



UNIVERSIDAD DE BUENOS AIRES

FACULTAD DE CIENCIAS EXACTAS Y NATURALES

Departamento de Biodiversidad y Biología Experimental

**Caracterización de las LEAs 4 y 5 en la respuesta a estrés
salino en *Chenopodium quinoa* Willd.**

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en
el área de Ciencias Biológicas

Lic. AXEL JOEL RIZZO

Director: **Dr. Hernán Pablo Burrieza**

Co-directora: **Dra. Alicia Mercedes Zelada**

Consejero de Estudios: **Dr. Norberto Iusem**

Lugar de trabajo: **Laboratorio de Biología del Desarrollo de las Plantas**

Departamento de Biodiversidad y Biología Experimental (DBBE)

Instituto de Biodiversidad y Biología Experimental Aplicada (IBBEA-UBA-CONICET)

Ciudad Autónoma de Buenos Aires, Argentina – 28 de julio del 2023

Resumen

La salinización de los suelos tiene un impacto devastador en los cultivos, ocasionando pérdidas económicas significativas debido a la reducción de la productividad y calidad de las cosechas. *Chenopodium quinoa* Willd (quinua) es un cultivo originario de los Andes, valorado por su alto perfil nutricional y adaptabilidad a diferentes condiciones climáticas. Una de las características destacadas de la quinua es su notable tolerancia a varios tipos de estrés abiótico. Esto ha despertado un gran interés en la investigación científica, ya que la quinua puede proporcionar valiosos recursos genéticos y mecanismos adaptativos que pueden ser utilizados en el mejoramiento de cultivos. Las plantas han desarrollado estrategias y mecanismos de adaptación frente al estrés, incluyendo la síntesis de proteínas específicas involucradas en la tolerancia al estrés como las proteínas LEAs (del inglés, *Late Embryogenesis Abundant*). Las LEAs son proteínas altamente conservadas y ampliamente distribuidas en el reino vegetal que, a pesar de su nombre pueden ser halladas en otras etapas del ciclo de vida de las plantas y en diversos tejidos. Se caracterizan por su acumulación en respuesta a diferentes tipos de estrés abiótico y tienen la capacidad de proteger las células vegetales contra el daño inducido por el estrés, preservando la estructura y/o función de diferentes biomoléculas.

El **objetivo general** del trabajo fue el identificar proteínas LEA presentes en *C. quinoa* Willd. y caracterizar candidatas relacionadas con la respuesta al estrés salino y el desarrollo de la semilla madura.

El objetivo del Capítulo 1 fue caracterizar el proteoma de semillas de distintos cultivares de quinua provenientes de regiones con condiciones edafoclimáticas contrastantes. Los resultados obtenidos brindan una visión sobre las vías metabólicas más importantes para las semillas maduras de quinua. Interesantemente, el perfil de LEAs observado sugiere un rol de las LEA_4, LEA_5 y SMP (del inglés *seed maturation protein*) en la viabilidad de las semillas, probablemente relacionado con la capacidad de tolerar la desecación inherente al proceso de formación de las semillas ortodoxas. Los cultivares Sajama, Amarilla de Maranganí y Nariño presentaron un patrón de expresión de estas proteínas compatible con el observado en semillas ortodoxas, por el contrario, el cultivar Chadmo, de origen chileno, mostró menores niveles de abundancia de estas mismas LEAs. Esto podría reflejar otra característica distintiva entre los dos eventos de domesticación recientemente sugeridos para *C. quinoa*.

El objetivo del Capítulo 2 fue identificar todas las proteínas LEA presentes en el genoma de *C. quinoa*, y caracterizar estructural y filogenéticamente a los grupos LEA_3 y LEA_4. Se identificaron 43 proteínas LEAs distribuidas de la siguiente forma: LEA_1 (5), LEA_2 (12), LEA_3 (5), LEA_4 (4), LEA_5 (4), LEA_6 (2) y SMPs (11). Para las LEA_3 se identificaron 3 subgrupos, dos de los cuales conforman grupos ortólogos separados con

características físico-químicas similares, mientras que el tercer subgrupo no posee información suficiente para verificar su ancestralidad. En quinua se hallaron representantes de los tres subgrupos de LEA_3 definidos. Para las LEA_4 se identificaron 3 subgrupos, uno de los cuales comprendía únicamente especies dicotiledóneas, otro solo con monocotiledóneas, ambos conformando grupos ortólogos separados, y un tercer subgrupo sin suficiente evidencia para verificar su ancestralidad. Todos los subgrupos de LEA_4 mostraron diferencias en varias de las características físico-químicas estudiadas. Un candidato de cada grupo de LEAs fue seleccionado para realizar estudios de expresión.

El objetivo del Capítulo 3 fue caracterizar la expresión de las LEA_3 y LEA_4 en respuesta al estrés salino.

Las LEA_3 mostraron estar presentes en plántulas de cuatro genotipos (Cahuil, BO25, Villarrica y Sajama) y en hojas, raíces y flores de plantas adultas del genotipo BO25. No se detectaron cambios significativos en la expresión de este gen por tratamiento salino en ninguno de los estadios de desarrollo analizados.

Las LEA_4 se encuentran presentes en plántulas y muestran una inducción bajo tratamiento salino en tres de los cuatro genotipos estudiados. En plantas adultas en condición control el transcripto resultó indetectable en hojas y raíces, mientras que en flores se encontraba presente. Se detectó una inducción de la LEA de interés en raíces de plantas del genotipo BO25 sometidas a estrés salino.

Se procedió con el clonado de la LEA_4 de interés para su estudio funcional mediante expresión en sistemas heterólogos, obteniendo hasta el momento la secuencia codificante de una de sus isoformas en un vector pGEM-T Easy.

Palabras clave

Quinua - Estrés salino - Proteínas de embriogénesis tardía - Proteómica - Análisis de expresión

Characterization of *Chenopodium quinoa* Willd. LEAs 4 and 5 in response to saline stress

Abstract

Soil salinization has a devastating impact on crops, causing significant economic losses due to reduced productivity and crop quality. *Chenopodium quinoa* Willd (quinoa) is a crop native to the Andes, valued for its high nutritional value and adaptability to different climatic conditions. One of the outstanding characteristics of quinoa is its remarkable tolerance to abiotic stress. This has raised great interest in scientific research, since quinoa can provide valuable genetic resources and adaptive mechanisms that can be used in crop improvement. Plants have developed strategies and mechanisms to adapt to stress, including the synthesis of specific proteins involved in stress tolerance such as LEAs (Late Embryogenesis Abundant) proteins. LEAs are highly conserved and widely distributed proteins in the plant kingdom that, despite their name, can be found in other stages of the plant life cycle and in various tissues. They are characterized by their accumulation in response to different types of abiotic stress and have the ability to protect plant cells against stress-induced damage, preserving the structure and/or function of different biomolecules.

The general objective of the work was to identify LEA proteins present in *C. quinoa* Willd. and to characterize candidates related to the response to saline stress and the development of the mature seed.

The objective of Chapter 1 was to characterize the seed proteome of different quinoa cultivars from regions with contrasting edaphoclimatic conditions. The results obtained provide insight into the most important metabolic pathways for mature quinoa seeds. Interestingly, the observed LEA profile suggests a role for LEA_4, LEA_5 and SMP (seed maturation protein) in seed viability, probably related to the ability to tolerate desiccation inherent to the seed maturation process of orthodox seeds. Sajama, Amarilla de Maranganí and Nariño cultivars presented an expression pattern of these proteins compatible with that observed in orthodox seeds, on the contrary, the Chadmo, of Chilean origin, showed lower levels of abundance of these same LEAs. This could reflect another distinguishing feature between the two suggested domestication events for *C. quinoa*.

The objective of Chapter 2 was to identify all the LEA proteins present in the *C. quinoa* genome, and to characterize the groups LEA_3 and LEA_4 structurally and phylogenetically. A total of 43 LEA proteins were identified, distributed as follows: LEA_1 (5), LEA_2 (12), LEA_3 (5), LEA_4 (4), LEA_5 (4), LEA_6 (2) and SMPs (11). For LEA_3, 3 subgroups were identified, two of which make up separate orthologous groups with similar physicochemical

characteristics, while the third subgroup does not have sufficient information to verify its ancestry. In quinoa, representatives of the three subgroups of LEA_3 defined were found. For LEA_4, 3 subgroups were identified, one of which included only dicotyledonous species, another only with monocots, both forming separate orthologous groups, and a third subgroup without sufficient evidence to verify their ancestry. All LEA_4 subgroups showed differences in several of the physicochemical characteristics studied. One candidate from each group of LEAs was selected to perform expression studies.

The objective of Chapter 3 was to characterize the expression of LEA_3 and LEA_4 in response to salt stress.

The LEA_3 of interest was shown to be present in seedlings of four genotypes (Cahuil, BO25, Villarica and Sajama) and in leaves, roots and flowers of adult plants of the BO25 genotype. No significant changes in the expression of this gene were detected by saline treatment in any of the development stages analyzed.

The LEA_4 of interest is present in seedlings and show an induction under saline treatment in three of the four genotypes studied. In adult plants in control condition, the transcript was undetectable in leaves and roots, while it was present in flowers. An induction of the LEA of interest was detected in roots of plants of the BO25 genotype subjected to saline stress.

The LEA_4 of interest was cloned for its functional study by means of expression in heterologous systems, thus far obtaining the coding sequence of one of its isoforms in a pGEM-T Easy vector.

Keywords

Quinoa – Salt stress – Late embryogenesis abundant proteins - Proteomics – Expression analysis

Agradecimientos

A mi madre y mi padre. Por bancarme en cada paso que doy en la vida. Por ofrecerme apoyo en todas mis decisiones y mis luchas. Por preocuparse y festejarme. Por calmarme, cuidarme y quererme. Por estar siempre. Por ser quienes son. Por la vida.

A mi hermana, Li, porque a pesar de las distancias y los desencuentros conseguimos encontrarnos y bancarnos en algunos de los pasos más importantes de nuestras vidas.

A Fátima, mi compañera, por todos los años de compañía en el camino de la vida. Por ayudarme a superarme todos los días. Por permitirme estar al lado suyo durante todos los años que fuimos creciendo y creciendo en tantos aspectos. Por ayudarme a abrazar cosas nuevas sin culpa. Por la hermosa sensación de saber que contamos el uno con el otro cuando nos necesitamos. Por todo el amor.

A mi familia. Por renovar mi confianza en mí mismo en cada encuentro y saber brindarme esas palabras de aliento que hacen que todo sea más liviano. Por las risas, los encuentros y por estar.

Al CONICET. Por ofrecer la oportunidad de transitar esta etapa de mi formación.

A la Universidad de Buenos Aires, y la Facultad de Ciencias Exactas y Naturales, por darme la oportunidad de formarme en ciencia en una Universidad pública, gratuita y de primera calidad.

A Sara Maldonado y al Laboratorio de Biología del Desarrollo de las Plantas por abrirme sus puertas y darme la oportunidad de profundizar mis estudios y formarme en el ámbito académico.

A Belén y Alejandra. Grandes compañeras de laboratorio que además de trabajar a la par durante todos estos años, me dieron el lujo de su amistad. Por todas las veces que nos apoyamos entre nosotros para avanzar en esto que es hacer ciencia. Por la inspiración que me da ver el esfuerzo que ponen día a día en su trabajo. Por las meriendas, los experimentos compartidos, las pláticas sobre ciencia y la vida.

A Leito. Por ponerle espíritu festivo al laboratorio. Por el compañerismo y la amistad. Por escuchar todo berrinche. Por la autenticidad.

A Leo. Por ser la primera persona que confió en mi capacidad en la investigación y me presentó a las personas con las que trabajé todos estos años. Por las enseñanzas. Por las charlas de ciencia, filosofía y humanidad. Por la música y los haikus. Por la generosidad.

A Vane y Laino. Por las bancadas mutuas, la colaboración, la buena onda y la ayuda.

A Andrea. Por estos años compartidos de trabajo, charlas y apoyo.

A la gente del DBBE, IBBEA e INMIBO. Investigadores y administrativos, que pusieron sus recursos y trabajo a disposición para ayudarme a transitar esta etapa. Mención especial para Pilar, que además de ocuparse de los papeles, se preocupó por la persona.

A Ezequiel Petrillo, que además de ver crecer este trabajo siempre dio un extra e hizo lo posible para que yo pudiera elevar la calidad de mi investigación. A Flor Rodríguez e integrantes de IFIBYNE que me dieron acceso a varios recursos que fueron fundamentales para llevar adelante este trabajo.

A Dani Hochbaum. Por poner a disposición equipo y recursos importantísimos para esta tesis. Por las palabras de aliento.

A Sebas. Que además de su amistad desde tiempos ya inmemoriales de la cursada, hoy todavía me muestran que unx no tiene por qué transitar nada solx. Por sus aportes a la mente, al cuerpo y al alma (bioinformática, sonadas de espalda y amor, respectivamente).

A Aldana. Prueba irrefutable de que el tiempo pasa, pero la amistad siempre está. Por las palabras de aliento, la confianza y por estar ahí en los momentos más turbulentos del proceso.

A Fleer. Por las tortas fritas, la ayuda en bioinformática, por sacrificar la oreja cuando no quedaba otra que despotricar. No sé si están en orden de importancia...

A Santi, Hora y Car. Amigxs, científicxs y personas de fierro. Por los momentos de amistad compartidos. Por las perspectivas, las anécdotas y el apoyo.

A Sango, Palo, Mari y Mica. Compañerxs de cursada y de trabajo de estos últimos años que hicieron que cursar y estudiar tuviera ese plus. Por los saludos de los jueves.

A lxs compañerxs de CM1 y afines, que se preocuparon y compartieron almuerzos y charlas durante todos estos años.

A Paula, Jess, Tito, Rodri, Agus, Diana y Gonza. Personas con capacidades increíbles, con un corazón enorme y con una fuerza imparable. Gracias por tantos años de amistad, de juntadas, de escuchar penas y glorias. Por las palabras de apoyo, la inspiración y la tranquilidad que me transmiten en cada encuentro. Por el invaluable apoyo en cada proyecto.

A Guido, Caro, Dabo, Santi, Yami, Belén y Eve. Un grupo que rebalsa de cariño, amor y apoyo. Gracias por existir.

A Mar, Cato y Javi, y a lxs Cryptids. Por darle de comer a la parte de mi cabeza que me dice "no todo es ciencia".

A Natalia. Por ayudarme, desde el 2020, a atravesar épocas de desesperanza, desesperación, tensión, nervios, ansiedad e incertidumbre con toda la gracia y elegancia posible. Por marcarme el camino para ver aquello que no veía. Por darme herramientas para entender las cosas que me pasan. Por escucharme, explicarme, entenderme y por ayudarme a reconstruirme y redefinirme.

A Hernán, mi director. Por confiarme proyectos importantes y confiar en mis capacidades. Por entender a la persona atrás del científico en formación. Por las charlas sobre la vida. Por darme libertades en la mesada y en mis objetivos.

A Alicia, mi codirectora. Por ser parte de mi formación. Por ayudarme a adquirir herramientas y conocimientos fundamentales. Por la confianza. Por los esfuerzos. Por las palabras y los consejos.

A todos los que ayudaron desinteresadamente en este proyecto de investigación y de vida.

Gracias totales.

A mamá y papá...

Índice

Resumen	2
Palabras clave.....	3
Abstract.....	4
Keywords	5
Agradecimientos	6
Objetivos e hipótesis.....	13
Introducción general	14
1. La quinua como cultivo (<i>Chenopodium quinoa</i> Willd.).....	14
2. El potencial de la quinua.	16
3. Estrés en plantas: Factores de estrés y estrés abiótico	17
SECCIÓN I	19
Capítulo I	
Proteómica de semillas de <i>Chenopodium quinoa</i> Willd.: contexto fisiológico y expresión de LEAs	20
I.1 Introducción.....	20
I.1.1 Generalidades de las semillas de quinua.	20
I.1.2 Aspectos nutricionales	21
I.1.3 Las proteínas LEA en semilla.....	22
I.2 Materiales y Métodos	24
I.2.1 Material vegetal	24
I.2.2 Extracción y digestión	24
I.2.3 Espectrometría de masa	24
I.2.4 Análisis de datos de proteómica.....	25
I.2.5 Análisis de secuencias	26
I.3 Resultados	28
I.3.1 Metabolismo de hidratos de carbono	31
I.3.2 Metabolismo y procesamiento de proteínas de reserva	34
I.3.3 Metabolismo y reserva de lípidos	36
I.3.4 Estrés oxidativo	36
I.3.5 Mecanismos de expresión.....	37
I.3.6 Proteínas LEA, HSPs y chaperonas.....	38
I.4 Discusión y conclusiones.....	44
I.4.1 Comentarios generales	44
I.4.2 Metabolismo de hidratos de carbono	44

I.4.3 Metabolismo y almacenamiento de proteínas	45
I.4.4 Metabolismo y almacenamiento de lípidos	46
I.4.5 Estrés oxidativo	47
I.4.6 Mecanismos de expresión	47
I.4.7 Proteínas LEA, Heat shock y chaperonas	48
I.4.8 Conclusiones generales	49
SECCIÓN II	50
Introducción	51
1. Estrés abiótico	51
1.1 Salinidad	51
1.2. Efectos de la salinidad	51
1.3. Mecanismos de tolerancia	52
1.3.1 Respuesta general	52
1.3.2 Respuesta fisiológica y molecular	52
2. Estrés abiótico en quinoa	54
2.1 Efectos observados	54
2.2 Estadios de desarrollo y tolerancia	55
3. Las proteínas LEA	56
3.1 Descubrimiento y nomenclatura	56
3.2 Expresión y regulación de las LEAs	57
Capítulo II	
Análisis bioinformático de LEAs en C. quinoa con foco en las LEA_3 y LEA_4: diversidad y evolución	59
II.1 Introducción	59
II.1.1 Diversidad de las proteínas LEA	59
II.1.2 Las LEA_4 y LEA_3	60
II.2 Materiales y Métodos	62
II.2.1 Obtención de secuencias de interés	62
II.2.2 Análisis físico-químico de las secuencias de interés	62
II.2.3 Búsqueda de secuencias conservadas mediante MEME suite.	63
II.2.4 Alineamientos y construcción de árboles	63
II.2.5 Análisis manual de microsinténia	63
II.3 Resultados	64
II.3.1 Construcción de las bases de secuencias curadas para las bases de datos de LEAs	64
II.3.2 Características físico-químicas de las LEAs de <i>C. quinoa</i>	64

II.3.3 Análisis de las LEA_3.....	71
II.3.3.1 Búsqueda de motivos conservados en las LEA_3	71
II.3.3.2 Distribución taxonómica de los motivos hallados	73
II.3.3.3 Análisis filogenético y microsinténico	74
II.3.4 Análisis de las LEA_4.....	79
II.3.4.1 Búsqueda de motivos conservados en las LEA_4 y su distribución.....	79
II.3.4.2 Análisis filogenético, físico-químico y microsinténico de las LEA_4.....	82
II.4 Discusión	88
II.4.1 La clasificación de las LEAs	88
II.4.2 El grupo LEA_3	89
II.4.3 El grupo LEA_4	92
II.4.4 Las LEAs de los grupos LEA_3 y LEA_4 en <i>C. quinoa</i>	93
II.5 Conclusiones	95
Capítulo III	
Expresión de dos proteínas LEA de interés (LEA_4 y LEA_3) en diferentes estadios de desarrollo y tejidos de <i>C. quinoa</i> sometidos a estrés salino	97
III.1 Introducción.....	97
III.1.1 Estadios de desarrollo y expresión de LEAs	97
III.1.2 Las funciones de las LEA en el estrés abiótico	97
III.2 Materiales y métodos	99
III.2.1 Material vegetal	99
III.2.2 Esterilización	99
III.2.3 Siembra y ensayos de crecimiento de plántula	99
III.2.4 Ensayos de estrés salino en plantas adultas.....	99
III.2.5 Morfometría	100
III.2.6 Extracción de ARN de plántula.....	100
III.2.8 Extracción de ARN de planta adulta.....	101
III.2.9 Precipitación de ARN con acetato de sodio.	101
III.2.10 Verificación de integridad del ARN por gel de agarosa	101
III.2.11 Obtención de cDNA.....	102
III.2.12 Real time qPCR y <i>primers</i>	102
III.2.13 Geles de agarosa para verificación de qPCR.....	102
III.3 Resultados	103
III.3.1 Efecto de la salinidad sobre la morfología de las plántulas.....	103

III.3.2 Expresión de LEA_3 (<i>late embryogenesis abundant protein</i> Lea5-like) y LEA_4 (<i>late embryogenesis abundant protein</i> ECP63-like) en plántulas de cuatro genotipos de <i>C. quinoa</i>	104
III.3.3 Expresión de LEA_3 y LEA_4 en tejidos de planta adulta	105
III.4 Discusión y conclusiones.....	108
Bibliografía.....	110
Anexo I – Herramienta de análisis y manipulación de secuencias en Python.122	
1. <i>Script</i> para análisis de secuencias proteicas por medio de ProtParam	122
2. <i>Script</i> para tratamiento de headers de archivos FASTA.....	122
3. <i>Script</i> para análisis de FoldIndex de secuencias proteicas.	124
Anexo II – Protocolos y <i>primers</i> de <i>Real Time</i> PCR	125
A. Protocolo de ciclado utilizado para RT-qPCR MyCycler BIO-RAD	125
B. <i>Primers</i> utilizados para clonados y <i>Real time</i> qPCR	125
Anexo III – Listado de accesiones y encabezados de LEA_3 y LEA_4.....	127
1. Encabezados de LEA_3.....	127
2. Encabezados de LEA_4.....	133

Objetivos e hipótesis

OBJETIVO GENERAL

Identificar proteínas LEA presentes en *C. quinoa* Willd. y caracterizar candidatas relacionadas con la respuesta al estrés salino y al desarrollo de la semilla madura.

OBJETIVOS ESPECÍFICOS

1. Caracterizar el proteoma de las semillas de diferentes cultivares de *C. quinoa* provenientes de regiones con condiciones edafoclimáticas contrastantes, centrándose específicamente en las proteínas tipo LEA.

2. Identificar las proteínas LEA en el genoma de *C. quinoa* Willd. y analizar las características físico-químicas de cada uno de los grupos.

3. Caracterizar en profundidad la diversidad genética de las proteínas tipo LEA_3 y LEA_4 y examinar su respuesta al estrés salino tanto en plántulas como en plantas adultas de diferentes variedades de *C. quinoa*.

HIPÓTESIS

1. Los proteomas de semillas maduras de *C. quinoa* provenientes de ambientes edafoclimáticos contrastantes presentarán diferencias que están relacionadas con su origen.

2. La diversidad genética presente dentro de cada grupo LEA determinado por criterios filogenéticos y dominios conservados, correlaciona con proteínas de diferentes características físico-químicas.

3. La respuesta al estrés salino de las LEAs_3 y 4 depende del genotipo estudiado y de la etapa de desarrollo de la planta.

Capítulo II

Análisis bioinformático de LEAs en *C. quinoa* con foco en las LEA_3 y LEA_4: diversidad y evolución

II.1 Introducción

II.1.1 Diversidad de las proteínas LEA

Dentro de la familia de proteínas de embriogénesis tardía, ha habido pocos grupos que han sido objeto de estudios detallados. Entre estos grupos, las dehidrinas han sido las proteínas que más atención han recibido a lo largo de los años. Sin embargo, se ha observado que otras proteínas LEA pertenecientes a diferentes grupos poseen características y dinámicas de expresión similares a las de las dehidrinas (Cheng et al., 2002; Gu et al., 2012; Olvera-Carrillo et al., 2010; Rorat, 2006)

Algunas publicaciones que abordaron con diferentes perspectivas la clasificación de las LEA lograron generar diferentes sistemas que dejan en evidencia no solo la complejidad y dificultad de su clasificación, sino también de la diversidad de LEAs existentes. Si bien algunas de ellas pueden ser identificadas consistentemente por tener, además de determinadas propiedades físico-químicas, una secuencia conservada, otras presentan bloques repetitivos que permiten su identificación, pero podrían no abarcar la totalidad de su diversidad (Battaglia et al., 2008; Dure et al., 1989; Tunnacliffe & Wise, 2007; Wise, 2003). Un análisis exhaustivo contemplando la totalidad de las LEAs halladas en *Arabidopsis thaliana* (Bies-Ethève et al., 2008) realizan una búsqueda de motivos cuyos resultados justificaban una reclasificación. Sin embargo, al estar el análisis limitado a las LEAs halladas en *A. thaliana*, limita la aplicabilidad de los nuevos criterios propuestos a todas las especies de plantas.

El trabajo realizado por Battaglia et al. (2008) en términos de clasificación y descripción proporciona una visión general importante sobre las diferencias existentes entre los diferentes grupos de proteínas LEAs. A lo largo de los años, otros estudios han intentado complementar este análisis en mayor profundidad mediante la utilización de métodos bioinformáticos (Singh & Graether, 2020; Tunnacliffe & Wise, 2007; Wise, 2003). Algunos de ellos permiten observar la información disponible desde otra perspectiva, asociando la clasificación a “clases” definidas por propiedades grupales, y no por grupos con secuencias conservadas. Al mismo tiempo, ponen un mayor foco en regiones de baja complejidad, en general ignoradas a la hora de llevar adelante clasificaciones de proteínas (Tunnacliffe & Wise, 2007; Wise, 2003).

En cuanto al abordaje de las LEAs desde una perspectiva evolutiva, el trabajo de Artur et al., (2019), aborda la diversificación de los diferentes grupos de LEAs con un foco en angiospermas. Las principales conclusiones de su trabajo incluyen un análisis profundo de

las LEA_1, LEA_2 y dehidrinas, haciendo referencia a la diversificación ancestral, las duplicaciones en tándem y el surgimiento de nuevos motivos. Por otra parte, en un análisis sobre las LEA_4, presenta el rol de la dinámica sinténica en el enriquecimiento del grupo y en su función en la tolerancia en la desecación, al observar comunidades sinténicas sumamente distantes que parecen no guardar relación entre sí. Sin embargo, su estudio sobre las LEA_4 se limita a las observaciones evolutivas, y a pesar de estar incluido, no presenta información relacionada con estos aspectos sobre el grupo de las LEA_3. Otro aporte en el trabajo sobre *A. thaliana* (Bies-Ethève et al., 2008), es el hallazgo de repeticiones duplicaciones y repeticiones en tándem de LEAs de varios grupos en el genoma, indicando que los grupos LEA_2, LEA_4 y LEA_4 poseen duplicaciones producto de una duplicación total del genoma (WGD), y los grupos de dehidrinas, SMPs y LEA_6 son los únicos que presentan duplicaciones en tándem.

II.1.2 Las LEA_4 y LEA_3

Hoy día una de las clasificaciones más utilizadas es la de Pfam (Mistry et al., 2021) que mediante modelos ocultos de Markov (HMM) desarrollados a partir de secuencias curadas logra identificar proteínas correspondientes a cada grupo. Con fines prácticos nos referiremos a las proteínas de interés de este trabajo como LEA_4 y LEA_3 haciendo uso de la clasificación según Pfam.

Las proteínas LEA_3 (PF003242) poseen cierta variación en la nomenclatura utilizada para su clasificación en la bibliografía, dificultando el análisis de los datos experimentales obtenidos por distintos grupos de investigación. Su actual clasificación corresponde con una subclasificación según la clasificación de Battaglia (2008) y se las denomina LEA 5B y corresponde a las D-73 según Dure (Dure et al., 1989), mientras que Bies-Etheve (2008) las incluye en el grupo que denomina grupo 6. Otros nombres que pueden atribuirse estas proteínas es el de AtD121, Sag21 y lea5. Al mismo tiempo poseen una expresión vinculada a sequía, luz UV, salinidad, frío y heridas (Kim et al., 2005; Kiyosue et al., 1992; Maitra & Cushman, 1994; Park et al., 2003; Stacy et al., 1999; Zegzouti et al., 1997). Su clasificación en Pfam incluye una secuencia conservada asociada de 93 aminoácidos, aunque solo ciertas secciones de la misma presentan altos niveles de conservación en el HMM, llamando la atención la conservación de un residuo triptófano en la posición 65 del mismo.

