Article

FEASIBILITY STUDY ON THE USE OF VISIBLE-NEAR INFRARED SPECTROSCOPY FOR THE SCREENING OF INDIVIDUAL AND TOTAL GLUCOSINOLATE CONTENTS IN BROCCOLI

José Miguel Hernández-Hierro, Juan Valverde, Salvador Villacreces, Kim Reilly, Michael Gaffney, Maria Lourdes González-Miret, Francisco José Heredia, and Gerard Downey

Just Accepted

ACS Publications
High quality. High impact.

Feasibility Study on the Use of Visible-Near Infrared Spectroscopy for the Screening of Individual and Total Glucosinolate Contents in Broccoli

José Miguel Hernández-Hierro[†], Juan Valverde[‡], Salvador Villacreces[‡], Kim Reilly[§], Michael Gaffney[§], Maria Lourdes González-Miret[†], Francisco José Heredia[†], Gerard Downey[‡,*]

[†] Food Colour & Quality Laboratory, Department of Nutrition & Food Science, Universidad de Sevilla, Facultad de Farmacia, 41012 Sevilla, Spain.

[‡] Teagasc, Ashtown Food Research Centre, Dublin 15, Dublin, Ireland.

[§] Horticulture Development Unit, Teagasc Research Centre, Kinsealy, Dublin 17, Ireland.

* Corresponding author: Gerard Downey

Phone: (+353 1) 805 9500

Fax: (+353 1) 805 9550

E-mail: Gerard.Downey@teagasc.ie

1 **ABSTRACT**

2 The potential of visible-near infrared spectroscopy to determine selected individual and

3 total glucosinolates in broccoli has been evaluated. Modified partial least squares

4 (MPLS) regression was used to develop quantitative models in order to predict

5 glucosinolate contents. Both the whole spectrum and different spectral regions were

6 separately evaluated to develop the quantitative models, in all cases the best results were

7 obtained using the NIR zone between 2000 and 2498 nm. These models have been

8 externally validated for the screening of glucoraphanin, glucobrassicin, 4-

9 methoxyglucobrassicin, neoglucobrassicin and total glucosinolates contents. In addition,

10 discriminant partial least squares was used to distinguish between two possible broccoli

11 cultivars and showed a high degree of accuracy. In the case of the qualitative analysis,

12 best results were obtained using the whole spectrum (i.e. 400-2498 nm) with a correct

13 classification rate of 100% in external validation being obtained.

14

15 **KEYWORDS:** Glucosinolates; broccoli; visible spectroscopy; near infrared

16 spectroscopy; chemometrics

17

18   **INTRODUCTION**

19   Plant bioactive compounds, commonly referred to as phytochemicals, are components

20   that can affect cells in ways other than by provision of nutrients. Some of them may

21   have functional properties, meaning that they provide health benefits beyond basic

22   nutrition. It has been suggested that the health-promoting properties of plant foods are

23   due to their content of bioactive components with health-promoting effects [1-3].

24   Broccoli (*Brassica oleracea* L. var *Italica*) has been produced and consumed in Europe

25   since early times and is a characteristic constituent of European diets. This vegetable

26   contains significant amounts of glucosinolates, which are small organic molecules not

27   essential in the primary metabolism of the plant but with specific functions within its

28   secondary metabolism. Although their role in plants is unclear, their potent odour and

29   pronounced taste suggests a role in herbivore and microbial defence. Structurally,

30   glucosinolates (β-thioglucoside-N-hydroxysulfates) (**Figure 1A**) are characterised by

31   the presence of nitrogen and sulphur groups and they are derived from glucose and an

32   amino acid [1]. Glucosinolates are not bioactive until they have been enzymatically

33   hydrolysed to the associated isothiocyanate by an endogenous myrosinase enzyme that

34   may be released by disruption of the plant cell through harvesting, processing or

35   mastication. Epidemiological studies have consistently reported a reduction in incidence

36   of chronic diseases such as cancer and myocardial infarction as a result of induction of

37   detoxifying enzymes and reduction of oxidative stress, although the anti-nutritive

38   effects of both glucosinolates and hydrolysis products have also been studied [1, 4, 5]. It is

39   therefore important to characterise the content of bioactive compounds in new or re-

40   introduced cultivars.

41   To this end, the use of rapid methods of analysis with minimal or no sample preparation

42   could be important. Such a capability would significantly enhance the ability of primary

43   processors to select samples for the market based on a high content of bioactive

44   compounds in a rapid and inexpensive manner.

45   Near infrared spectroscopy (NIRS) provides fast and non-destructive analysis. It allows

46   qualitative and quantitative analysis to be performed in different matrices thereby

47   reducing costs when compared to wet chemical analysis and without generating waste.

48   In the recent past, the use of NIR spectroscopy in the determination of bioactive

49   compounds in foods and plant natural products has increased considerably [6, 7]. With

50   particular regard to glucosinolates, it has been used to determine these compounds in

51   both seeds [8-13] and leaves [14, 15] from a wide range of *Brassica* species. The standard

52   errors of the glucosinolate's predictions reported in the aforementioned studies ranged

53   from 2.65 to 15.65 μmol g^{-1} DM for three cultivars of rapeseeds [13] and Indian mustard

54   seeds [11] respectively. The results reported in all these works reveal that NIRS is able to

55   predict glucosinolates in a wide range of matrices. Therefore, it may be expected that

56   this technique coupled with chemometric tools could provide an alternative method to

57   undertake the analysis of glucosinolates in broccoli.

58   The aim of this study was to evaluate the potential of visible-NIR technology to

59   determine   both   selected   individual   (glucoraphanin,   glucobrassicin,   4-

60   methoxyglucobrassicin, neoglucobrassicin) and total glucosinolates in broccoli. In

61   addition to this, its potential for the discrimination between two broccoli cultivars was

62   evaluated. To our knowledge, this is the first time that vis-NIRS technology has been

63   applied to broccoli for these purposes. The work is part of a wider study aimed at

64   maximising the concentrations of phytochemicals in vegetables by adopting a farm to

65   fork approach. The projected final output of the wider study is the production of

66   vegetables with optimal levels of selected phytochemical groups.

