

**UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA**

FACULTAD DE AGRONOMÍA



**“GENÓMICA COMPARATIVA DE LAS RUTAS DE FLORACIÓN
EN FABÁCEAS DE INTERÉS ECONÓMICO Y SU USO
EN EL MEJORAMIENTO GENÉTICO”**

TESIS PARA OPTAR EL TÍTULO DE:

INGENIERO AGRÓNOMO

JONATHAN VÁSQUEZ REGALADO

LIMA – PERÚ

2021

UNIVERSIDAD NACIONAL AGRARIA LA MOLINA
FACULTAD DE AGRONOMÍA

**“GENÓMICA COMPARATIVA DE LAS RUTAS DE FLORACIÓN
EN FABÁCEAS DE INTERÉS ECONÓMICO Y SU USO
EN EL MEJORAMIENTO GENÉTICO.”**

JONATHAN VÁSQUEZ REGALADO

**Tesis para optar el Título de:
INGENIERO AGRÓNOMO**

Sustentada y Aprobada ante el siguiente jurado:

.....
Dr. Erick Espinoza Nuñez
PRESIDENTE

.....
Dr. Jorge Jiménez Davalos
ASESOR

.....
PhD. Gastón Zolla Benites
CO-ASESOR

.....
M. Jorge Tobaru Hamada
MIEMBRO

.....
Dr. Ricardo Borjas Ventura
MIEMBRO

LIMA -PERÚ
2021

DEDICATORIA

Les dedico el presente trabajo a mis padres, Oscar Vasquez Aguilar y Mercedes Juana Regalado Milla, porque con su entrega y sacrificio por un mejor futuro, me demostraron que las metas se construyen con trabajo duro y que la familia siempre será lo mas importante.

A mis hermanos Elvis, Bruno y Marcel, por los momentos de goce y algarabía que me han permitido percibir los formas y matices que dibujan la felicidad en el mundo.

A mi pareja Melissa Quispe Poma “La Señora Mapache”, por demostrarme que la vida es una lucha constante y que los sueños solo se convierten en felicidad cuando los compartes con las personas que mas amas.

AGRADECIMIENTOS

Al Dr. Gaston Zolla Benites, por su estricta guía y sincera amistad, por sus lecciones académicas y no académicas, que han construido caminos hacia nuevos y mejores horizontes donde todo conocimiento es posible.

Al Dr. Jorge Jimenez Davalos, por su amistad, confianza y profesionalismo, que me han demostrado que no existe labor mas loable y satisfactoria que ser un profesor.

Al Mg. Sc. Jorge Tobaru, por su predisposición, conocimiento y perspectivas, aportando una mejor estructura, diseño y enfoque en la investigación.

Dr. Erick Espinoza y Dr. Ricardo Borjas, por sus recomendaciones y correcciones en la redacción de esta investigación.

ÍNDICE GENERAL

I. INTRODUCCIÓN	1
II. REVISIÓN DE LITERATURA	4
2.1. El cambio climático y sus consecuencias en el mundo	4
2.1.2. Efecto del cambio climático en los cultivos	5
2.1.3. Las Fabáceas y su capacidad de adaptación frente al cambio climático	6
2.1.4. Área cultivada, producción y potencial nutricional del cultivo de fabáceas	7
2.2. Fisiología y genética de la floración en <i>A. thaliana</i>	9
2.2.1. Importancia de los genes de floración en el proceso de mejoramiento genético	13
2.3. Arabidopsis y su papel en el mejoramiento genético de leguminosas.....	15
2.4. La genómica comparativa y la floración de las plantas	15
2.5. Homología de genes	15
2.5.1. La poliploidía vegetal y los parálogos	17
2.5.2. Filogenética molecular en la diversidad de especies y mejoramiento genético .	19
2.5.3. BLAST y la búsqueda de secuencias similares entre especies comparadas	20
2.5.4. Métodos para la identificación de genes ortólogos.....	21
2.5.5. ProteinOrtho y su aplicación en el análisis de ortólogos	23
2.5.6. Herramientas para la construcción de árboles filogenéticos.....	24
2.5.7. El programa MEGA en el análisis filogenético	25
2.6. Genómica funcional.....	25
III. MATERIALES Y MÉTODOS	26
3.1. Obtención de los genes involucrados en las rutas de floración en Arabidopsis thaliana.....	26
3.2. Obtención de los proteomas de las leguminosas con importancia económica ...	26
3.3. Anotación de las secuencias de proteínas	26
3.4. Detección de los Co-ortólogos.....	27
3.5. Análisis filogenético	27
IV. RESULTADOS Y DISCUSIÓN	28
4.1. Genes de secuencias de floración similares de <i>A. thaliana</i> en las 13 leguminosas de interés económico mediante el programa BLAST.....	28
4.2. Genes ortólogos únicos de Arabidopsis thaliana en 13 leguminosas de interés económico en las rutas de floración significativas	31

4.2.1. Ruta de procesos generales y procesos autónomos (GPAP)	31
4.2.2. Ruta de desarrollo floral e identidad del meristemo (FDMI)	34
4.2.3. Ruta de vernalización (VP).....	36
4.3. Interacción de los genes ortólogos únicos de floración en las 13 leguminosas de interés económico para estudios en mejoramiento genético.....	41
V. CONCLUSIONES	44
VI. RECOMENDACIONES	44
VII. BIBLIOGRAFÍA	46
VIII. ANEXOS.....	71

ÍNDICE DE TABLAS

Tabla 1. Colecciones de germoplasma cultivadas, herbáceas y silvestres de frijoles, garbanzos, caupí, habas, garbanzos, lentejas, guisantes y gandules conservadas en los bancos de germoplasma del CGIAR (consultative group on international agricultural research).....	7
Tabla 2. Contenido nutricional de legumbres secas en 100 g de granos	8
Tabla 3. Contenido de macronutrientes y micronutrientes en legumbres secas en 100 g de grano	9
Tabla 4. Clasificación de los genes homólogos.....	17

ÍNDICE DE FIGURAS

Figura 1. Evolución de las emisiones de CO ₂ por parte de combustibles fósiles en el ambiente en gigatoneladas de carbono y las emisiones per cápita. Fuente:hardy, (2003).....	4
Figura 2. Modelo de cultivos cuadrículados globales (ggmc) y el cuarto informe de evaluación del grupo intergubernamental de expertos sobre el cambio climático (ipcc ar4) sobre el cambio medio relativo del rendimiento de los cultivos.	6
Figura 3. Área y producción mundial de legumbres secas en el mundo.	8
Figura 4. Red del tiempo de floración y las rutas involucradas en <i>A. thaliana</i>	9
Figura 5. Ruta de fotoperiodo.	11
Figura 6. Proceso de poliploidización especies <i>glycine max</i> y <i>medicago truncatula</i> poliploidización en el genoma de (a) <i>glycine max</i> y (b) <i>medicago truncatula</i>	18
Figura 7. Cladograma de la familia fabácea. Fuente: smykal <i>et al.</i> , (2015).....	20
Figura 8. Comparación de poder computacional en el análisis de ortología.....	24
Figura 9. Genes involucrados en la floración de leguminosas.	29
Figura 10. Análisis filogenético de los genes ortólogos en la ruta autónoma y procesos generales.	32
Figura 11. Análisis filogenético de los genes ortólogos en la ruta de desarrollo floral e identidad del meristemo.	35
Figura 12. Análisis filogenético de los genes ortólogos en la ruta de vernalización – proteínas represoras de <i>flc</i>	37
Figura 13. Análisis filogenético de los genes ortólogos en la ruta de vernalización – proteínas promotoras de <i>flc</i>	39
Figura 14. Modelo en base a ortólogos únicos en las rutas de floración con relevancia agronómica en 13 leguminosas con importancia económica.	43

ÍNDICE DE ANEXOS

Anexo 1. Variabilidad de los parámetros de cobertura e identidad de proteínas homologas a la ruta de floración.....	71
Anexo 2. Variabilidad de los parámetros de bit score e e-value de proteínas homologas a la ruta de floración.....	72
Anexo 3. Genes ortólogos únicos de la floración en 13 leguminosas de interés económico.....	73
Anexo 4. Secuencias de referencia para el Anexo 3.....	74
Anexo 5. Genes de floración de <i>A. thaliana</i> en sus rutas respectivas.	83

RESUMEN

La secuenciación de nueva generación (NGS) ha permitido reconocer distintos tipos de secuencias de nucleótidos y aminoácidos en diversas especies a una escala jamás imaginada. *Arabidopsis thaliana* ha sido muy útil para anotar genes en otras especies vegetales a través de la genómica comparativa, herramienta diseñada para la identificación de genes ortólogos, basado en la conservación de sus funciones biológicas. Sin embargo, no existe una anotación completa a nivel genético en diversos procesos fisiológicos como la floración. En leguminosas, la respuesta de vernalización en *Pisum sativum* y las rutas de floración han sido parcialmente descritas como el caso de *Glycine max*, *Medicago truncatula* y *Lotus japonicus*. En tal sentido, el presente trabajo tuvo como objetivo la identificación de los genes involucrados en las rutas de floración (AP: Ruta de envejecimiento, PP: Ruta de fotoperiodo, CCP: Ruta del reloj circadiano, FDMI: Desarrollo floral e identidad del meristemo, TP: Ruta de temperatura, HP: Ruta hormonal, SP: Ruta de azúcar, VP: Ruta de vernalización y GIF: Genes relacionados en floración) de 13 leguminosas de interés económico usando como referencia a *Arabidopsis thaliana*. El análisis permitió identificar 316 genes homólogos para 13 leguminosas distribuidas en 10 rutas de floración, siendo 216 genes comunes entre todas las especies estudiadas, sin embargo 135 genes en promedio no presentaron homólogos para floración. Posteriormente, se logró identificar a 21 genes ortólogos únicos, los cuales no solo regulan el proceso de floración, sino también están involucrados en la producción agrícola, tolerancia/resistencia al estrés biótico y abiótico. Por otro lado, solo 11 de estos genes ortólogos únicos presentaron coherencia taxonómica en los árboles filogenéticos construidos. La identificación de estos genes abre una ventana de oportunidades en el mejoramiento genético de leguminosas, con la finalidad de enfrentar al cambio climático y garantizar la seguridad alimentaria a través de un alimento nutritivo y accesible a todos los estratos económicos.

Palabras clave: Genómica comparativa, genes homólogos, genes ortólogos, mejoramiento de plantas, floración.

ABSTRACT

On a scale never imagined before, next-generation sequencing (NGS) allows to identify different types of nucleotide and amino acid sequences in several species. *Arabidopsis thaliana* has been very useful to annotate genes in other plant species through comparative genomics, a tool designed to identify orthologous genes, based on the conservation of their biological functions. However, there is no complete annotation at the genetic level in various physiological processes such as flowering. In legumes, the vernalization response in *Pisum sativum* and the flowering pathway have been partially described in *Glycine max*, *Medicago truncatula* and *Lotus japonicus*. In this sense, the present work aimed to identify the genes involved in the flowering pathways (AP: Aging Pathway, PP: Photoperiod pathway, CCP: Circadian clock pathway, FDMI: Flower development and meristem identity, TP: Ambient temperature pathway, HP: Hormone pathway, SP: Sugar pathway, VP: Vernalization pathway, GIF: other genes involve in flowering) of 13 legumes of economic interest using *Arabidopsis thaliana* as a reference. The analysis allowed to identify 316 homologous genes for 13 legumes distributed in 10 flowering pathways with 216 genes common among all species, but 135 genes on average did not have homologues for flowering. Moreover, it was possible to identify 21 unique orthologous genes, which are involved in flowering process as well as agriculture production, tolerance/resistance to biotic and abiotic stress. On the other hand, only 11 of these unique orthologous genes showed taxonomic coherence in the constructed phylogenetic trees. The identification of these genes opens a window of opportunities in the genetic improvement of legumes, in order to guarantee food security through accessible and healthy food to all socioeconomic strata in face of climate change.

Key words: Comparative genomic, homologous genes, orthologous genes, plant breeding, flowering

I. INTRODUCCIÓN

El aumento del rendimiento en los cultivos es uno de los principales retos de la ciencia frente al cambio climático, debido a las condiciones edafoclimáticas adversas y al aumento de enfermedades vegetales, en un contexto de alto crecimiento demográfico y la disminución progresiva de los rendimientos de cultivos en el mundo (Jimenez-Lopez *et al.*, 2020).

La familia Fabácea es la segunda de mayor importancia para la humanidad, ocupando entre el 12 – 15 % de la superficie arable del mundo, correspondiendo al 27% de la producción primaria de cultivos (Graham y Vance, 2003). La adaptabilidad de sus especies y sus características intrínsecas como simbiosis con bacterias nitrificantes, captación de CO₂, producción de granos y forraje para la alimentación humana y animal (Jensen *et al.*, 2012) y su bajo consumo hídrico frente a otros productos como las carnes avícolas y ganaderas (FIA, 2017) destacan su importancia en el proceso de cambio climático. Además, su alto contenido de proteínas, minerales y vitaminas lo convierten en un alimento balanceado de fácil acceso a múltiples familias de bajos ingresos y contribuye a una adecuada alimentación en lugares como América Latina y el Caribe, salvaguardando su seguridad alimentaria (FAO, 2016). Además son consideradas productos de alta calidad alimenticia, destacando su contribución de 33% nitrógeno proteico en la dieta de humana (Graham y Vance, 2003). Los precios bajos facilitan su ingreso en la canasta familiar de millones de personas en condición de pobreza, contribuyendo con una alimentación sana (FAO, 2016) en el Perú y el mundo, con tendencia a un gran potencial de exportación.

La demanda de granos en el Perú se ha incrementado y ha tenido un crecimiento de su exportación, cuyo registro se muestra en el 2018 con la cifra record de US\$ 53 000 000.00, beneficiando a más 140 000 familias de agricultores (MINAGRI, 2020). Según Leon, (2021), el Viceministro del MIDAGRI mencionó que la agricultura familiar reporta el 82% de sus tierras están dedicadas a la producción de leguminosas, convirtiéndose en un garante de la seguridad alimentaria del país.

En todos los cultivos de grano, la producción es dependiente de la floración de la planta, que es regulada principalmente por genes (Jung y Müller, 2009). Identificar los genes ligados a la floración en leguminosas es vital para la adaptación de sus cultivares en un contexto de cambio climático. Por ejemplo, en cultivos como el naranjo trifoliado (*Poncirus trifoliata* L. Raf) se logró disminuir el tiempo de transición de la etapa de vegetativa a reproductiva, con cambios morfológicos de floración y fructificación temprana, gracias a los genes 35S::LFY y 35S::AP1 de *Arabidopsis thaliana* (Endo *et al.*, 2005) y en manzano (*Malus domestica* Borkh.) donde se usó el gen BpMADS4 de abedul plateado (*Betula pendula* Roth), obteniendo transgénicos de corta etapa juvenil con una fructificación precoz, sin afectar la viabilidad del polen o la germinación del tubo polínico (Flachowsky *et al.*, 2007), y permiten conocer la importancia de la identificación de los genes de floración de las especies vegetales para su manipulación en el proceso de mejoramiento genético, que en las leguminosas, permitiría la obtención de cultivares con mayor producción y calidad frente a condiciones edafoclimáticas adversas.

Sin embargo, es necesario entender de manera adecuada las variables fisiológicas y genéticas que permiten mejorar la producción de granos, como es el mecanismo del proceso de floración. El entendimiento del mecanismo genético y ecofisiológico de la floración es clave en los sistemas de producción agrícola en un contexto de cambio climático (Nelson *et al.*, 2010), con la finalidad de sincronizar y afinar el uso de recursos y el manejo agronómico del cultivo.

El estudio del genoma vegetal, a partir de *A. thaliana* ha permitido dilucidar el papel que cumplen múltiples genes en los procesos y ciclos biológicos de la planta, pudiendo ser comparadas a nivel genómico con plantas cultivadas, estableciendo la oportunidad de mejoramiento de los cultivos al identificar genes homólogos (Kaul *et al.*, 2000).

La tesis presentada tiene como objetivo realizar un análisis genómico comparativo de los genes involucrados en la floración de *Arabidopsis thaliana* con los genes de las leguminosas de importancia económica como son *Arachis duranensis*, *Arachis hypogaea*, *Arachis ipaensis*, *Cajanus cajan*, *Cicer arietinum*, *Glycine max*, *Lupinus angustifolius*, *Glycine max subsp. soja* (Siebold y Zucc.) H. Ohashi (Basónimo: *Glycine soja* Siebold y Zucc), *Vigna angularis*, *Vigna radiata* y *Vigna unguiculata* con la finalidad de identificar sus genes homólogos, que pueden poseer funciones similares en las plantas contrastadas.

OBJETIVO GENERAL

- Identificar genes homólogos en leguminosas de importancia económica para su uso en los programas de mejoramiento genético.

OBJETIVOS ESPECÍFICOS

- Seleccionar los genes homólogos de las rutas de floración de Arabidopsis en las leguminosas de importancia económica en función de los parámetros del software BLAST para su análisis de ortología en ProteinOrtho5.
- Relacionar los genes ortólogos únicos de las rutas de floración de Arabidopsis en las leguminosas estudiadas con características agronómicas de producción y capacidad de respuesta al estrés biótico y abiótico.
- Evaluar la coherencia filogenética de los genes ortólogos únicos con la taxonomía de las Fabáceas.

II. REVISIÓN DE LITERATURA

2.1. El cambio climático y sus consecuencias en el mundo

El cambio climático es uno de los principales problemas en el mundo moderno, producto de las diversas actividades antropogénicas desarrolladas por el hombre como la extracción y quema de carbono fosilizado (carbón, petróleo y gas) para combustible, tala y quema de bosques, cultivo de arroz en humedales, cría de ganado, vertederos de desechos sólidos y fertilización con nitrógeno de la agricultura entre otros (Hardy, 2003), cuyo impacto se observa en la acumulación constante de CO₂ en la atmósfera (Figura 1), con consecuencias sobre todos los sistemas ecológicos de la tierra (McCarty, 2001). Esto ha provocado problemas diversos como sequías, ondas de calor, lluvias torrenciales e inundaciones, con efectos adversos sobre los ecosistemas (Jentsch y Beierkuhnlein, 2008) y en actividades económicas como la agricultura, la pesca, ganadería, etc. Los estudios sobre el futuro que nos depara respecto a este problema, sugiere consecuencias alarmantes sobre la biodiversidad y los peores escenarios conducen a una sexta extinción masiva sobre la tierra (Bellard *et al.*, 2012).

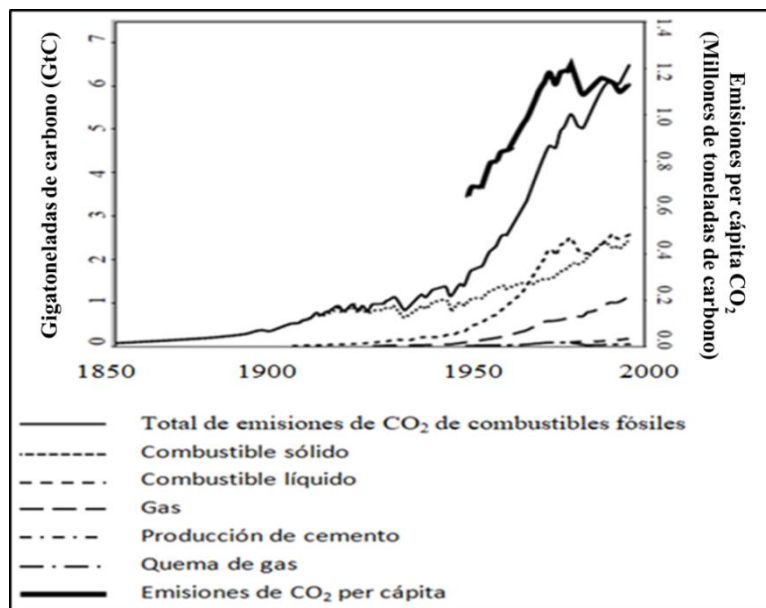


Figura 1. Evolución de las emisiones de CO₂ por parte de combustibles fósiles en el ambiente en gigatoneladas de carbono y las emisiones per cápita. Fuente: Hardy, (2003).

2.1.2. Efecto del cambio climático en los cultivos

Las tierras empleadas en la agricultura abarcan aproximadamente entre 1.2 – 1.5 millones de hectáreas, con otro 3.5 millones de hectáreas para pastos (Howden *et al.*, 2007). Estas zonas productivas conocidas como agroecosistemas son uno de los pilares de la sociedad moderna, resguardando la seguridad alimentaria para una correcta nutrición y salud en la sociedad (Jones y Ejeta, 2015). La actividad agrícola que hace uso del agroecosistema, es una labor ejecutada por el 39% de la población mundial, contribuyendo en sus ingresos económicos (Bhadouria *et al.*, 2019).

El cambio climático al tener un efecto desestabilizador sobre las condiciones medioambientales, modifican las zonas agroecológicas por efecto de la temperatura, precipitación, la tasa fotosintética por aumento de la concentración de CO₂ atmosférico, una disminución de la disponibilidad del recurso hídrico y la elevada frecuencia de eventos climáticos extremos (Kurukulasuriya y Rosenthal, 2013). Estos problemas ambientales incrementan de manera considerable la temperatura, radiación, sequía e inundaciones, provocando desordenes fisiológicos y fenológicos en los cultivos (Lobell y Gourджи, 2012) y cambios en el comportamiento, distribución y transmisión de enfermedades por insectos y patógenos de plantas (Coakley *et al.*, 1999), disminuyendo el rendimiento de los cultivos (Figura 2). Las consecuencias sobre la seguridad alimentaria complicarían la disponibilidad, estabilidad, utilización y acceso de los alimentos para la población mundial (Schmidhuber y Tubiello, 2007).

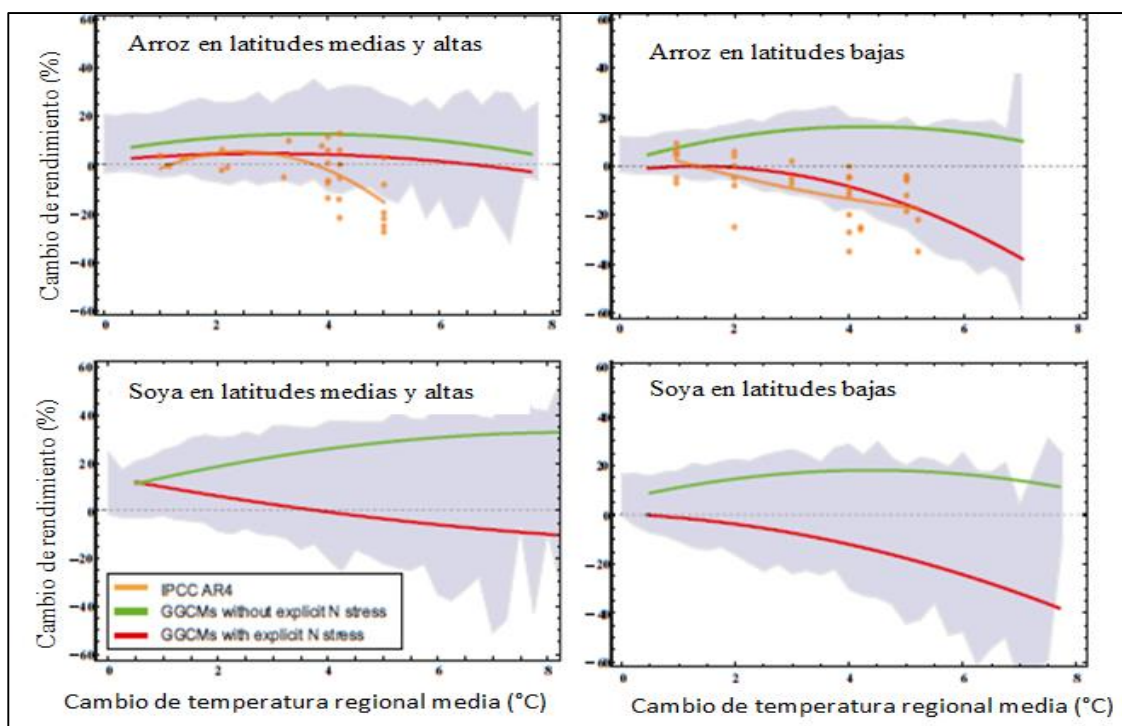


Figura 2. Modelo de cultivos cuadriculados globales (GGMC) y el Cuarto Informe de evaluación del Grupo Intergubernamental de Expertos sobre el Cambio Climático (IPCC AR4) sobre el cambio medio relativo del rendimiento de los cultivos. Fuente: Rosenzweig *et al.*, (2013).

2.1.3. Las Fabáceas y su capacidad de adaptación frente al cambio climático

El cambio climático, al tener un efecto negativo sobre el proceso de desarrollo de los cultivos y en consecuencia, sobre la calidad y rendimiento de las cosechas, hace necesaria la búsqueda de alternativas para evitar los problemas de escases de alimentos.

Bajo este contexto, se debe emplear cultivos con alta capacidad de tolerancia al estrés abiótico y biótico, adicional a una reducción de las prácticas de producción agrícola perjudiciales para el medio ambiente.

Los cultivos de leguminosas poseen características intrínsecas esenciales para su adaptación en el proceso de cambio climático, entre las que destacan : a) menor emisión de CO₂ y N₂O comparados con sistemas agrícolas con uso de fertilizantes sintéticos, b) reducción de la energía fósil usada en la producción de alimento y forraje, c) contribuye en el secuestro de carbono en suelos, d) fuente viable de biomasa para la generación de biocombustibles (Jensen *et al.*, 2012), e) bajo consumo hídrico en el manejo del cultivo (FIA, 2017) y f) su gran diversidad genética, cuyo germoplasma constituye aproximadamente el 15% de accesiones conservadas en bancos internacionales como AVRDC (Asian Vegetable Research Center), CIAT (Centro Internacional De Agricultura Tropical), ICARDA (International Center for Agricultural Research in the Dry Areas), ICRISAT (International

Crops Research Institute for Semi-Arid Tropics), IITA (International Institute for the Tropical Agriculture) (Tabla 3), sin agregar las especies silvestres (Upadhyaya *et al.*, 2011), proporcionan un mayor número de genotipos capaces de tolerar condiciones adversas en el proceso de adaptación. Estas características proporcionan a los cultivos una gran plasticidad y resiliencia frente a las adversidades provocadas por el cambio climático en el mundo, sin perjudicar la producción y calidad de las cosechas.

Tabla 1: Colecciones de germoplasma de legumbres conservadas en los bancos de germoplasma del CGIAR (Consultative Group on International Agricultural Research)

Instituto	Colección	Cultivado	Silvestre	Maleza	Estatus desconocido	Número total de accesiones
AVRDC	Vigna	15			10806	10821
CIAT	Frejol	33384	1858	715	167	36124
	Garbanzo	11988	270		1560	13818
	Haba	7316	2940	181	4949	15386
ICARDA	Guisante de hierba	506	1409	32	1368	3315
	Lenteja	6252	587		4165	11004
	Guisante	1893	212	1	3972	6078
ICRISAT	Garbanzo	19726	224		190	20140
	Gandul	13058	555		19	13632
	Caupi	13216	17	37	1727	14997
IITA	Vigna (Wild)				1522	1522
Total		107354	8072	966	30445	146837
		73.11%	5.50%	0.66%	20.73%	100.00%

Fuente: Evaluado el 27 de enero de 2011; <http://singer.cgiar.org/>

2.1.4. Área cultivada, producción y potencial nutricional del cultivo de fabáceas

La familia Fabácea es el segundo grupo más importante para la humanidad después de las gramíneas (Graham y Vance, 2003), contribuyendo con el 33% de la dieta de nitrógeno consumida por la humanidad (Vance *et al.*, 2000) razón que ha promovido un aumento del área cultivada y la producción en el tiempo (Figura 3).

Las fabáceas poseen un elevado contenido nutricional, donde la soya y los altramuces destacan en proteínas, grasas y energía (Tabla 2), así como en su contenido de macronutrientes y micronutrientes como el calcio, hierro, magnesio y selenio (Tabla 3), demostrando un gran aporte en la alimentación, reduciendo la probabilidad de contraer problemas de salud como la diabetes, obesidad y enfermedades coronarias (Calles *et al.*,

2019). Además presentan un alto grado de conservación pos cosecha con ausencia de sistemas de almacenamiento sofisticados, sin perder sus propiedades nutricionales y por sus bajos costos de producción y comercialización en los mercados es ideal en la dieta de personas de escasos recursos, garantizando la seguridad alimentaria de los sectores más vulnerables de la sociedad (FAO, 2016).

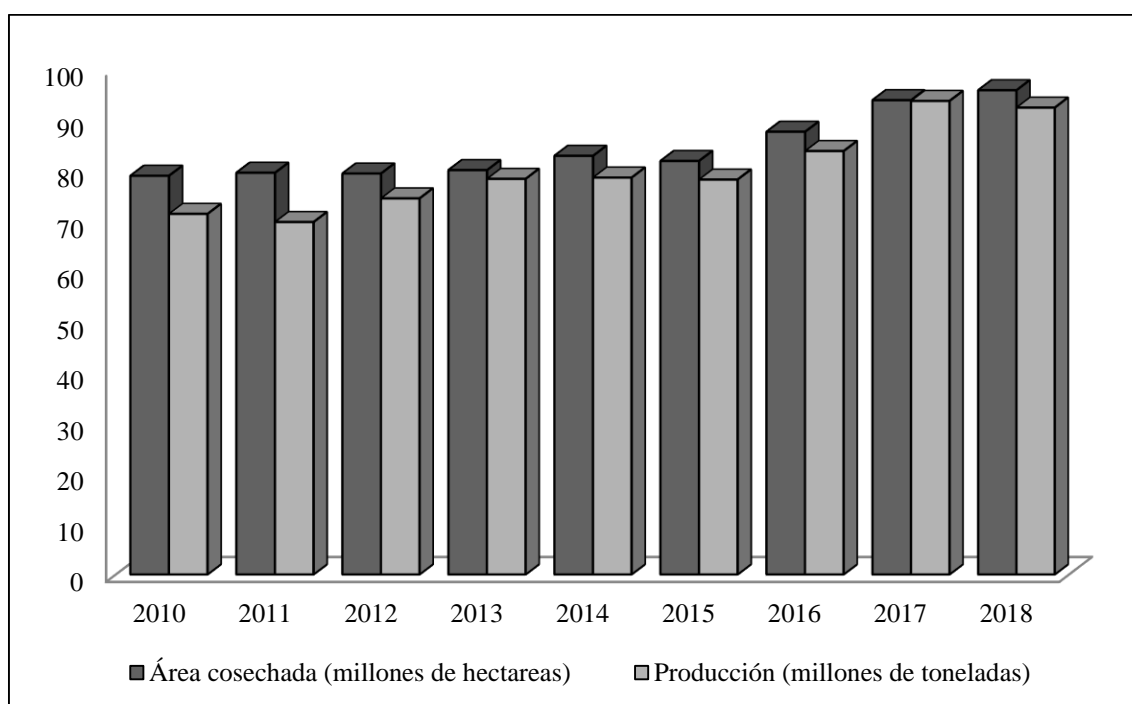


Figura 3. Área y producción mundial de legumbres secas en el mundo.

Fuente: (FAOSTAT) base de datos (<http://www.fao.org/faostat/es/#data/QC>)

Tabla 2: Contenido nutricional de legumbres secas en 100 g de granos

Nombre	Agua (g)	Energía (kcal)	Proteínas (g)	Grasas (g)	Carbohidratos (g)	Fibra (g)
Altramuces	10.4	394	36.2	9.74	40.4	-
Frijol	1.7	364	22.7	1.6	55.6	18.4
Garbanzos	5.6	373	19.4	5	55	15
Guisantes	3.4	365	21.6	2.3	56	16.7
Habas	0	372	23	2	56	19
Soja	14	406	35.9	18.6	15.8	15.7

Fuente: Adaptado de Perales *et al.*, (2017)

Tabla 3: Contenido de macronutrientes y micronutrientes en legumbres secas en 100 g de grano

Nombre	Ca (mg)	Fe (mg)	Mg (mg)	Zn (mg)	Na (mg)	K (mg)	P (mg)	Se (µg)
Altramuces	176	4.36	198	4.75	15	1013	440	8.2
Frijol	134	7.1	138	2.6	8	1464	415	-
Garbanzos	145	6.7	160	0.8	26	797	375	2
Guisantes	72	5.3	123	3.5	38	990	300	1
Habas	115	8.5	140	3.5	3.5	1030	590	8.2
Soja	240	9.7	250	4.3	5	1730	660	14

Fuente: Adaptado de Perales *et al.*, (2017)

2.2. Fisiología y genética de la floración en *A. thaliana*

La planta modelo *A. thaliana* ha permitido conocer la floración, proceso vital en el desarrollo de las plantas pues genera diversidad genética para la adaptación de la especie y su continuidad en el tiempo. Los procesos genéticos y fisiológicos presentes en esta etapa de desarrollo, así como las rutas involucradas se presentan en la Figura 4.

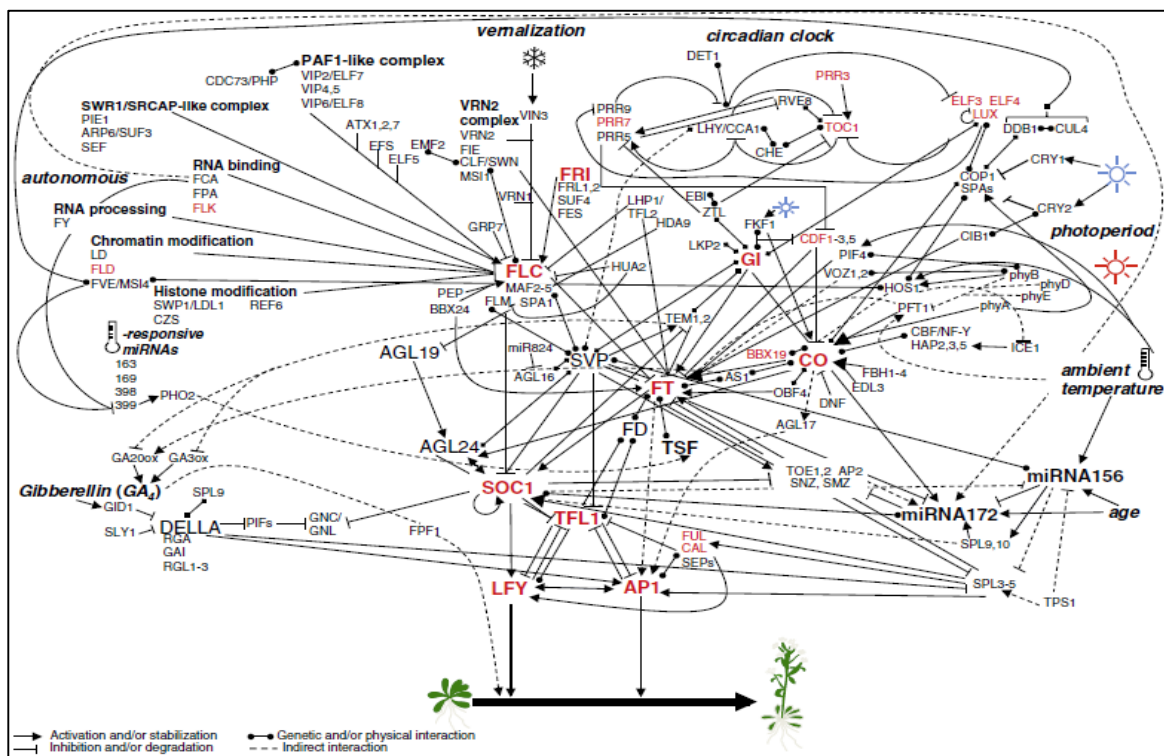


Figura 4. Red del tiempo de floración y las rutas involucradas en *A. thaliana*. Fuente: Blumel *et al.*, (2015).

El proceso de transición floral inicia con estímulos externos, señales percibidas por las hojas como los cambios de luz, temperatura, tamaño y arquitectura de la planta e internos como los cambios en la concentración de reservas de azúcares (Tremblay y Colasanti, 2006). Estas

señales son dirigidas a los meristemos de los brotes apicales (SAM) de la planta donde participan los genes *PENNYWISE (PNY)*, *BELI-LIKE (BELL)* y *POUND-FOOLISH (PNF)*, modificando la señalización genética del brote y la naturaleza celular. La menor expresión de estos genes mantiene al SAM en estado vegetativo (Ung *et al.*, 2011). El desarrollo de la floración involucra las siguientes rutas: fotoperiodo, ritmo circadiano, envejecimiento, autónoma, hormonal, azúcar, temperatura ambiente, vernalización y desarrollo floral (Bouche *et al.*, 2016).

Ruta del fotoperiodo

Los cambios de duración entre el día y la noche en función de un ciclo de 24 horas se comprende como fotoperiodo. Las proteínas fotorreceptores como *PHYTOCHROME A (PHYA)*, *PHYTOCHROME B (PHYB)*, *PHYTOCHROME C (PHYC)*, *PHYTOCHROME D (PHYD)* y *PHYTOCHROME E (PHYE)* que perciben la luz roja y roja lejana y los *CRYPTOCHROME 1 (CRY1)* y *CRYPTOCHROME 2 (CRY2)* para la luz azul y UV-A, mientras la calidad de la luz que ingresa es regulada por el gen *PHYTOCHROME AND FLOWERING TIME 1 (PFT1)*, son quienes perciben los cambios de fotoperiodo.

Ruta del reloj circadiano

Los cambios identificados por la planta como la luz y oscuridad, así como la temperatura obedece al ritmo circadiano, mecanismo interno de la planta que se mantiene inalterable frente a diversas condiciones externas para regular correctamente los cambios metabólicos y fisiológicos de la planta. Los genes que participan del reloj circadiano cuentan con tres bucles de retroalimentación interconectados: el primero con los genes *CIRCADIAN AND CLOCK ASSOCIATED 1 (CCA1)* y *LATE ELONGATED HYPOCOTYL (LHY)*, el primer ciclo con *TIMING OF CAB EXPRESSION 1 (TOC1)* y sus pseudoreguladores de respuesta (PRR), mientras sus parálogos PRR5, PRR7 y PRR9 para el segundo ciclo y finalmente el tercer ciclo por un factor de transcripción *LUX ARRHYTHMO (LUX)*, como parte de la programación de la planta en los periodos de transición floral (McClung, 2006). Adicionalmente, se ha demostrado que el reloj circadiano regula la expresión de *GIGANTEA (GI)*, quien a su vez regula y estabiliza el gen *CONSTANTS (CO)* en la planta (Hwang *et al.*, 2019), factor de transcripción que se expresa en condiciones de luz y se degrada bajo oscuridad (Figura 5).

