is mainly composed of histone tail modifying proteins, but some members bind methylated DNA, playing all key roles in epigenetic control of gene expression (Marmorstein, 2003; Lachner & Jenuwein, 2002). Our findings not only identify a new regulator of trichome density in tomato but also shed light on the role of SuvH3 Histone Methyl Transferases in transcriptional control mediated by epigenetic modification.

III.2. Methods

Plant material and growth conditions

The *hairplus* mutant was identified as part of a mutant collection obtained by chemical mutagenesis with ethylmethanesulfonate (EMS) in *S. lycopersicum* cultivar Moneymaker (MM). Seeds were incubated in 250 ml of a solution containing 0.7% of EMS (Sigma-Aldrich) at 30ºC with gentle shaking for 16 hours. Afterwards, seeds were thoroughly washed in distilled water, sown in seedling beds and germination was assessed prior to transplant. M_2 seeds were obtained by self-pollination of M_1 plants under greenhouse growing conditions. Wild relative *S. pimpinellifolium* accession LA1589 obtained from Tomato Genetics Resource Center (http://tgrc.ucdavis.edu/), was used for generating F_2 mapping population. Phenotypic characterisation of *hap* mutant was carried out in M_2 populations grew along with control MM plants. All experiments were conducted under greenhouse conditions and following standard management practices, including regular fertilization.

Scanning Electron Microscopy

Trichomes density and morphology was assessed in stems, inflorescences and leaves of the *hairplus* mutant and MM wild type plants. Five mutant and five control samples of all these tissues were analysed by Scanning Electron Microscopy conducted as

previously described (Lozano et al., 1998). Briefly, samples were fixed in FAEG (3.7% formaldehyde, 5% acetic acid, 50% absolute ethanol and 0.5% glutaraldehyde) and after 72 hours of incubation were stored in 70% ethanol. Prior to analyses, samples were dehydrated in increasing ethanol concentrations and completely dried by CO_2 critical point with a Bal-Tec CPD 030 CO_2 critical point dryer. Finally, samples were gold coated in a Bal-Tec SCD005 sputter coater and visualised in a Hitachi S-3500N scanning electron microscope. Trichome counts were performed on the images obtained. Type I, III, V, VI and VII as well as stomata were counted. Data are presented as the mean ± standard deviation in boxplot graphs created using the online program BoxPlotR (http://shiny.chemgrid.org/boxplotr/). Variance analyses was carried out using the least significance difference (LSD) test of Fisher for mean comparison. A probability of $P < 0,05$ was considered statistically significant.

Genetic mapping of the *hairplus* mutation

With the aim to determine the chromosomic location of the *hap* mutation, a total number of 242 F_2 plants derived from a cross between *S. pimpinellifolium* accession LA1589 and *hap* mutant were individually genotyped using codominant markers distributed along the genome (Capel et al., 2015). Leaves from parents and F_2 plants where frozen in liquid nitrogen and grounded using a Retsch MM301 mixer mill shaker. Genomic DNA was isolated from approximately 300 mg of powdered leaf tissue using DNAzol® Reagent Kit (Invitrogen Life Technologies, USA) and following manufacturer´s instructions. DNA concentration was estimated by DNA quantitation using a Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, USA) and by comparison with DNA standards after electrophoresis. Genetic linkage and distances were determined using JoinMap® 4 software (Van Ooijen, 2006).

Whole genome sequencing and allele frequency analysis

DNA for whole genome sequencing was isolated using DNAzol® Reagent Kit as previously described. DNA isolation was performed individually in 14 F_2 mutant plants, then a pool was constructed using equal amounts of DNA from each plant. Illumina TruSeq DNA

protocol was used for library generation and sequencing was performed in an Illumina HiSeq2000 platform (Illumina, Inc., USA) with 150 pb paired ends. Alignment of the obtained reads to the tomato genome reference sequence version 2.5 (ITAG2.4) was performed using Bowtie2 version 2-2.0.0-b5 with default parameters (Langmead & Salzberg, 2012). Picard version 1.65 was used to remove duplicated reads, whereas indels were realigned using GATK v2.2-8 under default settings (DePristo et al., 2011). Also, variant calling was performed using GATK v2.2-8 and VCFtools (Danecek et al., 2011) was used for variant filtering according to the following parameters: --min-alleles 2 --max-alleles 2 --min-meanDP 20 --max-meanDP 40. Next, pileup from SAMtools 1.2 (Li et al., 2009) was used for obtaining reference and non-reference allele counts for each position with the aim to determine the allele frequency ratio (i.e. non-reference allele counts / total allele counts) for bi-allelic variants. At last, determination of the chromosomic region where the *hap* mutation is located was performed by plotting the average allele frequencies determined for each chromosome using a custom script in the R environment for statistical computing (R Development Core Team, 2011) that uses a sliding window and step size of 1000 and 100 variants, respectively.

Transgenic analysis

A RNAi approach was used to obtain *HAP* (*Solyc10g077070*) silencing lines. A 185 pb sequence was amplified using the primers listed in Supplementary Table 2. The PCR product was cloned in sense and antisense orientation into de pKannibal vector and then digested with *NotI* restriction enzyme and cloned into the binary vector pART27 (Gleave, 1992; Campos et al., 2016). Also, full-length coding sequence of *HAP* from MM wild type plants (amplified with primers listed in **Table 5**) was cloned into the pGWB402 vector (ThermoFisher Scientific, USA) driven by the CaMV 35S promoter for generation of overexpression lines. CRISPR-Cas9 lines were also obtained following the protocol described by Vazquez-Vilar et al. (2016). sgRNA target sequence was designed within the coding sequence of *HAP* using the Breaking-Cas software (Oliveros et al., 2016) and the sgRNA selected was the one with the lowest off-target score. The construct contained the *A. thaliana* U626 promoter for constitutive expression of sgRNA, as well

as the transcriptional unit for Cas9 expression in plants under the control of the CaMV 35S promoter. The kanamycin resistance gene was used as a selective marker. Primers covering the SgRNA recognition site (**Table 5**) were designed and T_0 lines were genotyped. PCR products were cloned into the pGEMT easy vector (Promega, USA) and eight clones per each PCR product were sequenced for further allele characterisation. Also, quantitative assessment of genome editing was performed using the program Tide (https://tide.deskgen.com/). Finally, genetic transformation of all constructs was performed using *Agrobacterium tumefaciens* strain LBA4404 (Gelvin, 2003) and T_0 ploidy analysis was determined by flow cytometry (Atarés et al., 2011).

Table 5: Primers used for expression analysis and functional characterisation of *HAP*

Name	Primer
RNAi::Hap	F: tctagactcgagGCACTAGTCTGATGCATAAAAAGAA
	R: atcgatggtaccAAGTACTTAGGATTGTCATTGACTGG
35S::Hap	F: ggtaccGGGAGAAGTGCTGACCTTTG
	R: gtcgacATGTCCAAAACCTTGCTGCT
sgRNA-Hap	F: ATTGAGCATGCATCGTGGCAATG
	R: AAACCATTGCCACGATGCATGCT
Sequencing sgRNA-Hap	F: TTGTGCCTGGTTGGTTTACA
	R: CAGGTTGACTCTCTGCACCA
qRT-PCR Hap	F: TTATACTGGCCAGGGTGGAG
	R: CAATTGCCCAATTTGCTTTT

RNA isolation and gene expression analysis

Total RNA was isolated from target tissues with TRIzol® Reagent (Invitrogen Life Technologies, USA) and following manufacturer´s protocol. DNA contamination was avoided by treatment with DNA-free™ DNA removal kit (Invitrogen Life Technologies, USA). RNA integrity was determined by denaturing agarose gel electrophoresis and quantitation was performed using a Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, USA). First-strand cDNA synthesis was carried out with M-MuLV reverse transcriptase (ThermoFisher Scientific, USA), using a mixture of random hexamer and oligo$(dT)_{18}$ primers. Gene expression analysis was conducted by qRT-PCR using the 7300 Real-Time PCR System (Applied Biosystems, ThermoFisher Scientific, USA) and SYBR

Green PCR Master Mix (Applied Biosystems, ThermoFisher Scientific, USA). In all experiments, three biological replicates and two technical replicates were employed. Sequence of primers used for qRT-PCR are shown in **Table 5**. Raw data was analysed using the 7300 System Sequence Detection Software v1.2 (Applied Biosystems, ThermoFisher Scientific, USA). Sample normalization was performed to the housekeeping gene *UBIQUITIN3* and Ct calculation method (Livak & Schmittgen, 2001) was carried out to quantitation of relative gene expression. Finally, DNA contamination was assessed using a tomato intron specific primer.

RNA sequencing

RNA isolation from inflorescence stems of Moneymaker and *hap* mutant plants was performed with TRIzol® Reagent (Invitrogen Life Technologies, USA). Samples consisted of three biological replicates per each genotype (Hrdlickova et al., 2017). Preparation of libraries was performed following Illumina TruSeq RNA protocol and sequenced on the Illumina HiSeq2000 platform (Illumina, Inc., USA) with paired-end 150 pb. Quality control and adapter trimming of paired-end short reads were carried out by means of FastQC (Andrews, 2010) and Trim-Galore (Krueger, 2017). Then, the filtered paired-end short reads were aligned to the reference genome (*Solanum lycopersicum* ITAG 2.50) by means of STAR (Dobin, 2013). The Bioconductor R packages Rsamtools (Morgan et al., 2019), GenomicFeatures and GenomicAlignments (Lawrence et al., 2013) and SystemPipeR (Backman, 2016) were used for transcript assembly and abundance estimation. Differentially expressed genes were identified by means of custom python scripts (Gómez-Martin, 2017). We have selected genes with |FC|>2 in all mutant vs wild type comparisons. RNAseq variant calling was carried out by means of bcftools mpileup and bcftools isec utilities (Li et al., 2011). Finally, Differentially Methylated Cytosines (DMCs) analyses was performed by means of the logistic regression function implemented in *methylKit* (Akalin et al., 2012) when replicates were available considering a q-value<0.01 and percent methylation difference larger than 25%.

HAP phylogenetic tree construction tools

Database searches of HAP coding sequence was performed using Sol Genomics Network (https://solgenomics.net/), as well as Tair (https://www.arabidopsis.org/) for *A. thaliana* homologues search. Phylogenetic tree was constructed by computing first a Multiple Sequence Comparison by Log-Expectation (*MUSCLE*) (https://www.ebi.ac.uk/Tools/msa/muscle/) and finally by construction of a neighbour joining tree (Madeira et al., 2019).

III.3. Results

hairplus mutant plants exhibit higher trichome density

The *hap* mutant was characterised as part of a mutant collection obtained in *S. lycopersicum* cv. Moneymaker (MM) using ethylmethanesulfonate (EMS) as the mutagen agent. The most conspicuous phenotype of this mutant is a higher trichome density in stems (**Figure 22**b) and inflorescence stems (Fig. 22d) when compared to wild type plants (Fig. 22a, 22c). Also, mutant plants exhibit a lower growth than the observed in wild type plants (Fig. 22e, 22f). In addition, *hap* plants develop plant shoots on the older leaf petioles (Fig. 22g, 22h). The segregation ratio observed in a M_2 segregating population was consistent with a monogenic recessive inheritance for the mutant phenotype (202 WT: 76 mut; χ^2=0.36, *P*=0.81).

HAP directs type I trichome development

In order to determine the trichome types responsible for the *hap* phenotype, trichomes identity and density of wild type MM and *hap* mutant plants were examined using Scanning Electron Microscopy (SEM). Mutant plants exhibit a higher density of type I trichomes than wild type plants, and significant differences were observed in the inflorescence stem (**Figure 23**a, 23b, 23e), stems (Fig. 23c, 23d, 23e) and in the adaxial side of leaves (Fig. 23e). Additionally, significant differences were also observed when it comes to type III trichome density but in the opposite way, since density is increased in the inflorescence stems of wild type plants compared with *hap* mutant plants (Fig. 23f). Mutant plants also showed significant differences in type VI trichome density in the abaxial side of leaves (Supplementary Fig. 1a) and in the inflorescence stem and adaxial

side of leaves for type VII trichome density (Supplementary Fig. 1b). Finally, *hap* mutant plants do not differ of wild type plants in type V trichome density nor in stomata density (Supplementary Fig. 1c, 1d). All these results bring together suggests that *HAP* controls type I trichome initiation particularly in stems and inflorescence stems.

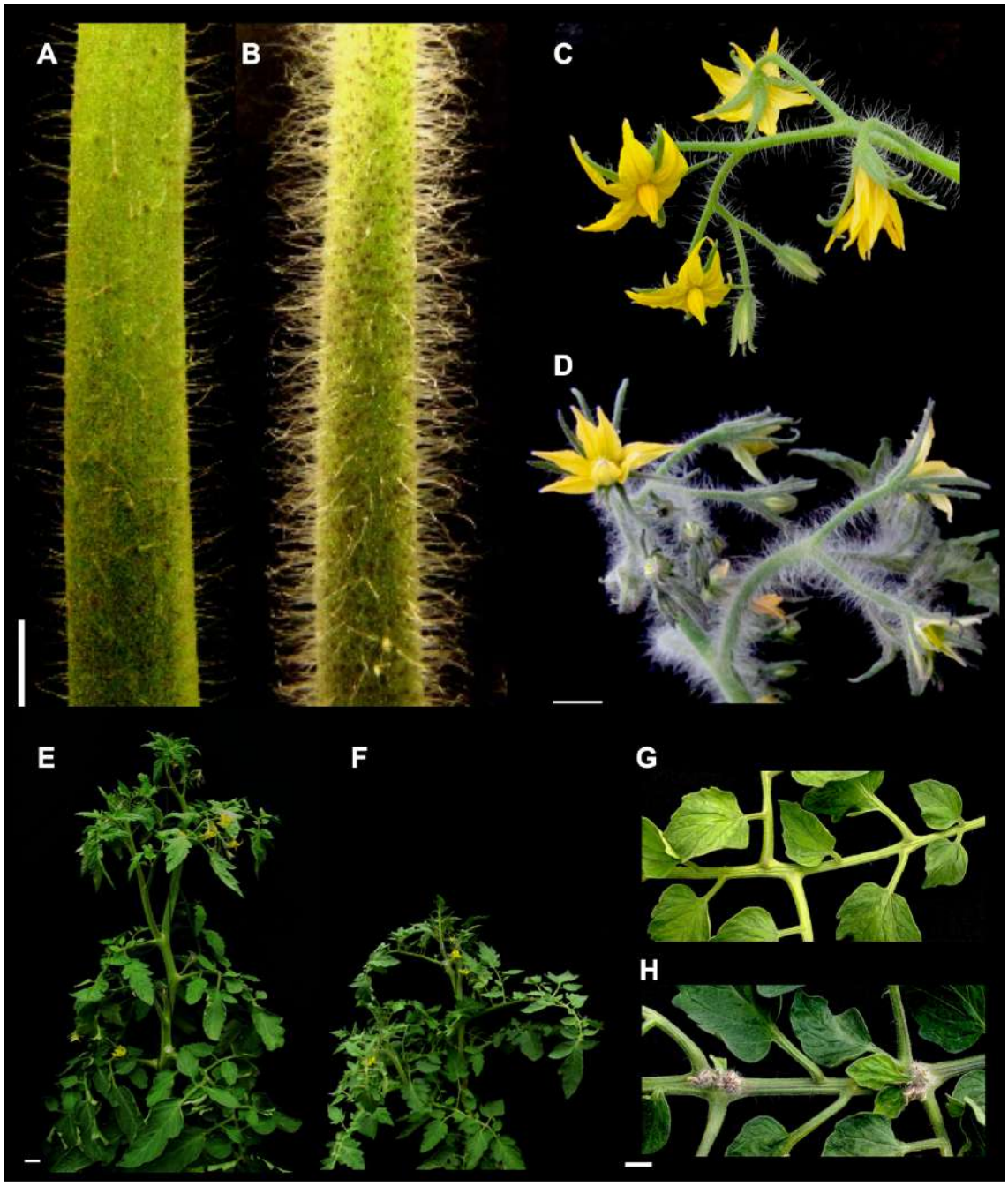


Figure 22: The *hap* mutant phenotype. A. Stem of wild type Moneymaker (MM) plants. **B.** Stem of a *hap* mutant plant where higher trichome density can be appreciated. **C.** Inflorescence stem

of a wild type plant. **D.** Inflorescence stem of a *hap* mutant plant showing higher trichome density. **E.** Wild type plant. **F.** *hap* mutant plant exhibiting lower growth. **G.** Petiole of a MM plant and **H.** Petiole of a *hap* mutant plant showing developing plant shoots. All scale bars apply to 1 cm.

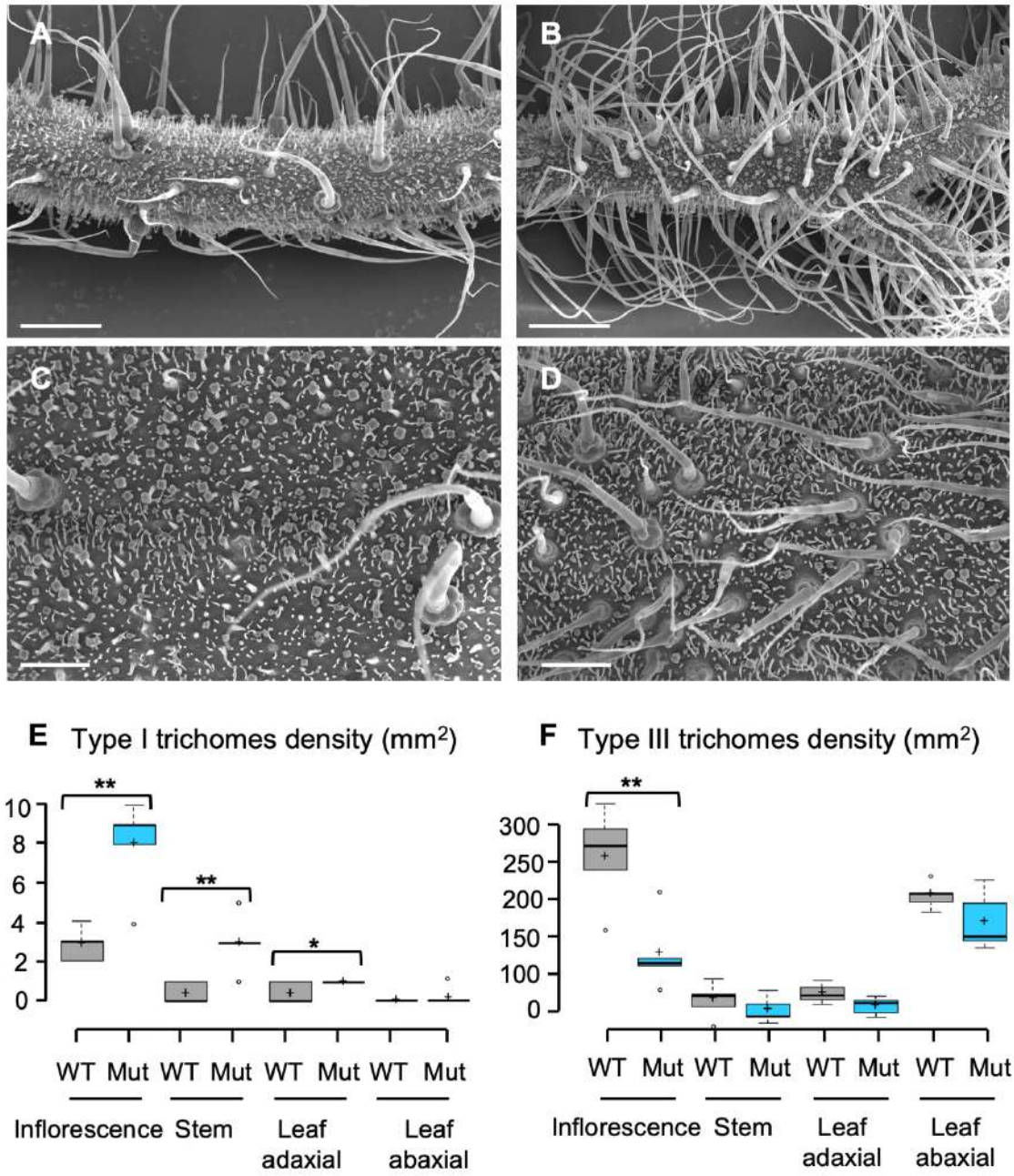


Figure 23: Trichome density in MM and *hap* mutant plants. A. Inflorescence stem of wild type MM plants. **B.** Inflorescence stem of a *hap* mutant plant where higher trichome density can be observed. **C.** Stem of a wild type plant. **D.** Stem of a *hap* mutant plant showing higher trichome

density. **E.** Type I trichome accounts in inflorescences, stems and leaf adaxial and abaxial side of wild type and *hap* mutant plants. **F.** Type III trichome accounts in inflorescences, stems and leaf adaxial and abaxial side of wild type and *hap* mutant plants. In **E.** and **F.** centre lines show the average values, whereas box limits indicate the 25th and 75th percentiles as determined by R software. Whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles and outliers are represented by dots. * Significant differences at $P < 0.05$. ** Significant differences at $P < 0.01$. Scale bars apply to 1 mm (a, b) and to 0,5 mm (c, d).

Characterisation of the *HAP* gene

Mapping of the *HAP* gene began with the generation of an F_2 segregating population derived from the cross of a *hap* mutant plant with a plant of the wild relative species *S. pimpinellifolium* accession LA1589 as pollen donor. The F_1 was self-pollinated and the F_2 population was characterised. Mapping was performed using codominant markers distributed along the genome (Capel et al., 2015), which allowed us to locate *HAP* to an approximately 0.6 Mb fragment in chromosome 10 (**Figure 24**a). In order to define the exact chromosomic location of the mutation, we sequenced two DNA pools, one consisting in equimolar amounts of DNA from 14 F_2 mutant plants and other of wild type plants. Analysis of the segregating allele frequencies confirmed the location of the mutation in chromosome 10 (Supplementary Fig. 2) and allowed to identify two mutation in the coding sequence of *Solyc10g077070* gene, where a deletion followed by a single-nucleotide variant (SNV) cause a premature stop codon (Fig. 24b). This gene codes for a Histone Lysine N-Methyltransferase homologue of the SET-domain containing proteins Su(var)3-9 (SUVH) of *A. thaliana*, and as a result of these mutations, the HAP mutant protein lacks all functional domains characteristics of these group of proteins, including the SAD-SRA and the SET-domain, which are crucial for the activity of the SUVH proteins (Lachner & Jenuwein, 2002) (Fig. 24c).

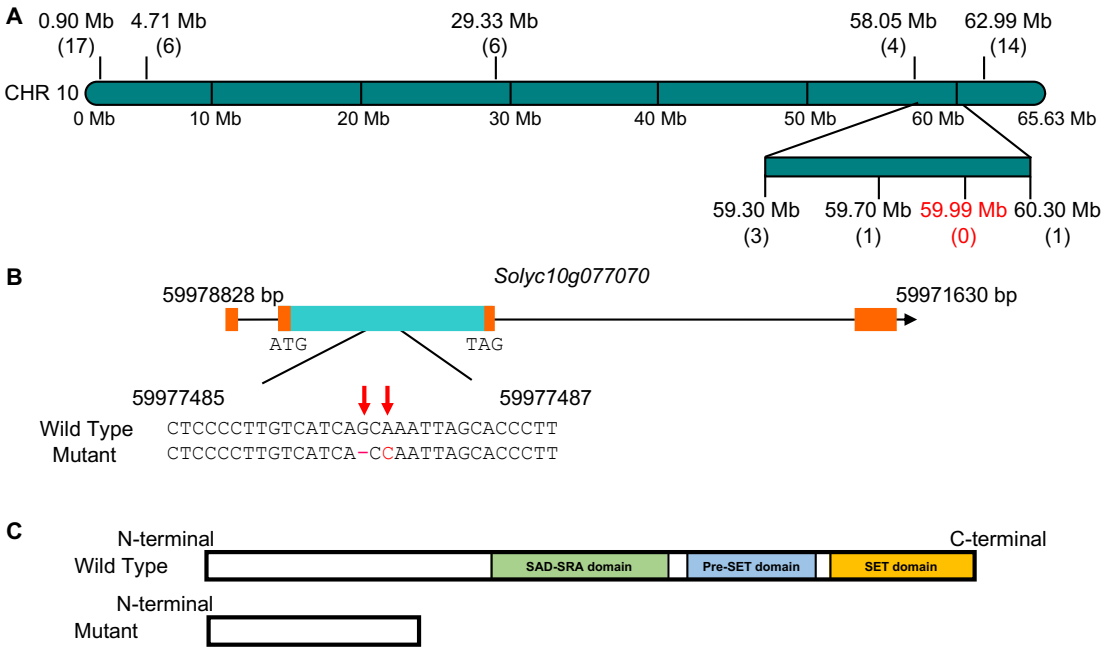


Figure 24: Map based cloning of *HAP*. A. Fine mapping delimited *HAP* to an interval of 602 Kb of chromosome 10. Numbers in parenthesis indicate the number of recombinant chromosomes identified between the *HAP* gene and each genetic marker analysed. **B.** Sequencing data obtained from a pool of DNA from *hap* mutant plants allowed the identification of a deletion followed by a point mutation in *Solyc10g077070*, a gene that encodes for a Histone-lysine N-methyltransferase. **C.** Schematic representation of the HAP protein domains. Mutations in *HAP* give rise to a premature stop codon and, as a result, the mutant protein lacks the functional domains characteristics of this gene family.

Silencing *Solyc10g077070* phenocopied the *hap* mutant phenotype whereas knocking it down increases even more type I trichome density

In order to confirm that the mutations in *Solyc10g077070* are responsible for the *hap* mutant phenotype, first silencing the gene by means of an RNA interference (RNAi) strategy was carried out. Eleven independents diploid transformants were obtained and used for phenotypic characterisation. Expression analysis performed by quantitative RT-PCR showed that *Solyc10g077070* expression levels was drastically reduced in all RNAi T_0 lines ranging from 0.04 to 0.27 relative expression of that observed in WT plants (Supplementary Fig. 3). All T_0 RNAi lines showed increased trichome density in

inflorescence stems (**Figure 25**c) and all have a similar phenotype to that of the *hap* mutant (Fig. 25b).

