



UNIVERSITÀ DI PARMA

ARCHIVIO DELLA RICERCA

University of Parma Research Repository

Influence of fermentation level and geographical origin on cocoa bean oligopeptide pattern

This is a pre print version of the following article:

Original

Influence of fermentation level and geographical origin on cocoa bean oligopeptide pattern / Caligiani, Augusta; Marseglia, Angela; Prandi, Barbara; Palla, Gerardo Giovanni; Sforza, Stefano. - In: FOOD CHEMISTRY. - ISSN 0308-8146. - 211:(2016), pp. 431-439. [10.1016/j.foodchem.2016.05.072]

Availability:

This version is available at: 11381/2809018 since: 2021-11-15T15:24:05Z

Publisher:

Elsevier Ltd

Published

DOI:10.1016/j.foodchem.2016.05.072

Terms of use:

Anyone can freely access the full text of works made available as "Open Access". Works made available

Publisher copyright

note finali coverpage

(Article begins on next page)

18 April 2025

Manuscript Number:

Title: Influence of Fermentation Level and Geographical Origin on Cocoa Beans Oligopeptide Pattern

Article Type: Research Article (max 7,500 words)

Keywords: Cocoa beans; oligopeptides; vicilin; 21kDa cocoa seed protein; geographic origin; fermentation level

Corresponding Author: Dr. Augusta Caligiani, Ph.D.

Corresponding Author's Institution: University of Parma

First Author: Augusta Caligiani, Ph.D.

Order of Authors: Augusta Caligiani, Ph.D.; Angela Marseglia; Barbara Prandi; Gerardo Palla; Stefano Sforza

Abstract: Peptides and amino acids generated during cocoa bean fermentation are the most important precursors for the development of cocoa aroma. Although fermentation, amino acids and aroma development in cocoa have been extensively studied, the cocoa oligopeptide fraction is under-investigated and studies regarding the identification of peptides in different cocoa beans are completely lacking. In this study we perform a deep investigation on the presence of oligopeptides in well fermented cocoa beans coming from all the main producing countries, in unfermented (slaty) and underfermented (violet) beans. Oligopeptide pattern was determined by UPLC/ESI-MS, 35 low-molecular weight peptides (Mw range 202-1244) were identified, semiquantified and considered for data elaboration. Principal Component Analysis was used as a helpful tool for a better interpretation of dataset. The main factor influencing the peptide content was the fermentation level, however, a variability of peptide pattern was observed among well fermented cocoa samples of different geographic origin.

Dr. Augusta Caligiani
Dip. Food Science
Università degli Studi di Parma
Parco Area delle Scienze 59A
43124-Parma, Italy
Tel: +39 0521-905407
Fax: +39 0521 905472
E-mail: augusta.caligiani@unipr.it

June, 22 2015

Dear Editor of Food Chemistry,

we would submit the manuscript entitled “**Influence of Fermentation Level and Geographical Origin on Cocoa Beans Oligopeptide Pattern**” for the eventual publication in ‘Food Chemistry’ as a research article.

Autors: Augusta Caligiani, Angela Marseglia, Barbara Prandi, Gerardo Palla, Stefano Sforza

Recommended reviewers:

1- Jean Daniel Coisson, E-mail: jeandaniel.coisson@pharm.unipmn.it

2- Miguel Ángel Sentandreu, E-mail: ciesen@iata.csic.es

3- Tagro Simplicie Guehi, E.mail: g_tagro@hotmail.com

We would greatly appreciate Your suggestions about this work.

Thank You in advance,

Yours sincerely

Augusta Caligiani

*Highlights (for review)

- An overview on the peptide pattern of different cocoa samples is provided
- Peptide pattern and amount are strictly related to the fermentation process
- Ratio of peptides from vicilin and from cocoa albumin can be proposed as fermentation marker
- Some peptides are characteristic of specific geographic origin/varieties

1 **Influence of Fermentation Level and Geographical Origin on Cocoa Beans Oligopeptide**
2 **Pattern**

3 Augusta Caligiani*, Angela Marseglia, Barbara Prandi, Gerardo Palla, Stefano Sforza

4 Department of Food Science, University of Parma, Parma, Italy.

5 E-mail: augusta.caligiani@unipr.it

6 *Corresponding author:

7 Augusta Caligiani

8 Department of Food Science

9 University of Parma

10 Parco Area delle Scienze 59A

11 43124-Parma, Italy

12 tel: +39 0521 905407; fax: +39 0521 905472; E-mail: augusta.caligiani@unipr.it

13 **Short Title:** oligopeptides in cocoa beans

14

15

16

17

18

19

20

21

22

23

24

25

26 **Abstract**

27 Peptides and amino acids generated during cocoa bean fermentation are the most important
28 precursors for the development of cocoa aroma. Although fermentation, amino acids and aroma
29 development in cocoa have been extensively studied, the cocoa oligopeptide fraction is under-
30 investigated and studies regarding the identification of peptides in different cocoa beans are
31 completely lacking. In this study we perform a deep investigation on the presence of oligopeptides
32 in well fermented cocoa beans coming from all the main producing countries, in unfermented
33 (slaty) and underfermented (violet) beans. Oligopeptide pattern was determined by UPLC/ESI-MS,
34 35 low-molecular weight peptides (Mw range 202-1244) were identified, semiquantified and
35 considered for data elaboration. Principal Component Analysis was used as a helpful tool for a
36 better interpretation of dataset. The main factor influencing the peptide content was the
37 fermentation level, however, a variability of peptide pattern was observed among well fermented
38 cocoa samples of different geographic origin.

39

40

41 **Keywords**

42 Cocoa beans, oligopeptides, vicilin, 21kDa cocoa seed protein, geographic origin, fermentation
43 level

44

45

46

47

48

49

50

51

52

53

54 **1. Introduction**

55 Cocoa beans, from the fruit of the cocoa tree (*Theobroma cacao L.*), are transformed into chocolate
56 and other cocoa products by a complex process involving fermentation, drying and roasting. The
57 cocoa bean quality and characteristic aroma are determined primarily by genetic factors related to
58 cocoa tree variety, but agricultural practices and fermentation, both characteristic processes of the
59 origin country, can modify this genetically inherent flavour potential (Lopez & Dimick, 1995). The
60 two major botanical groups of cocoa are Criollo and Forastero. Criollo (native) represents only 5%
61 of world's production, corresponds to the subspecies 'cacao' and represents the original cocoa, it
62 ferments easily and has a pleasant and penetrating aroma, an optimum taste and therefore is
63 considered of excellent quality. The cacao Forastero (foreigner) resulted from dissemination toward
64 the Amazon Valley in Northern Brazil and to African countries, the most important of which are
65 now the Ivory Coast and Nigeria. These countries are now the largest producers of cocoa in the
66 world. Forastero represents about 85% of the world's cocoa production because it is very strong and
67 resistant to disease. The seeds have a strong flavor, not very aromatic and of lower quality. Some
68 Forastero plants growing in particular areas give cocoa of excellent quality (*cru*) such as Arriba, the
69 national cocoa of Ecuador (Caligiani, Marseglia & Palla, 2015).

70 Fermentation represents a crucial stage in the development of the aromatic precursors and bioactive
71 compounds characteristic of chocolate and cocoa products (Schwan & Wheals, 2004). Cocoa beans
72 are extracted from the fruit together with the surrounding sugary mucilaginous pulp, which is the
73 real substrate for fermentation. Fermentation usually lasts 5–7 days, except for some
74 varieties/hybrids, as for example Criollo and Arriba subjected to shorter fermentation (Fowler,
75 1999).

76 Cocoa fermentation is generally a spontaneous phenomenon which takes place without the purpose
77 addition of any starter bacteria. It is operated by a microbial succession of a wide range of yeasts

78 (Kloeckera and Saccharomyces spp.) and of lactic-acid and acetic-acid bacteria (Lactobacillus,
79 Bacillus, Pediococcus, Acetobacter and Gluconobacter), producing a wide range of metabolic end
80 products, in particular alcohol and organic acids (Schwan et al., 2004). The most important aspects
81 of fermentation, having fundamental consequences on cocoa quality, are the biochemical changes
82 inside the cocoa beans. Acids that are synthesized from pulp sugars move into the beans and lower
83 the internal pH. Acetic acid in particular, diffuses through the beans and causes a breakdown of the
84 cell membranes and the cell contents become mixed, allowing various enzymatic reactions to take
85 place. For example, oxidation of polyphenols and their conversion to insoluble forms due to the
86 reactions with proteins are responsible for the removal of the bitter taste and the change of bean
87 colour, from violet to brown (Jinap, Jamilah, & Nazamid, 2003). In fact, fully fermented cocoa
88 beans have brown colour, while unfermented and underfermented cocoa beans appear respectively
89 slaty and violet in colour, and represent the beans dried without previously being fermented or
90 partly fermented or by using improper procedures. For the manufacturer of chocolate or cocoa
91 powders the degree of fermentation of the beans is the principal quality criterion. In fact, too high
92 contents of unfermented or underfermented beans result in a lack of cocoa flavour in the end-
93 product because they do not develop enough chocolate flavour when roasted. The slaty beans cause
94 a very acid and astringent flavour profile, whereas the violet beans cause a bitter and harsh flavour
95 (Kattenberg & Kemmink, 1993; Puziah, Jinap, Sharifah & Asbi, 1998).

