



UNIVERSIDAD DE JAÉN
Facultad de Ciencias Experimentales

Trabajo Fin de Grado

Identificación y análisis transcriptómico de genes de determinación floral en olivo

Alumno: Manuel Rueda Barranco

Julio, 2018



UNIVERSIDAD DE JAÉN
Facultad de Ciencias Experimentales

Trabajo Fin de Grado

Identificación y análisis transcriptómico de genes de determinación floral en olivo

Una firma manuscrita en tinta azul, que parece leer "Manuel", con trazos fluidos y extendidos.

Alumno: Manuel Rueda Barranco

Julio, 2018

ÍNDICE

RESUMEN - Palabras clave	7
ABSTRACT- Keywords.....	7
1. INTRODUCCIÓN.....	9
2. OBJETIVOS.....	16
3. MATERIAL Y MÉTODOS.....	17
3.1. Secuencias de aminoácidos de las proteínas que intervienen en la floración de Arabidopsis.....	17
3.2. Obtención de los genes en <i>O. europaea</i> L. cv. Picual, homólogos de los genes seleccionados de Arabidopsis.....	18
3.3. RNAseq de órganos/tejidos del olivo.....	18
3.4. Preprocesamiento de los datos de RNAseq y análisis transcriptómico...	18
4. RESULTADOS.....	20
4.1. Obtención de genes de olivo homólogos a los genes de floración de Arabidopsis.....	20
4.2. Análisis transcriptómico de los genes seleccionados.....	21
4.3. Expresión de los genes estudiados en olivo.....	22
4.3.1. Expresión de los genes <i>SEP3</i> del olivo.....	22
4.3.2. Expresión de los genes <i>CAL</i> del olivo.	23
4.3.3. Expresión de los genes <i>FT</i> del olivo.	23
4.3.4. Expresión de los genes <i>AP1</i> del olivo.	24
4.3.5. Expresión de los genes <i>SOC1</i> del olivo.	25
4.3.6. Expresión de los genes <i>AP3</i> del olivo.	26
4.3.7. Expresión de los genes <i>AG</i> del olivo.	27
4.3.8. Expresión de los genes <i>TFL1</i> del olivo.	28
4.3.9. Expresión de los genes <i>FLC</i> del olivo.	29
4.3.10. Expresión de los genes <i>PI</i> del olivo.	30
4.3.11. Expresión de los genes <i>LFY</i> del olivo.	31
4.3.12. Expresión de los genes <i>FD</i> del olivo.	32
5. DISCUSIÓN.....	34
6. CONCLUSIONES.....	38
7. BIBLIOGRAFÍA.....	39

RESUMEN

La floración es un proceso complejo, que conlleva una serie de cambios en el desarrollo de la planta. Actualmente gracias al desarrollo de la tecnología y las técnicas bioinformáticas, es posible identificar y relacionar las bases moleculares que llevan a cabo la floración. Hoy día, la planta base para el estudio de las bases moleculares de la floración es *Arabidopsis thaliana*, debido al gran conocimiento de todo su mapa genético, por tanto a partir de esta planta y utilizando herramientas bioinformáticas podemos identificar genes homólogos en otras plantas, en nuestro caso *Olea Europaea* var. 'Picual' (olivo picual). Además con experimentos RNAseq podemos ver la expresión de dichos genes en los diferentes lugares de la planta, con la finalidad de relacionar su actividad con el proceso de floración.

Finalmente, podemos concluir que identificar las bases moleculares de la floración del olivo nos aportaría información para mejorar la floración del cultivo y obtener mayores beneficios.

Palabras clave: Floración- *Arabidopsis thaliana*- Genes homólogos- Olivo- Bases moleculares.

ABSTRACT

Flowering is a complex process, which involves a series of changes in the development of the plant. Currently grace to the development of technology and bioinformatics techniques, it is possible to identify and relate these molecular bases of flowering. Today, the base plant for the study of the molecular bases of flowering is *Arabidopsis thaliana*, due to the great knowledge of all its genetic map, therefore from this plant and using bioinformatic tools can identify homologous genes in other plants in our case, *Olea Europaea* var. 'Picual'; (picual olive). In addition with RNAseq experiments we can see the expression of these genes in the different places of plant, in order to relate their activity in the flowering process. Finally, we can conclude that identify the molecular bases of flowering of the olive tree would bring us information to improve the crop flowering and get more benefits.

.Keywords: Flowering - *Arabidopsis thaliana* - homologous Genes -olive - molecular Bases

1. INTRODUCCIÓN

La floración es un proceso complejo que ha despertado fascinación en la humanidad, principalmente a científicos y pensadores, desde tiempos remotos; pero solo desde hace veinte años se han empezado a determinar las bases moleculares que rigen este fenómeno. Conocer el momento en que florecen las plantas de un cultivo resultó crucial en la antigüedad para optimizar la recolección de frutos y semillas, y hoy día sigue siendo un carácter primordial para la mejora de la producción agrícola. La floración conlleva en la planta un cambio de programa de desarrollo por otro. (Blázquez et al., 2011)

La inducción floral es el proceso mediante el cual una serie de estímulos que se originan fuera del ápice del brote floral induce la formación de primordios florales. La inducción fotoperiódica de la floración fue descubierta hace 86 años por Julien Tornois en lúpulo (Tournois, 1914). Posteriormente una serie de experimentos adicionales sugirieron que el control de floración por fotoperiodo era un fenómeno general que controlaba la floración en la mayoría de las plantas (Garner & Allard, 1920). Más tarde experimentos de luz enfocada mostró que las hojas son las estructuras que perciben la señal del fotoperiodo (Knott, 1934). Estos estudios y junto a numerosos experimentos de injerto, indican que la producción del estímulo floral (Bernier et al., 1993) inducida por el fotoperiodo se produce en las hojas en una gran variedad de plantas (Chailakhyan, 1968; Lang, 1965).

Para llevar a cabo la floración las plantas se han adaptado y han generado una complicada red de señalización molecular que les sirve para percibir condiciones ambientales (sobre todo, luz y temperatura) y decidir el mejor momento del año para florecer. Los últimos conocimientos aportan información sobre el modo en que las plantas memorizan el paso del invierno, identifican la estación según la longitud del día e integran señales internas con la finalidad de activar la floración. Actualmente la principal proteína activadora de la floración es conocida como FT, así como sus funciones en la actividad de la floración. Esta sustancia se genera en las hojas y se desplaza por los haces vasculares hasta los ápices, donde desencadena la formación de las flores. La plasticidad de la floración, se refleja a nivel molecular, en la integración de los cambios ambientales (Blázquez et al., 2011).

En 1970, Goethe propuso que los órganos florales representan hojas modificadas (Goethe, 1790). Esta idea de un mecanismo subyacente común ha sido considerablemente reforzada por hallazgos recientes que muestran que la acción de un meristemo floral en la formación de órganos florales tiene considerables similitudes a la de la yema apical para la producción de hojas (Carles & Fletcher, 2003; Sablowski, 2007). No obstante, también es evidente que hay una serie de productos génicos que operan específicamente durante el desarrollo de flores.

El proceso de floración ha sido posible entenderlo gracias al moderno estudio de mutantes florales en plantas de *Arabidopsis thaliana* (Blázquez et al., 2011). Actualmente la base del genoma más popular es la de *Arabidopsis* (TAIR), ya que esta proporciona una secuencia actualizada del genoma y diversa información genómica (García-Hernandez et al., 2002), además de varias redes de información genética y de interacción de proteínas (Schoof et al., 2002).

En *Arabidopsis* al igual que en la mayoría de las plantas, el meristemo apical da lugar a un meristemo floral reproductivo cuando percibe las señales apropiadas para este cambio de programación (Amasino, 2010). En *Arabidopsis* se forma un meristemo de florescencia, es decir a partir de un meristemo floral se producen una serie de meristemos secundarios adicionales, dando lugar a una inflorescencia o conjunto de flores.

Un meristemo floral difiere de otros meristemos en una serie de aspectos importantes, principalmente porque el órgano floral produce: Sépalos, pétalos, estambres y carpelos, creando así una serie de capas organizadas en anillos concéntricos o verticilos (Steeves & Sussex, 1989).

Las flores de *Arabidopsis* están formadas por cuatro sépalos en su parte más externa o primer verticilo, 4 pétalos que se alternan con las posiciones de los sépalos en el segundo verticilo, seis estambres, que constan de un filamento y una antera en la punta que produce el polen en el tercer verticilo, y la parte reproductora femenina, el gineceo, que se compone de dos carpelos y contiene los óvulos en el cuarto verticilo.

Un meristemo floral de *Arabidopsis*, al igual que otros meristemos apicales, está compuesto por tres capas de células clonalmente distintas. Las capas externa L1 y la capa subepidérmica, que son capas monocelulares, mantienen su organización a través de divisiones celulares anticlinales (Steeves & Sussex, 1989). La capa L3 es subyacente a estas y se compone de varias subcapas de células que se dividen en todas las direcciones. A pesar de esta disposición regular de las divisiones celulares orientadas, existen desviaciones ocasionales de esta regularidad lo que indica que existe una señalización hacia las células del meristemo floral que son esenciales y críticas para producir la floración (Jenik & Irish, 2000; Reddy et al., 2004; Kwiatkowska, 2006).

Nuestro conocimiento sobre la señalización molecular que lleva a cabo la inducción floral, se ha visto muy favorecida por el considerable avance de la Biología Molecular en las últimas décadas. Dicho avance ha permitido la identificación de genes concretos y sus productos génicos, de tipos celulares específicos, así como de procesos biológicos y moleculares (Niklinski et al., 2017). Para entender más la complejidad que lleva a cabo el proceso de la floración se recurre a la Biología de Sistemas, la cual integra y relaciona disciplinas como la Biología Molecular, la Bioinformática, las Matemáticas y la Física (Babtie et al., 2017).

De modo general, podemos decir que la Biología de Sistemas trata del estudio de los sistemas biológicos complejos en todos sus niveles, incluyendo interconexiones entre diferentes disciplinas ómicas (Genómica, Proteómica, Transcriptómica, Metabolómica y Epigenómica). El progreso y empleo de las tecnologías de alto rendimiento conlleva que, para el análisis de la multitud de datos que pueden obtenerse en la investigación interdisciplinaria de la Biología de Sistemas, sea el modelado matemático un instrumento poderoso para su comprensión (You, 2004).

Como ejemplo de aplicación de los procedimientos propios de la Biología de Sistemas nos centraremos en el estudio y análisis de proteínas que participan en la floración de *Arabidopsis*, ya que como planta modelo, es nuestra planta elegida como referencia para encontrar genes homólogos que participen en la floración de *Olea europaea* cv. Picual.

Una de los principales reguladores a lo largo de todo el ciclo de vida de *Arabidopsis* es el gen *FLOWERING LOCUS C (FLC)*, además tiene un papel clave en el momento de la iniciación de la floración. *FLC* se une al ADN y reprime a dos genes que promueven la floración, *FLOWERING LOCUS T (FT)* y *SUPPRESSOR OF CONSTANS1 (SOC1)*, de los cuales hablaremos con posterioridad. También ha quedado demostrado que *FLC* se une a muchos otros genes, lo que indica que tiene varias dianas en la represión de la floración. Cabe destacar que *FLC* se une a las regiones promotoras de los genes, en una secuencia diana denominada MADS box. El gen *FLC* está involucrado en el desarrollo reproductivo, vegetativo e incluso a nivel embrionario a lo largo de la vida de la planta (Weiwei et al, 2011).

