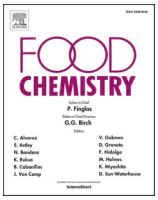
Food Chemistry 398 (2023) 133895

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem



Protein profiling and classification of commercial quinoa grains by MALDI-TOF-MS and chemometrics

Rocío Galindo-Luján [a], Laura Pont [a,b,*], Victoria Sanz-Nebot [a], Fernando Benavente [a]

[a] Department of Chemical Engineering and Analytical Chemistry, Institute for Research on Nutrition and Food Safety (INSA·UB), University of Barcelona, 08028 Barcelona, Spain
[b] Serra Húnter Programe, Generalitat de Catalunya, 08007 Barcelona, Spain

ARTICLE INFO

Keywords:
MALDIquant
MALDI-TOF-MS
Multivariate analysis
Profiling
Proteins
Quinoa

ABSTRACT

Quinoa is an Andean grain that is attracting attention worldwide as a high-quality protein-rich food. Nowadays, quinoa foodstuffs are susceptible to adulteration with cheaper cereals. Therefore, there is a need to develop novel methodologies for protein characterization of quinoa. Here, we first developed a matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) method to obtain characteristic mass spectra of protein extracts from 4 different commercial quinoa grains, which group different varieties marketed as black, red, white (from Peru) and royal (white from Bolivia). Then, data preprocessing and peak detection with MALDIquant allowed detecting 47 proteins (being 30 tentatively identified), the intensities of which were considered as fingerprints for multivariate data analysis. Finally, classification by partial least squares-discriminant analysis (PLS-DA) was excellent, and 34 out of the 47 proteins were critical for differentiation, confirming the potential of the methodology to obtain a reliable classification of quinoa grains based on protein fingerprinting.

1. Introduction

Quinoa (*Chenopodium quinoa* Willd.) is an herbaceous flowering plant, the seeds of which offer remarkable nutritional and functional properties (Aloisi et al., 2016). Quinoa consumption is rapidly increasing globally (Angeli et al., 2020), as this protein-rich food provides several benefits for human health. Among them, quinoa offers a gluten-free alternative for individuals with celiac disease (Niro et al., 2019). Additionally, quinoa flour is being studied as a substitute for wheat flour in bread making processes due to its immuno-nutritional properties (Laparra & Haros, 2016, 2018). This increased interest in quinoa has propelled demand and cost, making quinoa foodstuffs susceptible to adulteration with cheaper cereals (Rodríguez et al., 2019; Shotts et al., 2018).

As food adulteration may result in serious health problems for consumers (Bansal et al., 2017), there is a need for developing novel and robust methodologies for the characterization of foodstuff, as part of quality control and authentication programs. In food analysis, targeted and non-targeted analytical approaches are typically used (Cavanna et al., 2018). Targeted analysis requires a priori knowledge of the contaminants and it is focused on the detection of one or more classes of fraud markers. In contrast, non-targeted analysis, also known as fingerprinting, is based on obtaining a global profile of certain components by analytical techniques, including spectroscopic, spectrometric, chromatographic or electromigration techniques (Álvarez et al., 2018; Hong et al., 2017).

Due to its excellent performance, liquid chromatography with ultraviolet absorption detection (LC-UV) stills being one of the most

Abbreviations: ACN, acetonitrile; CE-UV-DAD, capillary electrophoresis with ultraviolet absorption-diode array detection; LC-MS, liquid chromatography-mass spectrometry; LC-MS/MS, liquid chromatography-tandem mass spectrometry; LC-UV, liquid chromatography with ultraviolet absorption detection; LOWESS, locally weighted scatterplot smoothing; LV, latent variable; MAD, median absolute deviation; MALDI-TOF-MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; PC, principal component; PCA, principal component analysis; PLS-DA, partial least squares-discriminant analysis; SA, sinapinic acid; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SNIP, statistics-sensitive non-linear iterative peak-clipping; SNR, signal-to-noise ratio; TFA, trifluoroacetic acid; TIC, total ion current; VIP, variable importance in the projection.

* Corresponding author at: Department of Chemical Engineering and Analytical Chemistry, Institute for Research on Nutrition and Food Safety (INSA·UB), University of Barcelona, 08028 Barcelona, Spain.
E-mail address: laura.pont@ub.edu (L. Pont).

https://doi.org/10.1016/j.foodchem.2022.133895
Received 2 April 2022; Received in revised form 28 July 2022; Accepted 6 August 2022
Available online 10 August 2022

applied fingerprinting techniques to analyze food products (Gan et al., 2019; Jablonski et al., 2014). In a recent publication, we demonstrated the usefulness of capillary electrophoresis with ultraviolet absorption-diode array detection (CE-UV-DAD) and advanced chemometrics for protein profiling and classification of commercial quinoa grains (Galindo-Luján et al., 2021). However, these UV-based methods are only appropriate for detection of highly absorbing compounds and they lack molecular mass confirmation, which may lead to ambiguous results. To overcome these major drawbacks, different mass spectrometry (MS) approaches have been proposed, including matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS), which represents an ideal option for the reliable detection of food adulteration due to the ease of use, speed of analysis, appropriate sensitivity and wide applicability (Zambonin, 2021). The potential of MALDI-TOF-MS in food fraud detection has been demonstrated analyzing a wide range of compounds, from small molecules (e.g. lipids, carbohydrates, sugars, phenols, etc.) to large biomolecules such as proteins, in milk and dairy products, meat, fish, seafood, oils, vegetables and fruit (Kiran et al., 2016; Kuo et al., 2019; Sassi et al., 2015; Stahl & Schröder, 2017).

Despite the usefulness of MALDI-TOF-MS for fingerprinting approaches, mass spectra processing for detection of peaks with appropriate signal-to-noise ratio (SNR), so they can be discriminated from the background noise (i.e. peak detection), still remains a challenging task. This is especially true for proteomics data, as mass spectra usually contain altering baseline and noise, which can negatively affect the interpretation of the results (Yang et al., 2009). In the past years, several efforts have been made to provide open-source data analysis software for proteomics studies, such as PROcess (Li, 2005), OpenMS (Kohlbacher et al., 2007) or MALDIquant (Gibb & Strimmer, 2017). These software solutions offer a collection of procedures for MALDI-TOF mass spectra, from data preprocessing to peak detection, in order to obtain accurate lists of peaks, with characteristic mass-to-charge ratio (m/z) and intensity values. Specifically, MALDIquant has proven to be useful for a wide range of proteomics applications, including protein profiling of serum from individuals with pancreatic cancer (Fiedler et al., 2009), peptide profiling of virus-causing insects (Uhlmann et al., 2014) or protein profiling of nasal swabs from individuals with SARS-CoV-2 (Nachtigall et al., 2020).

The novelty of the present study relies on the development, for the first time to the best of our knowledge, of a comprehensive workflow for protein profiling, peak detection and classification of commercial quinoa grains based on the combination of MALDI-TOF-MS analysis of protein extracts, MALDIquant and chemometrics. The proposed methodology enhances our previous method using CE-UV-DAD (Galindo-Luján et al., 2021). It allows, not only to efficiently classify and differentiate the most widely commercially available quinoa grains, which group different varieties marketed as black, red, white (from Peru) and royal (white from Bolivia), but also to tentatively identify the most important proteins for discrimination.

2. Material and methods

2.1. Chemicals and samples

All the chemicals used were of analytical reagent grade or better. Sodium hydroxide (≥99.0 %, pellets), hydrochloric acid (37 % (v/v)), boric acid (≥99.5 %), acetonitrile (ACN, LC-MS grade), water (LC-MS grade), trifluoroacetic acid (TFA, 99.0 %), acetone (99.8 %) and sinapinic acid (SA, ≥99.0 %) were supplied by Merck (Darmstadt, Germany). Black (B, 7 samples), red (R, 5 samples) and white (W, 7 samples) quinoa grains from Peru, as well as royal white (RO, 5 samples) quinoa grains from Bolivia were acquired in local supermarkets from Barcelona.

