Tutor/s

Dra. Laura Pont Villanueva
Departament d'Enginyeria Química i Química Analítica

Grau de Química

Treball Final de Grau

Identification of bioactive compounds in quinoa grains following a chemometrics-assisted metabolomics approach
Identificació de compostos bioactius en grans de quinoa seguint una estratègia metabolòmica assistida per quimiometria

Josep Compte Peña

June 2024

UNIVERSITAT DE BARCELONA

B:KC Barcelona Knowledge Campus
Campus d'Excel·lència Internacional

Aquest 2024 no ha estat el millor any a nivell personal ni mental, per tant les persones que m'han ajudat durant tot aquest semestre es mereixen uns agraïments.

Primer de tot la Laura ha estat la millor tutora que es pot demanar, m'ha ajudat en tot moment (tenint en compte el meu desconeixement inicial a la part experimental) i m'ha estat guiant i solucionant tots els dubtes el més ràpid possible. Per altra part, mil gràcies a la meva família per estar sempre disposada a ajudar-me en tot i a la meva psicòloga per contribuir a la organització mental de tots els pensaments negatius que m'han sorgit.

Mil gràcies a tots.

REPORT

IDENTIFICATION AND REFLECTION ON THE SUSTAINABLE DEVELOPMENT GOALS (SDG)

In 2015, the United Nations devised a plan to transform the entire world based on an action plan to combat poverty, climate change, and ensure the prosperity and peace of all people now and in the future. This transformation plan is called “The 2030 Agenda for Sustainable Development”. It comprises a total of 17 Sustainable Development Goals that can be divided into the 5P's: People, Prosperity, Planet, Peace, and Partnership.

The main objective of this work is to provide more information to all existing studies that highlight quinoa as one of the most relevant foods today. This fact allows this work to be categorized within two major realms of sustainable development: People and Planet. Quinoa could become one of the main international foods in the future thanks to its versatility in cultivation and its high protein and polyphenols content. This adaptability makes quinoa a viable alternative to excessive meat consumption and other foods that contribute to greenhouse gas emissions and environmental degradation, thus offering benefits to both human health and the planet.

Although this work is not entirely focused on verifying the effects of quinoa consumption on individuals, the fact that quinoa contains bioactive compounds with anticarcinogenic, antioxidant, and cardiovascular benefits, along with its essential amino acids and its gluten-free nature (making it suitable for individuals with celiac disease), categorizes this work as a contribution to the third goal of the 2030 Agenda: Good Health and Well-being. This goal aims to ensure healthy lives and promote well-being for all at all ages.

GOOD HEALTH AND WELL-BEING LOGO

CONTENTS

1. SUMMARY

Quinoa (*Chenopodium Quinoa* Willd.) is an Andean grain renowned for its diverse array of bioactive compounds, harbouring considerable potential for various biological activities. This food has been the subject of study in recent years due to its high-quality protein content, as well as an excellent balance of essential amino acids. In addition, the bioactive compounds present in quinoa, mainly polyphenols, present relevant medicinal attributes such as potent antioxidant and anticarcinogenic properties, alongside a remarkable ability to mitigate cardiovascular diseases.

This work has been focused on the identification of polyphenols within four different commercially available quinoa grains: black (B, from Peru), red (R, from Peru), white (W, from Peru) and royal (RO, white quinoa from Bolivia). The identification process employed a non-targeted metabolomics strategy based on liquid chromatography coupled to mass spectrometry (LC-MS). The obtained chromatograms were cross-referenced with a database of polyphenols created from an exhaustive review of literature on low molecular mass bioactive compounds. The investigation unveiled a total of 25 polyphenols distributed among the four quinoa grains, predominantly from the quercetin and kaempferol families**.** To assess the polyphenolic composition variance among the quinoa grains, a multivariate chemometric analysis was conducted utilizing principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). The results exhibited distinct clustering of the grains into three primary groups: B, R, and W-RO. Surprisingly, minimal differentiation was observed between the white quinoas (W and RO), indicating that the origin of the white quinoa had negligible impact on its polyphenol content. Among the identified polyphenols, quercetins emerged as the most significant compounds for discriminating between the quinoa grain classes. These flavonols are renowned for their potent bioactive properties, as extensively documented in the literature.

Keywords: Quinoa, bioactive compounds, food, multivariate data analysis, metabolomics

2. RESUM

La quinoa (*Chenopodium Quinoa* Willd.) és un gra andí conegut per la seva diversa gamma de compostos bioactius, que té un considerable potencial per a diverses activitats biològiques. Aquest aliment ha estat objecte d'estudi en els últims anys a causa del seu alt contingut proteic de qualitat, així com d'un excel·lent equilibri d'aminoàcids essencials. A més, els compostos bioactius presents a la quinoa, principalment polifenols, presenten atributs medicinals rellevants com a propietats antioxidants i anticarcinogèniques potents, a més d'una notable capacitat per mitigar les malalties cardiovasculars.

Aquest treball s'ha centrat en la identificació de polifenols en quatre grans de quinoa comercials diferents: quinoa negra (B, del Perú), vermella (R, del Perú), blanca (W, del Perú) i reial (RO, quinoa blanca de Bolívia). El procés d'identificació va utilitzar una estratègia de metabolòmica no dirigida basada en la cromatografia de líquids acoblada a l'espectrometria de masses (LC-MS). Els cromatogrames obtinguts es van creuar amb una base de dades de polifenols creada a partir d'una revisió exhaustiva de la literatura sobre compostos bioactius de baixa massa molecular. La investigació va revelar un total de 25 polifenols distribuïts entre les quatre classes de quinoa, predominantment de les famílies de la quercetina i del kaempferol. Per avaluar la variabilitat de la composició de polifenols entre els grans de quinoa estudiats, es va realitzar un anàlisi quimiomètric multivariant utilitzant l'anàlisi de components principals (PCA) i l'anàlisi discriminant de mínims quadrats parcials (PLS-DA). Els resultats van mostrar un agrupament distintiu dels grans en tres grups principals: B, R i W-RO. Sorprenentment, es va observar una diferenciació mínima entre les quinoes blanques (W i RO), indicant que l'origen de la quinoa blanca té un impacte negligible en el seu contingut de polifenols. D'entre els polifenols identificats, les quercetines van emergir com els compostos més significatius per discriminar entre les classes de grans de quinoa. Aquests flavonols són coneguts pels seus potents efectes bioactius, com està àmpliament documentat en la literatura.

Paraules clau: Quinoa, compostos bioactius, alimentació, anàlisi multivariant, metabolòmica

3. INTRODUCTION

Currently, drastic climate changes such as drought or high temperatures are affecting crop yields and the traditional agricultural practices. In response, experts have been forced to seek out healthy alternatives that meet environmental adaptability requirements and provide essential nutritional properties. One of the main imbalances of chemical species within our body is the alteration of the normal redox state of cells, commonly referred to as oxidative stress [1]. This imbalance, triggered by oxidative metabolism, can inflict direct damage on cellular structures, including DNA bases affected by reactive oxygen species, and can disrupt cellular signalling [2]. Notably, oxidative stress is identified as a potential contributor to the development of cancer and degenerative diseases, such as Alzheimer's [3]. Addressing these complexities can be partially achieved through the consumption of foods rich in healthy bioactive compounds with antioxidative properties, such as quinoa [4].

