Posted on 3 Dec 2022 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.167005135.55252518/v1 | This a preprint and has not been peer reviewed. Data may be preliminary.

Shotgun proteomics of quinoa seeds reveals chitinases enrichment under rainfed conditions

Maria Reguera[1], Laura Poza-Viejo[1], Miguel Redondo-Nieto[1], Javier Matías[2], Verónica Cruz[2], Sara Granado-Rodríguez[1], Isaac Maestro-Gaitán[1], Enrique Olmos[3], and Luis Bolanos[1]

[1]Universidad Autonoma de Madrid Departamento de Biologia
[2]Centro de Investigaciones Científicas y Tecnológicas de Extremadura (CICYTEX
[3]Centro de Edafologia y Biologia Aplicada del Segura

December 3, 2022

Abstract

Quinoa is an Andean crop whose cultivation has been extended to many different parts of the world in the last decade. It shows a great capacity for adaptation to diverse climate conditions, including environmental stressors, and moreover, the seeds are very nutritious in part due to their high protein content which is rich in essential amino acids. They also contain good amounts of other nutrients such as unsaturated fatty acids, vitamins, or minerals and are gluten-free seeds. Also, the use of quinoa hydrolysates and peptides has been linked to numerous health benefits. Altogether, these aspects have situated quinoa as a crop able to contribute to food security worldwide. Aiming to deepen our understanding of the protein quality and function of quinoa seeds and how they can vary when this crop is subjected to water-limiting conditions, a shotgun proteomics analysis was performed to obtain the proteomes of quinoa seeds harvested from two different water regimes in field: rainfed and irrigated conditions. Overrepresented proteins in seeds from each field condition were analysed, and the enrichment of chitinase-related proteins in seeds harvested from rainfed conditions was found. These proteins are described as pathogen-related proteins and can be accumulated under abiotic stress. Thus, our findings suggest that chitinase-like proteins in quinoa seeds can be potential biomarkers of drought. Also, this study points to the need for further research to unveil their role in conferring tolerance when coping with water-deficient conditions.

1. Introduction

Chenopodium quinoa Willd., commonly known as quinoa, is an allotetraploid species (2n = 4x = 36) belonging to the Amaranthaceae family and taxonomically related to beet, spinach, and amaranth (Chase*et al.* , 2016). The quinoa genome was recently sequenced enabling a better genomic understanding of this underutilized crop, which possesses a huge genetic diversity (with more than 6000 accessions described) linked to a great capacity for adaptation to a wide variety of environments (including those with high salinity or low water supply) (Jarvis *et al.* , 2017; Zou *et al.* , 2017; Yasui *et al.* , 2016; González *et al.* , 2015; Rojas *et al.* , 2010). In fact, quinoa has emerged as a promising crop whose cultivation has been expanded from its traditional agronomical areas, located in the Andean region, to more than 120 countries with very different climatic conditions, including Spain, France, Morocco, India or Pakistan, although Bolivia and Peru are still the largest producers (Pulvento*et al.* , 2010; Jacobsen *et al.* , 2013; Bazile *et al* ., 2016; Choukr-Allah *et al* ., 2016; Angeli *et al.* , 2020; Alandia *et al* ., 2020; Granado-Rodríguez *et al.* , 2021).

Furthermore, quinoa seeds have a remarkable nutritional profile with a high-quality protein composition that provide all the essential amino acids (including the most limiting amino acids in cereals and pulses, which are lysine and methionine, respectively) (Ando *et al.* , 2002). The most abundant proteins in quinoa seeds

Posted on 3 Dec 2022 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.167005135.55252518/v1 | This a preprint and has not been peer reviewed. Data may be preliminary.

are the storage proteins 2S albumins and 11S globulins (Janssen *et al.* , 2017), this last described as a specific type in quinoa called chenopodin (Brinegar and Goundan, 1993). Interestingly, neither prolamins nor other typically present celiac epitopes are found among the quinoa seed profile, giving nutritional value to the seeds as gluten-free food products that can be consumed by celiacs. In addition, quinoa seeds´ hydrolysates and peptides show bioactive properties including antioxidant capacity, antidiabetic, anti-inflammatory or ACE-related antihypertension activities (Guo *et al.* , 2021). Besides, quinoa seeds also provide polyunsaturated fatty acids, dietary fiber, minerals and vitamins (Abugoch James, 2009; Vega-Gálvez *et al.* , 2010; Gordillo-Bastidas *et al.* , 2016).

On the other hand, within the current climate context, extensive cultivation areas are expected to suffer from long drought episodes, especially those located in arid or semi-arid regions, such as the Mediterranean region (Araus, 2004; Jacobsen *et al.* , 2013; Tramblay *et al.* , 2020). This, together with the high global demand for food and feed for livestock, requires the selection of climate-resilient and nutritious crops, such as quinoa, which can contribute to global food security (Jacobsen *et al.* , 2013).

Understanding how plants perceive abiotic factors and adapt to adverse environmental conditions (abiotic stresses) is crucial to dealing with environmental and food future scenarios. Plant responses to abiotic stresses comprise complex molecular networks (at transcriptomic, proteomic, and metabolic levels) that result in morphological, physiological, and molecular adjustments that can lead to protection mechanisms for ensuring plant adaptation and survival under environmental constraints (Farooq *et al.* , 2009; Zhang *et al.* , 2022). The extent to which these responses can cause molecular changes usually depends on the type of stress (or the combination of stressors), the duration, and the intensity (Zhu, 2016; Zhang *et al* ., 2022). In line with this, large genomes, with high gene copy number, redundancy, and diversification of gene functions, as occurs in quinoa, shape plasticity of the plant genome architecture, which may contribute to dealing with unfavourable environmental conditions (Hinojosa *et al.* , 2018; Grenfell-Shaw and Tester, 2021; Zhang*et al.* , 2022).

Eventually, plant strategies can converge in the use of the same protein families to face different stresses and diversify individual functions in order to respond to specific conditions (Zhang *et al.* , 2022). In this regard, plant chitinases shape large gene families which are expressed under different biotic but also abiotic stresses (Grover, 2012). Although plant chitinases are the major and best-characterized pathogen-related (PR) proteins due to their hydrolase activity that enables them to clave chitin coming from arthropods or fungi, they are also involved in abiotic stress signalling functioning at different stages of plant development (Grover, 2012; Ben-Amar *et al.* , 2022). More specifically, chitinases hydrolyze ß-1,4 bonds that link long-chain polymers of N-acetyl-D-glucosamine, which conforms chitin´s structure, the second amplest biopolymer in nature only after cellulose (Oyeleye and Normi, 2018). This cleavage generates small lipo-chito-oligosaccharides (LCOs) which can act as plant resistance elicitors under biotic and abiotic stress in plants, although their functions are still not well characterized (Singh and Subudhi, 2014).

Plant chitinases are generally classified into six classes (class I – class VI) based on their genomic sequence, and are divided into Glycosyl Hydrolase family 18 (GH18) or Glycosyl Hydrolase family 19 (GH19), depending on their characteristic catalytic domain (Li and Greene, 2010; Kesari *et al.* , 2015). Both families have evolved from different ancestral genes, thus, their genomic sequences and 3D protein structures are strongly different (Tyler *et al.* , 2010). GH18 chitinases (classes III and V) have typically enzymatic triose-phosphate isomerase (TIM)-barrel fold structure, while GH19 (classes I, II, IV, and VI) have mainly helicoidal protein structure. Also, GH19 are pretty similar to other catalytic enzymes such as chitosanases and lysozymes (Santos*et al.* , 2008; Takenaka *et al.* , 2009). Besides, chitinases from class I GH19 possess a chitin-binding domain (ChtBD) at the N-terminal region (Tang *et al.* , 2004). Plant chitinases usually are targeted to the vacuolar compartment or are secreted to the apoplast and are expressed in a tissue-specific way along the plant (Oyeleye and Normi, 2018).

Beyond catalytic active chitinases, a large number of genes transcribing chitinase-like proteins (CLPs) are described along plant genomes. CLPs are "inactive" chitinases that share a strong similarity in their genomic sequence and structure to GH18 or GH19 chitinases. However, they have lost their catalytic activity or

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

their ChtBD, thus providing a source of functional diversification as emerging enzymes able to bind other polysaccharides and/or new catalytic activities hydrolysing diverse substrates (Kesari *et al.,* 2015).

Previous quinoa proteomic profiles have been published during the last years describing a discrete number of proteins accumulated in quinoa seeds. However, the lack of an accurate genome annotation or proteome information for *C. quinoa* greatly limited the outcomes of these studies. Thus, Capriotti et al. (2015) were only able to identify four specific proteins accumulated in quinoa seeds. In 2019, Burrieza and collaborators improved quinoa seed proteomic research utilizing the sequenced genome of the crop (Yasui *et al.* , 2016; Jarvis *et al.* , 2017; Zou *et al.* , 2017), identifying novel seed storage proteins of quinoa which contribute to the characteristic high-lysine content of the seeds. Recently, a descriptive proteomic study identified a total of 1211 seed proteins among four commercial quinoa varieties (Galindo-Luján *et al.* , 2021). However, none of the studies mentioned above have analysed the impact of abiotic stress changing the proteomic profile of quinoa seeds.

Here, aiming at analysing the quinoa seed proteome by using a shotgun proteomic approach, we evaluated changes associated with water limitation (rainfed conditions) when compared to full irrigation (irrigated conditions) in quinoa seed samples obtained from the field. In this regard, as far as we know, we report here the most complete quinoa seed proteome to date, finding putative quinoa seed chitinases as an overrepresented protein family in quinoa seeds under water limitation. Overall, our data highlight a potential role of chitinases in water stress responses in quinoa and the possibility of using these group of proteins as water stress biomarkers which can be useful for quinoa breeding programmes and crop improvement strategies.

2. Materials and methods

2.1. Plant material and experimental conditions

Quinoa (*Chenopodium quinoa* Willd.) seeds belonging to the cultivar F14 provided by Algosur SL (Seville, Spain), were grown in the field under two environmental conditions at two experimental stations belonging to the *Center for Scientific and Technological Research of Extremadura* (CICYTEX, Extremadura, Spain): under irrigated conditions (by applying drip irrigation) (latitude 38° 51'10" N; longitude 6deg 39'10" W) and under rainfed conditions (latitude 38deg 23' 29" N; longitude 5deg 42' 28" W). Both locations were nearly located and their monthly mean temperatures and precipitations were similar (Supplementary Fig. S1).

Sowing was conducted in February 2019 at a dose of 6 kg ha^{-1} using a mechanical plot drill. Harvesting was conducted at physiological maturity of the plants. The sampling area was 3 m^2 per elemental plot. Plants were manually cut at ground level and the seeds were separated using a stationary thresher (Wintersteiger LD 352, Ried, Austria).

2.2. Protein extraction and quantitative label free proteomic analysis (LC-MS/MS)

2.2.1. Protein precipitation

Three biologically independent pools of quinoa seeds obtained from rainfed and irrigated conditions were dried and milled. Fifty mg were solubilized in urea 8 M and filtrated to obtain 1 ml of solubilized protein suspension for each sample before starting the precipitation protocol. Proteins were precipitated by adding cold chloroform/methanol 1/3 (v/v) to each sample, followed by 10 min vortex at 4o C, and the addition of 3 volumes of milliQ water. Later, samples were incubated for 10 mins at 4o C and centrifuged at 1000 g for 2 mins to discard the supernatant. To solubilize the precipitated proteins, 3 volumes of methanol were added and mixed by vortex for 10 mins. Then, samples were centrifuged at 10000 g for 5 mins. Supernatants were discarded and proteins were resuspended in 2 ml urea 8 M.

