



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**Variación en la vía de respuesta a estrés genotóxico en
plantas (Viridiplantae)**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGA

P R E S E N T A:

GABRIELA GUZMÁN FAVILA



**DIRECTOR DE TESIS:
DRA. MA. DE LA PAZ SÁNCHEZ JIMÉNEZ
Cd. Mx. 2023**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

- | | |
|--|---|
| <p>1. Datos del alumno
Apellido paterno
Apellido materno
Nombre(s)
Teléfono
Universidad Nacional Autónoma de México
Facultad de Ciencias
Carrera
Número de cuenta</p> <p>2. Datos del tutor
Grado
Nombre(s)
Apellido paterno
Apellido materno</p> <p>3. Datos del sinodal 1
Grado
Nombre(s)
Apellido paterno
Apellido materno</p> <p>4. Datos del sinodal 2
Grado
Nombre(s)
Apellido paterno
Apellido materno</p> <p>5. Datos del sinodal 3
Grado
Nombre(s)
Apellido paterno
Apellido materno</p> <p>6. Datos del sinodal 4
Grado
Nombre(s)
Apellido paterno
Apellido materno</p> <p>7. Datos del trabajo escrito
Título
Subtítulo
Número de páginas
Año</p> | <p>1. Datos del alumno
Guzmán
Favila
Gabriela
5539841296
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
313290707</p> <p>2. Datos del tutor
Dra.
Ma. De La Paz
Sánchez
Jiménez</p> <p>3. Datos del sinodal 1
Dr.
César Luis
Cuevas
Velázquez</p> <p>4. Datos del sinodal 2
Dr.
Sol
Cristians
Niizawa</p> <p>5. Datos del sinodal 3
Dra.
María Del Rocío
Cruz
Ortega</p> <p>6. Datos del sinodal 4
Dr.
Francisco Javier
Plasencia
de la Parra</p> <p>7. Datos del trabajo escrito
Variación en la vía de respuesta a estrés genotóxico en plantas (Viridiplantae)
64 p
2023</p> |
|--|---|

Agradecimientos académicos

Esta tesis se realizó bajo la dirección de la Dra. Ma De La Paz Sánchez Jiménez, en el Laboratorio de Genética Molecular, Epigenética, Desarrollo y Evolución de Plantas del Instituto de Ecología, de la Universidad Nacional Autónoma de México, bajo la coordinación académico-científica de las Dras. Elena Alvarez-Buylla Roces, Adriana Garay Arroyo y Berenice García Ponce de León.

Agradezco la coordinación administrativa y logística, así como el apoyo técnico de la Dra. Diana Belén Sánchez Rodríguez.

Además, quiero agradecer al Dr. Marco Tulio Solano De la Cruz por su apoyo en enseñarme a utilizar todas las herramientas y técnicas bioinformáticas que utilicé en este trabajo.

Por último, quisiera agradecer por su tiempo a los miembros del jurado al revisar este trabajo

El financiamiento para esta investigación provino de proyectos de investigación:

UNAM-DGAPA-PAPIIT: IN206223, IN203223, IN211721.

CONACYT: CF-2019-102987, CF-2019-102959,

Agradecimientos personales

No soy buena expresándome, ni escribiendo, así que estos agradecimientos serán breves (pero sinceros).

A mis papás, por apoyarme siempre, en todos los sentidos. Sin ustedes no sólo no habría concluido la carrera y esta tesis, si no que tampoco sería lo que soy ahorita. A mi hermano, porque aunque peleemos siempre, mi vida no sería igual sin ti.

A mis tías y abuelitos, que siempre han estado ahí cuando necesito algo. A mi abuelita, estoy segura de que te hubiera gustado estar el día de mi titulación. A mi me hubiera gustado.

A Gris, por haberme apoyado tanto estos últimos años. Sin ti hubiera sido muy difícil encontrar motivación para seguir con la tesis y entrar a la maestría. Agradezco mucho haber coincidido contigo y espero sigas en mi vida por mucho tiempo más.

A Miguel, por aguantar mis crisis cada que quería cambiar de tema de tesis y por haberme descrito de la mejor forma que alguien pudo hacerlo en sus agradecimientos. A Diana, por escuchar cada una de mis crisis existenciales y acompañarme con las suyas; gracias por los proyectos concluidos e inconclusos, me daban motivación para seguir con la carrera.

A Mónica y a los cieguitos, espero que sigamos creciendo juntos. Gracias, Sandra.

A mis compañeros del laboratorio, por hacer los días más divertidos y aguantables, y por obligarme a hacer cosas que debía, pero no quería (como los trámites de titulación).

A mi tutora, Marypaz, por estar siempre presente, por no dejar a sus alumnxs solos con sus dudas y preocuparse por nosotrxs. Gracias por no echarme después de haberme tardado tanto con esta tesis.

A Paty Rivera, que contribuyó mucho a mi crecimiento académico sin tener la obligación de hacerlo. Gracias por aceptarme en tu laboratorio y perdón por arruinarte tantos experimentos.

Gracias a la música y al arte en general.

Índice

Resumen.....	1
Introducción.....	2
1. El clado Viridiplantae.....	3
2. Estrés genotóxico en plantas.....	4
2.1 Daño al DNA en plantas.....	4
2.1.1 Señalización de la respuesta a daño a DNA.....	5
2.1.2 Daño al DNA por Single Strand Break (SSB).....	5
2.1.3 Daño al DNA por Double Strand Break (DSB).....	7
2.2 Mecanismos de reparación de DSBs.....	8
2.2.1 Reparación Homóloga (HR).....	8
2.2.2 Reparación no homóloga (NHEJ).....	10
2.2.3 Muerte celular programada (PCD).....	11
3. Evolución de los mecanismos de reparación a DSBs.....	12
3.1 Variación en las rutas de reparación.....	12
3.2 Variación en las rutas de reparación en plantas.....	13
4. Planteamiento del problema.....	16
5. Hipótesis.....	16
6. Objetivo General.....	16
6.1 Objetivos particulares.....	16
7. Materiales y métodos.....	17
7.1 Identificación y caracterización de homólogos en Viridiplantae.....	17
7.2 Construcción de árboles filogenéticos y búsqueda de motivos.....	17
8. Resultados y Discusión.....	18
8.1 Análisis de las proteínas que participan en Single Strand break (SSB).....	20
8.1.1 Complejo 9-1-1.....	20
8.1.2 Replication protein 1 (RPA1).....	22
8.1.3 ATR interacting protein (ATRIP).....	23
8.2 Análisis de las proteínas que participan en Recombinación no homóloga (NHEJ)	24
7.2.1 Ku80/Ku70.....	24
8.3 Análisis de las proteínas que participan en la Recombinación Homóloga (HR)	25

8.3.1 Conservación de los principales señalizadores de la vía: ATM, ATR.....	26
8.3.2 Suppressor of gamma response 1 (SOG1).....	29
8.3.3 Breast cancer susceptibility gene 1 (BRCA1).....	32
8.4 Análisis de proteínas con funciones específicas dentro de la ruta de H	
HR	35
8.4.1 El complejo proteico MRN (MRE11, RAD50 y NBS1).....	35
8.4.2 Nijmegen breakage syndrome (NBS1).....	37
8.4.3 RAD50.....	38
8.5 Análisis de proteínas que participan en la búsqueda de cadenas	
homólogas.....	40
8.5.1 Breast cancer susceptibility gene 2 (BRCA2).....	40
8.5.2 RAD51.....	42
8.5.3 WEE1.....	44
9. Discusión general.....	45
10. Conclusiones.....	50
11. Perspectivas	50
12. Referencias.....	51

Abreviaturas

ATM	Ataxia- telangiectasia mutated
ATR	Ataxia- telangiectasia mutated rad-3 related
MRE11	Meiotic recombination 11
NBS1	Nijmegen breakage syndrome
PCD	Muerte celular programada
DDR	Respuesta a daño a DNA
HR	Recombinación homóloga
NHEJ	Unión de extremos no homólogos
SSB	Ruptura de cadena simple
DSB	Ruptura de doble cadena
BRCA1	Breast cancer susceptibility 1
BRCA2	Breast cancer susceptibility 1
MRN	MRN-NBS1-RAD50
SOG1	Suppressor of gamma response 1
WEE1	Wee1 kinase
RPA1	Replication protein A 1
ATRIP	ATR interacting protein
9-1-1	RAD9-RAD1-HUS
ROS	Especies reactivas de oxígeno
IR	Radiación ionizante

Nota: Debido a que toda la bibliografía consultada está en inglés, los acrónimos se pusieron en ese idioma, pero su significado se escribió en español, con excepción de los nombres de proteínas.

Resumen

Al ser organismos sésiles, las plantas son susceptibles a factores que pueden generar varios tipos de daño a su DNA, siendo la ruptura de doble cadena el más severo. En la respuesta a daño a DNA, participan diversas proteínas en el reconocimiento, señalización, activación de puntos de control del ciclo celular y la reparación del daño. Aunque las DSB se reparan por las vías de Recombinación Homóloga o Unión de Extremos no Homólogos, comparten factores con otras vías en los diferentes puntos de la DDR, habiendo proteínas con diferentes grados de participación. Si bien, en general los mecanismos de reparación de daño a DNA están altamente conservados entre eucariontes, existen variaciones que van desde la ausencia o presencia de factores específicos, hasta diferencias estructurales entre proteínas presentes en todos los taxa. La alta plasticidad de las plantas hace relevante el estudio de especificidades que expliquen en cierto grado su gran respuesta ante estímulos estresores. Si bien, existen estudios comparativos de algunas proteínas que participan en la reparación de DSB, no hay estudios en conjunto. En el presente trabajo se realizaron árboles filogenéticos y análisis de comparación de motivos de las proteínas que participan en las vías de reparación de daño al DNA de diferentes organismos representantes del grupo Viridiplantae . Se encontró que la mayoría de las proteínas están altamente conservadas en cuanto a presencia o ausencia de ortólogos en los organismos de Viridiplantae; no obstante, existe variación en las secuencias de aminoácidos, por lo que existen motivos que son específicos de organismos evolutivamente más cercanos. Así pues, la mayoría de los motivos, tanto específicos como conservados, se encuentran dentro de los dominios reportados de las proteínas. Diferencias mayores se observaron entre los motivos de las algas y las plantas terrestres; además de la ausencia de proteínas en Chlorophyta.

Introducción

Las plantas cuentan con una alta plasticidad fenotípica y dado que son organismos sésiles, presentan una gran cantidad de adaptaciones para responder a cambios en su ambiente (Arnold, Kruuk & Nicotra, 2019; Gratani, 2014). Estos cambios incluyen estímulos estresores, definidos como aquellos que pueden afectar el correcto desarrollo de los organismos y que se dividen en bióticos o abióticos (Gull et al., 2019). Dentro de estos últimos se encuentran la sequía, el frío, altas temperaturas, exposición a metales pesados, salinidad y radiación UV (Mareri, Parrotta & Cai, 2022). Algunos de estos estresores pueden ocasionar diversos daños a componentes celulares, entre ellos el DNA, generando lo que se conoce como estrés genotóxico (Spampinato, 2017). Este tipo de estrés también suele ocurrir por factores intrínsecos propios del desarrollo de las plantas (Ossowski et al., 2010; Schmidt, Pacher & Puchta; 2019).

En las plantas el daño al DNA generado por el estrés genotóxico tiene menor efecto en comparación con el daño que se genera en animales, esto es debido a las características fisiológicas propias de las plantas, que, a diferencia de los animales, son capaces de regenerar diferentes órganos y tejidos con gran facilidad; además de contar con metabolitos que disminuyen los efectos de estresores como la radiación UV (Kimura & Sakaguchi, 2006; Liu et al., 2023).

Los mecanismos que utilizan las plantas en respuesta a estrés genotóxico generan cambios genéticos y fisiológicos para evitar, por una parte, una propagación masiva del daño y en consecuencia permitir su adaptación (Burke et al., 2020; Macovei et al., 2016). A pesar de que los mecanismos están altamente conservados en las plantas, estos pueden tener variaciones dependiendo del grupo del que se hable (Čížková et al., 2019, Singh et al., 2010; Vágnerová et al., 2017). Para conocer las causas de la variación en los diferentes mecanismos de respuesta a estrés genotóxico en las plantas es importante analizarlos tanto en las algas, como en las plantas terrestres que comprenden el clado Viridiplantae.

1 El clado Viridiplantae

Debido a la gran diversidad de organismos eucariontes fotosintéticos y sus diferencias con las plantas terrestres (Embryophyta) existe debate acerca de la definición de plantas. Una de las clasificaciones de plantas es en su sentido más amplio, el cual las toma como aquellas pertenecientes al grupo monofilético Archaeplastida (Bowles et al., 2023; Delwiche & Time, 2011); no obstante, algunos autores usan una clasificación de plantas que incluye a las embriofitas y a las algas más cercanas a ellas: Viridiplantae (Bowles et al., 2023; One Thousand Plant Transcriptomes Initiative, 2019).

Viridiplantae es uno de los tres clados pertenecientes a Archaeplastida (siendo Glaucophyta y Rhodophyta los otros dos), grupo formado por la interacción endosimbiótica entre una cianobacteria y un organismo eucarionte (Bowles et al., 2023). Viridiplantae comprende a todas las plantas verdes y está dividido en dos grandes grupos: Chlorophyta (algas verdes), y Streptophyta, que incluye a las plantas terrestres (Embryophyta) y las algas carofitas (Charophyta) (Bowles et al., 2023; Li et al., 2020). Así pues, recientemente se propuso la división Prasinodermophyta, como un tercer subgrupo de Viridiplantae, el cual está formado por dos clases (Prasinodermophyceae y Palmophyllophyceae) y surgió antes de la divergencia entre Streptophyta y Chlorophyta (Li et al., 2020).

Se estima que el clado Viridiplantae surgió entre el Mesoproterozoico medio y el Neoproterozoico (Bowles et al., 2023) y contiene más de 450 mil especies (Bowles et al., 2023; Corlett, 2016) las cuales presentan innovaciones biológicas tales como el uso de la clorofila b, la síntesis de almidón intraplastidial, paredes celulares complejas, etc.. Esto, en conjunto con otras adaptaciones, permitieron la colonización de una gran diversidad de hábitats, entre ellos, el terrestre (Bowles et al., 2023; One Thousand Plant Transcriptomes Initiative., 2019).

2 Estrés genotóxico en plantas

2.1 Daño al DNA en plantas

Diversos factores pueden generar daño al DNA de los organismos, estos incluyen factores endógenos y exógenos, así como bióticos y abióticos. Los agentes que provocan lesiones al DNA se denominan genotóxicos y pueden repercutir en el desarrollo normal y supervivencia de los seres vivos (Giglia-Mari, Zotter & Vermeulen, 2011; Kimura y Sakaguchi, 2006). Las plantas son especialmente susceptibles a estos agentes, por lo que deben contar con mecanismos eficaces para responder a condiciones cambiantes en su ambiente y mantener la integridad de su genoma (Manova & Gruszka, 2015). Estos mecanismos en conjunto se denominan Respuesta a Daño a DNA (DDR, *DNA Damage Response*) e incluyen el reconocimiento del daño, la iniciación de cascadas de señalización, la entrada a puntos de verificación (*checkpoints*) del ciclo celular, la reparación del daño y, en caso de que esta última falle, la transición a la endorreduplicación o en última instancia la muerte celular programada (PCD, *Programmed Cell Death*) (DiRuggiero & Robb, 2007; Lehmann & Taylor, 2003).

Es importante mencionar que el mecanismo de reparación empleado depende del tipo de daño que se presente, ya que hay diversos tipos de daño al DNA. Las lesiones al DNA incluyen: cambios químicos o anomalías en los nucleótidos, que pueden ser provocados por agentes alquilantes que metilan las bases o por la oxidación de éstas; por aductos de DNA; entrecruzamientos intracatenarios entre guaninas adyacentes; enlaces covalentes entre pirimidinas adyacentes; enlaces cruzados entre hebras o rupturas de los enlaces fosfodiéster (**Figura 1**) (Giglia-Mari et al., 2011; Manova & Gruszka, 2015; Szurman-Zubrzycka, Jędrzejek & Szarejko, 2023). La ruptura del enlace fosfodiéster se puede presentar en una sola hebra del DNA (SSB, *Single strand break*) o en ambas (DSB, *Double strand break*), siendo esta última el tipo de lesión más grave dado que puede conllevar a la pérdida de información cromosómica (Bétermier et al., 2019).

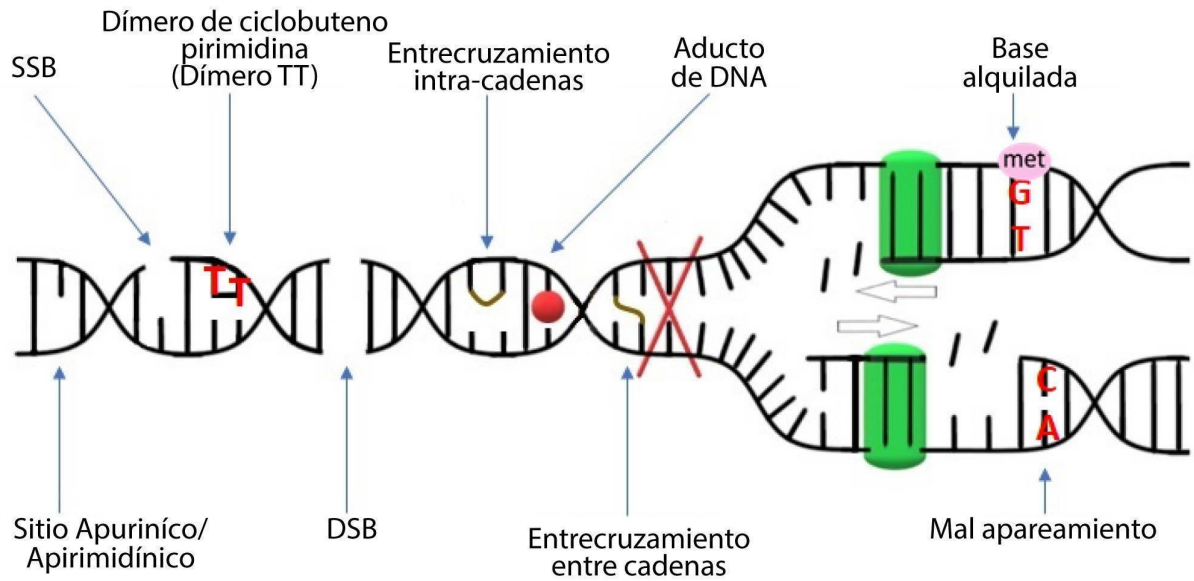


Figura 1. Tipos de Daño al DNA. La “X” roja representa bloqueos en la horquilla de replicación; los rectángulos verdes representan a la polimerasa llevando a cabo la replicación. SSB: Single Strand Break; DSB: Double Strand Break. Imagen modificada de Szurman-Zubrzycka, Jędrzejek & Szarejko (2023).

2.1.1 Señalización de la respuesta a Daño a DNA

Tanto para la ruptura de una cadena, como de ambas, la reparación del daño involucra una cascada de señalización (**Figura 2**), en donde participan las principales proteínas señalizadoras de la respuesta a daño ATM RAD-3 RELATED (ATR) y ATAXIA TELANGIECTASIA MUTATED (ATM), para SSBs y DSBs, respectivamente. Esta señalización converge en la activación del factor de transcripción SUPPRESSOR OF GAMMA RADIATION (SOG1) (Culligan et al., 2006; Nisa et al., 2019) seguida de una baja en la actividad de las polimerasas, detención del ciclo celular, reconocimiento del daño y activación de proteínas de señalización (Nisa et al., 2019; Saldivar, Cortez & Cimprich, 2017).

1.1.2 Daño al DNA por Single Strand Break (SSB)

La ruptura de una sola cadena de DNA (SSB), es un tipo de lesión que implica la ruptura de un enlace fosfodiéster entre un nucleótido y otro. Esta lesión se produce con más frecuencia que la ruptura de ambas cadenas y los agentes que la generan pueden ir desde especies reactivas de Oxígeno (ROS, *Reactive Oxygen Species*), propias del metabolismo celular y errores en la maquinaria de replicación,

hasta factores externos como la radiación UV-B y compuestos químicos como la Hidroxiurea (HU) (Caldecott, 2008; Hu, Cools & De Veylder, 2016; Macovei et al., 2016).

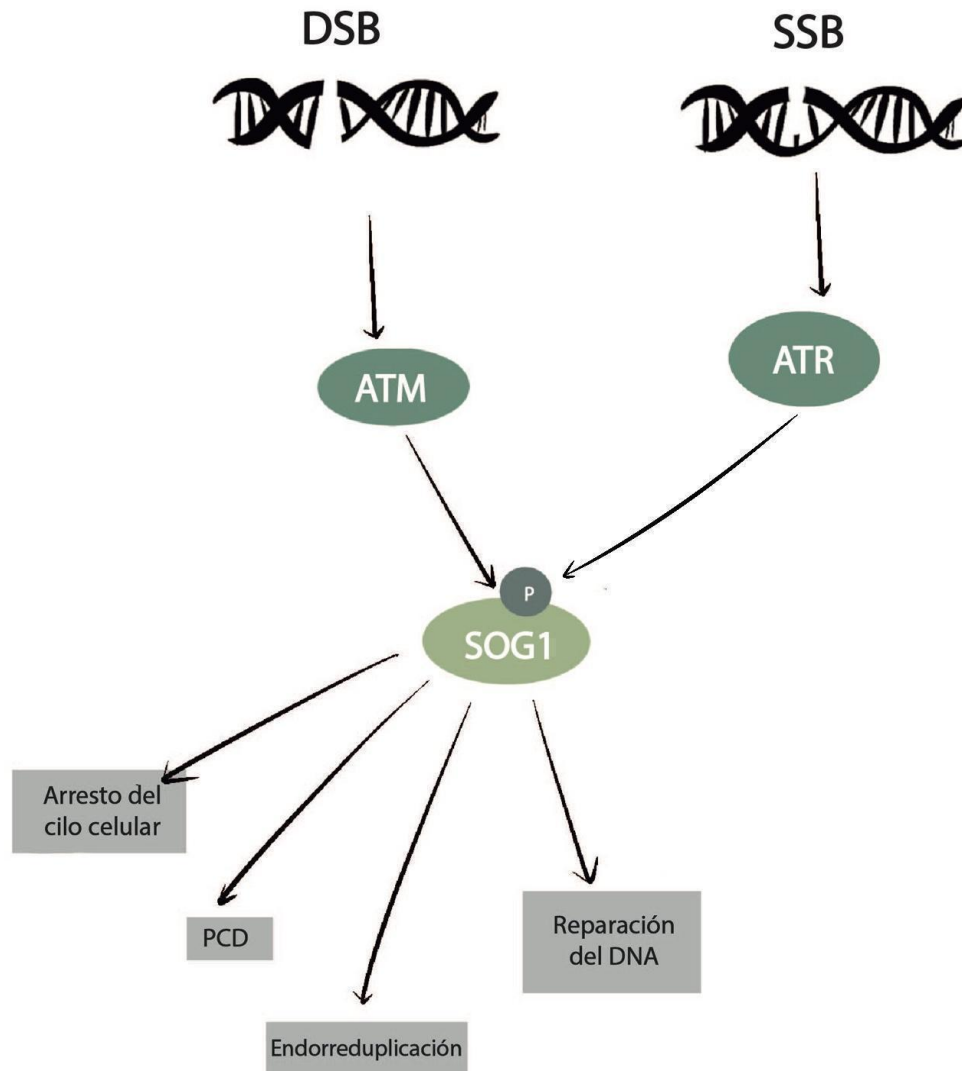


Figura 2. Principales componentes de la señalización de las rutas de reparación a SSB y/o DSB y los procesos que dirigen. Modificado de Nisa et al., (2019).

