

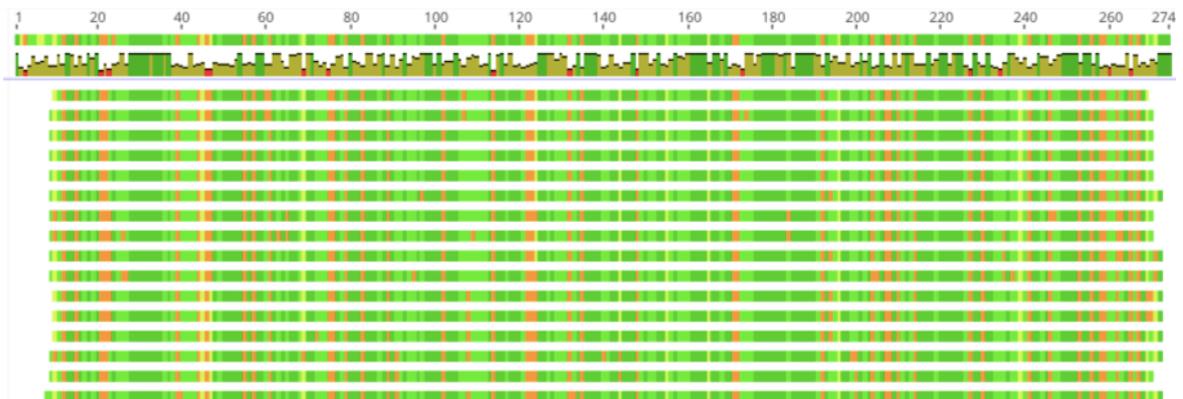


CENTRO INTERNACIONAL DE POSTGRADO

Máster en Biotecnología Aplicada a la Conservación y Gestión
Sostenible de Recursos Vegetales

TRABAJO FIN DE MÁSTER

ESTRIGOLACTONAS Y RAMIFICACIÓN EN *Pinus pinaster*



Pablo González Suárez
24 de julio de 2019

Máster en Biotecnología Aplicada a la Conservación y Gestión
Sostenible de Recursos Vegetales

TRABAJO FIN DE MÁSTER

ESTRIGOLACTONAS Y RAMIFICACIÓN EN *Pinus pinaster*

STRIGOLACTONES AND BRANCHING IN *Pinus pinaster*

Pablo González Suárez

TUTOR/ES

José Manuel Álvarez Díaz

Ricardo Javier Ordás Fernández

AGRADECIMIENTOS

Este trabajo ha sido posible gracias a la colaboración entre el Área de Fisiología Vegetal de la [Universidad de Oviedo](#) y el [Servicio Regional de Investigación y Desarrollo Agroalimentario del Principado de Asturias \(SERIDA\)](#), y está enmarcado dentro del proyecto INIA RTA 2017-00063-C04-04. Es el resultado del año que más he trabajado y aprendido, hasta ahora, en mi vida, y jamás hubiera sido posible sin el apoyo de mis tutores, el [Dr. Ricardo Ordás](#) y el [Dr. José Manuel Álvarez](#), a quien agradezco su paciencia y toda la ayuda que me ha ofrecido siempre que lo he necesitado.

Quiero aprovechar para agradecerle a [Lara García](#) por enseñarme gran parte de lo que sé sobre el trabajo en un laboratorio. A [Ana Álvarez](#) por ser lo mejor que me ha pasado este año y a [Victor Roces](#) por ser lo segundo. A todos los miembros del equipo [EPIPHYSAGE](#) por ser mis compañeros y amigos durante todo este año. A mi tutora de prácticas, la [Dr. Isabel Feito](#), le quiero agradecer una vez más el haberme acogido tan calurosamente en el SERIDA, y toda la atención que me ha dado durante estos meses, animándome siempre a ser crítico con lo que hago. Así como también a todo su equipo y especialmente a [Lucía Rodríguez](#), la mejor maestra de espectrometría de masas posible y, ahora, mi técnica de laboratorio favorita.

A nivel personal, me gustaría dar las gracias a mi familia por haberme apoyado incondicionalmente en el proceso que me ha traído hasta aquí, por haberlo hecho posible y no haber tirado nunca la toalla: no es fácil tener un hijo biólogo. A mis amigos, y especialmente a [Abraham](#), por haber aguantado mis interminables monólogos sobre estrigolactonas durante todo el año.

Finalmente, a la [Dr. Candela Cuesta](#) quiero decirle que, como ya sabe, los días sólo tienen veinticuatro horas. Por encontrar siempre entre ellas un rato para mí, por animarme a ser mejor científico y persona, y por acompañarme durante este año, gracias infinitas.

RESUMEN

La **ramificación** es un proceso clave del desarrollo vegetal que incide directamente sobre la arquitectura del tallo. En especies de gran interés ecológico y forestal, como *Pinus pinaster*, el número de ramas de la planta tiene una enorme influencia sobre la **calidad de la madera**. Por este motivo, comprender los mecanismos que regulan la ramificación vegetal y su base molecular es un paso clave a la hora de diseñar programas de mejora que aseguren un mejor aprovechamiento de nuestros recursos naturales. Gracias al estudio de especies modelo como *A. thaliana*, sabemos que la ramificación está controlada en gran medida por las **estrigolactonas**, un grupo de hormonas vegetales recientemente descubierto. Sin embargo, poco se sabe acerca de la función de estas fitohormonas en coníferas, así como de los procesos que llevan a su percepción y señalización. En este contexto, el objetivo principal de este trabajo es profundizar en el mecanismo de percepción de estrigolactonas en coníferas e identificar el gen homólogo a su receptor en *P. pinaster*. Todo ello permite sentar una base molecular de cara al estudio de la arquitectura del tallo en especies de interés forestal, un primer paso hacia el desarrollo de indicadores de madera de calidad.

ABSTRACT

In plants, **branching** is a key developmental process that has a direct impact on the shoot architecture. This is especially important in *Pinus pinaster*, a species with a great ecological and economic relevance, since the number of branches influences the **quality of its wood**. Due to this, understanding the mechanisms that control branching and their molecular basis is a fundamental step in order to design plant improvement programs and assure a better exploitation of our natural resources. Thanks to the study of model species such as *A. thaliana*, we now know that branching is highly controlled by **strigolactones**, a novel group of plant hormones. However, little is known about their role in conifers and the processes that lead to their perception and signaling. In this context, the main objective of this project is to deepen into the mechanisms that underlie strigolactone perception in conifers and to identify the homolog of the strigolactone receptor in *P. pinaster*. All of this will create a molecular basis for the study of shoot architecture in forest species, with the final goal of developing indicators of high quality wood.

ABREVIATURAS

- **ADNc.** ADN copia.
- **AM.** Arbuscular Mycorrhiza.
- **BLASTp.** Basic Local Alignment Search Tool (protein).
- **BLOSUM62.** Blocks of Amino Acid Substitution Matrix.
- **BRC1.** *BRANCHED1*.
- **CDS.** Coding Sequence, region codificante de un gen.
- **CK.** Citoquininas.
- **D14.** *DWARF14*.
- **D14-SCF^{MAX2}.** Complejo de interacción formado por D14 y el complejo multiproteico SCF con MAX2 como proteína F-box.
- **IAA.** Indole acetic acid.
- **IPA1.** *IDEAL PLANT ARCHITECTURE1*.
- **KAI2.** *KARRIKIN INSENSITIVE 2*.
- **KL.** Karrikin-Like, compuestos similares a karrikinas.
- **KR.** Karrikinas.
- **LB.** Luria Bertani, medio de cultivo.
- **MAFFT.** Multiple Alignment using Fast Fourier Transform.
- **MAX2.** *MORE AXILLARY GROWTH2*.
- **NJ.** Neighbor-Joining.
- **pb.** Pares de bases.
- **PCR.** Polymerase Chain Reaction.
- **PIN1.** *PIN-FORMED1*, auxin transporter.
- **PpD14.** *Pinus pinaster DWARF14*.
- **SCF.** Skp-Cullin-F-box, complejo ubiquitin-ligasa multiproteico E3.
- **SL.** Estrigolactonas.
- **SMAX1.** *SUPPRESSOR OF MAX2 1*.
- **SMXL.** *SMAX1-LIKE*.
- **SOC.** Super Optimal broth with Catabolite repression, medio de cultivo.
- **Ub.** Ubiquitina.
- **UPGMA.** Unpaired Pair Group Method with Arithmetic mean.

ÍNDICE DE CONTENIDOS

	PÁGINA
1. INTRODUCCIÓN	1
1.1. ESTADO DEL ARTE	1
1.1.1. Arquitectura vegetal - ¿Cómo se construye una planta?	1
1.1.2. ¿Cómo se regula la arquitectura del tallo?	1
1.1.3. ¿Qué son las estrigolactonas?	2
1.1.4. Percepción y señalización de estrigolactonas	3
1.1.5. La evolución de las estrigolactonas en el reino vegetal	4
1.2. PLANTEAMIENTO	5
1.3. OBJETIVOS	5
1.3.1. Objetivo principal	5
1.3.2. Objetivos específicos	5
2. MATERIAL Y MÉTODOS	6
2.1. BÚSQUEDA BIOINFORMÁTICA DE GENES CANDIDATOS PARA LA PERCEPCIÓN DE ESTRIGOLACTONAS EN CONÍFERAS	6
2.1.1. Obtención de secuencias candidatas	6
2.1.2. Filtrado de secuencias aminoácídicas candidatas	6
2.2. ESTUDIO FILOGENÉTICO DE LOS RECEPTORES DE ESTRIGOLACTONAS EN CONÍFERAS	7
2.2.1. Alineamiento de secuencias candidatas filtradas	7
2.2.2. Estimación filogenética	7
2.3. IDENTIFICACIÓN Y CLONACIÓN DE <i>PPD14</i>	8
2.3.1. Amplificación del gen D14 de <i>P. pinaster</i>	8
2.3.2. Clonación en el vector pJET	9
2.3.3. Secuenciación del vector pJET/ <i>PpiD14</i>	9
2.3.4. Clonación en el sistema Gateway®	10
3. RESULTADOS Y DISCUSIÓN	13
3.1. BÚSQUEDA BIOINFORMÁTICA DE GENES CANDIDATOS PARA LA PERCEPCIÓN DE ESTRIGOLACTONAS EN CONÍFERAS	13
3.2. ESTUDIO FILOGENÉTICO DE LOS RECEPTORES DE ESTRIGOLACTONAS EN CONÍFERAS	13
3.2.1. D14 y KAI2 se agrupan en clados taxón-específicos	16

3.2.2. La especificidad de los receptores D14 y KAI2 podría haber surgido con las gimnospermas	16
3.2.3. Las estrigolactonas podrían contribuir al estudio de la evolución vegetal	18
3.2.4. El receptor KAI2 se duplica en la base de las gimnospermas	19
3.3. IDENTIFICACIÓN Y CLONACIÓN DE <i>PpD14</i>	19
3.3.1. Obtención de la CDS del gen <i>PpD14</i>	19
3.3.2. Clonación del gen <i>PpD14</i> y construcción de un vector binario de sobreexpresión mediante el sistema Gateway®	21
4. BIBLIOGRAFÍA	24

ÍNDICE DE FIGURAS

	PÁGINA
Figura 1. Modelos propuestos para el control de la ramificación.	2
Figura 2. Estructura química de la estrigolactona 5-deoxystrigol.	3
Figura 3. Modelo aceptado actualmente para la percepción de SL.	3
Figura 4. Modelo propuesto actualmente para la percepción y señalización de estrigolactonas (SL) y karrikinas (KL) en gimnospermas en contraposición a angiospermas.	5
Figura 5. (a) Composición de la mezcla de reactivos utilizados para amplificar el gen candidato (<i>PpD14</i>), (b) Programa del termociclador empleado para la PCR.	8
Figura 6. Vector pJET1.2/blunt tras la reacción de ligación de la CDS del gen <i>PpD14</i> .	9
Figura 7. Esquema general de la estrategia de clonación llevada a cabo para introducir el gen de interés.	12
Figura 8. Alineamiento de las secuencias D14 y KAI2 de <i>A. thaliana</i> , <i>P. pinaster</i> , <i>T. baccata</i> y <i>W. mirabilis</i> .	14
Figura 9. Filograma del árbol más probable para los clados D14 y KAI2 en los principales grupos de gimnospermas, implementado mediante el algoritmo UPGMA.	15
Figura 10. Filograma del árbol más probable para los clados D14 y KAI2 en los principales grupos de gimnospermas, implementado mediante el algoritmo NJ.	17
Figura 11. Hipótesis más populares de la posición filogenética de los gnetales.	18
Figura 12. Representación gráfica de los principales dominios proteicos anotados en la secuencia <i>PpD14</i> .	20
Figura 13. Modelos 3D predichos a partir de las secuencias aminoacídicas de <i>PpD14</i> (<i>P. pinaster</i>) y <i>AtD14</i> (<i>A. thaliana</i>).	20
Figura 14. Vector de sobreexpresión generado a partir del plásmido pMDC32 y la secuencia <i>PpD14</i> tras la reacción LR.	22

ÍNDICE DE TABLAS

	PÁGINA
Tabla 1. Residuos aminoacídicos de la proteína D14 de <i>A. thaliana</i> esenciales para la percepción de estrigolactonas.	6
Tabla 2. Secuencia en sentido 5' a 3' de los cebadores utilizados para amplificar el gen DWARF14 de <i>P. pinaster</i> .	8
Tabla 3. Mezcla de reactivos empleados para la clonación de D14 en el vector pJET1.2/blunt.	9
Tabla 4. Secuencia en sentido 5' a 3' de los cebadores utilizados para la secuenciación del plásmido pJET1.2/blunt con la CDS de <i>PpD14</i> .	10
Tabla 5. Secuencia en sentido 5' a 3' de los cebadores con las etiquetas attB1 (F) y attB2 (R) utilizados para introducir <i>PpD14</i> en el sistema Gateway®.	10
Tabla 6. Número de secuencia candidatas KAI2/D14 de diferentes especies obtenidas para cada grupo de coníferas.	13

1. INTRODUCCIÓN

1.1. ESTADO DEL ARTE

1.1.1. ARQUITECTURA VEGETAL – ¿CÓMO SE CONSTRUYE UNA PLANTA?

Las plantas son **organismos sésiles**, incapaces de desplazarse para obtener nutrientes, defenderse de depredadores o hacer frente a condiciones adversas (Domagalska y Leyser, 2011). Sin embargo, esto no quiere decir que no posean estrategias para garantizar su supervivencia (Leyser, 2018). Una de las herramientas más poderosas que poseen es su **desarrollo post-embriionario**, que les permite adaptar su crecimiento al ambiente que las rodea. Cómo son las plantas capaces de integrar la información ambiental para tomar decisiones acerca de su propio desarrollo y cuáles son las señales que participan en este proceso son sólo algunas de las preguntas que trata de responder el estudio de la arquitectura vegetal (Walker and Bennett, 2018).

