

Original Research Article

Non-targeted metabolomic profiling of maize landraces (*Zea mays* L.) combined with chemometric tools

Abstract

Grain samples of maize landraces were collected and subjected to Fourier Transform Infrared spectroscopy (FTIR) analysis combined with chemometric tools in order to discriminate them regarding the chemical composition. Principal component analysis (PCA) and hierarchical clustering (HCA) were applied on selected peaks of the spectral data. The most important chemical groups found in all maize landrace samples were monoterpenes, sesquiterpenes, tetraterpenes, aminoacids, polysaccharides, lipids and proteins. Trace signals of secondary metabolites were also found in cultivars according to year of harvest.

Keywords: Maize landraces, chemical composition, FTIR, non-supervised techniques, R software.

1. Introduction

Commonly referred to as maize or corn, *Zea mays* ssp. *mays* is one of the world's most important crop plants, achieving a multibillion dollar annual revenue. In addition to its agronomic importance, maize has been a keystone model organism for basic research for nearly a century. (Dwivedi et al. 2016; Strable & Scanlon, 2009). According to Dwivedi et al. (2016), the current industrial agriculture system may be the single most

25 important threat to maize biodiversity. A serious consequence of biodiversity loss is
26 the displacement of locally adapted maize landraces with adaptation traits to future
27 climates by mono-cropping with genetically uniform hybrids and improved cultivars.

28

29 Maize landraces represent heterogeneous, local adaptations of domesticated species,
30 and thereby provide genetic resources that meet current and new challenges for
31 farming in stressful environments. These local ecotypes can show variable phenology
32 and low-to-moderate edible yield, but are often highly nutritious. The main
33 contributions of maize landraces to plant breeding have been traits for more efficient
34 nutrient uptake and utilization, as well as useful genes for adaptation to stressful
35 environments such as water stress, salinity, and high temperatures (Dwivedi et al.
36 2016; Vasconcelos et al. 2013; Bellon & Hellin, 2011; Mercer et al. 2008).

37

38 Systematic maize landraces evaluation may define patterns of diversity, which will
39 facilitate identifying alleles for enhancing yield and abiotic stress adaptation, thus
40 raising the productivity and stability of staple crops in vulnerable environments
41 (Dwivedi et al. 2016; Vasconcelos et al. 2013; Bellon & Hellin, 2011; Mercer et al.

42 2008). In other hand, regarding the evaluation tools, metabolomics has recently been
43 claimed as a promising concept and valuable tool in biotechnology, given its extensive
44 range of applications in functional genomics and, more globally, in the characterization
45 of biological systems (Costa et al. 2016). Indeed, the possible tasks include studying
46 metabolic systems, measuring biochemical phenotypes, understanding and
47 reconstructing networks and discriminating between samples (Costa et al. 2016). In
48 the targeted metabolomics approach, specific metabolites of known identity are

49 profiled; good quantitative precision is typically obtained. On the other hand,
50 untargeted metabolomics aims to simultaneously measure as many metabolites as
51 possible in a biological specimen (Yi et al. 2016).

52 Taking into account that maize landraces are currently being replaced by new
53 improved cultivars with a narrow genetic basis in Brazilian agriculture and with a
54 purpose to add value of those important genetic materials, the present study was
55 designed with the main goals of investigating the chemical composition of maize
56 landraces using non-targeted metabolomic approach (Fourier Transformed Infrared
57 Spectroscopy-FTIR) combined with chemometric tools (PCA and HCA) as a rapid
58 diagnostic tool to ascertain changes in chemical composition.

59 **2. Materials and methods**

60 *2.1 Maize landraces*

61 Maize landraces (*Zea mays* L.) were kindly provided by small farmers of the Anchieta
62 County (the western part of Santa Catarina state, southern Brazil, 26°31'11" S,
63 53°20'26" W) and produced under agro-ecological management practices during three
64 different harvests (2007, 2008, 2009—see Table 1 for details of the cultivars used each
65 year). Maize grains were collected, dried and crushed to obtain fine powder for later
66 analysis by Fourier transform infrared spectroscopy (FT-IR).

67 *2.2. Flour sample preparation*

68 Maize grains (250g – dry weight) were selected from each maize genotype and ground
69 to pass a 0.5mm sieve using a laboratory cyclone mill (MB Braeski C.Q).

