Journal of Food Composition and Analysis 138 (2025) 107018

Contents lists available at ScienceDirect

Journal of Food Composition and Analysis

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

Classification of farming systems by NMR widely targeted metabolomics: A cauliflower case study

Federico I. Brigante [a,b], Pavel Solovyev [a], Simona Fabroni [c], Gabriele Campanelli [d], Francesco Montemurro [e], Luana Bontempo [a,b,*]

[a] Traceability Unit, Research and Innovation Center – Fondazione Edmund Mach (F.E.M), Via E. Mach 1, San Michele all'Adige, TN 38098, Italy
[b] Fondazione OnFoods, via Università 12, Parma 43121, Italy
[c] Council for Agricultural Research and Economics (CREA), Research Center for Olive, Fruit and Citrus Crops, Corso Savoia 190, Acireale 95024, Italy
[d] Council for Agricultural Research and Economics (CREA) - Research Centre for Vegetable and Ornamental Crops, Via Salaria 1, Monsampolo del Tronto 63077, Italy
[e] Council for Agricultural Research and Economics (CREA) - Research Centre for Agriculture and Environment, Via Celso Ulpiani, 5, Bari 70125, Italy

ARTICLE INFO

Keywords:
Cultivation systems
Cauliflower corymbs
Nuclear Magnetic Resonance
Metabolomics

ABSTRACT

Organic and conventional farming systems follow a set of specific rules and consumers usually pay an extra price for organic products, since they are associated with a higher quality. Nuclear magnetic resonance allows the analysis of the metabolome in a fast, non-destructive, and automatic way, with high coverage of chemical families and accurate quantification. Cauliflower corymbs obtained from two farming systems in two consecutive years were analyzed by ^{1}H NMR and cultivation systems were correctly classified by this approach. Eleven and 3 metabolites showed significantly higher concentrations in organic corymbs in the first and second year, respectively, indicating the impact of cultivation year in the search for markers in organic produce. These included amino acids (alanine), purines (xanthine), organic acids, and N-containing compounds (choline), demonstrating their higher nutritional quality.

1. Introduction

Food quality is a broad concept that is closely connected both with physicochemical aspects of food and with aspects associated with consumer perception. The former includes organoleptic traits such as color, aroma, consistency, and flavor, as well as internal composition in microbiological and chemical terms. The combination of these factors has an impact on the nutritional aspect of food, making it a topic of increasing interest among consumers, regulation authorities, and governments. The latter is linked to the individual perception of each consumer with respect to a product, as they will pay higher prices for products considered of good or premium quality (Akkerman et al., 2010, Yang et al., 2019).

Over the past years, as the global population has constantly increased, so has the demand for food that is both safe and nutritious. In this context, the advent of conventional intensification allowed the production of more food to ensure food security (guaranteed access to food by the population). However, severe environmental consequences associated with this practice led to the emergence of more environmentally friendly practices such as sustainable intensification, ecological intensification, agroecological farming, and organic farming. Compared to conventional farmers, the organic and agroecological ones do not use synthetic fertilizers; they do not allow the use of genetically modified organisms (GMOs); they are more labor-dependent; encourage spatial heterogeneity; and exploit ecosystem services. Furthermore, compared to organic, the agroecological farming also relies on the integration of livestock and a more explicit focus on traditional knowledge (Garibaldi et al., 2017, Hird, 2017, Briones Alonso et al., 2018).

Brassica oleracea is a group of Brassicaceous vegetables including broccoli, cabbage, and cauliflower, among others. This group is well-known for its micronutrient contribution to the diet, particularly for their content of vitamins and minerals and other bioactive compounds like polyphenols, carotenoids, and sulfur-containing glucosinolates, the latter being also responsible for their characteristic bitter flavor (Li et al., 2018). To study its composition after a given treatment (farming system) and draw conclusions about quality, modern integrative approaches are required. Metabolomics is an omics science that studies the changes in the metabolome (this is, all the metabolites between 80 and 1000 Da of

* Corresponding author at: Traceability Unit, Research and Innovation Center – Fondazione Edmund Mach (F.E.M), Via E. Mach 1, San Michele all'Adige, TN 38098, Italy.
E-mail address: luana.bontempo@fmach.it (L. Bontempo).

https://doi.org/10.1016/j.jfca.2024.107018
Received 12 September 2024; Received in revised form 4 November 2024; Accepted 23 November 2024
Available online 24 November 2024

molecular mass) as a response to external conditions. It is the most preferred approach for the study of tissues, living organisms and food matrices, given the combination of advanced analytical techniques that allow the identification and quantification of the metabolites and chemometrics to find classification patterns among different groups of samples (Li et al., 2020, Yuliana et al., 2021). Three different classifications have emerged within this area, namely, targeted metabolomics, untargeted metabolomics, and widely targeted metabolomics. The latter is the most recent one, and it encompasses a combination of the other two, showcasing the wide coverage of chemical families of metabolites (untargeted metabolomics), and their quantification with very good accuracy (targeted metabolomics) (Wang et al., 2021, Qi et al., 2024). Mass spectrometry has been the primary technique for the application of this approach given its sensitivity, high dynamic range, and ease of coupling to separation techniques such as liquid and gas chromatography. Nuclear magnetic resonance (NMR) offers a series of advantages including fast analysis times, little to no sample preparation, it is non-destructive, and highly reproducible. Moreover, the integration of two-dimensional techniques allows the unambiguous identification of compounds (Consonni & Cagliani, 2022, Powers et al., 2024).

In our previous study, cauliflower (*Brassica oleracea* L. *var. botrytis*) samples from agronomic trials performed in two subsequent years (2018–2019 and 2019–2020) and 3 different farming systems (conventional, organic, agroecological) were analyzed by stable isotope ratio mass spectrometry (IRMS) and a series of physicochemical parameters were also determined. The previous results included the correct classification of the dataset. Total polyphenols, the nitrogen (N) isotope ratio of the soil and the corymbs were the most important discriminant variables to differentiate organic from conventional farming (Campanelli et al., 2024). However, no previous research studied the metabolites in the aqueous extracts that classify the farming systems in these cultivars with simultaneous contribution from different treatments with different amounts of external inputs, as well as the complete characterization of the corymbs by NMR. So, the objective of the present study is to characterize cauliflower corymbs from the same crops including identification and quantification of the metabolome and to propose discriminant metabolites of the farming systems by ^{1}H NMR widely targeted metabolomics.

2. Materials and methods

2.1. Plant materials

The complete description of the experimental design for plant harvesting and sample collection can be seen in detail in our previous work (Campanelli et al., 2024). The samples were grouped according to the six fertilization strategies (1 – organic; 2 – mix-organic; 3 – conventional; 4 – mix conventional a; 5 – mix conventional b; 6 – organic + agroecological practices). A summary of the identity, quantity, and the proportion of each component of the fertilizer used (in parentheses) is presented in Table 1. The samples of cauliflower corymbs (*Brassica oleracea* L. *var. botrytis* HF1 Triomphant (Clause seed company)) used for this study were obtained from the same collections and kept powdered and freeze-dried until further analysis. From this point, samples from the first collection (2018–2019) were labeled as Year 1, and samples from the second (2019–2020) as Year 2, to avoid confusion. A total of 21 samples for Year 1 and 23 samples for year 2 were analyzed considering 4 biological replicates for each treatment or 3 in some treatments due to limited sample availability.

Table 1
Summary of the six treatments used for the agronomical trials.

