

**IDENTIFICACIÓN Y CARACTERIZACIÓN DE GENES  
IMPLICADOS EN LA VARIACIÓN NATURAL PARA EL  
PATRÓN DE TRICOMAS EN ARABIDOPSIS**

**TESIS DOCTORAL**

**Noelia Arteaga Ramos**

**Madrid, 2019**

Programa de Doctorado en Biociencias Moleculares

**Universidad Autónoma de Madrid**

**IDENTIFICACIÓN Y CARACTERIZACIÓN DE GENES  
IMPLICADOS EN LA VARIACIÓN NATURAL PARA EL  
PATRÓN DE TRICOMAS EN ARABIDOPSIS**

TESIS DOCTORAL

**Noelia Arteaga Ramos**

DIRECTOR

**Dr. Carlos Alonso Blanco**

**Madrid, 2019**

Programa de Doctorado en Biociencias Moleculares

**Universidad Autónoma de Madrid**



**IDENTIFICACIÓN Y CARACTERIZACIÓN DE GENES  
IMPLICADOS EN LA VARIACIÓN NATURAL PARA EL  
PATRÓN DE TRICOMAS EN ARABIDOPSIS**

**Dr. Carlos Alonso Blanco**

Director

**Dr. Luis E. Hernández Rodríguez**

Tutor

**Noelia Arteaga Ramos**

Licenciada en Biología

Carlos Alonso Blanco, Investigador Científico en el Centro Nacional de Biotecnología (CNB), Consejo Superior de Investigaciones Científicas (CSIC), CERTIFICA que el trabajo presentado en esta Memoria de Investigación ha sido realizado por Noelia Arteaga Ramos bajo mi dirección.

Departamento de Biología Molecular

Facultad de Ciencias

**Universidad Autónoma de Madrid**

**Centro Nacional de Biotecnología (CNB-CSIC)**

## **AGRADECIMIENTOS**

En primer lugar, quiero agradecer a Carlos que me diese la oportunidad de realizar esta tesis doctoral bajo su dirección. También darle las gracias por su inagotable paciencia y dedicación, así como por la pasión que transmite y contagia por cada proyecto que lleva a cabo.

Igualmente, mi más sincero agradecimiento a Belén por ser mi ángel de la guarda durante estos cuatro años, siempre estaré agradecida por su inmensa ayuda. También quiero agradecer a Merche todo el apoyo técnico y personal recibido, trabajar a su lado es un verdadero placer. Gracias a Almudena y Alberto por la buena disposición en todo momento, tanto dentro como fuera del laboratorio.

Gracias a todos mis compañeros del departamento de Genética Molecular de Plantas por su ayuda siempre que la he necesitado, y por haber hecho del CNB mi segunda casa. Agradecer también el soporte técnico por parte de los servicios del centro, principalmente al servicio de invernadero y de cultivo in vitro.

No puedo terminar sin agradecer a mi familia y a Adrián todo su apoyo y comprensión a lo largo de estos años.

Por último, agradecer la financiación recibida por parte del Ministerio de Ciencia, Innovación y Universidades mediante la Ayuda para Contratos Predoctorales para la Formación de Doctores BES-2014-069809, asociada al proyecto BIO2013-45407-P.

## RESUMEN

*Arabidopsis thaliana* es la principal especie modelo vegetal, ampliamente utilizada para determinar los mecanismos genéticos y moleculares del desarrollo y la fisiología de las plantas anuales. Además, recientemente, gracias a su amplia distribución geográfica y su gran diversidad genética natural, *A. thaliana* se ha convertido también en una planta modelo para el estudio de procesos ecológicos y evolutivos. En particular, nuestro laboratorio está interesado en determinar las bases genéticas, moleculares y evolutivas de la variación natural para el patrón de tricomas, ya que este carácter se ha implicado en la adaptación de las plantas a numerosos factores de estrés abiótico y biótico. Con este fin se ha caracterizado previamente una colección regional de 174 accesiones de *A. thaliana* procedentes de la península ibérica a nivel genómico y ambiental. Interesantemente, a partir de esta colección se han identificado varias accesiones que desarrollan tricomas en el fruto, y su análisis genético previo ha mostrado que viene determinado por cinco loci denominados *Fruit Trichome Density (FTD1-5)*.

Esta tesis se propone identificar nuevos componentes genéticos y moleculares de la variación natural para el patrón de tricomas en *A. thaliana*, tanto en fase vegetativa como reproductiva, así como en respuesta a la herbivoría. Para alcanzar este objetivo general se han llevado a cabo las siguientes actividades:

1. Se ha realizado un análisis de asociación genómica para el patrón de tricomas en distintos órganos de la fase vegetativa y reproductiva a partir de la caracterización fenotípica de las 174 accesiones de la colección regional de la península ibérica. Este estudio ha permitido identificar varios genes candidatos conocidos (ej. *ETC1*, *MYC1*) y nuevos (ej. *At2g04925*, *At2g40130*, *At4g04985* y *At5g53420*) que probablemente contribuyan a la variación natural para el patrón de tricomas en uno o varios órganos, o en su respuesta al tratamiento con ácido jasmónico como un estímulo análogo a la herbivoría.

2. Se ha llevado a cabo un mapeo de alta precisión del locus *FTD1* utilizando una amplia población experimental derivada del cruzamiento entre las accesiones Ler y Doñ-0. Este análisis ha identificado a *EGL3* como único gen candidato para *FTD1*, para el cual la accesión Doñ-0 presenta un alelo de ganancia de función.

3. Se han identificado los genes correspondientes a los loci *FTD2*, *FTD3* y *FTD5* mediante la caracterización fenotípica y molecular de numerosas líneas transgénicas portadoras de diferentes alelos naturales para genes candidatos. De esta forma se ha demostrado que *TCL1*, *GL1* y *TRY* son los genes responsables del desarrollo de tricomas en los frutos de *A. thaliana*.

## ABSTRACT

*Arabidopsis thaliana* is the main research model plant, which is widely used to determine the genetic and molecular mechanisms involved in the development and physiology of annual plants. In addition, thanks to its broad geographic distribution and its large amount of natural genetic diversity, *A. thaliana* has also become recently a model plant to study ecological and evolutionary processes. In particular, our laboratory is interested in understanding the genetic, molecular and evolutionary bases of the natural variation for trichome patterning because this trait has been involved in adaptation to numerous abiotic and biotic stress factors. To this end, we are exploiting a regional collection of 174 *A. thaliana* accessions from the Iberian Peninsula, which has been previously characterized at the genomic and environmental level. Interestingly, several accessions have been identified in this collection as developing trichomes in fruits, and previous genetic analyses have shown that this trait is determined by five loci named as *Fruit Trichome Density* (*FTD1-5*).

This thesis aims to identify new genetic and molecular components contributing to the natural variation for trichome patterning in *A. thaliana*, both in vegetative and reproductive phases, and in response to herbivory. To achieve this goal, the following tasks have been carried out:

1. Genome-wide association analyses for trichome patterning in different organs of the vegetative and reproductive phases have been done by analysing phenotypically 174 accessions of the regional collection from the Iberian Peninsula. This study has identified several known (eg. *ETC1*, *MYC1*) and new (eg. *At2g04925*, *At2g40130*, *At4g04985* y *At5g53420*) candidate genes that likely contribute to the natural variation for trichome patterning in one or several plant organs, or in response to jasmonic acid treatment, as a proxy of herbivory.

2. A high-resolution fine mapping has been done for *FTD1* locus using a large experimental population derived from the cross between accessions *Ler* and *Doñ-0*. This analysis has identified *EGL3* as the only candidate gene for *FTD1*, for which *Doñ-0* accession carries a gain-of-function allele.

3. The genes underlying *FTD2*, *FTD3* and *FTD5* loci have been identified by means of phenotypic and molecular characterization of numerous transgenic lines carrying different natural alleles of candidate genes. Thus, it has been demonstrated that *TCL1*, *GL1* and *TRY* are the genes causing the development of trichomes in fruits of *A. thaliana*.

# ÍNDICE DE CONTENIDO

AGRADECIMIENTOS

RESUMEN

ABREVIATURAS.....	1
ÍNDICE DE FIGURAS.....	5
ÍNDICE DE TABLAS.....	9
INTRODUCCIÓN.....	11
1. Arabidopsis como especie modelo para el estudio del patrón de tricomas .....	13
2. Regulación del patrón de tricomas en Arabidopsis.....	14
2.1. Fases del desarrollo de los tricomas.....	15
2.2. Bases genéticas del desarrollo de los tricomas.....	16
2.3. Modelo de regulación del patrón de tricomas en Arabidopsis.....	18
2.4. Efecto de las hormonas y otros factores sobre la regulación del patrón de tricomas.....	20
3. Variación natural para el patrón de tricomas en Arabidopsis.....	22
3.1. Análisis genéticos de la variación cualitativa y cuantitativa en Arabidopsis.....	22
3.2. Análisis genómicos de la variación del patrón de tricomas.....	24
4. Análisis regional de la variación natural en la península ibérica.....	26
4.1. La península ibérica como escenario ideal para el estudio de los mecanismos genéticos y moleculares de la adaptación de Arabidopsis.....	26
4.2. Diversidad genética para el patrón de tricomas de Arabidopsis en la península ibérica.....	27
OBJETIVOS.....	29
MATERIALES Y MÉTODOS.....	33
1. Material vegetal.....	35
1.1. Acciones naturales para el análisis de asociación genómica.....	35
1.2. Líneas de introgresión de los loci <i>FTD2</i> , <i>FTD3</i> y <i>FTD5</i> .....	35

1.3. Población de mapeo de <i>FTD1</i> .....	37
1.4. Mutantes de genes del patrón de tricomas.....	37
1.5. Líneas transgénicas.....	38
2. Condiciones de cultivo.....	39
2.1. Mapeo fino de <i>FTD1</i> .....	39
2.2. Análisis fenotípicos y de expresión génica en líneas de introgresión o transgénicas.....	40
2.3. Caracterización fenotípica de las accesiones naturales y su tratamiento hormonal.....	40
3. Análisis fenotípico del patrón de tricomas.....	41
4. Genotipado de marcadores moleculares.....	43
4.1. Genotipado de las líneas de introgresión.....	44
4.2. Genotipado de la población de mapeo de <i>FTD1</i> .....	44
4.3. Genotipado de líneas mutantes.....	45
4.4. Genotipado de líneas transgénicas.....	46
5. Secuenciación del gen <i>EGL3</i> .....	47
6. Construcciones genómicas y transformación.....	49
7. Análisis de expresión génica.....	50
8. Análisis de asociación genómica.....	51
9. Otros análisis estadísticos.....	53
10. Análisis de diversidad nucleotídica .....	54
RESULTADOS.....	55
1. Análisis de asociación genómica para el patrón de tricomas en <i>Arabidopsis</i> .....	57
1.1. Variación genética para el patrón de tricomas en <i>Arabidopsis</i> y su respuesta al ácido jasmónico.....	57



1.1.1. Variación natural para la densidad de tricomas en fase vegetativa.....	57
1.1.2. Variación natural para el patrón de tricomas en fase reproductiva.....	60
1.1.3. Variación natural para el patrón de tricomas vegetativo en plantas tratadas con ácido jasmónico.....	62
1.1.4. Respuesta del patrón de tricomas vegetativo al ácido jasmónico.....	63
1.2. Bases genómicas de la variación natural para el patrón de tricomas y su respuesta al ácido jasmónico en Arabidopsis.....	64
1.2.1. Análisis de asociación genómica para la densidad de tricomas en fase vegetativa y su respuesta al ácido jasmónico.....	65
1.2.1.1. Asociación genómica para la densidad de tricomas en la hoja.....	65
1.2.1.2. Asociación genómica para la densidad de tricomas en la hoja de plantas tratadas con ácido jasmónico.....	67
1.2.1.3. Asociación genómica para la respuesta del patrón de tricomas en la hoja al ácido jasmónico.....	70
1.2.2. Análisis de asociación genómica para el patrón de tricomas en fase reproductiva.....	71
1.2.2.1. Asociación genómica para el patrón de tricomas en el tallo.....	71
1.2.2.2. Asociación genómica para el patrón de tricomas en el pedicelo.....	73
2. Caracterización genética y molecular de <i>FTD1</i> .....	76
2.1. Mapeo fino de <i>FTD1</i> .....	76
2.2. Diversidad nucleotídica y análisis filogenético de <i>EGL3</i> en la península ibérica...78	
2.3. Análisis de expresión de <i>EGL3</i> en accesiones parentales.....	83
2.4. Análisis de líneas mutantes de <i>EGL3</i> .....	84
2.5. Análisis funcional de los alelos de <i>EGL3</i> de Ler y Doñ-0 mediante complementación en plantas transgénicas.....	86
3. Caracterización genética y molecular de <i>FTD2</i> , <i>FTD3</i> y <i>FTD5</i> .....	87

3.1. Caracterización genética de los loci <i>FTD2</i> , <i>FTD3</i> y <i>FTD5</i> .....	87
3.2. Aislamiento y caracterización molecular de <i>FTD2</i> .....	89
3.2.1. Aislamiento de <i>FTD2</i> .....	89
3.2.2. Identificación del polimorfismo funcional de <i>TCL1/FTD2</i> .....	94
3.2.3. Caracterización funcional de la variación natural Ler/Doñ-0 de <i>TCL1</i> .....	95
3.3. Aislamiento y caracterización molecular de <i>FTD5</i> .....	99
3.3.1. Aislamiento de <i>FTD5</i> .....	99
3.3.2. Identificación del polimorfismo funcional de <i>TRY/FTD5</i> .....	103
3.3.3. Caracterización funcional de la variación natural Ler/Doñ-0 de <i>TRY</i> .....	104
3.4. Aislamiento y caracterización molecular de <i>FTD3</i> .....	108
3.4.1. Aislamiento de <i>FTD3</i> .....	108
3.4.2. Caracterización funcional de la variación natural Ler/Doñ-0 de <i>GL1</i> .....	110
3.5. Análisis de expresión de los genes <i>TCL1</i> , <i>TRY</i> y <i>GL1</i> en líneas de introgresión.....	112
DISCUSIÓN.....	117
1. Variación natural para el patrón de tricomas a lo largo del desarrollo.....	119
1.1. Arabidopsis muestra una gran variación para el patrón de tricomas en todos sus órganos.....	119
1.2. Bases genéticas del patrón de tricomas en distintos órganos y fases del desarrollo.....	121
2. Mecanismos moleculares del desarrollo de tricomas en los frutos.....	123
2.1. Alelos naturales de ganancia de función en los genes promotores del patrón de tricomas <i>EGL3</i> y <i>GL1</i> .....	124
2.2. Alelos naturales de pérdida de función parcial en los genes represores del patrón de tricomas <i>TCL1</i> y <i>TRY</i> .....	127

<b>2.3. Efecto de la variación natural sobre la regulación del patrón de tricomas</b>	
en el fruto.....	129
<b>3. Evolución del patrón de tricomas en los frutos de Arabidopsis.....</b>	<b>129</b>
CONCLUSIONES.....	133
BIBLIOGRAFÍA.....	137
ANEXO .....	151

## CLAVE DE ABREVIATURAS

**ADN:** ácido desoxirribonucleico

**ADNc:** ácido desoxirribonucleico complementario

**AH:** Anchura de la hoja

**AJ:** Ácido jasmónico

**ANOVA:** Análisis de varianza -*ANalysis Of VAriance*-

**ARN:** Ácido ribonucleico

**CAPS:** Secuencia polimórfica amplificada y cortada -*Cleaved Amplified Polymorphic Sequence*-

**CKs:** Citoquininas

**DC:** Día corto

**dCAPS:** CAPS degenerado -*degenerated Cleaved Amplified Polymorphic Sequence*-

**DL:** Día largo

**DTH:** Densidad de tricomas en la hoja

**DTHAJ:** Densidad de tricomas en la hoja de plantas tratadas con ácido jasmónico

**EE:** Error estándar de la media

**FT:** Factor de transcripción

**FTD:** *Fruit Trichome Density*

**GAs:** Giberelinas

**GWAS:** Análisis de asociación genómica -*Genome-Wide Association Studies*-

**$h^2_b$ :** Heredabilidad en sentido amplio

**Indel:** polimorfismo de inserción/delección

**kb:** Kilobase

**LB:** Borde izquierdo -*Left Border*-

**LH:** Longitud de la hoja

**LI:** Línea de introgresión

**MAF:** Frecuencia del alelo menor -*Minor Allele Frequency*-

**Mb:** Megabase

**miARN:** microARN

**MLM:** Modelo linear mixto

**NTF:** Número de tricomas en el fruto

**NTH:** Número de tricomas en la hoja

**NTHAJ:** Número de tricomas en la hoja de plantas tratadas con ácido jasmónico

**NTS:** Número de tricomas en los sépalos

**pb:** Pares de bases

**PCR:** Reacción en cadena de la polimerasa -*Polymerase Chain Reaction*-

**PDTH:** Plasticidad de la densidad de tricomas en la hoja

**PNTH:** Plasticidad del número de tricomas en la hoja

**PSH:** Plasticidad de la superficie de la hoja

**PTF:** Patrón de tricomas en el fruto

**PTP:** Patrón de tricomas en el pedicelo

**PTT:** Patrón de tricomas en el tallo

**QTL:** Locus de carácter cuantitativo -*Quantitative Trait Locus*-

**R<sup>2</sup>:** Proporción de varianza explicada.

**RT-PCR:** Reacción en cadena de la polimerasa a tiempo real -*Real Time Polymerase Chain Reaction*-

**RILs:** Líneas recombinantes consanguíneas -*Recombinant Inbred Lines*-

**SEC:** Marcador genotipado por secuenciación

**SH:** Superficie de la hoja

**SHAJ:** Superficie de la hoja en plantas tratadas con ácido jasmónico

**SNP:** Polimorfismo de un único nucleótido -*Single Nucleotide Polymorphism*-

**UTR:** Región no traducida -*Untranslated Region*-

**V<sub>G</sub>**: Variación genética

**V<sub>AJ</sub>**: Variación debida al efecto del tratamiento de ácido jasmónico

**V<sub>GxAJ</sub>**: Variación debida a la interacción genotipo x tratamiento de ácido jasmónico

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Variación natural de <i>Arabidopsis</i> .....	14
<b>Figura 2.</b> Fases del desarrollo de los tricomas en <i>Arabidopsis</i> .....	16
<b>Figura 3.</b> Modelo de la regulación del patrón de tricomas en <i>Arabidopsis</i> .....	19
<b>Figura 4.</b> Genes y QTLs implicados en la variación natural para el patrón de tricomas en <i>Arabidopsis</i> .....	24
<b>Figura 5.</b> Variación natural para el patrón de tricomas en el fruto de <i>Arabidopsis</i> en la península ibérica.....	28
<b>Figura 6.</b> Esquema de la obtención de las LIs para los loci <i>FTD2</i> , <i>FTD3</i> y <i>FTD5</i> .....	36
<b>Figura 7.</b> Planta madre de la población de mapeo de <i>FTD1</i> .....	37
<b>Figura 8.</b> Caracteres del patrón de tricomas fenotipados.....	43
<b>Figura 9.</b> Variación fenotípica para la densidad de tricomas de la hoja de <i>Arabidopsis</i> .....	58
<b>Figura 10.</b> Distribución de frecuencias de los caracteres relacionados con el patrón de tricomas en fase vegetativa.....	59
<b>Figura 11.</b> Variación fenotípica para el patrón de tricomas en fase reproductiva en <i>Arabidopsis</i> .....	61
<b>Figura 12.</b> Distribución de frecuencias del patrón de tricomas en fase reproductiva...	61
<b>Figura 13.</b> Normas de reacción de los caracteres del patrón de tricomas vegetativo en relación al tratamiento con ácido jasmónico.....	63
<b>Figura 14.</b> Distribución de frecuencias de la plasticidad al tratamiento de ácido jasmónico en fase vegetativa.....	64
<b>Figura 15.</b> Gráficos de tipo Manhattan de los análisis de asociación genómica para el patrón de tricomas de la hoja.....	69
<b>Figura 16.</b> Distribución geográfica de polimorfismos en algunos de los genes más relevantes detectados en los análisis de asociación genómica.....	71

<b>Figura 17.</b> Gráficos de tipo Manhattan de los análisis de asociación genómica para el patrón de tricomas en fase reproductiva.....	75
<b>Figura 18.</b> Mapeo fino de <i>FTD1</i> .....	77
<b>Figura 19.</b> Comparación de la secuencia de <i>EGL3</i> en Ler, Doñ-0 y Col.....	78
<b>Figura 20.</b> Diversidad nucleotídica y árbol filogenético de la región promotora de <i>EGL3</i> .....	79
<b>Figura 21.</b> Diversidad nucleotídica y cambios de aminoácidos en la región codificante de <i>EGL3</i> en la península ibérica.....	82
<b>Figura 22.</b> Árbol filogenético de la región codificante de <i>EGL3</i> en la península ibérica..	83
<b>Figura 23.</b> Análisis de la expresión del gen <i>EGL3</i> .....	84
<b>Figura 24.</b> Líneas mutantes en <i>EGL3</i> .....	85
<b>Figura 25.</b> Patrón de tricomas de las líneas de introgresión para los loci <i>FTD</i> .....	88
<b>Figura 26.</b> Construcciones genómicas de <i>TCL1</i> .....	89
<b>Figura 27.</b> Patrón de tricomas de las líneas transgénicas de <i>TCL1</i> en el fondo genético <i>tcl1-Salk</i> .....	91
<b>Figura 28.</b> Patrón de tricomas de las líneas transgénicas de <i>TCL1</i> en el fondo genético LI-235.....	93
<b>Figura 29.</b> Relación entre el nivel de expresión de <i>TCL1</i> y el patrón de tricomas de las líneas transgénicas.....	98
<b>Figura 30.</b> Construcciones genómicas de <i>TRY</i> .....	99
<b>Figura 31.</b> Patrón de tricomas de las líneas transgénicas de <i>TRY</i> en el fondo genético <i>try-Salk</i> .....	100
<b>Figura 32.</b> Patrón de tricomas de las líneas transgénicas de <i>TRY</i> en el fondo genético LI-235.....	102
<b>Figura 33.</b> Relación entre el nivel de expresión de <i>TRY</i> y el patrón de tricomas de las líneas transgénicas.....	107



<b>Figura 34.</b> Patrón de tricomas de las líneas transgénicas de <i>GL1</i> en el fondo genético <i>gl1-1</i> .....	109
<b>Figura 35.</b> Patrón de tricomas de las líneas transgénicas de <i>GL1</i> en el fondo genético <i>tcl1-Salk try-Salk</i> .....	110
<b>Figura 36.</b> Expresión de <i>TCL1</i> , <i>TRY</i> y <i>GL1</i> en las líneas de introgresión.....	113
<b>Figura 37.</b> Relación entre el nivel de expresión de <i>TCL1</i> , <i>TRY</i> o <i>GL1</i> y el patrón de tricomas de las líneas de introgresión.....	115
<b>Figura 38.</b> Mecanismos moleculares de la variación natural para el patrón de tricomas en el fruto y la hoja en Ler y Doñ-0.....	125
<b>Figura 39.</b> Distribución geográfica de los polimorfismos funcionales Ler/Doñ-0 de <i>TCL1</i> , <i>GL1</i> y <i>TRY</i> .....	131

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> Líneas transgénicas para los genes candidatos de los loci <i>FTD2</i> , <i>FTD3</i> y <i>FTD5</i> .....	39
<b>Tabla 2.</b> Marcadores utilizados para genotipar las líneas de introgresión.....	44
<b>Tabla 3.</b> Marcadores desarrollados para el mapeo fino del locus <i>FTD1</i> .....	45
<b>Tabla 4.</b> Marcadores utilizados para el genotipado de las líneas mutantes.....	46
<b>Tabla 5.</b> Marcadores utilizados para el genotipado de los transgenes.....	46
<b>Tabla 6.</b> Oligonucleótidos utilizados para la secuenciación y clonación de <i>EGL3</i> .....	48
<b>Tabla 7.</b> Construcciones genómicas de los genes <i>EGL3</i> , <i>TCL1</i> , <i>GL1</i> y <i>TRY</i> .....	50
<b>Tabla 8.</b> Marcadores utilizados para estimar la expresión de los genes estudiados....	51
<b>Tabla 9.</b> Correlaciones entre los caracteres del patrón de tricomas en fase vegetativa..	60
<b>Tabla 10.</b> Correlación entre el patrón de tricomas en fase vegetativa y reproductiva....	62
<b>Tabla 11.</b> Heredabilidades, número de SNPs y genes asociados con las variables del patrón de tricomas.....	66
<b>Tabla 12.</b> Diversidad nucleotídica en el promotor y la región codificante de <i>EGL3</i> en la península ibérica.....	79
<b>Tabla 13.</b> Asociación genética entre polimorfismos en el promotor de <i>EGL3</i> y el número de tricomas en el fruto.....	81
<b>Tabla 14.</b> Patrón de tricomas de las líneas mutantes de <i>EGL3</i> .....	85
<b>Tabla 15.</b> Expresión del gen <i>TCL1</i> en líneas transgénicas.....	96
<b>Tabla 16.</b> Expresión del gen <i>TRY</i> en líneas transgénicas.....	105
<b>Tabla 17.</b> Expresión del gen <i>GL1</i> en líneas transgénicas.....	111
<b>Tabla 18.</b> Efecto de los loci <i>FTD</i> sobre la expresión de <i>TCL1</i> , <i>TRY</i> y <i>GL1</i> en fase reproductiva y vegetativa.....	114
<b>Tabla Suplementaria 1.</b> Genes implicados en el desarrollo de tricomas.....	153
<b>Tabla Suplementaria 2.</b> Resumen de los análisis de asociación genómica.....	157

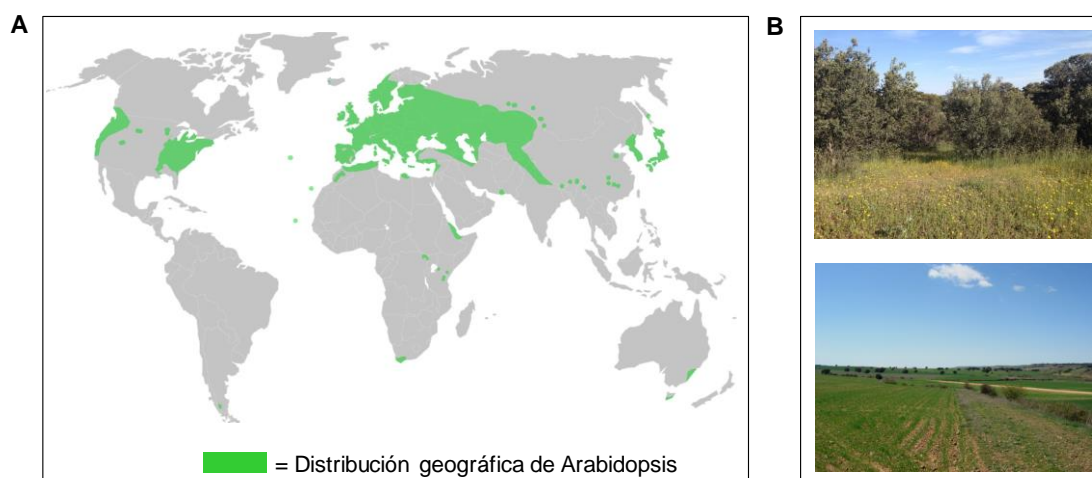
# INTRODUCCIÓN

## 1. *Arabidopsis* como especie modelo para el estudio del patrón de tricomas

Las plantas son organismos sésiles, por lo que su adaptación, definida como el ajuste entre el fenotipo y su ambiente local, es un proceso evolutivo fundamental para sobrevivir en el tiempo y a diferentes ambientes naturales (Orr *et al.*, 2005). En particular, el incipiente cambio climático global ha impulsado recientemente el interés por conocer los mecanismos genéticos, genómicos y moleculares de la adaptación de las plantas a distintos factores climáticos (Franks & Hoffmann, 2012; Savolainen *et al.*, 2013; Bigot *et al.*, 2018).

En las últimas dos décadas, el estudio de miles de mutantes artificiales generados en plantas modelo como *Arabidopsis thaliana* (a partir de ahora *Arabidopsis*) o arroz, han revelado los mecanismos moleculares de numerosos caracteres adaptativos (Shimamoto & Kyojuka, 2002; Koornneef & Meinke, 2010; Provar *et al.*, 2016; Li *et al.*, 2018a). Sin embargo, en la última década, *Arabidopsis* también se ha convertido en una especie modelo para estudios evolutivos y ecológicos debido principalmente a su enorme variación genética natural (Alonso-Blanco *et al.*, 2009; Weigel, 2012; Weigel & Nordborg, 2015; Krämer, 2015). *Arabidopsis* presenta una amplia distribución geográfica en el hemisferio norte (Figura 1A) donde se distribuye como especie nativa a través de Eurasia y África (Weigel, 2012; Brennan *et al.*, 2014; Lee *et al.*, 2017; Fulgione & Hancock, 2018) y como especie introducida recientemente en Australia, Norte América (Expósito-Alonso *et al.*, 2018) y Sudamérica (Kasulin *et al.*, 2017). Esta amplia distribución geográfica cubre una enorme diversidad climática y ecológica que incluye desde ambientes naturales (ej. bosques mediterráneos) a ambientes muy antrópicos (ej. campos agrícolas) (Figura 1B). Actualmente se dispone de más de 6.000 accesiones procedentes de distintas poblaciones naturales recogidas a través de todo su área de distribución. El estudio de estas colecciones de accesiones ha mostrado una gran diversidad genética para todo tipo de caracteres, la cual probablemente refleja la adaptación de *Arabidopsis* a dichos ambientes (Alonso-Blanco *et al.*, 2009; Weigel *et al.*, 2012). Además, la reciente secuenciación del genoma de más de mil accesiones llevada a cabo por el proyecto 1.001 genomas de *Arabidopsis* (The 1001 Genomes Consortium, 2016; Durvasula *et al.*, 2017), ha supuesto un hito para poder abordar el análisis de caracteres adaptativos a nivel genómico. Estos recursos están permitiendo conocer no sólo los mecanismos genéticos que determinan la variación natural, sino también su relevancia para la adaptación a distintos factores ambientales (Horton *et al.*, 2016; Dittberner *et al.*, 2018; Tabas-Madrid *et al.*, 2018; Wan *et al.*, 2018; Ferrero-Serrano & Assmann, 2019).

La variación natural de *Arabidopsis* se ha analizado extensivamente para un pequeño número de caracteres modelo implicados en la adaptación al clima, como son el tiempo de floración, la germinación y la dormición de las semillas (Méndez-Vigo *et al.*, 2011; 2016; Barua *et al.*, 2012; Kronholm *et al.*, 2012; Vidigal *et al.*, 2016; Bloomer & Dean, 2017; Tabas-Madrid *et al.*, 2018). Por el contrario, se desconocen casi por completo las bases genéticas y moleculares de la variación natural para numerosos caracteres cuantitativos más complejos que son esenciales para la adaptación de las plantas a distintos factores climáticos (Godfray *et al.*, 2010; Jamieson *et al.*, 2012; Mir *et al.*, 2012) como es el patrón de distribución y densidad de tricomas en los distintos órganos vegetales. Estudios realizados en diversas especies han mostrado que los tricomas están implicados en la adaptación de las plantas a diversos factores ambientales abióticos como son el exceso de luz, la radiación ultravioleta, las altas y bajas temperaturas, la sequía, los metales o el ozono (Hancock *et al.*, 2011; Swapna *et al.*, 2015; Ning *et al.*, 2016; Li *et al.*, 2018b). Igualmente, la presencia de tricomas proporciona una barrera de defensa contra factores bióticos como son los ataques de insectos herbívoros u hongos (Hauser *et al.*, 2014). Los recursos existentes actualmente en *Arabidopsis* permiten abordar los mecanismos genéticos, moleculares y evolutivos de la variación natural para este complejo carácter adaptativo escasamente estudiado.



**Figura 1. Variación natural de *Arabidopsis*.** A) Mapa de distribución global de *Arabidopsis*. B) Hábitats representativos de las poblaciones naturales de *Arabidopsis* en la península ibérica.

## 2. Regulación del patrón de tricomas en *Arabidopsis*

Los tricomas son células especializadas de la epidermis de las plantas, al igual que los pelos radiculares o las células guarda de los estomas (Esau, 1965). Los tricomas pueden ser unicelulares o pluricelulares, glandulares o no glandulares y simples o ramificados, de manera que sus características morfológicas se han utilizado a menudo

como caracteres taxonómicos para la identificación de especies.

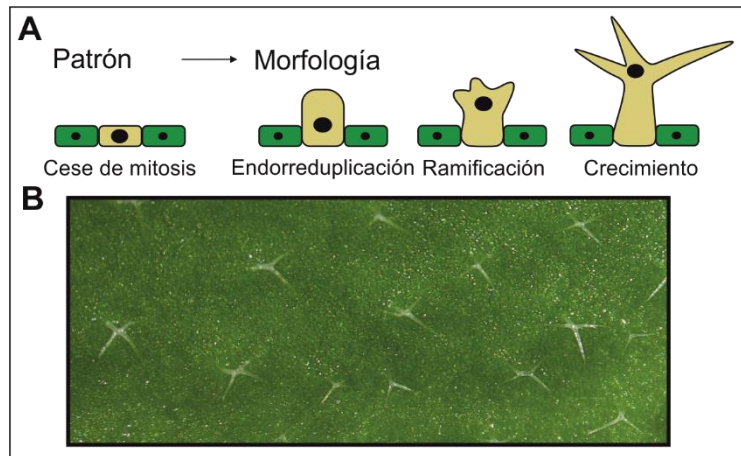
En particular, los tricomas de *Arabidopsis* son unicelulares y no glandulares y se localizan en la mayoría de las partes aéreas de la planta: hojas de la roseta, tallo, hojas caulinares y sépalos. En cuanto a su morfología, suelen ser ramificados en las hojas de la roseta mientras que en el tallo, las hojas caulinares y los sépalos no suelen tener ramificaciones (Hülskamp & Schnittger, 1998). Además, la distribución de los tricomas difiere en las distintas fases del desarrollo de *Arabidopsis*: en la fase vegetativa juvenil se localizan sólo la cara adaxial; en la fase vegetativa adulta se desarrollan en ambas caras de las hojas; y en la fase reproductiva se observa una reducción de la densidad de tricomas a lo largo del crecimiento del tallo, junto con una localización en ambas caras o sólo en la cara abaxial de las hojas caulinares (Telfer *et al.*, 1997).

### **2.1. Fases del desarrollo de los tricomas**

El desarrollo de los tricomas ha sido ampliamente estudiado como modelo de la expansión celular (Hülskamp *et al.*, 1999) y se divide en dos fases: el patrón de iniciación y la morfogénesis (Figura 2). La primera fase consiste en el cambio de identidad de la célula protodérmica para ser tricoma, el cual se lleva a cabo mediante la parada de su división celular y el comienzo de un primer ciclo de endorreducción, frente a las células adyacentes en las que la mitosis se sigue produciendo con regularidad (Hülskamp *et al.*, 1994). Durante el primer ciclo de endorreducción el tamaño del núcleo aumenta, lo que supone el primer cambio morfológico detectable en el desarrollo del tricoma (Figura 2A). Esta primera fase determina el patrón de distribución de los tricomas, definido como el patrón espacial y la densidad de los tricomas en los distintos órganos. El patrón de los tricomas a nivel celular no es aleatorio, sino que cada tricoma está inicialmente separado de los adyacentes por tres o cuatro células epidérmicas no especializadas, cuyo número incrementa posteriormente con las sucesivas divisiones celulares (Hülskamp *et al.*, 1994).

En la segunda fase del desarrollo del tricoma, continúan varios ciclos de endorreducción, con una media de cuatro ciclos, por lo que los tricomas poseen un contenido de ADN medio de 32C (Schnittger & Hülskamp, 2002). Durante la endorreducción la célula comienza a expandirse hacia exterior de la epidermis y posteriormente se forman varias protuberancias en la región más distal donde se inicia la formación de las ramas del tricoma. Por último tiene lugar la expansión de las ramas y la elongación del pie del tricoma (Figura 2A), procesos que ocurren a la vez que se

produce una rápida vacuolización de la célula (Hülkamp *et al.*, 1994; Schwab *et al.*, 2000).



**Figura 2. Fases del desarrollo de los tricomas en *Arabidopsis*.** A) Esquema de los cambios morfológicos que experimenta un tricoma en las fases de su desarrollo. B) Patrón y morfología de los tricomas de una hoja adulta de *Arabidopsis*.

## 2.2. Bases genéticas del desarrollo de los tricomas

Los mecanismos de regulación del desarrollo de los tricomas se han establecido principalmente mediante el análisis de mutantes en más de 100 genes de *Arabidopsis* (Tabla Suplementaria 1). Estos mecanismos son muy similares a los de la formación de los pelos radiculares, ya que comparten gran parte de las bases genéticas (Larkin *et al.*, 2003; Ishida *et al.*, 2008; Tominaga-Wada *et al.*, 2011; Grebe, 2012). Estos mutantes se han clasificado en cuatro clases atendiendo al proceso principal que tienen afectado y a los fenotipos de sus alelos de pérdida de función: 1) Mutantes del patrón de iniciación; 2) mutantes de la endorreduplicación; 3) mutantes de la ramificación; y 4) mutantes de la direccionalidad del crecimiento (Schwab *et al.*, 2000; Schellmann & Hülkamp, 2005).

La primera clase de mutantes ha permitido identificar los genes implicados en la regulación del patrón de tricomas (Tabla Suplementaria 1), a partir del cual se ha establecido un modelo molecular de dicho proceso que se detalla en el siguiente apartado (apartado 2.3). Brevemente, dentro de esta clase se distinguen reguladores positivos y reguladores negativos. Los mutantes de genes reguladores positivos como *GLABRA1* (*GL1*; Oppenheimer *et al.*, 1991) o *GLABRA3* (*GL3*; Koornneef *et al.*, 1982), muestran un fenotipo glabro o una reducción de la densidad de tricomas, respectivamente. Por el contrario, los mutantes de genes que actúan como reguladores negativos, como *CAPRICE* (*CPC*) o *TRIPTYCHON* (*TRY*), muestran un aumento del número de tricomas o una alteración de la distribución de tricomas, respectivamente (Schellmann *et al.*, 2002).

La segunda clase incluye los mutantes con alteraciones en los procesos de endorreduplicación (Tabla Suplementaria 1), donde uno de los genes más estudiados es *SIAMESE (SIM)* que regula la endorreduplicación mediante la supresión de la mitosis (Walker *et al.*, 2000). Puesto que el número de endorreduplicaciones correlaciona con el grado de ramificación, esta clase de mutantes ha permitido identificar genes implicados en la ramificación de los tricomas (Tabla Suplementaria 1). Así, mutaciones en genes como *KAKTUS (KAK)*, implican un contenido de ADN de 64C y provocan un incremento del número de ramificaciones. Por el contrario, mutantes del gen *GL3*, que también está implicado en el patrón de iniciación, muestran tricomas que no son ramificados (simples) y con un menor contenido de ADN (16C) (Hülkskamp *et al.*, 1994; Perazza *et al.*, 1999; Downes *et al.*, 2003). Igualmente, mutantes del gen *CONSTITUTIVE PATHOGEN RESPONSE 5 (CPR5)* que codifica para una proteína de muerte celular programada, presentan una reducción del número de endorreduplicaciones y, consecuentemente, un número menor de ramificaciones (Kirik *et al.*, 2001a). No obstante, también hay mutantes que muestran alteraciones de la ramificación de los tricomas sin modificaciones en la endorreduplicación. Estos constituyen la tercera clase de mutantes (Tabla Suplementaria 1) que presentan alteraciones de la organización intracelular de los microtúbulos, proceso que es necesario para la correcta formación de un citoesqueleto de tubulina durante la ramificación de los tricomas (Mathur *et al.*, 1999). Algunos de los genes implicados en este proceso son *KIESEL (KIS)* o *PORCINO (POR)* que codifican para los cofactores de tubulina A y C, respectivamente, y sus mutantes muestran tricomas sin ramificar (Kirik *et al.*, 2002a; 2002b).

La cuarta clase corresponde a los mutantes afectados en la direccionalidad o expansión del tricoma que ocurre tras iniciarse la ramificación (Tabla Suplementaria 1). Los mutantes más estudiados en este proceso son aquellos que muestran un fenotipo distorsionado y retorcido, lo que les da su nombre “*distorted*” (Szymanski *et al.*, 1999). Entre los genes implicados en este proceso destacan *DISTORTED1 (DIS1)*, *DIS2* y *DIS3*, que codifican componentes del complejo ARP2/3, el cual promueve la formación de filamentos de actina (Mullins *et al.*, 1998; Svitkina & Borisy, 1999), necesarios para la elongación de los tricomas (Mathur *et al.*, 2003a; El-Din El-Assal *et al.*, 2004; Basu *et al.*, 2005).



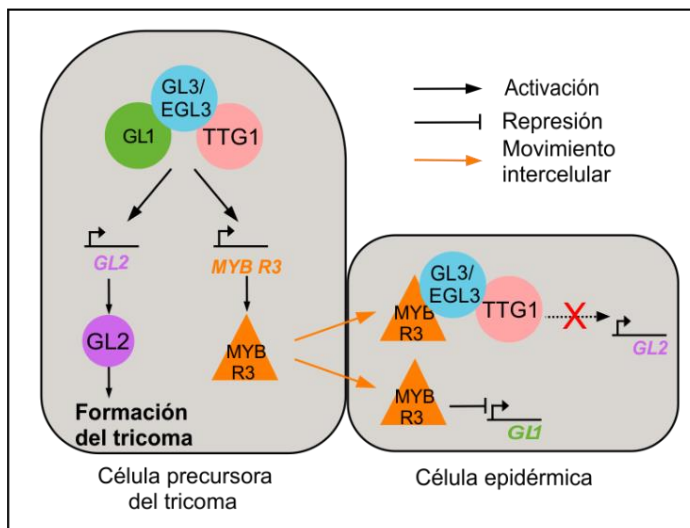
### 2.3. Modelo de regulación del patrón de tricomas en *Arabidopsis*

Como se ha descrito en el apartado anterior, el patrón de tricomas está regulado principalmente por numerosos genes que actúan como reguladores positivos (promotores) o negativos (represores) y que codifican principalmente para factores de transcripción. Además, en esta etapa participan otros reguladores, incluyendo microARNs (miARN) y hormonas vegetales.

Los promotores del patrón de tricomas son fundamentalmente factores de transcripción de tres tipos diferentes. En primer lugar, se ha descrito la función promotora de varios factores de transcripción de tipo MYB R2R3, destacando *GL1*. Otros factores de transcripción MYB R2R3 cuya implicación en este proceso ha sido demostrada son *MYB23* (Kirik *et al.*, 2001b) y *MYB82* (Liang *et al.*, 2014) que parecen llevar a cabo un papel redundante con *GL1* sobre el patrón de tricomas, aunque *MYB23* también está implicado en fases posteriores (Kirik *et al.*, 2005). En segundo lugar, varios promotores pertenecen a la familia de factores transcripcionales bHLH, entre los que se encuentran dos proteínas homólogas, *GL3* y ENHANCER OF *GLABRA3* (*EGL3*), que pueden formar homodímeros y heterodímeros (Payne *et al.*, 2000; Zhang *et al.*, 2003); *MYC1* que presenta un papel redundante o divergente con *GL3* y *EGL3* dependiendo del órgano (Zhao *et al.*, 2012); y *TRANSPARENT TESTA 8* (*TT8*) que controla el desarrollo de tricomas en los márgenes de las hojas (Maes *et al.*, 2008) (Figura 3). Por último, *TRANSPARENT TESTA GLABRA1* (*TTG1*) codifica para un factor de transcripción de tipo WD40 que afecta no sólo al desarrollo de tricomas sino también a la biosíntesis de antocianinas (Walker *et al.*, 1999). Además de estos factores de transcripción, recientemente se ha demostrado la implicación de *COP9 SIGNALOSOME 5A* (*CSN5A*) en el desarrollo de tricomas, probablemente mediante la regulación transcripcional de varios de estos reguladores positivos (Wei *et al.*, 2018).

Las proteínas *GL1*, *GL3/EGL3* y *TTG1* interaccionan físicamente para formar el llamado complejo activador (*GL1-GL3/EGL3-TTG1*) (Figura 3). Este complejo se forma mediante la unión de *GL1* y *TTG1* a la región N-terminal y región media (aminoácidos 212-401) de la proteína *GL3*, respectivamente. Además, la región C-terminal de *GL3* permite la formación de homodímeros y heterodímeros con *EGL3* (*GL3/EGL3*) (Payne *et al.*, 2000; Zhang *et al.*, 2003). El complejo activador promueve la expresión del gen *GLABRA2* (*GL2*; Rerie *et al.*, 1994) que codifica para un factor de transcripción con un homeodominio que es necesario para la morfogénesis del tricoma (Figura 3).

**Figura 3. Modelo de la regulación del patrón de tricomas en Arabidopsis.** El esquema representa la función de los principales genes y proteínas implicados en la regulación del inicio del tricoma en la célula precursora y en las células adyacentes. Por simplificación, los siete factores de transcripción MYB R3 se han representado de forma conjunta.



Por otra parte se distinguen los siguientes siete reguladores negativos del patrón de tricomas en Arabidopsis, todos ellos factores de transcripción de tipo MYB R3: *TRY* (Schellmann *et al.*, 2002), *CPC* (Wada *et al.*, 1997), *ENHANCER OF TRY AND CPC1*, (*ETC1*), *ETC2* y *ETC3* (Esch *et al.*, 2004; Kirik *et al.*, 2004a; Kirik *et al.*, 2004b; Wester *et al.*, 2009), *TRICHOMELESS1* (*TCL1*; Wang *et al.*, 2007) y *TCL2* (Gan *et al.*, 2011). Estos factores transcripcionales se caracterizan por presentar conservado un motivo WxM, el cual les permite el movimiento desde la célula precursora del tricoma a las células vecinas (Kurata *et al.*, 2005). El complejo activador GL1-GL3/EGL3-TTG1 no sólo promueve la expresión de *GL2*, sino que también induce la expresión de varios represores MYB R3, como son *TRY*, *CPC*, *ETC1* y *ETC3* (Wang *et al.*, 2008) (Figura 3). No obstante, actualmente no hay evidencias de que *TCL1*, *TCL2* y *ETC2* estén regulados por este complejo. A su vez, estos represores se desplazan a las células adyacentes para reprimir en ellas la formación de tricomas (Kurata *et al.*, 2005; Wester *et al.*, 2009) (Figura 3). Además, aunque la mayoría de los estudios del patrón de tricomas se han realizado en las hojas, varios de estos represores también se han implicado en la regulación del patrón de tricomas en fase reproductiva. En concreto, análisis fenotípicos de mutantes simples o múltiples han demostrado la contribución de *TCL1*, *CPC*, *ETC1* y *ETC3* a la reducción progresiva del desarrollo de tricomas a lo largo del tallo y los pedicelos de los frutos (Wang *et al.*, 2007; 2008).

Estos reguladores negativos llevan a cabo su función represora a través de dos mecanismos diferentes (Figura 3). En primer lugar, ciertos represores MYB R3 inhiben la formación del complejo activador (GL1-GL3/EGL3-TTG1) ya que las proteínas MYB R3 compiten con GL1 por unirse a GL3/EGL3 (Bernhardt *et al.*, 2003; Esch *et al.*, 2003; Payne *et al.*, 2000). El segundo mecanismo consiste en la represión directa de la

expresión de *GL1* por parte de los reguladores negativos *TCL1* y *TCL2* (Wang *et al.*, 2007; Gan *et al.*, 2011).

Además de estos reguladores negativos se han descrito otros represores del desarrollo de los tricomas de *Arabidopsis*, aunque su función se ha descrito principalmente en el tallo y el pedicelo. Cabe destacar SQUAMOSA PROMOTER BINDING PROTEIN LIKE 9 (SPL9) que regula de forma directa la expresión de *TCL1* y *TRY*, y cuyo gen es diana del microARN miR156 (Yu *et al.*, 2010). Las proteínas SPLs están reguladas negativamente por los factores de transcripción LOST MERISTEMS (LOM) cuyos genes también son regulados por otros miARNs (miR171) que afectan al desarrollo de tricomas (Xue *et al.*, 2014). Por último, recientemente se ha descrito el factor de transcripción de la familia NAC llamado NTM1-LIKE8 (NTL8) como un activador transcripcional de los reguladores negativos *TRY* y *TCL1* (Tian *et al.*, 2017). Aunque SPL9 y NTL8 llevan a cabo funciones similares, ambas proteínas no interactúan entre sí, por lo que se ha propuesto que participan de forma paralela en la regulación del desarrollo de tricomas (Tian *et al.*, 2017).

#### **2.4. Efecto de las hormonas y otros factores sobre la regulación del patrón de tricomas**

Varias hormonas vegetales tienen un papel muy importante en la regulación del patrón de tricomas, incluidas las giberelinas (GAs), las citoquininas (CKs), el ácido jasmónico (AJ), y en menor medida el ácido salicílico (Traw & Bergelson, 2003) (Tabla Suplementaria 1).

Las giberelinas son reguladores positivos de la formación de tricomas ya que su aplicación exógena aumenta la expresión de *GL1/MYB23*, *GL3* y *EGL3*, a la vez que reduce la expresión de los reguladores negativos *TRY*, *ETC1* y *ETC2* (Perazza *et al.*, 1998; Maes *et al.*, 2008). Además, se ha descrito que proteínas DELLA como GIBBERELLIC ACID INSENSITIVE (GAI) o REPRESSOR OF *ga1-3* (RGA) que son represoras de la señalización de las giberelinas, interactúan físicamente con *GL3*, *EGL3* y *GL1*, reprimiendo así la función transcripcional del complejo activador. En presencia de GAs las proteínas DELLA son degradadas, dejando libre el complejo activador (Qi *et al.*, 2014; Tian *et al.*, 2016). Además, recientemente se ha descrito que las proteínas TEMPRANILLO1 (TEM1) y TEM2 reprimen la iniciación de la formación de tricomas de dos formas: inhibiendo la expresión de genes como *GL1*, *GL3*, *EGL3* y *GL2*; y reprimiendo la biosíntesis y distribución de GAs (Matías-Hernández *et al.*, 2016).

Por otro lado, las CKs, junto con las GAs, activan la expresión de varios genes codificantes de factores de transcripción de tipo C2H2 que están implicados en la formación de tricomas en fase reproductiva, incluyendo *GLABROUS INFLORESCENS STEM (GIS)* (Gan *et al.*, 2006), *GIS2* (Gan *et al.*, 2007), *GIS3* (Sun *et al.*, 2015), *ZINC FINGER PROTEIN8 (ZFP8)* (Gan *et al.*, 2007), *ZFP5* (Zhou *et al.*, 2011) y *ZFP6* (Zhou *et al.*, 2013). Se ha demostrado que los genes *GIS*, *GIS2* y *ZFP8* inducen la formación de tricomas en el tallo, las hojas caulinares y las inflorescencias, mientras que *GIS3* y *ZFP5* inducen la expresión de *GIS*, *GIS2* y *ZFP8* (Zhou *et al.*, 2011; Sun *et al.*, 2015). Además, recientemente se ha propuesto que *TCL1* podría actuar por encima de *GIS* en la regulación de la formación de tricomas en la fase reproductiva (Zhang *et al.*, 2018).

El ácido jasmónico regula varias respuestas de la planta a las heridas o el estrés biótico, incluido el desarrollo de tricomas. En concreto, la síntesis de AJ se induce por herida (Koo & Howe, 2009) y la simulación mecánica de heridas, o la aplicación exógena de AJ, inducen la formación de tricomas (Traw & Bergelson, 2003). Varios mecanismos contribuyen a esta regulación del desarrollo de tricomas. Por un lado, análisis de expresión tras la aplicación exógena sugieren que el AJ regula positivamente la expresión de los genes *GL3* y *GL1* (Qi *et al.*, 2011). Además, se ha demostrado que *GL3*, *EGL3* y *GL1* interaccionan con las proteínas JASMONATE-ZIM (JAZ), reguladores negativos de la señalización del AJ, bloqueando así la interacción de *GL1-GL3/EGL3* con *TTG1* e inhibiendo la formación del complejo activador. Las proteínas JAZ son degradadas en presencia de AJ, de manera que *TTG1* puede entonces unirse al resto del complejo activador para promover el desarrollo de tricomas (Qi *et al.*, 2011; Tian *et al.*, 2016).

Por último, cabe destacar varios genes implicados en la regulación del tiempo de floración y de la transición entre las fases vegetativas, y que también afectan al patrón de tricomas, como son *FLOWERING LOCUS T (FT)*, *SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1)* o *FLOWERING LOCUS C (FLC)* (Willmann & Poething, 2011). No obstante, el efecto de estos genes sobre el patrón de tricomas es indirecto ya que viene determinado, principalmente, por su efecto primario sobre la duración de las fases vegetativas juvenil y adulta, que difieren cuantitativamente en su patrón de tricomas.

### 3. Variación natural para el patrón de tricomas en Arabidopsis

Como se ha descrito anteriormente, con el fin de determinar las bases genéticas, y moleculares de la adaptación de las plantas a distintos ambientes, en los últimos años se ha iniciado el análisis de la variación natural para el patrón de tricomas en Arabidopsis (Hauser *et al.*, 2001; Handley *et al.*, 2005; Mauricio, 2005; Symonds *et al.*, 2005; Hilscher *et al.*, 2009; Traw & Bergelson, 2003). Estos estudios han mostrado que hay una enorme diversidad genética para este carácter en Arabidopsis, la cual se está diseccionando a través de análisis genéticos cualitativos (monogénicos), cuantitativos y genómicos. De esta forma se han identificado los primeros loci y genes responsables de la variación natural para el patrón de tricomas (Figura 4).

#### 3.1. Análisis genéticos de la variación cualitativa y cuantitativa en Arabidopsis

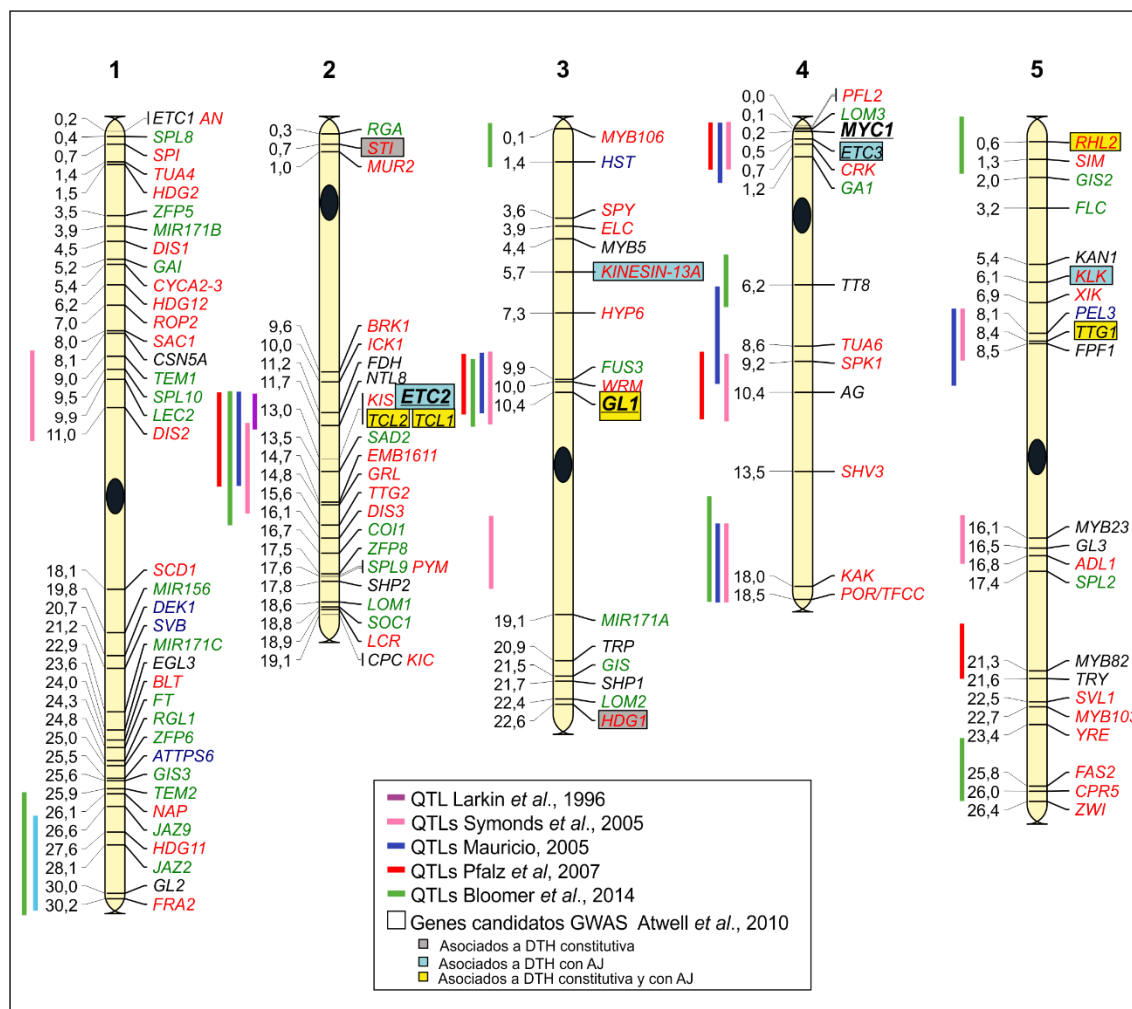
La caracterización fenotípica de las colecciones globales de variedades naturales de Arabidopsis ha identificado más de 30 accesiones glabras, cuyo análisis genético ha demostrado que todas ellas están producidas por el gen *GL1* (Hauser *et al.*, 2001; Bloomer *et al.*, 2012). De esta forma se han identificado numerosas mutaciones naturales independientes del dominio MYB de *GL1* que causan su pérdida de función. Sin embargo, la mayor parte de la variación natural para el patrón de tricomas de Arabidopsis es de naturaleza cuantitativa, por lo que viene determinada por numerosos genes y sus interacciones, así como por los efectos ambientales (Koorneef *et al.*, 2004).

El análisis genético de los caracteres cuantitativos clásicamente se ha realizado mediante el mapeo de loci de carácter cuantitativo (QTLs; "*Quantitative Trait Locus*"). Estos análisis tienen como objetivo determinar el número, la posición genética y el efecto de los QTLs o regiones genómicas que contribuyen a dicha variación cuantitativa. Para ello, básicamente se analizan genotípicamente y fenotípicamente poblaciones segregantes obtenidas experimentalmente a partir de accesiones que difieren en el carácter cuantitativo estudiado (Koorneef *et al.*, 2004). El principal tipo de poblaciones utilizadas en Arabidopsis son las líneas recombinantes consanguíneas (RILs; "*Recombinant Inbred Lines*") que se obtienen por autofecundación de una población F<sub>2</sub> durante 6-8 generaciones (Alonso-Blanco *et al.*, 2006). Debido a esta autofecundación consecutiva, las RILs son genotipos prácticamente homocigóticos, por lo que se pueden analizar para el mismo o distintos caracteres numerosas veces, permitiendo mapeos de QTLs

comparativos. La comparación de los fenotipos y los genotipos de las RILs para cientos de marcadores a través del genoma permite identificar los correspondientes QTLs.

En *Arabidopsis* se ha analizado el patrón de tricomas de las hojas de siete poblaciones de RILs diferentes (*Ler-0* x *Col-4*; *Ler-0* x *Cvi*; *Bay* x *Sha*; *Ler-2* x *No-0*; *Da-12* x *Ei-2*; *Hi-0* x *Ob-0*; *St-0* x *Sf-2*) (Larkin *et al.*, 1996; Mauricio, 2005; Symonds *et al.*, 2005; Pfalz *et al.*, 2007; Bloomer *et al.*, 2014). Estos estudios han abordado no sólo el patrón de tricomas en la hoja sino también la variación natural entre las fases vegetativas juvenil y adulta (Mauricio, 2005) y los efectos de la herbivoría sobre el patrón de tricomas (Pfalz *et al.*, 2007; Bloomer *et al.*, 2014). De esta forma se ha identificado un total de 14 regiones genómicas que contienen QTLs que contribuyen a la variación natural para el patrón de tricomas (Figura 4). La mayoría de estos análisis han mapeado QTLs en una región del cromosoma 2 de *Arabidopsis* que contiene los genes *ETC2*, *TCL1*, *TCL2*, *TTG2* y *SAD2*. Además, varios QTLs solapan con otros genes conocidos incluyendo *GL1*, *GL2*, *TTG1* y *GL3*.