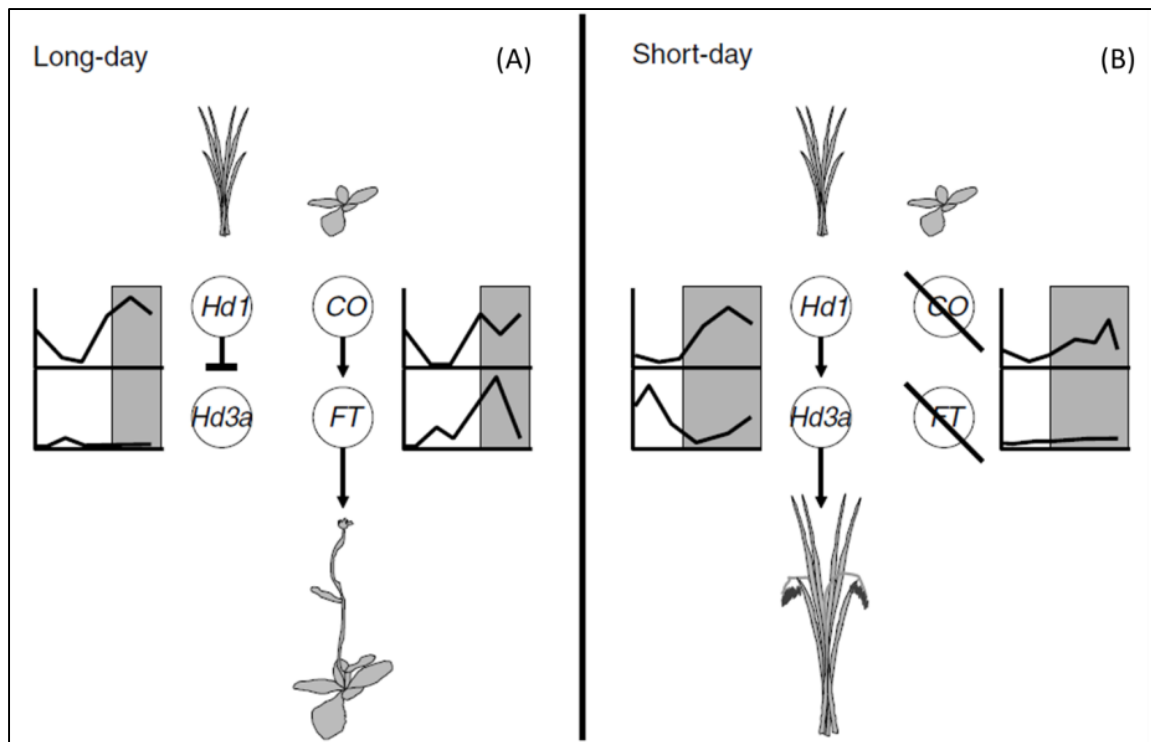


Figura 5. Ruta de fotoperiodo. (A) Proceso de floración de *Arabidopsis*, donde se observa la influencia del fotoperiodo largo en la concentración de la FT (FLOWERING LOCUS T) y CO (CONSTANS), al aumentar la concentración de FT y disminuir la de CO, logrando la floración de la planta. (B) Proceso de floración del arroz, donde se observa la influencia del fotoperiodo corto en la concentración de la de Hd1 (ortólogos de CO) y disminuir la de Hd3a (ortólogo de FT), logrando la floración de la planta. Fuente: Tremblay y Colasanti, (2006).

La generación del estímulo floral el fotoperiodo viaja desde las hojas hasta el SAM. El factor de transcripción *CO* tiene como objetivo la regulación downstream de los genes *FLOWERING LOCUS T (FT)* y *SUPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1)*. El *FT* y *SOC1* a su vez son promotores de la floración, regulando de manera positiva el factor de transcripción *LEAFY (LFY)*, quien finalmente regula los genes *CAULIFLOWER (CAL)* y *APETALA 1 (API)* relacionados a la identidad del meristemo floral (William *et al.*, 2004).

Ruta de vernalización

La vernalización por su parte, identifica los descensos de temperatura en un periodo de tiempo como el invierno con la finalidad de permitir la floración, esta característica es esencial en el ciclo biológico de la planta, pues identifica el momento óptimo para florecer. Los genes requeridos para la vernalización son *FRIGIDA (FRI)* y *FLOWERING LOCUS C (FLC)*, siendo la primera reguladora de los niveles de mRNA *FLC*, mientras el segundo es un supresor del proceso de floración (Geraldo *et al.*, 2009). De manera adicional, la

vernalización posee genes supresores de *FLC* para poder promover el periodo reproductivo. Los genes *VERNALIZATION 1 (VRN1)* y *VERNALIZATION 2 (VRN2)* reprimen a *FLC* bajo condiciones de vernalización, mientras al alcanzar temperaturas cálidas pierden su control. Otro inactivador de la *FLC* es *VERNALIZATION INSENSITIVE 3 (VIN3)*, el cual pierde su capacidad de regulación de la *FLC* al prolongarse los periodos de frío (Tremblay y Colasanti, 2006).

Ruta autónoma y procesos generales

La ruta autónoma a diferencia de las anteriores, logra la floración sin la necesidad de la longitud del día, es un proceso de floración independiente, el cual involucra a los siguientes represores de la floración según cita Cheng, J. Z. *et al.*, (2017) reprimen los genes *FLC*, *FLOWERING LOCUS CA (FCA)*, *FLOWERING LOCUS D (FLD)*, *FLOWERING LOCUS KH DOMAIN (FLK)*, *FLOWERING LOCUS PA (FLP)*, *FLOWERING LOCUS VE (FVE)*, *FLOWERING LOCUS Y (FLY)* y *LUMINIDEPENDENS (LD)*.

Ruta del ácido giberélico

La ruta hormonal se caracteriza por el ácido giberélico (GA), pues su presencia ha demostrado promover el proceso de floración. Sin embargo, experimentos posteriores esclarecieron que la generación de la señal en la ruta de giberélico y no la hormona en sí, es el promotor de la inducción floral. El GA afecta dos represores florales el *GIBBERELLIC ACID INSENSITIVE (GAI)* y *REPRESSOR OF GAI-3 (RGA)*, ambos con un efecto represor sobre la *SOCI* y los microRNA como el microRNA159 y los factores de transcripción MYB (myeloblastosis), donde una sobreexpresión del primero disminuye la concentración del segundo, retrasando la floración (Tremblay y Colasanti, 2006).

Ruta de temperatura ambiental

La ruta de temperatura ambiental influye de manera significativa en el proceso de floración, diferenciándose de la vernalización por presentar un rango óptimo de temperatura para cada especie sin ser sometida al estrés (Capovilla *et al.*, 2015). En estos procesos interviene los genes *PHYTOCHROME INTERACTING FACTOR 3 (PIF3)*, *PIF4* y *PIF5* promoviendo la floración a través de la inducción de la *FT* y su parálogo *TWIN SISTER OF FT (TSF)* en respuesta a noches cálidas e independiente de la *FT* (Thines *et al.*, 2014; Fernández *et al.*, 2016).

Ruta de envejecimiento

La ruta de envejecimiento impide la floración precoz de la planta mientras no adquiera una biomasa adecuada y habilita la floración de manera independiente a las condiciones climáticas (Wei *et al.*, 2017). Los componentes claves en esta ruta son los miRNA156, que dirige a *SQUAMOSA PROMOTOR BINDING PROTEIN-LIKE (SPL)* para una regulación downstream sobre la *FT/ FD* siguiendo una ruta de floración ausente de estímulos externos como luz o temperatura (Wang, J. W. *et al.*, 2009). La reducción del contenido de miRNA, aumenta la concentración de *SPL*, induciendo la expresión de los genes integradores de la ruta floral *SOCl*, *API*, *FRUITFULL (FUL)* y *LFY* en el SAM (Wei *et al.*, 2017).

Ruta de azúcares

La ruta de azúcar es esencial para el proceso de floración, donde azúcares como glucosa y sacarosa juegan el papel de señalizadores, mediando posteriormente la regulación de la inducción floral (Wahl *et al.*, 2013). La sacarosa juega un papel resaltante en la activación de genes como *TOCl*, *GI* y *CCAI*, claves en la ruta del reloj circadiano y en consecuencia, en el proceso de floración (Knight, *et al.*, 2008).

Desarrollo floral e identidad del meristemo

Finalmente el desarrollo floral y la identidad del meristemo se deben a los genes *LFY*, *FUL*, *CAL1* y *AGAMOUS LIKE-24 (AGL24)*, quienes responden a los estímulos dados por las rutas anteriores para modificar las funciones y propiedades de las células del SAM, para obtener de esta manera el brote floral.

2.2.1. Importancia de los genes de floración en el proceso de mejoramiento genético

El cambio de los procesos reproductivos de las plantas por las modificaciones climáticas locales y globales es un tema habitual en el fitomejoramiento, con la finalidad de alterar el crecimiento y la floración de los cultivos para beneficio del mejorador de plantas (Jung y Muller, 2009).

La floración es un proceso clave dentro de la producción de granos, debido a que la variación de tamaño de los granos se produce entre la etapa de floración y el llenado de granos (Munier-Jolain y Ney, 1998). La causa es la formación del endospermo y el embrión, productos de la fecundación doble de la planta e importantes en el desarrollo del grano, donde el endospermo presenta un impacto significativo (Bremner *et al.*, 1963), controlando

el tamaño de la semilla y transfiriendo las reservas en forma de carbohidratos, lípidos y proteínas al embrión (Li y Berger, 2012). Estas características son esenciales en la agricultura, pues el tamaño y la calidad de las reservas mejoran el rendimiento y calidad de las cosechas de granos.

Algunos genes involucrados en la rendimiento y calidad de los granos en función del tamaño, inhibición, germinación, número de granos y entre otras características son *LYSINE-SPECIFIC HISTONE DEMETHYLASE 1 (LDL1)*, *LYSINE-SPECIFIC HISTONE DEMETHYLASE 2 (LDL2)*, *REPRESSOR OF GA 1 (RGA1)*, *GIBBERELLIC ACID 1 (GA1)*, *HISTONE DEACETYLASE 1 (HDA1)*, *HISTONE DEACETYLASE 6 (HDA6)*, *CENTRORADIALIS (ATC)*, *CURVY1 (CVY1)*, *APETALA 1 (AP1)*, *APETALA 2 (AP2)* y *MICRO RNA 156 (miR156)* y *MICRO RNA 159 (miR159)* entre otros (Zhao, M. *et al.*, 2015; Ashikari *et al.*, 1999; Bradford y Nonogaky, 2007; Chen *et al.*, 2010; Bi *et al.*, 2019; Gachomo *et al.*, 2014; Shah *et al.*, 2018; Ohto *et al.*, 2005; Li *et al.*, 2017). Los cuales pueden ingresar en un programa de mejoramiento genético de leguminosas.

Por otra lado, los genes *TEMPRANILLO 1 (TEM1)*, *TEMPRANILLO 2 (TEM2)*, *DICER-LIKE1 (DCL1)*, *DICER-LIKE 3 (DCL3)*, *AGAMOUS LIKE 6 (AGL6)*, *AGAMOUS LIKE 10 (AGL10)*, *DE-ETIOLATED 1 (DET1)*, *MODIFIER OF SNC1 (MOS1)* *LUX ARRHYTHMO (LUX)* y *ACTIN-RELATED PROTEIN6 (ARP6)* están involucrados en estrés abiótico como sequía, salino, temperatura, radiación y metales pesados (Fu *et al.*, 2014; Gan *et al.*, 2017; Kim *et al.*, 2012; Zhang *et al.*, 2020; Khraiwesh *et al.*, 2012; Liang *et al.*, 2014), y su empleo podría mejorar la respuesta de las plantas frente al cambio climático.

Los genes de floración involucrados en la respuesta inmune de las plantas frente a patógenos como *TGACG MOTIF-BINDING FACTOR 4 (TGA4)*, *FLOWERING LOCUS D (FLD)*, *TGACG MOTIF-BINDING FACTOR 4 (TOE)*, *UBIQUITIN DOMAIN-CONTAINING PROTEIN (UBP12)* y *UBIQUITIN DOMAIN-CONTAINING PROTEIN (UBP13)* pueden regular de manera positiva o negativa el ataque de patógenos (Büttner y Singh, 1997; Singh *et al.*, 2014; Zou *et al.*, 2018; Ewan *et al.*, 2011). Su manipulación podría mejorar la resistencia de las plantas frente al estrés biótico.

2.3. Arabidopsis y su papel en el mejoramiento genético de leguminosas

Arabidopsis thaliana es una especie ampliamente estudiada, razón que ha permitido conocer su genoma, descifrando el papel que cada gen cumple dentro de los diversos procesos biológicos de la planta, con un impacto relevante en múltiples áreas científicas como la biología, agricultura, evolución, bioinformática, química combinatoria, genómica comparativa y funcional y medicina molecular (Kaul *et al.*, 2000).

2.4. La genómica comparativa y la floración de las plantas

Estudios de comparación genómica han logrado determinar la conservación de la organización cromosómica de las especies vegetales (Moore *et al.*, 1995). Liu, D. *et al.*, (2018) compararon y entendieron los genes de especies vegetales a través de la sintenia (región de la secuencia del genoma que abarca varios genes que son ortólogos y co-arreglados con otro genoma) y ha permitido encontrar genes con funciones homólogas de una especie en otra. Esto se ve demostrado en los análisis de genómica comparativa y filogenético en cultivos como las gramíneas, brasicáceas y fabáceas, donde se identificaron zonas conservadas comunes entre las especies comparadas a nivel del genoma (Feuillet y Keller, 2002; Acarkan *et al.*, 2000; Choi *et al.*, 2004).

La floración es un proceso regulado por genes, que afecta el vigor híbrido, con una respuesta en el crecimiento y aumento de la biomasa vegetal (Ni *et al.*, 2009). El uso de la genómica comparativa para la evaluación de genes de floración, ha permitido conocer los genes asociados a este proceso, cuya evidencia se ve respaldada por los estudios comparativos entre *A. thaliana* y *G. max*, donde se logró identificar 491 genes envueltos en la floración (Jung *et al.*, 2012). Esta información permitirá una mejor selección dentro del mejoramiento genético, desarrollando cultivares con las características deseadas (Jung y Müller, 2009), en el contexto del cambio climático.

2.5. Homología de genes

La homología se define como el grado de similitud entre dos secuencias o estructuras más allá de la esperada por casualidad, siendo la explicación más simple que ambas no surgieron de forma independiente, sino de un ancestro en común (Pearson, 2013). Cabe destacar que el elevado nivel de similaridad entre dos o más secuencias no garantiza necesariamente que sean homólogos, por ello el uso de modelos de distribución de probabilidad estadística con niveles de significancia al comparar dos o más secuencias son necesarios (Pearson, 2013).

Los homólogos se dividen en ortólogos y parálogos, cuyas definiciones fueron planteadas por Fitch, (1970): “Cuando la homología es el resultado de la duplicación de genes de modo que ambas copias han descendido una al lado de la otra durante la historia de un organismo, los genes deben llamarse parálogos. Cuando la homología es el resultado de la especiación, de modo que la historia del gen refleja la historia de la especie, los genes deberían llamarse ortólogos” *Systematic Biology* 19 (2). Páginas 99-113. Sin embargo, existen genes que no pueden ser clasificados como ortólogos o parálogos por causa de algunas variantes dentro de sus conceptos originales, haciendo necesario el uso de un nuevo sistema de clasificación (Tabla 4). La correcta definición y clasificación de los homólogos tiene como objetivo la identificación de ortólogos, genes cuyo ancestro puede tener propiedades similares a uno o más predecesores, contribuyendo en la predicción de funciones bioquímicas o sus roles biológicos (Storm y Sonnhammer, 2002) y para evaluar el proceso evolutivo de las especies en función de su filogenia (Dufayard *et al.*, 2005). Esto permite el desarrollo de investigación en las áreas de genómica comparativa, filogenética, anotación de funciones de proteínas y reordenamiento del genoma (Altenhoff y Dessimoz, 2009), logrando identificar y validar, mediante genómica funcional, genes de interés en el mejoramiento genético de las especies de interés económico.

Tabla 4: Clasificación de los genes homólogos

Homólogos	Genes compartidos de un origen común
Ortólogos	Genes que se originan a partir de un solo gen ancestral en el último ancestro común de los genomas comparados.
Pseudoortólogos	Genes que en realidad son parálogos pero que parecen ser ortólogos debido a la pérdida de genes diferenciales específicos del linaje.
Xenólogos	Genes homólogos adquiridos a través de XGD (desplazamiento de genes xenólogos) por una o ambas especies comparadas, pero que parecen ser ortólogos en las comparaciones de genomas por pares.
Co-ortólogos	Dos o más genes en un linaje que son, colectivamente, ortólogos a uno o más genes en otro linaje debido a una duplicación específica del linaje. Los miembros de un conjunto de genes co-ortólogos son parálogos relativos al evento de especiación respectivo.
Parálogos	Genes relacionados por duplicación
Parálogos internos o simparálogos	Genes parálogos resultantes de una duplicación específica de linaje posterior a un evento de especiación dado (definido solo en relación con un evento de especiación, sin significado absoluto).
Parálogos externos o aloparálogos	Genes parálogos resultantes de una duplicación que precede a un evento de especiación dado (definido solo en relación con un evento de especiación, sin significado absoluto).
Pseudoparálogos	Genes homólogos que aparecen como parálogos en un análisis de un solo genoma, pero que en realidad terminan en el genoma dado como resultado de una combinación de herencia vertical y HGT (transferencia horizontal de genes).

Fuente: Koonin, (2005)

2.5.1. La poliploidía vegetal y los parálogos

La poliploidización es el aumento del genoma por dos o más juegos cromosómicos en las células somáticas, causado por fallas en la reducción meiótica, esperma u óvulo no reducido, hibridación interespecífica y adicionalmente en el caso de las plantas, por duplicación somática (Alcántar-Vázquez, 2014). Este proceso genera regiones conservadas que albergan genes parálogos que pueden perder su función o permanecer en el proceso de desarrollo del organismo, así como participar de forma parcial para mantener la función original (Veitia, 2005). Estos genes pueden permanecer en la planta y provocar errores en la detección de genes ortólogos, dando problemas en el proceso de selección. Por ello, es deseable un menor número de poliploides en la comparación de especies. Sin embargo, la frecuencia de poliploides en las plantas angiospermas o con flores es recurrente y parte de su proceso de

evolución y establecimiento sobre la tierra (Soltis *et al.*, 2009). Esto ha permitido jugar roles importantes en los procesos de diversificación de plantas, evolución de genes y la domesticación de cultivos, este último destacando en el desarrollo de rasgos con importancia agronómica (Zhang, K. *et al.*, 2019).

En la familia Fabácea se ha observado que la frecuencia de poliploides es la cuarta parte de todos los géneros, entre quienes se ubican *Arachis*, *Cajanus*, *Medicago*, *Cicer*, *Glycine*, *Lupinus*, *Phaseolus* y *Vigna* (Doyle, 2012), dado por su proceso de evolución (Figura 6).

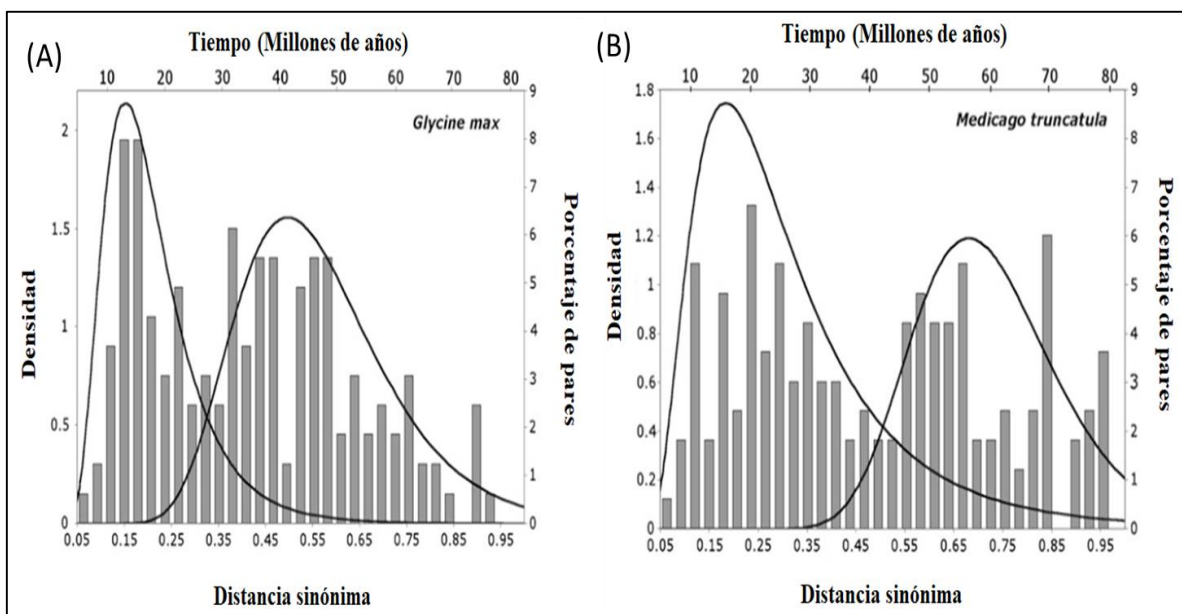


Figura 6. Proceso de poliploidización especies *Glycine max* y *Medicago truncatula*
Poliploidización en el genoma de (A) *Glycine max* y (B) *Medicago truncatula*. Las Figuras trazan el número de pares de secuencias parálogas ("densidad" o "porcentaje de pares") frente a clases agrupadas de Ks ("distancias sinónimas"). No se trazaron los pares con divergencia muy baja (producida por duplicaciones recientes en curso). Las curvas se ajustaron a los datos de divergencia agrupados y se interpretan como grupos de genes duplicados simultáneamente en eventos genómicos a gran escala como la poliploidía; los modos de picos proporcionan una edad máxima para los eventos aloploidios (Doyle y Egan 2010). El tiempo de divergencia se estimó a partir de distancias sinónimos utilizando métodos de reloj estándar; observe las diferentes edades estimadas (modos de curvas) para el evento más antiguo en las dos especies. Los datos son de Schlueter *et al.* (2004), quienes utilizaron etiquetas de secuencia expresada (EST). Fuente: Compilado de Doyle, (2012).

2.5.2. Filogenética molecular en la diversidad de especies y mejoramiento genético

La filogenética molecular es la rama que estudia la filogenia y el proceso de evolución entre las especies por análisis de secuencias de ADN y aminoácidos (Whelan et al., 2001). Para lograr su objetivo, la filogenética molecular hace uso de genes ortólogos, pues un árbol filogenético de ortólogos es por definición igual al árbol de especies que le corresponde (Altenhoff y Dessimoz, 2009). En consecuencia, la correcta identificación de genes ortólogos dará un mayor grado de certeza al árbol filogenético. Su aplicación puede ser usada para develar los procesos evolutivos en el tiempo de las especies y su diversificación por múltiples variables climáticas en la historia de las plantas (Fiz-Palacios *et al.*, 2011), donde especies que comparten un nicho ecológico no necesariamente se adaptan en todo su linaje (Donoghue, 2008). Esto explica los resultados obtenidos por Davis *et al.*, (2010), donde determinaron que especies relacionadas por su filogenia y con respuesta fenológicas compartidas a pesar de estar en zonas geográficas distintas, dan como resultado un patrón filogenético sesgado en beneficio de especies no nativas. Esta información mejora la selección de especies con objetivos específicos en el mejoramiento genético. Un ejemplo claro fue realizado por Smýkal *et al.*, (2015) sobre el origen, historia de domesticación y conservación del germoplasma (13 especies de 6 tribus) de cultivos de legumbres con elevado potencial para la agricultura y la ciencia (Figura 7). Para mejorar los sistemas de clasificación filogenética en leguminosas según The Legume Phylogeny Working Group, (2013) es necesario: (1) muestreo completo de taxón a nivel de género, especialmente en Papilionoideae, (2) aumentar el muestreo de taxones a nivel de especie para perfeccionar la delimitación genérica, (3) implementar técnicas de secuenciación de próxima generación para recopilar más datos de la secuencia de ADN, especialmente del genoma nuclear, (4) recopilar conjuntos de datos morfológicos más completos en toda la familia, (5) probar y optimizar enfoques y procedimientos de supermatriz y / o superárbol para integrar estudios de alto nivel y el número creciente de conjuntos de datos de secuencia a nivel de especie cada vez más densamente muestreados para mejorar los nodos mal resueltos en el árbol de leguminosas y (6) integrar conjuntos de datos moleculares y morfológicos.

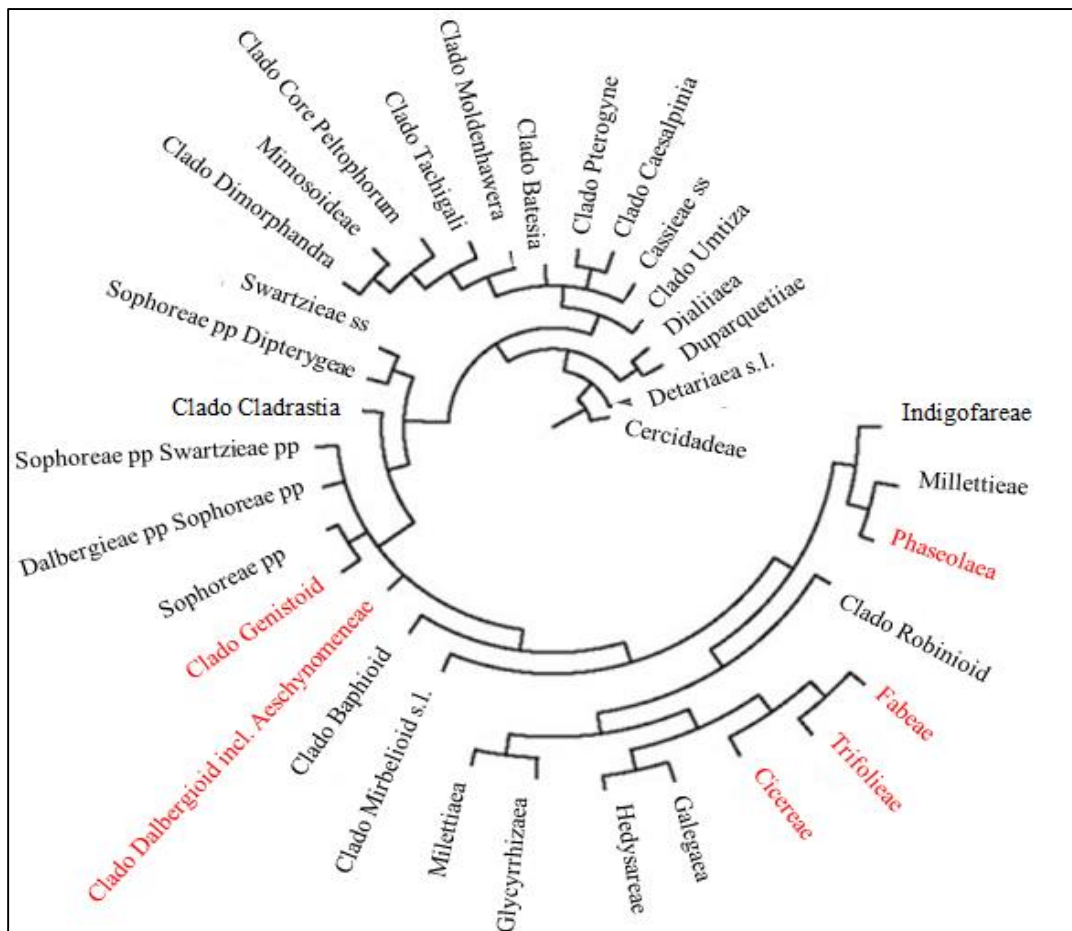


Figura 7. Cladograma de la familia Fabácea. Fuente: Smykal *et al.*, (2015).

2.5.3. BLAST y la búsqueda de secuencias similares entre especies comparadas

BLAST (Basic Local Alignment Search Tool) es un programa diseñado para identificar regiones de secuencias similares en ADN o proteínas en función a una prueba de significancia estadística (Maddan, 2013). Los análisis desarrollados por el programa están en función a una distribución aleatoria de valor extremo usando variables aleatorias independientes distribuidas de forma idéntica (Gumbel, 1958). Para poder calificar la fiabilidad de los resultados en las comparaciones de secuencias se usa un método de alta puntuación cuyo resultado se da en función de un bit score, relacionado a la calidad de la alineación o E-value que da la significancia estadística de una determinada alineación por pares reflejando el tamaño de la base de datos y la puntuación del sistema utilizado (Syngai *et al.*, 2013). Mediante el método de agrupación de proteínas en función de su grado de similitud se recomienda un E-value de corte de 1.0×10^{-5} para los análisis del programa (Tatusova *et al.*, 2014). Pese a ello, no existe consenso alguno sobre el límite del E-value para estudios de genómica comparativa (Rubin, 2000; Fulton, 2002; Ikeda *et al.*, 2003; Fei

et al., 2004; Bachhawat, 2006; Li *et al.*, 2010; Sharma y Chauhan, 2012; Xue *et al.*, 2012; Jung *et al.*, 2012; Yang *et al.*, 2016; Nowell *et al.*, 2016; Kono *et al.*, 2018). QIAGEN (2021), en su manual sugiere que un correcto E-value se debe analizar en función de la longitud de la secuencia y el tamaño de la base de datos, dando como secuencias confiables E-value $\geq 1.0 \times 10^{-50}$. Sin embargo, los métodos de análisis BLAST están limitados a un alto nivel de similitud, descartando los niveles bajos de conservación, por ello el uso de herramientas alternativas son necesarios para la detección de ortólogos (Wagner *et al.*, 2004).

2.5.4. Métodos para la identificación de genes ortólogos

Realizar la identificación de ortólogos resulta una labor complicada, debido a una serie de dificultades propias de los sistemas de evaluación, entre las que destacan diferenciar ortólogos de otros homólogos (Tabla 4), la elevada capacidad computacional para los análisis y los experimentos no prácticos, debido a la cantidad de genomas dilucidados (Galperin y Koonin, 2004). Para resolver estos problemas, múltiples programas han sido desarrollados, cada uno con sus propias características (Tabla 5). La aplicación de estos programas para la predicción de ortólogos se dividen en cuatro categorías: (1) diseñados para lograr más exactitud, (2) mayor velocidad de detección, (3) mayor flexibilidad para manipular un amplio grupo de datos y (4) métodos de integración (conductos) en análisis de ortólogos (Nichio *et al.*, 2017), cada uno con sus propias ventajas y desventajas en el proceso de evaluación (Tabla 6). El investigador priorizará el programa de acuerdo a las categorías, el acceso a banco de datos y la capacidad computacional de su laboratorio de bioinformática.

Tabla 5. Softwares para estudios de ortología.

Softwares	Principales características	Plataforma
Hieroned	Combina una metodología eficiente basada en gráficos con aspectos de métodos basados en árboles de cálculo intensivo para inferir ortología.	Linux / Unix; servidores web
MorFeus	Calcula una puntuación de red para los ortólogos de red resultantes para encontrar proteínas de ortólogos almacenadas de forma remota	Linux / Unix; servidores web
OrthAgogue	Alta velocidad de las relaciones de homología en grandes conjuntos de datos	Linux/Unix
OrthoInspector	Detección ágil de ortología e inparalogía incorporando un algoritmo único	Crossplatform (Java)
OrthoFinder	Resuelve los sesgos fundamentales en las comparaciones del genoma completo y mejora la precisión de la inferencia en los grupos de ortólogos	Linux/Unix
Ortholog-Finder	Identifica ortólogos genuinos entre especies relacionadas lejanamente mediante análisis filogenético utilizando datos ORF	Linux/Unix
Orthograph	Desarrollado para un gran conjunto de datos que mantiene la alta sensibilidad y precisión con AMBOS enfoques	Linux/Mac OS X
Orthonome	Diseñado para aumentar la precisión de múltiples especies en las predicciones de ortólogos y reducir la compensación entre las tasas de captura de ortólogos	Servidor web
OrthoVenn	Establece relaciones de agrupaciones ortólogas en varias especies mediante el diagrama de Venn.	Servidor web
PanOCT	Automatiza grupos de ortólogos para el análisis pangenómico de cepas bacterianas y especies estrechamente relacionadas	Linux/Unix
PhosphOrtholog	Desarrollado para el mapeo de ortólogos por proteínas PTM por especies cruzadas	Servidor web
PorthoDom	Desarrollado para acelerar la detección de proteínas de ortólogos utilizando secuencias de dominio	Crossplatform
PorthoMCL	Diseñado para encontrar ortólogos en una gran cantidad de genomas	Linux and Unix (OS X)
ProteinOrtho	Lidiando con cientos de especies bacterianas en un conjunto que contiene millones de proteínas usando poca memoria de computadora	Linux/Unix 64 bits
ReMark	Identifica ortólogos automáticamente mediante un ajuste de parámetros según el interés del usuario	Crossplatform (Java)
SPOCS	La predicción de ortólogos en el método gráfico basado en generar una tabla puede proporcionar una visualización de las relaciones entre los grupos de ortólogos	Servidor web; Linux/Mac OS X

Fuente: Adaptado de Nichio *et al.*,(2017)

Tabla 6. Ventajas y desventajas de los software empleados en ortología.

Softwares	Ventajas	Desventajas
OrthoVenn	Visualización utilizando el diagrama de Venn interactivo en las vistas de clústeres generados. Aporta información de gene ontology con cada función de proteína.	Solo servidor web, limitación de consultas.
PorthoMCL	Capacidad de identificación de ortología en gran número de genomas.	Aunque es una herramienta rápida y fácil, requiere los paquetes BLAST, PERL y Python.
OrthoFinder	Comando sencillo que utiliza como entrada un archivo multiFASTA (uno por especie) minimizando el sesgo de la longitud, gen previamente no detectado en el ortogrupo.	Necesita muchas dependencias, incluidos los algoritmos BLAST y MCL para ejecutarse.
Ortholog-Finder	Un programa que identifica ortólogos genuinos entre especies distantes utilizando filtros HGF para análisis filogenético	Muchas dependencias y el programa no admite los métodos de máxima verosimilitud o Bayes.
PanOCT	Usos procarióticos, relaciones ortólogos y coortólogos.	Depende de los paquetes PERL, BLAST + y se limita a un análisis de hasta 25 genomas.
ReMark	Automatiza el proceso mediante el ajuste de acuerdo con el interés del usuario.	La herramienta no está actualizada desde marzo de 2011. Necesita dependencias BLAST y JAVA.
Orthoinspector	Incorpora un algoritmo original, facilita la consulta de datos y procesa automatización.	Creación de una base de datos en Postgresql o MySQL y BLAST dependencias.
PorthoDom	Utiliza el dominio de proteínas para acelerar proteinOrtho. Utiliza la anotación Pfam para aumentar la precisión.	Es un poco laborioso de realizar, necesita muchos paquetes, base de datos Pfam y paquete HMMER, además de la herramienta ProteinOrtho.

Fuente: Adaptado de Nichio *et al.*,(2017)

2.5.5. ProteinOrtho y su aplicación en el análisis de ortólogos

Es un programa de detección de ortólogos que optimiza la velocidad en un grupo de datos extensos y provee una reducción drástica de uso de memoria (Figura 8), empleando una variante de detección de la agrupación de gráficos , ya que los conjuntos de co-ortólogos

corresponden a los subgráficos multipartitos completos máximos, implementando un enfoque basado en BLAST para determinar agrupaciones de secuencias de proteínas o genes co-ortólogos que generaliza la mejor alineación recíproca heurística (Lechner *et al.*, 2011). Un ejemplo del potencial del programa es el estudio de genómica comparativa del neem contra 23 especies vegetales, dando como resultado una muestra 5832 genes únicos con expresión en varios tejidos de la planta (Kuravadi y Gowda, 2019) y en la evolución independiente de las familias *BAX INHIBITOR (BI)* y *LIFEGUARD (LFG)* en la superfamilia *TRANSMEMBRANE BAX INHIBITOR MOTIF (TMBIM)* de las plantas, obteniendo ortólogos en 48 especies de plantas con 96 en las familias *BI* y 277 en las familias *LFG* respectivamente, usando opciones determinadas del programa (Gamboa-Tuz *et al.*, 2018).

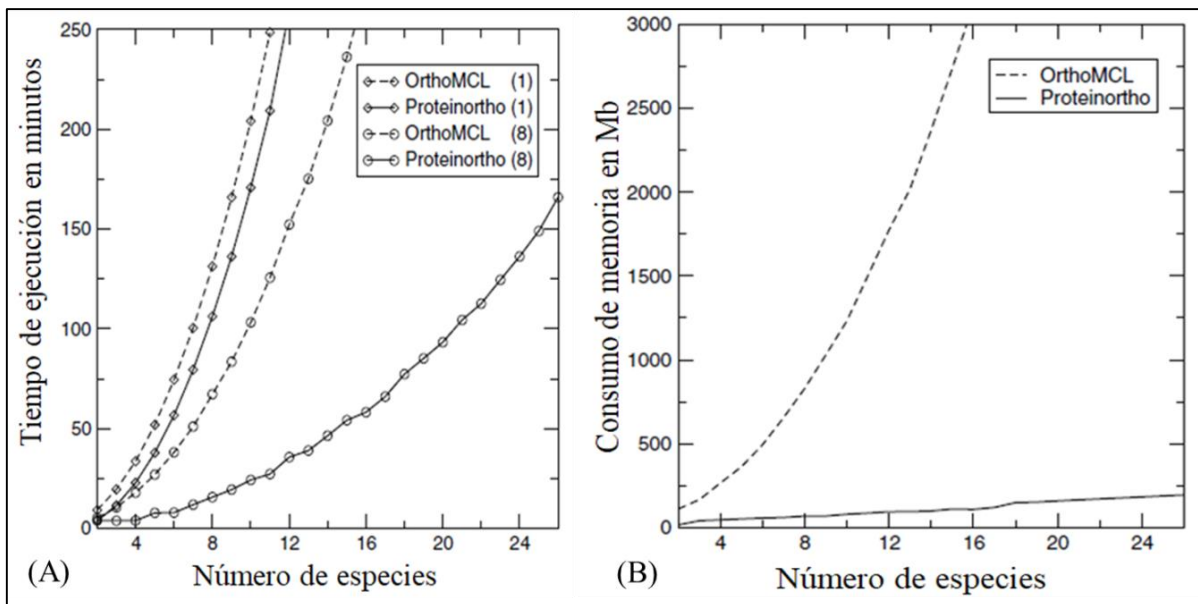


Figura 8. Comparación de poder computacional en el análisis de ortología. (A) Velocidad de análisis (B) Consumo de memoria. Fuente: Lechner *et al.*, (2011).

2.5.6. Herramientas para la construcción de árboles filogenéticos

El orden, inclusión y exclusión de especies se realizan a través de árboles filogenéticos, cuyo uso ha logrado proveer nuevas pistas sobre los sistemas de mejoramiento y la biología de la polinización (Weller y Sakai, 1999). Con la finalidad de detectar un correcto sistema de selección de árboles filogenéticos, se han empleado múltiples métodos de evaluación como Fitch-Margoliash, máxima parsimonia, máxima verosimilitud, mínima evolución y método de unión de vecinos, cada uno con sus asunciones correspondientes (Saitou y Imanishi, 1989). Estos métodos han sido incluidos en programas computacionales actuales, entre

quienes destacan por su frecuencia de uso MEGA, Dendroscope, FigTree, Phylotree.js, y ggtree, un paquete de R y se diferencian en su nivel de robustez para diversos formatos y versatilidad en la presentación y manipulación del árbol y sus ramas (Faiza, 2018).

