



Departamento de Genética

Aplicación de metodologías proteómicas al estudio de la estructura, modificaciones postraduccionales y diferenciación genética de la proteína faseolina en semilla de la judía común (*Phaseolus vulgaris* L.)

Memoria presentada por:

María Dolores López Pedrouso

para optar al Grado de Doctora en Biología

Santiago de Compostela, 2015





Departamento de Genética

José Carlos Zapata Babío, Catedrático del Departamento de Genética de la Universidad de Santiago de Compostela

Informa:

Que el trabajo titulado “**Aplicación de metodologías proteómicas al estudio de la estructura, modificaciones postraduccionales y diferenciación genética de la proteína faseolina en semilla de la judía común (*Phaseolus vulgaris* L.)**” que presenta **María Dolores López Pedrouso** para la obtención del Grado de Doctora ha sido realizado bajo su dirección, que lo consideran concluido y autorizan su presentación al tribunal calificador.

Y para que así conste, se firma el presente informe en Santiago de Compostela a 7 Septiembre 2015.

Fdo.: José Carlos Zapata Babío



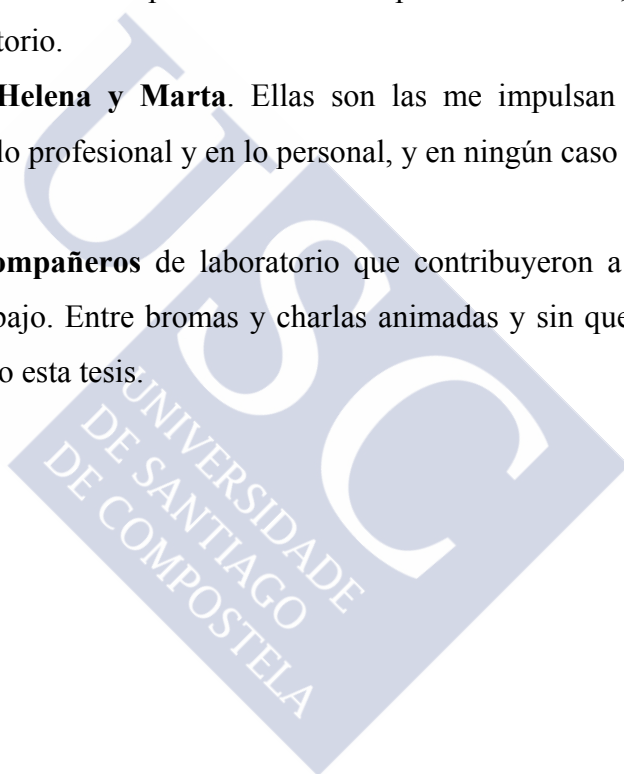
Agradecimientos

A mi director de tesis, **Carlos Zapata**, pues fue el principal motor de este trabajo de investigación. Una tesis doctoral es siempre fruto de ideas, proyectos y esfuerzos previos que corresponden a un investigador y que disfruta compartiéndolas y desarrollándolas con el doctorando. En mi caso, fue un placer poder compartir todo este tiempo y dedicación de mi director de tesis.

A **Daniel**, porque sin su motivación habría sido imposible. Su apoyo aunque de forma indirecta fue decisivo para salvar todo tipo de obstáculos, tanto dentro como fuera del laboratorio.

A mis hijas, **Helena y Marta**. Ellas son las que me impulsan a luchar por conseguir mejoras en lo profesional y en lo personal, y en ningún caso constituyeron un impedimento.

Al resto de **compañeros** de laboratorio que contribuyeron a mantener un buen ambiente de trabajo. Entre bromas y charlas animadas y sin que yo me diese cuenta, se fue haciendo esta tesis.





| | |
|--------------------------|--|
| 1-DE: | Electroforesis monodimensional |
| 2-DE: | Electroforesis bidimensional |
| AC: | Andina y cultivada |
| AW: | Andina y silvestre |
| Cv.: | Variedad cultivada |
| kDa: | kDalton (unidad de masa atómica) |
| HF: | Ácido fluorhídrico |
| HF-P: | Ácido fluorhídrico-Piridina |
| IEF: | Isoelectroenfoque o enfoque isoeléctrico |
| IMAC: | Cromatografía de afinidad de iones metálicos inmovilizados |
| IPG: | Gradiente inmovilizado de pH |
| m/z: | Relación masa/carga |
| MALDI: | Ionización/desorción por láser asistida por matriz |
| MC: | Mesoamericana y cultivada |
| MDS: | Escalado multidimensional no paramétrico |
| M_r: | Masa molecular relativa |
| MS: | Espectrometría de masas |
| MS/MS: | Espectrometría de masas en tándem |
| MW: | Mesoamericana y silvestre |
| NCBI: | Centro Nacional para la Biotecnología y la Información |
| PAGE: | Electroforesis en gel de poliacrilamida |
| PCA: | Análisis de componentes principales |
| pI: | Punto isoeléctrico |
| PMF: | Huella de la masa peptídica |
| PNGasa F: | N-Glicosidasa F |
| PR: | Tasa de fosforilación |
| Pro-Q DPS: | Tinción de fosfoproteínas Pro-Q Diamond (<i>Pro-Q Diamond phosphoprotein stain</i>) |
| RE: | Retículo endoplasmático |
| SDS-PAGE: | Electroforesis en gel de poliacrilamida con dodecilsulfato sódico (<i>sodium dodecyl sulfate polyacrylamide gel electrophoresis</i>) |

Abreviaturas

| | |
|-------------|--|
| SNP: | Polimorfismo de un solo nucleótido (<i>single nucleotide polymorphism</i>) |
| SRP: | Partícula reconocimiento de señal (<i>signal recognition particle</i>) |
| SSP: | Proteínas de reserva de la semilla(<i>seed storage proteins</i>) |
| TCA: | Tricloroacético |
| TOF: | Tiempo de vuelo (<i>time of flight</i>) |



INDICE

| | |
|--|-----------|
| RESUMEN | 11 |
| SUMMARY | 15 |
| 1. INTRODUCCIÓN | 19 |
| 1.1. Importancia económica y nutricional de la judía común | 21 |
| 1.2. Composición química y valor nutricional de la semilla de la judía común | 24 |
| 1.3. Proteómica de <i>Phaseolus vulgaris</i> L..... | 26 |
| 1.4. Faseolina | 29 |
| 1.4.1. Estructura y base genética | 29 |
| 1.4.2. Biosíntesis de la faseolina | 30 |
| 1.5. Fosfoproteoma | 33 |
| 1.6. Germinación de la semilla..... | 34 |
| 1.7. Origen y domesticación de la judía común..... | 36 |
| 2. OBJETIVOS | 41 |
| 3. RESULTADOS Y DISCUSIÓN | 45 |
| 3.1. Análisis proteómico de las semillas de judía común (<i>P. vulgaris</i> L.) basado en la electroforesis bidimensional (2-DE) | 47 |
| 3.1.1. Optimización del método de extracción proteica y de las condiciones electroforéticas para el proteoma de <i>P. vulgaris</i> L..... | 47 |
| 3.1.2. Artículo original: 2-DE-based proteomic analysis of common bean (<i>Phaseolus vulgaris</i> L.) seeds..... | 57 |
| 3.2. Caracterización de la diversidad de la proteína faseolina de la judía común (<i>P. vulgaris</i> L.) basada en la electroforesis bidimensional y en la espectrometría de masas..... | 61 |
| 3.2.1. Optimización de los perfiles 2-DE de la faseolina | 61 |
| 3.2.2. Descripción de los patrones 2-DE de la faseolina | 65 |
| 3.2.3. Diversidad de los patrones 2-DE de la faseolina..... | 67 |

| | |
|--|------------|
| 3.2.4. Interpretación molecular de los patrones 2-DE de la faseolina..... | 71 |
| 3.2.5. Artículo original: In-depth characterization of the phaseolin protein diversity of common bean (<i>Phaseolus vulgaris</i> L.) based on two-dimensional electrophoresis and mass spectrometry..... | 75 |
| 3.3. Evidencia de fosforilación de la principal proteína de reserva en la judía común y su degradación dependiente de la fosforilación durante la germinación..... | 79 |
| 3.2.1. Polipéptidos fosforilados de faseolina en semillas durmientes | 79 |
| 3.2.2. Cambios de fosforilación de la faseolina en semillas en germinación | 83 |
| 3.2.3. Identificación de los fosfopolipéptidos de faseolina mediante MS..... | 87 |
| 3.2.4. Artículo original: Evidence for phosphorylation of the major seed storage protein of the common bean and its phosphorylation-dependent degradation during germination | 93 |
| 3.4. Evaluación de los perfiles 2-DE de la faseolina como marcadores de diferenciación genética y calidad proteica en la judía común (<i>P. vulgaris</i> L.)..... | 97 |
| 3.4.1. Diversidad de patrones 2-DE de la faseolina..... | 97 |
| 3.4.2. Distancia Proteómica..... | 100 |
| 3.4.3. Distancia Proteómica y calidad de las proteínas de las semillas. | 104 |
| 3.4.4. Artículo original: Evaluating two-dimensional electrophoresis profiles of the protein phaseolin as markers of genetic differentiation and seed protein quality in common bean (<i>Phaseolus vulgaris</i> L.) | 109 |
| 4. CONCLUSIONES | 113 |
| 5. BIBLIOGRAFÍA | 117 |



RESUMEN



La judía común (*Phaseolus vulgaris* L.) es la legumbre de grano más importante para el consumo humano. Las proteínas de reserva de la semilla son la principal fuente de aminoácidos y de nitrógeno reducido que desempeñan un papel clave durante el proceso de la germinación y el crecimiento temprano de la plantas. La faseolina es la principal proteína de reserva de la semilla de la judía común, representando hasta un 40-60% de la proteína total. Décadas de estudios han permitido obtener un conocimiento cada vez más profundo de la base genética, expresión, regulación, estructura, diferenciación genética entre variedades y evolución de la faseolina. Por tanto, se puede considerar que la faseolina representa la proteína de reserva de semilla modelo. La faseolina está codificada por una familia génica localizada en el cromosoma 7. Comprende las subfamilias génicas α y β que dan lugar a la síntesis de polipéptidos glicosilados del tipo α y β , respectivamente. El estudio de la diversidad de la faseolina en poblaciones silvestres y cultivadas, mediante el uso de la electroforesis monodimensional, ha permitido la identificación de dos acervos génicos principales de *P. vulgaris* (mesoamericano y andino) que fue posteriormente confirmada con una diversidad de marcadores de ADN. Sin embargo, los avances recientes en las metodologías proteómicas proporcionan nuevas herramientas para profundizar en el conocimiento de esta compleja familia génica. En este estudio se abordan diversos aspectos sobre la estructura, expresión y diferenciación genética de la faseolina aplicando la electroforesis bidimensional (2-DE) y la espectrometría de masas (MALDI-TOF y MALDI-TOF/TOF MS) en conjunción con otras metodologías proteómicas. La optimización del método de extracción proteica y de los protocolos de la 2-DE ha permitido obtener mapas de alta resolución de la proteína faseolina que suministraron información nueva de gran relevancia. Los resultados del estudio proporcionan la primera evidencia sobre la existencia de isoformas de la faseolina diferencialmente fosforiladas. Además, isoformas con elevados niveles de fosforilación desempeñan un papel clave en la movilización de la proteína faseolina dado que experimentan un proceso de degradación preferente durante la germinación de la semilla. Cabe subrayar que este sería el primer papel funcional asignado a las isoformas fosforiladas detectadas hasta la fecha en las proteínas de reserva de semilla. Por otra parte, el análisis de las distancias proteómicas obtenidas a partir de los patrones 2-DE muestra que la faseolina es un marcador fiable de diferenciación genética y calidad de la semilla en variedades silvestres y cultivadas de los acervos mesoamericano y andino. Por tanto, la faseolina parece ser un marcador de utilidad para la mejora genética de la judía común. Globalmente, el estudio revela que las metodologías proteómicas son una herramienta importante para avanzar en el conocimiento de las complejas proteínas de reserva de semilla.





SUMMARY



Common bean (*Phaseolus vulgaris* L.) is the most important grain legume for human consumption. The seed storage proteins are the main source of amino acids and reduced nitrogen that play a fundamental role during the germination and the early plant growth. Phaseolin is the major seed storage protein of common bean accounting for 40-60% of the total seed protein. In the past decades, numerous studies have contributed to the enormous progress in our knowledge about the structure, genetic basis, expression, regulation and evolution of the phaseolin. Therefore, the phaseolin can be considered as the model protein to the study of seed storage proteins. Phaseolin is encoded by a gene family located on chromosome 7. Members of the gene family can be subdivided into two distinct gene types termed α and β , that give rise to α - and β - glycosylated polypeptides. The study of the diversity of phaseolin in wild and domesticated populations of common bean through monodimensional electrophoresis led to the identification of two major gene pools (mesoamerican and andean), later confirmed by DNA molecular markers. However, recent advances in the proteomic methodologies provide new tools for a better knowledge of this complex gene family. In this work, relevant aspects on the structure, expression and genetic differentiation of phaseolin were addressed by applying the current bidimensional electrophoresis (2-DE) and mass spectrometry (MALDI-TOF y MALDI-TOF/TOF MS) together with other proteomic methodologies. Optimization of the protein extraction protocol and 2-DE conditions allowed us to obtain high-resolution bidimensional maps of the protein phaseolin that provided a very relevant information. The results of this study revealed for the first time the existence of differentially phosphorylated phaseolin isoforms. In addition, phosphorylation plays an important role in mobilization of the phaseolin because highly phosphorylated isoforms are preferently degraded in germinating seeds. It is noteworthy that this would be the first function assigned to phosphorylated isoforms of seed storage proteins. On the other hand, the analysis of proteomic distances obtained from 2-DE patterns showed that the phaseolin is a reliable marker of genetic differentiation and quality of seed protein in wild and domesticated common bean populations belonging to the mesoamerican and andean gene pools. Accordingly, the phaseolin seems to be a useful marker for the genetic improvement of common bean populations. Overall, the study reveals that the proteomic methodologies are an important tool to achieve a deeper knowledge of the seed storage proteins.





1. INTRODUCCIÓN



1.1. Importancia económica y nutricional de la judía común

La judía común (*Phaseolus vulgaris* L.) es una planta anual herbácea perteneciente a la familia Fabaceae (o Leguminosae) originaria de América cuyas múltiples variedades se cultivan en todo el mundo para el consumo de las vainas verdes y de los granos secos. El género *Phaseolus* incluye 35 especies, de las cuales se cultivan cuatro: *P. vulgaris* L.; *P. lunatus* L.; *P. coccineus* L. y *P. acutifolius* L. En la primera etapa del desarrollo de la semilla, el sistema radical está formado por la radícula del embrión que se convierte posteriormente en la raíz primaria y a los pocos días surgen las raíces secundarias. El tallo se origina a partir del meristemo apical del embrión de la semilla durante la germinación y en las primeras etapas del desarrollo de la planta. En este proceso se van generando los nudos donde se insertan las hojas. La flor es de tipo papilionácea y su color puede ser blanco, rosado o púrpura (Figura 1). El fruto es una vaina con dos valvas, las cuales provienen del ovario comprimido. Las vainas pueden ser de diversos colores, uniformes o con rayas, dependiendo de la variedad (Arias et al., 2007; CIAT, 2014).



Figura 1. Planta y semilla de la judía común

Las leguminosas son uno de los cultivos básicos para la alimentación humana junto con el arroz y el maíz. En el año 2012, la producción total de judía en grano y en verde ascendió a 23,6 y 20,7 millones de toneladas, respectivamente. La

mayor producción a nivel mundial de judía en grano se obtiene por orden de importancia en países de Asia (46,0%), de América (30,6%) y África (21,0%) (Figura 2A); mientras que para la judía verde, Asia produce el 91,4% (Figura 2B). Considerando la producción de judía en grano y en verde conjuntamente, la mayor parte de la producción mundial procede del continente asiático (67,2%), seguido del continente americano (16,8%) y el africano (12,9%) (FAOSTAT, 2013; Figura 2C).

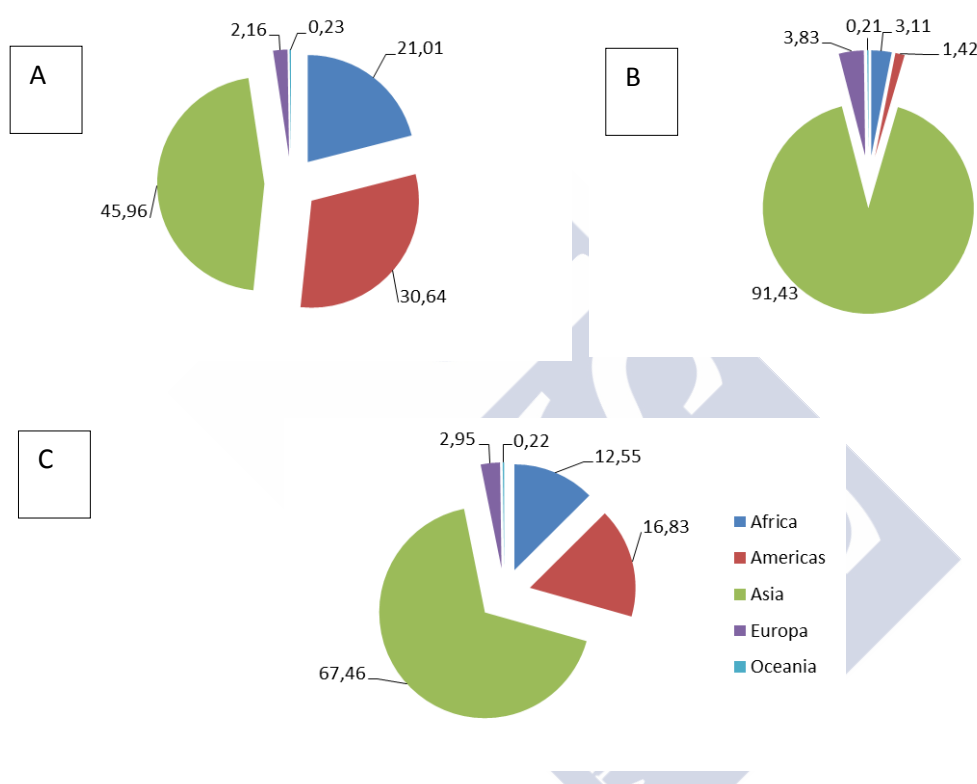


Figura 2. Porcentaje de la producción mundial de judía en grano (A), judía en verde (B) y total (C) en los distintos continentes

Los principales países productores son Birmania, India, Brasil, China, Estados Unidos y México (FAOSTAT 2013, Tabla 1). Para países como Birmania, el principal productor mundial, la judía representa el 10% del total de exportaciones. Las áreas cosechadas en los diferentes países varían de forma notable, estando la producción muy afectada por el rendimiento que varía significativamente desde 407,7 Kg/Ha en la India

hasta los 2116,8 Kg/Ha en países desarrollados como USA. Es evidente que la mejora de la productividad agrícola que se está produciendo en África y en los países sudamericanos en vías de desarrollo pueden suponer un mecanismo para invertir la tendencia de un incremento de la pobreza en esas regiones, por lo que la identificación y posterior optimización de los factores limitantes del rendimiento agronómico son actualmente una prioridad en la mayoría de los programas de mejora de la judía.

Tabla 1. Principales países productores de judía en grano (año 2012)

| | Área cosechada (Ha) | Rendimiento (Kg/Ha) | Producción (t) |
|-----------------|----------------------------|----------------------------|-----------------------|
| Birmania | 2.750.000 | 1418,2 | 3.900.000 |
| India | 9.100.000 | 407,7 | 3.710.000 |
| Brasil | 2.709.485 | 1031,5 | 2.794.854 |
| China | 971.000 | 1607,3 | 1.560.698 |
| USA | 684.081 | 2116,8 | 1.448.095 |
| México | 1.558.992 | 693,3 | 1.080.857 |

Sin embargo, se estima que las estadísticas globales sobre la producción de judía son una infraestima de la producción real dado que en los países en vías de desarrollo los cultivos de la judía se mezclan con otros cultivos, crecen en zonas remotas y/o áreas donde hay conflictos políticos o bélicos, etc., lo que dificulta la toma de datos precisos.

La judía es considerada la leguminosa para la alimentación humana de mayor importancia a nivel mundial, representando la mitad de las leguminosas de grano consumidas mundialmente. Debido al limitado acceso a proteína de origen animal en los países en vías de desarrollo, la judía común (*P. vulgaris* L.) es la fuente proteica de mayor importancia. Así en Ruanda aporta cerca del 60% y 30% de todas las proteínas y calorías ingeridas respectivamente, siendo el segundo en importancia después del maíz (Ferris et al., 2002). Incluso en Brasil, tercer país productor, el 9% de la proteína alimenticia proviene de la judía.

1.2. Composición química y valor nutricional de la semilla de la judía común

La judía común ocupa un lugar importante en la nutrición mundial y en especial en muchas poblaciones de África y América latina, donde es considerada un elemento básico debido a su valor nutritivo y a su bajo coste. Por otro lado el consumo en Europa es más bajo que en otras regiones del mundo aunque está aumentando de forma paulatina en los últimos años (Schneider, 2002). Los factores limitantes del consumo de judía en países de primer nivel son el escaso desarrollo de productos leguminosos adaptados a nuestra vida moderna, una limitada producción local y competencia con otras importaciones más baratas. Otros inconvenientes son los trastornos gástricos que provoca su baja calidad sensorial y que requiere mucho tiempo de cocción (Krupa, 2008).

Desde el punto de vista nutricional, la judía es rica en proteína, fibra, carbohidratos, vitaminas y minerales como K, Mg, P, Fe y Zn (Meiners et al., 1976). Además de su valor nutricional, está demostrado que el consumo regular de judía, tiene un efecto beneficioso sobre la salud, pues influye en la prevención y tratamiento de enfermedades cardiovasculares, diabetes de tipo II y cáncer (Mann y Chisholm, 1999; Bassano et al., 2001; Mathers, 2002). También su ingesta regular puede contribuir a disminuir el nivel de colesterol y constituyen una de las mejores fuentes de ácido fólico junto con guisantes, lentejas y garbanzos (Krupa, 2008).

Una desventaja de la proteína de judía radica en su bajo contenido en aminoácidos azufrados (metionina y cisteína), así como en triptófano cuando se compara con los requerimientos mínimos de aminoácidos recomendados por la FAO/WHO (2007). El contenido total de metionina y cisteína para distintos cultivares y variedades silvestres varía entre 8 y 11 mg/g de proteína (Tabla 2; Montoya et al., 2008a), siendo la cantidad diaria recomendada por la FAO para niños entre 2 a 5 años de 26 mg/g (FAO/WHO, 2007). Además, Montoya et al. (2008a) encontró que las diferencias en el contenido en aminoácidos se correlacionan con el tipo de faseolina en semillas de variedades cultivadas y silvestres.

Tabla 2. Composición de aminoácidos (mg/g de proteína) para distintos tipos de semillas de variedades cultivadas y silvestres (Montoya et al., 2008a).

| | Cultivated beans | | | | | | | | Wild beans | | | | | | | |
|--------------|------------------|------------|------------|------------|------------|------------|----------------|-----------|------------|------------|------------|------------|------------|------------|-----------------|-----------|
| | I | M6 | PA | S | T | T01 | X ^a | SD | H2 | J1 | J3 | J4 | K | L | X ^{ab} | SD |
| ARG | 51 | 54 | 59 | 50 | 53 | 59 | 54 | 4 | 57 | 57 | 51 | 56 | 59 | 57 | 57 | 2 |
| HIS | 32 | 36 | 39 | 33 | 35 | 39 | 36 | 3 | 38 | 39 | 35 | 38 | 40 | 37 | 38 | 2 |
| ILE | 41 | 44 | 48 | 44 | 44 | 48 | 45 | 3 | 47 | 48 | 44 | 46 | 48 | 45 | 46 | 2 |
| LEU | 71 | 76 | 81 | 78 | 74 | 81 | 77 | 4 | 79 | 81 | 74 | 79 | 81 | 77 | 78 | 2 |
| LYS | 61 | 66 | 70 | 68 | 65 | 70 | 67 | 3 | 68 | 69 | 63 | 68 | 61 | 67 | 68 | 3 |
| MET | 8 | 6 | 8 | 8 | 8 | 7 | 8 | 1 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 1 |
| PHE | 57 | 64 | 69 | 58 | 60 | 67 | 63 | 5 | 66 | 67 | 60 | 65 | 68 | 65 | 65 | 2 |
| THR | 31 | 33 | 36 | 26 | 34 | 34 | 32 | 4 | 34 | 34 | 26 | 73 | 35 | 34 | 34 | 3 |
| VAL | 45 | 48 | 49 | 47 | 48 | 49 | 48 | 2 | 50 | 51 | 47 | 49 | 51 | 48 | 49 | 2 |
| ALA | 38 | 43 | 44 | 34 | 37 | 42 | 40 | 4 | 44 | 45 | 40 | 43 | 46 | 44 | 44 | 2 |
| ASP | 101 | 108 | 119 | 116 | 105 | 116 | 111 | 7 | 111 | 114 | 107 | 112 | 115 | 110 | 111 | 3 |
| CYS | 2 | 2 | 2 | 3 | 3 | 2 | 2 | 1 | 2 | 2 | 3 | 2 | 3 | 2 | 3 | 1 |
| GLU | 122 | 146 | 157 | 147 | 146 | 157 | 146 | 13 | 151 | 148 | 137 | 148 | 155 | 144 | 148 | 5 |
| GLY | 32 | 34 | 38 | 36 | 35 | 36 | 35 | 2 | 37 | 36 | 34 | 36 | 39 | 37 | 36 | 2 |
| PRO | 34 | 41 | 45 | 39 | 36 | 44 | 40 | 4 | 43 | 43 | 36 | 41 | 44 | 44 | 43 | 3 |
| SER | 49 | 55 | 59 | 57 | 52 | 57 | 55 | 4 | 55 | 56 | 55 | 56 | 58 | 56 | 56 | 1 |
| TYR | 32 | 36 | 40 | 30 | 29 | 37 | 34 | 4 | 38 | 34 | 28 | 34 | 36 | 45 | 35 | 3 |
| Total | 806 | 891 | 963 | 874 | 864 | 946 | 891 | 57 | 935 | 936 | 850 | 919 | 959 | 910 | 920 | 31 |

^a Valores medios (mg/g de proteína).

^b Valores medios que incluyen las variedades M10, M2, M16, M17, M23 y M25 no mostradas en la tabla.

