

**IDENTIFICACIÓN Y CARACTERIZACIÓN DE ISOFORMAS DE LA ENZIMA
DESATURASA DE ÁCIDOS GRASOS 2 (FAD2) EN *Ricinus communis* MEDIANTE
MÉTODOS MOLECULARES Y COMPUTACIONALES - PRIMERA FASE**

JHON ALEXANDER VÉLEZ ZAPATA

Trabajo de grado realizado para optar al título de Biólogo

Diego Fernando Villanueva Mejía, Biólogo, M.Sc. Ph.D.

Director del trabajo de grado

Danna Melissa Susunaga Gómez, Bióloga, M.Sc.

Co-directora del trabajo de grado

Universidad EAFIT

Escuela de Ciencias Aplicadas e Ingeniería

Medellín, Colombia

2022

Contenido

1. PLANTEAMIENTO DEL PROBLEMA:	3
2. JUSTIFICACIÓN	4
3. OBJETIVOS	5
3.1. Objetivo general	5
3.2. Objetivos específicos	5
4. MARCO TEÓRICO	5
5. METODOLOGÍA	9
5.1 Búsqueda <i>in silico</i> de isoformas del gen <i>FAD2</i> en bases de datos públicas	9
5.2 Material vegetal	9
5.3 Cultivo <i>in vitro</i> de <i>R. communis</i>	9
5.4 Extracción de DNA de hojas de <i>R. communis</i>	10
5.5 Diseño de <i>primers</i> orientados al gen <i>FAD2</i> de <i>R. communis</i>	10
5.6 Estandarización de la PCR para la amplificación del gen <i>FAD2</i> de <i>R. communis</i>	11
5.7 Secuenciación y modelación de <i>FAD2</i> de <i>R. communis</i>	12
6. RESULTADOS	13
6.1 Búsqueda de isoformas del gen <i>FAD2</i>	13
6.2 Cultivo <i>in vitro</i> de <i>R. communis</i>	18
6.3 Extracción de DNA de hojas de <i>R. communis</i>	20
6.4 Diseño de <i>primers</i> orientados al gen <i>FAD2</i> de <i>R. communis</i>	22
6.5 Estandarización de la PCR para la amplificación del gen <i>FAD2</i> de <i>R. communis</i>	27
6.6 Secuenciación y modelación de <i>FAD2</i> en <i>R. communis</i>	33
7. DISCUSIÓN	40
7.1 Búsqueda <i>in silico</i> de isoformas del gen <i>FAD2</i>	40
7.2 Cultivo <i>in vitro</i> de <i>R. communis</i>	41
7.3 Extracción de DNA de hojas de <i>R. communis</i>	42
7.4 Diseño de <i>primers</i> orientados al gen <i>FAD2</i> de <i>R. communis</i>	43
7.5 Estandarización de PCR para la amplificación del gen <i>FAD2</i> de <i>R. communis</i>	43
7.6 Secuenciación y modelación de <i>FAD2</i> de <i>R. communis</i>	44
8. CONCLUSIONES	45
9. PERSPECTIVAS	45
10. BIBLIOGRAFÍA	45

1. PLANTEAMIENTO DEL PROBLEMA:

La higuierilla o ricino (*Ricinus communis*) es una planta de la familia Euphorbiaceae, nativa de Abisinia o Etiopía (Perry, 1943; Sujatha et al., 2008). Esta especie es muy cultivada en varios países como India y China por su fácil crecimiento en regiones tropicales y subtropicales (Sujatha et al., 2008). La higuierilla almacena en sus semillas grandes cantidades de aceite, llamado aceite de ricino, el cual está compuesto en su mayor parte por ácido ricinoleico, insumo vital en la industria química por ser una materia prima ideal para muchas sustancias como pigmentos, plásticos, pinturas, lubricantes, resinas, solventes y fertilizantes gracias a sus propiedades físicas y químicas únicas que favorecen gran cantidad de procesos (Ogunniyi, 2006). También se fabrican con él productos de alto valor añadido como anticongelantes, líquidos hidráulicos y fibras ópticas, entre otros materiales (Sujatha et al., 2008). Junto con el ácido ricinoleico, que compone un 89.15% del aceite de ricino (Berman et al., 2011), hay en menor cantidad otros tipos de ácidos grasos como el ácido oleico y el ácido linoleico, los cuales pueden acelerar el proceso de ranciamiento del aceite cuando se almacena, y reducir su vida útil, lo que hace necesario un proceso de purificación para enriquecer el ácido ricinoleico antes de llevarlo a sus múltiples aplicaciones (Ogunniyi, 2006).

Para omitir este proceso e incrementar la proporción de ácido ricinoleico en el aceite obtenido a partir de las semillas de ricino, estudios anteriores desarrollados por científicos del grupo de investigación CIBIOP de la Universidad EAFIT han buscado crear variedades de *R. communis* genéticamente modificadas enfocadas en maximizar la síntesis de ácido ricinoleico por encima de los otros lípidos (Figueroa-Varela, 2020; Susunaga-Gomez, 2022). Se centraron estos esfuerzos específicamente en la ruta metabólica de la planta que fabrica diversos ácidos grasos a partir del ácido oleico. Para promover la conversión de ácido oleico en oleoyl-CoA y ultimadamente en ácido ricinoleico, se intentó bloquear el gen de la enzima desaturasa FAD2 (Fatty Acid Desaturase 2), la cual se ubica en la membrana del cloroplasto y del retículo endoplasmático, y transforma el ácido oleico en ácido linoleico, desviándolo por una ruta alterna no deseada y reduciendo la proporción de ácido ricinoleico en el aceite (Figueroa-Varela, 2020; Miquel & Browse, 1992).

La edición genética en este estudio fue exitosa a nivel molecular (datos en proceso de publicación), pero también incluyó una serie de efectos secundarios en el desarrollo vegetativo *in vitro* de la planta, como una elongación lenta de las semillas germinadas. Los investigadores del grupo CIBIOP han hipotetizado que esto puede deberse a dificultades en las plantas para sintetizar suficientes lípidos en el momento de construir sus membranas celulares, por lo que se requiere realizar otros estudios y desarrollos que permitan soportar o refutar esta hipótesis (Figueroa-Varela, 2020). A pesar de esto, lo interesante es que el gen *FAD2* ya ha sido editado en otras especies como soya (Al Amin et al., 2018, 2019; Wu et al., 2020), algodón (Chen et al., 2021; Liu et al., 2002, 2017) y arroz (Abe et al., 2018; Bahariah et al., 2021), estudios en los que la producción de ácidos grasos fue aumentada exitosamente y se lograron obtener nuevas variedades viables.

De ser cierta esa hipótesis, los investigadores del grupo CIBIOP buscarían permitir la expresión de *FAD2* en las etapas tempranas y bloquearla en las tardías en las semillas, por medio del silenciamiento selectivo de las isoformas de *FAD2* que existan en la planta. Para ello, se requiere distinguir sus propiedades fisicoquímicas, su actividad enzimática, su expresión en distintos tejidos y edades, y en consecuencia, apagar solo algunas de las versiones del gen *FAD2*, evadiendo aquellas involucradas directamente en el desarrollo temprano y enfocándose en aquellas que actúan principalmente en las semillas (Figueroa-Varela, 2021).

Dar et al., (2017) clasifican el *FAD2* de las plantas oleaginosas en 4 isoformas, las cuales se encuentran presentes en muchas especies y, aunque sus secuencias comparten una alta similitud, se diferencian en sus funciones metabólicas (Kongcharoensuntorn, 2001). Al mismo tiempo, estas enzimas demuestran una expresión diferencial, es decir, son sintetizadas en distintas cantidades dependiendo del tejido, órgano y estadio de desarrollo de la planta (Dar et al., 2017; Xue et al., 2017). No obstante, hasta este momento no se ha realizado un estudio que determine cuáles son las isoformas de *FAD2* en *R. communis*, además no se cuenta con un genoma completo ensamblado sino un grupo de *scaffolds*. Es por esto que en este trabajo de investigación, se pretendió caracterizar las isoformas de la enzima desaturasa *FAD2* en *R. communis*, en qué órganos de la planta se expresan en mayor medida y cuáles son las diferencias entre la función biológica de cada una de las isoformas. Se hipotetiza que *R. communis* tiene las mismas 4 isoformas de *FAD2* reportadas para sus especies afines, presenta una expresión diferencial en tejidos de hoja, tallo, raíz y semilla, así como en los distintos estadios de desarrollo de la planta, similar a la reportada en sus especies afines y su acción enzimática predicha se ajusta a los requerimientos fisiológicos de cada tejido y estadio de desarrollo.

En esta primera fase del proyecto se abarcó la identificación y caracterización del gen *FAD2* mediante métodos computacionales y experimentales, lo cual sentará las bases para una segunda fase, en la que se abordará el estudio de las isoformas de *FAD2* a nivel de expresión diferencial.

2. JUSTIFICACIÓN

El hecho de apenas poder obtener un 89.15% de ácido ricinoleico del aceite de ricino extraído de las semillas de *R. communis* significa que se pierden recursos en términos de tiempo y dinero cada que se requiere pasar por un proceso de purificación del ácido ricinoleico, cada que hay que descartar los otros ácidos grasos del extracto y cada que se ve reducido el tiempo de vida útil del aceite debido a la presencia de PUFAs. Si se pudiera contar con un inventario de las isoformas de *FAD2* en *R. communis*, se tendría la posibilidad de fijar nuevos blancos de un proceso mejoramiento genético, de forma que se produzcan líneas vegetales con una mayor probabilidad de ofrecer todas las ventajas de las plantas *wild type* junto con las de la edición. Esta línea podría incluso ser lo suficientemente competente como para introducirse

en el mercado. De esta forma se beneficiarían las industrias involucradas en cuanto a producción, refinamiento, almacenamiento y procesamiento del ácido ricinoleico. Este trabajo de investigación también representa una oportunidad para la formación académica y científica de otras personas como jóvenes investigadores estudiantes de pregrado que participan del desarrollo de la revisión bibliográfica, procedimientos y análisis de datos. La comunidad científica internacional también será beneficiada por la contribución que realiza este trabajo en términos de aporte al conocimiento de la especie como modelo, profundizando en la estructura del gen *FAD2*, la expresión de la proteína y sus isoformas, sentando bases valiosas para futuras investigaciones en mejoramiento genético de *R. communis* y otras plantas oleaginosas.

3. OBJETIVOS

3.1. Objetivo general

Caracterizar el gen de la enzima desaturasa de ácidos grasos 2 (*FAD2*) en *Ricinus communis* de la variedad VERC-03 a partir de métodos moleculares, así como sus posibles isoformas usando métodos computacionales.

3.2. Objetivos específicos

1. Identificar las posibles isoformas de *FAD2* en *R. communis* contrastando con datos existentes en bases de datos públicas y empleando métodos *in silico*.
2. Validar los métodos de cultivo *in vitro* de *R. communis* estandarizados en estudios anteriores.
3. Estandarizar métodos moleculares para la amplificación del gen *FAD2* a partir de DNA de *R. communis* óptimo para secuenciación por el método de Sanger.
4. Predecir, caracterizar y comparar *in silico* los modelos 3D de la enzima *FAD2* de referencia y de la variedad VERC-03.

4. MARCO TEÓRICO

Ricinus communis tiene varios nombres comunes, como son higuera, castor, ricino, tártao, higuera infernal, mosquitera, entre otros (Lim, 2016). Está clasificada como una hierba, en su estado adulto puede llegar a medir 4 metros de altura, posee entrenudos de color azul, sus hojas son verdes y sus márgenes son ásperas (Langeland et al., 2008). A pesar de que se cultiva por su aceite, en muchos países también se le considera una planta altamente invasora. Por su rápido crecimiento y su alta capacidad de reproducirse, dispersarse y establecerse en nuevos ambientes puede poner en riesgo el ecosistema, apoderándose del espacio y los recursos de las especies nativas (Gordon et al., 2010). *R. communis* es más conocida por su facultad de expresar una proteína de tipo albúmina extremadamente tóxica para los humanos

llamada ricina, que se encuentra en las semillas más no en el aceite. La ricina, si es ingerida por medio de sus semillas, puede traer síntomas como dolor abdominal, vómito y diarrea y eventualmente generará daño a los riñones y al hígado antes de causar la muerte (Ogunniyi, 2006; Soto-Blanco et al., 2002).

R. communis es aprovechada para ser cultivada en ambientes pobres donde es difícil sembrar otros cultivos, ya que posee resistencia a la sequía y se adapta fácilmente a diversos ambientes, sobre todo los tropicales (Baquero, 2017). Por esta razón, países como China, India y Brasil han promovido su cultivo y han tratado de mejorar su productividad por medio de la investigación científica (Sánchez et al., 2016; Susunaga-Gomez, 2022).

Las semillas de *R. communis* contienen un aproximado de 46-55% del peso total en aceite, más que muchas otras especies como el girasol, la soja y el algodón, las cuales presentan porcentajes de 38%, 40% y 15% respectivamente (Solís Bonilla et al., 2017). A su vez, el aceite está compuesto en un 89% de ácido ricinoleico, 4.2% de ácido linoleico, 3% de ácido oleico, 1% de ácido esteárico, 1% de ácido palmítico, 0.7% de ácido dihidroxiesteárico, 0.3% de ácido linolénico y 0.3% de ácido eicosanoico (Ogunniyi, 2006). El ácido ricinoleico consiste a nivel molecular en una cadena alifática de carbono (18 C), un grupo carboxilo, un hidroxilo y un enlace doble. La principal limitación del aceite de ricino es que los grupos carboxilo e hidroxilo de las moléculas pueden reaccionar para formar polímeros llamados estólidos. Esto tiene como consecuencia que se reduzcan sus valores de hidroxilo y ácido paulatinamente, de forma que luego de tan solo 90 días de almacenamiento, ya se han reducido ambos valores en un 10% (Ogunniyi, 2006).

En la industria química, el aceite de ricino es apreciado por su contenido de ácido ricinoleico (Mutlu & Meier, 2010), el cual tiene una gran diversidad de aplicaciones debido a la proximidad que existe a nivel molecular entre el grupo hidroxilo y el doble enlace, los cuales le facilitan reaccionar químicamente para transformarse y dar lugar a múltiples productos como materiales funcionales, insumos para la industria y elementos de diversas áreas (Mubofu, 2016; Ramchiary & Kole, 2017; Susunaga-Gomez, 2022). Este aceite tiene más de 700 usos en el mercado internacional, por lo que se producen entre 270.000 y 360.000 toneladas cada año (Mutlu & Meier, 2010; Saadaoui et al., 2017).

El ácido ricinoleico es sintetizado a partir de la siguiente ruta metabólica: el ácido oleico se transforma en oleoyl-CoA por la acción de la enzima acyl-CoA sintetasa, luego es llevado al retículo endoplasmático y se fusiona el oleoyl-CoA con la fosfatidilcolina en la posición 2-^{na} de ésta última, formando 2-oleoyl-PC por la acción de la lisofosfatidilcolina aciltransferasa (LPCAT). Después, el 2-oleoyl-PC se transforma en 2-ricinoleyl-PC por la acción de la oleoyl-12 hidroxilasa (FAH12). Finalmente, el 2-ricinoleyl-PC es dividido en ácido ricinoleico y lisofosfatidilcolina (LPC) con ayuda de la fosfolipasa A2 (PLA2). A partir de este punto, el ácido ricinoleico, que no suele encontrarse en estado puro en la naturaleza, puede seguir la ruta que lo transforma en ricinoleyl-CoA para finalmente pasar a integrar un triglicérido por reacciones de acilación (Lin & Arcinas, 2007). Una vista gráfica de la ruta metabólica del ácido ricinoleico y las enzimas involucradas en su síntesis a partir de ácido oleico se observa en la Figura 1 (Susunaga-Gomez, 2022; L. Wang et al., 2019).

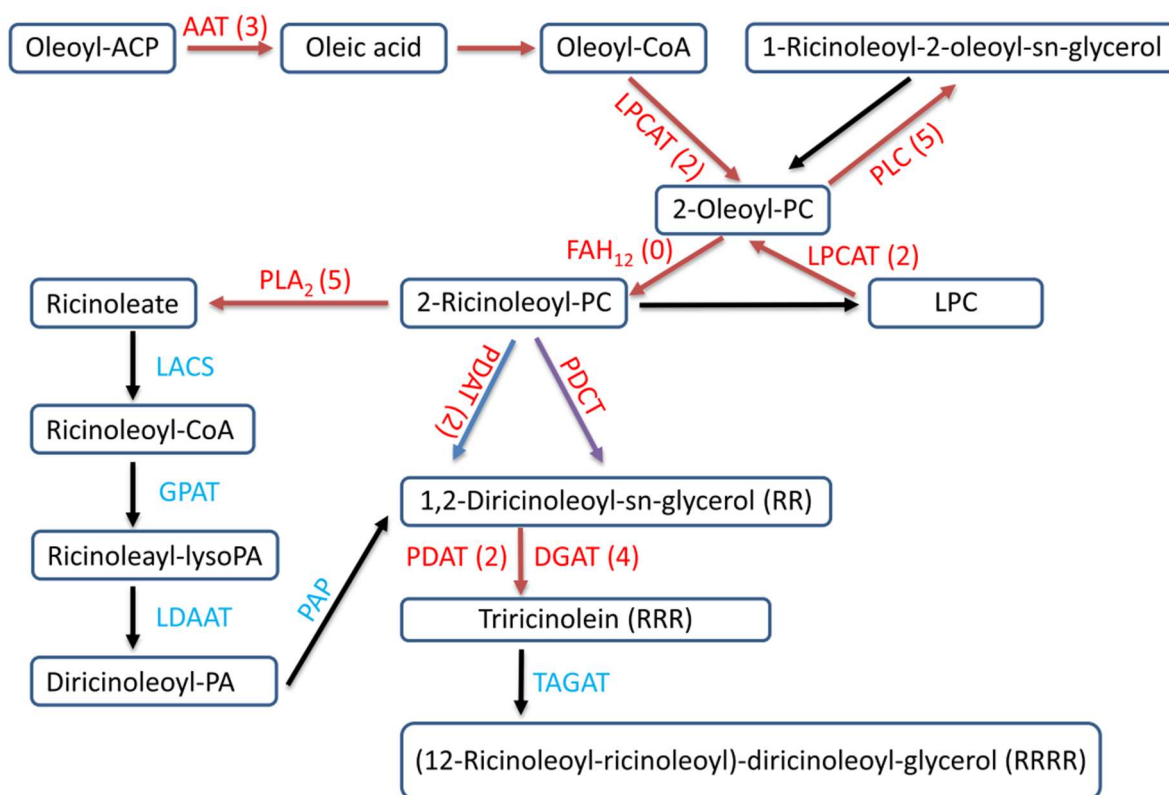


Figura 1. Ruta metabólica para la producción de ácido ricinoleico en *R. communis*. Tomado de L. Wang et al., (2019).

R. communis es una planta hexaploide, cuenta con 10 cromosomas y el tamaño total de su DNA es alrededor de 320 Mb, pero se cuenta realmente con poca información sobre su genoma (Chan et al., 2010). Hasta la fecha, solo se dispone de un ensamblaje incompleto con 25,828 contigs de un tamaño total de 350.6 Mb, los cuales alcanzan una cobertura de alrededor de 4.6x. Por medio de una búsqueda por homología, se logró predecir la existencia de 31,237 genes, distribuidos en 3020 familias proteicas, con un promedio de 4.2 exones por gen (Chan et al., 2010).