67   **MATERIAL AND METHODS**

68   **Samples**

69   Samples were generated at Teagasc, Kinsealy Research Centre, located in the north of

70   County Dublin, Ireland (53° 25' 10" N Lat 6° 10' 45" W). The soil is a grey brown

71   podzolic and presents a loam to clay loam topsoil. The experimental design was 2^3

72   factorial split plot design with four plot replicates. Briefly, eight treatments in 5.5m x

73   3.4 m plots were applied with four replicates combining conventional or organic soil

74   management with organic or conventional pest control. The four combinations were

75   sown with Belstar or Fiesta broccoli cultivars. **Table 1** shows the different agricultural

76   managements applied in this study. The aforementioned factorial trial was carried out

77   during the 2009 and 2010 years and provided 64 samples.

78   Broccoli samples were harvested, frozen and freeze-dried in a large scale freeze drier

79   (Frozen in Time Ltd. York, UK). The specimen chamber was kept below 0 °C during

80   the whole freeze-drying process. Once freeze-dried, samples were milled (Blixer 4,

81   Robotcoupe, France), vacuum packed in polypropylene bags and stored at -20 °C prior

82   to analysis. Two aliquots were taken from each sample, one for the micellar

83   electrokinetic capillary chromatography (MEKC) analysis and the other for visible-near

84   infrared analysis.

85   **Chemical Analysis**

86   *Sulphatase Purification Procedure.* Sulphatase (Type H-1 from Helix pomatia, Sigma,

87   MO, USA) was purified by dissolving the sulphatase powder (70 mg) in deionised

88   water (3 mL) and adding ethanol (3 mL). This solution was centrifuged (12000 rpm, 10

89   min, room temperature) and ethanol (9 mL) was added to the supernatant after which

90   the solution was centrifuged again (12000 rpm, 10 min, room temperature). The

91   resulting pellet was dissolved in deionised water (2 mL) and this sulphatase solution

92   was subsequently passed through a 0.5 mL DEAE Sephadex A-25 and a 0.5 mL SP

93   Sephadex C-25 column. This solution was collected in a vial and kept at -20 °C until

94   use.

95   *Sampling and Extraction.* Freeze-dried broccoli powder was extracted using pressurised

96   liquid extraction with an ASE200 instrument (Dionex; Sunnyvale, CA,USA) with

97   attached solvent controller. Extraction of 1.00 g freeze-dried broccoli was carried out in

98   22mL steel cartridges. Extraction conditions were slightly modified from Mohn *et al* [16].

99   In brief, preheat time: 1 min; static extraction per cycle: 5 min; flush: 60% of cell

100 volume; purge: 60 s with nitrogen; pressure: 120 bar. Glucotropaeolin (Phytolab GmbH,

101 Vestenbergsgreuth, Germany) was used as an internal standard - 100μL of a solution of

102 0.5 μmol of potassium salt of glucotropaeolin in Milli –Q water was spiked onto the

103 sample packed in the ASE cell. The collected extract (ca 30 mL) was dried under a

104 constant N_2 flow and re-dissolved in deionised water (7 mL). An aliquot (1 mL) of this

105 crude extract was applied to a DEAE Sephadex A-25 column (0.5 mL) and the unbound

106 material was removed by washing with deionised water (2 x 1 mL) and sodium acetate

107 buffer (2 x 0.5 mL; 20 mM, pH 5.0). After washing, the purified sulphatase (75 μL) was

108 added and the columns were incubated overnight at room temperature. After overnight

109 incubation, the desulphoglucosinolates (dGLS) were eluted from the columns with

110 deionised water (3 x 1 mL). The collected eluate was dried under constant N_2 flow, re-

111 dissolved in deionised water (200 μL) and centrifuged prior to analysis. An aliquot of

112 50 μL was used for the MEKC analysis.

113 *Micellar Electrokinetic Capillary Chromatography.* Analyses were performed using a

114 CE capillary electrophoresis system (Agilent, Waldbronn, Germany) equipped with

115 diode array detector. All separations were performed on a fused silica capillary (Agilent,

116 Stevens Creek, CA; 75 lm ID, 64.5 cm total length, 56 cm effective length). Samples

117 were injected from the anodic end of the capillary (vacuum injection, 50 mbar, 1 s). The

118   separation buffer consisted of sodium chlorate (250 mM) and boric acid (200 mM) at

119   pH 8.5; the separation was carried out at 12 kV and 60 °C. The capillary was

120   conditioned between each run sequentially with 1.0 M NaOH (3 min), 0.1 M NaOH (1

121   min), water (1 min) and separation buffer (5 min). Detection was performed on column

122   at 230 and 280 nm. Data processing was carried out with 3D-CE Chemstation software

123   (Agilent Waldbronn, Germany). The quantity of the dGLS was estimated as the average

124   of quantities calculated from the internal standards, taking into account the relative

125   response factors revised by Clarke *et al* [17].

126   **Near Infrared Spectroscopy Analysis**

127   Each aliquot of freeze-dried broccoli powder was tempered at room temperature for 16

128   hours prior to spectroscopy analysis. A Foss NIRSystems 6500 spectrometer (NIR

129   Systems Inc., Silver Spring MD) was used; samples were placed in small ring cups (3

130   cm diameter screen size) with a disposable backing disc and spectral measurements

131   were made in reflectance mode over the 400-2498 nm wavelength range. Spectra were

132   recorded in duplicate at intervals of 2 nm with 10 scans being performed for the

133   reference tile and 25 for samples; the sample cell was rotated through 180 degrees

134   between spectral collections. To minimise sampling error, all the samples were analysed

135   in duplicate with re-filling of the ring cup. Therefore, the four recorded spectra of each

136   sample were averaged to obtain each sample spectrum. Samples were scanned in

137   random order. The spectrum of sinigrin standard powder ((−)-sinigrin hydrate from

138   horseradish, Sigma-Aldrich Ireland Ltd, Arklow, Ireland) was also recorded in

139   reflectance mode over the wavelength range 1100-2498 nm at intervals of 2 nm. The

140   software used was Win ISI® (v1.50) (Infrasoft International, LLC, Port. Matilda, PA,

141   USA). This software allowed not only the instrument control and spectral acquisition

142   but also the data pre-treatment and development of qualitative and quantitative models.

