

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
FACULTAD DE QUÍMICA**

**RELACIÓN DE LAS CARACTERÍSTICAS FÍSICOQUÍMICAS DE PROTEÍNAS  
INTRÍNSECAMENTE DESORDENADAS CON LOS PATRONES DE  
EXPRESIÓN TRANSCRIPCIONAL EN ORGANISMOS TOLERANTES A LA  
DESECACIÓN**

**TESIS**

**QUE PARA OBTENER EL TÍTULO DE:  
QUÍMICO FARMACÉUTICO BIÓLOGO**

**P R E S E N T A:  
DIEGO PÉREZ VILLANUEVA**

**TUTOR:  
DR. CÉSAR LUIS CUEVAS VELÁZQUEZ**

**CIUDAD UNIVERSITARIA , CD.MX. 2023**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

**Jurado asignado:**

Presidente: **DRA. GAVILANES RUIZ MARINA**

Vocal: **DRA. DIMITROVA DINKOVA TZVETANKA**

Secretario: **DR. CUEVAS VELAZQUEZ CESAR LUIS**

Suplente 1: **DR. PLASCENCIA DE LA PARRA FRANCISCO JAVIER**

Suplente 2: **DR. GUITIERREZ AGUILAR MANUEL**

**Sitio donde se desarrolló el tema:**

Departamento de Bioquímica, Facultad de Química, Universidad Autónoma de México,  
Conjunto E, laboratorio 103

**ASESOR DEL TEMA:** DR. CÉSAR LUIS CUEVAS VELÁZQUEZ

**SUPERVISOR TÉCNICO:** DRA. MARÍA FERNANDA CORNEJO GRANADOS

**SUSTENTANTE:** DIEGO PÉREZ VILLANUEVA

## **Agradecimientos**

Este material está basado en trabajo realizado gracias al financiamiento del Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la UNAM IA203422; por lo que agradezco a la DGAPA-UNAM la beca recibida.

De igual manera agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el donativo de Ciencia de Frontera 2019 CONACYT con clave CF-2019 / 252952 al Dr. César Luis Cuevas Velázquez. Doy gracias por la beca recibida con numero de becario 30955 de la convocatoria titulada “Ciencia de Frontera 2019”.

Agradezco al financiamiento del University of California Institute for Mexico and the United States (UC MEXUS) y el Consejo Nacional de Ciencia y Tecnología de México (CONACYT). Proyecto CN-20-113 titulado “Sensibilidad al ambiente de proteínas intrínsecamente desordenadas de plantas inducidas por deshidratación”.

Al Programa de Apoyo a la Investigación y el Posgrado (PAIP) de la Facultad de Química con la clave 5000-9182 al Dr. César Luis Cuevas Velázquez.

Adicionalmente doy las gracias a la Dra. Fernanda Cornejo y al Dr. César Cuevas por guiarme en mi desarrollo tanto personal como profesional para poder obtener mi tesis de licenciatura. Por último, le agradezco a mi familia por el apoyo incondicional que me han dado siempre.

## Índice

|              |  |           |
|--------------|--|-----------|
| <b>I.</b>    | <b>Resumen</b>   | <b>1</b>  |
| <b>II.</b>   | <b>Introducción</b>  | <b>3</b>  |
|              | II.I. Estrés hídrico en los seres vivos  | 3         |
|              | II.II. Proteínas inducidas en respuesta a la desecación                              | 5         |
|              | II.III. Proteínas intrínsecamente desordenadas                                       | 7         |
|              | II.IV. Proteínas abundantes en la embriogénesis tardía                               | 8         |
|              | II.V. Proteínas tipo-LEA   | 9         |
|              | II.VI. <i>Hypsibius dujardini</i> , un animal (tardígrado) tolerante a la desecación | 12        |
|              | II.VII. <i>Cronobacter sakazakii</i> , una bacteria tolerante a la desecación        | 13        |
| <b>III.</b>  | <b>Antecedentes</b>  | <b>14</b> |
| <b>IV.</b>   | <b>Hipótesis</b>   | <b>17</b> |
| <b>V.</b>    | <b>Objetivo general</b>  | <b>17</b> |
| <b>VI.</b>   | <b>Objetivos particulares</b>  | <b>17</b> |
| <b>VII.</b>  | <b>Metodología</b>   | <b>18</b> |
|              | VII.I. Datos   | 18        |
|              | VII.II. Procesamiento de las lecturas de RNA-seq                                     | 19        |
|              | VII.III. Análisis bioinformático de las secuencias de proteínas                      | 20        |
|              | VII.IV. Determinación de la expresión de proteínas tipo-LEA                          | 21        |
| <b>VIII.</b> | <b>Resultados</b>  | <b>22</b> |
| <b>IX.</b>   | <b>Discusión</b>   | <b>34</b> |
| <b>X.</b>    | <b>Conclusión</b>  | <b>38</b> |
| <b>XI.</b>   | <b>Perspectivas</b>  | <b>39</b> |
| <b>XII.</b>  | <b>Referencias</b>   | <b>40</b> |

## Abreviaturas

**ADN** Ácido desoxirribonucleico

**ARN** Ácido ribonucleico

**ROS** Especies reactivas de oxígeno (*reactive oxygen species*)

**IDPs** Proteínas intrínsecamente desordenadas (*intrinsically disordered proteins*)

**LEA** Proteínas abundantes en la embriogénesis tardía (*late embryogenesis abundant proteins*)

**HSPs** Proteínas de choque térmico (*heat shock proteins*)

**sHSPs** Pequeñas proteínas de choque térmico (*small heat shock proteins*)

**ELIP** Proteínas tempranamente inducidas por luz (*early light-induced proteins*)

**TDPs** Proteínas intrínsecamente desordenadas específicas del tardígrado (*tardigrade-specific intrinsically disordered proteins*)

**CAHS** Proteínas abundantes en el citosol solubles al calor (*Cytosolic abundant heat soluble proteins*)

## I. Resumen

Todas las formas de vida en el planeta necesitan agua para vivir. Algunos organismos pueden tolerar la pérdida casi total del contenido de agua intracelular. A estos organismos se les conoce como organismos tolerantes a la desecación. Entre los organismos tolerantes a la desecación se encuentran la mayoría de las semillas de las plantas. Para tolerar la desecación, las semillas acumulan un grupo de proteínas llamadas proteínas abundantes en la embriogénesis tardía (LEA, *Late Embryogenesis Abundant*). Una de las funciones de las proteínas LEA durante condiciones de desecación es la de ser chaperonas no clásicas que estabilizan la estructura de proteínas (y otras macromoléculas) durante la pérdida de agua. Una característica que distingue a las proteínas LEA de otras proteínas, y que podría ser de gran importancia para desempeñar sus funciones durante la desecación, es su desorden estructural intrínseco. Dado que las proteínas LEA carecen de una estructura tridimensional estable, se ha propuesto que las propiedades fisicoquímicas derivadas de su secuencia de aminoácidos son cruciales para realizar sus funciones. Previamente en nuestro laboratorio se encontró que valores particulares de algunos parámetros fisicoquímicos permiten diferenciar a las proteínas LEA del resto de proteínas del proteoma de algunas especies plantas. Específicamente, las proteínas LEA contienen una alta composición de aminoácidos promotores del desorden, una alta hidrofiliidad, así como una carga neta media baja. A las proteínas que cumplen con dichos parámetros se les denominó proteínas tipo-LEA. Interesantemente, proteínas que cumplen con las características de las proteínas tipo-LEA existen en organismos tolerantes a la desecación, lo que sugiere que las proteínas tipo-LEA de estos organismos tienen una actividad tipo-chaperona, sin embargo, no sabemos si este tipo de proteínas se acumulan durante condiciones de desecación en dichos organismos.

En este trabajo nos enfocamos en determinar si las proteínas tipo-LEA se acumulan en organismos tolerantes durante condiciones de desecación. Para ello, buscamos y analizamos los transcriptomas de organismos tolerantes en condiciones húmedas y secas. Usando los datos de RNA-seq públicamente disponibles de muestras hidratadas y desecadas del tardígrado *Hypsibius dujardini* y de la bacteria Gram-negativa *Cronobacter sakazakii*, obtuvimos los niveles de expresión de transcritos

que codifican para proteínas tipo-LEA. Confirmamos que al igual que las proteínas LEA canónicas en plantas, existen transcritos que codifican para proteínas tipo-LEA que se acumulan durante la desecación en los organismos tolerantes. Con base en la similitud con las proteínas LEA, específicamente a nivel de composición de aminoácidos, así como en el patrón de expresión de los transcritos que las codifican, sugerimos que las proteínas tipo-LEA cuyos transcritos se acumulan en desecación tendrán una función protectora tipo-chaperona *in vitro*. En general, las propiedades fisicoquímicas como lo son la carga neutra, la hidrofiliidad y el desorden, en conjunto con los niveles de acumulación durante la desecación representan una forma innovadora de encontrar proteínas con posibles funciones protectoras relevantes durante la desecación en organismos tolerantes.



## II. Introducción

### II.I. Estrés hídrico en los seres vivos

La vida como la conocemos es imposible sin agua y en efecto, su ausencia en el ambiente limita la presencia y el desarrollo de los seres vivos. Una pérdida de agua a nivel celular desestabiliza la estructura de las macromoléculas (ADN, ARN, lípidos y proteínas) en el interior de las células, lo que puede llevar a la muerte (Hibshman et al., 2020). Cuando un organismo experimenta falta de agua se dice que este se encuentra bajo condiciones de estrés hídrico. El daño que puede llegar a causar este estrés depende de la severidad de la pérdida de agua y dos términos son los más utilizados para diferenciarla: deshidratación y desecación. La deshidratación corresponde a una pérdida de agua leve, mientras que la desecación define la pérdida extrema de agua (disminución de más del 90% del contenido de agua) (Oliver et al., 2020). Existen variaciones de estos términos dependiendo del autor. En este trabajo se hace referencia a las definiciones antes mencionadas.

Si bien los seres vivos han desarrollado maneras de resistir a estas condiciones inadecuadas del ambiente, no todos poseen la misma capacidad de resistencia a la deshidratación o a la desecación. Para el primer caso (deshidratación) se habla de una tolerancia a la falta de agua, donde los organismos que la presentan mantienen un desequilibrio constante entre el agua intracelular y la que se encuentra en el ambiente. Este es el caso de los organismos sensibles a la desecación (Alpert et al., 2006). En el caso extremo de la pérdida de agua, Bewley define a la tolerancia a la desecación como la aptitud de “revivir de un estado *air-dry* (donde la humedad relativa del aire es baja)”, o de manera más simple, la capacidad de revivir a un porcentaje de agua celular igual al del aire (Bewley, 1979). Existen organismos tolerantes a la desecación en todos los reinos de la vida. Algunos ejemplos los encontramos en bacterias generalmente Gram-positivas (Grzyb & Skłodowska., 2022), algas (Trainor & Gladych, 1995), levaduras (Ren et al., 2020), semillas (Maia et al., 2014), plantas de resurrección (Xiao et al., 2015), animales como los tardígrados (Boothby et al., 2017), entre otros. Los organismos

tolerantes a la desecación han desarrollado distintos mecanismos de resistencia contra el daño causado por dichas condiciones. Entre estos mecanismos destacan la modificación de la composición y/o re arreglo de las biomembranas para mejorar su estabilidad, estrategias de evasión del daño generado por especies reactivas de oxígeno (ROS), la producción de disacáridos no reductores (e.g., trehalosa) y la síntesis de proteínas específicas (e.g., proteínas abundantes en la embriogénesis tardía), que al igual que los disacáridos, previenen la agregación de proteínas y membranas (Oliver et al., 2020; Rabert et al., 2020; Tapia et al., 2015). En esta tesis nos centraremos en las proteínas que previenen la agregación proteica.

## II.II. Proteínas inducidas por la desecación

Como respuesta a la desecación, los organismos tolerantes inducen la biosíntesis de una serie de proteínas específicas. Entre ellas encontramos a las proteínas inducidas tempranamente por luz (ELIPs, *Early Light-Induced Proteins*), proteínas de choque térmico (HSPs, *Heat-Shock Proteins*), pequeñas proteínas de choque térmico (sHSPs, *small Heat-Shock Proteins*), enzimas antioxidantes, y algunos ejemplos en organismos puntuales son las proteínas localizadas en islas LEA (PvLIL, *Polypedilum vanderplanki LEA-Island Located*) del insecto *Polypedilum vanderplanki* y las proteínas intrínsecamente desordenadas específicas de tardígrados (TDP, *Tardigrade-specific intrinsically Disordered Proteins*) que, como su nombre indica, se acumulan en tardígrados (Oliver et al., 2020; Mizrahi et al., 2010; Boothby et al., 2017; Voronina et al., 2020). Todas estas proteínas tienen en común la sobreexpresión de sus genes en respuesta a la deshidratación y/o a la desecación (Oliver et al., 2020; Boothby et al., 2017; Mizrahi et al., 2010; Liu et al., 2020; Voronina et al., 2020). Además, al eliminar o mutar los genes de algunas de estas proteínas, se observa una disminución de la capacidad de tolerar la desecación en distintos modelos de estudio (Boothby et al., 2017, Cocotl-Yañez et al., 2014). Aunque estas proteínas comparten su patrón de acumulación en desecación, se propone que son funcionalmente diferentes. Además, sus características de desorden estructural e hidrofobicidad son variadas (Tabla 1).

