UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
DOCTORADO EN CIENCIAS BIOMÉDICAS
INSTITUTO DE ECOLOGÍA

ACONDICIONAMIENTO MÁTRICO EN SEMILLAS DE *Ceiba aesculifolia*: EL ESTÍMULO HÍDRICO COMO FACTOR DETERMINANTE EN LAS RESPUESTAS FISIOLÓGICA, METABÓLICA Y GENÉTICA EN SEMILLAS ALMACENADAS

TESIS
QUE PARA OBTENER EL GRADO DE:
DOCTORA EN CIENCIAS

PRESENTA:
XIMENA GÓMEZ MAQUEO BRIBIESCA

DIRECTORA DE TESIS
DRA. ALICIA GAMBOA DE BUEN
INSTITUTO DE ECOLOGÍA

COMITÉ TUTOR
DR. ELEAZAR MARTÍNEZ BARAJAS
FACULTAD DE QUÍMICA
DR. LUIS DAVID ALCARAZ PERAZA
FACULTAD DE CIENCIAS

CIUDAD DE MÉXICO, 3 DE NOVIEMBRE DE 2021

Agradecimientos institucionales

Agradezco la beca otorgada por el Consejo Nacional de Ciencia y Tecnología (CONACyT, No. 280483).

Agradezco la beca otorgada por la Dirección General de Asuntos del Personal Académico (DGAPA UNAM, proyecto: IN203218).

Este proyecto fue financiado por recursos otorgados a la Dra. Gamboa de Buen por los siguientes programas:

- Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT UNAM) con el proyecto: IN203218.
- Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT UNAM) con el proyecto: IN210621.
- Dirección General del Sector Primario y Recursos Naturales Renovables de la Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) y por la Comisión Nacional para la Biodiversidad (CONABIO) con el proyecto: OE001, "*Ceiba aesculifolia* como especie modelo para el estudio molecular del envejecimiento de las semillas en relación al desempeño germinativo".

Agradezco al Dr. Noé Velázquez Rosas del Centro de Investigaciones Tropicales, Universidad Veracruzana, por su asistencia en campo para la colecta del material vegetal utilizado en la tesis, así como sus aportaciones conceptuales para el análisis de los resultados de germinación.

Agradezco a los Dres. Alejandro Sánchez-Flores, Ricardo Grande, Verónica Jiménez-Jacinto y Leticia Vega-Alvarado de la Unidad de Secuenciación Masiva y Bioinformática del Instituto de Biotecnología, UNAM, por el servicio de secuenciación y ensamble, así como la asesoría relacionada con el manejo y procesamiento bioinformático de los datos transcriptómicos.

Agradezco al Ing. Rodrigo García-Herrera de la unidad de supercómputo del LANCIS-Instituto de Ecología y al M. en I. José Miguel Baltazar de la unidad de cómputo del Instituto de Ecología UNAM por su asistencia en los análisis bioinformáticos.

Agradezco a la M. en C. Karina Jiménez Durán de la Unidad de Microscopía de la Unidad de Servicios de Apoyo a la Investigación y a la Industria (USAII) UNAM, por la preparación y procesamiento de las muestras para microscopía óptica.

Agradezco al Dr. Roberto Benech Arnold de la Facultdad de Agronomía de la Universidad de Buenos Aires, Argentina por sus asistencia en la cuantificación del contenido de ABA presente en las semillas de *C. aesculifolia*.

Agradecimientos personales

Agradezco a mi hermosa familia, Ceci, Antoine y Toño quienes han sido mi apoyo y fuente de inspiración para siempre dar el máximo en cada faceta de mi desarrollo personal y profesional. Los amo infinito y siempre me he sentido orgullosa de formar parte de este equipo tan creativo, resiliente, guapo, brillante, amoroso y reflexivo. Esta tesis es otra ofrenda que les hago a ustedes y al universo como agradecimiento por haberme otorgado tan hermoso regalo.

A mi tutora, Alicia Gamboa de Buen por esta increíble aventura que ha sido el posgrado, tanto en la maestría como el doctorado. Me siento orgullosa de haber tenido la oportunidad de contar la biología de las ceibas desde la perspectiva biológica en compañía de Alicia. Mis dos posgrados han sido no solamente una oportunidad de aprender y prepararme como investigadora, sino que además ha sido una de las experiencias más gozosas que he experimentado. A pesar de las dificultades que inevitablemente surgieron a lo largo de este trabajo, mi estancia en el laboratorio de Alicia ha sido un verdadero privilegio y me permitió observar cómo hacerle frente a estas dificultades con creatividad, racionalidad y entusiasmo. Aunque esta tesis marca el final de una trayectoria alumno-maestro, en realidad es sólo un paso más en un viaje de colaboración y de avance científico que todavía sigue dando frutos.

A mi colega Diana Soriano Fernández, que además de sus contribuciones directas a la ciencia presentada en esta tesis, también fue una maestra indirecta para mí. A pesar del inicio algo turbulento de nuestra relación como colegas, Diana me dio la oportunidad de aprender a relacionarme de otra manera y de inicia mi carrera como docente. Esa lección de vida y la confianza que recibí fueron fundamentales para mí y para darle forma a este trabajo.

A mis colaboradores Jorge Martínez Villegas, Laura Figueroa Corona, Sandra Alvarado, Marimar Garciadiego, que en diferentes ocasiones contribuyeron a los resultados presentados en este trabajo. La experiencia personal y profesional que cada uno trajo a este trabajo me ayudó a madurar las ideas presentadas aquí, además de que la interacción que tuvimos en persona hizo que mi permanencia en el lab fuera lo más amena y emocionante que pudiera ser. Trabajar con todos siempre fue divertido y sumamente inspirador.

A mis entrañables compañeros del laboratorio, Erik, Alexis, Ale Ro, Elsa, Humberto, Ángel, Esther, Alfredo y Martín, amigos, cómplices en nuestras pequeñas aventuras y travesuras que hicieron todos estos años mágicos y llenos de emoción. Tenerlos cerca y convivir tanto en el laboratorio como fuera de él fue una hermosa experiencia que siempre llevo cerca del corazón.

A mis alumnos de licenciatura, Karina, Edwin y Leslie, que a pesar de mi inexperiencia como tutora han sido entusiastas, hemos podido generar un excelente equipo de trabajo y han adoptado la historia de Ceiba como suya. El ver su entusiasmo y su esfuerzo por aprender me inspira en seguir aprendiendo a ser el mejor ejemplo posible de cómo desarrollarse como profesionistas. Me han dado la oportunidad de ser lo que Alicia es para mí, y espero poder transmitirles la pasión y el amor por la ciencia que yo he experimentado de mis maestros. La docencia es un privilegio y un honor que muchas veces no recibe la consideración que debería; ahora que es mi

turno de pasarles la estafeta a la siguiente generación, espero poder ofrecerles la mejor experiencia posible.

Al Dr. Roberto Benech por darme la oportunidad de realizar una estancia de investigación en su laboratorio. Las reflexiones que surgieron del trabajo que realicé durante mi estancia fueron muy importantes no sólo para los fines de la tesis, sino también para mi propia formación como científica. El espacio que pude tener para enfrentarme a las distintas aristas que tiene la ciencia fue un gran tesoro que me acompañará por el resto de mi carrera.

Tabla de Contenidos

Índice de figuras.

RESUMEN

La germinación es una etapa crítica en el ciclo de vida de las plantas ya que permite la continuidad entre las diferentes generaciones de una especie. Las semillas a su vez fungen como estructuras de resistencia a condiciones adversas y permiten la diseminación de los individuos en el tiempo y en el espacio. Estas características se ven afectadas por el proceso de maduración en la planta madre, así como la permanencia prolongada en almacenamiento en condiciones subóptimas de temperatura y humedad. El estudio de estas características y del proceso de germinación se ha estudiado ampliamente en especies modelo o en especies domesticadas. Sin embargo, estas especies pertenecen a un grupo reducido de familias de plantas, principalmente herbáceas de las familias Poaceae y Brassicaseae que tienen un ciclo de vida anual, que presentan latencia fisiológica y su hábitat original es de clima templado.

En años recientes se ha hecho énfasis en la necesidad de ampliar la diversidad de especies estudiadas para tener una mejor representación de los procesos fisiológicos que ocurren durante la germinación en especies pertenecientes a otros ambientes y a otras familias de espermatofitas. En este trabajo se estudió la dinámica de germinación de las semillas de *Ceiba aesculifolia*, una especia arbórea de la familia Malvaceae y que habita en el bosque tropical caducifolio. Esta especie además presenta dos fenotipos de respuesta al tratamiento de acondicionamiento mátrico, el cual emula los ciclos de hidratación/deshidratación que ocurren en campo debidos a las lluvias esporádicas durante la estación seca. Se seleccionaron lotes colectados en distintos años y con respuesta contrastante al tratamiento para evaluar la dinámica de germinación así como los cambios asociados a la respuesta al acondicionamiento, a nivel transcriptómico y bioquímico. Se diseñó una estrategia de muestreo en función del contenido relativo de agua (CRA) de las semillas para la selección de estadios discretos durante el proceso de germinación. Esta estrategia consistió en pesar individualmente cada semilla durante la hidratación, usando el incremento en el peso fresco como referencia para estimar el CRA y registrar a su vez el CRA en el cual ocurren eventos característicos del proceso de germinación en semillas endospérmicas: ruptura de testa (TR) y ruptura de endospermo (G). Se seleccionaron cinco momentos fisiológicos representativos del proceso de hidratación asociados al cambio en el CRA (semilla seca, 16% CRA, 50% CRA, TR y G). Para el análisis transcriptómico se recurrió a un ensamble *de novo* por medio de secuenciación Illumina Hiseq2000 y del análisis de enriquecimiento de procesos biológicos comparando contra las categorías funcionales descritas en *Arabidopsis thaliana*. Los análisis bioquímicos incluyeron la cuantificación de lípidos almacenados, del contenido de ABA y la medición de la actividad relativa de pectin metil esterasas (PMEs) en diferentes estadios fisiológicos durante la germinación.

La respuesta favorable al acondicionamiento (fenotipo positivo) se reflejó en una mejoría en la dinámica de germinación con respecto a semillas no tratadas. Esta mejoría estuvo acompañada de un reajuste a nivel de abundancia de transcritos asociados a diversos procesos biológicos fundamentales para la germinación. Sin embargo, el acondicionamiento no genera un programa de desarrollo distinto al programa de germinación, sino que favorece que ciertos procesos biológicos ocurran de manera coordinada. Asimismo, se detectaron 61 genes expresados diferencialmente en ciertos estadios del proceso, los cuales participan en diversas respuestas a estímulos ambientales, como estrés por patógenos, estrés osmótico y estrés oxidativo. A nivel bioquímico se observó una movilización temprana de los lípidos almacenados y una disminución

del contenido de ABA en las semillas secas post-tratamiento. Conforme progresó el proceso de germinación, las diferencias entre semillas tratadas y no tratadas disminuyeron. La actividad relativa de PME no fue afectada por el acondicionamiento; sin embargo, la dinámica de actividad relativa concuerda con lo reportado para otras especies modelo, indicando que el fenotipo positivo al acondicionamiento puede ser considerado como el patrón de germinación óptimo de *C. aesculifolia*.

La respuesta negativa al acondicionamiento (fenotipo negativo) no mostró una mejoría en el desempeño germinativo con respecto a semillas no tratadas e incluso su desempeño fue menor comparado con los lotes con fenotipo positivo. En este fenotipo también se observó reajuste del transcriptoma en respuesta al tratamiento; sin embargo, estos cambios se ven opacados por el transcriptoma inicial de las semillas con fenotipo negativo. Este transcriptoma inicial fue el factor determinante del desempeño observado en ensayos subsecuentes. En comparación con el fenotipo positivo, las semillas con repuesta negativa al acondicionamiento muestran un transcriptoma similar a semillas en estadios tempranos de hidratación, además de presentar un enriquecimiento de procesos biológicos asociados a respuestas a estrés y metabolismo secundario. A nivel bioquímico, las semillas no tratadas con fenotipo negativo muestran un mayor contenido de ABA, el cual aumenta en estadios de hidratación previos al 50% de CRA. Este aumento podría estar asociado a una permanencia más prolongada en estadios de hidratación bajos, fase que fue aproximadamente 45% más prolongada en el fenotipo negativo con respecto al positivo. Asimismo, no se observó una movilización temprana de lípidos en respuesta al acondicionamiento. Estos resultados sugieren que la cuantificación de lípidos en la semilla seca post-tratamiento, en conjunto con los ensayos de germinación, pueden ser técnicas fáciles y rápidas para determinar si un lote será adecuado o no para el diseño de experimentos subsecuentes o su uso en la propagación de plantas requeridas para programas de restauración.

Finalmente, en contraste con especies modelo o domesticadas, el estudio y conservación de especies silvestres debe priorizar la variabilidad genética y fisiológica de los lotes colectados para una especie dada. Esta variabilidad presenta una desventaja en el estudio de la germinación desde la estrategia de tiempo cronológico. La aproximación utilizada en este estudio, basada en el CRA, permitió el contraste de estadios discretos equivalentes entre lotes de semillas con dinámicas de germinación altamente variables. En otras especies silvestres también se han utilizado aproximaciones enfocadas en la selección de estadios discretos para asegurar la comparabilidad de los grupos evaluados. El uso del CRA como una estrategia de muestreo durante la germinación podría favorecer el diseño de ensayos posteriores a nivel fisiológico, genético y molecular en otras especies.

ABSTRACT

Germination is a critical stage in the life cycle of plants as it allows continuity between the different generations of a species. The seeds in turn act as resistance structures to adverse conditions and allow the dissemination of individuals in time and space. These characteristics are affected by the maturation process in the mother plant, as well as prolonged storage in suboptimal conditions of temperature and humidity. The study of these characteristics and the germination process has been widely studied in model species or in domesticated species. However, these species belong to a reduced group of plant families, mainly herbaceous of the Poaceae and Brassicaseae families that have an annual life cycle, present physiological dormancy, and thrive in temperate habitats.

In recent years, emphasis has been placed on the need to expand the diversity of species studied to have a better representation of the physiological processes that occur during germination in species belonging to other environments and to other families of spermatophytes. In this work, the germination dynamics of the seeds of *Ceiba aesculifolia*, an arboreal species of the Malvaceae family that inhabits the tropical deciduous forest, was studied. This species also presents two phenotypes of response to the matric conditioning treatment, which emulates the hydration/dehydration cycles that occur in the field due to sporadic rains during the dry season. Lots collected in different years and with contrasting response to treatment were selected to evaluate the germination dynamics as well as the changes associated with the response to conditioning, at the transcriptomic and biochemical level. A sampling strategy was designed based on the seeds relative water content (RWC) for the selection of discrete stages during the germination process. This strategy consisted of individually weighing each seed during imbibition, using the increase in the fresh weight as a reference to estimate the RWC and in turn record the RWC in which two characteristic events of the germination process of endospermic seeds take place: the moment of testa rupture (TR) and endosperm rupture (G). Five representative physiological moments of the hydration process associated with the change in CRA (dry seed, 16% RWC, 50% RWC, TR and G) were selected. For the transcriptomic analysis, a de novo assembly was used by means of Illumina Hiseq2000 sequencing and the enrichment analysis of biological processes, comparing against the functional categories described in Arabidopsis thaliana. The biochemical analyzes included the quantification of stored lipids, ABA content and the measurement of the relative activity of pectin methyl esterases (PMEs) at different physiological stages during germination.

The favorable response to priming was reflected in an improvement in the germination performance with respect to untreated seeds. This improvement was accompanied by a fine adjustment of the transcriptome of the germinating seeds. This fine tuning of the transcriptome generated a readjustment of several fundamental biological processes for germination; however it does not generate a different development program. Likewise, 61 differentially expressed genes were detected in certain stages of the process, which participate in responses to environmental stimuli, such as stress by pathogens, osmotic stress and oxidative stress. At the biochemical level, early mobilization of stored lipids and a decrease in ABA content were observed in the treated dry seeds. As the germination process progressed, the differences between treated and untreated seeds decreased. The relative activity of PME was not affected by priming, however the PME activity agrees with that reported for other model species, indicating

that the positive response to priming can be considered as the optimal germination pattern of *C. aesculifolia* seeds.

The negative response to priming did not show an improvement in the germination performance with respect to untreated seeds and its performance was even lower compared to the lots with a positive phenotype. In this phenotype, a fine adjustment of the transcriptome was also observed in response to treatment; however these changes are overshadowed by the initial transcriptome of seeds with a negative phenotype. This initial transcriptome was the determining factor for the performance observed in subsequent trials. In contrast to the positive phenotype, seeds with a negative response to priming show a transcriptome similar to seeds in early stages of hydration, in addition to presenting an enrichment of biological processes associated with responses to stress and secondary metabolism. At the biochemical level, untreated seeds with a negative phenotype show a higher content of ABA, which increases in hydration stages prior to 50% RWC. This increase could be associated with a longer stay in low hydration stages, a phase that was approximately 45% longer in the negative phenotype compared to the positive one. Also, early lipid mobilization in response to priming was not observed. These results suggest that lipid quantification in the post-treatment dry seed, in conjunction with germination tests, can be quick and easy techniques to determine whether or not a seed batch will be suitable for the design of subsequent experiments or its use for the plant propagation required in restoration programs.

Finally, in contrast to model or domesticated species, the study and conservation of wild species must prioritize the genetic and physiological variability of the seed batches collected for a given species. This variability presents a disadvantage in the study of germination based on sample collection over chronological time. The approach used in this study, based on the seed RWC, allowed the contrast of equivalent discrete stages between seed batches with highly variable germination performances. In other wild species, approaches focused on the selection of discrete stages have also been used to ensure comparability of the groups evaluated. The use of the RWC as a sampling strategy during germination could favor the design of subsequent tests at the physiological, genetic and molecular levels in other species.

INTRODUCCIÓN GENERAL

La germinación es un proceso complejo en el cual intervienen una amplia gama de factores genéticos, endógenos y ambientales (Joosen *et al.*, 2013). Este proceso comienza con la absorción de agua, que desencadena de manera progresiva los procesos celulares necesarios para completar la germinación y el establecimiento de la nueva planta (Rosental *et al*., 2014).

En general, las semillas son fundamentales para facilitar la producción de especies de interés agrícola, medicinal, textil, etc. (Vázquez-Yañes y Batis, 1996). Las semillas proveen la continuidad entre las generaciones de plantas y contribuyen en la distribución de nuevos individuos en el tiempo y el espacio (Moreira de Carvahlo y Nakagawa, 1988; Fenner y Thompson, 2005). Algunas semillas pueden permanecer por periodos largos de tiempo como parte del banco de semillas o almacenadas *ex situ*, con posibilidad de ir perdiendo viabilidad. El desempeño germinativo de estas semillas puede mejorarse con tratamientos pre-germinativos de acondicionamiento, que consisten en someter a las semillas a uno o varios ciclos de hidratación controlada seguida de un proceso de deshidratación (Varier *et al.,* 2010). Los tratamientos de acondicionamiento han sido ampliamente utilizados para mejorar el desempeño germinativo de especies de interés agrícola y con fines de restauración. Sin embargo, las semillas almacenadas de manera prolongada pueden perder la capacidad de responder favorablemente a este tratamiento. La comprensión de los procesos fisiológicos que ocurren en las semillas antes y durante la germinación y en respuesta a los estímulos del medio permite mejorar el manejo de las especies de interés.

Las semillas son particularmente útiles para el inicio del análisis del transcriptoma de plantas silvestres utilizando técnicas de secuenciación masiva, ya que tanto los transcritos almacenados en semilla como los inducidos durante el proceso germinativo abarcan un espectro muy amplio de categorías funcionales y programas de desarrollo de la planta. Las técnicas de secuenciación masiva también permitirán analizar con mayor precisión el cambio en la expresión génica durante la toma de agua de las semillas con diferente grado de deterioro y establecer una correlación en función del desempeño germinativo.

La germinación es un proceso trifásico dirigido por la absorción de agua en los tejidos de la semilla (Bewley, 1997). La fase 1 de la germinación se caracteriza por una absorción rápida de agua. En la fase 2 se mantiene relativamente constante el contenido de agua, hasta la transición a la fase 3, donde ocurre nuevamente un incremento en el contenido de agua hasta que protruye la radícula a través de las cubiertas seminales (Bewley, 1997). En las semillas endospérmicas de angiospermas la germinación se presenta en dos etapas, siguiendo el mismo patrón trifásico de absorción de agua. La primera etapa comprende la fase 1 de absorción rápida de agua hasta la ruptura de la testa, en un punto intermedio de la fase 2 descrita por Bewley. La ruptura de la testa marca el inicio de la segunda etapa continuando hasta la ruptura del endospermo, proceso que determina el final de la germinación (protrusión de la radícula; Dekkers *et al.*, 2013).

A nivel transcripcional, la germinación depende de la calidad y abundancia de los transcritos acumulados durante el desarrollo de la semilla, que a su vez dependen de las condiciones de crecimiento experimentadas por la planta madre (Rajjou *et al.*, 2012). En transcriptomas de semillas secas, principalmente de especies modelo y/o domesticadas como *Arabidopsis thaliana* (*Arabidopsis*), *Hordeum vulgare* y *Oryza sativa* se han reportado entre 12,000 y 17,000 mRNAs (Nakabayashi *et al.*, 2005; Narsai *et al.*, 2011; Rajjou *et al.*, 2012; Dekkers *et al.*, 2013; Howell *et al.*, 2009). Los análisis comparativos del transcriptoma entre semillas secas y sometidas a distintos tiempos de hidratación indican que los cambios en la

abundancia de mRNAs y metabolitos durante la germinación ocurren en función del tiempo de hidratación de la semilla, el tejido estudiado y el estado fisiológico de la semilla (Howell *et al.*, 2009; Maia *et al.*, 2011; Dekkers *et al.*, 2013).

Dekkers y colaboradores (2013) describieron las diferencias en categorías funcionales y procesos biológicos entre los transcriptomas del eje embrionario y del endospermo en semillas de *Arabidopsis*, analizados en diferentes tiempos de hidratación. Este trabajo sugiere dos momentos fisiológicos importantes en la semilla. El primero comprende la transición del estado seco a la reactivación del metabolismo previo al rompimiento de la testa. El segundo momento fisiológico está asociado con la transición hacia plántula, donde se inicia la regulación transcriptómica de los procesos biológicos que no son fundamentales para la germinación, pero que permitirán el establecimiento de la plántula. El estadio fisiológico correspondiente a la ruptura de testa fue considerado por los autores como el momento fisiológico que presenta los cambios más relevantes para la germinación. Durante el proceso de germinación ocurre un recambio importante de transcritos: los genes relacionados con el desarrollo y maduración de la semilla se degradan, mientras se induce la transcripción de otros genes necesarios durante la germinación (Dekkers *et al.*, 2013). Por otro lado, el tiempo en almacenamiento favorece la degradación de los transcritos contenidos en las semillas, causando alteraciones en el desempeño germinativo (Rajjou *et al.*, 2012).

Asimismo, el desempeño germinativo puede modificarse a través de tratamientos pre-germinativos conocidos colectivamente como acondicionamiento (Paparella *et al.*, 2015). Existen diversas estrategias para implementar este acondicionamiento, pero fundamentalmente implican la hidratación de las semillas para inducir el programa de germinación, seguida de una deshidratación para prevenir la protrusión de la radícula. Posterior a este tratamiento, las semillas se colocan en condiciones adecuadas para la germinación, favoreciendo una germinación rápida y sincrónica en comparación con semillas no tratadas (Paparella *et al.*, 2015). El análisis del transcriptoma de semillas de *Brassica oleracea* sometidas a osmoacondicionamiento (Soeda *et al.*, 2005) reveló que durante este tratamiento se inducen algunos de los procesos tempranos de la germinación correspondientes al primer momento fisiológico descrito por Dekkers y colaboradores (2013). La implementación de los tratamientos de acondicionamiento es un proceso laborioso especie-específico que implica ensayo y error para poder optimizar la técnica, ya que una misma técnica puede tener efectos diferentes dependiendo de la especie o el lote utilizando. Asimismo, se desconoce cuál es el mecanismo por medio del cual el acondicionamiento favorece un cambio en los procesos que ocurren durante la germinación y su relación con los efectos que se observan en la dinámica de germinación: sincronía, menor tiempo medio de germinación, porcentajes finales de germinación más elevados, entre otros (Paparella *et al.*, 2015). A pesar de estas incógnitas, los tratamientos de acondicionamiento se utilizan habitualmente para mejorar el establecimiento de especies de interés agrícola. Además, muchas especies silvestres experimentan estos ciclos de hidratación y deshidratación de manera natural durante la permanencia de las semillas en el suelo (denominado como acondicionamiento natural). Los efectos benéficos del acondicionamiento natural han sido demostrados en diversas especies nativas de México como son *Wigandia urens* (Gamboa deBuen *et al.*, 2006; González Zertuche *et al.*, 2001), *Tecoma stans*, *Cordia megalantha* (Alvarado López *et al.*, 2014), *Albizia saman*, *Cedrela odorata*, *Enterolobium cyclocarpum* y *Swietenia macrophylla* (Peraza Villarreal *et al.*, 2018), así como en *Ceiba aesculifolia* (Alvarado López, 2012). En el caso de *C. aesculifolia* también se ha reportado que las semillas almacenadas en el laboratorio pueden

perder la capacidad de responder de manera favorable al acondicionamiento mátrico (Gómez Maqueo, 2014).

Ceiba aesculifolia (Kunth) Britten y Baker F. (Malvaceae), conocido como ceiba o pochote es un árbol de hasta 15m de altura, hermafrodita y caducifolio. Es una especie originaria de las regiones tropicales subhúmedas de América, y produce abundante algodón sedoso color marfil en el cual se encuentran inmersas las semillas (Pennington y Sarukhán, 1998; Niembro *et al.*, 2010). A diferencia de otras especies del género *Ceiba*, *C. aesculifolia* florece al final de la temporada de secas y los frutos se retienen en la planta madre casi un año, hasta la siguiente estación seca (Lobo *et al*., 2003). La liberación de las semillas hasta la siguiente temporada de secas implica que las semillas permanecen en el suelo varios meses antes de la temporada de lluvias, tiempo durante el cual pueden estar sometidas a ciclos de acondicionamiento natural debidos a lluvias esporádicas (Alvarado-López, 2012). Las semillas de esta especie muestran una respuesta al almacenamiento de tipo ortodoxo (Royal Botanic Gardens Kew, 2020) y son relativamente grandes, permitiendo obtener diversas mediciones a nivel individual (Garciadiego, 2014).

Dado que es una especie pionera-secundaria se considera adecuada para restauración de zonas perturbadas (Alvarado-López, 2012). Asimismo, diversos estudios han mostrado que *C. aesculifolia* es resistente a condiciones estresantes del suelo como con las altas temperaturas y la sequía debido a que tiene la capacidad de retener agua en sus tejidos (Encino Ruiz *et al*., 2013; Hernández Gómez, 2010; Valle Díaz *et al*., 2009).

Los estudios a nivel genético y molecular en esta especie han reportado cambios en la abundancia y diversidad de transcritos almacenados en semillas sometidas a acondicionamiento mátrico, utilizando microarreglos heterólogos para comparar la abundancia y diversidad de estos transcritos entre semillas tratadas y no tratadas, usando como referencia el genoma completo de *Arabidopsis* (Gómez-Maqueo, 2014). Este tipo de análisis se considera adecuado para especies cuyo genoma no ha sido secuenciado. El estudio de Gómez-Maqueo (2014) mostró que las semillas de *C. aesculifolia* deterioradas por almacenamiento prolongado, estudiadas en el estado seco previo al inicio del programa de germinación, presentaron diferentes respuestas al acondicionamiento mátrico. Los lotes almacenados de semillas que responden favorablemente a este tratamiento presentaron una mejoría en el porcentaje final de germinación, el cual también fue equivalente al porcentaje reportado cuando fueron colectadas. Sin embargo, el desempeño germinativo no es afectado por el tratamiento en todos los lotes analizados de semillas deterioradas, presentándose casos en los que un lote que respondía favorablemente perdió la capacidad de responder al tratamiento. Dado que el mismo estímulo generó diferencias notables en la respuesta de las semillas de *C. aesculifolia* tanto a nivel fisiológico como bioquímico y molecular, es importante determinar si estos cambios también ocurren durante el proceso de hidratación. *C. aesculifolia* ofrece la ventaja de tener semillas grandes de fácil manipulación que permite conocer el contenido de humedad semilla por semilla. Garciadiego (2014) observó en un estudio preliminar que las semillas acondicionadas se hidratan en menor tiempo que las semillas no tratadas y que el CRA se puede utilizar como un indicador del momento fisiológico de la semilla. Estos cambios en el CRA podrían ser independientes del tiempo cronológico de hidratación, permitiendo el contraste de lotes con velocidades distintas de progresión del programa de germinación.

HIPÓTESIS

Dado que el contenido relativo de agua de la semilla aumenta de manera controlada durante la germinación, la selección de muestras de semillas en función del contenido relativo de agua de cada semilla reflejará la progresión del programa de germinación a nivel genético y molecular. Esta aproximación permitirá el análisis de las diferencias existentes entre lotes con diferente desempeño germinativo a través del contraste de estadios fisiológicos equivalentes, a pesar de la variabilidad en el tiempo cronológico del programa de germinación que se puedan derivar del año de colecta o del pre-tratamiento aplicado a las semillas.

OBJETIVO GENERAL

Correlacionar los cambios en el desempeño germinativo generados a partir del acondicionamiento mátrico con cambios a nivel molecular, por medio de análisis de la abundancia de transcritos presentes en las semillas de *Ceiba aesculifolia* en distintos estados de hidratación (contenido relativo de agua, CRA), descriptivos de los estados fisiológicos que marcan la transición entre las fases de la germinación, ruptura de la testa y de endospermo.

OBJETIVOS PARTICULARES

1. Establecer si las diferentes fases de hidratación ocurren en función de valores específicos y reproducibles de CRA.
2. Correlacionar el cambio en la abundancia de transcritos de genes asociados a la modificación de la pared celular con cambios en la actividad de pectin metil esterasas, durante la germinación.
3. Correlacionar el cambio en la abundancia de transcritos de genes asociados al metabolismo de lípidos con cambios en el contenido de éstos durante la germinación.
4. Comprobar los cambios en abundancia de transcritos de los genes relacionados con la movilización de lípidos de reserva y modificación de la pared celular en las semillas de *Ceiba aesculifolia* en respuesta al acondicionamiento mátrico.
5. Determinar si la pérdida de respuesta favorable al acondicionamiento mátrico de las semillas almacenadas de *C aesculifolia* está relacionada con un cambio en la dinámica de hidratación.

MATERIAL Y MÉTODOS

Especie de estudio

Ceiba aesculifolia (Kunth) Britten y Baker F. (Malvaceae), conocido como ceiba o pochote es un árbol monopódico, caducifolio de hasta 15m de alto y hermafrodita. Presenta hojas digitado-compuestas, flores terminales, solitarias o geminadas en las puntas de ramas cortas. Es una especie pionera-secundaria, originaria de las regiones tropicales subhúmedas de América, y su distribución en México abarca desde la costa de Sinaloa hasta Oaxaca y el sur de Chiapas, en la cuenca del Balsas en Puebla, México, Morelos, Guerrero y Michoacán, hasta Yucatán y

Quintana Roo (Pennington y Sarukhán, 1998; Niembro *et al.*, 2010). La especie es resistente a condiciones ambientales como alta radiación solar y suelos poco profundos con poca retención de agua, situaciones comunes en zonas perturbadas por actividad humana (Herrerías-Diego et al., 2005; Velázquez-Rosas et al., 2017). Sus semillas muestran una respuesta al almacenamiento de tipo ortodoxo (Royal Botanic Gardens Kew Seed Information Database, 2020) y son capaces de responder a tratamientos pre-germinativos de acondicionamiento (Alvarado-López, 2012; Garciadiego, 2014; Gómez Maqueo, 2014). Actualmente en México la fibra de las diferentes especies de Ceiba se utiliza localmente como material para el relleno de almohadas o cojines y las semillas son consumidas por la población (Avendaño *et al.*, 2006). Las fibras provenientes de diferentes especies del género Ceiba pueden tener usos importantes en la industria (Zheng *et al.*, 2015). Estas fibras son resistentes y con alta capacidad de flotación por lo que pueden ser utilizadas como aislantes térmicos y sonoros (Dewey, 1943). Particularmente la fibra de *C. pentandra* ha sido muy estudiada; cada fibra es una única célula hueca y lignificada que le confiere las propiedades hidrofóbicas y oleofílicas adecuadas para ser utilizada en diferentes procesos industriales para equipo de flotación y como aislante, y más recientemente, en la absorción de compuestos hidrofóbicos como petróleo y aceites. También se ha propuesto la aplicación de la fibra en producción de bioetanol (Zheng *et al.*, 2015).

Sitio de colecta y manipulación del material vegetal

Se colectaron frutos maduros de 10 a 13 individuos de *C. aesculifolia* por año, en febrero de 2012, 2014, 2015 y diciembre 2015 en la localidad de "Trapiche del Rosario y Chicuasén", dentro del municipio de Actopan, en el estado de Veracruz (México, latitud: 19.5426, longitud: -96.7401, 479 m s.n.m.). Los frutos fueron trasladados a la Ciudad de México, donde se esperó hasta que los frutos secaran y las semillas fueran liberadas. Los frutos colectados en febrero de 2015 se subdividieron en función de la morfología del fruto: frutos compactos y esféricos (lote 2015-1) y frutos ovalados y alargados (lote 2015-2). Ese mismo año, en diciembre, se colectó un segundo lote de frutos, debido a una floración y fructificación temprana. Este lote se consideró, y será referido en el texto, como la colecta de la temporada correspondiente al año 2016. Así mismo, este lote de frutos se subdividió en función del morfotipo de los frutos (esféricos: lote 2016-1; alargados: lote 2016-2). Todos los lotes se almacenaron en el laboratorio a temperatura ambiente (23-27 °C), en sobres de papel dentro de contenedores de cartón corrugado, hasta el momento de realizar los experimentos correspondientes a partir del año 2015, momento en el cual los lotes 2012 y 2014 habían permanecido en almacenamiento cinco y un año respectivamente. Los lotes de 2015 y 2016 se analizaron el mismo año de colecta, después de un periodo de almacenamiento en seco de 30 días post-liberación de las semillas, el cual se requiere para eliminar los remanentes de latencia primaria que presentan las semillas (Gómez-Maqueo, 2014). Finalmente, el lote de 2014 se analizó en una segunda ocasión, al cabo de tres años en almacenamiento y será referido en el texto como "lote 2014-3y". En el cuadro 1 se muestra un listado de los lotes utilizados en esta tesis.

Cuadro 1. Listado de lotes referidos en el texto, año de colecta, características relevantes y año de procesamiento.

Nombre del lote	Año de colecta	Características relevantes	Año de procesamiento
2012	Febrero 2012	Semillas caracterizadas en Garciadiego (2012). Permanecieron cinco años en almacenamiento al momento de realizar los análisis pertinentes a la	2017

		presente tesis doctoral.	
2014	Febrero 2012	Semillas caracterizadas en Garciadiego (2012). Permanecieron un año en almacenamiento al momento de realizar los análisis pertinentes a la presente tesis doctoral.	2015
2014-3y	Febrero 2012	Semillas caracterizadas en Garciadiego (2012). Permanecieron tres años en almacenamiento al momento de realizar los análisis pertinentes a la presente tesis doctoral.	2017
2015-1	Febrero 2015	Semillas colectadas y caracterizadas para los fines de la presente tesis doctoral. Semillas provenientes de frutos compactos y esféricos	2015
2015-2	Febrero 2015	Semillas colectadas y caracterizadas para los fines de la presente tesis doctoral. Semillas provenientes de frutos largos y alargados	2015
2016-1	Diciembre 2015	Semillas colectadas y caracterizadas para los fines de la presente tesis doctoral. Semillas provenientes de frutos compactos y esféricos	2016
2016-2	Diciembre 2015	Semillas colectadas y caracterizadas para los fines de la presente tesis doctoral. Semillas provenientes de frutos largos y alargados	2016
2017-1	Diciembre 2016	Semillas colectadas y caracterizadas para los fines de la presente tesis doctoral. Semillas provenientes de frutos compactos y esféricos	2017
2017-2	Diciembre 2016	Semillas colectadas y caracterizadas para los fines de la presente tesis doctoral. Semillas provenientes de frutos largos y alargados	2017
2018-1	Febrero 2018	Semillas colectadas y caracterizadas para los fines de la presente tesis doctoral. Semillas provenientes de frutos compactos y esféricos	2018
2018-2	Febrero 2018	Semillas colectadas y caracterizadas para los fines de la presente tesis doctoral. Semillas provenientes de frutos largos y alargados	2018

Para generar las muestras utilizadas en todos los ensayos descritos en este trabajo se combinaron las semillas provenientes de cada árbol muestreado por año de colecta, procurando que tuvieran la misma proporción de semillas y asegurar la representatividad de cada planta parental. Todas las semillas usadas se revolvieron en una charola amplia para asegurar que las semillas no se agruparan por tamaño al momento de la selección. Después de mezclarlas se separaron entre tratamientos (control y acondicionadas) y se dosificaron en función de los requerimientos de cada experimento.

Acondicionamiento mátrico

Las semillas de los seis lotes descritos anteriormente se colocaron dentro de bolsas de organza, las cuales se enterraron durante dos días en macetas con suelo estéril, humedecido a capacidad de campo (Gómez Maqueo, 2014). Las macetas se mantuvieron en oscuridad y a 25°C. Posteriormente se exhumaron las semillas y se dejaron secar entre hojas de papel durante dos días más, en oscuridad y a 25°C. Una vez concluido el tratamiento, se congelaron muestras de semillas con y sin acondicionar para los análisis posteriores y se denominaron como tiempo cero "T0" de cada ensayo. Estas muestras se mantuvieron a -70 °C hasta el momento de su utilización.

Prueba de germinación

Se utilizaron 150 semillas por tratamiento repartidas en 5 réplicas de 30 semillas. Se comparó el desempeño germinativo de semillas sometidas a el tratamiento de acondicionamiento mátrico

contra semillas sin tratar (control) y de determinó si los lotes presentaban un fenotipo de respuesta “positiva”, “sin respuesta”, o “negativa” al tratamiento. Las semillas se germinaron en placas de agar al 1%, incubadas en una cámara de germinación LAB-LINE Biotronette (LAB-LINE Instruments Inc.) a 25 °C, con fotoperiodo 12h luz/12h oscuridad. Se registró la germinación acumulada por día, durante 20-25 días, considerando a una semilla como germinada en el momento de protrusión de la radícula a través del endospermo (Gómez Maqueo, 2014). En la figura 1 se muestra una síntesis del proceso general utilizado para manejar los tratamientos usados en cada lote colectado.

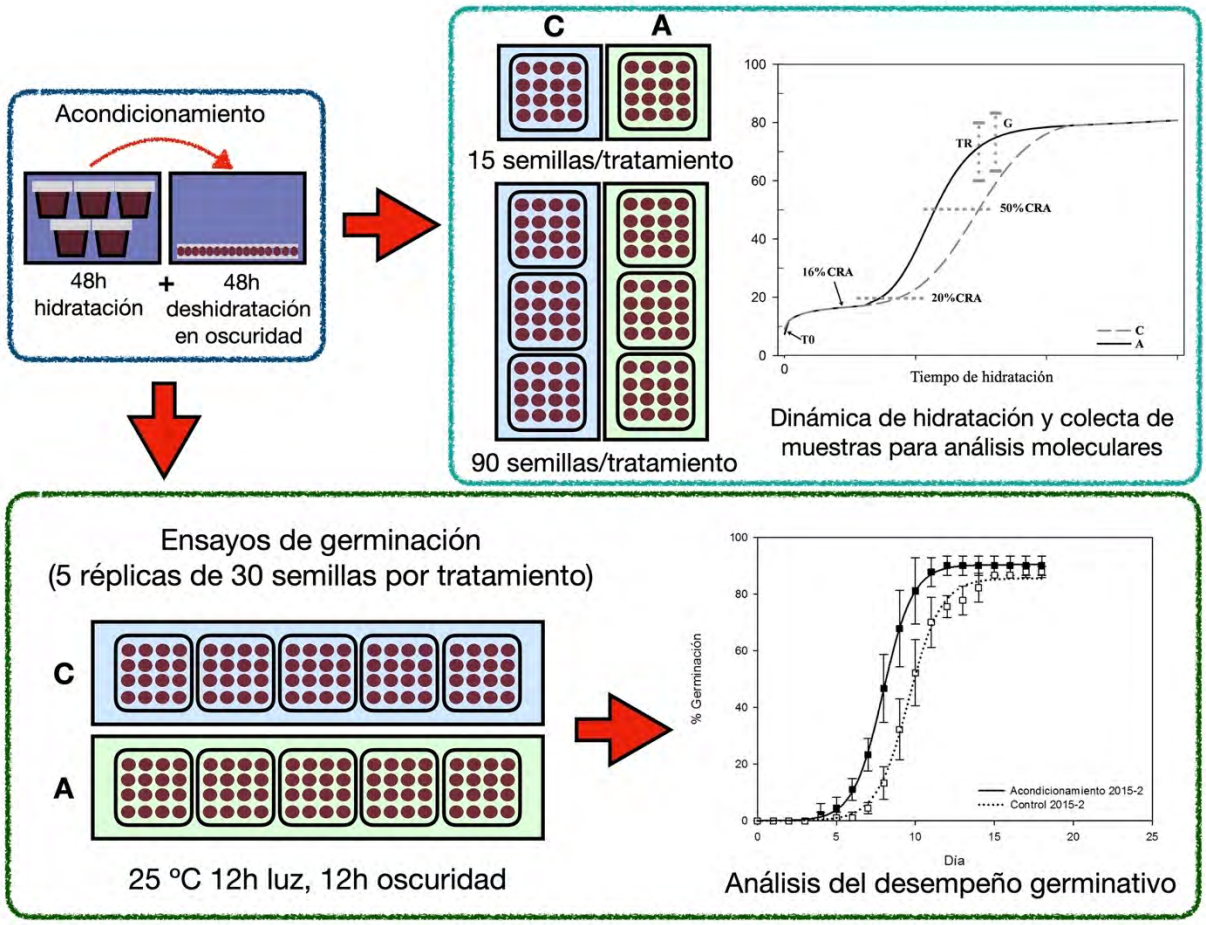


Figura 1. Diagrama del procedimiento general utilizado para el tratamiento de acondicionamiento mátrico (panel superior izquierdo), la colecta de muestras utilizadas en los análisis de la dinámica de hidratación y los análisis moleculares (panel superior derecho) y el desempeño germinativo (panel inferior). Esta estrategia se repitió con cada uno de los lotes mencionados en este trabajo. Abreviaturas: C, control; A: acondicionamiento; CRA, contenido relativo de agua; T0, semilla seca; TR, ruptura de testa; G, germinación. Los estadios fisiológicos mostrados en el gráfico de la dinámica de hidratación corresponden a los puntos de muestreo.

Análisis estadísticos de la germinación

Se hizo una transformación arcoseno del porcentaje de germinación. Posteriormente se ajustó a una curva sigmoide con la función $y=a/(1+b^{(-cx)})$. Para el análisis de germinación se obtuvieron los siguientes valores: inicio (día al que se presenta el 1% de germinación), índice de germinación (primera derivada máxima), tiempo medio (día al cual se alcanza el 50% de germinación) y porcentaje final de germinación (Alvarado López, 2012). Se evaluó si existían

diferencias significativas entre semillas con y sin acondicionamiento mátrico, por medio de una prueba de t de Student. Todos los análisis estadísticos se hicieron usando los programas Table Curve 2D v. 5.01 y SigmaPlot v. 11.0.

Análisis histológico

Se seleccionaron al azar tres semillas de cada lote y tratamiento y se fijaron en una solución FAA (formaldehído, alcohol, ácido acético, agua proporción 2:10:1:7) durante al menos 24h. Las muestras fijadas en FFA se trasladaron a la Unidad de Microscopía de la Unidad de Servicio de Apoyo a la Investigación y a la Industria (USAII) UNAM. Posteriormente se deshidrataron en series de alcohol y se embebieron en paraplast (Para Pro, Paraffin). Se realizaron secciones longitudinales y transversales de 10 µm en microtomo (RMC, MR3). Las secciones se tiñeron con safranina-verde rápido (Márquez *et al*., 2016). Las imágenes se obtuvieron usando un microscopio Olympus BX51. Se seleccionaron dos imágenes representativas de la anatomía de la semilla de *C. aesculifolia*.

Análisis de la dinámica de hidratación de las semillas

Se determinó el contenido de humedad inicial de la semilla seca en cada lote a partir de 15 semillas seleccionadas al azar. Las semillas se pesaron de manera individual y posteriormente se secaron en estufa a 100 °C durante 24 horas. El contenido de humedad se calculó con la siguiente fórmula:

(1) $$CH = ((i-d)/d)\ 100$$

Donde *i* es el peso inicial de la semilla y *d* es el peso después del secado. Estos contenidos de humedad se promediaron y se utilizaron como la línea base para calcular el contenido relativo de agua (CRA) de cada semilla durante la hidratación. Posteriormente, se seleccionaron otras 15 semillas por lote y tratamiento para generar las curvas modelo del proceso de hidratación. Para estas semillas, se registró el peso inicial y se colocaron placas de agar al 1% y se incubaron en una cámara de germinación a 25 °C y fotoperiodo 12:12 h luz:oscuridad. Se realizaron registros periódicos del cambio en el peso de cada semilla durante el tiempo necesario para completar la germinación. Dado que la dinámica de hidratación de las semillas de *C. aesculifolia* presenta varias fases en función de la velocidad a la que se hidratan, las mediciones se realizaron en intervalos de tiempo que variaron entre una y tres horas en las fases de hidratación rápidas, y entre cuatro y 12 horas en las fases de hidratación lenta. El CRA se calculó con la siguiente fórmula:

(2) $$CRA = CH + ((w-i)/i)\ 100$$

Donde *CH* es el promedio del contenido de humedad de la semilla seca de un lote dado, *i* es el peso inicial de una semilla dada y *w* es el peso dinámico de esa semilla durante el proceso de hidratación. También se registró el CRA en el momento en el que se observó la ruptura de testa (TR) y cuando protruyó la radícula (G). Para generar la curvas modelo se utilizó el programa Table Curve 2D para seleccionar la ecuación que se ajustara mejor a la dinámica promedio de cada lote y tratamiento.

Colecta de muestras en función del CRA

Se seleccionaron 90 semillas para cada lote-tratamiento y se dejaron hidratar en las condiciones descritas anteriormente. Se realizaron mediciones periódicas del cambio en el CRA y se colectaron las semillas que presentaron un CRA específico (10%, 16%, 20% y 50%), en función del patrón observado en las curvas modelo de la dinámica de hidratación. Dado que no es posible

determinar el contenido de humedad específico de cada semilla sin destruir la muestra por medio de secado, se utilizó el *CH* promedio ± una desviación estándar y el cambio en el peso de cada semilla como el criterio para determinar si una semilla dada era elegible para la colecta. Asimismo, se colectaron semillas en el momento fisiológico de TR o G y se registró el CRA. Cada semilla colectada se inspeccionó visualmente, para esto se removió manualmente la testa y se observó que no presentaran signos de deterioro o degradación por actividad microbiana. Finalmente, las semillas se congelaron en nitrógeno líquido y se almacenaron a -80 °C hasta el momento de ser procesadas.

Extracción de RNA

Se utilizó el protocolo de extracción de Li y Trick (2005), el cual fue desarrollado para extracción de RNA a partir de semillas con contenido elevado de almidón (Anexo 1). Se cuantificó el rendimiento de cada extracción en un espectrofotómetro NanoDrop y se verificó que el grado de pureza de cada muestra estuviera en el intervalo del 1.98-2.01 de la relación 260/280 nm en el espectro de absorción. Asimismo, se evaluó la integridad del RNA por medio de una electroforesis en gel de agarosa al 1% para observar la presencia de las dos bandas correspondientes al RNA ribosomal.

Secuenciación y ensamble de los transcriptomas de la semilla de *C. aesculifolia*.

Las muestras de RNA se trasladaron a la Unidad Universitaria de Secuenciación Masiva y Bioinformática (UUSMB) del Instituto de Biotecnología UNAM donde fueron procesadas por el personal de dicha unidad. Se construyeron bibliotecas de cDNA para secuenciación por medio de tecnología Illumina HiSeq2000 según las instrucciones del fabricante (Illumina, San Diego, CA, USA). Los análisis de calidad inicial se realizaron con kits Qubit2.0 y con un Bioanalyzer 2100, previo a la construcción de cada biblioteca. Se realizaron dos ensambles *de novo* en dos corridas independientes usando Trinity v2.1.1 (Grabherr *et al.*, 2011; Haas *et al.*, 2013). Para ambas corridas se utilizó una aproximación multiplex para cargar las bibliotedas correspondientes en las laminillas utilizadas. Se obtuvo un total de 1,165,201,460 lecturas pareadas de 100pb entre los dos ensambles. Posteriormente se eliminaron los adaptadores de Illumina usando la herramienta Trimmomatic y se verificó la calidad de las lecturas usando FASTQC. A partir de este análisis de calidad, se decidió recortar las lecturas a 75 pb para asegurar que todas las bases presentaran una calidad mínima de Q20. Los transcriptomas globales presentaron 117,920 y 398,598 secuencias ensambladas, respectivamente. Los resultados de calidad de la secuenciación y ensamble se presentan en los anexos 2 y 3.

Identificación de los transcritos para su clasificación funcional

Se utilizaron los programas Trinotate y transdecoder. En Trinotate se realizó una búsqueda por homología contra bases de datos de proteínas (BLAST+/SwissProt) y dominios de proteínas (HMMER/PFAM). Se identificaron tanto secuencias propias de *C. aesculifolia* como de organismos contaminantes dentro de cada biblioteca. Se usó transdecoder para obtener regiones codificantes de al menos 100 aminoácidos, en seis marcos de lectura. Se realizó un BLASTP para seleccionar el marco de lectura correcto para cada secuencia, utilizando una base de datos local que contenía secuencias de proteínas de tres especies de la familia Malvaceae: *Herrania umbratica*, *Durio zibethinus* y *Theobroma cacao*, así como de la especie modelo *Arabidopsis.* También se realizó un BLASTP contra especies representativas de los contaminantes detectados con Trinoate: *Mus musculus*, los pangenomas de *Pseudomonas aeruginosa* y *Saccharomyces*

cerevisiae, *Trichoderma reesei*, *Fusarium fujikuroi*, *F. verticillioides* y *Diplodia corticola*. Se eliminaron aquellas secuencias que tuvieron mayor similitud a alguno de los organismos contaminantes. Al resto de las secuencias se les asignó el nombre de la mejor secuencia correspondiente a alguna de las especies de malváceas, y también un locus-tag correspondiente al gen con mejor similitud en *Arabidopsis*. Los criterios de asignación comprendieron un porcentaje de similitud del 40% en al menos el 60% de la longitud total del transcrito de *C. aesculifolia*. En total se detectaron 166,491 secuencias (34,274 en el primer ensamble, 132,271 en el segundo ensamble) de *C. aesculifolia*, de las cuales alrededor del 90% presentaron similitud con *Arabidopsis*.

Posteriormente se realizó una validación de la identificación de cada transcrito realizada con Trinnotate y transdecoder, así como de la integridad estructural de cada transcrito. Para esto se corroboró que el marco de lectura seleccionado para cada secuencia estuviera completo, es decir, que contara con los codones de inicio y término, flanqueados por posibles regiones UTR en el extremo 5' y 3'. Se realizó una inspección manual de las secuencias parciales que presentaban un extremo 5' "abierto" (secuencias donde el algoritmo de transdecoder no había logrado distinguir entre el inicio de la región codificante y la región UTR). Para este paso se realizó un BLASTP contra la base de datos no-redundante de proteínas del NCBI y se buscó evidencia de la presencia del condón de inicio en los alineamientos resultantes. Se consideraron como secuencias estructuralmente completas todas aquellas secuencias 5' abiertas para las que se encontró evidencia de codón de inicio. El resto de las secuencias 5', así como secuencias con extremo 3' abierto o secuencias parciales se descartaron de los análisis de expresión diferencial. Finalmente, el conjunto de las secuencias se agrupó en función de los *loci* tags de *Arabidopsis* y se agruparon las lecturas asignadas a cada secuencia para los análisis posteriores. El análisis de la proporción de lecturas por millón, por gen, se muestra en el anexo 4.

Análisis de expresión diferencial y clasificación funcional

Se utilizaron dos estrategias para evaluar el cambio en la expresión durante la germinación. Por un lado, se realizó un análisis exploratorio de los datos, considerando cada estadio fisiológico como unidades discretas, utilizando la plataforma en línea IDEAMEX (Jiménez-Jacinto *et al*., 2019). Esta plataforma permitió implementar de manera simultánea cuatro algoritmos distintos: EdgeR, Limma, DESeq2 y NOISeq. Primero se realizó un análisis de componentes principales (PCA) sobre los datos normalizados. Estos análisis previos permitieron observar los patrones generales de agrupamiento de las muestras en función de los lotes, tratamientos y momentos fisiológicos estudiados. Se eligió DESeq2 como el mejor algoritmo para las comparaciones pareadas por momento fisiológico y entre tratamientos. Se utilizaron los valores de $P<0.05$ y FDR <0.05 (corrección por pruebas repetidas), así como una diferencia de al menos 1 en escala log2 para considerar un gen como diferencialmente expresado. Por otro lado, se utilizó un modelo de series de tiempo en DESeq2 (Love *et al.*, 2014), en conjunto con el paquete MFUZZ (Kumar y Futschik, 2007) para generar un modelo multifactorial de la expresión diferencial y probar el efecto conjunto del fenotipo y del acondicionamiento sobre la dinámica de germinación. Los perfiles trasncriptómicos generados en el paso anterior se agruparon con el software STEM (Ernst *et al.*, 2005; Ernst y Bar-Joseph, 2006). Se utilizaron las herramientas de R heatmap.2, gplots y ggplot2 para generar los mapas de calor y gráficos de barras (R v3.6.0 MacOS; R Core Team, 2016).

La clasificación funcional de los genes detectados en cada tratamiento y estado fisiológico se realizó por medio del análisis de ontologías de genes (GO-terms) en distintas

plataformas, incluyendo DAVID Bioinformatics Resources (v6.8; Huang *et al.*, 2009a, b), las vías genéticas de la base de datos KEGG (Kanehisa y Goto, 2000; Kanehisa, *et al.*, 2019; Kanehisa *et al.*, 2021), y Metascape (Zhou *et al.*, 2019). Las vías genéticas y GO-terms obtenidos se clasificaron en 32 categorías principales: Metabolismo del ADN, Transcripción, Traducción, Modificaciones postraduccionales, Transporte de proteínas, Transporte (por vesículas, Golgi), Metabolismo primario/intermedio, Metabolismo secundario, Proteólisis, Catabolismo de proteínas vía proteosoma, Fotosíntesis/fotorrespiración y cloroplasto, Mitocondrias, Desarrollo/crecimiento celular, Ciclo celular, Ciclo circadiano, Procesos de la pared celular, Procesos relacionados con auxinas, citoquininas y brasinosteroides, Procesos relacionados con etileno, jasmonatos y ácido salicílico, Procesos relacionados con ABA, Procesos relacionados con Giberelinas, Respuesta a Karrikinas, Señalización, Procesos relacionados con la luz, Germinación de la semilla, Estrés biótico, Estrés abiótico, Estrés oxidativo, Envejecimiento, Muerte celular, Transporte de iones, Metilación, Otros procesos.

Disponibilidad de los datos bioinformáticos.

Los archivos de las lecturas pareadas obtenidas para cada biblioteca, sin procesar, se encuentran depositadas en la base de datos SRA del NCBI GenBank con el código Bioproject PRJNA561202, el cual comprende las accesiones SAMN12611291 y SAMN12611291. Las secuencias ensambladas se depositaron en la base de datos TSA del NCBI GenBank con los códigos de accesión contemplados entre el GHVB01000001 y GHVB01166545.

Medición de la actividad de PME en extractos de proteína total

Para cada estadio fisiológico y tratamiento analizado se utilizaron tres muestras independientes.
La extracción de proteína se hizo de acuerdo con la metodología de Downie *et al.* (1998) y modificada para semillas de *C. aesculifolia* por Garciadiego (2014). Se utilizó una solución 50mM de Na_2HPO_4, 1M NaCl y 12.5mM de ácido cítrico, pH6.5. El extracto se filtró para eliminar restos de semilla, lípidos y carbohidratos. Se cuantificó la concentración de proteínas totales por el método de Bradford, a 595nm y usando una curva patrón con albúmina. Una vez conocida la concentración de cada muestra se hicieron las diluciones correspondientes para estandarizar las muestras a 30μg en 20μl.

Se prepararon placas de agarosa al 1% con pectina esterificada de Citrus (®SIGMA, Saez-Aguayo *et al.*, 2013). En cada placa se colocó una muestra del grupo control y una del grupo con acondicionamiento mátrico. Cada réplica biológica se midió por triplicado (réplicas técnicas). Se utilizó PME comercial (®SIGMA) como control positivo. Las placas se incubaron a 29 °C durante 16h y posteriormente se revelaron con una solución 0.05% de rojo de Rutenio. Se midió el área de los halos de actividad de PME y los datos se procesaron con el software ImageJ. La actividad de las PMEs de *C. aesculifolia* se reportan en relación a la actividad de la PME comercial. El análisis estadístico se hizo por medio una prueba no paramétrica de ANOVA mixto de tres vías con pruebas de rangos alineados (Wobbrock *et al.*, 2011; Elkin *et al.*, 2021). Se analizó el efecto del lote, del tratamiento y del estadio fisiológico.

Análisis de contenido de lípidos.

La cuantificación del porcentaje de lípidos totales se hizo por el método de Bligh y Dyer (1959, anexo 5) en las muestras de semillas en diferentes CRAs y estados fisiológicos tanto de los grupos control y con acondicionamiento mátrico. Las cuantificaciones se realizaron en tres réplicas independientes de tres semillas cada una. El análisis estadístico se realizó en cada lote

por medio de un ANOVA de dos vías, contrastando el efecto del tratamiento como primer factor y el estadio fisiológico como segundo factor.

