Journal of Food Composition and Analysis 123 (2023) 105532

Contents lists available at ScienceDirect

Journal of Food Composition and Analysis

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

Qualitative analysis of broccoli (Brassica oleracea var. italica) glucosinolates: Investigating the use of mid-infrared spectroscopy combined with chemometrics

Faye Langston [a,*,1], Ali Ali Redha [b,1], Geoffrey R. Nash [a], John R. Bows [d], Luciana Torquati [b], Michael J. Gidley [c], Daniel Cozzolino [c]

[a] Natural Sciences, Faculty of Environment, Science and Economy, University of Exeter, Exeter EX4 4QF, United Kingdom
[b] The Department of Public Health and Sport Sciences, University of Exeter Medical School, Faculty of Health and Life Sciences, University of Exeter, Exeter EX1 2 LU, United Kingdom
[c] Centre for Nutrition and Food Sciences, Queensland Alliance for Agriculture and Food Innovation (QAAFI), The University of Queensland, Brisbane, QLD 4072, Australia
[d] PepsiCo R&D, Leicester LE4 1ET, United Kingdom

ARTICLE INFO

Keywords:
Broccoli
Mid infrared
Chemometrics
Glucosinolates
Glucobrassicin

ABSTRACT

Glucosinolates are phytochemicals with important health and nutritional benefits. This study reports the use of high-performance liquid chromatography (HPLC) and mid-infrared (MIR) spectroscopy to characterise and differentiate between broccoli varieties and systems of production (organic vs. non-organic) depending on their glucosinolate content and infrared fingerprint. Broccoli samples (n = 53) from seven varieties were analysed using MIR spectroscopy and HPLC. Differences in the MIR spectra of the individual broccoli varieties were observed in the carbohydrate fingerprint region (950–1100 cm^{-1}) and between 1340 and 1615 cm^{-1} assigned to specific glucosinolates. Principal component analysis (PCA) of the MIR fingerprint spectra enabled the differentiation between samples with relatively high (200–500 mg/100 g DW) and low (< 200 mg/100 g DW) glucobrassicin content. Linear discriminant analysis (LDA) and PCA-LDA were used to classify broccoli varieties according to the system of production (organic vs. non-organic) and variety (common vs. Tenderstem® broccoli). The classification rates indicated that > 70 % of the samples were correctly classified as organic and non-organic, while > 90 % of the samples were correctly classified as common broccoli and Tenderstem®. This study demonstrates that MIR spectroscopy could be used as a potential tool to classify and monitor broccoli samples according to their variety and system of production.

1. Introduction

Glucosinolates are nitrogen sulphur-containing compounds found almost exclusively in the *Brassicaceae* family of vegetables. They are characterised by a thiohydroximate-*O*-sulphate group conjugated to a β-D-thioglucose with an alkyl, aralkyl or indole R-side chain (Fig. 1a). Over the past few decades, many studies have reported the potential health benefits of glucosinolate bioactive metabolites (isothiocyanates) in chemoprevention (Dinkova-Kostova and Kostov, 2012; Mitsiogianni et al., 2019), neuroprotection (Jaafaru et al., 2018), and metabolic syndrome treatment (Esteve, 2020). However, these compounds are also responsible for the astringent taste and sulphurous aroma of some of the *Brassica* vegetables commonly consumed (e.g., broccoli, cabbage, and kale). Sinigrin and progoitrin are aliphatic glucosinolates, abundant in cabbage (e.g., Napa cabbage) and brown mustard, that are especially noted for their pungent bitter taste which has been linked to consumer rejection (Van Doorn, 1999). Prolonged or high ingestion of glucosinolates present in animal feeds has also been shown to cause adverse side effects in livestock, particularly in non-ruminant animals (e.g. pigs) (Tripathi and Mishra, 2007).

Notably, the progoitrin hydrolysis product, goitrin, has been shown to restrict iodine metabolism in the thyroid gland inducing an iodine deficiency, which in turn causes issues with thyroid function, and the potential onset of goiter (Tripathi and Mishra, 2007; Wallig et al., 2002).

* Corresponding author.
E-mail address: fmal201@exeter.ac.uk (F. Langston).
[1] Authors will equal contribution.

https://doi.org/10.1016/j.jfca.2023.105532
Received 26 April 2023; Received in revised form 10 July 2023; Accepted 11 July 2023
Available online 12 July 2023

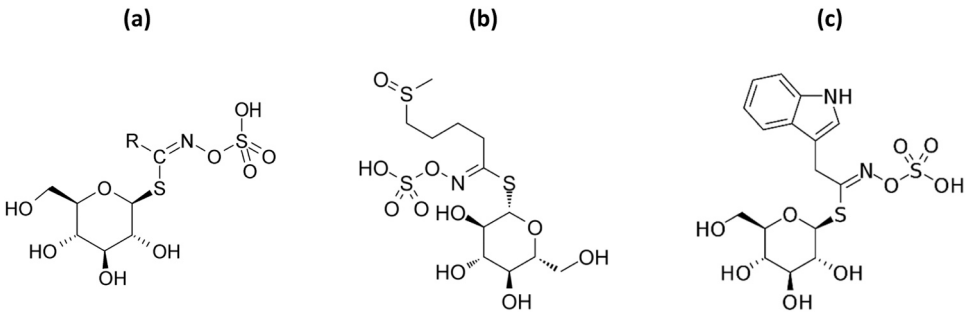

Fig. 1. Chemical structure of (a) general glucosinolate, (b) glucoraphanin, and (c) glucobrassicin.

These deleterious attributes have been shown to be either compound- or dose-dependent, as well as species-dependent (Tripathi and As, 2017) where no serious adverse effects have been reported in humans. To either optimise their medicinal benefit and consumer palatability, or to reduce possible adverse effects on livestock, understanding the glucosinolate content and profiles of *Brassica* vegetables is critical for both the food manufacturing industry and consumers.

Broccoli (*Brassica oleracea*) is one of the most consumed *Brassica* vegetables (Zhong et al., 2015). It is often referred to as a '*superfood*' as it is a good source of minerals (potassium, phosphorous, calcium and sodium), protein and fibre (Campas-Baypoli et al., 2009), and it is particularly rich in the glucosinolates, glucoraphanin and glucobrassicin (Fig. 1b and c).