143   From the whole data set, 25% of the samples (16 samples) were randomly allocated into

144   the validation set and the remaining 75% (48 samples) into the calibration set.

145   **Chemometric Techniques**

146   A supervised pattern recognition technique-Discriminant Partial Least Squares (DPLS)-

147   was used for qualitative analysis. The calibration was conducted by performing a

148   regression on the spectral information and the associated dummy sample group values,

149   in this case defined as 1 (variety 1) or 2 (variety 2). The regression method applied to

150   this procedure was MPLS, which is a modification of a normal PLS 1 [18, 19].

151   Prior to quantitative analysis, an unsupervised pattern recognition technique, principal

152   component analysis (PCA), was used in order to provide information about the latent

153   structure of spectral data. This method provides not only information related to spectral

154   outliers and the distribution of samples in the newly-created space but is also an

155   important source of knowledge with which to create cross-validation groups used in the

156   calibration process. PCA is also a useful tool to identify whether unknown samples

157   belong to the spectral space created by the samples from which the equations were

158   developed. Should this not be the case, the equations should not be used to make any

159   prediction [20, 21]. Using the raw spectral data and testing different spectral pre-treatments

160   and allocating the corresponding MEKC-DAD glucosinolate values to each sample,

161   calibrations were performed by modified partial least squares regression (MPLS). In this

162   method, the set of calibration samples is divided into a series of subsets in order to

163   perform cross-validation to set the number of PLS factors, reduce the likelihood of

164   overfitting [20] and remove chemical outliers. Using the $T \geq 2.5$ criterion, samples that

165   presented a high residual value when they were predicted were eliminated from the set.

166   Finally, validation errors are combined into a single figure, the standard error of cross-

167   validation (SECV).

8

168   Spectral pre-treatments are usually applied to NIR raw data; scattering effects were

169   removed using multiplicative scatter correction (MSC), standard normal variate (SNV),

170   and detrending [22, 23]. Moreover, the effect of derivatisation and variations in spectral

171   ranges were tested in the development of the NIRS calibrations.

172   **RESULTS AND DISCUSSION**

173   **Chemical Analysis**

174   Glucosinolates were determined by MEKC-DAD. The common glucosinolate skeleton

175   structure is shown in **Figure 1A**, leading to the expectation that their spectral signatures

176   should also be very similar. Sinigrin's reflectance spectrum (**Figure 1B**) shows

177   considerable spectral detail in the wavelength range above about 2000 nm; it was the

178   only pure glucosinolate available for spectral analysis but, given that its structure is

179   similar to the others being analysed in these broccoli samples, it is expected that all

180   glucosinolates will have similar information above 2000 nm. **Table 2** shows the range,

181   mean value and standard deviation of the selected individual glucosinolates and the total

182   glucosinolates in the broccoli samples analysed. These values were used in the

183   development of the quantitative models as reference values. It is noticeable that the

184   precision of the laboratory methods is quite poor with standard deviation between

185   replicate figures of between 15.9% (neoglucobrassicin) and 27.2% (glucoraphanin) of

186   mean contents.

187   **Near Infrared Spectroscopy Analysis and Chemometric Techniques**

188   *Near Infrared Spectra.* **Figure 1** shows the average and standard deviation spectra of

189   broccoli powders over the 400-2498 nm range (**Figure 1C**) together with the sinigrin

190   powder spectrum between 1100-2498 nm (**Figure 1B**).

191 Standard deviation spectra have been multiplied by a factor of 10 for display reasons.

192 Spectral intensities were low and well within the linear response range of the instrument

193 detector range; little noise was evident in these traces. A strong feature of the sample

194 spectra was the absorbance pattern in the visible wavelength range i.e. 400-780 nm.

195 This arose from the pigment remaining in the freeze-dried samples, all of which

196 displayed a green tinge.

197 **Figure 1D** shows the second derivative spectra of broccoli and sinigrin and presents

198 considerable spectral similarities over the 2000-2498 nm wavelength range. It may be

199 expected that this range would be useful for quantitative purposes of glucuosinolates.

200 *Qualitative analysis.* An SNV 2,4,4,1 spectral pre-treatment was applied to whole vis-

201 NIR spectra of samples from the calibration set and then a principal component analysis

202 was carried out. **Figure 2** shows the scores of the broccoli samples in the space defined

203 by the first and second principal components which described 52.3% (PC1) and 23.7%

204 (PC2) of the variability in the data. In this plot, slight differences between the 2009 and

205 2010 harvest years (**Figure 2A**) and also between the Belstar and Fiesta cultivars

206 (**Figure 2B**) are apparent. The main difference observed in these plots was between

207 harvest years although the separation between 2009 and 2010 samples was not

208 complete. This trend in harvest year locations was discernible on the basis of PC2.