2.2. Sample preparation

Quinoa grains were dried in an air-current oven at 40 °C for 24 h, ground in a coffee grinder and stored at room temperature in a desiccator. Quinoa proteins were extracted as described in our previous studies (Galindo-Luján et al., 2021; Galindo-Luján et al., 2021). Quinoa protein extracts were prepared in triplicate for the different quinoa grains (72 quinoa protein extracts). Before MALDI-TOF-MS analyses, quinoa protein extracts were desalted using MF-Millipore® membrane filters (Merck). Briefly, 10 μL of sample solution were deposited onto the membrane filter. After dialyzing with water for 45 min at room temperature, aliquots of the samples were collected and stored at −20 °C.

2.3. MALDI-TOF-MS

MALDI-TOF mass spectra were obtained using a 4800 MALDI TOF/TOF mass spectrometer (Applied Biosystems, Waltham, MA, USA). Mass spectra were acquired over a range of 5,500–25,000 m/z using the mid mass positive mode, which is recommended by the instrument manufacturer for the analysis of proteins with relative molecular masses (M_r) in the range of 5,000–25,000. Data acquisition and data processing were performed using the 4000 Series Explorer™ and Data Explorer® software (Applied Biosystems), respectively. Sample-MALDI matrix mixtures were freshly prepared as described in a previous work with some modifications (Pont et al., 2020). Briefly, the preparation consisted on depositing onto a stainless steel MALDI plate the following layers: 1 μL of SA in 99:1 (v/v) acetone:water (final SA concentration 27 mg·mL^{-1}), 1 μL of sample solution (desalted quinoa protein extract), again 1 μL of sample solution (to increase sample homogeneity) and, finally, 1 μL of SA acid in 50:50 (v/v) ACN:water with 0.1 % (v/v) of TFA (final SA concentration 10 mg·mL^{-1}). Spots were allowed to dry at room temperature between each layer addition in order to ensure maximum homogeneity and, therefore, reproducibility in the MALDI-TOF-MS analyses. All the samples were analyzed in triplicate (72 quinoa protein extracts × 3 spots).

2.4. Data analysis

Experimental data were analyzed combining MALDIquant and chemometrics (i.e. principal component analysis, PCA, followed by partial least squares-discriminant analysis, PLS-DA). First, MALDIquant was used to detect protein peaks, with characteristic m/z and intensity values in the mass spectra. Then, PCA followed by PLS-DA were applied to perform multivariate analysis and classify the different commercial quinoa grain samples according to their protein composition. Data processing and graphical representation were performed under R platform (version 4.0.4, https://www.R-project.org/) (R Development Core Team, 2020). MALDIquant was run under MALDIquant R package (version 1.19.3) (Gibb & Strimmer, 2017) and mdatools R package (version 0.12.0) was used for PCA and PLS-DA (Kucheryavskiy, 2020).

2.4.1. MALDIquant

First, raw mass spectra were converted to text format (.txt) using a macro available with the Data Explorer® software and, then, imported into the R environment using MALDIquantForeign R package (version 0.12) (Gibb, 2014). Subsequently, imported data were transformed for variance stabilization and then smoothed. Next, a baseline correction was applied to remove background noise. Denoised data was normalized for a proper comparison of intensity values across different mass spectra and, then, aligned. After that, the 3 preprocessed mass spectra obtained for the different protein extracts (3 spots for each protein extract were analyzed by MALDI-TOF-MS) were averaged to obtain a mean mass spectrum for each protein extract. Next, a peak detection algorithm was applied to accurately identify potential protein features. Finally, a peak binning procedure allowed correcting m/z shifts across mass spectra.

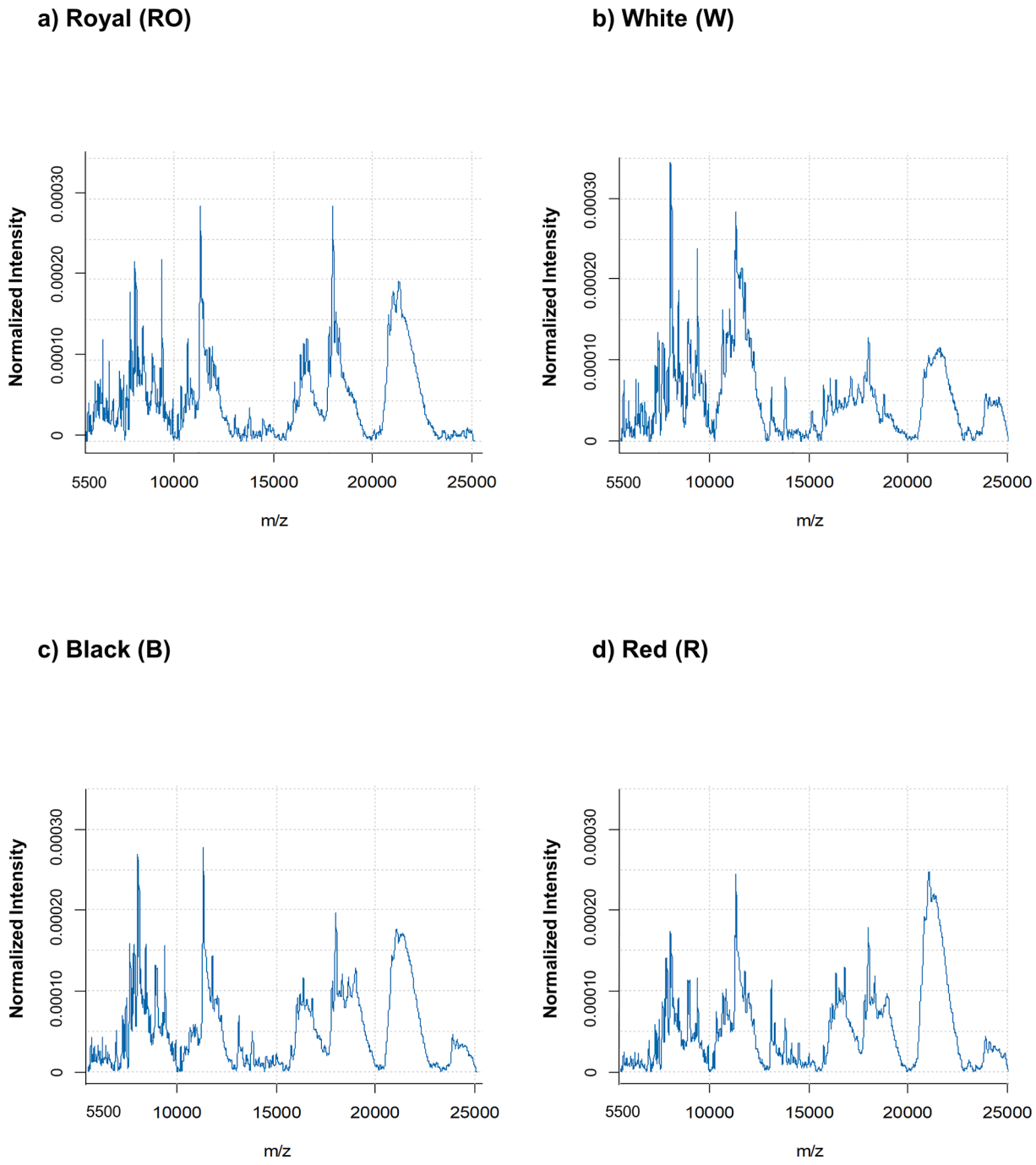


Fig. 1. MALDI-TOF mass spectra obtained after MALDIquant processing for (a) RO, (b) W, (c) B and (d) R quinoa protein extracts.

2.4.2. Multivariate data analysis

The intensities of the detected protein peaks by MALDIquant were considered for PCA and PLS-DA analysis. PCA was first used for the unsupervised identification of trends and clustering of the data, as well as for the detection of outliers (Joliffe & Morgan, 1992). PLS-DA was then applied to build a classification model with improved class separation (Barker & Rayens, 2003) and to reveal the importance of the different protein peaks for discrimination taking into account their variable importance in the projection (VIP) scores (Wold et al., 2001). A leave-one-out cross validation of the PLS-DA model was performed during model optimization (Wold et al., 2001).