Tabla 1. Características funcionales y estructurales de las proteínas inducidas en respuesta a la desecación.

| <b>Proteínas</b>      | <b>Función</b>  | <b>Características</b>      | <b>Referencia</b>  |
|-----------------------|---|-----------------------------|--|
| LEA                   | Evitar la agregación de otras proteínas   | hidrofílicas y desordenadas | Oliver et al., 2020  |
| PvLIL                 | No conocida   | hidrofóbicas y plegadas     | Voronina et al., 2020  |
| TDPs                  | vitrificación   | hidrofílicas y desordenadas | Boothby et al., 2017   |
| ELIP                  | Foto-protección   | hidrofóbicas y plegadas     | Oliver et al., 2020; Yamagata & Bowler., 1997                    |
| sHSP                  | Estabilizar proteínas que perdieron su conformación tridimensional y evitar su agregación   | hidrofóbicas y plegadas     | Oliver et al., 2020; Holt et al., 2019                           |
| HSP                   | Degradación y tráfico intracelular. Además, estabilizan y pliegan proteínas que perdieron su conformación tridimensional y/o que experimentaron agregación. | hidrofóbicas y plegadas     | Mizrahi et al., 2010; Javid et al 2007                           |
| Enzimas antioxidantes | Desintoxicación de especies reactivas de oxígeno  | hidrofóbicas y plegadas     | Oliver et al., 2020; Sugio et al., 2000; Al-Madboly et al., 2020 |

### **II.III. Proteínas intrínsecamente desordenadas**

proteínas intrínsecamente desordenadas (IDPs, *Intrinsically Disordered Proteins*) son proteínas que carecen de una estructura tridimensional definida y estable (Wright & Dyson et al., 2015). Las IDPs se caracterizan por tener baja complejidad de secuencia, baja proporción de aminoácidos hidrofóbicos y alta proporción de aminoácidos cargados e hidrofílicos (Wright & Dyson, 2015). Sin embargo, auncon la ausencia parcial o total de una configuración tridimensional fija, las IDPs desempeñan funciones de gran importancia. Entre sus funciones encontramos las de reguladores de la transcripción y traducción, almacenamiento de moléculas pequeñas, ensamblaje de complejos multi-proteicos, tolerancia a la desecación, entre otras (Boothby et al., 2017 ; Dyson & Wright, 2005). Por ejemplo, el grupo de IDPs correspondiente a las proteínas LEA protege la actividad de enzimas *in vitro* al las cuales se les evaporó más de un 98% del agua usando un concentrador de vacío (Cuevas-Velazquez et al., 2016). Se han propuesto dos mecanismos por los que las proteínas LEA protegen y en ambos mecanismos el desorden estructural juega un papel importante. El primer mecanismo tiene el nombre de escudo molecular. En este mecanismo las proteínas LEA se organizan entre sí y rodean como grupo a sus proteínas cliente para estabilizar su forma nativa. En el segundo mecanismo propuesto, las proteínas LEA, gracias a su hidrofiliidad y desorden presentan un radio hidrodinámico grande, lo que les permite interactuar, cubrir y estabilizar gran parte de la proteína cliente de forma individual (Cuevas-Velazquez et al., 2014). También se ha planteado que la actividad protectora está relacionada con la ganancia de orden estructural y de adquirir un plegamiento en alfa-hélice, esto en respuesta al amontonamiento macromolecular y al déficit hídrico (Cuevas-Velazquez et al., 2016).

## II.IV. Proteínas abundantes en la embriogénesis tardía

Las proteínas abundantes en la embriogénesis tardía (LEA, *Late Embryogenesis Abundant*) fueron descritas por primera vez en semillas de algodón (*Gossypium hirsutum* L.) por Galau y colaboradores en 1986. Se identificó que estas proteínas presentan una gran abundancia en semillas secas, cuando estas adquieren tolerancia a la desecación (Galau et al., 1986). Actualmente se sabe que las proteínas LEA son ubicuas en el reino vegetal y que no solamente se expresan en semillas, sino que pueden acumularse en tejidos de plantas adultas y otros organismos durante condiciones de falta de agua, choques hiperosmóticos y temperaturas bajas (Toxopeus et al., 2014; Olvera-Carrillo et al., 2010; Hanin et al., 2011). Adicionalmente, se ha reportado que las proteínas LEA tienen la capacidad de conferir tolerancia a la desecación si se expresan heterológamente en organismos sensibles (Czernik et al., 2020).

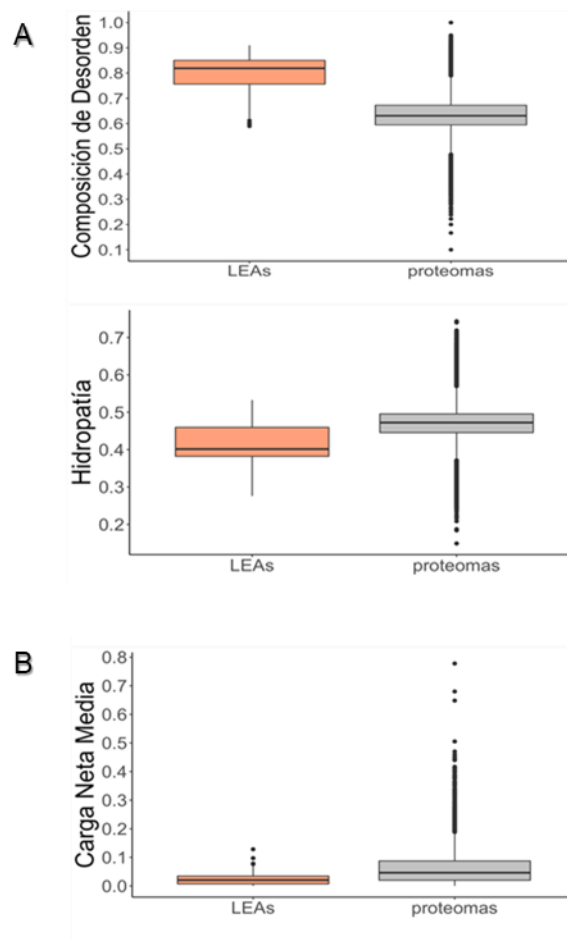
Las proteínas LEA asisten a los organismos durante condiciones de estrés mediante diversas funciones, entre ellas la de tipo-chaperona (Goyal et al., 2005), la de estabilizadores de membranas lipídicas, la de interacción con azúcares para facilitar el estado vítreo en el citoplasma, y las de protección y reparación del ADN (Oliver et al., 2020; Goyal et al., 2005).

A las proteínas LEA se les clasifica en familias o grupos por la presencia de motivos conservados en la secuencia de aminoácidos (Dure et al., 1989; Battaglia et al., 2008; Hundertmark & Hinch., 2008). Las diferentes familias de proteínas LEA poseen bajos niveles de conservación de secuencia de aminoácidos entre ellas, sin embargo, existen otras características que relacionan a todas las familias. Entre ellas está el carácter hidrofílico, el enriquecimiento de aminoácidos pequeños como Ala, Gly y Ser, y el bajo contenido de Cys y Trp (Baker et al., 1988; Dure, 1993). Además, mediante análisis estructurales se ha encontrado un alto grado de desorden de estas secuencias en solución acuosa, lo que las ha llevado a considerarse como proteínas intrínsecamente desordenadas (IDPs, *Intrinsically*

*Disordered Proteins*) (Rivera- Najera et al., 2014; Shih et al., 2010; Hughes & Graether, 2011; Battaglia et al., 2008).

## **II.V. Proteínas tipo-LEA**

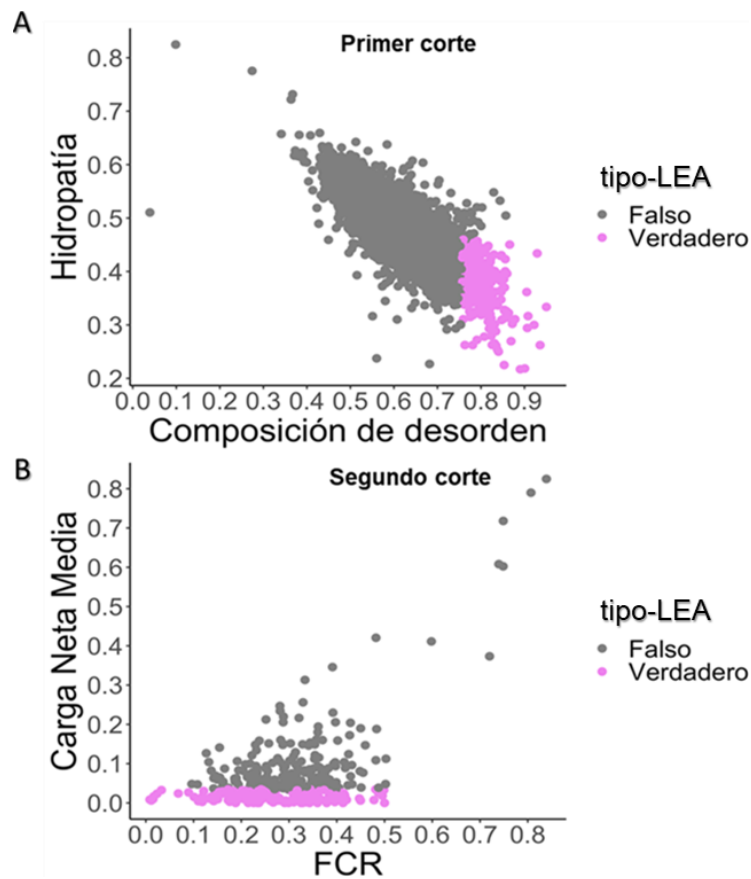
Como se mencionó antes en esta sección, las proteínas LEA cuentan con características fisicoquímicas compartidas entre ellas. En la tesis de Licenciatura de Michelle Lozano Ortega (2023), integrante de nuestro laboratorio, se presenta la caracterización de los parámetros fisicoquímicos que mejor diferencian a las proteínas LEA de plantas del resto de proteínas en sus respectivos proteomas. En primera instancia, se determinó que la composición de aminoácidos promotores del desorden estructural y la hidropatía son los mejores parámetros diferenciadores. Se realizó un análisis estadístico individual para determinar los intervalos de valores (de ambos parámetros) en los que quedaban incluidos el 75% de las proteínas LEA de plantas de sus respectivos proteomas. Como resultado de los análisis individuales, se definió un primer corte considerando la composición de aminoácidos promotores del desorden estructural mayor a 0.76 y un valor de hidropatía menor a 0.46. Este corte concentró a un nuevo grupo de proteínas que incluyó al 66.2 % de las proteínas LEA y solo al 3.7% de proteínas de los proteomas de plantas (Figura 1). Con el fin de diferenciar mejor a las proteínas LEA de ese nuevo grupo de proteínas del 3.7% de los proteomas, se caracterizaron otros parámetros fisicoquímicos en el grupo de proteínas que pasaron el primer corte. En la nueva caracterización, se encontró que la carga neta media es el mejor parámetro para diferenciar a las proteínas LEA en el nuevo grupo de proteínas. Al realizar un nuevo análisis estadístico se definió un segundo corte de carga neta media menor a 0.035 (Figura 2). El nuevo corte agrupó al 74.8% de las proteínas LEA que pasaron el primer corte, lo que corresponde al 49.6% de todas las proteínas LEA de plantas analizadas. Al emplear ambos cortes se separó a ese 49.6% de proteínas LEA del 98.6% de las proteínas del proteoma. A las proteínas que pasaron ambos cortes se les denominó “proteínas tipo-LEA”.



**Figura 1.** Distribución de los niveles de composición de aminoácidos promotores del desorden estructural, hidropatía y carga neta media característicos de las proteínas tipo- LEA. En color salmón se muestran las proteínas LEA de *A. thaliana*, *Oryza sativa* (arroz), *Triticum aestivum* (trigo), y *Zea mays* (maíz) mientras que en color gris se muestra el resto de las proteínas de los proteomas de las 4 especies. A) Análisis de composición de aminoácidos que promueven el desorden y la hidropatía de las proteínas LEA y los proteomas de las 4 especies de plantas. Los gráficos de caja de composición de desordene hidropatía se utilizaron para determinar el primer corte (composición de desorden  $> 0.76$ e hidropatía  $< 0.46$ ), eliminando al 25% de las proteínas LEA. B) Análisis de la carga neta media del grupo de proteínas LEA y los proteomas de las 4 especies de plantas resultantes al aplicar el primer corte. El gráfico de caja de carga neta media se utilizó para determinar el segundo corte (carga neta media  $< 0.035$ ). Para el segundo corte se eliminaron al 25% de las proteínas LEA restantes después de aplicar el primer corte. Modificado de la tesis delicenciatura de Michelle Lozano Ortega (2023).



