

Resumen

Las semillas del olivo constituyen un material de elevado potencial agroalimentario y biotecnológico que está comenzando a ser analizado con vistas a su potencial explotación. En la actualidad, disponemos de información proteómica sobre la semilla, así como de un transcriptoma ensamblado *de novo* correspondiente al endospermo y al embrión, y diversas secuenciaciones de tipo Illumina aún en proceso.

Como parte del análisis de esta información, se ha iniciado una categorización funcional de los transcritos expresados y de los péptidos identificados con objeto de incrementar nuestro conocimiento sobre la biología de esta semilla, sus características nutricionales, la presencia de rutas metabólicas únicas o de interés, y la presencia de componentes de utilidad como son las proteínas de almacenamiento de distintos tipos, sistemas enzimáticos, antioxidantes, componentes alergénicos, nutracéuticos etc. que permitan diseñar nuevos usos para este material con rendimiento económico.

Los productos génicos elegidos para su estudio en el presente trabajo son dos isoformas de las albuminas 2S (proteínas de almacenamiento de semillas o SSPs): la 2S *albumin-like* y la 2S *sulfur-rich seed storage protein*, y el factor de elongación 1-alfa. A partir de los transcriptomas generados, se realizaron análisis bioinformáticos completos, identificándose 49 secuencias de 2S de olivo de las cuales 12 de ellas se han usado en el alineamiento y el análisis filogenético. La estructura secundaria y terciaria predice que las albuminas 2S son prolaminas con función de proteínas de almacenamiento de semilla y transferencia de lípidos. También constan de dos dominios: el primero, un dominio hidrófobo helicoidal típico de la función de almacenamiento de semillas y el segundo, un dominio estructural que consiste en 4 hélices con una topología de hoja plegada o lámina beta, y que forma una superhélice diestra. En este dominio están integrados los inhibidores de la alfa amilasa. La 2S *albumin-like*, igual que la 2S *sulfur-rich seed storage protein*, contienen 15 átomos de azufre, 8 cisteínas (participando en 4 enlaces disulfuro) y contienen también 7 metioninas. El análisis de secuencias indica que la 2S *sulfur-rich seed storage protein* tiene potencial alergénico. Como validación experimental inicial, se generaron oligonucleótidos frente a las secuencias más relevantes, que fueron utilizados en experimentos de PCR cuantitativa en tiempo real. Los resultados muestran la expresión de los genes indicados con diferencias.

Por otra parte, se han identificado 36 secuencias de factor de elongación 1-alfa de olivo, de las cuales se han usado 30 en el alineamiento y el análisis filogenético, con 19 secuencias de EF-1- alfa de especies heterólogas. Después de alinear las secuencias y generar el árbol filogenético, el factor de elongación 1-alfa resulta estar altamente conservado en las especies de plantas analizadas.

El factor de elongación pertenece a la familia de los reguladores de proteínas de unión a GTP. Consta tres dominios: el dominio 1 que es el dominio N-terminal, también conocido como dominio GTP-regulador; el dominio 2 que adopta una estructura de barril beta, y está implicado en la unión al ARNt; y el dominio 3 representa el dominio C-terminal. El análisis de la estructura secundaria y terciaria de la proteína predice que el factor de elongación pertenece a la familia de los reguladores de proteínas de unión a GTP. Consta de tres dominios, entre los que destaca el dominio 1 (N-terminal), que incluye los motivos de unión a GTP. Se detectó de forma predictiva la presencia de 17 átomos de azufre, 6 cisteínas, pero sin formar enlaces disulfuros y la presencia adicional de 11 metioninas. El porcentaje de puentes salinos detectado en la proteína fue del 7.19%.

