Valcárcel, José Vicente; Peiró, Rosa María; Pérez-de-Castro Ana; Díez, María José

Morphological characterization of the cucumber (*Cucumis sativus* L.) collection of the COMAV's genebank

Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV), Universitat Poltècnica de València, Camino de Vera s/n, 46022 Valencia, SPAIN

Corresponding author: Díez MJ

mdiezni@btc.upv.es

Telephone number: 34 963879421

ORCID Díez, MJ: 0000 0002 1422 2144

ORCID Pérez-de-Castro, A: 0000-0002-4949-3323

ORCID Peiró, RM: 0000-0002-3009-2343

ABSTRACT

The cucumber (*Cucumis sativus* L.) is an important crop worldwide. In the present study the morphological characterization of 206 cucumber accessions, 195 from Spain and 11 outgroups from other countries, was carried out. One hundred and seventy eight of them came from the COMAV's Genebank, 116 collected by the COMAV and the others 62 maintained at this institution as safety duplicates of the BGHZ collection. Seventeen more accessions supplied by BGHZ were included in the present research. Five plants per accession were characterized, with 17 qualitative and nine quantitative descriptors, eight of them referred to plant traits and 18 related to the fruit. Fruit descriptors were evaluated in at least 25 fruits per accession. The accessions were classified in five groups: 'White', 'Short', 'French', 'Long' and 'Very long', based on the morphology of their fruits and their similarity to commercial types. Principal Component Analysis showed that, with few exceptions, the accessions grouped to the previously established groups. Variability found among and within groups displayed the potential of these plant materials in breeding programs for different traits. The morphological characterization allowed the selection of the 67.2 % of the collection, eliminating the most similar accessions.

Keywords: cucumber landraces, genebank rationalization, phenotyping.

INTRODUCTION

The cucumber (*Cucumis sativus* L.) belongs to the Cucurbitaceae family which comprises about 960 species (Jeffrey 2005; Schaefer *et al*., 2009). The genus *Cucumis* is considered of Asian origin (Schaefer *et al*. 2009). *Cucumis sativus* (2n = 14) was likely domesticated in the Indian subcontinent (Sebastian *et al*. 2010) and China is considered a secondary centre of diversity for this species. Cucumber's dissemination westward from India started very early. It was brought to Greece and Italy by the Romans by the 2nd century BC where became very popular as a food and medicinal product. The Romans took it all over Europe, and it was introduced into America by the Spanish settlers. Nowadays, cucumbers are cultivated in nearly all countries in temperate zones. It is cultivated either for fresh consumption, as slicing cucumber, or as pickling cucumber, marinated with vinegar, salt, dill or other spices (Staub *et al*. 2008).

China is by far the first country producing cucumber in the world, followed by Russian Federation and Iran. Spain occupies the third place regarding cucumber yield in Europe after the Russian Federation and Ukraine (FAOSTAT 2017). In Spain, cucumber production has been increasing during the last 15 years, although with fluctuations. Approximately, the 90% of the Spanish production of cucumbers comes from Andalusia, mainly from Almeria (53.2%) and Granada (33.2%) provinces; they are also grown in the Canary Islands (4.9%), Madrid (1.5%), Catalonia (1.4%), Murcia (1.1%), the Valencian Community (0.7%) and Extremadura (0.7%) and to a lesser extent in other Communities (MAPAMA 2017). The cultivation of cucumbers for early and late season production in Mediterranean countries such as Southern Spain takes place in unheated greenhouses during the winter months. As the mild winter temperatures are still not optimal, the varieties used must be able to grow and produce fruits at suboptimal temperatures.

Fruit skin traits including spine size and colour, fruit warts, dull and uniform colour are some of the most important external quality traits together with the size and shape that determine the commercial types. All these traits are related to the market value of cucumber (Lower and Edwards 1986; Pollack 2001).

Most cultivated varieties of cucumber are hybrids. They can be grouped into the following types: short cucumber or 'Spanish' type, medium length cucumber or 'Slicer' or 'French' type and long cucumber or 'Dutch' type, also called 'Almeria' type or 'European' type. Short cucumbers ('Spanish' type) are small-fruited varieties (maximum fruit length 15 cm), oblong-ellipsoid with rounded ends, with green skin and yellow or white stripes. They can be used for fresh consumption or for pickling. 'Slicer' or 'French' type includes varieties with fruits of medium length (20-25 cm). There are different types of 'Slicer' cucumber, those whose fruits have spines and other which have a smooth pericarp or minicucumber (similar to

77 'Almeria' type but shorter). Usually they produce parthenocarpic fruits. Long cucumber ('Dutch' type) are
78 varieties with spineless fruits exceeding 25 cm in length. Its thin and tender skin contributes to its excellent
79 eating quality, but causes quick dehydration (Rubatzky and Yamaguchi 1997). They are often found in
80 supermarkets shrink-wrapped in plastic. This type of cucumbers is usually grown in greenhouses. The leaf
81 size is much larger than the ones of the other types.

82 Around 80% of the cucumber grown in Spain is 'Dutch' type. Although it is also consumed in Spain,
83 most of it (90-95%) is destined for export to countries like Germany and United Kingdom. 'Spanish' type
84 represents 10% of the Spanish production and it is traditionally consumed in the domestic market. The rest
85 of the production belongs to the 'French' type.

86 Apart from the commercial scale production, many high quality landraces still exist in Spain, which
87 have been cultivated by farmers for centuries for self-consumption and to be sold in local markets. These
88 landraces constitute a valuable resource for cucumber breeders to increase the variability concerning
89 quantitative and qualitatively inherited characters and adaptation to specific growing conditions. Thus, this
90 germplasm can be used as a source of genes to develop new cucumber varieties.

91 Many germplasm collections of cucumber exist along the world. According to the germplasm database
92 GENESYS, the number of cucumber accessions stored in the genebanks worldwide is estimated in 9,808,
93 although some important collections (China, India and Japan) are not included in this database
94 (https://www.genesys-pgr.org/es/welcome). In the European genebanks, according to the European Central
95 Cucurbits Database consulted (Díez *et al*. 2008) in March 2017, the number is 7,177. The Spanish National
96 Inventory on Genetic Resources estimates in 401 the number of accessions stored in Spanish genebanks.
97 The two major collections of cucumber in Spain are the one stored at the Banco de Hortícolas de Zaragoza
98 (Vegetables Genebank BGHZ, Zaragoza, Spain) and the one conserved at the Instituto de Conservación y
99 Mejora de la Agrodiversidad Valenciana (Institute of Conservation and Improvement of Valencian
100 Agrodiversity, COMAV, Valencia, Spain). The first one holds a collection of 403 accessions, 248 being of
101 Spanish origin, and the second one conserves a collection of 217 accessions, of which 198 come from
102 Spain. Both collections are complementary. Northern Spanish Autonomous Communities are better
103 represented in the BGHZ collection, while Southeastern Communities are in the COMAV collection
104 (http://wwwx.inia.es/inventarionacional/). During several years both genebanks have participated in
105 projects funded by the National Government with the objective of regenerate and characterize their
106 collections. As a measure of safety many accessions are duplicated in both collections.

107 Genetic resources have been proven to be essential for plant breeding. However, their use depends on
108 the degree of knowledge of the stored materials as well as on an efficient system of databases that allow
109 the access to the information by breeders and other users. Also, the efficiency on the management of the
110 collections can be improved by rationalizing them through the characterization, both with morphological
111 and molecular data, and evaluation of the materials. In order to rationalize the collection and identify
112 duplicates, the COMAV has conducted phenotyping assays as a first step. We have found a high variability
113 in fruit and agronomic attributes than can be used in breeding programs for external appearance, fulfilling
114 the requirements of different commercial types, and for adaptation to specific growing conditions.

115 **MATERIAL AND METHODS**

116 *Material*

117 Two hundred and six accessions of cucumber were studied in this assay (Table S1). One hundred and
118 seventy eight of them came from the COMAV's Genebank, 116 collected by the COMAV and the others
119 62 maintained at this institution as safety duplicates of the BGHZ collection. Seventeen accessions were
120 provided by the BGHZ and six and five were supplied by the Center for Genetic Resources, The
121 Netherlands, CGN, and the Chinese Academy of Agricultural Science, China, CAAS, respectively and used
122 as outgroups in the analysis. The 73% of the Spanish accessions were collected from 1981 to 1992, being

the most active years in collecting expeditions from 1984 to 1986. In this three-year period there were collected 51% of the accessions evaluated in the present work.

Five plants per accession were cultivated in glasshouses at the COMAV (GPS coordinates: 39º28´46´´N; 0º20´06´´W) during the autumn season. The accessions were randomly distributed in the glasshouse. Plants were cultivated in 20 L pots, filled with coco fiber, trained with stakes and pruned to two stems. To increase the efficiency of fruit set bumblebees were used. Phytosanitary treatments were applied when needed. Fruits were harvested at commercial market stage.