2.5.7. El programa MEGA en el análisis filogenético

MEGA (Molecular Evolutionary Genetic Analysis) es un programa desarrollado para estimar la distancia de evolución, la reconstrucción filogenética de árboles y para calcular cantidades estadísticas básicas a partir de datos moleculares y su fácil formato de manejo lo convierten en una herramienta de uso sencillo (Kumar *et al.*, 1994). Sus aplicaciones en el reino vegetal destacan en estudios como el de Zhang, Y. B. *et al.*, (2019) donde se reconstruyó la filogenética y estimó el tiempo de divergencia de *Blumea* DC., miembro de la familia Asteraceae, donde se determinó que el periodo de máxima diferenciación fue durante el Oligoceno y Mioceno, mientras Kadam *et al.*, (2016) lograron extraer información filogenética relevante de la soja en relación al nemátodo del quiste de la soja en pos de identificar material vegetal resistente y prevenir la pérdida de rendimiento.

2.6. Genómica funcional

La genómica funcional, encargada del estudio de la función de elementos codificados como genes, ARN y proteínas y su rol en la biología de la planta (Fernandez *et al.*, 2015), ha permitido la validación funcional de genes obtenidos por genómica comparativa. Por ejemplo el estudio de las glicosiltransferasas y su respuesta al estrés biótico y abiótico en Brassicaceas, donde se identificó los mejores candidatos para un estudio UDP-glicosiltransferasas para la manipulación de las vías de estrés biótico y abiótico (Rehman *et al.*, 2018) y la identificación de 67 genes biosintéticos de carotenoides en *Brassica rapa* al comparar con *A. thaliana*, elucidando los genes involucrados en los procesos de producción de carotenoides y la posibilidad de producir mutantes o transgénicos con un contenido superior de carotenoides (Li *et al.*, 2015). Sin embargo, los transgénicos no son el único método para determinar la función de un gen. El TILLING (técnica de inducción de lesiones locales objetivo en genomas) y el VIGS (silenciamiento genético por inducción de virus) han sido empleados en estudio de desarrollo contra el estrés biótico y abiótico (Hussain, 2015) y su aplicación en el mejoramiento genético, ha permitido el desarrollo de nuevos cultivares.

III. MATERIALES Y MÉTODOS

3.1. Ubicación del experimento

Para la búsqueda, descarga y análisis de metadatos se usó la computadora de sistema operativo LINUX del Programa de Investigación y Proyección Social (PIPS) en el Departamento de Cereales y Granos Nativos de la Universidad Nacional Agraria La Molina (UNALM) y una laptop Lenovo G50-80 Intel Core i5 5200U con 8 GB RAM de sistema operativo WINDOWS de uso personal.

3.2. Obtención de los genes involucrados en las rutas de floración en *Arabidopsis thaliana*

La obtención de los genes involucrados en la floración de *A. thaliana* será desde la base de datos del Max Planck Institute (https://www.mpipz.mpg.de/14637/Arabidopsis_flowering_genes) y FLOR ID (<http://www.phytosystems.ulg.ac.be/florid/>). Las secuencias de proteínas de los genes de floración de *A. thaliana* serán obtenidas de TAIR (<https://www.arabidopsis.org/tools/bulk/sequences/index.jsp>).

3.3. Obtención de los proteomas de las leguminosas con importancia económica

Todas las secuencias de proteínas de los genomas (proteomas) de *A. duranensis*, *A. hypogaea*, *A. ipaensis*, *C. cajan*, *C. arietinum*, *G. max*, *L. angustifolius*, *Glycine max subsp. soja* (Siebold y Zucc.) *H. Ohashi*, *V. angularis*, *V. radiata* y *V. unguiculata* serán descargadas NCBI (<ftp://ftp.ncbi.nlm.nih.gov/genomes/refseq/plant/>).

3.4. Anotación de las secuencias de proteínas

La anotación de las secuencias de proteína se realizará por medio del programa BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastDocs&DOC_TYPE=Download) en la plataforma LINUX. Se anotarán las proteínas homólogas a la ruta de floración de *A. thaliana* en las leguminosas *A. duranensis*, *A. hypogaea*, *A. ipaensis*, *C. cajan*, *C. arietinum*, *G. max*, *L. angustifolius*, *G. max spp soja*, *V. angularis*, *V. radiata* y *V. unguiculata*, usando como criterio de selección un E-value $\leq 1.0 \times 10^{-50}$ y un porcentaje de cobertura e identidad $\geq 50\%$. La finalidad es la agrupación de los genes homólogos en las

siguientes rutas de floración: GPAP: Ruta autónoma y procesos generales, AP: Ruta de envejecimiento, PP: Ruta de fotoperiodo, CCP: Ruta del reloj circadiano, FDMI: Desarrollo floral e identidad del meristemo, TP: Ruta de temperatura, HP: Ruta hormonal, SP: Ruta de azúcar, VP: Ruta de vernalización y GIF: Genes relacionados en floración.

3.5. Detección de los Co-ortólogos

Los resultados de la anotación anterior serán utilizados para detectar los grupos ortólogos mediante el programa *proteinortho5* (Lechner *et al.*, 2011) con un E-value $\leq 1.0 \times 10^{-50}$. Los ortólogos únicos en cada especie fueron utilizados en la identificación de caracteres agronómicos y tolerancia/ resistencia al estrés biótico y abiótico a través de un meta análisis.

3.6. Análisis filogenético

El alineamiento múltiple de las secuencias de aminoácidos y los árboles filogenéticos serán construidos utilizando el software MEGA6 (Tamura *et al.*, 2012) basado en el método de método de máxima verosimilitud y el modelo de corrección de Poisson (Zukerkandl y Pauling, 1965) con 1000 bootstraps de soporte al modelo en cada nodo. El alineamiento múltiple y los árboles filogenéticos se basarán en las secuencias ortólogas únicas identificadas por *proteinortho5*.

IV. RESULTADOS Y DISCUSIÓN

4.1. Genes de secuencias de floración similares de *A. thaliana* en las 13 leguminosas de interés económico mediante el programa BLAST.

Según Altschul *et al.*, (1990), el algoritmo de búsqueda de BLAST contrasta las secuencias locales de proteínas y genes a través del método de máxima puntuación. El método hace uso de segmentos de secuencias máximos de carácter local (útiles para secuencias diferenciadas) con niveles de puntuación máxima, que permiten calificar a la secuencia en función a la cantidad de coincidencias de la consulta (Leguminosas) al compararla con la secuencia de referencia (*Arabidopsis*), obteniendo datos de los errores y vacíos en las secuencias de consulta, seleccionando las secuencias que tengan los valores más significativos, sin necesidad de que ambas presenten la misma longitud. La función del método es que las secuencias seleccionadas tengan un alto nivel de alineación que maximice los valores del bit score e identidad con un menor valor de divergencia (errores y vacíos de secuencia).

Esto permite distinguir significativamente secuencias similares y poder comprender su función biológica. Los resultados obtenidos por BLAST, fueron filtrados en función de los siguientes parámetros: cobertura (%) , identidad (%), E-value y el bit score. La variable cobertura, bit score y E-value, presentaron elevados niveles de dispersión, destacando el porcentaje cobertura, mientras la variable identidad y divergencia, que contempla los errores y vacíos en la secuencia de proteína, presentaron una menor cantidad de datos dispersos (Anexo 1 y 2). El criterio de selección (cobertura (%) e identidad (%) $\geq 50\%$ y un E-value $\leq 1.0 \times 10^{-50}$) permitió la identificación de genes de secuencias similares relacionados a la ruta de floración de *A. thaliana*, hubo un total de 103 genes ausentes, 236 genes eran compartidos por todas las leguminosas, 69 genes estaban distribuidas en grupos de 2 a 12 especies y 11 genes estaban presentes solo en una especie: *FERTILIZATION INDEPENDENT SEED 1 (FIS1)* y *PSEUDO-RESPONSE REGULATOR 3 (PRR3)* en *P. vulgaris*, *RADICAL INDICED CELL DEATH 1 (RCD1)* y *SIMILAR TO RCD ONE 1 (SRO1)* en *M. truncatula*, *GOLDEN 2-LIKE 1 (GLK1)* en *A. hypogea*, *SQUAMOSA PROMOTOR BINDING PROTEIN LIKE-9 (SPL9)* y *MODIFIER OF SNC 1 (MOS1)* en *C. arietinum*, *MYB DOMAIN PROTEIN 65 (MYB65)* y

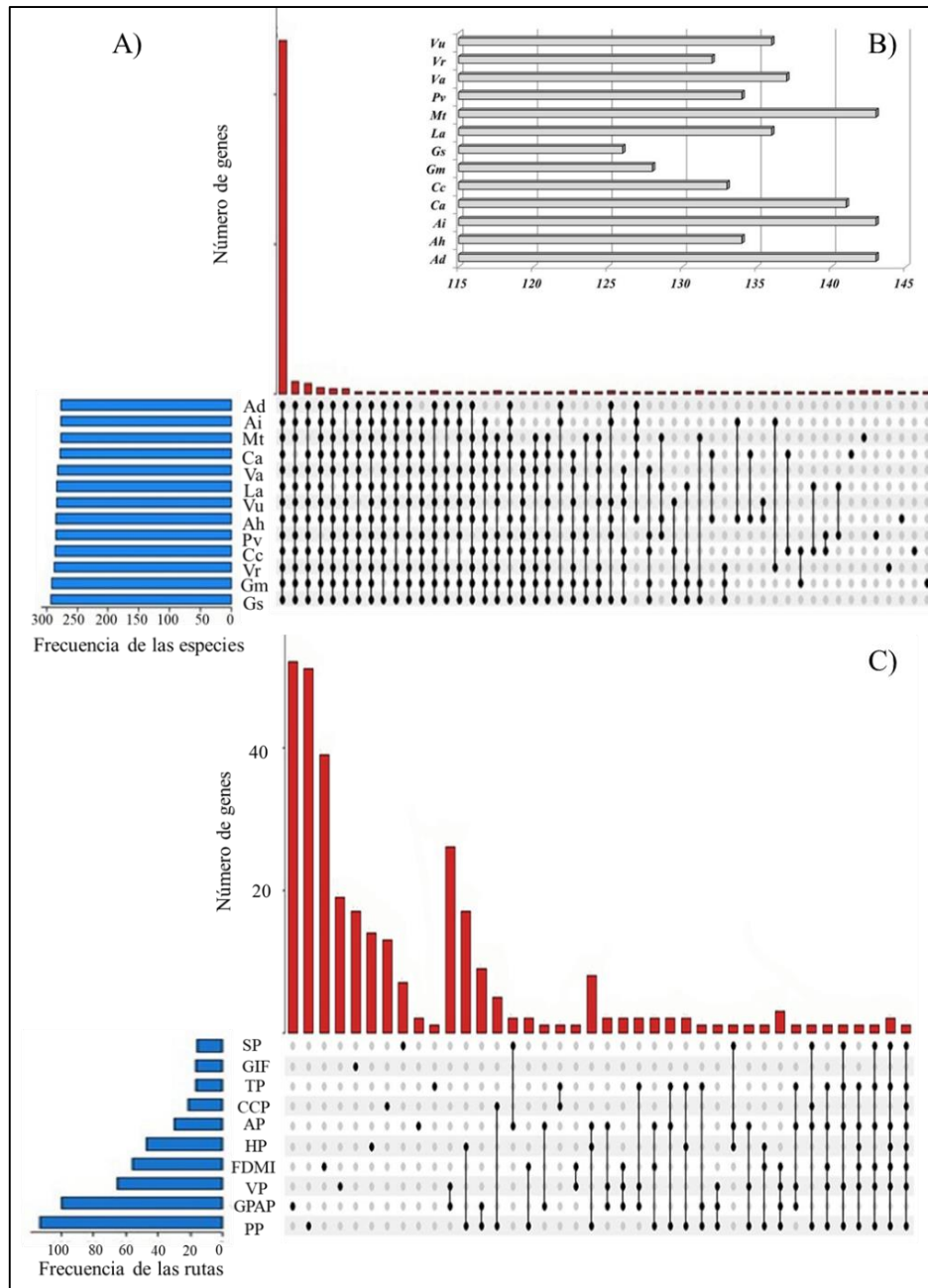


Figura 9. Genes involucrados en la floración de leguminosas. (A) Número de genes homólogos identificados en una o más leguminosas del estudio, donde se puede apreciar una elevada conservación de genes entre las leguminosas. Ad: *Arachis duranensis*, Ai: *Arachis ipaensis*, Ah: *Arachis hypogea*, Mt: *Medicago truncatula*, Pv: *Phaseolus vulgaris*, Cc: *Cicer arietinum*, Ca: *Cajanus cajan*, La: *Lupinus angustifolius*, Gs: *Glycine soja*, Gm: *Glycine max*, Vr: *Vigna radiata*, Va: *Vigna angularis*, Vu: *Vigna unguiculata*. (B) Número de genes en floración que no poseen homólogos en las leguminosas. (C) Número de genes homólogos por ruta de floración. GPAP: Ruta autónoma y procesos generales, AP: Ruta de envejecimiento, PP: Ruta de fotoperiodo, CCP: Ruta del reloj circadiano, FDMI: Desarrollo floral e identidad del meristemo, TP: Ruta de temperatura, HP: Ruta hormonal, SP: Ruta de azúcar, VP: Ruta de vernalización, GIF: Genes relacionados en floración.

MYB DOMAIN PROTEIN 33 (MYB33) en *V. radiata*, *SCHNARCHZAPFEN (SNZ)* en *G. max* y finalmente *HOMEODOMAIN PROTEIN 16 (HB16)* en *C. cajan*, ver Figura 9A. El número de homólogos por gen de *A. thaliana* encontrados tuvo un rango de 1 a 17, siendo el gen *SENSITIVE TO RED LIGHT REDUCED 1 (SRR1)* en la especie *C. cajan* quien presentó 17 secuencias de genes similares. La Figura 9B muestra la variación entre los genes de leguminosas no identificados como homólogos de *Arabidopsis* en la ruta de floración.

Los resultados obtenidos se respaldan en los estudios de Feuillet y Keller, (2002), quienes en su trabajo de genómica comparada en la familia de gramíneas, identificaron marcadores moleculares y genes conservados en el genoma de estas especies, pese a las diferencias en la ploidía y el número de cromosomas de las especies; Acarkan *et al.*, (2000) identificaron un alto nivel de colinealidad a nivel genético y molecular entre las especies *A. thaliana* y *Capsella rubella*, miembros de la familia Brassicaceae y finalmente Choi *et al.*, (2004) en sus análisis moleculares y filogenéticos de leguminosas cultivadas, identificaron macroestructuras del genoma conservado, en especial en el clado galegoide. En función a los resultados, existe una gran conservación para los genes de floración en leguminosas identificados a través de *Arabidopsis*. Las diferencias observadas entre secuencias similares podrían sugerir un proceso evolutivo en estas Fabaceas.

Sin embargo, los genes homólogos identificados y conservados en las especies en estudio no explican en que rutas se encuentran involucradas y como los procesos externos e internos gobiernan su expresión genética. En tal sentido, se utilizó la base de datos FLOR ID (Bouche *et al.*, 2016) para agrupar los genes en las siguientes rutas: (1) Fotoperiodo - PP, (2) Vernalización - VP, (3) Envejecimiento - AP, (4) Hormonas - HP, (5) Temperatura ambiente - TP, (7) Azúcar - SP, (8) Autónoma y procesos generales - GPAP y (9) Desarrollo floral e identidad del meristemo - FDMI. Los resultados encontrados están representados en la Figura 9C, donde de un total de 316 genes homólogos identificados, 215 son genes que están distribuidas en diferentes rutas, pero no son puntos de *crosstalk* con otras rutas. Por otro parte, 64 genes hacen *crosstalk* en dos rutas, 23 en tres rutas, 5 en cuatro rutas, *SHORT VEGETATIVE PHASE (SVP)* y *FRUITFUL (FUL)* en cinco rutas, *APETALA 1 (API)* y *TARGET OF EARLY ACTIVATION TAGGED 1 (TOE1)* en seis rutas, *LEAFY (LFY)* y *SUPPRESSOR OF CONSTANS 1 (SOC1)* en siete rutas y *FLOWERING LOCUS T (FT)* en 8 rutas. Finalmente, la ruta *GIF* incluyó 17 genes.

4.2. Genes ortólogos únicos de *Arabidopsis thaliana* en 13 leguminosas de interés económico en las rutas de floración significativas y su coherencia taxonómica en la familia Fabacea

Se realizó el análisis de ortología a los 316 genes homólogos mediante el software ProteinOrtho5, donde se identificó un total de 21 genes ortólogos únicos en floración para cada una de las leguminosas (Anexo 3). A través de su identificación, se procederá a realizar los análisis filogenéticos y su relación a los caracteres agronómicos de importancia agrícola.

4.2.1. Ruta de procesos generales y procesos autónomos (GPAP)

La ruta autónoma de las plantas responde a procesos endógenos que promueven el desarrollo floral de manera independiente de las características ambientales promotoras de la floración (Anexo 5). En las especies de leguminosas en estudio se logró identificar las siguientes proteínas ortólogas únicas a *Arabidopsis*: *SHK1 BINDING PROTEIN 1 (SKB1)*, *BROTHER OF FT AND TFL 1 (BFT)*, *JUMONJI 32 (JMJ32)*, *MEDIATOR 16 (MED16)*, *SEEDLINGS HYPERRESPONSIVE TO LIGHT (SHL1)* y *HISTONE H2A PROTEIN 9 (HTA9)*. Estos genes ortólogos únicos se usaron para la construcción de árboles filogenéticos con la finalidad de observar el nivel de coherencia taxonómica en la agrupación de especies. Se encontró que los genes *SKB1*, *JMJ32* y *SHL1* se agruparon correctamente, mientras los genes *MED16*, *HTA9* y *BFT* no se ajustaron de forma correcta según la base de datos de GRIN TAXONOMY (Wiersema, 2019).

Según Fu *et al.*, (2013) el gen *SKB1/CAUI/PRMT5* es un supresor epigenético de la señalización de calcio e interviene en el cierre estomático en respuesta al calcio extracelular. La mutación de *SKB1* en *Arabidopsis* incrementó la sensibilidad a la salinidad; afectó la tasa de crecimiento y evasión de la floración. De igual forma, el mediador de la floración que responde al estrés salino es el gen *BROTHER OF FT AND TFL 1 (BFT)*, el cual proporciona una estrategia adaptativa que asegura el éxito reproductivo bajo condiciones de estrés (Ryu *et al.*, 2011). Las proteínas *SKB1* y *BFT* relacionan la ruta de la floración con la sensibilidad salina. Dentro de las especies estudiadas, la salinidad tiene un efecto adverso en *Cajanus cajan*, disminuyendo la producción de flores, vainas, número de semillas y peso de semillas (Azeem *et al.*, 2016). Este patrón también se observa en *Glycine max*, donde la salinidad disminuyó el rendimiento, número de vainas, número y peso de semillas, tamaño, dimensiones de la hoja, número de entrenudos, número de ramas y el peso de 1000 semillas (Ghassemi-Golezani *et al.*, 2009 ; Phang *et al.*, 2008 ; Katerji *et al.*, 2003). En tal sentido *SKB1* y *BFT* permitiría desarrollar material resistente a la salinidad en leguminosas.

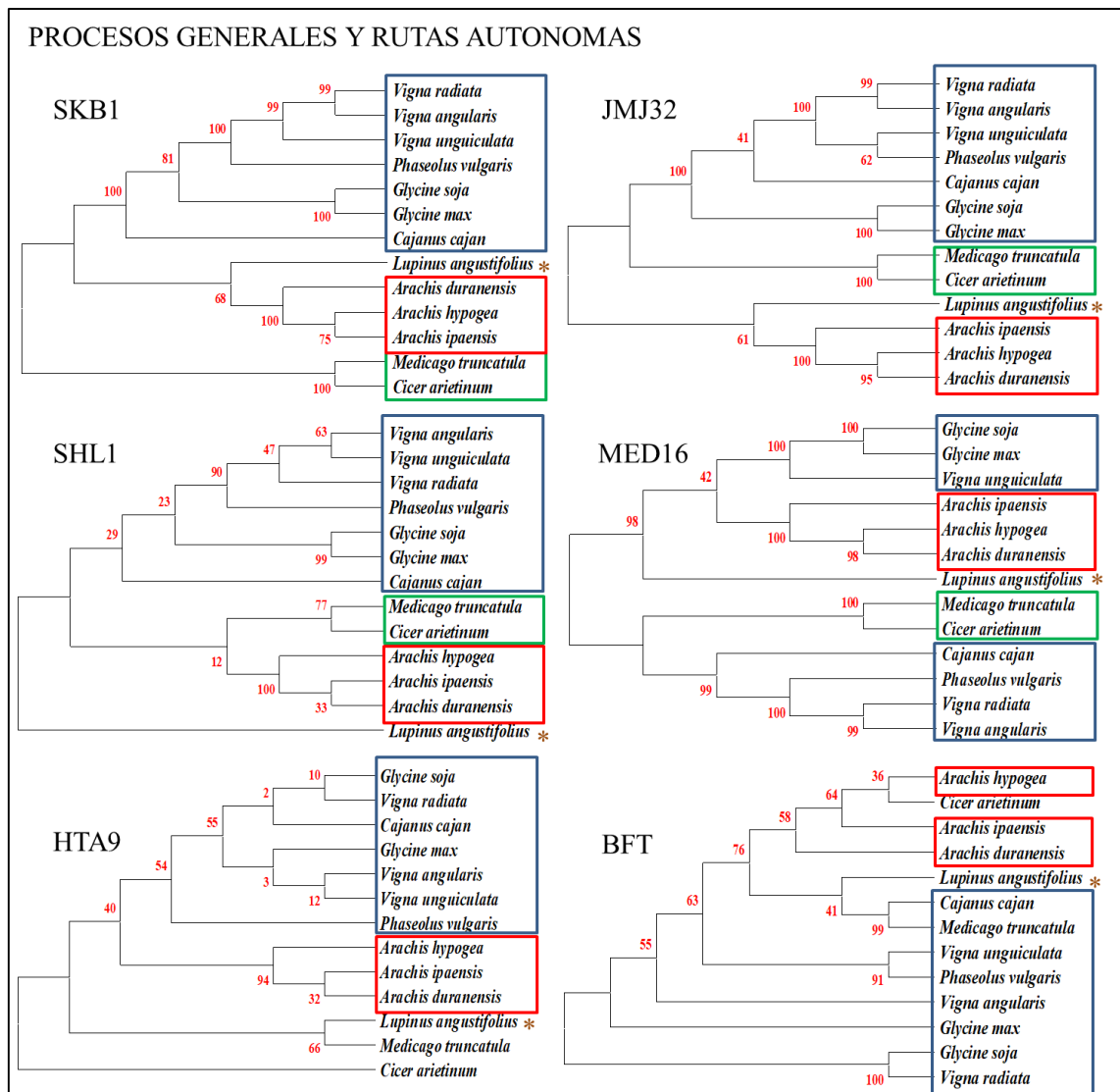


Figura 10. Análisis filogenético de los genes ortólogos en la ruta autónoma y procesos generales. Las cuadrículas de colores encierran las especies de una misma tribu taxonómica. Celda azul: Tribu Phaseoleae, Celda roja: Tribu Dalbergieae, Celda verde: Tribu Cicereae & Tribu Trifolieae y Asterisco: Tribu Genisteae.

El cambio climático incrementa de manera considerable la temperatura, radiación, sequía e inundaciones, provocando desórdenes fisiológicos y fenológicos en los cultivos (Lobell y Gourdji, 2012), se hace necesario encontrar mecanismos de interacción entre floración y control de temperatura con la finalidad de mantener y/o incrementar los rendimientos. El incremento de la temperatura afecta de manera dañina los rendimientos, abscisión de flores y senescencia foliar temprana en *Cicer arietinum*, *Glycine max*, *Lupinus angustifolius*, *Phaseolus vulgaris* y *Vigna unguiculata* (Lake y Sadras, 2014; Jansen, 2008; Konsens *et al.*, 1991; Monterroso y Wien, 1990; Beebe *et al.*, 2011; Bagnall y King, 1987). La Figura 10 muestra el gen ortólogo *JUMONJI 32* (*JMJ32*) para 13 leguminosas, el cual es un regulador del mecanismo de equilibrio en el control de la floración a temperaturas cálidas para prevenir

la floración temprana (Gan *et al.*, 2014). Este gen abre la posibilidad de seleccionar material genético con tolerancia al estrés térmico en leguminosas. Por otro lado, la exposición al frío *Cicer arietinum* disminuye su biomasa, la tasa y duración del llenado de la semilla, el tamaño de la semilla, el peso de la semilla, las vainas por planta y el índice de cosecha disminuyeron mucho en el tipo de macrocarpa (Nayyar *et al.*, 2007). Según Knight *et al.*, (2009), *MEDIATOR 16 (MED16)* es un componente clave en la aclimatación al frío y se ha podido identificar el ortólogo de este gen (Ver Figura 10) en las 13 leguminosas en estudio. La identificación de *MED16* y *JMJ32* permitirá estudiar en detalle los mecanismos de tolerancia al frío y calor en leguminosas y la respuesta de sus órganos reproductivos.

El estrés biótico es otro de los componentes clave en la pérdida de rendimiento en los cultivos. A nivel mundial, las epidemias por *Pseudomonas* han causado una disminución del rendimiento entre 80 – 100% en regiones frías y con alta humedad (Van Schoonhoven y Voysest, 1989). Los cultivos susceptibles a *Pseudomonas syringae pv. phaseolicola* tienen pérdidas considerables, siendo en China un problema epidémico para la especie *Vigna radiata* (Sun *et al.*, 2016). Se ha encontrado que las especies *Cajanus cajan*, *Phaseolus vulgaris*, *Glycina max*, *Vigna angularis*, *Vigna radiata* y *Vigna unguiculata* son susceptibles a *Pseudomonas* (Taylor *et al.*, 1996; Hunter y Taylor, 2005). Según March-Diaz *et al.*, 2007, el gen *HISTONE H2A PROTEIN 9 (HTA9)*, confiere tolerancia a la bacteria *Pseudomonas syringae*. La Figura 11 muestra el gen ortólogo único a *HTA9* en las 13 leguminosas. Las aplicaciones en el mejoramiento a través de la proteína *HTA9* permitiría reducir el grado de incidencia y severidad por *Pseudomonas*, disminuyendo la pérdida de rendimientos en leguminosas.

Finalmente se ha podido identificar el gen ortólogo *SEEDLINGS HYPERRESPONSIVE TO LIGHT (SHL1)* involucrado en el desarrollo post-embriónico, regulación del fotoperiodismo, floración y regulación negativa de la acetilación de histonas (UniProt, 2021; ver Figura 10). Las histonas desacetilasas (HDAC) desempeñan un papel importante en los procesos de crecimiento, desarrollo y defensa de las plantas y son una de las causas principales de modificaciones epigenéticas en un genoma (Liu *et al.*, 2012; Ma *et al.*, 2013). Esto se puede ver en el estudio de Melmaiee *et al.*, (2015), donde los análisis de expresión génica indicaron que el patrón de expresión del gen *PvSRT1* es similar a la tendencia de las *HDAC* activas totales, mostrando una expresión aumentada en las plantas inoculadas durante los puntos de tiempo medidos.

4.2.2. Ruta de desarrollo floral e identidad del meristemo (FDMI)

La transición de la planta del estado vegetativo a estado reproductivo es controlado por la ruta de desarrollo floral e identidad del meristemo (Anexo 5), (Simon *et al.*, 1996). La Figura 11 muestra las siguientes proteínas órtologas únicas identificadas para las leguminosas en estudio: *HUA ENHANCER 3 (HEN3)*, *ENHANCER OF AG-4 1 (HUA1)*, *LEAFY (LFY)* y *BLADE ON PETIOLE (BOP1)*. Solo los genes *LFY* y *HUA1* fueron los únicos en desarrollar arboles filogenéticos coherentes según la base de datos de GRIN TAXONOMY (Wiersema, 2019).

La identidad del meristema es clave para el inicio de la floración y múltiples genes están involucrados en la etapa de transición floral, entre los que destaca el gen integrador *LFY*, donde un estudio realizado por Nilsson *et al.*, (1998), determinó que la sobreexpresión de *LFY* producía flores precoces y en consecuencia, una floración temprana. Además se ha podido demostrar su papel en el desarrollo y formación de semillas en *Glycine max* y *Raphanus sativus* (Meng *et al.*, 2007; Oshima y Nomura, 2008). El presente estudio identificó un gen ortólogo único a *LFY* en 13 leguminosas (Figura 12). La manipulación de este gen permitiría la transición floral y el tamaño de las semillas en leguminosas.

El gen ortólogo *HUA1* (Figura 11) está encargado en el desarrollo de órganos reproductores como los estambres y carpelos en *Arabidopsis* (UniProt, 2021). Para maximizar la producción de semillas en leguminosas, se hace necesario reducir los efectos por sequía en la formación de órganos reproductivos y esterilidad de grano de polen, ya que se observó una disminución de 46 – 71% en *Glycine soja*, 45 – 69% en *Cicer arietinum*, 60% *Vigna unguiculata*, 58 – 87% en *Phaseolus vulgaris*, 40 – 55% en *Cajanus cajan* y 26% en *Vigna radiata* a nivel de la fase reproductiva (Nadeem *et al.*, 2019). La floración y producción de granos es una etapa crítica para los cultivos bajo sequía. Identificar los genes involucrados en la formación de estambres y gineceo son importantes para el proceso de desarrollo de la estructura floral y la formación de semillas, ya que son importantes para incrementar la producción (Smith y Zhao, 2016; Basheer-Salimia *et al.*, 2021).

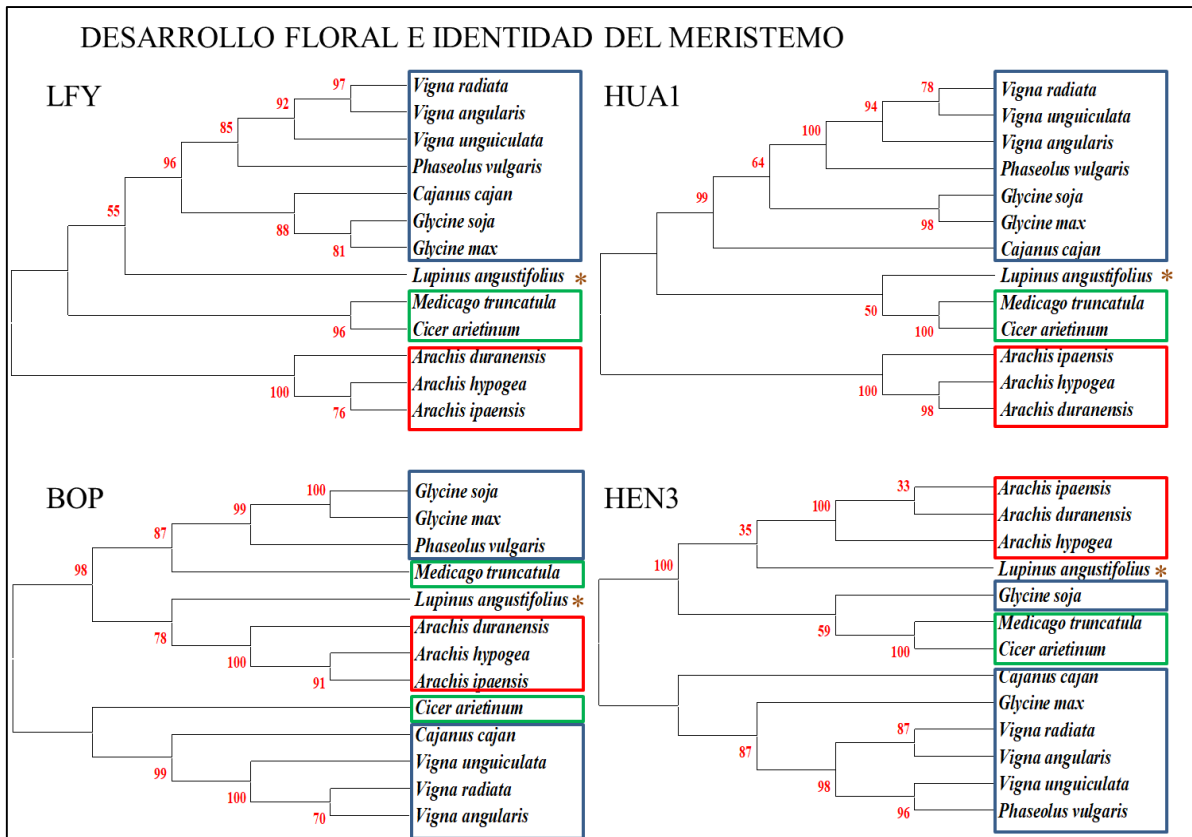


Figura 11. Análisis filogenético de los genes ortólogos en la ruta de desarrollo floral e identidad del meristemo. Las cuadrículas de colores encierran las especies de una misma tribu taxonómica. Celda azul: Tribu Phaseoleae, Celda roja: Tribu Dalbergieae, Celda verde: Tribu Cicereae & Tribu Trifolieae y Asterisco: Tribu Genisteeae.

Por otro lado, la arquitectura de la inflorescencia tiene un fuerte impacto en la producción de frutos y semillas que es un rasgo agronómico de gran relevancia (Benlloch *et al.*, 2015). Donde *HvBOP1* es responsable del control de la longitud de entrenudos y cambios en la inflorescencia (Jost *et al.*, 2016). Se ha podido identificar el gen ortólogo único *BOP1*, componente genético que controla la arquitectura de la inflorescencia en leguminosas (Figura 11) y cómo este gen contribuye en la arquitectura de la inflorescencia, pieza clave para el desarrollo de variedades de alto producción. Esto se sustenta con los estudios sobre *Phaseolus vulgaris* (Zilio *et al.*, 2013) y *Lupinus albus* (Milford *et al.*, 1993).

Los programas de mejoramiento genético en leguminosas se han enfocado entender los mecanismos de tolerancia relacionados al cierre estomático y la sensibilidad al ácido abscísico bajo estrés por sequía (Darai *et al.*, 2016). Es así que Zhu *et al.*, (2020) demostró que el gen *HEN3/CDK8* es un regulador crítico de la señalización del ácido abscísico y de la respuesta a sequía en *Arabidopsis*. Los mutantes de este gen durante sequía, mostraron una hipersensibilidad al ácido abscísico y función alterada en la apertura estomática. Como

las leguminosas demuestran una alta sensibilidad al estrés por sequía, en las etapas vegetativa y reproductiva (Nadeem *et al.*, 2019). La identificación del gen ortólogo *HEN3/CDK8* en leguminosas (Figura 11), permitirá seleccionar material vegetal con un mejor desempeño productivo y manejo del agua para *Arachys hypogea*, *Cajanus cajan*, *Cicer arietinum*, *Glycine max*, *Lupinus angustifolius*, *Medicago truncatula*, *Phaseolus vulgaris*, *Vigna radiata* y *Vigna unguiculata*; ya que estas especies mostraron bajo estrés hídrico, una disminución severa de la floración, reproducción, antesis, llenado de vainas y llenado de granos con pérdidas en el rendimiento entre 27 – 69% (Nadeem *et al.*, 2019).

4.2.3. Ruta de vernalización (VP)

La ruta de vernalización (Anexo 5) requiere el silenciamiento epigenético de *FLC* a través de la metilación de histonas (Bastow *et al.*, 2004). En tal sentido, la ruta fue dividida en promotoras y represoras del gen *FLC*.

4.2.3.1. Represoras del gen FLC

Bajo temperaturas frías prolongadas se da el proceso de vernalización que permite la floración (Kim y Sung, 2014). En *Arabidopsis thaliana* se han identificado 65 genes involucrados en la ruta de vernalización (Bouche *et al.*, 2016). Sin embargo, hemos identificado las siguientes proteínas órtologas únicas en 13 leguminosas con capacidad represora de FLC (Figura 12) : *FERTILIZATION-INDEPENDENT ENDOSPERM 1 (FIE1/FIS3)*, *HISTONE DEACETYLASE 6 (HDA6/RTS1)*, *MULTICOPY SUPPRESSOR OF IRA1 4 (MSI4/FVE)* y *FLOWERING LOCUS D (FLD/RS11)*. Estos genes tienen capacidad negativa en la expresión de FLC.

El tamaño y número de semillas son componentes claves en el rendimiento (Savadi, 2017), donde el volumen del endospermo es fundamental para determinar el tamaño final de la semilla, siendo un constituyente esencial de esta, donde se almacena las reservas que son fuente de alimento para la humanidad (Li y Berger, 2012). Se ha identificado el gen órtologo único *FIE1/FIS3* (Figura 12), donde la pérdida de función de este gen detuvo el desarrollo del embrión en *Arabidopsis* (Sun *et al.*, 2010).

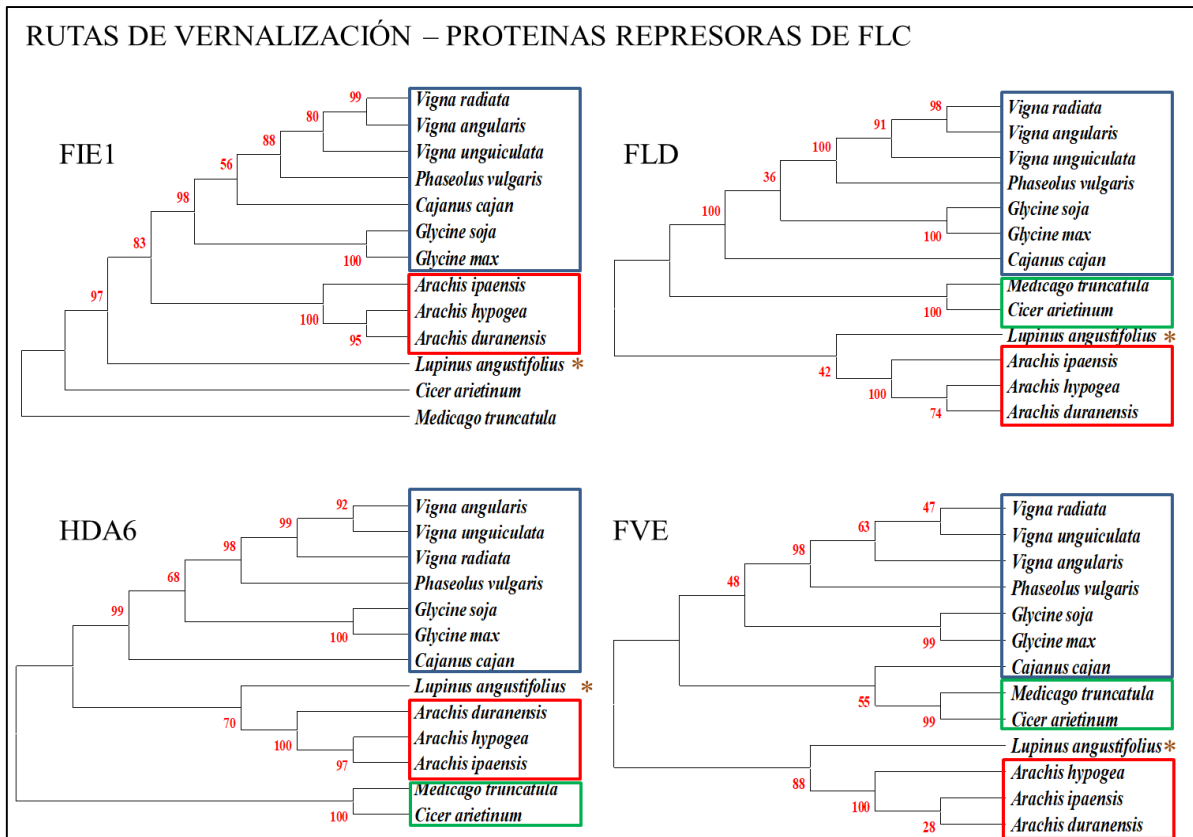


Figura 12. Análisis filogenético de los genes ortólogos en la ruta de vernalización – Proteínas represoras de FLC. Las cuadrículas de colores encierran las especies de una misma tribu taxonómica. Celda azul: Tribu Phaseoleae, Celda roja: Tribu Dalbergieae, Celda verde: Tribu Cicereae & Tribu Trifolieae y Asterisco: Tribu Genisteae.

