



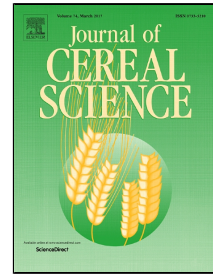
<b>Title</b>	Rice quality profiling to classify germplasm in breeding programs
<b>Author(s)</b>	Ferreira, Ana Rita; Oliveira, Jorge C.; Pathania, Shivani; Almeida, Ana S.; Brites, Carla
<b>Publication date</b>	2017-05-11
<b>Original citation</b>	Ferreira, A. R., Oliveira, J., Pathania, S., Almeida, A. S. and Brites, C. (2017) 'Rice quality profiling to classify germplasm in breeding programs', Journal of Cereal Science, 76, pp. 17-27. doi:10.1016/j.jcs.2017.05.007
<b>Type of publication</b>	Article (peer-reviewed)
<b>Link to publisher's version</b>	<a href="http://dx.doi.org/10.1016/j.jcs.2017.05.007">http://dx.doi.org/10.1016/j.jcs.2017.05.007</a> Access to the full text of the published version may require a subscription.
<b>Rights</b>	© 2017, Elsevier Ltd. All rights reserved. This manuscript version is made available under the CC-BY-NC-ND 4.0 license. <a href="http://creativecommons.org/licenses/by-nc-nd/4.0/">http://creativecommons.org/licenses/by-nc-nd/4.0/</a>
<b>Embargo information</b>	Access to this article is restricted until 12 months after publication by request of the publisher.
<b>Embargo lift date</b>	2018-05-11
<b>Item downloaded from</b>	<a href="http://hdl.handle.net/10468/4600">http://hdl.handle.net/10468/4600</a>

Downloaded on 2018-08-23T19:45:53Z

# Accepted Manuscript

Rice quality profiling to classify germplasm in breeding programs

Ana Rita Ferreira, Jorge Oliveira, Shivani Pathania, Ana Sofia Almeida, Carla Brites



PII: S0733-5210(17)30101-7  
DOI: 10.1016/j.jcs.2017.05.007  
Reference: YJCRS 2359  
To appear in: *Journal of Cereal Science*  
Received Date: 30 January 2017  
Accepted Date: 09 May 2017

Please cite this article as: Ana Rita Ferreira, Jorge Oliveira, Shivani Pathania, Ana Sofia Almeida, Carla Brites, Rice quality profiling to classify germplasm in breeding programs, *Journal of Cereal Science* (2017), doi: 10.1016/j.jcs.2017.05.007

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**RICE QUALITY PROFILING TO CLASSIFY GERMPLASM IN BREEDING PROGRAMS**

Ana Rita Ferreira<sup>1</sup>, Jorge Oliveira<sup>2,3</sup>, Shivani Pathania<sup>1</sup>, Ana Sofia Almeida<sup>1</sup>, Carla Brites<sup>1</sup>

<sup>1</sup>Instituto Nacional de Investigação Agrária e Veterinária (INIAV), Portugal. 2780-157

<sup>2</sup>University College Cork, School of Engineering, Ireland

<sup>3</sup>Ernesto Morgado, S.A., Figueira da Foz, Portugal

Email: carla.brites@iniav.pt

Tel: (+351)214403684

Fax: (+351) 214416011

**Highlights**

- A quality space was defined with five dimensions (water absorption, composition, visual inspection, texture and rheology)
- The selection of axis of the space was based on measuring 50 different parameters of 11 varieties, which included the widest variability of behaviour of long grain rice, and applying variable reduction techniques
- The scores of a rice variety in the proposed quality space provides a fingerprint of behaviour that can be used in breeding programmes to identify directions for improvement, which is illustrated with 4 new varieties under development

27

28 **Abstract**

29 The objective of this work was to define a quality space for assessing rice varieties. Eleven long grain  
30 varieties, seven commercial and four new advanced lines were assessed to obtain complete quality profile  
31 considering appearance, physicochemical parameters, water absorption behaviour, pasting profile and  
32 textural attributes. Commercial varieties were chosen to provide the widest variation in properties,  
33 applying the variability analysis concepts of the Taguchi method, including *Japonicas*, *Indicas*, hybrids  
34 and aromatics. Quality parameters were measured in five different dimensions of quality space (totalling  
35 50). Variable reduction techniques were applied to chose 3 parameters in each dimension (totalling 15  
36 quality indicators) that would describe the whole space with greatest orthogonality, accuracy and yet  
37 explaining a significant proportion of the whole variance of data. The analysis of the quality space thus  
38 defined and similarities between varieties is illustrated with the conclusion of how the 4 new advanced  
39 lines perform in terms of quality behaviour, where it is concluded that one of them is very promising as an  
40 improvement over European (*Indica*) towards the behaviour of a pure Guyana (*Indica*), whereas 3 others  
41 have significant shortcomings in various aspects of the quality space compared to all others, albeit their  
42 greater closeness to *Japonicas*.

43

44 **Keywords:** Rice, Quality Profile, Multivariate analysis, Breeding

45

46

47

48

49

50

51

52

## 53 1. Introduction

54 Rice is the principal staple food for half of the world's population is consumed from ancient times  
55 and there are several varieties grown around the world, which have been adapted to cooking and  
56 consuming styles in various cultures. The quality of rice is a multidimensional concept including  
57 nutritional composition, cooking, physical and textural characteristics (Bocevaska et al., 2009). Eating and  
58 cooking qualities are recognised as important traits affecting consumer acceptability of rice (Bao, 2014).  
59 Due to growing consumer demands, the global rice breeding programs target improvement in grain  
60 quality (Calingacion et al., 2014).

61 Many aspects of grain quality, such as milling behaviour, appearance, nutritional properties and  
62 cooking qualities, have been routinely evaluated. Starch is one of the main components in rice grain and  
63 so its physicochemical parameters such as apparent amylose content, gelatinization temperature, gel  
64 consistency, and pasting viscosity have been used as indicators of cooking and eating quality (Kong et al.,  
65 2015). Rice is consumed largely as cooked whole grains and consequently, texture is an essential attribute  
66 for palatability (Zhou et al., 2002). The elongation of grains, volume expansion as well as water  
67 absorption characteristics, are also important characteristics of cooked rice quality (Ge et al., 2005).  
68 Protein and lipid content are also important factors used to define rice quality (Zhou et al., 2002).

69 There is, therefore, no standard index of rice grain quality because there are many characteristics  
70 to be considered. The objective of this work was to provide a simple fingerprint of the multidimensional  
71 nature of rice quality that could be used to compare different varieties objectively. Different aspects of  
72 rice quality behaviour were deemed relevant: (i) appearance; (ii) chemical composition; (iii) water  
73 absorption; (iv) pasting parameters, and; (v) texture. In order to define a quality space with axis in each of  
74 these dimensions, it is important to define the possible variability that may exist so that the scales are  
75 sufficiently representative of the wide variety of targets that could be defined, depending on the variety  
76 and cooking uses of the rice itself. Therefore, this is not a problem that is addressed properly by using a  
77 large number of varieties chosen randomly, because this could give excessive weight to the average  
78 behaviour and not capture sufficiently well the extremes. An efficient way to quantify maximum

79 variability within a system has been proposed by the Taguchi method, which is a popular method for  
80 robust engineering design (Ross, 1988). In order to obtain an effective assessment of variability,  
81 replicates are planned with outer arrays to ensure that the system is tested under possible extremes of  
82 variability, and with a balance that does not give more weight to the average behaviour that would bias  
83 the results towards the trivial solution. Even with just two conditions, provided they are extreme, the  
84 assessment of variability is more likely to be correct than by testing a large number of conditions that  
85 occur randomly. It is noted that estimating variability with concepts like standard deviation or variance is  
86 more difficult than estimating an average, and if observations were randomly obtained, one might need at  
87 least 50 for a normal distribution, up to 100 for other distributions, which is not effective. Although the  
88 Taguchi suggestion of outer arrays and testing the system under extremes of the noise factors is not based  
89 on statistical proof (Montgomery, 2009), it has a very good body of evidence at this stage of success in  
90 analysing (and handling) variability in engineering systems (Taguchi, Chowdhury and Wau, 2005).  
91 Giving an example of what this could mean in practice, if we wished to describe amylose content  
92 variation, for instance, instead of testing the amylose content by analysing 100 different varieties of rice  
93 to obtain a standard deviation; analysing only two, a very low and a very high variety, would define the  
94 scale of variability efficiently. As the quality domain, in this case, is multidimensional and several  
95 characteristics that can vary significantly among varieties are important, a larger number of extremes was  
96 considered. *Basmati* (from India), *Indica* (from Guyana), *Indica* (from a European variety), *Japonica*  
97 (Ariete variety), *Japonica* (Carnaroli variety) and *Jasmine* (from Thailand) were considered to provide the  
98 widest variation of behaviour for commercially relevant cultivars in Europe. It was further decided to add  
99 a processed grain to this mix of extremes, parboiled (from European *Indica*), simply because from a  
100 consumer point of view, this is a rice that can be bought instead of any of the others as if it was a different  
101 type. Due to the parboiling process causing a pre-gelatinisation of starch, it is known that these grains will  
102 show a very different behaviour, but from a cooking point of view, relevant to the user, they are important  
103 extremes, so the scale of the quality space to be defined needed to be sure of covering this behaviour as  
104 well.

105 In order to illustrate how a quality space defined by extremes of behaviour of the rice varieties  
106 that the European consumer chooses from can be used, four advanced lines from a breeding program,  
107 ongoing in Portugal, were added, totalling 11 different varieties to be characterised.