Las proteínas correspondientes al grupo 4 (PF002987), por su parte, cuentan con una secuencia de 44 aminoácidos con alto grado de conservación (KDYAAEKAKEAKDATAEKAGEYKDYAAEKAGEAKDAKAAEKAG), que inicialmente era

considerado una repetición de una secuencia de 11 aminoácidos que se utiliza para identificación mediante alineamientos. Este grupo según Pfam se corresponde con el llamado grupo de las LEA 3, con los grupos D-7 y D-29 según las primeras clasificaciones o ECP63, PAP240 y PM27 por sus nombres comunes (Battaglia et al., 2008; Bies-Ethève et al., 2008). A su vez, algunos autores como Battaglia (2008) proponen una subclasificación del grupo en al menos dos grupos dadas algunas variantes en el motivo de 11 aminoácidos mencionado. En este trabajo se abordará más adelante la posible subclasificación de las LEA_4 en detalle basándonos en la validación de nuevos motivos para identificar diferentes LEAs de este grupo y sus implicancias evolutivas. Este grupo de LEAs fue vinculado con la estabilización de actividad enzimática en pruebas *in vitro* (Goyal et al., 2005; Grelet et al., 2005; Nakayama et al., 2008; Reyes et al., 2005) y con la tolerancia a estrés abiótico en organismos bacterianos, algas y nematodos (Furuki et al., 2012; Honjoh et al., 2000; Yu et al., 2005; L. Zhang et al., 2000), y existe evidencia de la interacción proteína-proteína para algunas integrantes de este grupo (Dirk et al., 2020).

Algunos de los estudios que involucran el análisis de expresión de LEAs en diferentes tejidos y bajo diferentes estreses reportan la detección de LEA_3 y LEA_4 en semillas, raíces, tallos hojas y flores de *A. thaliana*, *Carica papaya* y *Solanum tuberosum*, así como su inducción bajo tratamiento salino, por sequía o de bajas temperaturas (Bies-Ethève et al., 2008; Y. Chen et al., 2019; Hundertmark & Hinch, 2008; Z. Zou et al., 2022). En algunos casos, como el de *C. papaya*, la inducción más importante se detecta en LEAs del grupo LEA_4 en raíces de plantas sometidas a estrés por sequía y en hojas de plantas sometidas a estrés salino (Z. Zou et al., 2022).

II.2 Materiales y Métodos

II.2.1 Obtención de secuencias de interés

Para obtener las secuencias de trabajo se utilizó como referencia el *seed* o conjunto de secuencias consenso alineadas que provee Pfam (Mistry et al., 2021) para la identificación de las proteínas de interés.

Las búsquedas iniciales fueron realizadas utilizando la base de datos de proteínas de NCBI (Sayers et al., 2022) y Phytozome (Goodstein et al., 2012). Los resultados de las búsquedas para cada especie fueron utilizados para obtener además información como: identificador taxonómico de NCBI (taxid), accesión de proteína y su secuencia, accesión de mensajero y su secuencia.

Las secuencias obtenidas fueron filtradas utilizando nuevamente HMMScan (Potter et al., 2018), que permite asociar secuencias proteicas a dominios, familias o grupos de proteínas. Se preparó un set de datos exclusivo compuesto solo por aquellas accesiones que se asociaron automáticamente con el grupo de interés (LEA_3 o LEA_4 según Pfam).

Este procedimiento fue aplicado para la totalidad de las secuencias obtenidas de las especies seleccionadas para el análisis con el objetivo de evitar utilizar secuencias que podrían asociarse espuriamente con el grupo de interés a la hora de buscar motivos y caracterizar físico-químicamente los grupos.

II.2.2 Análisis físico-químico de las secuencias de interés

Se utilizaron las secuencias curadas para obtener información físico-química sobre las mismas. Se extrajo la masa teórica (MW), punto isoeléctrico (pI), índice de hidrofobicidad (GRAVY) mediante el uso de un *script* escrito para Python (ver. 3.1) que utiliza un API de Expasy ProtParam (Gasteiger et al., 2005) y puede ser hallado en el Anexo I. Información relacionada con la formación de diferentes estructuras secundarias fue obtenida mediante el uso de FIELDS (Piovesan et al., 2017). Finalmente se escribió un script de Python (Anexo I) para obtener los índices de plegamiento de las proteínas de interés mediante Fold Index (Prilusky et al., 2005). La estadística fue realizada mediante pruebas de ANOVA de un factor o pruebas no-paramétricas de Kruskal-Wallis, dependiendo de los supuestos cumplidos por los datos. El código utilizado para el procesamiento de las secuencias puede encontrarse en el Anexo I.

II.2.3 Búsqueda de secuencias conservadas mediante MEME suite.

La herramienta MEME de la suite de herramientas de MEME (Bailey & Elkan, 1994) fue utilizada para realizar búsquedas que hallaran cualquier número de repeticiones (*Any number of repeats*) de secuencias proteicas con longitudes variando entre 15 y 80 aminoácidos sobre la totalidad de las secuencias obtenidas luego del su procesamiento para un máximo de 10 secuencias conservadas (MEMEs). Al mismo tiempo, la misma suite incluye la herramienta *Motif alignment & search tool* (MAST) que permite comparar los motivos hallados entre sí y observar su distribución en las secuencias utilizadas para la búsqueda.

II.2.4 Alineamientos y construcción de árboles

Los alineamientos necesarios fueron realizados utilizando MAFFT (*Multiple alignments using Fast Fourier Transform*) de EMBL (Bateman et al., 2021), ClustalW (Madeira et al., 2022) o MUSCLE (Edgar, 2004) dependiendo de los objetivos de alineamiento. Se construyeron árboles a partir de las secuencias obtenidas utilizando el mejor modelo según MEGA (Tamura K, Stecher G, and Kumar S 2021) y se utilizaron los servicios de PhyML (Guindon et al., 2010), NGPhylogeny.fr (Lemoine et al., 2019) utilizando alineamientos MAFFT de las secuencias provistas y graficados utilizando iTOL (Letunic & Bork, 2021) mediante árboles Newick y los resultados de MEME.

II.2.5 Análisis manual de microsintenia

Luego de analizar los árboles, se seleccionaron especies a base a su diversidad de secuencias y taxonomía. Se obtuvo evidencia del contexto genético visualizando y analizando las tablas de genes de cada una de las accesiones de interés. Comparando las anotaciones de NCBI para estos genes se identificaron pares sinténicos.

II.3 Resultados

II.3.1 Construcción de las bases de secuencias curadas para las bases de datos de LEAs

Se realizó una búsqueda por alineamiento local (BLASTp) utilizando el conjunto de secuencias de Pfam usadas para construir los modelos ocultos de Markov correspondientes a cada grupo de LEAs en la base de datos del NCBI. Se analizaron las secuencias proteicas anotadas para los genomas de distintas especies representativas del clado Viridiplantae. Con el objetivo de filtrar las secuencias, se las sometió a una búsqueda contra bases de datos de motivos conservados utilizando el puntaje mínimo requerido por Pfam como umbral de detección para cada uno de los grupos de interés mediante la herramienta HMMscan. Esta metodología asegura que las secuencias contenidas pertenecen efectivamente a los grupos de interés con mayor confianza que elegir una línea de corte estadística arbitraria, que podría derivar en una base de secuencias conteniendo falsos positivos. Se removieron duplicados para evitar enriquecimiento de secuencias que pudieran alterar los resultados del análisis. Archivos FASTA para cada uno de estos pasos fueron almacenados para futura referencia. El FASTA final, conteniendo las proteínas de interés fue utilizado para todos los análisis descriptos a continuación.

II.3.2 Características físico-químicas de las LEAs de *C. quinoa*

Utilizando los *seeds* de cada grupo de LEAs disponible en Pfam se obtuvieron secuencias pertenecientes exclusivamente a la base de datos de proteínas de *C. quinoa* (taxid: 63459) de NCBI. Un árbol ilustrando las 43 accesiones de proteínas LEA identificadas en quinua y su distribución en grupos puede observarse en la Figura II.1. Las mismas secuencias utilizadas para ensamblar este árbol fueron utilizadas para realizar un análisis de los parámetros físico-químicos de cada proteína (Tabla II.1)

Al evaluar cada grupo y sus características físico-químicas (Tabla II.1) hallamos que el grupo de las LEA_2 es el que cuenta con mayor número de representantes (12) seguido por el de las SMPs (11). El grupo de las LEA_6 posee solo dos representantes. El tamaño de las proteínas es variable entre grupos abarcando un rango entre 79 (proteína más pequeña del grupo LEA_1) y 498 aminoácidos (proteína de mayor tamaño del grupo LEA_4), con pesos teóricos de 8,95 (LEA_5) a 53,52 kDa (LEA_4). Las proteínas de mayor peso y tamaño se encuentran en el grupo de las LEA_4.

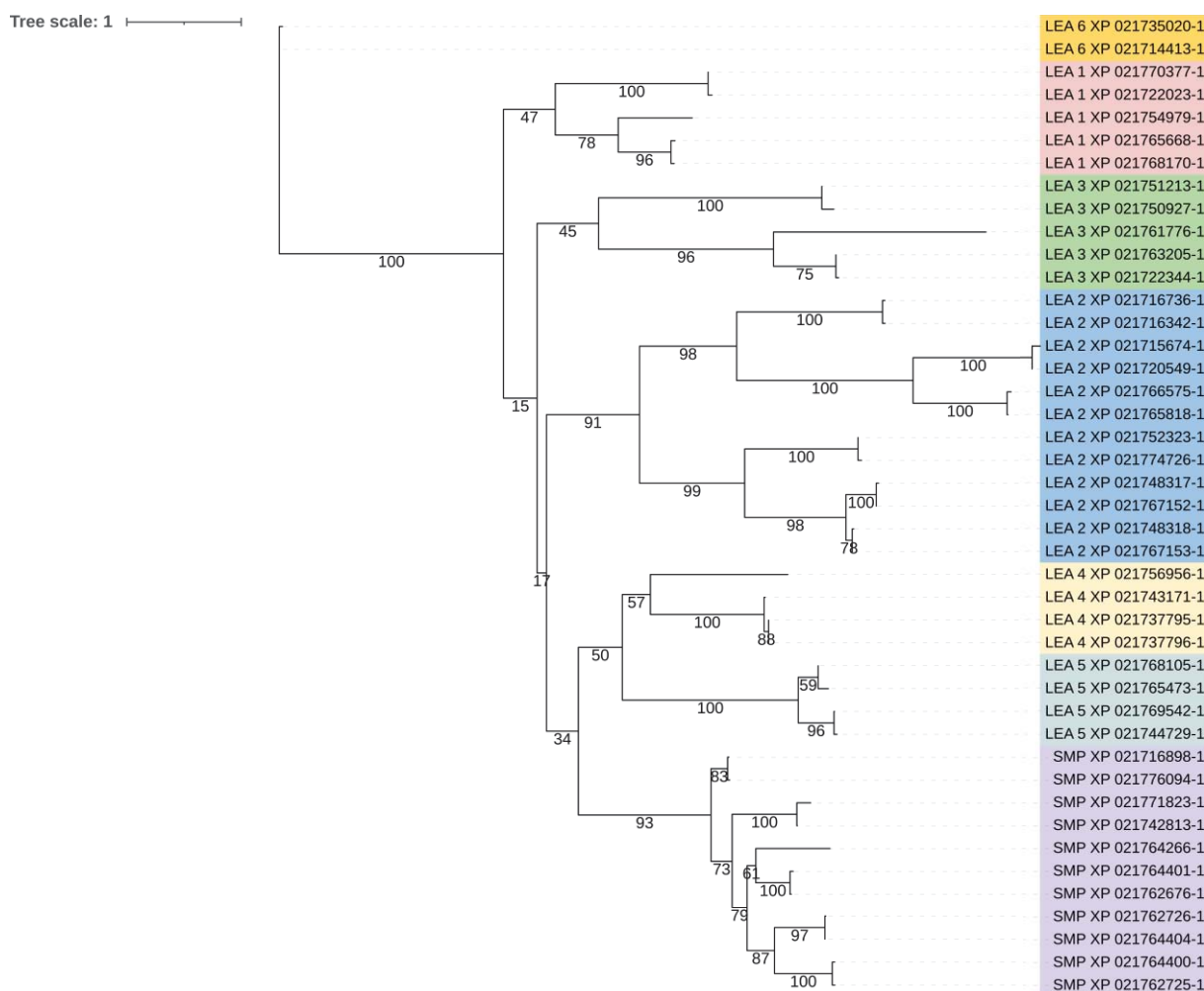


Figura II.1 Árbol construido a partir de todas las secuencias de proteínas LEA (grupos 1 a 6 y SMP) halladas en *C. quinoa*. Se utilizaron diferentes colores para agrupar LEAs de cada grupo. Árbol de máxima verosimilitud construido a partir de un alineamiento mediante MAFFT. Bootstrap $n = 100$.

Los grupos LEA del 1 al 4 están formados por proteínas que abarcan rangos de pI tanto ácidos como básicos, mientras que los grupos LEA_5, LEA_6 y SMP poseen pIs ácidos (Tabla II.1, Figura II.2). Al evaluar sus índices de hidropaticidad (GRAVY), se observa que todos los grupos presentan valores negativos. Las proteínas del grupo LEA_5 poseen los índices GRAVY más bajos, indicando que estas proteínas poseen una mayor afinidad por el agua que otros grupos, mientras que el grupo de las LEA_2 cuenta con el índice de mayor valor. Finalmente, los valores de índice de plegamiento (Fold Index) indican que solo los grupos LEA_2, SMP y algunas LEA_3 poseen un índice positivo, aunque pequeño, que difícilmente justifica la formación de una estructura globular. El resto de los grupos posee índices negativos, siendo los más bajos los de las proteínas LEA_5 y LEA_6, característica propia de las IDPs.

Graficando la distribución de los datos para cada grupo y sus parámetros físico-químicos (Figura II.2), se puede observar que el grupo de las LEA_3 posee una distribución bimodal del índice de todos los parámetros analizados (plegamiento, GRAVY, pI y masa teórica). Esto sugiere que dentro del grupo de las LEA_3 podrían existir dos subgrupos con características diferenciales en cuanto a su secuencia y posible estructura, ya que más allá de su tamaño, los índices de plegamiento e hidropaticidad alto de tres de las accesiones (XP_021761776.1, XP_021722344.1, XP_021763205.1) podría estar indicando que se trata de proteínas con mayor componente hidrofóbico que las demás.

El grupo de las LEA_4 posee un índice GRAVY y Fold Index homogéneo, con un puntaje promedio de -1,04 y -0,08 en promedio, respectivamente. Sin embargo, los pIs y las masas poseen distribución bimodal (Figura II.2), nuevamente sugiriendo, aunque con menor fuerza, la posibilidad de que existan subgrupos dentro de las LEA_4 de *C. quinoa*.

Tabla II.1 Resumen de características físico-químicas de las proteínas LEA halladas en *C. quinoa* y filtradas según el umbral de identificación por defecto de Pfam. Se presentan los datos ordenados por grupo. Se indica el número de secuencias encontradas y los máximos, mínimos y promedios de masa molecular (MW) en kDa, número de aminoácidos (Naa), punto isoeléctrico (pI) e índice GRAVY. Se agregó un gradiente de color rojo-blanco-verde en cada columna donde el rojo indica los valores más pequeños, blanco los intermedios y verde los mayores.

Grupo	N.º	MW (kDa)			Naa			pI			GRAVY		
		Min	X	Max	Min	X	Max	Min	X	Max	Min	X	Max
LEA_1	5	9.09	14.28	18.87	79	143	185	5.95	8.69	9.52	-1.18	-0.93	-0.79
LEA_2	12	17.88	24.17	36.17	159	217	324	4.73	6.80	9.54	-0.49	-0.20	0.00
LEA_3	5	9.97	12.89	17.33	93	113	148	5.88	8.25	9.89	-1.09	-0.65	-0.03
LEA_4	4	26.60	46.27	53.52	244	429	498	5.29	7.05	8.44	-1.10	-1.04	-1.01
LEA_5	4	8.95	10.72	12.43	84	99	114	5.77	5.94	6.16	-1.58	-1.46	-1.38
LEA_6	2	9.09	9.11	9.13	87	87	87	4.46	4.52	4.59	-1.18	-1.18	-1.17
SMP	11	15.07	23.87	30.69	146	230	290	4.25	4.90	6.33	-0.40	-0.32	-0.22

Los análisis de estructuras secundarias presentes en las secuencias de las LEA de quinua indican que la mayoría posee una gran componente de giros (*coils*), siendo también importante la proporción de posibles α -hélices (Figura II.3). El grupo de las LEA_2 es el que mayor representación de hojas plegadas- β posee.

Dentro de los grupos de interés, se destaca que LEA_3 poseen predicciones estructurales para α -hélice que apenas superan el 50% de los aminoácidos involucrados e inclusive poseen

una pequeña fracción vinculada con la estructura de hojas plegadas- β , pero aproximadamente el 50% de todas las proteínas analizadas estaría compuesto por coils, sugiriendo un mayor grado de desorden.

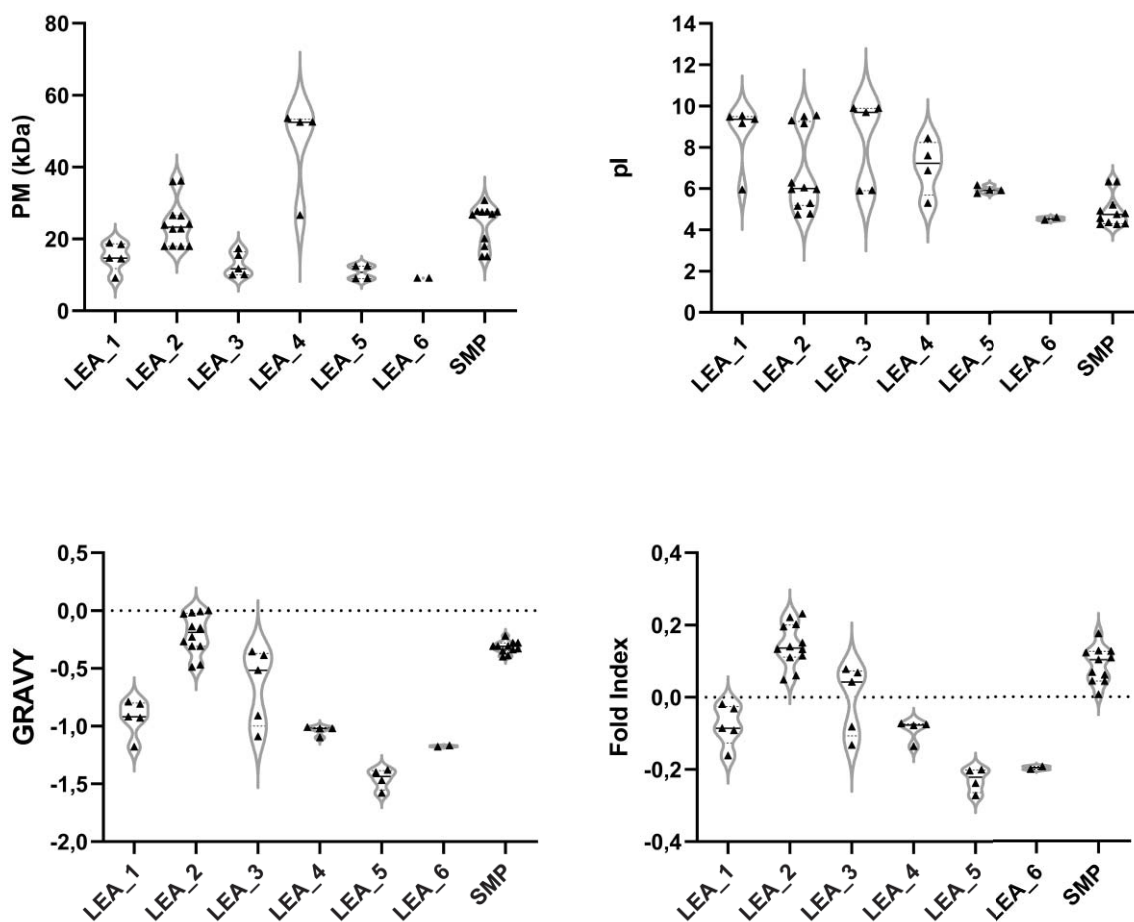


Figura II.2 Gráficos de violín de masa molecular (PM), punto isoeléctrico (pI), índice de hidrofobicidad (GRAVY) y de plegamiento (FoldIndex) para los grupos de LEAs de *C. quinoa* analizados. En cada caso se ilustró la distribución de los grupos para cada parámetro por separado.

Las proteínas correspondientes a las LEA_4, por su parte, poseen una gran cantidad de aminoácidos vinculados con estructuras secundarias de α -hélice. Esta información se ve respaldada por bibliografía que sugiere que, en determinadas condiciones, la proteína adquiere una estructura tridimensional que contempla α -hélices que en general no presenta cuando se encuentra en condiciones normales (Knox-Brown et al., 2020 y citas). En el caso de *C. quinoa*, las predicciones para esta estructura oscilan entre el 60 y 70% para todas las accesiones con la excepción de una accesión que alcanza el 90% de aminoácidos asociados con la misma.

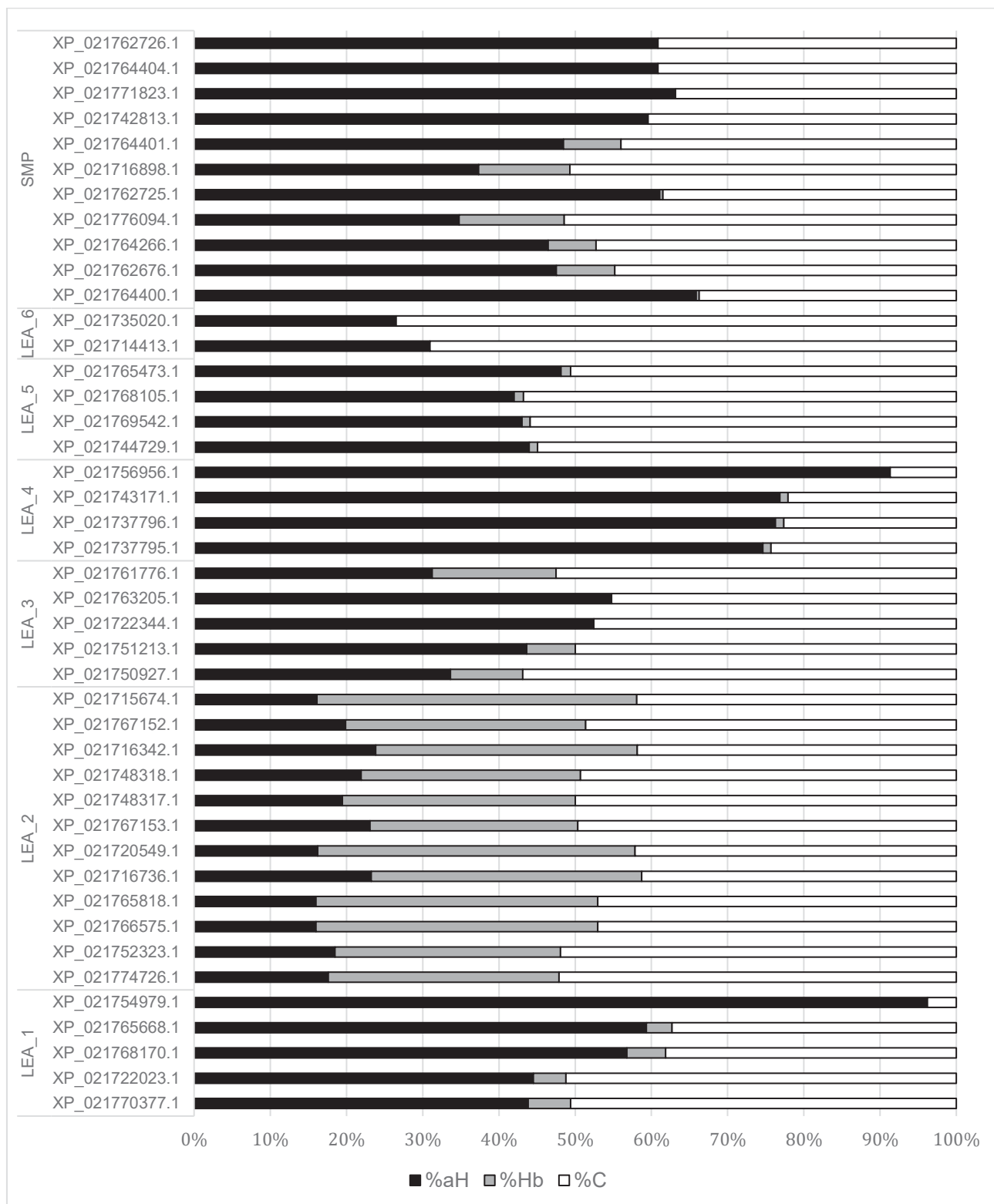


Figura II.3 Gráfico de barras indicando el porcentaje de aminoácidos asociados a diferentes estructuras secundarias de las proteínas LEA de *C. quinoa*. Las proteínas se clasificaron por grupo de LEA y se proveen las accesiones de cada una de ellas. Leyenda: %aH, porcentaje de α -hélice; %Hb, porcentaje de hoja plegada- β ; %C, porcentaje de coils (giros).

Las predicciones de desorden (Figura II.4) sitúan a los grupos LEA_5 y LEA_6 como los más propensos al desorden. Grupos como LEA_2, LEA_3 y SMPs presentan variabilidad en las predicciones de desorden, mientras que las LEA_4 y las LEA_1 parecen ser grupos más homogéneos. Las LEA_4 tienen una predicción de alrededor del 70% de desorden, y las LEA_3 varían entre un 10 y un 55% aproximadamente.

Vale la pena remarcar que se puede apreciar una relación entre los valores del GRAVY y las predicciones de desorden de las LEA. Los grupos LEA_5 y LEA_6, además de poseer los valores más altos de desorden, también poseen los valores más bajos de GRAVY. Lo opuesto es válido para las LEAs del grupo LEA_2, que poseen los menores valores de desorden y los mayores valores de GRAVY.