Descriptors used for characterization

A total of 17 qualitative and nine quantitative descriptors based on the ones developed by International Plant Genetic Resources Institute (IPGRI; Esquinas-Alcázar and Gulick 1983) and International Union for the Protection of New Varieties of Plants (UPOV 2007) were recorded (Table 1). Most descriptors were modified according to the variation observed in the Spanish collection. Eight of them corresponded to plant characteristics and 18 were taken on the fruits. Quantitative plant descriptors were measured at plant level and quantitative fruits descriptors were recorded in at least 25 fruits per accession.

Data analysis

Accessions were separated in groups after a visual inspection according to their similarity to commercial types or based on some phenotypic outstanding characteristics, mainly skin colour and fruit length. The groups established were: 'White' (four accessions), being the common characteristic of this group the skin colour; 'French' (21 accessions), including the accessions with smooth dark green skin; 'Short' (101 accessions), characterized by fruits with a mean length of 15 cm; 'Long' (70 accessions), with fruits of about 20 cm, and 'Very long' (10 accessions), with fruits longer and with a mean value of 27 cm. Inside each group a great variability was observed for many traits and hence sub-groups were established according to them.

To analyze jointly qualitative and quantitative variables, qualitative characters were expressed on a quantitative scale in which the highest value represented the highest intensity of the character. Afterward, a PCA (Principal Component Analysis) was used to examine association between traits and show the similarity among accessions. The total variation explained was calculated as the sum of extracted eigenvalues. For this analysis, the software NTSYS 2.02 was used.

Additionally, histograms were constructed with the qualitative traits. For the quantitative traits, descriptive statistics, such as mean, standard deviation, minimum and maximum values and coefficient of variation were calculated for the whole collection and for each of the groups (Little and Hills 1978).

RESULTS

Correlations among traits

Phenotypic correlation coefficients measured among the traits were significant in many cases ($P<0.01$). The highest correlations were found among fruit characteristics, including qualitative and quantitative ones (Figure 1). The highest correlation was observed between length/width rate (Lewirate) and fruit length (Frlength), with a value of r=0.96. Lewirate was the trait with the highest number of correlations with other traits. A positive correlation was found between Lewirate with Frlength (r=0.96), fruit shape (Frshape, r=0. 67), fruit predominant shape at blossom end (Blossend; r=0.55) and fruit weight (Frweight; r=0.51) and negative with fruit width (Frwidth; r=-0.65), core diameter (Cordiam; r=-0.54) and fruit predominant shape at stem end (Stemend; r=-0.52). Frlength showed also positive correlation with Blossend (r=0.50), Frshape and Frweight (r=0.70 for both) and negative correlations with Stemend (r=-0.51). Frweight showed also positive correlations with intensity of fruit texture (Inttext) and Frshape (r=0.58 and 0.51, respectively). Flowering and harvest stage earliness were also positively correlated

(r=0.85) and positive correlations between other traits related to presence and distribution of dots, type of vestiture and colour were also found (Figure 1).

Principal Component Analysis

Principal Component Analysis based on standardized phenotypic means was done using qualitative and quantitative traits to show which of them accounted for the greater variability and were the most discriminant among accessions. Nine accessions were excluded of the analysis for missing data on flowering and harvest stage earliness traits. The first three principal components (PCs) of the PCA accounted for 50.5% of the total variation observed. The traits more correlated with the first component (values higher than 0.70) were those related with the fruit. Fruit length (Frlength) and length/width ratio (Lewirate) showed a negative correlation whereas fruit width (Frwidth) and core diameter (Cordiam) showed a positive correlation (Table 2). The second component was positively correlated with other two fruit characteristics, Inttext and Frweight. The flowering and commercial harvest earliness (Floearly, Fruearly), the type of vestiture (Typeves), and the density of texture (Denstext) were correlated with the third principal component, although with lower values ranging from 0.50 to 0.70.

The scatterplot of accessions onto the two main components allowed the grouping of the accessions according to their phenotypic similarities (Figure 2A). Accessions of 'French' type were located separately mainly due to its characteristic soft fruit skin texture, as the trait Intext was highly correlated with the second PC. Most part of the accessions belonging to 'Very long' type were also separated from the rest due to their longer fruits. The four accessions with white skin were spread and grouped either with the accessions of the 'Long' or to the 'Short' types according to the length of their fruits. 'Short' and 'Long' accessions were distributed along the second PC, some of them overlapping in the middle. The accessions included in this study like outgroups spread by all graphic, according to their morphological characteristics. Thus, several accessions coming from China (VL234, VL235 and VL237), with very long fruits ('Oriental Chinese Long' type, up to 35 cm) were located highly separated from the rest, while the accession coming from The Netherlands (Shs228) positioned in the low part of the graphic due its soft skin, similar to the accessions of 'French' type. The highly remarkable differences between the 'French' and 'Very long' types with respect to the rest of the accessions gave rise to a partial grouping of the 'Short' and 'Long' types, make more difficult a detailed analysis of these two groups. For this reason, a new PCA including only the 'White', 'Short' and 'Long' types was performed.

In the PCA carried out including only the 'White', 'Short' and 'Long' types, the first three PCs described 45.4% of the total variation. The first PC was positively and highly associated with traits related to the size and shape of the fruits, Frlength, Lewirate, Frshape and Frweight (Table 2B). The second PC was positive correlated with the female flowering and commercial harvest earliness (Floearly and Fruearly, respectively). In this analysis, the accessions with short and long fruits separated along the first PC, having inside each group a remarkable variability for earliness traits (Figure 2B). Although the mean values of Floearly and Fruearly were very similar for the groups 'Short' and 'Long' (Table 3), accessions within the 'Short' group showed a higher variability for this characteristic specially for the commercial harvest stage earliness (Fruearly), ranging from 48.2 to 81.0 days in the 'Short' group and from 51.0 to 70.7 in the 'Long' one. On the other hand, the spread of both groups along the first PC showed the high variation for fruit characteristics related to fruit size and shape. The four accessions of 'White' type appeared grouped with the 'Short' or 'Long' groups according to the length of their fruits.

The results shown by the PCAs supported the groups made after the initial visual inspection of the studied accessions except for the 'White' group. However, due to the remarkable characteristic of the fruit colour of the accessions belonging to this group, we decided maintain it as a separate group for further descriptions and analysis. Accessions representative of each group are shown in Figure 3.

Description of the variability of the whole set of accessions and of the established groups

The group of accessions studied exhibited some common characteristics. All accessions were monoecious, of indeterminate growth type and held fruits with tubercular skin texture except for one accession included in the present work as an outgroup that had smooth skin. Pistillate flowers did not develop on the main stem in most part of the accessions. The exceptions were the five accessions of Chinese origin. The phenotypic characterization of Spanish accessions showed the great variation for most of the fruit and agronomic traits recorded (Table 3). Some of the most variable traits were the length/width range (from 0.3 to 0.6), the fruit weight, ranging from 122.5 g for a cucumber accession of 'Short' type to 298.6 g for an accession of 'Very Long' type, the length of the fruit (ranging from 12.5 cm to 25.6 cm), the harvest stage earliness (measured as the number of days between the transplant and the harvest of the first fruit at marketable stage), with a variation of 47.0 to 81.0 days and the number of days from anthesis to commercial harvest (from 11.8 to 28.7 days). When considering the whole group of accessions, the phenotypic coefficient of variation (CV) values ranged from 6.3 for Frwidth to 23.8 for Lewirate (Table 3). Excluding the accessions used as outgroup of the analysis, the CV values for the Spanish accessions decreased only for Lewirate (from 23.8 to 18.1) and for Frlength (from 18.7 to 14.3) as expected, considering that many of the accessions included in the outgroup were varieties of oriental origin with very long fruits. The accessions studied were landraces and, consequently, some extent of variation intra-accession was observed. However, the CV values of quantitative traits (weight, length, width, core diameter and rate length/width) after excluding the outliers, were in the range of other works in all cases and for all traits, ranging from 0.04 to 0.41 (data not shown). Some variation about presence, size and colour of the spines and on the presence and length of stripes was observed. This variation was considered as normal in landraces and accessions were not sub-divided.

With the aim of identifying the main differences among groups and to study the variability inside each group, frequency histograms with the qualitative traits and a statistical description with the quantitative ones were done for each group (Online resource S1, Table 3). Here, we describe the most outstanding results.

'White' type. The group called 'White' was made up of four accessions, including one accession that belongs to the outgroup from Sri Lanka. Fruits were very variable in weight and length for the Spanish accessions, ranging from 156.3 g to 192.6 g per fruit and 14.3 to 18.6 cm in length (Table 3). The common characteristics of this group were the white-cream colour of the skin, obtuse fruit stem end and absence of stripes. However, the accessions differed in many other characteristics, mainly in spines colour, predominant shape at blossom end, fruit length, presence and density of dots, intensity of fruit texture and dot distribution (Online Resource S1). In spite of the low number of accessions of this type, it was found a high variability in harvest stage earliness (from 47.0 to 65.6 days). The high variability found in this group is reflected by the highest values of CV for the traits Frlength, Lewirate, Floearly and Fruearly among the different groups established, even higher than the CV values for the whole group of Spanish accessions (Table 3). Two out of the three Spanish accessions had one flower per node.