108 The selection of a limited number of parameters that can define the quality space  
109 comprehensively, yet effectively, was done by profiling quality in its possible totality by measuring all  
110 parameters that could be used to represent the behaviour of the rice samples in all dimensions of the  
111 quality space. In each dimension, parameters were then selected considering their differentiating capacity  
112 of the different rice varieties (generating a high number of homogenous groups), their repeatability (low  
113 variation between replicate measurements), and that in each dimension the selected parameters would  
114 have a high orthogonality between them and describe a high portion of the total variance of the data (high  
115 loadings in different principal components).

116 The data analysis techniques used varied depending on the needs for the quantity of data under  
117 analysis in each case, including partial correlations, Analysis of Variance, multilinear regression and  
118 principal component analysis.

119

## 120 **2. Materials and methods**

### 121 *2.1 Materials*

122 Different commercial types of rice (*Oryza sativa*) available in the domestic market but from  
123 different sources were obtained via the international supply chain of the Ernesto Morgado company  
124 (Figueira da Foz, Portugal): two aromatics (a *Basmati* from India and a *Jasmine* from Thailand), two  
125 *Indica* (a pure *Indica* from Guyana, and a European *Indica* grown in Portugal, variety Gladio), and two  
126 *Japonica* (an *Ariete* and a *Carnaroli*, both grown locally - the former is the most widely used cultivar for  
127 the commercial “carolino” type and the latter for *risotto* cooking). All varieties complied with the  
128 requirements for long grain classification (grains in excess of 6 mm long). A seventh commercial type  
129 was added, parboiled rice, grown and processed in Greece from European *Indica*.

130 For comparison and illustration of the use of the quality space in breeding programs, four  
131 varieties under development by the Instituto Nacional de Investigação Agronomica e Veterinaria (INIAV,  
132 Oeiras, Portugal) were added. These are denoted by the code numbers OP1001, OP1109, OP1212,  
133 OP1203. The former is a variety to be registered under the name Maçarico and the latter is Ceres (the  
134 other two are not in the process of being registered). These varieties result from crossbreeding, containing  
135 the genotypes Basmati x Lido, Estrela x IR72, Valtejo x Onda x Estrela, and Regina x Delta x Estrela x  
136 Suweon x Basmati, respectively

137

## 138 2.2 *Methods*

### 139 2.2.1 *Grain appearance*

140 Rice grain length (L), rice grain width (W), their ratio (L/W), chalkiness (C), total whiteness  
141 (TW) and vitreous whiteness (VW) were evaluated in 50 g samples by image processing (S21 model and  
142 software, Suzuki, Brazil).

### 143 2.2.2 *Chemical composition*

144 Starch, protein, fat and ash content were assessed by a Near Infrared (NIR) analyser (MPA,  
145 Bruker) and calibrations provided by Bruker. The amylose content (in percent) was determined by a  
146 colorimetric technique according to the ISO 6647-2:2015 method.

### 147 2.2.3 *Water absorption behaviour*

148 Weight variation in soaking (WVS) was based on the method used by Medcalf and Gilles (1965)  
149 with slight modifications. Ten g of rice were immersed in 150 mL of water in an agitated container,  
150 without heating, for 20 hours at room temperature. Weight variation in cooking (WVC) was measured  
151 with a similar method but for temperature and time. Samples of rice (150 g) were placed in 500 mL of  
152 boiling water in an agitated container for 20 min. Rice was blotted with absorbent paper to remove excess  
153 water prior to weighing.

154 Weight variation in adapted cooking (WVAC) refers to the quality cooking test used in the  
155 industrial company Ernesto Morgado SA. For real culinary purposes, rice must be cooked in an



156 appropriate quantity of water and for the appropriate time to obtain a similar end result. The cooking time  
157 and the amount of water added in each case were that considered by the industry mentioned to provide a  
158 similar standard of sensory quality, as advised to consumers for home cooking. For *Basmati*, Guyana  
159 (*Indica*), European (*Indica*), Ariete (*Japonica*) and Thai (*Jasmine*) the proportion rice:water was 1:2, and  
160 the cooking time 14 minutes. For Parboiled (*Indica*) the proportion rice:water was 1:2.5, and the cooking  
161 time 17 minutes. For Carnaroli (*Japonica*) the proportion rice:water was 1:3, and the cooking time 17  
162 minutes.

163 The weight variation is just the difference between the final and the initial weights divided by the  
164 initial. It is noted that weight variation is not a measure of water absorption alone, as the rice may also  
165 leach out starch and other soluble solids. Volume-based analysis of water absorption was performed with  
166 the standard Borasio apparatus to measure the hydration capacity (MIAG BUHLER-spA). The Borasio  
167 test (Borasio, 1935) gives the volume increase (VI, in percent) of the rice due to water absorption and the  
168 volume of water that was absorbed (WA (%)) by the rice.

#### 169 2.2.4 Pasting parameters

170 The determination of rice paste gelatinization and viscosity characteristics was performed by a  
171 rapid viscosity analyzer (RVA-4, Newport Scientific, Warriewood, Australia) following AACC  
172 International Approved Method 61-02.01. Viscosity curves determined 8 parameters: peak viscosity  
173 (VP1), trough (VT1), time for 1<sup>st</sup> peak (VPt), temperature at 1<sup>st</sup> peak (VPT), setback of 1<sup>st</sup> peak (VS1),  
174 setback of 2<sup>nd</sup> peak (cooling) (VS2), breakdown (VB) and final viscosity (VF).

#### 175 2.2.5 Texture attributes

176 The texture profile analysis (TPA) of 5g samples of rice cooked in excess water (as in WVC  
177 measurements) was determined using a Texture Analyser TA-XT2 (Stable Micro Systems, London, UK).  
178 A two-cycle compression, the force-versus-time program was used with a test speed of 2 mm/s and a rate  
179 of 80% strain using a cylinder plunger with a 20 mm diameter. Twenty-five parameters were taken from  
180 each test: Peak force of 1<sup>st</sup> (P1) and 2<sup>nd</sup> (P2) positive cycles and 1<sup>st</sup> negative cycle (P-1); Area under the  
181 curve for the 1<sup>st</sup> (A1) and 2<sup>nd</sup> (A2) positive cycle and the first negative cycle (A-1); time to 1<sup>st</sup> positive

182 (tP1) and 1<sup>st</sup> negative (tP-1) peaks; Peak ratios 1<sup>st</sup> negative / 1<sup>st</sup> positive (RP-1/1) and 2<sup>nd</sup> positive / 1<sup>st</sup>  
 183 positive (RP2/1); Area ratios 1<sup>st</sup> negative / 1<sup>st</sup> positive (RA-1/1) and 2<sup>nd</sup> positive / 1<sup>st</sup> positive (Ac1);  
 184 Cohesiveness (Coh); Chewiness (Che); Apparent modulus (AM); Adhesiveness (Adh); Work done to  
 185 hardness of 1<sup>st</sup> cycle (WH1) and of 2<sup>nd</sup> cycle (WH2); Recoverable work done in 1<sup>st</sup> cycle (RW1) and in 2<sup>nd</sup>  
 186 cycle (RW2); Gumminess (Gum); Chewiness index (C Ind); Springiness (Spr); Springiness index (Spr I);  
 187 Mean load (ML).

#### 188 2.2.6 Statistical analysis

189 All analysis were made using the Statistica software (v7, Statsoft, Tulsa OK, USA). Pair-wise  
 190 partial correlations were assessed with Pearson's correlation coefficients. Homogenous groups were  
 191 determined with post-hoc analysis of one-way ANOVA for the parameter in question, with Tukey's  
 192 Honest Significant Differences. Multiparametric correlations were determined with least squares  
 193 regression. Principal Component Analysis (PCA) was done with normalised Varimax rotation. Data were  
 194 normalised prior to processing to ensure equal weights to each parameter:

$$195 \quad X_{norm,i} = \frac{X_i - X_{min}}{X_{max} - X_{min}}$$

196 where  $X_i$  is one sampling point of parameter  $X$ ;  $X_{min}$  the minimum of all  $X$  value readings;  $X_{max}$  the  
 197 maximum of all  $X$  value readings, and thus  $X_{norm,i}$  is the normalised value, necessarily having a value  
 198 between 0 and 1.

199 The overall similarity between different varieties was quantified from the mean squared deviation  
 200 (SD) of the selected quality parameters, by calculating the Euclidean distance (ED) of the normalised  
 201 parameters chosen:

$$202 \quad ED = \frac{\sqrt{\sum_{k=1}^{n_p} (X_{norm,k,j} - X_{norm,k,ref})^2}}{n_p}$$

203

204 where  $X_{\text{norm},k,j}$  denotes the normalised value of chosen parameter  $k$  for rice variety  $j$  and  $X_{\text{norm},k,\text{ref}}$   
205 the corresponding value for the same parameter for the variety against which the others are being  
206 compared, and  $n_p$  is the total number of parameters chosen to describe the 5 dimensions of the quality  
207 space ( $n_p=15$ ).