La faseolina, lectinas e inhibidores de α -amilasas, son las principales proteínas de reserva de la semilla de la judía común (Islam et al., 2002; Krupa et al., 2008). La faseolina representa entre un 40-50% del total de la fracción proteica, siendo por tanto la proteína mayoritaria (Emani and Hall, 2008). Por este motivo, sus características afectan en gran medida al valor nutricional de la semilla. Uno de los factores que merma el valor nutricional de la semilla de *P. vulgaris* reside en la composición química en aminoácidos de la faseolina dado su bajo contenido en metionina, cisteína y triptófano. Además del perfil de aminoácidos, existen otros parámetros para cuantificar la calidad de la proteína como son la digestibilidad y la disponibilidad de la misma. En el caso de la faseolina esta ofrece una pobre digestibilidad, inferior a otras proteínas de origen

animal. En consecuencia, se requiere un tratamiento térmico para que se produzca la desnaturalización de las proteínas, favoreciendo su susceptibilidad a la tripsina (Montoya et al., 2008a). La resistencia a la digestión depende de las características estructurales de la proteína dado que un alto porcentaje de láminas β limita el acceso a las enzimas proteolíticas (Yu, 2005), y la presencia de otros constituyentes como carbohidratos puede aumentar la resistencia a la hidrólisis (Deshpande y Nielsen, 1987). Esto unido al hecho de que es una proteína poco hidrofílica limita el acceso de las proteasas. Para caracterizar fidedignamente el valor nutricional de la judía común se han realizado estudios de proteólisis *in vitro* e *in vivo* usando distintos tipos de faseolina en la dieta de ratones (Montoya et al., 2008b). En definitiva, numerosos estudios se han orientado a la mejora de la composición de los aminoácidos y la digestibilidad de las proteínas con el objetivo de incrementar la calidad nutricional de la semilla de *P. vulgaris*.

Las lectinas y los inhibidores de α -amilasas contienen importantes componentes anti-nutricionales que pueden también ser eliminados con un tratamiento térmico adecuado. En las semillas de *P. vulgaris*, la lectina más abundante es la fitohemaglutinina (PHA) cuya toxicidad en humanos está demostrada, pudiendo causar gastroenteritis, náusea o diarrea (De Oliveira et al., 1988; Nasi et al., 2009). Por ese motivo, los concentrados de judía se utilizan en tratamientos para el control del peso en obesos. Por otra parte, la α -amilasa descompone los hidratos de carbono de cadena larga para que puedan ser absorbidos en el intestino. Por tanto, la presencia de inhibidores específicos de α -amilasas disminuyen también la digestibilidad de semilla de judía.

1.3. Proteómica de *Phaseolus vulgaris* L.

El término “proteoma” fue acuñado por primera vez por el doctor Wilkins durante un congreso de electroforesis bidimensional celebrado en el año 1994 en la ciudad de Siena. Es una adaptación del término “genoma”, que simplifica la idea de “el equivalente PROTeico de un genOME”. El proteoma se define por tanto como una mezcla compleja de proteínas presentes en una unidad biológica para un estado

específico del desarrollo bajo unas determinadas condiciones. La proteómica abarca el estudio de los perfiles proteicos, la cuantificación de proteínas, las modificaciones post-traduccionales, así como también las interacciones entre proteínas que se puedan dar en estas unidades biológicas.

La técnica de la electroforesis bidimensional fue inicialmente desarrollada por O'Farrel (1975) para resolver mezclas complejas de proteínas. Las mejoras paulatinas en la técnica como la introducción de los *Immobilized pH Gradient* (IPG) strips incrementaron notablemente su reproducibilidad y facilitaron en análisis comparativo de los perfiles proteicos entre muestras y la posterior identificación por espectrometría de masas (MS). El uso de tintes fluorescentes ha supuesto otro importante avance en la detección de las proteínas debido a su gran sensibilidad, linealidad y homogeneidad. Por otra parte, la mejora de los métodos de ionización de MS que permiten una ionización suave de los péptidos, hace posible la identificación de proteínas con mayor precisión y exactitud.

La proteómica vegetal ha evolucionado de forma exponencial en los últimos años. Cabe destacar la caracterización del proteoma de *Arabidopsis thaliana* que es usado como un organismo modelo en genética clásica (Weckwerth et al., 2008). Actualmente, existen también estudios proteómicos de las legumbres más importantes económicamente como son la soja (*Glycine max*) (Hajdich et al., 2005), el guisante (*Pisum sativum*) (Curto et al., 2006) y la judía común (*P. vulgaris*) (Marsolais et al., 2010; De la Fuente et al., 2011). También se utilizan también especies modelo como son *Medicago truncatula* (Gallardo et al., 2003) y *Lotus japonicus* (Nautrup-Pedersen et al., 2010) para el estudio del desarrollo de la semilla en leguminosas.

P. vulgaris es una especie diploide con una dotación cromosómica $2n=22$ en la mayoría de las especies y un tamaño del genoma haploide que varía entre 588 y 637 Mb. Debido a la importancia de las leguminosas, se han concluido distintos proyectos de secuenciación del genoma en especies modelo como *M. truncatula* y *L. japonicus*. Recientemente, ha concluido el proyecto de secuenciación del genoma de *P. vulgaris* (Schmutz et al., 2014) lo que es particularmente relevante para una mejor identificación de las proteínas mediante MS.

La extracción de proteínas a partir de muestras de tejidos supone el paso más crítico en cualquier estudio de proteómica vegetal. En este sentido el análisis proteómico incluye una serie de etapas que son más problemáticas en tejidos vegetales que en otro tipo de organismos. La matriz vegetal está en su mayor parte formada por pared celular y vacuolas mientras que la parte correspondiente al citosol únicamente representa el 1-2% de la célula (Carpentier et al., 2005), lo que explica la baja concentración de proteína. La complicación aumenta si a esto le sumamos la presencia de sustancias interferentes como compuestos fenólicos, enzimas oxidativas y proteolíticas, terpenos, pigmentos, ácidos orgánicos, inhibidores iónicos y carbohidratos. Esta realidad interfiere notablemente con la capacidad de resolución de 2-DE pues provoca *streaking* (horizontal y vertical) y *smearing*, disminuyendo el número de *spots* proteicos que se pueden resolver. En definitiva, la etapa de extracción de proteínas requiere un intenso estudio de optimización dada su influencia posterior en la separación de las proteínas mediante 2-DE. Uno de los métodos de extracción de proteínas más utilizado se basa en la precipitación de proteínas con tricloroacético (TCA)/Acetona que minimiza la degradación proteica al disminuir la actividad enzimática, así como eliminación de los iones salinos (Wang et al., 2007). Otro método de extracción, alternativo al clásico “método del fenol”, consiste en que las proteínas una vez disueltas con fenol se precipitan posteriormente con metanol y amonioacetato, seguido por una redisolución en tampón propio del isoelectroenfoco. Con este método de extracción proteica se obtiene un precipitado con una mínima concentración de contaminantes, incluso en aquellos casos en los que el extracto proteico provenga de tejidos recalcitrantes (Saravanan y Rose, 2004).

1.4. Faseolina

1.4.1. Estructura y base genética

Las proteínas de reserva (SSP, *seed storage proteins*) representan la fracción proteica mayoritaria en la semilla de las plantas. La semilla se divide en tres partes: (1) el embrión, (2) el tejido donde se almacenan los nutrientes de reserva y (3) la cubierta que protege la semilla. Las SSP se pueden clasificar en función de su solubilidad, siendo las más abundantes las prolaminas (solubles en mezclas de agua/alcohol), las gluteninas (solubles en ácidos o bases) para semillas monocotiledóneas y las globulinas (solubles en disolución salina) para semillas dicotiledóneas. Las globulinas a su vez se pueden dividir en varios tipos atendiendo a sus coeficientes de sedimentación: el tipo 7S o vincilina y el tipo 11S o legumina. En la semilla de *P. vulgaris*, la faseolina representa hasta un 40-60% de la proteína total y es una globulina soluble en soluciones salinas (NaCl 0,5M) en un rango amplio de pH (Osborne, 1894; Bollini et al., 1983).

Las globulinas 7S son proteínas que se caracterizan por ser proteínas triméricas con una masa molecular comprendida entre 150-190 kDa y que tienen pocos residuos de cisteína. La faseolina es una globulina 7S, que consiste en distintas subunidades glicosiladas de peso molecular entre 43 y 53 kDa. (Bollini y Chrispeels, 1978). La heterogeneidad de los distintos polipéptidos que constituyen las distintas subunidades vienen dada por una divergencia en la secuencia de ADN y por los procesos de modificaciones post- y co-traduccionales. La base genética de la faseolina no está totalmente definida pero se estima que está codificada por una familia génica integrada por 6 a 10 genes codominantes por genoma haploide, organizados en un único clúster en el cromosoma 7. Los genes de la faseolina exhiben una estructura discontinua constituida por cinco intrones y seis exones y las moléculas de ARNm maduras son el resultado de cinco o más fenómenos de *splicing* alternativo (Slightom et al., 1985). Estudios de secuenciación de nucleótidos realizados a partir de la clonación de cDNA indican que la familia génica de la faseolina se subdivide en dos subfamilias génicas, α

y β , constituidas por dos tipos de genes α y β , respectivamente, de distinto peso molecular, que dan nombre a las dos subfamilias de la faseolina (Slightom *et al.*, 1985). Los genes del tipo α y β muestran un alto grado de homología (98%) en sus secuencias de nucleótidos. Esta baja diversidad en la secuencia de nucleótidos sugiere que los genes del tipo α y β surgieron de un gen ancestral común a lo que siguió una escasa divergencia debido a los mecanismos de evolución concertada (Slightom *et al.*, 1983; 1985; Talbot *et al.*, 1984; Anthony *et al.*, 1990; Kami and Gepts, 1994; Gepts *et al.*, 1998). La diferencia más notable entre los dos tipos de genes reside en la existencia de dos diferentes repeticiones directas que se ubican en las regiones codificadoras (cuarto y sexto exón) de los genes del tipo α , produciendo diferencias de longitud entre las faseolinas de tipo α y β . La primera de las repeticiones tiene una longitud de 15 pb y se localiza entre las posiciones nucleotídicas 711 y 740, mientras que la segunda repetición de 27 pb se localiza entre las posiciones 1322 y 1375. El estudio comparativo de las secuencias de ADN de los genes del tipo α y β sugiere que los genes del tipo α evolucionaron de un gene ancestral de tipo β mediante dos eventos de duplicación y la acumulación de varias mutaciones puntuales (Slightom *et al.*, 1985). También se han descrito polipéptidos de tipo γ , pero parece que proceden más bien de una modificación post-traducciona (glicosilación) de los polipéptidos de tipo β . Cabe subrayar que los estudios disponibles sustentan el hecho de que la mayor parte de la heterogeneidad entre los polipéptidos de tipo α y β se debe a modificaciones post-traduccionales y no a sustituciones de aminoácidos (Brown *et al.*, 1981a)

1.4.2. Biosíntesis de la faseolina

La faseolina es una proteína de la vía secretora que se acumula en las vacuolas de las células cotiledóneas. La biosíntesis de este tipo de proteínas se inicia en los ribosomas unidos al retículo endoplasmático (RE) rugoso. Una secuencia de aminoácidos llamada secuencia señal es reconocida por la partícula receptora de la señal (SRP, *signal recognition particle*), la cual, a su vez se une a un receptor en la membrana del RE rugoso. La SRP hace posible que las proteínas recién sintetizadas adopten una

conformación correcta para permitir el transporte de las mismas al interior del RE, como se muestra en la Figura 3.

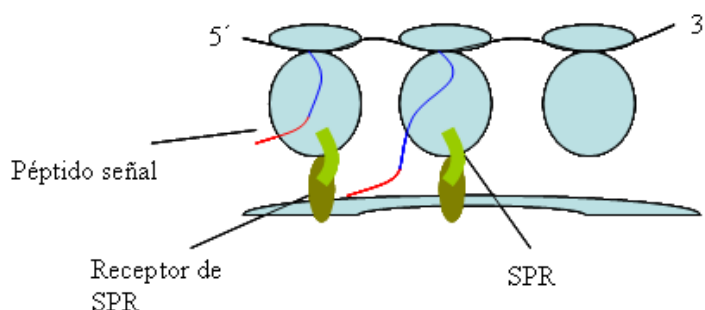


Figura 3. Diagrama esquemático del paso de nuevas proteínas a través del RE

Los polipéptidos precursores de la faseolina contienen un péptido señal de 24 aminoácidos que se rompe en el momento de su translocación al lumen del RE. Se han identificado tres polipéptidos distintos de la faseolina después de la ruptura del péptido señal: dos polipéptidos del tipo α de 411 y 412 aminoácidos y un polipéptido del tipo β de 397 aminoácidos (Slightom et al., 1983, 1985). La faseolina experimenta N-glicosilación en el lumen del RE (Bollini et al., 1983; Vitale et al., 1993). Se han descrito dos lugares posibles de glicosilación (Figura 4), uno de los cuales siempre está glicosilado mientras que en el otro es menos frecuente. Los polipéptidos que están doblemente glicosilados tienen dos glicanos $\text{Man}_7(\text{GlcNAc})_2$ en la posición Asn^{252} y la $\text{Man}_9(\text{GlcNAc})_2$ en la Asn^{341} mientras que los que están glicosilados en una sola posición tienen un glicano $\text{Xil-Man}_3(\text{GlcNAc})_2$ en la Asn^{252} . El procesamiento de los glicanos a través de la eliminación de residuos de glucosa se produce por la acción de enzimas situadas en el aparato de Golgi, donde se realizan las últimas transformaciones en las cadenas de oligosacáridos (Sturm et al., 1987).

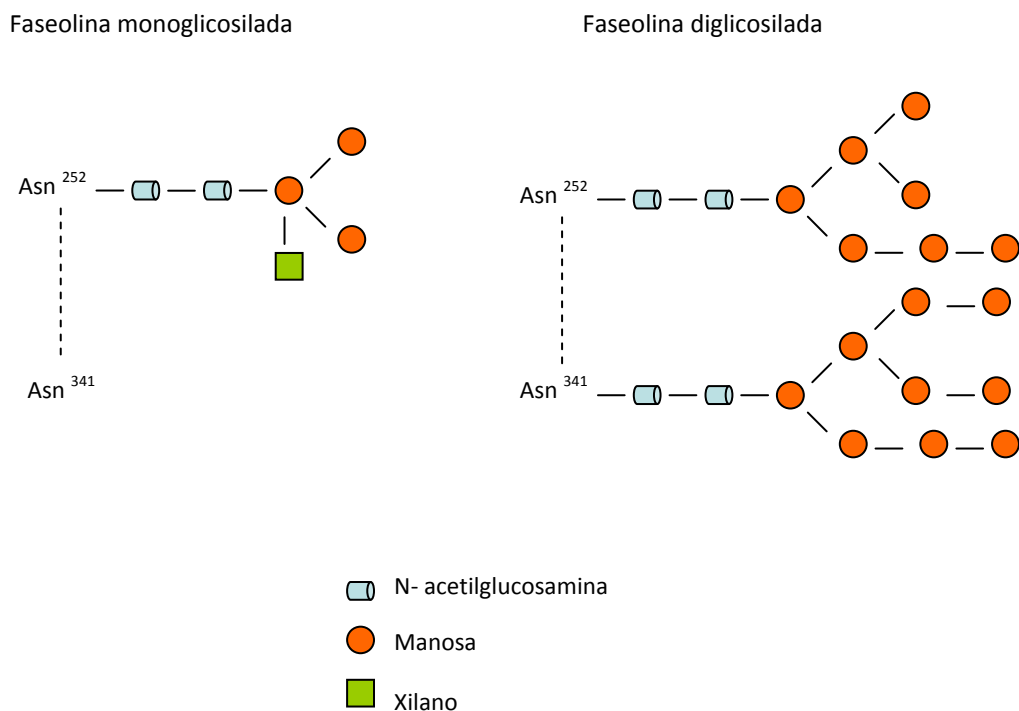


Figura 4. Esquema representativo de la faseolina mono y diglicosilada

Como se ha señalado anteriormente, la faseolina es un trímero de los tres polipéptidos glicosilados de tipo α (411 y 412 aminoácidos) y de tipo β (397 aminoácidos) que comprenden estructuras similares conteniendo gran cantidad de hélices alfa y una pequeña porción de láminas beta (Lawrence et al., 1990). El trímero ($M_r \sim 150$ kDa) se asocia en una estructura dodecamérica con un pH inferior a 4,5. Después de su síntesis y procesado, estas proteínas se transportan a los llamados cuerpos proteicos y vacuolas, donde se acumulan en grandes cantidades. Los cuerpos proteicos son pequeños (1-10 μm) y rellenan las vacuolas pero dentro de estas también se encuentran las lectinas y las hidrolasas ácidas. Las proteínas de reserva son hidrolizadas en el citoplasma por la acción de las proteasas endógenas en el momento de la germinación para ser utilizadas en la síntesis de proteínas.

1.5. Fosfoproteoma

Las proteínas pueden dar lugar a isoformas diversas como consecuencia de la existencia de distintas modificaciones post-traduccionales (PTMs), tales como la fosforilación, acetilación, glicosilación y la ubiquitinización. Las PTMs de las proteínas aumentan considerablemente la complejidad de los proteomas debido a que las proteínas pueden estar modificadas en distintos lugares y por distintos grupos funcionales. De entre todas las modificaciones post-traduccionales, una de las más estudiadas es la fosforilación por su reversibilidad e importancia en la regulación de su actividad (Agrawal y Thelen, 2006). Los procesos de fosforilación y desfosforilación están controlados por las quinasas y fosfatasas que además producen cambios conformacionales en la proteína haciendo que la correspondiente enzima se active o se desactive. El estudio de estos cambios dinámicos en respuesta a señales ambientales aumenta nuestro conocimiento de los procesos celulares. En vegetales se han encontrado más de 1.000 quinasas distintas, un número que dobla aproximadamente al existente en humanos. Sin embargo, se han identificado solo unos pocos sustratos (De la Fuente van Bentem et al., 2007).

La disciplina de la fosfoproteómica se ocupa de la identificación de las fosfoproteínas, de la localización y cuantificación de los lugares de fosforilación y de la unión de las proteínas a sus quinasas específicas. El análisis espacial y temporal de la fosforilación de la proteína es de gran interés para conocer el significado biológico ya que los ensayos *in vitro* no son reflejo de los procesos vegetales debido a la reversibilidad de la fosforilación, por lo que la experimentación tiene que ir más allá (Kline-Jonakin et al., 2011). Tampoco la identificación de lugares de fosforilación es suficiente para conocer los cambios dinámicos que se establecen en respuesta a señales ambientales. Además, la concentración de las fosfoproteínas es baja en muchas ocasiones lo que incrementa las dificultades de su caracterización e interpretación. El desarrollo de las técnicas actuales hace posible las identificaciones masivas de fosfoproteínas pero es necesario dilucidar los códigos de fosforilación que rigen los procesos celulares (Laugesen et al., 2006). Debido al conocimiento escaso que actualmente se tiene sobre las implicaciones de la fosforilación en plantas, es necesario

hacer un gran esfuerzo para avanzar en el conocimiento de los procesos celulares en los que están implicadas las fosfoproteínas y los mecanismos que las regulan.

En los sistemas eucariotas la fosforilación proteica se produce en los residuos de serina (S), treonina (T) y tirosina (Y), siendo más abundante en los dos primeros. Así, en un estudio de fosforilación a gran escala realizado en *Arabidopsis* se encontró que las abundancias relativas de fosforilación en S, T y Y fueron 85,7%, 10,7% y 4,4%, respectivamente (Sugiyama et al., 2008). Se estima que hasta el 30% de todas las proteínas en organismos multicelulares podrían estar fosforiladas en cualquier momento, aunque sólo algunos de estos lugares de fosforilación tienen importancia biológica (De la Fuente van Bentem y Hirt, 2007).

Uno de los métodos más utilizados para el análisis del fosfoproteoma consiste en llevar a cabo un enriquecimiento de la muestra en fosfopéptidos mediante la cromatografía de afinidad por iones metálicos inmovilizados (IMAC, *immobilized metal-ion affinity chromatography*) o columnas de dióxido de titanio (TiO₂) previo al análisis por MS (Nakagami et al., 2010). Un método complementario consiste en combinar la técnica de 2-DE con una tinción específica de fosfopéptidos fluorescente con la sal denominada Pro-Q® Diamond phosphoprotein stain (Pro-Q DPS). Esta tinción exhibe una alta sensibilidad y linealidad y permite realizar un análisis cuantitativo global de los fosfopéptidos de las proteínas correlacionando la fracción de fosfato con el volumen de spot (Agrawal y Thelen, 2005).

1.6. Germinación de la semilla

La semilla completa el ciclo reproductivo de la planta que comienza con la floración y la polinización. El desarrollo de la semilla se inicia en la embriogénesis y termina unos días después de la fecundación. La maduración de la semilla es una etapa importante dentro de su desarrollo en la cual el crecimiento del embrión se detiene, los componentes de reserva (aceites, polisacáridos y proteínas) se acumulan, el tegumento protector se diferencia y aumenta la tolerancia al estrés hídrico, lo que conduce a un estado de dormancia. La dormancia es un mecanismo que bloquea la germinación para

favorecer la supervivencia de la semilla en condiciones adversas. La etapa de dormancia se interrumpe como respuesta a una serie de señales o estímulos físicos o químicos como son la temperatura, luz y equilibrio hormonal (ácido abscísico y giberelinas). En estados tempranos durante el crecimiento del embrión se produce la acumulación de compuestos de reserva y en particular de las SSP como la faseolina. El promotor de la faseolina representa uno de los promotores más estudiados específicos de la semilla de judía. En otros tejidos vegetales el promotor está silenciado, mientras que en la semilla, durante la embriogénesis se potencia la activación del promotor de la expresión de ABA (ácido ábscísico) a través de un remodelado de la cromatina (Ng y Hall, 2008). A nivel molecular la dormancia fisiológica es un estado activo con complejas interacciones que continuamente integran las señales ambientales y sus respuestas. La relación dinámica entre ABA y GA (ácido giberélico) dirige los mecanismos de dormancia y germinación que alteran la sensibilidad de la semilla a las condiciones ambientales. (Finch-Savage et al., 2006). Los eventos de fosforilación/defosforilación juegan un papel crucial en la señalización de ABA lo cual implica la acción de muchas quinasas y fosfatasas (Ghelis et al., 2008). Además el ABA está relacionado con diversos procesos fisiológicos como la germinación, la maduración de la semilla así como la adaptación de los tejidos vegetales a condiciones ambientales estresantes.

Durante la germinación se produce la hidrólisis de las proteínas de reserva mediante las distintas enzimas proteolíticas, denominadas proteasas. A medida que avanza la germinación las SSP se transforman en péptidos de menor peso molecular, siendo por tanto una fuente importante de nitrógeno. Estos aminoácidos pueden ser utilizados para la síntesis de proteínas en la plántula y así poder comenzar su vida de forma autótrofa.

Estudios a gran escala del fosfoproteoma en cuanto a su identificación y a su cuantificación fueron llevados a cabo en soja (*Glycine max*), colza (*Brassica napus*) y *Arabidopsis* (*A. thaliana*) (Agrawal y Thelen, 2006; Meyer et al, 2012). Han et al. (2014) comprobaron en arroz que la fosforilación de proteínas está relacionada con la modificación de otras proteínas y con la regulación transcripcional, además se producen cambios temporales en los niveles de fosforilación de las proteínas durante la germinación en el caso del arroz. Por este motivo se puede pensar que la fosforilación

de las proteínas parece que tiene una gran implicación en el proceso de germinación, sin embargo el papel funcional de estas proteínas fosforiladas es todavía desconocido.

1.7. Origen y domesticación de la judía común

La domesticación de las plantas es un proceso que combina la evolución natural junto con una selección practicada por el hombre y mediante este proceso se obtienen los cultivos domesticados a partir de sus progenitores silvestres. Es de máximo interés identificar los sitios originales de domesticación desde el punto de vista evolutivo, así como de gran utilidad para la mejora genética y la conservación de las especies.

La diversidad de las poblaciones silvestres y cultivadas de la judía común ha sido estudiada en base a caracteres morfológicos, agronómicos, marcadores bioquímicos (alozimas y faseolina) y polimórficos del ADN nuclear y de los orgánulos citoplasmáticos: mitocondrias y cloroplastos (microsatélites y SNPs). La faseolina ha sido utilizada como marcador molecular para el estudio de la diferenciación genética y la evolución de las poblaciones silvestres y cultivadas de *P. vulgaris*.

El género *Phaseolus* comprende alrededor de 30 especies de las cuales sólo cinco están domesticadas (*P. acutifolius* A. Gray, *P. coccineus* L., *P. lunatus* L., *P. polyanthus* Greenman y *P. vulgaris* L.). Los primeros estudios sobre el origen de las primeras poblaciones silvestres de *P. vulgaris* están basados en datos arqueológicos, botánicos e históricos y situaban su origen en Mesoamérica, concretamente en México desde Jalisco hasta Oaxaca. Desde esta zona mesoamericana las formas silvestres migraron hacia Sudamérica donde se constituyó el acervo andino antes de la domesticación (Bitochi et al., 2012; Desiderio et al., 2013). Estas poblaciones silvestres experimentan posteriormente múltiples domesticaciones en tiempo y espacio, de forma independiente en estas dos zonas del continente americano. De esta manera se origina el acervo mesoamericano situado en México, América Central y Colombia, y acervo andino en el Sur de Perú, Bolivia y Argentina (Figura 5; Gepts y Debouck, 1991; Singh, 2001). La variabilidad genética se debe principalmente a las múltiples domesticaciones y en menor medida a los ocasionales cruzamientos entre las formas cultivadas y las

silvestres, de este modo el acervo mesoamericano es de mayor diversidad que el andino. Dentro de estos dos acervos como se muestra en la Figura 5 se pueden distinguir seis razas: tres razas se originadas en Mesoamérica (Durango, Jalisco y Mesoamérica) y otras tres en los Andes (Chile, Nueva Granada y Perú) (Singh et al., 1991). Posteriormente los distintos cultivares de *P. vulgaris* se fueron introduciendo en Europa a partir del siglo XVI como consecuencia del descubrimiento de América, lo que produjo una diversificación secundaria de la especie en el continente europeo.



Figura 5. Distribución geográfica de las principales razas de *Phaseolus vulgaris* L. en las zonas mesoamericana y andina

Los estudios mediante electroforesis en geles monodimensionales permitieron establecer una asociación entre el tipo de faseolina con el origen geográfico de la semilla. En América Central, los cultivares se caracterizaban principalmente por la presencia de faseolina de tipo “S” (Sanilac) y semillas pequeñas. Mientras en la región andina, la faseolina mayoritaria era de tipo “T” (Tendergreen) y las semillas eran más grandes (Singh et al., 2001). La faseolina T y S representan el 90% dentro de los cultivares de judía común (Figura 6). La distribución de ambos tipos de faseolinas es similar tanto en los cultivares como en las formas silvestres.

