



**INSTITUTO POTOSINO DE INVESTIGACION CIENTIFICA Y  
TECNOLOGICA, A.C.**

**POSGRADO EN CIENCIAS EN BIOLOGIA MOLECULAR**

**Análisis proteínico del grano de amaranto  
(*Amaranthus hypochondriacus* y *A. cruentus*) y  
obtención del mapa proteómico de *Amaranthus  
cruentus***

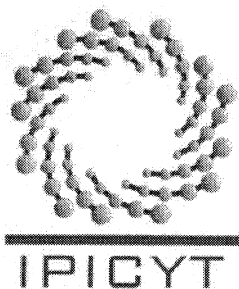
**M. en C. Enrique Maldonado Cervantes**

Para obtener el grado de:  
**Doctor en Ciencias en Biología Molecular**

Director de tesis:

**Dra Ana Paulina Barba de la Rosa**

San Luis Potosí, S.L.P., 28 de febrero de 2014



## Constancia de aprobación de la tesis

La tesis "**Análisis proteínico del grano de amaranto (*Amaranthus hypochondriacus* y *A. cruentus*) y obtención del mapa proteómico de *Amaranthus cruentus***" presentada para obtener el Grado de Doctor en Ciencias en Biología Molecular fue elaborada por **Enrique Maldonado Cervantes** y aprobada el **tres de marzo del dos mil catorce** por los suscritos, designados por el Colegio de Profesores de la División de Biología Molecular del Instituto Potosino de Investigación Científica y Tecnológica, A.C.

**Dra. Ana Paulina Barba de la Rosa**  
Director de la tesis

**Dr. J. Sergio Casas Flores**  
Miembro del Comité Tutorial

**Dr. Samuel Lara González**  
Miembro del Comité Tutorial

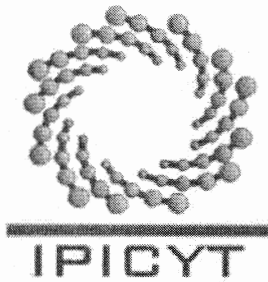
**Dra. Ana Silvia Pérez Martínez**  
Miembro del Comité Tutorial



## **Créditos Institucionales**

Esta tesis fue elaborada en el Laboratorio de Proteómica y Biomedicina Molecular de la División de Biología Molecular del Instituto Potosino de Investigación Científica y Tecnológica, A.C., bajo la dirección del la Dra. Ana Paulina Barba de la Rosa.

Durante la realización del trabajo el autor recibió una beca académica del Consejo Nacional de Ciencia y Tecnología 233301 y del Instituto Potosino de Investigación Científica y Tecnológica, A. C.



# Instituto Potosino de Investigación Científica y Tecnológica, A.C.

## Acta de Examen de Grado

El Secretario Académico del Instituto Potosino de Investigación Científica y Tecnológica, A.C., certifica que en el Acta 066 del Libro Primero de Actas de Exámenes de Grado del Programa de Doctorado en Ciencias en Biología Molecular está asentado lo siguiente:

En la ciudad de San Luis Potosí a los 3 días del mes de marzo del año 2014, se reunió a las 17:00 horas en las instalaciones del Instituto Potosino de Investigación Científica y Tecnológica, A.C., el Jurado Integrado por:

<b>Dr. J. Sergio Casas Flores</b>	<b>Presidente</b>	<b>IPICYT</b>
<b>Dr. Samuel Lara González</b>	<b>Secretario</b>	<b>IPICYT</b>
<b>Dra. Ana Paulina Barba de la Rosa</b>	<b>Sinodal</b>	<b>IPICYT</b>
<b>Dra. Ana Silvia Pérez Martínez</b>	<b>Sinodal externo</b>	<b>UASLP</b>

a fin de efectuar el examen, que para obtener el Grado de:

**DOCTOR EN CIENCIAS EN BIOLOGÍA MOLECULAR**

sustentó el C.

**Enrique Maldonado Cervantes**

sobre la Tesis intitulada:

*Análisis proteínico del grano de amaranto (*Amaranthus hypochondriacus* y *A. cruentus*) y obtención del mapa proteómico de *Amaranthus cruentus**

que se desarrolló bajo la dirección de

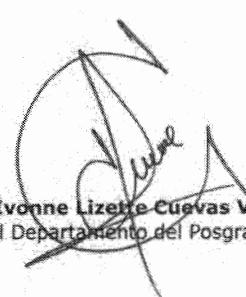
**Dra. Ana Paulina Barba de la Rosa**


El Jurado, después de deliberar, determinó

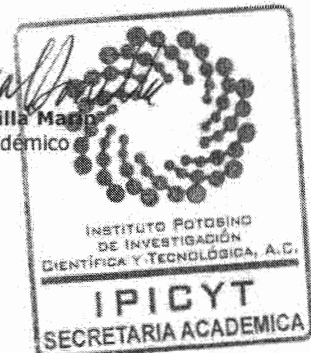
**APROBARLO**

Dándose por terminado el acto a las 18:53 horas, procediendo a la firma del Acta los integrantes del Jurado. Dando fe el Secretario Académico del Instituto.

A petición del interesado y para los fines que al mismo convengan, se extiende el presente documento en la ciudad de San Luis Potosí, S.L.P., México, a los 3 días del mes de marzo de 2014.

  
**Mtra. Ivonne Lizette Cuevas Vélez**  
Jefa del Departamento del Posgrado

  
**Dr. Marcial Bonilla Magaña**  
Secretario Académico



Dedico esta tesis a mis seres queridos que siempre han estado a mi lado a lo lado de todos mis proyectos:  
A mi esposa Julia, a mis padres y hermanos

## **Agradecimientos**

A la doctora Ana Paulina Barba de la Rosa, por la asesoría, la confianza y paciencia otorgada.

A la Dra. Gabriela Margarita Montero y al Dr. Angel Huerta por la asesoría proporcionada en este trabajo.

Al MC Alberto Barrera por su apoyo, asesoría y sobretodo amistad otorgada durante la realización de este trabajo.

A los chicos del Laboratorio 2, por sus consejos para la realización del trabajo de tesis, pero sobre todo por su amistad.

A mis revisores de tesis por sus comentarios.

A el Consejo Nacional de Ciencia y Tecnología por la beca otorgada.

Al IPICYT.

## CONTENIDO

Índice de Tablas .....	ix
Índice de Figuras .....	x
Resumen .....	xi
Abstract .....	xii
I. Introducción .....	1
1.1 El grano de amaranto .....	1
1.1.1 <i>Amaranthus hypochondriacus</i> .....	2
1.1.2 <i>Amaranthus cruentus</i> .....	5
1.3 Dormancia y las plantas tolerantes a estrés abiótico .....	7
1.4 Semillas y proteínas de reserva .....	9
1.4.1 Proteínas funcionales .....	15
1.5 Proteómica .....	16
II. Justificación .....	19
III. Objetivos .....	20
3.1 Objetivo general .....	20
3.2 Objetivos particulares .....	20
IV. Metodología .....	21
4.1 Semillas de amaranto .....	21
4.2 Extracción de proteínas .....	21
4.2.1 Extracción de proteínas de reserva .....	21
4.2.2 Extracción de proteína total .....	22
4.3 Cuantificación de proteínas .....	22
4.4 Análisis Western blot .....	23
4.4.1 Detección de la proteína tipo lunasin .....	24
4.4.2 Detección de residuos fosforilados .....	24
4.5 Electroforesis Bidimensional (2-DE) .....	25
4.5.1 Primera dimensión, isoelectroenfoque .....	25
4.5.2 Segunda dimensión, SDS-PAGE .....	26

4.6 Tinción de geles SDS-PAGE .....	26
4.6.1 Coomassie coloidal.....	26
4.6.2 ProQ Diamond .....	26
4.7 Digestión en gel .....	27
4.8 Espectrometría de masas .....	27
4.9 Identificación bioinformática de proteínas .....	28
4.10 Análisis estadístico.....	29
4.11 Predicción estructural de proteína <i>in silico</i> por analogía de secuencia .....	29
V. Resultados y Discusión.....	30
5.1 Extracción y electroforesis de proteínas en SDS PAGE .....	30
5.1.1 Proteínas de reserva.....	30
5.1.2 Identificación de la proteína tipo lunasin de amaranto .....	33
5.1.3 Extracción y electroforesis de proteínas totales del grano.....	34
5.1.3 Identificación de proteínas fosforiladas en proteína total .....	36
5.2 Identificación de proteínas totales de grano de amaranto a través de 2-D SDS-PAGE .....	39
5.3 Predicción estructural de la proteína LEA 3 por homología secuencial .....	52
VI. Conclusiones .....	53
VII. Anexos .....	55
Bibliografía.....	60
ARTICULO CIENTÍFICO ACEPTADO .....	70
Correo de aceptación de artículo en la revista Journal of Cereal Science .....	70



# Índice de Tablas

	Página
<b>Tabla 1.</b> Distribución y actividad de las proteínas de las semillas de plantas monocotiledóneas y dicotiledóneas.	19
<b>Tabla 2.</b> Proteínas de reserva más abundantes en el grano.	22
<b>Tabla 3.</b> Soluciones de extracción de las diferentes fracciones proteínicas.	30
<b>Tabla 4.</b> Preparación de la curva de calibración para la cuantificación de proteínas.	31
<b>Tabla 5.</b> Programación del equipo Ettan IPGphor II.	34
<b>Tabla 6.</b> Clasificación e identificación de las proteínas representativas expresadas en granos maduros de <i>A. cruentus</i> var. amaranteca.	56

# Índice de Figuras

	Página
<b>Figura 1.</b> Registro histórico del uso del amaranto en las culturas precolombinas.	9
<b>Figura 2.</b> <i>Amaranthus hypochondriacus</i> cultivo Nutrisol.	11
<b>Figura 3.</b> <i>Amaranthus hypochondriacus</i> cultivo revancha.	12
<b>Figura 4.</b> <i>Amaranthus cruentus</i> cultivo Amaranteca.	14
<b>Figura 5.</b> <i>Amaranthus cruentus</i> cultivo Dorada.	15
<b>Figura 6.</b> Semilla endospermo dominante (Maíz).	18
<b>Figura 7.</b> Semilla embriodominante (Frijol).	18
<b>Figura 8.</b> Curva de calibración de la proteína BSA en solución de rehidratado por el método <i>Dc Protein Assay</i> .	31
<b>Figura 9.</b> Análisis del contenido de proteína en las fracciones proteínicas de las variedades de amaranto.	39
<b>Figura 10.</b> Inmunodetección de la proteína lunasin en las fracciones proteínicas del grano de amaranto.	40
<b>Figura 11.</b> Electroforograma de proteína total de los 5 cultivos de amaranto teñidas con Coomassie Coloidal.	43
<b>Figura 12.</b> Electroferograma de proteínas totales de los 5 cultivos de amaranto teñidas con ProQ Diamond.	43
<b>Figura 13.</b> Inmunodetección de residuos fosforilados en extractos de proteína total de los 5 cultivos del grano de amaranto.	44
<b>Figura 14.</b> Patrón electroforético de las proteínas totales del grano de <i>Amaranthus cruentus</i> cv amaranteca.	47
<b>Figura 15.</b> Estructura predicha de la proteína LEA 3 de amaranto.	59

## Resumen

### **Análisis proteínico del grano de amaranto (*Amaranthus hypochondriacus* y *A. cruentus*) y obtención del mapa proteómico de *Amaranthus cruentus***

El amaranto es un grano que ha resurgido en este nuevo milenio con relevancia internacional debido a sus propiedades nutracéuticas y nutricionales. Una característica muy importante que posee el grano del amaranto es su capacidad de tolerar condiciones ambientales adversas a lo largo de su desarrollo desde su estado embrionario, pasando por la germinación y el desarrollo. Los mecanismos por los cuales esta planta resiste dichas condiciones se fundamentan en la función e interacción de las proteínas que permiten el equilibrio metabólico de las células. Se tienen escasos reportes de la presencia de proteínas que permiten la tolerancia al estrés en el grano de amaranto, específicamente en *A. cruentus*. Una de las modificaciones postraduccionales involucradas en el procesamiento y transferencia de información celular es la fosforilación de las proteínas, es por eso que se rastreó la presencia de esta modificación en las proteínas de reserva del grano de *A. hypochondriacus* y *A. cruentus*; donde reportamos una presencia significativa de proteínas fosforiladas en grano maduro de ambas especies de amaranto. Se encontraron diferencias en los patrones electroforéticos de las proteínas totales de 5 cultivos de las especies de *A. hypochondriacus* y *A. cruentus*, también se reportó la presencia de la proteína tipo lunasin en ambas especies. En el presente trabajo se realizó por primera vez una aproximación al proteoma del grano de *A. cruentus* y se describió la presencia de proteínas involucradas en los mecanismos de dormancia.

**Palabras clave:** Amaranto, semilla, proteínas de reserva, fosforilación, proteómica, dormancia

## Abstract

### **Amaranth seed protein analysis (*Amaranthus hypochondriacus* y *A. cruentus*) and amaranthus cruentus's proteomic map**

Amaranth is a grain that has gained international importance due to their nutraceutical and nutritional properties. A very important feature of amaranth is its ability to tolerate adverse weather conditions throughout its development from its embryonic state, through its germination and development. The mechanisms by which this plant resists such conditions are dictated by the function and interaction of proteins that allow the metabolic balance of cells. From our knowledge, no reports, regarding *Amaranthus cruentus*, exist about the presence of proteins that allow stress tolerance of amaranth grain. Differences in the total protein profile of three cultivars of *A. hypochondriacus* and two of *A. cruentus* were determined. as one of the post-translational modifications involved in the processing and transfer of cellular information is phosphorylation of proteins, we report the presence of this modification in storage proteins of *A. hypochondriacus* and *A. cruentus*, where a significant presence of tyrosine-phosphorylated proteins were found. Also the presence of lunasin-like protein in both species was reported. This is the first study dealing with the *Amaranthus cruentus* grain proteome which describes the presence of proteins involved in the mechanisms of dormancy.

**Keywords:** Amaranth, seed, storage proteins, phosphorylation, proteomics, dormancy.

# I. Introducción

## 1.1 El grano de amaranto

El amaranto, junto al frijol, el maíz y el chile, fue uno de los principales cultivos en la meseta mesoamericana antes de la llegada de los españoles, del cual se consumía tanto su grano como hojas y tallos (Sauer 1967). Dentro de las culturas precolombinas el amaranto tenía una importancia religiosa debido a que se formaron figurillas hechas a base de amaranto y miel, donde representaban a su deidad Huitzilopochtli (**Figura 1**). A la llegada de los españoles su cultivo se vio restringido debido a su asociación con prácticas paganas (Barros y Buenrostro, 1997).



**Figura 1. Registro histórico del uso del amaranto en las culturas precolombinas**

La familia *Amaranthaceae* cuenta con 70 géneros y alrededor de 800 especies, de origen tropical, pero que se adaptan muy bien a climas templados. Sus principales centros de distribución se encuentran en los trópicos, tanto en América, África, Asia y Oceanía (Feine y cols 1979, Sauer 1967). El amaranto ha evolucionado en regiones cálidas, secas y con suelos salinos; lo que le ha permitido reconocer ciertas características de adaptación como estructuras de secreción de sales (tricomas

globuliformes), cutícula gruesa, espinas en la base de su hoja en algunas especies y fijación del carbono atmosférico por la vía C<sub>4</sub>.

Dentro del género *Amaranthus* se encuentran las especies: *A. caudatus*, *A. hypochondriacus* y *A. cruentus*, que son las principales productoras de grano comestible; las dos últimas con una gran importancia económica en México. El grano de amaranto es usado en la elaboración de dulces, granolas y hojuelas. Además la harina obtenida de estas semillas se utiliza para la elaboración de galletas, cremas, pastas, etc. (Sanchez, 2010).

### **1.1.1 *Amaranthus hypochondriacus***

*A. hypochondriacus* es una planta herbácea anual de tallo simple o ramificada, en ocasiones alcanza los 3 metros de altura. Sus hojas son elípticas u ovalo-oblongas con el ápice agudo o acuminado y la base cuneada o aguda. La inflorescencia es de gran tamaño, con espigas y panículas laterales; es muy densa, erecta y espinosa. El tamaño de las brácteas es igual al de las ramificaciones del estilo, son muy puntiagudas y es esto precisamente lo que da a la inflorescencia su apariencia y tacto espinoso. La flor de esta especie presenta 5 tépalos, ligeramente curvados y más largos que los tépalos de otras especies para la producción de granos, sus ápices son acuminados, los tépalos internos son más cortos que los externos. El utrículo es circunsésil con una cápsula grande. Las ramificaciones del estilo son curvadas y se juntan en la base en una pequeña hendidura. Las semillas son de color blanco, dorado, café y negro. También es utilizado de manera ornamental. A continuación se mencionan las variedades trabajadas en el presente trabajo.

**Nutrisol.**- Es una variedad de alto rendimiento para Valles Altos, se obtuvo a partir de la raza Azteca. Se caracteriza por presentar pigmentación en la raíz, en el hipocotilo y en el tallo. Las inflorescencias son de crecimiento indeterminado con ramificaciones largas por lo que presentan un tacto espinoso a la madurez. La semilla es de color blanco marfil y ligeramente aplanada. La madurez fisiológica ocurre en promedio en 175 días, llega a crecer entre 1.65 y 2.5 metros.

Los beneficios que presenta son un mayor rendimiento del grano, uniformidad en madurez y en altura de la planta, es apta para cosecha mecánica y adaptable a diferentes ambientes. Tiene buena capacidad de reventado, por lo que es usada en diferentes productos alimenticios (**Figura 2**).



**Figura 2.** *Amaranthus hypochondriacus* cultivo Nutrisol.

**Revancha.**-Es la primer variedad mejorada en México, pertenece a la raza Mercado. Se caracteriza porque solo presenta pigmentación en la raíz, en el hipocotilo y en la

base del tallo, el resto de la planta es de color verde. La inflorescencia es de crecimiento indeterminado con ramificaciones cortas, las flores tienen brácteas cortas, por lo que no presentan el tacto espinoso como las variedades criollas. La semilla es de color blanco marfil y ligeramente aplanada. Es una variedad de ciclo vegetativo intermedio, la madurez fisiológica ocurre en promedio en 130 días, este genotipo es de porte intermedio, su altura varía de 1.01 a 1.62 metros. Los beneficios que presenta son un alto rendimiento de grano, uniformidad en la madurez y en la altura de la planta, sustentabilidad del cultivo y aptitud para cosecha mecánica. El grano de esta variedad tiene buena capacidad de reventado por lo que puede ser usada en productos alimenticios (**Figura 3**).



**Figura 3. *Amaranthus hypochondriacus* cultivo revancha**



### **1.1.2 *Amaranthus cruentus***

Es una planta herbácea de crecimiento erecto que alcanza hasta los 2 metros de altura; su tallo es simple y en ocasiones ramificado. Las hojas son elípticas, rombo-ovatinadas u ovalo-lanceoladas con el ápice agudo, obtuso o acuminado y la base cuneada o aguda. La inflorescencia cuando está completamente desarrollada presenta en la parte inferior espigas suaves y laxas y en la parte superior panículas. Las brácteas son pequeñas, con una punta delgada que se extiende más allá de las láminas, pero nunca tan largas como las ramificaciones del estilo; sin embargo, en algunas ocasiones sobrepasan los tépalos. Las flores presentan 5 tépalos rectos, oblongos u oblongos-obovatinados con ápices agudos; los tépalos internos son más cortos que los extremos. El utrículo es circuncesil con un ápice en forma de torre. Las ramificaciones del estilo son delgadas y erectas. Las semillas pueden ser negras, marrones, blancas o amarillas. Los colores claros son los que se utilizan para la producción del grano, mientras que las plantas que presentan semillas de color oscuro son utilizadas como verduras u ornato (Grubben 1976; Grubben y Solten 1981 y Feine 1979).

**Amaranteca.-** Esta variedad pertenece a *A. cruentus* raza Mexicana. La planta es de tallo y forraje verde y un bajo porcentaje presenta manchas rojas. La inflorescencia es de crecimiento indeterminado con ramificaciones cortas, las flores tienen brácteas cortas, por lo que presentan un tacto espinoso; es de color verde y un bajo porcentaje de plantas presentan inflorescencia variegada de rojo sobre verde. Los beneficios que presenta esta variedad son su maduración precoz (110-140 días) uniformidad en madurez y en altura de planta, aptitud para cosecha mecánica y

potencial para zonas semitropicales, inclusive presenta buen comportamiento en siembras de invierno (marzo) en regiones libres de heladas (**Figura 4**).



**Figura 4.** *Amaranthus cruentus* cultivo Amaranteca

**Dorada.-** Esta variedad pertenece a *A. cruentus* raza Mexicana. La planta es de tallo y follaje de color dorado. La inflorescencia es de crecimiento indeterminado con ramificaciones cortas, las flores tienen brácteas cortas, por lo que presentan tacto espinoso. Los beneficios que presentan esta variedad son su maduración precoz (115-150 días), uniformidad en madurez y en altura de planta, aptitud para cosecha mecánica y potencial para zonas semitropicales, inclusive presenta buen comportamiento en siembra en invierno (marzo) en regiones libres de heladas ya que es de fotoperiodo neutro (**Figura 5**).



**Figura 5. *Amaranthus cruentus* cultivo Dorada**

### **1.3 Dormancia y las plantas tolerantes a estrés abiótico**

Las semillas se pueden dividir con base en su tolerancia a la deshidratación en ortodoxas y recalcitrantes; las semillas que toleran la deshidratación pertenecen al primer grupo, en éste encontramos a la mayoría de las semillas conocidas (el 100% de las semillas de plantas monocotiledóneas); y las semillas que no toleran la deshidratación se clasifican dentro del grupo de las recalcitrantes; de las cuales existen ciertas semillas de plantas dicotiledóneas. Esta tolerancia a la deshidratación es la clave para que la semilla entre en un estado donde la germinación está restringida; y a esta fase se denomina dormancia. La dormancia en las semillas es la etapa en la que el embrión se prepara para la germinación y es mantenida bajo diversos mecanismos moleculares que ayudan a bloquear las señalizaciones necesarias para la germinación; así como mecanismos protectores que mantienen un metabolismo basal latente y la integridad de las proteínas de la semilla (Roberts 1997).

Se han sugerido diversos procesos o mecanismos que confieren tolerancia a la desecación, incluyendo la acumulación de las moléculas protectoras, por ejemplo proteínas abundantes de embriogénesis tardía (LEA), represores metabólicos y la presencia de sistemas reparadores durante la rehidratación de las semillas, como proteínas de choque térmico.

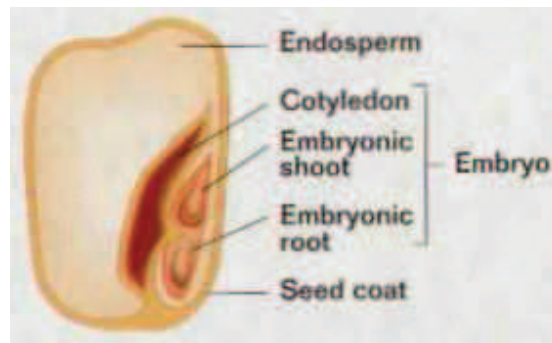
Cuming (1999) encontró que la presencia de proteínas LEA esta relacionada con la tolerancia a la deshidratación, y se han reportado altos niveles de proteínas LEA en el desarrollo de la maduración tardía y en la deshidratación de tejidos vegetativos de plantas de resurrección. Grelet y cols. (2005) identificaron la proteína LEA del grupo 3 (PsLEAm) en la matriz mitocondrial de la semilla del chícharo (*Pisum sativum*) y demostraron que protege a dos enzimas de este organelo (fumarasa y rodanasa) durante el secado. Boucher y cols. (2010) mostraron que MtPM25 (proteína LEA) ayuda a prevenir y disociar agregados proteicos formados durante varios tipos de estrés incluyendo la desecación. También se ha reportado que la ausencia de varias proteínas LEA conlleva a la pérdida de la tolerancia de la desecación durante la germinación (Bartels y Sunkar 2005).

El proceso de germinación comienza cuando la dormancia de la semilla termina, en este momento la semilla seca y quiescente pasa a un estado de metabolismo acelerado y una compleja red de vías de señalización es activada para iniciar la germinación y satisfacer los requerimientos para el desarrollo de la planta. Se ha sugerido que hay receptores comunes para los agentes del inicio de la germinación dentro de la membrana plasmática de las células embrionarias. Cuando son activados, estos receptores inician una cascada de señalización, que quizá involucre

la síntesis de hormonas promotoras de la germinación o sensibilicen al embrión para la acción de dichas hormonas; tales como las giberilinas, citocininas y auxinas (Bewley 1997; Lara-Núñez 2008). Se ha reportado que existen cambios en el estado de fosforilación en una cinasa CDPK dependiente de calcio asociada a membrana, lo cual conlleva al fin de la dormancia (Trewavas 1998). Estas evidencias nos indican que la acción de las cinasas y fosfatasa juegan un papel esencial en el proceso de transición de la dormancia y la germinación; y estos datos finalmente nos indican qué tan aptas para la germinación serán las semillas que se encuentran en dormancia.