Cabe destacar que las SSBs ocasionalmente pueden llevar a la inducción de cambios genómicos como deleciones, inserciones o secuencias repetidas en tándem, si se encuentran varias rupturas a determinadas distancias entre ellas, lo cual podría tener un papel importante en la evolución del genoma de las plantas (Schiml, Fauser & Puchta, 2016). A su vez, cuando la lesión no es reparada antes de la replicación del DNA, se genera una DSB (Amiard et al., 2013).

Una vez que se produce la ruptura de la cadena, diversos factores participan en el reconocimiento del daño (**Figura 3**). El extremo expuesto del DNA es cubierto por RPA, el cual recluta a la proteína ATRIP (ATR INTERACTING PROTEIN) que a su vez, ayuda al reclutamiento de ATR (Aklilu & Culligan, 2016; Nisa et al., 2019; Szurman-Zubrzycka, Jędrzejek & Szarejko, 2023). Al mismo tiempo, el complejo 9-1-1 (conformado por RAD9, RAD1 y HUS1) es reclutado al sitio de daño para activar a ATR (Nisa et al., 2019).

Si bien, generalmente las SSBs se reparan por simple ligación, diversos estudios en plantas y animales han reportado la participación, aunque con cierta variación, de las rutas HR (*Homologous Recombination*) y NHEJ (*Non-homologous end joining*), dos vías para la respuesta a DSB (la ruptura de ambos enlaces fosfodiéster), durante los cambios genómicos producidos por esta lesión (Davis & Maizels, 2014; Fauser, Smichl & Puchta, 2014; Nakajima et al., 2018; Wolter et al., 2021).

2.1.3 Daño al DNA por Double Strand Break (DSB)

La DSB consiste en la ruptura de los enlaces fosfodiéster en ambas hebras de DNA y es el tipo de daño más severo, dado que puede llevar a la pérdida de información cromosómica (Bétermier et al., 2019). Esta se puede generar tanto por factores endógenos, como exógenos. Los primeros incluyen a las especies reactivas de oxígeno producidas por el metabolismo celular, fallas en los mecanismos de reparación de otros tipos de daño, errores en la replicación, escisión de elementos transponibles, entre otros; mientras que dentro de los factores exógenos se encuentra la radiación ionizante (IR, *Ionizing Radiation*) y diversos compuestos químicos, como la zeocina (Schröpfer et al., 2014; Nisa et al., 2019, Mehta & Haber, 2014; Waterworth et al., 2011).

Existen dos vías principales para la reparación de DSBs: la reparación no homóloga (NHEJ) y la reparación homóloga (HR) (**Figura 3**), que difieren entre ellas en la necesidad de una secuencia homóloga para reparar el sitio dañado, siendo NHEJ el mecanismo más propenso a generar mutaciones (Amiard et al., 2013;

Macovei et al., 2016). La decisión de qué vía elegir cuando se presenta una DSB depende del organismo, del tipo celular y la fase del ciclo celular en la que se encuentra el mismo (Amiard 2013, Gallego & White, 2013; Giglia-Mari, Zotter & Vermeulen, 2011).

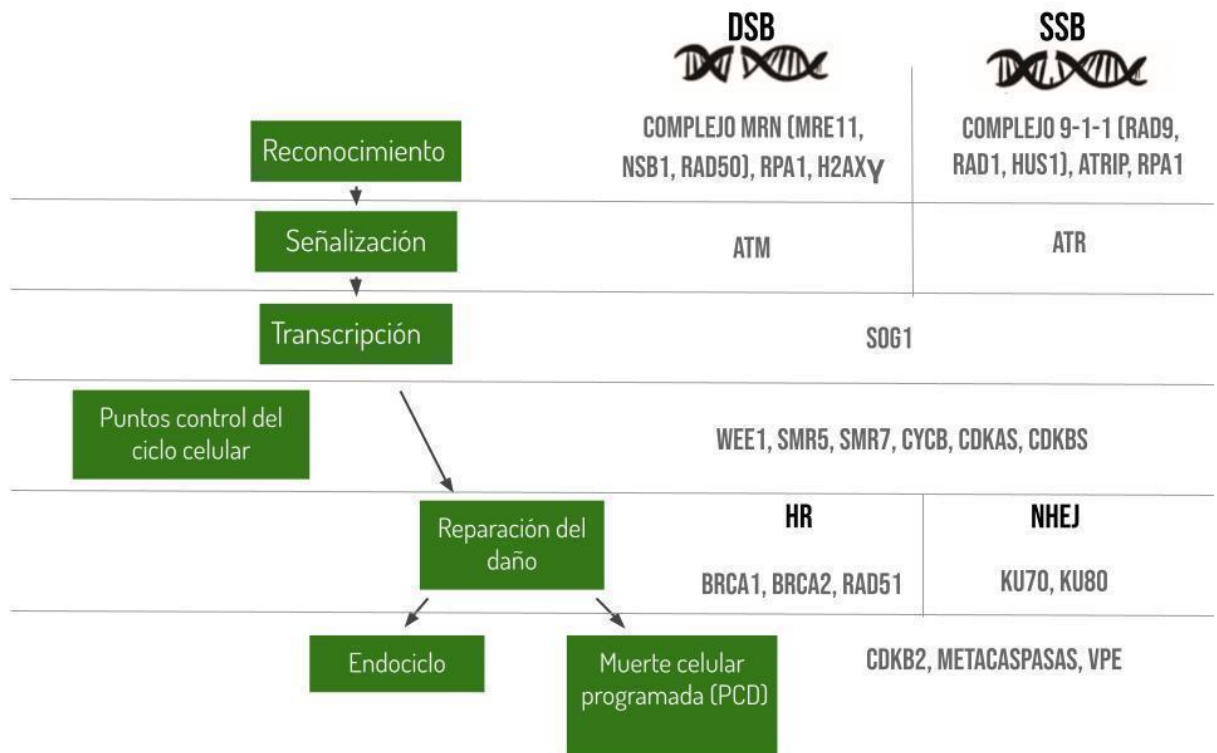


Figura 3. Procesos involucrados en las vías de reparación del daño por DSB y SSB, donde se muestran las principales proteínas involucradas en los diferentes puntos de la respuesta a daño.

2.2 Mecanismos de reparación de DSBs

2.2.1 Reparación Homóloga (HR)

La Reparación Homóloga (**Figura 3**) es el mecanismo predominantemente usado por organismos eucariontes unicelulares (Trapp, Seeliger & Puchta, 2011). Debido a que requiere de la presencia de una cromátida hermana homóloga, se presenta únicamente en las fases S y G2 del ciclo celular (Giglia-Mari, Zotter & Vermeulen2011; Howell, 2014).

Como se mencionó antes, una proteína clave para la señalización de la reparación homóloga es la proteína ATM, perteneciente a la familia PIKK

(PHOSPHOINOSITIDE 3-KINASE-RELATED KINASE) (**Figura 2**). Después de la doble ruptura, los extremos expuestos de ssDNA son cubiertos por el complejo trimérico de RPA (REPLICATION PROTEIN A), para ser protegidos de la degradación (Aklilu & Culligan, 2016). De igual forma, el complejo MRN (conformado por MRE11, RAD50 y NSB1) reconoce el sitio dañado y, por medio del dominio C-terminal de NSB1 (NIJMEGEN BREAKAGE SYNDROME 1), recluta a ATM, a estos sitios provocando que se autofosforile y active (Nisa et al., 2019). A su vez, ATM se encarga de la fosforilación de múltiples proteínas importantes para la reparación del daño a DNA. Entre ellas, la fosforilación de la variante de histona H2AX gamma (γ H2AX) en el sitio dañado, que sirve como señal para el reclutamiento de proteínas de reparación (Waterworth, 2011) (**Figura 3**). Cabe destacar que ésta señalización también es llevada a cabo por ATR, la otra proteína PIKK, y de hecho en *Arabidopsis thaliana* se ha reportado que aproximadamente el 10% de la fosforilación de la histona H2AX depende de ATR (Friesner et al., 2005).

De igual forma, ATM se encarga de la fosforilación y activación consecuente de la proteína SOG1, después de la inducción del daño (Yoshiyama, Sakaguchi & Kimura, 2013). SOG1 es un factor de transcripción perteneciente a la familia de proteínas NAC (NAM, ATAF1/2 y CUC2), cuya función ocurre predominantemente en células en proliferación activa y que regula la transcripción de múltiples genes relacionados con el arresto celular dependiente de daño a DNA, la reparación, la entrada al endociclo y a la muerte celular programada (PCD) (Mahapatra & Roy, 2020; Waterworth, 2011; Yoshiyama, 2016). Ogita et al. (2018) identificaron 146 genes directamente regulados por SOG1 durante la inducción de DSBs, entre los cuales destacan: factores de transcripción; reguladores negativos del ciclo celular como KRP6, SMR4, SMR5, SMR7 y WEE1; y una gran cantidad de genes relacionados con la reparación homóloga, entre ellos RFC, RPA, RAD51, BRCA1 y RAD17.

El arresto del ciclo celular puede ocurrir por diferentes vías, donde destaca la participación de SOG1 regulando directa o indirecta la actividad de algunos miembros de las familias KRP/ICK y SIM/SMR, inhibidores de complejos CDK/CYC (CIK) encargados del arresto en G1 y G2, respectivamente, así como de otras proteínas NAC involucradas en la regulación de Rep-MYB, responsable del arresto

en G2/M; o la proteína WEE1, que controla el arresto celular en G2/M a través de varios procesos de forma dependiente de ATM/SOG1 y/o ATR/SOG 1 (Culligan, Tissier & Britt, 2004; De Schutter et al., 2007; Gentric, Genschik & Noir, 2021; Mahapatra & Roy, 2020).

Cuando se activan los procesos de reparación, la proteína BREAST CANCER SUSCEPTIBILITY 1 (BRCA1) juega un papel clave. En animales, esta proteína se encuentra en forma de heterodímero junto con BARD1 y en 3 complejos distintos: A, B y C; los cuales varían dependiendo de las otras proteínas asociadas, siendo el primero el único involucrado en reparación homóloga (Trapp, Seeliger & Puchta, 2011). Cabe destacar que ortólogos de estas proteínas han sido reportados en plantas (Trapp, Seeliger & Puchta, 2011; Pfeffer, Ho & Singh, 2017). El complejo BRCA1 A, que incluye a la proteína BRCA2 (BREAST CANCER SUSCEPTIBILITY 2), recluta a RAD51 y permite su posterior intercambio con RPA1 en el sitio dañado (Pfeffer, Ho. & Singh, 2017; Seeliger et al., 2012), para que ésta, junto con otras proteínas RAD51-like, lleven a cabo la búsqueda de homología y el apareamiento homólogo (Kobayashi et al., 2019; Waterworth et al., 2011).

2.2.2 Reparación no homóloga (NHEJ)

La reparación de extremos no homólogos (NHEJ) se lleva a cabo principalmente en la fase G1 del ciclo celular (Giglia-Mari, Zotter & Vermeulen, 2011) y se puede dividir en dos procesos diferentes: la ruta canónica y la alternativa (aNHEJ o MMEJ) (Amiard, 2013) (**Figura 3**). La ruta canónica depende de la unión del heterodímero ku70/ku80 a los extremos del DNA, para posteriormente ser ligados directamente por la ligasa 4 y X-ray Repair Cross Complementing 4 (XRCC4) (Howell, 2014). Por otra parte, la ruta alternativa (aNHEJ) o mediada por micro homologías (*Microhomology mediated end joining*, MMEJ) es independiente del heterodímero ku70/ku80. En ésta, las proteínas tipo poly-ADP ribose (PARP) se unen al sitio dañado y reclutan al complejo MRN, el cual interactúa con el complejo formado por la ligasa 3 y XRCC1 (Que et al., 2019).

En caso de que la reparación resulte fallida, como mecanismo para evitar la propagación del daño, las células somáticas entrarán en endoreduplicación, que

implica pasar por la replicación del DNA varias veces, sin que exista una división mitótica (Adachi et al., 2011); mientras que las células troncales llevarán a cabo la muerte celular programada (Furukawa et al., 2010).

2.2.3 Muerte celular programada (PCD)

La muerte celular programada (PCD) juega un papel importante en procesos normales del desarrollo y senescencia, pero también puede ser activada como último mecanismo para proteger al organismo de células dañadas a causa de factores estresantes bióticos o abióticos (Burke et al., 2020). En animales existen tres grandes tipos de muerte celular, definidos por sus patrones morfológicos y moleculares: Apoptosis, Autofagia y Necrosis (Kroemer et al. 2009). En plantas, la clasificación morfológica de los tipos de muerte celular no está del todo clara (Rantong & Gunawardena, 2015). Más aún, los principales reguladores de la apoptosis en animales no están conservados en plantas, por lo que en estas últimas se habla de PCD, la cual puede dividirse en aquella provocada por factores internos y la estimulada por factores externos (Locato & De Gara, 2018; Rantong & Gunawardena, 2015).

En animales, el factor de transcripción p53 se encuentra río arriba de la cascada de señalización para la entrada a apoptosis, regulando la expresión de múltiples genes, entre los que destacan las proteínas caspasas. Si bien, no se han encontrado secuencias ortólogas para estas proteínas en plantas, el factor de transcripción SOG1 presenta funciones similares a p53 (Yoshiyama et al., 2014). De igual forma, se han encontrado distintas familias de proteasas que podrían tener las mismas funciones que las caspasas o que estarían involucradas en la vía de PCD, entre ellas: metacaspasas, la familia VPE (*Vacuolar Processing Enzyme*), y subtilasas (Burke et al., 2020; Kabagge et al., 2017; Rantong & Gunawardena, 2015). Aunque se han identificado algunos genes relacionados con PCD durante el desarrollo, hace falta información para elucidar los mecanismos moleculares relacionados con cada uno de los diferentes tipos de muerte celular (Olvera-Carrillo et al., 2015).

En células troncales, el daño a DNA y el fallo en el proceso de reparación induce la entrada a PCD como mecanismo para evitar la propagación del daño a células hijas, lo cual es de especial importancia ante DSB (Fulcher & Sablowsky, 2009). Se ha observado que tanto la exposición a UV-B, como a radiación ionizante, que producen SSBs y DSBs respectivamente, activan la muerte celular programada orquestada por SOG1 de manera dependiente de ATM y/o ATR (Furukawa et al., 2010). Si bien, se han reportado algunos de los genes blanco de SOG1 (Ogita et al., 2018), se necesitan más estudios para reconocer aquellos involucrados en los procesos de muerte celular.

3 Evolución de los mecanismos de reparación a DSBs

3.1 Variación en las rutas de reparación

Proteger el genoma ante distintos daños es fundamental para la supervivencia e integridad de los seres vivos, por lo que la vía de reparación de daño a DNA está presente tanto en procariontes como en eucariontes (**Tabla 1**). No obstante, el nivel de conservación depende de las rutas específicas a cierto daño y de cada taxón. Se ha observado que todas las rutas de reparación están conservadas en bacterias y arqueas, aunque no se han reportado ortólogos para todos los genes descritos en animales y levaduras. Para la vía de HR, se han encontrado genes de arqueas como: RADA, homólogo de RecA/RAD51; genes del complejo MRE11/RAD51; NurA, tercer componente del complejo MRN y RADB, posible homólogo funcional del complejo RAD51/RAD55 (DiRuiuggiero & Robb, 2007).

Por otra parte, para los organismos eucariontes existen respuestas y elecciones de ruta de reparación diferentes entre aquellos pluricelulares y unicelulares. Ante DSB la ruta de reparación comúnmente usada por eucariontes unicelulares es HR; mientras que ésta es menos usada por los organismos pluricelulares (Schmidt, Pacher & Puchta, 2019; Trapp, Seeliger & Puchta, 2011).

Aunque las rutas de reparación en eucariontes pluricelulares están altamente conservadas, es importante destacar diferencias entre animales y plantas, tales

como ausencia de ciertos ortólogos, genes únicos de un taxón, número de copias distintas de cada uno de los genes y diferencias en las secuencias o dominios (Spampinato, 2017; Eisen & Hanawalt, 1999; Yoshiyama et al., 2014). Más aún, Singh et al. (2010) reportaron que la ruta de NHEJ es una de las menos conservadas, de acuerdo con la presencia o ausencia de los genes, entre *Homo sapiens* y plantas. Otra diferencia para destacar es la plasticidad de las plantas, ya que una gran variedad de mutantes de pérdida de función que en animales tienen fenotipos letales, en plantas no son letales, entre ellos: RAD50, RAD51, MRE11, BRCA1 y BARD1 (Singh et al., 2010).

Por otra parte, en animales la actividad de las cinasas ATM y ATR converge en la activación del factor de transcripción p53; el cual regula múltiples genes, entre ellos las proteínas caspasas; mientras tanto, ATM y ATR también regulan la actividad de las cinasas CHK1 y CHK2, involucradas en el arresto celular. Si bien esta vía de señalización está conservada en plantas, a excepción de ATM y ATR, no se han encontrado ortólogos de estas proteínas en plantas, aunque en estas se presenta el factor de transcripción SOG1, que tiene funciones similares a p53 (Spampinato, 2017, Yoshiyama et al., 2013).

3.2 Variación en las rutas de reparación en plantas

Dentro de plantas y algas también se han observado variaciones en las respuestas a daño a DNA. Por una parte, la variación se puede dar en el fenotipo observado. Un ejemplo de esto es la respuesta de *Chlamydomonas reinhardtii* que, al contrario de las plantas terrestres, anula el *checkpoint* de replicación cuando se combinan cafeína con zeocina, un agente genotóxico; lo que podría indicar la presencia de otra ruta de reparación (Čížková et al., 2019). De igual forma, Vágnerová y colaboradores (2017), compararon las respuestas ante DSB en algas verdes carófitas de los géneros *Zygnema* y *Klebsormidium* con *Physcomitrella patens*, reportando una mayor tasa de reparación y menor daño a DNA en *Zygnema*.

Otra forma en que la variación se refleja es en el número de copias o parálogos de un gen en cada organismo. Se sabe que las plantas han pasado por múltiples duplicaciones del genoma completo en diferentes tiempos y clados, además de

procesos loci específicos, como pérdida de genes (Jiao et al., 2011; Proost et al., 2011; Tang et al., 2010). Así pues, se ha observado que el número de copias de ciertos genes influye en las adaptaciones y formas de vida de un organismo; por ejemplo, la asociación positiva entre el número de copias de proteínas PARP, involucradas en la reparación de SSBs, y la longevidad de las plantas (Aoyagi, Kusumi & Satake, 2021). De igual forma, se ha reportado que genes importantes para funciones vitales, como los que forman parte de la respuesta a daño a DNA, muestran “una resistencia a la duplicación”, manteniéndose como copias únicas a lo largo de la evolución de las plantas (De Smet et al., 2013). Singh y colaboradores (2010) compararon diversas rutas de reparación de daño entre *A. thaliana* y *Oryza sativa*; si bien, en general éstas se encuentran altamente conservadas, se mostró que casi todas las proteínas involucradas en la reparación homóloga tienen el mismo número de copias en ambos organismos, exceptuando a MRE11 (MEIOTIC RECOMBINATION 11).

Por último, la variación se ve reflejada en el grado de conservación de la secuencia y dominios proteicos. Si bien, los dominios de las proteínas que constituyen el complejo MRN se encuentran altamente conservados en diferentes grupos, existen diferencias importantes en las secuencias. Así mismo, las proteínas RAD50 y MRE11 están presentes en los tres dominios (Archaea, Bacteria y Eukarya) y tienen porcentajes de similitud, entre plantas, mayores al 80% (De Bustos, Pérez & Jouve, 2007; Pérez et al., 2011); el tercer componente del complejo, Nsb1, sólo se encuentra en eucariotas y está menos conservado, teniendo un porcentaje de similitud entre especies de trigo y *Arabidopsis thaliana* de aprox. 34% (Pérez et al., 2011).

Para el caso de las PIKK, se ha observado que en general, ATM presenta mayor grado de conservación que ATR (Martens et al., 2020), pero ambas se encuentran en algas y plantas. En contraste, el factor de transcripción SOG 1 se ha caracterizado sólo en angiospermas; aunque se han encontrado proteínas similares que conservan el dominio NAC, así como algunos motivos importantes, en gimnospermas y briofitas (Yoshiyama et al., 2014). Por otra parte, se observa mayor variación en RPA1, ya que algunas plantas tienen deleciones de los dominios DBD-F y DBD-C en algunos de sus ortólogos. Además, los miembros de la familia

Brassicaceae tienen dos ortólogos más (RPA1D y E) (Aklilu & Culligan, 2016). Otra proteína donde se ve una diferencia en el número de ortólogos es RAD51, ya que *P. patens* destaca por ser el único organismo con dos copias (Markmann-Mulisch, 2002).