208

### 209 3. Results and discussion

210

#### 211 3.1 Water absorption

212 The number of Tukey homogeneous groups (95% confidence level) and the average standard  
213 deviation for each parameter are shown in table 1. The partial correlation coefficients between all pairs  
214 are not shown, the lowest correlation (0.022) was between WVC (weight variation in cooking) and WVS  
215 (weight variation in soaking). These were also the parameters with greater distinguishable capacity  
216 (higher Tukey groups) and repeatability (lower average standard deviation) and therefore were chosen to  
217 define this quality dimension. WVAC (weight variation in adapted cooking) was highly correlated with  
218 these 2: a simple multilinear model  $\text{WVAC}(\text{normalised}) = a + b \cdot \text{WVS}(\text{normalised}) +$   
219  $c \cdot \text{WVC}(\text{normalised})$  had a 0.91 correlation coefficient with the experimental data, therefore WVAC can  
220 be considered largely defined by the other 2 parameters and removed. VI (volume increase) and WA  
221 (water absorption) had a strong correlation between them (0.73 correlation coefficient). VI had a lower  
222 correlation with WVS and WVC (0.49 correlation coefficient for a multilinear model) than WA (0.70)  
223 and a lower average standard deviation, and therefore VI was chosen as a 3rd parameter needed to define  
224 this quality dimension fully, and WA can be excluded: a multilinear model for WA as a function of WVS,  
225 WVC and VI has a high correlation coefficient (0.83).

226 In conclusion, weight variation in soaking and in cooking and volume increase in the Borasio test  
227 are the parameters needed to describe the behaviour regarding water absorption. These parameters can be  
228 used to identify similarities and differences between cultivars. For instance, figure 1 shows the 2-D  
229 projections of the quality space thus defined, where it can be seen that 3 of the new varieties, OP1109,

230 1203 and 1212, are very different from all others, whereas OP1001 is quite similar to European indica  
231 (Gladio variety).

232 It can also be seen that commercial varieties and advanced lines showed mean water absorption  
233 of 30% and 28% of their weight in cold water, and indica (Parboiled) particularly had high WVS, which  
234 could be due to hydrothermal treatment leading to lower glass transition temperature of rice, hence more  
235 hydration. Water binding during soaking depends on diffusion properties of rice, which rely on grain  
236 structure, composition (protein and amylose content), post-harvest processing, soaking temperature and  
237 initial moisture content of rice (Oli, Ward, Adhikari, & Torley, 2014). The presence of proteins in surface  
238 layers and their higher water diffusion coefficient than starch aids in water absorption during soaking,  
239 however, it has been reported that high protein rice requires more water and time to cook (Husain, Hwa,  
240 & Ahmad, 1981).

241 Water absorption increases significantly when the temperature exceeds gelatinization  
242 temperature, with WVC values being about 10 times greater than WVS. The longer varieties Basmati,  
243 Jasmine (Thai) and Indica (Guyana) ( $L = 6.82\text{mm}, 6.46\text{mm}, 6.83\text{ mm}$ ) were characterised by higher  
244 WVC values of 2.48, 2.44 and 2.12, respectively. High amylose content has been related to high WVC in  
245 literature (Juliano, 1993), however, in this study effect of length was more pronounced than effect of  
246 amylose content (amylose content = Basmati: 31.19%; Jasmine (Thai): 16.73%; Indica (Guyana):  
247 29.43%) on WVC values. It has been reported that high amylose rice has a higher capacity to absorb  
248 moisture during cooking (Juliano, 1993; Yadav and Jindal, 2007). Yadav and Jindal (2007) observed that  
249 the changes in dimensions were greater in lower amylose content rice varieties than the higher amylose  
250 content types. Hydration properties of high amylose varieties Indica (Guyana) and Japonica (Carnaroli)  
251 are different irrespective of similar amylose content which is attributable to the potential effect of  
252 amylopectin structure and other factors (Tran et al. 2011; Traore, McClung, Fjellstrom, & Futakuchi,  
253 2011). High protein rice has been shown to absorb water at a slower rate during cooking. It was also  
254 reported that high protein rice required more water and longer time to cook (Husain, Hwa, & Ahmad,

255 1981). Protein is supposed to form a thicker barrier around the starch granule, thus slowing water uptake,  
256 retarding gelatinization and swelling (Bocevaska, Aldabas, Andreevska, & Ilieva, 2009).

257

### 258 3.2 *Pasting parameters*

259 The typical RVA test applied subjects the sample to a gradual heating, holds it at the high temperature for  
260 some time, and then cools it down gradually again, measuring the apparent viscosity continuously, which  
261 will show a first peak during the heating part of the cycle, then a trough and a second peak with the  
262 cooling part of the cycle. From these viscosity curves 8 different parameters can be determined: Value of  
263 the 1st (heating) viscosity peak (VP1), value of the 1st (heating) trough (VT1), time at which the 1st peak  
264 occurs (VPt), temperature when the 1st peak occurred (VPT), setback of the 1st peak (VS1), setback of  
265 the 2nd peak (cooling) (VS2), breakdown (VB) and final viscosity (VF). Each variety was tested 3 times

266 In order to seek a minimum number of parameters with high orthogonality and high probability of  
267 explaining a high degree of the total variance of all data, the first approach was to cluster the parameters  
268 with a Principal Component Analysis. The results are shown in table 2, together with the number of  
269 Turkey groups at 95% confidence level in one-way ANOVA for each parameter and the normalised  
270 average standard deviation. Parameters with high loadings in one principal component have a high degree  
271 of interdependency whereas parameters with high loadings in different PC's will have a high degree of  
272 orthogonality. In problems with a large number of parameters, such as chemometry, it is common that the  
273 individual parameters are replaced by the scores of the principal components. However, it was not desired  
274 to do so in this case because it is important to have parameters that have a low error (high repeatability).  
275 Scores of principal components will inevitably have a much higher error than individual parameters for  
276 two reasons, the propagation of error of all parameters into the score, and the error of the loading factors  
277 themselves. Thus, it was considered more appropriate to use the PCA to select key parameters from each  
278 PC, with a high loading factor and a low standard deviation, and from among each group consider also the  
279 distinguishability capacity (number of Tukey groups) and a low pairwise correlation coefficient between  
280 all parameters thus selected (high orthogonality). It can be seen that just 3 principal components explain

281 most of the variance of the data (96%). VP1 is the most repeatable parameter (lowest standard deviation),  
282 but it is the only that does not have a clear allocation to a PC, which means that it is bound to generate too  
283 high correlations with others. This was indeed verified with the partial correlation coefficients (above 0.5  
284 for all parameters, except 0.27 for VS1). A multilinear model using VP1, VS1 and the lowest correlated  
285 in the PC1 group with these two (that would be VF) was indeed found to be unable to explain a  
286 sufficiently high variance of two parameters, (VPT and VPt), thus, VP1 is not a good choice for reasons  
287 of orthogonality with others.

288 The maximum orthogonality will be obtained by selecting first one of the parameters with high  
289 loadings in PC1, the best choice is VT1 which has the highest loading factor (hence the greatest  
290 probability of high orthogonality with others) and high differentiating capacity (highest number of Tukey  
291 groups). The next choice from PC2 will be VS1 which also has a high differentiating ability, low error,  
292 and lowest partial correlation coefficient with VT1 (0.25). Finally, there is only one parameter in the 3rd  
293 PC, VPt, and it has lower correlation coefficients with VT1 (0.5) and VS1 (0.067) than VP1 would have  
294 (0.81 and 0.27, respectively).

295 Table 2 also shows the percent of the variance of the data of each parameter that is explained by a  
296 simple multilinear model of the 3 chosen parameters, VT1, VS1 and VPt, with are respectively the trough  
297 of peak 1 (heating), the setback of peak 1 (heating) and the time when peak 1 (heating) occurs. The %  
298 variance explained is the square of the correlation coefficient in the case of the other parameters, and the  
299 explained variance of the one-way ANOVA in the case of the chosen parameters. It can be concluded that  
300 this choice of 3 provides a very good description of the whole set of data of all 8 parameters, with high  
301 orthogonality.

302 The locations of the 11 varieties in this dimension of the quality space are shown in fig. 1.

303 The effect of amylose content on peak viscosity is not evident. Works have been reported where  
304 amylose content is positively (Singh et al., 2006; Tran et al., 2001) and negatively (Chung, Liu, Lee, &  
305 Wei, 2011; Tan & Corke, 2002) correlated with peak viscosity. Differences in the viscosity profiles are  
306 also attributed to characteristics of the starch granules i.e. amylose/amylopectin ratio and chain length

307 distribution of amylopectin. Basmati had the highest value for trough, final viscosity and set back.  
308 Japonica varieties showed higher values for breakdown and lower values for final viscosity and setback.  
309 The low final viscosity in intermediate to low amylose varieties is explained by fewer amylose chains.  
310 The linear nature of the amylose molecule allows it to retrograde, forming a gel matrix in which  
311 amylopectin molecules and starch granules are embedded and sometimes take several days to retrograde.  
312 Non-waxy varieties retrograde faster, resulting in higher final viscosity than the waxy varieties as seen in  
313 this study (Cueves & Fitzgerald, 2012). High values for breakdown and low values for setback are  
314 desirable for processed rice and are indicative of high cooking quality since neither the cooked rice  
315 retrogrades nor becomes stiff upon cooling (Kong, Zhu, Sui, & Bao, 2015). For commercial varieties,  
316 setback viscosity increased with amylose content which is not due to peak viscosity, but rather because  
317 final viscosity is affected by amylose content. Advanced lines (except OP1001) had significantly different  
318 behaviour as they showed high peak viscosity, breakdown, intermediate holding strength (trough) and low  
319 set back values. OP1001, however, had a pasting profile similar to Basmati, Indica (Guyana) and Indica  
320 (European). Parboiled rice had minimum values for all the pasting parameters. Elevation of final viscosity  
321 and the decline in the breakdown and total setback time are the major changes due to parboiling (Oli,  
322 Ward, Adhikari, & Torley, 2014).