2.2.2. Protein concentration and trypsin digestion

Firstly, 50 µl of each protein sample were loaded on a 10% acrylamide gel using a Mini-PROTEAN® Tetra Cell (Bio-Rad). Protein electrophoresis was performed in Laemmli buffer (Laemmli, 1970) at 100 V. Then, gels were fixed in methanol 50% (v/v) and phosphoric acid 2% (v/v) for 30 mins and then washed, rinsing the gel twice with milliQ water. Later, gels were incubated in methanol 33% (v/v), ammonium sulphate 17%

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

(v/v) and phosphoric acid 3% (v/v) for 45 mins. Protein bands were visualized after incubating the gel in colloidal Coomasie (G-250) and methanol (6.6 mg/ml) overnight and rinsing the excess of Coomasie solution with milliQ water.

Remains of Coomasie solution were removed by rinsing protein gels twice with pure acetonitrile (ACN) and ammonium bicarbonate 25 mM. Disulphide bonds were reduced using dithiothreitol (DTT) 20 mM in ammonium bicarbonate 25 mM, 56º C, for 30 mins and then blocked with iodoacetamide 22.5 mM in ammonium bicarbonate 25 mM, 15 mins, in darkness. Two more washes were performed with ACN before completing dehydration of the gel using the SpeedVac (Thermo Fisher Scientific, Massachusetts, United States) for 30 mins.

Finally, protein bands were cut, and trypsin (Roche, Mannheim, Germany) 1:100 (v/v) in ammonium bicarbonate 25 mM was added for digestion at 37º C overnight. Digested peptides were recovered from the supernatant and dried using the SpeedVac (Thermo Fisher Scientific, Massachusetts, United States) for 30 mins and resuspended in 31 µL ACN 2% (v/v) and formic acid 0.1% (v/v). One µL of each protein extraction was used to determine sample concentration using Invitrogen Qubit 3 (Thermo Fisher Scientific, Massachusetts, United States).

2.2.3. Reversed-phase liquid chromatography (LC) for peptide separation

One µg of each protein extraction was injected into a nano-HPLC Easy-nLC 1000 (Thermo Fisher Scientific, Massachusetts, United States). Firstly, samples were concentrated using a precolumn PEPMAP100 C18 NanoViper Trap (Thermo Fisher Scientific, Massachusetts, United States). Then, samples were separated through a 50 cm column PEPMAP RSLC C18 (Thermo Fisher Scientific, Massachusetts, United States) on a gradient of ACN 5% to 40% (v/v) and formic acid 0.1% (v/v) for 120 min.

2.2.4. Data-dependent acquisition (DDA) for shotgun proteomics

Peptide fractions were electrospray ionized in positive mode and analyzed by a quadrupole Orbitrap mass spectrometer (Q Exactive HF, Thermo Fisher Scientific, Massachusetts, United States) in DDA mode. From each mass spectrometry (MS) scan (between 390 and 1700 Da), the 15 most intense precursors (charged between 2+ and 5+) were selected for their high collision energy dissociation (HCD) fragmentation. Then, the corresponding tandem mass spectrometry (MS/MS) spectra was acquired.

2.3. Quantitative proteomic analysis

2.3.1. Protein identification

Data generated by LC-MS/MS for each quinoa seed sample was analyzed using Proteome Discoverer 2.4 (Thermo Fisher Scientific, Massachusetts, United States). Each MS/MS spectra was identified by peptide-spectrum matches (PSMs) comparing them to theoretical masses obtained from the original precursor mass fragmentation, using JGI-Phytozome database (https://phytozome-next.jgi.doe.gov/; Phytozome genome ID: 392) taxonomically restricted to *Chenopodium quinoa* v1.0. Identified peptides were assigned to the annotated *C. quinoa* proteins. Whether a peptide may be assigned to different proteins, the software used parsimony principle to generate a master protein. Percolator algorithm was used to estimate false discovery rate (FDR). High-confidence proteins were identified filtering by $p\text{-}adj < 0.05$.

2.3.2. Peptide and protein normalization

Proteome Discoverer 2.4 (Thermo Fisher Scientific, Massachusetts, United States) was used to determine peptide and protein abundance. Firstly, mass recalibration was performed with Sequest HT comparing data base and identified proteins, getting a chromatography alignment of the samples with a tolerance up to 10 min. Then, an alignment of the retention time of all samples was performed to quantify precursor ions (considering unique peptides that were present in, at least, two of the three replicates). Finally, total protein amount was normalized among samples using peptide total abundance.

2.3.3. Sample pooling and relative protein quantification

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Three biological replicates were analyzed for each treatment (irrigated and rainfed conditions) using a No Nested/Pairwise design. Quantified proteins were obtained from peptide ratios calculated as a geometric median of the peptide ratio in each biological replicate (Supplementary Table S1).

Analysis of variance (ANOVA) was performed to estimate differentially abundant proteins between quinoa seeds harvested from irrigated and rainfed conditions with a significance level of *0.05* .*P-values* obtained with this analysis were corrected (*p-adj*), taking into account the False Discovery Rate (FDR) applying Benjamin&Hochberg (BH) test (Supplementary Table S3 and Table S4).

2.3.4. Protein annotation

C. quinoa Willd accession PI 614886 coding sequences (CDSs) from JGI last annotation version comes from 2017. In order to improve gene ontology (GO) terms associated to each protein, a new reannotation was carried out in our laboratory as follows. CDS sequences were downloaded and blasted against NCBI non redundant database (January 2022). Then, BLAST output was processed with Blast2GO software (*https://www.blast2go.com*) (Conesa *et al.* , 2005) to get a tabular file with the corresponding new functional annotation.

2.4. Gene ontology (GO) enrichment analysis

Functional enrichment was studied in different groups: proteins that appeared exclusively in seeds from plants grown under irrigated or rainfed conditions, proteins that were enriched in seeds from plants grown under rainfed conditions ($\log_2$FC [?] 1, *p-adj*[?] *0.05* ; n = 3) and proteins enriched in seeds from plants grown under irrigated conditions ($\log_2$FC [?] -1, *p-adj* [?]*0.05* ; n = 3). For each cluster of proteins, a GO term enrichment analysis was performed in R (R Fundation for Statistical Computing, 2020) using as annotation the list of terms obtained with Blast2GO using the topGO package (Alexa and Rahnenfuhrer, 2022).

2.5. Phylogenetic analysis

Phylogenetic analyses were performed with protein sequences using the online platform NGPhylogeny.fr (Lemoine *et al.* , 2019), following the FastTree/OneClick workflow (*https://ngphylogeny.fr/)*: MAFFT 7.407 for multiple alignment (Katoh and Standley, 2013), BMGE 1.12_1 for alignment curation (Criscuolo and Gribaldo, 2010), FastTree 2.1.11 for approximately maximum likelihood phylogenetic tree inference (Price*et al.* , 2009) and Newick Display 1.6 for tree rendering (Junier and Zdobnov, 2010).

2.6. Domain prediction, protein representations and sequence alignment

Protein-protein BLAST (BLASTp-NCBI;*https://blast.ncbi.nlm.nih.gov/Blast.cgi*) using clustered nr database (nr database clustered at 90% identity and 90% sequence length) was performed to deepen into chitinase-related proteins. NCBI Batch Web CD-search tool (*https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi*) was used for protein domain prediction. FASTA sequences for each protein were upload to Batch CD-search using automatic search mode which directly launch live search as sequences were submitted explicitly in FASTA format. CD database (CDD) was selected to perform the search, setting 0.01 as statistical significance threshold (*E-value*). Protter (*http://wlab.ethz.ch/protter/start/)*was used as a plotting tool to graphically represent proteins (Omasits*et al.* , 2014). Multiple sequence alignment by CLUSTALW (*https://www.genome.jp/tools-bin/clustalw*) was performed to compare protein homology.

3. Results and discussion

3.1. Proteomic analysis in quinoa seeds harvested from irrigated and rainfed conditions

In this study, seed protein extracts harvested from quinoa plants grown in the field in rainfed and irrigated conditions were analyzed to identify, quantify, and estimate protein abundance and to compare protein enrichment between the two water regimes. After raw data collection, Proteome Discoverer 2.4^{TM} (Thermo Fisher Scientific, Massachusetts, United States) was used to quantify peptides comparing *C. quinoa* v1.0

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

data available at NCBI (*https://www.ncbi.nlm.nih.gov/)*. Detected proteins were filtered at a *p-adj < 0.05* , as shown in Supplementary Table S1 and biological samples distribution analyzed by principal component analysis (PCA) confirmed substantial variation between them (Supplementary Fig. S2). A total number of 2577 proteins were identified in seeds harvested from irrigated and rainfed conditions (Fig. 1A and Supplementary Table S2). When compared to the proteomic results obtained by Galindo-Luján et al., 2021 that used quinoa seeds from four different commercial varieties, in which 1211 proteins were identified by using LC-MS/MS, our current analysis yielded a significantly higher number of proteins. Besides, from the total of 2577 proteins identified, 2388 proteins (93% of the total) appeared in both conditions, 103 proteins (4% of the total) were exclusively found in seeds harvested from irrigated conditions, and 86 proteins (3% of the total) were exclusively identified in seeds harvested from rainfed conditions (Fig. 1A and Supplementary Table S2). Thus, although a great number of proteins were found in both water conditions, proteins exclusively represented in each condition depicted a low percentage from total.

To determine whether there were quantitative differences regarding protein overrepresentation in seeds harvested from irrigated or rainfed conditions, a differential statistical analysis of the proteins found in rainfed compared to irrigated conditions was performed to search for quantitative changes. As seen in figure 1B, the number of proteins overrepresented in seeds harvested from rainfed conditions (196 proteins) was higher than the number of proteins overrepresented in seeds harvested from irrigated conditions (142 proteins). Accordingly, the dot´s distribution in the volcano plot appeared shifted to the right side, which represents protein overabundance in rainfed conditions (log2(FC) [?] 1).

3.2. Seed Storage Proteins (SSP) and other seed related proteins

Among the shared proteins obtained from our study in both water conditions, different seed storage proteins (SSP) were found, including the 2S albumins and two 11S globulins, also known as chenopodins (Table 1). Both classes of proteins are the major storage proteins found in quinoa seeds, as described in previous works (Brinegar and Goundan, 1993; Burrieza *et al.* , 2019; Galindo-Lujan *et al.* , 2021). Interestingly, in addition to presenting substantial amounts of essential amino acids in their composition, chenopodins have been recently linked to anti-inflammatory properties in mice (Pompeu *et al.* , 2021). Alternatively, the SSP 2S albumin, one of the major proteins class found in quinoa seeds, as firstly described by Brinegar *et al.* , (1996), possesses significant contents of sulfur amino acids such as cysteine and also, histidine, and arginine. Both types of proteins have been identified in quinoa seed samples using new approaches based on shotgun proteomics (Galindo-Lujan *et al.* , (2021) and this work). In addition, several 7S globulins and 13S globulins (Table 1) appeared in the seed protein samples obtained from rainfed and irrigated conditions, and were also present in previous proteomic analysis carried out by Burrieza *et al.* , (2019). Since these SSP were consistently found in seeds obtained from different quinoa varieties, this, and previous studies, suggest a homogeneous and conserved distribution of SSP among different quinoa cultivars. Furthermore, our results confirm that the presence or abundance of these SSP does not vary depending on the water regime, rainfed or irrigated conditions, at least in seeds harvested from the quinoa cultivar used in this study (Supplementary Table S2).