Los organismos vegetales experimentan una gran cantidad de procesos que repercuten en su morfología, como la organogénesis o la senescencia. Cuando hablamos de la **arquitectura del tallo**, algunos de los caracteres de mayor importancia son la altura de la planta y su patrón de ramificación. Y, de hecho, la **ramificación del tallo** es un proceso extraordinariamente complejo que condiciona el número, longitud, ángulo y posición de las ramas laterales que desarrolla una planta (Kebrom, 2017).

Pero, ¿cómo se produce exactamente esta ramificación? La mayoría de organismos vegetales presenta un tronco o eje principal que posee varias **yemas axilares**. Las yemas, cargadas de células meristemáticas, cuentan con toda la maquinaria molecular y celular necesaria para desarrollar una nueva rama, siempre y cuando se den las condiciones necesarias para liberarlas de su dormición, un proceso denominado **activación de yemas** (Barbier *et al.*, 2019). La salida de la dormición, que ocurre típicamente en las épocas más favorables del año, está regulada por señales tanto externas como endógenas, y tiene una enorme influencia sobre la estructura final del tallo.

A nivel más aplicado, la ramificación tiene un gran impacto sobre el rendimiento de los cultivos agrícolas y plantaciones forestales. En las especies de interés silvícola, incide directamente sobre la cantidad y calidad de la madera producida (Calderón *et al.*, 2016; Hevia, Álvarez-González y Majada, 2016). Todo ello hace que el estudio de la ramificación sea un proceso clave de cara a mejorar el rendimiento de las plantaciones forestales. Y en el caso concreto de los pinos, en los que la explotación se lleva a cabo años después de su establecimiento, entender los mecanismos que modulan el crecimiento vegetal es aún más crítico a la hora de diseñar programas de mejora (Jordy, 2004).

1.1.2. ¿CÓMO SE REGULA LA ARQUITECTURA DEL TALLO?

La ramificación está controlada por toda una serie de señales **endógenas** y **ambientales** que son percibidas y procesadas por la planta, dando lugar a su patrón de crecimiento. Esta regulación está coordinada tanto espacial como temporalmente por **fitohormonas**,

compuestos de transporte sistémico que permiten integrar los estímulos ambientales en el desarrollo vegetal (Walker y Bennett, 2017). Como tal, las fitohormonas son las principales encargadas de regular la activación de yemas axilares y la ramificación (Brewer *et al.*, 2015). A lo largo de los años se han propuesto dos posibles modelos para explicar su control de la arquitectura del tallo.

El **modelo de la canalización de auxina** (Figura 1a) se basa en la idea de que una yema debe exportar auxina para poder desarrollarse (Leyser, 2018). La auxina, que es sintetizada principalmente en las yemas, se transporta basipétalmente hacia la raíz. Puesto que la capacidad de una yema para exportar la auxina que produce depende de fuerzas relativas fuente-sumidero, existe una competencia entre diferentes yemas axilares para exportar su auxina. Como resultado de este fenómeno, sólo algunas de las yemas terminan desarrollándose y formando ramas.

El **modelo del segundo mensajero** (Figura 1b) propone que las auxinas, que de por sí solas no pueden entrar en las yemas, actúan a través de otras fitohormonas como las citoquininas y las recientemente caracterizadas **estrigolactonas** (Kebrom, 2017).

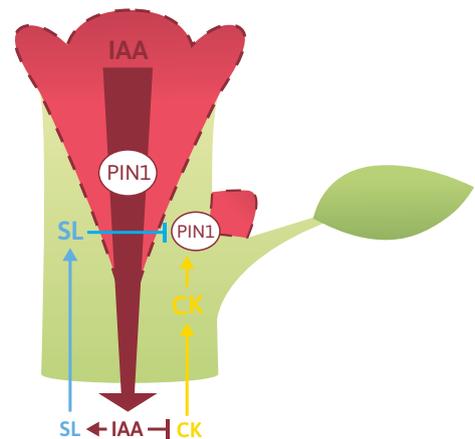
En la actualidad aún existen muchas dudas en relación a este mecanismo de regulación, aunque todo parece indicar que gran parte de la regulación es independiente de las auxinas, como se ha observado en trabajos con guisante (Brewer *et al.*, 2015). En especies perennes no modelo, como las coníferas, los misterios que rodean el control de la ramificación aún son más notables, dado que estos organismos experimentan unas etapas de dormición adicionales ausentes en plantas anuales (Barbier *et al.*, 2019).

1.1.3. ¿QUÉ SON LAS ESTRIGOLACTONAS?

Las **estrigolactonas** (SL) son fitohormonas con estructura heterocíclica derivadas del caroteno (Végh *et al.*, 2017). Pertenecen al diverso grupo de los **apocarotenoides**, junto con los isoprenoides, el ABA y otros compuestos volátiles (Lopez-Obando *et al.*, 2015). Su estructura química consta de dos regiones diferenciadas (Figura 2):

- Un núcleo de **lactona tricíclica** (ABC), con residuos variables según el compuesto.
- Un **anillo D**, unido al núcleo tricíclico mediante un enlace enol-éter. Se trata de una región invariable de la molécula que se considera indispensable para los efectos de las SL sobre la ramificación.

(a) Modelo de la canalización de auxina



(b) Modelo del segundo mensajero

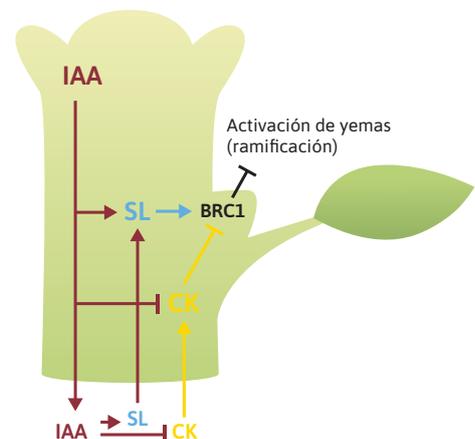


Figura 1. Modelos propuestos para el control de la ramificación. Modificado de Waters *et al.*, 2017. IAA: Ácido indolacético (auxinas). CK: Citoquininas. SL: Estrigolactonas. PIN1: PIN-FORMED1. BRC1: BRANCHED1.

Las estrigolactonas se definen como un extenso grupo de fitohormonas capaces de regular gran cantidad de aspectos relacionados con el desarrollo vegetal, como la ramificación del tallo, la estatura de la planta y el crecimiento radicular, entre otros (Bennett y Leyser, 2014; Marzec y Brewer, 2019). En lo que respecta a la arquitectura del tallo, actúan **inhibiendo la ramificación**, de forma análoga a las auxinas y antagónica a las citoquininas (Lopez-Obando *et al.*, 2015). Aunque aún no existe un consenso acerca de su mecanismo exacto de actuación, las SL se han integrado en los dos modelos propuestos hasta la fecha, ya sea regulando el transporte de auxinas (Figura 1a) o actuando como segundos mensajeros (Figura 1b).

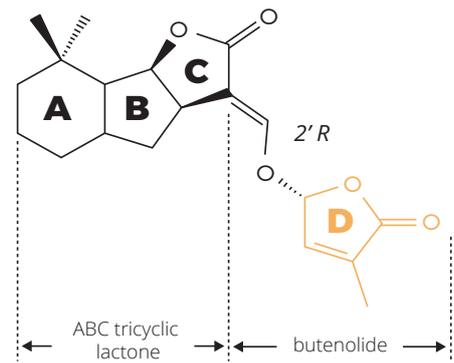


Figura 2. Estructura química de la estrigolactona 5-deoxystrigol.

En años recientes, trabajos con la conífera *Pinus pinaster* han confirmado que las estrigolactonas juegan un papel importante en el desarrollo de gimnospermas, concretamente en su patrón de ramificación (Meijón *et al.*, 2016). Esto ha puesto en alza el valor de estos compuestos como indicadores de calidad de la madera y ha remarcado la necesidad de investigar sus mecanismos de señalización en pino.

1.1.4. PERCEPCIÓN Y SEÑALIZACIÓN DE ESTRIGOLACTONAS

La percepción de estrigolactonas está basada en la **ubiquitinación y degradación de represores**, como se puede observar en la Figura 3. Es un proceso muy similar al que se observa en otras fitohormonas con la particularidad de que el receptor de SL, DWARF14, presenta una función dual muy poco convencional: por una parte la percepción y, por otra, la hidrólisis de su ligando (Marzec, 2016; Yao *et al.*, 2018; Marzec y Brewer, 2019).

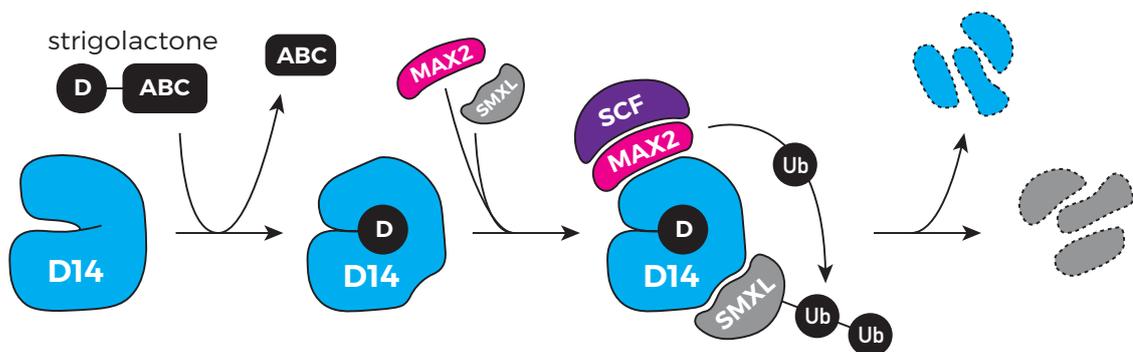


Figura 3. Modelo aceptado actualmente para la percepción de SL. Modificado de Waters *et al.*, 2017. D14: DWARF14, MAX2: MORE AXILLARY GROWTH2. SMXL: SMAX1-LIKE family. SCF: Skp1-Cullin1-F-box Ubiquitin ligase complex.

El receptor **DWARF14** (D14) codifica una α/β -hidrolasa que se une a la estrigolactona y posteriormente la hidroliza, separando el núcleo ABC del anillo D, el cual queda covalentemente unido al receptor (Marzec y Brewer, 2019). La unión de la hormona a D14 provoca un cambio conformacional en la proteína que posibilita su interacción con **MORE AXILLARY GROWTH2** (MAX2), el cual forma parte de un complejo de ubiquitinación de tipo SCF (Skp1-Cullin1-F-box). La interacción entre la estrigolactona, D14 y MAX2 pone en

funcionamiento el complejo de ubiquitinación, que marca para la proteólisis a represores de la familia **SMAX1-LIKE** (SMXL) (Bennett y Leyser, 2014). Como consecuencia de todo este proceso, tanto el receptor (D14) como el represor (SMXL) son degradados, y se pone en marcha la maquinaria molecular que provocará la **inhibición de la ramificación**. Aún se desconocen los procesos exactos que tienen lugar *downstream*, es decir, tras la degradación del represor, aunque se ha propuesto que podrían estar mediados por **factores de transcripción** como IDEAL PLANT ARCHITECTURE1 (IPA1) en monocotiledóneas o BRANCHED1 (BRC1) en dicotiledóneas (Yao *et al.*, 2018). Otra posibilidad sería que los efectos en la planta se debiesen a cambios en el **transporte de auxinas**, vía la polarización de transportadores de auxina PIN1 (Bennett y Leyser, 2014).

1.1.5. LA EVOLUCIÓN DE LAS ESTRIGOLACTONAS EN EL REINO VEGETAL

Las estrigolactonas son compuestos evolutivamente muy antiguos (Lopez-Obando *et al.*, 2018). De hecho, trabajos con mutantes muy ramificados en distintas especies parecen indicar que las rutas de señalización de SL están **muy conservadas** entre la mayoría de organismos vegetales (Yao, Li y Xie, 2018), al menos en lo que se refiere a la percepción de la molécula por parte del complejo D14-SCF^{MAX2}. Sin embargo, la situación cambia si centramos la atención en especies más basales. En el briófito *Physcomitrella patens*, en el que ya se ha descrito la síntesis de SL, el mecanismo de percepción no está del todo claro. La mayoría de autores coincide en que la ruta de señalización convencional de SL, mediada por D14 (Figura 3), parece ser una innovación relativamente reciente propia de plantas con semilla (Walker y Bennett, 2017).