323

### 324 3.3 *Texture attributes*

325 Texture is a multidimensional attribute that governs palatability of cooked rice. A very  
326 comprehensive test is a two-cycle compression test resulting in a TPA (texture profile analysis), which  
327 provides a large number of parameters, 23 in total. Table 2 shows the results of a principal component  
328 analysis of the normalised values of all parameters, as well as the normalised standard deviations of  
329 repeats and the number of homogenous groupings with one-way ANOVAs in each case. The peak force  
330 of the 1<sup>st</sup> cycle (P1) had the highest loading in PC1 and is also a quantification of hardness, perhaps the  
331 most fundamental texture characteristic, and, therefore, was chosen as a key parameter. Considering then  
332 a key parameter from PC2, the highest loading was for Adhesiveness, but this parameter did not have a

333 very good capacity to differentiate between varieties, having generated only 3 Tukey groups in its one-  
334 way ANOVA and therefore the negative peak force in the release part of the 1<sup>st</sup> cycle was preferred as key  
335 parameter for the 2<sup>nd</sup> PC. It also has a clear physical meaning, being the maximum adhesive force, and  
336 was called "stickiness" for further reference. As expected, the partial correlation coefficient between these  
337 2 chosen parameters is very low (0.071). A third parameter needs to be chosen, but those with high  
338 loadings in the 3<sup>rd</sup> and the 4<sup>th</sup> PCs have a very poor capacity to differentiate between varieties, with only 2  
339 Tukey groups for springness and springness index and no significant differences detected between any  
340 variety in the case of the ratio of the 2<sup>nd</sup> to the 1<sup>st</sup> peak. Thus, it was preferred to choose as a 3<sup>rd</sup> parameter  
341 one of those without a high loading factor in any of the PCs. The mean load also had no capacity for  
342 differentiation, so cohesiveness and ratio of the negative to positive areas of the first cycle would be better  
343 options. However, cohesiveness had a high partial correlation coefficient with one of the parameters  
344 already chosen (0.722 with P-1), whereas the area ratio between negative and positive peaks of the first  
345 cycle had low correlations (0.271 and 0.13 with P1 and P-1, respectively). It is, therefore, a good choice  
346 for a 3<sup>rd</sup> key parameter with high orthogonality and differentiating capacity. It quantifies the ratio of  
347 energies involved with adhesion and compression (energy released by the mechanical recovery over the  
348 energy spent in the initial compression) and was thus designated "recovery" for future reference.

349 In order to assess whether 3 parameters would suffice to quantify all the texture data, a  
350 polynomial 2<sup>nd</sup> order surface was used to fit the data for each of the other parameters. There were only 4  
351 were a simple combination of the chosen 3 key parameters could not explain satisfactorily the variance of  
352 the data: springiness, springiness index, ratio of 2<sup>nd</sup> to 1<sup>st</sup> peak forces and cohesiveness. However, these  
353 are also parameters with a very poor usability to differentiate between varieties, with only 2 or 1 Tukey  
354 groups, and therefore it was considered that the added complexity of bringing a 4<sup>th</sup> parameter to  
355 accommodate this data better was not worthwhile.

356  
357



358 Fig. 1 shows the location of the 11 varieties in the quality space defined by the 3 parameters  
359 chosen.

360 Hardness is the most commonly measured parameter (Zhout, Robards, Helliwell, & Blanchard,  
361 2002). *Indica* (Parboiled) and *Indica* (Guyana) showed the highest values with lowest being *Jasmine*  
362 (Thai) and *Basmati* varieties. Several studies have demonstrated that parboiling improves textural  
363 properties such as hardness, cohesiveness, gumminess, chewiness and resilience in cooked grain (Bello,  
364 Baeza, & Tolaba., 2006; Zhout, Robards, Helliwell, & Blanchard, 2002). Bello et al. (2006) reported  
365 harder texture in cooked parboiled rice than the untreated counterpart. The hardness value of other  
366 varieties was in a similar range, yet the results had some significant difference among them. Bocevaska,  
367 Aldabas, Andreevsha, & Ilieva (2009) stated that amylose content has a strong influence on rice texture.  
368 Higher amylose content corresponds to harder texture in general. This was not observed in these data, as  
369 basmati, which was the softest, had a high amylose content. Generally, rice varieties with high amylose  
370 content (over 25%) cook dry, are less tender and become hard upon cooling while those with low amylose  
371 content (below 20%) cook moist and are sticky (Dipti, Bari, & Kabir, 2003). Low protein rice samples are  
372 tender than high protein samples of the same cultivar (Hamaker and Griffin, 1993). It suggested that the  
373 existence of a protein barrier and the hydration of the protein together influenced viscosity curves, which  
374 would determine the texture of the cooked rice (Derycke et al., 2005).

375 There is a clear differentiation in cohesiveness values for commercial varieties, *Japonica* (Ariete)  
376 and the advanced lines. Cohesiveness values were positively correlated with the grain length, L/W,  
377 protein content ( $p < 0.01$ ) and amylose content ( $p < 0.05$ ). Results were similar to works documented by  
378 Mohapatra & Bal, 2006. However, the values of springiness are similar for almost all the samples.  
379 Odenigbo et al., (2014) reported that springiness of cooked parboiled and non-parboiled rice was similar.  
380 Gumminess is the energy required to disintegrate a semi-solid food to a readiness for swallowing. The  
381 value of gumminess for parboiled rice was substantially higher than commercial and advanced lines. Non-  
382 parboiled high amylose rice has been known to have low gumminess with a high value for adhesiveness  
383 (Odengibo et al., 2014), which is similar to the results shown here. Grain length, L/W and amylose

384 content were found to have a positive correlation with adhesiveness ( $p < 0.01$ ). However, Mohapatra and  
385 Bal (2006) reported that raw rice having high amylose content had lower adhesiveness values compared  
386 to low amylose and medium amylose varieties. This was attributed to the fact that high amylose rice has a  
387 dry and fluffy texture when cooked, compared to a lower amylose variety (Mohapatra and Bal, 2006).

388

### 389 3.4 Grain appearance

390 There were 6 parameters measured for visual appearance: 3 biometric - length (L), width (W), the  
391 length-to-width ratio (L/W) and 3 colour - total whiteness (TW), vitreous whiteness (VW) and chalkiness  
392 (C). Obviously, the 3 biometric parameters are linearly dependent (from 2 of them the 3rd is immediately  
393 given).

394 A Principal Component Analysis showed that the 6 parameters clustered in just 2 principal  
395 components explaining 87% of the total variance of the data, with the 3 colour parameters having high  
396 loadings in the first principal component, and the 3 biometric ones with high loadings in the second. Size  
397 and white/chalk parameters are therefore highly orthogonal. Length showed the lowest correlations with  
398 total and vitreous whiteness (0.04 and 0.008) and also a very low correlation with chalkiness (0.175) and  
399 was, therefore, the first choice. Vitreous and total whiteness were very highly correlated (0.97). Vitreous  
400 whiteness had the lowest correlation with the biometric parameters and therefore either one is a possible  
401 choice for the colour parameters. However, choosing only 2 parameters (one from each principal  
402 component) would not describe the data sufficiently (a multilinear model with only these 2 gives low  
403 correlation coefficients for chalkiness, 0.58). All combinations were tried and the best was to choose as  
404 defining parameters of the visual appearance the length, the width and the total whiteness. A multilinear  
405 model of these 3 has a high correlation coefficient with chalkiness (0.81), and obviously also with L/W  
406 and with vitreous whiteness. Thus, these 3 parameters represent the full set very well. Figure 2 shows  
407 how the 11 varieties spread in this dimension of the quality space.

408 Consumer preference of rice is invariably dependent on appearance and includes whiteness,  
409 chalkiness, grain length, grain width and length-to-width (L/W). Whiteness and translucency

410 (vitreousness) are important parameters as it has been studied that consumers prefer white and translucent  
411 grains and are prepared to pay a premium for it (Rani et al., 2006). The rice varieties studied had  
412 whiteness values ranging from 76.99 (Indica (Parboiled)) to 163.60 (Japonica (Carnaroli)). In general,  
413 Indicas had lower values than Japonicas for whiteness, vitreousness and chalkiness. The total whiteness of  
414 advanced lines was higher and vitreousness was comparable to Indicas. White and translucent grains are  
415 rated more highly by consumers (Fofana et al., 2011; Adu-Kwarteng, Ellis, Oduro, & Manful, 2003).  
416 Chalkiness is a major concern in rice breeding because it is one of the key factors in determining quality  
417 and price. The chalky endosperm consists of loosely packed, round and large compound starch granules  
418 while the translucent endosperm comprises tightly packed, polyhedral and small single starch granules.  
419 The chalky grains show significantly different physicochemical, morphological, thermal, cooking and  
420 textural properties from translucent grains. Percentage of grains with chalkiness is one of the main indices  
421 of rice-determining appearance quality (Bao, 2014).

422         The range for grain length, grain width and L/W ranged from 5.56 mm to 6.83mm, 2.26 mm to  
423 4.31 mm and 2.26 to 4.31 respectively. Indicas are characterised by longer grains and slender appearance.  
424 Grain appearance is an important parameter but it has been stated that it is not easy to recognise which  
425 rice varieties grains really belong to according to their appearance (Bocevaska, Aldabas, Andreevska, &  
426 Ilieva, 2009).