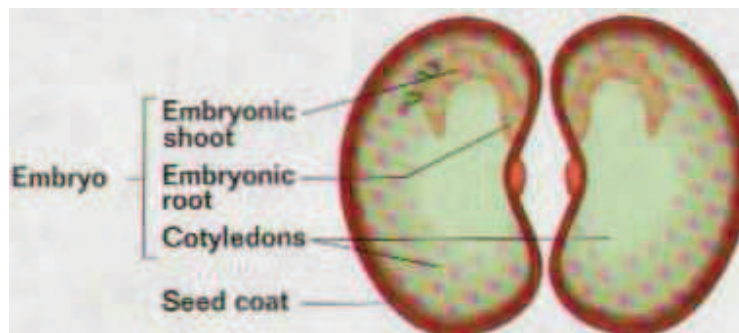
#### **1.4 Semillas y proteínas de reserva**

Las semillas se componen de tres partes básicas: el embrión, propiamente dicho (De Smet y cols. 2010); un tejido que contiene nutrientes de reserva para el desarrollo de la futura planta (endospermo); y una cubierta de la semilla (Haughn y Chaudhury 2005). Las semillas pueden clasificarse debido a la proporción de endospermo y embrión en: endospermo-dominante y embrión-dominante. Las semillas endospermo-dominantes (**Figura 6**) pertenecen generalmente a las plantas monocotiledóneas, dentro de los cuales se encuentran muchos de los cereales de mayor importancia económica (maíz, trigo, arroz, cebada, etc.) (Haughn y Chaudhury 2005; Baubec y Mittelsten Scheid 2006). Por análisis de proteomas de maíz, trigo, arroz y cebada se identificaron 1496 proteínas las cuales fueron clasificadas en 10 grupos acorde a la función de las proteínas (Miernyk y cols. 2011, Bevan y cols. 1998).



**Figura 6. Semilla endospermo dominante (Maíz)**

Por otra parte, las semillas embrión-dominantes (**Figura 7**) abarcan principalmente las semillas de plantas dicotiledóneas, como es el caso del amaranto, y como su nombre lo menciona, la parte principal de la semilla es el embrión, y una pequeña cantidad de endospermo (en algunos casos inexistente, como *Arabidopsis* (Haughn y Chaudhury 2005)).



**Figura 7. Semilla embriodominante (Frijol)**

Este grupo se subdivide en dos: ricas en almidón y ricas en aceites. El amaranto pertenece a las plantas embrión-dominantes ricas en almidón. Se analizaron y compararon los proteomas de *Glycine max* (Hajduch y cols. 2005; Agrawal y cols. 2008), *Medicago truncatula* (Gallardo y cols. 2003, Gallardo y cols. 2007), y *Lotus*

*japonicus* (Dam y cols. 2007), las cuales al igual que el amaranto poseen semillas ricas en almidón, las proteínas analizadas también fueron clasificadas en 10 grupos funcionales. La distribución obtenida de los análisis proteómicos de los granos monocotiledóneos y dicotiledóneos se encuentran resumidos en la **Tabla 1**.

**Tabla 1. Distribución y actividad de las proteínas de las semillas de plantas monocotiledóneas y dicotiledóneas**

Grupo	FUNCION CELULAR	Distribución de proteínas en semillas	
		Monocotiledóneas (%)	Dicotiledóneas (%)
1	Metabolismo central	34	49
2	Estructura celular	12	19
3	Marcaje de proteínas	7	3
4	Respuesta a estrés	5	3
5	Plegamiento de proteínas	5	5
6	Metabolismo de ácidos nucleicos	2	2
7	Síntesis de proteínas	2	3
8	Vía de señalización de hormonas	2	3
9	Transporte membranal	2	4
10	Proteínas de función desconocida	29	10

Miernyk y Hajduch (2011) realizaron un análisis de las proteínas fosforiladas presentes en la semilla y del conjunto total (1723 proteínas) de las 3 especies (*G.*

*max*, *M. truncatula* y *L. japonicus*) se identificaron 234 proteínas con al menos un residuo fosforilado; estas proteínas pertenecían en su mayoría a las categorías de metabolismo central y de las vías de señalización hormonal; aunque también se identificaron proteínas de reserva con residuos fosforilados (Agrawal and Thelen 2009).

Uno de los criterios de selección para el cultivo de plantas productoras de granos es el alto aporte energético proporcionado por las semillas; uno de los polímeros que aportan mayor energía son las proteínas, especialmente las proteínas de reserva, las cuales fueron clasificadas por Osborne (Osborne 1924), con base en su solubilidad en agua (albuminas), soluciones salinas (globulinas), soluciones alcohólicas (prolaminas) y soluciones alcalinas o ácidas (glutelinas). Agronómicamente hablando el grupo de mayor interés son las albúminas, globulinas y prolaminas (Gibbs y cols. 1989, Xu y Messing 2009, Shewry y cols. 1995); como rasgo general las albúminas están presentes en todas las semillas, las prolaminas y las glutelinas en monocotiledóneas y globulinas en dicotiledóneas (Xu y Messing 2009, Weber y cols. 2005).

Las globulinas se clasifican según su velocidad de sedimentación en 7S y 11S (Shewry y cols. 1995), aunque existen trabajos científicos que reportan globulinas 3S y 12S, y en algunos trabajos son más estrictos y reportan globulinas 2.2S y 11.3S (Templeman y cols. 1987).

En un principio las proteínas de reserva se les asignaba un nombre basado en la raíz latina de la planta de la que provenían; por ejemplo, las zeínas del maíz (*Zea mays*) (Lawton 2002), las hordeínas de la cebada (*Hordeum vulgare*) (Bietz 1982), las



secalinas del centeno (*Secale cereale*) (Kreis 1985), etc. En el caso del trigo (*Triticum aestivum*) sus prolaminas fueron denominadas gliadinas (Shewry y Tatham 1990, Qi y cols. 2009), las globulinas 7S son denominadas vicilinas y las globulinas 11S, leguminas (Häger y cols. 1996). Finalmente existen algunos casos donde los nombres fueron designados acorde al nombre de la planta como los casos de las cactinas que es usado para nombrar a las albuminas 2S de *Cereus jamacaru* (Aragão TCFR y cols. 2000).

Las proteínas de reserva son procesadas en el retículo endoplasmico y almacenadas en las vacuolas de proteínas de reserva, estas vacuolas tienen dos rutas diferentes dependiendo del tipo de proteínas que contienen; en el caso de las prolaminas estas se dirigen directamente a la vacuola de almacenamiento; y para el caso del resto de las proteínas de reserva (albúminas, globulinas y glutelinas) pasan al aparato de golgi para finalmente compartimentalizarse en las vacuolas de almacenamiento (Frigerio y cols. 2008).

Una característica que comparten las proteínas de reserva 2S, 7S y 11S es que son sintetizadas como prepro-poliproteínas; y es en el retículo endoplásmico donde empieza la maduración de la semilla, una de las modificaciones que sufren son la N-glicosilación, que generalmente son removidas en el proceso de maduración. El procesamiento y ensamblaje de las proteínas se lleva a cabo durante la ruta secretoria. En el caso de las dicotiledóneas las proteínas de reserva son almacenadas en las vacuolas de almacenamiento ligadas a membrana en las células del parénquima de los cotiledones así como también en el eje embrionario (Dunwell y cols. 2001). Y en el caso de las monocotiledóneas las proteínas se encuentran en la

parte muerta del endospermo; en células que murieron por una muerte programada después del depósito de proteínas.

En cuanto al patrón electroforético de las proteínas de reserva se sabe que éstas comparten características generales, **Tabla 2**.

**Tabla 2. Proteínas de reserva más abundantes en el grano**

Grupo		Descripción
Albuminas 2S		Peso molecular de aproximadamente de 18-21 kDa; una característica de sus secuencias es que cuenta con 8 residuos de cisteína que forman 4 puentes disulfuro; algunas albúminas son procesadas en 2 cadenas.
Globulinas 7S		Homo o heterotrimeros de dos tipos de subunidades homologas: vicilinas (45-50 kDa) y las convicilinas (60-70 kDa).
Globulinas 11S		Heterohexadímeros de precursores de cadena sencilla unidos por enlaces disulfuro; formados por una cadena ácida (30-40 kDa) y una cadena básica (17-20 kDa)
Prolaminas	Prolaminas ricas en sulfuro	Dominios de repetidos en prolina y glicina, seguidos de dominios ricos en cisteína homólogos a los dominios de albuminas 2S; estas cisteínas forman enlaces disulfuro.
	Prolaminas pobres en sulfuro	Dominios de repetidos de prolina y glicina
	Prolaminas de alto peso molecular	Dominios de repetidos en prolina y glicina; y repetidos de tri, hexa y nona glutamatos (en partes centrales de la secuencia)

#### **1.4.1 Proteínas funcionales**

Una de las principales características del grano de amaranto es su alto contenido de proteínas (17%, peso seco), superior al contenido de cualquier cereal; así como un buen balance de aminoácidos; teniendo un contenido alto de lisina; las proteínas mayoritarias en la semilla del amaranto son las fracciones de albúminas, globulinas y glutelinas; lo que lo hace un candidato ideal para la elaboración de preparados proteínicos o como componente de formulaciones alimenticias de alto valor energético. Así también se ha comprobado en diversas especies y sus respectivas variedades que las proteínas de la semilla de amaranto no son alergénicas (Chakraborty y cols. 2000, Rascón-Cruz y Sinagawa-García 2004). Y finalmente, en cuanto a las características generales de las proteínas de amaranto, se ha comprobado que estas son libres de gluten; lo cual representa una excelente opción alimenticia para las personas celíacas (Cinzia Ballabio y cols. 2011).

Silva-Sánchez y cols. (2008) reportaron a través de un análisis *in silico* que las proteínas de reserva del grano de amaranto, poseen encriptados en sus secuencias péptidos bioactivos como inhibidores de la enzima DPP IV e inhibidores de la enzima convertidora de la angiotensina. En este mismo estudio se descubrió la existencia de una proteína que era detectada por los anticuerpos anti-lunasin. El lunasin es un péptido de soya (Jeong y cols. 2003), el cual se ha demostrado que posee propiedades cáncer preventivas (Hernández Ledezma y cols. 2009); el cual consta de 43 aminoácidos, con un peso aproximado de 5.8 kDa; este péptido deriva del procesamiento de una albúmina 2S. Estos péptidos y proteínas relacionadas al lunasin han sido reportados en dicotiledóneas como *Solanum nigrum* y cereales como trigo, cebada, arroz y recientemente en tricale (Mitchell y cols. 2013).

Aparentemente la función del lunasin en la semilla, es detener la citocinesis en la etapa en la que la semilla almacena polímeros como grasas, almidón y proteínas; el péptido lunasin no evita que se duplique el material genético pero si evita que la célula se divida en dos células hijas, y esta es una característica muy común en las células de la semilla que se encuentra en esta fase de maduración (Gálvez y De Lumen 1999).

En un estudio reciente se reportó que la proteína tipo lunasin descubierta en el grano del amaranto posee características nutracéuticas cáncer preventivas similares a las reportadas del péptido lunasin de soya, esta proteína se encuentra presente en diversas fracciones del grano de amaranto con un peso aproximado de 20 kDa (Maldonado Cervantes y cols. 2009); la presencia de este péptido, nos habla de que las proteínas de reserva pueden poseer funciones reguladoras alternas a su simple función de reservorio de aminoácidos para la germinación. Estas propiedades son un campo de estudio en el cual poco se ha indagado y que puede llegar a ser una oportunidad de redescubrir propiedades de las proteínas de reserva útiles para su utilización en el área científico-tecnológica.

### **1.5 Proteómica**

La proteómica, por definición, es el estudio de las proteínas a gran escala, en particular de su estructura y función. Esta disciplina, ha tenido avances significativos a lo largo de los 20 años de su aparición; esto se debe a que durante este periodo se han mejorado las técnicas de extracción y purificación de proteínas; pero sobretodo a que la instrumentación con la que trabaja la proteómica ha evolucionado para darnos mayor sensibilidad, precisión de masas, y fragmentación de péptidos mucho más

eficiente; estos avances hacen que lo análisis proteómicos sean mucho más precisos, exactos y por ende fiables (Desiderio y cols. 2010, Rabilloud y cols. 2010, Brewis y Brennan 2010). Otro factor que ha permitido el desarrollo de esta disciplina es la secuenciación masiva de genomas de diversos organismos; ya que la identificación de las proteínas estudiadas con la proteómica resulta más eficiente al tener una referencia genómica (Qureshi y cols. 2007). En el caso particular del amaranto, aún se desconoce la secuencia de su genoma; pero se han realizado transcriptomas de diversos tejidos y bajo diversas condiciones que han ayudado a la identificación de proteínas de esta planta (Huerta-Ocampo y cols 2009, Aguilar-Hernández y cols. 2011, Délano-Frier y cols. 2011).

La proteómica provee una eficiente solución para los estudios a gran escala del desarrollo de la semilla. Comparada con estrategias moleculares y genéticas tradicionales, la proteómica involucra estudios comparativos globales, permitiendo la identificación a gran escala de las proteínas y análisis de expresión de proteínas activas en compartimientos específicos de la semilla o regiones en ciertos estadíos de desarrollo (Wienkoop y cols. 2010). Esta capacidad de la proteómica hace posible contestar múltiples preguntas que son incontestables analizando una sola proteína o gen; además, aporta información nueva de proteínas relacionadas con el desarrollo que no pueden ser detectadas por métodos genéticos (Deng y cols. 2013).

El análisis de la población de ARNs no proporciona la información de lo que realmente ocurre metabólicamente en una célula en un estado determinado; ya que las efectoras de dichos procesos son las proteínas y estas pueden poseer diversas modificaciones postraduccionales, como son metilaciones, ubiquitinaciones,

fosforilaciones glicosilaciones y procesamientos proteolíticos que pueden cambiar la ubicación, función, actividad y/o función de las proteínas. La presencia de diversas isoformas de las proteínas pueden ser el resultado de procesos antes mencionados o por empalme alternativo y pueden ser analizados a través de métodos proteómicos, no genómicos, ni transcriptómicos (Peck 2005). Otra ventaja de analizar directamente las proteínas es que los niveles de ARN no son directamente proporcionales a la cantidad de proteínas presentes en un determinado momento de desarrollo de la semilla. Finalmente, se hace mucho más fácil correlacionar la activación de varios genes en respuesta a un estímulo interno o externo analizando el total de proteínas presentes y/o activas en ese preciso momento (Deng y cols. 2013).

La proteómica, en combinación con otras disciplinas “ómicas”, ha aportado avances significativos en el estudio de proteínas presentes en diversas semillas. Por otra parte, estudios de proteómica comparativa han llegado a ser un método de básico para entender los mecanismos básicos de respuesta de diversos tejidos de la planta, entre ellos los del grano (Deng y cols. 2013).

## **II. Justificación**

La identidad y diversidad de las proteínas contenidas en un grano determina las características nutricionales; la eficiencia en la germinación, y la utilidad de estas proteínas en el área biotecnológica. La falta de un mapa proteómico del grano de amaranto y la falta de caracterización de los perfiles proteómicos de las variedades más utilizadas en nuestro país; acarrear una falta de información que frena el desarrollo y óptima utilización de este grano. Es por eso que el presente trabajo se dio a la tarea de analizar desde un punto de vista proteómico las proteínas constituyentes del grano de amaranto, para ampliar el conocimiento de este grano que ha cobrado importancia a nivel mundial en los últimos años.

## **III. Objetivos**

### **3.1 Objetivo general**

Ampliar el conocimiento de la composición proteínica de variedades de Amarantho cultivadas en la región del centro de México a través del estudio de las proteínas de reserva del grano y proteínas totales bajo las técnicas de electroforesis en geles de poliacrilamida en una y dos dimensiones.

### **3.2 Objetivos particulares**

- Realizar un mapa proteómico del grano de amaranto, para llevar un estudio comparativo de las proteínas de reserva de diversas variedades del grano de amaranto.
- Comparar los estados de fosforilación de estas proteínas con el fin de determinar la presencia de proteínas metabólicamente activas en el grano.
- Determinar la presencia del péptido lunasin en las fracciones proteínicas del grano.



## IV. Metodología

### 4.1 Semillas de amaranto

Se utilizaron semillas maduras de *Amaranthus hypochondriacus* L. variedades Nutrisol, Revancha y Annapurna y *Amaranthus cruentus* variedades Amaranteca y Dorada para el desarrollo de este trabajo, facilitadas por el Dr. Eduardo Espitia Rangel del Instituto Nacional de investigaciones Forestales Agrícolas y Pecuarias, campo experimental bajo. Los granos de amaranto fueron molidos en un molino para granos de café, y tamizados en una malla del número 100 y desengrasados con una mezcla de hexanos.

### 4.2 Extracción de proteínas

#### 4.2.1 Extracción de proteínas de reserva

El método de extracción se llevó a cabo según el método de Barba de la Rosa y cols. (1992). Una suspensión de la harina y la solución de extracción correspondiente a cada fracción de proteína de reserva en una proporción 1:10. En la **Tabla 3** se muestra la solubilidad de las fracciones proteínicas en las diversas soluciones de extracción. Para la extracción de las proteínas de la harina de amaranto se agitó la mezcla de harina-extrayente por 30 min con sonicación de 1.5 min a 35% de amplitud (Ultrasonic Processor, Newtown, Connecticut, USA). Después de la extracción la suspensión se centrifugó por 20 min a 13000 rpm a 4 °C. La pastilla se empleó para la extracción de la siguiente fracción proteínica, hasta extraer el total de las fracciones. Las muestras obtenidas fueron almacenadas a -80°C hasta su uso posterior.

**Tabla 3. Soluciones de extracción de las diferentes fracciones proteínicas.**

<b>Fracción proteínica</b>	<b>Extrayente</b>
<b>Albuminas</b>	Agua desionizada
<b>Globulinas 7S</b>	NaCl 0.1M , NaH <sub>2</sub> PO <sub>4</sub> 10 mM , EDTA 10 mM, pH 7.5
<b>Globulinas 11S</b>	NaCl 0.8M, NaH <sub>2</sub> PO <sub>4</sub> 10 mM, EDTA 10 mM, pH 7.5
<b>Glutelinas</b>	Tris 100 mM pH 8

#### **4.2.2 Extracción de proteína total**

Para la extracción de proteínas totales del grano de amaranto, el harina se homogenizó con solución de extracción (0.1M NaCl, 35 mM Na<sub>2</sub>HPO<sub>4</sub>, 1% triton x-100, 0.1%  $\beta$ -ME) en una relación 1:10, y se agitó por 10 minutos a 4 °C, con una subsecuente sonicación de 1.5 min en hielo, a una amplitud de 35% (Ultrasonic Processor). La suspensión se centrifugó 20 min a 13000 rpm a 4 °C. El sobrenadante se colectó y se precipitó con metanol-cloroformo (4:1), la pastilla de proteínas se lavó con metanol frio, cada lavado consistió de mezclado con agitación en vortex y centrifugación por 10 min a 13000 rpm a 4 °C. La pastilla resultante se almacenó a -80 °C hasta su uso posterior.

#### **4.3 Cuantificación de proteínas**

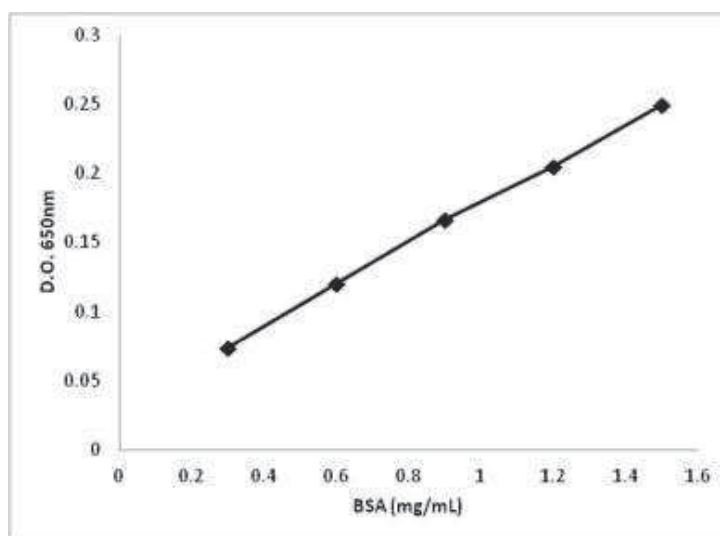
La cuantificación de las proteínas se realizó usando el reactivo *Dc Protein Assay* de BioRad (Hercules, California, USA). Para las curvas de calibración se preparó un stock de de BSA a una concentración de 5 mg/ml disuelto en solución de rehidratado para tiras IPG (8M urea, 2% CHAPS, 0.002% azul de bromofenol) y se realizó una curva estándar como se indica en la **Tabla 4**.

**Tabla 4. Preparación de la curva de calibración para la cuantificación de proteínas**

Estándar	Conc. Proteína [mg/ml]	5 mg/ml BSA [ $\mu$ l]	Solución de rehidratado [ $\mu$ l]
1	0.0	0	100
2	0.3	6	94
3	0.6	12	88
4	0.9	18	82
5	1.2	24	76
6	1.5	30	70

Se midió la absorbancia a una longitud de onda de 650 nm dentro de los 30 min de estabilidad del compuesto; la curva de calibración obtenida está representada en la

**Figura 8.**



**Figura 8. Curva de calibración de la proteína BSA en solución de rehidratado por el método *Dc* protein assay.** Los parámetros de la línea de tendencia fueron:  $y = 0.1457x + 0.032$ ,  $R^2 = 0.9991$

#### **4.4 Análisis Western blot**

Las proteínas separadas en geles de poliacrilamida fueron transferidas a membranas PVDF (Sequi-Blot PVDF, Bio-Rad). Las membranas se activaron con metanol al 80%

y se lavaron con agua MiliQ. La transferencia se llevó a cabo en la cámara de transferencia semi-seca (Trans-Blot SD, Biorad ®) para minigeles (11x7 cm) 45 min a 15 V, para geles de 13.3 x 8.5 cm de 1 h a 15V.

#### **4.4.1 Detección de la proteína tipo lunasin**

La membrana se bloqueó por 90 min con una solución de leche desengrasada al 5% disuelta en TBS 1%(1% tween-20, 500 mM NaCl, 20 mM tris base), y se lavó con solución 1% TBS-T. Después del lavado, se incubó por 1 h con el anticuerpo primario anti-lunasin a una dilución 1:3000 en la solución con leche desengrasada al 3% disuelta en 1% TBS 1%; se lavó con la solución TBS-T 1%y se incubó por otra hora con el anticuerpo secundario (Sigma Aldrich, St Louis, Missouri USA) acoplado a fosfatasa alcalina a una dilución de 1:10000 en la solución con de leche desengrasada al 3% disuelta en TBS 1%, después se lavó con la solución TBS-T 1%, y finalmente se reveló con los sustratos BCIP/NBT (Boehringer Mannheim, Ingelheim, Germany) disueltos en tris tris 0.1M, NaCl 0.01M, pH 9.5; y MgCl<sub>2</sub> 50 mM agregado al momento, la reacción se detuvo con agua destilada.

#### **4.4.2 Detección de residuos fosforilados**

La membrana se bloqueó por 90 min con una solución de BSA al 3% disuelta en TBS 1%(tween-20 1%, NaCl 500 mM, 20 mM tris base 20 mM), y se lavó con solución TBS-T 1%. Después de lavado, se incubó por 1 h con el anticuerpo primario anti-serina fosforilada, anti-treonina fosforilada o anti-tirosina fosforilada (Abcam, Cambridge, England) a diluciones de 1:700, 1:5000 y 1:500, respectivamente en la solución con BSA al 1% disuelta en TBS 1%; se lavó con la solución TBS-T 1%y se

incubó por otra hora con el anticuerpo secundario acoplado a la enzima peroxidasa en una dilución de 1:5000 en la solución con BSA 1% disuelta en TBS 1%, después se lavó con la solución TBS-T 1%, y finalmente se reveló con el sustrato 3,3'-diaminobencidina, DAB disueltos en Tris 0.05 M, NiCl<sub>2</sub> 0.03%, y H<sub>2</sub>O<sub>2</sub> 0.03% agregado al momento. Un precipitado color marrón fue el indicativo de la detección de las proteínas fosforiladas; la reacción se detuvo con agua destilada.

#### 4.5 Electroforesis Bidimensional (2-DE)

##### 4.5.1 Primera dimension, isoelectroenfoque

Para este estudio se utilizaron tiras de 13 cm de gradiente de pH inmovilizado con un intervalo de 3-10 (BioRad). Las tiras fueron rehidratadas con la muestra de proteína a una concentración de 1mg/mL disueltas en la solución de rehidratación, agregando DTT a una concentración final de 0.018 M y anfolitos (Biorad, Hercules, California USA) a una concentración final de 0.5%; por un lapso 16 horas a temperatura ambiente. El electroenfoque se realizó en el equipo Ettan IPGphor II (GE Healthcare) bajo las condiciones señaladas en la **Tabla 5**, con un paso previo de desalado a 150 V por 2 h.

**Tabla 5. Programación del equipo Ettan IPGphor II**

<b>Paso</b>	<b>Modo de voltaje</b>	<b>Voltaje máximo</b>	<b>Voltz-Hora (kVh)</b>
1	Gradiente	500	0.01
2	Gradiente	4000	3.4
3	Constante	8000	13.5

Las tiras fueron almacenadas a -20°C para su posterior uso.

#### **4.5.2 Segunda dimensión, SDS-PAGE**

Posterior a la separación por isoelectroenfoque las tiras se equilibraron en la solución con Tris-HCl 50 mM, pH 8.8, Urea 6M, Glicerol 30%, SDS 2% y Azul de bromofenol 0.002% por 15 minutos con agitación constante. Las tiras fueron colocadas en geles de acrilamida al 12% de la cámara Ruby cell (GE Healthcare) y se corre la electroforesis a 25 Volts por aproximadamente 5 h.

