

Informe final* del Proyecto OE001
Ceiba aesculifolia como especie modelo para el estudio molecular del envejecimiento de las semillas en relación al desempeño germinativo*

Responsable: Dra. Alicia Gamboa de Buen
Institución: Universidad Nacional Autónoma de México
Instituto de Ecología
Departamento de Ecología Funcional y Aplicada
Dirección: Av. Universidad # 3000, Ciudad Universitaria, Coyoacán, Ciudad de México, 04510, México
Correo electrónico: agamboa@ecologia.unam.mx
Teléfono/Fax: 5622-8222 ext. 47897
Fecha de inicio: Marzo 31, 2017.
Fecha de término: Abril 6, 2020
Principales resultados: Base de datos, informe final.
Forma de citar el informe final y otros resultados:** Gómez-Maqueo, X., Velázquez-Rosas, N. y A. Gamboa-deBuen. 2020. Ceiba aesculifolia como especie modelo para el estudio molecular del envejecimiento de las semillas en relación al desempeño germinativo. Universidad Nacional Autónoma de México. Instituto de Ecología. **Informe final SNIB-CONABIO, Proyecto No. OE001.** Ciudad de México.

Resumen:

Proyecto financiado con recursos provenientes de la Dirección General del Sector Primario y Recursos Naturales Renovables de la SEMARNAT.

En los últimos años se ha señalado la importancia del estudio de la diversidad genética de las especies para comprender la fisiología de las plantas dentro de un contexto ecológico. En este proyecto hemos seleccionado a Ceiba aesculifolia, una especie silvestre con un gran potencial para análisis genéticos y fisiológicos, así como por su importancia en procesos de restauración y el aprovechamiento de sus frutos (fibra y semilla) en las diferentes localidades donde se distribuye. En particular, estamos interesados en llevar a cabo un análisis transcriptómico del proceso de germinación en función del contenido relativo de agua de las semillas como un parámetro fisiológico adecuado para comparar los diferentes lotes de semillas, provenientes de diferentes años de colecta, y para comparar los diferentes tratamientos que pueden modificar el desempeño germinativo. Los datos obtenidos por medio de las técnicas de secuenciación masiva del transcriptoma de las diferentes muestras nos permitirán obtener una base con datos genéticos de la semilla de Ceiba aesculifolia que podrá ser usada como referencia para otro tipo de estudios. Por un lado, la combinación de los transcriptomas asociados a diferentes tratamientos de acondicionamiento materno y de envejecimiento, junto con sus respectivos grupos control, nos permitirá obtener una base de datos de posibles transcritos asociados a un amplio espectro de procesos biológicos pertenecientes a diversos programas de desarrollo. Por otro lado, el análisis comparativo de los transcriptomas de cada tratamiento nos permitirá discernir entre los distintos procesos biológicos que están asociados con el mejoramiento de la respuesta germinativa y el deterioro de las semillas por el proceso de envejecimiento. Estos análisis a su vez permitirán plantear nuevas hipótesis de trabajo para el manejo de especies mexicanas no-modelo. Este proyecto constituye una parte de la tesis doctoral de la alumna Ximena Gómez Maqueo Bribiesca, bajo la tutoría de los Dres. Alicia Gamboa de Buen (tutor principal) y Luis David Alcaraz Peraza.

-
- * El presente documento no necesariamente contiene los principales resultados del proyecto correspondiente o la descripción de los mismos. Los proyectos apoyados por la CONABIO así como información adicional sobre ellos, pueden consultarse en www.conabio.gob.mx
 - ** El usuario tiene la obligación, de conformidad con el artículo 57 de la LFDA, de citar a los autores de obras individuales, así como a los compiladores. De manera que deberán citarse todos los responsables de los proyectos, que proveyeron datos, así como a la CONABIO como depositaria, compiladora y proveedora de la información. En su caso, el usuario deberá obtener del proveedor la información complementaria sobre la autoría específica de los datos.

INFORME FINAL

Proyecto OE001 “*Ceiba aesculifolia* como especie modelo para el estudio molecular del envejecimiento de las semillas en relación al desempeño germinativo”

**Responsable: Dra. Alicia Gamboa de Buen,
Instituto de Ecología, Universidad Nacional Autónoma de México**

RESUMEN

La distribución en el tiempo de la germinación en un lote de semillas es un rasgo ecológico, es decir existe una dispersión temporal del proceso germinativo. Esta característica es muy variable inclusive en las plantas domesticadas a pesar de estar seleccionadas para tener una germinación sincrónica. Esta variación dificulta especialmente el estudio de este proceso en las plantas silvestres, las cuales presentan una gran plasticidad y variabilidad entre lotes derivados de diferentes poblaciones y/o diferentes años de producción de semillas.

Los estudios fisiológicos que se realizan en plantas modelo como *Arabidopsis thaliana* se basan en la selección de plantas en el mismo estado de desarrollo tomando en cuenta la morfología de la planta entre otros aspectos. Sin embargo, el análisis del proceso germinativo generalmente está basado en función del tiempo, frecuentemente determinado por el investigador. En años recientes, algunos investigadores han descrito el momento de la ruptura de la testa, para las semillas endospermicas, como un estado clave del desarrollo y marcador de la transición entre la germinación temprana y la germinación tardía.

Los resultados obtenidos en este proyecto apoyan la idea de seleccionar estados fisiológicos o de desarrollo apropiados para la comparación entre organismos individuales, poblaciones o respuestas a diferentes estímulos ambientales, especialmente cuando se trabaja con plantas silvestres. El objetivo principal de este proyecto fue establecer el contenido relativo de agua (CRA) como la aproximación a un estado discreto de desarrollo durante el proceso germinativo para la toma de muestras en lugar de utilizar el tiempo de germinación. Esto nos permitió utilizar momentos del desarrollo equivalentes para poder comparar diferentes lotes de semillas con importantes diferencias en el tiempo de imbibición y germinación resultado de historias de vida diferentes debido a su colecta en años distintos y el tiempo de almacenamiento. Las semillas con diferentes contenidos de agua y estados de desarrollo (ruptura de testa y ruptura de endospermo) fueron analizadas utilizando perfiles transcriptómicos y, con base en estos perfiles, fue posible establecer el contenido relativo de agua como un parámetro fisiológico de la semilla durante el proceso de germinación. El análisis de los resultados también nos está permitiendo establecer los mecanismos involucrados en la germinación en respuesta a diversas condiciones ambientales.