70 *2.3 Fourier Transform infrared spectroscopy*

71 FTIR spectra of maize flours were recorded in a Bruker IFS-55 (Model Opus v. 5.0,
72 Bruker Biospin, Germany) spectrometer with a DTGS detector equipped with a golden
73 gate single reflection diamond attenuated total reflectance (ATR) accessory (45°
74 incidence-angle). A background spectrum of the clean crystal was acquired and
75 samples (100 mg) were spread and measured directly after pressing them on the
76 crystal. The spectra were recorded at the absorbance mode from 4000 to 500 cm⁻¹ at
77 the resolution of 4cm⁻¹. Five replicate spectra were collected for each sample. For pre-
78 processing, the spectra were normalized, baseline-corrected in the region of interest
79 by drawing a straight line before resolution enhancement (*k* factor of 1.7). The
80 assumed line shape was Lorentzian with a half width of 19 cm⁻¹ (Uarrota et al. 2013;
81 Kuhnen et al. 2010; Schuls & Baranska 2007; Baranska et al. 2006).

82

83 *2.4 Chemometric analysis*

84 FT-IR spectra were acquired and subjected to chemometric analysis using the R
85 software (R core team, 2017). Non-targeted metabolomics was used to find variations
86 in the chemical composition of different cultivars, cultivated under different harvests
87 and environmental conditions. FT-IR spectra were acquired in transmittance mode and
88 then transformed to absorbance mode, as spectra objects in the ChemoSpec package
89 (Hanson, 2014) using the equation 1.

90
$$\left[A = -\log_{10}(T) \right]$$

91 (Eq. 1)

92 where A is the absorbance and T the transmittance

93

94 The spectra object was then converted to a hyperSpec (Beleites & Sego, 2016) object
 95 using the bridging package hyperChemoBridge (McManus, 2014). Before applying
 96 chemometric tools, the spectra were baseline corrected, intensity vector normalized
 97 and smoothed for subsequent analysis of PCA and HCA (Costa et al. 2015) for sample
 98 classification according to their biochemical status.

99

100 **3. Results and discussion**

101 **3.1 Peak selection and multivariate analysis**

102 Results of the assignments for the most characteristic FTIR bands found in maize
 103 landraces and related chemical group are summarized in Table 2. The most important
 104 chemical groups found in all maize landrace samples were monoterpenes,
 105 sesquiterpenes, tetraterpenes, aminoacids, polysaccharides, lipids and proteins (Table
 106 2). The intensity of peaks is also shown in the Figure 1 (A-C) for maize samples from
 107 2007, 2008 and 2009 respectively. The selected peaks found in figure 1 were subjected
 108 for further multivariate analysis aiming to find similarities or differences between
 109 maize samples.

110

111 Principal component analysis of the 2007 maize landraces shown in the Figure 2A.
 112 Differences between cultivars were observed. Cultivars (11 and 12) and (7, 17)
 113 grouped together. The total variance explained by the first 2 PCs was 94.4%, being
 114 80.42% for PC1 and 13.98% for PC2. Cultivars 7 and 17 grouped together due to
 115 similarities in amylose, amylopectin and lutein contents. Cultivar 10 grouped alone due
 116 to its content in fatty acids.

117

PCA of 2008 samples is represented in the Figure 2B. PC1 explained 43.09% and PC2 39.81%, respectively. A good separation was found between PRO, MPA10 and LP0 cultivars. PRO was most correlated to lipids or fatty acids and LP0 correlated to amylose, amylopectin, monoterpenes and proteins. PCA on 2009 maize samples is represented in Figure 2C. PC1 explained 62.64% and PC2 (27.31%) of the total variance. The loading plot indicated that cultivars MG1, LP1, RX1 and RXE1 are similar in chemical composition due to their values of amylose, amylopectin and fatty acids. Cultivars R8C1 and PR1 showed to be similar due to their lower levels of proteins and lipids.