209 Cultivars were completely overlapped in this hyperspace.

210 D-PLS modelling of the groups was carried out initially using the complete vis-NIR raw

211 spectra and one dummy variable, the values of which were defined as 1 (variety

212 1=Belstar) or 2 (variety 2=Fiesta). The model was developed using 6 PLS factors and

213 presented an RSQ of 0.752, SEC of 0.27 and SECV 0.34. The spectral regions between

214 600-700 nm and 2000-2498 nm showed important contributions to the model loadings

215 (**Figure 3A**) and are mainly related to green pigments such as chlorophyll [14, 15, 24] and C-

216   H, O-H and N-H overtones present in a number of bonds [25, 26]. The DPLS developed

217   model predicted a dummy value for each sample and then the samples were allocated

218   according to their predicted values ±0.5 in the corresponding cultivar with a dummy

219   variable breakpoint of 1.5. The predicted dummy variable plot for the qualitative model

220   is shown in **Figure 3B**. Samples of Belstar cultivar cluster around a dummy value of

221   1.0, while those belonging to Fiesta cultivar around 2.0. A prediction rate of 100%

222   samples correctly classified was obtained in external validation. This demonstrates that

223   vis-NIR spectroscopy coupled to discriminant DPLS analysis permit a clear

224   differentiation between the two aforementioned broccoli cultivars. Regarding the

225   studied compound, the classification of samples according to their cultivar could be

226   useful since they present significant differences in the 4-methoxyglucobrassicin

227   (p<0.001) and neoglucobrassicin (p<0.001) contents.

228   *Quantitative analysis.* Although different spectral ranges and the whole spectrum have

229   been separately evaluated, in all cases the best results were obtained using the

230   wavelength range 2000-2498 nm. As part of the quantitative analysis, a SNV (2,4,4,1)

231   spectral pre-treatment was applied to this spectral range of samples in the calibration set

232   and then principal component analysis was carried out in order to look for spectral

233   outliers and create cross-validation groups. Overall, the spectral variability explained

234   was 98% using 13 principal components and Mahalanobis distances for each sample

235   were calculated. Samples were ranked in order of their H (Mahalanobis) distance from

236   the mean spectrum of the entire sample set and the H > 3 criterion was applied. No H-

237   outliers were found.

238   The average spectrum of the best of the different mathematically pre-treated spectra are

239   shown in **Figure 4**. The statistical parameters of the final calibration equations are

240   shown in **Table 3** where N is the number of samples used to obtain the calibration

241 equation after eliminating samples for chemical reasons (T criterion). The best of the

242 different mathematical treatments, concentration range, and standard deviations are also

243 shown. We checked the robustness of the method by applying NIRS technology to 16

244 samples that did not belong to the calibration group. **Table 3** also shows the results

245 obtained in the external validation and the SEP values are presented. The percentage of

246 error for total glucosinolates is similar to the errors previously reported for other

247 matrices such as Indian mustard seeds[11].

248 The magnitude of these errors indicate the potential of the NIR spectroscopy for

249 prediction of glucosinolates in freeze-dried broccoli and is satisfactory for the screening

250 of glucoraphanin, glucobrassicin, 4-methoxyglucobrassicin, neoglucobrassicin and total

251 glucosinolates. However, individual glucosinolates biosynthesis may follow essentially

252 the same course so inter-correlations among them could be expected. It is not possible

253 to ascertain if the results of NIRS models for predicting the composition of these

254 glucosinolates were due to their real absorbance or the correlation between them.

255 **Figure 5** shows the loading and the beta regression coefficients plots of the MPLS

256 model for total glucosinolates. The spectral regions around 2072 nm, 2220 nm, 2300 nm

257 and 2446 nm show important contributions to the model loadings. These could be

258 related to N-H bonds (2072 nm), to C-H bonds (2220 nm), to first overtones of O-H

259 alcohol functional group, to stretching-bending of CH-CH$_2$ bonds and C-O bonds (2300

260 nm) and to C-N bonds (2446 nm). These can be attributed to the chemical structure of

261 the compounds analyzed [25, 26]. This confirms previous studies that showed important

262 contributions in the aforementioned spectral zones for determining glucosinolates in

263 other matrices [11, 14, 15, 27].

264 The potential of vis-NIRS for determination of glucosinolates in freeze-dried broccoli

265 was examined. The procedure reported here presents a good potential for a fast and

266   reasonably inexpensive screening of almost all individual glucosinolates and total

267   glucosinolates present in broccoli and also to distinguish between broccoli cultivars.

268   Although different spectral ranges and the whole spectrum have been separately

269   evaluated to develop the quantitative models, in all cases the best results (i.e. the highest

270   RSQ and the lowest SECV and SEP values) were obtained using the NIR region

271   between 2000 and 2498 nm. In the case of the qualitative analysis, best results were

272   obtained used the whole spectrum (i.e. 400-2498 nm). Nonetheless, a comprehensive

273   study should be made in order to evaluate other factors such as different production

274   areas and varieties, in the complete development of these models.

281

282 **REFERENCES**

283

284 (1) Crozier, A.; Clifford, M. N.; Ashihara, H., *Plant secondary Metabolites.*
285 *Occurrence, Structure and Role in the Human Diet*. Blacwell Publishing: Garsington
286 Road, Oxford, UK, 2006.

287 (2) Sloan, A. E., Top 10 functional food trends. *Food Technol-Chicago* **2008,** *62*,
288 25-44.

289 (3) Wildman, R. E. C., *Handbook of Nutraceuticals and Functional Foods*. CRC
290 Press LLC: Boca Raton, FL, USA, 2001.

291 (4) Shahindi, F., *Canola and rapeseed production, chemistry and processing*
292 *technology*. Van Nostrand Reinhold: New York, USA, 1990.

293 (5) Verkerk, R.; Schreiner, M.; Krumbein, A.; Ciska, E.; Holst, B.; Rowland, I.; De
294 Schrijver, R.; Hansen, M.; Gerhäuser, C.; Mithen, R.; Dekker, M., Glucosinolates in
295 Brassica vegetables: The influence of the food supply chain on intake, bioavailability
296 and human health. *Mol. Nutr. Food Res.* **2009,** *53*, S219-S219.

297 (6) McGoverin, C. M.; Weeranantanaphan, J.; Downey, G.; Manley, M., The
298 application of near infrared spectroscopy to the measurement of bioactive compounds in
299 food commodities. *J. Near Infrared Spectrosc.* **2010,** *18*, 87-111.

300 (7) Cozzolino, D., Near Infrared Spectroscopy in Natural Products Analysis. *Planta*
301 *Med.* **2009,** *75*, 746-756.