3. Results and discussion

3.1. MALDI-TOF-MS analysis

Some preliminary MALDI-TOF-MS experiments were performed using a RO quinoa sample to select the most appropriate conditions to obtain the characteristic mass spectra profiles of protein extracts from quinoa. Special attention was given to the preparation of the sample-MALDI matrix mixtures and spot deposition. In a previous work (Pont et al., 2020), we demonstrated the usefulness of a sandwich method to prepare reproducible spots for the analysis by MALDI-TOF-MS of protein extracts from barley and malt. In this sandwich method, a sample droplet was applied on top of a fast-evaporated matrix (e.g. SA in 99:1 (v/v) acetone:water), followed by the deposition of a second layer of

Table 1
Proteins detected by MALDI-TOF-MS used as PLS-DA variables with their corresponding experimental M_r, VIP score values (for discrimination of RO, W, B and R quinoa samples) and tentative identifications (theoretical M_r, accession number (ID) and protein name based on an experimental quinoa seed proteome map obtained by shotgun LC-MS/MS proteomics (Galindo-Luján et al., 2021)).

PLS-DA protein variables[a]						Tentative identifications[b]	
Protein	Experimental M_r[c]	VIP Scores[d]				Theoretical M_r	Accession number (ID)[e] and protein name
		RO	W	B	R		
1	6,279	0.61	**1.05**	0.86	0.73	–	–
2	6,392	0.55	**1.09**	0.84	0.72	6,413	XP_021764391.1 40S ribosomal protein S29
							XP_021764390.1 40S ribosomal protein S29
							XP_021762716.1 40S ribosomal protein S29
							XP_021762714.1 40S ribosomal protein S29
3	6,719	**1.21**	**1.26**	**1.11**	**1.27**	–	–
4	7,217	0.85	0.72	0.82	0.80	–	–
5	7,432	0.72	0.39	0.72	0.60	–	–
6	7,623	0.80	**1.11**	0.85	0.92	–	–
7	7,722	0.60	0.35	0.59	0.51	7,687	XP_021750374.1 LOW QUALITY PROTEIN: protein transport protein Sec61 subunit gamma-1-like
							XP_021727999.1 protein transport protein Sec61 subunit gamma-1-like
							XP_021713805.1 protein transport protein Sec61 subunit gamma-1-like
8	7,771	0.97	0.59	0.93	0.85	–	–
9	7,855	**1.70**	**1.07**	**1.57**	**1.52**	–	–
10	8,019	0.62	0.85	0.73	0.68	7,986	XP_021768098.1 ATP synthase subunit epsilon, mitochondrial
						8,039	XP_021772297.1 60S ribosomal protein L38
							XP_021734344.1 60S ribosomal protein L38
11	8,072	**1.20**	**1.03**	0.80	**1.26**	8,039	XP_021772297.1 60S ribosomal protein L38
							XP_021734344.1 60S ribosomal protein L38
12	8,173	0.64	0.51	0.57	0.62	8,163	XP_021745172.1 cytochrome *b*-c1 complex subunit 6-like
							XP_021717737.1 cytochrome *b*-c1 complex subunit 6-like
13	8,414	0.71	0.98	0.75	0.82	–	–
14	8,501	0.45	0.71	0.59	0.52	–	–
15	8,640	0.77	**1.34**	**1.03**	0.96	–	–
16	8,902	**1.11**	0.84	0.64	**1.16**	8,943	XP_021769286.1 40S ribosomal protein S21-like
17	8,937	0.75	**1.37**	**1.09**	0.93	8,943	XP_021769286.1 40S ribosomal protein S21-like
						8,951	XP_021765473.1 em-like protein GEA6
						8,977	XP_021733643.1 protein deletion of SUV3 suppressor 1-like
18	8,966	**1.22**	0.53	**1.20**	**1.00**	8,977	XP_021733643.1 protein deletion of SUV3 suppressor 1-like
19	9,049	0.73	**1.39**	**1.07**	0.94	9,076	XP_021768105.1 late seed maturation protein P8B6-like
20	9,266	0.71	**1.37**	**1.06**	0.91	–	–
21	9,293	0.82	**1.38**	**1.12**	0.98	–	–
22	9,383	**1.21**	**1.39**	**1.28**	**1.26**	–	–
23	9,773	0.92	**1.09**	**1.02**	0.96	–	–
24	10,656	**1.71**	0.96	**1.45**	**1.55**	–	–
25	10,808	**1.09**	0.99	0.79	**1.15**	10,858	XP_021759005.1 small ubiquitin-related modifier 1
							XP_021755691.1 small ubiquitin-related modifier 1
26	11,322	0.31	0.42	0.38	0.34	11,270	XP_021771595.1 sm-like protein LSM3A
							XP_021761024.1 sm-like protein LSM3A
						11,285	XP_021725656.1 lectin-C-like
						11,308	XP_021716413.1 60S acidic ribosomal protein *P*2-4-like
						11,339	XP_021723219.1 non-specific lipid-transfer protein-like
						11,343	XP_021748130.1 60S acidic ribosomal protein *P*1-like
						11,345	XP_021748130.1 60S acidic ribosomal protein *P*1-like
						11,348	XP_021764768.1 60S acidic ribosomal protein P2A-like
							XP_021728700.1 60S acidic ribosomal protein P2A-like
						11,366	XP_021722096.1 60S acidic ribosomal protein *P*2-2-like
27	11,540	**1.05**	**1.12**	0.96	**1.11**	–	–
28	11,617	**1.45**	**1.17**	**1.52**	**1.31**	–	–
29	11,775	0.96	**1.01**	**1.14**	0.90	11,797	XP_021763483.1 small ubiquitin-related modifier 1-like
							XP_021720059.1 small ubiquitin-related modifier 1-like
30	11,892	0.98	**1.02**	0.79	**1.06**	11,902	XP_021771529.1 60S acidic ribosomal protein *P*3-like
31	12,046	0.74	**1.12**	0.95	0.83	11,992	YP_009380236.1 ribosomal protein S18
							YP_009380152.1 ribosomal protein S18
						12,050	XP_021776279.1 peptidyl-prolyl cis–trans isomerase FKBP12-like
32	12,216	0.62	**1.05**	0.84	0.75	12,163	XP_021774134.1 peptidyl-prolyl cis–trans isomerase FKBP12-like
33	13,115	**1.23**	0.68	**1.11**	**1.09**	13,082	XP_021765265.1 peptidyl-prolyl cis–trans isomerase Pin1-like
							XP_021722042.1 peptidyl-prolyl cis–trans isomerase Pin1-like
34	16,060	0.63	0.54	0.71	0.55	15,995	XP_021730553.1 40S ribosomal protein S17-like
							XP_021773485.1 40S ribosomal protein S17-like
							XP_021765576.1 40S ribosomal protein S17-like
							XP_021728631.1 40S ribosomal protein S17-like
						15,997	XP_021732055.1 actin-depolymerizing factor 4-like
						16,093	XP_021752156.1 oleosin 1-like
						16,096	XP_021731590.1 glycine-rich RNA-binding, abscisic acid-inducible protein-like
						16,101	XP_021726936.1 probable prefoldin subunit 2
							XP_021773911.1 probable prefoldin subunit 2
						16,125	XP_021738617.1 glycine-rich RNA-binding protein-like

(*continued on next page*)

Table 1 (*continued*)