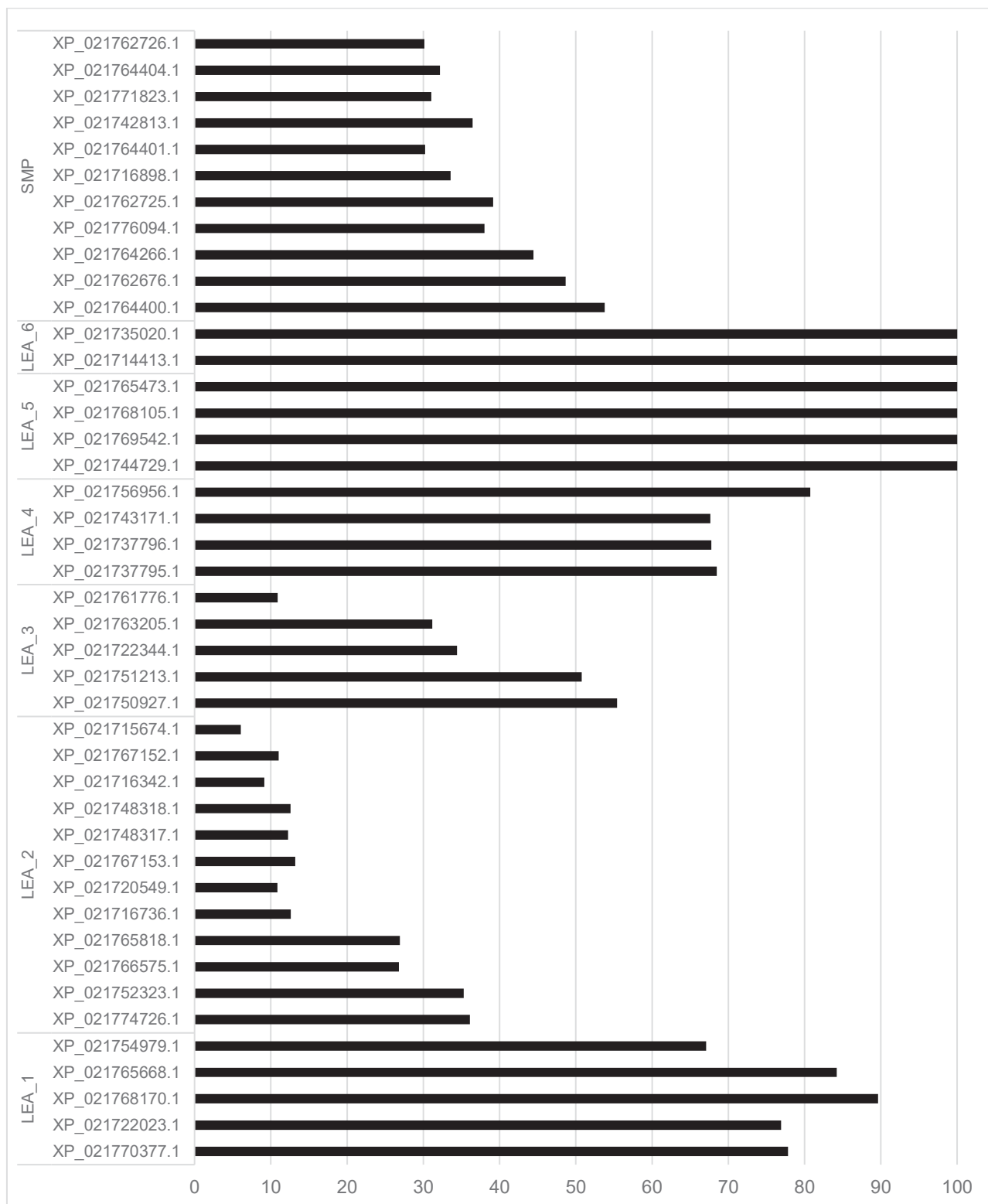


Figura II.4 Porcentaje de aminoácidos vinculados a estructuras desordenadas en las proteínas LEA de *C. quinoa*. Las proteínas se clasificaron por grupo y se proveen las accesiones de cada una de ellas.

II.3.3 Análisis de las LEA_3

II.3.3.1 Búsqueda de motivos conservados en las LEA_3

Utilizando la metodología descrita para la búsqueda de secuencias LEA en *C. quinoa* Willd. Se obtuvieron secuencias correspondientes a LEAs del grupo LEA_3 de distintas especies de plantas representativas del clado Viridiplantae, con un foco en Angiospermas. Estas secuencias fueron sometidas a una búsqueda de motivos conservados utilizando el software MEME suite (del inglés, *Multiple Em for Motif Elicitation*). El software MEME Suite es una colección de herramientas y algoritmos utilizados para el descubrimiento y análisis de motivos de secuencia en ADN, ARN y proteínas. Se basa en métodos estadísticos y de aprendizaje automático y proporciona información sobre la posición, frecuencia y características específicas de los motivos conservados. Cada uno de los motivos que entrega el MEME Suite se denomina "motivo MEME". Estos motivos son de gran importancia en la investigación biológica, ya que pueden revelar información sobre la estructura, función y regulación de las moléculas biológicas, ya que pueden ser indicativos de regiones funcionales o reguladoras. La Figura II.5 A presenta un resumen del resultado de este análisis.

Las primeras observaciones indican que de los 10 MEMEs buscados solo cuatro (1, 2, 3 y 5) poseen una distribución que abarca varios grupos taxonómicos de interés sugiriendo patrones conservados a lo largo de la evolución de las plantas. El resto de los MEMEs posee distribuciones limitadas a un solo grupo taxonómico o inclusive a una sola especie.

De las 229 secuencias identificadas en las especies seleccionadas para el estudio, 212 poseen el MEME 1 (WVPDPVTGYYRPENRANEIDAAELRAMLL), algunas de ellas (un total de 5) parecen poseer dos repeticiones de este motivo, aunque lo normal es encontrar este MEME una vez por secuencia, al igual que para el resto de los MEMEs hallados en este grupo. El MEME 1 contiene una de las regiones más comunes y conservadas que se utiliza para identificar proteínas que pertenecen al grupo LEA_3 y que en este trabajo llamamos **W** (Figura II.5 A y B). Unas pocas secuencias presentan en lugar de este MEME el MEME 5 (WVPHPRRTGIYFPEGHEWVMDDVPEGAASL), cuya similitud con el MEME 1 es evidente, pero muestra diferencias en la conservación de residuos en toda su extensión y al que denominamos **W'** (Figura II.5 A y B).

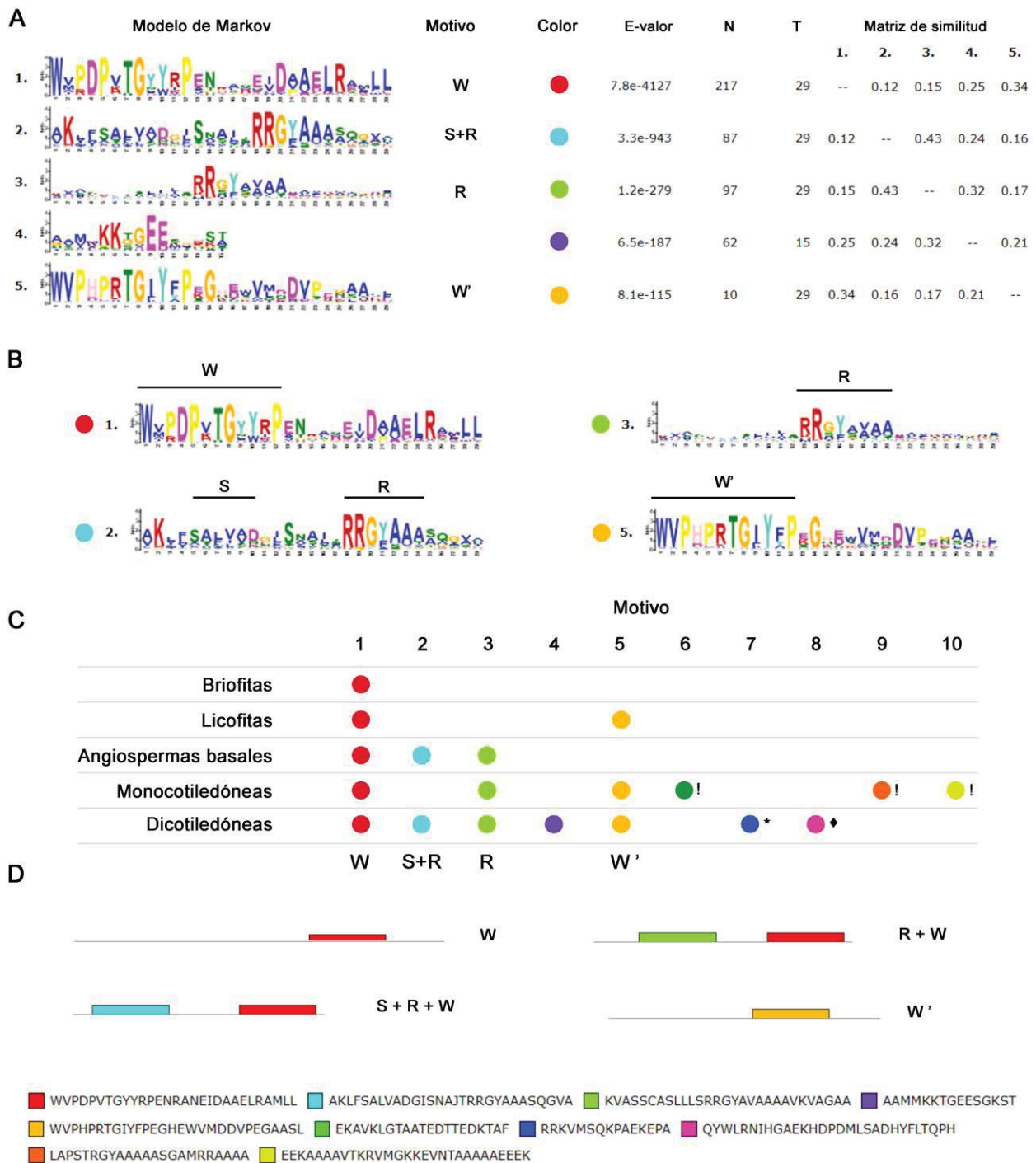


Figura II.5. Resultados del análisis de las secuencias conservadas en LEA_3. A. Resultado de MEME y MAST mostrando: modelo de Markov para cada MEME hallado, nombre asignado, color asignado, E-valor, número de ocurrencias (N), tamaño del MEME (T) y matriz de similitud comparando los MEMES de interés. B. Regiones conservadas identificadas dentro de los MEMES. C. Distribución de los MEMES hallados en los grupos taxonómicos representados en el análisis. Signos de exclamación (!), indican que los motivos son exclusivos de Poáceas. Asterisco (*), indica que el motivo fue hallado exclusivamente en *Kalanchoe laxiflora*. Diamante (♦) Indicando que el motivo pertenece exclusivamente a Caryophyllales. D. Arquitecturas halladas en el análisis MEME y MAST. La leyenda de colores en la parte inferior corresponde a los colores utilizados en las figuras A, C y D.

También se puede apreciar que la región C-terminal del MEME 2 es idéntica a la secuencia observada en el MEME 3, siendo la diferencia entre ambas, una secuencia conservada en el MEME 2 que llamaremos **S**, por su secuencia conservada (SALVAD) que se indica en la Figura II.5 B. La región que poseen en común (RRGYAAA), que nombramos **R**, también es una región muy conservada dentro de las LEA_3 de angiospermas. Es aparente que las secuencias conservadas S y R se encuentran enriquecidas diferencialmente en el conjunto de secuencias analizadas. El resto de los MEMEs no compartía similitudes o posiciones que indicaran una superposición que precisara removerles del análisis.

II.3.3.2 Distribución taxonómica de los motivos hallados

Analizando la distribución de los motivos en las especies estudiadas (Figura II.5 C) detectamos que el motivo W (MEME 1, rojo) es común a todos los grupos taxonómicos involucrados en el análisis. En *S. moellendorffii* (Lycophyta), se da la aparición del motivo W' en varias secuencias. Las angiospermas basales, *A. trichopoda* y *N. colorata*, muestran la presencia de los motivos S y R (MEMEs 2, celeste y 3, verde). En monocotiledóneas, podemos observar la presencia del motivo W, R y W', aunque el análisis por MAST no indica que este grupo contenga el MEME 2 (correspondiente con los motivos S y R en conjunto), es posible que entre ellas se encuentren secuencias que poseen un proto-motivo S. Finalmente las dicotiledóneas presentan todos los motivos de interés (W, S, R y W') distribuidos entre las secuencias estudiadas. EL MEME 7 (verde oscuro) fue hallado exclusivamente en *Kalanchoe laxiflora*, y el MEME 8 (rosa) corresponde a una región N-terminal hallada únicamente en *C. quinoa* y *B. vulgaris*, ambas pertenecientes a la familia Amaranthaceae. .

Las arquitecturas observadas se pueden reducir a 3 grupos de interés (Figura II.5 D) definidos por la combinación de motivos presentes: S, R, W y/o W'. Cabe aclarar que las secuencias que presentan motivos W' carecen de otros motivos característicos. Si bien parecía haber una excepción a esta regla en dos secuencias de *Selaginella moellendorffii*, donde MEME indica la presencia del motivo R, la observación de las secuencias individuales no permite distinguir que se trate de una identificación confiable del motivo y los e-valores correspondientes a los alineamientos son los de los más bajos en comparación al resto de las secuencias. Aquellas secuencias que presentan el motivo W pueden carecer o no del motivo R. Por otra parte, las secuencias que presentan el motivo S, siempre poseen los motivo R y W. Finalmente, las secuencias que presentan el motivo W' no responden a un ordenamiento taxonómico, habiendo sido encontradas en licofitas, monocotiledóneas y dicotiledóneas.

II.3.3.3 Análisis filogenético y microsinténico

La sintenia se refiere a la conservación de la organización y disposición relativa de genes o secuencias genéticas a lo largo de diferentes especies o linajes evolutivos (los llamaremos marcadores genéticos). Especifica la co-localización física de genes o secuencias en el mismo cromosoma o en regiones cromosómicas cercanas. La sintenia puede observarse en diferentes niveles, desde pequeñas regiones genómicas (microsintenia) hasta genomas completos. La conservación de la sintenia entre especies sugiere que las secuencias o genes involucrados tienen una función importante o están sujetos a restricciones evolutivas que mantienen su organización espacial relativa. El estudio de la sintenia proporciona información valiosa sobre la evolución genómica y la relación funcional entre genes y secuencias en diferentes organismos. La naturaleza modular y repetitiva de las proteínas LEAs acota las conclusiones que puede derivarse de la construcción de un árbol filogenético debido a que los alineamientos de secuencia por homología no son extensos. En este caso complementar el estudio con un análisis de microsintenia refuerza las conclusiones que se puedan obtener en referencia a la historia evolutiva de las LEAs.

Al estudiar comparativamente la totalidad de las secuencias analizadas mediante la construcción de árboles, detectamos a grandes rasgos tres subgrupos con motivos diferentes dentro del grupo de las LEA_3 (Figura II.6). En particular, aquellas secuencias que presentan el MEME 3 (que contiene el motivo R, pero no el S), tienen mayor tendencia a agruparse que aquellas que sí presentan el motivo S (Figura II.6 MEME 2, celeste). Las secuencias con el motivo W' forman dos grupos, uno incluyendo únicamente dos secuencias de *S. moellendorffii* y otro conteniendo el resto de las secuencias y más alejado del resto (Figura II.6.; MEME 1, rojo).

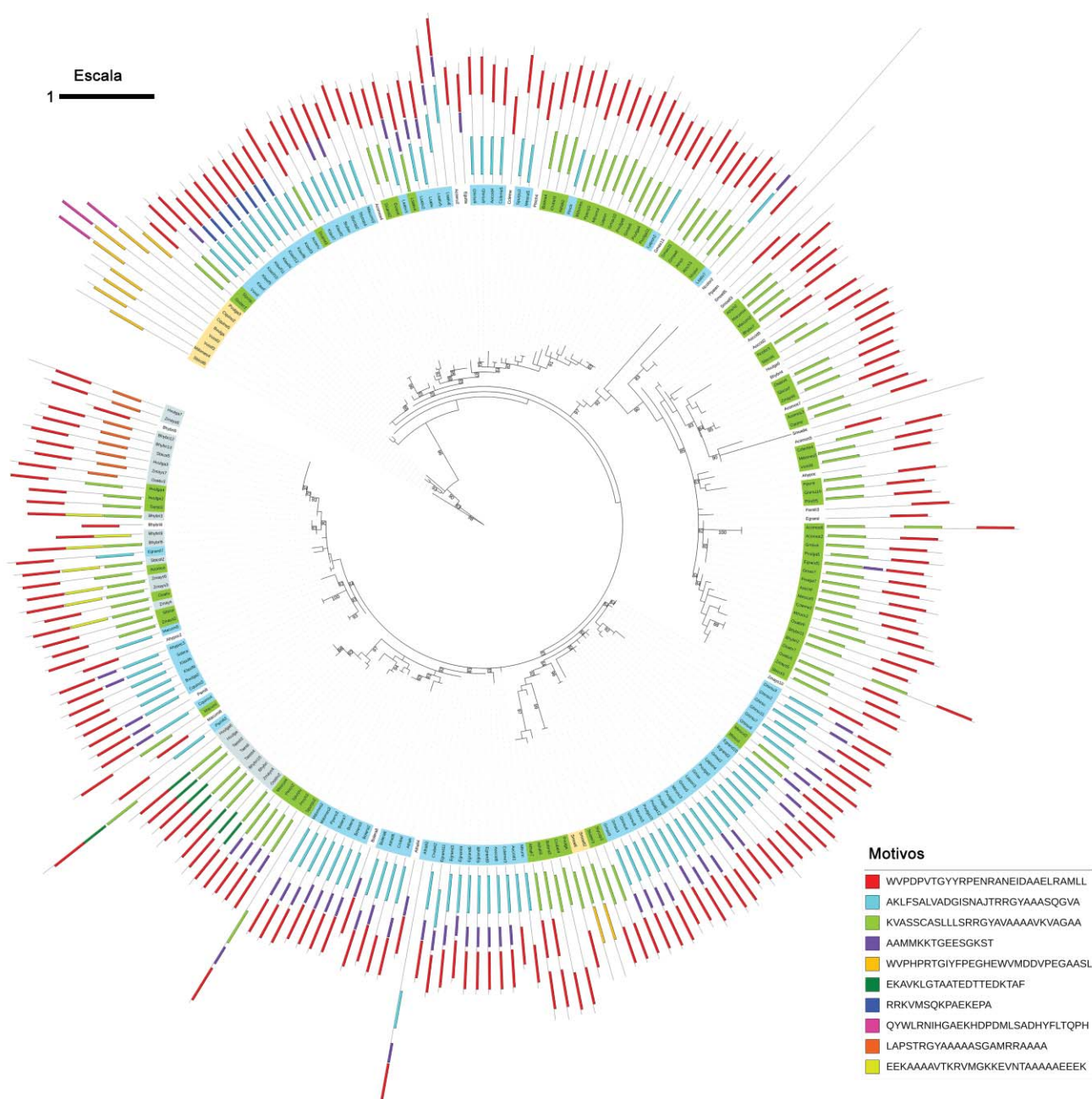


Figura II.6 Árbol de la totalidad de las secuencias halladas y curadas mediante la búsqueda con las secuencias *seed* que componen el HMM de LEA_3 según Pfam y su análisis mediante MEME-suite. Se incluyen los resultados gráficos de los motivos hallados por MEME-suite junto a su accesión correspondiente. Los taxones de las secuencias con arquitecturas similares fueron coloreados. La leyenda presenta, en orden, los MEMEs con su color representativo en las arquitecturas presentadas en el árbol. Valores estadísticos de cada rama utilizando aLRT se presentan en las ramas. La escala de distancia del árbol se encuentra en la parte superior izquierda. La leyenda para las secuencias utilizadas puede encontrarse en el Anexo III.1.

Se separaron las secuencias en aquellas que presentan los motivos R y W (grupo RW), aquellas que presentan el motivo S (SRW) y aquellas que presentan solamente el motivo W' y se estudiaron los parámetros fisicoquímicos de todos los grupos en busca de posibles características diferenciales entre los mismos (Figura II.7) que justificaran las observaciones iniciales realizadas sobre las LEA_3 de *C. quinoa*.

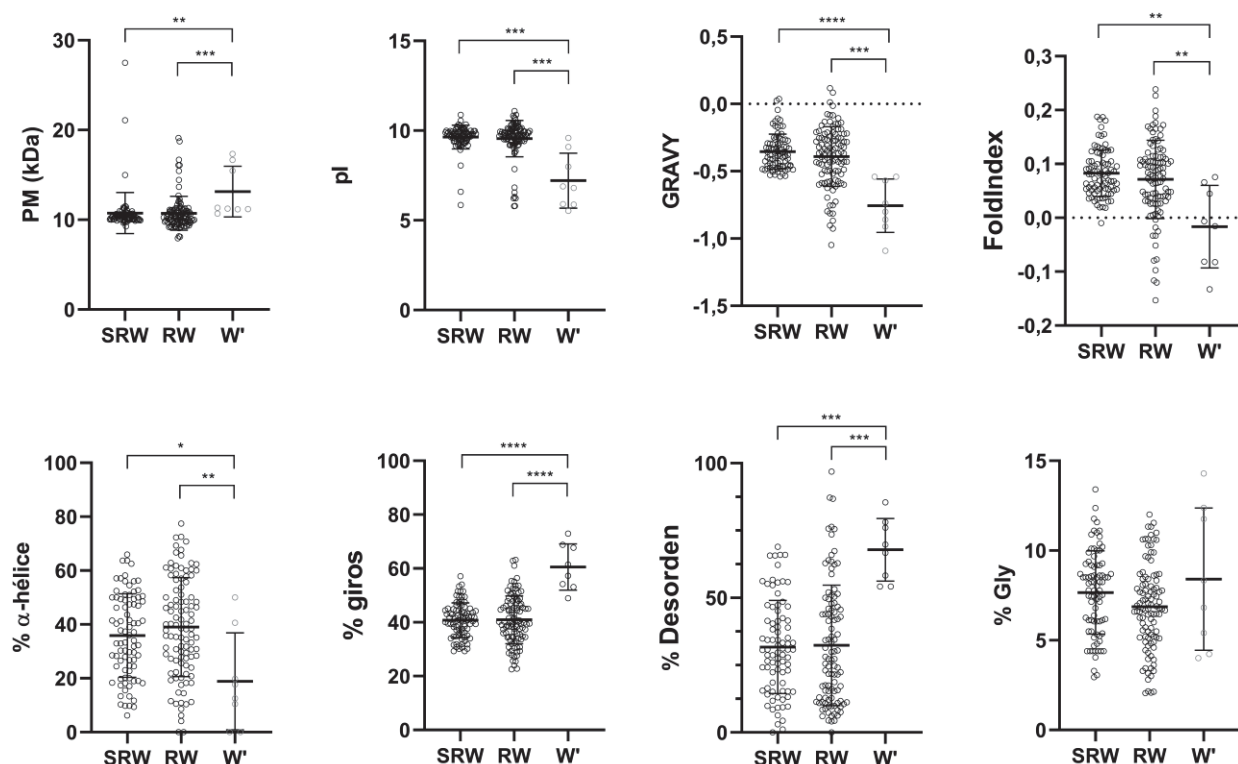


Figura II.7. Características fisicoquímicas de los subgrupos de LEA_3 discriminados Gráficos para peso teórico (PM), punto isoelectrico (pl), índice de hidrofobicidad (GRAVY), índice de plegamiento (FoldIndex), porcentaje de aminoácidos asociados a α-hélices, giros y desorden y porcentaje de glicinas en las secuencias. Cada punto representa una de las proteínas utilizadas para el análisis, la raya horizontal en cada columna representa el promedio. Las barras de error representan el desvío estándar (SD). *, $p < 0,05$; **, $p < 0,01$, ***, $p < 0,001$, ****, $p < 0,0001$.

Los datos indican inicialmente que no existe una diferencia significativa en las características de los grupos SRW y RW. Sin embargo, estos difieren significativamente de las secuencias que presentan el motivo W' en varios parámetros. Aunque con pocos representantes, las secuencias que presentan el motivo W' son de tamaño significativamente mayor que las de los otros grupos, presentan un punto isoelectrico menor ($X = 7,21$) y un índice de hidrofobicidad (GRAVY) menor, que sugiere una proteína más soluble. Al mismo tiempo el índice de plegamiento es menor que el de los otros grupos, lo que sugiere que se trata de proteínas de estructura más desordenada que los grupos SRW y RW. Reforzando

este resultado, los análisis por medio de FIELDS, que predicen el porcentaje de aminoácidos vinculados con posibles estructuras secundarias, indican que el grupo W' posee en promedio una cantidad significativamente menor de aminoácidos vinculados a α -hélices, y una mayor cantidad de aminoácidos vinculados a *coils* o giros, dando en consecuencia una proteína con un mayor grado de desorden. Finalmente, al evaluar la presencia de glicina, un aminoácido común en las hidrofílicas como las LEAs, no se detectaron diferencias significativas.

La información obtenida, en conjunto, sugiere la posibilidad de que existen al menos dos subgrupos dentro de las LEA_3 que, si bien pueden ser identificados mediante los criterios de Pfam, presentan regiones conservadas diferentes en su región N-terminal (R o SR), y un tercer grupo con motivo canónico de las LEA_3 (W) ligeramente modificado y sin ningún otro motivo característico acompañándole, cuyo origen evolutivo es desconocido hasta el momento.

Una selección de especies abarcando angiospermas basales, monocotiledóneas y dicotiledóneas que incluyen secuencias de todos los subgrupos identificados (SRW, RW y W') fue utilizada para realizar un análisis microsintético con el objetivo de identificar posibles orígenes evolutivos de cada uno de los subgrupos. Estudiando marcadores conservados en el contexto genómico de cada una de las secuencias se obtuvo el resultado observado en la Figura II.8.

Uno de los primeros hallazgos surgió de una comparación realizada entre las anotaciones presentes en NCBI y Phytozome para las LEA_3 de *Amborella trichopoda*. Se trata de una duplicación en tándem que figura en las bases de Phytozome, pero no en NCBI. Inicialmente se sospechó que tal duplicación podría tratarse de un error de anotación. Esta hipótesis fue rápidamente descartada cuando se observaron otros eventos de duplicación en otras especies analizadas (Figura II.8 A) tanto en NCBI como en Phytozome. Aún más, aquellas LEAs que se hallaron en tándem compartían marcadores de sintenia con otras donde ocurría el mismo evento. Esto dio indicios de que se trata de un grupo ortólogo que llamaremos LEA_3A (SRW), donde la duplicación ocurrió en un ancestro común previo a las angiospermas basales. En *N. colorata*, dos de las secuencias identificadas presentan en la región N-terminal una secuencia que guarda un alto grado de similitud con el motivo S (SAASEL y SAVSDL). Esta observación se repite en *Medicago truncatula*, donde dos de las secuencias identificadas y que poseen marcadores en común con este grupo, presentan un proto-motivo similar al S (SAVVLE). Cabe remarcar que una tercera secuencia de *M. truncatula* presenta el conjunto de motivos SRW, pero que no presenta marcadores que hayan sido encontrados en *A. trichopoda*, aunque sí presente algunos en común con las observadas en *Sorghum bicolor* (Sbicol y Sbicol2). Dos de las secuencias LEAs identificadas en *C. quinoa* corresponden a

este grupo (LEA_3A) según los marcadores presentes, pero no se detectaron duplicaciones en tándem.