Besides, seed oil body oleosins, a dehydrin family protein, late embryogenesis abundant (LEA) and LEA-related proteins, seed maturation family proteins and embryonic cell LEA-related proteins were found among the identified seed proteins obtained from both irrigated and rainfed conditions (Table 1), all of them related to characteristic desiccation and maturation processes occurring on seeds (Wang *et al.* , 2015; Rahman *et al.* , 2021).

3.3. Biological and functional significance of irrigated and rainfed quinoa seeds' proteomic profiles

In order to decipher the biological functions attributed to the proteins identified in quinoa seeds harvested from both irrigated and rainfed conditions, a gene ontology (GO) analysis was performed. Shared and exclusive protein enrichment was analyzed to evaluate Biological Process GO terms association (Fig. 2).

3.3.1. Biological Process GO terms enrichment in proteins annotated simultaneously in seeds

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

from irrigated and rainfed conditions

A total of 1960 proteins out of the 2388 shared proteins previously identified were associated with Biological Process (BP) GO terms (Fig. 2A). A large number of proteins were assigned to two main BP categories:*metabolic process* (GO:0008152; 1328 proteins) and *cellular process* (GO:0009987; 1476 proteins). Among them, there was a great number of GO terms related to the *primary* (GO:0044238; 963 proteins) and *organic substance metabolic process* (GO:0071704; 966 proteins), metabolic processes of nitrogenous compounds (*nitrogen compound metabolic process* GO:0006807, 732 proteins;*cellular nitrogen compound metabolic process* (GO:0034641, 478 proteins), *cellular metabolic process* (GO:0044237, 683 proteins), *biosynthetic process* (GO:0009058, 525 proteins),*cellular aromatic compound metabolic process* (GO:0006725; 356 proteins), and to the *heterocycle metabolic process* (GO:0046483; 356 proteins) (Fig. 2A). These categories were followed, in protein number, by the BP category *response to stimulus* (GO:0050896; 318 proteins) in which the *response to stress* (GO:0006950; 229 proteins) was the one presenting a larger protein number (Fig. 2A). The category *developmental process* (GO:0032502; 74 proteins) only involved GO terms related to *anatomical structure development*(GO:0048856; 73 proteins), detailed in Fig. 2A.

3.3.2. Biological Process GO terms enrichment of seed proteins from plants harvested from irrigated conditions

On one hand, a total of 81 proteins out of the 103 proteins exclusively found in irrigated conditions were associated to BP GO terms (Fig. 2B). BP GO categories such as *cellular biosynthetic process*(GO:0044249, 6 proteins), *cellular nitrogen compound metabolic process* (GO:0034641; 19 proteins), *cellular aromatic metabolic process* (GO:0006725; 17 proteins) and *heterocycle metabolic process* (GO:0046483; 17 proteins) belonging to *cellular metabolic process* (GO:0044237; 27 proteins), and also included in *cellular process* GO term (GO:0009987; 63 proteins, not detailed, as they were same categories), were categories exclusively present among proteins from seeds harvested from irrigated conditions. Also, metabolic processes such as *organic substance biosynthetic process*(GO:1901576; 6 proteins) and *cellular biosynthetic process*(GO:0044249; 6 proteins) were exclusively represented in samples obtained from this water condition. GO categories related to*localization* (GO:0051179; 20 proteins), *establishment of localization* (GO:0051234; 20 proteins) and *transport*(GO:0006810; 20 proteins) were unique for this water condition (Fig. 2B).

3.3.3. Biological Process GO terms enrichment of seed proteins from plants harvested from rainfed conditions

On the other hand, within the 86 proteins exclusively found in rainfed conditions, 81 were associated to BP GO terms (Fig. 2C). We remarkably found the subcategories *carbohydrate metabolic process*(GO:0005975; 7 proteins, not shown) and *protein metabolic process*(GO:0019538; 18 proteins, not shown), belonging to *primary metabolic process* (GO:0044238; 33 proteins), enriched in seeds under rainfed condition. Some proteins were assigned to the category*response to stimulus* (GO:0050896; 12 proteins) that fell out into the subcategories *response to stress* (GO:0006950; 12 proteins),*biotic stimulus* (GO:0009607; 4 proteins) and *external stimulus* (GO:0009605; 4 proteins), these last two categories only found under water limiting conditions (Fig. 2C).

Although proteins listed in the mentioned above sections were exclusive to each water condition, some of them fell into the same BP category (including GO terms assigned to *nitrogen compound metabolic process* , *catabolic process* , *organic substance metabolic process* , and *response to stress*). This result might imply that, despite being different proteins, they might share functionality (*eg* . AUR62024052 annotated as a peroxidase from rainfed seeds and AUR62013045 annotated as L-ascorbate peroxidase 3, were both classified into the *catabolic process* term, GO:0009056, respectively); or they can also belong to the same BP category without sharing similarities in their function (*eg* . AUR62006492 annotated as a mitogen-activated protein kinase 3 (MPK3) from rainfed seeds, and AUR62032691 annotated as a glutamate dehydrogenase B (GDHB), were both classified into exclusive rainfed or irrigated *response to stress*term, GO:0006950, respectively) (Supplementary Table S2). Therefore, these results suggest that although differences may appear in the protein that is synthesized, similar or dissimilar cellular or metabolic processes might have concurred.

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

3.3.4. GO terms showed differential enrichment of antioxidant-related proteins in overrepresented seed proteins from plants harvested from rainfed conditions

As the number of proteins found exclusively in each condition was limited, to deepen the understanding of possible mechanisms related to altered protein profile in seeds harvested from rainfed conditions, GO terms were assigned to proteins that showed statistically larger abundance in seeds harvested from rainfed conditions compared to irrigated conditions (Supplementary Tables S3 and S4). The GO analysis (including Biological Process, BP, Molecular Function, MF and Cellular Component, CC, terms) revealed interesting differences among water conditions (Fig. 3A-C and Supplementary Figs. S4-S9). Protein enrichment under irrigated conditions was found related to *transport*(GO:0006810) regarding BP-GO terms (Fig. 3A and Supplementary Fig. S4); and *protein binding* (GO:0005515), *nucleotide binding*(GO:0000166), *nucleic acid binding* (GO:0003676) and *DNA binding* (GO:0003677) within the enriched MF-GO terms (Fig. 3B and Supplementary Fig. S5). On the other hand, BP GO terms enrichment in seeds harvested from rainfed conditions presented a large number of overrepresented proteins related to *response to stress*(GO:0006950), *response to biotic stimulus* (GO:0009607),*response to external* (GO:0009605) and *endogenous stimulus*(GO:0009719), and *response to chemicals* (GO:0042221) (Fig. 3A and Supplementary Fig. S6). Along with this striking representation of proteins responding to stress and stimuli, a remarkable number of them were also related to *catabolic process* (GO:0009056),*carbohydrate metabolism* (GO:0005975) and *protein metabolic process* (GO:0019538). This enrichment coincided with MF GO terms involved in *binding* (GO:0005488), *hydrolase activity*(GO:0016787), and *catalytic activity* (GO:0003824) (Fig. 3B and Supplementary Fig. S7), standing out the importance of catalytic mechanisms triggered under rainfed conditions. These results are also supported by the enrichment of proteins under this conditions that fell into the *generation of precursor, metabolites, and energy*(GO:0006091) BP term.

A characteristic systemic drought-response mechanism in quinoa is the synthesis of reactive oxygen species (ROS) scavengers, together with the accumulation of osmolytes and antioxidants. Particularly those synthesized in the ornithine and raffinose pathways but also the accumulation of soluble sugars and proline, which also contribute to the cellular osmotic adjustment (reviewed by Bascunan-Godoy *et al.* , (2016) and Hinojosa *et al.* , (2018)). This enhanced accumulation of ROS detoxification enzymes has been recently described in 4-weeks old quinoa seedlings subjected to salinity stress (Ma *et al.* , 2021). In line with this, numerous antioxidant enzymes overrepresented in seeds harvested from rainfed conditions were categorized into *catabolic process* BP-GO term, such as L-ascorbate peroxidases (AUR62044027-RA, AUR62003342-RA), peroxidase (POD) (AUR62024052-RA), cytochrome C peroxidase (AUR62003343-RA), peroxidase C1C (AUR62026666-RA), peroxidase 4 (AUR62012343-RA, AUR62009723-RA), chatepsin B (AUR62001249-RA), plastidial pyruvate kinase 2 (AUR62021072-RA), peroxiredoxin-2E (AUR62037884-RA), fructose-bisphosphate aldolases 3 (AUR62033531-RA, AUR62028580-RA), glutathione S-transferase (GST) (AUR62008599-RA) and Cu/Zn superoxide dismutase (SOD) (AUR62000929-RA). Under water deficiency, plant tissues accumulate ROS (Apel and Hirt, 2004; Wang et al., 2015). As a consequence, plants respond by triggering ROS scavenging systems to avoid the oxidation of biomolecules that could hinder cellular homeostasis (Apel and Hirt, 2004). In our experiment, seeds from quinoa grown under rainfed conditions accumulated these types of enzymes. Similarly, other crops such as maize induce ROS scavenging enzymes (such as SOD, POD, and GST) as an early response mechanism to drought (Jiang *et al.* , 2019). Moreover, ROS molecules play a fine-tuning role in regulating seed dormancy release and germination, although they could trigger seed deterioration when produced in high concentrations causing DNA/RNA damage, lipid peroxidation or protein carbamylation (reviewed by Li *et al.* , 2022). Nonetheless, the regulatory mechanisms controlling ROS balance under stress are still not-well defined, although one can speculate that seeds promote dormancy to avoid tissue damage as a result of ROS accumulation, while the activation of ROS scavenging systems can be an effective response to reduce ROS concentrations when reaching extremely harmful levels.

Other enzymes differentially present in quinoa seeds harvested from rainfed conditions were the aspartic proteinases (AUR62006817-RA, AUR62000476-RA), nicastrin (AUR62040737-RA), cathepsin (AUR62001249-RA) and cysteine proteinase inhibitors (AUR62021845-RA, AUR62012808-RA), related to *protein metabolic process* BP-GO term. Moreover, enzymes such as the fructose-bisphosphate aldolase 3 (AUR62033531-

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

RA, AUR62028580-RA), the cytochrome C (AUR62027049–RA, AUR62027048-RA), the plastidial pyruvate kinase 2 (AUR62021072-RA), the NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 9 (AUR62010388-RA), the NADP isocitrate dehydrogenase (AUR62002238-RA), and the plastocyanin like domain (AUR62013468-RA, AUR62026803-RA) were overrepresented in seeds harvested from rainfed conditions. Overall, these results showed the accumulation of characteristic abiotic stress response proteins, antioxidant enzymes, and proteins involved in energy metabolism. Supporting these results it was previously observed that desiccation is also able to induce the accumulation of these types of proteins in tea (*Camellia sinensis*) recalcitrant seeds (desiccation sensitive seeds) (Chen *et al.* , 2011).