The 20 accessions included in the 'French' type had fruits quite uniform, elliptical elongate, predominantly obtuse at the stem end at rounded at the blossom end, dark green tubercular skin, with spines and superficial and medium density warts. Spines were predominantly white, but some accessions had black ones. Stripes and dots were variable among the accessions. Fruits were of medium size, with a mean of 189.8 g and a length of 18.7 cm, similar to that of the 'Long' type. All accessions had one flower per node. This group was split in two sub-groups, one of them including the accessions more closely similar to the 'French' type (Fty, 14 accessions) and another with the accessions similar to the 'French' type but with certain variations in the size and colour of their fruits (Fnt, six accessions). In spite of the visible differences both types distributed together in the PCA performed (data not shown). The CV for the quantitative traits were moderate, especially for the traits related to fruit length (Frlength), width (Frwidth) and core diameter (Cordiam) and the two traits related to earliness (Table 3).

Fruits of 'Short' type showed an average of 189.9 g weight and 15.4 cm in length. It is the most numerous group including 100 accessions. The most common traits in this group were the fruit shape at stem and blossom ends, obtuse and truncate respectively, medium green skin, tubercular skin texture, black

spines, and sparse dots (Online Resource S1). Most of the accessions having more than one flower per node were included in this group, 16 accessions had all plants with more than one flower per node and 12 accessions showed variation among plants for this characteristic. Due to the high variability found for the other traits some sub-divisions were made up according to the fruit shape, skin texture, and type of vestiture (presence of hairs and spines or only spines). The sub-types established were named 'Short typical' (Sty, 44 accessions), 'Short tubercular pronounced' (Stp, 33 accessions), 'Short ellipsoid fruit' (Sel, 12 accessions) and 'Short hairs and spines' (Shs, 11 accessions). We conducted a new PCA with the accessions included in this type (Figure 2C). The three first PCs explained the 45.1% of the variation, being the fruit traits the most correlated with the first PC, some of them positively related like length (Frlength), length/width rate (Lewirate), intensity of texture (Inttext) and other ones with a negative correlation like type of vestiture (Typeves) and density of texture (Denstext). The second PC was positively correlated with the fruit width (Frwidth) and the core diameter (Cordiam). In the projection of the accessions onto the two first PCs there was an apparent separation 'between the sub-types Sel and Shs on one side, and the sub-types Stp and Sty on the other. The two groups (Sty+Stp versus Sel+Shs) were primarily distinguished by the type of vestiture (Typeves) and intensity (Inttext) and density of texture (Denstext). The earliness of harvest stage (Fruearly) was also different between both groups, being earlier the group Stp+Sty (data not shown). Although four sub-groups were established, the PCA conducted separated the 'Short' type accessions in two groups. Stripes and dots were very variable in presence, density and length in all sub-groups. Regarding the coefficients of variation, the values for the traits related to fruit size were low (4.0 for Frwidth to 7.8 for Frlength), as the accessions included in this group were selected for the characteristics of their fruits. The CV for the other traits were higher but with moderate values, ranging from 11.0 for Floearly to 15.3 for Flowfru (Table 3).

'Long' type. This type included 68 accessions. Fruits had a medium length of 19.0 cm and a width of 49.3 mm (Table 3). In general, fruit shape was elliptical elongate, with medium green skin, predominantly with spines of black colour and tubercular. Most plants of these accessions had one flower per node, only four accessions had two flowers per node in all their plants. The accessions included in this group were subdivided in groups according to their differences in skin texture, being tubercular pronounced (Ltp, 34 accessions) in 50% of the accessions and softer (Lti) in 32 accessions. Two accessions were different having hairs and spines (Lhs) and one had soft skin without hairs or spines (Lsf). Other traits like size, stripes, dots and fruit shape showed a high variability in the Ltp and Lti sub-groups. The 'Long' group showed lower variability than the other groups (Table 3). Traits related to fruit had CV values below 10.5. The highest value was for the period of development of fruits, measured as the number of days between the anthesis of the flower and the harvest of the fruit at marketable stage, Flowfru.

'Very long' type. Four accessions were from Spain and six from Asia (five from China and the other one from Japan). It is the most heterogeneous group, with a remarkable variability for fruit shape, colour, type of vestiture, stripes and dots. Three out of five Chinese accessions included in this group, VL234, VL235 and VL237, were of Oriental Chinese Long type, with very long and thin fruits, dense spines and furrowed. Besides, the accession VL237 was the only accession with all plants producing two flowers per node. The accession VL227 from Japan and VL238 from China had a high density of yellow dots from the blossom end until near the stem end, giving a tonality greenish-yellow to the fruit skin (Figure 3). The Chinese VL236 was more similar to the 'Long' type. Considering the Spanish accessions, fruits showed an average length and width of 24.2 and 43.3 mm, respectively. The Spanish accession VL7 was similar to Chinese type and accessions VL151 and VL163 to Spanish type, although they differed in some traits. The accession VL151 had darker green and softer skin with very weak stripes and the accession VL163 had strong tubercular skin with marked stripes. The high variability obtained both in the Spanish and the whole group was reflected in the high values of CV.

Selection of accessions for further molecular analysis

The morphological characterization conducted has provided valuable information about the differences and similarities of the studied accessions, and the different types they can be grouped. It also allows the identification of the most similar accessions which can be discarded as a first step to the rationalization of

313 the collection. The remaining accessions should be molecularly characterized and this information, together
314 with the one coming from the phenotyping assay, can be used to rationalize the collection.

315 Based on the similarity of the accessions and their origin, a 32.8% of the accessions have been excluded
316 for the molecular analysis. The percentages of excluded accessions varied among groups, ranging from the
317 23.5% of the Ltp sub-type and the 75% of the VL type. The phenotypic values of the 131 Spanish selected
318 accessions were similar to the whole Spanish collection (195 accessions), including mean, range and
319 coefficient of variation for the studied quantitative traits, and also for the qualitative ones (data not shown).

320 **DISCUSION**

321 Breeding vegetable crops is a very competitive activity and it is subject to the strict varietal
322 requirements of the market. The commercial varieties developed have a very short life and have to fit with
323 high quality standards, both for agronomical characteristics and also for the characteristics that define each
324 variety. Genetic variability is essential for succeeding in plant breeding and this variability can be found in
325 genebanks. The Genebank of the COMAV holds a cucumber collection with high variability for fruit
326 characteristics and also for adaptation to different agroclimatic conditions and soil characteristics of all
327 growing Spanish areas of this crop.

328 The descriptors used to perform the morphological characterization have allowed the grouping of the
329 accessions according to their phenotypic similarity, and they have been useful to identify the variability
330 content in each group. According to the PCAs, the traits more variable considering the complete collection
331 are the ones related to shape and size of the fruit: length, width, length/width ratio, and the core diameter.
332 According to the second component, the fruit weight and the intensity of texture are also very important.
333 When considering the two predominant types in our collection grouped as 'Short' and 'Long' ones, again
334 the traits related to weight and shape of fruit are keys to explain the variability found among the accessions
335 included in each group. There was also a considerable variability for earliness, an economically important
336 trait. The 'Short' group was the most diverse, being responsible for the great variability found in traits
337 related to the type of vestiture, intensity and density of texture, and also for earliness. The most outstanding
338 characteristics that allowed to divide all accessions in two groups ('Short typical' + 'Short tubercular
339 pronounced' and 'Short hairs and spines' + 'Short ellipsoid') were the vestiture and texture of fruits and
340 the length and rate length/width. Besides, for these two sub-types, a considerable variation for width and
341 core diameter was present. This allows breeders to combine the different characteristics to fit the
342 requirements of each varietal type. The 'Long' type is not as variable as the 'Short' one. However,
343 differences also in length, length/width rate and core diameter explain the phenotypic variation of fruits.
344 Also the earliness was very variable. Several authors have carried out characterization trials with different
345 types of materials, including landraces, commercial hybrids and breeding lines. The objective was in many
346 cases to identify the best genotypes to be used in breeding programs aimed at increasing yield or to study
347 correlations between yield and other traits to be used in indirect selection. Thus, Gaikwad *et al*. (2011)
348 working with 18 genotypes collected in India found comparable variation range to the ones obtained in our
349 study for days to appearance to first female flower (39 to 59), fruit weight (127 to 237 g), fruit length (11,5
350 cm to 27,8 cm) and fruit diameter (3.5 cm to 5.2 cm). In the same way, Golabadi *et al*. (2012) studying a
351 group of 20 varieties coming from several countries found similar levels of variation for fruit traits. Khan
352 *et al*. (2015) selected a set of 24 genotypes to represent the maximum variation and obtained ranges of
353 variation similar to the ones obtained in our study (days to edible maturity: 57.5 to 71.5; fruit length: 12.8
354 to 20.4; fruit width: 4.4 to 6.4). In our case, the set of accessions studied were not selected for a specific
355 purpose, it was the collection of the Spanish landraces of our genebank. The amount of variability found
356 by other authors between groups of selected accessions according to diverse criteria was comparable to the
357 one detected in our collection highlighting the interest of the Spanish materials.