A continuación se investigó si otros organismos tolerantes a la desecación tenían proteínas con las propiedades fisicoquímicas consideradas. Aplicando ambos cortes al proteoma del tardígrado tolerante *Hypsibius dujardini*, se observó que el 1.3% de las proteínas de su proteoma son tipo-LEA (Figura 2), lo cual era sugerente de que las proteínas tipo-LEA encontradas podrían tener una función tipo-chaperona durante condiciones de desecación.



**Figura 2. Determinación de proteínas tipo-LEA de *H. dujardini*.** A) Distribución de las proteínas del tardígrado según los valores de hidropatía y composición de aminoácidos promotores del desorden. B) Distribución de las proteínas del tardígrado según su carga neta media y fracción de residuos cargados (FCR, *Fraction of Charged Residues*). En color violeta se muestran las proteínas de tardígrado que caen dentro del intervalo que corresponde al primer y segundo corte, mientras que en color gris se muestra el resto de las proteínas del proteoma de *H. dujardini*, las cuales no cumplen con dicho criterio (modificado de Lozano Ortega, 2023).

## II.VI. *Hypsibius dujardini*, un animal (tardígrado) tolerante a la desecación

La anhidrobiosis se define como la capacidad de “revivir” de un estado ametabólico resultado de la desecación a nivel celular (Grzyb & Skłodowska., 2022). Los términos anhidrobiosis y tolerancia a la desecación usualmente se emplean como sinónimos (Grzyb & Skłodowska., 2022). Pocos son los animales anhidrobiontes conocidos y todos ellos son invertebrados. Entre los invertebrados más interesantes para el estudio de la anhidrobiosis están los tardígrados, debido a la cantidad de información genómica, transcriptómica y molecular que se ha obtenido recientemente (Yoshida & Tanaka, 2022). Los tardígrados son animales microscópicos que podemos encontrar en todos los ambientes posibles. Los tardígrados, además de tolerar la desecación tienen la capacidad de soportar condiciones extremas como temperaturas arriba del punto de ebullición del agua, temperaturas congelantes, presiones altas y bajas, y hasta sobrevivir en el espacio (Hibshman et al., 2020). Para tolerar algunos de los tipos de estrés antes mencionados, incluida la desecación, los tardígrados necesitan entrar a un estado quiescente llamado “estado de tonel” en donde se producen cambios estructurales y morfológicos que se sugiere disminuyen la pérdida de agua (Hibshman et al., 2020). Existen variaciones en la capacidad de algunos tardígrados para tolerar la desecación cuando esta se impone de forma abrupta. Los tardígrados que no pueden tolerar un cambio abrupto requieren de una pérdida gradual de agua para ser capaces de tolerar la desecación. *Hypsibius dujardini* es una especie de tardígrado que requiere de este pre-acondicionamiento, mientras que otros como *Ramazzottius varieornatus* no lo necesitan. Interesantemente, se ha observado que ambas especies utilizan efectores similares para adquirir tolerancia, pero difieren en la forma en que los expresan. *R. varieornatus* lo hace constitutivamente e *H. dujardini* los induce solo en respuesta a una pérdida de agua (Yoshida & Tanaka, 2022). Entre los mecanismos que utilizan estos organismos para tolerar la desecación se encuentra la expresión de las TDPs, proteínas que se sugiere contribuyen a la tolerancia vitrificando el citoplasma y evitando la desnaturalización de otras proteínas (Boothby et al., 2017).

## II.VII. *Cronobacter sakazakii*, una bacteria tolerante a la desecación

Al igual que los eucariontes, algunos procariontes tienen la capacidad de entrar y salir de un estado ametabólico para tolerar la desecación. Se conocen arqueas tolerantes a la desecación y también existen ejemplos de bacterias tanto Gram-positivas como Gram-negativas (Grzyb & Skłodowska., 2022). Un ejemplo de bacterias Gram-negativas con tolerancia a la desecación es *Cronobacter sakazakii*. *C. sakazakii* es un patógeno neonatal anaerobio facultativo que normalmente se encuentra en la fórmula infantil en polvo. A esta bacteria Gram-negativa se le conoce principalmente por causar septicemia y meningitis neonatal con una alta tasa de mortalidad (Gurtler & Beuchat., 2007). La fórmula infantil en polvo tiene una actividad del agua de entre 0.25 y 0.35, por lo que la mayoría de microorganismos no pueden sobrevivir en este ambiente seco. En cambio, la capacidad de tolerar la desecación de *C. sakazakii* es la que le permite ser un contaminante de importancia en este producto comercial. Mediante estudios de secuenciación de RNA se sugiere que la síntesis de trehalosa y otros osmoprotectores son los principales mecanismos que usa *C. sakazakii* para contender con los daños asociados a la desecación (Srikumar et al., 2019).

### III. Antecedentes

El paradigma secuencia-estructura-función en las proteínas establece que la secuencia de aminoácidos de una proteína determina su estructura tridimensional y la estructura determina su función (Babu, 2016). Las proteínas intrínsecamente desordenadas (*IDPs, Intrinsically Disordered Proteins*) desafían el paradigma ya que tienen una o más funciones, aún cuando carecen de una estructura tridimensional bien definida (Babu, 2016). Existen propiedades derivadas de la secuencia de aminoácidos que están asociadas a la falta de orden estructural de las IDPs. Entre ellas destacan una baja complejidad en la diversidad de aminoácidos de la secuencia, baja proporción de aminoácidos hidrofóbicos y alta proporción de aminoácidos cargados e hidrofílicos (Wright & Dyson, 2015). Dentro de las IDPs se encuentra un grupo de proteínas conocidas como proteínas abundantes en la embriogénesis tardía (LEA, *Late Embryogenesis Abundant*). Las proteínas LEA poseen una elevada hidrofiliidad (afinidad por el agua) y un alto contenido de aminoácidos pequeños como Ala, Gly y Ser (Battaglia et al., 2008). Las proteínas LEA se acumulan cuando las semillas se encuentran en un estado dormante, que es cuando adquieren tolerancia a la desecación. Además, las proteínas LEA se acumulan en tejidos vegetativos y reproductivos en respuesta a deshidratación (Battaglia et al., 2008). De forma similar, las proteínas LEA se acumulan en otros organismos en respuesta a los mismos estímulos que en las plantas (Voronina et al., 2020; Toxopeus et al., 2014). Múltiples estudios demuestran el papel de las proteínas LEA en los organismos que las acumulan durante condiciones limitantes de agua. Por ejemplo, al silenciar los genes que codifican para proteínas LEA en el organismo tolerante a la desecación *Artemia franciscana*, la resistencia a la pérdida de agua y temperaturas bajas se ve disminuida (Toxopeus et al., 2014). Sin embargo no todos los organismos tolerantes a la desecación tienen proteínas LEA o son escasas, por lo que posiblemente utilicen otro mecanismo de protección o cuenten con proteínas con propiedades fisicoquímicas similares a las de las proteínas LEA, pero con una estructura primaria totalmente distinta.

Algunos autores han hipotetizado que las propiedades fisicoquímicas de las proteínas LEA son esenciales para ejercer su función en condiciones limitantes de

agua. Garay-Arroyo y colaboradores definieron hace 23 años al grupo de las “hidrofilinas”, el cual contiene la mayoría de las proteínas LEA. El criterio que define a las hidrofilinas no toma en cuenta los motivos conservados de las proteínas LEA pero si sus propiedades fisicoquímicas. el criterio es un alto índice de hidrofiliidad (mayor a 1) y un alto contenido de Gly (mayor al 6%). Al analizar los proteomas de la bacteria *Escherichia coli* y de la levadura *Saccharomyces cerevisiae*, los autores encontraron una asociación de la expresión de las hidrofilinas de estos organismos con condiciones de estrés osmótico. Un análisis posterior realizado por Battaglia y colaboradores (Battaglia et al., 2008) muestra que el 92% de un total de 378 proteínas LEA encontradas en bases de datos son hidrofilinas. Además, se encontró que las proteínas LEA con patrones de expresión conocidos responden a condiciones de falta de agua. En otro análisis de carácter comparativo entre las propiedades fisicoquímicas de las proteínas LEA, hidrofilinas y un dominio de proteínas que responde a estrés hídrico, se concluyó que las hidrofilinas más que ser un grupo que incluye a todas las proteínas LEA, está relacionado con las dehidrinas (grupo LEA 2) y otras proteínas. Además, en ese mismo estudio se propone el papel en común que pueden tener las hidrofilinas, dehidrinas y el dominio de proteínas que responde a estrés hídrico, durante condiciones limitantes de agua (Jaspard & Hunault, 2014).

Como ya se mencionó, se hizo un análisis masivo para comparar los parámetros fisicoquímicos de las proteínas LEA y el resto de proteínas de proteomas de distintas especies de plantas (Lozano Ortega, 2023). En dicho trabajo se encontró que una alta composición de aminoácidos promotores del desorden ( $\geq 0.76$ ), una baja hidropatía (alta hidrofiliidad) ( $\leq 0.46$ ), y niveles bajos de carga neta media ( $\leq 0.035$ ) separaban hasta en un 98.6% a un grupo de proteínas LEA de plantas de otras proteínas del resto de sus proteomas. A las proteínas que cumplían con estas características se les nombró proteínas tipo-LEA. Lo anterior les permitió identificar proteínas tipo-LEA, y sugerir su posible función en condiciones limitantes de agua al evaluar el proteoma de organismos tolerantes a la desecación como el tardigrado *Hypsibius dujardini*. Adicionalmente a las propiedades fisicoquímicas, otra característica compartida entre las proteínas LEA es su sobreexpresión durante

condiciones de deshidratación y desecación (Toxopeus et al., 2014; Olvera-Carrillo et al., 2010; Hanin et al., 2011). Las proteínas LEA no son las únicas que se sobreexpresan bajo tales condiciones, por lo que la sobreexpresión solo nos habla de la posible participación de una proteína en una determinada condición. Sin embargo, al analizar conjuntamente los niveles de expresión y las propiedades fisicoquímicas particulares de las proteínas LEA, nos damos cuenta que existen proteínas parecidas a las LEA que se acumulan en respuesta a la falta de agua en otros organismos. Un ejemplo son las proteínas CAHS pertenecientes al grupo de las proteínas intrínsecamente desordenadas específicas de tardígrados (TDP, *Tardigrade-specific intrinsically Disordered Proteins*), las cuales se sobreexpresan en condiciones de desecación, son desordenadas, muestran actividad tipo-chaperona *in vitro* y aumentan la tolerancia a la desecación de las células donde las han expresado (Boothby et al., 2017).

En este trabajo proponemos que los estudios transcriptómicos en combinación con el análisis de las propiedades fisicoquímicas de las proteínas LEA son útiles para establecer un criterio que permita encontrar nuevas proteínas que actúen de forma similar a las proteínas LEA en otros organismos tolerantes a la desecación, como en el caso del tardígrado *Hypsibius dujardini* y de la bacteria *Cronobacter sakazakii*.

#### **IV. Hipótesis**

Los transcritos de proteínas con características fisicoquímicas semejantes a las proteínas LEA de organismos tolerantes a la desecación (*Hypsibius dujardini* y *Cronobacter sakazakii*) se acumulan durante condiciones de desecación.

#### **V. Objetivo general**

Determinar los niveles de expresión de transcritos de proteínas tipo-LEA en respuesta a condiciones de desecación en los organismos tolerantes *Hypsibius dujardini* y *Cronobacter sakazakii*.

#### **VI. Objetivos particulares**

- Buscar transcriptomas reportados en condiciones húmedas y condiciones de desecación de los organismos tolerantes a la desecación *Hypsibius dujardini* y *Cronobacter sakazakii*.
- Realizar el análisis de expresión diferencial de transcritos de los organismos tolerantes a la desecación *Hypsibius dujardini* y *Cronobacter sakazakii*.
- Determinar si existen genes tipo-LEA cuyos transcritos se sobreexpresan en respuesta a condiciones de desecación en los organismos tolerantes a la desecación *Hypsibius dujardini* y *Cronobacter sakazakii*.

## VII. Metodología

### VII.I. Datos

#### Lecturas de RNA-seq

Con el fin de determinar el tipo de transcritos acumulados en condiciones de desecación por organismos tolerantes realizamos una búsqueda bibliográfica. Se buscaron estudios que utilizaron RNA-seq en organismos tolerantes a la desecación. La búsqueda se realizó en la base de datos Scopus en Septiembre de 2020 utilizando los términos “RNA-seq”, “survive” y “desiccation”. De esta búsqueda se seleccionaron 3 estudios con datos depositados en el NCBI. Los organismos se seleccionaron tomando tres condiciones: tener al menos 2 o más réplicas biológicas para cada condición (hidratado y desecado), poseer un genoma anotado y tener un proteoma disponible (Tabla 2). Requerimos de estas condiciones para posteriormente relacionar las propiedades fisicoquímicas del proteoma con los patrones de expresión.