La enzima *FAD2* es miembro de la familia de las desaturasas, las cuales llevan el proceso de transformación del ácido esteárico por ácido oleico, linoleico y gamma-linolénico por medio de desaturaciones continuas. A esta familia también pertenecen las enzimas *SAD*, *FAD3*, *FAD6*, *FAD7* y *FAD8*, entre otras (Bhunja et al., 2016; D. Zhang et al., 2009). A su vez, para *FAD2* se han reportado en especies distintas a higuera algunas isoformas como la *FAD2-1* que se expresa principalmente en semillas jóvenes y botones florales, *FAD2-2* que es la principal sintetizadora de ácido linoleico y *FAD2-3* y *FAD2-4* que se expresan uniformemente por todos los tejidos y poseen mucha similitud en sus secuencias (Dar et al., 2017). De los alrededor de 20 motifs que poseen las otras desaturasas, las *FAD2* conservan 6, correspondientes a las regiones transmembranales junto con 3 cajas de histidina, y tienen también 4 residuos de aminoácidos rodeando el sitio activo (Dar et al., 2017). Por

último, su regulación viene estimulada por factores como la temperatura, la luz solar, el oxígeno y las heridas, y está controlada tanto por la intervención de los promotores de cada gen como por los intrones de su mRNA (Dar et al., 2017; Martínez-Rivas et al., 2000; Xiao et al., 2014).

En el estudio realizado por Hernández et al. (2005) se pudo evidenciar una vez más, por medio de un análisis filogenético, que es posible clasificar las isoformas de *FAD2* de una o más especies en las categorías de “mantenimiento del individuo” y “semillas”, lo cual favorece este estudio y le da un motivo para esperar resultados satisfactorios. Comparando con los hallazgos de otros estudios en otras especies, ellos encontraron que, en plantas de olivo, la variedad *FAD2-1* se expresa fuertemente en semillas muy jóvenes y en hojas, con bajos niveles en el mesocarpo. Mientras tanto, el *FAD2-2* se expresó de forma moderada en las semillas durante las etapas más tardías del desarrollo de estas, y en hojas. Al mismo tiempo, altos niveles de ácido linoleico se observaron correlacionados a la alta transcripción de *FAD2-1* en las etapas tempranas y a *FAD2-2* en las etapas más tardías y el mesocarpo (Hernández et al., 2005).

Del mismo modo, otros estudios similares realizados en otras especies como soya, arroz, *Arabidopsis thaliana* e *Hyptage benghalensis* muestran resultados similares, aun empleando metodologías diferentes (Chi et al., 2011; Kishore et al., 2007; Li et al., 2008; Pham et al., 2011; Zhiguo et al., 2019; Zhou et al., 2013).

A la fecha, el grupo de investigación CIBIOP de la universidad EAFIT en Colombia ha aportado también con múltiples avances en la comprensión de este modelo y su potencial, en cuanto a cultivo *in vitro*, regeneración y edición genética de *R. communis*. Se estandarizaron los procesos de desinfección y germinación de las semillas, elongación y propagación de las plántulas, aislamiento y transfección de protoplastos, inducción y propagación de callo, extracción de DNA y RNA de hoja, amplificación de regiones de *FAD2* por PCR convencional y cuantitativa y edición genética por la técnica CRISPR/Cas9 (Figuroa-Varela, 2020; Susunaga-Gomez, 2022). Varios de los resultados y conclusiones de estas investigaciones son útiles para este trabajo y todos los que se realicen en el futuro, pues sientan las bases para lanzarse a profundizar todavía más en el mejoramiento de la especie y su manejo, logrando cada vez más alcance e impacto en el ámbito científico y económico. Por este motivo, los métodos empleados en este trabajo tienen su fundamento en aquellas investigaciones previas, como por ejemplo el uso de medio WPM suplementado con hormonas y nutrientes para cultivar *R. communis*, los protocolos de extracción de DNA de hoja y de amplificación del gen *FAD2* y los programas de PCR (Figuroa-Varela, 2020; Susunaga-Gomez, 2022).

5. METODOLOGÍA

5.1 Búsqueda *in silico* de isoformas del gen *FAD2* en bases de datos públicas

Para la búsqueda de isoformas del gen *FAD2*, se utilizó como genoma de referencia el ensamble con número de accesión ASM1957865v1, el cual fue realizado por el Jardín Botánico de Wuhan. A partir de este, se identificó la posición del gen *FAD2* y sus tres exones a través de la anotación realizada. Luego se hizo un análisis en la base de datos de NCBI en busca de genes homólogos de *FAD2* reportados en otras especies cercanas a *R. communis*. Con estas secuencias se hizo una búsqueda de copias de *FAD2* en el genoma de *R. communis* para encontrar posibles isoformas por medio de la herramienta BLAST de NCBI, optimizando para secuencias más disímiles.

Asimismo, se recopilaron las librerías de DNA y RNA-seq de proyectos referenciados con el fin de alinearlos usando como referencia la secuencia de *FAD2* anotada del ensamble mencionado anteriormente. Esto con el objetivo de buscar posibles variantes o mutaciones en la secuencia de DNA como SNPs o indels que estén presentes en la especie y que puedan marcar una diferencia entre la proteína que expresa la variedad colombiana de *R. communis* (VERC-03) y las demás del resto del mundo. Para los alineamientos se usó el software BWA (REF), para filtrar y ordenar los *reads* se usó Samtools (Danecek et al., 2021) y para encontrar variantes se usó Freebayes (Garrison & Marth, 2012). Adicionalmente, a la secuencia de referencia se le añadieron 300 pb antes y después para facilitar la unión de los *reads*.

5.2 Material vegetal

Se utilizaron semillas de *R. communis* de la variedad VERC-03, provenientes del banco de germoplasma de la Corporación Colombiana de Investigación Agropecuaria (AGROSAVIA). Las condiciones de almacenamiento de las semillas en el Laboratorio de Biotecnología Vegetal de la Universidad EAFIT, Medellín (Colombia) fueron: temperatura de 4°C y oscuridad.

5.3 Cultivo *in vitro* de *R. communis*

Para el cultivo *in vitro* de *R. communis* se emplearon los métodos estandarizados por Gil-Correal et al., (2019). Se seleccionaron 30 semillas sanas y se sumergieron en agua destilada con jabón yodado por una hora. Posteriormente, se les retiró la testa y se dejaron en agua destilada por una hora para hidratar la cutícula. Luego se retiró también la cutícula y se procedió con la desinfección.

Las semillas fueron puestas en un frasco estéril y se lavaron con una solución de HgCl₂ al 0.2% por 3 minutos. Se realizaron 3 lavados consecutivos con agua destilada estéril para retirar el mercurio residual. Luego se sumergieron las semillas en 50 mL de una solución de cefotaxima a 10 mg/L y vancomicina a 5 mg/L por 3 horas. Al final, se lavaron las semillas una vez más con agua destilada estéril y se colocaron sobre papel absorbente estéril para dejarlas secar. Una vez secas, se colocaron en medio de cultivo de germinación, el cual contenía WPM (2,41 g/L), azúcar (30g/L), caseína (100 mg/L), glutamina (50 mg/L), adenina (25 mg/L), arginina (15 mg/L), phytigel (1,8 g/L) y carbón activado (0,8 g/L), suplementado con ácido giberélico (2 mg/L).

En un plazo de entre una y dos semanas, se rescataron las semillas que germinaron sin contaminación visible y se sembraron en un medio de elongación el cual contenía WPM (2,41 g/L), azúcar (30g/L), caseína (150 mg/L), glutamina (50 mg/L), adenina (25 mg/L), arginina (15 mg/L), phytigel (1,8 g/L) y carbón activado (0,5 g/L), suplementado con 6-bencilaminopurina (2 mg/L) y ácido 1-naftalenacético (0,5 mg/L). Este medio tiene como función lograr que la planta crezca y se desarrolle rápidamente. Al cabo de los dos meses, se cortó la plántula por encima de los cotiledones y se sembró este epicótilo en un medio de propagación, el cual contenía WPM (2,41 g/L), azúcar (30g/L), caseína (100 mg/L), glutamina (50 mg/L), adenina (25 mg/L), arginina (15 mg/L), phytigel (1,8 g/L) y carbón activado (0,8 g/L), suplementado con 6-bencilaminopurina (2 mg/L), ácido 1-naftalenacético (0,7 mg/L) y ácido giberélico (0,7 mg/L). Este medio hace que la planta genere una raíz nueva a partir del corte en el epicótilo y que multiplique el número de hojas verdaderas.

5.4 Extracción de DNA de hojas de *R. communis*

Para las extracciones de DNA se usó el método del CTAB (Doyle & Doyle, 1987). Se precalentó el buffer de extracción con CTAB 2X por 20 minutos a 65 °C. Se maceró una porción pequeña de hoja fresca en el mortero con nitrógeno líquido, se depositó una porción pequeña de la muestra homogénea en un tubo estéril, se agregó 400 µL de CTAB y 4 µL de 2-mercaptoetanol. Se agitó y se incubó a 65 °C por 30 minutos. Luego se añadió 300 µL de cloroformo, se agitó un minuto y se centrifugó a 13000 rpm por 15 minutos a 4°C. Luego de centrifugar, se transfirió el sobrenadante a otro tubo y se agregó el mismo volumen en cloroformo, se agitó por un minuto y se centrifugó nuevamente con las mismas condiciones. Se transfirió otra vez el sobrenadante a otro tubo y se agregó el mismo volumen en isopropanol, se agitó y se almacenó al menos 2 horas a -20°C. Pasado este tiempo, se centrifugó a 13000 rpm por 15 minutos a 4 °C. Se descartó el sobrenadante y se hicieron dos lavados con etanol, que consisten en añadir etanol al 70% y centrifugar a 13000 rpm por 10 minutos a 4°C. Luego de los dos lavados se descartó el sobrenadante y se dejó secar el pellet por 20 o 30 minutos. Se resuspendió el DNA en 50 µL de agua libre de nucleasas estéril o en TE. Se agregó 1 µL de RNasa y se incubó a 37°C por una hora. Se almacenó el DNA a -20 °C.

Finalmente, para evaluar la integridad del DNA, se preparó un gel de agarosa al 1,2% y se usa TBE 1X como buffer. Se emplearon 2 µL de *loading dye* y 3 µL de muestra en cada carga. Se corrió la electroforesis con 70 V por 90 minutos. La calidad del DNA fue evaluada mediante la medición en Espectrofotómetro NanoDrop, donde se estimaron las relaciones A260/230 y A260/280, así como la concentración medida en ng/µL.

5.5 Diseño de *primers* orientados al gen *FAD2* de *R. communis*

Para secuenciar el CDS (región codificante) del gen de *FAD2* de *R. communis* se diseñaron tres pares de *primers*, cada par abarcando una región de alrededor de 800 bp, de forma que cubran todo el gen, aunque se sobrelapen. Los *primers* fueron diseñados con el software Primer3 (Kõressaar et al., 2018; Koressaar & Remm, 2007; Untergasser et al., 2012), escogiendo de entre todas las opciones ofrecidas las que tenían una temperatura de melting más similar, una longitud de amplicón adecuada para secuenciación con el método de Sanger, menos estructuras secundarias y menos *mismatches*. Los *mismatches* se analizaron con la

herramienta PrimerBLAST de NCBI (Ye et al., 2012) y las estructuras secundarias se predijeron con OligoAnalyzer (<https://www.idtdna.com/calc/analyzer>). Los *primers* diseñados fueron adquiridos a través de la empresa Purificación y Análisis de Fluidos SAS y se resuspendió el liofilizado hasta alcanzar una concentración de 100 ng/ μ L.

5.6 Estandarización de la PCR para la amplificación del gen *FAD2* de *R. communis*

Para la estandarización de la PCR para amplificar el CDS del gen *FAD2*, se emplearon inicialmente dos *master mix* distintos con el fin de evaluar su eficacia (Tabla 1). Asimismo, se realizó un gradiente de temperaturas de “*annealing*” entre 70°C y 50°C, para abarcar un rango amplio en el que se obtuviera amplificación (Tabla 2).

Tabla 1. Protocolos de PCR v1 y v2 para la amplificación del gen *FAD2*.

Master Mix v1	Volumen para 1 reacción (μ L)
DNA	1
Primer forward	0,5
Primer reverse	0,5
Buffer Taq	2,5
dNTPs	1,25
Taq Polimerasa	0,125
H ₂ O	19,12
Volumen total	25
Master Mix v2	Volumen para 1 reacción (μ L)
DNA	1
Primer forward	0,5
Primer reverse	0,5
Buffer Taq	5
dNTPs	0,75
Taq Polimerasa	0,25
H ₂ O	17
Volumen total	25

Tabla 2. Programa de PCR establecido para la amplificación del gen *FAD2*.

Etapa	Temperatura (°C)	Tiempo (min)	
Desnaturalización inicial	94	2:00	
Desnaturalización	50-68	0:30	x35 ciclos
Alineamiento	72	0:30	
Extensión	72	1:00	
Extensión final	72	5:00	
Enfriamiento	4	∞	

Posteriormente, se realizó una PCR empleando el *master mix* más óptimo con el fin de crear un gradiente de temperaturas para cada par de *primers* diseñados, y así evaluar en qué temperatura amplificaba mejor cada uno, esta vez abarcando un rango más estrecho y más bajo, entre 58°C y 45°C. Habiendo definido una temperatura mínima y máxima en la que se amplifica el tamaño esperado y sin inespecificidades, se repitió el gradiente de temperaturas, esta vez entre el rango que mejores resultados haya arrojado en las PCRs anteriores, para determinar con la mayor resolución posible la temperatura de *annealing* más adecuada.

Luego, para determinar la concentración de cloruro de magnesio óptima, se estableció un rango entre 1 y 4 mM, abarcando 7 tratamientos diferentes. Adicionalmente, se realizó un gradiente de concentración de DNA entre 25 y 125 mg/μL, abarcando 5 tratamientos. Finalmente, se hizo un último gradiente, esta vez variando la concentración de *primers* (entre 0.12 ng/μL y 0.2 ng/μL).

A partir de estos resultados, se diseñó un nuevo protocolo de PCR que unificara las condiciones y concentraciones establecidas para cada una de las variables evaluadas. Este protocolo se aplicó a cada par de *primers*, cada uno con dos repeticiones.

5.7 Secuenciación y modelación de *FAD2* de *R. communis*

A partir del DNA extraído de hojas de cuatro individuos de *R. communis* (réplicas biológicas), se amplificaron las tres regiones del gen *FAD2* con los tres pares de *primers* diseñados. Se ajustó un total de 20 muestras (Tabla 3), las cuales se secuenciaron mediante el método Sanger en la compañía Macrogen (Corea del Sur), mediante prestación de servicios técnicos.

Tabla 3. Distribución de las 20 muestras de DNA de individuos de *R. communis* destinadas a secuenciación por método Sanger, según los *primers* empleados para su amplificación (*primers* A, B y C).

Individuo	Primers empleados para amplificación				
1	A11	A12	B11	B12	C1
2	A21	A22	B2	C21	C22
3	A3	B31	B32	C31	C32
4	A41	A42	B41	B42	C4

Los reads en formato .ab1 se “trimaron” manualmente con el programa Geneious (*Geneious Prime 2022.2.2*, 2022) para retirar los extremos de las secuencias que contienen resultados mixtos en el cromatograma. Para el proceso de alineamiento, identificación de variantes, filtro de variantes de baja calidad y demás análisis de los archivos .ab1 trimados, se usó el software Tracy (Rausch et al., 2020). Para esto se usó como referencia el gen completo *FAD2* de *R. communis* de la anotación del ensamblaje ASM1957865v1, no solo el CDS. Luego se compararon entre sí los reportes de variantes para identificar aquellas que presentaran una calidad buena (Phred score > 10) con el fin de que exista una alta confianza en que la variante es real.

Finalmente, se creó una secuencia consenso para el gen *FAD2* incluyendo las variantes halladas y a partir de ella se modeló su estructura terciaria con ayuda del software Alphafold Colab (Jumper et al., 2021). Se predijo también la ubicación del sitio activo por homología con COACH (Yang et al., 2013b, 2013a) y se hizo análisis de “docking” molecular estimando la afinidad de la enzima *FAD2* de referencia y la de VERC-03 por los sustratos ácido oleico y ácido linoleico, empleando el software Autodock VINA (Eberhardt et al., 2021; Trott & Olson, 2009) con una exhaustividad de 15 y realizando 19 réplicas para cada caso.

6. RESULTADOS

6.1 Búsqueda de isoformas del gen *FAD2*

La tabla 4 muestra la estructura del gen *FAD2* de *R. communis* según la anotación realizada sobre el ensamblaje ASM1957865v1, las posiciones de sus exones, de su CDS y la extensión del mRNA producto de este. El gen *FAD2*, tal y como se reporta en NCBI, se extiende entre las posiciones 18470818 y 18475317 del cromosoma 10 de *R. communis*. En este gen se encuentran tres exones, dos pequeños al principio y uno grande y principal al final, el cual es el único que contiene CDS y se traduce a proteínas. El mRNA puede tomar dos isoformas, una uniendo los exones 1 y 3 y otra uniendo los exones 2 y 3 y recortando una pequeña porción del extremo 3'. El CDS para ambas isoformas de mRNA es exactamente el mismo. Las posiciones exactas de inicio y fin del gen completo, de las dos isoformas de mRNA, de cada uno de los tres exones y del CDS están señaladas en la Tabla 4.

Tabla 4. Anotación del gen *FAD2* en el ensamblaje de referencia ASM1957865v1.

ID	Tipo	Inicio	Fin
NC_063265.1	gene	18470818	18475317
NC_063265.1	mRNA	18470818	18475317
NC_063265.1	exón	18470818	18470951
NC_063265.1	exón	18473913	18475317
NC_063265.1	CDS	18473924	18475075
NC_063265.1	mRNA	18473470	18475265
NC_063265.1	exón	18473470	18473494
NC_063265.1	exón	18473913	18475265
NC_063265.1	CDS	18473924	18475075

De la base de datos de NCBI se tomó 86 secuencias del gen *FAD2* de 14 diferentes especies, incluyendo la propia *R. communis*. Algunas de estas especies reportaban varias secuencias distintas correspondientes a *FAD2*, ya sea por corresponder a isoformas o por haber sido obtenidas de múltiples variedades de esa planta. La Tabla 5 muestra los números de accesoión, la especie y el nombre que se le da al gen, para cada una de las secuencias usadas en el BLAST.

Tabla 5. Secuencias del gen *FAD2* procedente de diversas especies y variedades usadas en el BLAST contra el genoma de *R. communis*.