358 The high correlation values among fruit traits reflect the main types of cucumber. It is outstanding the
359 cluster of high correlation among fruit weight (Frweight), fruit shape (Frshape), fruit length (Frlength),
360 length/width rate (Lewirate) and texture intensity (Inttext). The highest correlation value was between

Lewirate and Frlength followed by the correlation between Frlength and Frweight (r=0.70; P<0.01). The low correlation between Frweight and Frwidth (r=0.23; data not shown), indicates that the weight of the fruit is determined in a greater extent by its length than by its width. The results obtained by Innark *et al*. (2013) and by Khan *et al*. (2015) are on the same line. They found a correlation of 0.92 and 0.60, respectively, between fruit length and fruit weight, supporting our results. In our assay, the heaviest fruits were the elliptical elongate shaped and also the ones with the most pronounced texture ($r_{(Frweight - Inttext)}$ = 0.58). Interestingly this description corresponds with the fruits included in the 'Long' type. Consequently, the fruits included in the 'Short' type have the alternate values for these characteristics, shorter, oblong ellipsoidal fruits and with softer skin. Additionally, the low correlation values found among the other traits indicated that for the defined groups it exists a high variability including traits related to fruit and earliness.

The coefficient of variation compares the relative amount of variability between crop plant traits. In the present assay, differences in the CV values of traits were observed, being the CV for Lewirate the highest (23.8) and the one for Frwidth the lowest (6.3). Gaikwad *et al*. (2011) and Yadav *et al*. (2012), working with a group of 18 landraces of Asian origin and 20 experimental materials respectively, found similar range of variation for fruit traits as well as similar values of CV. These results imply that for Lewirate, Frlength, Frweight and Flowfru, the traits with the highest CV values, have higher amounts of exploitable variability among these attributes and greater potential for advances compared to others with lower values. Conversely, the lowest CV recorded for Frwidth indicates low exploitable variability for this trait. However, these values change according to the established groups. So, the conclusions have to be referred to each group in particular.

The collection stored at the genebank of the COMAV comes from a big proportion of the Spanish provinces (39 out of 50), which include climates with extreme temperatures like the ones reached in the center of Spain (Castilla La Mancha, Castile and Leon, Madrid, Extremadura), Atlantic climate like the provinces of the Cantabrian Coast (Galicia, Cantabria, Basque Country), and the Mediterranean ones (Catalonia, Valencian Community, Murcia and part of Andalusia), as well as the colder climates of Aragon, La Rioja and Navarra. The adaptation of these local varieties to these different environmental conditions added a specific interest to this collection. Additionally, the adaptation to specific conditions like the rich limestone soil of the locality of Huete (Cuenca) has given rise to a high quality cucumber called 'Pepino de Huete', of soft flavor and crispy texture, highly appreciated in this area. The accession Sty239 belongs to this type.

The maintenance of this collection is critical, as the high variability stored in it is not conserved in other European collections and it is a valuable source of genes of interest for cucumber breeding.

ACKNOWLEDGEMENTS

This work has been partially funded by the projects PAID-06-10-2408 (Universitat Politècnica de València) and RFP2013-00011-00-00 (INIA, Ministerio de Ciencia e Innovación). The authors acknowledge E. Solbes, J. Torres, E. Muñoz and A.Rodríguez for their technical assistance.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

402 **REFERENCES**

403 Díez MJ, van Dooijeweert W, Maggioni L, Lipman E (compilers.) (2008) First meeting of the ECPGR
404 working group on cucurbits. 1-2 September 2005, Plovdiv, Bulgaria. Bioversity International,
405 Rome, Italy:67 pp

406 Esquinas-Alcázar J, Gulick PJ (1983). Genetic resources of cucurbitaceae. International Board for Plant
407 Genetic Resources:101 pp

408 FAOSTAT (2017) No Title. Retrieved February 27, 2017. http://faostat3.fao.org

409 Gaikwad AG, Musmade AM, Dhumal SS, Sonawane HG (2011) Variability studies in cucumber (*Cucumis*
410 *sativus* L.). Eco Env & Cons 17:799–802

411 Golabadi M, Golkar P, Eghteda AR (2012) Assessment of genetic variation in cucumber (*Cucumis sativus*
412 L.) genotypes. Euro J Exp Bio 2:1382–1388

413 Innark P, Khanobdee C, Samipak S, Jantasuriyarat C (2013) Evaluation of genetic diversity in cucumber
414 (*Cucumis sativus* L.) germplasm using agro-economic traits and microsatellite markers. Sci Hortic
415 162: 278–284

416 Jeffrey C (2005) A new system of Cucurbitaceae. Botanicheskii Zhurnal. Moscow & Leningrad St.
417 Petersburg 90:332–335.

418 Khan Z, Shah AH, Gul R, Majid A, Khan U, Ahmad H (2015) Morpho-agronomic characterization of
419 cucumber germplasm for yield and yield associated traits. Int J Agron Agric Res 6:1–6

420 Little T, Hills J (1978). Agricultural experimentation: Design and analysis. Wiley, New York

421 Lower RL, Edwards MD (1986) Cucumber breeding. In: Basset MJ (ed.) Breeding vegetable Crops. AVI
422 Publishing Co., Westport Connecticut, pp 173–207

423 MAPAMA (2017) Statistical Yearbook 2015. Retrieved February 27, 2017. http:// www.mapama.gob.es

424 Pollack S (2001) Consumer Demand for Fruit and Vegetables: The US Example. In: Regmi A (ed)
425 Changing Structure of Global Food Consumption and Trade. U.S. Department of Agriculture,
426 Economic Research Service. Agricultural and Trade Report, WRS-01-1

427 Rubatzky VE, Yamaguchi M (1997) World Vegetables: Principles, Production, and Nutritive Values.
428 Chapman and Hall, New York

429 Schaefer H, Heibl C, Renner SS (2009) Gourds afloat: a dated phylogeny reveals an Asian origin of the
430 gourd family (Cucurbitaceae) and numerous oversea dispersal events. P Roy Soc B-Biol Sci 276:
431 843–851

432 Sebastian P, Schaefer H, Telford IR, Renner SS (2010) Cucumber (*Cucumis sativus*) and melon (*C. melo*)
433 have numerous wild relatives in Asia and Australia, and the sister species of melon is from
434 Australia. P Natl Acad Sci USA 107:14269–14273

435 Staub JE, Robbins MD, Wehner TC (2008) Cucumber, in: Prohens, J., Nuez, F. (Eds.), Vegetables I:
436 Asteraceae, Brassicaceae, Chenopodiaceae, and Cucurbitaceae. Springer, New York:241–282

437 UPOV (2007) Guidelines for the conduct of tests for distinctness, uniformity and stability: Cucumber,
438 gherkin. Geneva, Switzerland

439 Yadav YC, Kumar S, Singh R (2012) Studies on genetic variability, heritability and genetic advance in
440 cucumber. HortFlora Research Spectrum 1:34–37

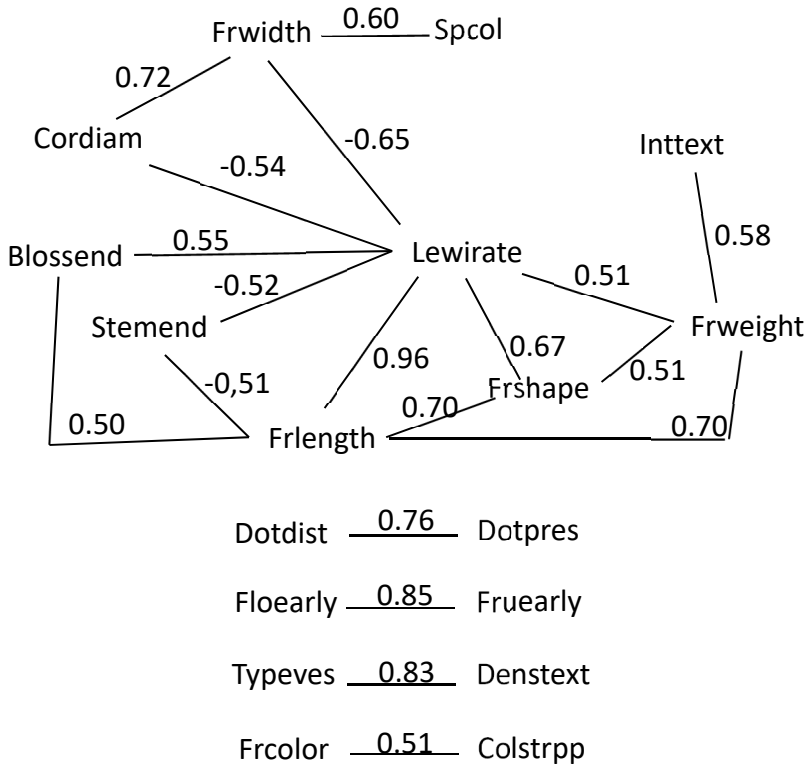


Fig. 1 Correlation among traits higher than 0.5 analyzing 197 cucumber accessions (all values significant at P<0.01; Blossend: Fruit predominant shape at blossom end, Colstrpp: Stripes color, Cordiam: Core diameter, Denstext: Density of fruit texture, Dotdist: Dots distribution, Dotpres: Dots presence and density; Floearly: Female flowering earliness; Frcolor: Predominant fruit skin color at market stage, Fruearly: Commercial harvest earliness; Frlength: Fruit length, Frshape: Fruit shape, Frweight: Fruit weight, Frwidth: Fruit width; Inttext: Intensity of fruit texture, Lewirate: Length/width rate, Espcol: Spines colour, Stemend: Fruit predominant shape at stem end, Typeves: Type of vestiture)