PLS-DA protein variables[a]						Tentative identifications[b]	
Protein	Experimental M_r[c]	VIP Scores[d]				Theoretical M_r	Accession number (ID)[e] and protein name
		RO	W	B	R		
						16,134	XP_021716351.1 ferredoxin, root R-B2-like
35	16,188	**1.41**	**1.21**	0.92	**1.49**	16,125	XP_021738617.1 glycine-rich RNA-binding protein-like
						16,134	XP_021716351.1 ferredoxin, root R-B2-like
						16,200	XP_021717733.1 high mobility group B protein 3-like
						16,215	XP_021716749.1 ferredoxin, root R-B2-like
						16,216	XP_021754488.1 high mobility group B protein 3-like
						16,239	XP_021762815.1 uncharacterized protein At5g48480-like
							XP_021733518.1 uncharacterized protein At5g48480-like
						16,250	XP_021766528.1 40S ribosomal protein S14-2
							XP_021765192.1 40S ribosomal protein S14-2
36	16,360	**1.38**	**1.37**	**1.17**	**1.45**	16,289	XP_021721762.1 oleosin 1-like
						16,304	XP_021756410.1 nucleoside diphosphate kinase 1
							XP_021756411.1 nucleoside diphosphate kinase 1
						16,318	XP_021755225.1 nucleoside diphosphate kinase 1-like
						16,431	XP_021747488.1 uncharacterized protein LOC110713339
							XP_021716984.1 uncharacterized protein LOC11068485
37	16,516	**1.36**	0.98	**1.19**	**1.28**	16,469	XP_021746531.1 60S ribosomal protein L27a-3-like
							XP_021743225.1 60S ribosomal protein L27a-3-like
						16,474	XP_021769235.1 glycine cleavage system H protein 2, mitochondrial-like
							XP_021732532.1 glycine cleavage system H protein 2, mitochondrial-like
38	16,679	**1.32**	0.85	**1.18**	**1.21**	16,616	XP_021755504.1 2S albumin-like
						16,624	XP_021751394.1 60S ribosomal protein L26-1
							XP_021714459.1 60S ribosomal protein L26-1
							XP_021769238.1 60S ribosomal protein L26-2-like
							XP_021732535.1 60S ribosomal protein L26-2-like
						16,625	XP_021730224.1 probable calcium-binding protein CML13
							XP_021725345.1 probable calcium-binding protein CML13
						16,651	XP_021731588.1 glycine-rich RNA-binding, abscisic acid-inducible protein-like
						16,676	XP_021732772.1 60S ribosomal protein L28-1-like
						16,685	XP_021735190.1 ubiquitin-conjugating enzyme E2 variant 1D-like
							XP_021724506.1 ubiquitin-conjugating enzyme E2 variant 1D-like isoform X2
							XP_021724505.1 ubiquitin-conjugating enzyme E2 variant 1C-like isoform X1
						16,693	XP_021774210.1 60S ribosomal protein L28-1-like
						16,702	XP_021717270.1 blue copper protein-like isoform X2
							XP_021717265.1 mavicyanin-like isoform X1
							XP_021761125.1 mavicyanin-like
						16,742	XP_021720407.1 17.4 kDa class III heat shock protein-like
							XP_021720406.1 17.4 kDa class III heat shock protein-like
39	16,807	0.66	**1.04**	0.88	0.76	16,742	XP_021720407.1 17.4 kDa class III heat shock protein-like
							XP_021720406.1 17.4 kDa class III heat shock protein-like
						16,775	XP_021749320.1 uncharacterized protein LOC110715055
						16,812	XP_021758596.1 2S albumin-like
						16,833	XP_021733717.1 40S ribosomal protein S16-like
							XP_021719033.1 40S ribosomal protein S16-like
							XP_021741014.1 40S ribosomal protein S16-like
							XP_021724150.1 40S ribosomal protein S16
						16,834	XP_021776507.1 calmodulin-7-like
							XP_021746004.1 calmodulin-7-like
						16,860	XP_021754554.1 calmodulin
							XP_021749259.1 calmodulin
							XP_021745669.1 esterase CG5412-like
							XP_021754747.1 calmodulin-2/4
						16,877	XP_021749775.1 peptidyl-prolyl cis–trans isomerase FKBP15-1-like
							XP_021720138.1 peptidyl-prolyl cis–trans isomerase FKBP15-1-like
40	17,811	0.71	0.37	0.72	0.59	17,762	XP_021756280.1 40S ribosomal protein S18
							XP_021748256.1 40S ribosomal protein S18-like
						17,803	XP_021769395.1 40S ribosomal protein S11-like
							XP_021739192.1 40S ribosomal protein S11-like
						17,855	XP_021773311.1 60S ribosomal protein L12-1
							XP_021746006.1 60S ribosomal protein L12-1
							XP_021735009.1 60S ribosomal protein L12-1
							XP_021730745.1 60S ribosomal protein L12-1
						17,901	XP_021732395.1 17.4 kDa class I heat shock protein-like
							XP_021718604.1 17.8 kDa class I heat shock protein-like
41	17,996	0.87	0.34	0.89	0.68	17,939	XP_021749487.1 MLP-like protein 43
							XP_021745405.1 MLP-like protein 43
							XP_021745404.1 MLP-like protein 43
						17,975	XP_021767152.1 desiccation protectant protein Lea14 homolog
						18,084	XP_021769109.1 17.8 kDa class I heat shock protein-like
42	18,173	**1.27**	0.41	**1.32**	0.98	18,084	XP_021769109.1 17.8 kDa class I heat shock protein-like
						18,127	XP_021767915.1 uncharacterized protein LOC110732302
							XP_021757753.1 uncharacterized protein LOC110722767

(*continued on next page*)

Table 1 (*continued*)

PLS-DA protein variables[a]						Tentative identifications[b]	
Protein	Experimental M_r[c]	VIP Scores[d]				Theoretical M_r	Accession number (ID)[e] and protein name
		RO	W	B	R		
						18,168	XP_021749113.1 14 kDa zinc-binding protein-like
							XP_021745801.1 14 kDa zinc-binding protein-like
						18,195	XP_021733566.1 oleosin 16.4 kDa-like
						18,214	XP_021746976.1 mitochondrial fission 1 protein A-like
							XP_021744566.1 mitochondrial fission 1 protein A-like
						18,221	XP_021738830.1 oleosin 16 kDa
						18,227	XP_021732378.1 18.3 kDa class I heat shock protein
							XP_021769094.1 18.3 kDa class I heat shock protein-like
						18,238	XP_021765145.1 60S ribosomal protein L24-like
							XP_021754139.1 60S ribosomal protein L24
							XP_021730109.1 60S ribosomal protein L24
							XP_021735902.1 60S ribosomal protein L24-like
						18,240	XP_021753128.1 peptidyl-prolyl cis–trans isomerase 1-like
						18,254	XP_021775867.1 peptidyl-prolyl cis–trans isomerase 1
43	18,317	**1.35**	0.86	**1.38**	**1.16**	18,238	XP_021765145.1 60S ribosomal protein L24-like
							XP_021754139.1 60S ribosomal protein L24
							XP_021730109.1 60S ribosomal protein L24
							XP_021735902.1 60S ribosomal protein L24-like
						18,240	XP_021753128.1 peptidyl-prolyl cis–trans isomerase 1-like
						18,254	XP_021775867.1 peptidyl-prolyl cis–trans isomerase 1
						18,271	XP_021730326.1 universal stress protein PHOS32
						18,276	XP_021744114.1 17.3 kDa class II heat shock protein-like
						18,348	XP_021738936.1 17.3 kDa class II heat shock protein-like
44	18,640	0.79	**1.16**	0.92	0.92	18,577	XP_021775345.1 SKP1-like protein 1A
							XP_021743046.1 SKP1-like protein 1A
							XP_021728238.1 SKP1-like protein 1A
							XP_021714061.1 SKP1-like protein 1A
						18,607	XP_021770354.1 60S ribosomal protein L21-1
							XP_021721611.1 60S ribosomal protein L21-1
						18,641	XP_021743659.1 universal stress protein PHOS34-like
45	18,994	**1.34**	**1.45**	**1.24**	**1.42**	18,937	XP_021742166.1 peptidyl-prolyl cis–trans isomerase CYP19-3-like
							XP_021742165.1 peptidyl-prolyl cis–trans isomerase CYP19-3-like
							XP_021742164.1 peptidyl-prolyl cis–trans isomerase CYP19-3-like
						18,997	XP_021756574.1 peptidyl-prolyl cis–trans isomerase E-like
							XP_021718372.1 peptidyl-prolyl cis–trans isomerase E-like
						19,030	XP_021769148.1 oleosin 18.2 kDa-like
						19,037	XP_021762572.1 uncharacterized protein LOC110727319
							XP_021725318.1 uncharacterized protein LOC110692594
						19,083	XP_021732650.1 oleosin 18.2 kDa-like
46	20,816	0.62	0.59	0.71	0.56	20,781	XP_021775343.1 translocator protein homolog
							XP_021775344.1 translocator protein homolog
							XP_021763024.1 translocator protein homolog
						20,799	XP_021753718.1 60S ribosomal protein L11-1
							XP_021732924.1 60S ribosomal protein L11-1
						20,844	XP_021763208.1 60S ribosomal protein L18-3-like
							XP_021732173.1 60S ribosomal protein L18-3-like
							XP_021722341.1 60S ribosomal protein L18-3-like
							XP_021772839.1 60S ribosomal protein L18-2
						20,852	XP_021763370.1 monothiol glutaredoxin-S10-like
							XP_021722128.1 glutaredoxin-C5, chloroplastic-like
						20,903	XP_021771897.1 uncharacterized GPI-anchored protein At3g06035-like
							XP_021742916.1 uncharacterized GPI-anchored protein At3g06035-like
						20,909	XP_021739907.1 nucleolin-like
							XP_021730379.1 nucleolin-like
47	21,028	0.67	**1.30**	**1.00**	0.87	20,942	YP_009380200.1 AtpF (chloroplast)
							YP_009380116.1 AtpF (chloroplast)
						21,121	XP_021733985.1 glycine-rich RNA-binding protein 3, mitochondrial-like

[a] PLS-DA variables correspond to the detected protein peaks by MALDIquant.