Code	Treatment	Fertilizer[a]
T1	Organic	Animal pellet (3–0–0)
T2	Mix-Organic	Animal pellet (3–0–0) Ammonium nitrate (26–0–0)
T3	Conventional	Multielement synthetic "YaraMila Blustar" (12–12–17) Ammonium nitrate (26–0–0)
T4	Mix-Conventional a	Animal pellet (3–0–0) Ammonium nitrate (26–0–0)
T5	Mix-Conventional b	Animal pellet (3–0–0) Ammonium nitrate (26–0–0)
T6	Organic + Agroecological practices	Organic vegetable amendment "Vegand" (4.1,4–0)

a proportion of N-P_2O_5-K_2O, respectively, between parentheses.

2.2. Sample preparation

A 50 mg aliquot of the freeze-dried material were weighed in an Eppendorf tube and 900 µL of Milli-Q water (18.2 MΩ.cm^{-1} at 25°C, 3 ppb TOC) (Q-POD®, Millipore Advantage A10) and 100 µL of D_2O with the standard 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt, or TMSP-d_4, 0.03 % by weight (producer-specified) (Deutero GmbH, Kastellaun, Germany) for internal reference were added. Then, the samples were ultrasonicated for 30 min and they centrifuged for 15 min at 14,000 *g* (Eppendorf *minispin Plus*, Germany). Finally, 600 µL of the supernatant were transferred to NMR tubes (Merck, Norell® Standard Series™ 5 mm).

2.3. NMR spectroscopy

NMR spectra were obtained using a Bruker Avance Neo 600 spectrometer, which operated at a base frequency of 600.13 MHz for ^{1}H nuclei. The system was equipped with a Bruker BBI 600 MHz S3, 5 mm with Z gradient broadband probe and a 24-position refrigerated SampleCase autosampler (Bruker BioSpin GmbH, Rheinstetten, Germany). Spectra acquisition and processing were performed automatically with Topspin 4.3.9 and Icon NMR 6.1.0 software. Line broadening was set to 0.3 Hz for exponential window function. For the deuterium lock signal, a 9:1 mixture of H_2O and D_2O (v/v) was optimized. The proton NMR spectra utilized experimental parameters like those described for the SGF profiling routine (Spraul et al., 2009). Specifically, the noesygppr1d pulse sequence was used, with a pulse power level set to 51.16 dB (corresponding to a 25 Hz suppression window). Acquisition time was 5.25 s, dwell time was 40 us, the spectral width (SW) was set to 20.83 ppm, and data points in the time domain (TD) were 131,072 (128 K). Each spectrum consisted of 64 scans (NS) and 4 dummy scans (DS), with a relaxation delay (D1) of 10 s. The receiver gain (RG) was fixed at 16 for all spectra, and baseopt digitization mode was applied. Prior to each spectrum acquisition, automatic probe tuning (ATMA routine) and automatic shimming (TOPSHIM) were performed. All spectra were processed automatically using the TopSpin software with the apk0.noe phase correction program.

2.4. Identification and quantification of compounds by NMR

Quantitative analysis was performed using AssureNMR software (Colson et al., 2012), employing the external standard method with the ERETIC2 technique (electronic reference to access in vivo concentrations) (Akoka et al., 1999, Hong et al. 2013). This technique is based on the PULCON (pulse length-based concentration determination) principle (Watanabe et al. 2016). Two liquid samples dissolved in a 9:1 mixture of H_2O and D_2O (v/v) were analyzed: a standard sample of known concentration (2 mmol sucrose solution in water, provided by the manufacturer) and the sample of interest (which may not necessarily be the same compound as in the standard), using identical experimental parameters. Validation was conducted by periodically measuring one manufacturer standard (2 mmol sucrose in water) against another (20 mmol sucrose and hippuric acid in water each), again with the same experimental parameters. The accuracy of the external standard method used is reported to be at least 95 % (Cullen et al., 2013). For identification and quantification, peak shapes were meticulously inspected.

Identification was performed by comparison with in-house, external public databases like the Biological Magnetic Resonace Databank (BMRB) (Hoch et al., 2023) and previous research studies (Barclay et al., 2011, Lucarini et al., 2020, Ingallina et al., 2023, Ramos-Figueroa et al., 2023). The assignment, chemical shift in ppm, coupling constant in Hz when available and possible, and multiplicity were reported for each signal. Regarding quantification, only clearly identified variables with minimal or no overlap and without variation in chemical shifts due to pH changes across the aligned spectra were quantified.

2.5. Statistical analyses

All the statistical analyses were performed in R Studio (R Core Team, 2023). Samples from treatments 1, 2, and 6 were all included under the

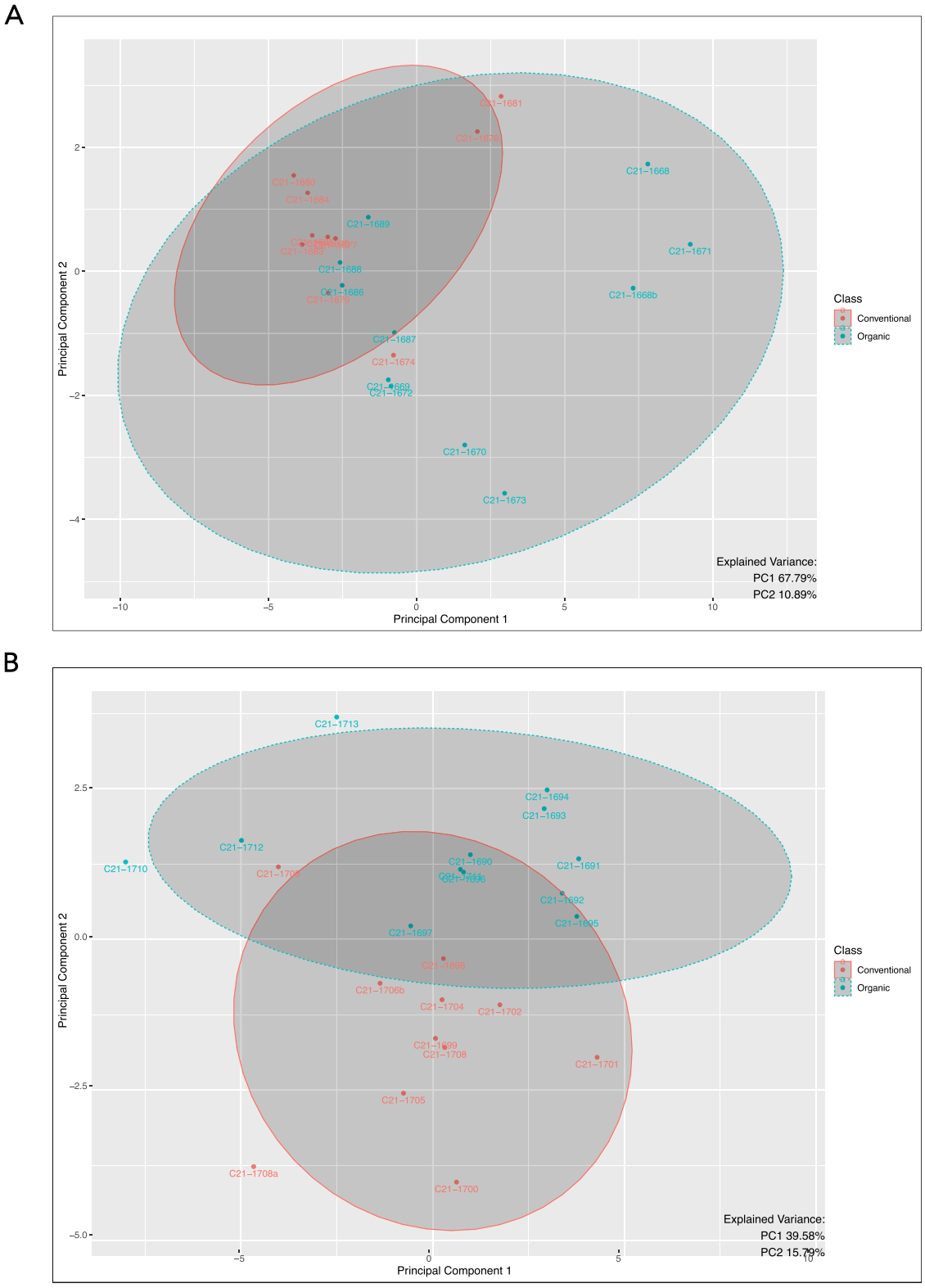

Fig. 1. Principal component analysis scores plots of both years of agronomical trials. A: Year 1; B: Year 2.