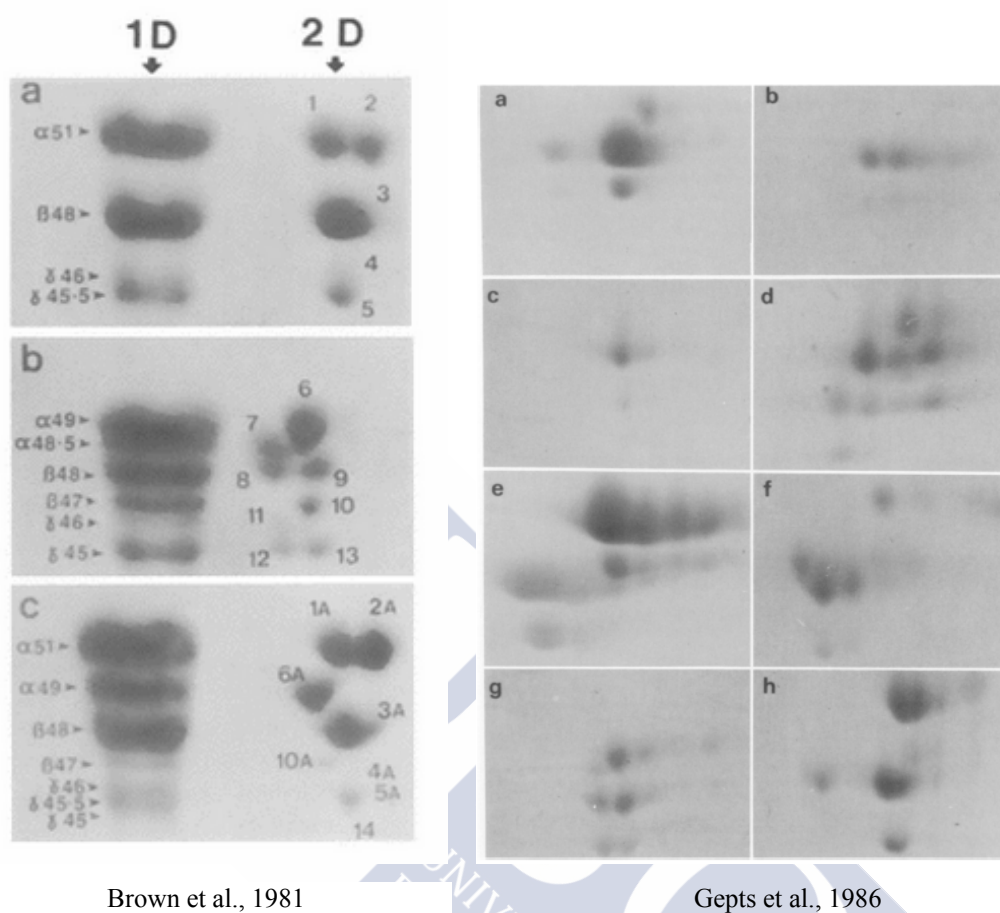


Figura 6. Patrones electroforéticos bidimensionales de la faseolinas S y T

Se han identificado otros patrones de faseolina, como la faseolina “H” (Huevo de Huanchaco) y “C” (Contender) de origen andino o la faseolina “B” (Boyaca 22) de origen mesoamericano (Gepts et al., 1986). Hasta la actualidad, más de 40 tipos de faseolina han sido estudiadas utilizando 1-D SDS-PAGE. Los distintos patrones electroforéticos monodimensionales muestran las diferencias entre los distintos tipos de faseolina en cuanto al número de subunidades (2-6 polipéptidos) y a los distintos pesos moleculares (40-54 kDa) de las distintas subunidades. Estos polipéptidos también varían en cuanto a su punto isoeléctrico desde 5,2 a 5,7 (Brown et al., 1981b). También se han realizado algunos estudios empleando electroforesis bidimensional con pobres resultados desde el punto de vista de resolución de los spots. Afortunadamente, hoy en día se han incorporado numerosas innovaciones, mejoras y optimizaciones en la

técnicas 2-DE, destacando las tiras con gradiente de pH inmovilizado (*IPG strips*). Estas mejoras en la técnica permiten mejorar la resolución y además son compatibles con el posterior análisis de las proteínas por MS.







2. OBJETIVOS



1. Obtención de un mapa de alta resolución del conjunto complejo de isoformas de la proteína faseolina en semilla de la judía común (*P. vulgaris* L.) basado en la electroforesis bidimensional (2-DE) y la espectrometría de masas (MALDI-TOF y MALDI-TOF/TOF MS), mediante la optimización del método de extracción de proteínas y los protocolos experimentales de la 2-DE.
2. Identificación de isoformas de la faseolina con glicosilaciones únicas y dobles en el mapa 2-DE de alta resolución de los cultivares (cvs.) Sanilac (S) y Tendergreen (T) perteneciente a los acervos mesoamericano y andino, respectivamente, de la judía común.
3. Búsqueda de isoformas fosforiladas de la faseolina en semillas durmientes y en germinación de los cvs. S y T de *P. vulgaris*.
4. Estudio del posible papel funcional de las isoformas fosforiladas de la faseolina durante la germinación.
5. Evaluación de los perfiles 2-DE de la faseolina como marcadores de diferenciación genética en poblaciones silvestres y cultivadas de los acervos mesoamericano y andino de *P. vulgaris*.
6. Análisis de la relación de las distancias proteómicas entre faseolinas de variedades silvestres y cultivadas de los acervos mesoamericano y andino con la calidad de la semilla estimada mediante el contenido en metionina, un aminoácido esencial deficiente en la semilla de la judía común.





3. RESULTADOS Y DISCUSIÓN



3.1. Análisis proteómico de las semillas de judía común (*P. vulgaris* L.) basado en la electroforesis bidimensional (2-DE)

3.1.1. Optimización del método de extracción proteica y de las condiciones electroforéticas para el proteoma de *P. vulgaris* L.

En esta parte del estudio se llevó a cabo un análisis del método más adecuado para la extracción de proteínas de la semilla de la judía común (*P. vulgaris* L.) y el estudio de su proteoma. Se ensayaron un total de tres métodos de extracción de proteínas diferentes: ácido tricloroacético-acetona (TCA-acetona) (Santoni et al., 1994), fenol (Saravanan y Rose, 2004) y el kit Clean-Up (GE Healthcare). La eficacia relativa de los tres métodos de extracción se evaluó con el fin de conseguir a una mayor resolución de las proteínas en los geles bidimensionales y posterior identificación de los *spots* proteicos mediante la técnica de espectrometría de masas MALDI-TOF/TOF. El estudio comparativo entre los métodos de extracción se realizó utilizando semillas del genotipo ICA Pijao, un genotipo domesticado de Colombia perteneciente al acervo genético mesoamericano. Para cada uno de los protocolos de extracción se analizaron cuatro réplicas técnicas a partir del material biológico resultante de la mezcla de tres semillas diferentes.

En paralelo, se procedió a una optimización de la técnica de la 2-DE para la obtención de perfiles proteicos de la semilla de *P. vulgaris* de alta resolución (i.e., incremento del número y resolución de los *spots* así como su reproducibilidad), variando una diversidad de parámetros electroforéticos (rango de pH en la primera dimensión, tamaño de los geles, voltaje y tiempo de corrido electroforético en la primera y segunda dimensión, etc.).

La Figura 7 muestra geles 2-DE representativos obtenidos a partir de extractos de proteína total de semilla para cada uno de los tres protocolos de extracción. Con los tres métodos se obtuvieron geles de gran resolución, con poco *smearing* y *streaking* tanto vertical como horizontal.

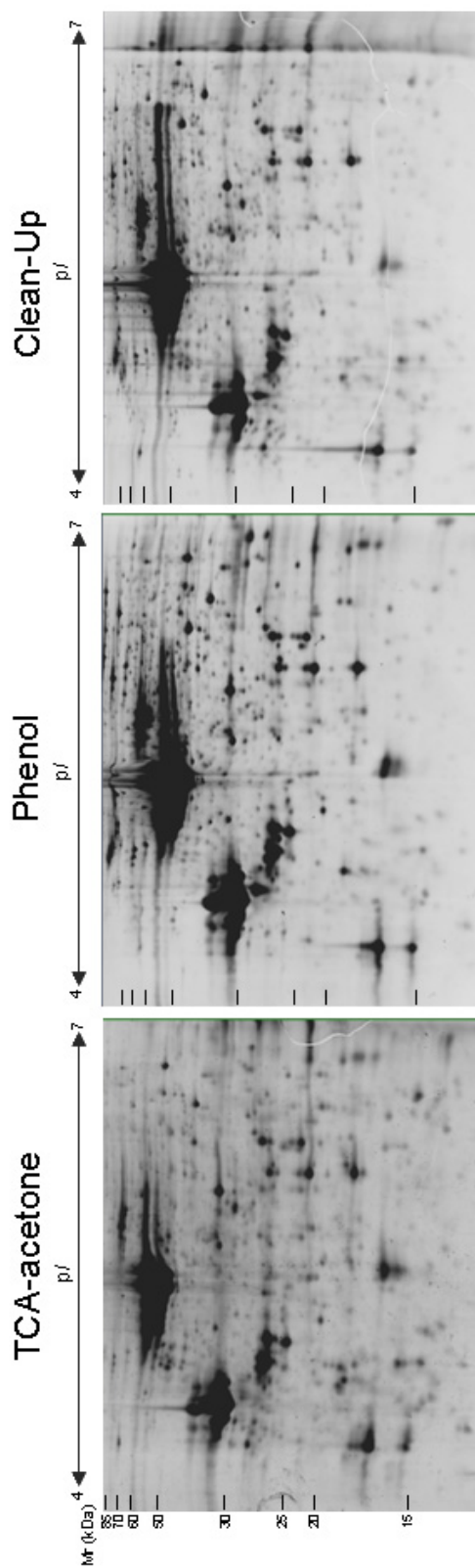


Figure 7. Representative 2-DE gels of proteins from bean seed extracted by three different protocols: TCA-acetone, phenol and clean-up. Proteins were separated on a linear pH 4-7 gradient in the first dimension and visualized using Sypro Ruby® staining. The approximate positions of molecular mass markers (kDa) are indicated on the left of the figure (Source: De la Fuente et al., 2011).

Además con el método del fenol se redujo el *background* de los geles, aumentando la nitidez de los *spots* en comparación con los otros dos métodos. Mediante el análisis de imagen con el software PDQuest™ v8.0.1 (Bio-Rad Laboratories) se pudieron detectar un total de 571 *spots* proteicos entre los distintos protocolos.

Las diferencias entre los tres protocolos de extracción mediante un análisis de componentes principales, a partir de los volúmenes normalizados de los *spots*, se muestra en la Figura 8. Este análisis reveló que los dos primeros componentes principales PC1 y PC2 explican el 27,4% y el 15,0% de la variación, respectivamente, del volumen de los *spots* entre protocolos. Cabe subrayar que el PC1 separa claramente las muestras de fenol de los otros dos tipos de muestras (TCA-acetona y Clean-Up), mientras que estas dos últimas son separadas por el PC2.

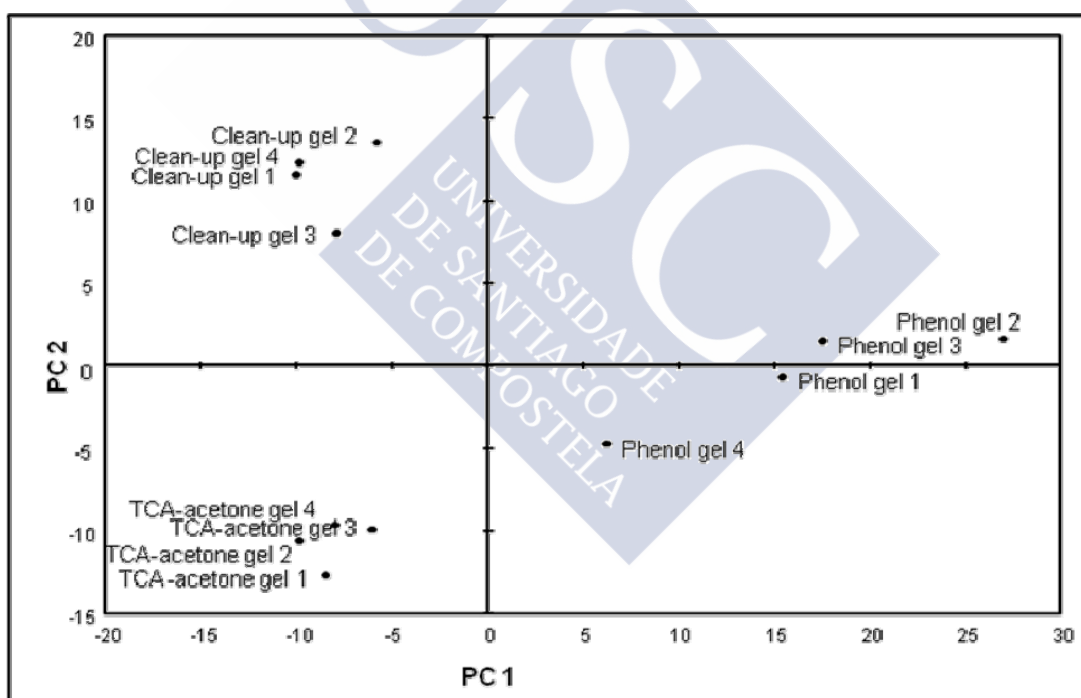


Figure 8. Principal component analysis for normalized volume of spots. Score plot for principal components (PCs) one and two (Source: De la Fuente et al., 2011)

El método de fenol dió lugar a un mayor número de *spots* únicos que los otros dos métodos de extracción: 150 *spots* únicos con el método del fenol en comparación

con los 25 *spots* únicos en TCA-acetona y los 39 en Clean-Up. Cuando las imágenes de los geles se dividieron en cuatro cuadrantes (Q1-4), numerados del 1 al 4 en el sentido de las agujas del reloj, se observó que la gran mayoría de *spots* únicos quedaron situados en el cuadrante Q2 correspondiente a las proteínas de mayor pI y M_r . Sin embargo, el método del fenol proporciona un incremento estadísticamente significativo del número de *spots* en el cuadrante Q2 en comparación con los otros métodos (ANOVA: $F= 7,27$, $P= 0,013$). Por tanto, es un método con una mayor capacidad de extracción de proteínas de pH más básico y de mayor M_r que los métodos TCA-acetona y Clean-Up. Resultados similares fueron encontrados en extractos proteicos obtenidos a partir del algodón, manzana, tomate y el euhalophyte *S. europea* (Carpentier et al., 2005; Wang et al., 2007; Xie et al., 2009). Sin embargo, Saravanan y Rose (2004) no detectaron un patrón diferencial consistente para pI o M_r entre los métodos del TCA-acetona y fenol. También, se llevó a cabo un análisis estadístico con el fin de detectar diferencias cuantitativas en los volúmenes normalizados de los *spots* entre los tres protocolos de extracción. El análisis reveló que 322 de 571 *spots* (56,4%) presentaron *fold change* superiores a 2 entre los protocolos siendo estas diferencias estadísticamente significativas (test U de Mann-Whitney, $P < 0,05$), mientras que los 249 *spots* restantes (43,6%) estaban compartidos en los tres protocolos de extracción. En general, nuestros resultados sugieren al método de fenol como la primera opción para la extracción de proteínas en semillas de judía, ya que da lugar a los geles 2-DE de más calidad, una mayor cantidad de *spots* para las proteínas menos ácidas y de alto peso molecular, y el mayor número de *spots* de mayor volumen. Sin embargo, la aplicación de los métodos de extracción TCA-acetona y Clean-Up unidos al método del fenol enriquece el extractoma.

La calidad del extractoma influye en el nivel de éxito de la identificación de los *spots* proteicos mediante las técnicas de MS. Por otra parte, los extractomas obtenidos mediante los diferentes protocolos de extracción pueden manifestar diferencias cualitativas y cuantitativas. En este estudio, se seleccionaron un total de 50 *spots* en geles 2-DE, diferencialmente o no diferencialmente extraídos por los tres protocolos de extracción ensayados, para su análisis mediante MS con el fin de evaluar el nivel

diferencial de calidad de los tres tipos de extractomas y sus diferencias en los perfiles proteicos. La identificación de proteínas se realizó a partir de *spots* escindidos de los geles 2-DE seguido de la digestión peptídica con tripsina e identificación de las proteínas a partir de las digestiones tripticas, mediante las técnicas de MALDI-TOF MS y MALDI-TOF/TOF MS, usando el espectrómetro de masas 4800 MALDI-TOF/TOF de Applied Biosystems. La identificación proteica se obtuvo a partir de la combinación de los datos de la huella peptídica (PMF) y de los espectros de fragmentación de péptidos de cada muestra, con ayuda del software GPS Explorer v3.6 (Applied Biosystems) usando el Mascot software v2.1. (Matrix Science), para la búsqueda en la base de datos de proteínas (NCBI nr release data 20100526).

Los datos obtenidos en nuestro estudio demuestran que los tres tipos de extractoma (TCA-acetona, fenol y Clean-Up) y su posterior separación por 2-DE, conducen a una identificación óptima de proteínas de semilla de judía mediante MS (Tabla 3). Así, se llegaron a identificar 35 *spots*, el 70% de los *spots* seleccionados. De los 35 *spots* se correspondieron con un total de 40 proteínas diferentes identificadas, debido a que cuatro *spots* (26, 66, 83 y 486) dieron identificaciones múltiples. La mayor parte de las proteínas identificadas correspondieron a *P. vulgaris* (67,5%) o a otras especies de leguminosas (15%): *P. acutifolius*, *L. japonicus*, *Pisum sativum*, *Daucus carota* y *Vigna unguiculata*. La proporción de identificaciones obtenida en este estudio es similar a estudios proteómicos previos llevados a cabo en otras semillas de especies de plantas como es el caso de la especie modelo *A. thaliana*. Además las identificaciones poseen un nivel de confianza elevado en base al *Mascot score* y al *coverage* que se muestran en la Tabla 3. A pesar del limitado número de secuencias disponibles en ese momento en las bases de datos, debido a que los proyectos de secuenciación del genoma de la judía todavía estaban en marcha en ese momento. Actualmente se dispone de una información más completa para la identificación de proteínas por MS, debido a que los proyectos de secuenciación del genoma de *P. vulgaris* han concluido recientemente (Schmutz et al., 2014). Sin embargo varios factores contribuyeron a aumentar el número de identificaciones a pesar de no conocer la secuencia del genoma completo. En primer lugar un gran número de *spots* (40%)

RESULTADOS Y DISCUSIÓN

fueron identificados como faseolinas, lectina fitohemaglutinina o como lectinas relacionadas con el inhibidor de α -amilasa, todas ellas proteínas de la judía bien caracterizadas y que conforman aproximadamente el 70% de la proteína total de la misma (Marsolais et al., 2010). En segundo lugar, la información sobre los genomas de otras especies de plantas en ese momento, fueron también muy útiles para la identificación de enzimas del metabolismo ya que están muy conservadas. Finalmente, la mayoría de las identificaciones se realizaron utilizando espectros de MS/MS.

Table 3. List of 2-DE-based identified *P. vulgaris* seed proteins by MALDI-TOF and MALDI-TOF/TOF. Source: De la Fuente et al., 2011.

| Spot No. | Protein identity | Mascot score | Match/ % Cov. ^a | M_r (kDa) | | pI | |
|------------------------|--|--------------|-------------------------------|-------------|------|-------|------|
| | | | | Theo. | Exp. | Theo. | Exp. |
| Storage protein | | | | | | | |
| 62 | Phaseolin, α -type (<i>P. vulgaris</i>) | 563 | 13/33 | 49.2 | 18.3 | 5.3 | 4.9 |
| 66 | Phaseolin - (<i>P. vulgaris</i>) | 174 | 9/23 | 48.4 | 26.4 | 5.4 | 4.6 |
| 81 | Phaseolin, α type - (<i>P. vulgaris</i>) | 556 | 20/60 | 49.2 | 39.4 | 5.3 | 5.0 |
| 83 | Phaseolin - (<i>P. vulgaris</i>) | 251 | 15/43 | 47.5 | 51.1 | 5.0 | 4.9 |
| 105 | Phaseolin - (<i>P. vulgaris</i>) | 60 | 5/14 | 47.5 | 19.6 | 5.4 | 5.3 |
| 110 | Phaseolin, chain A - (<i>P. vulgaris</i>) | 97 | 9/32 | 45.0 | 17.7 | 5.2 | 5.0 |
| 111 | Phaseolin, chain A - (<i>P. vulgaris</i>) | 420 | 12/30 | 45.0 | 18.5 | 5.2 | 5.0 |
| 224 | Phaseolin - (<i>P. vulgaris</i>) | 388 | 12/36 | 48.4 | 20.7 | 5.3 | 5.7 |
| 296 | Phaseolin - (<i>P. vulgaris</i>) | 444 | 12/29 | 47.5 | 25.3 | 5.4 | 6.1 |
| 311 | Phaseolin - (<i>P. vulgaris</i>) | 335 | 12/34 | 47.5 | 30.4 | 5.4 | 5.9 |
| 354 | Phaseolin, α -type - (<i>P. vulgaris</i>) | 391 | 16/36 | 49.2 | 21.1 | 5.3 | 6.3 |
| 355 | Phaseolin, α type - (<i>P. vulgaris</i>) | 302 | 12/34 | 49.2 | 23.8 | 5.3 | 6.3 |

| Spot No. | Protein identity | Mascot score | Match/ % Cov. ^a | <i>M_r</i> (kDa) | | <i>pI</i> | |
|--------------------------------|---|--------------|-------------------------------|----------------------------|------|-----------|------|
| | | | | Theo. | Exp. | Theo. | Exp. |
| 356 | Phaseolin – (<i>P. vulgaris</i>) | 509 | 12/34 | 47.5 | 19.9 | 5.4 | 6.1 |
| 368 | Phaseolin, α-type - (<i>P. vulgaris</i>) | 348 | 15/31 | 49.2 | 26.0 | 5.3 | 6.1 |
| 496 | Phaseolin – (<i>P. vulgaris</i>) | 69 | 5/14 | 48.8 | 25.6 | 5.4 | 6.5 |
| Carbohydrate metabolism | | | | | | | |
| 121 | α-1,4 glucan phosphorylase L isozyme, chloroplastic/amyloplastic — (<i>Triticum aestivum</i>) | 81 | 6/13 | 52.1 | 29.7 | 5.0 | 5.3 |
| 183 | Glucose and ribitol dehydrogenase – (<i>Daucus carota</i>) | 64 | 3/8 | 31.7 | 18.6 | 6.3 | 5.6 |
| 325 | Malate dehydrogenase [NADP], chloroplastic– (<i>Zea mays</i>) | 162 | 16/44 | 35.9 | 39.8 | 5.8 | 6.0 |
| 404 | Ribulose 1,5-bisphosphate carboxylase-oxygenase, large subunit – (<i>Callaeum septentrionale</i>) | 754 | 25/53 | 52.1 | 64.1 | 6.1 | 6.2 |
| 460 | Ribulose 1,5-bisphosphate carboxylase-oxygenase, large subunit – (<i>Piptanthus nepalensis</i>) | 136 | 9/24 | 50.8 | 60.5 | 6.3 | 6.3 |
| 472 | Ribulose 1,5-bisphosphate carboxylase-oxygenase, large subunit – (<i>Dipogon lignosus</i>) | 676 | 22/53 | 49.9 | 63.4 | 6.3 | 6.1 |
| 426 | Granule-bound starch synthase I – (<i>P. vulgaris</i>) | 316 | 5/11 | 67.6 | 18.6 | 6.4 | 6.5 |
| 427 | Granule-bound starch synthase I – (<i>P. vulgaris</i>) | 414 | 9/19 | 67.6 | 18.0 | 6.4 | 6.5 |
| 457 | Phosphoglycerate kinase - (<i>Arabidopsis thaliana</i>) | 63 | 4/13 | 42.2 | 43.5 | 5.5 | 6.3 |
| Defence | | | | | | | |
| 5 | α -amylase inhibitor β subunit – (<i>P. vulgaris</i>) | 359 | 6/69 | 15.4 | 15.1 | 4.7 | 4.4 |
| 26 | α -amylase inhibitor β subunit – (<i>P. vulgaris</i>) | 134 | 6/68 | 15.4 | 27.2 | 4.7 | 4.4 |
| 83 | α -amylase inhibitor like protein – (<i>P. vulgaris</i>) | 95 | 3/17 | 28.9 | 51.1 | 4.95 | 4.9 |
| 26 | Phytohemagglutinin - (<i>P. vulgaris</i>) | 443 | 11/53 | 29.7 | 27.2 | 4.8 | 4.4 |

RESULTADOS Y DISCUSIÓN

| Spot No. | Protein identity | Mascot score | Match/ % Cov. ^a | M_r (kDa) | | pI | |
|-------------------------------|---|--------------|-------------------------------|-------------|------|-------|------|
| | | | | Theo. | Exp. | Theo. | Exp. |
| 66 | Phytohemagglutinin - (<i>P. vulgaris</i>) | 104 | 9/43 | 29.8 | 26.4 | 5.0 | 4.6 |
| 27 | Lectin - (<i>P. vulgaris</i>) | 462 | 8/42 | 29.6 | 26.8 | 4.8 | 4.6 |
| 66 | Lectin - (<i>P. vulgaris</i>) | 100 | 4/20 | 29.6 | 26.4 | 4.8 | 4.6 |
| Stress response | | | | | | | |
| 360 | Heat-shock cognate - (<i>D. carota</i>) | 136 | 4/17 | 18.5 | 18.2 | 10.6 | 6.1 |
| 523 | Pv42p - (<i>P. vulgaris</i>) | 105 | 6/21 | 41.6 | 43.1 | 6.5 | 6.5 |
| 486 | Superoxide dismutase [Mn] - (<i>Pisum sativum</i>) | 98 | 4/20 | 26.7 | 25.6 | 7.2 | 5.7 |
| 554 | Dehydrin - (<i>Vigna unguiculata</i>) | 92 | 1/10 | 26.5 | 27.3 | 6.0 | 6.9 |
| Detoxification | | | | | | | |
| 43 | Putative glutathione S-transferase - (<i>Phaseolus acutifolius</i>) | 86 | 2/13 | 24.8 | 38.0 | 5.6 | 6.0 |
| 486 | NAD-dependent formate dehydrogenase - (<i>Oryza sativa</i>) | 87 | 4/10 | 41.4 | 25.6 | 6.9 | 6.5 |
| Growth and development | | | | | | | |
| 441 | IAA-protein conjugate - (<i>P. vulgaris</i>) | 413 | 23/53 | 35.5 | 36.4 | 6.2 | 6.4 |
| Protein transport | | | | | | | |
| 500 | GTP-binding nuclear protein Ran1A - (<i>Lotus japonicus</i>) | 260 | 10/52 | 24.1 | 27.4 | 6.7 | 6.7 |
| Nitrogen metabolism | | | | | | | |
| 524 | Glutamine synthetase β 2, cytosolic - (<i>P. vulgaris</i>) | 57 | 3/42 | 39.3 | 41.1 | 6.1 | 6.4 |

^a % Cov.: percentage of the protein sequence that corresponds to matched peptides; Theo.: theoretical value; Exp.: experimental value; Spot numbers exhibiting statistically significant quantitative differences among the different extraction protocols assayed.