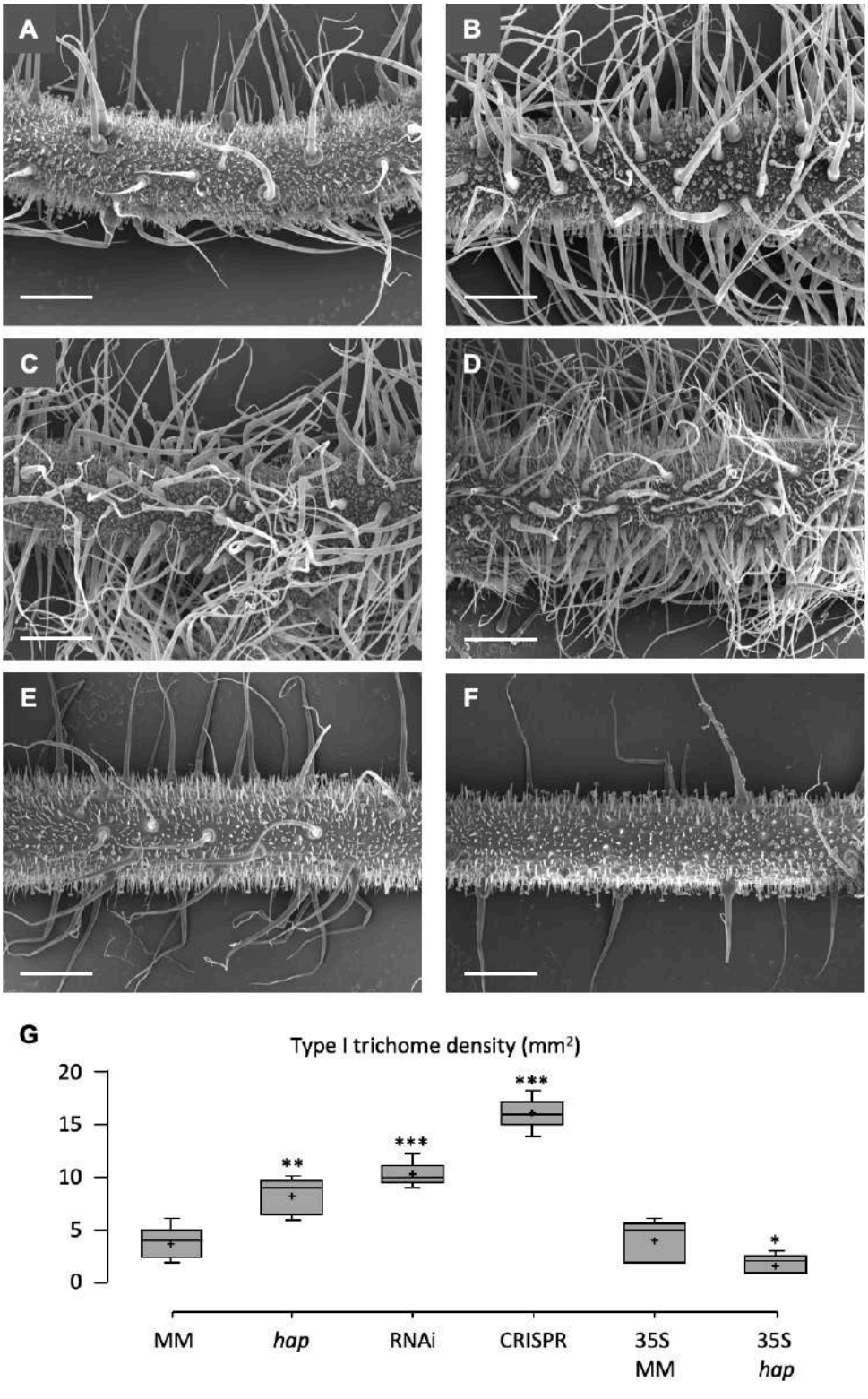


Figure 25: Modulation of type I trichome density by *HAP* gene expression. a, Inflorescence stem of a MM wild type plant. **b**, Inflorescence stem of a *hap* mutant plant. **c**, Phenotype of a *HAP* RNAi silencing line showing increased type I trichome density. **d**, Phenotype of the inflorescence stem of a CRISPR-CAS9 knock-down line where the increase in type I trichome density is even more evident. **e**, Inflorescence stem of an overexpression line obtained in a MM background shows no difference when compared to MM. **f**, Inflorescence stem of an overexpression line obtained in a *hap* mutant background showing less type I trichome density than the observed in MM wild type plants. **g**, Type I trichome accounts performed in transgenic lines, lines show the average values whereas box limits indicate the 25th and 75th percentiles as determined by R software. Whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles and outliers are represented by dots. * Significant differences at $P < 0.05$. ** Significant differences at $P < 0.01$. *** Significant differences at $P < 0.001$. All scale bars apply to 1 mm.

We also performed a CRISPR-Cas9 genome editing of the *Solyc10g077070* gene and the four diploid T_0 lines obtained were characterised by sequencing the region were the sgRNA was located (Supplementary Fig. 4). The four T_0 lines were biallelic mutants and their phenotypes either resembled the *hap* mutant phenotype or even had a higher type I trichome density than the original *hap* mutant plants (Fig. 25d). Finally, overexpression of the *Solyc10g077070* gene produced no phenotypic changes in MM plants (Fig. 25e) but restored the wild type phenotype in *hap* mutant plants (Fig. 25f), demonstrating that a WT copy of the gene complemented the mutant phenotype. Accounts of type I trichome density in these transgenic lines confirmed the significant increase observed in RNAi silencing and CRISPR-Cas 9 knock down lines when compared to wild type and even with *hap* mutant plants (Fig. 25g). On the other hand, a reduction in type I trichome density was observed in overexpression lines obtained in a *hap* mutant background (Fig. 25g). All these results bring together support that *Solyc10g077070* is responsible for the *hap* mutant phenotype.

Spatial expression pattern of *HAP*

With the aim to determine the expression pattern of *HAP*, we performed quantitative RT-PCR using total RNA extracted from several tissues of wild type MM plants, which

included root, stem, inflorescence stem, flower buds, whorls of anthesis day flowers and fruits at different stage of maturation. Expression pattern confirmed that *HAP* is constitutively expressed in all analysed tissues, with the highest expression level in adult leaf followed by the inflorescence stem. The lower levels of transcripts of the gene were detected in sepals, a tissue with trichomes, as well as in fruits in breaker state, which do not have trichomes on their surface (Supplementary Fig. 5a).

On the other hand, the expression in *hap* mutant plants of genes related to type I multicellular trichome development in tomato was also analysed. The expression analyses included genes involved in type I trichome development such as the previously described *Hairless* (*Hl*) (Kang et al., 2016), *Woolly* (*Wo*) (Yang et al., 2011) and *Hair* (*H*) (Chang et al., 2018). We have also analysed the expression of *SlCycB2*, a gene located in chromosome 10 that codes for a B type cyclin that physically interacts with *Woolly* (Yang et al., 2011). We found *Woolly*, *SlCycB2* and *Hair* to be slightly induced in *hap* mutant plants whereas *Hairless* showed expression values close to those observed in wild type plants, being *HAP* the only gene repressed in *hap* mutant plants (Supplementary Fig. 5b).

HAP homologues are conserved trough evolution but have diverged with unknow function

The predicted coding sequence of *HAP* was 2145 bp and the protein encoded by this sequence is 715 aa and as mentioned above belongs to a group of SET-domain proteins linked to epigenetic control of gene expression. In the model species *A. thaliana* 29 genes coding for SET-domain containing proteins have been identified (Baumbusch et al., 2001). With the aim to characterise the phylogenetic relationship among these proteins and the tomato SET-domain proteins including HAP, we first computed a Multiple Sequence Comparison then deriving a neighbour joining tree (Madeira et al., 2019). Among the most closely related proteins to HAP, there is one from tomato coded by the *Solyc09g082050* gene located in chromosome 9, and four *Arabidopsis* proteins from the SU(VAR)3-9 group which have all been grouped into one clade (Supplementary Fig. 6). Notably, multiple alignment showed that proteins in this clade share a small

region of 8 amino acids within the SET-domain (GWGLRSWD) that is exclusive to this clade (Supplementary Fig. 7).

Transcriptomic changes in *hap*

To reveal differentially expressed genes (DEGs) between wild type and *hap* mutant plants, we performed a high-through-output RNA-seq analysis using three biological replicates of each, mutant and wild type plants. This analysis allowed us to identify 84 downregulated, and 8 upregulated genes in mutant plants (Supplementary Tables 2 and 3). Among the upregulated genes, we found *Solyc12g096570* which is homologue of the *Arabidopsis ARGOS*, a gene highly induced by auxin that plays a key role in aerial organs size control (Hu et al., 2003). Another upregulated gene is *Solyc02g085910*, that codes for a homologue of the LOB-domain containing proteins. Interestingly, ectopic expression of *LOB* in *Arabidopsis* leads to alterations in the size and shape of leaves and floral organs as wells as to male and female sterility (Shuai et al., 2002). Among the genes almost not detected in wild type plants and induced in *hap* mutant plants is *Solyc03g082550*, that codes for a Homeobox leucine zipper protein homolog of a *Arabidopsis* protein family implicated in the ABA response network and cell growth control (Hur et al., 2015; Song et al., 2016). On the other hand, the *HAP* gene is also downregulated in the *hap* mutant which supports the results obtained by quantitative expression analyses. Finally, five genes (*Solyc09g092490*, *Solyc10g085870*, *Solyc10g085880*, *Solyc12g088700*, *Solyc12g098600*) of the numerous gene family that code for UDP-glucosyltransferase belonging to the family 1 of Glycosyltransferases protein (Vogt & Jones, 2000), were also found repressed in the *hap* mutant.

HAP induces several epigenetic changes

Two of the *Arabidopsis* SUVH proteins homologs of HAP have not been characterized yet but the other two bind methylated DNA and enhance proximal gene expression promoting demethylation and transcriptional activation of genes located near transposons. The epigenetic modifications caused by loss of function of *HAP* were analysed by Whole-Genome Bisulfite Sequencing of three DNA replicates of each, *hap*

and wild type plants. In the results obtained, Differentially Methylated Cytosines (DMCs) between wild type and mutant phenotypes were obtained by means of the logistic regression function implemented in *methylKit* (Akalin et al., 2012) by comparison of the sequences from the 3 replicates of wild type to the 3 replicates of *hap* mutant DNA. A total amount of 2,113 DMCs were obtained (q-value<0.01 and percent methylation difference larger than 25%), from which 270 were hyper- and 1843 hypo-methylated cytosines in mutant plants. The DMCs identified in the genome of *hap* mutants are not randomly located since their distribution shows a correlation with the well-known non-random distribution of genes, being DMCs mainly located in the euchromatic telomeric portion of the 12 tomato chromosomes (**Figure 26**a; supplementary Fig. 8). Chromosome 10 contains the highest amount of hypo- and hypermethylated DMCs, whereas chromosome 11 is where fewer DMCs have been located. The rest of chromosomes contain a similar amount of DMCs (Fig. 26b). Although DMCs are mainly located in the gene-rich portion of the genome, most of them (64 %) are located in intergenic regions, being the 25 % located in promoters and 11 % in transcribed regions from which 5 % were found in exons and 6 % in introns (Fig. 26c). Most of the DMCs were found far away from DEGs, except for *Solyc03g082550* that codes for a Homeobox leucine zipper protein and holds a DMC in its promoter which is demethylated in *hap* and hypermethylated in wild type plants (Fig. 26d). Among the infra-expressed genes in *hap* there are 5 genes with at least one DMC in their transcribed sequences (*Solyc01g068460*, *Solyc03g097870*, *Solyc04g007400*, *Solyc07g006680*, *Solyc11g071290* and *Solyc12g098600*) and 3 genes (*Solyc04g052980*, *Solyc06g052010* and *Solyc10g077070-HAP*) with DMC in their promoters. *HAP* gene itself (*Solyc10g077070*) has a DMC 2 Kb 5' upstream in a region that can be considered its promoter. This DMC is located close to an LTR and hypermethylated in the *hap* mutant (Fig. 26e), suggesting that it can be a direct target of HAP.

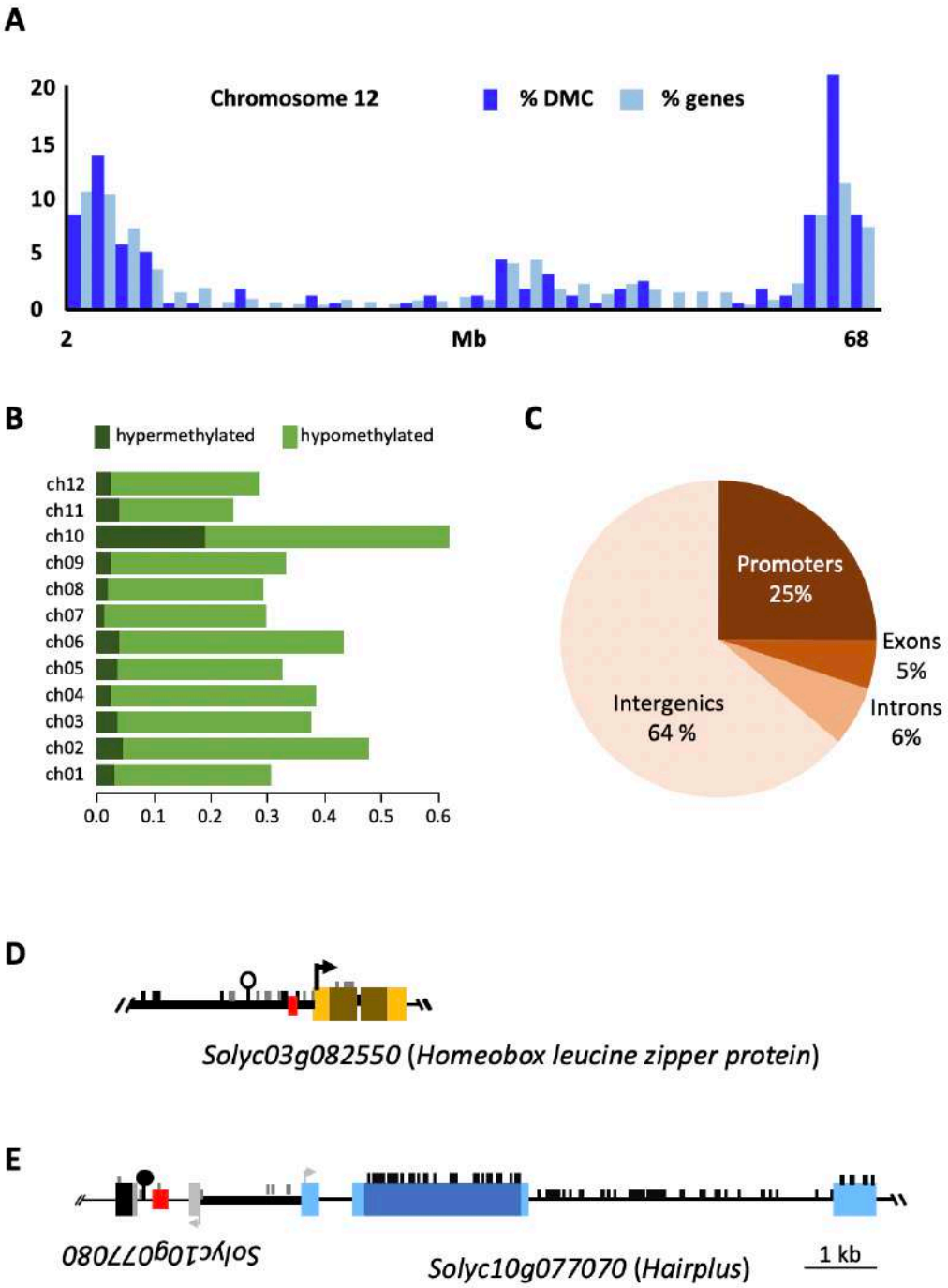


Figure 26: Location of DMCs in *hap* mutant plants. A. Density of DMCs and genes over chromosome 12. Each bar represents the percentage of DMCs or genes in 2 Mb. **B.** Percentage of hyper (dark green) or hypo- (light green) methylated DMCs founded in the *hap* genome. **C.** Differential methylation annotation of DMCs CpG. **D.** Genomic organization of the *Solyc03g082550* gene and **E.**, of the *HAP* gene. In **D.** and **E.** boxes represent exons and thin horizontal lines represent introns; promoter sequences are depicted as tick lines; annotated LTR

are represent as red boxes; methylated CpC are represented as vertical lines. Hypomethylated DMC is represented as a white circle whereas hypermethylated CpG is represented as a black circle.

III. 4. Discussion

Trichomes are protuberances of epithelial cells with multiple functions, with pest defence being one of the most proven. Thus, an increase in glandular trichome density has been correlated with resistance to plagues in tomato species. Such is the case of type IV glandular trichomes, that have been related with an increase in entrapment of some Lepidoptera plagues (Simmons et al., 2004) or with resistance to spider mite *Tetranychus urticae* (Koch) (Carter & Snyder, 1985; Salinas et al., 2013). In *S. galapagense*, a negative correlation has been observed between the presence of glandular trichomes of types I, IV and VI and adult survival, oviposition rate and pre-adult survival of the whitefly *Bemisia tabaci* (Lucatti et al., 2013). In all of these cases the key factor is the exudates composition. Among the broad spectrum of chemical compounds produced by glandular trichomes, acyl sugars followed by terpenes are the most important ones when it comes to pest resistance. Even more, removal of glandular trichomes exudates increases survival and longevity of pests (Simmons & Gurr, 2005). Given that the *hap* mutant exhibits a higher density of glandular type I trichomes in vegetative and reproductive stems, it would be expected an increase in pest resistance. Preliminary data obtained while assessing *hap* mutant resistance to the tomato fruit borer *Helicoverpa armigera* (Hübner) suggests that in fact damaged is reduced in mutant plants when compared to wild type ones (Supplementary Fig. 10). Future works should be focused to demonstrate the pest resistance displayed by the *hap* mutant plants, and if it is related to a higher amount of exudates or to the physical barriers that the trichomes represent in mutant plants.

Even tough trichomes are epidermal structures common in plants, gene networks controlling their development seems to be quite different among different species (Serna & Martin, 2006). Several studies have been carried out over the last decades in the model species *A. thaliana*, allowing the characterisation of a molecular complex of transcription factors that controls trichome development (Esch et al., 2003; Payne et al.,

2000). As part of this complex, a WD4 transcription factor named TTG1 interacts with GL3 as well as with EGL3 (both bHLH transcription factors) and with the R2R3-MYB transcription factor GL1. Finally, the complex activates trichome initiation by binding to GL2. Also, single-repeat MYBs named CPC, TRY and ETC1 repress trichome formation by competing with GL1 for binding to bHLH proteins (Serna & Martin, 2006). Nevertheless, very few information exists when it comes to trichome development in tomato and other Solanaceae species. *Hair* and *Woolly* are among the few genes characterised in tomato, both related to type I trichome differentiation. Physical interaction among both transcription factors as well as interaction of Woolly with SlCYCB2 has been proved and is a new evidence for the existence of a novel gene network controlling trichome development in tomato (Yang et al., 2011; Chang et al., 2018). *Hairless,* a gene also involved in multicellular type I trichome development, has also been characterised. In this case, *hairless* mutant plants exhibit aberrant trichome morphology as a result of mutations in *SRA1*, a tomato homologue involved in actin dynamics control (Kang et al., 2016). Here, we have characterised *HAP*, a novel gene involved in type I trichome formation thought epigenetic modification of the tomato genome.

Regarding functional characterisation, HAP contains three conserved domains. First, there is a SAD-SRA-domain known for recognizing the hemi-methylated sites of the DNA and for directing the DNA methyltransferase to these sites for the methylation of newly synthesized CpG sequences, helping thus to the inheritance of the epigenetic methylation patterns (Arita et al., 2008). Next, there is a Pre-SET-domain followed by a SET-domain. The Pre-Set-domain is a zinc binding motif that contains 9 conserved cysteines that coordinate three zinc ions (Min et al., 2002). This motif is exclusive of the SUVH3-9 family and it is thought to play a key role in stabilizing the SET-domain and in providing the specificity needed for it to methylate the lysine 9 residue of H3 histone (Kouzarides, 2002). H3K9 methyltransferases thus may act as epigenetic gene regulators.

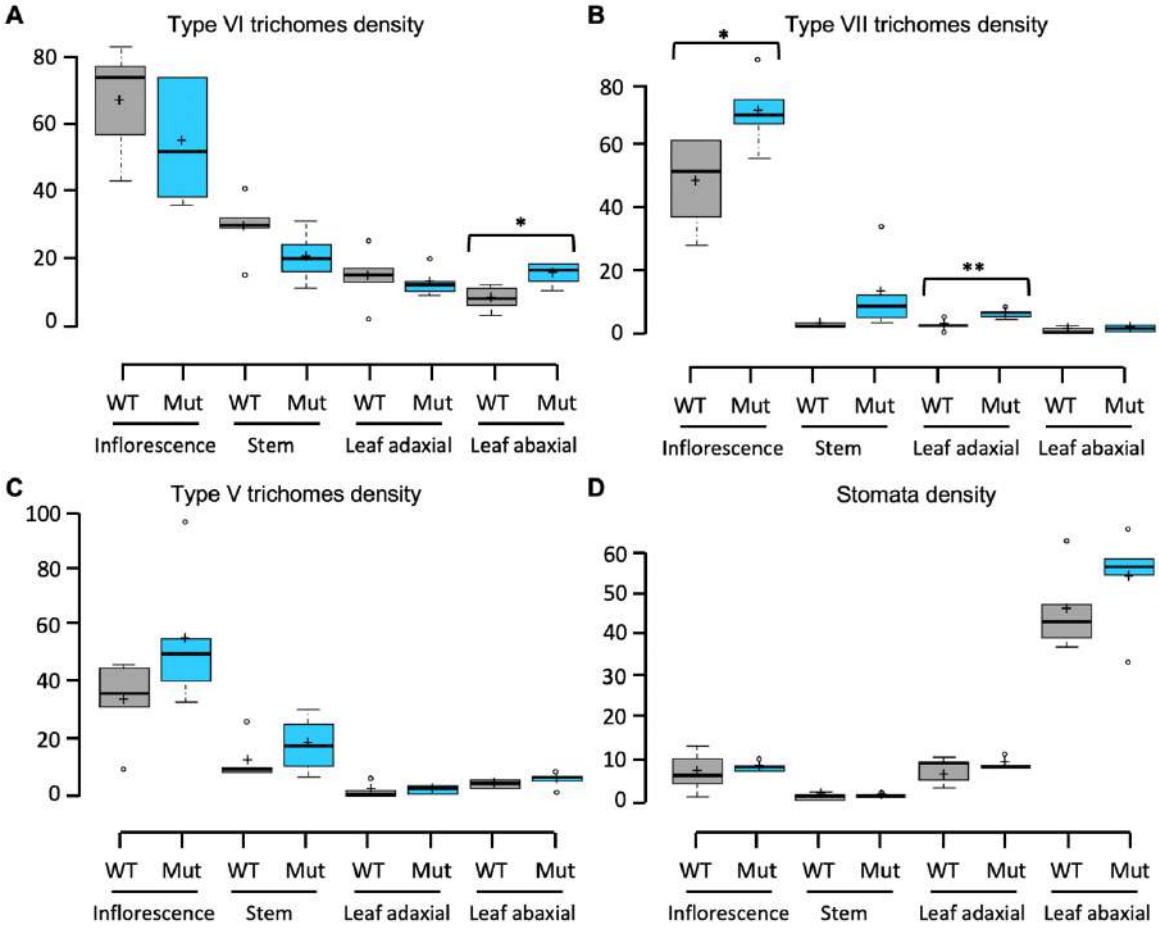
The differential methylation of histone lysine residues by the SET-domain holds a key role in regulating the dynamics between the two different states of chromatin. Euchromatin is characterised by methylation of histone 3 at K4, K36 and K79, whereas

heterochromatin contains methylated K9 and K27 residues (Ebert et al., 2004). Also, H3K9 methylation can appear in mono, bi or trimethylated forms, been trimethylation the hallmark of constitutive heterochromatin in mammals (Peters et al., 2003). It has been demonstrated that the silencing potential of H3K9 defective mutants of *Droshophila melanogaster* was correlated with the amount of HMTase activity they hold (Kouzarides, 2002). In *A. thaliana* the methyltransferase H3K9 *KRYPTONITE* (*KYP*) has been characterised as a suppressor of the *SUPERMAN* (*SUP*) locus (Jackson et al., 2002). Mutant plants of one of the alleles of *SUP* named *clk-s*t show defects in the number of floral organs due to the extreme methylation of *SUP*. The *kyp* loss-of-function mutants did not exhibit morphological alterations but a reduction in *SUP* methylation level is observed in CpG and especially CpHpG methylation. Thus, it could be inferred that H3K9 methyltransferases are related with promoting gene silencing.

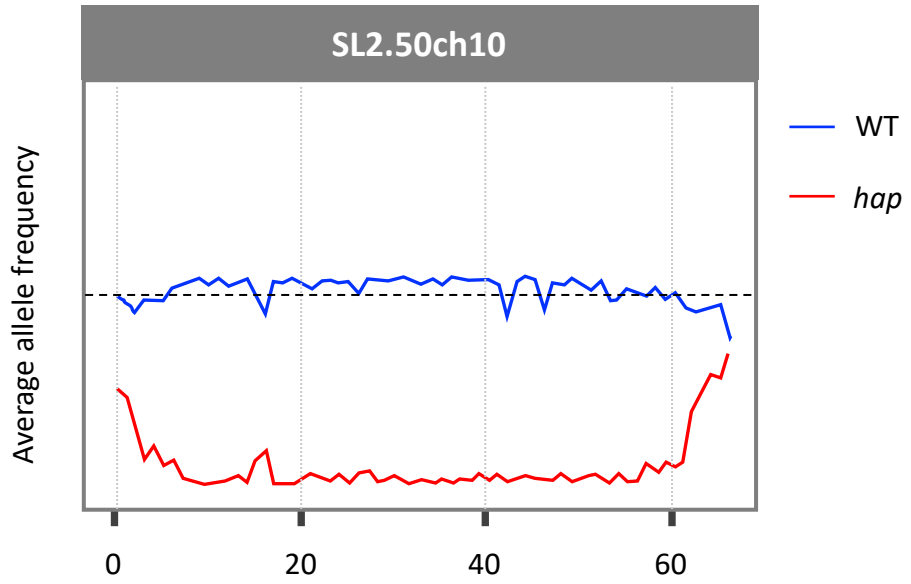
In this study we have provided solid evidence of the role of *HAP* in trichome development control. Nevertheless, the closest *HAP* homologue in *Arabidopsis* (*At1g73100*) is not related with trichome initiation in this model species. *At1g73100* gene codes for a SUVH3 protein and its homologue SUVH1 (*At5g04940*) have a partially redundant function and interact with at least three DNAJ-domain containing proteins forming a complex. This complex then binds to RNA-dependent methylated sites in the genome near Transposable Elements (TEs) located in the promoter sequence of some genes. The complex induces the demethylation and transcription of the genes proximal to the TEs (Zhao et al., 2019). The presence of a methylated DNA recognition domain (SAD-SRA) in HAP could suggest a similar function. Interestingly, no *in vitro* histone methyltransferase activity has been detected for neither *suvh1* nor *suvh3* mutants and their methylation levels appear unaltered (Harris et al., 2018). Since the number of DMCs in mutant plants is higher than the observed in wild type plants and given the large number of down-regulated genes (84) we found in *hap* is probable for HAP to form a SUVH-DNAJ complex similar to that described for SUVH3-DNAJ, inducing the transcription of the promoter methylated genes instead of functioning as a Histone methyltransferase. Further studies will be focused on dissecting whether if HAP is capable of recognition of RNA-dependent methylated regions and if recruits DNAJ-

domain containing proteins to these sites of the genome. A candidate genomic region where HAP could modulate CpG methylation is its own promoter, where there is a DMC hypermethylated in the mutant located in the proximity of an LTR (Fig. 5e). This latter DMC, together with the DMC located in the promoter of the Homeobox coding gene *Solyc03g082550*, could considered as CpG traffic light (Lioznova et al., 2019), since their methylated status are correlated to the expression levels of the surrounding genes.