Several analytical techniques have been developed and utilised to both identify and quantify glucosinolates in plant tissues (e.g., leaf and flower). These techniques included capillary electrophoresis, gas chromatography (GC), and enzyme-linked immunosorbent assays (ELISA) (Śmiechowska et al., 2010). By far the most used method is high-performance liquid chromatography (HPLC) coupled with ultraviolet (UV) or diode-array detection (DAD) (Almushayti et al., 2021), as these are robust and well-validated. However, this technique is time-consuming, costly, and requires complex pre-processing steps such as the utilization of a sulfatase pre-treatment to prepare the sample for effective separation during reversed-phase chromatography. Liquid chromatography in tandem with mass spectrometry (LC-MS/MS) has become an effective method for glucosinolate characterisation and quantification (Śmiechowska et al., 2010). This technique provides high selectivity and sensitivity, whilst reducing sample preparation and analysis time in comparison to HPLC. Nevertheless, LC-MS/MS still requires various levels of pre-processing of the sample prior to analysis, such as extraction, the use of solvents, columns, and labour input which can be a great drawback when handling large sample sets. Therefore, high throughput, cost-effective and rapid quantification methods of glucosinolates are needed.

An attractive method to address the short-falls of HPLC and LC-MS/MS is infrared (IR) spectroscopy, which has been explored in recent years to analyse the glucosinolate composition of vegetables such as broccoli (Chen et al., 2014; Renner and Fritz, 2020; Sahamishirazi et al., 2017; Toledo-Martin et al., 2017), kale (Chen et al., 2014), cabbage and Brussels sprouts (Renner and Fritz, 2020). A recent systematic review highlighted that among the current IR spectroscopy methods used to evaluate glucosinolates, near-infrared (NIR) spectroscopy is the most widely used (Ali Redha et al., 2023). However, most of the NIR applications were not efficient in predicting the amount of different glucosinolates in broccoli and it was suggested that mid-infrared (MIR) spectroscopy could be considered as an alternative method (Ali Redha et al., 2023).

MIR spectroscopy measures the characteristic absorption bands relating to fundamental vibrations between bonds of different molecular groups in a sample (Ozaki, 2021; Porep et al., 2015). Specifically, it contains the so-called 'fingerprint region' between 400 and 1500 cm^{-1} – a distinctive region in the IR spectrum that contains the main information about the main molecules which are characteristic of a sample (Ozaki, 2021; Porep et al., 2015). The sharper peaks and fingerprint region of the MIR spectra is more favourable for identify specific molecules within complex samples such as broccoli (Ozaki, 2021).

This study reports the use of both HPLC and MIR spectroscopy combined with chemometrics to characterise and differentiate between broccoli varieties (common vs. Tenderstem® broccoli) and the system of production (organic vs. non-organic) depending on their glucosinolate content and infrared fingerprint.

2. Materials and methods

2.1. Samples

A total of 51 fresh broccoli samples were purchased from local markets (Central England, UK), over the course of a year (2022) in different seasons: winter - February (n = 20), summer – June (n = 18), and autumn – October (n = 13). The samples comprised common broccoli (n = 20), organic broccoli (n = 9), common Tenderstem® broccoli (n = 11), organic Tenderstem® broccoli (n = 4), purple sprouting broccoli (n = 4), and Amaranthine broccoli (n = 3) that were grown in different parts of the UK, Spain, Morocco, and other countries. Two additional samples from Beneforte® broccoli were sourced from the John Innes Centre (Norwich, UK) bringing the total number of samples analysed in this study to 53. Both Tenderstem® and Beneforte® are trademark varieties that have been selectively bred for their milder and sweeter taste (Tenderstem®) (Martínez-Hernández et al., 2011) or high glucoraphanin content (Beneforte ®) (Mithen, 2013). The most commonly available and consumed broccoli varieties in the UK were selected for this study.

2.2. Sample preparation

Broccoli heads were removed from their stalks and cut into < 50 mm florets. The florets were then freeze-dried (Labconco™ FreeZone™ 4.5 L −50 °C Benchtop Freeze Dryer, USA) at − 50 °C and 0.054 bar for 96 h until a moisture content of approximately 4–6 %. The samples were then turned into a fine powder (US Standard 40 mesh) using a commercial grinder (Shardor CG-619, Bad Pyrmont, Germany). Powdered samples were stored at − 18 °C in air-tight plastic bags until analysis.

2.3. Determination of glucosinolates by HPLC

The glucosinolates (glucoraphanin, glucoiberin, glucobrassicin, neoglucobrassicin, 4-hydroxyglucobrassicin, and 4-methoxyglucobrassicin) were measured using the method proposed by Saha et al. (2012). This method converts the glucosinolates to the equivalent desulfoglucosinolates. Briefly, broccoli powder samples (40–50 mg) were extracted with a hot aqueous methanol mixture (70 % v/v, 10 mL) following the addition of 50 μL of the internal standard (sinigrin). Samples were mixed by vortexing and incubated at 70 °C for 20–30 min

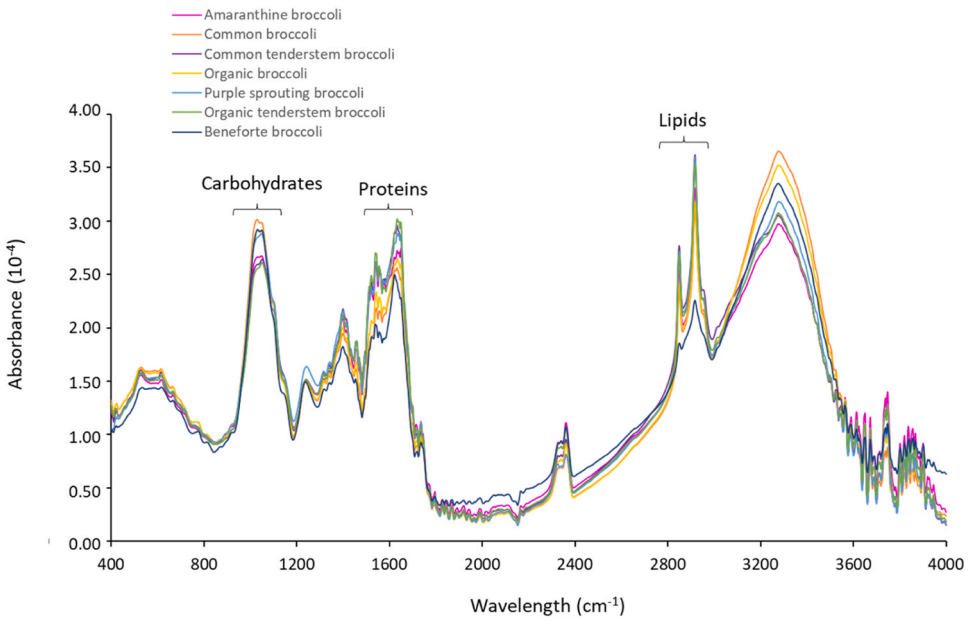

Fig. 2. Averaged spectra of different types of broccoli samples analysed using attenuated total reflectance mid-infrared spectroscopy after normalisation and smoothing.