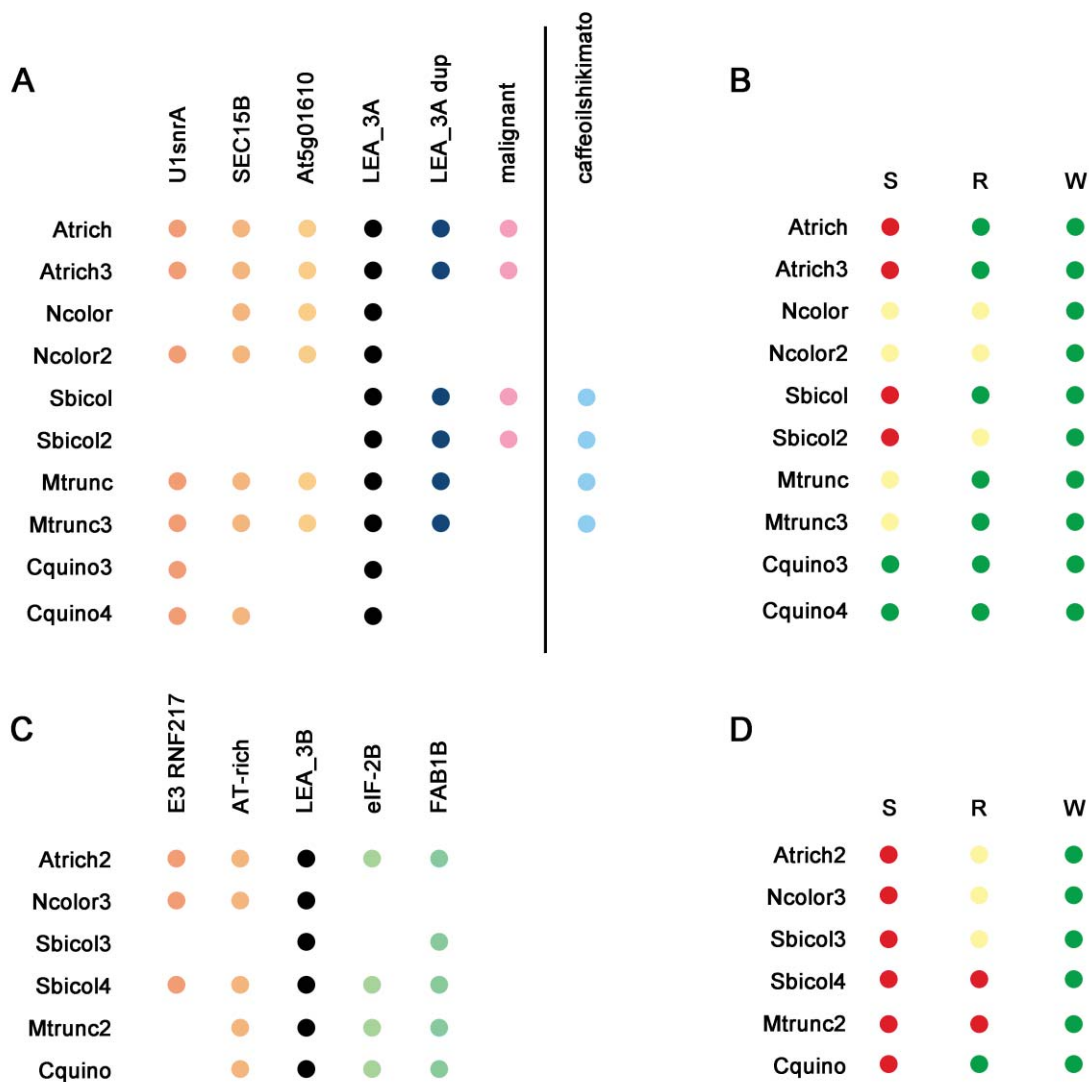


Figura II.8. Marcadores microsintéticos conservados de secuencias seleccionadas de LEA_3. **A** y **C**. Conservación de marcadores microsintéticos para los grupos ortólogos hallados. Se utiliza el orden de marcadores observado en *Amborella trichopoda*, aunque el mismo puede variar en otras especies. **B** y **D**. Identificación de motivos conservados S, R y W para las secuencias observadas en los gráficos A y C, respectivamente. Se indica presencia del motivo con verde, proto-motivos con amarillo y ausencias con rojo. U1snrA, U1 small nuclear ribonucleoprotein A; SEC15B, exocyst complex component SEC15B; At5g01610, uncharacterized protein At5g01610-like; LEA_3A, late embryogenesis abundant protein; LEA_3A dup, duplicación en tandem de LEA_3A; malignant, malignant T-cell-amplified sequence 1 homolog; caffeoilshikimato, caffeoylshikimate esterase; E3 RNF217, E3 ubiquitin-protein ligase RNF217; AT-rich, AT-rich interactive domain-containing protein 4; eIF-2B, translation initiation factor eIF-2B subunit alpha; FAB1B, 1-phosphatidylinositol-3-phosphate 5-kinase FAB1B. La leyenda para las secuencias de LEA_3 utilizadas puede encontrarse en el Anexo III.1.

El segundo conjunto de secuencias muestra las relaciones entre las secuencias que presentan los motivos R y W. Los marcadores observados permiten asumir que existe un segundo grupo de ortólogos que llamaremos LEA_3B (RW). Ninguna de las secuencias que se correspondieron con este grupo de ortólogos presentó la secuencia conservada S. Una de las secuencias identificadas en quinua corresponde a este grupo ortólogo.

Finalmente, pero no menos importante, aquellas secuencias que mostraron la presencia del motivo W' no poseen marcadores en común con ninguno de los otros grupos ortólogos, y la evidencia de marcadores que las vincula entre sí es débil. Para este conjunto que llamamos LEA_3C (W'), una de las especies evaluadas, *Vitis vinifera*, posee una duplicación en tándem de las dos accesiones que posee con este motivo. Esta evidencia sugiere que la duplicación en tándem puede no estar limitada a las LEA_3A, sino que pudo haber ocurrido antes de que los grupos A y C se diferenciaron. Dos de las secuencias identificadas en quinua corresponden a este grupo ortólogo.

II.3.4 Análisis de las LEA_4

II.3.4.1 Búsqueda de motivos conservados en las LEA_4 y su distribución

Utilizando las mismas metodologías que para LEA_3 se analizaron las secuencias obtenidas y curadas de LEA_4. El análisis de MEME arroja los resultados visibles en la Figura II.9. El motivo de mayor representación (MEME 1, rojo) corresponde con el motivo canónico de identificación de las LEAs de este grupo. Al mismo tiempo, los análisis de MAST indican que los MEMEs 3 y 7 tienen un alto grado de similitud con el MEME 1. Estos MEMEs se distribuyen homogéneamente entre todas las secuencias estudiadas, por lo que se los considera comunes a todas las secuencias LEA_4 y suelen ocupar la región central de la secuencia ordenándose en bloques repetitivos.

Entre los motivos restantes, se destacan motivos N-terminales y C-terminales que parecen variar entre las LEA_4. Los MEMEs 5 y/o 9 suelen formar parte de la región N-terminal de las secuencias estudiadas, a veces acompañados por el MEME 2. El MEME 5 se presenta por primera vez en angiospermas basales y se lo halla tanto en monocotiledóneas como dicotiledóneas, mientras que el MEME 9, que comparte una similitud del 45% con el MEME 6, se presenta únicamente en monocotiledóneas y dicotiledóneas (Figura II.9 C).

Otro conjunto de secuencias no presenta los MEMEs 5 y 9, y en su lugar presentan el MEME 8, con una posición más alejada del extremo. Este se presenta por primera vez en

Physcomitrella patens, representante de las briófitas, aunque su posición difiere de la que se observa en las secuencias de monocotiledóneas y dicotiledóneas donde también fue detectado.

En cuanto a la región C-terminal, hay dos MEMEs que llaman la atención. El MEME 10, por un lado, fue encontrado exclusivamente en un grupo de secuencias pertenecientes a las monocotiledóneas. Otro conjunto de secuencias, que incluyen especies monocotiledóneas y dicotiledóneas, presenta en el extremo C-terminal el MEME 4, a veces por duplicado. Finalmente, las secuencias que poseen el MEME 8 en su región N-terminal no presentan ninguna secuencia conservada en su región C-terminal.



Figura II.9 Resumen de resultados de MEME y MAST para las secuencias de LEA_4. **A.** Resultados de MEME y MAST informando Modelo de Markov, color asignado, E-valor, número de sitios hallados

(N), tamaño del MEME en aminoácidos (T) y matriz de similitud entre todos los MEMEs hallados ilustrando el grado de similitud con valores entre 0 y 1. EL resaltado en rojo indica un valor superior a 0,6 en la matriz de distancia. **B.** Alineamiento de los MEMEs 1, 3 y 7 (indicados con puntos rojos por su código de color) contra las secuencias *seed* de Pfam para las LEA_4, verificando similitud a la secuencia canónica de identificación. **C.** Resumen de la presencia de los MEMEs hallados en los diferentes grupos que forman parte del análisis. Se utilizó el mismo código de colores y numeración que en (A). Asterisco (*) indica que el motivo está presente con bajo valor estadístico. Leyenda con la secuencia más frecuente de cada MEME con su código de color en I aparte inferior.

II.3.4.2 Análisis filogenético, físico-químico y microsinéptico de las LEA_4

Un árbol de máxima verosimilitud fue realizado luego de alinear las secuencias utilizando MAFFT. Se observan a grandes rasgos tres grupos de secuencias claramente diferenciados por la presencia de determinados motivos y arquitecturas descriptas en la sección II.3.4.1 (Figura II.9, Figura II.10).

Se reconocen 3 grupos con arquitecturas definidas. El resultado del árbol, presenta algunas relaciones taxonómicas que obedecen a la clasificación filogenética según APG IV. Especies pertenecientes a las Caryophyllales (*Spinacia oleracea*, *Beta vulgaris* y *Amaranthus hypochondriacus*), se agrupan. Al mismo tiempo, la rama que las contiene surge de una rama conteniendo LEAs con una arquitectura N-terminal similar y especies dicotiledóneas. También aparecen agrupadas en el árbol secuencias correspondientes a la familia Fabaceae (*Glycine max*, *Lotus japonicus*, *Phaseolus vulgaris* y *Medicago truncatula*). Finalmente, aquellas secuencias de monocotiledóneas en las que se había observado la presencia C-terminal del MEME 10, también se encuentran agrupadas.

La distribución de las especies y las arquitecturas observadas sugiere que podrían existir tres subgrupos de LEA_4 que correlacionan con la presencia de los MEMEs 4, 8 y 10. En particular las secuencias que poseen el MEME 4 (violeta) en su región C-terminal, las que poseen el MEME 8 (fucsia) en su región N-terminal y, por último, las secuencias que poseen el MEME 10 (lima) en su región C-terminal. Estas secuencias fueron seleccionadas para realizar un análisis de sus características físico-químicas con énfasis en aquellas propiedades relacionadas a las proteínas IDP e hidrofílicas.

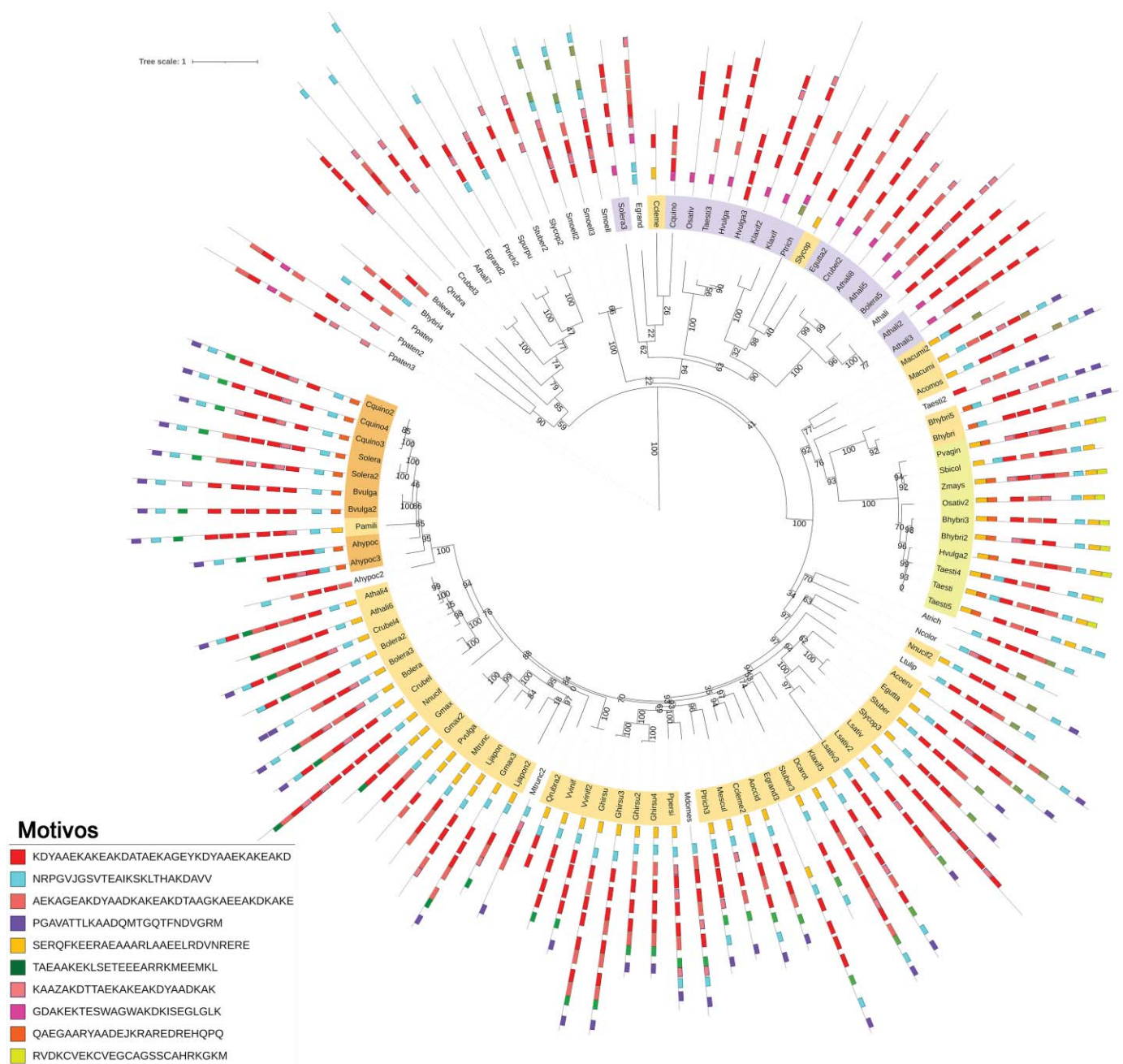


Figura II.10 Árbol de la totalidad de las secuencias halladas y curadas mediante la búsqueda con las secuencias *seed* que componen el HMM de LEA_4 según Pfam y su análisis mediante MEME-suite. Se incluyen los resultados gráficos de los motivos hallados por MEME-suite junto a su accesión correspondiente. Los taxones de las secuencias con arquitecturas similares fueron coloreados. Los MEMEs 13 y 7 fueron coloreados con tonalidades de rojo dada su similitud. La escala de distancia para el árbol se encuentra en la parte superior izquierda. La leyenda para las secuencias utilizadas puede encontrarse en el Anexo IV.2.

Se realizó una comparación de sus características físico-químicas con el objetivo de identificar posibles características que resalten sus similitudes y diferencias, obteniendo los resultados observados en la Figura II.9.

Los resultados indican que estos tres grupos difieren en diversas características. Las proteínas que forman parte del grupo V (MEME 4, violeta) son proteínas que al compararlas con el grupo L (MEME 10, lima) muestran ser de mayor tamaño, menor punto isoeléctrico, con índices de plegamiento e hidropaticidad más bajos, que indican una mayor probabilidad de que se trate de proteínas desordenadas. Esta observación se ve respaldada también por la presencia de mayor número de aminoácidos vinculados con giros y la menor tendencia a α -hélices. Las proteínas del grupo F (MEME 8, fucsia) no difieren en peso y punto isoeléctrico de las que poseen motivo violeta (MEME 4), aunque sus índices de hidropaticidad y plegamiento, así como su contenido de α -hélices, sugiere que se trata de un conjunto de proteínas que podría presentar estructuras secundarias con mayor probabilidad. Un dato que llamó la atención, dentro del grupo de MEME 8, es que algunas de sus integrantes mostraron un pI básico. Al buscar a que proteínas corresponden estos valores se encuentran 3 secuencias de monocotiledóneas (*Hordeum vulgare* XP_044982732.1, *Triticum aestivum* XP_044376522.1 y *Oryza sativa* XP_015628236.1) y una de *Eucalyptus grandis* (XP_010024874.2). El resto de las secuencias, con un pI ácido, corresponden únicamente a dicotiledóneas. Este fue el único caso donde se notó una distribución notoriamente asimétrica de los datos físico-químicos. Cabe remarcar que estas secuencias no muestran diferencias en otros parámetros.

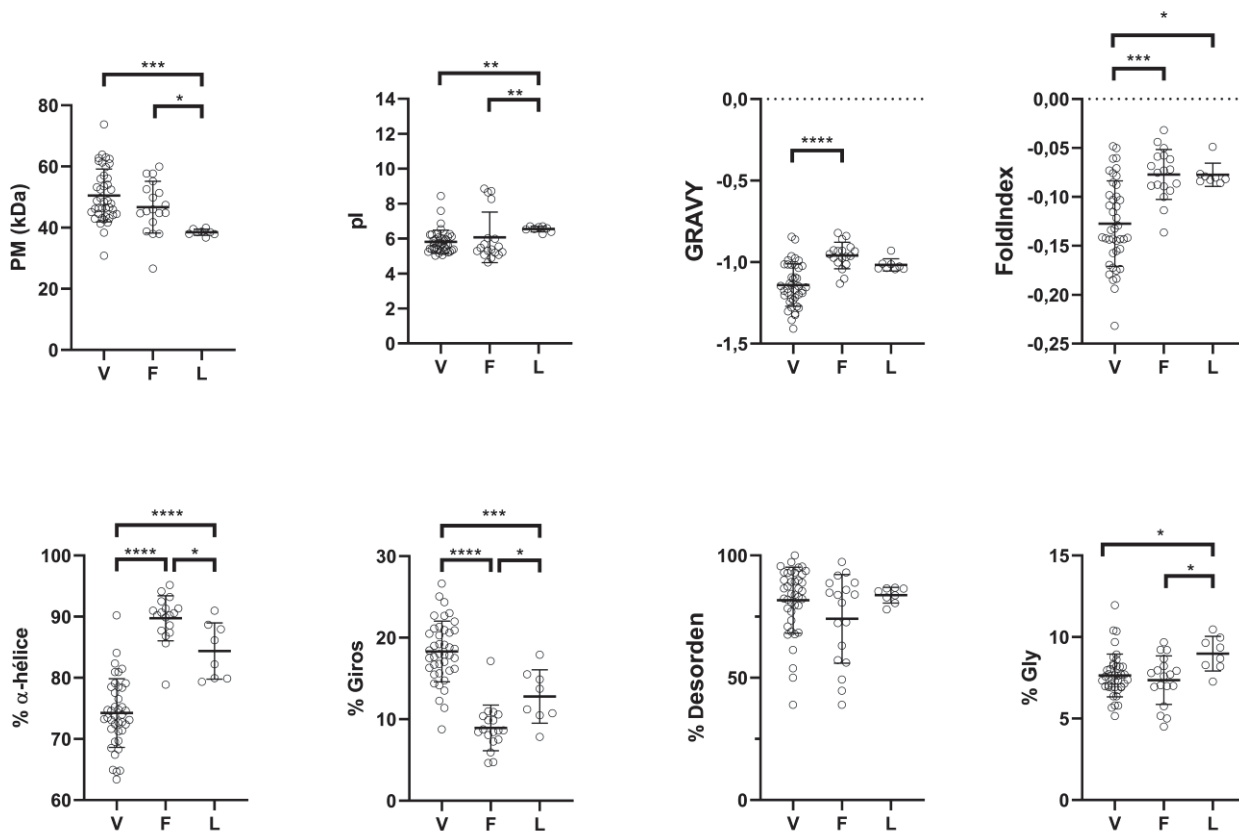


Figura II.9 Características fisicoquímicas de los subgrupos de LEA_4 discriminados. Se provee información sobre peso molecular (PM), punto isoeléctrico (pI), índice de hidropaticidad (GRAVY), índice de plegamiento (Fold Index), porcentaje de aminoácidos que pueden conformar α -hélices, giros y estructuras desordenadas y porcentaje de glicina, en orden. V, secuencias con motivo violeta (MEME 4); F, secuencias con motivo fucsia (MEME 8); L, secuencias con motivo lima (MEME 10). Cada punto es un valor individual. La marca central indica promedio. Barras de error: desvío estándar. *, $p < 0,05$; **, $p < 0,01$, ***, $p < 0,001$, ****, $p < 0,0001$.

Por su parte, las proteínas que poseen el MEME 10 poseen pesos moleculares, pI, GRAVY y Fold Index homogéneos. Son proteínas en promedio más pequeñas y cuyas propiedades sugieren una mayor probabilidad de estructuras secundarias que el resto de los grupos.

Dadas las diferencias observadas entre los grupos formados se decidió profundizar el análisis mediante un estudio microsintético tomando varias especies de referencia que presentaran secuencias en varios subgrupos, con el objetivo de identificar posibles ortólogos.

Los resultados indican que dentro de las LEAs_4 se identifican secuencias pertenecientes a dos grupos ortólogos bien definidos (Figura II.10). El primero (Figura II.10 A)

está conformado por aquellas proteínas que poseen el MEME 4 (violeta) en su posición C-terminal y que llamativamente todas pertenecen al grupo de las dicotiledóneas. Ninguna de las monocotiledóneas analizadas presentó marcadores que las vincularan con este grupo ortólogo. Por otra parte, un segundo grupo ortólogo (Figura II.10 B) comprende únicamente secuencias de monocotiledóneas que poseen el MEME 10 (lima) en su región C-terminal. Ninguna de las dicotiledóneas analizadas presentó marcadores que las vincularan con este grupo ortólogo.

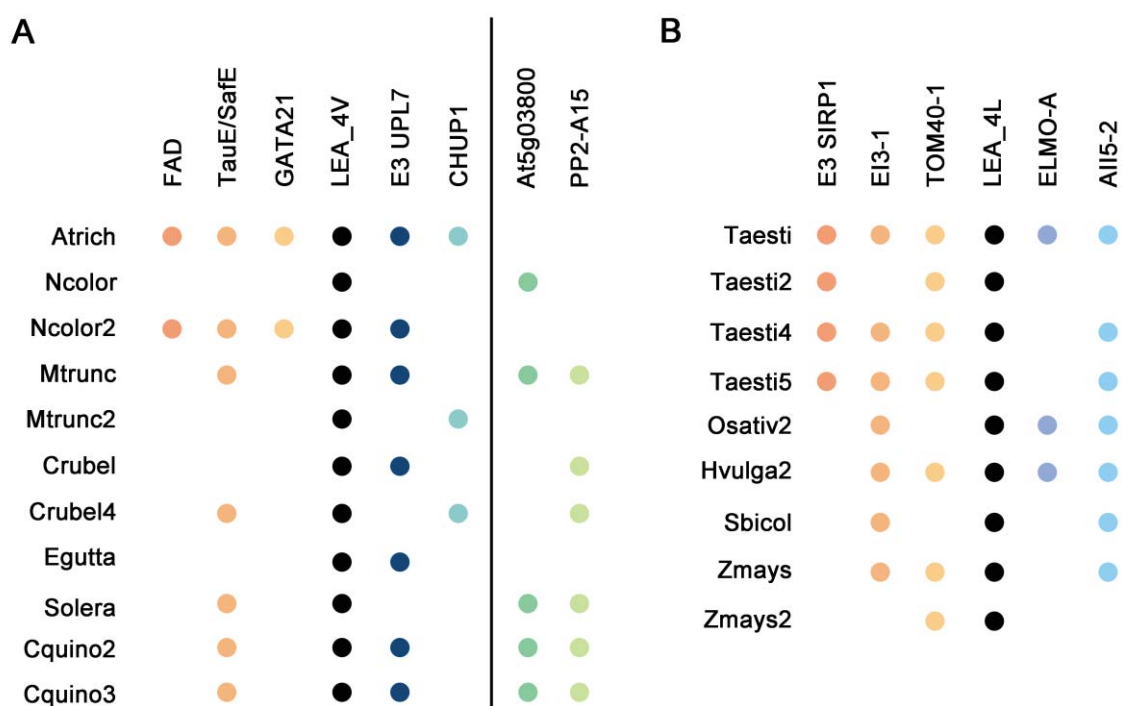


Figura II.10 Marcadores microsintéticos conservados de secuencias seleccionadas de LEA_4 formando dos grupos ortólogos (**A** y **B**). Cada punto coloreado indica la presencia del marcador indicado en la parte superior. Se indican las secuencias involucradas en la parte izquierda con su nombre resumido. El orden de los marcadores corresponde a *Amborella trichopoda* (**A**) y *Triticum aestivum* (**B**), y sus posiciones podrían variar entre gen y gen. Las secuencias posicionadas luego de la línea vertical (**A**) no fueron hallados en *A. trichopoda*, sino en las secuencias posteriores. FAD, FAD synthetase; TauE/SafE, sulfite exporter TauE/SafE family protein 4; GATA21, GATA transcription factor 21; LEA_4V, late embryogenesis abundant protein presentando MEME 4 (violeta); E3 UPL7, E3 ubiquitin-protein ligase UPL7-like; At5g03800, pentatricopeptide repeat-containing protein At5g03800; PP2-A15, F-box protein PP2-A15; E3 SIRP1, E3 ubiquitin-protein ligase SIRP1-like; EI3-1, ETHYLENE INSENSITIVE 3-like 1 protein; TOM40-1, mitochondrial import receptor subunit TOM40-1-like; LEA_4L, late embryogenesis abundant protein presentando MEME 10 (lima) ELMO-A, ELMO domain-containing protein A; All5-2, ABSCISIC ACID-INSENSITIVE 5-like protein 2.

El primer grupo ortólogo también incluye secuencias cuyos motivos N-terminales parecían presentar diferencias, pero el análisis revela que se trata de un mismo grupo y sugiere el origen ancestral de los motivos N-terminales correspondientes a los MEMEs 5 y 9.

Por otro lado, estudiando la microsinetia de las proteínas que presentan el MEME 8 en su región N-terminal, llamativamente no se hallaron marcadores que indicaran que pertenecieran al mismo grupo ortólogo. Los únicos casos que fueron una excepción corresponden a tres secuencias de monocotiledóneas que comparten un solo marcador. Casualmente, estas tres secuencias corresponden con aquellas LEA_4 de este subgrupo que presentaron un pI básico respecto del resto del conjunto, nuevamente sugiriendo un origen ancestral de las mismas, aunque con evidencia débil.

II.4 Discusión

II.4.1 La clasificación de las LEAs

Uno de los puntos centrales para la clasificación según Battaglia et al. (2008) es cuáles son los criterios a utilizar para generar los grupos. Inicialmente el criterio tomado fue el de tamaño y similitud de secuencias (Dure et al., 1989). Algunos trabajos hacen el esfuerzo de reclasificar las LEAs pasando a un sistema de “clases”, en lugar de los clásicos grupos (Tunnacliffe & Wise, 2007; Wise, 2003). Estos realizan una búsqueda de similitud de secuencias entre grupos, intentando discriminar *keywords* o pequeños péptidos que sean válidos para hallar y clasificar las LEAs en estas “clases”. Si bien esta búsqueda podría arrojar luz sobre las similitudes y disimilitudes de los diversos grupos de LEAs existentes hasta la fecha, Battaglia et al., propone que en este caso particular la búsqueda debe centrarse en regiones conservadas, y tiene en cuenta observaciones experimentales sobre las características físico-químicas de cada grupo. Para realizar su clasificación, Battaglia et al., toma como punto de partida las clasificaciones dadas a todas las LEAs siguiendo el surgimiento de los grupos según Dure et al. (1989) y compara su clasificación con la de otros trabajos y bases (Bies-Ethève et al., 2008; Mistry et al., 2021). En su trabajo, logra generar grupo coherentes basados en conservación de algunos motivos y características particulares. Una de las limitantes más importante de esta clasificación es que no tiene en cuenta el origen evolutivo de las proteínas. Al basarse en pequeños motivos conservados determinados mediante alineamiento local de secuencias, a diferencia del análisis de motivos mediante MEME suite, pierde la información almacenada en los motivos repetitivos que pueden o no guardar conservación en su posición y distancia en la secuencia proteica. A su vez no contempla la información evolutiva presente en las secuencias “no conservadas” (secuencias entre motivos conservados) que como se ha descrito recientemente para el grupo de las dehidrininas aporta características distintivas a las proteínas que permiten determinar que proteínas que conservan un módulo de secuencia conservado se diferencian en las regiones no conservadas y esto correlaciona directamente con un origen evolutivo distinto y en algunos casos propiedades físico químicas distintas (grupo ortólogo) (Melgar and Zelada, 2021).