Quinoa (*Chenopodium quinoa* Willd.) meets all these requirements. It is a high-protein food that contains essential amino acids such as lysine, arginine, or cysteine, while also boasting bioactive compounds currently being studied for their significant role as antioxidants [5]. Moreover, quinoa is gluten-free [6] and highly versatile, with plantations spanning across Europe, Asia, and America. Recognizing its potential, the FAO (Food and Agriculture Organization) has identified quinoa as one of the most promising crops for sustainable food production [7].

Quinoa grains are the most commercially traded part of the plant, primarily due to their high content of protein, starch, essential amino acids, minerals, and bioactive compounds [8]. The presence of these bioactive compounds classifies quinoa as a functional food [9], meaning it contains biologically active components that provide added health benefits beyond its nutritional value, reducing the risk of certain diseases. Among the main bioactive compounds found in quinoa grains are carotenoids, saponins, terpenes, sterols, and notably polyphenols, including phenolic acids (such as ferulic, caffeic, p-coumaric, benzoic, and vanillic acids) [10] and flavonoids (such as apigenin, kaempferol, myricetin, quercetin, and rutin) [11].

In recent years, numerous studies have supported the beneficial effects of polyphenol intake on health. These compounds, known for their immunomodulatory, antimicrobial, antihypertensive, anti-inflammatory, anticarcinogenic, and particularly antioxidant properties, have shown potential in preventing various diseases, including diabetes, cancer, obesity, infections, cardiovascular, and neurodegenerative diseases, among others [12], [13]. Polyphenols are compounds that feature one or more aromatic rings with hydroxyl groups in their structure [11]. Among the various types of polyphenols, two main groups stand out: flavonoids and phenolic acids [13-15]. More than 8000 phenolic structures are currently known, being classified by their source of origin, biological function, and chemical structure [17]. Among them, over 4000 flavonoids have been identified in different foods, making them the most important group in polyphenols [18]. Flavonoid compounds can be divided into different subgroups, although in quinoa, flavonols are the most common, especially kaempferol and quercetin, along with their respective derivatives [7]. These two flavonols exhibit significant antioxidant activities, helping to prevent stress-related diseases, as well as cardiovascular diseases and diabetes [15], [16]. Additionally, it is known that flavonoids are also responsible for the colour of quinoa [19].

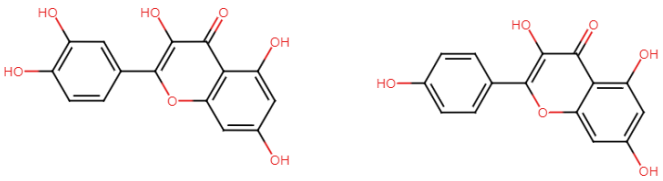

Figure 1. (A) Quercetin and (B) kaempferol structures (drawn with Reaxys, https://www-reaxys-com.sire.ub.edu/#/search/quick/query).

Despite the widely recognized beneficial effects of polyphenols in quinoa grains, their abundance may differ between grain classes, primary influenced by cultivation strategies, location, and growth factor conditions, among others. This study aimed to identify low molecular mass bioactive compounds, particularly polyphenols, present in four commercially available quinoa grains: black (B, from Peru), red (R, from Peru), white (W, from Peru), and royal (RO, white quinoa from Bolivia). A metabolomics approach using liquid chromatography coupled to mass spectrometry (LC-MS) was employed. Metabolomics involves the study of small molecules produced by the body from food decomposition or chemical substances, among others [19].

There are two strategies that can be applied in metabolomics: targeted analysis and untargeted analysis. Targeted metabolomics identifies a defined set of molecules [19]. However, this approach severely limits the identification of previously unknown compounds, as the experimental process focuses on a reduced number of bioactive compounds. For a general determination of polyphenols in quinoa grains, an untargeted analysis is required, allowing a broader identification of metabolites. This approach is guided by the experimental workflow, which involves the use of appropriate separation and detection techniques. LC-MS is the most commonly used technique in this type of strategy due to its high sensitivity and ability to identify many compounds in a single analysis.

In our work, after obtaining the LC-MS fingerprints of quinoa bioactive compounds, polyphenols were identified through an extensive database search. Subsequently, multivariate data analysis tools, such as principal component analysis, PCA, and partial least squares-discriminant analysis, PLS-DA, were used to assess metabolomic differences in the polyphenol content between the different quinoa grain classes.

4. OBJECTIVES

The objective of this study was to identify low molecular mass bioactive compounds, particularly polyphenols, in four different commercially available quinoa grains: B, R, W and RO. This was achieved through an LC-MS untargeted metabolomics approach combined with an extensive database searching. A total of 12 samples, 3 independent replicates for each quinoa grain class, were analysed. Finally, advanced multivariate data analysis tools, including PCA and PLS-DA, were utilized to classify the four quinoa classes and asses metabolomic differences between them.

5. EXPERIMENTAL SECTION

5.1. SAMPLE TREATMENT

As previously mentioned, a total of 12 samples were analysed, i.e., three independent replicates for each quinoa grain class (B, R, W, RO). Briefly, quinoa grains were ground into a fine powder (50 mg) and homogenized with a methanol/H_2O solution (4:1, v/v) containing 0.1% (v/v) of formic acid (2.5 mL). After resting for 20 min at room temperature, the mixture underwent centrifugation at 4°C for 20 min at 1000 x g. The resulting supernatant was then subjected to a second extraction using the same solution (1 mL) and centrifugation method. The combined supernatants were filtered through a 0.20 µm nylon filter before LC/MS analysis.

Table 1. Code for the studied quinoa samples (classified by colour).

Black (B)			Red (R)			White (W)			Royal (RO)		
B.1	B.2	B.3	R.1	R.2	R.3	W.1	W.2	W.3	RO.1	RO.2	RO.3

5.2. LC-MS ANALYSIS

LC-MS experiments were performed in a 1260 Infinity liquid chromatograph coupled to a 6546 LC/QTOF mass spectrometer with an orthogonal electrospray ionization (ESI) interface (Agilent Technologies, Waldbronn, Germany). For the separation, a Zorbax SB-C18 column (150 mm total length (L_T) × 2.1 mm internal diameter (ID), 5 µm particle size, 90 Å pore diameter, Agilent Technologies) was used. Experiments were carried out at a flow rate of 350 µL/min. Mobile phase solvents were (A) water and (B) acetonitrile (both with 0.1% (v/v) of formic acid). The optimized elution gradient of solvent B was: 5% (v/v) for 1 min and from 5% to 95% (v/v) in 15 min, followed by cleaning and reequilibration steps at 95% (v/v) for 2 min, from 95% to 5% (v/v) in 2 min, and finally 5% (v/v) for 5 min. The injection volume was 5 µL. The mass spectrometer was operated in negative ESI mode using the following parameters:

Table 2. MS parameters for LC-MS analyses.

Capillary Voltage (V)	Drying gas T (°C)	Flow rate (L/min)
3500	350	8

Nebulizer gas (psi)	Fragmentor voltage (V)	Skimmer voltage (V)	OCT 1 RF Vpp Voltage (V)
30	150	60	300

MassHunter software version B.10.0 (Agilent Technologies) was used for instrument control, data acquisition, and data processing.

5.3. DATA ANALYSIS

5.3.1. Creation of a database of polyphenols

A comprehensive database of polyphenols was created after an exhaustive search in the literature [2-12], [21-48] .This database included information about the molecular formula and the theoretical molecular mass of a total of 277 polyphenols, including phenolic acids (81), flavonoids (153) and other compounds that can be classified as polyphenols (43, including phenolic aldehydes, methoxyphenols, and phenolic glucosides).