organic class, and treatments 3, 4, and 5 under the conventional class. Firstly, an unpaired Welch's t-test was performed with the rstatix package to evaluate significant differences between the averages of both classes in both years of collection. Then, the quantitative data were scaled (unit variance scaling) and a principal component analysis was performed treating each collection year separately to explore preliminary differences between the samples. To find classification patterns between farming systems in both years, orthogonal partial least squares discriminant analysis (OPLS-DA) was performed with the ropls package. An independent validation set comprised the 20 % of the total samples

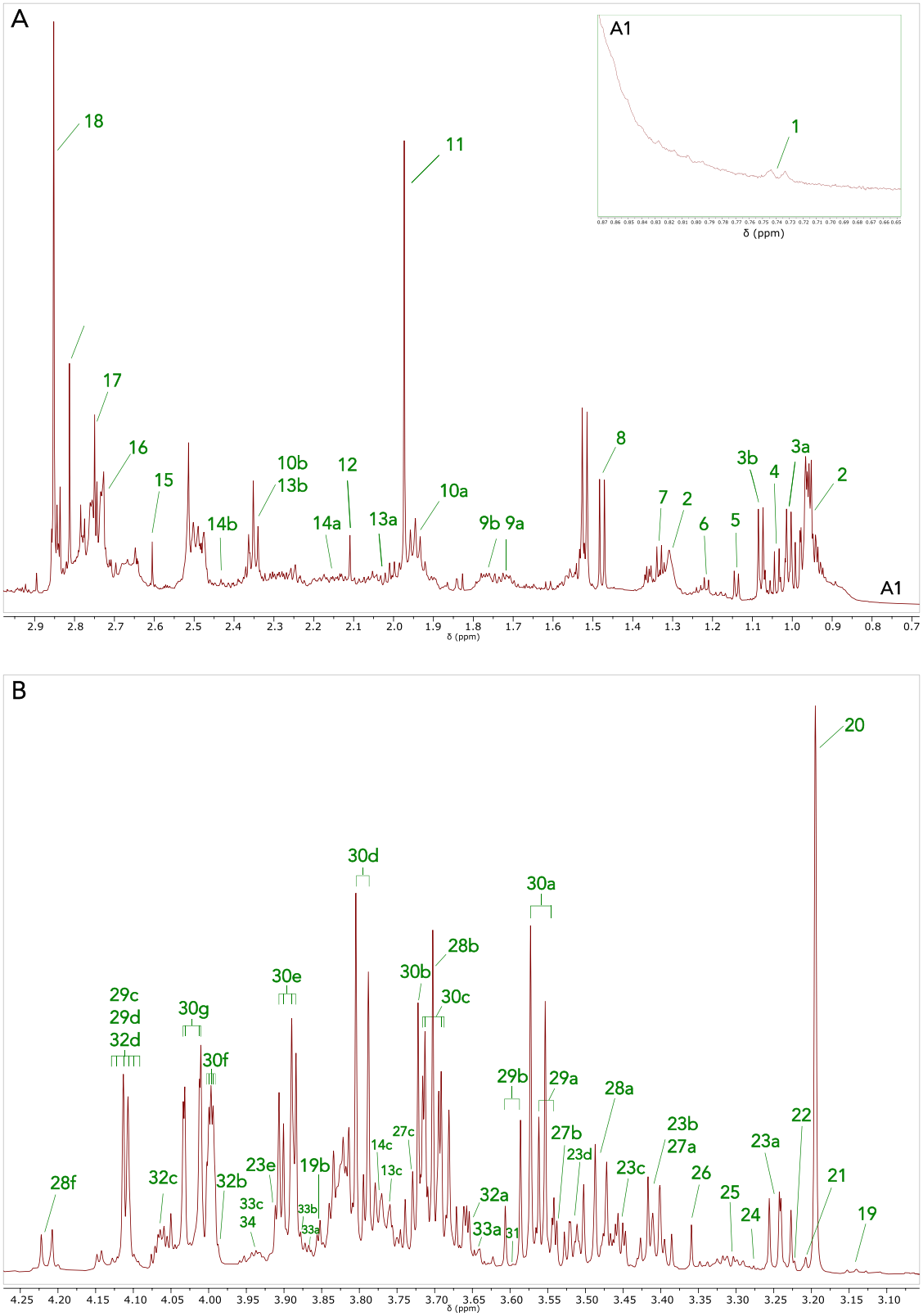

Fig. 2. General ^{1}H NMR spectrum of cauliflower corymbs.

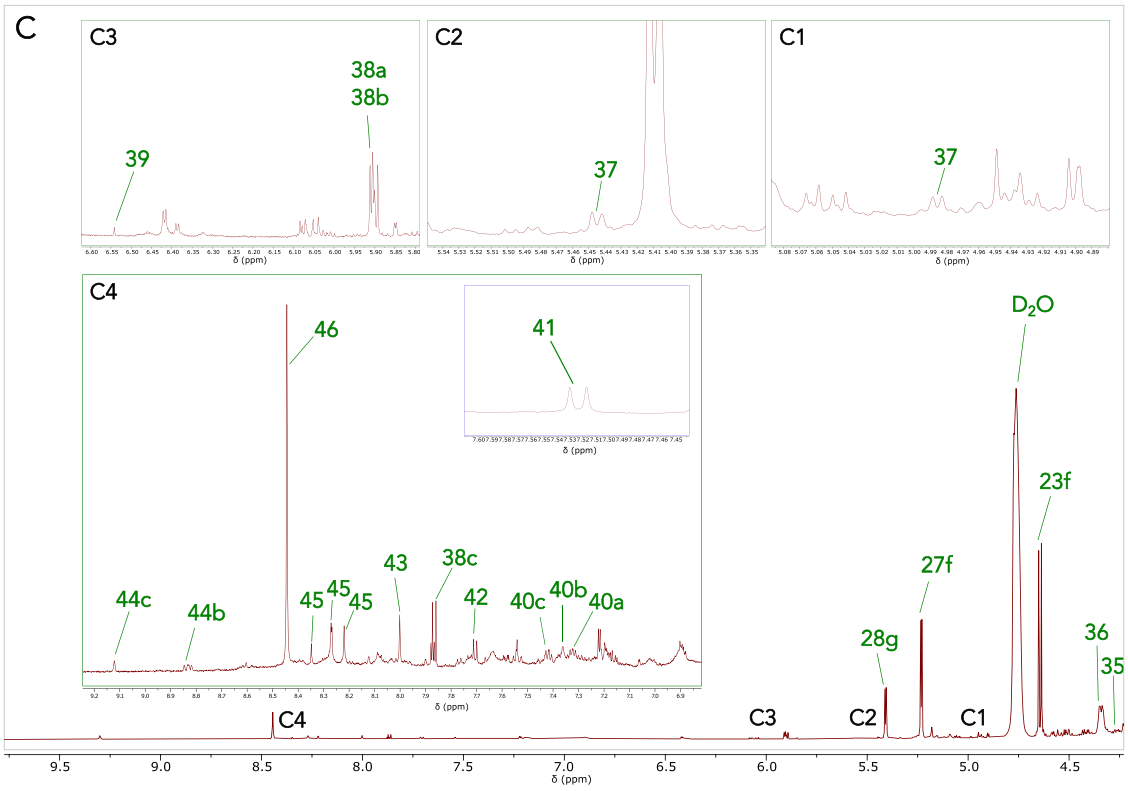

Fig. 2. (*continued*).

and these were not used to build the models. A confusion matrix was obtained for each of the models with the classification of the test set. An observation plot was obtained to search for potential outliers, an inertia plot was obtained with the number of predictive and orthogonal components obtained and the score plot with the separation of samples. The k-fold cross-validation (k=5) of the models gave the parameters R2X (goodness of fit), R2Y (accuracy), and Q2Y (predictability), accordingly reported in the plot. To evaluate overfitting, a permutation test (n=100) was performed by randomly changing the class labels and building new models, the R2Y and Q2Y parameters were measured for the permuted models and the original model.