#### **4.6 Tinción de geles SDS-PAGE**

##### **4.6.1 Coomassie coloidal**

Para la tinción de los geles se utilizó el colorante Coomassie G-250 empleando una variante del coomassie coloidal. Los geles se fijaron en una solución con metanol 50% y ácido fosfórico 2%. Después el gel se sensibilizó en una solución de sulfato de amonio 17%, metanol 33% y ácido fosfórico 3%; se agregó coomassie G-250 para obtener una concentración final del 0.06%, el gel se destiñó por 1 h con agua desionizada. Las imágenes fueron visualizadas en el equipo Pharos FX plus BioRad.

##### **4.6.2 ProQ Diamond**

Los geles se fijaron en una solución con metanol 50% y ácido acético 10%; se realizaron lavados con agua desionizada. El gel se tiñó con el colorante ProQ-Diamond (Molecular Probes, Eugene, Oregon, USA) en una proporción 1:3 con agua, se incubó y se destiñó con una solución con acetonitrilo 20% y acetato de amonio 50 mM. Las imágenes fueron visualizadas en el equipo Pharos FX plus (BioRad, Hercules, California, USA) a una longitud de excitación de 532 nm y usando filtros de 580 nm.

#### **4.7 Digestión en gel**

Los puntos seleccionados para la identificación fueron escindidos de los geles y lavados exhaustivamente con agua milliQ y  $\text{NH}_4\text{HCO}_3$  25 mM. Posteriormente fueron deshidratados con acetonitrilo para remover contaminantes y retirar el colorante. Las muestras fueron sometidas a una reducción con DTT 10 mM en una solución de  $\text{NH}_4\text{HCO}_3$  25 mM, seguida de una alquilación con 55 mM iodoacetamida. Se realizaron lavados con  $\text{NH}_4\text{HCO}_3$ /acetonitrilo. La digestión fue realizada toda la noche a 37 °C con tripsina grado secuenciación (Pro mega, Madison, WI, USA). Los fragmentos tripticos resultantes fueron extraídos dos veces con una solución de acetonitrilo y agua (3:2 v/v) conteniendo ácido trifluoroacético 0.1%. Los dos sobrenadantes fueron concentrados y los péptidos tripticos fueron desalados y concentrados a un volumen final de 12  $\mu\text{l}$ .

#### **4.8 Espectrometría de masas**

Los análisis de espectrometría de masas fueron realizados en el equipo Synapt-HDMS (Waters Corp, Milford, MA, USA) acoplado en línea a el nano-ACQUITY-UPLC (Waters). Los péptidos fueron cargados en la pre-columna Symetry C18 (5  $\mu\text{m}$ , 20 mm x 180  $\mu\text{m}$ ). La separación de los péptidos se realizó en la columna analítica BEH130 C18 (1.7  $\mu\text{m}$ , 10 cm x 100  $\mu\text{m}$ ) y eluidos a 35 °C con gradiente lineal de fase móvil B (0.1% ácido fórmico en acetonitrilo) por 30 min con un flujo de 600 nl/min, seguido de un lavado con 85% de la fase móvil B. La columna fue nuevamente equilibrada a las condiciones iniciales por 15 minutos con la fase móvil A (ácido fórmico 0.01% en agua). Como referencia en el spray de referencia se utilizó

[Glu1]-fibrinopeptide B, inyectado por la bomba auxiliar del sistema a 200 nl/min a una concentración de 100 fmol/ $\mu$ l.

El RF aplicado al analizador de masas de cuádruplo fue ajustado para transmitir los iones eficientemente con una relación  $m/z$  de 50-3000. El espectrómetro de masa fue operado en modo V para todas las mediciones y los análisis se realizaron en modo ESI. Los datos MS/MS fueron obtenidos para más de tres cargas de iones 2+, 3+,4+ detectados y obtenidos en una rampa de colisión de energía de 15 a 40 eV, la adquisición fue cambiada del modo MS a MS/MS cuando la intensidad del pico base excedía el umbral de 10 cuentas y regresaba al modo MS después de 5 s.

#### **4.9 Identificación bioinformática de proteínas**

La identificación de las proteínas fue realizada con los datos de los espectros MS/MS usando el motor de búsqueda MASCOT (Matrix Science, London, UK, <http://www.matrixscience.com>). La base de datos usada fue NCBI nr (<http://www.ncbi.nih.gov>). Se realizó la búsqueda de los péptidos monoisotópicos bajo los siguientes parámetros: tripsina como proteasa, tolerancia de péptidos de 10 ppm, tolerancia de masa/carga (MS/MS) de 0.1 Da, péptidos con carga de +2,+3,+4 y contemplando como modificación fija la carbamidometilación de las cisteínas y como modificación variable la oxidación de la metionina. La identificación fue considerada acertada cuando se obtenía un score significativo en MASCOT con un valor de  $P < 0.05$ .



#### **4.10 Análisis estadístico**

Para conocer las diferencias significativas de nuestros análisis se utilizó un análisis de varianza de una vía (one way ANOVA) con un límite de confianza de 95% usando el programa Statistica 7.

#### **4.11 Predicción estructural de proteína *in silico* por analogía de secuencia**

Adicionalmente a este trabajo se realizó la amplificación del gen y predicción estructural de la proteína LEA 3 (Maldonado y cols. 2014), reportada en el presente trabajo, por lo que se dio a la tarea de realizar un modelaje *in silico* de la estructura secundaria de la proteína LEA 3 (**Figura S1**) usando el programa en línea Phyre engine V 2.0 del servidor en línea: <http://www.sbg.bio.ic.ac.uk> (Kelley y Sternberg 2009).

## V. Resultados y Discusión

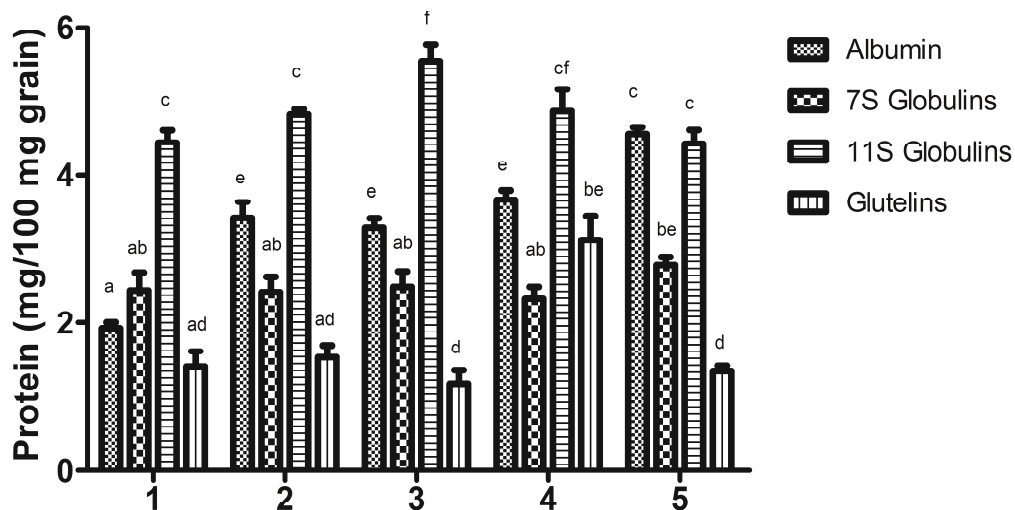
### 5.1 Extracción y electroforesis de proteínas en SDS PAGE

#### 5.1.1 Proteínas de reserva

La extracción y cuantificación de las proteínas totales de los cultivos de amaranto (*A. hypochondriacus* y *A. cruentus*) reveló diferencias tanto cuantitativas como cualitativas. La cuantificación de fracciones proteínicas de reserva en los granos de los diversos cultivos de amaranto reveló que las fracciones mayoritarias en los 5 cultivos analizados resultaron ser las albúminas y las globulinas, especialmente las globulinas 11S, las cuales en los cinco cultivos fue la fracción mayoritaria y las concentraciones se encuentran en el rango de 4.43 – 5.54 % (m/m) con respecto al peso total de la muestra; la siguiente fracción mayoritaria fueron las albuminas, excepto en el cultivo de amaranteca (*A. cruentus*), en el cual la segunda fracción más abundante fue globulina 7S (2.44%); y en el caso del cultivo revancha (*A. hypochondriacus*) las albuminas poseyeron una cantidad equiparable (4.74 %) a las cantidades de globulinas 11S de los cultivos analizados. De manera general las concentraciones de albúminas están dentro de un rango de 1.92 – 4.96 % (m/m). La tercera fracción predominante fueron las globulinas 7S, cuyas cantidades relativas se mostraron muy homogéneas y no hubo diferencias significativas entre los cultivos analizados; la cantidad de proteína tiene un rango de 2.33 – 2.78 % (m/m). Finalmente, la fracción minoritaria fueron las glutelinas las cuales tampoco mostraron diferencias significativas, excepto en el cultivo nutrisol (*A. hypochondriacus*) del cual se recuperó un 3.12% equiparable a la cantidad de albúmina obtenida en los 5 cultivos analizados; el rango general de porcentaje que representan estas fracciones

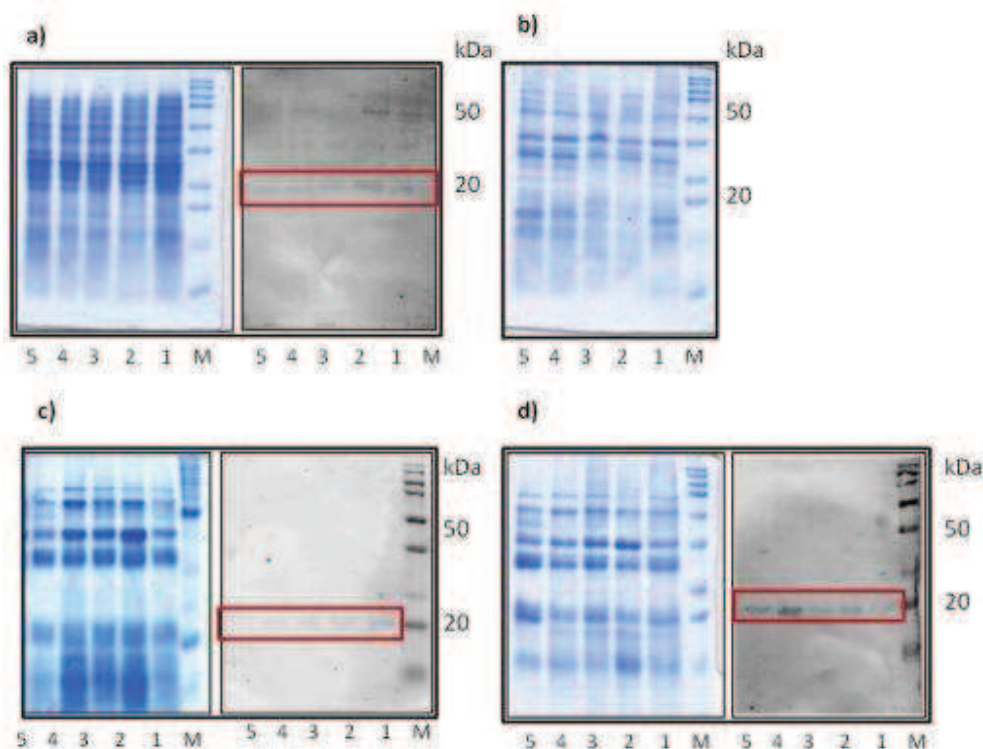
es del 1.17 – 3.12%. El gráfico que representa las cantidades de proteínas de reserva está representado en la **Figura 9**.

Se ha reportado en la literatura que la fracción de globulinas 11S es la más abundante en el grano de amaranto; así también se sabe que la fracción de albuminas representa una de las fracciones mayoritarias en este grano (Barba de la Rosa AP y cols. 1992). Sin embargo no se tenían datos de que las glutelinas fueran la fracción menos abundante, esto podría deberse a la gran variabilidad de métodos de extracción de este tipo de proteínas ya que existen reportes en los que determinan que la fracción mayoritaria en el grano de amaranto son las albúminas y que las prolaminas poseen el tercer sitio por arriba de las globulinas (Seguro-Nieto y cols. 1992) y en la actualidad existen múltiples reportes que notifican que las globulinas 11S son la fracción mayoritaria y que las prolaminas están en mínima cantidad de manera casi imperceptible (Barba de la Rosa y cols. 1992, Chen y Paredes-López 2007).



**Figura 9. Análisis del contenido de proteína en las fracciones proteínicas de las variedades de amaranto.** 1: *A. cruentus* cv. Amaranteca, 2: *A. hypochondriacus* cv. Annapurna, 3: *A. cruentus* cv. Dorada, 4: *A. hypochondriacus* cv. Nutrisol, 5: *A. hypochondriacus* cv. Revancha.

Los patrones electroforéticos de las proteínas de reserva de los 5 cultivos de amaranto analizados muestran patrones electroforéticos muy similares a los reportados previamente (Barba de la Rosa y cols. 1992), y dentro de las fracciones de albúminas y globulinas 7S no muestran diferencias notables dentro de los 5 cultivos; por otra parte, en las fracciones de globulinas 11S y glutelinas las proteínas con un peso molecular menor a 20 kDa mostraron diferencias en el patrón electroforético. Esto se debe al polimorfismo presente en este grupo de proteínas. Estas diferencias aún siendo sutiles pueden ser de suma importancia ya que los péptidos y proteínas biológicamente activos se sabe que tienen un peso menor a 10 kDa (Figura 10).



**Figura 10. Inmunodetección de la proteína lunasin en las fracciones proteínicas del grano de amaranto.** Electroferogramas y western blot de 4 fracciones proteínicas de las 5 variedades del grano de amaranto: a) albuminas, b) globulinas 7S, c) globulinas 11S y d) glutelinas. (1) *A. cruentus* var. Amaranteca, (2) *A. hypochondriacus* var. Annapurna, (3) *A. hypochondriacus* var. Dorada, (4) *A. hypochondriacus* var. Nutrisol, (5) *A. hypochondriacus* var. Revancha. Se usó la reacción de la fosfatasa alcalina sobre el NBT/BCIP para la detección.

### **5.1.2 Identificación de la proteína tipo lunasin de amaranto**

El péptido lunasin fue reportado originalmente en las proteínas de reserva del grano de soya, y forma parte del procesamiento de la albumina 2S (de Lumen y Galvez 1999); en los últimos años se ha reportado su presencia en otros granos dicotiledóneos y también en granos monocotiledóneos (Mitchell y cols. 2013). En este trabajo se rastreó la proteína tipo lunasin de amaranto en las proteínas de reserva de los 5 cultivos de *A. cruentus* y *A. hypochondriacus*, y se detectó la

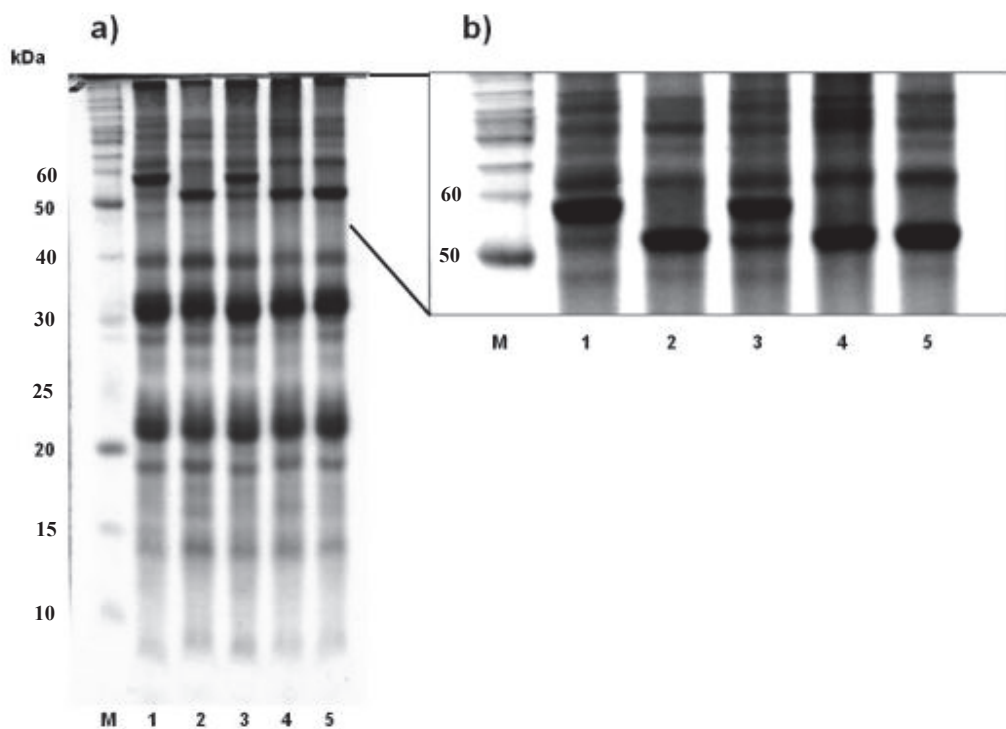
presencia de esta proteína en las fracciones correspondientes a albúminas, globulinas 11S y glutelinas, con un peso molecular aproximado de 20 kDa, en todos los casos (**Figura 10**). En las fracciones de las globulinas 7S y prolaminas no se detectó la presencia de la proteína tipo lunasin. Es la primera ocasión que se reporta la presencia de esta proteína en *A. cruentus* (cultivos Amaranteca y Dorada). También es el primer trabajo donde se reporta la presencia de esta proteína en los cultivos de Revancha y Annapurna; ya que solamente se había reportado en los cultivos Criolla, Nutrisol, DGTA y Gabriela (Silva-Sánchez y cols. 2008; Maldonado-Cervantes y cols. 2010).

La presencia de la proteína tipo lunasin en todos los cultivos analizados y en las fracciones más predominantes del grano demuestra que este péptido posee una gran importancia en el proceso de maduración y germinación del grano de especies muy distantes evolutivamente (Mitchell y cols. 2013). La función específica de este péptido bioactivo aun permanece desconocida; pero se especula que participa en el arresto celular que sufren las células del grano cuando entran en la etapa de acumulación de macromoléculas; este proceso es de suma importancia para la supervivencia del organismo al momento de la germinación (De Lumen y Hernandez 2008). Esta propiedad de inducir arresto celular le ha conferido propiedades anticancer preventivas en células de mamíferos; lo que lo convierte en un potencial biofármaco con importantes aplicaciones.

### **5.1.3 Extracción y electroforesis de proteínas totales del grano**

Se encontraron diferencias dentro del patrón electroforético de *A. cruentus* y *A. hypochondriacus* sus proteínas en geles SDS-PAGE de una dimensión se

encontraron diferencias evidentes en el rango de 50 a 60 kDa (**Figura 11**), específicamente de un par de proteínas de peso aproximado de 53 kDa y una de mayor peso de 58 kDa, lo interesante de estas proteínas es que la proteína de 58 kDa se encontraba en una mayor cantidad en los dos cultivos correspondientes al grano de *A. cruentus* (Amaranteca y Dorada) y en cambio en los cultivos correspondientes a *A. hypochondriacus* (Annapurna, Nutrisol y Revancha) se encontraba mayoritariamente la proteína de 53 kDa. Estos patrones electroforéticos diferenciales nos podrían ayudar a discernir la identificación rápida de los granos correspondientes a dos especies de amaranto ya que sus granos no poseen diferencias evidentes una vez libres de cascarilla.

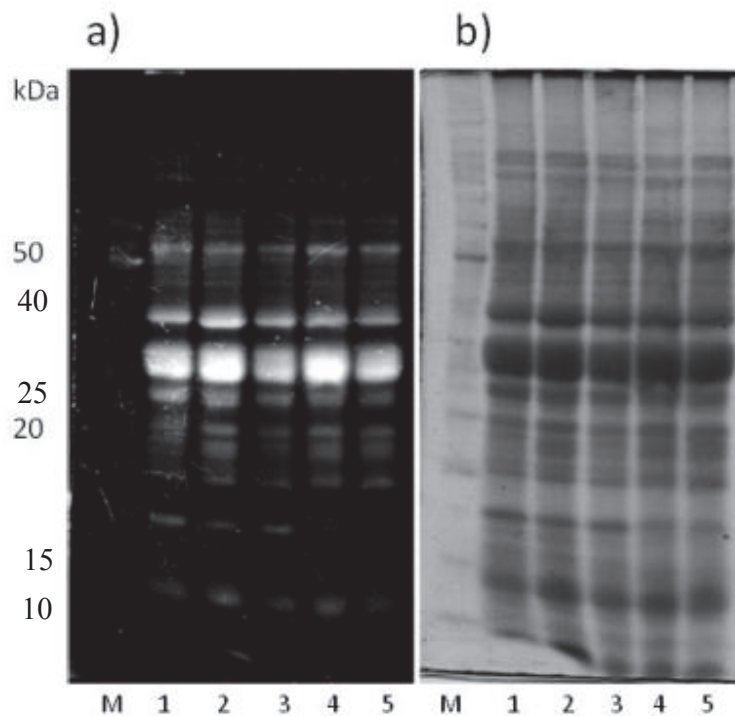


**Figura 11. Electroferograma de proteína total de los 5 cultivos de amaranto teñidas con coomassie coloidal.** a) electroferograma completo, b) acercamiento de la zona con mayor polimorfismo en proteínas. (1) *A. cruentus* var. Amaranteca, (2) *A. hypochondriacus* var. Annapurna, (3) *A. cruentus* var. Dorada, (4) *A. hypochondriacus* var. Nutrisol, (5) *A. hypochondriacus* var. Revancha, (M) marcador de peso molecular.

### **5.1.3 Identificación de proteínas fosforiladas en proteína total**

Las proteínas totales de grano fueron teñidas con un colorante fluorescente específico para identificar residuos fosforilados y se encontró que en los cinco cultivos analizados existe una gran cantidad de proteínas que tienen al menos un residuo fosforilado, esto concuerda con estudios donde reportan que al menos un tercio de las proteínas de plantas se encuentran en un estado fosforilado (Smith y cols. 2005). Dentro de los cinco cultivos analizados se observan diferencias en los patrones de proteínas fosforiladas, principalmente en la banda correspondiente a 15 kDa (**Figura 12**).

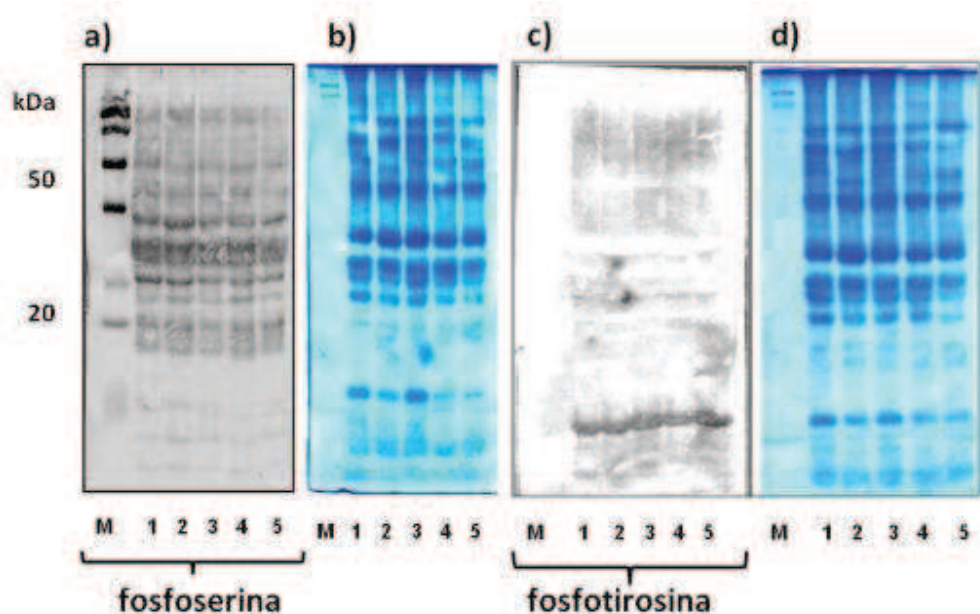




**Figura 12. Electroferograma de proteínas totales de los 5 cultivos de amaranto teñidas con ProQ Diamond.** a) Gel de proteínas teñidas con el colorante específico para fosfoproteínas ProQ Diamond y mismo gel teñido con coomassie coloidal. (1) *A. cruentus* cv. Amaranteca, (2) *A. hypochondriacus* cv. Annapurna, (3) *A. hypochondriacus* cv. Dorada, (4) *A. hypochondriacus* cv. Nutrisol, (5) *A. hypochondriacus* cv. Revancha.

Para corroborar que presencia de en estado fosforilado del grano de amaranto, se realizaron 3 diferentes western blots para identificar fosforilaciones en los residuos de serina, tirosina y treonina. No fue posible detectar residuos de treonina fosforilados en ninguna de las repeticiones realizadas (datos no mostrados). Para el caso de la identificación de las proteínas que contenían residuos de serina fosforilados, se encontró una gran cantidad de proteínas que concuerda con los resultados de la tinción con ProQ Diamond y no se observan diferencias evidentes en el patrón electroforético. Cabe hacer notar que en esta detección se observaron proteínas con un peso aproximado de 20 kDa que apenas son perceptibles con la

tinción del gel con azul de Coomassie; en cambio, aquellas proteínas que fueron evidentes en el gel teñido con azul de Coomassie de aproximadamente 15 kDa no fueron detectadas en el western blot (**Figura 13**).