Puesto que los mapeos de QTLs de *Arabidopsis* incluyen al menos 1.000 genes (Koorneef *et al.*, 2004), una vez localizado un QTL, su aislamiento molecular, es decir, la identificación del gen causal subyacente, requiere en primer lugar llevar a cabo su mapeo fino. Para ello se utilizan líneas genéticas que segregan sólo para la región genómica del QTL, denominadas líneas cuasi-isogénicas, las cuales tendrán una segregación monogénica. Finalmente, la demostración de un gen candidato y del polimorfismo nucleotídico que afecta a la función del mismo y causa la variación fenotípica (denominado polimorfismo funcional) implica la combinación de diversas estrategias funcionales, incluyendo el análisis de mutantes artificiales, estudios de expresión génica y experimentos de complementación con plantas transgénicas (Koorneef *et al.*, 2004; Alonso-Blanco *et al.*, 2005). Hasta la fecha sólo tres QTLs implicados en el patrón de tricomas han sido aislados en *Arabidopsis* (Hilscher *et al.*, 2009; Symonds *et al.*, 2011; Bloomer *et al.*, 2012). Uno de ellos ha identificado el gen regulador negativo *ETC2*, demostrando que el cambio de aminoácido Lys/Glu<sup>19</sup> es el polimorfismo funcional (Hilscher *et al.*, 2009). El segundo locus corresponde al regulador positivo *MYC1* que codifica para otro factor de transcripción en el que un cambio de aminoácido se ha asociado con las diferencias en el patrón de tricomas (Symond *et al.*, 2011). Por último, varios cambios de aminoácido del gen *GL1* se han asociado también con la variación cuantitativa para el patrón de tricomas, aunque la contribución de estas variantes alélicas no se ha demostrado de forma conclusiva (Bloomer *et al.*, 2012).



**Figura 4. Genes y QTLs implicados en la variación natural para el patrón de tricomas en Arabidopsis.** La figura muestra los mapas físicos de los cinco cromosomas de Arabidopsis con la posición de todos los genes (Mb) cuyos mutantes afectan a la regulación del patrón de distribución de los tricomas en la hoja (negro), la morfología de los tricomas (rojo), la respuesta del patrón de tricomas a hormonas u otros factores (verde), u otros genes cuya función sobre el patrón de tricomas no está caracterizada (azul). A la izquierda de cada cromosoma se representan los intervalos de confianza de los QTLs mapeados para la densidad de tricomas en la hoja en distintos estudios (Larkin *et al.*, 1996; Symonds *et al.*, 2005; Mauricio, 2005; Pfalz *et al.*, 2007; Bloomer *et al.*, 2014). Los genes candidatos identificados mediante análisis de asociación genómica (GWAS) por Atwell *et al.* (2010) se muestran enmarcados. Los genes cuya contribución a la variación natural del patrón de tricomas ha sido demostrada se muestran en negrita y subrayados.

### 3.2. Análisis genómicos de la variación del patrón de tricomas

La secuenciación completa del genoma de 1.200 accesiones naturales de Arabidopsis procedentes de todo su área de distribución geográfica ha posibilitado el estudio de las bases genéticas de la variación natural a través de análisis de asociación genómica (GWAS; “Genome-Wide Assosiation Studies”) (The 1001 Genomes Consortium, 2016). Para ello, se analiza fenotípicamente una colección de accesiones naturales para el carácter cuantitativo de interés y estos datos se comparan estadísticamente con los

genotipos de millones de polimorfismos de un único nucleótido (SNPs; “*Single Nucleotide Polymorphisms*”) segregantes a lo largo del genoma. Esta comparación permite la identificación de regiones genómicas de pequeño tamaño (1-100 genes) asociadas a la variación natural del carácter cuantitativo. A diferencia del mapeo de QTLs en poblaciones segregantes experimentales como las RILs, el uso de colecciones de accesiones naturales explota los eventos de recombinación acumulados en las accesiones de forma natural a lo largo de miles de años y generaciones (Bergelson & Roux, 2010; Bazakos *et al.*, 2017). Esta elevada recombinación reduce la extensión del desequilibrio de ligamiento entre marcadores próximos y permite aumentar la precisión del mapeo genético. No obstante, hay dos limitaciones que reducen el potencial de los análisis de asociación genómica. Por un lado, la existencia de numerosos alelos naturales independientes y de baja frecuencia en genes que contribuyen a la variación natural, dificulta la asociación de polimorfismos en estos genes. Por otra parte, las complejas relaciones genéticas entre las accesiones determinan la presencia de una fuerte estructura genética en *Arabidopsis*, es decir, de grupos de accesiones que comparten parte del genoma. Esta estructura genética produce un exceso de SNPs que aparecen asociados con el carácter cuantitativo debido a las relaciones genéticas entre las accesiones y que son falsos positivos. Para reducir la primera limitación se ha propuesto utilizar colecciones de accesiones a escala geográfica regional en lugar de global, de forma que se reduzca el número de alelos en cada gen y se aumente la frecuencia de los mismos (Horton *et al.*, 2012; Brachi *et al.*, 2013; Tabas-Madrid *et al.*, 2018). Para reducir el efecto de la estructura genética de las poblaciones de *Arabidopsis* se han desarrollado distintos modelos estadísticos que integran esta estructura en la detección de asociaciones SNP-fenotipo (Bergelson & Roux, 2010; François *et al.*, 2016; Klasen *et al.*, 2016; Bazakos *et al.*, 2017).

Hasta la fecha, sólo un estudio ha utilizado el análisis de asociación genómica para determinar las bases genéticas de la variación natural para el patrón de tricomas en *Arabidopsis* (Atwell *et al.*, 2010). Este estudio se ha basado en una pequeña colección global de 95 accesiones cuyos genomas no se han secuenciado completamente, sino que se han genotipado para 250.000 SNPs. Este análisis de asociación genómica para la densidad de tricomas, en plantas tratadas con AJ y sin tratar (Figura 4), detectó 13 genes candidatos que podrían contribuir a la variación natural del patrón de tricomas. Cinco de estos genes se han implicado previamente en la regulación del desarrollo de tricomas y aparecen asociados con ambas variables analizadas (*TCL1*, *TCL2*, *GL1*, *TTG1* y *ROOT HAIRLESS 2 -RHL2*; Sugimoto-Shirasu *et al.*, 2002-). Otros seis genes también corresponden a genes conocidos que afectan



al desarrollo de tricomas pero que sólo se detectaron asociados con el patrón de tricomas con o sin AJ (*ETC2*, *ETC3*, *KINESIN-13A* -Lu *et al.*, 2005-, *STICHEL* -*STI*; Ilgenfritz *et al.*, 2003-, *KLUNKER* -*KLK*; Saedler *et al.*, 2004- y *HOMEODOMAIN GLABROUS 1* -*HDG1*; Wu *et al.*, 2011-), lo que sugiere que podrían estar implicados en la variación natural para la respuesta a esta hormona y a la herbivoría. Por último, se identificaron *ACCELERATED CELL DEATH 6* (*ACD6*) y *FLOWERING WAGENINGEN* (*FWA*) como dos genes que no se han implicado previamente en este carácter.

## **4. Análisis regional de la variación natural en la península ibérica**

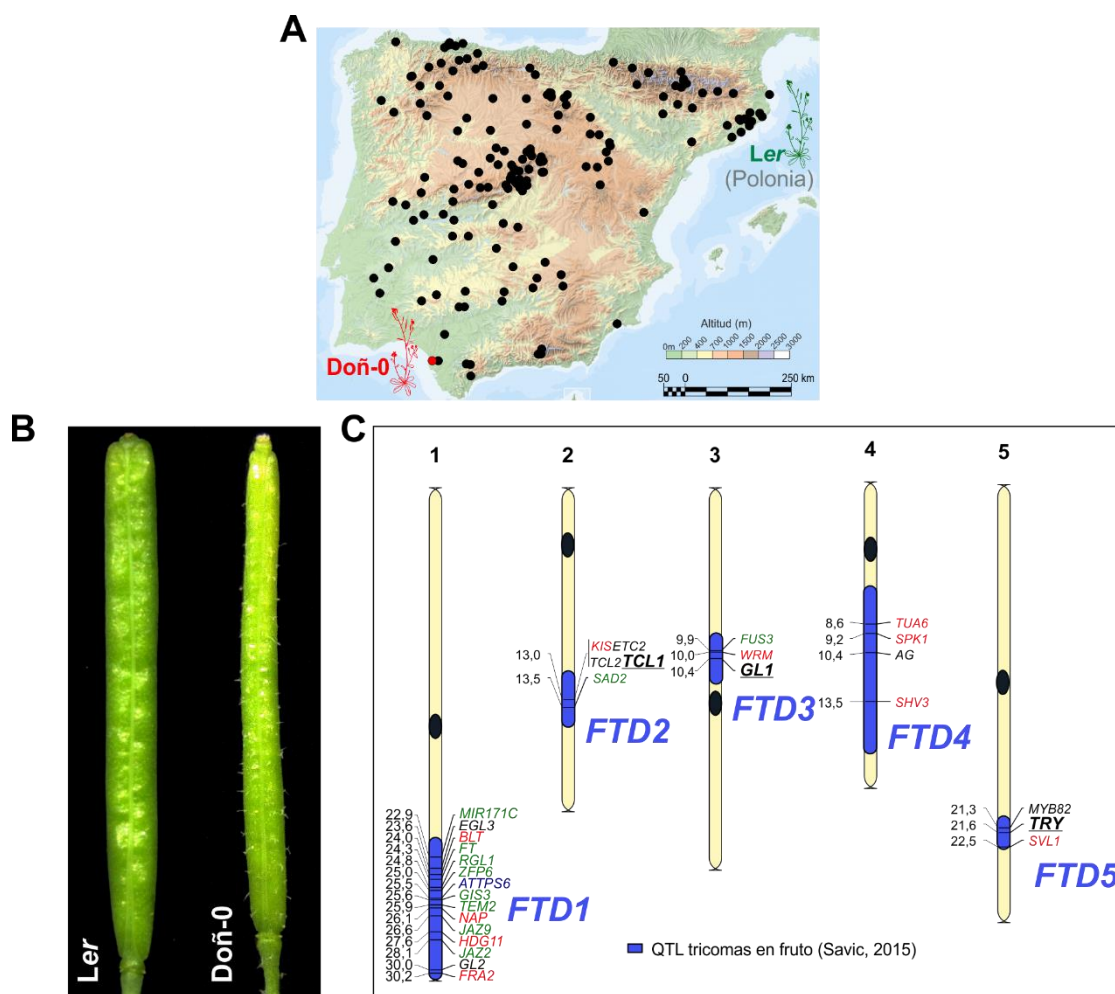
### **4.1. La península ibérica como escenario ideal para el estudio de los mecanismos genéticos y moleculares de la adaptación de *Arabidopsis***

La mayor parte de los estudios de asociación genómica realizados hasta la fecha se han basado en colecciones globales que incluyen accesiones de todo el mundo. Sin embargo, debido a la enorme diversidad genética de *Arabidopsis*, las colecciones globales impiden la identificación de genes con múltiples alelos de baja frecuencia (diversidad alélica). Análisis geográficos de la diversidad genética han mostrado la existencia de una fuerte estructura geográfica (Cao *et al.*, 2011; Horton *et al.*, 2012), de manera que la mayor parte de los alelos naturales se distribuyen principalmente en áreas regionales menores de 600 km (Tabas-Madrid *et al.*, 2018; Ferrero-Serrano & Asmann, 2019). Por ello, se han desarrollado varias colecciones regionales, como las de Suecia y Francia (Brachi *et al.*, 2013; Long *et al.*, 2013), con las que aumentar la frecuencia de los alelos raros y, consecuentemente, su probabilidad de detección en los análisis de asociación genómica. En particular, en nuestro laboratorio se ha desarrollado una colección de la península ibérica con más de 300 accesiones caracterizadas en detalle a nivel genético, ambiental y genómico (Picó *et al.*, 2008; Cao *et al.*, 2011; Manzano-Piedras *et al.*, 2014). Estos análisis han mostrado que la península ibérica es la región de Eurasia con mayor diversidad genética y fenotípica (Cao *et al.*, 2011; Méndez-Vigo *et al.*, 2011) y cubre una amplia diversidad altitudinal, climática y ecológica (Manzano-Piedras *et al.*, 2014). Además, el genoma de 174 accesiones de esta colección se ha secuenciado recientemente como parte del proyecto 1.001 genomas de *Arabidopsis* (The 1001 Genomes Consortium, 2016; Tabas-Madrid *et al.*, 2018) (Figura 5A). En definitiva, la península ibérica proporciona un escenario ideal para estudiar las bases genéticas, genómicas y moleculares de la adaptación de *Arabidopsis* a distintos ambientes (Tabas-Madrid *et al.*, 2018).

## 4.2. Diversidad genética para el patrón de tricomas de *Arabidopsis* en la península ibérica

Estudios previos de la colección regional de la península ibérica realizados en nuestro laboratorio han identificado una gran cantidad de variación para de la densidad de tricomas en las hojas (Savic, 2015). Además, se han identificado ~15 accesiones ibéricas que presentan tricomas en los frutos, un carácter que no se ha descrito previamente en *Arabidopsis* (Figura 5). Con el fin de determinar las bases genéticas de este carácter, en el laboratorio se llevó a cabo un mapeo de QTLs en una población de RILs derivada del cruzamiento entre una accesión con tricomas en el fruto recogida en el Parque Nacional de Doñana (Doñ-0) y la cepa referencia *Ler* (Figura 5). Este análisis identificó cinco QTLs a los que se ha denominado *Fruit Tricome Density 1 (FTD1)* a *FTD5*, que contribuyen a la variación para el patrón de tricomas en el fruto entre *Ler* y Doñ-0 (Figura 5C). En conjunto, los cinco QTLs explican el 45,1% de la variación fenotípica, siendo *FTD5* el locus de mayor efecto (20%), seguido de *FTD2* y *FTD3* (10-15%), y *FTD4* y *FTD1* que tienen menor efecto (<5%) (Savic, 2015). Como cabría esperar, el alelo de Doñ-0 incrementa el número de tricomas del fruto en los cinco QTLs. La comparación de los loci *FTD* con los QTLs para el patrón de tricomas en las hojas detectados en otros estudios (Figura 4) muestra que todos los loci excepto *FTD5* solapan con regiones genómicas previamente identificadas. Estos resultados sugieren que *FTD5* podría estar implicado específicamente en la variación natural para el patrón de tricomas en el fruto, mientras que los loci *FTD1-4* podrían contribuir a la variación observada tanto en las hojas como en los frutos.

Con el fin de identificar genes candidatos para los tres QTLs de efecto mayor, *FTD2*, *FTD3* y *FTD5*, se buscaron los genes conocidos de la regulación del desarrollo de tricomas que solapan con el intervalo de confianza de cada locus. De esta forma se encontró que los tres genes *ETC2*, *TCL1* y *TCL2*, que están estrechamente ligados en una región de 9,5 kb del cromosoma 2, solapan con *FTD2*. Posteriormente, *TCL1* fue seleccionado como el mejor candidato para *FTD2* en base a análisis de mutantes artificiales, estudios de expresión génica y comparación de las secuencias genómicas (Savic, 2015). El análisis de *FTD3* identificó *GL1* como el único gen candidato localizado dentro de su intervalo de confianza. Por último, el mapeo fino de *FTD5* permitió acotar este locus a una región genómica de menos de 10 kb que sólo contiene el gen regulador del patrón de tricomas *TRY* (Figura 5C) (Savic, 2015).



**Figura 5. Variación natural para el patrón de tricomas en el fruto de Arabidopsis en la península ibérica.** A) Distribución geográfica de las 174 accesiones de la colección regional de la península ibérica, cuyos genomas han sido secuenciados. En color rojo se indica la localización de la accesión parental Doñ-0 con tricomas en el fruto y utilizada para el mapeo de QTLs. B) Frutos de las accesiones Ler y Doñ-0, mostrando los tricomas en Doñ-0. C) Mapeo de QTLs para la densidad de tricomas en los frutos en una población de RILs Ler/Doñ-0 (Savic, 2015). En color azul se muestra el intervalo de confianza de los cinco QTLs, *FTD1* a *FTD5*, identificados. Sobre los cromosomas también se muestran los genes conocidos del desarrollo de tricomas que solapan con los intervalos de mapeo de los QTLs. Estos genes se presentan en distinto color en función del proceso del desarrollo de tricomas en que participan: el patrón espacial de los tricomas (negro), la morfología de los tricomas (rojo), la respuesta del patrón a hormonas u otros factores (verde), o con función indefinida (azul). Los genes candidatos seleccionados para *FTD2*, *FTD3* y *FTD5* se muestran subrayados y en negra.

Finalmente, las secuencias nucleotídicas de *TCL1*, *GL1* y *TRY* fueron analizadas en las cepas parentales Ler y Doñ-0 para identificar posibles polimorfismos funcionales de estos genes. A partir de estos polimorfismos se realizó un estudio de asociación con la densidad de tricomas del fruto usando una colección de 30 accesiones segregantes para este carácter. Este estudio permitió seleccionar una delección de la región 3' no traducida (UTR) de *TCL1*, un codón de parada prematuro de *GL1* presente sólo en Doñ-0, y un SNP intrónico en posición 701 de *TRY*, como polimorfismos funcionales candidatos para los loci *FTD2*, *FTD3* y *FTD5*, respectivamente (Savic, 2015).

# OBJETIVOS

El objetivo general de esta tesis es determinar las bases genéticas, genómicas y moleculares de la variación natural para el patrón de tricomas en *Arabidopsis*, tanto en fase vegetativa como reproductiva. Para ello se ha utilizado la colección de accesiones de la península ibérica cuyos genomas están disponibles como parte del proyecto 1.001 genomas de *Arabidopsis* (The 1001 Genomes Consortium, 2016; Tabas-Madrid *et al.*, 2018).

En particular, esta tesis se propone estudiar dos cuestiones generales relacionadas con las bases genéticas de la variación natural para el patrón de tricomas. Por un lado, se pretende determinar si la variación del patrón de tricomas en distintos órganos de la fase vegetativa y reproductiva de la planta está determinada por los mismos o distintos genes. Además, se propone cuantificar la variación natural para la respuesta del patrón de tricomas de las hojas al ácido jasmónico, como un marcador de herbivoría, e identificar las bases genéticas subyacentes. Por otra parte, se pretende averiguar los mecanismos moleculares de la variación natural para el patrón de tricomas en los frutos de *Arabidopsis*. Estas preguntas se han abordado utilizando una combinación de estrategias genéticas y moleculares, incluyendo el análisis de asociación genómica, el mapeo genético de alta precisión, y el análisis fenotípico y molecular de plantas transgénicas.

Para alcanzar estos objetivos generales, en esta tesis se han propuesto los siguientes objetivos específicos:

1. Identificación de genes candidatos para la variación natural del patrón de tricomas en distintas fases del desarrollo y su respuesta al ácido jasmónico, mediante un análisis de asociación genómica a escala regional.
2. Caracterización genética y molecular del locus *FTD1* implicado en el patrón de tricomas del fruto.
3. Aislamiento y caracterización de los genes correspondientes a *FTD2*, *FTD3* y *FTD5* del patrón de tricomas en fruto.

# **MATERIALES Y MÉTODOS**

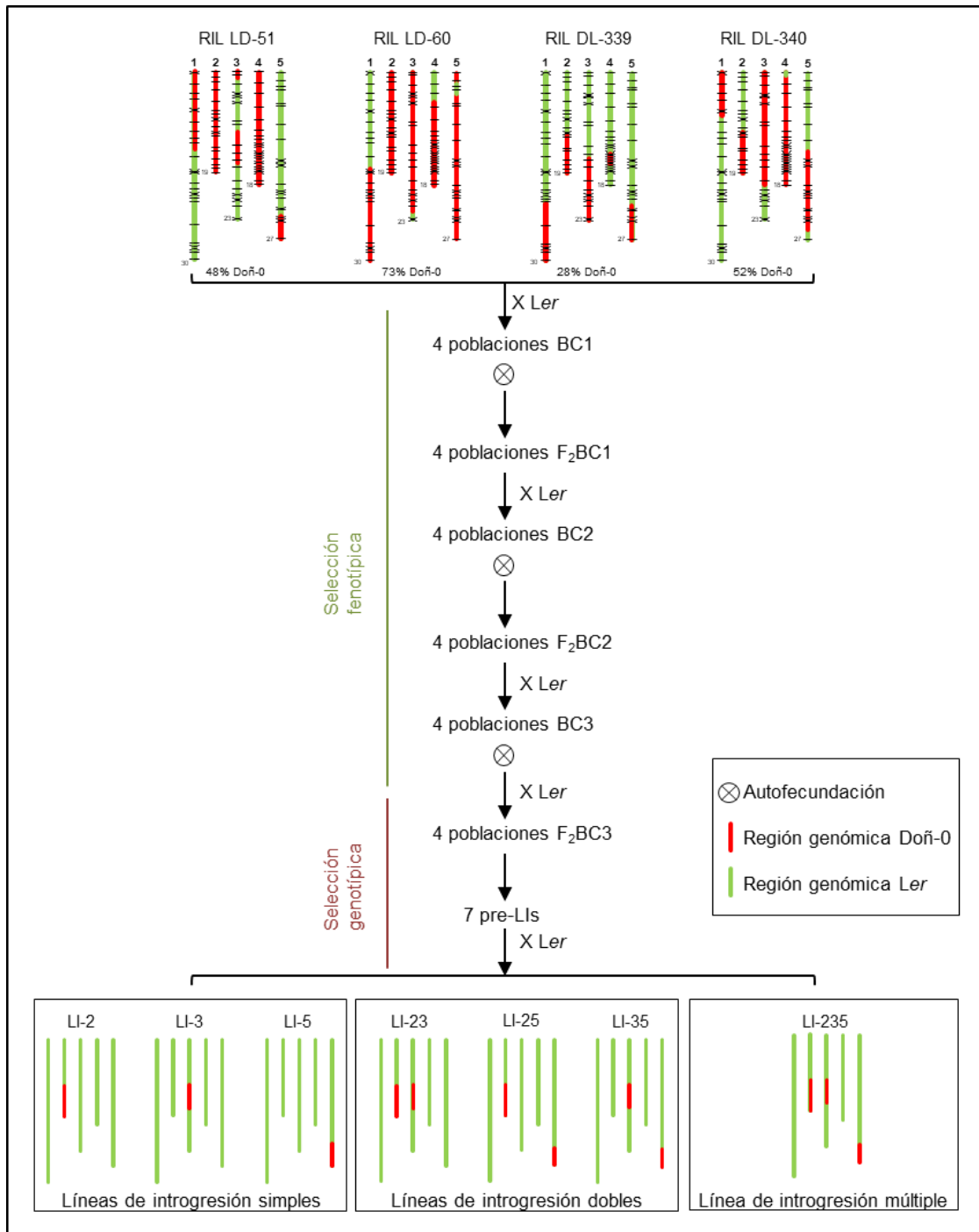
## 1. Material vegetal

### 1.1. Accesiones naturales para el análisis de asociación genómica

Para llevar a cabo el análisis de asociación genómica (GWAS) se ha utilizado una colección de 174 accesiones procedentes de distintas poblaciones naturales de la península ibérica (Figura 5A; Tabas-Madrid *et al.*, 2018). Estas poblaciones han sido previamente georreferenciadas para su altitud, longitud y latitud, y se han caracterizado en detalle a nivel ambiental (Manzano-Piedras *et al.*, 2014). Todas las accesiones son genéticamente distintas en base a la secuencia de sus genomas (The 1001 Genomes Consortium, 2016).

### 1.2. Líneas de introgresión de los loci *FTD2*, *FTD3* y *FTD5*

En esta tesis se han utilizado 21 líneas de introgresión (LIs) desarrolladas previamente en el laboratorio como portadoras de alelos de Doñ-0 en uno a tres loci *FTD*, en un fondo genético *Ler* (Savic, 2015). Las líneas de introgresión se obtuvieron a partir de cuatro líneas recombinantes consanguíneas (RILs), derivadas del cruzamiento de las accesiones Doñ-0 y *Ler* (LD-51, LD-60, DL-339, DL-340) (Figura 6), y que fueron seleccionadas por tener una alta densidad de tricomas en el fruto y alelos de Doñ-0 en los loci *FTD2*, *FTD3* y *FTD5*. En la Figura 6 se muestra el esquema de cruzamientos y selección fenotípica o genotípica utilizado para obtener las LIs a partir de las RILs. Brevemente, estas RILs se retrocruzaron con *Ler* durante tres generaciones sucesivas y cada generación de retrocruzamiento fue seguida de una generación de autofecundación. De esta forma se generaron cuatro familias de retrocruzamientos independientes, en cuyas poblaciones  $F_2BC_1$  y  $F_2BC_2$  se fenotiparon unas 100 plantas para seleccionar en cada población y generación, el individuo que presentó mayor número de tricomas en los frutos y que fue cruzado con *Ler* para obtener la siguiente generación de retrocruzamiento. Posteriormente, en la generación  $F_2BC_3$  se seleccionaron unas 40 plantas por población para su análisis genotípico con unos 300 SNPs que cubren todo el genoma. De esta manera se obtuvieron siete líneas, denominadas pre-LIs, como portadoras de alelos Doñ-0 en dos a tres loci *FTD* introgresados en fondo genético *Ler*. Estas pre-LIs se cruzaron entre ellas o con *Ler* para obtener siete familias segregantes en los loci *FTD* que fueron genotipadas para seleccionar las LIs definitivas. Finalmente se seleccionaron tres LIs independientes de cada uno de los siete genotipos diferentes correspondientes a todas las combinaciones de alelos Doñ-0 en uno a tres loci *FTD2*, *FTD3* y *FTD5* (Figura 6).



**Figura 6. Esquema de la obtención de las LIs en los loci *FTD2*, *FTD3* y *FTD5*.** Esquema de cruzamientos y selección fenotípica o genotípica realizados para obtener las LIs. La figura muestra el genotipo gráfico de las RILs utilizadas para generar las LIs (parte superior) y de las LIs obtenidas (parte inferior). Las regiones genómicas de Doñ-0 y Ler se muestran en color rojo y verde, respectivamente.

Las líneas de introgresión se nombraron en función de la localización cromosómica de los loci *FTD* introgresados, de manera que los genotipos LI-2, LI-3 y LI-5 contienen alelos Doñ-0 en un solo locus *FTD*, mientras que los genotipos LI-23, LI-25, LI-35 y LI-235 contienen alelos Doñ-0 en dos a tres loci *FTD*.



### 1.3. Población de mapeo de *FTD1*

Para realizar el mapeo fino del locus *FTD1* se seleccionó una línea de introgresión de fondo genético Ler, homocigótica para los alelos Doñ-0 en las regiones genómicas de *FTD2* y *FTD5*, y heterocigótica para la región de *FTD1* (Figura 7). A partir de esta línea se generó por autofecundación una población de 1.800 plantas que fue analizada con marcadores moleculares para seleccionar los individuos recombinantes entre las posiciones físicas 23,2 y 27 Mb del cromosoma 1 (ver apartado 4.2). Todas las plantas recombinantes fueron autofecundadas y 15-20 plantas de su descendencia por autofecundación, fueron nuevamente genotipadas para seleccionar individuos homocigóticos para el correspondiente evento de recombinación. De esta manera se generaron 110 líneas recombinantes homocigóticas que fueron fenotipadas en base al número de tricomas en el fruto (NTF) para determinar su genotipo en el locus *FTD1*. Para ello, se fenotiparon de 15-30 plantas descendientes por autofecundación de cada línea recombinante y la media del NTF de cada línea se comparó estadísticamente con el de las líneas homocigóticas Ler y Doñ-0 para toda la región de *FTD1*.



**Figura 7. Genotipo gráfico de la planta madre de la población de mapeo de *FTD1*.** Las regiones genómicas homocigóticas Doñ-0 o Ler se muestran en rojo y verde, respectivamente. La región heterocigótica de *FTD1* se muestra en amarillo.

### 1.4. Mutantes de genes del patrón de tricomas

En esta tesis se han utilizado líneas mutantes de pérdida de función de los genes implicados en el patrón de tricomas *TCL1*, *TRY*, *GL1* y *EGL3*, obtenidas a partir de los bancos internacionales de germoplasma de *Arabidopsis*. Los mutantes *tcl1-Salk* (línea Salk-055460C) y *try1-Salk* (línea N6518) son portadores de una inserción de T-ADN en los correspondientes genes, en un fondo Columbia (Col). El doble mutante *tcl1-Salk try-Salk* se ha obtenido previamente en el laboratorio a partir del cruzamiento de ambos mutantes simples (Savic, 2015). El mutante *gl1-1* (línea CS3748) es completamente glabro por lo que probablemente también es un alelo nulo causado por un cambio de aminoácido en la posición Ser<sup>92</sup> de la proteína GL1, en un fondo Col (Savic, 2015). Además, el mutante simple *egl3* (N446312) que posee una inserción de T-ADN en el gen *EGL3*, también en un fondo Col. Por último, la línea doble mutante

*gl3-1 egl3-1* (línea N66490) de fondo genético *Ler*, es portadora de una mutación inducida por EMS en los genes *GL3* y *EGL3* que provocan un codón de parada prematuro en cada gen (Koorneef *et al.*, 1982; Zhang *et al.*, 2003). Todas las líneas mutantes fueron analizadas genotípicamente mediante marcadores específicos para cada mutación (ver apartado 4.3).

### 1.5. Líneas transgénicas

En esta tesis se ha obtenido un total de 181 líneas transgénicas homocigóticas T<sub>3</sub>/T<sub>4</sub> para distintos alelos naturales de los transgenes *TCL1*, *TRY* y *GL1* (Tabla 1), a partir de las semillas T<sub>2</sub> generadas previamente en el laboratorio (Savic, 2015). Para los genes *TCL1* y *TRY* se obtuvieron líneas transgénicas con cuatro variantes del transgén, correspondientes a los alelos naturales de *TCL1* o *TRY* de *Ler* y Doñ-0, y a dos transgenes quiméricos *Ler*/Doñ-0 o Doñ-0/*Ler* (Tabla 1; ver apartado 6). Cada uno de los cuatro transgenes de *TCL1* y *TRY* se transformó en dos fondos genéticos: el mutante *tcl1-Salk* o *try-Salk*, y la línea de introgresión LI-235 portadora de alelos Doñ-0 en los loci *FTD2*, *FTD3* y *FTD5*. Para el gen *GL1* se obtuvieron líneas transgénicas de los alelos naturales de *Ler* y Doñ-0 en los dos fondos genéticos mutantes *gl1-1* y *tcl1-Salk try-Salk*.

Para obtener las líneas transgénicas analizadas en este trabajo, se sembraron las semillas T<sub>2</sub> ya disponibles (Savic, 2015) en medio sólido Murashige-Skoog (MS) de cultivo *in vitro* suplementado con higromicina (40 mg/L) y se seleccionaron las líneas portadoras de un único transgén en base a la resistencia al antibiótico. Las líneas T<sub>2</sub> con un único inserto se multiplicaron en el invernadero y su descendencia T<sub>3</sub> se sembró de nuevo en condiciones *in vitro* con higromicina para seleccionar las plantas homocigóticas para el transgén. Finalmente, las líneas transgénicas T<sub>3</sub> homocigóticas se genotiparon para verificar tanto el fondo genético como el transgén (ver apartado 4) y se multiplicaron para obtener la generación T<sub>4</sub>. Todos los experimentos de caracterización fenotípica y molecular de las líneas transgénicas para *TCL1*, *TRY* y *GL1* se llevaron a cabo utilizando las semillas T<sub>3</sub> o T<sub>4</sub> homocigóticas.

Por otra parte, en esta tesis se ha iniciado la obtención de líneas transgénicas T<sub>1</sub> para los alelos de *Ler* y Doñ-0 del gen *EGL3* en las líneas *gl3-1 egl3-1* y LI-25 portadora de alelos Doñ-0 en los loci *FTD2* y *FTD5* (ver apartado 6).

Gen	Fondo genético	Alelo transgén	Nº líneas T <sub>3</sub> /T <sub>4</sub> independientes
TCL1	tcl1-Salk	Ler	15
		Doñ-0	12
		Ler/Doñ-0	13
		Doñ-0/Ler	10
	LI-235	Ler	8
		Doñ-0	7
		Ler/Doñ-0	3
		Doñ-0/Ler	6
GL1	gl1-1	Ler	12
		Doñ-0	11
	tcl1try	Ler	9
		Doñ-0	6
TRY	try-Salk	Ler	14
		Doñ-0	12
		Ler/Doñ-0	16
		Doñ-0/Ler	9
	LI-235	Ler	3
		Doñ-0	8
		Ler/Doñ-0	3
		Doñ-0/Ler	4

**Tabla 1. Líneas transgénicas para los genes candidatos de los loci *FTD2*, *FTD3* y *FTD5*.** Líneas transgénicas T<sub>3</sub>/T<sub>4</sub> generadas para *TCL1*, *GL1* y *TRY* a partir de material T<sub>2</sub> previo. Para cada tipo de línea transgénica se indica el transgén, el fondo genético, la variante genética del transgén y el número de líneas T<sub>3</sub>/T<sub>4</sub> independientes y homocigóticas.

## 2. Condiciones de cultivo

Para obtener la mayoría de las plantas analizadas en esta tesis, las semillas se sembraron en placas Petri sobre papel de filtro húmedo y se incubaron en una cámara húmeda a 4 °C durante 4 días para romper su dormición. A continuación, las placas con semillas se trasladaron a una cámara climática a 21 °C y con fotoperiodo de día largo (DL) (16 horas de luz, 8 horas de oscuridad) durante 4 días, con el fin de promover la germinación. Finalmente, las plántulas se trasladaron en macetas o bandejas de alveolos conteniendo una mezcla de tierra y vermiculita en proporción 3:1, y fueron cultivadas en distintos tipos de tiestos y condiciones ambientales dependiendo del experimento.

### 2.1. Mapeo fino de *FTD1*

Para llevar a cabo el mapeo fino del locus *FTD1* las plantas se cultivaron en bandejas de 28 alveolos con cinco plántulas por alveolo, en un invernadero a 21 °C suplementado con lámparas que proporcionan un fotoperiodo de DL.

## **2.2. Análisis fenotípicos y de expresión génica en líneas de introgresión o transgénicas**

Para caracterizar fenotípicamente las líneas de introgresión o transgénicas se cultivaron 18 plantas por línea distribuidas en tres macetas de 8x8x8 cm (6 plantas por maceta), en una cámara climática visitable a 21 °C y con fotoperiodo de DL. Todas las líneas de introgresión y todas las líneas transgénicas de cada gen se cultivaron simultáneamente en el mismo experimento con tres bloques completos organizados aleatoriamente (un tiesto por bloque) y situados en diferentes zonas de la misma cámara.

Para los análisis de expresión génica en fase vegetativa se sembraron tres tiestos de 8x8x8 cm por cada línea de introgresión o transgénica, con unas 100 semillas cada uno. Los tiestos se colocaron en una cámara a 4 °C y fotoperiodo de día corto (DC) (8 horas de luz/16 horas de oscuridad) durante 4 días. Posteriormente, estos tiestos se pasaron a una cámara climática a 21 °C y con fotoperiodo de DL donde se organizaron en tres bloques aleatorios en diferentes zonas de la cámara. Para los análisis de expresión génica en fase reproductiva se plantaron otros tres tiestos de 8x8x8 cm por línea, con nueve plántulas por tiesto, que también se organizaron en tres bloques aleatorios y separados en la misma cámara climática. Para reducir efectos ambientales, las muestras para los análisis de expresión en fase vegetativa y reproductiva de todas las líneas de introgresión o de todas las líneas transgénicas de cada gen se cultivaron simultáneamente en el mismo experimento. Todas las plántulas de las tres réplicas biológicas vegetativas de cada línea se recogieron en un único tubo a los 14-16 días de crecimiento. De igual forma, los ápices de todas las inflorescencias principales (yemas florales sin fructificar) de las tres réplicas de cada línea se recogieron en un único tubo, 3-5 días después del inicio de la floración. Ambos tubos se utilizaron posteriormente para la extracción de ARN (ver apartado 7).

## **2.3. Caracterización fenotípica de las accesiones naturales y su tratamiento hormonal**

La caracterización fenotípica del patrón de tricomas de las 174 accesiones naturales y su respuesta al ácido jasmónico (AJ) se realizó en una cámara climática a 21 °C y con fotoperiodo de día corto (DC). Este experimento se realizó en DC porque la mayoría de las accesiones florecen más tarde y producen más hojas en DC que en DL. De esta forma se extendió y sincronizó el periodo de la fase vegetativa en el que poder evaluar

el efecto del AJ sobre la densidad de tricomas. Para cada accesión se cultivaron 24 plantas en cuatro tiestos de 8x8x8 cm (6 plantas por tiesto) organizados en dos bloques aleatorios formados por dos tiestos cada uno. A los 30-40 días después de la siembra, una de las macetas de cada bloque fue tratada con AJ, cuando las plantas tienen un diámetro de la roseta de 3-5 cm, con 6-8 hojas completamente expandidas de un total de 8-11 hojas. El tratamiento de AJ se aplicó mediante la dispensión de 80-100  $\mu$ l por planta de AJ (0,15 mM en dimetilformamida; Duchefa Biochemie) con una micropipeta de 1 ml, distribuidos sobre las cuatro hojas mayores y el meristemo. La otra maceta de cada bloque fue tratada con una solución control de dimetilformamida pero sin AJ, aplicada de forma similar. El momento de la fase vegetativa, la cantidad y la concentración del tratamiento con AJ se establecieron en un experimento preliminar previo en el que se ensayaron distintas cantidades y tiempos de aplicación. De esta forma se eligió la concentración utilizada por no producir daños en la hoja inoculada, pero mostrar un ligero efecto sobre el crecimiento de las hojas posteriores. Para determinar el efecto sistémico del AJ sobre las hojas desarrolladas con posterioridad al tratamiento, 20-22 días después del mismo se recogió la hoja en posición 14-16 de cada planta y ésta se utilizó para estimar la densidad de tricomas (ver apartado 3). Las hojas para el análisis fenotípico se recolectaron cuando estaban completamente expandidas, lo que corresponde a plantas de un diámetro de roseta de 8-14 cm. Posteriormente, las plantas se trasladaron durante dos meses a una cámara a 4 °C y con fotoperiodo de DC para su vernalización. Transcurrido este periodo las plantas se llevaron a una cámara climática a 21°C y con fotoperiodo de DL hasta su floración y fructificación, y se caracterizaron fenotípicamente para el patrón de tricomas en fase reproductiva (ver apartado 3).

### **3. Análisis fenotípico del patrón de tricomas**

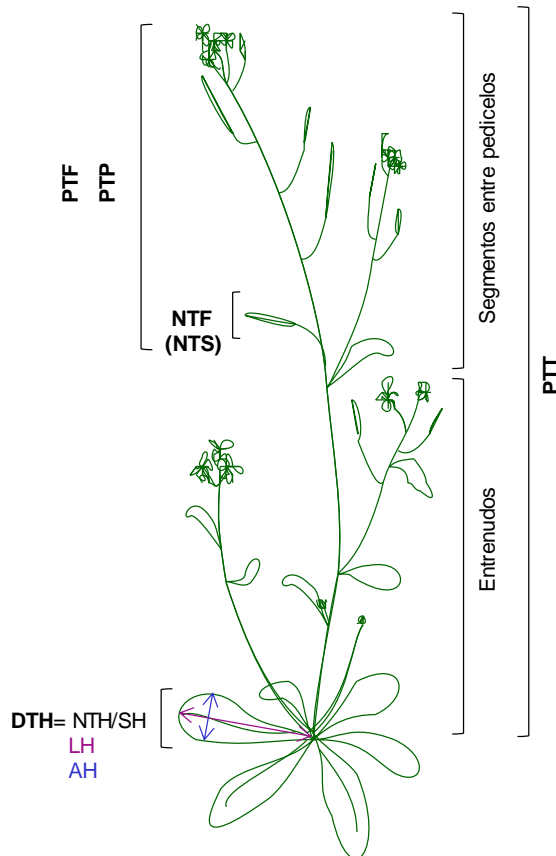
En esta tesis se ha analizado el patrón y la densidad de tricomas en fase vegetativa (hojas) y reproductiva (tallo, pedicelo y fruto) (Figura 8).

El patrón de tricomas en fase vegetativa se estimó cuantificando la densidad de tricomas en la hoja (DTH). En el caso de la caracterización de las líneas transgénicas o las líneas de introgresión cultivadas en condiciones de DL, este parámetro se midió en la hoja en posición 5-6 de la roseta, cuando está totalmente expandida. Para la caracterización de las accesiones se utilizó la hoja en posición 14-16 ya que este experimento se llevó a cabo en condiciones de DC. Las hojas recogidas se extendieron inmediatamente y se pegaron por su cara abaxial, sobre un folio, mediante

cinta adhesiva de doble cara y a continuación se fotografiaron. Estas fotografías se utilizaron para contar el número de tricomas (NTH) en la cara adaxial mediante el contador manual del programa de análisis de imagen ImageJ (<http://imagej.net>). La superficie de la hoja (SH) se estimó a partir de la medida de la longitud (LH) y la anchura (AH) de su limbo, utilizando la fórmula de la superficie de una elipse ( $0,5LH \times 0,5AH \times \pi$ ). La densidad de tricomas de la hoja se calculó como la relación NTH/SH (Figura 8). Además, en el experimento de caracterización de accesiones naturales se estimó la plasticidad o la respuesta al AJ para el tamaño, el número y la densidad de tricomas en la hoja, a través de un índice de plasticidad relativo que compara los valores medios de las plantas tratadas con AJ (SHAJ, NTHAJ y DTHAJ) y no tratadas (NTH, SH y DTH) (Valladares *et al.*, 2006). La plasticidad del tamaño de la hoja al AJ (PSH) se estimó como  $PSH = (SHAJ - SH) / (SHAJ + SH)$ ; la del número de tricomas (PNTH) como  $PNTH = (NTHAJ - NTH) / (NTHAJ + NTH)$ ; y la de la densidad de tricomas (PDTH) como  $PDTH = (DTHAJ - DTH) / (DTHAJ + DTH)$ . Los parámetros de plasticidad varían entre -1 y 1 (máxima plasticidad), con valor de 0 para la ausencia de plasticidad.

El patrón de tricomas en el tallo (PTT) se estimó como la longitud del tallo de la inflorescencia principal que presenta tricomas, midiendo dicha longitud en número de entrenudos (segmentos entre hojas caulinares) más el número de segmentos entre pedicelos. De forma similar se cuantificó el patrón de tricomas de los pedicelos (PTP) y de los frutos (PTF) a través del número total de pedicelos o frutos de la inflorescencia principal con tricomas (Figura 8).

El número de tricomas por fruto (NTF) se estimó como el número total de tricomas en el primer fruto fértil de la inflorescencia principal, que corresponde al fruto en posición uno a tres. Este valor se determinó cuando el fruto está completamente desarrollado, contando el número de tricomas visibles con una lupa binocular de 30 aumentos (30x). De forma similar se estimó el número de tricomas en los sépalos (NTS), correspondiente al número total de tricomas en los cuatro sépalos de la primera o segunda flor de la inflorescencia principal (Figura 8).



**Figura 8. Caracteres del patrón de tricomas fenotipados.** En la figura se muestra un esquema de *Arabidopsis* indicando los distintos caracteres analizados en las hojas y en la inflorescencia principal. DTH: densidad de tricomas en la hoja; LH: longitud de la hoja; AH: anchura de la hoja; SH: superficie de la hoja; PTT: patrón de tricomas en el tallo; PTP: patrón de tricomas en el pedicelo; PTF: patrón de tricomas en el fruto; NTF: número de tricomas en el fruto; NTS: número de tricomas en los sépalos (NTS se indica entre paréntesis ya que no está representada la primera flor).

#### 4. Genotipado de marcadores moleculares

En esta tesis se han utilizado numerosos marcadores moleculares de tipo inserción/delección (indel), microsatélite, CAPS (“*Codominant Amplified Polymorphic Sequence*”), dCAPS, así como marcadores genotipados por secuenciación (SEC) (Tablas 2-5).

La extracción de ácido desoxirribonucleico (ADN) se realizó a partir de muestras de tejido vegetativo (2-4 hojas de la roseta), o reproductivo (inflorescencia principal o lateral), utilizando el protocolo descrito por Bernatzky & Tanksley (1986) sin mercaptoetanol. El tejido se recogió en tubos de 2 ml, se congeló en nitrógeno líquido o en hielo seco y se almacenó a  $-80^{\circ}\text{C}$  hasta su utilización.

El genotipado de todos los marcadores moleculares se realizó a través de la amplificación de fragmentos de ADN mediante la reacción en cadena de la polimerasa (PCR) y su posterior separación y visualización en geles de agarosa (Konieczny & Ausubel, 1993). Las PCRs se realizaron en un volumen final de 50  $\mu\text{l}$ , con una mezcla de 5 ng de ADN, dNTPs (0,2 mM), oligonucleótidos del marcador (100 ng de cada uno), buffer de PCR (1x),  $\text{MgCl}_2$  (2 mM) y 0,3  $\mu\text{l}$  del enzima AmpliTaq polimerasa

(Perkin, Applied Biosystems, USA). En el caso de los marcadores de tipo CAPS, el fragmento amplificado fue digerido con el correspondiente enzima de restricción siguiendo las indicaciones del fabricante (New England Biolabs, Beverly, EE.UU.). En los marcadores genotipados por secuenciación, el fragmento amplificado fue purificado mediante tratamiento con Illustra ExoProStar (GE Healthcare Life Science, Reino Unido) antes de ser secuenciado con los mismos oligonucleótidos utilizados para la PCR.

#### 4.1. Genotipado de las líneas de introgresión

Para determinar el genotipo de las LIs se han utilizado 11 marcadores moleculares de tipo indel y microsatélite localizados en las regiones genómicas de los loci *FTD* (Tabla 2).

Locus	Nombre del marcador	Cromosoma	Posición (Mb)	Clase de marcador	Fragmento Ler (pb)	Fragmento Doñ-0 (pb)	Secuencia del oligonucleótido
<i>FTD1</i>	Don23,4Dinser	1	23,416	indel	229	250	F-gggttataggagactgactg R-tcgactgtgatattgtgctc
	Don23,59del	1	23,598	indel	510	480	F-ttgaccagtaaggcagtaacata R-tgtccaagcaccactctcc
	Don24,17del	1	24,175	indel	370	300	F-ccttgatctcctctgcttat R-acgttaacggctgctacatc
<i>FTD2</i>	TCL1-5 <sup>1</sup>	2	12,971	indel	500	250	F-agaaatcaaatcagggcgtaaaaa R-cccctagattgtgaaagttaa
	TCL1-3 <sup>1</sup>	2	12,969	indel	800	650	F-cgatgacttgcctaacaattgt R-tagacccttgcataccacgg
<i>FTD3</i>	MSAT3.32 <sup>2</sup>	3	11,208	microsatélite	148	250	F-gcacttgcagcttaactt R-cgtgactgtcaaaccg
<i>FTD4</i>	TRIPLET <sup>1</sup>	4	7,807	indel	300	280	F-gatgtttggctcctggtgctt R-ccggttgcacgcttatcag
	CIW7 <sup>2</sup>	4	11,524	microsatélite	120	100	F-aatttgagattagctggaat R-ccatgttgatgataagcacia
	nga1139 <sup>2</sup>	4	16,444	microsatélite	120	90	F-tagccggatgagttgtacc R-ttttccttgtgtgattcc
<i>FTD5</i>	TRY-5 <sup>1</sup>	5	21,581	indel	150	140	F-gctccaaactaaagacacatttc R-ggaagcctagtgtgtaaaatactc
	TRY-3 <sup>1</sup>	5	21,584	indel	500	1000	F-tctacgacttttatccacaacta R-tcaaggactaaacttaatacacia

**Tabla 2. Marcadores utilizados para genotipar las líneas de introgresión.** Para cada marcador se muestra el locus *FTD*, el nombre del marcador, el cromosoma, la posición física, el tipo de marcador, el tamaño de los fragmentos amplificados en Ler y Doñ-0, y la secuencia de los oligonucleótidos. <sup>1</sup> Marcadores previamente desarrollados en el laboratorio. <sup>2</sup> Microsatélites obtenidos de Bell & Ecker (1994), Loudet *et al.* (2002) y Lukowitz *et al.* (2000).

#### 4.2. Genotipado de la población de mapeo de *FTD1*

Para llevar a cabo el mapeo fino del locus *FTD1* se desarrollaron 22 marcadores de tipo indel, CAPS y marcadores genotipados por secuenciación entre las posiciones 23,2 y 27 Mb del cromosoma 1 (Tabla 3).



Nombre del marcador	Posición física (Mb)	Clase de marcador	Enzima de restricción	Fragmento Ler (pb)	Fragmento Doñ-0 (pb)	Secuencia del oligonucleótido
Don27-del	27,115	indel	-	315	280	F-gacaataagagagacatggcggg R-gctcatcatctcttctctca
Don26,4del	26,400	indel	-	560	500	F-ccagcataacaaaaccttctcc R-cctttttaccttcccaccgctc
Don25,4del	25,462	indel	-	400	355	F-gtgggtgaagattggataagatgt R-ttacaatggctcgaacaactgg
Don24,9-del	24,942	indel	-	460	400	F-gattcggaaaacacacgcaag R-gggctgtttttactactctcacctc
Don24,17-del	24,175	indel	-	370	300	F-ccttgatctctctgctgtat R-acgttaacggctgctacatc
Don23,80Del	23,802	indel	-	134	124	F-aaaaaaagtcaggatggccga R-cccaaaaacttatcttgaacaaat
Don23,6del	23,657	indel	-	150	120	F-ccagcataacaaaaccttctcc R-cctttttaccttcccaccgctc
1-23,64seq	23,641	SEC	-	450	450	F-gcaagaactatgagaagggtg R-gcgaagaagaagaggagcc
1-23,62seq	23,623	SEC	-	715	715	F-gccttactctgtgatgc R-caaatgttctatgtcccttc
1-23,63seq	23,621	SEC	-	790	790	F-gcatttcacctactttggg R-ccagcttgatgcacgtg
EGL3	23,601	indel	-	223	200	F-caccagaggaaaccgggaa R-tgtagaatcagcaagagactcaag
Don23,59del	23,598	indel	-	510	480	F-tttgaccagtaaggcagtaacata R-tgtccaagcaccactctcc
pEGL3-6seq	23,595	SEC	-	550	550	F-cgagaagaccctaagcaaatg R-gattcaaaactccgcctcc
1-23,589seq	23,589	SEC	-	720	720	F-ccgccttattgagtggttc R-tgtgctggcgaagggaatc
Don23,580seq	23,581	SEC	-	565	565	F-ttaccgtaaccctactctgct R-tgttagatccttggtcagtc
1-23,57A	23,570	CAPS	<i>Ava</i> II	800	700/100	F-cgtttgtctcaggtgctgct R-tacaaagacggaaccaatagat
Don23,534	23,534	CAPS	<i>Hae</i> III	510/490	1000	F-ctggcttaatgggcttcta R-ctggcttaatgggcttcta
Ler23,474	23,474	CAPS	<i>Dra</i> I	400	550	F-tgcctgtttggctatgga R-ggaagagaacgagaaggcaa
Don23,4Dinser	23,416	indel	-	229	250	F-gggttataggagactgactg R-tcgactgtgatattgctc
Don23,31del	23,310	indel	-	137	120	F-aatggtttgacacacatg R-tatctctgtctttgactctcc
Don23,28del	23,283	indel	-	300	100	F-caagagaacaacaagaccggg R-gggtttcgggttcactca
Don23,2del	23,253	indel	-	280	230	F-atttgagtggtatctgtgtg R-tgtgttactgcacgccc

**Tabla 3. Marcadores desarrollados para el mapeo fino del locus *FTD1*.** Para cada marcador se muestra el nombre del marcador, la posición física en el cromosoma 1, el tipo de marcador, el enzima de restricción usado en los CAPS, el tamaño de los fragmentos amplificados y la secuencia de los oligonucleótidos. SEC: marcador genotipado por secuenciación.

#### 4.3. Genotipado de líneas mutantes

Para genotipar los mutantes del patrón de tricomas se utilizaron seis marcadores que distinguen específicamente los alelos mutantes de los alelos silvestres (*Ler* para *gl3-1* *egl3-1* y *Col* para *egl3*, *tcl1-Salk*, *try-Salk* y *gl1-1*) (Tabla 4). Los marcadores de tipo CAPS o CAPS degenerados (dCAPS) se diseñaron usando el programa dCAPS Finder (<http://helix.wustl.edu/dcaps>). El genotipado de los mutantes de inserción de T-ADN Salk se realizó utilizando dos oligonucleótidos diseñados sobre el gen que

amplifican el alelo silvestre, y un oligonucleótido del borde izquierdo (“left border” -LB-) de la inserción de T-ADN que junto con los otros dos amplifica el alelo mutante. En cuanto al mutante *egl3* se utilizó un oligonucleótido específico de la inserción de T-ADN Gabi Kat (GK8474) y un oligonucleótido diseñado sobre el gen *EGL3*, que amplifican el alelo mutante pero no el alelo silvestre.

Gen	Alelo mutante	Nombre del marcador	Clase de marcador	Enzima de restricción	Fragmento alelo silvestre (pb)	Fragmento alelo mutante (pb)	Secuencia del oligonucleótido
<i>EGL3</i>	<i>egl3-1</i>	<i>egl3-MnlI</i>	dCAPS	<i>MnlI</i>	200	220	F-cacttctgctttttctgtttctc R-agacagaccagaagattccatacct
	<i>egl3</i>	<i>egl3 mutant</i>	inserción T-ADN	-	na	1000	F-ataataacgctgcgacatctacattt <sup>3</sup> R-cactgtctgaagcgaagca
<i>GL3</i>	<i>gl3-1</i>	<i>gl3-AciI</i>	CAPS	<i>AciI</i>	280/150	280/250	F-attcgtcttttctctatcag R-acactttactatccgccg
<i>TCL1</i>	<i>tcl1-Salk</i>	<i>tcl mutant</i> <sup>1</sup>	inserción T-ADN	-	700	500	F-aattcggatgctgtgagagtc R-cgtcccaccactcttctgc F-ggcaatcagctgtgcccgtctcactggg <sup>2</sup>
							<i>TRY</i>
<i>GL1</i>	<i>gl1-1</i>	<i>GL1-HaeIII</i> <sup>1</sup>	CAPS	<i>HaeIII</i>	153/17	170	F-atatttttccccaaattcagatggc R-cagtgatgaacaatgacgggtga

**Tabla 4. Marcadores utilizados para el genotipado de las líneas mutantes.** Para cada marcador se indica el gen mutado, el nombre del mutante, el nombre y la clase de marcador, el enzima de restricción utilizado en los CAPS o dCAPS, el tamaño del fragmento amplificado en los alelos silvestre y mutante, y las secuencias de los oligonucleótidos. <sup>1</sup>Marcadores diseñados previamente en el laboratorio. <sup>2</sup> Oligonucleótido del T-ADN Salk. <sup>3</sup> Oligonucleótido de T-ADN Gabi Kat. na: no amplifica.

#### 4.4. Genotipado de líneas transgénicas

El transgén de las líneas transgénicas homocigóticas se genotipó con marcadores moleculares que distinguen los alelos de *Ler* y *Doñ-0* de *TCL1*, *TRY* o *GL1* (Tabla 5). El fondo genético de cada línea transgénica se verificó con marcadores específicos de los mutantes *tcl1-Salk*, *try-Salk* y *gl1-1* utilizados (Tabla 4) y con los marcadores de las regiones *FTD* introgradadas (Tabla 2).

Transgén	Nombre del marcador	Clase de marcador	Enzima de restricción	Fragmento <i>Ler</i> (pb)	Fragmento <i>Doñ-0</i> (pb)	Secuencia del oligonucleótido
<i>TCL1</i>	<i>TCL1-5'</i>	indel	-	500	250	F-agaaatcaaatcgaggcgtaaaaa R-cccctagattgtgaaagtga
	<i>TCL1-3'</i>	indel	-	800	650	F-cgatgacttgccaacaattgt R-tagacccttgataccacgg
<i>GL1</i>	<i>GL1-AT</i>	indel	-	116	123	F-ctcactttgtacgtagacgcttt R-ggacctaattatatatgtgtgtgc
	<i>GL1-DraI</i>	CAPS	<i>DraI</i>	600/114	400/200/114	F-tgagaacaagacatatggaaaggag R-ttctcttttcatctcttcttatt
<i>TRY</i>	<i>TRY-5'</i>	indel	-	150	140	F-gctccaaactaaagacacatttc R-ggaagcctagtgtgtaaaatctc
	<i>TRY-3'</i>	indel	-	500	1000	F-tctacgacttttatccacaacta R-ttcaaaggactaaactaatacaca

**Tabla 5. Marcadores utilizados para el genotipado de los transgenes.** Para cada marcador se indica el transgén, el nombre y la clase de marcador, el enzima de restricción utilizado en el

CAPS, el tamaño del fragmento amplificado en *Ler* y Doñ-0, y las secuencias de los oligonucleótidos.

## 5. Secuenciación del gen *EGL3*

El gen *EGL3* de las accesiones *Ler* y Doñ-0 se ha secuenciado completamente incluyendo 4 kb de la región promotora, toda la región codificante (2,9 kb) y 1,2 kb de la región 3'. Para ello se diseñaron 15 pares de oligonucleótidos que amplifican fragmentos de 0,5-1 kb y que muestran un solapamiento de 100-250 pb entre fragmentos adyacentes (Tabla 6). Estos fragmentos se secuenciaron con los mismos oligonucleótidos usados para su amplificación, así como con oligonucleótidos diseñados específicamente en las regiones genómicas que contienen repeticiones poliA y poliT de difícil secuenciación (Tabla 6). Las secuencias obtenidas se alinearon y analizaron utilizando DNASTAR v.15 (DNASTAR Inc., Madison, WI, EE.UU.).

Siguiendo esta misma metodología también se ha secuenciado un fragmento de 2 kb de la región promotora más cercana al codón de inicio en otras 22 accesiones naturales procedentes de la península ibérica. Además, se ha secuenciado una región de 6 kb correspondiente a todo el promotor y parte de la región codificante de *EGL3* (2 kb), en tres líneas recombinantes homocigóticas obtenidas durante el mapeo fino de *FTD1*.