Tabla 2. Lecturas de RNA-seq utilizadas en este trabajo.

| Lecturas de RNA-seq                |              |                        | Genoma de referencia |
|------------------------------------|--------------|------------------------|----------------------|
| Organismo                          | Bioproject   | Referencia Bioproject  | Acceso de GenBank    |
| <i>Arabidopsis thaliana</i>        | PRAJNA314076 | Klepikova et al., 2016 | GCA_000001735.2      |
| <i>Cronobacter sakazakii</i> SP291 | PRJNA484789  | Srikumar et al., 2019  | GCA_000339015.1      |
| <i>Hypsibius dujardini</i>         | PRJNA369152  | Boothby et al., 2017   | GCA_002082055.1      |



### **Proteínas LEA de *Arabidopsis thaliana***

Para confirmar la acumulación de las proteínas LEA en respuesta a la desecación, buscamos las secuencias LEA descritas para *Arabidopsis thaliana*. Ya que no existe una base de datos oficial, se consideraron las 51 proteínas identificadas por Hundertmark & Hinch a como proteínas LEA (Hundertmark & Hinch a., 2008). Las 51 proteínas LEA son las que se muestran en la Figura 2.

### **Proteomas de referencia**

Para seleccionar a las proteínas tipo-LEA de los organismos tolerantes a la desecación, se calcularon los parámetros fisicoquímicos de los proteomas completos de las especies de interés depositados en Uniprot (Tabla 3). Los proteomas tienen identificadores en común con las anotaciones del genoma escogido.

Tabla 3. Proteomas de referencia tomados de Uniprot

| <b>Organismo</b>   | <b>Identificador Uniprot</b> | <b>Proteínas totales</b> |
|--|------------------------------|--------------------------|
| <i>Arabidopsis thaliana</i> (Mouse-ear cress)<br>(Strain: cv. Columbia)                    | UP000006548                  | 39,337                   |
| <i>Hypsibius dujardini</i> (Water bear)<br>( <i>Macrobiotus dujardini</i> ) (Strain: Z151) | UP000192578                  | 14,867                   |
| <i>Cronobacter sakazakii</i> SP291   | UP000011279                  | 4178                     |

### **VII.II. Procesamiento de las lecturas de RNA-seq**

Los datos crudos de secuenciación se limpiaron, mapearon y anotaron al genoma para realizar posteriormente el análisis de expresión diferencial de genes de los organismos tolerantes a la desecación.

### **Limpieza de lecturas**

Se descargaron los archivos FASTQ de cada artículo seleccionado utilizando **SRA Toolkit v2.11.0**. Los archivos de cada estudio se limpiaron de manera independiente con **fastp v0.20.0**. Se removieron adaptadores, lecturas menores a 33 pb y las que tuvieran una calidad menor a Q 20 en el 40% o más de su secuencia.

### **Mapeo de lecturas y cuenta de los alineamientos**

Las lecturas depuradas se mapearon con **Bowtie2 v2.4.4** al genoma de referencia de cada organismo. El mapeo se realizó de manera local con los parámetros default de la herramienta y se conservaron las lecturas únicas mapeadas en cada mapeo.

Después se intersecaron las anotaciones de cada genoma con las posiciones de las lecturas mapeadas con **featureCounts v2.0.0** desviándonos de las opciones default en el caso de los datos pareados de *H. dujardini*, donde se usaron fragmentos en lugar de lecturas para generar los reportes. También se modificó la opción de “strandness” a stranded en *C. sakazakii* debido a las características de los datos.

### **Análisis de expresión diferencial**

Se utilizó DESeq (Versión de Galaxy 2.11.40.6+galaxy2 ) en la plataforma de red Galaxy (usegalaxy.org) para analizar las lecturas mapeadas a los genes de cada organismo con las opciones default de la plataforma exceptuando la opción para filtrar “outliers”. Se llevo a cabo un test de Wald y se tomaron como diferencialmente expresados los genes con un valor de  $p < 0.05$ .

La paquetería ggplot2 de **RStudio v1.3.959** se utilizó para generar los gráficos de toda esta sección.

### **VII.III. Análisis bioinformático de las secuencias de proteínas**

Se realizó el cálculo de los valores de los parámetros fisicoquímicos de composición de aminoácidos promotores del desorden, hidropatía y carga neta media de los proteomas seleccionados en la sección VI.I. Datos/ proteínas de referencia.

### **Limpieza de secuencias**

Se utilizó **Protfasta v0.1.8** para remover las secuencias con residuos ambiguos (e.g., “X”, “Z”, “N”, etc) de los proteomas mencionados en la Tabla 4.

### **Cálculo de propiedades fisicoquímicas**

Después de convertir los archivos multifasta a archivos individuales usando **NextGen Sequence Workbench v3.2.3** se utilizó la versión local de CIDER (Holehouse et al., 2017) **localCIDER v0.1.18** para calcular los parámetros fisicoquímicos utilizados para definir al grupo de proteínas tipo-LEA naturales de los proteomas de interés. Los siguientes parámetros fisicoquímicos son los que utilizamos en este trabajo:

- Hidropatía de Uversky: Rango que va del 0 al 1 donde valores cercanos a 0 indican una mayor proporción de residuos hidrofílicos en la secuencia (hidrofilicidad) y valores cercanos a 1 denotan alta proporción de aminoácidos hidrofóbicos (hidrofobicidad).
- composición de aminoácidos promotores del desorden: Muestra la fracción de residuos promotores del desorden en una secuencia. Los aminoácidos considerados que promueven el desorden son: Thr, Ala, Gly, Arg, Asp, His, Gln, Lys, Ser, Glu, y Pro.
- Carga neta media: Rango de valores del 0 al 1 que representan la carga neta promedio absoluta de la secuencia evaluada.

### **VII.IV. Determinación de la expresión de proteínas tipo-LEA**

Con los parámetros calculados, seleccionamos a las proteínas con alta composición de aminoácidos promotores del desorden ( $\geq 0.76$ ), baja hidropatía ( $\leq 0.46$ ) y baja carga neta media ( $\leq 0.035$ ) de los proteomas de los organismos tolerantes a la desecación y se identificaron en el análisis de expresión diferencial respectivo.

## VIII. Resultados

### **Análisis de expresión diferencial de transcritos que codifican para proteínas LEA en semillas de *Arabidopsis thaliana***

Debido a que en esta tesis los patrones de expresión de las proteínas LEA en condiciones limitantes de contenido de agua son la base para el planteamiento de nuestra hipótesis, buscamos confirmar su acumulación (enriquecimiento y abundancia) en semillas secas de *A. thaliana*. Con ese fin, realizamos un análisis de expresión diferencial de transcritos con datos de RNA-seq de semillas hidratadas (SD1 y SD3 en Bioproject; semillas de silicua) y secas (SDd en Bioproject; semillas dormantes) (Klepikova et al., 2016). El número promedio de lecturas totales de las semillas hidratadas y secas fue de 33.9 y 27.1 millones, respectivamente, de las cuales en promedio 60.86% y 84% mapearon de manera única al genoma de referencia (Tabla 4). Además, una media del 96.46% de esas lecturas se asignaron dentro de los 38,308 genes anotados de *Arabidopsis*.

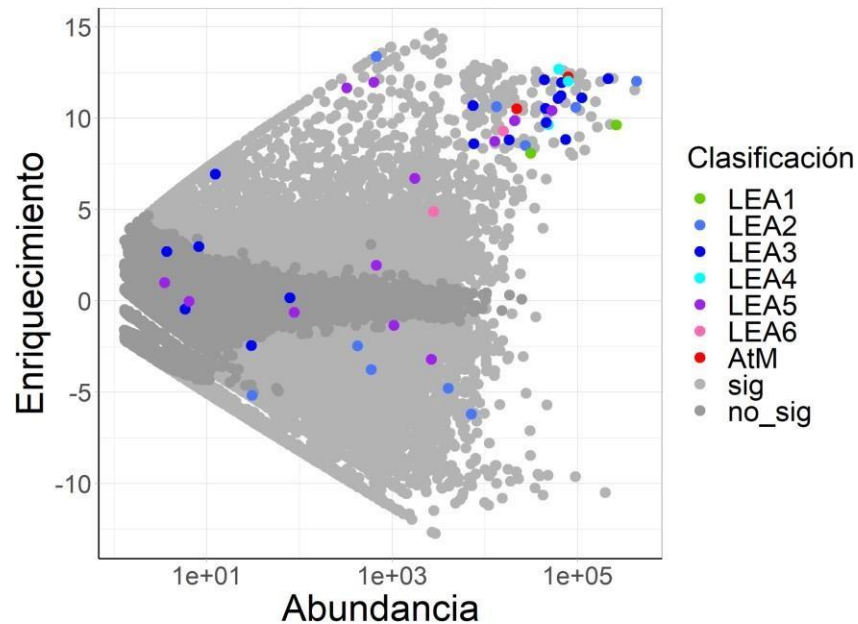
Tabla 4. Resumen de los datos de RNA-seq de semillas de *Arabidopsis thaliana*.

| Organismo                   | Muestras        | Número total de lecturas | Porcentaje de lecturas mapeadas de manera única | Número de lecturas mapeadas a genes |
|-----------------------------|-----------------|--------------------------|---|-------------------------------------|
| <i>Arabidopsis thaliana</i> | SD1 (réplica 1) | 33,496,354               | 63.38%  | 19,996,830                          |
|                             | SD1 (réplica 2) | 30,273,141               | 65.60%  | 19,037,986                          |
|                             | SD3 (réplica 3) | 32,319,064               | 66.32%  | 20,228,357                          |
|                             | SD3 (réplica 4) | 39,525,451               | 48.15%  | 17,812,411                          |
|                             | SDd (réplica 1) | 26,702,121               | 84.45%  | 21,376,899                          |
|                             | SDd (réplica 2) | 27,508,132               | 83.56%  | 22,061,679                          |

SD1 y SD3: semillas de silicua (hidratadas). SDd: semillas dormantes (secas)

El análisis de expresión diferencial mostró a un 40.98% de los genes diferencialmente expresados ( $p < 0.05$ ) en las semillas secas de *Arabidopsis*.

Además, la mayoría de proteínas LEA se encontraron acumuladas en semilla seca (Figura 3).

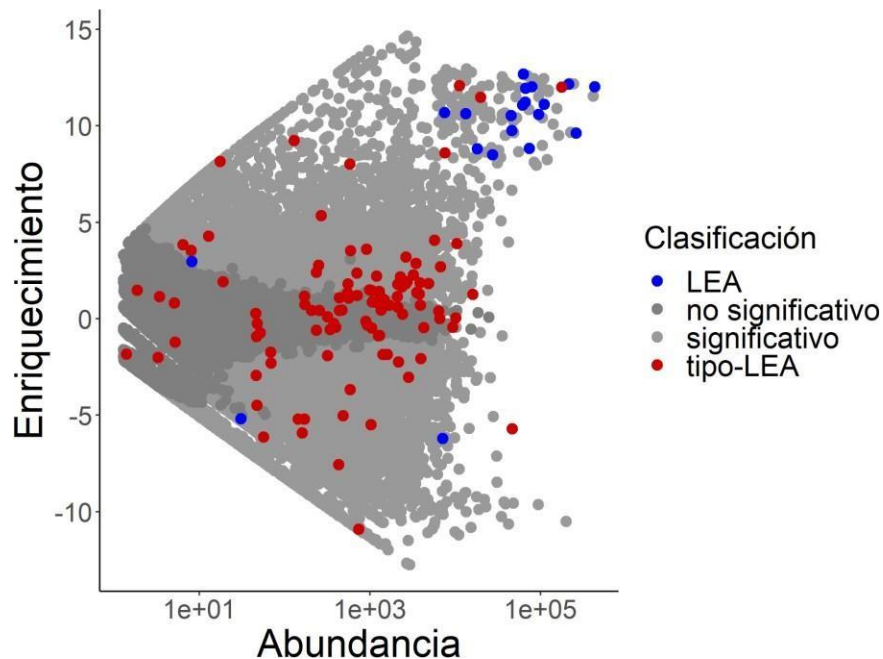


**Figura 3. Análisis de expresión diferencial de genes de proteínas LEA en semillas secas e hidratadas de *Arabidopsis thaliana*.** Cada punto representa un gen. Los ejes muestran el enriquecimiento ( $\log_2FC$ ) y la abundancia (media de cuentas normalizadas) de genes expresados en las semillas. El código de colores representa a los genes de proteínas LEA separados por grupos, genes no significativos para expresión diferencial en gris oscuro y significativos en gris claro ( $p < 0.05$ ). Las proteínas LEA se designaron como AtM al no tener un grupo asignado. 2 proteínas LEA (At2g33690 y At3g53770) de las 51 no se muestran por ausencia de resultados en el análisis.