**Figura 13. Inmunodetección de residuos fosforilados en extractos de proteína total de los 5 cultivos del grano de amaranto.** a) inmunodetección a través de western blot de los residuos fosforilados usando los anticuerpos primarios: a) anti-fosfoserina y c) anti-fosfotirosina; a la derecha de cada western blot se muestra el electroferograma de las proteínas totales (b y d). (1) *A. cruentus* cv. Amaranteca, (2) *A. hypochondriacus* cv. Annapurna, (3) *A. cruentus* cv. Dorada, (4) *A. hypochondriacus* cv. Nutrisol, (5) *A. hypochondriacus* cv. Revancha. Se usó la reacción de la fosfatasa alcalina sobre el NBT/BCIP para la detección.

En el western blot que identificó las proteínas con fosforilaciones en los residuos de tirosina se detectó una banda de un peso aproximado de 15 kDa; lo cual es sumamente interesante ya que las fosforilaciones en los residuos de tirosina son poco habituales en las proteínas de plantas (Kersten y cols. 2009).

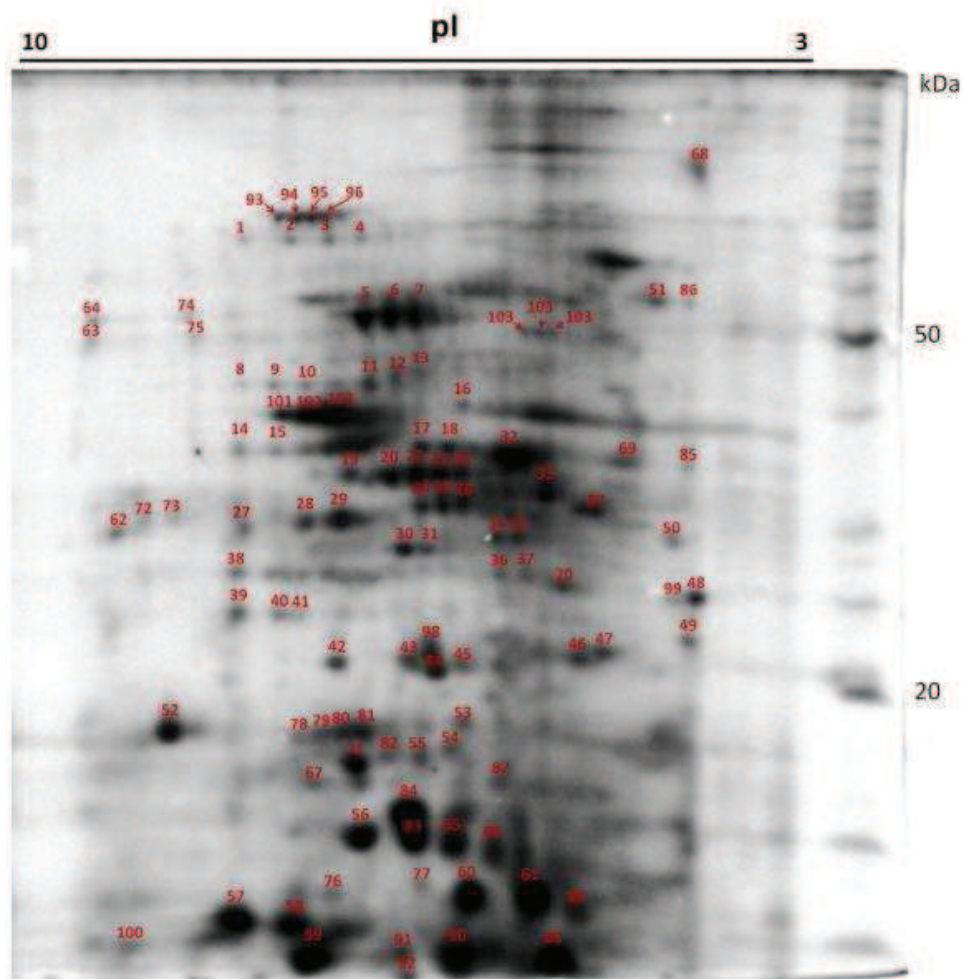
El hecho que se encontraran proteínas fosforiladas en el grano del amaranto no es de extrañarse debido a que esta modificación postraducciona posee una gran importancia en los procesos celulares, y a pesar de que la célula se encuentra en un estado de metabolismo basal, aun mantiene activos ciertos procesos involucrados en la germinación y mantenimiento del estado de dormancia; también se sabe que la dinámica de fosforilación/desfosforilación puede repercutir en la estructura de las proteínas, estabilidad, localización subcelular e interacción con otras biomoléculas; por lo tanto la fosforilación no solo participa en procesos de suma importancia como es la transducción de señales y la regulación de procesos metabólicos, sino que también da estabilidad y estructura a las proteínas en el estado de dormancia que se caracteriza por un estrés generalizado debido a la falta de agua en el sistema (De la Fuente y cols. 2007).

La presencia de proteínas fosforiladas en el grano de amaranto nos indica que estas pueden estar participando activamente en procesos biológicos importantes para el funcionamiento del metabolismo, en la estructura y localización de las proteínas y en la funcionalidad de las proteínas al momento de entrar en contacto con un estímulo como es el término de la dormancia.

## **5.2 Identificación de proteínas totales de grano de amaranto a través de 2-D SDS-PAGE**

Con el objetivo de realizar un mapa proteómico del grano de amaranto se extrajeron las proteínas totales del grano de *A. cruentus* cultivo amaranteca; por la razón que este cultivo ofrece una resistencia innata a estrés por sequía y por lo tanto, los resultados obtenidos nos ayudarían a determinar la maquinaria proteómica que

cuentan los granos de amaranto para conferirle resistencia a estreses abióticos. El mapa proteómico de *A. cruentus* se muestra en la **Figura 14** en el cual están enumeradas las proteínas que fueron sometidas al proceso de identificación a través de espectrometría de masas. Ya que se trataba de representar la mayor población posible de proteínas se trabajó con un rango de pH de 3-10, las proteínas del grano poseen un rango muy amplio de pesos moleculares y diversidad en sus puntos isoeléctricos.



**Figura 14.** Patrón electroforético de las proteínas totales del grano de *Amaranthus cruentus* cv amaranteca. El gel fue teñido con coomassie coloidal, el porcentaje de acrilamida fue del 12% y la longitud del gel 13 cm. Se utilizaron tiras de 13 cm con un rango de pH de 3-10.

La etapa de dormancia en la que se encuentran las semillas antes de la germinación es un estado de estrés generalizado por la ausencia de agua, tanto metabólicamente como estructuralmente. Lo que lleva a las células tanto del embrión como del endospermo a expresar proteínas que puedan ayudar a tolerar dicho estrés y mantener su metabolismo latente. Dentro de las proteínas expresadas en el grano de amaranto se encuentran las proteínas tipo LEA, como LEA 3, LEA 5 y Proteína EM. Las proteínas LEA están involucradas en la tolerancia a la deshidratación bajo mecanismos que van en la preservación de las estructuras celulares, retención de agua, secuestro de iones y protección directa de proteínas y/o membranas; o renaturalización de proteínas desnaturalizadas. Las proteínas LEA se encuentran principalmente en plantas, encontradas originalmente en semillas, así como en otros tejidos (Wise 2003). Las proteínas LEA han sido clasificadas en 6 grupos principalmente, con base en las proteínas descritas en *Gossypium hirsutum*; las características de las proteínas LEA encontradas en este trabajo son descritas a continuación:

- LEA grupo 1. Son proteínas relacionadas secuencialmente con las proteínas Em y solo se han encontrado en plantas. Son proteínas sin estructura secundaria en solución.
- Las proteínas EM (Early methionine labelled, por su nombre en inglés) son las proteínas de mayor abundancia en el citosol de las células del embrión de semillas maduras. La expresión de los genes EM está regulada positivamente por el ácido abscísico. La expresión de este tipo de proteínas es inducida como una respuesta a un estrés osmótico, tanto en los embriones inmaduros

como durante la germinación posiblemente con el ácido abscísico como intermediario en esta respuesta. Este tipo de proteínas está ampliamente distribuido en muchas especies y se ha observado su presencia en las etapas finales de la embriogénesis y en respuesta a estrés por falta de agua.

- LEA grupo 3 (spot 3). Este grupo de proteínas poseen un motivo de repetidos de 11-mer aminoácidos, que tienen una secuencia consenso que es:  $\Phi\Phi E/QX\Phi KE/QK\Phi XE/D/Q$  (donde  $\Phi$  representa un residuo hidrofílico). Lo interesante de este grupo es que se han reportado homólogos de este tipo de proteínas en otros organismos poco relacionados en plantas como nemátodos y procariontes.
- LEA 5 (spot 52). Este tipo de proteínas posee características hidrofóbicas. al contrario del resto de las proteínas LEA; se ha reportado su presencia en plantas sometidas a estrés por sequía y en hojas que sufrieron daño mecánico (Goyal y cols 2005).

Por otra parte se encontró la presencia de la cadena alfa del complejo NAC (Nascent polypeptide associated complex, spot 50); el cual es un heterodímero constituido por las cadenas alfa y beta, y puede unirse reversiblemente a los ribosomas. La función exacta del complejo NAC permanece aún incierta. Se ha sugerido que su función es clasificar y traslocar a las proteínas recién sintetizadas; y que promueve la entrada de las proteínas al retículo endoplásmico. Se ha reportado que la sobreexpresión de este complejo en arroz (*Oryza sativa*) aumenta la resistencia de la planta al estrés por salinidad (Yan y cols 2005).

Otra de las proteínas que identificamos en el grano de amaranto fue la proteína de unión a elementos de respuesta a la deshidratación (dehydration-responsive element-binding protein, DRE, spot 48); este tipo de proteínas participa en la respuesta de la planta a estrés abiótico y biótico. Su presencia está relacionada con órganos que han sufrido desecación, y se encuentra en la mayoría de las semillas ortodoxas (Yamaguchi-Shinozaki y Shinozaki 2006).

Una proteína identificada fue la proteína relacionada a la patogénesis (Pathogenesis-related protein, PR, spot 67), esta proteína es expresada después de un ataque de un agente biológico principalmente en las hojas. Sin embargo, se han reportado este tipo de proteínas en diversos granos, principalmente cereales. Dentro del grupo de este tipo de proteínas se han encontrado quitinasas,  $\beta$  1,3-glucanasas, inhibidores de proteasas, peroxidasas. Se cree que las proteínas relacionadas a la patogénesis en semillas inducen la producción de ácido abscísico durante la germinación y también forman parte del sistema de defensa de las semillas ante los depredadores (Van Loon y Van Strien 1999).

También se encontró la presencia de una proteína relacionada a la dormancia (spots 19 y 20), la cual como su nombre lo indica está encargada de mantener a la semilla en un estado de dormancia, preparándola para su posterior germinación. La dormancia cobra mucha importancia en las semillas de plantas tolerantes a diversos tipos de estrés abiótico ya que determina el momento propicio para que la planta crezca en las mejores condiciones ambientales (Oliver y cols. 2000, Dolferus y cols. 2011).

También se encontró la presencia de la cistatina (spots 30 y 31), este inhibidor de cistatina-proteasas ha sido reportado en un gran número de especies vegetales; específicamente la cistatina de amaranto ha sido reportada y caracterizada por Valdez-Rodríguez y cols (2007 y 2010); en estos reportes nos indican que la expresión de RNA se eleva en raíz y hojas cuando la planta es sometida a diversos tipos de estrés: hídrico, salino y estrés por temperaturas bajas; en un estrés por calor, los niveles de transcritos de esta proteína se elevan en hoja únicamente. Otra de las características de la cistatina de amaranto es que inhibe el crecimiento de hongos fitopatógenos. En cuanto a su expresión en el grano se ha encontrado que posiblemente esté relacionada con la resistencia al estrés abiótico que sufre la semilla durante la etapa de la dormancia; una evidencia que nos reafirma esta idea es que los niveles de cistatina disminuyen en el grano durante su germinación (Valdez-Rodríguez y cols. 2007).

Finalmente dentro de las proteínas encontradas que podrían estar confiriendo una protección ante algún tipo de estrés, encontramos a la enzima dehidriascorbato reductasa (DHAR, spot 57); la cual cataliza la conversión del ácido dehidroascórbico a ácido ascórbico, en presencia de glutatión. El ácido ascórbico es el mayor agente antioxidante con el que cuentan las células en general. El ácido ascórbico está involucrado en la desintoxicación de especies reactivas de oxígeno, las cuales se producen por ejemplo en procesos como en la cadena respiratoria y en la fotosíntesis. Debido a la importancia de la enzima DHAR en el reciclaje del ácido ascórbico; esta proteína se encuentra en prácticamente todas las plantas superiores. La expresión de esta enzima es muy elevada en los primeros estadios de formación



de la semilla y posteriormente va desapareciendo en conforme la semilla entra en un estado de desecación; hasta finalmente permanecer en estados basales en el grano maduro (Yang y cols. 2009). Estudios han comprobado que la sobreexpresión de la DHAR en plantas como tabaco, tomate y arabidopsis confiere resistencia a diversos tipos de estrés como son el salino; resistencia a la acción del metil viologeno y peróxido de hidrogeno; y bajas temperaturas (Eltayeb y cols. 2006; Kwon y cols. 2003).

En síntesis, se encontró la presencia de diversas proteínas que ayudan sobrellevar el estrés que supone el estado de dormancia en una semilla madura ya que en este estado, el metabolismo se encuentra latente y el agua disponible para las células es escasa. Estas condiciones llevan a la semilla a proteger la integridad tanto del embrión como de sus componentes moleculares a través de proteínas como las proteínas LEA, proteínas del complejo NAC, la proteína relacionada a la dormancia, la proteína de unión a elementos de la deshidratación y las proteínas EM; la presencia de estas proteínas son el resultado de la adaptación de las plantas que han sido sometidas a un estrés, generalmente abiótico; en este caso en particular, se trata de las semillas maduras latentes o dormantes. Por otro lado también se encontraron proteínas que son expresadas a raíz de un estrés biótico como son las proteínas relacionadas a la patogénesis y las cistatinas; esto nos indica que no solo la semilla está preparada para el estrés biótico que conlleva la desecación propia de la maduración, sino que también está preparada para el ataque de patógenos y de posibles depredadores. Así como también se encontró una proteína que ayuda a lidiar con las especies reactivas de oxígeno como es el caso de la DHRA.

En otro grupo, también muy relacionado al estado de estrés de las semillas en el estado de dormancia de la semilla, encontramos a las proteínas que ayudan a plegar de manera correcta a otras proteínas que han sido desnaturalizadas por efecto de diversas condiciones desfavorables. Un par de proteínas encontradas corresponden a proteínas de choque térmico; en el caso de las proteínas pequeñas de choque térmico (spot 82) se encuentran a niveles indetectables en los tejidos vegetativos; pero cuando la planta entra en estrés por altas temperaturas este tipo de proteínas se convierten en las proteínas más abundantes (Vierling 1997); en el caso de las proteínas pequeñas de choque térmico clase I y II se localizan en el citoplasma y se especula que su función es la de chaperonas (Lee y cols. 1995, 1997). A pesar de que este tipo de proteínas son expresadas a partir de un estímulo por altas temperaturas; también son expresadas en el desarrollo de la semilla, generalmente aparecen antes de la desecación (Zur Nieden y cols. 1995), también son relativamente abundantes durante la germinación y en los primeros días de desarrollo de la planta y después desaparecen rápidamente. Algunos estudios especulan que estas pequeñas proteínas de choque térmico tienen la función de proteger componentes celulares durante la desecación y subsecuente rehidratación (Wehmeyer y Vierling 2000).

Otro grupo de proteínas que identificamos en nuestro estudio son las encargadas del marcaje de proteínas para ser procesadas debido a su estado estructural alterado; la primera de ellas es la unidad beta del proteasoma (spot 37), el cual está compuesto por una gran cantidad de proteínas que se encargan de la degradación de proteínas ubiquitinadas, así como también en otros procesos celulares como: el progreso del

ciclo celular, germinación de polen y respuesta a estrés y adaptación (Miyawaki y cols. 1997). Dentro de este grupo de proteínas encargadas del marcaje de proteínas encontramos a la calreticulina (spot 51), la cual ha sido identificada en pocas plantas, encontrada originalmente en hojas de espinaca y posteriormente en Arabidopsis, cebada, chícharo, tabaco, maíz, nabo, *Ricinus communis* L, y arroz; su función principal es la de chaperona y sensor de  $Ca^{2+}$ ; existen evidencias que se expresa en niveles elevados durante la embriogénesis de la cebada, *Nicotiana plumbaginifolia*, y maíz. Existe evidencia de que este tipo de proteínas ayuda a mediar diversos tipos de estrés, por ejemplo señalizaciones generadas por patógenos, graviestimulación, y otros factores de estrés; sin embargo no se conoce a ciencia cierta los mecanismos por los cuales estas proteínas confieren resistencia a las plantas o ayudan a sobrellevar el estado de latencia en las semillas (Jia y cols 2008).

Así que observamos nuevamente proteínas que participan activamente en el estado de latencia o dormancia de las semillas, en este caso proteínas que ayudan a reciclar a las proteínas que han sido dañadas por el estrés (proteasoma) y proteínas que ayudan al correcto plegamiento de las proteínas (calreticulinas).

También tenemos la presencia de la proteína glutatión S transferasa (spot 38), la cual ha sido ubicada en la categoría del metabolismo central por su función en la desintoxicación de compuestos xenobióticos; pero también se sabe que participa en las vías de señalización de las MAP ciansas (Marrs 1996).

Otro grupo nutrido de las proteínas identificadas en este trabajo se trata de las que participan en el metabolismo de carbohidratos: gliceraldehido 3P deshidrogenasa (10), SORBIDRAFT\_06g011080 (spot 24); de lípidos y esterios: hidroximiristoil-[acil-

portador-proteína] deshidratasa (spot 27) y fosfatasa del ácido púrpura (spot 104), ácidos nucleicos: nucleolisin TIAR (spot 64) y urilidato cinasa (spot 25); ciclo de Krebs: glutamato descarboxilasa (spots 17, 101 y 103), generación de compuestos de alta energía y fuerza reductora: ubiquinol-citocromo C reductasa (spot 18), NADH deshidrogenasa (spot 21), NADH oxidasa (spot 41) y nucleótido difosfato cinasa I (spot 55) y síntesis de proteínas: proteína ribosomal 40S (spot 76) y factor de elongación SPOC (spot 85). La presencia de este tipo de proteínas tiene mucho sentido que se encuentren en la semilla del grano de amaranto ya que el endospermo de la semilla se encuentran diversas moléculas de reserva que ayudaran a germinar la planta obteniendo de dichas macromoléculas los aminoácidos, azúcares y ácidos grasos necesarios para el anabolismo que sugiere el proceso de germinación.

Cabe mencionar que la única proteína de reserva identificada fue la globulina 11S de amaranto (Barba de la Rosa y cols. 1992), esto se debe probablemente a la falta de caracterización del resto de las proteínas de reserva del grano de amaranto; particularmente en este estudio encontramos las 2 subunidades de la globulina 11S con un peso de 21.3 y 35.6 kDa.

A manera de resumen las proteínas identificadas en el análisis del mapa proteómico del grano de amaranto están resumidas en la **Tabla 6**.

**Tabla 6. Clasificación e identificación de las proteínas representativas expresadas en granos maduros de *A. cruentus* cv. Amaranteca.**

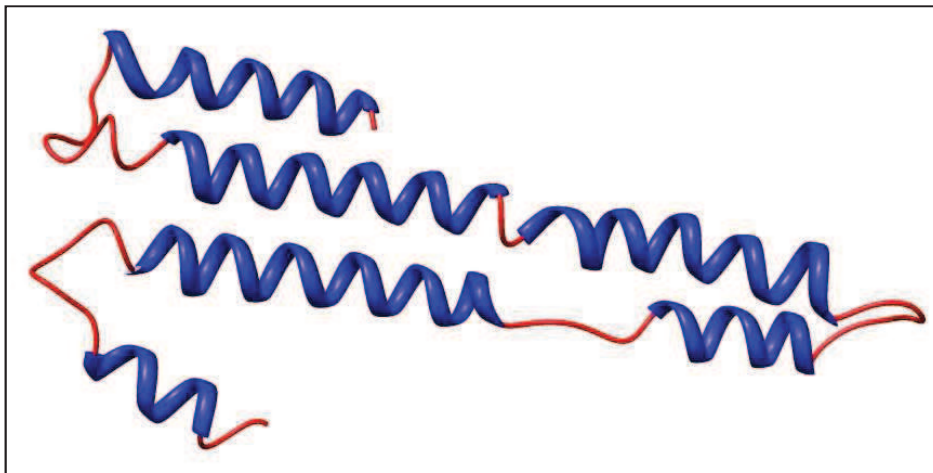
RESPUESTA A ESTRÉS							
No.	NOMBRE DE PROTEINA	ORTOLOGO MAS PROBABLE	NUMERO DE ACSESO	EXP KDA/PI	TEORICO KDA/PI	MS <sup>a</sup>	LOCALIZACION
3	LEA protein group 3	<i>Brassica napus</i>	<a href="#">isoTiG16059</a>	53.3/6.3	43.5/5.64	51	
50	Alpha chain of nascent polypeptide associated complex	<i>Theobroma cacao</i>	<a href="#">gil508727650</a>	29.8/4.09	22/4.3	157	<a href="#">cytosolic ribosome</a>
84	EM protein	<i>Vigna radiata</i>	<a href="#">gil1141782</a>	16.6/ 6.16	12.2/ 8.86	55	
52	LEA 5	<i>Vitis amurensis</i>	<a href="#">gil324103765</a>	19/8	16.4/6.2	117	
48	dehydration-responsive element-binding protein 1F-like	<i>Vitis vinifera</i>	<a href="#">gil225440970</a>	25.5/3.9	25.9/9	40	<a href="#">nucleus</a>
67	pathogenesis-related protein	<i>Spinacia oleracea</i>	<a href="#">gil444792485</a>	17.4/ 6.88	17.9/4.72	94	
19	dormancy related protein, putative	<i>Arabidopsis thaliana</i>	<a href="#">gil12322163</a>	35.5/6.6	31.4/5.92	79	
20	dormancy related protein, putative	<i>Arabidopsis thaliana</i>	<a href="#">gil12322163</a>	35.1/6.32	31.4/5.92	56	
30	Cystatin	<i>Amaranthus hypochondriacus</i>	<a href="#">gil110748606</a>	29.1/6.21	18.5/6	214	<a href="#">Secreted.</a>
31	Cystatin	<i>Amaranthus hypochondriacus</i>	<a href="#">gil110748606</a>	29.1/6.21	27.8/5.67	41	<a href="#">Secreted.</a>
57	dehydroascorbate reductase	<i>Nicotiana tabacum</i>	<a href="#">gil28192427</a>	11.6/7.46	23.6/7.7	40	<a href="#">cytosol plasma membrane</a>
METABOLISMO CENTRAL							
10	glyceraldehyde-3-phosphate dehydrogenase	<i>Salvia elegans</i>	<a href="#">gil343131368</a>	45.4/6.94	21.4/6.21	52	<a href="#">Cytoplasm</a>
24	hypothetical protein SORBIDRAFT_06g011080	<i>Sorghum bicolor</i>	<a href="#">gil242075416</a>	32.6/6.07	49/7.15	38	

27	PREDICTED: (3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase-like	<i>Glycine max</i>	<a href="#">gi 255640918</a>	30.7/7.42	19.9/9.72	32	<a href="#">cytoplasm</a>
55	Nucleoside diphosphate kinase 1	<i>Spinacia oleracea</i>	<a href="#">gi 400404</a>	16/6.12	16.3/6.42	103	<a href="#">apoplast</a> <a href="#">chloroplast</a> <a href="#">peroxisome</a> <a href="#">plasma</a> <a href="#">membrane</a> <a href="#">vacuole</a>
104	probable inactive purple acid phosphatase 29-like	<i>Glycine max</i>	<a href="#">gi 356573279</a>	52.5/5.24	42.8/8.1	130	
38	glutathione S-transferase-like protein	<i>Dianthus caryophyllus</i>	<a href="#">gi 390979559</a>	27.3/7.48	24.4/6.06	118	<a href="#">cytosol</a>
18	ubiquinol-cytochrome c reductase	<i>Plantago major</i>	<a href="#">gi 106879619</a>	38.4/5.85	8.4/9.46	39	<a href="#">mitochondrial</a>
21	NADH dehydrogenase	<i>Picea sitchensis</i>	<a href="#">gi 116792579</a>	35.4/6.12	15/9.34	39	<a href="#">Mitochondrial</a>
41	NADPH oxidase	<i>Triticum aestivum</i>	<a href="#">gi 51860180</a>	24.3/7	30/8.6	43	<a href="#">Membrane</a>
17	glutamate decarboxylase 5	<i>Arabidopsis thaliana</i>	<a href="#">gi 15229517</a>	38.4/6.07	56.2/5.42	41	<a href="#">plasmodesma</a>
101	glutamate decarboxylase 5	<i>Arabidopsis thaliana</i>	<a href="#">gi 15229517</a>	41.9/7.13	56.2/5.42	41	<a href="#">plasmodesma</a>
103	glutamate decarboxylase 5	<i>Arabidopsis thaliana</i>	<a href="#">gi 15229517</a>	41.3/6.78	56.2/5.42	41	<a href="#">plasmodesma</a>
<b>PROTEINAS DE RESERVA</b>							
22	11S globulin seed storage protein	<i>Amaranthus hypochondriacus</i>	<a href="#">gi 558149</a>	35.6/5.92	57/6.69	49	<a href="#">protein body</a>
42	11S globulin seed storage protein	<i>Amaranthus hypochondriacus</i>	<a href="#">gi 558149</a>	21.3/6.7	57/6.69	127	<a href="#">protein body</a>
88	11S globulin seed storage protein	<i>Amaranthus hypochondriacus</i>	<a href="#">gi 558149</a>	11.9/4.8	57/6.69	44	<a href="#">protein body</a>
<b>METABOLISMO DE ACIDOS NUCLEICOS</b>							
64	nucleolysin TIAR	<i>Vitis vinifera</i>	<a href="#">gi 296085190</a>	54/8.64	16.6/9.43	32	Nucleus
25	Uridylate kinase	<i>Ricinus communis</i>	<a href="#">gi 255571119</a>	32.5/5.9	23.5/5.24	32	<a href="#">Cytoplasm</a>

SINTESIS DE PROTEINAS							
76	40S ribosomal protein S21-2-like isoform 2	<i>Solanum lycopersicum</i>	<a href="#">gi 116778639</a>	12.8/ 6.74	9.1/7.8	41	<a href="#">cytosolic ribosome</a>
85	SPOC dand transcription elongation factor S-II domain protein	<i>Arabidopsis thaliana</i>	<a href="#">gi 42573473</a>	35.5/3.97	109.6/ 5.13	39	
VIAS DE SEÑALIZACION							
2	S-receptor kinase-like	<i>Oryza sativa</i>	<a href="#">gi 56202248</a>	79.41/7.06	57/6.79	41	<a href="#">Membrane</a>
PLEGAMIENTO DE PROTEINAS							
53	18.5 kDa class I heat shock protein-like	<i>Citrullus lanatus</i>	<a href="#">gi 315932720</a>	18/6.53	18.2/5.88	34	<a href="#">Cytoplasm</a>
82	small heat shock protein 17.4BCII	<i>Boea hygrometrica</i>	<a href="#">gi 496684331</a>	18.2/6.34	17.5/6.32	77	<a href="#">Cytoplasm</a>
MARCAJE DE PROTEINAS							
37	PREDICTED: proteasome subunit beta type-6-like	<i>Glycine max</i>	<a href="#">gi 351721274</a>	27/5.23	25.2/5.3	84	<a href="#">cytoplasm</a>
51	Calreticulin	<i>Beta vulgaris</i>	<a href="#">gi 11131631</a>	58/4.2	48.1/4.45	65	<a href="#">Endoplasmic reticulum</a>
PROTEINAS DE FUNCION DESCONOCIDA							
29	PREDICTED: uncharacterized protein LOC101246315	<i>Solanum lycopersicum</i>	<a href="#">isoTiG16947</a>	31.9/6.69	27.3/6.36	83	
40	hypothetical protein MTR_5g033270	<i>Medicago truncatula</i>	<a href="#">gi 357485713</a>	24.4/7.15	109.8/5.88	38	
45	predicted protein	<i>Populus trichocarpa</i>	<a href="#">gi 224131924</a>	21.2/5.74	22.3/6.16	39	
100	Uncharacterized protein	<i>Lotus japonicus</i>	DN608968	10.3/8.33	9,348/-	57	

### 5.3 Predicción estructural de la proteína LEA 3 por homología secuencial

El modelaje de la estructura tridimensional predijo una estructura constituida principalmente por cadenas alfa hélices conectadas por regiones cortas de estructura aleatoria que le otorgan flexibilidad a dicha estructura; lo cual concuerda con las características estructurales de la proteína (Zhang y Zhao 2003) (**Figura 15**).