Gen	Nombre del oligonucleótido	Tamaño del fragmento (pb)	Secuencia del oligonucleótido
<i>EGL3</i>	pEGL3-6seq	550	F-cgagaagaccctaagcaaatg R-gattcaaacttcgcctcc
	pEGL3-5seq	890	F-aagccataaggagaagaagaa R-tattcttcgttctctgacctgatt
	pEGL3-4seq	700	F-cgttaacatcgaattccacca R-taaaacaagcaatagatactaaatacaact
	pEGL3-3seq	860	F-gagataaaaagcaagagaatgtgt R-agagtgatcgaacagtcattttt
	PoliT-1 <sup>1</sup>		F-tttttgtgtgtgtgtgtca R-aaaaacaagaactgatcaaatatg
	pEGL3-2seq	950	F-ttagatatgaatttggtggttta R-aggtaagggattgataatgggtt
	PoliT-2 <sup>1</sup>		F-ttttttatcaacagtagactct
	pEGL3-1seq	650	F-ggagatgcgaagaattaattaagagc R-acttacaatccaaccaaaccaat
	EGL3-1seq	790	F-ggttggttgatttgaagtg R-gacattggttggttctctg
	PoliT-3 <sup>1</sup>		F-agacactactttttttttttacc R-caagatgtaaaaaaaaaaaaaaagtagtg
	EGL3-2seq	950	F-caccgtcctttctctcaac R-catctcccactccaacac
	EGL3-3seq	740	F-ctggtctgtctctgcttctc R-tggaagaagaaatgaaggagac
	EGL3-4seq	795	F-cttcgtctcaacatcgggt R-cactgtctgaagcgaagca
	EGL3-5seq	800	F-ccgccgatagcaaatgctt R-gctgatgcggtgtttga
	PoliT-4 <sup>1</sup>		F-ttttttttttcggttggcgtttg R-caaacgaaaaaaaaaaaaaaaaagagcg
	EGL3-6seq	750	F-gggagacttgcttgcgat R-gctggtcatctctctctaca
	EGL3-7seq	660	F-acagagacacggataacgatg R-attctattgtctctctctacg
	EGL3-8seq	520	F-gttttgtatttgaagtttgg R-ctcagtgccgtagatgc
	PoliT-5 <sup>1</sup>		F-gaactttttttgtgttg
	EGL3-9seq	935	F-ctcgtagtagtctgtcac R-gcattattatctacctgtgttac
	PoliT-6 <sup>1</sup>		F-aaaaaatatttttagaaaag R-ctaaaaatatttttaatttc
	EGL3-BamHI	6700 <sup>2</sup>	F-caggatcctaattcattctagtgaaaaagattgagt
	EGL3-Sall		R-caagtcgacggatcttctaatggccacctacc

**Tabla 6. Oligonucleótidos utilizados para la secuenciación y clonación de *EGL3*.** Para cada fragmento se indica el nombre, el tamaño y la secuencia de los oligonucleótidos. <sup>1</sup> Oligonucleótidos diseñados específicamente para la secuenciación de regiones genómicas ricas en poliA y poliT. <sup>2</sup> Fragmento utilizado para la clonación de *EGL3*. La secuencia de los oligonucleótidos marcada en rojo corresponde a la diana de restricción.

## 6. Construcciones genómicas y transformación

En esta tesis se han obtenido dos construcciones genómicas del gen *EGL3* correspondientes a los alelos naturales de *Ler* y *Doñ-0* (Tabla 7). Para ello se clonó un fragmento de 6,7 kb que incluye 2,9 kb de la región promotora, la región codificante completa (2,9 kb) y 0,9 kb de la región 3', en el vector binario pCAMBIA 2300 (CAMBIA, Canberra, Australia). Dicho fragmento se amplificó a partir de ADN de *Ler* o *Doñ-0* usando una ADN polimerasa de alta fidelidad (*Phusion*, NEB, Beverly, EE.UU.) y oligonucleótidos específicos de *EGL3* a los que se introdujo la secuencia de restricción de *BamH I* o *Sal I* (Tabla 6), y el producto amplificado por PCR se purificó (*QIAquick PCR Purification Kit*; QIAGEN). El fragmento purificado y el vector pCAMBIA 2300 se digirieron con *BamH I* y *Sal I* (NEB, Beverly, EE.UU.) y se ligaron (*Rapid DNA Ligation Kit*; Roche, Mannheim, Alemania). El producto de ligación se usó para transformar bacterias competentes de la cepa DH5 $\alpha$  de *Escherichia coli* mediante el protocolo de choque térmico y las bacterias transformadas se seleccionaron con el antibiótico kanamicina (50 mg/L). Las construcciones se verificaron mediante la secuenciación completa de las dos hebras del inserto usando los oligonucleótidos descritos previamente (Tabla 6).

Las dos construcciones genómicas de *Ler* y *Doñ-0* se transfirieron por electroporación a la cepa AGLO de *Agrobacterium tumefaciens* (Lazo, 1991) y los clones transformantes se seleccionaron por su resistencia a rifampicina (25 mg/L) y kanamicina (50 mg/L). Posteriormente, *A. tumefaciens* se utilizó para transformar plantas de *Arabidopsis* del fondo genético *Ler* con las mutaciones *gl3-1 egl3-1*, y la línea de introgresión LI-25 siguiendo el método de inmersión floral (Clough & Bent, 1998).

Por otra parte, en esta tesis se han analizado líneas transgénicas obtenidas a partir de 10 construcciones genómicas de los genes *TCL1*, *GL1* y *TRY* generadas previamente en el laboratorio (Tabla 7). Cuatro construcciones incluyen variantes del gen *TCL1*, de las cuales dos corresponden a los alelos naturales de *Ler* y *Doñ-0*, y las otras dos son quiméricas entre ambas accesiones (Tabla 7). Éstas últimas se desarrollaron mediante el intercambio recíproco de un fragmento de 1,3 kb de la región 3' del gen *TCL1* de *Ler* y *Doñ-0*. Otras cuatro construcciones son del gen *TRY*, dos correspondientes a los alelos de *Ler* y *Doñ-0* y otras dos son quimeras obtenidas mediante el intercambio recíproco de un fragmento de 2,5 kb de la región 3' de *TRY*. Por último, dos construcciones son portadoras de los alelos naturales de *GL1* de *Ler* y *Doñ-0* (Tabla 7).

Gen	Alelo transgén	Tamaño genómico (kb)			Tamaño total (kb)	Enzimas de restricción
		Región promotora	Región codificante	Región 3'		
<i>EGL3</i>	Ler	2,90	2,90	0,90	6,70	<i>Bam</i> H I/ <i>Sal</i> I
	Doñ-0	2,90	2,90	0,90	6,70	<i>Bam</i> H I/ <i>Sal</i> I
<i>TCL1</i> <sup>1</sup>	Ler	2,05	1,20	0,55	3,80	<i>Kpn</i> I/ <i>Bam</i> H I
	Doñ-0	1,80	1,20	0,37	3,37	<i>Kpn</i> I/ <i>Bam</i> H I
	Ler/Doñ-0	2,05	1,20	0,37	3,62	<i>Xba</i> I/ <i>Bgl</i> II
	Doñ-0/Ler	1,80	1,20	0,55	3,55	<i>Xba</i> I/ <i>Bgl</i> II
<i>GL1</i> <sup>1</sup>	Ler	1,60	1,50	1,90	5,00	<i>Bam</i> H I/ <i>Sal</i> I
	Doñ-0	1,60	1,50	1,90	5,00	<i>Bam</i> H I/ <i>Sal</i> I
<i>TRY</i> <sup>1</sup>	Ler	3,87	1,05	1,32	6,24	<i>Sac</i> I/ <i>Sal</i> I
	Doñ-0	3,87	1,05	1,82	6,74	<i>Sac</i> I/ <i>Sal</i> I
	Ler/Doñ-0	3,87	1,05	1,82	6,74	<i>Kpn</i> I/ <i>Sal</i> I
	Doñ-0/Ler	3,87	1,05	1,32	6,24	<i>Kpn</i> I/ <i>Sal</i> I

**Tabla 7. Construcciones genómicas de los genes *EGL3*, *TCL1*, *GL1* y *TRY*.** Para cada construcción se indica el gen, la variante alélica, el tamaño de las regiones genómicas incluidas y los enzimas de restricción utilizados para el clonaje. <sup>1</sup> Construcciones previamente desarrolladas en el laboratorio.

## 7. Análisis de expresión génica

La expresión de los genes *EGL3*, *TCL1*, *GL1* y *TRY* se cuantificó mediante PCR cuantitativa a tiempo real (RT-PCR), tanto en fase vegetativa como reproductiva. Para ello se aisló el ácido ribonucleico (ARN) de tejido vegetativo (plántulas de 15 días) y tejido reproductivo (inflorescencia principal) de plantas cultivadas a 21 °C con fotoperiodo de DL (ver apartado 2). El ARN se extrajo usando el reactivo TRIzol (Invitrogen, Carlsbad, CA, EE.UU.) y a continuación se purificó ("*High Pure RNA Isolation Kit*"; Roche, Mannheim, Alemania). La calidad del ARN se determinó mediante visualización en geles de agarosa al 1,4% y MOPS 1X y su concentración se estimó con un espectrofotómetro (NanoDrop 1000, Thermo Fisher Scientific, Wilmington, DE, EE.UU.). A partir de este ARN se sintetizó el ADN complementario (ADNc) mediante reacción con transcriptasa inversa (AMV de Promega; Madison, WI, EE.UU.). Este ADNc se usó como molde para la amplificación de regiones génicas usando una mezcla de reactivos con colorante fluorescente *SYBR® Green*, (Applied Biosystems Life Technologies) y oligonucleótidos específicos de cada gen (Tabla 8) en un sistema de RT-PCR 7300 (Applied Biosystems). En cada amplificación se cuantificó la expresión del gen correspondiente por el método de la curva estándar y se estimó la media y el error estándar (EE) de tres amplificaciones independientes (réplicas técnicas). Los niveles de expresión de cada gen se normalizaron utilizando como control endógeno de las muestras de ARN el gen *UBIQUITIN-CONJUGATING ENZYME 21 (UBC)*, seleccionado por presentar una expresión constante entre tejidos y ambientes (Czechowski *et al.*, 2005). Para evitar diferencias de amplificación causadas por polimorfismos nucleotídicos entre accesiones naturales, los

oligonucleótidos utilizados para la cuantificación de la expresión génica se diseñaron específicamente en regiones génicas que no presentaban polimorfismos entre Ler y Doñ-0.

Gen	Tamaño del fragmento (pb)	Secuencia del oligonucleótido
<i>EGL3</i>	90	F-gcttaactgtcaattgcaagca R-ccttaacatatccatgcaaccct
<i>TCL1</i>	100	F-catactgtttcgtgttccaa <sup>1</sup> R-tgagagctatgagtgaactggg <sup>1</sup>
<i>GL1</i>	90	F-gtggagacgaataatgagg R-gttgtagactgagatggatg
<i>TRY</i>	100	F-ccatgactctgaagaagtgagc <sup>1</sup> R-cccacctatcaccgacaagt <sup>1</sup>
<i>UBC</i>	60	F-ctgcgactcaggaatcttctaa <sup>1</sup> R-ttggtccattgaattgaacc <sup>1</sup>

**Tabla 8. Oligonucleótidos utilizados para estimar la expresión de los genes estudiados.** La tabla contiene el nombre del gen, el tamaño del fragmento amplificado y la secuencia de los oligonucleótidos utilizados. <sup>1</sup> Oligonucleótidos desarrollados previamente en el laboratorio.

## 8. Análisis de asociación genómica

Los genomas de las 174 accesiones utilizadas en esta tesis han sido secuenciados como parte del proyecto de los 1.001 genomas de Arabidopsis (The 1001 Genomes Consortium, 2016; <http://1001genomes.org>). Estos genomas se han analizado previamente para detectar un total de 1,88 millones de SNPs que presentan el alelo de frecuencia menor (MAF) en cinco o más accesiones (MAF $\geq$ 3%) y con más del 55 % de las accesiones genotipadas (media de muestras sin datos del 7,2%) (Tabas-Madrid *et al.*, 2018). Estos SNPs se localizan en 32.147 genes de Arabidopsis lo que supone un 95,7% de los genes anotados en TAIR10. De media, los 174 genomas contienen un SNP cada 64 pb, lo que corresponde a 58 SNPs informativos por gen. Estos SNPs se han clasificado funcionalmente en cuatro categorías: codones de parada prematuros con un posible efecto funcional fuerte (0,2 %); cambios de aminoácido con posible efecto moderado (9,6 %); mutaciones sinónimas con efecto potencialmente bajo (13,0 %); y polimorfismos en regiones no codificantes con un posible efecto regulador (77,2 %) (Tabas-Madrid *et al.*, 2018).

Los análisis de asociación se realizaron utilizando los valores medios de cada accesión para las siguientes variables fenotípicas del patrón de tricomas: DTH, DTHAJ, PDTH, PTT y PTP. Las dos accesiones glabras (Hue-3 y Pdl-1) fueron eliminadas de estos análisis. Además, la accesión Pib-1 que mostró un valor extremo

de DTH y DTHAJ, fuera de la distribución del resto de la población, fue transformada al valor máximo de cada distribución. Por último, para reducir el número de falsas asociaciones derivadas de a la falta de normalidad de las variables de tricomas de la fase reproductiva, estas variables se transformaron de forma específica. En el caso de PTT, los valores mayores de 8 fueron transformados a dicho valor, mientras que la variable PTP fue convertida en una variable binaria (sin tricomas=0; con tricomas=0,1).

La asociación de cada SNP con cada uno de los caracteres fenotípicos estudiados (análisis de GWAS) se evaluó utilizando el modelo lineal mixto (MLM) implementado en el programa TASSEL v.5 (Bradbury *et al.*, 2007). En este modelo, la variable fenotípica es la variable dependiente que se explica por el efecto fijo de un SNP y por el efecto aleatorio de una matriz de parentesco entre las accesiones. De esta forma, el modelo MLM reduce la detección de falsos SNPs positivos que pueden aparecer asociados con los fenotipos debido a las relaciones genéticas entre accesiones, ya que incluye como covariable la matriz de parentesco. Esta matriz se estimó a partir de la proporción de alelos compartidos utilizando todos los SNPs disponibles (Atwell *et al.*, 2010). El umbral de significación estadística aplicado para detectar SNPs potencialmente asociados con la mayoría de los caracteres fenotípicos fue de  $-\log(P)=4$ , tal y como se ha descrito para análisis de grandes cantidades de SNPs sobre un mismo carácter (Kooke *et al.*, 2016; Togninalli *et al.*, 2018). Sólo en el caso de PTP se aumentó el umbral de significación estadística a  $-\log(P)=6$ , debido al exceso de asociaciones significativas que se obtuvieron derivadas del uso de una variable binaria.

El número de genes detectado en cada análisis de asociación genómica se estimó a partir de la posición de los SNP en la versión TAIR10 del genoma de Arabidopsis, pero incluyendo los dos genes flanqueantes cuando un SNP se localizó en una región intergénica.

Las heredabilidades de los caracteres fenotípicos en sentido amplio ( $h^2_b$ ) se estimaron a partir de modelos lineales generales que incluyen la variable de tricomas como variable dependiente y accesión (genotipo) como factor explicativo. La heredabilidad se estimó como la proporción total de la varianza fenotípica ( $V_P$ ) que es explicada por la varianza entre genotipos ( $V_G$ ) ( $h^2_b = V_G/V_P$ ). La varianza genotípica explicada por los 10 SNPs asociados más fuertemente en los análisis de asociación genómica de cada carácter se estimó de forma similar a partir de modelos lineales generales en los que se incluyeron como factores explicativos los 10 SNPs. La proporción de varianza explicada por estos 10 SNPs se le denominó heredabilidad de



los 10 loci máximos ( $h^2_{10 \text{ loci max}}$ ). Los modelos lineales generales se aplicaron utilizando el programa STATISTICA v.8. Además, la varianza genotípica de cada carácter explicada por las relaciones de parentesco existentes entre las accesiones (la estructura genética) se estimó como la varianza explicada por la matriz de parentesco y se le denominó heredabilidad del parentesco ( $h^2_{parentesco}$ ). Esta heredabilidad se obtuvo utilizando el método de la mejor predicción lineal implementada en TASSEL v.5 (Bradbury *et al.*, 2007).

Con el fin de identificar genes implicados en el desarrollo de tricomas que puedan contribuir a la variación natural del patrón de tricomas se elaboró una lista de 115 genes conocidos (Tabla Suplementaria 1), a partir de los datos descritos en la literatura y la base de datos TrichOME (<http://www.plantrichome.org>). Esta lista se utilizó para buscar SNPs situados 10 kb por encima y por debajo del codón de inicio o de parada, respectivamente.

## 9. Otros análisis estadísticos

Las relaciones entre los fenotipos y las expresiones génicas de distintos genotipos (accesiones, líneas de introgresión o líneas transgénicas) se analizaron mediante coeficientes de correlación de Pearson.

Las diferencias fenotípicas o de expresión entre los distintos genotipos, así como la asociación de los polimorfismos de *EGL3* con el patrón de tricomas se contrastaron estadísticamente mediante análisis de la varianza (“ANalysis Of VAriance” -ANOVA-) de uno, dos o tres factores, dependiendo del experimento. Las diferencias fenotípicas entre las líneas transgénicas portadoras de distintos transgenes se contrastaron mediante ANOVAs mixtos incluyendo el transgén como factor fijo y las líneas como factor aleatorio. El porcentaje de varianza explicado por cada factor se determinó a través de análisis de componentes de varianza. Las diferencias fenotípicas entre las distintas líneas se determinaron mediante contrastes por pares con la corrección de Bonferroni. Por último las diferencias entre las accesiones tratadas con y sin AJ se analizaron estadísticamente mediante el test *t*-Student. Todos estos análisis se realizaron usando los paquetes estadísticos STATISTICA v.8 o SPSS v.24.

## 10. Diversidad nucleotídica y análisis filogenéticos

El análisis de la diversidad nucleotídica de *EGL3* se ha llevado a cabo a partir de la secuencia nucleotídica de la región promotora y la región codificante de *EGL3* en 23 y 145 accesiones de la península ibérica, respectivamente, utilizando el programa DNAsp v.6 (Rozas *et al.*, 2017). El análisis filogenético de estas secuencias nucleotídicas y de las secuencias de la proteína EGL3 de las 145 accesiones de *Arabidopsis thaliana* y *Arabidopsis lyrata* se realizó mediante el método “*Neighbour joining*” utilizando los programas DNASTAR v.15 (DNASTAR Inc, Madison Wi. EE.UU.) y Mega 6 (Tamura *et al.*, 2013).

# RESULTADOS

## **1. Análisis de asociación genómica para el patrón de tricomas en Arabidopsis**

Con el fin de determinar las bases genéticas de la variación natural para el patrón de tricomas en Arabidopsis, mediante análisis de asociación genómica (GWAS), se ha estudiado el patrón de tricomas de 174 accesiones tanto en fase vegetativa como reproductiva. Cada una de las accesiones procede de una población natural distinta de la península ibérica (Figura 5A; ver Materiales y Métodos, apartado 1.1 y 2.3) y se han analizado previamente como parte del proyecto de los 1.001 genomas de Arabidopsis (The 1001 Genomes Consortium, 2016).

En la fase vegetativa se midió la superficie de la hoja (SH), el número de tricomas en la hoja (NTH) y a partir de estos se estimó la densidad de tricomas en la hoja (DTH). En la fase reproductiva se determinó el patrón de tricomas en el tallo (PTT), el patrón de tricomas en el pedicelo (PTP) y el patrón y el número de tricomas en el fruto (PTF y NTF, respectivamente).

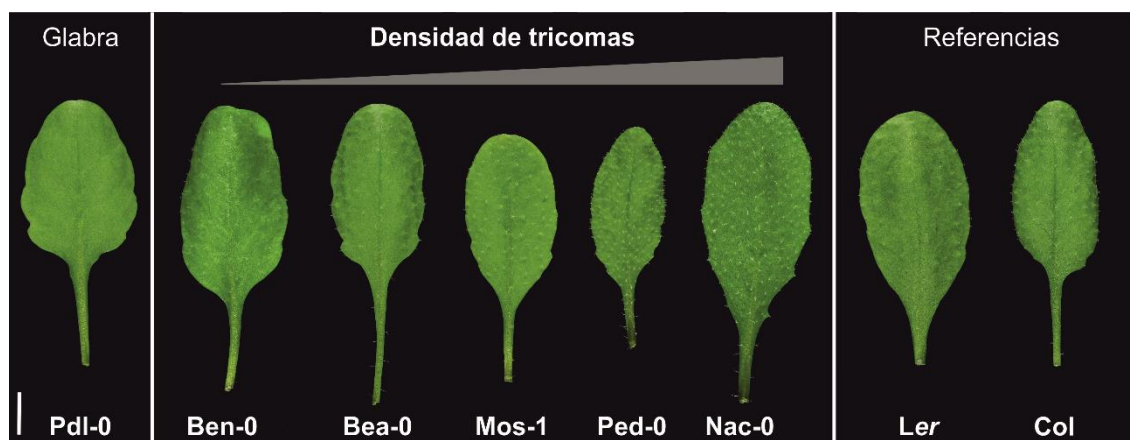
Además, para estimar la variación natural del patrón de tricomas en respuesta al ataque de herbívoros, las 174 accesiones fueron tratadas con ácido jasmónico (AJ). Este tratamiento se ha utilizado en estudios previos (Traw & Bergelson, 2003; Atwell *et al.*, 2010) para emular las respuestas de Arabidopsis causadas por la herbivoría, ya que el daño provocado por insectos herbívoros o por patógenos necrotróficos produce un incremento del ácido jasmónico que desencadena la formación de tricomas que actúan como barrera física (Koo & Howe, 2009; Pieterse *et al.*, 2012). Para ello, las primeras hojas de las 174 accesiones fueron tratadas con ácido jasmónico (0,15 mM; ver Materiales y Métodos, apartado 2.3) y posteriormente se cuantificaron los tres caracteres relacionados con el patrón de tricomas de fase vegetativa: SHAJ, NTHAJ y DTHAJ.

### **1.1. Variación genética para el patrón de tricomas en Arabidopsis y su respuesta al ácido jasmónico**

#### **1.1.1. Variación natural para la densidad de tricomas en fase vegetativa**

Los tres caracteres analizados en la fase vegetativa, SH, NTH y DTH, mostraron una gran variación cuantitativa en la península ibérica (Figura 9), con unos rangos de variación de 5-10 órdenes de magnitud entre las accesiones con fenotipos extremos (Figura 10). En general, las distribuciones de frecuencia se ajustaron a una distribución

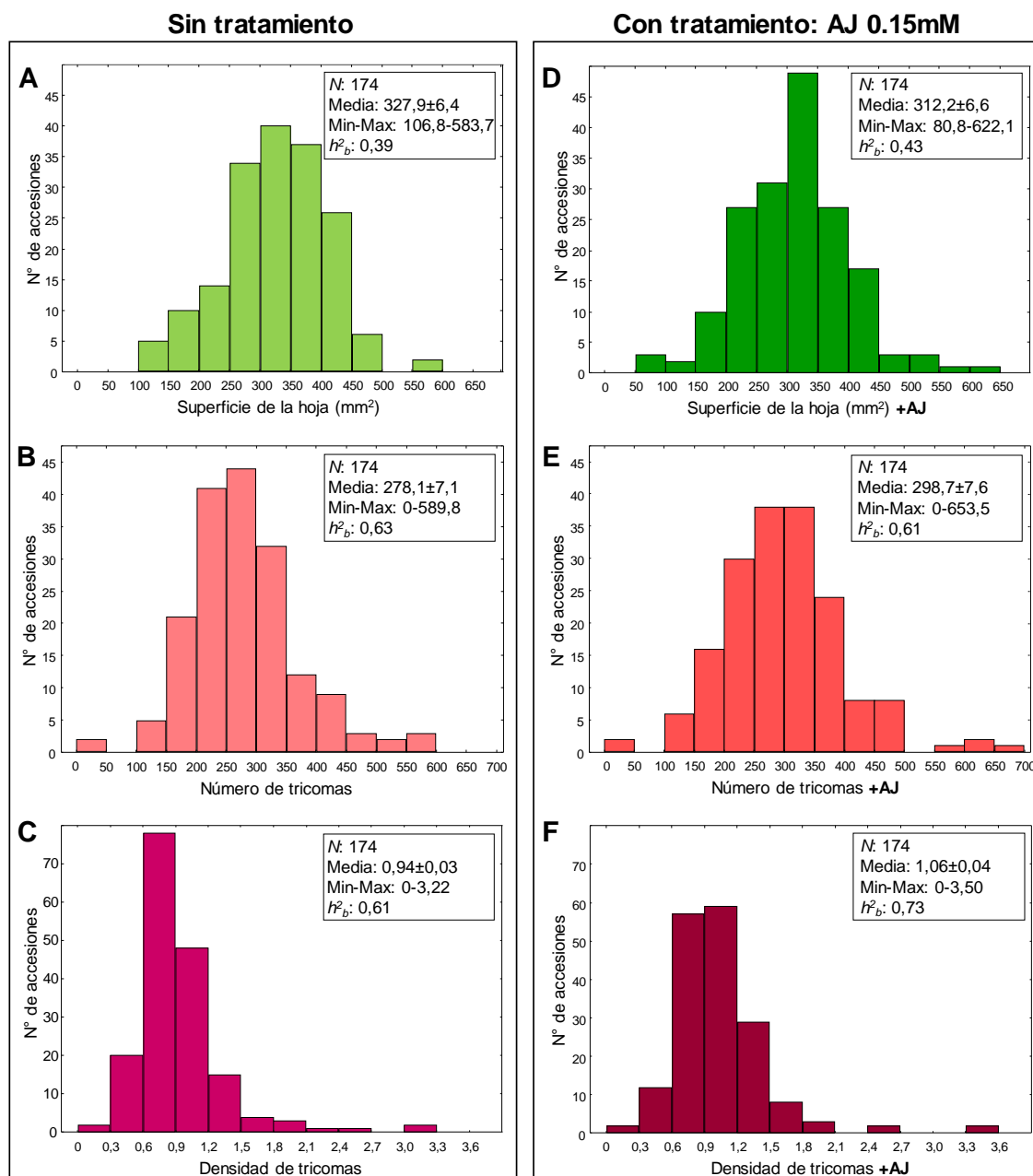
normal, aunque la densidad de tricomas mostró una cierta asimetría positiva debido a las cinco accesiones con valores de DTH más altos (Figura 10C).



**Figura 9. Variación fenotípica para la densidad de tricomas de las hojas en Arabidopsis.** La figura muestra hojas representativas de accesiones de la península ibérica que presentan diferente densidad de tricomas incluyendo una accesión glabra, cinco accesiones que varían de baja a alta densidad de tricomas y las accesiones de referencia Ler y Col. En la esquina inferior izquierda se representa una escala de 0,5 cm.

Los tres caracteres vegetativos mostraron heredabilidades moderadas (Figura 10A-10C), aunque estas fueron sustancialmente mayores para el número y la densidad de tricomas (0,61-0,63) que para la superficie de la hoja (0,39). Este resultado indica que alrededor del 60% de la variación fenotípica de los caracteres de tricomas en la hoja está determinada genéticamente.

Para determinar la relación entre los tres caracteres vegetativos se estimaron los coeficientes de correlación de Pearson entre todos los de pares de caracteres (Tabla 9). Como cabría esperar, la densidad de tricomas correlaciona negativamente con la superficie de la hoja ( $r=-0,47$ ;  $P<0,01$ ) y positivamente con el número de tricomas ( $r=0,55$ ;  $P<0,01$ ) (Tabla 9). Sin embargo, la superficie de la hoja también mostró una correlación moderada pero significativa con el número de tricomas ( $r=0,37$ ;  $P<0,01$ ). Estos resultados indican que cuanto mayor es la hoja, mayor es su número de tricomas, reduciendo su efecto sobre la densidad de tricomas. No obstante, los mayores valores de correlación observados entre la densidad de tricomas y el número de tricomas o el tamaño de la hoja indican que parte del control genético del número de tricomas y de la densidad de tricomas es independiente del tamaño de la hoja.



**Figura 10. Distribución de frecuencias de los caracteres relacionados con el patrón de tricomas en fase vegetativa.** A, B y C) Distribuciones de frecuencias de la superficie de la hoja (A), el número de tricomas en la hoja (B) y la densidad de tricomas (C) en plantas no tratadas con ácido jasmónico (AJ). D, E, y F) Distribuciones de frecuencias de la superficie de la hoja (D), el número de tricomas en la hoja (E) y la densidad de tricomas (F) en plantas tratada con AJ. Dentro de cada panel se indica el número de accesiones ( $N$ ), la media  $\pm$  EE, los valores mínimo y máximo y la heredabilidad en sentido amplio estimada ( $h^2_b$ ).

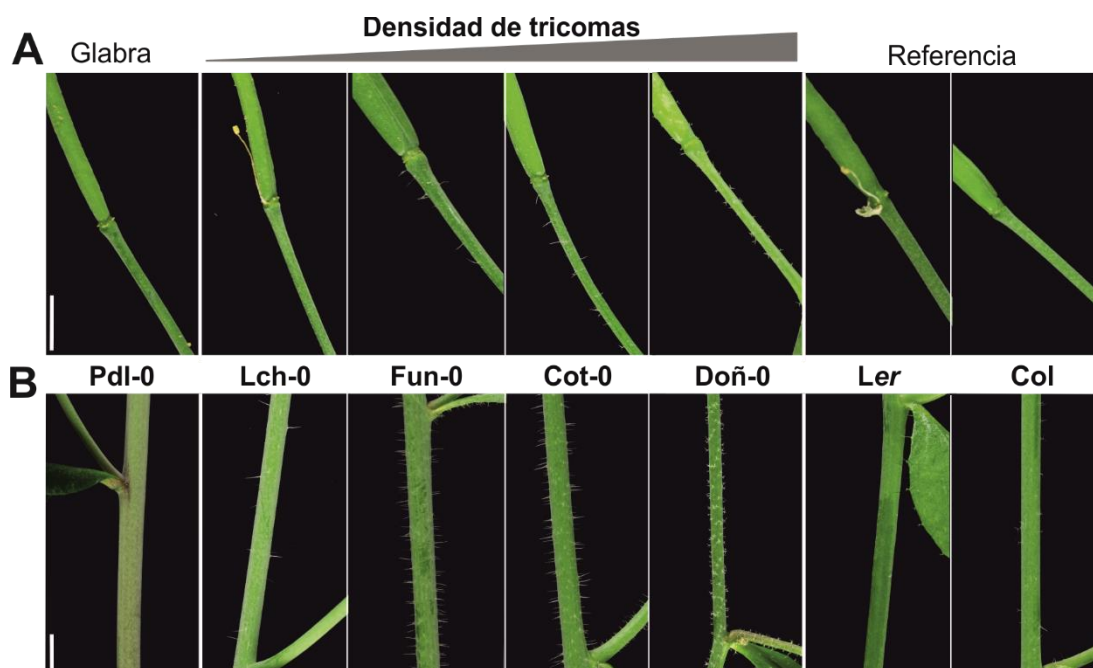
	SH	NTH	DTH	SHAJ	NTHAJ	DTHAJ	PSH	PNTH	PDTH
SH		**	**	**	**	**	*	ns	*
NTH	<b>0,37</b>		**	**	**	**	ns	*	ns
DTH	<b>-0,47</b>	<b>0,55</b>		**	**	**	ns	ns	**
SHAJ	<b>0,80</b>	<b>0,30</b>	<b>-0,38</b>		**	**	**	**	**
NTHAJ	<b>0,34</b>	<b>0,91</b>	<b>0,51</b>	<b>0,41</b>		**	ns	**	ns
DTHAJ	<b>-0,42</b>	<b>0,51</b>	<b>0,91</b>	<b>-0,48</b>	<b>0,51</b>		*	ns	ns
PSH	<b>-0,18</b>	-0,03	0,10	<b>0,42</b>	0,18	<b>-0,15</b>		**	**
PNTH	-0,02	<b>-0,18</b>	-0,12	<b>0,27</b>	<b>0,22</b>	-0,01	<b>0,51</b>		**
PDTH	<b>0,17</b>	-0,14	<b>-0,27</b>	<b>-0,21</b>	-0,01	0,13	<b>-0,60</b>	<b>0,33</b>	

**Tabla 9. Correlaciones entre los caracteres del patrón de tricomas en fase vegetativa.** En la diagonal inferior de la tabla se muestran los coeficientes de correlación de Pearson ( $r$ ) entre todos los pares de caracteres relacionados con el patrón de tricomas en fase vegetativa, estimados sin y con tratamiento de ácido jasmónico. Los coeficientes de correlación significativos se muestran en negrita. En la diagonal superior se indica la significación estadística (\*\*:  $P < 0,01$ ; \*:  $P < 0,05$ ; ns:  $P > 0,05$ ). SH: superficie de la hoja ( $\text{mm}^2$ ); NTH: número de tricomas por hoja; DTH: densidad de tricomas en la hoja; SHAJ: superficie de la hoja tratada con AJ; NTHAJ: número de tricomas de hojas tratadas con AJ; DTHAJ: densidad de tricomas de hojas tratadas con AJ; PSH: plasticidad de la superficie de la hoja al tratamiento con AJ; PNTH: plasticidad del número de tricomas por hoja al tratamiento con AJ; PDTH: plasticidad de la densidad de tricomas al tratamiento con AJ.

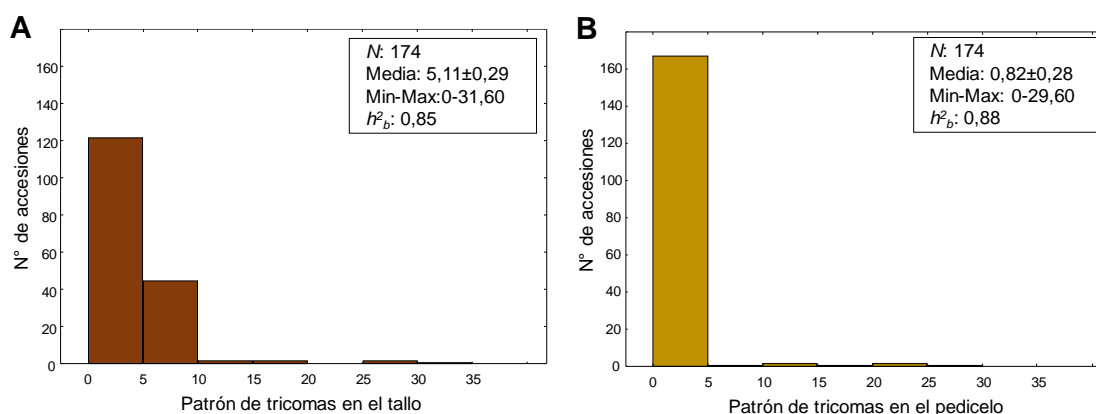
### 1.1.2. Variación natural para el patrón de tricomas en fase reproductiva

Al igual que en la fase vegetativa, los patrones de tricomas en el tallo (PTT) y en el pedicelo (PTP) mostraron una amplia variación fenotípica en la colección de 174 accesiones de la península ibérica (Figura 11). Sin embargo, el 60 y el 90% de las accesiones presentaron valores bajos de PTT y de PTP, respectivamente, indicando que la diversidad genética en fase reproductiva es inferior a la observada en fase vegetativa. Esta menor diversidad dio lugar a distribuciones de frecuencias con una elevada asimetría ya que solo una proporción pequeña o moderada presentó valores elevados de PTT o PTP (Figura 12). No obstante, los patrones de tricomas en fase reproductiva mostraron heredabilidades muy altas (0,85-0,89) (Figura 12), indicando que la mayor parte de la variación fenotípica en ambos caracteres está determinada genéticamente.

La reducción de diversidad genética fue especialmente extrema para los dos caracteres del fruto, PTF y NTF, ya que sólo cinco accesiones (2,9%) presentaron tricomas en los frutos. Por ello, el patrón y el número de tricomas en el fruto no se incluyeron en los análisis genómicos posteriores.



**Figura 11. Variación fenotípica para el patrón de tricomas en fase reproductiva en *Arabidopsis*.** La figura muestra pedicelos (A) y tallos (B) representativos de accesiones procedentes de la península ibérica con distinto patrón de tricomas incluyendo una accesión glabra, cuatro accesiones que varían de baja a alta densidad de tricomas en el pedicelo (A) y en el tallo (B) y las accesiones de referencia Ler y Col. En la esquina inferior izquierda de cada panel se representa una escala de 0,5 cm.



**Figura 12. Distribución de frecuencias del patrón de tricomas en fase reproductiva.** A y B) Distribuciones de frecuencias del patrón de tricomas en el tallo (A) y en pedicelo (B). Dentro de cada panel se indica el número de accesiones (N), la media ± EE, los valores mínimo y máximo y la heredabilidad en sentido amplio estimada ( $h^2_b$ ).

El análisis de correlación entre los patrones de tricomas en el tallo y el pedicelo (Tabla 10) mostró un coeficiente de Pearson muy elevado ( $r= 0,89$ ;  $P<0,01$ ), lo que indica que ambos caracteres comparten la mayor parte de sus bases genéticas. Por el contrario, la mayoría de los coeficientes de correlación entre los patrones de tricomas en fase reproductiva y los caracteres de la fase vegetativa no fueron significativos



( $P>0,05$ ). Solamente se detectó una baja correlación entre el patrón de tricomas en pedicelo y la superficie de la hoja ( $r=-0,19$ ;  $P<0,05$ ). Por lo tanto, los patrones de tricomas en fase vegetativa y fase reproductiva presentan bases genéticas prácticamente independientes.

	SH	NTH	DTH	PTT	PTP
SH		**	**	ns	*
NTH	<b>0,35</b>		**	ns	ns
DTH	<b>-0,46</b>	<b>0,59</b>		ns	ns
PTT	-0,11	0,02	0,13		**
PTP	<b>-0,19</b>	-0,06	0,14	<b>0,89</b>	

**Tabla 10. Correlación entre el patrón de tricomas en fase vegetativa y reproductiva.** En la diagonal inferior de la tabla se muestran los coeficientes de correlación de Pearson ( $r$ ) entre los caracteres relacionados con el patrón de tricomas en fase vegetativa (SH, NTH y DTH) y fase reproductiva (PTT y PTP). Los coeficientes de correlación significativos se muestran en negrita. En la diagonal superior se indica la significación estadística (\*\*:  $P<0,01$ ; \*:  $P<0,05$ ; ns:  $P>0,05$ ). SH: superficie de la hoja ( $\text{mm}^2$ ); NTH: número de tricomas por hoja; DTH: densidad de tricomas en la hoja; PTT: patrón de tricomas en el tallo; PTP: patrón de tricomas en el pedicelo.

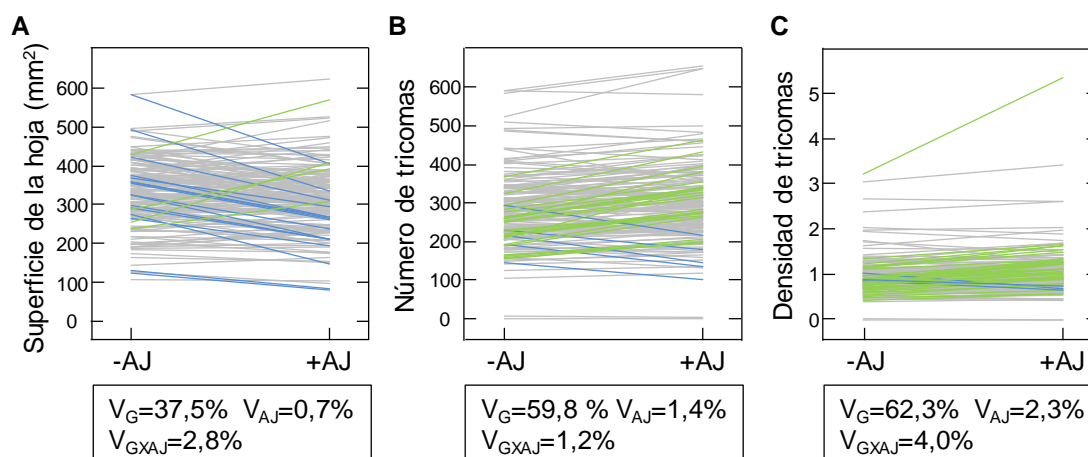
### 1.1.3. Variación natural para el patrón de tricomas vegetativo en plantas tratadas con ácido jasmónico

Las distribuciones de frecuencias de los tres caracteres estimados tras el tratamiento con AJ fueron muy similares a las obtenidas sin dicho tratamiento. Como en el caso anterior, los tres caracteres mostraron una gran cantidad de variación cuantitativa, con una diferencia de 10 a 12 órdenes de magnitud entre los valores extremos de cada carácter (Figura 10D-10F). Igualmente, los tres caracteres mostraron una heredabilidad moderada o alta. Especialmente el carácter DTHAJ presentó la mayor heredabilidad (0,73), siendo ligeramente superior a la estimada para el mismo carácter sin el tratamiento de AJ. Este resultado indica que más del 70% de la variación fenotípica de este carácter está determinada genéticamente.

Los análisis de correlación entre los caracteres obtenidos tras el tratamiento con AJ (Tabla 9) mostraron valores muy similares a los estimados sin dicho tratamiento. Básicamente, el carácter DTHAJ presentó una correlación positiva con NTHAJ ( $r=0,51$ ;  $P<0,01$ ) y negativa con SHAJ ( $r=-0,48$ ;  $P<0,01$ ), mientras que SHAJ correlacionó positivamente con NTHAJ ( $r=0,41$ ;  $P<0,01$ ). Por lo tanto, como en ausencia del tratamiento con AJ, los tres caracteres presentan un control genético parcialmente independiente.

### 1.1.4. Respuesta del patrón de tricomas vegetativo al ácido jasmónico

Para determinar la respuesta del patrón de tricomas al tratamiento de AJ se estimó la norma de reacción de cada accesión (Figura 13) y se comparó estadísticamente su fenotipo en las dos condiciones ambientales. Este análisis muestra que un 11,5, 17,2 y 24,4% de las accesiones presentaron un fenotipo distinto para SH, NTH y DTH, respectivamente, en ambas condiciones ( $P < 0,05$ ). Además, el análisis de las normas de reacción de estas accesiones indica que el 80, 83,3 y el 92,3% de las accesiones que respondieron significativamente al AJ lo hicieron en el mismo sentido, reduciendo SH y aumentando NTH y DTH, respectivamente (Figura 13). Estos resultados indican que gran parte de las accesiones se comportaron de forma similar con y sin tratamiento de AJ. En concordancia, las tres variables mostraron coeficientes de correlación de Pearson muy elevados ( $r = 0,80-0,91$ ;  $P < 0,01$ ) entre ambos ambientes (Tabla 9).

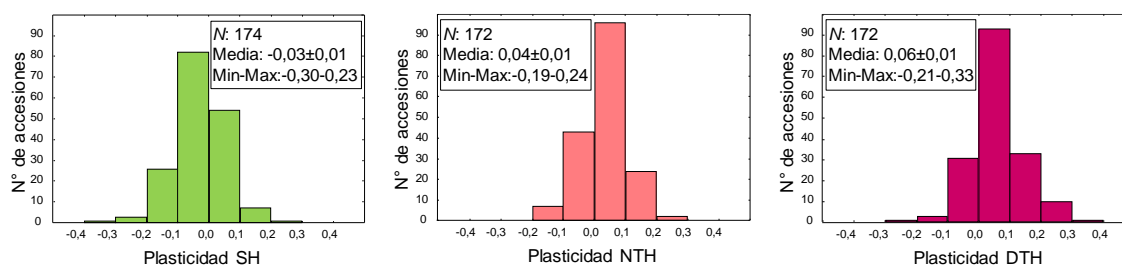


**Figura 13. Normas de reacción de los caracteres del patrón de tricomas vegetativo en relación al tratamiento con ácido jasmónico.** A, B y C) Normas de reacción para la superficie de la hoja (A), el número de tricomas en la hoja (B) y la densidad de tricomas en la hoja (C). Los colores de las líneas representan el comportamiento diferencial de las accesiones: accesiones que no difieren estadísticamente ( $P > 0,05$ ) entre tratamientos (gris); accesiones que aumentan (verde) o reducen (azul) su valor significativamente en el tratamiento con AJ ( $P < 0,05$  en el test de  $t$  de Student). En la parte inferior de cada panel se muestran los porcentajes de varianza explicada por los genotipos ( $V_G$ ), por el tratamiento con AJ ( $V_{AJ}$ ) y por la interacción entre genotipo y tratamiento de AJ ( $V_{G \times AJ}$ ).

Para cuantificar la proporción de la variación total observada para cada uno de los caracteres que se debe a las diferencias genéticas entre accesiones ( $V_G$ ), al efecto del tratamiento con AJ ( $V_{AJ}$ ), o a la interacción entre ambos factores ( $V_{G \times AJ}$ ), se llevó a cabo un análisis de componentes de varianza a partir del conjunto de datos generados en ambos ambientes (Figura 13). De esta forma se estimó que el 37,5, 59,8 y el 62,3% de la variación observada para SH, NTH y DTH, respectivamente, viene determinada por las diferencias genéticas entre las accesiones. Igualmente, este análisis indica que

sólo entre el 0,7 y el 2,3 % de las diferencias son debidas al tratamiento con AJ. Por último, la interacción entre el genotipo y el ambiente (tratamiento con AJ) explicó el 2,8, 1,2 y el 4% de la variación para SH, NTH y DTH, respectivamente. La varianza explicada por la interacción entre ambos factores corresponde a la variación genética para la respuesta o plasticidad al AJ, es decir, a la variación genética que depende del tratamiento. Por lo tanto, aunque la mayor parte de la variación observada se debe a las diferencias genéticas entre las accesiones que se expresan independientemente del tratamiento con AJ, existe una pequeña, pero significativa, proporción de la variación para el patrón de tricomas en fase vegetativa que está determinada por la variación genética para la propia respuesta al AJ.

Por último, la respuesta o plasticidad de cada accesión al tratamiento con AJ se calculó mediante un índice de plasticidad relativa (ver Materiales y Métodos, apartado 3) que varía entre 0 y -1 o 1, en función de la menor o mayor respuesta y su sentido (Figura 14). En este análisis, las dos accesiones completamente glabras (Hue-3 y Pdl-0) fueron descartadas. En concordancia con los análisis anteriores, el 78, 80 y 71% de las accesiones mostraron un índice de plasticidad muy bajo (entre -0,1 y 0,1) para SH, NTH y DTH, respectivamente (Figura 14). Por lo tanto, el patrón de tricomas de la hoja presentó una plasticidad general al tratamiento con AJ relativamente baja, ya que sólo entre el 20 y el 29% de las accesiones mostraron una respuesta sustancial a este tratamiento.



**Figura 14. Distribución de frecuencias de la plasticidad al tratamiento de ácido jasmónico en fase vegetativa.** A, B y C) Índices de plasticidad al ácido jasmónico para la superficie de la hoja (SH) (A), el número de tricomas en la hoja (NTH) (B) y la densidad de tricomas en la hoja (DTH) (C). Dentro de cada panel se indica el número de accesiones (*N*), la media ± EE y los valores mínimo y máximo.

## 1.2. Bases genómicas de la variación natural para el patrón de tricomas y su respuesta al ácido jasmónico en *Arabidopsis*

Con el fin de determinar las bases genómicas de la variación natural para el patrón de tricomas de *Arabidopsis*, se ha realizado un análisis de asociación genómica (GWAS) de cada una de las variables cuantificadas para la densidad de tricomas en fase

vegetativa (DTH, DTHAJ y su plasticidad medida como PDTH) y en fase reproductiva (PTT y PTP). Estos análisis se han llevado a cabo utilizando 1,88 millones de SNPs que han sido previamente genotipados mediante el análisis de la secuencia genómica de las 174 accesiones de la península ibérica (Tabas-Madrid *et al.*, 2018). La asociación de cada SNP con cada uno de los caracteres fenotípicos fue contrastada estadísticamente utilizando el modelo lineal mixto (MLM) implementado en el programa TASSEL v.5 (ver Materiales y Métodos, apartado 8).

Previo al análisis de asociación genómica, se estimó la proporción de la variación para el patrón de tricomas que está asociada con la estructura genética, es decir, con las relaciones genéticas entre las 174 accesiones. Para ello se cuantificó la varianza de cada variable que es explicada por la matriz de parentesco existente entre todas las accesiones (Tabla 11). Este análisis muestra que una proporción moderada o baja, de entre el 28% y el 3% de la variación para los caracteres del patrón de tricomas en fase vegetativa (DTH, DTHAJ y PDTH) está determinada por la estructura genética. Por el contrario, la estructura explicó una proporción mucho más alta de la varianza para el patrón de tricomas en tallo y pedicelo (52, 69% para PTT y PTP, respectivamente). Estos resultados indican que la variación natural para el patrón de tricomas en la fase reproductiva está más asociada con la historia demográfica de *Arabidopsis* que la variación de tricomas en la fase vegetativa.

### **1.2.1. Análisis de asociación genómica para la densidad de tricomas en fase vegetativa y su respuesta al ácido jasmónico**

#### *1.2.1.1. Asociación genómica para la densidad de tricomas en la hoja*

Para identificar regiones del genoma que puedan afectar al patrón de tricomas vegetativo se realizó un análisis de asociación genómica de la densidad de tricomas medida en condiciones normales (DTH). Este análisis detectó un total de 326 SNPs asociados por encima del umbral estadístico de  $-\log(P)=4$ , y localizados en 247 genes distribuidos a través de los cinco cromosomas de *Arabidopsis* (Tabla 11; Figura 15A). Con el fin de seleccionar los genes candidatos más relevantes, en primer lugar se analizaron en detalle las 10 regiones que presentaron una mayor asociación estadística (Figura 15A; Tabla Suplementaria 2) y que en conjunto explicaron el 62% de la variación para DTH (Tabla 11). Aunque ningún SNP presentó una asociación por encima de  $-\log(P)=7,6$  (correspondiente al 5% de la corrección de Bonferroni), dos regiones localizadas en el cromosoma 2 mostraron valores próximos de  $-\log(P)>7$ . Una de estas regiones destaca por tener la mayor significación estadística del análisis ( $-\log(P)=7,2$ ) y

corresponde al gen *At2g04925* que codifica para una proteína tipo defensina de función desconocida. La otra región incluyó un 22% del total de SNPs asociados (73 SNPs) y los valores de asociación máximos se detectaron en un clúster de cuatro genes implicados en el desarrollo de tricomas que están estrechamente ligados en una región genómica de 10 kb. En concreto, el gen *ETC2* (*At2g30420*) presentó la mayor asociación, incluyendo tres cambios de aminoácido asociados significativamente con DTH. Interesantemente, uno de estos SNPs es el polimorfismo Lys/Glu<sup>19</sup> de este factor de transcripción de la familia MYB que se ha demostrado previamente que contribuye a la variación natural para la densidad de tricomas en las hojas (Hauser *et al.*, 2009), por lo que esta asociación indica que *ETC2* es uno de los genes causantes de la variación para DTH en la península ibérica. El alelo menor de este polimorfismo aparece distribuido por toda la península ya que está presente en el 40% de las accesiones analizadas (Figura 16A). Además, varios SNPs de los genes *KIS* (*At2g30410*) y *TCL2* (*At2g30424*) mostraron asociaciones altamente significativas ( $-\log(P)>6$ ), mientras que el gen *TCL1* (*At2g30432*) presentó una asociación inferior ( $-\log(P)=4,3$ ). Por lo tanto, otros genes del clúster del cromosoma 2 también podrían contribuir a la variación para DTH. Por último, entre las diez regiones más fuertemente asociadas cabe destacar el gen *At4g04985* ya que un SNP que elimina el codón de inicio de este gen y dos cambios de aminoácido presentaron asociación con DTH (Figura 15A; Tabla Suplementaria 2). El primero de estos SNPs es de esperar que tenga un alto impacto sobre la función del gen y su alelo menor, presente en el 21% de las accesiones, aparece distribuido principalmente en el noroeste de la península ibérica (Figura 16B). Este gen codifica para una proteasa hipotética con función desconocida que no se ha implicado en el desarrollo de tricomas.

Carácter fenotípico	$h^2_b$	$h^2_{parentesco}$	$h^2_{10\ loci\ max}$	Nº SNPs $-\log(P)>4$	Nº Genes $-\log(P)>4$	Nº SNPs en genes de tricomas	Nº Genes de tricomas
DTH	0,61	0,25	0,62	326	247	31	8
DTHAJ	0,73	0,28	0,53	182	129	35	6
PDTH		0,03	0,38	38	49	0	0
PTT	0,85	0,52	0,39	650	419	18	9
PTP	0,88	0,69	0,43	3638/291*	1893/243*	26	9

**Tabla 11. Heredabilidades, número de SNPs y genes asociados con las variables del patrón de tricomas.** Para cada variable analizada se muestra la heredabilidad en sentido amplio ( $h^2_b$ ), la heredabilidad explicada por las relaciones genéticas de parentesco entre las accesiones ( $h^2_{parentesco}$ ), la heredabilidad explicada por los 10 loci más significativos ( $h^2_{10\ loci\ max}$ ), el número total de SNPs y los genes asociados por encima del umbral de significación estadística de  $-\log(P)=4$ , así como el número de SNPs y genes conocidos del desarrollo de tricomas que aparecen asociados. \*: Debido al exceso de significaciones obtenidas para PTP también se indica el número de SNPs y genes detectados con un umbral de  $-\log(P)=6$ . DTH: densidad de tricomas en la hoja; DTHAJ: densidad de tricomas en la hoja de plantas tratadas con AJ; PDTH:

plasticidad de la densidad de tricomas en la hoja al AJ; PTT: patrón de tricomas en el tallo; PTP: patrón de tricomas en el pedicelo.

En segundo lugar, para identificar genes previamente implicados en el desarrollo de tricomas de *Arabidopsis* que podrían contribuir a la variación natural para DTH, se realizó una búsqueda de SNPs localizados en 10 kb por encima y por debajo del conjunto de los 115 genes descritos en este proceso (Tabla Suplementaria 1). De esta forma se identificaron 8 genes candidatos (Tabla Suplementaria 2), incluyendo los cuatro genes del clúster del cromosoma 2, así como los genes *ZFP8* (*At2g41940*), *PFL2* (*At4g00100*), *LOM3* (*At4g00150*) y *MYC1* (*At4g00480*). En particular, el gen *MYC1* también ha sido previamente descrito como otro de los genes que contribuyen a la variación natural de *Arabidopsis* a través de una mutación que produce el cambio de aminoácido Pro/Ala<sup>189</sup> (Symonds *et al.*, 2011). Sin embargo, este polimorfismo no segregó entre las 174 accesiones analizadas y sólo se encontró asociación significativa con SNPs de la región intergénica adyacente.

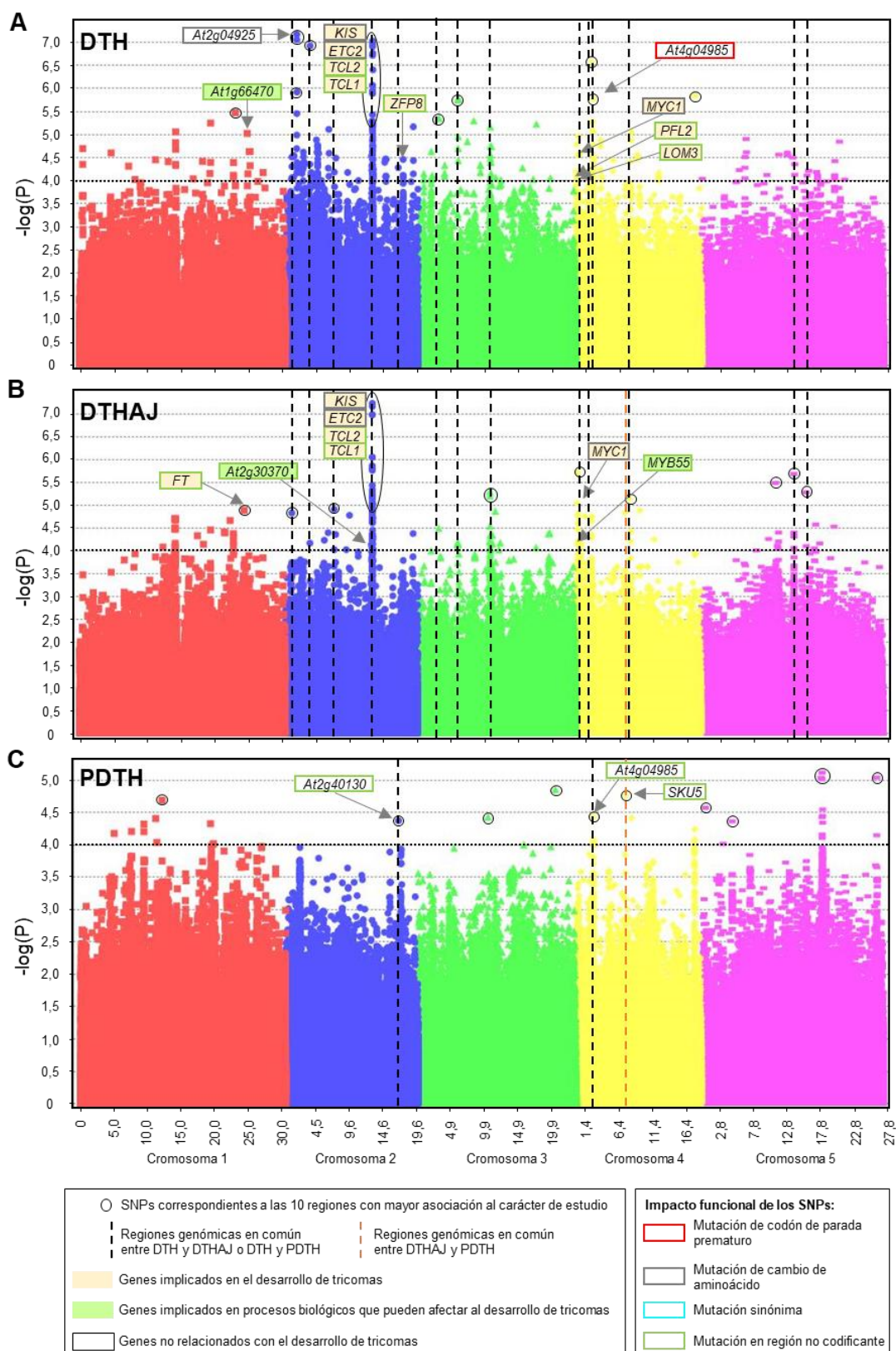
Por último, para identificar otros genes candidatos que podrían contribuir a la variación natural del carácter DTH también se analizó la función de los de 247 genes asociados, buscando aquellos que estén implicados en procesos que han sido indirectamente relacionados con el patrón de tricomas. Este análisis identificó el gen *ROOT HAIR DEFECTIVE6* (*RHD6*; *At1g66470*) ( $-\log(P)=5,02$ ) que codifica para un factor de transcripción con un dominio bHLH implicado en el desarrollo de pelos radiculares (Masucci & Schiefelbein, 1994), proceso que comparte un gran número de genes con el desarrollo de tricomas (Ishida *et al.*, 2008).

#### *1.2.1.2. Asociación genómica para la densidad de tricomas en la hoja de plantas tratadas con ácido jasmónico*

El análisis de asociación genómica de la densidad de tricomas en plantas tratadas con AJ detectó un número de asociaciones inferior a las encontradas con DTH. En total se identificaron 182 SNPs asociados por encima del umbral de  $-\log(P)=4$ , los cuales afectaron a 129 genes (Tabla 11). La comparación de DTH y DTHAJ mostró que el 51% de los genes asociados con DTHAJ también lo están con DTH. Como en el carácter anterior, para seleccionar los genes candidatos más importantes, en primer lugar se analizaron las 10 regiones que mostraron mayor significación estadística y que juntas explicaron el 53% de la variación fenotípica (Tabla 11). Nuevamente, la región que mostró mayor asociación corresponde al clúster de cuatro genes del cromosoma 2 ya

que este pico contuvo el 35% de los SNPs asociados (Figura 15B; Tabla Suplementaria 2). Los genes *KIS* y *ETC2* presentaron las significaciones estadísticas mayores ( $-\log(P) > 7$ ), próximas pero por debajo del valor equivalente al 5% después de la corrección de Bonferroni ( $-\log(P) = 7,6$ ). Como en el caso de DTH, *TCL2* también presentó significaciones elevadas de  $-\log(P) = 6,1$ , mientras que *TCL1* se detectó con asociaciones inferiores de  $-\log(P) = 4,8$ . Entre las regiones más significativas también se identificó el gen *FT* (*At1g65480*) implicado en la regulación de la floración, pero también en el patrón de tricomas durante la transición de la fase juvenil a adulta de las hojas (Williamn & Poething, 2011). Además, otras 12 de las 20 regiones más asociadas con DTH o DTHAJ fueron detectadas en ambos análisis, lo que refuerza su contribución al patrón de tricomas en la hoja (Figura 15). No obstante, todas estas regiones comunes aparecen entre las 10 más asociadas sólo en uno de los caracteres, pero no en el otro, por lo que podrían tener una contribución diferencial sobre DTH y DTHAJ y, consecuentemente, podrían también afectar a la respuesta al AJ.