Cuantificación del contenido de ácido abscísico (ABA)

Se utilizó la técnica de radio-inmuno-ensayo descrita por Steinbach *et al*. (1995). Las muestras de semillas colectadas en distintos CRAs se liofilizaron y pulverizaron. Posteriormente se re-suspendieron en agua destilada junto con el anticuerpo monoclonal MAC252. El análisis estadístico se hizo por medio una prueba no paramétrica de ANOVA mixto de tres vías con pruebas de rangos alineados (Wobbrock *et al.*, 2011; Elkin *et al.*, 2021). Se analizó el efecto del fenotipo, del tratamiento y del estadio fisiológico.

CAPÍTULO I. Las Ceibas y sus historias, desde lo sagrado hasta lo genético.

El género *Ceiba* pertenece a la familia Malvaceae y presenta alrededor de 17 especies, de las cuales dos se distribuyen en México. Estas especies se distribuyen de manera natural en América, África y Asia, sin embargo, algunas ceibas han sido introducidas en diversos países debido a su interés ornamental y los usos potenciales de su fibra. En México, las ceibas (*C. pentandra* y *C. aesculifolia*) además forman parte del patrimonio cultural debido a la importancia místico-religiosa de ambas especies para las culturas maya y azteca. Las ceibas eran veneradas como árboles sagrados, capaces de conectar el mundo terrenal con el divino debido a su gran altura y a sus raíces profundas.

En este capítulo se retoma la importancia cultural e industrial de las ceibas y se complementa con información biológica del género. Se espera generar en el lector una perspectiva global sobre la importancia ecológica-social de la especie, así como del estudio de las especies silvestres.

México está lleno de historias increíbles, muchas de ellas relacionadas con la gran diversidad de seres vivos que habitan el territorio mexicano. Basta con salir a dar un paseo por el parque más cercano para encontrarse con unos gigantes silenciosos, los árboles. Aunque los árboles no puedan contarnos las historias de las que han sido testigos, los humanos sí cuentan historias sobre ellos. En la cosmogonía mesoamericana, los árboles representan las vías de comunicación o los vínculos de los seres humanos con los dioses, con los seres sobrenaturales y con los antepasados. Entre las historias más antiguas, contadas por nuestros antepasados mayas y aztecas, se explicaba el origen del cosmos y del mundo, donde destaca en particular la presencia de un grupo de árboles con espinas en sus troncos, conocidos comúnmente en México como ceibas o pochotes. Los mayas relatan en su libro, El Popol Vuh, cómo los dioses habían plantado cuatro ceibas (*Ceiba pentandra*) para indicar los cuatro rumbos principales del cosmos: al Norte, una Ceiba blanca (Sac Imix Che), al Sur una Ceiba amarilla (Kan Imix Che), al Este una ceiba roja (Chac Imix Che), y al Sur un ceiba negra (Ek Imix Che). En el centro, donde los rumbos se unen, colocaron una ceiba verde (Yaax Imix Che, de ésta se desprende el otro nombre con el que conocen a la ceiba en la península de Yucatán, ya'axche). Esta ceiba era la más grande de todas, y tenía una labor muy especial, ya que sus raíces llevaban al Xibalba, el reino de los muertos, en su base se sostenía el Kab, el mundo donde habitamos los seres vivos, mientras que el tronco y las ramas era la morada de los dioses y conectaban con los 13 cielos. En la cima de esta ceiba habita el origen de todos los dioses, en forma de un majestuoso Quetzal. Los aztecas también contaban una historia similar, aunque en su relato eran cuatro árboles diferentes en cada uno de los rumbos. Según el relato plasmado en el Códice Fejérváry-Mayer, es un árbol de Pochote ("Quetzalpóchotl", *C. aesculifolia*), el representante del rumbo del Norte.

Dada la importancia de las ceibas en el origen del cosmos, nuestros antepasados procuraron plantar árboles de ceiba cerca de sus poblados, y actualmente siguen teniendo un lugar protagónico en muchas plazas y jardines. La ceiba es el árbol nacional de Guatemala, donde también goza de un papel importante, ya que es en este árbol donde habita el dios del agua Hunapú. Allí también conocen a la ceiba como "árbol viejo", pues en su base se reunían los ancianos a celebrar consejo; era donde se hacían las leyes del pueblo.

Pero las ceibas no solamente habitan en México y Centroamérica, también se encuentran varias especies de ceiba en Suramérica, donde comúnmente les llaman "Paineira" o "Palo borracho" (*C. speciosa* y *C. chodatii*), debido a que sus troncos tienen forma abombada como una botella. En la figura 2 se muestran algunas representaciones gráficas de las ceibas, en

algunas de las cuales se puede observar esta forma abombada de sus troncos. También se han reportado especies de ceiba nativas de África y Asia, pero en la actualidad se han introducido en distintos países alrededor del mundo como árboles ornamentales.

Se han reportado 17 especies de ceibas. La mayoría son árboles de entre 5 a 20m de altura, pero algunas especies (*C. pentandra*) pueden alcanzar hasta los 50m. Una característica de las ceibas que las distinguen de otras especies de árboles es la presencia espinas en el tronco. Sus hojas son digitadas, con entre 5-7 foliolos y sus flores pueden ser blancas, crema, amarillas o magentas. La mayoría de las especies abren sus flores en la noche y son polinizadas por murciélagos o esfíngidos (mariposas nocturnas); las especies que abren sus flores en el día normalmente son polinizadas por mariposas o colibríes.

Las ceibas pertenecen a la misma familia de la planta de algodón. De hecho, los frutos de ceiba contienen abundantes fibras que ayuda a la diseminación de las semillas en el ecosistema. Cuando los frutos abren y la fibra queda expuesta, esta se seca con el viento y arrastra consigo a las semillas, es como si cada semilla llevara consigo un paracaídas que las hace volar con estilo. Estas fibras tienen varios usos en la industria y comúnmente se le denomina como "kapok". Se usa en la fabricación de salvavidas o como aislantes térmicos gracias a que las fibras son tubulares, atrapando aire dentro de ellas. El kapok es una fibra muy corta y por lo tanto requiere mezclarse con fibras de algodón para poder generar hilos resistentes para la industria textil.

En algunas especies también se han reportado usos medicinales. Los extractos obtenidos de la corteza de *C. speciosa* muestran actividad antioxidante, mientras que los extractos de *C. pentandra* y *C. aesculifolia* tienen actividad antibacterial. Asimismo, en algunas regiones de México, las semillas se consumen tostadas, y los frutos tiernos se usan para preparar salsas y otros acompañamientos para la comida.

Así, las ceibas tienen un papel importante en la vida de los seres humanos, desde la parte religiosa-espiritual hasta lo práctico-utilitario. Sin embargo hay un papel fundamental sin el cual no podríamos disfrutar de los beneficios que nos proporcionan este género de plantas: su papel dentro del ecosistema. Todas las especies del planeta juegan un papel dentro del ambiente en el que viven; pueden ser alimento, refugio, competidores o facilitadores para otras especies. El estudio de las ceibas ha demostrado que hay varias especies que son más resistentes que otras ante condiciones ambientales adversas o que podrían ser muy estresantes para otras plantas. Por ejemplo, *C. aesculifolia* es una especie que puede resistir mejor los periodos prolongados de sequía en comparación con otras plantas, esto gracias a que pueden almacenar agua en sus troncos y en sus raíces. Asimismo, sus semillas se pueden almacenar durante muchos años y pueden germinar en suelos poco profundos y con poca retención agua. Esto ha promovido el estudio de varias especies de ceiba como especies potencialmente útiles para restaurar ecosistemas impactados por la actividad humana.

Sin embargo, a pesar de ser más resistentes a condiciones adversas, las ceibas también se ven amenazadas por el deterioro de sus hábitats y de los efectos del cambio climático. Algunos estudios han observado como las poblaciones "se mueven", alejándose de los sitios donde normalmente habitaban, hacia las zonas altas de los cerros o las montañas donde hace menos calor. Las plantas no pueden moverse como los animales, pero la dispersión de sus semillas es la que favorece que los individuos que germinan y se establezcan en lugares cada vez más elevados, alejándose de las ciudades. Este cambio en la distribución de las ceibas se debe, en parte, a que en las ciudades no hay muchos árboles o áreas verdes, favoreciendo que en las ciudades la temperatura sea hasta 4 °C más elevada que la temperatura de las áreas rurales de los alrededores. El deterioro y la fragmentación de los hábitats naturales no sólo afecta a las ceibas, también

cambia el hábitat de sus polinizadores y, por lo tanto, cambian las posibilidades de poder producir nuevos individuos con características que les permitan adaptarse a un ambiente cambiante.

La variabilidad genética de los individuos es la que permite que las especies se adapten a nuevos ambientes y evolucionen. Esta variabilidad se puede estudiar a distintos niveles para entender por qué y cómo logran sobrevivir las especies en los ambientes en los que habitan. Así como el estilo de vida actual de los seres humanos ha favorecido el deterioro de muchos ambientes, también se ha favorecido el desarrollo de nueva técnicas para entender a los seres vivos. En la actualidad podemos conocer nuestra composición genética a partir de una muestra de saliva. Con esa información podemos conocer más sobre la historia de nuestras familias, saber si tenemos predisposición a padecer ciertas enfermedades, o diseñar terapias específicas para atender a un paciente en particular. En el caso de las plantas, esta información genética se puede obtener a partir de una hoja, un tallo, o incluso de una semilla. Con este tipo de información genética podemos entender cómo logran sobrevivir y persistir en su entorno a pesar de que no se puedan mover o escapar del lugar donde están creciendo. Normalmente las técnicas más sofisticadas, como es el caso de las técnicas moleculares, se desarrollan para entender y aprovechar de manera más eficiente los recursos que, en nuestra forma de pensar, son más importantes. Actualmente conocemos la composición genética de muchas especies de animales y de plantas domésticas que consumimos a diario; pero desconocemos a muchas otras especies, como a las ceibas, que podrían ayudarnos a hacer frente a los desafíos del cambio climático y frenar el deterioro de los ecosistemas. En años recientes ha habido esfuerzos para revertir esta falta de información genética de las especies silvestres del país. Para el caso de las ceibas, actualmente se cuenta con información genética contenida en las semillas de *C. aesculifolia*. Esta información se está utilizando para entender cómo las semillas son capaces de percibir el ambiente y germinar en las condiciones más favorables para asegurar el establecimiento de nuevas plantas. En el largo plazo, esta información también puede ayudar a generar estrategias más eficientes de propagación de *C. aesculifolia* (y de otras ceibas), así como favorecer su uso para restaurar sitios perturbados.

Los estudios multidisciplinarios de las especies silvestres también son una lección importante para el ser humano. La vida se ha abierto paso durante millones de años utilizando distintas estrategias, muchas de las cuales apenas estamos empezando a entender. Si queremos hacer frente a los desafíos del mañana, es fundamental que reconozcamos la importancia de cada especie que habita este planeta y con las cuales compartimos nuestra existencia. Así como en nuestro país nuestros antepasados han contado historias sobre las ceibas durante cientos de años, es nuestro turno de aportar a esas historias y procurar que se sigan contando durante muchos siglos más.

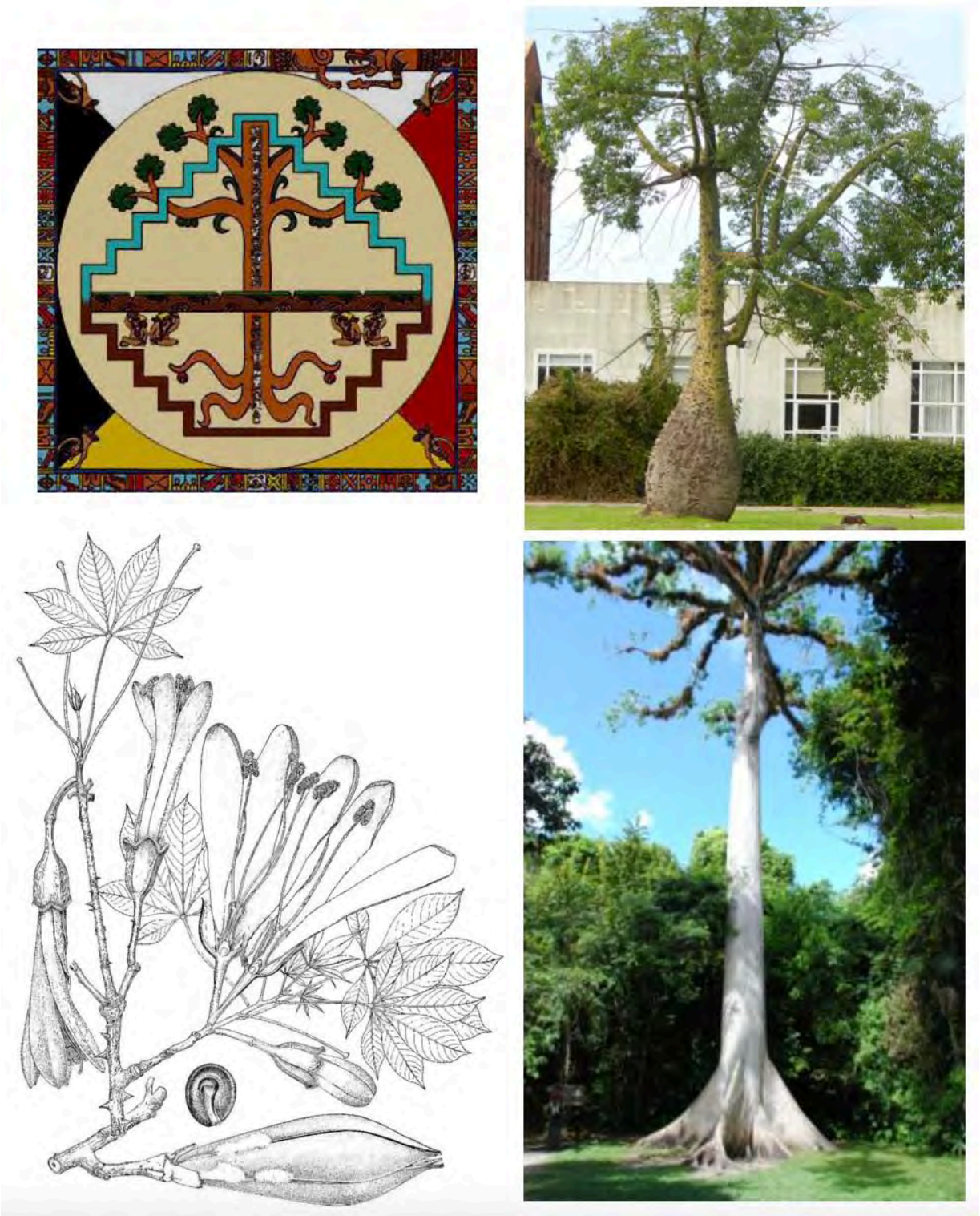


Figura 2. Representaciones destacadas de las especies del género *Ceiba*. Esquina superior izquierda: representación de *C. pentandra* como árbol sagrado de los Mayas. Esquina superior derecha: ejemplar de *C. speciosa* o Palo borracho, se observa el tronco abombado en la base o con forma de botella. Esquina inferior izquierda: Ilustración científica de *C. aesculifolia*, se observa el detalle de los tallos con espinas, las hojas, flores, fruto y semilla. Esquina inferior derecha: ejemplar de *C. pentandra* en su hábitat natural, se observa la altura superior a los 40m y la presencia de contrafuertes en la base del tallo.

CAPÍTULO II. Regulación transcriptómica del proceso de germinación, el contenido relativo de agua (CRA) como una aproximación a fases del desarrollo discretas durante la germinación.

Las semillas constituyen una etapa fisiológica clave en el ciclo de vida de las plantas. Durante la germinación existe un patrón espacio-temporal del proceso de hidratación de los tejidos, que correlaciona con diversos procesos fisiológicos que permitirán la protrusión de la radícula y el desarrollo post-germinativo. El estudio de la progresión de la germinación usualmente se realiza en función del tiempo de hidratación. Sin embargo, dentro de una población de semillas se pueden observar variaciones amplias en el tiempo de hidratación de cada semilla. Esta situación es particularmente evidente en especies silvestres. Desde una perspectiva ecológica, la variación en el tiempo de hidratación constituye una estrategia para distribuir el establecimiento de los nuevos individuos en el tiempo, previniendo la competencia intraespecífica o la germinación de toda la cohorte en un episodio corto de condiciones favorables. En un estudio realizado en *Arabidopsis*, se demostró que el momento de ruptura de testa, es una etapa crítica asociada con un perfil transcriptómico distinto al de semillas de la misma edad cronológica sin ruptura de testa. Este estudio en *Arabidopsis* sugiere que es posible establecer estadios fisiológicos discretos durante la germinación, sin la necesidad de recurrir al tiempo de imbibición como indicador de la progresión del proceso. En este capítulo se explora en *C. aesculifolia* la posibilidad de utilizar el contenido relativo de agua (CRA) como una aproximación a un estadio de desarrollo discreto, en lugar del tiempo de hidratación. Se realizó un seguimiento de la hidratación semilla por semilla durante la germinación para homogeneizar el muestreo en lotes colectados en años distintos y con tiempos de hidratación variables. Se analizó el transcriptoma asociado a cada CRA elegido para detectar las transiciones fisiológicas que ocurren en estos lotes de semillas que, además, presentaron respuestas fenotípicas contrastantes a un pre-tratamiento de acondicionamiento mátrico. El CRA reflejó las transiciones transcripcionales que ocurren durante la germinación, independientemente del tiempo de hidratación o del año de recolección. Asimismo se observó un conjunto de procesos biológicos que ocurren en la semilla seca y durante la germinación temprana que están asociados con la respuesta fenotípica al acondicionamiento. En este capítulo, demostramos que el uso de rasgos fisiológicos, específicos para una etapa de desarrollo particular, es un enfoque confiable independiente del tiempo de hidratación para *C. aesculifolia* a pesar de la variabilidad en la velocidad de hidratación observable entre lotes. Las figuras suplementarias que se referencian en esta publicación se incluyen en el anexo 6.

SCIENTIFIC REPORTS
nature research

ncias.unam.mx; agamboa@
ecologia.unam.mx

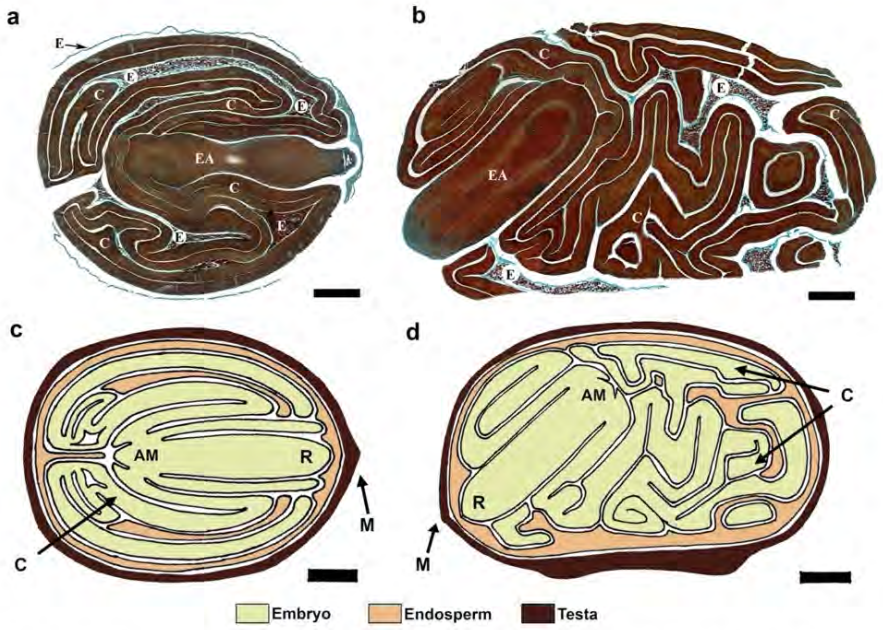


Figure 1. Seed anatomy. In (**a**) transversal, and (**b**) longitudinal sections (10 µm thickness) showing the embryo and remnants of endospermic tissue surrounding the embryo and allocated within the folds of the cotyledons. The sections were stained with Safranin-green fast[37]. EA, embryonary axis; C, cotyledons; E, endosperm. Section scale denoted by the black bars (1 mm). In (**c**) transversal, and (**d**) longitudinal diagrams depicting the full structure of the seed, with the embryo (light green) surrounded by the endosperm (orange), and the testa (brown). AM: apical meristem; C: cotyledons; M: micropyle area; R: radicle tip.

for sample collection instead of time-specific collection points provides with more homogeneous biological replicates[13,17].

Seed-water relations are fundamental in ecosystems with marked seasonality, where seeds are exposed to hydration-dehydration cycles due to stochastic rains, before the actual rainy season sets in, termed as natural priming[18,19]. Like natural priming, the diverse priming treatments performed under controlled conditions can improve germination and optimize field management of many crop species[20]. Here, we characterize the germination process and response to priming across different seed batches collected in different years, from the same population of the wild tree *Ceiba aesculifolia* (*Malvaceae*). This species is resistant to the environmental conditions common in disturbed ecosystems by human activity, such as higher exposure to sunlight, and poor, shallow soils with low water retention[21,22], and produces "kapok" fibres with potential industrial uses[23]. It has relatively large, orthodox seeds, allowing for individual seed tracking of the relative water content (RWC) during imbibition. Could a specific RWC value reflect a physiological stage useful for comparisons between seed batches and phenotypic responses to priming? We selected a transcriptomic analysis approach to answer this question.

Results and discussion

Characteristics of *Ceiba aesculifolia* seeds and germination performance. The *C. aesculifolia* seeds present an embryo with large and tightly folded cotyledons, which comprises the majority of the seeds volume (Fig. 1). The testa was removed for microtome sections in Fig. 1a,b. A schematic representation of the whole seed is presented in Fig. 1c,d. The mature seed presents a thin layer of endosperm surrounding the whole embryo that thickens near the micropyle area, where the radicle tip is located. However, the endosperm is not entirely consumed and remnants of endospermic tissue can be found enclosed within the folds of the cotyledons.

Ceiba aesculifolia thrives in dry tropical forests that have a stationary rainy pattern, and seeds are released at least four months before the rainy season (Supplementary Fig. S1). However, during the dry period, occasional rains can occur and the seeds are subjected to hydration/dehydration cycles that can be considered as natural priming since no germination occurs at that moment[18,19,24]. In order to study the response to these cycles on transcriptional processes, we determined two phenotypes related to the priming response. The first phenotype corresponds to seed batches that have a positive response to priming (PR, determined as an improvement of the germination parameters tested). The non-responsive/negative phenotype (NR) was comprised by the batches that had no response to priming (2012-5 y and 2014-3 y) and the 2016 batch, which had a negative response to priming by displaying a significant reduction in the germination rate and a slight decrease in the final germination percentage (Supplementary Table S1). Three independent batches were considered for each category. The priming

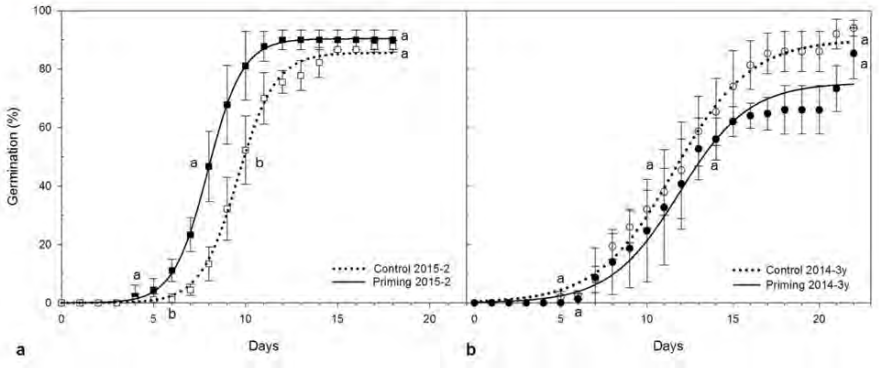


Figure 2. Seed germination performance of two seed batches that represent the studied phenotypes. Doted lines indicate the germination dynamic of the control (untreated) seeds, while bold lines denote the primed seeds. In (**a**) seeds collected in 2015, (**b**) seeds collected in 2014 and stored for 3 years. The letters indicate differences found between primed and control seeds for germination initiation (letters near the x axis), time to 50% (letters at the middle of the sigmoid curves), and final germination (letters near the top end of the curve).

test was performed as a means to classify the seed batches, thus the transcriptomic analyses and results presented correspond only to untreated (control) seeds.

The PR phenotype was represented by one seed batch from 2014 (named 2014) and two 2015 batches from round fruits (2015-1) and elongated fruits (2015-2) from different trees (Fig. 2 and Supplementary Fig. S2). In the untreated seeds, radicle protrusion started between day 5 and 7 after sowing, and took up to 20 days for the seed batch to attain its maximum germination percentage, while primed seeds started at day four in all seed batches. In all cases, un-germinated seeds (about 8 to 15% of seeds in each replicate) died within a 25-day timeframe due to microorganism infection. The positive response to the priming treatment was reflected in either a reduction of the time needed for germination initiation (at day four after sowing) or the time to attain 50% germination, which varied among seed batches. The primed seeds from the 2014 seed batch had a higher maximum germination rate compared to its respective control, while the 2015-1 seed batch was the only batch that showed a significantly slower rate in comparison to its respective control (2014: control 8.95%·day^{-1}, primed 10.65%·day^{-1} $P=0.02$; 2015-1: control 12.4%·day^{-1}, primed: 7.96%·day^{-1} $P=0.03$; 2015-2: control 13.08%·day^{-1}, primed 14.69%·day^{-1}, $P=0.31$). There was no significant effect on the final germination percentage except for the 2014 seed batch (2014: control 86.45%, primed 92.33%, $P=0.03$; 2015-1: control 87.78%, primed 85.56%, $P=0.34$; 2015-2: control 87.78%, primed 90%, $P=0.37$). Full measurements are presented in Supplementary Table S1.

The NR phenotype was represented by two batches that were stored for 5 years (2012-5 y) and 3 years (2014-3 y), and one batch recently collected from trees with elongated fruits (2016). Figure 2 and Supplementary Fig. S2 show that there is no effect of priming on the germination parameters for these batches, except for germination rate in 2016 primed seeds, which was significantly lower than its control (Supplementary Table S1). Final germination percentage of 2012-5 y (71.43%) and 2014-3 y (94.0%) batches were similar to the germination performance of those seed batches when freshly collected and had a positive response to priming (80% and 86.45%, respectively). These results suggest that, although these batches had lost the positive response priming, storage has not had a deleterious effect on germination capacity. The 2016 seed batch displays the lowest germination capacity at 53.33%, but is not the result of storage-mediated deterioration or from the induction of secondary dormancy, but rather from a higher mortality during imbibition. This higher mortality could be the result of stress experienced by the mother plants during seed production and maturation. The 2016 batch is different from all other batches (both NR and PR) due to a change in the phenology in that year, the un-timely flowering season in turn caused an early release of the mature seeds on December 2015, two months before its regular season.

The imbibition dynamic of *C. aesculifolia* seeds is not affected neither by collection year nor response to priming

Seed imbibition dynamic follows a specific pattern that correlates to the seed RWC. There is a species-specific basal seed water content that functions as a threshold that needs to be exceeded in order for radicle protrusion to occur. This was observed in some domesticated plant species, like cotton, sorghum, chickpea, and maize[25]. Consequently, we asked whether there are other critical thresholds along the imbibition curve. We tracked seed imbibition during the germination process to test this hypothesis.

The variation of seed RWC over a specific time of measurement on both PR- and NR-batches is shown Fig. 3a and Supplementary Figs. S3 and S4. A randomized sampling as a function of time could lead to a mixture of different imbibition or physiological stages. However, by tracking seed-by-seed changes in fresh weight, a similar pattern of imbibition can be observed, which is summarized in Fig. 3b. The seeds have an initial RWC of 7–8% (T0, or the dry state), and imbibe rapidly until they almost double their RWC. At this point, the rate of water uptake slows down; the duration of this phase appreciably varied between seed batches, and treatments. The next RWC value at which the imbibition rate changed again coincided with a 20%RWC. The priming treatment

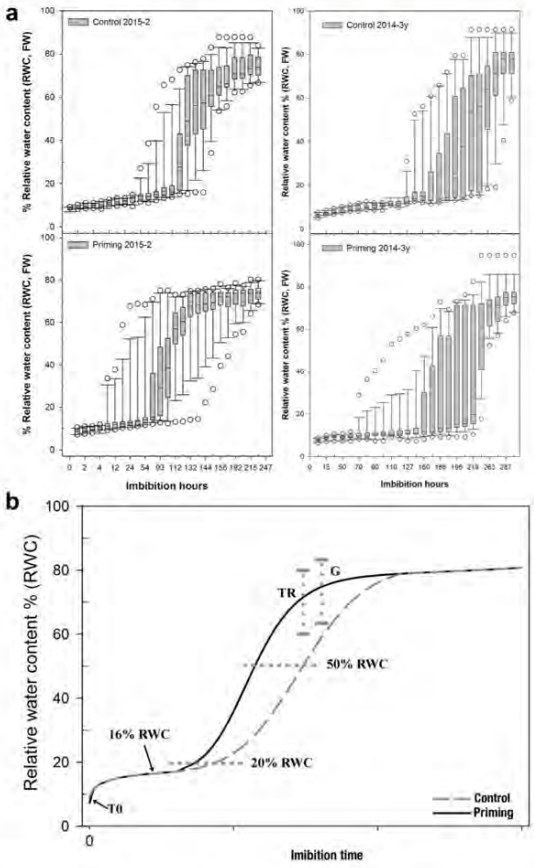
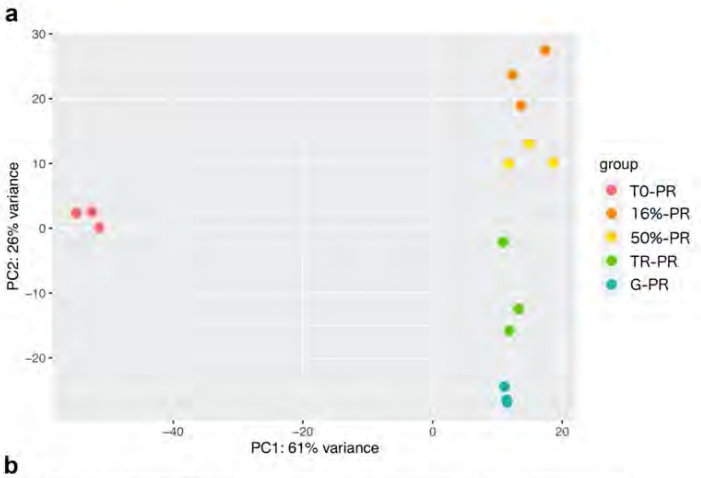


Figure 3. Seed imbibition dynamic. In (**a**) boxplot graphs depicting the variation of seed RWC over a specific time of measurement of the seed batches of 2015-2 (left) and 2014-3 y (right): the RWC dispersion by time of measurement in the control seeds (upper) and the RWC dispersion by time of measurement in primed seeds (lower). Note: The X axis is represented as a category rather than a continuous scale to allow the even distribution of all boxes. In (**b**) a general imbibition dynamic model proposed for control and primed *C. aesculifolia* seeds. Testa rupture (TR) can be most frequently observed when the seed has attained a RWC between 72–85%. Radicle protrusion (G) can occur almost simultaneously to TR, so its RWC interval (75–93%) overlaps with TR.

decreased the average imbibition time needed to reach the 20%RWC. The next imbibition phase showed the major change in seed fresh weight. Although there was not a specific RWC value at which all seeds presented testa rupture (TR) or radicle protrusion (G), the RWC interval at which both processes occurred was similar in all seed batches.

In order to establish a physiological trait that allow the comparison between seeds from different batches or treatments for molecular studies, we determined to collect the seeds of each category and batches before TR based on the RWC value for the transcriptional analyses. For PR-batches, we selected 16%RWC as the representative value for the slow imbibition phase and 50%RWC for the rapid imbibition phase. The 50%RWC was selected as the representative value for the early germination phase in the NR batches.

To explore the mechanism underlying the differences in priming responses of the two categories of *C. aesculifolia* seed batches, transcriptome analyses of each phenotype during germination were performed using Illumina clean reads for *de novo* transcriptome assembly. Considering that the germination process involves an important removal of stored transcripts related to seed development, we used only the transcripts with a complete coding sequence (see methods) for further analyses. A total of 54,793 complete transcripts was obtained, which then were identified using a local data base that included three annotated species from the *Malvaceae* family and *Arabidopsis*. We were able to associate a specific *Arabidopsis* locus-tag to 91% of the transcripts, which were used for GO-term enrichment of biological functions.



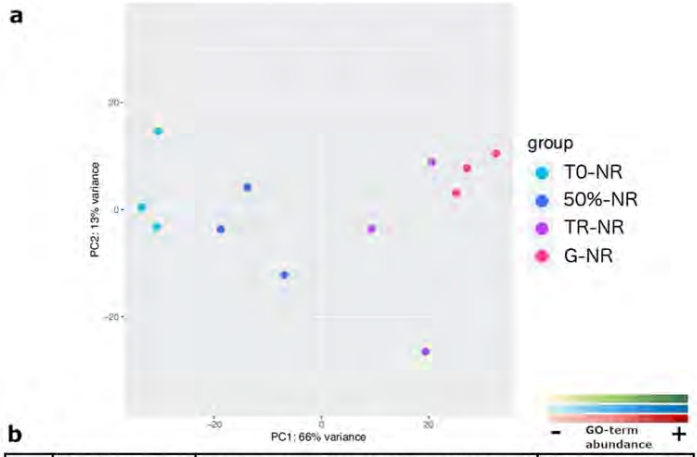
b

Present in T0	Up during germination: 16% RWC	Up during germination: 50% RWC	Up during germination: TR	Up during germination: G	Biological processes with enriched GO term categories that showed differential expression with respect to the previous physiological stage	Down during germination: 16% RWC	Down during germination: 50% RWC	Down during germination: TR	Down during germination: G
					DNA metabolism				
					Transcription				
					Translation				
					Post-traslational modifications				
					Protein trasport				
					Transport (vessicle-mediated, Golgi)				
					Primary/intermediary metabolism				
					Secondary metabolism				
					Proteolysis				
					Protein catabolism via proteasome				
					Photosynthesis/Photorrespiration and Chloroplast				
					Mitochondria				
					Development, cell growth				
					Cell cycle				
					Circadian cycle/rythm				
					Cell wall-related				
					Auxin, Cytokinin and Brassinosteroid-related processes				
					Ethylene, Jasmonates and Salicylic acid-related processes				
					ABA-related processes				
					Giberellin-related processes				
					Response to Karrikin				
					Signaling				
					Light-related processes				
					Seed/germination				
					Stress (biotic)				
					Stress (abiotic)				
					Oxydative stress				
					Aging				
					Cell Death				
					Ion transport				
					Methylation				
					Other processes				

Figure 4. Transcriptome analyses of seed batches with a positive response to priming. In (**a**) PCA analysis of the transcriptome profiles of three replicates for each physiological stage indicated by colour. PC1 explains 61% of the variation between the dry stage and the imbibed stages. PC2 explains 26% of the variation between imbibed stages. In (**b**) Main GO-term categories enriched in each physiological stage. The green colour represents the enriched categories present in the dry seed while the red and blue colours represent the enriched categories that had differential expression with respect to the previous stage. Up-regulated categories are shown in red and down-regulated categories in blue.

The RWC is a physiological trait that reflects the transcriptional transitions that occur during germination

The *C. aesculifolia* seed batches with positive response to priming present the two main transcriptional phases, as described for Arabidopsis: early and late germination. A principal component analysis (PCA) was performed to evaluate the RWC as a physiological trait during germination (Figs. 4a and 5a). For the PR phenotype, PC1 could explain 61% of the variance between dry and imbibed seeds, and PC2 could explain 26% of the variance between the imbibed stages (16%RWC, 50%RWC, TR and G; Fig. 4a). We obtained differentially expressed genes that show significant changes between different stages and in the overall course of germination, and between physiological stages. The overall expression changes from the dry stage (T0) until radicle protrusion (G) resulted in an up-regulation of 6,635 genes and a down-regulation of 2,457 genes (Supplementary Table S2). The most important changes were detected between T0 and 16% (2,727 up-regulated and 2,375 down-regulated genes); slight changes in gene expression were detected between 16 and 50% (69



Present in T0	Up during germination			Biological processes with enriched GO term categories that showed differential expression with respect to the previous physiological stage	Down during germination		
	50% RWC	TR	G		50% RWC	TR	G
				DNA metabolism			
				Transcription			
				Translation			
				Post-traslational modifications			
				Protein trasport			
				Transport (vessicle-mediated, Golgi)			
				Primary/intermediary metabolism			
				Secondary metabolism			
				Proteolysis			
				Protein catabolism via proteasome			
				Photosynthesis/Photorrespiration and Chloroplast			
				Mitochondria			
				Development, cell growth			
				Cell cycle			
				Circadian cycle/rythm			
				Cell wall-related			
				Auxin, Cytokinin and Brassinosteroid-related processes			
				Ethylene, Jasmonates and Salicylic acid-related processes			
				ABA-related processes			
				Giberellin-related processes			
				Response to Karrikin			
				Signaling			
				Light-related processes			
				Seed/germination			
				Stress (biotic)			
				Stress (abiotic)			
				Oxydative stress			
				Aging			
				Cell Death			
				Ion transport			
				Methylation			
				Other processes			

Figure 5. Transcriptome analyses of seed batches with negative/no response to priming. In (**a**) PCA analysis of the transcriptome profiles of three replicates for each physiological stage indicated by colour. PC1 explains 66% of the variation between the dry stage and the imbibed stages. PC2 explains 13% of the variation between replicates. In (**b**) Main GO-term categories enriched in each physiological stage. The green colour represents the enriched categories present in the dry seed while the red and blue colours represent the enriched categories that had differential expression with respect to the previous stage. Up-regulated categories are shown in red and down-regulated categories in blue.

up-regulated and 26 down-regulated), but a considerable change between 50% and TR (645 up-regulated and 453 down-regulated) was also detected. A small amount of differentially expressed genes was detected between TR and G (193 up-regulated and 148 down-regulated, Supplementary Table S2). These results suggest that *C. aesculifolia* seed germination presents the early and late transcriptional phases as described for *Arabidopsis*[13]. The GO-term enrichment analyses showed that the biological groups enriched in 16%RWC with respect to T0 were those related to DNA metabolism, transcription and translation, and primary/intermediary metabolism. The response to abscisic acid (ABA), ethylene and cytokinin were also up-regulated. Abiotic stress responses, that included cadmium, cold, heat and salt stress, were upregulated. The response to ABA and abiotic stress were

down-regulated. In contrast, only proteolysis related to mobilization of reserve proteins, transmembrane transport and response to water deprivation were enriched at 50%RWC and the down-regulated biological process were related to seed maturation and embryo development (Fig. 4b, Supplementary Table S3).

The enriched biological processes that represented the TR were primary/intermediary and secondary metabolism, development and cell growth, and cell wall-related processes. The response to cytokinin and auxin were also regulated as well as the response to abiotic and oxidative stresses. The main down regulated category was related to ABA. During G, the main up-regulated categories also included cell wall related processes. The primary metabolism was represented by the up-regulation of lipid metabolism and hormone response included auxin homeostasis and gibberellin response. The carbon fixation and cell cycle processes were strongly up-regulated. Enriched down-regulated biological processes were transcription, lipid storage, sucrose biosynthesis and response to ethylene and auxin.

In the seed batches with negative or no response to priming, the dry seed transcriptome is similar to the imbibed seeds transcriptome in the early germination stages of both phenotypes. For NR phenotype, PC1 could explain 66% of the variation between the different stages (T0, 50%, TR and G) and PC2 could explain 13% between replicates (Fig. 5a). In these seed batches, expression analyses showed moderate changes between stages and in the overall course of germination, in comparison to the observed changes in PR-batches (Supplementary Fig. S5). The overall expression changes from the dry stage (T0) until radicle protrusion (G) showed an up-regulation of 2,437 genes and a down- regulation of 1,856 genes (Supplementary Table S2). In the transition from T0 to 50%RWC a total of 559 up-regulated and 247 down-regulated genes were observed, but a similar amount of expressed genes as the PR-batches between 50%RWC and TR were also detected (762 up-regulated and 470 down-regulated genes). Finally, a small amount of differentially expressed genes was detected between TR and G (21 up-regulated and 56 down-regulated, Supplementary Table S2). The biological groups that were mainly enriched in 50%RWC were those related to primary/intermediary and secondary metabolism (Fig. 5b, Supplementary Table S3). The synthesis of jasmonic acid, and biotic and abiotic stress responses were also upregulated. In contrast to PR-batches, there was not a significant up-regulation of the proteolysis category. Down-regulated biological groups were related to ABA response and embryo development.

The enriched biological processes that represented the TR were primary/intermediary and secondary metabolism, photosynthesis related processes including carbon fixation, development and cell growth and cell wall related processes. The response to light, auxin and gibberellic acid were also up-regulated as well as the response to abiotic and oxidative stresses. The main down-regulated categories were related to ABA and abiotic stress. There were also a down-regulation of biological processes related to post-translational modification and primary metabolism. During G, no significantly enriched up-regulated categories were detected. The enriched down-regulated biological process was related to lipid storage.

The expression and the GO-term enrichment analyses indicated that the germination process in PR-batches occurred accordingly to what has been described in previous studies[11,13,26] (Fig. 4b), suggesting that for *C. aesculifolia* seeds, the transcriptomic profile could be the reference for the proper transitions that arise during germination. However, in the NR-batches, the differential expression and the GO-term enrichment profiles did not reflect, at the same extent, the transitions between physiological stages. We performed time-series analyses and detected eight clusters of gene expression patterns in which the main trends in the PR batches reflected an up- or down-regulation from T0 onwards, while in the NR-batches those same clusters showed subtle or no changes from T0 onwards (Supplementary Fig. S6). These major trends from T0 to 50%RWC suggested that the transcripts levels already present in the dry seeds of the NR-batches were more similar to the observed transcript levels in imbibed seeds. To verify this observation and to test if the RWC could distinguish between physiological stages, despite the underlying differences in the phenotypes, we performed a global PCA and a heatmap with hierarchical clustering analyses (Fig. 6).

For PR and NR samples, PC1 could explain 43% of the variation between the dry and imbibed seeds except for T0-NR that was grouped with the imbibed seeds, confirming the observation from the clustered expression patterns that the NR-T0 profile is more similar to that of low RWC samples. Meanwhile, PC2 could explain 35% between the different imbibed stages and T0-NR seeds (Fig. 6a). This analysis also demonstrated that the imbibed samples of PR- and NR-batches grouped by their respective physiological stage during the germination process, including 50%RWC, TR, and G. The hierarchical clustering performed with the global transcriptomic data also confirmed that the two transcriptional phases that occur during germination are distinguishable independently of the phenotype as well as the resemblance of the NR-T0 seeds with the low RWCs seeds, by its grouping with the PR-16%RWC samples (Fig. 6b).

In *Arabidopsis*, the dry seed contains about 12,600 different mRNAs stored during development and maturation[27]. In the combined transcriptomes of both PR- and NR-dry seeds there was an overlap of about 63% (8,064) of the locus tags reported for *Arabidopsis*[27], and about 37% (3,017) of those were shared between the two phenotypes and *Arabidopsis*. An incomplete overlap is expected due to the large differences in life histories between the two species. Like *Arabidopsis*, *C. aesculifolia* produces endospermic orthodox seeds and presents a two-stage germination, but it is a perennial tree that thrives in warm climates where water availability is the limiting factor for germination to occur. A GO-term analysis of the non-overlapping *Arabidopsis* locus tags confirmed that almost all those genes belonged to the same gene families within the biological categories presented in Figs. 4 and 5 for *C. aesculifolia*, especially stress responses, metabolism, and protein transport/modification, thus the differences between the two species is associated with their particular physiology and responses to the environment.

The results confirm that it is possible to compare seed batches of a particular species despite their differences in life histories. Using the transcriptomic profiles of PR-batches during germination as reference, we determined that the NR-dry seeds do present a transcriptomic profile corresponding to the transcriptional phase I of the germination process[13]. However, they also present some other enriched functional categories that reflect a disarray of

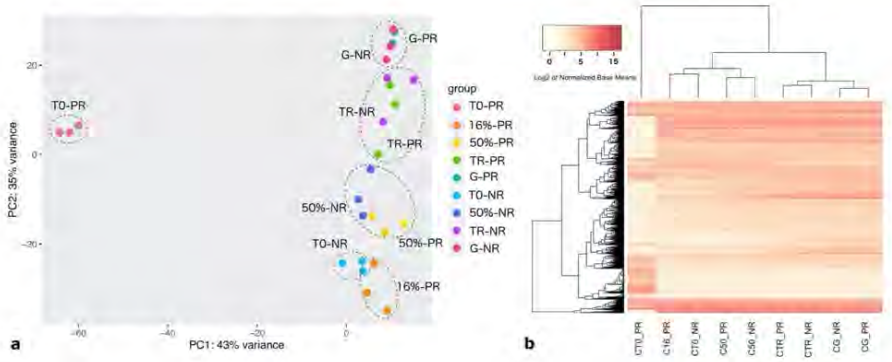


Figure 6. Global analysis of the transcriptome profiles of the two phenotypes during germination. In (**a**) PCA analysis of the transcriptome profiles of three replicates for each physiological stage indicated by colour. PCA was performed with the top-700 genes that accumulated the most variance across groups. PC1 indicates that the transcriptome profile of the dry stage (T0) in the PR seeds had the most important differences with respect to the other physiological stages, including the dry stage in the NR seeds (43% of variance). PC2 explains 35% of the variance between the remaining physiological stages for both phenotypes. In (**b**) heatmap of the normalized reads per transcript of all 12,683 *Arabidopsis* locus tags present in the global transcriptome assembly, and detected in at least one physiological stage. The hierarchical clustering distinguishes the two transcriptional phases[13] indicated by low RWCs (16% and 50%) and by TR and G physiological stages. The NR-T0 (dry seed) profile is clustered among the low RWCs, been most similar to 16% transcriptomic profile.

the early transcriptional stage, like an overrepresentation of stress-related processes. The premature expression of these processes could reflect a costly and inefficient use of the seed resources, precluding the ability of NR-batches to properly respond to the priming treatment or other external stimuli.

The NR-batches are capable of adjusting their transcriptomic profiles during the germination process, despite their initial differences with respect to the PR-batches. The overlap of the six samples at the moment of radicle protrusion (G), and the clustering of samples by physiological stages suggested that the transcriptomic profiles of PR and NR-batches gradually become more similar during germination (Fig. 6). This is also observable in the clusters in Supplementary Fig. S6, since none of the eight clusters showed a trend in which all the genes from that particular cluster had a reversed expression trend between PR- and NR-batches (*i.e.* that the PR-batches had a trend of up-regulation at G while the NR-batches displayed a down-regulation trend at G). This trend of PR- and NR-batches to converge in transcriptomic profiles by the end of the germination process was observed even in the simulations run with a higher number of clusters or cluster membership stringency. The differential expression analyses between PR- and NR-batches for each physiological stage confirmed a reduction in the total number of differentially expressed genes during late germination (Fig. 7).

For T0 79% of the present transcripts showed differential expression, while only a 4%, 0.03% and 0.2% of transcripts present in 50%RWC, TR and G respectively were affected. In the NR-T0 seeds it was detected a GO-term enrichment in 24 of the 32 categories of biological processes in Figs. 4 and 5, of which primary/intermediary metabolism, followed by abiotic stress, translation and, DNA metabolism were the categories which presented the most abundance of terms and up-regulated genes (Fig. 7b; Supplementary Table S3). In the case of the down-regulated genes, the main categories were again the primary/intermediary metabolism as well as abiotic stress, which presented the most abundance of GO-terms and genes; but also vesicle transport, protein transport and transcription showed an important enrichment. At the 50%RWC stage, primary/intermediary metabolism, biotic and abiotic stresses were the categories that presented GO-term enrichment, while transcription was the main down-regulated category, followed by seed/germination-related processes, translation, post-translational modifications, circadian rhythm, and ethylene/jasmonic acid-related processes. For TR and G stages, no GO-term enrichment was detected. This result suggests that, although the transcriptional profile of NR-T0 is more similar to the early germination profile, these seeds are capable of adjusting their transcript levels during imbibition in order to maintain the germination program.

Finally, transcript profiles and differential expression between PR- and NR-batches and during the germination process was verified by RT-PCR in five randomly selected genes from the data set (Fig. 8). The data was normalized using ACT7 (*At5g09810*) as reference[28]. Two of these selected genes correspond to a large gene family associated to primary cell wall modification (PME3, *At3g14310*, Fig. 8a,f; an inhibitor of PME activity *At5g20740*, Fig. 8b,g), and have been described in germination as playing important roles in testa and endosperm rupture during germination in many species[29–31]. In accordance with other studies, our RNAseq data and the RT-PCR verification showed that these two genes were consistently up-regulated in both PR- and NR- batches during germination, being at the moment of TR and G where the signal was more prominent. The gene GDPD1 (Fig. 8e,j) is associated to phospholipid metabolism, and represents the down-regulation trend in the overall gene expression during germination. No differential expression was detected between PR- and NR- batches at any stage in

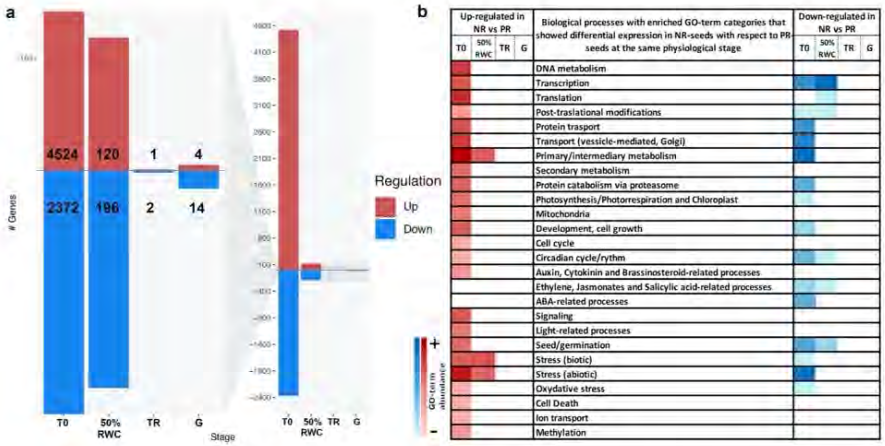


Figure 7. Differentially expressed genes in NR-batches with respect to PR-batches in each physiological stage. In (**a**) total gene count of differentially expressed genes in each physiological stage from the dry seed (T0) to the moment of radicle protrusion (G); the y axis represents total gene count. The left panel corresponds to a zoomed image of the graph in the right panel. In (**b**) GO-term enrichment in each category of biological processes represented in the gene set. The darker colours in the blue or red scales represents abundance of GO-terms and gene count within each category. The physiological stages TR and G had no particular enrichment of any category due to the small amount of genes detected.

the RNAseq data set, and this was verified by the PCR assay. This gene has been reported in *Arabidopsis* to be down-regulated during germination, with transcript levels being undetected as soon as 3 hours after imbibition[32,33]. In *C. aesculifolia*, a similar trend was detected, although the gene is still detectable by the end of germination. The two genes that represent the differential expression between PR- and NR- batches correspond to HDH (*At4g04320*, Fig. 8c,h), involved in degradation of valine for energy production[34], and XERICO (*At2g04240*, Fig. 8d,i), a master regulator of ABA metabolism and participates in osmotic and drought stress responses in seedlings and adult plants[35]. These two genes are up-regulated in T0 of NR-batches with respect to PR-batches, but their expression profiles converge by the end of germination. Overall, these genes reflect the main trends in expression changes detected in the data set and confirm that despite the initial differences in the dry seed between the two phenotypes, the germination program occurs in an orderly and stage-specific manner.

Ecophysiological implications of the changes in the transcriptomic profiles of dry seeds. Even in a metabolic quiescent state such as the mature dry seed, there are a series of changes that occur at low water contents affecting diverse cellular components and macromolecules[36–38]. Seed after ripening is a process that occurs in the dry stage that can widen or increase the seeds sensitivity and perception of environmental conditions promoting germination[39], and has been demonstrated to be a discrete developmental process[40] that allows for reversible changes in the overall transcriptomic and proteomic profiles. This allows the embryo to adjust in response to environmental stimuli before the commitment to initiate the germination program[41]. From an ecological perspective, seed release from the mother plant in the adequate season allows for the after ripening and natural priming to occur and proper synchronization of germination with the rainy season. The PR-batches displayed a phenotype that could readily incorporate the external signal produced by the priming treatment by allowing a gradual advancement of the germination program as a fine-tuning strategy to reduce the seed-batch average time needed to complete germination once the water stimulus became stable.

In the NR-batches there are two different life histories, associated with time in storage, that converge in a loss of the capability to respond to the priming treatment and resemble an "imbibed" seed. In the 2016 seed-batch, the failure to response to priming could be a result of stressful environmental conditions experienced by the mother plants, which generated an untimely release of seeds, and a change in the transcriptome that disrupted the proper transition from seed development/maturation to the quiescent stage in which after ripening occurs[38,41]. However, for the stored batches, the seeds mRNA degrades continuously during storage[37,38], and in a recent study, it was demonstrated that those changes can be detected even before notably changes in germination performance are perceived by standard longevity and vigour tests. These changes in the transcriptomic profile are especially notorious in longer transcripts, which can degrade faster than shorter transcripts[42]. This finding is in accordance with our rationale of studying the shifts in the transcriptome by analysing only those assembled sequences that have the features of a complete coding sequence, since the total read count would reflect only those sequences with the highest probability of still being fully functional in the seeds context. Thus, the germination capacity of the stored NR-batches in our tested conditions, along with their respective transcriptomic profiles, indicate that these seed batches might have some extent of the deleterious effects associated to storage[24,37,38], but the transcript profiles points to a scenario of an extended and disrupted after-ripening process rather than a deleterious effect on the

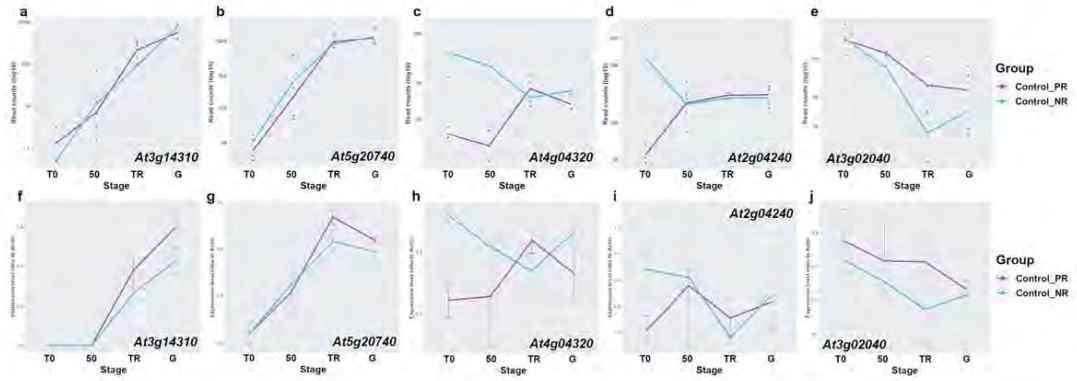


Figure 8. RNAseq expression profiles (**a–e**), and semi-quantitative RT-PCR validation (**f–j**) of the differential expression of five selected genes. RT-PCR data is represented as the ratio with respect to ACT7 (*At5g09810*) expression. Purple lines depict the average read counts (**a–e**) and average ratios (**f–j**) of the seed batches with positive response to priming (PR), while the turquoise lines depict the seed batches with no response/negative response to priming (NR). The dots in the upper row of panels depict the normalized read counts for each library, while the vertical bars in the lower row panels represent standard deviation of three independent PCR replicates.

germination program. Overall, the advancement of the transcriptomic profile in the NR-T0 seeds could reflect a risky gamble for survival ("bet-hedging")[43], by decoupling the germination program from the hydric signal in order to favour germination in a wider range of conditions, but the risk of survival could be re-assigned to the next developmental phase and have cascading effects on fitness.

Phenotypic variability within and among individuals of a particular species or group of species could be a positive feature for ensuring organismal fitness and can be studied at different levels[44]. Phenology transitions are varying as climate shifts, so these changes in the timing of developmental events, important for determining organismal fitness[45], pose another challenge to the comprehension of the basic molecular and physiological processes driving the interaction between organisms and their environment. We demonstrated that the use of physiological traits, specific to a particular developmental stage, is a reliable time-independent approach to detect common patterns between groups with wide variability and plasticity.

Material and Methods

Plant material. *Ceiba aesculifolia* (Kunth) Britten & Baker (*Malvaceae*) fruits were collected from 10–13 trees located within a human-activity disturbed ecosystem in the locations of "Trapiche del Rosario y Chicuasén" in Actopan Veracruz, Mexico (latitude 19.5426, longitude −96.7401, altitude 479 m a.s.l.), in February 2012, 2014 and 2015. The seed batch considered as "2016" was collected in December 2015 due to an un-timely flowering/fructification season in that year. These locations are conformed of fragments of the original tropical dry forest vegetation, accompanied by secondary and riparian vegetation, as well as several crop fields[22]. Two independent fruit batches were collected in February 2015, based on fruit morphology, which will be referred as 2015-1 (round to slightly oval) and 2015-2 (oval and elongated). The 2014 seed batch had a round to slightly oval shape as the 2015-1 batch. The fruits were transported to the laboratory, and were allowed to dry at room temperature and release the seeds. Immature or damaged seeds (*i.e.* incomplete testa formation, discoloration, or with visible cracks) were discarded. The seeds were stored in a dry room at 23–25 °C until used. For each initial batch, equal number of seeds from each mother plant was randomly selected. The 2012 and 2014 seed batches remained in storage for 5 and 1 years respectively, whereas the 2015 and 2016 seed batches were allowed a 30-day period of after-ripening to alleviate any primary dormancy remnants before any tests were performed.