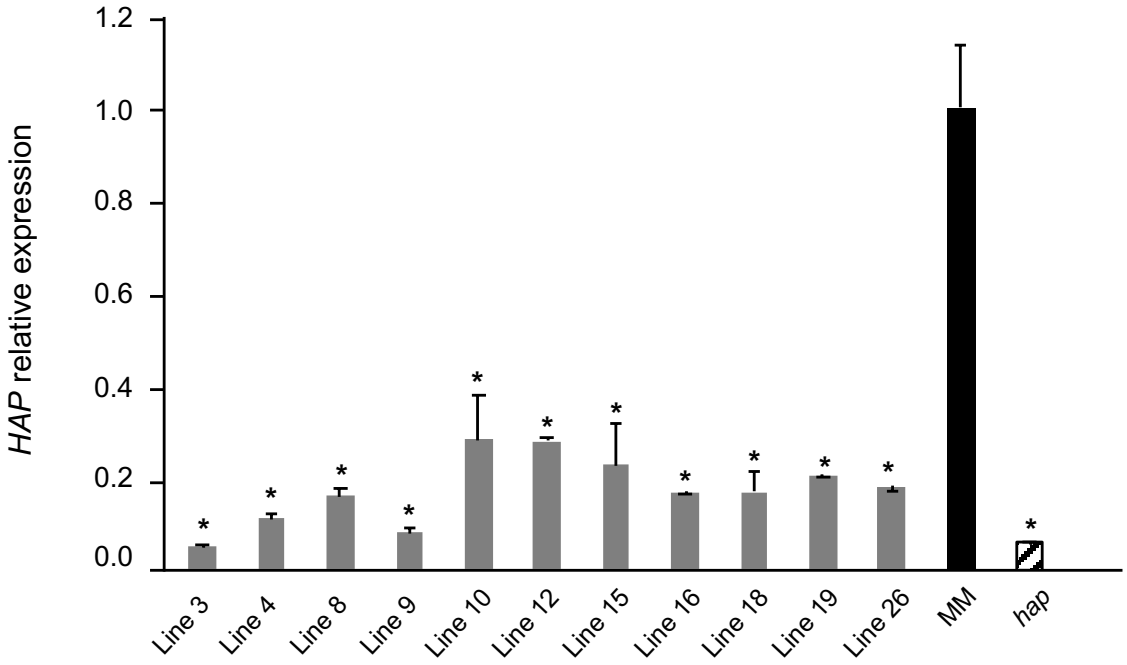
ANEXO II. Supplementary figures and tables



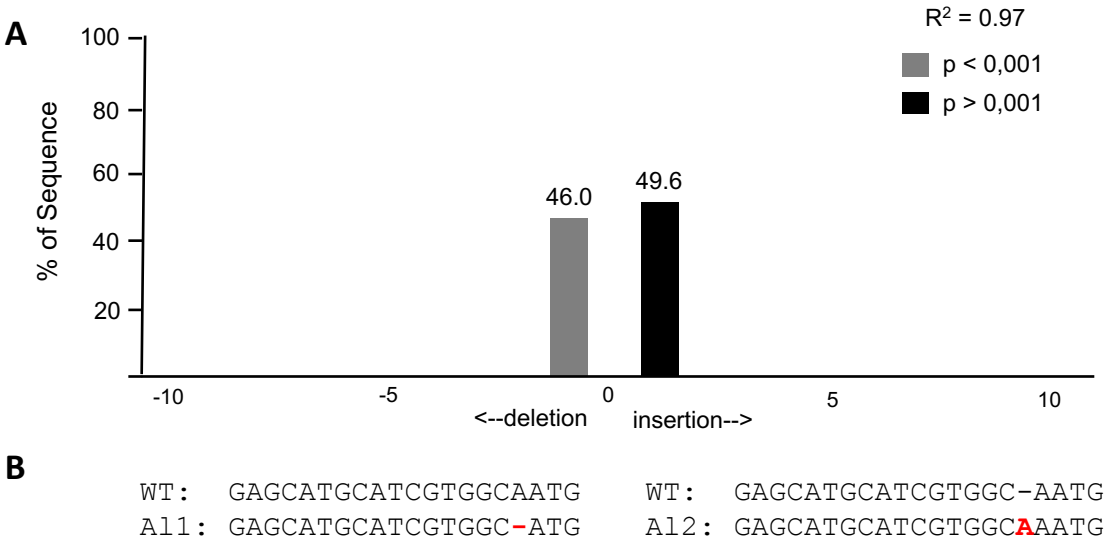
Supplementary Fig.1: Trichome and stomata density in inflorescence stems, stems and leaves of WT and *hap* mutant plants. A. Type VI trichome density. **B.** Type VII trichome difference. Significant differences are observed in the inflorescence stem and the adaxial side of leaves. **C.** Type V trichomes density. **D.** Stomata density. Center lines show the average values, whereas box limits indicate the 25th and 75th percentiles as determined by R software. Whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles and outliers are represented by dots. Statistical analyses were performed with Fischer´s LSD method (P = 0,05). * Significant differences at $P < 0.05$. ** Significant differences at $P < 0.01$.



Supplementary Fig. 2: Mapping by sequencing identification of *HAP*. Comparison of the allele frequencies obtained from the sequence of a DNA pool from 14 mutant plants (*hap*) and 14 wild type (WT) from a F_2 population derived from the cross between a *hap* mutant plant and a plant of the wild relative species *S. pimpinellifolium* accession LA1589. The centromeric region of chromosome 10 shown a drop in the average allele frequency indicating that *HAP* is located in a portion of chromosome 10.



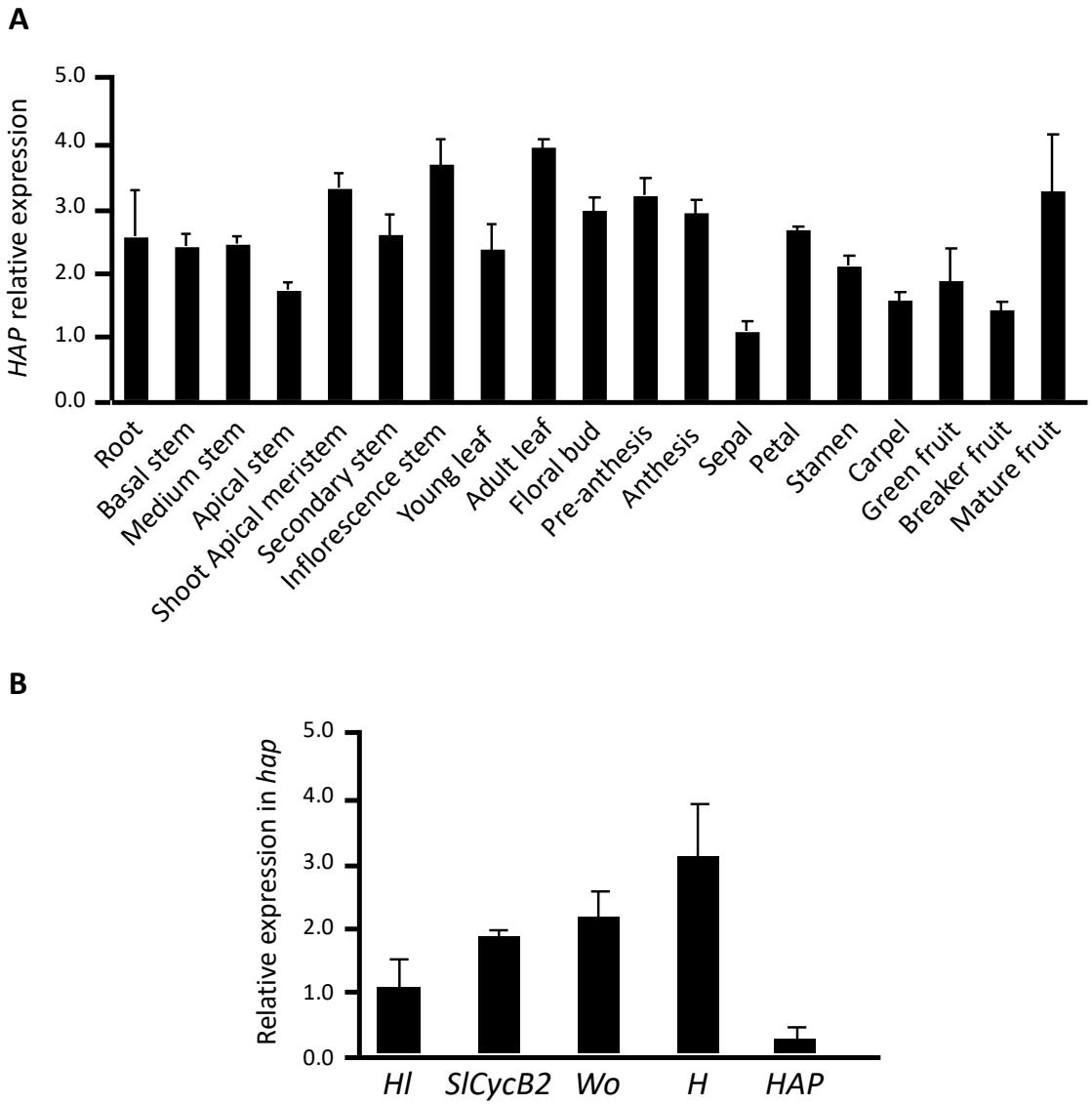
Supplementary Fig. 3: Expression analysis of *HAP* in RNAi silencing lines. Expression in *HAP* silencing lines (RNAi) decreased 4- to 25-fold (relative expression levels of 0.27 in line 12 and 0.04 in line 13, respectively) when compared to MM wild type plants. Asterisks account for significant differences at $P < 0.05$.



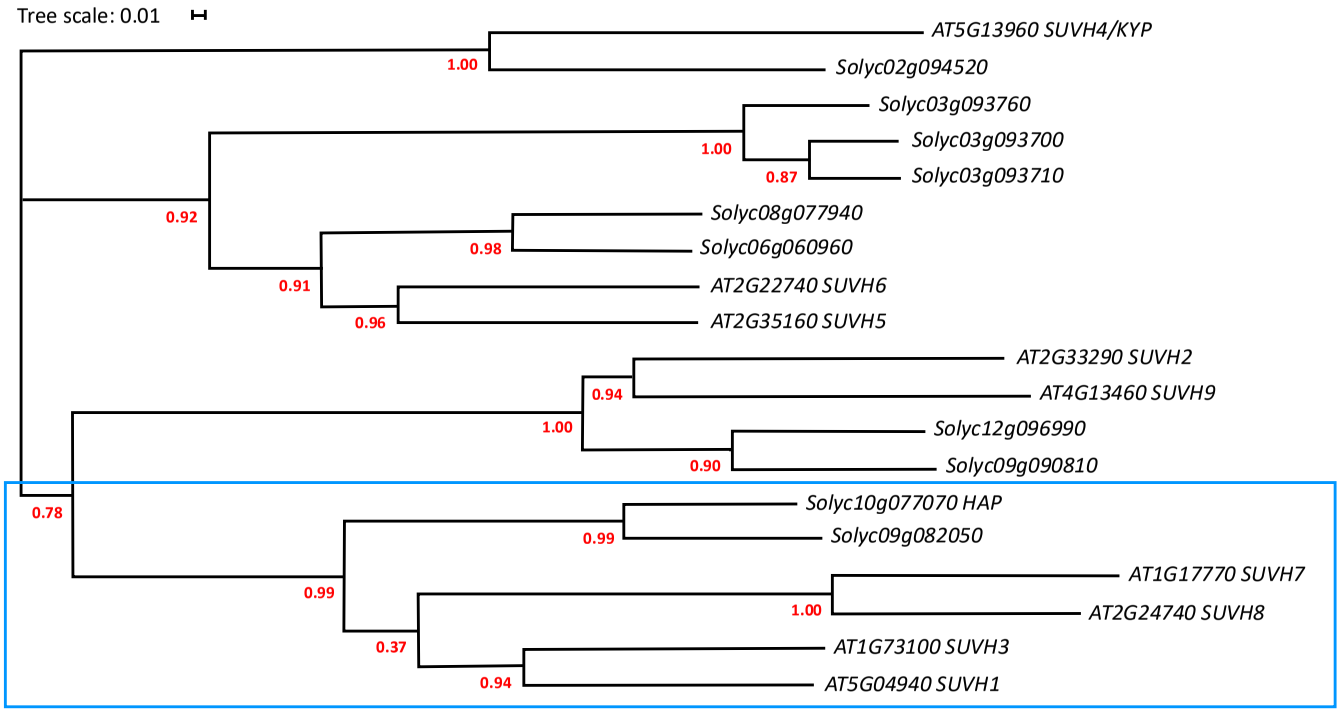
B

```
WT:  GAGCATGCATCGTGGCAATG    WT:  GAGCATGCATCGTGGC-AATG
Al1: GAGCATGCATCGTGGC-ATG    Al2: GAGCATGCATCGTGGCAAATG
```

Supplementary Fig. 4: Allele characterisation in a CRISPR-Cas9 knock-down line. **A.** Quantitative assessment of genome editing performed with the program Tide (https://tide.deskgen.com/) in a biallelic line. **B.** Allele sequence of the target sgRNA in wild type plants and in the CRISPR line characterised.



Supplementary Fig. 5: Spatial expression pattern of *HAP.* A. Relative expression was determined by real time RT-PCR using three biological replicates in different tissues from wild type MM plants and **B.** Expression analysis performed in leaves from *hap* mutant plants demonstrates that *HAP* is a constitutively expressed gene in MM plants, but it is repressed in *hap* mutant.



Supplementary Fig. 6: Phylogenetic relationships between *Arabidopsis* and tomato SET domain containing proteins. The tree was constructed using the alignment program MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle/). Bootstrap values are represented in red. Outlined by blue box is the clade that groups HAP and its most closely related proteins.

```
AT2G33290_SUVH2         ------------------------------------------------------------
AT4G13460_SUVH9         ------------------------------------------------------------
Solyc12g096990          ------------------------------------------------------------
Solyc09g090810          MYFINKILKIYKISNFTCISTKKRKVKFIILILLL-------------------------
AT1G17770_SUVH7         ------------------------------------------------------------
AT2G24740_SUVH8         ------------------------------------------------------------
Solyc10g077070_HAP      ------------------------------------------------------------
Solyc09g082050          ------------------------------------------------------------
AT1G73100_SUVH3         ------------------------------------------------------------
AT5G04940_SUVH1         ------------------------------------------------------------
AT5G13960_SUVH4/KYP     ------------------------------------------------------------
Solyc02g094520          ------------------------------------------------------------
Solyc03g093760          --------------------MVDTESPSTFKRLKIH------------------------
Solyc03g093700          MSVLAETSPSEMLGKKKLHQMVDTKSPSTFKRVKVD------------------------
Solyc03g093710          -----------MLGKKKLHQMVDTKSPSTVKRLKVD------------------------
Solyc08g077940          -----MVSF-----------------------------------------------------
Solyc06g060960          -----MIYWTELDINSTIWMVMIHIENEFFCRMISIILNTDLLGQRRNRYQMSSFQFQDS
AT2G22740_SUVH6         --------------MEMGVMENLMVHTEISKVK--------------------------
AT2G35160_SUVH5         ---------------------MVHSESSILSSLR--------------------------

AT2G33290_SUVH2         ---------------------------------------MSTLLPFPDLNLMPDSQSSTA
AT4G13460_SUVH9         ------------------MGSSHIPLDPSLNPSPSLIPKLEPVTESTQ-NLAFQLP----
Solyc12g096990          ----------------MEMGSVV-GLGDVN--FSTEPKTPTPTMIFPK-I----------
Solyc09g090810          --HLFTLSPYEHTCHFCKTSKLCSLFDFLK--TQNPIFDLGSLVPFQDLNLQPESTNFTS
AT1G17770_SUVH7         ----------------------------MDKSIPIKAIPVACVRP-DLVDDVTKNTS
AT2G24740_SUVH8         -----MVSTPPTLLMLFDDGDAGPSTGLVHREKSDAVNEEAHATSVPP-HAPPQTLWLLD
Solyc10g077070_HAP      ----------------MEQGFGSDSVPPAGPIDKSKVLDVKPLRCLVP-VFPS-------
Solyc09g082050          ----------------MEQGFGSDPSGSTI--DKTRVLDVKPLRCLSP-VFPSASE----
AT1G73100_SUVH3         --------------MQGVPGFNTVPNPNHY--DKSIVLDIKPLRSLKP-VFPN-------
AT5G04940_SUVH1         --------------MERNGGHYT---------DKTRVLDIKPLRTLRP-VFPS-------
AT5G13960_SUVH4/KYP     ----------------MAGKRKRANAP----DQTERRSSVRVQKVRQKALDEKAR-----
Solyc02g094520          ---------MVVPCVAELSDPVNDAMVPRR--CSARIKKLKSEQEAQR-ERESQRVRCRS
Solyc03g093760          --ATRNFPENCGPFVCQNNGSRK-IYPEFP--SNTKRVKVDSRRSFPE-NCGPQKRD---
Solyc03g093700          --ATRNFPENCGSFVYQNNGSKD-IYPEFR--SNSKRVKVNSTRSFPK-NCGPCVPEKKK
Solyc03g093710          --ATRNFPENCGPFVGENDGTGD-KYPEFP--SATKPVKVETTRNYPE-NCGPCVLQKKN
Solyc08g077940          --SNDGLSDQCVKKRSSVNGYHL-LDSGTM--SKHKVRIVCGEQDLPP-GCSRNAPKVDL
Solyc06g060960          NVSKDGLSNKSVKKRLLENGCHS-SYLGII--PKYKIRKVSAVRDFPP-GCGRTSLKVDL
AT2G22740_SUVH6         --SQSNGEVEKRGVSVLENGGVCKLDRMSG--LKFKRRKVFAVRDFPP-GCGSRAMEVKI
AT2G35160_SUVH5         --GGDGGGIPCSKDELAINGSYTDPMGR----RKSKRFKVAAESEFSP-DFGSITRQLRS

AT2G33290_SUVH2         ----------GTTAGDTVVTGKLEVKSEPI-------------------------------
AT4G13460_SUVH9         ----------NTNPQALISS-AVSDFNEAT-------------------------------
Solyc12g096990          ------------------------EPKLEPL-------------------------------
Solyc09g090810          ----------STTPNPRIIP-KIEPKLEPL--------------------------------
AT1G17770_SUVH7         ----------TIPTMVSPVLTNMPSATSPL--------------------------------
AT2G24740_SUVH8         ----------NFNIEDSYDR-DAGPSTGPV--------------------------------
Solyc10g077070_HAP      --------------PNGMAS-GTTPQPSPF--------------------------------
Solyc09g082050          -----------------MSS-ITTPQPSPF--------------------------------
AT1G73100_SUVH3         ----------------------GNQGPPF--------------------------------
AT5G04940_SUVH1         ----------------------GNQAPPF--------------------------------
AT5G13960_SUVH4/KYP     ---------------------LVQERVKLL--------------------------------
Solyc02g094520          ----------NDDSVLGKKT-KVYKKSKLV--------------------------------
Solyc03g093760          ----------GSDTQCSVDA-DNNSCSEVE--------------------------------
Solyc03g093700          ----------GSDTPCSVDS-EIKSCSDVDMNVVESAEPLSVFEP-----------------
Solyc03g093710          ----------GCDTQSSANV-DIGSCSEVEMDVVELGDPLSVFVPK-----------------
Solyc08g077940          ----------NQNENAMVSI-S-ENMADTLVAHGDNGPNTGVEFCSVEVAS----------
Solyc06g060960          ----------NHVQNAEVST-NIEDMTNIILVDGVKETNIEVKSQSVEVVNDLINLENQE
AT2G22740_SUVH6         ----------ACENGN-VVE-DVKVVESLV--------------------------------
AT2G35160_SUVH5         RRMQKEFTVETYETRNVSDVCVLSSQADVELIPGEIVAERDSF-----------------

AT2G33290_SUVH2         --------------------------------EEWQTPPSSTSDQSANTDLIAEFIRISE
AT4G13460_SUVH9         ---------------------------DFSSDYNTVAESARSAFAQRLQRHDDVAVLDSL
Solyc12g096990          -------------------DEFTPQSMNPNSNFSYNSGFRNTTTPQQQQLNATSSQTPSS
Solyc09g090810          ------------DEYTQADLQTPAFFSNPSPNFNTSSGSAFRRNPQLATHEADSQSPSSI
AT1G17770_SUVH7         -------------------LMVPPLRTIWPSNKEWYDGDAGPSSTGPIKREASDNTNDTA
AT2G24740_SUVH8         --------------------HRERSDAVNEEAHATSIPPHAPPQTLWLLDNFNIEDSYDR
Solyc10g077070_HAP      ---------------------VCVPPSGPFPPGVSPFYPFLSPNESGRSAENQDG------
Solyc09g082050          --------------------LCITPTGPFPSGVTPIFPFLSPDEPVRMGESSQQTPNQVP
AT1G73100_SUVH3         --------------------VGCPPFGPSSSEYSSFFPFGAQQPTHDTPDLNQT------
AT5G04940_SUVH1         --------------------VCAPPFGPFPPGFSSFYPFSSSQANQHTPDLNQAQYPPQH
AT5G13960_SUVH4/KYP     -------------------------------SDRKSEICVDDTELHEKEEENVDGSP----
```

```
Solyc02g094520          ----------------------------TPSQAQTQAPNNDVTVATVDNDDVTITNVG--
Solyc03g093760          --------------------------------SAESCNFEATGNQPLKLKEENVIYDESTQ
Solyc03g093700          ------------EDDLAATVVCPKEAGDSSHQNTSCQPANGNQQHEVLVNLVLQNPSID-
Solyc03g093710          ----------DMQFDLDATGVCEEEGGDSSHLNTSCQPVTNGNQVLTTKEVNLMYDDSTQ
Solyc08g077940          ----------ARTTNVIENGLEEPTSHDKSLRFELSKDHKNSEMSLLKKAKVIGYDELGT
Solyc06g060960          NVDRLAGEVMATNMSAIANGVGEKISDEKSTGFELPKDLKTSEMELSKETEDIQNDTSVK
AT2G22740_SUVH6         ---------------KEEESLGQRDASENVSDIRMAEPVEVQPLRICLPGGDVVRDLSVT
AT2G35160_SUVH5         ------------KSVDCNDMSVGLTEGAESLGVNMQEPMKDRNMPENTSEQNMVEVHPPS

AT2G33290_SUVH2         LF----------------------------------------------------------
AT4G13460_SUVH9         ----------------------------------------------------------
Solyc12g096990          IEAGV-------------------------------------------------------
Solyc09g090810          IPEVPPGCDRNNVYV---------------------------------------------
AT1G17770_SUVH7         HN----------------------------------------------------------
AT2G24740_SUVH8         DAGPSTSPIDREASHEVNEDAHA-------------------------------------
Solyc10g077070_HAP      ----------------------------------------------------------
Solyc09g082050          NQGT--------------------------------------------------------
AT1G73100_SUVH3         ----------------------------------------------------------
AT5G04940_SUVH1         QQPQNPPPVYQQQPP---------------------------------------------
AT5G13960_SUVH4/KYP     ----------------------------------------------------------
Solyc02g094520          ----------------------------------------------------------
Solyc03g093760          HHQVQKQ-----------------------------------------------------
Solyc03g093700          ----------------------------------------------------------
Solyc03g093710          LNEVLVNQILQKTSTD--------------------------------------------
Solyc08g077940          EVDVARHFFLVENVIGMYKDHVLHPGSMTDRVIPVCDSKTLSLPQCQIKNGSVEDNISPL
Solyc06g060960          EVD-EQGLPLVESING---------GHMTQKLISVMEHTSTS------------------
AT2G22740_SUVH6         ----------------------------------------------------------
AT2G35160_SUVH5         ISLPEEDM----------------------------------------------------

AT2G33290_SUVH2         -RSAFKPLQVKGLDGVSVYGLDSGA-------------------------------IVA
AT4G13460_SUVH9         -TGAIVPVEENPEPEPNPYSTSDSS-----------------------------------
Solyc12g096990          -HSEYNRISELFQTAFAQSVQRDGD-----------------------------------
Solyc09g090810          -YSEYNRISEMFKEAFTEKMQRYGD-----------------------VEVVGNQNQDSV
AT1G17770_SUVH7         -TFAPPPEMVIPLITIRPSDDSSNY-----------------------SCDAGAGPST
AT2G24740_SUVH8         -TSAPPHVMVSPLQNRRPFDQFNNQ-------------------------PYDASAGPST
Solyc10g077070_HAP      -LGFGTPISPVPLNSFRTPAANGDT-----------------------GPRRPGRPRASN
Solyc09g082050          -FGFGQPISPIPVNSFGNQTANGSS-------------------------------GHVN
AT1G73100_SUVH3         -QNTPIPSFVPPLRSYRTPTKTNGP-------------------------------SSS
AT5G04940_SUVH1         -QHASEPSLVTPLRSFRSPDVSNGN-----------------------------------
AT5G13960_SUVH4/KYP     -KRRSPPKLTAMQKGKQKLSVSLNG-----------------------------------
Solyc02g094520          -APIDCTDHPVPENSLNPQLSGNGT-----------------------------------
Solyc03g093760          -STDTFDWFIKDEPIENGPAIVSQE-------NLIDCQNDEPSKETCQSVHREEVSDDES
Solyc03g093700          -SGNTCDWFIKSEPIENEPELPAIV-----SQENLIQGRDEPSKETSKRVHYGEVPYDEY
Solyc03g093710          -TGNTCDWFINGDPIENGPELPSEE-------------------------TNKGFQYKEV
Solyc08g077940          PKKKYCRRGVFAVRDFPPFCGRNAPKSTKLDLLGGNEASKRAILLNKGVTENEVI--ETS
Solyc06g060960          PKNKYRKRRVSAVRDFPPFCGTKVPKSTEQNCFGVTEESKDVAGFGKAVTRNEVI--ETL
AT2G22740_SUVH6         -AGDECSNSEQIVAGSGVSSSSGTE----------------------NIVRDIVVYADES
AT2G35160_SUVH5         -MGSVCRKSITGTKELHGRTISVGR--------------DLSPNMGSKFSKNGKTAKRSI

AT2G33290_SUVH2         VPEKENRELIEPPPGFKDNRV---------------------------------STVVVS
AT4G13460_SUVH9         ------------------------------------------------------PSVATQ
Solyc12g096990          VEANEDLGCRAIVPVSNGSQV----------------------------------SDIVIT
Solyc09g090810          DVVMEDADARAIVPVSNNDTQV---------------------------------AEVVVA
AT1G17770_SUVH7         GPVKRGRGRPKGSKNST--------------------------------------PTEPKK
AT2G24740_SUVH8         GPGKRGRGRPKGSKNGSRKPKKPKAYDNNSTDASAGP-----------------SSGLGK
Solyc10g077070_HAP      GLAAEDDDSQNHSDQFGSGYSGHANDVEDTSTGKK-------------------RGRPRK
Solyc09g082050          NVGDSGSGKK--------------------------------------------KGGPKK
AT1G73100_SUVH3         SGTKRGVGRPKGTTSVKKKE----------------------------------KKTVAN
AT5G04940_SUVH1         -AELEGSTVK--------------------------------------------RRIPKK
AT5G13960_SUVH4/KYP     ----------------------------------------------------------
Solyc02g094520          ----------------------------------------------------------
Solyc03g093760          RSWVDDDDISILTCSEWNSLTSALKDGKKGGKEGEIIHK---------------CSDILE
Solyc03g093700          RSRVDNDEICILSCSESNSLKSGLKTLSAGKKGGK-------------------GEIVQE
Solyc03g093710          ADDESTSRVDNSSCSQSNSQNSGLKTPSASKKGGK-------------------GEIVQE
Solyc08g077940          KNVMDTGTLSLGLTASREADSWSKTEVTGSKCSLIERAT-VRVEDPEDVQDNYVRRSQLE
Solyc06g060960          REVTETGALPEKLIGSEDADSLKDRDVSSPKDRQLEQITMVRTEEQEGVQCDYDGRSQVE
AT2G22740_SUVH6         SLGMDNLDQTQPLEIEMSDVAVAKPRLVA------------------------GRKKAK
AT2G35160_SUVH5         SVEEENLVLEKSDSGDHLGPSPEVLELEKSE----------------------VWIITD

AT2G33290_SUVH2         PKFERPRELARIA-----------------------------ILG----------------
AT4G13460_SUVH9         RPRPQPRSSELVRIT---------------------------DVG----------------
```

```
Solyc12g096990          RRKYEKRSSELVRVT-------------------------DLK----------------
Solyc09g090810          RRKYQQRSSELVRVT-------------------------DLK----------------
AT1G17770_SUVH7         PKVYDPNSLKVTSRG-------------------------NFD----------------
AT2G24740_SUVH8         RRCGRPKGLKNRSRKPKKP---------------------KADDP--------------
Solyc10g077070_HAP      TRLGQPSSGNPATPPIEV----------------------DVD----------------
Solyc09g082050          PRKVPPENVVEV----------------------------DVE----------------
AT1G73100_SUVH3         EPNLDVQVVKKFSS--------------------------DFD----------------
AT5G04940_SUVH1         RPISRPENM-------------------------------NFE----------------
AT5G13960_SUVH4/KYP     ------------------------------------------------------------
Solyc02g094520          ------------------------------------------------------------
Solyc03g093760          DFKPLPDIIRPEQ---------------------------QYE----------------
Solyc03g093700          EAVSSPEPLHKCNV--------------------------IFE----------------
Solyc03g093710          EAVKCPEPLHKCKV--------------------------IFE----------------
Solyc08g077940          RTVMLPETMTKKERDDTGKFLLKESIVYSRNEREKATTARHGFGSGDK-------------
Solyc06g060960          RTVVMPEIMTKKG-SDAGP-VGKETLVYSENEREKLTSASSALGSGNEKQITKGAKPSGA
AT2G22740_SUVH6         KGIACHSSLKVVSR--------------------------EFG----------------
AT2G35160_SUVH5         KGVVMPSPVKPSEKRNG-----------------------DYG----------------


AT2G33290_SUVH2         ------------------------------------------------------------
AT4G13460_SUVH9         ------------------------------------------------------------
Solyc12g096990          ------------------------------------------------------------
Solyc09g090810          ------------------------------------------------------------
AT1G17770_SUVH7         ------------------------------------------------------------
AT2G24740_SUVH8         ------------------------------------------------------------
Solyc10g077070_HAP      ------------------------------------------------------------
Solyc09g082050          ------------------------------------------------------------
AT1G73100_SUVH3         ------------------------------------------------------------
AT5G04940_SUVH1         ------------------------------------------------------------
AT5G13960_SUVH4/KYP     ------------------------------------------------------------
Solyc02g094520          ------------------------------------------------------------
Solyc03g093760          ------------------------------------------------------------
Solyc03g093700          ------------------------------------------------------------
Solyc03g093710          ------------------------------------------------------------
Solyc08g077940          ------------------ITKPVVHGLMDERCSPWRQKKQTPRQIVQGLMAETNKDWRQK
Solyc06g060960          RKQGKQKSLDDPVSGNEIVVSQVESHLTKTAVNAFGSGHEIVKPIVQGLMAKPCCPWRQG
AT2G22740_SUVH6         ------------------------------------------------------------
AT2G35160_SUVH5         ------------------------------------------------------------


AT2G33290_SUVH2         ------------------------------------------------------------
AT4G13460_SUVH9         ------------------------------------------------------------
Solyc12g096990          ------------------------------------------------------------
Solyc09g090810          ------------------------------------------------------------
AT1G17770_SUVH7         ------------------------------------------------------------
AT2G24740_SUVH8         ------------------------------------------------------------
Solyc10g077070_HAP      ------------------------------------------------------------
Solyc09g082050          ------------------------------------------------------------
AT1G73100_SUVH3         ------------------------------------------------------------
AT5G04940_SUVH1         ------------------------------------------------------------
AT5G13960_SUVH4/KYP     ------------------------------------------------------------
Solyc02g094520          ------------------------------------------------------------
Solyc03g093760          ----------------SVFMKKQMDLGVPQEN----------------------------
Solyc03g093700          --------------DESVVMKNQIVLGVSQED----------------------------
Solyc03g093710          ---------------HESVVRKKQIDIGVSPED----------------------------
Solyc08g077940          EQTRLDGLMSRNQVPKPSMYRQRMSVVVARKSIPKPKFPETLFGRSRSGFVGEAVPEYPS
Solyc06g060960          EPTSLD---CGNQVEKDDFSGRKKAKAVTRKSNPRGKKKSVTLGEATDGLSSALVVFNDK
AT2G22740_SUVH6         --------------EGSRKKKSKKNLYWRDRE----------------------------
AT2G35160_SUVH5         ---------------EGSMRKNSERVALD-----------------------------


AT2G33290_SUVH2         ---------------------------------------------HEQRKELRQVMKRTRMT
AT4G13460_SUVH9         ---------------------------------------------PESERQFREHVRKTRMI
Solyc12g096990          ---------------------------------------------PEDVRYFRDLIRKTRML
Solyc09g090810          ---------------------------------------------VEDQLYFREAVRKTRML
AT1G17770_SUVH7         ------------------------SEITE-------------AETETGNQEIVDSVMMR
AT2G24740_SUVH8         -------------NSKMVISCPDFDSRITE-------------AERESGNQEIVDSILMR
Solyc10g077070_HAP      -------------PLLNQLLASFKLVEIDQ-------------VKKADGDKELSGRILLV
Solyc09g082050          -------------PLLNQLLMSFKLVDLDQ-------------AKKADGDKEVVRRILLV
AT1G73100_SUVH3         ------------------------SGISA-------------AEREDGNAYLVSSVLMR
AT5G04940_SUVH1         ------------------------SGINV-------------ADRENGNRELVLSVLMR
AT5G13960_SUVH4/KYP     ---------------------------------------------KDVNLEPHLKVTKCLRL
Solyc02g094520          ---------------------------------------------EKSSHARVTETLRI
Solyc03g093760          -------------SRNSAVMCGVSGHGFST-------------EYEHIHEVKQVRKTLKL
Solyc03g093700          -------------LRNSVVMCNVSGNGLLT-------------EHEHIQKVKEVRETLKL
```