Por otra parte, se analizaron los 115 genes conocidos del desarrollo de tricomas para buscar SNPs asociados situados en sus proximidades ( $\pm 10$  kb), lo que identificó los cuatro genes del clúster del cromosoma 2 y el gen *MYC1* situado en el extremo superior del cromosoma 4. Finalmente, la inspección de la función de los 129 genes detectados identificó otros dos genes que regulan procesos relacionados con el patrón de tricomas: el factor de transcripción *MYB55* (*At4g01680*) que se expresa en las células basales de los tricomas (Schliep *et al.*, 2010); y el gen *At2g30370* que inhibe la producción de otro tipo de células específicas de la epidermis vegetal como son los estomas (Shimada *et al.*, 2011) (Tabla Suplementaria 2). *FT* y *At2g30370* presentaron una significación en DTH cercana al umbral de selección ( $-\log(P) = 3,2$  y  $3,8$  respectivamente) por lo que también podrían afectar al patrón de tricomas independientemente del AJ. Por el contrario, la región de *MYB55* no mostró ninguna significación sobre DTH ( $P = 0,08$ ) sugiriendo un papel específico sobre la respuesta al AJ.



**Figura 15. Gráficos de tipo Manhattan de los análisis de asociación genómica para el patrón de tricomas de la hoja.** A, B y C) Gráficos de los análisis de asociación para la densidad de tricomas en la hoja (DTH; A), la densidad de tricomas en la hoja de plantas tratadas con ácido jasmónico (DTHAJ; B) y la plasticidad de la densidad de tricomas en la hoja al AJ (PDTH; C). Cada gráfico muestra el valor de  $-\log(P)$  de cada SNP (eje de ordenadas) y la posición de los SNPs (eje de abscisas) a lo largo de los mapas físicos de los cinco cromosomas de Arabidopsis.



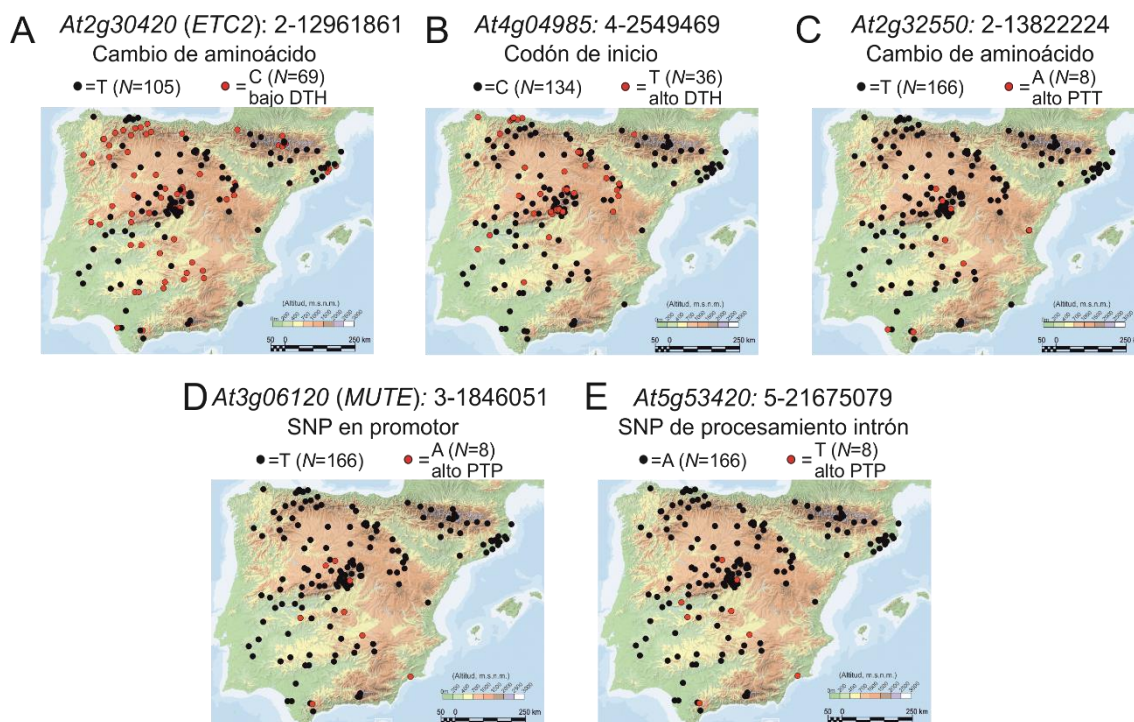
En cada panel se resaltan con círculos las 10 regiones detectadas con mayor significación estadística. Igualmente, cada panel muestra los nombres de los genes candidatos más relevantes, indicando su función biológica con distinto color: genes del desarrollo de tricomas en amarillo, genes que afectan otros procesos relacionados con el patrón de tricomas en verde; genes sin función conocida en el desarrollo de tricomas en blanco. El color del borde de la caja de los genes indica el mayor impacto funcional de los SNPs asociados. Las regiones genómicas de mayor significación detectadas en común entre DTH y DTHAJ, o entre PDTH y DTH o DTHAJ, se indican mediante líneas verticales discontinuas (ver leyenda). El nivel umbral de significación estadística ( $-\log(P)=4$ ) se muestra con la línea horizontal de puntos.

### 1.2.1.3. Asociación genómica para la respuesta del patrón de tricomas en la hoja al ácido jasmónico

Para identificar regiones genómicas que podrían afectar a la respuesta del patrón de tricomas en la hoja al ácido jasmónico se realizó un análisis de asociación genómica para la plasticidad de la densidad de tricomas (PDTH). Como cabría esperar por la escasa variación observada para este carácter (Figura 14C) sólo se detectaron 38 SNPs asociados significativamente, los cuales afectaron a 49 genes (Tabla 11; Figura 15C, Tabla Suplementaria 2). La comparación de este análisis con los resultados de DTH identificó dos regiones en común en los cromosomas 2 y 4 alrededor de los genes *At2g40130* y *At4g04985*. Igualmente, la comparación entre PDTH y DTHAJ detectó una única región del cromosoma 4 que incluye el gen *SKU5* (*At4g12420*). Estas tres regiones compartidas presentaron valores de significación entre los 10 máximos en los dos análisis comunes correspondientes, pero no fueron significativas para el tercer carácter vegetativo, lo que refuerza que estas regiones podrían estar implicadas en la respuesta del patrón de tricomas al AJ. *At2g40130* codifica para una proteína con homología a chaperoninas implicadas en termotolerancia, mientras que *SKU5* codifica para una proteína de función desconocida implicada en el crecimiento de la raíz. Los SNPs asociados en ambos genes se localizaron en las regiones intergénicas adyacentes. Por el contrario, como se describió para DTH, el gen *At4g04985* codifica para una proteasa de función desconocida que contiene un SNP de alto impacto potencial predicho y varios cambios de aminoácido asociados. A diferencia de DTH y DTHAJ, para PDTH no se detectó ningún gen conocido del desarrollo de tricomas, ni genes implicados en procesos biológicos relacionados.

En conjunto, los análisis genómicos de las tres variables DTH, DTHAJ y PDTH han identificado varios genes candidatos que podrían contribuir a la variación natural para el patrón de tricomas de la hoja y/o su respuesta al ácido jasmónico. Esta lista incluye dos genes ya demostrados como implicados en la variación natural para el patrón de tricomas, *ECT2* y *MYC1*; siete genes previamente conocidos por su función en el desarrollo de tricomas pero que se desconoce su contribución a la variación natural

(*KIS*, *TCL1*, *TCL2*, *ZFP8*, *PFL2*, *LOM3* y *FT*); varios genes implicados en otros procesos biológicos que podrían afectar indirectamente al patrón de tricomas (*MYB55*, *At2g30370* y *At1g66470*); y varios genes sin efecto previamente documentado sobre el desarrollo de tricomas (*At2g40130*, *SKU5* y *At4g04985*).



**Figura 16. Distribución geográfica de polimorfismos en algunos de los genes más relevantes detectados en los análisis de asociación genómica.** Cada mapa muestra la distribución del SNP más asociado en el gen *ETC2* (A), *At2g04985* (B), *At2g32550* (C), *MUTE* (D) y *At5g53420* (E). En cada panel se indica el nombre del gen, la posición del SNP y su efecto funcional. El alelo de menor frecuencia se muestra en color rojo y el efecto de la asociación con DTH (A, B), PTP (C) o PTP (D, E) se indica en la correspondiente leyenda. Los SNPs mostrados se seleccionaron para ilustrar un gen conocido del desarrollo de tricomas previamente implicado en la variación natural (A); dos genes implicados en procesos indirectamente relacionados con el desarrollo de tricomas y con SNPs asociados altamente (C, D); y dos genes de función desconocida, pero con SNPs de alto impacto funcional predicho y con fuerte asociación (B, E).

## 1.2.2. Análisis de asociación genómica para el patrón de tricomas en fase reproductiva

### 1.2.2.1. Asociación genómica para el patrón de tricomas en el tallo

Para identificar regiones genómicas y genes candidatos que contribuyan al patrón de tricomas en el tallo también se realizó un análisis de asociación genómica usando la variable PTT (Figura 17A). Como cabría esperar por la falta de normalidad de la distribución de este carácter (Figura 12A), este análisis detectó un mayor número de polimorfismos y genes que las variables de la fase vegetativa (Tabla 11). En total, se

identificaron 650 SNPs localizados en 419 genes (Tabla 11) de los que sólo el 3% (13) fueron comunes con los detectados para la densidad de tricomas en hoja. Este resultado concuerda con la falta de correlación observada entre las variables de tricomas en fase vegetativa y reproductiva, lo que refuerza que la variación natural para el patrón de tricomas en ambas fases tiene bases genéticas ampliamente diferentes.

Como en los análisis genómicos anteriores, se analizaron las 10 regiones genómicas más fuertemente asociadas y que en conjunto explicaron el 39% de la varianza fenotípica (Tabla 11; Tabla Suplementaria 2). Nuevamente, la región con mayor asociación se localizó en el cromosoma 2, pero los valores más significativos se detectaron en una zona situada alrededor de una megabase por debajo del pico detectado en la fase vegetativa y fuera del clúster de los cuatro genes conocidos del patrón de tricomas de esta región. Igualmente, el valor de significación estadística máximo ( $-\log(P)=7,4$ ) se detectó próximo pero por debajo del valor  $-\log(P)=7,6$ . Este valor se encontró en un SNP que produce un cambio de aminoácido en el gen *At2g32550*, el cual codifica para una proteína de función desconocida que afecta a la diferenciación celular ([www.arabidopsis.org](http://www.arabidopsis.org)). Este polimorfismo mostró un elevado desequilibrio de ligamiento con otros seis SNPs del mismo gen también asociados. El alelo menor de este polimorfismo presenta una frecuencia baja (5%) en la península ibérica y está distribuido principalmente en el centro y el sureste de la misma (Figura 16C). Además, el gen de desarrollo de tricomas *SAD2* (*At2g31660*) situado en la misma región, pero unas 360 kb por encima del anterior también apareció asociado, por lo que podría contribuir a la variación de PTT. Otra región genómica con una asociación similar a la anterior ( $-\log(P)=7,3$ ) se localizó en el cromosoma 4, sobre el gen  $\beta$ -*GAL3* (*At4g36360*) que codifica para una  $\beta$ -galactosidasa.

El análisis de los 115 genes del desarrollo de tricomas conocidos detectó nueve genes candidatos adicionales. Tres de ellos corresponden a *KIS*, *TCL2* y *TCL1* situados en el clúster del cromosoma 2. Sin embargo, a diferencia de la densidad de tricomas en hoja, *TCL1* fue el gen del clúster que presentó mayor asociación con PTT, mientras que ningún polimorfismo de *ETC2* mostró asociación significativa (Tabla Suplementaria 2). Además, el gen *SAD2* situado en el pico de máxima significación detectado, así como los genes *GL1* (*At3g27920*), *LOM2* (*At3g60630*), *TRY* (*At5g53200*), *SOC1* (*At2g4660*) y *miR156* (*At1g07867*) contienen SNPs asociados con PTT. Ninguno de estos últimos seis genes presentó asociación con las variables del patrón de tricomas de las hojas, por lo que probablemente afectan específicamente al patrón de tricomas reproductivo.

Por último, el análisis funcional de los 419 genes asociados con PTT identificó otros dos genes implicados en otros procesos biológicos que podrían afectar indirectamente al patrón de tricomas: el gen de floración *ELF3* (*At2g25930*) (Zagotta *et al.*, 1996) y el gen regulador de la diferenciación de las células estomáticas *MUTE* (*At3g06120*) (Pillitteri *et al.*, 2008) donde seis SNPs localizados en regiones no codificantes presentaron asociación significativa.

#### 1.2.2.2. Asociación genómica para el patrón de tricomas en el pedicelo

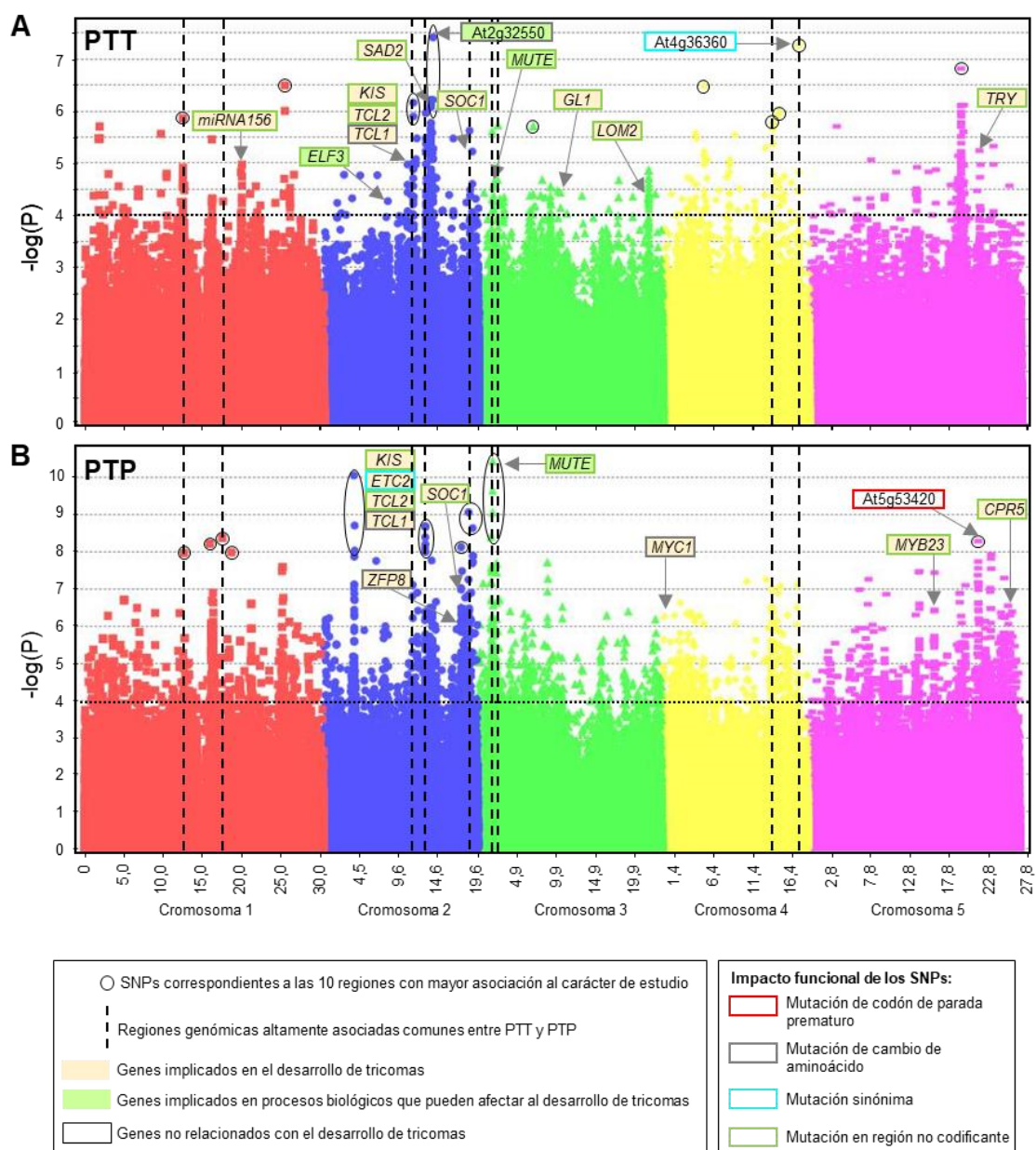
Debido a la baja proporción de accesiones con pelos en el pedicelo (Figura 12B), para realizar el análisis de asociación genómica se convirtió PTP en una variable binaria de presencia y ausencia de tricomas. Esta transformación redujo sustancialmente la proporción de SNPs asociados significativamente en comparación con la variable sin transformar, pero debido a la falta de normalidad de la distribución binaria, todavía produjo un mayor número de SNPs asociados que el esperado y que el obtenido para el resto de los caracteres analizados. En total se identificaron 3.638 SNPs asociados con un valor umbral de  $-\log(P)=4$  y localizados en 1.893 genes (Tabla 11), 40 de los cuales presentaron un umbral superior a  $-\log(P)=7,57$ . Por ello, para comparar el análisis de asociación genómica de PTP con las restantes variables, se utilizó un umbral de  $-\log(P)=6$ . Este umbral identificó un total de 291 SNPs asociados, los cuales se localizaron en 243 genes. Como en el caso de PTT, la comparación entre PTP y DTH mostró que solo 5,3% de los genes de PTP (13) son detectados en ambos análisis, apoyando el control genético diferencial entre la fase vegetativa y reproductiva. Por el contrario, la comparación entre los dos caracteres de la fase reproductiva indicó que el 25% de los genes asociados con PTP (61) también lo están con PTT. Además, 9 de las 20 regiones más asociadas con PTP y/o PTT fueron detectadas en ambos análisis (Figura 17), en concordancia con la amplia base génica común sugerida por la correlación entre ambas variables.

Las 10 regiones genómicas más asociadas explicaron el 43% de la varianza de este carácter (Tabla 11; Tabla Suplementaria 2). La región con la significación más alta ( $\log(P)=10,4$ ) identificó al gen *MUTE*, ya que ocho SNPs de su región promotora con fuerte desequilibrio de ligamiento están asociados (Figura 17B). Estos polimorfismos presentan una baja frecuencia en la península ibérica (5%) donde se encuentran distribuidos en el centro y sureste de la región (Figura 16D). Este resultado apoya el papel de *MUTE* sobre el patrón de tricomas en la fase reproductiva puesto que estos polimorfismos también fueron detectados para PTT. Además, como en los caracteres

anteriores, una de las regiones más significativamente asociada correspondió al clúster de cuatro genes conocidos del desarrollo de tricomas en el cromosoma 2. Sin embargo, sólo los genes *TCL1* y *TCL2* de este conjunto presentaron altas significaciones estadísticas ( $-\log(P) > 8$ ), mientras que *ETC2* y *KIS* mostraron asociaciones menores ( $-\log(P) < 6$ ). Por último, otra de las regiones más asociadas contiene el gen *At5g53420* donde aparecen asociados varios cambios de aminoácido y un SNP que afecta a un nucleótido de procesamiento intrónico que podría tener un alto impacto funcional, todos ellos en desequilibrio de ligamiento completo. El alelo menor de los SNPs asociados presenta una frecuencia baja (5%) y mostró también un desequilibrio de ligamiento muy alto con los polimorfismos de *MUTE*, por lo que la distribución geográfica de ambos genes es muy similar (Figura 16D y 16E). Este gen codifica para una proteína con un dominio CCT de función desconocida y también apareció asociado con PTT pero no con DTH. Por lo tanto, este gen también podría afectar a la variación natural para el patrón de tricomas en la fase reproductiva, pero este análisis no puede separar su efecto del de *MUTE*.

El análisis específico de los genes conocidos de tricomas identificó, además de los cuatro genes del clúster del cromosoma 2, los genes *MYC1*, *ZFP8*, *MYB23* (*At5g40330*), *CPR5* (*At5g64930*) y *SOC1* (Tabla Suplementaria 2). Solo el gen *SOC1* fue detectado en el análisis del tallo y sólo *ZFP8* y *MYC1* también aparecieron asociados en fase vegetativa.

En resumen, los análisis genómicos del patrón de tricomas del tallo y el pedicelo también han identificado una amplia lista de genes candidatos que podrían contribuir a la variación natural de estos caracteres. Algunos de ellos aparecen en común con la fase vegetativa, como son los genes conocidos *TCL1*, *TCL2*, *ETC2*, *KIS1*, *ZFP8* o *MYC1*. Sin embargo, ambas fases muestran un control genético ampliamente independiente ya que la mayoría de los genes identificados son distintos. Estos incluyen genes reguladores del desarrollo de tricomas (*miR156*, *SAD2*, *GL1*, *LOM2*, *TRY*, *CPR5* y *MYB23*), pero también genes que afectan otros procesos relacionados indirectamente como la floración (*ELF3*, *SOC1*), la diferenciación celular (*MUTE*, *At2g32550*) o genes sin función conocida en el desarrollo de tricomas ( $\beta$ -*GAL3*, *At5g53420*).



**Figura 17. Gráficos de tipo Manhattan de los análisis de asociación genómica para el patrón de tricomas en fase reproductiva.** A y B) Gráficos de los análisis de asociación genómica para el patrón de tricomas del tallo (PTT; A) y el patrón de tricomas del pedicelo (PTP; B). Cada gráfico muestra el valor de  $-\log(P)$  de cada SNP (eje de ordenadas) y la posición de los SNPs (eje de abscisas) a lo largo de los mapas físicos de los cinco cromosomas de *Arabidopsis*. En cada panel se resaltan con círculos las 10 regiones detectadas con mayor significación estadística. Igualmente, cada panel muestra los nombres de los genes candidatos más relevantes, indicando su función biológica con distinto color: genes del desarrollo de tricomas en amarillo, genes que afectan otros procesos relacionados con el desarrollo de tricomas en verde; genes sin función conocida en el desarrollo de tricomas en blanco. El color del borde de la caja de los genes indica el impacto funcional de los SNPs asociados. Las regiones genómicas de mayor significación detectadas en común entre PTT y PTP se indican mediante líneas verticales discontinuas. El nivel umbral de significación estadística ( $-\log(P)=4$ ) se muestra con la línea horizontal de puntos.

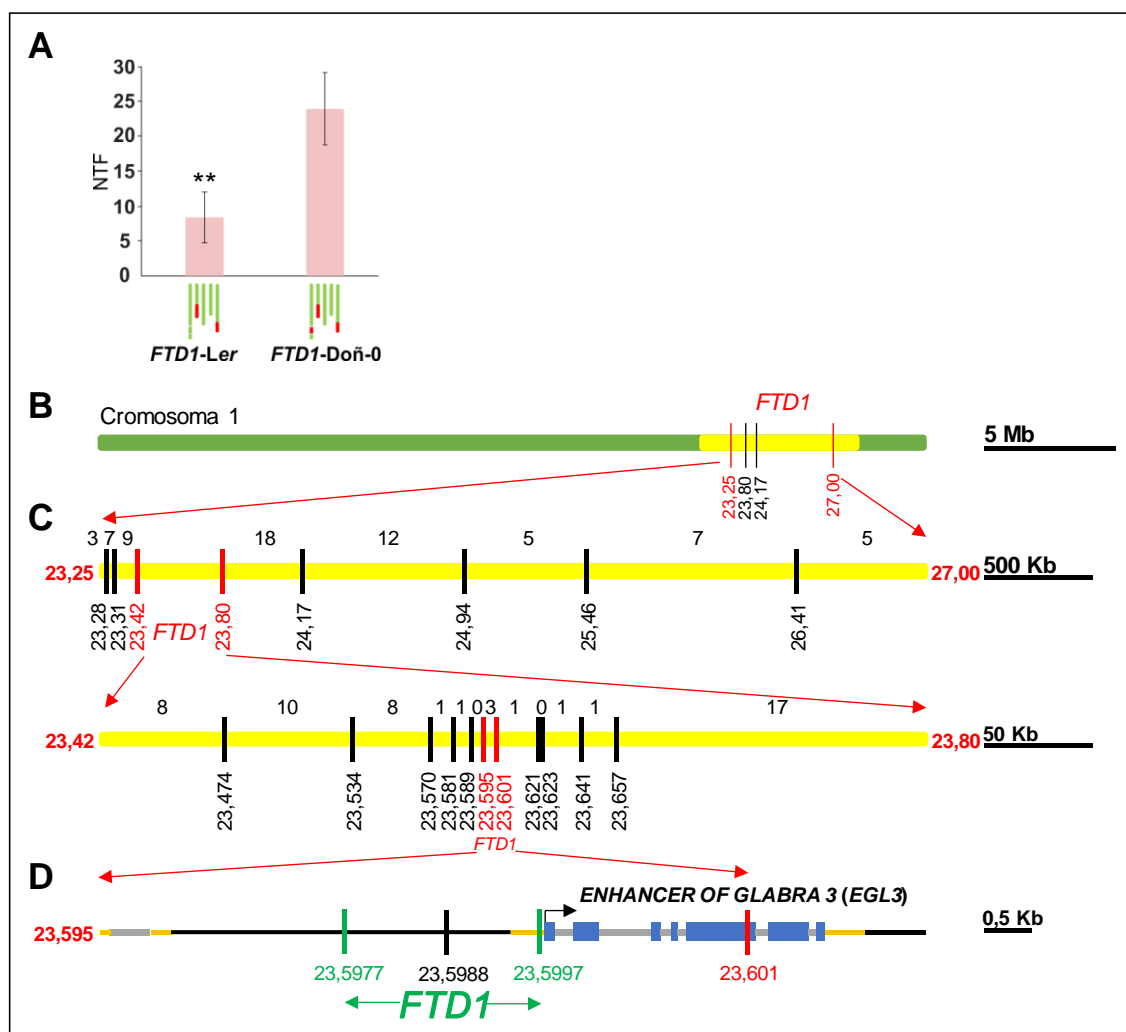
## 2. Caracterización genética y molecular de *FTD1*

*FTD1* es uno de los cinco loci identificados en la población de RILs Ler/Doñ como responsables del desarrollo de tricomas en los frutos de Arabidopsis. Estudios previos han demostrado que a pesar de que el efecto aditivo de *FTD1* es relativamente pequeño, este locus es necesario para el desarrollo de una alta densidad de tricomas en el fruto en combinación con los loci *FTD2*, *FTD3* y *FTD5* (Savic, 2015). Para determinar las bases moleculares de *FTD1*, en esta tesis se ha llevado a cabo su mapeo fino y su caracterización molecular.

### 2.1. Mapeo fino de *FTD1*

Para realizar el mapeo fino de *FTD1* se seleccionó una línea de introgresión de fondo genético Ler, homocigótica para los alelos de Doñ-0 de las regiones genómicas de *FTD2* y *FTD5*, y heterocigótica para una región alrededor de *FTD1* de ~5,5 Mb (Figura 7). A partir de esta línea se obtuvo por autofecundación una población de mapeo de 1.800 plantas que segregan para *FTD1*, las cuales se analizaron con cuatro marcadores moleculares para seleccionar las plantas recombinantes en la región de *FTD1* (Figura 18B). De esta forma se obtuvieron 110 individuos recombinantes cuya descendencia por autofecundación se analizó genotípicamente para seleccionar líneas homocigóticas para cada evento de recombinación (ver Materiales y Métodos, apartado 1.3). Los eventos de recombinación de todas las líneas homocigóticas se localizaron de forma precisa mediante análisis con un total de 22 marcadores moleculares (Tabla 3). El genotipo de cada línea para *FTD1* se determinó mediante el análisis fenotípico del número de tricomas en el fruto (NTF) ya que las líneas homocigóticas *FTD1*-Doñ-0 y *FTD1*-Ler difieren muy significativamente (Figura 18A). La comparación del genotipo de *FTD1* y los eventos de recombinación en las 110 líneas permitió localizar *FTD1* en una región de 6 kb delimitada por dos marcadores entre los que se identificaron tres plantas recombinantes (Figura 18C). Posteriormente, el evento de recombinación de cada una de estas tres líneas se localizó de forma precisa mediante secuenciación de toda esta región (ver Materiales y Métodos, apartado 5), lo que permitió reducir la posición de *FTD1* a un intervalo final de 2 kb, entre las posiciones 23,5977 y 23,5997 Mb (Figura 18D). La secuencia del genoma de Arabidopsis predice que la región de 6 kb donde se localiza *FTD1*, contiene solamente el gen *EHNANCER OF GLABRA3* (*EGL3*) (*At1g63650*), el cual codifica para un factor de transcripción de tipo bHLH previamente implicado mediante análisis de mutantes, en la regulación del patrón de tricomas en la hoja (Zhang *et al.*, 2003). Además, el intervalo final de mapeo de 2 kb se localiza en la

región promotora de *EGL3*, a 50 pb del codón de inicio de dicho gen. Estos resultados demuestran que el gen *FTD1* es *EGL3* y sugieren que el polimorfismo natural que afecta a la función de *EGL3* en Doñ-0 es un polimorfismo regulador situado en el promotor. Además, puesto que *EGL3* funciona como un regulador positivo (promotor) del desarrollo de tricomas (Zhang *et al.*, 2003) y el alelo de Doñ-0 en *FTD1* aumenta la densidad de tricomas, la accesión Doñ-0 debe de ser portadora de un alelo natural de ganancia de función en *EGL3* en comparación con el alelo silvestre de Ler.

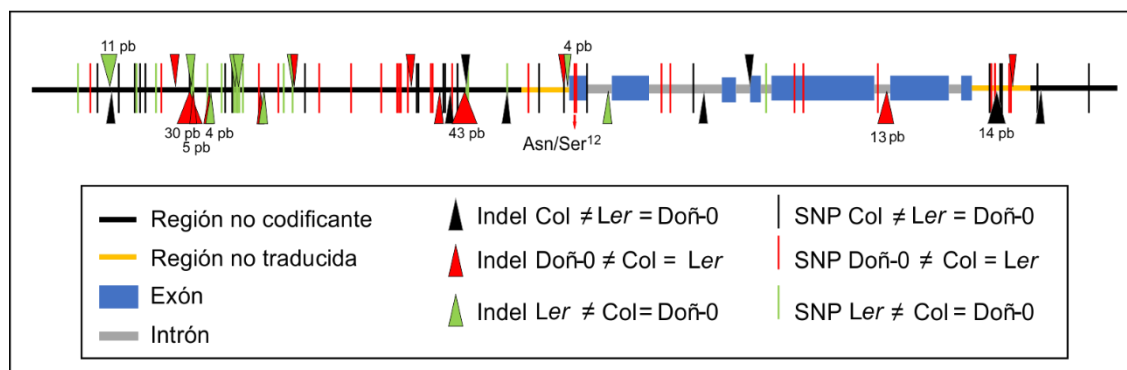


**Figura 18. Mapeo fino de *FTD1*.** A) Efecto fenotípico de *FTD1* sobre el número de tricomas en el fruto estimado en líneas homocigóticas Ler y Doñ-0 para la región original de mapeo. El genotipo gráfico de las líneas se muestra en la parte inferior del panel. La diferencia entre ambas líneas se contrastó mediante un test de *t* de Student (\*\*:  $P=0,003$ ). B) Esquema de la región del cromosoma 1 de *Arabidopsis* donde se localiza *FTD1*, mostrando en amarillo la región heterocigótica (5,5 Mb) de la planta madre usada para el mapeo fino y los cuatro marcadores utilizados inicialmente para la selección de recombinantes. C) Posición física de los 22 marcadores moleculares usados para el mapeo, y número de plantas recombinantes identificadas en cada intervalo. Las líneas verticales rojas corresponden a los marcadores que flanquean *FTD1* a lo largo del proceso de mapeo, mientras que las líneas negras representan el resto de los marcadores utilizados. D) Región genómica donde se localiza *FTD1* incluyendo la estructura del gen *EGL3* y las regiones no codificantes del gen adyacente. Las líneas verticales de color verde delimitan las posiciones físicas del intervalo final de mapeo de *FTD1*.



## 2.2. Diversidad nucleotídica y análisis filogenético de *EGL3* en la península ibérica

Con el fin de identificar polimorfismos nucleotídicos que pudieran afectar a la función del gen correspondiente al locus *FTD1*, *EGL3*, en primer lugar se secuenció este gen en las accesiones parentales *Ler* y *Doñ-0* y se comparó con la accesión de referencia *Col* disponible ([www.arabidopsis.org](http://www.arabidopsis.org)) (Figura 19). En concreto se secuenció un fragmento de 7,9 kb que incluye las 2,9 kb de la región codificante de *EGL3*, 4 kb de la región 5' no traducida (UTR) y el promotor, y 1 kb de la región 3' del gen.



**Figura 19. Comparación de la secuencia de *EGL3* en *Ler*, *Doñ-0* y *Col*.** La figura muestra la estructura del gen *EGL3* y los polimorfismos entre las accesiones *Ler*, *Doñ-0* y *Col*. Los SNPs e indels se representan con líneas verticales o triángulos respectivamente, utilizando diferentes colores en función de que el alelo minoritario sea el de *Col* (negro), *Doñ-0* (rojo) o *Ler* (verde). El tamaño de los indels mayor de 3 pb se indica junto al polimorfismo correspondiente.

El análisis comparativo de las secuencias muestra que la mayor parte de los polimorfismos entre *Ler* y *Doñ-0* se localizan en la región promotora de *EGL3*, donde encontramos 33 SNP y 16 indels (Figura 19). Entre estos polimorfismos destacan dos deleciones específicas de *Doñ-0*, de 30 y 43 pb, situadas respectivamente en posiciones -2.751 y -771 con respecto al codón de inicio. En particular, la deleción mayor se localiza dentro del intervalo de 2 kb definido por el mapeo fino de *FTD1* y corresponde a una secuencia de 43 pb repetida dos veces consecutivas en *Ler* y *Col* pero presente en una sola copia en *Doñ-0*. Además, *Ler* y *Doñ-0* difieren en 8 SNPs y 2 indels de la región codificante, aunque sólo un SNP afecta a la estructura de la proteína mediante el cambio de aminoácido Asn/Ser<sup>12</sup>. Por último, se identificaron 3 y 5 SNPs e indels polimórficos entre *Ler* y *Doñ-0* en las regiones 5' y 3' no traducidas, respectivamente (Figura 19).

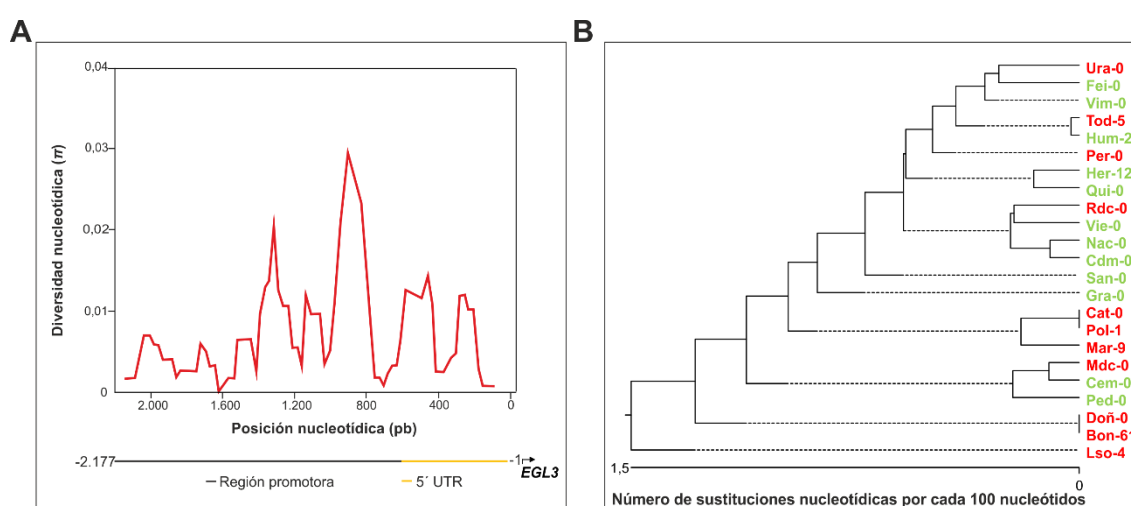
En conjunto, el mapeo fino de *FTD1* y el análisis de la secuencia nucleotídica de *EGL3* en las accesiones *Ler* y *Doñ-0* sugieren que la deleción en posición -771 podría ser el polimorfismo causal de *FTD1*. Por ello, a continuación se realizó un análisis más exhaustivo de la región de 2 kb del promotor de *EGL3* delimitada en el mapeo de *FTD1* (Tabla 12; Figura 20). Este fragmento se secuenció en 23 accesiones de la península

ibérica que difieren en el patrón de tricomas en el fruto, incluyendo 11 y 12 accesiones con y sin tricomas en el fruto respectivamente (ver Materiales y Métodos, apartado 5). En total se identificaron 36 SNPs y 20 indels informativos, y una diversidad nucleotídica media de 0,007, valor idéntico al descrito previamente para el resto del genoma (Nordborg *et al.*, 2005) (Tabla 12). Sin embargo, esta diversidad se distribuyó heterogéneamente a lo largo de este segmento ya que la región central presenta una diversidad dos veces mayor que las regiones adyacentes (Figura 20A).

Además, para determinar las relaciones genéticas entre las 23 accesiones de la península ibérica para la región promotora estudiada, se obtuvo un árbol filogenético (ver Materiales y Métodos, apartado 10). Este análisis no detectó haplogrupos altamente diferenciados, lo que indica que la diversidad nucleotídica del promotor de *EGL3* se distribuye de forma homogénea entre estas accesiones (Figura 20B).

	Longitud secuencia <sup>2</sup>	# de accesiones	# de haplogrupos	# de polimorfismos (SNPs/indels) <sup>3</sup>	# de cambios aminoacídicos	$\pi$ silenciosa	$\pi$ no sinónima
Promotor <sup>1</sup>	2.177	23	1	36/20	-	0,0070	-
Región Codificante	2.934	145	2	81/6	26	0,0145	0,0019

**Tabla 12. Diversidad nucleotídica en el promotor y la región codificante de *EGL3* en la península ibérica.** La tabla muestra los siguientes parámetros de la región promotora y la región codificante de *EGL3*: longitud de la secuencia nucleotídica analizada, número de accesiones, número de haplogrupos identificados, número de polimorfismos, número de cambios de aminoácidos y la diversidad nucleotídica ( $\pi$ ) silenciosa y no sinónima. Los datos del promotor y de la región codificante se basaron en grupos de secuencias diferentes de la península ibérica por lo que se presentan por separado. <sup>1</sup> La secuencia del promotor analizada incluye las 2 kb más próximas al codón de inicio de *EGL3*. <sup>2</sup> Longitud basada en el alineamiento de las secuencias. <sup>3</sup> No incluye los polimorfismos singleton.



**Figura 20. Diversidad nucleotídica y árbol filogenético de la región promotora de *EGL3*.** A) Diversidad nucleotídica en ventanas de 100 pb consecutivas a lo largo de 2 kb de la región promotora de *EGL3* situada más próxima a la región codificante. B). Árbol filogenético del promotor de *EGL3* de las 23 accesiones de la península ibérica, obtenido mediante el método

“*Neighbour joining*”. Las accesiones con y sin tricomas en el fruto se muestran en rojo y verde respectivamente.

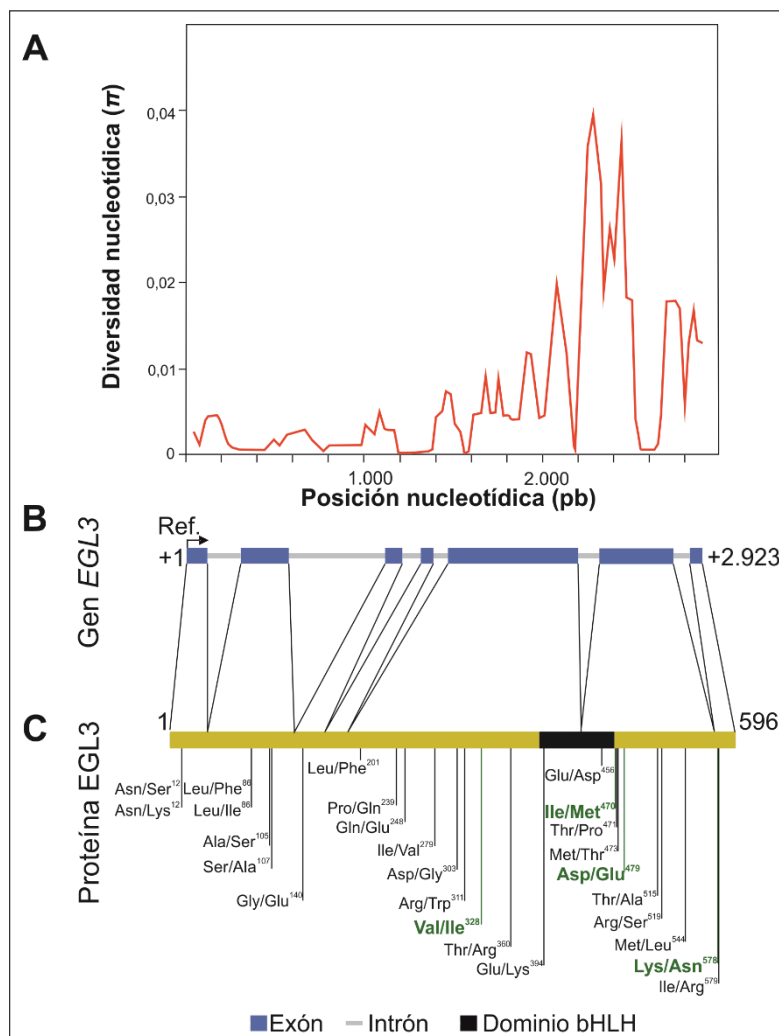
Por último, para determinar si alguno de los polimorfismos detectados en estas 23 accesiones podría afectar al patrón de tricomas en el fruto de *Arabidopsis*, se ha realizado un análisis de asociación entre estos polimorfismos y el carácter NTF (ver Materiales y Métodos, apartado 9). De esta forma se identificaron 4 SNPs situados en posiciones -1.002, -1.001, -839 y -299 con una asociación parcial significativa ( $P < 0,01$ ; Tabla 13). El alelo de Doñ-0 de estos SNPs se encontró en 3 o 6 accesiones con tricomas en el fruto, mientras que todas las accesiones sin tricomas presentaron el alelo de *Ler*. Esta asociación indica que los cuatro polimorfismos podrían contribuir a la alteración funcional de *EGL3* y ser responsables de *FTD1*. Por el contrario, la delección en posición -771, polimórfica entre *Ler* y Doñ-0, no aparece asociada en este análisis ya que el alelo de Doñ-0 de este indel está presente en accesiones con y sin tricomas en el fruto. No obstante, no se puede descartar que esta delección sea el polimorfismo causal de *FTD1* en *Ler* y Doñ-0 ya que combinaciones de distintos loci *FTD* podrían producir el mismo fenotipo de tricomas en el fruto. Así, accesiones que no poseen esta delección pueden presentar tricomas en el fruto debido a la presencia de alelos Doñ-0 en otros loci *FTD*, y accesiones con la delección en *FTD1* podrían no desarrollar tricomas en el fruto por carecer de alelos Doñ-0 en otros loci *FTD*. Además, distintos polimorfismos de *EGL3* en distintas accesiones podrían alterar de forma similar la función del gen *EGL3* (heterogeneidad alélica).

Por otra parte, aunque el mapeo fino indica que el polimorfismo funcional de *FTD1* se localiza en el promotor, también se analizó la diversidad nucleotídica de la región codificante de *EGL3* ya que al no detectarse ningún polimorfismo totalmente asociado con NTF, alguna otra mutación natural podría contribuir a la alteración funcional de *EGL3* en otras accesiones con tricomas en el fruto distintas de Doñ-0. Para ello se utilizó la secuencia de 145 accesiones de la península ibérica disponible como parte del proyecto de secuenciación de 1.001 genomas de *Arabidopsis* (The 1001 Genome Consortium, 2016). Este análisis mostró que la diversidad nucleotídica de la región codificante de *EGL3* en la península ibérica es alta, ya que el valor medio de 0,0145 en regiones no codificantes (diversidad silenciosa; Tabla 12) es superior a la media del genoma de *Arabidopsis* ( $\pi = 0,007$ ; Nordborg *et al.*, 2005). Sin embargo, esta diversidad no fue homogénea a lo largo del gen si no que se localiza principalmente en la mitad 3' (Figura 21A). Además, como cabría de esperar, la diversidad nucleotídica no sinónima ( $\pi_{\text{no sinónima}}$ ) mostró un valor inferior que la diversidad sinónima (Tabla 12). En total se identificaron 81 SNPs informativos, 26 de los cuales dan lugar cambios de

aminoácido en al menos una accesión, incluyendo el polimorfismo Asn/Ser<sup>12</sup> presente únicamente en Doñ-0 (Figura 21C).

Accesión	NTF	SNP -1.002	SNP -1.001	SNP -839	Delección (43 pb) -771	SNP -299
Doñ-0	151,1±7,1	Doñ-0	Doñ-0	Doñ-0	Doñ-0	Doñ-0
Mdc-0	96,4±6,3	Ler	Ler	Ler	Doñ-0	Ler
Cat-0	90,6±6,0	Doñ-0	Doñ-0	Doñ-0	Ler	Ler
Pol-1	86,4±5,7	Doñ-0	Doñ-0	Doñ-0	Ler	Ler
Lso-4	83,1±5,7	Ler	Ler	Doñ-0	Doñ-0	Doñ-0
Mar-9	70,8±8,5	Doñ-0	Doñ-0	Doñ-0	Ler	Ler
Bon-61	70,0±5,0	Doñ-0	Doñ-0	Doñ-0	Doñ-0	Doñ-0
Rdc-0	63,0±7,0	Ler	Ler	Ler	Ler	Ler
Per-0	12,9±5,1	Ler	Ler	Ler	Doñ-0	Ler
Tod-5	12,0±5,0	Ler	Ler	Ler	Ler	Ler
Ura-0	10,0±4,0	Ler	Ler	Ler	Ler	Ler
Cem-0	0	Ler	Ler	Ler	Doñ-0	Ler
Ped-0	0	Ler	Ler	Ler	Doñ-0	Ler
Gra-0	0	Ler	Ler	Ler	Ler	Ler
Her-12	0	Ler	Ler	Ler	Ler	Ler
Nac-0	0	Ler	Ler	Ler	Ler	Ler
Hum-2	0	Ler	Ler	Ler	Ler	Ler
Vim-0	0	Ler	Ler	Ler	Ler	Ler
San-0	0	Ler	Ler	Ler	Doñ-0	Ler
Vie-0	0	Ler	Ler	Ler	Ler	Ler
Qui-0	0	Ler	Ler	Ler	Ler	Ler
Fei-0	0	Ler	Ler	Ler	Ler	Ler
Cdm-0	0	Ler	Ler	Ler	Ler	Ler
		$P=0,001$	$P=0,001$	$P=0,0003$	$P=0,337$	$P=0,007$
		$R^2=0,378$	$R^2=0,378$	$R^2=0,453$		$R^2=0,261$

**Tabla 13. Asociación genética entre polimorfismos en el promotor de *EGL3* y el número de tricomas en fruto.** La tabla muestra el número de tricomas en el fruto (NTF) de las 23 accesiones del estudio y los polimorfismos candidatos para *FTD1*. Para cada uno de los polimorfismos se indica el alelo que poseen (Ler o Doñ-0), la significación estadística obtenida mediante ANOVA y la varianza explicada ( $R^2$ ) por los SNPs asociados ( $P<0,01$ ).

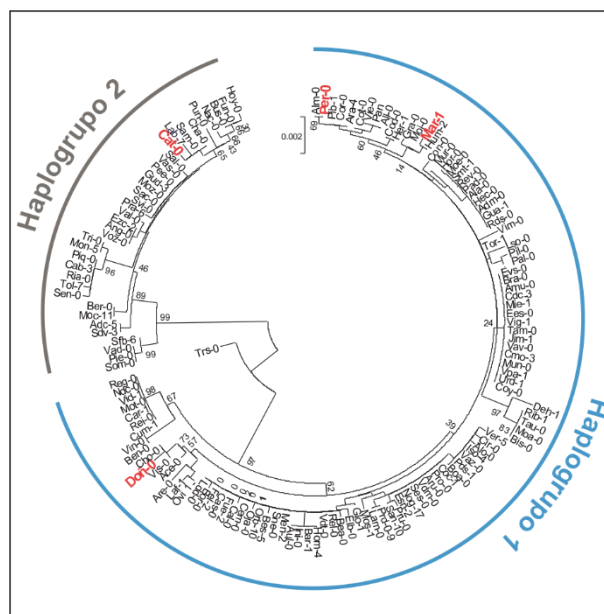


**Figura 21. Diversidad nucleotídica y cambios de aminoácidos en la región codificante de *EGL3* en la península ibérica.** La figura muestra la diversidad nucleotídica a lo largo de la región codificante de *EGL3* (A), la estructura de la región codificante de *EGL3* (B) y la estructura de su proteína (C) en la que se muestran los 26 aminoácidos que segregan en las 145 accesiones de la península ibérica. Los cambios aminoacídicos que diferencian los dos haplogrupos se resaltan en color verde. La longitud del gen y la proteína se expresan en nucleótidos y aminoácidos, respectivamente.

Para determinar las relaciones genéticas entre las 145 accesiones de la península ibérica para la región codificante de *EGL3* se realizó un análisis filogenético de estas secuencias. A diferencia de la región promotora, este análisis identificó dos haplogrupos mayoritarios diferenciados en 41 SNPs (Figura 22), de los cuales 27 se localizan en exones, incluyendo los siguientes cuatro cambios de aminoácido: Val/Ile<sup>328</sup>, Ile/Met<sup>470</sup>, Asp/Glu<sup>479</sup> y Lys/Asn<sup>578</sup> (Figura 21C). Este análisis sólo incluyó cuatro accesiones con tricomas en el fruto, tres de las cuales pertenecen al haplogrupo 1 y en el que también está la cepa referencia sin pelos en el fruto *Ler*. Por último, para determinar cuál de los dos haplogrupos es el ancestral también se realizó un análisis filogenético de la proteína de *EGL3* de las 145 accesiones de *Arabidopsis thaliana* junto con la secuencia de *Arabidopsis lyrata* (ver Materiales y Métodos, apartado 10). Sin embargo, debido a la elevada diferenciación genética de *EGL3* de *A. lyrata* no se pudo responder a esta pregunta.

En conjunto, los análisis de la diversidad nucleotídica del gen *EGL3* sugieren que además del polimorfismo funcional entre *Ler* y Doñ-0 localizado en la región promotora, podría haber otros polimorfismos que afecten a la función de este gen en otras accesiones de la de la península ibérica. Además, tanto la delección de 43 pb situada en posición -771, como los cuatro SNPs de la región promotora asociados con el NTF podrían contribuir al efecto del locus *FTD1*.

**Figura 22. Árbol filogenético de la región codificante de *EGL3* en la península ibérica.** El árbol mostrado fue obtenido a partir de las 145 accesiones procedentes de la península ibérica usando el método de “*neighbour joining*”. Las accesiones con tricomas en el fruto se muestran en color rojo.

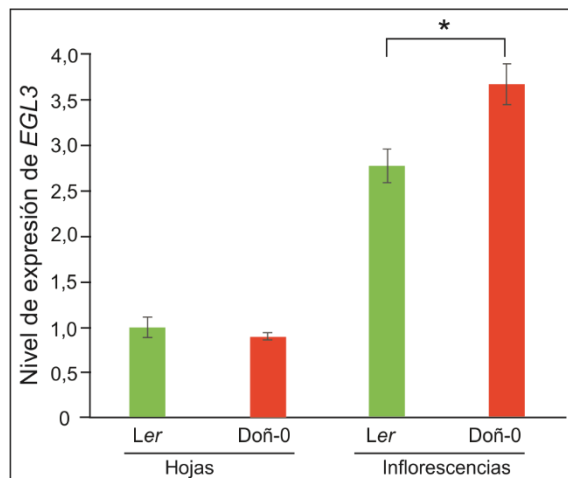


### 2.3. Análisis de expresión de *EGL3* en accesiones parentales

Debido a que el mapeo fino y el estudio de la diversidad nucleotídica sugieren que uno o varios polimorfismos en el promotor de *EGL3* son los causantes de *FTD1*, se estudió el patrón de expresión de este gen en las accesiones parentales *Ler* y Doñ. Para ello se analizó el nivel de expresión tanto en tejido vegetativo (hojas) como reproductivo (inflorescencias principales) mediante PCR cuantitativa a tiempo real (RT-PCR; ver Materiales y Métodos, apartado 7).

El análisis de las líneas parentales mostró que *Ler* y Doñ-0 difieren significativamente ( $P=0,012$ ) en el nivel de expresión de *EGL3* en la fase reproductiva pero no en tejido vegetativo ( $P>0,05$ ) (Figura 23). Las inflorescencias de Doñ-0 mostraron un nivel de expresión 25% superior que las de *Ler*, apoyando que algún polimorfismo entre *Ler* y Doñ-0 afecta a la regulación de *EGL3*, de manera que el alelo de Doñ-0 aumenta la expresión y consecuentemente conduce a una ganancia de función y a un aumento de la densidad de tricomas en el fruto. Además, este análisis mostró que *EGL3* se expresa 3-4 veces más en las inflorescencias que en las hojas de *Ler* y Doñ-0 ( $P<0,01$ ), lo que sugiere que *EGL3* podría estar implicado en la formación

de tricomas en los sépalos, ya que es el único órgano de la inflorescencia con tricomas en ambas accesiones.



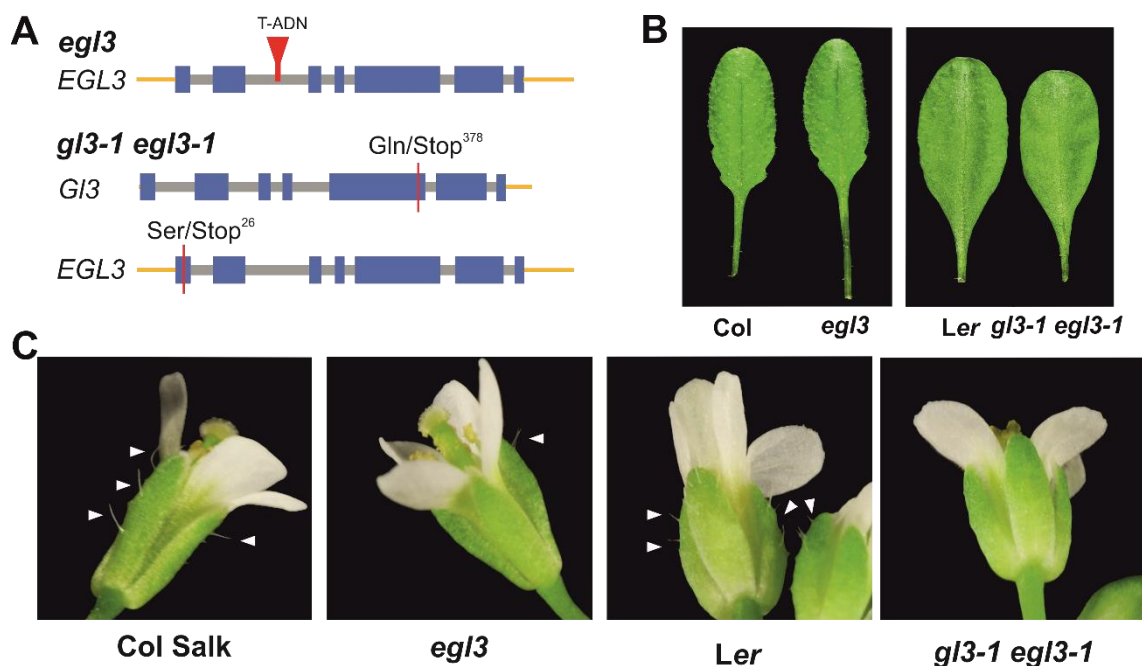
**Figura 23. Análisis de la expresión del gen *EGL3*.** Cada panel muestra el nivel de expresión del gen *EGL3* normalizado frente a la expresión del gen *UBC* en las líneas parentales (*Ler* y *Doñ-0*). El análisis estadístico se llevó a cabo mediante ANOVA (\*=  $P < 0,05$ ).

#### 2.4. Análisis de líneas mutantes de *EGL3*

Puesto que *EGL3* es el gen *FTD1* se ha llevado a cabo la caracterización fenotípica del patrón de tricomas en fase vegetativa y reproductiva de dos líneas mutantes en este gen: el mutante simple *egl3* portador de una inserción de T-ADN en un fondo genético Col; y el mutante doble *gl3-1 egl3-1* que posee mutaciones puntuales que provocan codones de parada y la pérdida de función de ambos genes en un fondo genético *Ler* (Figura 24A; ver Materiales y Métodos, apartado 1.4).

El análisis del patrón de tricomas en el mutante simple *egl3* mostró una reducción del 30% en la densidad de tricomas en hoja con respecto a Col ( $P < 0,01$ ; Tabla 14; Figura 24B), lo que concuerda con estudios previos (Zhang *et al.*, 2003). Igualmente, *egl3* y Col difieren en el patrón de tricomas en el tallo, siendo el PTT del mutante más bajo. Sin embargo, estas líneas no difieren en el patrón de tricomas en el fruto ya que ninguno de los dos genotipos desarrolló tricomas en este órgano. Además, el análisis del patrón de tricomas en los sépalos mostró que *egl3* difiere de Col (Tabla 14; Figura 24C), lo que indica que *EGL3* también es un regulador del patrón de tricomas en este órgano.

El análisis del patrón de tricomas de la línea doble mutante *gl3-1 egl3-1* mostró que estas plantas son completamente glabras tanto en fase vegetativa como reproductiva (Tabla 14; Figura 24B y 24C), en concordancia con estudios previos (Zhang *et al.*, 2003). Este resultado indica que *EGL3* interacciona genéticamente con *GL3* para regular el patrón de tricomas, ya que ambos mutantes simples desarrollan tricomas en las hojas (Koorneef *et al.*, 1982).



**Figura 24. Análisis de líneas mutantes en *EGL3*.** A) Posición de las mutaciones en *EGL3* y *GL3* en los mutantes de *egl3* y *gl3-1 egl3-1*. B, C) Patrón de tricomas en las hojas (B) y en los sépalos (C) de las líneas *egl3*, *gl3-1 egl3-1* y las líneas silvestres Col y Ler. Los triángulos blancos señalan los tricomas en los sépalos.

	Carácter fenotípico	Línea	Media ± EE	P	N
A	DTH	Col	0,59±0,02	**	41
		<i>egl3</i>	0,47±0,01		42
	PTT	Col	3,51±0,09	*	47
		<i>egl3</i>	3,18±0,09		49
NTS	Col	19,55±0,95	**	47	
	<i>egl3</i>	1,73±0,33		49	
B	DTH	Ler	0,21±0,01	**	44
		<i>gl3-1 egl3-1</i>	0,01±0,001		48
	PTT	Ler	1,84±0,09	**	49
		<i>gl3-1 egl3-1</i>	0		50
NTS	Ler	12,31±0,74	**	49	
	<i>gl3-1 egl3-1</i>	0		50	

**Tabla 14. Patrón de tricomas de las líneas mutantes de *EGL3*.** La tabla muestra la densidad de tricomas en la hoja (DTH), el patrón de tricomas en el tallo (PTT) y el número de tricomas en los sépalos (NTS) del mutante simple *egl3* (A) y el mutante doble *gl3-1 egl3-1* (B) junto a las líneas control. Los análisis estadísticos se realizaron con ANOVA (\*\*:  $P < 0,01$ ; \*:  $P < 0,05$ ; ns:  $P > 0,05$ ).