3. Results and discussion

3.1. Principal component analysis

The scores plot of the PCA for both years are presented in Fig. 1. In both years, no clear clustering was observed among the samples belonging to the six different treatments in each farming system, indicating that larger differences in the external input composition should be considered to further differentiate between different organic, the use of agroecological practices, or conventional treatments in cauliflower corymbs (Figs. 1A and 1B). This is why the samples were grouped considering only the farming system (organic or conventional) for the following quantification and classification analyses. In both years, a major preliminary clustering from of classes was detected, with a larger dispersion in the organic samples cluster and an overlapping with the conventional ones. For the samples from year 1 (Fig. 1A), the total explained variance was 78.68 % between the first two components, while for the second year this was of 55.37 %, indicating larger differences between the samples (hence low total variance was explained between the first two components and more total components were necessary). Possible causes of the higher dispersion of the organic samples could be mainly due to the different treatments, since we previously demonstrated the significant effect of agroecological management in the metabolome of corymbs or to differences in the mineral content of soils or climatic changes in daylight exposure of plants, or storage conditions (Campanelli et al., 2024). Similar results were found in the discrimination between organic and conventional farming systems by NMR in *Brassica oleracea* (var. *italica*) where a clear preliminary separation of both classes with a region of overlapping was observed. In this study, all the conditions were also controlled so the differences only came from the farming systems and the classification analysis allowed the discrimination despite the observed overlap (Lucarini et al., 2020).

3.2. Identification of compounds

The results of the identification of compounds are presented in Fig. 2 and Table 2. Forty-two compounds were individually identified in cauliflower corymbs from both years in both classes. Cursory inspection of the spectra did not reveal major qualitative differences between the spectra related to the number of signals, but clear differences were observed in terms of intensity of the signals that dominated the spectra. Signals associated with a chemical family (sterols, fatty acids, and purines) were labeled with the same number because clear identification was not possible. However, matches with compounds from these families were found in the consulted bibliography (see Materials and Methods section). Out of the identified compounds, carbohydrates and nitrogen compounds (excluding amino acids) were the most abundant families (11 compounds each), followed by amino acids and organic acids (9 and 8 compounds, respectively), then purines (2) and other compounds (1). The carbohydrate family in corymbs was composed of mono-, di-, and oligosaccharides. Monosaccharides primarily included the anomeric forms of glucose and all the tautomeric forms of fructose. As expected, the alpha and keto tautomers were present at lower relative concentrations than the beta forms, reflecting their natural distribution in aqueous solution (Barclay et al., 2011). However, at least one diagnostic signal for each tautomer was observed in corymb samples. Sucrose was the predominant disaccharide in cauliflower corymbs and

Table 2
Identification of metabolites in cauliflower corymbs by ^{1}H NMR.

Number	Compound	Assignment	Multiplicity[a]	Chemical shift (ppm)	Coupling constant, J [Hz]
1	Sterols	CH_3−18	m	0.74	-
2	Fatty acids	ω-CH_3	m	0.92–0.95	-
3a	Valine	γ-CH_3	d	1.00	7.0
4	Isoleucine	γ-CH_3	d	1.01	7.0
3b	Valine	γ'-CH_3	d	1.05*	7.0
5	Isobutyrate	CH_3	d	1.14*	6.3
6	Methylmalonate	α-CH_3	d	1.22*	6.5
2	Fatty acids	Aliphatic CH_2 chain	m	1.29	-
7	Lactic acid	β-CH_3	d	1.33*	6.9
8	Alanine	β-CH_3	d	1.48*	7.2
9a	Leucine	β-CH_2	m	1.73	-
9b	Leucine	γ-CH	m	1.70	-
10a	GABA	β-CH_2	qt	1.91	-[b]
11	Acetic acid	α-CH_3	s	1.93*	-
12	Dimethylsulfide	CH_3	s	2.10*	-
13a	Glutamate	β, β'-CH_2	m	2.13, 2.08	-
14a	Glutamine	β-CH_2	m	2.15	-
13b	Glutamate	γ-CH_2	m	2.35	-
10b	GABA	α-CH_2	t	2.36	7.4
14b	Glutamine	γ-CH_2	m	2.45	-
15	Methylamine	CH_3	s	2.60	-
16	Glucoraphanin	S-CH_3	s	2.72	-
17	Glucoiberin	S-CH_3	s	2.75	-
18	S-Methyl-L-cysteine-S-oxide (Methiin)	γ-CH_3	s	2.85*	-
19	Ethanolamine	CH_2-NH_2	m	3.15*	-
20	Choline	N-$(CH_3)_3$	s	3.19*	-
21	Phosphocholine	N-$(CH_3)_3$	s	3.20*	-
22	Glycerophosphocholine	N-$(CH_3)_3$	s	3.22	-
23a	β-D-Glucose	CH−2	dd	3.24	(9.4, 7.9)
24	Glycine betaine	N-$(CH_3)_3$	m	3.27	-
25	*Myo*-Inositol	CH−4	m	3.29	-
26	*Scyllo*-Inositol	CH−1,2,3,4,5,6	s	3.36*	-
23b	β-D-Glucose	CH−4	m	3.42	-
27a	α-D-Glucose	CH−4	m	3.42	-
23c	β-D-Glucose	CH−5	m	3.46	-
28a	Sucrose	CH-4g[c]	t	3.48	-[b]
23d	β-D-Glucose	CH−3	m	3.50	-
27b	α-D-Glucose	CH−2	m	3.55	-
29a	β-D-Fructofuranose	CH_2−1'	d	3.555	-[b]
30a	β-D-Fructopyranose	CH_2−1'	d	3.56	11.7
31	Glycine	α-CH_2	s	3.59	-
29b	β-D-Fructofuranose	CH_2−1	d	3.60	12.1
32a	α-D-Fructofuranose	CH_2−1'	d	3.64	-[b]
33a	α-D-Fructopyranose	CH_2−1'	d	3.65	-[b]
28b	Sucrose	CH-1f[c]	s	3.68	-
30b	β-D-Fructopyranose	CH−1	d	3.71	-[b]
30c	β-D-Fructopyranose	CH−6'	dd	3.71	-[b]
27c	α-D-Glucose	CH−3	m	3.72	-
13c	Glutamate	α-CH	m	3.77	-
27d	α-D-Glucose	CH_2−6	m	3.77, 3.84	
14c	Glutamine	α-CH	m	3.78	-
30d	β-D-Fructopyranose	CH−3	d	3.80	9.9
28c	Sucrose	CH-6g[c]	m	3.82	-
28d	Sucrose	CH-6f[c]	m	3.82	-
28e	Sucrose	CH-5g[c]	m	3.83	-
27e	α-D-Glucose	CH−5	m	3.84	-
33a	α-D-Fructopyranose	CH−6	m	3.87	-[b]
33b	α-D-Fructopyranose	CH−5	m	3.88	-[b]
30e	β-D-Fructopyranose	CH−4	dd	3.90	(10.1, 3.5)
23e	β-D-Glucose	CH_2−6	m	3.90, 3.75	-
34	keto-D-Fructose	CH−4	m	3.94	-
33c	α-D-Fructopyranose	CH−4	m	3.95	-
30 f	β-D-Fructopyranose	CH−5	m	4.00	-
32b	α-D-Fructofuranose	CH−4	m	4.00	-
30 g	β-D-Fructopyranose	CH−6	dd	4.02	(12.8, 1.4)
33d	α-D-Fructopyranose	CH−3	m	4.03	-
32c	α-D-Fructofuranose	CH−5	m	4.05	-
32d	α-D-Fructofuranose	CH−3	m	4.11	-
29c	β-D-Fructofuranose	CH−3	m	4.12	-
29d	β-D-Fructofuranose	CH−4	m	4.12	-
28 f	Sucrose	CH-3f[c]	d	4.22	8.7
35~~a~~	Inulin	CH-3f[c]	m	4.27	-
36	Malate	α-CH	dd	4.34*	(9.3, 2.9)
23 f	β-D-Glucose	CH−1	d	4.64*	7.9