96 During fermentation, another fundamental biochemical change occurs on the protein fraction.
97 Proteins make up 10-15% of the dry weight of cocoa seeds, the second most abundant constituent
98 after cocoa fat. Cocoa beans contain two main well-characterized proteins fractions, albumin and
99 globulin. Globulins are vicilin-like storage proteins mostly consisting of three subunits with
100 molecular masses of 47 kDa, 31 kDa, and 15 kDa, which come from a common 66-kDa precursor
101 (Spencer & Hodge, 1992). On the other side, the albumin fraction was mostly identified as a 21kDa
102 cocoa seed protein; its primary structure, together with its trypsin inhibitory properties, was
103 reported by Kochhar, Gartenmann & Juillerat (2000). Thus, basically only two proteins constitute

104 most of the protein fractions of cocoa. During fermentation, these proteins are cleaved to
105 hydrophilic and hydrophobic peptides as well as amino acids through autolysis by two endogenous
106 enzymes, aspartic endoprotease and carboxypeptidase (Amin, Jinap, Jamilah, Harikrisna, & Biehl,
107 2002; Voigt, Biehl, Heinrichs, Kamaruddin, Marsoner, & Hugi, 1994a), which are activated by
108 microbial metabolites (such as acetic acid) produced during fermentation. Aspartic endoprotease
109 was demonstrated active both at pH 5.2 or pH 3.5, while carboxypeptidase esplicates proteolytic
110 activity at pH > 5 (Voigt, Heinrichs, Voigt & Biel, 1994b), so the pH reached during fermentation
111 and the development of organic acids are key factors controlling the proteolysis. Globulin protein
112 fraction is the most degraded during fermentation and amino acids and oligopeptides formed
113 represent the essential precursors for the development of cocoa aroma (Amin, Jinap, & Jamilah,
114 1997; Voigt, Biehl, Kamaruddin, & Wazir, 1993).

115 Although the degradation of cocoa proteins into oligopeptides and amino acids is well documented,
116 the specific sequences of peptides formed during cocoa fermentation and autolysis and their role in
117 the development of cocoa flavour, are poorly characterized. Buyukpamukcu, Goodall, Hansen,
118 Keely, Kochhar, & Wille (2001) report the identification of peptides APLSPGDVF and SPGDVF
119 from vicilin, while other authors report HPLC profiles of complex peptide patterns without specific
120 identification and quantification (Rashidah, Jinap, Nazamid, & Jamilah, 2007; Amin et al., 2002;
121 Jinap, Nazamid & Jamilah, 2002; Jinap, Ikrawan, Bakar, Saari & Lioe, 2008). Recently we
122 performed a study aimed at characterizing the oligopeptide fraction in cocoa and we assigned 44
123 peptides based on their exact molecular mass, mass fragmentation pattern and comparison with
124 vicilin and 21 kDa albumin sequences (Marseglia, Sforza, Faccini, Bencivenni, Palla & Caligiani,
125 2014). Nevertheless, so far no comprehensive study focusing on the variation of the content and
126 distribution of oligopeptides of the commodity fermented cocoa beans has been made. In the
127 present work, we provide a comprehensive study on the content and pattern of oligopeptides in
128 commercially fermented and dried cocoa beans from different geographical origins and

129 fermentation levels. Results are discussed in respect to fermentation degree and origin specificity.
130 This study represents the first exhaustive dataset on peptide pattern in cocoa.

131

132

133 **2. Materials and Methods**

134 **2.1 Cocoa samples.**

135 We investigated 48 cocoa beans samples from 22 different geographical origins kindly provided by
136 Barry Callebaut (Lebeke-Wieze, Belgium). The samples analysed were of Forastero varieties,
137 fermented and dried. All fruits were harvested in 2012 and considered well fermented (brown
138 color). Countries of origin and numbers of lots of cocoa beans samples collected are the following:
139 Cuba (1), Mexico (1), Brazil (1), Grenada (1), Perù (6), Dominican Republic (3), Trinidad (1),
140 Venezuela (3), Indonesia (2), Ecuador (3), Sao Thome (4), Ivory Coast (5), Madagascar (1), Sierra
141 Leone (1), Nigeria (1), Tanzania (2), Uganda (1), Ghana (5), Congo (1), Java (1), Malaysia (1),
142 Papua New Guinea (2). A sample of Criollo variety from Mexico was also analysed.

143 In addition, Forastero cocoa beans with different fermentation level, determined by the visual
144 inspection of internal cocoa bean color (underfermented, violet color, 8 samples; unfermented, slaty
145 color, 5 samples) of the following origins were also considered: Ghana, Ecuador, Ivory Coast,
146 Grenada.

147 All the collected samples were stored frozen. For chemical analysis, the samples were treated with
148 liquid nitrogen and ground with a coffee mill.

149 **2.2 Chemicals.** All solvents and reagents were of HPLC grade and used as commercially available
150 without any further purification; deionized water was obtained by Millipore Alpha Q system
151 (Millipore Corporation, Billerica, MA); formic acid 99% was purchased from ACROS Organics
152 (Fair Lawn, NJ).

153 **2.3 Sample Preparation.** Peptides were extracted according to the method previously described
154 (Marseglia et al., 2014). A total of 10 g of finely grinded cocoa sample was suspended in 45 mL of

155 0.1 N HCl. (L,L)-phenylalanylphenylalanine (Phe-Phe) was added as an internal standard (2.25 mL
156 of a 1mM solution). The suspension was homogeneized for 1.5 min by Ultra Turrax T50 at
157 4000rpm (Janke and Hunkel Italabortechnik) and then centrifuged at 4000 rpm for 30 min at 4 °C
158 by an ALC 4237R centrifuge. The solution was filtered through paper filters (pore dimensions 15-
159 20 µm) and then extracted four times with 50 mL of ethyl ether. The solution was filtered again
160 with a Millipore 47 mm Steril Aseptic system through 0.45 µm HVLP Millipore filters. A total of
161 1.5 mL of the resulting solution were mixed with 0.5 mL of a formic acid solution (0.1%). The
162 solution was diafiltered through Sartorius Vivaspin 2 filters (nominal molecular cut-off 10 000 Da)
163 by using an Amicon Micropartition system MPS-1. The filtrate was dried under nitrogen,
164 redissolved in H₂O (0.1% HCOOH), and analyzed by HPLC.

165 **2.4 UPLC/ESI MS (WATERS, Milford, MA) analysis conditions:** eluent A: H₂O (0.2% CH₃CN
166 and 0.1% HCOOH); eluent B CH₃CN (0.1% HCOOH); gradient elution was performed according
167 to the following steps: 0–7 min isocratic 100% A, 7–50 min linear gradient from 100% A to 50% A,
168 50–52 min isocratic 50% A, 53– 58 min from 50% A to 0% A and reconditioning. Column:
169 AQUITY UPLC Beh C18 (1.7 mm, 2.1x150 mm). Flow rate: 0.2 ml/min. MS conditions: ESI,
170 positive ions, single quadrupole analyzer. Capillary voltage: 3.2 kV, cone voltage: 30V, source
171 temperature: 150 °C, desolvation temperature: 300 °C, cone gas flow (N₂): 100 L/h, desolvation gas
172 (N₂): 650 L/h, acquisition: 100:2000 m/z. All data were acquired and processed by the software
173 MassLynx 4.0 (Waters, Milford, MA).

174 **2.5 Peptides semi-quantification and statistical analysis**

175 Peptides were semi-quantified by comparison to the internal standard Phe-Phe. For the correct
176 integration of peaks, as the total ion chromatograms (TIC) were crowded, the eXtract Ion
177 Chromatogram (XIC) technique was applied. The oligopeptides were semi-quantified by measuring
178 the ratio between the XIC peptide area and the relative XIC area of Phe-Phe, as previously
179 described (Sforza, Galaverna, Schivazappa, Marchelli, Dossena, & Virgili, 2006), and expressed as
180 mg/kg of cocoa beans. Data are semi-quantitative because do not take into account the different

181 response factors, due to the unavailability of pure compounds for peptides. Data are provided for
182 inter-samples comparison purposes only and they represent the mean of two independent sample
183 extractions and analyses.

184 An unsupervised learning approach, Principal component analysis (PCA) was used to describe the
185 variation in the dataset. PCA was performed with IBM SPSS statistics (Version 22.0, IBM, New
186 York, USA), based on the correlation matrix, which is equivalent to using the covariance matrix on
187 mean centred and variance scaled variables.

188 **2.6. Determination of organic acids**

189 Lactic, acetic, citric and succinic acids were determined on cocoa aqueous extracts by quantitative
190 ^1H NMR , as previously reported (Caligiani, Acquotti, Cirilini & Palla, 2010)

191 **3. Results and Discussion**

192 Typical MS chromatograms (full scan acquisition) of cocoa bean samples of different origins and
193 fermentation levels are shown in Figure 1. The major differences in the peptide patterns were
194 observed among samples of different fermentation levels, but also some differences in samples from
195 different origins could be appreciated.