[b] The quinoa seed proteome map obtained by shotgun LC-MS/MS proteomics in our previous work (Galindo-Luján et al., 2021) was used as a reference for the tentative identification. Only a mass error ± 0.5 % between the theoretical and experimental M_r was considered acceptable for proposing an identity. This threshold value was established considering the mass error observed for the analysis of a ribonuclease A standard (from bovine pancreas) under the same instrumental conditions, $M_r = 13{,}690$.

[c] Experimental M_r were calculated from the *m/z* values considering the formation of single-charged molecular ions by MALDI-TOF-MS.

[d] VIP scores ≥ 1 were considered important for discrimination and are marked in bold red.

[e] Acession numbers (IDs) of the identified proteins correspond to the IDs of our previous work (Galindo-Luján et al., 2021).

matrix prepared in a hydroorganic acidic solution of lower volatility (e.g. SA in 50:50 (v/v) ACN:water with 0.1 % (v/v) of TFA). In a typical variant of this method, the sample is mixed 1:1 (v/v) with the hydroorganic acidic solution, and only the mixture is deposited after the fast-evaporated layer. With quinoa protein extracts, both method variants led to non-homogeneous crystal size distribution, bad quality mass spectra and unreproducible results. In order to improve spot homogeneity, it was necessary to add a second sample layer between the first sample layer and the last matrix layer (as described in section 2.3). Under these conditions, different laser intensities ranging between

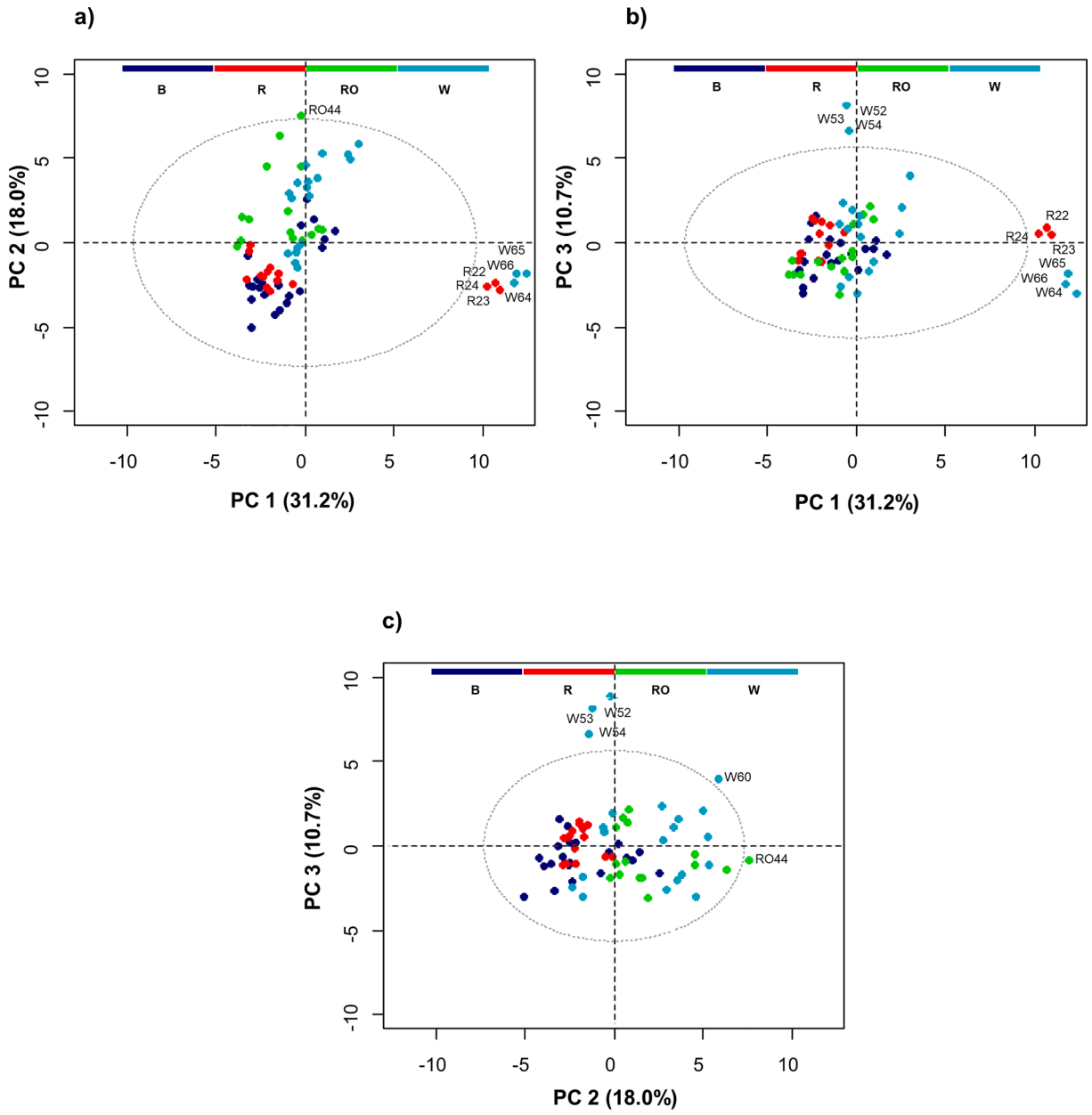


Fig. 2. Scores plots of the PCA model applied to the 72 quinoa protein extracts using the intensities of the 47 protein peaks detected by MALDIquant. Labelled samples were considered outliers (11), and were discarded before applying PLS-DA. (a) PC2 vs. PC1, (b) PC3 vs. PC1 and (c) PC3 vs. PC2.

6,000 and 7,000 were investigated to obtain an appropriate sensitivity. A value of 6,300 (80 % of maximum intensity) was selected to avoid excessive sample destruction, as no significant differences in sensitivity were observed with higher laser intensities. Fig. 1-a shows the MALDI-TOF mass spectrum acquired for a RO quinoa protein extract over a range of 5,500–25,000 *m/z* using the mid mass positive mode under optimized conditions. As can be observed, the mass spectrum presented single-charged molecular ions corresponding to different protein components with M_r comprising the full measured range. These results agree with our previous studies based on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Galindo-Luján et al., 2021), where we tentatively associated the most abundant protein bands detected between 5,000–15,000 and 15,000–25,000 M_r with 2S albumins and 11S globulins, respectively. Then, the rest of protein extracts from the different quinoa grains were analyzed under the selected conditions. As an example, Fig. 1 b-d show the MALDI-TOF mass spectra for the protein extracts of a W, a B, and a R quinoa sample. As can be observed, similarities and differences in the mass spectra profiles for the four quinoa grain samples were hard to distinguish at naked eye. In addition, direct peak detection for protein fingerprinting from this complex mass spectra profiles was extremely difficult because most of the peaks overlapped. As an alternative, we explored the use of MALDIquant for a reliable and improved peak detection before applying chemometrics (i.e. PCA and PLS-DA) for multivariate analysis and classification of the different quinoa grains.