302 (8) Biston, R.; Dardenne, P.; Cwikowski, M.; Marlier, M.; Severin, M.; Wathelet, J.
303 P., Fast Analysis of Rapeseed Glucosinolates by near-Infrared Reflectance
304 Spectroscopy. *J. Am. Oil Chem. Soc.* **1988,** *65*, 1599-1600.

305 (9) Velasco, L.; Becker, H. C., Analysis of total glucosinolate content and individual
306 glucosinolates in Brassica spp. by near-infrared reflectance spectroscopy. *Plant*
307 *Breeding* **1998,** *117*, 97-102.

308 (10) Hom, N.; Becker, H.; Möllers, C., Non-destructive analysis of rapeseed quality
309 by NIRS of small seed samples and single seeds. *Euphytica* **2007,** *153*, 27-34.

310 (11) Font, R.; del Rio, M.; Fernandez-Martinez, J. M.; de Haro-Bailon, A., Use of
311 Near-Infrared Spectroscopy for Screening the Individual and Total Glucosinolate
312 Contents in Indian Mustard Seed (*Brassica juncea* L. *Czern. & Coss.*). *J. Agric. Food*
313 *Chem.* **2004,** *52*, 3563-3569.

314 (12) Font, R.; del Río-Celestino, M.; de Haro-Bailon, A., The use of near-infrared
315 spectroscopy (NIRS) in the study of seed quality components in plant breeding
316 programs. *Ind. Crop. Prod.* **2006,** *24*, 307-313.

317 (13) Petisco, C.; García-Criado, B.; Vázquez-de-Aldana, B. R.; de Haro, A.; García-
318 Ciudad, A., Measurement of quality parameters in intact seeds of Brassica species using
319 visible and near-infrared spectroscopy. *Ind. Crop. Prod.* **2010,** *32*, 139-146.

320 (14) Font, R.; del Rio Celestino, M.; Rosa, E.; Aires, A.; de Haro-Bailon, A.,
321 Glucosinolate assessment in Brassica oleracea leaves by near-infrared spectroscopy. *J.*
322 *Agric. Sci.* **2005,** *143*, 65-73.

323   (15)   Font, R.; del Rio-Celestino, M.; Cartea, E.; de Haro-Bailon, A., Quantification
324   of glucosinolates in leaves of leaf rape (Brassica napus ssp. pabularia) by near-infrared
325   spectroscopy. *Phytochemistry* **2005,** *66*, 175-185.

326   (16)   Mohn, T.; Cutting, B.; Ernst, B.; Hamburger, M., Extraction and analysis of
327   intact glucosinolates--A validated pressurized liquid extraction/liquid chromatography-
328   mass spectrometry protocol for Isatis tinctoria, and qualitative analysis of other
329   cruciferous plants. *J. Chromatogr. A* **2007,** *1166*, 142-151.

330   (17)   Clarke, D. B., Glucosinolates, structures and analysis in food. *Anal. Methods*
331   **2010,** *2*, 310-325.

332   (18)   Lars, S.; Svante, W., Partial least squares analysis with cross-validation for the
333   two-class problem: A Monte Carlo study. *J. Chemometr.* **1987,** *1*, 185-196.

334   (19)   Massart, D. L.; Vandeginste, B. G.; Buydens, L. M.; Lewi, P. J.; Smeyers-
335   Verbeke, J.; Jong, S. D., *Handbook of Chemometrics and Qualimetrics*. Elsevier
336   Science Inc.: 1998; p 1616.

337   (20)   Shenk, J. S.; Westerhaus, M. O., *Routine Operation,Calibration, Development*
338   *and Network System Management Manual, NIRSystems*. Silver Spring: MD, USA,
339   1995; p 239.

340   (21)   Brereton, R. G., *Chemometrics : data analysis for the laboratory and chemical*
341   *plant*. J. Wiley: Chichester, West Sussex, England, 2003; p 489.

342   (22)   Geladi, P.; MacDougall, D.; Martens, H., Linearization and Scatter-Correction
343   for Near-Infrared Reflectance Spectra of Meat. *Appl. Spectrosc.* **1985,** *39*, 491-500.

344   (23)   Dhanoa, M. S.; Lister, S. J.; Barnes, R. J., On the Scales Associated with Near-
345   Infrared Reflectance Difference Spectra. *Appl. Spectrosc.* **1995,** *49*, 765-772.

346   (24)   Tkachuk, R.; Kuzina, F. D., Chlorophyll analysis of whole rapeseed kernels by
347   near infrared reflectance. *Can. J. Plant Sci.* **1982,** *62*, 875-884.

348   (25)   Osborne, B. G.; Fearn, T.; Hindle, P. T., *Practical NIR spectroscopy with*
349   *applications in food and beverage analysis*. Longman Scientific & Technical ; Wiley:
350   Harlow, Essex, England : New York :, 1993; p 224.

351   (26)   Siesler, H. W.; Ozaky, Y.; Kawata, S.; Heise, H. M., *Near infrared*
352   *spectroscopy: principles, instruments, applications*. Wiley-VCH: Weinheim, Germany,
353   2002; p 348.

354   (27)   Tkachuk, R., Oil and Protein-Analysis of Whole Rapeseed Kernels by near-
355   Infrared Reflectance Spectroscopy. *J. Am. Oil Chem. Soc.* **1981,** *58*, 819-822.

356

Figure captions

Figure 1. (A) Glucosinolete skeleton. (B) Sinigrin powder spectrum (1100-2498 nm). (C) Average and standard deviation spectra (400-2498 nm) of broccoli powders. (D) 2nd derivative spectra of the broccoli and sinigrin (1100-2000 nm).

Figure 2. Score plot of broccoli samples (calibration set) in the space defined by PC1 and PC2.

Figure 3. (A) Loading plot of the DPLS model. (B) Dummy variable plot of broccoli samples in the external validation of the DPLS method.

Figure 4. Average spectrum of different mathematical pre-treatments applied to broccoli powder spectra (2000-2498 nm).