## **2.5. Análisis funcional de los alelos de *EGL3* de *Ler* y *Doñ-0* mediante complementación en plantas transgénicas**

Para demostrar que los alelos de *EGL3* de *Ler* y *Doñ-0* difieren funcionalmente se ha iniciado un análisis de complementación funcional en líneas transgénicas. Para ello se han obtenido dos construcciones genómicas de 6,7 kb del gen *EGL3* de *Ler* y *Doñ-0*. Estas construcciones incluyen 2,9 kb de la región promotora de *EGL3*, la región codificante completa (2,9 kb) y 0,9 kb de la región 3' no codificante (ver Materiales y Métodos, apartado 6). Los fragmentos procedentes de las dos accesiones se han clonado en el vector pCAMBIA 2300 y posteriormente se han utilizado para transformar plantas de dos fondos genéticos distintos: la línea mutante *gl3-1 egl3-1* y la línea de introgresión LI-25 portadora de alelos *Doñ-0* sólo en la región de *FTD2* y *FTD5*. Actualmente se ha obtenido la semilla de la generación T1 que será procesada en el futuro para la obtención de líneas transgénicas T3 homocigóticas. Es de esperar, que el análisis de estas líneas transgénicas permita demostrar de forma conclusiva que *EGL3* es el gen correspondiente al locus *FTD1*.

### 3. Caracterización genética y molecular de *FTD2*, *FTD3* y *FTD5*

Los loci *FTD2*, *FTD3* y *FTD5* se han identificado previamente por su efecto mayor sobre el patrón de tricomas en el fruto (ver Introducción, apartado 4.2). Además, el mapeo genético y el análisis de secuencias realizados en el laboratorio identificaron los genes *TCL1*, *GL1* y *TRY* como candidatos para los loci *FTD2*, *FTD3* y *FTD5*, respectivamente. Con el fin de demostrar si estos genes corresponden a dichos loci y caracterizarlos molecularmente, se ha llevado a cabo su estudio en líneas de introgresión y en líneas transgénicas para cada uno de los genes candidatos.

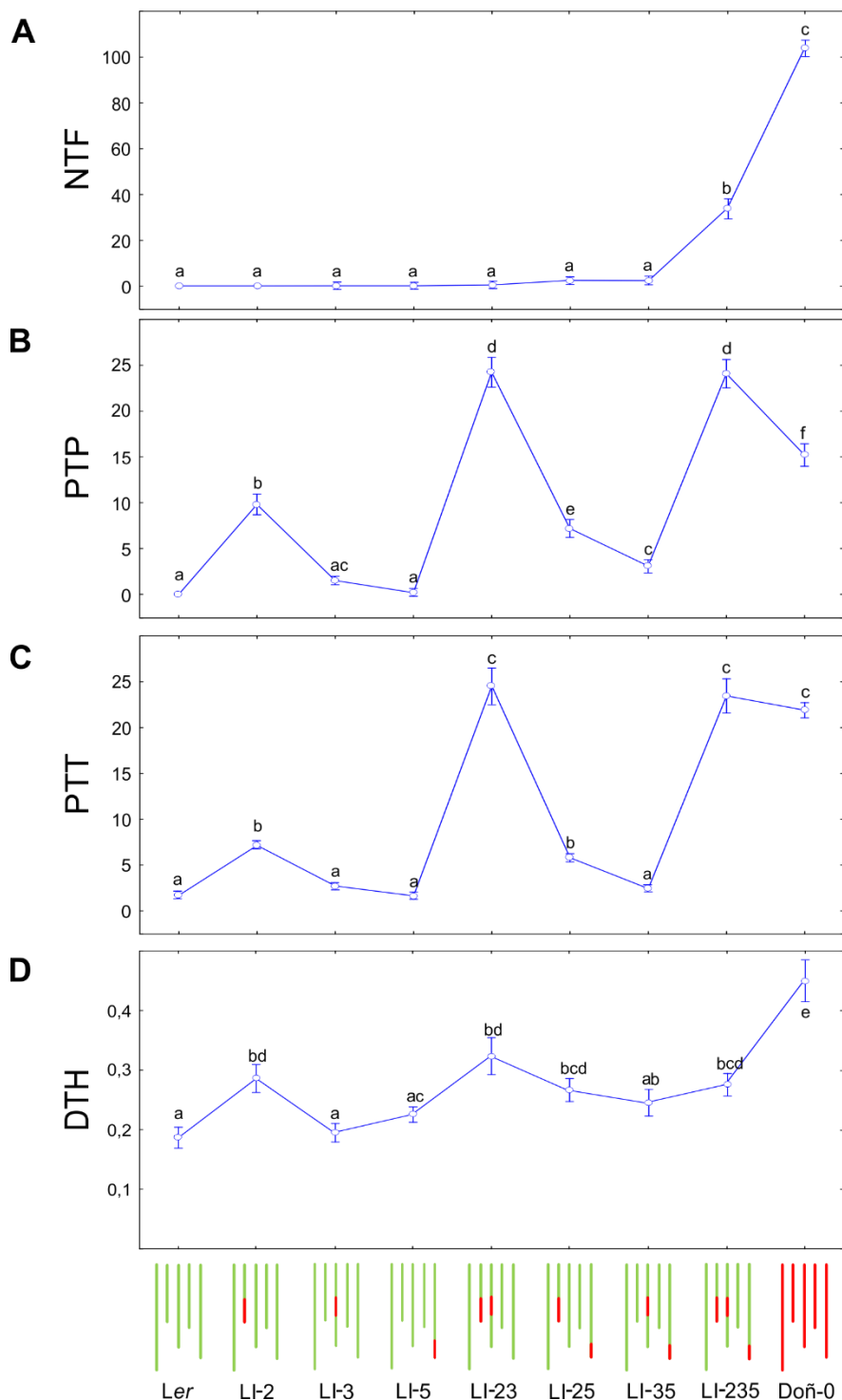
#### 3.1. Caracterización genética de los loci *FTD2*, *FTD3* y *FTD5*

Con el fin de validar el efecto los loci *FTD2*, *FTD3* y *FTD5* y determinar la interacción genética entre ellos, se ha realizado el análisis de líneas de introgresión (LIs) portadoras de alelos Doñ-0 para uno, dos o tres loci *FTD*. Estas líneas se obtuvieron previamente en el laboratorio a partir de cuatro RILs derivadas del cruzamiento entre las accesiones *Ler* y Doñ-0 (ver Materiales y Métodos, apartado 1.2). Este análisis se realizó con 21 líneas de introgresión correspondientes a siete genotipos distintos y las líneas parentales *Ler* y Doñ-0 (Figura 25), cuantificando su patrón de tricomas en fase vegetativa (DTH) y reproductiva (PTT, PTP y NTF) en condiciones de día largo.

El análisis del número de tricomas en el fruto (Figura 25A) mostró que todas las LIs con alelos Doñ-0 en un único locus *FTD* carecen de tricomas, lo que indica que un solo locus *FTD* no es suficiente para desarrollar tricomas en el fruto. Las LIs con alelos Doñ-0 en *FTD2* y *FTD5* o *FTD3* y *FTD5* sí produjeron tricomas, pero con una densidad muy baja ( $2,4 \pm 0,4$  y  $2,3 \pm 0,8$ , respectivamente) que no difiere significativamente de la ausencia de tricomas de *Ler*. Sin embargo, las LIs con alelos Doñ-0 en los tres loci *FTD* presentaron un número de tricomas elevado, lo que demuestra que estos tres loci interaccionan epistáticamente para regular el desarrollo de una alta densidad de tricomas en el fruto.

Análisis similares del patrón de tricomas en el pedicelo (Figura 25B) y en el tallo (Figura 25C) mostraron que todas las líneas de introgresión con alelos Doñ-0 en *FTD2*, (LI-2, LI-23, LI-25 y LI-235) tienen valores significativamente mayores que *Ler* ( $P < 0,01$ ). Por lo tanto, la presencia de alelos Doñ-0 en el locus *FTD2* incrementa los caracteres PTP y PTT con respecto a *Ler* de forma aditiva. Además, las LIs portadoras de alelos de Doñ-0 en los loci *FTD2* y *FTD3* (LI-23 y LI-235) presentaron un valor de PTP y PTT significativamente mayor que el de las líneas con alelos Doñ-0 sólo en *FTD2*, lo que

indica la interacción genética entre los dos loci para el control del desarrollo de tricomas en el pedicelo y en el tallo.



**Figura 25. Patrón de tricomas de las líneas de introgresión para los loci *FTD*.** A, B, C, D) Número de tricomas en el fruto (NTF; A), patrón de tricomas en el pedicelo (PTP; B), patrón de tricomas en el tallo (PTT; C) y densidad de tricomas en la hoja (DTH; D) cuantificado en LIs de siete genotipos distintos junto a las accesiones parentales Ler y Doñ-0. En cada panel se muestra la media  $\pm$  EE de cada genotipo. Dentro de cada panel se indica con letras diferentes los genotipos cuyos fenotipos difieren significativamente ( $P < 0,01$ ) y con la misma letra aquellos que no difieren de forma significativa. Las diferencias estadísticas entre las distintas líneas se contrastaron mediante ANOVA con la corrección de Bonferroni.

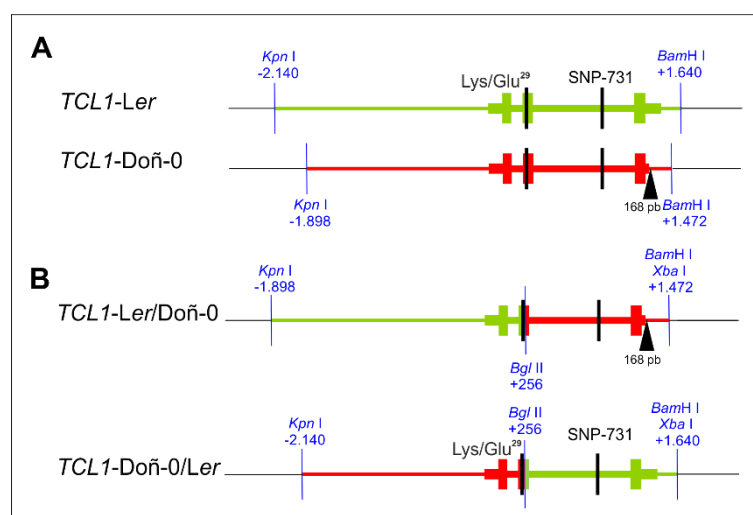
Finalmente, el análisis de la densidad de tricomas en la hoja (Figura 25D) mostró que todas las líneas que llevan alelos Doñ-0 en *FTD2* tienen mayor DTH que Ler. Por lo tanto, sólo el locus *FTD2* tiene un efecto aditivo independiente del resto de los loci *FTD*. No obstante, tanto para DTH como para NTF, Doñ-0 mostró un valor significativamente mayor que las líneas de introgresión analizadas, lo que indica que algún otro locus no incluido en el análisis también contribuye a la variación para estos caracteres.

En conjunto, estos resultados indican que *FTD2* es el único de los tres loci estudiados con efecto sobre DTH; *FTD2* y *FTD3* interaccionan para regular el desarrollo de tricomas en el tallo y el pedicelo; y el control de tricomas en el fruto está determinado por la interacción entre *FTD2*, *FTD3* y *FTD5*.

### 3.2. Aislamiento y caracterización molecular de *FTD2*

#### 3.2.1. Aislamiento de *FTD2*

Para demostrar si *TCL1* (*At2g30432*) es *FTD2* se utilizaron dos construcciones genómicas correspondientes a los alelos Ler y Doñ-0 de *TCL1*, de 3,8 y 3,4 kb respectivamente, obtenidas previamente (ver Materiales y Métodos, apartado 6; Figura 26A). Estas construcciones se usaron para generar plantas transgénicas en dos fondos genéticos: el mutante *tcl1-Salk* y la línea de introgresión LI-235 (ver Materiales y Métodos, apartado 1.5).

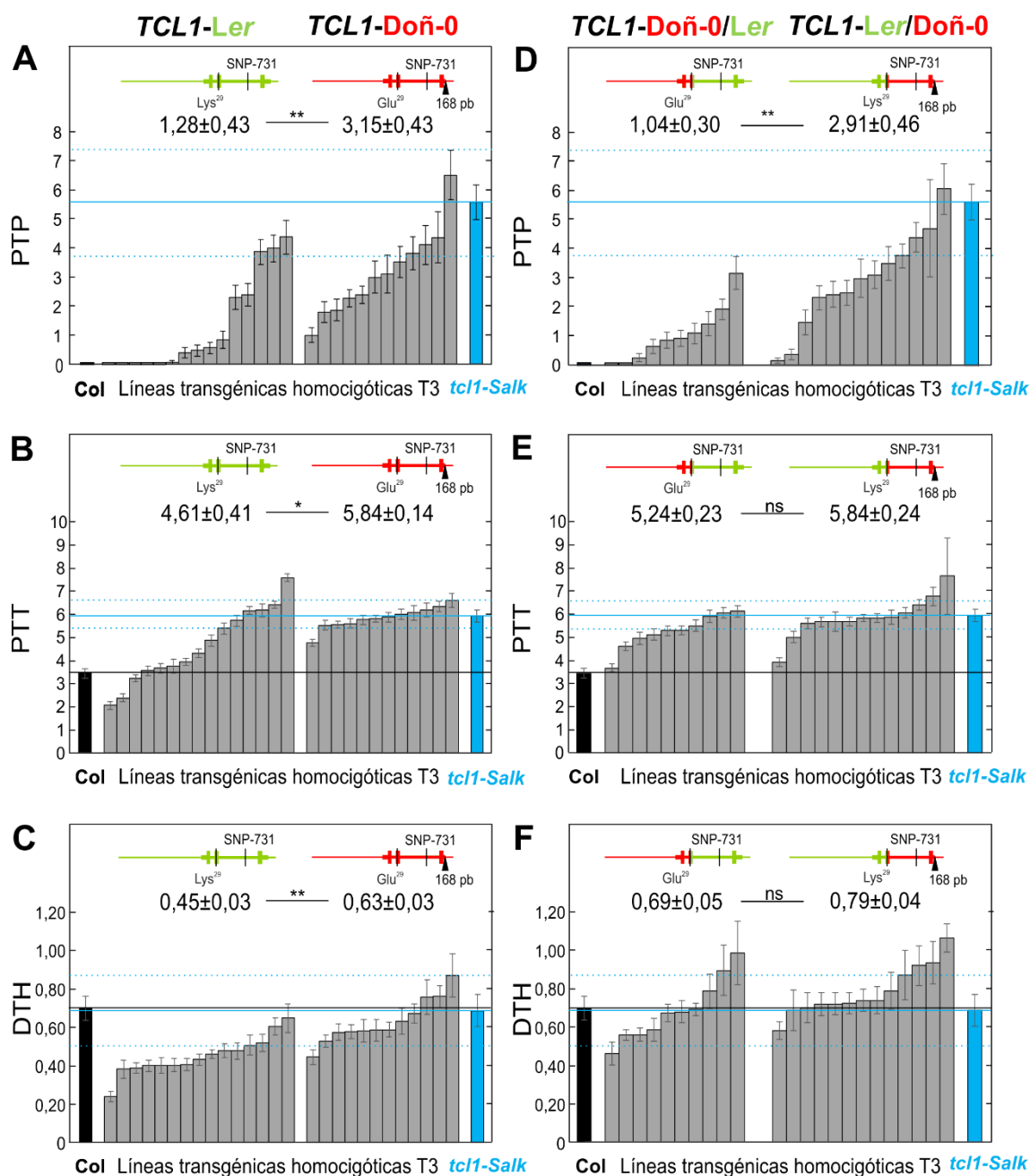


**Figura 26. Construcciones genómicas de *TCL1*.** A) Construcciones genómicas del gen *TCL1* de las líneas parentales Ler (verde) y Doñ-0 (rojo). B) Construcciones quiméricas Ler/Doñ-0 y Doñ-0/Ler. Las líneas verticales de color azul indican la posición de las dianas de restricción de los enzimas utilizados para la obtención de las construcciones. Los tres polimorfismos candidatos para *FTD2* se representan en negro.

El fondo genético mutante *tcl1-Salk* se seleccionó para determinar si los alelos de *TCL1* de *Ler* y Doñ-0 difieren funcionalmente ya que la pérdida de función de este gen se caracteriza por la presencia de tricomas en el pedicelo y por un incremento del valor de PTT (Wang *et al.*, 2007). En total se obtuvieron 15 y 12 líneas transgénicas independientes y homocigóticas con el transgén de *TCL1* de *Ler* y Doñ-0 respectivamente, que junto a las líneas Col y *tcl1-Salk* se analizaron para el patrón de tricomas en fase reproductiva (PTP y PTT) y fase vegetativa (DTH) (Figura 27).

El análisis del PTP muestra que todas las líneas que llevan el transgén *TCL1*-Doñ-0 tienen tricomas en el pedicelo, aunque 8 de las líneas analizadas presentaron un valor de PTP inferior al de las plantas mutantes *tcl1-Salk* sin transformar. Por el contrario, 10 de las 15 líneas con el transgén de *Ler* complementaron por completo el fenotipo de PTP del mutante *tcl1-Salk* ya que presentaron un valor inferior a uno (Figura 27A). Además, los dos tipos de líneas transgénicas difieren significativamente para PTP ( $P=0,005$ ) de manera que las líneas con el transgén *TCL1*-Doñ-0 tienen una mayor cantidad de tricomas en el pedicelo. Estos resultados indican que ambos alelos naturales son funcionales ya que poseen un valor de PTP menor que la línea *tcl1-Salk* sin transformar (Figura 27A). Sin embargo, la diferencia de complementación del fenotipo mutante de *tcl1-Salk* indica que los alelos naturales de *Ler* y Doñ-0 difieren funcionalmente, de manera que, teniendo en cuenta la función represora de *TCL1* sobre la formación de tricomas, el alelo de Doñ-0 es menos activo que el de *Ler*.

Para determinar si los alelos de *TCL1* de *Ler* y Doñ-0 afectan al patrón de tricomas en otros órganos de la planta, también se analizó el patrón de tricomas en tallo (Figura 27B) y la densidad de tricomas en la hoja (Figura 27C). En ambos análisis se observó que la mayoría de las líneas portadoras del transgén *TCL1*-Doñ-0 tienen valores de PTT y DTH similares a los del mutante *tcl1-Salk*. Por el contrario, la mayoría de las líneas con el transgén *TCL1*-*Ler* presentaron valores de PTT y DTH inferiores a los del mutante *tcl1-Salk*. Los dos grupos de líneas transgénicas difirieron estadísticamente en sus valores medios de PTT y DTH ( $0,001 < P < 0,02$ ).



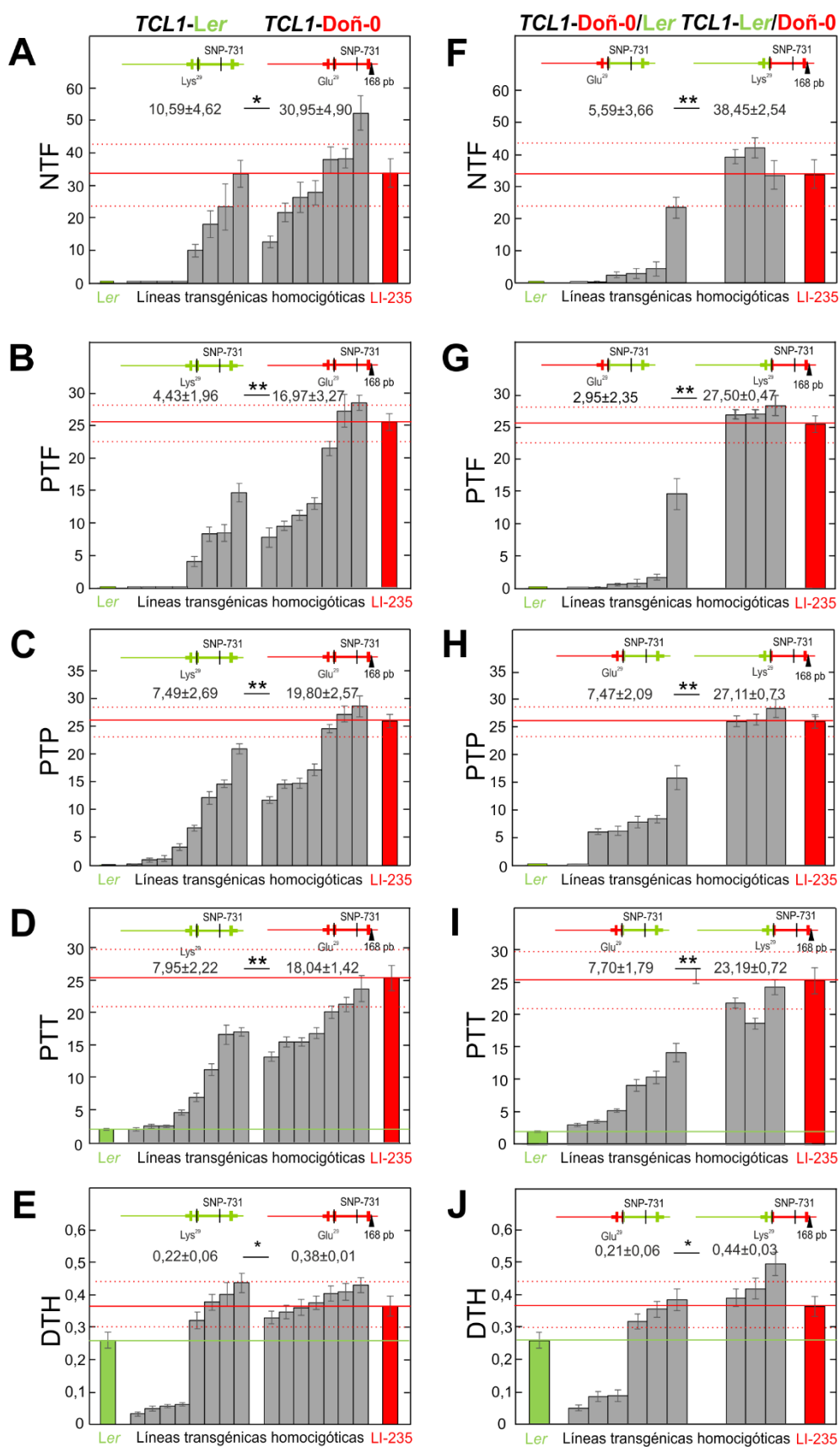
**Figura 27. Patrón de tricomas de las líneas transgénicas de *TCL1* en el fondo genético *tcl1-Salk*.** Patrón de tricomas en el pedicelo (PTP; A, D), patrón de tricomas en el tallo (PTT; B, E) y densidad de tricomas en la hoja (DTH; C, F) en líneas transgénicas homocigóticas T3 para las construcciones genómicas de *TCL1* de *Ler* y *Doñ-0* (A, B y C) y las construcciones quiméricas *Doñ-0/Ler* y *Ler/Doñ-0* (D, E y F) en fondo genético *tcl1-Salk*. Las líneas control sin transformar Col (negro) y *tcl1-Salk* (azul) se muestran a los lados de cada panel. Cada barra de cada panel representa la media  $\pm$  EE de cada línea, estimada a partir de 10-18 plantas. En la parte superior de cada panel se muestra el esquema del transgén y la media  $\pm$  EE de todas las líneas transgénicas para cada construcción indicando si difieren o no significativamente (\*\*:  $P < 0,01$ ; \*:  $P < 0,05$ ; ns:  $P > 0,05$ ). Las líneas horizontales continuas indican el valor medio del carácter de las líneas control Col (negro) y *tcl1-Salk* (azul) y las discontinuas limitan el intervalo de confianza del 95% de *tcl1-Salk*. Las diferencias estadísticas entre las distintas líneas se contrastaron mediante ANOVA con la corrección de Bonferroni.

Por otra parte, con el fin de determinar si *TCL1* es *FTD2*, también se transformó la línea de introgresión portadora de alelos Doñ-0 en *FTD2*, *FTD3* y *FTD5* (LI-235). Esta línea se eligió porque el desarrollo de alta densidad de tricomas en el fruto requiere alelos Doñ-0 en los tres loci *FTD*. Puesto que *TCL1* es un represor del desarrollo de tricomas, es de esperar que en este fondo genético el transgén *TCL1-Ler* reduzca el número y el patrón de tricomas en el fruto, en comparación con las líneas portadoras del transgén *TCL1-Doñ-0*. En total se obtuvieron 8 y 7 líneas independientes y homocigóticas para el transgén de *TCL1-Ler* y *TCL1-Doñ-0* respectivamente, las cuales se analizaron para su patrón de tricomas en fase reproductiva (NTF, PTF, PTP y PTT) y vegetativa (DTH) junto con las líneas *Ler* y LI-235 sin transformar (Figura 28).

El análisis del carácter NTF mostró que la mayoría de las líneas con el transgén de *TCL1-Doñ-0* poseen fenotipos similares al de la línea LI-235 sin transformar. Sin embargo, la mayoría de las líneas portadoras del transgén *TCL1-Ler* poseen valores de NTF significativamente inferiores a LI-235 (Figura 28A). Igualmente, el valor de PTF en la mayoría de las líneas con *TCL1-Doñ-0* es más alto que el de las líneas con *TCL1-Ler* (Figura 28B). De esta forma, ambos tipos de líneas transgénicas difirieron significativamente en su patrón y densidad de tricomas en el fruto ( $0,005 < P < 0,012$ ).

Así mismo, el análisis del PTP y el PTT (Figura 28C y 28D) mostró que las líneas con el transgén *TCL1-Doñ-0* y *TCL1-Ler* difieren significativamente ( $P < 0,01$ ) ya que, para ambos caracteres, la mayoría de las líneas con el transgén *TCL1-Doñ-0* tienen valores fenotípicos superiores a las líneas con el transgén *TCL1-Ler*. Finalmente, el análisis de DTH mostró que todas las líneas con el transgén de *TCL1-Doñ-0* poseen un fenotipo similar al de la línea LI-235, mientras que la mitad de las líneas con el transgén *TCL1-Ler* presentaron un fenotipo prácticamente glabro, ausente en las líneas con *TCL1-Doñ-0*.

En conjunto, el análisis de las líneas transgénicas de fondo LI-235 muestra que el transgén de *Ler* reduce el fenotipo de NTF, PTF, PTP, PTT y DTH de la línea LI-235 significativamente más que el transgén *TCL1-Doñ-0*. Estos resultados permiten concluir que *TCL1* es el gen correspondiente a *FTD2*, para el que Doñ-0 es portador de un alelo de pérdida de función parcial, y que su variación natural afecta al patrón de tricomas tanto en fase vegetativa como reproductiva.



**Figura 28. Patrón de tricomas de las líneas transgénicas de *TCL1* en el fondo genético LI-235.** Número de tricomas en el fruto (NTF; A, B), patrón de tricomas en el fruto (PTF; B, G), patrón de tricomas en el pedicelo (PTP; C, H), patrón de tricomas en el tallo (PTT; D, I) y densidad de tricomas en la hoja (DTH; E, J) en líneas transgénicas homocigóticas T3/T4 para las



construcciones genómicas de *TCL1* de *Ler* y Doñ-0 (A, B, C, D y E) y construcciones quiméricas Doñ-0/*Ler* y *Ler*/Doñ-0 (F, G, H, I y J) en el fondo genético LI-235. Las líneas control sin transformar *Ler* (verde) y LI-235 (rojo) se muestran a los lados de cada panel. Cada barra representa la media  $\pm$  EE de cada línea, estimada a partir de 10-18 plantas. En la parte superior de cada panel se muestra el esquema del transgén de las líneas y la media  $\pm$  EE de todas las líneas transgénicas para cada construcción indicando si difieren o no significativamente (\*\*:  $P < 0,01$ ; \*:  $P < 0,05$ ; ns:  $P > 0,05$ ). Las líneas horizontales continuas indican el valor medio del carácter fenotípico en las líneas control *Ler* (verde) y LI-235 (rojo) y las discontinuas limitan el intervalo de confianza al 95% de LI-235. Las diferencias estadísticas entre las distintas líneas se contrastaron mediante ANOVA con la corrección de Bonferroni.

### 3.2.2. Identificación del polimorfismo funcional de *TCL1*/FTD2

La comparación de las secuencias del gen *TCL1* entre *Ler* y Doñ-0 ha identificado previamente (Savic, 2015) tres polimorfismos asociados a la presencia de tricomas en el fruto: un cambio de aminoácido de Lys/Glu<sup>29</sup>, un SNP en el segundo intrón en posición 731 y una delección de 168 pb en la región 3' UTR de Doñ-0 (Figura 26). Para identificar el polimorfismo responsable de la alteración funcional de *TCL1* entre *Ler* y Doñ-0, se obtuvieron dos construcciones genómicas quiméricas derivadas por intercambio del fragmento *Kpn* I-*Bgl* II entre las construcciones parentales de *Ler* y Doñ-0 (Figura 26B). La construcción *TCL1-Ler/Doñ-0* contiene el promotor y la mitad 5' del gen *TCL1* de *Ler* (incluyendo el cambio de aminoácido en posición 29) y la mitad 3' de Doñ-0 (con el SNP-731 y la delección de la región 3'), mientras que la construcción *TCL1-Doñ-0/Ler* es recíproca. A partir de estas construcciones, se obtuvieron un total de 23 y 9 líneas transgénicas en los fondos genéticos *tcl1-Salk* y LI-235, respectivamente (ver Materiales y Métodos, apartado 1.5). Estas líneas se han analizado en detalle en esta tesis para su patrón de tricomas, tanto en fase reproductiva como vegetativa.

El análisis fenotípico de las líneas de fondo genético *tcl1-Salk* para el patrón de tricomas en pedicelo (PTP) mostró que la mayoría de las 13 líneas con el transgén quimérico de *TCL1-Ler/Doñ-0* tienen mayor valor que las 10 líneas con el transgén quimérico recíproco *TCL1-Doñ-0/Ler* (Figura 27D). Ambos grupos de líneas difirieron significativamente para PTP ( $P = 0,005$ ). Sin embargo, las líneas con el transgén *TCL1-Doñ-0* o *TCL1-Ler* no difieren de las líneas con el transgén *TCL1-Ler/Doñ-0* o *TCL1-Doñ-0/Ler*, respectivamente ( $P > 0,05$ ). Por lo tanto, las diferencias funcionales entre el gen *TCL1* de *Ler* y Doñ-0 se localizan en la región 3' del mismo. El análisis de PTT y DTH no mostró diferencias significativas entre los dos tipos de líneas transgénicas ( $P > 0,05$ ), a pesar de que las líneas con el transgén *TCL1-Ler/Doñ-0* presentaron valores de PTT y DTH superiores a las líneas con el transgén *TCL1-Doñ-0/Ler* (Figura 27E y 27F).

El estudio de las líneas obtenidas en el fondo LI-235, mostró que las tres líneas portadoras del transgén *TCL1-Ler/Doñ-0* tienen valores de NTF y PTF altos y similares

a la línea LI-235, mientras que la mayoría de las seis líneas con *TCL1-Doñ-0/Ler* presentaron valores de estas variables muy bajos o fueron glabras (Figura 28F y 28G). De esta forma, los dos tipos de plantas transgénicas difieren estadísticamente para ambos caracteres ( $P < 0,01$ ). Igualmente, el análisis de estas líneas para los caracteres PTP y PTT (Figura 28H-28I) mostró que las plantas con el transgén *TCL1-Ler/Doñ-0* tienen valores significativamente mayores que las líneas *TCL1-Doñ-0/Ler* ( $P < 0,01$ ). Los dos tipos de líneas transgénicas también mostraron diferencias significativas para el valor de DTH ( $P < 0,05$ ) (Figura 28J). De forma similar a lo observado en las líneas con el transgén de *Ler*, la mitad de las líneas con *TCL1-Doñ-0/Ler* presentaron una densidad de tricomas en la hoja muy baja y significativamente inferior al de las líneas con el transgén *TCL1-Ler/Doñ-0*. Al igual que lo observado en el fondo *tcl1-Salk*, las líneas con el transgén *TCL1-Ler* no difieren de las líneas con el transgén *TCL1-Doñ/Ler* para ninguno de los caracteres analizados ( $P > 0,05$ ) y las líneas con el transgén *TCL1-Doñ-0* y *TCL1-Ler/Doñ-0* mostraron valores similares para todos los caracteres. Los resultados del análisis de líneas transgénicas en ambos fondos genéticos demuestran que el polimorfismo responsable de las diferencias funcionales de *FTD2* se localiza en el segmento de 1,4 kb de la región 3' del gen *TCL1* situado entre la posición 256 del segundo intrón y los 431 pb posteriores al codón de parada. Puesto que *Ler* y *Doñ-0* no difieren en ningún cambio de aminoácido en este segmento, el polimorfismo funcional de *TCL1* tiene que ser regulador. Esta región incluye 4 SNPs y 6 indels localizados en intrones y la región 3' UTR que podrían afectar a la regulación de la expresión de *TCL1* (Savic, 2015). Sin embargo, el indel de 168 pb es el polimorfismo de la región 3' (Figura 26) asociado más significativamente con la densidad de tricomas en fruto (Savic, 2015), lo que sugiere que éste es el polimorfismo funcional de *FTD2* que afecta a la regulación de la expresión de *TCL1*.

### **3.2.3. Caracterización funcional de la variación natural *Ler/Doñ-0* de *TCL1***

Con el fin de caracterizar el polimorfismo regulador de *TCL1/FTD2* se analizó la expresión de *TCL1* en las líneas transgénicas de fondo genético *tcl1-Salk* y LI-235, tanto en fase vegetativa como reproductiva (Tabla 15). Estos análisis se realizaron mediante PCR cuantitativa a tiempo real (RT-PCR) a partir de ARN aislado de hojas e inflorescencias principales (ver Materiales y Métodos, apartado 7).

Fondo genético	Transgén	Expresión <i>TCL1</i>				
		Hoja	<i>P</i>	Inflorescencia	<i>P</i>	<i>N</i>
<i>tcl1-Salk</i>	Ler	1,9±0,3	**	44,4±13,1	**	6
	Doñ-0	0,9±0,1		8,0±1,8		6
	Doñ-0/Ler	1,6±0,4	**	35,8±6,2	**	6
	Ler/Doñ-0	0,8±0,2		2,5±1,1		6
<i>tcl1-Salk</i>	Col	1,0±0,2	**	4,7±0,7	**	3
	Col	3,8±0,2		36,7±1,8		3
LI-235	Ler	81,3±30,1	**	92,6±31,7	**	8
	Don	7,3±2,8		8,2±4,4		7
	Doñ-0/Ler	61,5±37,1	**	74,6±37,9	**	6
	Ler/Doñ-0	2,4±1,1		2,2±0,9		3
	Ler	6,7±0,6	**	14,3±1,1	**	3
	LI-235	1,0±0,4		1,7±0,2		4

**Tabla 15. Expresión del gen *TCL1* en líneas transgénicas.** La tabla muestra el fondo genético, el transgén, la expresión media  $\pm$  EE de *TCL1* en hoja e inflorescencia, y el número de líneas analizadas para cada tipo de transgén. Los niveles de expresión de cada línea en fondo *tcl1-Salk* y LI-235 y cada línea control están relativizados con respecto a la expresión media de las líneas control sin transformar *tcl1-Salk* o LI-235 en fase vegetativa, respectivamente. Las diferencias estadísticas entre las distintas líneas se contrastaron mediante ANOVA (\*\*:  $P < 0,01$ ).

El análisis de la expresión de *TCL1* en las líneas transgénicas de fondo *tcl1-Salk* se llevó a cabo con 6 líneas por cada construcción genómica (Tabla 15). Al igual que en las líneas control no transformadas (*tcl1-Salk* y Col) la expresión media de *TCL1* en los cuatro tipos de líneas transgénicas fue 5 a 20 veces mayor en la fase reproductiva que en la vegetativa. Además, en fase vegetativa, las líneas con el transgén *TCL1-Ler* mostraron el doble de expresión media que las líneas con el transgén *TCL1-Doñ-0* ( $P < 0,001$ ). De igual forma, en fase reproductiva, las líneas con los transgenes parentales difirieron significativamente en la expresión de *TCL1*, de manera que las líneas con el transgén *TCL1-Ler* presentaron una media cinco veces mayor que las líneas *TCL1-Doñ-0* (Tabla 15). Las líneas transgénicas para las construcciones quiméricas *TCL1-Doñ-0/Ler* y *TCL1-Ler/Doñ-0* mostraron resultados similares, con niveles y diferencias de expresión comparables a los de las líneas con los transgenes parentales (Tabla 15). Estos resultados indican que la expresión de *TCL1* tanto en tejido vegetativo como reproductivo difiere significativamente entre las líneas con el transgén de Doñ-0 y de Ler (Tabla 15). Por lo tanto, el polimorfismo funcional de *FTD2*, localizado en la mitad 3' del gen *TCL1*, afecta a la expresión general de *TCL1*.

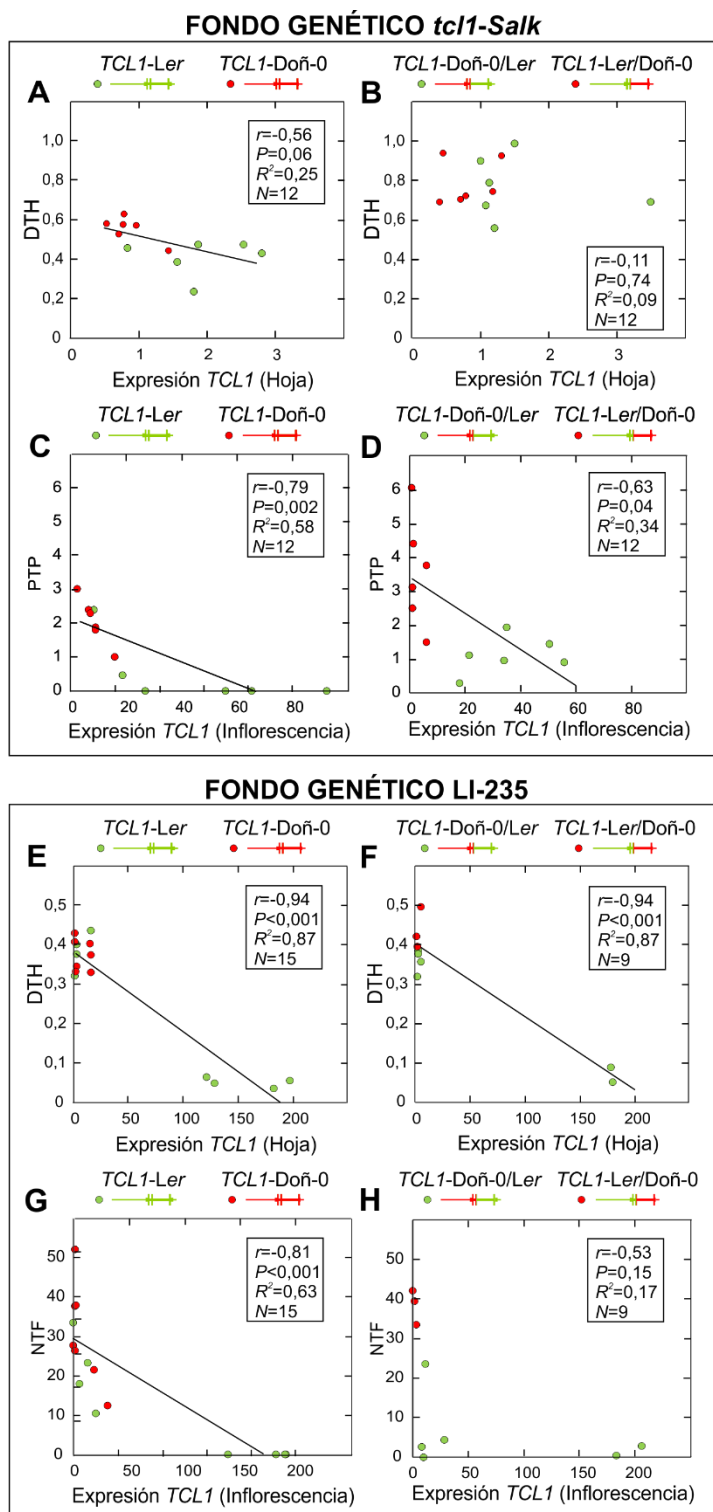
El análisis de expresión de *TCL1* en las líneas transgénicas de fondo LI-235 se realizó en 3 a 8 líneas para cada uno de los cuatro transgenes. Al igual que en el fondo *tcl1-Salk*, en el fondo genético LI-235, tanto las líneas control sin transformar (Ler y LI-235) como los cuatro tipos de líneas transgénicas, presentaron diferencias significativas entre la expresión de *TCL1* en la fase vegetativa y la reproductiva (Tabla 15). En

concreto, las líneas con los transgenes *TCL1-Doñ-0* o *TCL1-Ler/Doñ-0* presentaron una expresión media 10-20 veces inferior al de las líneas con los transgenes *TCL1-Ler* o *TCL1-Doñ-0/Ler* (Tabla 15). Además, las líneas con los transgenes *TCL1-Doñ-0* o *TCL1-Ler/Doñ-0* mostraron la misma expresión en fase vegetativa y reproductiva. Estos resultados confirman que los alelos de *TCL1* de *Ler* y *Doñ-0* difieren en su expresión, de manera que el polimorfismo funcional localizado en la región 3' del gen produce una reducción drástica de la expresión de *TCL1* en el alelo de *Doñ-0*. Además, en un fondo genético *Ler* pero no en fondo *Col*, el alelo *Doñ-0* de *TCL1* no cambia su nivel de expresión a lo largo del desarrollo.

Por último, para determinar la relación entre el nivel de expresión de *TCL1* y el patrón de tricomas en cada órgano se obtuvieron los coeficientes de correlación de Pearson entre el nivel de expresión de *TCL1* y el patrón de tricomas en la hoja, el pedicelo o el fruto (DTH, PTP y NTF) (Figura 29).

El análisis de las líneas de fondo genético *tcl1-Salk* mostró que la mayoría de las líneas portadoras de los transgenes *TCL1-Doñ-0* o *TCL1-Ler/Doñ-0* presentan una menor expresión de *TCL1* y mayor DTH o PTP que las líneas con los transgenes *TCL1-Ler* o *TCL1-Doñ-0/Ler*, especialmente en fase reproductiva (Figura 29A-29D). En conjunto, las líneas con los transgenes parentales o quiméricos mostraron correlaciones negativas entre la expresión de *TCL1* y la densidad de tricomas en hoja ( $-0,56 < r < -0,11$ ; Figura 29A y 29B) aunque la correlación sólo fue significativa para los transgenes parentales. Sin embargo, las correlaciones entre la expresión de *TCL1* y el patrón de tricomas en pedicelo fueron mayores ( $-0,79 < r < -0,63$ ; Figura 29C y 29D), de forma que la expresión de *TCL1* explica el 25-60% de la variación fenotípica de PTP en las líneas con los transgenes parentales o quiméricos.

Las líneas de fondo genético LI-235 mostraron correlaciones negativas y altas ( $-0,81 < r < -0,94$ ) entre los niveles de expresión de *TCL1* y el patrón de tricomas en la hoja o el fruto (Figura 29E-H). En particular, todas las líneas con los transgenes *TCL1-Doñ-0* o *TCL1-Ler/Doñ-0* presentaron expresión baja tanto en fase vegetativa como reproductiva y unos valores de DTH y NTF altos. Sin embargo, las líneas con los transgenes *TCL1 Ler* o *Doñ-0/Ler* presentaron mucha mayor variación tanto en la expresión de *TCL1* como en los fenotipos de DTH y NTF, aunque con una expresión media 10-20 veces mayor que las líneas con los otros dos transgenes (Figura 29E-29H). En total, el nivel de expresión de *TCL1* explicó el 87 y el 63% de la variación para el patrón de tricomas en la hoja y en el fruto, respectivamente.



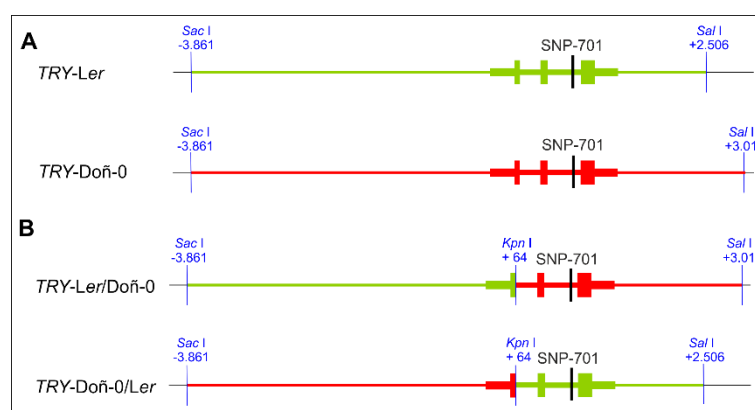
**Figura 29. Relación entre el nivel de expresión de *TCL1* y el patrón de tricomas de las líneas transgénicas.** Relación entre la expresión de *TCL1* en la hoja y DTH (A, B, E, F), o la expresión de *TCL1* en la inflorescencia y PTP o NTF (C, D, G, H) en las líneas transgénicas de fondo genético *tc1-Salk* (A, B, C, D) y LI-235 (E, F, G, H). A, C, E y G muestran las relaciones en las líneas con los transgenes *TCL1* parentales, mientras que B, D, F y H muestran las relaciones en las líneas con los transgenes quiméricos. Dentro de cada panel se indica el coeficiente de correlación de Pearson ( $r$ ), la significación estadística ( $P$ ), la proporción de la varianza explicada por la relación entre las dos variables comparadas ( $R^2$ ) y el número de líneas incluidas ( $N$ ). Los niveles de expresión de cada línea en fondo *tc1-Salk* y LI-235 están relativizados con respecto a la expresión media de las líneas control sin transformar *tc1-Salk* o LI-235 en fase vegetativa, respectivamente.

Estos resultados indican que la variación natural en la expresión de *TCL1* determina gran parte de la variación fenotípica para el patrón de tricomas en la hoja y en el fruto. Por lo tanto, el polimorfismo regulador de la región 3' de *TCL1* es el responsable de la variación natural causada por *FTD2* de manera que el alelo de Doñ-0 presenta una menor expresión y, consecuentemente, una menor actividad represora del patrón de tricomas en fase vegetativa y reproductiva.

### 3.3. Aislamiento y caracterización molecular de *FTD5*

#### 3.3.1. Aislamiento de *FTD5*

Con el fin de demostrar si *TRY* (*At5g53200*) es el gen *FTD5*, previamente se obtuvieron en el laboratorio dos construcciones genómicas de 6,2-6,7 kb correspondientes a los alelos naturales de *TRY* de Ler y Doñ-0, respectivamente (Figura 30A; ver Materiales y Métodos, apartado 6). A partir de estas construcciones se generaron líneas transgénicas en los fondos genéticos *try-Salk* y LI-235 (ver Materiales y Métodos, apartado 1.5).

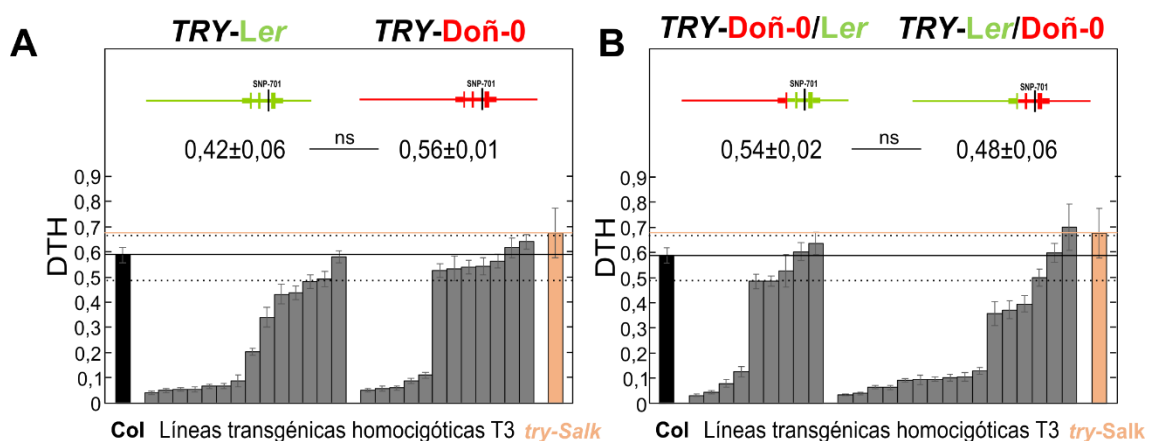


**Figura 30. Construcciones genómicas de *TRY*.** A) Construcciones genómicas de *TRY* de las accesiones Ler (verde) y Doñ-0 (rojo). B) Construcciones genómicas quiméricas Ler/Doñ-0 y Doñ-0/Ler. Las líneas verticales de color azul indican la posición de las dianas de restricción de los enzimas utilizados para generar las construcciones. La línea vertical negra indica la posición del SNP más asociado con NTF.

Para determinar si los alelos naturales de *TRY* de Ler y Doñ-0 son funcionales se obtuvieron líneas transgénicas en fondo genético *try-Salk*. Esta línea se seleccionó ya que posee un alelo de pérdida de función de *TRY* producido por una inserción de T-ADN que causa un patrón de tricomas en agregados en la superficie de la hoja (Schellmann *et al.*, 2002). En total, en esta tesis se generaron 14 y 12 líneas independientes y homocigóticas para los transgenes *TRY-Ler* y *TRY-Doñ-0* respectivamente, que se analizaron junto con las líneas control Col y *try-Salk* sin transformar.

El análisis fenotípico de estas líneas mostró que tanto las líneas portadoras del transgén *TRY-Ler* como del transgén *TRY-Doñ-0* presentan un patrón de tricomas en las hojas sin agregados similar a la línea control Col. Por lo tanto, ambos transgenes complementan el fenotipo de *try-Salk*, lo que demuestra que los dos alelos naturales de *TRY* de Ler y Doñ-0 son funcionales.

Con el fin de determinar si existen diferencias funcionales entre los dos alelos de *TRY* también se analizó el carácter DTH de forma cuantitativa. Este análisis mostró que el 50% de las líneas con el transgén *TRY-Ler* tienen un fenotipo prácticamente glabro (menos de 0,2 tricomas/mm<sup>2</sup>; Figura 31A), mientras que en las líneas con el transgén *TRY-Doñ-0* este porcentaje es inferior (40%). No obstante, la diferencia entre ambos porcentajes no fue estadísticamente significativa ( $P>0,05$ ). Además, los valores de DTH de ambos tipos de líneas transgénicas tampoco difirieron significativamente (Figura 31A). Por lo tanto, ambos alelos naturales de *TRY* no muestran diferencias funcionales en la fase vegetativa en este fondo genético.



**Figura 31. Patrón de tricomas de las líneas transgénicas de *TRY* en el fondo genético *try-Salk*.** A, B) Densidad de tricomas en la hoja (DTH) de las líneas transgénicas homocigóticas T3 portadoras de construcciones genómicas de *TRY-Ler* y *TRY-Doñ-0* (A) y de las construcciones quiméricas *TRY-Doñ-0/Ler* y *TRY-Ler/Doñ-0* (B) en el fondo genético *try-Salk*. Las líneas control Col (negro) y *try-Salk* (salmón) sin transformar se muestran a los lados de cada panel. Cada barra representa la media ± EE de cada línea estimada a partir de 10-18 plantas. En la parte superior de cada panel se muestra el esquema del transgén de las líneas y la media ± EE de las líneas transgénicas para cada construcción, indicando si difieren estadísticamente (ns:  $P>0,05$ ). Las líneas horizontales continuas indican el valor medio del carácter de la línea control Col (negro) y *try-Salk* (salmón) sin transformar y las discontinuas limitan el intervalo de confianza del 95% de Col. Las diferencias estadísticas entre las distintas líneas se contrastaron mediante ANOVA con la corrección de Bonferroni.

Para determinar si *TRY* es *FTD5*, también se obtuvieron líneas transgénicas en el fondo genético LI-235. Al igual que con *TCL1/FTD2*, la línea LI-235 se seleccionó

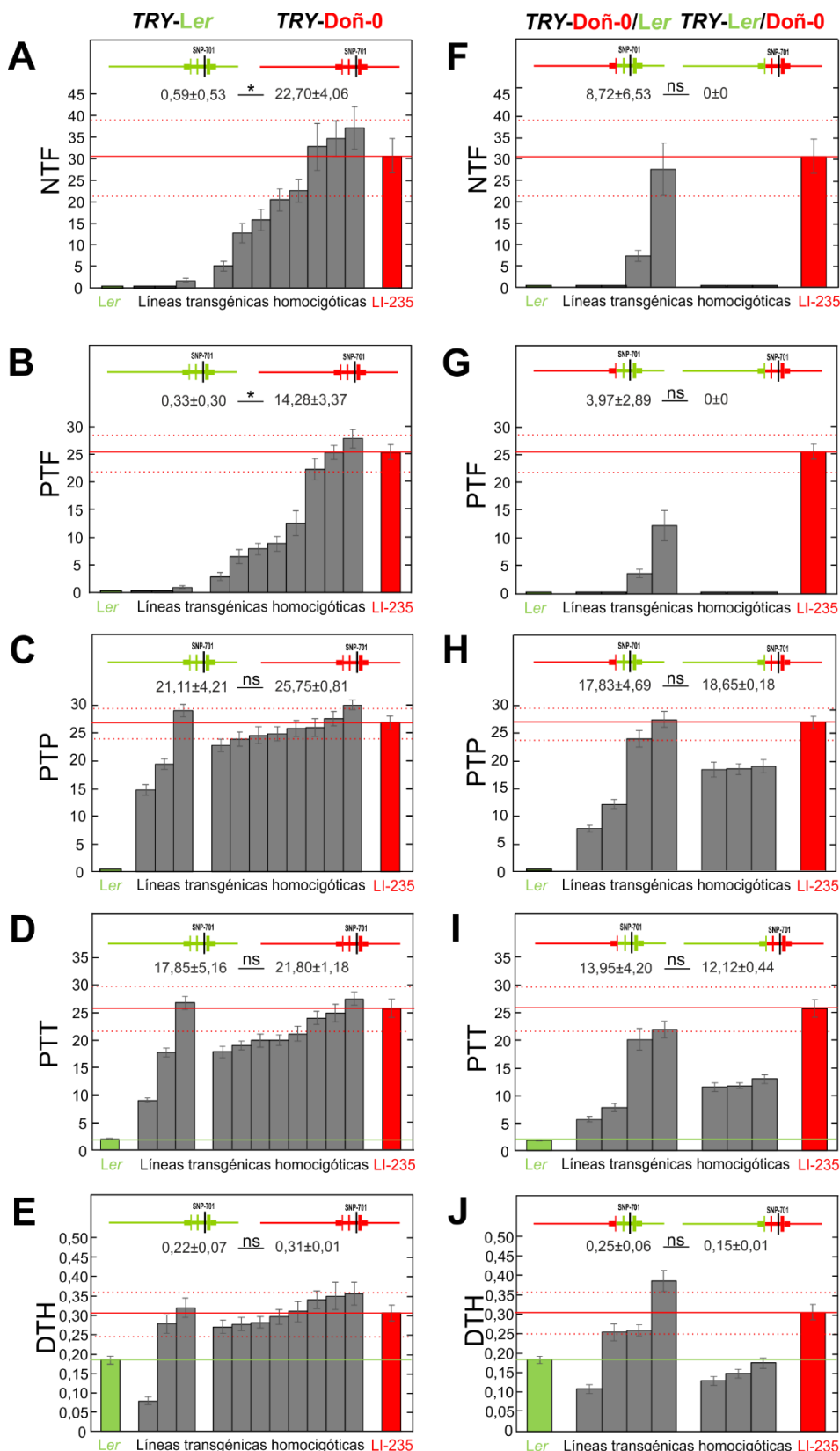
como fondo genético porque una alta densidad de tricomas en el fruto requiere alelos Doñ-0 en los tres loci *FTD2*, *FTD3* y *FTD5*. Puesto que *TRY* es un represor de la formación de tricomas, este fondo permitirá contrastar si las líneas con los transgenes *TRY-Ler* y *TRY-Doñ-0* difieren en la densidad de tricomas en el fruto. En total, en esta tesis se obtuvieron 3 y 8 líneas transgénicas portadoras de los transgenes *TRY-Ler* y *TRY-Doñ-0*, respectivamente, las cuales se analizaron para el patrón de tricomas junto con las líneas control *Ler* y *LI-235* sin transformar (Figura 32).

El análisis del carácter NTF mostró que la mitad de las líneas transgénicas con el transgén *TRY-Doñ-0* tienen un valor similar al de la línea *LI-235* sin transformar. Por el contrario, las 3 líneas con el transgén *TRY-Ler* no tienen tricomas en el fruto o su número es muy reducido (Figura 32A). El análisis de PTF mostró resultados similares, ya que todas las líneas con el transgén *TRY-Doñ-0* presentaron valores más altos que las líneas con el transgén *TRY-Ler* (Figura 32B). Ambas variables del patrón de tricomas del fruto, NTF y PTF, difirieron significativamente entre los dos tipos de líneas transgénicas ( $0,011 < P < 0,036$ ), y el efecto del transgén explicó el 48-46% de la varianza observada para estos caracteres. Estos resultados demuestran que los alelos naturales de *Ler* y *Doñ-0* de *TRY* difieren funcionalmente en fase reproductiva y que, por lo tanto, *TRY* es el gen correspondiente a *FTD5*.

Para determinar si ambos alelos naturales también afectan diferencialmente al patrón de tricomas en otros órganos se analizaron las variables PTP, PTT y DTH. El análisis de PTP mostró que los dos tipos de líneas transgénicas no difieren significativamente ( $P > 0,05$ ; Figura 32C) a pesar de que la mayoría de las líneas con *TRY-Doñ-0* no difirieron fenotípicamente de la línea *LI-235* sin transformar, mientras que 2 de las 3 líneas con el transgén *TRY-Ler* presentaron un valor de PTP significativamente inferior. De forma similar tampoco se observaron diferencias significativas entre los dos tipos de líneas transgénicas para los caracteres PTT y DTH ( $P > 0,05$ ; Figura 32D y 32E). Por lo tanto, las diferencias funcionales entre los dos alelos de *TRY* afectan únicamente al patrón de tricomas en el fruto.

Teniendo en cuenta que *TRY* es un represor del desarrollo de tricomas, estos análisis demuestran que *TRY* es el gen *FTD5*; que los dos alelos de *Ler* y *Doñ-0* son funcionales; y que el alelo de *Doñ-0* presenta una pérdida de función parcial en comparación con el de *Ler*.





**Figura 32. Patrón de tricomas de las líneas transgénicas de *TRY* en el fondo genético LI-235.** Número de tricomas en el fruto (NTF; A, F), patrón de tricomas en el fruto (PTF; B, G), patrón de tricomas en el pedicelo (PTP; C, H), patrón de tricomas en el tallo (PTT; D, I) y densidad de tricomas en la hoja (DTH; E, J) en líneas transgénicas homocigóticas T3/T4 portadoras de construcciones genómicas *TRY-Ler* y *TRY-Doñ-0* (A, B, C, D y E) y construcciones quiméricas *Doñ-0/Ler* y *Ler/Doñ-0* (F, G, H, I y J) en un fondo genético LI-235. Las líneas control sin

transformar *Ler* (verde) y LI-235 (rojo) se muestran a los lados de cada panel. Cada barra de cada panel representa la media  $\pm$  EE de cada línea estimada a partir de 10-18 plantas. En la parte superior de cada panel se muestra el esquema del transgén de las líneas y la media  $\pm$  EE de todas las líneas transgénicas para cada construcción, indicando si difieren o no estadísticamente (\*\*:  $P < 0,01$ ; \*:  $P < 0,05$ ; ns:  $P > 0,05$ ). Las líneas horizontales continuas indican el valor medio del carácter fenotípico en las líneas control *Ler* (verde) y LI-235 (rojo) y las discontinuas limitan el intervalo de confianza al 95% de LI-235. Las diferencias estadísticas entre las distintas líneas se contrastaron mediante ANOVA con la corrección de Bonferroni.

### 3.3.2. Identificación del polimorfismo funcional de *TRY/FTD5*

El análisis de las secuencias nucleotídicas del gen *TRY* en *Ler* y Doñ-0 realizado previamente en el laboratorio (Savic, 2015) identificó 23 SNPs y 13 indels de pequeño tamaño en la región promotora; 6 SNPs y 2 pequeños indels en la región codificante; y 11 SNPs, 2 pequeños indels y una inserción de 504 pb en Doñ-0 en la región 3' UTR. Cabe destacar que la región intergénica 5' de *TRY* es de gran tamaño ( $> 8$  kb) y que entre los polimorfismos detectados en la región codificante ninguno implica una modificación de la secuencia aminoacídica de *TRY*. Además, análisis previos de asociación entre los polimorfismos de la región codificante y la presencia de tricomas en el fruto identificó un polimorfismo del segundo intrón (posición 701) asociado muy significativamente (Savic, 2015).