Tabla 1. Proteínas involucradas en la reparación de SSB y DSB (por HR y NHEJ) en diferentes especies (Ayora et al., 2011; Chintapalli, et al., 2013; Connelly & Leach, 1996; Doherty, Jackson & Welley, 2011; Hefferin & Tomkinson, 2005; Lehmann & Taylor, 2003; Levin, 2020; Pfeffer, Ho & Singh, 2017; Shaik & Kirubakaran, 2020)

	<i>E. coli</i>	<i>S. cerevisiae</i>	<i>H. sapiens</i>	<i>A. thaliana</i>
HR	-----	Mec1p	HsATR	AtATR
	-----	Tel1p	HsATM	AtATM
	-----	-----	p53	-----
	-----	-----	-----	SOG1
	SbcC	Rad50	HsRAD50	AtRAD50
	SbcD	Mre11	HsMRE11	AtMRE11
	-----	Xrs2	HsNBS1	AtNBS1
	-----	-----	HsBRCA1	AtBRCA1
	-----	-----	HsBRCA2	AtBRCA2
	RecA	Rad51p	HsRAD51	AtRAD51
Reparación de SSB	-----	Ddc1p	HsRAD9	AtRAD9
	-----	Rad17p	HsRAD1	AtRAD1
	-----	Mec3p	HsHUS1	AtHUS1
	-----	Rad24p	HsRAD17	AtRAD17
NHEJ	ku	Hdf1p	HsKu70	AtKu70
	ku	Hdf2p	HsKu80	AtKu80

4 Planteamiento del problema

Hace falta mucha información acerca de las diferencias moleculares entre la respuesta a daño de los diferentes clados de plantas. Si bien, existen estudios filogenéticos o comparativos de pocas proteínas, no se han analizado en conjunto. Por lo tanto, en este trabajo se realizará un análisis de secuencias de las proteínas que participan en los procesos de reparación de daño al DNA de diferentes representantes del grupo Viridiplantae para evaluar diferencias que puedan tener un efecto en sitios importantes para la función de las proteínas.

5 Hipótesis

Dada la distancia filogenética y la variación en la respuesta a factores que provocan daño de doble cadena, las proteínas de la vía de respuesta a estrés genotóxico tendrán motivos particulares de cada grupo dentro de Viridiplantae.

6 Objetivo General

Analizar las diferencias en los motivos de las proteínas que participan en la vía de señalización de respuesta a estrés genotóxico en miembros de Viridiplantae.

6.1 Objetivos particulares

- Analizar la conformación de la vía de señalización de respuesta a estrés genotóxico y obtener las secuencias de las proteínas que participan en esta vía en Viridiplantae.
- Analizar la conservación y ancestría de las proteínas obtenidas en el objetivo anterior (Ku80, Ku70, RAD51, ATRIP, RAD9, RAD1, SOG1, WEE1, NSB1, MRE11, RAD50, RPA1, BRCA1, BRCA2, ATM y ATR), en Viridiplantae.
- Identificar los motivos más conservados y aquellos que difieren de las proteínas que participan en la vía de señalización de respuesta a estrés genotóxico.

7 Materiales y Métodos

7.1 Identificación y caracterización de homólogos en Viridiplantae

Se buscaron y descargaron las secuencias de aminoácidos, en formato fasta, homólogas de las proteínas Ku80, Ku70, RAD51, ATRIP, RAD9, RAD1, SOG1, WEE1, NSB1, MRE11, RAD50, RPA1, BRCA1, BRCA2, ATM y ATR, en al menos una especie representante de diferentes clados de Viridiplantae (Chlorophyta y Streptophyta; dentro de este último, Bryophyta, Pteridophyta, Gimnospermae y Magnoliophyta) a través del National Center of Biotechnology Information (NCBI)/GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>).

Para los genes no caracterizados o anotados en algunas especies, se realizó una búsqueda de homología comparando la secuencia de aminoácidos de cada proteína en *Arabidopsis thaliana* con la colección entera de secuencias de proteínas no redundantes del portal de NCBI/GenBank, usando BLASTP (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>). Se tomaron aquellas secuencias con los valores de porcentaje de similitud más altos.

7.2 Construcción de árboles filogenéticos y búsqueda de motivos

Con las secuencias de cada proteína se realizaron alineamientos múltiples con el algoritmo de MUSCLE (MULTiple Sequence Comparison by Log- Expectation) a través del programa MEGA X (Molecular Evolutionary Genetics Analysis) (Kumar et al., 2018) versión 10.1.8. Para cada alineamiento, se construyeron árboles filogenéticos con el algoritmo de Neighbor Joining (NJ) (Saitou & Nei, 1987), a través de MEGA X versión 10.1.8, empleando el modelo Poisson, No. de replicaciones de Bootstrap=10000, Distribución Gamma y Delección parcial (Núñez-Resendiz et al., 2019).

Para identificar motivos conservados en las proteínas de los diferentes organismos, se usó el programa de MEME suite (Bailey et al., 2009). Los resultados de los análisis anteriores se presentaron en forma de imágenes, que incluyen los árboles filogenéticos y las secuencias de motivos arrojadas por MEME suite.

8 Resultados y Discusión

Una forma de estudiar las diferencias y similitudes entre las respuestas a daño a DNA de diferentes organismos es a través de analizar las principales proteínas involucradas en este proceso. Por ello, se eligieron proteínas de los procesos de reparación por SSB y DSB que incluyen la reparación NHEJ y HR (**Tabla 1; Figura 2 y 3**), de diferentes especies representativas del clado Viridiplantae, para estudiar diferentes aspectos de sus secuencias y dominios que permitieran hacer un análisis comparativo sobre los mecanismos de reparación de daño al DNA entre las especies seleccionadas del clado. Para el análisis de cada proteína seleccionada, se eligieron diferentes especies, en función de las secuencias reportadas o disponibles en las bases de datos. Cabe señalar que de la proteínas utilizadas, algunas ya estaba caracterizada su función y otras estaban solo anotadas en las bases de datos. En los casos que no se encontró una secuencia anotada, se eligió una proteína cuya secuencia tenía un e-value positivo (<0.05) y el mayor porcentaje de similitud con la proteína de *A. thaliana*.

Con las secuencias obtenidas se realizaron análisis de *Neighbor joining*, el cual es un algoritmo que sirve para realizar árboles filogenéticos a partir de matrices de distancia (Sato & Nei, 1987; Studier & Kepler, 1998); esto para determinar la similitud o divergencia de las proteínas entre las diferentes especies. Además, se buscaron motivos conservados, así como únicos entre las proteínas de las diferentes especies, usando el Software de MEME suite (Bailey et al., 2009). Mientras que los motivos son secuencias pequeñas conservadas en la evolución de los organismos; los dominios tienen secuencias más grandes, con una organización globular estable y están formados por varios motivos (Bailey et al., 2015; Calderón-Vélez, 2006). Cabe resaltar que este software reconoce motivos, no dominios, por lo que algunas secuencias carecen de motivos que caen dentro de dominios ya reportados.

Para todos los casos se usó la secuencia ortóloga de *Homo sapiens* (Hominidae) para enraizar el árbol, con excepción de SOG1, ATRIP y BRCA1. Los resultados del análisis realizado de cada una de las proteínas se describen a continuación y para una mejor comprensión se presenta una breve discusión de

algunas de las características más importantes a resaltar de manera individual, y al final, una discusión general.

Tabla 1. Características de las proteínas de la vía de respuesta a daño a DNA elegidas para este estudio. Los colores indican a que vía pertenece cada proteína; Rojo= NHEJ, Verde= HR y Naranja= SSB.

Proteína	Peso molecular (kDa)	Locus	GenBank ID
ATM	341.534	AT3G48190	Q9M3G7
ATR	302.433	AT5G40820	NP_198898.2
SOG1	50.288	AT1G25580	NP_564238.2
RAD51	37.304	AT5G20850	OAO95923.1
Complejo MRN	MRE11	80.771	OAO95209.1
	RAD50	152.813	NP_565733.1
	NBS1	60.077	ABA54896.1
BRCA1	103921.0	AT4G21070	NP_193839.4
BRCA	BRCA2A	130.737	NP_001328739.1
	BRCA2B	127.254	NP_195783.3
SUV2 (ATRIP)	72.805	AT5G45610	NP_199374.4
Complejo 9-1-1	RAD1	107.633	NP_198931.1
	RAD9	49.932	NP_001319478.1
Ku80	76.687	AT3G05480	NP_564520.1
Ku70	70.290	AT1G16970	NP_564012.1
WEE1	56.529	AT1G02970	NP_171796.1
RPA1	RPA1A	72.242	NP_178690.1
	RPA1B	67.288	NP_196419.1
	RPA1C	94.393	NP_199353.1
	RPA1D	70,263	NP_200908.1
	RPA1E	86.122	AT4G19130

8.1 Análisis de las proteínas que participan en Single Strand break (SSB)

A continuación se describirán los análisis de cada una de las proteínas involucradas en la vía de reparación de daños de una sola cadena (SSB).

8.1.1 Complejo 9-1-1

Uno de los primeros complejos en actuar durante la respuesta a SSB es el complejo 9-1-1, que está conformado por las proteínas RAD9, RAD1, HUS1 y participa en la activación de ATR (Nisa et al., 2019; Szurman-Zubrzycka, Jędrzejek & Szarejko, 2023). Para RAD9, de forma general, el cladograma se dividió en dos grupos: uno que comprende a las plantas terrestres y otro, a las algas (**Figura 4A**). Si bien, las algas *Trebouxia sp.* (Trebouxiaceae) y *Chlamydomonas reinhardtii* (Chlamydomonadaceae) se agruparon juntas, esta última presenta mayor distancia genética, comparada con las demás plantas. Esto concuerda con los motivos observados (**Figura 4B**), ya que *Chlamydomonas* comparte solamente tres motivos con *Trebouxia* y cuatro con las embriofitas. La secuencia de *C. reinhardtii* no está definida como RAD9, sin embargo, fue una secuencia con una similitud razonable para considerarla como una posible homóloga a RAD9, por lo que harían falta estudios para definir si realmente conserva su función. Por otra parte, dentro de las plantas terrestres la filogenia se separó entre plantas con semilla y plantas sin semilla, encontrando los motivos 10, 13 y 15 presentes únicamente en plantas con semilla (**Figura 4A**).

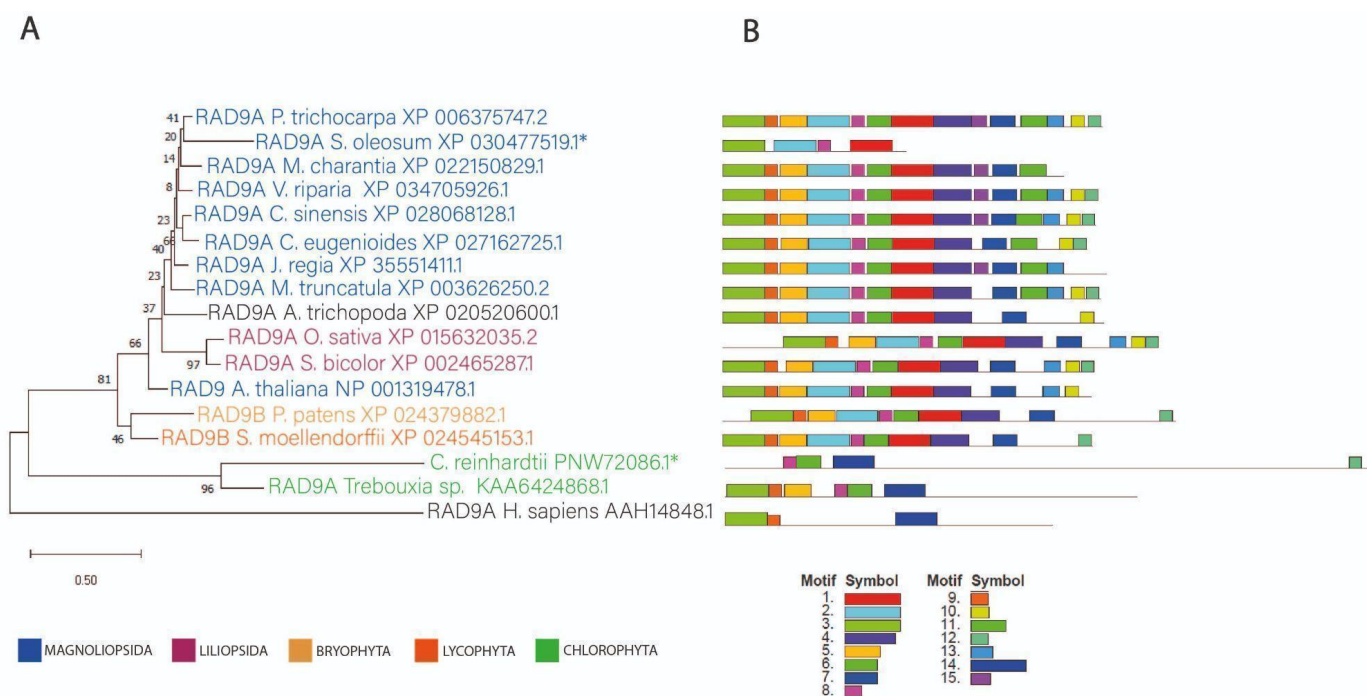


Figura 4. Cladograma de RAD9. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replications de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. El asterisco indica secuencias no anotadas como RAD9.

Por otro lado, RAD1 mostró una alta conservación de motivos en plantas con semilla, excluyendo a *Amborella trichopoda* (**Figura 5B**), la cual carece de todos los motivos encontrados en el extremo amino terminal. Las plantas sin semillas difieren de las plantas con semilla únicamente por la ausencia de los motivos 13 y 14; mientras que las secuencias de las algas carecen de varios motivos en el extremo C-terminal, pero guardan similitud con el extremo N-terminal de las plantas terrestres.

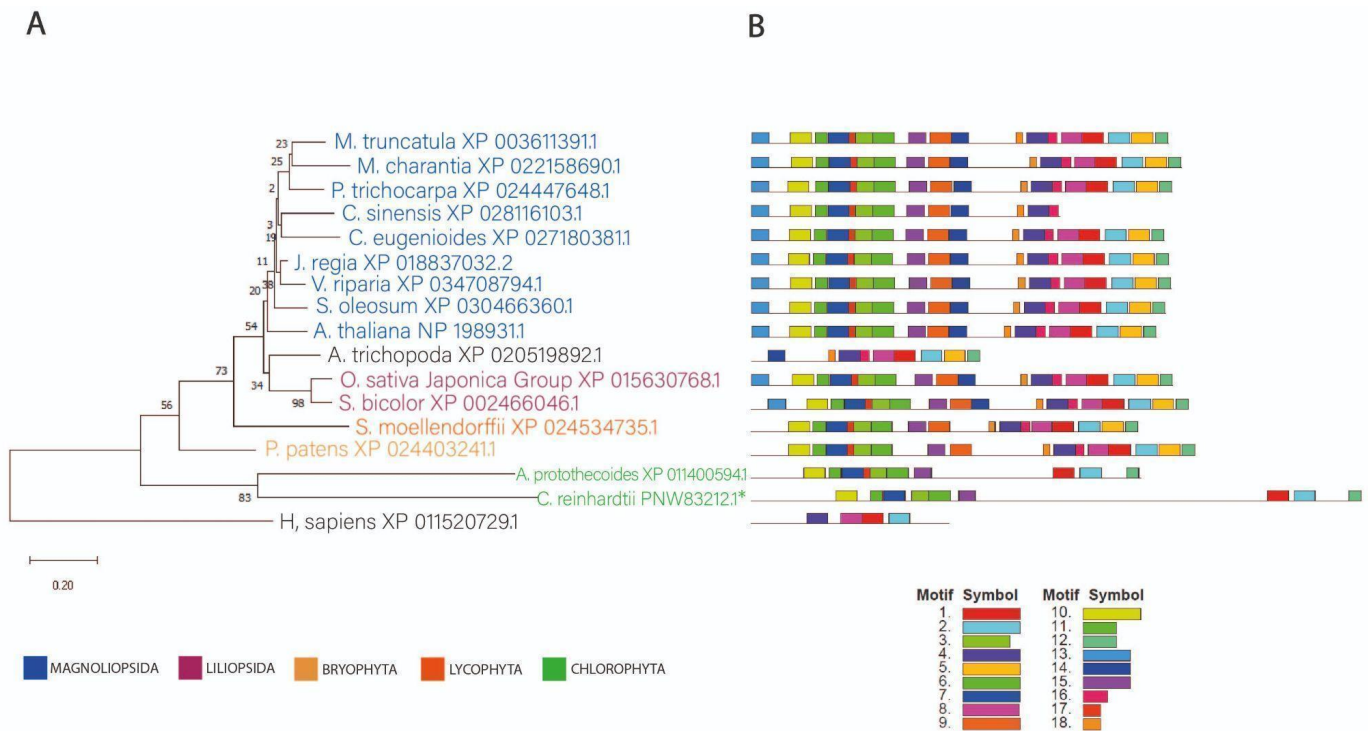


Figura 5. Cladograma de RAD1. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replications de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. El asterisco indica secuencias no anotadas como RAD1. *Auxenochlorella protothecoides*. (*A. protothecoides*).

8.1.2 Replication Protein 1 (RPA1)

En eventos de reparación del DNA, el complejo RPA, conformado por las subunidades RPA1, 2 y 3, se encarga de cubrir la cadena sencilla de DNA roto y reclutar a ATRIP y al complejo 9-1-1 (Aklilu & Culligan, 2016). Para realizar este cladograma, se tomaron los parálogos de RPA1 en algunas plantas (RPA1A-E) (**Figura 6A**). El árbol obtenido se divide en dos grupos: el primero contiene las secuencias de RPA1A, E y C; mientras que el segundo, las de RPA1B y D.

Por una parte, se puede observar que los parálogos se agrupan en clados definidos, a excepción de RPA1C, que se divide en dos grupos distintos, uno de ellos conteniendo las secuencias de las monocotiledóneas y otro, el de las algas. De forma inusual, este último también contiene a RPA1E de *Arabidopsis thaliana* (Brassicaceae). Por otra parte, resalta que RPA1E de *Coffea eugenioides* (Rubiaceae) presenta mayor distancia genética con respecto a los demás y queda fuera del clado de RPA1E.

Del lado de los motivos, se puede observar que las secuencias que se encuentran en el clado que incluye a RPA1, E y C comparten los mismos motivos. Esto a excepción de las proteínas de ambas especies de *Chlorella* (Chlorellaceae), las cuales carecen de un motivo en su extremo N-terminal (**Figura 6B**).

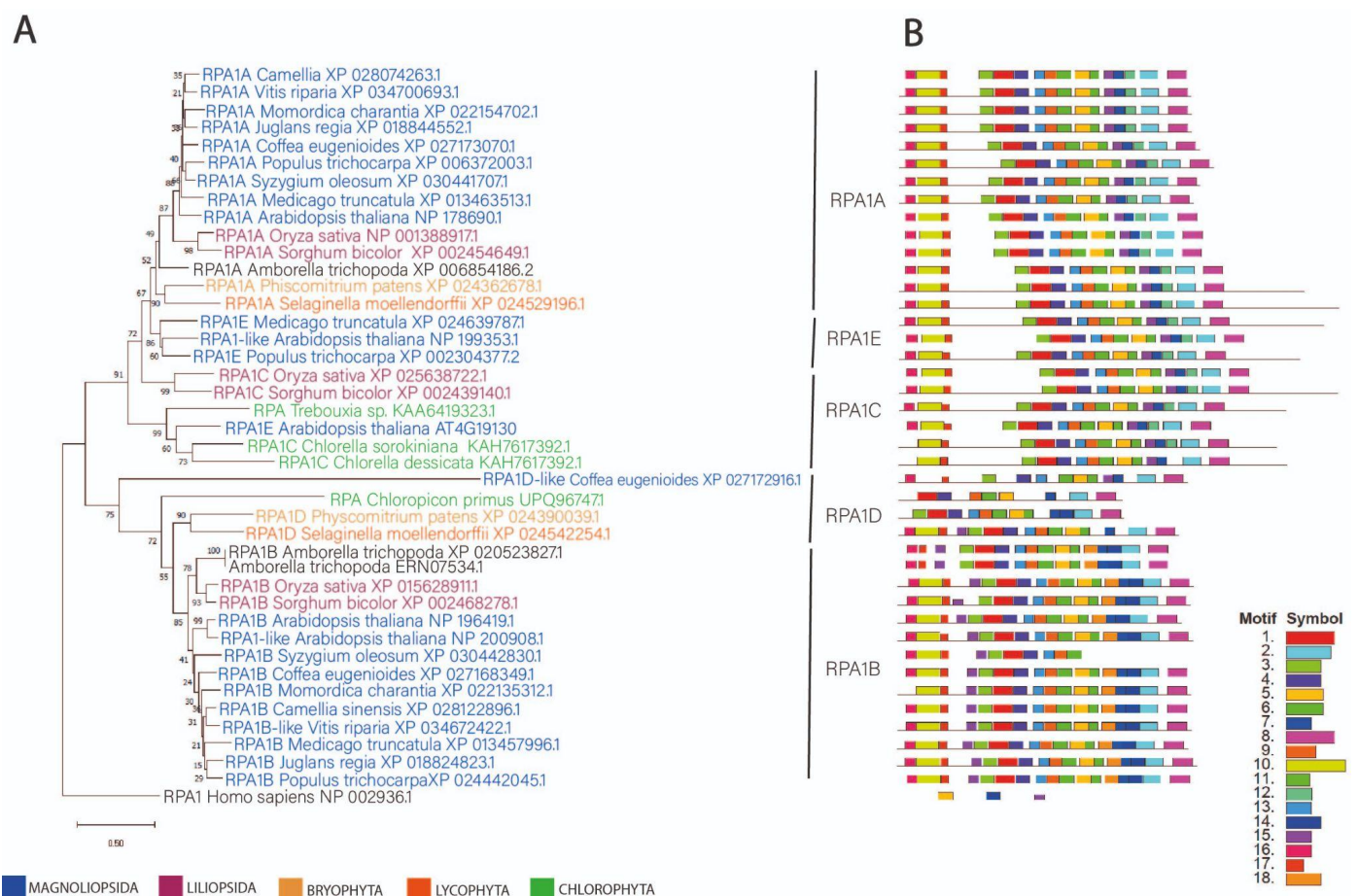


Figura 6. Cladograma de RPA1. A) Árbol de NJ con las secuencias obtenidas por BLAST; No. de replicaciones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias.

8.1.3 ATR interacting protein (ATRIP)

ATRIP (también llamada SUV2 en plantas) es una proteína que facilita el reclutamiento de ATR al sitio de daño (Nisa et al., 2019). Dentro de las plantas con semilla, la filogenia dividió en dos grupos a las monocotiledóneas y a las dicotiledóneas (**Figura 7A**). Así pues, la secuencia del musgo *Sphagnum magellanicum* (Sphagnaceae) mostró mayor distancia genética con respecto a las demás. Esta clasificación también se reflejó en los motivos encontrados en el

programa MEME (**Figura 7B**). Se encontraron dos motivos específicos de Dicotiledóneas (motivos 8 y 6), mientras que el motivo 15 también estuvo presente en *A. trichopoda* (Amborellaceae). De igual manera, se puede observar que las Monocotiledóneas presentan un motivo específico de su clado (motivo 19); además de que el motivo 18, en el extremo carboxilo terminal, sólo se encuentra en este grupo y en *Amborella*. Por otra parte, en cuanto a los motivos presentes, *S. magellanicum* tuvieron mayor similitud en su extremo C-terminal, que en el N-terminal, aunque compartió más motivos con las demás secuencias en comparación a otras proteínas evaluadas en este trabajo.