La información generada sobre los productos génicos analizados permite diseñar usos de interés para estos componentes de la semilla. En el caso de las proteínas 2S, se ha conseguido generar un anticuerpo específico mediante el diseño de péptidos sintéticos basados en su secuencia que servirá para validar experimentalmente su expresión en la semilla, digestibilidad, alergenicidad y otras propiedades. Se pretende continuar igualmente la validación experimental de los resultados obtenidos de forma predictiva para el factor de elongación en una vía experimental similar.

Palabras clave:

Proteínas de almacenamiento de semillas; SSP; 2S *albumin-like*; 2S *sulfur-rich seed storage protein*; Factor de elongación 1-alfa; análisis bioinformático.

Abstract

Olive seeds constitute a material of high agro-food and biotechnological potential that is beginning to be analyzed with a view to its potential exploitation. At present, we have proteomic information on the seed, as well as a de novo assembled transcriptome corresponding to the endosperm and embryo, and several Illumina-type sequences still in process.

As part of the analysis of this information, a functional categorization of the expressed transcripts and identified peptides has been initiated in order to increase our knowledge about the biology of this seed, its nutritional characteristics, the presence of unique or interesting metabolic pathways, and the presence of useful components such as storage proteins of different types, enzymatic systems, antioxidants, allergenic components, nutraceuticals, etc. that allow us to design new uses for this material with economic returns.

The gene products chosen for study in the present work are two isoforms of the 2S albumins (seed storage proteins or SSPs): the 2S albumin-like and the 2S sulfur-rich seed storage protein, and the elongation factor 1-alpha. From the generated transcriptomes, complete bioinformatics analyses were performed, identifying 49 olive 2S sequences of which 12 of them have been used in alignment and phylogenetic analysis. The secondary and tertiary structure predicts that the 2S albumins are prolamins with seed storage and lipid transfer protein function. They also consist of two domains: the first, a hydrophobic helical domain typical of seed storage function and the second, a structural domain consisting of 4 helices with a folded sheet or beta-sheet topology, and forming a right-handed superhelix. Alpha amylase inhibitors are integrated into this domain. The 2S albumin-like, like the 2S sulfur-rich seed storage protein, contains 15 sulfur atoms, 8 cysteines (participating in 4 disulfide bonds) and also contains 7 methionines. Sequence analysis indicates that the 2S sulfur-rich seed storage protein has allergenic potential. As an initial experimental validation, oligonucleotides were generated against the most relevant sequences, which were used in real-time quantitative PCR experiments. The results show the expression of the indicated genes with differences.

Moreover, 36 olive elongation factor 1-alpha sequences have been identified, of which 30 have been used in alignment and phylogenetic analysis, with 19 EF-1-alpha sequences from heterologous species. After aligning the sequences and generating the phylogenetic tree, the elongation factor 1-alpha turns out to be highly conserved in the analyzed plant species.

The elongation factor belongs to the family of GTP-binding protein regulators. It consists of three domains: domain 1 which is the N-terminal domain, also known as the GTP-regulatory domain; domain 2 which adopts a beta-barrel structure, and is involved in tRNA binding; and domain 3 represents the C-terminal domain. Analysis of the secondary and tertiary structure of the protein predicts that the elongation factor belongs to the family of GTP-binding protein regulators. It consists of three domains, most notably domain 1 (N-terminal), which includes the GTP-binding motifs. The presence of 17 sulfur atoms, 6 cysteines, but without forming disulfide bonds and the additional presence of 11 methionines were predicted. The percentage of salt bridges detected in the protein was 7.19%.

The information generated on the gene products analyzed allows us to design uses of interest for these seed components. In the case of the 2S proteins, a specific antibody has been generated by designing synthetic peptides based on their sequence, which will be used to experimentally validate their expression in the seed, digestibility, allergenicity and other properties. It is also intended to continue the experimental validation of the results obtained in a predictive way for the elongation factor in a similar experimental pathway.

Keywords :

Seed storage proteins ; SSP ; 2S *albumin-like*; 2S *sulfur-rich seed storage protein*; Elongation factor 1-alfa; bioinformatics analysis