FLC tiene su papel clave durante el periodo de vernalización (Sheldon et al., 1999). Antes de la vernalización *FLC* reprime la floración, pero tras un periodo prolongado de baja temperatura, la expresión de *FLC* es reprimida, lo que permite que se pueda iniciar la floración en la planta. La represión de *FLC* está relacionada con modificaciones de la cromatina, que impide la transcripción de dicho gen (Bastow et al., 2004; Sung & Amasino, 2004). La represión de *FLC*, libera los sitios de control de la transcripción de los genes *SOC1*, *FD*, *AP1*, *SEP3* y *FT*, lo cual permite su expresión, se producen transcritos y finalmente dichas proteínas, iniciando la floración (Dennis & Peacock, 2007; Searle et al., 2006; Helliwel et al., 2006).

LEAFY (LFY) es una proteína clave en la especificación del meristemo floral. En plantas con mutaciones que inactivan *LFY* se ha observado que en vez de producirse meristemos florales se producen ramas secundarias (Weigel et al., 1992).

En plantas mutantes para *LFY* se ha observado que aparecen meristemos axilares que poseen alguna identidad floral, causado por la actividad de *APETALA1 (AP1)* (Hueala and Sussex, 1992; Mandel et al., 1992; Bowman et al., 1993; Mandel & Yanofsky, 1995). En plantas con pérdida de función de *AP1* debido a mutaciones, se ha observado que el meristemo floral da lugar a una florescencia anormal, así como en plantas de doble mutantes para *LFY* y *AP1*, no se observa la formación de flores, lo cual indica que estos dos genes son en gran parte responsables de la especificación floral (Irish &

Sussex, 1990; Huala & Sussex, 1992; Bowman et al., 1993; Shannon & Meeks-Wagner, 1993). *AP1* codifica una proteína con dominio de unión a MADS box en otros genes (Maizel et al. 2005; Tanahashi et al. 2005), los cuales participan en la producción del meristemo floral, incluyendo parálogos para *AP1*, como *CAULIFLOWER (CAL)* (Bowman et al., 1993; Kempin et al., 1995; Ferrandiz et al., 2000a) u otros genes como *SOC1* (Gregis et al., 2008; Melzer et al., 2008).

En lugares donde no se deben formar flores se necesita otro regulador para reprimir la acción de *LFY* y *AP1*, de ello se encarga *TERMINAL FLOWER 1 (TFL1)*, ya que en plantas mutantes para *TFL1* se ha observado la conversión a meristemas florales de meristemas que no deberían ser florales. (Bradley et al., 1997; Ratcliffe et al., 1998). Adicionalmente, *LFY* es capaz de reprimir a *TFL1*, y suprimir el destino indeterminado del meristemo dando lugar a un meristemo floral (Weigel et al., 1992; Liljegren et al., 1999; Ratcliffe et al., 1999). Mientras que en otras zonas *TFL1* actúa para reprimir *LFY* y *AP1* en los meristemas de inflorescencia (Ratcliffe et al., 1998). Este equilibrio entre *TFL1* y los genes de identidad del meristemo floral regula la formación de flores en el lugar y momento adecuado (Prusinkiewicz et al., 2007).

Existen otros factores que regulan al alza la expresión de *AP1* en los primordios florales, dichos factores están regulados por el fotoperiodo. El factor de transcripción bZIP *FD* actúa formando un complejo con *FT* (*FT/FD*), el cual produce una activación directa de *AP1* (Wigge et al., 2005). La activación de *FT* es controlada por el fotoperiodo, principalmente los días largos o picos bajos de dióxido de carbono, activan la transcripción de *FT*, dando lugar a una transcripción constante de dicho gen (Putterill et al., 1995; Suárez-López et al., 2001). Una vez en el meristemo *FT* forma el complejo *FT/FD*, induciendo así la floración.

Como curiosidad y relacionando la proteína *TFL1* mencionada anteriormente junto a *FT*, son proteínas homólogas que tienen un ancestro común, en el cuál tenían una función similar pero debido a la separación evolutiva de gimnospermas y angiospermas, la función de *FT* cambió, siendo esta la que promueve la floración mientras que *TFL1* mantuvo su actividad inhibiendo la floración (Karlgrén et al., 2001; Klintenäs et al., 2012).

AP1 a su vez puede reprimir la expresión de *SOC1*, el cual actúa reprimiendo la expresión de genes con dominios de unión a MADS box, como es *SEPALLATA (SEP3)*. En este caso AP1 activa la expresión de *SEP3*, la cual puede interactuar con *LFY* y promover el desarrollo de la flor, activando genes que participan en la floración que se localizan en la MADS box (Honma & Goto, 2001; Castillejo et al., 2005; Immink et al., 2009; Liu et al., 2009).

Otro gen que interactúa y codifica factores de transcripción con dominios de unión a MADS box es *AGAMOUS (AG)*, la cual es fundamental para limitar la proliferación de células madre en el meristemo floral (Bowman et al., 1989; Yanofsky et al., 1990). *LFY* es un regulador de la expresión de *AG*. En meristemas florales *LFY* se une a sitios adyacentes de la región reguladora de *AG*, realizando una regulación positiva (Busch et al., 1999; Lenhard et al., 2001; Lohmann et al., 2001; Hong et al., 2003). A su vez la activación de *AG* provoca la activación de *WUS*, que provoca una retroalimentación negativa (Sablowski, 2007).

El papel fundamental de todas las relaciones que se dan entre todos estos genes de identidad del meristemo floral es activar a los genes de identidad de órgano floral. Para ello es necesario explicar el modelo clásico “ABC” de la identidad específica del órgano floral (Bowman et al., 1991; Coen & Meyerowitz, 1991; Weigel & Meyerowitz, 1994). Este modelo indica que los tres genes actúan de manera combinatoria para especificar de manera combinatoria cada tipo de órgano floral. Los genes de clase A inducen el desarrollo de los sépalos en el primer verticilo; los genes de la clase A junto a los genes de la clase B inducen el desarrollo de pétalos en el segundo verticilo; los genes de la clase B junto con los de la clase C inducen el desarrollo de los estambres en el tercer verticilo y los genes de la clase C inducen la formación de los carpelos en el cuarto verticilo. Además se ha propuesto que los genes A y C tienen una regulación negativa entre ellos.

Teniendo en cuenta el modelo “ABC” y los genes mencionados anteriormente, *AP1* Y *AP2* son necesarios para la formación de un sépalo normal y el desarrollo de pétalos, ambos genes confieren la función de A. *AP3* y *PISTILLATA (PI)*, ambos son necesarios para conferir la función de B, mientras que *AG* es necesaria para la función de C (Bowman et al., 1991).

La acción combinatoria de los genes ABC depende de su expresión en regiones concretas para el desarrollo de la flor y todo ello se consigue por una regulación entre los genes mencionados a lo largo de la introducción. *AP1* se expresa inicialmente en toda la flor debido a la actividad de *LFY*, posteriormente su actividad se restringe (Mandel et al., 1992; Parcy et al., 1998; Wagner et al., 1999). *AG* localizado en el tercer y cuarto verticilo reprime *AP1* en estas regiones (Gustafson-Brown et al., 1994). *AG* es inhibido en el primer y segundo verticilo por proteínas que se unen a sus elementos de regulación transcripcional, reprimiendo así *AG* en estas regiones (Sridhar et al., 2004,2006; Gregis et al., 2006). *AP2* cuya función solo se da en el primer y segundo verticilo, se expresa en toda la flor pero es reprimida por microARN en el tercer y cuarto verticilo (Aukerman & Sakai, 2003; Chen, 2004; Zhao et al., 2007). En el dominio B la activación de *AP3* depende directamente de la activación conjunta de *LFY* y *AP1* (Lee et al., 1997; Ng & Yanofsky, 2001; Chae et al., 2008). Estos dos últimos junto a *SEP3*, regulan la actividad de *PI* en el dominio B (Liu et al., 2009).

2. OBJETIVOS

El objetivo principal de este Trabajo Fin de Grado es identificar en olivo genes homólogos a genes de floración de Arabidopsis, y comprobar la expresión de dichos genes a partir de los resultados de RNAseq de tejidos/órganos de la variedad Picual, con la finalidad de ver si esa expresión es consistente y corroborativa de la función de floración de dichos genes en olivo.

3. MATERIAL Y MÉTODOS

3.1. Secuencias de aminoácidos de las proteínas que intervienen en la floración de Arabidopsis.

Los códigos de identificación de las proteínas codificadas por los genes de floración de Arabidopsis se obtuvieron de la base de datos UniProt, que es la colección de información funcional sobre proteínas (Uniprot database). En la tabla 1 se muestran las proteínas y genes elegidos para el estudio junto a su código de identificación.

Nombre del gen	Código de identificación (ID)
<i>FLC</i>	AT5G10140
<i>FT</i>	AT1G65480
<i>SOC1</i>	AT2G45660
<i>AP1</i>	AT1G69120
<i>LFY</i>	AT5G61850
<i>TFL1</i>	AT5G03840
<i>FD</i>	AT4G35900
<i>AG</i>	AT4G18960
<i>AP3</i>	AT3G54340
<i>SEP3</i>	AT1G24260
<i>PI</i>	AT5G20240
<i>CAL</i>	AT1G26310

Tabla 1. Proteínas codificadas por genes de floración de Arabidopsis junto a su ID.

Las secuencias de los genes se tomaron de la base de datos TAIR, de la cual se obtuvo el archivo Araport11 ([Araport11_genes.201606.pep.fasta.gz](http://araport11.genes.201606.pep.fasta.gz)). Araport es un portal de información de Arabidopsis, en el cual se encuentra la información molecular de los diferentes genes de Arabidopsis junto a la anotación asociada (Vivek et al.2015).

3.2. Obtención de los genes en *O. europaea* L. cv. Picual, homólogos de los genes seleccionados de Arabidopsis.

Se realizó el alineamiento de las secuencias de Arabidopsis con las secuencias génicas de olivo mediante un BLAST. Los programas BLAST son herramientas ampliamente utilizadas para buscar la similitud de secuencias de proteínas o ADN en diferentes bases de datos, obteniendo a su vez gran contenido de datos estadísticos. Concretamente se utilizó BLASTp que compara las secuencias peptídicas deducidas de los genes de Arabidopsis con todas las secuencias de peptídicas deducidas de los genes de olivo (Stephen et al.1997).

Para realizar el BLASTp, se creó una red un servidor del grupo de investigación que contiene las bases de datos del genoma de olivo ensamblado por el grupo de investigación, ensamblaje todavía no publicado.

3.3. RNAseq de órganos/tejidos del olivo

El RNAseq había sido realizado previamente por el grupo de investigación y en él se tomaron muestras de flores maduras, frutos en estado de envero, hojas jóvenes, tallos y raíces de varias plantas de la variedad Picual. De todas ellas se realizó un pool con varias muestras. La toma de las muestras se realizó mediante congelción inmediata en nitrógeno líquido y se conservaron a -80 °C hasta la extracción del ARN. La purificación y secuenciación del ARN se realizó siguiendo el procedimiento descrito en (Leyva-Pérez et al.,2018). Se secuenciaron dos muestras independientes de cada órgano/tejido, con dos réplicas técnicas en cada una de ellas y obteniendo al menos 20 millones de reads por cada muestra biológica independiente.

3.4. Preprocesamiento de los datos de RNAseq y análisis transcriptómico

Los datos no procesados habían sido previamente tratados por el grupo de investigación siguiendo el procedimiento que se describe en (Leyva-Pérez et

al., 2018). Los datos de RNAseq se analizaron usando el programa Arraystar Lasergene 15 y el transcriptoma de referencia obtenido por el grupo de investigación a partir del genoma ensamblado de Picual (transcriptoma no publicado).