Accesión	Especie	Gen
>NW_011550910.1:2474783-2475518	<i>Elaeis guineensis</i>	<i>FAD2</i>
>NM_112047.4	<i>Arabidopsis thaliana</i>	<i>FAD2</i>
>NC_038256.1:35315605-35319052	<i>Glycine max</i>	<i>FAD2-1B</i>
>NC_038255.1:40820238-40824909	<i>Glycine max</i>	<i>FAD2-2B</i>
>NC_038255.1:40814706-40815985	<i>Glycine max</i>	<i>FAD2-2A</i>
>NC_038246.1:50013588-50015460	<i>Glycine max</i>	<i>FAD2-1A</i>
>NC_038245.1:c22179154-22174478	<i>Glycine max</i>	<i>FAD2-2D</i>
>NC_038242.1:42049841-42050219	<i>Glycine max</i>	<i>FAD2-3</i>
>NC_025996.1:35241284-35246044	<i>Elaeis guineensis</i>	<i>FAD2</i>
>NC_016090.3:36014749-36020911	<i>Glycine max</i>	<i>FAD2-2</i>
>gi 94982470 gb DQ496227.1	<i>Zea mays</i>	<i>FAD2</i>
>gi 926459552 ref NM_001155650.2	<i>Zea mays</i>	<i>FAD2</i>
>gi 92380878 dbj AB257309.1	<i>Zea mays</i>	<i>FAD2</i>
>gi 8980834 gb AF272952.1	<i>Arachis hypogaea</i>	<i>FAD2</i>
>gi 8980832 gb AF272951.1	<i>Arachis duranensis</i>	<i>FAD2</i>
>gi 8886725 gb AF192486.1	<i>Sesamum indicum</i>	<i>FAD2</i>
>gi 604724428 gb KF741366.1	<i>Arachis hypogaea</i>	<i>FAD2-2</i>
>gi 604724426 gb KF741365.1	<i>Arachis hypogaea</i>	<i>FAD2-1</i>
>gi 60392579 gb AY802997.1	<i>Helianthus annuus</i>	<i>FAD2-2</i>
>gi 60392577 gb AY802996.1	<i>Helianthus annuus</i>	<i>FAD2-2</i>
>gi 60392575 gb AY802995.1	<i>Helianthus annuus</i>	<i>FAD2-2</i>
>gi 60392573 gb AY802994.1	<i>Helianthus annuus</i>	<i>FAD2-2</i>
>gi 60392571 gb AY802993.1	<i>Helianthus annuus</i>	<i>FAD2-2</i>
>gi 60392569 gb AY802992.1	<i>Helianthus annuus</i>	<i>FAD2-1</i>
>gi 60392567 gb AY802991.1	<i>Helianthus annuus</i>	<i>FAD2-1</i>
>gi 60392565 gb AY802990.1	<i>Helianthus annuus</i>	<i>FAD2-1</i>
>gi 60392563 gb AY802989.1	<i>Helianthus annuus</i>	<i>FAD2-1</i>
>gi 60392536 gb AY803008.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392534 gb AY803007.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392532 gb AY803006.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392530 gb AY803005.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392528 gb AY803004.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392526 gb AY803003.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392524 gb AY803002.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392522 gb AY803001.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392520 gb AY803000.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392518 gb AY802999.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>
>gi 60392516 gb AY802998.1	<i>Helianthus annuus</i>	<i>FAD2-3</i>

>gi 599328550 gb KC994458.1	<i>Arachis hypogaea</i>	FAD2
>gi 512127099 gb KC533766.1	<i>Hiptage benghalensis</i>	FAD2-4
>gi 512127097 gb KC533765.1	<i>Hiptage benghalensis</i>	FAD2-3
>gi 512127095 gb KC533764.1	<i>Hiptage benghalensis</i>	FAD2-2
>gi 512127093 gb KC533763.1	<i>Hiptage benghalensis</i>	FAD2-1
>gi 438450 gb L26296.1 ATHD12AAA	<i>Arabidopsis thaliana</i>	FAD2
>gi 322875052 gb HQ215949.1	<i>Arachis hypogaea</i>	FAD2
>gi 322875050 gb HQ215948.1	<i>Arachis hypogaea</i>	FAD2
>gi 322875048 gb HQ215947.1	<i>Arachis hypogaea</i>	FAD2
>gi 306976237 gb GU929424.1	<i>Arabidopsis lyrata</i>	FAD2
>gi 306032966 gb HM462476.1	<i>Arachis hypogaea</i>	FAD2
>gi 306032964 gb HM462475.1	<i>Arachis hypogaea</i>	FAD2
>gi 306032962 gb HM462474.1	<i>Arachis hypogaea</i>	FAD2
>gi 306032960 gb HM462473.1	<i>Arachis hypogaea</i>	FAD2
>gi 297541128 gb HM063066.1	<i>Arachis hypogaea</i>	FAD2
>gi 297540842 gb HM063065.1	<i>Arachis hypogaea</i>	FAD2
>gi 268527772 gb FJ768732.1	<i>Arachis hypogaea</i>	FAD2
>gi 225348405 gb FJ768953.1	<i>Oryza sativa</i>	FAD2
>gi 212275228 ref NM_001136762.1	<i>Zea mays</i>	FAD2
>gi 1898427443 ref XM_022144333.2	<i>Helianthus annuus</i>	FAD2
>gi 1898427442 ref XM_022144332.2	<i>Helianthus annuus</i>	FAD2
>gi 1898419917 ref XM_022141964.2	<i>Helianthus annuus</i>	FAD2
>gi 1898396379 ref XM_022127820.2	<i>Helianthus annuus</i>	FAD2
>gi 1898372801 ref XM_022114524.2	<i>Helianthus annuus</i>	FAD2
>gi 1898367519 ref XM_022181530.2	<i>Helianthus annuus</i>	FAD2
>gi 1898367364 ref XM_022181467.2	<i>Helianthus annuus</i>	FAD2
>gi 1898335447 ref XM_022116691.2	<i>Helianthus annuus</i>	FAD2
>gi 1897976731 ref XM_035967644.1	<i>Zea mays</i>	FAD2
>gi 1897976730 ref XM_020552996.3	<i>Zea mays</i>	FAD2
>gi 18418652 gb AF251844.1	<i>Helianthus annuus</i>	FAD2-3
>gi 18418650 gb AF251843.1	<i>Helianthus annuus</i>	FAD2-2
>gi 18418648 gb AF251842.1	<i>Helianthus annuus</i>	FAD2-1
>gi 1669383105 ref XM_025824044.2	<i>Arachis hypogaea</i>	FAD2
>gi 1666279011 ref XM_025762584.2	<i>Arachis hypogaea</i>	FAD2
>gi 157361223 gb EF687907.1	<i>Zea mays</i>	FAD2
>gi 1217042947 ref XM_021805835.1	<i>Hevea brasiliensis</i>	FAD2
>gi 1217042945 ref XM_021805834.1	<i>Hevea brasiliensis</i>	FAD2
>gi 1216274745 ref XM_021766109.1	<i>Manihot esculenta</i>	FAD2
>gi 1216274743 ref XM_021766108.1	<i>Manihot esculenta</i>	FAD2
>gi 1216274741 ref XM_021766107.1	<i>Manihot esculenta</i>	FAD2
>gi 1216274739 ref XM_021766106.1	<i>Manihot esculenta</i>	FAD2

>gi 1173803395 ref XM_020694947.1	<i>Sesamum indicum</i>	<i>FAD2</i>
>gi 1173803391 ref XM_020694946.1	<i>Sesamum indicum</i>	<i>FAD2</i>
>gi 1173803389 ref XM_011081925.2	<i>Sesamum indicum</i>	<i>FAD2</i>
>gi 1173803387 ref XM_011081924.2	<i>Sesamum indicum</i>	<i>FAD2</i>
>gi 1173789312 ref XM_011076843.2	<i>Sesamum indicum</i>	<i>FAD2</i>
>gi 1063710967 ref NM_001337975.1	<i>Arabidopsis thaliana</i>	<i>FAD2</i>
>gi 1021311890 ref NM_001323719.1	<i>Ricinus communis</i>	<i>FAD2</i>

Para el BLAST se emplearon como *queries* las 86 diferentes secuencias de *FAD2* de distintas especies cercanas y oleaginosas. El software arrojó dos coincidencias en el genoma de *R. communis* para la mayoría de las secuencias de *FAD2* presentadas, una en el cromosoma 10 y otra en el cromosoma 9 (Figura 2). La que se ubica en el cromosoma 10 corresponde al mismo locus de *FAD2* que ya se conoce. La que se ubica en el cromosoma 9 corresponde al gen de la enzima oleato hidrolasa FAH12, la cual tiene cierta similitud, por lo que es identificada por el BLAST como una secuencia parecida. Todas las secuencias de *FAD2* de distintas especies empleadas en la búsqueda tuvieron las mismas dos coincidencias en los mismos locus, por lo que en la figura 2 solo se muestra el resultado que arroja la secuencia de *FAD2* canónica de *R. communis*.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Ricinus communis isolate WT05 ecotype wild-type chromosome 10_ASM1957865v1	Ricinus communis	2078	2078	100%	0.0	100.00%	26623538	NC_063265.1
<input checked="" type="checkbox"/> Ricinus communis isolate WT05 ecotype wild-type chromosome 9_ASM1957865v1	Ricinus communis	541	541	93%	7e-152	71.44%	28210331	NC_063264.1

Figura 2. Resultado general del BLAST con la secuencia de *FAD2* de *R. communis* como *query*.

Las Tablas 6 y 7 presentan las mutaciones o variantes que tienen las otras variedades de *R. communis* sobre el gen *FAD2*, respecto al mRNA secuenciado de la variedad colombiana experimental VERC-03. En estas tablas se identifican de inmediato unos puntos en común en los cuales suceden los mismos cambios en la secuencia. Algunos de estos son 119, 283, 414, 827 y 1478. Estas posiciones se definen teniendo en cuenta que la referencia usada incluye 300 pb (antes y después del gen) para facilitar el alineamiento de los *reads*, por lo que, definiendo como inicio el comienzo del CDS, las posiciones serían -180, -16, 115, 528 y 1179. De estas, solo las posiciones 115 y 528 estarían dentro del CDS.

Tabla 6. Reporte de Freebayes de variantes del gen *FAD2* secuenciado de variedades de *R. communis* reportadas en NCBI en los proyectos PRJNA548999, PRJNA453634, PRJNA414256 y PRJNA513227. #CHROM indica la secuencia de referencia que se está usando para encontrar variantes, POS indica la posición de la referencia en la que ocurre la mutación, REF indica la base o las bases presentes en la referencia, ALT indica la base o las bases en su versión mutada presentes en los *reads* alineados y QUAL indica la confianza percibida sobre la existencia de la variante.

#CHROM	POS	REF	ALT	QUAL
PRJNA548999				
gi 2079614566 gb CM033710.1 :18473624-18475372	119	AAAGC	AAATCAGAAGC	5947.17
gi 2079614566 gb CM033710.1 :18473624-18475372	414	G	A	10434.9
gi 2079614566 gb CM033710.1 :18473624-18475372	827	C	T	8684.02
gi 2079614566 gb CM033710.1 :18473624-18475372	1478	A	G	9413.97
PRJNA453634				
gi 2079614566 gb CM033710.1 :18473624-18475372	283	TTT	CTA	825.196
gi 2079614566 gb CM033710.1 :18473624-18475372	414	G	A	22192.7
gi 2079614566 gb CM033710.1 :18473624-18475372	827	C	T	18129.4
gi 2079614566 gb CM033710.1 :18473624-18475372	1478	A	G	17828.3
gi 2079614566 gb CM033710.1 :18473624-18475372	1643	T	A	0
gi 2079614566 gb CM033710.1 :18473624-18475372	1653	A	C,T	5,76E-09
PRJNA414256				
gi 2079614566 gb CM033710.1 :18473624-18475372	707	C	T	0
PRJNA513227				
gi 2079614566 gb CM033710.1 :18473624-18475372	119	AAAGC	AAATCAGAAGC	1.96e-13
gi 2079614566 gb CM033710.1 :18473624-18475372	414	G	A	1,45E-08
gi 2079614566 gb CM033710.1 :18473624-18475372	827	C	T	0
gi 2079614566 gb CM033710.1 :18473624-18475372	919	G	C	3,27E-09
gi 2079614566 gb CM033710.1 :18473624-18475372	1478	A	G	12686.3

Tabla 7. Reporte de Freebayes de variantes del mRNA de *FAD2* secuenciado de la variedad de *R. communis* reportada en NCBI en el proyecto PRJNA453634. #CHROM indica la secuencia de referencia que se está usando para encontrar variantes, POS indica la posición de la referencia en la que ocurre la mutación, REF indica la base o las bases presentes en la referencia, ALT indica la base o las bases en su versión mutada presentes en los *reads* alineados y QUAL indica la confianza percibida sobre la existencia de la variante.

#CHROM	POS	REF	ALT	QUAL
PRJNA453634				
gi 2079614566 gb CM033710.1 :18473624-18475372	119	AAAGC	AAATCAGAAGC	979.241
gi 2079614566 gb CM033710.1 :18473624-18475372	266	CAC	CC	0
gi 2079614566 gb CM033710.1 :18473624-18475372	283	TTT	CTA	0
gi 2079614566 gb CM033710.1 :18473624-18475372	414	G	A	136855
gi 2079614566 gb CM033710.1 :18473624-18475372	827	C	T	114258
gi 2079614566 gb CM033710.1 :18473624-18475372	1478	A	G	65172.1
gi 2079614566 gb CM033710.1 :18473624-18475372	1608	T	C	0
gi 2079614566 gb CM033710.1 :18473624-18475372	1651	T	C	1,40E-09

6.2 Cultivo *in vitro* de *R. communis*

Luego de la desinfección, las semillas de *R. communis* fueron colocadas en medios de germinación, tal como se muestra en la Figura 3.



Figura 3. Semillas de *R. communis* VERC-03 en medio de germinación suplementado con ácido giberélico, luego de pasar por el proceso de desinfección.

Las semillas tardaron en germinar entre una y dos semanas, luego se transfirieron aquellas que no resultaron contaminadas y que germinaron exitosamente a medios de elongación para que desarrollaran su raíz y tallo como se observa en la Figura 4. Las semillas que pudieron germinar adecuadamente y sin contaminación fueron entre 20% y 30% del total de semillas desinfectadas, teniendo en cuenta que cada ciclo de desinfección incluía 30-35 semillas.



Figura 4. Plántula de *R. communis* VERC-03 en medio de elongación suplementado con 6-bencilaminopurina y ácido 1-naftalenacético.

Después de que la plántula alcanzaba una altura de unos 15 cm y tocaba la tapa del frasco, se transfería su epicótilo a medios de propagación para que generara hojas como insumo para las posteriores extracciones de DNA (Figura 5).



Figura 5. Plántula de *R. communis* VERC-03 en medio de propagación WPM suplementado con 6-bencilaminopurina, ácido 1-naftalenacético y ácido giberélico.

6.3 Extracción de DNA de hojas de *R. communis*

La extracción inicial de DNA proveniente de hojas de *R. communis* con el método CTAB (descrito anteriormente) permitió obtener una banda intensa y con poca degradación (Figura 6). Asimismo, la evaluación de la calidad del DNA mediante espectrofotometría (NanoDrop) mostró valores altos y óptimos para concentración y relaciones A260/280 y A260/230, respectivamente, como se muestra en la Tabla 8.

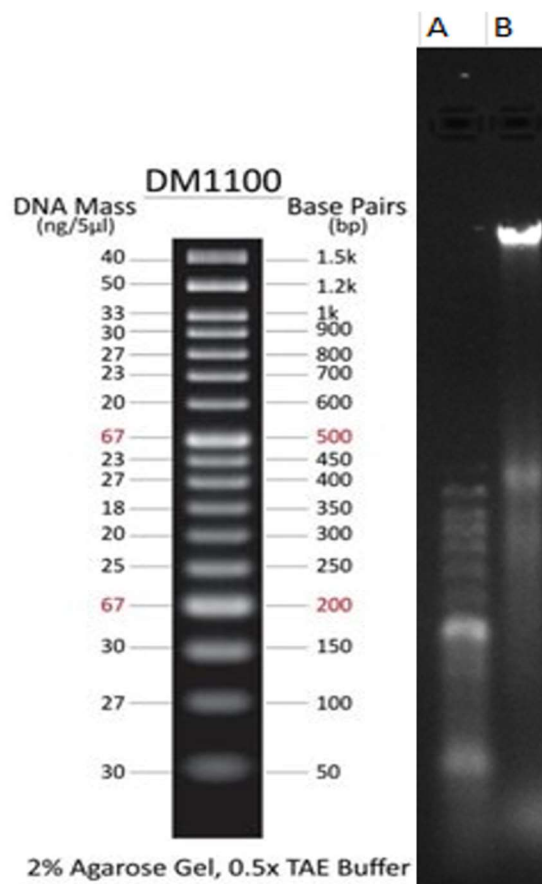


Figura 6. Prueba de integridad del DNA genómico de hojas de *R. communis* obtenido en la primera extracción. Carril A: Ladder 50 bp. Carril B: DNA genómico.

Tabla 8. Lectura de espectrofotometría en Nanodrop del DNA genómico de hojas de *R. communis*, obtenido en la primera extracción.

MUESTRA	CONCENTRACIÓN (ng/ μ L)	260/280	260/230
Hojas de <i>R. communis</i> (Primera extracción)	1296.0	1.75	1.16

Al realizar una segunda extracción de DNA, esta vez en hojas de cuatro individuos de *R. communis*, se observó una buena integridad del DNA genómico (Figura 7).

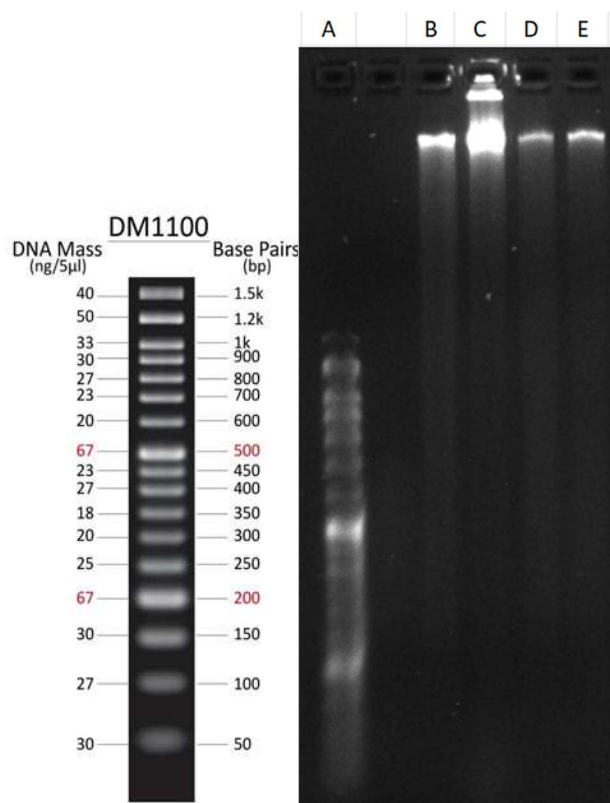


Figura 7. Prueba de integridad del DNA extraído de las 4 réplicas biológicas para la secuenciación del gen *FAD2* de *R. communis*. Carril A: Ladder 50 bp. Carril B: Individuo 1. Carril C: Individuo 2. Carril D: Individuo 3. Carril E: Individuo 4.

Por su parte, el espectrofotómetro mostró estadísticas favorables en las relaciones A260/280 y A260/230, aunque los individuos 3 y 4 muestran unas relaciones A260/230 por debajo de lo ideal, que es 1 (Tabla 9).

Tabla 9. Lecturas de espectrofotometría con Nanodrop para cada réplica biológica para la secuenciación del gen *FAD2* de *R. communis*.