En las identificaciones por MS un número de *spots* proteicos mostraron notables discrepancias entre los valores teóricos y experimentales en M_r y/o pI . Además diferentes *spots* fueron identificados como la misma proteína. Esas discrepancias en M_r y/o pI podrían corresponder a distintas isoformas producidas por modificaciones post-traduccionales, corte y empalme alternativo y a la presencia de familias multigénicas, así como a distintos fragmentos de la misma proteína. Para la faseolina y el inhibidor de la α -amilasa se aprecia que la proteólisis puede ser el factor principal que explique la

discordancia entre los valores teóricos y experimentales, de M_r y pI . La faseolina está codificada por una pequeña familia génica de 6-10 genes constituidas por dos subfamilias (α y β) que codifican dos tipos de polipéptido α y β . Estos polipéptidos se resolvieron en 2-DE dentro de un rango de M_r de 52,0-45,0 kDa y pI próximos a pH 5 (Slightom et al., 1985; Sturm et al., 1987), mientras que en nuestro estudio otras quince faseolinas (*spots* 62, 66, 81, 83, 105, 110, 111, 224, 296, 311, 354, 355, 356, 368, 496) fueron identificadas en otras zonas del gel, en el rango de M_r desde 51,1 hasta 17,7 kDa y en el de pI desde 6,5 a 4.6. Se sabe que cambios sustanciales en la estructura molecular de la faseolina no son detectables antes de los 7-10 días siguientes a la germinación de la semilla cuando la faseolina se degrada por la acción de enzimas proteolíticas dando lugar a fragmentos de 27-23 kDa (Boylan y Sussex, 1987). Por el contrario, nuestro estudio sugiere que existe una degradación parcial de la faseolina, anterior a la germinación de la semilla. En el caso del inhibidor de α -amilasa que se considera como una forma modificada de la faseolina. Esta se sintetiza a partir de un polipéptido precursor de aproximadamente 40 kDa y posteriormente se glicosila. Posteriormente mediante un proceso proteolítico se generan varios polipéptidos de M_r en el rango de 15 a 18 kDa (Ceriotti et al., 1989; Nasi et al., 2009). En nuestro estudio se han identificado *spots* como inhibidor de α -amilasa en el rango entre 27,2 y 15,1 kDa, lo que se podría corresponder a fragmentos de polipéptidos derivados de la proteólisis del polipéptido precursor. En general, la elevada reproducibilidad de los *spots* en las repeticiones y protocolos de extracción, junto con el hecho de que el método del fenol minimiza la proteólisis durante la extracción, hace que esta fragmentación in vitro sea poco probable.

Las proteínas identificadas por MS fueron clasificadas en diferentes grupos que se corresponden a su supuesta función biológica: proteínas de almacenamiento, metabolismo de los carbohidratos, defensa, respuesta al estrés, desintoxicación, crecimiento y desarrollo, proteínas de transporte y metabolismo del nitrógeno. La mayoría de las proteínas identificadas correspondían a proteínas de almacenamiento (37,5%), metabolismo de hidratos de carbono (22,5%), defensa (17,5%) y respuesta al estrés (10,0%). Sin embargo, con cada método de extracción obtuvimos un perfil

diferenciado de la semilla de judía. El protocolo TCA-acetona y Clean-Up extrajeron la mayor cantidad de proteínas de reserva y defensa mientras que el método del fenol extrajo una mayor cantidad de proteínas relacionadas con el metabolismo de los carbohidratos. También se detectaron diferencias notables entre los protocolos cuando las proteínas se clasificaron en glicosiladas (faseolinas y lectinas) y no-glicosiladas. Así, la proporción de proteínas glicosiladas / no glicosiladas, para cada método de extracción fue de 15/3, 1/8 y 13/1 para TCA-acetona, fenol y Clean-Up, respectivamente. Se deduce que TCA-acetona y Clean-Up proporcionaron extractos de semillas más ricos en glicoproteínas en comparación con fenol (test de Fisher, $P < 0.001$). Por el contrario, estudios anteriores con tomate y especies de frutas mostraron que el método basado en el fenol extrajo un mayor número de glicoproteínas que los métodos basados en TCA-acetona (Saravanan y Rose, 2004; Zheng et al., 2007). Sin embargo, nuestros resultados deben ser interpretados con cautela porque los valores experimentales de pI y M_r fueron generalmente más bajos para proteínas glicosiladas que para las no glicosiladas (test U de Mann-Whitney, $P < 0,05$ y $P < 0,001$ para el pI y M_r , respectivamente) y las observaciones antes mencionadas sugieren que el método de fenol fue menos eficiente para extraer las proteínas ácidas de bajo peso molecular. Por lo tanto, diferencias en pI y la M_r entre proteínas glicosiladas y no glicosiladas pueden ser subyacentes a los patrones de extracción, es decir dependiente de los protocolos. Además, muchas glicoproteínas identificadas (64%) eran en realidad productos de degradación de la faseolina con niveles desconocidos de glicosilación. En este sentido, se conoce que la faseolina tiene tres cadenas distintas de oligosacáridos unidos con enlaces n-glucosídicos a los residuos de asparagina y dos lugares de glicosilación en la posición 252 y uno en la posición 341 (la numeración empieza con la metionina N-terminal del péptido señal) (Sturm et al., 1987). Sin embargo, análisis con MS/MS revelaron que la mayoría (64,3%) de los fragmentos de faseolina (*spots* 105, 224, 296, 311, 354, 355, 356, 368, y 496) identificados en nuestro estudio se corresponden con la secuencia C-terminal (posiciones desde 257 hasta 421) que no contiene dos posiciones canónicas de glicosilación. Por tanto, los patrones de extracción diferencial de las glicoproteínas de semillas de judía detectados entre los protocolos podrían ser más aparentes que reales.

3.1.2. Artículo original: 2-DE-based proteomic analysis of common bean (*Phaseolus vulgaris* L.) seeds



<http://www.sciencedirect.com/science/article/pii/S1874391910002927>





3.2. Caracterización de la diversidad de la proteína faseolina de la judía común (*P. vulgaris* L.) basada en la electroforesis bidimensional y en la espectrometría de masas.

3.2.1. Optimización de los perfiles 2-DE de la faseolina

En primer lugar se llevó a cabo un estudio preliminar de optimización de la 2-DE para la caracterización de los perfiles de la faseolina en semilla a partir de los cultivares ICA Pijao (acervo mesoamericano) y Calima (acervo andino). Estos cultivares han sido utilizados frecuentemente como controles genéticos en diferentes estudios sobre la diversidad de la judía común (Kwak y Gepts, 2009). Para este estudio se utilizó el método del fenol para la extracción de proteína total de semilla de judía dado que proporciona geles 2-DE de alta resolución como se había establecido previamente (De la Fuente et al., 2011).

La Figura 9 muestra una imagen representativa de un gel 2-DE teñido con Sypro Ruby, obtenido a partir de 250 µg de proteína de semilla del cultivar ICA Pijao cargados en tiras IPG de 24 cm con un gradiente de pH 4-7 (primera dimensión) y en gel de poliacrilamida al 15% (segunda dimensión). En el gel se observa una zona muy saturada en el rango de pH 5,6 a 5,8 y con una masa molecular entre 45 y 51 kDa, descrita previamente como una zona que contiene polipéptidos de faseolina (Brown et al., 1981b; Bliss y Brown, 1982). Este extremo se confirmó en el presente estudio mediante análisis por Western blot (datos no mostrados).

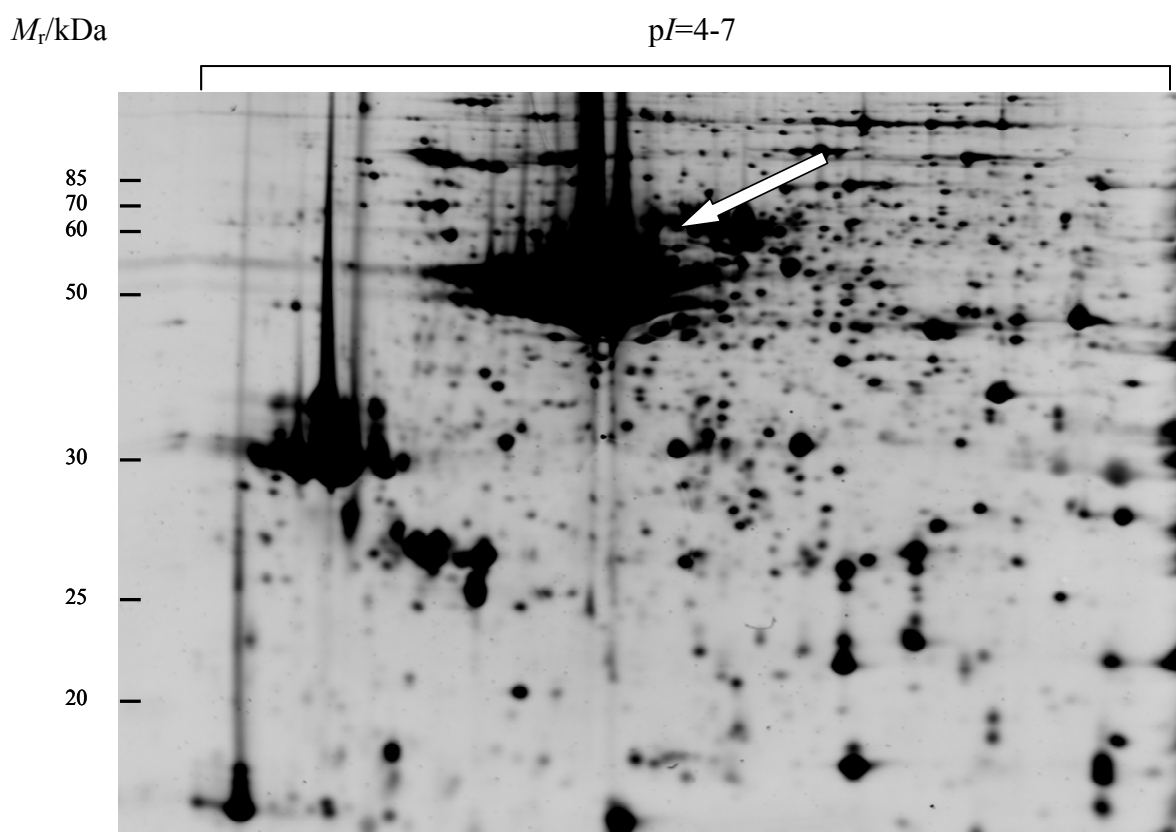


Figure 9. 2-DE gel image of the seed proteome of *P. vulgaris* for ICA Pijao cultivar. The arrow indicates the gel zone where the phaseolin protein is located. 2-DE was performed from 250 μ g of total protein loaded on 24-cm long IPG strips, pH=4-7, and 15 % (by mass) polyacrylamide gels. Source: De la Fuente et al., 2012.

Una vez localizada la zona del gel que contiene la proteína faseolina, se realizaron una serie de experimentos optimizando las condiciones de la 2-DE para obtener patrones de faseolina de alta resolución. En primer lugar, se redujo la concentración de proteína de semilla utilizada en la electroforesis, en distintas proporciones, para tratar de eliminar el fenómeno de saturación en la zona del gel de interés. En la Figura 10 se muestran los patrones 2-DE de la faseolina que surgieron con tan solo 50 μ g de proteína total de semilla de los cultivares ICA Pijao, Calima y una mezcla homogénea de ambos. La muestra compuesta de ambos tipos de cultivares es de gran utilidad para evaluar las posiciones relativas de los *spots* 2-DE de la faseolina, en los cultivares ICA Pijao, Calima, respecto al punto isoeléctrico y a la masa molecular.

Además, se muestra el resultado obtenido con la aplicación de la electroforesis de una única dimensión (1-DE). Se puede llegar a establecer una comparación directa entre los patrones de faseolina 2-DE y 1-DE porque la etapa SDS-PAGE se realizó a partir de tiras IPG que habían experimentado el isoelectroenfoque para obtener los patrones 2-DE y, simultáneamente, se introdujeron 50 µg de proteína total de semilla en un pocillo ubicado en un extremo del mismo gel para obtener los patrones de faseolina 1-DE bajo las mismas condiciones electroforéticas.

La Figura 10 revela que con 50 µg de proteína se elimina como se pretendía el fenómeno de saturación en la zona del gel 2-DE donde se ubica la proteína faseolina y emergen patrones claros de *spots* que continen los diferentes polipéptidos que componen la faseolina. El contraste entre los perfiles 2-DE y 1-DE ilustra la gran diferencia de información que suministran ambas técnicas sobre la diversidad polipeptídica de la faseolina. Cabe destacar que los perfiles de faseolina en geles monodimensionales obtenidos para los cultivares ICA Pijao y Calima coincidieron con los descritos previamente para cultivares S y C, los cuales consisten en patrones de dos y tres bandas principales, respectivamente (Brown et al., 1981b; Bliss y Brown, 1982; Montoya et al., 2008a y 2008b).

Finalmente, patrones de faseolina de más alta resolución se obtuvieron utilizando tiras IPG con un rango de pH (4,7–5,9) más estrecho y reduciendo la concentración de poliacrilamida al 10%, lo que contribuye a un menor solapamiento de los *spots* en la primera y segunda dimensión, respectivamente.

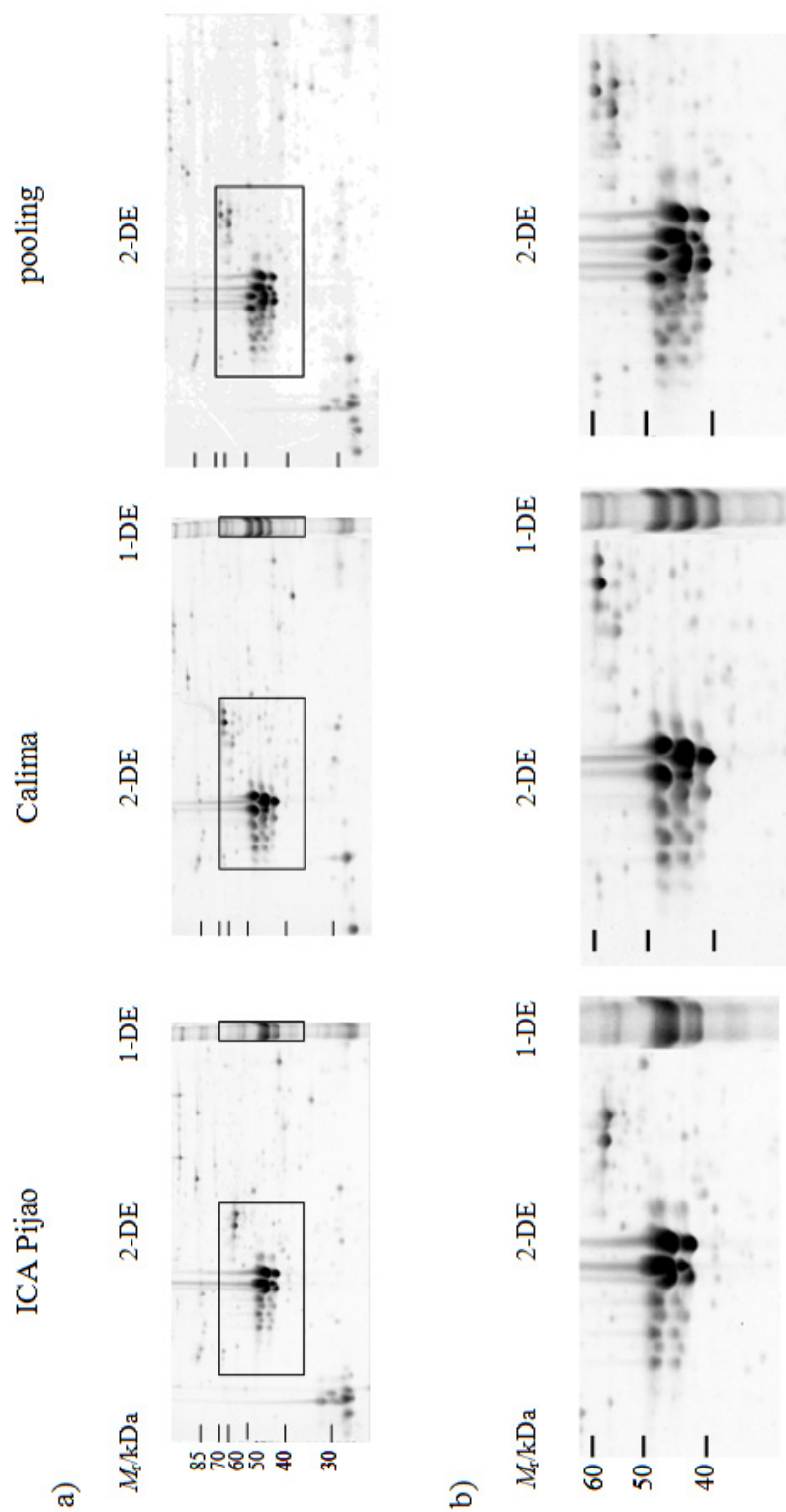


Figure 10. 2-DE phaseolin profiles of cultivars ICA Pijao (left) and Calima (centre) and both cultivars pooled (right). 1-DE phaseolin profiles of these two cultivars are also shown. Gel regions containing phaseolin spots and protein bands are enclosed in rectangles in (a) and enlarged in (b). 2-DE was performed from 50 μ g of total protein loaded on 24-cm long IPG strips, pH=4-7, and 15 % (by mass) polyacrylamide gels. Source: De la Fuente et al., 2012.

3.2.2. Descripción de los patrones 2-DE de la faseolina

La Figura 11 muestra los patrones de faseolina de alta resolución obtenidos mediante los protocolos de 2-DE optimizados para un total de seis variedades domesticadas de judía común: S (Sanilac) y B (Boyacá), del acervo mesoamericano; y T (Tendergreen), C (Contender), H (Huevo de Huanchaco) y A (Ayacucho), del acervo andino. Los geles de elevada resolución con *spots* definidos y muy reproducibles, permiten la descripción precisa y detallada de los patrones de faseolina 2-DE. En primer lugar, se puede observar que los patrones de faseolina se organizan como cadenas de *spots* horizontales que varían en función del pI y de la M_r . Por otro lado existe una disposición vertical de los *spots* en parejas que se corresponden con isoformas de distinta M_r . Las parejas de isoformas numeraron consecutivamente en función de su pI , desde posiciones más ácidas a las más alcalinas, de acuerdo con su ubicación en el gel máster y después de contrastar todos los *spots* en los diferentes cultivares con el PDQuest. Los *isospots* de cada pareja localizados arriba/abajo se denominaron como 1'/1, 2'/2, etc. En segundo lugar, se observa que los patrones de faseolina para los distintos cultivares contienen un número variable de *spots* intensos (por ejemplo, el *spot* 20' / 20 en la pareja del cultivar S) y otros más débiles (por ejemplo, el *spot* 5' / 5 en la pareja del cultivar S). Los *spots* intensos y débiles en la mayoría de los cultivares se denominaron como *major spots* y *minor spots*, respectivamente. Un análisis más detallado de las imágenes revela que el patrón de los *major spots* exhibe cierto grado de repetición dentro de cada variedad. Por ejemplo, los pares de *isospots* 7' / 7, 9' / 9, 11' / 11 'y 12' / 12, así como las pares 18' / 18, 19' / 19, 21' / 21 y 23' / 23 parecen ser repeticiones de un patrón principal en el cultivar H.

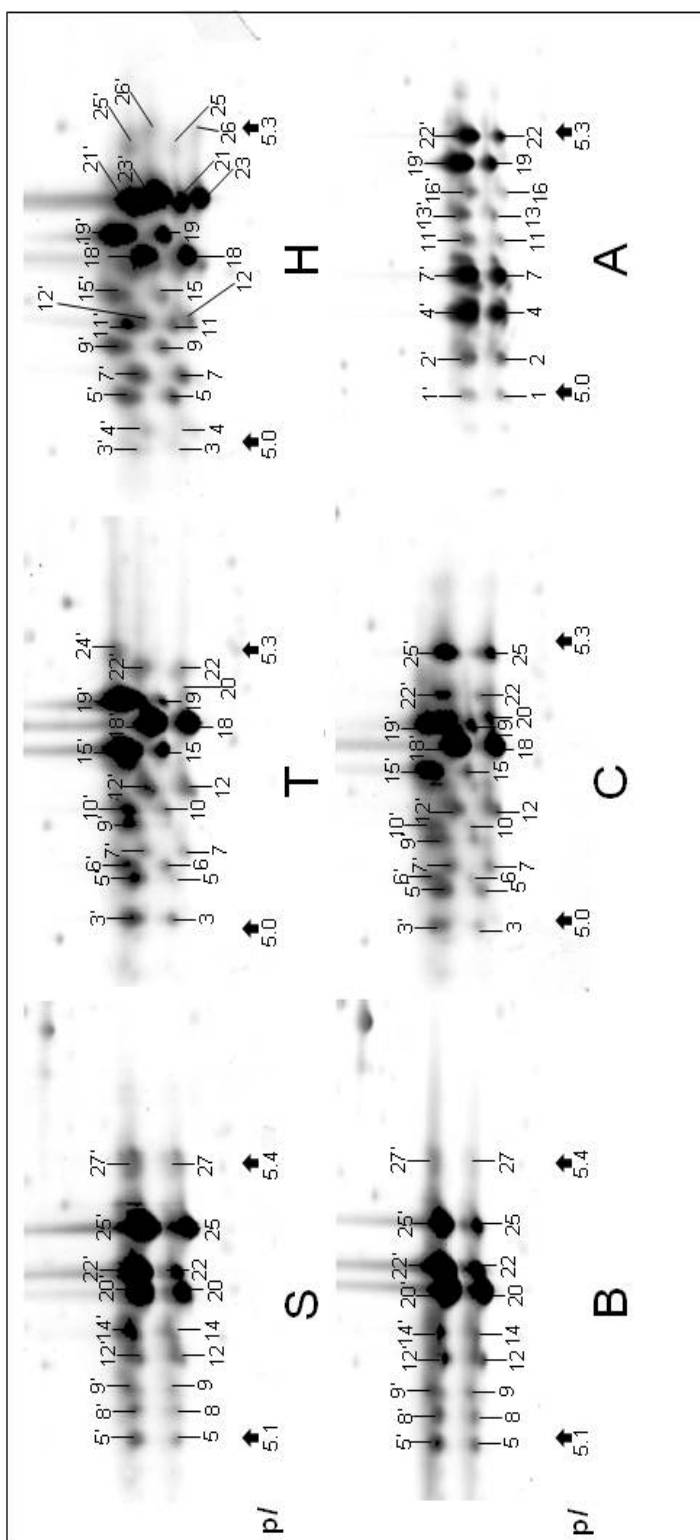


Figure 11. 2-DE spot patterns of different types of phaseolin (S, T, H, B, C and A). Spots were numbered after matching all spots across cultivars by PDQuest. Up and down isospots of different M_r and similar pI were denoted with and without apostrophe, respectively. 2-DE was performed from 50 μg of total protein loaded on 24-cm long IPG strips, pH=4.7-5.9, and 10 % (by mass) polyacrylamide gels. Source: De la Fuente et al., 2012.

En general, los perfiles 2-DE de faseolina observados en nuestro estudio exhiben grandes diferencias con los de estudios previos (Bollini y Vitale, 1981; Brown et al., 1981a, 1981b; Lioi y Bollini, 1984; Gepts y Bliss, 1986; Gepts et al., 1986; Lioi, 1989). En particular, se ha identificado por primera vez una importante cantidad adicional de *minor spots* situados esencialmente en las zonas menos básicas de los patrones 2-DE de faseolina. Esta nueva y mayor información sobre los perfiles de faseolina observada en nuestro estudio se puede atribuir a un número diverso de factores: como la mejora en el procedimiento de extracción proteica, el uso de tiras IPG, así como la mayor ampliación del rango de pH (4,7–5,9) en comparación con el usado en trabajos anteriores (pH 5,2–5,7) (Bollini y Vitale, 1981; Lioi y Bollini, 1984).

3.2.3. Diversidad de los patrones 2-DE de la faseolina

La diferenciación de la faseolina entre distintos cultivares de judía se puede detectar y cuantificar mediante la comparación de sus mapas de *spots* en geles 2-DE de alta resolución. Un estudio preliminar de la diversidad de la faseolina basada en estos patrones 2-DE se realizó a partir de los cultivares S y B (acervo mesoamericano) y C, T, H y A (acervo andino). Se han podido identificar un total de 51 *spots* diferentes entre estos cultivares, con una un número mayoritario de ellos (35) ocupando la misma posición relativa en el gel máster. El número de *major spots*, *minor spots* y *spots* únicos de faseolina detectados para cada cultivar se muestra en la Tabla 4. Se observa que el número de *spots* detectado fue elevado dentro de cada cultivar y bastante variable entre los distintos cultivares, oscilando entre 18 *spots* y 28 *spots*. El número de *spots* observado en nuestros patrones de faseolina es muy superior al de 6 *spots* encontrados en estudios previos (Lioi y Bollini, 1984; Brown et al., 1981b).

Table 4. Number of major, minor and unique 2-DE spots across different phaseolin types. Source: De la Fuente et al., 2012.

| Phaseolin type | Major spots | Minor spots | Unique spots | Percentage of unique spots* |
|----------------|-------------|-------------|--------------|-----------------------------|
| S | 6 | 12 | 0 | 0 |
| B | 6 | 12 | 0 | 0 |
| C | 13 | 11 | 0 | 0 |
| T | 11 | 12 | 1 | 2 |
| H | 16 | 12 | 5 | 10 |
| A | 8 | 10 | 8 | 16 |

*Referred to the total amount of spots detected with PDQuest software

Por lo tanto, la información sobre la cantidad de polipéptidos que conforman cada uno de los patrones de faseolina se incrementa sustancialmente. Así el número de *major spots* en los cultivares osciló entre 6 y 16, mientras que el de *minor spots* sólo varió de 10 a 12. El porcentaje de *spots* únicos y específicos de cada cultivar también varió mucho entre cultivares (desde 0 hasta 16%), siendo los cultivares H y A los que muestran los porcentajes más elevados de *spots* únicos (10 y 16%, respectivamente).