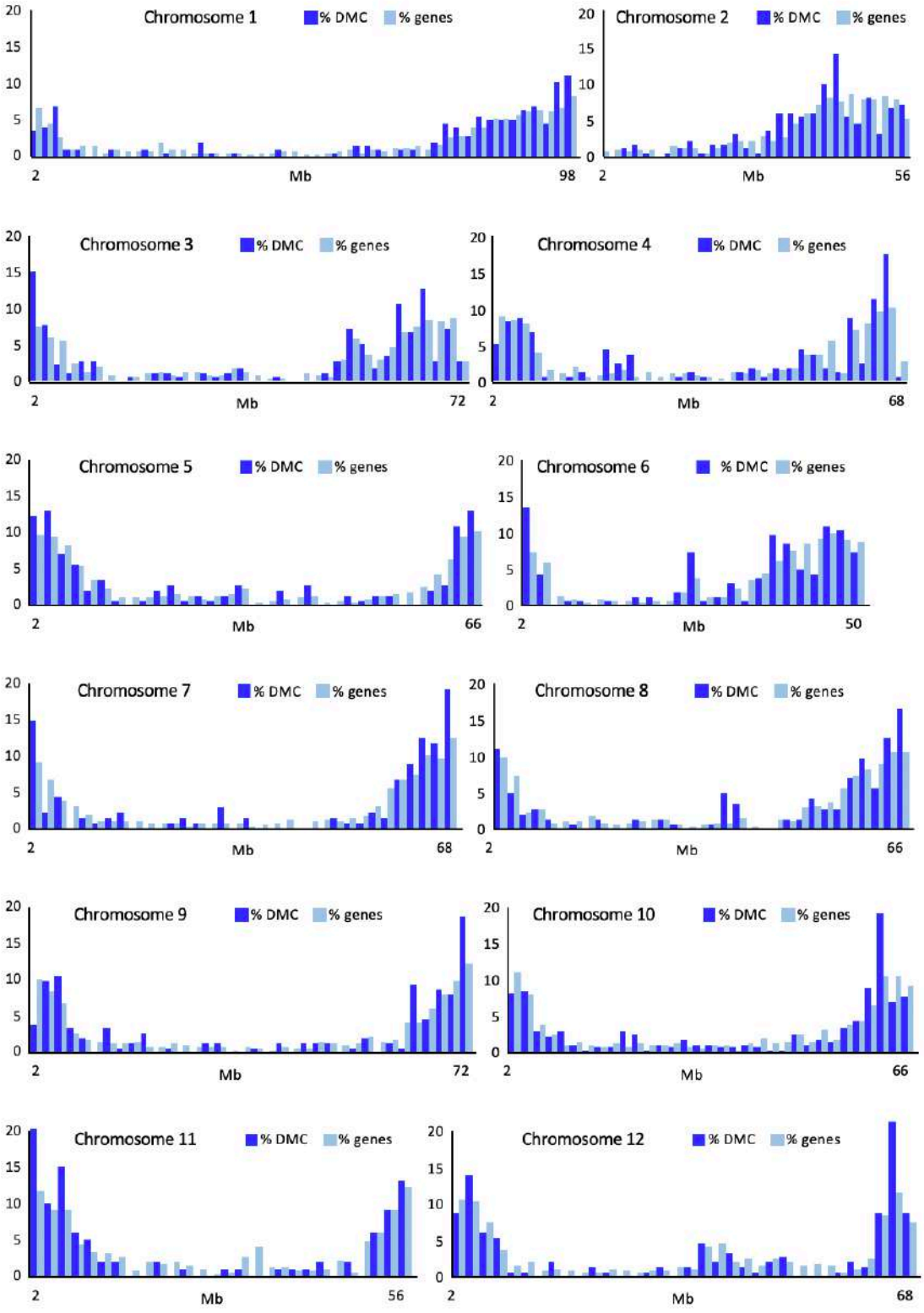
```
Solyc03g093710          -------------LRNSDVFCGASGNGLLM-------------EHENIQKVKEVKETLKL
Solyc08g077940          SP---FSKNDGIRNLNCEAQPKDSPIGQKKCEFDETRPPFGPKSSSRCDARSKVLETLRL
Solyc06g060960          GPGLWATSNDGACSLNREAVHEDSPVRRGQCDFDVTLPPFGPNSSSHGDARTKVRETLRL
AT2G22740_SUVH6         -------------SLDSPEQLRILGVGTSSG------------SSSGDSSRNKVKETLRL
AT2G35160_SUVH5         -------------KKRLASKFRLSNGGLPSC------------SSSGDSARYKVKETMRL


AT2G33290_SUVH2         YESLRIHLMAESM-----------------KNHVLGQGR--RRRSDMAAAYIMRDRGLWL
AT4G13460_SUVH9         YDSLRMFLMMEEA-----------------KRNGVGGRRARADGKAGKAGSMMRDCMLWM
Solyc12g096990          YDSLRIFVNLEDE-----------------NSQHLGSGRQTRARGDLKASQMMREHGLWL
Solyc09g090810          YDSLRILAMVEDD-----------------GSQHLGPYR--KPRGDLKACQILREHGLWM
AT1G17770_SUVH7         FDAVRRRLCQINH-----------------------------PEDILTTASGNCTKMGVKT
AT2G24740_SUVH8         FDAVRRRLCQLNY-----------------------------RKDKILTASTNCMNLGVRT
Solyc10g077070_HAP      YDLFRRRMTQIEE-----------------RRGETPGS---ARRPDLKGANLLMTRGART
Solyc09g082050          FDLFRRRMTQIDE-----------------PRYGAGS----GRRPDLKASKMMMLKGMRT
AT1G73100_SUVH3         FDAVRRRLSQVEF-----------------------------TKSATSKAAGTLMSNGVRT
AT5G04940_SUVH1         FDALRRRFAQLED-----------------AKEAVSGI---IKRPDLKSGSTCMGRGVRT
AT5G13960_SUVH4/KYP     FNKQYLLCVQAKL-----------------------------SRPDLKGVTEMIKAKAIL
Solyc02g094520          FNKHYLHFVQEEEIRCGRAQADQKTKKHSKSKEAEDDGKRSSKRPDLKAISKMISEKEVL
Solyc03g093760          FDDVYTKLLQEDK-----------------AENPEGRS---KRKIHIEAAMTLKNQKKWV
Solyc03g093700          FDDEYTKLLLEDR-----------------AEKHEGGP---KRSIHIEAAMALKKQKKWV
Solyc03g093710          FDDEYTKLLQEDK-----------------AKKHEGRS---KRRIHIEAAMNLKKQKKWV
Solyc08g077940          FQSHFRKILQGEE-------------SMSRSAGVNAKQKDKIRRIDLQAAKLVKDKGKQV
Solyc06g060960          FQGICRKLLQGEE---------------SKSKPEEAKSKQGPNRIDLHAAKIIKEKGKEV
AT2G22740_SUVH6         FHGVCRKILQEDE-----------------AKPEDQRRKGKGLRIDFEASTILKRNGKFL
AT2G35160_SUVH5         FHETCKKIMQEEE---------------ARPRKRDGG----NFKVVCEASKILKSKGKNL
                        :          .                                .


AT2G33290_SUVH2         NYDKHIVGPVTGVEVGDIFFYRMELCVLGLHGQTQAGIDCLTAERSATGE------PIAT
AT4G13460_SUVH9         NRDKRIVGSIPGVQVGDIFFFRFELCVMGLHGHPQSGIDFLTGSLSSNGE------PIAT
Solyc12g096990          NRDKRTVGPIPGVLVGDLFLYRMELCVVGLHGTPQAGIDYLPANQSSNGE------PIAT
Solyc09g090810          NRDKRIVGPIPGVLIGDVFFFRMELLVVGLHGQAQAGIDYVPASQSSNRE------PIAT
AT1G17770_SUVH7         NT-RRRIGAVPGIHVGDIFYYWGEMCLVGLHKSNYGGIDFFTAAESAVEG------HAAM
AT2G24740_SUVH8         NM-TRRIGPIPGVQVGDIFYYWCEMCLVGLHRNTAGGIDSLLAKESGVDG------PAAT
Solyc10g077070_HAP      NQ-TKRIGNVPGVEVGDIFFFRMELCLVGLHAPSMAGIDYMSVRLTGDEE------PIAV
Solyc09g082050          NQ-TKRIGNVPGIEVGDIFFFRMELCVVGLHAPTMSGIDYMSLKLTKDEE------PLAV
AT1G73100_SUVH3         NM-KKRVGTVPGIEVGDIFFSRIEMCLVGLHMQTMAGIDYIISKAGSDEE------SLAT
AT5G04940_SUVH1         NT-KKRPGIVPGVEIGDVFFFRFEMCLVGLHSPSMAGIDYLVVKGETEEE------PIAT
AT5G13960_SUVH4/KYP     YP-RKIIGDLPGIDVGHRFFSRAEMCAVGFHNHWLNGIDYMSMEYEKEYSNYKL--PLAV
Solyc02g094520          N--RERIGSLPGIDVGHQFFSRAEMVVAGFHNHWLNGIDCVGQSAGKKGEYKGYSLPLAV
Solyc03g093760          NC-EWTFGHVPGVQIGDRFRFRAELVMIGLHHQFMNGINYVNIGRK----------YVAT
Solyc03g093700          NC-EWTFGHVPGVQIGDQFRFRAELVMIGLHHQFIKGINYVTIGRK----------DVAS
Solyc03g093710          NC-EWTFGHVPGVQIGDQFRFRAELVAIGLHHQFIKGINYVTIGRK----------NVAS
Solyc08g077940          NTGTQILGEVPGVEVGDAFQYRVELSLVGVHRLYQAGIDSMYIKGGL---------LVAT
Solyc06g060960          NTGQHILGEVPGVEVGDEFQYRVELAIVGVHRLYQAGIDYMKQGGM----------LIAI
AT2G22740_SUVH6         NSGVHILGEVPGVEVGDEFQYRMELNILGIHKPSQAGIDYMKYGKA----------KVAT
AT2G35160_SUVH5         YSGTQIIGTVPGVEVGDEFQYRMELNLLGIHRPSQSGIDYMKDDGGE---------LVAT
                               * :.*: :*  *  . *:   *.*     **: .                *


AT2G33290_SUVH2         SIVVSGGYEDDEDTGDVLVYTGHGGQD--HQHK-----QCDNQRLVGGNLGMERSMHYGI
AT4G13460_SUVH9         SVIVSGGYEDDDDQGDVIMYTGQGGQD--RLGR-----QAEHQRLEGGNLAMERSMYYGI
Solyc12g096990          SIIASGGYEDDEDAGDVIIYTGQGGQD--KNSR-----QVVHQKLEGGNLALERSMYYGV
Solyc09g090810          SVIVSGGYEDDQDGGDVIIYTGHGGQD--KHSR-----QCVHQKLECGNLALERSMHYGI
AT1G17770_SUVH7         CVVTAGQYDGETEGLDTLIYSGQGGTDVYGNAR--------DQEMKGGNLALEASVSKGN
AT2G24740_SUVH8         SVVTSGKYDNETEDLETLIYSGHGG--------------KPCDQVLQRGNRALEASVRRRN
Solyc10g077070_HAP      SIVSSGGYDDEGDDGEVLIYTGQGGVQ--RRDG-----QMFDQKLERGNLALEKSMHRGN
Solyc09g082050          SIVSAGGYDDDGGDGDLLIYTGQGGVQ--RKDG-----QMFDQKLEKGNLALEKSVHRAN
AT1G73100_SUVH3         SIVSSGRYEGEAQDPESLIYSGQGGNA--DKNR-----QASDQKLERGNLALENSLRKGN
AT5G04940_SUVH1         SIVSSGYYDNDEGNPDVLIYTGQGGNA--DKDK-----QSSDQKLERGNLALEKSLRRDS
AT5G13960_SUVH4/KYP     SIVMSGQYEDDLDNADTVTYTGQGGHNLTGNKR-----QIKDQLLERGNLALKHCCEYNV
Solyc02g094520          SIVVSGQYEDDQDNYEEVVYTGQGGNDLLGNKR-----QIKDQVMERGNLGLKNCMEQSV
Solyc03g093760          SIVDSGRYDNEAISSETFIYVGQGGNPKVSINA-----RVEDQKLKGGNLALKNSMDMGC
Solyc03g093700          SIVDSGRYDNEAISSETFIYVGQGGNPKVSVNA-----RMEDQKLEGGNLALKNSMELGY
Solyc03g093710          SVVDSSRYDNEAISSETFIYVGQGGNPMVSLNG-----RVEDQKLEGGNLALKNSMDLGY
Solyc08g077940          SIVASGAYDDDDLGDADELIYSGQGGNV-VGKVK-----IPEDQKLVKGNLALKNSIRERN
Solyc06g060960          SIVSSGVYDDGLEDADVLIYSGQGGNV-VGKSK-----TPEDQKLERGNLALKNSISVKN
AT2G22740_SUVH6         SIVASGGYDDHLDNSDVLTYTGQGGNVMQVKKKGEELKEPEDQKLITGNLALATSIEKQT
AT2G35160_SUVH5         SIVSSGGYNDVLDNSDVLIYTGQGGNV--GKKKNNE--PPKDQQLVTGNLALKNSINKKN
                        .:: :. *:.     : . * *:**                * :  ** .:  .


AT2G33290_SUVH2         EVRVIRGIKYENSI---SSK-----VYVYDGLYKIVDWWFAVGKSG-FGVFKFRLVRIEG
AT4G13460_SUVH9         EVRVIRGLKYENEV---SSR-----VYVYDGLFRIVDSWFDVGKSG-FGVFKYRLERIEG
Solyc12g096990          EVRVIRGFKYVGSS---SGK-----VYVYDGLYRITESWFDVGKSG-FGVYKYKLVRIEN
Solyc09g090810          EVRVIRGFKYEGSGSA-SGK-----VYVYDGLYRIVECWFDVGKSG-FGVYKYKLVRIEN
AT1G17770_SUVH7         DVRVVRGVIHPHEN---NQK-----IYIYDGMYLVSKFWTVTGKSG-FKEFRFKLVRKPN
```

```
AT2G24740_SUVH8          EVRVIRGELYN------NEK-----VYIYDGLYLVSDCWQVTGKSG-FKEYRFKLLRKPG
Solyc10g077070_HAP       EVRVIRGVVDVQNGG--RGK-----IYMYDGLYRVQESWAEKSKLGNCSIFRYKLIRVPG
Solyc09g082050           DVRVIRGVKDVANP---TGK-----IYIFDGLYKIQGSWEEKIKTG-CNVFKYKLLRVPG
AT1G73100_SUVH3          GVRVVRGEEDAASK---TGK-----IYIYDGLYSISESWVEKGKSG-CNTFKYKLVRQPG
AT5G04940_SUVH1          AVRVIRGLKEASH----NAK-----IYIYDGLYEIKESWVEKGKSG-HNTFKYKLVRAPG
AT5G13960_SUVH4/KYP      PVRVTRGHNCKSSY---TKR-----VYTYDGLYKVEKFWAQKGVSG-FTVYKYRLKRLEG
Solyc02g094520           PVRVTRGHRCVNSY---VGK-----VYTYDGLYKVVNYWAEKGISG-FTVYKFRLKRIEG
Solyc03g093760           PVRVICGRKRVNGEKS-DIR------YIYDGLYTVTKCWEEIAPTG-KYVFKFELKRNPG
Solyc03g093700           PVRVICGRQRVNGEKS-DTR------YIYDGLYTVTKCWEERAPTG-KYIFKFELKRNLG
Solyc03g093710           PVRVICGRQRLNGEKS-DTR------YIYDGLYTVTKCWEERASTE-KYIFKFELKRNLG
Solyc08g077940           SVRVIRGSKEIRTPES-GGRPNVVTTYVYDGLYTVENYWKEKGPHG-KMVFMFKLVRIPG
Solyc06g060960           PVRVIRGSKETKNSDSVDGKGKLVTTYVYDGLYTVENYWTEQGTKG-KMVFMFKLVRVPG
AT2G22740_SUVH6          PVRVIRGKHKSTHDKS-KGG-----NYVYDGLYLVEKYWQQVGSHG-MNVFKFQLRRIPG
AT2G35160_SUVH5          PVRVIRGIKNTTLQSSVVAK-----NYVYDGLYLVEEYWEETGSHG-KLVFKFKLRRIPG
                          ***  *                  * :**:: :   *          : : * *  .

AT2G33290_SUVH2          QPMMGSAVMRFAQTL---------RNKPSMVRPTGYVSFDLSNKKENVPVFLYNDVDGDQ
AT4G13460_SUVH9          QAEMGSSVLKFARTL---------KTNPLSVRPRGYINFDISNGKENVPVYLFNDIDSDQ
Solyc12g096990           QPDMGSAILRFAESL---------RTRPLEVRPMGYISLDISRKKENVPVFLFNDIDNER
Solyc09g090810           QEEMGSAILRFAQNL---------RIRPLEARPTGYVTLDISRKKENVPVFLFNDIDDNH
AT1G17770_SUVH7          QPPAYAIWKTV-ENL---------RNHDLIDSRQGFILEDLSFGAELLRVPLVNEVDEDD
AT2G24740_SUVH8          QPPGYAIWKLV-ENL---------RNHELIDPRQGFILGDLSFGEEGLRVPLVNEVDEED
Solyc10g077070_HAP       QPEAYTLWKSV-QQW---------REG--TATRVGVILPDLTSGAESQPVCLVNDVDDEK
Solyc09g082050           QPEAFKVWKSI-QQW---------RDG--VVSRVGVILPDLTSGAESQAVCLVNDVDDEK
AT1G73100_SUVH3          QPPAFGFWKSV-QKW---------KEG--LTTRPGLILPDLTSGAESKPVSLVNDVDEDK
AT5G04940_SUVH1          QPPAFASWTAI-QKW---------KTG--VPSRQGLILPDMTSGVESIPVSLVNEVDTDN
AT5G13960_SUVH4/KYP      QPELTTDQVNF-VAG---------RIPTSTSEIEGLVCEDISGGLEFKGIPATNRVDDSP
Solyc02g094520           QPVLTTNQVHF-TRG---------CTPNSISEIRGLVCEDISGGLEDIPIPATNLVDDPP
Solyc03g093760           QPKLNREVVSRPTSLGKVDHFHVNKATKSIMESEFVVDNDVSQGKEKIPICVVNAIDDER
Solyc03g093700           QPKLARELVSRPAKLVKVNQFCVNKAKKSILQSEFVVDYDVSQGKEKIPILVVNAIDDER
Solyc03g093710           QPKLNRELVSRPAKLVKVTHSCVNKSTKSVMQSEFVVDYDVSQGKEKIPIRVVNAIDDER
Solyc08g077940           QPEL--TWKEV-QSS---------KN---SKARHGVCVPDITEGKESLPIAAVNTIDGEK
Solyc06g060960           QPEL--AWKEV-KSS---------RK---SKVRHGVCVHDITDGKETFAISAVNTIDGEK
AT2G22740_SUVH6          QPEL--SWVEV-KKS---------K----SKYREGLCKLDISEGKEQSPISAVNEIDDEK
AT2G35160_SUVH5          QPEL--PWKEV-AKS---------KK---SEFRDGLCNVDITEGKETLPICAVNNLDDEK
                         *                                    *::   *   :   * :*

AT2G33290_SUVH2          EPR--HYEYIAKAVFPP--GIFGQGGISRTGC-ECKLSCT--D--DCLCARKNGGEFAY-
AT4G13460_SUVH9          EPL--YYEYLAQTSFPP--GLFVQQSGNASGC-DCVNGCG----SGCLCEAKNSGEIAY-
Solyc12g096990           DPA--CYDYLLKTVFPP--YVY-QHVGNGSGC-ECTDGCG--NGTNCFCAMKNGGQFAY-
Solyc09g090810           DPA--YFEYLVKPIYPP--HVS-LNVHSGNGC-QCIDGCA--D--NCFCAMRNGGQFAY-
AT1G17770_SUVH7          KTIPEDFDYIPSQCHSGMMTHEFHFDRQSLGCQNCRHQPC--MHQNCTCVQRNGDLLPY-
AT2G24740_SUVH8          KTIPDDFDYIRSQCYSG--------MTNDVNV-DSQSLVQSYIHQNCTCILKNCGQLPY-
Solyc10g077070_HAP       GPA--YFTYIPSLKYSK--PFM--KSNPSVGC-QCLGGCQ-PGGTSCPCIQKNGGYLPF-
Solyc09g082050           GPA--YFTYIPSLKYSK--PFL--TPRPSLGC-QCIGGCQ-PDDTNCPCIQRNQGLLPY-
AT1G73100_SUVH3          GPA--YFTYTSSLKYSE--TFK--LTQPVIGC-SCSGSCS-PGNHNCSCIRKNDGDLPY-
AT5G04940_SUVH1          GPA--YFTYSTTVKYSE--SFK--LMQPSFGC-DCANLCK-PGNLDCHCIRKNGGDFPY-
AT5G13960_SUVH4/KYP      VSPTSGFTYIKSLIIEP--NVI--IPKSSTGC-NCRGSCT--DSKKCACAKLNGGNFPYV
Solyc02g094520           AAP-SGFTYSRDIVCAK--GIK--FPSAPTGC-NCHGSCL--DPRVCSCAKLNGSEFPYV
Solyc03g093760           LPS---FTYITSIRYPD--WYY---ISKPQGC-NCTSGCS--DSEQCSCASRNGGEIPF-
Solyc03g093700           PSP---FTYITSMQYPD--WYY---ISRPQGC-NCTSGCL--DSEQCSCASRNGGEIPF-
Solyc03g093710           LPP---FTYITNMQYPD--WYY---ISRPQGC-NCTSGCS--DSEQCSCASRNGGEIPF-
Solyc08g077940           PPP---FKYIKNMMYPV--GFR---PAPPRGC-DCIGRCS--DAERCSCAVKNGGEIPY-
Solyc06g060960           PPP---FNYIQKIIYPD--WFQ---PSPFKGC-DCIGRCS--DSKKCSCAVKNGGEIPY-
AT2G22740_SUVH6          PPL---FTYTVKLIYPD--WCR---PVPPKSC-CCTTRCTEAEARVCACVEKNGGEIPY-
AT2G35160_SUVH5          PPP---FIYTAKMIYPD--WCR---PIPPKSC-GCTNGCS--KSKNCACIVKNGGKIPY-
                          .    : *                    .   .           * *   * . :.:

AT2G33290_SUVH2          -DDNGHLLKGKHVVFECGEFCTCGPSCKSRVTQKGLRNRLEVFRSKETGWGVRTLDLIEA
AT4G13460_SUVH9          -DYNGTLIRQKPLIHECGSACQCPPSCRNRVTQKGLRNRLEVFRSLETGWGVRSLDVLHA
Solyc12g096990           -DTNGILLRGKPIIFECGPHCSCPPTCLNRVSQKGVRNRFEVFRSRETDWGVRSLDLLQA
Solyc09g090810           -DYNGILLRGKPLVFECGPHCRCPPTCRNRVTQKGLRNRFEVFRSRETGWGVRSLDLIQA
AT1G17770_SUVH7          --HNNILVCRKPLIYECGGSCPCPDHCPTRLVQTGLKLHLEVFKTRNCGWGLRSWDPIRA
AT2G24740_SUVH8          --HDNILVCRKPLIYECGGSCP------TRMVETGLKLHLEVFKTSNCGWGLRSWDPIRA
Solyc10g077070_HAP       -NPLGVLMSYKTLVYECGSACSCPPNCRNRITQAGPKARVEVFKTKNRGWGLRSWDPIRG
Solyc09g082050           -NSLGVLMTYKNLIHECGSACSCPANCRNRMSQGGPKVRMEVFKTKNKGWGLRSWDPIRG
AT1G73100_SUVH3          -LNGVILVSRRPVIYECGPTCPCHASCKNRVIQTGLKSRLEVFKTRNRGWGLRSWDSLRA
AT5G04940_SUVH1          -TGNGILVSRKPMIYECSPSCPC-STCKNKVTQMGVKVRLEVFKTANRGWGLRSWDAIRA
AT5G13960_SUVH4/KYP      DLNDGRLIESRDVVFECGPHCGCGPKCVNRTSQKRLRFNLEVFRSAKKGWAVRSWEYIPA
Solyc02g094520           HKDGGRLIEPKAVVFECGPNCGCGPACVNRTSQKGLRYRLEVFRTPNKGWGVRSWDYIPS
Solyc03g093760           -NTRGSIIRAQPLVYECGPSCKCPPSCKNRVSQHGPRDHLEVFKTESRGWGLRSRDRVSS
Solyc03g093700           -NTRGSIVRAQPLVYECGPSCKCPPSCKNRVSQHGPRYHLEVFKTESRGWGLRSRDYVTR
Solyc03g093710           -NTRGSIVRAQPLVYECGPSCKCPPSCKNRVSQHGPRYHLEVFKTESRGWGLRSRDHVSS
Solyc08g077940           -NRNGAIVEVKPLVYECGPHCKCPPSCYNRVSQHGIKIPLEIFKTDTRGWGVRALTSISS
Solyc06g060960           -NRNGAIVEVKPLVYECGPHCKCPPSCYNRVSQHGIKVPLEIFKTNSRGWGVRALTSIPS
```