(*continued on next page*)

Table 2 (*continued*)

Number	Compound	Assignment	Multiplicity[a]	Chemical shift (ppm)	Coupling constant, J [Hz]
37	Raffinose/Stachyose/Raffinose	CH−1Gal[c,,]	m	4.98–5.00	-
27 f	α-D-Glucose	CH−1	d	5.23*	3.7
28 g	Sucrose	CH-1g	d	5.42*	3.8
37	Raffinose/Stachyose/Raffinose	CH	m	5.44	-
38a	Uridine	CH−5	d	5.90*	7.8
38b	Uridine	CH−1′Rib[c]	d	5.91*	4.1
39	Fumaric acid	α, β-HC=CH	s	6.53*	-
40a	Phenylalanine	CH−2,6	m	7.34*	-
40b	Phenylalanine	CH−4	m	7.38*	-
40c	Phenylalanine	CH−3,5	m	7.43*	-
41	Tryptophan[d]	CH−7	d	7.55*	7.7
42	Indoleacetic acid	CH	d	7.72	8.0
38c	Uridine	CH−6	d	7.85*	8.1
43	Xanthine	CH−2	s	7.90	-
44a	Trigonelline	CH−3	m	8.13	-
45	Purines	CH	s	8.21	-
45	Purines	CH	s	8.26	-
45	Purines	CH	s	8.34	-
46	Formate	HCOO-	s	8.47*	-
44b	Trigonelline	CH−4,2	m	8.84	-
44c	Trigonelline	CH−6	s	9.12*	-

[a] s: singlet; d:doublet; t:triplet; q:quartet; qt: quintet; dd:doublet of doublets. [b] The coupling constant could not be calculated due to overlapping of signals or insufficient intensity. *:Signal used for quantification [c]g:glucose; f:fructose; Gal: galactose; Rib:ribose.[d]The signal was identified in other corymb samples, but it is not visible in the representative sample selected for the figure. The position of the signal was marked in Fig. 2C, Panel C4, in a separate box.

oligosaccharides included the raffinose family oligosaccharides (RFOs). These oligosaccharides included raffinose, stachyose, and verbascose with the characteristic signal at 4.98–5.00 ppm belonging to the galactose unit but they could not be individualized because the signal presented a large overlap. Compounds under the category of N-compounds included all those containing N, excluding amino acids. Amines including methylamine, ethanolamine, and choline with its derivatives, and purines with their diagnostic signals between 7.80 and 8.35 ppm from the indole ring system were included. Some purine signals could not be specifically assigned to individual compounds due to slight structural differences among matches that share the same indole skeleton. This was detected by the presence of a common resonance from the proton located between the two nitrogen atoms in the ring. The signals of the sterol family could belong to the stigmasterol or campesterol methyl group, that have been previously identified in cauliflower, but no other clear resonances could confirm this in the spectra (Ingallina et al., 2023). The fatty acids signals encompassed saturated fatty acids. As expected in a water solution, the signals from these family appeared as broad bands given the high structural similarity, sometimes differing only by a CH_2 residue, so the identification was based on the main functional groups with different chemical environments that give them diagnostic chemical shifts. Mono- and polyunsaturated fatty acids were not found in cauliflower corymbs by ^{1}H NMR analysis, this was confirmed by the absence of the resonances of the allylic and olefinic protons at 2.23 and 5.30 ppm, respectively.

3.3. Quantification of compounds

Twenty-five compounds met the criteria previously mentioned in Section 2.4 and were quantified with the external standard in the corymbs of *Brassica oleracea* L. *var. botrytis*. The results of the quantification for the corymbs from both farming systems of both years are presented

Table 3
Quantification of metabolites by ^{1}H NMR (mg/100 g).

Compound/System	Organic Year 1 (n=11)	Conventional Year 1 (n=10)	p-value	Organic Year 2 (n=11)	Conventional Year 2 (n=12)	p-value
Acetate	(5.45±2.62)	(2.87±1.36)	0.0117	(4.49±1.27)	(4.85±0.89)	0.0549
Alanine	(1.35±0.97)	(0.65±0.34)	0.0429	(2.50±1.12)	(1.61±0.38)	0.111
Choline	(6.52±1.13)	(4.65±1.41)	0.00403	(8.19±0.40)	(6.80±0.70)	0.0141
Ethanolamine	(3.59±1.18)	(2.44±1.13)	0.0358	(6.54±0.56)	(6.37±0.37)	0.0526
Formate	(1.13±0.22)	(0.96±0.13)	0.0535	(0.80±0.08)	(0.80±0.12)	0.109
Fumarate	(0.08±0.04)	(0.11±0.05)	0.286	(0.14±0.05)	(0.13±0.08)	0.916
α-D-Glucose	(15.50±9.13)	(23.62±8.61)	0.0497	(10.38±6.74)	(8.91±3.11)	0.762
β-D-Glucose	(24.86±10.51)	(37.81±4.25)	0.00233	(20.24±8.54)	(16.59±3.65)	0.457
Glycine	(3.29±1.29)	(3.13±1.15)	0.778	(3.23±0.43)	(2.71±0.36)	0.121
Isobutyrate	(0.27±0.09)	(0.14±0.06)	0.00168	(0.59±0.26)	(0.36±0.13)	0.00772
Lactate	(1.00±0.25)	(0.78±0.16)	0.0324	(1.47±0.17)	(1.44±0.18)	0.188
Malate	(36.66±7.44)	(27.37±10.23)	0.0312	(40.81±2.89)	(38.43±3.29)	0.4
Methiin	(3.48±1.17)	(2.92±0.61)	0.106	(4.04±0.47)	(2.73±0.37)	0.0000075
Methylamine	(0.08±0.03)	(0.04±0.01)	0.0289	(0.08±0.04)	(0.05±0.01)	0.4
Methylmalonate	(1.14±0.44)	(0.81±0.25)	0.0486	(1.64±0.27)	(1.29±0.18)	0.0265
Phenylalanine	(1.43±1.30)	(0.85±0.77)	0.286	(2.72±0.55)	(2.15±0.48)	0.0975
Phosphorylcholine	(0.61±0.14)	(0.45±0.11)	0.0109	(0.57±0.17)	(0.62±0.09)	0.0927
Scyllo-Inositol	(0.006±0.001)	(0.005±0.001)	0.0629	(0.007±0.001)	(0.006±0.002)	0.987
Dimethylsulfide	(1.68±0.58)	(1.33±0.44)	0.131	(2.33±0.30)	(2.45±0.29)	0.389
Sucrose	(22.03±9.50)	(31.54±12.62)	0.0419	(5.05±2.85)	(7.32±4.31)	0.141
Trigonelline	(0.16±0.04)	(0.13±0.02)	0.109	(0.19±0.02)	(0.17±0.02)	0.24
Tryptophan	(2.00±1.24)	(1.17±0.62)	0.0706	(2.80±0.49)	(2.55±0.28)	0.91
Uridine	(1.49±0.53)	(1.13±0.46)	0.113	(3.79±0.57)	(3.82±0.85)	0.235
Valine	(1.05±0.57)	(1.80±0.54)	0.313	(1.95±0.28)	(1.85±0.16)	0.67
Xanthine	(0.36±0.19)	(0.21±0.06)	0.0338	(0.25±0.15)	(0.48±0.17)	0.000694