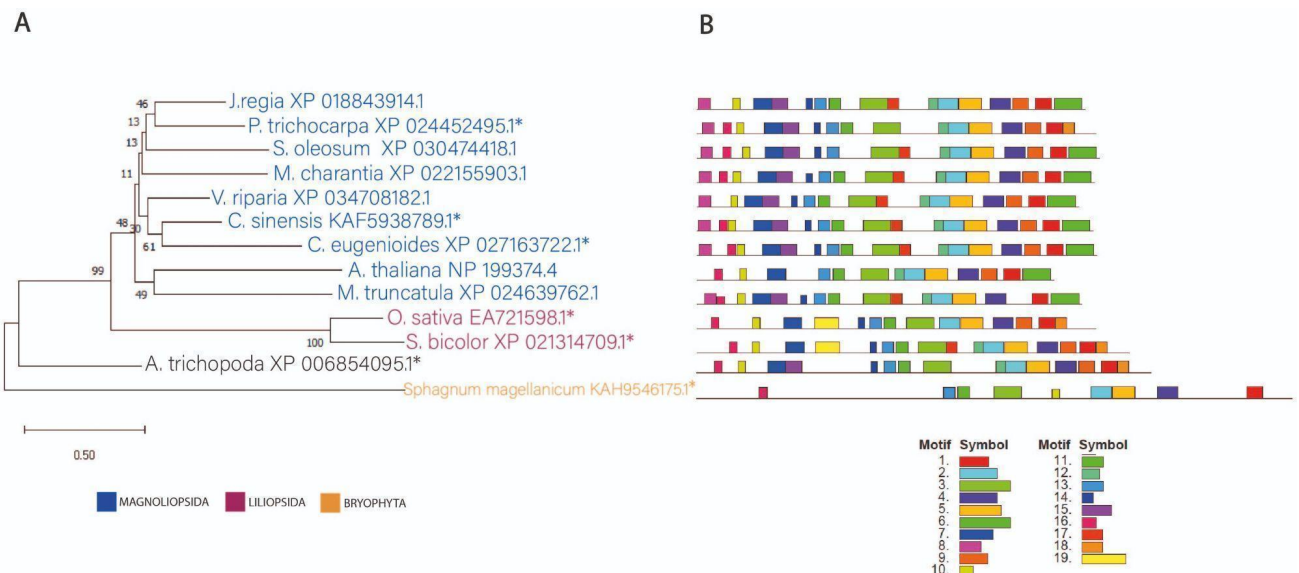


Figura 7. Cladograma de ATRIP. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replicaciones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. El asterisco indica secuencias no anotadas como ATRIP/SUV2.

8.2 Análisis de las proteínas que participan en Recombinación no homóloga (NHEJ)

A continuación se describirán los análisis de cada una de las proteínas involucradas en la vía de reparación de daño doble cadena por recombinación no homóloga.

8.2.1 Ku80/Ku70

Ku80 y Ku70 participan durante la recombinación no homóloga (NHEJ). El análisis de estas proteínas muestra un cladograma dividido en dos grupos: uno que comprende las secuencias homólogas a Ku80 y otro, a Ku70 (**Figura 8A**). No obstante, ambas secuencias de *H. sapiens* se agruparon más cerca de las plantas, que las secuencias utilizadas para las algas. Mientras que las secuencias de Ku80 de las algas presentan mayor similitud de motivos con las demás secuencias de plantas, que *H. sapiens* (**Figura 8B**). Esto podría indicar que hay una mayor divergencia de estas proteínas entre las algas y las plantas respecto a las de *H. sapiens*. Por otra parte, se puede observar que los motivos de ambas proteínas presentan un alto grado de conservación, incluso fuera de las angiospermas.

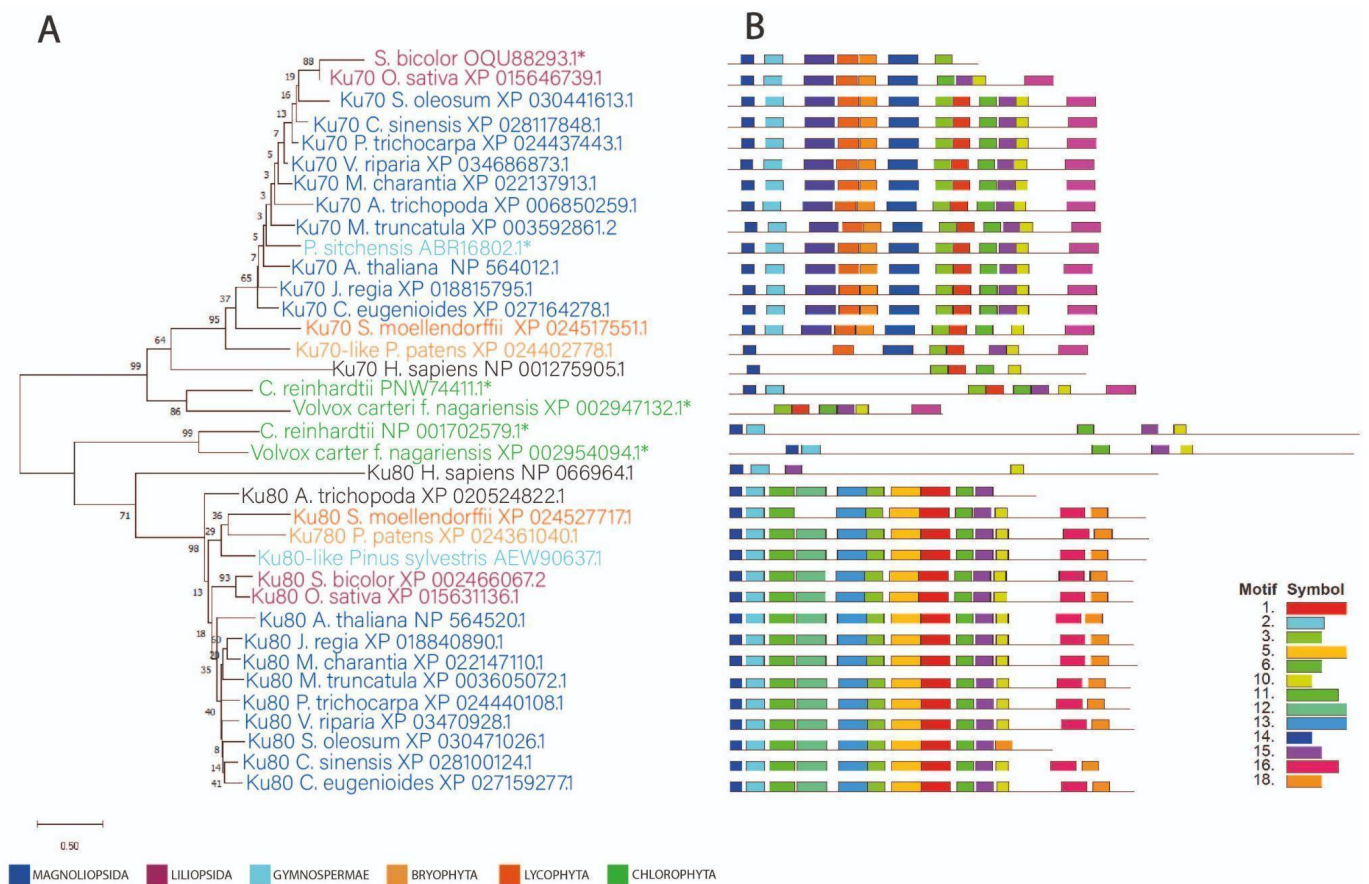


Figura 8. Cladograma de Ku70-Ku80. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replicaciones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. El asterisco indica secuencias no anotadas como Ku70-Ku80.

8.3 Análisis de las proteínas que participan en la Recombinación Homóloga (HR)

A continuación se describirán los análisis de cada una de las proteínas involucradas en la vía de reparación de daño de doble cadena por reparación homóloga.

8.3.1 Conservación de los principales señalizadores de la vía: ATM, ATR

En el mecanismo de señalización de daño al DNA, ATM y ATR en conjunto con SOG1, tienen un papel fundamental y se encuentran río arriba de la vía de señalización al daño. Dada su importancia, es de esperar que estas proteínas se encuentren altamente conservadas en las plantas, ya que cambios en sus secuencias o su ausencia podría provocar una respuesta alterada ante los agentes genotóxicos.

En humanos, ATM y ATR presentan los dominios FAT y FATC en su extremo carboxilo y entre ellos, un dominio catalítico PIKK; mientras que varios repetidos de motivos HEAT ocupan el extremo amino y varían entre las proteínas de la familia. Además, ATR presenta un dominio de unión de ATRIP en su extremo N-terminal (Shaik & Kirubakaran, 2020; Ueno, Sudo & Hirasawa, 2022).

Para evaluar el nivel de conservación de estas proteínas, se realizaron análisis de BLAST utilizando las secuencias de aminoácidos de las respectivas proteínas en *A. thaliana*. Los árboles de ambas proteínas comparten similitud en todos los clados observados (**Figura 9A y 10A**), estando las proteínas de todos los miembros de Magnoliopsida genéticamente cerca y agrupándose en clados monofiléticos las de Liliopsida. De igual forma, en ambos árboles el musgo *Physcomitrium patens* (Funariaceae) y la licofita *S. moellendorffii* (Selaginellaceae) conforman un clado monofilético. No obstante, los miembros de Chlorophyta parecen estar más alejados del resto.

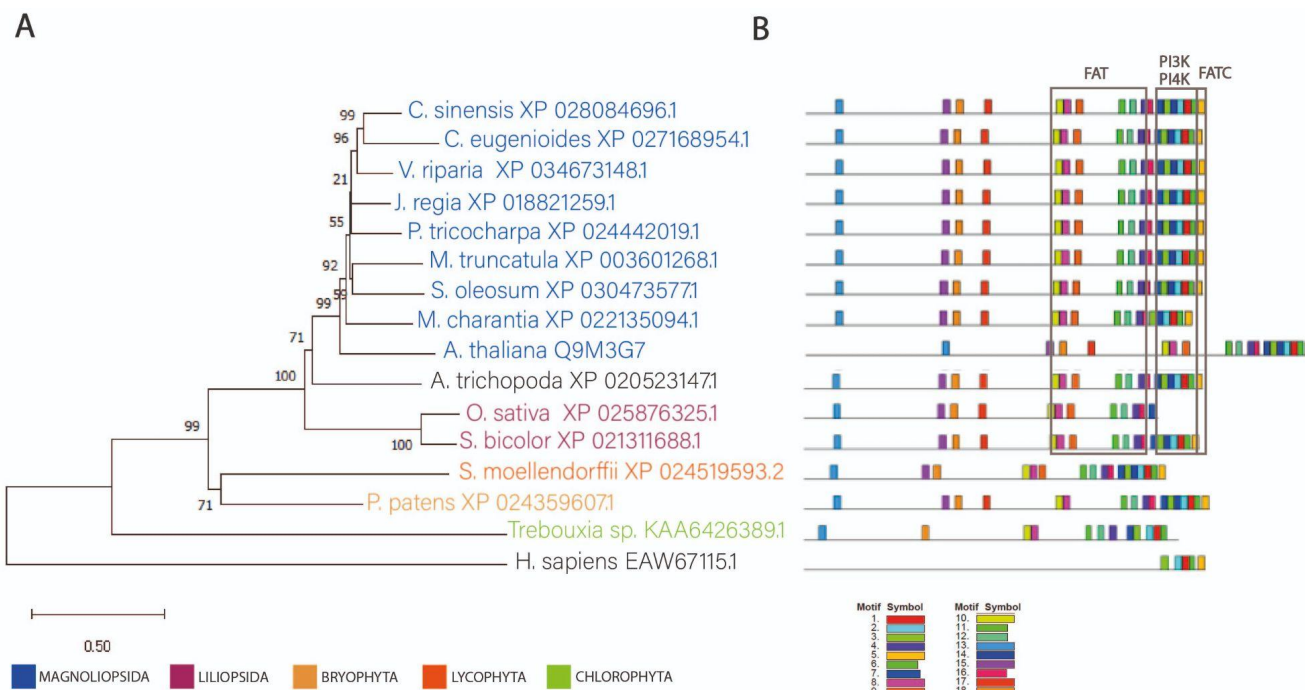


Figura 9. Filogenia de ATM. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replicaciones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. Los rectángulos grises marcan los dominios proteicos de ATM tomando como referencia la secuencia de *Camellia sinensis* (Theaceae).

Para evaluar el nivel de conservación de todas las secuencias analizadas, se realizó una búsqueda de motivos con el software en línea MEME suite. Dado que ATM y ATR son proteínas de gran tamaño que arrojaban más de 50 motivos con valor significativo, se estableció el número de 18 motivos para la búsqueda. En ambas proteínas, los 18 motivos parecen estar altamente conservados entre las secuencias y la mayoría de estos caen dentro de sus tres dominios proteicos (FAT, PI3K/PI4K y FATC) en su región carboxilo terminal (**Figura 9B y Figura 10B**). Así pues, se observaron motivos específicos de plantas (Motivos 8, 10, 11, 12, 13, 14, 16 y 18) o que al menos no se encuentran en *H. sapiens*, pero sí en todas las especies de Viridiplantae usadas. De la misma forma, el alga *Trebouxia sp.* y *H. sapiens* no presentan el motivo 17, estando presente únicamente en plantas superiores. Por último, el motivo 9 parece ser particular de plantas con semillas, ya que no se encuentra en *H. sapiens*, *Trebouxia sp.*, ni en *S. moellendorffii*.

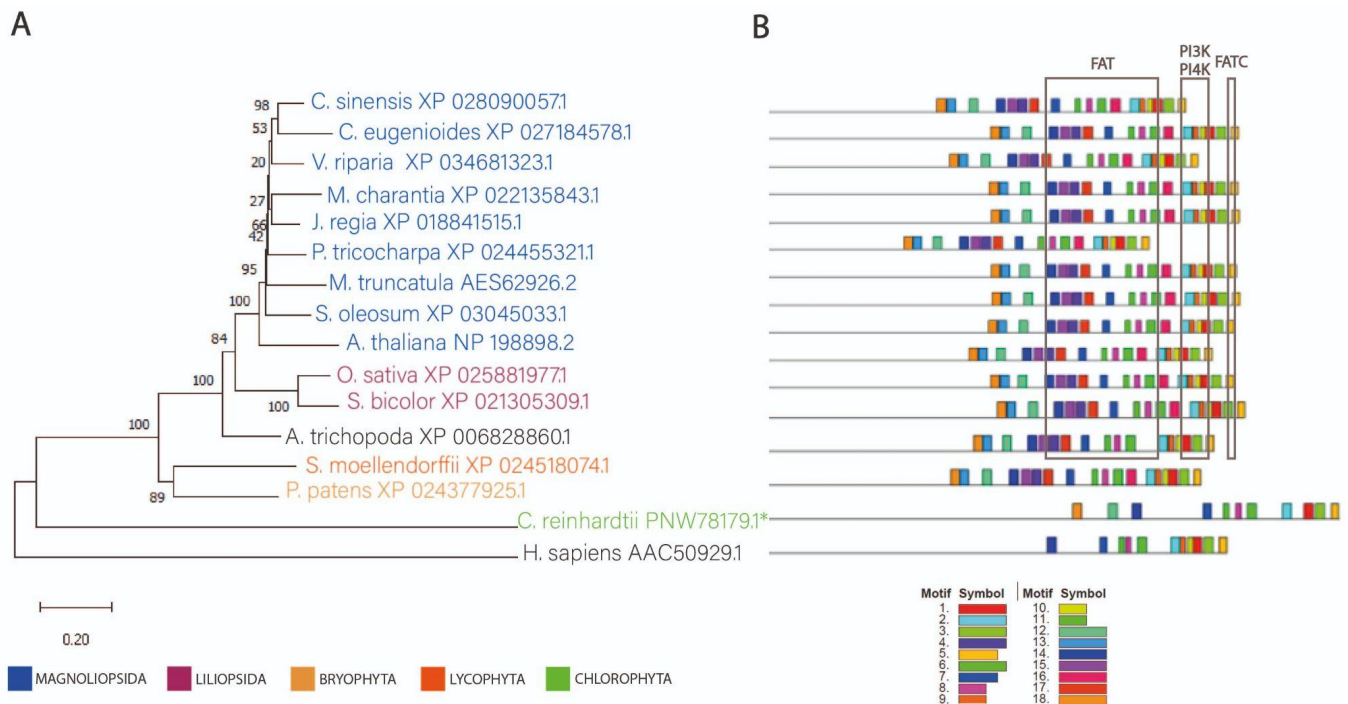


Figura 10. Filogenia ATR. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replications de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. Los rectángulos grises marcan los dominios proteicos de ATR tomando como referencia la secuencia de *A. thaliana* (Brassicaceae). El asterisco indica secuencias no anotadas como ATR.

El caso de *Oryza sativa* (Poaceae) resalta y es contrario a lo observado en las demás plantas con semilla, puesto que su proteína carece de los motivos situados en los dominios PI3K/PI4K y FATC, aunque no se puede descartar que la secuencia utilizada sea una isoforma de ATM que sólo tiene el dominio FAT.

Del mismo modo que ATM, la mayoría de los motivos de ATR caen en la región donde se encuentran sus dominios (FAT, PI3K/PI4K y FATC) (**Figura 10B**). Aunque la mayoría están bien conservados entre las plantas superiores, *S. moellendorffii* carece del motivo 16, que se encuentra entre el dominio FAT y el dominio PI3K/PI4K. Por otra parte, *Chlamydomonas reinhardtii* y *H. sapiens*, a diferencia de las plantas, no presentan los motivos 4, 13, 15, 16 y 17; mientras que el motivo 18 falta en *H. sapiens*. Interesantemente, aunque en animales sí está el motivo 9, *C. reinhardtii* carece del mismo y su motivo 12 se encuentra posicionado diferente al resto de secuencias.

En animales, mutaciones que afectan la función de ATM o ATR están asociadas a varios tipos de cáncer (Ueno, Tsudo & Hirasawa, 2022). Esto concuerda

con el hecho de que la mayoría de los motivos encontrados por MEME que caen dentro de los dominios funcionales de ambas proteínas estén altamente conservados. De igual forma, Martens et al., (2020) reportaron que ATR está más conservado entre varios organismos (incluyendo plantas y animales) en comparación con ATM. En concordancia con esto, HsATR compartió más motivos con las plantas que HsATM. Aunque existe un alto grado de conservación en la mayoría de los motivos, se ha observado que ATM y ATR podrían participar en procesos específicos en diferentes organismos, ejemplo de esto es la diferencia de fenotipos entre *A. thaliana* y el musgo *P. patens*; mientras que, en la primera, la ausencia de ATM o ATR no tiene efecto en el desarrollo vegetativo sin estrés, en *P. patens* las mutantes tienen claras anomalías aún en condiciones no estresantes (Martens et al., 2020). Quizá, los motivos encontrados por MEME que se encuentran en plantas con semilla, pero no en los demás organismos, estén relacionados con estas funciones específicas.

8.3.2 Suppressor of gamma response 1 (SOG1)

Otros señalizadores fundamentales para la HR que se encuentran río abajo de la cascada orquestada por ATM/ATR son SOG1 y BRCA1. Por un lado, SOG1 es una proteína que sólo se ha caracterizado en plantas y que podría llevar a cabo un papel similar al supresor de tumores p53, en mamíferos (Yoshiyama, 2016). Por medio de un BLAST, se obtuvieron secuencias de *S. moellendorffii*, *Pinus taeda* (Pinaceae) y *Picea sitchensis* (Pinaceae) con valor significativo para ser considerado como SOG1, pero que no están descritas como tal y se agregaron al árbol (**Figura 11**). Mientras que las proteínas de las gimnospermas *P. taeda* y de *P. sitchensis* conforman un grupo monofilético con las secuencias de las angiospermas, las secuencias de *S. moellendorffii*, así como las de *P. patens*, quedan más alejadas de las demás; no obstante, tienen más similitud con el SOG1 de angiospermas que otra proteína de la familia NAC de *A. thaliana* (base del árbol).

Por otra parte, MEME arrojó que las proteínas ortólogas a SOG1 comparten los mismos motivos (con ligeras excepciones) y en la misma disposición, mientras que las proteínas NAC de *P. sitchensis*, Lycophyta y Bryophyta sólo conservan los motivos que se encuentran dentro del dominio NAC (**Figura 11B**). De igual forma,

dentro de las angiospermas se puede ver que Magnoliopsida se diferencia por tres motivos que no se encuentran en Liliopsida (motivos 10, 14 y 19). Otra cosa por destacar es que, solo las angiospermas presentan el motivo 15 en su región N-terminal y que ambos SOG1 de *P. patens* presentan 4 motivos exclusivos de estas secuencias en su extremo carboxilo.

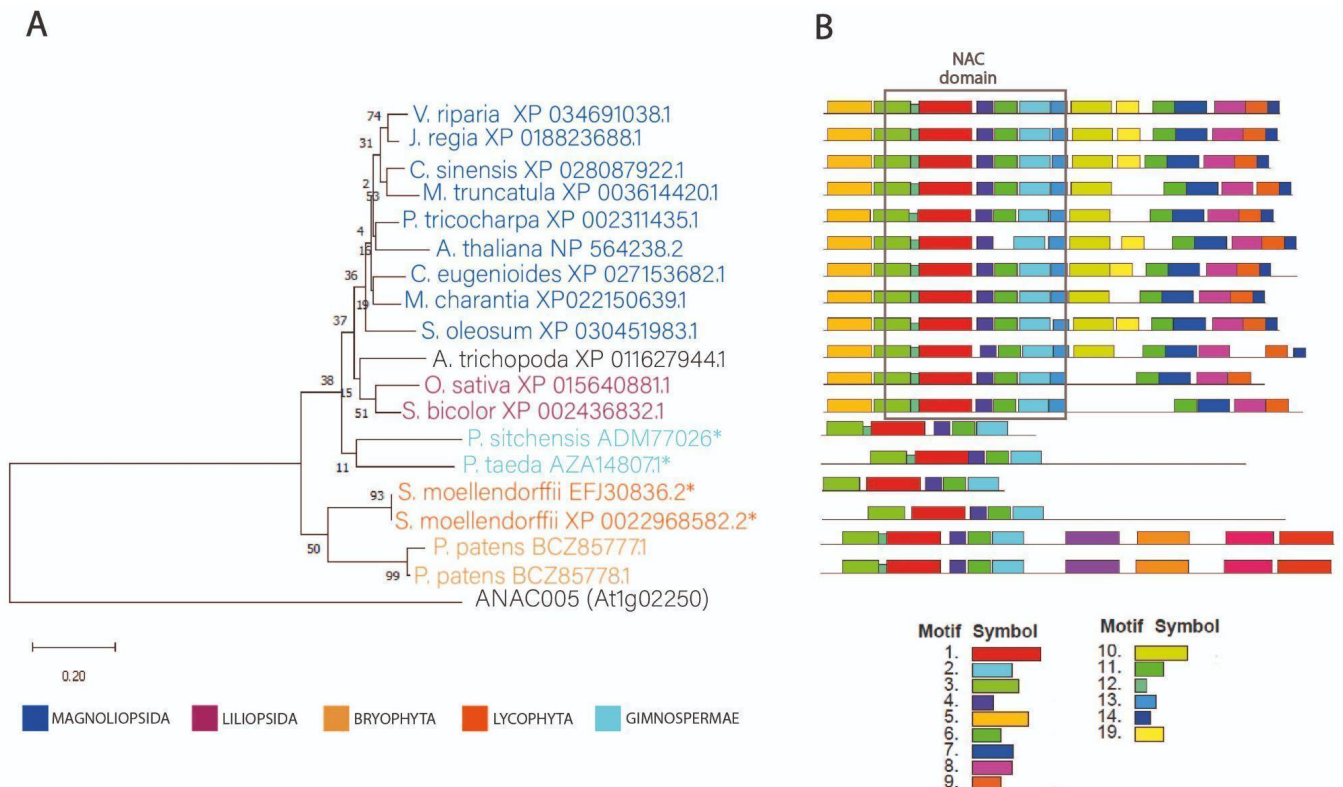


Figura 11. Filogenia de SOG1. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replicaciones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. El rectángulo gris marca el dominio NAC, tomando como referencia las secuencias de Magnoliopsida. El asterisco indica secuencias no anotadas como SOG1.