5.3.2. Identification of polyphenols through a database search

Raw LC-MS data files from the quinoa extracts were crossed against the database of polyphenols (which needed to be in a specific .csv format) using the Molecular Feature Extraction (MFE) tool of the MassHunter software. This tool performs mass spectra deconvolution among the chromatographic peaks, allowing to identify compounds (called molecular features) and compare them with those found in suitable databases. MFE has the particularity that locates ions that are covariant in the LC-MS dataset and that are logically related by charge-state envelope, isotopic distribution, and/or presence of adducts. Using this approach, the MFE algorithm can distinguish multiple co-eluting compounds that would appear as a single peak in a chromatogram. The MFE workflow parameters are described in Table 3.

Table 3. MFE workflow parameters.

Time range (min)	Mass error (ppm)	Adducts	Charges
0-18	20	hydrogen (-H)	One ([M-H]-)
Mass score	**Isotope abundance score**	**Isotope spacing score**	**Expected MS variations**
100	60	50	2.0 mDa + 5.6 ppm

5.3.3. Multivariate data analysis

After utilizing the MFE tool, a comprehensive list of molecular features was generated for each quinoa sample, detailing the compound name, the retention time (t_r), the experimental molecular mass, and the experimental *m/z* of the $[M-H]^-$ molecular ions. To enhance identification reliability, only features present in at least two out of the three replicates were selected for each quinoa sample. Subsequently, peak areas for the selected molecular features were meticulously extracted from the full scan LC-MS raw data, using the provided *m/z* value and the t_r as references. These areas were then utilized to construct a data matrix containing the peak area of each molecular feature (i.e., polyphenol) in every sample. This data matrix underwent PCA and PLS-DA analysis, both conducted with SOLO (Version 9.2.1, Eigenvector Research Incorporated, Wenatchee, WA, USA).

PCA was primarily utilized to observe similarities among the studied quinoa samples, facilitating an unsupervised identification of trends to interpret possible groupings and to detect outlier samples [50]. In contrast, PLS-DA enables the classification of quinoa varieties and assists in identifying the most influential variables contributing to class separation. In our work, the relevance of polyphenols for classification was assessed through their variable importance in projection (VIP) scores, offering a graphical representation of their significance. Additionally, PLS-DA provided quality parameters such as sensitivity and specificity, offering insight into the representativeness of the model [51].

6. IDENTIFICATION OF POLYPHENOLS

As previously mentioned, 12 samples of commercial quinoa were analysed by LC-MS, specifically three independent replicates of each quinoa grain class: B, R, W and RO. Chromatographic peaks were assigned using a comprehensive database of polyphenols and the MFE tool of the MAs shunter software [2-12], [21-48]. A list of molecular features was provided for each quinoa sample, and, after manually checking all the selected features in the full scan LC-MS raw data, a total of 25 polyphenols were identified considering the four quinoa grains.

Table 1 shows the peak number (from 1 to 25), the t_r (with the percentage of relative standard deviation, %RSD), the theoretical *m/z* of the $[M-H]^-$ molecular ion, the mass error (in ppm), the compound name, and the presence or not (tick or x, respectively) of the identified polyphenols in the different samples. It is worth mentioning that most of the identified polyphenols were classified as flavonoids (19), primarily belonging to the quercetin and kaempferol families, and phenolic acids (5). The only one that is considered as an “other polyphenol” is catechol (benzenediol).

Table 4. List of polyphenols identified by LC-MS-based metabolomics in the four commercially available quinoa grains.

Peak Number	t_r	RSD(%)	*m/z* theo $[M-H]^-$	Error (ppm)	Identification	RO	R	B	W
1	0.99	0.7	477.1402	13.3	3,5-Digallosylshikimic acid	☑	☑	☑	☑
2	1.01	1.5	431.0984	24.2	Apigenin 6-C-Glucoside	☑	☑	☑	☑
3	1.45	0.0	387.1085	15.2	5-Hydroxy-3,3',4',7,8-pentamethoxyflavone	⊗	☑	⊗	⊗
4	4.37	0.0	109.0295	16.5	Catechol	⊗	☑	⊗	⊗
5	6.48	0.5	137.0244	25.2	Salicylic acid	⊗	☑	☑	⊗
6	7.19	4.4	355.1035	0.6	1-O-feruloyl-beta-D-glucose	☑	☑	⊗	⊗
7	7.21	0.2	349.0565	14.3	Ethyl-m-digallate	☑	☑	☑	☑
8	7.27	0.2	305.0667	17.0	Epigallocatechin	☑	☑	☑	☑
9	7.93	0.2	755.2040	5.8	Quercetin-3-O-rutinoiside-7-O-rhamnoside	☑	☑	☑	☑
10	8.07	0.1	741.1884	0.4	Quercetin-3-O-(2''-O-arabinosyl)rutinoside	☑	☑	☑	☑
11	8.16	0.1	739.2091	0.2	Kaempferol-3-O-rutinoside-7-O-rhamnoside	☑	☑	☑	☑
12	8.22	0.3	769.2197	9.8	Tamarixetin-3-O-glucoside-7-O-rhamnoside	☑	☑	☑	☑
13	8.25	0.3	609.1461	1.1	Rutin	⊗	☑	☑	☑
14	8.35	0.1	631.0941	10.4	Myricetin	⊗	☑	☑	☑
15	8.34	0.1	725.1935	4.7	Kaempferol-3-O-(2-O-Xylosyl-6-O-Rhamnosyl)Glucoside	☑	☑	☑	☑
16	8.35	0.2	595.1305	2.6	Quercetin-3-O-sambubioside	☑	☑	☑	☑
17	8.41	0.3	609.1461	13.8	Rutin	⊗	☑	☑	☑
18	8.49	0.0	593.1512	4.5	Quercetin-3,7-Di-O-rhamnoside	☑	⊗	☑	☑
19	8.52	0.2	607.1305	20.4	Luteolin-7-O-glucuronide-5-O-rhamnoside	⊗	☑	⊗	⊗
20	8.54	0.3	758.1911	9.2	Delphidin	⊗	☑	☑	☑
21	8.65	0.2	579.1355	12.6	Quercetin-3-O-Rhamnosyl	☑	⊗	☑	☑
22	8.66	0.1	477.0675	2.8	Quercetin-3-O-glucuronide	⊗	☑	☑	☑
23	8.72	0.1	163.0401	21.5	Caffeic Aldehide	☑	☑	☑	☑
24	8.85	0.1	623.1618	8.9	Tamarixetin-3-O-rutinoside	⊗	☑	⊗	⊗
25	9.09	0.0	461.0725	3.0	Luteolin-7-O-glucuronide	⊗	☑	⊗	☑

As can be seen in the table, the %RSD in t_r was adequate for all the identified polyphenols, with values lower than 4.4%. Mass errors were also appropriate, between 0.5 and 25 ppm, confirming the reliability of the identifications. However, it can be observed that rutin (peak number 13 and 17) appeared at two different t_r (8.25 and 8.41). This is because, at this stage, and taking advantage of the highly accurate experimental M_r provided by the QT OF mass spectrometer in full scan mode, identifications were solely based on the measured experimental M_r. In the future, tandem-mass spectrometry (MS/MS) data will be used to improve the reliability of the identity assignments and unequivocally confirm the structure of the proposed compounds.