**Figura 15. Estructura predicha de la proteína LEA 3 de amaranto**

El grupo 3 de las proteínas LEA, posee un motivo de TAQAAKEKAGE que genera una estructura secundaria de alfa hélice; aunque existen reportes que mencionan que las proteínas LEA 3 poseen una estructura desordenada en solución pero que adquieren una estructura mas ordenada cuando se encuentran bajo condiciones de poco agua (Hong-Bo y cols. 2005). Posiblemente esta flexibilidad estructural le permite llevar a cabo su función de protección hacia otras proteínas en un estado de estrés como es la dormancia.



## VI. Conclusiones

En el presente trabajo se encontraron diferencias cualitativas y cuantitativas en las proteínas de reserva de los cinco cultivos del grano de amaranto de las especies *A. hypochondriacus* y *A. cruentus*. Demostrando que las cantidades relativas de proteínas de reserva pueden variar de especie a especie e inclusive en diversos cultivos de la misma especie. También se determinó a través de un patrón de electroforesis de una dimensión que la especie *A. cruentus* posee una banda característica de alrededor de 58 kDa que por su parte *A. hypochondriacus* no posee. Esta característica puede ser utilizada para realizar una identificación y clasificación rápida de estas dos especies de amaranto. En los últimos años se ha reportado la relevancia las proteínas de reserva, no solo por su valor nutricional, sino por ser un reservorio de importantes péptidos bioactivos; en el caso de este trabajo se detecto la presencia de la proteína tipo lunasin en los cinco cultivos analizados del grano de amaranto. Lo cual indica que esta proteína posee una función trascendental dentro del desarrollo del grano del amaranto, independientemente de la especie de la especie en cuestión.

Otro aspecto que refuerza la idea de la importancia de las proteínas del grano es la presencia de proteínas con residuos fosforilados; ya que esta modificación postraducciona forma parte de los procesos metabólicos esenciales en el metabolismo celular, la activación del metabolismo en general antes de la germinación es muy importante para la correcta germinación de la planta; así que las diferencias en los patrones de fosforilación de las proteínas del grano.

Finalmente, se realizó la primera aproximación del mapa proteómico del grano de *A. cruentus*, en el que, con base en el análisis e identificación de sus proteínas, se reportan por primera vez proteínas de suma importancia en la tolerancia del estrés (por ejemplo proteínas LEA), y un conjunto de proteínas que confieren una resistencia innata al grano de amaranto a sobrellevar la dormancia.

## VII. Anexos

**Tabla 7. Información ampliada de las proteínas identificadas en el grano de amaranto (*A. cruentus* cv Amaranteca)**

NO.	PROTEIN NAME	ACCESS NUMBER	EXP KD A/PI	ORTHO LOG	ACCESS NUMBER ACCORDING TO NCBI	THE ORIGINAL KDA/PI	MA SC OT SC ORE	P M/ S C	E-VA LUE	PEPTIDES
2	S-receptor kinase-like	-	79.41/7.06	Oryza sativa Japonica Group	<a href="#">gi 56202248</a>	57/6.79	41	1/1%		R.GVLDGGK.V
6	LEA protein group 3	<a href="#">isoTIG16059</a>	53.3/6.3	Brassica napus	<a href="#">gi 213053669 ACJ39155.1</a>	43.5/5.64	51	-	1e-42	K.DAAGYVADK.A R.EKGYDVEDEAAER.A K.GYDVEDEAAER.A R.SGEDKTVTVTVEESPPGAAA DLLK.N
10	glyceraldehyde-3-phosphate dehydrogenase	-	45.4/6.94	Salvia elegans	<a href="#">gi 343131368</a>	21.4/6.21	53	4/16%	-	R.SSIFDAK.A K.VVISAPSK.D K.VLPALDGK.L K.LTGMSFR.V + Oxidation (M) R.EGIEMTGK.F + Oxidation (M)
17	glutamate decarboxylase 5	-	38.4/6.07	Arabidopsis thaliana	<a href="#">gi 15229517</a>	56.2/5.42	41	1/1%		
18	ubiquinol-cytochrome c reductase	-	38.4/5.85	Plantago major	<a href="#">gi 106879619</a>	8.4/9.46	39	2/68%	-	K.AVTYALSPYNQKVMPLWK.D K.FSDHWLGAVCIFGPVFGTM QFAEWYTEREK.M
19	dormancy related protein, putative	-	35.5/6.6	Arabidopsis thaliana	<a href="#">gi 12322163</a>	31.4/5.92	79	1/5%		K.EGSSIINTTSVNAYK.G
20	dormancy related protein, putative	-	35.1/6.32	Arabidopsis thaliana	<a href="#">gi 12322163</a>	31.4/5.92	56	3/11%	-	K.EGSSIINTTSVNAYK.G K.GAIVAFTR.G R.GLALQLAEK.G
21	NADH dehydrogenase	-	35.4/6.12	Picea sitchensis	<a href="#">gi 116792579</a>	15/9.34	39	2/15%	-	.MASPLRQVGTGANSQSLEEA R.K M.ASPLRQVGTGANSQSLEEA R.K
22	11S globulin seed storage protein	-	35.6/5.92	Amaranthus hypochondriacus	<a href="#">gi 558149</a>	57/6.69	49	2/3%		R.NTGNIFR.G R.LTALEPTNR.I
24	hypothetical protein SORBIDRAFT_06g011080	-	32.6/6.07	Sorghum bicolor	<a href="#">gi 242075416</a>	49/7.15	38	2/10%		K.NRALETLLSVGTAGIGAGAD SQTDPFAAAMAADPASR.A + Oxidation (M) R.ALETLLSVGTAGIGAGADSQ TDPFAAAMAADPASRAAFVAE AVALAR.D + Oxidation (M)

25	Uridylate kinase	<a href="#">isoTiG14015</a>	32.5/5.9	Ricinus communis	<a href="#">gij255571119</a>	23.5/5.24	32	-	6e-105	K.SGSENGTMIQNMIE Oxidation (M) R.ALFFESITGIVPEIVLFFDCSEE EMEK.R Oxidation (M)
27	PREDICTED: (3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase-like	-	30.7/7.42	Glycine max	<a href="#">gij255640918</a>	19.9/9.72	32	2/37%	-	M.AASAFSNTLVSPISLPSKPKP KAISLQPK.F K.TTSLTTFCSDDAANAPQHDT PIELRYPAFPTVMDINQIR.D + Oxidation (M)
29	PREDICTED: uncharacterized protein LOC101246315	<a href="#">isoTiG16947</a>	31.9/6.69	Solanum lycopersicum	<a href="#">gij460376573</a>	27.3/6.36	83	6/47%	9e-72	K.SGLFVNPR.V R.VPEAVAKPELEK.F R.GDLLPIGEPALMMSPQGEVE GMVK.A 3 Oxidation (M) K.APLLQQR.D K.YGISSTK.L R.VDITGPTFINPNADYWK.K R.ESQGAANDAEIESLAR.F
30	cystatin	<a href="#">isoTiG19090</a>	29.1/6.21	Amaranthus hypochondriacus	<a href="#">gij110748606</a>	18.5/6	216	-	5e-141	K.ENALLEFAR.V K.ELQEFK.H K.HTEDSPSFTSSDLGAIR.E K.AEVVEDTAK.F K.SSDGNYNLNK.M K.MGNIQPEIENQ.-
31	cystatin	-	29.6/6.02	Amaranthus hypochondriacus	<a href="#">gij110748606</a>	27.8/5.67	144	3/14%	-	R.ESQGAANDAEIESLAR.F K.ENALLEFAR.V K.MGNIQPEIENQ.- + Oxidation (M)
33	hypothetical protein PRUPE_ppa024816m9	-	33.6/5.06	Prunus persica	<a href="#">gij462409013</a>	37.3/6.84	41	1/3%	-	K.NDETVER.L
37	PREDICTED: proteasome subunit beta type-6-like	<a href="#">isoTiG02916</a>	27/5.23	Glycine max	<a href="#">gij356507848</a>	25.2/5.3	84	-	8e-142	K.DGVVLGADSR.T R.TSTGVYVANR.A R.SGSAADSQIVSDYVR.H K.VAANLVR.L
38	glutathione S-transferase-like protein	<a href="#">isoTiG08680</a>	27.3/7.48	Dianthus caryophyllus	<a href="#">gij390979559</a>	24.4/6.06	118	-	4e-92	K.YIAYTYENQGTPLIYK.Q K.GMYGLQTDEAAVEEHTAK.L K.VLDVYEAR.L
40	hypothetical protein MTR_5g033270	-	24.4/7.15	Medicago truncatula	<a href="#">gij357485713</a>	109.8/5.88	38	2/68%	-	R.TPTLNNNDTLLSLFHFAS NTDLPILHLGKMK.V K.VSHASIHVVEEPTTDGMPD TAGGLSLR.I
41	NADPH oxidase	-	24.3/7	Triticum aestivum	<a href="#">gij51860180</a>	30/8.6	43	2/27%	-	.YMFVQCQPAVSPFEWHFHSITS APGDEYLSIHVRQLGDWTR.E R.NIEMHNYLTSVYEEGDARS

										ALITMLQALNHAK.N + 2 Oxidation (M)
4 2	11S globulin seed storage protein	-	21.3 /6.7	Amarant hus hypocho ndriacus	<a href="#">gil558149</a>	57/6. 69	127	4/ 9 %	-	R.LTTVNSFNPILR.H K.TSENAFQSLAGR.T + Oxidation (M) R.SLPIDVVSNIYQISR.E R.EEAFGLK.F
4 5	predicted protein	-	21.2 /5.7 4	Populus trichocar pa	<a href="#">gil2241319 24</a>	22.3/ 6.16	39	2/ 40 %	-	K.LALWHTRTFNPIMTHEEELP LMTTMGFVGLPPSSSSSWK.E K.IDGLHLYTYQAFIDSVNLYLE MSNISDLFHVGRMPLSR.N K.WVCELQRPR.K
4 8	dehydration -responsive element- binding protein 1F- like	-	25.5 /3.9	Vitis vinifera	<a href="#">gil2254409 70</a>	25.9/ 9	40	2/ 10 %	-	K.SICFPNFLTRLLYNLL.-
5 0	Alpha chain of nascent polypeptide associated complex	<a href="#">isoTiG18 763</a>	29.8 /4.0 9	Theobro ma cacao	<a href="#">gil5087276 50</a>	22/4. 3	157	-	3e- 60	K.SPTSETYVFFGEAK.I K.IEDLSSQLQTQAAEQFK.T
5 1	Calreticulin	<a href="#">isoTiG11 813</a>	58/4 .2	Beta vulgaris	<a href="#">gil1113163 1</a>	48.1/ 4.45	65	-	0	K.SGTLFDNVLITDDPEYAK.Q
5 2	LEA5	<a href="#">isoTiG02 577</a>	19/8	Vitis amurens is	<a href="#">gil3241037 65</a>	???? ?	117	-	5e- 33	K.DQTGSFASDK.A K.TGGILQR.T K.TGGILQR.T K.SMAQGAADAVK.N K.SMAQGAADAVK.N K.NTFGMGEPEEDDPIK.T -SAFANAR.M
5 3	18.3 kDa class I heat shock protein	<a href="#">isoTiG12 384</a>	19.3 /5.7 4	Chenopo dium rubrum	<a href="#">gil462322/ Q05832.1</a>	18.3/ 6.76	58	-	8e- 57	K.ADLPGVK.K
5 5	nucleoside diphosphat e kinase I	<a href="#">ConTiG0 0114</a>	18.2 /6.0 7	Mesemb ryanthe mum crystallin um	<a href="#">gil6225750</a>	16.3/ 6.3	75	-	4e- 95	K.IIGATNPLASEPGTIR.G R.NVIHGSDSVESATK.E
5 6	Nucleoside diphosphat e kinase 1	-	16/6 .5	Spinacia oleracea	<a href="#">gil400404</a>	16.3/ 6.42	103	2/ 16 %	-	K.LIGATNPLASEPGTIR.G R.GDFAIDIGR.N
5 7	dehydroasc orbate reductase	<a href="#">isoTiG14 118</a>	11.6 /7.4 6	Nicotian a tabacum	<a href="#">gil2819242 7</a>	23.6/ 7.7	40	-	1e- 12 1	K.DGNDGSEQALLDELK.A
6 4	nucleolysin TIAR	-	54/8 .64	Vitis vinifera	<a href="#">gil2960851 90</a>	16.6/ 9.43	32	2/ 27 %	-	.MLFCFSHLFGCKGYVGSSEDW AFK.R + Oxidation (M) K.MGGVHALMHPQGQHPFK.Q + Oxidation (M) K.VLPNVIK.S
6 7	pathogenes is-related protein	<a href="#">isoTiG12 599</a>	17.4 / 6.88	Spinacia oleracea	<a href="#">gil4447924 85</a>	17.9/ 4.72	94	-	2e- 59	K.VNELDVNNYVVK.F
7 1	18.5 kDa class I heat shock	-	18/6 .53	Citrullus lanatus	<a href="#">gil3159327 20</a>	18.2/ 5.88	34	4/ 16 %	-	K.ADLPLGLK.K K.EEVKVEVEDGK.V

	protein-like									K.VEVEDGK.V K.VIQISGER.S -IDLYIPR.K
76	40S ribosomal protein S21-2-like isoform 2	<a href="#">ConTiG25661</a>	12.8 / 6.74	Solanum lycopersicum	<a href="#">gil116778639</a>	9.1/7.8	41	-	8e-37	R.AQGDADSALDR.L
81	LEA5	<a href="#">isoTiG02577</a>	19.1 / 6.47	Vitis amurensis	<a href="#">gil324103765</a>	¿???	36	-	5e-33	K.SMAQGAADAVK.N Oxidation (M) K.NTFGMGEPEEDDPIK.T Oxidation (M)
82	small heat shock protein 17.4BCII	-	18.2 / 6.34	Boea hygrometrica	<a href="#">gil496684331</a>	17.5/6.32	77	2/19%	-	K.ADLPGLK.K K.EEVKVEVEDGK.V K.VEVEDGK.V K.VIQISGER.S
83	Nucleoside diphosphate kinase 1	<a href="#">ConTiG0114</a>	16/6.12	Mesembryanthemum crystallinum	<a href="#">gil62257501081372.1</a>	16.3/6.3	103	-	4e-95	R.GLIGDIISR.F K.IIGATNPLASEPGTIR.G R.GDFAIDIGR.N
84	EM protein	-	16.6 / 6.16	Vigna radiata	<a href="#">gil1141782</a>	12.2/8.86	55	2/18%	-	K.QGETVVPGGTGGK.S K.GIEIDESK.F
85	SPOC dand transcription elongation factor S-II domain protein	-	35.5 / 3.97	Arabidopsis thaliana	<a href="#">gil42573473</a>	109.6 / 5.13	39	1/0%	-	R.DSLIEVAK.S
88	11S globulin seed storage protein	-	11.9 / 4.8	Amaranthus hypochondriacus	<a href="#">gil558149</a>	57/6.69	44	2/4%	-	R.LTALEPTNR.I K.TSENAFMQSLAGR.T + Oxidation (M)
100	Uncharacterized protein	-	10.3 / 8.33	Lotus japonicus	<a href="#">DN608968</a>	9,348 / -	57	1/3%	-	R.ISSPATNISR.T
101	glutamate decarboxylase 5	-	41.9 / 7.13	Arabidopsis thaliana	<a href="#">gil15229517</a>	56.2/5.42	42	1/1%	-	R.EGIEMTGK.F + Oxidation (M)
103	glutamate decarboxylase 5	-	41.3 / 6.78	Arabidopsis thaliana	<a href="#">gil15229517</a>	56.2/5.42	42	1/1%	-	R.EGIEMTGK.F + Oxidation (M)
104	probable inactive purple acid phosphatase 29-like	<a href="#">isoTiG10344</a>	52.5 / 5.24	Glycine max	<a href="#">gil356573279</a>	42.8/8.1	130	-	1e-141	K.NTSLK.N K.SIVTWK.R K.DGSWGPVK.S R.EAVMEYIVGMK.N + 2 Oxidation (M) R.SMDEMFGPAIESNK.A + 2 Oxidation (M) R.VICPGVNSGFFDTMVTAGDIK.A + Oxidation (M)
106	probable inactive purple acid	<a href="#">isoTiG10344</a>	52.2 / 5.01	Glycine max	<a href="#">isoTiG10344</a>	42.8/8.1	78	-	1e-141	R.SMDEMFGPAIESNK.A 2 Oxidation (M)

phosphatas e 29-like										K.SIVTWK.R
-------------------------	--	--	--	--	--	--	--	--	--	------------

A

```

      10      20      30      40      50
GGAAAGACAA GAGGAGCAAC AGAAGCAATG AAGAACAAGG CCAACAAGC
  G K T R G A T E A M K N K A Q Q A

      60      70      80      90      100
AAAAGACAAG ACTATGGAAA CAGCCCAAGC AGCAAAAGAA AAGACTTCTG
  K D K T M E T A Q A A K E K T S
      -----

      110     120     130     140     150
AGAAAGCCTA CGAGGCTAAA GACAAAACAT CTAACATTGC CCAATCTGCT
  E K A Y E A K D K T S N I A Q S A

      160     170     180     190     200
AAGGGTCAGG CCAACAGTC CAAGGATCAA ACTGGATCCT TCGCTTCTGA
  K G Q A Q Q S K D Q T G S F A S D

      210     220     230     240     250
CAAGGCCCAA GCAGCCAAGG ACAAGGCTTA TCAAGCAGCC CAAGCCACCA
  K A Q A A K D K A Y Q A A Q A T

      260     270     280     290     300
AGGAGAAGGC GTCCGACATG ACGGAGTACG CTAAGGAGAC GCGGAAGCC
  K E K A S D M T E Y A K E T A E A

      310     320     330     340     350
GGAAAGGAGA AACTGGAGG AATTCTTCAA AGGACAGGTG AACAAGTCAA
  G K E K T G G I L Q R T G E Q V K

      360     370     380     390     400
GAGTATGGCT CAAGGTGCCG CTGATGCTGT TAAGAATACT TTTGGGATGG
  S M A Q G A A D A V K N T F G M

      410
GCGAACCTGA AGAAGATG
  G E P E E D

```

Figura 1S. Secuencia de aminoácidos y nucleótidos de la proteína LEA 3 de amaranto

## Bibliografía

- Agrawal GK, Hajduch M, Graham K, Thelen JJ. In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. *Plant Physiol* 2008;148:504–18.
- Aguilar-Hernández, H. S., Santos, L., León-Galván, F., Barrera-Pacheco, A., Espitia-Rangel, E., De León-Rodríguez, A., Barba de la Rosa, A. P. Identification of calcium stress induced genes in amaranth leaves through suppression subtractive hybridization. *Journal of plant physiology* 2011; 168(17): 2102-2109.
- Aragão TCFR, Souza PAS, Uchôa AF, Costa IR, Bloch Jr C, Campos FAP. Characterization of a methionine-rich protein from the seeds of *Cereus jamacaru* Mill. (Cactaceae). *Braz J Med Biol Res* 2000;33:897–903.
- Ballabio C, Uberti F, Di Lorenzo C, Brandolini A, Penas E, Restani P. Biochemical and immunochemical characterization of different varieties of amaranth (*Amaranthus L. ssp.*) as a safe ingredient for gluten-free products. *J. Agric. Food Chem.*, 2011; 59 (24): 12969–12974
- Barba de la Rosa, A. P., Paredes-Lopez, O., Gueguen, J. Characterization of amaranth globulins by ultracentrifugation and chromatographic techniques. *Journal of agricultural and food chemistry* 1992; 40(6), 937-940.
- Bartels D, Sunkar R. Drought and Salt Tolerance in Plants. *Crit Rev Plant Sci* 2005; 24: 23-58
- Baubec T, Mittelsten Scheid O. Medea in full self-control. *Trends Plant Sci* 2006;11:469–71.



- Bevan M, Bancroft I, Bent E, Love K, Goodman H, Dean C, Bergkamp R, Dirkse W, Van Staveren M, Stiekema W. Analysis of 19 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* 1998;391:485–8.
- Bewley JD. Seed germination and dormancy. *Plant Cell* 1997; 9:1055–1066
- Bietz JA. Cereal prolamin evolution and homology revealed by sequence analysis. *Biochem Genet* 1982;20:1039–53.
- Boucher V, Buitink J, Lin X, Boudet J, Hoekstra FA, Hundertmark M, Renard D, Leprince O. MtPM25 is an atypical hydrophobic late embryogenesis-abundant protein that dissociates cold and desiccation-aggregated proteins. *Plant Cell Environ.* 2010; 33(3):418-30
- Brewis IA, Brennan P. Proteomics technologies for the global identification and quantification of proteins. *Adv Protein Chem Struct Biol* 2010;80:1–44.
- Chakraborty S, Chakraborty N, Datta A. Increased nutritive value of transgenic potato by expressing a nonallergenic seed albumin gene from *Amaranthus hypochondriacus*. *Proc Natl Acad Sci* 2000; 97(7): 3724–3729.
- Cuming AC: LEA Proteins. In *Seed Proteins* Edited por: Peter R. Shewry and Rod Casey. Kluwer Academic Publishers; 1999:753-780.
- Dam S, Laursen BS, Ornfelt JH, Jochimsen B, Staerfeldt HH, Friis C, Nielsen K, Goffard N, Besenbacher S, Krusell L, Sato S, Tabata S, Thøgersen IB, Enghild JJ, Stougaard J. The proteome of seed development in the model legume *Lotus japonicus*. *Plant Physiol* 2009;149:1325–40.