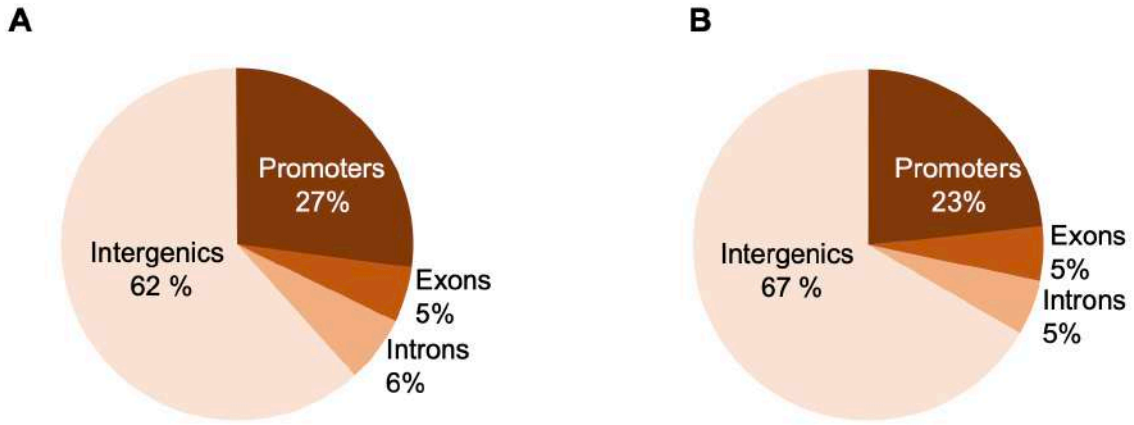
```
AT2G22740_SUVH6          -NFDGAIVGAKPTIYECGPLCKCPSSCYLRVTQHGIKLPLEIFKTKSRGWGVRCLKSIPI
AT2G35160_SUVH5          --YDGAIVEIKPLVYECGPHCKCPPSCNMRVSQHGIKIKLEIFKTESRGWGVRSLESIPI
                               ::  .  :.**.  *        .  :   .  .*:*.:   .*.:*    :

AT2G33290_SUVH2          GAFICEYAGVVVTRLQAE-ILSMNGDVMVYPG--RFT----D-QWRNWGDLSQVYPDFVR
AT4G13460_SUVH9          GAFICEYAGVALTREQAN-ILTMNGDTLVYPA--RFS----SARWEDWGDLSQVLADFER
Solyc12g096990           GSFICEYTGVVLTQEQAQ-IFTMNGDSLIYPS--HFA----E-RWAEWGDLSRIDSNYAR
Solyc09g090810           GSFICEYTGVVLTREQAQ-IFTMNGDSLVYPS--RFP----D-RWAEWGDLSQIYPNYER
AT1G17770_SUVH7          GTFICEFAGLRKTKEEVE-----EDDDYLFDTSKIYQ----RFRWNYEPELLLEDSWEQV
AT2G24740_SUVH8          GTFICEFTGVSKTKEEVE-----EDDDYLFDTSRIYH----SFRWNYEPELLCEDACEQV
Solyc10g077070_HAP       GGFVCEYAGEVIEESRVGEFGNDGDDDYIFDATRMYE----P----------LEAVRDYN
Solyc09g082050           GCFICEYAGEVRDIGYDR------DDNYIFDATRIYE----P----------LEAVHDYN
AT1G73100_SUVH3          GSFICEYAGEVKDNGNLR--GNQEEDAYVFDTSRVFN----SFKWNYEPELVDEDPSTEV
AT5G04940_SUVH1          GSFICIYVGEAKDKSKVQ--QTMANDDYTFDTTNVYN----PFKWNYEPGLADEDACEEM
AT5G13960_SUVH4/KYP      GSPVCEYIGVVRRTADVD---TISDNEYIFEIDCQQTMQGLGGRQRRLRDVAVPMNNGVS
Solyc02g094520           GATICEYTGLLKKTDQID---PAADNNYVFDIDCLQTMKGLDGRERRLREVSLPGYWHND
Solyc03g093760           GSFICEYVGELLDEKEAE--SRIDNDEYLFDVG-NYD----EEIPKRNPMRNNNLKVESD
Solyc03g093700           GRFICEYVGELLDEKEAE--NRIGHDEYLFDIG-NYD----EEIPKRNVARNNNLKVESN
Solyc03g093710           GSFICEYVGELLDEKEAE--NRIDNDEYLFDIG-NYD----EEIPKRNVARNNNLKVDSN
Solyc08g077940           GTFICEYTGQLLEDTEAE--RRIGMDEYLFDIGQNYG----GYTANSSGQANQNELVEEG
Solyc06g060960           GTFICEYVGELLEDKEAE--QRIGSDEYLFDIGQNYS----DCSVNSSRQAEVSEVVEEG
AT2G22740_SUVH6          GSFICEYVGELLEDSEAE--RRIGNDEYLFDIGNRYD----NSLAQGMSELMLGTQAGRS
AT2G35160_SUVH5          GSFICEYAGELLEDKQAE--SLTGKDEYLFDLG----------------------------
                         *  :* : *               :   :

AT2G33290_SUVH2          PNYPSLPPLDFSMDVSRMRNVACYISHSKEPNVMVQFVLH-DHNHLMFPRVMLFALENIS
AT4G13460_SUVH9          PSYPDIPPVDFAMDVSKMRNVACYISHSTDPNVIVQFVLH-DHNSLMFPRVMLFAAENIP
Solyc12g096990           PAYPSIPPLDFAMDVSRMRNLACYMSHSSSPNVLVQPVLY-DHNNVSFPHLMLFAMENIP
Solyc09g090810           PAYPSIPPLDFAMDVSRMRNVACYISHSSSPNALVQPVLY-DHNHVAFPHMMLFAMENIP
AT1G17770_SUVH7          SEFINLPTQ-VLISAKEKGNVGRFMNHSCSPNVFWQPIEY-ENRGDVYLLIGLFAMKHIP
AT2G24740_SUVH8          SEDANLPTQ-VLISAKEKGNVGRFMNHNCWPNVFWQPIEYDDNNGHIYVRIGLFAMKHIP
Solyc10g077070_HAP       DESKKVPYP-LVISAKKGGNVARFMNHSCSPNVYWQLVVR-EINNETFYHVAFFAIRHIP
Solyc09g082050           DESRKVPFP-LVISSKNGGNIARFMNHSCSPNVYWQLVVR-ESNNEAYYHIAFFAIRHIP
AT1G73100_SUVH3          PEEFNLPSP-LLISAKKFGNVARFMNHSCSPNVFWQPVIR-EGNGESVIHIAFFAMRHIP
AT5G04940_SUVH1          SEESEIPLP-LIISAKNVGNVARFMNHSCSPNVFWQPVSY-ENNSQLFVHVAFFAISHIP
AT5G13960_SUVH4/KYP      QSSEDENAPEFCIDAGSTGNFARFINHSCEPNLFVQCVLS-SHQDIRLARVVLFAADNIS
Solyc02g094520           SEKMSDGGPEYCIDAVSVGNVARFINHSCQPNLFVQCVLS-THHDIGLARVVLMAADNIP
Solyc03g093760           SLGRKDEDG-FALDAVRYGNVGRFINHSCSPNLYAQNVMY-YHGDRRVPHIMFFASKSIA
Solyc03g093700           SLTRKDEDG-FTLDALRYGNVGRFINHSCSPNLYAQNVMY-YHGDKKVPHIMFFASESIA
Solyc03g093710           SSMRKDEDG-FTLDAIRYGNVGRFINHSCSPNLYAQNVMY-YHGDKKVPHIMFFASESIA
Solyc08g077940           G---------YTIDAARYGNVGRFINHSCSPNLYAQNVVY-DHKDKRVPHIMLFAADNIP
Solyc06g060960           ----------YTIDAAQYGNIGRFINHSCSPNLYAQSVLY-DHEDKKMPHIMLFAADNIP
AT2G22740_SUVH6          MAEGDESSG-FTIDAASKGNVGRFINHSCSPNLYAQNVLY-DHEDSRIPHVMFFAQDNIP
AT2G35160_SUVH5          ----DEDDP-FTINAAQKGNIGRFINHSCSPNLYAQDVLY-DHEEIRIPHIMFFALDNIP
                                    :.     *.. ::.*.  **   * :            : ::*   *.

AT2G33290_SUVH2          PLAELSLDYGLADE-----------VNGKLAICN--------------------------
AT4G13460_SUVH9          PMTELSLDYGVVDD-----------WNAKLAICN--------------------------
Solyc12g096990           PLRELSIDYGMPDD-----------CTGKLAICN--------------------------
Solyc09g090810           PLKEISIDYGVADE-----------WTGNLKMMRQSCFSIPVKLKIVMTFIVVFLLQLPI
AT1G17770_SUVH7          PMTELTYDYGVSCVERSEEDEVLLYKGKKTCLCGSV-------------------------
AT2G24740_SUVH8          PMTELTYDYGISCVEKTGEDEVIY-KGKKICLCGSV-------------------------
Solyc10g077070_HAP       PMQELTFDYGMVPPDKAD-------RRRKKCLCGSL-------------------------
Solyc09g082050           PLQELTFDYGMDKAD----------HRRKKCLCGSF-------------------------
AT1G73100_SUVH3          PMAELTYDYGISPTSEARDESLL--HGQRTCLCGSE-------------------------
AT5G04940_SUVH1          PMTELTYDYGVSRPSGTQNGNPL--YGKRKCFCGSA-------------------------
AT5G13960_SUVH4/KYP      PMQELTYDYGYALDSVHGPDGK---VKQLACYCGAL-------------------------
Solyc02g094520           PLQELTYDYGYVLDSVMDREGK---VKQMACYCGAA-------------------------
Solyc03g093760           PFEEFTYHYNYG--HVYDKNSN---MKRKNCICGSQ-------------------------
Solyc03g093700           PLEELTYHYNYDVDQVSDKNGD---MKRKNCRCGSR-------------------------
Solyc03g093710           PLKELTYHYNYHIDHVYDKNGD---VKRKNCRCGSR-------------------------
Solyc08g077940           PLKELSYHYNYVVDQVYDSDGK---IKVKRCFCGSS-------------------------
Solyc06g060960           PLAELSYHYNYSVDQVHDSKGN---IKVKKCFCGSS-------------------------
AT2G22740_SUVH6          PLQELCYDYNYALDQVRDSKGN---IKQKPCFCGAA-------------------------
AT2G35160_SUVH5          PLQELSYDYNYKIDQVYDSNGN---IKKKFCYCGSA-------------------------
                         *: *:   *.

AT2G33290_SUVH2          ------------------------------------------------------------
AT4G13460_SUVH9          ------------------------------------------------------------
Solyc12g096990           ------------------------------------------------------------
Solyc09g090810           TFVLNLFVDYLRPISLKYCASLKFIWLWRLLDL----------------------------
AT1G17770_SUVH7          -----------------KCRGSFT--------------------------------------
AT2G24740_SUVH8          -----------------KCRGSFG--------------------------------------
Solyc10g077070_HAP       -----------------NCRGYFY--------------------------------------
Solyc09g082050           -----------------KCRGYFY--------------------------------------
```

```
AT1G73100_SUVH3          -----------------QCRGSFG-------------------------------------
AT5G04940_SUVH1          -----------------YCRGSFG-------------------------------------
AT5G13960_SUVH4/KYP      -----------------NCRKRLY-------------------------------------
Solyc02g094520           -----------------DCRKRLF-------------------------------------
Solyc03g093760           -----------------KCEGRMY-------------------------------------
Solyc03g093700           -----------------KCEGRMY-------------------------------------
Solyc03g093710           -----------------KCEGRMY-------------------------------------
Solyc08g077940           -----------------DCSGRMY-------------------------------------
Solyc06g060960           -----------------ECSGKNFHCNLFVLLTLMFAFDRINMYLFVSISSFPIAKKVGY
AT2G22740_SUVH6          -----------------VCRRRLY-------------------------------------
AT2G35160_SUVH5          -----------------ECSGRLY-------------------------------------


AT2G33290_SUVH2          --------------------
AT4G13460_SUVH9          --------------------
Solyc12g096990           --------------------
Solyc09g090810           --------------------
AT1G17770_SUVH7          --------------------
AT2G24740_SUVH8          --------------------
Solyc10g077070_HAP       --------------------
Solyc09g082050           --------------------
AT1G73100_SUVH3          --------------------
AT5G04940_SUVH1          --------------------
AT5G13960_SUVH4/KYP      --------------------
Solyc02g094520           --------------------
Solyc03g093760           --------------------
Solyc03g093700           --------------------
Solyc03g093710           --------------------
Solyc08g077940           --------------------
Solyc06g060960           IVGKVHTDFAGKEKFAQGGS
AT2G22740_SUVH6          --------------------
AT2G35160_SUVH5          --------------------
```

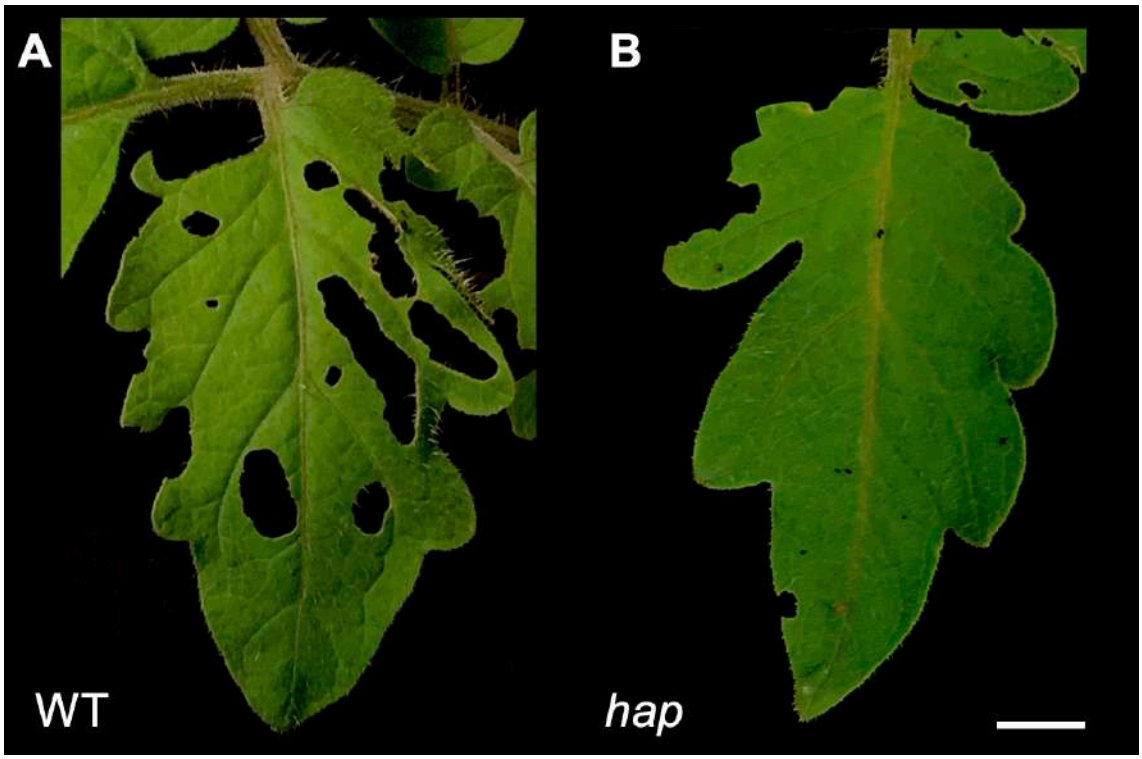
Supplementary Fig. 7: Multiple alignment of HAP and other Arabidopsis and tomato SET domain containing proteins. The blue box outlines the 8 amino acid region located inside the SET domain and shared exclusively by HAP and its closely related proteins.


Chromosome 1
% DMC
% genes
20
15
10
5
0
2
Mb
98
Chromosome 2
% DMC
% genes
20
15
10
5
0
2
Mb
56
Chromosome 3
% DMC
% genes
20
15
10
5
0
2
Mb
72
Chromosome 4
% DMC
% genes
20
15
10
5
0
2
Mb
68
Chromosome 5
% DMC
% genes
20
15
10
5
0
2
Mb
66
Chromosome 6
% DMC
% genes
20
15
10
5
0
2
Mb
50
Chromosome 7
% DMC
% genes
20
15
10
5
0
2
Mb
68
Chromosome 8
% DMC
% genes
20
15
10
5
0
2
Mb
66
Chromosome 9
% DMC
% genes
20
15
10
5
0
2
Mb
72
Chromosome 10
% DMC
% genes
20
15
10
5
0
2
Mb
66
Chromosome 11
% DMC
% genes
20
15
10
5
0
2
Mb
56
Chromosome 12
% DMC
% genes
20
15
10
5
0
2
Mb
68

Supplementary Fig. 8: Distribution in the 12 tomato chromosomes of the DMC caused by the loss of function of *HAP*. The dark blue histograms represent the percentage of DMC every 2 Mb whereas the light blue histograms show the density of genes.



Supplementary Fig. 9: Location of DMCs caused by the loss of function of *HAP*. A. Location of hypomethylated CpG, **B.** differential methylation annotation of hypermethylated CpG.



Supplementary Fig. 10: Damage caused by the tomato fruit borer *Helicoverpa armigera* (Hübner) after 15 days feeding in WT and *hairplus* mutant plants. A. WT plants show severe damage by the larvae feeding, which causes large lesions in the leaf. **B.** *hap* mutant plants on the other hand exhibits higher resistance levels, since larvae feed only in the leaf margin.

Supplementary Table 1. Trichome density accounts in leaves, stem and inflorescence stem of wild type and *hap* mutant plants.

Analysed Tissues	Type I	Type III	Type VI	Type VII	Type V	Stomata
Abaxial WT	0,0 ± 0,0	201,8 ± 18,0	8,0 ± 3,7	0,6 ± 0,9	3,6 ± 1,5	45,6 ± 9,9
Abaxial *hap*	0,2 ± 0,4	167,2 ± 39,2	15,0 ± 3,5 *	1,0 ± 1,0	5,2 ± 2,6	53,2 ± 12,0
Adaxial WT	0,4 ± 0,5	70,6 ± 12,9	14,4 ± 8,3	2,2 ± 1,8	1,4 ± 2,6	7,2 ± 3,0
Adaxial *hap*	1,0 ± 0,0 *	54,4 ± 12,0	12,8 ± 4,3	5,8 ± 1,5 **	1,6 ± 1,5	8,6 ± 1,3
Stem WT	0,4 ± 0,5	61,4 ± 24,2	29,4 ± 9,3	2,4 ± 0,5	12,0 ± 7,8	0,8 ± 0,8
Stem *hap*	3,0 ± 1,4 **	47,4 ± 17,6	20,4 ± 7,6	12,4 ± 3,9	17,6 ± 10,0	1,2 ± 0,4
Inflorescence Stem WT	2,8 ± 0,8	255,6 ± 64,9	66,8 ± 16,4	48,2 ± 15,2	33,4 ± 15,0	6,8 ± 4,8
Inflorescence Stem *hap*	8,0 ± 2,3 **	123,8 ± 49,3 **	54,0 ± 19,5	71,2 ± 11,7 *	55,0 ± 25,0	8,0 ± 1,2

Values are expressed as the mean ± standard deviation. Statistical analyses were performed with Fischer´s LSD method. * Significant differences at $P < 0.05$. ** Significant differences at $P < 0.01$.

Supplementary Table 2. Trichome density accounts in leaves, stem and inflorescence stem of wild type, *hap* mutant plants and transgenic lines.