4. RESULTADOS

4.1. Obtención de genes de olivo homólogos a los genes de floración de Arabidopsis

Con objeto de obtener del genoma del olivo los genes homólogos a los de floración de Arabidopsis, se realizó un alineamiento con BLASTp de la isoforma 1 de cada gen de Arabidopsis. Dado el carácter de poliploide reciente del olivo (Referencia de Turgay et al 2017) se tomaron las dos secuencias de olivo que presentaban un mejor alineamiento con la de Arabidopsis para su posterior estudio (Tabla 2). Tras el alineamiento con BLASTp se obtuvo 24 genes homólogos de olivo los cuales presentaban una gran similitud con los genes de floración de Arabidopsis, como muestran los datos estadísticos del alineamiento (Tabla 2). Se puede observar como todos los porcentajes de identidad de los genes de olivo seleccionados con los genes de Arabidopsis superan el 50%. Las longitudes de los genes alineados de Arabidopsis y Olivo son muy similares en la mayoría de los casos, para cada par de genes parálogos. También podemos observar en la tabla 2 como los genes *AP1* y *CAL* han alineado ambas con el mismo gen de olivo.

NOMBRE DEL GEN	ID EN ARATH	ID EN OLIVO	% IDENTITY	LONGITUD EN ARATH	LONGITUD EN OLIVO	LONGITUD ALINEAMIENTO	%ARATH QUE ALINEA	E VALUE	BITSCORE
AP3	AT1G24260.1	Oleu061Scf2874g06003.1	75.51	250	239	245	98	2E-121	350
	AT1G24260.1	Oleu061Scf0589g00023.1	75.51	250	239	245	98	2E-121	350
CAL	AT1G26310.1	Oleu061Scf2107g03003.1	57.03	255	242	256	100	4E-91	273
	AT1G26310.1	Oleu061Scf7465g03021.1	57.03	255	243	256	100	1E-89	270
FT	AT1G65480.1	Oleu061Scf5260g00017.1	69.11	175	196	191	97	1E-94	278
	AT1G65480.1	Oleu061Scf0872g04029.1	66.27	175	150	169	97	2E-76	229
AP1	AT1G69120.1	Oleu061Scf2107g03003.1	62.11	256	242	256	100	2E-108	318
	AT1G69120.1	Oleu061Scf3550g00008.1	60.94	256	242	256	100	3E-105	310
SOC1	AT2G45660.1	Oleu061Scf0398g05023.1	65.55	214	217	209	98	5E-81	245
	AT2G45660.1	Oleu061Scf6353g00013.1	65.55	214	217	209	98	5E-81	245
AP3	AT3G54340.1	Oleu061Scf8536g02002.1	59.48	232	227	232	100	1E-94	281
	AT3G54340.1	Oleu061Scf6496g00009.1	59.48	232	227	232	100	1E-94	281
AG	AT4G18960.1	Oleu061Scf1489g01008.1	70.29	252	257	239	94	5E-115	335
	AT4G18960.1	Oleu061Scf0498g00019.1	66.27	252	279	252	99	3E-114	333
TFL1	AT5G03840.1	Oleu061Scf2462g02004.1	70.18	177	172	171	97	3E-90	265
	AT5G03840.1	Oleu061Scf3186g04026.1	70.18	177	172	171	97	3E-90	265
FLC	AT5G10140.1	Oleu061Scf2231g05019.1	38.12	196	240	202	98	6E-36	129
	AT5G10140.1	Oleu061Scf0247g05013.1	43.79	196	235	169	82	6E-36	129
PI	AT5G20240.1	Oleu061Scf1966g00015.1	51.66	208	243	211	97	4E-65	205
	AT5G20240.1	Oleu061Scf4707g00009.1	51.18	208	243	211	97	2E-64	203
LFY	AT5G61850.1	Oleu061Scf5598g01043.1	63.3	420	381	406	95	1E-176	503
	AT5G61850.1	Oleu061Scf1518g03031.1	62.81	420	381	406	95	3E-169	484
FD	AT4G35900.1	Oleu061Scf0237g03012.1	40.89	285	223	269	78	4E-34	126
	AT4G35900.1	Oleu061Scf0830g14012.1	40.55	285	187	217	70	7E-28	108

TABLA 2. Genes de olivo con alta homología con los genes de floración de Arabidopsis obtenidos mediante BLASTp.

4.2. Análisis transcriptómico de los genes seleccionados

A partir de un experimento de RNAseq de órganos/tejidos realizado previamente en el laboratorio, analizamos la expresión de los genes seleccionados de cara a obtener un perfil de expresión que sirva de prueba confirmatoria de su anotación funcional. En la Figura 1 se muestra un mapa de calor general de todos los genes analizados, en el que ya se puede ver como algunas parejas de genes parálogos presentan perfiles similares mientras que otros casos hay diferencias en el perfil de expresión.

Como podemos observar en la figura 1, en cada columna se encuentra representada una zona de la planta en nuestro caso vamos a realizar el estudio en fruto, flor, hoja, meristemos, raíz y tallo. También podemos observar como en cada fila encontramos la ID en olivo de nuestro gen, y junto a ello el nombre de nuestro gen. Destacar que un color con una tonalidad cercana a la roja indica alta expresión, mientras que una tonalidad cercana a la azul indica una menor expresión del gen.

Los niveles y perfiles de expresión son variables entre distintos genes y se analizan de forma más detallada en los siguientes apartados.

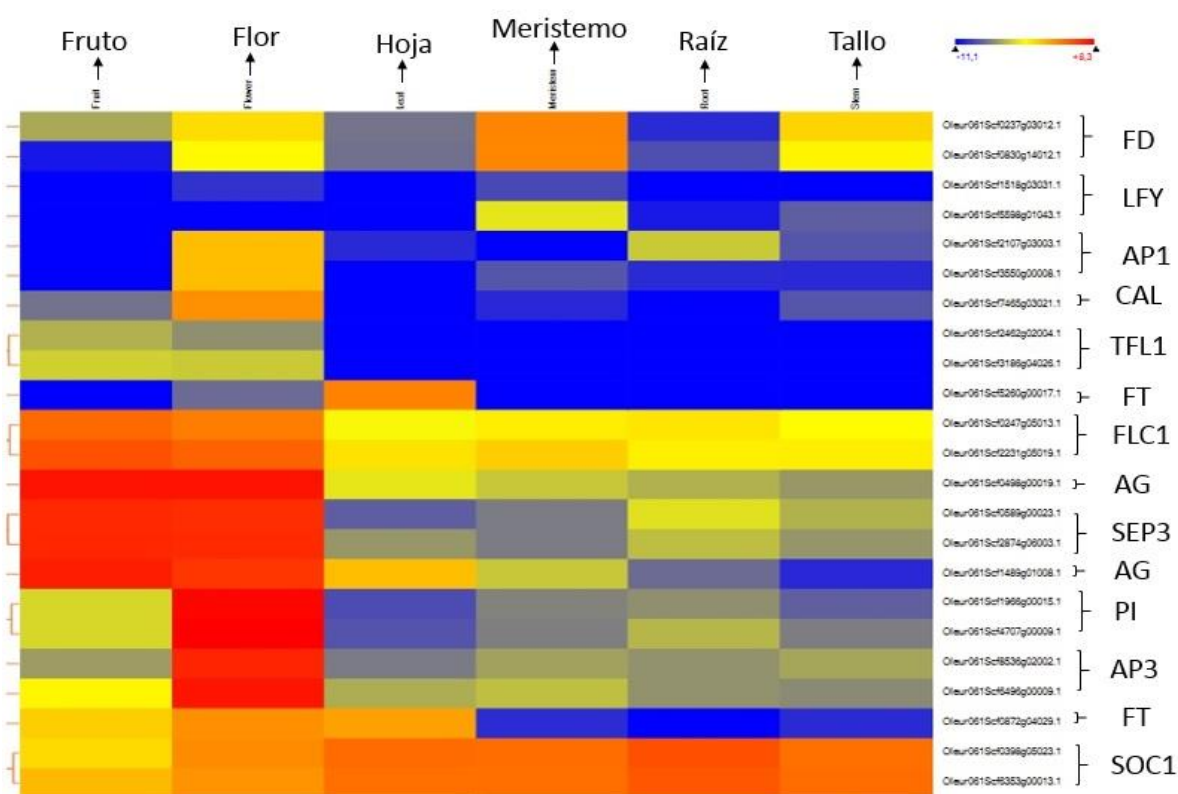


FIGURA1. Mapa de calor general.

4.3. Expresión de los genes estudiados en olivo.

4.3.1. Expresión de los genes *SEP3* del olivo.

El análisis de expresión de estos genes con alta homología con *SEP3* de *Arabidopsis*, será una prueba que apoye o no la esta identidad y su participación en la floración. Como podemos observar en la gráfica 1 vamos a estudiar la frecuencia de los genes parálogos *SEP3* en olivo. Observando dicha gráfica podemos ver que ambos genes parálogos presentan un perfil similar, pero una gran diferencia en el nivel de expresión, ya que el gen *Oleur061Scf0589g00023.1* apenas se expresa en la planta y presenta una mínima expresión en fruto y flor, mientras que el gen *Oleur061Scf2874g06003.1* tiene pequeña expresión en toda la planta y siendo muy alta en flor y fruto. Este perfil de expresión es compatible y esperable con la función de este gen.

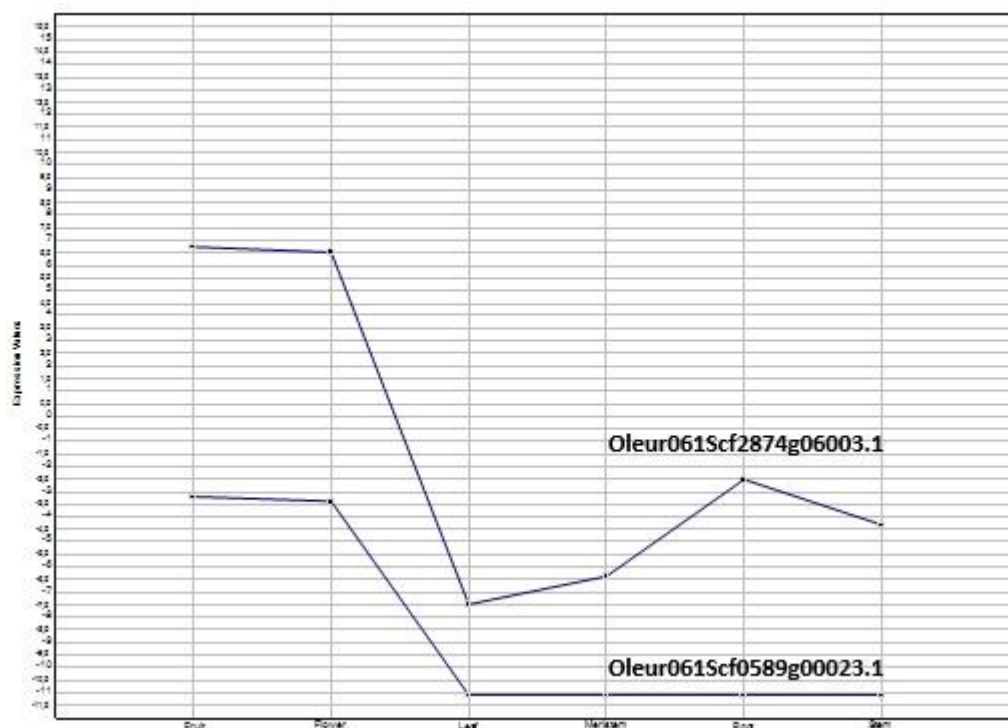


Gráfico 1. Frecuencia del transcrito *SEP3*. La expresión se representa como $\log_2$ de RPKM.