with occasional mixing. Extracts were allowed to cool at room temperature where a sample (3 mL) of the supernatant was applied to an ion exchange column that was subsequently washed first with water (2 × 0.5 mL) and then with 0.02 M sodium acetate (2 × 0.5 mL). The columns were then layered with purified sulfatase (75 μL) and incubated at room temperature and pressure overnight for desulfo-glucosinolates conversion. The desulfo-glucosinolates were eluted by sequential application of 0.5 and 0.25 mL water and analysed using HPLC. This consisted of a Waters Spherisorb ODS2 (250 mm × 4.6 mm i.d., 5 μm particle size) column connected to a model 1100 HPLC system (Agilent Technologies, Waldbronn, Germany) comprised of a binary pump, degasser, cooled autosampler, column oven, and diode array detector. Samples were eluted at 1.0 mL/min with a gradient of increasing acetonitrile using 0.1 % v/v formic acid containing water (solvent A) and 0.1 % v/v formic acid containing acetonitrile (solvent B). The gradient started at 0 % solution B increasing over 25 min to 50 % B and finally re-equilibrated to 0 % B for 7 min. The quantification of glucosinolates was done by comparing the internal standard (sinigrin) peak after the correction factor applied for each of the glucosinolates measured (absorbance at 229 nm).

2.4. MIR spectroscopy

The MIR spectra of the ground and freeze-dried broccoli samples were collected in triplicate using a laminated diamond attenuated total reflectance (ATR) cell attached to a Thermo Fisher™ Nicolet™ iS™5 FTIR instrument (Thermo Fisher Scientific, Massachusetts, USA). Approximately 30 mg of the sample was placed on top of the ATR crystal where the MIR (400–4000 cm^{-1}) spectrum was recorded using OMNIC™ software version 9.6 (Thermo Fisher Scientific, Massachusetts, USA). Each spectrum was computed using an average of 32 interferograms at a resolution of 8 cm^{-1}. Air, at atmospheric pressure, was used as the background, and the ATR crystal was wiped with an alcohol swab (70 % isopropyl, Klipspringer, Ipswich, UK) between each sample.

2.5. Data analysis

The MIR spectra were analysed using the VEKTOR DIREKTOR™ v1.1 software (KAX Group, Sydney, NSW, Australia). Prior to chemometric analysis, the MIR spectra of all samples (n = 53) were pre-processed using baseline correction and second derivative (second polynomial order, and 21 smoothing points) (Savitzky and Golay, 1964). Principal component analysis (PCA) was performed to visualise the data structure and identify patterns between different varieties of broccoli using MIR spectral data (800–1400 cm^{-1}) and the HPLC data (Cozzolino et al., 2019). Linear discriminant analysis (LDA) was used to classify the common broccoli samples according to the system of production to organic (n = 9) and non-organic (n = 20), and variety, i.e. common broccoli (n = 29) and Tenderstem® broccoli (n = 15). In order to overcome the constraint of requiring more samples than variables, PCA-LDA was used and compared with LDA. In this study, PCA-LDA was used to reduce the data dimensionality by developing a PCA (three principal components explaining 99 % of the variability) prior to running the LDA. In this way, the number of components used is less than the number of objects in each class. The LDA and PCA-LDA models were developed using the MIR spectral data between 800 and 1400 cm^{-1} with cross-validation (leave one out).

The glucosinolate composition of the broccoli samples was statically analysed by one-way analysis of variance (one-way ANOVA) and compared using Tukey's range test ($p < 0.05$) using IBM SPSS Statistics v28.0.0.0 software (SPSS Inc., Chicago, United States).

3. Results and discussion

3.1. MIR spectra interpretation

The average ATR-FTIR spectra after baseline correction and smoothing for the seven broccoli varieties analysed are shown in Fig. 2. Common features characteristic of the broccoli samples can be observed in the MIR region related to carbohydrates, lipids, and proteins. High peaks around 1099 cm^{-1} and 2916 cm^{-1} and the broad bands between 3255 and 3450 cm^{-1} have been attributed to aldehyde and ketone C-O stretching, alkyl C-H stretching, and alcohol O-H stretching of structural and non-structural carbohydrates, respectively (Jackson and Mantsch, 1996). The intense peaks typical of the carbohydrate fingerprint region (Hong et al., 2021), resulting from differences in glycosidic linkages, can be observed between 920 and 1185 cm^{-1}. Peaks between 1500 and 1700 cm^{-1} are related to the amide I and II groups where the peak at 1635 cm^{-1} and peaks around 1646 and 1652 cm^{-1} correspond to the N-H deformation (amide I groups) and C═O stretching of amide I,