Ceiba aesculifolia es una planta muy adecuada para restauración y para la producción de fibras naturales y este trabajo está contribuyendo con información relevante que permitirá profundizar en el conocimiento del género *Ceiba* para su conservación y aprovechamiento sostenible. Finalmente, se contribuyó al conocimiento de la biodiversidad genética ya que se generó una base de datos con 166545 secuencias de transcritos de *C. aesculifolia*, base que puede ser utilizada por diferentes grupos de investigación como referencia para la realización de estudios ecofisiológicos, genéticos y de interacción planta-ambiente específicamente para el género *Ceiba*.

INTRODUCCION

Las semillas son fundamentales para la conservación de los recursos genéticos de las plantas superiores. La determinación del efecto de diferentes condiciones ambientales sobre el desempeño germinativo y del proceso germinativo a nivel molecular de las semillas de especies de plantas silvestres es primordial dado que la mayor parte del conocimiento relacionado ha sido generado en plantas cultivadas. Desde el punto de vista productivo para estos cultivos, el vigor de las semillas se define con base a la velocidad y sincronía de la germinación y el establecimiento de la plántula y su emergencia en condiciones no favorables. La longevidad de las semillas, definida como su capacidad de de continuar siendo viables durante largos periodos de tiempo también es un factor importante en la conservación de los recursos genéticos de las plantas silvestres. Sin embargo, el incremento en la velocidad y la sincronía de la germinación no es un factor a considerar en el desempeño germinativo de la especie de planta silvestre a estudiar. De hecho, la variabilidad en el tiempo de germinación permite respuestas diferentes e individuales a condiciones ambientales similares y confiere a las poblaciones silvestres una ventaja al incrementar las posibilidades de germinar en ambientes con cambio constante (Mitchell *et al.*, 2016). La variabilidad, comprendida como las diferencias naturales que existen entre las diferentes muestras biológicas, se puede manifestar en las semillas a nivel interpoblacional e intrapoblacional. La variabilidad interpoblacional está determinada por diferentes factores que incluyen las características del ambiente materno en el que se desarrolló la semilla y se ve reflejada en los diferentes tiempos para completar la germinación en semillas expuestas a las mismas condiciones experimentales.

Las semillas ortodoxas, que pueden permanecer en el suelo o ser almacenadas por largos periodos de tiempo, presentan una latencia primaria adquirida durante el desarrollo de la semilla en la planta madre. El estado no latente o post-cosecha de la semilla es consecuencia de la remoción de la latencia primaria por condiciones ambientales como almacenamiento en seco. Solo las semillas que se encuentran en este estado son capaces de germinar cuando las condiciones ambientales son favorables. La germinación comienza con la absorción de agua, la cual desencadena de manera progresiva los procesos celulares necesarios para completar

la germinación y el establecimiento de la nueva planta (Rosental *et al.*, 2014). Sin embargo, cuando se inicia el proceso de imbibición en condiciones limitadas de agua, se ha propuesto que las semillas pueden revertir el proceso germinativo e iniciar mecanismos de reparación de macromoléculas y de defensa y, si las condiciones ambientales son inadecuadas, pueden entrar en un estado de segunda latencia. La latencia es uno de los rasgos adaptativos que permite a las semillas evitar la germinación en condiciones desfavorables. Se considera que estos ciclos pueden repetirse sucesivamente como resultado de la capacidad de las semillas de germinar solo cuando las condiciones son favorables (Catusse *et al.*, 2008). Los tratamientos de acondicionamiento que han sido utilizados para promover la sincronización de la germinación, especialmente en semillas de plantas cultivadas, se basan en esta capacidad de las semillas. Este tratamiento puede mejorar el desempeño germinativo de semillas en diferentes especies de plantas incluyendo plantas silvestres, pero también se ha reportado que el tratamiento puede no tener efecto o, inclusive, tener un efecto negativo en el desempeño germinativo.

La germinación depende de la calidad y abundancia de los transcritos acumulados durante el desarrollo de la semilla en función de las condiciones de crecimiento de la planta madre (Rajjou *et al.*, 2012). 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 tanto del tejido estudiado como del estado fisiológico de la semilla (Maia *et al.*, 2011; Dekkers *et al.*, 2013). En *Arabidopsis thaliana* existen dos etapas transcripcionales que corresponden a la etapa previa y a la etapa posterior a la ruptura de la testa. Los autores consideran que la etapa transcripcional previa a la ruptura de la testa es fundamental para la germinación, estableciendo la ruptura de la testa como un proceso fisiológico clave (Dekkers *et al.*, 2013). *Ceiba aesculifolia* es una especie originaria de las regiones tropicales subhúmedas de América y en México se encuentra en Sinaloa, Michoacán, Puebla, Veracruz, México, Morelos, Guerrero, Oaxaca, Chiapas, Yucatán y Quintana Roo. Las semillas son ortodoxas y presentan una gran variabilidad intrapoblacional. Los estudios realizados con semillas de árboles de una población de Actopan, Veracruz, utilizando cosechas de diferentes años muestran amplia heterogeneidad en los diferentes parámetros germinativos y en la respuesta germinativa al acondicionamiento mátrico que puede ser positiva, negativa o sin efecto aparente. Esta variabilidad de la respuesta germinativa de las semillas de *C. aesculifolia* impide la comparación en función del tiempo entre lotes de colectas distintas, que dificulta el análisis molecular (Figura 1).