Histological study. Seeds were fixed in FAA (Formaldehyde, ethanol, acetic acid, water 2:10:1:7) for 24 h, and dehydrated in an ethanol series. Afterwards, seeds embedded in paraplast (Paraffin), and transversally and longitudinally sectioned in a rotary microtome with a 10 µm thickness. The sections were stained with Safranin-green fast[46]. The images were obtained using Olympus BX51 microscope.

Germination performance and response to priming treatment. Seeds were submitted to a matric priming treatment. Seeds were placed in nylon mesh bags and buried in pots with a commercial soil mixture (METRO-MIX 300, Sun Gro Horticulture, MA, USA) moistened to full capacity during two days, and kept in a dark room at 23–25 °C. The upper section of the pots was covered with tin foil to avoid humidity loss. The seeds were then exhumed and allowed to dry at room temperature in darkness for two more days. *C. aesculifolia* seed germination is insensitive to light, thus imbibition in darkness does not inhibit nor promote germination in our priming conditions; however, the priming treatment was performed in total darkness in order to keep methodological consistency among seed batches collected and analysed in different years prior to the characterization of the seed response to light stimulus. Immediately after, both treated and untreated seeds were washed with 10%

hypochlorite for 1 min, surface dried with a paper towel, and placed on 1% agar plates. Five replicates (30 seeds each) of all treatments were incubated in a germination chamber at 25 °C and a photoperiod 12 h:12 h light:dark. Cumulative germination was registered daily for 20 days. A seed was considered to have germinated when the radicle had protruded through the endosperm layer. The germination performance was evaluated by fitting sigmoid curves to the arcsine-transformed germination percentages obtained for each seed batch and treatment using Table Curve 2D (version 5.01.01), and final graph construction was performed in SigmaPlot (version 11.0). The following parameters were obtained from each modelled curve for statistical analysis of germination performance: germination initiation (time at which 1% of the seeds had germinated), time to 50% germination, maximum germination rate^{-d}, and final germination percentage. These parameters were first tested for normality and equal variance assumptions using the Shapiro-Wilk and *F*-test respectively, and then for statistical differences between control and primed seeds per seed batch using a two-tailed Student's t-test, at a significance level of $\alpha = 0.05$. Priming response was deemed as favourable when at least one of the four parameters improved with respect to its control.

Imbibition curves. To determine the initial seed water content (*SWC*) in each batch and treatment, 25 seeds from each batch were weighed individually and then dried in an oven for 24 h at 100 °C. The initial *SWC* was calculated individually with the following equation:

$$SWC = ((i - d)/d)100 \quad (1)$$

where i is the initial seed weight and d is the seed weight after drying. These SWCs were then averaged for each seed batch and treatment (*SWCa*) and used as a baseline for the initial water content. Another 15 seeds from each batch and treatment were weighed individually and then placed in 1% agar plates, which were incubated in a germination chamber at 25 °C and 12/12 h photoperiod. All seeds were continuously weighed, until radicle protrusion occurred, at 1 h intervals for the first three hours, and then progressively increasing the time interval in about five-six hours or up to 12 h as the seeds maintained a constant weight. The *RWC* was calculated with equation:

$$RWC = SWCa + ((w - i)/i)100 \quad (2)$$

where *SWCa* is the averaged *SWC* obtained for a particular seed batch and treatment, i is the initial seed weight and w is the imbibed seed weight. We used Table Curve 2D to select the best fitting curve that reflected the average imbibition dynamic, and each modelled curve was used to track the changes in the slope. These data was then used to generate the model curves in Fig. 3b.

Sample collection at specific RWCs and physiological stages. Once the imbibition dynamic of each seed batch was assessed and modelled, another set of 90 seeds for control and primed seeds from each batch were individually tracked until reaching the desired RWC (16%, or 50%). These RWCs were selected based on the simplified imbibition curve generated for *C. aesculifolia* (Fig. 3b). We used the RWC ± standard deviation as the criteria for seed collection for any particular RWC at the time of measurement. The seeds exhibiting ruptured testa (TR), and germinated seeds (G) were collected and the RCW at the moment of occurrence was registered. The collection was random so at the same time of collection it was possible to collect seeds with different RWC values or different physiological states. Before the collected seeds were frozen in liquid nitrogen and stored, the testa was quickly removed and a visual inspection was performed to ensure the remainder of the seed did not show any signs of damage or decay. A total of 54 samples were stored at −80 °C until processed.

RNA extraction and *de novo* transcriptome assembly. Total RNA was extracted from 5 seeds per sample using a modified TRIZOL protocol by Li and Trick[47]. All samples were then sent to the Unidad Universitaria de Secuenciación Masiva y Bioinformática (UUSMB), where they were processed for cDNA synthesis and Illumina HiSeq 2000 library construction. Quality assessment was performed using Qubit2.0 and quantified with Bioanalyzer 2100 prior to library construction. The libraries were sequenced and *de novo* assembled in two independent runs. In both runs, a multiplex approach was used to load the corresponding samples in the flow cells. A total of 1,165,201,460 paired-end 100pb reads were obtained from both assemblies. Illumina adaptors were trimmed using Trimmomatic and quality assessment was performed using FASTQC. Reads were also trimmed to 75pb to ensure a minimum quality score of Q20 in all bases, and low quality reads were removed. For each run, a global transcriptome was assembled using Trinity v2.1.1[48,49]. For the first assembly, 117,920 sequences were found and for the second, 398,598 sequences. Library assembly quality scores are presented in the Supplementary RNAseq data file.

Transcript identification for functional classification. Functional annotation was determined using the Trinotate workflow to identify the main sources of contamination in each library, and the transdecoder tool to identify coding sequences of at least 100 amino acids in the 6 possible reading frames. A BLASTP was performed to select the proper reading frame for each transcript. The local base generated for this step included protein sequences from other assembled species in the *Malvaceae* family such as *Herrania umbratica*, *Durio zibethinus* and *Theobroma cacao*, and from the model plant *Arabidopsis*. We also performed a parallel filter for contaminants incorporating the organisms identified with the Trinnotate annotation report; the local base included model organisms like mouse, the *Pseudomonas aeruginosa* and the yeast pangenomes, as well as fungi such as *Trichoderma reesei*, *Fusarium fujikuroi*, *F. verticillioides* and *Diplodia corticola*. The identified proteins whose best hit matched one of the proteins in the contaminant database were filtered. We selected the best blast hit following

these criteria: 60% minimal length coverage, 40% minimal identity. About 90% of confirmed transcripts matched an *Arabidopsis* locus-tag.

We then performed a cross-check validation of the Trinnotate and the transdecoder methods to confirm the identity of the transcripts and to explore the structural characteristics of the assembled transcripts. The *de novo* assembly relies on de Bruijn graphs which provide higher sensitivity for splice-variant calling, however it can overestimate transcript diversity due to assembly artefacts or transcript incompleteness[50]. Thus, after the contaminant-filtering step, the resulting transcript database included 34,274 transcripts from the first assembly and 132,271 from the second assembly, and included fully assembled sequences with open reading frames (ATG and STOP codons) flanked by possible 5′ and 3′ UTR regions, and a mixture of incomplete transcripts lacking either the 5′ or the 3′ portion of the coding sequence (no ATG or STOP codon evidence), or lacking both ATG and STOP codons. Since the ultimate purpose of the assembly relies on analysing the biologically functional processes associated to germination using *Arabidopsis* as a proxy, we decided to select only the complete transcripts and their specific read counts for posterior analyses. The complete transcript pool also included those 5′ coding sequences which were manually examined using the non-redundant protein data base in the NCBI BLAST tool to verify if they had a start codon. This filtering step ensures that the read counts used for differential expression belong to functional units, and could avoid overestimation of transcripts or biased read inflation in some locus-tags due to similarity-miscalling of sequences within gene families[50], and takes into consideration the possibility that some of the incomplete transcripts could be a by-product of the dynamic degradation of stored mRNAs that occurs during germination[37,38,42].

Differential expression biological function categories analyses. We used the IDEAMEX website R-based tools for differential expression analysis, to simultaneously test differential expression with the R-workflows EdgeR, Limma, DESeq2 and NOISeq[51]. After an exploratory test, we decided to use DESeq. 2 for contrasts between physiological stages and time-series analysis using a factorial design to test the combined effects of stage and phenotype, and a reduced model to test for changes among physiological stages in each phenotype, using as cut-off values to consider a gene as differentially expressed a $P < 0.05$ and FDR < 0.05 and an expression change of at least 1 in log2 scale. The data sets of differentially expressed genes across germination were used to generate fuzzy clusters of expression patterns with the Mfuzz package, and a minimum membership threshold of 0.5. Simulations were run to test cluster stability and selected the simulation run with 8 clusters as the smallest number of clusters that reflect the major changes in the data set without excessive overlapping of clusters[52,53]. The heatmap plots were constructed with heatmap.2 (gplots v3.0.1 package) using default parameters (R v3.6.0 MacOS[54]). PCA plots were constructed with DESeq2, using the variance stabilizing transformation (vst) option for data normalization of gene counts[55]. Functional category classification was performed using the gene-ontology and KEGG-pathways options within the Functional annotation tool in the DAVID Bioinformatics Resources website (v6.8)[56–60]. The resulting GO-term and pathway clusters were then grouped in 32 main categories presented in Figs. 4b and 5b. The functional categories present in the dry seed (T0-PR and T0-NR) were ranked by abundance of GO-terms and genes contained within in order to generate the faux colour scale. The same process was performed for the up- and down-regulated categories, respectively.

Semi-quantitative RT-PCR expression validation. A set of five randomly selected locus tags (*At2g04240*, *At3g02040*, *At3g14310*, *At4g04320*, and *At5g20740*), plus ACT7 (*At5g09810*, used as a control gene for normalization[29]), were used to validate RNAseq expression data. RNA samples from both the PR and NR batches at T0, 50% RWC, TR and G were used to synthesize cDNA libraries. RT-PCR runs were performed in triplicate. A set of preliminary runs of 20, 25, 30, 32 and 35 cycles were performed in order to verify the best amplification conditions. At 32 cycles, ACT7 and *At2g04240* showed saturation, thus all runs were performed at 30 cycles. The primers used are shown in Supplementary information Table S4. Images with the PCR products were processed in GIMP (2.10, MACOS) and data was imported into excel and R for graph construction.

Data availability

Reads were deposited in the SRA platform within the NCBI GenBank website under Bioproject PRJNA561202, and accession numbers SAMN12611291 and SAMN12611291. Assembled sequences were deposited in the TSA platform under accession number GHVB00000000.

Received: 21 February 2020; Accepted: 26 May 2020;
Published online: 26 June 2020

References

1. Bewley, J. D., Bradford, K. J., Hilhorst, H. W. & Nonogaki, H. Germination. In *Seeds* pp. 133–181 (Springer New York, 2013).
2. Meyer, R. S. & Purugganan, M. D. Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* **14**, 840–852 (2013).
3. Gepts, P. The contribution of genetic and genomic approaches to plant domestication studies. *Curr. Opin. Plant Biol.* **18**, 51–59 (2014).
4. Jiménez-Alfaro, B., Silveira, F. A., Fidelis, A., Poschlod, P. & Commander, L. E. Seed germination traits can contribute better to plant community ecology. *J. Veg. Sci.* **27**, 637–645 (2016).
5. Bradford, K. J. Water stress and the water relations of seed development: a critical review. *Crop Sci.* **34**, 1–11 (1994).
6. Villela, F. A. Water relations in seed biology. *Sci. Agr.* **55**, 98–101 (1998).
7. Nonogaki, H., Bassel, G. W. & Bewley, J. D. Germination—still a mystery. *Plant Sci.* **179**, 574–581 (2010).
8. Manz, B., Müller, K., Kucera, B., Volke, F. & Leubner-Metzger, G. Water uptake and distribution in germinating tobacco seeds investigated *in vivo* by nuclear magnetic resonance imaging. *Plant Physiol.* **138**, 1538–1551 (2005).
9. Rathjen, J. R., Strounina, E. V. & Mares, D. J. Water movement into dormant and non-dormant wheat (*Triticum aestivum* L.) grains. *J. Exp. Bot.* **60**, 1619–1631 (2009).

10. Munz, E. *et al.* A functional imaging study of germinating oilseed rape seed. *New Phytol.* **216**, 1181–1190 (2017).
11. Howell, K. A. *et al.* Mapping metabolic and transcript temporal switches during germination in rice highlights specific transcription factors and the role of RNA instability in the germination process. *Plant Physiol.* **149**, 961–980 (2009).
12. Morris, K. *et al.* Regulation of seed germination in the close Arabidopsis relative *Lepidium sativum*: a global tissue-specific transcript analysis. *Plant Physiol.* **155**, 1851–1870 (2011).
13. Dekkers, B. J. *et al.* Transcriptional dynamics of two seed compartments with opposing roles in Arabidopsis seed germination. *Plant Physiol.* **163**, 205–215 (2013).
14. Dekkers, B. J., Pearce, S. P., van Bolderen-Veldkamp, R. P. M., Holdsworth, M. J. & Bentsink, L. Dormant and after-ripened *Arabidopsis thaliana* seeds are distinguished by early transcriptional differences in the imbibed state. *Front. Plant Sci.* **7**, 1323, https://doi.org/10.3389/fpls.2016.01323 (2016).
15. Narsai, R. *et al.* Extensive transcriptomic and epigenomic remodeling occurs during *Arabidopsis thaliana* germination. *Genome Biol.* **18**, 172, https://doi.org/10.1186/s13059-017-1302-3 (2017).
16. Yu, Y. *et al.* Transcriptome analysis during seed germination of elite Chinese bread wheat cultivar Jimai 20. *BMC Plant Biol* **14**, 20, https://doi.org/10.1186/1471-2229-14-20 (2014).
17. Maia, J., Dekkers, B. J. W., Provart, N. J., Ligterink, W. & Hilhorst, H. W. M. The re-establishment of desiccation tolerance in germinated *Arabidopsis thaliana* seeds and its associated transcriptome. *Plos One* **6**, e29123, https://doi.org/10.1371/journal.pone.0029123 (2011).
18. Dubrovsky, J. G. Seed hydration memory in Sonoran Desert cacti and its ecological implication. *Am. J. Bot.* **83**, 624–632 (1996).
19. González-Zertuche, L. *et al.* A. Natural priming of *Wigandia urens* seeds during burial: effects on germination, growth and protein expression. *Seed Sci. Res.* **11**, 27–34 (2001).
20. Paparella, S. *et al.* Seed priming: state of the art and new perspectives. *Plant Cell Rep.* **34**, 1281–1293 (2015).
21. Herrerías-Diego, Y., Quesada, M., Stoner, K. E. & Lobo, J. A. Effects of forest fragmentation on phenological patterns and reproductive success of the tropical dry forest tree *Ceiba aesculifolia*. *Conserv. Biol.* **20**, 1111–1120 (2006).
22. Velázquez-Rosas, N., Ruiz-Guerra, B., Sánchez-Coronado, M. E., Gamboa-de Buen, A. & Orozco-Segovia, A. Morphological variation in fruits and seeds of *Ceiba aesculifolia* and its relationship with germination and seedlings biomass. *Bot. Sci.* **95**, 81–91 (2017).
23. Zheng, Y., Wang, J., Zhu, Y. & Wang, A. Research and application of kapok fiber as an absorbing material: A mini review. *J. Env. Sci.* **27**, 21–32 (2015).
24. Kranner, I., Minibayeva, F. V., Beckett, R. P. & Seal, C. E. What is stress? Concepts, definitions and applications in seed science. *New Phytol.* **188**, 655–673 (2010).
25. Hadas, A. & Stibbe, E. An analysis of soil water movement towards seedlings prior to emergence. In Physical Aspects of Soil Water and Salts in Ecosystems (eds. Hadas, A., Swartzendruber, D., Rijtema, P. E., Fuchs, M. & Yaron, B.) 97–106 (Springer Berlin Heidelberg, 1973).
26. Weitbrecht, K., Müller, K. & Leubner-Metzger, G. First off the mark: early seed germination. *J. Exp. Bot.* **62**, 3289–3309 (2011).
27. Nakabayashi, K., Okamoto, M., Koshiba, T., Kamiya, Y. & Nambara, E. Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: epigenetic and genetic regulation of transcription in seed. *Plant J.* **41**, 697–709 (2005).
28. Graeber, K., Linkies, A., Wood, A. T. & Leubner-Metzger, G. A guideline to family-wide comparative state-of-the-art quantitative RT-PCR analysis exemplified with a Brassicaceae cross-species seed germination case study. *Plant Cell* **23**, 2045–2063 (2011).
29. Müller, K. *et al.* Demethylesterification of cell wall pectins in Arabidopsis plays a role in seed germination. *Plant Physiol.* **161**, 305–316 (2013).
30. Scheler, C. *et al.* Promotion of testa rupture during garden cress germination involves seed compartment-specific expression and activity of pectin methylesterases. *Plant Physiol.* **167**, 200–215 (2015).
31. Steinbrecher, T. & Leubner-Metzger, G. The biomechanics of seed germination. *J. Exp. Bot.* **68**, 765–783 (2017).
32. Winter, D. *et al.* An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PloS ONE* **2**, e718, https://doi.org/10.1371/journal.pone.0000718 (2007).
33. Bassel, G. *et al.* Elucidating the germination transcriptional program using small molecules. *Plant Physiol.* **147**, 143–155 (2008).
34. Schertl, P., Danne, L. & Braun, H. P. 3-Hydroxyisobutyrate dehydrogenase is involved in both, valine and isoleucine degradation in *Arabidopsis thaliana*. *Plant Physiol.* **175**, 51–61 (2017).
35. Ko, J. H., Yang, S. H. & Han, K. H. Upregulation of an Arabidopsis RING-H2 gene, XERICO, confers drought tolerance through increased abscisic acid biosynthesis. *Plant J.* **47**, 343–355 (2006).
36. Rajjou, L. *et al.* Proteome-wide characterization of seed aging in Arabidopsis: a comparison between artificial and natural aging protocols. *Plant Physiol.* **148**, 620–641 (2008).
37. Rajjou, L. *et al.* Seed germination and vigor. *Ann. Rev. Plant Biol.* **63**, 507–533 (2012).
38. Sano, N. *et al.* Staying alive: molecular aspects of seed longevity. *Plant Cell Physiol.* **57**, 660–674 (2016).
39. Finch-Savage, W. E. & Leubner-Metzger, G. Seed dormancy and the control of germination. *New Phytol.* **171**, 501–523 (2006).
40. Carrera, E. *et al.* Seed after-ripening is a discrete developmental pathway associated with specific gene networks in Arabidopsis. *Plant J.* **53**, 214–224 (2008).
41. Catusse, J., Job, C. & Job, D. Transcriptome-and proteome-wide analyses of seed germination. *CR Biol.* **331**, 815–822 (2008).
42. Zhao, L., Wang, S., Fu, Y. B. & Wang, H. Arabidopsis seed stored mRNAs are degraded constantly over aging time, as revealed by new quantification methods. *Front. Plant Sci.* **10**, 1764, https://doi.org/10.3389/fpls (2020).
43. Cohen, D. Optimizing reproduction in a randomly varying environment. *J. Theor. Biol.* **12**, 119–129 (1966).
44. Mitchell, J., Johnston, I. G. & Bassel, G. W. Variability in seeds: biological, ecological, and agricultural implications. *J. Exp. Bot.* **68**, 809–817 (2016).
45. Burghardt, L. T., Edwards, B. R. & Donohue, K. Multiple paths to similar germination behavior in *Arabidopsis thaliana*. *New Phytol.* **209**, 1301–1312 (2016).
46. Márquez-Guzmán, J., Wong, R., Pérez-Pacheco, M., López-Curto, L. & Murguía-Sánchez, G. Técnica de laboratorio para el estudio del desarrollo en angiospermas. (Facultad de Ciencias, Universidad Nacional Autónoma de México, 2016).
47. Li, Z. & Trick, H. Rapid method for high-quality RNA isolation from seed endosperm containing high levels of starch. *BioTechniques* **38**, 872–876 (2005).
48. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
49. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).
50. Ungaro, A. *et al.* Challenges and advances for transcriptome assembly in non-model species. *Plos One* **12**, e0185020, https://doi.org/10.1371/journal.pone.0185020 (2017).
51. Jimenez-Jacinto, V., Sanchez-Flores, A. & Vega-Alvarado, L. Integrative Differential Expression Analysis for Multiple EXperiments (IDEAMEX): a web server tool for integrated RNA-seq data analysis. *Front. Genet.* **10**, 279, https://doi.org/10.3389/fgene.2019.00279 (2019).
52. Futschik, M. & Carlisle, B. Noise robust clustering of gene expression time-course data. *J. Bioinf. Comput. Biol.* **3**, 965–988 (2005).
53. Kumar, L. & Futschik, M. Mfuzz: a software package for soft clustering of microarray data. *Bioinformation.* **2**, 5–7 (2007).

54. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL, https://www.R-project.org/ (2016).
55. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq. 2. *Genome Biol.* **15**, 550, https://doi.org/10.1186/s13059-014-0550-8 (2014).
56. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nat. Protoc.* **4**, 44–57 (2009).
57. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucl. Acids Res.* **37**, 1–13 (2009).
58. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* **28**, 27–30 (2000).
59. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucl. Acids Res.* **47**, D590–D595 (2019).
60. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).

Acknowledgements

This work was supported by Programa de Apoyo a Proyectos de Investigación e Inovación Tecnológica (PAPIIT UNAM) [grant IN203218 to AGB], and by the Project OE001 "*Ceiba aesculifolia* como especie modelo para el estudio molecular del envejecimiento de las semillas en relación al desempeño germinativo" financed by the Dirección General del Sector Primario y Recursos Naturales Renovables within Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) and by the Comisión Nacional para la Biodiversidad (CONABIO). X. Gómez-Maqueo was supported by the scholarship provided by the Consejo Nacional de Ciencia y Tecnología (CONACyT) [grant 280483] and by Dirección General de Asuntos del Personal Académico (DGAPA UNAM) [grant IN203218]. This paper constitutes a partial fulfilment of the Posgrado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM). We thank Drs. Alejandro Sánchez-Flores, Ricardo Grande, Verónica Jiménez-Jacinto, and Leticia Vega-Alvarado from Unidad de Secuenciación Masiva y Bioinformática, Instituto de Biotecnología, UNAM, for their support in the transcriptome assembly and bioinformatics. We thank Rodrigo García-Herrera and José Miguel Baltazar Galvez for his technical assistance with High-throughput computing services at the LANCIS-Instituto de Ecología and Instituto de Ecología UNAM respectively. We thank David Cruz-Sánchez for his technical assistance with the relational data base generated for this dataset. We thank Karina Macias-Ruiz for her assistance in primer design and RT-PCR analysis.

Author contributions

X.G.M., collected the initial biological material and produced the RNA samples, designed and performed experiments, analysed the data and wrote the manuscript. D.S. assisted with the examination of the assembled sequence data, analysed the data and reviewed the manuscript. S.A.L., performed experiments and analysed the germination data. N.V.R., collected the initial biological material, registered the climate variables, participated in the interpretation of the eco-physiological context associated to the data and reviewed the manuscript. M.M.G., collected the initial biological material, and performed experiments. K.J.D., performed the histological studies and reviewed the manuscript. A.G.B. collected the initial biological material, designed the experiments, analysed the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-66759-3.

Correspondence and requests for materials should be addressed to X.G.-M. or A.G.-d.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

CAPÍTULO III. El papel de los genes asociados a la pared celular durante el proceso de germinación

Resumen

La germinación de la semilla es un proceso complejo en el que el embrión, encerrado dentro de los tejidos circundantes, debe cambiar rápidamente de un programa de maduración a un proceso de desarrollo y crecimiento. Este cambio es impulsado por la germinación, etapa que preparará al embrión para establecimiento de la nueva plántula. El proceso de germinación inicia con la absorción de agua por la semilla seca y culmina, por lo general, con la protrusión de la radícula. La emergencia de la radícula de la semilla es un proceso altamente regulado que involucra cambios discretos y coordinados en la extensibilidad de la pared celular y la reorganización de sus componentes, entre otros procesos. En este capítulo se hizo una revisión del conocimiento actual del proceso fisiológico controlado de separación y expansión celular, que le dan a la pared celular primaria sus propiedades plásticas al "relajar" los componentes principales de la pared celular durante la germinación de la semilla. En esta revisión se incluyen los análisis de la regulación que ocurre a nivel transcriptómico y bioquímico de los genes de pared celular durante la germinación. También se hizo énfasis en la importancia fisiológica de la actividad *in muro* de una amplia variedad de enzimas modificadoras de la pared celular que incluyen hidrolasas y transglucosilasas, así como procesos no enzimáticos como el relajamiento mediado por expansinas durante la germinación de la semilla.

Chapter 7

The Dynamics of Plant Cell Wall *In Muro* Modifications and its Physiological Implications on Seed Germination

Ximena Gómez-Maqueo and
Alicia Gamboa-deBuen

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/64085

Abstract

Seed germination is a complex process in which the embryo, enclosed within the surrounding tissues, must quickly switch from a maturation program to a germination-driven developmental process that will prepare the embryo for seedling growth and establishment. The germination process initiates with water uptake by the dry seed and culminates, usually, with the radicle protrusion. The radicle emergence from the seed is a highly regulated process that involves discrete and coordinated changes in plant cell wall extensibility and rearrangements of its components, among other processes. In this chapter we will review current knowledge of the physiological process of controlled cell separation and expansion, which give the primary cell wall its plastic properties by "loosening" of the main components of the cell wall during seed germination. We will focus on the physiological importance of primary cell wall constitution and modification by the activity *in muro* of a broad variety of cell wall-modifying enzymes that include hydrolases and transglycosylases, as well as non-enzymatic processes such as expansin-mediated loosening during seed germination.

Keywords: cell wall modification, primary cell wall, seed germination

1. Introduction

Seeds constitute a critical stage in the life cycle of embryophytes. In this stage, the plant embryo remains in a quiescent state until the proper conditions of temperature, water availability, and, in some species, light are met in order for the processes of germination and seedling establishment to occur [1, 2]. The mature seed contains the embryo, which is surrounded by the seed coat

(testa) that is derived from the maternal tissues and in some species by one or more layers of storage tissue (endosperm) [2]. Seeds can function as resistance structures. Several mechanisms have evolved, in tight relation with the environment, to ensure the survival of the quiescent embryo [3]. Part of these mechanisms includes the modification of the structure and composition of plant cell walls.

One characteristic feature of plant cells is that they are enclosed in a polysaccharide and protein matrix, denominated as cell wall [4]. Plant cells can have two different types of wall. Primary walls, produced during cytokinesis, are flexible structures that regulate cell growth and shape. The secondary walls are deposited after the cell has achieved its final size and shape, by the inclusion of lignin and other phenolic compounds, thus making the cell wall rigid and usually impermeable. Cell walls have several functions that include the regulation of cell-cell adhesion and abscission, apoplastic transport, mechanical support and maintenance of turgor pressure, and defense against pathogens [2, 5]. In seeds, cell walls are modified in order to generate hard, and in some cases impermeable, coats that protect the embryo from the environmental conditions. Also, seed cell walls can store energy that can be mobilized to feed embryo growth and development. Finally, cell walls regulate the timing of seed germination by fine-tuning the processes of matrix polysaccharide loosening/breakage, as well as the integration of environmental cues with the hormonal and physiological status of the embryo [4, 6]. In this chapter we will focus only on primary cell walls and their importance on seed germination.

2. Seed germination

Seed germination is a physiological process initiated with water uptake and culminating with the emergence of the embryo through its protective tissues, which might include the testa, endosperm, perisperm, or pericarp [2]. The testa and the endosperm rupture must be coordinated with environmental seasonality to facilitate germination in the most favorable conditions [1, 6]. Several mechanisms have evolved to ensure proper synchronization of germination with environmental cues; among these is the interplay of hormonal signaling pathways via abscisic acid (ABA), gibberellins (GA), ethylene, and jasmonates [7–10]. These hormones exert their regulation on germination through different pathways including cell wall remodeling [7, 11].

In the classical model of seed germination described by Bewley et al. [12], the process of germination is divided into three phases, distinguished by the rate of water absorption by the seed tissues. The phase I, or imbibition phase, is characterized by a rapid water uptake rate driven by the difference in water potential between the seed and the environment. In this phase also the reactivation of primary metabolism and DNA repair pathways starts. Next, in phase II or activation phase, the imbibition rate decreases, water content remains stable, and major changes in the metabolic pathways and activation of other cellular processes take place. In this phase the integration of environmental cues with the internal status of the seed that will determine whether or not the seed will enter into the next phase occurs. Finally, in phase III there is another rapid water uptake driven by radicle protrusion and is mainly related to seedling growth. Germination is completed once the radicle has emerged at the onset of phase

http://dx.doi.org/10.5772/64085

III. This triphasic model of imbibition can be applied to all seeds analyzed thus far [12, 13]. The imbibition time needed for completion of germination is highly variable among species and even within seed lots, and it depends on several factors like seed history and environmental conditions experienced by the mother plant at the moment of seed dispersion and during the after-ripening period [12, 14, 15].

It is now generally accepted that radicle protrusion occurs by two nonexclusive processes [2, 13]. The first process involves a decrease in the mechanical resistance of the enclosing tissues, especially in the micropylar region of the testa and endosperm [2, 10]. The second process deals with an increasing growth potential of the embryo, driven by turgor pressure and cellular expansion in the embryonic axis [2, 13]. Most of the knowledge generated about the regulation of radicle protrusion comes from endospermic seeds, where testa and endosperm rupture can occur in two easily distinguishable stages (*Arabidopsis thaliana* –*Arabidopsis*–, *Chenopodium album*, *Lepidium* sp., *Nicotiana* sp., to mention a few) [2].

In recent years, with the advent of whole genome/transcriptome analysis, it has been possible to study the process of germination with high spatial-temporal resolution. Transcriptomic analysis allows a comprehensive view of seed germination by dissecting "early" or "late" germination processes, the first being the initial response to water and the second corresponding to the interval from the imbibed seed to the radicle protrusion [14, 15]. Also, in endospermic seeds, an important landmark is the distinction between the processes that occur prior to testa rupture and after it that leads to endosperm rupture [8, 16–18].

Several studies demonstrate that the main transcripts, enzymes, and other proteins accumulated in dry seeds participate in primary metabolism, starch and storage protein mobilization, reactive oxygen species (ROS) scavenging, and cell wall synthesis [14, 15]. Aside from providing building blocks to sustain protein production and cell growth, the reactivation of primary metabolism in the early stages of seed germination plays a major role in the generation of the proper redox state to promote the activity of different enzymes and produce energy to support processes essential for radicle protrusion [14, 19].

In *Arabidopsis*, the seed development and maturation programs are regulated by the LAFL transcription factor network (LEAFY COTYLEDON 1 (LEC1) and LEC1-LIKE (L1L), ABA INSENSITIVE 3 (ABI3), FUSCA 3 (FUS3), and LEC2), which activates other downstream transcription factor networks in concerted action of hormone, sugar, and light signalization pathways. Some target genes are involved in ABA, GA, ethylene, brassinoesteroids (BR), auxin, jasmonic acid (JA), and cytokinin (CK) signalization pathways [20]. The ABA signalization pathway participates in the regulatory networks of seed maturation, reserve accumulation, and desiccation tolerance acquisition [21]. GA blocks the LAFL and ABA networks during germination. The degradation of transcripts and enzymes related to seed maturation, which accumulated in the dry seed, has been described to occur in the first 6–12 h of seed imbibition in *Arabidopsis* [22] and within the first 24 h in *rice* and *barley* [14, 23].

Gibberellins play a major role in promoting a myriad of developmental programs, and its antagonistic role in ABA-mediated block of germination has been described [24]. GA stimulates seed germination by enhancing embryo growth; embryos of *Arabidopsis* GA-deficient

mutant seeds exhibit reduced growth rate phenotypes [25]. Also, GA enhances seed germination by overcoming the mechanical restraint to radicle protrusion of the surrounding tissues. In *Solanum lycopersicum* (tomato), GA-deficient embryos (unable to germinate unless incubated with exogenous GA) can grow into dwarf plants when the testa and the endosperm were removed mechanically [26]. This role of GAs in stimulating germination can be linked to the upregulation of several cell wall-modifying proteins (CWMPs) detected in whole-seed *Arabidopsis* transcriptomes of *ga1-3* mutants treated with GA_4 [24]. Jacobsen and Pressman [27] suggested that the embryo of celery (*Apium graveolens*) seeds does not secrete CWMPs but rather promote the activity of GA-inducible CWMPs in the endosperm. The depletion of the endosperm in this species generates a space where the embryo cells can expand and eventually penetrate the micropylar endosperm.

An overrepresentation analysis (ORA) of gene ontologies showed that transcription regulation is enriched in both the endosperm and the embryo transcriptomes of *Arabidopsis* seeds. The ORA analysis also showed that in the endosperm, the main biological processes are associated with cell wall metabolism, cell death, response to biotic stimulus, and defense and response to ABA. The main biological processes in the embryo include phosphate metabolic process, protein amino acid phosphorylation, hormone metabolic process (particularly auxin synthesis and transport), cell division and cell cycle, post-germination regulation of growth and organ development, and signaling [17].

3. Cell wall structure and composition

Plant cell walls are complex and highly dynamic structures composed of a variety of polysaccharides, proteins, and aliphatic or aromatic compounds [28, 29]. They are continually being modified throughout development and in response to environmental stimuli [30, 31]. Primary cell walls of flowering plants can be classified in two main groups depending on its general architecture and composition, as well as their biosynthetic processes [32, 33]. Type I cell walls are the most common, present in dicotyledonous and the non-commelinoid monocotyledonous plants (a more basal group of aroids, alismatids, and lilioids). Type II cell walls are found only in the commelinoid monocots that include the Poales (members of the families Poaceae, Bromeliaceae, and Cyperaceae) [32, 33].

3.1. Primary cell wall polysaccharides

Primary cell wall polysaccharides constitute the majority of the wall dry mass in land plants and can be grouped in three main classes: cellulose, hemicelluloses, and pectins [30]. Cellulose is a linear 1,4-β-D-glucan that assembles into partially crystalline microfibrils, each of which contains about 36 parallel polysaccharide chains [34]. Cellulose is synthesized *in muro* by the cellulose synthase complex (CSC), embedded within the plasma membrane and formed by 6 rosette subunits that contain 6 cellulose synthase proteins (CESA) [35]. Aside from cellulose, all other cell wall polysaccharides are synthesized and processed for wall targeting in the trans-Golgi system [5].

Hemicelluloses are polymers whose backbones consist of β-glucose, β-xylose, or β-mannose, with short side chains. In all vascular plants with type I walls, the most common hemicelluloses are xyloglucans (XyG), whereas type II cell walls contain less XyG, being the most abundant glucuronoarabinoxylans (GAX) and β1,3:β1,4 mixed glucans [33]. Hemicellulose chains adhere to cellulose microfibrils, in a rope-like manner, to restrain cell expansion [30, 34]. Also, in type I cell walls, hemicelluloses bind and cross-link with pectin and form the hydrated matrix [30].

The group generally known as pectins comprises over 30% of the cell wall total mass in dicots [31, 36]. Pectins are acidic heteropolymers that form a hydrated gel, in which cellulose and other molecules are embedded in the plant cell wall. Their main defining feature is 1,4-linked α-D-galacturonic acid residues (GalA). Pectins interact covalently and non-covalently with other pectin molecules or with hemicellulose xyloglucan or arabinogalactans [31]. Several studies support the hypothesis that the three major pectin classes, homoglacturonan (HG), rhamnogalacturonan I, and rhamnogalacturonan II, are covalently linked in the cell wall [29, 37], forming a hydrophilic macromolecular network. Pectin is deposited on the cell wall matrix in a highly methylesterified form [28]. The methyl group is removed by pectin methylesterases (PMEs) in muro, providing an anionically charged matrix and changing the mechanic properties of the cell wall. Increasing evidence shows that the regulation of the degree of methylesterification of the pectic matrix plays a fundamental role in plant growth, development, morphogenesis, cell-cell adhesion, cell expansion, seed hydration, and seed germination [5, 16, 36, 38].

Other polysaccharides present in primary cell walls of various species are the mannans, arabinoxylans, and arabinogalactans. Mannans are formed by mannosyl residues linked by β-1,4-glycosidic linkages. This mannosyl backbone can contain glucose residues (glucomannans) or be further substituted by single galactose residues with α-1,6-linkages (galactomannans). Arabinoxylan consists of a (1,4)-linked β-D-xylan backbone decorated with arabinose branches. Other residues, such as glucuronic acid and ferulic acid esters (FAE), are also attached in arabinoxylans that are particularly abundant in cereal grasses. Arabinogalactan and storage xyloglucans are used as reserves in cotyledons. The basic structure of storage xyloglucans differs from the primary wall xyloglucans in that it is not fucosylated [39].

3.2. Cell wall proteins

Primary cell walls are mainly constituted by polysaccharides; however, proteins account for about 10% of the total dry mass of the wall [40]. Proteins that contain a secretion signal peptide, which targets them to the secretory pathway and in most cases is excised to allow activation or proper protein function, are commonly referred as classical cell wall proteins [40–42]. In *Arabidopsis*, about 17% (~5000 genes) of the genome encodes for proteins targeted to the secretory pathway, and of this, about 1000–2000 genes could be cell wall proteins (CWPs). Cell wall proteins have several functions as structural, enzymatic, and defense and have been grouped in functional categories by different authors. The proteins with a structural function or those acting on polysaccharides are the two main functional categories [30, 42]. These

proteins are expressed in a tissue-specific and process-specific manner, contributing to the regulation of cell wall stabilization and rigidity [41].

3.2.1. Structural proteins

Structural proteins are usually classified by the predominant amino acids in their sequence, although some of them can belong to more than one category. The most common include the hydroxyproline-rich glycoproteins (HRGPs), the glycine-rich proteins (GRPs), the proline-rich proteins (PRPs), and the arabinogalactan proteins (AGPs). These proteins vary greatly in abundance within plant species, cell tissues, and environmental conditions [4]. Arabinogalactan proteins, that are widely distributed among plant families and comprise about 2–10% of the total protein in the wall, are highly glycosylated. Also, AGPs are rich in hydroxyproline, serine, alanine, threonine, and glycine, and resistant to proteolysis in their native state [41]. Extensins are a family of HRGPs particularly abundant in dicots that have been involved in modification of wall extensibility in elongating tissues [43].

3.2.2. Proteins acting on polysaccharides

Within the CWPs acting on polysaccharides, there are a broad variety of activities. For instance, the group of glycosyl hydrolases (GHs) include glycosidases (β-glucosidase, β-galactosidase, β-xylosidase, and α-xylosidase, and exo-polygalacturonases) and glycanases like β-mannanase, β-xylanase, (1→4)-β-glucanase "cellulase", endo-polygalacturonases, and xyloglucan endo-hydrolase (XEH). The combined activity of these kinds of enzymes is theoretically capable of hydrolyzing most of the glycosidic bonds in the cell wall polysaccharides but do not imply that all enzymes are active at the same time or tissue. The glycosyltransferases (GTs) activity involves the formation of a glycosyl-enzyme complex that is attacked by an acceptor substrate (another oligo/polysaccharide). This activity allows the integration of recently secreted polysaccharides into the matrix and the grafting of polysaccharides already present in the wall matrix [44]. This category includes the xyloglucan endotransglycosylase (XET). Both XEH and XET proteins are commonly grouped within the xyloglucan endotransglucosylase/hydrolase family (XTH) due to some of their members (like β-xylanase) that can have both GH and GT activities [30].

Polysaccharide lyases (PLs) promote cell separation by calcium-dependent de-polymerization of wall polygalacturonides. Plant pectate lyases are a group of enzymes that catalyze the cleavage of de-methylesterified pectin. PL activity has being described in cell wall degradation that occurs during fruit ripening [45], and Penfield et al. [7] report 34 pectate lyases that were downregulated in the endosperm of *Arabidopsis* imbibed seeds after treatment with ABA.

Carbohydrate esterases (CEs) include two enzyme families that have activity over pectins, the PMEs and pectin acetylesterases (PAEs), and the family of xylan acetylesterases. These enzymes cleave methyl or acetyl groups from the HG or Xyl backbone of polysaccharides [30].

PMEs catalyze the reaction by which methylesters are cleaved from a HG chain, producing a free carboxyl group and the release of a proton and methanol [46]. Plant PMEs are mainly alkaline isoforms bound to the wall matrix, while some isoforms are neutral and easily

solubilized or free apoplastic acidic isoforms. Alkaline isoforms seem to be the PMEs with most de-methylesterification activity, but the kinetics of PME activity is affected by the ionic composition of the matrix, thus influencing PME activity and mobility [47]. PME activity can lead to two different cell wall fates: the first one would be the formation of a rigid, stable structure by Ca^{2+} interaction with de-methylesterified GalA residues (>10) in the HG chains. The second fate of HG would be their degradation by polygalacturonases, where only small stretches or individual GalA residues are de-methylesterified, thus leading to a more relaxed matrix [28, 46]. Also, PME activity acidifies the cell wall; this acidification would allow expansin activity ("acid grow") [5]. PMEs are antagonistically regulated in the cell wall by proteinaceous PME inhibitors (PMEIs) meanwhile PGs by PG inhibitors (PGIPs) [28].

Expansins regulate cell wall loosening in a pH-dependent manner by disruption of the hydrogen bonds between xyloglucans and cellulose. Sequence analysis indicates that expansins contain an N-terminal domain slightly similar to the catalytic domain of the family-45 endoglucanases; however no catalytic activity has been reported.

In *Arabidopsis*, about 10% of the total CWPs described in cell wall proteomes from different tissues and plants correspond to gene families with domains of unknown function (DUF) [42]. Mewalal et al. [48] have pointed out the relevance of these DUF families on cell wall dynamics. In particular, the two plant-specific families DUF231 and DUF642 could be involved in pectin modification [49, 50].

3.3. Cell wall modification

The molecular modification of the wall network can result in the relaxation of wall stress or "wall loosening" by the controlled rearrangement of cellulose/matrix polymers, which involve sliding of a cross-link along a scaffold or the breakage of stress-bearing cross-links without substantial changes in wall dimensions. These rearrangements could include three processes: (a) the cleavage of the backbone of major matrix polymers, (b) the weakening of the noncovalent bonding between polysaccharides, and (c) the breakage of cross-links [5]. Following cell wall loosening, there are three main types of outcomes: cell expansion, cell separation, and wall stiffening. Cell wall enlargement occurs secondarily as a result of water uptake and the reduction of turgor pressure resulting from wall loosening [44].

Reactive oxygen species (ROS) like hydrogen peroxide (H_2O_2), hydroxyl radical (·OH), and superoxide radical (O_2^-) have been proposed to play a major role in germination by participating in defense against pathogens, signaling, and promotion of cell wall loosening [2, 9]. ROS can negatively affect germination by reacting with almost all macromolecules stored in the seed, causing oxidative damage and cleavage of polysaccharide chains in the cell walls [5, 9]. The participation of ROS in cell wall loosening and promotion of germination might be indirect, through the ethylene signaling pathways that involve ROS production and downstream activation of CWMPs [9]. Cosgrove [5] suggests the revision of ROS participation in the process of wall loosening, since in most studies reporting ROS-mediated extensibility comprises only a small fraction (about 1% extensibility) and the assays with higher ROS concentrations provoke wall breakage.

3.4. Role and regulation of cell wall enzymes and proteins during germination

The study of plant cell wall structure and physiology has achieved a major progress from the input of "-omics" technologies in the past two decades. These -omics technologies are able to capture the complexity of biological processes, like seed germination and cell wall modification, with high sensitivity and spatial-temporal resolution. A tissue-specific transcriptome analysis in *Arabidopsis* showed that both endosperm and embryo share gene expression patterns and biological processes during seed germination. About 10,800 transcripts (~84% of total genes expressed) are present in both endosperm and embryo. Endosperm-specific genes comprise about 415 genes that were highly expressed [17]. Of this gene set, 154 are cell wall-related genes, with most of them being expressed at the onset of testa rupture. Transcript abundance of several CWMPs shows a transient peak with a 6–24 h interval in tomato [51], *Arabidopsis* [17], and *Lepidium* seeds [16]. Although a change in transcript levels does not necessarily correlate to changes in protein abundance or enzymatic activity, it has been demonstrated that during *rice* germination, most cell wall-related transcripts, as well as the resulting metabolites from cell wall modification, accumulate about 12–24 h after imbibition (HAI, although some metabolites can be detected by 3 HAI) [23].

Cell wall modification can occur at five different stages during seed germination: (a) during the cellular expansion process triggered by rehydration of tissues, (b) at the onset of testa rupture, (c) during endosperm weakening and rupture, (d) during cellular expansion related to radicle elongation, and (e) during wall degradation and mobilization of stored reserves in both living and nonliving storage tissues.

3.4.1. Rehydration-driven cellular expansion

Seed imbibition is given by the difference in water potential between the seed and the environment. Nonviable seeds swell faster than viable seeds, as viable seeds develop turgor pressure that restricts further water uptake [2]. However, rapid imbibition can still occur and lead to solute leakage and damage of membranes. Gradual rehydration of seed tissues has been detected in legumes like peas and beans, where hydration starts in the tissues near to the micropyle. As water diffuses in the outermost tissues, a waterfront is formed between imbibed tissues and those about to be imbibed. The testa plays a significant role in modulating imbibition kinetics and the waterfront formation [2]. The seeds of mutants with altered testa structure or altered deposition of protecting substances (like flavonoids, cutin, suberin, and lignin) have increased permeability and lower longevity than the wild type [52]. Testa structure usually consists of several layers of highly compressed dead cells where protective substances are deposited during seed development and maturation. Plant cell walls of living cells can also function as an interface that modulates water intake by changing wall porosity, thus allowing a gradual swelling of all tissues. This regulation could be achieved by rapid changes in wall extensibility as the ones generated by expansins. In support of this view, in whole unstratified *Arabidopsis* seeds, the upregulation of *EXPA1, EXPA2, EXPA3, EXPA8, EXPA9, EXPA15,* and *EXPA20* transcripts from 0 to 12 h has been reported. This induction was evident in seeds imbibed in the light and during moist cold stratification at 4°C in the dark [11]. Also, in whole-seed transcriptomes, the expression of *AtXTH5, AtXTH6,* and *AtXTH33* transcripts

within the first 6HAI and *AtXTH3* and *AtXTH33* transcripts at 12HAI [24] was detected. The activity of XTHs from *chickpea* seeds has been detected in imbibed seeds from 1HAI until 24HAI, when the rate of radicle elongation slows down [53].

In many species from the Brassicaceae, Solanaceae, Linaceae, and Plantaginaceae, among others, the epidermis of the testa contains specialized cells that accumulate abundant pectins and heteroxylans, as well as some xyloglucans or arabinans during seed development. Upon imbibition these polysaccharides expand and burst out of the testa, generating a gel-like structure. This phenomenon, known as myxospermy, has been used as a model to study several hydrolases and PME activity [54, 55]. Although there is still uncertainty about the actual role of myxospermy, the proposed roles include regulating hydration, preventing desiccation, being an oxygen barrier, or allowing the seed to attach to the substrate and animals [54–56].

3.4.2. Testa rupture

In many plant species, testa rupture starts at the micropylar seed end. In tobacco (*Nicotiana tabacum*) seeds, the testa ruptures at the micropyle and follows predetermined breaking points due to the presence of channel-like structures underlying ridges in the testa [57]. In pea seeds the presence of xylogalacturonan in the inner walls of the testa was described, which coincides detached cells or with the junction sites between cells that are destined to detach, as described in other plant tissues [37]. In *Lepidium* seeds, HG composition shifts at the onset of testa rupture: while the seeds imbibe, non-esterified HG is ubiquitously distributed in all tissues, but by the time of testa rupture, this non-esterified HG is detected mostly on the endosperm and testa, meanwhile esterified HG is detected in the endosperm [38].

The seed testa is composed of several layers of nonliving cells, and thus the regulation and enzymes that facilitate cell separation in the testa must come from the living tissues underneath. At the onset of testa rupture (~25 HAI), it is possible to identify 90 cell wall-related genes from the 501 upregulated genes (~18%) in the micropylar endosperm and 58 from 282 genes (~20%) upregulated in the radicle. Also, about 8 (~8%) and 5 (~4%) genes were downregulated in both tissues, respectively [17]. In *Lepidium* seeds, tissue-specific transcript abundance patterns between the micropylar endosperm and the radicle accompany testa rupture, further supporting the view of this process as a decisive step in germination and in the regulation of cell wall-related genes [38]. In **Table 1**, these transcripts and its predicted biological/biochemical function are enlisted.

			Endosperm (HAI)				RA				Endosperm (HAI)				RA
Function	Gene ID	TS	0–12	16	25	31	25	Function	Gene ID	TS	0–12	16	25	31	25
EXPA10	AT1G26770	No	6–12		24		TR*	β-Gal	AT1G45130	EM	Not specified				
EXPA1	AT1G69530	No	3				TR*	β-Gal	AT5G08380	No	Not specified				
EXPA15	AT2G03090	No	12		24		TR*	β-Gal	AT1G77410	EN			24	31c	
EXPA6	AT2G28950	EM					TR	β-Gal	AT2G28470	No		16c	24		24

		Endosperm (HAI)				RA			Endosperm (HAI)				RA
EXPA4	AT2G39700	No	16				β-Gal	AT4G26140	No			31	
EXPA2	At5g05290	No	3		24		β-Glu	AT1G61820	EN	Not specified			
EXPA8	At2g40610	No	3–12		24		β-Glu	AT1G70710	No	16	TR*		
EXPA3	At2g37640	No	6–12		24		β-Glu	AT4G16260	EN	Not specified			
EXPA9	AT5G02260	No	12		24		β-Glu	AT1G26560	EM	Not specified			
EXPA20	AT4G38210	No	6–12		24		β-Glu	AT3G62750	No	3c			
EXPB1	At2g20750	No	12				β-Glu	AT2G44450	No		TR	31	
EXLA3	AT3G45960	No			TR		β-Glu	AT4G34480	No	12*			
EXLA1	AT3G45970	No		16	TR	TR	MAN5	AT4G28320	No	Not specified			
EXLA2	AT4G38400	No	6–12*		TR	TR	MAN6	AT5G01930	EN	Not specified			
EXT3	AT1G21310	EM	Not specified				MAN7	AT5G66460	No	6	24		
EXT10	AT5G06640	EM	Not specified				GH	AT5G49360	EN		TR		
EXTL	AT3G54590	EM	Not specified				GH	AT5G57560	No	16	TR		
EXTL	AT4G38770	EM	Not specified				GH	AT5G08370	No	16c	TR	31c	
EXTL	AT2G27380	RA	Not specified				GH	AT3G55430	No		TR	31	24*
XTH5	AT5G13870	No	6		24	24	GH	AT3G07320	No		24	31	24
XTH33	AT1G10550	No	12	16	TR	TR	GT	AT3G10320	EN			31c	
XTH	AT1G11545	No		16		TR	GH	AT3G13790	No	16		31	
XTH	AT1G32170	No			TR		GH	AT5G64570	No		24		TR
XTH17	AT1G65310	No			TR		GH-DUF3357	AT1G12240	No			31	TR
XTH	AT2G06850	No		16	TR	31 TR	GH	AT3G47010	EN		24		
XTH	AT2G36870	No				31	KOR2	AT1G65610	No			31	
XTH	AT3G23730	No			TR	TR	CESA5	AT5G09870	No		TR	31	24
XTH11	AT3G48580	No		16	24*	31	CSLC	AT4G07960	No		TR		TR
XTH	AT4G03210	No			TR	TR	AGP	AT3G11700	No		TR		TR
XTH	AT4G14130	No			TR	TR	AGP	AT5G44130	EN		TR		
XTH24	AT4G30270	EN		16	TR	TR	AGP	AT1G28290	No	16		31	
XTH18	AT4G30280	No		16	TR	TR	PRT	AT3G54400	No	16			24
XTH	AT4G30290	No		16		31	PRT	AT3G61820	No		TR*	31	24
XTH	AT4G37800	EM	Not specified				PRT	AT4G16563	No		TR		TR
XTH25	AT5G57550	No	6*		TR	TR	PL	AT3G24670	EM				24
XTR8	AT3G44990	EN		16c	24		PL	AT3G27400	No		TR		24
XTR6	AT4G25810	No			TR	TR	PL	AT4G13710	EN	Not specified			
PL	AT4G24780	No			TR*		PMEI	AT5G20740	No		TR		TR

http://dx.doi.org/10.5772/64085

			Endosperm (HAI)			RA				Endosperm (HAI)			RA
PL	AT5G48900	EM	Not specified				PMEI	AT5G46940	EN		TR		
PX	AT1G14540	EN		TR			PMEI	AT5G62340	No	16			TR
PX	AT1G14550	EN	Not specified				PMEI	AT5G64620	No			31c	
PX	AT1G30870	EM	Not specified				PG	AT3G59850	No	16	TR*	31	
PX	AT2G18980	No		TR	31		PG	AT3G61490	EM	Not specified			24
PX	AT2G43480	RA	Not specified				PG	AT4G23820	No	12	TR		24
PX	AT3G01190	EM	Not specified				Kinase	AT1G33590	No		TR	31	TR
PX	AT3G21770	EN			31c		Kinase	AT2G23770	No		TR		
PX	AT3G28200	No		TR	31		Kinase-DUF26	AT3G22060	MI	16		31	
PX	AT4G08770	EM	Not specified				Kinase	AT1G51940	No		TR		TR
PX	AT4G31760	EM	Not specified				LRR-p	AT4G26690	No		TR		TR
PX	AT5G05340	No			31		LRR-p	AT5G16590	No	16	TR	31	TR
PX	AT5G39580	No	3	TR			LRR-p	AT2G34930	No	16	TR	31	TR
PX	AT5G40150	EM	Not specified				DUF642	AT1G80240	EM	Not specified			
PX	AT5G64100	No	16	TR	31		DUF642	AT2G34510	EM	Not specified			
PX	AT5G64120	EN		TR	31		DUF642	AT2G41800	RA	Not specified			
PME	AT3G14310	No	6–12	24		24	DUF642	AT3G08030	No	3		31	
PME	AT1G04680	No			31	TR	DUF642	AT4G32460	No	16		31	
PME	AT1G57590	EM	Not specified			24	DUF642	AT5G11420	No	3		31	
PME	AT3G09410	EM	Not specified				DUF642	AT5G14150	EM	Not specified			
PME	AT3G10720	No	16	TR	31	TR	OX	AT1G62380	No	16c			
PME	AT3G62060	EM	Not specified			24	OX	AT1G76160	No		TR		
PME	AT4G19420	EN	Not specified				OX	AT2G46740	EN	16	TR	31	
PME	AT5G26670	RA	Not specified				OX	AT4G22010	No		TR		
PME	AT5G45280	No	3		31		OX	AT4G38420	No		TR		TR
PME	AT5G62330	EM	Not specified			24	OX	AT5G21105	No	16	TR	31	
PME2	AT1G02810	No	12	16	31		OX	AT5G44380	No		TR	31	
PME2	AT1G11580	No		16		24	PTRI	AT1G17860	No		TR	31	TR
PME2	AT2G26440	No		16	31	24	PTRI	AT2G38870	No			31	TR
PME2	AT3G47400	No			31	24	PTRI	AT4G22470	MI		TR	31	
PME2	AT3G49220	No		TR		TR	GT	AT1G64390	No	16	TR*	31	
PME2	AT4G02330	No	16	TR	31	TR	GT	AT2G02990	EN		TR		
PME2	AT4G33220	EM	Not specified				GT	AT2G14610	EN	Not specified			
PME2	AT5G64640	EM	12	24			GT	AT1G05170	EM				TR*

			Endosperm (HAI)		RA				Endosperm (HAI)		RA
PMEI	AT1G62770	RA	Not specified			GT	AT1G08280	No		31	
PMEI	AT2G47670	EM	3-12*	24*		PGIP	AT5G06860	No			TR
PMEI	AT4G00080	No		31	TR	PG	AT2G43860	MI	24	31	
PMEI	AT4G12390	EM	Not specified			PG	AT3G06770	No	TR		24

The endosperm expression profiles are subdivided by hours after imbibition (HAI) and testa rupture (TR, ~25HAI), whereas the radicle (Rad) expression only shows the moment of TR. *Abbreviations*: T-S, tissue-specific expression; EM, embryo; EN, endosperm; RA, radicle; MI, micropylar endosperm; EXPA, expansin; EXT, extensin; EXTL, extension-like protein; XTH, xyloglucan-transglycosylhydrolase; GH, glycosyl-hydrolase; β-Glu, β-glucosidase; β-Gal, β-galactosidase; AGP, arabinogalactan protein; CESA, cellulose synthase; PRT, protease; PL, pectate lyase; PX, peroxidase; CSLC, cellulose synthase-like; LRR-p, leucine-rich repeat protein; DUF, domain of unknown function; PME, pectin methylesterase; PME2, PME with an inhibitory domain; PMEI, PME inhibitor; PTRI, protease inhibitor; GT, transglycanase; PG, polygalacturonase; PGIP, PG inhibitor. The * means downregulation for that particular gene and time. The letter "c" beside a number means the expression was upregulated in the chalazal endosperm.

Table 1. Expression of some upregulated cell wall-modifying genes in *Arabidopsis* tissue-specific microarrays described by [17, 22] and expression profiles at http://bar.utoronto.ca/.

In *Arabidopsis* and *Lepidium* seeds, total PME activity increases gradually with imbibition time and peaks at the onset of testa rupture. ABA treatment does not affect testa rupture but endosperm rupture is delayed and PME activity fails to decrease following testa rupture [16, 38]. Two DUF642 genes, *BIIDXI* (*BDX*, *At4g32460*) and *At5g11420*, are expressed in the embryo and micropylar endosperm, respectively, during germination. In overexpression lines, testa and endosperm rupture of matrix-primed seeds occurred earlier compared to wild type. The germination performance of overexpression seeds was accompanied by an increase of total PME activity, compared to the wild type [50].

In non-endospermic seeds, testa rupture marks the end of germination. In this type of seeds, the testa rupture is accompanied by radicle elongation whose continued pressure in the inner face of the testa promotes cell separation [53].