Por otra parte, SOG1 presenta 5 motivos SQ en su región C-terminal, cuya fosforilación por parte de ATM/ATR es importante para su activación (Yoshiyama et al., 2013). Con la finalidad de evaluar si todas las proteínas evaluadas podrían ser SOG1-like, se buscaron motivos SQ en su extremo C-terminal. Si bien, todas las angiospermas evaluadas (a excepción de *A. trichopoda*) tienen los 5 motivos SQ, las monocotiledóneas presentan dos motivos más (**Figura 12**). Así pues, ninguna de las secuencias fuera de las angiospermas presenta los 5 motivos SQ en las posiciones exactas, teniendo menos, ninguno, igual o más, pero en diferentes posiciones. No obstante, *P. taeda* sí tiene cinco motivos, 3 de ellos conservados en la posición

(350SQ, 356SQ, 436SQ) y otros dos en distinta posición; sumado a esto, su secuencia presenta la extensión en el extremo amino, otra característica de SOG1.

Inicialmente, fue reportado por Yoshiyama et al. (2014) que, dado la dificultad para decir si las proteínas NAC de Lycophytas y Bryophytas correspondían o no a proteínas SOG-like, era posible que SOG1 hubiera surgido en las gimnospermas. No obstante, recientemente se encontraron dos homólogos de SOG1 en *P. patens* (Sakamoto et al., 2021), siendo estos los utilizados en el presente trabajo. Los árboles filogenéticos aquí obtenidos indican que estas secuencias podrían haber divergido mucho en comparación con el SOG1 de las plantas con semilla. Por otro lado, dos hechos que ponen en duda si las secuencias de *S. moellendorffii* podrían no corresponder a proteínas SOG-like son los siguientes: Ninguna de estas proteínas tiene la extensión del extremo N-terminal característica de SOG1; y no cuentan con los motivos SQ conservados. Cabe resaltar que las secuencias de *P. patens* únicamente tienen dos motivos SQ, uno en posición conservada y el otro, más hacia su extremo carboxilo. Dado esto, no podemos descartar que las secuencias de *S. moellendorffii* sean únicamente factores de transcripción tipo NAC o que sean proteínas SOG1-like con una regulación distinta que sea exclusiva de este clado.

El caso de *P. taeda* resulta interesante, en primer lugar, porque, aunque tiene la extensión N-terminal, MEME no lo reconoce como el mismo motivo que en angiospermas y en segundo, porque presenta 3 de 5 motivos SQ conservados. Si bien, en *A. thaliana* se ha visto que la fosforilación de los 5 motivos es necesaria para la completa activación de SOG1, el número de motivos fosforilados influye en el nivel de expresión de los genes regulados por el mismo (Yoshiyama et al., 2017) y el hecho de que *P. taeda* tenga más motivos podría indicar una regulación específica. Además, Yoshiyama et al. (2017) reportaron que es necesario que se fosforilen los dos primeros motivos (350SQ, 356SQ) para que se dé la fosforilación de los siguientes motivos (372, 430 y 436). Anteriormente, en *P. glauca* se reportó una secuencia con estos mismos motivos conservados (Yoshiyama et al., 2014; Yoshiyama, 2016), lo cual coincide con la secuencia de *P. taeda* aquí utilizada, indicando su homología con SOG1.

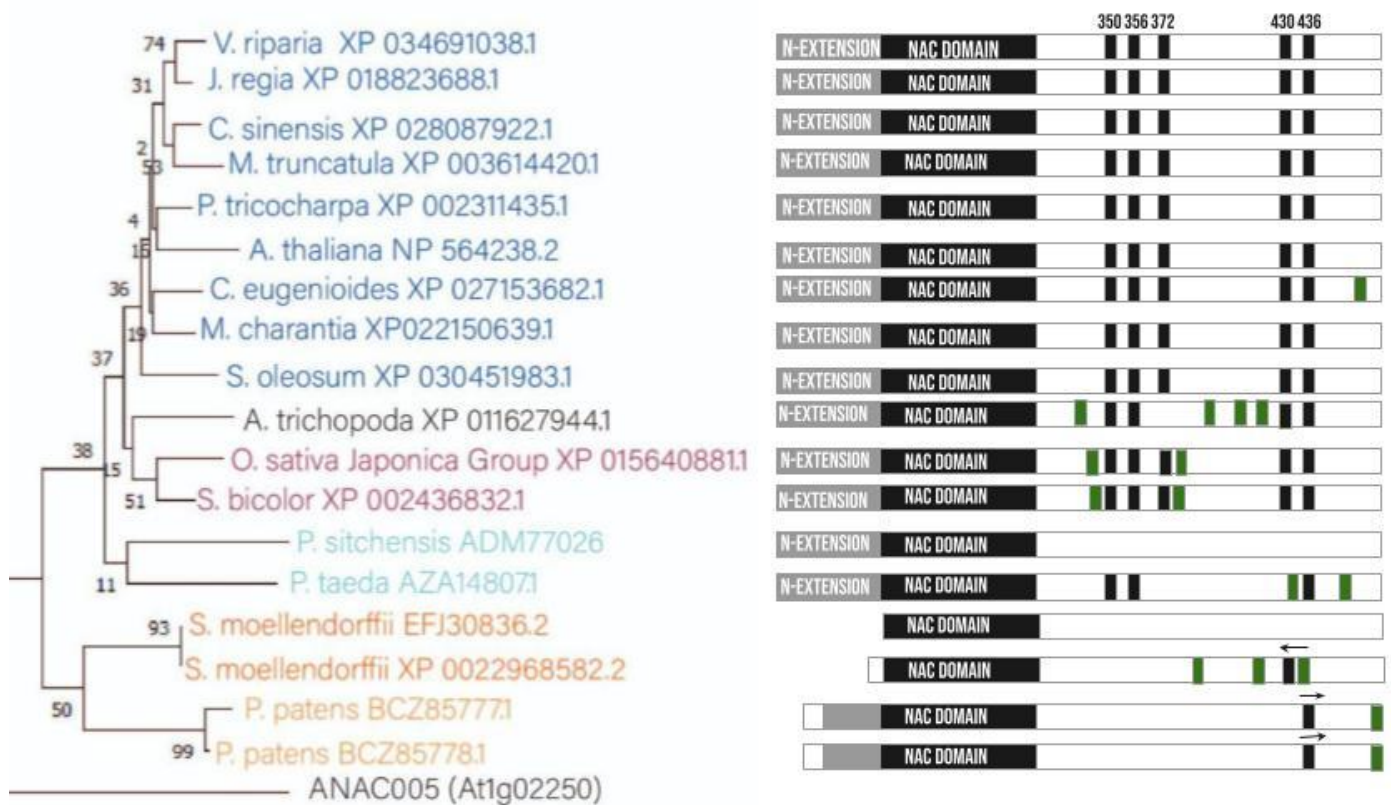


Figura 12. Conservación de los motivos SQ en SOG1. A) cladograma de las proteínas SOG1 y B) su asociación con los motivos SQ tomados de Yoshiyama et al., (2015). En negro se muestran los motivos en posiciones conservadas; en verde, posiciones no conservadas. Las flechas indican un cambio de posición de 1 nucleótido.

8.3.3 Breast cancer susceptibility gene 1 (BRCA1)

Durante la DSB, BRCA1 en complejo con otras proteínas, participa en la recombinación homóloga. En humanos, BRCA1 es una proteína de 22 kDa con un dominio RING en su extremo N-terminal y dos dominios BRCT, en el extremo carboxilo; entre ellos, un dominio de interacción de p300/CBP y uno *coiled coil* (Deng & Brodie, 2000; Hu et al., 2000). Un sólo homólogo de BRCA1 fue identificado en *A. thaliana*, el cual corresponde a una proteína de 104 kDa con los dominios RING, p300/CBP y ambos BRCT conservados (Lafarge & Montané, 2003). Además, presenta un dominio PHD y un dominio PHD extendido (Trapp, Seeliger & Putcha, 2011).

Dado que BRCA1 comparte ancestro común con BARD1 (Trapp, Seeliger & Putcha, 2011), se usaron secuencias de ambos para construir el árbol filogenético (**Figura 13**). Dentro de las angiospermas el árbol se dividió en dos grupos: un grupo con las secuencias de BRCA1 y otro grupo con las de BARD1 (**Figura 13A**). De forma interesante, la secuencia de BRCA1 de *A. trichopoda* se encontró en el grupo de BARD1; sin embargo, tanto el BRCA1 de *A. trichopoda* como el de *S. moellendorffii* presentan el motivo 13 (**Figura 13B**), el cual cae en la secuencia del dominio p300/CBP, dominio que se presenta en BRCA1 de plantas, pero no en BARD1 (Trapp, Seeliger & Putcha, 2011).

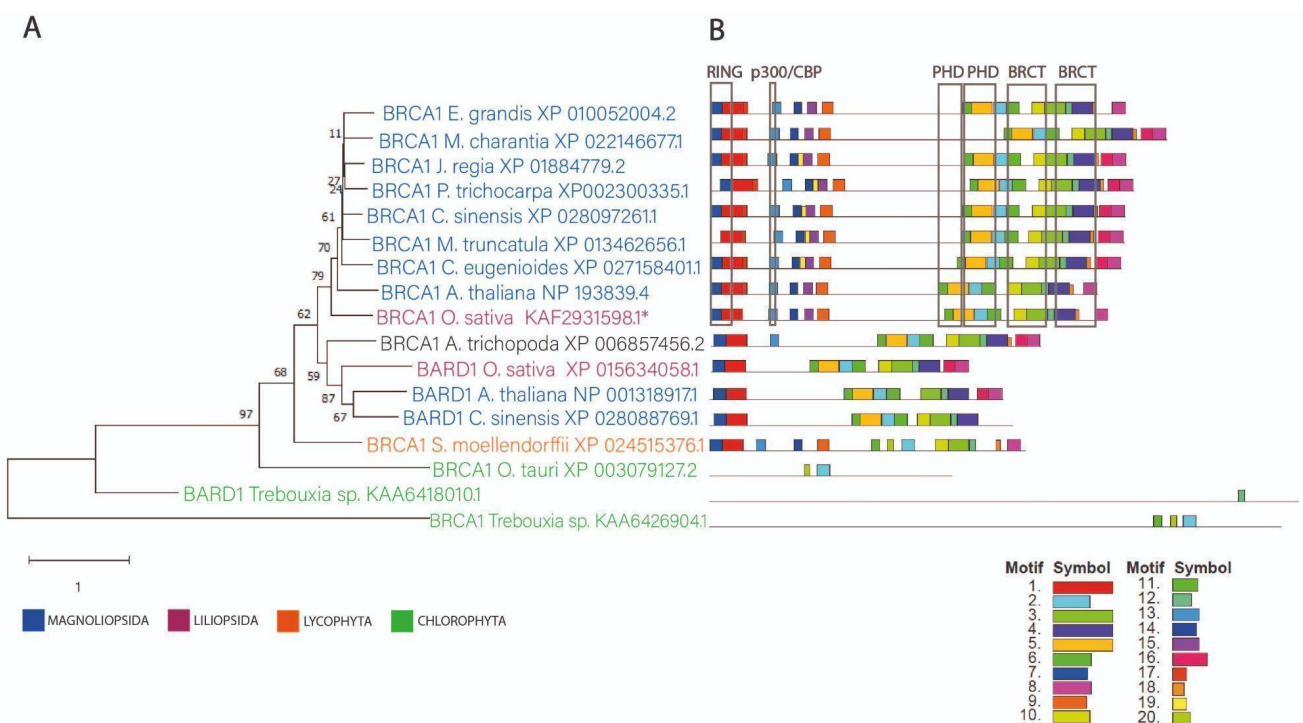


Figura 13. Filogenia BRCA1. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de repeticiones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. Los rectángulos grises marcan los dominios proteicos de BRCA1 tomando como referencia la secuencia de *Arabidopsis thaliana* (Brassicaceae). El asterisco indica secuencias no anotadas como BRCA1.

Así pues, se pueden observar 3 motivos específicos de BRCA1 (14, 15, 9) que no se encuentran en *A. trichopoda*, pero de los cuales, dos de ellos sí están presentes en BRCA1 de *S. moellendorffii*. Esta última presenta similitud con las otras secuencias de BRCA1 en los motivos que caen en el primer dominio BRCT y parte de los motivos del segundo dominio PHD y el segundo BRCT; sin embargo, su

extremo C-terminal también es similar a la secuencia de *Trebouxia sp.* (por el motivo 20).

Según Trapp, Seeliger & Putcha (2011) una hipótesis sobre la evolución de BRCA1 y BARD1 sería que ambas proteínas comparten un ancestro en común BRCA1/BARD1 en las algas, el cual sufrió una duplicación en algún punto entre la divergencia de Lycopodiopsida y Spermatophitina, pero antes de la separación entre monocotiledóneas y dicotiledóneas. Esto concuerda con el hecho de que *S. moellendorffii* presenta dos motivos que están en BRCA1, pero no en BARD1, los cuales se pudieron perder en esta última (**Figura 13**). También, Trapp, Seeliger & Putcha (2011) proponen que el dominio PHD se adquirió en las algas verdes, aunque no todas lo tienen. Los motivos obtenidos por el programa MEME indicaron poca conservación en esta región en las secuencias de las algas, lo cual no significa *per se* que no tengan el dominio, sino más bien que hay mucha variación en la secuencia.

El dominio PHD es capaz de reconocer marcas postraduccionales de las histonas (Mouriz et al., 2015), lo que podría correlacionar la función de BRCA1 y la regulación epigenética durante la reparación a daño de DNA; aunque la función exacta de este dominio en BRCA1 no se ha estudiado a profundidad, existe información sobre el papel de factores epigenéticos y ejemplo de esto es la participación de la metiltransferasa SUV2, de *Medicago truncatula* (Fabaceae), en la regulación de la compactación de la cromatina ante DSB y la creación de fases líquido-líquido donde se reclutan proteínas importantes para la HR, como RAD51 (Liu et al., 2022).

A pesar de que los dominios RING, PHD y BRCT están altamente conservados en diferentes clados, la zona intermedia es más variable y no presenta motivos en común. Aunque esto mismo se observa en especies de animales, no se han reportado casos donde BRCA1 cumpla una función distinta (Pfeffer, Ho & Singh, 2017). En chimpancés se ha visto que BRCA1 está bajo una constante selección (Lou et al., 2014). Dado que en las plantas han sufrido varios eventos de duplicación genómica y rearrreglos cromosómicos (One Thousand Plant Transcriptomes

Initiative, 2019), es posible que BRCA1 también se encuentre evolucionando y esto explique parte de la variación observada.

Además de la recombinación homóloga, en animales se ha visto que BRCA1 puede participar en procesos como checkpoints del ciclo celular en G1/S y G2/M, dependiendo de los complejos a los cuales se asocia (Trapp, Seeliger & Puchta, 2011). La variación en la historia de vida dentro de Viridiplantae también podría haber influido en la variación de la proteína.

8.4 Análisis de proteínas con funciones específicas dentro de la ruta de HR

8.4.1 El complejo proteico MRN (MRE11, RAD50 y NBS1)

Posteriormente, se decidió evaluar proteínas que participan en la HR, pero no como señalizadoras, si no con funciones más específicas. Entre ellas se encuentran las proteínas que conforman el complejo MRN: MRE11, NBS1 y RAD50 (Yoshiyama, Sakaguchi & Kimura, 2013). En primer lugar, se llevaron a cabo los análisis de MRE11 (MEIOTIC RECOMBINATION 11), la cual es una proteína que en animales presenta 5 motivos fosfoesterasa junto con un dominio de unión a NBS1 en su región N-terminal; mientras que en el extremo carboxilo presenta dos dominios de unión a DNA y entre ellos, un dominio de unión a RAD50 (Rupnik, Lowndes & Grenon, 2010).

El árbol filogenético de MRE11 dentro de las angiospermas se dividió en dos grupos: monocotiledóneas y dicotiledóneas (**Figura 14A**). Cabe destacar que Magnoliopsida presenta un motivo (17) que está ausente en Liliopsida (**Figura 14B**) y que cae dentro de la secuencia reportada para la unión de RAD51; este podría ser importante para la función de la proteína. De forma interesante, este motivo tampoco se encuentra en *P. patens*, *C. reinhardtii* y *H. sapiens*, pero sí en *S. moellendorffii*; por lo que este motivo pudo haber aparecido en Lycophyta y haberse perdido en los organismos que no lo presentan.

Cerca de la zona de unión de RAD51 hay otro motivo que parece ser específico de angiospermas (motivo 16), a excepción de *Camellia sinensis* (Theaceae), cuyo caso es particular dado que carece de una parte de ambos extremos N y C terminal de la proteína (**Figura 14B**). Esta característica y el tamaño puede ser indicativo de una pequeña isoforma de RAD51 en *C. sinensis*.

Los últimos motivos del extremo C-terminal parecen ser también específicos de angiospermas, con excepción del motivo 13, el cual cae dentro de la región reportada como motivo de unión al DNA. Si bien, este motivo está presente en casi todas las secuencias de plantas y la de *H. sapiens*, no se encuentra en la licofita *S. moellendorffii*. Otro motivo que cae en una secuencia reportada importante para unión al DNA es el motivo 6, el cual no está en el alga *C. reinhardtii*, ni en *H. sapiens*.

Es interesante que aunque la secuencia de *C. reinhardtii* presenta más distancia genética con las de las plantas superiores, está bastante conservada en su extremo N- terminal y tiene similitud con el MRE11 de animales.

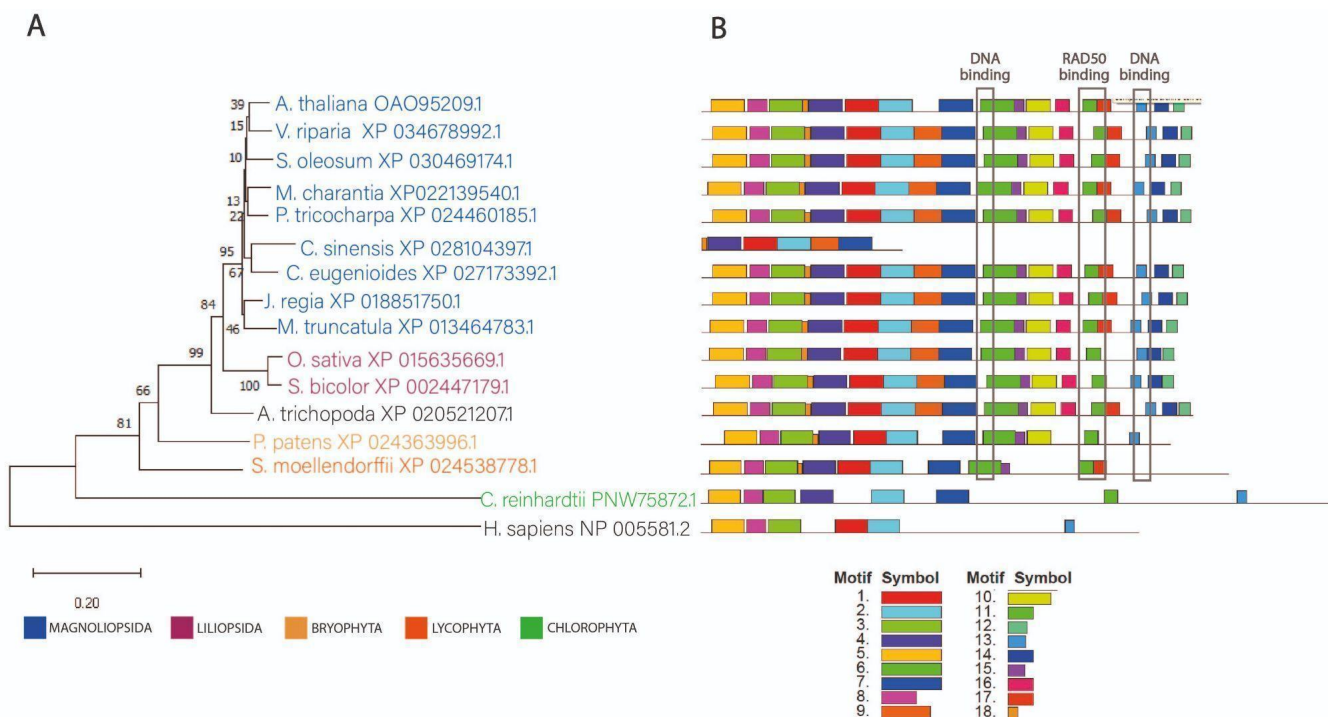


Figura 14. Filogenia MRE11. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replicaciones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. Los rectángulos grises indican los dominios proteicos de MRE11 tomando como referencia la secuencia de *Arabidopsis thaliana* (Brassicaceae).

8.4.2 Nijmegen breakage syndrome 1 (NBS1)

En animales, NBS1 presenta un dominio FHA (Fork-associated domain) y dos dominios BRCT en su extremo amino, así como sitios de unión para RPA, RAD18, MRE11, RNF20 y ATM, en su extremo carboxilo (Komatsu, 2016; Rupnik, Lowndes & Grenon, 2010).