HDA6 está involucrado en el desarrollo de flores, semillas y hojas en *Arabidopsis*. Mutantes de HDA6 han mostrado una floración tardía (Wang, Z. *et al.*, 2014; Wu, K. *et al.*, 2008). Se pudo identificar el gen ortólogo único *HDA6* en leguminosas, que según Tanaka *et al.*, (2008) es responsable de la modulación de la germinación de semillas con la represión indirecta o directa de propiedades embrionarias y además se relaciona con la tolerancia al estrés por temperatura, sequía y salinidad (Zhao *et al.*, 2020; Luo *et al.*, 2017; Luo *et al.*, 2012). Este gen podría ser clave en los procesos de germinación bajo condiciones de estrés abiótico, lo cual permitiría incrementar las zonas de producción.

El llenado de granos requiere los fotosintatos asimilables producto de la fotosíntesis. En *Arabidopsis* se ha reportado que los mutantes de *FVE/MSI4* tuvieron una mayor cantidad de biomasa y semillas; este gen es considerado como un componente multifactorial de biomasa y rendimiento (Morel *et al.*, 2008). La Figura 12 muestra el ortólogo único de este gen en 13 leguminosas. La selección genética permitiría aumentar el rendimiento y la cantidad de biomasa. Al ser incorporada la biomasa al suelo mejoraría la fertilidad de este ,

pues aumentaría el contenido de carbono orgánico, así como la disponibilidad de nitrógeno en el suelo en beneficio de los cultivos posteriores (Yuvaraj *et al.*, 2020).

Las leguminosas tienen un amplio número de patógenos que disminuyen la producción entre 35 – 70% y los patógenos que atacan a nivel foliar y deterioran el cultivo son *Erysiphe spp*, *Podosphaera spp*, *Peronospora spp*, *Uromyces spp* y *Puccinia spp*, haciendo necesario encontrar mecanismos genéticos de resistencia (Martins *et al.*, 2020), que puedan estar asociados al sistema inmune innato de las plantas. Este estudio identificó el gen ortólogo único *FLD/RSII* en las leguminosas (Figura 12), que participa en la regulación epigenética de la resistencia sistémica adquirida (SAR) para controlar los patógenos (Singh *et al.*, 2014). La manipulación de este gen podría llevar a la creación de cultivares tolerantes a procesos de infección ocasionados por patógenos, disminuyendo los niveles de incidencia y virulencia en los cultivos, mejorando el rendimiento y la calidad.

Los genes ortólogos únicos *FLD*, *HDA6* y *FVE* en los represores de *FLC* y *VIP3* y *SSRP1* en los promotores de *FLC* fueron los únicos en obtener árboles filogenéticos coherentes según la base de datos de GRIN TAXONOMY (Wiersema, 2019).

4.2.3.2. Promotoras del gen FLC

Mientras existen genes que impiden la expresión del gen *FLC*, existen aquellos que promueven su expresión, reprimiendo el tiempo y la transición al estado floral (Willmann y Poethig, 2011), lo que involucra una floración tardía. Los genes con características represoras del *FLC* han sido identificadas en *Arabidopsis* y este estudio reporta 5 genes ortólogos únicos para 13 leguminosas: *EARLY FLOWERING 7 (ELF7)*, *VERNALIZATION INDEPENDENT 3 (VIP3/SKI8)*, *CYCLIN DEPENDENT KINASE GROUP C2 (CDKC2)*, *HIGH MOBILITY GROUP (HMG/SSRP1/NFD)* y *YEAST ALL1 FUSED GENE FROM CHROMOSOME 9 (YAF9A/TAF14B)*.

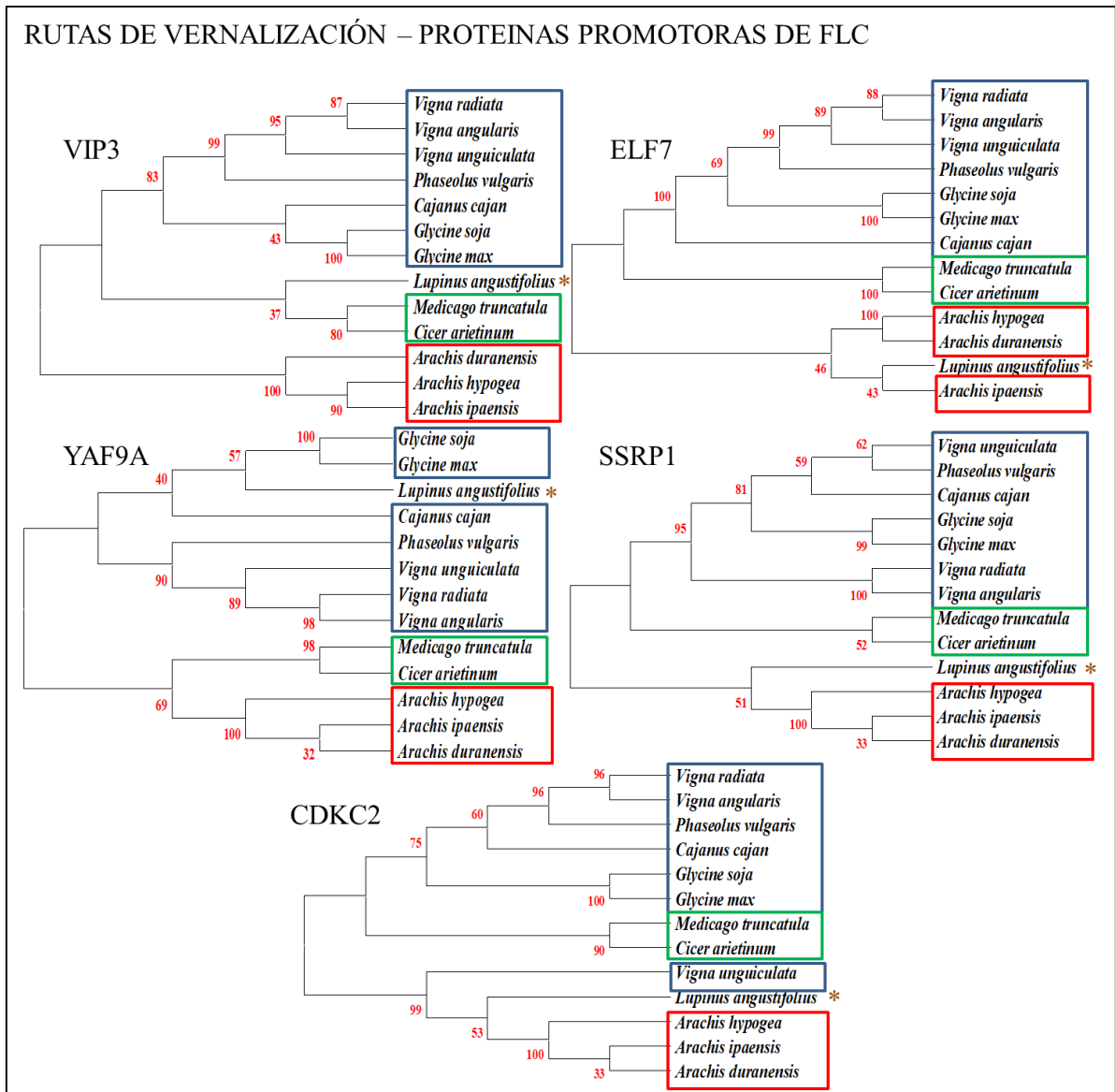


Figura 13. Análisis filogenético de los genes ortólogos en la ruta de vernalización – Proteínas promotoras de *FLC*. Las cuadrículas de colores encierran las especies de una misma tribu taxonómica. Celda azul: Tribu Phaseoleae, Celda roja: Tribu Dalbergieae, Celda verde: Tribu Cicereae & Tribu Trifolieae y Asterisco: Tribu Genisteae.

La Figura 13 muestra los ortólogos únicos para el gen *ELF7*, el cual codifica en plantas para un homólogo en el complejo *PAF1* en levaduras y mutaciones en este gen resultan en una floración temprana en *Arabidopsis thaliana* (He *et al.*, 2004). También se ha demostrado que mutaciones en *ELF7* afectan la dormancia de semillas (Liu *et al.* 2011), lo que demuestra la importancia del complejo *PAF1* en la latencia de semillas y floración. Además otro gen que afecta a estos dos procesos de forma similar es *VIP3*. Estudios con mutantes en este gen han demostrado que afecta los niveles de expresión de *FLC* (Yu y Michaels, 2010; Zhang *et al.*, 2003) y el fenotipo que se observa es una floración temprana bajo condiciones de días cortos y largos (Oh *et al.*, 2004; Yu y Michaels, 2010; Zhang *et al.*, 2003). *VIP3* es también

un regulador positivo del proceso de germinación (Auge *et al.*, 2017). El mejoramiento enfocado en *ELF7* y *VIP3* permitiría desarrollar un ideotipo precoz y de germinación temprana. Las modificaciones a nivel de la germinación podrían derivar en semillas adecuadas para condiciones específicas, como una latencia prolongada en almacenes de granos o una corta, para la siembra en momentos clave de la producción agrícola y el ingreso a ventanas comerciales de exportación.

CDKC2 es un gen que codifica para una subunidad CDKCquinasa (Bouche *et al.*, 2016) y es necesario para una eficiente transcripción de *COOLAIR* (Wang, Z-W. *et al.*, 2014), donde *COOLAIR* está asociado físicamente con *FLC*, acelerando la represión transcripcional de *FLC* (Csorba *et al.*, 2014). Los mutantes en *CDKC2* muestran una floración tardía bajo días largos y bajo condiciones de 12 horas de luz – 12 horas de oscuridad (Cui *et al.*, 2007; Wang, Z-W. *et al.*, 2014). Además, se ha demostrado que el gen *CDKC2* presenta una respuesta negativa bajo estrés por sequía (Zhao *et al.*, 2017). Por otro lado, *HMG* es un componente del complejo *FACT*, que actúa reorganizando el nucleosoma (Bouche *et al.*, 2016). Mutaciones de este gen muestran un efecto pleitrópico y los niveles de expresión de *FLC* son reducidos (Lolas *et al.*, 2010). El mutante muestra una floración temprana bajo condiciones de días largos y cortos. Sin embargo, el mutante fue capaz de producir mayor número de hojas que el parental bajo días cortos (Lolas *et al.*, 2010). La expresión de *SSRP1/HMG* se ve afectada por el estrés lumínico, modificando la biosíntesis de antocianinas en *Arabidopsis* (Pfab *et al.*, 2018). Además se ha podido demostrar que este gen *Betula platyphylla* tiene un rol crítico en la tolerancia a sales (Lei *et al.*, 2020). Esto expone la importancia que poseen los genes *CDKC2* y *SSRP1/HMG* para el mejoramiento genético a diferentes niveles y el descubrimiento de los ortólogos para estos genes en 13 leguminosas (Figura 13) abre un abanico de posibilidades para su aplicación en tolerancia a la radiación, sequía, la salinidad y la producción de antioxidantes.

El mutante simple *YAF9* muestra una floración temprana bajo condiciones de días largo y cortos (Bieluszewski *et al.*, 2015; Zacharaki *et al.*, 2012). Además, el rol de *YAF9* en el control de la floración puede ser a través de la regulación de la *FLC* (Zacharaki *et al.*, 2012; Bieluszewski *et al.*, 2015) o *GIGANTEA* (Su *et al.*, 2017). El gen *GIGANTEA* es un gen integrador cuya función biológica controla el ritmo circadiano, señalización luminosa, tolerancia al frío, señalización hormonal y floración fotoperiodica (Brandoli *et al.*, 2020). La identificación del ortólogo único *YAF9* en leguminosas (Figura 13), que es parte del complejo *SWRI* y controla los niveles de acetilación (Bouche *et al.*, 2016) permitiría estudiar

la regulación transcripcional, silenciamiento de la heterocromatina, progresión del ciclo celular y reparación del ADN (Espinosa- Cores *et al.*, 2020), con la finalidad de ver su aplicación en el mejoramiento genético de leguminosas.

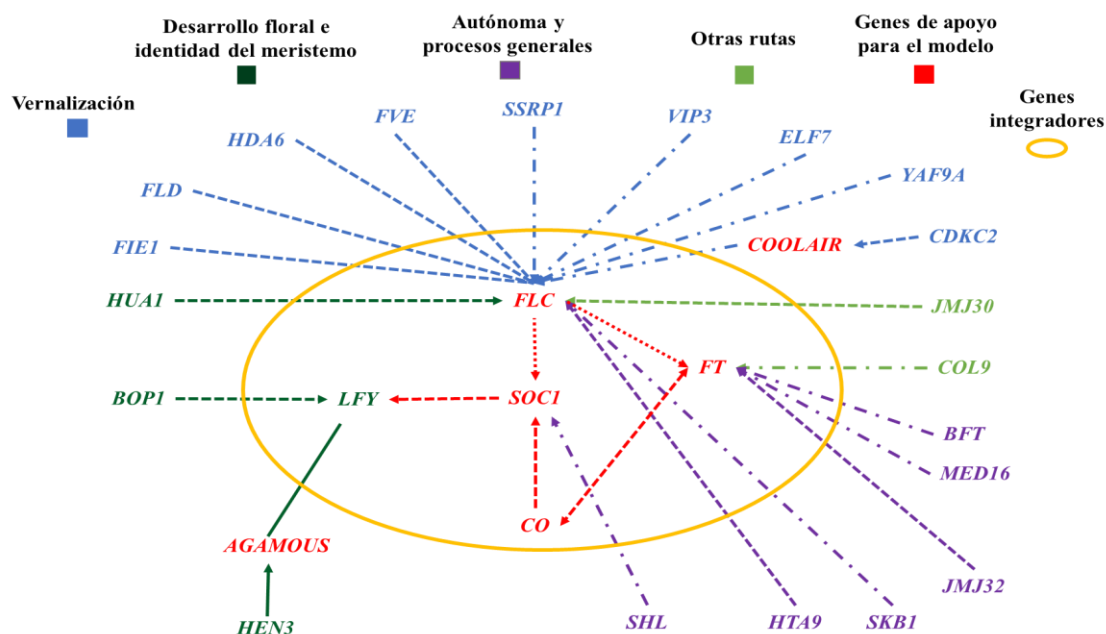
4.3. Interacción de los genes ortólogos únicos de floración en las 13 leguminosas de interés económico para estudios en mejoramiento genético

La Figura 14, muestra dos genes ortólogos únicos para leguminosas que no pertenecen a ninguna de las rutas de floración anteriormente señaladas. *COSTANS-LIKE 9 (COL9/BBX7)* es un factor de transcripción perteneciente a la familia de genes *CONSTANS-LIKE* que codifica para una proteína nuclear. La expresión de *COL9* es regulado por el ritmo circadiano en el fotoperiodo. Los niveles de ARNm de *CONSTANS* y *FT* se encuentran reducidos en líneas de *Arabidopsis* que sobreexpresan el gen *COL9* (Cheng y Wang, 2005). Bajo condiciones de frío, este factor de transcripción se encuentra sobreexpresado en el ecotipo Eringsboda de *Arabidopsis* (Barah *et al.*, 2013) y además es inducido por ácido jasmónico (Mandaokar *et al.*, 2006) y sobreexpresado por ABA (Sánchez *et al.*, 2004). Por otro lado, *JUMONJI 30 (JMJ30)* codifica para una histona desmetilasa involucrada en la regulación del ciclo circadiano (Lu *et al.*, 2011). La sobreexpresión de la *JMJ30* induce una floración tardía bajo condiciones de días largos y cortos (Yan *et al.*, 2014; Gan *et al.*, 2014; Lu *et al.*, 2011). El doble mutante *jmj30;jmj32* reduce la expresión de la *FLC* (Gan *et al.*, 2014). Además, el complejo *JMJ30* y *EFM* media la floración en respuesta de temperatura y luz (Yan *et al.*, 2014) y según Wu, J. *et al.*, (2019), este gen controla el crecimiento radicular en respuesta a la señalización ABA bajo estrés.

Las relaciones múltiples entre los genes ortólogos únicos y los genes integradores *FLOWERING LOCUS C (FLC)*, *FLOWERING LOCUS T (FT)*, *CONSTANS (CO)*, *SUPRESSOR OF CO 1 (SOC1)* y *LEAFY (LFY)*, descritos así por Van Dijk y Molenaar, (2017), se observan en la Figura 14. El primer gen integrador *FLC* está involucrado en evasión de la floración por represión de los genes *FT* y *SOC1* (Deng *et al.*, 2011). *FT* está involucrado en múltiples procesos como la floración, el cuajado, crecimientos vegetativo, control de estomas y tuberización (Pin y Nillson, 2012) y *SOC1* puede también ser inducido en estrés por sequía (dependiente de ABA) bajo condiciones de días largos (Castelán-Muñoz *et al.*, 2019). Ambos genes son regulados de forma positiva por *CO* (Yoo *et al.*, 2005), factor de transcripción que se expresa en condiciones de luz y se degrada bajo oscuridad (Tremblay y Colasanti, 2006). Por otro lado, *LFY* esta involucrada en la identidad del meristemo floral por inducción de los genes *CAL* y *API* (William *et al.*, 2004). Los genes ortólogos únicos se

interrelacionan con los integradores a través de sistemas promotores o represores de su expresión genética. En la ruta de vernalización, los genes ortólogos promotores de la *FLC* son *ELF7*, *VIP3/SKI8*, *CDKC2*, *HMG/SSRP1/NFD* y *YAF9A/TAF14B*, mientras los represores son *FIE1/FIS3*, *HDA6/RTS1*, *MSI4/FVE* y *FLD/RSI1*. *FLC* es un punto de *crossstalk* en la ruta autónoma, reprimiendo *FT* y *SOC1*. La interacción de *SHL* con *FLC* se da a través de *SOC1*. Por otro lado, la modulación positiva de *FLC* se da a través de *HTA9* y *SKB1* lo modula de forma negativa. De igual forma, *FT* es reprimido por los genes *BFT* y *MED16*, mientras *JMJ32* promueve su expresión. Finalmente, a nivel de la ruta de desarrollo floral e identidad del meristemo, *HUAL* promueve la expresión de *FLC*. Entretanto, *BOP1* y *HEN3* van a interactuar con *FLC* a través del *network LFY - SOC1*.

En la Figura 14, se muestra la red de conexión entre los ortólogos únicos y los genes integradores (*FLC*, *FT*, *SOC* y *LFY*). Esta información permite conectar los genes de floración con caracteres agronómicos tales como la arquitectura de la inflorescencia, la calidad y rendimiento de los granos, la tolerancia/resistencia al estrés abiótico (sequía, salinización, temperatura y frío) y factores bióticos. Por otro lado, el ritmo circadiano afecta la fotosíntesis, el metabolismo del almidón y hasta el 90% de la transcripción. Por lo tanto, una óptima regulación del ritmo circadiano es importante para incrementar la fijación de CO₂ y el crecimiento, su disrupción disminuye el vigor de la planta (Chen, 2021). Por ejemplo *FVE/MSI4* es considerado un componente multifactor de la biomasa y producción, con la capacidad de incrementar entre 3 – 8 veces la biomasa en función del fotoperiodo. Este tipo de genes son herramientas ideales en el mejoramiento genético porque permiten la reingeniería de la biomasa y rendimiento en función de las condiciones geográficas o climáticas (Morel *et al.*, 2008).



Genes	Ruta	Referencia (Floración)	Descriptor agronómico	Referencia (Agronómico)
SKB1		Wang <i>et al.</i> , 2007	Estrés abiótico (salinidad), crecimiento y evasión floral	Fu <i>et al.</i> , 2013
BFT	Autónoma y procesos generales	Ryu <i>et al.</i> , 2014	Estrés abiótico (salinidad)	Ryu <i>et al.</i> , 2011
JMJ32		Gan <i>et al.</i> , 2014	Estrés abiótico (temperaturas)	Gan <i>et al.</i> , 2014
MED16		Knight <i>et al.</i> , 2008	Estrés abiótico (frío)	Knight <i>et al.</i> , 2009
HTA9		March-Díaz <i>et al.</i> , 2007	Estrés biótico (<i>Pseudomonas syringae</i>)	March-Díaz <i>et al.</i> , 2007
SHL1		Lopez-Gonzales <i>et al.</i> , 2014	Desarrollo post-embriionario	UniProt, 2021
LFY	Desarrollo floral e identidad del meristemo	Wang <i>et al.</i> , 2004	Proceso de germinación	Meng <i>et al.</i> , 2007; Oshima y Nomura, 2008
HUA1		Doyle <i>et al.</i> , 2004	Desarrollo de estambres y carpelos	UniProt, 2021
BOP1		Xu <i>et al.</i> , 2010	Arquitectura de la inflorescencia	Jost <i>et al.</i> , 2016
HEN3		Gonzales <i>et al.</i> , 2007; Parcy <i>et al.</i> , 2002	Estrés abiótico (sequía)	Zhu <i>et al.</i> , 2020
FIE1	Vernalización (Represores FLC)	Jiang <i>et al.</i> , 2008	Desarrollo del embrión	Sun <i>et al.</i> , 2010
HDA6		Yu <i>et al.</i> , 2011	Estrés abiótico (temperatura, sequía y salinidad) y proceso de germinación	Zhao <i>et al.</i> , 2020; Luo <i>et al.</i> , 2012, 2017; Tanaka <i>et al.</i> , 2016
FVE		Kim <i>et al.</i> , 2004	Producción de biomasa y rendimiento	Morel <i>et al.</i> , 2009
FLD		He <i>et al.</i> , 2003	Estrés biótico (resistencia sistémica)	Singh <i>et al.</i> , 2014
ELF7	Vernalización (Promotores FLC)	He <i>et al.</i> , 2004	Proceso de germinación (dormancia)	Li <i>et al.</i> , 2011
VIP3		Zhang <i>et al.</i> , 2003	Proceso de germinación	Auge <i>et al.</i> , 2017
CDKC2		Wang <i>et al.</i> , 2014	Estrés abiótico (sequía)	Zhao <i>et al.</i> , 2017
SSRP1		Lolas <i>et al.</i> , 2010	Estrés abiótico (lumínico y sales)	Pfab <i>et al.</i> , 2018; Lei <i>et al.</i> , 2020
YAF9A		Zacharaki <i>et al.</i> , 2012	Regulación de FLC y GIGANTEA	Zacharaki <i>et al.</i> , 2012; Bieluszewski <i>et al.</i> , 2015; Su <i>et al.</i> , 2017
COL9	Otros	Cheng <i>et al.</i> , 2005	Estrés abiótico (frío) y Hormonal (inducido por ABA y JA)	Barah <i>et al.</i> , 2013
JMJ30		Gan <i>et al.</i> , 2014	Estrés abiótico (Temperatura)	Wu <i>et al.</i> , 2019

Figura 14. Modelo en base a ortólogos únicos en las rutas de floración con relevancia agronómica en 13 leguminosas con importancia económica.

V. CONCLUSIONES

- Se logró identificar un total de 236 genes de floración de *A. thaliana* homólogos a las 13 leguminosas, siendo 135 genes de floración en promedio ausentes de homología.
- Se encontró 21 genes ortólogos únicos que están relacionados a la ruta autónoma y procesos generales, ruta del desarrollo floral e identidad del meristemo y ruta de la vernalización.
- La identificación de los genes ortólogos únicos se relacionan con caracteres agronómicos (*SHL*, *LFY*, *HUA1*, *BOP1*, *FIE1*, *FVE*, *ELF7*, *VIP3* y *YAF9*), estrés abiótico (*SKB1*, *BFT*, *JMJ32*, *MED16*, *HEN3*, *HDA6*, *SSRP1*, *CDKC2*, *COL9* y *JMJ30*) y bióticos (*HTA9* y *FLD*).
- Solo para 11 genes ortólogos únicos (*LFY*, *HUA1*, *FLD*, *HDA6*, *FVE*, *VIP3*, *SSRP1*, *SKB1*, *SHL1*, *JMJ32* y *JMJ30*) los árboles filogenéticos mostraron coherencia con el agrupamiento según la tribu Phaseoleae (*Vigna*, *Phaseolus*, *Glycine* y *Cajanus*), Dalbergieae (*Arachis*), Ginesteae (*Lupinus*), Trifolieae (*Medicago*) y Cicereae (*Cicer*).

VI. RECOMENDACIONES

- Se va requerir la validación de esta información generada desde una perspectiva biológica.
- Se deben realizar estudios similares para mejorar la capacidad de absorción de nutrientes y la formación de fotosintatos en los cultivos que son de agroexportación y con potencial de serlo.

VII. BIBLIOGRAFÍA

- Acarkan, A., Robberg, M., Koch, M. y Schmidt, R. (2000). Comparative genome analysis reveals extensive conservation of genome organisation for *Arabidopsis thaliana* and *Capsella rubella*. *The plant journal* 23(1), 55 - 62. <https://doi.org/10.1046/j.1365-313x.2000.00790.x>
- Alcántar-Vázquez, J. P. (2014). La poliploidía y su importancia evolutiva. *Temas de Ciencia y Tecnología*, 18(54), 17–29.
- Altenhoff, A. M. y Dessimoz, C. (2009). Phylogenetic and functional assessment of orthologs inference projects and methods. *PLoS Computational Biology*, 5(1). <https://doi.org/10.1371/journal.pcbi.1000262>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. y Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [http://doi.org/10.1016/s0022-2836\(05\)80360-2](http://doi.org/10.1016/s0022-2836(05)80360-2)
- Ashikari, M., Wu, J., Yano, M., Sasaki, T. y Yoshimura, A. (1999). Rice gibberellin-insensitive dwarf mutant gene Dwarf 1 encodes the α -subunit of GTP-binding protein. *Proceedings of the National Academy of Sciences of the United States of America*, 96(18), 10284–10289. <https://doi.org/10.1073/pnas.96.18.10284>
- Auge, G. A., Blair, L. K., Neville, H. y Donohue, K. (2017). Maternal vernalization and vernalization-pathway genes influence progeny seed germination. *New Phytologist*, 216(2), 388–400. <https://doi.org/10.1111/nph.14520>
- Azeem, T. M., Qasim, M., Ahmed, N. y Ahmad, R. (2016). Salt stress responses of pigeon pea (*Cajanus cajan*) on growth, Yield and some biochemical attributes. *Pakistan Journal of Botany* 48(4), 1353-1360.
- Bachhawat, A. K. (2006). Comparative genomics. *Resonance*, 11(8), 22–40. <https://doi.org/10.1007/bf02855776>

- Bagnall, D.J. y King, R.W. (1987). Temperature and irradiance effects on yield in cowpea (*Vigna unguiculata*). *Field Crops Research* 16: 217-229. [https://doi.org/10.1016/0378-4290\(87\)90061-X](https://doi.org/10.1016/0378-4290(87)90061-X)
- Barah, P., Jayavelu, N., Rasmussen, S., Nielsen, H., Mundy, J. y Bones, A. M. (2013). Genome-scale cold stress response regulatory networks in ten *Arabidopsis thaliana* ecotypes. *BMC Genomics*, 14(1), 722. <https://doi.org/10.1186/1471-2164-14-722>
- Basheer-Salimia, R., Alowaidat, M. Y., Al-Salimiya, M. A., Hamdan, Y. A. y Sayara, T. A. (2021). Comparative Study of Five Legume Species under Drought Conditions. *Legume Research - An International Journal*. <https://doi.org/10.18805/LR-577>
- Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A. y Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature*, 427(6970), 164–167. <https://doi.org/10.1038/nature02269>
- Beebe, S., Ramirez, J., Jarvis, A., Rao, I. M., Mosquera, G., Bueno, J. M. y G., Blair, M. (2011). Genetic improvement of common beans and the challenges of climate change. En Yadav, S. S., Redden, R. J., Hatfield, J. L., Lotze-Campen, H., and Hall, A. E. (Ed.) *Crop Adaptation to Climate Change*. 1st ed (pp. 356–369). Wiley, New York.
- Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W. y Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, 15(4), 365–377. <https://doi.org/10.1111/j.1461-0248.2011.01736.x>
- Benlloch, R., Berbel, A., Ali, L., Gohari, G., Millán, T. y Madueño, F. (2015). Genetic control of inflorescence architecture in legumes. *Frontiers in Plant Science* 6. <https://doi.org/10.3389/fpls.2015.00543>
- Bhadouria, R., Singh, R., Singh, V.K., Borthakur, A., Ahamad, A., Kumar, G. and Singh, P. (2019). Agriculture in the Era of Climate Change : Consequences and Effects. En *Climate change and agricultural ecosystem - Current challenges and adaptation* (pp. 2–18). Chennai: Woodhead publishing.
- Bi, X., Esse, W. Van, Mulki, M. A., Kirschner, G., Zhong, J., Simon, R. y Von Korff, M. Von. (2019). *Centroradialis* interacts with flowering locus t-like genes to control floret development and grain number. *Plant Physiology*, 180(2), 1013–1030. <https://doi.org/10.1104/pp.18.01454>

- Bieluszewski, T., Galganski, L., Sura, W., Bieluszewska, A., Abram, M., Ludwikow, A., ... Sadowski, J. (2015). AtEAF1 is a potential platform protein for Arabidopsis NuA4 acetyltransferase complex. *BMC Plant Biology*, *15*(1), 75. <https://doi.org/10.1186/s12870-015-0461-1>
- Blumel, M., Dally, N. y Jung, C. (2015). Flowering time regulation crops - What did we learn from Arabidopsis?. *Current Opinion in Biotechnology* *32*, 121-129. <https://doi.org/10.1016/j.copbio.2014.11.023>
- Bouché, F., Lobet, G., Tocquin, P., y Périlleux, C. (2016). FLOR-ID: An interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Research*, *44*(D1), D1167–D1171. <https://doi.org/10.1093/nar/gkv1054>
- Bradford, K. J., y Nonogaki, H. (2007). Seed Development, Dormancy and Germination. *Seed Development, Dormancy and Germination*. <https://doi.org/10.1002/9780470988848>
- Brandoli, C., Petri, C., Egea-Cortines, M. y Weiss, J. (2020). GIGANTEA: Uncovering New Functions in Flower Development. *Genes*, *11*(10), 1142. <https://doi.org/10.3390/genes11101142>
- Bremner, P. M., Eckersall, R. N. y Scott, R. K. (1963). The relative importance of embryo size and endosperm size in causing the effects associated with seed size in wheat. *The Journal of Agricultural Science*, *61*(1), 139–145. <https://doi.org/10.1017/S0021859600013800>
- Büttner, M. y Singh, K. B. (1997). *Arabidopsis thaliana* ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(11), 5961–5966. <https://doi.org/10.1073/pnas.94.11.5961>
- Calles, T., Del Castello, R., Baratelli, M., Xipsiti, M. y Navarro, D. K. (2019). The International Year of Pulses: Final Report. Rome. FAO. 40 pp. Licence: CC BY-NC-SA 3.0 IGO.
- Capovilla, G., Schmid, M. y Posé, D. (2015). Control of flowering by ambient temperature. *Journal of Experimental Botany*, *66*(1), 59–69. <https://doi.org/10.1093/jxb/eru416>

- Castelán-Muñoz, N., Herrera, J., Cajero-Sánchez, W., Arrizubieta, M., Trejo, C., García-Ponce, B., ... Garay-Arroyo, A. (2019). MADS-box genes are key components of genetic regulatory networks involved in abiotic stress and plastic developmental responses in plants. *Frontiers in Plant Science*, *10*(July). <https://doi.org/10.3389/fpls.2019.00853>
- Chen (2021). Circadian Rhythms and Heterosis. <https://polyploidy.biosci.utexas.edu/circadian-rhythms-and-heterosis>
- Chen, L. T. y Wu, K. (2010). Role of histone deacetylases HDA6 and HDA19 in ABA and abiotic stress response. *Plant Signaling and Behavior*, *5*(10), 1318–1320. <https://doi.org/10.4161/psb.5.10.13168>
- Chen, L. T., Luo, M., Wang, Y. Y. y Wu, K. (2010). Involvement of Arabidopsis histone deacetylase HDA6 in ABA and salt stress response. *Journal of Experimental Botany*, *61*(12), 3345–3353. <https://doi.org/10.1093/jxb/erq154>
- Cheng, J. Z., Zhou, Y. P., Lv, T. X., Xie, C. P. y Tian, Ch. E. (2017). Research progress on the autonomous flowering time pathway in Arabidopsis. *Physiology and Molecular Biology of Plants*. <https://doi.org/10.1007/s12298-017-0458-3>
- Cheng, X.-F. y Wang, Z.-Y. (2005). Overexpression of COL9, a CONSTANS-LIKE gene, delays flowering by reducing expression of CO and FT in *Arabidopsis thaliana*. *The Plant Journal*, *43*(5), 758–768. <https://doi.org/10.1111/j.1365-313x.2005.02491.x>
- Choi, H. K., Mun, J. H., Kim, D. J., Zhu, H., Mudge, J., Roe, B., ...Cook, D. R.(2004). Estimating genome conservation between crop and model legume species. *PNAS Vol. 101*, N° 23, 15286 - 15294. <https://doi.org/10.1073/pnas.0402251101>
- Coakley, S. M. y Scherm, H. (1999). Climate change and plant disease mangament. *Annual Review of Phytopathology* (37), 399–426. <https://doi.org/10.1146/annurev.phyto.37.1.399>.
- Csorba, T., Questa, J. I., Sun, Q. y Dean, C. (2014). Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proceedings of the National Academy of Sciences*, *111*(45), 16160–16165. <https://doi.org/10.1073/pnas.1419030111>

- Cui, X., Fan, B., Scholz, J. y Chen, Z. (2007). Roles of Arabidopsis Cyclin-Dependent Kinase C Complexes in Cauliflower Mosaic Virus Infection, Plant Growth, and Development. *The plant cell online*, 19(4), 1388–1402. <https://doi.org/10.1105/tpc.107.051375>
- Darai, R., Ojha, B., Sarker, A. y Sah, R. (2016). Genetics and breeding for drought tolerance in food legumes. *International Journal of environment, agriculture and biotechnology* Vol. 1(4), 958-967. <https://doi.org/10.22161/ijeab/1.4.47>
- Davis, C. C., Willis, C. G., Primack, R. B. y Miller-Rushing, A. J. (2010). The importance of phylogeny to the study of phenological response to global climate change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1555), 3202–3213. <https://doi.org/10.1098/rstb.2010.0130>
- Deng, W., Ying, H., Helliwell, C. A., Taylor, J. M., Peacock, W. J. y Dennis, E. S. (2011). *FLOWERING LOCUS C (FLC)* regulates development pathways throughout the life cycle of Arabidopsis. *Proceedings of the National Academy of Sciences*, 108(16), 6680–6685. <https://doi.org/10.1073/pnas.1103175108>
- Donoghue, M. J. (2008). A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 105(SUPPL. 1), 11549–11555. <https://doi.org/10.1073/pnas.0801962105>
- Doyle, J. J. (2012). Polyploidy in Legumes. En Soltis, P. S. y Soltis, D. E. (Ed.), *Polyploidy and Genome Evolution*, (pp. 147 – 180). USA. Springer.
- Doyle, J. J. y Egan, A. N. (2009). Dating the origins of polyploidy events. *New Phytologist*, 186(1), 73–85. <https://doi.org/10.1111/j.1469-8137.2009.03118.x>
- Dufayard, J. F., Duret, L., Penel, S., Gouy, M., Rechenmann, F. y Perrière, G. (2005). Tree pattern matching in phylogenetic trees: Automatic search for orthologs or paralogs in homologous gene sequence databases. *Bioinformatics*, 21(11), 2596–2603. <https://doi.org/10.1093/bioinformatics/bti325>
- Endo, T., Shimada, T., Fujii, H., Kobayashi, Y., Araki, T. y Omura, M. (2005). Ectopic expression of an FT homolog from citrus confers an early flowering phenotype on trifoliolate orange (*Poncirus trifoliata* L. Raf.) *Transgenic research* 14(5), 703-12. <https://doi.org/10.1007/s11248-005-6632-3>.