3.4.3. Endosperm weakening for radicle protrusion

The endosperm functions as a barrier to control radicle protrusion as it can impose primary dormancy in many species like *Arabidopsis*, *Lepidium*, and yellow cedar, among others [2]. Endosperm structure of mature seeds varies greatly within species, where it can comprise one layer of cells as in *Arabidopsis*, *Lepidium*, and cucumber or to several layers as in tomato or tobacco [58, 59]. Structural studies in hard-seeded species like fenugreek (*Trigonella foenum-graecum*) and coffee (*Coffea arabica*) have demonstrated that near the micropyle a zone of thin-walled cells that can be a low-resistance area for radicle protrusion exists. Endosperm cell wall composition varies considerably among species: in the closely related species, *Arabidopsis* and *Lepidium* are rich in cellulose, non-esterified HG, arabinans, and XG. However, these polysaccharides are not uniformly distributed in the endosperm: in *Lepidium*, an epitope for arabinans (LM13) was localized in the inner and outer walls of the cells, but absent from the traverse walls. The endosperm of tomato seed contains mannans that have been shown to contribute

to the control of radicle protrusion and general endosperm hardness rather than a storage function [39].

The hydrolytic activity of β-glucanases or endo-β-mannanases can contribute to endosperm cell wall weakening in Brassicaceae and Solanaceae species, which have cell walls rich in mannans [59]. In *Arabidopsis* seeds the activity of *AtMAN5*, *AtMAN6*, and *AtMAN7* and regulation of their activity by the basic leucine-zipper 44 transcription factor, *AtbZIP44*, whose knockout mutants have delayed germination have been described [60]. In hard-seeded species, there was a negligible activity in the radicle and micropylar endosperm of endo-β-mannanases that could not be associated with endosperm weakening to allow radicle protrusion [61]. However it is still unexplored the activity of other CWMPs that could contribute to wall rearrangements during germination in these species. The upregulation of wall-related genes in the endosperm has been reported; this includes α-expansins (*AtEXPA2*, *AtEXPA8*, and *AtEXPA9*), β-expansin (*AtEXPB1*), expansin-like protein (*AtEXPL1*), cellulose synthase-like proteins (*CSLA2* and *CSLC4*), xyloglucan endotransglycosylases (*AtXTH11*, *AtXTH17*, *AtXTH18*, *AtXTH33*, *AtXTH31*, *AtXTH23*, and *AtXTH24*), and mannanase (*AtMAN7*) [62].

Several reports indicate that endosperm weakening and rupture are inhibited or delayed by ABA, in a dose-dependent manner, in some species of the Brassicaceae family [8, 16] and tobacco [63]. Microarray analysis of *Arabidopsis* seeds treated with exogenous ABA at the onset of radicle protrusion has shown a downregulation of several wall-related transcripts in the endosperm, including PMEs, AGPs, and PLs [7]. In tobacco, ABA delays the accumulation and activity of β-1,3-glucanase in the micropyle before radicle protrusion [63]. PMEs contribute to seed germination in several species by modulating the degree of methylesterification of pectins in the endosperm [16, 38, 64, 65]. In yellow cedar a loss of the internal structure of the megagametophyte surrounding the radicle during germination was described. The resulting decrease in the mechanical strength of the megagametophyte would allow radicle protrusion. There is a positive correlation between dormancy alleviation and PME activity, as well as with germination performance and PME activity in both the megagametophyte and the embryo. PME activity has also been demonstrated to positively correlate with germination performance in *Arabidopsis* [64]. The *LeXET4* gene transcripts, which are restricted to the endosperm cap, are detected in tomato seeds within 12 h of imbibition and reach a maximum at 24 h [66]; they decline after radicle emergence despite a continued degradation of the lateral endosperm cell walls. PG *At2g43860* is expressed within the endosperm cells of the seed adjacent to the site of the emerging radicle [67]. In germinating tomato seeds, several reports indicate the expression of expansins [51], PGs (*LeXPG1*) [68], and XTHs (*LeXET4*), which accumulate in the endosperm region adjacent to the expanding radicle (~40HAI). Transcripts are detected within 12HAI and generally peak by 24HAI, consistent with the endosperm weakening.

3.4.4. Embryonic axis elongation

Cell elongation, rather than cellular division, is the main process that drives embryo growth [14]. Cell division occurs after radicle protrusion and contributes to the rapid growth of the embryonic axis by generating new elongating cells [2]. Cell elongation that drives radicle

protrusion occurs at the transition zone, which comprises the cells between the last proximal root hair cell in the radicle and the lower basal cells of the hypocotyl [69]. In *Arabidopsis* Col seeds that have been previously stratified, the radicle protrusion can initiate as early as 32 HAI. By this time, and immediately prior to the radicle emergence through the endosperm, the cells in the transition zone had incremented their size by 44% while the cells in the radicle 10% and in the hypocotyl 30%. By 40HAI, the radicle has already protruded and the elongated cells in the seedling have increased their size by 15% in the radicle, 52% in the hypocotyl, and 108% in the transition zone. Elongation is often accompanied by an increase in DNA content without subsequent mitosis (endoreduplication) [69].

In the micropylar endosperm of tomato seeds, an important mobilization of protein bodies occurs, but it seems that there is no cell degradation as the radicle protrudes. Instead, a process similar to cell separation to allow radical protrusion was suggested [70]. A similar process was observed in celery seeds, where the radicle tip also seems to penetrate the micropylar endosperm by separating the endosperm cells, but, since the embryo needs to grow before germination is completed, cell degradation for storage mobilization occurs in the endosperm adjacent to the embryo [27]. The expression of *LeEXP8* and *LePG1* in the embryo elongation zone of tomato seeds has been reported at the onset of radicle protrusion [51, 68].

3.4.5. Cell wall participation in the mobilization of stored reserves

Major reserve mobilization occurs once germination has concluded, and these reserves are utilized to feed the growing seedling rather than to fuel radicle protrusion. However, in cereal grains, the preparation for starch and oligosaccharide mobilization occurs within the first hours of germination [15]. In cereals, the endosperm is a nonliving storage tissue, and the endosperm cell walls protect its contents from enzymatic attack. Accordingly, the degradation of cell walls is a limiting step in storage reserve mobilization that is induced by the GA produced by the embryo (at the scutellum) and secreted to the aleurone layer [2].

In most endospermic seeds this tissue is still living. Mannans in the endosperm cell walls of date palm (*Phoenix dactylifera*) and coffee are mobilized to support embryo development. It has been proposed that the mobilization of storage xyloglucans can be coupled to the growth rate of the seedling by transglycosylation. In legumes, endosperm galactomannans seems to function as reserves; they can constitute up to 30% of total seed dry weight. In fenugreek and *Schizolobium parahyba*, the cell walls of the endosperm are thickened with galactomannan and in some cases the cytoplasm is nonexistent. In *Tamarindus indica* and *Hymenaea courbaril*, reserve xyloglucans are stored between two primary walls and are degraded without hydrolyzing both walls [39]. During germination of celery seeds, the surrounding endosperm degrades leaving a small amount of un-degraded polymers of the cell wall, except for the micropylar endosperm, in which only some protein bodies are mobilized and the rest of the cells persist until the radicle pushes through; once radicle protrusion has started, these micropylar cells are degraded [27].

4. Concluding remarks

In-depth temporal screening of cell wall-related transcripts and proteins has provided an important overview of the possible actors involved in the five stages during germination where wall modification is involved, as described above. In *rice*, a valuable integrative effort using -omics approaches has been done to understand seed germination [23]. This analysis suggested that the changes in transcript levels during early germination (3–12 HAI) drive the subsequent changes in the metabolome (12–24 HAI) of germinating seeds, supporting that most of the changes observed at the transcriptional level are related to the cellular processes involved in germination. Other authors have associated transcript abundance with specific seed compartments and some enzymatic activity assays, demonstrating the relevance of understanding tissue-specific expression profiles [17, 38]. Much of the information available related to CWMPs still needs to be validated through enzymatic activity or in vivo interaction assays. Only about 121 (~12%) of the total cell wall-related genes are experimentally validated [48]. Also, many cell wall-related proteins belong to families of unknown function. The -omics approach can be useful to propose hypothesis of wall-modification complexes, whose activity could be regulated at several levels, and coordinated by unknown function proteins that could act as scaffolding proteins and direct this complex activity to specific polysaccharides. Cosgrove [44] proposed that CWMPs could be functionally classified into primary or secondary modifiers, but this idea has not being reflected in other studies. Following Cosgrove, the analysis of cell wall modification considering an alternating activity of primary or secondary modifiers could facilitate the understanding of the dynamics of cell wall modification during seed germination. For instance, expansins could be primary modifiers as they affect cell wall loosening and extensibility, but they do not remove or transfer polysaccharides into the wall during imbibition; other primary modifiers could be PMEs, as their activity precedes PGs and promotes cell expansion or cell separation, or the resulting exposed GAL residues can be cross-linked with Ca^{2+} and promote wall stiffening. Secondary modifiers would include GHs and GTs that would act on exposed residues either promoting cell expansion, separation, or stiffening. In assays to study mucilage properties, sequential treatment with different hydrolases allows solubilization of other components, which are masked to the activity of other enzymes [55, 71]. The alternate perspective of primary and secondary modifiers could help in identifying potential interactions in silico and tested in vivo.

Spatial transcriptomic analyses that include the different seed compartments and the analysis of cell wall composition changes using specific antibodies for in situ localization of the different polysaccharide epitopes in seed tissues provide valuable information. Although *Arabidopsis* is the best-known plant model, several authors demonstrate that comparative analysis allows higher resolution of tissue-specific cell wall microdomains that are not achievable in *Arabidopsis* [8, 10, 60]. As an example, Lee et al. [59] describe the presence of LM13 epitopes in the inner and outer cell walls, but absent in the transverse cell walls of the endosperm in *Lepidium* seeds; in tobacco, which has a thicker endosperm than *Arabidopsis* or *Lepidium*, XGs were abundant in the embryo, and at the micropyle (rich in heteromannans), these polysaccharides were only present in the middle lamella and intercellular regions. Thus, the analysis of cell wall-modification processes would benefit from the multispecies compar-

ison of in situ localization of polysaccharide epitopes in seed tissues. The characterization of wall microdomains could be combined with the valuable information generated by -omics technologies, to propose new hypothesis of regulation and coordinated activity of CWMPs. Ultimately, the activity of these CWMPs must be confirmed by in situ localization, in vivo protein interactions, and enzymatic activity. By combining the resources available for model species with the selection of other plant systems with bigger-easy-to-handle seeds, it could be possible to achieve a comprehensive view of seed-compartment functions and regulation during germination. The endosperm role during germination is fundamental in endospermic seeds; however, in non-endospermic seeds, this role must befall on either the embryo or the testa. Since the testa is a nonliving tissue, the radicle most certainly assumes part of this regulatory role, but a comparative analysis is needed to ascertain this supposition and to determine if some of the endosperm functions are developed by the testa while still in the maturation program. The occurrence of endospermic and non-endospermic seeds within the same taxa is relatively common in legumes such as *soybean*, which could offer a model for analyzing transcriptomic differences within embryo compartments comparable to the differences described between the endosperm and the radicle.

ROS participation in germination is supported by several reports and transcriptomic profiles of germinating seeds [9, 14, 17]. However, the actual role of ROS and ROS-related enzymes in promoting cell wall loosening needs to be further analyzed, since the physiological concentrations of ROS during germination do not seem to be sufficient to induce wall extension, and attempts of increasing ROS concentration lead to wall breakage [5]. Müller et al. [72] describe abnormal rupture of the micropylar endosperm of *Lepidium* seeds treated with H_2O_2, while the treatment with myrigalone A [73], which inhibits the hormone-mediated accumulation of ROS during germination, also induces abnormal endosperm breakage. These observations further support the notion of ROS as a signaling agent that induces downstream activation of CWMPs than inducing wall loosening on its own.

Acknowledgements

This work was supported by Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) IN207915 and SEP-CONACyT grant 155074. Ximena Gómez-Maqueo acknowledges the scholarship by the Consejo Nacional de Ciencia y Tecnología (CONACyT).

Author details

Ximena Gómez-Maqueo and Alicia Gamboa-deBuen*

*Address all correspondence to: agamboa@ecologia.unam.mx

Ecological Physiology Lab., Ecology Institute, National Autonomous University of Mexico, UNAM, Mexico City, Mexico

References

[1] Baskin CC, Baskin JM. Seeds: ecology, biogeography, and evolution of dormancy and germination. San Diego, Elsevier; 1998. 666 p. ISBN-13: 978-0120802609

[2] Bewley JD, Bradford KJ, Hilhorst HWM, Nonogaki H. Seeds, physiology of development, germination and dormancy. 3rd ed. New York, Springer; 2013. 392p. DOI 10.1007/978-1-4614-4693-4.

[3] Fenner M, Thompson K. The ecology of seeds. Cambridge, Cambridge University Press; 2005. 250p.

[4] Taiz L, Zeiger E. Plant physiology. 5th Ed. Sunderland: Sinauer Assoc; 2010. 782p. ISBN-13: 978-0878938667

[5] Cosgrove DJ. Growth of the plant cell wall. Nature Reviews Molecular Cell Biology. 2005;6:850-861. DOI: 10.1038/nrm1746

[6] Donohue K, Dorn L, Griffith C, Kim E, Aguilera A, Polisetty CR, Schmitt J. The evolutionary ecology of seed germination of *Arabidopsis thaliana*: variable natural selection on germination timing. Evolution. 2005;59:758-770. DOI: 10.1111/j.0014-3820.2005.tb01751.x

[7] Penfield S, Li Y, Gilday AD, Graham S, Graham IA. Arabidopsis ABA INSENSITIVE4 regulates lipid mobilization in the embryo and reveals repression of seed germination by the endosperm. The Plant Cell. 2006;18:1887-1899. DOI: http://dx.doi.org/10.1105/tpc.106.041277

[8] Müller K, Tintelnot S, Leubner-Metzger G. Endosperm-limited Brassicaceae seed germination: abscisic acid inhibits embryo-induced endosperm weakening of *Lepidium sativum* (cress) and endosperm rupture of cress and *Arabidopsis thaliana*. Plant and Cell Physiology. 2006;47:864-877. DOI: 10.1093/pcp/pcj059

[9] Linkies A, Leubner-Metzger G. Beyond gibberellins and abscisic acid: how ethylene and jasmonates control seed germination. Plant Cell Reports. 2012;31:253-270. DOI: 10.1007/s00299-011-1180-1

[10] Linkies A, Müller K, Morris K, Turečková V, Wenk M, Cadman CS, Corbineau F, Strnad M, Lynn JR, Finch-Savage WE, Leubner-Metzger G. Ethylene interacts with abscisic acid to regulate endosperm rupture during germination: a comparative approach using *Lepidium sativum* and *Arabidopsis thaliana*. The Plant Cell. 2009;21:3803-3822. DOI: http://dx.doi.org/10.1105/tpc.109.070201

[11] Kanai M, Nishimura M, Hayashi M. A peroxisomal ABC transporter promotes seed germination by inducing pectin degradation under the control of ABI5. The Plant Journal. 2010;62:936-947. DOI: 10.1111/j.1365-313X.2010.04205.x

[12] Bewley JD. Seed germination and dormancy. The Plant Cell. 1997;9(7):1055. Plant Cell. 1997;9:1055–1066. DOI: 10.1105/tpc.9.7.1055

[13] Nonogaki H, Bassel GW, Bewley JD. Germination—still a mystery. Plant Science. 2010;179:574-581. DOI: 10.1016/j.plantsci.2010.02.010

[14] Weitbrecht K, Müller K, Leubner-Metzger G. First off the mark: early seed germination. Journal of Experimental Botany. 2011;62:3289-3309. DOI: 10.1093/jxb/err030

[15] Rosental L, Nonogaki H, Fait A. Activation and regulation of primary metabolism during seed germination. Seed Science Research. 2014;24:1-5. DOI: http://dx.doi.org/10.1017/S0960258513000391

[16] Müller K, Levesque-Tremblay G, Bartels S, Weitbrecht K, Wormit A, Usadel B, Haughn G, Kermode AR. Demethylesterification of cell wall pectins in Arabidopsis plays a role in seed germination. Plant Physiology. 2013;161:305-316. DOI: http://dx.doi.org/10.1104/pp.112.205724

[17] Dekkers BJ, Pearce S, van Bolderen-Veldkamp RP, Marshall A, Widera P, Gilbert J, Drost HG, Bassel GW, Müller K, King JR, Wood AT. Transcriptional dynamics of two seed compartments with opposing roles in Arabidopsis seed germination. Plant Physiology. 2013;163:205-215. DOI: http://dx.doi.org/10.1104/pp.113.223511

[18] Yan D, Duermeyer L, Leoveanu C, Nambara E. The functions of the endosperm during seed germination. Plant Cell and Physiology. 2014;55:1521-1533. DOI: 10.1093/pcp/pcu089

[19] Narsai R, Law SR, Carrie C, Xu L, Whelan J. In-depth temporal transcriptome profiling reveals a crucial developmental switch with roles for RNA processing and organelle metabolism that are essential for germination in Arabidopsis. Plant Physiology. 2011;157:1342-1362. DOI: http://dx.doi.org/10.1104/pp.111.183129

[20] Jia H, Suzuki M, McCarty DR. Regulation of the seed to seedling developmental phase transition by the LAFL and VAL transcription factor networks. Wiley Interdisciplinary Reviews: Developmental Biology. 2014;3:135-145. DOI: 10.1002/wdev.126

[21] Kermode AR, Finch-Savage BE. Desiccation sensitivity in orthodox and recalcitrant seeds in relation to development. In: Black M, Pritchard HW, editors. Desiccation and survival in plants: drying without dying. New York: Cabi; 2002; p:149-184

[22] Nakabayashi K, Okamoto M, Koshiba T, Kamiya Y, Nambara E. Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: epigenetic and genetic regulation of transcription in seed. The Plant Journal. 2005;41:697-709. DOI: 10.1111/j.1365-313X.2005.02337.x

[23] Howell KA, Narsai R, Carroll A, Ivanova A, Lohse M, Usadel B, Millar AH, Whelan J. Mapping metabolic and transcript temporal switches during germination in rice highlights specific transcription factors and the role of RNA instability in the germi-

nation process. Plant Physiology. 2009;149:961-980. DOI: http://dx.doi.org/10.1104/pp.108.129874

[24] Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S. Gibberellin biosynthesis and response during Arabidopsis seed germination. The Plant Cell. 2003;15:1591-1604. DOI: http://dx.doi.org/10.1105/tpc.011650

[25] Silverstone AL, Chang CW, Krol E, Sun TP. Developmental regulation of the gibberellin biosynthetic gene GA1 in *Arabidopsis thaliana*. The Plant Journal. 1997;12:9-19. DOI: 10.1046/j.1365-313X.1997.12010009.x

[26] Groot SP, Karssen CM. Gibberellins regulate seed germination in tomato by endosperm weakening: a study with gibberellin-deficient mutants. Planta. 1987;171:525-531.

[27] Jacobsen JV, Pressman E. A structural study of germination in celery (*Apium graveolens* L.) seed with emphasis on endosperm breakdown. Planta. 1979;144:241-248.

[28] Pelloux J, Rusterucci C, Mellerowicz EJ. New insights into pectin methylesterase structure and function. Trends in Plant Science. 2007;12:267-277. DOI: 10.1016/j.tplants.2007.04.001

[29] Caffall KH, Mohnen D. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohydrate Research. 2009;344:1879-1900. DOI: 10.1016/j.carres.2009.05.021

[30] Franková L, Fry SC. Biochemistry and physiological roles of enzymes that 'cut and paste' plant cell-wall polysaccharides. Journal of Experimental Botany. 2013;64:3519–3550. DOI: 10.1093/jxb/ert201

[31] Levesque-Tremblay G Pelloux J, Braybrook SA. Tuning of pectin methylesterification: consequences for cell wall biomechanics and development. Planta. 2015; 242(4):791–811. DOI 10.1007/s00425-015-2358-5

[32] Carpita NC. Structure and biogenesis of the cell walls of grasses. Annual Review of Plant Biology. 1996;47:445-476. DOI: 10.1146/annurev.arplant.47.1.445

[33] Yokoyama R, Nishitani K. Genomic basis for cell-wall diversity in plants. A comparative approach to gene families in rice and Arabidopsis. Plant and Cell Physiology. 2004;45:1111-1121. DOI: 10.1093/pcp/pch151

[34] Reiter W. Biosynthesis and properties of the plant cell wall. Current Opinion in Plant Biology. 2002;5:536-542. DOI: 10.1016/S1369-5266(02)00306-0

[35] Bashline L, Li S, Gu Y. The trafficking of the cellulose synthase complex in higher plants. Annals of Botany. 2014;114:1059-1067. DOI: 10.1093/aob/mcu040

[36] Mohnen, D. Pectin structure and biosynthesis. Current Opinion in Plant biology. 2008;11:266-277. DOI: 10.1016/j.pbi.2008.03.006

[37] Willats WG, McCartney L, Steele-King CG, Marcus SE, Mort A, Huisman M, van Alebeek GJ, Schols HA, Voragen AG, Le Goff A, Bonnin E. A xylogalacturonan epitope

is specifically associated with plant cell detachment. Planta. 2004;218:673-681. DOI: 10.1007/s00425-003-1147-8

[38] Scheler C, Weitbrecht K, Pearce SP, Hampstead A, Büttner-Mainik A, Lee KJ, Voegele A, Oracz K, Dekkers BJ, Wang X, Wood AT. Promotion of testa rupture during garden cress germination involves seed compartment-specific expression and activity of pectin methylesterases. Plant Physiology. 2015;167:200-215. DOI: 10.1104/pp.114.247429

[39] Buckeridge MS. Seed cell wall storage polysaccharides: models to understand cell wall biosynthesis and degradation. Plant Physiology. 2010;154:1017-1023. DOI: 10.1104/pp.110.158642

[40] Watson BS, Lei Z, Dixon RA, Sumner LW. Proteomics of *Medicago sativa* cell walls. Phytochemistry. 2004;65:1709-1720. DOI: 10.1016/j.phytochem.2004.04.026

[41] Showalter AM. Structure and function of plant cell wall proteins. The Plant Cell. 1993;5:9-23. DOI: 10.1105/tpc.5.1.9

[42] Jamet E, Canut H, Boudart G, Pont-Lezica RF. Cell wall proteins: a new insight through proteomics. Trends in Plant Science. 2006;11:33-39. DOI: 10.1016/j.tplants.2005.11.006

[43] Dubreucq B, Berger N, Vincent E, Boisson M, Pelletier G, Caboche M, Lepiniec L. The Arabidopsis AtEPR1 extensin-like gene is specifically expressed in endosperm during seed germination. The Plant Journal. 2000;23:643-652. DOI: 10.1046/j.1365-313x.2000.00829.x

[44] Cosgrove DJ. Enzymes and other agents that enhance cell wall extensibility. Annual Review of Plant Biology. 1999;50:391-417. DOI: 10.1146/annurev.arplant.50.1.391

[45] Marín-Rodríguez MC, Orchard J, Seymour GB. Pectate lyases, cell wall degradation and fruit softening. Journal of Experimental Botany. 2002;53:2115-2109. DOI: 10.1093/jxb/erf089

[46] Wolf S, Mouille G. Pelloux J. Homogalacturonan methyl-esterification and plant development. Molecular Plant. 2009;2:851-860. DOI: 10.1093/mp/ssp066

[47] Bordenave M, Goldberg R. Immobilized and free apoplastic pectinmethylesterases in mung bean hypocotyl. Plant Physiology. 1994;106:1151-1160. DOI: http:/dx.doi.org/10.1104/pp.106.3.1151

[48] Mewalal R, Mizrachi E, Mansfield SD, Myburg AA. Cell wall-related proteins of unknown function: missing links in plant cell wall development. Plant and Cell Physiology. 2014;55:1031-1043. DOI: 10.1093/pcp/pcu050

[49] Manabe Y, Nafisi M, Verhertbruggen Y, Orfila C, Gille S, Rautengarten C, Knox JP. Loss-of-function mutation of REDUCED WALL ACETYLATION2 in Arabidopsis leads to reduced cell wall acetylation and increased resistance to *Botrytis cinerea*. Plant Physiology. 2011;155:1068-1078. DOI: http://dx.doi.org/10.1104/pp.110.168989

[50] Zúñiga-Sánchez E, Soriano D, Martínez-Barajas E, Orozco-Segovia A, Gamboa-deBuen A. *BIIDXI*, the *At4g32460* DUF642 gene, is involved in pectin methyl esterase regula-

tion during *Arabidopsis thaliana* seed germination and plant development. BMC Plant Biology. 2014;14:338. DOI: 10.1186/s12870-014-0338-8

[51] Chen F, Bradford KJ. Expression of an expansin is associated with endosperm weakening during tomato seed germination. Plant Physiology. 2000;124:1265-1274. DOI: http://dx.doi.org/10.1104/pp.124.3.1265

[52] Sano N, Rajjou L, North HM, Debeaujon I, Marion-Poll A, Seo M. Staying alive: molecular aspects of seed longevity. Plant and Cell Physiology. 2015 Dec 3:pcv186. DOI: 10.1093/pcp/pcv186.

[53] Hernández-Nistal J, Labrador E, Martín I, Jiménez T, Dopico B. Transcriptional profiling of cell wall protein genes in chickpea embryonic axes during germination and growth. Plant Physiology and Biochemistry. 2006;44:684-692. DOI: 10.1016/j.plaphy.2006.10.017

[54] Saez-Aguayo S, Ralet MC, Berger A, Botran L, Ropartz D, Marion-Poll A, North HM. PECTIN METHYLESTERASE INHIBITOR6 promotes Arabidopsis mucilage release by limiting methylesterification of homogalacturonan in seed coat epidermal cells. The Plant Cell. 2013;25:308-323. DOI: http://dx.doi.org/10.1105/tpc.112.106575

[55] Huang D, Wang C, Yuan J, Cao J, Lan H. Differentiation of the seed coat and composition of the mucilage of *Lepidium perfoliatum* L.: a desert annual with typical myxospermy. Acta Biochimica et Biophysica Sinica. 2015;47:775-787. DOI: 10.1093/abbs/gmv078

[56] Kreitschitz A, Vallès J. Achene morphology and slime structure in some taxa of Artemisia L. and Neopallasia L.(Asteraceae). Flora-Morphology, Distribution, Functional Ecology of Plants. 2007;202:570-580. DOI: 10.1016/j.flora.2006.12.003

[57] Leubner-Metzger G. Functions and regulation of β-1,3-glucanases during seed germination, dormancy release and after-ripening. Seed Science Research. 2003;13:17-34. DOI: http://dx.doi.org/10.1079/SSR2002121

[58] Salanenka YA, Goffinet MC, Taylor AG. Structure and histochemistry of the micropylar and chalazal regions of the perisperm–endosperm envelope of cucumber seeds associated with solute permeability and germination. Journal of the American Society for Horticultural Science. 2009;134:479-487.

[59] Lee KJ, Dekkers BJ, Steinbrecher T, Walsh CT, Bacic A, Bentsink L, Leubner-Metzger G, Knox JP. Distinct cell wall architectures in seed endosperms in representatives of the Brassicaceae and Solanaceae. Plant Physiology. 2012;160:1551-1566. DOI: http://dx.doi.org/10.1104/pp.112.203661

[60] Iglesias-Fernández R, Rodríguez-Gacio MC, Barrero-Sicilia C, Carbonero P, Matilla A. Three endo-β-mannanase genes expressed in the micropylar endosperm and in the radicle influence germination of *Arabidopsis thaliana* seeds. Planta. 2011;233:25-36. DOI: 10.1007/S00425-010-1257-z

[61] Gong X, Bassel GW, Wang A, Greenwood JS, Bewley JD. The emergence of embryos from hard seeds is related to the structure of the cell walls of the micropylar endosperm, and not to endo-β-mannanase activity. Annals of Botany. 2005;96:1165-1173. DOI: 10.1093/aob/mci269

[62] Endo A, Tatematsu K, Hanada K, Duermeyer L, Okamoto M, Yonekura-Sakakibara K, Saito K, Toyoda T, Kawakami N, Kamiya Y, Seki M. Tissue-specific transcriptome analysis reveals cell wall metabolism, flavonol biosynthesis and defense responses are activated in the endosperm of germinating *Arabidopsis thaliana* seeds. Plant and Cell Physiology. 2012;53:16-27. DOI: 10.1093/pcp/pcr171

[63] Leubner-Metzger G, Frundt C, Vogeli-Lange R, Meins Jr F. Class I [β]-1, 3-Glucanases in the endosperm of tobacco during germination. Plant Physiology. 1995;109:751-759. DOI: http://dx.doi.org/10.1104/pp.109.3.751

[64] Ren C, Kermode AR. An increase in pectin methyl esterase activity accompanies dormancy breakage and germination of yellow cedar seeds. Plant Physiology. 2000;124:231-242. DOI: http://dx.doi.org/10.1104/pp.124.1.231

[65] Martínez-Andújar C, Pluskota WE, Bassel GW, Asahina M, Pupel P, Nguyen TT, Takeda-Kamiya N, Toubiana D, Bai B, Górecki RJ, Fait A. Mechanisms of hormonal regulation of endosperm cap-specific gene expression in tomato seeds. The Plant Journal. 2012;71:575-86. DOI: 10.1111/j.1365-313X.2012.05010.x

[66] Chen F, Nonogaki H, Bradford KJ. A gibberellin-regulated xyloglucan endotransglycosylase gene is expressed in the endosperm cap during tomato seed germination. Journal of Experimental Botany. 2000;53:215-223. DOI: 10.1093/jexbot/53.367.215

[67] González-Carranza ZH, Elliott KA, Roberts JA. Expression of polygalacturonases and evidence to support their role during cell separation processes in *Arabidopsis thaliana*. Journal of Experimental Botany. 2007;58:3719-3730. DOI: 10.1093/jxb/erm222

[68] Sitrit Y, Hadfield KA, Bennett AB, Bradford KJ, Downie AB. Expression of a polygalacturonase associated with tomato seed germination. Plant Physiology. 1999;121:419-428. DOI: http://dx.doi.org/10.1104/pp.121.2.419

[69] Sliwinska E, Bassel GW, Bewley JD. Germination of *Arabidopsis thaliana* seeds is not completed as a result of elongation of the radicle but of the adjacent transition zone and lower hypocotyl. Journal of Experimental Botany. 2009;60:3587-3594. DOI: 10.1093/jxb/erp203

[70] Karssen CM, Haigh A, Van der Toorn P, Weges R. Physiological mechanisms involved in seed priming. In: Taylorson, R.B. Recent advances in the development and germination of seeds. 1989 (pp. 269-280). New York, Springer US.

[71] North HM, Berger A, Saez-Aguayo S, Ralet MC. Understanding polysaccharide production and properties using seed coat mutants: future perspectives for the

exploitation of natural variants. Annals of Botany. 2014;114:1251-1263. DOI: 10.1093/aob/mcu011

[72] Müller K, Linkies A, Vreeburg RA, Fry SC, Krieger-Liszkay A, Leubner-Metzger G. In vivo cell wall loosening by hydroxyl radicals during cress seed germination and elongation growth. Plant Physiology. 2009;150:1855-65. DOI: http://dx.doi.org/10.1104/pp.109.139204

[73] Oracz K, Voegele A, Tarkowská D, Jacquemoud D, Turečková V, Urbanová T, Strnad M, Sliwinska E, Leubner-Metzger G. Myrigalone A inhibits *Lepidium sativum* seed germination by interference with gibberellin metabolism and apoplastic superoxide production required for embryo extension growth and endosperm rupture. Plant and Cell Physiology. 2012;53:81-95. DOI: 10.1093/pcp/pcr124

Resultados

Actividad relativa de PME

Se registró un cambio en la actividad relativa de los extractos de proteína total obtenidos de las semillas de *C. aesculifolia*. Este cambio en la actividad relativa estuvo asociado con el estadio fisiológico observado y el lote analizado. El acondicionamiento no tuvo un efecto significativo en la actividad, independientemente del lote o estadio fisiológico observado (figura 3). Durante la germinación, la actividad de las PME aumentó paulatinamente, hasta llegar a un pico máximo de actividad en etapas intermedias del proceso; posteriormente hubo una disminución en la actividad hacia el momento de protrusión de la radícula. El estadio fisiológico en el cual se observa el pico de actividad es dependiente de la identidad del lote: en el caso del lote de 2014 el pico ocurrió en el momento de ruptura de testa, mientras que para los dos lotes colectados en 2015 el pico se observó en el 50% CRA.

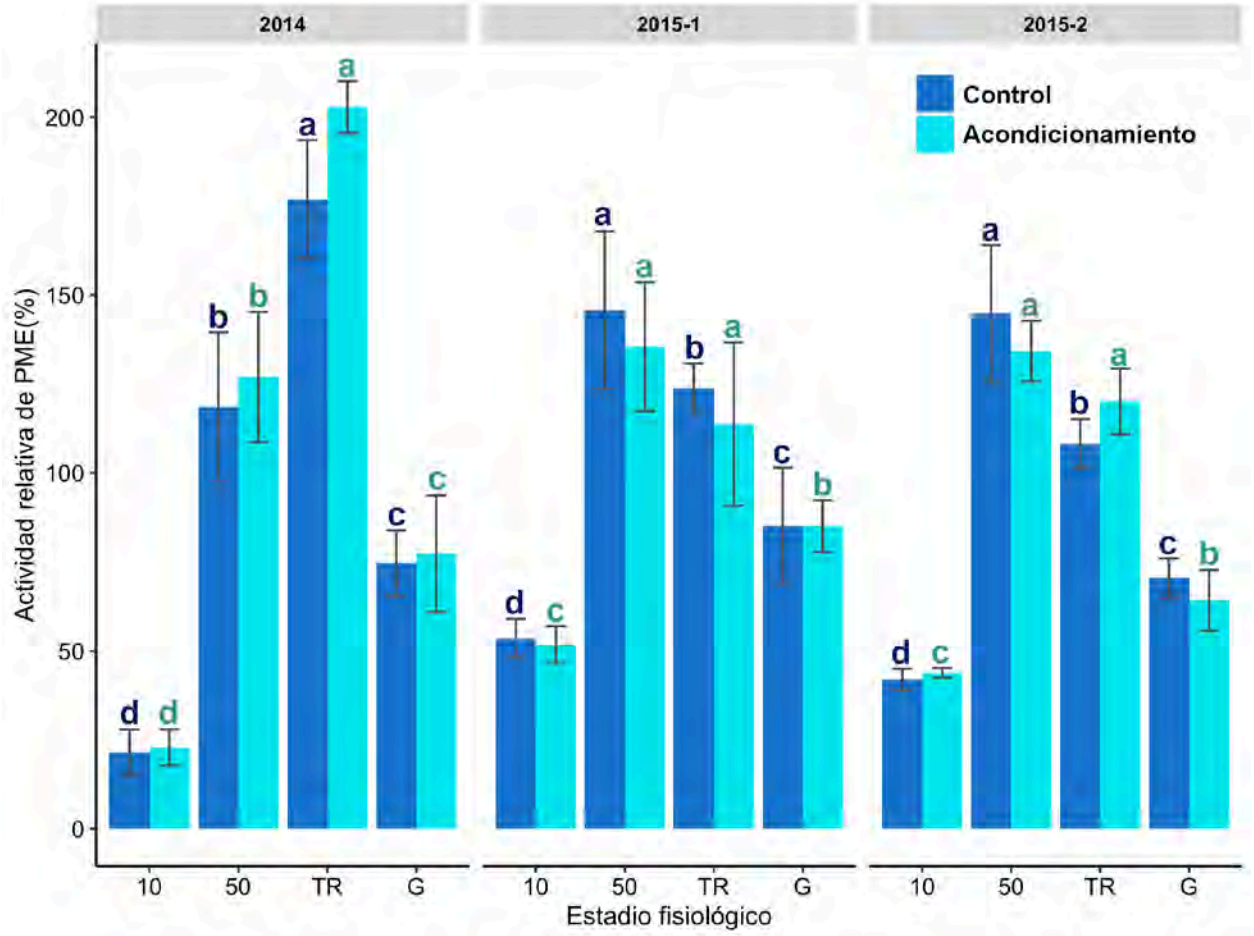


Figura 3. Actividad relativa de PME observada en los lotes 2014, 2015-1 y 2015-2. Los estadios fisiológicos reportados corresponden a 10 y 50 % de CRA, al momento de ruptura de testa (TR), y al momento de la germinación (G). Las letras azules y turquesas indican diferencias significativas entre estadios fisiológicos del grupo control y con acondicionamiento mátrico, respectivamente. No se detectaron diferencias en la actividad debidas al acondicionamiento.

Discusión

La pared celular es una estructura dinámica que se modifica de manera continua durante el desarrollo y crecimiento de las plantas (Levesque-Tremblay *et al.*, 2015; Cosgrove, 2016). Uno de los componentes más abundantes en la pared celular primaria de las dicotiledóneas es un grupo de polisacáridos complejos, conocidos colectivamente como pectinas (Levesque-Tremblay *et al.*, 2015). Las modificaciones de las pectinas en la pared celular, mediadas por las PME, le confieren la flexibilidad y/o rigidez necesaria para permitir la entrada de agua y dirigir el crecimiento celular durante la germinación (Scheler *et al.*, 2015).

La dinámica de la PME observada en las semillas de *C. aesculifolia* concuerda con estudios previo en *Arabidopsis* y *Lepidium sativum*, donde se demostró que existe un patrón de incremento en la actividad de PME en etapas intermedias del proceso de gemrinación, seguido de una disminución de la actividad hacia el momento de protrusión de la radícula (Müller *et al.*, 2013; Scheler *et al.*, 2015). En el caso de *Arabidosis*, el pico máximo de actividad de PME se registró en el momento fisiológico de ruptura de testa (Müller *et al.*, 2013), mientras que en *L. sativum*, la actividad de las PME aumenta paulatinamente en el endospermo micropilar, llegando a un máximo de actividad antes de la ruptura de testa, para después disminuir y mantenerse estable hacia el momento de protrusión de la radícula (Scheler *et al.*, 2015).

La actividad de las PME registrada en el lote colectado en 2014 es similar al patrón reportado para *Arabidopsis*, mientras que los lotes 2015-1 y 2015-2 se registró el pico de actividad en un estadio fisiológico previo a la ruptura de testa, comportamiento similar al reportado para *L. sativum*. Se requeriría replicar el experimento para corroborar las diferencias entre la dinámica de las PME en *C. aesculifolia* con respecto a las dos especies de Brasicáceas. Sin embargo, la actividad de las PME registrada en los lotes de *C. aesculifolia* con fenotipo de respuesta positiva al acondicionamiento mátrico es consistente con un patrón de actividad adecuado para culminar la germinación. En trabajos posteriores realizados en *C. aesculifolia*, esta dinámica en la actividad de PME podría se utilizada como referencia para contrastar con otros lotes colectados en años distintos o con fenotipo de respuesta negativa al acondicionamiento.

CAPÍTULO IV. El transcriptoma de la semilla seca y su relevancia en la capacidad de responder a estímulos ambientales, el caso de la respuesta al acondicionamiento mátrico.

Las especies tropicales que habitan en ecosistemas con estacionalidad marcada por la disponibilidad de agua deben de acoplar su ciclo de vida y sus programas de desarrollo con las fluctuaciones ambientales del recurso hídrico. La capacidad de sincronizar estos programas endógenos con el ambiente permitirá la sobrevivencia y evolución de una especie determinada en dicho ambiente. Los ciclos de hidratación/deshidratación que ocurren en el bosque tropical seco, a causa de precipitaciones esporádicas previas a la temporada de lluvias, pueden ser simulados en el laboratorio por medio del acondicionamiento mátrico. Las técnicas de acondicionamiento se utilizan para mejorar el desempeño germinativo de diversas especies. En el caso particular de *C. aesculifolia* se han detectado dos fenotipos de respuesta contrastante al tratamiento de acondicionamiento mátrico: un fenotipo donde el tratamiento mejora la germinación contra un fenotipo donde las semillas tratadas tienen un desempeño bajo con respecto a semillas no tratadas. En este capítulo se incluye un manuscrito en preparación sobre el efecto del acondicionamiento mátrico en *C. aesculifolia* durante el proceso de germinación, contrastando los cambios a nivel transcriptómico observados en los dos fenotipos de respuesta y en comparación con las semillas no tratadas.

El fenotipo de respuesta positiva al tratamiento mostró una mejoría en el desempeño germinativo en comparación con las semillas no tratadas. Asimismo se observó un ajuste fino del transcriptoma y del programa de germinación, que estuvo acompañado de una movilización temprana de lípidos y una disminución del contenido de ABA. En cambio, los lotes del fenotipo de respuesta negativa no mostraron mejoría en la germinación con respecto a semillas no tratadas, ni se observó una movilización temprana de lípidos. En estas semillas también se observó el ajuste fino del transcriptoma a causa del acondicionamiento; sin embargo, este ajuste generó cambios en la regulación de ciertos procesos, como el metabolismo de lípidos, que concuerda con los resultados mencionados anteriormente. Asimismo se observó un contenido de ABA más elevado en los lotes con fenotipo negativo en contraste al fenotipo positivo.

Los resultados obtenidos indican que el factor determinante del fenotipo de respuesta al acondicionamiento está asociado con el transcriptoma de la semilla seca. Este transcriptoma inicial de las semillas se ve afectado por las señales ambientales que percibe la planta madre durante la maduración de la semilla, o puede también alterarse por los procesos de post-maduración que ocurren en el estado seco durante el almacenamiento prolongado en condiciones sub-óptimas de temperatura y humedad. Los análisis realizados en este estudio podrían contribuir a un manejo más eficiente de los lotes de semillas de *C. aesculifolia*, especialmente considerando los análisis de contenido de lípidos y de respuesta al acondicionamiento mátrico como técnicas fáciles de implementar para determinar la calidad de un lote. Las figuras suplementarias que se referencian en esta publicación se incluyen en el anexo 7.

The transcriptomic profiles during germination reflect the coordination and adjustment of diverse processes that contribute to the response to priming

Keywords: *Ceiba aesculifolia*, germination, priming, transcriptome, wild species, response to water, lipid content.

Abstract
The proper timing of germination is a crucial step for plant survival. We tested the effects of a matric priming treatment, which simulates the hydration-dehydration cycles that occur due to sporadic rains prior to the rainy season in the tropical dry forest, in two contrasting priming phenotypes of *Ceiba aesculifolia* seeds collected in different years. The effects of the treatment are dependent on the initial transcriptomic landscape of the seed batch, which in turn is affected during seed maturation or prolonged storage. The positive response to priming was associated to a fine-tuning of the transcriptome during germination, as well as a reduction of ABA content and an the mobilization of a fraction of the stored lipids during early imbibition. In contrast, the seed batches with a negative response to the treatment present fluctuations in ABA content during early imbibition, and altered regulation of lipid metabolism that prevented the early mobilization observed in the positive phenotype. Understanding how environmental signals regulate developmental programs such as germination could aid in better management and use of the species in restoration programs in the tropical dry forest, which is at risk of alterations in the seasonal rain patterns due to climate change.

Keywords: *Ceiba aesculifolia*, germination, priming, transcriptome, wild species, response to water, lipid content.

Introduction
The seed constitutes a fundamental "pause" stage in the plant life cycle, and functions as a resistance structure to adverse conditions and as a unit for dispersal for individuals in both time and space (Bewley, Bradford, Hilhorst and Nonogaki, 2013). Humans have harnessed this inherent potential of seeds to utilize plant resources, and thus we now have a wealth of knowledge on how to manipulate germination performance in many species, especially of those of agronomic or industrial importance. Some of the most widespread methods to enhance germination performance are those collectively known as priming. Though the priming method may vary, the basic concept is to allow the imbibition of the seeds, while impeding the full germination process and radicle protrusion to occur by dehydrating the seeds (Paparella *et al.*, 2015).

It is generally assumed that the priming treatments does not alter the germination program *per se*, but rather it is an advancement of said process, allowing the primed seeds to progress the germination program to a certain point (Paparella *et al.*, 2015). The effects of priming are often reflected in a more synchronous germination and shorter times to complete germination, and in some species, it has been reported a benefit in the vigour of seedlings and young plants (Paparella *et al.*, 2015; Varier, Vari and Dadlani, 2010). While the literature is full of successful implementation of priming techniques in a wide variety of species, the majority of those correspond to domesticated plant species, for which the final aim is to retain the benefit of the treatment after certain time in storage, and to ensure or improve production on the field by enhancing vigour and survival of the seedlings, as well as resistance to different stresses. Thus,

there is still a lack of understanding of the cellular/molecular/genetic mechanisms that generate the priming response, with a less extended list of studies and reviews that have endeavoured into the matter (Chen and Arora, 2013; Cheng *et al.*, 2017; Gallardo *et al.*, 2001; Kubala *et al.*, 2015; Paparella *et al.*, 2015; Soeda *et al.*, 2005; Varier *et al.*, 2010; Yacoubi, Job, Belghazi, Chaibi and Job, 2011).

There are several studies that report positive responses to priming treatments in wild plant species, such as temperate annuals like *Ranunculus sceleratus* (Probert, Bogh, Smith and Wechsberg, 1991) and *Digitalis purpurea* (Butler, Hay, Ellis, Smith and Murray, 2009), and in tropical species such as the treelike shrubs *Psidium guineense* (Santos, Queiróz, Bispo and Dantas, 2015) and *Wigandia urens* (Gamboa-deBuen, Cruz-Ortega, Martínez-Barajas, Sánchez-Coronado and Orozco-Segovia, 2006; González-Zertuche *et al.*, 2001), in trees such as *Tecoma stans*, *Cordia megalantha* (Alvarado-López, Soriano, Velázquez, Orozco-Segovia, A. and Gamboa-deBuen, 2014), *Albizia saman*, *Cedrela odorata*, *Enterolobium cyclocarpum* and *Swietenia macrophylla* (Peraza-Villarreal *et al.*, 2018), as well as in desiccation-sensitive species such as *Harconia speciosa* (Masetto and Scalon, 2014), *Cupania glabra* and *Cymbopetalum bailloni* (Becerra-Vázquez, Sánchez-Nieto, Coates, Flores-Ortiz and Orozco-Segovia, 2020).

To this day, whether the priming treatment is implemented on wild or domesticated species, the process of treatment calibration still depends on trial and error, and even when a priming procedure resulted successful on a seed batch, replication over different seed batches of the same species may not yield similar results. We have particularly observed this phenomenon in seed batches of *Ceiba aesculifolia*, a wild tree species from tropical dry forests (TDF) within Mexico, for which we have observed both positive and negative responses to matrix priming (Gómez-Maqueo *et al.*, 2020). We also implemented a system to track the germination process based on physiological stages associated to the seed relative water content (RWC) instead of chronological time to complete germination. This system allowed us to implement a *de novo* transcriptomic analysis in which we observed wide and significant differences in the transcriptome from non-primed seeds from batches which had been identified as having a negative response (NR-phenotype), with respect to the transcriptomic profile of non-primed seeds from the seed batches with a positive response (PR-phenotype). The main transcriptomic differences between NR- and PR-seeds were observed in the dry state (T0), where the NR-phenotype had an overall transcriptomic profile resembling an early imbibition stage, and an overrepresentation of stress-related gene ontology terms. By tracking the transcriptomic differences at discrete physiological stages across the germination process, we also reported that those differences gradually subsided towards the moment of radicle protrusion. Although the germination program was not altered *per se* in the NR-batches, the overall response to the priming treatment was affected by these differences in the initial transcriptomic landscape. However, it is an open question whether the priming treatment failed to promote the advancement of the germination program in the NR-batches, or if there is evidence of altered mechanisms that overshadowed the response to the treatment.

Here we present the transcriptomic profiles of the PR- and NR-primed seeds and the contrasts with their respective controls across germination to determine if the positive response to the treatment reflects the expected advancement of the germination while the NR-batches do not. We also ask if the contrast of the transcriptomic landscapes between the PR- and NR-batches indicates an alteration in specific biological processes, and if they occur in specific physiological stages during germination. Finally, we discuss the ecological relevance of the proper integration of the hydric signal generated by priming into the germination program of *C.*

aesculifolia seeds, since hydration-dehydration cycles in the soil are an important part of its ecological context.

Materials and methods

Plant material. The plant material used was previously studied by Gómez-Maqueo *et al.* (2020), and correspond to seeds of *Ceiba aesculifolia* (Kunth) Britten & Baker (*Malvaceae*) collected in February 2012, 2014 and 2015, from a population located within a human-activity disturbed ecosystem in the location of "Trapiche del Rosario y Chicuasén" in Actopan Veracruz, Mexico (latitude 19.5426, longitude −96.7401, altitude 479 m a.s.l.). The climate is warm subhumid with an average temperature of 22 °C and a rainy season spanning from July to November (209 mm monthly precipitation), and a dry season (December-June, 22 mm monthly precipitation; García, 2004). The seed batch reported as "2016" was collected in December 2015 due to an un-timely fructification season in that year. This previously reported material was used for germination and priming tests, as well as for the transcriptome assembly. All sample collection and germination tests occurred after the seed batches had completed at a 30-day afteripening period. In the case of the 2012 and 2014 batches, sample collection, as well as secondary germination and priming tests were performed after the seeds had experienced a dry storage period of 5 and 3 years respectively (23-25 °C, 35-45% RH, in darkness). These stored seed batches will be presented as 2012-5y and 2014-3y to distinguish them from the freshly collected samples. All samples were stored at -80 °C until processed.

Germination and response to priming treatment. Seeds were submitted to a matric priming treatment as described in Gómez-Maqueo *et al.* (2020). In brief, seeds were buried in pots with a commercial soil mixture (METRO-MIX 300, Sun Gro Horticulture, MA, USA) moistened to full capacity during two days; they were covered with tin foil to avoid humidity loss and kept in a dark room at 23–25 °C. The seeds were then exhumed and allowed to dry at room temperature in darkness for two more days. Afterwards, treated and untreated seeds were washed with 10% hypochlorite for 1 min, surface dried with a paper towel, and placed on 1% agar plates. Five replicates (30 seeds each) of all treatments were incubated in a germination chamber at 25 °C and a photoperiod 12 h:12 h light:dark. Cumulative germination was registered daily for about 25 days. Germination percentages were arcsine-transformed and a sigmoid curve was fitted to the data. Statistical differences between control and primed seeds per seed batch were tested with a two-tailed Student's t-test, at a significance level of $\alpha = 0.05$. Priming response was deemed as favourable when at least one of the following characteristics improved with respect to its control: 1) germination initiation, 2) germination index (GI), 3) mean germination time (t_{50}), and 4) final germination percentage.

Sample collection for transcriptome assembly. For all batches and treatments, the samples consisted of five seeds that were collected at discrete physiological stages, based on the increment in the seed RWC during the early germination stages (for PR-batches: samples at 16% and 50% RWC; for NR-batches: at 50% RWC), at the moment of testa rupture (TR) and of radicle protrusion (G). For the transcriptomic analyses, each seed batch was considered as a biological replicate, thus the PR-phenotype consisted of three biological replicates (2014, 2015-1 and 2015-2), and the NR-phenotype consisted of another three biological replicates (2012-5y, 2014-3y and 2016). RNA extraction was performed with a TRIZOL protocol modified from Li

and Trick (2005). Transcriptome assembly and quality assessment were previously reported by Gómez-Maqueo *et al*. (2020).

Differential expression, time series, and functional analyses. Differential expression was analysed using a multifactorial design in R with the DESeq2 package to test the effect of the treatment, phenotype, and the physiological stage (Love, Huber and Anders, 2014; R Core Team, 2016). We performed the test in three independent runs, the primary model included all samples, except the 16%RWC, and another two runs for each phenotype in separate to be able to observe the transition from T0 to 16%RWC in the PR phenotype. The differentially expressed genes pertaining to the effect of the treatment at the 16%RWC were combined with the results from the primary model for downstream analyses of biological functions enrichment. We also performed time series analyses in the STEM suite (Ernst and Bar-Joseph, 2006; Ernst, Nau, and Bar-Joseph, 2005). First, we normalized the read counts with DESeq2 and then standardized the global expression profile of each gene in each treatment and phenotype using MFUZZ functions (Kumar and Futschik, 2007). These data were then imported into STEM software for profile classification and clustering. We also used the web-based tool Metascape for GO-term analyses of the STEM clusters (Zhou *et al.*, 2019).

Total lipid content quantification. Lipid extraction and quantification was performed with the methanol-chloroform protocol proposed by Bligh and Dyer (1959). Samples for all physiological stages within each seed batch and treatment were collected in triplicate. Lipid content was expressed as the mean percentage of the seed dry weight. We performed a two-way ANOVA on the arcsine-transformed percentages to test for differences in total lipid content between control and treated seeds, and between physiological stages within each seed batch. The significance level was set at $\alpha = 0.05$.

Lipid Extraction and Fatty Acid Analysis by Gas Chromatography-Mass Spectrometry (GC-MS). We performed a GC-MS on the extracts of representative samples of the PR- and NR-phenotypes, using a transesterification method (Cabrera Santos, *et al.*, 2021; Priestley, 1986). Fatty acid transesterification was done through evaporation of the chloroformic phase and reaction with BF3-CH3OH. After that, C_6H_{14}:H_2O (2:1) was added to recover the methyl esters of fatty acids from the organic phase. Heptadecanoic acid was used as the internal standard for fatty acid quantification (Priestley, 1986). For the analysis of the methyl esters of fatty acids, a gas chromatograph (Agilent Technologies 6850, Santa Clara, CA, USA) coupled with a mass spectrometer (Agilent Technologies 5975C VL MSD, Santa Clara, CA, USA) was used. A DB-1 (dimethylpolysiloxane) capillary column (30 m length 0.32 mm i.d., 5.00 m film thickness, part number:123-1035E, Agilent Technologies 6850, Santa Clara, CA, USA) was used for the GC system. The oven temperature was programmed as follows: from 100 °C; ramp 1: To 250C with 5 °C/min. The injector temperature was 200 °C in split mode. Helium was used as carrier gas at a linear flow velocity of 35 cm s^{-1} of 1.4 mL min^{-1}. Mass detector conditions were: transfer line at 250 °C, range from 20 to 400 m/z, positive polarity, the ionization energy of 70 eV, and temperature of 200 °C, with an injection volume of 2 µL (Cabrera Santos, *et al.*, 2021). The mass spectra were compared with the NIST/EPA/NIH Mass Spectral Library 2020 version (NIST, 2021). Fatty acid analyses were performed in triplicate. The dry seeds (T0) were the control for any of the treatments and phenotypes.

Abscisic acid (ABA) content quantification. ABA was quantified according to Steinbach, Benech-Arnold, Kristof Sánchez and Marcucci-Poltri (1995). In brief, a radio-immuno-assay, which uses the monoclonal antibody MAC 252, was performed in seed samples from both phenotypes and treatments. The samples were collected at five different RWCs between T0 and 50% RWC. All sample stages were collected in triplicate, freeze-dried and grounded in a mortar. Statistical analysis of ABA content was performed using a non-parametric three-way mixed ANOVA with aligned rank tests in R (Wobbrock, Findlater, Gergle, and Higgins, 2011), and was performed as indicated in the vignette for the ARTool package (Elkin, Kay, Higgins, and Wobbrock, 2021).

Data availability. Reads were deposited in the SRA platform within the NCBI GenBank website (www.ncbi.nlm.nih.gov/genbank) under Bioproject PRJNA561202, and accession numbers SAMN12611291 and SAMN12611291. Assembled sequences were deposited in the TSA platform under accession numbers between GHVB01000001 and GHVB01166545.

Results

Germination and the response to matric priming

The response to the priming treatment was tested in six independent seed batches (named 2014, 2015-1, 2015-2, 2014-3y, 2012-5y, and 2016), which in turn were classified into two phenotypes (PR-batches: 2014, 2015-1, and 2015-2; NR-batches: 2014-3y, 2012-5y, and 2016) after the evaluation of their response to the priming treatment. In Fig. 1 are shown the four variables registered to assess germination performance (data adapted from Gómez-Maqueo *et al*., 2020). The positive response phenotype (PR) was associated with a reduced time needed for the completion of the germination process, as seen by a shorter time for the onset of germination (Fig. 1A), as well as the time to attain 50% of germination (Fig. 1C). In the case of the 2014 batch, priming also improved final germination (Fig. 1D). The 2015-1 batch was the only PR-batch that displayed a reduced GI due to the treatment (Fig. 1B); however, the overall germination performance was deemed as favourable since it displayed a significant reduction in the onset of germination, and the germination window widened from 10 days in the control, to 14 days in priming. The negative phenotype (NR) displayed by the 2014-3y, 2012-5y and 2016 batches depicted a slower germination dynamic in comparison to the PR-batches and non-statistically significant differences between treated and control seeds. The 2014-3y and 2012-5y batches had been stored for 3 and 5 years respectively at the time of testing. The 2016 seed batch, despite being tested in the same collection year, had a germination performance similar to the stored seed batches, and is the seed batch with the lowest germination performance in all the germination characteristics we tested (Fig. 1). This germination performance, as well as a negative response to the priming treatment, was also observed in another seed batch (not part of the current study) collected the following year (2017, Supporting Information Fig. S1). The low performance of the seed batches collected in 2016 and 2017 could be associated to atypical weather conditions that occurred during those years (Supporting Information Table S1), which in turn could have generated stressful conditions for the mother plants. This germination trend seems to have been partially alleviated in the following years, since another seed batch collected in 2018 had a positive response to priming by reducing the time for the onset of germination with respect to the untreated seeds (Fig. S1).