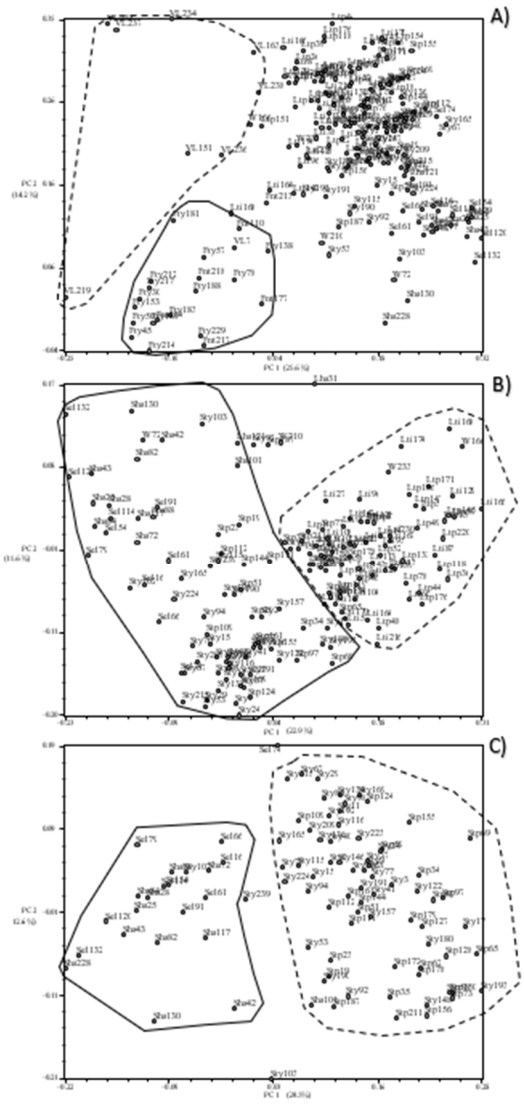


Fig. 2 Principal Component Analysis conducted with A) all accessions (solid lines includes accessions of 'French' type, dashed line includes accessions of 'Very long' type, B) all accessions except for 'French' and 'Very long' types (solid line includes accessions of 'Short' type, dashed line includes accessions of 'Long' type, C) accessions of 'Short' type (solid line includes accessions of Sty and Stp sub-types and dashed line includes accessions of Shs and Sel types)

459

460

461 **Fig. 3** Representatives of each of the groups and sub-groups established. W: White, Shs: Short hairs and
462 spines, Sel: Short ellipsoid, Stp: Short tubercular pronounced, Sty: Short typical, Fty: French typical, Fnt:
463 French non-typical, Ltp: Long tubercular pronounced, Lhs: Long hairs and spines, Lti: Long tubercular
464 intermediate, Lsf: Long soft skin, VL: Very Long. Accessions Lsf230, VL227, VL237 and VL238 are
465 outgroups. See text for more details

466

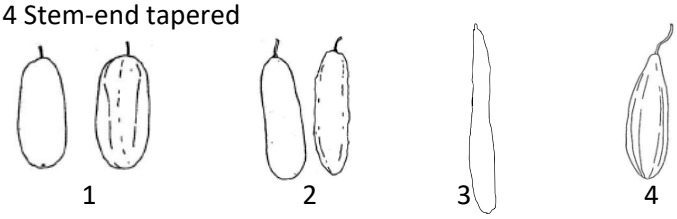
467 Table 1. Descriptors and their scales used in the phenotyping trial of cucumber accessions.
468

Descriptor (acronym)	Units/Scores
Plant, Quantitative	
Female flowering earliness (Floearly)	Measured as the number of days between the transplant until anthesis of the first female flower
Commercial harvest earliness (Fruearly)	Measured as the number of days between the transplant and the harvest of the first fruit at marketable stage
Time from anthesis to commercial harvest (Flowfru)	Measured as the number of days between the anthesis of the flower and the harvest of the fruit at marketable stage. Monitored at least in five fruits per plant
Plant height (Pltheigt)	Measured in centimeters until node 20th on the main stem in each plant
Plant, Qualitative	
Plant growth habit (Plgrowth)	3. Bushy, 5. Intermediate, 7. Prostrate
Sex expression (Sexexp)	1. Monoecious, 2. Andromonoecious, 3 Gynomonoecious
Appearance of female flowers in the main stem (Ffstem)	Monitored until the 20th node. 0 Absence of female flowers on the main stem, 1 Presence of female flowers on the main stem
Presence of female flowers per node (Ffnode)	Number of pistillate flowers per node. 1 One, 2 More than one at least in one plant
Fruit, Quantitative (measured at harvest stage)	
Fruit weight (Frweight)	Measured in grams in at least five fruits per plant
Fruit length (Frlength)	Measured in centimeters in at least five fruits per plant
Fruit width (Frwidth)	Measured in centimeters in at least five fruits per plant
Core diameter (Cordiam)	Measured in centimeters in at least five fruits per plant
Length/width rate (Lewirate)	Rate between the length and width of the

469

470

471 Table 1. (Continuation)

Descriptor (acronym)	Units/Scores
Fruit, Qualitative (measured at harvest stage)	
Fruit shape (Frshape)	1 Oblong ellipsoid, 2 Elliptical elongate, 3 Very long and thin, 4 Stem-end tapered 
Fruit predominant shape at stem end (Stemend)	1 Necked, 2 Acute, 3 Obtuse 
Fruit predominant shape at blossom end (Blossend)	3 Truncate, 5 Rounded, 7 Acute 
Predominant fruit skin color at market stage (Frcolor)	1 White, 5 Light green, 7 Medium green, 9 Dark green.
Type of vestiture (Typeves)	0 Spines only, 1 Hairs and spines
Fruit skin texture (Frsktext)	0 Smooth, 1 Tubercular
Intensity of fruit texture (Inttext)	3 Superficial (weak), 5 Intermediate, 7 Pronounced (strong)
Density of fruit texture (Denstext)	3 Sparse, 5 Medium, 7 Dense, 9 Very dense
Spines color (Spcol)	0 White, 1 Brown or black
Stripes: Presence and length (Stripes)	1 Absent, 3 Less than 1/3 of fruit length, 5 Approx. 1/2 of fruit length, 7 More than 2/3 of fruit length
Stripes color (Colstrpp)	1 Absent, 3 White, 5 Green, 7 Dark green
Dots: presence and density (Dotpres)	1 Absent, 3 Sparse, 5 Medium
Dots: distribution (Dotdist)	1 Absent, 3 Only in bands, 5 Mainly in bands, 7 Evenly distributed

472

473

Table 2. Correlation coefficients between the first three principal components (PC) and the morphological descriptors.

	All accessions			**All accessions except for 'French' and 'Very long' types**			**'Short' type**		
	PC 1	PC 2	PC 3	PC 1	PC 2	PC 3	PC 1	PC 2	PC 3
Spcol	0.61	0.47	0.06	0.01	0.17	0.22	0.01	0.24	0.60
Typeves	0.40	-0.41	**0.61**	-0.62	0.51	-0.30	**-0.81**	-0.27	0.09
Stemend	0.45	-0.19	-0.30	-0.27	-0.06	0.09	-0.15	-0.18	0.02
Blossend	-0.65	-0.10	0.22	0.31	0.08	-0.40	-0.23	-0.13	-0.22
Frshape	-0.64	0.35	0.30	**0.78**	0.25	0.01	0.55	-0.46	0.15
Frcolor	-0.40	-0.22	-0.04	0.05	-0.02	0.62	-0.03	-0.41	-0.45
Dotpres	0.65	0.30	0.09	-0.29	-0.11	0.49	-0.14	0.07	-0.35
Strippes	0.39	0.19	-0.28	-0.09	-0.30	0.60	0.27	0.04	-0.33
Frsktext	-0.06	0.20	-0.11	0.20	- [b]	-	0.30	0.17	0.45
Inttext	-0.07	**0.76**	-0.08	0.64	-0.11	0.30	**0.71**	0.12	0.04
Denstext	0.34	-0.17	**0.68**	-0.49	0.47	-0.21	**-0.70**	-0.27	0.19
Dotdist	0.48	0.35	0.33	-0.36	0.02	0.36	-0.52	0.07	-0.13
Colstrpp	-0.46	-0.43	-0.26	0.08	-0.12	0.25	0.21	-0.28	-0.09
Frweight	-0.32	**0.81**	0.13	**0.78**	0.08	0.26	0.68	0.22	0.36
Frlength	**-0.78** [a]	0.48	0.34	**0.90**	0.16	0.01	**0.79**	-0.26	0.21
Frwidth	**0.74**	0.38	-0.07	-0.21	0.06	0.53	-0.02	**0.62**	0.52
Cordiam	**0.73**	0.20	0.03	-0.58	-0.12	0.25	-0.30	**0.69**	0.29
Lewirate	**-0.85**	0.28	0.31	**0.89**	0.14	-0.02	**0.74**	-0.52	-0.04
Pltheigt	-0.40	0.25	-0.40	0.36	-0.47	0.08	0.35	0.39	-0.31
Floearly	0.21	-0.02	**0.53**	-0.13	**0.79**	0.34	-0.16	-0.58	0.67
Fruearly	0.08	-0.22	**0.59**	-0.17	**0.84**	0.26	-0.26	-0.57	0.66
Flowfru	-0.20	-0.43	0.07	-0.13	0.05	-0.14	-0.12	-0.23	0.09

a Numbers in bold indicate correlation values cited in the text

b The only accession showing variability for this trait was excluded in this analysis