in Table 3 with the p-value from the Welch's t-test indicating significant differences between the means. Eleven compounds showed significantly higher concentrations in organic samples from year 1, namely, acetate, alanine, choline, ethanolamine, isobutyrate, lactate, malate, methylamine, methylmalonate, phosphorylcholine, and xanthine. Methiin, isobutyrate, and choline were in significantly higher concentrations in organic samples from year 2. On the other hand, conventional samples showed higher concentrations of sugars including both anomers of glucose and sucrose in year 1, while xanthine was in higher concentration in year 2.

Direct comparison with previous studies is a very complex task in fruit and vegetable crops. Accurate comparisons require the same variety, cultivar, geographical location, growing media, external inputs, climate, and farming system. A previous study found that the differences in the chemical composition between organic and conventional cauliflower did not only depend on the growing management practices but also on the genotypic characteristics (e.g.: different time of ripening) (Picchi et al., 2012). In another study it was demonstrated that the levels of antioxidant metabolites and phytochemical compounds like glucosinolates were affected not only by the farming system but also by the environmental conditions and the genotype (Lo Scalzo et al., 2013). No previous studies carried out the ^{1}H NMR characterization of *Brassica oleracea* L. var. *botrytis* but some similar trends were found in other cultivars. A previous study performed the ^{1}H NMR metabolomics classification in the farming systems of *Brassica oleracea* L. var. *italica*. The results showed a higher concentration of malate and choline in organic samples, in agreement with our results (Lucarini et al., 2020). Malate also plays an essential role in plant development, as it is a substrate for bacteroid respiration in the N-fixing roots and is the source of carbon for the synthesis of amino acids with the N that plants fix. Previous research also shows that phosphorous deficiency during plant development increases the levels of malate, as it is also an important metabolite for the incorporation of phosphorous by roots. In the case of the organic treatments of our study, the animal pellet composition was enriched in N, as well as the vegetable amendment in the agroecological practices, thus showing such low-phosphorous conditions that could increase levels of malate (Schulze et al., 2002). In our previous work, significant differences in the levels of ascorbate were found, with higher concentrations in the organic samples. Now, more organic acids showed the same behavior, particularly acetate, isobutyrate, lactate, malate, and methylmalonate. Alanine and N-containing compounds with important biological roles in nutrition like choline and phosphorylcholine accompanied the trend, showing higher concentrations in the organic samples in the first year, except for xanthine that was higher in conventional samples in both years. We previously demonstrated that samples in an organic farming system showed higher values of $\delta(^{15}N)$ in corymbs, with the animal pellet used to fertilize the organic system plots being the highest in the $\delta(^{15}N)$ (‰ compared to air) (Campanelli et al., 2024). However, we observed then that the origin of the N source did not have a significant difference in total and inorganic N content between treatments and farming systems, so it was the fractionation processes of N in the plants that caused this result. Now, the content of N-containing molecules in organic samples supported our previous results, giving organic samples an improved nutritional profile. Carbohydrates (both anomers of glucose and sucrose) showed a different behavior, with higher concentrations in conventional cauliflower corymbs in the first year. This result supports the findings of our previous study where samples from conventional farming system were higher in total soluble solids compared with organic and organic + agroecological practices treatments. It is known that N from synthetic fertilizers in conventional farming systems is more readily available for plants since it is 100 % in its mineral form, but in the case of organic fertilizers, the availability depends on the mineralization and immobilization turnover (Luo et al., 2021). At the same time, the use of pesticides in conventional farming systems reduce pest stress. These factors promote a faster growth, allowing the plants to allocate more energy towards the production of more sugars. Plants grown under organic farming systems, however, are subjected to a lower N release from natural sources and pest stress that activate defense mechanisms and the production of other secondary metabolites. Higher sugar content was previously reported in conventionally grown fruits and vegetables like strawberries, apples, carrots, green and red bell peppers, and cabbages. However, the sugar content is always related to the maturity stage, the cultivar, and particular treatment (Rahman et al., 2021).

Finally, a note regarding pH values and classification. The differentiation of samples by statistical methods was not affected by slight chemical shifts due to pH values since we worked with the concentrations of the compounds obtained from the aligned spectra in this approach (widely targeted). Untargeted approaches suffer more from this issue if alignment cannot completely solve the shift variation in complex mixtures, leading to the same compound being present in different bins due to pH shifts.

3.4. Classification of farming systems

OPLS-DA models were built to find metabolites that could discriminate between farming systems in *Brassica oleracea* L. *var. botrytis* for two consecutive years. The results of the models are presented in Fig. 3. One predictive and two orthogonal components were used for the models, with a higher inertia explained by the first predictive component in year 2 (Fig. 3B). The classification rate for the independent test set was of 100 % for both models (See Supplementary Table 1) with a root mean square error of estimation of 0.199 and 0.131 for years 1 and 2, respectively. Both models did not show overfitting as shown by the Q2Y parameter (0.738 and 0.824, respectively) and the permutation test plots. Some metabolites were common to both years of the agronomical trails, while others were found only for each year showing a clear impact of the farming system. It is always important to take into consideration the possible contribution of the above-mentioned factors (soil composition, temperature, daylight exposure) in the metabolism of plants even though they were carefully controlled in our previous study. Also, the use of cover crops in the plot of the agroecological practices to promote soil health may have an impact on the soil composition given the variable N fixating capacity between those crops and cauliflower (Campanelli et al., 2024). The complete list of selected metabolites for each year, with their VIP values, is presented in Supplementary Table 2. Choline, isobutyrate, formate, and alanine were common variables chosen for both years of trials. These metabolites had higher contribution to the separation of the organic samples in the model, along with phosphorylcholine, malate, and lactate, for year 1 exclusively. The detection of farming practices in *Brassica oleracea* L. *var botrytis* could be based on the presence and concentration of this set of molecules, that were also selected regardless of the year of the trial. This demonstrated their resistance to the variety of external factors previously mentioned that affect the metabolism of plants but also to internal factors regarding the intrinsic metabolic cycles of the plants and their genotype. Both anomers of glucose were selected as discriminant metabolites of the conventional corymbs for year 1. The presence of sugars is ubiquitous in fruits and vegetables, making their potential as indicators of a farming system very limited. However, for year 2, a wider variety of metabolites was selected, including xanthine, sucrose, and phenylalanine for conventional samples. This allowed the proposal of these metabolites as potential markers of the farming systems in future studies in this cultivar and reaffirmed the ubiquitous character of sugars as indicators of conventional farming practices in this cultivar. For organic samples of year 2, methiin (S-methyl cysteine sulfoxide), methylamine, and methylmalonate were selected exclusively. Again, N containing metabolites and organic acids were preferred for the organic farming systems, following the trends previously found for the first year.