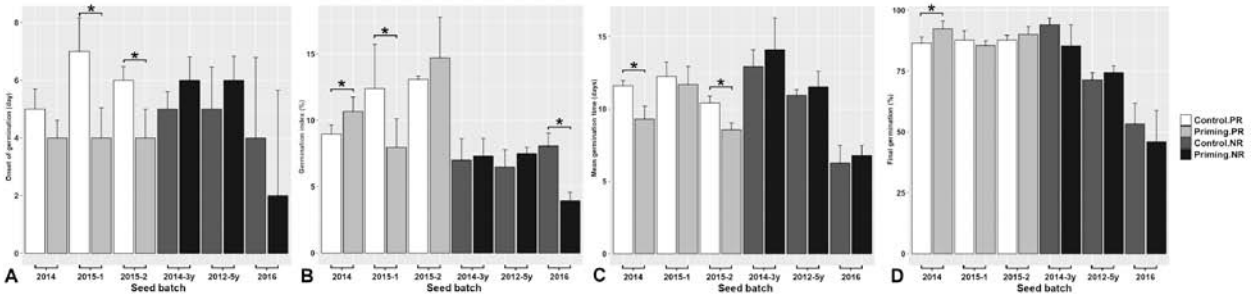


Figure 1. Germination performance and response to priming of the two contrasting phenotypes. The data of the four variables used to assess germination was adapted from Gómez-Maqueo *et al*. (2020). A) Days required for the onset of germination; B) Germination index (%); C) Mean germination time; and D) Final germination percentage. The vertical bars represent standard deviation. The asterisks at the top of the bars denote significant differences ($P < 0.05$) between control and primed seeds within each seed batch.

The transcriptomic profiles of PR- and NR-batches are highly dynamic and display important differences between control and primed seeds

The assembled transcripts were associated to a unique *Arabidopsis* locus tag, yielding a pool of 12,683 locus tags. Profile comparisons were performed independently in the PR- and NR-batches between control and primed seeds using a time series approach with the STEM suite (Ernst and Bar-Joseph, 2006), considering each physiological stage based on the RWC as the natural progression of the germination process (Gómez-Maqueo *et al.*, 2020). In the PR phenotype, this analysis produced 53 (control), and 55 (priming) possible profiles, from which 11 were significant in both control and primed seeds, and clustered 10,121 and 10,309 genes respectively (Fig. 2A, control: profiles denoted with the letter "a", priming: profiles denoted with the letter "b"). We detected four major trends in these 11 profiles: 1) an up-regulation trend, in which the maximum expression was attained at the moment of radicle protrusion (G); 2) an up-down regulation, which was primarily associated with a maximum expression at the 16%RWC stage; 3) a down-up regulation trend, with the lowest expression at the 16% RWC, followed by a peak expression at either testa rupture (TR) or G; and 4) a down-regulation trend, with its maximum expression in the dry seed (T0), and low expression values at later stages. Although the number of significant profiles was the same, their identity was not equal: of the 9,019 genes that were identical between the control and priming profiles, about 49% were assigned to the exact same profiles in both control and primed seeds (Fig. 2. dark green lines in each plot), indicating that over half of the genes changed profiles in response to the priming treatment (Fig. 2. light green lines in each plot). The percentage of identity among equivalent profiles between control and primed seeds was not homogeneous as well, with some profiles having as much as 70% of identical genes assigned (Fig. 2. Profiles 1-a and 1-b), and other profiles having as little as 9% of identity (Fig. 2. profiles 4-a and 4-b). A reference table of all alphanumerical equivalencies between comparison groups is shown in Supporting Information Table S2. Also, each treatment had a singular profile unique to them, with profile 7-a being significant in control seeds and having an up-down regulation with peak expression at TR, while profile 10-b was significant in the primed seeds, with a down-then-up regulation trend with a low peak at 16% RWC and a maximum peak at TR. In the case of profile 7-a, about 32% of its genes were assigned to profile 5-b in the primed seeds, another 5% to profile 6-b, and the rest were assigned to non-significant profiles in the primed seeds. Meanwhile, the priming-exclusive profile 10-b received about 40% of its genes from profile 1-a, and another 35% from non-significant profiles in the control treatment.

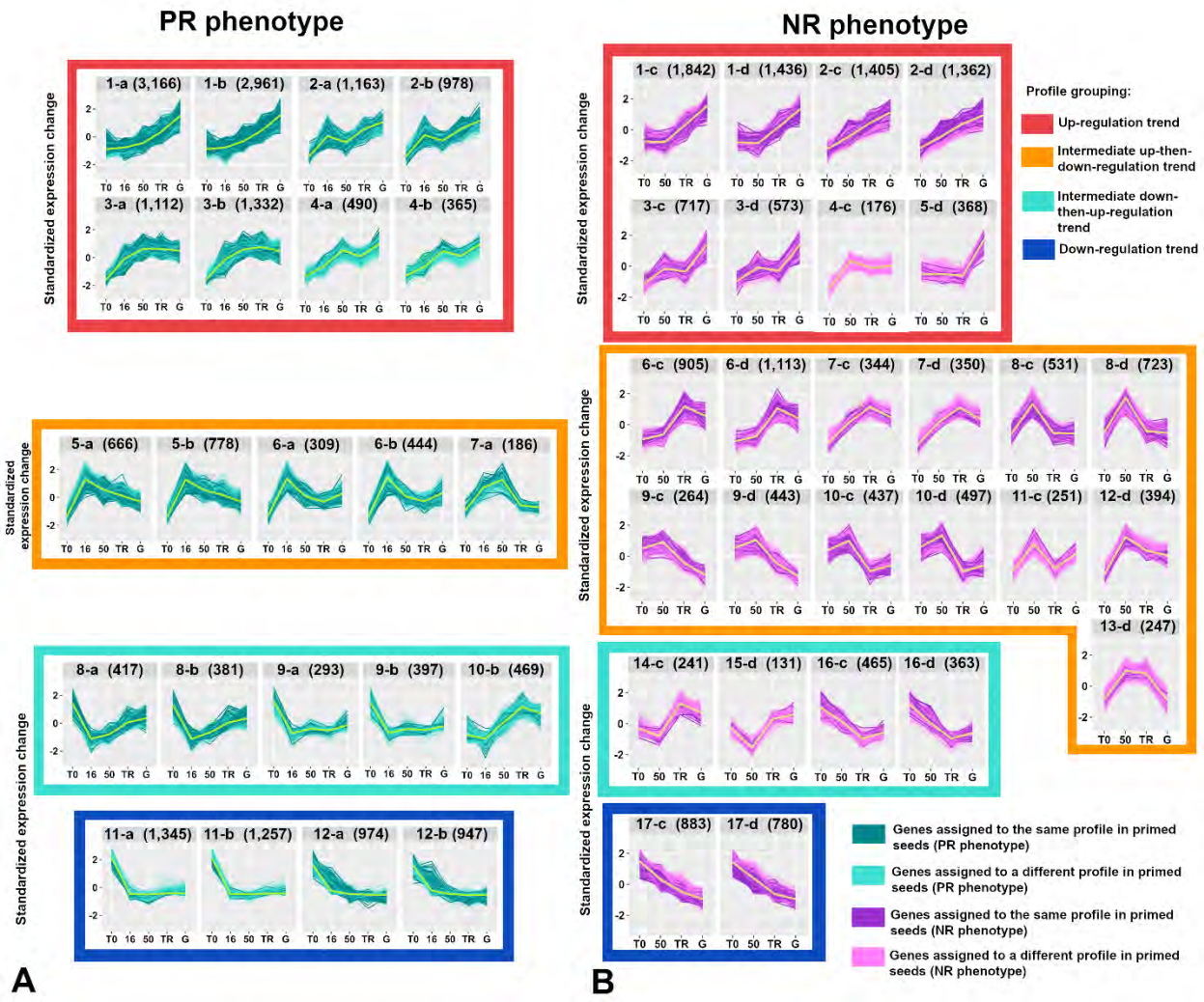


Figure 2. Expression profiles detected in the transcriptome of control and primed seeds in the PR phenotype (A) and the NR phenotype (B). Expression levels of each gene are represented as a standardized expression with mean zero. Profile graphs are grouped into four major trends (red, orange, turquoise and blue boxes). The numbers in each profile indicate profile identity within each figure panel (A or B), while the lowercase letters distinguish between treatments (a: PR-control, b: PR-priming, c: NR-control, and d: NR-priming). The numbers in parentheses indicate profile size. The dark green and dark purple lines depict the genes that were assigned to the same profile in both control and primed seeds within each phenotype. The light green and light purple lines depict those genes whose expression profile changed in response to the priming treatment and thus were assigned to a different profile with respect to its control.

In the NR phenotype, there was an increase in the number of possible profiles, up to 56 (control) and 60 (priming), from which 13 were significant in the control, and 14 in the primed seeds. Also, each treatment had three and four unique profiles respectively (Fig. 2B; control: 4-c, 11-c, and 14-c; priming: 5-d, 12-d, 13-d, and 15-d). Despite having a wider diversity of expression profiles, less genes were clustered within those profiles, with 8,461 and 8,780 genes for control and primed seeds, respectively. Profile 1 was the largest profile, but still had about half the genes in comparison to their counterpart in the PR phenotype, with 1,842 for the control seeds, and 1,436 genes for the primed seeds. Gene identity within equivalent profiles was also lower, with profiles 1-c and 1-d having the highest similarity at about 50% (Fig. 2B, dark purple lines), and a complete loss of identity in the case of profile 4-c, which was also exclusive to the control treatment (Fig. 2B, light purple lines). In contrast to the PR phenotype, the NR seeds had

a higher diversity of intermediate regulation trends, especially those related to peak expression ant 50%RWC or TR (up-down trend), with about eight significant profiles. This phenotype also showed a lower diversity of significant down-regulated profiles with just one profile in each treatment (Fig. 2B, 17-c and 17-d).

The physiological stages based on RWC reflect the progression of the germination process in both phenotypes

We performed a hierarchical clustering of all profiles and trend groups in both control and primed seeds to detect the major biological processes occurring within each treatment and phenotype. Based on abundance of terms, the "metabolism" category (GO:0008152) was the most important category in both phenotypes, and was present in almost all profiles, indicating a highly dynamic regulation (Fig. 3, and Supporting Information Figs. S2 and S3). The following categories in abundance were "localization" (GO:0051179), "organization or biogenesis of cellular components" (GO:0071840), and "response to stimulus" (GO:0050896), which were also enriched in about 75% of the profiles in both PR and NR phenotypes. Both priming phenotypes shared most of the categories shown in Fig. 3a, the exceptions being "detoxification" (GO:0098754), "rhythmic process" (GO:0048511), "reproduction" (GO:0000003, PR-control exclusive), "biological adhesion" (GO:0022610, PR-priming exclusive) for the PR phenotype, and "nitrogen utilization" (GO:0019740) for the NR phenotype. We also detected 1,253 and 1,211 GO-terms present in the PR and NR phenotypes respectively, of which 723 were shared between phenotypes. The shared terms included fundamental germination and seedling establishment processes such as DNA repair, energy metabolism and reserve mobilization, mitochondrion-related processes, response to light, ribosome/translation processes, lipid metabolism, as well as photosynthesis and chloroplast-chlorophyll related processes (Supporting Information Figs. S4 and S5, and Supporting Information Table S3). These processes were particularly represented in the up-regulation trend group (Fig. 3b), indicating a gradual induction of key mechanisms towards the end of germination in both phenotypes. The down-regulation trend was particularly enriched in genes associated to phase transitions of the meristem, chromatin organization, ribosome organization, negative regulation of transcription, and splicing; some of which have been reported as part of the developmental switch from the maturation program to the germination program during imbibition (Dekkers *et al*., 2013; Narsai, Law, Carrie, Xu, and Whelan, 2011). In the PR phenotype, the intermediate trends reflected a fine tuning of key germination processes, present in either the up or down trends. The up-down trend usually presented a peak expression at the 16%RWC and had an enrichment in ribosome biogenesis/organization, as well as the regulation of alternative splicing, and chromosome/chromatin organization. The down-up trend, which implicated a decrease in expression at the 16%RWC, fine-tuned diverse processes but mainly in the metabolism, signalling, and developmental categories. Meanwhile, the NR phenotype had a higher number of terms in the up-down-regulation group. The main differences among control and primed seeds, as well as with the PR phenotype, were present in the down-up trend group. There was a particular enrichment of development and hormone-mediated signalling in the up-down trend, while the down-up trend was particularly enriched in ribosome-translation processes, contrary to the observed enrichment in the PR phenotype. Only the alternative splicing category remained enriched in the up-down trend as it occurred in the PR phenotype. The term abundance in the NR seeds was lower within all parental categories, most notoriously in the metabolism category with respect to the PR seeds (Fig. 3a, Supporting Information Table S4).

Although the priming treatment induced a change of profile for over half of the analysed genes in contrast to their respective controls, such changes usually occurred among profiles within the same trend group. Thus, the physiological outcome of having either a peak expression at the dry seed stage, by the end of germination or in intermediate stages remained relatively the same within each phenotype (Fig. S4). Also, the PR phenotype had a more dynamic regulation than the NR phenotype despite having less profile diversity. This was particularly evident after analysing the resulting network of the top 20 GO-term clusters of each phenotype. While in the PR-phenotype all trend groups were represented in the GO-network (Fig. S4; Supporting Information Table S5), the top 20 functional clusters in the NR seeds were primarily present in the up-regulated trends (peak expression by the end of germination; Supporting Information Table S6; Fig. S5).

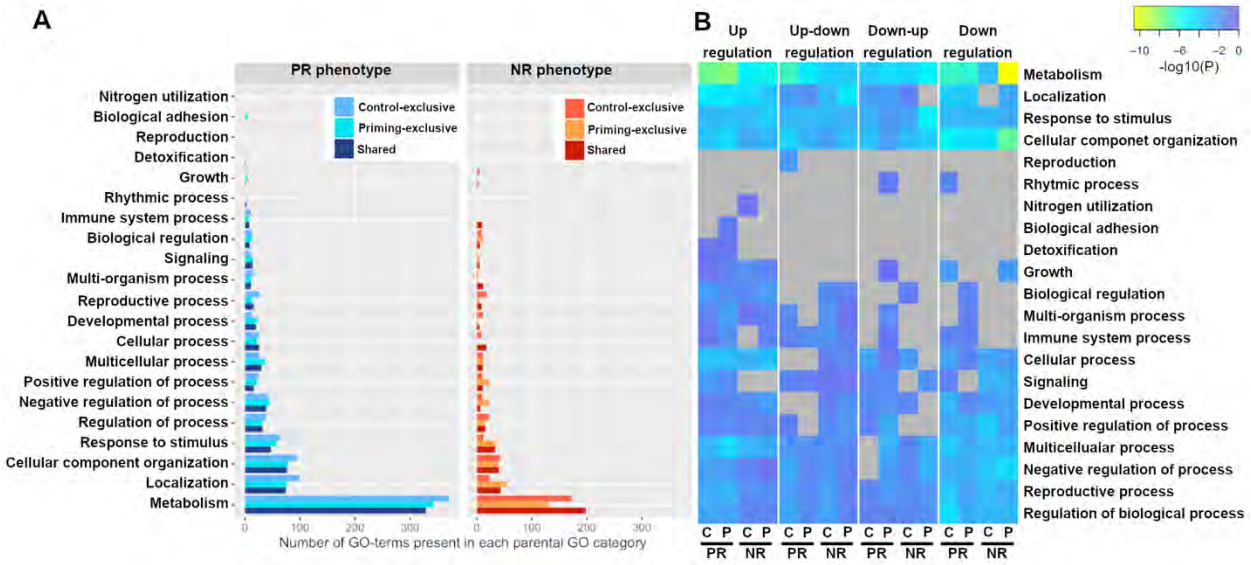


Figure 3. Gene ontology term abundance and distribution among expression trend groups in control and primed seeds of the two studied priming phenotypes. A) Total number of GO terms detected in both phenotypes, grouped by broader ("parental") GO categories within each phenotype. The "shared" bar depicts those GO-terms that were detected in both control and primed seeds, whereas the "control" and "priming" bars depict those terms which only occurred in either treatment. B) GO term distribution heatmap, depicting the presence and relevance (based on the negative logarithm of P-values) of terms within profile trend groups.

The differentially expressed genes in response to the priming treatment reflect a fine-tuning of the germination program

There was an important overlap of biological processes between control and primed seeds within phenotypes, with about 15% of the registered GO-terms been exclusive of either control or primed seeds (Fig. 3). These exclusive GO-terms usually included some genes that participated in related subprocesses (child terms) or modules within a pathway that were already detected in the shared terms between treatments (Tables S3 and S4).

The expression analysis in DESeq2 detected 7,604 differentially expressed (DE) genes due to the combined effect of the phenotype and treatment, in at least one stage (Supporting Information Table S7). This number of DE genes is very similar to the number of genes that changed profiles in response to priming, as evaluated in the time series analyses with STEM (7,883 genes in PR; 6,446 genes in NR), indicating that both approaches detected similar changes in the overall expression trend. We then performed a GO term enrichment of this DE dataset, grouped by their respective profiles. The DE genes were predominantly associated to

primary metabolism/energy production, DNA metabolism and repair, ribosome biogenesis, chromosome organization, response to stimulus and to ABA, and carbon fixation. We then tested the effect of priming on the expression at a specific stage, and the number of DE genes was reduced to a total of 61 genes for the PR phenotype, and seven genes for the NR phenotype (Fig. 4; Supporting Information Fig. S6, Table S8). The enrichment analysis detected four GO-terms associated to responses to abiotic stimulus (GO:0071214), to fungus (GO:0050832), oxidative stress (GO:0006979), and to fatty acids (GO:0070542). Combined, both DE tests and the profile clustering indicate that the priming treatment does not generate large changes on the overall germination program, but instead fine tunes the expression pattern of most genes, while altering net expression of a smaller set of genes in particular biological processes.

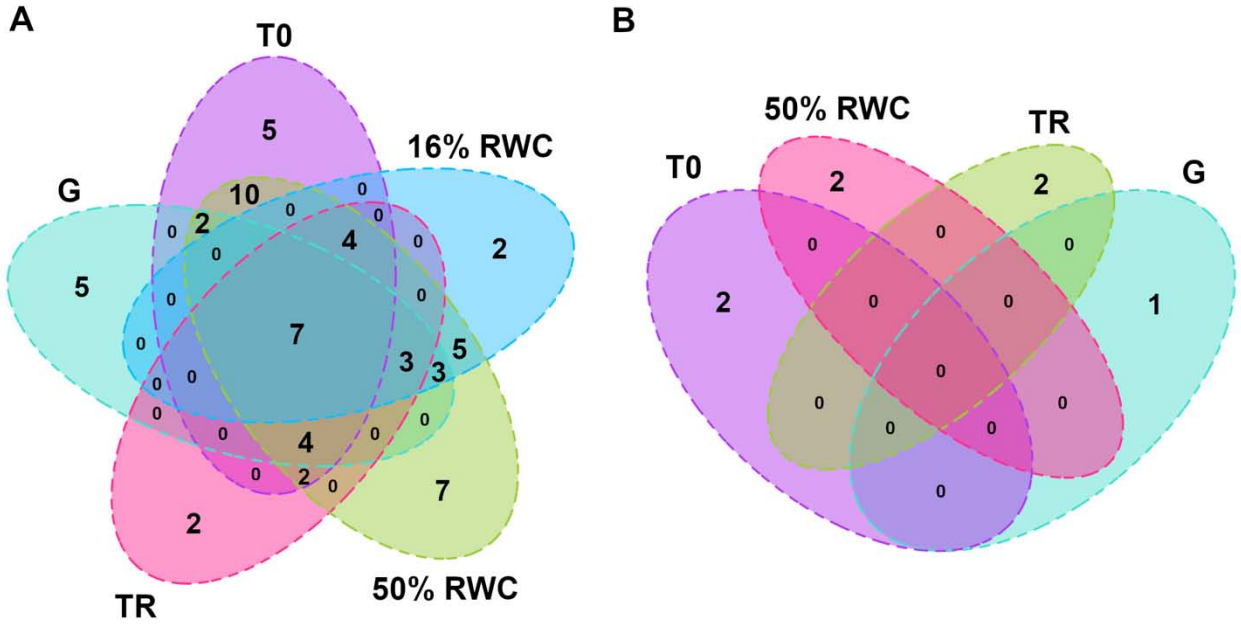


Figure 4. Differentially expressed genes in response to the priming treatment in the PR and NR phenotypes. A) Venn diagram depicting the overlap of DE genes among physiological stages in the PR phenotype. B) Venn diagram depicting the overlap of DE genes among physiological stages in the NR phenotype.

The C. aesculifolia seeds present fluctuations in ABA content during germination in response to the priming treatment

The differential expression analysis of the PR phenotype indicated that the response to priming was related to environmental or stress responses. These environmental responses are mediated by several hormones, including brassinosteroids, jasmonates, and ABA. The role of ABA in seed maturation, dormancy and germination has been extensively studied in other species (Finch-Savage and Leubner-Metzger, 2006). In *C. aesculifolia*, we observed that the duration of the low imbibition phase (between 10% and 20% RCW) is about 45% longer in NR-seeds in contrast to PR-seeds (Gómez-Maqueo *et al.*, 2020). We quantified ABA during the early stages of germination, between T0 and 50% RWC, since this period comprehends the longest phase of imbibition, and it is when ABA would exert an inhibitory activity over germination. The seeds of *C. aesculifolia* present ABA concentrations ranging from 30 to 70 pg·mg^{-1} of dry weight (Fig. 5). There was a statistically significant interaction between phenotypes, treatments and physiological stages, indicating that ABA content at a given stage depended on the phenotype and treatment observed. The major changes in ABA content occurred between the two RWCs that delimit the slow imbibition phase (10% and 20%), but at the 50% RWC, the ABA content

was similar in both phenotypes and treatments (Supporting Information Table S9). In the PR phenotype, the priming treatment reduced ABA content at T0, 16% and 20% RWC. Untreated seeds in the PR phenotype also reduced their ABA content during imbibition and remained relatively stable during later stages. Meanwhile in the NR phenotype, the priming treatment did not reduce ABA content at T0 but did reduce it at later stages (16% and 20% RWC) before returning to basal levels detected at T0. The untreated seeds in the NR phenotype showed a peak in ABA concentration towards the 16% RWC. The reduction in ABA content during imbibition in non-dormant seeds has been associated with ABA catabolism and the enhanced expression of said pathway, while in highly dormant seeds *de novo* synthesis has been registered, as well as an up-regulation of ABA synthesis genes (Finch-Savage and Leubner-Metzger, 2006). At the T0 stage in the NR-seeds, there was an up-regulation of four genes involved in ABA synthesis (Xerico, ZEP, AAO1, and NCED3), as well as one ABA catabolism gene (CYP707a2) with respect to PR-seeds (Supporting Information Fig. S7). These genes presented opposite expression trends between phenotypes during early germination (T0-50% RWC), but these differences diminished at later stages. Additionally, there was no enrichment of ABA metabolism during the germination process in any of the evaluated phenotypes or treatments. Although the major differences between ABA-metabolism genes were detected between T0 and 50% RWC, the change in ABA content at the 16% RWC observed in both phenotypes will require an in-depth analysis to detect the transcriptomic regulation that could be related to ABA metabolism during early imbibition in *C. aesculifolia*.

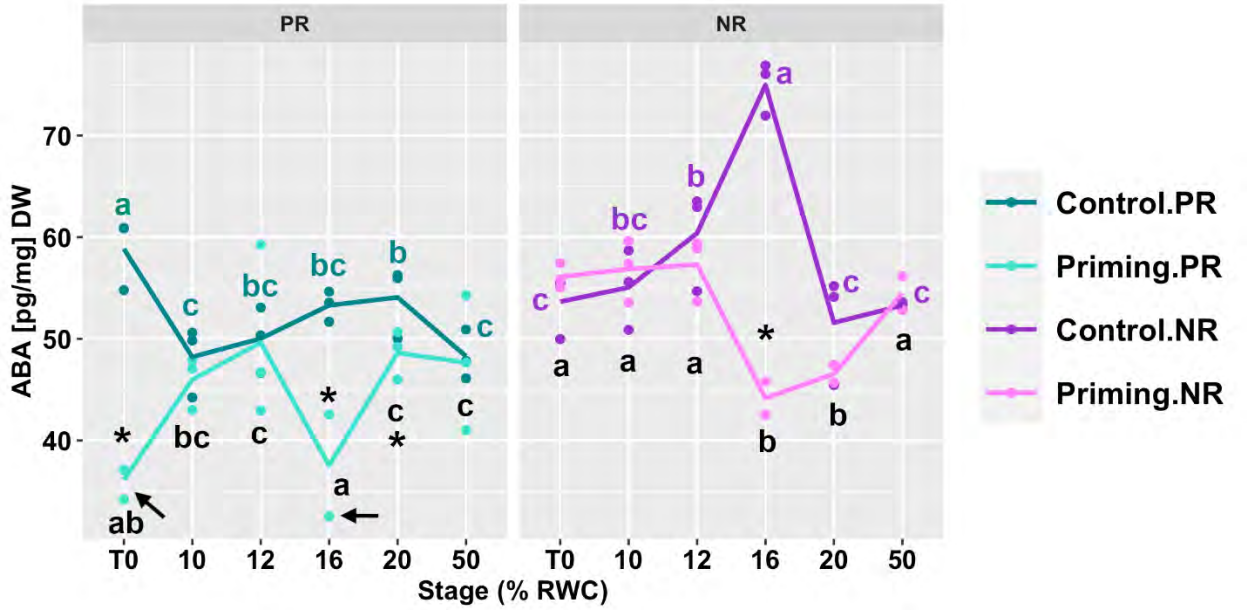


Figure 5. ABA content fluctuations in the two priming phenotypes during early imbibition (before the TR and G stages). Each dot represents an independent replicate. The arrows mark those replicates in which the estimation of ABA content was extrapolated from the quantification curve and should be interpreted with caution. There was a statistically significant interaction of the phenotypes×treatments×physiological stage ($p < 0.05$), indicating that ABA content at a given physiological stage depended on the phenotype and treatment observed. The asterisks denote significant differences between control and primed seeds within each phenotype, while the letters denote significant differences between physiological stages within each treatment in each phenotype (dark green: Control PR; dark purple: Control NR; black: primed seeds).

The positive response to priming is associated with proper regulation of lipid metabolism and an early mobilization of a small portion of storage lipids

Lipids are the main energy source of the *C. aesculifolia* seeds, representing 30% of its total weight (Fig. 6). The lipid profiling showed that the linoleic acid was the principal fatty acid contained in the seed (~41%), followed by oleic acid (~30%), palmitic acid (~27%), and minimal amounts of stearic acid (~2%). The GO-network analysis showed that lipid metabolism was one of the most important processes that were regulated in both phenotypes. However, this regulation occurred differently among treatments and phenotypes. For instance, in the PR phenotype, the control seeds had lipid catabolism and synthesis processes subdivided into two up-regulation profiles with peak expression at the end of germination (Fig. 6A lower panel, light pink and magenta circles, respectively). The catabolic pathways were present in profile 2-a, which has a small expression peak at 16%RWC; the synthesis pathways and generic lipid metabolism were present in profile 1-a, which has a slow but consistent increase in expression towards the end of germination. In their respective primed seeds, both synthesis and catabolism occurred mostly in profile 1-b (magenta circles in Fig. 6A, lower panel), but certain modules were subject of fine-tuning. The catabolism pathways had a module being down-regulated towards the end of germination (11-b), and the synthesis and generic metabolism had modules being either down-regulated (9-b) or with a slight peak (2-b) at the 16% RWC stage. Also, in contrast to its respective control, primed seeds had a better representation of other processes needed for proper reactivation of seed metabolism, such as the glyoxylate cycle and those involved in the regulation of the redox state and cell respiration (1-b), and the regulation of ROS metabolism was rapidly induced upon imbibition in the primed seeds (2-b, 3-b) in contrast to untreated seeds (1-a, Fig. S4).

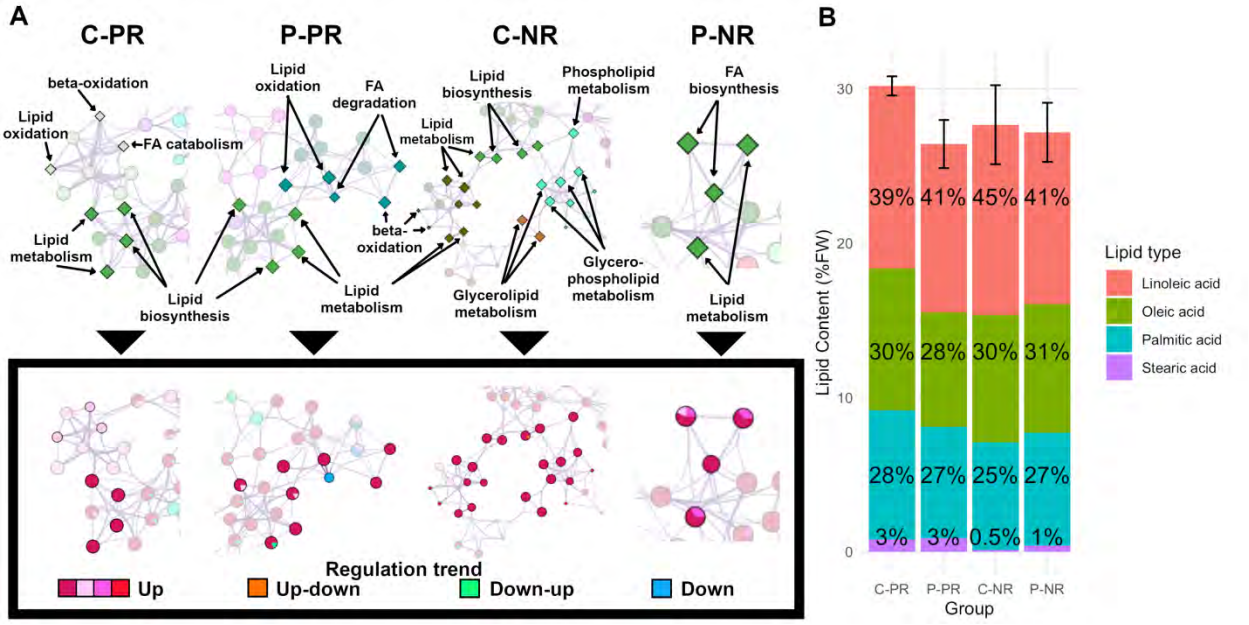


Figure 6. Lipid metabolism in *C. aesculifolia* seeds during germination. A) Extract of the GO-term networks detected in each treatment and phenotype in which lipid-related processes are allocated. The upper half of the panel depicts the GO-terms while the lower half of the panel depicts the regulation trend in which each term was detected. Non-lipid terms were blurred to ease interpretation. The full networks are shown in Figs. S4 and S5. The size of each diamond/circle reflects the number of genes involved in each functional category. B) Total lipid content and profiling of the seeds at the T0 stage. The black bars depict standard deviation; the percentage over each colour depicts the average percentage of each lipid type. Abbreviations: C: control, P: priming.

Meanwhile, in the NR phenotype, lipid metabolism was almost exclusively represented in profile 1-c in the control seeds, but primed seeds showed the lowest diversity of lipid-related

terms, which were also subdivided in three up-trends (1-d, 2-d and 3-d, Fig. 6A). In the control seeds, while a submodule of the β-oxidation was present in profile 1-c, the other satellite processes associated with ROS metabolism and oxidative stress were poorly represented in profile 14-c within the up-down-regulation trend (Tables S4 and S6). In the primed seeds there was no enrichment of lipid degradation through β-oxidation, and there was a delayed response to ROS as it was present in profiles 1-d and 15-d and had twice the number of allocated genes in comparison to the PR phenotype (Tables S4 and S6). Also, in the NR control seeds the metabolism of membrane lipids had a more prominent representation than in the other tested seeds, and had enriched categories associated to the catabolism of toxic derivates such as methylglyoxal, which are also indicative of stress responses (Tables S4 and S6).

The expression dynamism of lipid metabolism was also reflected in the total seed lipid content. In the PR phenotype, the primed seeds had an early mobilization of a small but significant portion of its storage lipids (Fig. 7A-C). This mobilization occurred at the T0 stage, and total lipids remained stable until the 20%RWC. The control seeds mobilized these lipids until the 16%RWC was reached. Afterwards, both treatments followed a similar trend towards the end of germination. Interestingly, the control seeds showed an initial peak expression at 16%RWC of the catabolism pathways (profile 2-b), especially β-oxidation, and coincides with the moment in which the reduction in lipid content occurs in these seeds. This peak is downregulated towards the 50% RWC stage, and then progressively increases again towards germination, thus emulating the following reduction of storage lipids. Instead, in the primed seeds, the β-oxidation pathway is subdivided into profiles 1-b (up) and 11-b (down). In profile 1-b, gene expression increases slowly in early stages until the 50% RWC stage, indicating a slower trend in contrast to profile 2-a. Since the lipid mobilization already occurred during the priming treatment, this “delayed” expression, altogether with the down-regulated module, could reflect a compensatory measure to prevent further mobilization of storage lipids until needed for the seed-seedling transition. In the NR phenotype we had lower resolution since not all stages were tested. However, these seeds did not display the early mobilization at T0, which is the one stage indicative of the effect of the priming treatment, and the transcriptomic regulation of the β-oxidation pathway was not enriched in any profile (Fig. 7D-F). On the later stages of germination, it is evident that lipid mobilization occurred as in the PR phenotype.

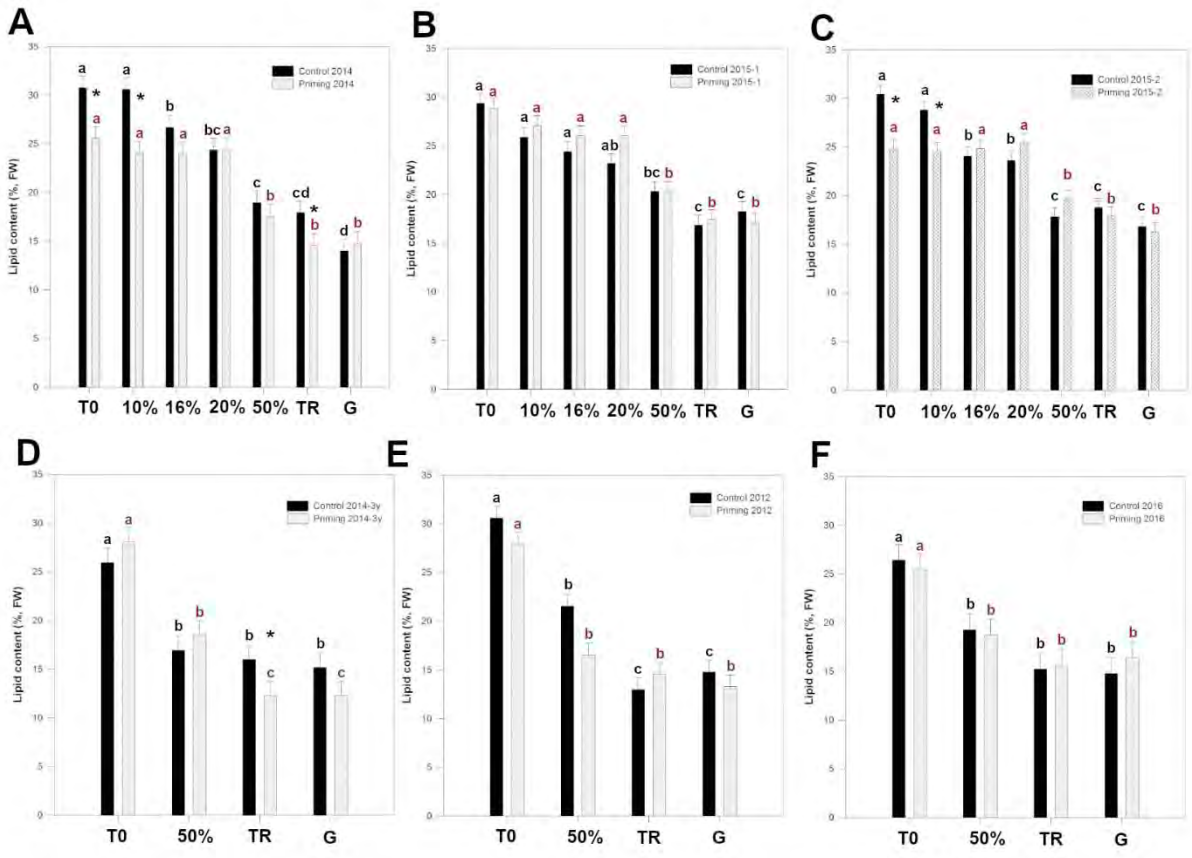


Figure 7. Total lipid content in *C. aesculifolia* seeds at different stages during germination. Panels A to C correspond to the PR phenotype, and panels D to F correspond to the NR phenotype. A two-way ANOVA was performed within each seed batch to test for differences in lipid content between control and primed seeds, as well as within physiological stages. Different letters depict statistically significant differences between physiological stages (black for control, red for priming), while the asterisks (*) depict differences between treatments at that specific stage.

Discussion

The proper timing of germination is a crucial step for plant survival. *C. aesculifolia* thrives in the TDF, characterized by a short rainy season during the summer. The seeds are usually dispersed during the dry season and are exposed to cycles of hydration-dehydration akin to the priming treatment, due to sporadic rains prior to the rainy season. In this study, we tested the effects of the priming treatment in two contrasting phenotypes, each consisting of three independent seed batches collected in different years. There are previous observations that the success of the priming treatment correlates to the tested species identity, genotype, physiology, seed batch and vigour, and the applied treatment (Paparella *et al.*, 2015). Our results align with these previous observations, as the main driver of the effect of the priming treatment was associated to the initial transcriptomic landscape of the tested batches (phenotypes, Fig. 3b). These results are also consistent with our previous description of the transcriptomic differences between untreated seeds of both phenotypes (Gómez-Maqueo *et al.*, 2020). In that previous study we observed that the main differences were allocated in the dry seed, and the transcriptomes in both phenotypes became similar as the germination program progressed. Here, the positive effects of the treatment were noticeable in the fine-tuning of gene expression patterns over the course of germination (Fig. 1), as well as in the decrease of ABA content (Fig. 5) and the early

mobilization of storage lipids at the T0 stage in the PR phenotype (Fig. 6). The NR phenotype did not show the stimulation of these processes. The profile clustering indicated subtle but important changes in expression trend in over half of the observed genes, but only a small portion of those had significant differences in net expression change (Fig. 4). The observation that the response to the treatment is species-specific and priming-method-dependent is also supported by a low overlap of specific genes in our dataset with those generated by previous studies in *B. oleracea* (Soeda *et al*., 2005), *B. napus* (Kubala *et al*., 2015) and in Arabidopsis (Sano *et al.*, 2017; Supporting Information Fig. S8). Although there was low overlap among studies, the biological functions associated to their respective gene identities showed more similarities. Thus, in the case of the whole gene set of *C. aesculifolia* that showed changes in the expression trend in response to the treatment, the enriched terms reflected key processes within the germination program: DNA repair, primary metabolism, antioxidant mechanisms, ribosome biogenesis, and the cell cycle (Dekkers *et al*., 2013; Kubala *et al.*, 2015; Paparella *et al.*, 2015; Weitbrecht *et al.*, 2011). Also, the gene set of DE genes in the PR phenotype seemed to be related to the species-specific environmental stresses that could have been experienced by the mother plants, as these specific GO-terms were not shared with the other published datasets. Thus, the general conclusion from these studies is that priming does not make major changes to the inherent germination program, but rather anticipates part of the germination program and promotes the regulation of key processes.

In *C. aesculifolia* seeds, the major transcriptome changes occur during the early stages of imbibition, when the seed water content is still low. Changes in ABA content occurred in both phenotypes at this stage, which is highly variable in duration between treatments and among phenotypes. In species with physiological dormancy such as *Arabidopsis*, barley, sunflower, and *Polygonum aviculare*, the reduction of the initial ABA content during imbibition is a required mechanism for dormancy release and germination to occur, although the total ABA content in the dry seed varies among species (Millar *et al.*, 2006; Benech-Arnold *et al*. 2006; Xia *et al.*, 2018; Laspina *et al*., 2020). This reduction was particularly noticeable in the untreated seeds of the PR-phenotype, in which the initial ABA content was rapidly reduced in the following stage and remained relatively stable afterwards. The PR primed seeds showed the lowest ABA content at T0, but there is a fluctuating increase in the following stages until reaching equivalent levels as in the untreated 50%RWC seeds. This suggests that in *C. aesculifolia* ABA levels need to be maintained to ensure proper ABA signalling and coordination of responses to environmental stimuli. The fluctuation in ABA partially corresponds with the overall expression trend of ABA-metabolic genes. In general, the PR-phenotype had an up-regulated expression towards the end of germination, and the NR-phenotype already displayed higher expression levels at T0 stage that decreased and became similar to the PR-seeds. Xerico, an important regulator of ABA synthesis during osmotic and drought stresses in vegetative tissues, is highly abundant in T0 NR-seeds and could be involved in the difference in ABA levels between the phenotypes. In Arabidopsis (Ko, Yang and Han, 2006) and rice (Zeng, Hou, Xiao and Liu, 2015) it has been demonstrated an enhancement of ABA levels in transgenic plants overexpressing Xerico. These plants also displayed enhanced transcript levels of NCED genes, which are known to catalyse a rate-limiting step in the ABA synthesis pathway (Thompson *et al*., 2007).

The regulation of lipid metabolism begins at the transcriptomic level during germination, though storage lipids are mainly mobilized after radicle protrusion, during seedling establishment (Graham, 2008). In *C. aesculifolia* we observed this tight regulation, in which lipid content was maintained at high levels during early stages, until reaching the 50% RWC (Fig. 7). In the PR-

phenotype, this pattern was accompanied by the up-regulation of diverse lipid-related processes, from which the most relevant was lipid catabolism, and regulatory processes such as protection against ROS and detoxification mechanisms (Fig. 6; Table S3). In the NR-phenotype lipid-related processes were also detected, however it was noticeable the underrepresentation of lipid catabolism and detoxification mechanisms. Instead, the untreated NR-seeds displayed an enrichment in the metabolism of membrane lipids (Fig. 6) and toxic lipid-derived compounds such as methylglyoxal (Li, 2016; Hoque *et al*., 2016; Table S4). Meanwhile, the primed NR-seeds displayed the lowest diversity of lipid-related processes, and no enrichment of lipid catabolism through β-oxidation (Fig. 6; Table S4). Several authors have observed that storage oil breakdown during germination is not a requirement for successful radicle protrusion, but peroxisomal β-oxidation does affect germination performance (Graham, 2008; Kelly *et al.*, 2011). In *C. aesculifolia*, the positive response to priming was associated to a faster germination, as well as a reduction of storage lipids at T0. This suggests that lipid quantification could be a fast and easy test alongside germination performance assessment in *C. aesculifolia* seed batches. These tests could aid in the decision-making process of whether to use a certain batch for the plant propagation needed in restoration programs. It is an open question whether the early mobilization of storage lipids is a species-specific phenomenon, of if it could be a common process of species with lipid-rich seeds.

Seed batch quality is fundamental for successful restoration programs and, in the case of *C. aesculifolia*, the effects of storage over germination take several years to be noticeable. The 2012-5y batch is the batch with the longest time in storage, but only the t_{50} has shown a statistically significant difference with respect to its germination performance when freshly collected (~11 and ~9 days, respectively Fig. S9). This seed batch showed a lower final germination when freshly collected in contrast to the 2014 and 2015 seed batches. These t_{50} values are within the observed variation across batches; thus, the combined testing of priming response with continuous testing over time could aid in better assessing the quality of a seed batch. A similar trend was observed in the 2014 seed batch, which showed a lower GI after 3 years of storage, but the other tested germination attributes were like freshly collected seeds (Fig. 1). However, the major difference between the assessments performed in this batch (2014, and 2014-3y) is associated to the loss of the response to priming, which in turn is associated to a shift in its overall transcriptomic landscape during its permanence in storage.

The 2016 seed batch was tested when freshly collected, but the negative response to the priming treatment and lack of lipid mobilization at T0 also correlated to a low germination performance. The 2012, 2016 and 2017 batches developed in dryer years, receiving about 70% of the expected annual rain for the region (García, 2004), which could have generated stressful conditions during seed development. In *C. aesculifolia*, water availability determines tree phenology and growth, particularly the timing of leaf flush and abscission. Also, unlike other related species in the Malvaceae family, this species flowers by the end of the dry season and its fruits are retained until the next dry season (Lobo *et al*., 2003). Thus, the integration of external signals into the germination program during maturation could impair the perception and response to the adequate conditions for germination and seedling establishment.

In species with physiological dormancy such as *Arabidopsis* it has been described a "germination checkpoint" mediated by ABA, which controls whether a seed will transition from primary dormant, to non-dormant, or to secondary dormancy prior to the commitment to continue the germination program (Catusse *et al.*, 2008). In *C. aesculifolia*, this checkpoint could deal with the proper ABA signalling related to the water imbibition process during the sporadic

rains, and seems to be influenced by the maturation environment. The successful timing of germination is dependent of the adequate perception of the rainy season in the TDF. However, this ecosystem is currently at risk of alterations in precipitation regimes, as well as in soil properties due to anthropogenic disturbances and climate change (Dantas *et al*., 2020; Pulla *et al.*, 2015). These alterations not only have been demonstrated to have negative impacts on germination, as it was reported in several species of the Brazilian TDF (Dantas *et al*., 2020), but also are expected to have cascading effects on plant physiology and survival at later developmental stages (Pulla *et al*., 2019). Thus, it will be necessary to further understand the physiological mechanisms that control the proper integration of external signals to ensure the survival of new individuals in a rapidly changing environment. Further analyses are also needed to understand the relevance of the response to the hydration-dehydration cycles in later developmental stages.

Acknowledgements
This work was supported by Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT UNAM) [grant IN210621 to AGB], and by the Project OE001 "*Ceiba aesculifolia* como especie modelo para el estudio molecular del envejecimiento de las semillas en relación al desempeño germinativo" financed by the Dirección General del Sector Primario y Recursos Naturales Renovables within Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) and by the Comisión Nacional para la Biodiversidad (CONABIO). This paper constitutes a partial fulfilment of the Posgrado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM). X. Gómez-Maqueo was supported by the scholarship provided by the Consejo Nacional de Ciencia y Tecnología (CONACyT) [grant 280483] and by Dirección General de Asuntos del Personal Académico (DGAPA UNAM) [grant IN203218]. We thank Drs. Alejandro Sánchez-Flores, Ricardo Grande, Verónica Jiménez-Jacinto, and Leticia Vega-Alvarado from Unidad de Secuenciación Masiva y Bioinformática, Instituto de Biotecnología, UNAM, for their support in the transcriptome assembly and bioinformatics. We thank Rodrigo García-Herrera and José Miguel Baltazar Galvez for their technical assistance with High-throughput computing services at the LANCIS-Instituto de Ecología, and Instituto de Ecología UNAM respectively. We thank David Cruz-Sánchez for his technical assistance with the relational data base generated for this dataset. We thank Laura Figueroa-Corona for her assistance in GO-term network construction.

Competing interests
The authors declare no competing interests.

Author contributions: XGM designed and performed experiments, analysed the data and wrote the manuscript. DS performed experiments, analysed the data and reviewed the manuscript. SAL, performed experiments and analysed the germination data. EMB assisted in the interpretation of the data and reviewed the manuscript. CFO assisted in lipid profiling and reviewed the manuscript. RBA assisted in the design of ABA experiments, the interpretation of the data, and reviewed the manuscript. AGB designed the experiments, analysed the data and wrote the manuscript.

CAPÍTULO V. El estudio de la biología de las semillas de especies silvestres desde la perspectiva molecular y ecofiosiológica.

La germinación representa la culminación del programa de desarrollo de las semillas y se ve afectada por las condiciones ambientales que experimenta la planta madre durante el proceso de maduración. Durante la maduración, se establecen tanto la condición de latencia como la tolerancia a la deshidratación. Estas características están moduladas por el medio al que están sometidas, teniendo un impacto importante sobre las especies silvestres. En este trabajo se realizó una revisión de las bases moleculares de la maduración, los procesos de imposición y pérdida de latencia, así como el proceso de germinación en diferentes especies silvestres con diferentes historias de vida, y provenientes de diversos hábitats. También se especifica cuáles de estas especies presentan cierto tipo de manejo o procesos de domesticación. Se discute el impacto que ha tenido el proceso de domesticación sobre determinadas características de la semilla, así como la importancia de determinar etapas fisiológicas con base a características morfológicas, para afrontar las complejidades del estudio de estas especies y preservar su diversidad genética y respuestas fisiológicas. Este ejercicio bibliográfico demostró que, hasta la fecha, los estudios genéticos y moleculares del proceso de maduración siguen siendo realizados en especies modelo como *Arabidopsis* y en especies domesticadas. Sin embargo, el advenimiento de técnicas de secuenciación masiva de nueva generación ha permitido el análisis de la germinación en especies silvestres, es en estas especies donde se observa un mayor interés por determinar estadios fisiológicos para poder hacer análisis comparativos de la fisiología de la semilla, en lugar de usar la edad cronológica. Asimismo, resalta la necesidad de promover el estudio de especies silvestres, debido a su baja representación en bancos de germoplasma ex situ y a la necesidad de promover la variabilidad genética de los lotes almacenados en estos bancos. El comprender los mecanismos fisiológicos asociados a la maduración y germinación de las semillas con énfasis en las especies silvestres será necesario para promover la adaptabilidad y supervivencia de las especies ante la crisis actual de cambio climático.

plants

MDPI

Review

The Relevance of a Physiological-Stage Approach Study of the Molecular and Environmental Factors Regulating Seed Germination in Wild Plants

Ximena Gómez-Maqueo [1], Laura Figueroa-Corona [1], Jorge Arturo Martínez-Villegas [1], Diana Soriano [2] and Alicia Gamboa-deBuen [1,*]

[1] Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; ximenagmb@ciencias.unam.mx (X.G.-M.); laurafc25@gmail.com (L.F.-C.); jorge.mtzvillegas@ciencias.unam.mx (J.A.M.-V.)

[2] Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; dsoriano@ciencias.unam.mx

* Correspondence: agamboa@ecologia.unam.mx

Abstract: Germination represents the culmination of the seed developmental program and is affected by the conditions prevailing during seed maturation in the mother plant. During maturation, the dormancy condition and tolerance to dehydration are established. These characteristics are modulated by the environment to which they are subjected, having an important impact on wild species. In this work, a review was made of the molecular bases of the maturation, the processes of dormancy imposition and loss, as well as the germination process in different wild species with different life histories, and from diverse habitats. It is also specified which of these species present a certain type of management. The impact that the domestication process has had on certain characteristics of the seed is discussed, as well as the importance of determining physiological stages based on morphological characteristics, to face the complexities of the study of these species and preserve their genetic diversity and physiological responses.

Keywords: germination; dormancy; non-model species; wild species; plant domestication; seed physiology; seed maturation

Citation: Gómez-Maqueo, X.; Figueroa-Corona, L.; Martínez-Villegas, J.A.; Soriano, D.; Gamboa-deBuen, A. The Relevance of a Physiological-Stage Approach Study of the Molecular and Environmental Factors Regulating Seed Germination in Wild Plants. *Plants* **2021**, *10*, 1084. https://doi.org/10.3390/plants10061084

Academic Editors: Raquel Iglesias Fernández and Jesús Vicente-Carbajosa

Received: 29 April 2021
Accepted: 24 May 2021
Published: 28 May 2021

1. Introduction

The seed plays a fundamental role in the plant life cycle as a "pause" stage between successive plant generations [1]. It also functions as a resistance structure to adverse environmental conditions and as a unit for dispersal of individuals in both time and space [1]. Thus, plant populations depend on the success of a series of biological processes that are regulated by environmental conditions and that will impact on the germination process and the seed-seedling transition [2,3]. These biological processes modulated by the environment include genetic and developmental programs such as seed maturation, dormancy, longevity, timing of seed release, and germination, but also include other factors such as dispersal strategies and the effect of biotic and abiotic interactions on germination and during the seed-seedling transition [3–5].

In angiosperms, the seeds' constituents include an outer cover or testa, the endosperm and the cotyledon or cotyledons, and the embryonic axis [1]. The characteristics of each component and the relationship between each component are highly variable between the seeds of different species [1]. In dicots, the endosperm may be present or absent in the dry, mature seeds. Endospermic seeds can have one or more cell layers of this tissue surrounding the embryo, as in the case of *Arabidopsis thaliana* (L.) Heynh (Arabidopsis), and, in the case of species with folded cotyledons, the endosperm may be located in-between the folds. In general, the endosperm is abundant in monocots [6]. In dicots, there are about 10 different types of embryo morphologies, five of which does not occur in monocots [7].

In general, the endosperm is very abundant in monocots due to a smaller embryo size with respect to total seed size [6,7]. Additionally, there are six types of embryos present in monocots, two of which are exclusive to the monocot clade. One of the most known embryo types is the lateral embryo type, which occurs in the Poaceae family. In gymnosperms, the embryo can be linear and small in proportion to the total seed size, or spatulate and proportionally large in comparison with total seed size [7].

Embryogenesis is characterized by cell division and organ formation. The time required for the development of a seed is regulated by multiple environmental factors. Therefore, chronological age is a less adequate approximation than physiological age. Physiological age has been determined from easily quantifiable characteristics such as weight, size and color, in addition to a series of species-specific morphological and anatomical characteristics [8].

In general, the embryogenesis concludes and embryo growth ceases during seed maturation before seed shattering. In some plant species, the embryogenesis and maturation finish after seed dispersion. At the maturation stage, in which major reserve substances accumulate, the desiccation tolerance and dormancy status are acquired. These processes have been extensively reported to occur in the so-called orthodox seeds, which enable them to persist for long periods in naturally occurring soil seed banks, as well as in ex situ seed banks. Orthodox seeds can be quiescent or dormant. Quiescent seeds have very little metabolic activity that is activated to germinate in the appropriate environmental conditions, mainly determined by the availability of water to hydrate the seed and by temperature [1]. Dormant seeds, where the depth of dormancy is determined by environmental conditions during maturation, present special requirements for germination [9]. Recalcitrant seeds occur among different taxa and can experience a slight dehydration process during seed maturation, but are sensitive to further desiccation below a species-specific threshold and cannot be stored for prolonged periods. In natural populations, these seeds may germinate within the fruit or soon after dehiscence, and they do not persist in the soil seed bank. These types of embryos rapidly lose viability if they are dried or chilled. The term recalcitrant is generally applied to seeds that have been systematically tested to determine their ability to tolerate desiccation [10]. Germination *sensu stricto* refers to the events that begin with the hydration of the mature seed and the elongation of the radicle until its protrusion takes place [1].

This article will review the molecular bases of different processes involved in the recruitment of new individuals from a community, including the effect of environmental conditions, with special emphasis on non-model plants. It is important to point out that the use of physiological stages rather than chronological age would benefit the elucidation of the molecular and physiological mechanisms involved in seed germination performance and seedling establishment. This approach also aims to achieve a proper contrast between plant populations across genotypes, cohorts, localities and experimental treatment.

2. Molecular Bases of Seed Maturation

The different processes that occur during seed maturation are related to storage products accumulation, desiccation tolerance and seed dormancy. Seed maturation is highly regulated by the parental environmental cues such as temperature, light, nitrate and water [9].

Stressful environmental conditions during seed development also affect seed germination. *Glycine max* (L.) Merr. (soybean) seeds that were developed in plants subjected to heat stress showed a decrease in the speed of germination and, in the three determined stages (dry) seed, 6 h of imbibition and germinated, an increase in the detection and expression of genes related to abiotic stress was also detected [11].

2.1. Accumulation of Reserves

The molecular and physiological mechanisms of reserve substance storage in the seed have been extensively studied in Arabidopsis and in dicots such as legumes, and in mono-

cots. The transcription factors FUSCA, ABI3 LEC1 and LEC2 ("LAFL" network) participate in the acquisition of desiccation tolerance, and reserve accumulation has been established by means of transcriptomic studies and the use of mutants [12]. Recently, it was described that the ABI3 homologous-gene in *Arachis hypogaea* L. (peanut) is also involved during seed filling under dark conditions [13]. This network was also described in gymnosperms [14]. The synthesis and accumulation of reserve proteins is similar in different plant species regardless of the botanical group (dicotyledonous or monocotyledonous, in angiosperms or gymnosperms), the storage behavior (orthodox or recalcitrant) [1] or the life history (perennials or annuals/biennials).

The accumulation of starch in the endosperm of the Poaceae family has been extensively studied in maize mutants, and important differences have been found with that of Arabidopsis. Loss of function of the GW2 and qSW5 genes (encoding a RING-type E3 ubiquitin ligase of unknown function, and a nuclear protein of unknown function (respectively) are involved in seed size increase in *Oryza sativa* L. (rice) [15,16]. These two proteins are involved in the positive regulation of GS3, a transmembrane protein with a cysteine-rich domain that participates in the negative regulation of seed size [17].

2.2. Tolerance to Desiccation

High tolerance to desiccation, associated to an important decrease on water content, was described in seeds for a wide variety of plant species [18]. Orthodox seeds can withstand a water content between 5% and 15% of total seed weight. Recalcitrant seeds usually present a high-water content at seed maturity, between 35% and 60% of total seed weight, indicating that this type of seed is also subject to stress during seed maturation [18]. Between 7.5% and 19.6% of the world's plant species present seeds that are sensitive to desiccation [19]. These plants are present in tropical environments, temperate regions and semi-arid ecosystems [20]. Tolerance to desiccation is a complex trait that is under the control of different genes, but potentially small changes would be needed to eliminate this trait and make the seeds of a species sensitive to desiccation. This characteristic may have appeared, independently, several times in plant evolution [21].

Different processes or mechanisms were described as important factors in tolerance to dehydration [20]. The factors identified include the accumulation of protective molecules, such as LEAs, sucrose and oligopolysaccharides, such as raffinose and the "switching off" of metabolism including cell cycle activity [20].

The LEA proteins are present in both orthodox and recalcitrant seeds, as well as in angiosperms and gymnosperms [22,23]. A comparative study of LEA proteins involved in seed function between recalcitrant seeds of *Castanospermum australe* A.Cunn. and C. Fraser (Moreton Bay chestnut or blackbean) and orthodox seeds of *Medicago truncatula* Gaertn. (barrelclover) indicated that both types of seeds present four common LEA proteins with similar levels, but 12 were less abundant in blackbean seeds. These results suggest that the developmental program involved in desiccation tolerance is present in both types of seeds [22]. The accumulation of sucrose and raffinose is also involved in desiccation tolerance. The presence of transcripts involved in the synthesis of these saccharides was described in mature seeds with either orthodox or recalcitrant behavior. The reduction of metabolic activity is also essential for seed tolerance to desiccation. In orthodox seeds, there is an important decrease in the expression of transcripts involved in energy production and respiration [24]. In orthodox seeds, most cells are arrested in the G2 phase and DNA replication, one of the first events to arrest during dehydration. Recalcitrant seeds have an active metabolism when they are dispersed. Consequently, germination can potentially start, and end, without external water. In recalcitrant seeds of *Avicennia marina* (Forssk.) Vierh. (gray mangrove), a transient arrest of replication was detected no more than 24 h prior to release from the mother plant and germination [20].