- De Smet I, Lau S, Mayer U, Jürgens G. Embryogenesis: the humble beginnings of plant life. *Plant J* 2010;61:959–70.
- Délano-Frier, John P., Hamlet Avilés-Arnaut, Kena Casarrubias-Castillo, Gabriela Casique-Arroyo, Paula A. Castrillón-Arbeláez, Luis Herrera-Estrella, Julio Massange-Sánchez et al. "Transcriptomic analysis of grain amaranth (*Amaranthus hypochondriacus*) using 454 pyrosequencing: comparison with *A. tuberculatus*, expression profiling in stems and in response to biotic and abiotic stress." *BMC genomics* 2011; 1(12): 363.
- Deng, Zhu Yun, Chun Yan Gong, and Tai Wang. "Use of proteomics to understand seed development in rice." *Proteomics* 2013; 1: 1-17.
- Desiderio C, Rossetti DV, Iavarone F, Messina I, Castagnola M. Capillary electrophoresis-mass spectrometry: recent trends in clinical proteomics. *J Pharm Biomed Anal* 2010;53: 1161–9.
- Dolferus, R., Ji, X., Richards, R. A. Abiotic stress and control of grain number in cereals. *Plant Science* 2011; 181(4): 331-341.
- Dunwell JM, Culham A, Carter CE, Sosa-Aguirre CR, Goodenough PW. Evolution of functional diversity in the cupin superfamily. *Trends Biochem Sci* 2001; 26: 740–746
- Eltayeb AE, Kawano N, Badawi G, Kaminaka H, Sanekata T, Morishima I. Enhanced tolerance to ozone and drought stresses in transgenic tobacco overexpressing dehydroascorbate reductase in cytosol. *Physiol. Plant.* 2006; 127: 57–65.

- Frigerio L, Hinz G, Robinson DG. Multiple vacuoles in plant cells: rule or exception? *Traffic* 2008;9:1564–70.
- Gallardo K, Firnhaber C, Zuber H, Hericher D, Belghazi M, Henry C, Küster H, Thompson RD. A combined proteome and transcriptome analysis of developing *Medicago truncatula* seeds. *Mol Cell Proteomics* 2007;6:2165–79.
- Gallardo K, Le Signor C, Vandekerckhove J, Thompson RD, Burstin J. Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiol* 2003;133:664–82.
- Galvez AF, De Lumen BO. A soybean cDNA encoding a chromatinbinding peptide inhibits mitosis of mammalian cells. *Nature* 1999; 17: 495-500.
- Gibbs PE, Strongin KB, McPherson A. Evolution of legume seed storage proteins—a domain common to legumins and vicilins is duplicated in vicilins. *Mol Biol Evol* 1989;6:614–23.
- Goyal, Kshamata, L. Walton, and Alan Tunnacliffe. "LEA proteins prevent protein aggregation due to water stress." *Biochem. J* 2005; 388: 151-157.
- Grelet J, Benamar A, Teyssier E, Avelange-Macherel MH, Grunwald D, Macherel D. Identification in Pea Seed Mitochondria of a Late-Embryogenesis Abundant Protein Able to Protect Enzymes from Drying. *Plant Physiology* 2005; 137: 157–167.
- Häger KP, Müller B, Wind C, Erbach S, Fischer H. Evolution of legumin genes: loss of an ancestral intron at the beginning of angiosperm diversification. *FEBS Lett* 1996;387:94–8.

- Hajduch M, Ganapathy A, Stein JW, Thelen JJ. A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol* 2005;137:1397–419.
- Haughn G, Chaudhury A. Genetic analysis of seed coat development in *Arabidopsis*. *Trends Plant Sci* 2005;10:472–7.
- Hernández-Ledesma, B., Hsieh, C.-C., de Lumen, B.O., 2009. Lunasin, a novel seed peptide for cancer prevention. *Peptides* 2009; 30: 426-430.
- Huerta-Ocampo, José Ángel, Erika Patricia Briones-Cerecero, Guillermo Mendoza-Hernández, Antonio De León-Rodríguez, and Ana Paulina Barba de la Rosa. "Proteomic analysis of amaranth (*Amaranthus hypochondriacus* L.) leaves under drought stress." *International Journal of Plant Sciences* 2009; 8(170): 990-998.
- Jeong, H.J., Park, J.H., Lam, Y., de Lumen, B.O. Characterization of lunasin isolated from soybean. *Journal of Agricultural and Food Chemistry* 2003; 51: 7901-7906.
- Jia, X. Y., Xu, C. Y., Jing, R. L., Li, R. Z., Mao, X. G., Wang, J. P., Chang, X. P. Molecular cloning and characterization of wheat calreticulin (CRT) gene involved in drought-stressed responses. *Journal of experimental botany*, 2008; 59(4): 739-751.
- K. Yamaguchi-Shinozaki, K. Shinozaki, Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses, *Annu. Rev. Plant Biol.* 2006; 57: 781–803.

- Kreis M, Forde BG, Rahman S, Mifflin BJ, Shewry PR. Molecular evolution of the seed storage proteins of barley, rye and wheat. *J Mol Biol* 1985;183:499–502.
- Kwon SY, Choi SM, Ahn YO, Lee HS, Lee HB, Park YM. Enhanced stress-tolerance of transgenic tobacco plants expressing a human dehydroascorbate reductase gene. *J. Plant Physiol.* 2003; 160: 347–353.
- Lara-Núñez A, De JN, Vázquez-Ramos JM. Maize D4;1 and D5 cyclin proteins in germinating maize. Associated kinase activity and regulation by phytohormones. *Physiol Plant* 2008; 132:79–88
- Lawton JW. Zein: a history of processing and use. *Cereal Chem* 2002;79:1–18.
- Lee G, Pokala N, Vierling E. Structure and in vitro molecular chaperone activity of cytosolic small heat shock proteins from pea. *J Biol Chem* 1995; 270: 10432–10438
- Lee G, Roseman A, Saibil H, Vierling E. A small heat shock protein stably binds heat denatured model substrates and can maintain a substrate in a folding-competent state. *EMBO J* 1997; 16: 659–671
- Maldonado-Cervantes, E., Jeong, H.J., León-Galván, F., Barrera-Pacheco, A., De León-Rodríguez, A., Gonzales de Maija, E., de Lumen, B., Barba de la Rosa, A.P. Amaranth lunasin-like peptide internalizes into the cell nucleus and inhibits chemical carcinogen-induced transformation of N1H-3T3 cells. *Peptides* 2010; 31: 1635e1642.

- Marrs, KA. The functions and regulation of glutathione S-transferases in plants. *Annual review of plant biology*, 1996; 47(1): 127-158.
- Miernyk JA, Preťová A, Olmedilla A, Klubicová K, Obert B, Hajduch M. Using proteomics to study sexual reproduction in angiosperms. *Sex Plant Reprod* 2011;24:9–22.
- Mitchell RAC, Lovegrove A, Shewry PR. Lunasin in cereal seeds: What is the origin? *J of Cereal Sci* 2013; 57: 267-269
- Miyawaki, M., Aito, M, Ito, N, Yanagawa, Y, Kendrick, RE., Tanaka, K, Nakagawa, H. Changes in proteasome levels in spinach (*Spinacia oleracea*) seeds during imbibition and germination. *Bioscience, biotechnology, and biochemistry* 1997; 61(6): 998-1001.
- Oliver, M. J., Tuba, Z., & Mishler, B. D. The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology* 2000; 151(1): 85-100.
- Osborne TB. *The Vegetable Proteins*, 2nd Edn. Longmans Green and Co., 1924 London
- Peck, S. C., Update on proteomics in Arabidopsis. Where do we go from here? *Plant Physiol.* 2005; 138: 591–599.
- Qi PF, Wei YM, Ouellet T, Chen Q, Tan X, Zheng YL. The gamma-gliadin multigene family in common wheat (*Triticum aestivum*) and its closely related species. *BMC Genomics* 2009;10:168.
- Qureshi MI, Qadir S, Zolla L. Proteomics-based dissection of stress-responsive pathways in plants. *J Plant Physiol* 2007;164:1239–60.

- Rabilloud T, Chevallet M, Luche S, Lelong C. Two-dimensional gel electrophoresis in proteomics: past, present and future. *J Proteomics* 2010;73:2064–77.
- Rascón-Cruz Q, Sinagawa-García S, Osuna-Castro JA, Bohorova N, Paredes-López O. Accumulation, assembly, and digestibility of amarantin expressed in transgenic tropical maize. *Theor Appl Genet.* 2004;108(2):335-42
- Roberts EH. Predicting the storage life of seeds. *Seed Sci Technol* 1973; 1:499–514.
- Shewry PR, Napier JA, Tatham AS. Seed storage proteins: structures and biosynthesis. *Plant Cell* 1995;7:945–56.
- Shewry PR, Tatham AS. The prolamin storage proteins of cereal seeds: structure and evolution. *Biochem J* 1990;267: 1–12.
- Silva-Sanchez C, Gonzalez-Castañeda J, De Leon-Rodriguez A, De La Rosa APB. 2004. Functional and rheological properties of amaranth albumins extracted from two Mexican varieties. *Plant Food Hum Nutr* 2004; 59:169–74.
- Trewayas AJ, Malho R. Signal perception and transduction: the origin of the phenotype. *Plant Cell* 1997; 9:1181–1195
- Valdes-Rodriguez, S., Cedro-Tanda, A., Aguilar-Hernandez, V., Cortes-Onofre, E., Blanco-Labra, A., Guerrero-Rangel, A. Recombinant amaranth cystatin (AhCPI) inhibits the growth of phytopathogenic fungi. *Plant Physiology and Biochemistry* 2010; 48(6): 469-475.
- Valdes-Rodriguez, S., Guerrero-Rangel, A., Melgoza-Villagómez, C., Chagolla-López, A., Delgado-Vargas, F., Martínez-Gallardo, N. Délano-Frier,

J. Cloning of a cDNA encoding a cystatin from grain amaranth (*Amaranthus hypochondriacus*) showing a tissue-specific expression that is modified by germination and abiotic stress. *Plant Physiology and Biochemistry* 2007; 45(10): 790-798.

- Van Loon, L. C., Van Strien, E. A. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiological and Molecular Plant Pathology* 1999; 55(2): 85-97.
- Vierling E. The small heat shock proteins in plants are members of an ancient family of heat induced proteins. *ACTA Physiol Planta* 1997; 19: 539–547.
- Weber H, Borisjuk L, Wobus U. Molecular physiology of legume seed development. *Annu Rev Plant Biol* 2005;56: 253–79.
- Wehmeyer, N., Vierling, E. 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* 2000; 122(4): 1099-1108.
- Wienkoop S, Baginsky S, Weckwerth W. *Arabidopsis thaliana* as a model organism for plant proteome research. *J Proteomics* 2010;73:2239–48.
- Wise, Michael J. "LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles." *Bmc Bioinformatics* 2003; 1 (4): 1-19
- Xu JH, Messing J. Amplification of prolamin storage protein genes in different subfamilies of the Poaceae. *Theor Appl Genet* 2009;119:1397–412.



- Yan, S., Tang, Z., Su, W., & Sun, W. Proteomic analysis of salt stress-responsive proteins in rice root. *Proteomics* 2005; 5(1), 235-244.
- Yang, H. L., Zhao, Y. R., Wang, C. L., Yang, Z. L., Zeng, Q. Y., Lu, H. Molecular characterization of a dehydroascorbate reductase from *Pinus bungeana*. *Journal of Integrative Plant Biology* 2009; 51(11): 993-1001.
- Zhang L.S., Zhao W.M., LEA protein functions to tolerance drought of the plant, *Plant Physiology Communications* 2003; 39 (1): 61–66.

## ARTICULO CIENTÍFICO ACEPTADO

**Correo de aceptación de artículo en la revista Journal of Cereal Science**

**From:** "Journal of Cereal Science" <[jcs@elsevier.com](mailto:jcs@elsevier.com)>

**Subject:** Decision on Your Submission JCS13-414R1

**Date:** February 17, 2014 9:05:28 AM CST

**To:** [apbarba@ipicyt.edu.mx](mailto:apbarba@ipicyt.edu.mx)

Ref.: Ms. No. JCS13-414R1

Title: Characterization of *Amaranthus cruentus* L. seed proteins by 2-DE and LC/MS-MS: Identification and cloning of a novel late embryogenesis-abundant protein

Authors: Enrique Maldonado-Cervantes, MsSc; José Á Huerta-Ocampo, Ph.D.;

Gabriela M Montero-Morán, Ph.D.; Alberto Barrera-Pacheco, MsSc; Eduardo Espitia-Rangel, Ph.D.; Ana Paulina Barba de la Rosa, Ph.D.

Article Type: Research Article

Dear Dr. Barba de la Rosa,

I am pleased to inform you that your work has now been accepted for publication in the Journal of Cereal Science.

Your paper will be passed to the Editor-in-Chief for final editing. You may receive your edited paper direct from him so we ask that you could please respond promptly to any queries he may have. This will ensure that your paper will be published without delay.

When your paper is published on ScienceDirect, you want to make sure it gets the attention it deserves. To help you get your message across, Elsevier has developed a new, free service called AudioSlides: brief, webcast-style presentations that are shown (publicly available) next to your published article. This format gives you the opportunity to explain your research in your own words and attract interest. You will receive an invitation email to create an AudioSlides presentation shortly. For more information and examples, please visit <http://www.elsevier.com/audioslides>.

Thank you for submitting your work to Journal of Cereal Science.

With kind regards,

Domenico Lafiandra, Laurea, Editor  
Journal of Cereal Science

Reviewer #2: After reading the new version of the article once the authors have made the corresponding changes according to the reviewers' comments, I believe that the quality of the manuscript has improved, and no additional changes are needed.

Manuscript Number:

Title: Characterization of *Amaranthus cruentus* L. seed proteins by 2-DE and LC/MS-MS: Identification and cloning of a novel late embryogenesis-abundant protein

Article Type: Research Article

Keywords: Amaranth; Late embryogenesis abundant proteins; proteomics; seed storage proteins

Corresponding Author: Dr. Ana Paulina Barba de la Rosa, Ph.D.

Corresponding Author's Institution: IPICYT, Instituto Potosino de Investigación Científica y Tecnológica A.C.

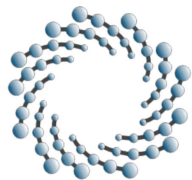
First Author: Enrique Maldonado-Cervantes, MsSc

Order of Authors: Enrique Maldonado-Cervantes, MsSc; José Á Huerta-Ocampo, Ph.D.; Gabriela M Montero-Morán, Ph.D.; Alberto Barrera-Pacheco, MsSc; Eduardo Espitia-Rangel, Ph.D.; Ana Paulina Barba de la Rosa, Ph.D.

Abstract: Amaranth is taking great attention as an important cereal crop that could fulfill food requirements for the growing population, especially in developing countries. However, still is not well known the protein composition of these seeds. We have used two-dimensional gel electrophoresis (2-DE) to characterize amaranth seed proteome. About 400 proteins spots were resolved on 2-DE gels, the most abundant protein spots were analyzed by LC-MS/MS. Identified proteins were related with stress and defense responses, metabolic processes, respiratory and oxide-reduction process, and nutrient reservoir. One abundant spot was identified as a Late Embryogenesis Abundant (LEA) protein, and then the gene from was cloned. The AcLEA cDNA contains a 418 bp ORF encoding a polypeptide of 139 amino acids with a calculated molecular mass of 14,433.8 kDa and a theoretical pI of 8.65. AcLEA corresponded to the Group 5 of LEAs, bioinformatics analysis revealed that is a protein highly hydrophilic with intrinsically ordered large  $\alpha$ -helices. In this study, we report for the first time the presence of late embryogenesis abundant proteins in amaranth seeds. The identification of these proteins gives a tool to start with strategies to improve the quality of amaranth production.

**DOI:10.1016/j.jcs.2014.02.008**

<http://dx.doi.org/10.1016/j.jcs.2014.02.008>



INSTITUTO POTOSINO  
DE INVESTIGACIÓN  
CIENTÍFICA Y TECNOLÓGICA, A.C.

IPICYT

San Luis Potosí, S. L. P., 14 november 2013

**Dr. F. MacRitchie**

Kansas State University,  
Manhattan, KS, USA

**Editor-in-Chief**

*Journal of Cereal Science*

**Dear Dr. MacRitchie**

Please find enclosed our manuscript entitled: “**Characterization of amaranth seed proteins by 2-DE and LC/MS-MS: Identification and cloning of a novel late embryogenesis-abundant protein**”, authors: Enrique Maldonado-Cervantes, José A. Huerta-Ocampo, Gabriela M. Montero-Morán, Alberto Barrera-Pacheco, Eduardo Espitia-Rangel, Ana P. Barba de la Rosa.

Amaranth seed proteins have been studied mainly from *A. hypochondriacus* cultivars, however, *A. cruentus* is taking great attention because has better response to abiotic stress. *A. cruentus* grows in semi-arid areas, but until now, no information in relation about the seed proteins composition is reported for this variety. In this work, we have analyzed the seed protein profile different of different cultivars from *A. cruentus* (Amaranteca and Dorada) and *A. hypochondriacus* (Nutrisol, Revancha, and Annapurna). Differences in protein profile were observed. The 2-DE protein profile of *A. cruentus* cv Amaranteca was obtained and a set of proteins were identified by LC-MS/MS. Among these proteins, a group of LEA, HSPs, DREB proteins were identified. The LEA 5 was cloned and characterized. This is the first LEA proteins are reported to be present in amaranth seeds.

Sincerely

IPICYT: <http://www.ipicyt.edu.mx>;

E-mail: [apbarba@ipicyt.edu.mx](mailto:apbarba@ipicyt.edu.mx)

*“2013, Año de la Lealtad Institucional y Centenario del Ejército Mexicano”*

## HIGHLIGHTS

- Differences in protein profiles were observed among *A. hypochondriacus* and *A. cruentus*
- 2-DE protein profile from *A. cruentus* revealed different LEAs, HSPs, and DREBs proteins
- AcLEA was cloned and characterized
- Isoforms of purple acid phosphatases were identified

1           **Characterization of *Amaranthus cruentus* L. seed proteins by 2-**  
2           **DE and LC/MS-MS: Identification and cloning of a novel late**  
3           **embryogenesis-abundant protein**

4           Enrique Maldonado-Cervantes<sup>a§</sup>, José A. Huerta-Ocampo<sup>a§</sup>, Gabriela M. Montero-  
5           Morán<sup>a</sup>, Alberto Barrera-Pacheco<sup>a</sup>, Eduardo Espitia-Rangel<sup>b</sup>, Ana P. Barba de la  
6           Rosa<sup>a,\*</sup>

7           <sup>a</sup>IPICyT, Instituto Potosino de Investigación Científica y Tecnológica A.C. Camino a la  
8           Presa San José No. 2055, Lomas 4<sup>a</sup> Sección, 78216 San Luis Potosí, S.L.P., México.

9           <sup>b</sup>INIFAP, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Campo  
10           Experimental Bajío, Km 6.5 Carretera Celaya-San Miguel de Allende, Celaya, Gto.,  
11           México.

12           <sup>§</sup>These authors contribute equally to this work

14           \*Corresponding Author: Dr. Ana Paulina Barba de la Rosa

15           Molecular Biology Division

16           Tel: +52 444 8342082; fax: +52 444 8342010

17           E-mail: [apbarba@ipicyt.edu.mx](mailto:apbarba@ipicyt.edu.mx)

18

19

20

21

22

23        **ABSTRACT**

24        Amaranth is taking great attention as an important cereal crop that could fulfill food  
25        requirements for the growing population, especially in developing countries.  
26        However, still is not well known the protein composition of these seeds. We have  
27        used two-dimensional gel electrophoresis (2-DE) to characterize amaranth seed  
28        proteome. About 400 proteins spots were resolved on 2-DE gels, the most  
29        abundant protein spots were analyzed by LC-MS/MS. Identified proteins were  
30        related with stress and defense responses, metabolic processes, respiratory and  
31        oxide-reduction process, and nutrient reservoir. One abundant spot was identified  
32        as a Late Embryogenesis Abundant (LEA) protein, and then the gene from was  
33        cloned. The *AcLEA* cDNA contains a 418 bp ORF encoding a polypeptide of 139  
34        amino acids with a calculated molecular mass of 14,433.8 kDa and a theoretical p/  
35        of 8.65. *AcLEA* corresponded to the Group 5 of LEAs, bioinformatics analysis  
36        revealed that is a protein highly hydrophilic with intrinsically ordered large  $\alpha$ -  
37        helices. In this study, we report for the first time the presence of late  
38        embryogenesis abundant proteins in amaranth seeds. The identification of these  
39        proteins gives a tool to start with strategies to improve the quality of amaranth  
40        production.

41

42

43

44

45        **KEYWORDS**

46        Amaranth; Late embryogenesis abundant proteins; proteomics; seed storage proteins

47

48        **ABBREVIATIONS**

49        2-DE: two-dimensional electrophoresis

50        CAN: acetonitrile

51        CBB: Coomassie Brilliant Blue

52        LC-MS/MS: liquid chromatography tandem mass spectrometry

53        MW: molecular weight

54        pI: isoelectric point

55        SDS-PAGE: sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

56

57

58

59

60

61

62

63

64



65 **1. INTRODUCTION**

66 Amaranth is an ancient crop belonging to the *Amaranthaceae* family with origin in  
67 the Americas. The genus *Amaranthus* L. contains more than 70 species of plants  
68 (Espitia-Rangel et al., 2012). Current interests in amaranth reside in the great  
69 amount of genetic diversity, phenotypic plasticity, and its extremely adaptability to  
70 adverse growing conditions (Rastogi and Shuckla, 2013). Resistance to several  
71 abiotic and biotic stresses such as pest, drought, heat, and salinity has been  
72 reported (Délano-Frier et al., 2011; Huerta-Ocampo et al., 2011). Besides  
73 amaranth cultivation has grown worldwide mainly due that the plant produces  
74 seeds with high protein content with an excellent essential amino acid profile  
75 (more lysine and methionine compared with other grains), and the unique starch  
76 and oil have composition (Rastogi and Shuckla, 2013). In addition, a renewed  
77 interest in amaranth consumption is due to the presence of encrypted peptides  
78 with several biological activities such as: antihypertensive, anticarcinogenic, and  
79 antidiabetic (Silva-Sánchez et al., 2008, Velarde-Salcedo et al., 2013), the seeds  
80 also contain several secondary metabolites (Barba de la Rosa et al., 2009) that are  
81 claimed as potent antioxidants. Amaranth has gained importance in the USA and  
82 is an upcoming product in Europe, especially as a substitute for wheat in diets of  
83 patients with coeliac disease (Ballabio et al., 2011). Therefore, amaranth is a  
84 promising crop to fulfill food requirements of the increasing global population  
85 mainly in developing countries (Rastogi and Shuckla, 2013).

86 Amaranth commercial species are represented by *A. hypochondriacus*, *A.*  
87 *cruentus*, and *A. caudatus*, the three species are native to America but are now

88 distributed throughout Asia and Africa. *A. cruentus* is taking great attention due to  
89 high tolerance to environmental stresses such as salinity (Espitia-Rangel et al.  
90 2012).

91 Nowadays, proteomics is widely used on the study of the seed proteome (Deng et  
92 al., 2013) focusing on identifying as many proteins as possible in order to create  
93 reference proteomic maps at specific developmental stages or to identify the  
94 changes due to environmental conditions (Miernyk and Hajduch, 2011). The  
95 increasing number of genomes and genes reported in databases has opened high  
96 possibilities for identifying proteins from non-sequenced ones such as amaranth.

97 In this work, the 2-DE map of *A. cruentus* L. cv Amaranteca was obtained and  
98 proteins spots were analyzed and identified by LC-MS/MS. An abundant protein  
99 spot was identified as LEA, and then the corresponding gene was cloned. The  
100 AcLEA has special features that classify it as a novel LEA belonging to the Group  
101 5. These results represent the first effort to characterize the amaranth seed protein  
102 profile using proteomics tools, generating new insights in the amaranth seed  
103 protein composition beyond the traditional analysis of SSPs.

104

## 105 **2. MATERIALS AND METHODS**

### 106 *2.1 Plant material*

107 Mature seeds from *Amaranthus cruentus* cv Amaranteca were obtained from  
108 INIFAP, Campo Experimental Bajío, Guanajuato, México. Amaranth seeds were  
109 ground into flour (80 mesh), the meal was defatted with hexane and defatted meal  
110 was kept in plastic bags at 4 °C until used.

111 *2.2 Total soluble protein extraction*

112 Amaranth flour was resuspended in the extraction buffer (100 mM Na<sub>2</sub>HPO<sub>4</sub>, 100  
113 mM NaCl, 2 mM 2-mercaptoethanol, and 2% Triton X-100 at pH 7.5), the  
114 suspension was stirred for 10 min and vortex for 1 min, then centrifuged for 30 min  
115 at 16200 x g at 4 °C. The supernatant containing the total soluble protein was  
116 precipitated using the methanol-chloroform method (Wesel and Flügge, 1984).  
117 Samples were re-suspended in rehydration solution (8 M urea, 2% CHAPS,  
118 0.002% bromophenol blue) and protein concentration was determined as  
119 mentioned above.

120

121 *2.3 2-DE seed protein analysis*

122 Protein samples (750 µg) were re-suspended in 250 µl of rehydration buffer and  
123 loaded onto a 13 cm, pH 3–10 linear immobilized pH gradient (IPG) strip (Bio-  
124 Rad). Passive rehydration was carried out at room temperature for 14–16 h.  
125 Isoelectric focusing was conducted at 20 °C using the Ettan IPGphor system (GE  
126 Healthcare, Piscataway, NJ, USA) at constant 50 mA per strip under the following  
127 conditions: (I) 500 V gradient until 0.01 kVh, (II) 4000 V gradient until 5.6 kVh, and  
128 (III) holding at 8000 V until 27.3 kVh. After isoelectric focusing, the IPG strips were  
129 stored at –20 °C or immediately equilibrated for 15 min in equilibration buffer (50  
130 mM Tris–HCl pH 8.8, 6 M Urea, 30% v/v glycerol, 2% w/v SDS, 0.002% w/v  
131 bromophenol blue) containing 1% w/v DTT. The second dimension was performed  
132 on 13% polyacrylamide gels using a SE 600 Ruby electrophoresis system (GE  
133 Healthcare). Gels were fixed for 1 h with a solution containing 50% v/v methanol

134 and 3% v/v phosphoric acid. Gels were stained with CBB G-250 (GE Healthcare)  
135 and images analysis was performed with PDQuest 2-D Analysis Software (Bio-  
136 Rad). Experimental molecular mass of each protein spot was estimated by  
137 comparison with molecular weight standards (BenchMark™ Protein Ladder,  
138 Invitrogen), while experimental *pI* was determined by migration of protein spots on  
139 the linear IPG strips.