- Espinosa-Cores, L., Bouza-Morcillo, L., Barrero-Gil, J., Jiménez-Suárez, V., Lázaro, A., Piqueras, R., ... Piñeiro, M. (2020). Insights Into the Function of the NuA4 Complex in Plants. *Frontiers in Plant Science*, 11. <https://doi.org/10.3389/fpls.2020.00125>
- Ewan, R., Pangestuti, R., Thornber, S., Craig, A., Carr, C., O'Donnell, L., ... Sadanandom, A. (2011). Deubiquitinating enzymes *AtUBP12* and *AtUBP13* and their tobacco homologue NtUBP12 are negative regulators of plant immunity. *New Phytologist*, 191(1), 92–106. <https://doi.org/10.1111/j.1469-8137.2011.03672.x>
- Faiza, M. (2018) Most widely used tools for phylogenetic tree customization. *Bioinformatics Reviews* 6 (12). <https://bioinformaticsreview.com/20180818/most-widely-used-tools-for-phylogenetic-tree-customization/>
- FAO (7 de marzo de 2016). Legumbres pueden ayudar a combatir el cambio climático, el hambre y la obesidad en America Latina y el Caribe. FAO. Recuperado de <http://www.fao.org/americas/noticias/ver/es/c/409624/#:~:text=Empleo-,Legumbres%20pueden%20ayudar%20a%20combatir%20el%20cambio%20clim%C3%A1tico%2C%20el%20hambre,Am%C3%A9rica%20Latina%20y%20el%20Caribeytext=Seg%C3%BAn%20la%20Organizaci%C3%B3n%20de%20las,Am%C3%A9rica%20Latina%20y%20el%20Caribe.>
- FAOSTAT (18/12/2020) Crops [Base de datos]. Recuperado de <http://www.fao.org/faostat/es/#data/QC>
- Fei, Z., Tang, X., Alba, R. M., White, J. A., Ronning, C. M., Martin, G. B., ... Giovannoni, J. J. (2004). Comprehensive EST analysis of tomato and comparative genomics of fruit ripening. *The Plant Journal*, 40(1), 47–59. <https://doi.org/10.1111/j.1365-313x.2004.02188.x>
- Fernandez, E., Alvarez, M. J., Podhajcer, O. y Stolovitzky, G. (2015). Genómica Funcional : En Busca De La Función De Los Genes. *Actas De La Academia Nacional De Ciencias. (Cordoba, Argentina)*, (November).
- Fernández, V., Takahashi, Y., Le Gourrierc, J. y Coupland, G. (2016). Photoperiodic and thermosensory pathways interact through *CONSTANS* to promote flowering at high temperature under short days. *The Plant journal : for cell and molecular biology*, 86(5), 426–440. <https://doi.org/10.1111/tpj.13183>

- Feuillet, C. y Keller, B. (2002) Comparative genomics in the Grass Family: Molecular Characterization of Grass Genome Structure and Evolution. *Annals of Botany* 89, 3-10. <https://doi.org/10.1093/aob/mcf008>.
- FIA (2017). Serie Estudios Para La Innovación Fia. <https://doi.org/10.1002/adma.201104551>
- Fitch, W.M. (1970). Distinguishing homólogos from analogous proteins. *Systematic Biology* 19 (2), 99-113. <https://doi.org/10.2307/2412448>
- Fiz-Palacios, O., Schneider, H., Heinrichs, J. y Savolainen, V. (2011). Diversification of land plants: Insights from a family-level phylogenetic analysis. *BMC Evolutionary Biology*, 11(1). <https://doi.org/10.1186/1471-2148-11-341>
- Flachowsky, H., Peil, A., Sopenan, T., Elo, A. y Hanke, V. (2007) Overexpression of *BpMADS4* from silver birch (*Betula pendula* Roth.) induces early-flowering in apple (*Malus domestica* Borkh.) *Plant Breeding*, 126, 137-145. <https://doi.org/10.1111/j.1439-0523.2007.01344.x>
- Fu, M., Kang, H. K., Son, S. H., Kim, S. K. y Nam, K. H. (2014). A subset of arabidopsis RAV transcription factors modulates drought and salt stress responses independent of ABA. *Plant and Cell Physiology* (Vol. 55). <https://doi.org/10.1093/pcp/pcu118>
- Fu, Y. L., Zhang, G. B., Lv, X. F., Guan, Y., Yi, H. Y. y Gong, J. M. (2013). Arabidopsis Histone Methylase *CAU1/PRMT5/SKB1* Acts as an Epigenetic Suppressor of the Calcium Signaling Gene *CAS* to Mediate Stomatal Closure in Response to Extracellular Calcium. *The Plant Cell* Vol. 25(8), 2878 – 2891. <https://doi.org/10.1105/tpc.113.113886>
- Fulton, T. M. (2002). Identification, Analysis, and Utilization of Conserved Ortholog Set Markers for Comparative Genomics in Higher Plants. *The plant cell online*, 14(7), 1457–1467. <https://doi.org/10.1105/tpc.010479>
- Gachomo, E. W., Jno Baptiste, L., Kefela, T., Saidel, W. M. y Kotchoni, S. O. (2014). The Arabidopsis *CURVY1* (*CVY1*) gene encoding a novel receptor-like protein kinase regulates cell morphogenesis, flowering time and seed production. *BMC Plant Biology*, 14(1), 1–9. <https://doi.org/10.1186/s12870-014-0221-7>

- Galperin, M. Y., y Koonin, E. V. (2004). "Conserved hypothetical" proteins: Prioritization of targets for experimental study. *Nucleic Acids Research*, 32(18), 5452–5463. <https://doi.org/10.1093/nar/gkh885>
- Gamboa-Tuz, S. D., Pereira-Santana, A., Zhao, T., Schranz, M. E., Castano, E. y Rodriguez-Zapata, L. C. (2018). New insights into the phylogeny of the *TMBIM* superfamily across the three of life: Comparative genomics and synteny networks reveal independent evolution of the BI and LFG families in plants. *Molecular Phylogenetics and Evolution*, 126(April), 266–278. <https://doi.org/10.1016/j.ympev.2018.04.032>
- Gan, D., Zhan, M., Yang, F., Zhang, Q., Hu, K., Xu, W., ... Liang, D. (2017). Expression analysis of argonaute, Dicer-like, and RNA-dependent RNA polymerase genes in cucumber (*Cucumis sativus* L.) in response to abiotic stress. *Journal of Genetics*, 96(2), 235–249. <https://doi.org/10.1007/s12041-017-0758-y>
- Gan, E.-S., Xu, Y., Wong, J.-Y., Geraldine Goh, J., Sun, B., Wee, W.-Y., ... Ito, T. (2014). Jumonji demethylases moderate precocious flowering at elevated temperature via regulation of *FLC* in Arabidopsis. *Nature Communications*, 5(1). <https://doi.org/10.1038/ncomms6098>
- Geraldo, N., Bäurle, I., Kidou, S. I., Hu, X. y Dean, C. (2009). *FRIGIDA* delays flowering in Arabidopsis via a cotranscriptional mechanism involving direct interaction with the nuclear cap-binding complex. *Plant Physiology*, 150(3), 1611–1618. <https://doi.org/10.1104/pp.109.137448>
- Ghassemi-Golezani, K., Taifeh-Noori, M., Oustan, S. y Moghaddam, M. (2009) Response of soybean cultivars to salinity stress. *Journal of Food, Agriculture and Environment* 7(2), 401-404.
- Graham, P. H. y Vance, C. P. (2003). Legumes: Importance and constraints to greater use. *Plant Physiology*, 131(3), 872–877. <https://doi.org/10.1104/pp.017004>
- Gumbel, E. J. (1958) "Statistics of extremes." Columbia University Press, New York, NY.
- Hardy, J. T. (2003) Chapter 1: Earth and greenhouse effect. En Hardy, J. T. (Ed.) *Climate change – Causes, effects, and solutions*. (pp. 3-20). Gran Bretaña. Wiley.

- He, Y., Doyle, M. R. y Amasino, R. M. (2004). *PAF1*-complex-mediated histone methylation of *FLOWERING LOCUS C* chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis. *Genes and Development*, 18(22), 2774–2784. <https://doi.org/10.1101/gad.1244504>
- Howden, S. M., Soussana, J. F., Tubiello, F. N., Chhetri, N., Dunlop, M. y Meinke, H. (2007). Adapting agriculture to climate change. *PNAS*, Vol 104, 19691 - 19696. https://doi.org/10.1073_pnas.0701890104
- Hunter, P. J. y Taylor, J. D. (2005). Patterns of interaction between isolates of three pathovars of *Pseudomonas syringae* and accessions of a range of host and nonhost legume species. *Plant Pathology Vol 55 (1)*, 46 - 53. <https://doi.org/10.1111/j.1365-3059.2005.01314.x>
- Hussain, B. (2015). Modernization in plant breeding approaches for improving biotic stress resistance in crop plants. *Turkish Journal of Agriculture and Forestry*, 39(4), 515–530. <https://doi.org/10.3906/tar-1406-176>
- Hwang, D. Y., Park, S., Lee, S., Lee, S. S., Imaizumi, T. y Song, Y. H. (2019). GIGANTEA Regulates the Timing Stabilization of CONSTANS by Altering the Interaction between FKF1 and ZEITLUPE. *Molecules and Cells*, 42 (10), 693 - 671. <https://doi.org/10.14348/molcells.2019.0199>
- Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., ... Ōmura, S. (2003). Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nature Biotechnology*, 21(5), 526–531. <https://doi.org/10.1038/nbt820>
- Jansen, G. (2008). Effects of temperature on yield and protein content of *Lupinus angustifolius* cultivars. Lupins for health and wealth. *Proceedings of the 12th International Lupin Conference, Fremantle, Western Australia, 14-18 September 2008* (pp.342-345) ref.11
- Jensen, E. S., Peoples, M. B., Boddey, R. M., Gresshoff, P. M., Henrik, H. N., Alves, B. J. R. y Morrison, M. J. (2012). Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. *A review. Agronomy for Sustainable Development (Vol. 32)*. <https://doi.org/10.1007/s13593-011-0056-7>

- Jentsch, A. y Beierkuhnlein, C. (2008). Research frontiers in climate change: Effects of extreme meteorological events on ecosystems. *Comptes Rendus - Geoscience*, 340(9–10), 621–628. <https://doi.org/10.1016/j.crte.2008.07.002>
- Jimenez-Lopez, J. C., Singh, K. B., Clemente, A., Nelson, M. N., Ochatt, S. y Smith, P. M. C. (2020). Editorial: Legumes for Global Food Security. *Frontiers in Plant Science*, 11(7), 1–5. <https://doi.org/10.3389/fpls.2020.00926>
- Jones, A. D. y Ejeta, G. (2015). A global agenda for nutrition and health. *Bull World Health Organ*, 94(9), 228–229. Recuperado de <http://dx.doi.org/10.2471/BLT.15.164509>
- Jost, M., Taketa, S., Mascher, M., Himmelbach, A., Yuo, T., Shahinnia, F., ... Stein, N. (2016). A homolog of Blade-On-Petiole 1 and 2 (*BOPI/2*) controls internode length and homeotic changes of the barley inflorescence. *Plant Physiology* 2, 1113 - 1127. <https://doi.org/10.1104/pp.16.00124>
- Jung, C. H., Wong, C. E., Singh, M. B. y Bhalla, P. L. (2012). Comparative genomic analysis of soybean flowering genes. *PLoS ONE*, 7(6). <https://doi.org/10.1371/journal.pone.0038250>
- Jung, C. y Müller, A. E. (2009). Flowering time control and applications in plant breeding. *Trends in Plant Science*, 14(10), 563–573. <https://doi.org/10.1016/j.tplants.2009.07.005>
- Kadam, S., Vuong, T. D., Qiu, D., Meinhardt, C. G., Song, L., Deshmukh, R., ... Nguyen, H. T. (2016). Genomic-assisted phylogenetic analysis and marker development for next generation soybean cyst nematode resistance breeding. *Plant Science*, 242, 342–350. <https://doi.org/10.1016/j.plantsci.2015.08.015>
- Katerji, N., Van Hoorn, J. W., Hamdy, A. y Mastrorilli, M. (2003) Salinity effect on crop development and yield, analysis of salt tolerance according to several classification methods. *Agricultural Water Management* 62, 37-66.
- Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., ... Somerville, C. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408(6814), 796–815. <https://doi.org/10.1038/35048692>

- Khraiweh, B., Zhu, J. K. y Zhu, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1819(2), 137–148. <https://doi.org/10.1016/j.bbagr.2011.05.001>
- Kim, D.-H. y Sung, S. (2014). Genetic and Epigenetic Mechanisms Underlying Vernalization. *The Arabidopsis Book*, 12, e0171. <https://doi.org/10.1199/tab.0171>
- Kim, E., Ly, V., Hatherell, A. y Schroeder, D. F. (2012). Genetic interactions between arabidopsis DET1 and UVH6 during development and abiotic stress response. *G3: Genes, Genomes, Genetics*, 2(8), 913–920. <https://doi.org/10.1534/g3.112.003368>
- Knight, H., Mugford, S. G., Ülker, B., Gao, D., Thorlby, G. y Knight, M. R. (2009). Identification of *SFR6*, a key component in cold acclimation acting post-translationally on *CBF* function. *The Plant Journal*, 58(1), 97–108. <https://doi.org/10.1111/j.1365-313x.2008.03763.x>
- Knight, H., Thomson, A. J. W. y McWatters, H. G. (2008). Sensitive to freezing6 integrates cellular and environmental inputs to the plant circadian clock. *Plant Physiology*, 148(1), 293–303. <https://doi.org/10.1104/pp.108.123901>
- Kono, T. J. Y., Lei, L., Shih, C.-H., Hoffman, P. J., Morrell, P. L. y Fay, J. C. (2018). Comparative Genomics Approaches Accurately Predict Deleterious Variants in Plants. *G3: Genes, Genomes, Genetics*, 8(10), 3321–3329. <https://doi.org/10.1534/g3.118.200563>
- Konsens, I., Ofir, M. y Kigel, J. (1991). The effect of temperature on the production and abscission of flowers and pods in snap bean (*Phaseolus vulgaris* L.). *Annals of Botany* 67, 391–399. <https://doi.org/10.1093/oxfordjournals.aob.a088173>
- Koonin, E. V. (2005). Orthologs, paralogs, and evolutionary genomics. *Annual Review of Genetics*, 39, 309–338. <https://doi.org/10.1146/annurev.genet.39.073003.114725>
- Kumar, S., Tamura, K. y Nei, M. (1994). MEGA: Molecular evolutionary genetics analysis software for microcomputers. *Bioinformatics*, 10(2), 189–191. <https://doi.org/10.1093/bioinformatics/10.2.189>

- Kuravadi, N. A. y Gowda, M. (2019) Comparison of gene families and synteny analysis from Neem Genome. En Gowda, M., Sheetal, A. y Kole, C. *The Neem Genome*. (pp. 93 - 98). Suiza. Springer.
- Kurukulasuriya, P. y Rosenthal, S. (2013). Climate Change and Agriculture - A Review of Impacts and Adaptations. *Climate Change Series, N° 91*. Editorial. Environment Department Papers. Whashington.
- Lake, L. y Sadras, V. O. (2014). The critical period for yield determination in chickpea (*Cicer arietinum* L.). *Field Crops Research* 168, 1–7. <https://doi.org/10.1016/j.fcr.2014.08.003>
- Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., y Prohaska, S. J. (2011). Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics*, 12. <https://doi.org/10.1186/1471-2105-12-124>
- Lei, X., Liu, Z., Li, X., Tan, B., Wu, J., y Gao, C. (2020). Screening and functional identification of salt tolerance HMG genes in *Betula platyphylla*. *Environmental and Experimental Botany*, 104235. <https://doi.org/10.1016/j.envexpbot.2020.104235>
- León, J. C. (11 de Febrero de 2021) Consumo de legumbres en nuestro país alcanza los 9 kilos por persona al año. Agencia Agraria de noticias. <https://agraria.pe/noticias/consumo-de-legumbres-en-nuestro-pais-alcanza-los-9-kilos-por-23646>
- Li, J. y Berger, F. (2012). Endosperm: Food for humankind and fodder for scientific discoveries. *New Phytologist*, 195(2), 290–305. <https://doi.org/10.1111/j.1469-8137.2012.04182.x>
- Li, P., Zhang, S., Zhang, S., Li, F., Zhang, H., Cheng, F., ... Sun, R. (2015). Carotenoid biosynthetic genes in *Brassica rapa*: Comparative genomic analysis, phylogenetic analysis, and expression profiling. *BMC Genomics*, 16(1), 1–11. <https://doi.org/10.1186/s12864-015-1655-5>
- Li, S., Castillo-González, C., Yu, B., y Zhang, X. (2017). The functions of plant small RNAs in development and in stress responses. *The Plant Journal*, 90(4), 654–670. [doi:10.1111/tbj.13444](https://doi.org/10.1111/tbj.13444)

- Li, X., Wu, H. X. y Southerton, S. G. (2010). Comparative genomics reveals conservative evolution of the xylem transcriptome in vascular plants. *BMC Evolutionary Biology*, 10(1), 190. <https://doi.org/10.1186/1471-2148-10-190>
- Liang, X., Qin, L., Liu, P., Wang, M. y Ye, H. (2014). Genes for iron-sulphur cluster assembly are targets of abiotic stress in rice, *Oryza sativa*. *Plant, Cell and Environment*, 37(3), 780–794. <https://doi.org/10.1111/pce.12198>
- Liu, D., Hunt, M. y Tsai, I. J. (2018). Inferring synteny between genome assemblies: A systematic evaluation. *BMC Bioinformatics*, 19(1), 1–13. <https://doi.org/10.1186/s12859-018-2026-4>
- Liu, X., Luo, M., Zhang, W., Zhao, J., Zhang, J., Wu, K., Tian, L. y Duan, J. (2012). Histone acetyltransferases in rice (*Oryza sativa* L.): phylogenetic analysis, subcellular localization and expression. *BMC Plant Biology* 12,145. <https://doi.org/10.1186/1471-2229-12-145>
- Liu, Y., Geyer, R., van Zanten, M., Carles, A., Li, Y., Hörold, A., ... Soppe, W. J. J. (2011). Identification of the Arabidopsis REDUCED DORMANCY 2 Gene Uncovers a Role for the Polymerase Associated Factor 1 Complex in Seed Dormancy. *PLoS ONE*, 6(7), e22241. <https://doi.org/10.1371/journal.pone.0022241>
- Lobell, D. B. y Gourdjji, S. M. (2012). The influence of climate change on global crop productivity. *Plant Physiology*, 160(4), 1686–1697. <https://doi.org/10.1104/pp.112.208298>
- Lolas, I. B., Himanen, K., Gronlund, J. T., Lynggaard, C., Houben, A., Melzer, M., ... Grasser, K. D. (2010). The transcript elongation factor *FACT* affects Arabidopsis vegetative and reproductive development and genetically interacts with *HUB1/2*. *The Plant Journal*, 61(4), 686–697. <https://doi.org/10.1111/j.1365-313x.2009.04096.x>
- Lu, S. X., Knowles, S. M., Webb, C. J., Celaya, R. B., Cha, C., Siu, J. P. y Tobin, E. M. (2011). The Jumonji C Domain-Containing Protein JMJ30 Regulates Period Length in the Arabidopsis Circadian Clock. *Plant physiology*, 155(2), 906–915. <https://doi.org/10.1104/pp.110.167015>

- Luo, M., Cheng, K., Xu, Y., Yang, S. y Wu, K. (2017). Plant Responses to Abiotic Stress Regulated by Histone Deacetylases. *Frontiers in Plant Science*, 8. <https://doi.org/10.3389/fpls.2017.02147>
- Luo, M., Wang, Y.-Y., Liu, X., Yang, S., Lu, Q., Cui, Y. y Wu, K. (2012). *HD2C* interacts with *HDA6* and is involved in ABA and salt stress response in Arabidopsis. *Journal of Experimental Botany*, 63(8), 3297–3306. <https://doi.org/10.1093/jxb/ers059>
- Ma, X., Lv, S., Zhang, C. y Yang, C. (2013). Histone deacetylases and their functions in plants. *Plant Cell Reports*, 32(4), 465–478. <https://doi.org/10.1007/s00299-013-1393-6>
- Maddan, T. (2013) The BLAST sequence analysis tool. NCBI (2019) *The NCBI Handbook second edition*, 361 – 370.
- Mandaokar, A., Thines, B., Shin, B., Markus Lange, B., Choi, G., Koo, Y. J., ... Browse, J. (2006). Transcriptional regulators of stamen development in Arabidopsis identified by transcriptional profiling. *The Plant Journal*, 46(6), 984–1008. <https://doi.org/10.1111/j.1365-313x.2006.02756.x>
- March-Díaz, R., García-Domínguez, M., Lozano-Juste, J., León, J., Florencio, F. J., y Reyes, J. C. (2007). Histone H2A.Z and homologues of components of the *SWR1* complex are required to control immunity in Arabidopsis. *The Plant Journal*, 53(3), 475–487. <https://doi.org/10.1111/j.1365-313x.2007.03361.x>
- Martins, D., de Sousa Araújo, S., Rubiales, D. y Vaz Patto, M. C. (2020). Legume Crops and Biotrophic Pathogen Interactions: A Continuous Cross-Talk of a Multilayered Array of Defense Mechanisms. *Plants* (11), 1460. <https://doi.org/10.3390/plants9111460>
- McCarty, J. P. (2001). Consecuencias Biológicas de Cambios Climáticos Recientes. *Conservation Biology*, 15(2), 320–331. <https://doi.org/10.1046/j.1523-1739.2001.015002320.x>
- McClung, C. R. (2006). Plant circadian rhythms. *Plant Cell*, 18(4), 792–803. <https://doi.org/10.1105/tpc.106.040980>
- Melmaiee, K., Kalavacharla, V., Brown, A., Todd, A., Thurston, Y. y Elevarthi, S. (2015). Quantification and Gene Expression Analysis of Histone Deacetylases in Common

Bean during Rust Fungal Inoculation. *International Journal of Genomics*.
<https://doi.org/10.1155/2015/153243>

Meng, Q., Zhang, C., Huang, F., Gai, J. y Yu, D. (2007). Molecular cloning and characterization of a LEAFY-like gene highly expressed in developing soybean seeds. *Seed Science Research* 17, 297–302.

Milford, G. F. J., Day, J. M., Leach, J. E., Stevenson, H. J., Huyghe, C. y Papineau, J. (1993). The effect of modifying plant structure on the yield and maturity of the white lupin (*Lupinus albus*). *Annals of Applied Biology*, 122(1), 113–122.
<https://doi.org/10.1111/j.1744-7348.1993.tb04019.x>

MINAGRI (9 de febrero de 2020). Minagri celebra el día mundial de las legumbres con exhibición de variedades en Agroferia campesina. Gob.pe. Recuperado de <https://www.gob.pe/institucion/minagri/noticias/81361-minagri-celebro-dia-mundial-de-las-legumbres-con-exhibicion-de-variedades-en-agroferia-campesina>

Monterroso, V. A. y Wien, H. C. (1990). Flower and pod abscission due to heat stress in beans. *Journal of the American Society for Horticultural Science* 115, 631–634.
<https://doi.org/10.21273/JASHS.115.4.631>

Moore, G., Devos, K.M., Wang, Z. y Gale, M. D. (1995). Grasses, line up and form a circle. *Current biology*, 5(7), 737–739. <https://doi.org/10.1029/88EO01170>

Morel, P., Tréhin, C., Breuil-Broyer, S. y Negrutiu, I. (2008). Altering FVE/MSI4 results in a substantial increase of biomass in *Arabidopsis*—the functional analysis of an ontogenesis accelerator. *Molecular Breeding*, 23(2), 239–257. <https://doi.org/10.1007/s11032-008-9229-8>

Munier-Jolain, N. G. y Ney, B. (1998). Seed growth rate in grain legumes II. Seed growth rate depends on cotyledon cell number. *Journal of Experimental Botany*, 49(329), 1971–1976. <https://doi.org/10.1093/jxb/49.329.1971>

Nadeem, M., Li, J., Yahya, M., Sher, A., Ma, C., Wang, X. y Qiu, L. (2019). Research Progress and Perspective on Drought Stress in Legumes: A Review. *International Journal of Molecular Sciences*, 20(10), 2541. <https://doi.org/10.3390/ijms20102541>

- Nayyar, H., Kaur, G., Kumar, S. y Upadhyaya, H. D. (2007). Low Temperature Effects during Seed Filling on Chickpea Genotypes (*Cicer arietinum* L.): Probing Mechanisms affecting Seed Reserves and Yield. *Journal of agronomy and crop science* Vol. 193(5), 336-344. <https://doi.org/10.1111/j.1439-037X.2007.00269.x>
- Nelson, M. N., Berger, J. D. y Erskine, W. (2010). Flowering time control in annual legumes: Prospects in a changing global climate. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 5(017). <https://doi.org/10.1079/PAVSNR20105017>
- Ni, Z., Kim, E. D., Ha, M., Lackey, E., Liu, J., Zhang, Y., ... Chen, Z. J. (2009). Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature*, 457(7227), 327–331. <https://doi.org/10.1038/nature07523>
- Nichio, B. T. L., Marchaukoski, J. N. y Raittz, R. T. (2017). New tools in orthology analysis: A brief review of promising perspectives. *Frontiers in Genetics*, 8(OCT), 1–12. <https://doi.org/10.3389/fgene.2017.00165>
- Nilsson, O., Lee, I., Blázquez, M. A. y Weigel, D. (1998). Flowering-Time Genes Modulate the Response to LEAFY Activity. *Genetics* Vol. 150 (1), 403 – 410.
- Nowell, R. W., Laue, B. E., Sharp, P. M. y Green, S. (2016). Comparative genomics reveals genes significantly associated with woody hosts in the plant pathogen *Pseudomonas syringae*. *Molecular Plant Pathology*, 17(9), 1409–1424. <https://doi.org/10.1111/mpp.12423>
- Oh, S., Zhang, H., Ludwig, P. y Van Nocker, S. (2004). A Mechanism Related to the Yeast Transcriptional Regulator Paf1c Is Required for Expression of the Arabidopsis FLC/MAF MADS Box Gene Family. *The plant cell online*, 16(11), 2940–2953. <https://doi.org/10.1105/tpc.104.026062>
- Ohto, M. A., Fischer, R. L., Goldberg, R. B., Nakamura, K. y Harada, J. J. (2005). Control of seed mass by APETALA2. *Proceedings of the National Academy of Sciences of the United States of America*, 102(8), 3123–3128. <https://doi.org/10.1073/pnas.0409858102>

- Oshima, S. y Nomura, K. (2008). *RsLFY*, a *LEAFY* homologue gene in radish (*Raphanus sativus*), is continuously expressed in vegetative, reproductive and seed development. *Plant Biotechnology* 25 (6). <https://doi.org/10.5511/plantbiotechnology.25.579>
- Pearson, W. R. (2013). An introduction to sequence similarity (“homology”) searching. *Current Protocols in Bioinformatics*, (SUPPL.42), 1–8. <https://doi.org/10.1002/0471250953.bi0301s42>
- Perales, L., Rodríguez, P., Valero, T., Ruiz, E., Ávila, J. M. y Valera, G. (2017). Informe sobre Legumbres, Nutrición y Salud. *Fundacion Española de Nutrición*, 94. Recuperado de http://www.aecosan.msssi.gob.es/AECOSAN/docs/documentos/noticias/2017/Informe_Legumbres_Nutricion_Salud.pdf
- Pfab, A., Breindl, M. y Grasser, K. D. (2018). The Arabidopsis histone chaperone *FACT* is required for stress-induced expression of anthocyanin biosynthetic genes. *Plant Molecular Biology*, 96(4-5), 367–374. <https://doi.org/10.1007/s11103-018-0701-5>
- Phang, TH., Shao, G. y Lam, HM. (2008) Salt tolerance in soybean. *Journal of Integrative Plant Biology* 50, 1196-1212.
- Pin, P. A. y Nilsson, O. (2012). The multifaceted roles of *FLOWERING LOCUS T* in plant development. *Plant, Cell y Environment*, 35(10), 1742–1755. <https://doi.org/10.1111/j.1365-3040.2012.02558.x>
- QIAGEN (2021) BLAST Search. En *Manual for CLC Genomics Workbench 21.0.4* (pp. 310 – 316). Dinamarca. QIAGEN Digital Insights
- Rehman, H. M., Nawaz, M. A., Shah, Z. H., Ludwig-Müller, J., Chung, G., Ahmad, M. Q., ... Lee, S. I. (2018). Comparative genomic and transcriptomic analyses of Family-1 UDP glycosyltransferase in three Brassica species and Arabidopsis indicates stress-responsive regulation. *Scientific Reports*, 8(1), 1–18. <https://doi.org/10.1038/s41598-018-19535-3>
- Rosenzweig, C., Elliot, J., Deryng, D., Ruane, A.C., Muller, C., Arneth, A., ... Jones, J. W. (2013). Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *PNAS Vol. 111*, 3268–3273. <https://doi.org/10.1073/pnas.1222463110>

- Rubin, G. M. (2000). Comparative Genomics of the Eukaryotes. *Science*, 287(5461), 2204–2215. <https://doi.org/10.1126/science.287.5461.2204>
- Ryu, J. Y., Park, C.-M. y Seo, P. J. (2011). The floral repressor BROTHER OF FT AND TFL1 (BFT) modulates flowering initiation under high salinity in Arabidopsis. *Molecules and Cells*, 32(3), 295–303. <https://doi.org/10.1007/s10059-011-0112-9>
- Saitou, N. y Imanishi, T. (1989). Relative Efficiencies of the Fitch-Margoliash, Maximum-Parsimony, Maximum-Likelihood, Minimum-Evolution, and Neighbor-joining Methods of Phylogenetic Tree Construction in Obtaining the Correct Tree. *Molecular Biology and Evolution*, (3). <https://doi.org/10.1093/oxfordjournals.molbev.a040572>
- Sánchez, J.-P., Duque, P. y Chua, N.-H. (2004). ABA activates ADPR cyclase and cADPR induces a subset of ABA-responsive genes in Arabidopsis. *The Plant Journal*, 38(3), 381–395. <https://doi.org/10.1111/j.1365-313x.2004.02055.x>
- Savadi, S. (2017). Molecular regulation of seed development and strategies for engineering seed size in crop plants. *Plant Growth Regulation*, 84(3), 401–422. <https://doi.org/10.1007/s10725-017-0355-3>
- Schlueter, J. A., Dixon, P., Granger, C., Grant, D., Clark, L., Doyle, J. J. y Shoemaker, R. C. (2004). Mining EST databases to resolve evolutionary events in major crop species. *Genome*, 47(5), 868–876. <https://doi.org/10.1139/g04-047>
- Schmidhuber, J. y Tubiello, F. N. (2007). Global food security under climate change. *PNAS*, Vol. 10, 19703–19708. <https://doi.org/10.1073/pnas.0701976104>
- Shah, S., Karunarathna, N. L., Jung, C., y Emrani, N. (2018). An APETALA1 ortholog affects plant architecture and seed yield component in oilseed rape (*Brassica napus* L.). *BMC Plant Biology*, 18(1), 1–12. <https://doi.org/10.1186/s12870-018-1606-9>
- Sharma, A. y Chauhan, R. S. (2012). In Silico Identification and Comparative Genomics of Candidate Genes Involved in Biosynthesis and Accumulation of Seed Oil in Plants. *Comparative and Functional Genomics*, 2012, 1–14. <https://doi.org/10.1155/2012/914843>
- Simon, R., Igeño, M. I. y Coupland, G. (1996). Activation of floral meristem identity genes in Arabidopsis. *Nature* 384 (6604), 59 – 62. <https://doi.org/10.1038/384059a0>

- Singh, V., Roy, S., Singh, D. y Nandi, A. K. (2014). Arabidopsis *FLOWERING LOCUS D* influences systemic-acquired-resistance- induced expression and histone modifications of *WRKY* genes. *Journal of Biosciences*, *39(1)*, 119–126. <https://doi.org/10.1007/s12038-013-9407-7>
- Smith, A. R. y Zhao, D. (2016). Sterility Caused by Floral Organ Degeneration and Abiotic Stresses in Arabidopsis and Cereal Grains. *Frontiers in Plant Science*, *7*. <https://doi.org/10.3389/fpls.2016.01503>
- Smýkal, P., Coyne, C. J., Ambrose, M. J., Maxted, N., Schaefer, H., Blair, M. W., ... Varshney, R. K. (2015). Legume Crops Phylogeny and Genetic Diversity for Science and Breeding. *Critical Reviews in Plant Sciences*, *34(January)*, 43–104. <https://doi.org/10.1080/07352689.2014.897904>
- Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., ... Soltis, P. S. (2009). Polyploidy and angiosperm diversification. *American Journal of Botany*, *96(1)*, 336–348. <https://doi.org/10.3732/ajb.0800079>
- Storm, C. E. V. y Sonnhammer, E. L. L. (2002). Trees and Calculation of Orthology Reliability. *Genomics*, *18(1)*, 92–99.
- Su, Y., Wang, S., Zhang, F., Zheng, H., Liu, Y., Huang, T. y Ding, Y. (2017). Phosphorylation of Histone *H2A* at Serine 95: A Plant-Specific Mark Involved in Flowering Time Regulation and *H2A.Z* Deposition. *The Plant Cell*, *29(9)*, 2197–2213. <https://doi.org/10.1105/tpc.17.00266>
- Sun, S., Zhi, Y. y Zhu, Z. (2016). An Emerging Disease Caused by *Pseudomonas syringae* pv. *phaseolicola* Threatens Mung Bean Production in China. *Plant disease* *101*, 95 - 102. <https://doi.org/10.1094/PDIS-04-16-0448-RE>
- Sun, X., Shantharaj, D., Kang, X. y Ni, M. (2010). Transcriptional and hormonal signaling control of Arabidopsis seed development. *Current Opinion in Plant Biology*, *13(5)*, 611–620. <https://doi.org/10.1016/j.pbi.2010.08.009>
- Syngai, G. G., Barman, P., Bharali, R. y Dey, S. (2013). BLAST: An introductory tool for students to bioinformatics applications. *Keanean Journal of Science Vol 2* ,67 – 76.

- Tamura, K., Battistuzzi, F. U., Billings-Ros, P., Murillo, O., Filipinski, A. y Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *PNAS* 109(47), 19333–19338. <https://doi.org/10.1073/pnas.1213199109>
- Tanaka, M., Kikuchi, A. y Kamada, H. (2008). The Arabidopsis Histone Deacetylases *HDA6* and *HDA19* Contribute to the Repression of Embryonic Properties after Germination. *Plant physiology*, 146(1), 149–161. <https://doi.org/10.1104/pp.107.111674>
- Tatusova, T., Zaslavsky, L., Fedorov, B., Haddad, D., Vatsan, A., Ako-adjei, D., ...Ghazal, H. (2019). Protein Clustering. En NCBI (Ed.) *The NCBI Handbook second edition*, (pp. 331 – 342).USA. NCBI
- Taylor, J. D., Teverson, D. M., Allenm D. J. y Pastor-Corrales, M. A. (1996). Identification and origin of races of *Pseudomonas syringae* pv. *phaseolicola* from Africa and other bean growing areas. *Plant Pathology Vol 45 (3)* , 469 - 478. <https://doi.org/10.1046/j.1365-3059.1996.d01-147.x>
- The Legume Phylogeny Working Group ; Doyle, J. J., Herendeen, P., Hughes, C., Kenicer, G., Lewis, G., Mackinder, B., ... Simon, M. F. (2013). Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon*, 62(2), 217–248. <https://doi.org/10.12705/622.8>
- Thines, B. C., Youn, Y., Duarte, M. I. y Harmon, F. G. (2014). The time of day effects of warm temperature on flowering time involve *PIF4* and *PIF5*. *Journal of Experimental Botany*, 65(4), 1141–1151. <https://doi.org/10.1093/jxb/ert487>
- Tremblay, R. y Colsanti, J. (2007). Floral induction. En Ainsworth, C. (Ed.), *Flowering and its Manipulation*, (pp. 28-48). Reino Unido. Blackwell Publishing.
- Ung, N., Lal, S. y Smith, H. M. S. (2011). The role of *PENNYWISE* and *POUND-FOOLISH* in the maintenance of the shoot apical meristem in arabidopsis. *Plant Physiology*, 156(2), 605–614. <https://doi.org/10.1104/pp.110.171462>
- UniProt (2021). Recuperado 25 de mayo de 2021. UniProtKB - Q941Q3 (C3H37_ARATH). <https://www.uniprot.org/uniprot/Q941Q3>
- UniProt. (2021). Recuperado 25 de mayo de 2021, de UniProtKB - Q9FEN9 (SHL_ARATH) website: <https://www.uniprot.org/uniprot/Q9FEN9>

- Upadhyaya, H. D., Dwivedi, S. L., Ambrose, M., Ellis, N., Berger, J., Smýkal, P., ... Gowda, C. L. L. (2011). Legume genetic resources: Management, diversity assessment, and utilization in crop improvement. *Euphytica*, *180*(1), 27–47. <https://doi.org/10.1007/s10681-011-0449-3>
- Van Dijk, A. D. J. y Molenaar, J. (2017). Floral pathway integrator gene expression mediates gradual transmission of environmental and endogenous cues to flowering time. *PeerJ*, *5*, e3197. <http://doi.org/10.7717/peerj.3197>
- Van Schoonhoven, A. y Voysest, O. (1989) Common beans in Latin America and their constraints. En CIAT (Ed.), *Bean production problems in the tropics*, (pp. 33 – 57). 2nd edition. Colombia.CIAT.
- Vance, C. P., Graham, P. H. y Allan, D. L. (2000). Biological Nitrogen Fixation: Phosphorus - A Critical Future Need?. En Pedrosa, F. O, Hungria, M., Yates, G. y Newton, W. E. (Ed.), *Nitrogen Fixation: From Molecules to Crop Productivity*, (pp. 509–514). USA. Kluwer Academic Publisher.
- Veitia, R. A. (2005). Paralogs in polyploids: One for all and all for one? *Plant Cell*, *17*(1), 4–11. <https://doi.org/10.1105/tpc.104.170130>
- Wagner, I., Volkmer, M., Sharan, M., Villaveces, J. M., Oswald, F., Surendranath, V. y Habbermann, B. H. (2004). morFeus: a web-based program to detect remotely conserved orthologs using symmetrical best hits and orthology network scoring. *BMC Bioinformatics*, 15-263. <https://doi.org/10.1186/1471-2105-15-263>
- Wahl, V., Ponnu, J., Schlereth, A., Arrivault, S., Langenecker, T., Franke, A., ... Schmid, M. (2013). Regulation of flowering by trehalose-6-phosphate signaling in *Arabidopsis thaliana*. *Science (February)*, 704–708. <https://doi.org/10.1126/science.1230406>.
- Wang, J. W., Czech, B. y Weigel, D. (2009). miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in *Arabidopsis thaliana*. *Cell*, *138*(4), 738–749. <https://doi.org/10.1016/j.cell.2009.06.014>
- Wang, Z., Cao, H., Chen, F., y Liu, Y. (2014). The roles of histone acetylation in seed performance and plant development. *Plant Physiology and Biochemistry*, *84*, 125–133. <https://doi.org/10.1016/j.plaphy.2014.09.010>

- Wang, Z.-W., Wu, Z., Raitskin, O., Sun, Q. y Dean, C. (2014). Antisense-mediated FLC transcriptional repression requires the P-TEFb transcription elongation factor. *Proceedings of the National Academy of Sciences*, *111*(20), 7468–7473. <https://doi.org/10.1073/pnas.1406635111>
- Wei, Q., Ma, C., Xu, Y., Wang, T., Chen, Y., Lü, J., ... Gao, J. (2017). Control of chrysanthemum flowering through integration with an aging pathway. *Nature Communications*, *8*(1). <https://doi.org/10.1038/s41467-017-00812-0>
- Weller, S. G. y Sakai, A. K. (1999). Using Phylogenetic Approaches for the Analysis of Plant Breeding System Evolution. *Annual Review of Ecology and Systematics*, *30*, 167–199. <https://doi.org/10.1146/annurev.ecolsys.30.1.167>
- Whelan, S., Liò, P. y Goldman, N. (2001). Molecular phylogenetics: state-of-the-art methods for looking into the past. *Trends in Genetics*, *17*(5), 262–272. [https://doi.org/10.1016/s0168-9525\(01\)02272-7](https://doi.org/10.1016/s0168-9525(01)02272-7)
- Wiersema, J. H. (2019). GRIN Taxonomy. US National Plant Germplasm System. Checklist dataset <https://doi.org/10.15468/ao14pp> accessed via GBIF.org on 2021-08-30.
- William, D. A., Su, Y., Smith, M. R., Lu, M., Baldwin, D. A. y Wagner, D. (2004). Genomic identification of direct targets genes of LEAFY. *PNAS* *101*(6), 1775 - 1780. <https://doi.org/10.1073/pnas.0307842100>
- Willmann, M. R. y Poethig, R. S. (2011). The effect of the floral repressor *FLC* on the timing and progression of vegetative phase change in Arabidopsis. *Development*, *138*(4), 677–685. <https://doi.org/10.1242/dev.057448>
- Wu, J., Yamaguchi, N. y Ito, T. (2019). Histone demethylases control root elongation in response to stress-signaling hormone abscisic acid. *Plant Signaling y Behavior*, 1–4. <https://doi.org/10.1080/15592324.2019.1604019>
- Wu, K., Zhang, L., Zhou, C., Yu, C.-W. y Chaikam, V. (2008). *HDA6* is required for jasmonate response, senescence and flowering in Arabidopsis. *Journal of Experimental Botany*, *59*(2), 225–234. <https://doi.org/10.1093/jxb/erm300>

- Xue, M., Yang, J., Li, Z., Hu, S., Yao, N., Dean, R. A., ... Peng, Y.-L. (2012). Comparative Analysis of the Genomes of Two Field Isolates of the Rice Blast Fungus *Magnaporthe oryzae*. *PLoS Genetics*, 8(8), e1002869. <https://doi.org/10.1371/journal.pgen.1002869>
- Yan, Y., Shen, L., Chen, Y., Bao, S., Thong, Z. y Yu, H. (2014). A MYB-Domain Protein EFM Mediates Flowering Responses to Environmental Cues in Arabidopsis. *Developmental Cell*, 30(4), 437–448. <https://doi.org/10.1016/j.devcel.2014.07.004>
- Yang, J., Chai, X.-Q., Zhao, X.-X. y Li, X. (2016). Comparative genomics revealed the origin and evolution of autophagy pathway. *Journal of Systematics and Evolution*, 55(1), 71–82. <https://doi.org/10.1111/jse.12212>
- Yoo, S. K. (2005). *CONSTANS* Activates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* through *FLOWERING LOCUS T* to Promote Flowering in Arabidopsis. *Plant physiology*, 139(2), 770–778. <https://doi.org/10.1104/pp.105.066928>
- Yu, X. y Michaels, S. D. (2010). The Arabidopsis *Paf1c* Complex Component *CDC73* Participates in the Modification of *FLOWERING LOCUS C* Chromatin. *Plant physiology*, 153(3), 1074–1084. <https://doi.org/10.1104/pp.110.158386>
- Yuvaraj, M., Pandiyan, M. y Gayathri, P. (2020). Role of Legumes in Improving Soil Fertility Status, Legume Crops - Prospects, Production and Uses, Mirza Hasanuzzaman, IntechOpen, <https://doi.org/10.5772/intechopen.93247>. Available from: <https://www.intechopen.com/books/legume-crops-prospects-production-and-uses/role-of-legumes-in-improving-soil-fertility-status>
- Zacharaki, V., Benhamed, M., Poulis, S., Latrasse, D., Papoutsoglou, P., Delarue, M. y Vlachonassios, K. E. (2012). The Arabidopsis ortholog of the *YEATS* domain containing protein *YAF9a* regulates flowering by controlling H4 acetylation levels at the *FLC* locus. *Plant Science*, 196, 44–52. <https://doi.org/10.1016/j.plantsci.2012.07.010>
- Zhang, H., Ransom, C., Ludwig, P. y Van Nocker, S. (2003). Genetic analysis of early flowering mutants in Arabidopsis defines a class of pleiotropic developmental regulator required for expression of the flowering-time switch flowering locus C. *Genetics* 164(1), 347 – 358.