Trichome Type/Tissue		WT	*hap*	RNAi 3	RNAi 15	CRISPR 1	OX::MM	OX::*hap*
Type I	Inflorescence stem	3,8 ± 1,3	8,2 ± 1,6 **	7,0 ± 1,4 **	10,2 ± 1,1 ***	16,0 ± 1,2 ***	4,2 ± 1,1	1,6 ± 0,9 *
	Stem	0,2 ± 0,4	0,2 ± 0,4	0,4 ± 0,5	1,4 ± 0,9 *	0,4 ± 0,5	0,2 ± 0,4	0,2 ± 0,4
	Leaf Adaxial	0,4 ± 0,5	0,0 ± 0,0	0,0± 0,0	0,2 ± 0,4	1,6 ± 2,5	1,8 ± 1,1 *	0,0 ± 0,0
	Leaf Abaxial	0,2 ± 0,4	1,0 ± 1,2	0,2 ± 0,2	1,2 ± 0,8*	0,6 ± 0,9	-	0,2 ± 0,4
Type III	Inflorescence stem	95,4 ± 18,4	28,6 ± 7,2 ***	93,0 ± 23,6	24,0 ± 6,6 ***	98,4 ± 14,4	132,6 ± 11,1 **	100 ± 20,8
	Stem	17,8 ± 4,7	40,0 ± 5,0 ***	37,2 ± 11,1**	47,2 ± 24,2 *	22,0 ± 2,0	19,0 ± 3,5	12,0 ± 3,6
	Leaf Adaxial	24,8 ± 8,2	7,6 ± 2,9 **	7,2 ± 5,8	69,2 ± 19,4 **	30,2 ± 3,0	33,0 ± 3,5	53,2 ± 6,2 ***
	Leaf Abaxial	54,8 ± 6,0	56,0 ± 5,1	43,4 ± 6,1 *	-	67,6 ± 14,6	-	67,2 ± 13,3
Type VI	Inflorescence stem	34,0 ± 4,9	46,6 ± 14,5	24,0 ± 6,1 *	29,6 ± 7,7	10,4 ± 1,5 ***	8,0 ± 2,7 ***	25,4 ± 11,9
	Stem	6,2 ± 2,2	9,6 ± 1,8 *	7,6 ± 1,7	5,6 ± 2,1	4,0 ± 1,0	3,8 ± 1,6	3,8 ± 1,1
	Leaf Adaxial	6,6 ± 4,0	2,4 ± 1,1	1,0 ± 1,0 *	1,0 ± 1,2 *	1,4 ± 1,1 *	2,0 ± 1,2 *	5,2 ± 3,6
	Leaf Abaxial	4,6 ± 2,3	7,8 ± 2,4	3,4 ± 2,7	2,2 ± 1,9	0,2 ± 0,4 **	-	1,4 ± 1,1 *
Type VII	Inflorescence stem	30,4 ± 5,0	35,2 ± 8,6	35,8 ± 7,2	38,6 ± 1,8 **	31,0 ± 5,1	5,6 ± 1,5 ***	4,2 ± 1,6 ***
	Stem	1,4 ± 0,5	3,0 ± 1,4 *	5,8 ± 0,8 ***	6,6 ± 2,6 **	0,6 ± 0,9	0,6 ± 0,9	1,6 ± 0,9
	Leaf Adaxial	0,6 ± 0,9	0,0 ± 0,0	0,8 ± 0,8	10,0 ± 2,0 ***	0,4 ± 0,9	0,8 ± 0,8	2,8 ± 1,6 *
	Leaf Abaxial	0,2 ± 0,4	1,0 ± 1,0	1,0 ± 1,0	1,6 ± 1,5	0,4 ± 0,5	-	0,0 ± 0,0
Type V	Inflorescence stem	12,6 ± 1,9	60,6 ± 17,4 ***	34,8 ± 5,4 ***	32,4 ± 9,3 **	52,6 ± 5,7 ***	2,2 ± 0,8 ***	3,2 ± 1,9 ***
	Stem	1,0 ± 1,2	5,6 ± 2,2 **	10,6 ± 0,9 ***	16,8 ± 4,8 ***	0,0 ± 0,0	0,4 ± 0,5	1,0 ± 1,2
	Leaf Adaxial	1,0 ± 1,0	0,0 ± 0,0	0,4 ± 0,5	2,4 ± 1,5	0,0 ± 0,0	0,2 ± 0,4	5,8 ± 1,9 **
	Leaf Abaxial	2,0 ± 2,3	4,4 ± 3,4	3,8 ± 3,6	1,4 ± 1,7	0,8 ± 0,8	-	2,4 ± 3,9

Values are expressed as the mean ± standard deviation. Statistical analyses were performed with Fischer´s LSD method.

* Significant differences at $P < 0.05$. ** Significant differences at $P < 0.01$. *** Significant differences at $P < 0.001$.

Supplementary Table 3. Summary of up regulated genes detected in a RNAseq transcriptomic analyses. RPKM values are listed for each of three replicates of wild type and hap mutant plants

Over-expressed genes	MT01	MT02	MT03	WT01	WT02	WT03	Description
Solyc02g077880	782.30	443.88	765.93	138.70	158.96	207.32	Auxin-repressed protein
Solyc02g085910	35.48	2.22	4.39	0.37	0.59	0.47	Lateral organ boundaries (LOB), LOB domain protein 42
Solyc03g020060	1.85	24.50	4.10	0	0	0.27	Proteinase inhibitor I20, Pin2
Solyc03g020070	1.58	28.87	6.78	0.14	0.08	0.15	Proteinase inhibitor I20, Pin2
Solyc03g020080	5.53	42.27	7.31	0	0	0.08	Proteinase inhibitor I20, Pin2
Solyc03g082550	8.56	3.32	3.34	0.57	1.15	0.85	Homeobox leucine zipper protein
Solyc08g065940	217.31	131.50	157.06	58.54	35.83	45.96	Zinc finger CCCH domain-containing protein 20
Solyc12g096570	143.37	50.40	54.20	10.92	20.34	14.02	ARGOS

Supplementary Table 4. Summary of down regulated genes detected in a RNAseq transcriptomic analyses. RPKM values are listed for each of three replicates of wild type and hap mutant plants

Infra-expressed genes	MT01	MT02	MT03	WT01	WT02	WT03	Description
Solyc01g006560	0.76	1.69	1.08	12.04	15.38	9.52	Lipoxygenase
Solyc01g008420	0.12	0.08	0.12	4.00	2.59	3.42	Mate efflux family protein
Solyc01g014320	0.80	0.45	1.30	5.88	4.68	4.71	SAM dependent carboxyl methyltransferase
Solyc01g068460	4.47	6.09	1.96	18.48	27.31	25.89	Pathogen-induced calmodulin-binding protein
Solyc01g079950	3.50	3.18	2.54	10.38	16.17	12.33	Xylanase inhibitor
Solyc01g080220	0.35	0.67	0.63	4.67	4.45	5.55	Dienelactone hydrolase
Solyc01g080790	0.00	0.13	0.00	5.07	3.79	2.95	Plant Basic Secretory Protein
Solyc01g090300	0.78	1.40	0.98	6.09	9.22	4.48	Ethylene responsive transcription factor 1b
Solyc01g104780	3.76	7.15	4.67	19.86	22.58	20.83	Nodulin-related integral membrane protein DUF125
Solyc02g005110	17.64	20.39	23.73	70.70	80.96	95.17	Unknown Protein
Solyc02g071380	0.36	0.26	0.35	4.93	4.38	6.39	1-aminocyclopropane-1-carboxylate oxidase 3
Solyc02g075620	0.00	0.40	0.36	2.54	2.50	3.02	Pectinesterase
Solyc02g079490	1.74	2.78	2.62	12.25	7.47	10.79	Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase
Solyc02g082450	0.31	0.55	0.77	3.29	2.74	6.28	Auxin efflux carrier family protein
Solyc03g005360	1.95	2.32	2.37	7.15	6.26	11.82	Unknown Protein
Solyc03g097870	3.91	8.39	3.84	49.84	23.40	38.85	MtN3-like protein_homologous SWEET11_sugar transporter
Solyc03g098300	2.96	4.13	3.57	22.99	42.11	20.29	Ornithine decarboxylase
Solyc04g007400	9.12	14.90	11.24	43.30	43.20	38.29	Tropinone reductase II
Solyc04g007790	33.60	14.70	13.46	307.90	270.20	305.10	Major latex-like protein
Solyc04g052980	1.00	2.84	1.57	9.38	12.04	8.90	Auxin-responsive protein
Solyc04g057940	0.95	1.69	0.59	5.06	4.79	5.27	U-box domain-containing protein
Solyc04g057980	7.62	19.39	4.38	50.37	42.33	51.36	NAD(P)H-quinone oxidoreductase subunit M
Solyc04g078030	0.37	0.89	0.45	2.93	5.11	3.98	Unknown Protein
Solyc05g015780	5.01	6.67	3.90	24.65	26.28	19.08	Unknown Protein
Solyc05g055290	0.00	0.09	0.00	4.54	3.53	3.02	Hydrolase alpha/beta fold family protein
Solyc06g009190	7.55	15.58	4.74	35.09	35.24	39.09	Pectinesterase
Solyc06g010240	0.00	0.80	0.82	3.09	6.39	5.83	Unknown Protein
Solyc06g049020	6.32	6.82	7.53	19.58	29.76	20.42	protein targeted either to mitochondria or chloroplast proteins
Solyc06g051560	0.96	3.28	2.24	7.83	15.43	9.52	Flavoprotein wrbA

Solyc06g052010	0.79	1.51	0.23	5.36	4.18	6.45	IQ calmodulin-binding region
Solyc06g069070	10.54	17.83	4.55	134.2	128.8	108.5	Lipid transfer protein
Solyc06g071830	1.60	4.01	2.83	13.58	11.49	13.32	BTB/POZ domain-containing protein
Solyc06g082300	0.80	1.76	1.77	7.46	5.46	8.39	UDP-glucosyltransferase
Solyc07g005370	13.57	21.25	21.2	143.9	86.15	121.4	Norcoclaurine synthase
Solyc07g006680	4.75	4.71	3.10	20.77	14.41	22.43	Hydroxycinnamoyl CoA quinate transferase
Solyc07g007260	16.42	66.06	18.95	138.60	134.30	136.40	Metallocarboxypeptidase inhibitor
Solyc07g021530	0.31	0.58	0.90	5.15	7.16	12.41	Unknown Protein
Solyc07g040960	2.11	1.98	1.95	7.45	12.36	19.34	Unknown protein
Solyc07g042630	3.06	6.42	2.31	50.07	38.24	43.04	Beta-Amyrin Synthase
Solyc07g048060	0.32	2.46	0.10	11.17	20.10	10.7	Auxin-induced protein-like
Solyc07g053550	4.85	4.38	1.69	12.89	46.41	12.2	Glutaredoxin
Solyc07g063770	0.56	0.81	0.62	3.83	4.91	2.79	Serine/threonine kinase receptor
Solyc08g007040	52.33	50.19	103.5	289.3	262.90	275.6	Glycine cleavage system H protein 1
Solyc08g007680	1.84	2.60	1.38	6.40	6.58	7.26	Subtilisin-like protease
Solyc08g007690	0.38	0.86	0.46	4.23	3.68	3.42	Subtilisin-like protease
Solyc08g008500	0.29	3.39	0.49	8.83	15.00	10.40	Dof zinc finger protein
Solyc08g028690	5.16	8.60	6.92	18.80	19.86	25.94	Tasselseed2-like short-chain dehydrogenase/reductase
Solyc08g078330	0.87	2.90	2.27	14.22	16.67	17.85	Oxidoreductase 2OG-Fe(II) oxygenase family
Solyc09g010570	2.60	3.49	1.04	11.34	9.92	10.85	Unknown Protein
Solyc09g065620	0.14	0.09	0.41	2.54	3.11	3.78	Chlorophyllase 1
Solyc09g075710	0.23	0.91	0.16	3.43	4.15	3.07	Gibberellin receptor GID1L2
Solyc09g082700	1.80	4.82	4.76	68.64	45.79	50.51	Early light-induced protein
Solyc09g092490	3.63	5.42	5.78	20.12	14.68	21.25	UDP-glucosyltransferase family 1 protein
Solyc10g006860	3.21	4.56	6.14	21.98	20.38	21.82	Glucose/ribitol dehydrogenase
Solyc10g006900	6.42	22.55	15.51	61.77	58.04	65.98	Protochlorophyllide reductase
Solyc10g007110	10.85	14.07	11.08	31.68	30.36	32.26	Tyrosine aminotransferase
Solyc10g007750	0.20	0.38	0.13	3.14	2.72	2.46	Plant Basic Secretory Protein
Solyc10g008160	0.75	1.24	0.45	3.71	4.46	5.6	Transcription factor
Solyc10g018140	0.44	0.56	0.07	9.09	7.89	12.05	Dihydroflavonol 4-reductase
Solyc10g051020	0.39	0.65	0.35	8.41	5.94	4.88	Cytochrome P450
Solyc10g077070	10.69	10.71	10.73	51.78	45.76	49.67	HAIRPLUS
Solyc10g078260	0.90	4.56	0.42	14.47	10.53	12.45	Peptidyl-prolyl cis-trans isomerase
Solyc10g080340	0.69	1.98	0.60	6.45	6.38	7.5	Multi antimicrobial extrusion protein MatE
Solyc10g085870	4.11	5.83	3.03	25.02	14.69	15.51	UDP-glucosyltransferase family 1 protein
Solyc10g085880	0.11	0.00	0.16	3.43	1.69	1.43	UDP-glucosyltransferase family 1 protein
Solyc10g086720	0.24	1.33	1.30	4.21	3.91	4.93	Fructose-1 6-bisphosphatase class 1
Solyc11g005340	4.91	8.03	5.40	28.24	22.62	24.56	PAP fibrillin family protein

Solyc11g007730	4.80	10.92	4.82	33.62	23.62	28.98	MTA/SAH nucleosidase / phosphatase
Solyc11g011740	0.21	0.20	0.31	3.84	2.08	4.35	Ethylene-responsive transcription factor 2
Solyc11g064840	1.14	1.40	0.51	5.97	6.59	5.10	C2 calcium-dependent membrane targeting
Solyc11g068380	0.78	1.90	0.86	6.74	9.83	7.17	Unknown Protein
Solyc11g071290	0.20	1.59	0.19	7.62	13.3	5.79	Alcohol dehydrogenase
Solyc12g009800	38.68	47.35	34.75	148.20	149.10	173.10	Purple acid phosphatase 3
Solyc12g010950	3.37	6.00	5.50	56.87	51.57	47.9	Alcohol dehydrogenase zinc-containing
Solyc12g011040	1.56	1.66	0.71	9.47	12.000	10.37	Lipoxygenase
Solyc12g042100	3.35	5.39	1.00	15.9	17.71	16.34	Unknown Protein
Solyc12g049470	0.55	1.04	1.59	7.42	7.63	4.4	Actin-6
Solyc12g056620	2.95	4.38	4.59	16.07	11.15	15.12	One-helix protein
Solyc12g088700	2.48	3.87	4.08	14.27	10.83	13.9	UDP-glucosyltransferase family 1 protein
Solyc12g089380	0.29	1.09	0.84	5.79	10.89	9.45	Expansin 45, endoglucanase-like
Solyc12g094460	7.59	4.34	5.52	18.30	19.15	18.25	Laccase-2
Solyc12g098600	3.76	7.11	3.34	20.07	16.21	18.9	UDP-glucosyltransferase family 1 protein
Solyc12g098900	1.42	10.91	1.10	25.25	45.78	41.32	Late embryogenesis abundant protein D-29
Solyc12g099430	3.15	2.40	5.00	22.10	21.81	24.69	Unknown

Conclusiones

En este trabajo se ha abordado la caracterización de una colección de mutantes de tomate obtenida mediante mutagénesis química con etil metano sulfonato (EMS) en el cultivar Moneymaker. Dicha caracterización ha permitido identificar nuevos reguladores del desarrollo vegetativo y reproductivo de tomate. Además, se ha llevado a cabo el aislamiento y caracterización funcional de los genes responsables de las mutaciones *succulent stamens2* (*sus2*) y *hairplus* (*hap*). El conjunto de resultados obtenidos nos ha permitido obtener las siguientes conclusiones:

PRIMERA. - Se ha generado una colección de mutantes en el cultivar Moneymaker de tomate mediante mutagénesis química con EMS. Un total de 8988 líneas mutagenizadas han sido caracterizadas y los mutantes detectados se han agrupado atendiendo a 6 categorías fenotípicas principales. Entre ellas destacan los mutantes afectados en el desarrollo vegetativo (36,57 %), mutantes de desarrollo reproductivo (17,83 %), mutantes partenocárpicos o que producen un menor número de semillas (33,76 %) y mutantes de calidad de fruto (11,84 %).

SEGUNDA. - Las variantes fenotípicas más frecuentes en dicha colección son las mutaciones que producen frutos partenocárpicos (33,76%) y aquellas que afectan al tamaño de la planta (25,30%), datos que indican que ambos fenotipos estarían controlados por un número de genes superior al resto de categorías fenotípicas, lo incrementaría la probabilidad de mutagénesis de alguno de ellos.

TERCERA. - Para la caracterización molecular de los mutantes identificados en esta colección se ha utilizado una estrategia de mapeo por secuenciación (*Whole Genome Sequencing)*, basada en el análisis de las frecuencias alélicas observadas en individuos de poblaciones F_2 derivadas del cruzamiento de plantas mutantes y una planta de la especie silvestre *S. pimpinellifolium* (acc. LA 1589). Esta estrategia ha permitido la identificación de los genes responsables de todos los mutantes seleccionados para este análisis.

CUARTA. - Entre los mutantes identificados destaca *succulent stamens2* (*sus2*), que se caracteriza por la formación de pétalos de pequeño tamaño y estambres de morfología aberrante y exhibe un patrón de herencia monogénica recesiva. El análisis fenotípico de todos los verticilos de la flor mediante microscopía electrónica de barrido (SEM) demostró cambios en la identidad de las células epidérmicas de los pétalos y que los estambres de plantas *sus2* presentan una conversión homeótica completa en carpelos. La caracterización molecular del mutante demostró que una mutación en el gen homeótico de clase B *TOMATO MADS-BOX 6* (*TM6*) es la responsable de este fenotipo.

QUINTA. - La pérdida de función de TM6 en plantas mutantes *sus2* da lugar a cambios en la expresión de otros genes homeóticos durante el desarrollo de las flores, ya que la expresión de los genes de clase A *MACROCALYX* (*MC*) y de clase C *TOMATO AGAMOUS LIKE1* (*TAGL1*) está inducida en casi la totalidad de los estadios analizados. Ello demuestra que *TM6* en particular y otros genes de clase B en general son esenciales para el mantenimiento del balance de la expresión de los diferentes genes homeóticos relacionados con la morfogénesis floral.

SEXTA. - Como parte de la caracterización de la colección se ha identificado un mutante de elevada densidad de tricomas multicelulares de tipo I al que hemos denominado *hairplus* (*hap*), cuya mutación sigue un patrón de herencia monogénica recesiva. La caracterización genética y molecular del mutante ha permitido demostrar que el gen *HAP* (*Solyc10g077070*) codifica para una posible Histona N-Lisina Metil Transferasa, proteínas de las que no se había descrito implicación en el control de la densidad de tricomas.

SÉPTIMA. - El análisis transcriptómico del mutante *hap* ha hecho posible la identificación de un conjunto de genes diferencialmente expresados, de los cuales la mayoría están inducidos y unos pocos reprimidos, entre ellos el propio *HAP*, en las plantas mutantes. Por otra parte, la secuenciación del epigenoma permitió identificar un elevado número

de Citosinas diferencialmente metiladas (CMCs) en el mutante *hap*. La mayoría de esas CMCs están localizadas en regiones intergénicas (64 %), si bien un 25 % se localiza en regiones promotoras y 11 % en regiones transcritas. Este conjunto de resultados demuestra la participación de *HAP* en la regulación epigenética de la expresión génica.

Bibliografía

Adibi, M., Yoshida, S., Weijers, D. y Fleck, C. (2016). Centering the organizing center in the Arabidopsis thaliana shoot apical meristem by a combination of cytokinin signaling and self-organization. PLoS ONE, 11(2).

Aida, M., Ishida, T. y Tasaka, M. (1999). Shoot apical meristem and cotyledon formation during *Arabidopsis* embryogenesis: interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. *Development*, *126*(8), 1563-1570.

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A. y Mason, C. E. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome biology*, *13*(10), R87.

Alba, J. M., Montserrat, M. y 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.

Ampomah-Dwamena, C., Morris, B. A., Sutherland, P., Veit, B. y Yao, J. L. (2002). Down-regulation of *TM29*, a Tomato *SEPALLATA* homolog, causes parthenocarpic fruit development and floral reversion. *Plant physiology*, *130*(2), 605-617.

Andrews S. (2010). FASTQC: A Quality Control Tool for High Throughput Sequence Data.

Arab, L. y Steck, S. (2000). Lycopene and cardiovascular disease. *The American journal of clinical nutrition*, *71*(6), 1691S-1695S.

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. y Shirakawa, M. (2008). Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature*, *455*(7214), 818-821.

Atarés, A., Moyano, E., Morales, B., Schleicher, P., García-Abellán, J. O., Antón, T. y Moreno, V. (2011). An insertional mutagenesis programme with an enhancer trap for the identification and tagging of genes involved in abiotic stress tolerance in the tomato wild-related species *Solanum pennellii*. *Plant cell reports*, *30*(10), 1865.

Backman, T. W. H. y Girke, T. (2016). systemPipeR: NGS workflow and report generationenvironment. *BMC Bioinformatics*, *17*(1), 388.

Bibliografía

Bai, Y., Kissoudis, C., Yan, Z., Visser, R. G., y van der Linden, G. (2018). Plant behaviour under combined stress: tomato responses to combined salinity and pathogen stress. *The Plant Journal*, *93*(4), 781-793.

Bai, Y. y Lindhout, P. (2007). Domestication and breeding of tomatoes: What have we gained and what can we gain in the future? *Annals of Botany*, *100*(5), 1085–1094.

Barnett, W. E. y DeBusk, A. G. (1960). Nitrous acid induced reverse mutation in *Neurospora crassa. Genetics*, *45*, 973-74.

Barry, C. S., Aldridge, G. M., Herzog, G., Ma, Q., McQuinn, R. P., Hirschberg, J. y Giovannoni, J. J. (2012). Altered chloroplast development and delayed fruit ripening caused by mutations in a zinc metalloprotease at the *lutescent2* locus of tomato. *Plant physiology*, *159*(3), 1086-1098.

Bauchet, G., y Causse, M. (2012). Genetic diversity in tomato (*Solanum lycopersicum*) and its wild relatives. *Genetic diversity in plants*, *133*, 162.

Baumbusch, L. O., Thorstensen, T., Krauss, V., Fischer, A., Naumann, K., Assalkhou, R., Schulz, I., Reuter, G. y Aalen, R. B. (2001). The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic acids research*, *29*(21), 4319-4333.

Behnke HD. (1984). Plant trichomes-structure and ultrastructure: general terminology, taxonomic applications, and aspects of Trichome bacterial interaction in leaf tips of Dioscorea. En: Rodriguez E, Healey PL, Mehta I, eds. Biology and chemistry of plant trichomes. New York: Plenum Press, 1–21.

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., Francis, D., Causse, M., van der Knaap, E. y Cañizares, J. (2015). Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics*, *16*(1), 257.

Bleeker, P. M., Diergaarde, P. J., Ament, K., Guerra, J., Weidner, M., Schütz, S., de Both, M. T. J., Haring, M. A. y Schuurink, R. C. (2009). The role of specific tomato volatiles in tomato-whitefly interaction. *Plant Physiology*, *151*(2), 925–935.

Bibliografía

Bouzroud, S., Gouiaa, S., Hu, N., Bernadac, A., Mila, I., Bendaou, N., Smouni, A., Bouzayen, M. y Zouine, M. (2018). Auxin Response Factors (ARFs) are potential mediators of auxin action in tomato response to biotic and abiotic stress (*Solanum lycopersicum*). *PLoS One*, *13*(2).

Bowman, J. L., Alvarez, J., Weigel, D., Meyerowitz, E. M. y Smyth, D. R. (1993). Control of flower development in *Arabidopsis thaliana* by *APETALA*1 and interacting genes. *Development*, *119*(3), 721-743.

Brunelle, D. C., Clark, J. K. y Sheridan, W. F. (2017). Genetic Screening for EMS-Induced Maize Embryo-Specific Mutants Altered in Embryo Morphogenesis. *G3: Genes, Genomes, Genetics*, *7*(11), 3559-3570.

Brutnell, T. P. (2002). Transposon tagging in maize. *Functional and Integrative Genomics*, *2*(1-2), 4–12.

Busi, M. V., Bustamante, C., D'angelo, C., Hidalgo-Cuevas, M., Boggio, S. B., Valle, E. M. y Zabaleta, E. (2003). MADS-box genes expressed during tomato seed and fruit development. *Plant molecular biology*, *52*(4), 801-815.

Campos, J. F., Cara, B., Pérez-Martín, F., Pineda, B., Egea, I., Flores, F. B., Fernandez-García, N., Capel, J., Moreno, V., Angosto, T., Lozano, R. y Bolarín, M. C. (2016). The tomato mutant *ars1* (*altered response to salt stress 1*) identifies an R1-type MYB transcription factor involved in stomatal closure under salt acclimation. *Plant biotechnology journal*, *14*(6), 1345-1356.

Candela, H., y Hake, S. (2008). The art and design of genetic screens: maize. *Nature Reviews Genetics*, *9*(3), 192.

Cao, X., Liu, X., Wang, X., Yang, M., van Giang, T., Wang, J., Liu, X., Sun, S., Wei, K., Wang, X., Gao, J., Du, Y., Qin, Y., Guo, Y. y Huang, Z. (2019). B-class MADS-box TM6 is a candidate gene for tomato male sterile-15 26. *Theoretical and Applied Genetics*, *132*(7), 2125-2135.

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. y 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.

Carter, C. D. y Snyder, J. C. (1985). Mite responses in relation to trichomes of *Lycopersicon esculentum* x *L. hirsutum* F_2 hybrids. *Euphytica 34*(1), 177-185.

Carvalho, R. F., Campos, M. L., Pino, L. E., Crestana, S. L., Zsögön, A., Lima, J. E., Benedito, V. A. y Peres, L. E. (2011). Convergence of developmental mutants into a single tomato model system: 'Micro-Tom' as an effective toolkit for plant development research. *Plant Methods*, *7*(1), 18.