Intriguingly, our data revealed a pathogen-related protein overrepresented and exclusively present protein in rainfed conditions, the germin-like protein (GLP) AUR62037551-RA (Supplementary Table S3). GLPs genes are described to be induced in quinoa under *Trichoderma* symbiotic interaction (Rollano-Penaloza *et al.* , 2021). However, plant genomes contain a large number of GLPs copies with putative diverse enzymatic activities as SOD or ADP-glucose pyrophosphatase/phosphodiesterase (AGPPase) activities, in addition to their canonical function as oxalate oxidases (OXO) that increase their activity under abiotic stresses in plants (Davidson *et al.* , 2009).

In this regard, as previously mentioned, our proteomic study has revealed an enhanced protein accumulation of enzymatic strategies related to ROS scavenging and cellular detoxification alleviation in seeds harvested from rainfed conditions. Therefore, GLP enzymatic activity may contribute to this putative drought-avoidance strategy that quinoa seeds developed under water-deficient conditions.

Besides, the overrepresented seed proteins from rainfed conditions were preferentially assigned to CC-GO terms as *extracellular region*(GO:0005576), *mitochondrion* (GO:0005739), and *ribosome*(GO:0005840) (Fig. 3C and Supplementary Fig. S8), where peroxidases and other catabolic enzymes above described for BP and MF GO terms were also found. Under well-watered conditions, seed proteins that belong to the*membrane* (GO:0016020), *endoplasmic reticulum* (GO:0005783), and *Golgi Apparatus* (GO:0005794) terms were significantly higher (Fig. 3C and Supplementary Fig. S9). According to these findings, the major groups of seed proteins overrepresented under irrigated conditions and downregulated in rainfed conditions were heat shock proteins (AUR62015029-RA, AUR62017485-RA, AUR62014325-RA, AUR62021118-RA, AUR62035682-RA, and AUR62017128-RA) and calnexin homologs (AUR62032201-RA and AUR62036970-RA), among others (Supplementary Table S3). Interestingly, calnexins are proteins related to endoplasmic reticulum (ER) stress, that can be triggered by abiotic and biotic stresses (Qian *et al.* , 2015). However, the two calnexin homologs overrepresented in seeds harvested from irrigated conditions were also found in previous published works analysing quinoa seeds not subjected to stress (Galindo-Lujan *et al.* , 2021). Moreover, within the overrepresented seed proteins obtained from rainfed conditions, we identified another ER stress response protein, the somatic embryogenesis receptor kinase 1 (SERK1), AUR62018453-RA, which has been described as a co-receptor kinase linked to ER-associated degradation (ERAD), induced to alleviate ER stress in plants (Chen *et al.* , 2020) that, in our case, could be induced by drought in seeds.

3.4. Chitinase-related proteins were accumulated in quinoa seeds harvested from rainfed conditions

As previously mentioned, among the most represented GO categories that included overrepresented proteins in rainfed conditions were found*hydrolase* and *catalytic activities* , *catabolic process* and *carbohydrate metabolism* and *response to stress* . When analysing the proteins assigned to those GO terms, the protein family chitinase appeared to be predominant under water limiting conditions.

Chitinases are chitin hydrolases which are expressed in plants in response to biotic stresses, during plant development or in response to abiotic stresses (Grover, 2012). Seed chitinases seem to play multiple roles in seed germination and seedling establishment as part of the defence response against microbes (Gomez *et al.* , 2002). However, the specific functions that these proteins possess have been little explored.

Here, we identified 9 chitinase-related proteins overrepresented in seeds harvested from rainfed conditions (Fig. 4A) among the total number of 76 chitinase-related proteins found in *C. quinoa* genome v1.0 (Phy-

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

tozome v13). Based on their peptide sequences, a phylogenetic tree was obtained including the 25 peptide sequences of *Arabidopsis thaliana* chitinases previously described by Grover, 2012 (Fig. 4B).

Plant chitinases are divided into two main families, GH18 and GH19, based on their protein structure, which determines their catabolic activity (Kesari *et al.* , 2015). Also, in plants, there are numerous copies of Chitinase-Like Proteins (CLPs) that conserve one of these two types of protein structures. Although some CLPs have lost their chitin-binding domain, they have diversified their catalytic activities and could bind other substrates (Kesari *et al.,* 2015). Regarding the quinoa chitinases overrepresented in seeds harvested from rainfed conditions, we found both types, GH18-like (AUR62021380-RA and AUR62021381-RA) and GH19-like (AUR62027403-RA and AUR62023849-RA) chitinases (Table 2). Additionally, a conserved N-terminal Chitin Binding Domain (ChtBD) followed by a GH19-like domain have been described in four of the identified proteins (AUR62002379-RA, AUR62031322-RA, AUR62027403-RA and AUR62023809-RA) resembling typical class I GH19 plant chitinases; and a ChtBD *solo* peptide (AUR62003220-RA) was also found (Table 2).

In *Oryza sativa* (Kezuka *et al.* , 2010), *Bryum coronatum* (Taira *et al.* , 2011), and *Picea abies*(Ubhayasekera *et al.* , 2009) the GH19 chitinase family was the most predominant and well-characterized chitinase family found in these plant species. Also, GH19 chitinases have been reported as the most important family representing seed chitinases (Henrissat *et al.* , 1991) and, indeed, they were the most abundant chitinase type found in our study (AUR62027403-RA, AUR62023849-RA AUR62002379-RA, AUR62031322-RA, AUR62027403-RA, and AUR62023809-RA) (Fig. 5 and Table 2).

However, from the 9 chitinase-related proteins overrepresented in seeds harvested from rainfed conditions, only AUR62021380-RA (GH18-like) and AUR62023849-RA (GH19-like) were exclusively detected under such conditions. This result pointed these two chitinases-like proteins as potential candidates to be used as drought molecular markers for quinoa seeds.

The annotated chitinase-related proteins in *C. quinoa* genome v1.0 were obtained as homologous of the chitinases described in other organisms (Table 2). Moreover, the domain prediction performed in the chitinase-related proteins found in *C. quinoa* , based on the results yielded by the NCBI Batch Web CD-search tool (*https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi)*displayed GH18, GH19 and ChtBD domains characteristic of these protein family (Fig. 5 and Table 2), confirming their conserved sequence and their possible role as chitinases or CLPs in response to drought stress in quinoa seeds. In line with this, recent results from Rasouli *et al.,* (2021) showed proteome profiles of guard cells in quinoa in response to osmotic stress regulated by ABA signalling. Among the proteins overrepresented under salinity stress, Rasouli and collaborators found an increase in chitinases-related proteins, coinciding with two out of the four chitinases found in seeds harvested from rainfed conditions (AUR62021381-RA GH18-like chitinases and AUR62023809-RA GH19-like chitinases + ChtBD). Moreover, AUR62021381-RA is highly similar to ChiA superfamily chitinases, whose homolog in pepper (*CaChi2*) is able to increase the tolerance to osmotic stress when overexpressed in *A. thaliana* (Hong and Hwang, 2006).

It is worth mentioning that no chitinase-related proteins have been previously detected in shotgun proteomics in quinoa seeds (Burrieza*et al.* , 2019; Galindo-Lujan *et al.* , 2021), reinforcing the idea that the differential accumulation of chitinase-related proteins in quinoa seeds appear in response to water constraint, since none of the previously published proteomic experiments worked with seed harvested from water stress conditions. Therefore, even though plant chitinases seem to show tissue-specificity, as reported in other plant species such as sugar cane (Su *et al.* , 2015), similar abiotic stress signalling pathways could occur in different types of cells or tissues, giving rise to the importance of some specific chitinase-related proteins as proteins that participate in signal-transduction networks that operate under abiotic stress.

Interestingly, chitinase-like proteins identified in rainfed samples were grouped according to their functional domains based on their homology with the *A. thaliana* chitinases (Fig. 6). In *A. thaliana* , chitinase transcripts were notably upregulated in seedlings, leaves, shoots and roots subjected to different drought conditions (Grover, 2012). These results were also supported by the work performed by Rasheed *et al.* ,

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

(2016), in which the chitinase gene*AT2G43570* was highly upregulated under drought stress in shoots and roots. These *A. thaliana* chitinases were closely related to the quinoa chitinase-related proteins detected in rainfed seeds (Fig. 6), suggesting conserved roles in response to drought among this taxonomically distant plant species and in the different tissues analysed. Other studies have shown increments of endochitinase protein abundance during vegetative and flowering stages under drought stress in common bean (*Phaseolus vulgaris* L.) (Gupta *et al.* , 2019), similar to the accumulation of diverse endochitinase-like proteins (AUR62021381-RA, AUR62002379-RA, AUR62031322-RA, AUR62031316-RA, AUR62023809-RA) found in quinoa seeds harvested from rainfed conditions (Table 2). In addition, several GH19-like chitinases from Manchurian wild rice (*Zizania latifolia* L.) increased their expression under abiotic stresses (Zhou *et al.* , 2020) and the accumulation of plant chitinases was found in roots of barley, corn, pea, soybean, and beans in response to heavy metal toxicity (Bekesiova *et al.* , 2008). Other environmental stresses also induced the accumulation of chitinases in agronomically important species such as tomato, bromegrass, or blueberry (Ernst *et al* ., 1992; Chen *et al.* , 1994; Nakamura *et al.* , 2008; Kikuchi and Masuda, 2009). Likewise, the overexpression of the *CHITINASE 2* (*LcCHI2*) from wheatgrass (*Leymus chinensis*) in transgenic tobacco and maize plants showed an increased tolerance to saline-alkali stress (Liu*et al.* , 2020) and tea (*C. sinensis*) desiccation-sensitive (recalcitrant) seeds accumulate a homolog of AUR62023849 (AAX83263 orthologous in *Triticum aestivum*) under redox-status alteration (Chen *et al.* , 2011). These and our findings highlight the potential role of plant chitinases in alleviating the effects of different stressors, not only playing protective activities against pathogens but also becoming promising tools for plant engineering abiotic stress mitigation or drought stress biomarkers.

4. Conclusions and future perspectives

Proteomics is considered the most accurate and efficient -omic approach, over genomic and transcriptomic studies, to obtain biological information of plant tissue-specific and cellular status during plant growth and development. In this study, the impact of two contrasting water regimes (rainfed and irrigated conditions) in the field on*C. quinoa* Willd. seed proteomics was evaluated. A total of 2577 proteins were identified resulting in the most complete quinoa seed proteome published till date, highlighting the presence of characteristic seed proteins also found in other plant species such as LEA proteins, oleosins or SSPs as albumins or globulins, including the quinoa specific 11S globulin chenopodin (Brinegar and Goundan, 1993). Moreover, exclusive proteins for each water condition represented a low percentage from the total proteins identified. Statistically significant differentially abundant proteins were analysed to unravel differences between water treatments. GO terms associated to the overrepresented proteins in each water condition revealed variations in protein functions including the upregulation of proteins involved in catalytic processes under rainfed conditions. Among these interesting proteins, we found 9 chitinase-related proteins that were overrepresented under limiting water availability. These proteins are well characterized pathogenesis-related (PR) proteins that act degrading chitin in different organisms including plants, animals, or bacteria (Grover, 2012). Nonetheless, previous works have shown an induced chitinase activity or the upregulation of chitinase-related gene expression in many plants (including crops) when subjected to various abiotic stresses (Ernst *et al.* , 1992; Chen *et al.* , 1994; Hong and Hwang, 2006; Bekesiova *et al.* , 2008; Nakamura *et al.* , 2008; Kikuchi and Masuda, 2009; Grover, 2012; Rasheed *et al.* , 2016; Gupta *et al.,* 2019; Zhou *et al.* , 2020). Indeed, chitinases represent a huge family of proteins in plants, that include a great number of gene copies and evolutionary divergent sequences that have allowed them to acquire new functionalities resulting in emerging chitinase-like proteins (CLPs) that possess the ability to catalyse or bind different molecules other than chitin (Kesari *et al.* , 2015). In here, we described 9 chitinase-related proteins in quinoa seeds in response to drought stress. Two of them appeared exclusively in seeds harvested from rainfed conditions. Therefore, these findings could help improving our understanding regarding quinoa strategies that may contribute to improving its adaptation and survival under drought and, possibly, to other abiotic stresses. Moreover, the results here presented open the possibility of utilizing these proteins as plant stress biomarkers for quinoa seeds.