El grado de similitud de los patrones de faseolina entre dos cultivares dados se estimó a partir del porcentaje de *spots* compartidos mediante el coeficiente F:

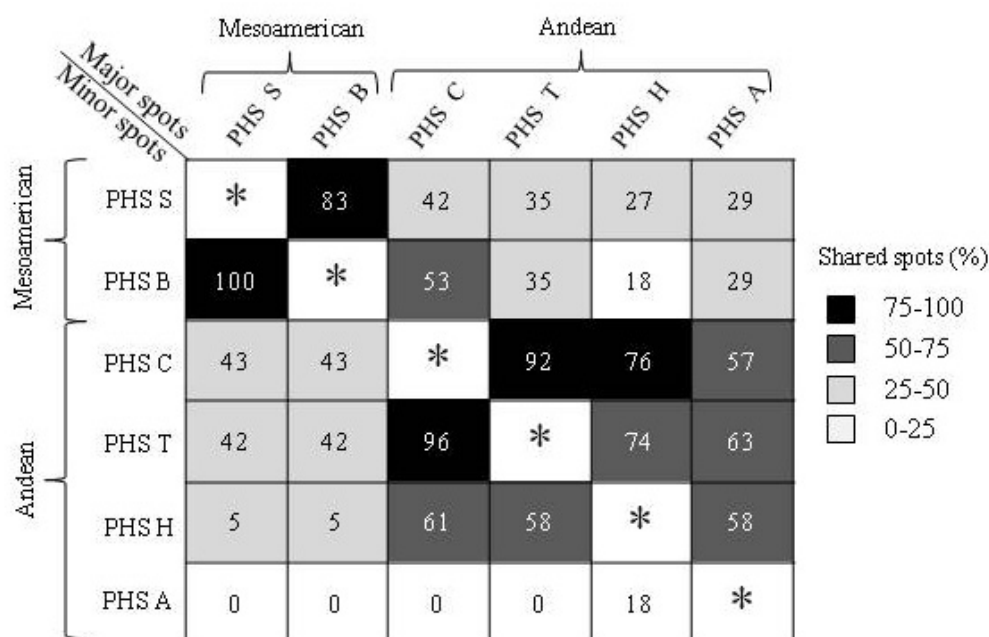
$$F=2n_{xy}/(n_x+n_y)$$

donde n_x y n_y son el número total de *spots* en los cultivares x e y, respectivamente, y n_{xy} es el número de *spots* compartidos por x e y (Figura 12a) (Biron et al., 2006). La distancia proteómica (D) entre dos patrones de faseolina se obtuvo a partir de la expresión 1-F (Biron et al., 2006). El estudio de la distancia proteómica entre cultivares se realizó para *major spots* y *minor spots*, separadamente. A partir de las distancias proteómicas, se construyó un mapa de las relaciones existentes entre los diferentes patrones de faseolina mediante la técnica del escalado multidimensional no-

métrico (MDS, *nonmetric multidimensional scaling*) (Manly, 1986) (Figura 12b). El análisis MDS se llevó a cabo mediante el programa XLSTAT v. 1.01 (Addinsoft, Andemach, Alemania). El análisis MDS reveló una clara diferenciación de los tipos de faseolina pertenecientes a los acervos mesoamericano y andino para los *major spots* y los *minor spots*. Sin embargo, una diferenciación más clara entre los patrones de faseolina de los acervos andino y mesoamericano se obtuvo a partir de los *major spots*. El mapa producido por MDS también mostró que el patrón de la faseolina tipo A era el más diferenciado. Se ha planteado la hipótesis de que la faseolina del tipo C pudiera ser el resultado de una combinación de los tipos T y S, en base a los patrones 1-DE y 2-DE, porque muestra polipéptidos comunes a ambos tipos de faseolina (Brown et al., 1981b). Por este motivo se podría decir que la faseolina C está en una posición intermedia entre el acervo mesoamericano y el andino. Este razonamiento contradice otros estudios posteriores que sitúan el cultivar C dentro del acervo andino (Gepts et al., 1986) basándose en datos del fenotipo de la semilla y de la arquitectura de la planta de esta variedad que presentan características claramente presentes en otras variedades del tipo andino. Los patrones 2-DE analizados en este trabajo sugieren que el patrón de la faseolina C pertenece al acervo andino y está estrechamente relacionado con el tipo T (porcentaje de *spots* compartidos superior al 80%).

Por otra parte, los perfiles de las faseolinas H y A son los que se diferencian en mayor grado del resto, lo que está de acuerdo con estudios iniciales de 2-DE (Gepts y Bliss, 1986).

a)



b)

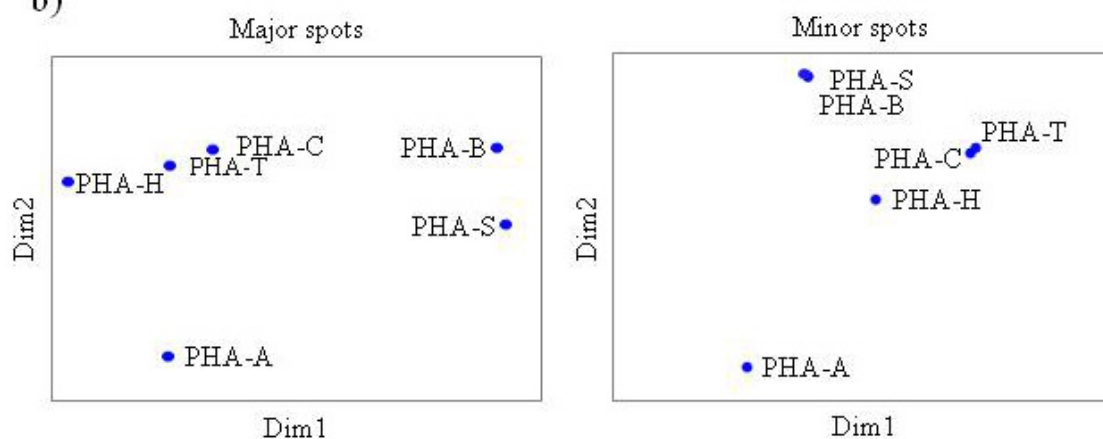


Figure 12. Percentage of major (upper) and minor (bottom) spots shared between phaseolin cultivars as measured by F-index (a); non-metrical multidimensional scaling (MDS) from proteomic distance (D=1-F) between phaseolin types (b). Source: De la Fuente et al., 2012.

3.2.4. Interpretación molecular de los patrones 2-DE de la faseolina

Los patrones 2-DE descritos anteriormente pueden ser interpretados molecularmente teniendo en cuenta la información disponible acerca de la composición de polipéptidos de faseolina y de sus modificaciones co- y post-traduccionales. La presencia y distribución de los dos tipos de polipéptidos (α y β) que componen la faseolina, a lo largo de las filas de *spots*, se determinó en cultivares S (acervo mesoamericano) y T (acervo andino), por MALDI-TOF y MALDI-TOF / TOF MS. Los 22 *spots* seleccionados para el análisis por MS se identificaron inequívocamente como polipéptidos α o β a través de PMF, datos espectrales de MS/MS y de sus características distintivas en la secuencia de nucleótidos. Los resultados de las identificaciones se muestran en la Tabla 5.

Los datos obtenidos revelaron varias características destacadas de los patrones de faseolina. En primer lugar, cada par de *isospots* individuales contenían sólo un tipo de polipéptido (α o β) de faseolina. En segundo lugar, se encontró que aquellos pares de *isospots* con mayor M_r (por ejemplo, el par 22' / 22 en el cultivar S) se correspondían con polipéptidos de tipo α , mientras que los pares con menor M_r (por ejemplo, el par 20' / 20 en el cultivar S) eran polipéptidos de tipo β . La mayor M_r en los polipéptidos de tipo α que en los de tipo β es consistente con la conocida divergencia genética entre sus regiones codificadoras. Así en los genes α existen diferentes repeticiones de nucleótidos directas (repetición de 15 pb en el cuarto exón y repetición de 27 pb en el sexto exón) que los hacen ser de un mayor tamaño que los genes β (Slightom et al., 1983, 1985; Anthony et al., 1990; Kami y Gepts, 1994; Kami et al., 1995). La observación de una mayor abundancia de polipéptidos de tipo α en los patrones de faseolina (Figura 11) también está de acuerdo con el hecho de que el número de sustituciones de aminoácidos parece ser mayor en los polipéptidos de tipo α que en polipéptidos de tipo β (Slightom et al., 1983, 1985).

RESULTADOS Y DISCUSIÓN

Table 5. Identification of the type of polypeptide (α or β) in 2-DE phaseolin spot patterns (cultivars S and T) by MALDI-TOF and MALDI-TOF/TOF MS*. Source: De la Fuente et al., 2012.

| Cultivar S | | Cultivar T | |
|------------|--------------------------------|------------|--------------------------------|
| No. spot | Phaseolin polypeptide identity | No. spot | Phaseolin polypeptide identity |
| 5' | α | 10' | α |
| 5 | α | 10 | α |
| 14' | α | 15' | α |
| 14 | α | 15 | α |
| 20' | β | 18' | β |
| 20 | β | 18 | β |
| 22' | α | 19' | α |
| 22 | α | 19 | α |
| 25' | β | 22' | β |
| 25 | β | 22 | β |
| 27' | α | | |
| 27 | α | | |

* Averages (\pm SE) of Mascot scores and percentages of the protein sequence for matched peptides were 833 ± 24 and 60.3 ± 1.6 , respectively.

La diferencia de masa molecular entre *isospots* de cada uno de los pares puede ser interpretada por el hecho ya conocido de que los polipéptidos α o β experimentan uno o dos glicosilaciones (Bollini et al., 1983). Con el fin de confirmar este supuesto se procedió a incubar extractos de proteínas de los cultivares estudiados con la enzima PNGasa F. Los datos obtenidos en la Figura 13 revelaron que los extractos de proteínas experimentaban desglucosilación parcial y total con la PNGasa F. Cabe señalar que las glicoproteínas no sólo experimentan cambios en su M_r sino también en su pI , ya que los residuos de asparagina se convierten en ácido aspártico.

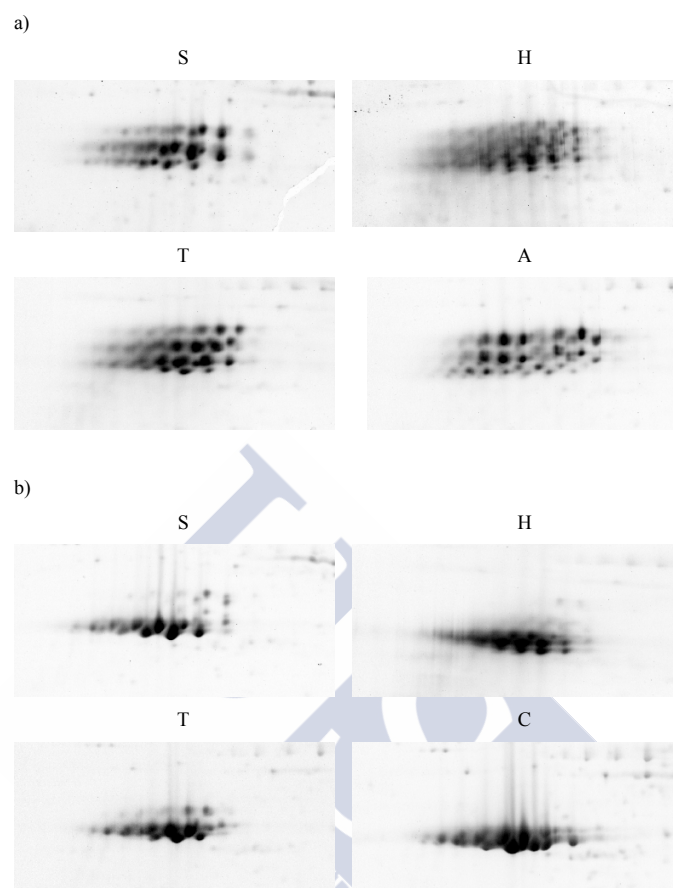


Figure 13. 2-DE spot patterns of phaseolin types partially (a) and totally (b) deglycosylated with PNGase F. 2-DE was performed from 50 μg of total protein loaded on 24-cm long IPG strips, pH=4.7-5.9, and 10 % (by mass) polyacrylamide gels. Source: De la Fuente et al., 2012.

Se observó que los polipéptidos de la faseolina fueron en efecto sensibles a la digestión parcial y total con la enzima PNGasa F (Figura 13). En primer lugar, los patrones de la faseolina desglucosilada parcialmente (Figura. 13a) mostraron cadenas de *spots* adicionales de una menor M_r en comparación con los patrones de la faseolina no tratada con la enzima, lo que indica que los polipéptidos de la faseolina contienen glicosilaciones. En segundo lugar, los patrones de la faseolina desglucosilada totalmente (Figura 13b) revelaron la existencia de sólo dos cadenas de *spots* con diferente M_r , pero de mayor movilidad electroforética en comparación a los patrones de faseolina no

tratados con la enzima, lo que sugiere que cada par de *isospots* contiene polipéptidos (α o β) con un grado diferente de glicosilación. Por tanto, los *spots* superiores (por ejemplo 1', 2', etc) y los *spots* inferiores (por ejemplo, 1, 2, etc) en las parejas de *isospots* contienen polipéptidos (α o β) mono y diglicosilados, respectivamente. Además, como se muestra en la Figura 13, las isoformas diglicosiladas de cada par de *spots* parecen en general más intensas que las isoformas monoglicosiladas que aparecen abajo, lo que está de acuerdo con observaciones anteriores que muestran que los polipéptidos del tipo α o β están mono o diglicosilados, aunque en la mayoría de las ocasiones aparecen dos veces glicosilados (Bollini et al., 1983). Finalmente, los geles 2-DE muestran que los pares de *isospots* del tipo α y β se manifiestan en múltiples repeticiones con variaciones en *pI* (por ejemplo, 5'/5 and 8'/8 en el cultivar S). Estas variaciones en *pI* son probablemente el resultado de modificaciones post-traduccionales de los polipéptidos α o β de la faseolina, tales como fosforilaciones y acetilaciones (Yamagata et al., 2002; Zhu et al., 2005). Sin embargo, la posible existencia de estos tipos de modificaciones pos-traduccionales en los polipéptidos que componen la faseolina apenas ha sido considerada (Paaren et al., 1987; Lawrence et al., 1990),

3.2.5. Artículo original: In-depth characterization of the phaseolin protein diversity of common bean (*Phaseolus vulgaris* L.) based on two-dimensional electrophoresis and mass spectrometry



<http://www.ftb.com.hr/index.php/component/content/article/79-volume-50-issue-no-3/433-in-depth-characterization-of-the-phaseolin-protein-diversity-of-common-bean-phaseolus-vulgaris-l-based-on-two-dimensional-electrophoresis-and-mass-spectrometry>





3.3. Evidencia de fosforilación de la principal proteína de reserva en la judía común y su degradación dependiente de la fosforilación durante la germinación

3.2.1. Polipéptidos fosforilados de faseolina en semillas durmientes

Los patrones 2-DE de referencia y fosforilados de la proteína faseolina en semillas maduras y secas (cvs. S y T) después de la tinción con SYPRO Ruby y Pro-Q DPS, respectivamente, se muestran en la Figura 14 a y b. Se puede apreciar que se detectó señal de fluorescencia con Pro-Q DPS para todos los *spots* de los patrones de referencia de las faseolinas S y T. Los marcadores moleculares de fosforilación (Peppermint) validan la especificidad del reconocimiento de fosfoproteínas con Pro-Q DPS bajo nuestras condiciones experimentales. Así, las fosfoproteínas de Peppermint (ovoalbúmina y β -caseína) dieron señal fluorescente con Pro-Q DPS, mientras que las proteínas no fosforiladas (β -galactosidasa, albúmina de suero bovino, avidina y lisozima), no mostraron fluorescencia. Por tanto, se puede concluir que las semillas de judía de los cvs. S y T contienen isoformas fosforiladas de faseolina en estado de dormancia. Además, el hecho de que los fosfopolipéptidos de faseolina exhiban variaciones en *pI* indica que son polipéptidos diferencialmente fosforilados.

Un estudio intensivo de las isoformas fosforiladas de la faseolina se realizó aplicando distintos métodos de desfosforilación de tipo tanto enzimático (fosfatasa alcalina) como químico (HF y HF-P) capaces de eliminar los residuos fosfato de los fosfopolipéptidos.

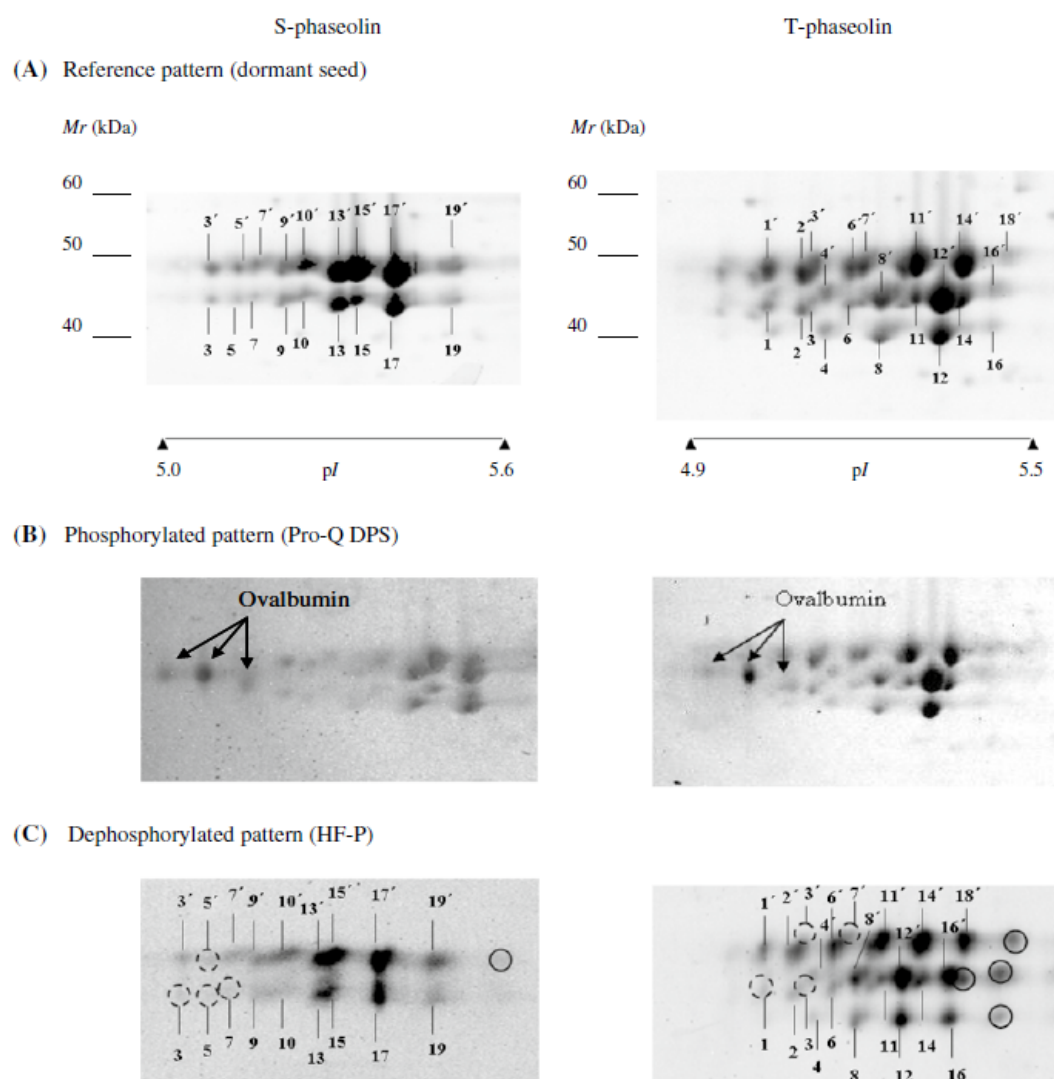


Figure 14. Identification of phosphorylated phaseolin isoforms in dry mature bean seeds of cvs. S and T. **(a)** 2-DE images of gel sections showing reference phaseolin patterns of cvs. S (left) and T (right) obtained from untreated samples of total seed protein. Gels were stained with SYPRO Ruby protein gel stain. **(b)** Same gels stained with the phosphoprotein-specific fluorescent dye Pro-Q DPS to detect phosphorylated phaseolin polypeptides. Arrows show the signal of fluorescence of the phosphoprotein marker ovalbumin. **(c)** Dephosphorylated phaseolin patterns after HF-P treatment followed by SYPRO Ruby gel staining. The dashed and closed circles represent missing and newly arisen phaseolin spots after dephosphorylation, respectively, as compared with reference patterns. Source: López Pedrouso et al. 2014a.

La comparación de los resultados obtenidos con los tres métodos de desfosforilación reflejó que el método de HF-P proporciona perfiles 2-DE más

informativos de la faseolina desfosforilada que los otros métodos (Figura 14c). En la Figura 14c se muestran los patrones 2-DE de la proteína faseolina desfosforilada en los cvs. S y T. Se puede observar que los *spots* altamente expresados (o *major spots*) en los patrones de referencia (i.e., los pares de *spots* 13'/13, 15'/15 y 17'/17 del cv. S y los pares 11'/11, 12'/12 y 14'/14 del cv. T) se disponen en las mismas posiciones en los patrones desfosforilados que en los de referencia, pero con un volumen relativo más bajo. De ello se deduce que los *major spots* en los patrones de faseolina de referencia contenían una mezcla de polipéptidos fosforilados y no fosforilados. La cantidad relativa de fosfopolipéptidos en cada *major spot*, se evaluó mediante el programa PDQuest, determinando la diferencia del volumen de un *spot* dado entre las muestras control y las tratadas con HF-P. La tasa de fosforilación relativa (*PR*) en cada *spot* se cuantificó mediante la fórmula: $PR = [(X - Y) / X] \times 100$, donde *X* y *Y* indican el volumen de un *spot* dado en las muestras control y tratadas con HF-P, respectivamente. Los valores medios de *PR* para los *major spots* en las distintas réplicas del cv. S oscilaron entre el 49 y el 81% con un valor promedio del $63 \pm 1,7\%$; mientras que para el cv. T, variaron entre el 13 y el 82% y la media fue de $46 \pm 3,9\%$. Además, se observa que los valores de *PR* para los *major spots* tienden a aumentar a medida que su *pI* disminuye. La relación entre los valores de *pI* y *PR* para los *major spots* en los cvs. S y T se muestra gráficamente en la Figura 15. Los valores de *pI* y *PR* muestran una relación negativa y estadísticamente significativa mediante el test de correlación no-paramétrica de Spearman en ambos cvs. (cv S: $r_s = -0.37$, $P < 0.05$, $n = 24$; cv. T: $r_s = -0.62$, $P < 0.01$, $n = 24$). Estos resultados demuestran que las semillas de judía en un estado de dormancia contienen una variedad de isoformas de la faseolina diferencialmente fosforiladas.

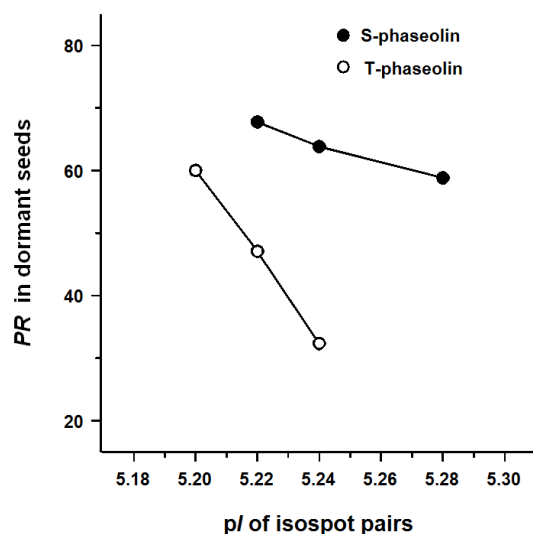


Figure 15. Plot of relative *PR* against *pI* for major isopot pairs of S- and T-phaseolin patterns in dry mature bean seeds. Lines connect mean values of *PR*-values across replicates for major isopot pairs. Source: López Pedrouso et al. 2014a.

La desfosforilación de proteínas debería provocar cambios predecibles en el *pI* y en la M_r de los *spots* (Zhu et al 2005; Irar et al 2006). Así, la desfosforilación afecta al *pI* de las proteínas porque se sustituyen los grupos fosfato cargados negativamente por grupos hidroxilo neutros. Como resultado, la desfosforilación generalmente induce un movimiento de los *spots* hacia zonas más básicas de los geles 2-DE y múltiples desfosforilaciones pueden causar un aumento significativo del *pI*. En cuanto a la M_r , la pérdida de grupos fosfato en residuos de fosfoserina, fosfotreonina y fosfotirosina se traduce en una disminución de la masa de 80 Da o múltiplos de este valor en el caso de múltiples desfosforilaciones. En consecuencia, nuestras observaciones muestran que algunos de los pares de *spots* débilmente teñidos o *minor spots* ubicados en posiciones más ácidas de los patrones de referencia (es decir, *spots* 3, 5', 5 y 7 en el cv S; y *spots* 1, 3' y 3 en cv. T) llegan a no detectarse después del tratamiento con HF-P (Figura 15c). Estos *minor spots* ausentes en los patrones desfosforilados probablemente contenían polipéptidos fosforilados de faseolina que cambiaron a posiciones más básicas en los geles después de la desfosforilación. De hecho, se encontraron *spots* nuevos en los patrones desfosforilados (es decir, un *spot* en el cv. S y tres *spots* en el cv. T) que

ocupaban más básicas en los geles 2-DE. Por otra parte, se observa que los pares de *isospots* para el mismo tipo de polipéptido también aparecieron a una M_r más homogénea que en las muestras control debido a la pérdida de los grupos fosfato.

3.2.2. Cambios de fosforilación de la faseolina en semillas en germinación

Los posibles cambios temporales en los niveles de fosforilación de la faseolina se estudiaron en los cvs. S y T durante la transición de la semilla desde un estado de dormancia hasta el cuarto día de germinación (Figura 16 a-c). Se observó que los patrones de referencia de la faseolina en geles 2-DE cambiaron notablemente en ambos cvs. durante la transición de la semilla a la fase de germinación (véanse Figura 14a y 16a).

En primer lugar, los volúmenes de los *spots* de la faseolina mostraron claras diferencias entre los dos estados de la semilla. El cambio en el volumen de los *major spots* desde el estado de dormancia al estado de germinación (cuatro réplicas) se evaluó con el programa PDQuest para un total de 12 *major spots* en cvs. S y T. El *fold change* de cada *spot* se calculó como la relación entre el volumen promedio de un *spot* dado en las semillas durmientes y en germinación. En total, once de los doce *major spots* analizados mostraron una disminución significativa de volumen del orden de 1,2 veces o superior ($P < 0.01$, $n = 12$, test de rangos de Wilcoxon a una cola). Para la mayoría (82%) de los *major spots*, los volúmenes decrecieron significativamente de acuerdo con el test de Mann-Whitney ($P < 0,05$). Estos resultados pueden explicarse por la proteólisis de la faseolina que se produce durante la germinación de la semilla. En segundo lugar, un número sustancial de nuevos *minor spots* (8 nuevos *minor spots* en cada variedad) aparecieron en las posiciones más ácidas de los patrones de faseolina de semillas germinadas en comparación con los perfiles de faseolina de semillas durmientes.