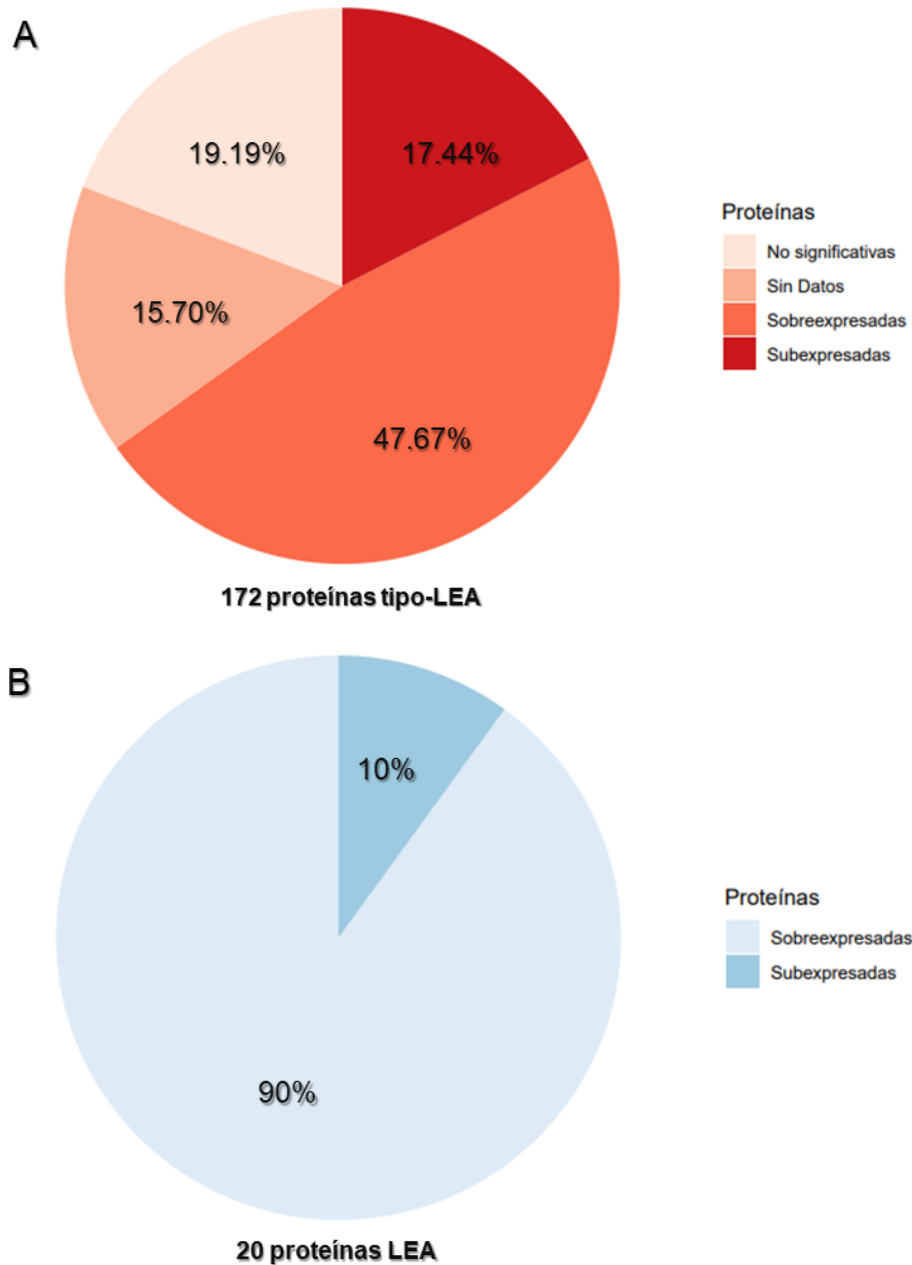
### **Análisis de expresión diferencial de transcritos proteínas tipo-LEA en semillas de *Arabidopsis thaliana***

Con los patrones de acumulación confirmados, decidimos utilizar el criterio de selección de las proteínas tipo-LEA generado en nuestro laboratorio para encontrar proteínas con características fisicoquímicas semejantes a las de las proteínas LEA y posteriormente determinar sus patrones de expresión transcripcional en semillas secas de *A. thaliana*. Para ello, aplicamos los cortes de composición de aminoácidos promotores del desorden ( $\geq 0.76$ ), hidropatía ( $\leq 0.46$ ) y carga neta media ( $\leq 0.035$ )

al proteoma de *Arabidopsis* e identificamos a los genes de las proteínas resultantes (tipo-LEA y LEA que pasaron los cortes) en el análisis de expresión diferencial de *Arabidopsis* mencionado anteriormente (Figura 3). En el proteoma de *Arabidopsis*, 192 proteínas pasan los cortes, de las cuales 172 son proteínas tipo-LEA y 20 corresponden a proteínas LEA. De esas proteínas el 47.67% y el 90% presentaron genes diferencialmente sobreexpresados en semillas secas, respectivamente (Figura 5). La mayor parte de las proteínas LEA se acumularon en semilla seca (Figura 4, puntos azules) y las proteínas tipo-LEA difirieron en sus patrones de expresión. (Figura 4, puntos rojos).



**Figura 4. Disposición de las proteínas tipo-LEA y LEA en el análisis de expresión diferencial de genes en semillas secas e hidratadas de *Arabidopsis thaliana*.** Análisis de expresión diferencial donde cada punto representa un gen y los ejes muestran el enriquecimiento ( $\log_2FC$ ) contra la abundancia (media de cuentas normalizadas) de genes expresados en semillas. Genes con valores positivos de enriquecimiento indican mayor expresión en semillas secas y genes con valores negativos indican mayor expresión en semillas hidratadas. El código de colores representa a los genes de proteínas LEA en azul, proteínas tipo-LEA en rojo, genes no significativos para expresión diferencial en gris oscuro y significativos en gris claro ( $p < 0.05$ ). Las proteínas que se muestran corresponden a las proteínas LEA que pasan cortes de composición de aminoácidos promotores del desorden, hidropatía y carga neta media.



**Figura 5. Clasificación de las proteínas tipo-LEA y LEA en función de su expresión diferencial de genes en semillas secas e hidratadas de *Arabidopsis thaliana*.** (A) Gráfico circular que indica la clasificación de proteínas tipo-LEA en función de su expresión en el análisis de expresión diferencial.

(B) Gráfico circular que indica la clasificación de proteínas LEA en función de su expresión en el análisis de expresión diferencial. Las proteínas LEA que se muestran corresponden a las proteínas LEA que pasan cortes de composición de aminoácidos promotores del desorden, hidropatía y carga neta media.

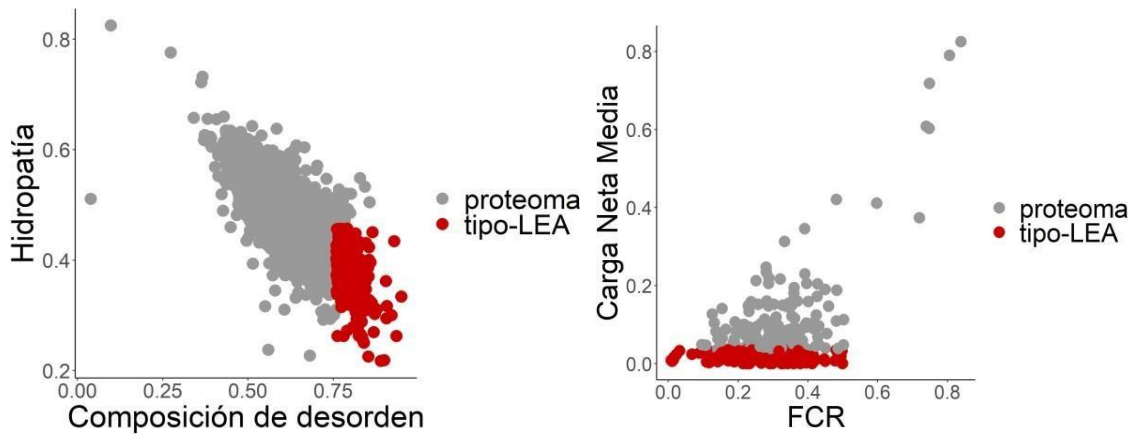
## **Análisis de expresión diferencial de transcritos que codifican para proteínas tipo-LEA en organismos tolerantes a la desecación**

Además de las semillas, existen otros organismos con la capacidad de tolerar la desecación como el tardígrado *Hypsibius dujardini* y la bacteria *Cronobacter sakazakii*. *Hypsibius dujardini* cuenta con muy pocos genes que codifican para proteínas LEA. *Cronobacter sakazakii* carece de genes que codifican para proteínas LEA. Este hecho nos llevó a plantear la hipótesis de que existen proteínas tipo-LEA que se sobreexpresan en condiciones de desecación en esos organismos. Buscando contestar esa pregunta realizamos un análisis de expresión diferencial con datos de RNA-seq de muestras hidratadas (HDH; muestra de *Hypsibius dujardini* hidratada) y secas (HDD; muestra de *Hypsibius dujardini* desecada) del tardígrado, y de muestras hidratadas (ESP; muestras en fase estacionaria temprana) y secas (DMS; muestras desecadas) de la bacteria (Srikumar et al., 2019; Boothby et al., 2017). También determinamos las proteínas tipo-LEA presentes en el proteoma del tardígrado (Figura 6) y la bacteria (Figura 7). Por último identificamos en el análisis de expresión diferencial a los genes que codifican para las proteínas tipo-LEA presentes en su proteoma (Figuras 8 y 10).

### **Búsqueda de proteínas tipo-LEA en el tardígrado *Hypsibius dujardini***

Michelle Lozano Ortega en su tesis de licenciatura estableció el criterio para definir a las proteínas tipo-LEA. Aplicó dicho criterio para determinar la existencia de proteínas tipo-LEA en el proteoma de *H. dujardini*. Como resultado encontró 190 proteínas tipo-LEA presentes en el proteoma del tardígrado, el 1.3% del total de proteínas (Figura 6).

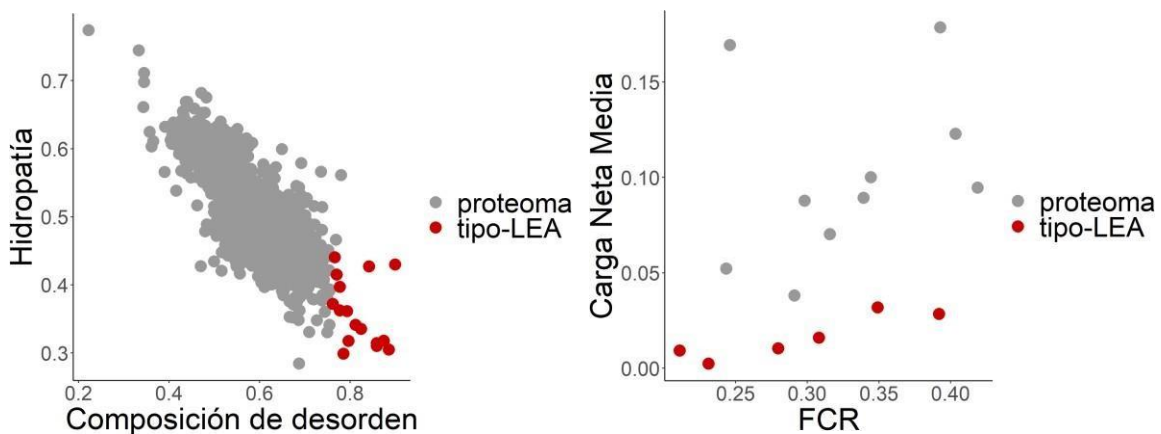




**Figura 6. Proteínas tipo-LEA de *Hypsibius dujardini*.** En color rojo se muestran las proteínas de tardígrado que caen dentro del intervalo que corresponde al primer y segundo corte, mientras que en color gris se muestra el resto de las proteínas del proteoma de *H. dujardini*, las cuales no cumplen con dicho criterio.

### Búsqueda de proteínas tipo-LEA en la bacteria *Cronobacter sakazakii*

En esta tesis buscamos determinar si existen proteínas tipo-LEA en otro organismo tolerante a la desecación, *Cronobacter sakazakii*. Aplicamos el criterio que define a las proteínas tipo-LEA al proteoma completo de la bacteria. Como resultado, encontramos 6 proteínas tipo-LEA presentes en el proteoma, el 0.15% del total de proteínas (Figura 7).



**Figura 7. Proteínas tipo-LEA de *Cronobacter sakazakii*.** En color rojo se muestran las proteínas de la bacteria que caen dentro del intervalo que corresponde al primer y segundo corte, mientras que en color gris se muestra el resto de las proteínas del proteoma de *C. sakazakii*, las cuales no cumplen con dicho criterio.