Por otra parte, es importante tener en cuenta que la ruta de señalización de las SL está íntimamente relacionada con la de otros compuestos químicamente similares pero con funciones fisiológicas distintas, las **karrikinas** (KR) (Waters *et al.*, 2012). Las únicas diferencias residen en el receptor implicado, que en este caso es **KARRIKIN INSENSITIVE 2** (KAI2), y el represor degradado, en este caso **SUPPRESSOR OF MAX2 1** (SMAX1). KR y SL tienen un mecanismo de percepción tan

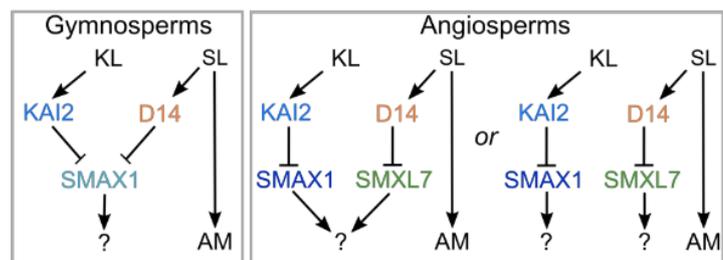


Figura 4. Modelo propuesto actualmente para la percepción y señalización de estrigolactonas (SL) y karrikinas y compuestos similares (KL) en gimnospermas en contraposición a angiospermas. Extraído de Walker and Bennett, 2017. KL: Karrikinas o compuestos similares. SL: Estrigolactonas. KAI2: KARRIKIN INSENSITIVE 2. D14: DWARF14. SMAX1: SUPRESOR OF MAX2 1. SMXL7: SMAX1-LIKE 7. AM: Arbuscular Mycorrhiza.

similar que a menudo se ha propuesto que ambas podrían compartir un **mismo origen evolutivo** (Bennett y Leyser, 2014). Ambas rutas podrían haber surgido a partir de una única ancestral que, a lo largo de la evolución, se habría duplicado dando lugar a la ruta D14-dependiente (SL) y KAI2-dependiente (KR) (Végh *et al.*, 2017).

La mayoría del conocimiento que se posee actualmente sobre la percepción de SL deriva del estudio de especies modelo angiospermas como *Arabidopsis* (*Arabidopsis thaliana*) o arroz (*Oryza sativa*). Por ello, aún no se ha determinado con exactitud el momento evolutivo en que ambas rutas se diferenciaron (Bythell-Douglas *et al.*, 2017) así como tampoco la

base molecular de la percepción de SL y KR en gimnospermas. A causa de la íntima relación entre ambos tipos de molécula y alta similitud entre sus receptores (D14 y KAI2, respectivamente), el estudio de la señalización de estrigolactonas en gimnospermas pasa inevitablemente por el estudio de la percepción de las karrikinas. Y, de hecho, todo parece indicar a que, en contraposición a lo que ocurre en angiospermas, donde SL y KR provocan la degradación de represores distintos; en gimnospermas ambas moléculas podrían compartir la misma diana, presumiblemente el represor SMAX1 (Walker y Bennett, 2017) (Figura 4).

1.2. PLANTEAMIENTO

Teniendo en cuenta todo lo comentado anteriormente, este Trabajo de Final de Máster se plantea como una primera aproximación al estudio de los mecanismos de percepción de SL en coníferas, con especial atención al receptor DWARF14 de la especie *Pinus pinaster* (*PpD14*). Identificar y clonar la secuencia candidata a D14 en pino podría suponer el primer paso de cara a validar la función de dicho gen utilizando a la especie modelo *Arabidopsis thaliana* como sistema heterólogo, mediante la complementación de mutantes *d14*.

Todo ello permitiría, en último término, identificar el posible receptor de estrigolactonas en coníferas, un hito indispensable para obtener marcadores de calidad de la madera, así como sentar la base de posibles programas de mejora genética en el panorama forestal.

1.3. OBJETIVOS

1.3.1. OBJETIVO PRINCIPAL

Se plantea como objetivo principal de este trabajo **identificar y caracterizar genes implicados en la percepción de estrigolactonas en coníferas**, especialmente en la especie de incuestionable importancia forestal *Pinus pinaster*.

1.3.2. OBJETIVOS ESPECÍFICOS

1. **Búsqueda bioinformática de genes candidatos** para la señalización de estrigolactonas y karrikinas en coníferas.
2. **Estudio filogenético** de los receptores de estrigolactonas y karrikinas en coníferas.
3. **Identificación y clonación del gen DWARF14** de *P. pinaster*.

2. MATERIAL Y MÉTODOS

2.1. BÚSQUEDA BIOINFORMÁTICA DE GENES CANDIDATOS PARA LA PERCEPCIÓN DE ESTRIGOLACTONAS EN CONÍFERAS

2.1.1. OBTENCIÓN DE SECUENCIAS CANDIDATAS

Con el objetivo de identificar posibles homólogos para los genes responsables de la percepción de estrigolactonas en gimnospermas, se utilizaron varias bases de datos públicas: Gymno PLAZA 3.0 (Proost et al., 2015), SustainPineDB (Fernández-Pozo et al., 2011) y 1,000 Plants (1KP) Project (Matasci et al., 2014). En los tres casos, las secuencias candidatas se obtuvieron mediante el algoritmo **BLASTp** a partir de la secuencia aminoacídica ya identificada del gen D14 de *Arabidopsis thaliana*, ya a diferencia de la secuencia nucleotídica las proteínas permiten una mayor conservación entre especies relativamente alejadas.

2.1.2. FILTRADO DE SECUENCIAS AMINOACÍDICAS CANDIDATAS

El filtrado de las secuencias se realizó teniendo en cuenta tres criterios principalmente. En primer lugar, sólo se tuvieron en cuenta los resultados con una **identidad superior al 50%** a nivel de aminoácidos. En segundo lugar, se consideró una posible homología siempre que el **E-valor** fuese inferior a 10^{-30} .

Por otra parte, en los últimos años se ha avanzado mucho en la caracterización de las interacciones que tienen lugar entre las estrigolactonas y las diferentes proteínas implicadas en su percepción. Es bien sabido que la actividad catalítica de las proteínas α/β -hidrolasa D14 y KAI2 requiere una **tríada catalítica Ser-His-Asp** (S-H-D) muy conservada (Waters et al., 2015; Yao et al., 2018). Además, el trabajo de Yao et al. (2016) identifica varios **aminoácidos esenciales**, necesarios para la interacción de D14 con la proteína MAX2 y, por tanto, la señalización de estrigolactonas. Como tal, se han seleccionado sólo aquellas secuencias candidatas que mantienen los aminoácidos críticos en las posiciones descritas en la **Tabla 1**, dado que la interacción con MAX2 es imprescindible tanto para la señalización mediada por D14 como la mediada por MAX2.

Tabla 1. Residuos aminoacídicos de la proteína D14 de *A. thaliana* esenciales para la percepción de estrigolactonas.

Residuo	Categoría	Función
S97	Tríada catalítica	Actividad enzimática D14
D218	Tríada catalítica	Actividad enzimática D14
H247	Tríada catalítica	Actividad enzimática D14
E174	Esencial	Contacto D14-MAX2
R177	Esencial	Contacto D14-MAX2
P161	Esencial	Contacto D14-MAX2
F180	Esencial	Contacto D14-MAX2
G154	Esencial	Interacción D14-MAX2
A160	Esencial	Interacción D14-MAX2
V164	Esencial	Interacción D14-MAX2

Todas las secuencias candidatas fueron procesadas manualmente utilizando el software **Geneious 11.1.5** (<http://www.geneious.com>).

Las proteínas ya caracterizadas D14 y KAI2 de la especie modelo *A. thaliana*, con los residuos clave anotados, se añadieron al conjunto de secuencias como referencia y se utilizaron como grupo externo para la mayoría de análisis filogenéticos posteriores.

2.2. ESTUDIO FILOGENÉTICO DE LOS RECEPTORES DE ESTRIGOLACTONAS EN CONÍFERAS

2.2.1. ALINEAMIENTO DE SECUENCIAS CANDIDATAS FILTRADAS

Tras el filtrado de secuencias aminoacídicas candidatas, éstas se alinearon utilizando el algoritmo **MAFFT** implementado en Geneious (Katoh *et al.*, 2002; Katoh and Standley, 2013), con la matriz de sustitución BLOSUM62. Para analizar detalladamente el grado de conservación de los distintos residuos aminoacídicos, algunas de las proteínas también se alinearon utilizando T-Coffee, un paquete en línea cuya herramienta M-Coffee permite combinar el resultado de algunos de los algoritmos de alineamiento más comunes (Mafft, ClustalW, etc.) (Notredame, Higgins and Heringa, 2000). Los resultados de los alineamientos se visualizaron mediante la herramienta en línea **MView** (<https://www.ebi.ac.uk/Tools/msa/mview/>) (Madeira *et al.*, 2019).

En algunos casos puntuales se recurrió a alineamientos por parejas para comprobar la similitud e identidad entre dos secuencias determinadas. Para ello se utilizó el algoritmo Needleman-Wunsch (matriz de sustitución EBLOSUM 62), implementado en la herramienta en línea **EMBOSS Needle** (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) (Madeira *et al.*, 2019).

2.2.2. ESTIMACIÓN FILOGENÉTICA

A continuación, el alineamiento obtenido se utilizó como *input* para generar un árbol filogenético usando dos algoritmos de estimación de filogenias distintos:

- a. **UPGMA** (Unweighted Pair Group Method with Arithmetic mean), que asume tasas de evolución idénticas en los distintos organismos analizados.
- b. **NJ** (Neighbor-Joining), que sí permite distintas tasas de evolución entre organismos a la hora de estimar la filogenia.

La fiabilidad de los árboles generados se evaluó mediante la técnica iterativa *bootstrap* con 500 réplicas. Finalmente, ambos árboles filogenéticos se visualizaron mediante la herramienta en línea **TreeDyn 198.3** (Chevenet *et al.*, 2006; Dereeper *et al.*, 2008, 2010).

2.3. IDENTIFICACIÓN Y CLONACIÓN DE *PpD14*

2.3.1. AMPLIFICACIÓN DEL GEN D14 DE *P. PINASTER*

Tal y como se describe en el apartado 2.1, la secuencia candidata para el homólogo de D14 en *P. pinaster* se obtuvo utilizando las bases de datos Gymno PLAZA 3.0 (Proost et al., 2015) y SustainPineDB (Fernández-Pozo et al., 2011). Una vez confirmado a través de diferentes aproximaciones bioinformáticas que dicha secuencia aminoacídica compartía una gran cantidad de características con el receptor caracterizado de *A. thaliana* (*AtD14*), se obtuvo la secuencia nucleotídica correspondiente (*PpiD14*) de la misma base de datos.

A continuación, se diseñaron cebadores específicos para la secuencia codificante (CDS) del gen utilizando el software **Geneious 11.1.5** (<http://www.geneious.com>) (Tabla 2).

Tabla 2. Secuencia en sentido 5' a 3' de los cebadores utilizados para amplificar el gen DWARF14 de *P. pinaster*.

Cebador	Sentido (<i>Forward</i> , F)	Anti-sentido (<i>Reverse</i> , R)
Secuencia	ATG GGT GTT GGC TCT GGA C	TCA CTG CAA CAC GCG TTT C

Para la amplificación, se empleó ADNc obtenido a partir de yemas apicales de plántulas como molde para la reacción en cadena de la polimerasa (PCR), utilizando la polimerasa de alta fidelidad Phusion (**Thermo Fisher Scientific, #E3540**). La mezcla de reactivos para la PCR y el protocolo del termociclador (Figura 5) se establecieron siguiendo las instrucciones del fabricante.

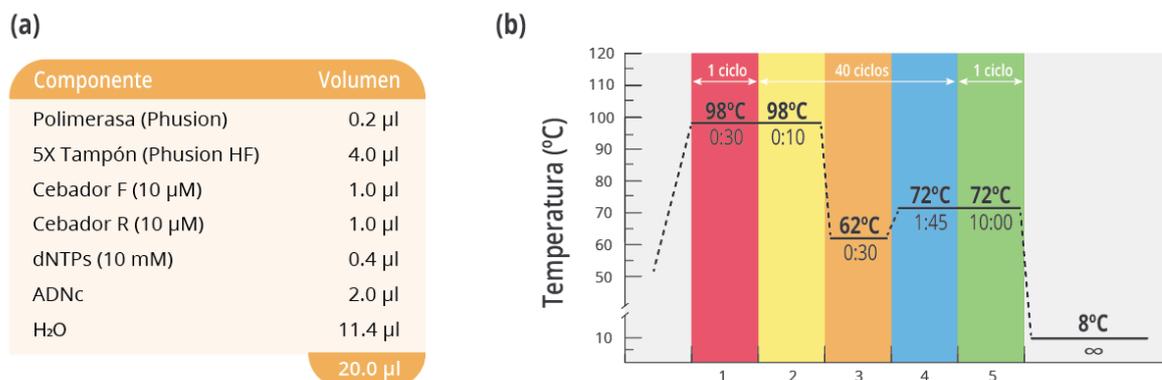


Figura 5. (a) composición de la mezcla de reactivos utilizados para amplificar el gen candidato (*PpD14*), (b) programa del termociclador empleado para la PCR.

El resultado de la PCR se comprobó mediante electroforesis en gel de agarosa. La banda con el tamaño correspondiente al amplicón deseado (*PpD14*), de 804 pb, se recortó con bisturí, se purificó con el kit comercial **GeneMATRIX Basic DNA Purification Kit** (EURx Molecular Biology Products, #E3545-01) y se cuantificó mediante espectrofotometría con el equipo **Pico100 Picodrop Spectrophotometer** (Picodrop, Ref. PICOPET).