3.2. *MALDIquant*

Once the raw data from the protein extracts from RO, W, B and R quinoa samples (72 protein extracts × 3 spots) were imported into the R environment, different preprocessing methods were applied to the mass spectra. First, variance stabilization using a square root transformation was performed (Purohit & Rocke, 2003). After that, a smoothing procedure employing the Savitsky-Golay filter was applied to reduce noise

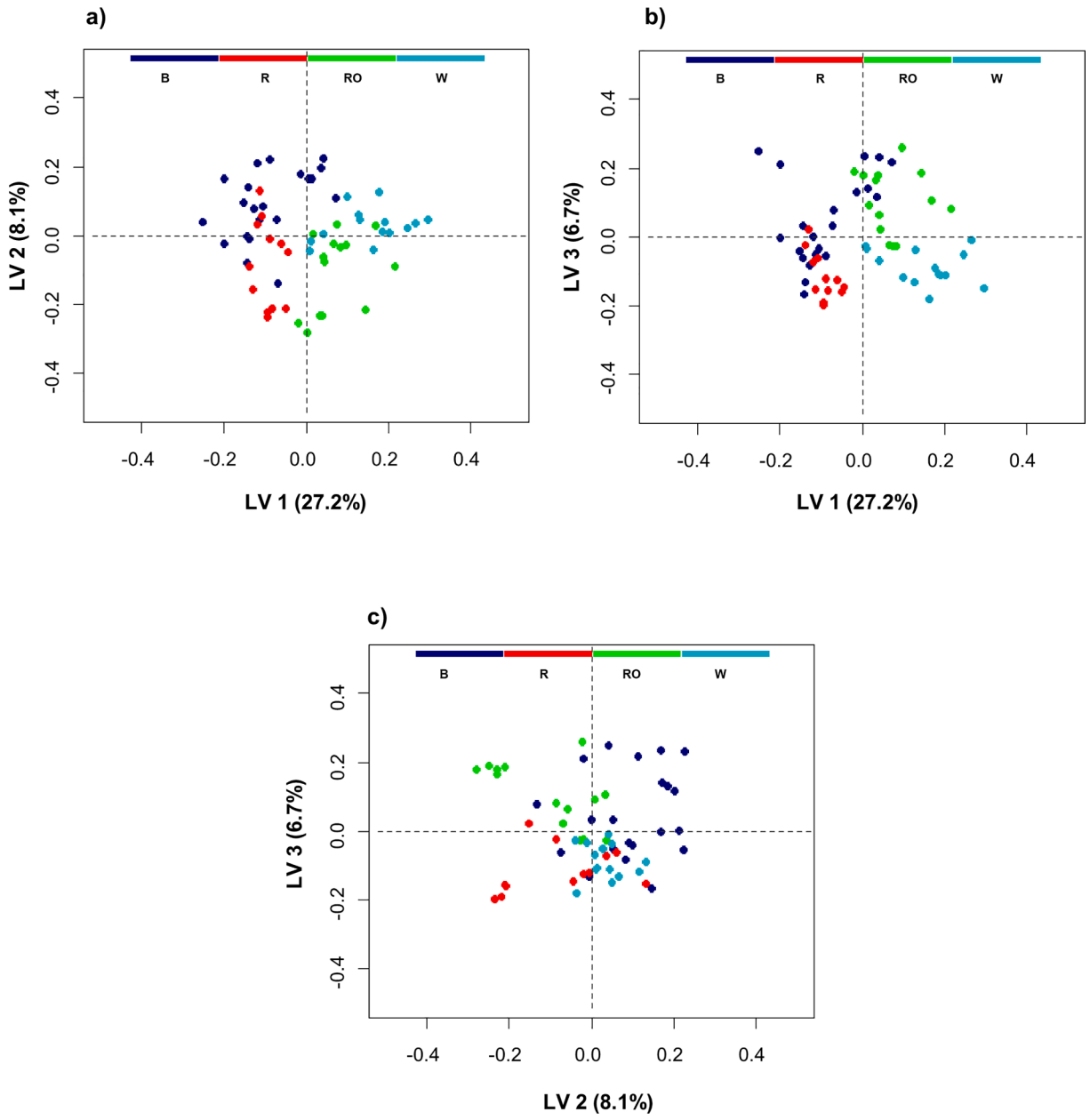


Fig. 3. Scores plots of the PLS-DA model applied to the 61 quinoa protein extracts after removing outliers using the intensities of the 47 protein peaks detected by MALDIquant. (a) LV2 vs. LV1, (b) LV3 vs. LV1 and (c) LV3 vs. LV2.

coming from artefacts and, consequently, to improve SNR (Savitzky & Golay, 1964). A baseline correction was also mandatory in order to remove background effects. With this purpose, we used the statistics-sensitive non-linear iterative peak-clipping (SNIP) algorithm (Ryan et al., 1988). After that, a normalization step using the total ion current (TIC) method, which is the most common normalization procedure for MALDI-TOF-MS data (Borgaonkar et al., 2010), was applied to allow proper comparison of intensity values across different preprocessed mass spectra. Finally, an alignment procedure was performed using the locally weighted scatterplot smoothing (LOWESS) warping algorithm, which performed better extracting local variability from data subsets than other regression methods (linear, polynomial, etc.) (Cleveland, 1979). After the alignment, the preprocessed mass spectra obtained for the 3 spots of the different protein extracts were averaged to obtain a mean mass spectrum for each protein extract. As an example, Supplementary Fig. S1 a-b and Fig. 1 c-d show the MALDI-TOF mass spectra obtained for a B and a R quinoa sample before and after mass spectra preprocessing, respectively.

Protein peaks were detected with the median absolute deviation (MAD) method (Friedman, 1984). In this procedure, a window is moved across the spectra and local maxima are detected. These local maxima are then compared against a noise baseline estimated by the MAD method. If a local maximum is above a given SNR, it is considered a peak, whereas local maxima below the SNR threshold are discarded. At this point, it is worth mentioning that a priori knowledge of mass spectra can help in selecting a suitable window size for data subsets and SNR. In our case, after investigating SNR values from 1 to 3, a value of 2 was selected, as a good compromise to detect a large set of characteristic protein peaks. Finally, in order to correct small m/z shifts across mass spectra from different protein extracts, a peak binning procedure was applied, so that the peak intensities were assigned to a set of common m/z values, based on the similarity of their original values. The mass spectra preprocessing and peak detection by MALDIquant allowed detecting 47 proteins in the mass spectra. The detected proteins were