Con el fin de localizar el polimorfismo responsable de las diferencias funcionales de *TRY* entre *Ler* y Doñ-0, se obtuvieron previamente otras dos construcciones quiméricas mediante el intercambio recíproco del segmento *Kpn* I/*Sal* I entre las construcciones parentales (Figura 30B; ver Materiales y Métodos, apartado 6). Estas construcciones se utilizaron para transformar los fondos genéticos *try-Salk* y LI-235 y en esta tesis se han obtenido y caracterizado 25 y 7 líneas independientes y homocigóticas para cada fondo respectivamente (ver Materiales y Métodos, apartado 1.5).

El análisis fenotípico de las líneas transgénicas de fondo *try-Salk* para el carácter DTH mostró un alto porcentaje de líneas con un fenotipo prácticamente glabro: 62% de las líneas con *TRY-Ler/Doñ-0* y 44% con *TRY-Doñ-0/Ler*. Sin embargo, como con el análisis de los transgenes parentales (*TRY-Ler* y *TRY-Doñ-0*), ambos tipos de líneas no difirieron ni en el porcentaje de plantas glabras ni en los valores de DTH (Figura 31B).

El análisis de las líneas transgénicas de fondo genético LI-235 para las variables NTF y PTF mostró que la mitad de las líneas con el transgén *TRY-Doñ-0/Ler* tiene tricomas en el fruto, mientras que todas las líneas con el transgén *TRY-Ler/Doñ-0* fueron glabras en este órgano (Figura 32F y 32G). No obstante, las diferencias entre los transgenes no fueron estadísticamente significativas ( $P=0,3$ ). La comparación de valor de NTF entre las líneas con los transgenes parentales y con las construcciones

quiméricas mostró que las líneas con el transgén *TRY-Ler* y *TRY-Ler/Doñ-0* no difieren entre ellas, pero las líneas con *TRY-Doñ/Ler* difirieron tanto de las líneas con *TRY-Ler* como con *TRY-Doñ-0* ya que mostraron un fenotipo intermedio. El análisis de PTP, PTT y DTH tampoco mostró diferencias significativas entre los dos tipos de líneas transgénicas ( $0,19 < P < 0,89$ ) (Figura 32H-32J).

Todos estos resultados indican que las diferencias funcionales en el gen *TRY* que afectan al patrón de tricomas en fase reproductiva se localizan tanto en la región 5' promotora como en la región 3' de la construcción (donde se ubica el SNP-701 previamente asociado), ya que las líneas transgénicas con los transgenes quiméricos presentan valores distintos de las líneas con el alelo de *Doñ-0*.

### **3.3.3. Caracterización funcional de la variación natural *Ler/Doñ-0* de *TRY***

Con el fin de caracterizar el efecto de los polimorfismos reguladores de *TRY/FTD5* se analizó la expresión de *TRY* en las líneas transgénicas de los fondo *try-Salk* y LI-235. Este análisis se realizó mediante RT-PCR siguiendo el procedimiento descrito para el gen *TCL1* (ver Materiales y Métodos, apartado 7).

El análisis de la expresión de *TRY* en las líneas transgénicas de fondo genético *try-Salk* se llevó a cabo con 4 a 8 líneas transgénicas de cada una de las cuatro construcciones genómicas y con las líneas control (Tabla 16). En general, Col sin transformar y las líneas transgénicas presentaron una mayor expresión de *TRY* en las inflorescencias que en las hojas, aunque estas diferencias fueron muy reducidas y no significativas para las líneas con el transgén *TRY-Doñ-0*. El análisis de las líneas transgénicas mostró que, tanto en fase vegetativa como en fase reproductiva, las líneas con el transgén de *Ler* presentaron más del doble de expresión que las líneas con el transgén de *Doñ-0* ( $P < 0,001$ ). Sin embargo, las líneas con los transgenes quiméricos no difirieron significativamente (Tabla 16).

El análisis de la expresión de *TRY* en las líneas transgénicas de fondo LI-235 se llevó a cabo con 3 y 4 líneas para cada transgén (Tabla 16). La comparación de estas líneas mostró que el nivel de expresión de *TRY* de las líneas con *TRY-Ler* es cinco veces mayor que el de las líneas con el transgén *TRY-Doñ-0* en fase vegetativa ( $P = 0,005$ ) y tres veces mayor en fase reproductiva ( $P < 0,001$ ). Por el contrario, las líneas quiméricas no mostraron diferencias significativas en la expresión de *TRY* en ninguna de las dos fases, en concordancia con lo observado en las líneas transgénicas de fondo *try-Salk*. Sin embargo, a diferencia de lo observado en el fondo *try-Salk*, las líneas con

el transgén parental *TRY-Ler* presentaron una expresión inferior en fase reproductiva que en fase vegetativa, mientras que las líneas con el transgén *TRY-Doñ-0* mostraron una expresión similar en ambas fases.

Fondo genético	Transgén	Expresión <i>TRY</i>				
		Hoja	<i>P</i>	Inflorescencia	<i>P</i>	<i>N</i>
<i>try-Salk</i>	<i>Ler</i>	58,17±16,00	**	73,07±20,79	**	7
	<i>Doñ-0</i>	20,74±10,36		23,95±18,08		4
	<i>Doñ-0/Ler</i>	47,87±15,03	ns	88,20±23,97	ns	5
	<i>Ler/Doñ-0</i>	58,97±12,94		67,49±16,02		8
<i>Col</i>		4,98±0,52	**	8,40±0,14	**	3
<i>try-Salk</i>		1,00±0,10		0,36±0,06		3
LI-235	<i>Ler</i>	10,33±5,58	**	5,96±4,12	**	3
	<i>Doñ-0</i>	2,20±1,10		2,04±0,70		4
	<i>Doñ-0/Ler</i>	9,69±4,51	ns	12,65±4,78	ns	4
	<i>Ler/Doñ-0</i>	11,82±3,12		16,51±2,71		3
<i>Ler</i>		1,61±0,19	*	2,13±0,15	**	3
LI-235		1,00±0,09		1,35±0,10		2

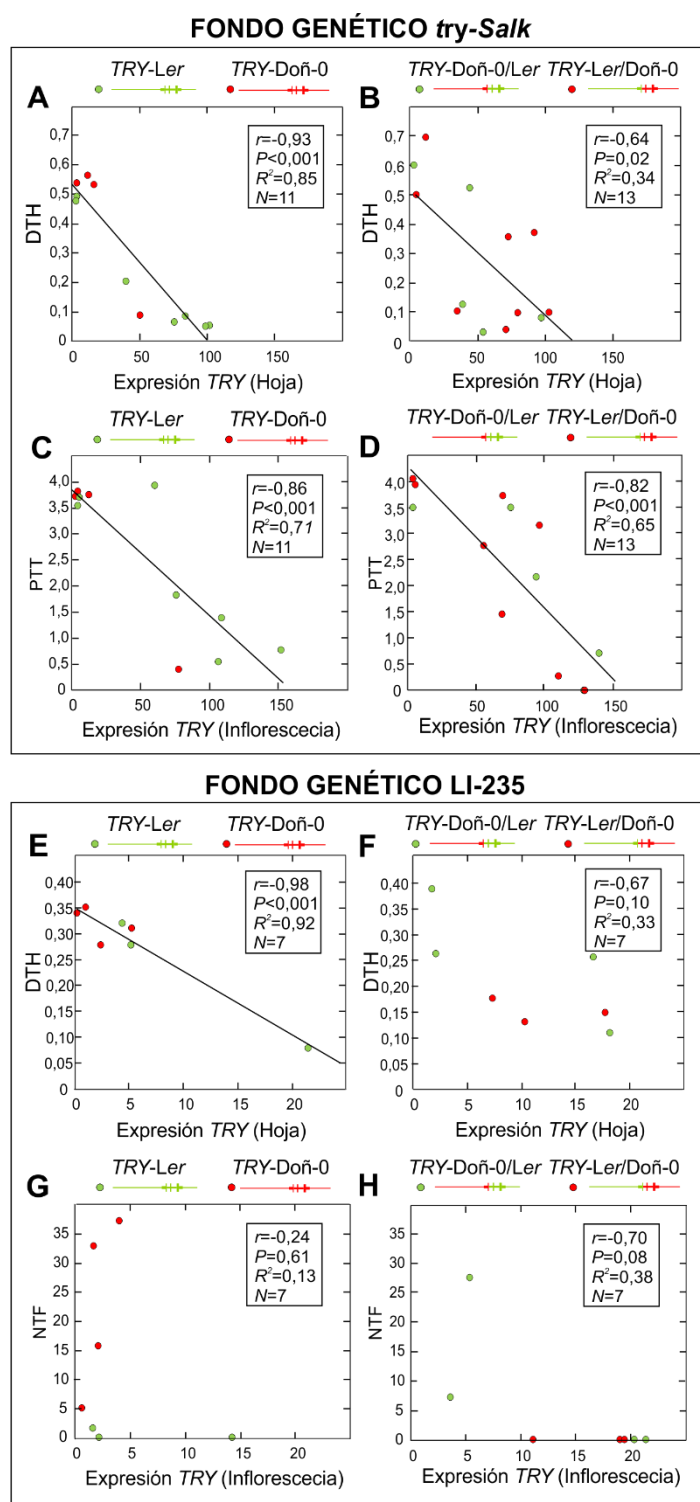
**Tabla 16. Expresión del gen *TRY* en líneas transgénicas.** La tabla muestra el fondo genético, el transgén, la expresión media  $\pm$  EE de *TRY* en la hoja y en la inflorescencia y el número de líneas analizadas para cada tipo de transgén. Los niveles de expresión de cada línea transgénica en los fondos *try-Salk* y LI-235 y las líneas control están relativizados con respecto a la expresión media de las líneas control sin transformar *try-Salk* o LI-235, respectivamente en fase vegetativa. Las diferencias estadísticas entre la expresión media en las líneas transgénicas o las líneas control se contrastaron mediante ANOVA (\*\*:  $P < 0,01$ ; \*:  $P < 0,05$ ; ns:  $P > 0,05$ ).

Estos resultados demuestran que los alelos *Ler* y *Doñ-0* de *TRY* difieren en su expresión, confirmando la existencia de polimorfismos que afectan a la regulación en *cis* de *TRY*. Además, el alelo de *Doñ-0* se caracteriza por una menor variación de la expresión a lo largo del desarrollo en comparación con el alelo de *Ler* cuya expresión es mayor en fase reproductiva. Por otro lado, las líneas con los transgenes quiméricos no mostraron diferencias de expresión y sus valores medios fueron similares a los de las líneas con el transgén de *Ler*. Como en el caso de los fenotipos del patrón de tricomas, el comportamiento diferente de las líneas con los transgenes parentales y quiméricos sugiere la existencia de polimorfismos reguladores que afectan a la expresión de *TRY*, tanto en el promotor como en la región 3' de *TRY*.

Finalmente, para determinar la relación entre la expresión de *TRY* y el patrón de tricomas se han estimado los coeficientes de correlación de Pearson entre la expresión de *TRY* en la hoja o en la inflorescencia y su patrón de tricomas correspondiente (Figura 33).

El análisis de las líneas de fondo *try-Salk* mostró que la mayoría de las líneas con el transgén *TRY-Doñ-0* tienen una expresión baja y valores de DTH o PTT altos, mientras que las líneas con *TRY-Ler* mostraron en su mayoría una expresión más alta y valores menores de DTH o PTT (Figura 33A y 33C). En las líneas con los transgenes quiméricos se observó una correlación similar pero menor. De esta forma, el nivel de expresión de *TRY* mostró una correlación negativa y significativa con el patrón de tricomas en la hoja y la inflorescencia ( $0,64 < r < 0,93$ ;  $0,001 < P < 0,02$ ), lo que explicó entre el 34 y el 85% de la variación fenotípica. Por lo tanto, la variación natural en la expresión de *TRY* afecta a la variación fenotípica provocada por *FTD5*.

De igual forma en el fondo genético LI-235 la mayoría de las líneas con el transgén *TRY-Doñ-0* presentaron una expresión de *TRY* más baja y un valor fenotípico más alto que las líneas con el transgén *TRY-Ler* (Figura 33E y 33G). Estos comportamientos explican la correlación negativa observada en las líneas con los transgenes parentales en fase vegetativa ( $r = -0,98$ ;  $P < 0,001$ ) (Figura 33E). En fase reproductiva se observó una tendencia similar, aunque esta correlación no fue significativa probablemente debido al bajo número de líneas transgénicas analizadas. Además, las líneas con los transgenes quiméricos no mostraron correlaciones significativas entre los niveles de expresión y los patrones de tricomas (Figura 33F y 33H). Por lo tanto, los polimorfismos de *TRY* entre *Ler* y *Doñ-0* explican parte de las diferencias fenotípicas observadas en las líneas transgénicas parentales a través de su efecto sobre la regulación en *cis* de la expresión de *TRY*.



**Figura 33. Relación entre el nivel de expresión de *TRY* y el patrón de tricomas de las líneas transgénicas.** Relación entre la expresión de *TRY* en la hoja y DTH (A, B, E, F) o la expresión de *TRY* en la inflorescencia y PTT o NTF (C, D, G, H) en las líneas transgénicas de fondo genético *try-Salk* (A, B, C, D) y LI-235 (E, F, G, H). A, C, E y G muestran la relación en las líneas con transgenes *TRY* parentales, y B, D, F, H la relación en las líneas transgénicas con los transgenes quiméricos. Dentro de cada panel se indica el índice de correlación de Pearson ( $r$ ), la significación estadística de la correlación ( $P$ ), la proporción de varianza explicada por la relación entre las dos variables comparadas ( $R^2$ ) y el número de muestras analizadas ( $N$ ). Los valores de expresión de cada línea transgénica en fondo *try-Salk* o LI-235 están relativizadas frente al valor medio de la línea control *try-Salk* o LI-235 sin transformar en fase vegetativa, respectivamente.

### 3.4. Aislamiento y caracterización molecular de *FTD3*

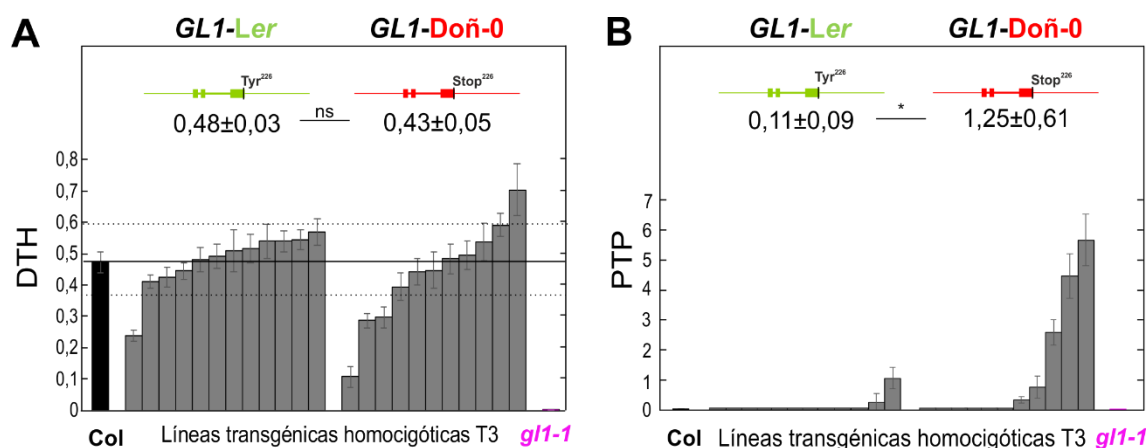
#### 3.4.1. Aislamiento de *FTD3*

Con el fin de demostrar si *GL1* (*At3g27920*) es *FTD3* se utilizaron dos construcciones genómicas obtenidas previamente, correspondientes a los alelos *Ler* y *Doñ-0* de *GL1* y que incluyen 1,6 kb del promotor, la región codificante completa (1,5 kb) y 1,9 kb de la región 3' (ver Materiales y Métodos, apartado 6). Estas construcciones se utilizaron para generar plantas transgénicas en dos fondos genéticos: el mutante simple *gl1-1* y el mutante doble *tcl1-Salk try-Salk* (ver Materiales y Métodos, apartado 1.5).

Para determinar si los alelos naturales de *Ler* y de *Doñ-0* de *GL1* son funcionales se obtuvieron las líneas transgénicas de fondo genético *gl1-1* ya que es un alelo de pérdida de función debido al cambio de aminoácido Ser<sup>92</sup> (Savic, 2015) que provoca un fenotipo glabro. En total se analizaron 12 y 11 líneas independientes y homocigóticas para los transgenes *GL1-Ler* y *GL1-Doñ-0*, respectivamente.

El análisis fenotípico de estas líneas transgénicas para el carácter DTH mostró que todas ellas, independientemente del transgén de *GL1*, desarrollaron tricomas en las hojas (Figura 34A). Por lo tanto, los alelos naturales de *GL1* de *Ler* y *Doñ-0* son funcionales ya que ambos transgenes complementan el fenotipo glabro de *gl1-1*. El análisis cuantitativo de DTH mostró que las líneas con el transgén *GL1-Doñ-0* y *GL1-Ler* no difieren significativamente ( $P > 0,05$ ) (Figura 34A), por lo que ambos alelos naturales de *GL1* no mostraron diferencias funcionales en fase vegetativa.

Para determinar si los alelos de *GL1* de *Ler* y *Doñ-0* afectan al patrón de tricomas en fase reproductiva también se analizaron las variables PTT y PTP. Los dos tipos de líneas transgénicas no difirieron en sus valores de PTT por lo que no mostraron diferencias funcionales en el tallo. Sin embargo, el análisis del patrón de tricomas en el pedicelo mostró que ambos tipos de líneas difieren significativamente ( $P = 0,048$ ; Figura 34B) y estas diferencias explicaron el 14% de la varianza fenotípica observada para el carácter PTP. Por lo tanto, los alelos de *Ler* y *Doñ-0* de *GL1* difieren funcionalmente en sus efectos sobre el valor de PTP. Teniendo en cuenta que *GL1* es un promotor del desarrollo de tricomas (Oppenheime *et al.*, 1991) estos resultados indican que el alelo de *GL1* de *Doñ-0* es más activo que el de *Ler* en fase reproductiva.



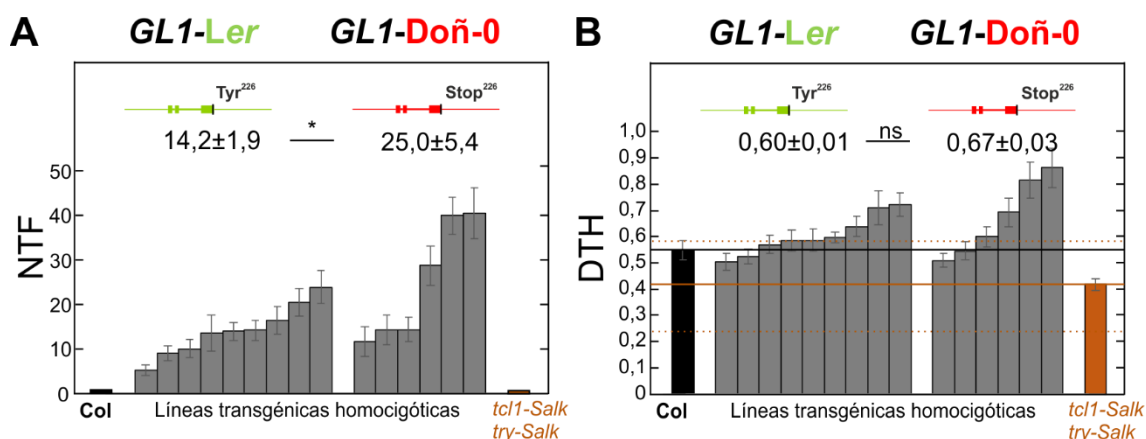
**Figura 34. Patrón de tricomas de las líneas transgénicas de *GL1* en el fondo genético *g/1-1*.** A) Densidad de tricomas en la hoja (DTH) y B) patrón de tricomas en el pedicelo (PTP) de las líneas transgénicas homocigóticas T3 para las construcciones genómicas de *GL1-Ler* y *GL1-Doñ-0* en el fondo genético *g/1-1*. Las líneas control Col (negro) y *g/1-1* (rosa) sin transformar se muestran a los lados de cada panel. Cada barra de cada panel representa la media  $\pm$  EE de una línea estimada a partir de 10-18 plantas. En la parte superior de cada panel se muestra el esquema del transgén de las líneas y la media  $\pm$  EE de todas las líneas transgénicas de cada construcción indicando si difieren o no estadísticamente (\*:  $P < 0,05$ ; ns:  $P > 0,05$ ). La línea horizontal continua indica el valor medio del carácter de la línea control Col (negro) y las discontinuas limitan el intervalo de confianza del 95% de Col. Las diferencias estadísticas entre las distintas líneas se contrastaron mediante ANOVA con la corrección de Bonferroni.

Para determinar si *GL1* es *FTD3* se transformó la línea doble mutante *tcl1-Salk try-Salk*. Esta línea se seleccionó ya que, como se ha demostrado en los apartados anteriores, los alelos de pérdida de función en los genes represores *TCL1* y *TRY* son necesarios para desarrollar tricomas en el fruto. Es de esperar que un fondo genético con alelos de pérdida de función de estos dos represores permita contrastar si hay diferencias funcionales entre los alelos de *Ler* y *Doñ-0* de *GL1* para el desarrollo de tricomas en el fruto.

En total se generaron 9 y 6 líneas independientes y homocigóticas para los transgenes *GL1-Ler* y *GL1-Doñ-0*, respectivamente. El análisis de estas líneas para el carácter NTF mostró que todas ellas desarrollan tricomas en el fruto, lo que indica que la sobreexpresión de *GL1* en ausencia de *TRY* y *TCL1*, es suficiente para inducir el desarrollo de tricomas en este órgano. Además, los dos tipos de líneas difirieron significativamente ( $P=0,04$ ; Figura 35A), ya que la mitad de las líneas con el transgén *GL1-Doñ-0* tienen mayor valor de NTF que las líneas con *GL1-Ler*. En conjunto, el tipo de transgén explicó el 18% de la varianza fenotípica observada para NTF, lo que indica que los alelos naturales de *Ler* y *Doñ-0* de *GL1* difieren en su función sobre la formación de tricomas en el fruto. Por lo tanto, *GL1* es el gen correspondiente a *FTD3* de manera que el alelo de *GL1* de *Doñ-0* es más activo que el de *Ler* en fase reproductiva, en concordancia con los resultados de las líneas transgénicas en fondo *g/1-1*.



El análisis del carácter PTP mostró que todas las líneas transgénicas presentaron un patrón similar con tricomas en el pedicelo, ya que son portadoras del alelo de pérdida de función *tcl1-Salk*. Igualmente, estas líneas se caracterizaron por su patrón de tricomas agregados en la hoja como consecuencia de la mutación *try-Salk*. Por otro lado, el análisis cuantitativo de la densidad de tricomas en la hoja de estas líneas no mostró diferencias significativas ( $P>0,05$ ; Figura 35B). Por lo tanto, los alelos naturales de *GL1* de Doñ-0 y *Ler* únicamente difieren funcionalmente para su efecto sobre el patrón de tricomas en el fruto.



**Figura 35. Patrón de tricomas de las líneas transgénicas de *GL1* en el fondo genético *tcl1-Salk try-Salk*.** A) Número de tricomas en el fruto (NTF) y B) densidad de tricomas en la hoja (DTH) en líneas transgénicas homocigóticas T3 para las construcciones genómicas de *GL1-Ler* y *GL1-Doñ-0* en el fondo genético *tcl1-Salk try-Salk*. Las líneas control Col (negro) y *tcl1-Salk try-Salk* (marrón) sin transformar se muestran a los lados de cada panel. Cada barra de cada panel representa la media ± EE de una línea estimada a partir de 15-20 plantas. En la parte superior de cada panel se muestra el esquema del transgén de las líneas y la media ± EE de todas las líneas transgénicas para cada construcción indicando la diferencia estadística (\*:  $P<0,05$ ; ns:  $P>0,05$ ). Las diferencias estadísticas entre las distintas líneas se contrastaron mediante ANOVA con la corrección de Bonferroni.

### 3.4.2. Caracterización funcional de la variación natural *Ler/Doñ-0* de *GL1*

El análisis de las secuencias nucleotídicas de *GL1*, realizado previamente en el laboratorio, mostró que las accesiones *Ler* y *Doñ-0* difieren en 3 polimorfismos que afectan a la estructura de la proteína *GL1*. Dos de ellos producen cambios de aminoácidos (Pro/Arg<sup>177</sup>, Val/Phe<sup>224</sup>), mientras que el tercero causa un codón de parada prematuro en *Doñ-0*, en el antepenúltimo aminoácido (Tyr<sup>226</sup>). Además, *Ler* y *Doñ-0* difieren en 6 SNPs y 1 indel de pequeño tamaño situados en la región codificante, y en 25 SNPs y 5 pequeños indels de las regiones promotora y 3' UTR (Savic, 2015). Para determinar si el polimorfismo funcional de *GL1/FTD3* es regulador se ha analizado la expresión génica de *GL1* en las líneas transgénicas de fondo genético *gl1-1* y *tcl1-Salk*

*try-Salk*, siguiendo el mismo protocolo descrito en los análisis de expresión anteriores (ver Materiales y Métodos, apartado 7).

La expresión de *GL1* en las plantas control de Col fue significativamente superior a la del mutante *gl1-1* en fase vegetativa ( $P < 0,001$ ), pero no en fase reproductiva donde *GL1* se expresa 5 veces menos que en las hojas (Tabla 17). De igual forma, las líneas Col y *try-Salk tcl1-Salk* difirieron en la expresión de *GL1* tanto en fase reproductiva como vegetativa ( $P < 0,001$ ) presentando también una menor expresión en inflorescencias que en las hojas. Sin embargo, el análisis de 12 líneas transgénicas para *GL1-Ler* y *GL1-Doñ-0* en fondo genético *gl1-1* mostró que ambos tipos de líneas no difieren en los niveles de expresión de *GL1*, ni en tejido vegetativo ni reproductivo ( $0,06 < P < 0,68$ ). Igualmente, el análisis de la expresión de *GL1* en 15 líneas transgénicas en fondo *try-Salk tcl1-Salk* tampoco mostró diferencias significativas en las inflorescencias, aunque sí en las hojas (Tabla 17).

Fondo genético	Transgén	Expresión <i>GL1</i>				N
		Hoja	P	Inflorescencia	P	
<i>gl1-1</i>	Ler	102,32±19,22	ns	185,05±30,6	ns	7
	Doñ-0	93,58±37,22		82,15±36,84		5
Col	<i>gl1-1</i>	3,93±0,38	**	0,80±0,13	ns	3
		1,00±0,11		0,67±0,07		3
<i>tcl1-Salk try-Salk</i>	Ler	2,61±0,24	**	1,97±0,56	ns	9
	Doñ-0	3,82±1,17		1,56±0,34		6
Col	<i>tcl1-Salk try-Salk</i>	0,44±0,04	**	0,02±0,004	**	3
		1,00±0,07		0,04±0,01		2

**Tabla 17. Expresión del gen *GL1* en líneas transgénicas.** La tabla muestra el fondo genético, el transgén, la expresión media  $\pm$  EE de *GL1* en hoja e inflorescencia y el número de líneas analizadas para cada tipo de transgén o línea sin transformar. Los niveles de expresión de las líneas transgénicas en fondo *gl1-1* o *tcl1-Salk try-Salk*, así como los controles, están relativizados con respecto a la expresión media de *gl1-1* o *tcl1-Salk try-Salk* en fase vegetativa, respectivamente. Las diferencias estadísticas entre las distintas líneas se contrastaron mediante ANOVA (\*\*:  $P < 0,01$ ; \*:  $P < 0,05$ ; ns:  $P > 0,05$ )

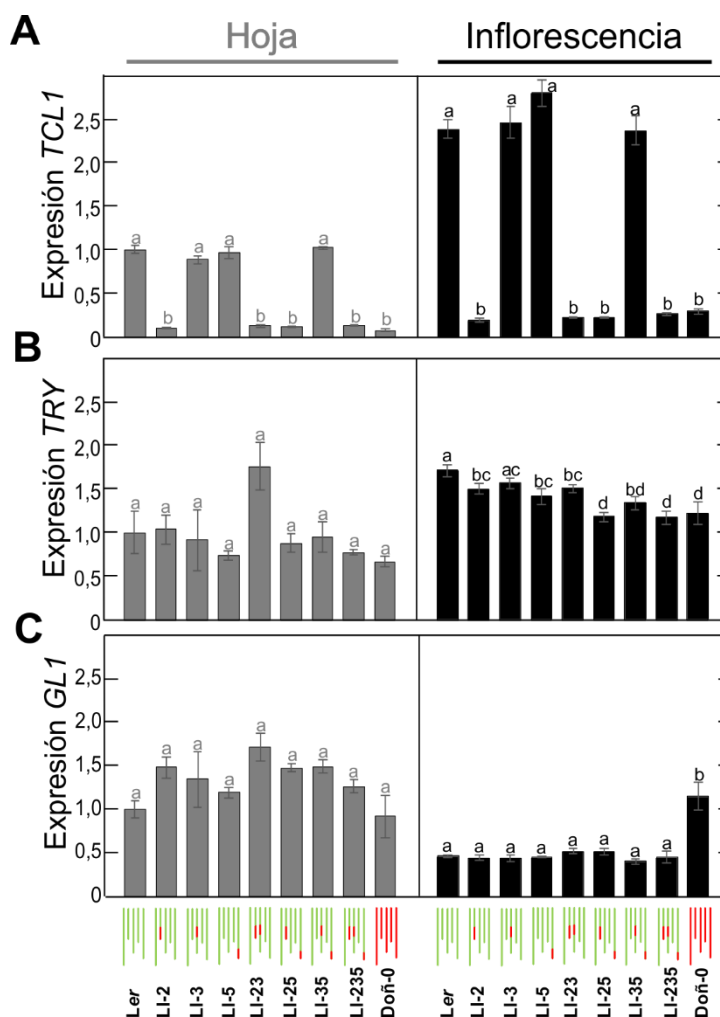
Por último, para determinar si existe relación entre el nivel de expresión de *GL1* y el patrón de tricomas en fase vegetativa o reproductiva se estimaron los coeficientes de correlación de Pearson entre el nivel de expresión de *GL1* y la densidad de tricomas en la hoja (DTH), el patrón de tricomas en el tallo (PTT), o el número de tricomas en el fruto (NTF) en las líneas transgénicas de los dos fondos genéticos. Estos análisis no detectaron ninguna correlación significativa entre la expresión de *GL1* y las variables fenotípicas en las líneas transgénicas, tanto en fondo *gl1-1* ( $0,08 < P < 0,14$ ) como fondo *try-Salk tcl1-Salk* ( $0,41 < P < 0,49$ ).

En conjunto, estos análisis indican que las diferencias funcionales entre los alelos de Ler y Doñ-0 de *GL1* responsable del locus *FTD3* sobre el patrón de tricomas en el fruto no se deben a polimorfismos que afecten a la regulación en *cis* de *GL1*. Por el contrario, estos resultados sugieren que alguno de los polimorfismos estructurales de *GL1* es responsable de las diferencias funcionales entre los dos alelos naturales. No obstante, no podemos descartar que ambos alelos difieran en algún polimorfismo que afecte a la regulación de su expresión en fase vegetativa, dependiente del fondo genético.

### **3.5. Análisis de expresión de los genes *TCL1*, *TRY* y *GL1* en líneas de introgresión**

Una vez demostrado que *TCL1*, *TRY* y *GL1* son *FTD2*, *FTD5* y *FTD3*, respectivamente, se ha estudiado si la variación natural en estos loci afecta a la propia expresión (regulación en *cis*) o a la expresión de los otros genes (regulación en *trans*). Para ello se analizó la expresión de los tres genes en 21 líneas de introgresión que corresponden a siete genotipos distintos portadores de alelos Doñ-0 en uno a tres loci *FTD2*, *FTD3* y *FTD5* en un fondo genético Ler. Este análisis se realizó tanto en hoja como en inflorescencia (ver Materiales y Métodos, apartado 7), e incluyó las líneas parentales Ler y Doñ-0 como controles (Figura 36). Además, para determinar el efecto que ejercen los distintos loci *FTD* sobre la expresión de cada gen se llevó a cabo un análisis de la varianza en el que se utilizaron los tres loci *FTD2*, *FTD3* y *FTD5* como factores y que incluyó los efectos aditivos de cada factor y las interacciones dobles o triples entre ellos (Tabla 18). Por último, también se analizó la correlación entre el fenotipo de las líneas de introgresión para los caracteres DTH y NTF (ver Resultados, apartado 3.1) y los niveles de expresión de los tres genes en hoja e inflorescencia, respectivamente (Figura 37).

El análisis de la expresión de *TCL1* mostró que se expresa de 8 a 15 veces menos en las líneas con alelos Doñ-0 en *FTD2* que en las líneas con alelos Ler, tanto en fase reproductiva como vegetativa (Figura 36A), en concordancia con los resultados del análisis de las líneas transgénicas de *TCL1*. Además, el análisis de la varianza con los tres loci como factores mostró que *FTD2* es el único locus que afecta significativamente a la expresión de *TCL1* ( $P < 0,001$ ), tanto en fase reproductiva como vegetativa, explicando el 98% de las diferencias de expresión observadas (Tabla 18).



**Figura 36. Expresión de *TCL1*, *TRY* y *GL1* en las líneas de introgresión.** A, B, C) Niveles de expresión génica de *TCL1* (A), *TRY* (B) y *GL1* (C) en fase vegetativa (panel izquierdo) y fase reproductiva (panel derecho), en las líneas de introgresión y las accesiones parentales Ler y Doñ-0. Cada barra representa el valor medio  $\pm$  EE de tres líneas de introgresión independientes con el mismo genotipo. Las líneas parentales Ler y Doñ-0 se presentan a los lados de cada panel. Los genotipos con niveles de expresión similares se indican con la misma letra y los que difieren significativamente con diferentes letras. Los análisis estadísticos se realizaron mediante ANOVA. En la parte inferior se muestra el genotipo gráfico de las líneas de introgresión, representando en color rojo las regiones de Doñ-0 y en color verde las de Ler.

El análisis de la relación entre el patrón de tricomas y la expresión de *TCL1* mostró que todas las líneas con alelos Ler en *TCL1/FTD2* tienen una expresión muy alta en inflorescencia y no tienen tricomas en el fruto, mientras que todas las líneas con alelos Doñ-0 en *TCL1/FTD2* tienen baja expresión. Puesto que sólo las líneas con alelos Doñ-0 en los tres loci *FTD2*, *FTD3* y *FTD5* presentan alta densidad de tricomas en el fruto no se detectó correlación entre la expresión de *TCL1* en inflorescencias y el valor de NTF (Figura 37A). Por el contrario, sí que se encontró una correlación negativa entre la expresión de *TCL1* en las hojas y el valor de DTH ( $r=-0,78$ ;  $P<0,001$ ) de manera que las líneas con alelos Doñ-0 en *FTD2* presentan menor expresión de *TCL1* y mayor DTH

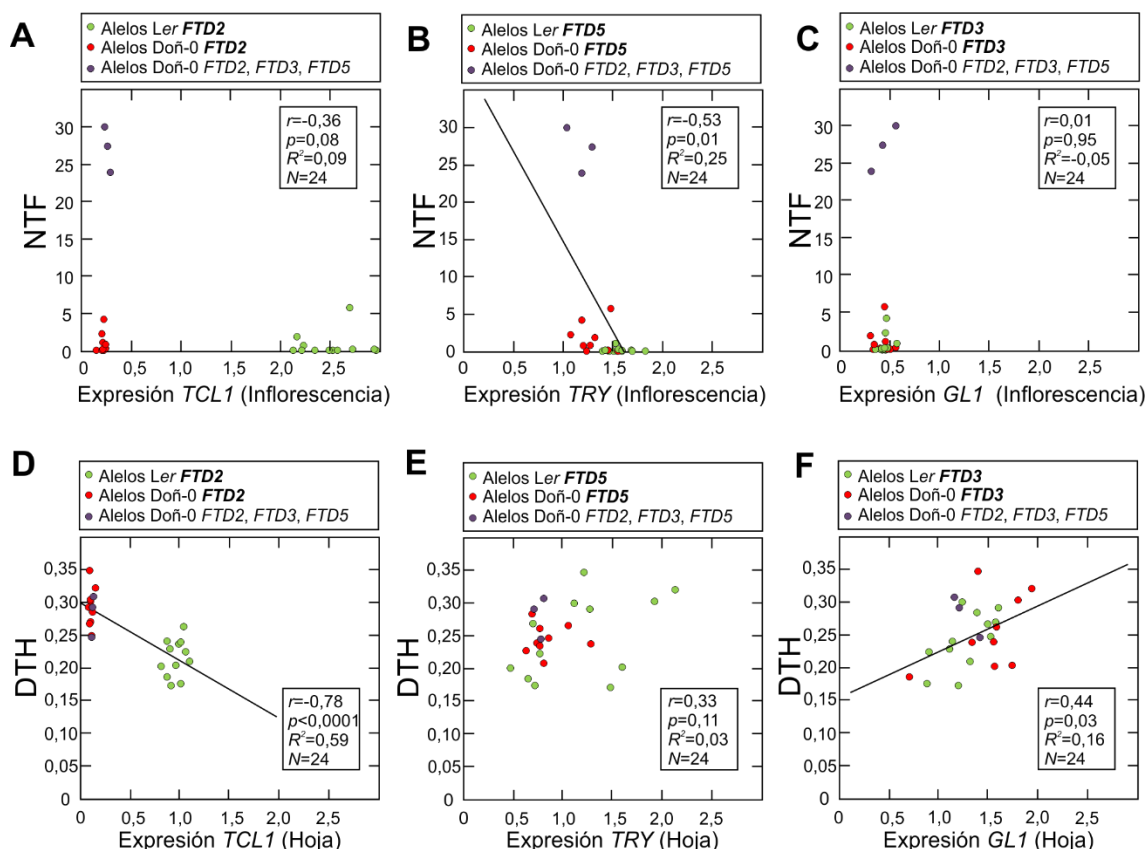
que las líneas con alelos *Ler* (Figura 37D). Estos resultados indican que sólo los polimorfismos entre *Ler* y Doñ-0 de *TCL1* afectan a la regulación de su expresión en *cis*, mientras que la variación natural *Ler*/Doñ-0 en *GL1/FTD3* y *TRY/FTD5* no contribuye a la regulación en *trans* de este gen.

	Expresión génica en inflorescencia						Expresión génica en hoja					
	<i>TCL1</i>		<i>TRY</i>		<i>GL1</i>		<i>TCL1</i>		<i>TRY</i>		<i>GL1</i>	
	<i>P</i>	% varianza explicada	<i>P</i>	% varianza explicada	<i>P</i>	% varianza explicada	<i>P</i>	% varianza explicada	<i>P</i>	% varianza explicada	<i>P</i>	% varianza explicada
<i>FTD2</i>	**	97,80%	**	17,10%	ns	-	**	98,10%	ns	-	ns	-
<i>FTD3</i>	ns	-	ns	-	ns	-	ns	-	ns	-	ns	-
<i>FTD5</i>	ns	-	**	58,40%	ns	-	ns	-	*	17%	ns	-
<i>FTD2</i> x <i>FTD3</i>	ns	-	ns	-	ns	-	ns	-	ns	-	ns	-
<i>FTD2</i> x <i>FTD5</i>	ns	-	ns	-	ns	-	ns	-	ns	-	ns	-
<i>FTD3</i> x <i>FTD5</i>	ns	-	ns	-	ns	-	ns	-	ns	-	ns	-
<i>FTD2</i> x <i>FTD3</i> x <i>FTD5</i>	ns	-	ns	-	ns	-	ns	-	ns	-	ns	-

**Tabla 18. Efecto de los loci *FTD* sobre la expresión de *TCL1*, *TRY* y *GL1*.** La tabla muestra el efecto aditivo de los loci *FTD* y de las interacciones de dos o tres loci *FTD* sobre la expresión de *TCL1*, *TRY* o *GL1*, tanto en fase reproductiva (panel izquierdo) como vegetativa (panel derecho). Los efectos de los loci *FTD* sobre la expresión de los genes fueron contrastados mediante un ANOVA completo con tres factores. Para cada efecto contrastado se muestra la significación (\*\*:  $P < 0,01$ ; \*:  $P < 0,05$ ; ns:  $P > 0,05$ ) y el porcentaje de varianza explicado en el caso de los efectos significativos.

El estudio similar de la expresión de *TRY* mostró que las líneas de introgresión con alelos Doñ-0 en *TRY/FTD5* presentan una reducción significativa de la expresión de *TRY* en fase reproductiva en comparación con las líneas con alelos *Ler* (Figura 36B), en concordancia con el análisis de las líneas transgénicas de *TRY*. Además, el análisis de varianza con los tres loci *FTD* como factores mostró que *TCL1/FTD2* también afecta a la expresión de *TRY* de forma aditiva con *TRY/FTD5* (Tabla 18), de manera que las líneas con alelos Doñ-0 en ambos loci tienen menor expresión que las líneas con alelos en un solo locus. *TRY/FTD5* y *TCL1/FTD2* explicaron el 58,4 y el 17,1% de la varianza para la expresión de *TRY* en fase reproductiva, pero sólo *TRY/FTD5* afectó a la expresión de *TRY* en fase vegetativa. Esta menor expresión de *TRY* en las líneas de introgresión con alelos Doñ-0 en estos dos loci indica que los polimorfismos de *TRY* entre *Ler* y Doñ-0 afectan a su regulación en *cis*; que *TCL1* es un promotor de la expresión de *TRY* que regula en *trans* su expresión; y que los polimorfismos *Ler*/Doñ-0 en *TCL1/FTD2* contribuyen a la regulación diferencial en *trans* de *TRY*. Por último, se observó una correlación negativa significativa ( $r = -0,53$ ;  $P = 0,01$ ) entre el valor de NTF y la expresión de *TRY* en fase reproductiva (Figura 37B), de forma que las líneas con menor expresión de *TRY* tienen mayor cantidad de tricomas en el fruto. Sin embargo, no se observó una correlación significativa entre la expresión de *TRY* y la densidad de

tricomas en fase vegetativa (Figura 37E). Estos resultados permiten concluir que los polimorfismos entre Ler y Doñ-0 en *TRY/FTD5* y *TCL1/FTD2* afectan a la regulación en *cis* y *trans* de *TRY*, respectivamente.



**Figura 37. Relación entre el nivel de expresión de *TCL1*, *TRY* o *GL1* y el patrón de tricomas de las líneas de introgresión.** Relación entre la expresión de *TCL1* (A, D), *TRY* (B, E) y *GL1* (C, F) en la inflorescencia y el valor de NTF (A, B, C) o la expresión en la hoja y el valor de DTH (D, E, F) en las líneas de introgresión. Dentro de cada panel se indica el índice de correlación de Pearson ( $r$ ), la significación estadística de la correlación ( $P$ ), la proporción de varianza explicada por la relación entre las dos variables comparadas ( $R^2$ ) y el número de muestras analizadas ( $N$ ). Los valores de expresión de las líneas de introgresión están relativizados frente al valor medio de la expresión de cada gen en Ler en las hojas.

Finalmente, el análisis de *GL1* en las líneas de introgresión mostró la mitad de expresión en las inflorescencias que en las hojas (Figura 36C), en concordancia con los resultados de las líneas transgénicas para el gen *GL1*. Tampoco se detectaron diferencias significativas en el nivel de expresión de *GL1* entre las líneas de introgresión, ni en fase vegetativa ni reproductiva (Figura 36C) y el análisis de varianza con los tres loci *FTD* como factores mostró que ninguno de ellos afecta significativamente a la expresión de *GL1* (Tabla 18). Además, no se detectó correlación entre los niveles de expresión de *GL1* y el patrón de tricomas en el fruto, aunque sí en la hoja (Figura 37C y 37F). No obstante, la accesión parental Doñ-0 presentó una expresión de *GL1* en la

inflorescencia y un valor de NTF superior al de las líneas de introgresión. Por lo tanto, Ler y Doñ-0 no difieren en la regulación en *cis* de *GL1* pero sí en su regulación en *trans* a través de algún locus distinto de los tres loci incluidos en las líneas de introgresión estudiadas.

# DISCUSIÓN



## 1. Variación natural para el patrón de tricomas a lo largo del desarrollo

### 1.1. Arabidopsis muestra una gran variación para el patrón de tricomas en todos sus órganos

En *Arabidopsis*, la variación natural para el patrón de tricomas se ha estudiado principalmente en las hojas de la fase vegetativa (Larkin *et al.*, 1996; Mauricio, 2005; Symonds *et al.*, 2005; Pfalz *et al.*, 2007; Atwell *et al.*, 2010; Bloomer *et al.*, 2012; Bloomer *et al.*, 2014). Igualmente, estudios realizados en otras especies han abordado de forma mayoritaria la variación de este carácter en esta fase del desarrollo, enfocándose principalmente en el estudio de las variantes naturales cualitativas, es decir, en las variedades glabras frente a las que presentan tricomas, y en la contribución del gen *GL1* a esta variación (Hauser *et al.*, 2001; Kärkkäinen & Ågren, 2002; Kivimäki *et al.*, 2007; Kawagoe *et al.*, 2011; Li *et al.*, 2013). A diferencia de los trabajos previos, esta tesis es uno de los pocos estudios en los que se ha analizado la variación natural para el patrón de tricomas a lo largo de todo el desarrollo de la planta, incluyendo las hojas de la fase vegetativa, así como los tallos, pedicelos y frutos de la fase reproductiva (Kawagoe *et al.*, 2011; Savic, 2015). El análisis de una amplia colección de accesiones naturales de *Arabidopsis* de la península ibérica demuestra que hay una gran variación genética intraespecífica en ambas fases del desarrollo y en todos estos órganos. Además, la identificación de *EGL3* como gen causal de *FTD1* sugiere que también hay diversidad natural para el patrón de tricomas en los sépalos, ya que mutaciones de pérdida de función de *EGL3* afectan al patrón de tricomas en este órgano.

El análisis genético de las 174 accesiones naturales ibéricas muestra que el patrón de tricomas en fase vegetativa y reproductiva tiene heredabilidades moderadas (0,61) y altas (0,85-0,89) respectivamente, en concordancia con los valores obtenidos en otros estudios realizados sólo en fase vegetativa (Symonds *et al.*, 2005; Mauricio, 2005; Atwell *et al.*, 2015; Bloomer *et al.*, 2014; Savic, 2015). Este resultado indica que la variación natural para el patrón de tricomas tiene un componente genético relativamente alto, aunque menor que el estimado para otros caracteres más estudiados como el tiempo de floración, o la resistencia a ciertos patógenos, donde el efecto ambiental es sustancialmente menor, lo que facilita su estudio (Alonso-Blanco & Méndez-Vigo, 2014). Sin embargo, las bases genéticas de la variación natural para el patrón de tricomas en las hojas y en los distintos órganos de la fase reproductiva son parcialmente diferentes ya que los caracteres de ambas fases presentan correlaciones bajas. Por el contrario, la mayor correlación entre las variables del patrón de tricomas

en el tallo y el pedicelo indica que estos órganos comparten una mayor proporción de las bases genéticas de su variación. Este resultado se ve apoyado por los análisis de las líneas de introgresión y las líneas transgénicas portadoras de alelos Doñ-0 para distintos loci *FTD*, puesto que sólo el locus *FTD2/TCL1* afecta a ambas fases, mientras que los loci *FTD5/TRY* y *FTD3/GL1* afectan al patrón de tricomas únicamente en la fase reproductiva.

La variación natural para el patrón de tricomas se ha asociado con diferentes factores ambientales como la herbivoría (Mauricio, 1998; Valverde *et al.*, 2001; Züst *et al.*, 2012), cuyo efecto se ha abordado en esta tesis de forma indirecta mediante el tratamiento exógeno con ácido jasmónico. En general, las accesiones de *Arabidopsis* mostraron una escasa respuesta del patrón de tricomas en las hojas al AJ, ya que las plantas tratadas presentaron una variación similar a la observada en las accesiones sin tratamiento. Esta respuesta fue inferior a la descrita previamente para las accesiones Col y Ws-0 (Traw & Bergelson, 2003). Es posible que esta menor respuesta se deba a la menor concentración de ácido jasmónico empleada en nuestro estudio en comparación con otros trabajos previos (Traw & Bergelson, 2003; Atwell *et al.*, 2010). Esta concentración se eligió a partir de un análisis preliminar donde se observó que concentraciones superiores a 0,15 mM producen un gran efecto sobre la reducción del tamaño y un daño físico en las hojas mucho mayor. Por ello, es probable que las mayores respuestas observadas en estudios previos no se deban directamente al tratamiento con AJ sino que sean consecuencia indirecta del menor tamaño de las hojas y de los daños provocados por la elevada concentración de AJ. Además, en concordancia con un estudio previo en el que se aplicó un daño mecánico en las hojas como tratamiento imitador de la herbivoría (Bloomer *et al.*, 2014), la heredabilidad de la densidad de tricomas de las accesiones ibéricas tratadas fue ligeramente superior al de las plantas no tratadas con AJ (0,61 y 0,73). Este resultado sugiere que la inducción del desarrollo de tricomas provocada por el AJ conlleva una reducción del efecto de otros factores ambientales sobre el patrón de tricomas (Bloomer *et al.*, 2014). No obstante, este aumento de la heredabilidad del patrón de tricomas en plantas tratadas con AJ no se ha observado en todos los estudios (Atwell *et al.*, 2010), por lo que probablemente otros factores ambientales afecten al patrón de tricomas de forma independiente de la herbivoría.

Las accesiones de *Arabidopsis* analizadas no sólo mostraron respuestas pequeñas sino también una limitada variación natural para dicha respuesta o plasticidad del patrón de tricomas al tratamiento con AJ, ya que sólo el 4% de la varianza de DTH fue explicada por la interacción entre el genotipo de las accesiones y el tratamiento con AJ. No obstante, este resultado indica la existencia de un pequeño

componente genético para la variación natural para la respuesta a la herbivoría, el cual también se ha podido diseccionar genéticamente para identificar posibles genes causales.

## **1.2. Bases genéticas del patrón de tricomas en distintos órganos y fases del desarrollo**

Los análisis de asociación genómica realizados a partir de la colección regional de accesiones ibéricas aportan varios aspectos novedosos sobre las bases genéticas del patrón de tricomas en *Arabidopsis*. En primer lugar, el número moderado de genes asociados con el patrón de tricomas en los diferentes órganos de la planta (entre 243 y 419) sugiere que la variación natural para el patrón de tricomas está determinada por un número relativamente bajo de genes, de forma similar a lo descrito para otros caracteres cuantitativos de *Arabidopsis* como el tiempo de floración (Tabas-Madrid *et al.*, 2018). En segundo lugar, en concordancia con las correlaciones observadas entre las distintas variables del patrón de tricomas en los distintos órganos analizados, la comparación de genes asociados en los análisis de GWAS muestra un menor solapamiento entre las bases genéticas de la fase vegetativa y reproductiva (3-5% de los genes asociados en común), que entre las bases de los distintos órganos de la fase reproductiva (25% de los genes asociados). Por último, como se describe a continuación, este análisis ha permitido identificar varios genes candidatos nuevos que podrían contribuir a la variación natural para el patrón de tricomas en los distintos órganos de la planta.

Los genes candidatos identificados en este estudio muestran la utilidad de esta colección regional para los estudios de asociación genómica, ya que incluyen: 1) genes conocidos como implicados en la variación natural para el patrón de tricomas; 2) genes conocidos implicados en la regulación del patrón de tricomas que podrían contribuir a su variación natural; 3) genes nuevos desconocidos que también podrían contribuir a dicha variación. Entre los primeros, cabe destacar la identificación de los genes *ETC2* y *MYC1* cuya contribución a la variación natural del patrón de tricomas ha sido previamente demostrada (Hilscher *et al.*, 2009; Symonds *et al.*, 2011). La identificación del polimorfismo de *ETC2* (Lys/Glu<sup>19</sup>) segregando en la península ibérica indica la contribución de este gen a la variación en esta región. Sin embargo, el polimorfismo Pro/Ala<sup>189</sup> candidato para el gen *MYC1* descrito previamente (Symonds *et al.*, 2011) no segrega en esta región, lo que sugiere la existencia de otros polimorfismos funcionales específicos de la península ibérica en dicho gen. Además, los genes *TCL1*, *TRY* y *GL1*, cuya variación natural se ha demostrado en esta tesis,

también aparecen como significativos en los análisis de asociación genómica. En concordancia con los resultados obtenidos a partir del análisis de plantas transgénicas, *TCL1* aparece asociado con el patrón de tricomas en todas las fases y órganos, mientras que *TRY* y *GL1* sólo aparecen asociados con las variables de la fase reproductiva. Además, varios de los genes conocidos como reguladores del patrón de tricomas detectados en este estudio regional, como *TCL2* y los ya mencionados *ETC2*, *TCL1*, *GL1* y *MYC1*, también se han identificado en otro análisis de GWAS realizado con una colección de accesiones mundial (Atwell *et al.*, 2010), lo que avala la utilidad de las colecciones regionales.

En el segundo grupo cabe destacar la identificación de varios nuevos genes candidatos cuya función sobre el desarrollo de tricomas ha sido previamente descrita, como son *ZFP8* y *KIS*, los cuales están asociados con ambas fases del desarrollo; *SOC1* asociado con el patrón de tricomas de los tallos y del pedicelo; los genes *PFL2* y *LOM3*, los genes *miR156*, *SAD2* y *LOM2*, y los genes *MYB23* y *CPR5*, asociados con el patrón de tricomas en la hoja, en el tallo o en el pedicelo, respectivamente; o el gen *SKU5* asociado con el patrón de tricomas en plantas tratadas con AJ y con la respuesta a dicho tratamiento. Por último, en esta tesis se han identificado otros nuevos genes candidatos desconocidos hasta la fecha como son: el gen *At2g04925* de función desconocida pero asociado fuertemente con el patrón de tricomas en las hojas; el gen *MUTE*, implicado en el desarrollo de estomas, y que aparece asociado con el patrón de tricomas en el tallo y en el pedicelo; y el gen *At4g04985* que codifica para una proteasa y que presenta un polimorfismo de alto impacto funcional que podría estar implicado en la respuesta del patrón de tricomas al AJ. La identificación de estos nuevos genes muestra la utilidad de la colección regional de la península ibérica, ya que probablemente los alelos de estos genes presenten una mayor frecuencia a escala regional que global, lo que facilita su detección estadística (Tabas-Madrid *et al.*, 2018).

A pesar de la utilidad de esta colección regional para los análisis de asociación genómica, hay varios problemas que todavía limitan la potencia de esta metodología para la detección de los genes causales de la variación natural. En primer lugar, aunque la heterogeneidad alélica dentro de los genes que contribuyen a la variación natural es menor a escala regional que a escala global, esta sigue limitando la detección de algunos genes responsables. Esto está ilustrado por el gen *EGL3* identificado en esta tesis como el gen correspondiente al locus *FTD1* implicado en el patrón de tricomas del fruto. Dada la correlación entre el patrón de tricomas de los distintos órganos de la fase reproductiva, es de esperar que *EGL3/FTD1* también afecte al patrón de tricomas en el pedicelo y/o el tallo. Sin embargo, es probable que

este gen no se haya detectado debido a la existencia de varios polimorfismos funcionales segregando en la península ibérica, como sugiere la asociación parcial de varios polimorfismos de *EGL3* con el patrón de tricomas en el fruto. Esta limitación también se ha observado previamente en esta colección regional para el tiempo de floración, donde *FRIGIDA*, uno de los genes principales responsables de su variación natural, no se detectó debido a su elevada heterogeneidad alélica (Tabas-Madrid *et al.*, 2018). En segundo lugar, la correlación entre la estructura genética y la variación fenotípica también limita la potencia de los análisis de asociación genómica. Esto es especialmente importante para el patrón de tricomas en el tallo y el pedicelo, ya que estos caracteres están muy correlacionados con la estructura genética general de la península ibérica (0,52-0,69). Esta correlación, junto con la falta de distribuciones normales de dichas variables, es probablemente responsable del mayor número de falsas asociaciones detectadas para estos caracteres. Por último, el fuerte desequilibrio de ligamiento entre genes ligados genéticamente limita la detección de los genes causales precisos. Esto se ilustra con el clúster de genes estrechamente ligados en el cromosoma 2, *KIS*, *ETC2*, *TCL1* y *TCL2*, donde todos ellos aparecen asociados aunque sólo parte de ellos probablemente contribuyan a la variación natural. No obstante, la identificación del gen *TCL1/FTD2* llevada a cabo en esta tesis, y del gen *ETC2* identificado previamente (Hilscher *et al.*, 2009), demuestra que varios genes estrechamente ligados pueden afectar simultáneamente al mismo carácter, lo que complica la disección genética de caracteres complejos como el patrón de tricomas.