Espcol: Spines color, Typeves: Type of vestiture, Stemend: Fruit predominant shape at stem end, Blossend: Fruit predominant shape at blossom end, Frshape: Fruit shape, Frcolor: Predominant fruit skin color at market stage, Dotpres: Dots presence and density, Strippes: Stripes Presence and length, Frsktext: Fruit skin texture, Inttext: Intensity of fruit texture, Denstext: Fruit skin texture, Dotdist: Dots distribution, Colstrpp: Stripes color, Frweight: Fruit weight, Frlength: Fruit length, Frwidth: Fruit width, Cordiam: Core diameter, Lewirate: Length/width rate, Ffnode: Number of female flowers per node, Pltheigt: Plant height, Floearly: Female flowering earliness, Fruearly: Commercial harvest earliness, Flowfru: Time from anthesis to commercial harvest.

Table 3. Descriptive statistics for quantitative traits of plant and fruit in each of the established groups and in the whole set of accessions (number of accessions).

	Frweight (g)	Frlength (cm)	Frwidth (mm)	Cordiam (mm)	Lewirate	Pltheigt (m)	Floearly (days)	Fruearly (days)	Flowfru (days)
Whole group (206)									
Mean	205.9	17.6	48.5	30.1	0.4	1.8	43.9	59.1	19.7
SD	32.9	3.3	3.1	3.1	0.1	0.2	4.3	6.1	3.1
CV	16.0	18.7	6.3	10.2	23.8	12.6	9.7	10.4	15.5
Range	122.5-312.2	12.5-35.6	40.2-55.6	21.5-35.5	0.3-0.9	1.1-2.4	37.0-58.0	47.0-81.0	11.8-28.7
Spanish (195)									
Mean	204.6	17.2	48.7	30.1	0.4	1.7	44	59.2	19.7
SD	30.8	2.5	2.9	3.1	0.1	0.2	4.2	6.1	3.1
CV	15.0	14.3	6.0	10.3	18.1	12.5	9.6	10.3	15.6
Range	122.5-298.6	12.5-25.6	40.3-55.6	21.5-35.5	0.3-0.6	1.1-2.4	37.0-58.0	47.0-81.0	11.8-28.7
French (20)									
Mean	189.8	18.7	43.7	25.3	0.4	1.9	42	58.9	21.9
SD	23.6	1.8	2.4	2.4	0.1	0.2	2.5	3.4	2.8
CV	12.5	9.8	5.6	9.4	12.0	10.8	6.0	5.8	12.8
Range	149.4-238.5	14.4-21.5	40.3-48.9	22.6-30.0	0.3-0.5	1.5-2.2	38.6-47.0	51.0-65.8	15.1-27.0
Long (68)									
Mean	231.4	19.0	49.3	29.7	0.4	1.8	44.2	59.3	19.2
SD	21.9	1.5	2.1	2.5	0.1	0.2	3.4	4.5	3.0
CV	9.5	7.7	4.3	8.4	10.1	10.4	7.6	7.6	15.6
Range	169.3-295.2	17.0-22.5	44.2-55.6	24.7-35.4	0.3-0.5	1.4-2.3	38.4-53.3	51.0-70.7	13.0-27.0

Frweight: Fruit weight, Frlength: Fruit length, Frwidth: Fruit width, Cordiam: Core diameter, Lewirate: Length/width rate, Pltheigt: Plant height, Floearly: Female flowering earliness, Fruearly: Commercial harvest earliness, Flowfru: Time from anthesis to commercial harvest.

Table 3. (Continuation)

	Frweight (g)	Frlength (cm)	Frwidth (mm)	Cordiam (mm)	Lewirate	Pltheigt (m)	Floearly (days)	Fruearly (days)	Flowfru (days)
Short (100)									
Mean	189.9	15.4	49.6	31.6	0.3	1.7	44.3	59.2	19.5
SD	22.8	1.2	2.0	2.2	0.0	0.2	4.9	7.3	3.0
CV	12.0	7.8	4.0	7.0	8.6	13.1	11.0	12.3	15.3
Range	122.5-245.0	12.5-17.8	42.9-55.6	25.2-35.5	0.3-0.4	1.1-2.1	37.2-58.0	48.2-81.0	11.8-28.7
Very Long (4)									
Mean	216.6	24.2	43.3	26.3	0.6	1.9	44.1	62.7	23.1
SD	56.1	1.1	4.1	4.5	0.1	0.4	4.0	4.8	3.1
CV	25.9	4.4	9.6	17.2	11.1	22.1	9.1	7.7	13.6
Range	178.8-298.6	23.0-25.6	40.3-49.3	21.5-30.3	0.5-0.6	1.4-2.4	40.0-49.5	58.0-68.0	19.9-27.3
White (3)									
Mean	168.9	15.9	45.2	30.0	0.4	1.7	40.8	54.2	19.2
SD	20.6	2.3	2.2	4.2	0.1	0.1	6.6	10.0	2.6
CV	12.2	14.7	4.8	14.0	19.3	5.2	16.1	18.4	13.6
Range	156.3-192.6	14.3-18.6	42.9-47.2	27.0-34.8	0.3-0.4	1.6-1.7	37.0-48.4	47.0-65.6	16.2-20.9

Frweight: Fruit weight, Frlength: Fruit length, Frwidth: Fruit width, Cordiam: Core diameter, Lewirate: Length/width rate, Pltheigt: Plant height, Floearly: Female flowering earliness, Fruearly: Commercial harvest earliness, Flowfru: Time from anthesis to commercial harvest.

Table S1. Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers).

Code	BGV	Accession name	Active Collection[a]	CTY[b]	Origin (Autonomous Community, province, locality)	Local name	Latitude	Longitude
Fnt110	BGV010290	NC019552	ESP027/ESP026	ESP	Andalusia, Granada, Moclín		37º 20' 38'' N	3º 47' 8'' W
Fnt177	BGV015607	V-C-203	ESP026	ESP	Valencian Community, Castellón, Sierra Engarcerán, Els Rosildos	corto del terreno	40º 18' 14'' N	0º 1' 53'' W
Fnt184	BGV014964	BGHZ2372	ESP027/ESP026	ESP	Basque Country, Álava, Iruraiz-Gauna, Ezkerekotxa	verde largo	42º 51' 4'' N	2º 26' 10'' W
Fnt213		NC044368	ESP027	ESP	Cantabria, Cantabria, Colindres	verde	43º 23' 43'' N	3º 26' 57'' W
Fnt217		NC055359(b)	ESP027	ESP	Basque Country, Vizcaya, Zalla, Otxaran	verde de Guernika	43º 12' 44'' N	3º 8' 3'' W
Fnt218		NC055479	ESP027	ESP	Basque Country, Vizcaya, Zaratamo	verde corto	43º 12' 41'' N	2º 52' 25'' W
Fty30	BGV003372	CL-C-12	ESP026	ESP	Castile and Leon, Valladolid, Simancas		41º 35' 30'' N	4º 49' 35'' W
Fty45	BGV013600	NC020546	ESP027/ESP026	ESP	Castilla La Mancha, Guadalajara, Valverde de los Arroyos		41º 7' 49'' N	3º 13' 56'' W
Fty50	BGV011159	1979	ESP027/ESP026	ESP	Aragon, Zaragoza, Sádaba	del terreno	42º 16' 56'' N	1º 16' 16'' W
Fty57	BGV011735	BGHZ2210	ESP027/ESP026	ESP	Aragon, Zaragoza, Ateca	del terreno	41º 19' 55'' N	1º 47' 32'' W
Fty78	BGV004851	V-C-16	ESP026	ESP	Valencian Community, Castellón, La Pobla de Benifassà		40º 39' 29'' N	0º 9' 26'' E
Fty138	BGV004903	V-C-69	ESP026	ESP	Valencian Community, Valencia, Chulilla		39º 39' 25'' N	0º 53' 31'' W
Fty153	BGV000519	AN-C-166	ESP026	ESP	Andalusia, Granada, Válor, Mecina Alfahar		39º 59' 52'' N	3º 4' 11'' W
Fty181	BGV014961	BGHZ2362	ESP027/ESP026	ESP	Valencian Community, Castellón, Todolella	de casa	40º 38' 55'' N	0º 14' 16'' W
Fty183	BGV014963	BGHZ2371	ESP027/ESP026	ESP	Basque Country, Álava, Salvatierra	verde	42º 51' 4'' N	2º 23' 15'' W
Fty188	BGV014969	BGHZ2389	ESP027/ESP026	ESP	Cantabria, Cantabria, Ramales de la Victoria		43º 15' 33'' N	3º 27' 46'' W
Fty189	BGV015121	BGHZ2390	ESP027/ESP026	ESP	Cantabria, Cantabria, Valderredible, Sobrepeña		42º 47' 17'' N	3º 57' 42'' W
Fty212		NC044054	ESP027	ESP	Navarra, Navarra, Aibar		42º 38' 30'' N	1º 20' 17'' W
Fty214		NC044369	ESP027	ESP	Cantabria, Cantabria, Villaescusa, La Concha	del país	43º 22' 19'' N	3º 51' 9'' W
Fty217		NC055359(a)	ESP027	ESP	Basque Country, Vizcaya, Otxaran, Zalla	verde de Guernika	43º 12' 44'' N	3º 8' 3'' W
Fty229		CGN19655	NLD037	USA		SC 53-B (6)		

(1) CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

(2) CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan, LKA: Sri Lanka, NLD: The Netherlands, USA: United States of America.