The Venn diagram in Figure 2 shows the relationships between the number identified polyphenols in the four quinoa grain classes. As can be seen in this figure, a similar total number of polyphenols were identified in B, R, W and RO quinoa (i.e., 19, 23, 19, and 14, respectively, from the total of 25 identified polyphenols). Among them, 12 polyphenols (48% of the total) were

identified in all the grains, while 13 polyphenols (52% of the total) were only present in some of them. Regarding polyphenols identified in only one class, 4 polyphenols were exclusively identified in R quinoa, corresponding to peak numbers 3, 4, 19, and 24) suggesting their importance as markers for discriminating between R and the rest of classes.

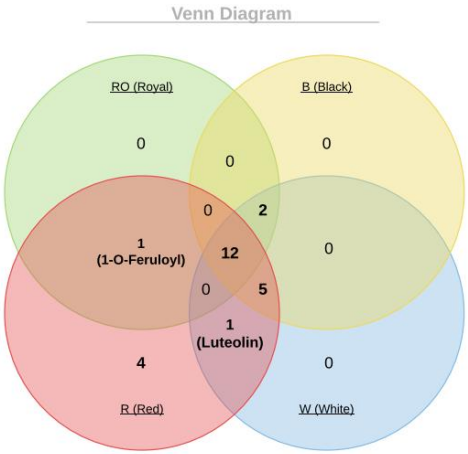

Figure 2. Venn diagram of the identified polyphenols in the four analysed quinoa grain classes, B, R, W, RO.

As an example, Figure 3 shows the base peak chromatograms (BPCs) obtained by LC-MS for the quinoa extracts of RO, R, W, and B quinoa. As it can be seen, the LC-MS profile for the four quinoa grain classes was very similar at naked eye, and, in all cases, polyphenols appeared at t_r lower than 10 min. However, subtle differences in the intensity of the four BPCs were discernible, suggesting a different concentration of polyphenols among the four quinoa grain classes.

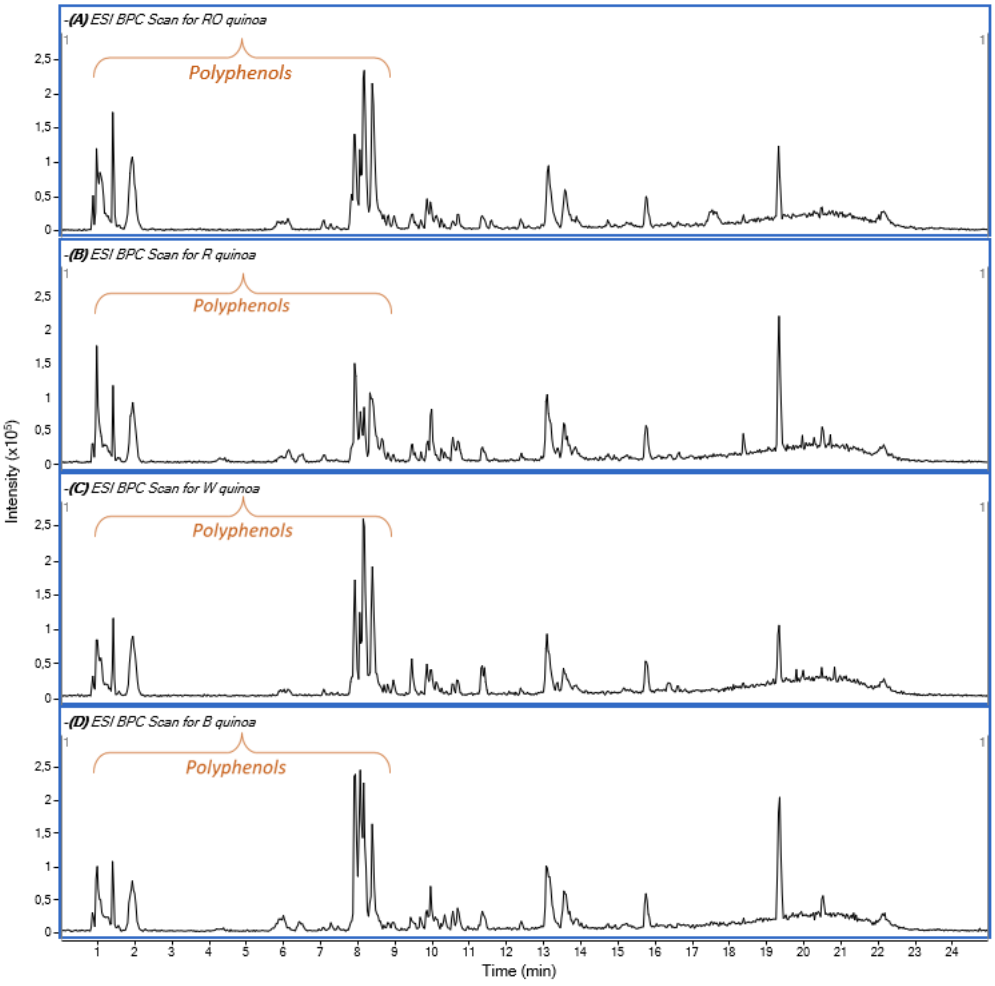

Figure 3: Base peak chromatograms (BPCs) obtained by LC-MS for the quinoa extracts of a (A) RO, (B) R, (C) W, and (D) B sample.

7. MULTIVARIATE DATA ANALYSIS

Despite the LC-MS profile and the number of identified polyphenols was similar between the four quinoa grains, it was necessary to consider differences at the concentration level for a more confident discrimination. For this purpose, a matrix containing the areas of the identified polyphenols (25 columns) in each sample (four quinoa grains per triplicate, in total, 12 rows) was built and subjected to multivariate data analysis, i.e., PCA and PLS-DA.

7.1. PCA

An initial exploratory analysis was conducted by PCA. This method enables a preliminary visualisation of the data by reducing its dimensionality, while preserving all relevant information from the matrix variables. These variables, known as principal components (PCs), are vectors derived from the linear combination of the original data, indicating the direction of maximum variation. Qualitative interpretation in PCA is based on two matrices generated from the model: the scores matrix and the loadings matrix [44]. The scores matrix describes the behaviour of the samples across the PCs, while the loadings matrix describes the behaviour and significance of each variable in contributing information to each PC. It is worth noting that an autoscaling pre-processing step is, in general, mandatory, alongside cross-validation (in our case, leave-one-out due to the number of replicates) to ensure the creation of a representative model [52].

As shown in the scores plot of Figure 4, two PCs explained 69% of the variance. Three groups were clearly separated by PC1 and PC2 (42% and 27% of the explained variance, respectively), namely R, B, and W-RO. In this context, minimal differentiation was observed between the white quinoas (W and RO), highlighting that their different origin (Peru and Bolivia, respectively) had a negligible impact on the polyphenol content. Notably, all the replicates (_1, _2, and _3) from a quinoa grain class were clustered together, suggesting the excellent repeatability of the extraction procedure and the LC-MS analysis. Only one replicate was clustered outside the 95% confidence

ellipse, which was considered as an outlier and discarded for the construction of the PLS-DA model.