2.3.2. CLONACIÓN DE *PpD14* EN EL VECTOR pJET

Una vez amplificada y purificada la CDS del gen D14 de *P. pinaster*, se procedió a clonar dicha secuencia en *E. coli* para contar con una mayor producción del amplicón, así como para secuenciarlo y comprobar que la PCR había tenido lugar sin ningún error.

Para la clonación, se escogió el kit comercial **CloneJET PCR Cloning Kit** (Thermo Fisher Scientific, #K1232) por su alta eficiencia y compatibilidad con las cepas de *E. coli* competentes disponibles en el laboratorio. Concretamente, el gen *PpD14* se clonó en el vector **pJET1.2/blunt** utilizando la mezcla de reactivos indicada en la **Tabla 3**. La cantidad de producto de PCR a añadir se determinó en base a las recomendaciones del fabricante. El resultado de la reacción fue un vector pJET1.2/blunt al que se había ligado la CDS del gen de interés, *PpD14* (**Figura 6**).

El siguiente paso consistió en transferir dicho plásmido a células competentes de *E. coli*, siguiendo un protocolo de transformación mediante choque térmico de la siguiente forma:

1. **Previa incubación en hielo** de 50 μ l de *E. coli* competentes (4°C)
2. **Adición del plásmido** (pJET1.2/blunt con el gen *PpD14*)
3. **Incubación en hielo** (4°C, 30 min)
4. **Choque térmico** (42°C, 45 seg)
5. **Incubación en hielo** (4°C, 2 min)
6. **Adición de 500 μ l de medio enriquecido SOC**
7. **Incubación en agitación** (37°C, 60 min, 2 00 rpm)

Finalmente, las células ya transformadas se plaquearon en medio nutritivo Luria Bertani (LB) que contenía el antibiótico de selección propio del vector pJET1.2/blunt, **ampicilina**, a una concentración de 100 μ g/ml. Las placas se incubaron durante 24 horas a 37°C.

2.3.3. SECUENCIACIÓN DEL VECTOR pJET1.2/PpD14

Tras la incubación a 37°C, se seleccionaron algunas de las colonias de *E. coli* crecidas en la placa con ampicilina, ya que a priori serían aquellas que habrían sido transformadas, incorporando el vector **pJET1.2/PpD14** y, con él, la resistencia a ampicilina. Una parte de cada colonia se utilizó para realizar una **PCR de colonia**, utilizando los cebadores de secuenciación incluidos en el propio kit CloneJET PCR Cloning Kit (**Tabla 4**). Esta PCR de colonia emplea la colonia seleccionada como molde para la amplificación y permite, de esta forma, determinar si las bacterias han incorporado el plásmido de interés.

Tabla 3. Mezcla de reactivos empleados para la clonación de *PpD14* en el vector pJET1.2/blunt.

Componente	Volumen
Vector pJET1.2/blunt	1.0 μ l
Tampón reacción 2X	10.0 μ l
Ligasa de ADN T4	1.0 μ l
Producto PCR	8.0 μ l
	20.0 μl

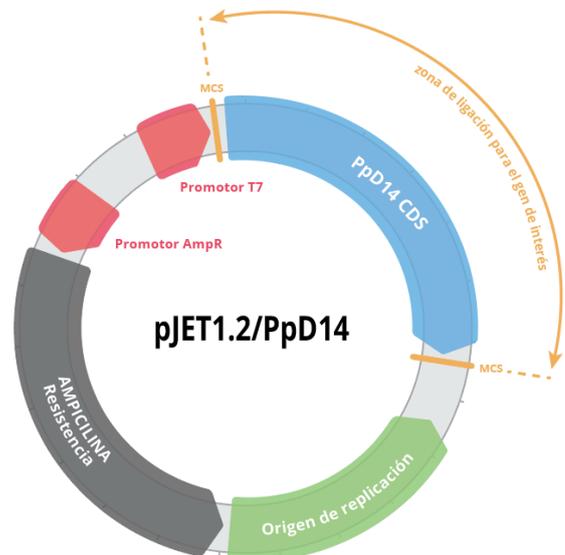


Figura 6. Vector pJET1.2/PpD14 tras la reacción de ligación del vector pJET1.2/blunt y la CDS del gen *PpD14*.

Tabla 4. Secuencia en sentido 5' a 3' de los cebadores utilizados para la secuenciación del plásmido pJET1.2/blunt con la CDS de *PpD14*.

Cebador	Sentido (<i>Forward</i> , F)	Anti-sentido (<i>Reverse</i> , F)
Secuencia	CGA CTC ACT ATA GGG AGA GCG GC	AAG AAC ATC GAT TTT CCA TGG CAG

Una vez finalizada la PCR, se verificó en gel de agarosa que a partir de las colonias seleccionadas se amplificaba una banda de un tamaño similar al esperado en caso de haber incorporado el plásmido con el gen *PpD14*, 804 pb de la CDS más los primers propios del vector. Para todas aquellas muestras positivas, se transfirió la mitad restante de la colonia a 5 ml de medio LB líquido con ampicilina (100 µg/ml) y se dejó crecer a 37°C.

Transcurridas aproximadamente 24 horas, se purificó el vector pJET1.2/PpiD14 del cultivo de *E. coli* transformadas utilizando el kit comercial **GeneMATRIX Plasmid Miniprep DNA Purification Kit** (EURx Molecular Biology Products, #E3500-02), basado en la utilización de columnas de sílice. Para eluir el vector purificado se usó un volumen final de 50 µl de agua. Al finalizar el proceso se cuantificó la concentración de plásmido y una parte de la muestra se secuenció utilizando el servicio externo Stab Vida (Portugal) con ayuda de los cebadores de secuenciación incluidos en el propio kit CloneJET PCR Cloning Kit (**Tabla 4**).

2.3.4. CLONACIÓN DE LA CDS DEL GEN *PPD14* EN EL SISTEMA GATEWAY

Una vez comprobado, mediante secuenciación del vector pJET1.2/PpiD14, que la CDS de *PpD14* carecía de errores derivados de la amplificación, se procedió a la construcción del vector binario mediante la **tecnología Gateway** (Thermo Fisher Scientific).

El sistema de clonación Gateway se basa en la recombinación sitio-específica de los denominados sitios att del fago lambda. Por consiguiente, el primer paso para clonar el gen de interés *PpD14* fue flanquearlo de las secuencias attB homólogas (attB1 y attB2) a los sitios attP del vector **pDONOR221** (attP1 y attP2) (Thermo Fisher Scientific, #12536017) para generar un vector de entrada. Esto se llevó a cabo mediante PCR utilizando una nueva pareja de cebadores que contenían las etiquetas attB1 y attB2 (**Tabla 5**).

Tabla 5. Secuencia en sentido 5' a 3' de los cebadores con las etiquetas attB1 (F) y attB2 (R) utilizadas para introducir *PpD14* en el sistema Gateway.

Cebador	Secuencia
<i>PpD14</i> -attB1 (F)	GGG GAC AAG TTT GTA CAA AAA AGC AGG CTG C ATG GGT GTT GGC TCT GGA C
<i>PpD14</i> -attB2 (R)	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTA TCA CTG CAA CAC GCG TTT C

Tras comprobar el resultado de la PCR en gel de agarosa, la banda deseada se recortó, purificó y cuantificó utilizando la metodología ya descrita previamente (apartado 2.4.1). Con la secuencia de *PpD14* flanqueada por los sitios att necesarios, se procedió a realizar la primera reacción de la tecnología Gateway, denominada **reacción BP**. Para ello se usó el kit Gateway® BP Clonase™ II Enzyme Mix (Thermo Fisher Scientific, #11789-020) siguiendo las instrucciones del fabricante. En esta reacción, la recombinación entre los sitios attB del gen de interés *PpD14* y los sitios attP del vector pDONOR221 da lugar a dos productos:

- Por una parte, un **derivado tóxico** que contiene la secuencia *ccdB*, que se desprende del vector pDONOR221.
- Por otra parte, el **vector de entrada**, **pDONOR221/PpD14**, que consta de los elementos propios del pDONOR221 (excepto la secuencia *ccdB*) más el gen *PpD14*, ahora flanqueado por los sitios attL1 y attL2.

Se utilizó este vector de entrada para transformar células de *E. coli* competentes mediante choque térmico, tal y como se ha descrito previamente (apartado 2.4.2). Nuevamente, se plaquearon las células transformadas en medio LB con el agente selectivo en bacteria, en este caso **kanamicina** a una concentración de 50 µg/ml, puesto que el vector pDONOR221 incluye un gen de resistencia a kanamicina. Las colonias positivas se comprobaron mediante PCR de colonia y se secuenciaron utilizando un servicio externo, tal y como se explica en el apartado 2.3.3).

Una vez verificado el resultado de este primer paso, se procedió a llevar a cabo la segunda reacción del sistema Gateway, la **reacción LR**, usando el kit Gateway® LR Clonase™ II Enzyme Mix (Thermo Fisher Scientific, #11791-020). En este caso, la recombinación tiene lugar entre los sitios attL que flanquean al gen de interés *PpD14* en el vector de entrada y los sitios attR del vector de destino pMDC32, y la reacción da lugar a:

- Un nuevo **derivado tóxico** con la secuencia *ccdB*, que se desprende del vector pMDC32.
- El **vector de expresión**, **pMDC32/PpD14**, que contiene: un promotor doble 35S, la CDS del gen de interés (*PpD14*) y genes de resistencia a higromicina (planta) y kanamicina (bacteria).

Como paso final del proceso de clonación, se utilizó el vector de expresión generado (pMDC32/PpD14) para transformar *E. coli* competentes siguiendo los mismos pasos que se habían llevado a cabo tras la primera reacción (BP) para verificar la construcción mediante PCR de colonia y secuenciación. Un esquema completo de toda la estrategia de clonación realizada puede consultarse en la **Figura 7**.

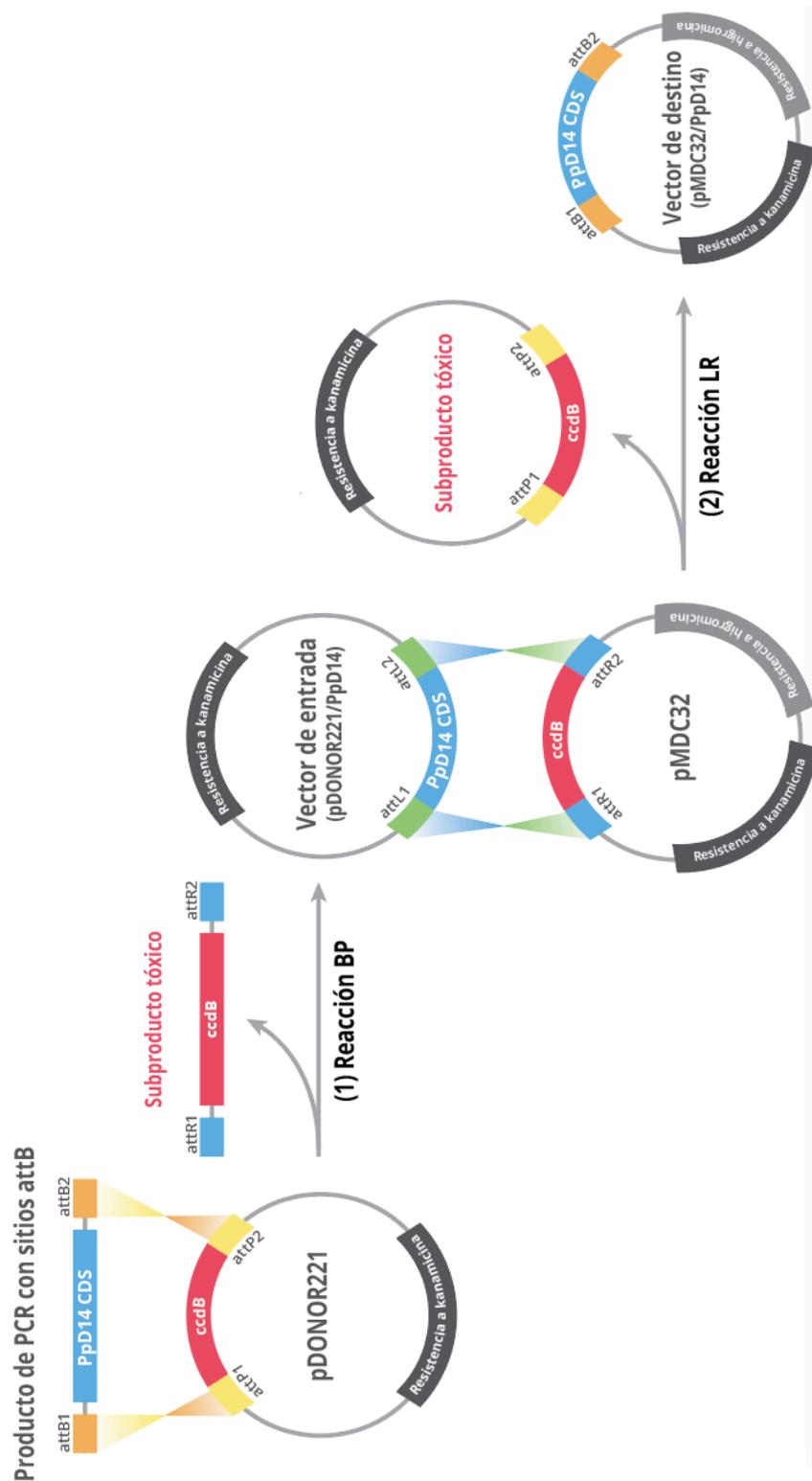


Figura 7. Esquema general de la estrategia de clonación llevada a cabo para introducir el gen de interés, *PpD14*, en el sistema Gateway (pDONOR 221) y generar un vector de sobre-expresión (pMDC32) apto para la complementación de mutantes de *Arabidopsis thaliana*.