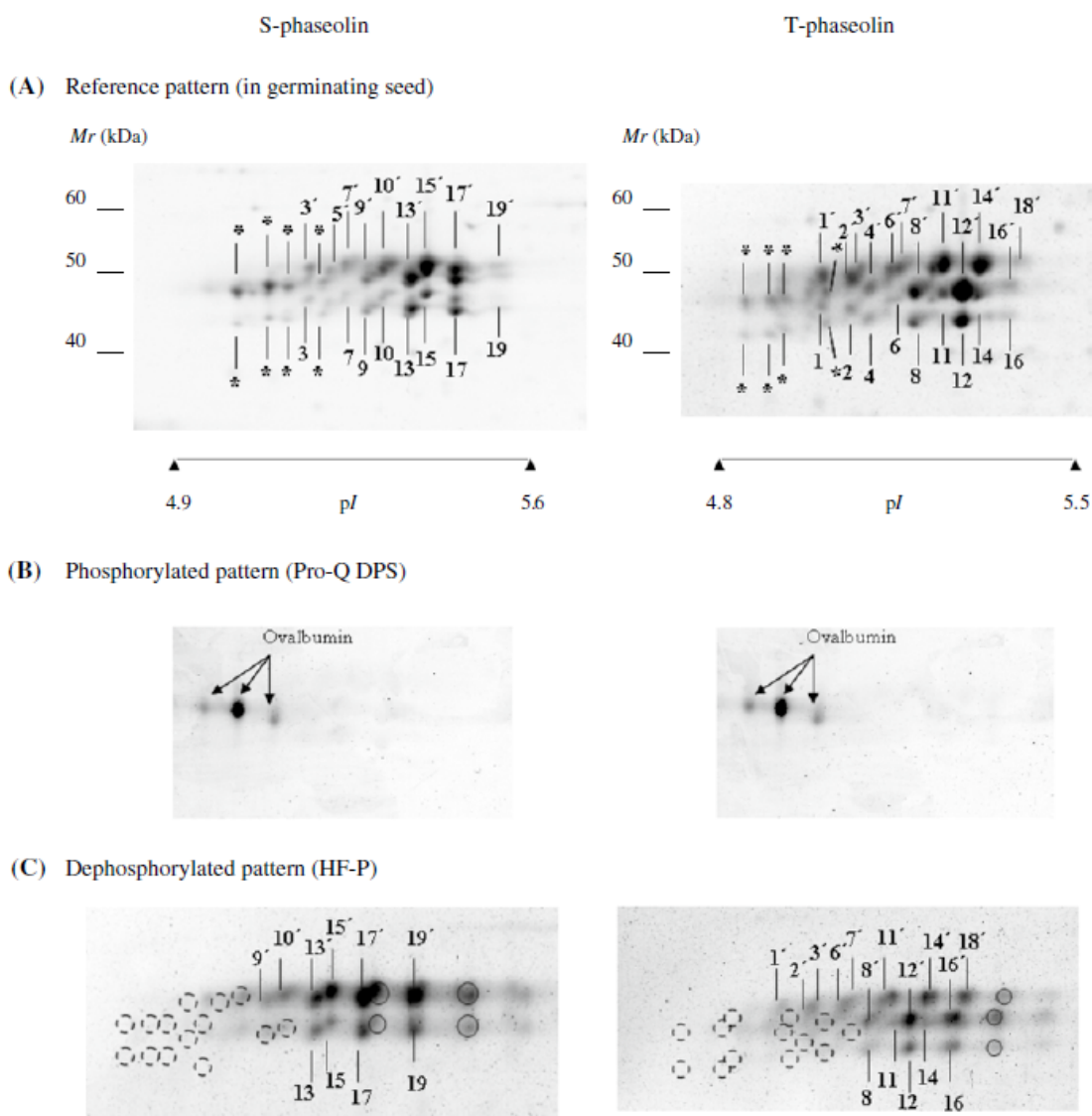


Figure 16. Degradation and identification of phosphorylated phaseolin isoforms in germinating bean seeds of cvs. S and T. **(a)** 2-DE images of gel sections showing reference phaseolin patterns of cvs. S (left) and T (right) from untreated samples of total seed protein. Asterisks indicate spots absent in reference phaseolin patterns of dormant seeds as shown in Figure 3. **(b)** Phosphorylated phaseolin patterns on 2-DE gels stained with Pro-Q DPS. **(c)** Dephosphorylated phaseolin patterns on 2-DE gels after HF-P treatment. Arrows show the signal of fluorescence of the phosphoprotein marker ovalbumin. The dashed and closed circles represent missing and newly arisen spots after dephosphorylation, respectively, as compared with reference patterns. Source: López Pedrouso et al. 2014a.

El análisis de los cambios en los niveles de fosforilación de la faseolina en la transición de semilla durmiente a semilla germinada se llevó a cabo utilizando en primer lugar la tinción Pro-Q DPS (Figura 16b). Es destacable que con Pro-Q DPS no se detectó faseolina fosforilada en los patrones 2-DE de semillas en germinación. Sin embargo, no puede descartarse la posibilidad de la existencia de niveles de fosforilación por debajo de la capacidad de detección del Pro-Q DPS. De hecho, los patrones de faseolina S y T en semilla en germinación fueron sensibles a la desfosforilación con HF-P (Figura 16c), lo que demuestra que existen formas fosforiladas de la faseolina en esta fase aunque a niveles más bajos que en semillas no germinadas. De estos resultados se deduce que el tratamiento con HF-P puede ser más eficaz para la detección de pequeñas cantidades de fosfopéptidos en geles 2-DE que el tratamiento con Pro-Q DPS. La Figura 16c también revela la desaparición de un número considerable de *minor spots* en la zona ácida del patrón de referencia (*spots* 14 y 13 en cvs. S y T, respectivamente) después de la desfosforilación con HF-P. De ello se deduce que estos *spots* sólo contenían polipéptidos fosforilados. Además, las imágenes de los geles de las muestras desfosforiladas muestran la aparición de nuevos *spots* de faseolina en posiciones más alcalinas de los patrones, en comparación con muestras control no tratadas. Los nuevos *spots* se corresponderían con fosfopolipéptidos que sufrieron una desfosforilación y por tanto una deriva hacia un *pI* más básico. En general, las observaciones sugieren que los polipéptidos de faseolina están diferencialmente fosforilados y coexisten con formas no fosforiladas en semillas de judía en germinación. Sin embargo, la proporción de isoformas fosforiladas es más baja que en las semillas en fase de dormancia, ya que no se pudieron detectar con la tinción Pro-Q DPS. En consonancia con estas observaciones, los valores de *PR* de los polipéptidos de faseolina disminuyeron notablemente en las semillas germinadas en comparación con semillas en fase de dormancia. La mayoría de los *major spots* (11 de los 12 *spots* de cvs. S y T) experimentaron cambios negativos de *PR* desde la fase de dormancia hasta la fase de germinación temprana ($P < 0.05$, $n = 12$, test de rangos de Wilcoxon a una cola; Figura 17). Es importante destacar que en las semillas en germinación, algunos *minor spots* alcanzan posiciones más ácidas en el gel, en comparación a los patrones de semilla no germinada, pero con menor peso molecular a pesar del hecho de que contienen polipéptidos altamente fosforilados. Esta

observación sugiere que estas isoformas de faseolina altamente fosforiladas son productos de degradación intermedios, resultado de la acción de enzimas proteolíticas.

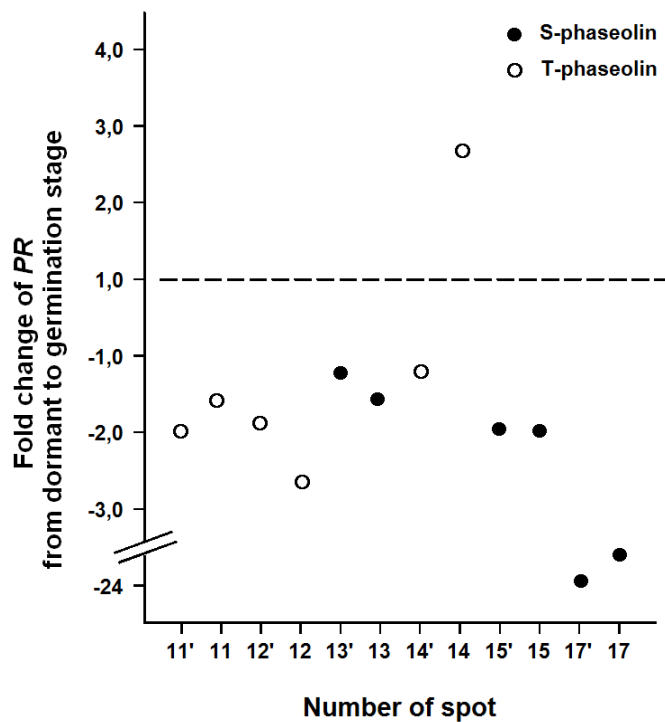


Figure 17 Fold change of the relative *PR* for major phaseolin isospot pairs (cvs. S and T) in the transition from dormant to germination. Gel position of assigned spots is shown in figure 8. Source: López Pedrouso et al. 2014a.

3.2.3. Identificación de los fosfopolipéptidos de faseolina mediante MS

Un total de 24 *spots* de faseolina de diferente pI y M_r en las semillas en fase de dormancia de los cvs. S y T se seleccionaron para su análisis por MALDI-TOF y MALDI-TOF/TOF MS. La mayoría de los *spots* seleccionados (23 de los 24 *spots*) fueron identificados inequívocamente (Tabla 6). Los resultados obtenidos muestran que cada par de *isospots* sólo contiene un único tipo polipéptido, α o β . Los pares de *isospots* con mayor M_r se corresponden con polipéptidos de tipo α , y los de menor M_r con polipéptidos de tipo β . La diferencia de M_r entre ambos tipos de polipéptidos se debe a que los genes de tipo α contienen una repetición de 15 pb en el cuarto exón y 27 pb en el sexto exón (Slightom et al., 1983, 1985; Kami y Gepts, 1994).

Un estudio preliminar de mapeado de fosfopéptidos en la proteína faseolina se llevó a cabo a partir de los datos espectrales obtenidos por MALDI-TOF. Para la detección de posibles fosfopéptidos, se tomó en cuenta la fosforilación como una posible modificación variable en el proceso de búsqueda en la base de datos Swiss-Prot. Como resultado, se detectaron un total de seis diferentes fosfopéptidos en los polipéptidos de tipo α y β de los cvs. S y T (Tabla 6).

Table 6 Phosphorylated peptides of the protein phaseolin in dry mature common bean seeds of cvs. S and T identified by MALDI-TOF

| Spot no. ^a | Exp. pI ^b | Polypeptide ^c | Submitted mass [M + H] ⁺ d | Theoretical mass [M + H] ⁺ e | Phosphopeptide ^f | Position ^g |
|-----------------------|----------------------|--------------------------|--|--|------------------------------|-----------------------|
| Cultivar S | | | | | | |
| 3' | 5.09 | α | 2,960.48 | 2,960.47 | ARVPLLLGILFLASLSASFATSLR | 4–28 |
| 3' | 5.09 | α | 1,932.98 | 1,932.95 | HILEASFNSKFEIINR | 189–204 |
| 10' | 5.19 | α | 2,960.52 | 2,960.47 | ARVPLLLGILFLASLSASFATSLR | 4–28 |
| 13' | 5.22 | β | 3,230.37 | 3,230.46 | TDNSLNVLISSIEKKEGALFVPHYYYSK | 257–283 |
| 15' | 5.24 | α | 2,960.45 | 2,960.47 | ARVPLLLGILFLASLSASFATSLR | 4–28 |
| 15 | 5.24 | α | 1,932.98 | 1,932.95 | HILEASFNSKFEIINR | 189–204 |
| 17' | 5.28 | β | 3,230.38 | 3,230.46 | TDNSLNVLISSIEKKEGALFVPHYYYSK | 257–283 |
| 17 | 5.28 | β | 2,103.92 | 2,103.91 | ATSNVNFVTFGINANNR | 336–354 |
| Cultivar T | | | | | | |
| 7' | 5.14 | α | 1,932.95 | 1,932.95 | HILEASFNSKFEIINR | 189–204 |
| 12' | 5.22 | β | 2,167.09 | 2,166.93 | EYFFLTSDNPIFSDHQQ | 119–135 |
| 12' | 5.22 | β | 2,296.05 | 2,296.05 | DVLGLTFSGSGDEVKMLINK | 376–395 |
| 12' | 5.22 | β | 3,230.46 | 3,230.35 | TDNSLNVLISSIEKKEGALFVPHYYYSK | 257–283 |
| 12' | 5.22 | β | 1,991.95 | 1,991.83 | HILEASFNSKFEIINR | 189–204 |
| 12 | 5.22 | β | 1,932.95 | 1,932.95 | HILEASFNSKFEIINR | 189–204 |

^a Spot code numbers according to Fig. 1

^b Experimental pI values

^c Type of phaseolin polypeptide

^d Tryptic peptide experimental masses (expressed in Da) of spots

^e Theoretical masses (expressed in Da) resulting from the data base search

^f Phosphopeptide sequence

^g Position of the phosphopeptide in the polypeptide sequence

Cada uno de los fosfopéptidos identificados contenía dos o más sitios posibles de fosforilación. Para algunos *spots*, se identificaron formas fosforiladas y no fosforiladas del mismo péptido lo que estaría de acuerdo con los análisis de fosforilación basados en 2-DE. Es importante subrayar que el fosfopéptido TDNSLNVLISSEIMKEGALFVPHYYSK, identificado en la posición 257-283 de la secuencia de la subunidad β , contiene el motivo de fosforilación X-T-D-X, un motivo frecuentemente representado en las SSPs (Meyer et al., 2012). Así, estudios previos indican que alrededor del 25% de las proteínas que contenían X-T-D-X en *Arabidopsis*, colza y soja son SSPs (Meyer et al., 2012). Por tanto, X-T-D-X podría ser un motivo de fosforilación clave durante el crecimiento de la semilla. En general, los estudios disponibles indican que la fosforilación de SSPs puede jugar un papel en la regulación de transición de semillas a un estado de reposo. Sin embargo, sería necesario abordar una caracterización más en profundidad de los fosfopéptidos y lugares de fosforilación de las subunidades α y β de la faseolina recurriendo a estrategias de enriquecimiento de fosfopéptidos en combinación con metodologías de MS de alta resolución. Estudios previos realizados en nuestro laboratorio, utilizando columnas de dióxido de titanio y cromatografía de afinidad de iones metálicos inmovilizados (IMAC) para la captura específica de fosfopéptidos, no dieron los frutos esperados (resultados no mostrados).

En este estudio, por tanto, se presenta la primera evidencia de isoformas de faseolina fosforiladas en semillas de judía común. Más específicamente, se encontró que formas fosforiladas y no fosforiladas de polipéptidos del tipo α y β coexisten en semillas en las fases de germinación y dormancia de las variedades S y T. De este modo, la faseolina se suma a la creciente relación de SSPs fosforiladas que ya habían sido descritas para otras especies de plantas (Schwenke et al., 2000; Agrawal y Thelen, 2006; Irar et al., 2006; Meyer et al., 2012). Estos hallazgos abren nuevas perspectivas para la comprensión de los mecanismos reguladores implicados en la síntesis y la degradación de la principal SSP de la judía común. En definitiva, las observaciones obtenidas en el presente trabajo sugieren que la fosforilación puede jugar un papel regulador importante en la síntesis de la faseolina al nivel de la transcripción, sino también en su posterior degradación actuando a un nivel post-traducciona

Estudios recientes del fosfoproteoma han permitido detectar cambios cuantitativos en los perfiles de fosfoproteínas expresadas en etapas secuenciales del desarrollo de la semilla en *Arabidopsis*, colza y soja (Agrawal y Thelen, 2006; Meyer et al., 2012.). Se encontró también que los perfiles de expresión de las fosfoproteínas siguen tendencias variables durante el crecimiento (*filling*) de la semilla dependiendo de las categorías funcionales de las proteínas. El estudio del fosfoproteoma de la semilla de colza realizado por Meyer et al. (2012) es particularmente apropiado para establecer una comparación con nuestros datos, ya que están representadas todas las etapas del desarrollo de la semilla, incluyendo la maduración tardía. Cabe destacar que la mayoría (74,5%) de los fosfopéptidos identificados en diferentes etapas secuenciales de desarrollo en la colza corresponden a la etapa de la maduración tardía. La mayoría de los fosfopéptidos identificados se asocian con proteínas conocidas abreviadamente como LEA (*Late Embryogenesis Abundant*) y con proteínas que participan en la estructura de la pared celular o en la síntesis de polisacáridos de la pared celular. En particular, se encontró que la expresión de algunas subunidades fosforiladas de la crucifera (SSP de colza, *Brassica napus*) aumentan a medida que avanza el desarrollo de la semilla, alcanzando niveles más altos de fosforilación durante la maduración tardía. Nuestras investigaciones también muestran que los fosfopéptidos de faseolina de tipo α y β se expresan abundantemente en semillas de judía común en la fase de dormancia.

La dinámica de la fosforilación proteica que tiene lugar en el tránsito de la etapa de dormancia a la de germinación es todavía poco conocida. El presente estudio muestra que el nivel de fosforilación de la faseolina es capaz de sufrir cambios dinámicos en respuesta a las variaciones fisiológicas y metabólicas, que tienen lugar desde la dormancia hasta la etapa de germinación temprana. Concretamente, se encontró que la cantidad de fosfopéptidos de tipo α y β se redujo considerablemente en las semillas en germinación en comparación con las semillas en fase de dormancia. Observaciones previas en el pino silvestre (*Pinus sylvestris* L.) también indican que los patrones de fosforilación de SSPs cambian notablemente en el momento de la germinación de las semillas (Kovaleva et al., 2013). Cambios en el grado de fosforilación proteica en respuesta a un estímulo apoyan la conclusión de que las fosforilaciones tienen un

significado funcional (Agrawal y Thelen, 2006; Myernik y Hajduch, 2011). En nuestro caso, los cambios en la cantidad de faseolina fosforilada desde la fase de dormancia a la de germinación no prueban necesariamente su importancia biológica. Es decir, los cambios en la abundancia de fosfopéptidos podrían deberse simplemente a una disminución de la cantidad total de faseolina que ocurre durante la germinación y crecimiento inicial para proporcionar los intermedios biosintéticos necesarios hasta que el desarrollo de la plántula. Sin embargo, en nuestro estudio encontramos que un subconjunto de fosfopéptidos de faseolina exhibe unos niveles de fosforilación más altos en las semillas en germinación que en las semillas en fase de dormancia. Este resultado no cabe interpretarlo simplemente por una disminución de la cantidad de faseolina durante la germinación. Los cambios en el grado de fosforilación de la faseolina detectados en respuesta a variaciones fisiológicas y metabólicas de la semilla sugieren que tienen importancia funcional. Además, se encontraron isoformas de faseolina altamente fosforiladas pero parcialmente degradadas en semillas en germinación. Lo que sugiere la hipótesis de que la faseolina se moviliza a través de la degradación de las isoformas altamente fosforiladas por acción de las proteasas. De hecho, diversos estudios demuestran que cuando las células se exponen a una variedad de estímulos extracelulares se produce una rápida fosforilación de las proteínas intracelulares y su posterior degradación vía la ruta ubiquitina-proteosoma (Karin y Ben-Neriah, 2000; Hellmann Estelle, 2002).

La fosforilación de proteínas es un evento transitorio y reversible que puede cambiar en intervalos muy cortos de tiempo lo que dificulta su identificación y caracterización (Schubert et al., 2006; Kersten et al., 2009). En este sentido, la biología de las semillas tiene muchos aspectos que permiten utilizarlas como un sistema modelo excelente para la identificación de eventos de fosforilación (Finch-Savage, 2006; Gutiérrez et al., 2007; Baud et al., 2002). La dormancia es una propiedad particularmente interesante de la semilla porque implica una falta de actividad metabólica y una actividad mínima con el fin de retrasar la germinación se retrasa en condiciones desfavorables para la supervivencia de las plántulas. Por tanto, la germinación se puede inducir bajo condiciones experimentales controladas lo que facilita la detección de los cambios rápidos que se pueden producir en los niveles de

fosforilación proteica. En el presente estudio, la estrategia experimental basada en la 2-DE de alta resolución fue fundamental para la identificación de isoformas fosforiladas de polipéptidos de faseolina así como sus cambios dinámicos en la transición desde la semilla en estado latente hasta la semilla germinada. Se ha demostrado que los niveles de fosforilación de los polipéptidos de tipo α y β varían en función del estado fisiológico de la semilla, disminuyendo de forma considerable en el caso de semillas germinadas. Sin embargo, más estudios prospectivos son claramente necesarios para abordar la regulación de la fosforilación dependiente de la expresión de la faseolina a nivel post-traducciona. Específicamente, se requiere más investigación para determinar los *phosphosites*, para determinar la estequiometría de la fosforilación en diferentes etapas de desarrollo de las semillas y en la germinación así como su relación con quinasas y fosfatasas específicas.



3.2.4. Artículo original: Evidence for phosphorylation of the major seed storage protein of the common bean and its phosphorylation-dependent degradation during germination



<http://link.springer.com/article/10.1007%2Fs11103-013-0141-1>





3.4. Evaluación de los perfiles 2-DE de la faseolina como marcadores de diferenciación genética y calidad proteica en la judía común (*P. vulgaris* L.).

3.4.1. Diversidad de patrones 2-DE de la faseolina

Los perfiles 2-DE de faseolina se analizaron para un total de 18 variedades de judía común (*P. vulgaris* L.) tanto silvestres como cultivadas de los acervos andino y mesoamericano. En la Tabla 7 se muestra la información general más relevante sobre las variedades estudiadas: código de registro, nombre común, tipo de faseolina basado sobre los patrones en 1-DE, raza, estatus, origen, peso y color de la semilla.

Los patrones 2-DE de faseolina en las 18 variedades silvestres y cultivadas de judía común se muestran en la Figura 18. Se observa que los perfiles de faseolina difieren significativamente entre las distintas variedades. Así, el número total de *spots* a lo largo de las variedades varió entre los 10 de Jalisco y los 28 de Huevo de Huanchaco (Figura 18) con un valor medio (\pm error estándar) de $20,4 \pm 1,1$. Además, se observa que los patrones de faseolina son manifiestamente menos complejos en el acervo mesoamericano (M) que en el Andino (A), siendo la media de *spots* para estos acervos de $16,8 \pm 1,4$ y $23,4 \pm 1,0$, respectivamente.

Cabe destacar que la mayor parte de estas variaciones se debieron a la presencia/ausencia de *spots* en la primera dimensión (*pI*). Estas variaciones en el *pI* de los *spots* de la faseolina se deben en gran medida a las diferencias existentes en el grado de fosforilación de los polipéptidos del tipo α y β (López-Pedrouso et al., 2014a). López-Pedrouso et al. (2014a) también demostraron que la movilización de la faseolina en las semillas germinadas parece ocurrir precisamente a través de la degradación de isoformas altamente fosforiladas. Por tanto, la diversidad de los patrones observada en el presente estudio sugiere que los mecanismos de regulación responsables de la síntesis y la posterior degradación de faseolina dependiente de la fosforilación podrían ser diferentes para los distintos tipos de faseolina.

Table 7. Information on the Different Cultivated and Wild Bean Accessions Used in This Study^a. Source: López Pedrouso et al., 2014b

| Accession identifier | Common name | phaseolin ^b | Gene pool ^c | race ^d | status | origin/ country | origin/ locality | seed/weight of 100 seeds (g) | seed/colour |
|----------------------|--------------------|------------------------|------------------------|-------------------|------------|--------------------|---------------------|---------------------------------|---------------------|
| PI 201011 | | S, Sb, Sd, M11 | M | | wild | Guatemala | Santa Rosa | 6.0 | cream, black, other |
| PI 268110 | Jamapa | S | M | M | landrace | Mexico | Veracruz | 18.0 | black |
| PI 313590 | Boyaca 22 | B, S | M | M | cultivated | Colombia | Boyaca | 16.0 | brown, black |
| PI 417780 | | M | M | | wild | Mexico | Jalisco | 10.5 | brown |
| PI 549695 | Sanilac | S | M | M | cultivated | USA | | 14.0 | white |
| PI 661723 | | | M | | wild | Costa Rica | Puntarenas | | brown |
| PI 661840 | | M4, Sb | M | | wild | Mexico | Chihuahua | 11.1 | cream, brown |
| G 5773 | Ica Pijao | B | M | M | cultivated | Colombia | Valle Del Cauca | 24.0 | black |
| G 12053 | Nuña | A | A | P | cultivated | Perú | Ayacucho | 36.6 | purple |
| G 12588 | Huevo de Huanchaco | H | A | P | cultivated | Perú | Cajamarca | 49.8 | white, red |
| PI 390770 | | T | A | | wild | Perú | Apurimac | 11.0 | cream, black |
| PI 474218 | Contender | C | A | NG | cultivated | USA | | 44.0 | cream, brown |
| PI 549633 | Tendergreen | T | A | NG | cultivated | USA | | 30.0 | cream, brown |
| PI 557465 | Coscoron | S, B | A | C | landrace | Chile | | 52.0 | white, other |
| PI 638874 | | | A | | wild | Argentina | Jujuy | | brown, black |
| PI 640968 | | | A | | wild | Argentina | Salta | | brown, black |
| PI 661817 | | | A | | wild | Bolivia | Chuquisaca | | cream |
| G 4494 | Calima | T | A | NG | cultivated | Colombia | | 54.8 | red, cream |

^aSource: Singh et al.¹⁹ and CIAT²⁰ germplasm collection. ^bPhaseolin types (and subtypes) based on 1-DE banding patterns^{11,20,28}. S, Sanilac (Sb, S Brazil, Sd, S Durango 222); M, Mesoamerica (M4, M11); B, Boyaca; A, Ayacucho; H, Huanchaco; T, Tendergreen. ^cGene pool: M, Mesoamerican; A, Andean. ^dRace: M, Mesoamerica; NG, Nueva Granada; P, Peru; C, Chile.