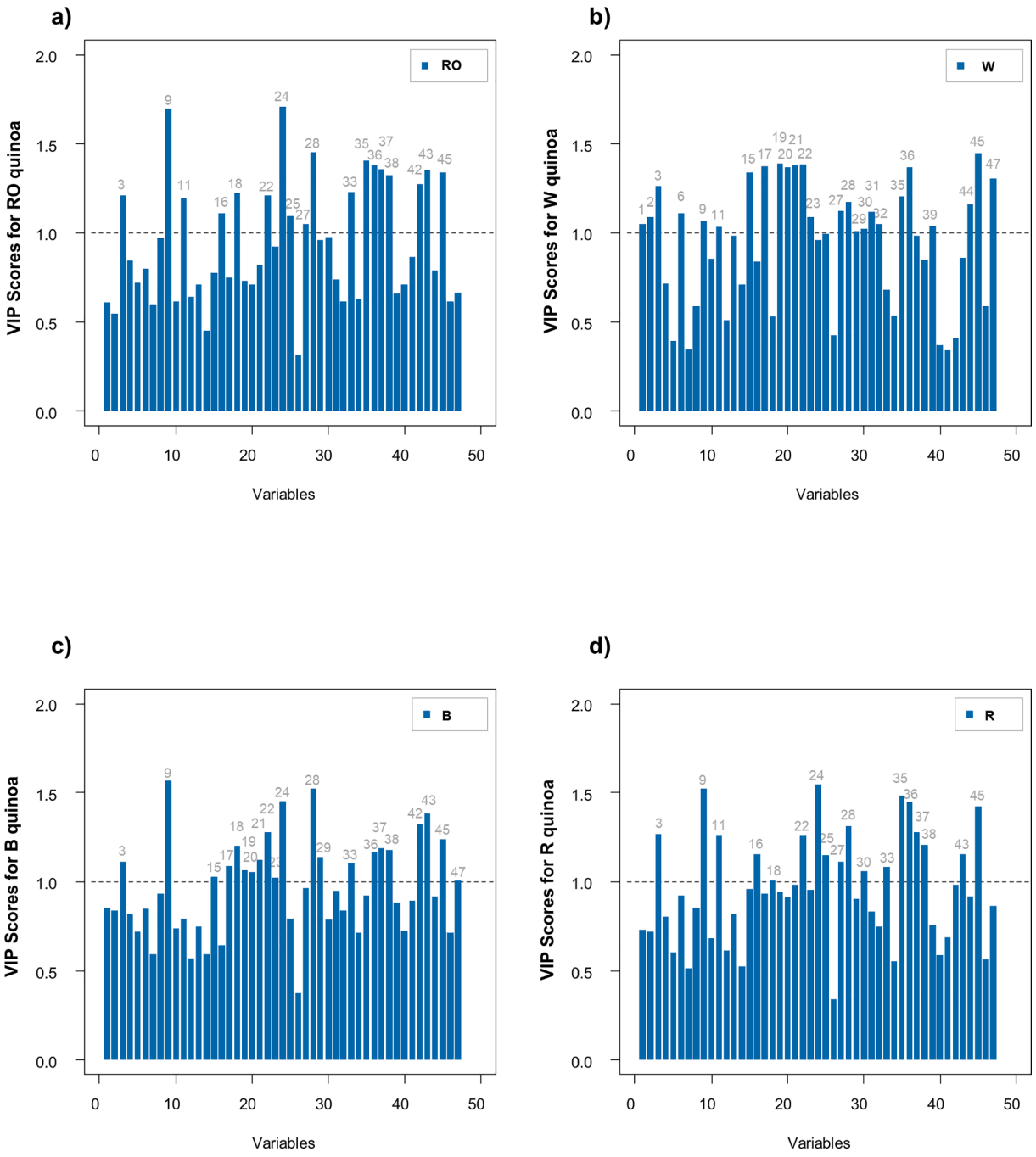


Fig. 4. VIP scores of the different protein variables when considering the separation of (a) RO, (b) W, (c) B and (d) R quinoa grain samples from the rest of classes. Protein variables with a VIP score value higher than 1 are numbered (see also Table 1).

tentatively identified, taking as a reference the quinoa grain proteome map obtained for similar protein extracts by shotgun liquid chromatography-tandem mass spectrometry (LC-MS/MS) proteomics in a previous study (Galindo-Luján et al., 2021). Table 1 shows the experimental M_r calculated for the detected proteins, as well as the theoretical M_r, the accession number (ID) and the name of the tentatively identified proteins. As can be seen in this table, we could propose an identity for 30 out of the 47 detected proteins.

3.3. Multivariate data analysis

Multivariate data analysis was performed considering as variables the intensities of the 47 detected proteins in the 72 quinoa protein extracts. First, we used PCA to explore the classes present in the data and the presence of outliers. Three principal components (PCs) allowed explaining 60 % of the variance (Fig. 2 a-c). As can be observed in the scores plots, PC1 (31 % of the explained variance) allowed a slight separation of W quinoa from the rest of classes, while PC2 (18 % of the explained variance) separated slightly B-R from W-RO quinoa samples. Additionally, the three PCs allowed detecting different outliers corresponding to W, R and RO protein extracts (see labels of the samples appearing alone or clustered outside the 95 % confidence ellipse in Fig. 2 a-c), which were discarded before applying PLS-DA. Once explored the data by PCA, four classes were defined (i.e. RO, W, B and R) to build a PLS-DA model with improved class separation and to reveal the importance of the different protein peaks for discrimination between

quinoa grain classes. As can be observed in the scores plots of Fig. 3 a-c, a PLS-DA model with three latent variables (LVs) (42 % of X-variance and 53 % of Y-variance explained) allowed a proper discrimination between the four quinoa classes, with most of the variance explained by LV1 and LV2 (Fig. 3). As can be seen in these score plots, LV1 (27 % of the explained variance) allowed separating B and R from W and RO quinoa, LV2 (8 % of the explained variance) allowed discriminating B and W from R and RO quinoa, and LV3 (7 % of the explained variance) allowed slightly separating B and RO from R and W quinoa. The contribution of the different variables (proteins) to the LVs can be observed in the loadings plots of Supplementary Fig. S2 a-c. However, the VIP scores are more informative because they facilitated a direct estimation of their influence on separation between the quinoa grain classes. The bar plots of Fig. 4 a-d show the VIP scores of the different proteins considering separation of RO, W, B and R quinoa from the rest of classes, respectively. Only those proteins with a VIP score over a particular threshold value (typically 1) were considered important for discrimination (Wold et al., 2001). As can be observed in this figure, 18, 25, 21 and 18 out of the 47 proteins were important for discriminating RO, W, B and R quinoa from the rest of classes, respectively, whereas 13 out of the 47 proteins were non-critical for differentiation. Table 1 shows in red the VIP scores values higher than 1 for the 34 detected proteins that were important for discrimination. As examples of these proteins that cover all the measured range (5,500–25,000 *m/z*) were found ribosomal protein S29 (protein 2, M_r experimental = 6,392 and M_r theoretical = 6,413, VIP score of 1.09 to discriminate W quinoa from the rest of classes), 60S acidic ribosomal protein *P*3-like (protein 30, M_r experimental = 11,892 and M_r theoretical = 11,902, VIP scores of 1.02 and 1.06 for W and R quinoa, respectively), and AtpF (chloroplast) and/or mitochondrial glycine-rich RNA-binding protein 3 (protein 47, M_r experimental = 21,029 and M_r theoretical = 20,942 / 21,121 for AtpF (chloroplast) and mitochondrial glycine-rich RNA-binding protein 3, respectively, VIP scores of 1.30 and 1.00 for W and B quinoa, respectively).

4. Conclusions

We have demonstrated that protein profiling by MALDI-TOF-MS combined with MALDIquant preprocessing and peak detection followed by multivariate data analysis is an efficient approach to classify commercial quinoa grains. After a rapid and simple protein extraction from different RO, W, B and R quinoa grains, MALDI-TOF-MS was applied to obtain mass spectra profiles of the protein extracts. Data processing with MALDIquant allowed the detection of the most relevant protein peaks in the mass spectra profiles, which were tentatively identified. The intensities of the 47 detected proteins were considered as fingerprints for multivariate data analysis. Classification of quinoa grain samples by PLS-DA was excellent, and 34 out of the 47 variables were critical for differentiation. The proposed methodology could find application in quality control and food fraud prevention programs. Furthermore, it could be also applied to protein profiling of other food products, presenting complex mass spectra profiles with highly overlapped peaks.

CRediT authorship contribution statement

Rocío Galindo-Luján: Investigation, Methodology, Formal analysis. **Laura Pont:** Supervision, Visualization, Writing – original draft, Writing – review & editing. **Victoria Sanz-Nebot:** Conceptualization, Supervision. **Fernando Benavente:** Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors are unable or have chosen not to specify which data has been used.

Acknowledgements

This study was supported by the Spanish Ministry of Economy and Competitiveness (RTI2018-097411-B-I00) and the Cathedra UB Rector Francisco Buscarons Úbeda (Forensic Chemistry and Chemical Engineering). RG thanks the National Agricultural Innovation Program and the Ministry of Education from Peru for a research stay and a PhD fellowship, respectively. We also thank Dr. Sergey Kucheryavskiy for his support using the mdatools R package.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2022.133895.

References

Aloisi, I., Parrotta, L., Ruiz, K. B., Landi, C., Bini, L., Cai, G., … Del Duca, S. (2016). New insight into quinoa seed quality under salinity: Changes in proteomic and amino acid profiles, phenolic content, and antioxidant activity of protein extracts. *Frontiers in Plant Science, 7*, 1–21. https://doi.org/10.3389/fpls.2016.00656

Álvarez, G., Montero, L., Llorens, L., Castro-Puyana, M., & Cifuentes, A. (2018). Recent advances in the application of capillary electromigration methods for food analysis and Foodomics. *Electrophoresis, 39*, 136–159. https://doi.org/10.1002/elps.201700321

Angeli, V., Silva, P. M., Massuela, D. C., Khan, M. W., Hamar, A., Khajehei, F., … 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*, 1–31. https://doi.org/10.3390/foods9020216

Bansal, S., Singh, A., Mangal, M., Mangal, A. K., & Kumar, S. (2017). Food adulteration: Sources, health risks, and detection methods. *Critical Reviews in Food Science and Nutrition, 57*, 1174–1189. https://doi.org/10.1080/10408398.2014.967834

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics, 17*, 166–173. https://doi.org/10.1002/cem.785

Borgaonkar, S. P., Hocker, H., Shin, H., & Markey, M. K. (2010). Comparison of normalization methods for the identification of biomarkers using MALDI-TOF and SELDI-TOF mass spectra. *OMICS A Journal of Integrative Biology, 14*, 115–126. https://doi.org/10.1089/omi.2009.0082

Cavanna, D., Righetti, L., Elliott, C., & Suman, M. (2018). The scientific challenges in moving from targeted to non-targeted mass spectrometric methods for food fraud analysis: A proposed validation workflow to bring about a harmonized approach. *Trends in Food Science and Technology, 80*, 223–241. https://doi.org/10.1016/j.tifs.2018.08.007

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association, 74*, 829–836. https://doi.org/10.1080/01621459.1979.10481038

Fiedler, G. M., Leichtle, A. B., Kase, J., Baumann, S., Ceglarek, U., Felix, K., … Thiery, J. (2009). Serum peptidome profiling revealed platelet factor 4 as a potential discriminating peptide associated with pancreatic cancer. *Clinical Cancer Research, 15*, 3812–3819. https://doi.org/10.1158/1078-0432.CCR-08-2701

Friedman, J. H. (1984). A variable span smoother. *Laboratory for Computational Statistics, Stanford University Technical Report No. 5*, *5*, 1–32.