## Expresión de genes tipo-LEA según el análisis de expresión diferencial en *Hypsibius dujardini*

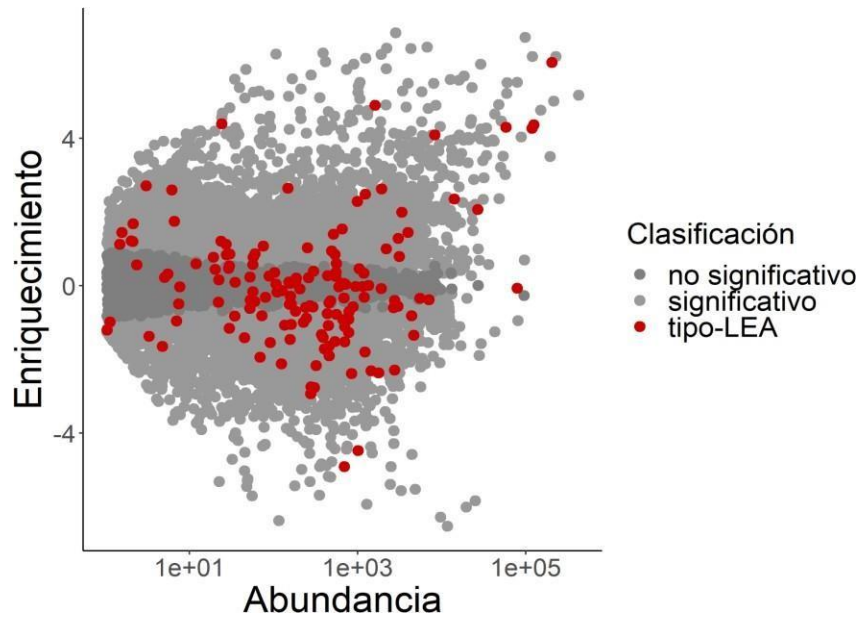
Para el tardígrado, el promedio de lecturas totales de las muestras desecadas e hidratadas fue de 20.5 y 34.5 millones respectivamente, de las cuales en promedio 44.60% y 67.89% mapearon de manera única al genoma de referencia (Tabla 5). Además, una media del 84.85% de esas lecturas se asignaron dentro de los 20,090 genes anotados de *H. dujardini*.

Tabla 5. Resumen de los datos de RNA-seq de *Hypsibius dujardini*.

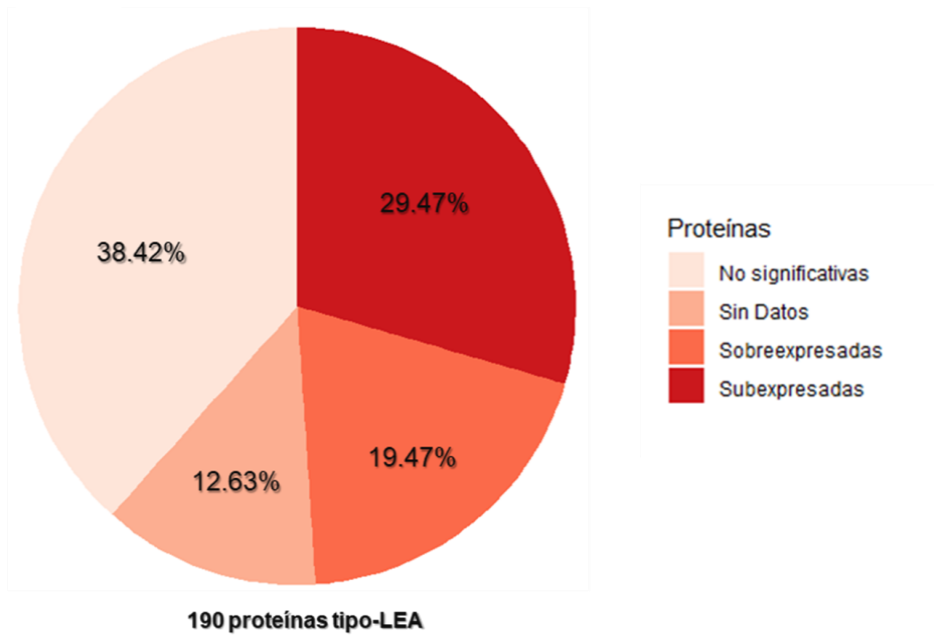
| Organismo                  | Muestras         | Número total de lecturas | Lecturas mapeadas de manera única | Número de lecturas mapeadas a genes |
|----------------------------|------------------|--------------------------|-----------------------------------|-------------------------------------|
| <i>Hypsibius dujardini</i> | HDHA (réplica 1) | 39,913,990               | 68.66%                            | 22,343,620                          |
|                            | HDHB (réplica 2) | 27,535,651               | 66.80%                            | 14,950,986                          |
|                            | HDHC (réplica 3) | 36,050,340               | 68.21%                            | 20,228,674                          |
|                            | HDHA (réplica 1) | 14,355,533               | 27.53%                            | 3,059,452                           |
|                            | HDDB (réplica 2) | 12,170,672               | 44.54%                            | 4,310,366                           |
|                            | HDCC (réplica 3) | 34,986,837               | 61.74%                            | 17,534,020                          |

HDH: muestra de *Hypsibius dujardini* hidratada. HDD: muestra de *Hypsibius dujardini* desecada.

Ya realizado el análisis de expresión diferencial de muestras desecadas e hidratadas de *Hypsibius dujardini*, identificamos en el a los genes tipo-LEA (genes que codifican para una proteína tipo-LEA). En el proteoma de *H. dujardini* encontramos 190 proteínas tipo-LEA (Figura 6), de las cuales el 19.5% presentaron genes diferencialmente sobreexpresados en condiciones de desecación (Figura 9). En conjunto, las proteínas tipo-LEA del tardígrado mostraron patrones de expresión variados (Figura 8, puntos rojos).



**Figura 8. Expresión diferencial de genes de proteínas tipo-LEA en muestras desecadas e hidratadas de *Hypsibius dujardini*.** Análisis de expresión diferencial donde cada punto representa un gen y los ejes muestran el enriquecimiento ( $\log_2FC$ ) contra la abundancia (media de cuentas normalizadas) de genes expresados en las muestras. Genes con valores positivos de enriquecimiento indican expresión diferencial en condiciones de desecación y genes con valores negativos indican expresión diferencial en condiciones húmedas. El código de colores representa a los genes de proteínas tipo-LEA en rojo, genes no significativos para expresión diferencial en gris oscuro y significativos en gris claro ( $p < 0.05$ ).



**Figura 9. Clasificación de las proteínas tipo-LEA en el análisis de expresión diferencial de genes en muestras desecadas e hidratadas de *Hypsibius dujardini*.** Gráfico circular que indica la clasificación de proteínas tipo-LEA en función de su expresión en el análisis de expresión diferencial.

## Expresión de genes tipo-LEA según el análisis de expresión diferencial en *Cronobacter sakazakii*

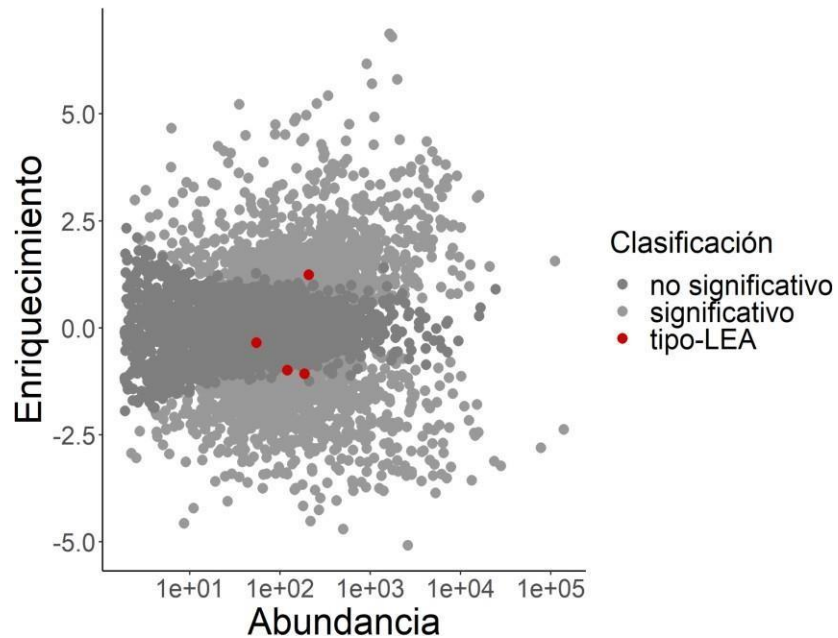
En la bacteria tolerante, las lecturas totales promedio de las muestras desecadas e hidratadas fueron de 27.8 y 24.8 millones respectivamente, de las cuales en promedio 8.41% y 19.40% mapearon de manera única al genoma de referencia (Tabla 6). Además, una media del 71.07% de esas lecturas se asignaron dentro de los 4312 genes anotados de *C. sakazakii*.

Tabla 6. Resumen de los datos de RNA-seq de *Cronobacter sakazakii*.

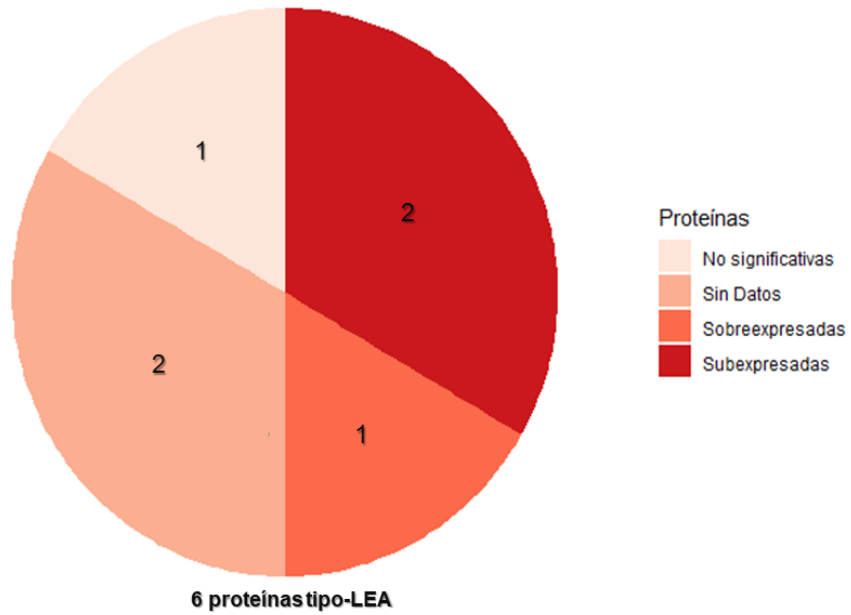
| Organismo                    | Muestras        | Número total de lecturas | Lecturas mapeadas de manera única | Número de lecturas mapeadas a genes |
|------------------------------|-----------------|--------------------------|-----------------------------------|-------------------------------------|
| <i>Cronobacter sakazakii</i> | ESP (réplica 1) | 24,787,002               | 22.97%                            | 3,487,347                           |
|                              | ESP (réplica 2) | 22,836,639               | 11.90%                            | 1,896,714                           |
|                              | ESP (réplica 3) | 26,902,127               | 23.34%                            | 3,994,100                           |
|                              | DMS (réplica 1) | 34,077,290               | 8.67%                             | 2,038,399                           |
|                              | DMS (réplica 2) | 24,264,019               | 8.59%                             | 1,454,104                           |
|                              | DMS (réplica 3) | 25,168,130               | 7.96%                             | 1,319,262                           |

ESP: muestras de *Cronobacter sakazakii* en fase estacionaria temprana. DMS: muestras de *Cronobacter sakazakii* desecadas.

Con el análisis de expresión diferencial de muestras desecadas e hidratadas de *Cronobacter sakazakii*, identificamos a los genes tipo-LEA. Existen 6 proteínas tipo-LEA presentes en el proteoma de *Cronobacter sakazakii* (Figura 7). Dentro del análisis de expresión diferencial de *Cronobacter* encontramos 4 proteínas tipo-LEA, de las cuales 1 presentó genes diferencialmente sobreexpresados en desecación (Figura 11). Las 4 proteínas mostraron patrones de expresión variados (Figura 10, puntos rojos).



**Figura 10. Análisis de expresión diferencial de genes de proteínas tipo-LEA en muestras desecadas e hidratadas de *Cronobacter sakazakii*.** Análisis de expresión diferencial en donde cada punto representa un gen y los ejes muestran el enriquecimiento ( $\log_2FC$ ) contra la abundancia (media de cuentas normalizadas) de genes expresados en las muestras. Genes con valores positivos de enriquecimiento indican expresión diferencial en condiciones de desecación y genes con valores negativos indican expresión diferencial en condiciones húmedas. El código de colores representa a los genes de proteínas tipo-LEA en rojo, genes no significativos para expresión diferencial en gris oscuro y significativos en gris claro ( $p < 0.05$ ).