Acknowledgements

The authors greatly thank Susana Vilarino (Algosur, Spain) for providing the quinoa seeds that were used

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

in the field experiments performed in this study. The authors would also like to especially thank Dr Felipe Clemente and his team at the Proteomic Unit from Biological Techniques Research Support Centre (CAI) at the Universidad Complutense de Madrid (Madrid, Spain) for their technical assistance and great support in the data analysis performed.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Abugoch James LE . 2009. *Quinoa (Chenopodium quinoa Willd.): Composition, chemistry, nutritional, and functional properties* . Elsevier Inc.

Alandia G, Rodriguez JP, Jacobsen SE, Bazile D, Condori B . 2020. Global expansion of quinoa and challenges for the Andean region. Global Food Security **26** , 100429.

Alexa A, Rahnenfuhrer J . 2022. topGO: Enrichment analysis for gene ontology. R package version 2.48.0.

Ando H, Chen YC, Tang H, Shimizu M, Watanabe K, Mitsunaga T . 2002. Food Components in Fractions of Quinoa Seed. Food Science and Technology Research **8** , 80–84.

Angeli V, Silva PM, Massuela DC, Khan MW, Hamar A, Khajehei F, Grae S, Piatti C . 2020. Quinoa (Chenopodium quinoa Willd.): An Overview of the Potentials of the "Golden Grain" and Socio-Economic and Environmental Aspects of Its Cultivation and Marketization. Foods**9** .

Apel K, Hirt H . 2004. Reactive oxygen species: Metabolism, oxidative stress, and signal transduction. Annual Review of Plant Biology **55** , 373–399.

Araus JL . 2004. The problems of sustainable water use in the Mediterranean and research requirements for agriculture. Annals of Applied Biology **144** , 259–272.

Bascunan-Godoy L, Reguera M, Abdel-Tawab YM, Blumwald E . 2016. Water deficit stress-induced changes in carbon and nitrogen partitioning in Chenopodium quinoa Willd. Planta **243** , 591–603.

Bazile D, Bertero D, Nieto C . 2015. FAO & CIRAD. 2015. State of the Art Report of Quinoa in the World in 2013

Bazile D, Pulvento C, Verniau A, *et al.* 2016. Worldwide evaluations of quinoa: Preliminary results from post international year of quinoa FAO projects in nine countries. Frontiers in Plant Science**7** .

Bekesiova B, Hraška Š, Libantová J, Moravčíková J, Matušíková I . 2008. Heavy-metal stress induced accumulation of chitinase isoforms in plants. Molecular Biology Reports **35** , 579–588.

Ben-Amar A, Allel D, Mliki A . 2022. Up-regulation of a stress-responsive endochitinase VvChit-IV in grapevine cell cultures improves in vitro stress tolerance. Protoplasma.

Brinegar C, Goundan S . 1993. Isolation and Characterization of Chenopodin, the 11S Seed Storage Protein of Quinoa (Chenopodium quinoa). Journal of Agricultural and Food Chemistry **41** , 182–185.

Brinegar C, Sine B, Nwokocha L . 1996. High-Cysteine 2S Seed Storage Proteins from Quinoa (Chenopodium quinoa) . Journal of Agricultural and Food Chemistry **44** , 1621–1623.

Burrieza HP, Rizzo AJ, Moura Vale E, Silveira V, Maldonado S . 2019. Shotgun proteomic analysis of quinoa seeds reveals novel lysine-rich seed storage globulins. Food Chemistry **293** , 299–306.

Capriotti AL, Cavaliere C, Piovesana S, Stampachiacchiere S, Ventura S, Zenezini Chiozzi R, Laganà A . 2015. Characterization of quinoa seed proteome combining different protein precipitation

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

techniques: Improvement of knowledge of nonmodel plant proteomics. Journal of Separation Science **38** , 1017–1025.

Chase MW, Christenhusz MJM, Fay MF, *et al.* 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Botanical Journal of the Linnean Society **181** , 1–20.

Chen Q, Yang L, Ahmad P, Wan X, Hu X . 2011. Proteomic profiling and redox status alteration of recalcitrant tea (Camellia sinensis) seed in response to desiccation. Planta **233** , 583–592.

Chen RD, Yu LX, Greer AF, Cheriti H, Tabaeizadeh Z . 1994. Isolation of an osmotic stress- and abscisic acid-induced gene encoding an acidic endochitinase from Lycopersicon chilense. Molecular & general genetics : MGG **245** , 195–202.

Chen Q, Yu F, Xie Q . 2020. Insights into endoplasmic reticulum-associated degradation in plants. New Phytologist**226** , 345–350.

Choukr-Allah R, Rao NK, Hirich A, Shahid M, Alshankiti A, Toderich K, Gill S, Butt KUR . 2016. Quinoa for marginal environments: Toward future food and nutritional security in MENA and central Asia regions. Frontiers in Plant Science **7** .

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M . 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21** , 3674–3676.

Criscuolo A, Gribaldo S . 2010. BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evolutionary Biology **10** .

Davidson RM, Reeves PA, Manosalva PM, Leach JE . 2009. Germins: A diverse protein family important for crop improvement. Plant Science**177** , 499–510.

Ernst D, Schraudner M, Langebartels C, Sandermann HJ . 1992. Ozone-induced changes of mRNA levels of beta-1,3-glucanase, chitinase and 'pathogenesis-related' protein 1b in tobacco plants. Plant molecular biology **20** , 673–682.

Farooq M, Wahid A, Kobayashi N, Fujita D, Basra SMA . 2009. Plant Drought Stress: Effects, Mechanisms and Management. Agronomy for sustainable Development **29** , 185–212.

Galindo-Lujan R, Pont L, Minic Z, Berezovski M V., Sanz-Nebot V, Benavente F . 2021. Characterization and differentiation of quinoa seed proteomes by label-free mass spectrometry-based shotgun proteomics. Food Chemistry **363** .

Gomez L, Allona I, Casado R, Aragoncillo C . 2002. Seed chitinases. Seed Science Research **12** , 217–230.

Gonzalez JA, Eisa SSS, Hussin SAES, Prado FE . 2015. Quinoa: An Incan Crop to Face Global Changes in Agriculture. Quinoa: Improvement and Sustainable Production, 1–18.

Gordillo-Bastidas E, Diaz-Rizzolo D, Roura E, Massanes T, Gomis R . 2016. Quinoa (Chenopodium quinoa Willd), from Nutritional Value to Potential Health Benefits: An Integrative Review. Journal of Nutrition & Food Sciences **06** .

Granado-Rodriguez S, Aparicio N, Matias J, *et al.* 2021. Studying the Impact of Different Field Environmental Conditions on Seed Quality of Quinoa: The Case of Three Different Years Changing Seed Nutritional Traits in Southern Europe. Frontiers in Plant Science**12** , 1–21.

Grenfell-Shaw L, Tester M . 2021. Abiotic Stress Tolerance in Quinoa.139–167.

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Grover A . 2012. Plant Chitinases: Genetic Diversity and Physiological Roles. Critical Reviews in Plant Sciences **31** , 57–73.

Guo H, Hao Y, Yang X, Ren G, Richel A . 2021. Exploration on bioactive properties of quinoa protein hydrolysate and peptides: a review. Critical Reviews in Food Science and Nutrition **0** , 1–14.

Gupta N, Zargar SM, Salgotra RK, Dar TA . 2019. Identification of drought stress-responsive proteins in common bean. Journal of Proteins and Proteomics **10** , 45–53.

Henrissat B, Vegetales M, Grenoble F- . 1991. A classification of glycosyl hydrolases based sequence similarities amino acid.**280** , 309–316.

Hinojosa L, Gonzalez JA, Barrios-Masias FH, Fuentes F, Murphy KM . 2018. Quinoa abiotic stress responses: A review. Plants **7** .

Hong JK, Hwang BK . 2006. Promoter activation of pepper class II basic chitinase gene , CAChi2 , and enhanced bacterial disease resistance and osmotic stress tolerance in the CAChi2-overexpressing Arabidopsis. , 433–448.

Jacobsen SE, Jensen CR, Liu F . 2013. Improving Crop Production in the Arid Mediterranean Climate. Improving Water and Nutrient-Use Efficiency in Food Production Systems, 187–209.

Janssen F, Pauly A, Rombouts I, Jansens KJA, Deleu LJ, Delcour JA . 2017. Proteins of Amaranth (Amaranthus spp.), Buckwheat (Fagopyrum spp.), and Quinoa (Chenopodium spp.): A Food Science and Technology Perspective. Comprehensive Reviews in Food Science and Food Safety**16** , 39–58.

Jarvis DE, Ho YS, Lightfoot DJ, *et al.* 2017. The genome of Chenopodium quinoa. Nature **542** , 307–312.

Jiang Z, Jin F, Shan X, Li Y . 2019. iTRAQ-Based Proteomic Analysis Reveals Several Strategies to Cope with Drought Stress in Maize Seedlings. , 1–17.

Junier T, Zdobnov EM . 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics **26** , 1669–1670.

Katoh K, Standley DM . 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution **30** , 772–780.

Kesari P, Patil DN, Kumar P, Tomar S, Sharma AK, Kumar P . 2015. Structural and functional evolution of chitinase-like proteins from plants. Proteomics **15** , 1693–1705.

Kezuka Y, Kojima M, Mizuno R, Suzuki K, Watanabe T, Nonaka T . 2010. Structure of full-length class I chitinase from rice revealed by X-ray crystallography and small-angle X-ray scattering. , 2295–2305.

Kikuchi T, Masuda K . 2009. Scientia Horticulturae Class II chitinase accumulated in the bark tissue involves with the cold hardiness of shoot stems in highbush blueberry (Vaccinium corymbosum L .). **120** , 230–236.

Laemmli UK . 1970. Cleavage of Structura Proteins during the Assembly of the Head of Bacteriophage T3. Nature Publishing Group**227** , 680–685.

Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S, Gascuel O . 2019. NGPhylogeny.fr: New generation phylogenetic services for non-specialists. Nucleic Acids Research**47** , W260–W265.

Li H, Greene LH . 2010. Sequence and structural analysis of the chitinase insertion domain reveals two conserved motifs involved in chitin-binding. PLoS ONE **5** .