427

### 428 *3.5 Chemical composition*

429         Six compositional parameters were measured: the contents of amylose, starch, moisture, ash, fat  
430 and protein. A principal component analysis clustered in the first component with high loadings starch,  
431 moisture and protein content, and in a second component the fat and amylose contents. Ash was not  
432 strongly bound to either. Amylose and moisture contents had the best repeatability, a good number of  
433 Tukey groups and a low correlation between themselves (0.36), being in different PC's. Thus, they are  
434 good choices for this quality dimension. These two parameters would replace starch, fat and protein  
435 content as a simple multilinear model of these as a function of the chosen 2 has a good correlation (0.86,

436 0.72 and 0.75, respectively), but not ash (correlation coefficient of the multilinear model 0.37). Ash was  
437 also not clustered in the two principal components and had very low correlation with either moisture or  
438 amylose content. Thus, it is necessary to choose this parameter as well to cover a high percentage of the  
439 variance of the data. A multilinear model of moisture, amylose and ash content explains a high percentage  
440 of the variance of the data of the other 3 parameters (correlation coefficients of 0.72, 0.86 and 0.91 for fat,  
441 starch and protein contents, respectively).

442 Different rice varieties encompass varied physicochemical properties which are the basis of rice  
443 grain quality and influence advanced quality attributes in relation to pasting and textural profiles. Within  
444 the compositional data, amylose content and protein content are most important parameters. Earlier  
445 studies have reported that amylose content and protein content directly influence the cooking, eating,  
446 pasting and textural qualities of milled rice (Champagne et al., 1997; Fofana et al., 2011; Kong, Zhu, Sui,  
447 & Bao, 2015; Bagchi, Sharma, & Chattopadhyay, 2016). Rice varieties had a wide variation in amylose  
448 content, ranging from 16.73% to 31.19% for the seven commercial varieties (19.42% to 33.70% for  
449 advanced lines). Previous studies reported that apparent amylose content of different rice genotypes  
450 ranged from 1.5% to 31.6% (Singh, Kaur, Sodhi & Sekhon, 2005; Yadav and Jindal, 2007; Kaur, Panesar,  
451 Bera, & Kumari, 2014) and Lawal et al. (2011) classified varieties as waxy (0-2 % amylose), very low  
452 amylose (2 -12 %), low amylose (12 -20 %), intermediate amylose (20-25 %) and high amylose (25-33  
453 %). The range obtained with the chosen 7 commercial varieties is, therefore, a good representation of rice  
454 with intermediate to high amylose content. As depicted in figure 2, Jasmine (Thai) with OP1109 and  
455 OP1212 had the lowest values for amylose content while OP1001 had maximum amylose content  
456 (33.70%). It is considered that amylose content is mainly linked with sensory properties of cooked rice  
457 while protein in rice is associated with textural properties of cooked rice (Juliano, 2003). The mean value  
458 of protein content in commercial varieties was 7.64%. With respect to milled rice, protein is the second  
459 most abundant constituent (after starch) ranging from 4 to 11% (Champagne, Wood, Juliano, & Bechtel,  
460 2004). According to the classification given by Matsue and Ogata (1998), Japonica varieties usually have  
461 6.0-7.0 % protein whereas Indica varieties have protein content in the range of 7.1-8.9 %. Protein content

462 in Indica rice was reported to be 10.00–20.00% higher than Japonica varieties (Sun, Hou, & Zhang,  
463 2008).

464 In spite of protein content having a very significant importance in rice behaviour, it is not part of  
465 the chosen parameters to define the compositional space, as its values are well predicted by the 3 chosen  
466 parameters (a simple linear combination of the values of moisture, ash and amylose content had a  
467 correlation coefficient for protein content of 0.91, and a simple quadratic surface would explain 93.9% of  
468 the variance of the protein data).

469

### 470 3.6. Overall quantification of differences

471 All scores of the chosen quality parameters were normalised between their observed minimum  
472 and maximum, and show a variability measured with 3 repeats. The overall similarity between different  
473 varieties can thus be quantified by a mean square difference, giving equal importance to each parameter.  
474 The mean square error is the same as the average Euclidean distance (ED) in the quality space, defined  
475 by:

476 This quantitative measure of similarity of the quality behaviour can be measured for one dimension of the  
477 space (3 parameters) or for the whole space, (15 parameters). The significance of the number obtained can  
478 be compared to the maximum ED between different repeat measurements of the same variety: when ED  
479 between two varieties is smaller than the maximum ED between repeats, there is no significant difference  
480 between the varieties. Triangular plots can be used to detail the difference.

481 The use of this analysis of the quality space for breeding programs is illustrated with the analysis  
482 of the 4 new varieties denoted with the codes OP. As the trials took place in Portuguese fields where the  
483 commercial varieties that are known to grow well are European Indica and Japonica (ariete), would the  
484 new varieties just be more of the same (in which case better yields would be their only interest if  
485 verified)? Or would any of them show a quality profile more similar to more interesting varieties, like a  
486 pure Indica (Guyana), or a basmati, or a carnaroli, in which case they would show potential for particular  
487 culinary uses?

488           There are of course many comparisons that could be shown. Figure 3 shows an example,  
489 comparing all varieties to the Guyana Indica variety. The bar chart at the centre shows that all varieties  
490 are significantly different overall (the error between repeats is just about 4% and the closest ED is 6%),  
491 with basmati being closest, followed by OP1001. This new advanced line has a closer quality space  
492 location to the pure Guyana indica than even the European indica, whereas the other 3 advanced lines are  
493 all much more different. The radar plots show the details for the 5 dimensions of the quality space. The  
494 darker lines (full for Guyana, dashed for basmati) show clearly that the main difference between these 2 is  
495 the texture behaviour, with basmati having much lower hardness, stickiness and recovery than Guyana,  
496 indeed, among the lowest in the first 2 cases observed for all varieties, whereas Guyana showed among  
497 the highest hardness and medium stickiness measured. Another significant difference is in the appearance  
498 parameters, where basmati showed the lower width. The OP1001 new variety (dashed grey line) is more  
499 similar to European Indica (full grey line) except in compositional data where it showed among the  
500 highest amylose contents and much lower ash content than European indica, and in texture, where it had a  
501 much lower recovery, lower even than that of basmati. This analysis shows in a very visual way that  
502 OP1001 is an improvement over European indica towards a closer quality behaviour to that of pure  
503 Guyana indica, and also shows the main areas of further improvement: higher water absorption  
504 (especially in cooling) and volume increase, lower moisture content.

505           It might be thought that maybe the other varieties would be similar to japonicas, but that was  
506 actually not the case. The other 3 show significant overall differences to carnaroli, and although they  
507 would be closer to ariete, even a European indica is closer even, so these varieties do not show a  
508 promising quality behaviour compared to any of the existing varieties. This is summarised by the mean  
509 squared deviations to the japonicas in figure 4. It may also be seen in the previous figures showing  
510 locations in the quality space that in many cases these 3 varieties were almost in a class of their own,  
511 significantly different from all others. It is also shown in figure 4 that OP1001 is the closest to a European  
512 indica, as already hinted by figure 3.

513

#### 514 4. Conclusions

515 A quantification of quality attributes of rice varieties was proposed that allows to visually and  
516 speedily compare different varieties, using parameters that can be measured with accuracy and that  
517 represent with conciseness a very complex quality space.

518 The definition of the rice quality profiling was based on grain appearance (grain length, width,  
519 and vitreous whiteness) chemical composition (moisture, ash, amylose), water absorption behaviour  
520 (absorption in soaking and cooking, volume increase), pasting profile (trough, setback and peak time) and  
521 textural attributes (hardness, stickiness, recovery) of different commercial varieties that had been chosen  
522 in order to provide a sufficiently wide variation of the quality parameters. Illustrating the use of this  
523 quality profiling in analysing breeding program outcomes, 4 advanced lines were compared, where it  
524 could be concluded that one of the four (Maçarico, coded OP1001) had unique characteristics with regard  
525 to the others, showing a behaviour that goes from that a European (*Indica*) towards a pure Guyana  
526 (*Indica*). Other advanced lines would be closer to Ariete (*Japonica*), but with many quality traits  
527 completely different from any other variety, so they cannot be expected to perform so well the existing  
528 varieties and for consumer acceptance.

529 This illustrates the applicability of a quality space analysis such as this in identifying which traits  
530 are more likely to produce a rice closer to commercially valuable varieties. This is particularly important  
531 as consumers are typically loyal to specific quality traits of the existing varieties, and often the result of a  
532 breeding program could be a hybrid that satisfies no one for not being either one or another of the existing  
533 varieties. Therefore, it is important to ensure that a new variety has quality behaviour that is sufficiently  
534 close to that of commercially successful varieties.

535

#### 536 Acknowledgements

537 Funding for this research has been received from the Portuguese Fundação para a Ciência e  
538 Tecnologia (FCT) under the grant agreement number RECI/AGR-TEC/0285/2012, BEST-RICE-4-LIFE  
539 project.