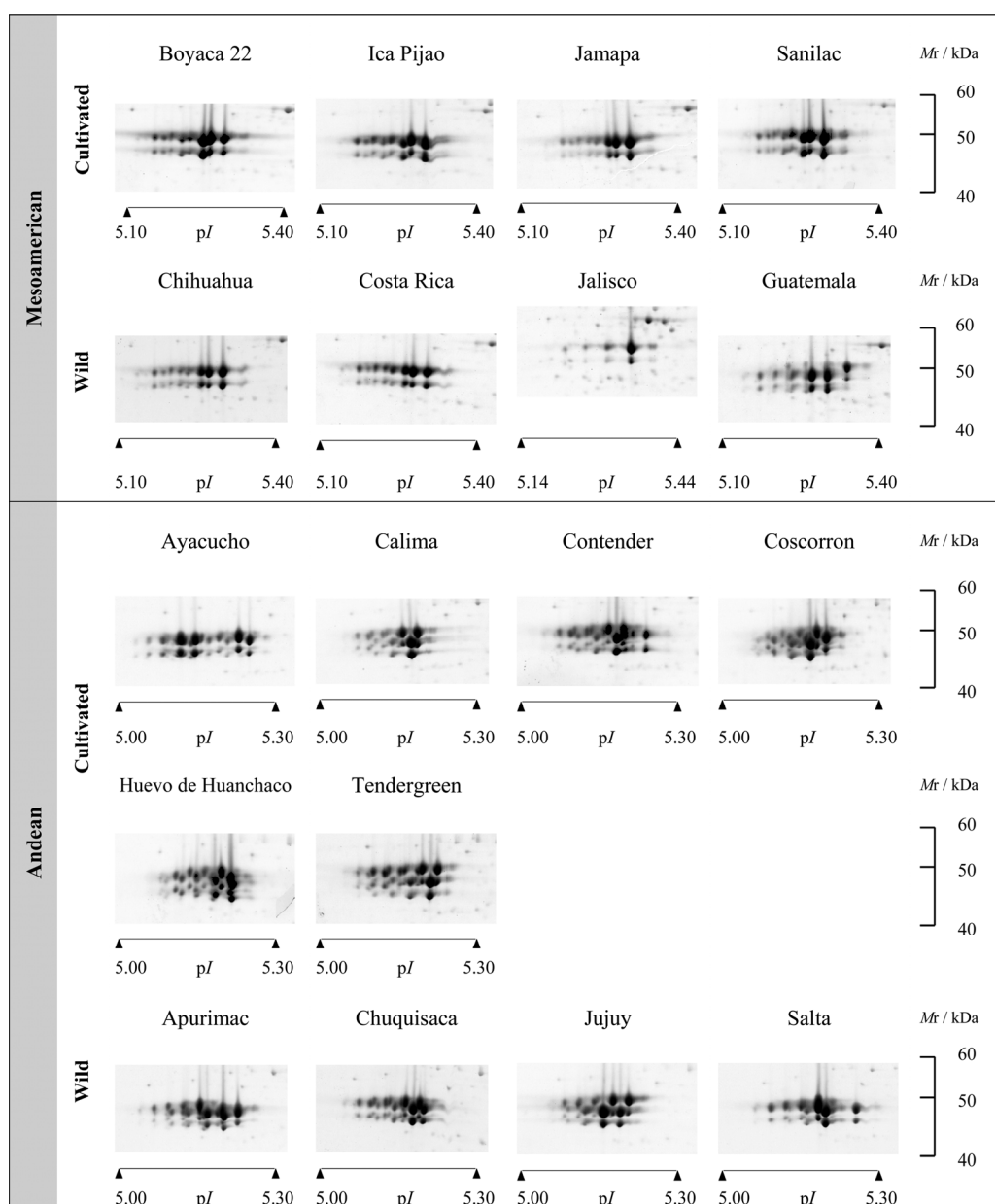


Figure 18. High-resolution 2-DE phaseolin profiles from seeds of 18 cultivated and wild accessions of common bean corresponding to the Mesomerican and Andean gene pools. Position of the spots according to their isoelectric point (pI) and molecular mass (Mr) are shown. Source: López Pedrouso et al., 2014b

3.4.2. Distancia Proteómica

Una estimación cuantitativa de las diferencias entre los perfiles 2-DE de la faseolina se obtuvo a partir de la matriz de distancias proteómicas calculadas mediante el estadístico $D = 1 - F$, $F = n_{xy} / (n_x + n_y)$, donde n_{xy} es el número de *spots* compartidos; y n_x , n_y son el número total de *spots* de la faseolina del tipo x e y, respectivamente. Los valores de D (en porcentaje) entre todos los posibles pares de poblaciones de judía común analizadas se muestran en la Figura 19. Además se empleó el *bootstrap* como técnica de remuestreo para el cálculo de la estimación de los intervalos de confianza (95 y 99%) de las medias de distancia proteómica. Se observa que las distancias proteómicas entre variedades son generalmente muy heterogéneas, con valores de D que oscilan entre 0% (tres comparaciones M × M) y el 100% (seis comparaciones M × A), siendo el valor de D promedio de $59,9 \pm 2,4\%$. La mayor parte de esta heterogeneidad se puede atribuir a las diferencias existentes entre acervos. De hecho, la distancia proteómica entre los dos acervos genéticos fue significativamente más alta que las distancias dentro de los acervos ($P < 0,01$). Dentro de cada acervo, las distancias proteómicas fueron bastante homogéneas con la excepción de las poblaciones mesoamericanas cultivadas (MC). Estas variedades exhibieron de hecho la distancia proteómica media más baja ($11,0 \pm 4,9$), significativamente diferente de las restantes distancias proteómicas promedio ($P < 0,01$). La menor diversidad observada en este trabajo para los tipos de faseolina de las poblaciones mesoamericanas es congruente con estudios previos basados en los patrones 1-DE (Gepts et al., 1986). En conjunto, la evidencia disponible apoyaría la hipótesis de que un solo suceso de domesticación de la judía pudo haber ocurrido en México, mientras que múltiples sucesos de domesticación tuvieron lugar en los Andes (Gepts et al., 1986). Sin embargo, el fenómeno de la domesticación de la judía común es complejo y sigue siendo un tema actualmente controvertido (Rossi et al., 2009).

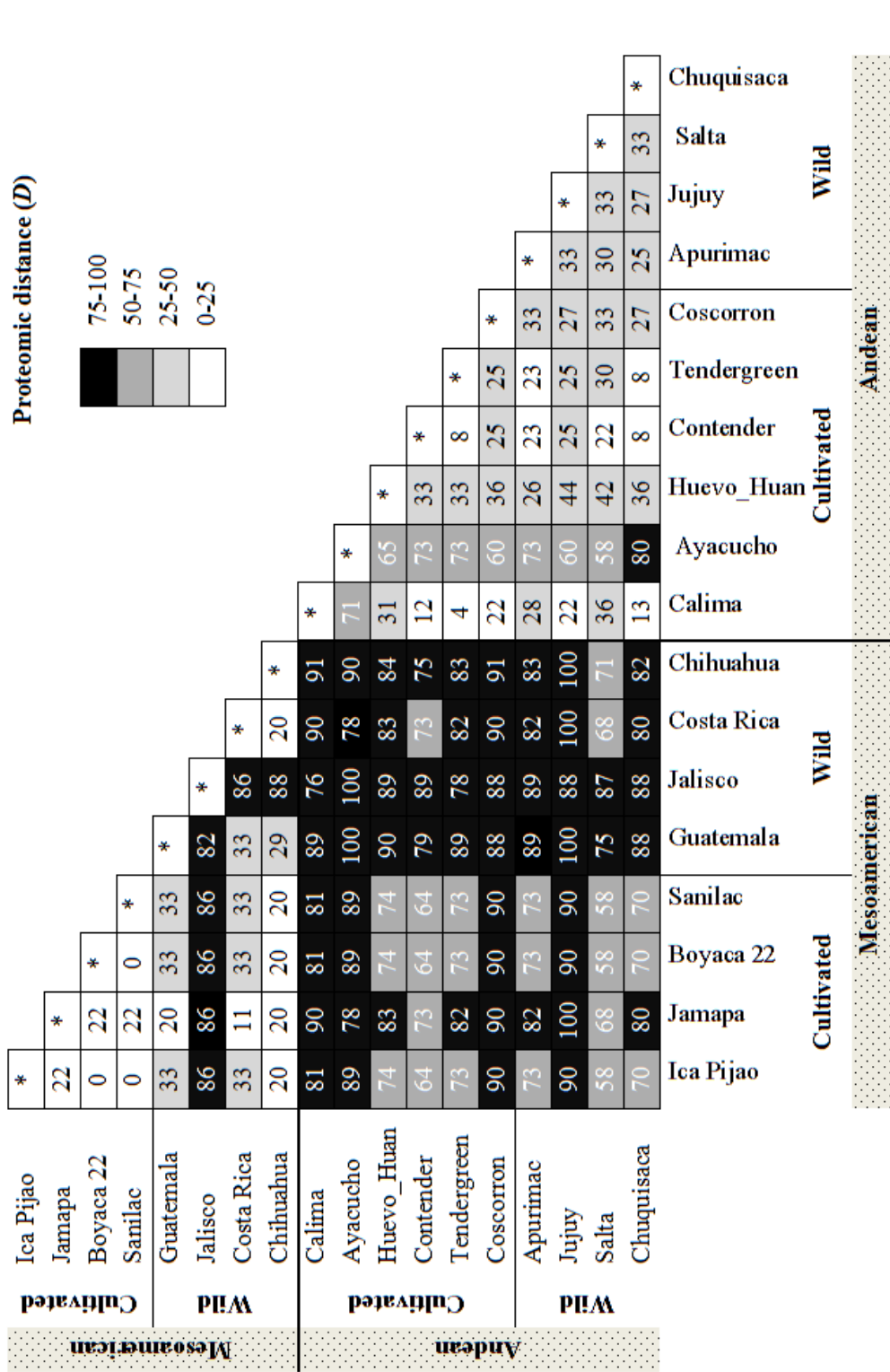


Figure 19. Pairwise proteomic distance (D), expressed as a percentage, based on high-resolution 2-DE phaseolin patterns from seeds of 18 cultivated and wild accessions of common bean. Source: López Pedrouso et al., 2014b.

Las relaciones existentes en el conjunto de las poblaciones de judía común estudiadas se analizaron aplicando el método de agrupamiento o *cluster* UPGMA (*Unweighted Pair Group Method with Arithmetic Mean*) a partir de la matriz de las distancias proteómicas (Figura 20). Es importante destacar a la luz de este análisis que las distancias proteómicas basadas sólo en los perfiles 2-DE de faseolina de alta resolución son capaces de separar adecuadamente las poblaciones de judía común en los dos grandes grupos (mesoamericano y andino) a los que pertenecen. Sólo la variedad silvestre Jalisco se separa de estos dos grandes agrupamientos en el dendrograma UPGMA. Además, el dendrograma UPGMA también fue capaz de desentrañar la heterogeneidad existente dentro de los dos grandes acervos y, en particular, la variedad Ayacucho mostró una clara diferenciación del resto de las poblaciones perteneciente al grupo andino.

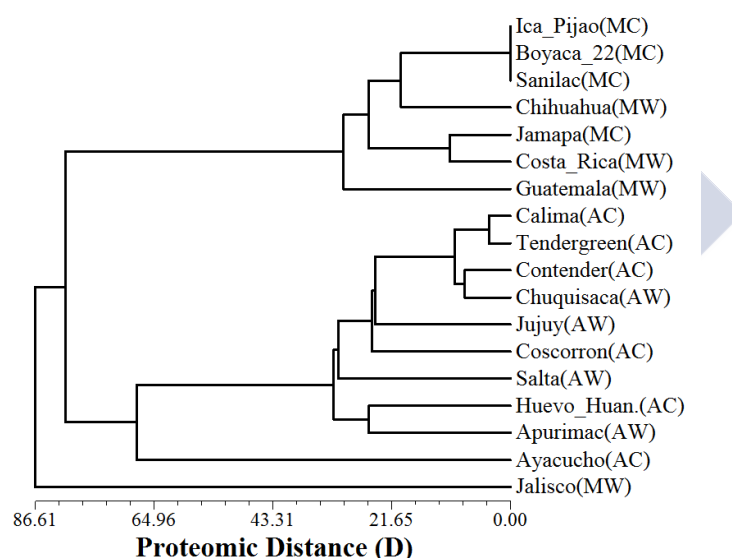


Figure 20. Unweighted pair-group method with arithmetic averaging (UPGMA) dendrogram from the matrix of pairwise proteomic distances (D) based on high-resolution 2-DE phaseolin profiles from seeds of 18 cultivated and wild accessions of common bean. Source: López Pedrouso et al., 2014b

La representación de las relaciones entre las poblaciones en un espacio bidimensional se obtuvo mediante la técnica de escalado multidimensional no-métrico (MSD, *non-metric multidimensional scaling*) (Figura 21).

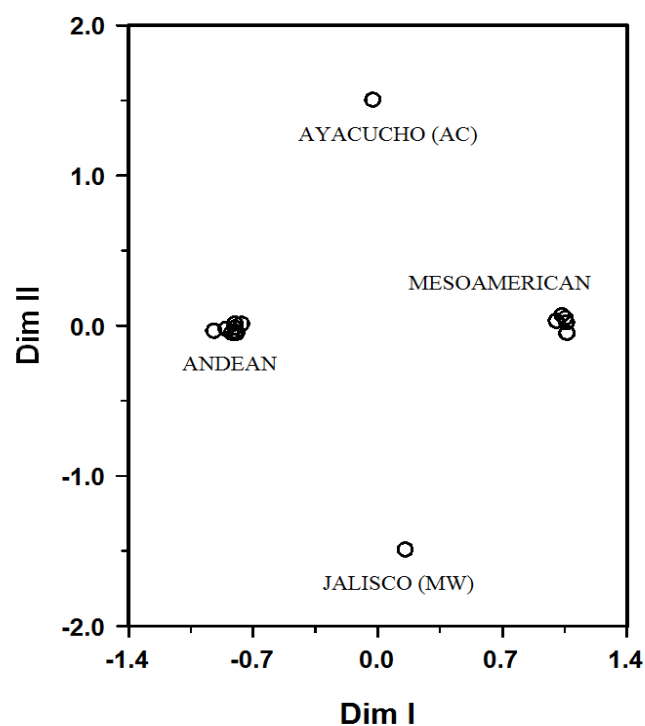


Figure 21. Multidimensional scaling (MDS) projection of the proteomic distance (D) between 18 cultivated and wild populations onto two-dimensional space based on high-resolution 2-DE phaseolin patterns. Source: López Pedrouso et al., 2014b.

El valor de la función de *stress* fue 0,074, lo que indica un buen ajuste de las distancias proyectadas en el plano con las distancias originales. El mapa bidimensional obtenido por MDS refleja esencialmente el mismo escenario que el obtenido por el dendrograma UPGMA. Así, la dimensión I revela las grandes diferencias existentes entre las poblaciones de los acervos mesoamericano y andino mientras que la dimensión II permitió separar claramente las variedades Ayacucho y Jalisco del resto de las poblaciones.

Investigaciones anteriores apoyan las grandes distancias proteómicas encontradas en el presente trabajo para las variedades Ayacucho y Jalisco (Tohme et al., 1995; Kwak y Gepts, 2009; Kwak et al., 2009; Rossi et al., 2009; Bitocchi et al., 2012; Desiderio et al., 2013). En primer lugar, es importante señalar que Ayacucho es una “nuña”, un tipo de cultivar limitado actualmente a ciertas zonas de montaña de Perú y

Bolivia, que tiene la propiedad particular de reventar como palomitas cuando se exponen al calor. Se considera que las semillas de “nuña” pueden representar antiguas variedades descendientes de las primeras judías andinas domesticadas (Tohme et al., 1995; Kwak y Gepts, 2009). Segundo, Jalisco es una posible variedad silvestre ancestral de *P. vulgaris*. Estudios recientes usando una variedad de marcadores de ADN nuclear y de cloroplastos sugieren que las formas silvestres de *P. vulgaris* se originaron en Mesoamérica y el acervo genético andino se originó a través de diferentes eventos de migración del centro de México (Rossi et al., 2009; Bitocchi et al., 2012; Desiderio et al., 2013). También parece que las semillas mesoamericanas fueron domesticadas en la región de Jalisco (Kwak y Gepts, 2009). Además, el análisis de la diversidad genética con marcadores microsatélites en un gran número de variedades silvestres y domesticadas de judía de los acervos mesoamericano y andino han demostrado que la raza Jalisco constituye un grupo separado, junto con la raza Durango, del resto de las poblaciones (Kwak y Gepts, 2009). En consecuencia, el presente estudio demuestra que las distancias proteómicas obtenidas a partir de los patrones 2-DE de faseolina son un marcador fiable de diferenciación genética entre las poblaciones cultivadas y silvestres de judía común.

3.4.3. Distancia Proteómica y calidad de las proteínas de las semillas.

La composición en aminoácidos esenciales y no esenciales en proteína total de semilla de judía se analizó a partir de dos réplicas biológicas de cada una de las 18 variedades silvestres y domesticadas de los acervos mesoamericano y andino, mediante la técnica de cromatografía líquida de alta resolución (HPLC, *high-performance liquid chromatography*). En la Figura 22 se proporciona el contenido medio en metionina (mg/g proteína total) para cada una de las variedades estudiadas.

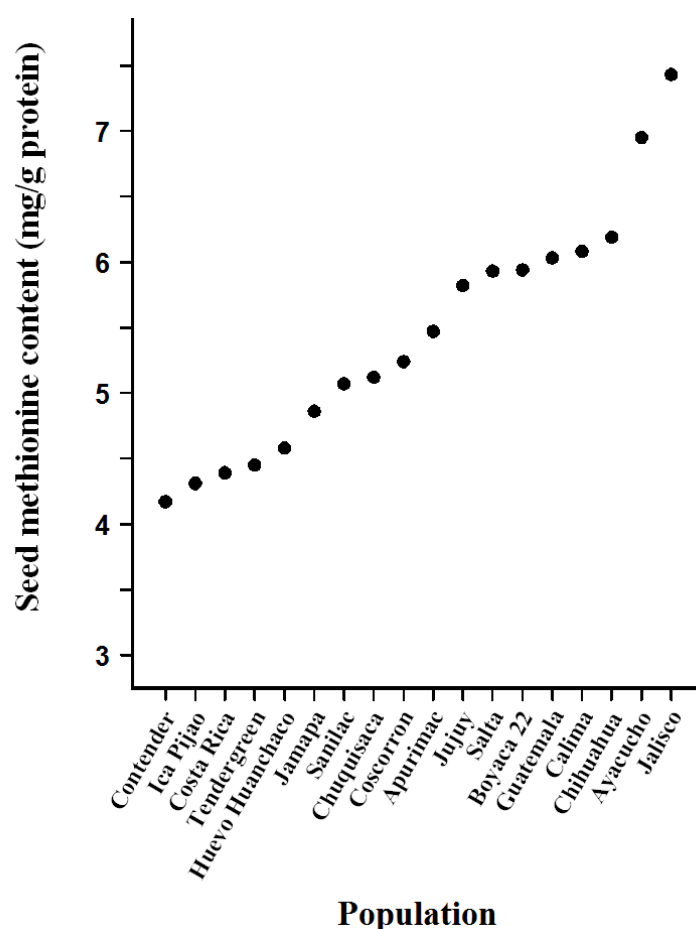


Figure 22. Mean methionine content (mg/g total seed protein) across 18 cultivated and wild populations of common bean. Source: López Pedrouso et al., 2014b.

Se observa que el contenido de metionina experimenta notables variaciones entre las distintas variedades pues los valores medios oscilaron entre 4,2 (Contender) hasta 7,4 (Jalisco) mg/g. A pesar de estas importantes diferencias, todas las variedades alcanzaron valores totales de metionina+cisteína inferiores a los niveles de referencia de estos aminoácidos recomendados por la FAO/WHO para la nutrición humana (FAO/WHO/UNU, 2007). Además, no se encontraron diferencias estadísticamente significativas en el contenido de metionina entre el conjunto de las poblaciones pertenecientes a los acervos mesoamericano y andino (test U de Mann-Whitney $p = 0,79$). Sin embargo, se detectó una débil relación entre la distancia proteómica para pares de poblaciones y su diferencia en el contenido en metionina en la proteína total de la semilla. Concretamente, se detectó una correlación positiva entre ambas variables

utilizando tanto el test de correlación de Pearson para evaluar la fuerza de la asociación y el test de Mantel para evaluar su probabilidad ($r = 0,176$, $P < 0,05$, $n = 153$). Esta débil correlación puede interpretarse en base a que la faseolina representa alrededor del 35-50% del total de la proteína en las semillas (Ma y Bliss, 1978). Estos resultados sugieren que es difícil predecir el contenido en metionina en la proteína total de la semilla a partir de los valores de distancia proteómica basados en los perfiles 2-DE de faseolina. El análisis de *cluster* UPGMA ofrece una perspectiva diferente para el estudio de la diversidad en el contenido de metionina a través de las distintas variedades de judía, debido a que su grado de semejanza/diferencia en contenido de metionina se puede analizar conjuntamente. El dendrograma UPGMA que resulta usando las diferencias en el contenido en metionina entre todos los pares de variedades de judía se muestra en la Figura 23.

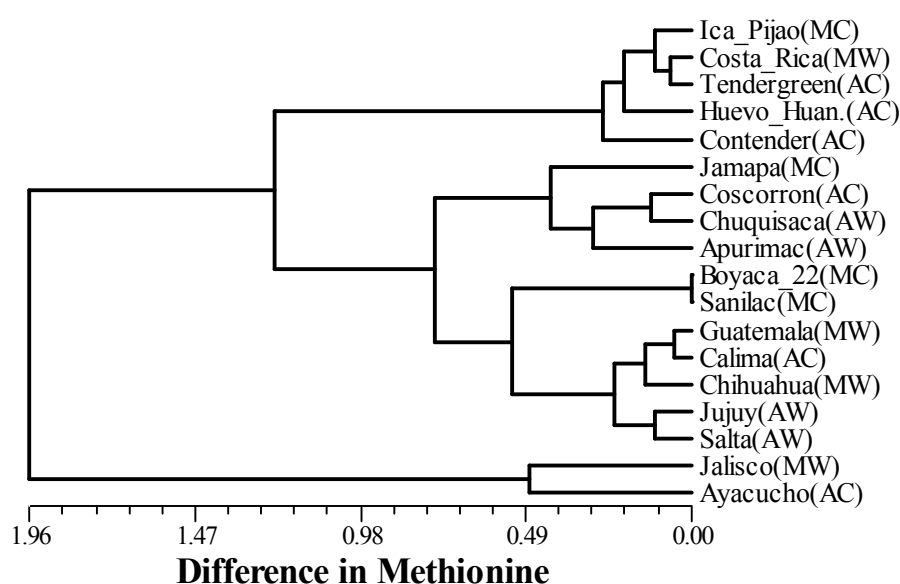


Figure 23. Unweighted pair-group method with arithmetic averaging (UPGMA) dendrogram based on pairwise differences in seed methionine content (mg/g total protein) from 18 cultivated and wild populations of common bean. Source: López Pedrouso et al., 2014b.

Es destacable la observación de que las variedades Jalisco y Ayacucho se comportan como claros *outliers* en relación al resto de las variedades estudiadas, de acuerdo con los análisis basados en las distancias proteómicas anteriormente

discutidos.. Además los niveles de metionina en las semillas de Jalisco y Ayacucho fueron significativamente más elevados que en el resto de las variedades ($P < 0,01$, Tabla 8), alcanzado valores que son un 70% más elevados que en la variedad Contender (Figura 22).

Table 8. Methionine Content (mg/g protein) in Total Seed Protein of Ayacucho and Jalisco and Other Accessions. Source: López Pedrouso et al., 2014b.

| accessions | n^a | methionine content ^b | 95% bootstrap CI (CL, CU) ^c | 99% bootstrap CI (CL, CU) |
|------------------|-------|---------------------------------|--|---------------------------|
| Ayacucho | 2 | 6.95 ± 0.15 | 6.94, 7.10 | 6.80, 7.10 |
| Jalisco | 2 | 7.44 ± 0.45 | 7.43, 7.88 | 6.99, 7.88 |
| Other accessions | 32 | 5.22 ± 0.13 | 4.98, 5.47 | 4.91, 5.52 |

^aNumber of values. ^bData are means \pm standard error. ^cCI, Confidence interval; CL, lower bound; CU, upper bound.

Estas observaciones sugieren, por tanto, que las distancias proteómicas basadas en perfiles 2-DE de faseolina son útiles para la identificación de *outliers* con un contenido de metionina diferencial. En consecuencia, las distancias proteómicas basadas en patrones de faseolina pueden establecer un vínculo entre la diferenciación genética y el contenido en metionina en proteína total, lo que tiene importantes implicaciones para la mejora genética de semillas de judía con una mayor calidad nutricional (Kami et al., 1995; Taylor et al., 2008; Montoya et al., 2010).

El estudio de la variabilidad natural de la faseolina basado en patrones 2-DE pueden proporcionar una valiosa información adicional para mejorar los distintos cultivares de judía. Estudios previos sugieren que la variación natural existente en diferentes tipos de faseolina podría ser utilizada para la mejora de la digestibilidad de la judía. Son muchos los estudios que apoyan el hecho de que el valor nutricional de la judía común se ve claramente afectado por la alta resistencia que ofrece la faseolina a su proteolisis por parte de las enzimas digestivas (Marquez y Lajolo, 1981; Jivotovskaya et al., 1996; Genovese y Lajolo, 1998, Marinello et al., 2007; Montoya et al., 2008a, 2010). Sin embargo, estudios *in vitro* basados en los patrones de faseolina 1-DE muestran que las características estructurales de la faseolina tienen una mayor influencia

en el grado de hidrólisis de la faseolina que su composición aminoacídica (Montoya et al., 2008a, 2010). Así, diversas variaciones estructurales en la proteína faseolina pueden alterar su susceptibilidad a la proteólisis como diferencias en la proporción del tipo de polipéptidos α y β en toda la molécula y modificaciones post-traduccionales tales como las glicosilaciones (Jivotovskaya et al., 1996; Genovese y Lajolo, 1998; Montoya et al., 2008a, 2010). En este sentido, los perfiles 2-DE de faseolina tienen la ventaja de que proporcionan una información más detallada acerca de las diferencias estructurales entre los tipos de faseolina que la disponible a través de los patrones de bandas de los geles monodimensionales. En particular, este estudio demuestra que los perfiles 2-DE nos permiten identificar polipéptidos del tipo α y β de faseolina diferencialmente glicosilados y fosforilados (De la Fuente et al., 2012; López-Pedrouso et al., 2014a). Sería interesante abordar en el futuro la medida en que la variación natural en los niveles de glicosilación y fosforilación de polipéptidos del tipo α y β afecta a la susceptibilidad de la proteólisis de la faseolina, tanto *in vitro* como *in vivo*. Por otra parte, también es probable que las distancias proteómicas basadas en patrones 2-DE de faseolina puedan proporcionar información relevante sobre otros caracteres en la judía, dado que la diversidad en la faseolina se ha asociado con varios caracteres de importancia agronómica como el peso de la semilla, la tolerancia a suelos de bajo pH, la resistencia a los parásitos, etc. (Gepts et al., 1986; Koenig et al., 1990; Johnson et al., 1996).

En general, los resultados obtenidos en el presente trabajo muestran que las distancias proteómicas de faseolina basadas en perfiles 2-DE de alta resolución son un marcador fiable de diferenciación genética de las poblaciones de judía común. Nuestro estudio también indica que las distancias proteómicas basadas en la faseolina pueden establecer una relación entre la diferenciación genética de *P. vulgaris* y la calidad de la semilla evaluada por su contenido en metonina. Estas asociaciones podrían asimismo afectar a otros rasgos de la judía de importancia agronómica. Por tanto, los patrones 2-DE de faseolina pueden ser de gran utilidad para acelerar el proceso de selección de germoplasma en futuros programas de mejora de las variedades de judía común.

3.4.4. Artículo original: Evaluating two-dimensional electrophoresis profiles of the protein phaseolin as markers of genetic differentiation and seed protein quality in common bean (*Phaseolus vulgaris* L.)