196 In our previous study (Marseglia et al. 2014), 44 peptides were identified , many of them belonging
197 to vicilin and to 21 kDa cocoa seed albumin. In the present study, we focused on 35 peptides,
198 among the ones previously identified, showing the most intense signals in our samples. The list of
199 peptides considered, ordinated on the basis of their molecular weight, is reported in Table 1 together
200 with the maximum, minimum and mean values. From the mean values reported in Table 1 it
201 appears well evident that small peptides are quantitatively the most abundant, indicating a general
202 extensive proteolytic activity during cocoa fermentation.

203 **3.1. Total peptide amounts and ratio vicilin/21kDa albumin peptides.**

204 In Figure 2a, the total content of peptides in each cocoa bean sample is reported, considering the
205 sum of all the peptides. Among well fermented beans, results showed that total peptide amount is
206 extremely variable in cocoa from different geographic origin, and the samples with lower peptide

207 content are those from Indonesia. As expected, the other samples containing low amount of peptides
208 are unfermented, slaty beans. Quite surprisingly, supposedly underfermented violet beans contain
209 total peptide amounts comparable with well fermented beans.

210 Well fermented cocoa samples of different geographic origins can be divided according to their
211 peptide content. Low amounts of oligopeptides, in the range of about 50-150 mg kg⁻¹ were found in
212 commercial samples from Indonesia that contained values between 43 and 54 mg kg⁻¹, Ecuador
213 (117-121 mg kg⁻¹), Ivory Coast (134-205 mg kg⁻¹), Madagascar (130 mg kg⁻¹), Malaysia (117 mg
214 kg⁻¹), Nigeria (106 mg kg⁻¹), Tanzania (134-160 mg kg⁻¹), Uganda (142 mg kg⁻¹), Venezuela (117-
215 165 mg kg⁻¹). Medium amounts of peptides, from 150 to 250 mg kg⁻¹, were observed in cocoa
216 from Brazil (198 mg kg⁻¹), Ghana (173-211 mg kg⁻¹), Perù (109-225 mg kg⁻¹). High level of
217 oligopeptides (superior to 250 mg kg⁻¹) were showed by all cocoa samples from Central
218 America/Carribbean Islands as Mexico, Cuba, Grenada, Santo Domingo, Trinidad, together with
219 African samples from Congo and Sao Thome, Papua New Guinea and also Criollo variety from
220 Mexico.

221 Intra-country specific variations of total peptides were comparatively low in Ecuador (117-121 mg
222 kg⁻¹), Ghana (173-211 mg kg⁻¹) Papua New Guinea (250-294 mg kg⁻¹) and Santo Domingo (314-
223 324 mg kg⁻¹), whereas large variations were found in cocoa from Perù (109-225 mg kg⁻¹).

224
225 Also the sum of the peptides resulting from vicilin and those arising from 21kda was considered,
226 excluding dipeptides because in most cases they were common to both proteins. The ratio
227 vicilin/21kDa peptides, shown in Figure 2b, is generally greater than 1, indicating a quantitative
228 prevalence of peptides from vicilin. The ratio of peptide amounts generated from vicilin and 21kDa
229 cocoa seed protein was variable and in general samples showing low content of peptides (less
230 fermented samples) have a higher ratio, indicating that peptides from vicilin are formed first in the
231 initial steps of fermentation. In fact, a negative correlation (correlation coefficient = -0.76, p<0.05)
232 was observed between total peptide amounts and the ratio of peptides from vicilin and 21 kDa

233 cocoa seed albumin. Generally, it was found that fermented cocoas can be differentiated from
234 unfermented cocoas by the higher total amount of oligopeptides and by the lower ratio of vicilin to
235 21 kDa cocoa seed albumin. So, in the case of genetically identical material, this ratio can be
236 proposed as a suitable indicator for the fermentation level.

237 **3.2. Exploration of cocoa peptide patterns by Principal Component Analysis (PCA)**

238 Semiquantitative amounts of the 35 peptides considered for each cocoa sample are reported in
239 supplementary materials (S1). The organic acids amounts were also calculated for each sample and
240 reported in S1. In order to obtain an easily interpretable quali-quantitative description of the
241 eventual differences present in the peptide pattern of cocoa beans, a PCA was performed on all the
242 bean samples, including all geographic origins and fermentation levels. PCA was performed on the
243 matrix of 35 peptides for 61 cocoa samples, and the principal components were obtained with the
244 correlation method. Figure 3 reports the 3D scatter plot of the scores of PC1 versus PC2 and PC3,
245 explaining more than 70 % of the total system variance, together with the corresponding
246 combinations of PCs in bidimensional plots. The 3D score plot (Figure 3a) shows that samples of
247 cocoa beans are divided into three principal groups, according to the fermentation level; the main
248 group corresponds to the fermented beans of different geographical origins that were considered
249 well fermented by supplier. Unfermented and underfermented beans (slaty and violet color,
250 respectively) are separated each other and from the group of brown, well fermented beans.

251 The loading values of the variables associated to the first three principal components are reported in
252 supplementary material (S2). PC2 appears as the component that better separates the groups of slaty
253 and violet beans from brown beans (Figure 3b and 3d). The variables with positive values on PC2,
254 that characterize the group of slaty and violet beans, are peptides of higher molecular weight as
255 APLSPGDVF, DEEGNFKIL, Mw1596, DNEWAW, NGTPVIF, while the variables with negative
256 values on PC2, related to the well fermented beans, are low molecular weight peptides, mainly
257 dipeptides as AI, FL(I), AW, FV, VF. This is consistent with the fact that during the course of

258 fermentation an intense proteolytic activity occurs inside the beans degrading longer peptides to
259 shorter ones (Amin, Jinap & Jamilah., 1998).

260 Slaty beans are well separated from all the other beans based on PC1 (figure 3b and 3c), having the
261 most negative values on this component. The corresponding variables are Mw1244, Mw1596,
262 Mw1090, VLE and TVWRLD. They correspond, except VLE, to high molecular weight peptides.
263 Results of PCA are confirmed in Figure 4 where the general mean (mg/kg) of each peptide in
264 brown, violet and slaty beans was reported. Slaty beans are very different from the others because
265 they are characterized by low amount of almost all peptides except for VLE, TVWRLD and two
266 high Mw peptides (1090, 1244). Brown and violet beans are more similar regarding total peptide
267 content, but generally brown beans contain higher amount of dipeptides, while intermediate Mw
268 peptides as GAGGGGL, KDQPL, SPGDVF characterize the group of violet beans.

269

270 Focusing only on well fermented cocoa, no specific grouping was observed according to
271 geographical origin, and generally also in the samples from the same country differences can be
272 observed. Only samples from Peru, Ghana and Ivory Coast showed a larger intra-country
273 similitudine. Sample from Congo seems to be very different from the others. Comparing samples on
274 the left side of PC1-PC2 score plot (Figure 3b) with the samples on the right side, indicated that all
275 the samples on the left, near the group of slaty beans and corresponding to negative values on PC1
276 and positive on PC2, have low amount of peptides (compare figure 2a): for example Nigeria L6
277 (106,48 mg/kg), Tanzania L3 (134,43 mg/kg), Ecuador L3 (116,92mg/kg).

278 On the contrary, the group of sample on the right, corresponding to positive values on PC1 and
279 negative on PC2, have high amount of peptides: Congo (312,20 mg/kg), Criollo (301,29), Papua
280 New Guinea (294,75 mg/kg), S. Domingo L2 (324,12), Sao Thome (302,82 mg/kg), Trinidad
281 (352,50 mg/kg). So it is possible to conclude that moving from negative to positive values on PC1
282 and from positive to negative on PC2, the degree of fermentation is increasing. As a consequence,
283 peptide pattern and amounts can be considered a very good index of cocoa fermentation.

284 Samples from Indonesia are in the same group of the slaty beans and showed the same peptide
285 pattern, indicating that they were poorly fermented. They showed also low amount of total peptides
286 and low amount of free amino acids, as previously reported (Marseglia, Palla & Caligiani, 2014b).
287 These data are consistent with the fact that cocoa fermentation in Indonesia is short or even omitted
288 in Sulawesi regions and Indonesian cocoa is of poor commercial value (Rohsius, Matissek, &
289 Lieberei, 2006).

290 Ecuador samples in the score plot is near the group of slaty beans, and this is compatible with the
291 shorter fermentation times utilized for this cocoa. Ecuador samples showed also low total amount
292 of peptides and high ratios vicilin/21 kDa peptides.

293 The specific distribution of peptides in some cocoa samples containing respectively low, medium
294 and high amount of peptides was also considered (Figures 5a, 5b and 5c). High-peptide samples
295 show a more uniform distribution of peptides, indicating that for a complete fermentation the
296 peptide profile tend to be similar without specific dependance on cocoa origin. More differences in
297 peptide pattern can be observed in medium- and, even more, low-peptide samples, indicating that
298 the peptide profile is linked to the moment in which fermentation stops. This characteristic makes
299 possible, in most cases, to find specific peptides markers of a particular cocoa origin. For example
300 samples from Ecuador, considered a fine cocoa as Criollo, contain, respect to the other Forastero
301 samples, the highest amount of peptide ASKDQPL and of peptide with Mw 1090. The same
302 peptides characterize Nigeria sample together with Mw1596. Malaysia cocoa is mainly
303 characterized by TVWRLD, IEF and DEEGNFKIL and the last two peptides are high also in Brazil
304 cocoa. APLSPGDVF and SPGDVF are most represented in Ghana samples.