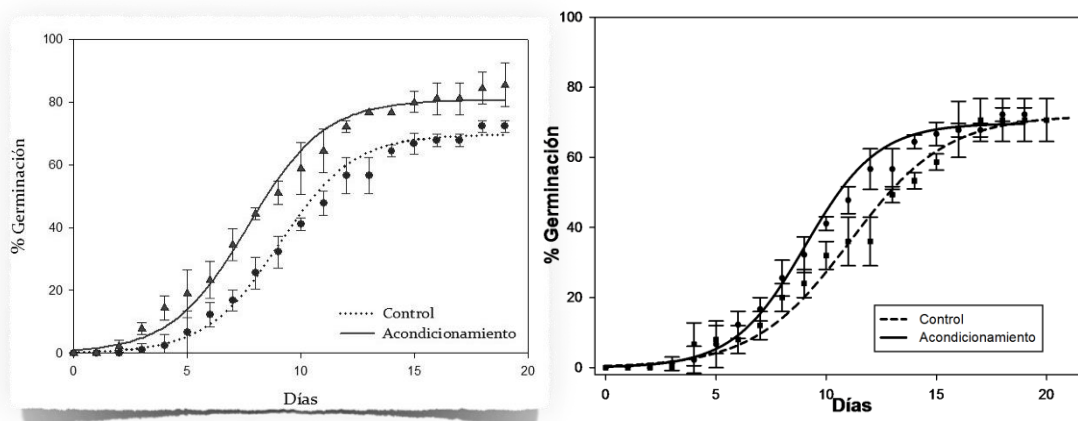


Figura 1. Efecto del tratamiento de acondicionamiento mátrico sobre la germinación de semillas de *Ceiba aesculifolia*. Lote de semillas que si responde al tratamiento: presentan una disminución en el tiempo de inicio, en el tiempo medio y/o en la velocidad de la germinación (panel izquierdo). Lote de semillas que no responde al acondicionamiento; no presentan cambios en ninguno de los parámetros germinativos (panel derecho).

La cinética de absorción de agua durante la imbibición de las semillas de *Ceiba* se puede determinar a través de la medición del contenido relativo de agua (CRA) de las semillas (figura 2).

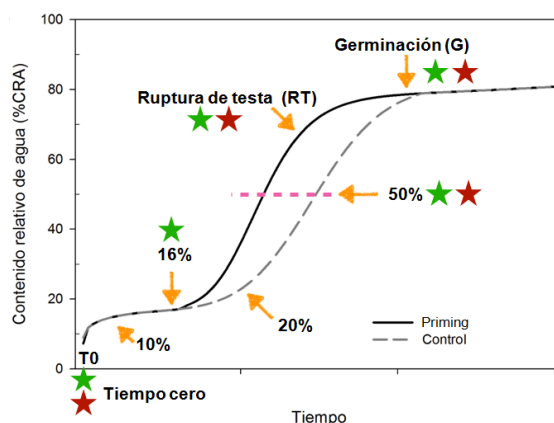


Figura 2. Curva de imbibición de las semillas de *Ceiba aesculifolia* control y tratadas con acondicionamiento de lotes de semillas que si responden al tratamiento. Las semillas de los lotes que no responden al tratamiento presentan una curva de imbibición similar.

OBJETIVOS

El objetivo principal de este trabajo fue determinar los patrones de expresión génica en función del grado de hidratación de la semilla (contenido relativo de agua) para comparar los cambios en el desempeño germinativo causados por el acondicionamiento mátrico y el deterioro a las semillas de *Ceiba aesculifolia*. Estos datos permitieron generar una base de datos del transcriptoma del proceso germinativo para esta especie.

Objetivos particulares

Obtener por medio de técnicas de ensamble de *novo* el transcriptoma de semillas de *C. aesculifolia* durante la germinación y en respuesta a diferentes tratamientos de acondicionamiento mátrico y/o envejecimiento.

Describir los procesos biológicos mayormente representados en los transcriptomas de semillas de *C. aesculifolia* con diferentes contenidos relativos de agua, germinadas en condiciones óptimas (control).

Describir los procesos biológicos mayormente representados en los transcriptomas de semillas tratadas con acondicionamiento mátrico o con deterioro de *Ceiba aesculifolia*.

Comparar los procesos biológicos de las semillas tratadas con respecto al control para validar el uso del contenido relativo de agua como un parámetro fisiológico para el estudio de la germinación y comparación de diferentes lotes (años de colecta y desempeño germinativo) en especies silvestres.

MÉTODOS

Muestras.

Los resultados entregados corresponden a 54 muestras de las cuales 6 muestras corresponden a un ensamble realizado previamente.

Los ensambles 1 (6 muestras) y 2 (48 muestras) corresponden a las siguientes muestras:

M6 (Ensamble 1): Semillas secas T0 Control, tres réplicas CaS2014 (S14_T0_C), CaS2015R (S15R_T0_C) y CaS2015L (S15L_T0_C).

Semillas secas T0 Con Acondicionamiento, tres réplicas CaS2014 (S14_T0_P), CaS2015R (S15R_T0_P) y CaS2015L (S15L_T0_P).

M48 (Ensamble 2): Semillas CRA 16% Control, tres réplicas CaS2014 (S14_16_C), CaS2015R (S15R_16_C) CaS2015L (S15L_16_P).

Semillas CRA16% Control Con Acondicionamiento, tres réplicas CaS2014 (S14_16_P), CaS2015R (S15R_16_P) y CaS2015L (S15L_16_P).

Semillas CRA 50% Control, tres réplicas CaS2014 (S14_50_C), CaS2015R (S15R_50_C) y CaS2015L (S15L_50_C).

Semillas CRA 50% Control Con Acondicionamiento, tres réplicas CaS2014 (S14_50_P), CaS2015R (S15R_50_P) y CaS2015L (S15L_50_P).

Semillas RT Control, tres réplicas CaS2014 (S14_TR_C), CaS2015R (S15R_TR_C) y CaS2015L (S15L_TR_C).

Semillas RT Control Con Acondicionamiento, tres réplicas CaS2014 (S14_TR_P), CaS2015R (S15R_TR_P) y CaS2015L (S15L_TR_P).

Semillas RE Control, tres réplicas CaS2014 (S14_ER_C), CaS2015R (S15R_ER_C) y CaS2015L (S15L_ER_C).

Semillas RE Control Con Acondicionamiento, tres réplicas CaS2014 (S14_ER_P), CaS2015R (S15R_ER_P) y CaS2015L (S15L_ER_P)

Semillas secas T0 que no responden al acondicionamiento, tres réplicas CaS2012 (N12_T0_C), CaS20143 (N143_T0_C) y CaS2016 (N16_T0_C).

Semillas secas T0 que no responden al acondicionamiento. Con acondicionamiento, tres réplicas CaS2012 (N12_T0_P), CaS20143 (N143_T0_P) y CaS2016 (N16_T0_P).

Semillas secas CRA 50% que no responden al acondicionamiento, tres réplicas CaS2012 (N12_50_C), CaS20143 (N143_50_C) y CaS2016 (N16_50_C).