Figure 5. (A) Loading and (B) beta regression coefficients plots of the MPLS model for total glucosinolate prediction.

16

Table 1. Different Agricultural Managements Applied in this Study.

	Organic	Conventional
Soil treatment	• Four year rotation ley: red clover-broccoli-onion-carrot • Additional organic fertilisation as indicated by soil test • Winter cover crop	• No rotation, plots, randomly allocated each year • Mineral fertilisers as indicated by soil test • No ley crop • No winter crop
Pest control	• Certified organic seed • Refuge area • *Brassica* collars on broccoli • Certified organic chemicals (eg. garlic spray) • Weed control by mechanical methods	• Chemically treated seed • Chemical weed control–herbicides • Chemical pest control–fungicides and insecticides

17

Table 2. Statistical Descriptors of Glucosinolate Contents Determined by MEKC-DAD (µmol g^{-1} DM).

	GBS[a]	GRA[b]	4MGBS[c]	NGBS[d]	TGS[e]
Min (n=64)	1.60	1.09	0.09	0.74	4.74
Max (n=64)	6.83	4.82	0.55	4.45	15.10
Mean (n=64)	3.92	2.83	0.28	1.97	9.53
SD (n=64)	1.09	0.87	0.10	0.86	1.87
SD replicates	0.88	1.01	0.08	0.59	1.80

[a]GBS: glucobrassicin; [b]GRA: glucoraphanin; [c]4MGBS: 4-methoxyglucobrassicin; [d]NGBS: neoglucobrassicin; [e]TGS: total glucosinolates

Table 3. Calibration Statistical Descriptors for the Models Developed in the NIR Zone Close to 2000-2498 nm.

Spectral pre-treatments	Compounds	T outliers	PLS factors	N[a]	Mean	Est. Min	SD[b]	Est. Max	SEC[c]	RSQ[d]	SECV[e]	SEP[f]
							(μmol g^{-1} DM)				(μmol g^{-1} DM)	
Detrend 1,4,4,1	GBS[g]	2	7	46	3.93	0.68	1.09	7.18	0.35	0.89	0.52	0.74
Detrend 1,4,4,1	GRA[h]	1	2	47	2.86	0.26	0.87	5.46	0.67	0.40	0.71	0.75
SNV 1,4,4,1	4MGBS[i]	0	5	48	0.28	0.00	0.11	0.62	0.06	0.69	0.07	0.05
MSC 2,8,6,1	NGBS[j]	0	5	48	1.93	0.00	0.86	4.51	0.48	0.68	0.66	0.74
None 2,10,10,1	TGS[k]	3	5	45	9.40	4.04	1.79	14.76	0.93	0.73	1.11	1.44