Por ejemplo, Battaglia et al (2008) definen al grupo 5, en dos subgrupos que donde un análisis evolutivo muestra que se trata de proteínas no homólogas donde el grupo 5A se corresponde con las SMP, 5B con LEA_3 y 5C con LEA_2. Por otra parte, refiere que las proteínas pertenecientes al grupo 5 poseen mayor proporción de aminoácidos hidrofóbicos, lo que las hace no solubles, y además que se acumulan en semilla. Si bien estas afirmaciones son correctas para los grupos SMP y LEA_2 (correspondientes con sus grupos 5A y 5C), no logramos hallar reportes que hagan referencia estricta a una LEA_3 que posea estas

características. Cabe agregar que, en los reportes de semilla que se utilizaron de referencia para los análisis proteómicos de semillas del Capítulo I no se hallaron LEAs pertenecientes a las LEA_3 (o grupo 5B, según Battaglia et al.), y que las referencias proporcionadas en el trabajo de Battaglia et al. (Kim et al., 2005; Kiyosue et al., 1992; Maitra & Cushman, 1994; Park et al., 2003; Zegzouti et al., 1997) hacen referencia a LEAs “atípicas” que coinciden únicamente con los grupos SMP y LEA_2 según Pfam, y ninguna de ellas corresponde al grupo LEA_3.

Por otro lado, las LEA_4 que se abordan en esta tesis se encuentran incluidas por Battaglia en el grupo 3A y 3B. Esta subclasificación está basada en la variabilidad hallada en el motivo de 11 aminoácidos (que se encuentra representado en los MEMEs 1, 3 y 7, identificados en este trabajo) que se repite en forma de bloques. En la base de Pfam, la identificación de las LEAs correspondiente a este conjunto consta de una secuencia de 44 aminoácidos compuesta por repeticiones de este péptido de 11 aminoácidos, pero no se realiza una discriminación de los subgrupos como realizan Battaglia et al. En contraste, en este trabajo determinamos que existen otros MEMEs que definen las diferencias entre subgrupos de LEA_4, y, por el contrario, los MEMEs 1, 3 y 7 definirían indiscriminadamente a cualquier LEA_4. De hecho, logramos discriminar grupos ortólogos de LEA_4 en base a dos motivos conservados (MEMEs) ubicados en la región C-terminal de la proteína, el MEME 4 (violeta) y el MEME 10 (lima) no identificados previamente como motivos conservados.

II.4.2 El grupo LEA_3

Las LEA_3 se presentan como un grupo con una diversidad intrínseca evidente al analizar las secuencias que son incluidas en el mismo. Se trata de un grupo de proteínas pequeñas (91 aminoácidos de moda) y de bajo peso (10,65 kDa en promedio), con un contenido de glicina bajo (4,4% de moda) punto isoeléctrico básico (9,9 de moda). Su índice de hidropaticidad (GRAVY) no es extremadamente bajo, aunque su índice de plegamiento (Fold Index) no las sitúa como proteínas con una estructura definida. Las predicciones sobre su estructura secundaria son variables, habiendo proteínas con mayor y menor contenido de α -hélices, aunque consistentemente presentan un gran porcentaje de aminoácidos vinculados a giros (aproximadamente 40% en promedio). El análisis reveló la existencia de motivos conservados cuya distribución entre los taxones estudiados forma principalmente cuatro estructuras básicas. Estas estructuras parecen basarse en 4 motivos conservados, S, R, W y W', que se combinan dando lugar a unas pocas arquitecturas que llamamos por sus siglas SRW, RW, W y W'. Al mismo tiempo, puede ocurrir que dentro de estos grupos se dé la presencia de proto-motivos S o R que guardan similitud con estos, pero con poca identidad.

Dentro de las especies que fueron incluidas en este análisis, se observa que los motivos novedosos (los MEMEs además del 1) se presentan a medida que se avanza evolutivamente sobre la filogenia de las plantas. En *P. patens* (una briofita tolerante a la desecación), solo se observa el motivo representativo de las LEA_4, que llamamos W. Algunas secuencias de *S. moellendorffii* presentan el motivo W'. También cabe agregar que mientras *S. moellendorffii* presenta el motivo W en otras de sus secuencias, la región C-terminal de este motivo, que suele estar conservado en otras especies, está menos conservada. En *P. patens* la aparición del motivo W es clara y refuerza la idea de que el motivo W (sobre todo sus primeros aminoácidos) son características ancestrales de las LEAs de este grupo (Singh & Graether, 2020), aunque nuestros resultados sugieren que una afirmación similar es posible para el motivo W'.

Al introducir las estructuras de las angiospermas basales (*A. trichopoda* y *N. colorata*), las secuencias muestran nuevos motivos o proto-motivos. En el caso de *A. trichopoda*, dos de las tres secuencias de LEA_3 halladas cuentan con el motivo R bien definido. Las secuencias de *N. colorata*, sugieren que el motivo S tiene un origen alrededor de las angiospermas basales o anterior a las mismas, ya que dos de sus secuencias fueron vinculadas con la combinación de motivos S y R (MEME 2). Reportes previos sobre las LEAs 3 sugieren que su surgimiento podría haber ocurrido en plantas terrestres, ya que no se hallan secuencias de este grupo en algas (Singh & Graether, 2020).

La teoría de que el motivo S se desarrolló paulatinamente a lo largo de la evolución de las LEA_3 está reforzada por los estudios microsintéticos, que demuestran además de los marcadores, duplicaciones en tándem conservadas en algunas de las secuencias analizadas: *A. trichopoda* (angiosperma basal), *S. bicolor* (monocotiledónea) y *M. truncatula* (dicotiledónea), una observación que fue previamente realizada en *A. thaliana* (Bies-Ethève et al., 2008). Al mismo tiempo, el análisis microsintético aporta información sobre la existencia de dos grupos ortólogos de LEA_3 que llamamos LEA_3A y LEA_3B y de un tercer grupo, conformado por las secuencias conteniendo el motivo W' que, si bien no comparte marcadores de sintenia, no se relaciona con ninguno de los otros dos grupos de manera directa, por lo que fue considerado el grupo LEA_3C. La conclusión más parsimoniosa es que este grupo de secuencias pertenezcan a un mismo grupo ortólogo, y que las huellas del proceso evolutivo se hayan perdido, en caso contrario los motivos deberían haber surgido en forma independiente en proteínas con distinto origen evolutivo.

La evolución de las plantas se caracteriza por la duplicación de genomas, también conocida como poliploidía, un proceso mediante el cual se produce un aumento en el número de juegos completos de cromosomas en una especie. Esto puede ocurrir de forma natural o

inducida y puede tener importantes implicaciones evolutivas (Artur et al., 2019). La poliploidía puede ocurrir de dos formas principales: duplicaciones en tándem y duplicaciones en bloques. Las duplicaciones en tándem se refieren a la repetición de un segmento cromosómico en la misma región, lo que da como resultado múltiples copias de genes en tándem. Por otro lado, las duplicaciones en bloques implican la duplicación de segmentos más grandes del genoma, lo que lleva a la aparición de bloques cromosómicos duplicados. Estas duplicaciones de genomas pueden proporcionar nuevas oportunidades evolutivas para las plantas. Las copias adicionales de genes permiten que algunos de ellos adquieran nuevas funciones o se diversifiquen en su expresión, lo que puede conducir a la aparición de nuevas características fenotípicas. Algunas duplicaciones de genomas también pueden desempeñar un papel importante en la adaptación de las plantas a diferentes condiciones ambientales o en la respuesta a estreses bióticos o abióticos. En resumen, la duplicación de genomas en las plantas, ya sea a través de duplicaciones en tándem o en bloques, es un proceso evolutivo clave que puede generar diversidad genética y proporcionar nuevas oportunidades para la adaptación y evolución de las especies vegetales. Si bien existen trabajos que abordan el estudio de las LEAs incluyendo algunos matices de estas perspectivas (Artur et al., 2019; Bies-Ethève et al., 2008; Z. Zou et al., 2022), ninguno de ellos ofrece un análisis en profundidad para el grupo de las LEA_3 como se presenta aquí.

Algunos de nuestros hallazgos están relacionados con aquellos publicados recientemente (Singh & Graether, 2020), aunque con diferencias con implicancias importantes. Una de ellas radica en el abordaje utilizado para clasificar las LEA_3, que usa principalmente diferencias halladas en su región N-terminal y no profundiza en las relaciones evolutivas entre los subgrupos formados, pudiendo estar agrupando proteínas con distinto origen evolutivo. Nuestro análisis determinó que los dominios conservados definidos (S, R, W, W') correlacionan perfectamente con los grupos de ortología, de manera que cada uno de estos grupos representa una combinación distinta de motivos (LEA_3A = SRW; LEA_3B = RW, LEA_3C = W').

Las observaciones sobre los orígenes de cada uno de los subgrupos de LEA_3 en conjunto con la demostración de que existen diferencias físico-químicas entre ellas, sugiere que su clasificación requiere una reevaluación que tome en cuenta las nuevas diferencias halladas dentro del grupo y estudie la posibilidad de que su expresión y función se de en estadios fenológicos, condiciones y tejidos distintos dependiendo del subgrupo estudiado.

II.4.3 El grupo LEA_4

Los análisis realizados sobre este grupo indican que son proteínas de mayor tamaño que las LEA_3 (45,2 kDa en promedio) y punto isoelectrico ácido en promedio ($X_{pi}=5,93$), un contenido de glicina que no suele superar el 10% (7,9 % en promedio). Sin embargo, analizando las secuencias que forman parte del grupo se halló diversidad entre las secuencias. Una de las observaciones de la bibliografía (Battaglia et al., 2008; Dure et al., 1989), indica que el motivo de 11 aminoácidos que se repite diferente número de veces a lo largo de las diferentes secuencias es el principal factor que influye sobre el peso final de la proteína. Nuestro análisis está de acuerdo con este hecho, ya que el MEME1 hallado, que contiene tres repeticiones de este motivo, y los MEMEs similares (3 y 7) se pueden repetir en conjunto entre 2 y 11 veces. Sin embargo, los motivos conservados hallados sugieren que la subclasificación podría ser más compleja e involucrar diversos grupos ortólogos de genes.

Luego de observar la diversidad de la arquitectura de las secuencias y sus relaciones filogenéticas, se trabajó bajo la hipótesis de que dentro de este grupo podríamos distinguir hasta tres subgrupos con características diversas, por lo que se analizaron conjuntos de secuencias basados en estos grupos, poniendo hincapié en la presencia de tres motivos particulares: el MEME 4 (violeta), el MEME 10 (lima) y el MEME 8 (fucsia), los primeros dos de localización C-terminal, y el tercero de localización N-terminal. Las propiedades físico-químicas de cada uno de estos grupos sugirieron que podría tratarse de proteínas con distintas propiedades, cuyo punto en común es el motivo canónico de las LEA_4.

Los análisis microsintéticos permitieron confirmar que las secuencias pertenecientes al subgrupo del MEME 4 corresponden a un mismo grupo de ortólogos, dentro del cual se hallan secuencias de especies dicotiledóneas. Por otra parte, otro grupo ortólogo comprendido únicamente por especies monocotiledóneas surgió al analizar el contexto genético de las accesiones correspondientes al subgrupo que poseía el MEME 10. Sin embargo, para el grupo cuya región N-terminal presenta el MEME 8, se halló poca evidencia que vinculara estas especies entre sí o con cualquiera de los otros grupos. La única evidencia que indica que estos poseen un ancestro común se limita a un único marcador microsintético compartido por tres accesiones correspondientes a monocotiledóneas. Al mismo tiempo, algunos motivos detectados por MEME en la región N-terminal de otras secuencias (aquellas que no poseían el MEME 8), caen dentro del primer grupo de ortólogos (MEME 4), por lo que es aparente que esta variabilidad en la región N-terminal de este grupo posee un origen ancestral común. Estas observaciones guardan similitudes con las hechas en el trabajo de Artur et al., (2019), donde se sugiere que estas LEAs presentan una sintenia dinámica, generando comunidades que se separan entre angiospermas en general, monocotiledóneas y dicotiledóneas y, dependiendo

de la distancia evolutiva entre las especies estudiadas, es posible que no se hallen marcadores sinténicos que indiquen su ancestralidad. El ejemplo provisto relaciona a *P. patens* con *Xerophyta viscosa*, ambas especies tolerantes a la desecación, donde de los ocho genes identificados en *X. viscosa*, solo uno comparte marcadores con otras angiospermas, y ninguno de ellos con *P. patens*. El resto de los genes codificantes de LEA_4 en *X. viscosa* parecen ser parálogos internos, sugiriendo que las duplicaciones de estos genes son importantes en la tolerancia a la desecación (Costa et al., 2017). En el caso de *C. quinoa*, no encontramos esta expansión en la familia de las LEA_4. Esto sugiere que más allá del número de copias del gen, en esta especie podría haber factores regulatorios que resultan de mayor importancia para aportar tolerancia a estrés abiótico.

A nivel funcional, las proteínas pertenecientes a este grupo han sido relacionadas directa o indirectamente con la tolerancia a la deshidratación, salinidad y bajas temperaturas (Cattivelli & Bartels, 1990; Harada et al., 1989; Hsing et al., 1995; Romo et al., 2001). Su expresión en sistema heterólogos ha resultado en incrementos significativos de la tolerancia a estrés salino, deshidratación y frío en *Oryza sativa* y en *A. thaliana* (Babu et al., 2004; NDong et al., 2002; Xu et al., 1996). Una de las características importantes dentro de este grupo de proteínas es que, si bien las predicciones computacionales predicen que estas proteínas poseen gran cantidad de α -hélices, estudios de dicroísmo circular o espectroscopía infrarroja, indican que las mismas no poseen una estructura definida en solución, pero que su estructura responde activamente al nivel de agua en el medio, pudiendo alterar su estructura a medida que el medio se deshidrata, perder la estructura durante la rehidratación o inclusive tomar una estructura diferencial dependiendo de la tasa de deshidratación a la que se sometida (Goyal et al., 2003; Tolleter et al., 2007; Wolkers et al., 2001). Este punto podría ser importante a la hora de estudiar los subgrupos de LEAs que surgieron de este análisis, ya que sus índices de plegamiento difieren en conjunto con otros parámetros relacionados con su estructura. Estas diferencias podrían afectar negativamente su función como protectores contra diversos tipos de estrés abiótico si implican cambios en su estructura tanto en un estado de hidratación como bajo condiciones de deshidratación. Cabe remarcar, que las predicciones de desorden son similares para los diferentes subgrupos, y que las diferencias observadas entre sus índices de plegamiento, aunque significativas, podrían no ser suficientes para justificar un cambio en la estructura tridimensional de las proteínas.

II.4.4 Las LEAs de los grupos LEA_3 y LEA_4 en *C. quinoa*

A partir de este análisis se desprende que dentro de las LEA_3, *C. quinoa* posee proteínas pertenecientes a los tres subgrupos formados. Las secuencias curadas halladas

corresponden a: dos LEA_3A (con motivos SRW), una a las LEA_3B (motivo RW), y dos a las LEA_3C (motivo W'). Una de las correspondientes al grupo LEA_3A fue estudiada previamente en nuestro laboratorio, obteniendo resultados que indican su aumento de expresión en condiciones de estrés salino y por frío en un genotipo tolerante (Amarilla de Marangani) y recibe el nombre de *late embryogenesis abundant protein* Lea5-like (A.J. Rizzo, 2016)

En cuanto a las LEA_4 presentes en quinua: con excepción de una única secuencia correspondiente al grupo de las que presentan el MEME 8 (F), el resto (3 accesiones de proteína) forma parte del grupo ortólogo que posee el MEME 4 (V) en su región C-terminal. Todas las LEAs detectadas en el estudio de semilla también corresponden a este grupo ortólogo y se preseleccionó de ellas, coincidente al mismo tiempo con reportes previos donde su expresión aumentó más de 100 veces en el genotipo de quinua R49 bajo condiciones de estrés por sequía (Morales et al., 2017), para realizar análisis de expresión en plántulas y plantas adultas de *C. quinoa*. Finalmente, se destaca que dos accesiones de proteínas pertenecientes a este mismo subgrupo fueron detectadas en semilla (XP_021743171.1, XP_021737796.1), sugiriendo que en semilla madura se expresan ambos loci codificantes de las proteínas ECP63 homólogas.

II.5 Conclusiones

En este capítulo se logra profundizar sobre la complejidad de la clasificación de las LEAs y se aporta información evolutiva que podría ser de gran interés para el estudio de las proteínas de embriogénesis tardía. Se detectaron tres subgrupos dentro de las LEA_3 (que llamamos provisoriamente A, B y C), dos de los cuales conforman grupo de ortólogos con ancestros comunes, mientras que un tercero no parece compartir marcadores internos ni con los demás grupos. También detectamos tres subgrupos de LEA_4 (que llamamos provisoriamente con el nombre del MEME o motivo conservado que poseían en común), uno correspondiente a monocotiledóneas y otro a dicotiledóneas y reforzados por análisis microsinténicos en cuanto a su origen evolutivo común. Nuevamente, un tercer grupo conteniendo secuencias tanto de monocotiledóneas como de dicotiledóneas, no presentó marcadores microsinténicos que permitieran dilucidar si poseen un antepasado común.

Aquí se presentaron motivos que, hasta donde hemos logrado estudiar, no han sido reportados previamente y que en conjunto con estudios microsinténicos han demostrado ser útiles para discriminar diferentes grupos ortólogos de proteínas dentro de cada uno de los grupos LEA estudiados. Cabe remarcar que, en ambos grupos estudiados, se dio la presencia de un subgrupo que a pesar de compartir características de secuencia con los criterios de identificación de Pfam, no pudieron ser respaldados por los estudios microsinténicos en cuanto a su ancestralidad y origen. Esto plantea la duda de si estos subgrupos corresponden realmente a sus respectivos grupos LEA, son simplemente artefactos de detección y anotación generados automáticamente o por similitud de secuencias, o inclusive si se trata de un subgrupo cuya reclasificación debería ser considerada ya sea dentro o fuera de las LEAs.

Si bien la clasificación de LEAs actual por Pfam posibilita la identificación de grupos de manera específica, este estudio demuestra que se puede observar una variabilidad intrínseca de cada uno de los grupos estudiados que atraviesa aspectos físico-químicos y evolutivos. Las características de las diferentes LEAs halladas dentro de los grupos LEA_3 y LEA_4 amerita un estudio en profundidad sobre su expresión en diferentes especies, estadios de desarrollo, tejidos y condiciones, para poder esclarecer si las propiedades de protección que le son atribuidas a las proteínas de cada uno de estos grupos son comunes a todas ellas o a subgrupos particulares y cuáles son las propiedades estructurales que resultan indispensables para cumplir con estas funciones, siendo un candidato importante poseer un alto grado de desorden en solución que disminuya en situaciones de deshidratación dando lugar a una estructura tridimensional. Los análisis de motivos conservados prueban ser una metodología eficiente para discriminar entre subgrupos de proteínas y, utilizada correctamente, puede arrojar luz sobre estos hechos al ser utilizadas comparativamente.

Un candidato perteneciente a *C. quinoa* de cada grupo LEA fue seleccionado en base a resultados preliminares o reportes previos para el estudio de su expresión en diferentes estadios de desarrollo de *C. quinoa* en nuestro laboratorio.

Capítulo III

Expresión de dos proteínas LEA de interés (LEA_4 y LEA_3) en diferentes estadios de desarrollo y tejidos de *C. quinoa* sometidos a estrés salino

III.1 Introducción

III.1.1 Estadios de desarrollo y expresión de LEAs

Las proteínas LEA pueden encontrarse en cualquier etapa de desarrollo de las plantas y también pueden presentarse en diferentes órganos y tejidos (Battaglia et al., 2008; Bies-Ethève et al., 2008; Chatelain et al., 2012; Close, 1997). Durante el desarrollo de la semillas ortodoxas, éstas son importantes para la tolerancia a la deshidratación que se produce sobre el final del programa de desarrollo (Leprince et al., 2017). En otros estadios de crecimiento y desarrollo, su expresión puede verse alterada dependiendo del genotipo y el tipo de estrés al que está sometida. Al mismo tiempo, la expresión de LEAs en plantas adultas sometidas a estrés salino puede variar dependiendo del genotipo estudiado (A.J. Rizzo, 2016)

La quinua es una planta de gran diversidad genética, presentan expresiones diferenciales de distintos genes debido a las variaciones genéticas inherentes a cada genotipo, así como a la selección artificial realizada por los agricultores a lo largo del tiempo. Estas diferencias en la expresión génica dan lugar a características fenotípicas distintivas en cada genotipo de quinua. Es por ello que una caracterización funcional de estas proteínas en quinua debe ser abordada teniendo en cuenta el genotipo, estadio fenológico y el estrés a evaluar.

III.1.2 Las funciones de las LEA en el estrés abiótico

Numerosos estudios sitúan a las proteínas LEA como candidatos en la tolerancia de diversas especies a varios tipos de estrés abiótico o inclusive biótico. Analizando los niveles de mensajero o en algunos casos la abundancia de proteína y su localización subcelular mediante el uso de anticuerpos (Burrieza et al., 2012), se hallaron proteínas LEA correspondientes a diferentes grupos cuya transcripción, traducción y localización varía al comparar situaciones de estrés como sequía y salinidad con su control (Bies-Ethève et al., 2008; Burrieza et al., 2012; NDong et al., 2002; Romo et al., 2001; Rorat, 2006).

Por otro lado, cuando estas proteínas se encuentran en medio acuoso, por lo general no poseen una estructura tridimensional definida. Sin embargo, diversos estudios y predicciones *in silico* indican la posibilidad de formación de estructuras secundarias de tipo α -

hélice (Amara et al., 2014; Goyal et al., 2003; Romo et al., 2001). Como proteínas desordenadas, no son fáciles de asociar a ninguna función particular. Sin embargo, ensayos *in vitro* sobre la actividad enzimática simulando situaciones de estrés hídrico, demuestran que la presencia de LEAs en el medio asiste al mantenimiento de la actividad enzimática (Goyal et al., 2005; Reyes et al., 2005), aparentemente evitando la agregación de proteínas. La hipótesis que se sostiene actualmente es que estas proteínas pueden envolver a las enzimas en un medio acuoso conservando la hidratación y estabilizando su estructura, y en consecuencia sosteniendo su actividad. En paralelo, las LEAs también han sido estudiadas en relación a otros tipos de moléculas y funciones que no están necesariamente relacionadas con la conservación de la actividad enzimática. Por ejemplo, existen estudios relacionando la presencia de LEAs con la estabilización de matrices de cristales de azúcar y estabilización de fases agua-agua, e inclusive en otros casos se propone la homo u heterodimerización de proteínas LEA en situaciones de estrés (Candat et al., 2014; He et al., 2021; Hinch et al., 2021; Koubaa et al., 2019; Thalhammer et al., 2010).

En este capítulo, haciendo uso de la técnica de RT-qPCR (PCR cuantitativa en tiempo real) se evaluó la expresión de dos genes correspondientes a una proteína LEA_3 y una proteína LEA_4 de interés en plántulas de quinua de 4 genotipos y diferentes órganos de plantas adultas de quinua del genotipo BO25 sometido o no a estrés salino. La LEA_3 fue escogida en base a resultados preliminares obtenidos en nuestro laboratorio que mostraron que la expresión de la misma en hojas de plantas adultas de quinua sometidas a estrés salino y a bajas temperaturas aumenta significativamente en el genotipo Amarilla de Maranganí (A.J. Rizzo, 2016). La LEA_4 escogida para este estudio fue seleccionada en base a los resultados obtenidos en proteómica de semilla (Capítulo I) y trabajos previos que muestran que su expresión se induce más de 100 veces en el genotipo R49 de quinua cuando este es sometido a estrés por sequía (Morales et al., 2017).

III.2 Materiales y métodos

III.2.1 Material vegetal

Se seleccionaron semillas de cuatro genotipos de quinua: Cahuil, BO25, Villarrica y Sajama. Para los ensayos en planta adulta, se escogieron los genotipos BO25 y Cahuil. Todo el material utilizado proviene de los bancos de germoplasma CIP (Lima, Perú) e INIA-Vicuña (La Serena, Chile).

III.2.2 Esterilización

Para evitar contaminaciones por hongos u otros organismos, las semillas fueron y esterilizadas antes de la siembra mediante una incubación en agitador de 10 minutos a temperatura ambiente en una solución 5% lavandina en agua bidestilada. Posteriormente se realizaron tres lavados de 5 minutos con agua bidestilada.

III.2.3 Siembra y ensayos de crecimiento de plántula

Para la germinación entre 50 y 60 semillas de cada cultivar fueron colocadas en placas de Petri cuadradas conteniendo medio Murashige Skoog (MS) 0,5% en agar planta (marca/empresa). Una vez comenzada la germinación se controló periódicamente el largo de la radícula y se contabilizaron aquellas que tenían al menos 2 mm de crecimiento radicular como semillas germinadas, una vez alcanzado un 70% de germinación en cada placa, se dio comienzo a los ensayos de crecimiento de plántula en medio salino.

Utilizando un máximo de 10 semillas distribuidas homogéneamente en una línea central de la placa se distribuyeron las semillas al azar entre placas conteniendo el mismo medio mencionado anteriormente y otras conteniendo el medio con un agregado de 250 mM de NaCl (Biopack). Las placas fueron incubadas en una cámara con temperatura (20 ± 1 °C) y fotoperíodo (16/8 horas luz/oscuridad) controlados dentro de una cámara de incubación (Sanyo). Para evitar problemas de deshidratación de los medios, se incluyó una bandeja de agua destilada en todas las incubaciones. Los ensayos contaron con triplicados o cuadruplicados técnicos para cada condición.

Las plántulas fueron cosechadas 7 días después de su traspaso a las placas de tratamiento.

III.2.4 Ensayos de estrés salino en plantas adultas

Plántulas previamente germinadas en germinadores fueron trasplantadas de forma individual a macetas de 14 cm de diámetro con mezcla de tierra GROWMIX (Terrafertil,

Argentina). Se siguió un cronograma de riego de tres veces por semana hasta que las plantas desarrollaron cuatro a cinco pares de hojas verdaderas. A partir de ese momento se dividieron en dos grupos al azar. Uno de los grupos fue regado periódicamente con agua bidestilada, mientras que el otro recibió un riego escalonado de 50 mM de NaCl cada 2 días hasta alcanzar la concentración de 500 mM NaCl. Una vez alcanzada se sostuvo un cronograma de riego a capacidad de campo por una semana. Las plantas fueron fertilizadas semanalmente desde su trasplante con HAKAPHOS (0.7 g/L) (COMPO EXPERT, Argentina) agregado al agua de riego.

Luego de la semana de tratamiento se cosecharon 2 hojas por planta de todas las plantas. La cosecha de flores y raíces se dio al alcanzar las etapas de floración, período durante el cual se sostuvo el cronograma de riego. Las plantas fueron cosechadas y se almacenaron inflorescencias y raíces. Las últimas cuidadosamente lavadas con agua corriente y enjuagadas finalmente en agua bidestilada. Se guardó el material a -80 °C para futuras extracciones de ARN.

III.2.5 Morfometría

Al momento de la cosecha las plántulas fueron fotografiadas para obtención de variables morfométricas. Los resultados fueron utilizados para evaluar el comportamiento de los cultivares. La masa fue obtenida utilizando una balanza de precisión Traveler (OHAUS, EE.UU.). Todo el material fotografiado fue analizado mediante software ImageJ (Schneider et al., 2012).