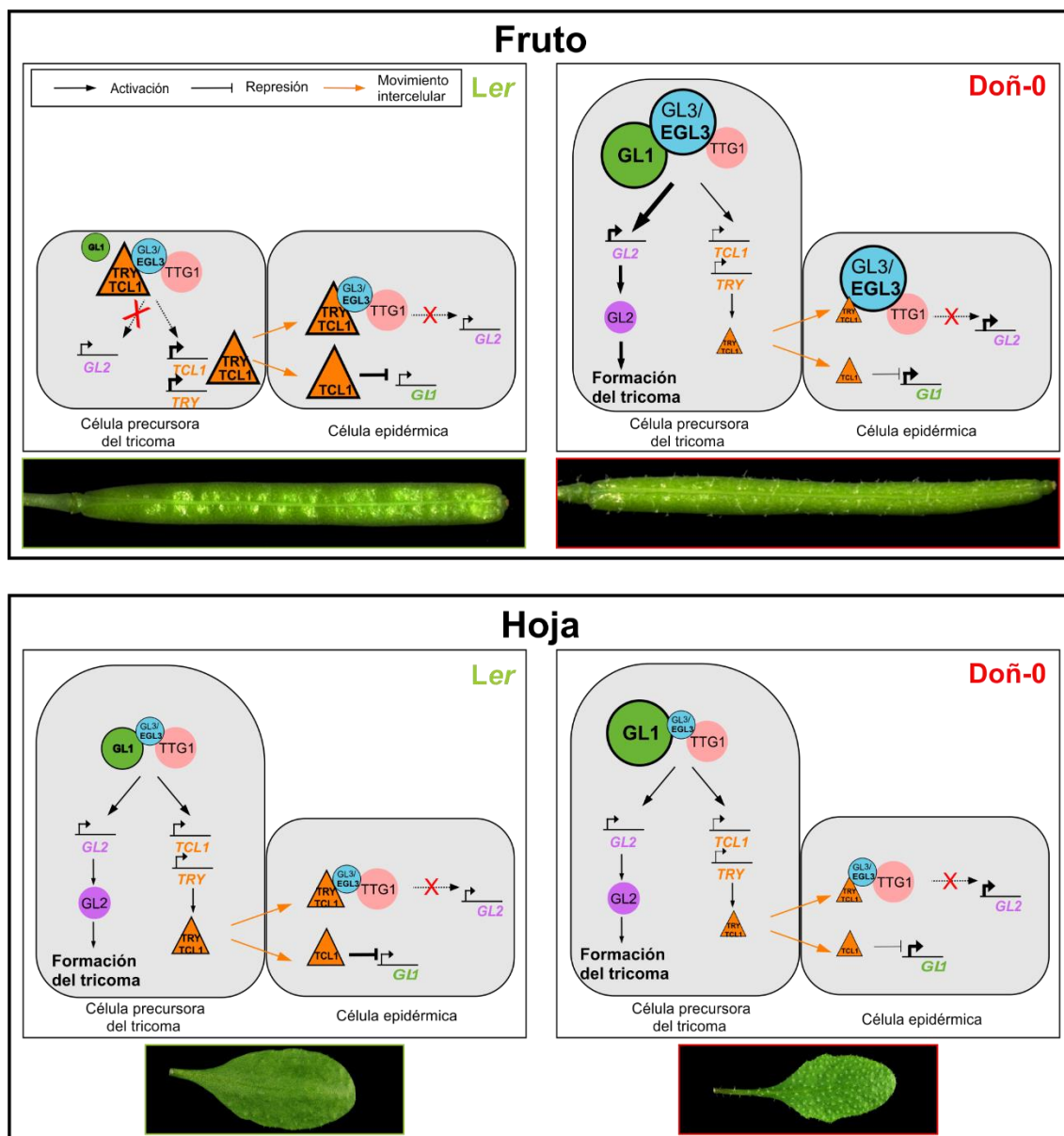
## 2. Mecanismos moleculares del desarrollo de tricomas en los frutos

Los análisis llevados a cabo en esta tesis han permitido identificar *EGL3*, *TCL1*, *GL1* y *TRY* como los genes subyacentes a *FTD1*, *FTD2*, *FTD3* y *FTD5* respectivamente. De esta forma, se han identificado cuatro nuevos genes implicados en la variación natural para el patrón de tricomas, ya que ninguno de ellos se había descrito en estudios previos (Hauser *et al.*, 2001; Hilscher *et al.*, 2009; Symonds *et al.*, 2011). Como se ha mostrado con el análisis de líneas de introgresión, los genes *TCL1/FTD2*, *GL1/FTD3* y *TRY/FTD5* son suficientes para el desarrollo de tricomas en el fruto, mientras que *EGL3/FTD1* contribuye al aumento de la densidad de tricomas en este órgano. El aislamiento de estos cuatro loci proporciona los mecanismos moleculares que determinan la variación natural para el patrón de tricomas en el fruto. Interesantemente, los cuatro genes codifican para factores de transcripción de distintas familias (bHLH, MYB R2R3 y MYB R3) que juegan un papel central en la

regulación del patrón de tricomas en *Arabidopsis* (Figura 3; Hauser, 2014; Pattanaik *et al.*, 2014). En concreto, *EGL3* y *GL1* son genes promotores del desarrollo de tricomas (Oppenheimer *et al.*, 1991; Zhang *et al.*, 2003) y ambos forman parte del complejo activador (GL1-GL3/EGL3-TTG1) (Payne *et al.*, 2000; Zhang *et al.*, 2003). Por el contrario, *TCL1* y *TRY* son dos de los genes represores del patrón de tricomas con función parcialmente redundante (Wang *et al.*, 2007; Schelmann *et al.*, 2002). La identificación de estos cuatro genes demuestra que el desarrollo de tricomas en el fruto en *Arabidopsis* ha requerido mutaciones en múltiples genes reguladores centrales, tanto promotores como represores, del patrón de tricomas. A continuación se describe la variación alélica identificada en los cuatro genes y su relevancia funcional para la variación natural en el patrón de tricomas.

### **2.1. Alelos naturales de ganancia de función en los genes promotores del patrón de tricomas *EGL3* y *GL1***

Varios de los resultados obtenidos en este estudio demuestran que *EGL3* es el gen *FTD1* y que la accesión Doñ-0 es portadora de un alelo de ganancia de función debido a una mutación reguladora en *cis* de la expresión de *EGL3* en fase reproductiva. En primer lugar, la alta definición del mapeo genético realizado para *FTD1* indica que el polimorfismo causal de este locus se localiza en la región promotora de *EGL3*. Segundo, el análisis de expresión de *EGL3* en las accesiones parentales muestra una mayor expresión en Doñ-0 que en *Ler* en fase reproductiva, así como una mayor expresión en inflorescencias y frutos que en hojas, en concordancia con las descripciones previas en la accesión Col (Bernhardt *et al.*, 2003). Además, el mutante artificial de pérdida de función *egl3* presenta una reducción de la cantidad de tricomas en los sépalos en comparación con Col, lo que indica que *EGL3* está implicado principalmente en la regulación del patrón de tricomas en la fase reproductiva. Así, esta mayor expresión de *EGL3* es uno de los mecanismos moleculares que aumenta la cantidad de tricomas en los frutos de Doñ-0 (Figura 38). Además, la menor expresión de *EGL3* en las hojas, tanto de *Ler* como Doñ-0, sugiere que este polimorfismo no contribuye a la variación natural para el patrón de tricomas en la fase vegetativa. No obstante, actualmente se desconoce el polimorfismo preciso que determina las diferencias funcionales de *EGL3* entre Doñ-0 y *Ler*, aunque las asociaciones detectadas (Tabla 13) sugieren la existencia de varios polimorfismos localizados en la kilobase del promotor adyacente al gen.



**Figura 38. Mecanismos moleculares de la variación natural para el patrón de tricomas en Ler y Doñ-0.** Cada panel representa el modelo de regulación del patrón de tricomas de Arabidopsis (ver Introducción, Figura 3) indicando los mecanismos moleculares determinados por los alelos de Ler (paneles de la izquierda) y Doñ-0 (paneles de la derecha), en los frutos (paneles superiores) y en las hojas (paneles inferiores). Las consecuencias moleculares de las variantes alélicas naturales de *EGL3*, *GL1*, *TCL1* y *TRY* se representan mediante cambios en el tamaño de las proteínas y de las flechas que representan la transcripción de los genes implicados en el patrón de tricomas, de manera que mayor tamaño indica mayor actividad de la proteína o mayor expresión del gen.

Igualmente, los análisis realizados demuestran que *GL1*, otro de los genes que codifican para un factor de transcripción del complejo *GL1-GL3/EGL3-TTG1*, es *FTD3*. Sin embargo, varios resultados muestran que su contribución a la regulación y la variación natural para el patrón de tricomas es distinta a la de *EGL3*. Por un lado, al contrario que *EGL3*, *GL1* muestra una mayor expresión en las hojas que en las inflorescencias, tanto en Ler como Doñ-0, lo que sugiere una mayor implicación en la

regulación del patrón de tricomas en la fase vegetativa. Además, el polimorfismo funcional de *GL1* entre *Ler* y Doñ-0 no afecta a la regulación de su expresión, como muestran los análisis de líneas portadoras del transgén de *Ler* o Doñ-0. No obstante, las diferencias de expresión de *GL1* en fase reproductiva entre la accesión parental Doñ-0 y las líneas de introgresión con alelos Doñ-0 en *FTD2*, *FTD3* y *FTD5*, indican la existencia de variación natural para algún regulador en *trans* de *GL1* distinto de *TCL1* y *TRY*. Estudios futuros de *GL1* en materiales genéticos que difieran en los alelos de *EGL3/FTD1*, permitirán determinar si *EGL3* u otro gen como el locus *FTD4* (Figura 5) podrían regular la expresión de *GL1*.

Estos resultados indican que el polimorfismo funcional de *GL1* entre *Ler* y Doñ-0 probablemente se deba a un cambio estructural de la proteína, el cual produce una ganancia de función en el alelo de Doñ-0. Estudios previos en nuestro laboratorio han mostrado que *Ler* y Doñ-0 difieren en el codón de parada prematuro Tyr/Stop<sup>226</sup> presente en Doñ-0 y que implica la pérdida de los dos últimos aminoácidos de la proteína (Cys<sup>227</sup> y Leu<sup>228</sup>). Actualmente creemos que este codón de parada es el polimorfismo funcional ya que los dos últimos aminoácidos de *GL1* están altamente conservados en la familia *Brassicaceae* (Savic, 2015) y el aminoácido Cys<sup>227</sup> podría estar implicado en la regulación funcional de *GL1* (Gould *et al.*, 2013). Puesto que *GL1* regula el patrón de inicio de los tricomas a través de varias funciones, la ganancia de función del alelo de Doñ-0 se podría generar a través de distintos mecanismos moleculares. Por un lado, *GL1* activa la expresión de los genes represores MYB R3 (Wang *et al.*, 2008), de forma que la alteración de la proteína *GL1* en Doñ-0 podría conducir a una menor activación. Sin embargo, esta hipótesis no se encuentra apoyada por los análisis de expresión de las líneas de introgresión, ya que la variación natural de *GL1* no afecta a la expresión de *TCL1* o *TRY*. Por otro lado, *GL1* activa la expresión de *GL2* a través de la unión de su región C-terminal con *GL3*, y de la unión de los dominios MYB y bHLH del complejo *GL1/GL3* a secuencias dianas del promotor de *GL2* (Wang & Chen, 2008). Por ello, es probable que el codón de parada prematuro de Doñ-0, al alterar la región C-terminal de *GL1*, permita una unión más fuerte o estable con *GL3* y, en consecuencia, una mayor activación de *GL2* (Figura 38). Este aumento de la función promotora de *GL1* presente en Doñ-0 proporciona el segundo mecanismo molecular necesario para el desarrollo de tricomas en el fruto.

La contribución de *GL1* a la variación natural cualitativa y cuantitativa se ha estudiado previamente en *Arabidopsis* y especies afines, donde se han identificado numerosos alelos de pérdida de función que provocan un fenotipo glabro en toda la planta (Hauser *et al.*, 2001; Kivimäki *et al.*, 2007; Kawagoe *et al.*, 2011; Bloomer *et al.*, 2012; Li *et al.*, 2013; Savic, 2015). Sin embargo, el alelo de *GL1* de Doñ-0 es el primer



alelo de ganancia de función descrito hasta la fecha, por lo que es un alelo nuevo que amplía sustancialmente la diversidad funcional de este gen. No obstante, el análisis de las secuencias de los genomas de *Arabidopsis* disponibles (The 1001 Genomes Consortium, 2016) muestra que el codón prematuro de Doñ-0 sólo está presente en la accesión Bon-61 procedente de una población geográficamente muy próxima al parque natural de Doñana (Figura 39C). Por lo tanto, *GL1* posiblemente presente otros alelos de ganancia de función distintos del de Doñ-0 ya que se conocen ~15 accesiones diferentes con tricomas en los frutos (Savic, 2015). En este sentido, análisis de las secuencias de *GL1* han identificado un haplogrupo diferenciado por otros tres aminoácidos que se han asociado con la mayor densidad de tricomas en las hojas y con el desarrollo de tricomas en el fruto (Bloomer *et al.*, 2012; Savic, 2015).

## **2.2. Alelos naturales de pérdida de función parcial en los genes represores del patrón de tricomas *TCL1* y *TRY***

Los análisis realizados en esta tesis han permitido demostrar que los genes *TCL1* y *TRY*, que codifican para dos factores de transcripción MYB represores del patrón de tricomas, corresponden a los loci *FTD2* y *FTD5*, respectivamente. La variación natural que altera la función de ambos genes y determina la formación de tricomas en el fruto de Doñ-0 comparte varias características. En primer lugar, los polimorfismos funcionales de ambos genes son mutaciones reguladoras en *cis*. En el caso de *TCL1*, el polimorfismo funcional se ha localizado en una pequeña región 3' del gen, en la que *Ler* y Doñ-0 no difieren en ningún polimorfismo que afecte a la estructura de la proteína (Savic, 2015), pero cuyos polimorfismos afectan fuertemente a los niveles de expresión de *TCL1*. En concreto, se ha identificado una delección de 168 pb de la región 3' UTR de Doñ-0, la cual se ha asociado previamente con el número de tricomas en el fruto de accesiones de la península ibérica (Savic, 2015), lo que sugiere que éste es el polimorfismo funcional de *TCL1*. En el caso del gen *TRY*, la ausencia de polimorfismos entre *Ler* y Doñ-0 que afecten a la estructura de la proteína (Savic, 2015), junto con las diferencias de expresión observadas entre las líneas parentales y transgénicas, indican que el polimorfismo funcional también ha de ser regulador. Sin embargo, a diferencia de *TCL1*, el gen *TRY* probablemente requiera varios polimorfismos funcionales para el desarrollo de tricomas en el fruto, ya que las líneas con transgenes quiméricos entre *Ler* y Doñ-0 presentaron patrones de tricomas diferentes de los de las líneas con los alelos parentales de Doñ-0. Actualmente se desconocen los polimorfismos de *TRY* que podrían alterar su función, pero el SNP-701 del segundo intrón probablemente sea uno de ellos, ya que también se ha asociado

previamente con el patrón de tricomas en el fruto de accesiones de la península ibérica (Savic, 2015).

En segundo lugar, los alelos de Doñ-0 de ambos genes, *TCL1* y *TRY*, presentan una pérdida de función sólo parcial, ya que los transgenes de estos alelos son funcionales puesto que son capaces de complementar los fenotipos de pérdida de función total causados por las mutaciones *tcl1-Salk* y *try-Salk*, respectivamente. Sin embargo, ambos alelos de Doñ-0 presentan una menor expresión que los alelos de *Ler* y producen un mayor número de tricomas en los frutos de plantas transgénicas, lo que demuestra su menor funcionalidad.

Además, los alelos *Ler* de *TCL1* y *TRY* se caracterizan por presentar una mayor expresión en las inflorescencias que en las hojas (Figura 38), en concordancia con trabajos previos realizados en la accesión *Col* (Wang *et al.*, 2007; Yu *et al.*, 2010; Gan *et al.*, 2011; Tian *et al.*, 2017). Por el contrario, los alelos de Doñ-0 de ambos genes mantienen una expresión constante y baja a lo largo del desarrollo, produciendo así una reducción general de su actividad represora del patrón de tricomas (Figura 38). Por lo tanto, ambos alelos de Doñ-0 han perdido la regulación de su expresión en *cis* a lo largo de la ontogenia de la planta, proporcionando así el tercer mecanismo molecular necesario para el desarrollo de tricomas en los frutos de Doñ-0.

La caracterización de los alelos naturales de *TCL1* y *TRY* también ha mostrado diversas diferencias entre ambos genes. Por un lado, *TCL1* y *TRY* difieren en su función como reguladores negativos del patrón de tricomas a lo largo del desarrollo de la planta, ya que *TCL1* afecta a todos los órganos de la fases vegetativa y reproductiva, mientras que *TRY* solo afecta a la fase reproductiva. En concordancia, estudios previos han mostrado que *TCL1* es el principal regulador negativo del patrón de tricomas en los tallos y pedicelos (Wang *et al.*, 2007), pero también está implicado en el desarrollo de tricomas en las hojas cuando se combina con mutaciones en otros genes represores MYB R3 (Wang *et al.*, 2008). Igualmente, los análisis de mutantes simples y múltiples de *TRY* han mostrado su función en la regulación del patrón de tricomas en las hojas y en el tallo (Schelmann *et al.*, 2002). Sin embargo, como muestran los análisis de plantas transgénicas y de líneas de introgresión con distintos alelos de *Ler* y Doñ-0, la variación natural de *TRY* sólo afecta de forma significativa al fruto. Este comportamiento diferencial de la variación natural de ambos genes se ve apoyado por los análisis de asociación genómica con las accesiones de la península ibérica ya que *TCL1* aparece asociado con el patrón de tricomas en todos los órganos, mientras que *TRY* sólo se detectó en el tallo. Por otra parte, los alelos naturales de *TCL1* y *TRY* también difieren en la regulación de su expresión en *trans*. Como muestran los análisis de expresión de las líneas de introgresión, la expresión de *TRY*

depende del propio alelo natural (regulación en *cis*), pero también del alelo de *TCL1* (regulación en *trans*), de forma que el alelo de *TCL1-Doñ-0* activa menos la expresión de *TRY* que el alelo *TCL1-Ler*. Por el contrario, la expresión de *TCL1* sólo se ve afectada por su regulación en *cis*. Estos resultados demuestran que *TCL1* es un regulador positivo de *TRY*, de manera que parte de su efecto represor se lleva a cabo a través de la activación de la expresión de *TRY*, lo que revela una relación funcional entre ambos genes no descrita previamente en *Arabidopsis*.

### **2.3. Efecto de la variación natural sobre la regulación del patrón de tricomas en el fruto**

En conjunto, la caracterización de los alelos naturales de *TCL1*, *TRY*, *EGL3* y *GL1* permite establecer un modelo de regulación diferencial del desarrollo de tricomas en *Ler* y *Doñ-0* (Figura 38). Básicamente, el desarrollo de tricomas en los frutos de *Doñ-0* se debe a los tres mecanismos moleculares descritos: 1) la menor expresión de *TCL1* y *TRY* producida por los alelos de pérdida de función parcial determina una menor cantidad de ambos represores. Por lo tanto, estas proteínas competirán menos con *GL1* para bloquear la formación del complejo activador, y producirán una menor represión de la expresión de *GL1* (Figura 38). 2) Una mayor expresión de *EGL3* causada por el alelo de ganancia de función aumentará la cantidad del complejo activador (*GL1-GL3/EGL3-TTG1*) y consecuentemente una mayor activación de *GL2*. 3) Una proteína *GL1* estructuralmente más activa producida por el alelo de ganancia de función de *Doñ-0* conducirá a una mayor estabilidad y/o actividad del complejo activador (Figura 38). Por el contrario, en *Ler* y probablemente en la mayoría de las accesiones de *Arabidopsis*, la mayor expresión de *TCL1* y *TRY*, junto con la menor expresión de *EGL3* y la menor estabilidad de *GL1* determinan la ausencia de tricomas en los frutos. Estos alelos naturales afectarán a la densidad de tricomas en las hojas en menor medida ya que ambas accesiones difieren menos en su expresión de los cuatro genes en fase vegetativa. No obstante, la mayor densidad de tricomas en las hojas de *Doñ-0* vendrá determinada, al menos en parte, por la menor expresión de *TCL1* (Figura 38).

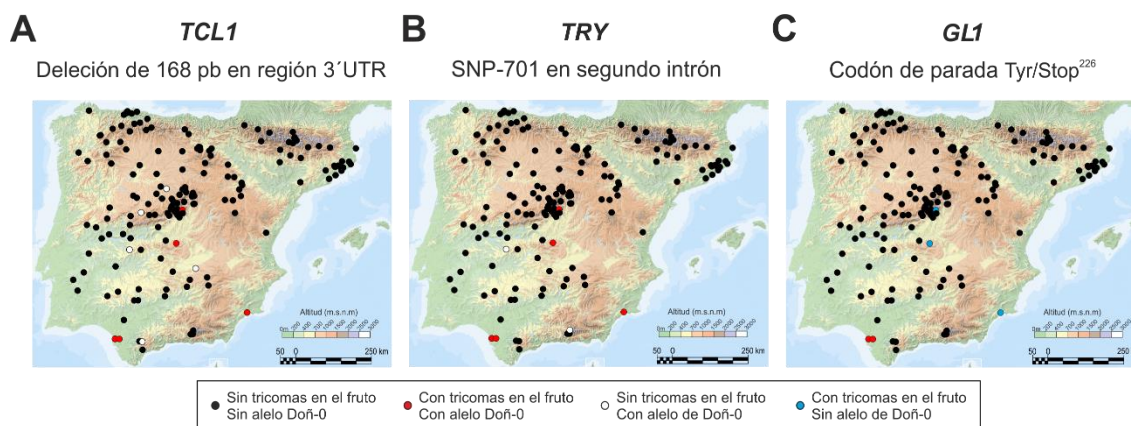
## **3. Evolución del patrón de tricomas en los frutos de *Arabidopsis***

Varios trabajos previos han conseguido desarrollar tricomas en los frutos de *Arabidopsis* mediante diferentes modificaciones genéticas artificiales, incluyendo la sobreexpresión de *GL1* en el fondo mutante *try* (Szymanski *et al.*, 1998), el cuádruple mutante *try cpc etc1 tcl1* (Wang *et al.*, 2008), líneas sobreexpresoras de los genes

*ZFP5*, *ZFP6*, *ZFP8*, *GIS*, *GIS2* o *GIS3* (Gan *et al.*, 2007; Sun *et al.*, 2015), o el silenciamiento de *AGAMOUS* (*AG*), tanto en un fondo genético *Ler* como en fondos genéticos con mutaciones en genes implicados en el patrón de tricomas (Ó'Maoiléidigh *et al.*, 2018). Sin embargo, este carácter no se ha descrito previamente en variedades naturales de *Arabidopsis*, por lo que se trata de una innovación evolutiva para esta especie. Interesantemente, la presencia de tricomas en los frutos sólo se ha encontrado en ~15 poblaciones naturales de la península ibérica, lo que sugiere que este carácter ha evolucionado específicamente en esta región geográfica. Además, todas las accesiones con tricomas en los frutos pertenecen a un mismo grupo genético, denominado grupo relicto por su alta diferenciación del resto de las poblaciones de la especie (The 1001 Genomes Consortium, 2016; Tabas-Madrid *et al.*, 2018). Análisis genómicos han mostrado que los relictos son el grupo genético más antiguo de *Arabidopsis*, el cual se originó en África desde donde colonizó la península ibérica antes de las últimas glaciaciones (Brennan *et al.*, 2014; Durvasula *et al.*, 2017; Fulgione & Hancock, 2018). No obstante, sólo el 23% de las accesiones del grupo relicto secuenciadas (5 de 22) producen tricomas en los frutos.

La identificación de los genes y polimorfismos funcionales responsables del desarrollo de tricomas en los frutos permite determinar si este carácter ha evolucionado por mutaciones originadas específicamente en esta región, o a partir de variación genética previamente existente en otras regiones del mundo. El análisis de las secuencias de ~1.200 genomas procedentes de todo el mundo (The 1001 Genomes Consortium, 2016) muestra que la delección de la región 3' UTR de *TCL1*, el SNP-701 de *TRY* y el codón de parada prematuro de *GL1* identificados en Doñ-0, no existen fuera de la península ibérica. La presencia del alelo de Doñ-0 en estos polimorfismos, exclusivamente en la península ibérica, indica que el desarrollo de tricomas en el fruto ha evolucionado en esta región a partir del grupo genético de los relictos. En particular, el codón de parada de Doñ-0 en *GL1* se ha encontrado sólo en 2 de las 5 accesiones con tricomas en el fruto secuenciadas, Doñ-0 y Bon-61, las cuales proceden de poblaciones genética y geográficamente muy próximas (Figura 39C). Por el contrario, todas las accesiones con tricomas en el fruto presentan la delección de *TCL1* y el alelo de Doñ-0 en el SNP-701 de *TRY*. Además, los alelos de Doñ-0 en *TCL1* y *TRY* también aparecen segregando a baja frecuencia en el grupo relicto independientemente de la presencia de tricomas, ya que también se encuentran en 2 o 5 accesiones de este grupo sin tricomas en los frutos (Figura 39). Estos resultados sugieren que la presencia de tricomas en los frutos evolucionó por combinación de la variación genética ya existente en *TCL1* y *TRY* del grupo relicto, y nuevas mutaciones surgidas en *GL1*, como el codón de parada prematuro de Doñ-0.

No obstante, otras mutaciones en *GL1* u otros genes son necesarias para explicar el desarrollo de tricomas en los frutos de las accesiones sin este codón de parada (Figura 39C), ya que como muestra el análisis de las líneas de introgresión (Figura 25), los alelos de Doñ-0 en *TCL1* y *TRY* no son suficientes para producir este carácter.



**Figura 39. Distribución geográfica de los polimorfismos funcionales de Ler/Doñ-0.** Cada panel muestra la distribución de los polimorfismos candidatos identificados en *TCL1* (A), *TRY* (B) y *GL1* (C), en las 174 accesiones ibéricas de *Arabidopsis* secuenciadas (The 1001 Genomes Consortium, 2016). Las accesiones sin el alelo de Doñ-0 y sin tricomas en los frutos se muestran en color negro. Las accesiones con el alelo de Doñ-0 se muestran en color rojo o blanco en función de la presencia o no, respectivamente, de tricomas en el fruto. Las accesiones sin alelo Doñ-0 y con tricomas en el fruto se representan en color azul.

Estudios previos han mostrado que el patrón de tricomas en las hojas de *Arabidopsis* está implicado en la adaptación a diversos factores ambientales como la herbivoría (Mauricio, 1998; Züst *et al.*, 2012; Hauser, 2014; Hancock *et al.*, 2011; Swapna *et al.*, 2015; Ning *et al.*, 2016; Li *et al.*, 2018b) por lo que la presencia de tricomas en el fruto probablemente también sea un carácter adaptativo. Apoyando esta hipótesis, las accesiones con tricomas en el fruto y los alelos de Doñ-0 en los polimorfismos funcionales de *TCL1*, *TRY* y *GL1* se distribuyen principalmente en el centro y sur de la península ibérica (Figura 39), lo que sugiere que este carácter podría estar implicado en la adaptación a algún factor ambiental específico de este área geográfica. Sin embargo, las bases genéticas parcialmente diferentes detectadas como causantes de la variación natural para el patrón de tricomas en fase vegetativa y reproductiva sugieren que distintos factores ambientales podrían actuar sobre los distintos órganos. En particular, la presencia de tricomas en los frutos de *Arabidopsis* se ha asociado previamente con bajas precipitaciones y altas temperaturas, sugiriendo su implicación en la adaptación al clima mediterráneo (Savic, 2015). No obstante, actualmente se desconocen los factores ambientales precisos que contribuyen a mantener la variación genética para este carácter en *Arabidopsis*. Estudios futuros similares a los realizados para otros caracteres (Tabas-Madrid *et al.*, 2018; Ferrero-

Serrano & Assmann, 2019), dirigidos a comparar los polimorfismos funcionales identificados en *TCL1*, *TRY* y *GL1* con distintas variables ambientales, permitirán identificar los factores que dirigen la adaptación de *Arabidopsis* a través de modificaciones del patrón de tricomas en el fruto. Es de esperar que estos estudios sean relevantes para numerosas plantas, ya que se ha descrito variación natural para el patrón de tricomas en los frutos de diversas especies silvestres de la familia Brassicaceae, como *Cardamine hirsuta* o *Brassica nigra* (Castroviejo, 1996), y en plantas domesticadas de otras familias botánicas como el pepino o el melocotón (Vendramin *et al.*, 2014; Yang *et al.*, 2014; Xue *et al.*, 2019). Futuros análisis genómicos comparativos entre distintas especies permitirán determinar el grado de conservación de este carácter adaptativo, es decir, si la evolución de este carácter se debe a los mismos o diferentes genes.

# CONCLUSIONES

1. *Arabidopsis* muestra una gran variación natural para el patrón de tricomas en todos los órganos de la planta, con bases genéticas parcialmente distintas en las fases vegetativa y reproductiva del desarrollo.
2. El análisis de asociación genómica a escala regional de la península ibérica complementa los análisis a escala global, ya que permite identificar genes conocidos implicados en la variación natural para el patrón de tricomas (ej. *ETC2* y *MYC1*), genes candidatos identificados en estudios previos (ej. *TCL1*, *TCL2* y *GL1*), pero también nuevos genes candidatos no identificados en los análisis globales.
3. La adquisición de tricomas en los frutos de *Arabidopsis* ha requerido mutaciones en cuatro genes que regulan distintos mecanismos de promoción y represión del patrón de tricomas.
4. *EGL3* es un nuevo gen que contribuye a la variación natural para el patrón de tricomas en el fruto, en el que la accesión Doñ-0 es portadora de un alelo de ganancia de función determinado por un polimorfismo regulador en *cis*.
5. La accesión Doñ-0 es portadora de un nuevo alelo de ganancia de función de *GL1* que es necesario para el desarrollo de tricomas en el fruto y que es regulado por otros genes que también presentan variación natural.
6. La accesión Doñ-0 es portadora de alelos de pérdida de función parcial de *TCL1* y *TRY* causados por polimorfismos reguladores en *cis*.
7. *TCL1* es un regulador en *trans* de *TRY*, por lo que parte de su efecto represor del patrón de tricomas en el fruto está mediado por *TRY*.
8. El desarrollo de tricomas en los frutos de *Arabidopsis* ha evolucionado exclusivamente en el grupo de accesiones relicto de la península ibérica a partir de variación genética existente en los genes *TCL1* y *TRY* y una nueva mutación en *GL1*.



# BIBLIOGRAFÍA

- Abe T., Thitamadee S. & Hashimoto T. (2004). Microtubule defects and cell morphogenesis in the *lefty1lefty2* tubulin mutant of *Arabidopsis thaliana*. *Plant Cell Physiol*, 45, 211-220.
- Alonso-Blanco C. & Méndez-Vigo B. (2014). Genetic architecture of naturally occurring quantitative traits in plants: an updated synthesis. *Curr Opin Plant Biol*, 18, 37-43.
- Alonso-Blanco C., Méndez-Vigo B. & Koornneef M. (2005). From phenotypic to molecular polymorphisms involved in naturally occurring variation of plant development. *Int J Dev Biol*, 49, 717-732.
- Alonso-Blanco C., Koornneef M., & van Ooijen J. W. (2006). QTL analysis. *Methods Mol Biol*, 323, 79-99.
- Alonso-Blanco C., Aarts M. G., Bentsink L., Keurentjes J. J., Reymond M., Vreugdenhil D. & Koornneef M. (2009). What has natural variation taught us about plant development, physiology, and adaptation? *Plant Cell*, 21, 1877-1896.
- Atwell S., Huang Y. S., Vilhjalmsson B. J., Willems G., Horton M., Li Y., Meng D., Platt A., Tarone A. M., Hu T. T., Jiang R., Mulyati N. W., Zhang X., Amer M. A., Baxter I., Brachi B., Chory J., Dean C., Debieu M., de Meaux J., Ecker J. R., Faure N., Kniskern J. M., Jones J. D., Michael T., Nemri A., Roux F., Salt D. E., Tang C., Todesco M., Traw M. B., Weigel D., Marjoram P., Borevitz J. O., Bergelson J. & Nordborg M. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*, 465, 627-631.
- Barua D., Butler C., Tisdale T. E. & Donohue K. (2012). Natural variation in germination responses of *Arabidopsis* to seasonal cues and their associated physiological mechanisms. *Ann. Bot*, 109, 209-226.
- Basu D., Le J., El-Essal Sel D., Huang S., Zhang C., Mallery E. L., Koliantz G., Staiger C. J. & Szymanski D. B. (2005). DISTORTED3/SCAR2 is a putative arabidopsis WAVE complex subunit that activates the Arp2/3 complex and is required for epidermal morphogenesis. *Plant Cell*, 17, 502-524.
- Bazakos C., Hanemian M., Trontin C., Jiménez-Gómez J. M. & Loudet O. (2017). New Strategies and Tools in Quantitative Genetics: How to Go from the Phenotype to the Genotype. *Annu Rev Plant Biol*, 68, 435-455.
- Bell C. J. & Ecker J. R. (1994). Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics*, 19, 137-144.
- Bergelson J. & Roux F. (2010). Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nat Rev Genet*, 11, 867-879.
- Bernatzky R. & Tanksley S. D. (1986). Genetics of actin-related sequences in tomato. *Theor Appl Genet*, 72, 314-321.
- Bernhardt C., Lee M. M., González A., Zhang F., Lloyd A. & Schiefelbein J. (2003). The bHLH genes *GLABRA3* (*GL3*) and *ENHANCER OF GLABRA3* (*EGL3*) specify epidermal cell fate in the *Arabidopsis* root. *Development*, 130, 6431-6439.
- Bigot S., Buges J., Gilly L., Jacques C., Le Boulch P., Berger M., Delcros P., Domergue J. B., Koehl A., Ley-Ngardigal B., Tran Van Canh L. & Couee I. (2018). Pivotal roles of environmental sensing and signaling mechanisms in plant responses to climate change. *Glob Chang Biol*, 24, 5573-5589.
- Bloomer R. H. & Dean C. (2017). Fine-tuning timing: natural variation informs the mechanistic basis of the switch to flowering in *Arabidopsis thaliana*. *J Exp Bot*, 68, 5439-5452.
- Bloomer R. H., Juenger T. E. & Symonds V. V. (2012). Natural variation in *GL1* and its effects on trichome density in *Arabidopsis thaliana*. *Mol Ecol*, 21, 3501-3515.
- Bloomer R. H., Lloyd A. M. & Symonds V. V. (2014). The genetic architecture of constitutive and induced trichome density in two new recombinant inbred line populations of *Arabidopsis thaliana*: phenotypic plasticity, epistasis, and bidirectional leaf damage response. *BMC Plant Biol*, 14, 119.
- Brachi B., Faure N., Bergelson J., Cuguen J. & Roux F. (2013). Genome-wide association mapping of flowering time in *Arabidopsis thaliana* in nature: genetics for underlying components and reaction norms across two successive years. *Acta Bot Gallica*, 160, 205-219.
- Bradbury P. J., Zhang Z., Kroon D. E., Casstevens T. M., Ramdoss Y. & Buckler E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23, 2633-2635.

- Brennan A. C., Méndez-Vigo B., Haddioui A., Martínez-Zapater J. M., Picó F. X. & Alonso-Blanco C. (2014). The genetic structure of *Arabidopsis thaliana* in the south-western Mediterranean range reveals a shared history between North Africa and southern Europe. *BMC Plant Biol*, 14, 17.
- Burk D. H., Liu B., Zhong R., Morrison W. H. & Ye Z. H. (2001). A Katanin-like Protein regulates normal cell wall biosynthesis and cell elongation. *Plant Cell*, 13, 807-827.
- Cao J., Schneeberger K., Ossowski S., Gunther T., Bender S., Fitz J., Koenig D., Lanz C., Stegle O., Lippert C., Wang X., Ott F., Muller J., Alonso-Blanco C., Borgwardt K., Schmid K. J. & Weigel D. (2011). Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat Genet*, 43, 956-963.
- Castroviejo S. (1996). Flora iberica 1-8, 10-15, 17-18, 21. Real Jardín Botánico, CSIC, Madrid.
- Chary S. N., Hicks G. R., Choi Y. G., Carter D. & Raikhel N. V. (2008). Trehalose-6-phosphate synthase/phosphatase regulates cell shape and plant architecture in *Arabidopsis*. *Plant Physiol*, 146, 97-107.
- Clough S. J. (2005). Floral dip: agrobacterium-mediated germ line transformation. *Methods Mol Biol*, 286, 91-102.
- Curaba J., Moritz T., Blervaque R., Parcy F., Raz V., Herzog M. & Vachon G. (2004). *AtGA3ox2*, a key gene responsible for bioactive gibberellin biosynthesis, is regulated during embryogenesis by *LEAFY COTYLEDON2* and *FUSCA3* in *Arabidopsis*. *Plant Physiol*, 136, 3660-3669.
- Czechowski T., Stitt M., Altmann T., Udvardi M. K. & Scheible W. R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol*, 139, 5-17
- Dill A. & Sun T. (2001). Synergistic derepression of gibberellin signaling by removing *RGA* and *GAI* function in *Arabidopsis thaliana*. *Genetics*, 159, 777-785.
- Djakovic S., Dyachok J., Burke M., Frank M. J. & Smith L. G. (2006). *BRICK1/HSPC300* functions with SCAR and the ARP2/3 complex to regulate epidermal cell shape in *Arabidopsis*. *Development*, 133, 1091-1100.
- Dittberner H., Korte A., Mettler-Altmann T., Weber A. P. M., Monroe G. & de Meaux J. (2018). Natural variation in stomata size contributes to the local adaptation of water-use efficiency in *Arabidopsis thaliana*. *Mol Ecol*, 27, 4052-4065.
- Downes B. P., Stupar R. M., Gingerich D. J. & Vierstra R. D. (2003). The HECT ubiquitin-protein ligase (UPL) family in *Arabidopsis*: UPL3 has a specific role in trichome development. *Plant J*, 35, 729-742.
- Durvasula A., Fulgione A., Gutaker R. M., Alacakaptan S. I., Flood P. J., Neto C., Tsuchimatsu T., Burbano H. A., Picó F. X., Alonso-Blanco C. & Hancock A. M. (2017). African genomes illuminate the early history and transition to selfing in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A*, 114, 5213-5218.
- El-Din El-Assal S., Le J., Basu D., Mallery E. L. & Szymanski D. B. (2004). *DISTORTED2* encodes an ARPC2 subunit of the putative *Arabidopsis* ARP2/3 complex. *Plant J*, 38, 526-538.
- Esau, K. (1965). *Plant Anatomy*. John Wiley, New York.
- Esch J. J., Chen M., Sanders M., Hillestad M., Ndkium S., Idelkope B., Neizer J. & Marks M. D. (2003). A contradictory *GLABRA3* allele helps define gene interactions controlling trichome development in *Arabidopsis*. *Development*, 130, 5885-5894.
- Esch J. J., Chen M. A., Hillestad M. & Marks M. D. (2004). Comparison of *TRY* and the closely related *At1g01380* gene in controlling *Arabidopsis* trichome patterning. *Plant J*, 40, 860-869.
- Exner V., Taranto P., Schonrock N., Gruissem W. & Hennig L. (2006). Chromatin assembly factor CAF-1 is required for cellular differentiation during plant development. *Development*, 133, 4163-4172.
- Expósito-Alonso M., Becker C., Schuenemann V. J., Reiter E., Setzer C., Slovak R., Brachi B., Hagmann J., Grimm D. G., Chen J., Busch W., Bergelson J., Ness R. W., Krause J., Burbano H. A. & Weigel D. (2018). The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genet*, 14, e1007155.
- Falbel T. G., Koch L. M., Nadeau J. A., Segui-Simarro J. M., Sack F. D. & Bednarek S. Y. (2003). *SCD1* is required for cytokinesis and polarized cell expansion in *Arabidopsis thaliana*. *Development*, 130, 4011-4024.
- Ferrero-Serrano A. & Assmann S. M. (2019). Phenotypic and genome-wide association with the local environment of *Arabidopsis*. *Nat Ecol Evol*, 3, 274-285.

- Folkers U., Kirik V., Schöbinger U., Falk S., Krishnakumar S., Pollock M. A., Oppenheimer D. G., Day I., Reddy A. S. Jürgens G. & Hülskamp M. (2002). The cell morphogenesis gene *ANGUSTIFOLIA* encodes a CtBP/BARS-like protein and is involved in the control of the microtubule cytoskeleton. *EMBO J*, 21, 1280-1288.
- François O., Martins H., Caye K. & Schoville S. D. (2016). Controlling false discoveries in genome scans for selection. *Mol Ecol*, 25, 454-469.
- Franks S. J. & Hoffmann A. A. (2012). Genetics of climate change adaptation. *Annu Rev Genet*, 46, 185-208.
- Fulgione A. & Hancock A. M. (2018). Archaic lineages broaden our view on the history of *Arabidopsis thaliana*. *New Phytol*, 219, 1194-1198.
- Gan L., Xia K., Chen J. G. & Wang S. (2011). Functional characterization of TRICHOMELESS2, a new single-repeat R3 MYB transcription factor in the regulation of trichome patterning in *Arabidopsis*. *BMC Plant Biol*, 11, 176.
- Gan Y., Kumimoto R., Liu C., Ratcliffe O., Yu H. & Broun P. (2006). *GLABROUS INFLORESCENCE STEMS* modulates the regulation by gibberellins of epidermal differentiation and shoot maturation in *Arabidopsis*. *Plant Cell*, 18, 1383-1395.
- Gan Y., Liu C., Yu H. & Broun P. (2007). Integration of cytokinin and gibberellin signalling by *Arabidopsis* transcription factors *GIS*, *ZFP8* and *GIS2* in the regulation of epidermal cell fate. *Development*, 134, 2073-2081.
- Gao Y., Gong X., Cao W., Zhao J., Fu L., Wang X., Schumaker K. S. & Guo Y. (2008). *SAD2* in *Arabidopsis* functions in trichome initiation through mediating *GL3* function and regulating *GL1*, *TTG1* and *GL2* expression. *J Integr Plant Biol*, 50, 906-917.
- Godfray H. C., Beddington J. R., Crute I. R., Haddad L., Lawrence D., Muir J. F., Pretty J., Robinson S., Thomas S. M. & Toulmin C. (2010). Food security: the challenge of feeding 9 billion people. *Science*, 327, 812-818.
- Gould N., Doulias P. T., Tenopoulou M., Raju K. & Ischiropoulos H. (2013). Regulation of protein function and signaling by reversible cysteine S-nitrosylation. *J Biol Chem*, 288, 26473-26479.
- Grebe M. (2012). The patterning of epidermal hairs in *Arabidopsis*--updated. *Curr Opin Plant Biol*, 15, 31-37.
- Hancock A. M., Brachi B., Faure N., Horton M. W., Jarymowycz L. B., Sperone F. G., Toomajian C., Roux F. & Bergelson, J. (2011). Adaptation to climate across the *Arabidopsis thaliana* genome. *Science*, 334, 83-86.
- Handley R., Ekbom B. & Ågren J. (2005) Variation in trichome density and resistance against a specialist insect herbivore in natural populations of *Arabidopsis thaliana*. *Ecol Entomol*, 30, 284-292.
- Hase Y., Trung K. H., Matsunaga T. & Tanaka A. (2006). A mutation in the *uvi4* gene promotes progression of endo-reduplication and confers increased tolerance towards ultraviolet B light. *Plant J*, 46, 317-326.
- Hauser M. T. (2014). Molecular basis of natural variation and environmental control of trichome patterning. *Front Plant Sci*, 5, 320.
- Hauser M. T., Harr B. & Schlotterer C. (2001). Trichome distribution in *Arabidopsis thaliana* and its close relative *Arabidopsis lyrata*: molecular analysis of the candidate gene *GLABROUS1*. *Mol Biol Evol*, 18, 1754-1763.
- Hayashi S., Ishii T., Matsunaga T., Tominaga R., Kuromori T., Wada T., Shinozaki K. & Hirayama T. (2008). The glycerophosphoryl diester phosphodiesterase-like proteins SHV3 and its homologs play important roles in cell wall organization. *Plant Cell Physiol*, 49, 1522-1535.
- Hilscher J., Schlotterer C. & Hauser M. T. (2009). A single amino acid replacement in *ETC2* shapes trichome patterning in natural *Arabidopsis* populations. *Curr Biol*, 19, 1747-1751.
- Horton M. W., Hancock A. M., Huang Y. S., Toomajian C., Atwell S., Auton A., Mulyati N. W., Platt A., Sperone F. G., Vilhjalmsson B. J., Nordborg M., Borevitz J. O. & Bergelson J. (2012). Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nat Genet*, 44, 212-216.
- Horton M. W., Willems G., Sasaki E., Koornneef M. & Nordborg M. (2016). The genetic architecture of freezing tolerance varies across the range of *Arabidopsis thaliana*. *Plant Cell Environ*, 39, 2570-2579.
- Hülskamp M., Misra S. & Jürgens G. (1994). Genetic dissection of trichome cell development in *Arabidopsis*. *Cell*, 76, 555-566.
- Hülskamp M. & Schnittger A. (1998). Spatial regulation of trichome formation in *Arabidopsis thaliana*. *Semin Cell Dev Biol*, 9, 213-220.

- Hülskamp M., Schnittger A. & Folkers U. (1999). Pattern formation and cell differentiation: trichomes in Arabidopsis as a genetic model system. *Int Rev Cytol*, 186, 147-178.
- Ilgenfritz H., Bouyer D., Schnittger A., Mathur J., Kirik V., Schwab B., Chua N. H., Jurgens G. & Hülskamp M. (2003). The Arabidopsis *STICHEL* gene is a regulator of trichome branch number and encodes a novel protein. *Plant Physiol*, 131, 643-655.
- Imai K. K., Ohashi Y., Tsuge T., Yoshizumi T., Matsui M., Oka A. & Aoyama T. (2006). The A-type cyclin *CYCA2;3* is a key regulator of ploidy levels in Arabidopsis endoreduplication. *Plant Cell*, 18, 382-396.
- Ishida T., Kurata T., Okada K. & Wada T. (2008). A genetic regulatory network in the development of trichomes and root hairs. *Annu Rev Plant Biol*, 59, 365-386.
- Ito T., Kim G. T. & Shinozaki K. (2000). Disruption of an Arabidopsis cytoplasmic ribosomal protein S13-homologous gene by transposon-mediated mutagenesis causes aberrant growth and development. *Plant J*, 22, 257-264.
- Jakoby M. J., Falkenhan D., Mader M. T., Brininstool G., Wischnitzki E., Platz N., Hudson A., Hülskamp M., Larkin J. & Schnittger A. (2008). Transcriptional profiling of mature Arabidopsis trichomes reveals that *NOECK* encodes the MIXTA-like transcriptional regulator MYB106. *Plant Physiol*, 148, 1583-1602.
- Jamieson M. A., Trowbridge A. M., Raffa K. F. & Lindroth R. L. (2012). Consequences of climate warming and altered precipitation patterns for plant-insect and multitrophic interactions. *Plant Physiol*, 160, 1719-1727.
- Johnson C. S., Kolevski B. & Smyth D. R. (2002). *TRANSPARENT TESTA GLABRA2*, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. *Plant Cell*, 14, 1359-1375.
- Kang B. H., Busse J. S. & Bednarek S. Y. (2003). Members of the Arabidopsis dynamin-like gene family, *ADL1*, are essential for plant cytokinesis and polarized cell growth. *Plant Cell*, 15, 899-913.
- Kärkkäinen K. & Ågren J. (2002). Genetic basis of trichome production in *Arabidopsis lyrata*. *Hereditas*, 136, 219-226.
- Kasulin L., Rowan B. A., Leon R. J. C., Schuenemann V. J., Weigel D. & Botto J. F. (2017). A single haplotype hyposensitive to light and requiring strong vernalization dominates *Arabidopsis thaliana* populations in Patagonia, Argentina. *Mol Ecol*, 26, 3389-3404.
- Kawagoe T., Shimizu K. K., Kakutani T. & Kudo H. (2011). Coexistence of trichome variation in a natural plant population: a combined study using ecological and candidate gene approaches. *PLoS One*, 6, e22184.
- Kerstetter R. A., Bollman K., Taylor R. A., Bomblies K. & Poethig R. S. (2001). *KANADI* regulates organ polarity in Arabidopsis. *Nature*, 411, 706-709.
- Kim S. Y., Hyoung S., So W. M. & Shin J. S. (2018). The novel transcription factor TRP interacts with ZFP5, a trichome initiation-related transcription factor, and negatively regulates trichome initiation through gibberellic acid signaling. *Plant Mol Biol*, 96, 315-326.
- Kirik V., Bouyer D., Schobinger U., Bechtold N., Herzog M., Bonneville J. M. & Hülskamp M. (2001a). *CPR5* is involved in cell proliferation and cell death control and encodes a novel transmembrane protein. *Curr Biol*, 11, 1891-1895.
- Kirik V., Schnittger A., Radchuk V., Adler K., Hülskamp M. & Baumlein H. (2001b). Ectopic expression of the Arabidopsis *AtMYB23* gene induces differentiation of trichome cells. *Dev Biol*, 235, 366-377.
- Kirik V., Grini P. E., Mathur J., Klinkhammer I., Adler K., Bechtold N., Herzog M., Bonneville J. M. & Hülskamp M. (2002a). The Arabidopsis *TUBULIN-FOLDING COFACTOR A* gene is involved in the control of the  $\alpha/\beta$ -tubulin monomer balance. *Plant Cell*, 14, 2265-2276.
- Kirik V., Mathur J., Grini P. E., Klinkhammer I., Adler K., Bechtold N., Herzog M., Bonneville J. M. & Hülskamp M. (2002b). Functional analysis of the Tubulin-Folding Cofactor C in *Arabidopsis thaliana*. *Curr Biol*, 12, 1519-1523.
- Kirik V., Simon M., Hülskamp M. & Schiefelbein J. (2004a). The *ENHANCER OF TRY AND CPC1* gene acts redundantly with *TRIPTYCHON* and *CAPRICE* in trichome and root hair cell patterning in Arabidopsis. *Dev Biol*, 268, 506-513.
- Kirik V., Simon M., Wester K., Schiefelbein J. & Hülskamp M. (2004b). *ENHANCER of TRY and CPC 2 (ETC2)* reveals redundancy in the region-specific control of trichome development of Arabidopsis. *Plant Mol Biol*, 55, 389-398.
- Kirik V., Lee M. M., Wester K., Herrmann U., Zheng Z., Oppenheimer D., Schiefelbein J. & Hülskamp M. (2005). Functional diversification of *MYB23* and *GL1* genes in trichome morphogenesis and initiation. *Development*, 132, 1477-1485.

- Kivimäki M., Karkkainen K., Gaudeul M., Loe G. & Agren J. (2007). Gene, phenotype and function: *GLABROUS1* and resistance to herbivory in natural populations of *Arabidopsis lyrata*. *Mol Ecol*, 16, 453-462.
- Klasen J. R., Barbez E., Meier L., Meinshausen N., Buhmann P., Koornneef M., Busch W. & Schneeberger K. (2016). A multi-marker association method for genome-wide association studies without the need for population structure correction. *Nat Commun*, 7, 13299.
- Konieczny A. & Ausubel F. M. (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant J*, 4, 403-410.
- Koo A. J. & Howe G. A. (2009). The wound hormone jasmonate. *Phytochemistry*, 70, 1571-1580.
- Kooke R., Kruijjer W., Bours R., Becker F., Kuhn A., van de Geest H., Buntjer J., Doeswijk T., Guerra J., Bouwmeester H., Vreugdenhil D. & Keurentjes J. J. (2016). Genome-wide association mapping and genomic prediction elucidate the genetic architecture of morphological traits in *Arabidopsis*. *Plant Physiol*, 170, 2187-2203.
- Koornneef M. & Meinke D. (2010). The development of *Arabidopsis* as a model plant. *Plant J*, 61, 909-921.
- Koornneef M., Dellaert L. W. & van der Veen J. H. (1982). EMS- and radiation-induced mutation frequencies at individual loci in *Arabidopsis thaliana* (L.) Heynh. *Mutat Res*, 93, 109-123.
- Koornneef M., Alonso-Blanco C. & Vreugdenhil D. (2004). Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu Rev Plant Biol*, 55, 141-172.
- Krämer U. (2015). Planting molecular functions in an ecological context with *Arabidopsis thaliana*. *Elife*, 4, e06100.
- Kronholm I., Picó F. X., Alonso-Blanco C., Goudet J. & de Meaux J. (2012). Genetic basis of adaptation in *Arabidopsis thaliana*: local adaptation at the seed dormancy QTL *DOG1*. *Evolution*, 66, 2287-2302.
- Kurata T., Kawabata-Awai C., Sakuradani E., Shimizu S., Okada K. & Wada T. (2003). The *YORE-YORE* gene regulates multiple aspects of epidermal cell differentiation in *Arabidopsis*. *Plant J*, 36, 55-66.
- Kurata T., Ishida T., Kawabata-Awai C., Noguchi M., Hattori S., Sano R., Nagasaka R., Tominaga R., Koshino-Kimura Y., Kato T., Sato S., Tabata S., Okada K. & Wada T. (2005). Cell-to-cell movement of the CAPRICE protein in *Arabidopsis* root epidermal cell differentiation. *Development*, 132, 5387-5398.
- Larkin J. C., Oppenheimer D. G. & Marks M. D. (1994). The *GL1* gene and the trichome developmental pathway in *Arabidopsis thaliana*. *Results Probl Cell Differ*, 20, 259-275.
- Larkin J. C., Young N., Prigge M. & Marks M. D. (1996). The control of trichome spacing and number in *Arabidopsis*. *Development*, 122, 997-1005.
- Larkin J. C., Brown M. L. & Schiefelbein J. (2003). How do cells know what they want to be when they grow up? Lessons from epidermal patterning in *Arabidopsis*. *Annu Rev Plant Biol*, 54, 403-430.
- Lazo G. R., Stein P. A. & Ludwig R. A. (1991). A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. *Biotechnology (N Y)*, 9, 963-967.
- Leasure C. D., Fiume E. & Fletcher J. C. (2009). The essential gene *EMB1611* maintains shoot apical meristem function during *Arabidopsis* development. *Plant J*, 57, 579-592.
- Lee C. R., Svardal H., Farlow A., Exposito-Alonso M., Ding W., Novikova P., Alonso-Blanco C., Weigel D. & Nordborg M. (2017). On the post-glacial spread of human commensal *Arabidopsis thaliana*. *Nat Commun*, 8, 14458.
- Li F., Zou Z., Yong H. Y., Kitashiba H. & Nishio T. (2013). Nucleotide sequence variation of *GLABRA1* contributing to phenotypic variation of leaf hairiness in Brassicaceae vegetables. *Theor Appl Genet*, 126, 1227-1236.
- Li S. F., Milliken O. N., Pham H., Seyit R., Napoli R., Preston J., Koltunow A. M. & Parish R. W. (2009). The *Arabidopsis* MYB5 transcription factor regulates mucilage synthesis, seed coat development, and trichome morphogenesis. *Plant Cell*, 21, 72-89.
- Li S., Tosens T., Harley P. C., Jiang Y., Kanagendran A., Grosberg M., Jaamets K. & Niinemets U. (2018b). Glandular trichomes as a barrier against atmospheric oxidative stress: Relationships with ozone uptake, leaf damage, and emission of LOX products across a diverse set of species. *Plant Cell Environ*, 41, 1263-1277.
- Li Y., Sorefan K., Hemmann G., & Bevan M. W. (2004). *Arabidopsis* *NAP* and *PIR* regulate actin-based cell morphogenesis and multiple developmental processes. *Plant Physiol*, 136, 3616-3627.

- Li Y., Xiao J., Chen L., Huang X., Cheng Z., Han B., Zhang Q. & Wu C. (2018a). Rice Functional Genomics Research: Past Decade and Future. *Mol Plant*, 11, 359-380.
- Liang G., He H., Li Y., Ai Q. & Yu D. (2014). MYB82 functions in regulation of trichome development in Arabidopsis. *J Exp Bot*, 65, 3215-3223.
- Lid S. E., Olsen L., Nestestog R., Aukerman M., Brown R. C., Lemmon B., Mucha M., Opsahl-Sorteberg H. G. & Olsen O. A. (2005). Mutation in the Arabidopsis thaliana DEK1 calpain gene perturbs endosperm and embryo development while over-expression affects organ development globally. *Planta*, 221, 339-351.
- Long Q., Rabanal F. A., Meng D., Huber C. D., Farlow A., Platzer A., Zhang Q., Vilhjalmsson B. J., Korte A., Nizhynska V., Voronin V., Korte P., Sedman L., Mandakova T., Lysak, M. A., Seren U., Hellmann I. & Nordborg M. (2013). Massive genomic variation and strong selection in Arabidopsis thaliana lines from Sweden. *Nat Genet*, 45, 884-890.
- Loudet O., Chaillou S., Camilleri C., Bouchez D. & Daniel-Vedele F. (2002). Bay-0 x Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in Arabidopsis. *Theor Appl Genet*, 104, 1173-1184.
- Lu L., Lee Y. R., Pan R., Maloof J. N. & Liu B. (2005). An internal motor kinesin is associated with the Golgi apparatus and plays a role in trichome morphogenesis in Arabidopsis. *Mol Biol Cell*, 16, 811-823.
- Lukowitz W., Gillmor C. S. & Scheible W. R. (2000). Positional cloning in Arabidopsis. Why it feels good to have a genome initiative working for you. *Plant Physiol*, 123, 795-805.
- Maes L., Inze D. & Goossens A. (2008). Functional specialization of the TRANSPARENT TESTA GLABRA1 network allows differential hormonal control of laminal and marginal trichome initiation in Arabidopsis rosette leaves. *Plant Physiol*, 148, 1453-1464.
- Manzano-Piedras E., Marcer A., Alonso-Blanco C. & Picó F. X. (2014). Deciphering the adjustment between environment and life history in annuals: lessons from a geographically-explicit approach in Arabidopsis thaliana. *PLoS One*, 9, e87836.
- Marks M. D., Wenger J. P., Gilding E., Jilk R. & Dixon R. A. (2009). Transcriptome analysis of Arabidopsis wild-type and gl3-sst sim trichomes identifies four additional genes required for trichome development. *Mol Plant*, 2, 803-822.
- Masucci J. D. & Schiefelbein J. W. (1994). The rhd6 Mutation of Arabidopsis thaliana alters root-hair initiation through an auxin- and ethylene-associated process. *Plant Physiol*, 106, 1335-1346.
- Mathur J., Spielhofer P., Kost B. & Chua N. (1999). The actin cytoskeleton is required to elaborate and maintain spatial patterning during trichome cell morphogenesis in Arabidopsis thaliana. *Development*, 126, 5559-5568.
- Mathur J., Mathur N., Kernebeck B. & Hülskamp M. (2003a). Mutations in actin-related proteins 2 and 3 affect cell shape development in Arabidopsis. *Plant Cell*, 15, 1632-1645.
- Mathur J., Mathur N., Kirik V., Kernebeck B., Srinivas B. P. & Hülskamp M. (2003b). Arabidopsis CROOKED encodes for the smallest subunit of the ARP2/3 complex and controls cell shape by region specific fine F-actin formation. *Development*, 130, 3137-3146.
- Matias-Hernandez L., Aguilar-Jaramillo A. E., Osnato M., Weinstain R., Shani E., Suarez-Lopez P. & Pelaz S. (2016). TEMPRANILLO reveals the mesophyll as crucial for epidermal trichome formation. *Plant Physiol*, 170, 1624-1639.
- Mauricio R. (1998). Costs of resistance to natural enemies in field populations of the annual plant Arabidopsis thaliana. *Am Nat*, 151, 20-28.
- Mauricio R. (2005). Ontogenetics of QTL: the genetic architecture of trichome density over time in Arabidopsis thaliana. *Genetica*, 123, 75-85.
- Méndez-Vigo B., Picó F. X., Ramiro M., Martínez-Zapater J. M. & Alonso-Blanco C. (2011). Altitudinal and climatic adaptation is mediated by flowering traits and FRI, FLC, and PHYC genes in Arabidopsis. *Plant Physiol*, 157, 1942-1955.
- Méndez-Vigo B., Savic M., Ausín I., Ramiro M., Martín B., Picó F. X. & Alonso-Blanco C. (2016). Environmental and genetic interactions reveal FLOWERING LOCUS C as a modulator of the natural variation for the plasticity of flowering in Arabidopsis. *Plant Cell Environ*, 39, 282-294.
- Melzer S., Kampmann G., Chandler J. & Apel K. (1999). FPF1 modulates the competence to flowering in Arabidopsis. *Plant J*, 18, 395-405.
- Mir R. R., Zaman-Allah M., Sreenivasulu N., Trethowan R. & Varshney R. K. (2012). Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. *Theor Appl Genet*, 125, 625-645.