MUESTRA	CONCENTRACIÓN (ng/µL)	A260/280	A260/230
Individuo 1	356	1.81	1.04
Individuo 2	1409	1.94	1.50
Individuo 3	592	1.82	0.88
Individuo 4	246	1.77	0.86

6.4 Diseño de *primers* orientados al gen *FAD2* de *R. communis*

Los tres pares de *primers* diseñados en Primer3 se nombraron A (primera parte del gen), B (segunda parte) y C (tercera parte). Los *primers* A generan un amplicón de 836 bp, los B generan uno de 785 bp y los C generan uno de 847 bp (Tabla 10). La ubicación de los *primers* en el gen *FAD2* de *R. communis* se muestra gráficamente en la Figura 8.

Tabla 10. Secuencias de *primers* empleados para la amplificación completa del gen *FAD2* de *R. communis*.

PRIMERS	SECUENCIA
AF	TTCCTGTGTAAGAAACTGCCT
AR	CCAAGTGTAAGTGTGACGGC
BF	TTCCACCTCCTTCCTGAACC
BR	GCCTTTGTAGCCTCCATTGC
CF	TTCAATGTTTCAGGCAGGCC
CR	AACCAAATTGACGGATGCGC

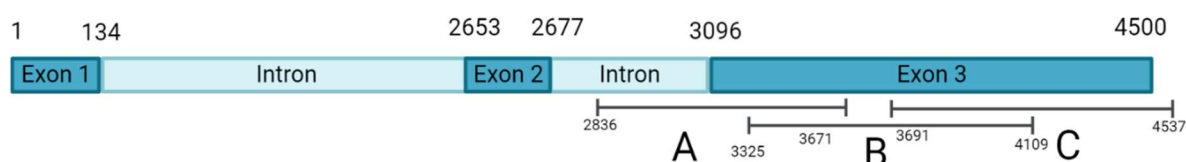


Figura 8. Representación gráfica de las regiones amplificadas con los *primers* diseñados en el gen *FAD2* de *R. communis*. Las dimensiones de cada exón, intrón y amplicón no están a escala.

El programa Primer3 arrojó los resultados que se muestran en la Figura 9. Como se observa, los tres pares de *primers* presentan estadísticas favorables predichas. La T_m es similar en todos, los porcentajes de guanina y citosina son adecuados y el tamaño de los amplicones que generan es idóneo para las secuenciaciones por el método de Sanger.

PRIMERS A

```

OLIGO      start  len  tm  gc%  any_th  3' th  hairpin  seq
LEFT PRIMER  231   21  57.14  42.86  0.00  0.00  0.00  TTCCTGTGTAAGAAACTGCCT
RIGHT PRIMER 1066  20  59.13  55.00  0.00  0.00  0.00  CCAAGTGTAAGTGTGACGGC
SEQUENCE SIZE: 2149
INCLUDED REGION SIZE: 2149

```

PRODUCT SIZE: 836, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

PRIMERS B

```

2 LEFT PRIMER  720  20  58.94  55.00  0.00  0.00  0.00  TTCCACCTCCTTCCTGAACC
RIGHT PRIMER  1504  20  59.54  55.00  0.00  0.00  0.00  GCCTTTGTAGCCTCCATTGC
PRODUCT SIZE: 785, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

```

PRIMERS C

```

OLIGO          start  len  tm  gc%  any_th  3'_th  hairpin  seq
LEFT PRIMER    1086  20  59.03  50.00  0.97  0.97  0.00  TTCAATGTTTCAGGCAGGCC
RIGHT PRIMER   1932  20  60.11  50.00  6.51  0.00  0.00  AACCAAATTGACGGATGCGC
SEQUENCE SIZE: 2149
INCLUDED REGION SIZE: 2149

PRODUCT SIZE: 847, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00


```

Figura 9. Características de *primers* diseñados para amplificación del gen *FAD2*, proporcionadas por el software Primer3.






Por otra parte, el programa OligoAnalyzer mostró algunas estructuras secundarias que son teóricamente posibles, principalmente en el par de primers A (Figura 10). Se reporta también que para los tres pares de *primers*, estas estructuras secundarias tienen una estabilidad demasiado baja como para formarse en una cantidad apreciable en los rangos de temperatura típicos de una PCR, por lo que no interfieren en el proceso de amplificación.

PRIMERS A

Forward:

structure	Image	ΔG (kcal.mole ⁻¹)	T _m (°C)	ΔH (kcal.mole ⁻¹)	ΔS (cal.K ⁻¹ .mole ⁻¹)	Output
1		-0.04	25.5	-21.2	-70.98	<input type="checkbox"/> Ct <input type="checkbox"/> Det
2		0.01	24.8	-16.2	-54.37	<input type="checkbox"/> Ct <input type="checkbox"/> Det
3		0.95	11.7	-20.4	-71.62	<input type="checkbox"/> Ct <input type="checkbox"/> Det

Reverse:





structure	Image	ΔG (kcal.mole ⁻¹)	T_m (°C)	ΔH (kcal.mole ⁻¹)	ΔS (cal.K ⁻¹ .mole ⁻¹)	Output
1		0.86	10.8	-17.3	-60.92	<input type="button" value="Ct"/> <input type="button" value="Det"/>
2		1.32	6.2	-19.6	-70.17	<input type="button" value="Ct"/> <input type="button" value="Det"/>
3		1.44	-13.6	-9.7	-37.37	<input type="button" value="Ct"/> <input type="button" value="Det"/>
4		1.48	-425.1	0.5	-3.29	<input type="button" value="Ct"/> <input type="button" value="Det"/>
5		1.57	-2.1	-15.7	-57.91	<input type="button" value="Ct"/> <input type="button" value="Det"/>
6		1.58	-16	-9.9	-38.5	<input type="button" value="Ct"/> <input type="button" value="Det"/>
7		1.76	-25.8	-8.6	-34.76	<input type="button" value="Ct"/> <input type="button" value="Det"/>
8		1.77	-9.4	-13.6	-51.56	<input type="button" value="Ct"/> <input type="button" value="Det"/>
9		1.86	-20.9	-10.2	-40.44	<input type="button" value="Ct"/> <input type="button" value="Det"/>
10		1.86	-24.8	-9.3	-37.44	<input type="button" value="Ct"/> <input type="button" value="Det"/>

PRIMERS B

Forward:



structure	Image	ΔG (kcal.mole ⁻¹)	T_m (°C)	ΔH (kcal.mole ⁻¹)	ΔS (cal.K ⁻¹ .mole ⁻¹)	Output
1		1.22	3.2	-15.5	-56.08	<input type="button" value="Ct"/> <input type="button" value="Det"/>
2		1.35	1.5	-15.8	-57.53	<input type="button" value="Ct"/> <input type="button" value="Det"/>
3		1.52	-0.4	-16.3	-59.76	<input type="button" value="Ct"/> <input type="button" value="Det"/>
4		1.73	-25.4	-8.5	-34.31	<input type="button" value="Ct"/> <input type="button" value="Det"/>
5		1.96	-15.4	-12.5	-48.5	<input type="button" value="Ct"/> <input type="button" value="Det"/>

Reverse:

structure	Image	ΔG (kcal.mole ⁻¹)	T_m (°C)	ΔH (kcal.mole ⁻¹)	ΔS (cal.K ⁻¹ .mole ⁻¹)	Output
1		0.63	13.7	-15.9	-55.43	<input type="button" value="Ct"/> <input type="button" value="Det"/>
2		0.83	9.1	-14.8	-52.43	<input type="button" value="Ct"/> <input type="button" value="Det"/>
3		1.21	2.3	-14.7	-53.37	<input type="button" value="Ct"/> <input type="button" value="Det"/>
4		1.53	-9	-11.9	-45.05	<input type="button" value="Ct"/> <input type="button" value="Det"/>

PRIMERS C

Forward:

structure	Image	ΔG (kcal.mole ⁻¹)	T _m (°C)	ΔH (kcal.mole ⁻¹)	ΔS (cal.K ⁻¹ .mole ⁻¹)	Output
1		-0.15	27.5	-18.7	-62.2	<input type="checkbox"/> Ct <input type="checkbox"/> Det
2		0.51	17.5	-19.7	-67.79	<input type="checkbox"/> Ct <input type="checkbox"/> Det

Reverse:



structure	Image	ΔG (kcal.mole ⁻¹)	T _m (°C)	ΔH (kcal.mole ⁻¹)	ΔS (cal.K ⁻¹ .mole ⁻¹)	Output
1		-1.04	39.5	-22.4	-71.65	<input type="checkbox"/> Ct <input type="checkbox"/> Det
2		-0.27	29.4	-18.4	-60.82	<input type="checkbox"/> Ct <input type="checkbox"/> Det

Figura 10. Estructuras secundarias de los *primers* diseñados para la amplificación del gen *FAD2*, predichas por el software Primer3.

Mediante el PrimerBLAST se logró identificar tres posibles *mismatch* por parte del par de primers B, dos en el cromosoma 9 y otro en el cromosoma 7 (Figura 11). Los *mismatch* son regiones del genoma de *R. communis* que pueden amplificarse accidentalmente, resultando en inespecificidades.

PRIMERS B

>[JAABNL010000009.1](#) Ricinus communis chromosome 9 Chr09, whole genome shotgun sequence

```
product length = 2198
Forward primer 1 TTCCACCTCCTTCCTGAACC 20
Template 23471649 .A.AC..C.....C.... 23471668

Forward primer 1 TTCCACCTCCTTCCTGAACC 20
Template 23473846 G..TG..C.T..... 23473827
```

```
product length = 3634
Reverse primer 1 GCCTTTGTAGCCTCCATTGC 20
Template 10056873 A.....CA.T.....T. 10056854

Reverse primer 1 GCCTTTGTAGCCTCCATTGC 20
Template 10053240 .TA....G..A.....A 10053259
```

>[CM033707.1](#) Ricinus communis isolate WT05 ecotype wild-type chromosome 7, whole genome shotgun sequence

```
product length = 3284
Forward primer 1 TTCCACCTCCTTCCTGAACC 20
Template 6695216 C...T.....AAC..... 6695197

Reverse primer 1 GCCTTTGTAGCCTCCATTGC 20
Template 6691933 A..A....T.A.....A. 6691952
```

Figura 11. Características de *primers* diseñados para amplificación del gen *FAD2*, proporcionadas por el software PrimerBLAST.

6.5 Estandarización de la PCR para la amplificación del gen *FAD2* de *R. communis*

Después de obtener los resultados del proceso de estandarización de la PCR, se observa que el *master mix* v2 tiene menos inespecificidades en las temperaturas que mejor amplifican en ambos gradientes (57.5°C y 53.8°C), por lo que se concluye que es más eficaz en comparación con el *master mix* v1 (Figura 12). No obstante, se observa como la PCR deja de amplificar correctamente en temperaturas superiores a 65.9°C. Adicionalmente, los amplicones tienen un tamaño de entre 1000 y 1500 pb y se presenta en algunas de las muestras una inespecificidad con un tamaño de entre 600 y 700 pb.

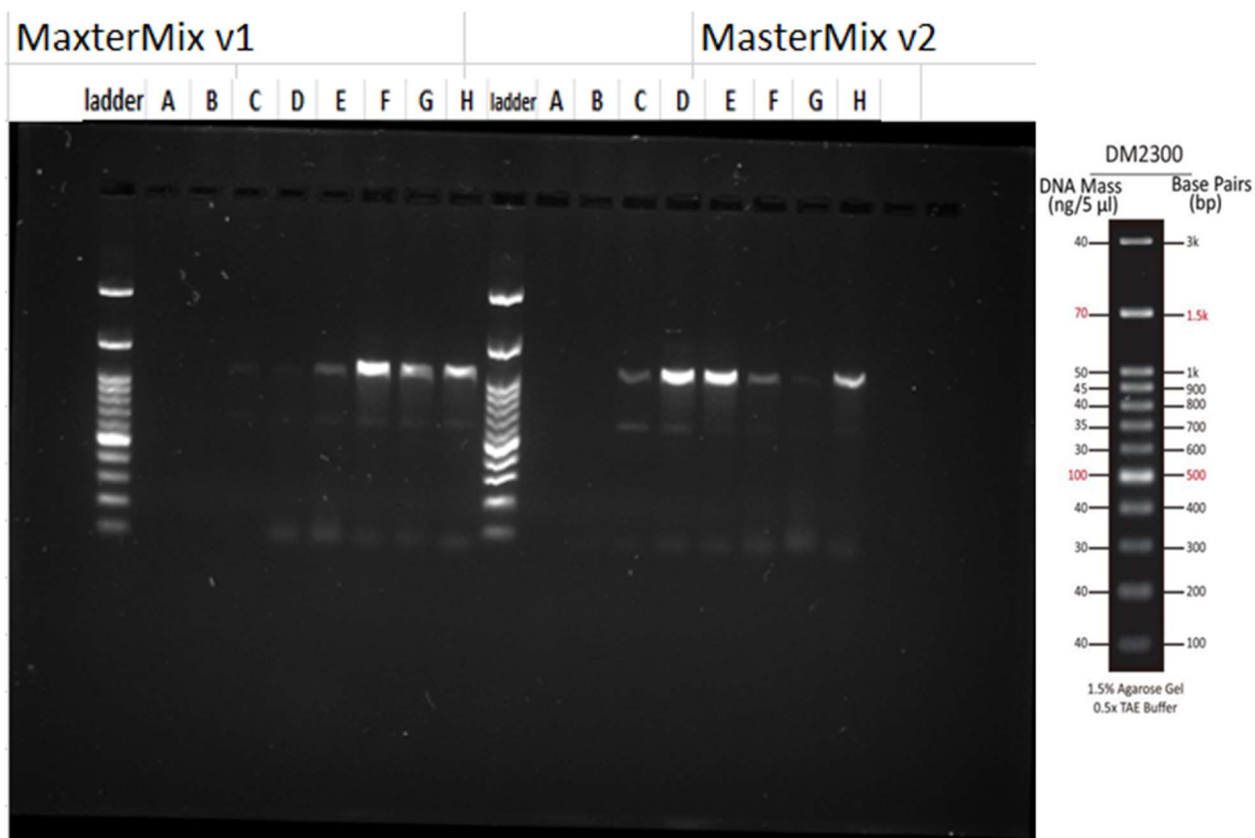


Figura 12. Electroforesis comparativa de los gradientes de temperatura entre 70°C y 50°C resultantes de usar los *master mix* v1 y v2 con el par de *primers* A. El tamaño de los amplicones esperados para los *primers* A, B y C son 836, 785 y 847, respectivamente. Las temperaturas correspondientes a cada carril son las siguientes: Ladder: 1 kb, A: 70 °C, B: 68.5 °C, C: 65.9 °C, D: 62 °C, E: 57.5 °C, F: 53.8 °C, G: 51.4 °C, H: 50 °C.

Los resultados del segundo gradiente de temperaturas (Figura 13) muestran que temperaturas demasiado bajas favorecen la aparición de inespecificidades. En este sentido, son preferibles las temperaturas superiores a 55.3°C.

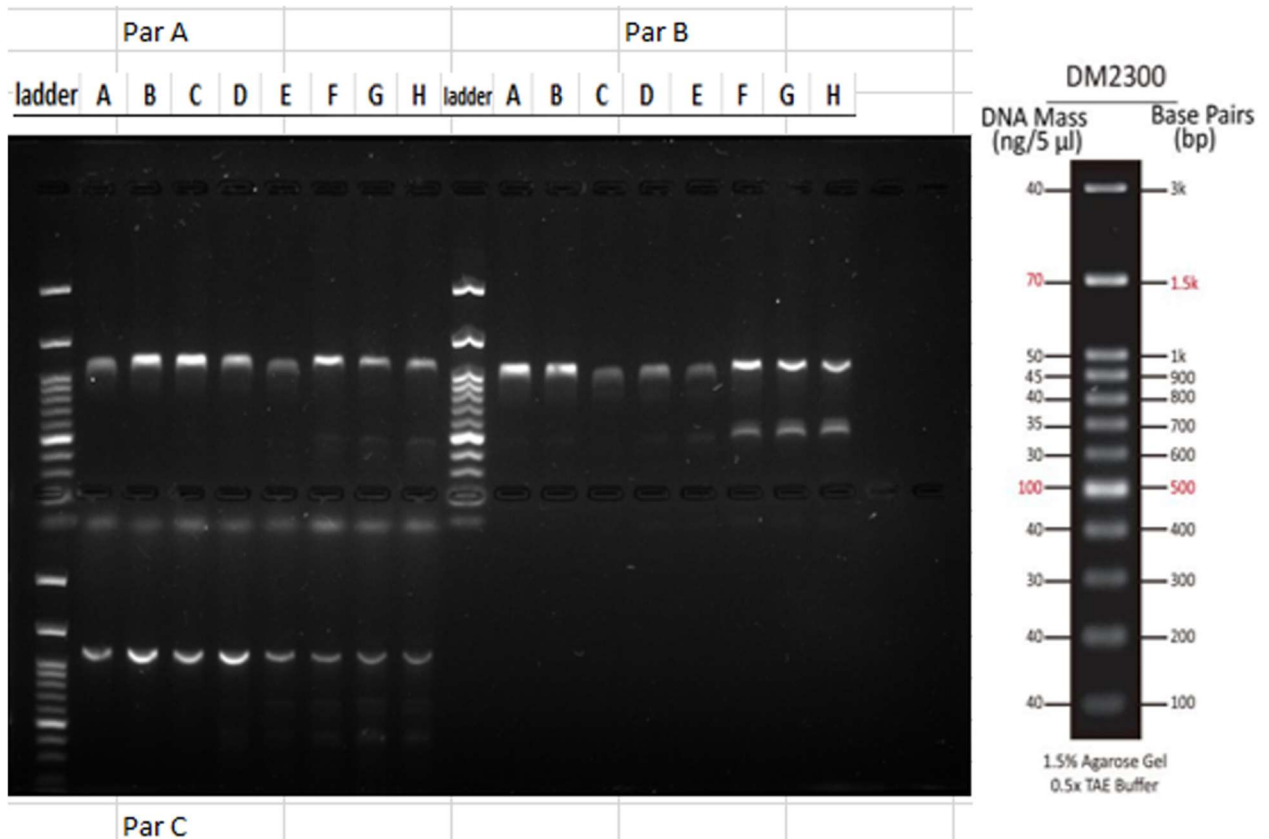


Figura 13. Gradientes de temperatura entre 58°C y 45°C para los pares de *primers* A, B y C. Las temperaturas de los carriles son: Ladder: 1 kb, A: 58 °C, B: 57 °C, C: 55.3 °C, D: 53 °C, E: 50 °C, F: 47.5 °C, G: 45.9 °C, H: 45 °C.

Se repite luego la PCR (Figura 14) pero esta vez entre 56°C y 66°C. Las bandas más intensas se encuentran entre las temperaturas 59.8 y 57.9, lo que sugiere una mayor eficiencia en la PCR. Se decidió, por lo tanto, y razones prácticas, elegir 60°C como temperatura de *annealing* para los tres pares de *primers* en todas las PCRs en adelante.