305 Sample of Criollo variety (origin Mexico) is separated from all the others probably because it
306 represent another variety; it occupies the PCs score plot zone associated with well fermented cocoa,
307 even if Criollo fermentation is generally short. Criollo separation is mainly based on PC3, where it
308 shows the highest negative values (figure 3c and 3d); the associate variables are peptides of
309 different molecular weight as GAGGGGL, RLD, VI, RRSDDL, EVL. All these peptides, except

310 EVL, derive from 21 kDa albumin. This is confirmed by the specific peptide pattern of criollo
311 compared with the forastero variety of the same region (figure 5d).

312 In order to better understand the proteolytic activity giving origin to the different oligopeptides
313 pattern observed, a correlation with content of organic acids in cocoa beans was performed. Organic
314 acids (reported in supplementary material) showed slight positive correlations (correlation
315 coefficient 0.6-0.7, data not shown) with cocoa dipeptides. In general, citric acid level is higher in
316 samples having low amount of peptides, while lactic acid and acetic acid are generally more
317 abundant in samples with higher amounts of peptides, as for example Congo, Criollo and Sao
318 Thome. Criollo variety, having a peptide pattern different from other cocoas, showed similar
319 amounts of organic acids.

320 **4. Discussion**

321 Many parameters affect the composition of the fermented cocoa beans: variety, geographical origin,
322 climate during ripening, degree of ripeness, and the processing steps of fermentation and drying. A
323 better knowledge of the biochemical reactions occurring during fermentation, aimed at a better
324 control of cocoa primary processing, is a strategic and challenging issue for cocoa industry. In this
325 context, the knowledge of the structure and cinetic of formation of oligopeptides is very important
326 because proteolysis is the main reaction involved in cocoa aroma formation and, as a consequence,
327 in cocoa quality. The contribute of peptides in generating typical cocoa aroma is currently unknown
328 but it is underlined by a study showing that no cocoa aroma was obtained when synthetic mixtures
329 of amino acids, resembling the spectrum of free amino acids present in fermented cocoa seeds, were
330 roasted in the presence of reducing sugars (Voigt et al., 1994a).

331 The results in this work clearly showed that peptide amount and profile are strongly influenced by
332 the cocoa fermentation level, which also made possible to appreciate origin-related differences.

333 PCA demonstrates that the main grouping factor is the fermentation level and this indicates that
334 fermentation influences the cocoa bean composition more than botanical or geographical factors do.

335 An effect of the botanical variety (Criollo vs Forastero) has also been registered, probably because a
336 different cocoa chemical composition influence the modification occurring during fermentation.

337 Regarding cocoa samples with different fermentation level, lower amounts of peptides were
338 observed in slaty beans, as expected, while similar contents were registered in violet and brown
339 beans, with a slight higher amount of peptides in violet beans. This was a trend observed in all the
340 series of cocoa beans and might be explained considering that progressing fermentation peptides are
341 further hydrolyzed to amino acids. However, in a previous work, we found that also amino acids are
342 higher in violet than in brown beans (Caligiani, Palla, Acquotti, Marseglia & Palla, 2014),
343 indicating that probably the proteolytic activity is higher in the first days of fermentation and then
344 tends to decrease. Biehl & Passern (1982) and Jinap et al. (2008), followed the development of free
345 amino acids during cocoa bean fermentation, reporting that different amino acids reached the
346 maximum value after different days of fermentation. The high content of peptides and amino acids
347 in violet beans suggests that violet beans are able to generate specific cocoa aroma during roasting,
348 but according to Hii et al. (Hii, Law, Cloke & Suzannah, 2009), due to their higher polyphenols
349 content, they probably cause excessive astringency to the final products that masks the chocolate
350 flavour.

351 Variable amounts of peptides were observed not only in samples with known different fermentation
352 levels, but also in well fermented samples of different geographical origin. In general, different
353 amounts of peptides correspond to different ratios of vicilin/21 kDa peptides and often to different
354 peptides distribution, indicating that sample-specific mechanism/kinetic of proteolysis occur.

355 Although clearly due to the differences in the fermentation processes and thus on the enzymatic
356 activity brought in by the different microorganisms, the reasons for such differences in peptide
357 content and composition remain largely unclear, since commercial samples usually lack information
358 on fermentation and drying practices as well as planting material used. But, even though sample-
359 specific information is lacking, overall information on the common fermentation practices in use is
360 available for many countries and regions. For example, countries and regions like Ecuador,

361 Tanzania and Sulawesi are known for traditionally short or, in the case of Sulawesi, often omitted
362 fermentation processes. Malaysia normally makes use of beans imported from Indonesia that are
363 usually blend with fully fermented beans, to obtain the desired flavour characteristics and to reduce
364 the excessive astringency and bitterness (Puziah et al., 1998). Ecuadorian cocoa is usually poorly
365 fermented: traditional fermentations of the local Ecuadorian cocoa type Nacional, with its fine
366 flavor, are carried out in boxes and on platforms for a short time (Beckett, 2009; Papalexandratou,
367 Falony, Romanens, Jimenez, Amores, Daniel & De Vuyst, 2011). Accordingly, all the samples
368 analysed from these countries contained low amounts of oligopeptides.

369 African countries are known to produce bulk cocoa with an average quality through long
370 fermentation times, typically 3 to 6 days in heaps (Rohsius et al., 2006; Wood & Lass, 1993). High
371 amounts of oligopeptides were found in Congo and Sao Thome samples but not in the other West
372 African countries as Ivory Coast, Ghana, Nigeria. In Ivory Coast, cocoa beans were commonly
373 fermented 4 - 5 days without turning, and fermented cocoa beans were dried by solar drying method
374 (Guehi, Konan, Koffi-Nevry, N'Dri, & Manizan, 2007). Turning ensures uniform fermentation and
375 frequent mixing, at 6–12 hours intervals, produces a higher number of well-fermented beans
376 (Senanayake. Jansz & Buckle, 1997).

377 In the case of selected cocoa varieties as Criollo, even if fermentation is short, oligopeptide content
378 is one of the highest, so it is possible that the acidification and the induction of proteolysis occurs
379 earlier. Criollo contains more peptides arising from 21 kDa cocoa albumin, respect to all the other
380 cocoa analysed. However, this protein should be more resistant to hydrolysis, since it has been
381 reported that degradation of vicilin-class globulin was about 88% while that of albumin was 47% at
382 the end of fermentation (Amin at al., 1998). Moreover, albumin does not originate specific cocoa
383 aroma, but Criollo is one of the most aromatic cocoa (Voigt et al.,1994b). A previous work showed
384 that no differences were present in HPLC oligopeptide profile obtained by the action of aspartic
385 endoprotease on partially purified vicilin from various cocoa genotypes, comprising Criollo, but no
386 data about peptides from albumin were provided (Amin et al., 2002). It was reported that proteolytic

387 activity in cocoa is higher immediately after the bean death (1-2 days of fermentation) and then
388 tends to decrease (Biehl et al., 1982).

389 It is well known that formation and degradation of peptides in cocoa is related to the action of two
390 endogenous enzymes, a carboxypeptidase and an aspartic endoprotease, both activated by organic
391 acids of fermentative origin that penetrate cocoa beans lowering the inside pH.

392 Aspartic endoprotease is able to release peptides but only very low amounts of free amino acids,
393 which were instead formed when globulin peptides were treated with carboxypeptidase, releasing
394 preferentially hydrophobic amino acids (Voigt et al 1994b). The predominant amino acids released
395 from the albumin-derived oligopeptides by carboxypeptidase treatment were instead acidic amino
396 acids such as aspartic acid, glutamic acid and asparagine. Aspartic endoprotease was demonstrated
397 active both at pH 5.2 or pH 3.5, while carboxypeptidase esplicates proteolytic activity at pH > 5
398 (Voigt et al 1994b). All these data indicates that different peptides patterns can depend on the pH
399 value reached at different steps of fermentation. Amin et al. (1998) showed that the pH during
400 fermentation determines the rate of aspartic endoproteinase and carboxypeptidase activity. The pH
401 range of 5.0-5.5 during fermentation is thus important in producing fermented and dried cocoa
402 beans with high aroma potential, whereas a pH range of 4.0-4.5 will result in low aroma potential
403 (Biehl, Brunner, Passern, Quesnel & Adomako, 1985).

404 Regarding proteases specificity, it has been found that the cleavage specificity of the cocoa aspartic
405 endoprotease was essential for the formation of the cocoa-specific aroma precursors (Voigt, Voigt,
406 Heinrichs, Wrann, & Biehl, 1994c) while cocoa carboxypeptidases can be substituted by other
407 carboxypeptidases as porcine pancreas carboxypeptidase A (Bytof, Biehl, Heinrichs & Voigt,
408 1995).