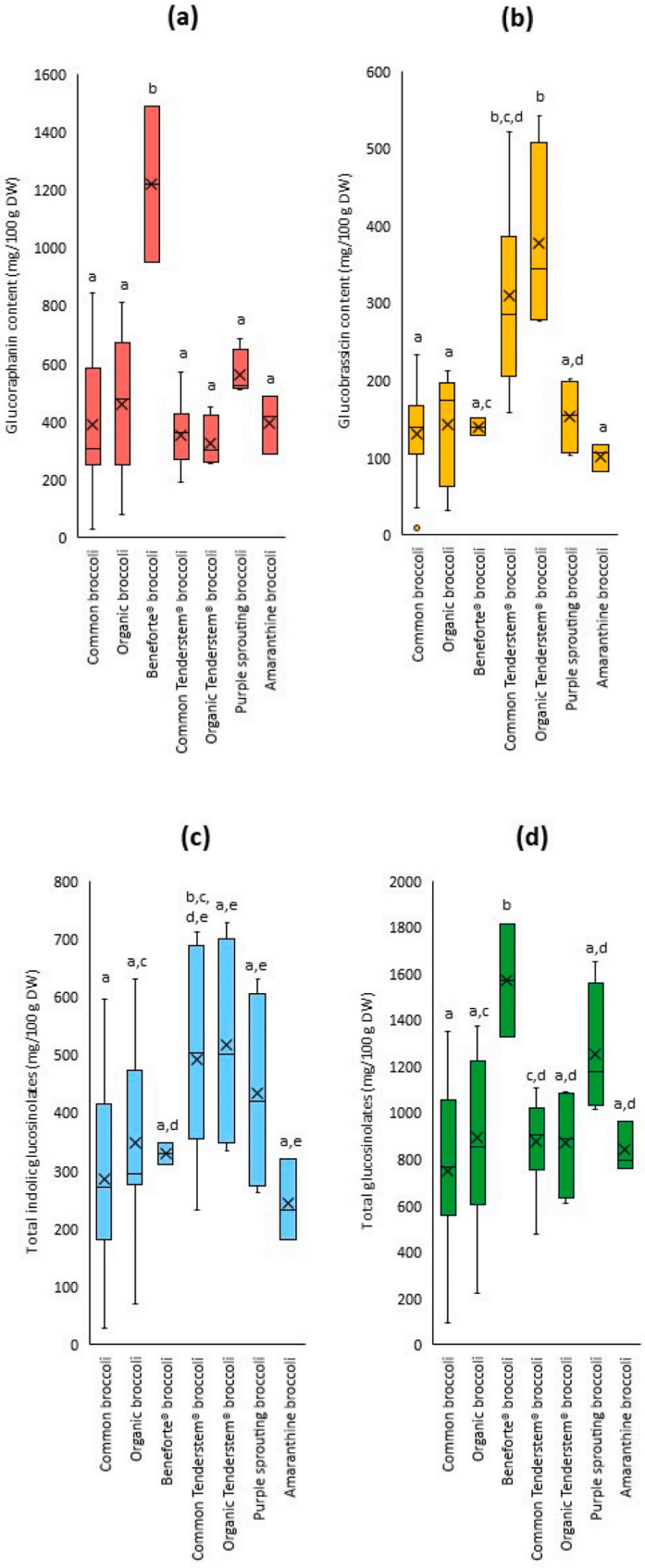

Fig. 3. Amount of (a) glucoraphanin, (b) glucobrassicin, (c) total indolic glucosinolates, and (d) total glucosinolates of different types of broccoli samples in mg/100 g dry weight (DW). Different lowercase letters in each plot mean statistical difference at $p < 0.05$.

respectively (Turker-Kaya and Huck, 2017). The peak around 1538 cm^{-1} is assigned to the N-H bending and C-N stretching of amide II (Turker-Kaya and Huck, 2017). Furthermore, the CH_2 symmetric and asymmetric stretching derived from lipids can be observed by the distinct peaks around 2847 and 2916 cm^{-1}. Peaks at 1099 and 1241 cm^{-1} were related to the symmetrical and asymmetrical stretching of the C-O-C ester bonds where the peaks at 1733 and 1749 cm^{-1} correspond to the C═O stretch of ester bonds (Jackson and Mantsch, 1996).

Vo and collaborators have reported the characteristic peaks in the MIR spectra of individual glucosinolates (Vo et al., 2013, 2014). These authors have reported specific bands associated with glucoraphanin at 3316, 2976, 2868, 1651, 1495, 1265, 1063 cm^{-1} and at 3378, 1600, 1236, 1041 cm^{-1} associated with glucobrassicin (Vo et al., 2013, 2014). It is well known that glucosinolates contain a rare C═N-sulphate moiety, which result in strong IR bands between 1630 and 1690 cm^{-1} that superimpose with the amide I group (Larkin, 2018). In addition, the β-D-thioglucose has characteristic peaks around 1000–1200 cm^{-1} (C-O), 2900 cm^{-1} (C-H), and 3200–3500 cm^{-1} (O-H) (Ma et al., 2018). The main peak observed around 3200 cm^{-1} can be associated with the broad absorption of compounds containing alcohols, phenols and carboxylic groups and might be associated with the indole group (Ma et al., 2018). This peak is also reported to be associated with the N-H of the amide group (Larkin, 2018). Distinct differences between the MIR spectra of aliphatic glucosinolates, which have an open hydrocarbon chain, compared to indole glucosinolates which have a closed chain structure in the form of an indole ring have been also observed (Coates, 2007). In particular, indole glucosinolates have been reported to be located around 1350–1650 cm^{-1} (C-H vibrations) (Coates, 2007).

3.2. Glucosinolate composition

The broccoli samples were analysed for their glucoraphanin, glucobrassicin, total indolic glucosinolates (sum of glucobrassicin, neoglucobrassicin, 4-hydroxyglucobrassicin, and 4-methoxyglucobrassicin content), and total glucosinolate content (sum of glucoraphanin, glucoiberin, and total indolic glucosinolate content) using HPLC-UV (Fig. 3). The data for minor glucosinolates (neoglucobrassicin, 4-hydroxyglucobrassicin, 4-methoxyglucobrassicin, and glucoiberin) is not shown.

In terms of glucoraphanin content (Fig. 3a), most samples had a glucoraphanin content between 200 and 600 mg/100 g dried weight (DW). The highest glucoraphanin content was observed in the Beneforte® broccoli samples (between 950 and 1490 mg/100 g DW) with an average of 1221 ± 380 mg/100 g DW and statistically significantly greater than all the other samples ($p < 0.05$) analysed. The Beneforte® broccoli has been developed through genome introgression from *Brassica villosa* to produce a high content of glucoraphanin in comparison with other Brassica cultivars (Traka et al., 2013). The purple sprouting broccoli samples also had a high glucoraphanin content in the range of 510–690 mg/100 g DW. A wide variation in the glucoraphanin content in common broccoli (± 240 mg/100 g DW) and its organic variety (± 260 mg/100 g DW) was also observed. The lowest glucoraphanin content in common broccoli was 30 mg/100 g DW, whilst the highest was 850 mg/100 g DW. However, the range in composition was relatively narrower in the Tenderstem® broccoli samples where in most of the samples the content of glucoraphanin was between 260 and 540 mg/100 g DW. A similar trend was observed in the total glucosinolate content of the broccoli samples analysed (Fig. 3d). This was expected since glucoraphanin accounts for a considerable fraction of the total glucosinolate content of broccoli. The highest amount of glucosinolates was observed in the Beneforte® broccoli (1300–1800 mg/100 g DW) which was significantly higher compared to the other varieties ($p < 0.05$).

In terms of glucobrassicin content, remarkable differences between common broccoli and Tenderstem® broccoli were observed (Fig. 3b). Glucobrassicin content of both common and organic broccoli ranged between 30 and 260 mg/100 g DW, while Tenderstem® broccoli (including organic and non-organic) ranged between 160 and 540 mg/100 g DW. Organic Tenderstem® broccoli had significantly higher glucobrassicin content compared to all the other varieties ($p < 0.05$), yet the difference in comparison to common Tenderstem® broccoli was not statistically significant ($p > 0.05$). The Beneforte® broccoli samples had a glucobrassicin content in the range of the common broccoli (130–150 mg/100 g DW) samples, while having a rich glucoraphanin content. The total indolic glucosinolates content of both common and Tenderstem® broccoli samples (including organic and non-organic) also followed the same trend as observed with glucobrassicin content. This was expected since glucobrassicin is the main indolic glucosinolate in broccoli and accounts for the highest glucosinolate fraction in the samples analysed.