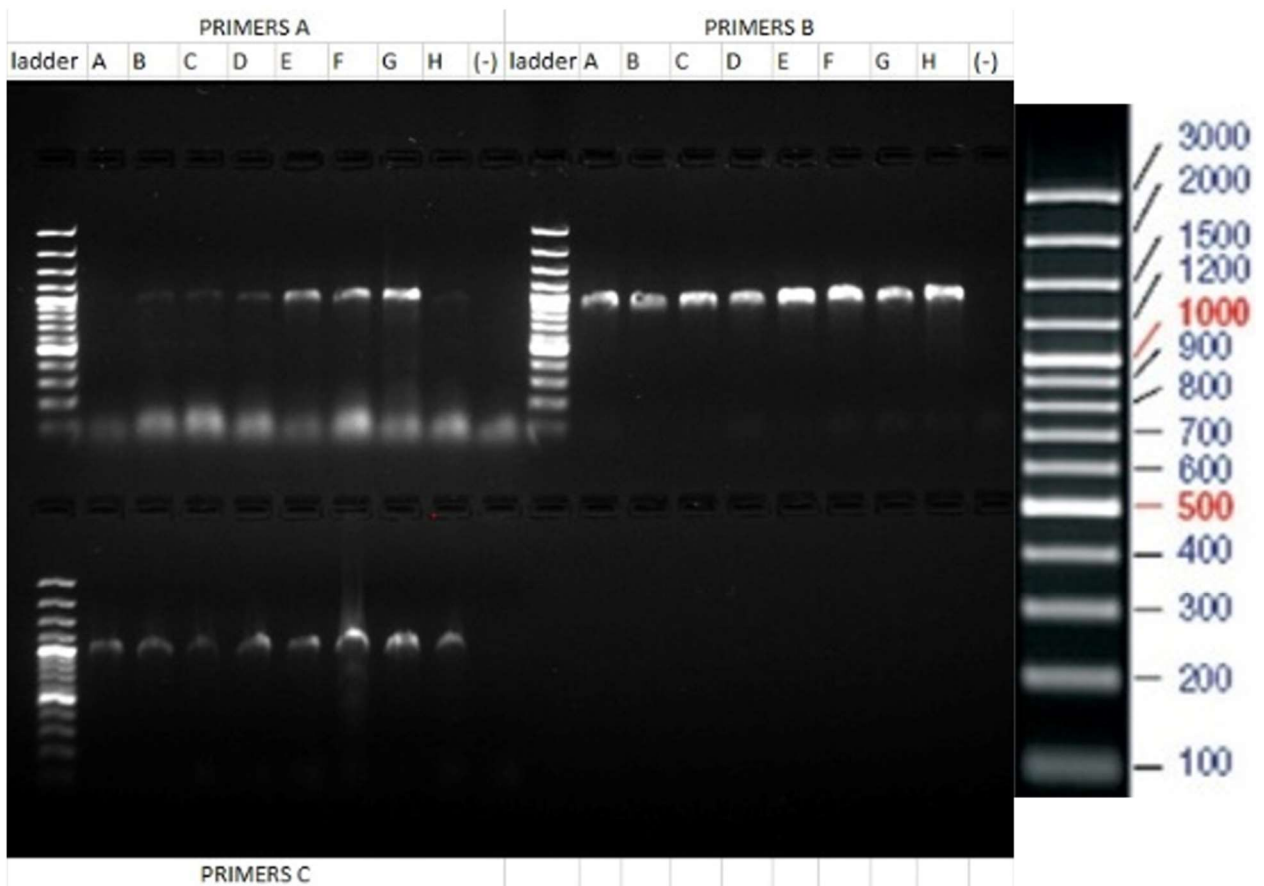


Figura 14. Gradientes de temperatura entre 66°C y 56°C para los pares de *primers* A, B y C. Las temperaturas de cada carril son: Ladder: 1 kb, A: 66 °C, B: 65.3 °C, C: 64 °C, D: 62.1 °C, E: 59.8 °C, F: 57.9 °C, G: 56.7 °C, H: 56 °C, (-): control negativo.

En los resultados del gradiente de cloruro de magnesio se puede observar que las bandas más intensas y sin inespecificidades para los tres pares de *primers* se encuentran en el carril correspondiente a la concentración de $MgCl_2$ de 2 mM (Figuras 15 y 16). Esto sugiere que la concentración de $MgCl_2$ de 2 mM es la más adecuada para amplificar las tres regiones del CDS de *FAD2*.

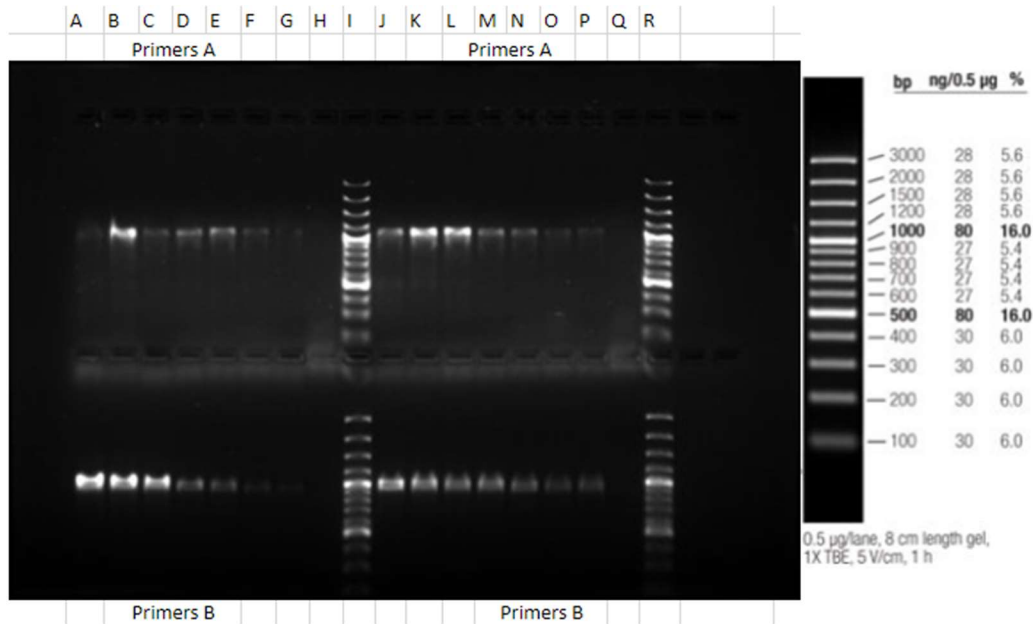


Figura 15. Gradiente de concentración de cloruro de magnesio entre 1 mM y 4 mM para los pares de *primers* A y B. El Ladder usado es de 1 kb. Las concentraciones de cloruro de magnesio (mM) en las muestras de cada carril son: A: 1, B: 1.5, C: 2, D: 2.5, E: 3, F: 3.5, G: 4, H: control negativo, I: ladder, J: 1, K: 1.5, L: 2, M: 2.5, N: 3, O: 3.5, P: 4, Q: control negativo, R: ladder.

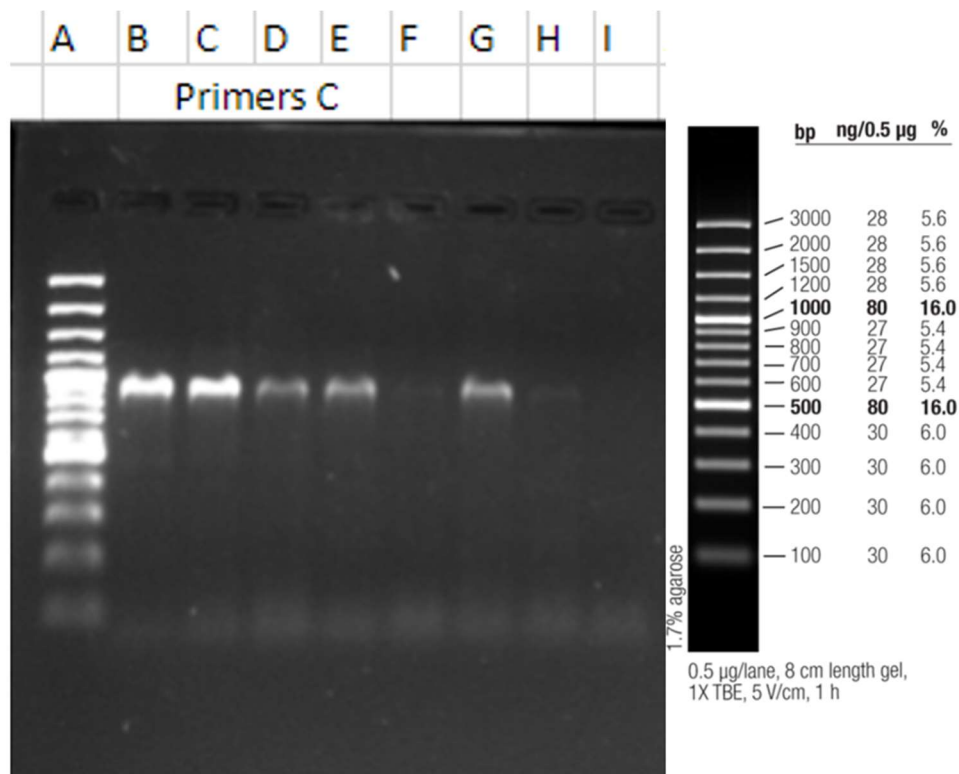


Figura 16. Gradiente de concentración de cloruro de magnesio entre 1 mM y 4 mM para el par de *primers* C. El Ladder usado es de 1 kb. Las concentraciones de cloruro de magnesio

(mM) en las muestras de cada carril son: A: ladder, B: 1, C: 1.5, D: 2, E: 2.5, F: 3, G: 3.5, H: 4, I: control negativo.

Se evidencia en la Figura 17 que, para las concentraciones usadas en el experimento, la que resulta en bandas más intensas es 125 ng/μL, y a juzgar por la tendencia que se muestra, es probable que concentraciones más altas de DNA resulten en una amplificación todavía mayor. A pesar de esto, se decidió continuar usando una concentración de 100 ng/μL, ya que también da una buena intensidad de banda.

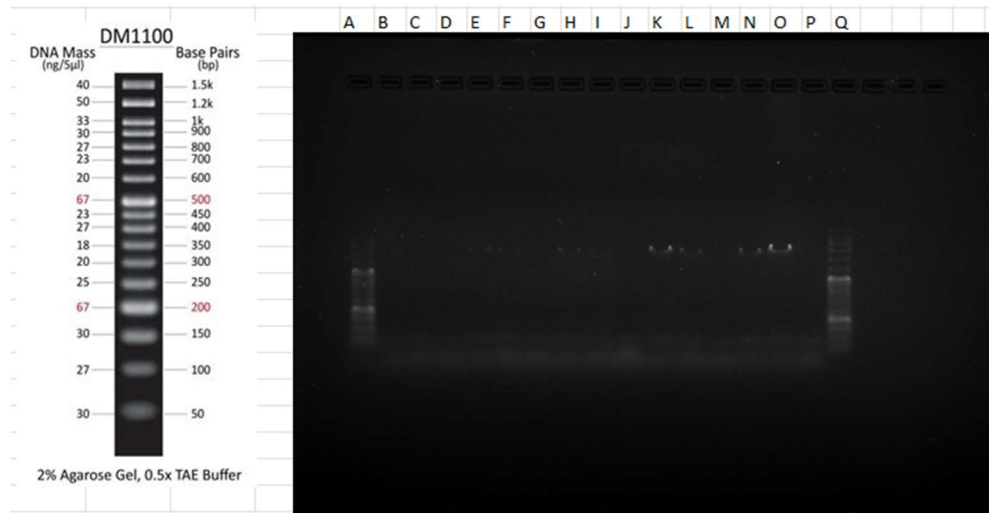


Figura 17. Gradiente de concentración de DNA entre 25 ng/μL y 125 ng/μL para el par de *primers* C. Las concentraciones de DNA (ng/μL) en las muestras de cada carril son: A: ladder, B: 25, C: 25, D: control negativo, E: 50, F: 50, G: control negativo, H: 75, I: 75, J: control negativo, K: 100, L: 100, M: control negativo, N: 125, O: 125, P: control negativo, Q: ladder.

Los resultados del experimento de la Figura 18 muestran que la concentración de *primers* más adecuada para las PCRs es 0.16 ng/μL.

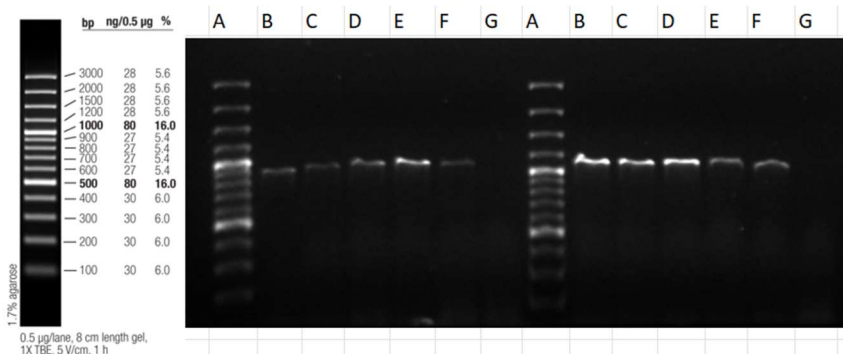


Figura 18. Gradiente de concentración de *primers* entre 0.12 ng/μL y 0.2 ng/μL para el par de *primers* C. Las concentraciones de primers (ng/μL) en las muestras de cada carril son: A: ladder, B: 0.12, C: 0.14, D: 0.16, E: 0.18, F: 0.2, G: control negativo.

De los experimentos propios de la estandarización de la PCR para amplificar las regiones de *FAD2* a partir de DNA de hoja de *R. communis*, se puede concluir que para los tres pares de *primers*, las condiciones ideales de la reacción son temperatura de *annealing* de 60 °C, concentración de cloruro de magnesio de 2 mM, concentración del stock de DNA de 100 ng/μL y concentración de *primers* de 0.16 ng/μL, empleando el *master mix* v2. Estas condiciones resultan en una amplificación eficiente y específica, tal como se evidencia en la Figura 19.

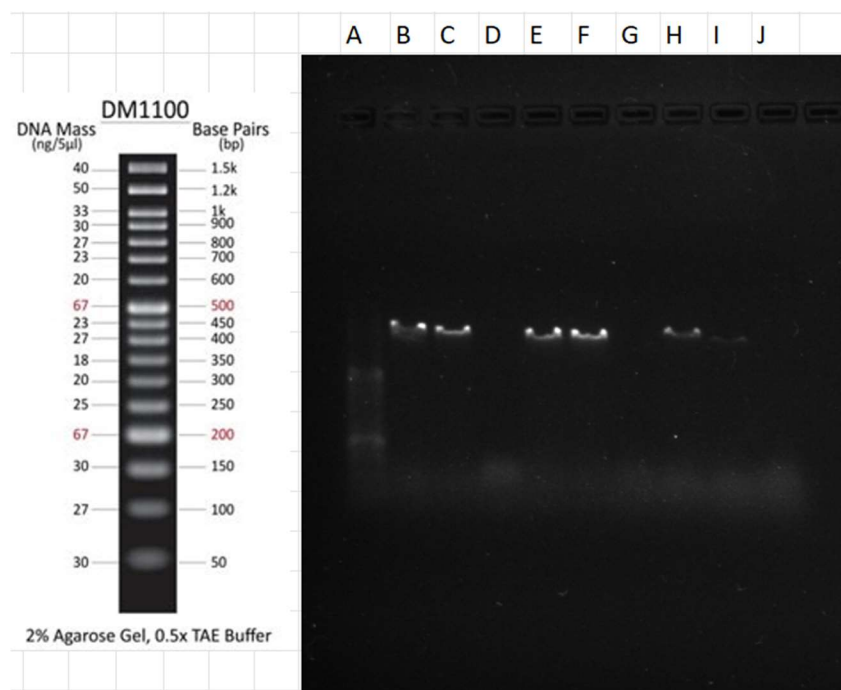


Figura 19. PCR final con las condiciones sugeridas por los experimentos anteriores empleando los pares de *primers* A, B y C. Las muestras en cada carril son: A: ladder, B: primers A, C: primers A, D: control negativo, E: primers B, F: primers B, G: control negativo, H: primers C, I: primers C, J: control negativo.

Con base en las condiciones estandarizadas, se llevó a cabo una PCR para amplificar el DNA extraído de los cuatro individuos de *R. communis* con los *primers* A, B y C (Figura 20). Posteriormente, estos amplicones fueron sometidos a secuenciación.

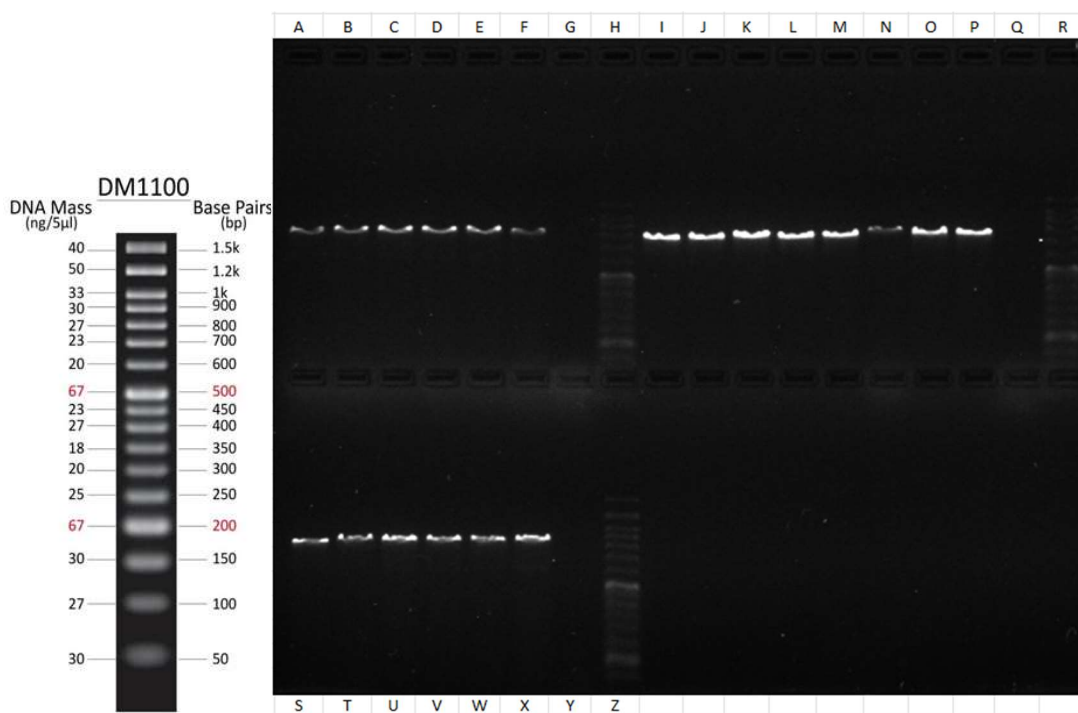


Figura 20. Amplificación de las tres regiones del gen *FAD2* de *R. communis* con los *primers* A, B y C, para 20 muestras destinadas a secuenciación. Las concentraciones de cloruro de magnesio (mM) en las muestras de cada carril son: A: A11, B: A12, C: A21, D: A22, E: A3, F: A41, G: control negativo, H: ladder, I: A42, J: B11, K: B12, L: B2, M: B31, N: B32, O: B41, P: B42, Q: control negativo, R: ladder, S: C1, T: C21, U: C22, V: C31, W: C32, X: C4, Y: control negativo, Z: ladder.

6.6 Secuenciación y modelación de *FAD2* en *R. communis*

Luego de ser secuenciados los amplicones, el resultado fueron 40 archivos en formato .ab1 que corresponden a los cromatogramas que arroja la prueba (20 *forward* y 20 *reverse*), cada uno con una interpretación automática en formato .txt donde se extraen los nucleótidos correspondientes a cada señal. No obstante, algunos de los *reads* generados de las muestras B *forward* tuvieron baja calidad (B11/BF, B12/BF, B2/BF, B31/BF, B32/BF, B41/BF, B42/BF), por lo que fueron descartados para el alineamiento (Tabla 11).