When FTIR data of 2007, 2008 and 2009 were subjected to hierarchical clustering (Figure 3A-C) aiming to find similarities or differences between samples and look for important wave bands related to sample clustering, an interesting result can be observed and the main bands responsible for sample differences were found. Differences in the 2007 dataset were mainly due to the region of carbohydrates (998, 1012 and 3000). Cultivar 15 grouped alone, and the most similar in chemical composition were cultivars 7 and 17, 11 and 12 respectively (Figure 3A). For samples of 2008 harvest, the region of proteins, amylose and amylopectin was most important in sample dissimilarity. PRO and MPA10 were the most dissimilar cultivars (Figure 3B). Two visible groups were found; being the first composed by LP0 and RXE0 and second group by RX0, MG0, RJO and R8C0. For maize landraces of 2009 (Figure 3C), proteins derived the sample clustering. MPA11 and LP1 were the most dissimilar cultivars confirming that observed previously in PCA analysis.

142 **Conclusions**

143 FT-IR technique coupled with chemometric tools (PCA, HCA) were capable in finding
 144 changes in chemical composition of maize landraces produced over different years.
 145 Similarities in maize samples were also found due to carbohydrate composition. The
 146 major bands in FT-IR are related to carbohydrates, lipids and proteins. Trace signals of
 147 secondary metabolites were also found in cultivars according to year of harvest.

148

149 **References**

150 Baranska, M., Schulz, H., Joubert, E. and Manley, M. (2006). In Situ Flavonoid Analysis
 151 by FT-Raman Spectroscopy: Identification, Distribution, and Quantification of
 152 Aspalathin in Green Rooibos (*Aspalathus linearis*). *Analytical Chemistry*, 78(22),
 153 pp.7716-7721.

154 Beleites, C.; Sergo, V. (2016) `hyperSpec: a package to handle hyperspectral data sets
 155 in R', R package version 0.98-20150304. <http://hyperspec.r-forge.r-project.org>

156 Bellon, M. and Hellin, J. (2011). Planting Hybrids, Keeping Landraces: Agricultural
 157 Modernization and Tradition Among Small-Scale Maize Farmers in Chiapas, Mexico.
 158 *World Development*, 39(8), pp.1434-1443.

159 Costa, C., Maraschin, M. and Rocha, M. (2016). An R package for the integrated
 160 analysis of metabolomics and spectral data. *Computer Methods and Programs in*
 161 *Biomedicine*, 129, pp.117-124.

- 162 Dwivedi, S., Ceccarelli, S., Blair, M., Upadhyaya, H., Are, A. and Ortiz, R. (2016).
163 Landrace Germplasm for Improving Yield and Abiotic Stress Adaptation. *Trends in Plant*
164 *Science*, 21(1), pp.31-42.
- 165 Hanson, B.A. (2016) ChemoSpec: Exploratory Chemometrics for Spectroscopy. R
166 package version 4.3.28, github.com/bryanhanson/ChemoSpec
- 167 Hellin, J., Bellon, M. and Hearne, S. (2014). Maize Landraces and Adaptation to Climate
168 Change in Mexico. *Journal of Crop Improvement*, 28(4), pp.484-501.
- 169 Kuhnen, S., Bernardi Ogliari, J., Dias, P., da Silva Santos, M., Ferreira, A., Bonham, C.,
170 Wood, K. and Maraschin, M. (2010). Metabolic Fingerprint of Brazilian Maize
171 Landraces Silk (Stigma/Styles) Using NMR Spectroscopy and Chemometric Methods. *J.*
172 *Agric. Food Chem.*, 58(4), pp.2194-2200.
- 173 McManus, C. (2014). hyperChemoBridge: Bridge between hyperSpec &ChemoSpec. R
174 package version 1.1.3. <https://github.com/Chathurga/HyperChemoBridge>
- 175 Mercer, K., Martínez-Vásquez, Á. and Perales, H. (2008). Asymmetrical local adaptation
176 of maize landraces along an altitudinal gradient. *Evolutionary Applications*, 1(3),
177 pp.489-500.
- 178 Mercer, K., Perales, H. and Wainwright, J. (2012). Climate change and the transgenic
179 adaptation strategy: Smallholder livelihoods, climate justice, and maize landraces in
180 Mexico. *Global Environmental Change*, 22(2), pp.495-504.
- 181 R Core Team (2016). R: A language and environment for statistical computing. R
182 Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- 183 Schauer, N. And Fernie, A. (2006). Plant metabolomics: towards biological function and
184 mechanism. *Trends in Plant Science*, 11(10), pp.508-516.
- 185 Schulz, H. and Baranska, M. (2007). Identification and quantification of valuable plant
186 substances by IR and Raman spectroscopy. *Vibrational Spectroscopy*, 43(1), pp.13-25.
- 187 Strable, J. and Scanlon, M. (2009). Maize (Zea mays): A Model Organism for Basic and
188 Applied Research in Plant Biology. *Cold Spring Harbor Protocols*, 2009(10),
189 pp.pdb.emo132-pdb.emo132.
- 190 Uarrota, V., Amante, E., Demiate, I., Vieira, F., Delgadillo, I. and Maraschin, M. (2013).
191 Physicochemical, thermal, and pasting properties of flours and starches of eight
192 Brazilian maize landraces (Zea mays L.). *Food Hydrocolloids*, 30(2), pp.614-624.
- 193 Vasconcelos, A., Bonatti, M., Schlindwein, S., D'Agostini, L., Homem, L. and Nelson, R.
194 (2013). Landraces as an adaptation strategy to climate change for smallholders in
195 Santa Catarina, Southern Brazil. *Land Use Policy*, 34, pp.250-254.
- 196 Yi, L., Dong, N., Yun, Y., Deng, B., Ren, D., Liu, S. and Liang, Y. (2016). Chemometric
197 methods in data processing of mass spectrometry-based metabolomics: A review.
198 *Analytica Chimica Acta*, 914, pp.17-34.
- 199