**Figura 11. Clasificación de las proteínas tipo-LEA en el análisis de expresión diferencial de genes en muestras desecadas e hidratadas de Cronobacter sakazakii.** Gráfico circular que indica la clasificación de proteínas tipo-LEA en función de su expresión en el análisis de expresión diferencial.

## **IX. Discusión**

La mayoría de las proteínas LEA son hidrofílicas y carecen de una estructura tridimensional estable. Además de sus propiedades fisicoquímicas particulares, las proteínas LEA se caracterizan por acumularse principalmente en semillas en condiciones de desecación (Battaglia et al., 2008). Las plantas no son las únicas con este comportamiento, pues existen otros organismos tolerantes a la desecación que expresan proteínas LEA (Toxopeus et al., 2014). No obstante, existen otros organismos que toleran a la desecación con poca o nula presencia de proteínas LEA en su proteoma. Con base en ello, en este proyecto nos planteamos la existencia de proteínas con características fisicoquímicas similares a las proteínas LEA acumulándose en esos organismos tolerantes en respuesta a la desecación.

En esta tesis confirmamos con el análisis de expresión diferencial de genes en semillas de *Arabidopsis* la acumulación en desecación que presentan los transcritos que codifican para proteínas LEA. La mayoría de proteínas LEA se acumulan en semillas secas, no obstante, algunas proteínas LEA de los grupos LEA 2, LEA 3 y LEA 5 sesubexpresan y/o se expresan de forma similar en semillas húmedas y secas. Estos resultados se relacionan con los observados al usar RT-PCR cuantitativa en el análisis de expresión de semillas de *Arabidopsis* (Hundertmark & Hinch., 2008). En el mismo artículo se observa como las proteínas LEA pertenecientes a los grupos LEA 2, LEA 3 y LEA 5 se sobreexpresan en respuesta a estímulos diferentes a la desecación. La razón de la variabilidad de los patrones de expresión puede que se deba precisamente al papel que pueden tener las proteínas LEA en respuesta a distintos estímulos ambientales.

Particularmente al definir al grupo de proteínas LEA con alto desorden, baja hidropatía y baja carga neta media observamos que casi en su totalidad muestran patrones de acumulación en semillas secas de *Arabidopsis*. La acumulación de este grupo de proteínas LEA habla de lo relevantes que pueden ser las propiedades fisicoquímicas para desempeñar una función protectora, similar a lo que sucede con



otros fenómenos relacionados a las IDPs (Theillet et al., 2014). En semillas de *Arabidopsis*, existen más genes de proteínas tipo-LEA con una expresión alta en condiciones secas. Dentro del grupo de proteínas tipo-LEA existe variabilidad en los patrones de expresión entre condiciones secas y húmedas, pero podemos observar una mayoría de proteínas tipo-LEA con abundancia y sobreexpresión de sus transcritos en semilla seca. Esto resalta la importancia de las proteínas con características tipo-LEA durante la desecación de las semillas. También podemos afirmar que las proteínas LEA no son las únicas proteínas con alto desorden, hidrofiliidad y baja carga neta media con transcritos acumulados en semillas maduras. Además, podemos sugerir la posible función tipo-chaperona de estas proteínas durante condiciones de desecación por la similitud en sus propiedades fisicoquímicas con las de las LEA. Por otro lado, existe también otro grupo de proteínas tipo-LEA que o se expresan constitutivamente o se subexpresan. Para las proteínas tipo-LEA de este grupo posiblemente sucede lo que a algunas proteínas LEA, las cuales se ven inducidas en otros estadios del desarrollo de las plantas y/o por algún otra condición ambiental distinta a la falta de agua (e.g., elevadas temperaturas) (Hundertmark & Hinch, 2008).

Con el análisis de expresión diferencial de los organismos tolerantes a la desecación *Hypsibius dujardini* y *Cronobacter sakazakii* constatamos que algunos transcritos que codifican para proteínas tipo-LEA se acumulan en desecación, aunque muchos otros no muestran expresión diferencial o tienen acumulación significativamente menor en ese estado. Para las proteínas tipo-LEA acumuladas en desecación podemos sugerir, al igual que en *Arabidopsis*, la posible función tipo-chaperona. En el tardígrado, dentro de las proteínas acumuladas en desecación existen algunas con un mayor enriquecimiento que otras. Este patrón de enriquecimiento hace un símil a lo sucedido con las proteínas LEA en semilla seca respecto a las proteínas tipo-LEA. Probablemente, las proteínas tipo-LEA de tardígrado con mayor enriquecimiento sean las que tienen un comportamiento funcional más apegado al de las proteínas LEA. También debemos tener en cuenta que a diferencia de *Arabidopsis* el tardígrado no presenta acumulación preferencial de las proteínas

tipo-LEA en desecación y puede indicar un funcionamiento no tan apegado a el de las LEA de *Arabidopsis*. Distinto a lo que sucede en *Arabidopsis*, la mayoría de proteínas tipo-LEA del tardígrado se expresan constitutivamente en ambas condiciones. Al expresar constitutivamente proteínas tipo-LEA esperaríamos que el tardígrado respondiera rápidamente a condiciones de desecación. Por el contrario, se ha observado que el tardígrado *Hypsibius dujardini* requiere de un pre-acondicionamiento (deshidratación lenta) para tolerar la desecación (Boothby et al., 2017). En concordancia a estas necesidades del tardígrado, potencialmente las proteínas tipo-LEA que se acumulan en desecación son las principales responsables de la tolerancia.

Contrastando con el número de proteínas tipo-LEA de *A. thaliana* e *H. dujardini*, la bacteria *C. sakazakii*, solo presenta 4 proteínas tipo-LEA en su proteoma. La baja presencia de las proteínas tipo-LEA en la bacteria podría indicar el uso de otros mecanismos de protección contra la desecación. Por ejemplo, la acumulación del disacárido trehalosa en las células (Breeuwer et al., 2003). La trehalosa evita la agregación de proteínas y protege del daño a las membranas en desecación (Tapia et al., 2015). Aunado a esto, al interrumpir la síntesis de trehalosa, se ha observado una disminución en la supervivencia de *C. sakazakii* a la desecación, confirmando la importancia del disacárido en la tolerancia (Srikumar et al., 2019). No obstante, la capacidad de sobrevivir a la desecación no se elimina completamente. Es posible que, además de la trehalosa, existan otras moléculas (azúcares o proteínas) que contribuyan a la tolerancia.

Por último, en los tardígrados la trehalosa se encuentra ausente o en muy bajas cantidades (Boothby et al., 2017). En la bacteria, la trehalosa se acumula en desecación (Breeuwer et al., 2003). El tardígrado presenta un gran número de proteínas tipo-LEA, mientras que en la bacteria son casi nulas. La discrepancia en la cantidad de proteínas tipo-LEA entre el tardígrado y la bacteria, se vuelve un ejemplo perfecto de la hipótesis que menciona que los organismos tolerantes a la desecación requieren de un conjunto de proteínas con propiedades fisicoquímicas

distintas dependiendo de la presencia o ausencia de trehalosa (Janis et al., 2018). Por ejemplo, las proteínas específicas de tardígrado CAHS protegen de la desecación a enzimas *in vitro* y confieren tolerancia a la desecación en ausencia de trehalosa a *Saccharomyces cerevisiae* al sobreexpresarle CAHS (Boothby et al., 2017). Esto sugiere que las proteínas CAHS pueden ser un mecanismo que sustituye a la trehalosa para tolerar la desecación en los tardígrados. Nuestros resultados sugieren que propiedades como el desorden, la hidrofobicidad y la carga neutra en las proteínas inducidas en desecación pueden ser relevantes para conferir tolerancia en los organismos que no producen trehalosa o tienen poca presencia de esta.

## **X. Conclusiones**

Realizamos un análisis bioinformático a nivel proteómico de las propiedades fisicoquímicas correspondientes a composición de aminoácidos promotores del desorden, hidrofiliidad y carga neta media en conjunto con el análisis de expresión diferencial de genes en *Hypsibius dujardini*, *Cronobacter sakazakii* y semillas de *Arabidopsis thaliana* en desecación. Mediante el traslape de las propiedades fisicoquímicas y los patrones de expresión génica confirmamos la existencia de transcritos de proteínas tipo-LEA sobreexpresados por los organismos mencionados en condiciones de desecación. La proporción de proteínas tipo-LEA sobreexpresadas en desecación por ambos organismos difiere entre si. Esto sugiere diferencia en cuanto a los mecanismos para contender con la falta de agua. En semillas de *Arabidopsis* constatamos la sobreexpresión de los transcritos que codifican para proteínas LEA en semillas de *Arabidopsis* reportada en la literatura. Las proteínas tipo-LEA con transcritos sobreexpresados por los organismos tolerantes en desecación son candidatos para desempeñar una actividad protectora tipo-chaperona, similar a la función de las proteínas LEA canónicas. Esta tesis aporta información sobre cómo los parámetros fisicoquímicos de las proteínas intrínsecamente desordenadas y los patrones de expresión génica en respuesta a estrés, en conjunto, pueden utilizarse para encontrar proteínas con posibles funciones relevantes en un contexto determinado.

## **XI. Perspectivas**

Probar la actividad protectora tipo-chaperona *in vitro* de las proteínas con transcritos acumulados en desecación por los organismos tolerantes. Además, sería de utilidad observar si el patrón observado de pocas proteínas tipo-LEA en organismos tolerantes a la desecación con acumulación de trehalosa se repite al analizar otros organismos. Lo anterior otorgaría un valor extra a las proteínas tipo-LEA como sustituyentes de la trehalosa para promover tolerancia. Por último, sería importante generar nuevos datos genómicos y transcriptómicos de un organismo tolerante a la desecación con el fin de encontrar proteínas nuevas con la capacidad de conferir tolerancia.

## XII. Referencias

- Al-Madboly, L. A., Ali, S. M., Fakharany, E., Ragab, A. E., Khedr, E. G., & Elokely, K. M. (2020). Stress-based production, and characterization of glutathione peroxidase and glutathione s-transferase enzymes from *Lactobacillus plantarum*. *Frontiers in bioengineering and biotechnology*, 8, 78. <https://doi.org/10.3389/fbioe.2020.00078>
- Alpert P. (2006). Constraints of tolerance: why are desiccation-tolerant organisms so small or rare?. *The Journal of Experimental Biology*, 209(Pt 9), 1575–1584. <https://doi.org/10.1242/jeb.02179>
- Breeuwer, P., Lardeau, A., Peterz, M., & Joosten, H. M. (2003). Desiccation and heat tolerance of *Enterobacter sakazakii*. *Journal of applied microbiology*, 95(5), 967–973. <https://doi.org/10.1046/j.1365-2672.2003.02067.x>
- Babu M. M. (2016). The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochemical Society Transactions*, 44(5), 1185–1200. <https://doi.org/10.1042/BST20160172>
- Baker, J., Van Dennsteele, C., & Dure, L., 3rd (1988). Sequence and characterization of 6 Lea proteins and their genes from cotton. *Plant Molecular Biology*, 11(3), 277–291. <https://doi.org/10.1007/BF00027385>
- Battaglia, M., Olvera-Carrillo, Y., Garcarrubio, A., Campos, F., & Covarrubias, A. A. (2008). The enigmatic LEA proteins and other hydrophilins. *Plant Physiology*, 148(1), 6–24. <https://doi.org/10.1104/pp.108.120725>
- Bewley, J D (1979). Physiological aspects of desiccation tolerance. *Annual Review of Plant Physiology*, 30(1), 195–238. <https://doi.org/10.1146/annurev.pp.30.060179.001211>
- Boothby, T. C., Tapia, H., Brozena, A. H., Piszkiwicz, S., Smith, A. E., Giovannini, I., Rebecchi, L., Pielak, G. J., Koshland, D., & Goldstein, B. (2017). Tardigrades use intrinsically disordered proteins to survive desiccation. *Molecular Cell*, 65(6), 975–984.e5. <https://doi.org/10.1016/j.molcel.2017.02.018>
- Cocotl-Yañez, M., Moreno, S., Encarnación, S., López-Pliego, L., Castañeda, M., & Espín, G. (2014). A small heat-shock protein (Hsp20) regulated by RpoS is essential

for cyst desiccation resistance in *Azotobacter vinelandii*. *Microbiology* (Reading, England), 160(Pt 3), 479–487. <https://doi.org/10.1099/mic.0.073353-0>

Cuevas-Velazquez, C. L., Rendón-Luna, D. F., & Covarrubias, A. A. (2014). Dissecting the cryoprotection mechanisms for dehydrins. *Frontiers in Plant Science*, 5, 583. <https://doi.org/10.3389/fpls.2014.00583>