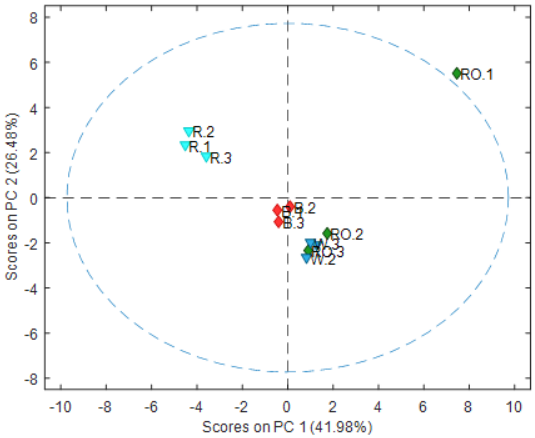

Figure 4. Scores plot of the PCA model applied to the 12 quinoa samples using the areas of the 25 identified polyphenols.

7.2. PLS-DA

Once the data was explored by PCA, three classes were defined (i.e., R, B, and W-RO) to build a PLS-DA model with improved class separation and to reveal the importance of the different polyphenols for discrimination between the quinoa grain classes. The matrix contained the areas of the identified polyphenols (25 columns) in each sample (12 samples - 1 outlier, 11 rows). As can be seen in the scores plot of Figure 5, latent variable 1 (LV1) (46% of the explained variance) allowed separating R from W-RO quinoa samples, whereas LV2 (18% of the explained variance) allowed discriminating B from R and W-RO quinoa samples.

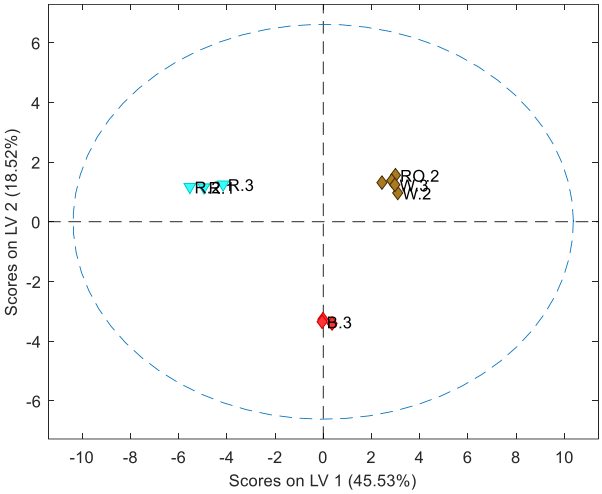

Figure 5. Scores plot of the PLS-DA model applied to the 11 quinoa samples using the areas of the 25 identified polyphenols.

In addition to the explained variance, the quality of a PLS-DA model can be assessed using specific parameters, namely, specificity and sensitivity. Sensitivity measures the model's ability to correctly classify samples from the case class, while specificity measures its capacity to predict samples from the control class [53]. In fact, a model is considered effective when both parameters are close to 1. In our work, both specificity and sensitivity were 1, confirming the strong classification performance. Another way to assess the model's classification was through the confusion matrix (Table 5), where we can observe that all the samples were correctly assigned to their corresponding class (3 in B, 3 in R, and 5 in W-RO).

Table 5. Confusion matrix of the PLS-DA model.

Confusion Table (CV):	Actual Class		
	B.1	R.1	W-RO
Predicted as B.1	3	0	0
Predicted as R.1	0	3	0
Predicted as W-RO	0	0	5
Predicted as Unassigned	0	0	0

The loadings plot provided insights into variable correlations, helping to identify the most contributing variables to each quinoa grain class (Figure 6). By comparing this information with the previously observed scores plot, it becomes evident that all the variables associated with the W-RO quinoa samples were clustered int the upper right quadrant (LV1 between 0 and 0.2, LV2 between 0 and 0.3). Variables related to R quinoa were clustered in the upper left quadrant (LV1 between 0 and -0.3, LV2 between 0 and 0.3), while those related to B quinoa were grouped in the lower part of the graph (LV1 between -0.2 and 0.2, LV2 between 0 and -0.4).

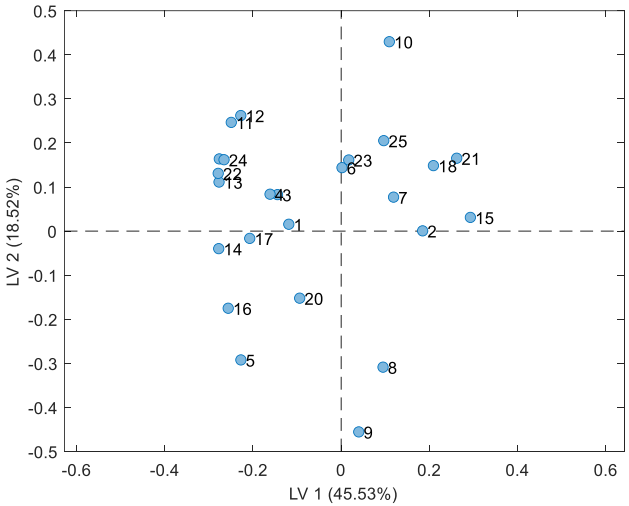

Figure 6: Loadings plot of the PLS-DA model applied to the 11 quinoa samples using the areas of the 25 identified polyphenols.

To identify the most relevant variables (polyphenols) for classifying the quinoa grain samples, the VIP scores were used. Quantitatively, only variables with VIP values > 1 are deemed significant for classification [54]. The plots of Figure 7 show the VIP scores of the different variables when considering separation of W-RO (A), B (B), and R (C) quinoa from the rest of classes, respectively.

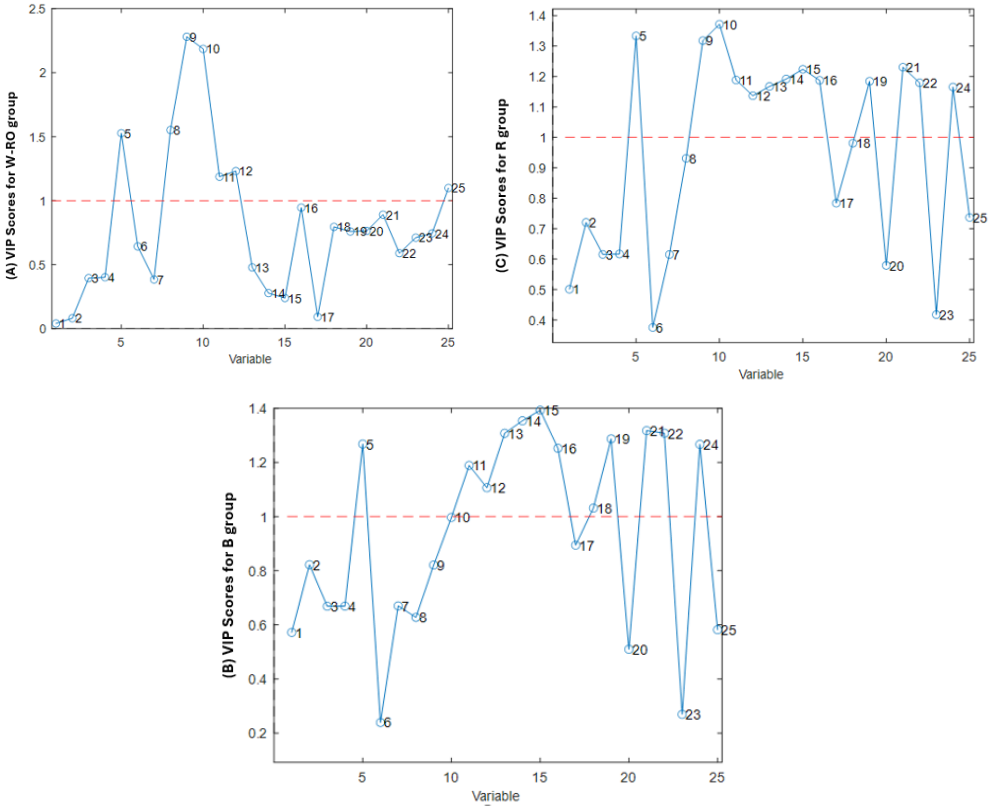

Figure 7: VIP scores of the different variables (polyphenols) when considering the separation of (A) W-RO, (B) B, and (C) R samples from the rest of classes. Polyphenols with a VIP score value higher than 1 are significant for the differentiation.