409 The preferred cleavage sites for cocoa aspartic endoprotease is aspartic acid N-terminal (Biehl et al.,
410 1982), and it cleaves preferentially protein substrates at hydrophobic amino acids residues, although
411 its exact cleavage specificity has not yet been characterized (Watson, Preedy & Zibadi, 2012).
412 Bytof et al (1995) showed that the specificity of carboxypeptidase is influenced by the pH-value,

413 and liberation of acidic amino acids was favoured by low pH-values, but in general peptides with
414 carboxyterminal arginine, lysine or proline residues are resistant against degradation by the cocoa
415 seed carboxypeptidase. The enzyme preferentially liberates hydrophobic amino acids, whereas
416 acidic amino acids are released very slowly. The rate of hydrolysis is not only determined by the
417 carboxyterminal, but is also affected by the neighboring amino acid residue.

418 Amin et al. (1998) showed that the aspartic endoproteinase levels fell sharply in the first 2 days of
419 fermentation and then increased to 116% of the initial value at 3 days, falling to zero by the end of
420 the fermentation (6 days). The carboxypeptidase activity fell to approximately one third of the
421 initial value during the first 3 days of fermentation and then rose to 157% of the initial value at 4
422 days and fell to zero by the end of the fermentation (6 days). The changes in both enzyme activities
423 throughout the fermentation could be due to pH changes in the cotyledon during fermentation, with
424 optimum pH values for these enzymes occurring around 72 h for the aspartic endoproteinase and 96
425 h for the carboxypeptidase. Therefore, the proteolytic activity is higher immediately after the bean
426 death and then tends to decrease, probably due to inhibition by the oxidation compounds of
427 polyhydroxyphenol products which were released from storage cells (Forsyth, 1958; Biehl et al.,
428 1982).

429 Jinap et al. (2008) followed the development of free amino acids during cocoa bean fermentation,
430 from 0 days (unfermented) to 3 days (underfermented), reporting that some hydrophobic amino
431 acids reached the maximum value after 3 days of fermentation (Tyr, Ile), and others after 2 days
432 (Ala, Leu) or 1 day (Val, Phe). It is possible to infer that a similar behaviour is followed by
433 peptides, and the specific pattern observed in the final cocoa sample is a sort of 'snapshot' of the
434 moment in which fermentation and proteolytic activity is stopped. Further work is in progress to
435 confirm these hypothesis by analysing cocoa beans sampled at different days of fermentation

436 Thus, all the literature data demonstrate that pH values reached during fermentation and their
437 variations are key factors for the activation of proteolytic enzymes and as a consequence for the
438 development of specific peptides and amino acids. The variations in pH of fermented cocoa are the

439 results of different factors: the loss of citric acid concentration in the pulp (Ardhana & Fleet, 2003),
440 the migration of lactic acid, acetic acid and other many organics acids produced by microbial
441 activities from the outside to the inside of cocoa seeds, the loss of volatile acetic acid. Some authors
442 (Guehi, Dadie, Koffi, Dabonne, Ban-Koffi, Kedjebo & Nemlin, 2010) demonstrate that the acidity
443 of fermented cocoa beans varied on the duration and on the method of fermentation. Acidity of
444 cocoa decreased on fermentation duration and partial fermented cocoa beans presented higher
445 acidity than well-fermented, because of volatility of acetic acid. Since the mechanism of
446 fermentation of cocoa beans is very complex in nature, and factors such as temperature mass and
447 aeration affecting the pH during fermentation were not controlled, the correlation between protease
448 activities, fermentation level, peptide pattern with pH and organic acids were not easy to be
449 interpreted. In the present paper, only slight positive correlations were observed among lactic acid
450 and acetic acid with cocoa dipeptides and they are generally more abundant in samples with higher
451 amounts of peptides, while citric acid level is higher in samples having low amount of peptides, a
452 further confirmation that they were poorly fermented. However, the relationship among the
453 production of organic acids during fermentation and the activation of cocoa proteases remain
454 unclear, for example Criollo variety, having a peptide pattern different from other cocoas, showed
455 similar amounts of organic acids.

456 Another important point that needs to be clarified is the reason why peptides derive only from
457 specific zones of vicilin and 21 kDa cocoa albumin and are not spread in all proteins sequences
458 (Marseglia et al., 2014). This phenomenon is not completely unknown in fermented products, since,
459 for example, only peptides from specific casein regions are usually observed in cheese, and this
460 might be due to a decreased accessibility of certain part of proteins, which upon cleavage become
461 insoluble and aggregate, precluding any further cleavage (Sforza, Cavatorta, Lambertini, Galaverna,
462 Dossena & Marchelli, 2012).

463

464 All the data presented in this paper are useful to understand how peptide amount and profile can be
465 used to predict cocoa flavour quality. The data here presented clearly demonstrated this possibility
466 and with larger databases, an effective control, through molecular markers, of environmental
467 factors affecting cocoa quality as geographical origin, harvesting, fermentation and processing can
468 be reached. Moreover, further work is in progress aimed at modeling Maillard reactions on peptide
469 fractions isolated from cocoa sample, in order to identify the contribute of peptides in the
470 development of cocoa aroma. Of particular interest could be to understand the contribute of
471 peptides from 21 kDa cocoa albumin to cocoa aroma, as they seem to be most abundant in Criollo
472 variety, considered a fine flavor cocoa.

473

474

475

476 **References**

477 Amin, I., Jinap, S., & Jamilah, B. (1997). Vicilin-class globulins and their degradation during cocoa
478 fermentation. *Food Chemistry*, 59, 1–5.

479

480 Amin, I., Jinap, S., & Jamilah, B. (1998). Proteolytic Activity (Aspartic Endoproteinase and
481 Carboxypeptidase) of Cocoa Bean during Fermentation. *Journal of the Science of Food and
482 Agriculture*, 76, 123-128.

483

484 Amin, I., Jinap, S., Jamilah, B., Harikrisna, K., & Biehl, B. (2002). Oligopeptide patterns produced
485 from *Theobroma cacao* L of various genetic origins. *Journal of the Science of Food and
486 Agriculture*, 82, 733–737.

487

488 Ardhana, M.M. & Fleet, G.H. (2003). The microbial ecology of cocoa bean fermentations in
489 Indonesia. *International Journal of Food Microbiology*, 86, 87–99.

490

491 Beckett, S.T. 2009. Industrial chocolate manufacture and use, 4th ed. John Wiley & Sons, Ltd.,
492 Chichester, United Kingdom.

493

494 Biehl, B., & Passern, D. (1982). Proteolysis during fermentation- like incubation of cocoa seeds.
495 *Journal of the Science of Food and Agriculture*, 33, 1280-90.

496

497 Biehl, B., Brunner, E., Passern, D., Quesnel, V.C., Adomako D. (1985). Acidification, proteolysis
498 and flavour potential in fermenting cocoa beans. *Journal of the Science of Food and Agriculture*,
499 36, 583-598.

500

501 Buyukpamukcu, E., Goodall, D.M., Hansen, C.E., Keely, B.J., Kochhar, S., & Wille, H. (2001).
502 Characterization of Peptides Formed during Fermentation of Cocoa Bean. *Journal of Agricultural*
503 *and Food Chemistry*, 49, 5822-5827.

504

505 Bytof, G., Biehl, B., Heinrichs, H., & Voigt, J. (1995). Specificity and stability of the
506 carboxypeptidase activity in ripe, ungerminated seeds of *Theobroma cacao* L. *Food Chemistry*, 54,
507 15-22.

508

509 Caligiani, A., Acquotti, D., Cirlini, M. & Palla, G. (2010). ¹H NMR Study of Fermented Cocoa
510 (*Theobroma Cacao* L.) Beans. *Journal of Agricultural and Food Chemistry*, 58, 12105–12111.

511

512 Caligiani, A., Palla, L., Acquotti, D., Marseglia, A., Palla, G. (2014). Application of 1H NMR for
513 the Characterization of Cocoa Beans of Different Geographical Origins and Fermentation Levels.
514 *Food Chemistry*, 157, 94-99.

515

516 Caligiani, Marseglia & Palla, (2015) Cocoa: Production, Chemistry, and Use. In: Encyclopedia of
517 Food and Health (Edited by B. Caballeros, P.M. Finglas and F. Toldrà), Academic Press.
518

519 Forsyth, W.G.C., Quesnel, V.C., Robert, J.B. (1958). The interaction of polyphenols and proteins
520 during cocoa curing. *Journal of the Science of Food and Agriculture*, 9, 181-184.
521

522 Fowler, M.S. (1999). Cocoa beans: from tree to factory. In: Industrial Chocolate Manufacture and
523 Use (Edited by S.T. Beckett). Pp. 8–35. Oxford: Blackwell Science.
524

525 Guehi, T.S., Konan, Y..M., Koffi-Nevry, R., N’Dri, D.Y. & Manizan, N.P. (2007). Enumeration
526 and Identification of Main Fungal Isolates and Evaluation of Fermentation’s Degree of Ivorian Raw
527 Cocoa Beans. *Australian Journal of Basic and Applied Sciences*, 1, 479–486.
528

529 Guehi, T. S., Dadie A.T., Koffi, K.P.B., Dabonne, S., Ban-Koffi, L., Kedjebo, K.D., & Nemlin G.J
530 (2010). Performance of different fermentation methods and the effect of their duration on the
531 quality of raw cocoa beans. *International Journal of Food Science and Technology*, 45, 2508–2514.
532

533 Hii, C.L., Law, C.L., Cloke, M. & Suzannah, S. (2009). Thin layer drying kinetics of cocoa and
534 dried product quality. *Biosystems Engineering*, 102, 153–161),
535