3. RESULTADOS Y DISCUSIÓN

3.1. IDENTIFICACIÓN DE GENES HOMÓLOGOS PARA LA PERCEPCIÓN DE ESTRIGOLACTONAS EN CONÍFERAS

Tras la búsqueda en las bases de datos públicas mencionadas y el posterior filtrado de secuencias aminoacídicas candidatas, se terminaron seleccionando un total de **73 secuencias** distintas de **37 especies** de gimnospermas diferentes, con representantes de los principales grupos de coníferas (**Tabla 6**). Todas las secuencias obtenidas poseen la **tríada catalítica** Ser-His-Asp necesaria para la actividad de los receptores KAI2 y D14, así como los **aminoácidos descritos como críticos para la interacción** entre KAI2/D14 y la proteína MAX2.

Un alineamiento preliminar de las secuencias con el algoritmo CluscalW (BLOSUM) revela una similitud elevada entre todas las secuencias obtenidas, con una identidad por parejas promedio del 70.1% y un porcentaje de 25.8% sitios completamente idénticos. Por otra parte, la matriz de distancias por pares que resulta de comparar las secuencias dos a dos permite comprobar que el porcentaje de identidad (aminoácidos idénticos en ambas secuencias) más bajo es del 51%.

En conjunto, estos resultados demuestran el **alto grado de conservación** que existe en la familia proteica de α/β -hidrolasas D14/KAI2, al menos a nivel de estructura primaria. Como se observa, las secuencias tienden a mantener residuos idénticos, especialmente en las posiciones correspondientes a la tríada catalítica Ser-His-Asp y el centro de interacción con la proteína MAX2, necesaria para la percepción y señalización de estrigolactonas y compuestos relacionados con la ruta de las karrikinas (**Figura 8**).

3.2. ESTUDIO FILOGENÉTICO DE LOS RECEPTORES DE ESTRIGOLACTONAS EN CONÍFERAS

El análisis de las secuencias obtenidas y filtradas del apartado 3.1 mediante algoritmos de estimación de filogenias distintos permitió obtener dos estimaciones relativamente similares que nos permiten arrojar algo de luz sobre los procesos que han podido dirigir la evolución de la percepción de estrigolactonas en gimnospermas.

Tabla 6. Número de secuencias candidatas KAI2/D14 de diferentes especies obtenidas para cada grupo de coníferas.

Taxón	Clado KAI2	Clado D14
Araucariaceae	3	3
Cephalotaxaceae	2	2
Cupressaceae	7	7
Cycadaceae	5	1
Ginkgoaceae	2	1
Gnetaceae	2	
Pinaceae	10	10
Podocarpaceae	4	4
Sciadopityaceae	1	1
Taxaceae	4	4

		COV	PID		
1	<i>Arabidopsis thaliana</i>	D14	100.0%	100.0%	1 [M-SQHNI [●] EALNVRVVG [●] TGDRILFLAHGFGTDQSAWHLILPYFTQNYRVV : 50
2	<i>Arabidopsis thaliana</i>	KAI2	99.3%	50.7%	---MGVVEEAHN [●] VKVI [●] GSGEATIVLGHGFGTDQSVWKHLVPHLVDDYRVV
3	<i>Pinus pinaster</i>	D14	99.6%	64.4%	MGVSGGLEALNVRVVGSGRRILVLSHGFGTDQSVWQRILPYFVRDFRIV
4	<i>Pinus pinaster</i>	KAI2	98.9%	52.1%	----MYLTNAHN [●] VHVLGSGHELIVLAHGFGTDQSVWKHVVP [●] ILEGRYRLV
5	<i>Taxus baccata</i>	D14	98.5%	62.5%	---GSGLLNALN [●] VKVI [●] TGRRILVLSHGFGTDQSVWQRILPHVDFRQIV
6	<i>Taxus baccata</i>	KAI2	97.8%	54.6%	-----VEAHNVHVLGSGQ [●] EVII [●] LAHGFGTDQSVWKHVVP [●] HLVDSYRII
7	<i>Welwitschia mirabilis</i>	KAI2	96.6%	54.8%	-----AHNVHVLGFGCEVIVLAHGFGTDQSVWRHVLPHLVDSYKLV
1	<i>Arabidopsis thaliana</i>	D14	100.0%	100.0%	51 LYDLVCAGSVNPDYDFDNRYTTLD [●] PYVDDLLNIVDSLGIQNCAYVGH [●] SVS
2	<i>Arabidopsis thaliana</i>	KAI2	99.3%	50.7%	LYDNMGAGTTNPDYDFDRYSNLEGYSFDLIAILEDLKI [●] ESCIFVGH [●] SVS
3	<i>Pinus pinaster</i>	D14	99.6%	64.4%	LYDLMCAGSVNPDNDFQRYSSLYAYADDLITILEELNVDN [●] CFVGH [●] SLS
4	<i>Pinus pinaster</i>	KAI2	98.9%	52.1%	MFDNMGAGTTNPEFFDFERYSTLYGYAYD [●] LLAILEELQVDSCIFVGH [●] SVS
5	<i>Taxus baccata</i>	D14	98.5%	62.5%	LFDLICAGSVNPDYDFQRYGCVEAYAEDLITILDELGVENC [●] VFVGH [●] SLS
6	<i>Taxus baccata</i>	KAI2	97.8%	54.6%	IFDNMGAGTTNPDFFDFDRYSTLHGAYDMLAILDELQIESCI [●] FLGH [●] SVS
7	<i>Welwitschia mirabilis</i>	KAI2	96.6%	54.8%	IFDNMGAGTTNPEYDFERYSSLYGYAFD [●] LLAILEELQVDSCIFVGH [●] SVS
1	<i>Arabidopsis thaliana</i>	D14	100.0%	100.0%	101 AMIGIIASIRRP [●] ELFSKLILIGFS [●] PRFLNDEDYHGGFEEGEIEK [●] VFSAME
2	<i>Arabidopsis thaliana</i>	KAI2	99.3%	50.7%	AMIGVLASLN [●] RPDFLFSKIVMISASPRYVNDVDYQGGF [●] EQEDLNQ [●] LFEAIR
3	<i>Pinus pinaster</i>	D14	99.6%	64.4%	AMIGCMASIKQ [●] PDLFSKLVLF [●] GASPRYSNDNEYEGGFDKQ [●] IEKVF [●] GAME
4	<i>Pinus pinaster</i>	KAI2	98.9%	52.1%	GLVGMLASIER [●] PDLFSKIIAISASPRYLNDIDYFGG [●] FQEDLNQ [●] LF [●] EAMQ
5	<i>Taxus baccata</i>	D14	98.5%	62.5%	AMVGCMA [●] SIKRP [●] SLFTKLVLF [●] GASPRYLNDNEYQGGFERIEIQ [●] VF [●] GAME
6	<i>Taxus baccata</i>	KAI2	97.8%	54.6%	GLIGCLASIEK [●] PLLSK [●] LIMI [●] SASPRYLNDVDYFGG [●] FQEDLDQ [●] LF [●] EAMQ
7	<i>Welwitschia mirabilis</i>	KAI2	96.6%	54.8%	GMVGC [●] LASLQ [●] KPYLFSK [●] MVLI [●] GASPRYLNDVGYHGGF [●] EQEDLEQ [●] LF [●] QAMQ
1	<i>Arabidopsis thaliana</i>	D14	100.0%	100.0%	151 ANYEAWVHG [●] FAPLAVGADVP--AAVREFS [●] RTLF [●] NMRPD [●] ISL [●] FVSRT [●] VFNSD
2	<i>Arabidopsis thaliana</i>	KAI2	99.3%	50.7%	SNYKAWL [●] CGFAPLAVGGDMSIAVQ [●] EFS [●] RTLF [●] NMRPD [●] IALSVG [●] QTIF [●] QSD
3	<i>Pinus pinaster</i>	D14	99.6%	64.4%	SNYA [●] AWVSGFAPLAVGGDVP--VAVREFS [●] RTLF [●] NMRPD [●] I [●] ALYVAK [●] TVFETD
4	<i>Pinus pinaster</i>	KAI2	98.9%	52.1%	SNFKAWVSGFAPLAVGADLDSIAVQ [●] EFS [●] RTLF [●] NMRPD [●] IALSVAK [●] TIFQSD
5	<i>Taxus baccata</i>	D14	98.5%	62.5%	SNYKDWVSGFAPLLV [●] GADVP--VAVREFS [●] RTLF [●] NMRPD [●] I [●] ALYVAK [●] IVFESD
6	<i>Taxus baccata</i>	KAI2	97.8%	54.6%	SNFKAWVSGFAPLAVGADLDSIAVQ [●] EFS [●] RTLF [●] NMRPD [●] IALSVAK [●] TIFQSD
7	<i>Welwitschia mirabilis</i>	KAI2	96.6%	54.8%	SNYKAWVSGFAPLAVGSDMSDAVQ [●] EFS [●] RTLF [●] NMRPD [●] I [●] ALCVAK [●] TIFQSD
1	<i>Arabidopsis thaliana</i>	D14	100.0%	100.0%	201 LRGVGLV [●] RVPTCVIQTAKD [●] VSPASVAEYLRSHLGGDTT [●] VETLKTEGHL
2	<i>Arabidopsis thaliana</i>	KAI2	99.3%	50.7%	MRQILPFV [●] TPCHILQSVKDLAVP [●] VVSEY [●] LHANLGCESVVEV [●] IPSDGHL
3	<i>Pinus pinaster</i>	D14	99.6%	64.4%	MRGILGQV [●] KVPCCVIQTSKDMAV [●] PLGVAIYVKKHLGCNATLEILDTEGHL
4	<i>Pinus pinaster</i>	KAI2	98.9%	52.1%	MRSLLPHV [●] TPCHILQSSKDLAVP [●] VAVAEYIHQNLGGKSI [●] VEILPSEGLH
5	<i>Taxus baccata</i>	D14	98.5%	62.5%	MRGILKQV [●] KVPCCVIQTSKDMAV [●] PLAVAIYVQKNLGGKSTLEILDTEGHL
6	<i>Taxus baccata</i>	KAI2	97.8%	54.6%	LRGILPHV [●] TPCHILQSSKDLAVP [●] VVSEYMQEHLGGKSI [●] VEILPTEGHL
7	<i>Welwitschia mirabilis</i>	KAI2	96.6%	54.8%	MRALLPHV [●] TPCHVLSCKDLAVP [●] IFVSEYIRQNLGGKTIVEILPAEGHL
1	<i>Arabidopsis thaliana</i>	D14	100.0%	100.0%	251 PQLSAPAQLAQFLRRALPR----
2	<i>Arabidopsis thaliana</i>	KAI2	99.3%	50.7%	PQLSSPDSVIPVILRRHNRDIAM
3	<i>Pinus pinaster</i>	D14	99.6%	64.4%	PHLSQPAKIQLLKRVLQ-----
4	<i>Pinus pinaster</i>	KAI2	98.9%	52.1%	PQLSSPAIVLPVLLRHIEGY--I
5	<i>Taxus baccata</i>	D14	98.5%	62.5%	PHLSQPSKMIQLLRRI-L-----
6	<i>Taxus baccata</i>	KAI2	97.8%	54.6%	PQLSSPDIMVVVILRHIRG----
7	<i>Welwitschia mirabilis</i>	KAI2	96.6%	54.8%	PQLSSPDVVIPVLLRHIR-----

Figura 8. Alineamiento de las secuencias D14 y KAI2 de *A. thaliana* (angiosperma), *P. pinaster* (gimnosperma pinácea), *T. baccata* (gimnosperma taxácea) y *W. mirabilis* (gimnosperma gnetal). Se han resaltado los residuos aminoacídicos más relevantes para la actividad proteica. ● Triada catalítica, ● Aminoácidos descritos como críticos para la interacción con MAX2 y ● Aminoácidos descritos como esenciales para la interacción con MAX2. COV = Cobertura del alineamiento, PID = Porcentaje de identidad.

El árbol filogenético resultante de aplicar el algoritmo UPGMA (Figura 9) muestra una clara división de las proteínas analizadas en dos grandes grupos: el clado **DWARF14** (D14), más similar al receptor de estrigolactonas (SL) descrito en *A. thaliana*; y el clado **KARRIKIN INSENSITIVE2** (KAI2), más próximo al receptor de karrikinas (KR), también descrito en la especie modelo *A. thaliana*.