Table S1 (cont.). Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers).

Code	BGV	Accession name	Active Collection[a]	CTY[b]	Origin (Autonomous Community, province, locality)	Local name	Latitude	Longitude
Lhs31	BGV004926	V-C-94	ESP026	ESP	Valencian Community, Castellón, Argelita	verde del terreno	40º 3' 18'' N	0º 20' 57'' W
Lhs121	BGV000381	AN-C-27	ESP026	ESP	Andalusia, Málaga, Ronda		36º 44' 19'' N	5º 9' 51'' W
Lsf230		CGN21616	NLD037	IRN		Rasht		
Lti27	BGV011737	2322	ESP027/ESP026	ESP	Castile and Leon, Ávila, Cebreros	largo	40º 27' 21'' N	4º 27' 47'' W
Lti38	BGV011537	2522	ESP027/ESP026	ESP	Extremadura, Badajoz, Herrera del Duque	enano	39º 10' 0'' N	5º 2' 54'' W
Lti47	BGV000040	A-C-8	ESP026	ESP	Aragon, Teruel, Gea de Albarracín		40º 24' 43'' N	1º 20' 49'' W
Lti57	BGV011735	BGHZ2210	ESP027/ESP026	ESP	Aragon, Zaragoza, Ateca	del terreno	41º 19' 55'' N	1º 47' 32'' W
Lti60	BGV011724	2444	ESP027/ESP026	ESP	Aragon, Teruel, Castellote, Ladruñán	de Los Llanos	40º 44' 29'' N	0º 24' 3'' W
Lti63	BGV002505	CA-C-33	ESP026	ESP	Canary Islands, Santa Cruz de Tenerife (La Palma), Garafía, El Tablado		28º 49' 58'' N	17º 52' 38'' W
Lti68	BGV004936	V-C-104	ESP026	ESP	Valencian Community, Valencia, Torrebaja		40º 5' 49'' N	1º 15' 17'' W
Lti71	BGV004939	V-C-107	ESP026	ESP	Valencian Community, Valencia, Bocairent		38º 46' 3'' N	0º 36' 36'' W
Lti87	BGV000469	AN-C-115	ESP026	ESP	Andalusia, Granada, Güejar Sierra		37º 9' 39'' N	3º 26' 15'' W
Lti90	BGV000453	AN-C-99	ESP026	ESP	Andalusia, Jaén, Cazorla		37º 54' 48'' N	3º 0' 0'' W
Lti93	BGV010299	357	ESP027/ESP026	ESP	Castilla La Mancha, Guadalajara, Corduente, Torete		40º 48' 57'' N	2º 3' 21'' W
Lti95	BGV000419	AN-C-65	ESP026	ESP	Andalusia, Cádiz, Grazalema	del país	36º 46' 9'' N	5º 21' 52'' W
Lti96	BGV000518	AN-C-165	ESP026	ESP	Andalusia, Almería, Láujar de Andarax		36º 59' 20'' N	2º 54' 37'' W
Lti99	BGV004883	V-C-49	ESP026	ESP	Valencian Community, Alicante, Alcoleja		38º 40' 36'' N	0º 19' 47'' W
Lti100	BGV001774	C-C-4	ESP026	ESP	Catalonia, Barcelona, Torelló		42º 2' 53'' N	2º 15' 52'' E
Lti104	BGV000415	AN-C-61	ESP026	ESP	Andalusia, Cádiz, Benaocaz	del país	36º 42' 5'' N	5º 25' 12'' W
Lti108	BGV002489	CA-C-17	ESP026	ESP	Canary Islands, Santa Cruz de Tenerife (La Palma), Barlovento	del país	28º 49' 38'' N	17º 48' 14'' W
Lti113	BGV002494	CA-C-22	ESP026	ESP	Canary Islands, Santa Cruz de Tenerife (La Palma), Puntallana, Santa Lucía		28º 43' 38'' N	17º 44' 48'' W
Lti123	BGV000377	AN-C-23	ESP026	ESP	Andalusia, Málaga, Jimera de Líbar		36º 39' 8'' N	5º 16' 24'' W
Lti126	BGV000372	AN-C-18	ESP026	ESP	Andalusia, Granada, Ugíjar		36º 57' 51'' N	3º 3' 10'' W

(1) CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

(2) CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan, LKA: Sri Lanka, NLD: The Netherlands, USA: United States of America.

Table S1 (cont.). Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers)..

Code	BGV	Accession name	Active Collection[a]	CTY[b]	Origin (Autonomous Community, province, locality)	Local name	Latitude	Longitude
Lti129	BGV000459	AN-C-105	ESP026	ESP	Andalusia, Jaén, Alcalá la Real		37º 27' 40'' N	3º 55' 41'' W
Lti131	BGV000445	AN-C-91	ESP026	ESP	Andalusia, Jaén, Canena	verde	38º 3' 0'' N	3º 28' 49'' W
Lti142	BGV001310	AS-C-1	ESP026	ESP	Asturias, Asturias, Siero, La Carrera		43º 23' 1'' N	5º 41' 53'' W
Lti164	BGV015697	NC094842	ESP026	ESP	Valencian Community, Valencia, Anna		39º 1' 23'' N	0º 38' 35'' W
Lti166	BGV015698	NC094962(a)	ESP026	ESP	Valencian Community, Castellón, Sierra Engarcerán, Els Rosildos	del terreno	40º 17' 34'' N	0º 2' 48'' W
Lti167	BGV015702	C-C-51	ESP026	ESP	Catalonia, Girona, Olot	del país	42º 10' 55'' N	2º 29' 17'' E
Lti168	BGV015229	BGHZ4778	ESP027/ESP026	ESP	Basque Country, Vizcaya, Sopelana	de casa	43º 22' 44'' N	2º 58' 58'' W
Lti170	BGV015700	C-C-49	ESP026	ESP	Cataluña, Girona, Llagostera		41º 49' 45'' N	2º 53' 38'' E
Lti173	BGV012101	Z-02-036	ESP027/ESP026	ESP	Aragon, Zaragoza, Gallur	de piel anaranjada	41º 52' 16'' N	1º 19' 1'' W
Lti174	BGV015699	C-C-48	ESP026	ESP	Catalonia, Girona, Mieres	del país	42º 7' 37'' N	2º 38' 21'' E
Lti216		NC055008	ESP027	ESP	Galicia, Orense, Barbadas, Lamas	de Valenzana	42º 18' 57'' N	7º 53' 20'' W
Lti221		NC074362	ESP027	ESP	Castile and Leon, Burgos, Castrillo de Riopisuerga		42º 30' 53'' N	4º 15' 8'' W
Lti231		CGN21691	NLD037	COG		N2/81		
Ltp4	BGV011738	DT83	ESP027/ESP026	ESP	Aragon, Huesca, Alquézar	del terreno	42º 10' 24'' N	0º 1' 44'' E
Ltp5	BGV004302	MU-C-48	ESP026	ESP	Region of Murcia, Murcia, Torre-Pacheco	del terreno	37º 44' 39'' N	0º 57' 8'' W
Ltp8	BGV004308	MU-C-54	ESP026	ESP	Region of Murcia, Murcia, Murcia, Churra	rugoso corto	38º 1' 35'' N	1º 8' 5'' W
Ltp12	BGV004309	MU-C-55	ESP026	ESP	Region of Murcia, Murcia, Molina de Segura	medio largo del país	38º 3' 6'' N	1º 12' 38'' W
Ltp13	BGV004303	MU-C-49	ESP026	ESP	Region of Murcia, Murcia, Murcia, Santa Cruz	murciano	38º 1' 41'' N	1º 2' 50'' W
Ltp14	BGV004923	V-C-91	ESP026	ESP	Valencian Community, Castellón, Ludiente, La Giraba		40º 6' 6'' N	0º 22' 34'' W
Ltp18	BGV004307	MU-C-53	ESP026	ESP	Region of Murcia, Murcia, Murcia, Guadalupe		38º 0' 4'' N	1º 10' 15'' W
Ltp22	BGV011741	F37	ESP027/ESP026	ESP	Aragon, Teruel, Andorra		40º 58' 34'' N	0º 26' 33'' W
Ltp36	BGV004893	V-C-59	ESP026	ESP	Valencian Community, Valencia, Ademuz		40º 3' 45'' N	1º 17' 8'' W

(1) CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

(2) CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan; LKA: Sri Lanka, NLD: The Netherlands, USA: United States of America.