4.3.2. Expresión de los genes CAL del olivo.

La expresión de los genes de olivo con homología con *CAL* se presenta en la gráfica 2. Ambos genes parálogos (Oleur061Scf2107g03003.1 y Oleur061Scf7465g03021.1), presentan una expresión muy similar, siendo ésta elevada en flor, mientras que en el resto de la planta apenas presenta expresión. Destacar que el gen Oleur061Scf2107g03003.1 tiene una ligera expresión en raíz, se trata de un nivel de expresión muy bajo poco significativo. Este perfil de expresión es compatible y esperable con la función de este gen.

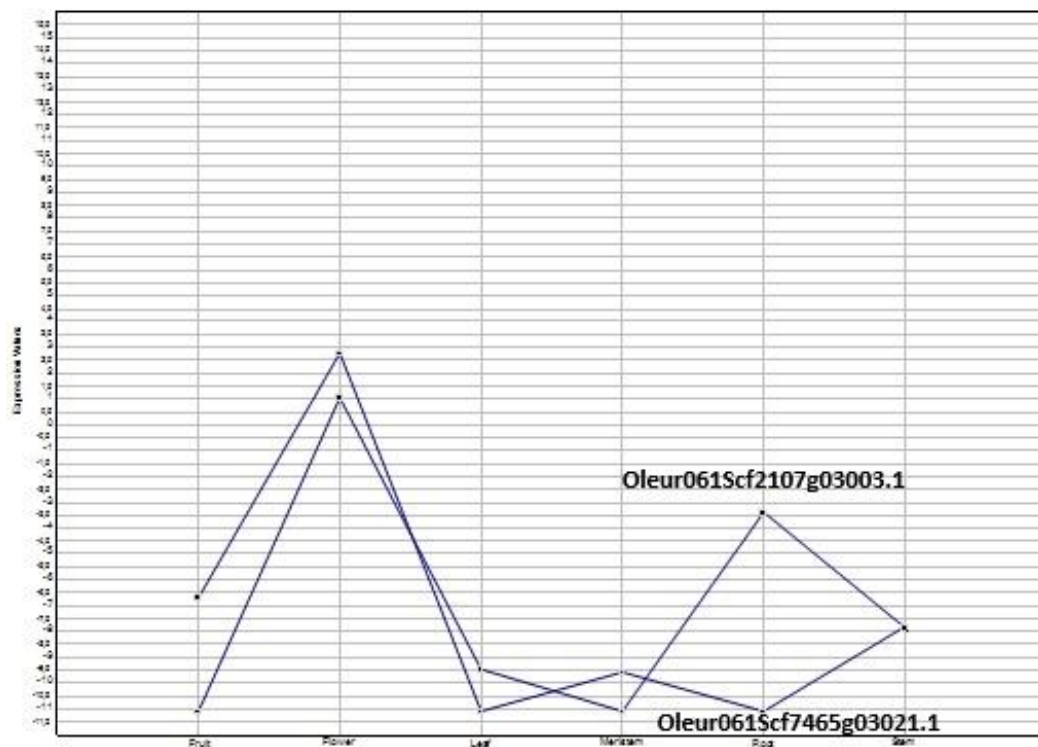


Gráfico 2. Frecuencia del transcrito *CAL*. La expresión se representa como $\log_2$ de RPKM.

4.3.3. Expresión de los genes FT del olivo.

La expresión de los genes de olivo con homología con *FT* se muestra en la gráfica 3. En este caso se observa que ambos genes parálogos presentan perfiles de expresión muy diferentes, debido a que el gen Oleur061Scf5260g00017.1 se expresa solamente en hoja, pero no en flor, lo que no es lo esperable para este gen, mientras que el gen Oleur061Scf0872g04029.1 tiene una gran expresión en hoja, flor y fruto, siendo

muy baja en el resto de las zonas de la planta. Por lo tanto, el gen Oleur061Scf0872g04029.1 es el que probablemente tiene la función de *FT* en el olivo.

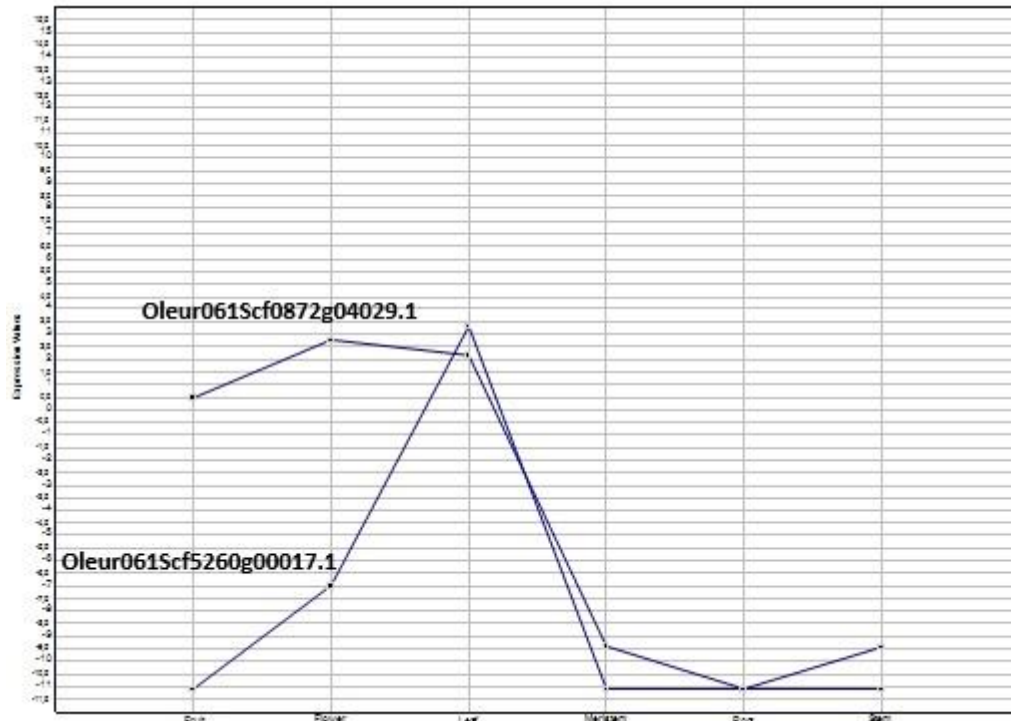


Gráfico 3. Frecuencia del transcrito *FT*. La expresión se representa como $\log_2$ de RPKM.

4.3.4. Expresión de los genes *AP1* del olivo.

La expresión de los genes de olivo con homología con *AP1* se muestra en la gráfica 4. Se puede observar que ambos genes parálogos (Oleur061Scf2107g03003.1 y Oleur061Scf3550g00008.1), presentan un perfil de expresión muy similar, Expresándose solamente en flor, mientras que en el resto de la planta apenas hay expresión, salvo el gen Oleur061Scf2107g03003.1, que tiene una ligera expresión en raíz, siendo un nivel muy bajo poco representativo. Esta expresión específica de la flor es la esperable para la función de *AP1*.

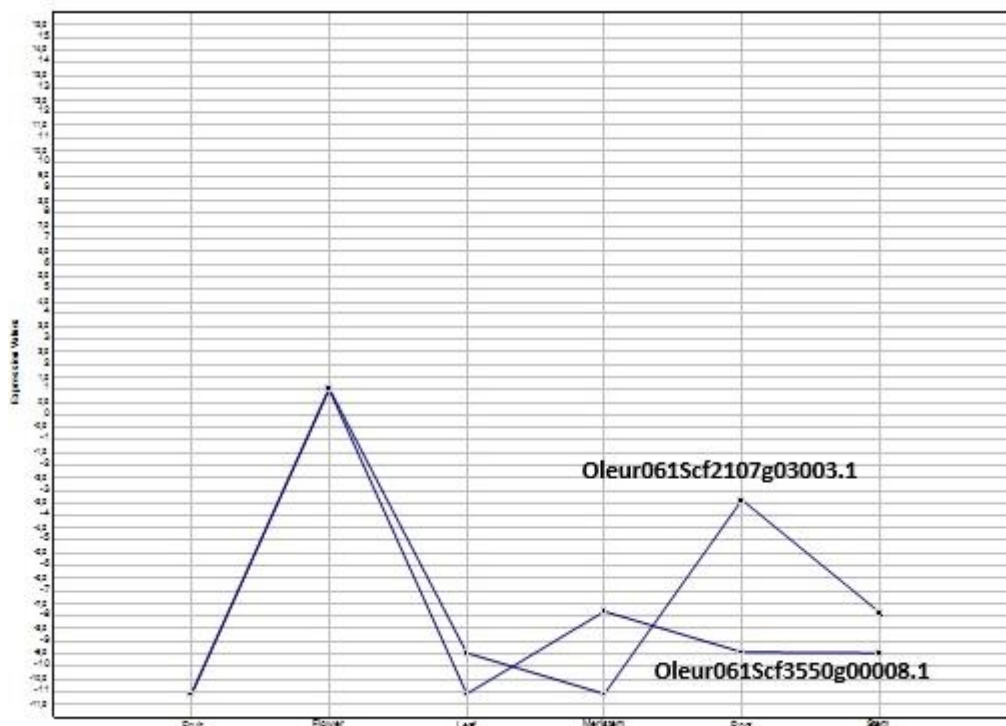


Gráfico 4. Frecuencia del transcrito *AP1*. La expresión se representa como $\log_2$ de RPKM.

4.3.5. Expresión de los genes *SOC1* del olivo.

La expresión de los genes de olivo con homología con *SOC1* se muestra en la gráfica 5. Ambos genes parálogos (Oleur061Scf0398g05023.1y Oleur061Scf6353g00013.1), presentan un perfil de expresión idéntico, tratándose de genes que se expresan en todas las zonas de la planta. Su mayor frecuencia la encontramos en raíz, mientras que la menor frecuencia la encontramos en fruto, en el resto de la planta la frecuencia es muy similar. Este perfil de expresión genera dudas sobre que estos genes sean los que realmente realizan la función de *SOC1* en el olivo.

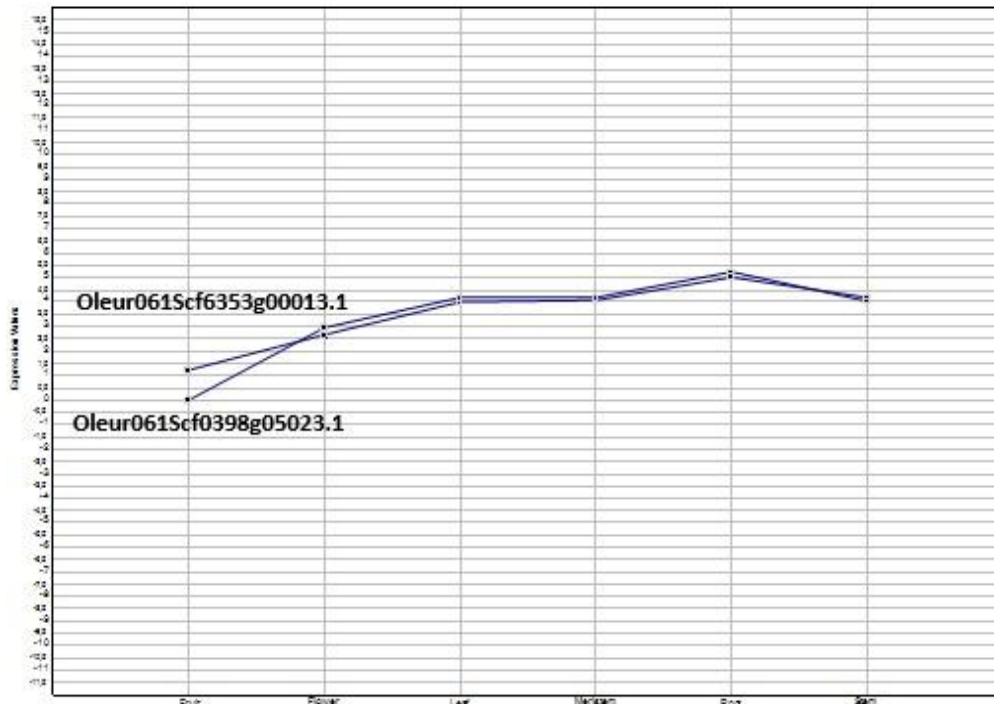


Gráfico 5. Frecuencia del transcrito *SOC1*. La expresión se representa como $\log_2$ de RPKM.