Semillas secas CRA 50% que no responden al acondicionamiento. Con Acondicionamiento, tres réplicas CaS2012 (N12_50_P), CaS20143 (N143_50_P) y CaS2016 (N16_50_CP)

Semillas secas RT que no responden al acondicionamiento, tres réplicas CaS2012 (N12_TR_C), CaS20143 (N143_TR_C) y CaS2016 (N16_TR_C)

Semillas secas RT que no responden al acondicionamiento. Con Acondicionamiento, tres réplicas CaS2012 (N12_TR_P), CaS20143 (N143_TR_P) y CaS2016 (N16_TR_P).

Semillas secas RE que no responden al acondicionamiento, tres réplicas CaS2012 (N12_ER_C), CaS20143 (N143_ER_C) y CaS2016 (N16_ER_C)

Semillas secas RE que no responden al acondicionamiento. Con Acondicionamiento, tres réplicas CaS2012 (N12_ER_P), CaS20143 (N143_ER_P) y CaS2016 (N16_ER_P).

Extracción de RNA y ensamble de novo del transcriptoma.

El RNA total se extrajo de 5 semillas para cada muestra utilizando el protocolo modificado de Li y Trick (2005). Todas las muestras se enviaron a la Unidad Universitaria de Secuenciación Masiva y Bioinformática (UUSMB) del Instituto de Biotecnología, UNAM, donde fueron procesadas para sintetizar el cDNA y construir las librerías de Illumina HiSeq2000. La calidad de las muestras de RNA se determinó utilizando el kit Qubit2.0 y se cuantificaron usando un equipo Bioanalyzer 2100. Las librerías se construyeron de acuerdo a las instrucciones del fabricante (Illumina, San Diego, CA, USA). M6 y M48 fueron secuenciadas *de novo* y ensambladas de manera independiente. Tanto para M6 como M48 se utilizó una aproximación “multiplex” para cargar las muestras correspondientes en la laminilla. Se obtuvieron un total de 1,165,201,460 lecturas pareadas de 100pb incluyendo

M6 y M48. Los adaptadores de Illumina fueron cortados utilizando Trimmomatic y la calidad fue evaluada utilizando FASTQC. Después de la evaluación de la calidad, las lecturas se recortaron a 75pb para asegurar el valor mínimo de calidad de Q20 en todas las bases y las lecturas de baja calidad fueron removidas. Para M6 y M48, se ensambló un transcriptoma global utilizando Trinity v2.1.1 (Grabherr *et al.*, 2011; Haas *et al.*, 2013). Para M6 se encontraron 117,920 secuencias y para M48 398,598 secuencias. Los valores de calidad para el ensamble de las librerías se presentan en el anexo 1.

Identificación de los transcritos para la clasificación funcional

La anotación funcional fue determinada utilizando el programa Trinotate y la herramienta transdecoder que se encuentran dentro del paquete Trinity. Trinotate depende de la búsqueda de homologías en bases de datos de secuencias (BLAST+/SwissProt), e identificación de dominios de proteínas (HMMER/PFAM), lo que nos permitió detectar las fuentes más importantes de contaminantes en las librerías. Para identificar las secuencias codificantes de al menos 100pb en los 6 marcos de lectura se utilizó la herramienta transdecoder. Para identificar el marco correcto de lectura se llevó a cabo un BLASTP. Para este paso se utilizó una base local que incluyó las secuencias de proteínas de otras especies de la familia Malvaceae como *Herrania umbratica*, *Durio zibethinus* y *Theobroma cacao* y de la planta modelo *A. thaliana*. Paralelamente se realizó un filtro para detectar las secuencias contaminantes de los organismos identificados con Trinotate y se generó una base local que incluyó organismos modelo como ratón, los pangenomas de *Pseudomonas aeruginosa* y de levadura, así como secuencias de los hongos *Trichoderma reesei*, *Fusarium fujikuroi*, *F. verticillioides* y *Diplodia corticola*. Se removieron las proteínas que presentaron como mejor alineamiento una de las proteínas presentes en esta base. El mejor valor de alineamiento se seleccionó con los siguientes criterios: una cobertura mínima de 60% y al menos un 40% de identidad. Alrededor del 90% de los transcritos confirmados de plantas pudieron ser asociados a un locus de *A. thaliana*. Finalmente se realizó una validación cruzada de los métodos Trinotate y transdecoder para confirmar la identidad de los transcritos y para explorar las características estructurales de los transcritos ensamblados.

Posteriormente se realizó una validación cruzada utilizando los métodos Trinotate y transdecoder para confirmar la identidad de los transcritos. La base resultante después del paso para eliminar los contaminantes incluye 34,274 transcritos derivados del primer ensamble y 132,271 del segundo ensamble (corresponden a los 166545 transcritos incluidos en la tabla relacional). Estos números incluyen secuencias ensambladas con marcos de lectura abiertos basados en la presencia de codones de inicio y finalización (ATG y STOP) y flanqueados por posibles regiones UTR 5' y 3' UTR, y una mezcla de transcritos incompletos.

Análisis de la expresión diferencial de las categorías funcionales.

Para estos análisis se utilizó la base “IDEAMEX website R-based tools” (<http://zazil.ibt.unam.mx/ideamex/>; Jiménez-Jacinto *et al.*, 2019). Después de una prueba exploratoria, se utilizó DESeq2 para comparaciones pareadas entre estados fisiológicos y fenotipos usando como valores de corte, para considerar expresión diferencial, $P < 0.05$, una FDR < 0.05 y un cambio en la expresión de al menos uno en una escala logarítmica base 2. Para los análisis de PCA, se utilizó el paquete DESeq2 (Love *et al.*, 2014). La clasificación de las categorías funcionales se llevó a cabo en las listas de genes que incrementaba su expresión y en las que disminuía la expresión entre las condiciones de manera individual. Se utilizó la herramienta de anotación funcional en “DAVID Bioinformatics Resources website” (v6.8; <https://david.ncifcrf.gov/>; Huang *et al.*, 2009a, b).

Datos y disponibilidad del material

Las lecturas se depositaron en la plataforma SRA dentro del sitio NCBI GenBank Bioproject PRJNA561202 y números de acceso SAMN12611291 (ensamble 1) y SAMN12611291 (ensamble 2). El transcriptoma ensamblado fue depositado en la plataforma TSA con el número de acceso GHVB000000000.