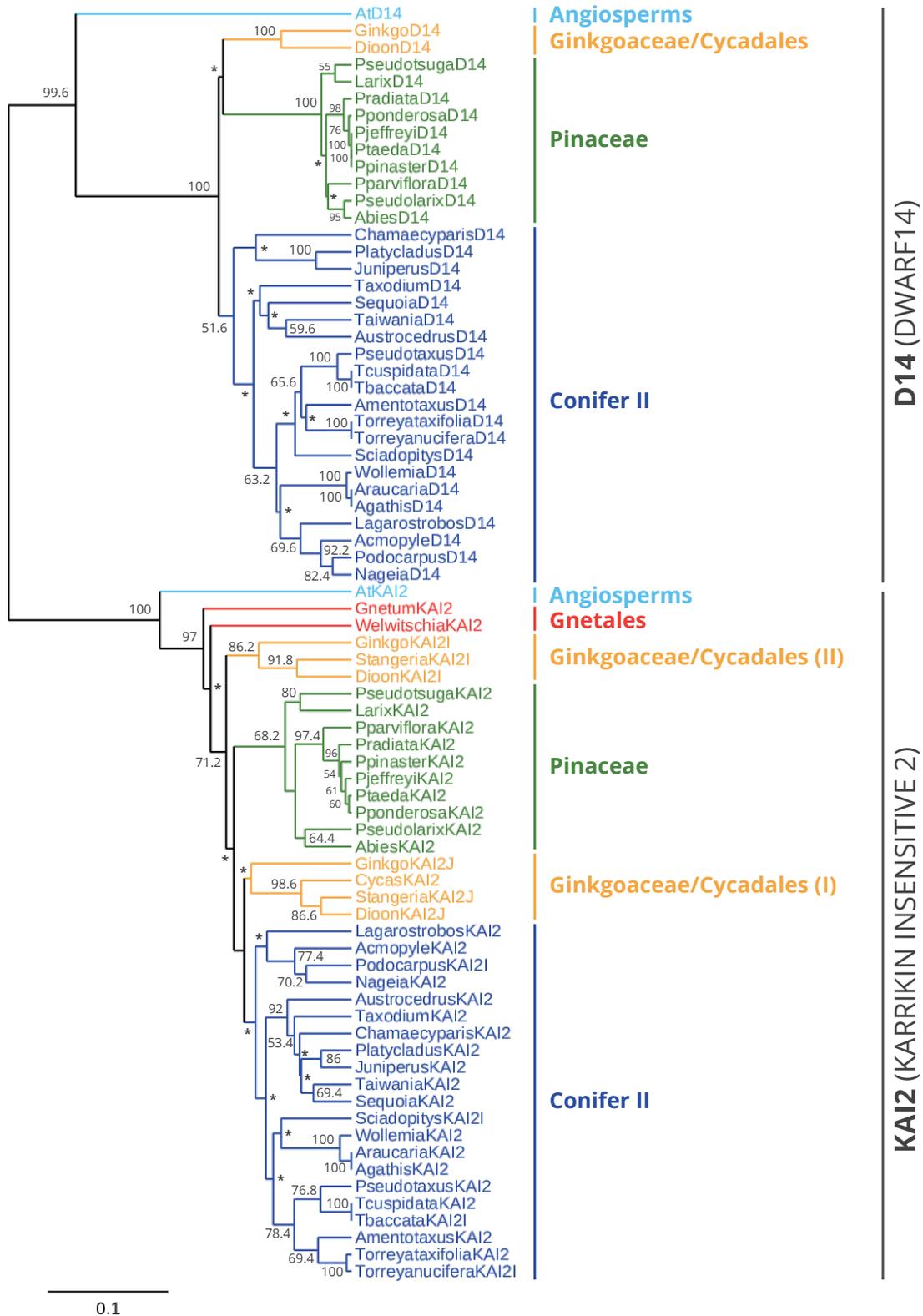


Figura 9. Filograma del árbol más probable para los clados D14 y KAI2 en los principales grupos de gimnospermas, implementado mediante el algoritmo **UPGMA** utilizando la familia D14/KAI2 al completo (73 secuencias de 37 especies). Las secuencias caracterizadas de la especie angiosperma modelo *Arabidopsis thaliana* se han añadido para mayor claridad (*AtD14*, *AtKAI2*). Los números asociados a ramas internas denotan fiabilidad del *bootstrap* (%); valores inferiores a 50% están indicados por *.

3.2.1. D14 Y KAI2 SE AGRUPAN EN CLADOS TAXÓN-ESPECÍFICOS

Al mismo tiempo, se puede observar una buena agrupación de cada uno de estos dos receptores, D14 y KAI2, **en función del grupo de gimnospermas** al que pertenecen: gnetales, ginkgoáceas y cicadáceas, pináceas y el supergrupo Coníferas II, que ya ha sido previamente descrito en los estudios filogenéticos más recientes llevados a cabo sobre gimnospermas y aglutina al resto de familias (Ran *et al.*, 2018). Como se puede comprobar en la **Figura 10**, este patrón de agrupación se repite en la estimación filogenética obtenida con el algoritmo NJ, aunque en este caso el gen D14 de *A. thaliana* se ha utilizado como grupo externo a la hora de construir el filograma. En conclusión, y al igual que ocurría en otros análisis filogenéticos previos (Bythell-Douglas *et al.*, 2017), los miembros de la familia D14/KAI2 tienden a colocarse en **clados taxón-específicos**.

Los principales modelos aceptados actualmente acerca del origen de la familia proteica D14/KAI2 indican que ambos receptores surgieron por la duplicación de un gen KAI2 ancestral. A continuación, los dos genes resultantes de la duplicación se habrían especializado y adquirido especificidad hacia las estrigolactonas (D14) o karrikinas (KAI2) (Waters *et al.*, 2017). Sin embargo, aunque sabemos que D14 y KAI2 provienen de un mismo gen ancestral, el momento exacto en el que esta duplicación ocurrió aún es desconocido. La agrupación tan clara que observamos de D14 y KAI2 en clados taxón-específicos apoya la idea general de que la separación entre ambos receptores es ancestral y habría ocurrido mucho antes del origen de las gimnospermas, dando así tiempo evolutivo suficiente para que ambos receptores siguieran patrones distintos según el grupo de organismos.

3.2.2. LA ESPECIFICIDAD DE LOS RECEPTORES D14 Y KAI2 PODRÍA HABER SURGIDO CON LAS GIMNOSPERMAS

La mayoría de las hipótesis que existen acerca de la evolución de las estrigolactonas en el reino vegetal sugieren que estas fitohormonas surgieron en **organismos muy primitivos**, como en el grupo de los cariófitos (Delaux *et al.*, 2012). Este modelo afirma que las SL podrían estar implicadas en la elongación de rizoides, que incrementa la capacidad de anclaje de los organismos al suelo, y por ende este grupo de hormonas podrían haber jugado un papel clave en la colonización de la tierra por parte de las plantas. Aproximaciones más recientes como la de Bythell-Douglas *et al.* (2017) respaldan esta idea, situando la separación entre los receptores KAI2 y D14 y, por tanto, el origen de la percepción de estrigolactonas, mucho antes del origen de las plantas con semilla.

En relación a ello, uno de los aspectos más relevantes de este análisis es la **ausencia de D14 en el grupo de los gnetales**, algo muy similar a lo que se observa en grupos de plantas basales como, precisamente, los cariófitos. Con las herramientas bioinformáticas y los criterios de filtrado empleados, no se ha podido identificar más de una secuencia de la familia D14/KAI2 en este grupo de gimnospermas. Esto podría ser debido a la pérdida del gen D14 en este grupo de gimnospermas, o bien a que la que dio lugar a los receptores D14 y KAI2 que conocemos hoy en día podría haber surgido mucho después de lo que se afirma en trabajos anteriores. Si este fuera el caso y la especificidad de D14 y KAI2 hacia SL y KR, respectivamente, hubiera surgido con las gimnospermas, el papel de estas hormonas en la evolución del reino vegetal podría ser muy distinto al descrito anteriormente.

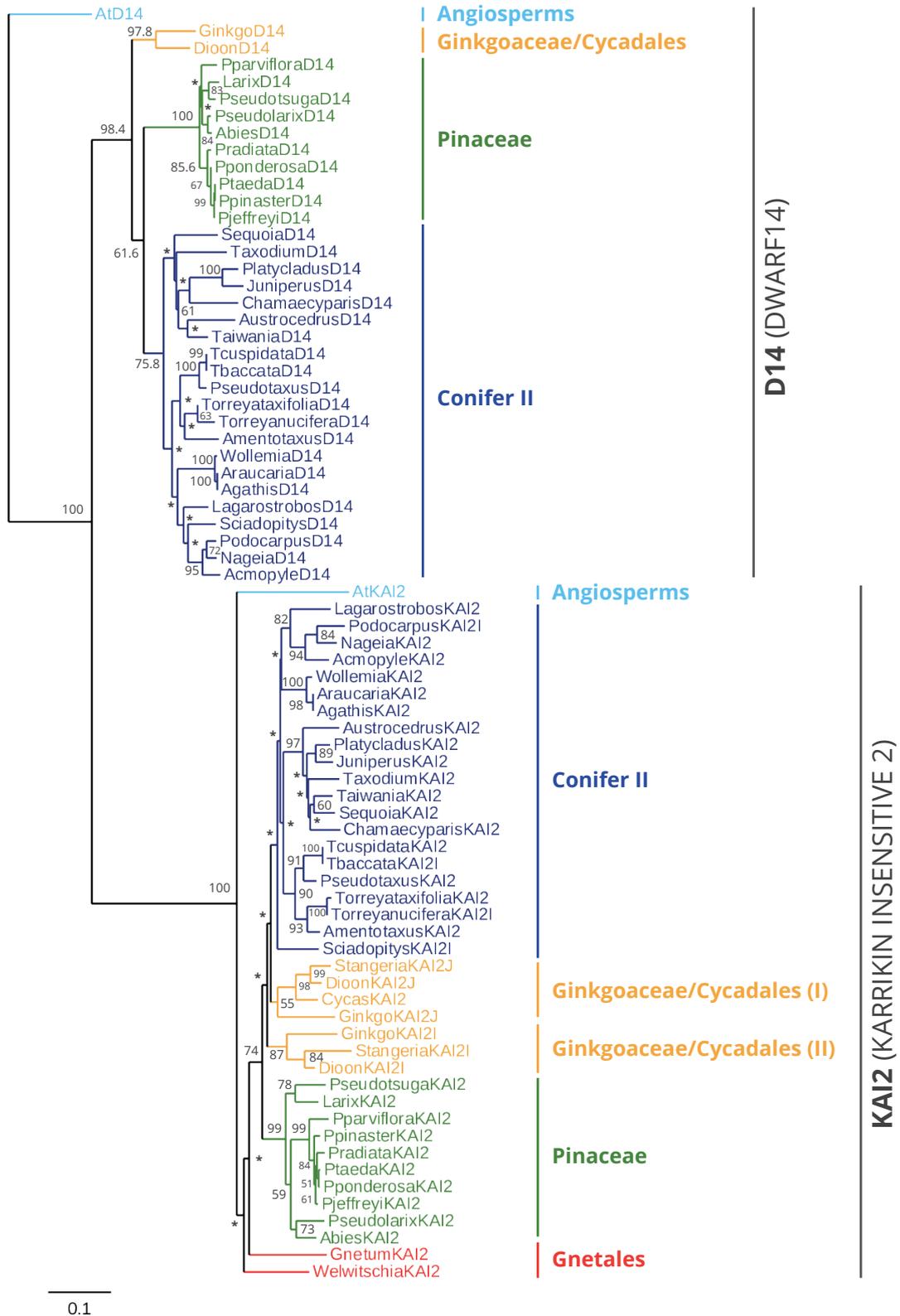


Figura 10. Filograma del árbol más probable para los clados D14 y KAI2 en los principales grupos de gimnospermas, implementado mediante el algoritmo **Neighbor-Joining** utilizando la familia D14/KAI2 al completo (73 secuencias de 37 especies). Las secuencias caracterizadas de la especie angiosperma modelo *Arabidopsis thaliana* se han añadido para mayor claridad (*AtD14*, *AtKAI2*). Los números asociados a ramas internas denotan fiabilidad del *bootstrap* (%); valores inferiores a 50% están indicados por *.

Toda esta discrepancia respecto a trabajos anteriores se debe a que aproximaciones como la de [Bythell-Douglas *et al.* \(2017\)](#) tratan de desvelar el momento en que D14 y KAI2 se separaron sin considerar la especificidad y capacidad de estas proteínas para funcionar. Al tener en cuenta los residuos aminoacídicos necesarios para desencadenar las respuestas a karrikinas y estrigolactonas nos acercamos, no ya a conocer el origen de las proteínas, sino el origen de la especificidad de dichas proteínas por sus respectivos ligandos.

En conclusión, aunque la divergencia que dio lugar a D14 y KAI2 pudiese ocurrir en grupos de plantas basales, no existe evidencia experimental de que la especificidad de ambas proteínas (y su preferencia hacia SL y KR respectivamente) se estableciese en un momento tan temprano de la evolución. Existe una gran incertidumbre acerca del papel que ambos compuestos podrían jugar en organismos como las gimnospermas, así como su mecanismo de percepción exacto. Esta aproximación filogenética a la cuestión podría sugerir que **la especificidad hacia karrikinas y estrigolactonas hubiese surgido dentro de las gimnospermas**. No obstante, para confirmar esta hipótesis sería necesario llevar a cabo un estudio posterior que se escapa del alcance de este trabajo.

No obstante, hay que tener en cuenta que las bases de datos relativas a gimnospermas son escasas y la mayoría de especies no cuenta con un genoma secuenciado. Al mismo tiempo, la mayoría de aproximaciones transcriptómicas tienen una baja cobertura, lo que dificulta la recuperación de transcritos completos. Todo ello impide afirmar irrefutablemente que el D14 de los organismos gnetales pudiera corresponder a un transcrito incompleto que no haya superado los criterios de filtrado.

3.2.3. LAS ESTRIGOLACTONAS PODRÍAN CONTRIBUIR AL ESTUDIO DE LA EVOLUCIÓN VEGETAL

De hecho, toda esta dificultad a la hora de llevar a cabo análisis genómicos en gimnospermas ha dado lugar a **problemas a la hora de evaluar su evolución**. En el caso concreto de los gnetales, su gran parecido morfológico a las angiospermas les ha convertido en un grupo polémico con varias hipótesis propuestas respecto a su posición dentro de las gimnospermas ([Jiao and Guo, 2014](#)) ([Figura 11](#)).