In order to differentiate between W-RO and the rest of the classes, the most relevant polyphenols were salicylic acid (peak number 5), quercetin-3-O-rutinoiside-7-O-rhamnoside (peak number 9), quercetin-3-O-(2"-O-arabinosyl)rutinoside (peak number 10) and 1-O-feruloyl-beta-D-glucose (peak number 6). Similarly, for discriminating R quinoa, the same polyphenols were relevant, together with Quercetin-3,7-Rhamnosyl (peak number 21) and Kaempferol-3-O-(2-O-Xylosyl-6-O-Rhamnosyl)Glucoside (peak number 15). Many of those compounds have hight intensity in the LV2 direction (as we can see in figure 6), which allowed the separation of W-RO and R (groups with almost no contribution in LV2) from B quinoa as I said before.

In the case of B quinoa, the most relevant polyphenols to differentiate between the rest of the classes were kaempferol-3-O-(2-O-xylosyl-6-O-rhamnosyl)glucoside, myricetin (peak number 14), quercetin-3-O-glucuronide (peak number 22) and Quercetin-3,7-Rhamnosyl (peak number 21), compounds with hight intensity in LV1 direction allowing the discrimination of B from the other groups. Those compounds mentioned for each class are the most relevant besides every compound with VIP>1 is considered important for the classification. Interestingly, most of the abovementioned polyphenols belong to the family of quercetins [49] , which have been described in the literature as potential antioxidant bioactive compounds. Quercetins are recognized for their therapeutic potential in the treatment of various diseases, such as cancer, inflammation, and cardiovascular disorders, among others [16], underscoring the significance of studying these metabolites.

Although valuable information was obtained through the application of PLS-DA and VIPs, a more comprehensive study utilizing the areas of identified compounds, along with statistical methods to identify significant differences between them, will be conducted. This approach will not only enable us to determine which compounds are critical for differentiation, but also to ascertain whether they are up-regulated, down-regulated, or remain unaltered between classes.

8. CONCLUSIONS

In this work, we analysed samples from four different commercially available quinoa grains: B, R, W, and RO. The main objective was to employ an untargeted metabolomics strategy based on LC-MS to identify low molecular mass bioactive compounds, particularly, polyphenols. Following LC-MS analysis and database searching, a total of 25 polyphenols, comprising phenolic acids and flavonoids, were identified. To effectively discriminate between the four quinoa grain classes, a matrix containing the areas of the identified compounds in all the samples was constructed. Subsequently, multivariate data analysis techniques, including PCA and PLS-DA, were applied.

Through the application of these chemometric tools, it became evident how the quinoa grains were categorized based on their polyphenol content, resulting in three distinct groups: R, B, and W-RO. According to the VIP values, the most relevant polyphenols for discriminating W-RO samples were salicylic acid, quercetin-3-O-rutinoiside-7-O-rhamnoside, quercetin-3-O-(2"-O-arabinosyl)rutinoside and 1-O-feruloyl-beta-D-glucose. Similarly, for discriminating R quinoa, the same polyphenols were relevant, together with Quercetin-3,7-Rhamnosyl and Kaempferol-3-O-(2-O-Xylosyl-6-O-Rhamnosyl)Glucoside. In the case of B quinoa, the most relevant polyphenols to differentiate between the rest of the classes were kaempferol-3-O-(2-O-xylosyl-6-O-rhamnosyl)glucoside, myricetin, quercetin-3-O-glucuronide and Quercetin-3,7-Rhamnosyl. Remarkably, most of these polyphenols belong to the family of quercetins, widely known for their anticancerogenic, anti-inflammatory and, specially, antioxidant properties.

In the future, a more comprehensive study will be conducted, utilizing MS/MS data to enhance the reliability of the identity assignments. In addition, in-vivo studies will be performed to precisely assess the bioavailability of the list of identified polyphenols. Through these studies, we aim not only to evaluate the impact of polyphenols on human health, but also to recommend an optimal daily dosage of quinoa to ensure noticeable positive effects.

9. REFERENCES AND NOTES

[1] M. M. Al-Qabba, M. A. El-Mowafy, S. A. Althwab, H. A. Alfheeaid, T. Aljutaily, and H. Barakat, "Phenolic profile, antioxidant activity, and ameliorating efficacy of chenopodium quinoa sprouts against ccl4-induced oxidative stress in rats," Nutrients, vol. 12, no. 10, pp. 1–15, Oct. 2020, doi: 10.3390/nu12102904.

[2] A. Vazquez-Luna, V. P. Cortés, F. F. Carmona, and R. Díaz-Sobac, "Quinoa leaf as a nutritional alternative," Cienc Investig Agrar, vol. 46, no. 2, pp. 137–143, 2019, doi: 10.7764/rcia.v46i2.2098.

[3] M. J. Koziol' Latinreco, "Chemical Composition and Nutritional Evaluation of Quinoa (Chenopodium quinoa Willd.)," 1992.

[4] J. V. Gil, A. Esteban-Muñoz, and M. T. Fernández-Espinar, "Changes in the polyphenolic profile and antioxidant activity of wheat bread after incorporating quinoa flour," Antioxidants, vol. 11, no. 1, Jan. 2022, doi: 10.3390/antiox11010033.

[5] M. García-Parra, D. Roa-Acosta, V. García-Londoño, B. Moreno-Medina, and J. Bravo-Gomez, "Structural characterization and antioxidant capacity of quinoa cultivars using techniques of ft-mir and uhplc/esi-orbitrap ms spectroscopy," Plants, vol. 10, no. 10, 2021, doi: 10.3390/plants10102159.

[6] F. M. V. Farinazzi-Machado, S. M. Barbalho, M. Oshiiwa, R. Goulart, and O. Pessan Junior, "Use of cereal bars with quinoa (Chenopodium quinoa W.) to reduce risk factors related to cardiovascular diseases," Food Science and Technology, vol. 32, no. 2, pp. 239–244, Apr. 2012, doi: 10.1590/s0101-20612012005000040.

[7] G. Ren et al., "Nutrient composition, functional activity and industrial applications of quinoa (Chenopodium quinoa Willd.)," Food Chemistry, vol. 410. Elsevier Ltd, Jun. 01, 2023. doi: 10.1016/j.foodchem.2022.135290.

[8] C. Alasalvar, S. K. Chang, B. Bolling, W. Y. Oh, and F. Shahidi, "Specialty seeds: Nutrients, bioactives, bioavailability, and health benefits: A comprehensive review," Compr Rev Food Sci Food Saf, vol. 20, no. 3, pp. 2382–2427, May 2021, doi: 10.1111/1541-4337.12730.

[9] A. M. M. Filho, M. R. Pirozi, J. T. D. S. Borges, H. M. Pinheiro Sant'Ana, J. B. P. Chaves, and J. S. D. R. Coimbra, "Quinoa: Nutritional, functional, and antinutritional aspects," Crit Rev Food Sci Nutr, vol. 57, no. 8, pp. 1618–1630, May 2017, doi: 10.1080/10408398.2014.1001811.