A pesar de que la filogenia no logró resolver bien el clado de las dicotiledóneas (**Figura 15A**), se puede observar una clara separación entre ellas y las monocotiledóneas, así como de las angiospermas con *P. patens* y *H. sapiens*. No obstante, resalta la secuencia de *Vitis riparia* (Vitaceae), la cual parece tener mayor distancia genética con respecto a las demás y presenta únicamente los motivos del extremo Carboxilo (**Figura 15B**). Lo que podría ser indicativo de una isoforma pequeña de la proteína.

Si bien, los motivos que caen dentro de los dominios reportados están bastante conservados en las plantas con semilla, las monocotiledóneas presentan un motivo extra dentro de la secuencia de FHA (motivo 14). Además, el musgo *P. patens* carece de varios motivos presentes en las angiospermas en los dominios BRCT y la región de unión a ATM.

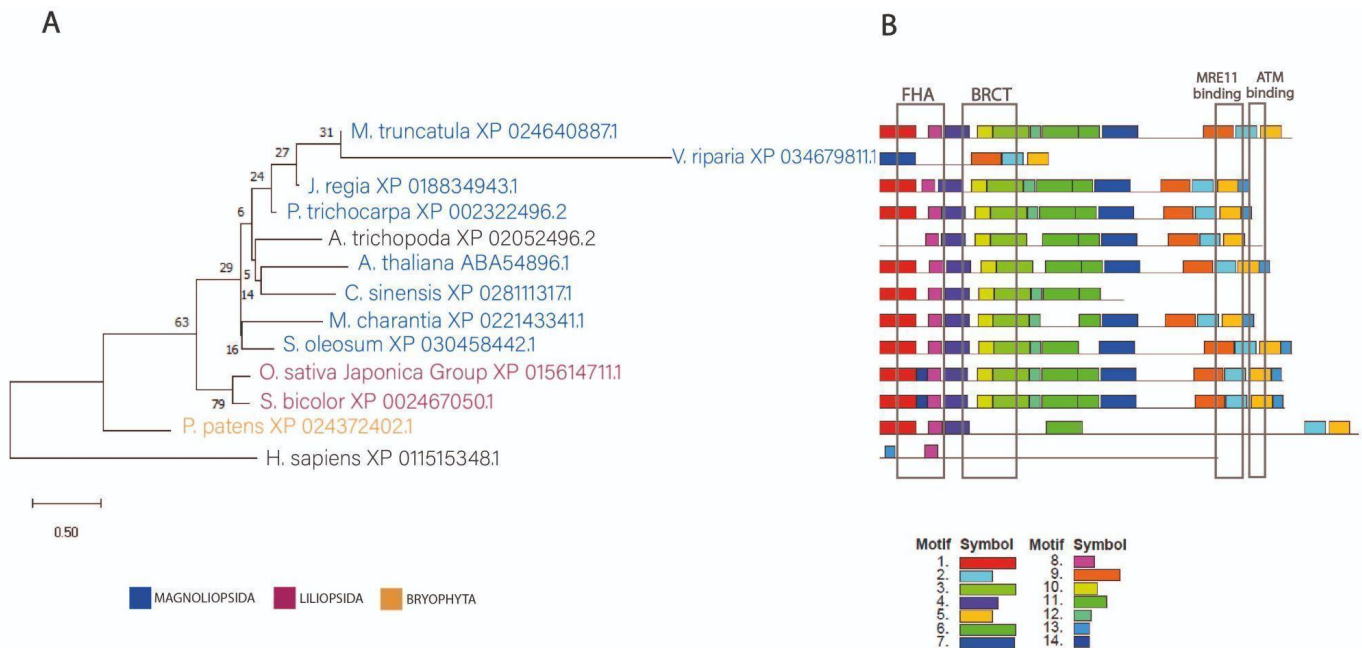


Figura 15. Filogenia NBS1. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replicaciones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. Los rectángulos grises marcan los dominios proteicos de NBS1 tomando como referencia la secuencia de *Arabidopsis thaliana* (Brassicaceae).

8.4.3 RAD50

Otra proteína que forma parte del complejo MRN es RAD50. Esta proteína presenta un dominio Walker en cada extremo, los cuales participan a su función como ATPasa, mientras que en la parte central tienen un dominio bisagra con un zinc hook (Rupnik, Lowndes y Grenon, 2010).

La filogenia en el grupo de las angiospermas separó las secuencias de Magnoliopsida y Liliopsida; no obstante, *C. sinensis* quedó fuera de estos dos clados (**Figura 16A**). La secuencia de *C. sinensis* es pequeña y podría ser una isoforma de la proteína. RAD50 parece estar altamente conservada, en plantas vasculares, en los motivos encontrados por MEME que caen dentro de los dominios reportados (**Figura 16B**); siendo la única excepción *C. sinensis*. Por otra parte, las angiospermas presentan 3 motivos entre los dominios de ATPasa y Zinc hook que no están en los otros organismos (motivos 10,17 y 18), así como otro motivo después del zinc hook (motivo 14). De igual manera, los motivos 9, 11, 15 y 16

parecen ser exclusivos de plantas terrestres, ya que no se encuentran en *C. reinhardtii*, ni en *H. sapiens*.

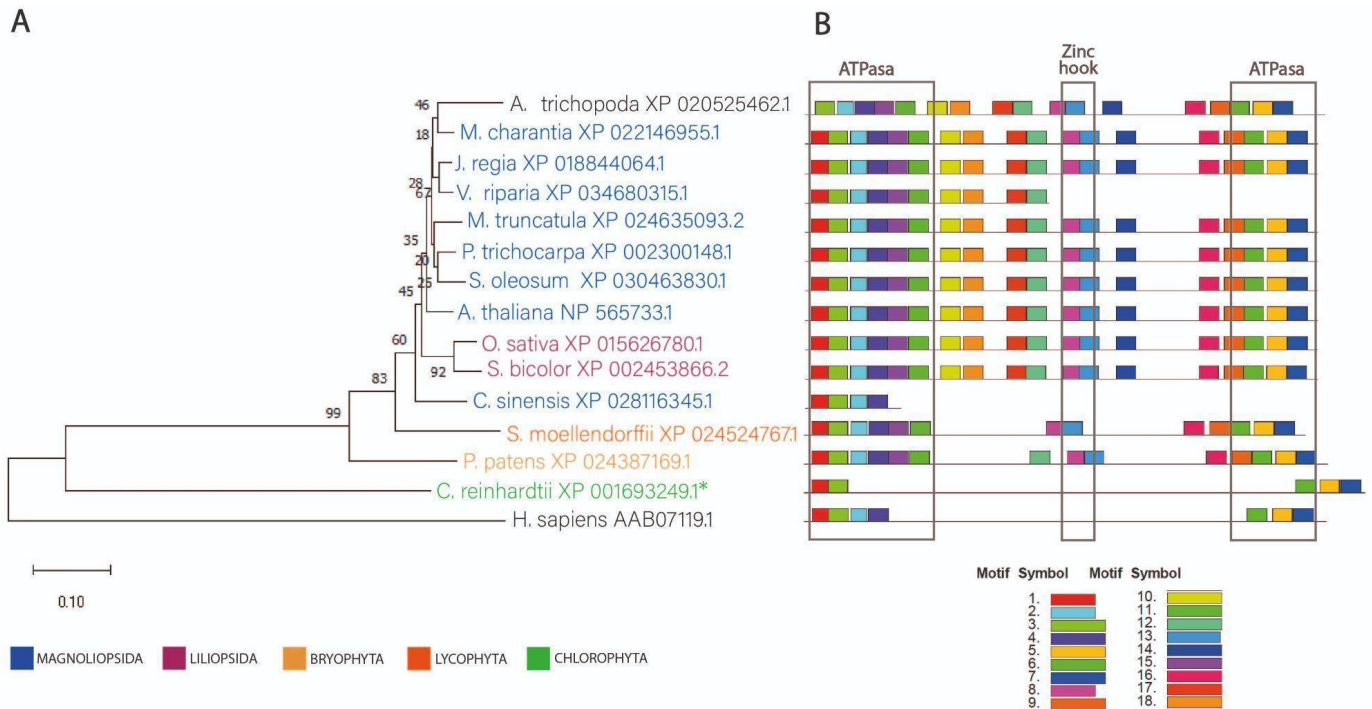


Figura 16. Filogenia Rad50. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de repeticiones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. Los rectángulos grises marcan los dominios proteicos de Rad50 tomando como referencia la secuencia de *Arabidopsis thaliana* (Brassicaceae). El asterisco indica secuencias no anotadas como Rad50.

Mientras que MRE11 y RAD50 se han encontrado en todos los taxa, NSB1 sólo se encuentra en eucariontes (Rupnik, 2008). MRE11 y RAD50 están altamente conservados entre los organismos, tan solo el RAD50 de *A. thaliana* y del arroz (*O. sativa*) tienen una similitud de 64% (Pérez et al., 2011), en tanto el MRE11 del trigo y *Arabidopsis* muestran 84% de identidad (De Bustos, Pérez & Jouve, 2007); lo que concuerda con el nivel de similitud de los motivos de MEME entre secuencias. Así pues, se ha reportado que MRE11 presenta mayor grado de conservación en el extremo N-terminal, dentro de donde se encuentran los motivos de fosfoesterasa (De Bustos, Pérez & Jouve, 2007), lo cual de nuevo concuerda con lo observado por MEME.

Por otra parte, NBS1 presenta un menor grado de conservación dado que los homólogos de diferentes organismos tienen menor similitud (Pérez et al., 2011). Incluso, se ha observado que el NBS1 de *A. thaliana* y el de *O. sativa* comparten 57% de similitud (Akutsu et al., 2007), lo cual podría explicar el problema de

resolución del árbol de las angiospermas, así como la gran diferencia encontrada entre los motivos presentes en plantas y aquellos de *H. sapiens*.

Llama la atención las diferencias en los motivos del extremo C-terminal entre las plantas con semilla y *P. patens*. Cabe resaltar que se ha observado en mamíferos que el motivo de reconocimiento de MRE11 en NBS1 es suficiente para la viabilidad celular en ausencia de la proteína (Paull et al., 2018). Esto podría explicar porque *P. patens* difiere tanto en otros motivos, además de que en ella, se ha reportado que NBS1 no es esencial para la HR durante el gene targeting, mientras que MRE11 y RAD50 sí lo son (Kamisugi et al., 2012) .

También, el hecho de que las mutantes de pérdida de función de RAD50 y MRE11 tengan fenotipos más fuertes tanto en *Arabidopsis*, como en *Physcomitrium*, que las mutantes de NBS1 (Kamisugi, 2011), refleja la importancia de su función y puede ser que esté correlacionado con las diferencias en el nivel de conservación de las 3 proteínas.

8.5 Análisis de proteínas que participan en la búsqueda de cadenas homólogas

8.5.1 Breast cancer susceptibility gene 2 (BRCA2)

Breast cancer susceptibility protein 2 (BRCA2) se encarga de dirigir a RAD51 a las cadenas dañadas durante la DSB (Pfeffer, Ho & Singh, 2017). El BRCA2 de humano está conformado por 8 repetidos BRC, un DBD y un dominio de unión a PALB2 en su extremo N-terminal, y un extremo carboxilo con sitios de unión DDS1 y RAD51 dentro de un dominio helicoidal y tres pliegues OB (Andreassen et al., 2021). Por otra parte, en *A. thaliana* se presentan dos proteínas homólogas a BRCA2, que cuentan con 4 repetidos BRCA y al igual que en animales, sitios DBD homólogos a los de hBRCA2, un dominio helicoidal y tres pliegues OB (Trapp, Seeliger & Puchta, 2011).

En el árbol filogenético se puede observar que, si bien, en la parte de las angiospermas la filogenia se divide en dos grandes grupos (dicotiledóneas y monocotiledóneas), *A. thrichopoda* queda dentro del grupo de las monocotiledóneas (**Figura 17A**). Otra cosa interesante que muestra el árbol es que los BRCA2 tipo A y B no se agrupan, es decir, puede ser que hayan surgido de eventos de duplicación independientes y que, por lo tanto, todos los BRCA2A no sean homólogos entre sí. Esto concuerda con lo reportado en *A. thaliana*, donde ambas copias de BRCA2 comparten 96.8% de identidad y es probable que hayan surgido de una duplicación reciente (Siaud et al., 2004).

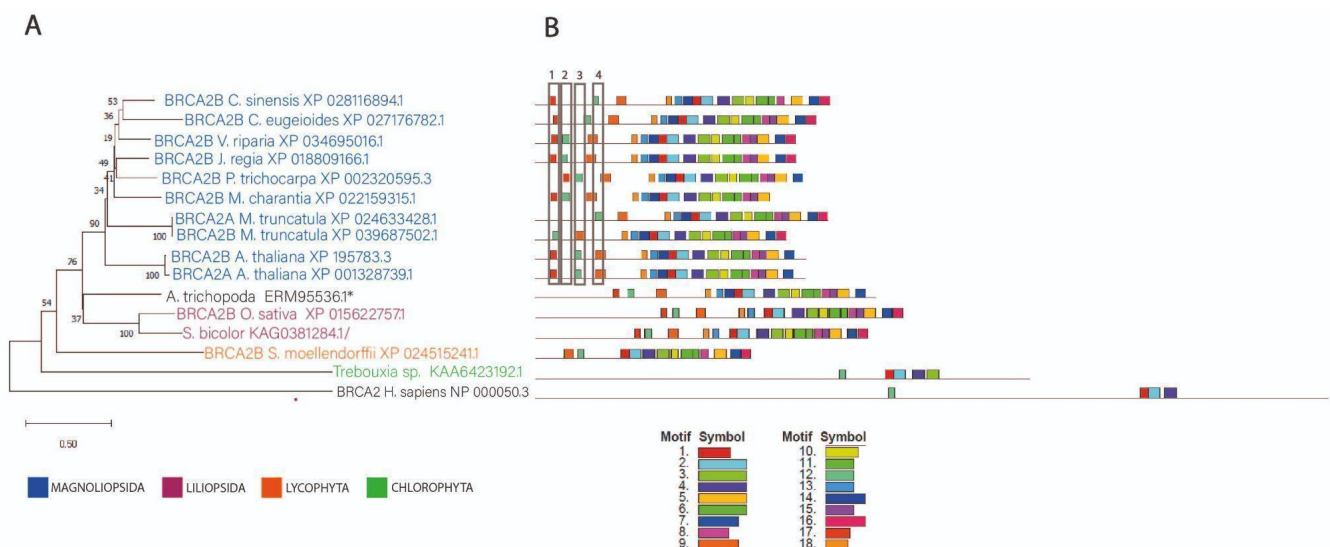


Figura 17. Filogenia BRCA2. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replicaciones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. Los rectángulos grises marcan los dominios proteicos de BRCA2 tomando como referencia la secuencia de *Arabidopsis thaliana* (Brassicaceae). El asterisco indica secuencias no anotadas como BRCA2. Los números indican los dominios BRC1, 2, 3 y 4, respectivamente.

En cuanto a los motivos arrojados por MEME, se puede observar que hay una alta conservación de estos en las secuencias de plantas (**Figura 17B**). No obstante, los motivos 13, 15, 17 y 14 parecen ser exclusivos de plantas con semilla, dado que no se encuentran en la licofita *S. moellendorffii*. También se puede observar que varias plantas carecen de un motivo en el extremo C-terminal, pero su presencia o ausencia no es particular de un solo clado. Las proteínas de *Trebouxia* sp. y *H. sapiens* guardan poca similitud con el resto y sólo comparten 3 motivos en su extremo carboxilo terminal. La conservación del extremo C-terminal también se ha

visto entre animales, hongos y plantas (Kojic et al., 2002), lo que indica que este segmento es sumamente importante para la función de BRCA2.

A diferencia de BRCA1 que sólo se encuentra en animales y plantas, BRCA2 también está presente en hongos. En los tres grupos, BRCA2 conserva la función de interactuar con RAD51 (Pffefer, Ho & Singh, 2017) y algunos de los sitios a través de los cuales lo hace, es por medio de los dominios BRC (Roy, Chun & Powell, 2016). Es interesante que los motivos reportados por MEME que caen dentro de los dominios BRC están conservados en las plantas (**Figura 17B**), pero no aparecen en *Trebouxia sp.* y *H. sapiens*. Esto puede ser debido a que las secuencias de los 8 BRCs de *H. sapiens* difieren mucho de los de plantas. Interesantemente, el número de BRC varía entre organismos, hay algunos que tienen sólo uno (como en *Ustilago maydis*) o hasta 15 (*Trypanosoma*) (Hartley & Mcculloch, 2008; Kojic et al., 2002).

8.5.2 RAD51

Durante la HR, el filamento de nucleoproteína formado por monómeros de RAD51 se encarga de la invasión y alineamiento del ssDNA dañado con la cadena homóloga del otro cromosoma (Bonilla et al., 2020). En general, RAD51 presenta dos regiones conservadas: un motivo HuH de unión a DNA y un dominio central α/β ATPasa (marcado en uniprot como FtsK domain, en Arabidopsis), que contiene un motivo Walker A y un motivo Walker B (Bonilla et al., 2020; Lin et al., 2006). Los análisis realizados mostraron que las proteínas de RAD51 están bastante conservadas (**Figura 18**). Sin embargo, en el árbol filogenético se puede observar que la secuencia del alga *C. reinhardtii* está más alejada genéticamente de las secuencias de plantas a comparación de *H. sapiens* (**Figura 18A**). Asimismo, *C. reinhardtii* carece de los motivos 2, 5 y 8 (**Figura 18B**), que se encuentran en las demás secuencias, incluyendo la de *H. sapiens*, y dos de las cuales entran en el dominio FtsK de RAD51. Por otra parte, se puede observar que sólo el grupo de Magnoliopsida presenta el motivo 10 en su extremo amino terminal, a excepción de *V. riparia* y que *A. trichopoda* carece del motivo 9, el cual se encuentra en el extremo C-terminal de todas las demás secuencias.

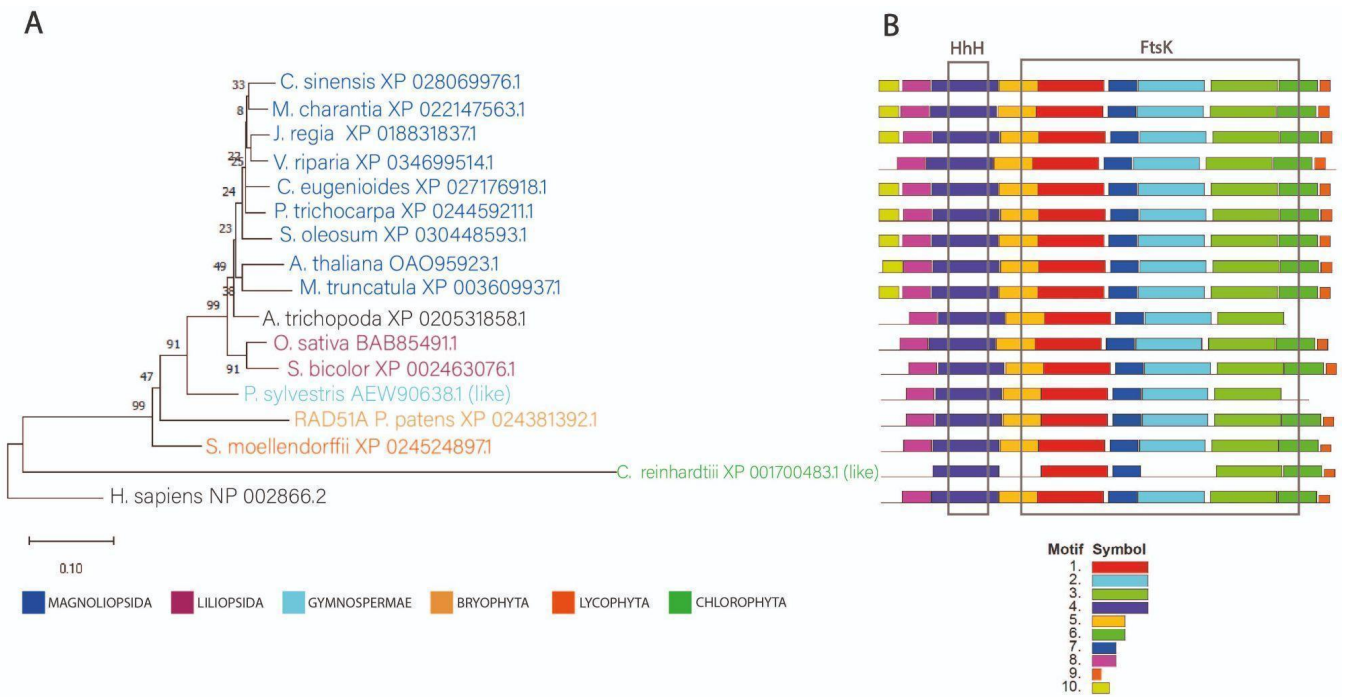


Figura 18. Filogenia RAD51. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replications de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. Los rectángulos grises marcan los dominios proteicos de RAD51 tomando como referencia la secuencia de *A. thaliana* (Brassicaceae).

RAD51 es una proteína altamente conservada no sólo dentro de las plantas, se ha observado que AtRAD51 y HsRAD51 comparten 69% de similitud (Kobayashi et al., 2019), lo que concuerda con el gran parecido de motivos arrojados por MEME. Esta alta conservación de RAD51 podría ser explicada por su importancia en el mantenimiento del genoma. En vertebrados, mutantes de pérdida de función resultan letales (Sonoda et al., 1998; Tsuzuki et al., 1996). Aunque en angiospermas la mutación homocigota de RAD51 no muestra defectos durante el desarrollo vegetativo, en *P. patens* las doble mutantes de RAD51 presentan problemas de crecimiento y en el desarrollo de los gametóforos (Markmann-Mulisch et al., 2007); mientras que en el maíz las dobles mutantes presentan un menor número de semillas y los machos son estériles (Li et al., 2007).

Si bien, RAD51 tienen una enorme cantidad de parálogos en diferentes especies, hasta el momento no se ha observado que alguno de estos tenga las mismas funciones que RAD51 (Bonilla et al., 2020; Bleuyard et al., 2005). RAD51 pertenece a la familia de proteína Rec A, la cual ha sufrido una gran diversificación debido a diferentes eventos de duplicación, pero también debido a la transferencia

horizontal de genes (Chintapalli et al., 2013; Hofstatter et al., 2016; Lin et al., 2006). Varios de estos parálogos participan en conjunto con RAD51 durante la alineación de cadenas homólogas (Bonilla et al., 2020; Bleuyard et al., 2005). Uno de los parálogos de RAD51 es DMC1. Sin embargo, se ha visto que RAD51 funciona durante la mitosis y meiosis, mientras que DMC1 únicamente durante la meiosis, es decir, no son funcionalmente redundantes (Kobayashi et al., 2019).