The evergreen oak *(Quercus ilex* L.) is one of the most representative trees in Mediterranean forests and presents recalcitrant seeds. Dehydrin accumulation was reported in mature seeds, which are desiccated up to 38%. Additionally, there is an important accu-

mulation of transcripts related to raffinose synthesis in both evergreen and northern red oak (*Q. rubra* L.) [25,26]. In order to study the maturation and germination processes as a continuum, a proteomic study was carried out in acorns at different stages of development (early development M4–M6 and late development M8–M9, where M9 corresponds to the mature acorn) and during germination (early G1–G3 and late G4–G5). These stages were determined based on morphology, weight and water content (the lowest water content was 60% and 35% in the embryo and cotyledons, respectively, in the M9 stage). Proteomic analysis of the different stages corresponding to maturation and germination showed that in these seeds there was a metabolic continuum in the embryo between the late stages of maturation and early germination. These two stages share a highly similar proteomic profile suggesting that there is no metabolic switch between them. As it was described, the recalcitrant seeds are released with an active metabolism, possibly associated with their sensitivity to dehydration [27].

3. Molecular Bases of Dormancy

Dormancy refers to a state that is characterized by a temporary arrest of the growth and development of a plant or some of its organs. In seeds, the definition implies the absence of germination in a viable, hydrated, seed that is in favorable conditions for germination [1,28]. Dormancy is a trait that arose in many species belonging to different taxa through their evolutionary history by natural selection in environments that present unfavorable conditions such as heat, cold or drought [1]. The number of species that present dormant seeds tends to increase with geographical distance from the equator, as seasonal variation in the pattern of precipitation and temperatures increases [7]. The biological function of dormancy is mainly to disperse germination in time, which can be achieved in two ways. The first one deals with the dependence of dormancy breakage on environmental factors that, in turn, are distributed differently through the time of year. This is one of the most known mechanisms of dormancy, as it implies the requirement of an environmental factor such as a chilling period during winter for summer annuals to germinate and develop during the most favorable season [1,29]. Within a seed population, different depths of dormancy can be detected across individual seeds, but in some species these differences are more intense, leading to the second type of temporal dispersal, through the expression of polymorphisms, heteromorphy or heteroblasty. In this type of dormancy, the differences in morphology and degree of dormancy are distinguishable between seeds of the same plant or among plants of the same species. Examples of this phenomenon are mainly registered in the Asteraceae and Chenopodiaceae families and allow for different portions of the population to germinate at different times [30]. Recent studies, using two different germplasms of *Leymus chinensis* (Trin.) Tzvelev (sheepgrass) with high and low germination performance, suggest that the level of ABA at the middle and late maturation stages is involved in the germination performance [31].

Finally, a non-dormant seed is one which can germinate under the widest range of environmental conditions that are compatible with the germination requirements for a particular species [4,32].

There are several classifications and sub-classifications of dormancy based on the various mechanisms by which the block of germination can occur. The most common types of dormancy reported in the literature are physiological dormancy, morphological dormancy, morphophysiological dormancy and physical dormancy. If based on the time at which dormancy is imposed, there are two main types of dormancy: primary dormancy, which occurs as part of seed development to prevent germination in the mother plant and to favor dispersal over a longer period after its detachment from the mother plant [32,33]; and secondary dormancy, which is a state induced in a mature seed during imbibition, as a consequence of adverse environmental conditions and that can usually be expressed in seeds with non-deep physiological dormancy [1,32]. Here, we present a short description of the main types of dormancy reported in the literature. However, these types have been known to occur in different combinations in some species.

Physical dormancy. For some species, the seed coat is impermeable to water and/or gas exchange, thus the germination process is blocked. The seed coat impermeability is usually associated with the presence of one or more layers of impermeable palisade cells, and the natural openings of a seed, such as the micropyle, hilum and chalazal area, also become impermeable to water. Specialized openings, valves or plugs have evolved, and it is not until these plugs are dislodged, usually by fluctuations in soil temperature, that water can enter the seed. In some cases, ingestion by animals can also promote germination of this type of seeds [7,34]. Physical dormancy has not been described for gymnosperms but was reported in at least 18 families of angiosperms, but not all members of a family present this type of dormancy [7,35]. These families include members of the Cucurbitaceae, Malvaceae, Fabaceae, Nelumbonaceae, Rhamnaceae and Sapindaceae, among others [35].

Physiological dormancy. This is one of the most widespread and abundant type of dormancy, it deals with a hormonal regulation that prevents germination, and it is present in gymnosperms and major angiosperm clades. It is most common in species of temperate climates and in the model species in which dormancy and germination are usually studied, which include Arabidopsis, *Solanum lycopersicum* Lam. (tomato), *Nicotiana tabacum* L. (tobacco) and several cereals [4]. In species such as Arabidopsis, the depth of dormancy has been demonstrated to be affected by temperature, by nitrate deficiency, as well as by light intensity and photoperiod [9,36]. Additionally, the effect of the prevalent maturation environment was higher than the effect of the genotype [9].

Morphological dormancy. In some species, mature seeds contain a fully differentiated, but underdeveloped, embryo. In others, the embryo did not complete the process of histodifferentiation and, thus, needs to develop after its detachment from the mother plant. In either situation, the embryo needs to grow in order to protrude from the enclosing layers, and germination is delayed by this process. This type of dormancy occurs in many plant families from both temperate and tropical regions. Some representative families include Amborellaceae, Cycadaceae, Ginkgoaceae, Hydrophyllaceae, Liliaceae, Loranthaceae, Magnoliaceae, Orchidaceae, Orobancheaceae, and Ranunculaceae, among others [7].

Morphophysiological dormancy. This type of dormancy usually occurs in species with rudimentary or linear embryos in combination with a physiological block to germination. Thus, two processes must take place during germination: (1) embryo differentiation and growth, and (2) dormancy alleviation. Baskin and Baskin [7] reported that the order in which both processes occur may be species-specific.

Hilhorst [28] discussed that dormancy is not an all-or-nothing property of the seed, but a relative phenomenon whose expression varies with the environment. He pointed out some considerations that should be taken in the implementation of the term dormancy. The term depends on the determination of the "absence of germination", which is intrinsically associated with germination time: a seed can be classified as dormant or non-dormant depending on the (arbitrary) time of germination measurement. A second situation stems from the characterization of "the appropriate conditions for germination". When dormancy is characterized in a particular species, erroneous conclusions can be drawn if the analysis is performed within or outside the optimal temperature range or other species-specific requirements. Finally, it is important to understand that the different sub-classifications of the two most common types of primary dormancy, physiological and morphophysiological, have their usefulness in the conceptualization of questions from a phylogenetic, biogeographic or evolutionary perspective. However, they are not suitable for physiological and molecular studies since they are mostly arbitrary, and do not distinguish the underlying mechanisms that drive the expression of the dormancy status. This is particularly relevant since the criteria for determining the level of dormancy are subjective.

3.1. Regulation of Dormancy: The Role of Hormones and Dormancy-Specific Regulators

Seed dormancy is an important adaptive trait in the early-life history of a particular species, and it is modulated by endogenous and environmental factors [32,37]. The role of model species such as Arabidopsis and other members of the Brassicaceae family has been

fundamental for the understanding of the regulation mediated by the ABA/GA balance and antagonism, and its integration with other dormancy-specific regulators [38–40]. These specific regulators include the proteins Delay of Germination 1 (DOG1), and Reduced Dormancy 5, which are reviewed by Chahtane et al. [32], and Nee et al. [39]. The current model of the integration of the dormancy status requires both ABA and DOG1, the absence of either one leads to the loss of dormancy [37]. The ABA signaling and the DOG1 pathway converge in the inhibition of the activity of the clade A PP2C phosphatases, which regulate downstream responses that will determine whether a seed germinates or not [37]. The DOG1 gene has been studied in other species such as *Triticum aestivum* L. (wheat) [41], and *Lactuca sativa* L. (lettuce) [42], but a protein blast in the NCBI database indicates that the DOG1 and DOG-like genes have been identified in the assembled genomes available for a wider range of species, including members of the Cleomaceae, Euphorbiaceae, Malvaceae, and Rubiaceae families [43,44]. Further studies of the signaling pathway mediated by DOG1 are required in non-model species.

The participation of ABA in the induction and maintenance of dormancy has been extensively studied in different plant species [45,46]. ABA biosynthesis is enhanced in seeds with deeper dormancy phenotypes, as observed in Arabidopsis Cape Verde Island ecotype, where an important increase in genes involved in biosynthesis of ABA, such as NCED, was described during dormancy induction (reviewed in [4]). Additionally, dormancy imposition and maintenance has been observed to depend on an intrinsic balance between both the synthesis and catabolism of GA and ABA, which will determine the dominance of either hormone and its downstream signaling cascades [4,47]. The early ABA perception and signaling pathway requires the participation of three main components: the ABA receptors (PYR/PYL/RCAR); the SnRK2s protein kinases; and the PP2Cs protein phosphatases [40]. The receptors and the SnRK2.2, 2.3, and 2.6 are positive regulators of ABA signaling. The receptors bind and inhibit the activity of the PP2Cs (negative regulators of ABA signaling). This, in turn, allows for the activation of the SNRKs, which phosphorylate downstream regulators such as ABI5 and the AREB/ABFs transcription factors of the bZIP class [40]. ABI5 functions are the main inhibitor of germination [40,48]. The ABI3 transcription factor is known to physically interact with ABI5, and this interaction is important for the synergistic activation of gene expression [48]. Another positive regulator of ABI5 expression during germination is the ABI4 transcription factor, which, in turn, represses lipid breakdown in the embryo [40].

Loss of dormancy is a typical domestication syndrome that allows for the majority of the seeds in a batch to germinate upon imbibition and with little input from other stimuli. The specific genes related to loss of dormancy have been described in several models including rice, soybean, and barley. The selection of a dormancy gene during the domestication process in the Poaceae family was also described. The loss of function of the G gene, which encodes for a CAAX-amino-terminal protease involved in the regulation of ABA synthesis, determines the loss of dormancy in Poaceae, Solanaceae and legume crops [49]. Dormancy loss during domestication and cultivar management could result in a viviparous phenotype commonly known as preharvest sprouting (PHS, [50]). The control of PHS is a result of a complex signaling network that includes internal and external cues. In cereals, ABA metabolism and signaling play an important role in PHS regulation. The involvement of GA and other phytohormones, carbohydrate metabolism, reactive oxygen species (ROS), nitric oxide (NO) and microRNAs were also demonstrated [51]. Many cultivars of *Chenopodium quinoa* C.L. Willdenow (Willd.) (quinoa) are susceptible to PHS, so the effect of an exogenous ABA treatment on germination and gene expression was studied. The expression of the transcription factors from B3, bZIP, GATA and LBMD families, involved in ABA mediated transcriptional regulation during seed dormancy and germination, was increased [52].

The dormancy release process depends on seed structure or composition and environment conditions. For example, it was described that ethylene is involved in dormancy release by overcoming the inhibitory release of ABA in different plant species [53]. Alter-

nating temperatures are also relevant for dormancy release in some plant species. This procedure inhibits ABA synthesis and decreases ABA sensitivity. In *Cynara cardunculus* L. *var. sylvestris* Lam. (wild cardoon), alternating temperatures during imbibition increase germination and there is a stimulation of the expression patterns related with ethylene and ROS signaling. ABA catabolism and cell wall loosening was also increased by alternating temperatures [54].

3.2. Dormancy Break by Different Treatments

For the study of germination processes in wild plant seeds, an important obstacle can be the depth and type of dormancy. The different treatments to break dormancy can include stratification, the use of hormones such as GA3, and the use of chemicals such as NO and ethanol. In recent years, the effect of these treatments at the transcriptome level has begun to be studied to understand the mechanisms behind dormancy breakage. The mature seeds of *Paris polyphylla* Sm. (love apple) present a globular-stage embryo with morphophysiological dormancy. No germination occurs up to 18 months, but dormancy can be broken by stratification treatment. The participation of ABA in dormancy was described and high levels of ABI5 were detected in dormant seeds [55].

A transcriptomic analysis to determine the mechanisms related to primary dormancy release by stratification was performed in *Cunninghamia lanceolata* (Lamb.) Hook (chinese fir) seeds. The induction of a secondary dormancy using a heat treatment was also studied. These two processes are highly regulated by ABA/GA balance in this species. For primary dormancy release, an important decrease in the expression of RGA, a gene related to negative regulation to GA sensitivity was detected. In contrast, an important increase in the expression of ABA1, a gene related to ABA synthesis, was observed during secondary dormancy induction [56].

The seeds of *Ginkgo biloba* L. (ginko) are recalcitrant and have morphophysiological dormancy. The freshly released seeds present an embryo at heart-stage and require a cold and wet stratification process to complete development, and for the seed to germinate. For the transcriptomic study, three stages of embryo development were selected: induction of dormancy, maintenance of dormancy and release of dormancy. The comparison between transcriptomes indicates that ABA and GA3 are highly involved in dormancy maintenance and release, respectively. The phytohormones auxin and brassinosteroids (BR), associated to morphological development of the embryo, are also involved [57].

The seeds of *Fraxinus hupehensis* S.Z.Qu, C.B.Shang and P.L.Su (ash tree) present dormancy, and no germination was observed during the first year. Differential expression analyses between germinated seeds treated with GA3 and without treatment suggest that GA3 increased the production of energy. Key genes involved in the glycolytic pathway are strongly expressed in the germinated seeds with ethanol treatment [58].

For *Aconitum heterophyllum* Wall. ex Royle (atish), that grows at altitudes up to 4000 m above sea level, the seeds present morphophysiological dormancy which can be broken with an ethanol treatment. The proteomic analyses indicated that ethanol induces the germination of these seeds by increasing the synthesis of GA via LYTB-like protein and a probable acetyl-CoA acetyl transferase 2 like protein, proteins detected only in the seeds treated with ethanol. The results suggest ethanol treatment enhanced the growth potential of embryonic axis through an induction of metabolism and cell wall remodeling [59].

4. Molecular Bases of Germination

Arabidopsis, tomato and cereal crops have been the most studied plants for the elucidation of the germination process at the molecular level. The state-of-the-art of the molecular bases of germination can be consulted in Carrera-Castaño et al. [40]. Therefore, the majority of the knowledge that we now have about germination comes from annual plants. In perennial plants, these types of studies were carried out in: the poplar tree (*Poplar simonii* Carrière × *Poplar nigra* L. [60]); in cultivated species with recalcitrant seeds, such as the evergreen oak [27]; and, more recently, in wild tree species, such as *Ceiba aesculifolia*

(Kunth) Britten and Baker f. (pochote) [61], and wild orchids, such as *Calanthe tsoongiana* Tang and F.T. Wang [62].

A chronological approach has been extensively used to study the germination process at a molecular level in model plants. However, during Arabidopsis seed germination, it was demonstrated that there are two distinct morphological stages at 25 h of imbibition. At this time point, 50% of seeds presented a ruptured testa stage associated with a distinct transcriptomic profile in contrast to the other 50% of imbibed seeds, which did not have a ruptured testa yet [63]. This approach, based on morphology rather than chronological time, has been frequently applied to study germination in the seeds of several species of orchids, but also in trees such as poplar and pochote [60,61,64]. In the tropical tree pochote, a physiological stage approach instead of a chronological progression of germination was implemented by sampling seeds at specific relative water contents (rwc), and morphological cues such as testa rupture and radicle protrusion. This approach allowed the authors to describe and track the transcriptional changes that occur in seed batches collected in different years, and with distinct germination performance, thus detecting the main molecular processes associated to the germination performance observed across phenotypes, and in response to pre-germinative treatments such as priming [61].

The desiccation of the seed and the subsequent rehydration causes damage to the different macromolecules contained within. The imbibition of the seed promotes the reactivation of different DNA and protein repair mechanisms. Damage to DNA, mainly by strand breakage, affects the viability of the seeds. In the first transcriptomic phase described for Arabidopsis, there is an upregulation of the genes related to the DNA repair process, such as DNA ligases and DNA polymerases [5,63]. The seed repair mechanisms that allow proper genome maintenance was extensively reviewed by Waterworth et al. [65].

It has been determined that the mature seeds of different plant species present about 12,000 to 18,000 stored transcripts in several species including Arabidopsis, rice, and pochote [61,63,66], some of which can be subject to degradation during the early stages of imbibition. These transcripts usually correspond to genes expressed during the development and maturation of the seeds.

Likewise, during the start of water intake, the transcription of genes related to the activation of primary metabolism, including the metabolism of carbohydrates, amino acids and hormones, as well as the genes involved in the modifications of the cell wall, have been detected. In recent years, it was also described that the alternative editing of these transcripts plays a very important role during germination [67]. In pochote, the early up-regulation of genes related to spliceosome was detected mainly in the first stage of imbibition [61].

Gemination is highly regulated by the phytohormones ABA and GA3, which have antagonistic roles during this process. During germination, water uptake promotes the expression of genes related to ABA catabolism and to GA3 synthesis. The transcripts that are upregulated during seed imbibition include ABA hydrolases, that inactivate ABA, and GA oxidases, involved in GA3 synthesis [4].

The proper timing of germination is critical for the successful propagation and survival of seed plant species and is determined by external and internal factors. A molecular network that integrates different environmental stimuli, such as light or cold, with hormone signaling, mainly GA3, that result in the expression of cell wall proteins involved in germination was described in Arabidopsis [68]. In recent years, molecular studies of seed germination have been carried out with plants that require factors other than those described for model plants. Light is a fundamental factor that regulates germination as it provides information related to the photoperiod, the season of the year, the position of the seed in the soil and the presence of competitors. In general, light is perceived by phytochromes which regulate hormone levels. In Arabidopsis, red light induces the expression of GA biosynthesis genes and ABA degradation genes to induce germination, whereas the expression of the genes responsible for the degradation of GA and the synthesis of ABA decrease. The *Aethionema arabicum* Blum. (stonnecress) is an annual plant that

presents accessions with different responses to light. In particular, the germination of the CYP accession is inhibited by light. Transcriptomic studies carried out in stonnecress showed that inhibition by light involves the same regulatory elements described for a positive response to light in Arabidopsis. In the CYP accession, light promotes the synthesis of ABA and increases the degradation of GA [69].

The seeds can have an associated microbiota (mainly fungi and bacteria). The endophytic microbiota is present in the internal tissues of the seeds and is transmitted from the mother plant to the seeds and from the seeds to the seedlings. The epiphytic microbiota, found on the surface of the seeds, are obtained by their interaction with the surrounding environment after dispersal, which generally occurs by horizontal transmission. The endophytic microbiome is often distinct from the microbiome associated with the soil in which the plants have grown [70,71]. Microbiome diversity can vary from species to species, genotype to genotype, geographic location, and the presence of other microorganisms. However, endophytic bacteria can be highly conserved in some plant species [70–73].

Seed microbiome can directly impact their quality, influencing biomass accumulation, the production of metabolites, desiccation tolerance, and resistance to pathogens. The microbiome also can promote germination by releasing seeds from dormancy through cytokinin production [71].

The domestication of plants is an important determinant of the microbial diversity of seeds. The globalization of the seed trade could lead to the homogenization of the seed microbiome on a global scale. To what extent and how domestication affects the microbiome of seeds is something that needs to be studied [74,75].

In the Orchidaceae family, seed germination is promoted by mycorrhizae colonization. Orchid seeds are very small and have very few reserves, so colonization with a compatible fungus is essential for the seed-seedling transition. In the orchid *Anoectochilus roxburghii* (Wall.) Lindl. (jinxianlian), a transcriptomic analysis was carried out to determine the effect of symbiosis on the development of the first true leaf, defined by the authors as the culmination of germination. The results obtained suggested that the interaction with the fungus induced the expression of genes related to the GA signaling pathway, promoting germination [14]. For the orchid *Dendrobirum officinalis* Kimura and Migo (gold plant), the level of transcripts related to ABA and jasmonic synthesis was lower during the first true leaf stage from colonized seeds compared with non-colonized seeds [76]. In another study performed on gold plant, the colonized seeds required less time to reach the testa rupture stage with respect to the control seeds. This advancement in the timing of the transition to the other physiological stages was also observed between the colonized and non-colonized seeds. At the testa rupture stage, when comparing the molecular profile, it was concluded that the germination process was not affected by the symbiosis with the fungus, although it did have a positive influence on the subsequent developmental stages, the protocorm and first true leaf development [64]. In general, the three studies describe that the symbiotic interaction between mycorrhizal fungi and orchid seeds induce higher expression of some key proteins involved in lipid and carbohydrate metabolism, and thus improves the efficiency of utilization of stored substances present in the embryo [14,64,76].

Figure 1 shows the resulting network from the co-expression and predicted relations between the differentially expressed genes reported by Chen et al. [64] and Wang et al. [76]. In accordance with Chen et al. [64], the expression is similar under asymbiotic and symbiotic seeds during testa rupture (Figure 1A) and major changes begin during the protocorm formation stage (Figure 1B). In stage 3, the fungal interaction induced a higher and earlier expression of several disulfide-isomerase and Cucumisin proteins involved in the posttranscriptional process (blue cluster), and a higher expression of 6-phosphogluconolactonase 4 or Beta-D-xylosidase 4, enzymes related to carbohydrate metabolism (pink cluster). Additionally, the downregulation of genes, such as 2,3-bisphosphoglycerate-independent phosphoglycerate mutase and Enolase 2, was described as improving the efficiency of the utilization of stored substances.

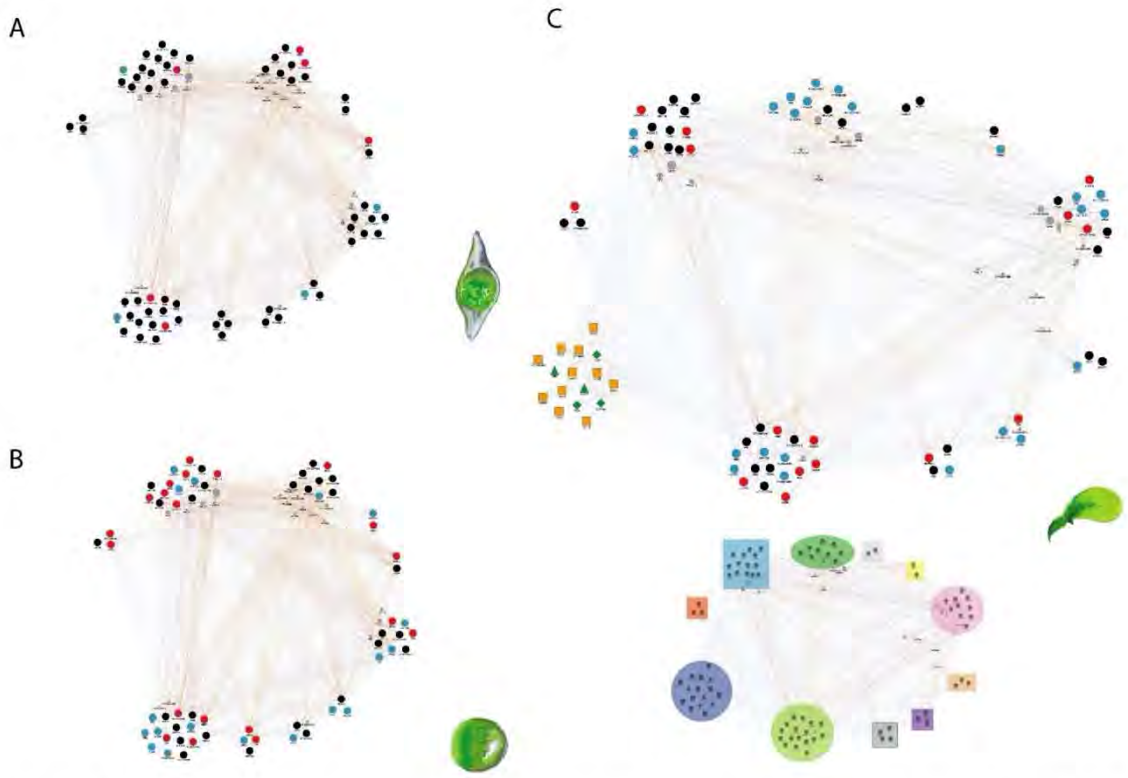


Figure 1. Network comparison between symbiotic germination (up-regulated) and asymbiotic germination (down-regulated) in three different developmental stages in *Dendrobium officinale*. The network was constructed using the differential expressed genes reported by Chen et al. [64] using the geneMANIA app (v3.5.2; [77]) in Cytoscape (v3.8.2; [78]) predicting the function on Arabidopsis gene sets using the information reported in Ronemus et al. [79] and Lee et al. [80]. (**A**) Stage 2: testa rupture (germination). (**B**) Stage 3: protomeristem appearance (protocorm). (**C**) Stage 4: emergence of the first leaf (seedling), at this stage we incorporate the phytohormonal regulation described by Wang et al. [76]. In the bottom left corner we present the color guide of the functional clustering: red, inorganic ion transport; blue, posttranslational modification; dark green, translation and ribosomal structure and biogenesis; lilac, lipid transport and metabolism; yellow, coenzyme transport and metabolism; pink, carbohydrate transport and metabolism; orange, nucleotide transport and metabolism; purple, amino acid transport and metabolism; gray, energy production and conversion; pale green, unknown/general function prediction; and dark blue, phytohormonal regulation (upregulated, orange, and downregulated, green; squares, SL regulators; triangles, JA regulators; and diamonds, ABA regulators).

At the seedling stage (first leaf, Figure 1C), Chen et al. [64] reported a downregulation of the activities in translation (green), and amino acid transport (purple), and metabolism pathways (present in different clusters). An upregulation of four genes related to photosynthesis was detected at the first leaf stage (pale green), when the transition from heterotrophy to photoautotrophy occurs. In the work presented by Wang et al. [76], they described phytohormonal regulation at the first leaf stage, thus their data were included in the network for this stage (Figure 1C). An important downregulation for ABA and JA pathways occurred in the first leaf from the symbiotic germinated seeds. In contrast, the SL pathway was upregulated, as it has been observed for other seed interactions. Figure 1C integrated the data from two different studies at the same physiological stage, the first leaf stage. This allows for other exploratory analyses based on predicted and confirmed genetic co-expressions in order to generate new hypothesis of the genetic regulation mediated by the fungal interaction.

About 1% of plant species are parasitic and are present in 19 families. These plants obtain resources, such as water, carbohydrates and minerals, from a host plant. Some parasitic species can present chlorophyll and carry out photosynthesis to a certain degree (hemiparasites), while others do not (holoparasites; [81]). In both types of parasitism, germination is characterized by the requirement of a molecular recognition signal from the host, so that the interaction can take place [82]. This molecular signal is given, mainly, by strigolactones (SLs) released by the host to the rhizosphere. The germination of parasitic plants was prominently studied in the Orobanchaceae family [83]. The seeds of broomrape species (*Orobanche spp.* L.) contain few reserves and can survive only for a few days after germination unless they reach a host root to establish a xylem connection. Germination in parasitic species requires a period of conditioning (priming), or an imbibition period, during which sensitivity to SLs is acquired. Without this conditioning period, the seeds cannot germinate, even in the presence of SLs or GAs [83]. Using a synthetic SL (GR24), studies suggest that SL generates an increase in the production of gibberellins and cytokinins, and a decrease in the production of ABA [83,84].

Transcriptomic changes associated to conditioning and to SL stimulation using GR24 were studied in *Phelipanche aegyptiaca* (Pers.) Pomel (Egyptian broomrape) seeds. DNA reparation, protein synthesis and carbohydrate metabolism processes were induced during conditioning. The expression levels of the GA20OX and GA3 genes were also increased during this treatment. After stimulation, a promotion of protein and carbohydrate metabolism and an important regulation of GA and BR biosynthesis was observed. For ABA, the gene NCED, involved in ABA synthesis, was down-regulated and the CYP707A1 gene, which participates in ABA catabolism, was upregulated [30].

5. Domestication, Seed Management and the Need for Rewilding Seed Science

With the rise of human settlements across the world, so began the management of plants through the domestication process about 12,000–6000 years ago [85]. To date, it is estimated that about 1000–2500 plant species around the world, belonging to 120–160 taxonomic families, have undergone some level of domestication or management by humans [86]. The spread of crop species outside of their wild ancestors original distribution range is a statement of a successful plant-animal mutual relationship that has increased the fitness of domesticated plants [87]. Several traits have evolved during domestication, that are distinguishable between the cultivated plants from their wild ancestors [85]; an example of this divergence in traits can be observed between the cultivated wheats and their wild ancestors in the genus Aegilops [88]. The term "domestication syndrome" was initially used to describe traits described in crop cereals, but, in recent years, its use has expanded to other domesticated species [89]. Many traits present in domesticated plants show signs of convergence across taxa, but the subjacent mechanisms from which they arose are part of the specific taxa, cultivar or landrace evolutionary history [85]. In this section, we will also refer to "domesticated plants" and to plants that have had some sort of management, selection or cultivation by humans as "managed" plants, in order to distinguish them from true wild species. This distinction is important, managed species have a wild origin, but in practice they can be subject to selection, breeding or variability reduction that does not necessarily reflect the original diversity of the population. These practices include the selection of a single plant or the "single seed descent" for subsequent seed production [90,91].

Seed research has depended heavily on crop and model species, but it is known that there are several seed characteristics that show signs of "selective sweeps" or a domestication signature. There has been a particular interest in fast and uniform germination, better germination in increased soil disturbance or depth in the soil, increased seed size and reduced seed shattering [87]. Seed traits are majorly selected at the onset of the domestication history of a species [85,86]. In Table 1 we present some of the representative seed-related genes with these signatures, and the varied molecular mechanisms that generated the change (from punctual mutations to loss of function, to changes in regulation). The frequency of the major crops of the Poaceae family is noticeable, although this could

be partially explained by the historical importance of said family. Still, some of these genes have also been reported in members of other families, especially the Brassicaceae and the Fabaceae. Thus, it is evident that there are still unexplored areas of seed science that could benefit from taking into consideration that many of the seed traits observed in the majority of the managed species might have been subjected to systematical (if unconscious) selection at some point. Additionally, as reported in a review by Kilian and Graner [90] about the status of ex situ seed banks across the globe, there is a clear need for turning efforts towards the collection of wild species, since about 50% of the total seed accessions stored in about 1700 seed banks correspond to only 10 major crop species with multiple, redundant accessions. Aside from the cultivated species present in germplasms, there has been an increasing interest in the ex situ conservation of crop wild relatives (CWR), as they are useful in breeding programs to enhance crop production and resistance to environmental factors [88,92]. Still, these efforts only amount to about 4% of European accessions present in ex situ seed banks. This estimate reflects that, despite the potential benefits of CWRs in crop management, and even their ecological importance, wild species are profoundly neglected from conservation efforts, both ex situ and in situ [88,92]. Ex situ conservation efforts pose a challenge with no straight-forward solutions, and, for the case of CWRs, several strategies have been proposed to prioritize and focus conservation efforts [88,92], but these strategies should be implemented and adjusted for the other non-crop-relative wild species, as they could fall even far behind in the priority listings. Pedrini and Dixon [91] propose a set of standard practices that can be implemented for the specific handling of wild seeds, which include prioritizing population diversity and variability, key elements to tackle environmental crises, such as the restoration of disturbed habitats and species adaptation to climate change.

Another emerging line of evidence that should be considered in the decision of studying the germination process in wild species, and their encompassing variability, comes from whole genome and transcriptome analyses through next generation sequencing technologies. In rapeseed, the expansion and diversification of spring, winter and semi-winter cultivars is tightly associated with the selection of the genetic network that controls the responses to stress and developmental processes associated to flowering time [93]. As shown in previous sections, seed development and maturation (as well as proper timing and synchronization of developmental switches, such as flowering time) are fundamental for the success of germination and seedling establishment. In recent years there has been an increasing interest in comparative transcriptomic analyses between a domesticated crop and its wild relatives to improve breeding programs. In a study performed in *Zea mays* L. (maize) and *Zea mexicana ssp. Parviglumis* (H. H. Iltis and Doebley) Greb. (teosinte), it was demonstrated that domestication has caused a considerable reshaping of the maize transcriptome in comparison to teosinte. This reshaping also caused an important reduction in the complexity of the genetic network of maize [94]. A similar result was obtained in a comparative study in tomato and its wild relatives [95], in which both transcriptome reshaping and reduction of complexity impacted in important process related to the response to environmental stimuli and stress tolerance. Although none of these studies have focused on seed germination, these studies are a statement of the impact that human management has had on plant phenotypes and environmental responses. Thus, it is important to consider the importance of these comparative studies that include wild relatives, to plan and perform seed-oriented analyses to test the extent and real impact of global genomic regulation changes on germination traits and performance.

Table 1. Representative genes associated to seed characteristics with evidence of selection during domestication and plant management. Several genes are conserved among taxa, displaying similar functions, which were also characterized in the model plant Arabidopsis through mutant screening. Some domesticated phenotypes are products of complex and lineage/cultivar-specific domestication histories; thus, the same phenotype can arise from multiple mutations. The domestication phase refers to the moment in which such characteristic would have most-likely arisen based on the discussion presented by the authors and on the four-phase model proposed by Meyer and Purugganan [85]: phase 1 onset of domestication; phase 2 diversification; phase 3 dispersion from original distribution; and phase 4 variety improvement.

Species	Gene Name	Characteristic	Type/Function	Type of Mutation	Domestication Phase	Reference
Brassica rapa	TT8	Lighter-colored testa	Transcription factor	Insertion, LOF	Diversification and dispersion	[85]
Glycine max	GmG	Reduced dormancy, lighter-colored testa	CAAX amino-terminal protease protein	Splicing defect, premature stop	Onset of domestication	[76]
Oryza sativa	qSD7-1/qPC7	Dormancy	Transcription factor	Deletion	Onset of domestication	[96,97]
Oryza sativa	Sdr4	Reduced dormancy	Not characterized	Insertion	Onset of domestication	
Oryza sativa	OsG	Reduced dormancy	CAAX amino-terminal protease protein	Non-synonymous SNP	Onset of domestication	[76]
Oryza sativa	GIF1	Reserve accumulation	Cell wall invertase	Cis-regulatory, restricted expression	Onset of domestication	[98]
Oryza sativa	Bh4	Seed color	Amino acid transporter	Deletion/LOF, premature stop	Onset of domestication	[99]
Oryza sativa	GS3	Seed size and weight increase	Transmembranal protein	Premature stop	Diversification and dispersion	[17]
Oryza sativa	OsGRF4	Seed size increase	Transcription factor	Mutation in regulator target site	Variety improvement	[100]
Oryza sativa	GS5	Seed size increase	Serine-carboxypeptidase	Indels, cis-regulatory	Diversification and improvement	[101]
Oryza sativa	GW2	Seed size increase	E3-ubiquitin ligase	Deletion, premature stop	Unknown	[15]
Oryza sativa	GW6a	Seed size/weight increase, plant biomass	Histone H4-acetyltransferase, transcription regulator	Cis-regulatory	Variety improvement	[102]
Oryza sativa	Ae1	Amylose properties	Alpha-amylase	SNPs	Onset of domestication	[103]
Panicum miliaceum	PmGBSSI	Starch properties	Granule-bound starch synthase	Deletion, LOF, frameshift, missense	Diversification and dispersion	[85]
Setaria italica	WAXY	Starch properties	Granule-bound starch synthase	Insertion LOF	Diversification and dispersion	[104]
Solanum lycopersicum	SolyG	Reduced dormancy	CAAX amino-terminal protease protein	Non-synonymous SNP	Onset of domestication	[76]

Table 1. *Cont.*

Species	Gene Name	Characteristic	Type/Function	Type of Mutation	Domestication Phase	Reference
Triticum aestivum	WAP2 (Q)	Reduced seed shattering, plant height and other pleiotropic effects	Transcription factor	Missense, cis-regulatory, premature stop	Onset of domestication	[105]
Zea mays	Bt2 (SSU)	Amylopectin properties	ADP-glucose pyrophosphorylase	Signal peptide, expression site	Onset of domestication and diversification	[106] and references therein
Zea mays	Vgt1	Flowering time QTL	AP2-like transcription factor	Cis-regulatory	Diversification and dispersion	[107]
Zea mays	HEX9	Glycolysis pathway, trans eQTL effects	Hexokinase 9	Cis-regulatory	Not determined	[76]
Zea mays	Adh2	Resistance to hypoxia	Alcohol dehydrogenase	Small tandem repeats, indels	Onset of domestication	[108]
Zea mays	Su1	Starch biosynthesis, sweetness	Isoamylase	Insertion LOF, missense, premature stop	Onset of domestication	[108]
Zea mays	DULL1	Starch synthesis	soluble starch synthase	Insertion, reduced expression	Diversification	[94]

In recent years, the number of "omics" studies performed in non-model species have increased considerably. In this review we selected some representative studies to reflect this new approach to the molecular bases of seed dormancy and germination, as well as the seed-seedling transition. In Table 2 we summarize the main characteristics of the species presented in this review, focusing on life history and habitat, storage behavior, dormancy, and management status. Of the 31 species reported, 10 correspond to wild species which reflect a wide diversity of families, habitats, life histories, storage behavior and dormancy types. This contrasts to the pattern observed for the model species, which usually belong to a few families, present orthodox seeds with physiological dormancy, come from temperate climates and have an herbaceous, annual life history. Nine species also belong to the monocot clade, but five of them are in fact from the Poaceae family. It will be important to include a wider variety of families and genera within this clade to have a better understanding of the germination process in the monocot clade.

Table 2. Origin and management of the studied species reported in this review. Abbreviations: D, dicotyledons; G, gymnosperms; M, monocotyledons; Cu, cultivated; Do, domesticated; Ma, managed; W, wild; PHS, preharvest sprouting.

Name	Family	Clade	Management	Original Climate	Life History	Storage Behavior	Dormancy
Aconitum heterophyllum	Ranunculaceae	D	W	Sub-alpine	Herbaceous, perennial	Possibly orthodox	Morpho-physiological
Aethionema arabicum	Brassicaceae	D	Ma	Semi-arid	Herbaceous, annual	Orthodox	Morphotype-dependent, physiological
Arabidopsis thaliana	Brassicaceae	D	Ma	Temperate	Herbaceous, annual	Orthodox	Physiological
Arachis hypogaea	Fabaceae	D	Do	Alpine	Herbaceous, annual	Orthodox	Physiological

Table 2. *Cont.*

Name	Family	Clade	Management	Original Climate	Life History	Storage Behavior	Dormancy
Avicennia marina	Acanthaceae	D	W	Subtropical	Shrub, tree, perennial	Recalcitrant	Viviparous
Brassica rapa	Brassicaceae	D	Do	Semi-arid	Herbaceous, annual	Orthodox	Physiological
Castanospermum australe	Fabaceae	D	Ma	Tropical, subtropical	Tree, perennial	Recalcitrant	Not determined
Ceiba aesculifolia	Malvaceae	D	W	Tropical, subtropical	Tree, perennial	Orthodox	Non dormant
Chenopodium quinoa	Amaranthaceae	D	Do	Temperate	Herbaceous, annual	Orthodox	PHS
Cynara cardunculus var. *sylvestris*	Asteraceae	D	W	Mediterranean	Herbaceous, perennial	Possibly orthodox	Physiological
Fraxinus hupehensis	Oleaceae	D	W	Temperate, no dry season	Tree, perennial	Not determined	Not determined
Glycine max	Fabaceae	D	Do	Subtropical	Herbaceous, annual	Orthodox	Non-dormant
Lactuca sativa	Asteraceae	D	Do	Mediterranean, temperate	Herbaceous, annual	Orthodox	Physiological dormancy
Medicago truncatula	Fabaceae	D	Do	Mediterranean	Herbaceous, annual	Orthodox	Physical
Phelipanche aegyptiaca	Orobanchaceae	D	W	Temperate, desertic	Herbaceous	Not determined	Not determined
Populus nigra x Populus simonii	Salicaceae	D	Do	Temperate	Tree, perennial	Possibly Orthodox	Possibly non dormant
Quercus ilex	Fagaceae	D	Cu	Mediterranean	Tree, perennial	Recalcitrant	Non dormant
Quercus robur	Fagaceae	D	Cu	Temperate	Tree, perennial	Recalcitrant	Non dormant
Nicotiana tabacum	Solanaceae	D	Cu	Tropical	Herbaceous, annual	Orthodox	Physiological
Solanum lycopersicum	Solanaceae	D	Do	Tropical	Herbaceous, annual	Orthodox	Physiological
Anoectochilus roxburghii	Orchidaceae	M	W	Temperate, no dry season	Herbaceous, perennial	Orthodox	Not determined
Calanthe tsoongiana	Orchidaceae	M	W	Temperate, subtropical	Herbaceous, perennial	Not determined	Not determined
Dendrobium officinale	Orchidaceae	M	W	Tropical, subtropical	Herbaceous, perennial	Orthodox	Not determined
Leymus chinensis	Poaceae	M	Ma?	Temperate	Herbaceous, perennial	Orthodox	Possibly physiological
Oryza sativa	Poaceae	M	Do	Temperate and tropical	Herbaceous, annual	Orthodox	Cultivar-dependent. Physiological to PHS susceptible
Paris polyphylla	Melanthiaceae	M	Cu?	Temperate	Herbaceous, perennial	Not determined	Morpho-physiological

Table 2. *Cont.*

Name	Family	Clade	Management	Original Climate	Life History	Storage Behavior	Dormancy
Triticum aestivum	Poaceae	M	Do	Semi-arid	Herbaceous, annual	Orthodox	Non-dormant, PHS susceptible
Zea mays	Poaceae	M	Do	Tropical	Herbaceous, annual	Orthodox	Non-dormant, PHS susceptible
Zea mexicana ssp. parviglumis	Poaceae	M	W	Tropical	Herbaceous, annual	Orthodox	Physiological
Cunninghamia lanceolata	Cupressaceae	G	Cu?	Temperate	Tree, perennial	Probably orthodox	Physiological
Ginkgo biloba	Ginkgoaceae	G	Cu	Mediterranean	Tree, perennial	Recalcitrant	Morpho-physiological

6. Concluding Remarks

Climate change and ecosystem deterioration are the major challenges to tackle in the coming years. As mentioned by various authors presented in this review, our capacity to ameliorate and delay the progression of these crises will depend on the understanding of the mechanisms that allow organismal adaptation and responses to the environment. Germination in non-model species represents a challenge for their study due to the wide genetic and physiological variability they present, which must also be conserved in order to face climate change. In seed maturation studies, physiological age is used to study and compare the processes between different species and maturation conditions. Considering the great variability in the germination performance of non-model species, the establishment of morphological parameters and physiological stages is essential. These approaches are exemplified for the cases of evergreen oak, poplar, pochote and gold plant. Likewise, these parameters allow the integration of the information generated by different authors for the same species, since the definition of germination may vary. To date, most of the characterization of developmental switches and phase transitions related to dormancy and germination have been studied in different model species, such as Arabidopsis. Still, the physiological/developmental-stage approach used in many of these studies can be implemented in wild species to explore the relevance of the signaling cascades reported for Arabidopsis and other systems. This characterization will be fundamental to understanding dormancy mechanisms, aside from the non-deep physiological dormancy that is displayed by the majority of the model plants studied to date.

The domestication process has reduced the diversity and complexity of genetic interactions in the species studied. Therefore, global studies that include wild species are needed to be able to study germination in relation to its interaction with the environment. Transcriptomics and proteomics comparisons across species will help provide a better understanding of seed-related processes, such as maturation, the transitions needed for commitment to germination, as well as seedling establishment. This knowledge will also provide a better approach for breeding programs and/or in biotechnological studies to improve crop yields, as well as adaptation to climate change.

Author Contributions: Conceptualization, X.G.-M., A.G.-d.; network construction and analysis, L.F.-C., X.G.-M.; writing, review and editing, X.G.-M., A.G.-d., L.F.-C., J.A.M.-V., D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT UNAM) (grant IN203218 to AGB), and by the Project OE001 "*Ceiba aesculifolia*

como especie modelo para el estudio molecular del envejecimiento de las semillas en relación al desempeño germinativo" financed by the Dirección General del Sector Primario y Recursos Naturales Renovables within Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) and by the Comisión Nacional para la Biodiversidad (CONABIO). X. Gómez-Maqueo was supported by the scholarship provided by the Consejo Nacional de Ciencia y Tecnología (CONACyT) (grant 280483) and by Dirección General de Asuntos del Personal Académico (DGAPA UNAM) (grant IN203218).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: This paper constitutes a partial fulfilment of the Posgrado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM), as part of the doctoral thesis presented by X. Gómez-Maqueo.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bewley, J.D.; Bradford, K.J.; Hilhorst, H.W.M.; Nonogaki, H. *Seeds*; Springer: New York, NY, USA, 2013; ISBN 978-1-4614-4692-7.
2. Fenner, M.; Thompson, K. *The Ecology of Seeds*, 1st ed.; Cambridge University Press: Cambridge, UK, 2005; ISBN 978-0-521-65311-4.
3. Lambers, H.; Chapin, F.S.; Pons, T.L. *Plant Physiological Ecology*; Springer: New York, NY, USA, 2008; ISBN 978-0-387-78340-6.
4. Finch-Savage, W.E.; Leubner-Metzger, G. Seed Dormancy and the Control of Germination. *New Phytol.* **2006**, *171*, 501–523. [CrossRef]
5. Long, R.L.; Gorecki, M.J.; Renton, M.; Scott, J.K.; Colville, L.; Goggin, D.E.; Commander, L.E.; Westcott, D.A.; Cherry, H.; Finch-Savage, W.E. The Ecophysiology of Seed Persistence: A Mechanistic View of the Journey to Germination or Demise: The Ecophysiology of Seed Persistence. *Biol. Rev.* **2015**, *90*, 31–59. [CrossRef]
6. McDonald, M.B.; Copeland, L.O. Seed Formation and Development. In *Seed Production*; Springer: Boston, MA, USA, 1997; pp. 19–28, ISBN 978-1-4613-6825-0.
7. Baskin, C.C.; Baskin, J.M. *Seeds: Ecology, Biogeography, and Evolution of Dormancy and Germination*; Academic Press: San Diego, CA, USA, 1998; ISBN 978-0-12-080260-9.
8. Miernyk, J.A. Seed Proteomics. In *Plant Proteomics, Methods in Molecular Biology (Methods and Protocols)*; Jorrin-Novo, J.V., Komatsu, S., Weckwerth, W., Wienkoop, S., Eds.; Humana Press: Totowa, NJ, USA, 2014; Volume 1072, pp. 361–377. ISBN 978-1-62703-630-6.
9. He, H.; Willems, L.A.J.; Batushansky, A.; Fait, A.; Hanson, J.; Nijveen, H.; Hilhorst, H.W.M.; Bentsink, L. Effects of Parental Temperature and Nitrate on Seed Performance Are Reflected by Partly Overlapping Genetic and Metabolic Pathways. *Plant Cell Physiol.* **2016**, *57*, 473–487. [CrossRef]
10. Farnsworth, E. The Ecology and Physiology of Viviparous and Recalcitrant Seeds. *Annu. Rev. Ecol. Syst.* **2000**, *31*, 107–138. [CrossRef]
11. Gillman, J.D.; Biever, J.J.; Ye, S.; Spollen, W.G.; Givan, S.A.; Lyu, Z.; Joshi, T.; Smith, J.R.; Fritschi, F.B. A Seed Germination Transcriptomic Study Contrasting Two Soybean Genotypes That Differ in Terms of Their Tolerance to the Deleterious Impacts of Elevated Temperatures during Seed Fill. *BMC Res. Notes* **2019**, *12*, 522. [CrossRef]
12. Carbonero, P.; Iglesias-Fernández, R.; Vicente-Carbajosa, J. The AFL Subfamily of B3 Transcription Factors: Evolution and Function in Angiosperm Seeds. *J. Exp. Bot.* **2017**, *68*, 871–880. [CrossRef]
13. Liu, H.; Liang, X.; Lu, Q.; Li, H.; Liu, H.; Li, S.; Varshney, R.; Hong, Y.; Chen, X. Global Transcriptome Analysis of Subterranean Pod and Seed in Peanut (*Arachis hypogaea* L.) Unravels the Complexity of Fruit Development under Dark Condition. *Sci. Rep.* **2020**, *10*, 13050. [CrossRef]
14. Liu, S.-S.; Chen, J.; Li, S.-C.; Zeng, X.; Meng, Z.-X.; Guo, S.-X. Comparative Transcriptome Analysis of Genes Involved in GA-GID1-DELLA Regulatory Module in Symbiotic and Asymbiotic Seed Germination of *Anoectochilus roxburghii* (Wall.) Lindl. (Orchidaceae). *Int. J. Mol. Sci.* **2015**, *16*, 30190–30203. [CrossRef]
15. Song, X.-J.; Huang, W.; Shi, M.; Zhu, M.-Z.; Lin, H.-X. A QTL for Rice Grain Width and Weight Encodes a Previously Unknown RING-Type E3 Ubiquitin Ligase. *Nat. Genet.* **2007**, *39*, 623–630. [CrossRef]
16. Shomura, A.; Izawa, T.; Ebana, K.; Ebitani, T.; Kanegae, H.; Konishi, S.; Yano, M. Deletion in a Gene Associated with Grain Size Increased Yields during Rice Domestication. *Nat. Genet.* **2008**, *40*, 1023–1028. [CrossRef]
17. Yan, S.; Zou, G.; Li, S.; Wang, H.; Liu, H.; Zhai, G.; Guo, P.; Song, H.; Yan, C.; Tao, Y. Seed Size Is Determined by the Combinations of the Genes Controlling Different Seed Characteristics in Rice. *Theor. Appl. Genet.* **2011**, *123*, 1173–1181. [CrossRef] [PubMed]
18. Marques, A.; Buijs, G.; Ligterink, W.; Hilhorst, H. Evolutionary Ecophysiology of Seed Desiccation Sensitivity. *Funct. Plant Biol.* **2018**, *45*, 1083–1095. [CrossRef]
19. Wyse, S.V.; Dickie, J.B. Predicting the Global Incidence of Seed Desiccation Sensitivity. *J. Ecol.* **2017**, *105*, 1082–1093. [CrossRef]
20. Berjak, P.; Pammenter, N.W. From *Avicennia* to *Zizania*: Seed Recalcitrance in Perspective. *Ann. Bot.* **2008**, *101*, 213–228. [CrossRef]
21. Dickie, J.B.; Pritchard, H.W. Systematic and evolutionary aspects of desiccation tolerance in seeds. In *Desiccation and Survival in Plants: Drying without Dying*; Black, M., Pritchard, H.W., Eds.; CABI: Wallingford, UK, 2002; pp. 239–259, ISBN 978-0-85199-534-2.