[10] A. Itoh et al., "Hepatoprotective Effect of Syringic Acid and Vanillic Acid on Concanavalin A-Induced Liver Injury," 2009.

[11] M. M. Al-Qabba, M. A. El-Mowafy, S. A. Althwab, H. A. Alfheeaid, T. Aljutaily, and H. Barakat, "Phenolic profile, antioxidant activity, and ameliorating efficacy of chenopodium quinoa sprouts against ccl4-induced oxidative stress in rats," Nutrients, vol. 12, no. 10, pp. 1–15, Oct. 2020, doi: 10.3390/nu12102904.

[12] J. H. Park, Y. J. Lee, Y. H. Kim, and K. S. Yoon, "Antioxidant and antimicrobial activities of Quinoa (Chenopodium quinoa Willd.) Seeds Cultivated in Korea," Prev Nutr Food Sci, vol. 22, no. 3, pp. 195–202, Sep. 2017, doi: 10.3746/pnf.2017.22.3.195.

[13] S. Navruz-Varli and N. Sanlier, "Nutritional and health benefits of quinoa (Chenopodium quinoa Willd.)," Journal of Cereal Science, vol. 69. Academic Press, pp. 371–376, May 01, 2016. doi: 10.1016/j.jcs.2016.05.004.

[14] R. Tsao, “Chemistry and biochemistry of dietary polyphenols,” Nutrients, vol. 2, no. 12. MDPI AG, pp. 1231–1246, 2010. doi: 10.3390/nu2121231.
[15] A. Periferakis et al., “Kaempferol: Antimicrobial Properties, Sources, Clinical, and Traditional Applications,” International Journal of Molecular Sciences, vol. 23, no. 23. MDPI, Dec. 01, 2022. doi: 10.3390/ijms232315054.
[16] Deepika and P. K. Maurya, “Health Benefits of Quercetin in Age-Related Diseases,” Molecules, vol. 27, no. 8. MDPI, Apr. 01, 2022. doi: 10.3390/molecules27082498.
[17] D. M. Kopustinskiene, V. Jakstas, A. Savickas, and J. Bernatoniene, “Flavonoids as anticancer agents,” Nutrients, vol. 12, no. 2. MDPI AG, Feb. 01, 2020. doi: 10.3390/nu12020457.
[18] L. Bravo, “Polyphenols: Chemistry, Dietary Sources, Metabolism, and Nutritional Significance,” 1998. [Online]. Available: https://academic.oup.com/nutritionreviews/article/56/11/317/1901722
[19] G. Qian et al., “Metabolomics analysis reveals the accumulation patterns of flavonoids and phenolic acids in quinoa (Chenopodium quinoa Willd.) grains of different colors,” Food Chem X, vol. 17, Mar. 2023, doi: 10.1016/j.fochx.2023.100594.
[20] A. M. Gómez-Caravaca, A. Segura-Carretero, A. Fernández-Gutiérrez, and M. F. Caboni, “Simultaneous determination of phenolic compounds and saponins in quinoa (Chenopodium quinoa Willd) by a liquid chromatography-diode array detection-electrospray ionization-time-of-flight mass spectrometry methodology,” J Agric Food Chem, vol. 59, no. 20, pp. 10815–10825, Oct. 2011, doi: 10.1021/jf202224j.
[21] N. Kumar and N. Goel, “Phenolic acids: Natural versatile molecules with promising therapeutic applications,” Biotechnology Reports, vol. 24. Elsevier B.V., Dec. 01, 2019. doi: 10.1016/j.btre.2019.e00370.
[22] Y. Brend, L. Galili, H. Badani, R. Hovav, and S. Galili, “Total Phenolic Content and Antioxidant Activity of Red and Yellow Quinoa (<i>Chenopodium quinoa</i> Willd.) Seeds as Affected by Baking and Cooking Conditions,” Food Nutr Sci, vol. 03, no. 08, pp. 1150–1155, 2012, doi: 10.4236/fns.2012.38151.
[23] R. A. Carciochi, L. Galván-D’Alessandro, P. Vandendriessche, and S. Chollet, “Effect of Germination and Fermentation Process on the Antioxidant Compounds of Quinoa Seeds,” Plant Foods for Human Nutrition, vol. 71, no. 4, pp. 361–367, Dec. 2016, doi: 10.1007/s11130-016-0567-0.
[24] P. Paśko, P. Zagrodzki, H. Bartoń, J. Chłopicka, and S. Gorinstein, “Effect of Quinoa Seeds (Chenopodium quinoa) in Diet on some Biochemical Parameters and Essential Elements in Blood of High Fructose-Fed Rats,” Plant Foods for Human Nutrition, vol. 65, no. 4, pp. 333–338, Dec. 2010, doi: 10.1007/s11130-010-0197-x.
[25] S. Sohaimy, S. Mohamed, M. Shehata, T. Mehany, and M. Zaitoun, “Compositional Analysis and Functional Characteristics of Quinoa Flour,” Annu Res Rev Biol, vol. 22, no. 1, pp. 1–11, Jan. 2018, doi: 10.9734/arrb/2018/38435.
[26] I. Bettaieb, S. Bourgou, W. A. Wannes, I. Hamrouni, F. Limam, and B. Marzouk, “Essential oils, phenolics, and antioxidant activities of different parts of cumin (Cuminum cyminum L.),” J Agric Food Chem, vol. 58, no. 19, pp. 10410–10418, Oct. 2010, doi: 10.1021/jf102248j.
[27] R. Repo-Carrasco-Valencia, J. K. Hellström, J. M. Pihlava, and P. H. Mattila, “Flavonoids and other phenolic compounds in Andean indigenous grains: Quinoa (Chenopodium quinoa), kañiwa (Chenopodium pallidicaule) and kiwicha (Amaranthus caudatus),” Food Chem, vol. 120, no. 1, pp. 128–133, May 2010, doi: 10.1016/j.foodchem.2009.09.087.
[28] Y. Tang et al., “Assessing the fatty acid, carotenoid, and tocopherol compositions of amaranth and quinoa seeds grown in Ontario and their overall contribution to nutritional quality,” J Agric Food Chem, vol. 64, no. 5, pp. 1103–1110, Feb. 2016, doi: 10.1021/acs.jafc.5b05414.