Li W, Niu Y, Zheng Y, Wang Z . 2022. Advances in the Understanding of Reactive Oxygen Species-Dependent Regulation on Seed Dormancy, Germination, and Deterioration in Crops. Frontiers in Plant

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Science **13** , 1–9.

Liu X, Yu Y, Liu Q, *et al.* 2020. A Na2CO3-Responsive Chitinase Gene From Leymus chinensis Improve Pathogen Resistance and Saline-Alkali Stress Tolerance in Transgenic Tobacco and Maize. Frontiers in Plant Science **11** , 1–12.

Ma Q, Su C, Dong C-H . 2021. Genome-Wide Transcriptomic and Proteomic Exploration of Molecular Regulations in Quinoa Responses to Ethylene and Salt Stress. Plants **10** .

Nakamura T, Ishikawa M, Nakatani H, Oda A . 2008. Characterization of Cold-Responsive Extracellular Chitinase in Bromegrass Cell Cultures and Its Relationship to Antifreeze Activity. Plant Physiology **147** , 391–401.

Omasits U, Ahrens CH, Muller S, Wollscheid B . 2014. Protter: Interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics **30** , 884–886.

Oyeleye A, Normi YM . 2018. Chitinase: Diversity, limitations, and trends in Engineering for suitable applications. Bioscience Reports**38** , 1–21.

Pompeu DG, Cordeiro HG, Tonelli FCP, *et al.* 2021. Chenopodin as an anti-inflammatory compound. Natural Product Research, 1–4.

Price MN, Dehal PS, Arkin AP . 2009. Fasttree: Computing large minimum evolution trees with profiles instead of a distance matrix. Molecular Biology and Evolution **26** , 1641–1650.

Pulvento C, Riccardi M, Lavini A, D'Andria R, Iafelice G, Marconi E . 2010. Field Trial Evaluation of Two Chenopodium quinoa Genotypes Grown Under Rain-Fed Conditions in a Typical Mediterranean Environment in South Italy. Journal of Agronomy and Crop Science**196** , 407–411.

Qian D, Tian L, Qu L . 2015. Proteomic analysis of endoplasmic reticulum stress responses in rice seeds. Scientific Reports **5** , 1–15.

R Fundation for Statistical Computing . 2020. R Core Team (2020). European Environment Agency.

Rahman M, Guo Q, Baten A, Mauleon R, Khatun A, Liu L, Barkla BJ . 2021. Shotgun proteomics of Brassica rapa seed proteins identifies vicilin as a major seed storage protein in the mature seed. PLoS ONE**16** , 1–23.

Rasheed S, Bashir K, Matsui A, Tanaka M, Seki M . 2016. Transcriptomic analysis of soil-grown arabidopsis thaliana roots and shoots in response to a drought stress. Frontiers in Plant Science**7** .

Rasouli F, Kiani-Pouya A, Shabala L, *et al.* 2021. Salinity effects on guard cell proteome in chenopodium quinoa. International Journal of Molecular Sciences **22** , 1–22.

Rojas W, Pinto M, Soto JL . 2010. *Distribucion geografica y variabilidad genetica de los granos andinos* .

Rollano-Penaloza OM, Mollinedo PA, Widell S, Rasmusson AG . 2021. Transcriptomic Analysis of Quinoa Reveals a Group of Germin-Like Proteins Induced by Trichoderma. Frontiers in Fungal Biology **2** , 1–14.

Santos P, Fortunato A, Ribeiro A, Pawlowski K . 2008. Chitinases in root nodules. Plant Biotechnology **25** , 299–307.

Singh A, Subudhi E . 2014. Expression of a chitinase family protein at4g01700 from Arabidopsis thaliana. Journal of Chemical and Pharmaceutical Sciences (ISSN:0974-2115), 23–30.

Su Y, Xu L, Wang S, Wang Z, Yang Y, Chen Y, Que Y . 2015. Identification, phylogeny, and transcript of chitinase family genes in sugarcane. Scientific Reports **5** , 1–15.

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Taira T, Mahoe Y, Kawamoto N, Onaga S, Iwasaki H, Ohnuma T, Fukamizo T . 2011. Cloning and characterization of a small family 19 chitinase from moss (Bryum coronatum). Glycobiology **21** , 644–654.

Takenaka Y, Nakano S, Tamoi M, Sakuda S, Fukamizo T . 2009. Chitinase gene expression in response to environmental stresses in arabidopsis thaliana: Chitinase inhibitor allosamidin enhances stress tolerance. Bioscience, Biotechnology and Biochemistry **73** , 1066–1071.

Tang CM, Chye ML, Ramalingam S, Ouyang SW, Zhao KJ, Ubhayasekera W, Mowbray SL . 2004. Functional analyses of the chitin-binding domains and the catalytic domain of Brassica juncea chitinase BjCHI1. Plant Molecular Biology **56** , 285–298.

Tramblay Y, Koutroulis A, Samaniego L, *et al.* 2020. Challenges for drought assessment in the Mediterranean region under future climate scenarios. Earth-Science Reviews **210** , 103348.

Tyler L, Bragg JN, Wu J, Yang X, Tuskan GA, Vogel JP . 2010. Annotation and comparative analysis of the glycoside hydrolase genes in Brachypodium distachyon. BMC Genomics **11** .

Ubhayasekera W, Rawat R, Ho SWT, Wiweger M, Arnold S Von, Chye M-L, Mowbray SL . 2009. The first crystal structures of a family 19 class IV chitinase : the enzyme from Norway spruce. Plant Molecular Biology **71** , 277–289.

Vega-Galvez A, Miranda M, Vergara J, Uribe E, Puente L, Martinez EA . 2010. Nutrition facts and functional potential of quinoa (Chenopodium quinoa willd.), an ancient Andean grain: A review. Journal of the Science of Food and Agriculture **90** , 2541–2547.

Wang WQ, Liu SJ, Song SQ, Moller IM . 2015. Proteomics of seed development, desiccation tolerance, germination and vigor. Plant Physiology and Biochemistry **86** , 1–15.

Yasui Y, Hirakawa H, Oikawa T, *et al.* 2016. Draft genome sequence of an inbred line of Chenopodium quinoa, an allotetraploid crop with great environmental adaptability and outstanding nutritional properties. DNA Research **23** , 535–546.

Zhang H, Zhu J, Gong Z, Zhu JK . 2022. Abiotic stress responses in plants. Nature Reviews Genetics **23** , 104–119.

Zhou N, An Y, Gui Z, Xu S, He X, Gao J, Zeng D, Gan D, Xu W . 2020. Identification and expression analysis of chitinase genes in Zizania latifolia in response to abiotic stress. Scientia Horticulturae**261** .

Zhu JK . 2016. Abiotic Stress Signaling and Responses in Plants. Cell **167** , 313–324.

Zou C, Chen A, Xiao L, *et al.* 2017. A high-quality genome assembly of quinoa provides insights into the molecular basis of salt bladder-based salinity tolerance and the exceptional nutritional value. Cell Research **27** , 1327–1340.

Tables

Table 1 . SSPs and other characteristic seed-related proteins simultaneously found in seeds harvested from irrigated and rainfed conditions.

C. quinoa **ID**	**Description**
AUR62015663-RA	2S albumin
AUR62020540-RA	2S albumin
AUR62011869-RA	11S globulin (chenopodin)
AUR62024712-RA	11S globulin (chenopodin)
AUR62028591-RA	7S globulin
AUR62032318-RA	7S globulin

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

C. quinoa **ID**	**Description**
AUR62034727-RA	7S globulin
AUR62033661-RA	7S globulin
AUR62015569-RA	13S globulin
AUR62022853-RA	Seed oil body oleosin
AUR62012221-RA	Seed oil body oleosin
AUR62040213-RA	Seed oil body oleosin
AUR62008167-RA	Seed oil body oleosin
AUR62036943-RA	Seed oil body oleosin
AUR62002243-RA	Seed oil body oleosin
AUR62004102-RA	Dehydrin family protein
AUR62011287-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62034707-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62043549-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62032331-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62028605-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62028603-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62014787-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62002497-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62023689-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62018728-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62007271-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62014840-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62017037-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62011567-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62022650-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62037387-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62042308-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62002551-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62029965-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62012039-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62022623-RA	Late Embryogenesis Abundant (LEA) and LEA-related protein
AUR62032329-RA	Seed Maturation family protein
AUR62028604-RA	Seed Maturation family protein
AUR62037914-RA,	Embryonic Cell LEA-related protein
AUR62040165-RA	Embryonic Cell LEA-related protein

Table 2 . Results from NCBI Batch Web CD-search tool for protein domain prediction and *C. quinoa* v1.0 annotation from Phytozome13 (From. . . to: range of amino acids in the query protein sequence to which the domain model aligns; E-Value: expected value, statistical significance of the hit as the likelihood the hit was found by chance; Accession: accession number of the hit, cd = conserved domain from NCBI, cl = superfamily cluster; Superfamily: specific accession number of the superfamily to which the domain model belongs; Short name: defining name for the conserved domain). Underlined AUR codes show chitinase-related proteins exclusively identified in seeds harvested from rainfed conditions.

Query	**Description** *C. Quinoa* **v1.0**	**From**	**To**	**E-Value**	**Accession**	**Superfamily**	**Short name**

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

AUR62021380-RA	Acidic mammalian chitinase	123	246	2,61E-06	cd02879	cl10447	GH18_plant_chitinase_class_V
AUR62021381-RA	Endochitinase 46	27	355	5,61E-116	cd02879	cl10447	GH18_plant_chitinase_class_V
		70	351	1,11E-33	cl34587	-	ChiA superfamily
AUR62002379-RA	Acidic endochitinase SP2	83	282	2,21E-80	cd00325	cl00222	chitinase_GH19
		83	282	2,21E-80	cl00222	-	Lyz-like superfamily
		32	59	1,72E-02	cd00035	cl16916	ChtBD1
AUR62031322-RA	Endochitinase EP3	75	270	6,24E-77	cd00325	cl00222	chitinase_GH19
		75	270	6,24E-77	cl00222	-	Lyz-like superfamily
		29	52	1,24E-02	cd00035	cl16916	ChtBD1
AUR62031316-RA	Basic endochitinase C	47	242	4,98E-77	cd00325	cl00222	chitinase_GH19
		47	242	4,98E-77	cl00222	-	Lyz-like superfamily
AUR62027403-RA	Chitinase 4	76	271	5,92E-76	cd00325	cl00222	chitinase_GH19
		76	271	5,92E-76	cl00222	-	Lyz-like superfamily
		29	53	2,82E-02	cd00035	cl16916	ChtBD1
AUR62003220-RA	Antimicrobial peptide 2	31	52	8,46E-02	cd00035	cl16916	ChtBD1
AUR62023809-RA	Acidic endochitinase SP2	84	283	5,58E-79	cd00325	cl00222	chitinase_GH19
		84	283	5,58E-79	cl00222	-	Lyz-like superfamily
		32	59	1,50E-02	cd00035	cl16916	ChtBD1
AUR62023849-RA	Chitinase 3	10	240	1,67E-105	cd00325	cl00222	chitinase_GH19
		10	240	7,03E-132	cl00222	-	Lyz-like superfamily

Figure legends

Fig. 1. Total proteins quantified in quinoa seeds harvested from irrigated and rainfed conditions. A, From the 2577 identified proteins in quinoa seeds, 103 appeared, exclusively in seeds harvested

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

from irrigated conditions and 86 appeared exclusively in seeds harvested from rainfed conditions. **B,** Volcano plot representing all the identified proteins in seeds harvested from irrigated and rainfed conditions. Different colours show two-fold statistically significant overrepresented proteins for each condition (| $\log_2$FC [?] 1 |, *p-adj* [?] *0.05* ; n = 3). Red dots: rainfed conditions; blue dots: irrigated conditions; black dots: no statistically significant differentiated proteins.