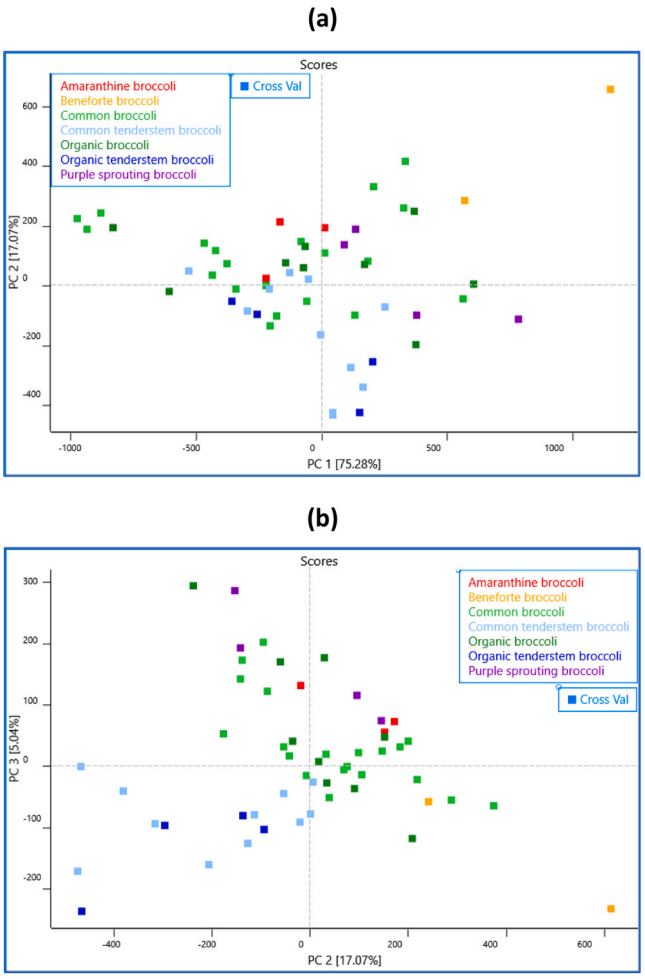

Fig. 4. Principal component (PC) analysis score plots of glucosinolate composition of different types of broccoli varieties analysed using HPLC-DAD: (a) PC1 vs. PC2, and (b) PC2 vs. PC3.

Tenderstem® broccoli is a hybrid between conventional broccoli (*Brassica oleracea* var. *italica*) and kai-lan, a Chinese broccoli (*Brassica oleracea* var. *Alboglabra)*. It is also referred to as broccolini or the tradename Bimi®. The few studies that have characterised the glucosinolate content of this kai-lan hybrid broccoli (Baenas et al., 2019; Martínez-Hernández et al., 2013; Nieto et al., 2023), reported glucobrassicin content between 200 and 480 mg/100 g DW - a similar trend to this study. Significantly lower quantities (30–100 mg/100 g DW) of this compound were also reported in the mature Tenderstem® broccoli samples grown in a controlled glasshouse environment (Langston et al., 2022).

No statistical differences between the glucosinolate content of the

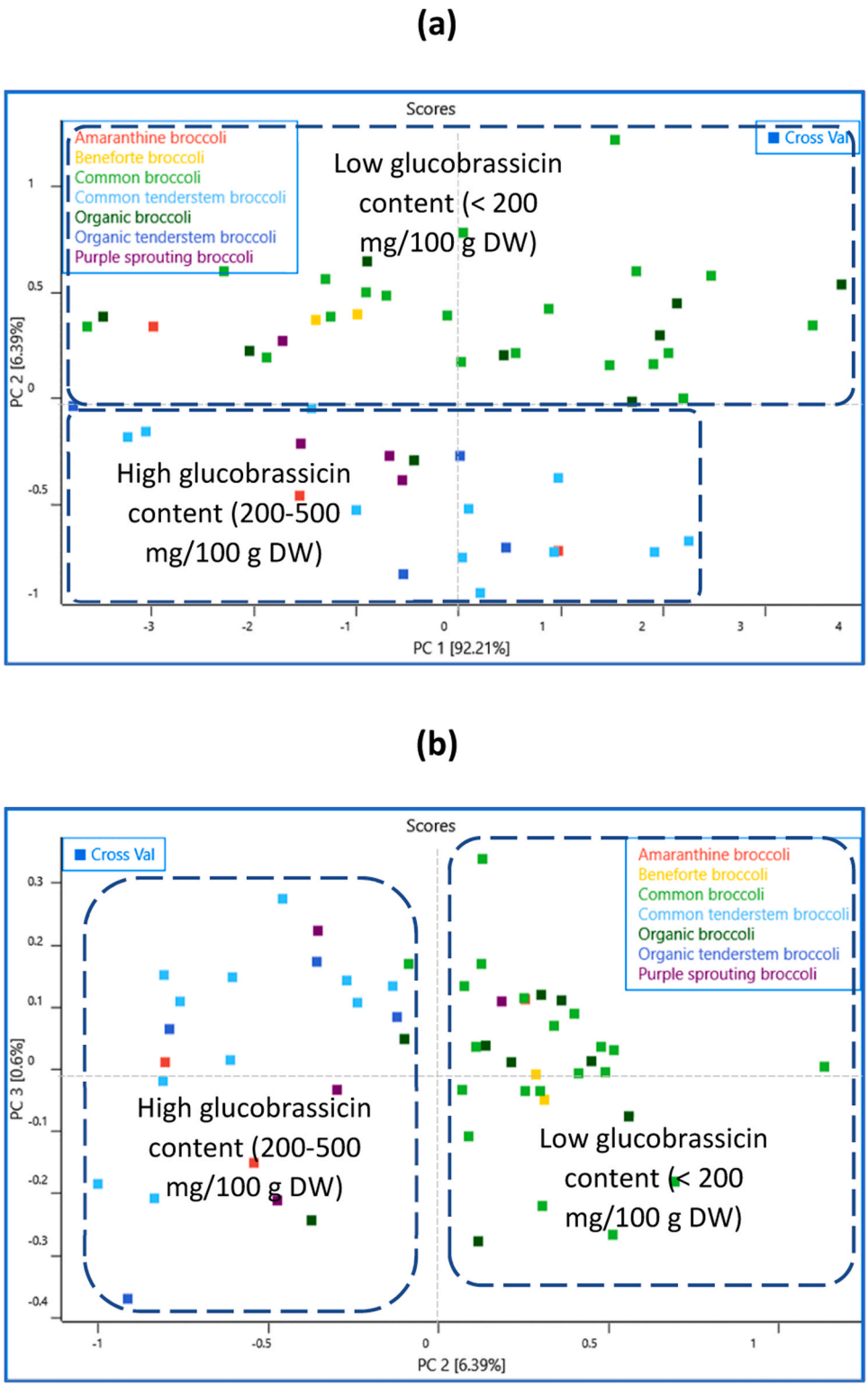

Fig. 5. Score plots of principal component (PC) analysis of different types of broccoli samples analysed using pre-attenuated total reflectance mid-infrared spectroscopy: (a) PC1 vs. PC2, and (b) PC2 vs. PC3.