Castelán-Muñoz, N., Herrera, J., Cajero-Sánchez, W., Arrizubieta, M., Trejo, C., García-Ponce, B., Sánchez, M. de la P., Álvarez-Buylla, E. R. y Garay-Arroyo, A. (2019). MADS-box genes are key components of genetic regulatory networks involved in abiotic stress and plastic developmental responses in plants. *Frontiers in Plant Science*, *10*, 853.

Causse, M. y Grandillo, S. (2016). Gene mapping in tomato. En: The tomato genome, 23–37. Springer, Berlin, Heidelberg.

Chamarro, J. (1995). Anatomía y fisiología de la planta. En: El cultivo del tomate, pp. 43-91. Mundi-Prensa.

Chang, J., Yu, T., Yang, Q., Li, C., Xiong, C., Gao, S., Xie, Q., Zheng, F., Li, H., Tian, Z., Yang, C. y Yang, C. (2018). *Hair*, encoding a single C2H2 zinc-finger protein, regulates multicellular trichome formation in tomato. *The Plant Journal*, *96*(1), 90-102.

Chang, J., Yu, T., Gao, S., Xiong, C., Xie, Q., Li, H., Ye, Z. y Yang, C. (2016). Fine mapping of the *dialytic* gene that controls multicellular trichome formation and stamen development in tomato. *Theoretical and applied genetics*, *129*(8), 1531-1539.

Channarayappa, C., Shivashankar, G., Muniyappa, V. y Frist, R. H. (1992). Resistance of *Lycopersicon* species to *Bemisia tabaci*, a tomato leaf curl virus vector. *Canadian Journal of Botany*, *70*(11), 2184-2192.

Chen, S., Jin, W., Wang, M., Zhang, F., Zhou, J., Jia, Q., Wu, Y., Liu, F. y Wu, P. (2003). Distribution and characterization of over 1000 T-DNA tags in rice genome. *The Plant Journal*, *36*(1), 105–113.

Bibliografía

Chen, X., Goodwin, S. M., Boroff, V. L., Liu, X. y Jenks, M. A. (2003). Cloning and characterization of the *WAX2* gene of Arabidopsis involved in cuticle membrane and wax production. *The Plant Cell*, *15*(5), 1170-1185.

Chung, M. Y., Vrebalov, J., Alba, R., Lee, J., McQuinn, R., Chung, J. D., Klein, P. y Giovannoni, J. (2010). A tomato (Solanum lycopersicum) APETALA2/ERF gene, SlAP2a, is a negative regulator of fruit ripening. *The Plant Journal*, *64*(6), 936-947.

Clark, S. E., Williams, R. W. y Meyerowitz, E. M. (1997). The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in arabidopsis. *Cell*, *89*(4), 575–585.

Clark, S. E., Running, M. P. y Meyerowitz, E. M. (1995). *CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the same processes as *CLAVATA1*. *Development*, *121*(7), 2057-2067.

Coen, E. S. y Meyerowitz, E. M. (1991). The war of the whorls: Genetic interactions controlling flower development. *Nature*, *353*(6339), 31–37.

Colombo, M., Brambilla, V., Marcheselli, R., Caporali, E., Kater, M. M. y Colombo, L. (2010). A new role for the *SHATTERPROOF* genes during *Arabidopsis* gynoecium development. *Developmental Biology*, *337*(2), 294–302.

Cubero J.I. (2003). Introdución a la Mejora Genética de Plantas. Ediciones Mundi-Prensa.

Daminato, M., Masiero, S., Resentini, F., Lovisetto, A., y Casadoro, G. (2014). Characterization of *TM8*, a MADS-box gene expressed in tomato flowers. *BMC plant biology*, *14*(1), 319.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R. y 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158.

De Azevedo, S. M., Ventura Faria, M., Maluf, W. R., Barneche De Oliveira, A. C. y De Freitas, J. A. (2003). Zingiberene-mediated resistance to the South American tomato pinworm derived from *Lycopersicon hirsutum* var. *hirsutum*. *Euphytica*, *134*(3), 347–351.

De Candolle, A. (1882). L'Origine des Plantes Cultivées. Appleton, New York.

de Martino, G., Pan, I., Emmanuel, E., Levy, A. y Irish, V. F. (2006). Functional analyses of two tomato *APETALA3* genes demonstrate diversification in their roles in regulating floral development. *The Plant Cell*, *18*(8), 1833-1845.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis1, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D. y Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics*, *43*(5), 491.

Dickinson, H. (1995). Dry stigmas, water and self-incompatibility in Brassica. *Sexual Plant Reproduction*, *8*(1), 1-10.

Díez, M. J. y Nuez, F. (2008). Tomato. En: Vegetables II, 249-323. Springer, New York, NY.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. y Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15-21.

Drews, G. N., Bowman, J. L. y Meyerowitz, E. M. (1991). Negative regulation of the Arabidopsis homeotic gene *AGAMOUS* by the *APETALA2* product. *Cell*, *65*(6), 991-1002.

Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T. y Reuter, G. (2004). Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. *Genes & development*, *18*(23), 2973-2983.

Eigenbrode, S. D., Trumble, J. T., Millar, J. G. y White, K. K. (1994). Topical toxicity of tomato sesquiterpenes to the beet armyworm and the role of these compounds in resistance derived from an accession of *Lycopersicon hirsutum* f. *typicum*. *Journal of Agricultural and Food Chemistry*, *42*(3), 807–810.

Emmanuel, E. y Levy, A. A. (2002). Tomato mutants as tools for functional genomics. *Current opinion in plant biology*, *5*(2), 112-117.

Endrizzi, K., Moussian, B., Haecker, A., Levin, J. Z. y Laux, T. (1996). The *SHOOT MERISTEMLESS* gene is required for maintenance of undifferentiated cells in

Arabidopsis shoot and floral meristems and acts at a different regulatory level than the meristem genes *WUSCHEL* and *ZWILLE*. *Plant Journal*, *10*(6), 967–979.

Esch, J. J., Chen, M., Sanders, M., Hillestad, M., Ndkium, S., Idelkope, B., Neizer, J. y Marks, M. D. (2003). A contradictory *GLABRA3* allele helps define gene interactions controlling trichome development in *Arabidopsis*. *Development*, *130*(24), 5885-5894.

Evans, M. M. y Barton, M. K. (1997). Genetics of angiosperm shoot apical meristem development. *Annual Review of Plant Biology*, *48*(1), 673-701.

Feldman, A. B., Murchie, E. H., Leung, H., Baraoidan, M., Coe, R., Yu, S. M., Lo, S. F. y Quick, W. P. (2014). Increasing leaf vein density by mutagenesis: laying the foundations for C4 rice. *PLoS One*, *9*(4).

Feldmann, K. A. (1991). T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. *The Plant Journal*, *1*(1), 71–82.

Fernández-Muñoz, R., Salinas, M., Álvarez, M., y Cuartero, J. (2003). Inheritance of resistance to two-spotted spider mite and glandular leaf trichomes in wild tomato *Lycopersicon pimpinellifolium* (Jusl.) Mill. *Journal of the American Society for Horticultural Science*, *128*(2), 188-195.

Ferrario, S., Immink, R. G. H. y Angenent, G. C. (2004). Conservation and diversity in flower land. *Current Opinion in Plant Biology, 7*(1), 84–91.

Fletcher, J. C. (2002). Shoot and floral meristem maintenance in *Arabidopsis*. *Annual review of plant biology*, *53*(1), 45-66.

Fletcher, J. C. y Meyerowitz, E. M. (2000). Cell signaling within the shoot meristem. *Current opinion in plant biology*, *3*(1), 23-30.

Fobes, J. F., Mudd, J. B. y Marsden, M. P. (1985). Epicuticular lipid accumulation on the leaves of *Lycopersicon pennellii* (Corr.) D'Arcy and *Lycopersicon esculentum* Mill. *Plant Physiology*, *77*(3), 567-570.

Folkers, U., Berger, J. y Hülskamp, M. (1997). Cell morphogenesis of trichomes in Arabidopsis: Differential control of primary and secondary branching by branch initiation regulators and cell growth. *Development*, *124*(19), 3779–3786.

Foolad, M. R. (2007). Genome mapping and molecular breeding of tomato. *International Journal of Plant Genomics*, *2007*.

Garcia, V., Bres, C., Just, D., Fernandez, L., Tai, F. W. J., Mauxion, J. P., Le Paslier, M. C., Bérard, A., Brunel, D., Aoki, K., Alseekh, S., Fernie, A. R., Fraser, P. D. y Rothan, C. (2016). Rapid identification of causal mutations in tomato EMS populations via mapping-by-sequencing. *Nature protocols*, *11*(12), 2401.

Gelvin, S. B. (2003). *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiology and Molecular Biology Reviews, 67*(1), 16-37.

Geuten, K. y Irish, V. (2010). Hidden variability of floral homeotic B genes in Solanaceae provides a molecular basis for the evolution of novel functions. *The Plant Cell*, *22*(8), 2562-2578.

Giménez, E., Castañeda, L., Pineda, B., Pan, I. L., Moreno, V., Angosto, T. y Lozano, R. (2016). TOMATO AGAMOUS1 and *ARLEQUIN/TOMATO AGAMOUS-LIKE1* MADS-box genes have redundant and divergent functions required for tomato reproductive development. *Plant molecular biology*, *91*(4-5), 513-531.

Giménez, E., Pineda, B., Capel, J., Antón, M. T., Atarés, A., Pérez-Martín, F., García-Sogo, B., Angosto, T. y Lozano, R. (2010). Functional analysis of the *Arlequin* mutant corroborates the essential role of the *Arlequin/TAGL1* gene during reproductive development of tomato. *PloS one*, 5(12).

Glas, J. J., Schimmel, B. C., Alba, J. M., Escobar-Bravo, R., Schuurink, R. C. y 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.

Gleave, A. P. (1992). A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Molecular Biology*, *20* (6), 1203-1207.

Glover, B. J., Bunnewell, S. y Martin, C. (2004). Convergent evolution within the genus *Solanum*: The specialised anther cone develops through alternative pathways. *Gene*, *331*(1–2), 1–7.

Bibliografía

Gómez-Martín, C. RNAseqScripts in python (2017). https://github.com/cris12gm/rnaseqScripts

Gómez, P., Jamilena, M., Capel, J., Zurita, S., Angosto, T. y Lozano, R. (1999). *STAMENLESS*, a tomato mutant with homeotic conversions in petals and stamens. *Planta*, 209(2), 172-179.

Gorguet, B., Van Heusden, A. W. y Lindhout, P. (2005). Parthenocarpic fruit development in tomato. *Plant Biology*, *7*(02), 131-139.

Greene, E. A., Codomo, C. A., Taylor, N. E., Henikoff, J. G., Till, B. J., Reynolds, S. H., Enns, L. C., Burtner, C., Johnson, J. E., Odden, A. R., Comai, L. y Henikoff, S. (2003). Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. *Genetics*, *164*(2), 731–740.

Gulfishan, M., Bhat, T. A. y Oves, M. (2015). Mutants as a genetic resource for future crop improvement. En: Advances in plant breeding strategies: breeding, biotechnology and molecular tools, 95-112. Springer, Cham.

Ha, C. M., Jun, J. H. y Fletcher, J. C. (2010). Shoot apical meristem form and function. *Current topics in developmental biology*, *91*, 103-140.

Hanano, S. y Goto, K. (2011). *Arabidopsis TERMINAL FLOWER1* is involved in the regulation of flowering time and inflorescence development through transcriptional repression. *The Plant Cell*, *23*(9), 3172-3184.

Harris, C. J., Scheibe, M., Wongpalee, S. P., Liu, W., Cornett, E. M., Vaughan, R. M., Li, X., Chen, W., Xue, Y., Zhong, Z., Yen, L., Barshop, W. D., Rayatpisheh, S., Gallego-Bartolome, J., Groth, M., Wang, Z., Wohlschlegel, J. A., Du, J., Rothbart, S. B., Butter, F. y Jacobsen, S. E. (2018). A DNA methylation reader complex that enhances gene transcription. *Science*, *362*(6419), 1182-1186.

Heuermann, M. C., Rosso, M. G., Mascher, M., Brandt, R., Tschiersch, H., Altschmied, L. y Altmann, T. (2019). Combining next-generation sequencing and progeny testing for rapid identification of induced recessive and dominant mutations in maize M2 individuals. *The Plant Journal*, *100*(4), 851-862.

Hirochika, H., Guiderdoni, E., An, G., Hsing, Y. I., Eun, M. Y., Han, C. D., Upadhyaya, N., Ramachandran, S., Zhang, Q., Pereira, A., Sundaresan, V. y Leung, H. (2004). Rice mutant resources for gene discovery. *Plant molecular biology*, *54*(3), 325-334.

Hobson, G. y Grierson, D. (1993). Tomato. En: Biochemistry of fruit ripening, 405-442. Springer, Dordrecht.

Hrdlickova, R., Toloue, M. y Tian, B. (2017). RNA-Seq methods for transcriptome analysis. *Wiley Interdisciplinary Reviews RNA, 8*(1), e1364.

Hu, Y., Xie, Q. y Chua, N. H. (2003). The *Arabidopsis* auxin-inducible gene *ARGOS* controls lateral organ size. *Plant Cell*, *15*(9), 1951-1961.

Huang, Z., Shi, T., Zheng, B., Yumul, R. E., Liu, X., You, C., Gao, Z., Xiao, L. y Chen, X. (2017). *APETALA2* antagonizes the transcriptional activity of *AGAMOUS* in regulating floral stem cells in *Arabidopsis thaliana*. *New Phytologist*, *215*(3), 1197–1209.

Huang, P. C. y Paddock, E. F. (1962). The time and site of the semidominant lethal action of "*wo*" in *Lycopersicon esculentum*. *American Journal of Botany*, *49*(4), 388–393.

Huchelmann, A., Boutry, M. y Hachez, C. (2017). Plant glandular trichomes: natural cell factories of high biotechnological interest. *Plant physiology*, *175*(1), 6-22.

Hülskamp, M. (2004). Plant trichomes: a model for cell differentiation. *Nature Reviews Molecular Cell Biology*, *5*(6), 471-480.

Hülskamp, M., Schnittger, A. y Folkers, U. (1999). Pattern formation and cell differentiation: trichomes in *Arabidopsis* as a genetic model system. *International Review of Cytology*, *186,* 147–178.

Hur, Y. S., Um, J. H., Kim, S., Kim, K., Park, H. J., Lim, J. S., Ohme-Takagi, M., Kim, D., Park, J., Kim, G. T. y Cheon, C. I. (2015). *Arabidopsis thaliana* homeobox 12 (ATHB 12), a homeodomain-leucine zipper protein, regulates leaf growth by promoting cell expansion and endoreduplication. *New Phytologist*, *205*(1), 316-328.

Ingram, G. y Nawrath, C. (2017). The roles of the cuticle in plant development: organ adhesions and beyond. *Journal of experimental botany*, *68*(19), 5307-5321.

Irish, V. (2017). The ABC model of floral development. *Current Biology*, *27*(17), R887-R890.

Bibliografía

Irish, V. F. y Sussex, I. M. (1990). Function of the *apetala-1* gene during *Arabidopsis* floral development. *The Plant Cell*, *2*(8), 741-753.

Jackson, J. P., Lindroth, A. M., Cao, X. y Jacobsen, S. E. (2002). Control of CpNpG DNA methylation by the *KRYPTONITE* histone H3 methyltransferase. *Nature*, *416*(6880), 556-560.

Jain, S. M. (2010). Mutagenesis in crop improvement under the climate change. *Romanian biotechnological letters*, *15*(2), 88-106.

Jenkins, J. A. (1948). The origin of the cultivated tomato. *Economic Botany*, *2*(4), 379-392.

Jeon, J. S., Lee, S., Jung, K. H., Jun, S. H., Jeong, D. H., Lee, J., Kim, C., Jang, S., Lee, S., Yang, K., Nam, J., An, K., Han, M. J., Sung, R. J., Choi, H. S., Yu, J. H., Choi, J. H., Cho, S. Y., Cha, S. S., Kim, S. I. y An, G. (2000). T-DNA insertional mutagenesis for functional genomics in rice. *Plant Journal*, *22*(6), 561–570.

Jupe, F., Rivkin, A. C., Michael, T. P., Zander, M., Motley, S. T., Sandoval, J. P., Slotkin, R. K., Chen, H., Castañon, R., Nery, J., y Ecker, J. R. (2019). The complex architecture and epigenomic impact of plant T-DNA insertions. *PLoS genetics*, *15*(1).

Kalloo, G. (1991). Interspecific and intergeneric hybridization in tomato. En: Genetic improvement of tomato, 73-82. Springer, Berlin, Heidelberg.

Kang, J. H., Campos, M. L., Zemelis-Durfee, S., Al-Haddad, J. M., Jones, A. D., Telewski, F. W., Brandizzi, F. y Howe, G. A. (2016). Molecular cloning of the tomato *Hairless* gene implicates actin dynamics in trichome-mediated defense and mechanical properties of stem tissue. *Journal of experimental botany*, *67*(18), 5313-5324.

Kang, J. H., Shi, F., Jones, A. D., Marks, M. D. y Howe, G. A. (2010). Distortion of trichome morphology by the *hairless* mutation of tomato affects leaf surface chemistry. *Journal of Experimental Botany*, *61*(4), 1053–1064.

Kaudewitz, F. (1959). Production of bacterial mutants with nitrous acid. *Nature*, *183*(4678), 1829–1830.

Karlova, R., Rosin, F. M., Busscher-Lange, J., Parapunova, V., Do, P. T., Fernie, A. R., Fraser, P. D., Baxter, C., Angenet, G. C. y de Maagd, R. A. (2011). Transcriptome

and metabolite profiling show that *APETALA2a* is a major regulator of tomato fruit ripening. *The Plant Cell*, *23*(3), 923-941.

Khan, M. R. G., Ai, X. Y. y Zhang, J. Z. (2014). Genetic regulation of flowering time in annual and perennial plants. *Wiley Interdisciplinary Reviews: RNA*, *5*(3), 347–359.

Kikkert, J. R., Vidal, J. R. y Reisch, B. I. (2005). Stable transformation of plant cells by particle bombardment/biolistics. En: Transgenic plants: methods and protocols, 61-78. Humana Press.

Kim, Y. J. y Zhang, D. (2018). Molecular control of male fertility for crop hybrid breeding. *Trends in plant science*, *23*(1), 53-65.

Kirik, V., Simon, M., Hülskamp, M. y Schiefelbein, J. (2004). The *ENHANCER of TRY* and *CPC1* gene acts redundantly with *TRIPTYCHON* and *CAPRICE* in trichome and root hair cell patterning in Arabidopsis. *Developmental Biology*, *268*(2), 506–513.

Kissoudis, C., Chowdhury, R., van Heusden, S., van de Wiel, C., Finkers, R., Visser, R. G., Bai, Y. y van der Linden, G. (2015). Combined biotic and abiotic stress resistance in tomato. *Euphytica*, *202*(2), 317-332.

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M. y Araki, T. (1999). A pair of related genes with antagonistic roles in mediating flowering signals. *Science*, *286*(5446), 1960–1962.

Kodym, A. y Afza, R. (2003). Physical and chemical mutagenesis. *Methods in Molecular Biology*, *236*, 189-204.

Koornneef, M. (1981). The complex syndrome of *ttg* mutants. *Arabidopsis Information Service*, *18*, 45-51.

Kouzarides, T. (2002). Histone methylation in transcriptional control. *Current opinion in genetics & development*, *12*(2), 198-209.

Kramer, E. M., Jaramillo, M. A. y Di Stilio, V. S. (2004). Patterns of gene duplication and functional evolution during the diversification of the *AGAMOUS* subfamily of MADS box genes in angiosperms. *Genetics*, *166*(2), 1011-1023.

Kramer, E. M., Dorit, R. L. y Irish, V. F. (1998). Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the *APETALA3* and *PISTILLATA* MADS-box gene lineages. *Genetics*, *149*(2), 765-783.

Bibliografía

Krieg, D. R. (1963). Ethyl methanesulfonate-induced reversion of bacteriophage T4rII mutants. *Genetics*, *48*(4), 561.

Krizek, B. A. y Meyerowitz, E. M. (1996). The *Arabidopsis* homeotic genes *APETALA3* and *PISTILLATA* are sufficient to provide the B class organ identity function. *Development*, *122*(1), 11-22.

Krueger, F. *Taking appropriate QC measures for RRBS-type or other -Seq applications with Trim Galore* (2017).

Kuromori, T., Hirayama, T., Kiyosue, Y., Takabe, H., Mizukado, S., Sakurai, T., Akiyama, K., Kamiya, A., Ito, T. y Shinozaki, K. (2004). A collection of 11.800 single-copy *Ds* transposon insertion lines in *Arabidopsis*. *The Plant Journal*, *37*(6), 897–905.

Lachner, M. y Jenuwein, T. (2002). The many faces of histone lysine methylation. *Current opinion in cell biology*, *14*(3), 286-298.

Langmead, B. y Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, *9*(4), 357.

Larkin, J. C., Brown, M. L. y Schiefelbein, J. (2003). How do cells know what they want to be when they grow up? Lessons from epidermal patterning in Arabidopsis. *Annual Review of Plant Biology*, *54*(1), 403-430.

Laskar, R. A., Chaudhary, C., Khan, S. y Chandra, A. (2018). Induction of mutagenized tomato populations for investigation on agronomic traits and mutant phenotyping. *Journal of the Saudi society of agricultural sciences*, *17*(1), 51-60.

Laux, T., Mayer, K. F., Berger, J. y Jurgens, G. (1996). The *WUSCHEL* gene is required for shoot and floral meristem integrity in Arabidopsis. *Development*, *122*(1), 87-96.

Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T. y Carey, V. J. (2013). Software for computing and annotating genomic ranges. *PLoS computational biology*, *9*(8).

Li, M., Zhang, D., Gao, Q., Luo, Y., Zhang, H., Ma, B., Chen, C., Whibbey, A., Zhang, Y., Cao, Y., Li, Q., Guo, H., Li, J., Song, Y., Zhang, Y., Copsey, L., Li, Y., Li, X., Qi, M., Wang, J., Chen, Y., Wang, D., Zhao, J., Liu, G., Wu, B., Yu, L., Xu, C., Li, J., Zhao, S., Zhang, Y., Hu, S., Liang, C., Yin, Y., Coen, E y Xue, Y. (2019). Genome structure and evolution of *Antirrhinum majus* L. *Nature plants*, *5*(2), 174-183.

Li, H. (2011). Improving SNP discovery by base alignment quality. *Bioinformatics*, *27*(8), 1157-1158.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. y 1000 Genome Project Data Processing Subgroup. (2009). The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics 25*(16), 2078-2079.

Li, X. y Zhang, Y. (2002). Reverse genetics by fast neutron mutagenesis in higher plants. Functional y integrative genomics, 2(6), 254-258.

Li, X., Song, Y., Century, K., Straight, S., Ronald, P., Dong, X., Lassner, M. y Zhang, Y. (2001). A fast neutron deletion mutagenesis-based reverse genetics system for plants. *The Plant Journal*, *27*(3), 235-242.

Lioznova, A. V., Khamis, A. M., Artemov, A. V., Besedina, E., Ramensky, V., Bajic, V. B., Kulakovskiy, I. V. y Medvedeva, Y. A. (2019). CpG traffic lights are markers of regulatory regions in human genome. *BMC genomics*, *20*(1), 102.

Lippman, Z. B., Cohen, O., Alvarez, J. P., Abu-Abied, M., Pekker, I., Paran, I., Eshed, Y. y Zamir, D. (2008). The making of a compound inflorescence in tomato and related nightshades. *PLoS biology*, *6*(11).

Litt, A. y Kramer, E. M. (2010). The ABC model and the diversification of floral organ identity. *Seminars in cell & developmental biology* 21(1), 129-137.

Livak, K. J. y Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*, *25*(4), 402-408.

Lloyd, A. M., Walbot, V. y Davis, R. W. (1992). Arabidopsis and Nicotiana anthocyanin production activated by maize regulators R and C1. *Science*, *258*(5089), 1773–1775.

Long, J. A. y Barton, M. K. (1998). The development of apical embryonic pattern in *Arabidopsis*. *Development*, *125*(16), 3027-3035.

Lozano, R., Giménez, E., Cara, B., Capel, J., y Angosto, T. (2009). Genetic analysis of reproductive development in tomato. I*nternational Journal of Developmental Biology*, *53*(8-9-10), 1635-1648.

Lozano, R., Angosto, T., Gómez, P., Payán, C., Capel, J., Huijser, P., Salinas, J. y Martínez-Zapater, J. M. (1998). Tomato flower abnormalities induced by low

temperatures are associated with changes of expression of MADS-box genes. *Plant Physiology*, *117*(1), 91-100.

Lucatti, A. F., van Heusden, A. W., de Vos, R. C., Visser, R. G. y Vosman, B. (2013). Differences in insect resistance between tomato species endemic to the Galapagos Islands. *BMC evolutionary biology*, *13*(1), 175.

Luckwill, L. C. (1943). The genus Lycopersicon: an historical, biological, and taxonomic survery of the wild and cultivated tomatoes. Aberdeen University Press.

Mabuchi, T. y Arnason, T. J. (1969). Absence of chlorophyll mutations in Arabidopsis treated with LSD. *Arabidopsis Information Service*, *6*, 27.

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C. y Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic acids research*, *47*(W1), W636-W641.

Mao, L., Begum, D., Chuang, H. W., Budiman, M. A., Szymkowiak, E. J., Irish, E. E. y Wing, R. A. (2000). *JOINTLESS* is a MADS-box gene controlling tomato flower abscission zone development. *Nature*, *406*(6798), 910-913.

Marks, M. D. y Feldmann, K. A. (1989). Trichome development in *Arabidopsis thaliana*. I. T-DNA tagging of the *GLABROUS1* gene. *The plant cell*, *1*(11), 1043-1050.

Marmorstein, R. (2003). Structure of SET domain proteins: a new twist on histone methylation. *Trends in biochemical sciences*, *28*(2), 59-62.

Martín, B., Ramiro, M., Martínez-Zapater, J. M. y Alonso-Blanco, C. (2009). A high-density collection of EMS-induced mutations for TILLING in Landsberg *erecta* genetic background of *Arabidopsis*. *BMC Plant biology*, *9*(1), 147.

Martin, C. y Glover, B. J. (2007). Functional aspects of cell patterning in aerial epidermis. *Current opinion in plant biology*, *10*(1), 70-82.

Martin, C. y Paz-Ares, J. (1997). MYB transcription factors in plants. *Trends in Genetics*, 13(2), 67-73.

May, B. P. y Martienssen, R. A. (2003). Transposon mutagenesis in the study of plant development. *Critical Reviews in Plant Sciences, 22*(1), 1–35.