Se realizó y se entregó una tabla relacional que incluye 166545 transcritos completos e incompletos que pudieron ser identificados como secuencias de ‘plantas (objetivo 1).

ÁREA DE COLECTA

Los frutos de *Ceiba aesculifolia* se colectaron de 10 a 13 árboles localizados en un ecosistema perturbado con actividad humana en las localidades de “Trapiche del Rosario y Chicuasén” en Actopan Veracruz, México (latitud 19.5426, longitud -96.7401, altitud 479 m a.s.l.).

RESULTADOS

El Contenido Relativo de Agua (CRA) es un parámetro fisiológico que permite la comparación entre lotes de semillas de Ceiba aesculifolia que han sido colectadas en diferentes años y sometidas a diferentes tratamientos.

La germinación es un proceso que implica tanto la síntesis de ARN mensajeros como la depuración de los ARN mensajeros que corresponden al plan de desarrollo de maduración de la semilla. Por lo tanto, consideramos importante utilizar para los siguientes análisis únicamente los transcritos identificados como completos tal como se describió en materiales y métodos. En la figura 3 se muestra el resultado del análisis de componentes principales. El primer componente corresponde al 79% de la variación y resuelve las diferencias entre los lotes de semillas secas control y con acondicionamiento que si responden al tratamiento y todos los lotes de semillas durante el proceso de imbibición y los lotes de semillas control y

con acondicionamiento que no responden al tratamiento. En general, las seis réplicas correspondientes a cada lote con y sin tratamiento se agrupan en función del valor de CRA, de la ruptura de testa y de la ruptura de endospermo (PC2: 17% de la variación) (Figura 3).

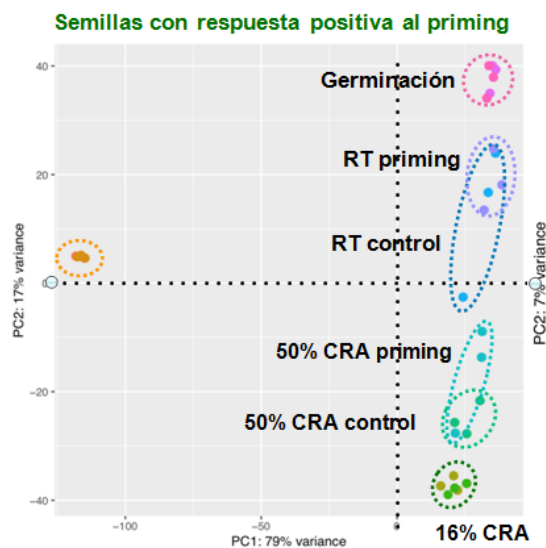


Figura 3. Análisis de componentes principales de las 30 muestras correspondientes a los tres lotes de semillas que sí responden al acondicionamiento o “priming”.

Los lotes de semillas de *C. aesculifolia* que sí responden al acondicionamiento presentan las dos fases transcripcionales descritas previamente para semillas de *A. thaliana* considerando que la ruptura de la testa es el momento fisiológico que determina el paso a la segunda etapa. La primera fase para las semillas de *C. aesculifolia* incluye a las semillas imbibidas hasta un CRA 16% y consiste en una importante síntesis de nuevos transcritos (2727 para las semillas control y 464 para las semillas con acondicionamiento) y en una importante posible depuración de transcritos (2375 para las semillas control y 785 para las semillas con acondicionamiento). Los cambios transcripcionales entre el 16 y el 50% también corresponden a esta fase. La segunda fase transcripcional iniciaría con la ruptura de la testa, proceso en el cual se expresan 645 genes para las semillas control y 79 para las semillas con acondicionamiento. Considerando esta fase transcripcional como otra etapa del desarrollo es congruente la posible depuración de 453 transcritos en las semillas control y 82 en las semillas con tratamiento. Durante la ruptura del endospermo o germinación también se detectan nuevos transcritos aunque en menor proporción (Figura 5).

La primera fase transcripcional de la germinación descrita por Dekkers et al. (2013) está representada por la expresión de genes pertenecientes a procesos involucrados en iniciar el metabolismo y continuar con la expresión de genes y síntesis de proteínas. En los últimos años también se ha señalado la importancia de la edición alternativa en el proceso germinativo. Los grupos funcionales, que se incrementan en comparación con el tiempo T0 y que son correspondientes a esta primera fase en el proceso germinativo de los lotes de semillas de *C. aesculifolia* que si responden al acondicionamiento incluyen, con respecto a la síntesis de biomoléculas, a los siguientes grupos: replicación del ADN, procesamiento y edición alternativa de ARN mensajero, procesamiento de ARN ribosomal, biogénesis del ribosoma, traducción de proteínas y su regulación, ensamblaje y transporte de proteínas y proteólisis. Asimismo, la expresión de genes relacionados con el metabolismo se incrementa, especialmente los correspondientes al ciclo del ácido tricarboxílico (TCA) procesos glicolíticos, fotosíntesis y fosforilación de carbohidratos. El incremento en la expresión de genes relacionados con transducción de señales involucra la respuesta a fitohormonas como ABA y etileno y a azúcares y a factores ambientales como cadmio y calor. Con respecto al desarrollo de la plántula, se expresan genes involucrados con el desarrollo de la raíz. Considerando a los mARN que posiblemente se depuran se han identificado los genes relacionados con desarrollo del embrión y maduración de la semilla. Los grupos funcionales antes descritos corresponden a la comparación entre las semillas con un CRA de 16% con respecto al T0 y estos resultados sugieren que estas semillas ya iniciaron el proceso germinativo. Aunque no podemos concluir de manera contundente si el proceso germinativo es reversible cuando la semilla alcanza el valor de 16% CRA, los resultados correspondientes a la comparación entre 16% CRA y 50% CRA sugieren que el proceso germinativo cuando las semillas alcanzan el 16% CRA es irreversible puesto que solo se incrementa la expresión de un bajo porcentaje de genes específicamente los relacionados con proteólisis.