<http://pubs.acs.org/doi/abs/10.1021/jf502439u>







4. CONCLUSIONES



1. El estudio comparativo de los métodos de extracción de proteína total de semilla de *P. vulgaris* con fenol, TCA-acetona y el kit Clean-up reveló que el método de extracción con fenol es el más eficiente en base a la capacidad de separación de las proteínas de semilla y el nivel de resolución obtenido en geles bidimensionales (2-DE) así como en el porcentaje de éxito en la identificación proteica posterior mediante espectrometría de masas (MALDI-TOF y MALDI-TOF/TOF).
2. La optimización del método de extracción de proteína total de semilla y de los protocolos de la 2-DE ha permitido obtener patrones bidimensionales de la faseolina de alta resolución para los cvs. Sanilac y Tendergreen (T) de *P. vulgaris* que mejoran muy notablemente los de estudios previos. El análisis de los patrones 2-DE y la identificación proteica mediante MS (MALDI-TOF y MALDI-TOF/TOF) muestran que la faseolina engloba un conjunto complejo de isoformas de los polipéptidos α y β que varían en su punto isoelectrico (pI) y masa molecular (M_r). Los experimentos de desglicosilación parcial y total con la enzima PNGasa F mostraron que las variaciones en M_r se explican por una glicosilación diferencial de los polipéptidos α y β de la faseolina.
3. La aplicación de la tinción fluorescente específica de fosfoproteínas Pro-Q DPS y la desfosforilación química mediante FH-P en combinación con la 2-DE y la MS permitieron identificar por primera vez isoformas fosforiladas de la faseolina en los estados de dormancia y germinación de la semilla de los cvs. S y T. Isoformas no fosforiladas y diferencialmente fosforiladas de los polipéptidos α y β coexisten en ambos cvs. de *P. vulgaris*.
4. El análisis de los resultados del estudio de fosforilación/desfosforilación mostró que las isoformas de faseolina con elevados niveles de fosforilación parecen desempeñar un papel clave en la movilización de la proteína dado que experimentan un proceso de degradación preferente durante la germinación de la semilla. Por tanto, este

estudio permite asignar por primera vez un papel funcional a las fosforilaciones que experimentan las proteínas de reserva de semilla..

5. Los perfiles 2-DE de la faseolina de alta resolución son marcadores fiables de diferenciación genética en poblaciones silvestres y cultivadas de los acervos mesoamericano y andino. El análisis de cluster (UPGMA) basado en las distancias proteómicas (D) mostró que los perfiles 2-DE son eficaces en asignar las poblaciones analizadas al acervo que les corresponde, así como, en la identificación de poblaciones *outliers* distantemente relacionadas con el resto de las poblaciones.
6. Las distancias proteómicas basadas en los perfiles 2-DE de la faseolina suministran información valiosa sobre la calidad de la semilla al identificar poblaciones *outliers* en el contenido de metionina, un aminoácido esencial deficiente en la semilla de *P. vulgaris*. Por tanto, los perfiles 2-DE de la faseolina son marcadores potencialmente importantes para acelerar el proceso de selección del germoplasma en futuros programas de mejora genética de la calidad de la semilla en la judía común.
7. Los resultados obtenidos en el presente trabajo demuestran que la aplicación de las metodologías proteómicas son de gran utilidad para avanzar en el conocimiento de las proteínas complejas de reserva en las semillas de las plantas.



5. BIBLIOGRAFÍA



- Agrawal, G.K., Thelen, J.J. (2005). Development of a simplified, economical polyacrylamide gel staining protocol for phosphoproteins. *Proteomics* 5: 4684–4688.
- Agrawal, G.K., Thelen, J.J. (2006). Large-scale identification and quantitative profiling of phosphoproteins expressed during seed filling in oilseed rape. *Mol. Cell. Proteomics* 5: 2044–2059.
- Anthony, J.L., Vonder Haar, R.A., Hall, T.C. (1990). Nucleotide sequence of an a-phaseolin gene from *Phaseolus vulgaris*. *Nucleic Acids Res.* 18: 3396.
- Arias, J.H., Jaramillo, M., Rengifo, T. (2007). Manual: Buenas Prácticas Agrícolas, en la Producción de Frijol Voluble, 1-168.
- Bassano, L.A., Jiang, H., Ogden, L.G., Loria, C., Vupputuri, S., Myers, L., Whelton, P.K. (2001). Legume consumption and risk of coronary heart disease in US men and women. *Arch. Intern. Med.* 161: 2573–2578.
- Baud, S., Boutin, J., Miquel, M., Lepiniec, L., Rochat, C. (2002). An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. *Plant Physiol. Biochem.* 40: 151–160.
- Biron, D.G., Loxdale, H.D., Ponton, F., Moura, H., Marché, L., Brugidou, C., Thomas, F. (2006). Population proteomics: An emerging discipline to study metapopulation ecology. *Proteomics* 6: 1712–1715.
- Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P. S., Logozzo, G., Stougaard, J., McClean, P. et al. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *Proc. Natl. Acad. Sci. USA* 109: 5148–5149.
- Bliss, F.A., Brown, J.W.S. (1982). Genetic control of phaseolin protein expression in seeds of common bean *Phaseolus vulgaris* L. *Plant Foods Hum. Nutr.* 31: 269–279.
- Bollini, R., Chrispeels, M.J. (1978). Characterization and subcellular localization of vicilin and phytohemagglutinin, the two major reserve proteins of *Phaseolus vulgaris* L. *Planta (Berl.)* 142: 291-298.

- Bollini, R., Vitale, A. (1981). Genetic variability in charge microheterogeneity and polypeptide composition of phaseolin, the major storage protein of *Phaseolus vulgaris* and peptide maps of its three major subunits. *Physiol. Plant.* 52: 96–100.
- Bollini, R., Vitale, A., Chrispeels, M.J. (1983). In vivo and in vitro processing of seed reserve protein in the endoplasmic reticulum: Evidence for two glycosylation steps. *J. Cell Biol.* 96: 999–1007.
- Boylan, M.T., Sussex, I.M. (1987). Purification of an endonuclease involved with storage-protein degradation in *Phaseolus vulgaris* L. cotyledons. *Planta* 170: 343–52.
- Brown, J.W.S., Bliss, F.A., Hall, T.C. (1981a). Linkage relationships between genes controlling seed proteins in french bean. *Theor. Appl. Genet.* 60: 251–259.
- Brown, J.W.S., Ma, Y., Bliss, F.A., Hall, T.C. (1981b). Genetic variation in the subunits of globulin-1 storage protein of French bean. *Theor. Appl. Genet.*, 59: 83–88.
- Carpentier, S.C., Witters, E., Laukens, K., Deckers, P., Swennen, R., Panis, B. (2005). Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics* 5: 2497–2507.
- Ceriotti, A., Vitale, A., Bollini, R. (1989). Lectin-like proteins accumulate as fragmentation products in bean seed protein bodies. *FEBS Lett.* 250: 157–60.
- CIAT (2014) Centro Internacional de Agricultura Tropical. Genetic Resources Program: bean collection. URL <http://isa.ciat.cgiar.org/urg/beancollection.do>; visitado agosto de 2014.
- Curto, M., Camafeita, E., López, J.A., Maldonado, A.M., Rubiales, D., Jorrín, J.V. (2006). A proteomic approach to study pea (*Pisum sativum*) responses to powdery mildew (*Erysiphe pisi*). *Proteomics* 6: 163–74.
- De la Fuente, M., Borrajo, A., Bermúdez, J., Lores, M., Alonso, J., Santalla, M., de Ron, A.M., Zapata, C., Alvarez, G. (2011). 2-DE-based proteomic analysis of common bean (*Phaseolus vulgaris* L.) seeds. *J. Proteomics* 74: 262-267.

- De la Fuente, M., López-Pedrouso, M., Alonso, J., Santalla, M., de Ron, A.M., Alvarez, G., Zapata, C. (2012). In-depth characterization of the phaseolin protein diversity of common bean (*Phaseolus vulgaris* L.) based on two-dimensional electrophoresis and mass spectrometry. *Food Technol. Biotechnol.* 50: 315–325.
- De la Fuente van Bentem, S., Hirt, H. (2007). Using phosphoproteomics to reveal signalling dynamics in plants. *Trends Plant Sci.* 9: 404–411.
- De Oliveira, J.T.A., Pusztai, A., Grant, G. (1988). Changes in organs and tissues induced by feeding of purified kidney bean (*Phaseolus vulgaris*) lectins. *Nutr. Res.* 8: 943.
- Deshpande, S.S., Nielsen, S. (1987). In vitro digestibility of dry bean (*Phaseolus vulgaris* L.) proteins: The role of heat-stable protease inhibitor. *J. Food Sci.* 52: 1330–1334.
- Desiderio, F., Bitocchi, E., Bellucci, E., Rau, D., Rodríguez, M., Attene, G., Papa, R., Nanni, L. (2013). Chloroplast microsatellite diversity in *Phaseolus vulgaris*. *Plant Sci.* 3: 1–15.
- Emani, C., Hall, T.C. (2008). Phaseolin: Structure and evolution. *Open Evol. J.* 2: 66–74.
- FAO/WHO/UNU. Protein and amino acid requirements in human nutrition. WHO Technical Rep. Ser. 2007, 935.
- FAOSTAT (2013). FAO statistical data base for agriculture. Food and Agriculture Organization of the United Nations (FAO) URL <http://faostat3.fao.org/faostat-gateway/go/to/download/Q/QC/E>; visitado agosto de 2014
- Ferris, R.S.B., Tuyisenge, J., Rucibango, M., Munkankubana, D., Ngagoyisonga, C., Gatarayiha, C., Butare, C., Kanyange, L.C., Uwantege, C., et al. (2002). Atdt-Ciat/Isar/Iita-Foodnet And Pearl Project – Ruanda. International Institute Of Tropical Agricultura (IITA - Foodnet), Institute Des Sciences Agronomique Du Ruanda (ISAR), PEARL– Project., Rwanda, Africa
- Finch-Savage, W.E., Leubner-Metzger, G. (2006). Seed dormancy and the control of germination. *New Phytol.* 171: 501–523.
- Frigerio, L., Virgilio, M., Prada, A., Faoro, F., Vitale, A. (1998). Sorting of phaseolin to the vacuole is saturable and requires a short C-terminal peptide. *Plant Cell.* 10: 1031–1042

- Gallardo, K., Le Signor, C., Vandekerckhove, J., Thompson, R.D., Burstin, J. (2003). Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiol.* 133: 664–82.
- Genovese, M.I., Lajolo, F.M. (1998). Influence of naturally acid-soluble proteins from beans (*Phaseolus vulgaris* L.) on in vitro digestibility determination. *Food Chem.* 62: 315–323.
- Gepts, P., Bliss, F.A. (1986). Phaseolin variability among wild and cultivated common beans (*Phaseolus vulgaris*) from Colombia. *Econ. Bot.* 40: 469–478.
- Gepts, P., Debouck, D.G. (1991). Origin, domestication, and evolution of the common bean, *Phaseolus vulgaris*. In: Voysest O, Van Schoonhoven A (eds.), Common beans: research for crop improvement. CAB, Oxon, UK: pp. 7-53.
- Gepts, P., Osborn, T., Rashka, K., Bliss, F. (1986). Phaseolin protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris* L.): evidence for multiple centers of domestication. *Econ. Bot.* 40: 451-468.
- Ghelis, T., Bolbach, G., Clodic, G., Habricot, Y., Miginiac, E., Sotta, B., Jeannette, E. (2008). Protein Tyrosine Kinases and Protein Tyrosine Phosphatases Are Involved in Abscisic Acid-Dependent Processes in *Arabidopsis* Seeds and Suspension. *Cell. Plant Physiol.* 148: 1668–1680.
- Givalisco, P., Nordhoff, E., Kreitler, T., Kloppel, K.D., Lehrach, H., Klose, J. (2005). Proteome análisis of *Arabidopsis thaliana* by two-dimensional gel electrophoresis and matrix-assisted laser desorption/ionisation-time of flight mass spectrometry. *Proteomics* 5: 1902-1913.
- Gutierrez, L., Van Wuytswinkel, O., Castelain, M., Bellini, C. (2007) Combined networks regulating seed maturation. *Trends Plant Sci.* 12: 294–300.
- Hajduch, M., Ganapathy, A., Stein, J.W., Thelen, J.J. (2005). 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.* 137: 1397–1419.
- Han, C., Yang, P., Sakata, K., Komatsu, S. (2014). Quantitative Proteomics Reveals the Role of Protein Phosphorylation in Rice Embryos during Early Stages of Germination. *J. Proteome Res.* 13: 1766-1782.

- Hellmann, H., Estelle, M. (2002). Plant development: regulation by protein degradation. *Science* 297: 793–797
- Irar, S., Oliveira, E., Pagès, M., Goday, A. (2006). Towards the identification of late-embryogenic-abundant phosphoproteome in Arabidopsis by 2-DE and MS. *Proteomics* 6: S175–S185.
- Islam, F.M.A., Basford, K.E., Redden, R.J., Gonzalez, A.V., Kroonenberg, P.M., Beebe, S. (2002). Genetic variability in cultivated common bean beyond the two major gene pools. *Genet. Res. Crop Evol.* 49:271-283.
- Jivotovskaya, A.V., Senyuk, V.I., Rotari, V.I., Horstmann, C., Vaintraub, I.A. (1996). Proteolysis of phaseolin in relation to its structure. *J. Agric. Food Chem.* 44: 3768–3772.
- Johnson, W.C., Menéndez, C., Nodari, R., Koinange, E.M.K., Magnusson, S., Singh, S.P., Gepts, P. (1996). Association of a seed weight factor with the phaseolin seed storage protein locus across genotypes, environments, and genomes in *Phaseolus-Vigna* spp.: Sax (1923) revisited. *J. Agric. Genomics (J. Quant. Trait Loci)* 2: 5
- Kami, J., Gepts, P. (1994). Phaseolin nucleotide sequence diversity in *Phaseolus*. I. Intraspecific diversity in *Phaseolus vulgaris*. *Genome* 37: 751–757
- Kami, J., Velasquez, V.B., Debouck, D.G., Gepts, P. (1995). Identification of presumed ancestral DNA sequences of phaseolin in *Phaseolus vulgaris*. *Proc. Natl. Acad. Sci.* 92: 1101-1104.
- Karin, M., Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: the control of NF- κ B activity. *Annu Rev. Immunol.* 18: 621–663.
- Kersten, B., Agrawal, G.K., Durek, P., Neigenfind, J., Schulze, W., Walther, D., Rakwal, R. (2009). Plant phosphoproteomics: an update. *Proteomics* 9: 964–988
- Kline-Jonakin, K.G., Barrett-Wilt, G.A., Sussman, M.R. (2011). Quantitative plant phosphoproteomics. *Curr. Opin. Plant Biol.* 14: 507–511.
- Koenig, R.L., Singh, S.P., Gepts, P. (1990). Novel phaseolin types in wild and cultivated common bean (*Phaseolus vulgaris*, *Fabaceae*). *Econ. Bot.* 44: 50–60
- Kovaleva, V., Cramer, R., Krynytskyy, H., Gout, I., Gout, R. (2013). Analysis of tyrosine phosphorylation and phosphotyrosine-binding proteins in germinating seeds from Scots pine. *Plant Physiol. Biochem.* 67: 33–40.

- Krupa, U. (2008). Main nutritional and antinutritional compounds of bean seeds—a review. *Pol. J. Food Nutr. Sci.* 58: 149-155.
- Kwak, M., Gepts, P. (2009). Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). *Theor. Appl. Genet.* 118: 979–992.
- Kwak, M., Kami, J. A., Gepts, P. (2009). The putative Mesoamerican domestication center of *Phaseolus vulgaris* is located in the Lerma- Santiago basin of Mexico. *Crop Sci.* 49: 554–563
- Laugesen, S., Messinese, E., Hem, S., Pichereaux, C., Grat, S., Ranjeva, R., Rossignol, M., Bono, J-J. (2006). Phosphoproteins analysis in plants: A proteomic approach. *Phytochem.* 67: 2208-2214.
- Lawrence, M.C., Suzuki, E., Varghese, J.N., Davis, P.C., Van Donkelaar, A., Tulloch, P.A., Colman, P.M. (1990). The three-dimensional structure of the seed storage protein phaseolin at 3 Å resolution. *EMBO J.*, 9: 9–15.
- Linneo, C. (1753). *Species Plantarum*. Ed. 1. Stockholmiae
- Lioi, L. (1989). Variation of the storage protein phaseolin in common bean (*Phaseolus vulgaris* L.) from the Mediterranean area. *Euphytica* 44: 151–155.
- Lioi, L., Bollini, R. (1984). Contribution of processing events to the molecular heterogeneity of four banding types of phaseolin, the major storage protein of *Phaseolus vulgaris* L. *Plant Mol. Biol.* 3: 345–353.
- López-Pedrouso, M., Alonso, J., Zapata, C. (2014a). Evidence for phosphorylation of the major seed storage protein of the common bean and its phosphorylation-dependent degradation during germination. *Plant. Mol. Biol.* 4-5: 415-428.
- López-Pedrouso, M., Bernal, J., Franco, D., Zapata, C. (2014b). Evaluating two-dimensional electrophoresis profiles of the protein phaseolin as markers of genetic differentiation and seed protein quality in common bean (*Phaseolus vulgaris* L.). *J. Agr. Food Chem.* 62: 7200-7208.
- Ma, Y., Bliss, F. (1978). Seed proteins of common bean. *Crop Sci.* 18: 431–437
- Manly, B.F.J.: *Multivariate statistical methods: A primer*, Chapman & Hall, London, UK (1986).

- Mann, J., Chisholm, A. Cardiovascular diseases. (1999). in: Essentials of Human Nutrition (eds. J. Mann, S. Truswell). Oxford University Press, Oxford, pp. 282–312.
- Marinello, L., Giosaffato, C. V. L., Di Pierro, P., Sorrentino, A., Porta, R. (2007). Synthesis and resistance to in vitro proteolysis of transglutaminase cross-linked phaseolin, the major storage protein from *Phaseolus vulgaris*. *J. Agric. Food Chem.* 55: 4717–4721.
- Marquez, U. M., Lajolo, F. M. (1981). Composition and digestibility of albumin, globulins and glutelins from *Phaseolus vulgaris*. *J. Agric. Food Chem.* 29: 1068–1074.
- Marsolais, F., Pajak, A., Yin, F., Taylor, M., Gabriel, M., Merino, D.M. Ma, V., Kameka, A., Vijayan, P. et al. (2010). Proteomic analysis of common bean seed with storage protein deficiency reveals up-regulation of sulfur-rich proteins and starch and raffinose metabolic enzymes and down-regulation of the secretory pathway. *J Proteomics* 73: 1587–1600.
- Mathers, J.C. (2002). Pulses and carcinogenesis: potential for the prevention of colon, breast and other cancers. *Brit. J. Nutr.* 88: 273–279.
- Meiners, C.R., Derise, N.L., Lau, C., Ritchey, S.J., Murphy, E.W. (1976). Proximate composition and yield of raw and cooked mature dry legumes. *J. Agric. Food Chem.* 24: 1122-1126.
- Meyer, L.J., Gao, J., Xu, D., Thelen, J.J. (2012) Phosphoproteomic analysis of seed maturation in Arabidopsis, rapeseed, and soybean. *Plant Phys.* 159: 517–528.
- Montoya, C.A.; Lallès, J-P.; Beebe, S.; Leterme, P. (2010). Phaseolin diversity as a possible strategy to improve the nutritional value of common beans (*Phaseolus vulgaris*). *Food Res. Int.* 43: 443–449.
- Montoya, C.A., Leterme, P., Beebe, S., Souffrant, W.B., Molle, D., Lalles, J-P., (2008b) Phaseolin type and heat treatment influence the biochemistry of protein digestion in the rat intestine. *British J. Nutrition* 99: 531–539.
- Montoya, C.A., Leterme, P., Victoria, N.F., Toro, O., Souffrant, W.B., Beebe, S., Lallès, J-P. (2008a). Susceptibility of phaseolin to in vitro proteolysis is highly variable across common bean varieties (*Phaseolus vulgaris*). *J. Agric. Food Chem.* 56: 2183–2191.

- Myernik, J.A., Hajduch, M. (2011). Seed proteomics. *J. Proteomics*, 74: 389–400.
- Nakagami, H., Sugiyama, N., Mochida, K., Daudi, A., Yoshida, Y., Toyoda, T., Tomita, M., Ishihama, Y., Shirasu, K. (2010). Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. *Plant Physiol.* 153: 1161–1174.
- Nasi, A., Picariello, G., Ferranti, P. (2009). Proteomic approaches to study structure, functions and toxicity of legume seeds lectins. Perspectives for the assessment of food quality and safety. *J. Proteomics* 72: 527–38.
- Nautrup-Pedersen, G., Dam, S., Laursen, B.S., Siegumfeldt, A.L., Nielsen, K., Goffard, N., Stærfeldt, H.H., Friis, C., Sato, S. et al. (2010). Proteome analysis of pod and seed development in the model legume *Lotus japonicus*. *J. Proteome Res.* 9: 5715-5726.
- Ng, D.W-K., Hall, T.C. (2008). PvALF and FUS3 activate expression from the phaseolin promoter by different mechanisms. *Plant Mol. Biol.* 66: 233–244.
- O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250: 4007–4021.
- Osborne, T.B. (1894). The proteids of the kidney bean. *J. Am. Chem. Soc.* 16: 703–712.
- Paaren, H.E., Slightom, J.L., Hall, T.C., Inglis, A.S., Blagrove, R.J. (1987). Purification of a seed glycoprotein: N-terminal and deglycosylation analysis of phaseolin. *Phytochemistry* 26: 335–343.
- Rossi, M., Bitocchi, E., Bellucci, E., Nanni, L., Rau, D., Attene, G., Papa, R. (2009). Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L. *Evol. Appl.* 2: 504–522.
- Santoni, V., Bellini, C., Caboche, M. (1994). Use of 2-dimensional protein pattern analysis for the characterization of *Arabidopsis thaliana* mutants. *Planta* 192: 557–66.
- Saravanan, R.S., Rose, J.K.C. (2004). A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4: 2522–2532.
- Schmutz, J., McClean, P.E., Mamidi, S., Wu, G.A., Cannon, S.B., Grimwood, J., Jenkins, J., Shu, S. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* 46: 707-713.

- Schneider, A.V.C. (2002). Overview of the market and consumption of pulses in Europe. *Brit. J. Nutr.* 88: 243–250.
- Schubert, P., Hoffman, M.D., Sniatynski, M.J., Kast, J. (2006). Advances in the analysis of dynamic protein complexes by proteomics and data processing. *Anal Bioanal Chem.* 386: 482-493
- Schwenke, K.D., Mothes, R., Dudek, S., Görtniz, E. (2000) Phosphorylation of the 12S globulin from rapeseed (*Brassica napus* L.) by phosphorous oxychloride: chemical and conformational aspects. *J. Agric. Food Chem.* 48: 708–715.
- Singh, S.P. (2001). Broadening the Genetic Base of Common Bean Cultivars: A Review. *Crop Sci.* 41: 1659–1675.
- Singh, S.P., Gutierrez, J.A., Molina, A., Urrea, C., Gepts, P. (1991). Genetic diversity in cultivated common bean: II. Marker-based analysis of morphological and agronomic traits. *Crop Sci.* 31: 19–23.
- Slightom, J.L., Drong, R.F., Klassy, C., Hoffman, L.M. (1985). Nucleotide sequences from phaseolin cDNA clones: the major storage proteins from *Phaseolus vulgaris* are encoded by two unique gene families. *Nucleic Acids Res.* 13: 6483–6498.
- Slightom, J.L., Sun, S. M., Hall, T.C. (1983). Complete nucleotide sequence of a French bean storage protein gene: Phaseolin. *Proc. Natl. Acad. Sci. USA* 80: 1897–1901.
- Strasburger, E., Nol, F., Schenck, H., Schimper, A.F.W. (1994). *Tratado de Botánica*. Barcelona (España): OMEGA. 1068.
- Sturm, A., Van Kuik, J.A., Vliegenthartll, J.F.G., Chrispeels, M.J. (1987). Structure, position, and biosynthesis of the high mannose and the complex oligosaccharide side chains of the bean storage protein phaseolin. *J. Biol. Chem.* 262: 13392–13403.
- Sugiyama, N., Nakagami, H., Mochida, K., Daudi, A., Tomita, M., Shirasu, K., Ishihama, Y. (2008). Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in *Arabidopsis*. *Mol. Sys. Biol.* 4: 1-7.
- Talbot, D.R., Adang, M.J., Slightom, J.L., Hall, T.C. (1984). Size and organisation of a multigene family encoding phaseolin, the major seed storage protein of *Phaseolus vulgaris* L. *Molecular General Genet.* 198: 42-49.

- Taylor, M., Chapman, R., Beyaert, R., Hernández-Sebastià, C., Marsolais, F. (2008). Seed storage protein deficiency improves sulphur amino acid content in common bean (*Phaseolus vulgaris* L.): redirection of sulphur from γ -glutamyl-S-methylcysteine. *J. Agric. Food Chem.* 56: 5647–5654.
- Tohme, J., Toro, O. C., Vargas, J., Debouck, D. G. (1995). Variability in Andean nuña common beans (*Phaseolus vulgaris*, *Fabaceae*). *Econ. Bot.* 49: 78–95.
- Vitale, A., Bielli, A., Ceriotti, A. (1995). The binding protein associates with monomeric phaseolin. *Plant Physiol.* 107:1411–1418.
- Wang, X.C., Li, X.F., Deng, X., Han, H.P., Shi, W.L., Li, Y.X. (2007). A protein extraction method compatible with proteomic analysis for the euhalophyte *Salicornia europaea*. *Electrophoresis* 28: 3976–87.
- Weckwerth, W., Baginsky, S., van Wijk, K., Heazlewood, J. L., Millar, H. (2008). The multinational *arabidopsis* steering subcommittee for proteomics assembles the largest proteome database resource for plant systems biology. *J. Proteome Res.* 7: 4209–4210.
- Xie, C., Wang, D., Yang, X. (2009). Protein extraction methods compatible with proteomic analysis for the cotton seedling. *Crop Sci.* 49: 395–402.
- Yamagata, A., Kristensen, D.B., Takeda Y., Miyamoto Y., Okada K., Inamatsu M. et al., (2002). Mapping of phosphorylated proteins on two-dimensional polyacrylamide gels using protein phosphatase. *Proteomics* 2: 1267–1276.
- Yu, P. (2005). Protein secondary structures (α -helix and β -sheet) at a cellular level and protein fractions in relation to rumen degradation behaviours of protein: A new approach. *Br. J. Nutr.* 94: 655-665.
- Zheng, Q., Song, J., Doncaster, K., Rowland, E., Byers, D.M. (2007). Qualitative and quantitative evaluation of protein extraction protocols for apple and strawberry fruit suitable for two-dimensional electrophoresis and mass spectrometry analysis. *J. Agric Food Chem.* 55: 1663–73.
- Zhu, K., Zhao, J., Lubman, D.M. (2005). Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins. *Anal. Chem.* 77: 2745–2755