same variety of common and organic broccoli were observed ($p > 0.05$). This agrees with a previous study that reported glucoraphanin content of ecologically (organic) grown broccoli was not significantly different in comparison to conventionally grown broccoli ($p > 0.01$) (Meyer and Adam, 2007). Nevertheless, the glucobrassicin content of their ecologically grown samples (5.2 mmol/kg) was significantly greater than the conventionally grown broccoli (3.7 mmol/kg) ($p < 0.01$). In another study, the concentrations of glucoraphanin, glucobrassicin and neoglucobrassicin were comparable between the conventional and organic broccoli samples (Renaud et al., 2014).

3.3. Principal component analysis (PCA) of the HPLC data

The HPLC data of the broccoli samples was analysed using PCA

(Fig. 4a and b). The first three principal components (PCs) explained 75.3 %, 17.1 % and 5.0 % of the variation in the HPLC data, respectively. The first plot displaying PC1 vs. PC2 (Fig. 4a), showed limited clustering, with most of the broccoli samples overlapping. However, when PC2 vs. PC3 are plotted (Fig. 4b) both varieties of the Tenderstem® appeared to be grouped in the third quadrant, whilst the remaining samples clustered centrally, with a slight skew to the second and fourth quadrant. Overall, the PCA analysis of the HPLC data was able to differentiate between the Tenderstem® broccoli from the other broccoli samples.

3.4. Principal component analysis (PCA) of the MIR data

The MIR spectra of the broccoli samples were also analysed using PCA. In this way, PCA was used to visualise the differences or trends between the different varieties of broccoli samples using the MIR region between 800 and 1400 cm^{-1}. This MIR region has been reported to contain the key bands associated with glucosinolates (Vo et al., 2013, 2018). The first three PCs explained more than 99 % of the variance in the dataset. PC1 accounts for 92.2 % of the variability while PC2 and PC3 accounted for 6.4 % and 0.6 % of the variability in the MIR spectra of broccoli samples, respectively.

As seen in Fig. 5a (PC1 vs. PC2), most of the common and organic broccoli samples were grouped in the first and second quadrants, including the Beneforte® broccoli samples. On the other hand, most of the common Tenderstem® and organic Tenderstem® broccoli samples were grouped in the third and fourth quadrants, including the amaranthine broccoli samples. In Fig. 5b (PC2 vs. PC3), the common and organic broccoli samples were grouped in the first and fourth quadrants, including the Beneforte® broccoli samples. While most of the common Tenderstem® and organic Tenderstem® broccoli samples were grouped in the second and third quadrants, including the amaranthine broccoli samples. The non-organic and organic samples of the same variety of broccoli tend to overlap. This can be explained by the lack of differences in the composition between organic and non-organic broccoli samples as discussed in the previous sections. In comparison with the PCA developed using the HPLC data, the PCA developed using the MIR spectra provided a better separation between the different varieties of broccoli. This can be explained due to the fact that the MIR spectra contain more information (e.g., sugars, proteins and other compounds containing aromatic groups) about the chemical composition of the samples beyond the glucosinolate content.

The loadings of the first three PCs were analysed to assess the key regions of the MIR spectra that contribute to explaining the differences between the samples analysed (data not shown). The highest loadings in both the first and second PCs were observed in the carbohydrate fingerprint region between 950 and 1100 cm^{-1}. Previous studies have shown that the total sugar content, as well as the ratio between sugars (e.g. glucose, fructose, and sucrose) might be influenced by the variety of broccoli (Rosa et al., 2001). Absorption bands in this region might be associated with the C-C and C-O bonds of the sugars (Cadet et al., 1997). The third PC shows a complex array of peaks between 1340 and 1615 cm^{-1}, which might be related to differences in the glucosinolate profile of the broccoli varieties. The aromatic glucosinolates (e.g. glucobrassicin) show a stronger contribution, with two sets of intense peaks located between 1475 and 1505 and 1560–1615 cm^{-1} denoting the C═C stretch of the aromatic ring from the indole (Mossoba et al., 1989). Peaks between 1385 and 1425 cm^{-1} are associated with the S═O stretch. Weaker peaks characteristic of aliphatic glucosinolates can be located between 1150 and 1210 cm^{-1} and 1340–1420 cm^{-1} for the symmetric and asymmetric stretch of S═O, where peaks between 1410 and 1480 cm^{-1} might be associated with the C-H bend and CH_2/CH_3 stretching (Coates, 2007).

Table 1
Linear discriminant analysis (LDA) and principal component analysis-LDA (PCA-LDA) confusion matrix for the classification of broccoli samples according to variety and the production system.

LDA			Overall classification (%)
Varietal classification	Common broccoli	Tenderstem® broccoli	
Common broccoli	39	0	
Tenderstem® broccoli	2	12	96
System of production classification	Organic	Non-organic	
Organic	9	4	
Non-organic	6	34	79
PCA-LDA			
Varietal classification	Common broccoli	Tenderstem® broccoli	
Common broccoli	38	2	
Tenderstem® broccoli	3	10	90
System of production classification	Organic	Non-organic	
Organic	7	10	
Non-organic	5	31	71

Note: Common broccoli variety represents a low glucobrassicin content variety, while Tenderstem® broccoli represents a high glucobrassicin content variety.

3.5. Linear discriminant analysis (LDA) and PCA-LDA

The LDA confusion matrix for the classification of variety (common vs. Tenderstem® broccoli) and system of production (organic vs. non-organic) is shown in Table 1. The classification rates indicated that 79 % of the samples were correctly classified as organic and non-organic while 96 % of the samples were correctly classified as common broccoli and Tenderstem® broccoli. The PCA-LDA results are also shown in Table 1. In this case, the classification rates indicated that 71 % of the samples were correctly classified as organic, while 90 % of the samples were correctly classified as common broccoli and different from Tenderstem® broccoli. It has been reported that pre-harvest factors such as growing environment (e.g., temperature and soil properties), harvest season, agricultural practices as well as maturity can modulate the content of glucosinolates in broccoli (Ilahy et al., 2020). The results showed that the MIR spectra have information that can be used to either classify samples according to variety or system of production. This information can be utilised to better understand trends in composition as well as assist the food manufacturing industry in the selection of varieties.