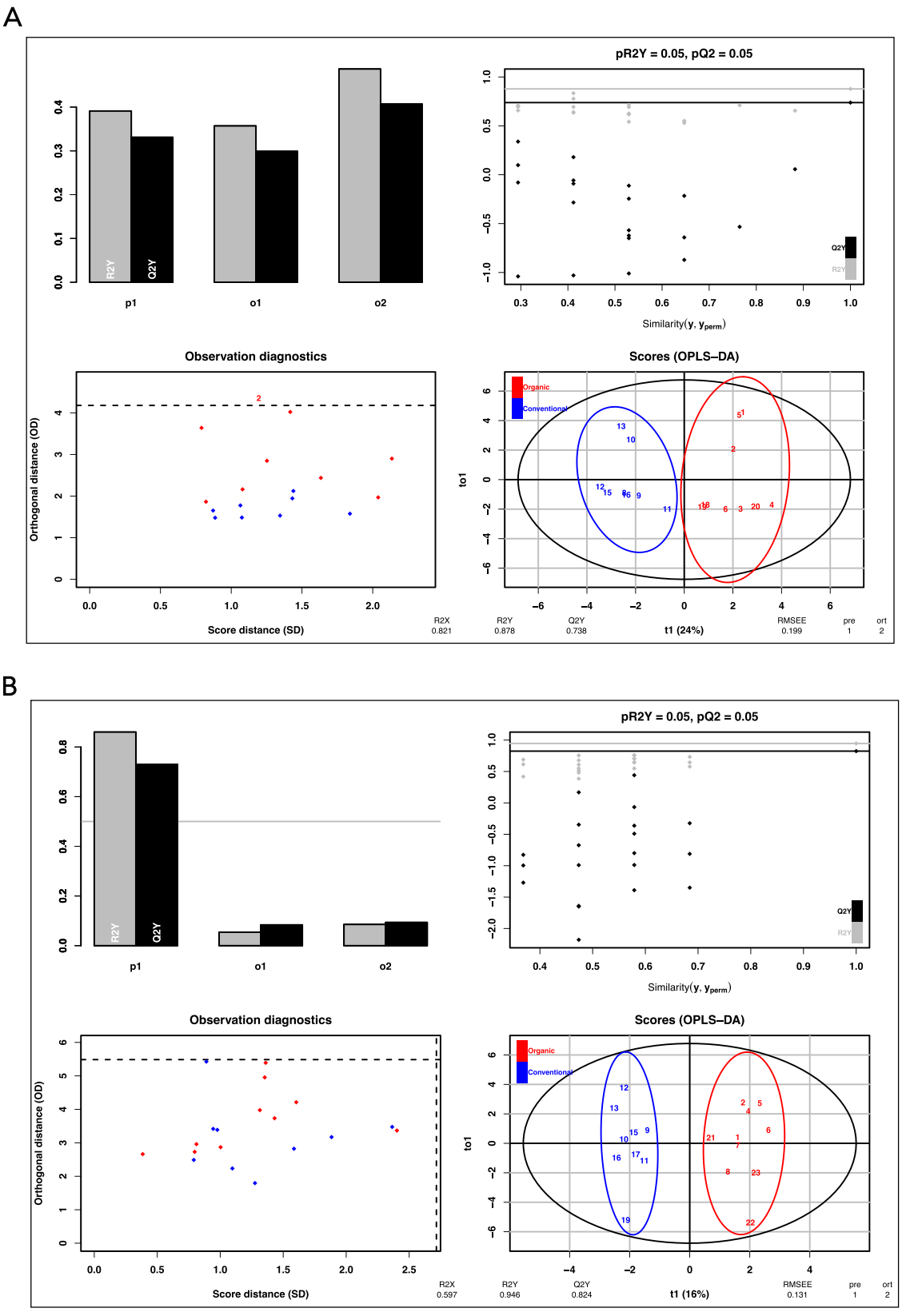

Fig. 3. OPLS-DA classification of farming systems of cauliflower corymbs of both years of agronomical trials. A: Year 1; B: Year 2. Top-right: inertia plot; Top-left: permutation test; Bottom-left: observation diagnostics; Bottom-right: scores plot.

4. Conclusions

NMR widely targeted metabolomics allowed the characterization and classification of *Brassica oleracea* L. *var. botrytis* corymbs in a fast and efficient way, making this approach very suitable for the analysis of cultivation systems in food samples. Organic cauliflower corymbs of this cultivar represented a more nutritious option for consumers now in terms of amino acids, organic acids, and N-containing metabolites, expanding the findings of our previous study. The application of multivariate statistics allowed, with excellent performance measures, the classification of farming systems of this species regardless of the year of the trial and the mixture of external inputs in the different treatments. This was achieved using 9 and 11 discriminant molecules for years 1 and 2, respectively. Particularly, 4 metabolites, namely choline, isobutyrate, formate, and alanine were resistant to the external and internal variations in plants and were proposed as potential indicators of the farming practices in cauliflower. The quantification results in both years indicated a significant impact of the cultivation year in the nutritional profile of cauliflower corymbs. The obtained results also consider the different combinations of external inputs for both farming systems, making them of potential application in future studies in the area. The main challenge in these studies remains the control of the conditions of the soil environment to allow better comparisons with previous studies and the validation of the proposed discriminant metabolites including the analysis of subsequent years of trials.

Funding

The work was supported by the Ministero dell'Università e della Ricerca [PE00000003 ON FOODS]; Provincia Autonoma di Trento [FRUITOMICS, FESR 2014–2020]; the Italian Ministry of Agriculture, Food Sovereignty and Forestry, with grant number 93173/2017, under the "Application of innovative methods for the traceability of organic farming products"—INNOVABIO project.

CRediT authorship contribution statement

Simona Fabroni: Writing – review & editing, Investigation, Funding acquisition, Conceptualization. **Pavel Solovyev:** Writing – review & editing, Methodology, Formal analysis, Data curation, Conceptualization. **Federico I. Brigante:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. **Luana Bontempo:** Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization. **Francesco Montemurro:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Gabriele Campanelli:** Writing – review & editing, Methodology, Investigation, Conceptualization.

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

Acknowledgements

The authors acknowledge the support of "ON Foods - Research and innovation network on food and nutrition Sustainability, Safety and Security - Working ON Foods" to Fondazione Edmund Mach, funded by the Italian Ministry of University and Research, PNRR, Mission 4 Component 2 Investment 1.3 - Project code PE00000003, CUP B43C22000770006. This study has also received funding from the Autonomous Province of Trento, with EU co-financing (FRUITOMICS, FESR 2014–2020 Program of the Autonomous Province of Trento, Italy) and from the Italian Ministry of Agriculture, Food Sovereignty and Forestry, with grant number 93173/2017, under the "Application of innovative methods for the traceability of organic farming products"—INNOVABIO project.

Disclaimer

This study was carried out within the ON Foods – Research and innovation network on food and nutrition Sustainability, Safety and Security – Working ON Foods" project and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – Mission 4 Component 2 Investment 1.3 – D.D. 1550 11/10/2022, PE0000003, CUP B43C22000770006). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jfca.2024.107018.

Data Availability

Data will be made available on request.

References

Akkerman, R., Farahani, P., Grunow, M., 2010. Quality, safety and sustainability in food distribution: a review of quantitative operations management approaches and challenges. OR Spectr. 32, 863–904. https://doi.org/10.1007/s00291-010-0223-2.

Akoka, S., Barantin, L., Trierweiler, M., 1999. Concentration measurement by proton NMR using the ERETIC method. Anal. Chem. 71 (13), 2554–2557. https://doi.org/10.1021/ac981422i.

Barclay, T., Ginic-Markovic, M., Johnston, M.R., Cooper, P., Petrovsky, N., 2011. Observation of the keto tautomer of d-fructose in D_2O using ^1H NMR spectroscopy. Carbohydr. Res. 347 (1), 136–141. https://doi.org/10.1016/j.carres.2011.11.003.