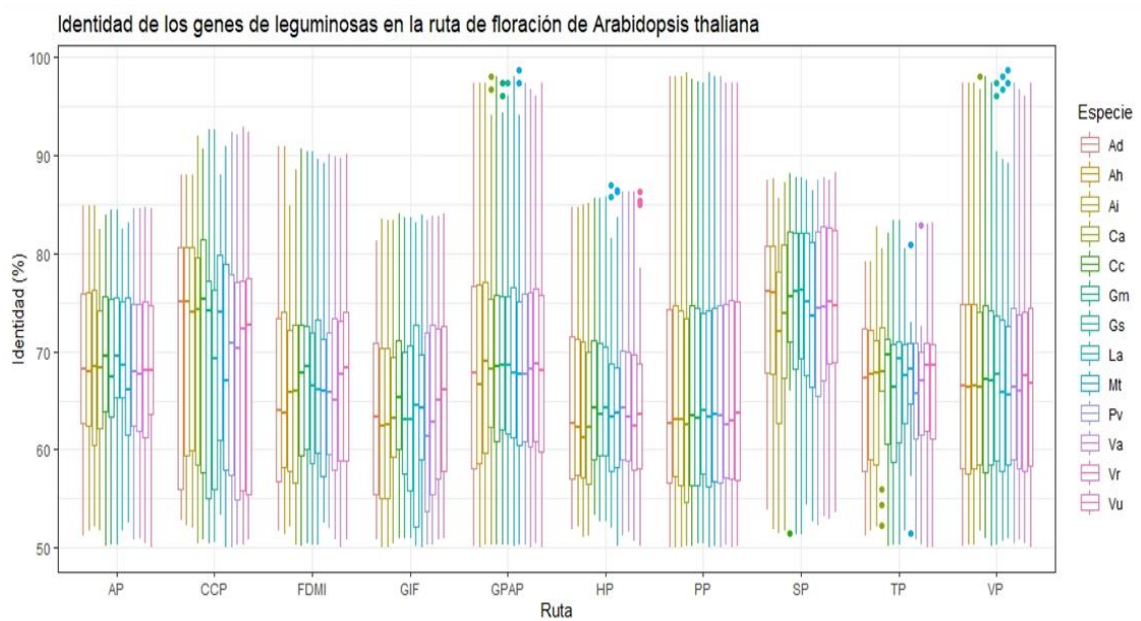
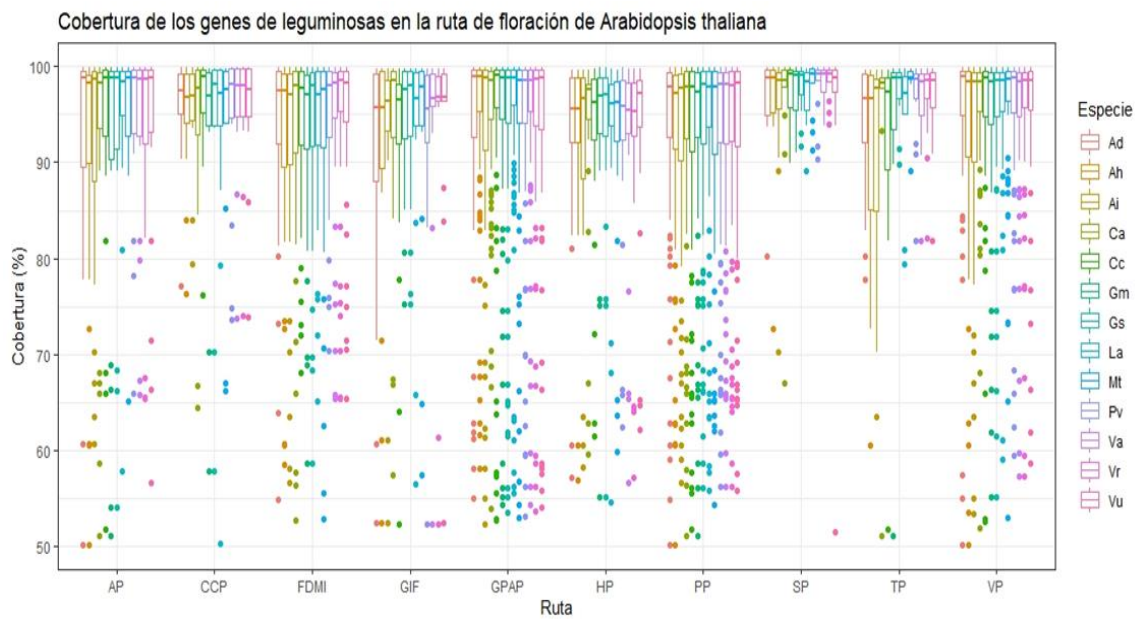
- Zhang, K., Wang, X. y Cheng, F. (2019). Plant Polyploidy: Origin, Evolution, and Its Influence on Crop Domestication. *Horticultural Plant Journal*, 5(6), 231–239. <https://doi.org/10.1016/j.hpj.2019.11.003>
- Zhang, N., Wang, M., Huang, J., Yang, L., Wang, Z., Wu, D. y Shu, X. (2020). *Mos1* negatively regulates sugar responses and anthocyanin biosynthesis in Arabidopsis. *International Journal of Molecular Sciences*, 21(19), 1–16. <https://doi.org/10.3390/ijms21197095>
- Zhang, Y. B., Yuan, Y., Pang, Y. X., Yu, F. L., Yuan, C., Hu, X. y Wang, D. (2019). Phylogenetic reconstruction and divergence time estimation of blumea dc. (Asteraceae: Inuleae) in China Based on nrDNA its and cpDNA trnL-F sequences. *Plants*, 8(7). <https://doi.org/10.3390/plants8070210>
- Zhao, J., Lu, Z., Wang, L. y Jin, B. (2020). Plant Responses to Heat Stress: Physiology, Transcription, Noncoding RNAs, and Epigenetics. *International Journal of Molecular Science* 22(1). <https://doi.org/10.3390/ijms22010117>
- Zhao, L., Li, Y., Xie, Q. y Wu, Y. (2017). Loss of *CDKC ;2* increases both cell division and drought tolerance in Arabidopsis thaliana. *The Plant Journal*, 91(5), 816–828. <https://doi.org/10.1111/tpj.13609>
- Zhao, M., Yang, S., Liu, X. y Wu, K. (2015). Arabidopsis histone demethylases *LDL1* and *LDL2* control primary seed dormancy by regulating *DELAY OF GERMINATION 1* and ABA signaling-related genes. *Frontiers in Plant Science*, 6(MAR), 1–10. <https://doi.org/10.3389/fpls.2015.00159>
- Zhu, Y., Huang, P., Guo, P., Chong, L., Yu, G., Sun, X., ... Zhu, J. (2020). *CDK8* is associated with *RAP2.6* and *SnRK2.6* and positively modulates abscisic acid signaling and drought response in Arabidopsis. *New Phytologist*. <https://doi.org/10.1111/nph.16787>
- Zilio, M., Souza, C. A., Coelho, C. M. M., Miquelluti, D. J. y Michels, A. F. (2013). Cycle, canopy architecture and yield of common bean genotypes (*Phaseolus vulgaris*) in Santa Catarina State. *Acta Scientiarum. Agronomy*, 35(1), 21 – 30. <https://doi.org/10.4025/actasciagron.v35i1.15516>

Zou, Y., Wang, S., Zhou, Y., Bai, J., Huang, G., Liu, X., ... Lu, D. (2018). Transcriptional regulation of the immune receptor *FLS2* controls the ontogeny of plant innate immunity. *Plant Cell*, 30(11), 2779–2794. <https://doi.org/10.1105/tpc.18.00297>

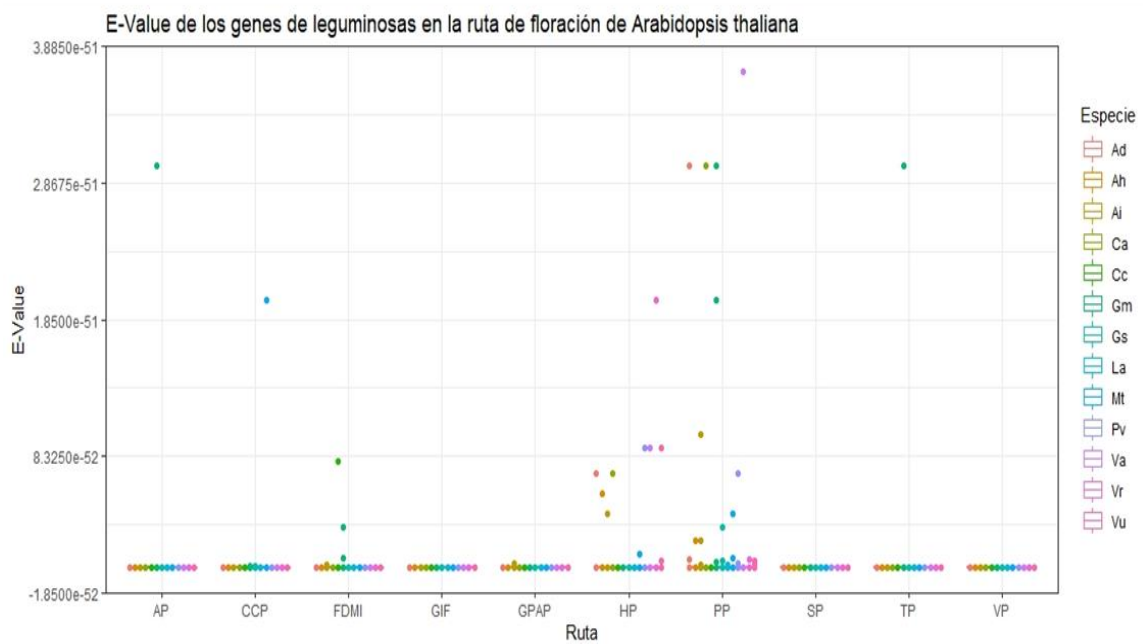
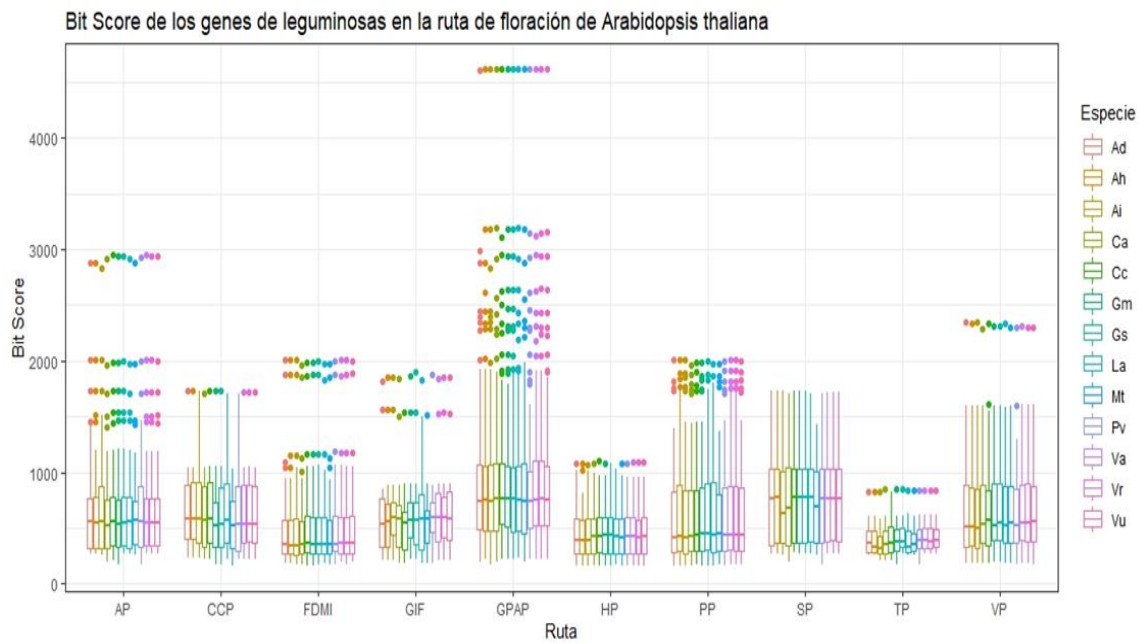
Zuckerkindl, E. y Pauling, L. (1965). Evolutionary divergence and convergence in proteins. En Bryson, V. y Vogel, H. J. (Ed.), *Evolving genes and proteins* (pp. 97-166) . USA. Academic Press.

VIII. ANEXOS

Anexo 1. Variabilidad de los parámetros de cobertura e identidad de proteínas homologas a la ruta de floración.



Anexo 2. Variabilidad de los parámetros de bit score e E-value de proteínas homologas a la ruta de floración.



Anexo 3. Genes ortólogos únicos de la floración en 13 leguminosas de interés económico

<i>Arabidopsis thaliana</i>	<i>Arachis ipaensis</i>	<i>Arachis hypogea</i>	<i>Arachis duranensis</i>	<i>Vigna radiata</i>	<i>Vigna angularis</i>	<i>Vigna unguiculata</i>
AT3G20810.2	XP_016187154.1	XP_025642133.1	XP_015952169.1	XP_014493266.1	XP_017431559.1	XP_027910196.1
AT3G07650.1	Araip_0046	Arahy_0059	Aradu_0023	Vigra_0043	Vigan_0038	Vigun_0040
AT5G61850.1	XP_016169153.1	XP_025672113.1	XP_015937724.1	XP_022633651.1	XP_017441945.1	XP_027905350.1
AT3G12680.1	XP_016170465.1	XP_025615097.1	XP_015936930.1	XP_014490221.1	XP_017433251.1	XP_027910843.1
AT3G57130.1	XP_016205093.1	XP_025660632.1	XP_015969225.1	XP_014514599.1	XP_017412674.1	XP_027934922.1
AT5G63610.1	XP_016170961.1	Arahy_0045	XP_015937545.1	XP_014509222.1	Vigan_0033	XP_027919065.1
AT4G31120.1	XP_016171040.1	XP_025672235.1	Aradu_0028	XP_014502375.1	XP_017419811.1	XP_027903914.1
AT3G45880.1	XP_016199001.1	Arahy_0034	XP_015963431.1	XP_014520966.1	XP_017427378.1	XP_027913594.1
AT4G39100.1	XP_016166727.1	Arahy_0014	XP_015932760.1	XP_014502099.1	XP_017422602.1	XP_027903212.1
AT4G04920.1	XP_016189784.1	XP_025690083.1	XP_015956477.1	XP_022639304.1	XP_017435676.1	XP_027932443.1
AT1G52740.1	XP_016162353.1	Arahy_0078	XP_015972006.1	XP_014494381.1	XP_017422643.1	XP_027904762.1
AT5G62040.1	XP_016200998.1	XP_025701974.1	XP_015967277.2	XP_014503123.1	XP_017438938.1	XP_027923489.1
AT3G20740.1	XP_016168099.1	Arahy_0016	XP_015936492.1	XP_014510715.1	XP_017409201.1	XP_027922072.1
AT2G19520.1	Araip_0037	Arahy_0042	XP_015954449.1	XP_014502046.1	XP_017421558.1	XP_027904592.1
AT5G45600.1	XP_016186887.1	Arahy_0093	XP_015951909.1	Vigra_0072	Vigan_0068	Vigun_0069
AT3G10390.1	Araip_0029	Arahy_0031	Aradu_0011	XP_014513819.1	XP_017415090.1	XP_027935477.1
AT1G79730.1	XP_020964125.1	XP_025617781.1	XP_015938431.1	XP_014516734.1	XP_017407652.1	XP_027907846.1
AT5G63110.1	XP_016190895.1	XP_025629059.1	XP_015970912.1	XP_014490556.1	XP_017427570.1	XP_027922126.1
AT4G29830.1	XP_016194581.1	Arahy_0064	XP_015962490.1	XP_014505934.1	XP_017427983.1	XP_027922288.1
AT5G64960.1	XP_016166475.1	Arahy_0079	XP_015973469.1	XP_014522149.1	XP_017439681.1	XP_027940749.1
AT3G28730.1	XP_016180395.1	XP_025624058.1	XP_015946008.1	XP_014493594.1	Vigan_0067	XP_027924856.1

<i>Arabidopsis thaliana</i>	<i>Lupinus angustifolius</i>	<i>Phaseolus vulgaris</i>	<i>Cajanus cajan</i>	<i>Glycine max</i>	<i>Glycine soja</i>	<i>Medicago truncatula</i>	<i>Cicer arietinum</i>
AT3G20810.2	XP_019413308.1	XP_007131982.1	XP_020214973.1	XP_003539394.1	XP_028193886.1	XP_013456231.1	XP_004507119.1
AT3G07650.1	Lupan_0036	XP_007141372.1	XP_020231086.1	Glyma_0052	Glyso_0048	Meditru_0032	Cicar_0041
AT5G61850.1	XP_019458009.1	XP_007137848.1	XP_020224595.1	XP_025984699.1	XP_028227085.1	XP_003602745.1	XP_004501760.1
AT3G12680.1	XP_019449304.1	XP_007133464.1	XP_020207375.1	XP_006585719.1	XP_028204321.1	XP_013445292.1	XP_004517028.1
AT3G57130.1	XP_019458812.1	XP_007162101.1	XP_020224348.1	NP_001276282.1	XP_028216047.1	XP_003625012.1	XP_004514968.1
AT5G63610.1	Lupan_0026	Phavul_0010	XP_029125272.1	XP_003525137.1	XP_028229594.1	XP_003602688.2	XP_004502994.1
AT4G31120.1	Lupan_0049	XP_007137105.1	XP_020228398.1	Glyma_0066	XP_028228109.1	XP_003603517.1	XP_004501256.1
AT3G45880.1	XP_019445141.1	XP_007158056.1	XP_020207278.1	XP_003538923.1	XP_028189597.1	XP_003612456.1	XP_004512306.1
AT4G39100.1	XP_019438180.1	XP_007135939.1	XP_020218862.1	NP_001236094.1	XP_028227287.1	XP_003603985.1	XP_004500782.1
AT4G04920.1	Lupan_0068	XP_007159313.1	XP_020222881.1	XP_003543060.2	XP_028188040.1	XP_013457064.1	XP_012572579.1
AT1G52740.1	Lupan_0055	XP_007138323.1	XP_020207935.1	NP_001235400.1	XP_028228890.1	Meditru_0052	XP_004508547.1
AT5G62040.1	XP_019450450.1	XP_007152319.1	Cajca_0065	Glyma_0076	Glyso_0068	XP_013444336.1	Cicar_0062
AT3G20740.1	XP_019458295.1	Phavul_0028	Cajca_0033	XP_006588358.1	XP_028185981.1	Meditru_0061	XP_004498406.1
AT2G19520.1	XP_019415490.1	XP_007138727.1	XP_020224017.1	XP_003533752.1	XP_028180790.1	XP_003595171.1	XP_004488042.1
AT5G45600.1	XP_019413672.1	XP_007131738.1	Cajca_0072	XP_003540935.1	XP_028192689.1	Meditru_0039	Cicar_0069
AT3G10390.1	XP_019449823.1	Phavul_0006	Cajca_0023	XP_003520261.1	XP_028207900.1	XP_013467388.1	Cicar_0021
AT1G79730.1	XP_019430327.1	XP_007135633.1	XP_020227957.1	XP_025984926.1	XP_028238892.1	XP_003627041.2	XP_012574096.1
AT5G63110.1	XP_019457189.1	XP_007155467.1	XP_020213802.1	XP_003525556.1	XP_028231550.1	XP_003601202.1	XP_004501884.1
AT4G29830.1	XP_019463682.1	XP_007154616.1	XP_020226934.1	XP_003551163.1	XP_028211279.1	XP_013458438.1	XP_004507970.1
AT5G64960.1	XP_019460261.1	XP_007146195.1	XP_020231471.1	XP_003555471.1	XP_028221949.1	XP_013459476.1	XP_004489978.1
AT3G28730.1	Lupan_0067	XP_007151197.1	XP_020232756.1	XP_003517023.1	XP_028224253.1	XP_003616631.1	XP_004499164.1

XP:	Secuencia de referencia de NCBI
NP:	Secuencia de referencia de NCBI
	Ver anexo 4

Anexo 4. Secuencias de referencia para el anexo 3.

>Aradu_0011

MDSPSEFSSFPSPQTQLPPQONPNPIVFLDPNSDSFPNWAENPNPNLQNYLSFTVPKKRRRGR
SQRNPASFRIPDPNDASNSNNNNNGFSSSAAAVPVSAPLSKVVIGNPPPLSMQYSVPDNDSEII
VINKEPKTEALIALSAGFPADSLTEEEIEAGVLPVIGGIEQVNYTLIRNHIIAKWRENVSQWVSKEMFV
DYIPQHYHALLDSAYNYLVSHGYINFGVAPLIKEKIPLMSKPGVIIIAGLAGLAAARQLMRFGFKV
TVLEGRKRAGGRVYTKKMEGGNRVCAAADLGGSVLTGTLGNPLGIVARQLGDVHLKVRDKCPLY
SLDGKVPDPMVVKVENAFNRLLDKASRLRQLMGEVSMDSVSLGAALETFRQVYKDAVNDEEMNL
FNWHLANLEYANAGLLSHLSLAFWDQDDPYDMGGDHCFLPGGNGKLVQALAENVLSILYEKTVHT
VRYSGDGVQVIAGNQVFEGDMALCTVPLGVLLKKSIFPELPQRKLDGIKRLGFGLLNKVAMLFPH
VFWETDLDTFGHLTDDPSRRGEFFLFYSYAPVAGGPLLIALVAGEAAHKFESMPPTDAVTRVLQILK
GIYEPKGISVPEPIQTVCTRWGNDFCFGSYSNVAVGASGDDYDILAETVGDGRLFFAGEATRRYP
ATMHGAFSLGREAANMAHYADIRSRRLKVDRAPSNAYSASLLADLFREPDIEFGSFAIFARKNA
DSKSPAILRVTFSEHRKKSHEVAKQDQQQHSNKLFLQQLQSHFNQQQLHVYTLRSRQQILDREVR
GGDEMRLNYLCEKIGVKLVGRKGLGPSGDSVIASIKAEGRNRKPLSTSAIKPGVSKLKAGIIRKLR
KAKVVKTHGSLHTSTNVVNTTKVTEENRTDQALPGVPVLGNSQSGLSNP

>Aradu_0023

MGYLCDFCGEQRSLVYCRSDAASLCLSCDRNVHSANALSKRHSRTLVCERCNSQPAFIRCA
EEKISLQCNCDWLGHGTTPTSTSGHKRQTINLYSGCPSAAEFSSIWSFFSDIPAMGETCEQELGLMSISE
NSEKSAWVPPENQNASSAQVTNLPKDKSVWGTSSMPESSEPRIMDQPPGPSNACVPKLYCPVKT
TPALCEDDDNLYYDFNMDEVDLDLENYEELFGVALSHSEELFENGIDSLFGTKEMSASAGDSNCQ
GAVAAEGSTGLINATQPACSNAAASDSMMSTKTEPICTGRQSQSNISFSGATKDGSGGGGGDYQD
CGASSMLLMGEPWPCCPESSMQSANRSNAVVMRYKEKKKTRKFEKRVRYASRKARADVRRRVK
GRFVKAGDIYDYPMSQTRSY

>Aradu_0028

MPLGERQGDKSESRYCGVETDFNDDMPHVLAFLNLSSGGFDFVIAPLMDPAYRPSLVQKGS
AGSGVPPFAGSDLVMSPSQWSSHVVGKISSWIDLSEDETIRMSETTLKQEIWASHLSLQACLLP
APKGTSCGNARCVNQLQGLSNMQLWLRIPLVKPDADSADGKNDTLVDSWETWNSFRLLCEHHS
LLSVALDILGTLPSANSLGRWFGESVRAAIINTEAFLTNARGYPCLSKRHQRLVTGFFNHSIQIIISGKS
VHPQANVAADASQNSHDEGLQRHALRPYLDYVGYLYQKMDPLPEQERFELGYRDFLQSPLOPLMD
NLEAQTYESFEKDKAMKVIQYQRAVSKALLDRVPDEEAFTKTTVLMVVGAGRGPLVRASLQAAEET
GRKLVFAVEKNPNAVVTLHALVKLEGWEDSVTIVSSDMRHWNAPKADILVSELLGSFGDNELSP
ECLDGAQRFLKEDGISPSSYTSFLQPVTASKLHNDIKAHKDLVHFETAYVVKIHNVARLAPTQPVFT
FTHPKHSDKESNQRYKKLQFVIPNDTGSAMVHGFAGYFDATLYKDVHLGIEPSTATPNMFSWFAIFF
PLRTPICLSPGSNLDVHFWRCCGPTKVWYEWVSTSPSSPIHNTNGRSYVWGL

>Arahy_0014

MAKPKAPRRTLDSYTVKHINKTFRAGDCVLMRPSDPKPSYVARIERIEADSRGANVVKVHV
RWYYRPEESIGRRRQFHGSKEVFLSDHFDVQSADTIEAKCTVHSFKSYTKLDAVGNDDFFCRFEYNS
STGAFNPDRVAVYCKCEMPYNPDDLVMVQCEGCSDWFHPACIDMTVEEAKRLDHFFCESCSAEGQK
KLQNSHSTSRHSDTKVETKRRRRS

>Arahy_0016

MAKVFLGSEPVVGSULTPSKKRDYRVTNRLQEGKRPLYAVVFNFIDSRYLNVFATVGGNRI
TVYQCLEGGVIAVLQSYVDEKDESFYTVSWASNVDGAPFLVAGGINGVIRVIDTGSERIYKSFVGH
GDSINEIRTQALKSSLVISASKDESURLWNVHTGICILIFAGAGGHRNEVLSVDFHPSDIYKIASCGMD
NTVKIWSMKEFWTYVEKSFTWTDLPSKFPTKYVQFPVFNASVHSNYVDCNRWLGDFILSKSVNEI
VLWEPKVKEQTPGEGVVDILQKYPVPECDIWFIFSCDFHYNQAAIGNREGKIFVWELQSSPPVLIAR
LSHPQSKSPIRQTAVSFDGSILSCCEDGTIWRWDDVTNPPA

>Arahy_0031

MDSPSEFSSFPSPQTQLPPQONPNPIVFLDPNSDSFPNWAENPNPNLQNYLSFTVPKKRRRGR
SQRNPASFRIPDPNDASNSNNNNNGFSSSAAAVPVSAPLSKVVIGNPPPLSMTMQYSVPDNDSEII
VINKEPKTEALIALSAGFPADSLTEEEIEAGVLPVIGGIEQVNYTLIRNHIIAKWRENVSQWVSKEMFV
DYIPQHYHALLDSAYNYLVSHGYINFGVAPLIKEKIPLMSKPSVIIIAGLAGLAAARQLMRFGFKV
TVLEGRKRAGGRVYTKKMEGGNRVCAAADLGGSVLTGTLGNPLGIVARQLGDVHLKVRDKCPLY

SLDGKPVDPMDVKVENAFNRLLDKASRLRQLMGEVSMDEVSLGAALETFRQVYKDAVNDEEMNL
FNWHLANLEYANAGLLSHLSLAFWDQDDPYDMGGDHCFPLPGNGKLVQALAENVLSILYEKTVHT
VRYSGDGVQVIAGNQVFEGDMALCTVPLGLVKKGSIKFIPELPQRKLDGIKRLGFGLLNKVAMLFPH
VFWETDLDTFGHLTDDPSRRGEFFLFYSYAPVAGGPLLIALVAGEAAHKFESMPPTDAVTRVLQILK
GIYEPKGLSVPEPIQTVCTRWGNDFCFGSYSNVAVGASGDDYDILAETVGDGRLFFAGEATTRYP
ATMHGAFLSGLREANMAHYADIRSRLKVDRAPSNAYSASLLADLFREPDIEFGSFSIAIFARKNA
DSKSPAILRVTFSEHRKKSHEVAKQDQQQHSNKLLFQQLQSHFNQQQLHVYTLRSRQQILDREVR
GGDEMRLNYLCEKIGVKLVGRKGLGPSGDSVIASIKAEGRNRKPLSTSIIKPGVSKLKAGIIRKLIR
KAKVVKTHGSLTSTNVVNTTKVTEENRTDQALPGVPVLGNSQSGLSNP

>Arahy_0034

MEEVERLWNEVRELSLGNNSKVERLQAPPTALQFLRDFVNPKNPCII SNAINHWPALSLWA
QAQPNQAQHSYLSRSLSSSSVSLHLTPSGRADALAPLSSPSSLCFASAYVQRTPPFDALRLIASPDHN
PAVVAYAQQQNDCFRSEYSALAADCDPHIDWATEAFGGALPEAVNLWIGNQRSHTSFHKDHYENL
YAVVSGEKHFILLPPTDVHRLYIRQYPAATYRYSSTGFEFDLEELPIRYVPWCSVDPFPPEKLHEE
VSKFPLYFNGPKPFECTVKAGEILYLPMSWFHHRVQSSGDDGCTIAVNYWYDMQFDIKYAYFNFLK
SIHYRSPSLLLPEKFSEETDSDPDCL

>Arahy_0042

MESPPSAGAASPTVGKKKETRGRKPKPKDEQHTKTIKEGKTPHQHHQHHHHQPSV
DDKYTQWKSLLVPLVDWLANHNLVWPSLSCRWGPQLEQATYKNRQRLYLSEQTDGSPNTLVIA
NCEVVKPRVAAAHEISQFNEEARSPFKKYKTIHPGEVNRIRELPQNSKIVATHTDSDPDLIWDVES
QPNRHAVLGATNSRPDLVLTGHQDNAEFALAMCPTPEYVLSGGKDKTVVLWSIEDHITSSATDSKS
GGSIVKSNKSGEGNDKSADGPSVGPRIYCGHDETVEDVTFPCSSAQEFCSVGDDSCILWDARV
SSPAVKVEKAHNADLHCVDWNPHDNLILTGSADHSVRMFDNRNLTSNGVGTPIHKFECHKAAVL
CVQWSPDKASVFGSSAEDGLLNWDYEVGKKIERSGKS VSSPGLFFQHAGHRDKVDFHWNAQ
DPWTIVSVSDDCESTGGGGTLQIWRMSDLIYRPEEEVLAELEKFKSHVVACASKTEK

>Arahy_0045

MGDGGSGNRGGNSNNSGGGSNKPEWLQYDLIGKIGEGTYGLVFLARIKSQTNRGKSIAIK
KFKQSKDGDGVSPTAIREIMLLREITHENVVCLVNVHINHADMSLYLAFDYAEHDLYEIRHHRDKV
NHSINQYTVKSLLWQLLNLNLYLHSNWIIHRDLKPSNILVMGEGEEHGVVKIADFLARIYQAPLKP
LSENGVVVTIWRAPPELLGAKHYTSAVDMWAVGCIFAELLTLKPLFQGVVVKATPNPFQLDQLDK
IFKVLGHPTLEKWPSLANLPHWQQDVQHIQGHKYDNNSLYNVHLSPKSPAFDLSKMLEYDPRKR
ITAAQALEHEYFKIEPQGRNALVPCQPGETFVNYPTRPVDTTDFEGTTNMQPPQPSQVTSGA AVS
GSMPPGGHVSNRSPRPMNVVGMQRMPPHQAQAYNLTSQAGMGGGMNPPGMPMQRGVPPHQ
QQLRRRDQMGMPGYPPQKSRRI

>Arahy_0059

MGYLCDFCGEQRSLVYCRSDAASLCLSCDRNVHSANALSKRHSRTLVCERCNSQPAFIRCA
EEKISLCQNCDWLGHGTTSTSGHQRQTINLYSGCPSAAEFSSIW SFFSDIPAMGETCEQELGLMSISE
NSEKSAWVPPENQNASSAQVTNLPGKDKSWVGTSSMPESSEPRIMDQPPGPSNACVPKLYCPVKT
TPALCEDDDNL YDDFNMDVDLDLENYEELFGVALSHSEELFENGIDSLFGTKEMSASAGDSNCQ
GAVAAEGSTGLINATQPACSNAAASADSMSTKTEPIICFTGRQSQSNISFSGATKDGGGGGGGGDYQ
DCGASSMLLMGEPWPCCPESSMQSANRSNAVMRYKEKKKTRKFEKRVRYASRKARADVRRRV
KGRFVKAGDIYDYDPMSQTRSY

>Arahy_0064

MKLAGIKSVENAHDDSIWAATWLPSTASRPPLLLTGSLDETVMWRSDLVLERSNTCHCL
GVASVAAPLGSIAASSLDSFVRVFDVDSNATIATLEAPPSEVWQMRFPKGA VLAVAGGGSASV
KLWDTNTWELVATLSIPRPEGSKPTDKSGSKFVLSVAWSPDGKRLACGSMDGTISVFDVQRGKFL
HHLEGHYMPVRSLSVSPYDPRLLFTASDDTNVHMYDAEGKALVGNMSGHASWVLCVDVSPDGAA
IATGSSDKTVRLWDLAMRASVQTMSNHTDQVWGVAFRPDGRGGRLASVSDDKSVSLYDYS

>Arahy_0078

MSGKGAKGLITAKTPAPNKDKDKKPKVSRSSRAGLQFPVGRHRLKSRVTANRVRGATA
AVYSAALEYLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELDTLIKGTIAGGGVIPHIKSLIN
KSSKE

>Arahy_0079

MSIAAPGQLNVNESPSWGSRSVDCFEKLEQIGEGTYGQVYMAKEIKTGEIVALKKIRMDNE
REGFPITAIREIKILKKLHHENVIKLKEIVTSPGPEKDEQGRPDGNKYKGGIYMVFEYMDHDLTGLAD
RPGMRFTVPQIKCYMRQLLTGLHYCHVNQVLRDRIKGSNLLIDNEGNLKLADFLARSFSDHNAN
LTNRVITLWYRPELLLGTTKYGPAVDMWSVGCIFAELLHGKPIFPKGDEPEQLNKIFELCGAPDEV
NWPVGVSKTPWYNQFKPTRPMKRRLREVFRHFDRHALELLERMLTLDPAQRITAKDALDAEYFWTD
PLPCDPKSLPKYESSHEFQTKKKRQQQRQNEEMAKRQKMQHPPQHTRLPPIQQPGQHAQMRSGPN
QNLHGSQPQVAAGPSHHYGKPRGPGSGGPRYPPSGNPGGGYNHPNRGGQGGGGYGS GPYPPQGRG
APYGSSGMPGGGPRGGGGSGYGVGAPTYPQQGGPYGSGTGRGSNMMGGNRNQYGWQQ

>Arahy_0093

MTNSSSSKKHGQDQPELSGPTPKSHRTKMGKSEDTDKKNLGKCLKDVEISVPIVYGNIADF
LGKKASEYQSHKWTVYVRGASNEDLGAIVKRAVFLHSSFNPNTRVVESSPPFELSEAGWGEFEIVIT
LHFHSDVCDKPLNLYHHLKLYPEDENSSMSTKKPVVVELYDEIVFPDPSEAFLSRVQNHPAVNVPRL
PPGVTLPPPIVEDASKRRKGDTKDHPLSQWFTNFSEADELLQLAAARQQVQAHAIAKLRRQISLIDGQ
HHQYFKSSSDQ