540 **References**

- 541 AACC, International. Approved Methods of Analysis, 11th Ed. Method 61-02.01. Determination of the  
542 Pasting Properties of Rice with the Rapid Visco Analyser. Approved November 3, 1999. AACC  
543 International, St. Paul, MN, U.S.A.
- 544 Bao, J., 2014. Genes and QTLs for Rice Grain Quality Improvement. In: Rice - Germplasm, Genetics  
545 and Improvement. Dr. Wengui Yan (Ed.), pp-239-278. ISBN: 978-953-51-1240-2, InTech, DOI:  
546 10.5772/56621. Available from: [http://www.intechopen.com/books/rice-germplasm-genetics-and-](http://www.intechopen.com/books/rice-germplasm-genetics-and-improvement/genes-and-qtls-for-rice-grain-quality-improvement)  
547 [improvement/genes-and-qtls-for-rice-grain-quality-improvement](http://www.intechopen.com/books/rice-germplasm-genetics-and-improvement/genes-and-qtls-for-rice-grain-quality-improvement).
- 548 Bello, M., Baeza, R., Tolaba, M., 2006. Quality characteristics of milled and cooked rice affected by  
549 hydrothermal treatment. J. Food Engg. 72, 124-133.
- 550 Bocevaska, M., Aldabas, I., Andreevska, D., Ilieva, V., 2009. Gelatinization behavior of grains and flour  
551 in relation to physico- chemical properties of milled rice (*Oryza sativa* L.). J. Food Quality. 32,108-  
552 124.
- 553 Borasio, L., 1935. Metodi di Ricerca e di Analisi del Riso, ANNO 9 (11). Tipografia Gallardi. Vercelli,  
554 Itália.
- 555 Calingacion M., Laborte, A., Nelson, A., Resurreccion, J.C., Concepcion, V.D., et al. 2014: Diversity of  
556 global rice markets and the science required for consumer-targeted rice breeding. PLoS ONE. 9,  
557 e85106. doi:10.1371/journal.pone.0085106.
- 558 Champagne, E. T., Bett, K. L., Vinyard, B. T., Webb, B. D., McClung, A. M., Barton, F. E., et al.  
559 (1997). Effects of drying conditions, final moisture content and degree of milling on rice flavor. J.  
560 Cereal Chem. 74, 566–570.
- 561 Champagne, E.T., Wood, D.F., Juliano, B.O., Bechtel, D.B., 2004. The Rice Grain and its Gross  
562 Composition. 3<sup>rd</sup> Ed<sup>n</sup>. In: Champagne E.T (Ed.).Rice: Chemistry and Technology. American  
563 Association of Cereal Chemists, Inc. St. Paul, Minnesota, U.S.A, Pp 77-107.
- 564 Correa, M.J., Ferrero, C., Puppo, C., Brites, C., 2013. Rheological properties of rice-locust bean gum  
565 gels from different rice varieties, Food Hydrocolloid. 31, 383-391.



- 566 Fofana, M., Futakuchi, K., Manful, J.T., Bokossa Y.I., Dossou J., & Bleoussi, R.T.M. (2011). Rice grain  
567 quality: a comparison of imported varieties, local varieties with new varieties adopted in Benin. Food  
568 Control. 22, 1821-1825.
- 569 Ge, X. J., Xing, Y. Z., Xu, C. G., He, Y. Q., 2005. QTL analysis of rice grain elongation, volume  
570 expansion and water absorption using are combinant inbred population. Plant Breeding. 124, 121–  
571 126.
- 572 ISO 6647-2:2015 Rice -- Determination of amylose content -- Part 2: Routine methods
- 573 Juliano, B. O., 1993. Rice in human nutrition. Rome, Italy: FAO (html form freely available from  
574 <http://www.fao.org/inpho/vlibrary/t0567e/T0567E0a.htm>).
- 575 Kong, X., Zhu, P., Sui, Z., Bao, J., 2015. Physicochemical properties of starches from diverse rice  
576 cultivars varying in apparent amylose content and gelatinisation temperature combinations. Food  
577 Chem. 172, 433-440.
- 578 Lawal, O.S., Lapasin, R., Bellich, B., Olayiwola, T.O., Cesaro, A., Yoshimura, M., Nishinari, K., 2011.  
579 Rheology and functional properties of starches isolated from five improved rice varieties from West  
580 Africa. Food Hydrocolloid. 25, 1785-1792.
- 581 Matsue, Y., Ogata, T., 1998. Studies on palatability of rice grown in northern Kyushu – comparison of  
582 palatability between old and new types of rice cultivars. Jpn J. Crop Sci. 67, 312–317.
- 583 Medcalf, D. G., Gilles., K. A., 1965. Wheat Starches .I. Comparison of Physicochemical Properties.  
584 Cereal Chem. 42, 558-568.
- 585 Mohapatra, D., Bal, S., 2006. Cooking quality and instrumental textural attributes of cooked rice for  
586 different milling fractions. J. Food Engg. 73, 253–259.
- 587 Montgomery, D. (2009). Design and Analysis of Experiments, 7<sup>th</sup> Edition. J. Wiley & Sons Inc., New  
588 Jersey (USA)
- 589 Odenigbo, A. M., Ngadi, M., Ejebe, C., Woin, N., Ndindeng, S. A., 2014. Physicochemical, Cooking  
590 Characteristics and Textural Properties of TOX 3145 Milled Rice. J. Food Res. 3, 82-92.

- 591 Oli, P., Ward, R., Adhikari, B., Torley, P., 2014. Parboiled rice: Understanding from a materials science  
592 approach. *J. Food Engg.* 124, 173-183.
- 593 Ross, P.J. (1988). *Taguchi Techniques for Quality Engineering*. McGraw-Hill Inc, New York (USA)
- 594 Sun, J., Hou, C., Zhang, S., 2008. Effect of protein on the rheological properties of rice flour. *J. Food*  
595 *Process. Preserv.* 32, 987–1001.
- 596 Taguchi, G., Chowdhury, S. and Wu, Y. (2005). *Taguchi's Quality Engineering Handbook*. John Wiley  
597 & Sons, Inc. New Jersey (USA)
- 598 Tran, N.A., Daygon, V.D., Resurreccion, A.P., Cuevas, R.P., Corpuz, H.M., Fitzgerald, M.A., 2011. A  
599 single nucleotide polymorphism in the Waxy gene explains a significant component of gel  
600 consistency. *Theor. Appl. Genet.* 123, 519-525.
- 601 Traore, K., McClung, A.M., Fjellstrom, R., & Futakuchi, K. (2011). Diversity in grain physico-chemical  
602 characteristics of West African rice, including NERICA genotypes, as compared to cultivars from the  
603 United States of America. *Int. Res. J. Agric. Sci. Soil Sci.* 1, 435-448.
- 604 Zhou, Z., Robards, K., Helliwell, S., Blanchard, C., 2002. Ageing of stored rice: changes in chemical and  
605 physical attributes. *J. Cereal Sci.* 35, 65–78.

**Table 1**

Homogenous groups and average standard deviations of the values obtained for the 6 and 5 parameters defining compositional data and water absorption behaviour of 11 varieties

<b>Parameter</b>	<b>Number of homogeneous groups</b>	<b>Average standard deviation (normalised)</b>
Amylose	6	0.01
Starch	5	0.0706
Moisture	7	0.0111
Ash	5	0.0889
Fat	5	0.0432
Protein	7	0.029
WVS	8	0.0207
WVC	5	0.0224
WVAC	3	0.0672
VI	4	0.0785
WA	4	0.0989

**Table 2**

Factor loadings of a PCA with normalised varimax rotation, homogenous groups and average SD of the values defining pasting behaviour and texture profile of varieties

	Factor loadings				N. Tukey groups	Normalised st.dev.	Variance explained
	PC1	PC2	PC3	PC4			
<b>Rheological behaviour</b>							
<b><u>VT1</u></b>	<b><u>0.951</u></b>	<b><u>-0.264</u></b>	<b><u>0.053</u></b>		<b>9</b>	<b>0.0128</b>	<b>99.4%</b>
<b>VF</b>	<b><u>0.933</u></b>	-0.067	0.353		7	0.0064	96.6%
<b>VS2</b>	<b><u>0.779</u></b>	0.129	0.584		7	0.0114	89.5%
<b>VB</b>	0.146	<b><u>-0.945</u></b>	0.190		8	0.0125	89.9%
<b><u>VS1</u></b>	<b><u>0.476</u></b>	<b><u>0.825</u></b>	<b><u>0.294</u></b>		<b>9</b>	<b><u>0.0119</u></b>	<b>99.8%</b>
<b>VPT</b>	-0.240	<b><u>0.817</u></b>	-0.358		5	0.0334	74.1%
<b><u>VPt</u></b>	<b><u>-0.315</u></b>	<b><u>0.521</u></b>	<b><u>-0.761</u></b>		<b>6</b>	<b><u>0.0354</u></b>	<b>99.8%</b>
<b>VP1</b>	0.642	-0.748	0.150		8	0.00567	95.8%
<b>% Var.exp</b>	40.0%	39.5%	16.5%				
<b>Texture profile</b>							
<b><u>P1 (hardness)</u></b>	<b><u>0.988</u></b>	<b><u>0.029</u></b>	<b><u>0.072</u></b>	<b><u>-0.077</u></b>	<b>5</b>	<b>0.115</b>	<b>94.10%</b>
<b>P2</b>	<b>0.981</b>	0.032	0.061	-0.150	4	0.130	99.8%
<b>WH2</b>	<b>0.956</b>	0.221	0.174	0.020	4	0.154	97.6%
<b>Ac1</b>	<b>0.955</b>	0.124	0.078	0.075	4	0.143	93.7%
<b>A2</b>	<b>0.945</b>	0.263	0.167	0.048	3	0.148	97.6%
<b>WH1</b>	<b>0.944</b>	0.068	0.057	0.053	4	0.130	92.7%
<b>A1</b>	<b>0.944</b>	0.063	0.058	0.073	4	0.131	92.2%
<b>Gum</b>	<b>0.939</b>	0.159	0.147	-0.069	5	0.156	98.2%
<b>AM</b>	<b>0.931</b>	0.024	0.086	-0.232	5	0.170	96.0%
<b>RW2</b>	<b>0.911</b>	0.349	0.149	0.107	3	0.125	96.4%
<b>RW1</b>	<b>0.899</b>	0.356	0.159	0.161	3	0.122	95.6%
<b>Che</b>	<b>0.838</b>	0.181	0.436	0.015	3	0.109	91.4%
<b>CInd</b>	<b>0.838</b>	0.181	0.436	0.015	3	0.109	91.4%
<b>tP1</b>	<b>0.799</b>	-0.140	-0.006	0.290	3	0.176	91.6%
<b>Adh</b>	0.009	<b>0.937</b>	-0.109	-0.031	3	0.055	82.4%
<b><u>P-1 (stickiness)</u></b>	<b><u>0.082</u></b>	<b><u>0.920</u></b>	<b><u>0.153</u></b>	<b><u>-0.131</u></b>	<b>5</b>	<b>0.053</b>	<b>95.6%</b>
<b>A-1</b>	0.049	<b>0.854</b>	0.254	0.077	6	0.053	97.8%
<b>tP-1</b>	0.432	<b>0.768</b>	0.167	0.036	2	0.084	81.9%
<b>RP-1/1</b>	-0.499	<b>-0.768</b>	-0.226	0.095	5	0.090	99.6%
<b>Spr</b>	0.236	0.186	<b>0.924</b>	0.066	2	0.114	53.4%
<b>Spr.I</b>	0.236	0.186	<b>0.924</b>	0.066	2	0.114	53.4%
<b>RP2/1</b>	-0.016	0.068	-0.104	<b>-0.942</b>	1	0.130	48.0%
<b>ML</b>	0.105	-0.182	-0.277	0.207	1	0.243	49.0%
<b><u>RA-1/1 (recovery)</u></b>	<b><u>-0.171</u></b>	<b><u>-0.741</u></b>	<b><u>-0.270</u></b>	<b><u>-0.089</u></b>	<b>4</b>	<b>0.103</b>	<b>89.6%</b>
<b>Coh</b>	0.215	0.454	0.358	-0.122	5	0.093	76.4%
<b>% Var.exp</b>	48.8%	12.8%	13.6%	9.5%			