22. Delahaie, J.; Hundertmark, M.; Bove, J.; Leprince, O.; Rogniaux, H.; Buitink, J. LEA Polypeptide Profiling of Recalcitrant and Orthodox Legume Seeds Reveals ABI3-Regulated LEA Protein Abundance Linked to Desiccation Tolerance. *J. Exp. Bot.* **2013**, *64*, 4559–4573. [CrossRef]
23. Artur, M.A.S.; Zhao, T.; Ligterink, W.; Schranz, E.; Hilhorst, H.W.M. Dissecting the Genomic Diversification of Late Embryogenesis Abundant (LEA) Protein Gene Families in Plants. *Genome Biol. Evol.* **2019**, *11*, 459–471. [CrossRef]
24. Kijak, H.; Ratajczak, E. What Do We Know About the Genetic Basis of Seed Desiccation Tolerance and Longevity? *Int. J. Mol. Sci.* **2020**, *21*, 3612. [CrossRef]
25. Miguel, A.; de Vega-Bartol, J.; Marum, L.; Chaves, I.; Santo, T.; Leitão, J.; Varela, M.C.; Miguel, C.M. Characterization of the Cork Oak Transcriptome Dynamics during Acorn Development. *BMC Plant Biol.* **2015**, *15*, 158. [CrossRef] [PubMed]
26. Romero-Rodríguez, M.C.; Jorrín-Novo, J.V.; Castillejo, M.A. Toward Characterizing Germination and Early Growth in the Non-Orthodox Forest Tree Species *Quercus ilex* through Complementary Gel and Gel-Free Proteomic Analysis of Embryo and Seedlings. *J. Proteom.* **2019**, *197*, 60–70. [CrossRef]
27. Sghaier-Hammami, B.; Hammami, S.B.M.; Baazaoui, N.; Gómez-Díaz, C.; Jorrín-Novo, J.V. Dissecting the Seed Maturation and Germination Processes in the Non-Orthodox *Quercus ilex* Species Based on Protein Signatures as Revealed by 2-DE Coupled to MALDI-TOF/TOF Proteomics Strategy. *IJMS* **2020**, *21*, 4870. [CrossRef]
28. Hilhorst, H.W.M. Definitions and Hypotheses of Seed Dormancy. In *Seed Development, Dormancy and Germination*; Bradford, K.J., Nonogaki, H., Eds.; Blackwell Publishing Ltd: Oxford, UK, 2007; pp. 50–71, ISBN 978-0-470-98884-8.
29. Benech-Arnold, R.L.; Sánchez, R.A.; Forcella, F.; Kruk, B.C.; Ghersa, C.M. Environmental Control of Dormancy in Weed Seed Banks in Soil. *Field Crops Res.* **2000**, *67*, 105–122. [CrossRef]
30. Yao, Z.; Tian, F.; Cao, X.; Xu, Y.; Chen, M.; Xiang, B.; Zhao, S. Global Transcriptomic Analysis Reveals the Mechanism of *Phelipanche aegyptiaca* Seed Germination. *IJMS* **2016**, *17*, 1139. [CrossRef]
31. Li, X.; Liu, S.; Yuan, G.; Zhao, P.; Yang, W.; Jia, J.; Cheng, L.; Qi, D.; Chen, S.; Liu, G. Comparative Transcriptome Analysis Provides Insights into the Distinct Germination in Sheepgrass (*Leymus chinensis*) during Seed Development. *Plant Physiol. Biochem.* **2019**, *139*, 446–458. [CrossRef]
32. Chahtane, H.; Kim, W.; Lopez-Molina, L. Primary Seed Dormancy: A Temporally Multilayered Riddle Waiting to Be Unlocked. *EXBOTJ* **2016**, erw377. [CrossRef]
33. Penfield, S. Seed Dormancy and Germination. *Curr. Biol.* **2017**, *27*, R874–R878. [CrossRef] [PubMed]
34. Baskin, J.M.; Baskin, C.C. Evolutionary Considerations of Claims for Physical Dormancy-Break by Microbial Action and Abrasion by Soil Particles. *Seed Sci. Res.* **2000**, *10*, 409–413. [CrossRef]
35. Gama-Arachchige, N.S.; Baskin, J.M.; Geneve, R.L.; Baskin, C.C. Identification and Characterization of Ten New Water Gaps in Seeds and Fruits with Physical Dormancy and Classification of Water-Gap Complexes. *Ann. Bot.* **2013**, *112*, 69–84. [CrossRef] [PubMed]
36. He, H.; de Souza Vidigal, D.; Snoek, L.B.; Schnabel, S.; Nijveen, H.; Hilhorst, H.; Bentsink, L. Interaction between Parental Environment and Genotype Affects Plant and Seed Performance in Arabidopsis. *J. Exp. Bot.* **2014**, *65*, 6603–6615. [CrossRef]
37. Soppe, W.J.J.; Bentsink, L. Seed Dormancy Back on Track; Its Definition and Regulation by DOG1. *New Phytol.* **2020**, *228*, 816–819. [CrossRef]
38. Graeber, K.; Linkies, A.; Steinbrecher, T.; Mummenhoff, K.; Tarkowská, D.; Turečková, V.; Ignatz, M.; Sperber, K.; Voegele, A.; de Jong, H.; et al. DELAY OF GERMINATION 1 Mediates a Conserved Coat-Dormancy Mechanism for the Temperature- and Gibberellin-Dependent Control of Seed Germination. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E3571. [CrossRef]
39. Née, G.; Xiang, Y.; Soppe, W.J. The Release of Dormancy, a Wake-up Call for Seeds to Germinate. *Curr. Opin. Plant Biol.* **2017**, *35*, 8–14. [CrossRef]
40. Carrera-Castaño, G.; Calleja-Cabrera, J.; Pernas, M.; Gómez, L.; Oñate-Sánchez, L. An Updated Overview on the Regulation of Seed Germination. *Plants* **2020**, *9*, 703. [CrossRef]
41. Ashikawa, I.; Mori, M.; Nakamura, S.; Abe, F. A Transgenic Approach to Controlling Wheat Seed Dormancy Level by Using Triticeae *DOG1*-like Genes. *Transgenic Res.* **2014**, *23*, 621–629. [CrossRef]
42. Huo, H.; Wei, S.; Bradford, K.J. *DELAY OF GERMINATION1* (*DOG1*) Regulates Both Seed Dormancy and Flowering Time through MicroRNA Pathways. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E2199. [CrossRef]
43. BLAST. National Library of Medicine (US) National Center for Biotechnology Information Bethesda (MD). 2004. Available online: https//ncbi.nlm.nih.gov/gene/ (accessed on 18 May 2021).
44. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
45. Barrero, J.M.; Jacobsen, J.V.; Talbot, M.J.; White, R.G.; Swain, S.M.; Garvin, D.F.; Gubler, F. Grain Dormancy and Light Quality Effects on Germination in the Model Grass *Brachypodium distachyon*. *New Phytol.* **2012**, *193*, 376–386. [CrossRef]
46. Cantoro, R.; Crocco, C.D.; Benech-Arnold, R.L.; Rodríguez, M.V. In Vitro Binding of *Sorghum bicolor* Transcription Factors ABI4 and ABI5 to a Conserved Region of a GA 2-OXIDASE Promoter: Possible Role of This Interaction in the Expression of Seed Dormancy. *J. Exp. Bot.* **2013**, *64*, 5721–5735. [CrossRef]
47. Topham, A.T.; Taylor, R.E.; Yan, D.; Nambara, E.; Johnston, I.G.; Bassel, G.W. Temperature Variability Is Integrated by a Spatially Embedded Decision-Making Center to Break Dormancy in *Arabidopsis* Seeds. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6629–6634. [CrossRef] [PubMed]

48. Holdsworth, M.J.; Bentsink, L.; Soppe, W.J.J. Molecular Networks Regulating *Arabidopsis* Seed Maturation, After-ripening, Dormancy and Germination. *New Phytol.* **2008**, *179*, 33–54. [CrossRef] [PubMed]
49. Wang, M.; Li, W.; Fang, C.; Xu, F.; Liu, Y.; Wang, Z.; Yang, R.; Zhang, M.; Liu, S.; Lu, S.; et al. Parallel Selection on a Dormancy Gene during Domestication of Crops from Multiple Families. *Nat. Genet.* **2018**, *50*, 1435–1441. [CrossRef] [PubMed]
50. Gubler, F.; Millar, A.A.; Jacobsen, J.V. Dormancy Release, ABA and Pre-Harvest Sprouting. *Curr. Opin. Plant Biol.* **2005**, *8*, 183–187. [CrossRef] [PubMed]
51. Tai, L.; Wang, H.-J.; Xu, X.-J.; Sun, W.-H.; Ju, L.; Liu, W.-T.; Li, W.-Q.; Sun, J.; Chen, K.-M. Pre-Harvest Sprouting in Cereals: Genetic and Biochemical Mechanisms. *J. Exp. Bot.* **2021**, *72*, 2857–2876. [CrossRef]
52. Wu, Q.; Bai, X.; Wu, X.; Xiang, D.; Wan, Y.; Luo, Y.; Shi, X.; Li, Q.; Zhao, J.; Qin, P.; et al. Transcriptome Profiling Identifies Transcription Factors and Key Homologs Involved in Seed Dormancy and Germination Regulation of *Chenopodium quinoa*. *Plant Physiol. Biochem.* **2020**, *151*, 443–456. [CrossRef] [PubMed]
53. Xia, Q.; Saux, M.; Ponnaiah, M.; Gilard, F.; Perreau, F.; Huguet, S.; Balzergue, S.; Langlade, N.; Bailly, C.; Meimoun, P.; et al. One Way to Achieve Germination: Common Molecular Mechanism Induced by Ethylene and After-Ripening in Sunflower Seeds. *IJMS* **2018**, *19*, 2464. [CrossRef] [PubMed]
54. Huarte, H.R.; Puglia, G.; Prjibelski, A.; Raccuia, S.A. Seed Transcriptome Annotation Reveals Enhanced Expression of Genes Related to ROS Homeostasis and Ethylene Metabolism at Alternating Temperatures in Wild Cardoon. *Plants* **2020**, *9*, 1225. [CrossRef] [PubMed]
55. Qi, J.; Sun, P.; Liao, D.; Sun, T.; Zhu, J.; Li, X. Transcriptomic Analysis of American Ginseng Seeds during the Dormancy Release Process by RNA-Seq. *PLoS ONE* **2015**, *10*, e0118558. [CrossRef]
56. Cao, D.; Xu, H.; Zhao, Y.; Deng, X.; Liu, Y.; Soppe, W.J.J.; Lin, J. Transcriptome and Degradome Sequencing Reveals Dormancy Mechanisms of *Cunninghamia lanceolata* Seeds. *Plant Physiol.* **2016**, *172*, 2347. [CrossRef]
57. Jia, Z.; Zhao, B.; Liu, S.; Lu, Z.; Chang, B.; Jiang, H.; Cui, H.; He, Q.; Li, W.; Jin, B.; et al. Embryo Transcriptome and MiRNA Analyses Reveal the Regulatory Network of Seed Dormancy in *Ginkgo biloba*. *Tree Physiol.* **2021**, *41*, 571–588. [CrossRef]
58. Song, Q.; Cheng, S.; Chen, Z.; Nie, G.; Xu, F.; Zhang, J.; Zhou, M.; Zhang, W.; Liao, Y.; Ye, J. Comparative Transcriptome Analysis Revealing the Potential Mechanism of Seed Germination Stimulated by Exogenous Gibberellin in *Fraxinus hupehensis*. *BMC Plant Biol.* **2019**, *19*, 199. [CrossRef]
59. Rana, B.; Sreenivasulu, Y. Protein Changes during Ethanol Induced Seed Germination in *Aconitum heterophyllum*. *Plant Sci.* **2013**, *198*, 27–38. [CrossRef]
60. Qu, C.; Zuo, Z.; Cao, L.; Huang, J.; Sun, X.; Zhang, P.; Yang, C.; Li, L.; Xu, Z.; Liu, G. Comprehensive Dissection of Transcript and Metabolite Shifts during Seed Germination and Post-Germination Stages in Poplar. *BMC Plant Biol.* **2019**, *19*, 279. [CrossRef]
61. Gómez-Maqueo, X.; Soriano, D.; Velázquez-Rosas, N.; Alvarado-López, S.; Jiménez-Durán, K.; del Mar Garciadiego, M.; Gamboa-deBuen, A. The Seed Water Content as a Time-Independent Physiological Trait during Germination in Wild Tree Species Such as *Ceiba aesculifolia*. *Sci. Rep.* **2020**, *10*, 10429. [CrossRef] [PubMed]
62. Jiang, Y.; Tian, M.; Wang, C.; Zhang, Y. Transcriptome Sequencing and Differential Gene Expression Analysis Reveal the Mechanisms Involved in Seed Germination and Protocorm Development of *Calanthe tsoongiana*. *Gene* **2021**, *772*, 145355. [CrossRef]
63. Dekkers, B.J.W.; Pearce, S.; van Bolderen-Veldkamp, R.P.; Marshall, A.; Widera, P.; Gilbert, J.; Drost, H.-G.; Bassel, G.W.; Müller, K.; King, J.R.; et al. Transcriptional Dynamics of Two Seed Compartments with Opposing Roles in Arabidopsis Seed Germination. *Plant Physiol.* **2013**, *163*, 205. [CrossRef] [PubMed]
64. Chen, J.; Liu, S.S.; Kohler, A.; Yan, B.; Luo, H.M.; Chen, X.M.; Guo, S.X. ITRAQ and RNA-Seq Analyses Provide New Insights into Regulation Mechanism of Symbiotic Germination of *Dendrobium officinale* Seeds (Orchidaceae). *J. Proteome Res.* **2017**, *16*, 2174–2187. [CrossRef]
65. Waterworth, W.M.; Bray, C.M.; West, C.E. The Importance of Safeguarding Genome Integrity in Germination and Seed Longevity. *J. Exp. Bot.* **2015**, *66*, 3549–3558. [CrossRef] [PubMed]
66. Howell, K.A.; Narsai, R.; Carroll, A.; Ivanova, A.; Lohse, M.; Usadel, B.; Millar, A.H.; Whelan, J. Mapping Metabolic and Transcript Temporal Switches during Germination in Rice Highlights Specific Transcription Factors and the Role of RNA Instability in the Germination Process. *Plant Physiol.* **2009**, *149*, 961–980. [CrossRef]
67. Tognacca, R.S.; Servi, L.; Hernando, C.E.; Saura-Sanchez, M.; Yanovsky, M.J.; Petrillo, E.; Botto, J.F. Alternative Splicing Regulation During Light-Induced Germination of *Arabidopsis thaliana* Seeds. *Front. Plant Sci.* **2019**, *10*, 1076. [CrossRef]
68. Xu, H.; Lantzouni, O.; Bruggink, T.; Benjamins, R.; Lanfermeijer, F.; Denby, K.; Schwechheimer, C.; Bassel, G.W. A Molecular Signal Integration Network Underpinning Arabidopsis Seed Germination. *Curr. Biol.* **2020**, *30*, 3703–3712.e4. [CrossRef]
69. Mérai, Z.; Graeber, K.; Wilhelmsson, P.; Ullrich, K.K.; Arshad, W.; Grosche, C.; Tarkowská, D.; Turečková, V.; Strnad, M.; Rensing, S.A.; et al. *Aethionema arabicum*: A Novel Model Plant to Study the Light Control of Seed Germination. *J. Exp. Bot.* **2019**, *70*, 3313–3328. [CrossRef] [PubMed]
70. Nelson, E.B. The Seed Microbiome: Origins, Interactions, and Impacts. *Plant Soil* **2018**, *422*, 7–34. [CrossRef]
71. Shade, A.; Jacques, M.-A.; Barret, M. Ecological Patterns of Seed Microbiome Diversity, Transmission, and Assembly. *Curr. Opin. Microbiol.* **2017**, *37*, 15–22. [CrossRef]
72. Gundel, P.E.; Rudgers, J.A.; Ghersa, C.M. Incorporating the Process of Vertical Transmission into Understanding of Host–Symbiont Dynamics. *Oikos* **2011**, *120*, 1121–1128. [CrossRef]
73. Frank, A.; Saldierna Guzmán, J.; Shay, J. Transmission of Bacterial Endophytes. *Microorganisms* **2017**, *5*, 70. [CrossRef]

74. Van der Heijden, M.G.; de Bruin, S.; Luckerhoff, L.; van Logtestijn, R.S.; Schlaeppi, K. A Widespread Plant-Fungal-Bacterial Symbiosis Promotes Plant Biodiversity, Plant Nutrition and Seedling Recruitment. *ISME J.* **2016**, *10*, 389–399. [CrossRef] [PubMed]
75. Berg, G.; Raaijmakers, J.M. Saving Seed Microbiomes. *ISME J.* **2018**, *12*, 1167–1170. [CrossRef]
76. Wang, T.; Song, Z.; Wang, X.; Xu, L.; Sun, Q.; Li, L. Functional Insights into the Roles of Hormones in the *Dendrobium officinale-tulasnella* sp. Germinated Seed Symbiotic Association. *Int. J. Mol. Sci.* **2018**, *19*, 3484. [CrossRef] [PubMed]
77. Warde-Farley, D.; Donaldson, S.L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C.T.; et al. The GeneMANIA Prediction Server: Biological Network Integration for Gene Prioritization and Predicting Gene Function. *Nucleic Acids Res.* **2010**, *38*, W214–W220. [CrossRef] [PubMed]
78. Shannon, P. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef]
79. Ronemus, M.; Vaughn, M.W.; Martienssen, R.A. MicroRNA-Targeted and Small Interfering RNA–Mediated MRNA Degradation Is Regulated by Argonaute, Dicer, and RNA-Dependent RNA Polymerase in *Arabidopsis*. *Plant Cell* **2006**, *18*, 1559–1574. [CrossRef]
80. Lee, I.; Ambaru, B.; Thakkar, P.; Marcotte, E.M.; Rhee, S.Y. Rational Association of Genes with Traits Using a Genome-Scale Gene Network for *Arabidopsis thaliana*. *Nat. Biotechnol.* **2010**, *28*, 149–156. [CrossRef] [PubMed]
81. Irving, L.J.; Cameron, D.D. Chaper 3 You are What You Eat: Interactions Between Root Parasitic Plants and Their Hosts. In *Advances in Botanical Research*; Academic Press: Cambridge, MA, USA, 2009; Volume 50, pp. 87–138, ISBN 0065-2296.
82. Westwood, J.H.; Yoder, J.I.; Timko, M.P.; de Pamphilis, C.W. The Evolution of Parasitism in Plants. *Trends Plant Sci.* **2010**, *15*, 227–235. [CrossRef]
83. Brun, G.; Braem, L.; Thoiron, S.; Gevaert, K.; Goormachtig, S.; Delavault, P. Seed Germination in Parasitic Plants: What Insights Can We Expect from Strigolactone Research? *J. Exp. Bot.* **2018**, *69*, 2265–2280. [CrossRef]
84. Toh, S.; Kamiya, Y.; Kawakami, N.; Nambara, E.; McCourt, P.; Tsuchiya, Y. Thermoinhibition Uncovers a Role for Strigolactones in Arabidopsis Seed Germination. *Plant Cell Physiol.* **2012**, *53*, 107–117. [CrossRef] [PubMed]
85. Meyer, R.S.; Purugganan, M.D. Evolution of Crop Species: Genetics of Domestication and Diversification. *Nat. Rev. Genet.* **2013**, *14*, 840–852. [CrossRef]
86. Purugganan, M.D. Evolutionary Insights into the Nature of Plant Domestication. *Curr. Biol.* **2019**, *29*, R705–R714. [CrossRef]
87. Purugganan, M.D.; Fuller, D.Q. The Nature of Selection during Plant Domestication. *Nature* **2009**, *457*, 843–848. [CrossRef]
88. Perrino, E.V.; Wagensommer, R.P.; Medagli, P. *Aegilops* (Poaceae) in Italy: Taxonomy, Geographical Distribution, Ecology, Vulnerability and Conservation. *Syst. Biodivers.* **2014**, *12*, 331–349. [CrossRef]
89. Allaby, R.G. Domestication Syndrome in Plants. In *Encyclopedia of Global Archaeology*; Smith, C., Ed.; Springer: New York, NY, USA, 2014; pp. 2182–2184, ISBN 978-1-4419-0426-3.
90. Kilian, B.; Graner, A. NGS Technologies for Analyzing Germplasm Diversity in Genebanks. *Brief. Funct. Genom.* **2012**, *11*, 38–50. [CrossRef]
91. Pedrini, S.; Dixon, K.W. International Principles and Standards for Native Seeds in Ecological Restoration. *Restor. Ecol.* **2020**, *28*. [CrossRef]
92. Heywood, V.; Casas, A.; Ford-Lloyd, B.; Kell, S.; Maxted, N. Conservation and Sustainable Use of Crop Wild Relatives. *Agric. Ecosyst. Environ.* **2007**, *121*, 245–255. [CrossRef]
93. Wu, D.; Liang, Z.; Yan, T.; Xu, Y.; Xuan, L.; Tang, J.; Zhou, G.; Lohwasser, U.; Hua, S.; Wang, H.; et al. Whole-Genome Resequencing of a Worldwide Collection of Rapeseed Accessions Reveals the Genetic Basis of Ecotype Divergence. *Mol. Plant* **2019**, *12*, 30–43. [CrossRef] [PubMed]
94. Swanson-Wagner, R.; Briskine, R.; Schaefer, R.; Hufford, M.B.; Ross-Ibarra, J.; Myers, C.L.; Tiffin, P.; Springer, N.M. Reshaping of the Maize Transcriptome by Domestication. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11878. [CrossRef] [PubMed]
95. Koenig, D.; Jiménez-Gómez, J.M.; Kimura, S.; Fulop, D.; Chitwood, D.H.; Headland, L.R.; Kumar, R.; Covington, M.F.; Devisetty, U.K.; Tat, A.V.; et al. Comparative Transcriptomics Reveals Patterns of Selection in Domesticated and Wild Tomato. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E2655. [CrossRef] [PubMed]
96. Gu, X.-Y.; Foley, M.E.; Horvath, D.P.; Anderson, J.V.; Feng, J.; Zhang, L.; Mowry, C.R.; Ye, H.; Suttle, J.C.; Kadowaki, K.; et al. Association Between Seed Dormancy and Pericarp Color Is Controlled by a Pleiotropic Gene That Regulates Abscisic Acid and Flavonoid Synthesis in Weedy Red Rice. *Genetics* **2011**, *189*, 1515–1524. [CrossRef]
97. Sugimoto, K.; Takeuchi, Y.; Ebana, K.; Miyao, A.; Hirochika, H.; Hara, N.; Ishiyama, K.; Kobayashi, M.; Ban, Y.; Hattori, T.; et al. Molecular Cloning of Sdr4, a Regulator Involved in Seed Dormancy and Domestication of Rice. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 5792–5797. [CrossRef] [PubMed]
98. Wang, E.; Wang, J.; Zhu, X.; Hao, W.; Wang, L.; Li, Q.; Zhang, L.; He, W.; Lu, B.; Lin, H.; et al. Control of Rice Grain-Filling and Yield by a Gene with a Potential Signature of Domestication. *Nat. Genet.* **2008**, *40*, 1370–1374. [CrossRef] [PubMed]
99. Poets, A.M.; Fang, Z.; Clegg, M.T.; Morrell, P.L. Barley Landraces Are Characterized by Geographically Heterogeneous Genomic Origins. *Genome Biol.* **2015**, *16*, 173. [CrossRef]
100. Duan, P.; Ni, S.; Wang, J.; Zhang, B.; Xu, R.; Wang, Y.; Chen, H.; Zhu, X.; Li, Y. Regulation of OsGRF4 by OsmiR396 Controls Grain Size and Yield in Rice. *Nat. Plants* **2015**, *2*, 15203. [CrossRef] [PubMed]
101. Li, Y.; Fan, C.; Xing, Y.; Jiang, Y.; Luo, L.; Sun, L.; Shao, D.; Xu, C.; Li, X.; Xiao, J.; et al. Natural Variation in GS5 Plays an Important Role in Regulating Grain Size and Yield in Rice. *Nat. Genet.* **2011**, *43*, 1266–1269. [CrossRef]

102. Song, X.J.; Kuroha, T.; Ayano, M.; Furuta, T.; Nagai, K.; Komeda, N.; Segami, S.; Miura, K.; Ogawa, D.; Kamura, T.; et al. Rare Allele of a Previously Unidentified Histone H4 Acetyltransferase Enhances Grain Weight, Yield, and Plant Biomass in Rice. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 76–81. [CrossRef] [PubMed]
103. Zhu, B.-F.; Si, L.; Wang, Z.; Jingjie Zhu, Y.Z.; Shangguan, Y.; Lu, D.; Fan, D.; Li, C.; Lin, H.; Qian, Q.; et al. Genetic Control of a Transition from Black to Straw-White Seed Hull in Rice Domestication. *Plant Physiol.* **2011**, *155*, 1301. [CrossRef] [PubMed]
104. Fukunaga, K.; Kawase, M.; Kato, K. Structural Variation in the Waxy Gene and Differentiation in Foxtail Millet [*Setaria italica* (L.) P. Beauv.]: Implications for Multiple Origins of the Waxy Phenotype. *Mol. Gen. Genom.* **2002**, *268*, 214–222. [CrossRef]
105. Simons, K.J.; Fellers, J.P.; Trick, H.N.; Zhang, Z.; Tai, Y.-S.; Gill, B.S.; Faris, J.D. Molecular Characterization of the Major Wheat Domestication Gene Q. *Genetics* **2006**, *172*, 547. [CrossRef] [PubMed]
106. Corbi, J.; Debieu, M.; Rousselet, A.; Montalent, P.; Le Guilloux, M.; Manicacci, D.; Tenaillon, M.I. Contrasted Patterns of Selection since Maize Domestication on Duplicated Genes Encoding a Starch Pathway Enzyme. *Theor. Appl. Genet.* **2011**, *122*, 705–722. [CrossRef]
107. Salvi, S.; Sponza, G.; Morgante, M.; Tomes, D.; Niu, X.; Fengler, K.A.; Meeley, R.; Ananiev, E.V.; Svitashev, S.; Bruggemann, E.; et al. Conserved Noncoding Genomic Sequences Associated with a Flowering-Time Quantitative Trait Locus in Maize. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11376. [CrossRef]
108. Hufford, M.B.; Xu, X.; van Heerwaarden, J.; Pyhäjärvi, T.; Chia, J.-M.; Cartwright, R.A.; Elshire, R.J.; Glaubitz, J.C.; Guill, K.E.; Kaeppler, S.M.; et al. Comparative Population Genomics of Maize Domestication and Improvement. *Nat. Genet.* **2012**, *44*, 808–811. [CrossRef] [PubMed]

CONCLUSIONES

- El transcriptoma de las semillas puede ser alterado durante el proceso de maduración debido a condiciones adversas experimentadas por la planta madre o por el tiempo en almacenamiento en condiciones subóptimas. Este perfil transcriptómico de la semilla seca es el determinante principal de la capacidad para responder e incorporar los estímulos ambientales al programa de germinación.
- El contenido relativo de agua (CRA) refleja la progresión del programa de germinación, mostrando un ajuste importante de dicho programa en estadios tempranos de hidratación. Estos cambios sugieren que, en *C. aesculifolia*, el mecanismo molecular que regula la decisión de culminar el proceso de germinación ocurre en etapas tempranas de hidratación, previo al estadio fisiológico de ruptura de testa.
- El CRA es una estrategia confiable y reproducible para rastrear los eventos asociados al proceso de germinación en *C. aesculifolia*. Esta estrategia permite contrastar lotes con desempeño germinativo distintos, independientemente del origen de dichas diferencias.
- La actividad de las PME en las semillas de C. aesculifolia muestra un incremento paulatino durante la germinación, alcanzando un punto máximo de actividad cercano al momento de ruptura de la testa. Esta dinámica concuerda con lo observado en otras especies modelo.
- Las semillas de *C. aesculifolia* presentan un contenido alto de lípidos, los cuales son movilizados en etapas tardías de la germinación, para proveer de energía a la nueva plántula durante el establecimiento y la transición hacia la autotrofía.
- La regulación transcriptómica de las vías de metabolismo de lípidos correlaciona con los estadios fisiológicos en los que ocurre la movilización de lípidos durante la germinación.
- La respuesta positiva al acondicionamiento está asociada con una regulación del catabolismo de lípidos, potencialmente por medio de la vía de β-oxidación. Esta regulación genera una movilización temprana de lípidos, observable en la semilla seca post-tratamiento.
- Las semillas que no responden al acondicionamiento no mostraron un enriquecimiento de la vía de β-oxidación durante la germinación temprana, en cambio, mostraron un enriquecimiento importante de vías de metabolismo de lípidos de membrana.
- Las semillas que incorporan exitosamente el estímulo hídrico del acondicionamiento con el programa de germinación presentan una movilización temprana de lípidos. La cuantificación de esta movilización temprana podría ser un indicador fácil y rápido en su implementación para determinar la calidad y el potencial de germinación de un lote dado.
- El acondicionamiento mátrico genera un ajuste fino de las vías genéticas que constituyen el programa de germinación, generando únicamente un cambio neto en la expresión de un número reducido de genes.
- Los genes que muestran un cambio en la expresión neta en respuesta al acondicionamiento están asociados a respuestas a estímulos externos como defensa contra hongos, estrés osmótico o estrés oxidativo.
- Las semillas que no responden favorablemente al acondicionamiento muestran una dinámica de hidratación similar a las semillas con respuesta favorable al tratamiento. Sin embargo, la fase de baja hidratación presenta mayor variabilidad en su duración, llegando a ser más prolongada.

- Las semillas con respuesta positiva al acondicionamiento mostraron una disminución de los niveles de ABA endógenos durante la hidratación temprana, mientras que el fenotipo negativo mostró un aumento incluso superior a los niveles de ABA observados en la semilla seca. Esta dinámica podría constituir parte del mecanismo regulador de la toma de decisión de concretar el proceso de germinación.
- El estudio de la germinación a nivel genético y molecular en especies silvestres aún es escaso. Sin embargo, los estudios en estadios fisiológicos discretos en lugar de tiempo cronológico han permitido el avance de la disciplina en diversas especies. Es fundamental la promoción del estudio de más especies silvestres, debido a que la mayoría de los sistemas modelo pertenecen a unas cuantas especies domesticadas, anuales, de ambientes templados y con latencia fisiológica.

REFERENCIAS

- Alvarado-López, S., Soriano, D., Velázquez, N., Orozco-Segovia, A., & Gamboa-deBuen, A. (2014). Priming effects on seed germination in *Tecoma stans* (Bignoniaceae) and *Cordia megalantha* (Boraginaceae), two tropical deciduous tree species. *Acta oecologica*, *61*, 65-70.
- Becerra-Vázquez, Á. G., Sánchez-Nieto, S., Coates, R., Flores-Ortiz, C. M., & Orozco-Segovia, A. (2018). Seed longevity of five tropical species from south-eastern Mexico: changes in seed germination during storage. *Tropical Conservation Science*, *11*, 1940082918779489.
- Benech-Arnold, R. L., Gualano, N., Leymarie, J., Côme, D., & Corbineau, F. (2006). Hypoxia interferes with ABA metabolism and increases ABA sensitivity in embryos of dormant barley grains. *Journal of Experimental Botany*, *57*, 1423-1430.
- Bewley, J. D., Bradford, K. J., Hilhorst, H. W. M., & Nonogaki, H. (2013). Seeds: physiology of development, germination and dormancy., 3rd ed.(Springer: New York).
- Bewley, J.D. 1997. Seed germination and dormacy. *The plant cell*. 9:1055-1066.
- Bligh, E.G. y W.J. Dyer. 1959. A rapid method of total lipid extraction and purification. *Canadian journal of biochemistry and physiology*, 37(8): 911-917
- Butler, L. H., Hay, F. R., Ellis, R. H., Smith, R. D., & Murray, T. B. (2009). Priming and re-drying improve the survival of mature seeds of Digitalis purpurea during storage. *Annals of Botany*, *103*(8), 1261-1270.
- Cacao Genome Database. Disponible en: www.cacaogenomedb.org. Fecha de visita: marzo 2017
- Catusse, J., Job, C., & Job, D. (2008). Transcriptome-and proteome-wide analyses of seed germination. *Comptes rendus biologies*, *331*(10), 815-822.
- Chen, K., & Arora, R. (2013). Priming memory invokes seed stress-tolerance. *Environmental and Experimental Botany*, *94*, 33-45.
- Cheng, J., Wang, L., Zeng, P., He, Y., Zhou, R., Zhang, H., & Wang, Z. (2017). Identification of genes involved in rice seed priming in the early imbibition stage. *Plant Biology*, *19*(1), 61-69.
- Climate Prediction Center Internet Team, NOAA/National Weather Service. Disponible en: https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php?fbclid=IwAR3IYm1KGN2XlD5f9bLLXvl39Cx2IK-ms2O1bDvc1vGjsIBqtZs9v9JTvhA última fecha de visita: mayo 2021
- Cosgrove, D. J. (2016). Plant cell wall extensibility: connecting plant cell growth with cell wall structure, mechanics, and the action of wall-modifying enzymes. *Journal of experimental botany*, *67*(2), 463-476.
- Cotton Database Resources. Disponible en: www.cottongen.org. Fecha de visita: marzo 2017.
- Dantas, B. F., Moura, M. S., Pelacani, C. R., Angelotti, F., Taura, T. A., Oliveira, G. M., ... & Seal, C. E. (2020). Rainfall, not soil temperature, will limit the seed germination of dry forest species with climate change. *Oecologia*, *192*(2), 529-541.
- DAVID Bioinformatics Resources 6.8. Disponible en: https://david.ncifcrf.gov/ última fecha de visita: mayo 2020

- Dekkers, B. J., Pearce, S., van Bolderen-Veldkamp, R. P., Marshall, A., Widera, P., Gilbert, J., ... & Bentsink, L. (2013). Transcriptional dynamics of two seed compartments with opposing roles in Arabidopsis seed germination. *Plant physiology*, *163*(1), 205-215.
- Encino-Ruiz, L., Lindig-Cisneros, R., Gómez-Romero, M., & Blanco-García, A. (2013). Desempeño de tres especies arbóreas del bosque tropical caducifolio en un ensayo de restauración ecológica. *Botanical Sciences*, *91*(1), 107-114.
- Elkin, L. A., Kay, M., Higgins, J. J., & Wobbrock, J. O. (2021). An Aligned Rank Transform Procedure for Multifactor Contrast Tests. *arXiv preprint arXiv:2102.11824*.
- Ernst, J., & Bar-Joseph, Z. (2006). STEM: a tool for the analysis of short time series gene expression data. *BMC bioinformatics*, *7*(1), 191.
- Ernst, J., Nau, G. J., & Bar-Joseph, Z. (2005). Clustering short time series gene expression data. *Bioinformatics*, *21*(suppl_1), i159-i168.
- Fenner, M. y K. Thompson. 2005. *The ecology of seeds*. Cambridge University Press. 260p.
- Finch-Savage, W. E., & Leubner-Metzger, G. (2006). Seed dormancy and the control of germination. *New phytologist*, *171*(3), 501-523.
- Gallardo, K., Job, C., Groot, S. P., Puype, M., Demol, H., Vandekerckhove, J., & Job, D. (2001). Proteomic analysis of Arabidopsis seed germination and priming. *Plant physiology*, *126*(2), 835-848.
- Gamboa-deBuen, A., Cruz-Ortega, R., Martínez-Barajas, E., Sánchez-Coronado, M. E., & Orozco-Segovia, A. (2006). Natural priming as an important metabolic event in the life history of *Wigandia urens* (Hydrophyllaceae) seeds. *Physiologia Plantarum*, *128*(3), 520-530.
- Garciadiego San Juan M. (2014). Expresión génica y actividad de pectin metil esterasa en semillas de *Ceiba aesculifolia* tratadas con acondicionamiento mátrico. Tesis de licenciatura. Facultad de Ciencias, UNAM. 50p.
- Gómez-Maqueo X. (2014). Efecto del acondicionamiento mátrico sobre vías genéticas de semillas almacenadas de especies silvestres. Tesis de Maestría, Instituto de Ecología, UNAM. 86p.
- Gómez-Maqueo, X., Soriano, D., Velázquez-Rosas, N., Alvarado-López, S., Jiménez-Durán, K., del Mar Garciadiego, M., & Gamboa-deBuen, A. (2020). The seed water content as a time-independent physiological trait during germination in wild tree species such as Ceiba aesculifolia. *Scientific reports*, *10*(1), 1-14.
- González-Zertuche, L., Vázquez-Yanes, C., Gamboa, A., Sánchez-Coronado, M. E., Aguilera, P., & Orozco-Segovia, A. (2001). Natural priming of *Wigandia urens* seeds during burial: effects on germination, growth and protein expression. *Seed Science Research*, *11*(1), 27.
- Graham, I. A. (2008). Seed storage oil mobilization. *Annual Review of Plant Biology*, *59*, 115-142.
- Hernández-Gómez M. 2010. Efectos en el desempeño de *Ceiba aesculifolia* en condiciones de estrés hídrico con fines de res- tauración ecológica. Tesis de licenciatura, Universidad Michoacana de San Nicolás de Hidalgo. 43 pp.
- Hoque, T. S., Hossain, M. A., Mostofa, M. G., Burritt, D. J., Fujita, M., & Tran, L. S. P. (2016). Methylglyoxal: an emerging signaling molecule in plant abiotic stress responses and tolerance. *Frontiers in plant science*, *7*, 1341.

- Howell, K. A., Narsai, R., Carroll, A., Ivanova, A., Lohse, M., Usadel, B., ... & Whelan, J. (2009). Mapping metabolic and transcript temporal switches during germination in rice highlights specific transcription factors and the role of RNA instability in the germination process. *Plant physiology*, *149*(2), 961-980.
- Huang D. W., Sherman B.T. & Lempicki R. A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research. 37,* 1-13.
- Huang D. W., Sherman B.T. & Lempicki R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protocols. 4,* 44-57.
- IDEAMEX. Disponible en: http://zazil.ibt.unam.mx/ideamex/ última fecha de visita: mayo 2020
- Joosen, R. V. L., Arends, D., Li, Y., Willems, L. A., Keurentjes, J. J., Ligterink, W., ... & Hilhorst, H. W. (2013). Identifying genotype-by-environment interactions in the metabolism of germinating Arabidopsis seeds using generalized genetical genomics. *Plant physiology*, *162*(2), 553-566.
- Kanehisa, M. (2019). Toward understanding the origin and evolution of cellular organisms. *Protein Science*, *28*(11), 1947-1951.
- Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, *28*, 27-30.
- Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., & Tanabe, M. (2021). KEGG: integrating viruses and cellular organisms. *Nucleic acids research*, *49*, D545-D551.
- Kelly, A. A., Quettier, A. L., Shaw, E., & Eastmond, P. J. (2011). Seed storage oil mobilization is important but not essential for germination or seedling establishment in Arabidopsis. *Plant Physiology*, *157*(2), 866-875.
- Ko, J. H., Yang, S. H., & Han, K. H. Upregulation of an Arabidopsis RING-H2 gene, XERICO, confers drought tolerance through increased abscisic acid biosynthesis. *Plant J.* **47,** 343-355 (2006).
- Kubala, S., Garnczarska, M., Wojtyla, Ł., Clippe, A., Kosmala, A., Żmieńko, A., ... & Quinet, M. (2015). Deciphering priming-induced improvement of rapeseed (Brassica napus L.) germination through an integrated transcriptomic and proteomic approach. *Plant Science*, *231*, 94-113.
- Kumar L., & Futschik M. Mfuzz: a software package for soft clustering of microarray data. *Bioinformation.* **2,** 5-7 (2007).
- Kyoto Encyclopedia of Genes and Genomes (KEGG). Disponible en: https://www.genome.jp/kegg/pathway.html última fecha de visita: mayo 2021
- Laspina, N. V., Batlla, D., & Benech-Arnold, R. L. (2020). Dormancy cycling in *Polygonum aviculare* seeds is accompanied by changes in ABA sensitivity. *Journal of Experimental Botany*.
- Levesque-Tremblay, G., Müller, K., Mansfield, S. D., & Haughn, G. W. (2015). HIGHLY METHYL ESTERIFIED SEEDS is a pectin methyl esterase involved in embryo development. *Plant physiology*, *167*(3), 725-737.
- Li, Z. & Trick, H. (2005). Rapid method for high-quality RNA isolation from seed endosperm containing high levels of starch. *BioTechniques* **38,** 872-876.

- Li, ZG. Methylglyoxal and Glyoxalase System in Plants: Old Players, New Concepts. *Bot. Rev.* **82,** 183–203 (2016). https://doi.org/10.1007/s12229-016-9167-9
- Lobo, J. A., Quesada, M., Stoner, K. E., Fuchs, E. J., Herrerías-Diego, Y., Rojas, J., & Saborío, G. (2003). Factors affecting phenological patterns of bombacaceous trees in seasonal forests in Costa Rica and Mexico. *American Journal of Botany*, *90*(7), 1054-1063.
- Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550; 10.1186/s13059-014-0550-8 (2014).
- Maia, J., Dekkers, B. J., Provart, N. J., Ligterink, W., & Hilhorst, H. W. (2011). The re-establishment of desiccation tolerance in germinated *Arabidopsis thaliana* seeds and its associated transcriptome. *PloS one*, *6*(12), e29123.
- Masetto, T. E., & Scalon, S. D. P. Q. (2014). Drying and osmotic conditioning in *Hancornia speciosa* Gomes seeds. *Floresta e Ambiente*, *21*(1), 62-68.
- Millar, A. A., Jacobsen, J. V., Ross, J. J., Helliwell, C. A., Poole, A. T., Scofield, G., ... & Gubler, F. (2006). Seed dormancy and ABA metabolism in Arabidopsis and barley: the role of ABA 8′-hydroxylase. *The Plant Journal*, *45*(6), 942-954.
- Moreira de Carvahlo, N. y Nakagawa, J. (1988). *Semillas ciencia, tecnología y producción*. Editorial Agropecuaria Hemisferio Sur S. R. L., Uruguay. 406pp.
- Müller, K., Levesque-Tremblay, G., Bartels, S., Weitbrecht, K., Wormit, A., Usadel, B., ... & Kermode, A. R. (2013). Demethylesterification of cell wall pectins in Arabidopsis plays a role in seed germination. *Plant physiology*, *161*(1), 305-316.
- Nakabayashi, K., Okamoto, M., Koshiba, T., Kamiya, Y., & Nambara, E. (2005). Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: epigenetic and genetic regulation of transcription in seed. *The plant journal*, *41*(5), 697-709.
- Narsai, R., Law, S. R., Carrie, C., Xu, L., & Whelan, J. (2011). In-depth temporal transcriptome profiling reveals a crucial developmental switch with roles for RNA processing and organelle metabolism that are essential for germination in Arabidopsis. *Plant physiology*, *157*(3), 1342-1362.
- Paparella, S., Araújo, S. S., Rossi, G., Wijayasinghe, M., Carbonera, D., & Balestrazzi, A. (2015). Seed priming: state of the art and new perspectives. *Plant cell reports*, *34*(8), 1281-1293.
- Peraza-Villarreal, H., . Sánchez-Coronado, M. E., Lindig-Cisneros, R., Tinoco-Ojanguren, C., Velázquez-Rosas, N., Cámara-Cabrales, L., & Orozco-Segovia, A. (2018). Seed priming effects on germination and seedling establishment of useful tropical trees for ecological restoration. *Tropical Conservation Science*, *11*, 1940082918817886.
- Priestley, D.A. 1986. *Seed aging: Implications for seed storage and persistence in the soil*. New York: Cornell University Press.
- Probert, R. J., Bogh, S. V., Smith, A. J., & Wechsberg, G. E. (1991). The effects of priming on seed longevity in Ranunculus sceleratus L. *Seed Science Research*, *1*(4), 243-249.
- Pulla, S., Ramaswami, G., Mondal, N., Chitra-Tarak, R., Suresh, H. S., Dattaraja, H. S., ... & Sukumar, R. (2015). Assessing the resilience of global seasonally dry tropical forests. *International Forestry Review*, *17*(2), 91-113.

- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ (2016).
- Rajjou, L., Duval, M., Gallardo, K., Catusse, J., Bally, J., Job, C., & Job, D. (2012). Seed germination and vigor. *Annual review of plant biology*, *63*, 507-533.
- Rosental, L., Nonogaki, H., & Fait, A. (2014). Activation and regulation of primary metabolism during seed germination. *Seed science research*, *24*(1), 1-15.
- Royal Botanic Gardens Kew. (2020) Seed Information Database (SID). Version 7.1. Disponible en: http://data.kew.org/sid/ última fecha de visita: enero 2020
- Sano, N., Kim, J. S., Onda, Y., Nomura, T., Mochida, K., Okamoto, M., & Seo, M. (2017). RNA-Seq using bulked recombinant inbred line populations uncovers the importance of brassinosteroid for seed longevity after priming treatments. *Scientific Reports*, *7*(1), 1-14.
- Santos, M. A. C. D., Queiróz, M. A. D., Bispo, J. D. S., & Dantas, B. F. (2015). Seed germination of Brazilian guava (Psidium guineense Swartz.). *Journal of Seed Science*, *37*(4), 214-221.
- Scheler, C., Weitbrecht, K., Pearce, S. P., Hampstead, A., Büttner-Mainik, A., Lee, K. J., ... & Leubner-Metzger, G. (2015). Promotion of testa rupture during garden cress germination involves seed compartment-specific expression and activity of pectin methylesterases. *Plant physiology*, *167*(1), 200-215.
- Soeda, Y., Konings, M. C., Vorst, O., van Houwelingen, A. M., Stoopen, G. M., Maliepaard, C. A., ... & van der Geest, A. H. (2005). Gene expression programs during *Brassica oleracea* seed maturation, osmopriming, and germination are indicators of progression of the germination process and the stress tolerance level. *Plant physiology*, *137*(1), 354-368.
- Steinbach, H. S., Benech–Arnold, R. L., Kristof, G., Sánchez, R. A., & Marcucci-Poltri, S. (1995). Physiological basis of pre-harvest sprouting resistance in *Sorghum bicolor* (L.) Moench. ABA levels and sensitivity in developing embryos of sprouting-resistant and-susceptible varieties. *Journal of Experimental Botany*, *46*(6), 701-709.
- Thompson, A. J., Mulholland, B. J., Jackson, A. C., McKee, J. M., Hilton, H. W., Symonds, R. C., ... & Taylor, I. B. (2007). Regulation and manipulation of ABA biosynthesis in roots. *Plant, cell & environment*, *30*(1), 67-78.
- Valle-Díaz, O., Blanco-García, A., Bonfil, C., Paz, H., & Lindig-Cisneros, R. (2009). Altitudinal range shift detected through seedling survival of Ceiba aesculifolia in an area under the influence of an urban heat island. *Forest Ecology and Management*, *258*(7), 1511-1515.
- Varier, A., Vari, A. K., & Dadlani, M. (2010). The subcellular basis of seed priming. *Current Science*, 450-456.
- Vázquez-Yanes, C. & Batis, M. A. I. (1996). La restauración de la vegetación, árboles exóticos vs árboles nativos. *Ciencias.* 43: 16-23.
- Weitbrecht, K., Müller, K., & Leubner-Metzger, G. (2011). First off the mark: early seed germination. *Journal of experimental botany*, *62*(10), 3289-3309.
- Wobbrock, J.O., Findlater, L., Gergle, D. and Higgins, J.J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. Proceedings of the ACM Conference on Human Factors in Computing

Systems (CHI '11). Vancouver, British Columbia (May 7-12, 2011). New York: ACM Press, pp. 143-146.

- Xia, Q., Saux, M., Ponnaiah, M., Gilard, F., Perreau, F., Huguet, S., ... & El-Maarouf-Bouteau, H. (2018). One way to achieve germination: common molecular mechanism induced by ethylene and after-ripening in sunflower seeds. *International journal of molecular sciences*, *19*(8), 2464.
- Yacoubi, R., Job, C., Belghazi, M., Chaibi, W., & Job, D. (2011). Toward characterizing seed vigour in alfalfa through proteomic analysis of germination and priming. *Journal of proteome research*, *10*(9), 3891-3903.
- Zeng, D. E., Hou, P., Xiao, F., & Liu, Y. (2015). Overexpression of Arabidopsis XERICO gene confers enhanced drought and salt stress tolerance in rice (Oryza sativa L.). *Journal of plant biochemistry and biotechnology*, *24*(1), 56-64.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., ... & Chanda, S. K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications*, *10*(1), 1-10.

https://doi.org/10.1038/s41467-019-09234-6. Disponible en: http://metascape.org/ última fecha de visita: mayo 2021

ANEXOS

Anexo 1. Protocolo de extracción de RNA a partir de semillas (modificado de Li y Trick, 2005)

Buffer 1 100 mM Tris pH 8.0
150 mM LiCl, 50 mM EDTA
1.5% Dodecil sulfato de sodio (SDS)
1.5% 2-mercapto-etanol

Buffer 2 0.91 M citrato de sodio
2.4 M acetato de sodio
10% Laurylsarcosine
pH 4.0-5.0

1. Macerar la muestra en mortero con nitrógeno líquido y transferir a un tubo de 1.5 ml (libre de RNAsas) para microcentrífuga. La cantidad de polvo no debe rebasar la marca de 250 µl.
2. Agregar 400 µl de buffer 1 y mezclar perfectamente con una punta de micropipeta o por vórtex.
3. Agregar 250 µl de fenol-cloroformo-isoamílico (25:24:1) pH 8.0; mezclar por inversión y centrifugar a 11,700 rpm durante 15 min a 4°C.
4. Transferir con cuidado la fase acuosa a un tubo nuevo de 1.5 ml y agregar 250 µl de buffer 2; mezclar por inversión e incubar 10 min a temperatura ambiente.
5. Agregar 1ml de TRIZOL, mezclar por inversión e incubar a temperatura ambiente durante 5 min. Centrifugar la muestra a 11,700 rpm 15 min a 4°C.
6. Transferir la fase acuosa transparente a un tubo nuevo de 1.5 ml y agregar 1/5 del volumen de cloroformo. Centrifugar la muestra a 11,700 rpm 15 min a 4°C.
7. Pasar el sobrenadante a un tubo nuevo y agregar 1/3 de volumen de isopropanol, mezclar por inversión y agregar 1/10 de volumen (del sobrenadante) de acetato de sodio 3M pH 5.2. Incubar en hielo por 15 min.
8. Centrifugar la muestra a 11,700 rpm por 15 min a 4°C.
9. Eliminar el sobrenadante y lavar el pellet con etanol al 70%.
10. Secar el pellet y re-suspender en 50 µl de H_2O libre de RNAsas.
11. Almacenar el RNA a -70°C hasta su utilización.

Anexo 2. Calidad y nivel de lecturas pareadas de 100pb obtenidas por biblioteca de RNAseq.

Registro de la base SRA	Nombre de la biblioteca	Lecturas totales
SAMN12611291	AG_C1_1_3C_R1	4,498,548
SAMN12611291	AG_C1_1_3C_R2	4,498,548
SAMN12611291	AG_C2_1_1_R1	3,492,306
SAMN12611291	AG_C2_1_1_R2	3,492,306
SAMN12611291	AG_C3_3_1_R1	6,702,499
SAMN12611291	AG_C3_3_1_R2	6,702,499
SAMN12611291	AG_P1_1_2P_R1	5,383,072

Registro de la base SRA	Nombre de la biblioteca	Lecturas totales
SAMN12611292	POOL1_XIMENA_1_1	1,260,657
SAMN12611292	POOL1_XIMENA_1_2	1,260,657
SAMN12611292	POOL1_XIMENA_2_1	1,960,316
SAMN12611292	POOL1_XIMENA_2_2	1,960,316
SAMN12611292	POOL1_XIMENA_3_1	3,233,783
SAMN12611292	POOL1_XIMENA_3_2	3,233,783
SAMN12611292	POOL1_XIMENA_4_1	1,217,511
SAMN12611292	POOL1_XIMENA_4_2	1,217,511
SAMN12611292	POOL1_XIMENA_5_1	3,929,721
SAMN12611292	POOL1_XIMENA_5_2	3,929,721
SAMN12611292	POOL1_XIMENA_6_1	2,555,219
SAMN12611292	POOL1_XIMENA_6_2	2,555,219
SAMN12611292	POOL1_XIMENA_7_1	5,009,356
SAMN12611292	POOL1_XIMENA_7_2	5,009,356
SAMN12611292	POOL1_XIMENA_8_1	6,913,534
SAMN12611292	POOL1_XIMENA_8_2	6,913,534
SAMN12611292	POOL1_XIMENA_9_1	5,453,352
SAMN12611292	POOL1_XIMENA_9_2	5,453,352
SAMN12611292	POOL1_XIMENA_10_1	9,977,047
SAMN12611292	POOL1_XIMENA_10_2	9,977,047
SAMN12611292	POOL1_XIMENA_11_1	7,476,922
SAMN12611292	POOL1_XIMENA_11_2	7,476,922
SAMN12611292	POOL1_XIMENA_12_1	7,734,694
SAMN12611292	POOL1_XIMENA_12_2	7,734,694
SAMN12611292	POOL1_XIMENA_13_1	9,944,690
SAMN12611292	POOL1_XIMENA_13_2	9,944,690
SAMN12611292	POOL1_XIMENA_14_1	15,812,297
SAMN12611292	POOL1_XIMENA_14_2	15,812,297
SAMN12611292	POOL1_XIMENA_15_1	17,246,752
SAMN12611292	POOL1_XIMENA_15_2	17,246,752
SAMN12611292	POOL1_XIMENA_16_1	17,737,645
SAMN12611292	POOL1_XIMENA_16_2	17,737,645
SAMN12611292	POOL1_XIMENA_17_1	15,440,236
SAMN12611292	POOL1_XIMENA_17_2	15,440,236
SAMN12611292	POOL1_XIMENA_18_1	14,043,807
SAMN12611292	POOL1_XIMENA_18_2	14,043,807
SAMN12611292	POOL1_XIMENA_19_1	12,389,341
SAMN12611292	POOL1_XIMENA_19_2	12,389,341
SAMN12611292	POOL1_XIMENA_20_1	12,662,340
SAMN12611292	POOL1_XIMENA_20_2	12,662,340
SAMN12611292	POOL1_XIMENA_21_1	14,137,815
SAMN12611292	POOL1_XIMENA_21_2	14,137,815
SAMN12611292	POOL1_XIMENA_22_1	14,810,131
SAMN12611292	POOL1_XIMENA_22_2	14,810,131
SAMN12611292	POOL1_XIMENA_23_1	15,718,464
SAMN12611292	POOL1_XIMENA_23_2	15,718,464
SAMN12611292	POOL1_XIMENA_24_1	12,980,228
SAMN12611292	POOL1_XIMENA_24_2	12,980,228
SAMN12611292	POOL1_XIMENA_25_1	16,216,395
SAMN12611292	POOL1_XIMENA_25_2	16,216,395
Registro de la base SRA	Nombre de la biblioteca	Lecturas totales
SAMN12611292	POOL1_XIMENA_26_1	14,659,470
SAMN12611292	POOL1_XIMENA_26_2	14,659,470
SAMN12611292	POOL1_XIMENA_27_1	16,359,898
SAMN12611292	POOL1_XIMENA_27_2	16,359,898
SAMN12611292	POOL1_XIMENA_28_1	10,535,078
SAMN12611292	POOL1_XIMENA_28_2	10,535,078
SAMN12611292	POOL1_XIMENA_29_1	14,513,548
SAMN12611292	POOL1_XIMENA_29_2	14,513,548
SAMN12611292	POOL1_XIMENA_30_1	14,433,618
SAMN12611292	POOL1_XIMENA_30_2	14,433,618
SAMN12611292	POOL1_XIMENA_31_1	13,533,127
SAMN12611292	POOL1_XIMENA_31_2	13,533,127
SAMN12611292	POOL1_XIMENA_32_1	17,127,932
SAMN12611292	POOL1_XIMENA_32_2	17,127,932
SAMN12611292	POOL1_XIMENA_33_1	18,691,122
SAMN12611292	POOL1_XIMENA_33_2	18,691,122

SAMN12611292	POOL1_XIMENA_34_1	13,620,740
SAMN12611292	POOL1_XIMENA_34_2	13,620,740
SAMN12611292	POOL1_XIMENA_35_1	11,710,751
SAMN12611292	POOL1_XIMENA_35_2	11,710,751
SAMN12611292	POOL1_XIMENA_36_1	14,992,386
SAMN12611292	POOL1_XIMENA_36_2	14,992,386
SAMN12611292	POOL1_XIMENA_37_1	12,099,889
SAMN12611292	POOL1_XIMENA_37_2	12,099,889
SAMN12611292	POOL1_XIMENA_38_1	13,226,172
SAMN12611292	POOL1_XIMENA_38_2	13,226,172
SAMN12611292	POOL1_XIMENA_39_1	15,587,051
SAMN12611292	POOL1_XIMENA_39_2	15,587,051
SAMN12611292	POOL1_XIMENA_40_1	13,220,814
SAMN12611292	POOL1_XIMENA_40_2	13,220,814
SAMN12611292	POOL1_XIMENA_41_1	11,593,188
SAMN12611292	POOL1_XIMENA_41_2	11,593,188
SAMN12611292	POOL1_XIMENA_42_1	15,748,435
SAMN12611292	POOL1_XIMENA_42_2	15,748,435
SAMN12611292	POOL1_XIMENA_43_1	10,171,731
SAMN12611292	POOL1_XIMENA_43_2	10,171,731
SAMN12611292	POOL1_XIMENA_44_1	12,599,728
SAMN12611292	POOL1_XIMENA_44_2	12,599,728
SAMN12611292	POOL1_XIMENA_45_1	9,201,476
SAMN12611292	POOL1_XIMENA_45_2	9,201,476
SAMN12611292	POOL1_XIMENA_46_1	9,940,334
SAMN12611292	POOL1_XIMENA_46_2	9,940,334
SAMN12611292	POOL1_XIMENA_47_1	9,851,750
SAMN12611292	POOL1_XIMENA_47_2	9,851,750
SAMN12611292	POOL1_XIMENA_48_1	12,703,784
SAMN12611292	POOL1_XIMENA_48_2	12,703,784
SAMN12611292	POOL1_XIMENA_1_1	1,260,657
SAMN12611292	POOL1_XIMENA_1_2	1,260,657
SAMN12611292	POOL1_XIMENA_2_1	1,960,316
SAMN12611292	POOL1_XIMENA_2_2	1,960,316
SAMN12611292	POOL1_XIMENA_3_1	3,233,783
SAMN12611292	POOL1_XIMENA_3_2	3,233,783
SAMN12611292	POOL1_XIMENA_4_1	1,217,511
SAMN12611292	POOL1_XIMENA_4_2	1,217,511
SAMN12611292	POOL1_XIMENA_5_1	3,929,721
SAMN12611292	POOL1_XIMENA_5_2	3,929,721
SAMN12611292	POOL1_XIMENA_6_1	2,555,219
SAMN12611292	POOL1_XIMENA_6_2	2,555,219
SAMN12611292	POOL1_XIMENA_7_1	5,009,356
SAMN12611292	POOL1_XIMENA_7_2	5,009,356

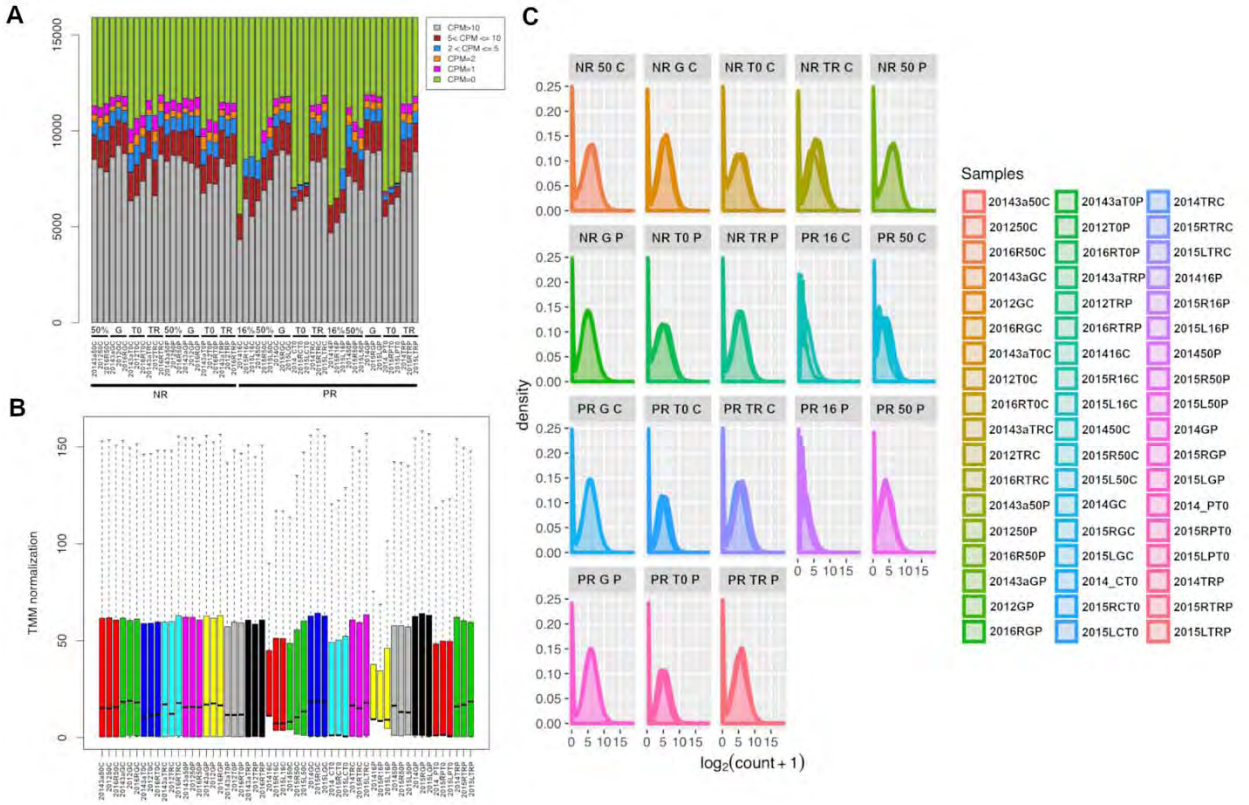
Anexo 3. Estadísticas de calidad del ensamble de cada biblioteca secuenciada.

Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library1	Total transcripts:	39806	library2	Total transcripts:	42511	library3	Total transcripts:	44044
	Percent GC	42.64		Percent GC	42.61		Percent GC	42.56
	Contig N50	1442		Contig N50	1444		Contig N50	1448
2014_CT0	Median contig length	876	2015-1_CT0	Median contig length	875	2015-2_CT0	Median contig length	883
	Average contig	1104.88		Average contig	1105.2		Average contig	1111.64
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library4	Total transcripts:	38326	library5	Total transcripts:	40623	library6	Total transcripts:	44355
	Percent GC	42.68		Percent GC	42.64		Percent GC	42.58
	Contig N50	1447		Contig N50	1439		Contig N50	1449
2014_PT0	Median contig length	881	2015-1_PT0	Median contig length	874	2015-2_PT0	Median contig length	881
	Average contig	1107.95		Average contig	1102.63		Average contig	1110.31
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library1	Total transcripts:	70164	library2	Total transcripts:	132430	library3	Total transcripts:	136296
	Percent GC	41.64		Percent GC	41.04		Percent GC	41.38
AG-1416C	Contig N50	1301	AG-15R16C	Contig N50	1245	AG-15L16C	Contig N50	1240
	Median contig	673		Median contig	663		Median contig	657

	length			length			length	
	Average contig	938.41		Average contig	912.81		Average contig	906.32
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library4	Total transcripts:	75625	library5	Total transcripts:	82829	library6	Total transcripts:	115498
	Percent GC	41.81		Percent GC	41.45		Percent GC	41.23
	Contig N50	1274		Contig N50	1319		Contig N50	1271
AG-1416P	Median contig length	661	AG-15R16P	Median contig length	684	AG-15L16P	Median contig length	673
	Average contig	919.93		Average contig	950.83		Average contig	927.49
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library7	Total transcripts:	124216	library8	Total transcripts:	175348	library9	Total transcripts:	212851
	Percent GC	41.36		Percent GC	40.81		Percent GC	40.94
	Contig N50	1264		Contig N50	1201		Contig N50	1146
AG-1450C	Median contig length	667	AG-15R50C	Median contig length	644	AG-15L50C	Median contig length	613
	Average contig	918.98		Average contig	884.56		Average contig	845.91
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library10	Total transcripts:	212119	library11	Total transcripts:	228706	library12	Total transcripts:	172823
	Percent GC	40.65		Percent GC	40.43		Percent GC	40.87
	Contig N50	1149		Contig N50	1124		Contig N50	1217
AG-1450P	Median contig length	614	AG-15R50P	Median contig length	601	AG-15L50P	Median contig length	655
	Average contig	847.59		Average contig	831.11		Average contig	894.62
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library13	Total transcripts:	254728	library14	Total transcripts:	288247	library15	Total transcripts:	294893
	Percent GC	40.35		Percent GC	40.2		Percent GC	40.29
	Contig N50	1097		Contig N50	1040		Contig N50	1035
AG-14TRC	Median contig length	591	AG-15RTRC	Median contig length	560	AG-15LTRC	Median contig length	559
	Average contig	814.73		Average contig	778.69		Average contig	776.39
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library16	Total transcripts:	275324	library17	Total transcripts:	273774	library18	Total transcripts:	304050
	Percent GC	40.43		Percent GC	40.46		Percent GC	40.21
	Contig N50	1066		Contig N50	1068		Contig N50	1021
AG-14TRP	Median contig length	573	AG-15RTRP	Median contig length	574	AG-15LTRP	Median contig length	550
	Average contig	793.69		Average contig	794.82		Average contig	765.94
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library19	Total transcripts:	284446	library20	Total transcripts:	286688	library21	Total transcripts:	294581
	Percent GC	40.34		Percent GC	40.35		Percent GC	40.32
	Contig N50	1050		Contig N50	1047		Contig N50	1036
AG-14GC	Median contig length	567	AG-15RGC	Median contig length	569	AG-15LGC	Median contig length	560
	Average contig	785.68		Average contig	785.95		Average contig	777.28
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library22	Total transcripts:	298547	library23	Total transcripts:	268600	library24	Total transcripts:	289530
	Percent GC	40.27		Percent GC	40.31		Percent GC	40.32
	Contig N50	1030		Contig N50	1031		Contig N50	1043
AG-14GP	Median contig length	555	AG-15RGP	Median contig length	557	AG-15LGP	Median contig length	566
	Average contig	772.24		Average contig	773.35		Average contig	782.54
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library25	Total transcripts:	241656	library26	Total transcripts:	249854	library27	Total transcripts:	293462
	Percent GC	40.47		Percent GC	40.57		Percent GC	40.18
	Contig N50	1099		Contig N50	1093		Contig N50	1026
AG-143aT0C	Median contig length	584	AG-12T0C	Median contig length	575	AG-16RT0C	Median contig length	546
	Average contig	812.41		Average contig	803.57		Average contig	766.2
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor

library28	Total transcripts:	219710	library29	Total transcripts:	268053	library30	Total transcripts:	273721
	Percent GC	40.5		Percent GC	40.18		Percent GC	40.25
	Contig N50	1130		Contig N50	1058		Contig N50	1051
AG-143aT0P	Median contig length	603	AG-12T0P	Median contig length	567	AG-16RT0P	Median contig length	559
	Average contig	834.2		Average contig	789.31		Average contig	782.5
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library31	Total transcripts:	293426	library32	Total transcripts:	279282	library33	Total transcripts:	275764
	Percent GC	40.2		Percent GC	40.34		Percent GC	40.49
	Contig N50	1031		Contig N50	1051		Contig N50	1061
AG-143a50C	Median contig length	555	AG-1250C	Median contig length	564	AG-16R50C	Median contig length	573
	Average contig	773.42		Average contig	785.11		Average contig	793.24
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library34	Total transcripts:	281279	library35	Total transcripts:	296322	library36	Total transcripts:	305560
	Percent GC	40.2		Percent GC	40.15		Percent GC	40.38
	Contig N50	1046		Contig N50	1028		Contig N50	1016
AG-143a50P	Median contig length	563	AG-1250P	Median contig length	551	AG-16R50P	Median contig length	542
	Average contig	783.41		Average contig	769.5		Average contig	759.97
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library37	Total transcripts:	254317	library38	Total transcripts:	200627	library39	Total transcripts:	288364
	Percent GC	40.42		Percent GC	40.97		Percent GC	40.4
	Contig N50	1099		Contig N50	1170		Contig N50	1044
AG-143aTRC	Median contig length	597	AG-12TRC	Median contig length	635	AG-16RTRC	Median contig length	566
	Average contig	819.61		Average contig	867.2		Average contig	783.72
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library40	Total transcripts:	258802	library41	Total transcripts:	269786	library42	Total transcripts:	251854
	Percent GC	40.38		Percent GC	40.42		Percent GC	40.45
	Contig N50	1089		Contig N50	1078		Contig N50	1100
AG-143aTRP	Median contig length	594	AG-12TRP	Median contig length	575	AG-16RTRP	Median contig length	600
	Average contig	815.03		Average contig	798.01		Average contig	822.32
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library43	Total transcripts:	250499	library44	Total transcripts:	285116	library45	Total transcripts:	277377
	Percent GC	40.54		Percent GC	40.22		Percent GC	40.48
	Contig N50	1105		Contig N50	1048		Contig N50	1063
AG-143aGC	Median contig length	605	AG-12GC	Median contig length	571	AG-16RGC	Median contig length	575
	Average contig	826.41		Average contig	788.11		Average contig	794.18
Nombre	Atributo	Valor	Nombre	Atributo	Valor	Nombre	Atributo	Valor
library46	Total transcripts:	250172	library47	Total transcripts:	245977	library48	Total transcripts:	266491
	Percent GC	40.61		Percent GC	40.61		Percent GC	40.97
	Contig N50	1104		Contig N50	1112		Contig N50	1078
AG-143aGP	Median contig length	601	AG-12GP	Median contig length	604	AG-16RGP	Median contig length	579
	Average contig	824.2		Average contig	828.49		Average contig	801.04

Anexo 4. **Gráficos del análisis de calidad de las 48 bibliotecas secuenciadas**. A) cuentas por millón de lecturas detectadas por gen, por biblioteca. B) cuentas por millón normalizadas por el método de medias recortadas. C) Densidad de cuentas por millón en escala logarítmica (base 2) por biblioteca, agrupadas por estadio fisiológico y tratamiento.



Anexo 5. Método de Bligh y Dyer (1959) para la cuantificación de lípidos totales.

- Para 100mg de muestra:

1) Agregar 1 ml de $CHCl_3$/MeOH (1:2) a la muestra de tejido y 0.27 ml de H_2O, agitar vigorosamente durante 5min.
2) Agregar 0.0.33 ml de $CHCl_3$ y agitar por un minuto.
3) Agregar 0.3 ml de H_2O y agitar por un minuto.
4) Centrifugar 1 min a 2000 rpm para separar las fases.
5) Los lípidos se encuentra en la fase orgánica (inferior). Transferir la fase orgánica a un tubo nuevo y dejar que se evapore cualquier rastro del solvente.

Anexo 6. Figuras suplementarias que acompañan los resultados presentados en el Capítulo II: Regulación transcriptómica del proceso de germinación, el contenido relativo de agua (CRA) como una aproximación a fases del desarrollo discretas durante la germinación.

Fig. S1. Climograph of the "Trapiche del Rosario y Chicuasén" locality within Actopan, Veracruz (Mexico). Blue bars depict the mean precipitation (mm) per month. The red, orange and yellow lines depict the maximum, mean and minimum average temperatures per month (ºC), respectively. The arrow above February indicates the moment at which *C. aesculifolia* seeds as shed on a regular season.

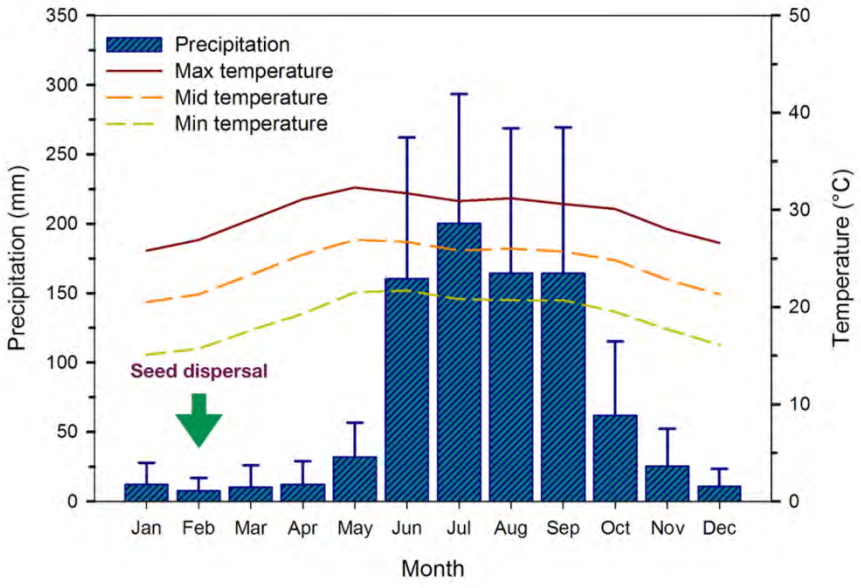


Fig. S2. Seed germination performance of seed lots with positive (2014, 2015-1) and negative (2012-5y, 2016) response to priming. In all graphs, letters denote statistical differences in the start of germination(near the X axis), time to 50% (middle), and final germination (top right corner). The asterisk (*) in the 2016 graph indicates a statistically significant decrease in the maximum germination rate in primed seeds with respect to control seeds (see table 1).

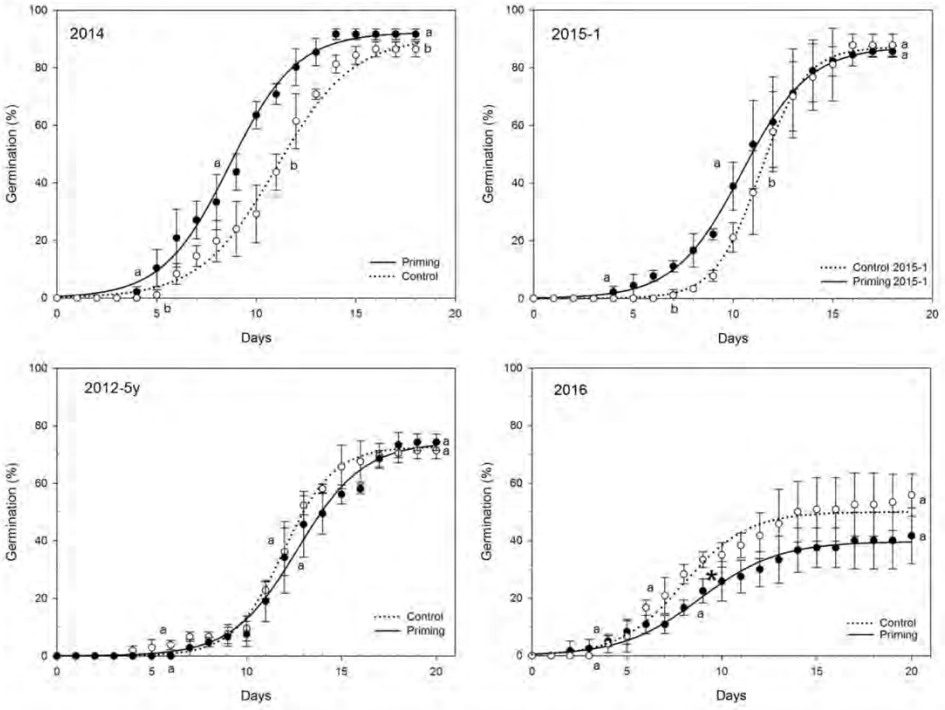


Fig. S3. Seed imbibition variability per hour in seed batches with positive response to priming. The imbibition axis is shown as a category rather than a continuous scale to facilitate interpretation.

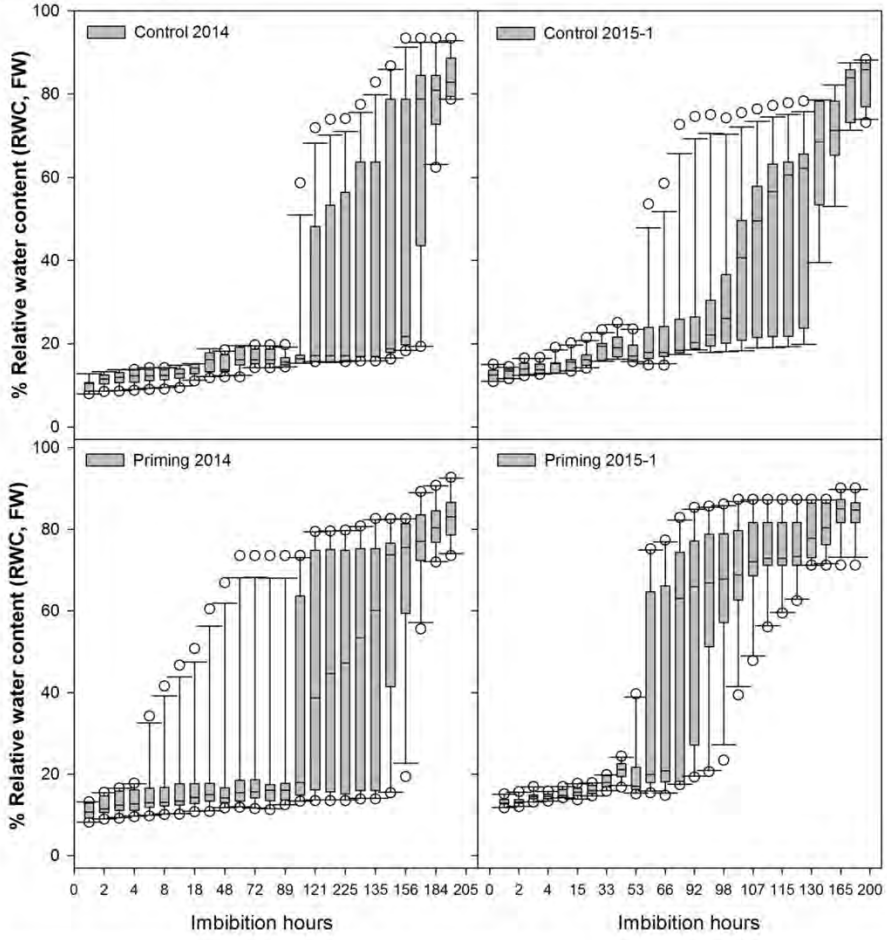


Fig. S4. Seed imbibition variability per hour in seed batches with negative response to priming. The imbibition axis is shown as a category rather than a continuous scale to facilitate interpretation.

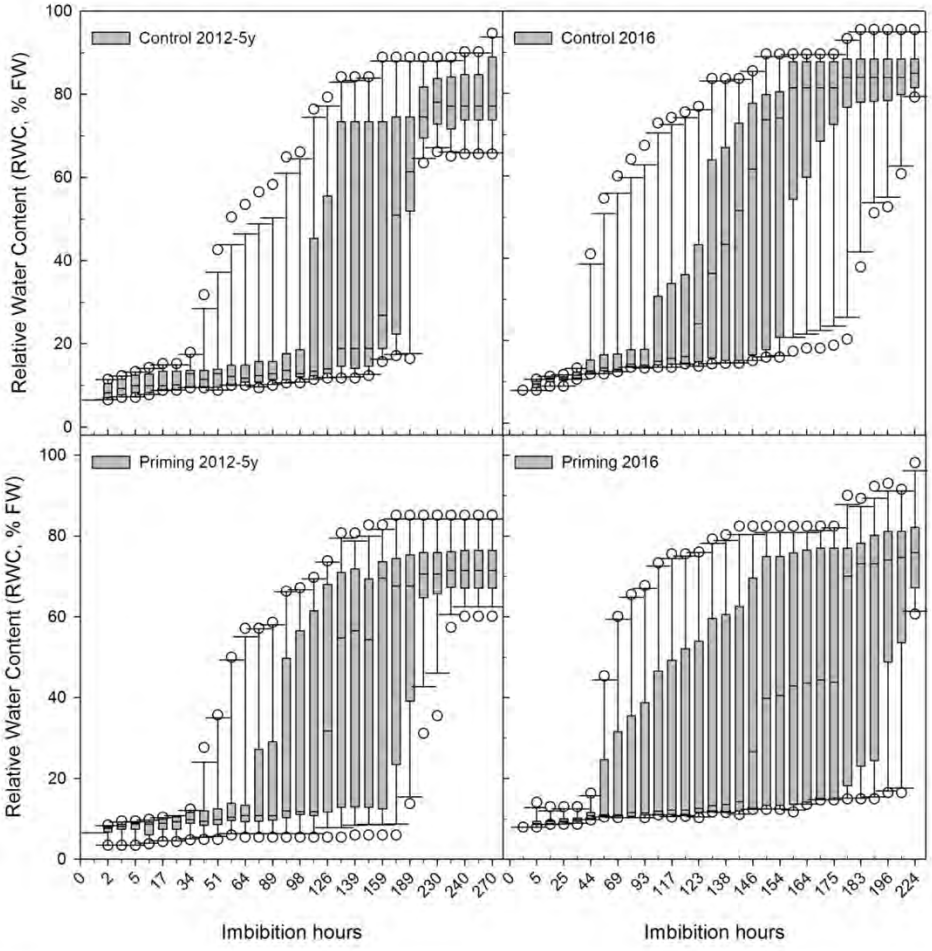


Fig. S5. Venn diagrams of the total number of genes differentially expressed in the two phenotypes. A) Venn diagram of the up- or down-regulated genes in at least one physiological stage during germination in each phenotype. Ellipses from left to right: up-regulated genes in the PR-batches, down-regulated genes in PR-batches, up-regulated genes in the NR-batches, and down-regulated genes in NR-batches. B) Comparison between the regulated genes (up- and down-regulated) in each phenotype and the overall differential expression pattern during of the NR-phenotype during germination in relation to the observed expression pattern in the PR-phenotype (*i.e.* the genes that on average had higher or lower read counts in NR in comparison to PR over the whole duration of the germination process).

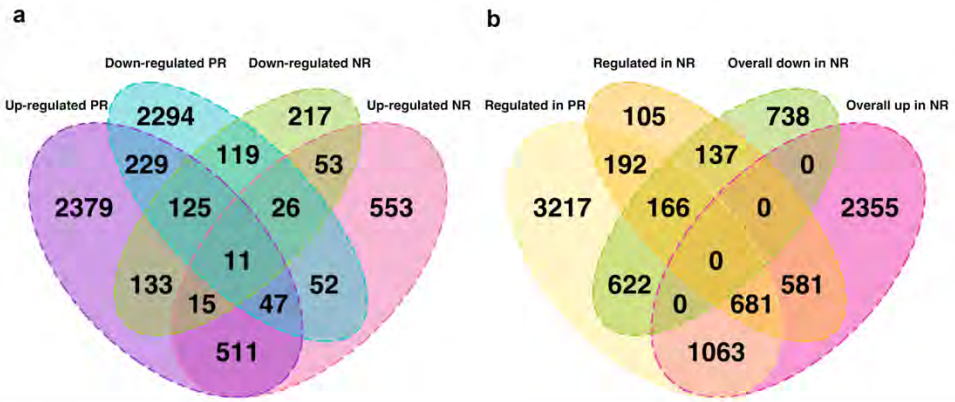


Fig. S6. Time-series clusters of the main gene expression patterns detected in the differentially expressed genes above a threshold of 0.35 of overall difference during germination between PR- and NR-batches. The total number of genes in each cluster are shown. From a to d, gene clusters that had an overall up-regulation in PR batches (purple clusters), and in NR-batches (turquoise batches). In a and d is notoriously different the average read counts at T0 in NR-batches. In e and f are the clusters with a down-regulation trend during germination, while in g and h the main pattern of expression changes by the 50% stage in PR- batches, but in NR-batches that trend is not notorious due to a total read count in T0 distinctively above or below the observed read counts in PR-batches.

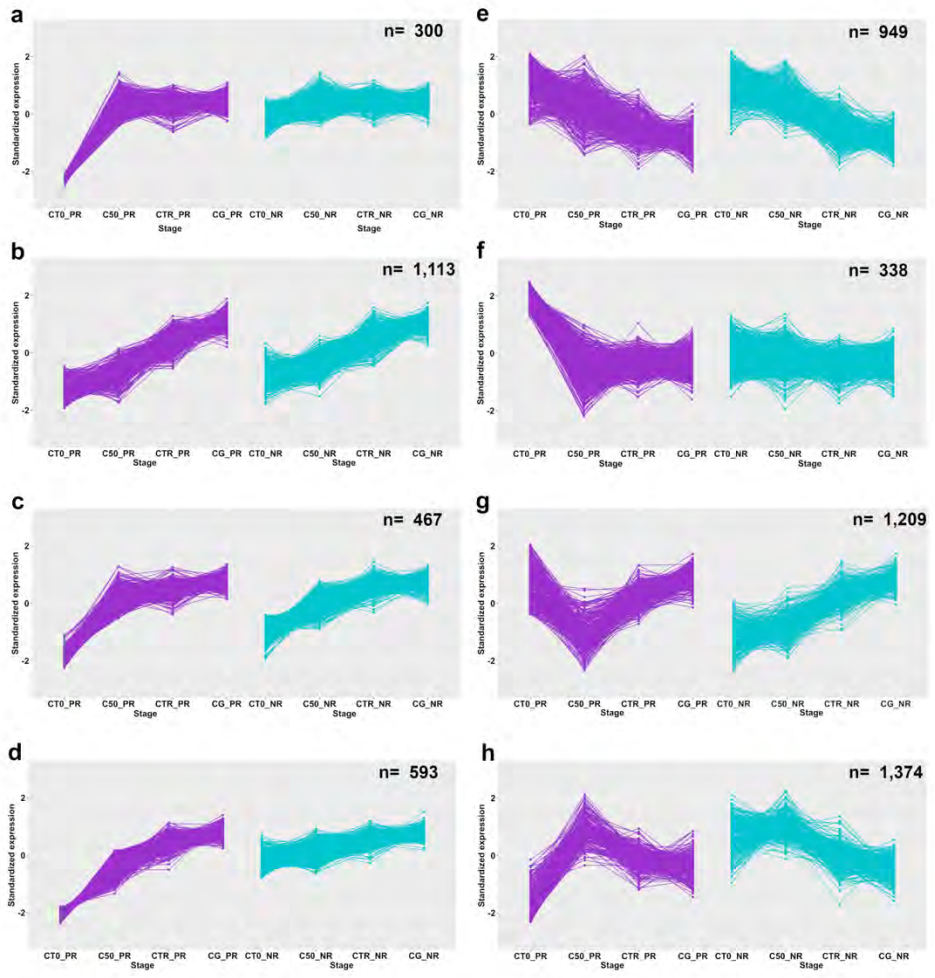


Table S1. Summarized germination performance of the six seed batches analysed, based on four germination features: germination initiation, maximum germination rate, day to attain 50% germination, and final germination. The numbers in bold denote significant differences for a two-tailed *t* test with 8 degrees of freedom and α=0.05 (*P*-values in parenthesis) of primed seeds in comparison to its respective control, for each seed batch. Results are presented as the average of 5 independent replicates and its respective standard deviation. The last column denotes the decision to consider a particular seed batch as having a positive response to priming (PR), or non-responsive to priming (NR), based on the germination test. The numbers with an asterisk (*) in each priming row denote those comparisons for which a Mann-Whitney test was performed due to a failed normality test at an α=0.05 threshold.

Seed Batch	Treatment	Germination initiation (day)	Maximum rate (germination %·d^{-1})	Day to 50% germination	Final germination %	Priming phenotype
2014	Control	5 ± 0.7	8.95 ± 0.68	11.6 ± 0.37	86.45 ± 2.6	PR
2014	Priming	4 ±0.61 (0.08)	**10.65 ± 1.1** (0.02)	**9.3 ± 0.89** (0.001)	**92.33 ± 3.34** (0.03)	PR
2015-1	Control	7 ±1.17	12.4 ± 3.3	12.24 ± 1	87.78 ± 3.85	PR
2015-1	Priming	**4 ± 1.05** (0.01)	**7.96 ± 2.14** (0.03)	11.7 ± 1.24 (0.6)	85.56 ± 1.92 (0.34)	PR
2015-2	Control	6 ± 0.48	13.08 ± 0.24	10.42 ± 0.46	87.78 ± 1.92	PR
2015-2	Priming	**4 ±1** (0.01)	14.69 ± 3.06* (0.31)	**8.55 ± 0.48** (0.01)	90.0 ± 3.33 (0.37)	PR
2014-3y	Control	5 ± 0.6	6.99 ± 1.62	12.93 ± 1.17	94.0 ± 2.79	NR
2014-3y	Priming	6 ± 0.81* (0.91)	7.3 ± 1.34 (0.75)	14.09 ± 2.19 (0.33)	85.33 ± 8.69 (0.07)	NR
2012-5y	Control	5 ± 1.46	6.48 ± 1.29	10.94 ± 0.4	71.43 ± 2.86	NR
2012-5y	Priming	6 ± 0.84 (0.12)	7.49 ± 0.49 (0.27)	11.52 ± 1.08 (0.43)	74.29 ± 2.86 (0.14)	NR
2016	Control	4 ± 2.79	8.05 ± 0.984	6.26 ± 1.22	53.33 ± 8.5	NR
2016	Priming	2 ± 3.65* (0.07)	**3.94 ± 0.63** (0.001)	6.77 ± 0.69 (0.5)	46.0 ± 12.78* (0.11)	NR

Table S4. Primers of the selected genes used for the semi-quantitative RT-PCR validation of RNAseq expression data.

Gene name	Locus tag	Primer sequence (5´- 3')	PCR product size (pb)	Tm (°C)	%GC
Actin 7	*At5g09810*	F- CTGCCATGTATGTTGCCATC R- ACGGAATCTCTCAGCTCCAA	380	59.96 59.95	50.00 50.00
PME3	*At3g14310*	F -ATCAACGGTTCCTGCTACGA R - CAGTTGCTGTGGTTGGTGAA	209	59.11 58.90	50.00 50.00
PMEI-like	*At5g20740*	F -GTCAAAGTCAGCCTGTCACG R -CCCCACTCACCCATGTCTG	315	59.14 59.70	55.00 63.16
MCD	*At4g04320*	F - GGAGTTGGTCGTCGTTGTTT R - CCAGTAGCAAGATCCCCAGA	464	60.01 60.21	50.00 55.00
GDPD1	*At3g02040*	F - CAGGGCAGCATCGTTTAAGT R - TGGATGAGATGGTCTTGCTG	492	60.27 59.79	50.00 50.00
Xerico	*At2g04240*	F - CGTCGTTATCACCATCCCCT R -ACCACTTTTCGAGGCACACC	249	59.25 60.82	55.00 55.00

Anexo 7. Figuras suplementarias que acompañan los resultados presentados en el capítulo IV: El transcriptoma de la semilla seca y su relevancia en la capacidad de responder a estímulos ambientales, el caso de la respuesta al acondicionamiento mátrico.

Table S1. Cold and Warm episodes by season, associated to the El Niño-Southern Oscillation (ONI). The estimates were obtained from the NOAA Center for Weather and Climate Prediction (https://origin.cpc.ncep.noaa.gov). The trimesters begin on December (D) of the previous year and end on January (J) of the following year. We present the years from 2011 (maturation season for the 2012 seed batch), until 2018 (last collection year presented in this study). The color scale indicates the severity of the temperature anomaly registered for that trimester (blue: colder than average; red: warmer than average). The numbers marked by a green box indicate the normal season in which seed batches are usually collected. The numbers in a yellow box indicate the collection of the 2016 and 2017 seed batches, which occurred out of the usual seed dispersal season. It is noticeable that the 2016 and 2017 seed batches matured in the two years with the most severe weather anomalies detected in the nine years presented.

Year	Average effect by trimester of the year											
	DJF	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ
2011	-1.4	-1.1	-0.8	-0.6	-0.5	-0.4	-0.5	-0.7	-0.9	-1.1	-1.1	-1
2012	-0.8	-0.6	-0.5	-0.4	-0.2	0.1	0.3	0.3	0.3	0.2	0	-0.2
2013	-0.4	-0.3	-0.2	-0.2	-0.3	-0.3	-0.4	-0.4	-0.3	-0.2	-0.2	-0.3
2014	-0.4	-0.4	-0.2	0.1	0.3	0.2	0.1	0	0.2	0.4	0.6	0.7
2015	0.6	0.6	0.6	0.8	1	1.2	1.5	1.8	2.1	2.4	2.5	2.6
2016	2.5	2.2	1.7	1	0.5	0	-0.3	-0.6	-0.7	-0.7	-0.7	-0.6
2017	-0.3	-0.1	0.1	0.3	0.4	0.4	0.2	-0.1	-0.4	-0.7	-0.9	-1
2018	-0.9	-0.9	-0.7	-0.5	-0.2	0	0.1	0.2	0.5	0.8	0.9	0.8

Scale	El Niño effect	Scale	La Niña effect
0.5 to 1	Weak	-0.5 to -1	Weak
1 to 1.5	Moderate	-1 to -1.5	Moderate
1.5 to 2	Strong	-1.5 to -2	Strong
above 2	Very Strong	below -2	Very strong

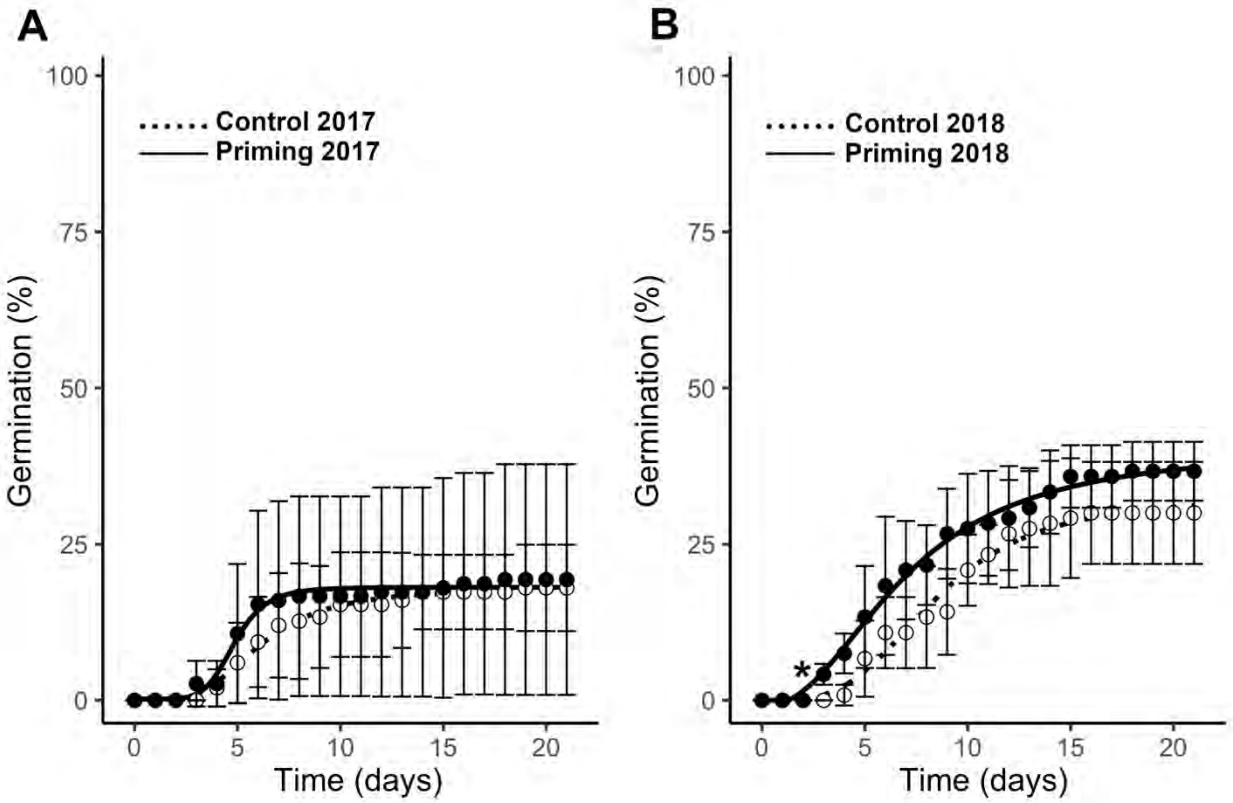


Figure S1. Germination curves of the seed batches collected in 2017 and 2018, which were tested after a 30-day period of after-ripening, and had no positive response to the priming treatment. A) seed batch collected in 2017; B) seed batch collected in 2018. The asterisk in B depicts a statistically significant difference in the timing of the onset of germination between primed seeds and its respective control.

Table S2. Alphanumerical codes used to identify profiles in control seed batches and its equivalency in the profiles detected in primed seeds for PR- and NR- batches, as well as the contrast between priming phenotypes. Numbers were assigned within each contrast based on profile size, from largest to smallest, and ordering profiles present in both treatments before treatment-specific profiles. Letters were assigned in order to distinguish controls (a and c) from primed seeds (b and d).

Profile trend in Figure 2A PR phenotype	Profile pairs in Figure 2A		Profile trend in Figure 2B NR phenotype	Profile pairs in Figure 2B	
	control	priming		control	priming
Up	**1-a**	**1-b**	Up	**1-c**	**1-d**
Up	**2-a**	**2-b**	Up	**2-c**	**2-d**
Up	**3-a**	**3-b**	Up	**3-c**	**3-d**
Up	**4-a**	**4-b**	Up-then-down	**6-c**	**6-d**
Up-then-down	**5-a**	**5-b**	Up-then-down	**7-c**	**7-d**
Up-then-down	**6-a**	**6-b**	Up-then-down	**8-c**	**8-d**
Down-then-up	**8-a**	**8-b**	Up-then-down	**9-c**	**9-d**
Down-then-up	**9-a**	**9-b**	Up-then-down	**10-c**	**10-d**
Down	**11-a**	**11-b**	Down-then-up	**16-c**	**16-d**
Down	**11-a**	**11-b**	Down	**17-c**	**17-d**
Unpaired profiles					
Up-then-down	**7-a**	-	Up	**4-c**	-
Down-then-up	-	**10-b**	Up	-	**5-d**
-	-	-	Up-then-down	**11-c**	-
-	-	-	Up-then-down	-	**12-d**
-	-	-	Up-then-down	-	**13-d**
-	-	-	Down-then-up	**14-c**	-
-	-	-	Down-then-up	-	**15-d**

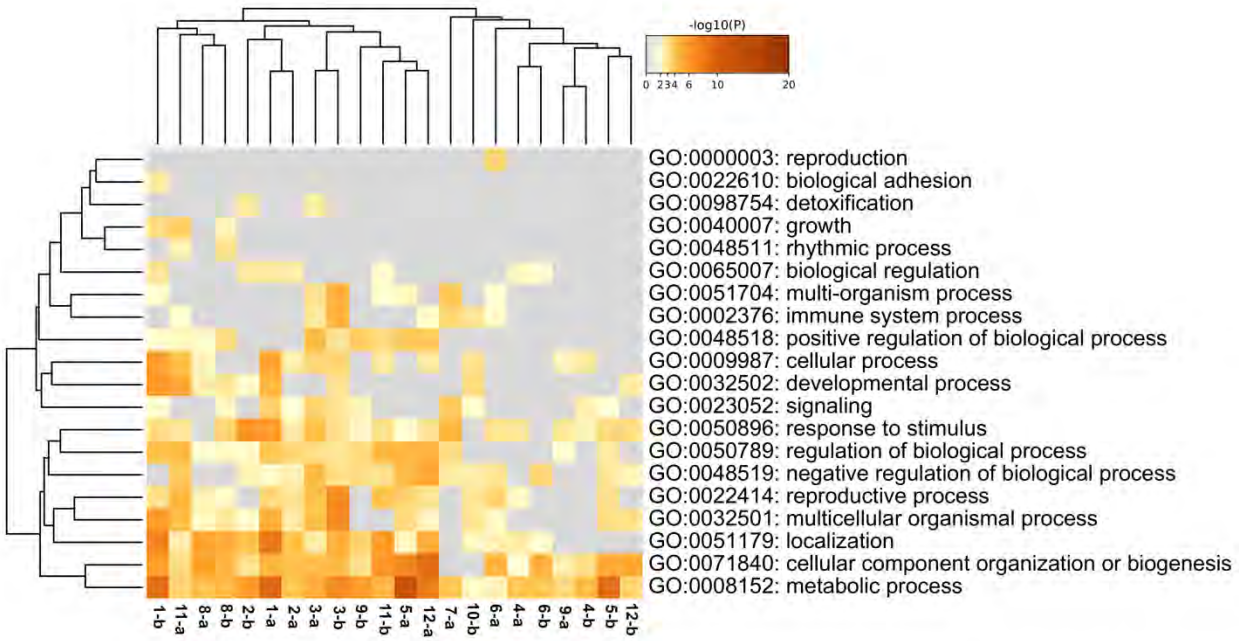


Figure S2. A summary of broad GO-term classes detected in all profiles of control ("a"), and primed ("b") batches with a positive response to priming (PR-phenotype). Filtering of the enriched categories was performed using an accumulative hypergeometric distribution of p-values. The remaining significant terms were then hierarchically clustered into a tree based on Kappa-statistical similarities among their gene memberships, and a kappa score of 0.3 was applied as the threshold to cast the tree into term clusters (Zhou *et al.*, 2019).

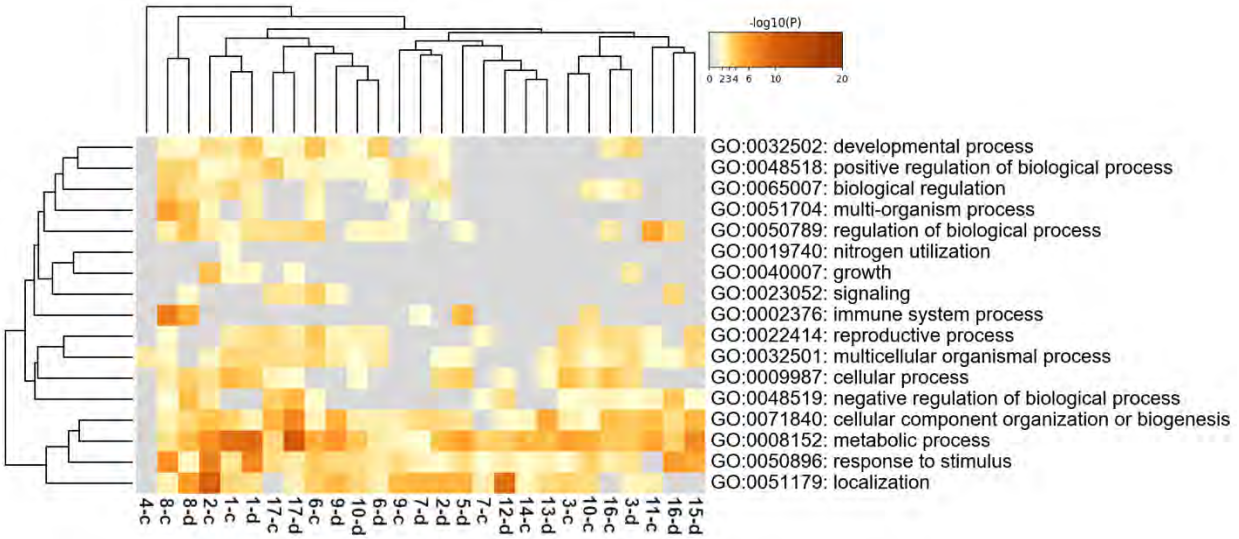


Figure S3. A summary of broad GO-term classes detected in all profiles of control ("c"), and primed ("d") batches with a negative response to priming (NR-phenotype). Filtering of the enriched categories was performed using an accumulative hypergeometric distribution of p-values. The remaining significant terms were then hierarchically clustered into a tree based on Kappa-statistical similarities among their gene memberships, and a kappa score of 0.3 was applied as the threshold to cast the tree into term clusters (Zhou *et al.*, 2019).

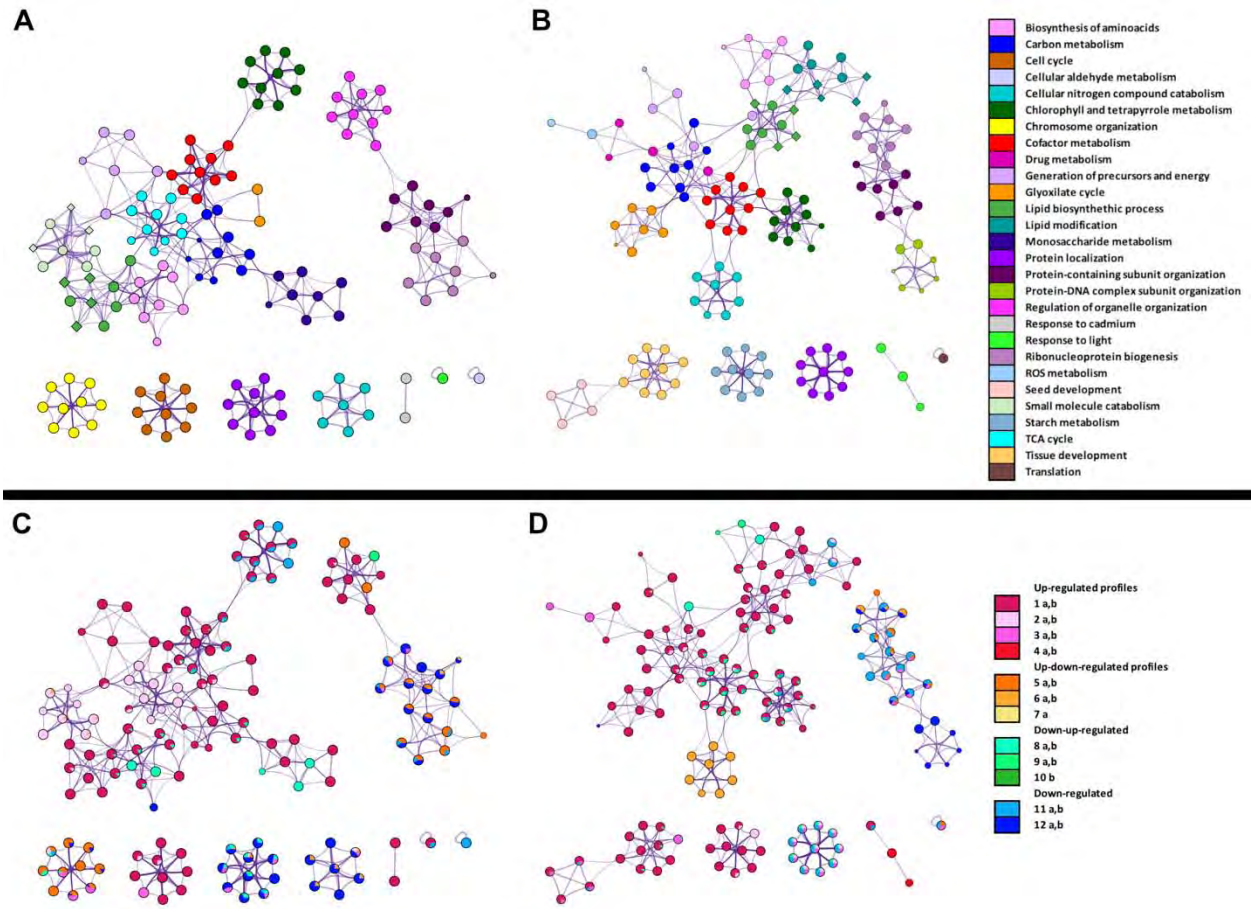


Figure S4. GO-term network of the top 20 clusters detected in control (A and C) and primed seeds (B and D) from the PR phenotype. Panels A and B depict each cluster with a different colour; the terms within a cluster are represented by nodes and the edges represent a similarity score > 0.3; node size represents the number of genes included within each term. Panels C and D are coloured by profile-trend groups. Red tones: Up-regulation trend; orange tones: up-down-regulation trend; turquoise tones: down-up-regulation trend; and blue tones: down-regulation trend.

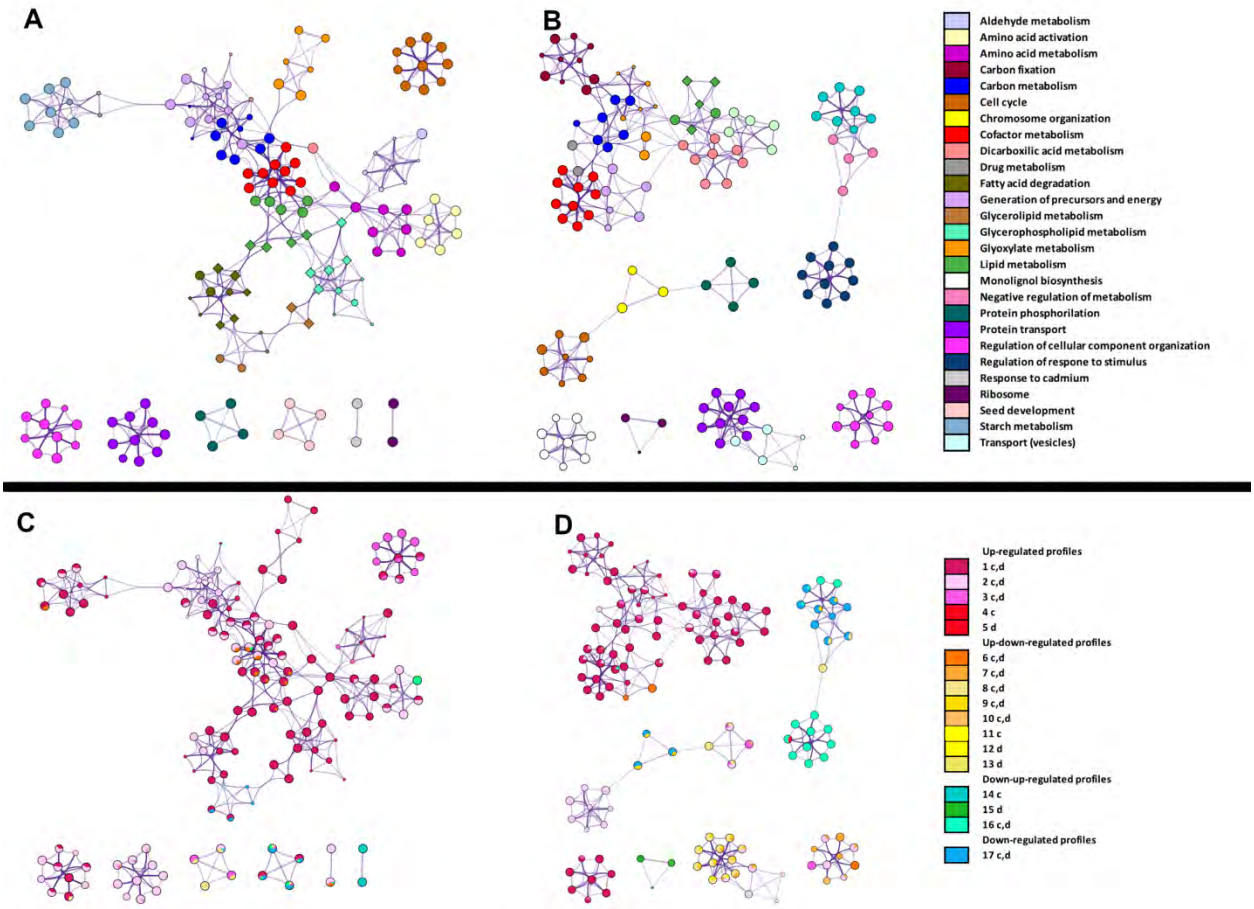


Figure S5. GO-term network of the top 20 clusters detected in control (A and C) and primed seeds (B and D) from the NR phenotype. Panels A and B depicts each functional cluster with a different colour; the terms within a cluster are represented by nodes and the edges represent a similarity score > 0.3; node size represents the number of genes included within each term. Panels C and D are coloured by profile-trend groups. Red tones: Up-regulation trend; yellow tones: up-down-regulation trend; turquoise tones: down-up-regulation trend; and blue tones: down-regulation trend.

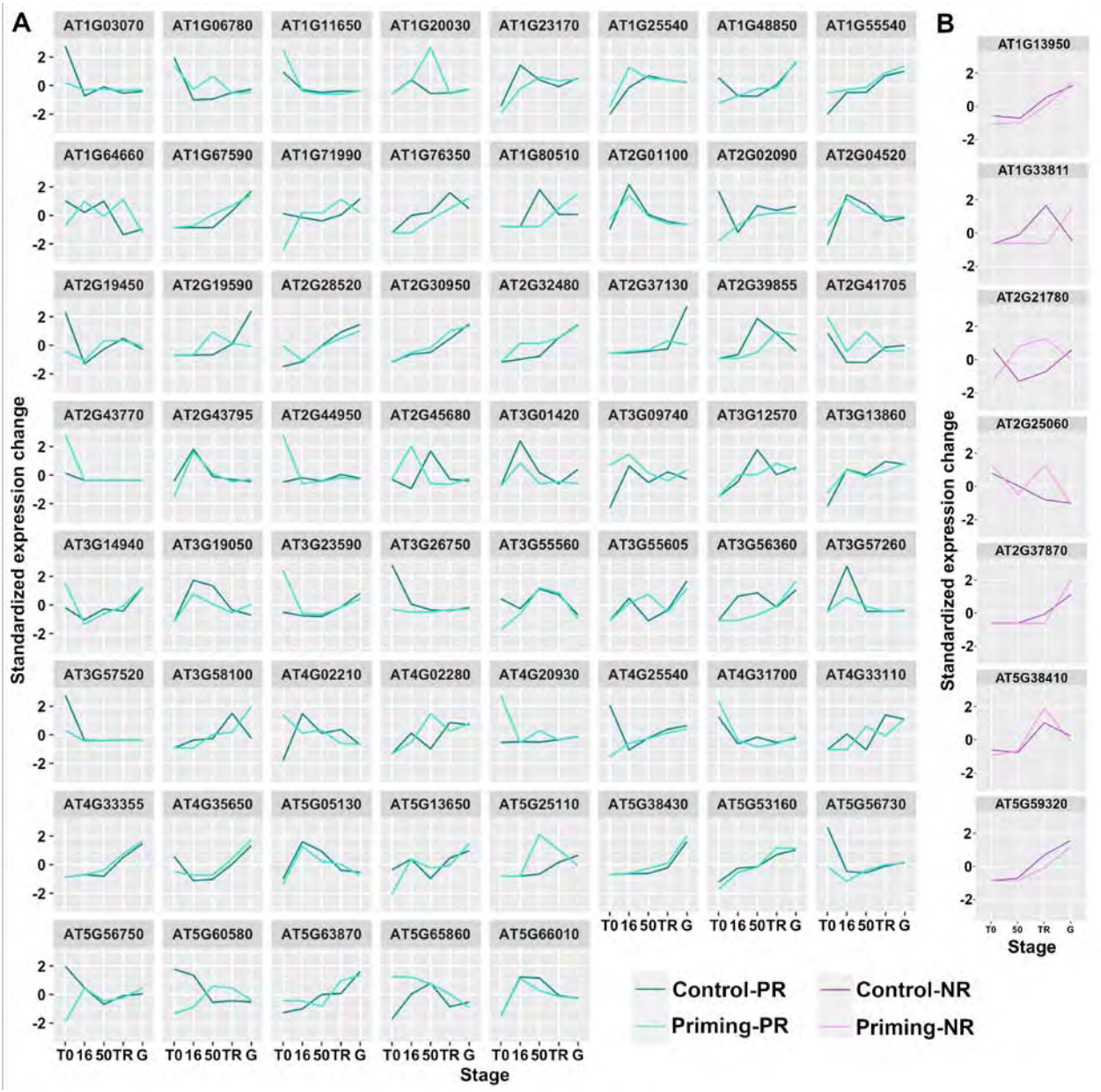


Figure S6. Expression trend during germination of the differentially expressed genes between control and primed seeds in the PR (A) and NR (B) phenotypes.

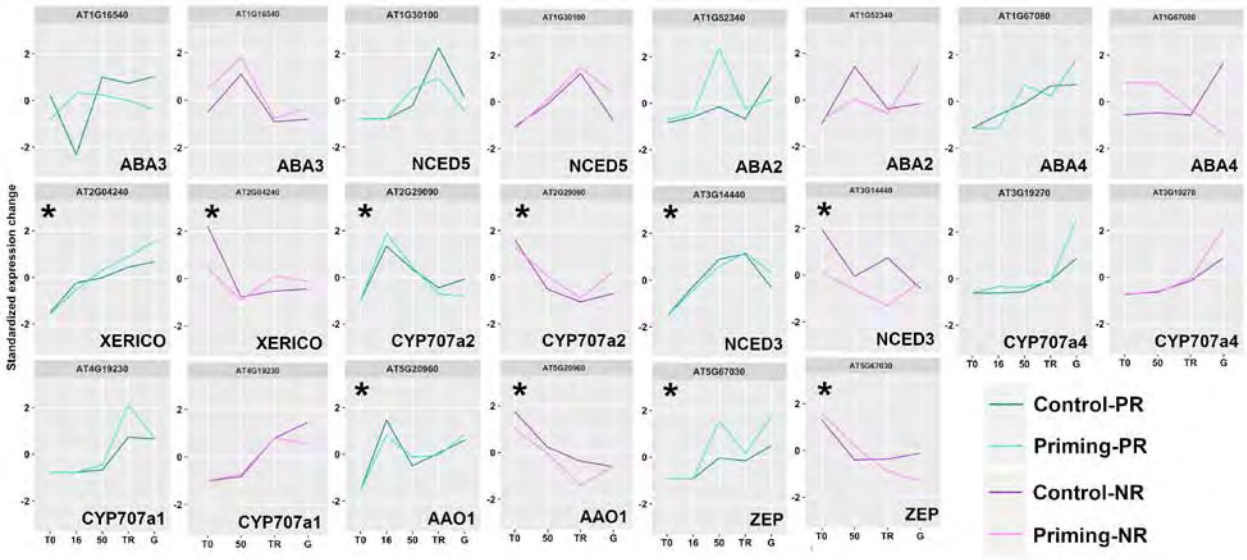


Figure S7. Expression pattern of the identified ABA-metabolism genes in PR- and NR-seeds. The asterisk (*) indicates those genes that had a differential expression at the T0 stage between priming phenotypes.

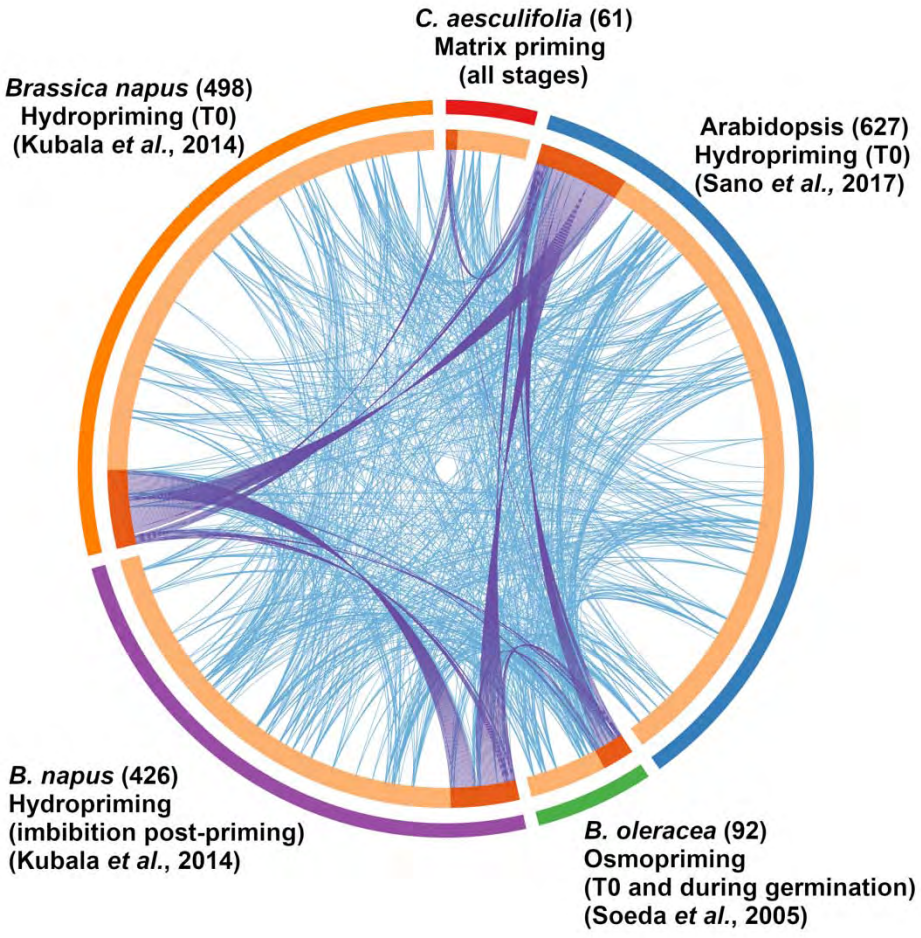


Figure S8. Gene set comparison of the DE genes detected in *C. aesculifolia* and three publicly available gene sets of DE genes in response to different priming methods. The purple lines connect identical genes between gene sets. The blue lines indicate those genes which have common biological functions.

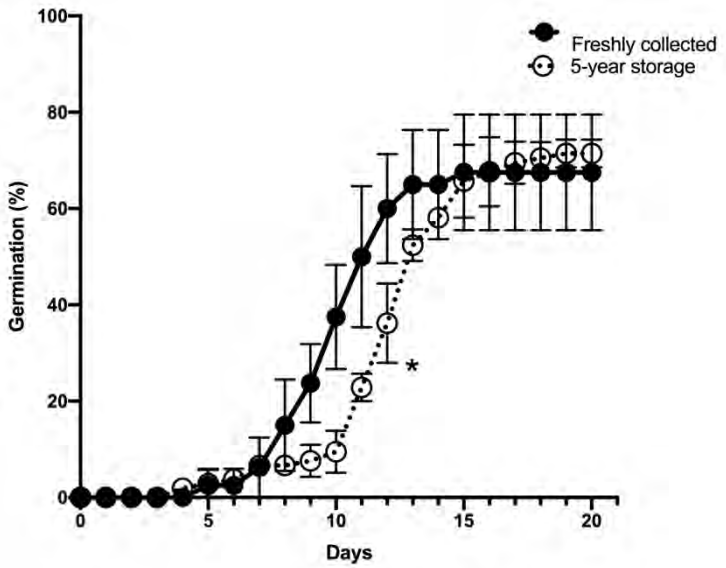


Figure S9. Germination performance of untreated seeds collected in 2012 and tested after 5 years in storage. The * at the middle of the graph corresponds to a statistically significant increase of t_{50} in the stored seeds.