Galindo-Luján, R., Pont, L., Sanz-Nebot, V., & Benavente, F. (2021). Classification of quinoa varieties based on protein fingerprinting by capillary electrophoresis with ultraviolet absorption diode array detection and advanced chemometrics. *Food Chemistry, 341*, Article 128207. https://doi.org/10.1016/j.foodchem.2020.128207

Galindo-Luján, 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*, 130250–130263. https://doi.org/10.1016/j.foodchem.2021.130250

Gan, Y., Xiao, Y., Wang, S., Guo, H., Liu, M., Wang, Z., & Wang, Y. (2019). Protein-based fingerprint analysis for the identification of ranae oviductus using RP-HPLC. *Molecules, 24*, 1–11. https://doi.org/10.3390/molecules24091687

Gibb, S. (2014). *MALDIquantForeign : Import / Export routines for MALDIquant.* 1–7. https://cran.r-project.org/package=MALDIquantForeign.

Gibb, S., & Strimmer, K. (2017). Mass spectrometry analysis using MALDIquant. In *Statistical Analysis of Proteomics, Metabolomics, and Lipidomics Data Using Mass Spectrometry* (pp. 101–124). https://doi.org/10.1007/978-3-319-45809-0_6.

Hong, E., Lee, S. Y., Jeong, J. Y., Park, J. M., Kim, B. H., Kwon, K., & Chun, H. S. (2017). Modern analytical methods for the detection of food fraud and adulteration by food

category. *Journal of the Science of Food and Agriculture, 97*, 3877–3896. https://doi.org/10.1002/jsfa.8364
Jablonski, J. E., Moore, J. C., & Harnly, J. M. (2014). Nontargeted detection of adulteration of skim milk powder with foreign proteins using UHPLC-UV. *Journal of Agricultural and Food Chemistry, 62*, 5198–5206. https://doi.org/10.1021/jf404924x
Joliffe, I. T., & Morgan, B. J. (1992). Principal Component Analysis and exploratory factor analysis. *Statistical Methods in Medical Research, 1*, 69–95. https://doi.org/10.1177/096228029200100105
Kiran, M., Naveena, B. M., Reddy, K. S., Shahikumar, M., Reddy, V. R., Kulkarni, V. V., … More, T. H. (2016). Understanding tenderness variability and ageing changes in buffalo meat: Biochemical, ultrastructural and proteome characterization. *Animal, 10*, 1007–1015. https://doi.org/10.1017/S1751731115002931
Kohlbacher, O., Reinert, K., Gröpl, C., Lange, E., Pfeifer, N., Schulz-Trieglaff, O., & Sturm, M. (2007). TOPP - The OpenMS proteomics pipeline. *Bioinformatics, 23*, e191–e197. https://doi.org/10.1093/bioinformatics/btl299
Kucheryavskiy, S. (2020). mdatools – R package for chemometrics. *Chemometrics and Intelligent Laboratory Systems, 198*, Article 103937. https://doi.org/10.1016/j.chemolab.2020.103937
Kuo, T. H., Kuei, M. S., Hsiao, Y., Chung, H. H., Hsu, C. C., & Chen, H. J. (2019). Matrix-assisted laser desorption/ionization mass spectrometry typings of edible oils through spectral networking of triacylglycerol fingerprints. *ACS Omega, 4*, 15734–15741. https://doi.org/10.1021/acsomega.9b02433
Laparra, J. M., & Haros, M. (2016). Inclusion of ancient Latin-American crops in bread formulation improves intestinal iron absorption and modulates inflammatory markers. *Food and Function, 7*, 1096–1102. https://doi.org/10.1039/c5fo01197c
Laparra, J. M., & Haros, M. (2018). Inclusion of whole flour from Latin-American crops into bread formulations as substitute of wheat delays glucose release and uptake. *Plant Foods for Human Nutrition, 73*, 13–17. https://doi.org/10.1007/s11130-018-0653-6
Li, X. (2005). Ciphergen SELDI-TOF processing. *Package PROcess, 10*, 1–18.
Nachtigall, F. M., Pereira, A., Trofymchuk, O. S., & Santos, L. S. (2020). Detection of SARS-CoV-2 in nasal swabs using MALDI-MS. *Nature Biotechnology, 38*, 1168–1173. https://doi.org/10.1038/s41587-020-0644-7
Niro, S., D'Agostino, A., Fratianni, A., Cinquanta, L., & Panfili, G. (2019). Gluten-free alternative grains: Nutritional evaluation and bioactive compounds. *Foods, 8*, 208–217. https://doi.org/10.3390/foods8060208
Pont, L., Compte, I., Sanz-Nebot, V., Barbosa, J., & Benavente, F. (2020). Analysis of hordeins in barley grain and malt by capillary electrophoresis-mass spectrometry. *Food Analytical Methods, 13*, 325–336. https://doi.org/10.1007/s12161-019-01648-8
Purohit, P. V., & Rocke, D. M. (2003). Discriminant models for high-throughput proteomics mass spectrometer data. *Proteomics, 3*, 1699–1703. https://doi.org/10.1002/pmic.200300518
R Development Core Team: (2020). *R: A language and environment for statistical computing R foundation for statistical computing*. http://www.r-project.org/.
Rodríguez, S. D., Rolandelli, G., & Buera, M. P. (2019). Detection of quinoa flour adulteration by means of FT-MIR spectroscopy combined with chemometric methods. *Food Chemistry, 274*, 392–401. https://doi.org/10.1016/j.foodchem.2018.08.140
Ryan, C. G., Clayton, E., Griffin, W. L., Sie, S. H., & Cousens, D. R. (1988). SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 34*, 396–402. https://doi.org/10.1016/0168-583X(88)90063-8
Sassi, M., Arena, S., & Scaloni, A. (2015). MALDI-TOF-MS platform for integrated proteomic and peptidomic profiling of milk samples allows rapid detection of food adulterations. *Journal of Agricultural and Food Chemistry, 63*, 6157–6171. https://doi.org/10.1021/acs.jafc.5b02384
Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*, 1639–1643. https://doi.org/10.1021/ac60214a048
Shotts, M. L., Plans Pujolras, M., Rossell, C., & Rodriguez-Saona, L. (2018). Authentication of indigenous flours (Quinoa, Amaranth and kañiwa) from the Andean region using a portable ATR-Infrared device in combination with pattern recognition analysis. *Journal of Cereal Science, 82*, 65–72. https://doi.org/10.1016/j.jcs.2018.04.005
Stahl, A., & Schröder, U. (2017). Development of a MALDI-TOF MS-based protein fingerprint database of common food fish allowing fast and reliable identification of fraud and substitution. *Journal of Agricultural and Food Chemistry, 65*, 7519–7527. https://doi.org/10.1021/acs.jafc.7b02826
Uhlmann, K. R., Gibb, S., Kalkhof, S., Arroyo-Abad, U., Schulz, C., Hoffmann, B., … Feltens, R. (2014). Species determination of Culicoides biting midges via peptide profiling using matrix-assisted laser desorption ionization mass spectrometry. *Parasites and Vectors, 7*, 3–18. https://doi.org/10.1186/1756-3305-7-392
Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*, 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1
Yang, C., He, Z., & Yu, W. (2009). Comparison of public peak detection algorithms for MALDI mass spectrometry data analysis. *BMC Bioinformatics, 10*, 1–13. https://doi.org/10.1186/1471-2105-10-4
Zambonin, C. (2021). Maldi-tof mass spectrometry applications for food fraud detection. *Applied Sciences, 11*, 3–18. https://doi.org/10.3390/app11083374