Mayer, K. F., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G. y Laux, T. (1998). Role of *WUSCHEL* in regulating stem cell fate in the *Arabidopsis* shoot meristem. *Cell*, *95*(6), 805-815.

Mazzucato, A., Taddei, A. R. y Soressi, G. P. (1998). The *parthenocarpic fruit* (*pat*) mutant of tomato (*Lycopersicon esculentum* Mill.) sets seedless fruits and has aberrant anther and ovule development. *Development*, *125*(1), 107-114

Mba, C. (2013). Induced Mutations Unleash the Potentials of Plant Genetic Resources for Food and Agriculture. *Agronomy*, *3*(1), 200–231.

Mba, C., Afza, R., y Shu, Q. Y. (2012). Mutagenic radiations: X-rays, ionizing particles and ultraviolet. *Plant mutation breeding and biotechnology*, 83-90.

McCallum, C. M., Comai, L., Greene, E. A. y Henikoff, S. (2000). Targeted screening for induced mutations. *Nature biotechnology*, *18*(4), 455-457.

Meissner, R., Chague, V., Zhu, Q., Emmanuel, E., Elkind, Y. y Levy, A. A. (2001). A high throughput system for transposon tagging and promoter trapping in tomato. *Plant Journal*, *22*(3), 265–274.

Meissner, R., Jacobson, Y., Melamed, S., Levyatuv, S., Shalev, G., Ashri, A., Elkind, Y. y Levy, A. (1997). A new model system for tomato genetics. *The Plant Journal*, *12*(6), 1465-1472.

Menda, N., Semel, Y., Peled, D., Eshed, Y. y Zamir, D. (2004). In silico screening of a saturated mutation library of tomato. *The Plant Journal*, *38*(5), 861-872.

Meyerowitz, E. M. (1997, January). Control of cell division patterns in developing shoots and flowers of *Arabidopsis thaliana*. En: *Cold Spring Harbor symposia on quantitative biology*, *62*, 369-375. Cold Spring Harbor Laboratory Press.

Miller, J. C. y Tanksley, S. D. (1990). RFLP analysis of phylogenetic relationships and genetic variation in the genus Lycopersicon. *Theoretical and applied genetics*, 80(4), 437-448.

Miller, P. (1731). The Gardeners Dictionary: containing the methods of cultivating and improving the kitchen, fruit and flower garden, as also the physick garden, wilderness, conservatory, and vineyard.

Bibliografía

Min, J., Zhang, X., Cheng, X., Grewal, S. I. y Xu, R. M. (2002). Structure of the SET domain histone lysine methyltransferase Clr4. *Nature structural biology*, *9*(11), 828-832.

Mizzotti, C., Mendes, M. A., Caporali, E., Schnittger, A., Kater, M. M., Battaglia, R. y Colombo, L. (2012). The MADS box genes *SEEDSTICK* and *ARABIDOPSIS* B_{sister} play a maternal role in fertilization and seed development. *Plant Journal*, *70*(3), 409–420.

Moco, S., Bino, R. J., Vorst, O., Verhoeven, H. A., de Groot, J., van Beek, T. A., Vervoot, J. y De Vos, C. R. (2006). A liquid chromatography-mass spectrometry-based metabolome database for tomato. *Plant physiology*, *141*(4), 1205-1218.

Mohapatra, T., Robin, S., Sarla, N., Sheshashayee, M., Singh, A. K., Singh, K., Singh N. K., Mithra, A. C. R. y Sharma, R. P. (2014). EMS induced mutants of upland rice variety Nagina22: generation and characterization. *Proceedings of the Indian National Science Academy*, *80*(1), 163-172.

Molinero-Rosales, N., Latorre, A., Jamilena, M. y Lozano, R. (2004). *SINGLE FLOWER TRUSS* regulates the transition and maintenance of flowering in tomato. *Planta*, *218*(3), 427-434.

Molinero-Rosales, N., Jamilena, M., Zurita, S., Gómez, P., Capel, J. y Lozano, R. (1999). *FALSIFLORA*, the tomato orthologue of *FLORICAULA* and *LEAFY*, controls flowering **time and floral meristem identity. *The Plant Journal*, *20*(6), 685-693.**

Morgan, M., Pagès, H., Obenchain, V. y Hayden, N. (2019). *Rsamtools: Binary Alignment (BAM), FASTA, Variant Call (BCF), and Tabix File Import*. (R package version 2.0.3, 2019). http://bioconductor.org/packages/Rsamtools.

Nadeau, J. A. y Sack, F. D. (2002). Control of stomatal distribution on the *Arabidopsis* leaf surface. *Science*, *296*(5573), 1697-1700.

Negi, J., Moriwaki, K., Konishi, M., Yokoyama, R., Nakano, T., Kusumi, K., Hashimoto-Sugimoto, M., Schroeder, J., Nishitani, K., Yanagisawa, S. y Iba, K. (2013). A Dof transcription factor, *SCAP1*, is essential for the development of functional stomata in Arabidopsis. *Current Biology*, *23*(6), 479-484.

Notaguchi, M., Abe, M., Kimura, T., Daimon, Y., Kobayashi, T., Yamaguchi, A., Tomita, Y., Dohi, K., Mori, M. y Araki, T. (2008). Long-Distance, Graft-Transmissible Action

of *Arabidopsis* FLOWERING LOCUS T Protein to Promote Flowering. *Plant and Cell Physiology*, *49*(11), 1645–1658.

Nuez, F. (1995). El cultivo del tomate. Mundi-Prensa. Madrid, España.

Okabe, Y., Asamizu, E., Saito, T., Matsukura, C., Ariizumi, T., Brès, C., Rothan, C., Mizoguchi, T. y Ezura, H. (2011). Tomato TILLING technology: development of a reverse genetics tool for the efficient isolation of mutants from Micro-Tom mutant libraries. *Plant and cell physiology*, *52*(11), 1994-2005.

Oliveros, J. C., Franch, M., Tabas-Madrid, D., San-León, D., Montoliu, L., Cubas, P. y Pazos, F. (2016). Breaking-Cas-interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes. *Nucleic acids research*, *44*(W1), W267-W271.

Oppenheimer, D. G., Herman, P. L., Sivakumaran, S., Esch, J. y Marks, M. D. (1991). A myb gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. *Cell*, *67*(3), 483–493.

Park, Y. H., Lee, Y. J., Kang, J. S., Choi, Y. W. y Son, B. G. (2009). A Gene-based dCAPS Marker for Selecting *old-gold-crimson* (*og^c*) Fruit Color Mutation in Tomato. *Journal of Life Science*, *19*(1), 152-155.

Parry, M. A. J., Madgwick, P. J., Bayon, C., Tearall, K., Hernandez-Lopez, A., Baudo, M., Rakszegi, M., Hamada, W., Al-Yassin, A., Ouabbou, H., Labhilili, M. y Phillips, A. L. (2009). Mutation discovery for crop improvement. *Journal of Experimental Botany*, *60*(10), 2817–2825.

Payne, C. T., Zhang, F. y Lloyd, A. M. (2000). *GL3* encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. *Genetics*, *156*(3), 1349-1362.

Payne, T., Clement, J., Arnold, D. y Lloyd, A. (1999). Heterologous myb genes distinct from *GL1* enhance trichome production when overexpressed in *Nicotiana tabacum*. *Development*, *126*(4), 671-682.

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E. y Yanofsky, M. F. (2000). B and C floral organ identity functions require *SEPALLATTA* MADS-box genes. *Nature*, *405*(6783), 200–203.

Bibliografía

Peralta, I. E., Knapp, S. y Spooner, D. M. (2005). New species of wild tomatoes (*Solanum* section *Lycopersicon*: Solanaceae) from Northern Peru. *Systematic Botany*, *30*(2), 424-434.

Pérez-Martín, F., Yuste-Lisbona, F. J., Pineda, B., Angarita-Díaz, M. P., García-Sogo, B., Antón, T., Sánchez, S., Giménez, E., Atarés, A., Fernández-Lozano, A., Ortíz-Atienza, A., García-Alcázar, M., Castañeda, L., Fonseca, R., Capel, C., Goergen, G., Sánchez, J., Quispe, J. L., Capel, J., Angosto, T., Moreno, V. y Lozano, R. (2017). A collection of enhancer trap insertional mutants for functional genomics in tomato. *Plant Biotechnology Journal*, *15*(11), 1439–1452.

Peters, A. H., Kubicek, S., Mechtler, K., O'Sullivan, R. J., Derijck, A. A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tacibana, M., Martens, J. H. y Jenuwein, T. (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Molecular cell*, *12*(6), 1577-1589.

Pino-Nunes, L. E., de O. Figueira, A. V., Tulmann Neto, A., Zsögön, A., Piotto, F. A., Silva, J. A., Bernardi, W.F. y Peres, L. E. P. (2008). Induced mutagenesis and natural genetic variation in tomato 'Micro-Tom'. *International Symposium on Tomato in the Tropics, 821*, 63-72).

Pnueli, L., Carmel-Goren, L., Hareven, D., Gutfinger, T., Alvarez, J., Ganal, M., Zamir, D. y Lifschitz, E. (1998). The *SELF-PRUNING* gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of *CEN* and *TFL1*. *Development*, *125*(11), 1979–1989.

Pnueli, L., Hareven, D., Rounsley, S. D., Yanofsky, M. F. y Lifschitz, E. (1994). Isolation of the tomato *AGAMOUS* gene *TAG1* and analysis of its homeotic role in transgenic plants. *The Plant Cell*, *6*(2), 163-173.

Pnueli, L., Hareven, D., Broday, L., Hurwitz, C. y Lifschitz, E. (1994). The *TM5* MADS box gene mediates organ differentiation in the three inner whorls of tomato flowers. *Plant Cell*, *6*(2), 175–186.

Pnueli, L., Abu-Abeid, M., Zamir, D., Nacken, W., Schwarz-Sommer, Z. y Lifschitz, E. (1991). The MADS box gene family in tomato: temporal expression during floral

development, conserved secondary structures and homology with homeotic genes from *Antirrhinum* and *Arabidopsis*. *The Plant Journal*, *1*(2), 255-266.

Poli, Y., Basava, R. K., Panigrahy, M., Vinukonda, V. P., Dokula, N. R., Voleti, S. R., Desiraju, S. y Neelamraju, S. (2013). Characterization of a Nagina22 rice mutant for heat tolerance and mapping of yield traits. *Rice*, *6*(1), 36.

Quinet, M., Bataille, G., Dobrev, P. I., Capel, C., Gómez, P., Capel, J., Lutts, S., Motyka, V., Angosto, T. y Lozano, R. (2014). Transcriptional and hormonal regulation of petal and stamen development by *STAMENLESS*, the tomato (*Solanum lycopersicum* L.) orthologue to the B-class *APETALA3* gene. *Journal of experimental botany*, *65*(9), 2243-2256.

Quinet, M., Dubois, C., Goffin, M. C., Chao, J., Dielen, V., Batoko, H., Boutry, M. y Kinet, J. M. (2006). Characterization of tomato (*Solanum lycopersicum* L.) mutants affected in their flowering time and in the morphogenesis of their reproductive structure. *Journal of experimental botany*, *57*(6), 1381-1390.

R Development Core Team. (2011). R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2011.

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. y Caicedo, A. (2020). Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution*.

Rick, C. M. (1978). The tomato. *Scientific American*, *239*(2), 76-89.

Rick, C. M., Quiros, C. F., Lange, W. H. y Stevens, M. A. (1976). Monogenic control of resistance in the tomato to the tobacco flea beetle: probable repellance by foliage volatiles. *Euphytica*, *25*(1), 521-530.

Rick, C. M. y Fobes, J. F. (1975). Allozyme Variation in the Cultivated Tomato and Closely Related Species. *Bulletin of the Torrey Botanical Club*, *102*(6), 376.

Rick, C. M. y Butler, L. (1956). Cytogenetics of the Tomato. *Advances in Genetics*, *8*, 267–382.

Riechmann, J. L. y Meyerowitz, E. M. (1997). MADS domain proteins in plant development. *Biological chemistry*, *378*(10), 1079-1102.

Bibliografía

Ries, G., Heller, W., Puchta, H., Sandermann, H., Seidlitz, H. K. y Hohn, B. (2000). Elevated UV-B radiation reduces genome stability in plants. *Nature*, *406*(6791), 98-101.

Rijpkema, A. S., Royaert, S., Zethof, J., van der Weerden, G., Gerats, T., y Vandenbussche, M. (2006). Analysis of the Petunia *TM6* MADS box gene reveals functional divergence within the *DEF/AP3* lineage. *The Plant Cell*, *18*(8), 1819-1832.

Saito, T., Ariizumi, T., Okabe, Y., Asamizu, E., Hiwasa-Tanase, K., Fukuda, N., Mizoguchi, T., Yamazaki, Y., Aoki, K. y Ezura, H. (2011). TOMATOMA: a novel tomato mutant database distributing Micro-Tom mutant collections. *Plant and cell physiology*, *52*(2), 283-296.

Salinas, M., Capel, C., Alba, J. M., Mora, B., Cuartero, J., Fernández-Muñoz, R., Lozano, R. y 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.

Schellmann, S., Schnittger, A., Kirik, V., Wada, T., Okada, K., Beerman, A., Thumfahrt, J., Jürgens, G. y Hülskamp, M. (2002). *TRIPTYCHON* and *CAPRICE* mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. *EMBO Journal*, *21*(19), 5036–5046.

Schnittger, A. y Hülskamp, M. (2002). Trichome morphogenesis: a cell–cycle perspective. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1422), 823-826.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F. X., Jürgens, G. y Laux, T. (2000). The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the *CLAVATA* and *WUSCHEL* genes. *Cell*, *100*(6), 635–644.

Schwab, B., Folkers, U., Ilgenfritz, H. y Hülskamp, M. (2000). Trichome morphogenesis in Arabidopsis. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1399), 879-883.

Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. y Sommer, H. (1990). Genetic control of flower development by homeotic genes in *Antirrhinum majus*. *Science*, *250*(4983), 931-936.

Sega, G. A. (1984). A review of the genetic effects of ethyl methanesulfonate. *Mutation Research/Reviews in Genetic Toxicology,* 134(2–3), 113–142.

Serna, L. y Martin, C. (2006). Trichomes: different regulatory networks lead to convergent structures. *Trends in plant science*, *11*(6), 274-280.

Shikata, M., Hoshikawa, K., Ariizumi, T., Fukuda, N., Yamazaki, Y. y Ezura, H. (2016). TOMATOMA update: phenotypic and metabolite information in the Micro-Tom mutant resource. *Plant and Cell Physiology*, *57*(1), e11-e11.

Shuai, B., Reynaga-Pena, C. G. y Springer, P. S. (2002). The lateral organ boundaries gene defines a novel, plant-specific gene family. *Plant physiology*, *129*(2), 747-761.

Sikora, P., Chawade, A., Larsson, M., Olsson, J. y Olsson, O. (2011). Mutagenesis as a tool in plant genetics, functional genomics, and breeding. *International Journal of Plant Genomics, 2011.*

Simmons, A. T. y Gurr, G. M. (2005). Trichomes of *Lycopersicon* species and their hybrids: effects on pests and natural enemies. *Agricultural and Forest Entomology*, *7*(4), 265-276.

Simmons, A. T., Gurr, G. M., McGrath, D., Martin, P. M. y Nicol, H. I. (2004). Entrapment of *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae) on glandular trichomes of *Lycopersicon* species. *Australian Journal of Entomology*, *43*(2), 196-200.

Soltis, D., Soltis, P., Endress, P., Chase, M. W., Manchester, S., Judd, W., Majure, L. y Mavrodiev, E. (2018). Phylogeny and evolution of the angiosperms: revised and updated edition. University of Chicago Press.

Somssich, M., Je, B. Il, Simon, R. y Jackson, D. (2016). CLAVATA-WUSCHEL signaling in the shoot meristem. *Development*, *143*(18), 3238–3248.

Song, L., Huang, S. S. C., Wise, A., Castanon, R., Nery, J. R., Chen, H., Watanabe, M., Thomas, J., Bar-Joseph, Z. y Ecker, J. R. (2016). A transcription factor hierarchy defines an environmental stress response network. *Science*, *354*(6312), aag1550.

Bibliografía

Sundström, J. F., Nakayama, N., Glimelius, K. y Irish, V. F. (2006). Direct regulation of the floral homeotic *APETALA1* gene by *APETALA3* and *PISTILLATA* in *Arabidopsis*. *The Plant Journal*, *46*(4), 593-600.

Szabados, L., Kovács, I., Oberschall, A., Ábrahám, E., Kerekes, I., Zsigmond, L., Nagy, R., Alvarado, M., Krasovskaja, I., Gál, M., Berente, A., Rédei, G. P., Haim, B. A. y Koncz, C. (2002). Distribution of 1000 sequenced T-DNA tags in the *Arabidopsis* genome. *The Plant Journal*, *32*(2), 233–242.

Tanksley, S. D. (2004). The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell*, *16* (Suppl. 1), S181-S189.

Tax, F. E. y Vernon, D. M. (2001). T-DNA-associated duplication/translocations in Arabidopsis. Implications for mutant analysis and functional genomics. *Plant Physiology*, *126*(4), 1527–1538.

Till, B. J., Cooper, J., Tai, T. H., Colowit, P., Greene, E. A., Henikoff, S. y Comai, L. (2007). Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biology*, *7*(1), 19.

Till, B. J., Reynolds, S. H., Weil, C., Springer, N., Burtner, C., Young, K., Bowers, E., Codomo, C. A., Enns, L. C., Odden, A. R., Greene, E. A., Comai, L. y Henikoff, S. (2004). Discovery of induced point mutations in maize genes by TILLING. *BMC plant biology*, *4*(1), 12.

Tomato Genome Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, *485*(7400), 635.

Uchida, N. y Torii, K. U. (2019). Stem cells within the shoot apical meristem: identity, arrangement and communication. *Cellular and Molecular Life Sciences*, *76*(6), 1067–1080.

van Der Krol, A. R. y Chua, N. H. (1993). Flower development in *Petunia*. *The Plant Cell*, *5*(10), 1195.

van Harten, A. M. (1998). Mutation breeding: theory and practical applications. Cambridge University Press.

van Houwelingen, A., Souer, E., Spelt, K., Kloos, D., Mol, J. y Koes, R. (1998). Analysis of flower pigmentation mutants generated by random transposon mutagenesis in

Petunia hybrida. *The Plant Journal*, *13*(1), 39–50.

Van Ooijen, J. W. (2006). JoinMap 4. Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen, Netherlands, 33.

Vandenbussche, M., Zethof, J., Royaert, S., Weterings, K. y Gerats, T. (2004). The duplicated B-class heterodimer model: whorl-specific effects and complex genetic interactions in *Petunia hybrida* flower development. *The Plant Cell*, *16*(3), 741-754.

Vazquez-Vilar, M., Bernabé-Orts, J. M., Fernandez-del-Carmen, A., Ziarsolo, P., Blanca, J., Granell, A. y Orzaez, D. (2016). A modular toolbox for gRNA–Cas9 genome engineering in plants based on the GoldenBraid standard. *Plant Methods*, *12*(1), 10.

Vercosa de Magalhaes, S. T., Jham, G. N., Picanço, M. C. y Magalhães, G. (2001). Mortality of second-instar larvae of *Tuta absoluta* produced by the hexane extract of *Lycopersicon hirsutum* f. *glabratum* (PI 134417) leaves. *Agricultural and Forest Entomology*, *3*(4), 297-303.

Vogt, T. y Jones, P. (2000). Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends in plant science*, *5*(9), 380-386.

Vrebalov, J., Pan, I. L., Arroyo, A. J. M., McQuinn, R., Chung, M., Poole, M., Rose, J., Seymour, G., Grandillo, S., Giovannoni, J. y Irish, V. F. (2009). Fleshy fruit expansion and ripening are regulated by the tomato *SHATTERPROOF* gene *TAGL1*. *The Plant Cell*, *21*(10), 3041-3062.

Vrebalov, J., Ruezinsky, D., Padmanabhan, V., White, R., Medrano, D., Drake, R., Schuch, W. y Giovannoni, J. (2002). A MADS-box gene necessary for fruit ripening at the tomato *ripening-inhibitor* (*rin*) locus. *Science*, *296*(5566), 343–346.

Wang, R., Tavano, E. C. da R., Lammers, M., Martinelli, A. P., Angenent, G. C. y de Maagd, R. A. (2019). Re-evaluation of transcription factor function in tomato fruit development and ripening with CRISPR/Cas9-mutagenesis. *Scientific Reports*, *9*(1).**Wagner, G. J., Wang, E. y Shepherd, R. W.** (2004). New approaches for studying and exploiting an old protuberance, the plant trichome. *Annals of Botany*, *93*(1), 3–11.

Wang, Y., Tang, X., Cheng, Z., Mueller, L., Giovannoni, J. y Tanksley, S. D. (2006).

Euchromatin and pericentromeric heterochromatin: comparative composition in the tomato genome. *Genetics*, *172*(4), 2529-2540.

Watanabe, S., Mizoguchi, T., Aoki, K., Kubo, Y., Mori, H., Imanishi, S., Yamazaki, Y., Shibata, D. y Ezura, H. (2007). Ethylmethanesulfonate (EMS) mutagenesis of *Solanum lycopersicum* cv. Micro-Tom for large-scale mutant screens. *Plant Biotechnology*, *24*(1), 33-38.

Weigel, D. y Meyerowitz, E. M. (1994). The ABCs of floral homeotic genes. *Cell*, *78*(2), 203–209.

Wester, K., Digiuni, S., Geier, F., Timmer, J., Fleck, C. y Hülskamp, M. (2009). Functional diversity of R3 single-repeat genes in trichome development. *Development*, *136*(9), 1487-1496.

Wilkins, T. A., Rajasekaran, K. y Anderson, D. M. (2000). Cotton biotechnology. *Critical Reviews in Plant Sciences*, *19*(6), 511–550.

Williams, W. G., Kennedy, G. G., Yamamoto, R. T., Thacker, J. D. y Bordner, J. (1980). 2-Tridecanone: a naturally occurring insecticide from the wild tomato *Lycopersicon hirsutum* f. *glabratum*. *Science*, *207*(4433), 888-889.

Wu, L., Tian, Z. y Zhang, J. (2018). Functional dissection of auxin response factors in regulating tomato leaf shape development. *Frontiers in plant science*, *9*(957).

Wu, J. L., Wu, C., Lei, C., Baraoidan, M., Bordeos, A., Madamba, M. R. S., Ramos-Pamplona, R., Mauleon, R., Portugal, A., Ulat, V. J., Bruskiewich, R., Wang, G., Leach, J., Khush, G. y Leung, H. (2005). Chemical-and irradiation-induced mutants of indica rice IR64 for forward and reverse genetics. *Plant molecular biology*, *59*(1), 85-97.

Yadav, R. K., Perales, M., Gruel, J., Girke, T., Jönsson, H. y Venugopala Reddy, G. (2011). WUSCHEL protein movement mediates stem cell homeostasis in the Arabidopsis shoot apex. *Genes and Development*, *25*(19), 2025–2030.

Yang, C., Li, H., Zhang, J., Luo, Z., Gong, P., Zhang, C., Li. J., Wang, T., Zhang, Y., Lu, Y. y Ye, Z. (2011). A regulatory gene induces trichome formation and embryo lethality in tomato. *Proceedings of the National Academy of Sciences*, *108*(29), 11836-11841.

Yang, C., Li, H., Zhang, J., Wang, T. y Ye, Z. (2011). Fine-mapping of the *woolly* gene controlling multicellular trichome formation and embryonic development in tomato. *Theoretical and applied genetics*, *123*(4), 625-633.

Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A. y Meyerowitz, E. M. (1990). The protein encoded by the *Arabidopsi*s homeotic gene *agamous* resembles transcription factors. *Nature*, *346*(6279), 35-39.

Yant, L., Mathieu, J., Dinh, T. T., Ott, F., Lanz, C., Wollmann, H., Chen, X. y Schmid, M. (2010). Orchestration of the floral transition and floral development in Arabidopsis by the bifunctional transcription factor *APETALA2*. *The Plant Cell*, *22*(7), 2156-2170.

Yu, H., Ito, T., Wellmer, F. y Meyerowitz, E. M. (2004). Repression of *AGAMOUS-LIKE 24* is a crucial step in promoting flower development. *Nature genetics, 36*(2), 157-161.

Yuan, Y., Mei, L., Wu, M., Wei, W., Shan, W., Gong, Z., Zhang, Q., Yang, F., Yan, F., Zhang, Q., Luo Y., Xu, X., Zhang, W., Miao, M., Lu, W., Li, Z. y Deng, W. Luo, Y. (2018). SlARF10, an auxin response factor, is involved in chlorophyll and sugar accumulation during tomato fruit development. *Journal of experimental botany*, *69*(22), 5507-5518.

Yuan, L., Dou, Y., Kianian, S. F., Zhang, C. y Holding, D. R. (2014). Deletion mutagenesis identifies a haploinsufficient role for γ-zein in *opaque2* endosperm modification. *Plant physiology*, *164*(1), 119-130.

Yuste-Lisbona, F. J., Quinet, M., Fernández-Lozano, A., Pineda, B., Moreno, V., Angosto, T. y Lozano, R. (2016). Characterization of *vegetative inflorescence* (*mc-vin*) mutant provides new insight into the role of *MACROCALYX* in regulating inflorescence development of tomato. *Scientific Reports*, *6*, 18796.

Zhang, J., Wang, Y., Naeem, M., Zhu, M., Li, J., Yu, X., Hu, Z. y Chen, G. (2019). An AGAMOUS MADS-box protein, SlMBP3, regulates the speed of placenta liquefaction and controls seed formation in tomato. *Journal of experimental botany*, *70*(3), 909-924.

Zhang, F., Gonzalez, A., Zhao, M., Payne, C. T. y Lloyd, A. (2003). A network of

redundant bHLH proteins functions in all *TTG1*-dependent pathways of *Arabidopsis*. *Development*, *130*(20), 4859-4869.

Zhao, Q. Q., Lin, R. N., Li, L., Chen, S. y He, X. J. (2019). A methylated-DNA-binding complex required for plant development mediates transcriptional activation of promoter methylated genes. *Journal of integrative plant biology*, *61*(2), 120-139.

Zimmermann, I. M., Heim, M. A., Weisshaar, B. y Uhrig, J. F. (2004). Comprehensive identification of Arabidopsis thaliana MYB transcription factors interacting with R/B-like BHLH proteins. *Plant Journal*, *40*(1), 22–34.

Zimmermann, F. K. (1977). Genetic effects of nitrous acid. *Mutation Research/Reviews in Genetic Toxicology*, *39*(2), 127–148.

Universidad de Almería

Grupo de Investigación Universidad de Almería

Genética y Fisiología del Desarrollo Vegetal