III.2.6 Extracción de ARN de plántula

Entre 70 y 100 mg de masa fresca de plántulas fueron molidos en morteros esterilizados y pretratados con solución inhibidora de RNAsas (RNaseZap, Thermo Fisher), utilizando nitrógeno líquido. Se utilizó 1 ml de TRIzol por cada 100 mg de tejido para realizar la extracción. Luego de una incubación a temperatura ambiente se agregaron 200 µl de Cloroformo (Anedra, Research AG) y se agitó vigorosamente por 15 segundos. Se incubó por 10 minutos a temperatura ambiente y se centrifugó a 12.000 g y 4 °C por 15 minutos. Se tomó el sobrenadante y se procedió con la extracción de ARN enriquecido en miRNAs mediante el uso del kit ARNmi mirVana (Ambion, Thermo Fisher Scientific) siguiendo el protocolo.

III.2.8 Extracción de ARN de planta adulta.

Para realizar las extracciones se utilizaron 100 mg de material de flor y hoja y 300 mg de raíz por muestra. Se molió la muestra en mortero con nitrógeno líquido hasta formar polvo y se agregó 1 ml de agente TRIzol. Se recogió el extracto en un tubo eppendorf de 2 ml y se agregaron 200 μ l de cloroformo. Se agitó vigorosamente por 15 segundos y se incubó a temperatura ambiente por 10 minutos. Se centrifugaron las muestras a 12.000 g y 4 °C por 15 minutos (Eppendorf) y se recuperó el sobrenadante en un tubo eppendorf de 1,5 ml. Se agregó un volumen de isopropanol 100%, se vortexeó suavemente y se incubó por 5 minutos a temperatura ambiente. Se centrifugaron las muestras a 12.000 g y 4 °C por 15 minutos. El precipitado producto de la centrifugación contiene el ARN. Se removió el sobrenadante y se agregó etanol 70% para enjuagar el pellet. Luego de esto, antes de descartar el sobrenadante se volvió a centrifugar la muestra en las mismas condiciones por 2 minutos. Se removió el sobrenadante y se permitió la evaporación del alcohol restante por 5 minutos a temperatura ambiente dejando los tubos abiertos. Se diluyó el pellet en 30 μ l de buffer TE, haciendo uso de un termobloque (Benchmark, Sayerville, USA) a 55 °C y agitación por vortex o a mano hasta que se observó la dilución total.

Las muestras fueron cuantificadas en NanoDrop (Thermo Fisher Scientific) para conocer su concentración aproximada y pureza. Aquellas de baja pureza (cociente 260/230 por debajo de 1,4) fueron posteriormente purificadas nuevamente mediante una precipitación con acetato de sodio (sección III.2.9).

III.2.9 Precipitación de ARN con acetato de sodio.

Un total de 10 μ g de ARN fue utilizado como material inicial. Se compuso una solución con concentración final de 0,3 M de acetato de sodio (NaAc) (Anedra, Research AG) y volumen final de 100 μ l en un tubo eppendorf de 1,5 ml. Se agregaron 300 μ l de isopropanol 100 %. Se incubó la mezcla a -20 °C por 1 hora y posteriormente se centrifugó a 12.000 g y 4 °C por 30 minutos. Se removió el sobrenadante y se realizó un lavado del pellet con etanol 75 %. Se centrifugó nuevamente en las mismas condiciones por 2 minutos, se removió el sobrenadante y se permitió la evaporación de solvente por 5 minutos a temperatura ambiente. Se disolvió el pellet en 30 μ l de agua bidestilada y se cuantificó utilizando NanoDrop (Thermo Fisher Scientific).

III.2.10 Verificación de integridad del ARN por gel de agarosa

Se utilizaron geles de agarosa 1,5 % con BrEt para verificar la integridad del ARN a utilizar. Se sembró 1 μ g de material diluido en 10 μ l finales y con agregado de *Loading*

Buffer (LB) con azul de bromofenol como colorante. La electroforesis fue realizada a 100 V por 45 minutos. El resultado fue visualizado y fotografiado utilizando luz UV dentro de una cámara G-box (Syngene) utilizando una exposición de 800 ms.

III.2.11 Obtención de cDNA

Partiendo de 1 ug de ARN total se realizó un tratamiento de DNAsa (NEB). Se agregó 1U/ul de enzima y se incubó a 37 °C por 30 minutos. Se agregó 1 ul de EDTA 50 mM y pH (6,8) y se incubó a 75 °C por 15 minutos para terminar la reacción.

La totalidad del ARN tratado fue incubado con una concentración final de 0,3 uM de oligo dT. Luego se utilizó la transcriptasa reversa M-MLV (NEB) de acuerdo al protocolo incubando 45 minutos a 42 °C e inactivando la reacción a 75 °C por 15 minutos.

III.2.12 Real time qPCR y primers

Se probaron varias diluciones de los cDNAs obtenidos en los ensayos de PCR cuantitativa con el objetivo de encontrar las mejores condiciones de eficiencia de amplificación y valores de Ct. Los cDNAs obtenidos se utilizaron en diluciones 1/50 para todos los tejidos y estadios de desarrollo.

Los ensayos contaron con triplicados biológicos y técnicos, y se midieron dos genes de referencia en simultáneo para obtener mayor confianza en las mediciones. Los genes de *housekeeping* escogidos fueron ubiquitina (UBI) y *Elongation factor 1 alpha* (EF1a) en base a ensayos anteriores evaluando exclusivamente genes de referencia.

Las amplificaciones y mediciones fueron realizadas utilizando Master Mix (BIORAD) y la cicladora MyCycler (BIORAD) manejada mediante el software iCycle (BIORAD). El protocolo de la cicladora y los primers utilizados se presentan en el ANEXO II.

Las herramientas del software se utilizaron para evaluar las curvas de *melting* de las reacciones y los datos de fluorescencia para cada punto de medición fueron extraídos del software para su evaluación mediante el método de regresión lineal con el software LinRegPCR (Ruijter et al., Nucleic Acids Research, 2009) y posteriormente graficados utilizando el software GraphPad Software (Prism, Dotmatics).

III.2.13 Geles de agarosa para verificación de qPCR

Geles de agarosa 2,5 % fueron utilizados para realizar electroforesis de los productos de PCR *Real time* en su punto final. Se realizó la electroforesis a 100 V por 40 minutos. Se visualizaron los resultados con G-box (Syngene) y SnapGene (Dotmatics).

III.3 Resultados

III.3.1 Efecto de la salinidad sobre la morfología de las plántulas.

Las mediciones sobre plántulas revelan que todos los genotipos estudiados se ven afectados por la presencia de NaCl en el medio (Figura III.1 y Figura III.2). Sin embargo, al comparar la proporción de los cambios en cada cultivar se observa que Cahuil y Villarrica presentaron una disminución menos marcada del vástago frente a BO25 y Sajama. En todos los casos la longitud de las raíces se vio severamente afectada, siendo Sajama el genotipo más afectado. Al mismo tiempo, esta diferencia se traslada al cociente vástago/raíz, donde Sajama es el único genotipo que presenta una diferencia significativa en el aumento de su cociente (Figura III.1).

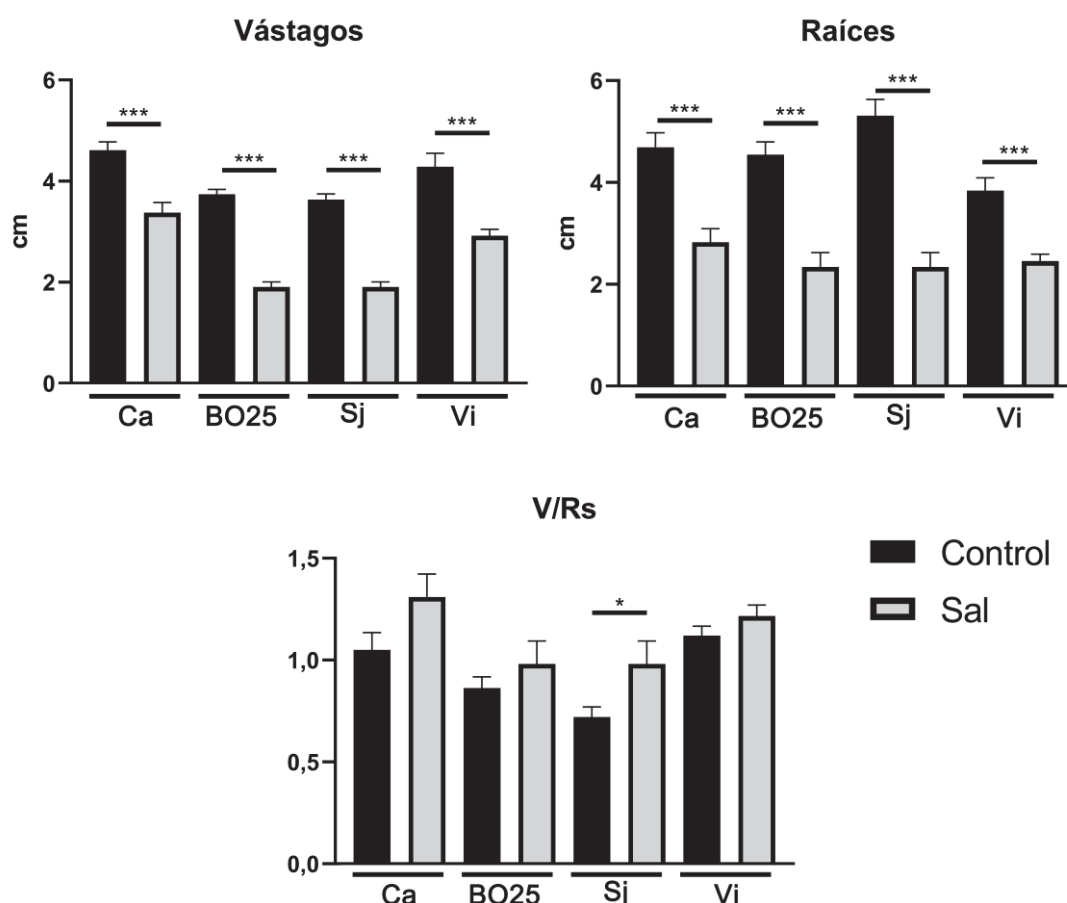


Figura III.1 Morfometría de plántulas de quinua sometidas o no a estrés salino (250 mM NaCl). Se presentan las medidas de vástago y raíz (cm) y el cociente de ambos (V/Rs). Los resultados significativos se encuentran indicados con asteriscos: (*) $p < 0.05$; (***) $p < 0.001$. Ca, Cahuil; Sj, Sajama; Vi, Villarrica.

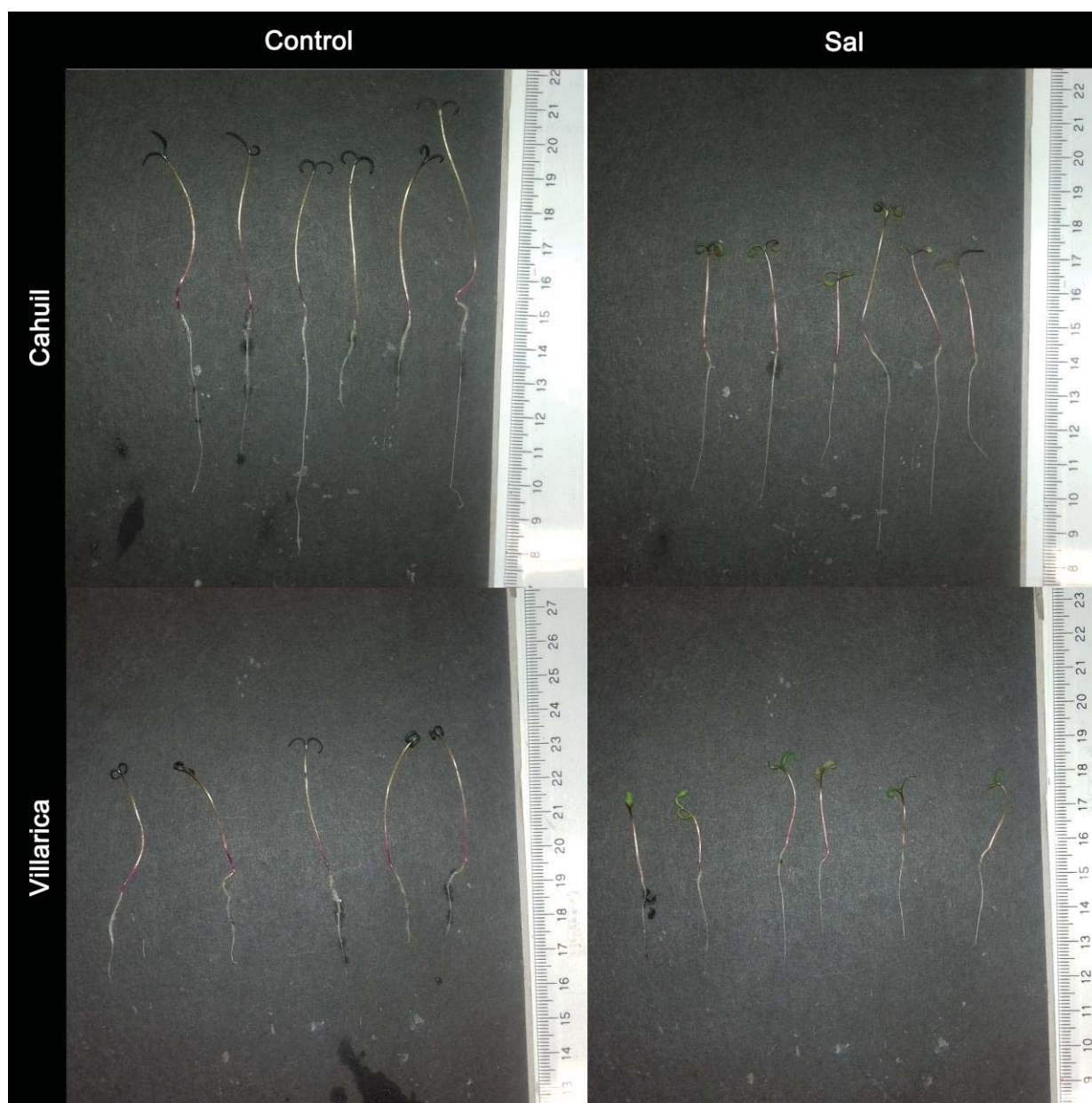


Figura III.2 Plántulas de quinua luego de 7 días de cultivo *in vitro*. Se presentan plántulas representativas de los resultados obtenidos para los tratamientos en los genotipos de Cahuil y Villarrica sometidas (derecha) o no (izquierda) a estrés salino. La escala se presenta en cada fotografía.

III.3.2 Expresión de LEA_3 (*late embryogenesis abundant protein* Lea5-like) y LEA_4 (*late embryogenesis abundant protein* ECP63-like) en plántulas de cuatro genotipos de *C. quinoa*

Al analizar la expresión del mensajero codificante de la LEA_3 mediante RTqPCR se encontró que ninguno de los cultivares utilizados mostró cambios significativos en su expresión (Figura III.3). Esto implica que, si bien existe un nivel de expresión basal para el transcripto, este no responde a la situación de estrés por salinidad en las plántulas de los genotipos estudiados.

Los resultados para LEA_4, en cambio, muestran una clara inducción por estrés salino en distintos genotipos. A excepción del cultivar Sajama, hay un aumento significativo en la abundancia del transcripto en las plántulas sometidas a estrés salino. Se encontró una diferencia de aproximadamente 13, 15 y 40 veces de expresión para Cahuil, BO25 y Villarica, respectivamente.

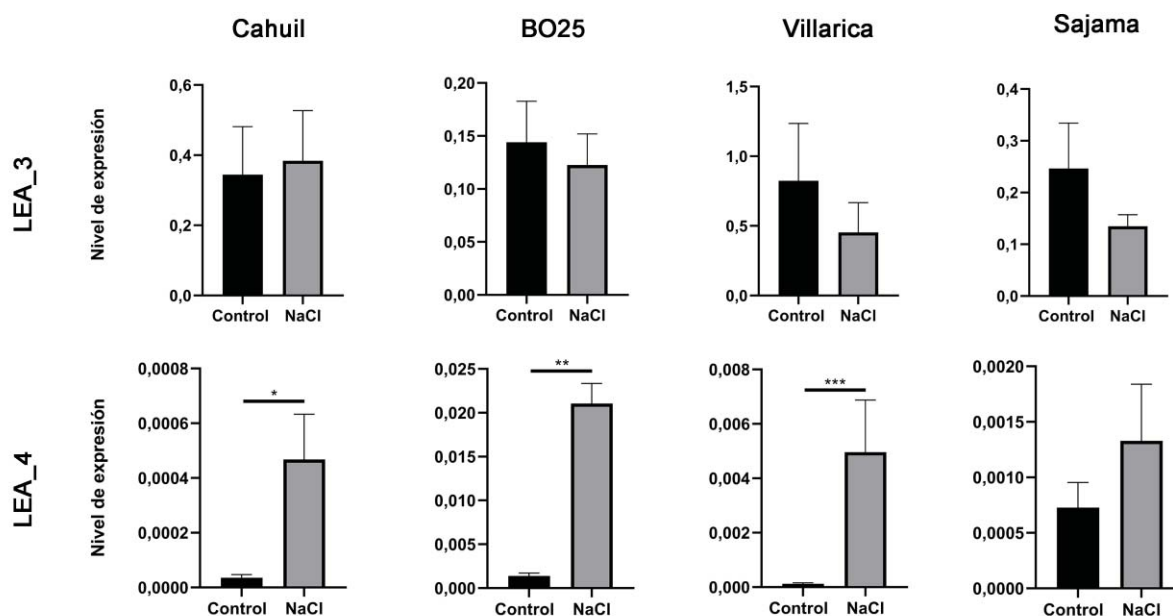


Figura III.3 Gráficos de expresión relativa de los mensajeros correspondientes a la LEA_3 (arriba) y LEA_4 (abajo) en plántulas de quinua de cuatro genotipos contrastantes en situación control y sometidas a 250 mM de NaCl. Se grafica el promedio del nivel de expresión acompañado por su desvío estándar. Las diferencias significativas se expresaron con * = $p < 0,05$, ** = $p < 0,01$, *** = $p < 0,001$ ($\alpha = 0,05$)

III.3.3 Expresión de LEA_3 y LEA_4 en tejidos de planta adulta

Para evaluar la expresión de las LEAs de interés en diferentes tejidos, plantas adultas de quinoa sometidas a estrés salino (NaCl 500 mM) fueron cosechadas separando raíz, hojas completamente desarrolladas e inflorescencias de las mismas. Para estos ensayos se utilizaron plantas del genotipo BO25, un cultivar con comportamiento sensible en experimentos de germinación proveniente del centro-sur de Chile. Si bien también se planificó inicialmente evaluar la expresión en tallo, este tejido presentó problemas a la hora de realizar las mediciones, no pudiendo obtener resultados reproducibles.

La expresión de LEA_3 no mostró diferencias significativas en ninguno de los tejidos al comparar los resultados de las plantas control con las tratadas (Figura III.4).

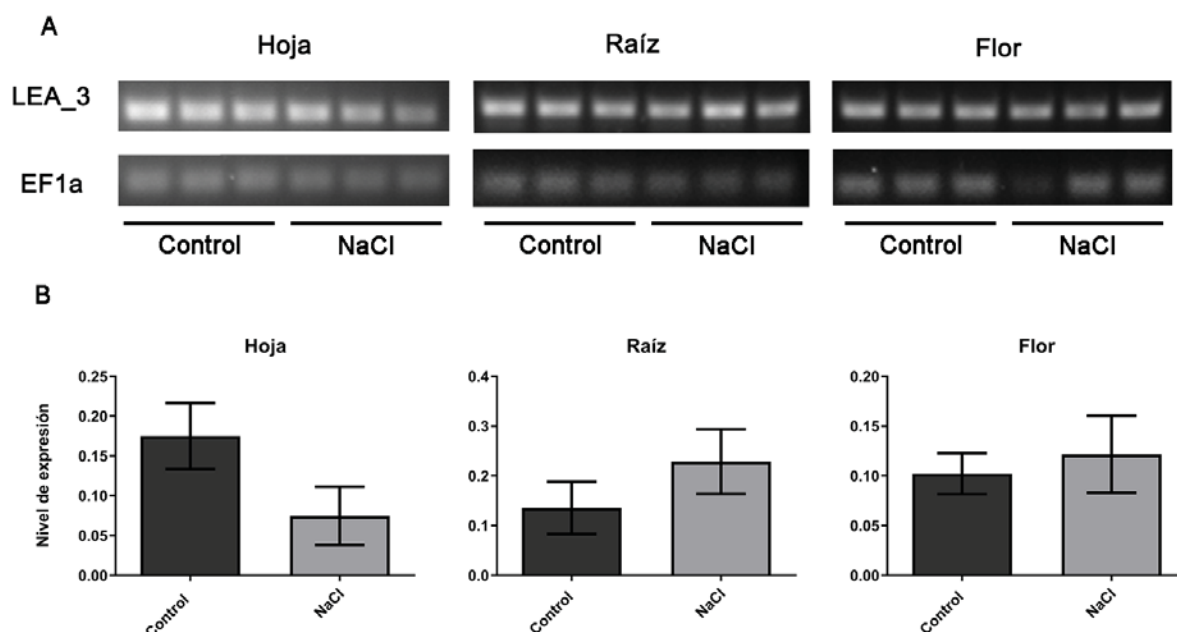


Figura III.4 Resultados de cuantificación de LEA_3 en muestras de hojas, raíces y flores de plantas adulta de *C. quinoa* del genotipo BO25. **A.** Geles de agarosa 2,5 % de los productos finales de Real Time para LEA_3. Se muestra el gen de interés arriba y el gen de *housekeeping* EF1- α abajo. Cada columna representa una réplica biológica. **B.** Gráficos de expresión para el mensajero de LEA_3 en hojas, raíces y flores de planta adulta de BO25 en condiciones control y bajo tratamiento salino (500 mM NaCl). Se expresa la media del nivel de expresión normalizado respecto del gen de referencia acompañado de su desvío estándar. Se presenta la media acompañada de su desvío estándar (SD).

En cuanto al mensajero de la LEA_4, no se obtuvieron resultados que indiquen niveles detectables del mismo en hoja. En raíz y flor, sin embargo, los mismos son detectados en condiciones basales, siendo más abundante en la inflorescencia (Figura III.5.B). Sin embargo, el análisis de la expresión en geles de agarosa mediante una caracterización semicuantitativa por RT-PCR de punto final sugiere una inducción de la expresión de LEA_4 en raíz. En cuanto a la inflorescencia no se observa inducción por estrés, pero los valores de expresión son mayores que en los otros tejidos (Figura III.5.A). Esto sugiere que existe una tendencia de la LEA_4 a expresarse en condiciones de estrés salino en raíz y que podría encontrarse en niveles basales indetectables o ausente en situaciones control.

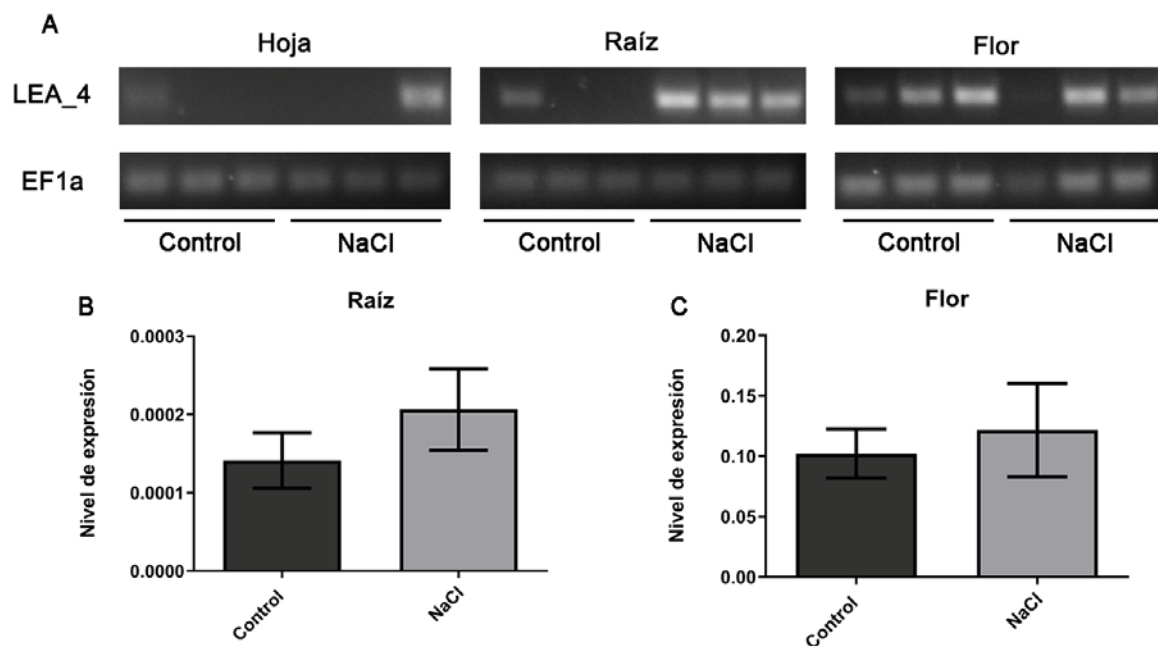


Figura III.5 Resultados de cuantificación de LEA_4 en muestras de hojas, raíces y flores de plantas adultas de *C. quinoa* del genotipo BO25. **A.** Geles de agarosa 2,5% de los productos finales de Real Time para LEA_4 de los tres tejidos evaluados en este trabajo. Se muestra el gen de interés arriba y el gen de *housekeeping* EF1- α abajo. Cada columna representa una réplica biológica. **B** y **C.** Niveles de expresión de LEA_4 en situación control y bajo estrés salino (500 mM NaCl) para raíces (B) y flores (C) de BO25. Se presenta la media acompañada de su desvío estándar (SD).

III.4 Discusión y conclusiones

Una de las primeras dificultades que se hallaron para la cuantificación de la expresión de los genes de interés es la gran similitud de secuencia entre las secuencias correspondientes al mismo grupo de ortólogos. Esto impide utilizar oligonucleótidos que discriminen entre las copias del gen presentes en *C. quinoa*. En el caso de la LEA_4 ocurre algo similar. Inclusive en el reporte hallado que indica que la LEA_4 aumenta su expresión en el plantas adultas sometidas a estrés por sequía (Morales et al., 2017), los oligonucleótidos reportados no son específicos de una única isoforma.

Los resultados obtenidos indican que la LEA correspondiente al grupo LEA_3 posee una expresión basal en plántulas de quinua, pero no posee respuesta ante el estrés por salinidad, independientemente del genotipo estudiado. En plantas adultas del genotipo sensible BO25, se detectaron niveles de transcripto en hojas, raíces e inflorescencias independientemente del tratamiento con 500 mM de NaCl. En conjunto con resultados preliminares de nuestro grupo analizando el genotipo tolerante Amarilla de Maranganí (A.J. Rizzo, 2016), esto indicaría que este gen responde de formas distintas dependiendo del genotipo estudiado. Será necesario extender este estudio a otros genotipos de quinua caracterizados en su tolerancia al estrés salino para determinar si existe una correlación.

Por otro lado, el mensajero correspondiente a la LEA_4 mostró un incremento de su expresión en plántulas en tres de los cuatro cultivares evaluados al someterlos a un tratamiento salino de 250 mM de NaCl, siendo Sajama el único cultivar que no mostró inducción de la expresión. Llamativamente, este genotipo proveniente de Bolivia pertenece a la ecorregión del altiplano, mientras los otros tres cultivares, Cahuil, BO25 y Villarrica, tienen su origen en tierras bajas de Chile. Notoriamente, Sajama fue la variedad más afectada por el estrés salino, en cuanto a la elongación de la raíz y a la relación vástago/raíz, sugiriendo que la expresión de LEA_4 en etapas tempranas del desarrollo de las plántulas podría tener un rol en la tolerancia al estrés salino en los genotipos del nivel del mar. En el mismo sentido, la LEA_4 mostró, en los ensayos de estrés en plantas adultas del genotipo BO25, una inducción de su expresión en raíz, y no en otros tejidos, apoyando un posible rol en los ecotipos del nivel del mar relacionado con la capacidad de la raíz de tolerar el estrés.