Suponiendo que las bases de datos empleadas en este análisis sean completas, la agrupación de los receptores KAI2 de gimnospermas podría tener algo que decir al respecto de la evolución de las gimnospermas, aportando **evidencia a favor de la hipótesis Gnetifera**, que ubica a los gnetales como clado hermano del resto de las coníferas. Por otra parte, estos resultados son opuestos a lo que se sugiere en

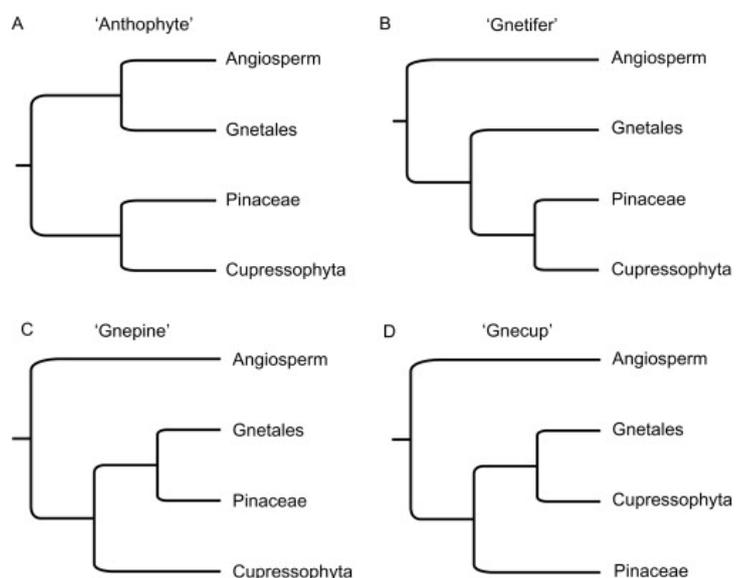


Figura 11. Hipótesis más populares de la posición filogenética de los gnetales. Extraído de [Jiao y Guo, 2014](#).

aproximaciones filo-transcriptómicas recientes, más partidarias de la hipótesis Gnepina (Ran *et al.*, 2018), la cual posiciona a los gnetales como grupo hermano de las pináceas, mucho más alejados del resto de coníferas (Figura 11).

3.2.4. EL RECEPTOR KAI2 SE DUPLICA EN LA BASE DE LAS GIMNOSPERMAS

Otro aspecto remarcable de las estimaciones filogenéticas obtenidas para la familia D14/KAI2 en coníferas es el hecho de que existen **dos secuencias de KAI2 distintas** en las cicadáceas y ginkgoáceas. Esto coincide con observaciones de trabajos previos (Bythell-Douglas *et al.*, 2017) y es probablemente debido a una **duplicación del receptor KAI2** que ocurrió en la base de las gimnospermas, puesto que estos son los únicos grupos en los que se observa.

Por este motivo, tanto cicadáceas como ginkgoáceas cuentan con un receptor D14 y dos KAI2: uno de ellos más próximo al KAI2 del resto de gimnospermas (KAI2J), y otro más alejado (KAI2I). Sería necesaria una aproximación experimental a esta cuestión para determinar si ambas secuencias mantienen la especificidad hacia el mismo ligando, presumiblemente las karrikinas, o habrían divergido también a nivel funcional.

3.3. IDENTIFICACIÓN Y CLONACIÓN DE *PpD14*

3.3.1. OBTENCIÓN DE LA CDS DEL GEN *PpD14*

La búsqueda mediante BLASTp en las bases de datos Gymno PLAZA 3.0 y SustainPineDB, utilizando la secuencia proteica del receptor D14 de *A. thaliana* permitió obtener una secuencia homóloga al gen *AtD14* en la especie *Pinus pinaster*, *PpD14*.

El **alineamiento** de ambos genes mediante el algoritmo Needleman-Wunsch (matriz de sustitución EBLOSUM 62) permitió comprobar que existe una **elevada similitud** entre ambos, incluso a nivel nucleotídico. Se observa una identidad del **61.9%** (nucleótidos idénticos entre ambas secuencias) y una similitud de también **61.9%** (residuos similares teniendo en cuenta posibles sustituciones, deleciones...). Todo ello apoya la homología entre estas secuencias y constituye un indicio de que ambas proteínas podrían tener funciones conservadas incluso tratándose de especies evolutivamente muy distantes.

Para confirmar esta hipótesis, se utilizó la herramienta en línea **InterPro** (<https://www.ebi.ac.uk/interpro/>) (Mitchell *et al.*, 2019), que permite identificar y anotar los principales **dominios proteicos** presentes en una secuencia aminoacídica. Tal y como se puede observar en la **Figura 12**, las anotaciones obtenidas corresponden a la familia proteica de α/β -hidrolasas o bien a dominios de pliegue atribuibles a la misma. Todo ello coincide con la información de la que se dispone acerca del receptor D14, así como los dominios anotados en la proteína D14 ya caracterizada de *A. thaliana*.

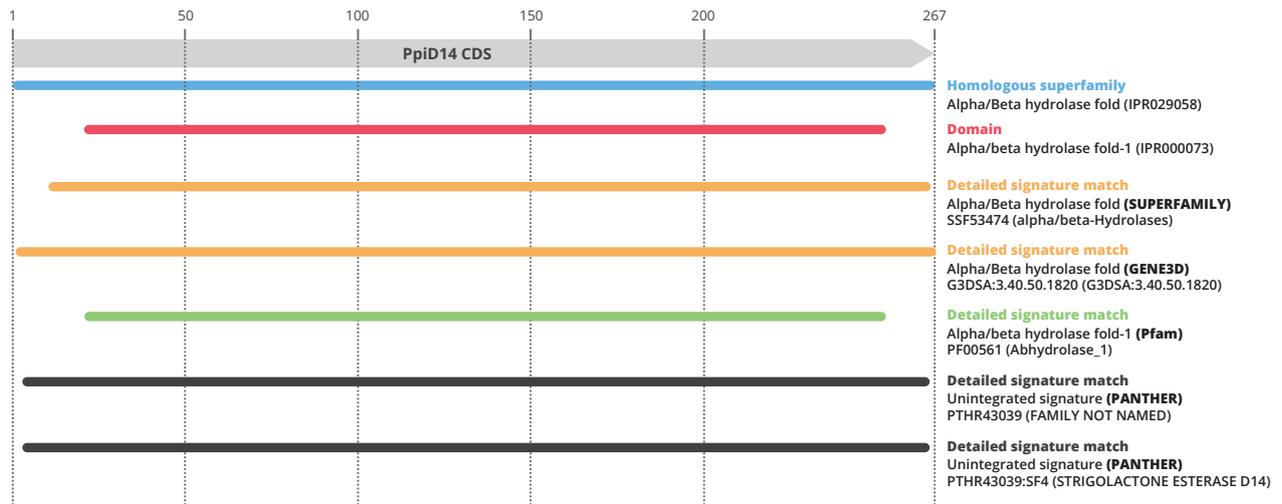


Figura 12. Representación gráfica de los principales dominios proteicos anotados en la secuencia *PpD14* con la herramienta InterPro. Se detalla el tipo de dominio, su código y la base de datos de la que se ha extraído la anotación.

Con el objetivo de seguir recabando indicios de que la secuencia candidata de *P. pinaster* obtenida sea la homóloga al receptor de estrigolactonas D14 de *A. thaliana*, se llevó a cabo una comparación de la **estructura tridimensional** predicha de ambas proteínas. Para ello, se utilizó la herramienta en línea **Phyre 2.0** (Kelley *et al.*, 2015), que permite obtener una predicción del modelado 3D de una proteína a partir de su secuencia aminoacídica. Los modelos generados para las proteínas *PpD14* y *AtD14* pueden observarse en la **Figura 13**.

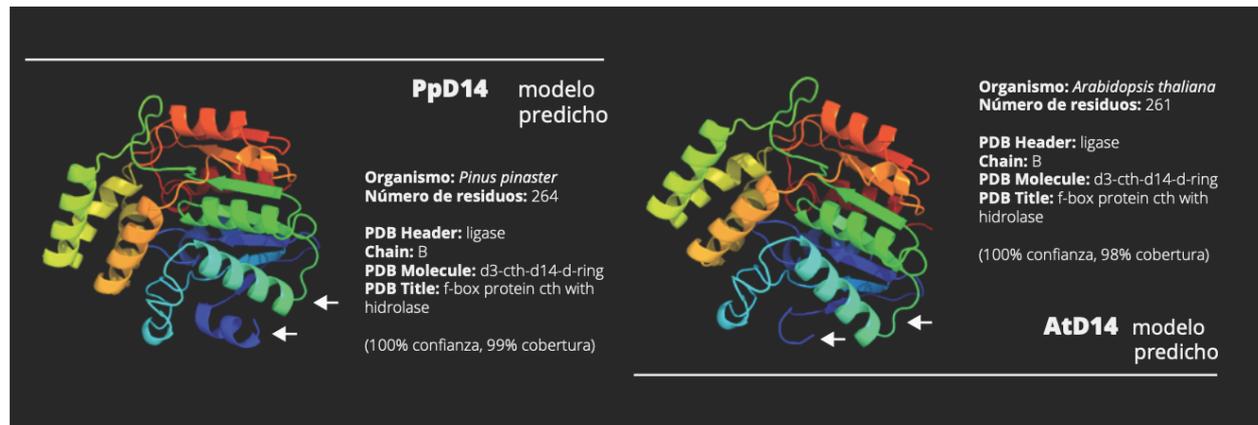


Figura 13. A la izquierda, modelo 3D predicho a partir de la secuencia aminoacídica de *PpD14* (*P. pinaster*). A la derecha, *AtD14* (*A. thaliana*). Se indican el número de residuos y los principales datos del modelo generado, así como el porcentaje de confianza de la predicción y el porcentaje de cobertura de la secuencia primaria de la proteína. Las zonas diferenciales aparecen resaltadas sobre el propio modelo.

Tal y como se podría esperar dada la alta similitud de las proteínas a nivel de secuencia, la estructura tridimensional es muy parecida entre ambas predicciones. La estructura secundaria es prácticamente idéntica, con el mismo número de hélices alfa y láminas beta, aunque en posiciones ligeramente distintas. Las diferencias más notables aparecen en las

posiciones relativas de estas estructuras secundarias a nivel tridimensional, así como en el extremo N-terminal, pero en ningún caso son apreciables en la parte central de la proteína, que contiene la tríada catalítica y la región de unión al ligando (estrigolactona).

Tomados en conjunto, todos estos resultados respaldan la hipótesis de que la secuencia obtenida es homóloga al gen D14 de *A. thaliana* y, por tanto, aportan un fuerte indicio de que dicha secuencia muy probablemente ejerce las funciones de **receptor de estrigolactonas** en especie *P. pinaster*.

3.3.2. CLONACIÓN DEL GEN *PpD14* Y CONSTRUCCIÓN DE UN VECTOR BINARIO DE SOBREENPRESIÓN MEDIANTE EL SISTEMA GATEWAY

A expensas de todo lo comentado anteriormente, una aproximación bioinformática no equivale ni sustituye a una **validación funcional de un gen**. Aunque existan abundantes indicios de que la secuencia *PpD14* obtenida es el receptor de estrigolactonas en la especie *P. pinaster*, es necesario un enfoque experimental, como la complementación heteróloga, para corroborar definitivamente la hipótesis. Aunque es evidente que una complementación heteróloga excede con creces la temporalidad de este TFM, si se planteó como objetivo del mismo la **clonación del gen *PpD14*** y la construcción de un vector binario de sobreexpresión que sirviese como base para futuros ensayos de validación del gen a nivel funcional.

Para generar el vector de sobreexpresión se escogió la tecnología Gateway® (Thermo Fisher Scientific) por su alta flexibilidad a la hora de crear distintos tipos de construcciones, así como por su alta eficiencia. El vector de destino escogido fue el **pMDC32**, que contiene un doble promotor 35S (2X35S), así como dos agentes de selección: una resistencia a higromicina para la selección en planta y una resistencia a kanamicina en bacteria. Un esquema completo del vector de expresión resultante (pMDC32 con el gen completo de *PpD14*) se puede consultar en la **Figura 14**.

Para confirmar la fiabilidad del proceso de clonación del gen a partir de ADNc conseguido de yemas apicales de plántulas, se realizaron secuenciaciones de control periódicas tras cada nuevo constructo generado (como se comentó en los apartados 2.4.2 y 2.4.3). En cada uno de los casos se comprobó que la secuencia nucleotídica insertada en los vectores correspondía a la CDS de *PpD14* con un 100% de identidad.

Una vez introducido *PpD14* en el sistema Gateway® y obtenido el vector de sobreexpresión (pMDC32/*PpD14*), el siguiente paso será utilizar esta construcción para transformar células competentes de *Agrobacterium tumefaciens*. Esto servirá como primer paso de cara a validar experimentalmente la función de *PpD14*, utilizando *Arabidopsis thaliana* como sistema heterólogo a pino y mediante una estrategia de complementación de mutantes. En definitiva, los indicios recabados con el análisis bioinformático de *PpD14*, así como la construcción de un vector de sobreexpresión permitirá, en el futuro, arrojar luz sobre la percepción de estrigolactonas en *P. pinaster*, una especie de gran interés económico y ecológico.