[29] D. Khang, T. Dung, A. Elzaawely, and T. Xuan, "Phenolic Profiles and Antioxidant Activity of Germinated Legumes," Foods, vol. 5, no. 4, p. 27, Apr. 2016, doi: 10.3390/foods5020027.
[30] M. García-Parra, D. Roa-Acosta, V. García-Londoño, B. Moreno-Medina, and J. Bravo-Gomez, "Structural characterization and antioxidant capacity of quinoa cultivars using techniques of ft-mir and uhplc/esi-orbitrap ms spectroscopy," Plants, vol. 10, no. 10, Oct. 2021, doi: 10.3390/plants10102159.
[31] G. Balakrishnan and R. G. Schneider, "Quinoa flavonoids and their bioaccessibility during in vitro gastrointestinal digestion," J Cereal Sci, vol. 95, Sep. 2020, doi: 10.1016/j.jcs.2020.103070.
[32] M. Pellegrini et al., "Bioaccessibility of polyphenolic compounds of six quinoa seeds during in vitro gastrointestinal digestion," J Funct Foods, vol. 38, pp. 77–88, Nov. 2017, doi: 10.1016/j.jff.2017.08.042.
[33] M. Pellegrini et al., "Chemical, fatty acid, polyphenolic profile, techno-functional and antioxidant properties of flours obtained from quinoa (Chenopodium quinoa Willd) seeds," Ind Crops Prod, vol. 111, pp. 38–46, Jan. 2018, doi: 10.1016/j.indcrop.2017.10.006.
[34] Y. Hirose, T. Fujita, T. Ishii, and N. Ueno, "Antioxidative properties and flavonoid composition of Chenopodium quinoa seeds cultivated in Japan," Food Chem, vol. 119, no. 4, pp. 1300–1306, Apr. 2010, doi: 10.1016/j.foodchem.2009.09.008.
[35] T. Kuljanabhagavad and M. Wink, "Biological activities and chemistry of saponins from Chenopodium quinoa Willd.," Phytochemistry Reviews, vol. 8, no. 2, pp. 473–490, Jun. 2009, doi: 10.1007/s11101-009-9121-0.
[36] S. Kumpun, A. Maria, S. Crouzet, N. Evrard-Todeschi, J. P. Girault, and R. Lafont, "Ecdysteroids from Chenopodium quinoa Willd., an ancient Andean crop of high nutritional value," Food Chem, vol. 125, no. 4, pp. 1226–1234, Apr. 2011, doi: 10.1016/j.foodchem.2010.10.039.
[37] U. Gawlik-Dziki, M. Świeca, M. Sułkowski, D. Dziki, B. Baraniak, and J. Czyz, "Antioxidant and anticancer activities of Chenopodium quinoa leaves extracts - In vitro study," Food and Chemical Toxicology, vol. 57, pp. 154–160, Jul. 2013, doi: 10.1016/j.fct.2013.03.023.
[38] Y. Tang et al., "Bound Phenolics of Quinoa Seeds Released by Acid, Alkaline, and Enzymatic Treatments and Their Antioxidant and α-Glucosidase and Pancreatic Lipase Inhibitory Effects," J Agric Food Chem, vol. 64, no. 8, pp. 1712–1719, Mar. 2016, doi: 10.1021/acs.jafc.5b05761.
[39] A. M. Gómez-Caravaca, A. Segura-Carretero, A. Fernández-Gutiérrez, and M. F. Caboni, "Simultaneous determination of phenolic compounds and saponins in quinoa (Chenopodium quinoa Willd) by a liquid chromatography-diode array detection-electrospray ionization-time-of-flight mass spectrometry methodology," J Agric Food Chem, vol. 59, no. 20, pp. 10815–10825, Oct. 2011, doi: 10.1021/jf202224j.
[40] Y. Tang, X. Li, B. Zhang, P. X. Chen, R. Liu, and R. Tsao, "Characterisation of phenolics, betanins and antioxidant activities in seeds of three Chenopodium quinoa Willd. genotypes," Food Chem, vol. 166, pp. 380–388, Jan. 2015, doi: 10.1016/j.foodchem.2014.06.018.
[41] L. Alvarez-Jubete, H. Wijngaard, E. K. Arendt, and E. Gallagher, "Polyphenol composition and in vitro antioxidant activity of amaranth, quinoa buckwheat and wheat as affected by sprouting and baking," Food Chem, vol. 119, no. 2, pp. 770–778, Mar. 2010, doi: 10.1016/j.foodchem.2009.07.032.
[42] H. J. Jung, Y. S. Song, C. J. Lim, and E. H. Park, "Anti-angiogenic, anti-inflammatory and anti-nociceptive activities of vanillyl alcohol," Arch Pharm Res, vol. 31, no. 10, pp. 1275–1279, Oct. 2008, doi: 10.1007/s12272-001-2106-1.

[43] P. Paśko, M. Sajewicz, S. Gorinstein, and Z. Zachwieja, “Analysis of selected phenolic acids and flavonoids in Amaranthus cruentus and Chenopodium quinoa seeds and sprouts by HPLC,” Acta Chromatogr, vol. 20, no. 4, pp. 661–672, 2008, doi: 10.1556/AChrom.20.2008.4.11.

[44] O. H. Abdelhafez et al., “Metabolomics analysis and biological investigation of three Malvaceae plants,” Phytochemical Analysis, vol. 31, no. 2, pp. 204–214, Mar. 2020, doi: 10.1002/pca.2883.

[45] L. Piñuel et al., “Production of white, red and black quinoa (Chenopodium quinoa willd var. real) protein isolates and its hydrolysates in germinated and non-germinated quinoa samples and antioxidant activity evaluation,” Plants, vol. 8, no. 8, 2019, doi: 10.3390/plants8080257.

[46] Y. Zhang, B. Bai, Y. Yan, J. Liang, and X. Guan, “Bound Polyphenols from Red Quinoa Prevailed over Free Polyphenols in Reducing Postprandial Blood Glucose Rises by Inhibiting α-Glucosidase Activity and Starch Digestion,” Nutrients, vol. 14, no. 4, Feb. 2022, doi: 10.3390/nu14040728.

[47] Y. Feng et al., “Effects of variety and origin on the metabolic and texture characteristics of quinoa seeds based on ultrahigh-performance liquid chromatography coupled with high-field quadrupole-orbitrap high-resolution mass spectrometry,” Food Research International, vol. 162, Dec. 2022, doi: 10.1016/j.foodres.2022.111693.

[48] Q. Zhang, B. Xing, M. Sun, B. Zhou, G. Ren, and P. Qin, “Changes in bio-accessibility, polyphenol profile and antioxidants of quinoa and djulis sprouts during in vitro simulated gastrointestinal digestion,” Food Sci Nutr, vol. 8, no. 8, 2020, doi: 10.1002/fsn3.1718.

[49] M. Lin, P. Han, Y. Li, W. Wang, D. Lai, and L. Zhou, “Quinoa secondary metabolites and their biological activities or functions,” Molecules, vol. 24, no. 13. MDPI AG, 2019. doi: 10.3390/molecules24132512.

[50] I. T. Joliffe and B. Morgan, “Principal component analysis and exploratory factor analysis,” Stat Methods Med Res, vol. 1, no. 1, pp. 69–95, 1992, doi: 10.1177/096228029200100105.

[51] W. Song, H. Wang, P. Maguire, and O. Nibouche, “Nearest clusters based partial least squares discriminant analysis for the classification of spectral data,” Anal Chim Acta, vol. 1009, pp. 27–38, Jun. 2018, doi: 10.1016/j.aca.2018.01.023.

[52] A. L. Pomerantsev and O. Y. Rodionova, “Procrustes Cross-Validation of short datasets in PCA context,” Talanta, vol. 226, May 2021, doi: 10.1016/j.talanta.2021.122104.

[53] E. Szymańska, E. Saccenti, A. K. Smilde, and J. A. Westerhuis, “Double-check: Validation of diagnostic statistics for PLS-DA models in metabolomics studies,” Metabolomics, vol. 8, pp. 3–16, Jun. 2012, doi: 10.1007/s11306-011-0330-3.

[54] A. L. Souza and G. J. Patti, “A Protocol for Untargeted Metabolomic Analysis: From Sample Preparation to Data Processing,” in Methods in Molecular Biology, vol. 2276, Humana Press Inc., 2021, pp. 357–382. doi: 10.1007/978-1-0716-1266-8_27.