Estos resultados indican presencia del transcripto en inflorescencias tanto en situación control como bajo tratamiento, y en raíces de plantas bajo tratamiento salino. Esto sugiere que el transcripto de LEA_4 podría inducirse por estrés salino en raíces de quinua, particularmente en las sensibles. No obstante, reportes previos (Morales et al., 2017) indican que al menos una variante de este transcripto se expresa diferencialmente en hojas de plantas del genotipo R49 sometidas a estrés por sequía.

Bibliografía

- Adolf, V. I., Jacobsen, S. E., & Shabala, S. (2013). Salt tolerance mechanisms in quinoa (*Chenopodium quinoa* Willd.). *Environmental and Experimental Botany*, 92, 43–54. <https://doi.org/10.1016/j.envexpbot.2012.07.004>
- Amara, I., Zaidi, I., Masmoudi, K., Dolores Ludevid, M., Pagès, M., Goday, A., & Brini, F. (2014). Insights into Late Embryogenesis Abundant (LEA) Proteins in Plants: From Structure to the Functions. *American Journal of Plant Sciences*, 5(5), 3440–3455. <https://doi.org/10.4236/ajps.2014.522360>
- Artur, M. A. S., Zhao, T., Ligterink, W., Schranz, E., & Hilhorst, H. W. M. (2019). Dissecting the genomic diversification of late embryogenesis abundant (LEA) protein gene families in plants. *Genome Biology and Evolution*, 11(2), 459–471. <https://doi.org/10.1093/gbe/evy248>
- Babu, R. C., Zhang, J., Blum, A., Ho, T. H. D., Wu, R., & Nguyen, H. T. (2004). HVA1, a LEA gene from barley confers dehydration tolerance in transgenic rice (*Oryza sativa* L.) via cell membrane protection. *Plant Science*, 166(4), 855–862. <https://doi.org/10.1016/j.plantsci.2003.11.023>
- Bailey, T. L., & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology*, 2, 28–36.
- Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., da Silva, A., Denny, P., Dogan, T., Ebenezer, T. G., Fan, J., Castro, L. G., ... Teodoro, D. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*, 49(D1), D480–D489. <https://doi.org/10.1093/nar/gkaa1100>
- Battaglia, M., Olvera-Carrillo, Y., Garcíarrubio, 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>
- Bazile, D. (2015). La dinámica de la expansión mundial de la quínoa. In A. Romero & A. Zurita-Silva (Eds.), *Tierra Adentro* (Issue 108, p. 12). www.inia.cl
- Bhargava, A., Shukla, S., & Ohri, D. (2006). *Chenopodium quinoa* — An Indian perspective. 23(April 2005), 73–87. <https://doi.org/10.1016/j.indcrop.2005.04.002>
- Bies-Ethève, N., Gaubier-Comella, P., Debures, A., Lasserre, E., Jobet, E., Raynal, M., Cooke, R., & Delseny, M. (2008). Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. *Plant Molecular Biology*, 67(1–2), 107–124. <https://doi.org/10.1007/s11103-008-9304-x>
- Bojórquez-Velázquez, E., Barrera-Pacheco, A., Espitia-Rangel, E., Herrera-Estrella, A., & De La Rosa, A. P. B. (2019). Protein analysis reveals differential accumulation of late embryogenesis abundant and storage proteins in seeds of wild and cultivated amaranth species. *BMC Plant Biology*, 19(1), 1–17. <https://doi.org/10.1186/s12870-019-1656-7>
- Bray, E. A. (1997). Plant responses to water deficit. *Cell*, 1385(96). <https://sci-hub.io/http://www.sciencedirect.com/science/article/pii/S1360138597825629>
- Brini, F., Hanin, M., Lumbreras, V., Amara, I., Khoudi, H., Hassairi, A., Pagès, M., & Masmoudi, K. (2007). Overexpression of wheat dehydrin DHN-5 enhances tolerance to salt and osmotic stress in *Arabidopsis thaliana*. *Plant Cell Reports*, 26(11), 2017–2026. <https://doi.org/10.1007/s00299-007-0412-x>
- Burrieza, H. P., Koyro, H. W., Tosar, L. M., Kobayashi, K., & Maldonado, S. (2012). High salinity induces dehydrin accumulation in *Chenopodium quinoa* Willd. cv. Hualhuas embryos. *Plant and Soil*, 354(1–2), 69–79. <https://doi.org/10.1007/s11104-011-1045-y>
- Burrieza, H. P., Rizzo, A. J., Moura Vale, E., Silveira, V., & Maldonado, S. (2019). Shotgun proteomic analysis of quinoa seeds reveals novel lysine-rich seed storage globulins. *Food Chemistry*, 293,

299–306. <https://doi.org/10.1016/j.foodchem.2019.04.098>

- Burrieza, H. P., Rizzo, A. J., & Pérez, O. E. (2020). Quinoa does not contain prolamins. Comments on “Quinoa protein: Composition, structure and functional properties”, Dakhili et al. (2019). *Food Chemistry*, 325(January), 126934. <https://doi.org/10.1016/j.foodchem.2020.126934>
- Burrieza, H. P., Sanguinetti, A., Michieli, C. T., Bertero, H. D., & Maldonado, S. (2016). Death of embryos from 2300-year-old quinoa seeds found in an archaeological site. *Plant Science*, 253, 107–117. <https://doi.org/10.1016/j.plantsci.2016.10.001>
- Cai-Hong Pang, & Wang, B.-S. (2010). Role of Ascorbate Peroxidase and Glutathione Reductase in Ascorbate–Glutathione Cycle and Stress Tolerance in Plants. In *Ascorbate-Glutathione Pathway and Stress Tolerance in Plants* (pp. 91–113). <https://doi.org/10.1007/978-90-481-9404-9>
- Cai, Z. Q., & Gao, Q. (2020). Comparative physiological and biochemical mechanisms of salt tolerance in five contrasting highland quinoa cultivars. *BMC Plant Biology*, 20(1). <https://doi.org/10.1186/s12870-020-2279-8>
- Candat, A., Paszkiewicz, G., Neveu, M., Gautier, R., Logan, D. C., Avelange-Macherel, M.-H. H., & Macherel, D. (2014). The ubiquitous distribution of late embryogenesis abundant proteins across cell compartments in Arabidopsis offers tailored protection against abiotic stress. *The Plant Cell*, 26(7), 3148–3166. <https://doi.org/10.1105/tpc.114.127316>
- Cattivelli, L., & Bartels, D. (1990). Molecular cloning and characterization of cold-regulated genes in barley. *Plant Physiology*, 93(4), 1504–1510. <https://doi.org/10.1104/pp.93.4.1504>
- Catusse, J., Strub, J. M., Job, C., Van Dorsselaer, A., & Job, D. (2008). Proteome-wide characterization of sugarbeet seed vigor and its tissue specific expression. *Proceedings of the National Academy of Sciences of the United States of America*, 105(29), 10262–10267. <https://doi.org/10.1073/pnas.0800585105>
- Causin, H. F., Bordón, D. A. E., & Burrieza, H. P. (2020). Salinity tolerance mechanisms during germination and early seedling growth in *Chenopodium quinoa* Wild. genotypes with different sensitivity to saline stress. *Environmental and Experimental Botany*, 172, 103995. <https://doi.org/10.1016/j.envexpbot.2020.103995>
- Chandra Babu, R., Shashidhar, H. E., Lilley, J. M., Thanh, N. D., Ray, J. D., Sadasivam, S., Sarkarung, S., O'Toole, J. C., & Nguyen, H. T. (2001). Variation in root penetration ability, osmotic adjustment and dehydration tolerance among accessions of rice adapted to rainfed lowland and upland ecosystems. *Plant Breeding*, 120(3), 233–238. <https://doi.org/10.1046/j.1439-0523.2001.00578.x>
- Chase, M. W., Christenhusz, M. J. M., Fay, M. F., Byng, J. W., Judd, W. S., Soltis, D. E., Mabberley, D. J., Sennikov, A. N., Soltis, P. S., Stevens, P. F., Briggs, B., Brockington, S., Chautems, A., Clark, J. C., Conran, J., Haston, E., Möller, M., Moore, M., Olmstead, R., ... Weber, A. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, 181(1), 1–20. <https://doi.org/10.1111/boj.12385>
- Chatelain, E., Hundertmark, M., Leprince, O., Gall, S. Le, Satour, P., Deligny-Penninck, S., Rogniaux, H., & Buitink, J. (2012). Temporal profiling of the heat-stable proteome during late maturation of *Medicago truncatula* seeds identifies a restricted subset of late embryogenesis abundant proteins associated with longevity. *Plant, Cell and Environment*, 35(8), 1440–1455. <https://doi.org/10.1111/j.1365-3040.2012.02501.x>
- Chen, R. gang, Jing, H., Guo, W. li, Wang, S. Bin, Ma, F., Pan, B. G., & Gong, Z. H. (2015). Silencing of dehydrin CaDHN1 diminishes tolerance to multiple abiotic stresses in *Capsicum annuum* L. *Plant Cell Reports*, 34(12), 2189–2200. <https://doi.org/10.1007/s00299-015-1862-1>
- Chen, Y., Li, C., Zhang, B., Yi, J., Yang, Y., Kong, C., Lei, C., & Gong, M. (2019). The role of the Late Embryogenesis-Abundant (LEA) protein family in development and the abiotic stress response: A comprehensive expression analysis of potato (*Solanum Tuberosum*). *Genes*, 10(2), 1–16. <https://doi.org/10.3390/genes10020148>
- Cheng, Z., Targolli, J., Huang, X., & Wu, R. (2002). Wheat LEA genes, PMA80 and PMA1959,

- enhance dehydration tolerance of transgenic rice (*Oryza sativa* L.). *Molecular Breeding*, 10(1–2), 71–82. <https://doi.org/10.1023/A:1020329401191>
- Chinnusamy, V., Schumaker, K., & Zhu, J. K. (2004). Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. *Journal of Experimental Botany*, 55(395), 225–236. <https://doi.org/10.1093/jxb/erh005>
- Close, T. J. (1997). Dehydrins: A commonality in the response of plants to dehydration and low temperature. *Physiologia Plantarum*, 100(2), 291–296. <https://doi.org/10.1034/j.1399-3054.1997.1000210.x>
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21(18), 3674–3676. <https://doi.org/10.1093/bioinformatics/bti610>
- Costa, M. C. D., Artur, M. A. S., Maia, J., Jonkheer, E., Derks, M. F. L., Nijveen, H., Williams, B., Mundree, S. G., Jiménez-Gómez, J. M., Hesselink, T., Schijlen, E. G. W. M., Ligterink, W., Oliver, M. J., Farrant, J. M., & Hilhorst, H. W. M. (2017). A footprint of desiccation tolerance in the genome of *Xerophyta viscosa*. *Nature Plants*, 3(March). <https://doi.org/10.1038/nplants.2017.38>
- Damerval, C., De Vienne, D., Zivy, M., & Thiellement, H. (1986). Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. *Electrophoresis*, 7(1), 52–54. <https://doi.org/10.1002/elps.1150070108>
- Delahaie, J., Hundertmark, M., Bove, J., Leprince, O., Rogniaux, H., & Buitink, J. (2013). LEA polypeptide profiling of recalcitrant and orthodox legume seeds reveals ABI3-regulated LEA protein abundance linked to desiccation tolerance. *Journal of Experimental Botany*, 64(14), 4559–4573. <https://doi.org/10.1093/jxb/ert274>
- Detjens, A. C., & Goldberger, J. R. (2015). *Results from a Statewide Survey about the Experiences and Perspectives of Certified Organic Producers GROWING QUINOA IN WASHINGTON STATE*.
- Dini, A., Rastrelli, L., Saturnino, P., & Schettino, O. (1992). A compositional study of *Chenopodium quinoa* seeds. *Food / Nahrung*, 36(4), 400–404. <https://doi.org/10.1002/food.19920360412>
- Dirk, L. M. A., Abdel, C. G., Ahmad, I., Neta, I. C. S., Pereira, C. C., Pereira, F. E. C. B., Unêda-Trevisoli, S. H., Pinheiro, D. G., & Downie, A. B. (2020). Late embryogenesis abundant protein–client protein interactions. In *Plants* (Vol. 9, Issue 7, pp. 1–35). <https://doi.org/10.3390/plants9070814>
- Distler, U., Kuharev, J., Navarro, P., Levin, Y., Schild, H., & Tenzer, S. (2014). Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nature Methods*, 11(2), 167–170. <https://doi.org/10.1038/nmeth.2767>
- Distler, U., Kuharev, J., Navarro, P., & Tenzer, S. (2016). Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nature Protocols*, 11(4), 795–812. <https://doi.org/10.1038/nprot.2016.042>
- Dure, L., Crouch, M., Harada, J., Ho, T. H. D., 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>
- Dure, L., & Galau, G. a. (1981). Developmental Biochemistry of Cottonseed Embryogenesis and Germination 1. *Plant Physiology*, 68(1), 187–194. <https://doi.org/10.1007/BF01578378>
- Eastmond, P. J. (2007). Monodehydroascorbate reductase4 is required for seed storage oil hydrolysis and postgerminative growth in *Arabidopsis*. *Plant Cell*, 19(4), 1376–1387. <https://doi.org/10.1105/tpc.106.043992>
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- FAO. (2022a). *Salt-affected soils*. <https://www.fao.org/soils-portal/soil-management/manejo-de-suelos-problematicos/suelos-afectados-por-salinidad/more-information-on-salt-affected-soils/es/>

- FAO. (2022b). *Technical issues of salt-affected soils*. <https://www.fao.org/soils-portal/soil-management/manejo-de-suelos-problematicos/suelos-afectados-por-salinidad/more-information-on-salt-affected-soils/technical-issues/es/>
- FAO, & CIRAD. (2015). *State of the Art Report of Quinoa in the World in 2013* (D. Bazile, D. Bertero, & C. Nieto (eds.)).
- Frandsen, G. I., Mundy, J., & Tzen, J. T. C. (2001). Oil bodies and their associated proteins, oleosin and caleosin. *Physiologia Plantarum*, 112(3), 301–307. <https://doi.org/10.1034/j.1399-3054.2001.1120301.x>
- Fuchs, S. (2020). *GRAVY calculator*. <http://www.gravy-calculator.de/>
- Fuentes, F. F., Bazile, D., Bhargava, A., & Martínez, E. A. (2012). Implications of farmers' seed exchanges for on-farm conservation of quinoa, as revealed by its genetic diversity in Chile. *Journal of Agricultural Science*, 150(6), 702–716. <https://doi.org/10.1017/S0021859612000056>
- Fuentes, F. F., Martinez, E. A., Hinrichsen, P. V., Jellen, E. N., & Maughan, P. J. (2009). Assessment of genetic diversity patterns in Chilean quinoa (*Chenopodium quinoa* Willd.) germplasm using multiplex fluorescent microsatellite markers. *Conservation Genetics*, 10(2), 369–377. <https://doi.org/10.1007/s10592-008-9604-3>
- Furuki, T., Shimizu, T., Chakrabortee, S., Yamakawa, K., Hatanaka, R., Takahashi, T., Kikawada, T., Okuda, T., Mihara, H., Tunnacliffe, A., & Sakurai, M. (2012). Effects of Group 3 LEA protein model peptides on desiccation-induced protein aggregation. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1824(7), 891–897. <https://doi.org/10.1016/j.bbapap.2012.04.013>
- Galau, G. A., Wang, H. Y. C., & Hughes, D. W. (1992). COTTON LEA4 (D19) AND LEA2 (D132) GROUP-1 LEA GENES ENCODING WATER STRESS-RELATED PROTEINS CONTAINING A 20-AMINO ACID MOTIF. *PLANT PHYSIOLOGY*, 99(2), 783–788. <https://doi.org/10.1104/pp.99.2.783>
- Galau, G. A., Wang, H. Y., & Hughes, D. W. (1993). Cotton Lea5 and Lea14 encode atypical late embryogenesis-abundant proteins. *Plant Physiology*, 101(2), 695–696. <https://doi.org/10.1104/pp.101.2.695>
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In J. M. Walker (Ed.), *Protein Identification and Analysis Tools on the ExPASy Server* (pp. 571–607). Humana Press.
- Ginsawaeng, O., Gorka, M., Erban, A., Heise, C., Brueckner, F., Hoefgen, R., Kopka, J., Skirycz, A., Hinch, D. K., & Zuther, E. (2021). Characterization of the heat-stable proteome during seed germination in arabidopsis with special focus on LEA proteins. *International Journal of Molecular Sciences*, 22(15). <https://doi.org/10.3390/ijms22158172>
- Goel, A., Goel, A. K., & Sheoran, I. S. (2003). Changes in oxidative stress enzymes during artificial ageing in cotton (*Gossypium hirsutum* L.) seeds. *Journal of Plant Physiology*, 160(9), 1093–1100. <https://doi.org/10.1078/0176-1617-00881>
- Gómez-Pando, L. R., Álvarez-Castro, R., & Eguiluz-de la Barra, A. (2010). Effect of salt stress on peruvian germplasm of chenopodium quinoa willd.: A promising crop. *Journal of Agronomy and Crop Science*, 196(5), 391–396. <https://doi.org/10.1111/j.1439-037X.2010.00429.x>
- González, J. A., Roldán, A., Gallardo, M., Escudero, T., & Prado, F. E. (1989). Quantitative determinations of chemical compounds with nutritional value from inca crops: *Chenopodium quinoa* ('quinoa'). *Plant Foods for Human Nutrition*, 39(4), 331–337. <https://doi.org/10.1007/BF01092070>
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1), 1178–1186. <https://doi.org/10.1093/nar/gkr944>
- Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talón, M., Dopazo, J., & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, 36(10), 3420–3435.

<https://doi.org/10.1093/nar/gkn176>

- Goyal, K., Tisi, L., Basran, A., Browne, J., Burnell, A., Zurdo, J., & Tunnacliffe, A. (2003). Transition from natively unfolded to folded state induced by desiccation in an anhydrobiotic nematode protein. *Journal of Biological Chemistry*, 278(15), 12977–12984. <https://doi.org/10.1074/jbc.M212007200>
- 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>
- Grelet, J., Benamar, A., Teyssier, E., Avelange-Macherel, M. H., Grunwald, D., & Macherel, D. (2005). Identification in pea seed mitochondria of a late-embryogenesis abundant protein able to protect enzymes from drying. *Plant Physiology*, 137(1), 157–167. <https://doi.org/10.1104/pp.104.052480>
- Gu, H., Jia, Y., Wang, X., Chen, Q., Shi, S., Ma, L., Zhang, J., Zhang, H., & Ma, H. (2012). Identification and characterization of a LEA family gene CarLEA4 from chickpea (*Cicer arietinum* L.). *Molecular Biology Reports*, 39(4), 3565–3572. <https://doi.org/10.1007/s11033-011-1130-6>
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321. <https://doi.org/10.1093/sysbio/syq010>
- Hager, A. S., Mäkinen, O. E., & Arendt, E. K. (2014). Amylolytic activities and starch reserve mobilization during the germination of quinoa. *European Food Research and Technology*, 239(4), 621–627. <https://doi.org/10.1007/s00217-014-2258-0>
- Harada, J. J., DeLisle, A. J., Baden, C. S., & Crouch, M. L. (1989). Unusual sequence of an abscisic acid-inducible mRNA which accumulates late in *Brassica napus* seed development. *Plant Molecular Biology*, 12(4), 395–401. <https://doi.org/10.1007/BF00017579>
- Hasegawa, P. M. (2013). Sodium (Na⁺) homeostasis and salt tolerance of plants. *Environmental and Experimental Botany*, 92, 19–31. <https://doi.org/10.1016/j.envexpbot.2013.03.001>
- He, C., Liu, X., Teixeira da Silva, J. A., Wang, H., Peng, T., Zhang, M., Si, C., Yu, Z., Tan, J., Zhang, J., Luo, J., & Duan, J. (2021). Characterization of LEA genes in *Dendrobium officinale* and one Gene in induction of callus. *Journal of Plant Physiology*, 258–259. <https://doi.org/10.1016/j.jplph.2020.153356>
- Hincha, D. K., Zuther, E., & Popova, A. V. (2021). Stabilization of dry sucrose glasses by four lea₄ proteins from *Arabidopsis thaliana*. *Biomolecules*, 11(5). <https://doi.org/10.3390/biom11050615>
- Honjoh, K. I., Matsumoto, H., Shimizu, H., Ooyama, K., Tanaka, K., Oda, Y., Takata, R., Joh, T., Suga, K., Miyamoto, T., Iio, M., & Hatano, S. (2000). Cryoprotective activities of group 3 late embryogenesis abundant proteins from *Chlorella vulgaris* C-27. In *Biosci. Biotechnol. Biochem* (Vol. 64, Issues 0916-8451 (Print), pp. 1656–1663). <https://doi.org/10.1271/bbb.64.1656>
- Hsing, Y. I. C., Chen, Z. Y., Shih, M. D., Hsieh, J. S., & Chow, T. Y. (1995). Unusual sequences of group 3 LEA mRNA inducible by maturation or drying in soybean seeds. *Plant Molecular Biology*, 29(4), 863–868. <https://doi.org/10.1007/BF00041175>
- Hundertmark, M., & Hincha, D. K. (2008). LEA (Late Embryogenesis Abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics*, 9(1), 1–22. <https://doi.org/10.1186/1471-2164-9-118>
- Jacobsen, S.-E., Mujica, A., & Jensen, C. R. (2003). The Resistance of Quinoa (*Chenopodium quinoa* Willd.) to Adverse Abiotic Factors. *Food Reviews International*, 19(1–2), 99–109. <https://doi.org/10.1081/FRI-120018872>
- Jarvis, D. E., Ho, Y. S., Lightfoot, D. J., Schmöckel, S. M., Li, B., Borm, T. J. A., Ohyanagi, H., Mineta, K., Michell, C. T., Saber, N., Kharbatia, N. M., Rupper, R. R., Sharp, A. R., Dally, N., Boughton, B. A., Woo, Y. H., Gao, G., Schijlen, E. G. W. M., Guo, X., ... Tester, M. (2017). The genome of *Chenopodium quinoa*. *Nature*, 542(7641), 307–312. <https://doi.org/10.1038/nature21370>
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI

- BLAST: a better web interface. *Nucleic Acids Research*, 36(Web Server issue), 5–9. <https://doi.org/10.1093/nar/gkn201>
- Joshi, R. (2018). Role of enzymes in seed germination. *International Journal of Creative Research Thoughts*, 6(2), 1481–1485.
- Keisham, M., Mukherjee, S., & Bhatla, S. C. (2018). Mechanisms of sodium transport in plants—Progresses and challenges. *International Journal of Molecular Sciences*, 19(3). <https://doi.org/10.3390/ijms19030647>
- Khan, S., Jabeen, R., Deeba, F., Waheed, U., Khanum, P., & Iqbal, N. (2021). Heat Shock Proteins: Classification, Functions and Expressions in Plants during Environmental Stresses. *Journal of Bioresource Management*, 8(2), 85–97. <https://corescholar.libraries.wright.edu/cgi/viewcontent.cgi?article=1281&context=jbm>
- Kijak, H., & Ratajczak, E. (2020). What do we know about the genetic basis of seed desiccation tolerance and longevity? *International Journal of Molecular Sciences*, 21(10). <https://doi.org/10.3390/ijms21103612>
- Kim, H.-S., Lee, J. H., Kim, J. J., Kim, C.-H., Jun, S.-S., & Hong, Y.-N. (2005). Molecular and functional characterization of CaLEA6, the gene for a hydrophobic LEA protein from *Capsicum annuum*. *Gene*, 344, 115–123. <https://doi.org/10.1016/j.gene.2004.09.012>
- Kiyosue, T., Yamaguchi-Shinozaki, K., Shinozaki, K., Higashi, K., Satoh, S., Kamada, H., & Harada, H. (1992). Isolation and characterization of a cDNA that encodes ECP31, an embryogenic-cell protein from carrot. *Plant Molecular Biology*, 19(2), 239–249. <https://doi.org/10.1007/BF00027345>
- Knox-Brown, P., Rindfleisch, T., Günther, A., Balow, K., Bremer, A., Walther, D., Miettinen, M. S., Hinch, D. K., & Thalhammer, A. (2020). Similar yet different—structural and functional diversity among *Arabidopsis thaliana* LEA₄ proteins. *International Journal of Molecular Sciences*, 21(8). <https://doi.org/10.3390/ijms21082794>
- Kotak, S., Larkindale, J., Lee, U., von Koskull-Döring, P., Vierling, E., & Scharf, K. D. (2007). Complexity of the heat stress response in plants. *Current Opinion in Plant Biology*, 10(3), 310–316. <https://doi.org/10.1016/j.pbi.2007.04.011>
- Kotak, S., Vierling, E., Bäumlein, H., & Von Koskull-Döring, P. (2007). A novel transcriptional cascade regulating expression of heat stress proteins during seed development of *Arabidopsis*. *Plant Cell*, 19(1), 182–195. <https://doi.org/10.1105/tpc.106.048165>
- Koubaa, S., Bremer, A., Hinch, D. K., & Brini, F. (2019). Structural properties and enzyme stabilization function of the intrinsically disordered LEA₄ protein TdLEA3 from wheat. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-019-39823-w>
- Koyro, H.-W., Ahmad, P., & Geissler, N. (2012). Abiotic Stress Responses in Plants: An Overview. *Environmental Adaptations and Stress Tolerance of Plants in the Era of Climate Change*, 1–515. <https://doi.org/10.1007/978-1-4614-0815-4>
- Koyro, H.-W., & Eisa, S. S. (2008). Effect of salinity on composition, viability and germination of seeds of *Chenopodium quinoa* Willd. *Plant and Soil*, 302(1–2), 79–90. <https://doi.org/10.1007/s11104-007-9457-4>
- Kozioł, M. J. (1992). Chemical Composition and Nutritional Evaluation of Quinoa (*Chenopodium quinoa* Willd.). *Journal of Food Composition and Analysis*, 5, 35–68.
- Kronzucker, H. J., Coskun, D., Schulze, L. M., Wong, J. R., & Britto, D. T. (2013). Sodium as nutrient and toxicant. *Plant and Soil*, 369(1–2), 1–23. <https://doi.org/10.1007/s11104-013-1801-2>
- Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S., & Gascuel, O. (2019). NGPhylogeny.fr: New generation phylogenetic services for non-specialists. *Nucleic Acids Research*, 47(W1), W260–W265. <https://doi.org/10.1093/nar/gkz303>
- Leprince, O., Pellizzaro, A., Berriri, S., & Buitink, J. (2017). Late seed maturation: Drying without dying. *Journal of Experimental Botany*, 68(4), 827–841. <https://doi.org/10.1093/jxb/erw363>