Los procesos que se presentan durante la ruptura de testa, correspondientes a la segunda fase transcripcional, son las diferentes vías metabólicas que incluyen el metabolismo de ácidos grasos, del glioxilato y el ciclo de los ácidos tricarboxílicos. También se incrementa la expresión de genes relacionados con procesos de oxidación y reducción y de catabolismo del peróxido de hidrógeno. El incremento de genes relacionados con fitohormonas incluye la respuesta a citocininas y a auxinas y la biosíntesis de brasinosteroides. Por otro lado, se induce la respuesta a karriquina, hipoxia, frío y escasez de agua. Con respecto al desarrollo de la plántula, se expresan genes relacionados con el desarrollo del tejido vascular de los cotiledones y el desarrollo del tallo; también hay un incremento importante en los genes relacionados con la reorganización de la pared celular incluyendo la formación de la placa celular durante la citocinesis.

La organización de la pared celular y la homeostasis del agua corresponden a procesos relevantes que se expresan durante la ruptura del endospermo junto con la fijación del carbono y el transporte de urea. Con respecto a las hormonas, se detectó un incremento en la expresión de genes relacionados con la homeostasis de auxinas y respuesta a giberelinas y

se induce la respuesta a daño por herida. Tres procesos importantes en los que se detecta una baja en la expresión de genes son el almacenamiento de lípidos, la formación de cuerpos oleicos y la respuesta a ABA.

En el cuadro 1 se presentan los procesos asociados a cada uno de los momentos fisiológicos estudiados. El avance del proceso germinativo se refleja en los diferentes procesos; a 16% CRA se incrementan los transcritos relacionados con la regulación y la señalización de las hormonas asociadas a estrés ABA y etileno. A 50% CRA, se incrementan los transcritos relacionados con las citocininas, las auxinas y los brassinosteroides, hormonas asociadas al desarrollo y finalmente, durante la ruptura de testa, los transcritos relacionados con las giberelinas. En relación al medio ambiente, se presentan transcritos asociados a diferentes estreses en cada etapa.

Cuadro 1. Procesos asociados a cada estado fisiológico durante la germinación de semillas que sí responden al acondicionamiento.

Procesos	16%	50%	Ruptura de testa	Ruptura de endospermo
Metabolismo relacionado con ADN, ARN y proteínas	<ul style="list-style-type: none"> ✓ Replicación ADN ✓ Síntesis ARN ✓ Edición alternativa ARN ✓ Traducción 	<ul style="list-style-type: none"> ✓ Proteolisis (degradación de proteínas de reserva) 		
Metabolismo intermediario	<ul style="list-style-type: none"> ✓ TCA ✓ Glucólisis ✓ Fotosíntesis 		<ul style="list-style-type: none"> ✓ TCA ✓ Metabolismo lípidos ✓ Vía del glioxilato ✓ Reacciones oxido-reducción 	<ul style="list-style-type: none"> ✓ Fijación de carbono ✓ Transporte de urea
Señalización	<ul style="list-style-type: none"> ✓ ABA ✓ Etileno ✓ Azúcares 		<ul style="list-style-type: none"> ✓ Citocininas ✓ Auxinas ✓ Brasinosteroides ✓ Karrikina 	<ul style="list-style-type: none"> ✓ Homeostasis auxinas ✓ Giberelinas
Desarrollo	<ul style="list-style-type: none"> ✓ Raíz ✓ Tilacoides 		<ul style="list-style-type: none"> ✓ Tejido vascular en cotiledones ✓ Tallo 	<ul style="list-style-type: none"> ✓

Procesos	16%	50%	Ruptura de testa	Ruptura de endospermo
Respuesta al ambiente	<ul style="list-style-type: none"> ✓ Cadmio ✓ Calor 		<ul style="list-style-type: none"> ✓ Hipoxia ✓ Frío ✓ Escasez de agua 	<ul style="list-style-type: none"> ✓ Heridas ✓ Homeostasis Agua
Pared celular			<ul style="list-style-type: none"> ✓ Organización de la pared ✓ Formación de placa celular 	

El tratamiento de acondicionamiento promueve cambios en los niveles de transcripción en todas las etapas fisiológicas del proceso germinativo

Se llevó a cabo un análisis comparativo de las semillas control y con tratamiento para cada estado fisiológico. Los cambios en los grupos funcionales detectados en las semillas tratadas con acondicionamiento correspondientes a T0 y 16% CRA no fueron significativos. Sin embargo, para el T0 se detectó un incremento significativo en la detección de 14 transcritos relacionados con regulación post-transcripcional, respuesta a estímulo mecánico y ubiquitinación de proteínas. La detección de transcritos relacionados con la síntesis de 15 transcritos relacionados con síntesis de rafnosa y transporte de auxinas fue menor. Para las semillas de 50% se incrementa la expresión de genes relacionados con estrés, metabolismo y transporte de lípidos, fotosíntesis y metabolismo de glutamina. El incremento en los genes de respuesta a estrés también ocurre durante la ruptura de testa incluyendo la señalización por jasmónico. Finalmente, durante la ruptura del endospermo hay una mayor respuesta a estrés por falta de agua y de genes relacionados con el ensamble de microtúbulos.

Las semillas secas (T0) de Ceiba aesculifolia que no responden al acondicionamiento presentan un perfil transcriptómico relacionado con el inicio de la germinación.

Los lotes de semillas control y con acondicionamiento que no responden al tratamiento y corresponden a las semillas secas, con un CRA de 6-8%, se agrupan con las semillas con un CRA de 16% (lotes de semillas control y con acondicionamiento que si responden al tratamiento). Estos resultados indican que estas semillas, que corresponden a lotes almacenados durante varios años (2012 y 2014-3) o lotes recién colectados con un bajo porcentaje de germinación (2016), presentan el perfil transcriptómico correspondiente al inicio del proceso germinativo (Figura 4).