8.5.3 WEE1

La proteína WEE1, se encuentra río abajo de la cascada iniciada por SOG1 y que tiene un papel fundamental en los puntos de chequeo o *checkpoint* del ciclo celular. Los resultados del árbol filogenético mostraron que la secuencia de la gimnosperma *P. sitchensis* es la que mayor distancia presenta con respecto a las otras, incluyendo a *H. sapiens*, **Figura 19A**. Así pues, tanto *P. sitchensis* como *S. moellendorffii* agrupan en un clado distinto al que comprende a las demás plantas. En los resultados arrojados por MEME se puede observar que las secuencias de estas dos especies presentan poca similitud en el extremo N-terminal de la proteína, careciendo de varios motivos que se encuentran ahí en las otras plantas (**Figura 19B**). Cabe resaltar que los mismos motivos que se encuentran ausentes en *S. moellendorffii*, faltan en el alga *Scenedesmus sp.* (Scenedesmaceae); dicho de otra forma, ambas tienen los mismos motivos. De forma extraña, esto mismo ocurre con *H. sapiens* y *Trebouxia sp.* Por último, se puede ver que el motivo 13, en el extremo amino, es exclusivo de Dicotiledóneas, aunque no todas lo presentan. Cabe destacar que mientras que en *Arabidopsis* las mutantes de WEE1 no muestran defectos en el crecimiento (De Schutter, 2007), en *Solanum lycopersicum* sí se altera el desarrollo (Gonzalez et al., 2007) y en ratones, la pérdida de función de WEE1 es letal embrionaria (Tominaga et al., 2006).

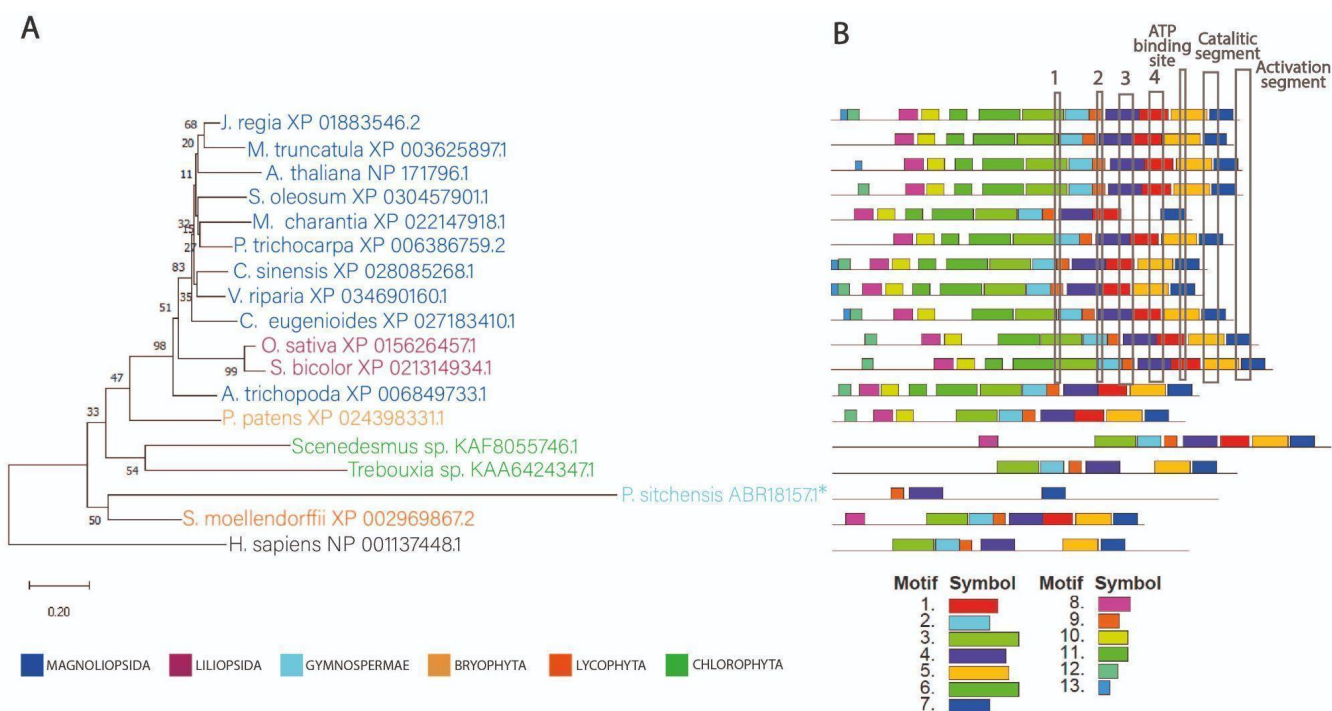


Figura 19. Filogenia WEE1. A) Árbol de NJ con las secuencias obtenidas por BLASTp; No. de replicaciones de Bootstrap = 10000. B) Motivos encontrados con MEME suite y sus respectivas secuencias. Los rectángulos grises marcan los dominios proteicos de WEE1 tomando como referencia la secuencia de *Arabidopsis thaliana* (Brassicaceae). El asterisco indica secuencias no anotadas como WEE1. 1- ATP binding site, 2- ATP binding site, 3- Catalytic segment, 4- Activation segment.

9 Discusión general

En el presente trabajo se evaluó la similitud de diferentes proteínas asociadas a DDR en plantas. La **Tabla 2** muestra la conservación de las vías de reparación con respecto a la presencia/ausencia de las proteínas involucradas. Resalta la falta de varios ortólogos en gimnospermas, inclusive en proteínas que se encuentran en grupos más basales, como en briofitas. No obstante, no se puede descartar que esto se deba más bien a la ausencia de información, dada la poca información que se tiene sobre las gimnospermas. También, las licofitas y las algas destacan por la ausencia de algunos ortólogos, hecho que puede ser tanto efecto de la falta de información, como de la ausencia del homólogo (como es el caso de SOG1 en algas). One Thousand Plant Transcriptomes Initiative (2019) reportó que de 119 especies de algas pertenecientes a Chlorophyta, sólo 12 presentaron duplicaciones

de genoma completo (WGD, por sus siglas en inglés) en su ancestría, mientras que ninguna duplicación fue observada en *Selaginella* y algunas hepáticas. Dado que las WGD se han relacionado con innovaciones en la función de los genes (Force et al., 1999; Ohno, 1970), esto podría estar correlacionado con la falta de algunos ortólogos en las algas y licofitas aquí usadas. Sumado a esto, la mayoría de las expansiones de familias de genes en Archaeplastida se han inferido entre los orígenes de Viridiplantae y de las plantas vasculares (One Thousand Plant Transcriptomes Initiative, 2019). Además, para tener más información es importante tomar en cuenta las características (ciclo de vida, hábitat, etc.) y las formas de respuesta de los diferentes organismos.

Aunque las vías de respuesta a estrés genotóxico están, en general, bien conservadas, los mecanismos empleados y las consecuencias del daño varían entre organismos. Por ejemplo, los efectos del daño a DNA en animales contrastan con los de las plantas, causando defectos en el desarrollo y en algunos casos, cáncer; mientras que en plantas afecta mayormente la producción de biomasa (Caplin & Willey, 2018). En concordancia con esto, se ha visto que células BY-2 de tabaco pueden aguantar mayor cantidad de DBS provocados por la radiación gamma, que las células CHO-K1 de ovario de hámster a las dosis letales 10.5 Gy y 1.4 Gy, respectivamente (Yokota et al., 2005).

En plantas, dentro de las angiospermas, la respuesta a daño a DNA varía, por ejemplo, las monocotiledóneas presentan mayor resistencia a la luz ultravioleta que la dicotiledóneas (Musil, 1995; Pal et al., 1997; Pal et al., 2006); en las primeras se acumula más CPDs (cyclobutane-pyrimidine dimers) ante la exposición a luz UV-B y, dentro de las dicotiledóneas, las plantas leñosas responden mejor que las herbáceas (Wang et al., 2016). Variaciones en la sensibilidad a radiación ionizante también se han observado entre ambos grupos, ejemplo de esto es la hipersensibilidad de los pinos de la especie *Pinus sylvestris*; los cuales murieron ante el accidente de Chérbobil en la zona de exclusión y fueron reemplazados por especies más tolerantes a la radiación, como pastos, arbustos y árboles caducifolios (Ludovici et al., 2020). Sumado a esto, en estudios de laboratorio, se ha observado que *A. thaliana* tolera más la radiación gamma que *P. sylvestris* (Blagojevic et al.,

2019). Aunque el incidente de Chérbobil representa un caso donde los organismos se han visto expuestos a una alta dosis de radiación crónica, la cual ha generado defectos en el desarrollo y la reproducción de las plantas de la zona, es incierto que parte de estos son producto directo de la radiación y cuales son causados por efectos indirectos, como el cambio en las interacciones bióticas (Mousseau, Møller & Pape, 2020). No obstante, Santos et al. (2019) reportaron que el efecto negativo de la radiación sobre las plantas de Chernobyl es menor al efecto positivo del abandono de la tierra.

Como ya se mencionó anteriormente, tanto las briofitas como las algas, presentan respuestas diferentes ante los estresores genotóxicos (Čížková et al., 2019; Vágnerová et al., 2017) y estas diferencias pueden ir acompañadas de variación en la función de los genes, como es el caso del complejo MRN en Briofitas (Kamisugi et al., 2012) o la ausencia de genes como *SOG1* en las algas (Yoshiyama et al., 2014). Además las diferencias en las tasas de evolución de cada taxa, son factores importantes que pueden afectar la variación en estas vías, se ha visto que tanto las gimnospermas como las briofitas tienen tasas de evolución más lentas que las angiospermas (Buschiazzo et al., 2012; Linde et al., 2021). Sería interesante que futuros estudios evaluaran la respuesta y sensibilidad a un factor que produzca DSB (por ejemplo, radiación ionizante) entre plántulas de las especies aquí utilizadas para los análisis.

En el presente trabajo encontramos que, dentro de las angiospermas, éstas presentan una alta similitud, a excepción de algunos motivos específicos que no se encontraban en todas. Los motivos conservados en todas o la mayoría de las secuencias evaluadas podrían ser aminoácidos importantes para su función, lo que concuerda con el hecho de que muchos están dentro de dominios reportados. Por otra parte, los motivos variables podrían estar asociados a una función o regulación específica de algún organismo. Será importante evaluar con métodos funcionales si las diferencias en los motivos encontrados entre monocotiledóneas y dicotiledóneas en ATRIP, *SOG1*, MRE11, NBS1 y RAD51, son relevantes para la función de la proteína y en consecuencia generen alguna característica particular en la capacidad de respuesta de un determinado organismo.

Al comparar los resultados entre angiospermas y los demás grupos, se observaron mayores diferencias. Dentro de las plantas con semilla se observó que las proteínas diferían mucho entre angiospermas y gimnospermas. De igual manera, en general, hubo grandes diferencias con las plantas sin semilla y las algas, lo cual también puede estar ligado con las diferencias fisiológicas entre estos organismos. No sólo la variación en las secuencias influye en los mecanismos de regulación de las plantas, si no que otros factores tales como el número de copias génicas, el ciclo de vida, la historia evolutiva y biogeográfica de cada taxa pueden influir en las adaptaciones ante el estrés. Las plantas conquistaron el medio terrestre hace aprox. 500 Ma (Schreiber, Rensing & Gould, 2022) y con la transición del medio acuático al hábitat terrestre, los ancestros de las embriofitas tuvieron que adaptarse a diferentes condiciones, algunas de las cuales implicaban daño a su DNA (Rensing et al., 2008). Más aún, la alta concentración de CO₂ en la atmósfera comparada con el medio acuático conllevó un incremento del Oxígeno intracelular y, en consecuencia, mayor fotorespiración y producción de especies reactivas de Oxígeno (ROS) (Delwiche & Cooper, 2015). El hecho de que no se hayan encontrado en algas algunas proteínas de DDR homólogas o con los mismos motivos que las plantas terrestres puede reflejar esto.

Tabla 3. Se muestra la conservación de las proteínas en el clado Viridiplantae. Los guiones indican que no se encontraron secuencias ortólogas en ese grupo.

Vía	Proteína	Magnoliopsida	Liliopsida	Gimnospermas	Lycophyta	Bryophyta	Algas
SSB	ATR	✓	✓	----	✓	✓	✓
	RAD1	✓	✓	----	✓	✓	✓
	RAD9	✓	✓	----	✓	✓	✓
	RPA1	✓	✓	----	✓	✓	✓
	ATRIP	✓	✓	----	----	✓	----
NHEJ	Ku70	✓	✓	✓	✓	✓	✓
	Ku80	✓	✓	✓	✓	✓	✓
HR	ATM	✓	✓	----	✓	✓	✓
	SOG1	✓	✓	✓	?	✓	----
	MRE11	✓	✓	----	✓	✓	✓
	NBS1	✓	✓	----	----	✓	----
	RAD50	✓	✓	----	✓	✓	✓
	BRCA1	✓	✓	----	✓	----	✓
	BRCA2	✓	✓	----	✓	----	✓
	RAD51	✓	✓	✓	✓	✓	✓
	WEE1	✓	✓	?	✓	✓	✓

10 Conclusiones

Las proteínas aquí evaluadas que participan en los puntos iniciales de la respuesta a ruptura de doble cadena (ATM y ATR) están altamente conservadas dentro de Viridiplantae.

Para algunas proteínas que participan en puntos más específicos dentro de la vía de respuesta a DSB, como SOG1, BRCA1 y BRCA2 no se encontraron ortólogos en todos los organismos; no obstante, sí en la mayoría de los clados.

La alta conservación de ku70 y ku80 indica que la vía de recombinación no homóloga (NHEJ) está conservada en Viridiplantae.

Aunque las proteínas de las angiospermas comparten gran cantidad de motivos, existen algunos motivos específicos de proteínas en monocotiledóneas, así como en dicotiledóneas.

Los motivos de las proteínas que participan en la vía de respuesta a daño de DNA están menos conservados entre Lycophyta, Briophyta y Chlorophyta, así como entre estos tres grupos y las angiospermas. Asimismo, existen motivos específicos para cada uno de estos clados.

En gimnospermas, hay un hueco de información que limita el estudio de la conservación de las proteínas en este grupo.

11 Perspectivas

Los resultados de este trabajo aportan información sobre la conservación de las proteínas que participan en la respuesta a ruptura de doble cadena de diferentes especies. Sin embargo, hace falta descifrar si las diferencias en los motivos tienen una relevancia estructural o funcional. Para lo cual será importante seleccionar proteínas claves que presenten diferencias en los motivos y realizar ensayos funcionales, complementando mutantes de *Arabidopsis* de pérdida de función con las secuencias proteicas de los organismos seleccionados.

De igual manera, en un futuro sería interesante poder obtener más información acerca de las secuencias de gimnospermas o de otras especies que no hayan sido utilizadas en este trabajo.

12 Referencias

- Adachi, S., Minamisawa, K., Okuschima, Y., Inagaki, S., Yoshiyama, K., Kondoub, Y., Kaminumac, E., Kawasaki, M., Toyoda, T., Matsui, M., Kurihara, D., Matsunaga, S. & Umeda, M. (2011) Programmed induction of endoreduplication by DNA double-strand breaks in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 108(24), 10004-10009.
- Aklilu, B. & Culligan, K. (2016) Molecular evolution and functional diversification of replication protein A1 in plants. *Frontiers in Plant Science*, 7(33).
- Akutsu, N., Iijima, K., Hinata, T. & Tauchi, H. (2007) Characterization of the plant homolog of Nijmegen breakage syndrome 1: Involvement in DNA repair and recombination. *Biochemical and Biophysical Research Communications*, 353, 394-398.
- Amiard, S., Gallego, M. & White, C. (2013) Signaling of double strand breaks and deprotected telomeres in arabidopsis. *Frontiers in Plant Science*, 4, 405.
- Andreassen, P. R., Seo, J., Wiek, C. & Hanenberg, H. (2021) Understanding BRCA2 function as a tumor suppressor based on domain-specific activities in DNA damage responses. *Genes*, 12, 1034.
- Aoyagi, Y., Kusumi, J., & Satake, A. (2021). Copy number analyses of DNA repair genes reveal the role of poly(ADP-ribose) polymerase (PARP) in tree longevity. *iScience*, 24(7), 102779.
- Arnold, P. A., Kruuk, L. E. B. & Nicotra, A. B. (2019) How to analyse plant phenotypic plasticity in response to a changing climate. *New Phytologist*, 222(3), 1235-1241.
- Ayora, S., Carrasco, B., Cárdenas, P. P., César, C. E., Cañas, C., Yadav, T., Marchisone, C. & Alonso, J. C. (2011). Double-strand break repair in bacteria: a view from *Bacillus subtilis*. *FEMS microbiology reviews*, 35(6), 1055–1081.
- Bailey, T. L., Boden, M., Buske, F., Frith, M., Grant, C., Clementi, L., Ren, J., Li, W. & Noble, W. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*, 37(2), W202–W208. <https://doi.org/10.1093/nar/gkp335>
- Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. (2015) The MEME suite. *Nucleic Acids Research*, 43(W1), W39–W49.
- Bétermier, M., Borde, V. & Villartay, J. (2019) Coupling DNA Damage and Repair: an Essential Safeguard during Programmed DNA Double-Strand Breaks? *Trends in Cell Biology*, 30(2), 87-96. <https://doi.org/10.1016/j.tcb.2019.11.005>
- Blagojevic, D., Lee, Y., Brede, D. A., Lind, O. C., Yakolev, I., Solhaug, K. A., Fossdal, C. G., Salbu, B. & Olsen, J. E. (2019) Comparative sensitivity to gamma radiation at the organismal, cell and DNA level in young plants of Norway spruce, Scots pine and *Arabidopsis thaliana*. *Planta*, 250, 1567-1590.
- Bleuyard, J., Gallego, M., Savigny, F. & White, C. (2005) Differing requirements for the Arabidopsis Rad51 paralogs in meiosis and DNA repair. *Plant Journal*, 41, 533-545.
- Bonilla, B., Hengel, S. R., Grundy, M. K. & Bernstein, K. (2020) RAD51 Gene Family Structure and Function. *Annual Review of Genetics*, 54, 25-46.
- Bowles, A. M. C., Williamson, C. J., Williams, T. A., Lenton, T. M. & Donogue, P. C. J. (2023) The origin and early evolution of plants. *Trends in Plant Science*, 28(3), 312-329.
- Burke, R., Schwarze, J., Sherwood, O., Jnaid, Y. McCabe, P. & Kacprzyk, J. (2020) Stressed to Death: The Role of Transcription Factors in Plant Programmed Cell Death Induced by Abiotic and Biotic Stimuli. *Frontiers in Plant Science*, 11.
- Buschiazzo, E., Ritland, C., Bohlmann, J., Ritland, K., (2012). Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. *BMC Evolutionary Biology*, 12(8).
- Caldecott, K. (2008) Single-strand break repair and genetic disease. *Nature Reviews Genetics*, 9(8), 619-631.

- Calderón Vélez, J. C. (2006) Proteínas: Redefiniendo algunos conceptos. *Revista de la Facultad de Medicina*, 54(2), 143-147.
- Caplin, N. & Willey, N. (2018) Ionizing radiation, higher plants, and radioprotection: From acute high doses to chronic low doses. *Frontiers in Plant Science*, 9.
- Chintapalli, S., V., Bhardwaj, G., Babu, J., Hadjiyianii, L., Hong, Y., Todd, G. K., Boosalis, C., Zhang, Z., Zhou, X., M, H., Anishkin, A., van Rossum, D. R. & Patterson, R. (2013) Reevaluation of the evolutionary events within recA/RAD51 phylogeny. *BMC Genomics*, 14(240).
- Čížková, M., Slavková, M., Vítová, M., Zachleder, V. & Bišová, K. (2019) Response of the Green Alga *Chlamydomonas reinhardtii* to the DNA Damaging Agent Zeocin. *Cells*, 8(7), 375.
- Connelly, J.C. and Leach, D.R. (1996) The sbcC and sbcD genes of *Escherichia coli* encode a nuclease involved in palindrome inviability and genetic recombination. *Genes Cells*, 1, 285–291.
- Corlett, R.T. (2016) Plant diversity in a changing world: status, trends, and conservation needs. *Plant Diversity*, 38, 10–16
- Culligan, K. Robertson, C., Foreman, J., Doerner, P. & Britt, A. (2006) ATR and ATM play both distinct and additive roles in response to ionizing radiation. *Plant Journal*, 48(6), 947-961.
- Culligan, K., Tissier, A. & Britt, A. (2004) ATR regulates a G2-phase cell-cycle checkpoint in *Arabidopsis thaliana*. *Plant Cell*, 16, 1091-1104.
- De Bustos, A., Pérez, R. & Jouve, N. (2007) Characterization of the gene Mre11 and evidence of silencing after polyploidization in *Triticum*. *Theoretical and Applied Genetics*, 114, 985-999.
- Deng, C. & Brodie, S. G. (2000) Roles of BRCA1 and its interacting proteins. *BioEssays*, 22, 728-737.
- Delwiche, C. F. & Cooper, E. D. (2015). The Evolutionary Origin of a Terrestrial Flora. *Current biology* 25(19), R899–R910.
- Davis, L., & Maizels, N. (2014). Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. *Proceedings of the National Academy of Sciences of the United States of America*, 111(10), E924–E932.
- Delwiche, C. F. & Timme, R. E. (2011) Plants. *Current Biology*, 21(11), 417-422.
- De Schutter, K., Joubès, J., Cools, T., Verkest, A., Corellou, F., Babiychuk, E., Van Der Schueren, E., Beeckman, T., Kushnir, S., Inze, D. & De Veylder, L. (2007) *Arabidopsis* WEE1 kinase controls cell cycle arrest in response to activation of the DNA integrity checkpoint. *Plant Cell*, 19(1), 211-225.
- De Smet, R., Adams, K. L., Vandepoele, K., Van Montagu, M. C. E., Maere, S. & Van De Peer, Y. (2013) Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, 110(8), 2898-2903.
- DiRuggiero, J. & Robb, T. (2007) Early Evolution of DNA repair mechanisms. En L. Ribas (ed.), *The Genetic Code and the Origin of Life* (pp. 169-182). Kluwer Academic/ Plenum Publishers.
- Doherty, A. J., Jackson, S. P. & Weller, G. R (2001) Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS letters*, 500(3), 186–188.
- Eisen, J. & Hanawalt, P. (1999) A phylogenomic study of DNA repair genes, proteins, and processes. *Mutation Research.*, 435(3), 171–213.
- Fausser, F., Smichl, S. & Puchta, G. (2014) Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant Journal*, 79(2), 348-359.
- Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L. & Postlethwait (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 151, 1531–1545 .
- Friesner, J. D., Liu, B., Culligan, K. & Britt, A. B. (2005). Ionizing radiation– dependent γ -H2AX focus formation requires ataxia telangiectasia mutated and ataxia telangiectasia mutated and Rad3-related. *Molecular Biology of the Cell* 16, 16(5), 2566–2576.
- Fulcher, N. & Sablowsky, R. (2009) Hypersensitivity to DNA damage in plant stem cell niches. *Proceedings of the National Academy of Sciences of the United States of America*, 106(49), 20984-20988.