200 **Figure and table captions**

201

202 Figure 1: Main peaks found in FTIR spectra of (A) maize landraces produced in 2007,
203 (B) 2008, (C) maize landraces produced during 2009 harvest , respectively

204

205 Figure 2: Score and loading plots of Principal component analysis (PCA) of the whole
206 spectral range ($600-3000\text{ cm}^{-1}$) showing the sample clustering according to their
207 similarities. (A) Maze landraces produced in 2007 year, (B) 2008, (C) 2009 cultivars.

208

209 Figure 3: Hierarchical cluster dendrogram (HCA) of the main peaks found in the
210 spectra. (A) 2007 year, (B) 2008, (C) maize landraces produced during 2009 year.

211

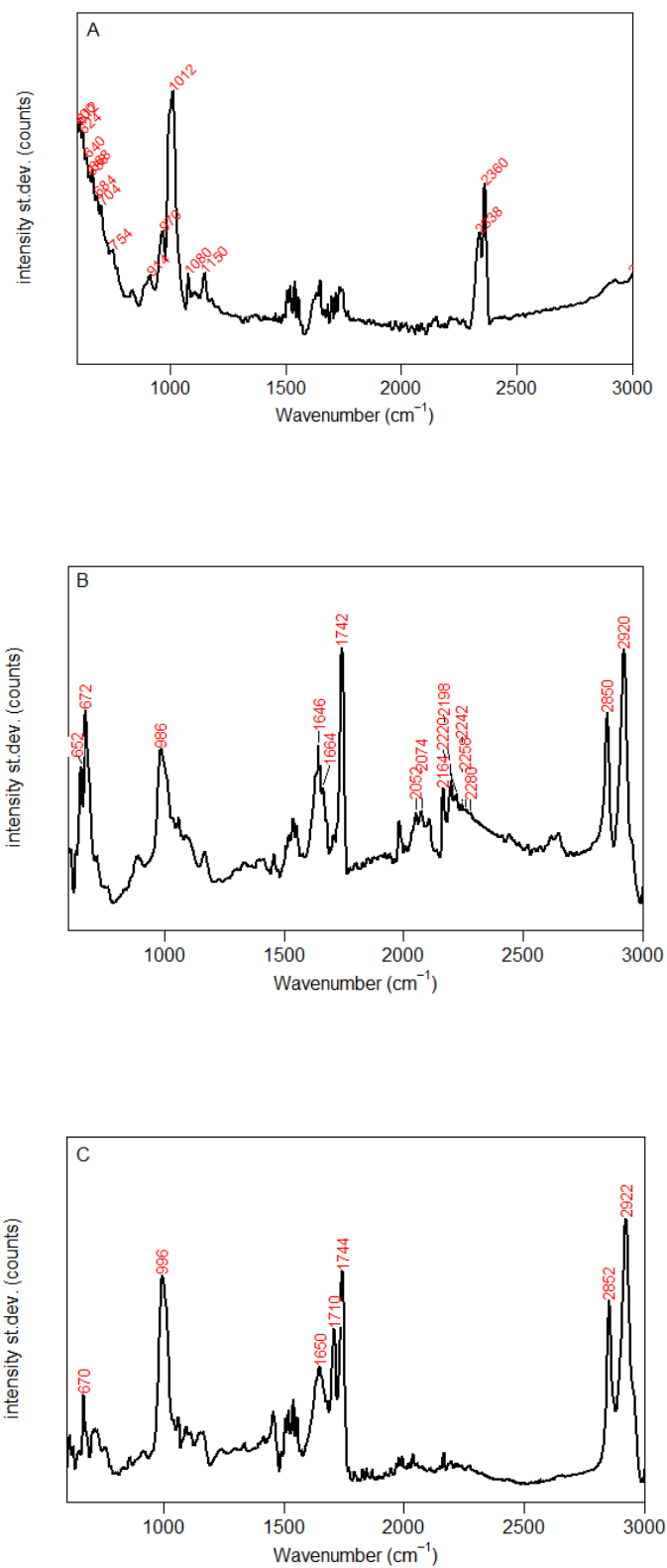
212 Table 1: Maize landrace cultivars studied in this research during the three years.