4. Conclusions

The results of this study demonstrated the ability of MIR spectroscopy to differentiate between different varieties of broccoli and the system of production. The PCA scores revealed that the broccoli samples can be differentiated based on their glucobrassicin content, as well as based on the MIR spectra as supported by the analysis of loadings (e.g., aromatic groups). Furthermore, LDA and PCA-LDA classification models showed that MIR spectroscopy can be used to differentiate common broccoli from Tenderstem® as well as broccoli from organic and non-organic production. Future work would investigate the capability of MIR spectroscopy to quantify individual glucosinolates in broccoli. With the bioactive effects of glucosinolates being pertinent in both animal agriculture and medicinal applications, these qualitative models provide the food manufacturing industry with the ability to distinguish and select between broccoli varieties and systems of production.

CRediT authorship contribution statement

Faye Langston: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original

draft, Writing – review & editing. **Ali Ali Redha:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Geoffrey R. Nash:** Conceptualization, Writing – review & editing, Supervision. **John R. Bows:** Resources, Writing – review & editing. **Luciana Torquati:** Writing – review & editing, Supervision. **Michael J. Gidley:** Writing – review & editing, Supervision. **Daniel Cozzolino:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Declaration of Competing Interest

Author Faye Langston Ph.D. is fully sponsored by PepsiCo, Inc. Author John R. Bows is an employee of PepsiCo, Inc. All other authors declare no conflicts of interest.

Data Availability

No data was used for the research described in the article.

Acknowledgements

The authors would also like to acknowledge the funding made available from PepsiCo, Inc. which supported this work. The views expressed in this report are those of the authors and do not necessarily represent the position of the policy of PepsiCo, Inc.

Financial support for this study was also provided by the QUEX Institute, a partnership between The University of Queensland and the University of Exeter.

References

Ali Redha, A., Torquati, L., Langston, F., Nash, G.R., Gidley, M.J., Cozzolino, D., 2023. Determination of glucosinolates and isothiocyanates in glucosinolate-rich vegetables and oilseeds using infrared spectroscopy: a systematic review. Crit. Rev. Food Sci. Nutr. 1–17.

Almushayti, A.Y., Brandt, K., Carroll, M.A., Scotter, M.J., 2021. Current analytical methods for determination of glucosinolates in vegetables and human tissues. J. Chromatogr. A 1643, 462060.

Baenas, N., Marhuenda, J., Garcia-Viguera, C., Zafrilla, P., Moreno, D.A., 2019. Influence of cooking methods on glucosinolates and isothiocyanates content in novel cruciferous foods. Foods 8 (7), 257.

Cadet, F., Robert, C.O., Bernard, 1997. Simultaneous determination of sugars by multivariate analysis applied to mid-infrared spectra of biological samples. Appl. Spectrosc. 51 (3), 369–375.

Campas-Baypoli, O.N., Sanchez-Machado, D.I., Bueno-Solano, C., Nunez-Gastelum, J.A., Reyes-Moreno, C., Lopez-Cervantes, J., 2009. Biochemical composition and physicochemical properties of broccoli flours. Int J. Food Sci. Nutr. 60 (Suppl 4), 163–173.

Chen, J., Li, L., Wang, S., Tao, X., Wang, Y., Sun, A., He, H., 2014. Assessment of glucosinolates in Chinese kale by near-infrared spectroscopy. Int. J. Food Prop. 17 (8), 1668–1679.

Coates, J., 2007. Interpretation of Infrared Spectra, a Practical Approach, Encyclopedia of Analytical Chemistry: Applications, Theory, and Instrumentation. John Wiley & Sons.

Cozzolino, D., Power, A., Chapman, J., 2019. Interpreting and reporting principal component analysis in food science analysis and beyond. Food Anal. Methods 12 (11), 2469–2473.

Dinkova-Kostova, A.T., Kostov, R.V., 2012. Glucosinolates and isothiocyanates in health and disease. Trends Mol. Med. 18 (6), 337–347.

Van Doorn, J.E., 1999. Development of Vegetables with Improved Consumer Quality: A Case Study in Brussels Sprouts, Ph.D. dissertation. Wageningen University & Research, Wageningen, The Netherlands.

Esteve, M., 2020. Mechanisms underlying biological effects of cruciferous glucosinolate-derived isothiocyanates/indoles: a focus on metabolic syndrome. Front Nutr. 7, 111.

Hong, T., Yin, J.Y., Nie, S.P., Xie, M.Y., 2021. Applications of infrared spectroscopy in polysaccharide structural analysis: Progress, challenge and perspective. Food Chem. X 12, 100168.

Ilahy, R., Tlili, I., Pek, Z., Montefusco, A., Siddiqui, M.W., Homa, F., Hdider, C., R'Him, T., Lajos, H., Lenucci, M.S., 2020. Pre- and post-harvest factors affecting glucosinolate content in broccoli. Front. Nutr. 7, 147.

Jaafaru, M.S., Karim, Abd, Enas, N.A., Rollin, M.E., Mazzon, P., Abdull Razis, A.F, E., 2018. Protective effect of glucosinolates hydrolytic products in neurodegenerative diseases (NDDs). Nutrients 10 (5).

Jackson, M., Mantsch, H., 1996. Biomedical infrared spectroscopy. In: Mantsch, H.H., Chapman, D. (Eds.), Infrared Spectroscopy of Biomolecules. Wiley-Liss, New York, pp. 311–340.

Langston, F.M.A., Monaghan, J.M., Cousins, O., Nash, G.R., Bows, J.R., Chope, G., 2022. Manipulation of the phytochemical profile of Tenderstem((R)) broccoli florets by short duration, pre-Harvest LED lighting. Molecules 27 (10), 3224.

Larkin, P.J., 2018. Infrared and Raman Spectroscopy: Principles and Spectral Interpretation, 2nd ed. Elsevier, Stamford.

Ma, Y., He, H., Wu, J., Wang, C., Chao, K., Huang, Q., 2018. Assessment of polysaccharides from Mycelia of genus *Ganoderma* by mid-infrared and near-infrared spectroscopy. Sci. Rep. 8 (1), 10.

Martínez-Hernández, G.B., Artés-Hernández, F., Gómez, P.A., Artés, F., 2013. Induced changes in bioactive compounds of kailan-hybrid broccoli after innovative processing and storage. J. Funct. Foods 5 (1), 133–143.