>Araip_0029

MDSPEFSSFPSSQTQLPPQQNPNPIVFLDPNSDSFPNWAENPNPNLHNYLSFTVPPKRRRGR
SQRNPALFRIPLNPNNDASNSNNNGFSSSAAAVPVSAPLSKVVIGNPPPPPLSMSMQYSVPDNSDEIIV
INKEPKTEALIALSAGFPADSLTEEEIEAGVLPVIGGIEQVNYTLIRNHIIAKWRENVSQWVTKEMFVD
YIPQHYHALLDSAYNYLVSHGYINFGVAPLIKEKIPLEMSKPSVIIIAGLAGLAAARQLMRFQKVT
VLEGRKRAGGRVYTKMEGGNRVCAAADLGGSVLTGTGLGNPLGIVARQLGDVLHKVRDKCPLYS
LDGKVPDPMDVKVENAFNRLLDKASRLRQLMGEVSMDVSLGAALETFRQVYKDAVNDEEMNLF
NWHLANLEYANAGLLSHLSLAFWDQDDPYDMGGDHCFLPGGNGKLVQALAENVSILYEKTVHTV
RYSGDGVQVIAGNQVFEGDMALCTVPLGVLLKKSIFPELQPKLDGIKRLGFGLLNKVAMLFPHV
FWETDLDTFGHLTDDPSRRGEFFLFYSYAPVAGGPLLIALVAGEAAHKFESMPPTDAVTRVLQILKGI
YEPKGISVPEPIQTVCTRWGNDFCFGSYSNVAVGASGDDYDILAETVGDGRLFFAGEATTRYPAT
MHGAFLSGLREANMAHYADIRSRRLKVDRAPSNAYSASLLADLFREPDIEFGSFAIFARKNADS
KSPAILRVTFSEHRKKSHEVAKQDQQQHSNKLLFQQLQSHFNQQQLHVYTLRSRQQILDREVRG
GDEMRLNYLCEKLGVKLVGRKGLGPSGDSVIASIKAEGRNRKPLSTSAIKPGVSKLKAGIIRKRLIR
KAKVVVKTHGSLHTSTNVVNTTNVTEENRTDQALPGVPVLGNTQSGLSNP

>Araip_0037

MESPPSAGAASPTVGKKKETRGRKPKPKPKDEQHTKTIKEGKKTTPHQHQQHHHHQPSV
DDKYTQWKS LVPVLYDWLANHNLVWPSLSCRWGPQLEQATYKNRQRLYLSEQTDGSPNTLVIA
NCEVVKPRVAAAHEHISQFNEEARSPFKYKTIHPGEVNRIRELPQNSKIVATHTDSPDVLWDVES
QPNRHAVLGATNSRDLVLTGHQDNAEFALAMCPTPEYVLSGGKDKTVVLWSIEDHITSSATDSKS
GGSIVKSNKSGEGNDKSADGPSVGPRGIYCGHDETVEDVTFCPSSAQEFCSVGGDSCLILWDARVG
SSPAVKVEKAHNADLHCVDWNPHDVNLILTGSADHSVRMFDRRNLTSNGVGTPIHKFECHKAAVL
CVQWSPDKASVFGSSAEDGLLNIWDYEKVGKKIERSGKSVSPPGLFFQHAGHRDKVVDHFHWAQ
DPWTIVSVSDDCESTGGGGTLQIWRMSDLIYRPEEEVLAELEKFKSHVACASKTEK

>Araip_0046

MGYLCDFCGEQRSLVYCRSDAASLCLSCDRNVHSANALSKRHSRTLVCERCNSQPAFIRCA
EEKISLCQNCDWLGHGTTASTSGHKRQTINLYSGCPSAAEFSSISFFSDIPAMGETCEQELGLMSISE
NSEKSAWVPPENQNASSAQAATNLPQKDKSWVGTSTMPESSEPRIMDQPPGPSNACVPKLYCPVK
TTPALCEDDDNLYDDFNMDVDLDLENYEELFGVALSHSEELFENGGIDSLFGTKEMSASAGDSNC
QGAVAAEGSTGLINATQPACSNAAASADSMSTKTEPIICFTGRQSQSNISFSGATKDGGGGGGDYQ
DCGASSMLLMGEPWPCCPESSMQSANRSTAVMRYKEKKKTRKFEKRVRYASRKARADVRRRV
KGRFVKAGDIYDYPMSQTRSY

>Cajca_0023

MDPPPQFPPQNPVFLDPTSDSIPNLLQPNHYLSFSVPKRRRGRSQRNPSSFRLPLPLPSD
HPPSSSRAPASDEIIVINKEPKTEALIALTAGFPADSLTEEEIDA AVLPIVIGGIEQVNYTLIRNHIAKW
RENVSNWVSKKAFLDYIPQHYHALLDSAYNYLVSHGYINFGV ASPVKEKIPAEPSTPAVIVIGAGLA
GLAAARQLMRFGFRVTVLEGRKRAGGRVYTKKMEGGNRLCAAADLGGSVLTGTGLGNPLGIVARQ
LGELLHKVRDKCPLYCVDGRPVPDMDVKVESAFNRLLDKASRLRQLMGEVSVVSVLGAALLETFR
QVFKDGVSDDEMNLFNWHLANLEYANAGLLSHLSLAFWDQDDPYDMGGDHCFLPGGNGKLVQA
LAENVPILEYKTVHTIRYSGDGVQVIAGSQAFEGDMALCTVPLGVLLKKSIFPELPQRKLDGIKRL
GFGLLNKVAMLFPHVFWEMDLDTFGHLNDDSSRRGEFFLFYSYATVAGGPLLIALVAGEAAHKFES
MPPTDAVTRVLQILKGIYEPKGI VVPEPIQTVCTRWGSDPFCFGSYSNVAVGASGDDYDILAESVGD
GRLFFAGEATTRRYPATMHGAFLSGLREANMAH HANIRTLKMKVDKAPSNAHSCASLLADLFRE
PDIEFGSFSVIFARKNTDPKSPAILRVTFSEPRKKCHEVAKQDQQHSNKLLFQQLQSHFNQQQLHVY
TLLSRQQVLDLREVRGGDEMRLNYLCEKLGVKLVGRKGLGMNADSVISIKAEGRNRKPVSTSLAL
KPGVSSKLKAGIMKRKLIRRAKVVRKSNGSINVGSAIKVSEEVKITDHVLPDVPVSGSDQNDLSNP

>Cajca_0033

MGGETGKSFGLGCDPVVGSLSSSKKREYRVTNRLQEGKRPLYAIVFNFIDSRVFNVFATVG
GNRVTVYQCLEGGVIAVLQSYVDEKDES FYTVSWACNV DGTPTYV VAGGINGVIRVIDAGNEKIHK
SFGVGHGDSVNEIRTQTLKPSLVISASKDES VRLWNVHTGVCILIFAGAGGHRNEVLSVDFHPSDIYRI
ASCGMDNTVKIWSMKEFWTYVEKSFTWTDLP SKFPTKYVQFPVYNASVHLNYVDCNRWLGDFILS
KSVDNEIVLWEPKVKEQTPGEGVVDVLQKYPVPECDIWFIFSCDFHFNMIAVGNREGKIFIWELQS
SPPVLIARLSHPQSKSPIRQTATSF DGSTILSCCEDGTIWRWDDVTNNSST

>Cajca_0065

MSRLMEPLAVGRVIGEVVDIFTPSVRMNVTYSTKQVANGHELMPSTIVAKPRVEIGDDMR
TAYTLIMTDPDAPSPDPYLRHLHWMVTDIPGTTDVSFGKEIVGYESP KPVIGIHRVYFILFKQRGR
QTVRPPSSRDYFNTRRFSEENGLGLPVA AVYFNAQRETAARRR

>Cajca_0072

MTNSSSSKKHGQDQPELSGPTPKSQRTKMGKSEDNDKKNLKGKLDVEINIPIVYGNIAFW
LGKKASEYQSHKWTVYVRGASNE DLGAIKRAVFQLHSSFNPNTRVVEAPPFELSEAGWGEFEVAIT
LYFHSDVCDKPLNLYHHLKLYPEDENSSMSTKKPVVVEFYDEXVFPDPSEAF LARMQGHPAVNLPR
LPAGLTLPPPVEDASKRRKGDTKDHLSLQWFTNFSEADELLQLAAARQQVQAHI AKLRRQISLIDG
QHQQFKSTSEQ

>Cicar_0021

MDPPPPSHVSSPQQNPNDNIQSDSIPNFDQNP NPNLSHYLSLSVPKRRRGRSQRNPSSFRS
PFTLNASSLNNDLTSSIRNPPSSSTQFHDFSD EIIIMINKEAKTEALIALSAGFPADSLSEEEIETGVIPVIG
GIEQVNYTLIRNHIAKWRENVSIWVNKKMFTDYIPQHYHSLLD SAYNYLLSHGYINFGVASPIKDKI
PTEPSKPGVIIIGAGLAGLAAARQLMRFGFKVTVLEGRKRAGGRVYTKKMEVGNRVGAAADLGG
VLTGTGLGNPLGIVARQLGDLLHKVRDKCPLYSDVGKPVNPDMDVKVESAFNRLLDKASRLRQLMG
EVSVVSVLGAALLETFRQVYKDAVNDEEMKLFNWHLANLEYANAGLLSHLSLAFWDQDDPYDMGG
DHCFLPGGNGKLVQALAE NVPILEYKTVHTIRYSGDGVQVNAGSQVFEGDMALCTVPLGVLLKKSIF
KFIPELPQRKLDGIKRLGFGLLNKVAMLFPHVFWEMDLDTFGHLSDDPSRRGEFFLFYSYATVAGGP
LLIALVAGEAAHKFESMPPTDAVTKVLHILKGIYEPKGINVPEPIQTVCTRWGSDPFCFGSYSNVAVG
ASGDDYDILAESVGDGRLFFAGEATTRRYPATMHGAFLSGLREANMAH HANIRSMNVKVEKPPSS
AHSCASLLADLFREPDIEFGSFSIIFAQKNADPKSPAILRVTFGEPKKKYHEVAKQDQQHSNKLLFQ
QLQSHFNQQQLHVYTL LSRQQVLDLREVRGGDEMRLNYLCEKLGVKLVGRKGLGLNADSVIASI
KAERGNRKPVSTSM LKPLSKIKAGIIRKAKVVKKSNESIPRASMNVGSASRVSEENRIIDQV
IPDVVVLGNNQNDLANSNP

>Cicar_0041

MGYVCDFCGQQRSMVYCRSDAACLCLSCDRNVHSANTLARRHSRTLLCERCSSQPALVRC
SEEKISLCQNCDWLGHGNTSSSTHKRQTINCYSGCPSSAELSSIWVFLDIPSLSETTCEQELGFM SINE
NRS AWVAPKNQNVSDSDQVTDLPDMDKSWAGTSSKPESSEPRLLDHSPGSTKECVPKLYCSATKC
PEEREDKDDLYGDFDMDEM DINIENYDELFGMALTHSEELFENGGFNSLFGAKEM SAGSDCQDAI
AAEGSSVGHFNATQPACSNAA ADSILSTKTEPNLCITAKQAQSNLSFSGINEESGGAGDYQDCGASS
MLLMGEPWLNTPENDLQLQSANRSSAVMRYKEKKKTRKFDKRVRYASRKERADVRRRVKGRF
VKAGDVYDYDPLSQTRSY

>Cicar_0062

MSRSLEPLSMGRVIGEVVDIFNPSVRMNVTYSTKQVANGHELLPSIVMKNKPRVEIGDDLR
TAYTLIMTDPDAPSPDPYLREHLHWMVTDIPGTTDVSFGKEIVEYENPKPVIGIHRVYFILFKQRGR
QTVRSPTFRDNFNTRSFSQQNSLGLPVAAYVFNAQRETAARRR

>Cicar_0069

MTNSSSSKKHGQDQSELSGPTSQSKSQRKMGKSEDNDKKNKGKLDVEISVPIVYGNMAF
WLGKKASEYQSHKWTVYVRGASNEGLGVIKRVVFLHSSFNPNTRVVDAPPFELSEAGWGEFEIAI
TLHFHSDVCDKPLNLYHHLKLYPEDENSSMSTKKPVVVEFYDEVVFPDPSEAFARVQNHPAVNFP
RLPAGHTLPPPIVEDANRRRKGDTKDNPLSQWFLSFSEADELLQLAAARQQVQAHIAKVKRQISLIE
GQHKQLKSSSDQ

>Glyma_0052

MGYICDFCGDQRSMVYCRSDAACLCLSCDRNVHSANALSKRHSRTLLCERCNSQPAFVRC
VEEKISLCQNCDWLGHGTSTSSSTHKRQAINCYSGCPSAAELSSIWSFVLDIAAISESTCEQELGLMSI
NENKSVGVPPEGQNVSGSDEVTDQPALDKSLVGTSSMPRESSKPRILDQPARPANECLSKLYCPATK
FPALCENDNLYDDFNMDDEVLDNLENYEELFGMALSHSEELFENGGIDSLFGTKDMSAGDFSCEDAI
AAEGSSVGQVNMVQACSNAAASADSILSTKTEPILCFTGRQAQSNLSFSGVTGDSSAGDYQDCGASS
MLLMGEPWFAPCPENSLQSANRSNAVMRYKEKKKTRMFDKKVRYASRKARADVRRRVKGRFVK
AGDVYDYDPLSTTRSF

>Glyma_0066

MPLGERAGDKSESRYCGVETDFNDDMPHVLHFNLYSGSDFVVAAPLMDHSYRPSLVQKDA
LGGVLPFAGSDLVLSQWSSHVVGKISSWIDLSEDETLRIDSETTLKQEIAWASHLSLQACLLPA
PKGTSCANYARCQVNLQGSNNMQLWLRVPLVKPDDDSMDAKSVALVDSWETWNSFRLLCEHHS
QLSVALDILSTLPSNSLGRWFGPEVRAAILHTDCLTNSHGCPCLSKRHQSLITGFFNHSIQIIVSGNS
VHPKASVDANDNSGNTNSQRHPLRPYLDYIGLYQRMPLPEQERFELGYRDFLQSPQLMDNLEA
QTYETFERDAVKYIQYQRAVSKALLDRVPDEEASVKTIVLMVVGAGRGPLVRASLQAAEETGRKLLK
VYAVEKNPNAVVTLHALVKLEGWEDTVTIVSSDMRHWSAPEKADILVSELLGSGFDNELSPECLDG
AQRFLKQDGISIPSSYTSFLQPVTASKLYNDVKAHKDLVHFETAYVVKMHNVAKLAPTQPVFTFTHP
KHSDKEINQRYKLNFIIPNDTGSAMVHGAFAGYFDATLYKDVHLGIEPSTATPNMFSWFAIFFPLRTP
ICVDPGSTLEVHFWRCCGSTKVWYEWCVASPASSPMHNSNGRSYWWGL

>Glyma_0076

MSRLMEPLVVGRVIGEVVDIFSPSVKMNVTYSTKQVANGHELMPSTIMAKPRVEIGGDDM
RTAYTLIMTDPDAPSPDPCLREHLHWMVTDIPGTTDVSFGKEIVGYESPVPVIGIHRVYFILFKQRGR
QTVRPPSSRDHFNTRRFSEENGLGLPVAAYVFNAQRETAARRR

>Glyso_0048

MGYICDFCGDQRSMVYCRSDAACLCLSCDRNVHSANALSKRHSRTLLCERCNSQPAFVRC
VEEKISLCQNCDWLGHGTSTSSSTHKRQAINCYSGCPSAAELSSIWSFVLDIAAISESTCEQELGLMSI
NENKSVGVPPEGQNVSGSDEVTDQPALDKSLVGTSSMPRESSKPRILDQPARPANECLSKLYCPATK
FPALCENDNLYDDFNMDDEVLDNLENYEELFGMALSHSEELFENGGIDSLFGTKDMSAGDFSCEDAI
AAEGSSVGQVNMVQACSNAAASADSILSTKTEPILCFTGRQAQSNLSFSGVTGDSSAGDYQDCGASS
MLLMGEPWFAPCPENSLQSANRSNAVMRYKEKKKTRMFDKKVRYASRKARADVRRRVKGRFVK
AGDVYDYDPLSTTRSF

>Glyso_0068

MSRLMEPLVVGRVIGEVVDIFSPSVKMNVTYSTKQVANGHELMPSTIMAKPRVEIGGDDM
RTAYTLIMTDPDAPSPDPCLREHLHWMVTDIPGTTDVSFGKEIVGYESPVPVIGIHRVYFILFKQRGR
QTVRPPSSRDHFNTRRFSEENGLGLPVAAYVFNAQRETAARRR

>Lupan_0026

MGDGSSNRTNSNSNKPEWLQYDLIGKIGEGTYGLVFLARIKSQANRGKSIKFKQSKD
GDGVSPTAIREIMLLREITHENVVCLNVHINHTDMSLYLAFDYAEHDLYEIIRHHRDKVNHSINQY
TVKSLLWQLNGLNLYLHSNWIIHRDLKPSNILVMGDGEEHGVVKIADFLARIYQAPLKPLSENGVV
VTIWRAPPELLGAKHYTSAVDMWAVGCFAELLTKPLFQGAEVKGTNPNFQLDQLDKIFKVLGH
PTLDKWPPLANLPHWQQDVQHIQGHKYDNTSLCNVHLSPKSPAYDLLSKMLEYDPRKRITAAQA
LEHEYFKMEPQGRNTLVPCQPGETFVNYPTRPVDTTTTDFEGTANLQPPVSSGAGSMPGGHGSNR
AVPRPMNVVGMQRMPHQPMQAYNLASQAGMGPMPNPGGIPMQRGVPAHQQQQLRRKDQMG
PGYPPQKSRRI

>Lupan_0036

MGYTCDFCGDQRSMVYCRSDAACLCLSCDRNVHSANLSKRHSRTLLCERCNSQPAVVRC
AEEKVSLCQNCDWLGHGSSTSSSTHKRQTINCYSGCPSSSEELSSIWSFILDISSMSGSSCEQELGLMSI
NENSNTSSRVPPENQNVGTGSDQVNNLPGKDKSGVGASSKSESSTESVDPDHPAGPSNQCLPKFYRPG
TKCPSICEDNMYPDFEMGEADINLENYDELFGMTLTHSEELFENGGFDSLFGIKDISAEDSDCLGAVV
AESSVELVNAMQPACSNAAESADSILSTKTEPIICFTARQAPSNISFSGIAGESNAGDYQDCGASSMRL
MGEPPWHSPFPENSLHSASNRSNVAVMRYKEKKKIRKFDKRVRYASRKERADVRRRVKGRFVKVGD
AYDYDPLGPTRSY

>Lupan_0049

MPLGERGGDKTESRYCGVETEFNDDMPHVLAFNLSSAGFDFVIAPLMDPAYRPSLVQKDIF
GSAVLPFAGSDLVLSQWSSHVVGKISPWLDLSEDEILRLDSETTLKQEISWASHLSLQACLLPAP
KGTTTCANYARCVNQILQGLSNMQLWLRIPVVKPDDDSMDANTDTLVDTWEMWNSFRLQCEHHSQ
LSVALDILNLTLPENSLARWFGESVRAAIINTDSFLTARGYPCLSKRHQKLITGFFNHSIQIIISGNSV
HPKTKVVADVSHNHSGADSQRHPLRPYLDYVGYLYQKMDPLPEQERFELGYRDFLQAPLQPLMDN
LEAQTYYETFEKDAMKYIQYQRAVSKALLDRVPDEEASVKVTVLMVVGAGRGPLVRASLQAAEETG
RKLKVYA VEKNPNAVVTLHALVKLEGWEDIVNIVSSDMRCWNAPEEADILVSELLGSFGDNELSPE
CLDGAERFLKQDGISIPSSYTSFLQPVTASKLYNDVKAHKDLAHFETAYVVKIHNAAARLAPSQPVFTF
THPKPKDESQRKYKLLHFVMPNDTGSAMVHGFAFYFDATLYKDVHLGIEPSTATPNMFSWFAIFFP
LRSPICVPGSTLEVNFWRCCGPKKVWYEWCVTSPSSSPIHNSNGRSYWVGL

>Lupan_0055

MSGKGAKGLITGKTSAAANKDKDKKKPTSRSRAGLQFPVGRHRLKTRTTAHGRVGATA
AVYSAALEYLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELDTLIKGTIAGGGVIPHIHSLIN
KSSKE

>Lupan_0067

MTEGHQFNITLGGRRGNTGQMKIYSEGILWRRQGGGKLEVDKDDIMGVTWMKVPR
NQLGVQIKDGLFYKFTGFRDQDVASLTTFQNTFGIEVQEKQLSVTGRNWGEVDLNGNMLAFMVSS
KQAFELSVADVAQTQLQGKNDVILEFHVDDTTGANEKDSLMEISFHIPNSNTQFIGDENCPPAQVFR
DQIMSMADVGVGGEDAVVTFEGIAILTPRGRYSVELHLSFLRLQGQANDFKIQYSSVRLFLPKSN
QPHTFVILSLDPPIRKQTLYPHIVMQFETDLVVQSELSMGEDLYNTKYKDKLEMSYKGLIHEVFTTI
LRGLSGAKVTKPGKFRSCQDGYAVKSSLKAEDGILYPLEKSFFFLPKPPTLILHEEIDYVEFERHAAG
GSNMHYFDLLVRLKSDTEHLFRNIQRNEYHNL YGFISSKGLKIMNIGDAQAPGGMAKVLETDDDDA
VDPHLERIKNAAGGGDESDEEDEDVLDKDDGGSPDSDSGGDDSDASQSGDDEKEKPAKKEPKKD
LSSKASTSKKSKDADEDGKKKKQKKKKDPNAPKRALSGFMFFSQMERENLKKTNPGISFTDVG
LGEKWKMSVVEEKEYEAKARADKIRYKDEISGYNKNPQPMNVDSGNESDSS

>Lupan_0068

MTQISSSTGTKDPEQEQTLDQDKELVLPNPKGSEPVSGEEQEVVASEENKEKPLVDQPVPM
EEESVNPATVFCIRLQPNNSLLHKMSVPEICRNFS AVSWCGKMNAIGCAAETCARIPSSTTNP
PIHIVIPERPTESVFNVIADSPRDSVQFIEWSPTSCSRALLIANFHGRVTIWTQPSQGPANLVHDTTCW
QREHEWRQDIAVVTKWLSGVSPYRWLSSKSSAPSNSKSTFEEKFLSQSQT SARWPNFLCACSVFSS
GSVQLHWSQWPPSQNGKTPKWFYTSKGLLGCGPSGIMAGDAITESGAMHVAGVPIVNPSTVVVVE
VTPGPGNGFQVTPKTSTTCGIPPLSPPNWAGFAPLAAYLFSWQDYLLSEAKQGRKQTDKNLGD
LYCSPVSNFSAYVTPEAAAQSAATTTWGSVTA VAFDPTCGGSVIAVVIVEGYMSPYDPDEGPSIT
GWRVQRWESSLQQVVLHPIFGNPTSSMGGQPPMQTVWQSKVDLSIPPTNDFKNHPAPGIITD
VQRVS VSGSDKSERVNFDPFDLPDVRTLARVVYSAHGGEIAIAFLRGGVHIFSGPNFSPVDNYQIN
VGSIAI VPAFSTSCCSASVWHDSSKDHTILKIIRVLPPAIPTSQVKANSSAWERAIAERFWWSLLV
GVDWWD AVGCTQSAEADGIVSLNSVIAVLDADFHSLPSSQHRLQYGPSLDRIKCRLLLEG
SNAQEVRAMVLDM QARLLDMLGKIESALINPSALVPEPWQASAEALNNIDSESMAIEPALVPTI
QAYVDSVLDLASHFIT RLRRYASFCRTLATHAVNSGTGSNNRNVTSTPTQSSATPATSQGGQ
NGSATSMSGSTQIAWIQEAIAK ISNTTDGASNPPTNPISGPPTFVPISINTGTFPGTAVR
LIGDCHFLHRLCQLLLFCYFFRRTQIPRYSGV ANRSADTNMQKPQPNASAPGKVEDITK
PASAVVKSDDGQTGRAGSIVPGAKVVEEASSARSRLGAG NAGQGYTFDEVKVLFMLLID
LCRRTAALQHPLPISQVGSNNIQVRLHYIDGNVTVLPEVVEASLGP MQNMPRPRGVDA
AAGLLLRELELYPPAEWHRRNMFGGPWTDPEVDVDFANDAPRLVSSDPHDFSSL EHC
DVYNGAQLRWPKRMSERDAAFGLNTSVGLGGYLGIMGSRDVTAMWKTGLEGVWYK
IRCMRQTSFTSPDATSSPSQNDNEIRWISRWAYGCPMCGGTWIRVV

>Meditru_0032

MGSLCDFCGDQSRSLVYCRSDAASLCLSCDRNVHSANELSKRHSRTLVCERCNLQPAYVRC
VEEKVSLCQNCDSWAHGTNPSSSTHQRQSINCFSGCPSASELSSIWPFSDIPSTGEACEHKLGLMSIN
ENSDNSARVPPEKSNVSGSAQVADLPSKNKSGVDTSSIPESAKPRILDQAPGSSNEFMPKLLCPSRK
APALCEDDKLLDDFNIDEMDFELENYSELFDALNHSEEFFENGGINSLFERKDMSASAGDSNCQGA
FAAEGSSARFVSAIQPECSNAASADSILSTKTEPVIYFATERQSNLSFSGVNKDASAGDYQECGTSSML
LTGEPPWCPCPENSIQSANRSNAVMRYKEKKKNRKFDDKVVRYASRKARADVRKRVKGRFVKAGE
TYDYDPLSQTRSC

>Meditru_0039

MLNTSSSKNHDQDQSQLSGPTPNSHPIKMGKSEDNDKKNKSGKLLKDVEISVPIVYGNVSWF
LGKKASEYQSHKWTIYVRGASNEDLGVIVKRVVFLHTSFNNPTRIVDAPPFELSEAGWGEFEIAITL
YFHSDVCDKPVNLYHQLRLYPDESSSYSTKPKVVVEFYDEVVFPDPSEAFARLQNHVVMNPR
PAGLTLPPPIEDASKRRRGDTKDNPLSQWFLNFSEADELLQLAAARQQVQAHISKLKRQISLIDGQ
HKQLKSSPDQ

>Meditru_0052

MSGKGAKGLTTGKTPANLKD KDKKKKPTSRSSRAGLQFPVGRHRLKSRRTMAHGRV GATA
AVYSA AILEYLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELDTLIGKTIAGGGVIPHHKSLIN
KSSKE

>Meditru_0061

MVKLFLSGTEPVVGNLNSRKEYRVTNRLQEGKRPLYAVVFNFIDSRYFNVFATVGGNR
VTIYQCLQGGVIAVLQSFVDDDKDESFTYVSWACNTNGAPLAVAGGVNGIIRVINGSDEKLHKS
GHGDSINEIRTQAMRPALVISASKDES VRLWNVHTGVCVLIFAGAGGHRNEVLSIDFHPSDIFK
GMDNTVKIWSMKDFWGYVEQSFTWTDIPSKFPTKYVQFPIFNASVHTNYVDCTRWLGEFILSKS
HEIILWEPKVKEQSPGEGAADILQKYPVPECDIWFIKFCDFHFKACAIGNREGKIFVWDLQSSP
ILN AKLVHAQSKSPIRQTALSHDGSTILCCDDGTIWRWDDANSEA

>Phavul_0006

MDPPPQFSSSLNPNPNPNVFLDPTSDSITNLIQNPNSNPNNPTTLSHYLSFSVPKRRRGRS
QRNPASFRLPLNLPTDTPSSSRTPASDEIIVINKEPKTEALIALTAGFPADSLTEEEIDA
AVLPVIGGIEQ VNYTLIRNHIAKWRENVSNWVSKKTFLDYIPQHYHALLDSA YNYLLSHGYIN
FGVASPIKERIPAEA SKPSVIIVGAGLAGLAAARQLMRFGFKVTVLEGRKRAGGRVYTKKMEG
GNRVCAAADLGGSVLTG TLGNPLGIVGRQLGELLHKVRDKCPLYCMDGRPVDPDTPDKVES
AFNRMLDKASRLRQLMGEVSV DVSLGAALETFFQVFKDSVSDEELSLFNWHLANLEYANAG
LLSNLSLAFWDQDDPYDMGGDHCF L PGGNGKLVQALAENVPILYEKTVHTIRYSGDGVQ
VIAGSQVFEGDMALCTVPLGV LKKGSIKFIPEL PQRKLDGIKRLGFLNKVAMLFPHVFW
EMDLDTFGHLSDDPSRRGEFFLFYSYATVAGGPLLIALV AGEAAHKFESMHPTDAVTRVL
QILKGIYEPKGINVPEPIQTVCTRWGSDFCFGSYSNVAVGASGDD YDILAESVGDGRLFF
AGEATRRYPATMHGAFVTGLREANMAHHANIRTLKMKVDKAPSNAHSC ASLLADLFR
EPDIEFGSFSIIFAQKNTDPKSPAILRVTFGEVRKCKNEVAKQDQQHSNKLLFQQLQSHF
NQQQLHVYTLTRQQVLDLREVRGGDEMRLNYLCEKLGVKLVGRKGLGMNADTVIDFIK
SERGN RKPVSTSLALKPGVSSKLGAGITKRKLIRRAKVVRKSNGSVDMGSATKISEV
KMTDLVRPDVTVSG SNQSDLSNP

>Phavul_0010

MGDGSRSRWSRAEWVQQYDLLGKIGEGTYGLVFLARTKTPSPSKSIAIKKFKQSKDGDGVS
PTAIREIMLLREITHENVVCLVNVHINHADMSLYLAFDYAEHDLYEIIRHHRDKLNHSINQY
TVKSLL WQLLNGLTYLHSNWMIHRLKPSNILVMGEGEEHG VVKIADFG LARIYQAPLKS
LSDNGVVVTIYWRAPELLGAKHYTSAVDMWAVGCIFAELLTLKPLFQGAEVKATSNPFQ
LDQLDKIFKVLGHPTLEK WPSLASLQHWQQDVQHIQGRKYDNTGLYNVVHLSTKSPAY
DLLSKMLEYDPRKRITAAQALEHEY FKIEPLPGRNALVPCQLGEKIVNYPTRPVDNTD
LEGTTNLPSSQAVTAVSGSMPGAHGSNRSVPRP MNVVGMPRLPPQAMAAYNLTSQA
AMGDGMNPGSISKQRGVPQGHPPQLRRKEQMGMPGYPSQQ KSRRI

>Phavul_0028

MVGETGKLFGLGCDPVVGSLSASKKREYRVTNRLQEGKRPLYAVVFNFIDSRYFNVFATV
GGNRVTVYQCLEEGVIAVLQSYVDEKDESFYTVSWACNVDGAPFVVAGGINGVIRVIDAGTEKIH
KSFVGHGDSVNEIRTQTLRPSLVISASKDESURLWNVHTGICILIFAGAGGHRNEVLSVDFHPSDIYRI
ASCGMDNTVKIWSMKEFWTYVEKSFTWTDLPKSFPTKYVQFPVYNASVHLNYVDCNRWLGDFFILS
KSDVDNEIVLWEPKVKEQTPGEGVVDILQKYPECDIWFIFKFSCDFHFDIAAVGNREGKIFIWELQSSP
PVLIALRSLHPQSKSPIRQTATSFDFGSTILSCCEDGTIWRWDDVTNSTT

>Vigan_0033

MGDGSRSRWSRAEWWQYDLLGKIGEGTYGLVFLARTKTPSPSKSIAIKKFKQSKDGDGVS
PTAIREIMLLREITHENVVKLNVNHADMSLYLAFDYAEHDLYEIRHHRDKLNHSINQYTVKSLL
WQLLNGLSYLHSNMIHRDLKPSNILVMGEGEEHG VVKIADFGFLARIYQAPLKPLSDNGVVVTI
RAPELLGAKHYTSAVDMWAVGCIFAELLTLKPLFQGAEVKATSNPFQLDQLDKIFKVLGHPTLEK
WPSLASLPHWQQDVQHIQGHKYDNTGLYNVHLSTKSPAYDLLSKMLEYDPRKRITAAQALEHEY
FKIEPLPGRNALVPCQLGEKIVNYPTRPVDNTDLEGTTLNPASQAVTAVSGSMPGAHGSNRVPRP
MNVVGMQRMPPQMAAYNLTSQAAMGDGMNPGNISKQRGVPQGHPPQLRRKEQMGMPPGYPSQQ
KSRRI

>Vigan_0038

MGYICDFCGDQRSMVYCRSDAACLCLSCDRNVHSANALSRRHSRTLLCERCNSQPAFVRC
VEEKISLCQNCDWLGHGVSTSSSTHQRGISCYSGCPSAAELSSIWSFVLDIPSIRESTCEQELGLMSIN
ENSESVGVPPENGNVSVSDQVTDLPALDRSFVGTSSIPESSETRILDQPDGPANECLPKLYCPATKCP
ALCEDDNLVYDDFDMDEVDLNLNENYEELFGMALSHSEELFENGGIDSLFGTKDMSAGDSSCQDAIAA
EGSSVGLVNATQPACSNAAASADSILSTKTEPILCFTGRQAQSNLSFSGITGESSAADYQDCGTTSMML
MGEPPWFAPCPENSLQSANRSNAVRLRYKEKKKARKFEKQVRYASRKARADIRRRVKGRFVKAGDV
YDYDPLSTTRSC

>Vigan_0067

MTDLHQFNNTLGGRRGGTNPQGMKLSHSGGIWKRQGGGKLEIVDKSDIEGVTWMKVPKT
NQLGVQIKDGLYKFTGFRDQDVVSLTNFFQSTCGITVREKQLSVSGRNWGEVDLNGNMLAFMVG
SKQAFEVSLADVSQTQLQGKNDVILEFHVDDTTGANEKDSLMEISFHIPSSNTQFVGDEKTPPAEVFR
SRIMSMADV GAGGEDAVVTFESIAILTPRGRYIVELHMSFLRLQGQANDFKIQYSSVVRLLPKSNQ
PHTFVIISLDPPIRKGQTLYPHIVMQFESDYVVQSELAMNEDLYNTKYKDKLELSYKALIEVFTTILR
GLSGAKVTKPGKFRSCQDGYAVKSSLKAEDGILYPLEKSFFFLPKPPTLILHEEIDYVEFERHAAGGS
NMHYFDLLIRLQSEQHLFRNIQRNEYHNL YEFISSKGLKIMNLGDGQPTAGIKKVLNDDDDAVDP
HLERIKNAAGDDESDEEDSDFVADKDDGSPDSDSGGDDSDASKSGEEKEKPAKKEIKKDLPPKAST
SKKKPKDDEDGKKKKKQKKKDPNAPKRAMSGFMFFSKLERENLKKTNPGISFTDVGRVLGEKWKK
MSAEEKEPYEAKAREDKQRYKDEISGYKNTQPMNIDSANVTDSA

>Vigan_0068

MTISSSSKKHGQDLPELSGPTPKSQRKMGKSDDNDKKNLGKCLKDVEISVPIVYGNIAFW
LGKKASEYQSHKWTVYVRGATNEDLGAIIKRVVFLHSSFNPNTRVVEPPFELSEAGWGEFEVAIT
LYFHSDVCDKPLNLYHHLKLYPEDENSSMSAKKPVVVEFYDEIVFPDPSEAFARLQSHPAVNLPRL
PAGLTLPPPVPVEDASKRRKGDTKDHSLSQWFTNFSEADELLQLAAARQQVQAHIAKLRRQINLIDG
QHQQFKSSSDQ

>Vigra_0043

MGYICDFCGDQRSMVYCRSDAACLCLSCDRNVHSANALSRRHSRTLLCERCNSQPAFVRC
VEEKISLCQNCDWLGHGVSTSSSTHQRGISCYSGCPSAAELSSIWSFVLDIPSIRESTCEQELGLMSIN
ENSESFGVPPENGNVSVSDQVTDLPALDRSFVGTSSIPVSNSETHILDQPAGPANECLPKLYCPATKCP
ALCEDDNLVYDDFDMDEVDLNLNENYEELFGMALSHSEELFENGGIDSLFGTKDMSAGDSSCQDAIAA
EGSSVGLVNATQPACSNAAASADSILSTKTEPILCFTGRQAQSNLSFSGITGESSAADYQDCGASSMML
MGEPPWFAPCPENSLQSANRSNAVRLRYKEKKKARKFEKQVRYASRKARADVRRRVKGRFVKAGD
VYDYDPLSTTRSC

>Vigra_0072

MTISSSSKKHGQDQPELSGPTPKSQRKMGKSDDNDKKNLGKCLKDVEISVPIVYGNIAFW
LGKKASEYQSHKWTVYVRGATNEDLGAIIKRVVFLHSSFNPNTRVVEPPFELSEAGWGEFEVAIT
LYFHSDVCDKPLNLYHHLKLYPEDENSSMSAKKPVVVEFYDEIVFPDPSEAFARLQSHPAVNLPRL
PAGLTLPPPVPVEDASKRRKGDTKDHSLSQWFTNFSEADELLQLAAARQQVQAHIAKLRRQINLIDG
QHQQFKSSSDQ

>Vigun_0040

MGYICDFCGEQRSMVYCRSDAACLCLSCDRSVHSANALSRRHSRTLLCERCNSQAAFVRC
VEEKISVCQNCDWLGDVSTSWTHKRQGISCYSGCPSAAELSSIWSFVLDIPSIRECTTCEQELGLMSI
NENSESVGVATNEKCPALCCEDFDMDEVDLNIEKYEELFGSEELFENGGIDSLFGTKDMSAEVNACS
NAASADSILSTKTEPIVCFSGITGDYQDCGASSSMVLMGEHPCPDNSLHSANRTNAVLRVYKEKKKAR
KFDKQVRYASRKARADVRRRVKGRFVKADYDPLTTTRSC

>Vigun_0069

MTNSSSSKKHGQDQPELSGPTPKSQRTKMGKSEDNDKKNLGGKLDVEISVPIVYGNIAFW
LGKKASEYQSHKWTVYVRGATNEDLGAVIKRVVFLHSSFNPNTRVVESEPPFELSEAGWGEFEVAI
TLYFHSDVCDKPLNLYHHLKLYPEDENSSMSAKKPVVVEFYDEIVFPDPSEAFLSRLQSHPAVNLPR
LPPGLTLPPPVPVEDASKRRRKGDTKDHSLSQWFTNFSEADELLQLAAARQQVQAHIAKLRRQINLID
GQHQQFKSSSDQ

Anexo 5. Genes de floración de *A. thaliana* en sus rutas respectivas. Las siglas representan la abreviación del nombre de las rutas que están en la parte superior del cuadro (Bouche *et al.*, 2016). Las siglas representadas en cada código de gen representa pertenencia, mientras la “X” es no presencia en la ruta.