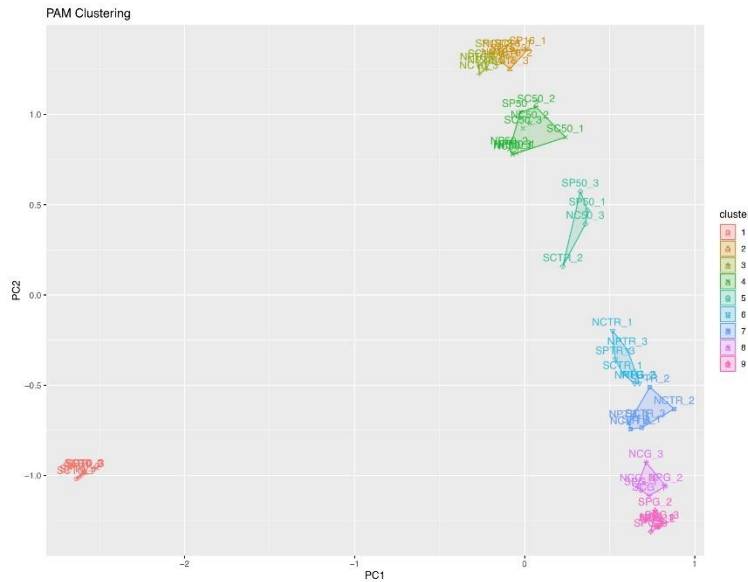


Figura 4. Análisis de componentes principales de las 54 muestras correspondientes a los tres lotes de semillas que si responden al acondicionamiento (SC y SP) y a los tres lotes que no responden al acondicionamiento (NC y NP). “Cluster” naranja, semillas T0 de los lotes que si responden. “Cluster” café, semillas con CRA16% de los lotes que, si responden, cluster verde musgo, semillas T0 de los lotes que no responden. “Clusters” verdes, semillas con CRA50% de los lotes que si responden y de los lotes que no responden al acondicionamiento. “Clusters” azules, semillas con RT de los lotes que si responden y de los lotes que no responden al tratamiento. “Clusters” rosas, semillas con RE de los lotes que si responden y de los lotes que no responden al tratamiento.

Para comprobar que las semillas T0 de los lotes que no responden al acondicionamiento presentan un perfil transcriptómico equivalente al inicio de la germinación (16% CRA) se llevo a cabo una comparación entre los dos perfiles transcriptómicos. Las diferencias se encontraron en la mayor expresión de genes relacionados con la respuesta al cadmio, el desarrollo de la hoja, la proteólisis dependiente de ubiquitina, la biosíntesis de almidón y el metabolismo de los aminoácidos aromáticos. Por lo tanto, no hubo diferencias en los procesos relacionados con el inicio del metabolismo y la expresión de genes y síntesis de proteínas que caracterizan a la primera fase transcripcional.

A diferencia de lo descrito para las semillas control con 50% CRA, las semillas que no responden al acondicionamiento presentaron un cambio importante en la expresión de genes perteneciendo la mayoría a grupos funcionales relacionados con la producción de metabolitos secundarios, síntesis de jasmónico y respuesta a estreses abióticos y bióticos.

El incremento en transcritos relacionados con metabolismo y con procesos de oxidación y reducción se presenta durante la ruptura de testa tal como se describió para las semillas control. Asimismo, hay un incremento en la expresión de genes relacionados con pared

celular y respuesta a diferentes tipos de estrés. Con respecto a las hormonas hay diferencias importantes ya que se detectaron transcritos relacionados con el ácido giberélico, pero no con auxinas o brasinosteroides. Finalmente, no hubo cambio durante la ruptura de endospermo. Estos resultados apoyan que el perfil transcriptómico de las semillas T0 que no responden al acondicionamiento corresponde al de la primera fase transcripcional y la segunda fase transcripcional si ocurre durante la ruptura de la testa tal como describió para las semillas control y las semillas de *A. thaliana*.

Cuadro 2. Procesos asociados a cada estado fisiológico durante la germinación de semillas que no responden al acondicionamiento

Procesos	50%	Ruptura de testa	Ruptura de endospermo
Metabolismo relacionado con ADN, ARN y proteínas			
Metabolismo intermediario	Procesos de oxidación-reducción	Fotosíntesis Metabolismo de lípidos Vía del glioxilato Procesos de oxidación-reducción Fijación de Carbono Metabolismo de carbohidratos	
Señalización		Transporte de auxinas Giberelinas Karrikina	
Desarrollo		Gineceo Hoja	
Respuesta al ambiente	Estrés oxidativo Escasez de agua Hipoxia Estrés salino Frío Hongos Bacterias	Frío Luz azul Estrés oxidativo	

Procesos	50%	Ruptura de testa	Ruptura de endospermo
	Herida		
Pared celular		Organización de la pared celular	

Las semillas que no respondieron al acondicionamiento presentan una importante respuesta a los estímulos ambientales tanto abióticos como bióticos durante el proceso germinativo como puede observarse en el cuadro 2. Es importante señalar que las vías metabólicas asociadas a la ruptura de la testa en estas semillas se comparten con las de las semillas que si responden al acondicionamiento.

El acondicionamiento si altera la expresión de genes durante el proceso germinativo de las semillas de los lotes que no responden al tratamiento

En las semillas imbibidas hasta un CRA 50% se incrementa la expresión de nuevos transcritos (559 para las semillas control y 862 para las semillas con acondicionamiento) y también se detecta una disminución en la detección de algunos transcritos (247 para las semillas control y 610 para las semillas con acondicionamiento). Los cambios transcripcionales entre el 50% y la ruptura de testa se detectaron cambios muy importantes en las semillas tratadas con acondicionamiento. Durante esta transición se incrementa la expresión de 762 genes para las semillas control y 1656 para las semillas con acondicionamiento y disminuye la expresión de 470 y 1092 transcritos respectivamente. Finalmente, al comparar la expresión entre la ruptura de testa y la ruptura del endospermo, se detectó un aumento de 21 para las semillas control y 160 para las semillas con tratamiento y una disminución de 56 y 24 respectivamente (figura 5).