Martínez-Hernández, G.B., Gómez, P.A., Pradas, I., Artés, F., Artés-Hernández, F., 2011. Moderate UV-C pretreatment as a quality enhancement tool in fresh-cut Bimi® broccoli. Postharvest Biol. Technol. 62 (3), 327–337.

Meyer, M., Adam, S.T., 2007. Comparison of glucosinolate levels in commercial broccoli and red cabbage from conventional and ecological farming. Eur. Food Res. Technol. 226 (6), 1429–1437.

Mithen, R.F., 2013. Development and commercialisation of 'Beneforté' broccoli and potential health benefits. Acta Hortic. 1005, 67–70. https://doi.org/10.17660/ActaHortic.2013.1005.4.

Mitsiogianni, M., Koutsidis, G., Mavroudis, N., Trafalis, D.T., Botaitis, S., Franco, R., Zoumpourlis, V., Amery, T., Galanis, A., Pappa, A., Panayiotidis, M.I., 2019. The role of isothiocyanates as cancer chemo-preventive, chemo-therapeutic and anti-melanoma agents. Antioxidants 8 (4), 106.

Mossoba, M.M., Shaw, G.J., Andrzejewski, D., Sphon, J.A., Page, S.W., 1989. Application of gas chromatography/matrix isolation/Fourier transform infrared spectrometry to the identification of glucosinolates from Brassica vegetables. J. Agric. Food Chem. 37 (2), 367–372.

Nieto, J.A., Hellín, P., Pérez, B., Viadel, B., Alapont, A., Agudelo, A., 2023. Fresh *Brassicaceae* sprouting broccoli (Bimi®) glucosinolates profile characterization and bioaccessibility through an in vitro dynamic digestion study. J. Food Compos. Anal. 115, 104941.

Ozaki, Y., 2021. Infrared spectroscopy-mid-infrared, near-infrared, and far-infrared/terahertz spectroscopy. Anal. Sci. 37 (9), 1193–1212.

Porep, J.U., Kammerer, D.R., Carle, R., 2015. On-line application of near infrared (NIR) spectroscopy in food production. Trends Food Sci. Technol. 46 (2), 211–230.

Renaud, E.N., van Bueren, Lammerts, Myers, E.T., Paulo, J.R., van Eeuwijk, M.J., Zhu, F. A., Juvik, J.A, N., 2014. Variation in broccoli cultivar phytochemical content under organic and conventional management systems: implications in breeding for nutrition. PLoS One 9 (7), e95683.

Renner, I.E., Fritz, V.A., 2020. Using Near-infrared reflectance spectroscopy (NIRS) to predict glucobrassicin concentrations in cabbage and Brussels sprout leaf tissue. Plant Methods 16, 136.

Rosa, E., David, M., Gomes, M.H., 2001. Glucose, fructose and sucrose content in broccoli, white cabbage and Portuguese cabbage grown in early and late seasons. J. Sci. Food Agric. 81 (12), 1145–1149.

Saha, S., Hollands, W., Teucher, B., Needs, P.W., Narbad, A., Ortori, C.A., Barrett, D.A., Rossiter, J.T., Mithen, R.F., Kroon, P.A., 2012. Isothiocyanate concentrations and interconversion of sulforaphane to erucin in human subjects after consumption of commercial frozen broccoli compared to fresh broccoli. Mol. Nutr. Food Res. 56 (12), 1906–1916.

Sahamishirazi, S., Zikeli, S., Fleck, M., Claupein, W., Graeff-Hoenninger, S., 2017. Development of a near-infrared spectroscopy method (NIRS) for fast analysis of total, indolic, aliphatic and individual glucosinolates in new bred open pollinating genotypes of broccoli (*Brassica oleracea* convar. *botrytis* var. *italica*). Food Chem. 232, 272–277.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36 (8), 1627–1639.

Śmiechowska, A., Bartoszek, A., Namieśnik, J., 2010. Determination of glucosinolates and their decomposition products—indoles and isothiocyanates in cruciferous vegetables. Crit. Rev. Anal. Chem. 40 (3), 202–216.

Toledo-Martin, E.M., Font, R., Obregon-Cano, S., De Haro-Bailon, A., Villatoro-Pulido, M., Del Rio-Celestino, M., 2017. Rapid and cost-effective quantification of glucosinolates and total phenolic content in rocket leaves by visible/near-infrared spectroscopy. Molecules 22 (5), 851.

Traka, M.H., Saha, S., Huseby, S., Kopriva, S., Walley, P.G., Barker, G.C., Moore, J., Mero, G., van den Bosch, F., Constant, H., Kelly, L., Schepers, H., Boddupalli, S., Mithen, R.F., 2013. Genetic regulation of glucoraphanin accumulation in Beneforte broccoli. New Phytol. 198 (4), 1085–1095.

Tripathi, M.K., As, M., 2017. Prospects and problems of dietary glucosinolates in animal feeding. Adv. Dairy Res. 5 (3), 1000180.

Tripathi, M.K., Mishra, A.S., 2007. Glucosinolates in animal nutrition: a review. Anim. Feed Sci. Technol. 132 (1–2), 1–27.

Turker-Kaya, S., Huck, C.W., 2017. A review of mid-infrared and near-infrared imaging: principles, concepts and applications in plant tissue analysis. Molecules 22 (1), 168.

Vo, Q.V., Rochfort, S., Nam, P.C., Nguyen, T.L., Nguyen, T.T., Mechler, A., 2018. Synthesis of aromatic and indole alpha-glucosinolates. Carbohydr. Res. 455, 45–53.

Vo, Q.V., Trenerry, C., Rochfort, S., Hughes, A.B., 2013. A total synthesis of (*R*, *S*) *S*-glucoraphanin. Tetrahedron 69 (41), 8731–8737.

Vo, Q.V., Trenerry, C., Rochfort, S., Wadeson, J., Leyton, C., Hughes, A.B., 2014. Synthesis and anti-inflammatory activity of indole glucosinolates. Bioorg. Med. Chem. 22 (2), 856–864.
Wallig, M.A., Belyea, R.L., Tumbleson, M.E., 2002. Effect of pelleting on glucosinolate content of Crambe meal. Anim. Feed Sci. Technol. 99, 205–214.
Zhong, X., Dolan, K.D., Almenar, E., 2015. Effect of steamable bag microwaving versus traditional cooking methods on nutritional preservation and physical properties of frozen vegetables: A case study on broccoli (*Brassica oleracea*). Innov. Food Sci. Emerg. Technol. 31, 116–122.