- Furukawa, T., Curtis, M., Tominey, C., Duonga, Y., Wilconx, B., Aggouneb, D., Hays, J. & Britt, A. (2010) A shared DNA-damage-response pathway for induction of stem-cell death by UVB and by gamma irradiation. *DNA Repair*, 9(9), 940-948.
- Gentric, N., Genschik, P. & Noir, S. (2021) Connections between the cell cycle and the dna damage response in plants. *International Journal of Molecular Sciences*, 22(17), 9558.
- Giglia-Mari G, Zotter A, Vermeulen W. (2011) DNA damage response. *Cold Spring Harb Perspect Biol*. 3(1), a000745
- Gonzalez, N., Gévaudant, F., Hernould, M., Chevalier, C. & Mouras, A. (2007) The cell cycle-associated protein kinase WEE1 regulates cell size in relation to endoreduplication in developing tomato fruit. *Plant Journal*, 51(4), 642-55.
- Goode, E., Ulrich, C. & Potter, J. (2002) Polymorphisms in DNA Repair Genes and Associations with Cancer Risk. *Cancer Epidemiology, Biomarkers and Prevention*, 11(12), 1513–1530.
- Gratani, L. (2014) Plant Phenotypic Plasticity in Response to Environmental Factors. *Advances in Botany*, 2014, 1-17.
- Gull, A., Ahmad, A. & Ul, N. (2019) Biotic and Abiotic Stresses in Plants. En Bosco de Oliveira, A. (Ed.) *Abiotic and Biotic Stress in Plants*. IntechOpen.
- Hartley, C. L. & McCulloch, R. (2008) *Trypanosoma brucei* BRCA2 acts in antigenic variation and has undergone a recent expansion in BRC repeat number that is important during homologous recombination. *Molecular Microbiology*, 68(5), 1237–1251.
- Hefferin, M. L. & Tomkinson, A. E. (2005). Mechanism of DNA double-strand break repair by non-homologous end joining. *DNA repair*, 4(6), 639–648.
- Hofstatter, P. G., Tice, A. K., Kang, S., Brown, M. & Lahr, D. J. G. (2016) Evolution of bacterial recombinase a (recA) in eukaryotes explained by addition of genomic data of key microbial lineages. *Proceedings of the Royal Society B: Biological Sciences*, 283(20161453).
- Hu, Y.-F., Miyake, T., Ye. Q. & R. Li. (2000). Characterization of a novel trans-activation domain of BRCA1 that functions in concert with the BRCA1 C-terminal (BRCT) domain. *Journal of Biological Chemistry*, 275(52),40910–40915.
- Hu, Z., Cools, T., De Veylder, L. (2016) Mechanisms Used by Plants to Cope with DNA Damage. *Annual Review of Plant Biology*, 67, 4.1–4.24.
- Huelsenbeck, J.P., & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics*, 17, 754-755.
- Jiao, Y., Wickett, N., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., Tomsho, L. P., Hu, Y., Liang, H., Soltis, P. S., Soltis, D. E., Clifton, S. W., Schlarbaum, S. E., Schuster, S. C., Ma, H., Leebens-Mack, J. & Depamphilis, D. W. (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature* 473, 97–100.
- Kabbage, M., Kessens, R., Bartholomay, L. & Williams, B. (2017) The Life and Death of a Plant Cell. *Annual Review of Plant Biology*, 68, 375-404.
- Kamisugi, Y., Schaefer, D., Kozak, J., Charlot, F., Vrielynck, N., Holá, M., Angelis, K., Cumberg, A., Nogué, F. (2012) MRE11 and RAD50, but not NBS1, are essential for gene targeting in the moss *Physcomitrella patens*. *Nucleic Acids Research*, 40(8), 3496-3510.
- Kimura, S. & Sakaguchi, K. (2006) DNA repair in plants. *Chemical Reviews*, 106(2), 753-766.
- Kobayashi, W., Liu, E., Ishii, H., Matsunaga, S., Schlögelhofer, P. & Kurumizaka, H. (2019) Homologous pairing activities of *Arabidopsis thaliana* RAD51 and DMC1. *Journal of Biochemistry*, 165(3), 289-295.
- Kojic, M., Kostrub, C., Buchman, A. & Holloman, W. (2002) BRCA2 homolog required for proficiency in DNA repair, recombination, and genome stability in *Ustilago maydis*. *Molecular Cell*, 10, 683-691.
- Komatsu, K. (2016) NBS1 and multiple regulations of DNA damage response. *Journal of Radiation Research*, 57(S1), i11-i17.
- Kroemer, G., Galluzzi, L., Vandenabeele, P., Abrams, J., Alnemri, E.S., Baehrecke, E.H., Blagosklonny, M.V., El-Deiry, W.S., Golstein, P., Green, D.R., Hengartner, M., Knight, R.A., Kumar, S., Lipton, S.A., Malorni, W., Nunez, G., Peter, M.E., Tschopp, J., Yuan, J., Piacentini,

- M., Zhivotovsky, B. & Melino, G., (2009). Classification of cell death: recommendations of the Nomenclature Committee on Cell Death. *Cell Death and Differentiation*, 16, 3–11.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Lafarge, S. & Montané, M. (2003) Characterization of *Arabidopsis thaliana* ortholog of the human breast cancer susceptibility gene 1: *AtBRCA1*, strongly induced by gamma rays. *Nucleic Acids Research*, 31(4), 1148-1155.
- Lehmann, A. & Taylor, E. (2003) Conservation of Eukaryotic DNA repair mechanisms. En J. Nickoloff & M. Hoekstra (Eds.) *DNA Damage and Repair* (pp. 377-401). Humana Press Inc.
- Levine, A.J. (2020) p53: 800 million years of evolution and 40 years of discovery. *Nature Reviews Cancer*, 20, 471–480.
- Li, J., Harper, L., Golubovskaya, I., Wang, R., Weber, D., Meeley, R., McElver, J., Bowen, B., Cande, Z. & Schnable, P. (2007) Functional analysis of maize RAD51 in meiosis and double-strand break repair. *Genetics*, 176, 1469-1482.
- Lin, Z., Hong, H., Nei, M. & Ma, H. (2006) Origins and evolution of the recA/RAD51 gene family: Evidence for ancient gene duplication and endosymbiotic gene transfer. *Proceedings of the National Academy of Sciences of the United States of America*, 103(27), 10328-10333.
- Li, L., Wang, S., Wang, H., Sahu, S. K., Marin, B., Li, H., Xu, Y., Liang, H., Li, Z., Cheng, S., Rede, T., Cebi, Z., Wittek, S., Petersen, M., Melkonian, B., Du., H., Yang, H., Wang, J., Wong, G. K. S., Xu, X., Liu, X., Van de Peer, Y., Melkonian, M. & Liu, H. (2020) The genome of *Prasinoderma coloniale* unveils the existence of a third phylum within green plants. *Nature Ecology & Evolution*, 4, 1220–1231.
- Linde, A., Eklund, M., Cronberg, N., Bowman, J. & Lagercrantz, U. (2021) Rates and patterns of molecular evolution in bryophyte genomes, with focus on complex thalloid liverworts, Marchantiopsida. *Molecular Phylogenetics and Evolution*, 165(107295).
- Liu, Q., Liu, P., Ji, T., Zheng, L., Shen, C., Ran, S., Liu, J., Zhao, Y., Niu, Y., Wang, T. & Dong, J. (2022) The histone methyltransferase SUV2 promotes DSB repair via chromatin remodeling and liquid–liquid phase separation. *Molecular Plant*, 15, 1157-1175.
- Liu, L., Qiu, L., Zhu, Y., Luo, L., Han, X., Man, M., Li, F., Ren, M. & Xing, Y. (2023) Comparisons between Plant and Animal Stem Cells Regarding Regeneration Potential and Application. *International Journal Molecular Sciences*, 24(5), 4392.
- Locato, V., & De Gara, L. (2018). Programmed Cell Death in Plants: An Overview. *Methods in molecular biology* (Clifton, N.J.), 1743, 1–8.
- Lou, D.I., McBee, R.M., Le, U.Q., Stone, A. C., Wilkerson, G. K., Demogines, A. M. & Sawyer, S. (2014) Rapid evolution of BRCA1 and BRCA2 in humans and other primates. *BMC Evol Biol* 14, 155.
- Ludovici, G. M., Oliveira de Souza, S., Chierici, A., Cascone, M. G., D'Errico, F. & Malizia, A. (2020) Adaptation to ionizing radiation of higher plants: From environmental radioactivity to chernobyl disaster. *Journal of Environmental Radioactivity*, 222, 106375.
- Macovei, A., Donà, M., Carbonera, D. & Balestrazzi, A. (2016) Plant Response to Genotoxic Stress: A Crucial Role in the Context of Global Climate Change. En Tuteja, N. & Gill, S. S. (Eds.) *Abiotic stress and response* (pp 13-26). Wiley-Vch.
- Mahapatra, K. & Roy, S. (2020) An insight into the mechanism of DNA response in plants- role of SUPPRESSOR OF GAMMA RESPONSE 1: An overview. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 819-820.
- Manova, V. & Gruzka, D. (2015) DNA damage and repair in plants- from models to crops. *Frontiers in Plant Science*, 6, 885.
- Markmann-Mulisch, U., Hadi, M.Z., Koepchen, K., Alonso, J.C., Russo, V.E.A., Schell, J. & Reiss, B. (2002) The organization of *Physcomitrella patens* RAD51 genes is unique among eukaryotic organisms. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 2959–2964.

- Markmann-Mulish, U., Wendeler, E., Zobel, O., Schween, G., Steinbiss, H. & Reiss, B. (2007) Differential requirements for RAD51 in *Physcomitrella patens* and *Arabidopsis thaliana* development and DNA damage repair. *Plant Cell*, 9, 3080-3089.
- Mareri, L., Parrotta, L. & Cai, G. (2022) Environmental Stress and Plants. *International Journal of Molecular Sciences*, 23(10), 5416..
- Martens, M., Horres, R., Wendeler, E. & Reiss, B. (2020) The importance of ATM and ATR in *Physcomitrella patens* DNA damage repair, Development, and Gene Targeting. *Genes*, 11(7), 752.
- Metha, A. & Haber, J. (2014) Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harbor Perspectives in Biology*, 6(9), a016428. .
- Mosseu, T. A. & Møller, A. P. (2020) Plants in the light of ionizing radiation: What have we learned from chernobyl, fukushima, and other “hot” places?. *Frontiers in Plant Science*, 11, 552.
- Mouriz, A., López-González, L., Jarillo, J. & Piñeiro, M. (2015) PHDs govern plant development. *Plant Signaling and Behavior*, 10(7).
- Musil, C. F. (1995) Differential effects of elevated ultraviolet-B radiation on the photochemical and reproductive performances of dicotyledonous and monocotyledonous arid environment ephemerals. *Plant, Cell & Environment*, 18, 844-854.
- Nakajima, K., Zhou, Y., Tomita, A., Hirade, Y., Gurumurthy, C. B. & Nakada, S. (2018). Precise and efficient nucleotide substitution near genomic nick via noncanonical homology-directed repair. *Genome research*, 28(2), 223–230.
- Nisa, M., Huang, Y., Benhamed, M. & Raynaud, C. (2019) The plant DNA damage response: Signaling pathways leading to growth inhibition and putative role in response to stress conditions. *Frontiers in Plant Science*, 10, 653.
- Núñez-Resendiz, M., Senties, A., Dreckmann, K., Meave, M. y Zamudio-Resendiz, M. (2019) *Sistemática Filogenética aplicada a la Ficología Marina*. Universidad Autónoma Metropolitana, México, Ciudad de México. 104 pp.
- Ogita, N. Okushima, Y., Tokizawa, M., Yamamoto, Y., Tanaka, M., Seki, M., Makita, Y., Matsui, M., Okamoto-Yoshiyama, K., Sakamoto, T., Kurata, T., Hiruma, K., Saijo, Y., Takahashi, N. & Umeda, M. (2018) Identifying the target genes of SUPPRESSOR OF GAMMA RESPONSE 1, a master transcription factor controlling DNA damage response in Arabidopsis. *Plant Journal*, 94(3), 439-453.
- Olvera-Carrillo, Van-Bel, M., Van-Haytegem, T. Ferndrych, M., Huysmans, M., Simaskova, M., Van-Durme, M., Buscaill, P., Rivas, S., Cool, N., Coppens, F., Maere, S. & Nowack, M. (2015) A conserved core of programmed cell death indicator genes discriminates developmentally and environmentally induced programmed cell death in plants. *Plant Physiology*, 169(4), 2684-2699.
- Ohno, S. (1970) *Evolution by Gene Duplication*. Springer-Verlag, New York. 160 pp.
- One Thousand Plant Transcriptomes Initiative. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, 574, 679–685.
- Pal M, Sharma A, Abrol YP, Sengupta UK (1997) Exclusion of UV-B radiation from normal solar spectrum on the growth of mung bean and maize. *Agriculture, ecosystems & environment*, 61, 29–34.
- Pal M, Zaidi PH, Voleti SR, Raj A (2006) Solar UV-B exclusion effects on growth and photosynthetic characteristics of wheat and pea. *Journal of New Seeds*, 8,19–34.
- Pagano, A., Gualtieri, C., Mutti, G., Raveane, A., Sincinelli, F., Semino, O., Balestrazzi, A. & Macovei, A. (2022) Identification and Characterization of SOG1 (Suppressor of Gamma Response 1) Homologues in Plants Using Data Mining Resources and Gene Expression Profiling. *Genes*, 13(667).
- Paull, T. T. (2018) 20 Years of Mre11 Biology: No End in Sight. *Molecular Cell*, 71, 419-427.
- Pérez, R., Cuadrado, A., Jouve, N. & De Bustos, A. (2011) Characterization of the Nbs1 gene and analysis of the expression of homologous and homoeologous MRN complex genes in

- meiocytes and somatic cells of different wheat species. *International Journal of Plant Sciences*, 172(8), 959-69.
- Pfeffer, C., Ho, B. & Singh, A. (2017) The evolution, functions and applications of the breast cancer genes BRCA1 and BRCA2. *Cancer Genomics and Proteomics*, 14(15), 293-298.
- Proost, S., Pattyn, P., Gerats, T. & Van de Peer, Y. (2011) Journey through the past: 150 million years of plant genome evolution. *Plant Journal: for cell and molecular biology*, 66(1), 58–65.
- Que, Q., Chen, Z., Kelliher, T., Skibbe, D., Dong, S. & Chilton, M. (2019) Plant DNA Repair Pathways and Their Applications in Genome Engineering. *Methods in molecular biology*, 1917, 3–24.
- Rangtong, G. & Gunawardena, A. (2015) Programmed cell death: Genes involved in signaling, regulation, and execution in plants and animals. *Botany*, 93(4), 193-210.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. & Huelsenbeck, J. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3), 539–542.
- Roy, R., Chun, J. & Powell, S. N. (2016) BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nature Review Cancer*, 12(1), 68-78.
- Rupnik, A., Grenon, M. & Lowndes, N. (2008) The MRN complex. *Current Biology*, 18, R455–R457
- Rubnik, A., Lowndes, N. F. & Grenon, M. (2010) MRN and the race to the break. *Chromosoma*, 119, 115-135.
- Saitou, N. & Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, 4(4), 406-425
- Sakamoto, A. N., Sakamoto, T., Yokota, Y., Teranishi, M., Yoshiyama, K. O. & Kimura, S. (2021) SOG1, a plant-specific master regulator of DNA damage responses, originated from nonvascular land plants. *Plant Direct*, 5(e370).
- Santos, P., Sillero, N., Boratyński, Z. & Teodoro, A. C. M. (2019) Landscape changes at Chernobyl. En Christopher; M. U. N. & Antonino, M, (Eds.) *Remote Sensing for Agriculture, Ecosystems, and Hydrology XXI* (pp. 509-526). SPIE.
- Saldivar, J., Cortez, D. & Cimprich, K. (2017) The essential kinase ATR: Ensuring faithful duplication of a challenging genome. *Nature Reviews Molecular Cell Biology*, 18(10), 622-636
- Seeliger, K., Dukowic-Schulze, S., Wurz-Wildersinn, R., Pacher, M. & Puchta, H. (2012) BRCA2 is a mediator of RAD51- and DMC1-facilitated homologous recombination in *Arabidopsis thaliana*. *New Phytologist*, 193(2), 364-375.
- Schimpl, S., Fauser, F. & Puchta, H. (2016) Repair of adjacent single-strand breaks is often accompanied by the formation of tandem sequence duplications in plant genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 113(26), 7266-7271.
- Schmidt, C., Pacher, M. & Puchta (2019) DNA break repair in plants and its application for Genome Engineering. En S. Kumar, P. Barone & M. Smith (Eds.) *Transgenic Plants: Methods and Protocols* (pp 237-266). Springer Nature.
- Schreiber, M., Rensing, S. A. & Gould, S. B. (2022) The greening ashore. *Trends in Plant Science*, 27(9), 847-857.
- Schröpfer, S., Knoll, A., Trapp, O. & Puchta, H. (2014) DNA repair and recombination in plants. En S. Howell (Ed.) *Molecular Biology, The Plant Sciences 2* (pp. 51-93). Springer Nature.
- Shaik, A. & Kirubakaran, S. (2020) Evolution of PIKK family kinase inhibitors: A new age cancer therapeutics. *Frontiers in Bioscience*, 25, 1510-1537.
- Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P-F., Lindquist, E. A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin, T., Kuroki, Y., Toyoda, A., Theodoulou, F. L., Verrier, P. J., Quatrano, R. S. & Boore, J. L. (2008) The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science*, 319(5859), 64-69.
- Siaud, N., Dray, E., Gy, I., Gérard, E., Takvorian, N. & Doutriaux, M. (2004) Brca2 is involved in meiosis in *Arabidopsis thaliana* as suggested by its interaction with Dmc1. *EMBO Journal*, 23(6), 1392-1401.

- Singh, S., Roy, S., Choudhury, S. & Sengupta, D. (2010) DNA repair and recombination in higher plants: insights from comparative genomics of arabidopsis and rice. *BMC Genomics*, 11(1).
- Spampinato, C. P. (2017) Protecting DNA from errors and damage: an overview of DNA repair mechanisms in plants compared to mammals. *Cellular and Molecular Life Sciences*, 74(9), 1693-1709.
- Studier, J. A., & Keppler, K. J. (1988). A note on the neighbor-joining algorithm of Saitou and Nei. *Molecular Biology and Evolution*, 5(6), 729–731.
- Szurman-Zubrzycka, M., Jędrzejek, P. & Szarejko, I. (2023) How Do Plants Cope with DNA Damage? A Concise Review on the DDR Pathway in Plants. *International Journal of Molecular Sciences*, 24(3), 2404.
- Tang, H., Bowers, J. E., Wang, X. & Paterson, A. H. (2010) Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. *Proceedings of the National Academy of Sciences of the United States of America*, 107(1), 472–477.
- Trapp, O., Seeliger, K. & Puchta, H. (2011) Homologs of breast cancer genes in plants. *Frontiers in Plant Science*, 2, 19.
- Tominaga, Y., Li, C., Wang, R. H., Deng, C. X. (2006) Murine Wee1 plays a critical role in cell cycle regulation and pre-implantation stages of embryonic development. *Int J Biol Sci*, 2(4), 161-70.
- Ueno, S., Sudo, T. & Hirasawa, A. (2022) ATM: Functions of ATM Kinase and Its Relevance to Hereditary Tumors. *International Journal of Molecular Sciences*, 23(523).
- Vágnerová, R., Lukešová, A., Lukeš, M., Rožnovská, P., Holá, M., Fulnečková, J., Fajkus, J. Angelis, K. J. (2017). Evolutionarily Distant Streptophyta Respond Differently to Genotoxic Stress. *Genes*, 8(11), 331.
- Wang, Q. Kamiyama, C., Hidema, J. & Hikosaka, K. (2016) Ultraviolet-B-induced DNA damage and ultraviolet-B tolerance mechanisms in species with different functional groups coexisting in subalpine moorlands. *Oecologia*, 181(4), 1069-1082.
- Waterworth, W., Drury, G., Bray, C. & West, C. (2011) Repairing breaks in the plant genome: the importance of keeping it together. *New Phytologist*, 192, 805–822.
- Wolter, F., Schindele, P., Beying, N., Scheben, A. & Puchta, H. (2021) Different DNA repair pathways are involved in single-strand break-induced genomic changes in plants. *The Plant Cell*, 33(11), 3454-3469.
- Yokota, Y., Shikazono, N., Tanaka, A., Hase, Y., Funayama, T., Wada, S. & Inoue, M. (2005) Comparative radiation tolerance based on the induction of DNA double-strand breaks in tobacco BY-2 cells and CHO-K1 cells irradiated with gamma rays. *Radiation Research*, 163(5), 520-525.
- Yoshiyama, K. O., Kobayashi, J., Ogita, N., Ueda, M., Kimura, S., Maki, H. & Umeda, M. (2013) ATM-mediated phosphorylation of SOG1 is essential for the DNA damage response in Arabidopsis. *EMBO Reports*, 14(9), 817-822.
- Yoshiyama, K. O., Sakaguchi, K. & Kimura, S. (2013) DNA damage response in plants: Conserved and variable response compared to animals. *Biology*, 2(4), 1338-1356. <https://doi.org/10.3390/biology2041338>
- Yoshiyama, K. O., Kimura, S., Maki, H., Britt, A. & Umeda, M. (2014) The role of SOG1, a plant-specific transcriptional regulator, in the DNA damage response. *Plant Signaling and Behavior*, 9(4), e28889.
- Yoshiyama, K. O. (2016) SOG1: a master regulator of the DNA damage response in plants. *Genes and Genetic Systems*, 90(4), 209-216.
- Yoshiyama, K. O., Kaminoyama, K., Sakamoto, T. & Kimura, S. (2017) Increased Phosphorylation of Ser-Gln Sites on SUPPRESSOR OF GAMMA RESPONSE1 Strengthens the DNA Damage Response in Arabidopsis thaliana. *Plant Cell*, 29, 3244-3268.
- Yoshiyama, K. O. & Kimura, S. (2018) Ser-Gln sites of SOG1 are rapidly hyperphosphorylated in response to DNA double-strand breaks. *Plant Signaling and Behavior*, 13(6). e1477904 (4 pages)