Parameters clustered in each PC are highlighted in bold and chosen ones are underlined. The % variance explained column refers to the ability of the 3 parameters chosen to explain the variance of each of the parameters with a simple multilinear model (rheological behaviour) and a quadratic surface (texture profile).

**List of figures:**

- 1) Location of the 11 varieties in the quality space defined with the 3 parameters chosen for water absorption, rheological data and texture of cooked samples.
- 2) Location of the 11 varieties in the quality space defined with the 3 parameters chosen for visual appearance and chemical compositional data.
- 3) Analysis of similarities comparing to pure Guyana (*Indica*).
- 4) Mean squared deviations (ED) of all 15 quality parameters chosen for the 5 dimensions of the quality space between all varieties and the 2 japonicas and the European (*Indica*). (a) Ariete, (b) Carnaroli, (c) European (*Indica*)

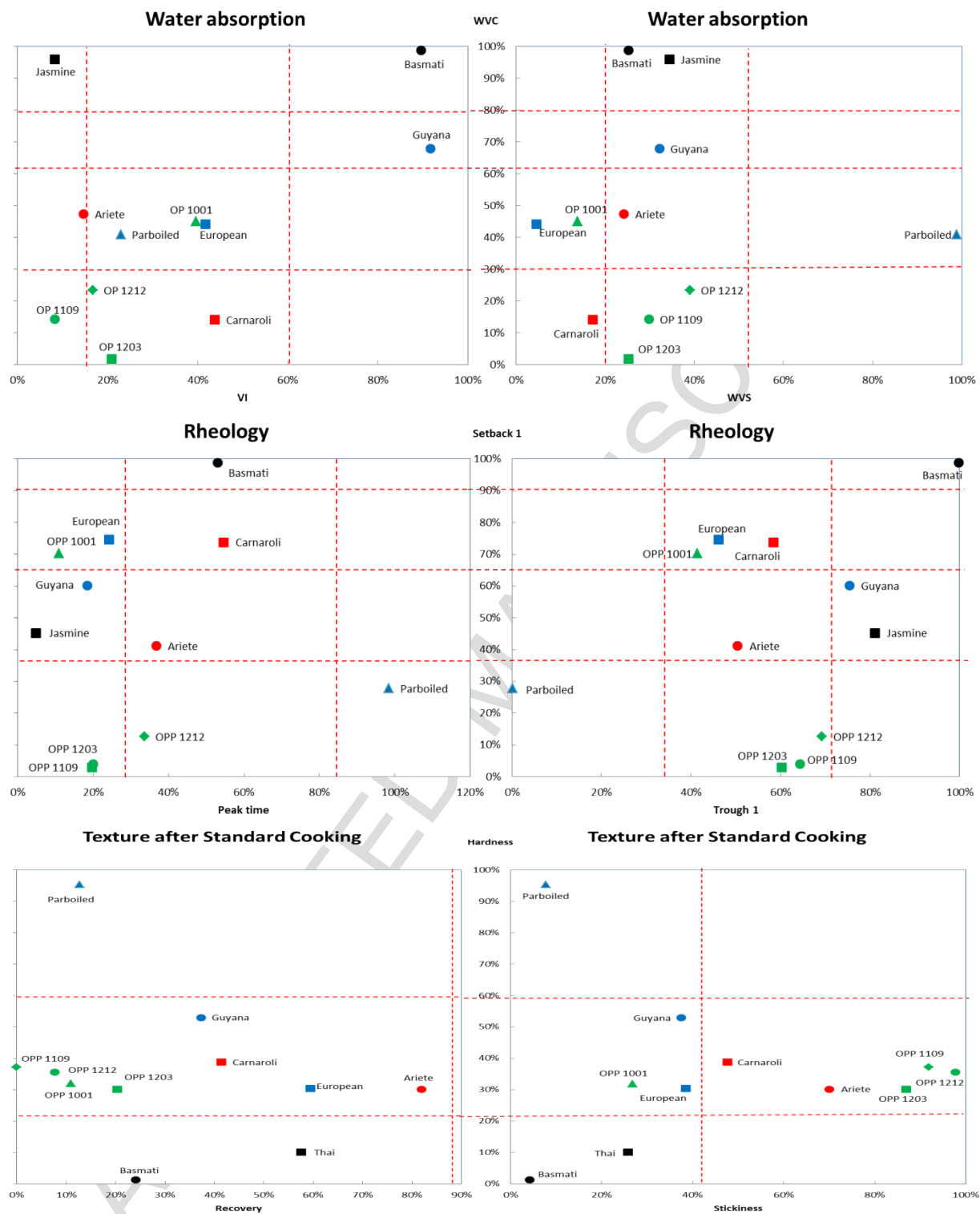


Fig. 1. Location of the 11 varieties in the quality space defined with the 3 parameters chosen for water absorption, rheological data and texture of cooked samples.

(The dotted lines summarise statistically significant differences expressed by the Tukey groups. 0% represents the minimum value observed of the parameter (0.2 for WVS, 1.35 for WVC, 140 for VI, 560 cP for VT1, -439 cP for VS1, 5.53 min for VPt, 351 N for P1, -10.5 N for P-1 and 0 for RA-1/1) and 100% the maximum observed (0.49 for WVS, 2.49 for WVC, 220 for VI, 2479 cP for VT1, 1732 cP for VS1, 7 min for VPt, 1930 N for P1, -74.5 N for P-1 and 0.016 for RA-1/1)