213

214 Table 2: Assignment for the most characteristic ATR-FTIR bands found in maize
215 landraces. Peaks and chemical group related is represented.

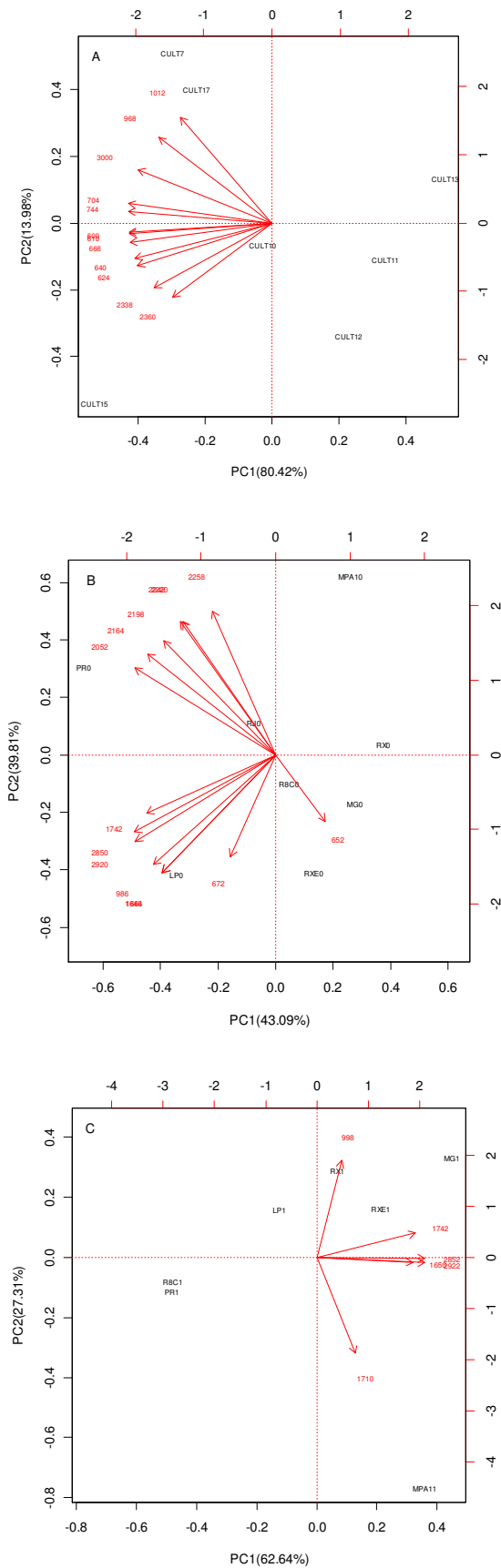
216

217 **Figure 1**

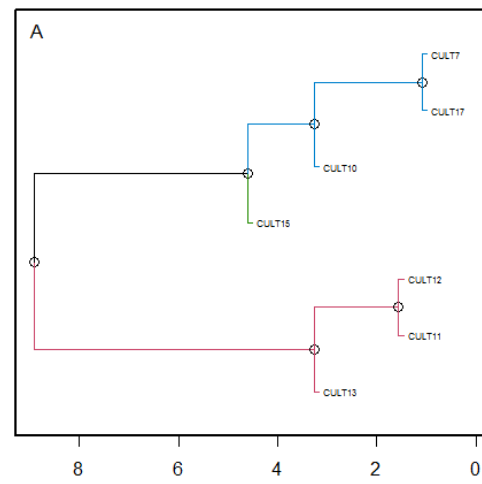


220
221
222

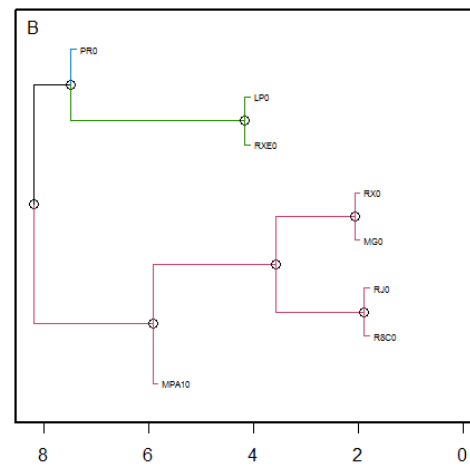
223 **Figure 2**



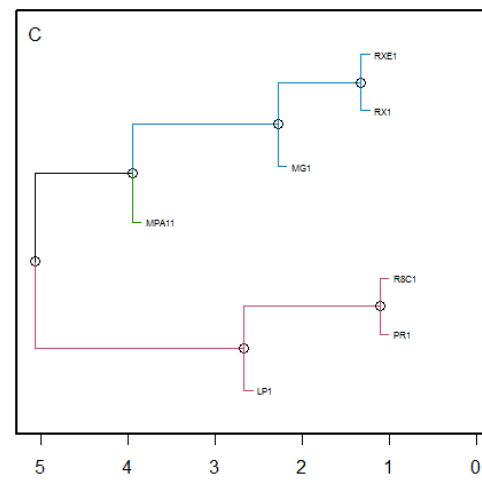
228 **Figure 3**



229



230



231

232 **Table 1**

Year	Maize cultivar analyzed	Acronym used
2007	Cultivar7	CULT7
	Cultivar10	CULT10
	Cultivar11	CULT11
	Cultivar12	CULT12
	Cultivar13	CULT13
	Cultivar15	CULT15
	Cultivar17	CULT17
2008	Lingua De Papagaio	LP0
	Mato Grosso-Palha Roxa	MG0
	Movimento De Pequenos Agricultores	MPA10
	Palha Roxa	PR0
	Rajado 8 Carreiras	R8C0
	Rajado	RJ0
	Roxo	RX0
	Roxo De Emílio	RXE0
2009	Lingua De Papagaio	LP1
	Mato Grosso-Palha Roxa	MG1
	Movimento De Pequenos Agricultores	MPA11
	Palha Roxa	PR1
	Rajado 8 Carreiras	R8C1
	Roxo	RX1
	Roxo De Emílio	RXE1

233 **Table 2**

Year of harvest	Chemical groups	Representative	ATR-IR wavenumber (1/cm)
2007	Monoterpenes	Lutein, beta-carotene	600
			610
			624
			640
			668
			684
			704
	Aminoacids	Methionine	744
2008	Polysacharides	Amylose, amylopectin	968
			998
			1012
			2338
			2360
			3000
	Lipids/fatty acids	Amylose, amylopectin	
2009	Monoterpenes	Amylose, amylopectin	652
			672
			986
			1646
			1664
			1742
			2052
	Lipids	Amylose, amylopectin	2164
			2198
			2220
			2242
			2258
			2850
			2920
2010	Polysacharides	Amylose, amylopectin	998
			1650
			1710
			1742
			2852
2011	Lipids	Amylose, amylopectin	2922

234