4.3.6. Expresión de los genes *AP3* del olivo.

La expresión de los genes de olivo con homología con *AP3* se muestra en la gráfica 6. Ambos genes parálogos (Oleur061Scf8536g02002.1 y Oleur061Scf6496g00009.1) se expresan solamente en la flor mientras que en el resto de la planta la expresión es muy baja, lo cual concuerda con lo esperable para este gen.

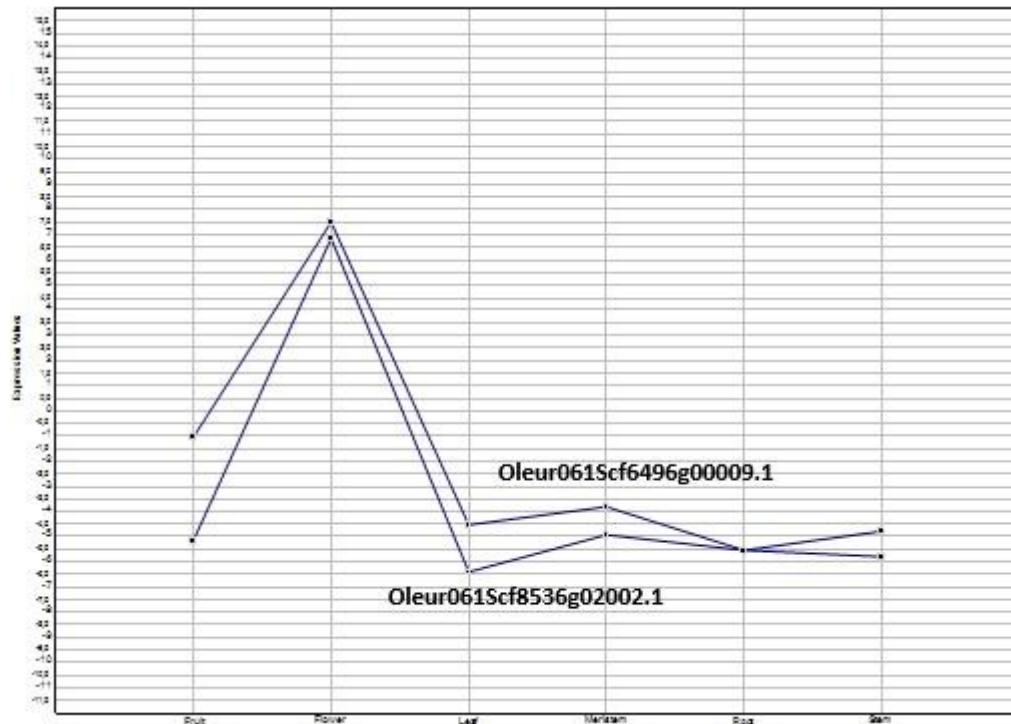


Gráfico 6. Frecuencia del transcrito *AP3*. La expresión se representa como $\log_2$ de RPKM.

4.3.7. Expresión de los genes AG del olivo.

La expresión de los genes de olivo con homología con AG se muestra en la gráfica 7. En este caso de nuevo ambos genes parálogos (Oleur061Scf1489g01008.1 y Oleur061Scf0498g00019.1), presentan un perfil de expresión muy similar, mostrando niveles altos en flor y fruto, mientras que en hoja y meristemo tienen una expresión baja y muy baja en tallo y raíz. Estos datos son también compatibles con que sean los genes AG del olivo.

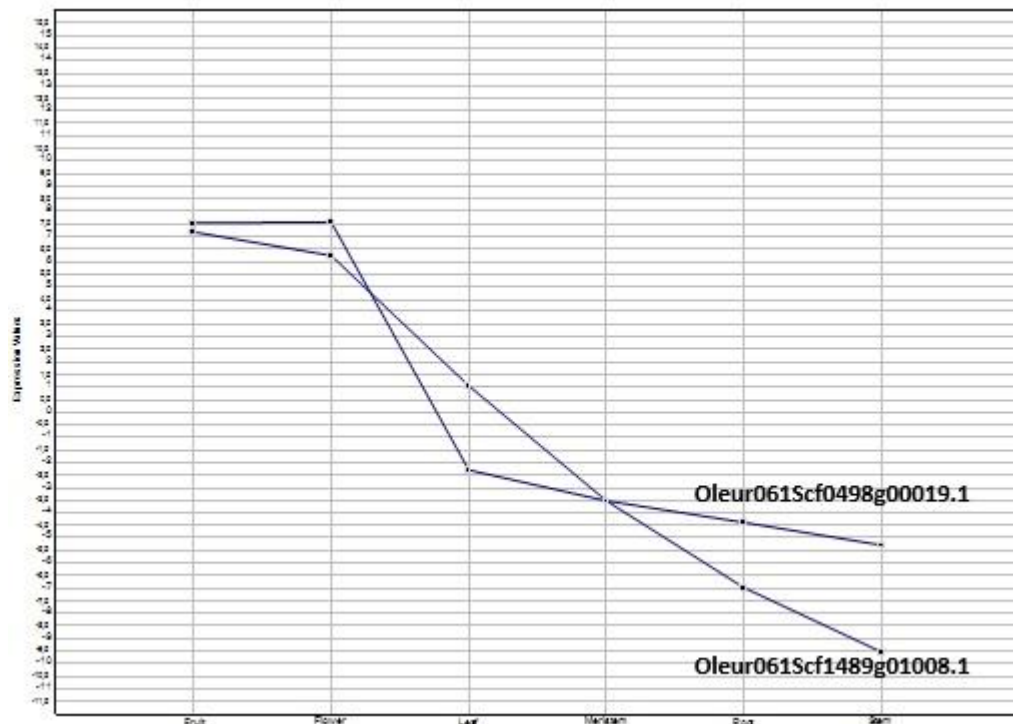


Gráfico 7. Frecuencia del transcrito AG. La expresión se representa como $\log_2$ de RPKM.

4.3.8. Expresión de los genes *TFL1* del olivo.

La expresión de los genes de olivo con homología con *TFL1* se muestra en la gráfica 8. Se observa que ambos genes parálogos (Oleur061Scf2462g02004.1 y Oleur061Scf3186g04026.1), presentan perfiles de expresión muy similares, la expresión es muy baja en flor y fruto mientras que en el resto de la planta la expresión es nula. Dado su papel de represión de la floración estos niveles muy bajos en flor y fruto y ausente en el resto son también esperables para este gen.

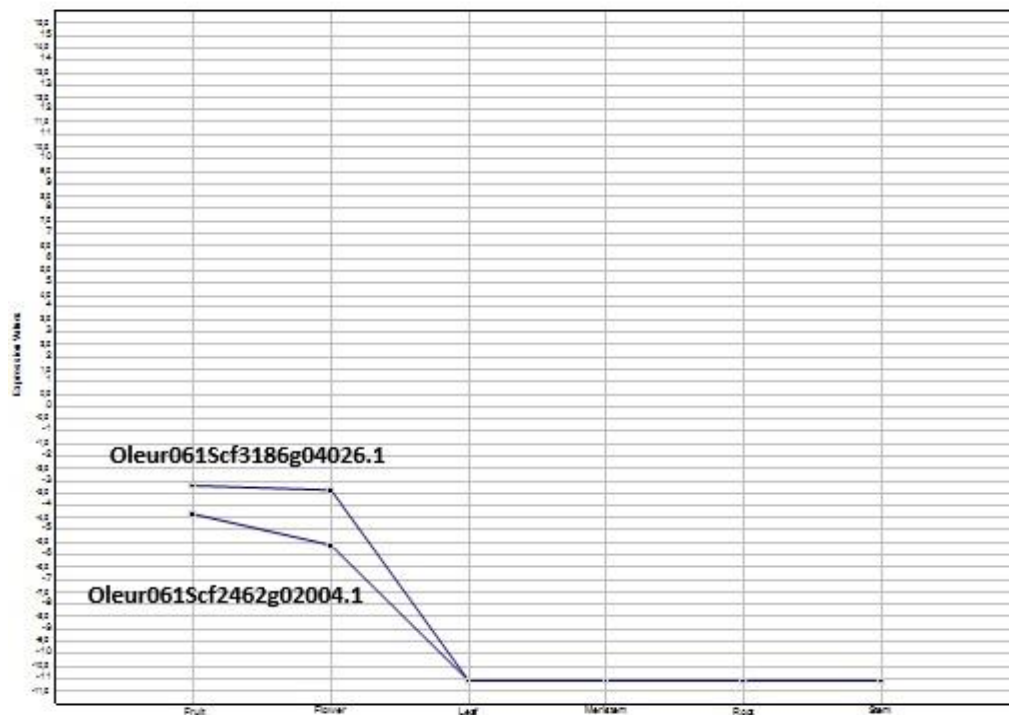


Gráfico 8. Frecuencia del transcrito *TFL1*. La expresión se representa como $\log_2$ de RPKM.

4.3.9. Expresión de los genes *FLC* del olivo.

La expresión de los genes de olivo con homología con *FLC* se muestra en la gráfica 9. De nuevo en este caso ambos genes parálogos (Oleur061Scf2231g05019.1 y Oleur061Scf0247g05013.1), presentan un perfil de expresión muy similar, teniendo una frecuencia muy alta en fruto y flor, y una frecuencia media en hoja, meristemo, tallo y raíz. En este caso el perfil de expresión es compatible con su función, pero dada la regulación anual de este gen hace difícil su interpretación.

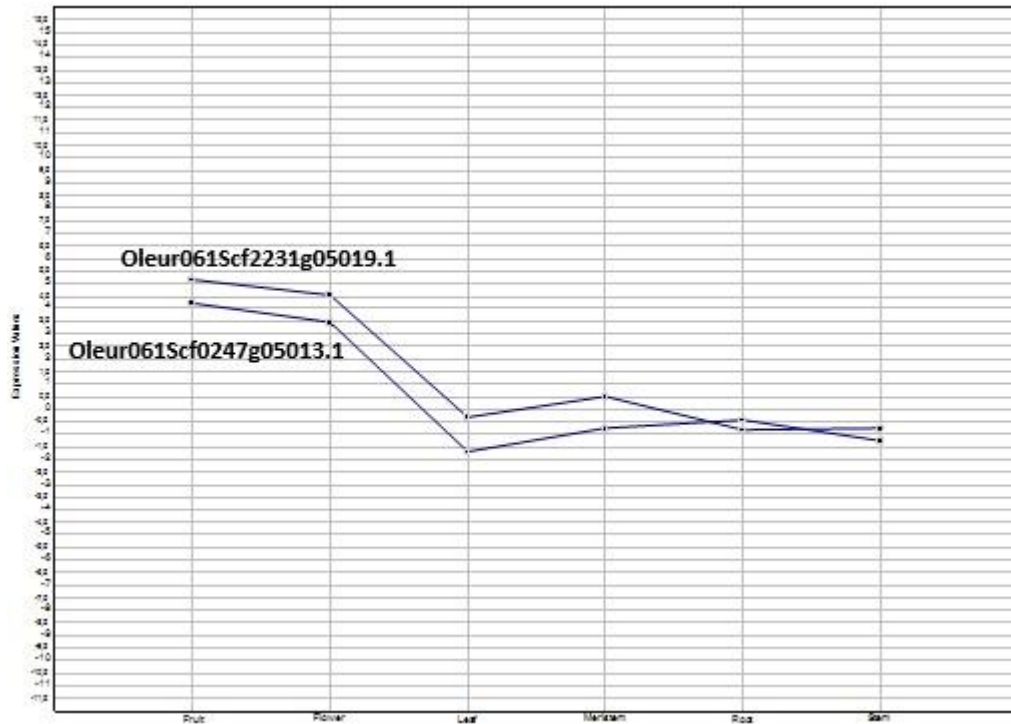


Gráfico 9. Frecuencia del transcrito *FLC*. La expresión se representa como $\log_2$ de RPKM.