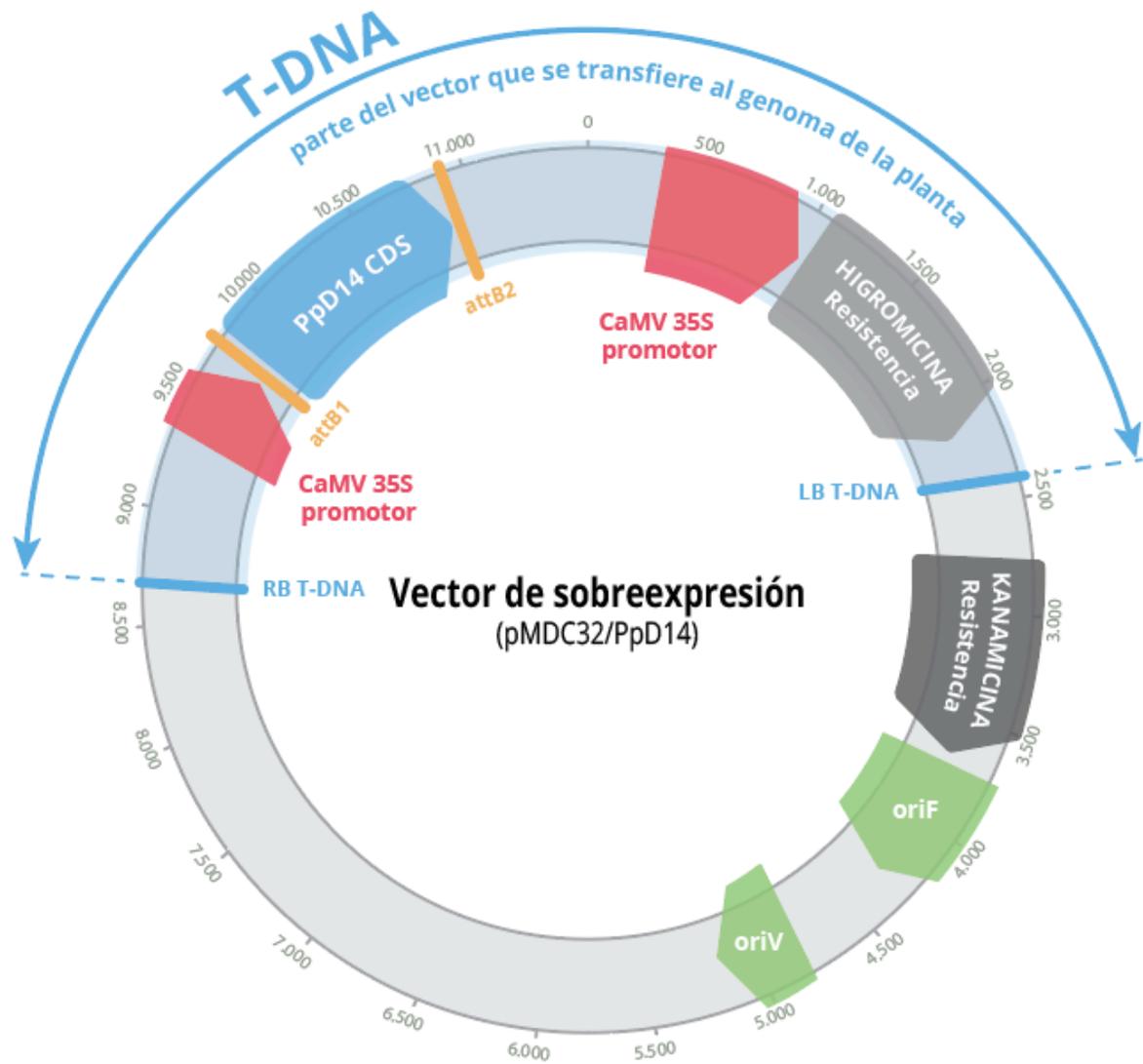


Figura 14. Vector de sobreexpresión generado a partir del plásmido pMDC32 y la secuencia *PpD14* tras la reacción LR de la tecnología Gateway.

4. CONCLUSIONES

1. El receptor de estrigolactonas *DWARF14* está altamente conservado en el reino vegetal, lo cual denota una alta conservación del mecanismo de percepción y señalización de estrigolactonas.
2. La mayoría de grupos de gimnospermas cuentan con una proteína homóloga al receptor de estrigolactonas *DWARF14* y otra homóloga al receptor de karrikinas *KAI2*, aunque las funciones y especificidad de las mismas aún son desconocidas y se escapan del alcance de este trabajo.
3. La ausencia de un receptor de estrigolactonas *D14* funcional en gnetales podría indicar que la especialización de los receptores *D14/KAI2* hacia sus respectivos ligandos, *SL/KR*, surgió con las gimnospermas.
4. La clonación del homólogo a *D14* en *P. pinaster* constituye un primer paso hacia la validación funcional del gen *PpD14* y la identificación del hasta ahora desconocido receptor de estrigolactonas en coníferas.

5. BIBLIOGRAFÍA

1. Barbier, F. F. *et al.* (2019) 'An update on the signals controlling shoot branching', *Trends in Plant Science*. Elsevier Ltd, 24(3), pp. 220–236. doi: 10.1016/j.tplants.2018.12.001.
2. Bennett, T. and Leyser, O. (2014) 'Strigolactone signalling: Standing on the shoulders of DWARFs', *Current Opinion in Plant Biology*. Elsevier Ltd, 22, pp. 7–13. doi: 10.1016/j.pbi.2014.08.001.
3. Brewer, P. B. *et al.* (2015) 'Strigolactone inhibition of branching independent of Polar Auxin Transport', *Plant Physiology*, 168(4), pp. 1820–1829. doi: 10.1104/pp.15.00014.
4. Bythell-Douglas, R. *et al.* (2017) 'Evolution of strigolactone receptors by gradual neofunctionalization of KAI2 paralogues', *BMC Biology*, 15(52). doi: 10.1186/s12915-017-0397-z.
5. Calderón, Á. *et al.* (2016) 'Nuevos marcadores de calidad de madera en *Pinus pinaster*', *Tecnología Agroalimentaria: Boletín informativo del SERIDA*, 17, pp. 21–27.
6. Chevenet, F. *et al.* (2006) 'TreeDyn: Towards dynamic graphics and annotations for analyses of trees', *BMC Bioinformatics*, 7, pp. 1–9. doi: 10.1186/1471-2105-7-439.
7. Delaux, P.-M. *et al.* (2012) 'Origin of strigolactones in the green lineage', *New Phytologist*, 195(4), pp. 857–871. doi: 10.1111/j.1469-8137.2012.04209.x.
8. Dereeper, A. *et al.* (2008) 'Phylogeny.fr: robust phylogenetic analysis for the non-specialist.', *Nucleic acids research*, 36(Web Server issue), pp. W465-9. doi: 10.1093/nar/gkn180.
9. Dereeper, A. *et al.* (2010) 'BLAST-EXPLORER helps you building datasets for phylogenetic analysis', *BMC Evolutionary Biology*, 10(1), pp. 8–13. doi: 10.1186/1471-2148-10-8.
10. Domagalska, M. A. and Leyser, O. (2011) 'Signal integration in the control of shoot branching', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 12(4), pp. 211–221. doi: 10.1038/nrm3088.
11. Fernández-Pozo, N. *et al.* (2011) 'EuroPineDB: A high-coverage web database for maritime pine transcriptome', *BMC Genomics*. BioMed Central Ltd, 12(1), p. 366. doi: 10.1186/1471-2164-12-366.
12. Hevia, A., Álvarez-González, J. G. and Majada, J. (2016) 'Effects of pruning on knotty core taper and form of *Pinus radiata* and *Pinus pinaster*', *European Journal of Wood and Wood Products*, 74(5), pp. 741–750. doi: 10.1007/s00107-016-1019-0.
13. Jiao, Y. and Guo, H. (2014) 'Prehistory of the Angiosperms: Characterization of the Ancient Genomes', *Advances in Botanical Research*, 69, pp. 223–245.
14. Jordy, M. N. (2004) 'Seasonal variation of organogenetic activity and reserves allocation in the shoot apex of *Pinus pinaster* ait', *Annals of Botany*, 93(1), pp. 25–37. doi: 10.1093/aob/mch005.
15. Katoh, K. *et al.* (2002) 'MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform.', *Nucleic acids research*, 30(14), pp. 3059–66.
16. Katoh, K. and Standley, D. M. (2013) 'MAFFT multiple sequence alignment software version 7: improvements in performance and usability.', *Molecular biology and evolution*, 30(4), pp. 772–80. doi: 10.1093/molbev/mst010.
17. Kebrom, T. H. (2017) 'A growing stem inhibits bud outgrowth – The overlooked theory of apical dominance', *Frontiers in Plant Science*, 8(October), pp. 1–7. doi: 10.3389/fpls.2017.01874.

18. Kelley, L. A. *et al.* (2015) 'The Phyre2 web portal for protein modelling, prediction, and analysis', *Nature Protocols*. Nature Publishing Group, 10(6), pp. 845–858. doi: 10.1038/nprot.2015-053.
19. Leyser, O. (2018) 'Auxin signaling', *Plant Physiology*, 176, pp. 465–479. doi: 10.1016/j.pbi.2006.07.006.
20. Lopez-Obando, M. *et al.* (2015) 'Strigolactone biosynthesis and signaling in plant development', *The Company of Biologists, Development*, 142, pp. 3615–3619. doi: 10.1242/dev.120006.
21. Lopez-Obando, M. *et al.* (2018) '*Physcomitrella patens* MAX2 characterization suggests an ancient role for this F-box protein in photomorphogenesis rather than strigolactone signalling', *New Phytologist*, 219(2), pp. 743–756. doi: 10.1111/nph.15214.
22. Madeira, F. *et al.* (2019) 'The EMBL-EBI search and sequence analysis tools APIs in 2019', *Nucleic Acids Research*, 47(April), pp. 636–641. doi: 10.1093/nar/gkz189.
23. Marzec, M. (2016) 'Perception and signaling of strigolactones', *Frontiers in Plant Science*, 23(7). doi: 10.3389/fpls.2016.01260.
24. Marzec, M. and Brewer, P. (2019) 'Binding or hydrolysis? How does the strigolactone receptor work?', *Trends in Plant Science*. doi: 10.1016/j.tplants.2019.05.001.
25. Matasci, N. *et al.* (2014) 'Data access for the 1,000 Plants (1KP) project', *GigaScience*, 3(17). doi: 10.1186/2047-217X-3-17.
26. Meijón, M. *et al.* (2016) 'Exploring natural variation of *Pinus pinaster* Aiton using metabolomics: Is it possible to identify the region of origin of a pine from its metabolites?', *Molecular Ecology*, 25(4). doi: 10.1111/mec.13525.
27. Mitchell, A. L. *et al.* (2019) 'InterPro in 2019: Improving coverage, classification and access to protein sequence annotations', *Nucleic Acids Research*, 47(D1), pp. D351–D360. doi: 10.1093/nar/gky1100.
28. Notredame, C., Higgins, D. G. and Heringa, J. (2000) 'T-coffee: A novel method for fast and accurate multiple sequence alignment', *Journal of Molecular Biology*, 302(1), pp. 205–217. doi: 10.1006/jmbi.2000.4042.
29. Proost, S. *et al.* (2015) 'PLAZA 3.0: An access point for plant comparative genomics', *Nucleic Acids Research*, 43(D1), pp. D974–D981. doi: 10.1093/nar/gku986.
30. Ran, J. H. *et al.* (2018) 'Phylogenomics resolves the deep phylogeny of seed plants and indicates partial convergent or homoplastic evolution between Gnetales and angiosperms', *Proceedings of the Royal Society Biological Sciences*, 285(1881). doi: 10.1098/rspb.2018.1012.
31. Végh, A. *et al.* (2017) 'Comprehensive analysis of DWARF14-LIKE2 (DLK2) reveals its functional divergence from strigolactone-related paralogs', *Frontiers in Plant Science*, 8(September), pp. 1–14. doi: 10.3389/fpls.2017.01641.
32. Walker, C. and Bennett, T. (2017) 'Reassessing the evolution of strigolactone synthesis and signalling', *bioRxiv*, p. 228320. doi: 10.1101/228320.
33. Walker, C. H. and Bennett, T. (2018) 'Forbidden fruit: Dominance relationships and the control of shoot architecture', *Annual Plant Reviews*. doi: 10.1002/9781119312994.apr0640.
34. Waters, M. T. *et al.* (2012) 'Specialisation within the DWARF14 protein family confers distinct responses to karrikins and strigolactones in *Arabidopsis*', *Development*, 139(7), pp. 1285–1295. doi: 10.1242/dev.074567.
35. Waters, M. T. *et al.* (2015) 'Substrate-induced degradation of the α/β -fold hydrolase KARRIKIN INSENSITIVE2 requires a functional catalytic triad but is independent of MAX2', *Molecular Plant*, 8(5), pp. 814–817. doi: 10.1016/j.molp.2014.12.020.
36. Waters, M. T. *et al.* (2017) 'Strigolactone signaling and evolution', *Annual Review of Plant Biology*. Annual Reviews , 68(1), pp. 291–322. doi: 10.1146/annurev-arplant-042916-040925.

37. Yao, R. *et al.* (2016) 'DWARF14 is a non-canonical hormone receptor for strigolactone', *Nature*, 536(7617), pp. 469–473. doi: 10.1038/nature19073.
38. Yao, R. *et al.* (2018) 'Rice DWARF14 acts as an unconventional hormone receptor for strigolactone', *Journal of Experimental Botany*, 69(9), pp. 2355–2365. doi: 10.1093/jxb/ery014.
39. Yao, R., Li, J. and Xie, D. (2018) 'Recent advances in molecular basis for strigolactone action', *Science China Life Sciences*, 61(3), pp. 277–284. doi: 10.1007/s11427-017-9195-x.