536 Jinap, M.S., Nazamid, S. & Jamilah. B. (2002) Activation of remaining key enzymes in dried
537 under-fermented cocoa beans and its effect on aroma precursor formation. *Food Chemistry*, 78,
538 407–417.
539

540 Jinap, M.S., Jamilah, B., & Nazamid, S. (2003). Effects of incubation and polyphenol oxidase
541 enrichment on colour, fermentation index, procyanidins and astringency of unfermented and partly
542 fermented cocoa beans. *International Journal of Food Science and Technology*, 38, 285–295.
543

544 Jinap, S., Ikrawan, Y., Bakar, J., Saari, N. & Lioe, H.N. (2008). Aroma Precursors And
545 Methylpyrazines In Underfermented Cocoa Beans Induced By Endogenous Carboxypeptidase.
546 *Journal of Food Science*, 73 (7), 141-147.
547

548 Kattenberg, H. R. & Kemmink, A. (1993). The flavor of cocoa in relation to the origin and
549 processing of the cocoa beans. *Development in Food Science*, 32, 1-22.
550

551 Kochhar, S., Gartenmann, K. & Juillerat, M.A. (2000). Primary Structure of the Abundant Seed
552 Albumin of Theobroma cacao by Mass Spectrometry. *Journal of Agricultural and Food Chemistry*,
553 48, 5593-5599.
554

555 Lopez, A.S. & Dimick, P.S. (1995). Cocoa fermentation. In: Enzymes, Biomass, Food and Feed.
556 2nd edn (Edited by G. Reed & T.W. Nagodawithana). Pp. 561–577. Weinheim: VCH
557

558 Marseglia, A., Sforza, S., Faccini, A., Bencivenni, M., Palla, G. & Caligiani, A. (2014). Extraction,
559 identification and semi-quantification of oligopeptides in cocoa beans. *Food Research*
560 *International*, 63(C), 382-389.
561

562 Marseglia, A., Palla, G. & Caligiani, A. (2014b). Presence and variation of γ -aminobutyric acid and
563 other free amino acids in cocoa beans from different geographical origins. *Food Research*
564 *International*, 63(C), 360-366.
565

566

567 Papalexandratou, Z., Falony, G., Romanens, E., Jimenez, J.C., Amores, F., Daniel, H.M. & De
568 Vuyst L. (2011). Species Diversity, Community Dynamics, and Metabolite Kinetics of the
569 Microbiota Associated with Traditional Ecuadorian Spontaneous Cocoa Bean Fermentations.
570 *Applied And Environmental Microbiology*, 11, 7698–7714.

571

572 Puziah, H., Jinap, S., Sharifah, K.S.M., & Asbi, A. (1998). Changes in free amino acids, peptide-N,
573 sugar and pyrazine concentration during cocoa fermentation. *Journal of the Science of Food and*
574 *Agriculture*, 78, 535–542.

575

576 Rohsius, C., Matissek, R. & Lieberei, R. (2006). Free amino acid amounts in raw cocoas from
577 different origins. *European Food Research and Technology*, 222, 432–438.

578

579 Schwan, R.F. & Wheals, A.E. (2004). The Microbiology of Cocoa Fermentation and its Role in
580 Chocolate Quality. *Critical Reviews in Food Science and Nutrition*, 44(4), 205-221.

581

582 Senanayake. M, E. Jansz, K. & Buckle. (1997). Effect of different mixing intervals on the
583 fermentation of cocoa beans. *Journal of the Science of Food and Agriculture*, 74, 42-48.

584

585 Sforza, S., Galaverna, G., Schivazappa, C., Marchelli, R., Dossena, A. & Virgili, R. (2006). Effect
586 of Extended Aging of Parma Dry-Cured Ham on the Content of Oligopeptides and Free Amino
587 Acids. *Journal of Agricultural and Food Chemistry*, 54, 9422-9429.

588

589 Sforza, S., Cavatorta, V., Lambertini, F., Galaverna, G., Dossena, A. & Marchelli, R. (2012).
590 Cheese peptidomics: A detailed study on the evolution of the oligopeptide fraction in Parmigiano-
591 Reggiano cheese from curd to 24 months of aging. *Journal of Dairy Science*, 95, 3514–3526.

592

593 Spencer, M. E., Hodge, R. (1992). Cloning and sequencing of a cDNA encoding the major storage
594 proteins of *Theobroma cacao*. *Planta* 186(4), 567-576.

595

596 Voigt, J., Biehl, B., Kamaruddin, S. & Wazir, S. (1993). The major seed proteins of *Theobroma*
597 *cacao* L. *Food Chemistry*, 47, 145- 151.

598

599 Voigt, J., Biehl, B., Heinrichs, H., Kamaruddin, S., Gaim arsoner, G. & Hugi, A. (1994a). In-vitro
600 formation of cocoa-specific aroma precursors: aroma-related peptides generated from cocoa seed
601 protein by co-operation of an aspartic endoprotease and a carboxypeptidase. *Food Chemistry*, 49,
602 173-180.

603

604 Voigt, J., Heinrichs, H., Voigt, G. & Biehl., B. (1994b). Cocoa-specific aroma precursors are
605 generated by proteolytic digestion of the vicilin-like globulin of cocoa seeds. *Food Chemistry*, 50,
606 177-184.

607

608 Voigt, J., Voigt, G., Heinrichs, H., Wrann, D. & Biehl, B. (1994c). In vitro studies on the
609 proteolytic formation of the characteristic aroma precursors of fermented cocoa seeds: the
610 significance of endoprotease specificity. *Food Chemistry*, 51, 7-14.),

611

612 Watson, R., Preedy, V.R. & Zibadi, S. (2012). In 'Chocolate in health and nutrition', Ed. Springer
613 Science & Business Media, p.94.

614

615 Wood, G.A.R. & Lass, R.A. (1993) *Cocoa*, 4th edn. Longman Scientific & Technical, Harlow.

616 Zak, D. L., Keeney, P. G. (1976a). Changes in cocoa proteins during ripening of fruit, fermentation,
617 and further processing of cocoa beans. *Journal of Agricultural and Food Chemistry*, 24, 483–486.

618

619

620

621

622

623

624

625

626

627 **Figure Captions**

628 Figure 1: UPLC-ESI-MS chromatograms of peptide elution zone for different cocoa beans samples:

629 a) Criollo from Mexico, b) Forastero from Mexico, c) Forastero from Ecuador d) Forastero from

630 Ecuador (underfermented, violet beans), e) Forastero from Ecuador (unfermented, slaty beans).

631

632 Figure 2: a) Total peptide content (mg/kg) in each cocoa bean sample of different geographical

633 origin and fermentation levels; b) ratio vicilin/21 kDa peptides (light gray: well fermented cocoa

634 beans; dark gray: underfermented beans; black: unfermented beans)

635

636 Figure 3: Principal component analysis: (a) 3d-score plot of all cocoa bean samples on PC1, PC2

637 and PC3 and 2d-score plot on (b) PC1 vs PC2 (c) PC1 vs PC3 and (d) PC2 vs PC3.

638

639 Figure 4: mean peptide values of peptides (mg/kg) in cocoa beans with different fermentation

640 levels, calculated on all the cocoa samples

641

642 Figure 5: distribution of peptides in some representative cocoa samples having low (a), medium (b)
643 and high (c) total peptide amounts; d) comparison of peptide pattern of Criollo variety from Mexico
644 with a Forastero variety from Mexico

645

646

647

648

649

650

Table 1: Maximum, minimum and mean values of each peptide identified in cocoa beans (mg/kg). Values are calculated on all the samples of different origins (well fermented cocoa beans). Peptides are identified by their Mw, sequence and precursor protein (v=vicilin, 21= 21kDa albumin).

Mw	Identification	Protein	Max	Min	Mean
202,2	AI	v	103,78 (Congo)	3,43	37,50
230,2	VI(L)	v/21	39,59	1,51	16,21
264,2	FV	v/21	42,80	2,41	22,75
264,3	VF	v	21,92	2,37	14,15
275,1	AW	21	21,02	0,91	9,28
278,3	FL(I)	v	34,74	2,89	18,05
359,3	VLE	v	5,91	0,37	2,90
359,3	EVL	v	3,65	0,45	1,54
379,2	DVF	v	25,70	0,95	10,57
402,4	RLD	21	11,78	0,06	3,16
407,3	IEF	21	23,26	1,16	9,06
436,3	GDVF	v	4,35	0,01	1,41
474	Mw474		6,83	0,04	1,94
486,3	ANSPV	21	12,84	0,37	6,18
487,3	GAGGGGL	21	21,45	0,12	5,67
533,2	PGDVF	v	9,41	0,04	2,69
599,4	KDQPL	v	13,70	0,17	5,61
620,5	SPGDVF	v	19,43	0,66	7,30
633,3	VSTDVN	21	4,23	0,16	2,01
689,3	NGKGTIT	v	2,88	0,02	0,82
709,4	DEEGNF	v	6,24	0,10	2,66
746,5	NGTPVIF	21	6,91	0,00	2,03
757,4	ASKDQPL	v	18,33	0,95	6,36
760,4	RRSDL	21	2,27	0,00	0,29
788,6	TVWRLD	21	6,54	0,02	0,62
819,5	DNEWAW	21	1,59	0,01	0,44
837,4	DEEGNFK	v	1,78	0,05	0,62
861,4	SSISGAGGGGL	21	6,33	0,02	1,57
901,6	APLSPGDVF	v	1,61	0,01	0,47
932,6	DSKDDVVR	21	2,51	0,04	0,68
1063,5	DEEGNFKIL	v	5,95	0,05	1,84
1090	Mw1090		8,55	0,00	1,30
1204,5	SNADSKDDVVR	21	5,21	0,06	1,76
1244	Mw1244		3,81	0,02	0,93
1596	Mw1596		13,21	0,01	1,57