Fig. 2. Seed proteins harvested from irrigated and rainfed conditions classified by gene ontology (GO) terms related to Biological Process (BP). A, From the total of 2388 proteins quantified which were found in both conditions, 1960 were associated to BP-GO terms. In the graphs, the widest GO terms related to BP identified for these 1960 proteins were represented. **B,** From the 103 proteins quantified exclusively under irrigated conditions, 81 were associated to BP-GO terms. **C,** From the 86 proteins quantified exclusively under rainfed conditions, 81 were associated to BP-GO terms.

Fig. 3. Gene ontology (GO) annotation of overrepresented proteins in seeds harvested from irrigated and rainfed conditions . The graph represents the number of statistically significant overrepresented proteins in seeds harvested from rainfed and irrigated conditions (| $\log_2$FC [?] 1 |, *p-adj [?] 0.05* ; n = 3) assigned to Biological Process (BP) (**A**), Molecular Function (MF) or (**B**) Cellular Component (CC) (**C**) GO categories. From the total of 196 overrepresented proteins in seeds harvested from rainfed compared to irrigated conditions, 170 were assigned to GO terms that belong to the categories Biological Process (BP), Molecular Function (MF) or Cellular Component (CC). Seed proteins from irrigated conditions samples yielded 126 proteins, from a total of 142, that were assigned to BP, MF and CC-GO terms.

Fig. 4. Phylogenetic trees of chitinase-like proteins identified in quinoa. A, Chitinase-like proteins found in seeds harvested from rainfed conditions in *C. quinoa.* **B,** Phylogenetic tree containing 76 chitinase-like proteins annotated in *C. quinoa*genome v1.0 (Phytozome v13). Their peptide similarity was analysed including the 25 chitinases described in model plant *A. thaliana* , grouped according to their functional domains using NGPhylogeny.fr.

Fig. 5. Conserved domain (CD) prediction for the 9 chitinase-related proteins identified in quinoa seeds under rainfed conditions. Representation of predicted conserved domains for the 9 chitinase-related proteins identified in quinoa seeds harvested from rainfed conditions, using Protter (http://wlab.ethz.ch/protter/start/).

Fig. 6. Phylogenetic tree of *A. thaliana* chitinases and quinoa chitinase-related proteins found in seeds under rainfed condition. Twenty-five chitinases have been identified in *A. thaliana* which are mainly divided into three groups based on their catalytic and binding domains. The 9 chitinase-like proteins overrepresented under rainfed conditions in the quinoa seeds analyzed showed sequence similarities to protein domains of *A. thaliana*ones. In addition, quinoa chitinase-like proteins were closer to the ones that were highly expressed in a microarray data from leaves and seedlings of *A. thaliana* grown under drought conditions from two independent experiments summarized in Groover, 2012.

Supporting information

Table S1. Identification of total peptides from quinoa seeds (irrigated and rainfed conditions).

Table S2. Annotation of shared and exclusive proteins of quinoa seeds (irrigated and rainfed conditions).

Table S3. Overrepresented proteins of quinoa seeds harvested from irrigated conditions.

Table S4. Overrepresented proteins of quinoa seeds harvested from rainfed conditions.

Supplementary Fig. S1 . Rainfall and temperatures registered in irrigated and rainfed field experimental areas.

Supplementary Fig. S2 . Principal component analysis (PCA) of three biological replicates for each condition.

Supplementary Fig. S3 . Heat map representation for shared and exclusive proteins found in each replicate for seeds harvested from irrigated and rainfed conditions.

Supplementary Fig. S4 . Hierarchical Biological Process GO terms of overrepresented proteins in seeds under irrigated conditions.

Supplementary Fig. S5 . Hierarchical Molecular Function GO terms of overrepresented proteins in seeds under irrigated conditions.

Supplementary Fig. S6 . Hierarchical Biological Process GO terms of overrepresented proteins in seeds under rainfed conditions.

Supplementary Fig. S7 . Hierarchical Molecular Function GO terms of overrepresented proteins in seeds under rainfed conditions.

Supplementary Fig. S8 . Hierarchical Cellular Component GO terms of overrepresented proteins in seeds under rainfed conditions.

Supplementary Fig. S9 . Hierarchical Cellular Component GO terms of overrepresented proteins in seeds under irrigated conditions.

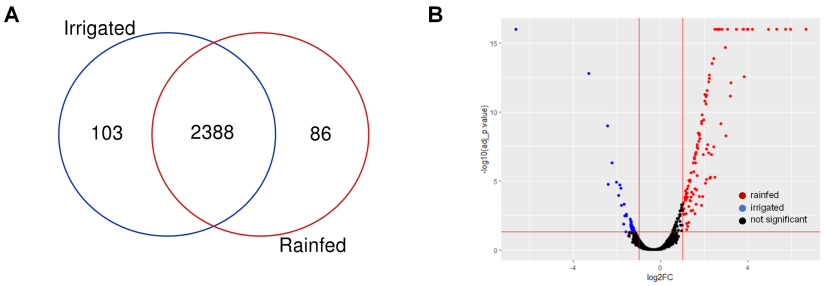

Figure 1. Total proteins quantified in quinoa seeds harvested from irrigated and rainfed conditions. A, From 2577 quantified proteins in seeds from both conditions, 103 appeared, exclusively, in seeds harvested from irrigated conditions and 86 appeared, exclusively, in seeds harvested from rainfed conditions. **B,** Volcano plot representing all proteins quantified in seeds harvested from irrigated and rainfed conditions. Different colours show two-fold statistically significant overrepresented proteins for each condition ($|\log_2 FC \geq 1|$, p-adj ≤ 0.05; n = 3). Red dots: rainfed conditions; blue dots: irrigated conditions; black dots: no statistically significant differentiated proteins.

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

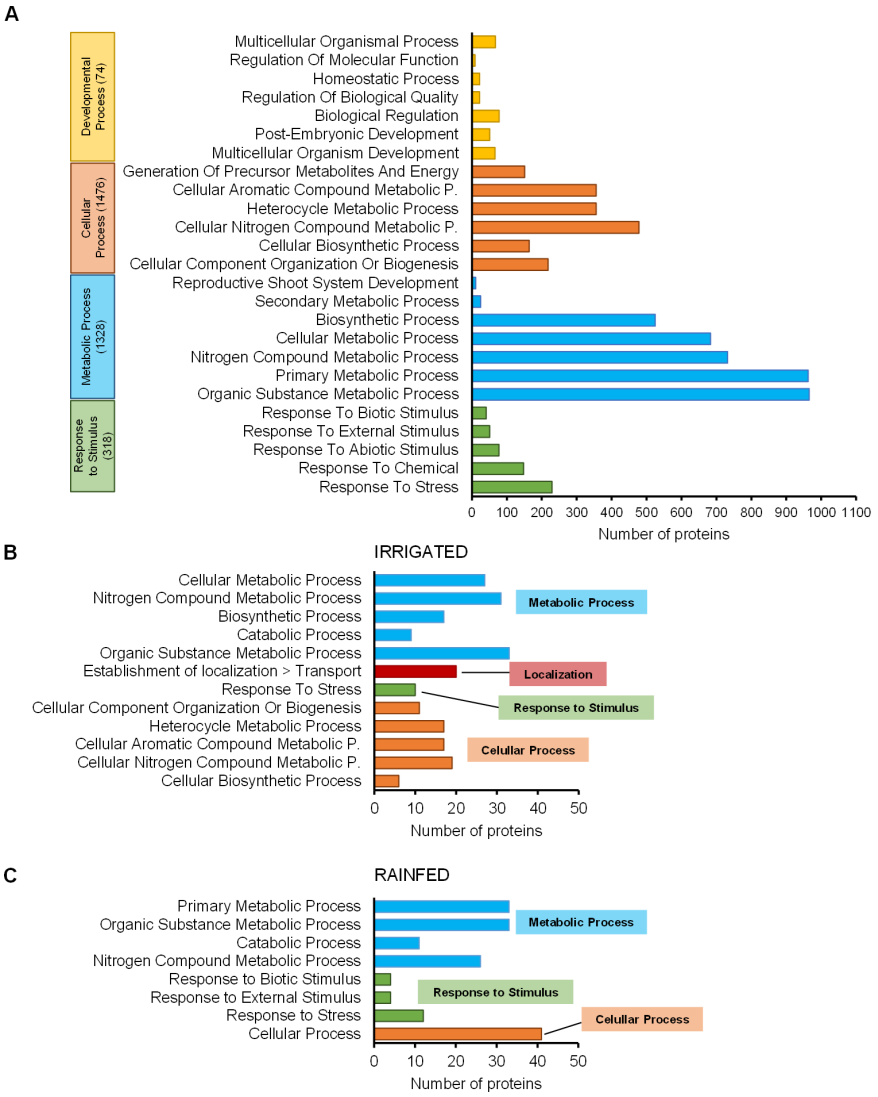

Figure 2. Proteins quantified in seeds harvested from irrigated and rainfed conditions classified by gene ontology (GO) terms related to Biological Process (BP). A, From the total of 2388 proteins quantified which were common in both conditions, 1960 were associated to BP-GO terms. In the graphs, the widest GO terms related to BP identified for these 1960 proteins were represented. **B,** From 103 proteins quantified exclusively under irrigated conditions, 81 were associated to BP-GO terms. **C,** From 86 proteins quantified exclusively under rainfed conditions, 81 were associated to BP-GO terms.