4.3.10. Expresión de los genes *PI* del olivo.

La expresión de los genes de olivo con homología con *PI* se muestra en la gráfica 10. Observando dicha gráfica podemos ver que ambos genes parálogos (Oleur061Scf1966g00015.1 y Oleur061Scf4707g00009.1), presentan una expresión casi idéntica y exclusivamente en la flor mientras que en el resto de la planta apenas presenta expresión. Estos datos son también compatibles con que sean los genes *PI* del olivo.

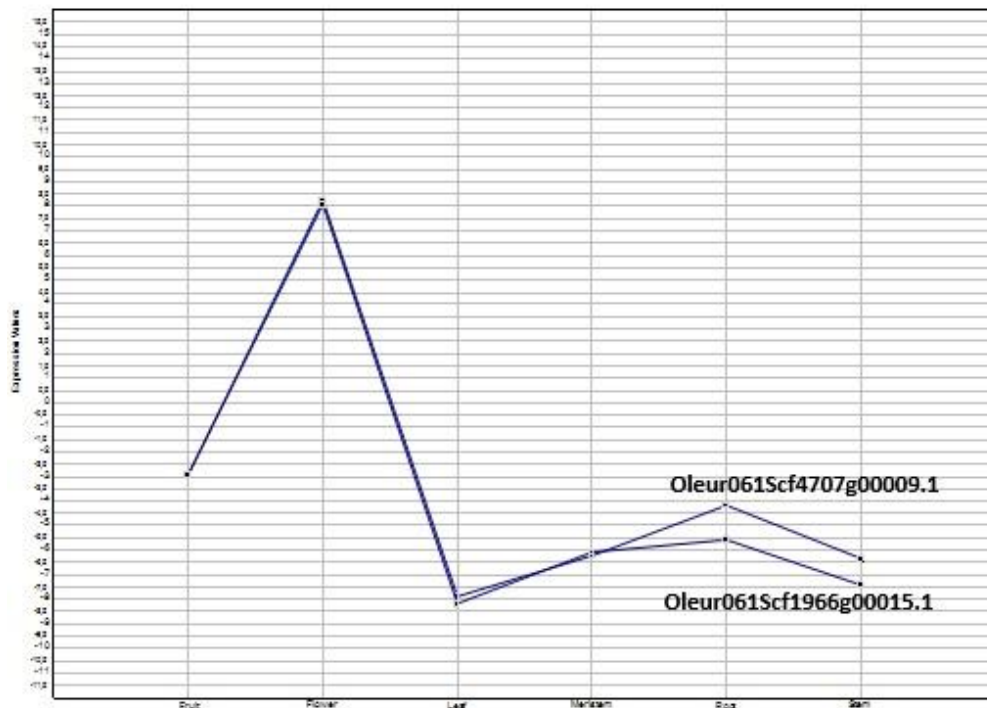


Gráfico 10. Frecuencia del transcrito *PI*. La expresión se representa como $\log_2$ de RPKM.

4.3.11. Expresión de los genes *LFY* del olivo.

La expresión de los genes de olivo con homología con *LFY* se muestra en la gráfica 11. Apenas se observa expresión en ambos genes. Se observa una muy ligera expresión en meristemo del gen Oleur061Scf5598g01043.1, mientras que el gen Oleur061Scf1518g03031.1 no se expresa en ningún tejido. Con estos resultados no queda claro que estos genes sean los verdaderos *LFY* de olivo.

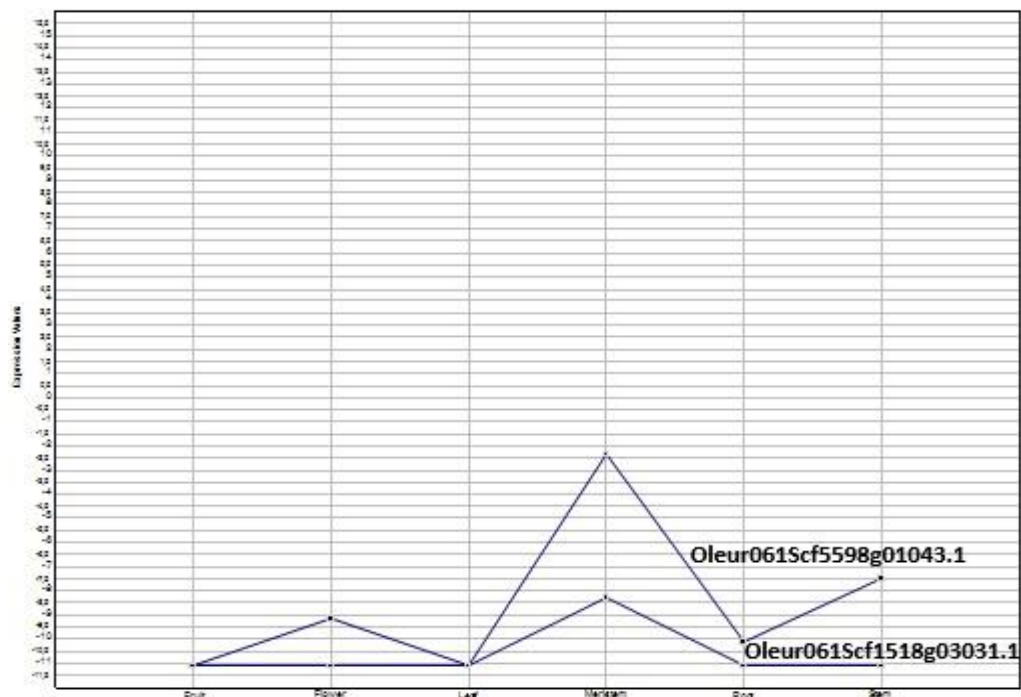


Gráfico 11. Frecuencia del transcrito *LFY*. La expresión se representa como $\log_2$ de RPKM.

4.3.12. Expresión de los genes *FD* del olivo.

La expresión de los genes de olivo con homología con *LFY* se muestra en la gráfica 12. Ambos genes parálogos (Oleur061Scf0237g03012.1 y Oleur061Scf0830g14012.1), presentan un perfil de expresión casi idéntico. En ambos casos se detecta expresión solamente en meristemo, flor y tallo. Este perfil de expresión es compatible con la función de *FD*.

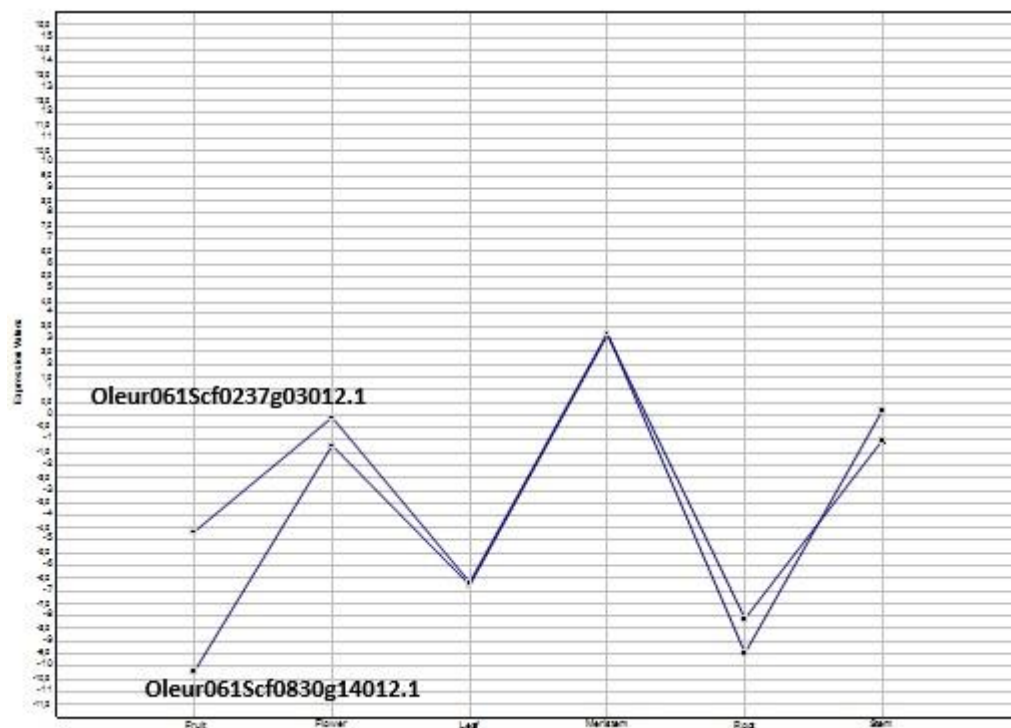


Gráfico 12. Frecuencia del transcrito *FD*. La expresión se representa como $\log_2$ de RPKM.

5. DISCUSIÓN

La floración es un proceso complejo que ha fascinado desde tiempo remotos a científicos y pensadores, pero solo desde hace veinte años se han empezado a determinar las bases moleculares de este proceso (Blázquez et al., 2011).

En este Trabajo Fin de Grado se han buscado los posibles genes de olivo implicados en la floración. Para ello se han buscado genes con alta homología con los genes de floración de Arabidopsis y se ha comprobado su perfil de expresión en distintos órganos/tejidos mediante un análisis de RNAseq. La utilización de Arabidopsis como planta modelo se debe a que fue la primera base de datos del genoma de angiospermas que se creó (2001), además de ser una de las bases de datos genómica más completa de angiospermas (Huala et al., 2001).

Al seleccionar los genes de Arabidopsis implicados en la floración observamos que en algunos casos codificaban varias isoformas, alcanzando hasta 5 isoformas diferentes, pero todas ellas muy similares entre sí. Fundamentalmente variaban en el número de aminoácidos, por lo que cogimos la isoforma codificada por cada gen, que presentaba una mayor longitud de aminoácidos, todo ello con la finalidad de buscar los genes de olivo que representen la versión más larga y completa de los genes.

Al realizar un alineamiento entre genes de floración de Arabidopsis y genes de olivo, escogimos los dos mejores alineamientos para cada gen de Arabidopsis, con lo que obtuvimos 24 genes de olivo. Destacar que se obtuvieron una alta cantidad de alineamientos, posiblemente debido al proceso evolutivo de la especie, ya que en el olivo se han producido eventos de poliploidía por duplicación del genoma relativamente recientes, concretamente se ha descrito que este proceso se producido dos veces, en torno hace 28 millones de años y otro hace 59 millones de años, lo que conlleva que se puedan encontrar varias copias del mismo gen en el genoma, que actualmente se han convertido en genes parálogos (Turgay Unver et al., 2017).