Figure1
[Click here to download high resolution image](#)

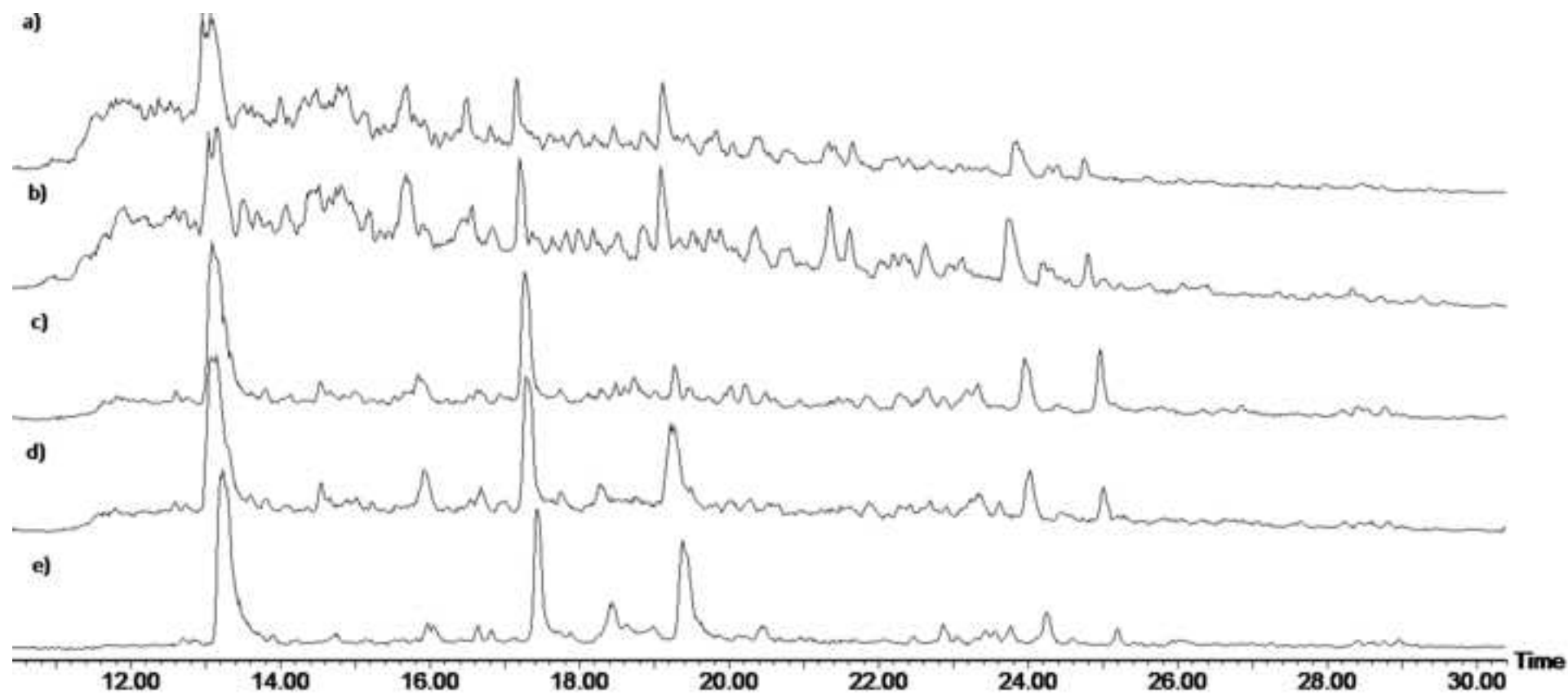


Figure 2

[Click here to download high resolution image](#)

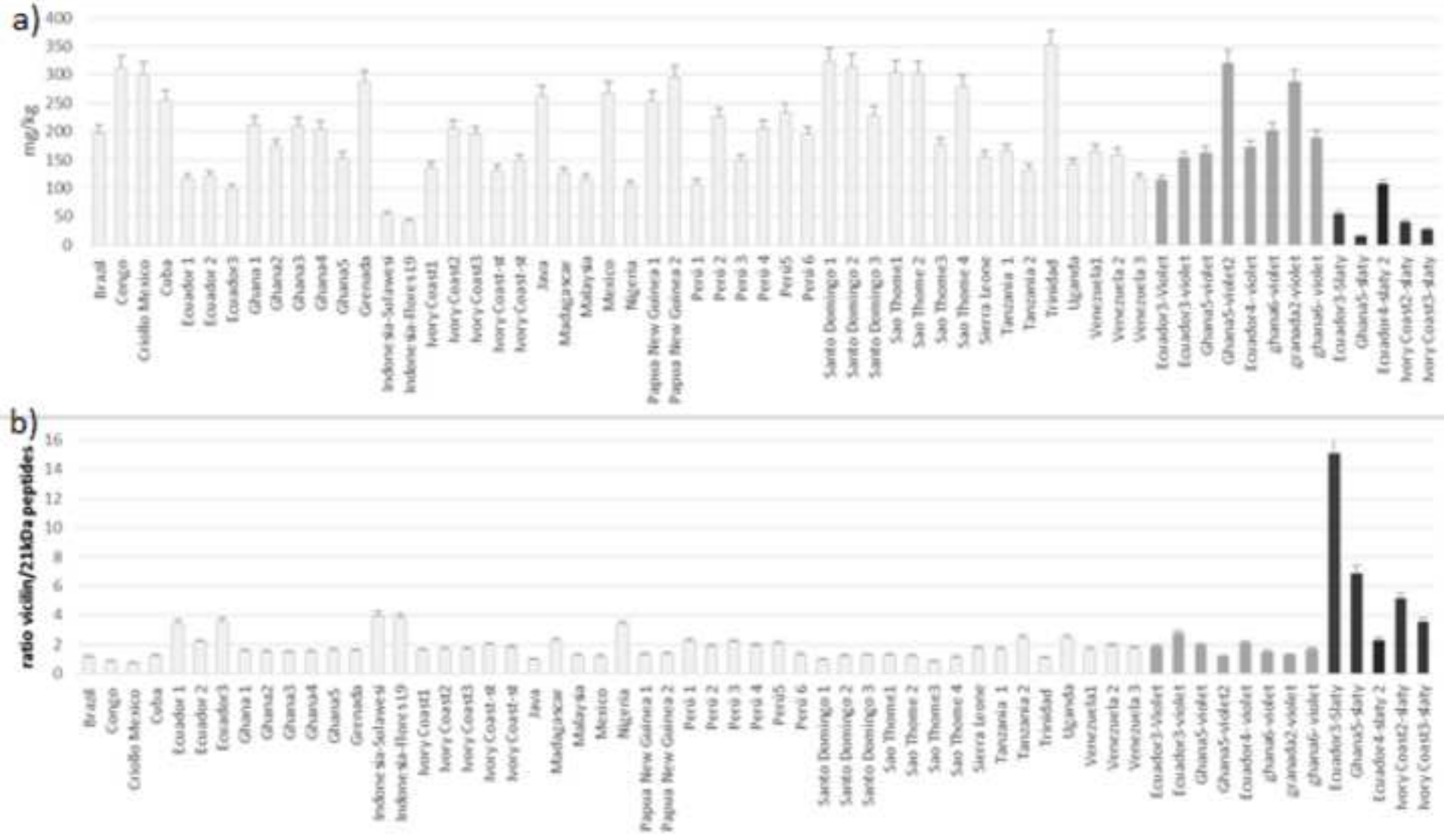


Figure3
[Click here to download high resolution image](#)