140

#### 141 *2.4 In-gel digestion and LC-MS/MS analysis*

142 Selected protein spots were excised from the 2-DE gels and destained. After  
143 destaining the pieces of gel were reduced with 10 mM DTT in 25 mM ammonium  
144 bicarbonate followed by protein alkylation with 55 mM iodoacetamide. Protein  
145 digestion was carried out overnight at 37 °C with sequencing grade trypsin  
146 (Promega, Madison, WI, USA).

147 Mass spectrometric analysis was carried out with a SYNAPT-HDMS (Waters Corp,  
148 Milford, MA, USA) coupled online to a nano-ACQUITY-UPLC (Waters). The  
149 peptides were loaded onto a Symmetry C<sub>18</sub> pre-column (5 μm, 20 mm x 180 μm).  
150 Peptide separation was carried out on a BEH130 C<sub>18</sub> (1.7 μm, 10 cm x 100 μm)  
151 analytical column (Waters). Mass spectrometric analysis (LC-MS/MS) was carried  
152 out in a SYNAPT-HDMS (Waters). The spectrometer was operated in V-mode and  
153 analyses were performed in positive mode ESI. The TOF analyzer was externally  
154 calibrated with [Glu1]-Fibrinopeptide B from *m/z* 50 to 2422.

155

#### 156 *2.5 Database searching and protein identification*

157 The MS/MS spectra were searched against the *Viridiplantae* nrNCBI protein  
158 database and against a private database created using the data obtained from the  
159 first large-scale transcriptomic analysis of *Amaranthus hypochondriacus* (Délano-  
160 Frier et al., 2011) using the MASCOT search algorithm v 2.3.0 (Matrix Science,  
161 London, UK). Trypsin was used as the specific protease and one missed cleavage  
162 was allowed. The mass tolerance for precursor and fragment ions was set to 10  
163 ppm and 0.1 Da respectively. Carbamidomethyl cysteine was set as fixed  
164 modification and oxidation of methionine was specified as variable modification.  
165 Identifications were considered successful when significant MASCOT scores (>35  
166 for *Viridiplantae* nrNCBI or > 30 for *A. hypochondriacus* transcriptome database)  
167 were obtained, indicating the identity or extensive homology at  $P < 0.05$ .

168

## 169 *2.6 RNA extraction and first-strand cDNA synthesis*

170 Immature seeds from *Amaranthus cruentus* plants were harvested, immediately  
171 frozen under liquid nitrogen and stored at -70 °C. Immature seeds were  
172 homogenized with a mortar and pestle in the presence of liquid nitrogen. Seeds  
173 were transferred into Eppendorf tubes and total RNA was extracted using TRizol®  
174 Reagent (Invitrogen, Grand Island, NY, USA) according to manufacturer's  
175 instructions. The first-strand cDNA was synthesized from total RNA (200 ng) using  
176 the oligonucleotide CDS PRIME IIA: (5'-  
177 AAGCAGTGGTATCAACGCAGAGTACTTT TTT TTT TTT TTT TTT TTT TTT  
178 TTT TVN-3') and SMART II: (5'-AAGCAGTGGTATCAACGCAGAGTACGCGGG-  
179 3') from the SMART™ PCR cDNA Synthesis Kit (Clontech, Palo Alto, CA), and

180 purified through Chroma-Spin-1000 columns (Clontech) according to  
181 manufacturer's protocols.

182

### 183 *2.7 Isolation and sequencing of the full-length cDNA of Amaranthus cruentus LEA* 184 *(AcLEA)*

185 Primers were designed based on the matched LEA protein sequence obtained by  
186 LC-MS/MS analysis and database search against the *Amaranthus*  
187 *hypochondriacus* transcriptome. Primers were used to amplify the full-length open  
188 reading frame of the *AcLEA* gene using the first-strand cDNA (8 ng) as a template  
189 and the following primers: LEA-fdw: (5'-GGAAAGACAAGAGGAGCAACAGAAGC-  
190 3') and LEA-rev: (5'-GTGCTCGAGCTAGGGCCTAGTAGTCTTAATTGGAT-3').  
191 PCR was carried out with pre-denaturation at 94 °C for 2 min, 30 cycles of 30 s at  
192 94 °C, 30 s at 60 °C and 1 min at 72 °C, and a final elongation step of 10 min at 72  
193 °C. A PCR product of 0.45 kb was amplified and purified from agarose gel 1.5%  
194 with QIAquick Gel Extraction Kit (Qiagen GmbH, Hilden, Germany) and  
195 sequenced (Lanbama-IPICyT, Mexico). *Amaranthus cruentus* LEA (*AcLEA*) gene  
196 sequence has been deposited at GenBank.

197

## 198 **3. RESULTS AND DISCUSSION**

### 199 *3.1 2-DE protein profile of amaranth total soluble proteins*

200 The 2-DE protein profile of *A. cruentus* cv *Amaranteca* is shown in Figure 1.  
201 Approximately 400 CBB-stained protein spots were detected on the 2-DE gel, 106  
202 spots were analyzed by LC-MS/MS and 43 proteins were identified by database

203 search against the *A. hypochondriacus* EST private database (Délano-Frier et al.,  
204 2011) and the public *Viridiplantae* nrNCBI database (Table 1 and Supplementary  
205 Table S1). Proteins were classified into functional categories: stress and defense  
206 response, metabolic process, respiratory chain and oxidation-reduction process,  
207 nutrient reservoir and miscellaneous. In five cases it was not possible to assign  
208 any function (unknown function).

209

### 210 *3.1.2 Proteins involved in stress and defense responses*

211 Desiccation tolerance is one of the most fundamental properties of seeds, it is  
212 acquired late in seed development and is considered necessary for the completion  
213 of the plant's life cycle, as an adaptive strategy that allows seed survival during  
214 storage or against environmental stress, ensuring better dissemination of the  
215 species (Leprince et al., 1993). Stress-responsive proteins such as small heat  
216 shock proteins (HSPs), Late Embryogenesis Abundant proteins (LEAs), seed  
217 storage proteins, and dormancy related proteins are known to play protective roles  
218 under environmental stresses, and up-regulation of their expression is an  
219 important mechanism for stress tolerance to desiccation in seeds by preventing  
220 protein aggregation and preserving enzymatic activity during dehydration (Huang  
221 et al., 2012). We have found three protein spots (6, 52, 81) belonging to LEAs and  
222 two spots (53, 82) identified as HSPs. LEAs are very hydrophilic protein members  
223 of a class of highly conserved proteins which may play an essential role in seed  
224 survival and in controlling water exchange during seed desiccation and imbibition  
225 (Huang et al., 2012). Plant small HSPs (sHSPs) are a structurally diverse protein

226 family, and the members with molecular masses between 15 and 30 kDa are  
227 grouped into five major classes. HSPs and sHSPs act as molecular chaperones  
228 and are thought to repair the stress-damaged proteins (Huang et al., 2012).  
229 sHSPs of 14.4 kDa (spot 82) are unusually abundant diverse proteins involved in  
230 various abiotic stresses, in *Nelumbo nucidera* *NnHSP17.5* was specifically  
231 expressed in seeds under normal conditions, and was strongly up-regulated in  
232 germinating seeds upon heat and oxidative stresses (Zhou et al., 2012).

233 The alpha chain of the Nascent polypeptide Associated Complex (NAC) was found  
234 in spot 50. NAC is a heterodimeric complex of  $\alpha$  and  $\beta$  chains that can reversibly  
235 bind to eukaryotic ribosomes, being probably the first cytosolic protein to contact  
236 with nascent polypeptide chains emerging from ribosome. The function of NAC is  
237 still unknown, but has been suggested that could be involved in protein sorting and  
238 translocation, preventing mistargeting of nascent polypeptide chains to  
239 endoplasmic reticulum and that represents a novel target of salt tolerance in plants  
240 (Lu et al., 2012). The proteasome is a multicatalytic proteinase complex that is  
241 involved in ATP/ubiquitin-dependent proteolytic pathways and can degrade  
242 unneeded or damaged proteins to protect the seed from damage (Huang et al.,  
243 2012). In amaranth seed the spot 37 corresponded to the proteasome subunit beta  
244 type-6-like.

245 Dehydration-Responsive Element-Binding protein (DREB) also referred as CB (C-  
246 repeat binding factor) proteins was found in spot 48. DREBs belong to the  
247 AP2/ERF family of transcription factors which confer stress endurance, in plants  
248 play important roles in promoting oxidative stress tolerance (Akhatar et al., 2012).



249 Spot 67 was identified as one pathogenesis-related protein and spots 30 and 31  
250 as cystatins. Pathogenesis-resistant proteins are present in desiccation-tolerant  
251 embryos after maturation drying and desiccation treatment in maize (Huang et al.,  
252 2012). Phytocystatins are cysteine proteinase inhibitors from plants implicated in  
253 defense mechanisms against insects and plant pathogens. The cystatin cDNA  
254 from *Amaranthus hypochondriacus* has been cloned and characterized (Valdes-  
255 Rodriguez et al., 2010).

256 Seeds of many plant species enter a period of dormancy when they fail to  
257 germinate under favorable conditions. Spots 19 and 20 were identified as  
258 dormancy related putative protein. Seed dormancy is controlled by the  
259 physiological or structural properties of a seed and the external conditions. It is  
260 induced as part of the genetic program of seed development and maturation.  
261 Seeds with deep physiological embryo dormancy can be stimulated to germinate  
262 by a variety of treatment including cold stratification (Pawlowski et al., 2009).

263

### 264 *3.1.3 Proteins related with metabolic process*

265 Glycolysis is a central metabolic pathway that provides energy and generates  
266 precursors for the synthesis of primary metabolites such as amino acids and fatty  
267 acids. Reactive oxygen species (ROS) are produced in plants in response to a  
268 wide variety of stresses and are signals that alter various cellular and physiological  
269 processes in plant growth and development. Spot 10 was identified as  
270 glyceraldehyde-3-phosphate dehydrogenase (GAPC). Recently, attention has

271 taken to the importance of GAPC as crucial role in the control of primary  
272 metabolism in plants (Muñoz-Bertomeu et al., 2010).

273 Lipid A biosynthesis begins with acylation of UDP-GlcNAc, using (3R)-  
274 hydroxymyristoyl-[acyl-carrier-protein]-dehydratase as the acyl donor, this  
275 dehydratase was found in spot 27. The two substrates of UDP-GlcNAc 3-O-  
276 acyltransferase, UDP-GlcNAc and (3R)-hydroxymyristoyl-ACP, are situated at  
277 important biosynthetic branch points, leading to several major cell envelope  
278 components (Mohan et al., 1994).

279 Nucleoside diphosphate kinase (NDPK) was found in spots 55, 56, 83. NDPK is a  
280 ubiquitous enzyme that catalyzes the transfer of the  $\gamma$ -phosphoryl group from a  
281 nucleoside triphosphate to a nucleoside diphosphate, it has been found in  
282 immature seeds and is considered as a housekeeping enzyme that regulates  
283 nucleotide pools within the cells, in plants are strongly related to signal  
284 transduction, differentiation, and development and as possible component of the  
285 ethylene signal transduction chain (Kihayra et al., 2011).

286 Glutamate decarboxylase (GAD) was found in spots 17, 101, and 103. GAD is a  
287 pyridoxal 5'-phosphate-dependent enzyme that catalyzes the irreversible  $\alpha$ -  
288 decarboxylation of L-glutamate to  $\gamma$ -aminobutyric acid (GABA). Proper regulation  
289 of the GABA metabolic pathway in plants is important for stress tolerance (Zik et  
290 al., 2006).

291 Phytases belonging to the HAP (histidine acid phosphatases) have been described  
292 only in plants but PAP (purple acid phosphatases), metallohydrolases with phytase  
293 activity, appears to be restricted only to plants. Phytases are of particular

294 importance during seed germination, they mobilize phosphate from phytate, the  
295 major reserve of phosphorus in plant seeds. In *Arabidopsis*, overexpression of  
296 PAP2 encoded by *AtPAP2* promotes plant growth by modulating carbon  
297 metabolism (Zhang et al., 2012) while PAP3 may play role in stress tolerance by  
298 enhancing ROS scavenging (Li et al., 2008). Different isoforms of PAPS were  
299 found in spots 103, 104, and 106.

300

#### 301 *3.1.4 Respiratory chain and oxidation-reduction process*

302 Under stress the enzyme activities are affected and the basic metabolism is  
303 disturbed. In order to maintain homeostasis under stress conditions, plants need to  
304 fortify the resistance mechanisms, such as ion transport, reactive oxygen species  
305 (ROS) scavenging and osmolyte synthesis, processes that require an extra energy  
306 supply. Cyt<sub>c</sub> (spot 18) is a hydrophilic metalloprotein found in the plant's  
307 mitochondria and probably facilitates energy generation through the respiratory  
308 chain under stress conditions (Yan et al., 2005). Oxidative signaling by ROS has  
309 been demonstrated to play a role in seed dormancy alleviation, but the detailed  
310 molecular mechanism underlying this process remains largely unknown. Key redox  
311 compounds in the soluble phase of the cell are NAD<sup>+</sup>, NADP<sup>+</sup>, glutathione, and  
312 ascorbate, which form an integrated network and interact strongly with ROS. Spots  
313 21 and 41 were identified as NADH dehydrogenase and NADPH oxidase,  
314 respectively, while spot 38 was identified as glutathione S-transferase.

315

#### 316 *3.1.5 Nutrient reservoir activity*

317 11S proteins considered as one of the main fractions of amaranth SSPs were  
318 identified in spots 22, 42, and 88. SSPs constitute the major source of nitrogen for  
319 germinating seeds, and are synthesized and accumulated during seed  
320 development. SSPs mobilization is accompanied by the extensive reduction of  
321 disulfide bonds of cysteine-rich proteins, this regulatory sulphhydryl/disulphide bond  
322 system, alters the structure and activity of proteins controlling an array of events in  
323 seed germination such as mobilization of storage proteins, activation of proteases,  
324 transcription, cell division, radical scavenging, and detoxification (Shahmoradi et  
325 al, 2013).

326

### 327 *3.1.6 Miscellaneous*

328 Spot 25 was identified as uridylyl transferase (UTPase). The UTPase (uridine  
329 monophosphate kinase) is the precursor for the biosynthesis of all pyrimidine  
330 nucleoside triphosphates and has been described as one of the candidate genes  
331 that have significant association with QTLs of grain yield, flowering time and leaf  
332 rolling under drought stress conditions (Sehgal et al., 2012). The response of cells  
333 to changes in their environment often requires co-regulation of gene networks, but  
334 little is known about how this can occur at the post-transcriptional level. Under  
335 conditions of limited nutrients, eukaryotic cells reprogram protein expression in a  
336 way that slow growth but enhance survival (Ivanov et al., 2011). T-cell intracellular  
337 antigen (TIA)-1 and TIA-1-related protein (TIAR, spot 64) are mRNA-binding  
338 proteins that can aggregate within granules under specific stress conditions and  
339 are involved in multiple aspects of cellular metabolism, including mRNA splicing,

340 mRNA translation and decay, cell survival, and the stress response (Gotschald et  
341 al., 2010).

342

### 343 3.2 Characterization of the full-length sequence of AcLEA protein

344 From all proteins identified by 2-DE, LC-MS/MS, and database search approach,  
345 special attention was taken to LEA's protein. First, because one LEA (spot 52)  
346 was one of the abundant proteins in 2-DE maps of *A. cruentus* and second  
347 because until now, there are not reports of the presence of LEAs in amaranth  
348 seed. The first LEAs proteins were identified in mature cotton (*Gossypium*  
349 *hirsutum*) seeds, but LEA's genes are found throughout higher plants, including  
350 monocots, dicots, and gymnosperms, and differences in temporal expression  
351 pattern and sequence similarity among the encoded proteins have allowed them to  
352 be subdivided into groups. Computational analysis of *LEA* genes from different  
353 plant species demonstrated that their classification is in fact not so clear and that  
354 some groups might be considered to be related, comparison of specific motifs  
355 specific to a unique LEA indicates that proteins with classical group 5 motifs are  
356 related to group 3 proteins (Bies-Ethève et al., 2008).

357 Based on matched LEA sequences obtained by LC-MS/MS and database search  
358 against the *Amaranthus hypochondriacus* transcriptome (Table 1, spot 52), a pair  
359 of primers was designed for amplification of *Amaranthus cruentus* LEA (*AcLEA*)  
360 cDNA, the 450 bp amplified fragment is shown in Supplementary Figure S1. The  
361 full-length cDNA of *AcLEA* (Figure 2A) contained an ORF of 418 bp that encodes a  
362 protein of 139 amino acids with a predicted

363 ([http://www.expasy.org/tools/pi\\_tool.html](http://www.expasy.org/tools/pi_tool.html)) molecular mass of 14,433.8 kDa and a  
364 pI of 8.65 (Supplementary Table S2). AcLEA5 is rich in alanine (19.5%), glycine  
365 (7.48%), lysine (15.65%) and threonine (10.20%), while it lacks cysteine,  
366 tryptophan, is rich in acidic amino acids (D, 6.12% and E 9.52%) phenylalanine  
367 and proline are present in low amounts (1.36 and 2.04, respectively)  
368 (Supplementary Table S2). AcLEA sequence shows the repeating motif of 11  
369 amino acids TAQAAKEK (Figure 2A) that are characteristics of LEA Group 3 (Liu  
370 et al., 2013), however, phylogenetic tree clearly shows that AcLEA belongs to the  
371 LEAs Group 5 closely related to the LEA5s from *Solanum lycopersicum* and  
372 *Brassica napus* (Figure 2B, Supplementary Figure S2) and is far from OsLEA5  
373 from *Oryza sativa* which has been reported as hydrophobic atypical LEA and has  
374 been associated with resistance to multiple abiotic stresses (He et al., 2012). The  
375 LEA5 from maize, the *Rab28* gene has been reported as candidate to enhance  
376 stress tolerance in maize plants (Amara et al., 2013).

377 Due to the amino acid composition, AcLEA is highly hydrophilic (Supplementary  
378 Figure S3A). In agreement with other LEAs structures (Wise and Tunnacliffe,  
379 2004), AcLEA was predicted to have secondary structure  $\alpha$ -helices  
380 (Supplementary Figure S3B). Using the Phyre<sup>2</sup> Engine ([www.sbg.bio.ic.ac.uk](http://www.sbg.bio.ic.ac.uk),  
381 Kelley and Sternberg, 2009) it was possible to predict the 3-D folding pattern of  
382 AcLEA  $\alpha$ -helices (Figure 3). The LEA structure seems to be very flexible that will  
383 an advantage in protein adaptation in function of environmental conditions, during  
384 seed dehydration-hydration.

385 In conclusion, information about the amaranth seed protein composition should  
386 benefit the improvement of the quality of this grain. There is a general consensus  
387 that desiccation tolerance involves the cellular mechanisms for membranes  
388 protection from the deleterious effect of water removal and the resultant necessity  
389 to maintain the cellular structure in the absence of an aqueous environment. We  
390 have shown for the first time the presence of LEAs, Hsp, and DREBs proteins in  
391 amaranth seeds, as well the presence of purple acid phosphatases (PAPs) related  
392 with phosphorous mobilization in seeds. 11S proteins could have further functions  
393 in the regulatory of seed germination events regulating the redox conditions  
394 between an oxidized state (disulphide) and a reduced state (thiol), and *vice-versa*,  
395 working together with Cyt<sub>c</sub>, NADP<sup>+</sup> and NADPH<sup>+</sup> in protection against ROS.

396

### 397 **Acknowledgements**

398 We thank CONACYT-MEXICO GRANT 56787 (Laboratory for Nanoscience and  
399 Nanotechnology Research-LINAN) and Waters-México for their technical  
400 assistance. E. Maldonado-Cervantes thanks CONACyT-Mexico for the fellowship  
401 No. 233301. Thanks to J. Délano (Cinvestav-Irapuato), Antonio Castellanos (CNS-  
402 IPICyT) for informatics support and Carlos J. Salas for his technical assistance.

403

### 404 **References**

- 405 Akhatar, M., Jaiswal, A., Taj, G., Jaiswal, J.P., Qureshi, M.I., Singh, N.K., 2012.  
406 DREB1/CBF transcription factors: their structure, function and role in abiotic stress  
407 tolerance in plants. *Journal of Genetics* 91,385-395.
- 408 Amara, I., Capellades, M., Ludevid, M.D., Pages, M., Goday, A., 2013. Enhanced water  
409 stress tolerance of transgenic maize plants over-expressing *LEA Rab28* gene.  
410 *Journal of Plant Physiology* 170,864-873.
- 411 Ballabio, C., Uberti, F., Di Lorenzo, Ch., Brandolini, A., Pens, E., Restani, P., 2011.

412 Biochemical and immunochemical characterization of different varieties of amaranth  
413 (*Amaranthus* L. ssp.) as a safe ingredient for gluten-free products. *Journal of*  
414 *Agricultural and Food Chemistry* 59,12969-12974.

415 Barba de la Rosa, A.P., Fomsgaard, I.S., Laursen, B., Mortensen, A.G., Olvera-  
416 Martínez, J.L., Silva-Sánchez, C., Mendoza-Herrera, A., De León-Rodríguez A.,  
417 González-Castañeda, J. 2009. Amaranth (*Amaranthus hypochondriacus*) as an  
418 alternative crop for sustainable food production: Phenolic acids and flavonoids with  
419 potential impact on its nutraceutical quality. *Journal of Cereal Science* 49,117-121.

420 Bies-Ethève, N., Gaubier-Comella, P., Debures, A., Lasserre, E., Jobet, E., Reynal, M.,  
421 Cooke, R., Delseny, M., 2008. Inventory, evolution and expression profiling diversity  
422 of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis*  
423 *thaliana*. *Plant Molecular Biology* 67,107-124.

424 Délano-Frier, J.P., Avilés-Arnaut, H., Casarrubias-Castillo, K., Casique-Arroyo, G.,  
425 Castrillón-Arbaláez, P.A., Herrera-Estrella, L., Massange-Sánchez, J., Martínez-  
426 Gallardo, N.A., Parra-Cota, F.I., Vargas-Ortiz, E., Estrada-Hernández, M.G., 2011.  
427 Transcriptomic analysis of grain amaranth (*Amaranthus hypochondriacus*) using 454  
428 pyrosequencing: comparison with *A. tuberculatus*, expression profiling in stems and in  
429 response to biotic and abiotic stress. *BMC Genomics* 12,363.

430 Deng, Z.Y., Yan, C., Gong, C.Y., Want, T., 2013. Use of proteomics to understand seed  
431 development in rice. *Proteomics* 12-13:1784-1800.

432 Espitia-Rangel, E., Escobedo-López, D., Mapes-Sánchez, C., de la O-Olán, M., Aguilar-  
433 Delgado, M., Hernández-Casillas, J.M., Ayala-Garay, A.V., Rivas-Valencia, P.,  
434 Martínez-Trejo, G., Ramírez-Vázquez, M.L., Morán-Bañuelos, S.H., 2012.  
435 Conservación de los recursos genéticos de amaranto (*Amaranthus* spp.) en México.  
436 In: *Amaranto: Ciencia y Tecnología*, Eduardo Espitia-Rangel (Ed.). INIFAP/SINAREFI  
437 Chapter XI pp. 147-163.

438 Gotschald, O.R., Malec, V., Krasteva, G., Hsan, D., Kamlah, F., Herold, S., Rose, F.,  
439 Seeger, W., Hänze, Jörg., 2010. TIAR and TIA-1 mRNA-binding proteins co-  
440 aggregate under conditions of rapid oxygen decline and extreme hypoxia and  
441 suppress the HIF-1 $\alpha$  pathway. *Journal of Molecular Cell Biology* 2,345-356.

442 He, S., Tan, L., Hu, Z., Chen, G., Wang, G., Hu, T., 2012. Molecular characterization and  
443 functional analysis by heterologous expression in *E. coli* under diverse abiotic  
444 stresses for OsLEA5, the atypical hydrophobic LEA protein from *Oryza sativa* L.  
445 *Molecular Genetics and Genomics* 287,39-54.

446 Huang, H., Moller, I.M., Song, S-Q., 2012. Proteomics of desiccation tolerance during  
447 development and germination of maize embryos. *Journal of Proteomics* 75,1247-  
448 1262.

449 Huerta-Ocampo, J.A., León-Galván, M.F., Ortega-Cruz, L.B., Barrera-Pacheco, A., De  
450 León-Rodríguez, A., Mendoza-Hernández, G., Barba de la Rosa, A.P., 2011. Water  
451 stress induces the up-regulation of DOF1 and MIF1 transcription factors and down-  
452 regulation of proteins involved in secondary metabolism in amaranth roots  
453 (*Amaranthus hypochondriacus* L.). *Plant Biology* 13,472-482.

454 Ivanov, P., Kedersha, N., Andersen, P., 2011. Stress puts TIA on TOP. *Genes and*  
455 *Development* 25,2119-2124.



456 Kelley, L.A., Strnberg, M.J.E., 2009. Protein structure prediction on the web: A case of  
457 study using the Phyre server. *Nature Protocols* 4,363-371.