Si observamos los datos de los alineamientos mediante BLASTp obtenidos entre los péptidos codificados por los genes de Arabidopsis y sus posibles homólogos en el olivo, en todos los casos se obtiene un porcentaje de identidad mayor del 50 %, por lo que la similitud entre los péptidos codificados por los genes de ambas plantas es suficientemente alta como para ser considerados como posibles genes ortólogos (Tabla 2). También la mayoría de los genes de Arabidopsis alinean en su secuencia de nucleótidos más de un 90 % con respecto a la secuencia del gen de olivo. Resaltar los altos valores de bitscore obtenidos en todos los casos. Todo este conjunto de datos nos lleva a postular que en el olivo encontramos genes homólogos a los que participan en la floración de Arabidopsis, y por lo tanto estos genes podrían participar en la floración del olivo.

Un dato interesante y muy característico lo tuvimos al alinear las proteínas codificadas por *AP1* y *CA*, resultando que ambas proteínas tenían en común el mismo gen del olivo, es decir, alineaban en olivo con el mismo gen, obteniéndose mejores datos de alineamiento para *AP1*. Esto es debido a que ambos genes son muy similares y tienen dominios parecidos en Arabidopsis, esto junto a la distancia evolutiva entre Arabidopsis y olivo, hace que el gen *CA* de Arabidopsis tenga mayor similitud de secuencia con *AP1* de olivo, que con *CA* de olivo. También es posible que *CA*, durante la evolución del olivo y debido al proceso de duplicación del genoma de olivo, haya perdido su función en olivo y que ésta la ejerza otro gen, como podría ser *AP1* ya que mediante un proceso de neofuncionalización producida en la evolución de olivo, dicho gen habría podido adquirir la función de *CA* (Turgay Unver et al., 2017).

La expresión de los genes seleccionados como probables genes de floración se analizó mediante un análisis transcriptómico a partir de datos de RNAseq que había sido realizado previamente en el laboratorio por el grupo de investigación. Este análisis transcriptómico incluía muestras de diversos órganos/tejidos de la planta, concretamente de la flor, fruto, hoja, meristemo, tallo y raíz. En general el análisis transcriptómico presenta perfiles de expresión que son compatibles y esperables para genes implicados en la floración ya que muestran una expresión más o menos elevada en las flores, aunque con

frecuencia se observa expresión también en fruto. Esto sucede con los homólogos del olivo que se corresponden con los genes *CAL*, *AP1*, *AP3*, *AG*, *FLC* y *PI* de *Arabidopsis*, en los cuales encontramos además una expresión similar para ambos homólogos. Por lo tanto, estos genes de olivo que hemos seleccionado manifiestan unos perfiles de expresión compatibles y esperables con su participación en el proceso de floración del olivo. Sin embargo, se han encontrado patrones de expresión en algunos genes que no apoyan su papel en la floración. En el caso del gen *LFY*, se observó que el gen Oleur061Scf1518g03031.1 no presentó expresión en ningún órgano/tejido de la planta, mientras que el gen Oleur061Scf5598g01043.1 mostró una leve expresión en meristemas y nula en el resto de muestras. Dado que *LFY* es un inductor general de la floración en *Arabidopsis* habríamos esperado encontrar expresión de sus posibles ortólogos en la flor del olivo. El hecho de no encontrar expresión de estos genes en la flor puede ser debido a que actúen en etapas muy iniciales de la inducción de la floración y que después su actividad no sea necesaria y dejen de expresarse o alternatively que ninguno de ellos sea el gen que ha retenido la función de *LFY* después de las duplicaciones genómicas ocurridas en la evolución del olivo. Determinar esto requeriría de posteriores análisis a este Trabajo. En cuanto a los genes con homología con *FT*, uno de ellos, concretamente el gen Oleur061Scf5260g00017.1 no se expresa en flor pero sí en hoja, mientras que el gen Oleur061Scf0872g04029.1 se expresa en fruto, flor y hoja, por lo que este último podría ser el verdadero gen que mantiene la función de *FT*. No obstante, la expresión que presenta en hojas genera dudas sobre que realmente sea el gen que ha mantenido la función *FT* en el olivo. Es conocido que en *Arabidopsis* *TFL1* y *FT* son genes homólogos pero con función inversa de uno con respecto al otro, ya que *TFL1* reprime la floración mientras que *FT* la activa. Ambos genes proceden de un ancestro común, que evolucionó para dar lugar a estos dos genes con homología pero funciones antagónicas (Turgay Unver et al., 2017). Si bien los resultados obtenidos con *FT* arrojan ciertas dudas sobre los homólogos de olivo que hemos analizado en este trabajo, la expresión de los homólogos de *TFL* en el olivo resultó muy baja en flor y fruto y nula en el resto de tejidos, lo que dado su papel de represión de la floración

son resultados esperables para este gen, aunque no suponen una confirmación clara de que realmente realicen esta función en el olivo.

En el caso de *SEP3*, hay un gen (Oleur061Scf2874g06003.1) que tiene una alta expresión en flor y en fruto, siendo expresión mínima en el resto de la planta, mientras que el gen Oleur061Scf0589g00023.1 tiene una expresión muy baja en flor y fruto y nula en el resto de la planta. Esto es frecuente en plantas que han pasado por un episodio de poliploidía reciente, donde uno de los homólogos se expresa más y el otro puede terminar dejando de expresarse, dando lugar a un proceso de degeneración del gen (Turgay Unver et al.,2017).

Con respecto a la expresión de los homólogos del gen *FD* en olivo, se expresa en meristemo, flor y tallo. Este gen codifica un factor de transcripción que no solamente está implicado en la inducción de la floración por lo que este perfil de expresión es compatible con su función como ortólogos del *FD* de *Arabidopsis*.

La expresión de los homólogos de *SOC1* es muy similar en toda la planta teniendo su valor más alto en raíz y el más bajo en fruto. Este perfil de expresión genera dudas sobre que estos genes sean los que realmente realizan la función de *SOC1* en el olivo.

6. CONCLUSIONES

1. En este trabajo hemos encontrado en el genoma del olivo genes homólogos de los que participan en la floración en Arabidopsis, concretamente los genes FD, LFY, AP1, CAL, TFL1, FT, FLC, AG, SEP3, PI, AP3 y SOC1.
2. Los genes homólogos encontrados en olivo de CAL, AP1, AP3, AG, FLC, FT, SOC1, PI y FD de Arabidopsis muestran una expresión en flor acorde con su actividad en la floración.
3. Los genes homólogos de FT encontrados en olivo, muestran un perfil de expresión en fruto, flor y hoja, siendo compatible con su función en el proceso de floración.
4. De los genes homólogos de SEP3, solo uno de ellos muestra un perfil de expresión en flor compatible con su función en la floración.
5. Los genes homólogos FD encontrados en olivo, muestran expresión en tallo, meristemo y flor, lo cual es compatible con su participación en el proceso de floración.
6. Los genes CA y AP1 de Arabidopsis, tienen en común el mismo gen homólogo (Oleur061Scf2107g03003.1) en olivo.

7. BIBLIOGRAFÍA

- Amasino, R. (2010). Seasonal and developmental timing of flowering. *Plant J.* 61, 1001–1013.
- Aukerman, M.J. and Sakai, H. (2003) Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell*, 15, 2730–2741.
- Babtie A, Kirk P, Stumpf M. 2014. Topological sensitivity analysis for systems biology Model space. *Proc Natl Acad Sci USA*. 111: 18507–18512.
- Bernier, G. et al. (1993) Physiological signals that induce flowering. *PlantCell* 5, 1147–1155
- Blázquez, M.A.; Piñeiro M. y Valverde F. 2011. Bases moleculares de la floración. *Temas 61. Investigación y Ciencia* 416: 29- 36.
- Bowman, J.L., Alvarez, J., Weigel, D., Meyerowitz, E.M. and Smyth, D.R. (1993) Control of flower development in *Arabidopsis thaliana* by APETALA1 and interacting genes. *Development*, 119, 721–743.
- Bowman, J.L., Smyth, D.R. and Meyerowitz, E.M. (1989) Genes directing flower development in *Arabidopsis*. *Plant Cell*, 1, 37–52.
- Bowman, J.L., Smyth, D.R. and Meyerowitz, E.M. (1991) Genetic interactions among floral homeotic genes of *Arabidopsis*. *Development*, 112, 1–20.
- Bradley, D., Ratcliffe, O., Vincent, C., Carpenter, R. and Coen, E. (1997) Inflorescence commitment and architecture in *Arabidopsis*. *Science*, 275, 80–83.
- Busch, M.A., Bomblies, K. and Weigel, D. (1999) Activation of a floral homeotic gene in *Arabidopsis*. *Science*, 285, 585–587.
- Carles, C.C. and Fletcher, J.C. (2003) Shoot apical meristem maintenance: the art of a dynamic balance. *Trends Plant Sci.* 8, 394–401.
- Castillejo, C., Romera-Branchat, M. and Pelaz, S. (2005) A new role of the *Arabidopsis* SEPALLATA3 gene revealed by its constitutive expression. *Plant J.* 43, 586–596.

- Chae, E., Tan, Q.K., Hill, T.A. and Irish, V.F. (2008) An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. *Development*, 135, 1235–1245.
- Chailakhyan, M.K. (1968) Internal factors of plant flowering. *Annu. Rev. Plant Physiol.* 19, 1–36
- Chen, X. (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science*, 303, 2022–2025.
- Coen, E.S. and Meyerowitz, E.M. (1991) The war of the whorls: genetic interactions controlling flower development. *Nature*, 353, 31–37.
- Dennis ES, Peacock WJ (2007) Epigenetic regulation of flowering. *Curr Opin Plant Biol* 10:520–527.
- García-Hernández M., Berardini TZ, Chen G., Crist D., Doyle A., Huala E., et al. . (2002). TAIR: un recurso para datos integrados de Arabidopsis. *Funct. Integr. Genomics* 2, 239-253. 10.1007 / s10142-002-0077-z
- Garner, W.W. and Allard, H.A. (1920) Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. *J. Agric. Res.* 18, 553–606
- Goethe, J.W. (1790) Goethe's botany: the metamorphosis of plants. Translated by A. Arber as Goethe's Botany (1946). *Chronica Botanica*, 10, 63–126.
- Gregis, V., Sessa, A., Colombo, L. and Kater, M.M. (2006) AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis. *Plant Cell*, 18, 1373–1382.
- Gregis, V., Sessa, A., Colombo, L. and Kater, M.M. (2008) AGAMOUS-LIKE24 and SHORT VEGETATIVE PHASE determine floral meristem identity in Arabidopsis. *Plant J.* 56, 891–902.
- Gustafson-Brown, C., Savidge, B. and Yanofsky, M.F. (1994) Regulation of the Arabidopsis homeotic gene APETALA1. *Cell*, 76, 131–143.