Tabla 11. Tamaño reportado para los *reads* resultado de la secuenciación por el método de Sanger. El nombre de cada *read* consiste en la muestra de la que proviene y tiene añadido AF, AR, BF, BR, CF o CR dependiendo si la reacción usó los *primers forward* o *reverse* para esa región. El largo se indica para la longitud total de los *reads* (normal) y por la longitud del *read* cuando se le retiran las porciones que no tienen buena calidad. Q20 indica 1 error por cada 100 nucleótidos.

Read	Largo	
	Normal	Q20
A11/AF	781	763
A12/AF	793	482

A21/AF	778	767
A22/AF	789	730
A3/AF	802	668
A41/AF	790	164
A42/AF	783	587
B11/BF	95	9
B12/BF	745	57
B2/BF	742	24
B31/BF	63	8
B32/BF	64	13
B41/BF	96	9
B42/BF	121	10
C1/CF	800	737
C21/CF	781	740
C22/CF	804	794
C31/CF	797	659
C32/CF	787	776
C4/CF	771	770
A11/AR	788	767
A12/AR	785	261
A21/AR	787	765
A22/AR	787	187
A3/AR	782	172
A41/AR	856	218
A42/AR	791	174
B11/BR	741	535
B12/BR	742	557
B2/BR	737	662
B31/BR	752	141
B32/BR	752	491
B41/BR	679	131
B42/BR	524	51
C1/CR	807	563
C21/CR	800	737
C22/CR	813	516
C31/CR	797	515
C32/CR	799	519
C4/CR	799	787

De la detección de variantes resultaron un total de 6 SNPs, de las cuales solo tres se ubicaban dentro del CDS (variantes en las posiciones 3308, 3353 y 3652). Los tres consisten en un

cambio de timina a adenina en la secuencia (tabla 12). Las primeras dos variantes generan un cambio en el codón que altera el aminoácido para el que codifican. En ambos casos pasa de codificar una tirosina (Y) a una asparagina (N) en los aminoácidos de las posiciones 68 y 83. Por otro lado, la tercera variante es sinónima, el cambio en el codón no tiene efecto sobre el aminoácido para el que codifica, por lo que no hay alteración en la estructura primaria de la proteína (Figura 21).

Tabla 12. SNPs presentes en la secuencia de DNA del gen *FAD2* de *R. communis* variedad VERC-03. Se denota la posición de la variante, la base presente en la referencia y la variante encontrada en la secuenciación, junto con su índice de calidad en *phred score*, la confianza que se tiene en que esta variante es real.

Posición	Referencia	Variante	Calidad
2973	T	C	37.6032
3308	T	A	32.2598
3353	T	A	19.4952
3652	T	A	84.5524
4483	T	A	17.2681
4488	G	T	11.7843

Score	Expect	Method	Identities	Positives	Gaps
791 bits(2043)	0.0	Compositional matrix adjust.	381/383(99%)	381/383(99%)	0/383(0%)
Query 1	1	MGAGGRMSVPPPSKKVESDDLKRAPSSKPPFTLGQIKKAIPPHCFKRSIPRSFSYVYDL			60
Sbjct 1	1			60
Query 61	61	TIAFLFYVATNYFHLLPEPLSYVAWPIYWALQGCVL TGWVIAHECGHHAFFSDYQLLDD			120
Sbjct 61	61N.....N.....			120
Query 121	121	VVGLILHSCLLVPYFSWKHSHRRHSNTGSLERDEVFVPKKKSSIRWYSKYLNNPPGRIM			180
Sbjct 121	121			180
Query 181	181	TIAVTLTLGWPLYLAFNVSGRPYDRFACHYDPYGPYNDRERIEIFISDAGVLAVTFGLY			240
Sbjct 181	181			240
Query 241	241	QLAIAGKLAWVVCVYGVPLLVNSFLVLITFLQHTHPALPHYDSSEWDWLRGALATVDRD			300
Sbjct 241	241			300
Query 301	301	YGILNKVFHNIITDTHVAHHLFSTMPHYHAMEATKAIKPILGEYYQFDGTSFYKAMWREAK			360
Sbjct 301	301			360
Query 361	361	ECIYVEKDDAEQNGGVFWYNNKF	383		
Sbjct 361	361	383		

Figura 21. Alineamiento entre la proteína *FAD2* canónica de *R. communis* (*query*) y la predicha a partir del DNA secuenciado de la variedad VERC-03 (*subject*). En rojo se resaltan las variantes presentes.

Al modelar la estructura 3D de ambas proteínas, la original y la predicha de la variedad VERC-03 de *R. communis*, se encontró que, en ambos casos, el software tuvo problemas para modelar con precisión los extremos de la proteína, pero pudo predecir con un alto grado de

confianza el cuerpo principal (Figura 22 y 23). Esto puede deberse a la falta de cobertura en los extremos durante el alineamiento de las secuencias de *FAD2* con las disponibles en las bases de datos consultadas (Figura 22). Era de esperar también que los resultados para la proteína de referencia y la de VERC-03 fueran casi idénticos, ya que solo difieren en dos residuos de aminoácidos. Estos dos aminoácidos están ubicados en las posiciones 68 y 83, las cuales tienen buena cobertura y una confianza alta en la predicción.

En la Figura 24 se ve a la izquierda el LDDT predicho por residuo, el cual es útil para estimar la confianza en el modelo a nivel de dominio. A la derecha se muestra el error alineado predicho, que sirve para estimar el error en la predicción en la interacción entre dominios. En ambos casos se evidencia una caída en la calidad de la predicción al llegar a los extremos de la cadena, aunque la caída es mayor en el extremo N terminal que en el C terminal. En el modelo 3D de la figura 23 se observa que este extremo N terminal toma la forma de una cadena larga y suelta, no incorporada al resto de la proteína, lo que probablemente no se acerque a la realidad.

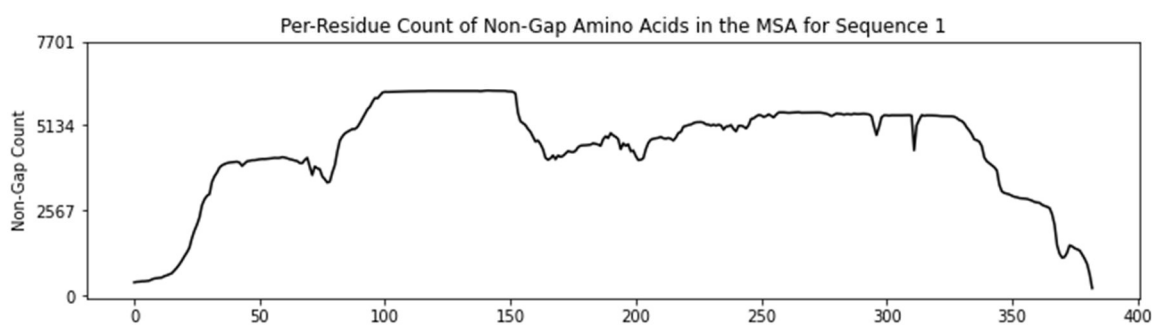
REFERENCIA

6761 unique sequences found in uniref90 for sequence 1

563 unique sequences found in smallbfd for sequence 1

501 unique sequences found in mgnify for sequence 1

7703 unique sequences found in total for sequence 1



VERC-03

6739 unique sequences found in uniref90 for sequence 1

554 unique sequences found in smallbfd for sequence 1

501 unique sequences found in mgnify for sequence 1

7677 unique sequences found in total for sequence 1

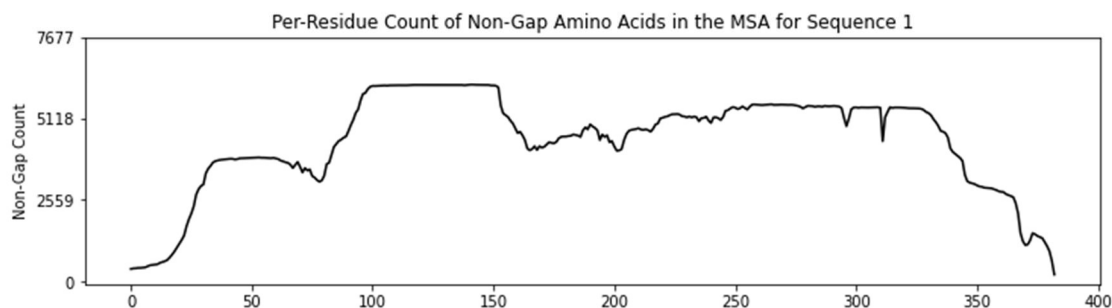
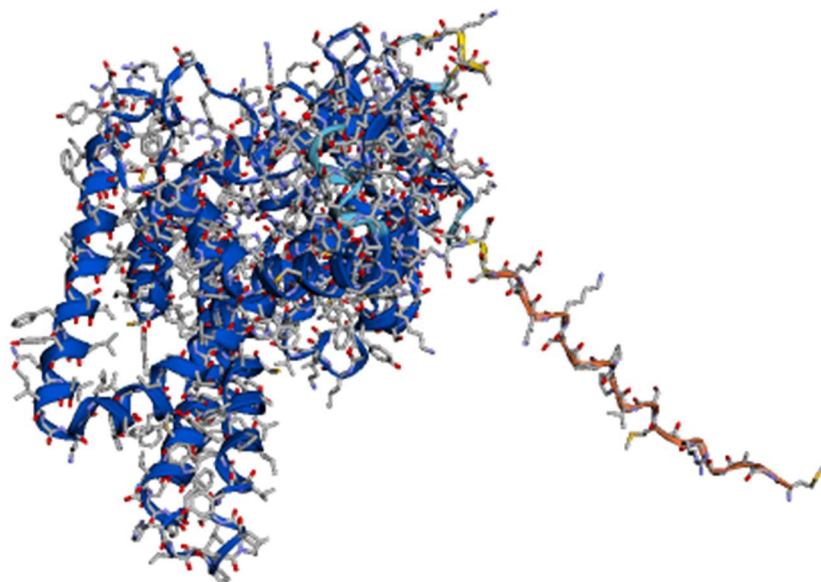


Figura 22. Cobertura de las secuencias proteicas de las bases de datos sobre las secuencias de la proteína *FAD2* de referencia y la de la variedad VERC-03 realizado por el software AlphaFold.

REFERENCIA



VERC-03

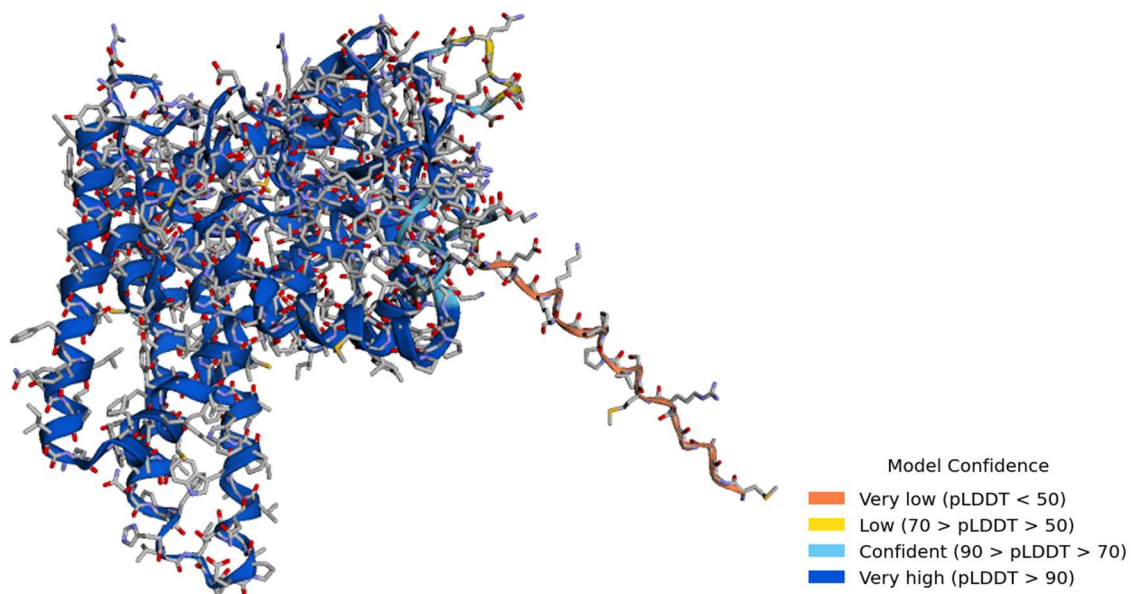
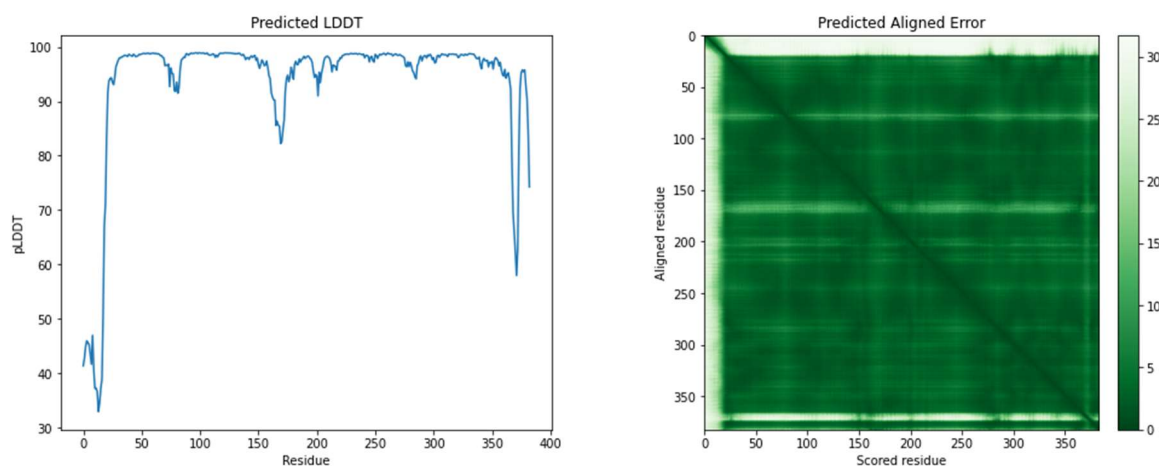


Figura 23. Comparación de los modelos 3D de la proteína *FAD2* de referencia y la de la variedad VERC-03 obtenidos con el software AlphaFold.

REFERENCIA



VERC-03

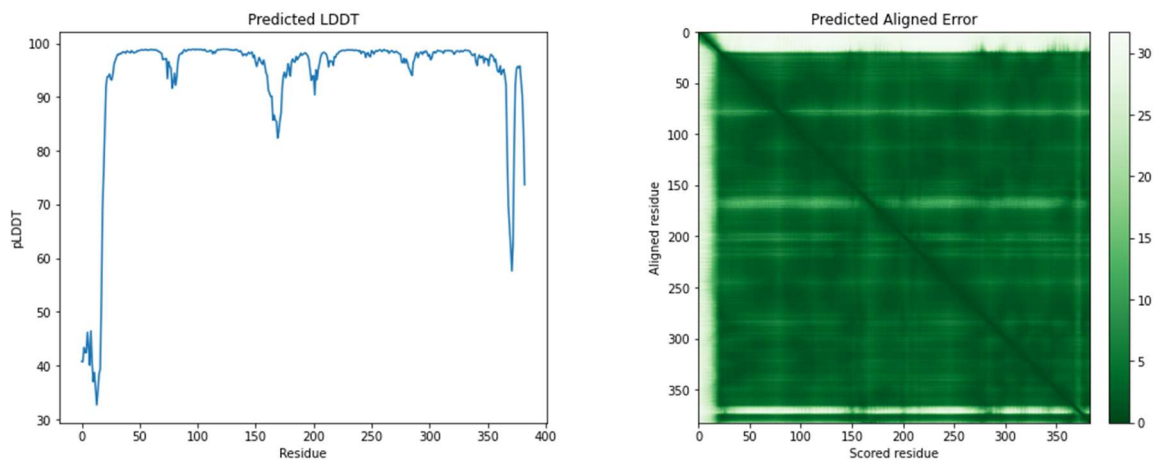


Figura 24. Comparación entre los grados de error predichos para los modelos 3D de la proteína *FAD2* de referencia y la de la variedad VERC-03 obtenidos con el software AlphaFold Colab. A la izquierda, LDDT predicho por residuo, a la derecha, error alineado predicho.

Usando el software COACH se predijo el sitio activo de la proteína con un puntaje de 0.11 y se encontró que éste está ubicado entre los residuos 105, 109, 141, 145 y 318, los cuales corresponden a una caja de histidinas (Figura 25).

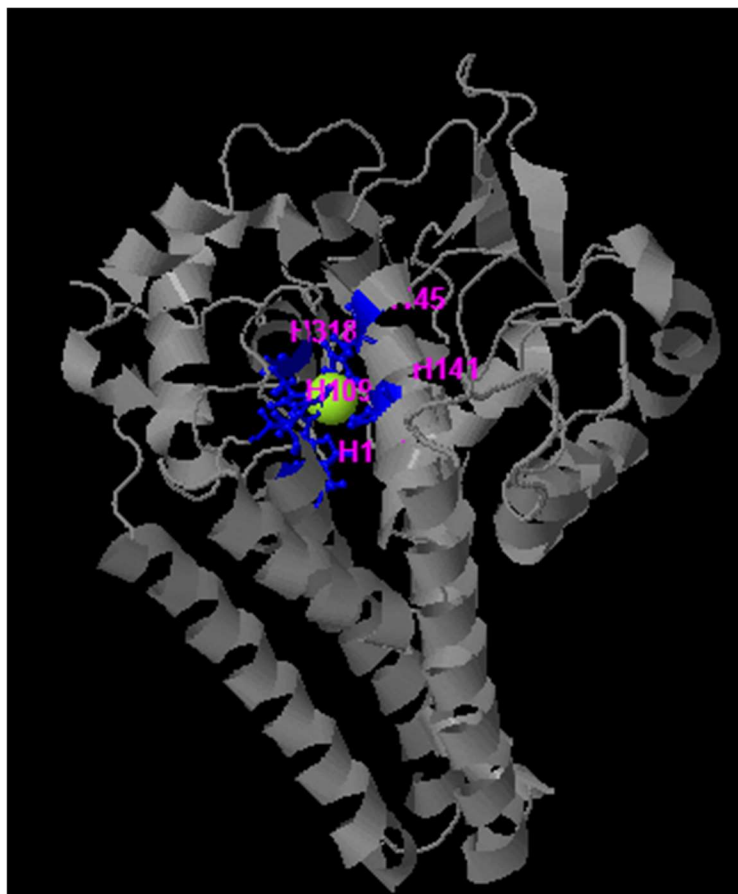


Figura 25. Ubicación del sitio activo de la proteína, predicho por COACH. Se resaltan en azul los residuos implicados.

Posteriormente se realizó el análisis de “docking” molecular con el software Autodock VINA, enfocándose especialmente en la región predicha por COACH para ser el sitio activo. Se encontró que el *FAD2* de referencia tiene una afinidad por el ácido oleico de alrededor de 5,8 y por el ácido linoleico de 6,4. Unos valores similares se dan para el *FAD2* de la variedad VERC-03 de *R. communis*, 5,6 de afinidad por el ácido oleico y 6,3 por el ácido linoleico. No se observa una diferencia significativa entre la afinidad de ambos *FAD2* por los sustratos y se observa también una mayor afinidad de la enzima por el ácido linoleico que por el ácido oleico (Tabla 13).