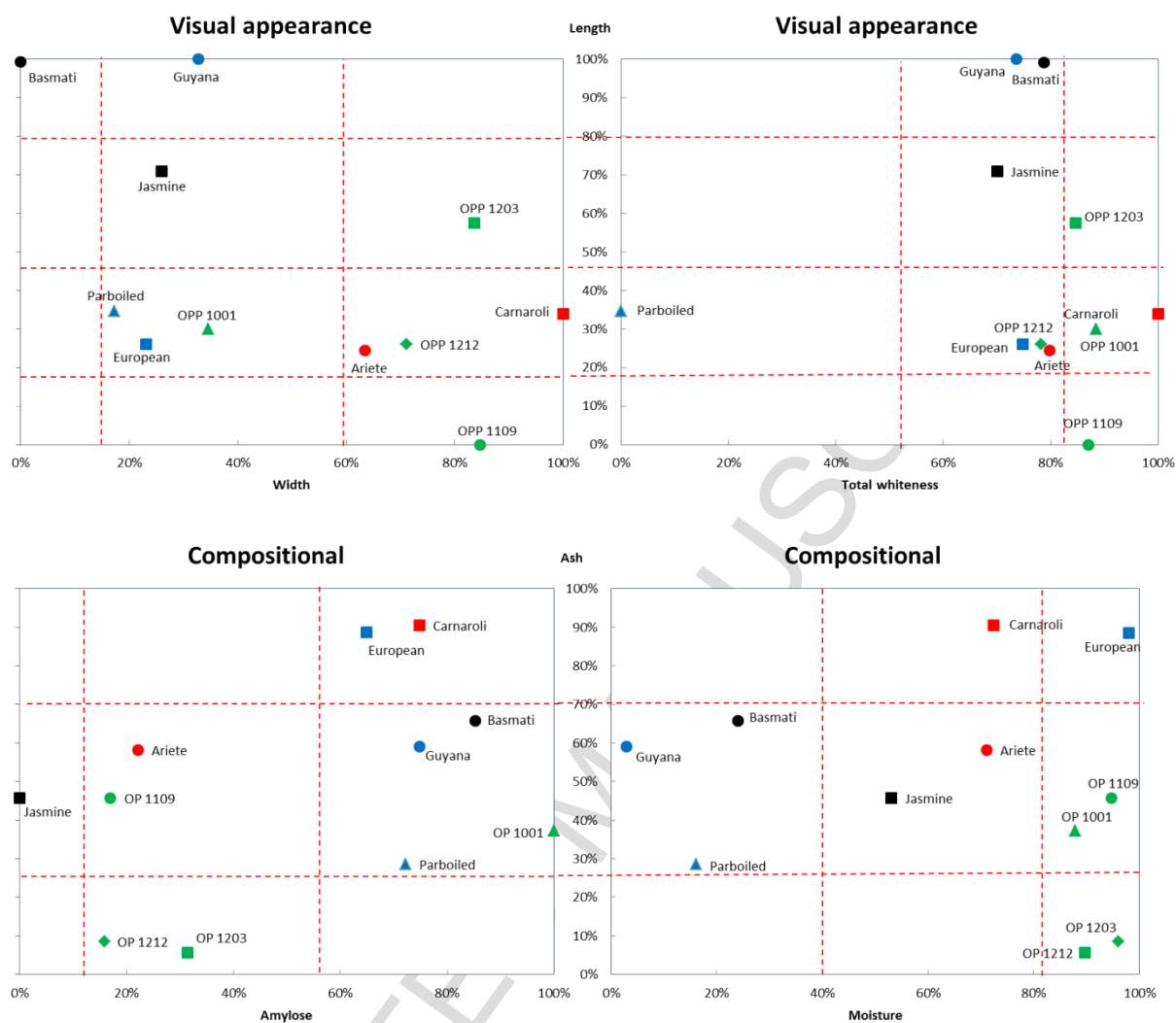


Fig. 2. Location of the 11 varieties in the quality space defined with the 3 parameters chosen for visual appearance and chemical compositional data.

(The dotted lines summarise statistically significant differences expressed by the Tukey groups. 0% represents the minimum value observed of the parameter (5.56 for L, 1.58 for W, 76.99 for TW, 16.7 for amylose, 12.7 for moisture and 0.29 for ash) and 100% the maximum observed (6.83 for L, 2.62 for W, 163.69 for TW, 33.7 for amylose, 15.3 for moisture and 0.64 for ash))

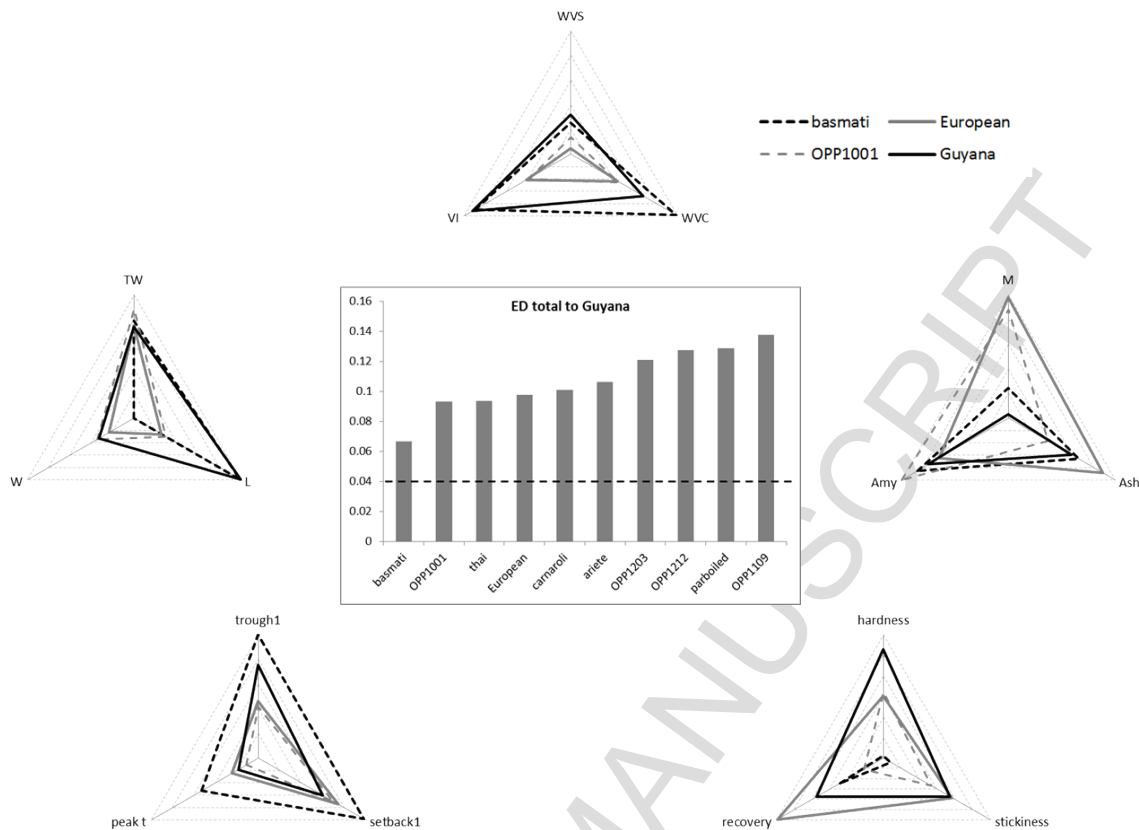


Fig. 3. Analysis of similarities comparing to pure Guyana (*Indica*).

(The dashed line in the bar plot indicates the maximum ED (mean squared deviation) between repeats. The radar plots show the average scores of 4 of the varieties being compared in detail (legend on the top right) in the 5 dimensions of the quality space defined. The radar plot gridlines mark 20% of the normalised scale, from minimum observed (0%) to maximum observed (100%) of the respective quality parameter.)



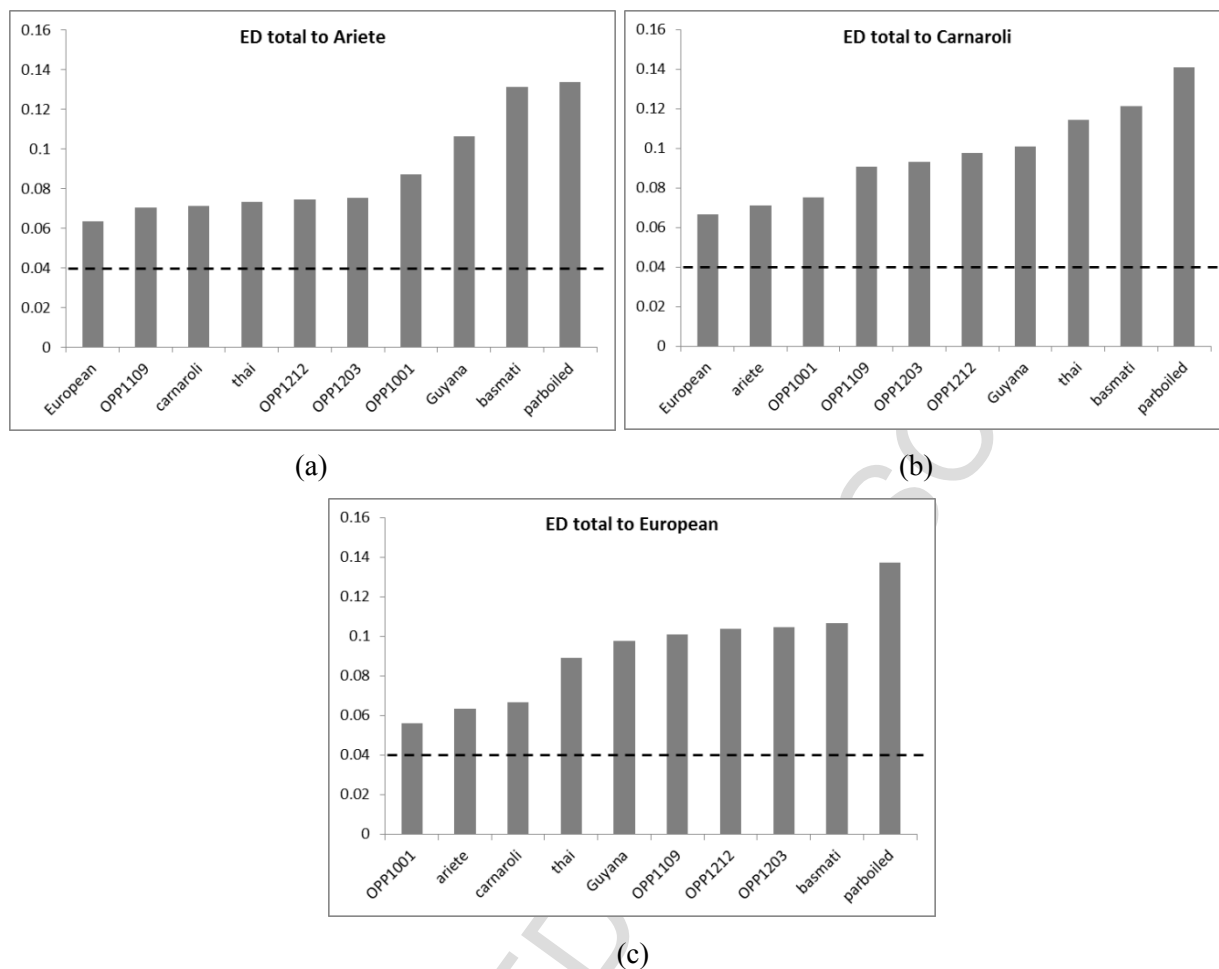


Fig. 4. Mean squared deviations (ED) of all 15 quality parameters chosen for the 5 dimensions of the quality space between all varieties and the 2 *Japonicas* and the European (*Indica*). (a) Ariete, (b) Carnaroli, (c) European (*Indica*)