[a]N: number of samples (calibration set); [b]SD: standard deviation; [c]SEC: standard error of calibration; [d]RSQ: coefficient of determination;
[e]SECV: standard error of cross-validation; [f]SEP: standard error of prediction; [g]GBS: glucobrassicin; [h]GRA: glucoraphanin;
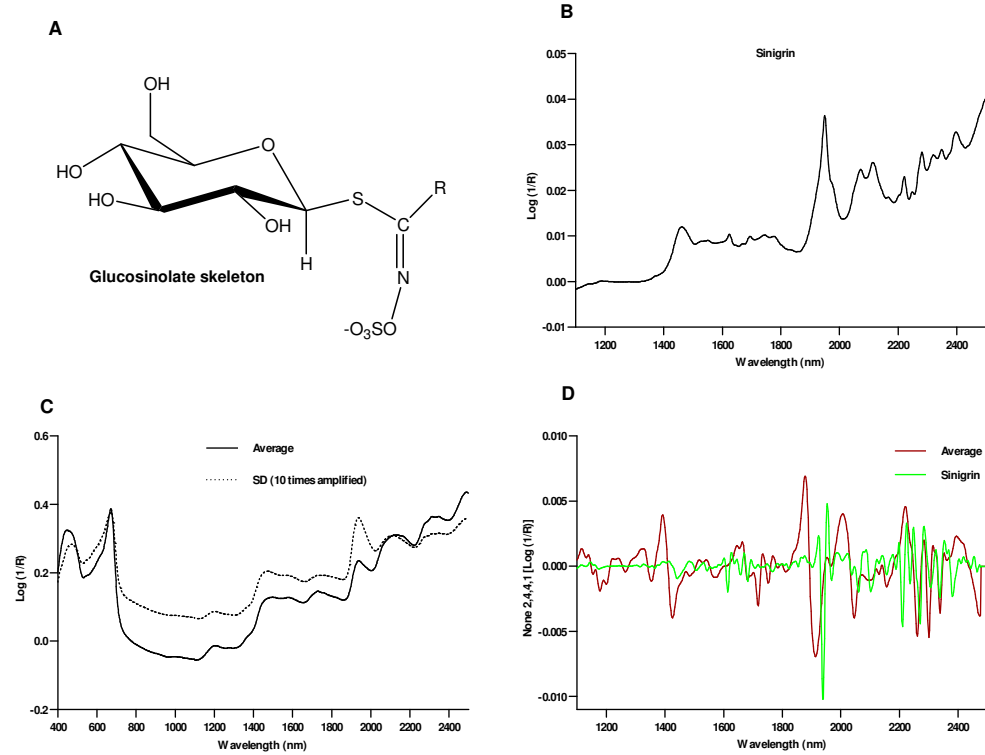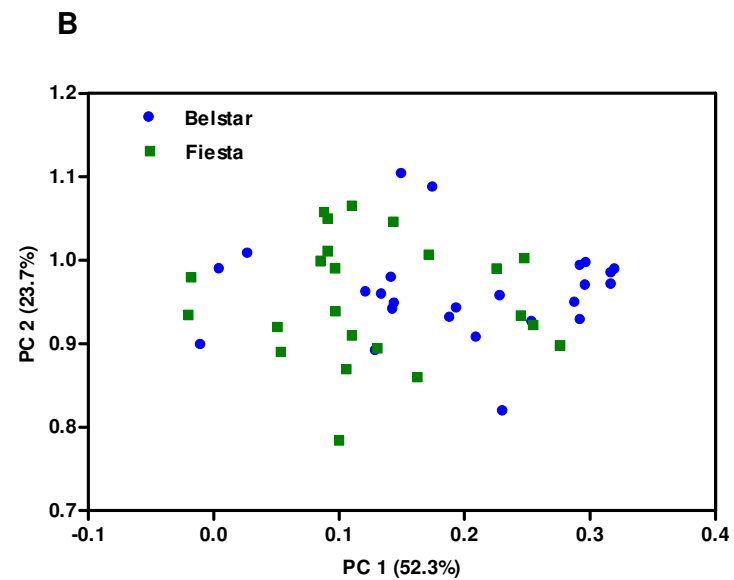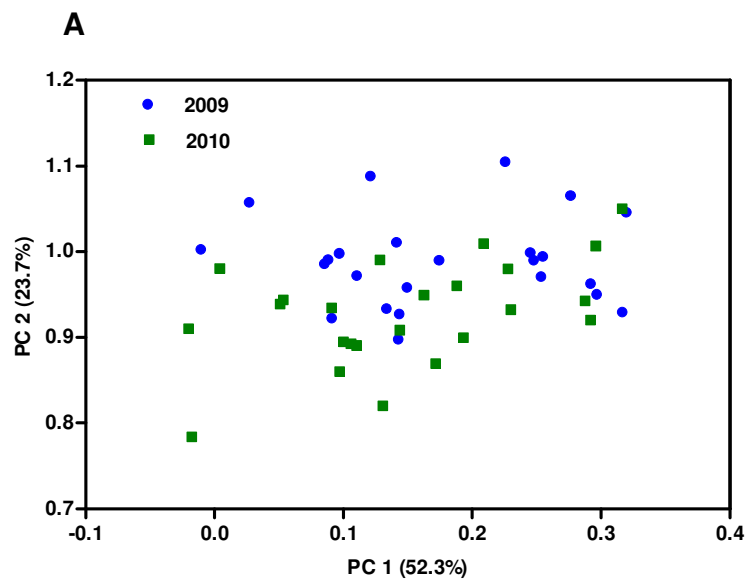[i]4MGBS: 4-methoxyglucobrassicin; [j]NGBS: neoglucobrassicin; [k]TGS: total glucosinolates
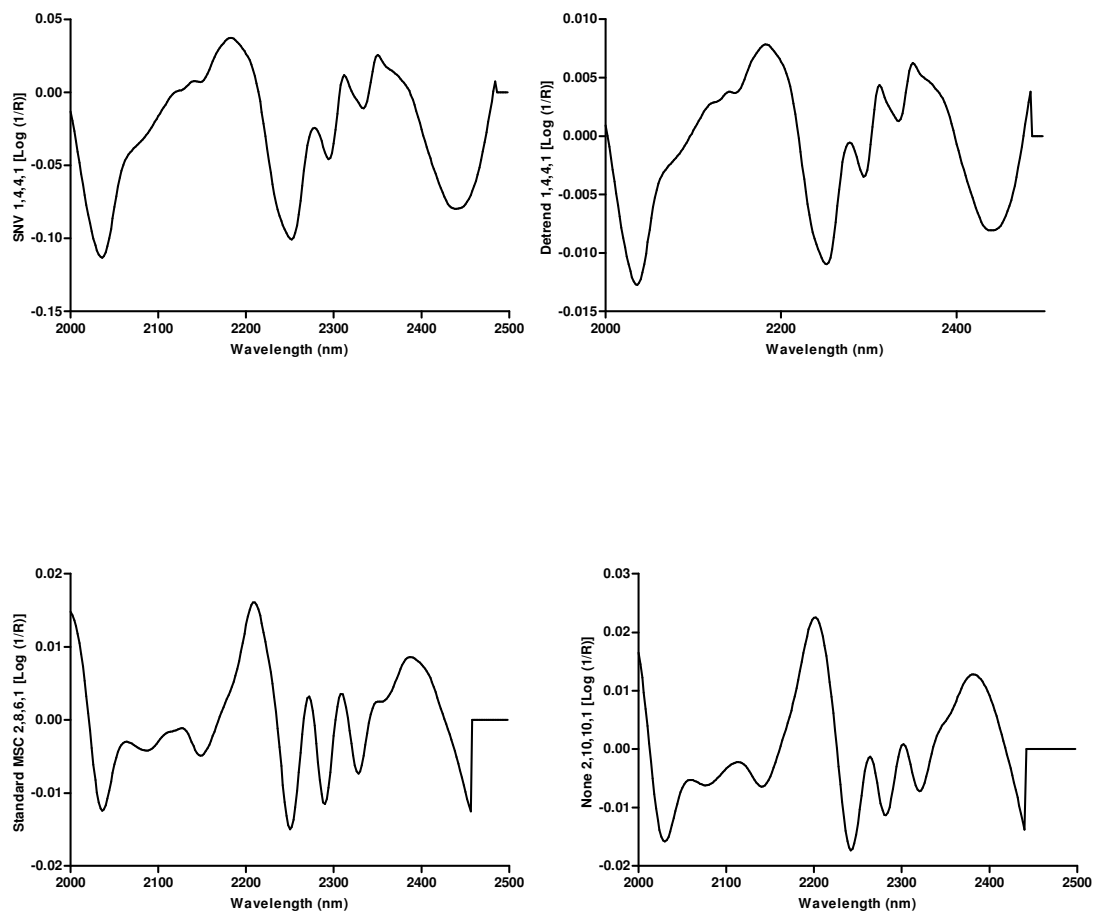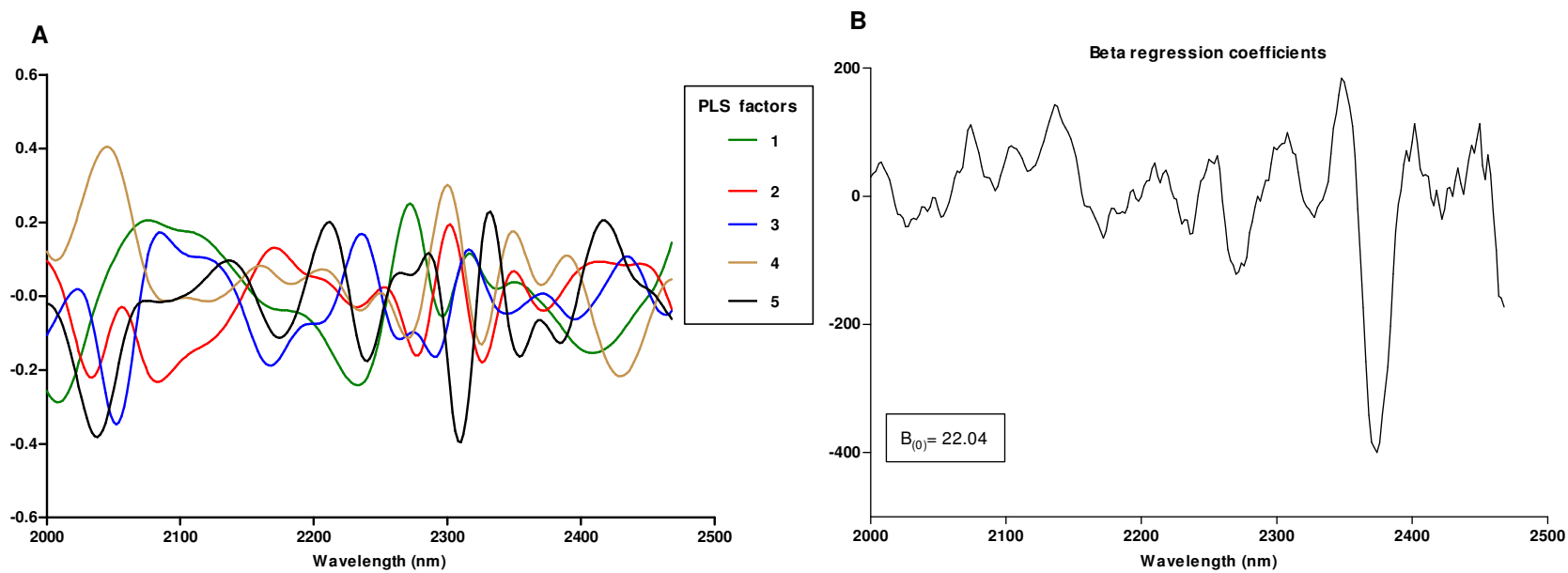
Figure 1.