Tabla 13. Afinidad entre la enzima *FAD2* de referencia y la de VERC-03 y los sustratos ácido oleico y ácido linoleico. Se muestran los valores arrojados por las 19 réplicas realizadas y el promedio ponderado.

Afinidad (kcal/mol)			
Ácido oleico		Ácido linoleico	
Referencia	VERC-03	Referencia	VERC-03
5,4	6,7	6,6	5,4
6,3	6,8	5,9	6,9

5,5	5,3	6,7	6,2
5,3	5	6,8	5,9
5,5	6,9	5,4	7,2
6,8	5,7	5,5	5,3
5,8	5,4	6,7	5,9
6,5	5	7,5	5,9
6,5	6,7	6	5,9
7,1	5,1	6,8	6,2
5,5	5,2	5,7	6,8
5,8	5	5,5	5,9
5,9	5,5	7,6	7,3
4,9	5,4	6,2	5,9
5,3	5	6,7	7,9
5,6	6,6	7,2	8,3
5,3	5,7	5,7	5,9
5,5	5,1	6,1	5,4
6,9	5,6	7	5,8
Promedio	Promedio	Promedio	Promedio
5,863158	5,668421	6,4	6,315789

7. DISCUSIÓN

7.1 Búsqueda *in silico* de isoformas del gen *FAD2*

En la primera fase de este trabajo de investigación se pretendió caracterizar el gen de la enzima desaturasa de ácidos grasos 2 (*FAD2*) en *Ricinus communis* de la variedad VERC-03 a partir de métodos moleculares, así como sus posibles isoformas usando métodos computacionales, lo cual serviría como base para una segunda fase, en la que se abordaría el estudio de las isoformas de *FAD2* a nivel de expresión diferencial.

El BLAST realizado usando múltiples secuencias de *FAD2* de diversas especies, variedades e isoformas contra el genoma de *R. communis* reportado en NCBI, mostró que no existen secuencias homólogas al gen *FAD2* en su genoma, que está ubicado únicamente en un locus y no tiene otras copias en el genoma. Lo más similar que se encontró a cualquiera de los genes *FAD2* de otras especies y sus isoformas, fue el gen *FAH12*, que cumple una función muy distinta y de todas formas comparte con *FAD2* una identidad baja. Por lo tanto, este experimento sugiere buscar las isoformas de *FAD2* de *R. communis* en ese mismo locus, en forma de alelos o productos de *splicing* alternativo. Esto contrasta con lo reportado para otras especies como soya, la cual posee dos copias del gen que se expresan de forma independiente y provienen de dos loci distintos, *FAD2-1* y *FAD2-2* (Heppard et al., 1996). A su vez, el gen *FAD2-1* se expresa de forma diferencial en dos isoformas, *FAD2-1A* y *FAD2-1B* (Tang et al., 2005).

La estructura del gen *FAD2* que se reporta en la anotación del ensamblaje ASM1957865v1 consiste en dos exones pequeños de 134 pb y 25 pb más un exón principal de 1405 pb que contiene todo el CDS. Dos isoformas que se reportan en esta anotación para el mRNA de *FAD2* son entre el exón principal y cada uno de los dos exones de forma independiente. Incluso en la isoforma que crea la unión del exón de 25 pb y el de 1405, este último recorta su extremo 3' para quedar de un tamaño de 1353 pb. No obstante, el CDS siempre es el mismo y la proteína generada viene a ser la misma, independientemente de las transformaciones que sufra el mRNA, por lo que la maduración de éste y el *splicing* no parecen tener en este caso un efecto sobre el resultado final de la proteína. Esta estructura difiere con lo reportado para especies como *Arabidopsis thaliana*, la cual también presenta una región grande 5' del CDS y que no se traduce pero sí se transcribe, con la diferencia de que esta no presenta los dos exones que indica la anotación de *R. communis* (Okuley et al., 1994).

Estos pequeños exones posiblemente correspondan a secuencias de señalización que permiten que la proteína pueda ser transportada y ubicada en un sitio específico de la célula. *FAD2* es una proteína que se ubica en la membrana del retículo endoplasmático y del cloroplasto, por lo que es posible que cada uno de estos exones pequeños ayude a la proteína a dirigirse a un organelo específicamente. Este tipo de señalización se reporta en las proteínas *FAD2* de otras especies como *Olea europaea* entre otras (Banilas et al., 2005; Dehghan Nayeri & Yarizade, 2014), mientras en especies como *Camellia chekiangoleosa* no existe (Z. Wang et al., 2020). Este precedente hace que sea posible que en *R. communis* esté presente también un péptido señal que ayude a *FAD2* a ubicarse ya sea en el retículo endoplasmático o en el cloroplasto.

Otra explicación válida es que estos exones no traducidos constituyen UTRs (regiones no traducidas) que se unen de forma diferencial al CDS. Este fenómeno se denomina IME (mejoramiento mediado por intrones), consiste en afectar la eficiencia de la traducción por medio de aprovechar los UTRs para cambiar la frecuencia de los eventos de retención de intrones, los cuales son una forma de *splicing* alternativo (Chaudhary et al., 2019). En humanos se ha encontrado por ejemplo que existen dos variantes de distinto tamaño de la región 5' UTR del gen *ARPC2* y la más larga contiene un sitio interno de entrada al ribosoma (IRES) del que carece la más corta. También se demostró que solo la variante larga que contiene este elemento es capaz de promover la traducción cap-independiente y posee una región rica en guanina que forma estructuras cuádruplex de guanina (G-quadruplex), las cuales contribuyen a la actividad del IRES (Al-Zeer et al., 2019). Es probable por lo tanto que en el gen *FAD2* de *R. communis* esté ocurriendo el mismo fenómeno de *splicing* alternativo de variantes del 5' UTR.

7.2 Cultivo *in vitro* de *R. communis*

La desinfección de las semillas de *R. communis* involucra el uso de una solución de HgCl_2 , la cual elimina de forma muy efectiva todo tipo de contaminación de la superficie de la semilla, ya sea esporas, hongos o bacterias. Por otro lado, este desinfectante tiene la desventaja de que, al contener un metal pesado, puede tener un efecto nocivo sobre la viabilidad de los embriones de las semillas. Esto se evidencia en que varias de las semillas

que fueron desinfectadas exitosamente luego no lograban germinar para convertirse en plántulas sanas, sino que se quedaban en un estado de atrofia, causada probablemente por la toxicidad del mercurio (Mrozek & Funicelli, 1982). También está claro que un mayor tiempo de exposición al HgCl_2 siempre resulta en pérdida de viabilidad para las semillas, como muestra el estudio de Susunaga-Gomez, (2022). Es por eso por lo que es importante encontrar un equilibrio entre la eficacia de la desinfección y la toxicidad sobre el embrión, de forma que se optimice el proceso y resulten más plántulas sanas (Tilahun et al., 2013). Este equilibrio probablemente varíe según otros aspectos como la carga microbiana inicial de las semillas, la edad de estas, la viabilidad y las condiciones de almacenamiento, ya que unas semillas con menos viabilidad serán más débiles a la hora de resistir la exposición al HgCl_2 . Adicionalmente, el lavado con antibióticos, en este caso Vancomicina y Cefotaxima, contribuye mucho a deshacerse de los últimos remanentes de contaminación que hayan quedado en la superficie de las semillas o en pequeñas grietas y pliegues.

La tasa de germinación encontrada en este trabajo para las semillas de *R. communis* es similar a la reportada por Susunaga-Gómez (2022), la cual empleó el mismo protocolo usado en este trabajo, que a su vez fue estandarizado por Gil-Correal et al. (2019). Varía un poco porque hay una separación en el tiempo entre las dos investigaciones y se usaron las mismas semillas, por lo que se pierde inevitablemente algo de viabilidad.

Los medios de elongación y propagación contienen hormonas del grupo de las auxinas y citoquininas, lo cual concuerda con otros estudios de cultivo *in vitro* de *R. communis* en los que se usan también TDZ y ácido naftalenacético para promover el crecimiento de tallos, epicótilos y raíces, adicional al ácido giberélico. Por otro lado, en cuanto al medio de macronutrientes, algunos autores (Bertozzo & Machado, 2010; J. X. Zhang et al., 2016) reportan el uso de medio Murashige & Skoog (MS) (Murashige & Skoog, 1962) como un método más estándar de cultivo de tejidos vegetales en vez de WPM (Lloyd & McCown, 1980) como en este trabajo, el cual es propuesto por (Correal et al., 2019) como opción de macronutrientes para propagar *R. communis*.

7.3 Extracción de DNA de hojas de *R. communis*

Las extracciones de DNA de hojas de *R. communis* resultaron en una buena concentración de ácidos nucleicos, esto en parte como consecuencia de una biomasa inicial grande y en parte por el desarrollo cuidadoso del protocolo de extracción con CTAB. La relación A260/280 de todas las muestras estuvo en el rango ideal para DNA, entre 1.7 y 1.9, lo cual nos indica pureza del DNA, aunque una de ellas, la del individuo 2 de las extracciones para secuenciación, tuvo un valor un poco superior, lo que indica trazas restantes de RNA que no fueron eliminadas por la RNasa al final del protocolo. La relación A260/230 de las muestras fue aceptable, ya que estuvo siempre rozando el límite del rango ideal, que es superior a 1, pero estuvo por debajo en dos de las muestras, los individuos 3 y 4 de las extracciones para secuenciación. Esto puede deberse a contaminación con carbohidratos, ya que este es un problema usual en extracciones de plantas, sobre todo en tejido de hojas.

7.4 Diseño de *primers* orientados al gen *FAD2* de *R. communis*

El software OligoAnalyzer mostró varias estructuras secundarias teóricamente posibles que se podían formar en los *primers* y que tenían una mayor probabilidad de formarse dependiendo de la temperatura. Como eran estructuras que se formaban más fácilmente en ambientes de baja energía o baja temperatura, es probable que no hayan tenido un impacto en el desempeño de las PCRs, las cuales se desarrollan a temperaturas superiores a los 50°C. No obstante, dado que eran muchas las estructuras teóricamente posibles para cada *primer*, es posible también que en algunos experimentos hayan dificultado una amplificación más eficiente.

El PrimerBLAST mostró dos sitios que teóricamente podían ser amplificados por el mismo par de *primers* que amplificaba también la región B del gen *FAD2*. Estas inespecificidades resultaban no obstante en unos amplicones demasiado grandes, de entre 2000 y 3000 pb, por lo que la enzima Taq polimerasa tendría más dificultad y tardaría más durante la elongación con estos amplicones que con los esperados de la región B de *FAD2*. Adicionalmente, los *primers* no eran totalmente complementarios a estas regiones inespecíficas, por lo que su unión al molde no iba a ser tan eficiente tampoco. Finalmente, los resultados de los geles de electroforesis muestran que no hubo en ningún momento una banda correspondiente a estos tamaños predichos, e incluso si hubiera aparecido una, se hubiera podido eliminar variando condiciones de la PCR como la temperatura de *annealing* o la concentración de cloruro de magnesio, ambos métodos fiables para ajustar los parámetros hasta un punto en el que se favorezca la amplificación de la región de interés.

7.5 Estandarización de PCR para la amplificación del gen *FAD2* de *R. communis*

Inicialmente, los geles de electroforesis mostraron unas bandas de entre 600 y 700 pb que parecían inespecificidades, pero con los siguientes experimentos dejaron de aparecer, por lo que es más probable que fueran contaminación. Las bandas observadas en todos los geles y que correspondían a las 3 regiones del gen *FAD2*, siempre mostraban un tamaño aparente de entre 1000 y 1200 pb al compararse con el marcador de peso molecular, lo cual es paradójico teniendo en cuenta que el tamaño esperado de cada uno de estos amplicones era de alrededor de 800 pb. La velocidad de migración de un fragmento a través del gel depende de factores como la concentración de agarosa, la conformación del DNA, el voltaje, el tipo de agarosa y el *buffer* de corrida, además del tamaño del fragmento (Sambrook, 2001). Sin embargo, todos estos factores siempre fueron usados bajo las mismas condiciones en todos los experimentos, por lo que no pudieron haber sido la causa del cambio en la velocidad de migración. Por lo tanto, se presume que el aumento en el tamaño aparente de las bandas puede deberse a que algunos *buffers* de carga añaden peso a la muestra y ralentizan su avance por el gel, causando una lectura errónea del peso de los fragmentos.

Por otra parte, durante este estudio se determinó que para los *primers* específicos que amplificaban las tres regiones del CDS de *FAD2* el protocolo v2 (2 mM de cloruro de magnesio, 100 ng/μL de DNA y 0.16 ng/μL de *primers*) fue el más efectivo. Aunque estas condiciones de trabajo son muy distintas a las propuestas por Sambrook (2001), que consisten en una concentración mayor de *primers* y DNA, y menor concentración de dNTPs, al final del proceso de estandarización de la PCR los resultados fueron óptimos.

7.6 Secuenciación y modelación de *FAD2* de *R. communis*

Los *reads* generados en la secuenciación por el método de Sanger resultaron en su mayoría con una alta calidad, a excepción de las reacciones realizadas con el *primer forward* de la región B. Dado que la calidad del DNA genómico extraído de hojas de plántulas de *R. communis* fue alta, las amplificaciones con cada par de *primers* fueron exitosas, y la electroforesis reveló que cada una de las 20 muestras enviadas a secuenciar fueron íntegras, es posible deducir que el fallo en producir *reads* de calidad se debió a eventos durante el envío de las muestras o a un error de la misma reacción de secuenciación. Sin embargo, es importante resaltar que esto no afectó la búsqueda de variantes ya que los *reads reverse* de la región B fueron de alta calidad y suplían suficientemente bien la cobertura de la región.

Se determinó que los dos cambios de aminoácidos en el gen *FAD2* de la variedad VERC-03 con respecto a la referencia, son de una Tirosina a una Asparagina y se ubican en las posiciones 68 y 83, ambos en las hélices alfa que anclan la proteína a la membrana lipídica. Es también un cambio de un aminoácido aromático a uno polar, por lo que la adherencia de la proteína a la membrana se vería reducida posiblemente, aunque no demasiado teniendo en cuenta que solo son dos aminoácidos de todos los que hay en las 5 hélices. El sitio activo de la proteína debería por lo tanto funcionar de la misma forma en la proteína de referencia y la de la variedad VERC-03.

Adicionalmente, los dos extremos de la proteína *FAD2* modelada mostraron una calidad baja en la predicción, siendo el N terminal el de peor calidad, con menos de 50 LDDT. Esta cola se ubica espacialmente en una zona de la proteína apartada del sitio activo, pero, aunque no tenga contacto con él, es posible que la mala predicción tenga alguna repercusión en el plegamiento y por lo tanto en la interacción con el sustrato durante la modelación.

La caja de histidinas que se predijo con COACH para ser el sitio activo de la proteína concuerda con otros estudios donde se modelan enzimas desaturasas como Hajiahmadi *et al.* (2020), por lo que la predicción es confiable y presenta ya un precedente reportado. La enzima *FAD2* funciona tomando ácido oleico y transformándolo en ácido linoleico. Según se observa en los promedios de los datos obtenidos en las 19 repeticiones, la enzima *FAD2* de referencia y la de la variedad VERC-03 de *R. communis* tienen aproximadamente la misma afinidad por el sustrato, no existe diferencia significativa entre los dos, por lo que es posible concluir que los cambios de aminoácidos que hay en VERC-03 no tienen un efecto apreciable en la actividad enzimática de la proteína. También cabe resaltar que el sitio activo de la enzima parece tener más afinidad por el ácido linoleico que por el oleico, lo que haría que la enzima retenga más tiempo el producto que el sustrato, ocupando el espacio y ralentizando su acción.

8. CONCLUSIONES

Este estudio permitió sondear las posibles isoformas de *FAD2* en *Ricinus communis* contrastando con datos existentes en bases de datos públicas y empleando métodos *in silico*. A partir de esta búsqueda fue posible encontrar que no existen otras copias de *FAD2* en el genoma y se hipotetizó acerca de la estructura de intrones y exones del gen. Adicionalmente, se validaron los métodos de cultivo *in vitro* estandarizados en estudios anteriores para *R. communis* y se estandarizaron los métodos moleculares para la amplificación del gen *FAD2* de una forma eficiente y apta para la secuenciación por el método de Sanger, pasando por la extracción de DNA a partir de hoja, diseño de *primers* y PCR convencional. En particular, se definieron las condiciones ideales de la reacción para los tres pares de *primers*, resultando en una amplificación eficiente y específica.

Finalmente, se secuenció el CDS de *FAD2* de *R. communis* y se pudo predecir, caracterizar y comparar *in silico* los modelos 3D de la enzima *FAD2* de referencia y de la variedad VERC-03. Se encontró que la variedad VERC-03 de *R. communis* tiene varios SNPs en la secuencia de DNA de *FAD2* que afectan luego la secuencia de aminoácidos. Se predijo la estructura terciaria y se determinó que los dos cambios en los aminoácidos no afectan de forma perceptible la función catalítica de la proteína, no hay una diferencia sustancial entre el *FAD2* de VERC-03 y la proteína de referencia reportada en NCBI. Estos resultados servirán como base para una eventual segunda fase en la que se ha de abordar el estudio de las isoformas de *FAD2* en *R. communis* a nivel de expresión diferencial de forma que se determine de forma definitiva el número y naturaleza de estas.

9. PERSPECTIVAS

Es imprescindible darle continuidad a esta investigación realizando en la segunda fase un estudio exhaustivo de la expresión diferencial y regulación de las posibles isoformas de *FAD2* en los distintos tejidos y etapas de desarrollo de la planta. Para el desarrollo de esta segunda fase ya se tienen diseñados los *primers* dirigidos a amplificar las dos isoformas de mRNA de *FAD2* y están en proceso de compra.

10. BIBLIOGRAFÍA

- Abe, K., Araki, E., Suzuki, Y., Toki, S., & Saika, H. (2018). Production of high oleic/low linoleic rice by genome editing. *Plant Physiology and Biochemistry*, 131(April), 58–62. <https://doi.org/10.1016/j.plaphy.2018.04.033>
- Al-Zeer, M. A., Dutkiewicz, M., von Hacht, A., Kreuzmann, D., Röhrs, V., & Kurreck, J. (2019). Alternatively spliced variants of the 5'-UTR of the ARPC2 mRNA regulate translation by an internal ribosome entry site (IRES) harboring a guanine-quadruplex motif. *RNA Biology*, 16(11), 1622–1632.