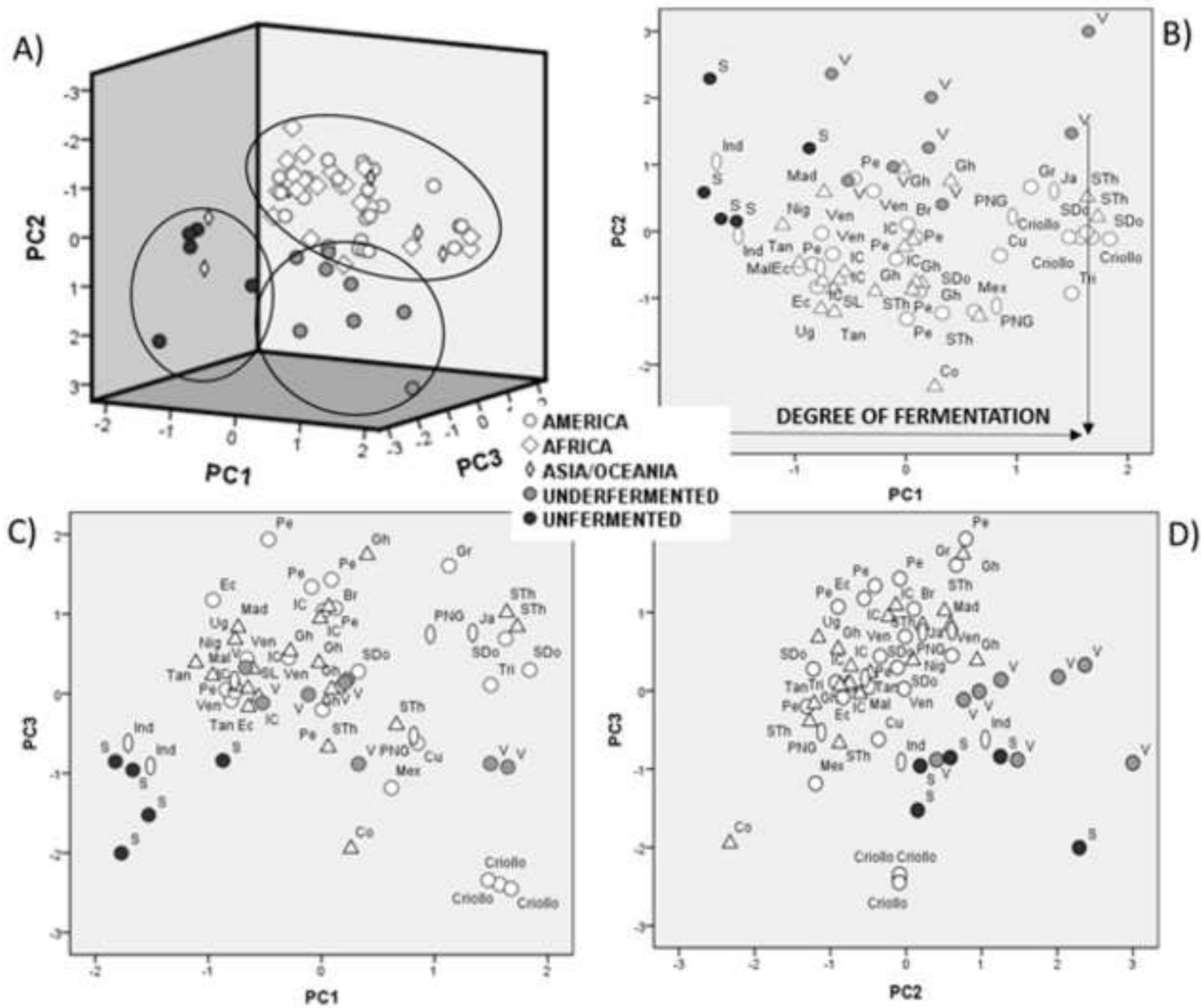


Figure4

[Click here to download high resolution image](#)

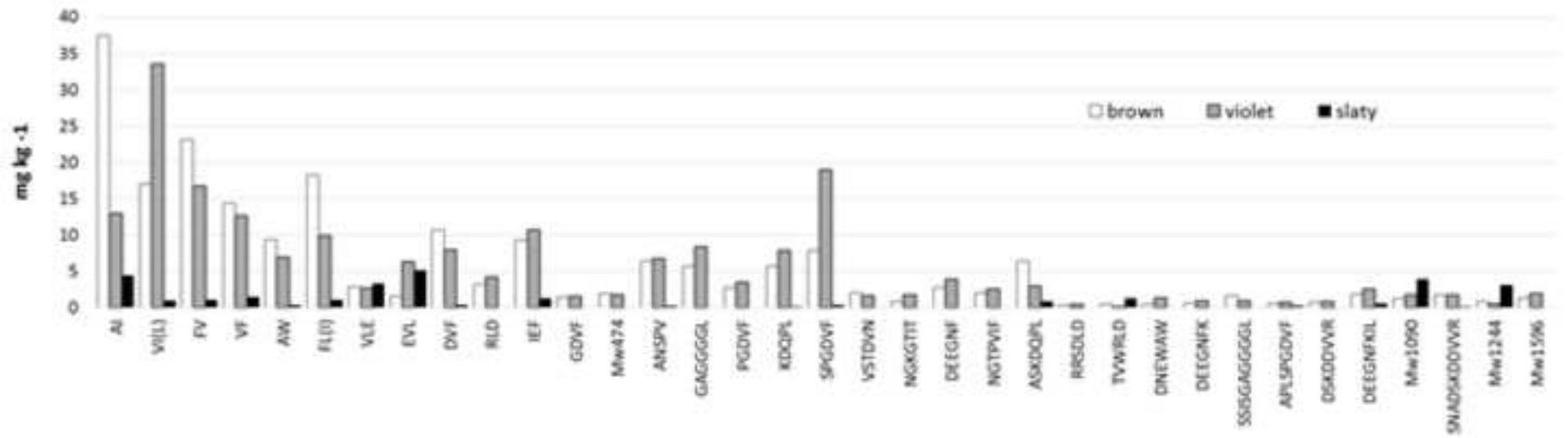
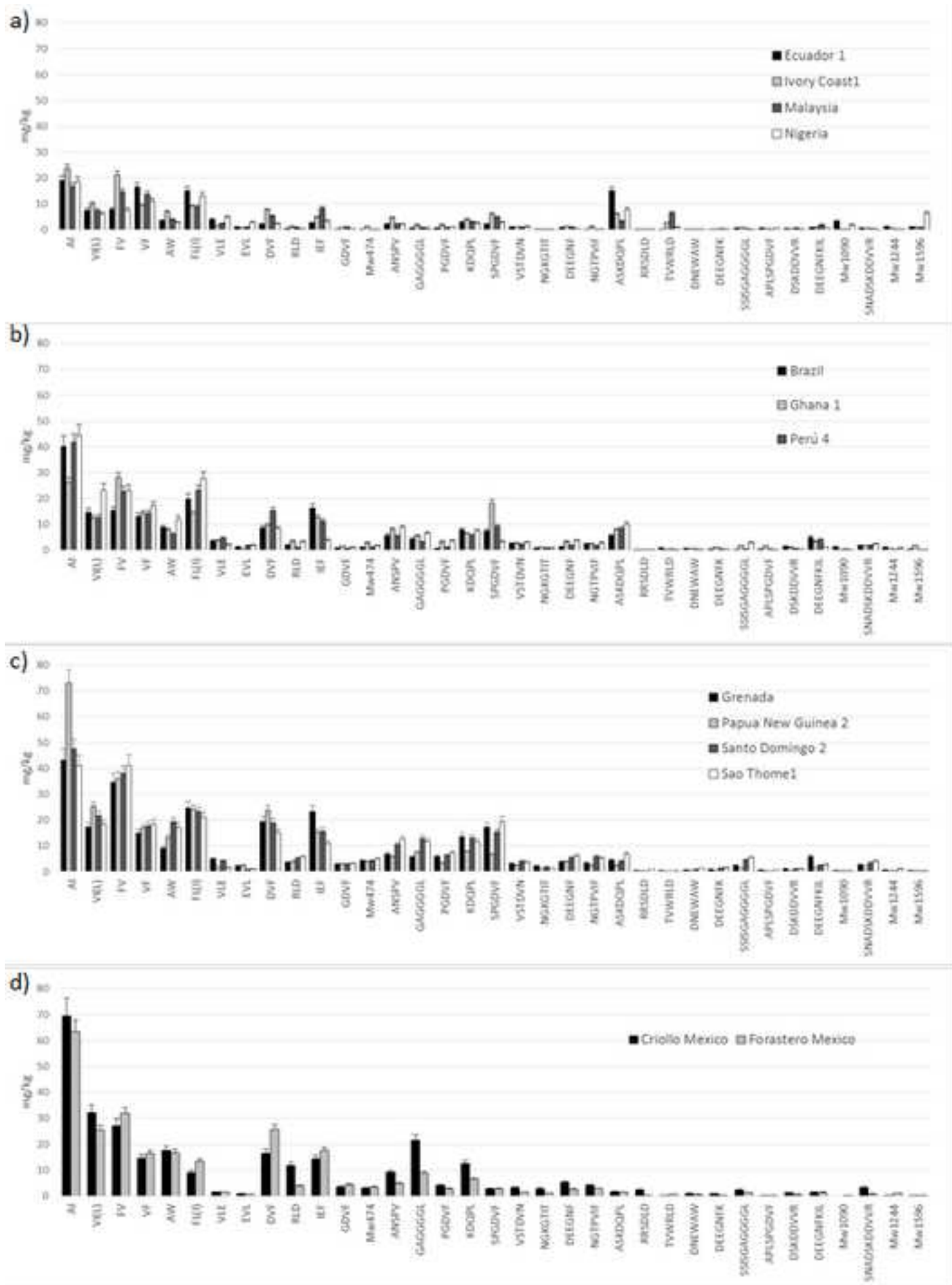


Figure5
[Click here to download high resolution image](#)



Peptides and organic acids

[Click here to download Supplementary Material: Supplementary material peptides and organic acids-S1.xlsx](#)

PCA loadings

[Click here to download Supplementary Material: Supplementary material loadings PCA-S2.docx](#)