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

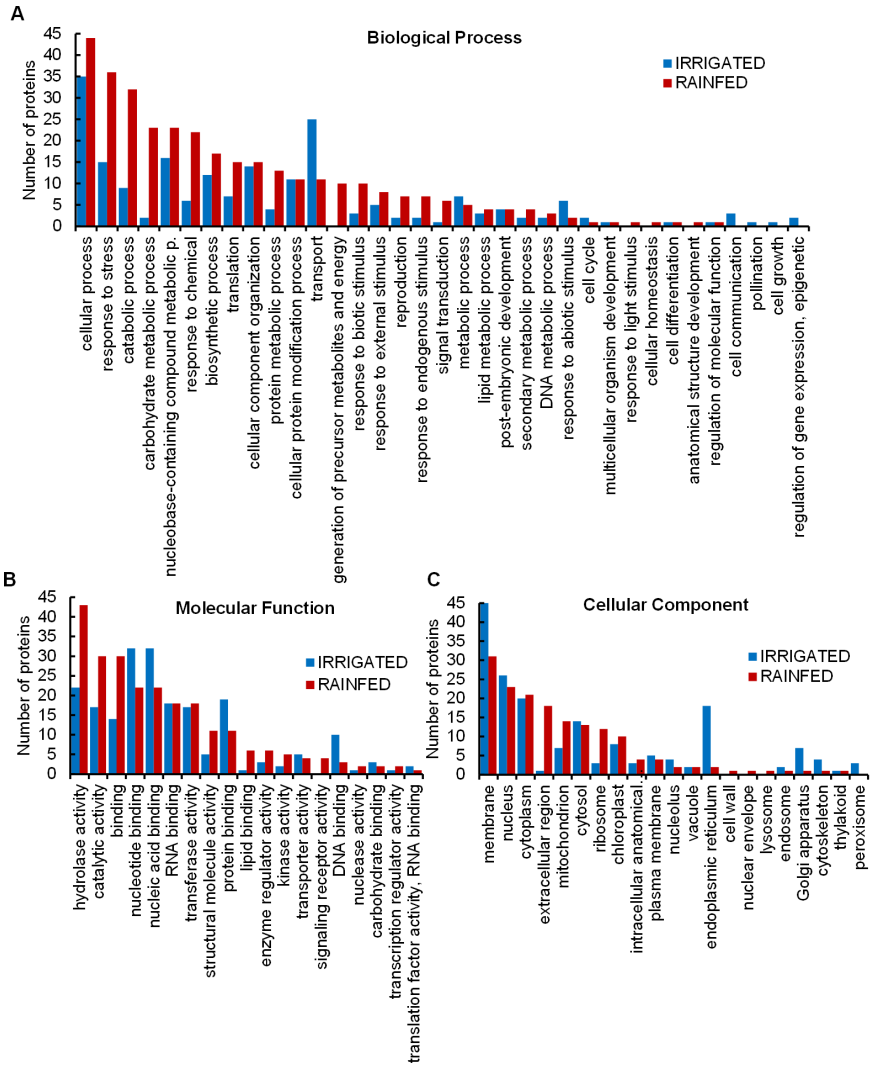

Figure 3. Gene ontology (GO) annotation of overrepresented proteins in seeds harvested from irrigated and rainfed conditions. The graph represents number of statistically significant overrepresented proteins in seeds harvested from rainfed and irrigated conditions (| $\log_2$FC ≥ 1 |, p-adj ≤ 0.05; n = 3) assigned to Biological Process (**A**), Molecular Function (**B**) and Cellular Component (**C**) GO categories. From the total of 196 overrepresented proteins in seeds harvested from rainfed compared to irrigated conditions, 170 were assigned to GO terms regarding Biological Process (BP), Molecular Function (MF) and Cellular Component (CC). For proteins from seeds harvested from irrigated conditions, 126 proteins from a total of 142 were assigned to BP, MF and CC-GO terms.

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

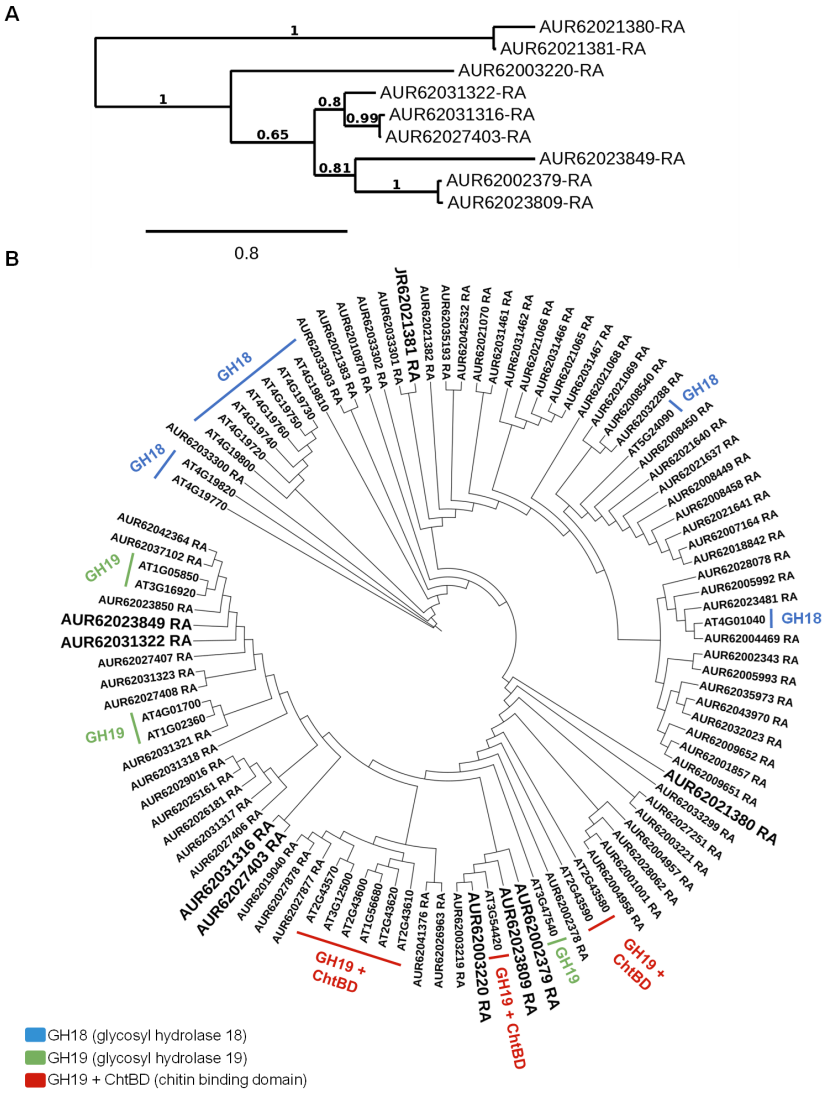

Figure 4. Phylogenetic trees of chitinase-like proteins identified in quinoa. A, Chitinase-like proteins found in seeds harvested from rainfed conditions in *C. quinoa*. **B,** Phylogenetic tree containing 76 chitinase-like proteins annotated in *C. quinoa* genome v1.0 (Phytozome v13). Their peptide similarity were analysed including 25 chitinases described in model plant *A. thaliana*, grouping in order to their functional domains with NGPhylogeny.fr.

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

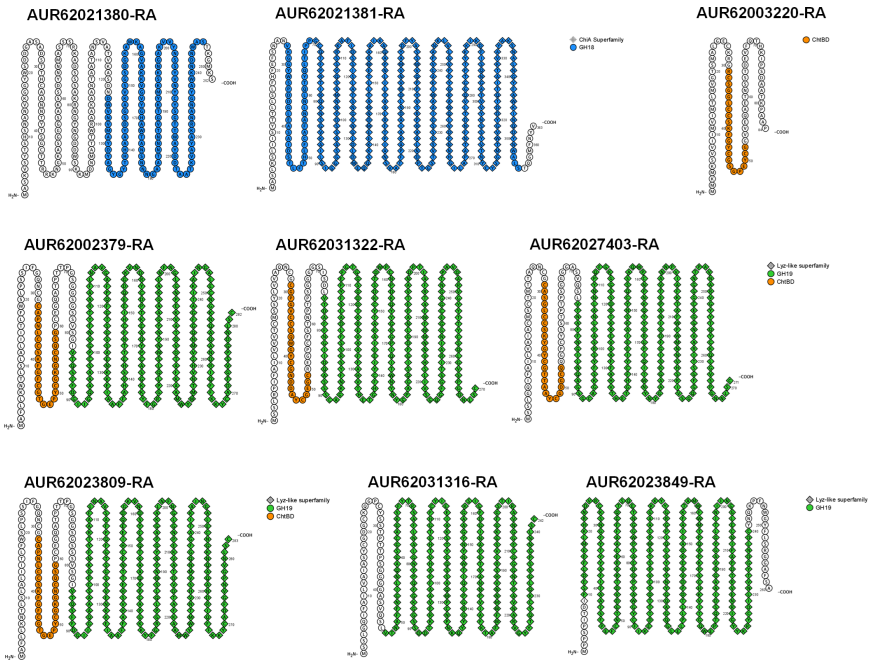

Figure 5. Conserved domain (CD) prediction for the 9 chitinase-related proteins identified in quinoa seeds under rainfed conditions. Representation of predicted conserved domains for the 9 chitinase-related proteins identified in quinoa seeds harvested from rainfed conditions, using Protter (http://wlab.ethz.ch/protter/start/).

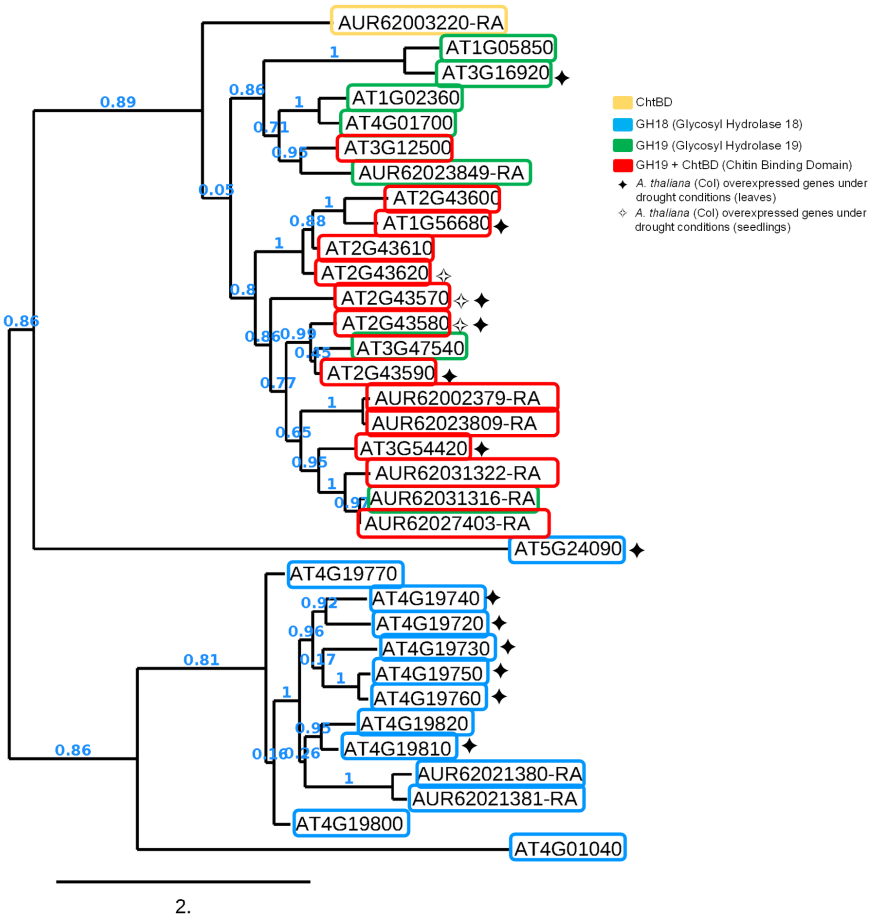

Figure 6. Phylogenetic tree of *A. thaliana* chitinases and quinoa chitinase-related proteins found in seeds under rainfed condition. 25 chitinase are identified in *A. thaliana* which are mainly divided into three groups regarding their catalitic and binding domains. The 9 chitinase-like proteins overrepresented under rainfed conditions in quinoa seeds showed sequence similarities to protein domains of *A. thaliana* ones. In addition, quinoa chitinase-like proteins were closer to ones that were highly expressed in microarray data from leaves and seedlings of *A. thaliana* grown under drought conditions from two independent experiments summarized in Grover, 2012.

Posted on 3 Dec 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.167005135.55252518/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.