<https://doi.org/10.1080/15476286.2019.1652524>

- Al Amin, N., Ahmad, N., Wu, N., Fu, X., Nan, W., Bo, X., Ma, T., & Wang, P. (2018). An efficient transient assay for crispr cas9 system delivering targeted mutation using synthetic oligo sgrna in soybean (*Glycine max*). *Pakistan Journal of Botany*, *50*(6), 2223–2230.
- Al Amin, N., Ahmad, N., Wu, N., Pu, X., Ma, T., Du, Y., Bo, X., Wang, N., Sharif, R., & Wang, P. (2019). CRISPR-Cas9 mediated targeted disruption of FAD2-2 microsomal omega-6 desaturase in soybean (*Glycine max*.L). *BMC Biotechnology*, *19*(1), 1–10. <https://doi.org/10.1186/s12896-019-0501-2>
- Bahariah, B., Masani, M. Y. A., Rasid, O. A., & Parveez, G. K. A. (2021). Multiplex CRISPR/Cas9-mediated genome editing of the FAD2 gene in rice: a model genome editing system for oil palm. *Journal of Genetic Engineering and Biotechnology*, *19*(1). <https://doi.org/10.1186/s43141-021-00185-4>
- Banilas, G., Moressis, A., Nikoloudakis, N., & Hatzopoulos, P. (2005). Spatial and temporal expressions of two distinct oleate desaturases from olive (*Olea europaea* L.). *Plant Science*, *168*(2), 547–555. <https://doi.org/10.1016/j.plantsci.2004.09.026>
- Baquero, M. (2017). *Modelo De Negocio Para El Cultivo De Higuierilla (Ricinus Communis L) En La Finca Capuriche, Ubicada En El Municipio De Hato Corozal, Departamento Del Casanare*. Universidad Distrital Francisco José De Caldas.
- Berman, P., Nizri, S., & Wiesman, Z. (2011). Castor oil biodiesel and its blends as alternative fuel. *Biomass and Bioenergy*, *35*(7), 2861–2866. <https://doi.org/10.1016/j.biombioe.2011.03.024>
- Bertozzo, F., & Machado, I. S. (2010). Meios de cultura no desenvolvimento de ápices caulinares de mamoneira (*Ricinus communis* L.) in vitro. *Ciencia e Agrotecnologia*, *34*(6), 1477–1482. <https://doi.org/10.1590/S1413-70542010000600018>
- Bhunja, R. K., Kaur, R., & Maiti, M. K. (2016). Metabolic engineering of fatty acid biosynthetic pathway in sesame (*Sesamum indicum* L.): assembling tools to develop nutritionally desirable sesame seed oil. *Phytochemistry Reviews*, *15*(5), 799–811. <https://doi.org/10.1007/s11101-015-9424-2>
- Chan, A. P., Crabtree, J., Zhao, Q., Lorenzi, H., Orvis, J., Puiu, D., Melake-Berhan, A., Jones, K. M., Redman, J., Chen, G., Cahoon, E. B., Gedil, M., Stanke, M., Haas, B. J., Wortman, J. R., Fraser-Liggett, C. M., Ravel, J., & Rabinowicz, P. D. (2010). Draft genome sequence of the oilseed species *Ricinus communis*. *Nature Biotechnology*, *28*(9), 951–956. <https://doi.org/10.1038/nbt.1674>
- Chaudhary, S., Khokhar, W., Jabre, I., Reddy, A. S. N., Byrne, L. J., Wilson, C. M., & Syed, N. H. (2019). Alternative splicing and protein diversity: Plants versus animals. *Frontiers in Plant Science*, *10*(June), 1–14. <https://doi.org/10.3389/fpls.2019.00708>
- Chen, Y., Fu, M., Li, H., Wang, L., Liu, R., Liu, Z., Zhang, X., & Jin, S. (2021). High-oleic acid content, nontransgenic allotetraploid cotton (*Gossypium hirsutum* L.) generated by knockout of GhFAD2 genes with CRISPR/Cas9 system. *Plant Biotechnology Journal*,

19(3), 424–426. <https://doi.org/10.1111/pbi.13507>

- Chi, X., Yang, Q., Lu, Y., Wang, J., Zhang, Q., Pan, L., Chen, M., He, Y., & Yu, S. (2011). Genome-wide analysis of fatty acid desaturases in soybean (*Glycine max*). *Plant Molecular Biology Reporter*, 29(4), 769–783. <https://doi.org/10.1007/s11105-010-0284-z>
- Correal, A. G., Restrepo-Osorio, C., Álvarez, J. C., & Villanueva-Mejía, D. F. (2019). Direct in vitro regeneration of castor bean plants (*Ricinus communis*) using epicotyls. *Bioscience Journal*, January, 347–355. <https://doi.org/10.14393/bj-v35n2a20198-41475>
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), 1–4. <https://doi.org/10.1093/gigascience/giab008>
- Dar, A. A., Choudhury, A. R., Kancharla, P. K., & Arumugam, N. (2017). The FAD2 gene in plants: Occurrence, regulation, and role. *Frontiers in Plant Science*, 8(October), 1–16. <https://doi.org/10.3389/fpls.2017.01789>
- Dehghan Nayeri, F., & Yarizade, K. (2014). Bioinformatics study of delta-12 fatty acid desaturase 2 (FAD2) gene in oilseeds. *Molecular Biology Reports*, 41(8), 5077–5087. <https://doi.org/10.1007/s11033-014-3373-5>
- Doyle, J., & Doyle, J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*, 19, 11–15.
- Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. (2021). AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling*, 61(8), 3891–3898. <https://doi.org/10.1021/acs.jcim.1c00203>
- Figueroa-Varela, P. (2020). *Desarrollo de una línea de Ricinus communis (Malpighiales: Euphorbiaceae) editada genéticamente con la tecnología CRISPR/Cas9 para mayor producción de ácido ricinoleico-fase I*.
- Garrison, E., & Marth, G. (2012). *Haplotype-based variant detection from short-read sequencing* (arXiv:1207.3907 [q-bio.GN]). arXiv preprint.
- Geneious Prime 2022.2.2. (2022). <http://www.geneious.com/>
- Gordon, D. R., Tancig, K. J., Onderdonk, D. A., & Gantz, C. A. (2010). Assessing the invasive potential of biofuel species proposed for Florida and the United States using the Australian Weed Risk Assessment. *Biomass and Bioenergy*, 35(1), 74–79. <https://doi.org/10.1016/j.biombioe.2010.08.029>
- Hajiahmadi, Z., Abedi, A., Wei, H., Sun, W., Ruan, H., Zhuge, Q., & Movahedi, A. (2020). Identification, evolution, expression, and docking studies of fatty acid desaturase genes in wheat (*Triticum aestivum* L.). *BMC Genomics*, 21(1), 1–20. <https://doi.org/10.1186/s12864-020-07199-1>

- Heppard, E. P., Kinney, A. J., Stecca, K. L., & Miao, G. H. (1996). Developmental and growth temperature regulation of two different microsomal ω -6 desaturase genes in soybeans. *Plant Physiology*, *110*(1), 311–319. <https://doi.org/10.1104/pp.110.1.311>
- Hernández, M. L., Mancha, M., & Martínez-Rivas, J. M. (2005). Molecular cloning and characterization of genes encoding two microsomal oleate desaturases (FAD2) from olive. *Phytochemistry*, *66*(12), 1417–1426. <https://doi.org/10.1016/j.phytochem.2005.04.004>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kishore, K., Sinha, S. K., Kumar, R., Gupta, N. C., Dubey, N., & Sachdev, A. (2007). Isolation and characterization of microsomal ω -6-desaturase gene (fad2-1) from soybean. *Indian Journal of Experimental Biology*, *45*(4), 390–397.
- Kongcharoensuntorn, W. (2001). Isolation and analysis of cotton genomic clones encompassing a fatty acid desaturase (FAD2) gene.
- Kõressaar, T., Lepamets, M., Kaplinski, L., Raime, K., Andreson, R., & Remm, M. (2018). Primer3-masker: Integrating masking of template sequence with primer design software. *Bioinformatics*, *34*(11), 1937–1938. <https://doi.org/10.1093/bioinformatics/bty036>
- Koressaar, T., & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, *23*(10), 1289–1291. <https://doi.org/10.1093/bioinformatics/btm091>
- Langeland, K. A., Cherry, H. M., McCormick, C. M., & Craddock Burks, K. A. (2008). *Identification and Biology of Non-native Plants in Florida's Natural Areas*. University of Florida.
- Li, L., Wang, X., Gai, J., & Yu, D. (2008). Isolation and characterization of a seed-specific isoform of microsomal omega-6 fatty acid desaturase gene (FAD2-1B) from soybean. *Mitochondrial DNA*, *19*(1), 28–36. <https://doi.org/10.1080/10425170701207208>
- Lim, T. K. (2016). Edible Medicinal and Non-Medicinal Plants. In *Edible Medicinal and Non-Medicinal Plants* (Vol. 10). Springer Netherlands. <https://doi.org/10.1007/978-94-017-7276-1>
- Lin, J. T., & Arcinas, A. (2007). Regiospecific analysis of diricinoleoylacylglycerols in castor (*Ricinus communis* L.) oil by electrospray ionization-mass spectrometry. *Journal of Agricultural and Food Chemistry*, *55*(6), 2209–2216. <https://doi.org/10.1021/jf063105f>
- Liu, Q., Singh, S., & Green, A. (2002). High-oleic and high-stearic cottonseed oils: nutritionally improved cooking oils developed using gene silencing. *Journal of the American College of Nutrition*, *21*(October 2014), 205S-211S. <https://doi.org/10.1080/07315724.2002.10719267>

- Liu, Q., Wu, M., Zhang, B., Shrestha, P., Petrie, J., Green, A. G., & Singh, S. P. (2017). Genetic enhancement of palmitic acid accumulation in cotton seed oil through RNAi down-regulation of ghKAS2 encoding β -ketoacyl-ACP synthase II (KASII). *Plant Biotechnology Journal*, 15(1), 132–143. <https://doi.org/10.1111/pbi.12598>
- Lloyd, G., & McCown, B. (1980). Commercially-Feasible Micropropagation of Mountain Laurel, *Kalmia latifolia*, by Use of Shoot-Tip Culture. *Proceedings of the International Plant Propagators Society*, 30, 421–427.
- Martínez-Rivas, J. M., García-Díaz, M. T., & Mancha, M. (2000). Temperature and oxygen regulation of microsomal oleate desaturase (FAD2) from sunflower. *Biochemical Society Transactions*, 28(6), 890–892. <https://doi.org/10.1042/bst0280890>
- Miquel, M., & Browse, J. (1992). Arabidopsis mutants deficient in polyunsaturated fatty acid synthesis: Biochemical and genetic characterization of a plant oleoyl-phosphatidylcholine desaturase. *Journal of Biological Chemistry*, 267(3), 1502–1509. [https://doi.org/10.1016/s0021-9258\(18\)45974-1](https://doi.org/10.1016/s0021-9258(18)45974-1)
- Mrozek, E., & Funicelli, N. A. (1982). Effect of zinc and lead on germination of *Spartina alterniflora* loisel seeds at various salinities. *Environmental and Experimental Botany*, 22(1), 23–32. [https://doi.org/10.1016/0098-8472\(82\)90005-3](https://doi.org/10.1016/0098-8472(82)90005-3)
- Mubofu, E. B. (2016). Castor oil as a potential renewable resource for the production of functional materials. *Sustainable Chemical Processes*, 4(1), 1–12. <https://doi.org/10.1186/s40508-016-0055-8>
- Murashige, T., & Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum*, 15, 474–497.
- Mutlu, H., & Meier, M. A. R. (2010). Castor oil as a renewable resource for the chemical industry. *European Journal of Lipid Science and Technology*, 112(1), 10–30. <https://doi.org/10.1002/ejlt.200900138>
- Ogunniyi, D. S. (2006). Castor oil: A vital industrial raw material. *Bioresource Technology*, 97(9), 1086–1091. <https://doi.org/10.1016/j.biortech.2005.03.028>
- Okuley, J., Lightner, J., Feldmann, K., Yadav, N., Lark, E., & Browse, J. (1994). Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell*, 6(1), 147–158. <https://doi.org/10.2307/3869682>
- Perry, B. A. (1943). Chromosome number and phylogenetic relationships in the Euphorbiaceae. *American Journal of Botany*, 30(7), 527–543. <https://doi.org/10.1002/j.1537-2197.1943.tb14796.x>
- Pham, A. T., Lee, J. D., Shannon, J. G., & Bilyeu, K. D. (2011). A novel FAD2-1 A allele in a soybean plant introduction offers an alternate means to produce soybean seed oil with 85% oleic acid content. *Theoretical and Applied Genetics*, 123(5), 793–802. <https://doi.org/10.1007/s00122-011-1627-3>
- Ramchiary, N., & Kole, C. (2017). Compendium of Plant Genomes The brassica Genome. In *Compendium of Plant Genomes: The brassica Genome*. <https://link.springer.com/content/pdf/10.1007%2F978-3-319-66117-9.pdf>

- Rausch, T., Fritz, M. H. Y., Untergasser, A., & Benes, V. (2020). Tracy: Basecalling, alignment, assembly and deconvolution of sanger chromatogram trace files. *BMC Genomics*, *21*(1), 1–9. <https://doi.org/10.1186/s12864-020-6635-8>
- Saadaoui, E., Martín, J. J., Tlili, N., & Cervantes, E. (2017). Castor bean (*Ricinus communis* L.): Diversity, seed oil and uses. *Oilseed Crops*, 19–33. <https://doi.org/10.1002/9781119048800.ch2>
- Sambrook, J. (2001). *Molecular Cloning : a Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- Sánchez, M., Castañeda, R., & Castañeda S., M. (2016). Usos y potencialidad de la Higuierilla (*Ricinus communis*) en sistemas agroforestales en Colombia. *Pubvet*, *10*(6), 507–512. <https://doi.org/10.22256/pubvet.v10n6.507-512>
- Solís Bonilla, J. L., Muñoz Orozco, A., Escalante Estrada, J. A. S., & Zamarripa Colmenero, A. (2017). Crecimiento de variedades y componentes del rendimiento de higuierilla (*Ricinus communis* L.) en Montecillo, Estado de México. *Revista Mexicana de Ciencias Agrícolas*, *7*(2), 311–323. <https://doi.org/10.29312/remexca.v7i2.346>
- Soto-Blanco, B., Sinhorini, I. L., Gorniak, S. L., & Schumacher-Henrique, B. (2002). *Ricinus communis* cake poisoning in a dog. *Veterinary and Human Toxicology*, *44*(3), 155–156.
- Sujatha, M., Reddy, T. P., & Mahasi, M. J. (2008). Role of biotechnological interventions in the improvement of castor (*Ricinus communis* L.) and *Jatropha curcas* L. *Biotechnology Advances*, *26*(5), 424–435. <https://doi.org/10.1016/j.biotechadv.2008.05.004>
- Susunaga-Gomez, D. (2022). Silenciamiento de leg de la desaturasa de ácidos grasos 2 (FAD2) en *Ricinus communis* (Malpighiales: Euphorbiaceae) mediante edición genética basada en el uso de la tecnología CRISPR/Cas9.
- Tang, G. Q., Novitzky, W. P., Carol Griffin, H., Huber, S. C., & Dewey, R. E. (2005). Oleate desaturase enzymes of soybean: Evidence of regulation through differential stability and phosphorylation. *Plant Journal*, *44*(3), 433–446. <https://doi.org/10.1111/j.1365-313X.2005.02535.x>
- Tilahun, M., Mulugeta, D., & Manju, S. (2013). An alternative safer and cost effective surface sterilization method for sugarcane (*Saccharum officinarum* L.) explants. *African Journal of Biotechnology*, *12*(44), 6282–6286. <https://doi.org/10.5897/ajb2013.12481>
- Trott, O., & Olson, A. J. (2009). AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *Journal of Computational Chemistry*, *31*, 455–461. <https://doi.org/DOI.10.1002/jcc.21334>
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Research*, *40*(15), 1–12. <https://doi.org/10.1093/nar/gks596>
- Wang, L., Jiang, X., Wang, L., Wang, W., Fu, C., Yan, X., & Geng, X. (2019). A survey of transcriptome complexity using PacBio single-molecule real-time analysis combined with Illumina RNA sequencing for a better understanding of ricinoleic acid biosynthesis

- in *Ricinus communis*. *BMC Genomics*, 20(1), 1–17. <https://doi.org/10.1186/s12864-019-5832-9>
- Wang, Z., Wang, B., Dong, L., Xu, L., Tang, S., & Wen, Q. (2020). Cloning of CcFAD2 gene from *camellia chekiangoleosa* and analysis on its expression during kernel development. *Journal of Plant Resources and Environment*, 29(5), 1–8. <https://doi.org/doi:10.3969/j.issn.1674-7895.2020.05.01>
- Wu, N., Lu, Q., Wang, P., Zhang, Q., Zhang, J., Qu, J., & Wang, N. (2020). Construction and analysis of GmFAD2-1A and GmFAD2-2A soybean fatty acid desaturase mutants based on CRISPR/Cas9 technology. *International Journal of Molecular Sciences*, 21(3), 1–12. <https://doi.org/10.3390/ijms21031104>
- Xiao, G., Zhang, Z. Q., Yin, C. F., Liu, R. Y., Wu, X. M., Tan, T. L., Chen, S. Y., Lu, C. M., & Guan, C. Y. (2014). Characterization of the promoter and 5'-UTR intron of oleic acid desaturase (FAD2) gene in *Brassica napus*. *Gene*, 545(1), 45–55. <https://doi.org/10.1016/j.gene.2014.05.008>
- Xue, Y., Yin, N., Chen, B., Liao, F., Win, A. N., Jiang, J., Wang, R., Jin, X., Lin, N. A., & Chai, Y. (2017). Molecular cloning and expression analysis of two FAD2 genes from chia (*Salvia hispanica*). *Acta Physiologiae Plantarum*, 39(4). <https://doi.org/10.1007/s11738-017-2390-0>
- Yang, J., Roy, A., & Zhang, Y. (2013a). BioLiP: A semi-manually curated database for biologically relevant ligand-protein interactions. *Nucleic Acids Research*, 41(D1), 1096–1103. <https://doi.org/10.1093/nar/gks966>
- Yang, J., Roy, A., & Zhang, Y. (2013b). Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics*, 29(20), 2588–2595. <https://doi.org/10.1093/bioinformatics/btt447>
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13, 134. <https://doi.org/10.1186/1471-2105-13-134>
- Zhang, D., Pirtle, I. L., Park, S. J., Nampaisansuk, M., Neogi, P., Wanjie, S. W., Pirtle, R. M., & Chapman, K. D. (2009). Identification and expression of a new delta-12 fatty acid desaturase (FAD2-4) gene in upland cotton and its functional expression in yeast and *Arabidopsis thaliana* plants. *Plant Physiology and Biochemistry*, 47(6), 462–471. <https://doi.org/10.1016/j.plaphy.2008.12.024>
- Zhang, J. X., Wang, X. Y., Feng, Z. Z., Geng, X. J., Mu, S. M., Huo, H. Y., Tong, H., Li, M. Z., Li, Y., Chi, Y., & Chen, Y. S. (2016). In vitro establishment of a highly effective method of castor bean (*Ricinus communis* L.) regeneration using shoot explants. *Journal of Integrative Agriculture*, 15(6), 1417–1422. [https://doi.org/10.1016/S2095-3119\(15\)61286-2](https://doi.org/10.1016/S2095-3119(15)61286-2)
- Zhiguo, E., Chen, C., Yang, J., Tong, H., Li, T., Wang, L., & Chen, H. (2019). Genome-wide analysis of fatty acid desaturase genes in rice (*Oryza sativa* L.). *Scientific Reports*, 9(1), 1–11. <https://doi.org/10.1038/s41598-019-55648-z>

Zhou, X. R., Singh, S. P., & Green, A. G. (2013). Characterisation of the FAD2 gene family from *Hiptage benghalensis*: A ricinoleic acid accumulating plant. *Phytochemistry*, 92, 42–48. <https://doi.org/10.1016/j.phytochem.2013.05.006>