458 Kihayra, A., Saburi, W., Wakuta, S., Kim, M.H., Hamada, S., Ito, H., Imai, R., Matsui, H.,  
459 2011. Physiological and biochemical characterization of three nucleoside diphosphate  
460 kinase isosymes from rice (*Oryza sativa* L.). *Bioscience, Biotechnology, and*  
461 *Biochemistry* 74,1740-1745.

462 Leprince, O., Hendry, G.A.F., McKersie, B.D., 1993. The mechanisms of desiccation  
463 tolerance in developing seeds. *Seed Science Research* 3,231-246.

464 Li, W.Y., Shao, G., Lam, H.M., 2008. Ectopic expression of GmPAP3 alleviates oxidative  
465 damage caused by salinity and osmotic stresses. *New Phytologist* 178,80-91.

466 Liu, Y., Wang, L., Xing, X., Sun, L., Pan, J., Kong, X., Zhang, M., Li, D., 2013. ZmLEA3,  
467 a multifunctional group 3 LEA protein from maize (*Zea mays* L.), is involved in biotic  
468 and abiotic stresses. *Plant Cell Physiology* 54,944-959.

469 Lu, M., Ying, S., Zhang, D-F., Shi, Y-S., Song, Y-Ch., Wang, T-Y., Li, Y., 2012. A maize  
470 stress-responsive NAC transcription factor, ZmSNAC1, confers enhanced tolerance  
471 to dehydration in transgenic *Arabidopsis*. *Plant Cell Reports* 31,1701-1711.

472 Miernyk, J.A., Hajduch, M., 2011. Seed proteomics. *Journal of Proteomics* 74,389-400.

473 Mohan, S., Kelly T.M., Eveland, S.S., Raetz, C.R.H., Anderson, M.S., 1994. An  
474 *Escherichia coli* gene (*FabZ*) encoding (3R)-hydroxymyristoyl acyl carrier protein  
475 dehydrase. *The Journal of Biological Chemistry* 269,32896-32903.

476 Muñoz-Bertomeu, J., Cascales-Miñana, B., Alaiz, M., Segura, J., Ros, R., 2010. A critical  
477 role of plastidial glycolytic glyceraldehyde-3-phosphate dehydrogenase in the control  
478 of plant metabolism and development. *Plant Signaling and Behavior* 5:67-69.

479 Pawlowski, T.A., 2009. Proteome analysis of Norway maple (*Acer platanoides* L.) seeds  
480 dormancy breaking and germination: influence of abscisic and gibberelic acids. *BMC*  
481 *Plant Biology* 9:48.

482 Rastogi, A., Shuckla, S., 2013. Amaranth: a new millennium crop of nutraceutical values.  
483 *Critical Reviews in Food Science and Nutrition* 63,109-125.

484 Sehgal, D., Rajaram, V., Armstead, I.P., Vadez, V., Yadav, Y.P., Hash, C.T., Yadav, R.  
485 S., 2012. Integration of gene-based markers in a pearl millet genetic map for  
486 identification of candidate genes underlying drought tolerance quantitative trait loci.  
487 *BMC Plant Biology* 12:9.

488 Shahmoradi, Z., Tamaskani, F., Sadeghipour, H.R., Abdolzadeh, A., 2013. Redox  
489 changes accompanying storage protein mobilization in moist chilled and warm  
490 incubated walnut kernels prior to germination. *Journal of Plant Physiology* 170,6-17.

491 Silva-Sánchez, C., Barba de la Rosa, A.P., León-Galván, M.F., de Lumen B.O., De  
492 León-Rodríguez, A., González de Mejía, E., 2008. Bioactive peptides in amaranth  
493 (*Amaranthus hypochondriacus*) seed. *Journal of Agricultural and Food Chemistry*  
494 56,1233-1240.

495 Valdes-Rodriguez, S., Cedro-Tanda, A., Aguilar-Hernandez, V., Cortes-Onofre, E.,  
496 Blanco-Labra, A., Guerrero-Rangel, A., 2010. Recombinant amaranth cystatin  
497 (AhCPI) inhibits the growth of phytopathogenic fungi. *Plant Physiology and*  
498 *Biochemistry* 48,469-475.

499 Velarde-Salcedo, A.J., Barrera-Pacheco, A., Lara-González, S., Montero-Morán, G.M.,

500 Díaz-Gois, A., González de Mejía, E., Barba de la Rosa, A.P. 2013. *In vitro* inhibition  
501 of dipeptidyl peptidase IV by peptides derived from the hydrolysis of amaranth  
502 (*Amaranthus hypochondriacus* L.) proteins. Food Chemistry 12,3204-3207.  
503 Wessel, D., Flügge, U.I., 1984. A method for the quantitative recovery of protein in dilute  
504 solution in the presence of detergents and lipids. Analytical Biochemistry 138:141-  
505 143.  
506 Wise, M.J., Tunnacliffe, A., 2004. POPP the question: what do LEA proteins do?. Trends  
507 Plant Science 9,13-17.  
508 Yan, S., Tang, Z., Su, W., Sun, W., 2005. Proteomic analysis of salt stress-responsive  
509 proteins in rice root. Proteomics 5,235-244.  
510 Zhang, Y., Yu, L., Yung, K-F., Leung, D.Y.C., Sun, F., Lim, B.L., 2012. Over-expression  
511 of AtPAP2 in *Camelina sativa* leads to faster plant growth and higher seed yield.  
512 Biotechnology for Biofuels 5,19.  
513 Zhou, Y., Chen, H., Chu, P., Li, Y., Tan, B., Ding, Y., Tsang, E.W.T., Jiang, L., Wu, K.,  
514 Huang, S., 2012. *NnHSP17.5*, a cytosolic class II small heat shock protein gene from  
515 *Nelumbo nucifera*, contributes to seed germination vigor and seedling  
516 thermotolerance in transgenic *Arabidopsis*. Plant Cell Reports 31,379-389.  
517 Zik, M., Fridmann-Sirkis, Y., Fromm, H., 2006. C-terminal residues of plant glutamate  
518 decarboxylase are required for oligomerization of a high-molecular weight complex  
519 and for activation by calcium/calmodulin. *Biochimica et Biophysica Acta (BBA)-*  
520 *Proteins and Proteomics* 1764,872-876.  
521

522

## 523 **Figure captions**

524

525 **Figure 1.** CBB-stained 2-DE gel of total *Amaranthus cruentus* cv *Amaranteca*  
526 seed proteins. Total soluble proteins (750 µg) were loaded on 13 cm IPG strips  
527 (pH 3-10 linear gradient) and separated using isoelectric focusing (IEF) in the first  
528 dimension, followed by 13% SDS-PAGE gels in the second dimension. Molecular  
529 mass in kDa and *pI* of proteins are indicated on the left and top, respectively.  
530 Spots numbers indicate the protein spots analyzed by LC-MS/MS and shown in  
531 Table 2.

532

533 **Figure 2.** Nucleotide and deduced amino acid sequence of cDNA encoding for  
534 *AcLEA* (submitted to GenBank) isolated from immature seeds of *Amaranth*  
535 *cruentus* cv *Amaranteca*. The putative conserved segments for LEA3 proteins are  
536 indicated with the line. B) Phylogenetic tree describing the relationship among LEA

537 proteins from different plants species constructed with ClustalW (ebi.ac.uk).  
538 Accessions numbers of published sequences in the GenBank are as follows:  
539 *Amaranthus cruentus*, AcLEA (submitted to GenBank); *Zea mays*, ZmLEA3  
540 (NP\_001146945); *Oryza sativa*, LEA5 (AEH76905); *Gossypium hirsutum*, D-29  
541 (CAA33194) and D7 (PI3939); *Glycine max*, LEA5 (NP\_001235600.1); *Vitis*  
542 *amurensis*, LEA5 (ADY17817.1); *Ricinus communis* LEA (XP\_002516942.1);  
543 *Medicago truncatula* LEA5 (XP\_003609687.1); *Brassica napus* LEA3  
544 (CAD59382.1); *Catharanthus roseus* LEA (AAY84145.1); *Solanum lycopersicum*  
545 LEA (NP\_001238798.1); *Triticum aestivum*, ABA inducible protein (BAF79927.1);  
546 *Arabidopsis thaliana* LEA3, (NP\_001190082.1).

547  
548 **Figure 3.** Tertiary structure prediction of AcLEA using the Phyre<sup>2</sup> Engine V2.0  
549 ([www.sbg.bio.ic.ac.uk](http://www.sbg.bio.ic.ac.uk)). Images are coloured by rainbow N- C terminus, the  
550 structures models have confidence and dimensions (Å): A) 97.6%, **X**:21.368,  
551 **Y**:79.450, **Z**:40.727, and B) 96%, **X**:30.261, **Y**:57.231, **Z**:26.133.

552  
553  
554  
555

**Table 1.** Identification by LC-MS/MS and database search of seed proteins in mature grain of *A. cruentus* L. cv Amaranteca.

Spot a	Protein	Plant species <sup>b</sup>	Accession Number <sup>c</sup>	Exper. Mr/pI <sup>d</sup>	Theor. Mr/pI <sup>e</sup>	Score <sup>f</sup>	Pm <sup>g</sup> / Sc <sup>h</sup>
<b>Stress and defense response</b>							
6	Late embryogenesis abundant protein group 3	<i>Brassica napus</i>	<a href="#">gij213053669</a>	53.3/6.3	43.5/5.6	51	4/-
19	Dormancy related protein, putative	<i>Arabidopsis thaliana</i>	<a href="#">gij12322163</a>	35.5/6.6	31.4/5.92	79	1/5
20	Dormancy related protein, putative	<i>Arabidopsis thaliana</i>	<a href="#">gij12322163</a>	35.1/6.3	31.4/5.92	56	3/11
30	Cystatin	<i>Amaranthus hypochondriacus</i>	<a href="#">gij110748606</a>	29.1/6.21	27.8/5.67	214	7/-
31	Cystatin	<i>Amaranthus hypochondriacus</i>	<a href="#">gij110748606</a>	29.6/6.02	27.8/5.67	144	4/14
37	PREDICTED: Proteasome subunit beta type-6-like	<i>Glycine max</i>	<a href="#">gij351721274</a>	27/5.2	25.2/5.3	84	4/-
48	Dehydration-responsive element-binding protein 1F-like	<i>Vitis vinifera</i>	<a href="#">gij225440970</a>	25.5/3.9	25.9/9.0	40	2/10
50	Alpha chain of nascent polypeptide associated complex	<i>Theobroma cacao</i>	<a href="#">gij508727650</a>	29.8/4.1	22.0/4.3	157	2/-
52	Late embryogenesis abundant protein group 5	<i>Vitis amurensis</i>	<a href="#">gij324103765</a>	19/8.0	16.4/6.2	117	5/-
53	18.3 kDa class I heat shock protein	<i>Chenopodium rubrum</i>	<a href="#">gij462322</a>	19/5.74	18.3/6.76	34	2/-
57	Dehydroascorbate reductase	<i>Nicotiana tabacum</i>	<a href="#">gij28192427</a>	11.6/7.46	23.6/7.7	40	1/4
67	Pathogenesis-related protein	<i>Spinacia oleracea</i>	<a href="#">gij444792485</a>	17.4/6.9	17.9/4.7	94	2/-
71	18.5 kDa class I heat shock protein	<i>Citrullus lanatus</i>	<a href="#">gij462322</a>	18./6.53	18.2/5.88	38	4/16
81	Late embryogenesis abundant protein group 5	<i>Vitis amurensis</i>	<a href="#">gij324103765</a>	19/6.47	16.4/6.2	36	2/-
82	Small heat shock protein 17.4BCII	<i>Boea hygrometrica</i>	<a href="#">gij496684331</a>	18.2/6.3	17.5/6.3	77	2/19
84	EM protein	<i>Vigna radiate</i>	<a href="#">gij1141782</a>	16.6/6.2	12.2/8.9	55	2/18
<b>Metabolic process</b>							
10	Glyceraldehyde-3-phosphate dehydrogenase	<i>Salvia elegans</i>	<a href="#">gij343131368</a>	45.4/6.9	21.4/6.2	52	4/16
17	Glutamate decarboxylase 5	<i>Arabidopsis thaliana</i>	<a href="#">gij115229517</a>	38.4/6.07	56.2/5.42	41	1/1
24	Hypothetical protein	<i>Sorghum bicolor</i>	<a href="#">gij242075416</a>	32.6/6.1	49/7.1	38	2/10

27	SORBDRAFT_06g011080 3R-hydroxymyristoyl-[ACP] dehydratase-like	<i>Glycine max</i>	<a href="#">gij2555640918</a>	30.7/7.4	19.9/9.7	42	2/37
55	Nucleoside diphosphate kinase 1	<i>Mesembryanthemum crystallinum</i>	<a href="#">gij6225750</a>	18.0/6.07	16.3/6.3	75	2/20
56	Nucleoside diphosphate kinase 1	<i>Spinacia oleracea</i>	<a href="#">gij400404</a>	16/6.1	16.3/6.4	103	2/-
83	Nucleoside diphosphate kinase 1	<i>Mesembryanthemum crystallinum</i>	<a href="#">gij6225750</a>	16/6.12	16.3/6.3	103	3/-
101	Glutamate decarboxylase 5	<i>Arabidopsis thaliana</i>	<a href="#">gij15229517</a>	41.9/7.13	56.2/5.42	42	1/-
103	Glutamate decarboxylase 5	<i>Arabidopsis thaliana</i>	<a href="#">gij15229517</a>	41.3/6.78	56.2/5.42	42	1/-
104	Probable inactive purple acid phosphatase 29-like	<i>Glycine max</i>	<a href="#">gij356573279</a>	52.5/5.2	42.8/8.1	130	6/-
106	Probable inactive purple acid phosphatase 29-like	<i>Glycine max</i>	<a href="#">gij356573279</a>	52.3/5.0	42.8/8.1	78	3/-

#### Respiratory chain and oxidation-reduction process

18	Ubiquinol-cytochrome c reductase	<i>Plantago major</i>	<a href="#">gij106879619</a>	38.4/5.9	8.4/9.5	39	2/68
21	NADH dehydrogenase	<i>Picea sitchensis</i>	<a href="#">gij116792579</a>	35.4/6.1	15.0/9.3	39	2/15
41	NADPH oxidase	<i>Triticum aestivum</i>	<a href="#">gij51860180</a>	24.3/7.0	30.0/8.6	43	2/27
38	Glutathione S-transferase-like protein	<i>Dianthus caryophyllus</i>	<a href="#">gij390979559</a>	27.3/7.5	24.4/6.1	118	3/-

#### Nutrient reservoir

22	11S globulin seed storage protein	<i>Amaranthus hypochondriacus</i>	<a href="#">gij558149</a>	35./5.92	57.0/6.69	49	2/3
42	11S globulin seed storage protein	<i>Amaranthus hypochondriacus</i>	<a href="#">gij558149</a>	21.3/6.7	57.0/6.69	127	4/9
88	11S globulin seed storage protein	<i>Amaranthus hypochondriacus</i>	<a href="#">gij558149</a>	11.9/4.8	57.0/6.69	44	2/4

#### Miscellaneous

25	Uridylate kinase	<i>Ricinus communis</i>	<a href="#">gij255571119</a>	32.5/5.9	23.5/5.2	32	2/-
33	Hypothetical protein PRUPE_ppa024816mg	<i>Prunus persica</i>	<a href="#">gij462409013</a>	33.6/5.06	37.3/6.84	41	1/3
64	Nucleolysin TIAR	<i>Vitis vinifera</i>	<a href="#">gij296085190</a>	54/8.64	16.6/9.43	36	2/27
76	40S ribosomal protein S21-2-like	<i>Solanum</i>	<a href="#">gij116778639</a>	12.8/6.7	9.1/7.8	41	4/-

isoform 2 lycopersicum

**Unknown function**

2	Conserved hypothetical protein [DUF1442]	<i>Ricinus communis</i>	gi 2555537091	79.4/7.06	23.5/8.9	40	1/-
29	Predicted uncharacterized protein LOC101246315	<i>Solanum lycopersicum</i>	gi 460376573	31.9/6.7	27.3/6.4	83	6/47
40	Hypothetical protein MTR_5g033270	<i>Medicago truncatula</i>	<a href="#">gi 357485713</a>	24.4/7.1	10.98/5.9	38	2/68
45	Predicted protein	<i>Populus trichocarpa</i>	<a href="#">gi 224131924</a>	21.2/5.7	22.3/6.2	39	2/40
100	Uncharacterized protein	<i>Lotus japonicas</i>	gi 388511639	10.3/8.33	9.3/8.08	57	1/8

<sup>a</sup> Spot numbers as indicated in Figure 5.

<sup>b</sup> Plant species represent the most likely orthologous organisms.

<sup>c</sup> Accessions numbers according to NCBI. When the best match was against the private *A. hypochondriacus* EST database, the most likely orthologous obtained after Translated BLAST (BLASTX) are reported.

<sup>d</sup> Experimental mass (kDa) and pI of identified protein spots.

<sup>e</sup> Theoretical mass (kDa) and pI of identified proteins retrieved from nrNCBI database or after calculation using the Compute pI/Mw tool (Available at [http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)).

<sup>f</sup> Mascot score reported after database search. Individual ion scores > 30 (for *Amaranthus hypochondriacus* EST), > 35 (for *Viridiplantae* nrNCBI) are statistically significant (P<0.05).

<sup>g</sup> Number of peptides matched

<sup>h</sup> Sequence coverage for the matched proteins, when the best matches were against the private *A. hypochondriacus* EST database the sequence coverage is not shown.

558

559

560

561

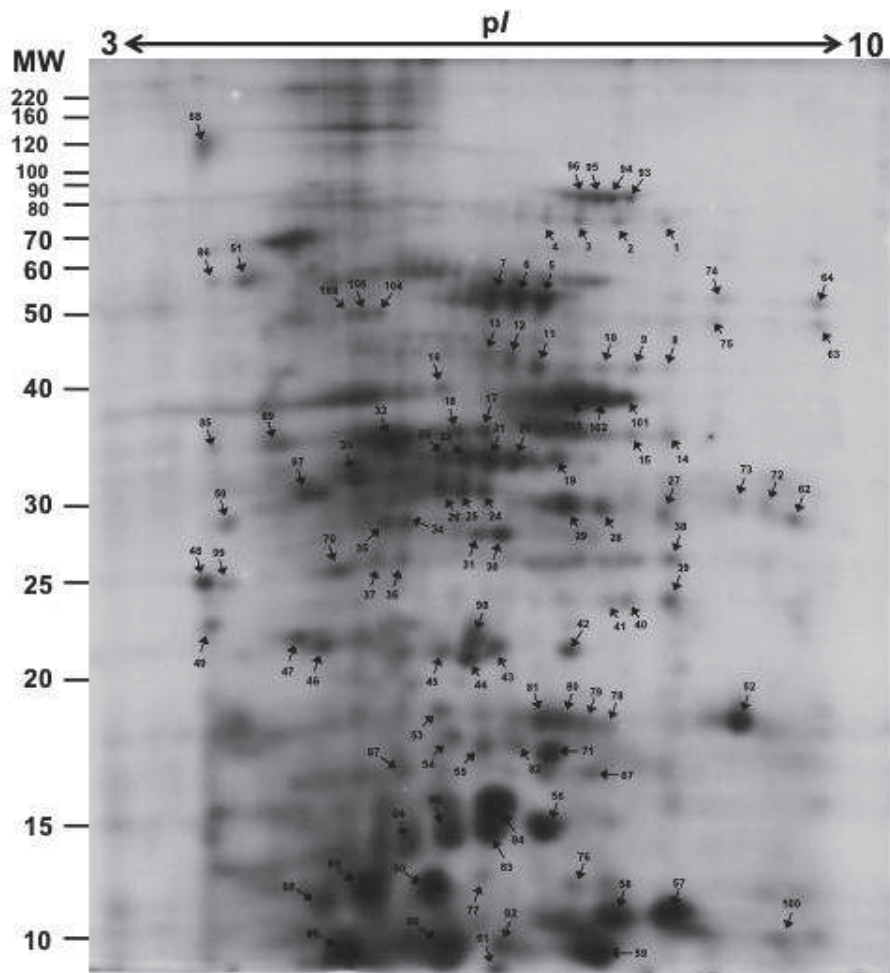
562

563

Figure 1

564

565



566

567

568

569

570

571

572

573

Figure 2

A

```

      10      20      30      40      50
GGAAAGACAA GAGGAGCAAC AGAAGCAATG AAGAACAAGG CCAACAACAGC
G K T R G A T E A M K N K A Q Q A

      60      70      80      90     100
AAAAGACAAG ACTATGGAAA CAGCCCAAGC AGCAAAAGAA AAGACTTCTG
K D K T M E T A Q A A K E K T S
      -----

      110     120     130     140     150
AGAAAGCCTA CGAGGCTAAA GACAAAACAT CTAACATTGC CCAATCTGCT
E K A Y E A K D K T S N I A Q S A

      160     170     180     190     200
AAGGGTCAGG CCAACAGTC CAAGGATCAA ACTGGATCCT TCGCTTCTGA
K G Q A Q Q S K D Q T G S F A S D

      210     220     230     240     250
CAAGGCCCAA GCAGCCAAGG ACAAGGCTTA TCAAGCAGCC CAAGCCACCA
K A Q A A K D K A Y Q A A Q A T

      260     270     280     290     300
AGGAGAAGGC GTCCGACATG ACGGAGTACG CTAAGGAGAC GCGGGAAGCC
K E K A S D M T E Y A K E T A E A

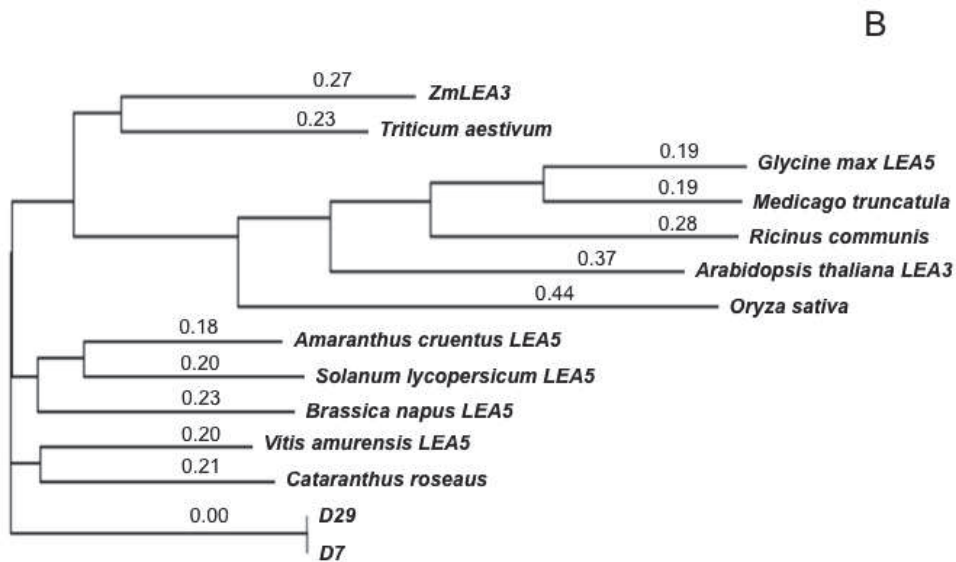
      310     320     330     340     350
GGAAAGGAGA AACTGGAGG AATTCTTCAA AGGACAGGTG AACAAGTCAA
G K E K T G G I L Q R T G E Q V K

      360     370     380     390     400
GAGTATGGCT CAAGGTGCCG CTGATGCTGT TAAGAATACT TTTGGGATGG
S M A Q G A A D A V K N T F G M

      410
GCGAACCTGA AGAAGATG
G E P E E D

```

574



575

576

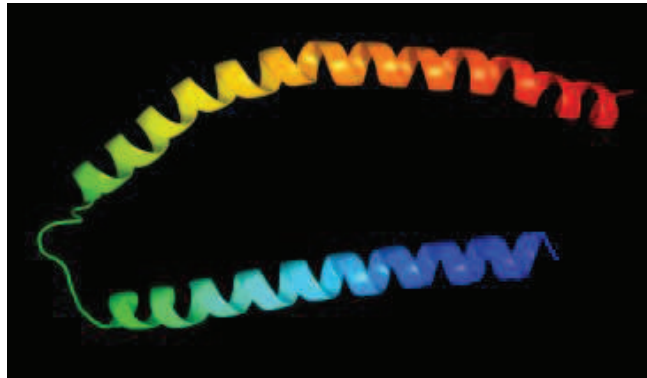
577



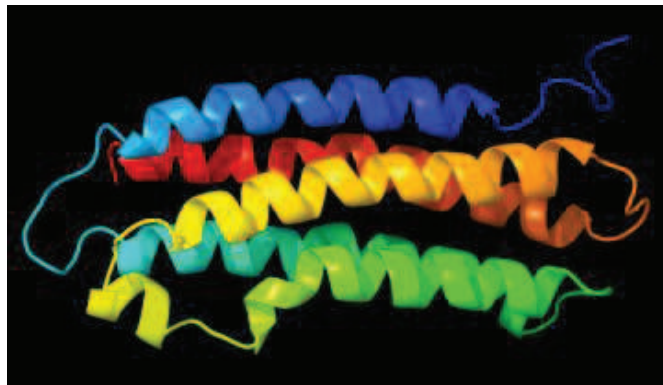
578

Figure 3

A



B



579

580

581

582

583

**Supplementary material for online publication only**

[Click here to download Supplementary material for online publication only: TableS1\\_JSC.docx](#)

**Supplementary material for online publication only**

[Click here to download Supplementary material for online publication only: Supplementary\\_JSC.docx](#)