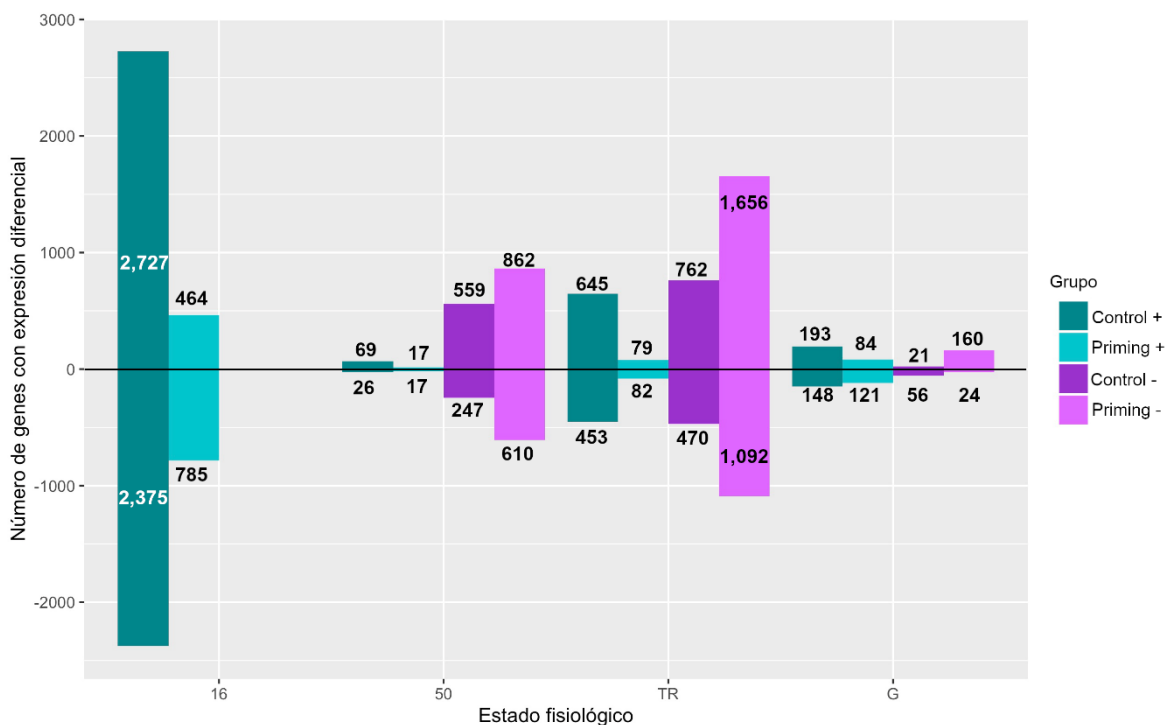


Figura 5. Perfiles transcriptómicos asociados a cada etapa de la germinación en las semillas de los lotes control (verde) y con acondicionamiento (azul) que, si responden al tratamiento, y lotes control (morado) y con acondicionamiento (rosa) que no responden al tratamiento. La diferencia en el número de transcritos se estableció con respecto al estado fisiológico anterior tal como fue realizado por Dekkers et al. (2013).

El análisis realizado para este informe es un análisis global que nos permite visualizar la importancia de este proyecto para el estudio de los mecanismos presentes en el proceso germinativo para responder a diferentes condiciones ambientales. El análisis de los mecanismos específicos para cada condición se presentará en las diferentes publicaciones en revistas nacionales e internacionales resultado de este proyecto.

En conclusión, este trabajo demuestra que el contenido relativo de agua puede ser utilizado como un indicador del estado fisiológico de la semilla para poder llevar a cabo comparaciones entre diferentes lotes y tratamientos (objetivo 2). Por otro lado, se demuestra que el perfil transcriptómico de las semillas secas de los lotes que no responden al tratamiento corresponde a un programa relacionado con los eventos tempranos de la germinación (objetivo 3 y 4). Es importante aclarar que dos de los lotes utilizados presentaron un desempeño germinativo óptimo aunque no respondieron al acondicionamiento. Estos resultados sugieren que estas semillas, aunque no han perdido vigor, presentan un estado de desarrollo diferente y que, posiblemente, ya no respondan adecuadamente a los ciclos de hidratación-deshidratación que ocurren en el suelo.

Finalmente, se contribuyó al conocimiento de la biodiversidad genética ya que se generó una base de datos con alrededor de 350,000 secuencias de transcritos de *C. aesculifolia* y esta base de datos puede ser utilizada por diferentes grupos de investigación como referencia para estudios ecofisiológicos, genéticos y de interacción planta-ambiente

REFERENCIAS

- Bewley JD (1997) Seed germination and dormancy. *Plant Cell* 9: 1055.
- Catusse J, Job C, Job D (2008) Transcriptome-and proteome-wide analyses of seed germination. *Comptes rendus biologiques* 331: 815-822.
- Dekkers BJ, Pearce S, van Bolderen-Veldkamp RP, Marshall A, Widera P, Gilbert J, Wood AT (2013) Transcriptional dynamics of two seed compartments with opposing roles in Arabidopsis seed germination. *Plant Physiol* 163: 205-215.
- Garza-Caligaris LE, Avendaño-Vázquez AO, Alvarado-López S, Zúñiga-Sánchez E, Orozco-Segovia A, Gamboa-deBuen A (2012) *At3g08030* transcript: a molecular marker of seed ageing. *Ann Bot* 1109: 1253-1260.
- Gómez Maqueo Bribiesca X (2014) Efecto del acondicionamiento sobre el metabolismo en semillas de especies nativas. Tesis de Maestría, Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech* 29: 644.
- Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber MD (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protocols* 8: 1494.
- Huang DW, Sherman BT, Lempicki RA (2009a) Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nat Protocols* 4: 44-57.
- Huang DW, Sherman BT, Lempicki RA (2009b) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37: 1-13.
- Jimenez-Jacinto V, Sanchez-Flores A, Vega-Alvarado L (2019) Integrative Differential Expression Analysis for Multiple EXperiments (IDEAMEX): A Web server tool for integrated RNA-seq data analysis. *Front Gen* 10: 279.
- Li Z, Trick H (2005) Rapid method for high-quality RNA isolation from seed endosperm containing high levels of starch. *BioTechniques* 38: 872-876.
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15.
- Maia J, Dekkers BJW, Provart NJ, Ligterink W, Hilhorst HWM. 2011. The re-establishment of desiccation tolerance in germinated *Arabidopsis thaliana* seeds and its associated transcriptome. *PLoS ONE* 6: e29123.

Mitchell J, Johnston IG, Bassel GW (2016) Variability in seeds: biological, ecological, and agricultural implications. *J Exp Bot* 68: 809-817.

Rajjou L, Duval M, Gallardo K, Catusse J, Bally J, Job C, Job D (2012) Seed germination and vigor. *Ann Rev Plant Biol* 63: 507-533.

Rosental L, Nonogaki H, Fait A (2014) Activation and regulation of primary metabolism during seed germination. *Seed Sci Res* 24:1-15.

Soeda Y, Konings MC, Vorst O, van Houwelingen AM, Stoopen GM, Maliepaard CA, van der Geest AH (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 Physiol* 137: 354-368.