- Letunic, I., & Bork, P. (2021). Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49(W1), W293–W296. <https://doi.org/10.1093/nar/gkab301>
- Li, H., Yu, W., Dhital, S., Gidley, M. J., & Gilbert, R. G. (2019). Starch branching enzymes contributing to amylose and amylopectin fine structure in wheat. *Carbohydrate Polymers*, 224(April), 115185. <https://doi.org/10.1016/j.carbpol.2019.115185>
- Liang, Jing, Zhou, M., Zhou, X., Jin, Y., Xu, M., & Lin, J. (2013). JcLEA, a novel LEA-like protein from *Jatropha curcas*, confers a high level of tolerance to dehydration and salinity in *Arabidopsis thaliana*. *PLoS ONE*, 8(12), e83056. <https://doi.org/10.1371/journal.pone.0083056>
- Liang, Junjun, Deng, G., Long, H., Pan, Z., Wang, C., Cai, P., Xu, D., Nima, Z. X., & Yu, M. (2012). Virus-induced silencing of genes encoding LEA protein in Tibetan hulless barley (*Hordeum vulgare* ssp. *vulgare*) and their relationship to drought tolerance. *Molecular Breeding*, 30(1), 441–451. <https://doi.org/10.1007/s11032-011-9633-3>
- López-Ribera, I., La Paz, J. L., Repiso, C., García, N., Miquel, M., Hernández, M. L., Martínez-Rivas, J. M., & Vicient, C. M. (2014). The evolutionary conserved oil body associated protein OBAP1 participates in the regulation of oil body size. *Plant Physiology*, 164(3), 1237–1249. <https://doi.org/10.1104/pp.113.233221>
- Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, 50(W1), W276–W279. <https://doi.org/10.1093/nar/gkac240>
- Magwanga, R. O., Lu, P., Kirungu, J. N., Lu, H., Wang, X., Cai, X., Zhou, Z., Zhang, Z., Salih, H., Wang, K., & Liu, F. (2018). Characterization of the late embryogenesis abundant (LEA) proteins family and their role in drought stress tolerance in upland cotton. *BMC Genetics*, 19(1), 1–31. <https://doi.org/10.1186/s12863-017-0596-1>
- Mahalingam, R. (2017). Shotgun proteomics of the barley seed proteome. *BMC Genomics*, 18(1), 1–11. <https://doi.org/10.1186/s12864-016-3408-5>
- Maitra, N., & Cushman, J. C. (1994). Isolation and characterization of a drought-induced soybean cDNA encoding a D95 family late-embryogenesis-abundant protein. *Plant Physiology*, 106(2), 805–806. <http://www.ncbi.nlm.nih.gov/pubmed/7991700>
- Mäkinen, O. E., Hager, A.-S. S., & Arendt, E. K. (2014). Localisation and development of proteolytic activities in quinoa (*Chenopodium quinoa*) seeds during germination and early seedling growth. *Journal of Cereal Science*, 60(3), 484–489. <https://doi.org/10.1016/j.jcs.2014.08.009>
- Melgar, A. E., & Zelada, A. M. (2021). Evolutionary analysis of angiosperm dehydrin gene family reveals three orthologues groups associated to specific protein domains. *Scientific Reports*, 11(1), 1–12. <https://doi.org/10.1038/s41598-021-03066-5>
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49(D1), D412–D419. <https://doi.org/10.1093/nar/gkaa913>
- Mönke, G., Seifert, M., Keilwagen, J., Mohr, M., Grosse, I., Hähnel, U., Junker, A., Weisshaar, B., Conrad, U., Bäumlein, H., & Altschmied, L. (2012). Toward the identification and regulation of the *Arabidopsis thaliana* ABI3 regulon. *Nucleic Acids Research*, 40(17), 8240–8254. <https://doi.org/10.1093/nar/gks594>
- Morales, A., Zurita-Silva, A., Maldonado, J., & Silva, H. (2017). Transcriptional responses of chilean quinoa (*Chenopodium quinoa* Willd.) under water deficit conditions uncovers ABA-independent expression patterns. *Frontiers in Plant Science*, 8(March), 1–13. <https://doi.org/10.3389/fpls.2017.00216>
- Munns, R., & Tester, M. (2008). Mechanisms of salinity tolerance. *Annual Review of Plant Biology*, 59, 651–681. <https://doi.org/10.1146/annurev.arplant.59.032607.092911>
- Murphy, D. J., & Vance, J. (1999). Mechanisms of lipid-body formation. *Trends in Biochemical*

- Sciences*, 24(3), 109–115. [https://doi.org/10.1016/S0968-0004\(98\)01349-8](https://doi.org/10.1016/S0968-0004(98)01349-8)
- Nakayama, K., Okawa, K., Kakizaki, T., & Inaba, T. (2008). Evaluation of the protective activities of a late embryogenesis abundant (LEA) related protein, Cor15am, during various stresses in Vitro. *Bioscience, Biotechnology and Biochemistry*, 72(6), 1642–1645. <https://doi.org/10.1271/bbb.80214>
- National Research Council. (1989). Lost Crops of the Incas: Little Known Plants of the Andes with promise for Worldwide Cultivation. In F. R. R. Bostid (Ed.), *Lost Crop of the Incas* (pp. 148–161). National Academy Press.
- NDong, C., Danyluk, J., Wilson, K. E., Pocock, T., Huner, N. P. A., & Sarhan, F. (2002). Cold-regulated cereal chloroplast late embryogenesis abundant-like proteins. Molecular characterization and functional analyses. *Plant Physiology*, 129(3), 1368–1381. <https://doi.org/10.1104/pp.001925>
- Negacz, K., Malek, Ž., de Vos, A., & Vellinga, P. (2022). Saline soils worldwide: Identifying the most promising areas for saline agriculture. *Journal of Arid Environments*, 203. <https://doi.org/10.1016/j.jaridenv.2022.104775>
- Olías, R., Eljakaoui, Z., Li, J., De Morales, P. A., Marín-Manzano, M. C., Pardo, J. M., & Belver, A. (2009). The plasma membrane Na⁺/H⁺ antiporter SOS1 is essential for salt tolerance in tomato and affects the partitioning of Na⁺ between plant organs. *Plant, Cell and Environment*, 32(7), 904–916. <https://doi.org/10.1111/j.1365-3040.2009.01971.x>
- Oliveros, J. C. (2015). *An Interactive Tool for Comparing Lists with Venn's Diagrams*. Csic. <https://bioinfogp.cnb.csic.es/tools/venny/index.html>
- Olvera-Carrillo, Y., Campos, F., Reyes, J. L., Garcíarrubio, 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>
- Panuccio, M. R., Jacobsen, S. E., Akhtar, S. S., & Muscolo, A. (2014). Effect of saline water on seed germination and early seedling growth of the halophyte quinoa. *AoB PLANTS*, 6, 1–18. <https://doi.org/10.1093/aobpla/plu047>
- Parihar, P., Singh, S., Singh, R., Singh, V. P., & Prasad, S. M. (2015). Effect of salinity stress on plants and its tolerance strategies: a review. *Environmental Science and Pollution Research*, 22(6), 4056–4075. <https://doi.org/10.1007/s11356-014-3739-1>
- Park, J. A., Cho, S. K., Kim, J. E., Chung, H. S., Hong, J. P., Hwang, B., Hong, C. B., & Kim, W. T. (2003). Isolation of cDNAs differentially expressed in response to drought stress and characterization of the Ca-LEAL1 gene encoding a new family of atypical LEA-like protein homologue in hot pepper (*Capsicum annuum* L. cv. Pukang). *Plant Science*, 165(3), 471–481. [https://doi.org/10.1016/S0168-9452\(03\)00165-1](https://doi.org/10.1016/S0168-9452(03)00165-1)
- Patiranage, D. S. R., Rey, E., Emrani, N., Wellman, G., Schmid, K., Schmöckel, S. M., Tester, M., & Jung, C. (2022). Genome-wide association study in the pseudocereal quinoa reveals selection pattern typical for crops with a short breeding history. *ELife*, 11, 2020.12.03.410050. <https://doi.org/https://doi.org/10.1101/2020.12.03.410050>
- Piovesan, D., Walsh, I., Minervini, G., & Tosatto, S. C. E. (2017). FELLs: Fast estimator of latent local structure. *Bioinformatics*, 33(12), 1889–1891. <https://doi.org/10.1093/bioinformatics/btx085>
- Planella, M. T., Lopez, M. L., & Bruno, M. C. (2015). Domestication and Prehistoric Distribution. In *State of the Art Report on Quinoa Around the World in 2013*. https://scholar.dickinson.edu/faculty_publications/1027
- Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., & Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Research*, 46(W1), W200–W204. <https://doi.org/10.1093/nar/gky448>
- Präger, A., Munz, S., Nkebiwe, P. M., Mast, B., & Graeff-Hönniger, S. (2018). Yield and quality characteristics of different quinoa (*Chenopodium quinoa* Willd.) cultivars grown under field conditions in southwestern Germany. *Agronomy*, 8(10), 197.

<https://doi.org/10.3390/agronomy8100197>

- Prego, I., Maldonado, S., & Otegui, M. (1998). Seed structure and localization of reserves in *Chenopodium quinoa*. *Annals of Botany*, 82(4), 481–488. <https://doi.org/10.1006/anbo.1998.0704>
- Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., Silman, I., & Sussman, J. L. (2005). FoldIndex©: A simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics*, 21(16), 3435–3438. <https://doi.org/10.1093/bioinformatics/bti537>
- R Core Team. (2016). *R: A language and environment for statistical computing*. 2, 1–12.
- Rafik, S., Chaoui, M., Assabban, Y., Jazi, S., Choukr-Allah, R., El Gharouss, M., & Hirich, A. (2020). Quinoa value chain, adoption, and market assessment in Morocco. *Environmental Science and Pollution Research*, 1–12. <https://doi.org/10.1007/s11356-020-11375-x>
- Ranhotra, G. S., Gelroth, J. A., Glaser, B. K., Lorenz, K. J., & Johnson, D. L. (1993). Composition and protein nutritional quality of quinoa. In *Cereal Chemistry* (Vol. 70, Issue 3, pp. 303–305).
- Reis, R. S., Vale, E. M., Sousa, K. R., Santa-Catarina, C., & Silveira, V. (2021). Pretreatment free of 2,4-dichlorophenoxyacetic acid improves the differentiation of sugarcane somatic embryos by affecting the hormonal balance and the accumulation of reserves. *Plant Cell, Tissue and Organ Culture*, 145(1), 101–115. <https://doi.org/10.1007/s11240-020-01995-z>
- Repo-Carrasco, R., Espinoza, C., & Jacobsen, S. E. (2003). Nutritional value and use of the andean crops quinoa (*Chenopodium quinoa*) and kañiwa (*Chenopodium pallidicaule*). *Food Reviews International*, 19(1–2), 179–189. <https://doi.org/10.1081/FRI-120018884>
- Reyes, J. L., Rodrigo, M. J., Colmenero-Flores, J. M., Gil, J. V., Garay-Arroyo, A., Campos, F., Salamini, F., Bartels, D., & Covarrubias, A. A. (2005). Hydrophilins from distant organisms can protect enzymatic activities from water limitation effects in vitro. *Plant, Cell and Environment*, 28(6), 709–718. <https://doi.org/10.1111/j.1365-3040.2005.01317.x>
- Rhee, S. Y., Beavis, W., Berardini, T. Z., Chen, G., Dixon, D., Doyle, A., Garcia-Hernandez, M., Huala, E., Lander, G., Montoya, M., Miller, N., Mueller, L. A., Mundodi, S., Reiser, L., Tacklind, J., Weems, D. C., Wu, Y., Xu, I., Yoo, D., ... Zhang, P. (2003). The Arabidopsis Information Resource (TAIR): A model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. In *Nucleic Acids Research* (Vol. 31, Issue 1, pp. 224–228). <https://doi.org/10.1093/nar/gkg076>
- Risi, J. ., & Galwey, N. (1984). The *Chenopodium* grains of the Andes: Inca crops for modern agriculture. *Adv Appl Bot*, 10, 145–216. <https://sci-hub.io/http://agris.fao.org/agris-search/search.do?recordID=US201301421301>
- Rizzo, A.J. (2016). *Clonado de una LEA-5 de Quinoa (Chenopodium quinoa Willd.): caracterización y expresión en condiciones de estrés abiótico*. Universidad de Buenos Aires.
- Rizzo, Axel Joel, Palacios, M. B., Vale, E. M., Zelada, A. M., Silveira, V., & Burrieza, H. P. (2023). Snapshot of four mature quinoa (*Chenopodium quinoa*) seeds: a shotgun proteomics analysis with emphasis on seed maturation, reserves and early germination. *Physiology and Molecular Biology of Plants*, 29(3), 319–334. <https://doi.org/10.1007/s12298-023-01295-8>
- Rodríguez, M., Canales, E., & Borrás-Hidalgo, O. (2005). Molecular aspects of abiotic stress in plants. *Biotecnología Aplicada*, 22(1), 1–10. <https://doi.org/10.1080/07352689709701950>
- Romo, S., Labrador, E., & Dopico, B. (2001). Water stress-regulated gene expression in *Cicer arietinum* seedlings and plants. *Plant Physiology and Biochemistry*, 39(11), 1017–1026. [https://doi.org/10.1016/S0981-9428\(01\)01318-3](https://doi.org/10.1016/S0981-9428(01)01318-3)
- Rorat, T. (2006). Plant dehydrins--tissue location, structure and function. *Cellular & Molecular Biology Letters*, 11(4), 536–556. <https://doi.org/10.2478/s11658-006-0044-0>
- Ruales, J., & Nair, B. M. (1992). Nutritional quality of the protein in quinoa (*Chenopodium quinoa*, Willd) seeds. *Plant Foods for Human Nutrition*, 42(1), 1–11. <https://doi.org/10.1007/BF02196067>

- Rubio, F., Nieves-Cordones, M., Horie, T., Shabala, S., Nieves-Cordones, M., Horie, T., Shabala, S., Nieves-Cordones, M., Horie, T., Shabala, S., Nieves-Cordones, M., Horie, T., Shabala, S., ... Shabala, S. (2020). Doing 'business as usual' comes with a cost: evaluating energy cost of maintaining plant intracellular K⁺ homeostasis under saline conditions. *New Phytologist*, 225(3), 1097–1104. <https://doi.org/10.1111/nph.15852>
- Ruiz-Carrasco, K., Antognoni, F., Coulibaly, A. K., Lizardi, S., Covarrubias, A., Martínez, E. A., Molina-Montenegro, M. A., Biondi, S., & Zurita-Silva, A. (2011). Variation in salinity tolerance of four lowland genotypes of quinoa (*Chenopodium quinoa* Willd.) as assessed by growth, physiological traits, and sodium transporter gene expression. *Plant Physiology and Biochemistry*, 49(11), 1333–1341. <https://doi.org/10.1016/j.plaphy.2011.08.005>
- Ruiz, K. B., Biondi, S., Martínez, E. A., Orsini, F., Antognoni, F., & Jacobsen, S. E. (2016). Quinoa – a Model Crop for Understanding Salt-tolerance Mechanisms in Halophytes. *Plant Biosystems*, 150(2), 357–371. <https://doi.org/10.1080/11263504.2015.1027317>
- Ruiz, Karina B., Aloisi, I., Del Duca, S., Canelo, V., Torrigiani, P., Silva, H., & Biondi, S. (2016). Salares versus coastal ecotypes of quinoa: Salinity responses in Chilean landraces from contrasting habitats. *Plant Physiology and Biochemistry*, 101, 1–13. <https://doi.org/10.1016/j.plaphy.2016.01.010>
- Ruiz, Karina B., Biondi, S., Oses, R., Acuña-Rodríguez, I. S., Antognoni, F., Martínez-Mosqueira, E. A., Coulibaly, A., Canahua-Murillo, A., Pinto, M., Zurita-Silva, A., Bazile, D., Jacobsen, S. E., & Molina-Montenegro, M. A. (2014). Quinoa biodiversity and sustainability for food security under climate change. A review. In *Agronomy for Sustainable Development* (Vol. 34, Issue 2, pp. 349–359). EDP Sciences. <https://doi.org/10.1007/s13593-013-0195-0>
- Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., ... Sherry, S. T. (2022). Database resources of the national center for biotechnology information. *Nucleic Acids Research*, 50(D1), D20–D26. <https://doi.org/10.1093/nar/gkab1112>
- Serrano, R., Gaxiola, R., Ríos, G., Forment, J., Vicente, O., & Ros, R. (2003). Salt Stress Proteins Identified by a Functional Approach in Yeast. *Monatshefte Fur Chemie*, 134(11), 1445–1464. <https://doi.org/10.1007/s00706-002-0606-4>
- Sestili, F., Janni, M., Doherty, A., Botticella, E., D'Ovidio, R., Masci, S., Jones, H. D., & Lafiandra, D. (2010). Increasing the amylose content of durum wheat through silencing of the SBEIIa genes. *BMC Plant Biology*, 10. <https://doi.org/10.1186/1471-2229-10-144>
- Seung, D. (2020). Amylose in starch: towards an understanding of biosynthesis, structure and function. *New Phytologist*, 228(5), 1490–1504. <https://doi.org/10.1111/nph.16858>
- Shabala, S., Hariadi, Y., & Jacobsen, S. E. (2013). Genotypic difference in salinity tolerance in quinoa is determined by differential control of xylem Na⁺ loading and stomatal density. *Journal of Plant Physiology*, 170(10), 906–914. <https://doi.org/10.1016/j.jplph.2013.01.014>
- Shimada, T. L., Hayashi, M., & Hara-Nishimura, I. (2018). Membrane dynamics and multiple functions of oil bodies in seeds and leaves. *Plant Physiology*, 176(1), 199–207. <https://doi.org/10.1104/pp.17.01522>
- Singh, K. K., & Graether, S. P. (2020). Conserved sequence motifs in the abiotic stress response protein late embryogenesis abundant 3. *PLoS ONE*, 15(8 August), 1–18. <https://doi.org/10.1371/journal.pone.0237177>
- Slade, A. J., McGuire, C., Loeffler, D., Mullenberg, J., Skinner, W., Fazio, G., Holm, A., Brandt, K. M., Steine, M. N., Goodstal, J. F., & Knauf, V. C. (2012). Development of high amylose wheat through TILLING. *BMC Plant Biology*, 12. <https://doi.org/10.1186/1471-2229-12-69>
- Stacy, R. A., Nordeng, T. W., Culiáñez-Macià, F. A., & Aalen, R. B. (1999). The dormancy-related peroxiredoxin anti-oxidant, PER1, is localized to the nucleus of barley embryo and aleurone cells. *The Plant Journal : For Cell and Molecular Biology*, 19(1), 1–8.

<http://www.ncbi.nlm.nih.gov/pubmed/10417721>

- Stanley, D., Rejzek, M., Naested, H., Smedley, M., Otero, S., Fahy, B., Thorpe, F., Nash, R. J., Harwood, W., Svensson, B., Denyer, K., Field, R. A., & Smith, A. M. (2011). The role of α -glucosidase in germinating barley grains. *Plant Physiology*, 155(2), 932–943. <https://doi.org/10.1104/pp.110.168328>
- Stikic, R., Glamoclija, D., Demin, M., Vucelic-Radovic, B., Jovanovic, Z., Milojkovic-Opsenica, D., Jacobsen, S. E., & Milovanovic, M. (2012). Agronomical and nutritional evaluation of quinoa seeds (*Chenopodium quinoa* Willd.) as an ingredient in bread formulations. *Journal of Cereal Science*, 55(2), 132–138. <https://doi.org/10.1016/j.jcs.2011.10.010>
- Su, P. H., & Li, H. M. (2008). Arabidopsis stromal 70-kD heat shock proteins are essential for plant development and important for thermotolerance of germinating seeds. *Plant Physiology*, 146(3), 1231–1241. <https://doi.org/10.1104/pp.107.114496>
- Suárez-Estrella, D., Bresciani, A., Iametti, S., Marengo, M., Pagani, M. A., & Marti, A. (2020). Effect of Sprouting on Proteins and Starch in Quinoa (*Chenopodium quinoa* Willd.). *Plant Foods for Human Nutrition*, 75(4), 635–641. <https://doi.org/10.1007/s11130-020-00864-6>
- Sun, J., Ke, J., Johnson, J. L., Nikolau, B. J., & Wurtele, E. S. (1997). Biochemical and molecular biological characterization of CAC2, the Arabidopsis thaliana gene coding for the biotin carboxylase subunit of the plastidic acetyl-coenzyme A carboxylase. *Plant Physiology*, 115(4), 1371–1383. <https://doi.org/10.1104/pp.115.4.1371>
- Sun, Z., & Henson, C. A. (1990). Degradation of native starch granules by barley α -glucosidases. *Plant Physiology*, 94(1), 320–327. <https://doi.org/10.1104/pp.94.1.320>
- Thalhammer, A., Hundertmark, M., Popova, A. V., Seckler, R., & Hinch, D. K. (2010). Interaction of two intrinsically disordered plant stress proteins (COR15A and COR15B) with lipid membranes in the dry state. *Biochimica et Biophysica Acta - Biomembranes*, 1798(9), 1812–1820. <https://doi.org/10.1016/j.bbamem.2010.05.015>
- Tolte, D., Jaquinod, M., Mangavel, C., Passirani, C., Saulnier, P., Manon, S., Teyssier, E., Payet, N., Avelange-Macherel, M. H., & Macherel, D. (2007). Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *Plant Cell*, 19(5), 1580–1589. <https://doi.org/10.1105/tpc.107.050104>
- Tunnacliffe, A., & Wise, M. J. (2007). The continuing conundrum of the LEA proteins. *Naturwissenschaften*, 94(10), 791–812. <https://doi.org/10.1007/s00114-007-0254-y>
- Van Breusegem, F., & Dat, J. F. (2006). Reactive oxygen species in plant cell death. *Plant Physiology*, 141(2), 384–390. <https://doi.org/10.1104/pp.106.078295>
- Vega-Gálvez, A., Miranda, M., Vergara, J., Uribe, E., Puente, L., & Martínez, E. A. (2010). Nutrition facts and functional potential of quinoa (*Chenopodium quinoa* Willd.), an ancient Andean grain: A review. In *Journal of the Science of Food and Agriculture* (Vol. 90, Issue 15, pp. 2541–2547). John Wiley & Sons, Ltd. <https://doi.org/10.1002/jsfa.4158>
- Verdier, J., Lalanne, D., Pelletier, S., Torres-Jerez, I., Righetti, K., Bandyopadhyay, K., Leprince, O., Chatelain, E., Vu, B. L., Gouzy, J., Gamas, P., Udvardi, M. K., & Buitink, J. (2013). A regulatory network-based approach dissects late maturation processes related to the acquisition of desiccation tolerance and longevity of medicago truncatula seeds. *Plant Physiology*, 163(2), 757–774. <https://doi.org/10.1104/pp.113.222380>
- Wang, M., Li, P., Li, C., Pan, Y., Jiang, X., Zhu, D., Zhao, Q., & Yu, J. (2014). SiLEA14, a novel atypical LEA protein, confers abiotic stress resistance in foxtail millet. *BMC Plant Biology*, 14(1), 290. <https://doi.org/10.1186/s12870-014-0290-7>
- Wehmeyer, N., & Vierling, E. (2000). The expression of small heat shock proteins in seeds responds to discrete developmental signals and suggests a general protective role in desiccation tolerance. *Plant Physiology*, 122(4), 1099–1108. <https://doi.org/10.1104/pp.122.4.1099>
- Wise, M. J. (2003). LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. *BMC Bioinformatics*, 4, 52.

<https://doi.org/10.1186/1471-2105-4-52>

- Wolkers, W. F., McCready, S., Brandt, W. F., Lindsey, G. G., & Hoekstra, F. A. (2001). Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses in vitro. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, 1544(1–2), 196–206. [https://doi.org/10.1016/S0167-4838\(00\)00220-X](https://doi.org/10.1016/S0167-4838(00)00220-X)
- Xiong, L., Schumaker, K. S., & Zhu, J. (2002). Cell Signaling during Cold, Drought, and Salt Stress. *The Plant Cell*, 14(1), 165–183. https://doi.org/10.1007/978-81-322-1542-4_11
- Xiong, L., & Zhu, J. K. (2001). Abiotic stress signal transduction in plants: Molecular and genetic perspectives. *Physiologia Plantarum*, 112(2), 152–166. <https://doi.org/10.1034/j.1399-3054.2001.1120202.x>
- Xu, D., Duan, X., Wang, B., Hong, B., Ho, T. H. D., & Wu, R. (1996). Expression of a late embryogenesis abundant protein gene, HVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. *Plant Physiology*, 110(1), 249–257. <https://doi.org/10.1104/pp.110.1.249>
- Yamaguchi-Shinozaki, K., & Shinozaki, K. (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annual Review of Plant Biology*, 57, 781–803. <https://doi.org/10.1146/annurev.arplant.57.032905.105444>
- Yang, A., Akhtar, S. S., Amjad, M., Iqbal, S., & Jacobsen, S. E. (2016). Growth and Physiological Responses of Quinoa to Drought and Temperature Stress. *Journal of Agronomy and Crop Science*, 202(6), 445–453. <https://doi.org/10.1111/jac.12167>
- Yasui, Y., Hirakawa, H., Oikawa, T., Toyoshima, M., Matsuzaki, C., Ueno, M., Mizuno, N., Nagatoshi, Y., Imamura, T., Miyago, M., Tanaka, K., Mise, K., Tanaka, T., Mizukoshi, H., Mori, M., & Fujita, Y. (2016). Draft genome sequence of an inbred line of *Chenopodium quinoa*, an allotetraploid crop with great environmental adaptability and outstanding nutritional properties. *DNA Research*, dsw037. <https://doi.org/10.1093/dnares/dsw037>
- Yu, J. H., Kim, K. P., Park, S. M., & Hong, C. B. (2005). Biochemical analysis of a cytosolic small heat shock protein, NtHSP18.3, from *Nicotiana tabacum*. *Molecules and Cells*, 19(3), 328–333. <https://doi.org/10.1007/s00234-005-1355-y>
- Zegzouti, H., Jones, B., Marty, C., Lelièvre, J.-M. M., Latché, A., Pech, J.-C. C., & Bouzayen, M. (1997). ER5, a tomato cDNA encoding an ethylene-responsive LEA-like protein: Characterization and expression in response to drought, ABA and wounding. *Plant Molecular Biology*, 35(6), 847–854. <https://doi.org/10.1023/A:1005860302313>
- Zhang, L., Ohta, A., Takagi, M., & Imai, R. (2000). Expression of plant group 2 and group 3 lea genes in *Saccharomyces cerevisiae* revealed functional divergence among LEA proteins. *Journal of Biochemistry*, 127(4), 611–616. <http://www.ncbi.nlm.nih.gov/pubmed/10739953>
- Zhang, X., Liu, S., & Takano, T. (2008). Two cysteine proteinase inhibitors from *Arabidopsis thaliana*, AtCYSa and AtCYSb, increasing the salt, drought, oxidation and cold tolerance. *Plant Molecular Biology*, 68(1–2), 131–143. <https://doi.org/10.1007/s11103-008-9357-x>
- Zhu, J. K. (2001). Plant salt tolerance. *Trends in Plant Science*, 6(2), 66–71. [https://doi.org/10.1016/S1360-1385\(00\)01838-0](https://doi.org/10.1016/S1360-1385(00)01838-0)
- Zhu, J. K. (2002). Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology*, 53, 247–273. <https://doi.org/10.1146/annurev.arplant.53.091401.143329>
- Zou, C., Chen, A., Xiao, L., Muller, H. M., Ache, P., Haberer, G., Zhang, M., Jia, W., Deng, P., Huang, R., Lang, D., Li, F., Zhan, D., Wu, X., Zhang, H. H., Bohm, J., Liu, R., Shabala, S., Hedrich, R., ... Zhang, H. H. (2017). A high-quality genome assembly of quinoa provides insights into the molecular basis of salt bladder-based salinity tolerance and the exceptional nutritional value. *Cell Research*, 27(11), 1327–1340. <https://doi.org/10.1038/cr.2017.124>
- Zou, Z., Guo, J., Zheng, Y., Xiao, Y., & Guo, A. (2022). Genomic Analysis of LEA Genes in *Carica papaya* and Insight into Lineage-Specific Family Evolution in Brassicales. *Life*, 12(9). <https://doi.org/10.3390/life12091453>