Cuevas-Velazquez, C. L., Saab-Rincón, G., Reyes, J. L., & Covarrubias, A. A. (2016). The unstructured N-terminal region of arabidopsis group 4 Late Embryogenesis Abundant (LEA) proteins is required for folding and for chaperone-like activity under water deficit. *The Journal of Biological Chemistry*, 291(20), 10893–10903. <https://doi.org/10.1074/jbc.M116.720318>

Czernik, M., Fidanza, A., Luongo, F. P., Valbonetti, L., Scapolo, P. A., Patrizio, P., & Loi, P. (2020). Late Embryogenesis Abundant (LEA) proteins confer water stress tolerance to mammalian somatic cells. *Cryobiology*, 92, 189–196. <https://doi.org/10.1016/j.cryobiol.2020.01.009>

Dure L (1993b) Structural motifs in LEA proteins. In TJ Close, EA Bray, eds, *Plant Responses to Cellular Dehydration during Environmental Stress*. American Society of Plant Physiologists, Rockville, MD, pp 91–103

Dure, L., 3rd, Crouch, M., Harada, J., Ho, T. H., Mundy, J., Quatrano, R., Thomas, T., & Sung, Z. R. (1989). Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Molecular Biology*, 12(5), 475–486. <https://doi.org/10.1007/BF00036962>

Dyson, H. J., & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nature Reviews. Molecular Cell Biology*, 6(3), 197–208. <https://doi.org/10.1038/nrm1589>

Grzyb, T., & Skłodowska, A. (2022). Introduction to Bacterial Anhydrobiosis: A general perspective and the mechanisms of desiccation-associated damage. *Microorganisms*, 10(2), 432. <https://doi.org/10.3390/microorganisms10020432>

Gurtler, J. B., & Beuchat, L. R. (2007). Survival of *Enterobacter sakazakii* in powdered infant formula as affected by composition, water activity, and temperature. *Journal of Food Protection*, 70(7), 1579–1586. <https://doi.org/10.4315/0362-028x-70.7.1579>

Galau, G. A., Hughes, D. W., & Dure, L., 3rd (1986). Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs. *Plant Molecular Biology*, 7(3), 155–170. <https://doi.org/10.1007/BF00021327>

Garay-Arroyo, A., Colmenero-Flores, J. M., Garcarrubio, A., & Covarrubias, A. A. (2000). Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *The Journal of Biological Chemistry*, 275(8), 5668–5674. <https://doi.org/10.1074/jbc.275.8.5668>

Goyal, K., Walton, L. J., & Tunnacliffe, A. (2005). LEA proteins prevent protein aggregation due to water stress. *The Biochemical Journal*, 388(Pt 1), 151–157. <https://doi.org/10.1042/BJ20041931>

Grzyb, T., & Skłodowska, A. (2022). introduction to bacterial anhydrobiosis: a general perspective and the mechanisms of desiccation-associated damage. *Microorganisms*, 10(2), 432. <https://doi.org/10.3390/microorganisms10020432>

Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O., & Pappu, R. V. (2017). CIDER: Resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophysical Journal*, 112(1), 16–21. <https://doi.org/10.1016/j.bpj.2016.11.3200>

Holt, C., Raynes, J. K., & Carver, J. A. (2019). Sequence characteristics responsible for protein-protein interactions in the intrinsically disordered regions of caseins, amelogenins, and small heat-shock proteins. *Biopolymers*, 110(9), e23319. <https://doi.org/10.1002/bip.23319>

Hanin, M., Brini, F., Ebel, C., Toda, Y., Takeda, S., & Masmoudi, K. (2011). Plant dehydrins and stress tolerance: versatile proteins for complex mechanisms. *Plant Signaling & Behavior*, 6(10), 1503–1509. <https://doi.org/10.4161/psb.6.10.17088>

Hibshman, J. D., Clegg, J. S., & Goldstein, B. (2020). Mechanisms of desiccation tolerance: themes and variations in brine shrimp, roundworms, and tardigrades. *Frontiers in Physiology*, 11, 592016. <https://doi.org/10.3389/fphys.2020.592016>

Hughes, S., & Graether, S. P. (2011). Cryoprotective mechanism of a small intrinsically disordered dehydrin protein. *Protein Science : a publication of the Protein Society*, 20(1), 42–50. <https://doi.org/10.1002/pro.534>



Hundertmark, M., & Hinch, D. K. (2008). LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics*, 9, 118. <https://doi.org/10.1186/1471-2164-9-118>

Janis, B., Belott, C., & Menze, M. A. (2018). Role of intrinsic disorder in animal desiccation tolerance. *Proteomics*, 18(21-22), e1800067. <https://doi.org/10.1002/pmic.201800067>

Javid, B., MacAry, P. A., & Lehner, P. J. (2007). Structure and function: heat shock proteins and adaptive immunity. *Journal of Immunology (Baltimore, Md. : 1950)*, 179(4), 2035–2040. <https://doi.org/10.4049/jimmunol.179.4.2035>

Jaspard, E., & Hunault, G. (2014). Comparison of amino acids physico-chemical properties and usage of late embryogenesis abundant proteins, hydrophilins and WHY domain. *PloS One*, 9(10), e109570. <https://doi.org/10.1371/journal.pone.0109570>

Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., Logacheva, M. D., & Penin, A. A. (2016). A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *The Plant Journal : for cell and molecular biology*, 88(6), 1058–1070. <https://doi.org/10.1111/tpj.13312>

Liu, X., Zhang, Y., Yang, H., Liang, Y., Li, X., Oliver, M. J., & Zhang, D. (2020). Functional Aspects of Early Light-Induced Protein (ELIP) genes from the desiccation-tolerant moss *Syntrichia caninervis*. *International Journal of Molecular Sciences*, 21(4), 1411. <https://doi.org/10.3390/ijms21041411>

Lozano-Ortega, M. (2023). Las propiedades fisicoquímicas de las proteínas LEA y su participación en la tolerancia a la desecación de diversos organismos. Tesis de licenciatura. [https://tesiunam.dgb.unam.mx/F/S8V32SJKBVCV5YQAQ9GQN3LT7MIB6JCCA3F2DLUQ8UPX1TTQ5Q6-41511?func=full-set-set&set\\_number=957773&set\\_entry=000001&format=999](https://tesiunam.dgb.unam.mx/F/S8V32SJKBVCV5YQAQ9GQN3LT7MIB6JCCA3F2DLUQ8UPX1TTQ5Q6-41511?func=full-set-set&set_number=957773&set_entry=000001&format=999)

Mizrahi, T., Heller, J., Goldenberg, S., & Arad, Z. (2010). Heat shock proteins and resistance to desiccation in congeneric land snails. *Cell Stress & Chaperones*, 15(4), 351–363. <https://doi.org/10.1007/s12192-009-0150-9>

Maia, J., Dekkers, B. J., Dolle, M. J., Ligterink, W., & Hilhorst, H. W. (2014). Abscisic acid (ABA) sensitivity regulates desiccation tolerance in germinated Arabidopsis seeds. *The New Phytologist*, 203(1), 81–93. <https://doi.org/10.1111/nph.12785>

Oliver, M. J., Farrant, J. M., Hilhorst, H., Mundree, S., Williams, B., & Bewley, J. D. (2020). Desiccation tolerance: avoiding cellular damage during drying and rehydration. *Annual Review of Plant Biology*, 71, 435–460. <https://doi.org/10.1146/annurev-arplant-071219-105542>

Olvera-Carrillo, Y., Campos, F., Reyes, J. L., Garcarrubio, A., & Covarrubias, A. A. (2010). Functional analysis of the group 4 late embryogenesis abundant proteins reveals their relevance in the adaptive response during water deficit in Arabidopsis. *Plant Physiology*, 154(1), 373–390. <https://doi.org/10.1104/pp.110.158964>

Rabert, C., Inostroza, K., Bravo, S., Sepúlveda, N., & Bravo, L. A. (2020). Exploratory study of fatty acid profile in two filmy ferns with contrasting desiccation tolerance reveal the production of very long chain polyunsaturated omega-3 fatty acids. *Plants (Basel, Switzerland)*, 9(11), 1431. <https://doi.org/10.3390/plants9111431>

Ren, Q., Brenner, R., Boothby, T. C., & Zhang, Z. (2020). Membrane and lipid metabolism plays an important role in desiccation resistance in the yeast *Saccharomyces cerevisiae*. *BMC Microbiology*, 20(1), 338. <https://doi.org/10.1186/s12866-020-02025-w>

Rivera-Najera, L. Y., Saab-Rincón, G., Battaglia, M., Amero, C., Pulido, N. O., García-Hernández, E., Solórzano, R. M., Reyes, J. L., & Covarrubias, A. A. (2014). A group 6 late embryogenesis abundant protein from common bean is a disordered protein with extended helical structure and oligomer-forming properties. *The Journal of Biological Chemistry*, 289(46), 31995–32009. <https://doi.org/10.1074/jbc.M114.583369>

Sugio, S., Hiraoka, B. Y., & Yamakura, F. (2000). Crystal structure of cambialistic superoxide dismutase from *Porphyromonas gingivalis*. *European Journal of Biochemistry*, 267(12), 3487–3495. <https://doi.org/10.1046/j.1432-1327.2000.01373.x>

Shih, M. D., Hsieh, T. Y., Lin, T. P., Hsing, Y. I., & Hoekstra, F. A. (2010). Characterization of two soybean (*Glycine max* L.) LEA IV proteins by circular dichroism and Fourier transform infrared spectrometry. *Plant & Cell Physiology*, 51(3), 395–407. <https://doi.org/10.1093/pcp/pcq005>

Srikumar, S., Cao, Y., Yan, Q., Van Hoorde, K., Nguyen, S., Cooney, S., Gopinath, G. R., Tall, B. D., Sivasankaran, S. K., Lehner, A., Stephan, R., & Fanning, S. (2019). RNA sequencing-based transcriptional overview of xerotolerance in *Cronobacter sakazakii* SP291. *Applied and Environmental Microbiology*, 85(3), e01993-18. <https://doi.org/10.1128/AEM.01993-18>

Tapia, H., Young, L., Fox, D., Bertozzi, C. R., & Koshland, D. (2015). Increasing intracellular trehalose is sufficient to confer desiccation tolerance to *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*, 112(19), 6122–6127. <https://doi.org/10.1073/pnas.1506415112>

Theillet, F. X., Binolfi, A., Frembgen-Kesner, T., Hingorani, K., Sarkar, M., Kyne, C., Li, C., Crowley, P. B., Gierasch, L., Pielak, G. J., Elcock, A. H., Gershenson, A., & Selenko, P. (2014). Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chemical Reviews*, 114(13), 6661–6714. <https://doi.org/10.1021/cr400695p>

Toxopeus, J., Warner, A. H., & MacRae, T. H. (2014). Group 1 LEA proteins contribute to the desiccation and freeze tolerance of *Artemia franciscana* embryos during diapause. *Cell Stress & Chaperones*, 19(6), 939–948. <https://doi.org/10.1007/s12192-014-0518-3>

Trainor F. R., Gladych R. (1995). Survival of algae in a desiccated soil: a 35-year study. *Phycologia* 34 191–192 <https://doi.org/10.2216/i0031-8884-34-3-191.1>

Voronina, T. A., Nesselov, A. A., Kondratyeva, S. A., Deviatiiarov, R. M., Miyata, Y., Tokumoto, S., Cornette, R., Gusev, O. A., Kikawada, T., & Shagimardanova, E. I. (2020). New group of transmembrane proteins associated with desiccation tolerance in the anhydrobiotic midge *Polypedilum vanderplanki*. *Scientific Reports*, 10(1), 11633. <https://doi.org/10.1038/s41598-020-68330-6>

- Wright, P. E., & Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nature reviews. Molecular Cell Biology*, 16(1), 18–29. <https://doi.org/10.1038/nrm3920>
- Xiao, L., Yang, G., Zhang, L., Yang, X., Zhao, S., Ji, Z., Zhou, Q., Hu, M., Wang, Y., Chen, M., Xu, Y., Jin, H., Xiao, X., Hu, G., Bao, F., Hu, Y., Wan, P., Li, L., Deng, X., Kuang, T., ... He, Y. (2015). The resurrection genome of *Boea hygrometrica*: A blueprint for survival of dehydration. *Proceedings of the National Academy of Sciences of the United States of America*, 112(18), 5833–5837. <https://doi.org/10.1073/pnas.1505811112>
- Yamagata, H., & Bowler, C. (1997). Molecular cloning and characterization of a cDNA encoding early light-inducible protein from soybean (*Glycine max L.*). *Bioscience, Biotechnology, and Biochemistry*, 61(12), 2143–2144. <https://doi.org/10.1271/bbb.61.2143>
- Yoshida, Y., & Tanaka, S. (2022). Deciphering the biological enigma-genomic evolution underlying anhydrobiosis in the phylum tardigrada and the chironomid *Polypedilum vanderplanki*. *Insects*, 13(6), 557. <https://doi.org/10.3390/insects13060557>