- Mullins R. D., Heuser J. A. & Pollard T. D. (1998). The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proc Natl Acad Sci U S A*, 95, 6181-6186.
- Nagawa S., Xu T., Lin D., Dhonukshe P., Zhang X., Friml J., Scheres B., Fu, Y. & Yang, Z. (2012). ROP GTPase-dependent actin microfilaments promote *PIN1* polarization by localized inhibition of clathrin-dependent endocytosis. *PLoS Biol*, 10, e1001299.
- Nakamura M., Katsumata H., Abe M., Yabe N., Komeda Y., Yamamoto K. T. & Takahashi T. (2006). Characterization of the class IV homeodomain-Leucine Zipper gene family in *Arabidopsis*. *Plant Physiol*, 141, 1363-1375.
- Ning P., Wang J., Zhou Y., Gao L., Wang J. & Gong C. (2016). Adaptational evolution of trichome in *Caragana korshinskii* to natural drought stress on the Loess Plateau, China. *Ecol Evol*, 6, 3786-3795.
- Nordborg M., Hu T. T., Ishino Y., Jhaveri J., Toomajian C., Zheng H., Bakker E., Calabrese P., Gladstone J., Goyal R., Jakobsson M., Kim S., Morozov Y., Padhukasahasram B., Plagnol V., Rosenberg N. A., Shah C., Wall J. D., Wang J., Zhao K., Kalbfleisch T., Schulz V., Kreitman M. & Bergelson J. (2005). The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol*, 3, e196.
- Ojangu E. L., Jarve K., Paves H. & Truve E. (2007). *Arabidopsis thaliana* myosin *XIK* is involved in root hair as well as trichome morphogenesis on stems and leaves. *Protoplasma*, 230, 193-202.
- Oppenheimer D. G., Herman P. L., Sivakumaran S., Esch, J. & Marks M. D. (1991). A myb gene required for leaf trichome differentiation in *Arabidopsis* is expressed in stipules. *Cell*, 67, 483-493.
- Oppenheimer D. G., Pollock M. A., Vacik J., Szymanski D. B., Ericson B., Feldmann K. & Marks M. D. (1997). Essential role of a kinesin-like protein in *Arabidopsis* trichome morphogenesis. *Proc Natl Acad Sci USA*, 94, 6261-6266.
- Ó'Maoiléidigh D. S., Stewart D., Zheng B., Coupland G. & Wellmer F. (2018). Floral homeotic proteins modulate the genetic program for leaf development to suppress trichome formation in flowers. *Development*, 145.
- Orr H. A. (2005). The genetic theory of adaptation: a brief history. *Nat Rev Genet*, 6, 119-127.
- Pattanaik S., Patra B., Singh S. K. & Yuan L. (2014). An overview of the gene regulatory network controlling trichome development in the model plant, *Arabidopsis*. *Front Plant Sci*, 5, 259.
- Payne C. T., Zhang F. & Lloyd A. M. (2000). *GL3* encodes a bHLH protein that regulates trichome development in *Arabidopsis* through interaction with *GL1* and *TTG1*. *Genetics*, 156, 1349-1362.
- Perazza D., Vachon G. & Herzog M. (1998). Gibberellins promote trichome formation by Up-regulating *GLABROUS1* in *Arabidopsis*. *Plant Physiol*, 117, 375-383.
- Perazza D., Herzog M., Hülskamp M., Brown S., Dorne A. M. & Bonneville J. M. (1999). Trichome cell growth in *Arabidopsis thaliana* can be derepressed by mutations in at least five genes. *Genetics*, 152, 461-476.
- Pfalz M., Vogel H., Mitchell-Olds T., & Kroymann J. (2007). Mapping of QTL for resistance against the crucifer specialist herbivore *Pieris brassicae* in a new *Arabidopsis* inbred line population, Da(1)-12 x Ei-2. *PLoS One*, 2, e578.
- Picó F. X., Méndez-Vigo B., Martínez-Zapater J. M. & Alonso-Blanco C. (2008). Natural genetic variation of *Arabidopsis thaliana* is geographically structured in the Iberian Peninsula. *Genetics*, 180, 1009-1021.
- Pieterse C. M., Van der Does D., Zamioudis C., Leon-Reyes A. & Van Wees S. C. (2012). Hormonal modulation of plant immunity. *Annu Rev Cell Dev Biol*, 28, 489-521.
- Pillitteri L. J., Bogenschutz N. L. & Torii K. U. (2008). The bHLH protein, MUTE, controls differentiation of stomata and the hydathode pore in *Arabidopsis*. *Plant Cell Physiol*, 49, 934-943.
- Provart N. J., Alonso J., Assmann S. M., Bergmann D., Brady S. M., Brkljacic J., Browse J., Chapple C., Colot V., Cutler S., Dangl J., Ehrhardt D., Friesner J. D., Frommer W. B., Grotewold E., Meyerowitz E., Nemhauser J., Nordborg M., Pikaard C., Shanklin J., Somerville C., Stitt M., Torii K. U., Waese J., Wagner D. & McCourt P. (2016). 50 years of *Arabidopsis* research: highlights and future directions. *New Phytol*, 209, 921-944.
- Qiu J. L., Jilk R., Marks M. D. & Szymanski D. B. (2002). The *Arabidopsis SPIKE1* gene is required for normal cell shape control and tissue development. *Plant Cell*, 14, 101-118.



- Qi T., Song S., Ren Q., Wu D., Huang H., Chen Y., Fan M., Peng W., Ren C. & Xie D. (2011). The Jasmonate-ZIM-domain proteins interact with the WD-Repeat/bHLH/MYB complexes to regulate Jasmonate-mediated anthocyanin accumulation and trichome initiation in *Arabidopsis thaliana*. *Plant Cell*, 23, 1795-1814.
- Qi T., Huang H., Wu D., Yan J., Qi Y., Song S. & Xie D. (2014). Arabidopsis DELLA and JAZ proteins bind the WD-repeat/bHLH/MYB complex to modulate gibberellin and jasmonate signaling synergy. *Plant Cell*, 26, 1118-1133.
- Reddy V. S., Day I. S., Thomas T. & Reddy A. S. (2004). KIC, a novel Ca<sup>2+</sup> binding protein with one EF-hand motif, interacts with a microtubule motor protein and regulates trichome morphogenesis. *Plant Cell*, 16, 185-200.
- Rerie W. G., Feldmann K. A. & Marks M. D. (1994). The *GLABRA2* gene encodes a homeo domain protein required for normal trichome development in Arabidopsis. *Genes Dev*, 8, 1388-1399.
- Rozas J., Ferrer-Mata A., Sánchez-DelBarrio J. C., Guirao-Rico S., Librado P., Ramos-Onsins S. E. & Sánchez-Gracia A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol Biol Evol*, 34, 3299-3302.
- Saedler R., Zimmermann I., Mutondo M. & Hülskamp M. (2004). The Arabidopsis *KLUNKER* gene controls cell shape changes and encodes the AtSRA1 homolog. *Plant Mol Biol*, 56, 775-782.
- Saedler R., Jakoby M., Marin B., Galiana-Jaime E. & Hülskamp M. (2009). The cell morphogenesis gene *SPIRRIG* in Arabidopsis encodes a WD/BEACH domain protein. *Plant J*, 59, 612-621.
- Savic M. (2015) A genetic, molecular and ecological study of the natural variation for trichome patterning in *Arabidopsis thaliana*. Tesis Doctoral.
- Savolainen O., Lascoux M. & Merila J. (2013). Ecological genomics of local adaptation. *Nat Rev Genet*, 14, 807-820.
- Schellmann S. & Hülskamp M. (2005). Epidermal differentiation: trichomes in Arabidopsis as a model system. *Int J Dev Biol*, 49, 579-584.
- Schellmann S., Schnittger A., Kirik V., Wada T., Okada K., Beermann A., Thumfahrt J., Jurgens G. & Hülskamp M. (2002). *TRIPTYCHON* and *CAPRICE* mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. *EMBO J*, 21, 5036-5046.
- Schliep M., Ebert B., Simon-Rosin U., Zoeller D. & Fisahn J. (2010). Quantitative expression analysis of selected transcription factors in pavement, basal and trichome cells of mature leaves from *Arabidopsis thaliana*. *Protoplasma*, 241, 29-36.
- Schnittger A. & Hülskamp M. (2002). Trichome morphogenesis: a cell-cycle perspective. *Philos Trans R Soc Lond B Biol Sci*, 357, 823-826.
- Schnittger A., Weinl C., Bouyer D., Schobinger U. & Hülskamp, M. (2003). Misexpression of the cyclin-dependent kinase inhibitor *ICK1/KRP1* in single-celled Arabidopsis trichomes reduces endoreduplication and cell size and induces cell death. *Plant Cell*, 15, 303-315.
- Schwab B., Folkers U., Ilgenfritz H. & Hülskamp M. (2000). Trichome morphogenesis in Arabidopsis. *Philos Trans R Soc Lond B Biol Sci*, 355, 879-883.
- Shikata M., Koyama T., Mitsuda N. & Ohme-Takagi M. (2009). Arabidopsis SBP-box genes *SPL10*, *SPL11* and *SPL2* control morphological change in association with shoot maturation in the reproductive phase. *Plant Cell Physiol*, 50, 2133-2145.
- Shimada T., Sugano S. S. & Hara-Nishimura I. (2011). Positive and negative peptide signals control stomatal density. *Cell Mol Life Sci*, 68, 2081-2088.
- Shimamoto K. & Kyojuka J. (2002). Rice as a model for comparative genomics of plants. *Annu Rev Plant Biol*, 53, 399-419.
- Spitzer C., Schellmann S., Sabovljevic A., Shahriari M., Keshavaiah C., Bechtold N., Herzog M., Müller S., Hanisch F. G. & Hülskamp M. (2006). The Arabidopsis *elch* mutant reveals functions of an ESCRT component in cytokinesis. *Development*, 133, 4679-4689.
- Sugimoto-Shirasu K., Stacey N. J., Corsar, J., Roberts, K. & McCann, M. C. (2002). DNA topoisomerase VI is essential for endoreduplication in Arabidopsis. *Curr Biol*, 12, 1782-1786.
- Sun L., Zhang A., Zhou Z., Zhao Y., Yan A., Bao S., Yu H. & Gan Y. (2015). *GLABROUS INFLORESCENCE STEMS3 (GIS3)* regulates trichome initiation and development in Arabidopsis. *New Phytol*, 206, 220-230.
- Svitkina T. M. & Borisy G. G. (1999). Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *J Cell Biol*, 145, 1009-1026.

- Swapna K. S., Salim N., Chandra R. & Puthur J. T. (2015). Structural changes in response to bioaccumulation of iron and mercury in *Chromolaena odorata* (L.) King & Robins. *Environ Monit Assess*, 187, 551.
- Symonds V. V., Godoy A. V., Alconada T., Botto J. F., Juenger T. E., Casal J. J. & Lloyd A. M. (2005). Mapping quantitative trait loci in multiple populations of *Arabidopsis thaliana* identifies natural allelic variation for trichome density. *Genetics*, 169, 1649-1658.
- Symonds V. V., Hatlestad G. & Lloyd A. M. (2011). Natural allelic variation defines a role for *ATMYC1*: trichome cell fate determination. *PLoS Genet*, 7, e1002069.
- Szymanski D. B. & Marks M. D. (1998). *GLABROUS1* overexpression and *TRIPTYCHON* alter the cell cycle and trichome cell fate in *Arabidopsis*. *Plant Cell*, 10, 2047-2062.
- Szymanski D. B., Marks M. D. & Wick S. M. (1999). Organized F-actin is essential for normal trichome morphogenesis in *Arabidopsis*. *Plant Cell*, 11, 2331-2347.
- Tabas-Madrid D., Méndez-Vigo B., Arteaga N., Marcer A., Pascual-Montano A., Weigel D., Picó F. X. & Alonso-Blanco C. (2018). Genome-wide signatures of flowering adaptation to climate temperature: Regional analyses in a highly diverse native range of *Arabidopsis thaliana*. *Plant Cell Environ*, 41, 1806–1820.
- Tamura K., Stecher G., Peterson D., Filipinski A. & Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*, 30, 2725-2729.
- Telfer A., Bollman K. M. & Poethig R. S. (1997). Phase change and the regulation of trichome distribution in *Arabidopsis thaliana*. *Development*, 124, 645-654.
- Telfer A. & Poethig R. S. (1998). *HASTY*: a gene that regulates the timing of shoot maturation in *Arabidopsis thaliana*. *Development*, 125, 1889-1898.
- The 1001 Genomes Consortium. (2016). 1,135 Genomes Reveal the Global Pattern of Polymorphism in *Arabidopsis thaliana*. *Cell*, 166, 481-491.
- Tian H., Qi T., Li Y., Wang C., Ren C., Song S. & Huang H. (2016). Regulation of the WD-repeat/bHLH/MYB complex by gibberellin and jasmonate. *Plant Signal Behav*, 11, e1204061.
- Tian H., Wang X., Guo H., Cheng Y., Hou C., Chen J. G. & Wang S. (2017). *NTL8* regulates trichome formation in *Arabidopsis* by directly activating R3 MYB genes *TRY* and *TCL1*. *Plant Physiol*, 174, 2363-2375.
- Thines B., Katsir L., Melotto M., Niu Y., Mandaokar A., Liu G., Nomura K., He S. Y., Howe G. A. & Browse J. (2007). JAZ repressor proteins are targets of the SCF(CO1) complex during jasmonate signalling. *Nature*, 448, 661-665.
- Togninalli M., Seren U., Meng D., Fitz J., Nordborg M., Weigel D., Borgwardt K., Korte A. & Grimm D. G. (2018). The AraGWAS Catalog: a curated and standardized *Arabidopsis thaliana* GWAS catalog. *Nucleic Acids Res*, 46, D1150-D1156.
- Tominaga-Wada R., Ishida T. & Wada T. (2011). New insights into the mechanism of development of *Arabidopsis* root hairs and trichomes. *Int Rev Cell Mol Biol*, 286, 67-106.
- Traw M. B. & Bergelson J. (2003). Interactive effects of jasmonic acid, salicylic acid, and gibberellin on induction of trichomes in *Arabidopsis*. *Plant Physiol*, 133, 1367-1375.
- Unte U. S., Sorensen A. M., Pesaresi P., Gandikota M., Leister D., Saedler H. & Huijser P. (2003). *SPL8*, an SBP-box gene that affects pollen sac development in *Arabidopsis*. *Plant Cell*, 15, 1009-1019.
- Valladares F., Sánchez-Gómez D. & Zavala M.A. (2006). Quantitative estimation of phenotypic plasticity: bridging the gap between the evolutionary concept and its ecological applications. *J Ecol*, 94, 1103-1116.
- Valverde P. L., Fornoni J. & Núñez-Farfán J. (2001). Defensive role of leaf trichomes in resistance to herbivorous insects in *Datura stramonium*. *J. Evol. Biol*, 14, 424-432.
- Vanzin G. F., Madson M., Carpita N. C., Raikhel N. V., Keegstra K. & Reiter W. D. (2002). The *mur2* mutant of *Arabidopsis thaliana* lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. *Proc Natl Acad Sci USA*, 99, 3340-3345.
- Vendramin E., Pea G., Dondini L., Pacheco I., Dettori M. T., Gazza L., Scalabrin S., Strozzi F., Tartarini S., Bassi D., Verde I. & Rossini L. (2014). A unique mutation in a MYB gene cosegregates with the nectarine phenotype in peach. *PLoS One*, 9, e90574.
- Vidigal D. S., Marques A. C., Willems L. A., Buijs G., Méndez-Vigo B., Hilhorst H. W., Bentsink L., Picó F. X. & Alonso-Blanco C. (2016). Altitudinal and climatic associations of seed dormancy and flowering traits evidence adaptation of annual life cycle timing in *Arabidopsis thaliana*. *Plant Cell Environ*, 39, 1737-1748.
- Wada T., Tachibana T., Shimura Y. & Okada K. (1997). Epidermal cell differentiation in *Arabidopsis* determined by a Myb homolog, *CPC*. *Science*, 277, 1113-1116.

- Walker A. R., Davison P. A., Bolognesi-Winfield A. C., James C. M., Srinivasan N., Blundell T. L., Esch J. J. Marks M. D. & Gray J. C. (1999). The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. *Plant Cell*, 11, 1337-1350.
- Walker J. D., Oppenheimer D. G., Concienne J. & Larkin J. C. (2000). *SIAMESE*, a gene controlling the endoreduplication cell cycle in *Arabidopsis thaliana* trichomes. *Development*, 127, 3931-3940.
- Wan T., Liu Z. M., Li L. F., Leitch A. R., Leitch I. J., Lohaus R., Liu Z. J., Xin H. P., Gong Y. B., Liu Y., Wang W. C., Chen L. Y., Yang Y., Kelly L. J., Yang J., Huang J. L., Li Z., Liu P., Zhang L., Liu H. M., Wang H., Deng S. H., Liu M., Li J., Ma L., Liu Y., Lei Y., Xu W., Wu L. Q., Liu F., Ma Q., Yu X. R., Jiang Z., Zhang G. Q., Li S. H., Li R. Q., Zhang S. Z., Wang Q. F., Van de Peer Y., Zhang J. B. & Wang X. M. (2018). A genome for gnetophytes and early evolution of seed plants. *Nat Plants*, 4, 82-89.
- Wang S. & Chen J. G. (2008). Arabidopsis transient expression analysis reveals that activation of *GLABRA2* may require concurrent binding of *GLABRA1* and *GLABRA3* to the promoter of *GLABRA2*. *Plant Cell Physiol*, 49, 1792-1804.
- Wang S., Kwak S. H., Zeng Q., Ellis B. E., Chen X. Y., Schiefelbein J. & Chen J. G. (2007). *TRICHOMELESS1* regulates trichome patterning by suppressing *GLABRA1* in *Arabidopsis*. *Development*, 134, 3873-3882.
- Wang S., Hubbard L., Chang Y., Guo J., Schiefelbein J. & Chen J. G. (2008). Comprehensive analysis of single-repeat R3 MYB proteins in epidermal cell patterning and their transcriptional regulation in *Arabidopsis*. *BMC Plant Biol*, 8, 81.
- Wei S., Li X., Gruber M. Y., Feyissa B. A., Amyot L. & Hannoufa A. (2018). *COP9 signalosome subunit 5A* affects phenylpropanoid metabolism, trichome formation and transcription of key genes of a regulatory tri-protein complex in *Arabidopsis*. *BMC Plant Biol*, 18, 134.
- Weigel D. (2012). Natural variation in *Arabidopsis*: from molecular genetics to ecological genomics. *Plant Physiol*, 158, 2-22.
- Weigel D. & Nordborg M. (2015). Population genomics for understanding adaptation in wild plant species. *Annu Rev Genet*, 49, 315-338.
- Wellesen K., Durst F., Pinot F., Benveniste I., Nettesheim K., Wisman E., Steiner-Lange S., Saedler H. & Yephremov A. (2001). Functional analysis of the *LACERATA* gene of *Arabidopsis* provides evidence for different roles of fatty acid omega -hydroxylation in development. *Proc Natl Acad Sci USA*, 98, 9694-9699.
- Wester K., Digiuni S., Geier F., Timmer J., Fleck C. & Hülskamp M. (2009). Functional diversity of R3 single-repeat genes in trichome development. *Development*, 136, 1487-1496.
- Willmann M. R. & Poethig R. S. (2011). The effect of the floral repressor *FLC* on the timing and progression of vegetative phase change in *Arabidopsis*. *Development*, 138, 677-685.
- Wu R., Li S., He S., Wassmann F., Yu C., Qin G., Schreiber L., Qu L. J. & Gu H. (2011). CFL1, a WW domain protein, regulates cuticle development by modulating the function of HDG1, a class IV homeodomain transcription factor, in rice and *Arabidopsis*. *Plant Cell*, 23, 3392-3411.
- Xue S., Dong M., Liu X., Xu S., Pang J., Zhang W., Weng, Y & Ren, H. (2019). Classification of fruit trichomes in cucumber and effects of plant hormones on type II fruit trichome development. *Planta*, 249, 407-416.
- Xue X. Y., Zhao B., Chao L. M., Chen D. Y., Cui W. R., Mao Y. B., Wang L. J. & Chen X. Y. (2014). Interaction between two timing microRNAs controls trichome distribution in *Arabidopsis*. *PLoS Genet*, 10, e1004266.
- Yang X., Zhang W., He H., Nie J., Bie B., Zhao J., Ren G. Li Y., Zhang D., Pan J. & Cai R. (2014). Tuberculate fruit gene *Tu* encodes a C2H2 zinc finger protein that is required for the warty fruit phenotype in cucumber (*Cucumis sativus* L.). *Plant J*, 78, 1034-1046.
- Yephremov A., Wisman E., Huijser P., Huijser C., Wellesen K. & Saedler, H. (1999). Characterization of the *FIDDLEHEAD* gene of *Arabidopsis* reveals a link between adhesion response and cell differentiation in the epidermis. *Plant Cell*, 11, 2187-2201.
- Yu N., Cai W. J., Wang S., Shan C. M., Wang L. J. & Chen X. Y. (2010). Temporal control of trichome distribution by microRNA156-targeted *SPL* genes in *Arabidopsis thaliana*. *Plant Cell*, 22, 2322-2335.
- Zagotta M. T., Hicks K. A., Jacobs C. I., Young J. C., Hangarter R. P. & Meeks-Wagner D. R. (1996). The *Arabidopsis* *ELF3* gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. *Plant J*, 10(4), 691-702.

- Zhang F., González A., Zhao M., Payne C. T. & Lloyd A. (2003). A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. *Development*, 130, 4859-4869.
- Zhang N., Yang L., Luo S., Wang X., Wang W., Cheng Y., Cai L. & Wang S. (2018). Genetic evidence suggests that *GIS* functions downstream of *TCL1* to regulate trichome formation in Arabidopsis. *BMC Plant Biol*, 18, 63.
- Zhang Z. B., Zhu J., Gao J. F., Wang C., Li H., Li H., Zhang, H. Q., Zhang S., Wang D. M., Wang Q. X., Huang H., Xia H. J. & Yang Z. N. (2007). Transcription factor *AtMYB103* is required for anther development by regulating tapetum development, callose dissolution and exine formation in Arabidopsis. *Plant J*, 52, 528-538.
- Zhao H., Wang X., Zhu D., Cui S., Li X., Cao Y. & Ma L. (2012). A single amino acid substitution in IIIf subfamily of basic helix-loop-helix transcription factor *AtMYC1* leads to trichome and root hair patterning defects by abolishing its interaction with partner proteins in Arabidopsis. *J Biol Chem*, 287, 14109-14121.
- Zhong R., Burk D. H., Nairn C. J., Wood-Jones A., Morrison W. H. & Ye Z. H. (2005). Mutation of *SAC1*, an Arabidopsis SAC domain phosphoinositide phosphatase, causes alterations in cell morphogenesis, cell wall synthesis, and actin organization. *Plant Cell*, 17, 1449-1466.
- Zhou Z., An L., Sun L., Zhu S., Xi W., Broun P., Yu H. & Gan Y. (2011). *ZINC FINGER PROTEIN5* is required for the control of trichome initiation by acting upstream of *ZINC FINGER PROTEIN8* in Arabidopsis. *Plant Physiol*, 157, 673-682.
- Zhou Z., Sun L., Zhao Y., An L., Yan A., Meng X. & Gan Y. (2013). *ZINC FINGER PROTEIN 6 (ZFP6)* regulates trichome initiation by integrating gibberellin and cytokinin signaling in *Arabidopsis thaliana*. *New Phytol*, 198, 699-708.
- Züst T., Heichinger C., Grossniklaus U., Harrington R., Kliebenstein D. J. & Turnbull L. A. (2012). Natural enemies drive geographic variation in plant defenses. *Science*, 338, 116-119.

# ANEXO

Nombre	Cromosoma	Función molecular	Función en el desarrollo de los tricomas	Fenotipo mutante	Referencia
<b>Genes implicados en el patrón de tricomas</b>					
<i>ETC1</i>	1	FT R3 MYB	Regulador negativo	↑ DTH en triples mutantes, ↓ pelos radicales	Kirik <i>et al.</i> , 2004a
<i>CSN5A</i>	1	Complejo COP9	Relacionado con el complejo MYB/bHLH/TTG1	↓ DT y anomalías en la ramificación	Wei <i>et al.</i> , 2018
<i>EGL3</i>	1	FT bHLH	Regulador positivo	↓ DT ↑ pelos radicales	Zhang <i>et al.</i> , 2003
<i>GL2</i>	1	FT homeobox	Regulador positivo	↓ DTH, ↑ pelos radicales	Rerie <i>et al.</i> , 1994
<i>SHP2</i>	2	Gen homeótico de floración	Inhibición de formación de tricomas en carpelo	Tricomas en el carpelo	Ó'Maoláidigh <i>et al.</i> , 2018
<i>FDH</i>	2	β-ketocil sintasa	Iniciación de tricomas	↓ número de tricomas	Yepremov <i>et al.</i> , 1999
<i>MTL8</i>	2	FT NAC	Regulador negativo	Sin defectos en la formación de tricomas	Tian <i>et al.</i> , 2017
<i>ETC2</i>	2	FT MYB R3	Regulador negativo	↑ DTH	Kirik <i>et al.</i> , 2004b
<i>TCL2</i>	2	FT MYB R3	Regulador negativo	↑ DT en tallo con inflorescencias y pedicelo	Can <i>et al.</i> , 2011
<i>TCL1</i>	2	FT MYB R3	Regulador negativo	↑ DT en tallo con inflorescencias y pedicelo	Wang <i>et al.</i> , 2007
<i>CPC</i>	2	FT MYB R3	Regulador negativo	Tricomas agrupados y sin pelos radicales	Wada <i>et al.</i> , 1997
<i>MYB5</i>	3	FT MYB R2R3	Regulador positivo	↓ DTH	Li <i>et al.</i> , 2009
<i>GL1</i>	3	FT MYB R2R3	Regulador positivo	Glabro	Koormneef <i>et al.</i> , 1982
<i>TRP</i>	3	FT WRKY C2H2 Dedo de Zinc	Regulador negativo	↑ número tricomas	Kim <i>et al.</i> , 2018
<i>SHP1</i>	3	Gen homeótico de floración	Inhibición de formación de tricomas en carpelo	Tricomas en el carpelo	Ó'Maoláidigh <i>et al.</i> , 2018
<i>AG</i>	4	Gen homeótico de floración	Inhibición de formación de tricomas en carpelo	Tricomas en el carpelo	Ó'Maoláidigh <i>et al.</i> , 2018
<i>MYC1</i>	4	FT bHLH	Regulador positivo	↓ DTH	Zhao <i>et al.</i> , 2012
<i>ETC3</i>	4	FT MYB R3	Regulador negativo	↓ DTH, ↑ ramificación	Wester <i>et al.</i> , 2009
<i>TT8</i>	4	FT bHLH	Desarrollo de tricomas marginales	↓ DTH	Maes <i>et al.</i> , 2008
<i>KAN1</i>	5	Gen homeótico de floración	Formación de tricomas en hoja	Producción prematura de tricomas en la cara abaxial de hojas en rosela	Kerstetter <i>et al.</i> , 2001
<i>TTG1</i>	5	Proteína WD40	Regulador positivo	Glabros, ↑ pelos radicales, ↓ antocianinas	Walker <i>et al.</i> , 1999
<i>MYB23</i>	5	FT MYB R2R3	Regulador positivo	↓ DTH	Kirik <i>et al.</i> , 2001b
<i>GL3</i>	5	FT bHLH	Regulador positivo	↓ DTH, ↓ ramificación, ↑ pelos radicales	Payne <i>et al.</i> , 2000
<i>MYB82</i>	5	FT MYB R2R3	Regulador positivo	La sobreexpresión provoca desarrollo anormal del tricoma	Liang <i>et al.</i> , 2014
<i>TRY</i>	5	FT MYB R3	Regulador negativo	Tricomas agrupados, ↑ ramificación, ↓ pelos radicales	Schellmann <i>et al.</i> , 2002
<b>Genes implicados en la endorreducción de los tricomas</b>					
<i>CYCA2-3</i>	1	Ciclina de tipo A2	Regulador de la plididia durante la endorreducción	↓ endorreducción	Imai <i>et al.</i> , 2006
<i>ICK1</i>	2	Inhibidor de quinasas dependiente de ciclina	Reduce la endorreducción	-	Schnittger <i>et al.</i> , 2003
<i>EMB1611</i>	2	Regulación de endorreducción	Incremento de endorreducción	Tricomas hiperramificados	Leasure <i>et al.</i> , 2009
<i>PYM</i>	2	Inhibidor del complejo APC/C	Regulación de la endorreducción	Tricomas hiperramificados	Hase <i>et al.</i> , 2006
<i>SPY</i>	3	N-acetil glucosamintransferasa	Señalización de GA	↑ DTH ↑ plididia, ↓ ramificación	Perazza <i>et al.</i> , 1998
<i>ELC</i>	3	Parte del complejo ESQRTI-III	Regulación de la endorreducción	Varios tricomas en un único pie y multinucleados	Spitzer <i>et al.</i> , 2006
<i>HYP6</i>	3	ADN topoisomerasa VI	Progresión a la endorreducción	↓ plididia, ↓ ramificación, ↓ pelos radicales	Sugimoto-Shirasu <i>et al.</i> , 2002
<i>KAK</i>	4	Ubiquitin ligasa	Represión de la endorreducción	↑ plididia, ↓ ramificación	Downes <i>et al.</i> , 2003
<i>RHL2</i>	5	DNA topoisomerasa VI	Progresión a la endorreducción	↓ plididia, ↓ pelos radicales	Sugimoto-Shirasu <i>et al.</i> , 2002
<i>SIM</i>	5	Inhibidor de quinasas de ciclo celular	Transición mitosis a endorreducción	Tricomas multicelulares	Walker <i>et al.</i> , 2000
<i>CPR5</i>	5	Proteína de membrana para la patogénesis	Regulador de la muerte celular	↓ plididia, ↓ ramificación	Kirik <i>et al.</i> , 2001a

**Tabla Suplementaria 1. Genes implicados en el desarrollo de los tricomas.** Genes identificados mediante análisis de mutantes que participan en el desarrollo de los tricomas, clasificados según el proceso de desarrollo en el que intervienen. La tabla muestra: el nombre del gen; el cromosoma donde se localiza; la función molecular y en el desarrollo de los tricomas; el fenotipo del mutante o de la sobreexpresión; la referencia. FT: Factor de transcripción. DT: Densidad de tricomas. ↓ Reducción. ↑ Incremento.

Nombre Genes implicados en la ramificación de los tricomas	Cromosoma	Función molecular	Función en el desarrollo de los tricomas	Fenotipo mutante	Referencia
AN	1	Nucleación de actina	Incremento en el número de microtúbulos	↓ ramificación	Folkers <i>et al.</i> , 2002
BLT	1	Función desconocida	Iniciación de la ramificación y regulador del ciclo celular	No ramificado	Marks <i>et al.</i> , 2009
TUA4	1	α-Tubulina	Organización de microtúbulos	↓ ramificación	Abe <i>et al.</i> , 2004
HGG12	1	FT HD-ZIP IV	Ramificación	Tricomas hiperramificados	Nakamura <i>et al.</i> , 2006
SCD1	1	Citocinesis y expansión celular	Morfogénesis	Tricomas no ramificados	Falbel <i>et al.</i> , 2003
HGG11	1	FT HD-ZIP IV	Ramificación	Tricomas hiperramificados	Nakamura <i>et al.</i> , 2006
FRA2	1	Katanina	Morfogénesis	↓ ramificación	Burk <i>et al.</i> , 2001
STI	2	Nucleación de actina	Estabilización de microtúbulos	Tricomas no ramificados	Igenfritz <i>et al.</i> , 2003
KIS	2	Complejo de ensamblaje de la tubulina	Factor de regulación de microtúbulos	Tricomas no ramificados	Kirik <i>et al.</i> , 2002a
TTG2	2	FT WQRKY	Desarrollo temprano de tricomas	Poca o nula ramificación	Johnson <i>et al.</i> , 2002
KIC	2	Proteína de la familia EF-hand	Ramificación	↓ Menor ramificación	Reddy <i>et al.</i> , 2004
MYB106	3	FT MYB	Regulador negativo	Tricomas hiperramificados	Jakoby <i>et al.</i> , 2008
KINESIN-13A	3	Unión a ATP	Morfogénesis	Tricomas hiperramificados	Lu <i>et al.</i> , 2005
TUA6	4	α-Tubulina	Organización de microtúbulos	↓ ramificación	Abe <i>et al.</i> , 2004
POR/TFCC	4	Proteína de unión a tubulina	Morfogénesis	Tricomas no ramificados	Kirik <i>et al.</i> , 2002b
ADL1	5	Proteína de unión a GTP	Morfogénesis	Defectos ramificación	Kang <i>et al.</i> , 2003
MYB103	5	FT MYB R2R3	Regulador de endorreducción	Tricomas hiperramificados	Zhang <i>et al.</i> , 2007
FAS2	5	Complejo CAF-1	Control de ramificación	Tricomas hiperramificados	Exner <i>et al.</i> , 2006
ZWI	5	Proteína de unión a calmodulina	Transporte a lo largo de los microtúbulos	↓ ramificación	Oppenheimer <i>et al.</i> , 1997
<b>Genes implicados en la expansión y direccionalidad de los tricomas</b>					
SPI	1	Proteína WD/BEACH	Morfogénesis	Ramificaciones más cortas	Saedler <i>et al.</i> , 2009
HGG2	1	FT HD-ZIP IV	Desarrollo de la pared celular	Tricomas con aspecto de cristal	Marks <i>et al.</i> , 2009
DIS1	1	Complejo proteico Arp2/3	Morfogénesis	Tricomas distorsionados	Mathur <i>et al.</i> , 2003a
ROP2	1	Rho GTPasa	Organización de actina	Tricomas distorsionados	Nagawa <i>et al.</i> , 2012
SAC1	1	Fosfatasa	Morfogénesis	Tricomas más pequeños	Zhong <i>et al.</i> , 2005
DIS2	1	Actividad ARPC2	Maduración	Tricomas distorsionados	El-Din El-Assal <i>et al.</i> , 2004
NAP	1	FT del complejo WAVE	Organización de actina	Tricomas defectuosos	Li <i>et al.</i> , 2004
LCR	2	Miembro de fitocromo p450	Diferenciación celular	Desarrollo más lento de los tricomas	Welleßen <i>et al.</i> , 2001
MUR2	2	Glucosil transferasa	Desarrollo de la pared celular	Tricomas con papilas arrugadas y colapsadas	Vanzin <i>et al.</i> , 2002
BRK1	2	Complejo proteico SCAR	Nucleación de actina	Tricomas distorsionados	Djakovic <i>et al.</i> , 2006
GRL	2	Complejo proteico SCAR	Nucleación de actina	Tricomas distorsionados	Hülkamp <i>et al.</i> , 1994
DIS3	2	Homólogo Scar/WAVE	Morfogénesis	Tricomas distorsionados	Basu <i>et al.</i> , 2005
WRM	3	Complejo proteico Arp2/3	Morfogénesis	Tricomas distorsionados	Djakovic <i>et al.</i> , 2006
HGG1	3	FT HD-ZIP IV	Correcta formación de la cutícula	La sobreexpresión provoca tricomas deformes y aplanados	Wu <i>et al.</i> , 2011
PFL2	4	Proteína ribosomal citoplasmática S13	Desarrollo de tricomas	Tricomas aberrantes	Ito <i>et al.</i> , 2000
CRK	4	Complejo APR2/3	Expansión celular	Tricomas distorsionados	Mathur <i>et al.</i> , 2003b
SPK1	4	Factor de cambio de guaninas	Organización de microtúbulos	Tricomas no ramificados	Qiu <i>et al.</i> , 2002
SHV3	4	Desarrollo pared celular	Morfogénesis	Tricomas colapsados	Hayashi <i>et al.</i> , 2008
KLK	5	Complejo proteico SCAR	Organización de actina	Tricomas distorsionados	Saedler <i>et al.</i> , 2004
XIK	5	Miosina de tipo XI	Morfogénesis	Tamaño y forma irregular de los tricomas	Qjangu <i>et al.</i> , 2007
SVL1	5	Desarrollo pared celular	Morfogénesis	Tricomas colapsados	Hayashi <i>et al.</i> , 2008
YRE	5	Proteína transmembrana	Morfogénesis	Tricomas más pequeños	Kurata <i>et al.</i> , 2003

**Tabla Suplementaria 1. Genes implicados en el desarrollo de los tricomas. (Continuación)**

Nombre	Cromosoma	Función molecular	Función en el desarrollo de los tricomas	Fenotipo mutante	Referencia
<b>Genes implicados en el patrón de tricomas a través de otros factores</b>					
<i>SPL8</i>	1	FT SBP	Regulador negativo	↓ DT en sépalos	Unite <i>et al.</i> , 2003
<i>MIR166</i>	1	miARN con diana en proteínas SPLs	Regula la actividad de las proteínas SPL	La sobreexpresión provoca tricomas ectópicos	Yu <i>et al.</i> , 2010
<i>ZFP5</i>	1	FT C2H2	Desarrollo de tricomas dependiente de GAs/CKs	Iniciación de tricomas en inflorescencia	Zhou <i>et al.</i> , 2011
<i>MIR171B</i>	1	miARN con diana en proteínas LOMs	Regulan la actividad de las proteínas LOM	La sobreexpresión provoca ↓ DT en tallo	Xue <i>et al.</i> , 2014
<i>GAI</i>	1	Señalización GAs	Reprime formación tricomas y su ramificación	-	Gan <i>et al.</i> , 2007
<i>TEM1</i>	1	FT RAV	Regulador negativo mediante el control de acumulación de GAs	↑DT en hojas y en inflorescencia <sup>1</sup>	Maitas-Hernández <i>et al.</i> , 2016
<i>SPL10</i>	1	FT SBP	Regulador negativo	↑ DT en inflorescencia	Shikata <i>et al.</i> , 2009
<i>LEC2</i>	1	FT con dominio B3	Desarrollo de tricomas dependiente de GAs	Tricomas ectópicos	Curaba <i>et al.</i> , 2004
<i>MIR171C</i>	1	miARN con diana en proteínas LOMs	Regulan la actividad de las proteínas LOM	La sobreexpresión provoca ↓ DT en tallo	Xue <i>et al.</i> , 2014
<i>FT</i>	1	Unión a fosfatidicolina	Formación de tricomas en cara abaxial de la hoja	↓ DT en la cara abaxial de la hoja	Willmann & Poething, 2011
<i>RGL1</i>	1	Señalización GAs	Iniciación en flores	↓ DT en fondo $\varnothing$ 1-3	Gan <i>et al.</i> , 2007
<i>ZFP6</i>	1	FT C2H2	Desarrollo de tricomas dependiente de GAs/CKs	↓ DT en sépalos, hojas caulinares y tallo	Zhou <i>et al.</i> , 2013
<i>GIS3</i>	1	FT C2H2	Desarrollo de tricomas dependiente de GAs/CKs	↓ DT en sépalos, hojas caulinares y tallo	Sun <i>et al.</i> , 2015
<i>TEM2</i>	1	FT RAV	Regulador negativo mediante el control de acumulación de GAs	↑DT en hojas de la roseta	Maitas-Hernández <i>et al.</i> , 2016
<i>JAZ9</i>	1	Proteína con dominio Ja-Zim	Señalización AJ	↑ DTH	Thines <i>et al.</i> , 2007
<i>JAZ2</i>	1	Proteína con dominio Ja-Zim	Señalización AJ	↑ DTH	Thines <i>et al.</i> , 2007
<i>RGA</i>	2	FT DELLA	Señalización de GAs	↑ DTH	Dill & Sun, 2001
<i>SAD2</i>	2	Señalización ABA	Iniciación de tricomas	↓ número de tricomas	Gao <i>et al.</i> , 2008
<i>CO1</i>	2	Proteína F-box	Señalización del AJ	↓ DT	Qi <i>et al.</i> , 2011
<i>ZFP8</i>	2	FT C2H2	Desarrollo de tricomas dependiente de GAs	Iniciación de tricomas en inflorescencia	Zhou <i>et al.</i> , 2013
<i>SPL9</i>	2	FT SBP	Regulador negativo	↑ DT en inflorescencia	Yu <i>et al.</i> , 2010
<i>LOM1</i>	2	FT GRAS	Regulan proteínas SPL	↓ DT en tallo y órganos florales	Xue <i>et al.</i> , 2014
<i>SOC1</i>	2	FT MADS-box tipo II	Formación de tricomas en cara abaxial de la hoja	↓ DT en la cara abaxial de la hoja	Willmann & Poething, 2011
<i>FUS3</i>	3	Biosíntesis de GAs	Desarrollo de tricomas dependiente de GAs	Tricomas ectópicos	Curaba <i>et al.</i> , 2004
<i>MIR171A</i>	3	microARN con diana en proteínas LOMs	Regulan la actividad de las proteínas LOM	La sobreexpresión provoca ↓ DT en tallo	Xue <i>et al.</i> , 2014
<i>GIS</i>	3	FT C2H2	Desarrollo de tricomas dependiente de GAs	Iniciación de tricomas en inflorescencia	Gan <i>et al.</i> , 2006
<i>LOM2</i>	3	FT GRAS	Regulan la actividad de las proteínas SPL	↓ DT en tallo y órganos florales	Xue <i>et al.</i> , 2014
<i>LOM3</i>	4	FT GRAS	Regulan proteínas SPL	↓ DT en tallo y órganos florales	Xue <i>et al.</i> , 2014
<i>GAI</i>	4	FT DELLA	Señalización de GAs	↑ DTH	Dill & Sun, 2001
<i>SPL2</i>	5	FT SBP	Regulador negativo	↑ DT en inflorescencia	Shikata <i>et al.</i> , 2009
<i>GIS2</i>	5	FT C2H2	Desarrollo de tricomas dependiente de GAs/CKs	Iniciación de tricomas en inflorescencia	Gan <i>et al.</i> , 2007
<i>FLC</i>	5	FT MADS-box	Formación de tricomas en cara abaxial de la hoja	↑ DT en la cara abaxial de la hoja	Willmann & Poething, 2011
<i>FPF1</i>	5	Señalización GAs	Afecta al patrón de distribución	-	Melzer <i>et al.</i> , 1999
<b>Genes implicados en el desarrollo de tricomas pero sin función determinada</b>					
<i>DEK1</i>	1	-	-	La sobreexpresión produce plantas glabras	Lid <i>et al.</i> , 2005
<i>SVB</i>	1	-	-	Tricomas mas pequeños con ramificaciones variables	Marks <i>et al.</i> , 2009
<i>ATP56</i>	1	Trealosa sintasa	-	↓ ramificación	Chary <i>et al.</i> , 2008
<i>HST</i>	3	Transición fase juvenil a adulta	-	Tricomas en la cara abaxial	Teller <i>et al.</i> , 1998
<i>PEL3</i>	5	Aciltransferasa	-	Tricomas entrecados	Marks <i>et al.</i> , 2009

Tabla Suplementaria 1. Genes implicados en el desarrollo de los tricomas. (Continuación). <sup>1</sup> Fenotipo sólo en mutante múltiple.



Carácter	Análisis	Locus altamente asociado	Cromosoma	Posición SNP	$-\log(P)$ MLM	# de SNPs asociados	Frecuencia alelo menor	Gen próximo	Nombre proteína	Función de la proteína	Impacto funcional
DTH	10 loci más asociados	8	1	23055167	5,5	1	0,03	At1g62340	ALE1	Formación de la superficie de la epidermis	1
		1	2	1729824	7,2	3	0,03-0,05	At2g04925		Proteína similar a defensinas	1,2,3
		1	2	1760013	5,9	1	0,03	At2g04990		Elemento transposable	3
		3	2	<b>3707395</b>	6,9	1	0,04	At2g09840		Proteína de unión de iones de zinc a ácidos nucleicos	3
		2	2	12862494	5,3	3	0,03-0,05	At2g30120	FLX	Proteína de floración	3
		2	2	<b>12960512</b>	6,9	4	0,34	At2g30410	KIS	Cofactor de plegamiento de tubulina	1,3
		2	2	<b>12961968</b>	7,0	9	0,27-0,40	At2g30420	ETC2	Regulador negativo del desarrollo de tricomas	1,2,3
		2	2	<b>12966654</b>	6,1	11	0,16-0,31	At2g30424	TCL2	Regulador negativo del desarrollo de tricomas	3
		2	2	<b>13008617</b>	6,8	3	0,30-0,46	At2g30520	RPT2	Respuesta de fototropismo de la raíz	3
		2	2	<b>13012551</b>	6,4	10	0,07-0,13	At2g30540		Glutarredoxina	2,3
		9	3	<b>3111527</b>	5,4	1	0,04	At3g10090			3
		7	3	<b>5952121</b>	5,8	2	0,03	At3g17380	MAT4	Proteína ribosomal 40S	3
		7	3	5952476	5,8	4	0,03	At3g17390		Proteína TRAF, sin función conocida	2,3
		7	3	5957124	5,8	2	0,03-0,04	At3g17410		S-adenosilmetionina sintasa	2,3
		7	3	5959634	5,8	2	0,03	At3g17420	GPK1	Kinasa	3
		10	3	8279766	5,3	1	0,03	At3g23190		Receptor quinasa	3
		4	4	<b>2341504</b>	6,6	2	0,04-0,05	At4g04632		Inducción de la respuesta hipersensible	3
		4	4	2379441	5,3	5	0,07-0,17	At4g04693		Kinasa	3
		6	4	<b>2549629</b>	5,8	17	0,06-0,35	At4g04985		Pseudogen	3
		5	4	17643412	5,8	1	0,03	At4g37540	LBD39	Proteasa	0,1,2,3
										Regulador negativo de la biosíntesis de antocianinas	3
	Genes tricomas										
			2	12960512	6,9	4	0,27-0,34	At2g30410	KIS	Cofactor de plegamiento de tubulina	1,3
			2	12961968	7,0	9	0,33-0,40	At2g30420	ETC2	Regulador negativo del desarrollo de tricomas	1,3
			2	12966654	6,0	11	0,16-0,31	At2g30424	TCL2	Regulador negativo del desarrollo de tricomas	3
			2	12970819	4,4	2	0,20-0,23	At2g30432	TCL1	Regulador negativo del desarrollo de tricomas	3
			2	17500826	4,5	2	0,06-0,08	At2g41940	ZFP8	Regulador de la formación de tricomas mediante GAs	3
			4	42176	4,1	1	0,08	At4g00100	PFL2	Morfología de los tricomas	3
			4	68463	4,0	1	0,05	At4g00150	LOM3	Mantenimiento de meristemas	3
			4	228354	4,6	1	0,07	At4g00480	MYC1	Regulador positivo del desarrollo de tricomas	1
	Otros genes asociados relevantes		1	24797237	5,0	1	0,08	At1g66470	RDH6	Formación de pelos radiculares	3

**Tabla Suplementaria 2: Resumen de los análisis de asociación genómica.** Para cada variable fenotípica (DTH, DTHAJ, PDTH, PTT y PTP) se muestran los 10 loci más asociados, los genes candidatos del desarrollo de tricomas asociados y los genes que indirectamente podrían alterar el patrón de tricomas. Para los loci que incluyen varias regiones genómicas estrechamente ligadas y asociadas al patrón de tricomas se muestran de 2-8 regiones altamente significativas. Los genes candidatos de tricomas corresponden a los 115 genes conocidos implicados en el patrón y la morfología de los tricomas ( $\pm 10$  kb) que contienen SNPs significativos. Para cada región genómica se muestran los siguientes parámetros: cromosoma, posición física, valor de  $-\log(P)$ , número de SNPs asociados en el mismo gen, frecuencia del alelo menor, nombre del gen más próximo, nombre y función de la proteína correspondiente y el impacto funcional predicho (0: mutaciones de codón de parate prematuro; 1: mutaciones de cambio de aminoácido; 2: mutaciones sinónimas; 3: mutaciones localizadas en regiones no codificantes). La posición cromosómica de las regiones genómicas detectadas en común para DTH y DTHAJ, PDTH y DTH/DTHAJ, o para PTT y PTP se indican con texto en negrita e itálica.

Carácter	Análisis	Locus altamente asociado	Cromosoma	Posición SNP	-log(P) MLM	# de SNPs asociados	Frecuencia alelo menor	Gen próximo	Nombre proteína	Función de la proteína	Impacto funcional		
DTHAJ	10 loci mas asociados	9	1	24339333	4,9	1	0,06	At1g65480	FT	Promotor de la floración	3		
		10	2	1019823	4,8	1	0,20	At2g03350		Proteína de función desconocida	2		
		8	2	7289608	4,9	1	0,09	At2g16810		Proteína F-box, función desconocida	3		
		1	2	12950206	5,2	2	0,28-0,29	At2g30380		Proteína de función desconocida	1,3		
		1	2	12953572	5,9	1	0,30	At2g30390		Ferredoxinase	2		
		1	2	12960512	7,2	4	0,27-0,34	At2g30410		KIS	1,3		
		1	2	12961968	7,0	9	0,33-0,40	At2g30420		ETC2	Regulador negativo del desarrollo de tricomas	1,2,3	
		1	2	12966654	6,1	11	0,16-0,31	At2g30424		TCL2	Regulador negativo del desarrollo de tricomas	3	
		1	2	12979218	5,2	11	0,18-0,22	At2g30470		VAL1	Represor de la maduración de semillas	3	
		1	2	13008617	5,0	2	0,04-0,15	At2g30520		RPT2	Respuesta de fototropismo de la raíz	3	
		1	2	13012551	5,0	4	0,07-0,13	At2g30540			Glutarredoxina	3	
		5	3	10725785	5,3	10	0,30-0,41	At3g28610			Hidrolasa	1,2	
		2	4	228482	5,1	2	0,07-0,45	At4g00520			AcilCoA tioesterasa	1,3	
		2	4	321193	5,7	3	0,04-0,10	At4g00752			Proteína con dominio UBX, función desconocida	3	
		7	4	7354699	4,8	1	0,05	At4g12420		SKU5	Crecimiento de la raíz	3	
		7	4	8051299	5,1	1	0,03	At4g13930		SHM4	Serina hidroximetiltransferasa	3	
		7	4	9518299	4,9	1	0,06	At4g16915			Proteína desconocida	3	
		4	5	11072444	5,5	1	0,03	At5g29030			Pseudogen	3	
		3	5	13797580	5,7	3	0,04-0,10	At5g35606			Elemento transponible	3	
		6	5	15632955	5,3	2	0,03-0,07	At5g39050		PMAT1	Detoxicación fenólica	3	
		Genes tricomas		1	1	24339333	4,9	1	0,06	At1g65480	FT	Promotor de la floración	3
				2	2	12960286	7,2	4	0,27-0,34	At2g30410	KIS	Cofactor de plegamiento de tubulina	1,3
				2	2	12960512	7,2	9	0,33-0,40	At2g30420	ETC2	Regulador negativo del desarrollo de tricomas	1,2,3
2	2			12966733	6,1	10	0,16-0,31	At2g30424	TCL2	Regulador negativo del desarrollo de tricomas	3		
2	2			12967408	5,8	9	0,18-0,30	At2g30432	TCL1	Regulador negativo del desarrollo de tricomas	3		
Otros genes asociados relevantes		2	2	228482	5,1	2	0,07-0,45	At4g00480	MYC1	Regulador positivo del desarrollo de tricomas	1,3		
		2	2	12944467	4,2	1	0,22	At2g30370	EPFL6	Inhibidor de la formación de estomas	3		
		4	4	722293	4,1	1	0,22	At4g01680	MYB55	Implicada en la morfogénesis de la hoja	3		

Tabla Suplementaria 2: Resumen de los análisis de asociación genómica. (Continuación)

Carácter	Análisis	Locus altamente asociado	Cromosoma	Posición SNP	-log(P) MLM	# de SNPs asociados	Frecuencia alelo menor	Gen próximo	Nombre proteína	Función de la proteína	Impacto funcional		
PDTH	10 loci mas asociados	5	1	11117949	4,4	1	0,04	At1g31140	GOA	Represión del crecimiento del fruto	3		
		5	1	12063708	4,7	1	0,41	At1g33250		Función desconocida	3		
		9	2	<b>16762209</b>	4,4	1	0,07	At2g40130		Hidrolasa	3		
		8	3	10328946	4,4	1	0,03	At3g27870		ATPasa	3		
		3	3	20358557	4,8	1	0,03	At3g54950	PLP7	Fosfolipasa	3		
		7	4	<b>2550880</b>	4,4	2	0,17-0,18	At4g04985		Proteasa	3		
		4	4	<b>7354572</b>	4,8	1	0,10	At4g12420	SKU5	Crecimiento de la raíz	3		
		4	4	8225183	4,4	1	0,04	At4g14280		Proteína ARM, función desconocida	3		
		6	5	631001	4,6	1	0,05	At5g02780	GSTL1	Glutación S-transferasa	3		
		10	5	4578259	4,4	1	0,34	At5g14200	IMD1	3-isopropilmalato deshidrogenasa	3		
		1	5	17813324	5,0	1	0,09	At5g44220		Proteína F-box, función desconocida	3		
		1	5	17868215	5,0	1	0,07	At5g44350		Proteína nuclear sensible al etileno	3		
		1	5	17890187	4,4	1	0,09	At5g44400		Proteína de unión a FAD	3		
		1	5	17929555	5,1	1	0,07	At5g44500		Ribonucleoproteína nuclear	3		
		1	5	17996609	4,5	3	0,07-0,11	At5g44620		Citocromo P450	3		
		2	5	26046181	5,0	2	0,18-0,21	At5g65180		Proteína ENTH/VHS, función desconocida	2,3		
		Genes tricomas											
		Otros genes asociados relevantes											
		No hay candidatos de genes de tricomas											
		El análisis no muestra otros genes asociados relevantes											

**Tabla Suplementaria 2: Resumen de los análisis de asociación genómica. (Continuación)**

Carácter	Análisis	Locus altamente asociado	Cromosoma	Posición SNP	-log(P) MLM	# de SNPs asociados	Frecuencia alelo menor	Gen próximo	Nombre proteína	Función de la proteína	Impacto funcional		
PTT	10 loci mas asociados	8	1	12422554	5,9	38	0,03-0,04	AT1g34120	IP5P1	Inositol polifosfatasa	1,2,3		
		4	1	25406936	6,5	2	0,03	AT1g67770	TEL2	Implicado en meiosis	3		
		6	2	11316035	5,9	3	0,4-0,5	At2g26590	RPN13	Subunidad reguladora del proteosoma 26S	3		
		6	2	<b>11319489</b>	6,2	3	0,05-0,06	At2g26610		Proteína WD-40, función desconocida	1,2		
		1	2	12937759	6,0	7	0,04	At2g30360	CIPK11	Serín/treonín proteína quinasa	1,2		
		1	2	13463562	6,1	3	0,03-0,04	At2g31660	SAD2	Desarrollo de tricomas mediante ABA	3		
		1	2	13529930	6,1	4	0,03-0,04	At2g31810	AHASS1	Acetolactato sintasa	3		
		1	2	13539422	6,1	4	0,03	At2g31840		Tiorredoxina	1,3		
		1	2	13546847	6,1	2	0,03	At2g31865	PARAG2	Glicohidrolasa poli(ADP-ribosa)	2,3		
		1	2	13549890	6,2	2	0,04	At2g31870	PARAG1	Glicohidrolasa poli(ADP-ribosa)	3		
		1	2	13554936	6,1	1	0,03	At2g31880	SOBIR1	Receptor serina/treonina/tirosina protein quinasa	1		
		1	2	13557411	6,1	3	0,03	At2g31890	RAP	Regulación de la respuesta de defensa	1		
		1	2	13565853	6,1	12	0,03-0,08	At2g31900	XI-F	Miosina	1,2,3		
		1	2	13822224	7,4	7	0,03-0,04	At2g32550		Proteína de diferenciación celular	1,2,3		
		10	3	<b>2364447</b>	5,7	1	0,04	At2g32550		Función desconocida	3		
		5	4	5167342	6,5	6	0,07-0,13	At4g08170	AIR12	Función desconocida	3		
		9	4	<b>13956962</b>	5,8	7	0,08-0,14	At4g28080	ITPK3	Inositol-trifosfato quinasa	3		
		7	4	14663861	5,9	1	0,03	At4g29970	REC2	Establecimiento del tamaño de compartimento cloroplástico	1,2,3		
		7	4	14671405	6,0	1	0,03	At4g30000		Proteína F-box, función desconocida	2		
		2	4	<b>17183772</b>	7,3	4	0,05-0,09	At4g36360	β-GAL3	Kinasa	1		
		3	5	19275201	6,1	5	0,08-0,10	At5g47510		β-galactosidasa	2,3		
		3	5	19279697	6,8	21	0,07-0,10	At5g47520	RABA5A	Fofatidilinositol/fosfatidilcolina transferasa	1,3		
		3	5	19281635	6,1	9	0,08-0,10	At5g47530		GTPasa	2,3		
		3	5	19290045	5,9	3	0,05-0,08	At5g47560	TDT	Proteína de respuesta a uxinas	1,2,3		
		3	5	19631003	6,1	3	0,05-0,08	At5g47560		Transportador de dicarboxilato	1,2,3		
		3	5	19631003	6,1	1	0,03	At5g48440		Oxireductasa dependiente de FAD	1		
		Genes tricomas											
		1	1	19809188	4,8	1	0,03	At1g07867	miR156	Regula la actividad de las proteínas SPL	3		
		2	1	12959279	4,0	1	0,09	At2g30410	KIS	Cofactor de plegamiento de tubulina	3		
		2	2	12966406	5,0	2	0,04-0,05	At2g30424	TCL2	Regulador negativo del desarrollo de tricomas	3		
		2	2	12969839	5,0	3	0,04-0,05	At2g30432	TCL1	Regulador negativo del desarrollo de tricomas	1,3		
		2	2	13463562	6,1	4	0,03-0,04	At2g31660	SAD2	Desarrollo de tricomas mediante ABA	3		
		2	2	18812711	5,2	3	0,05	At2g45660	SOC1	Control de la floración y transición entre fases vegetativas	3		
3	3	10365212	4,4	2	0,05	At3g27920	GL1	Regulador positivo del desarrollo de tricomas	3				
3	3	22413273	4,4	1	0,07	At3g60630	LOM2	Mantenimiento del meristemo	3				
5	5	21580654	5,2	1	0,03	At5g53200	TRY	Regulador negativo del desarrollo de tricomas	3				
Otros genes asociados relevantes													
2	2	11058127	4,4	1	0,05	At2g25930	ELF3	Control de la floración	3				
3	3	<b>1845693</b>	4,4	1	0,04	At3g06120	MUTE	Diferenciación en el desarrollo de estornas	3				

Tabla Suplementaria 2: Resumen de los análisis de asociación genómica. (Continuación)

Carácter	Análisis	Locus altamente asociado	Cromosoma	Posición SNP	-log(P) MLM	# de SNPs asociados	Frecuencia alelo menor	Gen próximo	Nombre proteína	Función de la proteína	Impacto funcional
PTP	10 loci mas asociados	10	1	<b>12629570</b>	8,0	5	0,06-0,09	At1g34530		Elemento transponible	3
		7	1	15926547	8,2	1	0,06	At1g42460		Elemento transponible	3
		5	1	<b>17530294</b>	8,4	1	0,03	At1g47655		Proteína con dedo de zinc, función desconocida	3
		9	1	18693494	8,0	1	0,05	At1g50460	HKL1	Implicada en la relación glucosa-etileno	3
		2	2	3862924	10,6	1	0,04	At2g10175		Elemento transponible	3
		2	2	3964695	8,7	4	0,05-0,07	At2g10300		Elemento transponible	3
		4	2	12917943	8,1	6	0,06-0,22	At2g30300		Implicada en facilitación	1,2,3
		4	2	<b>12966406</b>	8,7	4	0,04-0,05	At2g30424	TCL2	Regulador negativo del desarrollo de tricomas	3
		4	2	<b>12969839</b>	8,7	5	0,04-0,05	At2g30432	TCL1	Regulador negativo del desarrollo de tricomas	1,3
		8	2	17526761	8,1	4	0,09-0,10	At2g41980		Ubiquitinización	3
		3	2	<b>18471233</b>	9,1	4	0,03-0,07	At2g44798		ARN antisentido	3
		3	2	18991835	8,6	2	0,06-0,07	At2g46250	PLDPP2	Miosina de cadena pesada	2,3
		1	3	1637452	8,4	2	0,06	At3g05630	MUTE	Regulación del tráfico de vesículas	2,3
		1	3	<b>1846051</b>	10,4	8	0,04-0,07	At3g06120		Diferenciación en el desarrollo de estomas	3
		6	5	21670538	7,0	15	0,03-0,06	At5g53420		Proteína con motivo CCT, función desconocida	0,1,2,3
			2	12958564	4,7	5	0,03-0,10	At2g30410	KIS	Cofactor de plegamiento de tubulina	3
			2	12961609	5,7	3	0,07-0,08	At2g30420	ETC2	Regulador negativo del desarrollo de tricomas	2,3
			2	12966406	8,7	4	0,04-0,05	At2g30424	TCL2	Regulador negativo del desarrollo de tricomas	3
			2	12969839	8,7	5	0,03-0,05	At2g30432	TCL1	Regulador negativo del desarrollo de tricomas	1,3
			2	17497719	6,1	2	0,09	At2g41940	ZFP8	Regulador de la formación de tricomas mediante GAs	1,2
			2	18812711	6,9	4	0,04-0,08	At2g45660	SOC1	Control de la floración y transición entre fases vegetativas	3
			4	228354	6,3	1	0,07	At4g00480	MYC1	Regulador positivo del desarrollo de tricomas	1
			5	16120436	6,4	1	0,04	At5g40330	MYB23	Regulador positivo del desarrollo de tricomas	3
			5	25953580	6,4	1	0,08	At5g64930	CPR5	Proteína de membrana para la patogénesis	3
	Otros genes asociados relevantes										
										<i>El análisis no muestra otros genes asociados relevantes</i>	

**Tabla suplementaria 2: Resumen de los análisis de asociación genómica (Continuación)**