Briones Alonso, E., Cockx, L., Swinnen, J., 2018. Culture and food security. Glob. Food Secur. 17, 113–127. https://doi.org/10.1016/j.gfs.2018.02.002.

Campanelli, G., Amenta, M., Bontempo, L., Leteo, F., Montemurro, F., Platani, C., Timpanaro, N., Torrisi, B., Fabroni, S., 2024. Innovative tools for nitrogen fertilization traceability in organic farming products: a cauliflower case study. Horticulturae 10, 94. https://doi.org/10.3390/horticulturae10010094.

Colson, K.L., Hicks, J.M., & Fischer, C. (2012). Method and apparatus for automated raw material screening (USPTO Patent No. 8248072). In US Patent (No. 8248072). https://patentimages.storage.googleapis.com/c1/26/d9/65cbb13655277d/US8248072.pdf.

Consonni, R., Cagliani, L.R., 2022. Quality assessment of traditional food by NMR analysis. Food Control 142, 109226. https://doi.org/10.1016/j.foodcont.2022.109226.

Cullen, C.H., Ray, G.J., Szabo, C.M., 2013. A comparison of quantitative nuclear magnetic resonance methods: internal, external, and electronic referencing. Magn. Reson. Chem.: MRC 51 (11), 705–713. https://doi.org/10.1002/mrc.4004.

Garibaldi, L.A., Gemmill-Herren, B., D'Annolfo, R., Graeub, B.E., Cunningham, S.A., Breeze, T.D., 2017. Farming approaches for greater biodiversity, livelihoods, and food security. Trends Ecol. Evol. 32 (1), 68–80. https://doi.org/10.1016/j.tree.2016.10.001.

Hird, V., 2017. Farming systems and techniques that promote biodiversity. Biodiversity 18 (2–3), 71–74. https://doi.org/10.1080/14888386.2017.1351395.

Hoch, J.C., Baskaran, K., Burr, H., Chin, J., Eghbalnia, H.R., Fujiwara, T., Gryk, M.R., Iwata, T., Kojima, C., Kurisu, G., et al., 2023. Biological magnetic resonance data bank. Nucleic Acids Res 51, D368–D376. https://doi.org/10.1093/nar/gkac1050.

Hong, R.S., Hwang, K.H., Kim, S., Cho, H.E., Lee, H.J., Hong, J.T., Moon, D.C., 2013. Survey of ERETIC2 NMR for quantification. J. Korean Magn. Reson. Soc. 17 (2), 98–104. https://doi.org/10.6564/JKMRS.2013.17.2.098.

Ingallina, C., Di Matteo, G., Spano, M., Acciaro, E., Campiglia, E., Mannina, L., Sobolev, A.P., 2023. Byproducts of globe artichoke and cauliflower production as a new source of bioactive compounds in the green economy perspective: an NMR study. Molecules 28, 1363. https://doi.org/10.3390/molecules28031363.

Li, Z., Lee, H.W., Liang, X., Liang, D., Wang, Q., Huang, D., Ong, C.N., 2018. Profiling of phenolic compounds and antioxidant activity of 12 cruciferous vegetables. Molecules 23, 1139. https://doi.org/10.3390/molecules23051139.

Li, S., Tian, Y., Jiang, P., Lin, Y., Liu, X., Yang, H., 2020. Recent advances in the application of metabolomics for food safety control and food quality analyses. Crit. Rev. Food Sci. Nutr. 61 (9), 1448–1469. https://doi.org/10.1080/10408398.2020.1761287.

Lucarini, M., Di Cocco, M.E., Raguso, V., Milanetti, F., Durazzo, A., Lombardi-Boccia, G., Santini, A., Delfini, M., Sciubba, F., 2020. NMR-based metabolomic comparison of *Brassica oleracea* (Var. *italica*): organic and conventional farming. Foods 9, 945. https://doi.org/10.3390/foods9070945.

Luo, H., Robles-Aguilar, A.A., Sigurnjak, I., Michels, E., Meers, E., 2021. Assessing nitrogen availability in biobased fertilizers: effect of vegetation on mineralization patterns. Agriculture 11, 870. https://doi.org/10.3390/agriculture11090870.

Powers, R., Andersson, E.R., Bayless, A.L., Brua, R.B., Chang, M.C., Cheng, L.L., Clendinen, C.S., Cochran, D., Copié, V., Cort, J.R., Crook, A.A., Eghbalnia, H.R., Giacalone, A., Gouveia, G.J., Hoch, J.C., Jeppesen, M.J., Maroli, A.S., Merritt, M.E., Pathmasiri, W., Wishart, D.S., 2024. Best practices in NMR metabolomics: current state. TrAC Trends Anal. Chem. 171, 117478. https://doi.org/10.1016/j.trac.2023.117478.

R Core Team (2023) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. https://www.R-project.org/.

Rahman, S.M.E., Mele, M.A., Lee, Y.T., Islam, M.Z., 2021. Consumer preference, quality, and safety of organic and conventional fresh fruits, vegetables, and cereals. Foods 10 (1), 105. https://doi.org/10.3390/foods10010105.

Ramos-Figueroa, J.S., Tse, T.J., Shen, J., Purdy, S.K., Kim, J.K., Kim, Y.J., Han, B.K., Hong, J.Y., Shim, Y.Y., Reaney, M.J.T., 2023. Foaming with starch: Exploring faba bean aquafaba as a green alternative. Foods 12 (3391). https://doi.org/10.3390/foods12183391.

Schulze, J., Tesfaye, M., Litjens, R.H.M.G., Bucciarelli, G., Trepp, G., Miller, S., Samac, D., Allan, D., Vance, C.P., 2002. Malate plays a central role in plant nutrition. Plant Soil 247, 133–139. https://doi.org/10.1023/A:1021171417525.

Spraul, M., Schütz, B., Rinke, P., Koswig, S., Humpfer, E., Schäfer, H., Mörtter, M., Fang, F., Marx, U., Minoja, A., 2009. NMR-based multi parametric quality control of fruit juices: SGF profiling. Nutrients 1 (2), 148–155. https://doi.org/10.3390/nu1020148.

Qi, S., Zeng, T., Wu, P., Sun, L., Dong, Z., Xu, L., Xiao, P., 2024. Widely targeted metabolomic analysis reveals effects of yellowing process time on the flavor of vine tea (*Ampelopsis grossedentata*). Food Chem.: X 22, 101446. https://doi.org/10.1016/j.fochx.2024.101446.

Wang, H., Hua, J., Yu, Q., Li, J., Wang, J., Deng, Y., Yuan, H., Jiang, Y., 2021. Widely targeted metabolomic analysis reveals dynamic changes in non-volatile and volatile metabolites during green tea processing. Food Chem. 363, 130131. https://doi.org/10.1016/j.foodchem.2022.132982.

Watanabe, R., Sugai, C., Yamazaki, T., Matsushima, R., Uchida, H., Matsumiya, M., Takatsu, A., Suzuki, T., 2016. Quantitative nuclear magnetic resonance spectroscopy based on PULCON methodology: application to quantification of invaluable marine toxin, okadaic acid. Toxins 8 (10). https://doi.org/10.3390/toxins8100294.

Yang, T., Zhao, B., & He, L. (2019). Raman instruments for food quality evaluation, In J. Zhong, & X. Wang (Eds.) Evaluation Technologies for Food Quality. Woodhead Publishing Series in Food Science, Technology and Nutrition (pp. 119-143). Woodhead Publishing. https://doi.org/10.1016/b978-0-12-814217-2.00008-1.

Yuliana, N.D., Hunaefi, D., Goto, M., Ishikawa, Y.T., Verpoorte, R., 2021. Measuring the health effects of food by metabolomics. Crit. Rev. Food Sci. Nutr. 62 (23), 6359–6373. https://doi.org/10.1080/10408398.2021.1901256.