Figure 2.

21
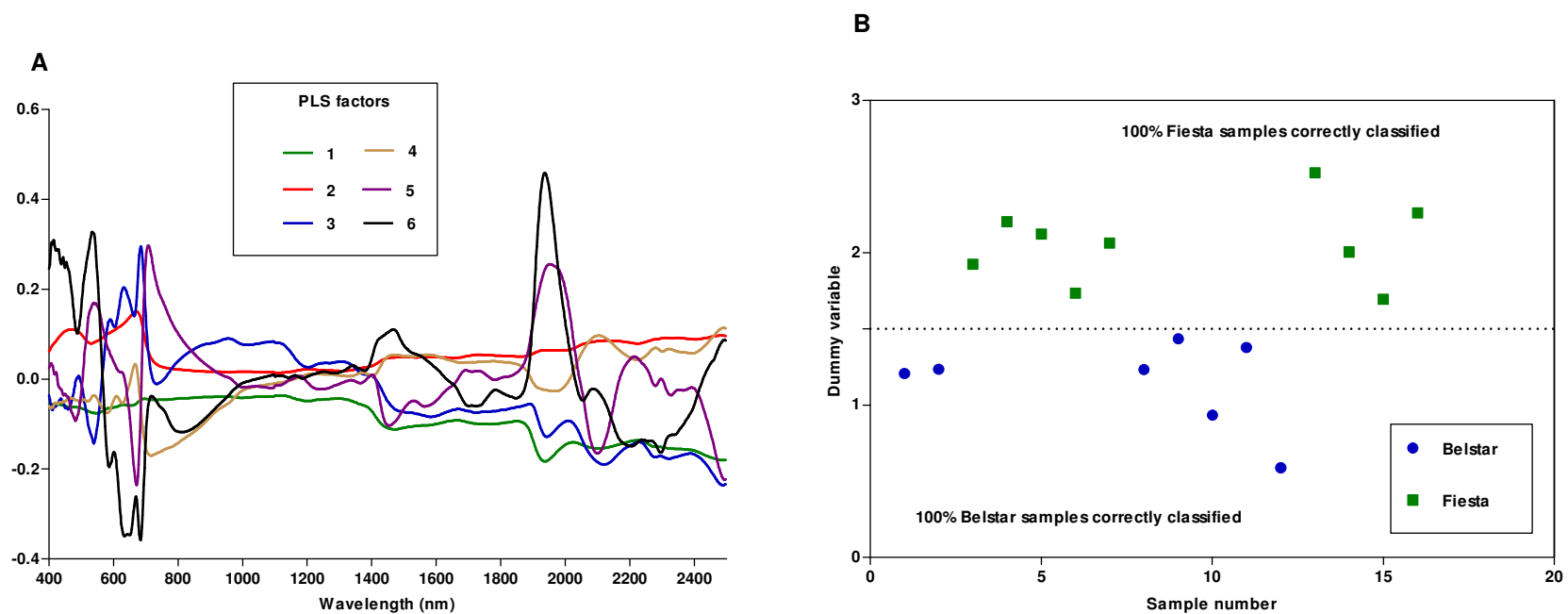
Figure 3.

Figure 4.

Figure 5.

Graphic for table of contents