Código	Nombre	General processes & autonomous pathway	Aging pathway	Circadian clock pathway	Flower development and meristem identity	Gene involve to Flowering	Hormonal pathway	Photoperiod pathway	Sugar pathway	Temperature pathway	Vernalization pathway
AT1G01040	DCL1	GPAP	AP	X	X	X	X	X	X	X	X
AT1G02065	SPL8	X	X	X	FDMI	X	X	X	X	X	X
AT1G02400	GA2ox5	X	X	X	X	X	HP	PP	X	X	X
AT1G02580	MEA, FIS1	X	X	X	X	X	X	X	X	X	VP
AT1G02740	MRC2	GPAP	X	X	X	X	X	PP	X	X	X
AT1G04400	CRY2	X	X	X	X	X	X	PP	X	X	X
AT1G04870	ATPRMT10	GPAP	X	X	X	X	X	X	X	X	X
AT1G05830	ATX2	GPAP	X	X	X	X	X	PP	X	X	VP
AT1G06040	STO	X	X	X	X	X	X	PP	X	X	X
AT1G06770	AtBMI1B, DRIP1	X	X	X	X	X	X	X	X	X	VP
AT1G08260	ESD7, TIL1, ABO4	GPAP	X	X	X	X	X	X	X	X	X
AT1G08970	NF-YC9, HAP5C	X	X	X	X	X	HP	PP	X	X	X
AT1G09570	PHYA	X	X	X	X	X	X	PP	X	X	X
AT1G09700	HYL1	X	AP	X	X	X	X	X	X	X	X
AT1G10120	CIB4	X	X	X	X	X	X	PP	X	X	X
AT1G10570	OTS2	GPAP	X	X	X	X	X	X	X	X	X
AT1G12910	LWD1, ATAN11	X	X	CCP	X	X	X	X	X	X	X
AT1G14400	UBC1	GPAP	X	X	X	X	X	X	X	X	VP
AT1G14920	GAL, RGA2	X	AP	X	X	X	HP	PP	X	X	X
AT1G15550	GA3OX1	X	X	X	X	X	HP	X	X	X	X
AT1G15750	TPL	X	AP	X	FDMI	X	X	PP	X	X	X
AT1G17720	PP2A-B55BETA	X	X	X	X	GIF	X	X	X	X	X
AT1G17760	CSTF77	GPAP	X	X	X	X	X	X	X	X	VP
AT1G18080	RACK1A, ATARCA	X	X	X	X	X	HP	X	X	X	X
AT1G18100	MFT	X	X	X	X	X	X	PP	X	X	X
AT1G18450	ARP4	GPAP	X	X	X	X	X	X	X	X	VP
AT1G19330	AFR2	X	X	X	X	X	X	PP	X	X	X
AT1G20900	AHL27, ESC	X	X	X	X	X	X	PP	X	X	X
AT1G22770	GI	X	AP	CCP	X	X	X	PP	SP	X	X
AT1G23420	INO	X	X	X	FDMI	X	X	X	X	X	X
AT1G24260	SEP3	X	X	X	FDMI	X	X	X	X	X	X
AT1G25540	PFT1, MED25	X	X	X	X	X	X	PP	X	X	X
AT1G25560	TEM1, EDF1	X	X	X	X	X	HP	PP	X	X	X
AT1G26310	CAL	X	X	X	FDMI	X	X	X	X	X	X
AT1G26830	CUL3A	X	X	X	X	X	X	X	X	X	VP
AT1G28210	ATJ1	X	X	X	X	GIF	X	X	X	X	X
AT1G28520	VOZ1	X	X	X	X	X	X	PP	X	X	X
AT1G30040	GA2ox2	X	X	X	X	X	HP	PP	X	X	X
AT1G30950	UFO	X	X	X	FDMI	X	X	X	X	X	X
AT1G30970	SUF4	X	X	X	X	X	X	X	X	X	VP
AT1G32230	RCD1	GPAP	X	X	X	X	X	X	X	X	X
AT1G43850	SEU	X	X	X	FDMI	X	X	X	X	X	X
AT1G47990	GA2ox4	X	X	X	X	X	HP	PP	X	X	X
AT1G48270	GCR1	X	X	X	FDMI	X	HP	PP	X	X	X
AT1G50370	FYPP1	X	X	X	X	X	X	PP	X	X	X
AT1G50700	CPK33	X	X	X	X	X	X	PP	X	X	X
AT1G51450	ASH2R, TRO	GPAP	X	X	X	X	X	X	X	X	VP
AT1G51690	PP2A-B55ALPHA	X	X	X	X	GIF	X	X	X	X	X
AT1G52520	FRS6	X	X	X	X	X	X	PP	X	X	X
AT1G52740	HTA9	GPAP	X	X	X	X	X	X	X	X	X
AT1G53090	SPA4	X	X	X	X	X	X	PP	X	X	X
AT1G54440	RRP6L1	GPAP	X	X	X	X	X	X	X	X	X
AT1G54830	NF-YC3	X	X	X	X	X	HP	PP	X	X	X
AT1G55250	HUB2	GPAP	X	X	X	X	X	X	X	X	VP
AT1G55325	MED13, GCT, MAB2	GPAP	X	X	X	X	X	X	X	X	X

Codigo	Nombre	General processes & autonomous pathway	Aging pathway	Circadian clock pathway	Flower development and meristem identity	Gene involve to Flowering	Hormonal pathway	Photoperiod pathway	Sugar pathway	Temperature pathway	Vernalization pathway
AT1G56170	ATHAP5B, HAP5B, NF-YC2	X	X	X	X	GIF	X	X	X	X	X
AT1G57820	VIM1, ORTH2	GPAP	X	X	X	X	X	X	X	X	X
AT1G60220	OTS1, ULP1D	GPAP	X	X	X	X	X	X	X	X	X
AT1G61040	VIP5	GPAP	X	X	X	X	X	X	X	X	VP
AT1G62360	STM, SHL, WAM1	X	X	X	FDMI	X	X	X	X	X	X
AT1G62830	LDL1	GPAP	X	X	X	X	X	PP	X	X	X
AT1G65480	FT	X	AP	CCP	FDMI	X	HP	PP	SP	TP	VP
AT1G66050	VIM2	GPAP	X	X	X	X	X	X	X	X	X
AT1G66350	RGL1	X	AP	X	X	X	HP	PP	X	X	X
AT1G68050	FKF1, ADO3	X	X	CCP	X	X	X	PP	X	X	X
AT1G68640	PAN, TGA8	X	X	X	FDMI	X	X	X	X	X	X
AT1G68840	TEM2, RAV2	X	X	X	X	X	HP	PP	X	TP	X
AT1G69120	AP1	X	AP	X	FDMI	X	X	PP	SP	TP	VP
AT1G71692	AGL12, XAL1	X	X	X	X	GIF	X	X	X	X	X
AT1G75060	AFR1	X	X	X	X	X	X	PP	X	X	X
AT1G76500	AHL29, SOB3	X	X	X	X	X	X	PP	X	X	X
AT1G76710	SDG26, ASHH1	GPAP	X	X	X	X	X	X	X	X	X
AT1G78440	GA2ox1	X	X	X	X	X	HP	X	X	X	X
AT1G78580	TPS1	X	AP	X	X	X	X	X	SP	X	X
AT1G79280	NUA	X	X	X	X	GIF	X	X	X	X	X
AT1G79430	APL, WDY, FE	X	X	X	X	X	X	PP	X	X	X
AT1G79460	GA2, ATKS1	X	X	X	X	X	HP	X	X	X	X
AT1G79730	ELF7	GPAP	X	X	X	X	X	X	X	X	VP
AT1G80010	FRS8	X	X	X	X	X	X	PP	X	X	X
AT1G80070	PRP8, SUS2, EMB33	GPAP	X	X	X	X	X	X	X	X	X
AT1G80340	GA3OX2	X	X	X	X	X	HP	PP	X	X	X
AT2G01570	RGA1	X	AP	X	X	X	HP	PP	X	X	X
AT2G02760	UBC2	GPAP	X	X	X	X	X	X	X	X	VP
AT2G03500	EFM	X	X	X	X	X	X	X	X	TP	X
AT2G03710	SEP4	X	X	X	FDMI	X	X	X	X	X	X
AT2G06210	ELF8, VIP6	GPAP	X	X	X	X	X	X	X	X	VP
AT2G13540	ABH1, CBP80	GPAP	AP	X	X	X	X	X	X	X	VP
AT2G17290	CPK6	X	X	X	X	X	X	PP	X	X	X
AT2G18000	YAF9B, TAF14	X	X	X	X	X	X	X	X	X	VP
AT2G18040	PIN1AT	X	X	X	FDMI	X	X	X	X	X	X
AT2G18790	PHYB	X	X	X	X	X	X	PP	X	X	X
AT2G18915	LKP2	X	X	CCP	X	X	X	PP	X	X	X
AT2G19520	FVE, MSI4	GPAP	X	X	X	X	X	X	X	TP	VP
AT2G20570	GLK1, GPR11	GPAP	X	X	X	X	X	X	X	X	X
AT2G20825	ULT2	X	X	X	FDMI	X	X	X	X	X	X
AT2G21070	FIO1	X	X	CCP	X	X	X	X	X	X	X
AT2G22370	MED18	GPAP	X	X	X	X	X	X	X	X	X
AT2G22540	SVP	X	AP	X	FDMI	X	X	PP	X	TP	VP
AT2G22630	AGL17	X	X	X	X	X	X	PP	X	X	X
AT2G23380	CLF, SDG1	GPAP	X	X	FDMI	X	X	PP	X	X	VP
AT2G23740	SUVR5	GPAP	X	X	X	X	X	X	X	X	X
AT2G25170	PKL, SSL2, CKH2	X	X	X	FDMI	X	X	X	X	X	X
AT2G27550	ATC	X	X	X	X	X	X	PP	X	X	X
AT2G28550	TOE1, RAP2.7	X	AP	X	FDMI	X	HP	PP	X	TP	VP
AT2G28830	PUB12	GPAP	X	X	X	X	X	X	X	X	X
AT2G30140	UGT87A2	GPAP	X	X	X	X	X	X	X	X	X
AT2G30580	AtBM11A, DRIP2	X	X	X	X	X	X	X	X	X	VP
AT2G31650	ATX1, SDG27	GPAP	X	X	FDMI	X	X	PP	X	X	VP
AT2G32950	COP1	X	X	X	X	X	X	PP	X	X	X
AT2G33380	RD20, PXG3, ATCLO3	X	X	X	X	X	HP	X	X	X	X
AT2G33860	ARF3, ETT	X	X	X	FDMI	X	X	X	X	X	X
AT2G34140	CDF4	X	X	X	X	X	X	PP	X	X	X
AT2G34555	GA2ox3	X	X	X	X	X	HP	PP	X	X	X
AT2G34720	NF-YA4	X	X	X	X	X	HP	PP	X	X	X

Codigo	Nombre	General processes & autonomous pathway	Aging pathway	Circadian clock pathway	Flower development and meristem identity	Gene involve to Flowering	Hormonal pathway	Photoperiod pathway	Sugar pathway	Temperature pathway	Vernalization pathway
AT2G34880	JMJ15, MEE27	GPAP	X	X	X	X	X	X	X	X	X
AT2G35510	SRO1	GPAP	X	X	X	X	X	X	X	X	X
AT2G37630	AS1	X	X	X	X	X	HP	PP	X	X	X
AT2G38810	HTA8	GPAP	X	X	X	X	X	X	X	X	X
AT2G38880	NF-YB1, HAP3A	X	X	X	X	X	HP	PP	X	X	X
AT2G39250	SNZ	X	AP	X	X	X	X	PP	X	TP	X
AT2G39360	CVY1	X	X	X	X	GIF	X	X	X	X	X
AT2G39810	ESD6, HOS1	X	X	X	X	X	X	PP	X	X	X
AT2G40360	BOP2	X	X	X	FDMI	X	X	X	X	X	X
AT2G42200	SPL9	X	AP	X	X	X	HP	X	SP	X	X
AT2G42280	FBH4, AKS3	X	X	X	X	X	X	PP	X	X	X
AT2G42400	VOZ2	X	X	X	X	X	X	PP	X	X	X
AT2G42830	SHP2, AGL5	X	X	X	FDMI	X	X	X	X	X	X
AT2G44150	SDG7, ASSH3	X	X	X	X	X	X	X	X	X	VP
AT2G44680	CKB4	X	X	CCP	X	X	X	X	X	X	X
AT2G44950	HUB1	GPAP	X	X	X	X	X	X	X	X	VP
AT2G45430	AHL22	GPAP	X	X	X	X	X	PP	X	X	X
AT2G45640	SAP18	X	X	X	FDMI	X	X	X	X	X	X
AT2G45650	AGL6	GPAP	X	X	X	X	X	X	X	X	X
AT2G45660	SOCI, AGL20	X	AP	X	FDMI	X	HP	PP	SP	TP	VP
AT2G46020	BRM, CHR2	GPAP	X	X	X	X	X	X	X	X	X
AT2G46260	LRB1	X	X	X	X	X	X	X	X	X	VP
AT2G46340	SPA1	X	X	X	X	X	X	PP	X	X	X
AT3G01090	AKIN10, SNRK1.1	X	X	X	X	X	X	X	SP	X	X
AT3G01890	SWP73A	X	X	X	FDMI	X	X	X	X	X	X
AT3G02310	SEP2	X	X	X	FDMI	X	X	X	X	X	X
AT3G03450	RGL2	X	AP	X	X	X	HP	PP	X	X	X
AT3G04030	MYR2	X	X	X	X	X	X	PP	X	X	X
AT3G04610	FLK	GPAP	X	X	X	X	X	X	X	X	X
AT3G04910	ATWNK1, WNK1, ZIK4	X	X	CCP	X	X	X	X	X	X	X
AT3G05040	HST1	X	AP	X	X	X	X	X	X	X	X
AT3G05120	GID1A	X	AP	X	X	X	HP	PP	X	X	X
AT3G07650	COL9	X	X	X	X	X	X	PP	X	X	X
AT3G09600	RVE8	X	X	CCP	X	X	X	X	X	X	X
AT3G10390	FLD, RSI1	GPAP	X	X	X	X	X	X	X	X	VP
AT3G10480	NAC050	X	X	X	FDMI	X	X	PP	X	X	X
AT3G10490	NAC052	X	X	X	FDMI	X	X	PP	X	X	X
AT3G11440	MYB65	X	X	X	X	X	HP	X	X	X	X
AT3G11910	UBP13	GPAP	X	X	X	X	X	X	X	X	X
AT3G12145	FLR1, FTM4	X	X	X	X	X	X	PP	X	X	X
AT3G12680	HUA1	X	X	X	FDMI	X	X	X	X	X	X
AT3G12810	PIE1	GPAP	X	X	X	X	X	X	X	X	VP
AT3G13682	LDL2	GPAP	X	X	X	X	X	PP	X	X	X
AT3G15354	SPA3	X	X	X	X	X	X	PP	X	X	X
AT3G18990	VRN1	X	X	X	X	X	X	X	X	X	VP
AT3G19980	FYPP3	X	X	X	X	X	X	PP	X	X	X
AT3G20740	FIE1, FIS3	GPAP	X	X	FDMI	X	X	X	X	X	VP
AT3G20810	JMJ30	GPAP	X	X	X	X	X	PP	X	TP	X
AT3G22590	PHP, CDC73	X	X	X	X	X	X	X	X	X	VP
AT3G23050	IAA7, AXR2	X	X	X	X	GIF	X	X	X	X	X
AT3G24440	VIL1, VRN5	X	X	X	X	X	X	X	X	X	VP
AT3G26640	LWD2	X	X	CCP	X	X	X	X	X	X	X
AT3G26744	ICE1, SCR1	X	X	X	X	X	X	X	X	X	VP
AT3G28730	SSRP1, HMG, NFD	GPAP	X	X	X	X	X	X	X	X	VP
AT3G28910	MYB30	GPAP	X	X	X	X	X	X	X	X	X
AT3G33520	ARP6, SUF3, ESD1	GPAP	X	X	X	X	X	X	X	TP	VP
AT3G43190	SUS4	X	X	X	X	X	X	X	SP	X	X
AT3G43920	DCL3	GPAP	X	X	X	X	X	X	X	X	X
AT3G44110	ATJ3	GPAP	X	X	X	X	X	X	X	X	X

Codigo	Nombre	General processes & autonomous pathway	Aging pathway	Circadian clock pathway	Flower development and meristem identity	Gene involve to Flowering	Hormonal pathway	Photoperiod pathway	Sugar pathway	Temperature pathway	Vernalization pathway
AT3G44680	HDA9	GPAP	X	X	X	X	X	X	X	X	X
AT3G45880	JMJ32, MJJD5	GPAP	X	X	X	X	X	X	X	X	X
AT3G46510	PUB13	GPAP	X	X	X	X	X	X	X	X	X
AT3G46640	LUX, PCL1	X	X	CCP	X	X	X	X	X	TP	X
AT3G47500	CDF3	X	X	X	X	X	X	PP	X	X	X
AT3G48430	REF6, MJJ12	GPAP	X	X	X	X	X	PP	X	X	X
AT3G48590	NF-YC1, HAP5A	X	X	X	X	X	X	PP	X	X	X
AT3G49600	UBP26	GPAP	X	X	X	X	X	X	X	X	VP
AT3G49660	WDR5A	X	X	X	X	X	X	X	X	X	VP
AT3G51060	STY1, SRS1	X	X	X	X	GIF	X	X	X	X	X
AT3G54340	AP3	X	X	X	FDMI	X	X	X	X	X	X
AT3G54560	HTA11	GPAP	X	X	X	X	X	X	X	X	X
AT3G55630	FPGS3, DFD	X	X	X	X	GIF	X	X	X	X	X
AT3G57130	BOP1	X	X	X	FDMI	X	X	X	X	X	X
AT3G57230	AGL16	GPAP	X	X	X	X	X	X	X	X	X
AT3G57300	INO80	GPAP	X	X	X	X	X	X	X	X	X
AT3G57390	AGL18	X	X	X	X	X	X	PP	X	X	X
AT3G58780	SHP1, AGL1	X	X	X	FDMI	X	X	X	X	X	X
AT3G60250	CKB3	X	X	CCP	X	X	X	X	X	X	X
AT3G60870	AHL18	X	X	X	FDMI	X	X	X	X	X	X
AT3G61250	LM12, MYB17	X	X	X	FDMI	X	X	X	X	X	X
AT3G61600	LRB2, POB1	X	X	X	X	X	X	X	X	X	VP
AT3G63010	GID1B	X	AP	X	X	X	HP	PP	X	X	X
AT4G00450	MED12, CCT, CRP	GPAP	X	X	X	X	X	X	X	X	X
AT4G00830	LIF2	GPAP	X	X	X	X	X	X	X	X	VP
AT4G02020	SWN	X	X	X	X	X	X	X	X	X	VP
AT4G02780	GA1	X	X	X	X	X	HP	X	X	X	X
AT4G04920	MED16, SFR6	GPAP	X	X	X	X	X	X	X	X	X
AT4G08920	CRY1	X	X	X	X	X	X	PP	X	X	X
AT4G10180	DET1, FUS2	GPAP	X	X	X	X	X	X	X	X	X
AT4G10710	SPT16	GPAP	X	X	X	X	X	X	X	X	VP
AT4G11110	SPA2	X	X	X	X	X	X	PP	X	X	X
AT4G11880	XAL2, AGL14	X	X	X	FDMI	X	X	X	X	X	X
AT4G14540	NF-YB3, HAP3C	X	X	X	X	X	HP	PP	X	X	X
AT4G15880	ESD4	GPAP	X	X	X	X	X	X	X	X	X
AT4G16110	RR2, ARR2	X	X	X	X	X	HP	X	X	X	X
AT4G16250	PHYD	X	X	X	X	X	X	PP	X	X	X
AT4G16280	FCA	GPAP	AP	X	X	X	X	X	X	TP	VP
AT4G16845	VRN2	X	X	X	X	X	X	X	X	X	VP
AT4G18130	PHYE	X	X	X	X	X	X	PP	X	X	X
AT4G18960	AG	X	X	X	FDMI	X	X	X	X	X	X
AT4G20370	TSF	X	X	X	X	X	HP	PP	X	TP	X
AT4G20400	JMJ14, PKDM7B, MJJ4	GPAP	X	X	X	X	X	PP	X	X	X
AT4G21200	GA2o8	X	X	X	X	X	HP	PP	X	X	X
AT4G22140	EBS	X	X	X	X	X	X	PP	X	X	X
AT4G22950	AGL19	X	X	X	X	X	X	X	X	X	VP
AT4G23100	CAD2, RML1, CSH1	GPAP	X	X	X	X	X	X	X	X	X
AT4G24210	SLY1	X	X	X	X	X	HP	X	X	X	X
AT4G24540	AGL24	X	X	X	FDMI	X	X	X	X	X	VP
AT4G24620	PGI1	X	X	X	X	X	X	X	SP	X	X
AT4G24680	MOS1	GPAP	X	X	X	X	X	X	X	X	X
AT4G25000	AMY1	X	X	X	X	GIF	X	X	X	X	X
AT4G25420	GA2o8x1	X	X	X	X	X	HP	PP	X	X	X
AT4G25490	CBF1, DREB1B	X	X	X	X	X	X	X	X	X	VP
AT4G26440	WRKY34, MSP3	X	X	X	X	X	X	X	X	X	VP
AT4G28190	ULT1	X	X	X	FDMI	X	X	X	X	X	X
AT4G29130	HXK1, GIN2	X	AP	X	X	X	X	X	SP	X	X
AT4G29830	VIP3, SKI8	GPAP	X	X	X	X	X	X	X	X	VP
AT4G30200	VIL2, VEL1	GPAP	X	X	X	X	X	X	X	X	VP

Codigo	Nombre	General processes & autonomous pathway	Aging pathway	Circadian clock pathway	Flower development and meristem identity	Gene involve to Flowering	Homonal pathway	Photoperiod pathway	Sugar pathway	Temperature pathway	Vernalization pathway
AT4G31120	SKB1, PRMT5	GPAP	X	X	X	X	X	X	X	X	X
AT4G32551	LUG	X	X	X	FDMI	X	X	X	X	X	X
AT4G32980	ATH1	X	X	X	X	X	HP	X	X	X	X
AT4G34530	CIB1	X	X	X	X	X	X	PP	X	X	X
AT4G34710	ADC2	X	X	X	X	X	HP	X	X	X	X
AT4G36920	AP2	X	AP	X	FDMI	X	X	PP	X	X	X
AT4G37280	MRG1	GPAP	X	X	X	X	X	PP	X	X	X
AT4G38130	HDA1, HDA19,	X	X	X	FDMI	X	X	X	X	X	X
AT4G38960	BBX19	X	X	X	X	X	X	PP	X	X	X
AT4G39100	SHL1, SHL	GPAP	X	X	X	X	X	X	X	X	X
AT4G39400	ATBRI1, BIN1, BRI1, CBB2, DWARF 2, DWF2	X	X	X	X	GIF	X	X	X	X	X
AT4G40060	HB16, ATHB16	X	X	X	X	X	X	PP	X	X	X
AT5G01040	LAC8	X	X	X	X	GIF	X	X	X	X	X
AT5G02030	PNY, RPL, BLR	X	X	X	FDMI	X	X	X	X	X	X
AT5G02810	PRR7	X	X	CCP	X	X	X	X	X	X	X
AT5G03680	PTL	X	X	X	FDMI	X	X	X	X	X	X
AT5G03790	LM11, HB51	X	X	X	FDMI	X	X	X	X	X	X
AT5G03840	TFL1	X	X	X	FDMI	X	X	X	X	X	X
AT5G04240	ELF6	GPAP	X	X	X	X	X	PP	X	X	X
AT5G05980	FPCS1, DFB	X	X	X	X	GIF	X	X	X	X	X
AT5G06100	MYB33	X	X	X	X	X	HP	X	X	X	X
AT5G06170	SUC9	X	X	X	X	X	X	X	SP	X	X
AT5G06600	UBP12	GPAP	X	X	X	X	X	X	X	X	X
AT5G06850	FTIP1	X	X	X	X	X	X	PP	X	X	X
AT5G07200	GA20OX3	X	X	X	X	X	HP	PP	X	X	X
AT5G08330	ATTCP21, TCP21	X	X	CCP	X	X	X	X	X	X	X
AT5G09740	HAM2, HAG5, HAC11	GPAP	X	X	X	X	X	X	X	X	X
AT5G10030	TGA4	X	X	X	X	X	X	PP	X	X	X
AT5G12840	NF-YA1, HAP2B	X	X	X	X	X	X	PP	X	X	X
AT5G13320	WIN3, GDG1, GH3.12	X	X	X	X	GIF	X	X	X	X	X
AT5G13480	FY	GPAP	X	X	X	X	X	X	X	X	VP
AT5G13790	AGL15	X	X	X	X	X	X	PP	X	X	X
AT5G14170	BAF60, CHC1	GPAP	X	X	X	X	X	X	X	X	X
AT5G15800	SEP1	X	X	X	FDMI	X	X	X	X	X	X
AT5G15840	CO	X	X	CCP	X	X	X	PP	X	X	X
AT5G17490	RGL3	X	AP	X	X	X	HP	PP	X	X	X
AT5G18240	MYR1	X	X	X	X	X	X	PP	X	X	X
AT5G20240	PI	X	X	X	FDMI	X	X	X	X	X	X
AT5G20320	DCL4	GPAP	X	X	X	X	X	X	X	X	X
AT5G24860	FPF1	X	X	X	X	X	HP	X	X	X	X
AT5G27320	GID1C	X	AP	X	X	X	HP	PP	X	X	X
AT5G35840	PHYC	X	X	X	X	X	X	PP	X	X	X
AT5G35910	RRP6L2	GPAP	X	X	X	X	X	X	X	X	X
AT5G37055	SEF, SWC6	GPAP	X	X	X	X	X	X	X	X	VP
AT5G37260	RVE2, CIR1	X	X	CCP	X	X	X	X	X	X	X
AT5G39550	VIM3, ORTH1	GPAP	X	X	X	X	X	X	X	X	X
AT5G40490	HLP1	GPAP	X	X	X	X	X	X	X	X	VP
AT5G44160	IDD8, NUC	X	X	X	X	X	X	X	SP	X	X
AT5G44200	CBP20	GPAP	AP	X	X	X	X	X	X	X	VP
AT5G45600	YAF9A, TAF14B, GAS41	GPAP	X	X	X	X	X	X	X	X	VP
AT5G46210	CUL4	GPAP	X	X	X	X	X	PP	X	X	X
AT5G47640	NF-YB2, HAP3B	X	X	X	X	X	HP	PP	X	X	X
AT5G48300	ADGL1, APS1	X	X	X	X	X	X	X	SP	X	X
AT5G49020	PRMT4A	GPAP	X	X	X	X	X	X	X	X	X
AT5G51230	EMF2, CYR1	GPAP	X	X	FDMI	X	X	X	X	X	VP
AT5G51810	GA20ox2	X	X	X	X	X	HP	PP	X	X	X
AT5G51820	PGM1	X	X	X	X	X	X	X	SP	X	X
AT5G51860	AGL72	X	X	X	FDMI	X	X	X	X	X	X
AT5G51870	AGL71	X	X	X	FDMI	X	X	X	X	X	X

Codigo	Nombre	General processes & autonomous pathway	Aging pathway	Circadian clock pathway	Flower development and meristem identity	Gene involve to Flowering	Hormonal pathway	Photoperiod pathway	Sugar pathway	Temperature pathway	Vernalization pathway
AT5G57360	ZTL	X	X	CCP	X	X	X	PP	X	X	X
AT5G57660	COL5	X	X	X	X	X	X	PP	X	X	X
AT5G58230	MS11	GPAP	X	X	FDMI	X	X	PP	X	X	VP
AT5G59560	SRR1	X	X	CCP	X	X	X	PP	X	X	X
AT5G59570	BOA	X	X	CCP	X	X	X	X	X	X	X
AT5G60100	PRR3	X	X	CCP	X	X	X	X	X	X	X
AT5G60120	TOE2	X	AP	X	X	X	X	PP	X	TP	X
AT5G60410	SIZ1	GPAP	X	X	X	X	X	X	X	X	X
AT5G60800	HIPP3	X	X	X	X	GIF	X	X	X	X	X
AT5G60910	FUL, AGL8	X	AP	X	X	X	X	PP	SP	TP	VP
AT5G61060	HDA5	GPAP	X	X	X	X	X	X	X	X	VP
AT5G61150	VIP4	GPAP	X	X	X	X	X	X	X	X	VP
AT5G61380	TOC1	X	X	CCP	X	X	X	X	X	X	X
AT5G61850	LFY	X	AP	X	FDMI	X	HP	PP	SP	TP	VP
AT5G62000	ARF2	X	X	X	X	X	HP	X	X	X	X
AT5G62040	BFT	GPAP	X	X	X	X	X	X	X	X	X
AT5G62165	AGL42, FYF	X	X	X	FDMI	X	X	X	X	X	X
AT5G63110	HDA6, RTS1	GPAP	X	X	X	X	X	X	X	X	VP
AT5G63470	NF-YC4	X	X	X	X	X	X	PP	X	X	X
AT5G63610	HEN3, CDKE;1	X	X	X	FDMI	X	X	X	X	X	X
AT5G63960	GIS5, EMB2780	GPAP	X	X	X	X	X	X	X	X	X
AT5G64610	HAM1	GPAP	X	X	X	X	X	X	X	X	X
AT5G64813	LIP1	X	X	X	X	GIF	X	X	X	X	X
AT5G64960	CDKC;2	GPAP	X	X	X	X	X	X	X	X	VP
AT5G67180	TOE3	X	AP	X	X	X	X	PP	X	X	VP
AT1G01060	LHY	X	X	CCP	X	X	X	X	X	X	X
AT1G09530	PIF3	X	X	CCP	X	X	X	X	X	X	X
AT1G18075	MIR159B	X	X	X	X	X	HP	X	X	X	X
AT1G22310	MBD8	X	X	X	X	GIF	X	X	X	X	X
AT1G26260	CIB5	X	X	X	X	X	X	PP	X	X	X
AT1G31814	FRL2	X	X	X	X	X	X	X	X	X	VP
AT1G35460	FBH1	X	X	X	X	X	X	PP	X	X	X
AT1G50960	GA2ox7	X	X	X	X	X	HP	PP	X	X	X
AT1G51140	FBH3, AKS1, BHLH122	X	X	X	X	X	X	PP	X	X	X
AT1G53160	SPL4	X	AP	X	X	X	HP	PP	SP	X	X
AT1G66650	AT-STUb4	X	X	X	X	X	X	PP	X	X	X
AT1G69570	CDF5	X	X	X	X	X	X	PP	X	X	X
AT1G71800	CSTF64	GPAP	X	X	X	X	X	X	X	X	VP
AT1G72390	PHL	X	X	X	X	X	X	PP	X	X	X
AT1G73687	MIR159A	X	X	X	X	X	HP	X	X	X	X
AT1G77080	FLM, MAF1, AGL27	X	X	X	X	X	X	X	X	TP	VP
AT1G77300	EFS, SDG8	GPAP	X	X	X	X	X	PP	X	X	VP
AT2G17770	FDP, BZIP27	X	AP	X	X	X	HP	PP	SP	TP	X
AT2G17950	WUS	X	X	X	FDMI	X	X	X	X	X	X
AT2G18880	VIL3, VEL2	X	X	X	X	X	X	X	X	X	VP
AT2G19425	MIR156G	X	AP	X	X	X	HP	X	SP	X	X
AT2G21660	GRP7, CCR2	GPAP	X	X	X	X	X	X	X	X	X
AT2G25095	MIR156A	X	AP	X	X	X	HP	X	SP	X	X
AT2G25930	ELF3, PYK20	X	X	X	X	X	X	X	X	X	X
AT2G27990	PNF, BLH8	X	X	X	FDMI	X	X	X	X	X	X
AT2G28056	MIR172A	X	AP	X	FDMI	X	HP	PP	SP	TP	X
AT2G30120	FLX	X	X	X	X	X	X	X	X	X	VP
AT2G33810	SPL3	X	AP	X	X	X	HP	PP	SP	X	X
AT2G33835	FES1	X	X	X	X	X	X	X	X	X	VP
AT2G35670	FIS2, FIE2	X	X	X	X	X	X	X	X	X	VP
AT2G40080	ELF4	X	X	CCP	X	X	X	PP	X	TP	X
AT2G43010	PIF4	X	X	CCP	X	X	HP	PP	X	TP	X
AT2G43410	FPA	GPAP	X	X	X	X	X	X	X	X	VP
AT2G46255	MIR159C	X	X	X	X	X	HP	X	X	X	X

Codigo	Nombre	General processes & autonomous pathway	Aging pathway	Circadian clock pathway	Flower development and meristem identity	Gene involve to Flowering	Hormonal pathway	Photoperiod pathway	Sugar pathway	Temperature pathway	Vernalization pathway
AT2G46790	PRR9	X	X	CCP	X	X	X	X	X	X	X
AT2G46830	CCA1	X	X	CCP	X	X	X	X	X	X	X
AT2G46970	PLI1	X	X	X	X	X	X	PP	X	X	X
AT2G47700	RFI2	X	X	X	X	X	X	PP	X	X	X
AT2G48160	HULK2	GPAP	X	X	X	X	X	X	X	X	X
AT3G01460	MBD9	GPAP	X	X	X	X	X	X	X	X	X
AT3G02885	GASA5	X	X	X	X	X	HP	X	X	X	X
AT3G11435	MIR172C	X	AP	X	FDMI	X	HP	PP	SP	TP	X
AT3G15270	SPL5	X	AP	X	X	X	HP	PP	SP	X	X
AT3G18550	BRC1, TCP18	X	X	X	FDMI	X	X	PP	X	X	X
AT3G19140	DNF	X	X	X	X	X	X	PP	X	X	X
AT3G22231	PCC1	X	X	CCP	FDMI	X	X	PP	X	X	X
AT3G22380	TIC	X	X	CCP	X	X	X	X	X	X	X
AT3G23060	AtBMI1C	GPAP	X	X	X	X	X	X	X	X	VP
AT3G23130	SUP, FLO10	X	X	X	FDMI	X	X	X	X	X	X
AT3G54500	LNK2	X	X	CCP	X	X	X	X	X	X	X
AT3G54990	SMZ	X	AP	X	X	X	HP	PP	X	TP	VP
AT3G57920	SPL15	X	AP	X	X	X	X	X	X	X	VP
AT3G59060	PIF5, PIL6	X	X	X	X	X	X	X	X	TP	X
AT3G63060	EDL3	X	X	X	X	X	HP	X	X	X	X
AT3G63070	HULK3	GPAP	X	X	X	X	X	X	X	X	X
AT4G00650	FRI	X	X	X	X	X	X	X	X	X	VP
AT4G02560	LD	GPAP	X	X	X	X	X	X	X	X	X
AT4G03090	AtINDX	GPAP	X	X	X	X	X	X	X	X	VP
AT4G09180	FBH2	X	X	X	X	X	X	PP	X	X	X
AT4G24415	MIR824	X	X	X	X	X	X	PP	X	X	X
AT4G25530	FWA, HDG6	GPAP	X	X	X	X	X	X	X	X	X
AT4G26150	GNL, CGA1	X	X	X	X	X	HP	X	X	X	X
AT4G27330	SPL, NZZ	X	X	X	FDMI	X	X	X	X	X	X
AT4G30972	MIR156B	X	AP	X	X	X	HP	X	SP	X	X
AT4G31877	MIR156C	X	AP	X	X	X	HP	X	SP	X	X
AT4G35900	FD	X	AP	X	FDMI	X	HP	PP	SP	TP	VP
AT4G38680	CSP2, GRP2	GPAP	X	X	X	X	X	X	X	X	X
AT5G04275	MIR172B	X	AP	X	FDMI	X	HP	PP	SP	TP	X
AT5G08230	HULK1	X	X	X	X	GIF	X	X	X	X	X
AT5G10140	FLC, AGL25, FLF	X	X	X	X	X	X	PP	X	X	VP
AT5G10944	FLC, AGL25, FLF	X	X	X	X	GIF	X	X	X	X	X
AT5G10945	MIR156D	X	AP	X	X	X	HP	X	SP	X	X
AT5G11530	EMF1	GPAP	X	X	FDMI	X	X	X	X	X	VP
AT5G11977	MIR156E	X	AP	X	X	X	HP	X	SP	X	X
AT5G14010	KNU	X	X	X	FDMI	X	X	X	X	X	X
AT5G15230	GASA4	X	X	X	X	X	HP	X	X	X	X
AT5G16320	FRL1, SUF8	X	X	X	X	X	X	X	X	X	VP
AT5G17690	TFL2, LHP1	GPAP	X	X	FDMI	X	X	PP	X	X	VP
AT5G23150	HUA2	GPAP	X	X	X	X	X	X	X	X	X
AT5G23730	RUP2, EFO2	X	X	X	X	X	X	PP	X	X	X
AT5G24470	PRR5	X	X	CCP	X	X	X	X	X	X	X
AT5G26147	MIR156F	X	AP	X	X	X	HP	X	SP	X	X
AT5G35770	SAP	X	X	X	FDMI	X	X	X	X	X	X
AT5G37770	CML24, TCH2	X	X	X	X	GIF	X	X	X	X	X
AT5G39660	CDF2	X	X	X	X	X	X	PP	X	X	X
AT5G42400	ATXR7, SDG25	GPAP	X	X	X	X	X	PP	X	X	VP
AT5G44190	GLK2	GPAP	X	X	X	X	X	X	X	X	X
AT5G48120	MMS19, MET18	X	X	X	X	GIF	X	X	X	X	X
AT5G48560	CIB2	X	X	X	X	X	X	PP	X	X	X
AT5G48890	LATE	X	X	X	X	X	X	PP	X	X	X
AT5G50080	ERF110	X	X	X	X	X	HP	X	X	X	X
AT5G52250	RUP1, EFO1	X	X	X	X	X	X	PP	X	X	X
AT5G55835	MIR156H	X	AP	X	X	X	HP	X	SP	X	X

Codigo	Nombre	General processes & autonomous pathway	Aging pathway	Circadian clock pathway	Flower development and meristem identity	Gene involve to Flowering	Hormonal pathway	Photoperiod pathway	Sugar pathway	Temperature pathway	Vernalization pathway
AT5G56860	GNC	X	X	X	X	X	HP	X	X	X	X
AT5G57380	VIN3	X	X	X	X	X	X	X	X	X	VP
AT5G61920	FLX4, FLL4	X	X	X	X	X	X	X	X	X	VP
AT5G62430	CDF1	X	X	X	X	X	X	PP	X	X	X
AT5G64170	LNK1	X	X	CCP	X	X	X	X	X	X	X
AT5G64390	HEN4	X	X	X	FDMI	X	X	X	X	X	X
AT5G65050	MAF2, AGL31	X	X	X	X	X	X	X	X	TP	VP
AT5G65060	MAF3, AGL70	X	X	X	X	X	X	X	X	X	VP
AT5G65070	MAF4, AGL69	X	X	X	X	X	X	X	X	X	VP
AT5G65080	MAF5, AGL68	X	X	X	X	X	X	X	X	X	VP