- Helliwell CA, Wood CC, Robertson M, James Peacock W, Dennis ES (2006) The Arabidopsis FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex. *Plant J* 46:183–192.
- Hong, R.L., Hamaguchi, L., Busch, M.A. and Weigel, D. (2003) Regulatory elements of the floral homeotic gene AGAMOUS identified by phylogenetic footprinting and shadowing. *Plant Cell*, 15, 1296–1309.
- Honma, T. and Goto, K. (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature*, 409, 469–471.
- Huala E., Dickerman A. W., Garcia-hernandez M., Weems D., Reiser L., Lafond F., et al. . (2001). The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucleic Acids Res.* 29, 102–105. 10.1093/nar/29.1.102.
- Huala, E. and Sussex, I.M. (1992) LEAFY interacts with floral homeotic genes to regulate Arabidopsis floral development. *Plant Cell*, 4, 901–913.
- Immink, R.G., Tonaco, I.A., de Folter, S., Shchennikova, A., van Dijk, A.D., Busscher-Lange, J., Borst, J.W. and Angenent, G.C. (2009) SEPALLATA3: the ‘glue’ for MADS box transcription factor complex formation. *Genome Biol.* 10, R24.
- Irish, V.F. and Sussex, I.M. (1990) Function of the *apetala-1* gene during Arabidopsis floral development. *Plant Cell*, 2, 741–753.
- Jenik, P.D. and Irish, V.F. (2000) Regulation of cell proliferation patterns by homeotic genes during Arabidopsis floral development. *Development* 127, 1267–1276.
- Karlgrén, A., Gyllenstrand, N., Kallman, T., Sundström, J.F., Moore, D., Lascoux, M., and Lagercrantz, U. (2011). Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. *Plant Physiol.* 156:1967–1977.

- Kempin, S., Savidge, B. and Yanofsky, M.F. (1995) Molecular basis of the cauliflower phenotype in *Arabidopsis*. *Science*, 267, 522–525.
- Klintenaˆs, M., Pin, P.A., Benlloch, R., Ingvarsson, P.K., and Nilsson, O. (2012). Analysis of conifer FLOWERING LOCUS T/TERMINAL FLOWER1-like genes provides evidence for dramatic biochemical evolution in the angiosperm FT lineage. *New Phytol.* 196:1260–1273.
- Knott, J.E. (1934) Effect of a localized photoperiod on spinach. *Proc. Am.Soc. Hortic. Sci.* 31, 152–154
- Kwiatkowska, D. (2006) Flower primordium formation at the *Arabidopsis* shoot apex: quantitative analysis of surface geometry and growth. *J. Exp.Bot.* 57, 571–580.
- Lang, A. (1965) The physiology of flower initiation. In *Handbuch d.Pflanzenphysiologie (Encyclopedia of Plant Physiology)* (Vol. XV/1)(Ruhland, W., ed.), pp. 1379–1536, Springer-Verlag
- Lee, I., Wolfe, D.S., Nilsson, O. and Weigel, D. (1997) A LEAFY co-regulator encoded by UNUSUAL FLORAL ORGANS. *Curr. Biol.* 7, 95–104.
- Lenhard, M., Bohnert, A., Juˆrgens, G. and Laux, T. (2001) Termination of Stem Cell Maintenance in *Arabidopsis* Floral Meristems by Interactions between WUSCHEL and AGAMOUS. *Cell*, 105, 805–814.
- Leyva-Pérez, M. de la O.; J. Jiménez-Ruiz, C. Gómez-Lama Cabanás, A. Valverde-Corredor, J.B. Barroso, F. Luque, J. Mercado-Blanco (2018). Tolerance of olive (*Olea europaea*) cv Frantoio to *Verticillium dahliae* relies on both basal and pathogen-induced differential transctomic responses. *New Phytologist*, 217:671-686.
- Liljegren, S.J., Gustafson-Brown, C., Pinyopich, A., Ditta, G.S. and Yanofsky, M.F. (1999) Interactions among APETALA1, LEAFY, and TERMINAL FLOWER1 specify meristem fate. *Plant Cell*, 11, 1007–1018.
- Liu, C., Xi, W., Shen, L., Tan, C. and Yu, H. (2009) Regulation of floral patterning by flowering time genes. *Dev. Cell*, 16, 711–722.

- Lohmann, J.U., Hong, R.L., Hobe, M., Busch, M.A., Parcy, F., Simon, R. and Weigel, D. (2001) A molecular link between stem cell regulation and floral patterning in *Arabidopsis*. *Cell*, 105, 793–803.
- Maizel, A., Busch, M.A., Tanahashi, T., Perkovic, J., Kato, M., Hasebe, M. and Weigel, D. (2005) The floral regulator *LEAFY* evolves by substitutions in the DNA binding domain. *Science*, 308, 260–263.
- Mandel, M.A. and Yanofsky, M.F. (1995) A gene triggering flower formation in *Arabidopsis*. *Nature*, 377, 522–524.
- Ng, M. and Yanofsky, M.F. (2001) Activation of the *Arabidopsis* B class Homeotic Genes by *APETALA1*. *Plant Cell*, 13, 739–753.
- Niklinski J, Kretowski A, Moniuszko M, Reszec J, Michalska A, Niemira M, Ciborowski M, Charkiewicz R, Jurgilewicz D, Kozłowski M, Ramlau R, Piwkowski C, Kwasniewski M, Kaczmarek M, Ciereszko A, Wasniewski T, Mroz R, Naumnik W, Sierko E, Paczkowska M, Kisluk J, Sulewska A, Cybulski A, Mariak Z, Kedra B, Szamatowicz J, Kurzawa P, Minarowski L, Charkiewicz AE, Mroczko B, Malyszko J, Manegold C, Pilz L, Allgayer H, Abba ML, Juhl H, Koch F. 2017. Systematic biobanking, novel imaging techniques, and advanced molecular analysis for precise tumor diagnosis and therapy: The Polish MOBIT Project. *Advances in Medical Sciences*. 62: 405-41.
- Parcy, F., Nilsson, O., Busch, M.A., Lee, I. and Weigel, D. (1998) A genetic framework for floral patterning. *Nature*, 395, 561–566.
- Prusinkiewicz, P., Erasmus, Y., Lane, B., Harder, L.D. and Coen, E. (2007) Evolution and development of inflorescence architectures. *Science*, 316, 1452–1456.
- Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80:847–857.

- Ratcliffe, O.J., Amaya, I., Vincent, C.A., Rothstein, S., Carpenter, R., Coen, E.S. and Bradley, D.J. (1998) A common mechanism controls the life cycle and architecture of plants. *Development*, 125, 1609–1615.
- Ratcliffe, O.J., Bradley, D.J. and Coen, E.S. (1999) Separation of shoot and floral identity in *Arabidopsis*. *Development*, 126, 1109–1120.
- Reddy, G.V., Heisler, M.G., Ehrhardt, D.W. and Meyerowitz, E.M. (2004) Realtime lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of *Arabidopsis thaliana*. *Development*, 131, 4225–4237.
- Sablowski, R. (2007) Flowering and determinacy in *Arabidopsis*. *J. Exp. Bot.* 58, 899–907.
- Schoof H., Zaccaria P., Gundlach H., Lemcke K., Rudd S., Kolesov G., y col. (2002). MIPS *Arabidopsis thaliana* Database (MAtdB): un recurso integrado de conocimiento biológico basado en el primer genoma completo de la planta. *Nucleic Acids Res.* 30, 91-93. 10.1093 / nar / 30.1.91
- Searle I, et al. (2006) The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in *Arabidopsis*. *Genes Dev* 20:898–912.
- Shannon, S. and Meekswagner, D.R. (1993) Genetic interactions that regulate inflorescence development in *arabidopsis*. *Plant Cell*, 5, 639.
- Sheldon CC, et al. (1999) The FLF MADS box gene: A repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. *Plant Cell* 11:445–458. Bastow R, et al. (2004) Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 427:164–167.
- Sridhar, V.V., Surendrarao, A. and Liu, Z. (2006) APETALA1 and SEPALLATA3 interact with SEUSS to mediate transcription repression during flower development. *Development*, 133, 3159–3166.
- Sridhar, V.V., Surendrarao, A., Gonzalez, D., Conlan, R.S. and Liu, Z. (2004) Transcriptional repression of target genes by LEUNIG and SEUSS, two

- interacting regulatory proteins for Arabidopsis flower development. *Proc. Natl Acad. Sci. USA*, 101, 11494–11499.
- Steeves, S.T. and Sussex, I.M. (1989) *Patterns in plant development* 2 edn. Cambridge: Cambridge University Press.
- Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, David J. Lipman; Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Research*, Volume 25, Issue 17, 1 September 1997, Pages 3389–3402
- Suárez-López, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature* 410:1116–1120.
- Sung S, Amasino RM (2004) Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature* 427:159–164.
- Tanahashi, T., Sumikawa, N., Kato, M. and Hasebe, M. (2005) Diversification of gene function: homologs of the floral regulator FLO/LFY control the first zygotic cell division in the moss *Physcomitrella patens*. *Development*, 132, 1727–1736.
- Tournois, J. (1914) Etudes sur la sexualite du houblon. *Ann. Sci. Nat. Bot.*(9th Series) 19, 49–189.
- Turgay Unver, Zhangyan Wu, Lieven Sterck, Mine Turktas, Rolf Lohaus, Zhen Li, Ming Yang, Lijuan He, Tianquan Deng, Francisco Javier Escalante, Carlos Llorens, Francisco J. Roig, Iskender Parmaksiz, Ekrem Dundar, Fuliang Xie, Baohong Zhang, Arif Ipek, Serkan Uranbey, Mustafa Erayman, Emre Ilhan, Oussama Badad, Hassan Ghazal, David A. Lightfoot, Pavan Kasarla, Vincent Colantonio, Huseyin Tombuloglu, Pilar Hernandez, Nurengin Mete, Ozgur Cetin, Marc Van Montagu, Huanming Yang, Qiang Gao, Gabriel Dorado, and Yves Van de Peer. Genome of wild olive and the evolution of oil biosynthesis. *PNAS* October 31, 2017. 114 (44) E9413-E9422.
- UniProt database. Disponible en: <http://www.uniprot.org>.

- Vivek Krishnakumar, Matthew R. Hanlon, Sergio Contrino, Erik S. Ferlanti, Svetlana Karamycheva, Maria Kim, Benjamin D. Rosen, Chia-Yi Cheng, Walter Moreira, Stephen A. Mock, Joseph Stubbs, Julie M. Sullivan, Konstantinos Krampis, Jason R. Miller, Gos Micklem, Matthew Vaughn, Christopher D. Town; Araport: the Arabidopsis Information Portal, *Nucleic Acids Research*, Volume 43, Issue D1, 28 January 2015, Pages D1003–D1009
- Wagner, D., Sablowski, R.W.M. and Meyerowitz, E.M. (1999) Transcriptional activation of *APETALA1* by *LEAFY*. *Science*, 285, 582–584.
- Weigel, D. and Meyerowitz, E.M. (1994) The ABCs of floral homeotic genes. *Cell*, 78, 203–209.
- Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F. and Meyerowitz, E.M. (1992) *LEAFY* controls floral meristem identity in *Arabidopsis*. *Cell*, 69, 843– 859.
- Weiwei Deng, Hua Ying, Chris A. Helliwell, Jennifer M. Taylor, W. James Peacock, and Elizabeth S. Dennis. *FLOWERING LOCUS C (FLC)* regulates development pathways throughout the life cycle of *Arabidopsis*. *PNAS* April 19, 2011. 108 (16) 6680-6685.
- Wigge, P.A., Kim, M.C., Jaeger, K.E., Busch, W., Schmid, M., Lohmann, J.U. and Weigel, D. (2005) Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science*, 309, 1056–1059.
- Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A. and Meyerowitz, E.M. (1990) The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature*, 346, 35–39.
- You L. 2004. Toward Computational Systems Biology. *Cell Biochemistry and Biophysics*. *Cell Biochemistry and Biophysics*. 40: 167-184.
- Zhao, L., Kim, Y., Dinh, T.T. and Chen, X. (2007) *miR172* regulates stem cell fate and defines the inner boundary of *APETALA3* and *PISTILLATA* expression domain in *Arabidopsis* floral meristems. *Plant J.* 51, 840–849.