Table S1 (cont.). Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers).

Code	BGV	Accession name	Active Collection[a]	CTY[b]	Origin (Autonomous Community, province, locality)	Local name	Latitude	Longitude
Ltp39	BGV004838	V-C-1	ESP026	ESP	Valencian Community, Alicante, Cocentaina	tendral	38º 44' 48'' N	0º 26' 20'' W
Ltp44	BGV004918	V-C-85	ESP026	ESP	Valencian Community, Valencia, Xàtiva		38º 59' 21'' N	0º 31' 3'' W
Ltp46	BGV011586	2457	ESP027/ESP026	ESP	Galicia, Orense, Barbadás, A Valenzá	Valenzana	42º 18' 57'' N	7º 53' 20'' W
Ltp48	BGV011540	2523	ESP027/ESP026	ESP	Extremadura, Badajoz, Herrera del Duque	largo del país	39º 10' 0'' N	5º 2' 54'' W
Ltp49	BGV004960	V-C-128	ESP026	ESP	Valencian Community, Alicante, Almoradí	medio largo verde	38º 6' 40'' N	0º 47' 27'' W
Ltp52	BGV004962	V-C-130	ESP026	ESP	Valencian Community, Alicante, Guardamar del Segura	pequeño verde	38º 5' 28'' N	0º 39' 13'' W
Ltp76	BGV000438	AN-C-84	ESP026	ESP	Andalusia, Jaén, Canena		38º 3' 0'' N	3º 28' 49'' W
Ltp78	BGV004851	V-C-16	ESP026	ESP	Valencian Community, Castellón, La Pobla de Benifassà		40º 39' 29'' N	0º 9' 26'' E
Ltp81	BGV000449	AN-C-95	ESP026	ESP	Andalusia, Jaén, Cazorla		37º 54' 48'' N	3º 0' 0'' W
Ltp118	BGV004840	V-C-3	ESP026	ESP	Valencian Community, Valencia, Valencia, La Punta	del terreno	39º 26' 45'' N	0º 20' 22'' W
Ltp119	BGV004899	V-C-65	ESP026	ESP	Valencian Community, Valencia, Casas Altas		40º 2' 27'' N	1º 15' 40'' W
Ltp133	BGV000421	AN-C-67	ESP026	ESP	Andalusia, Cádiz, Tarifa	andaluz	36º 1' 13'' N	5º 35' 53'' W
Ltp137	BGV000503	AN-C-149	ESP026	ESP	Andalusia, Huelva, Aracena		37º 53' 27'' N	6º 33' 37'' W
Ltp139	BGV004964	V-C-132	ESP026	ESP	Valencian Community, Alicante, Novelda	rugoso del país	38º 23' 14'' N	0º 45' 51'' W
Ltp143	BGV004976	V-C-144	ESP026	ESP	Valencian Community, Valencia, Casas Bajas		40º 1' 29'' N	1º 15' 34'' W
Ltp145	BGV000380	AN-C-26	ESP026	ESP	Andalusia, Málaga, Benaoján	del terreno	36º 43' 12'' N	5º 15' 9'' W
Ltp147	BGV014220	07-A30-01	ESP026	ESP	Valencian Community, Castellón, Tírig		40º 25' 31'' N	0º 4' 44'' E
Ltp151	BGV000489	AN-C-135	ESP026	ESP	Andalusia, Huelva, Calañas	del país	37º 39' 21'' N	6º 52' 39'' W
Ltp154	BGV015696	NC094818	ESP027/ESP026	ESP	Valencian Community, Alicante, Gaianes	del terreno	38º 48' 52'' N	0º 24' 23'' W
Ltp158	BGV000416	AN-C-62	ESP026	ESP	Andalusia, Cádiz, Benaocaz	amarillo	36º 42' 5'' N	5º 25' 12'' W
Ltp160	BGV000371	AN-C-17	ESP026	ESP	Andalusia, Granada, Pórtugos		36º 56' 34'' N	3º 18' 35'' W
Ltp171	BGV012114	Z-00-035	ESP027/ESP026	ESP	Aragon, Teruel, Cantavieja	grande verde	40º 31' 45'' N	0º 24' 16'' W
Ltp176	BGV015695	NC094812	ESP026	ESP	Valencian Community, Alicante, Cocentaina		38º 44' 48'' N	0º 26' 20'' W

(1) CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

(2) CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan, LKA: Sri Lanka: NLD: The Netherlands, USA: United States of America.

Table S1 (cont.). Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers).

Code	BGV	Accession name	Active Collection [a]	CTY[b]	Origin (Autonomous Community, province, locality)	Local name	Latitude	Longitude
Ltp186	BGV014967	BGHZ2387	ESP027/ESP026	ESP	Castilla La Mancha, Guadalajara, Miedes de Atienza	alargado	41º 16' 8'' N	0º 57' 47'' W
Ltp220		NC056334	ESP027	ESP	Region of Murcia, Murcia, Lorca, Zarzadilla de Totana		37º 52' 40'' N	1º 42' 25'' W
Sel16	BGV003693	CM-C-24	ESP026	ESP	Castilla La Mancha, Albacete, Montalvos		39º 10' 6'' N	2º 1' 35'' W
Sel33	BGV010550	1351	ESP027/ESP026	ESP	Castilla La Mancha, Cuenca, Pozorrubio		39º 49' 5'' N	2º 56' 57'' W
Sel54	BGV010381	680	ESP027/ESP026	ESP	Castilla La Mancha, Ciudad Real, Manzanares		38º 59' 58'' N	3º 22' 8'' W
Sel61	BGV004845	V-C-10	ESP026	ESP	Valencian Community, Valencia, Venta del Moro	conqueño	39º 29' 6'' N	1º 21' 19'' W
Sel66	BGV011740	MD90	ESP027/ESP026	ESP	Extremadura, Badajoz, Puebla de Alcocer	enano	38º 59' 11'' N	5º 15' 21'' W
Sel74	BGV003681	CM-C-12	ESP026	ESP	Castilla La Mancha, Albacete, Villatoya	del terreno	39º 20' 6'' N	1º 20' 14'' W
Sel79	BGV003706	CM-C-38	ESP026	ESP	Castilla La Mancha, Cuenca, Sotorribas, Villaseca		40º 19' 6'' N	2º 12' 50'' W
Sel91	BGV011739	JT88	ESP027/ESP026	ESP	Castile and León, Segovia, Martín Muñoz de las Posadas	antiguo	40º 59' 50'' N	4º 35' 40'' W
Sel106	BGV005022	V-C-190	ESP026	ESP	Valencian Community, Alicante, Villena		38º 38' 10'' N	0º 51' 58'' W
Sel114	BGV000502	AN-C-148	ESP026	ESP	Andalusia, Huelva, Cortegana		37º 54' 43'' N	6º 49' 6'' W
Sel120	BGV003703	CM-C-35	ESP026	ESP	Castilla La Mancha, Cuenca, Sotorribas, Sotos		40º 11' 47'' N	2º 9' 45'' W
Sel132	BGV005021	V-C-189	ESP026	ESP	Valencian Community, Alicante, Villena		38º 38' 10'' N	0º 51' 58'' W
Shs25	BGV010380	672	ESP027/ESP026	ESP	Castilla La Mancha, Ciudad Real, Alhambra		38º 54' 3'' N	3º 3' 11'' W
Shs28	BGV003713	CM-C-45	ESP026	ESP	Castilla La Mancha, Cuenca, Palomera, Molinos de Papel		40º 4' 52'' N	2º 4' 34'' W
Shs42	BGV010386	698	ESP027/ESP026	ESP	Castilla La Mancha, Ciudad Real, Malagón		39º 10' 10'' N	3º 51' 15'' W
Shs43	BGV003714	CM-C-46	ESP026	ESP	Castilla La Mancha, Cuenca, Palomera, Molinos de Papel		40º 4' 52'' N	2º 4' 34'' W
Shs72	BGV010629	2413	ESP027/ESP026	ESP	Castile and Leon, Valladolid, Peñafiel	del terreno	41º 35' 51'' N	4º 7' 2'' W
Shs82	BGV014501	V-C-214	ESP026	ESP	Valencian Community			
Shs84	BGV010291	NC019555	ESP027/ESP026	ESP	Castilla La Mancha, Cuenca, Zarzuela		40º 15' 24'' N	2º 7' 39'' W
Shs88	BGV011734	BGHZ2123	ESP027/ESP026	ESP	Castile and Leon, Valladolid, Peñafiel	el terreno	41º 35' 51'' N	4º 7' 2'' W

(1) CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

(2) CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan, LKA: Sri Lanka, NLD: The Netherlands, USA: United States of America.

Table S1 (cont.). Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers).

Code	BGV	Accession name	Active Collection[a]	CTY[b]	Origin (Autonomous Community, province, locality)	Local name	Latitude	Longitude
Shs101	BGV002495	CA-C-23	ESP026	ESP	Canary Islands, Santa Cruz de Tenerife (La Palma), Garafía, Don Pedro		28º 50' 32'' N	17º 53' 58'' W
Shs117	BGV010367	616	ESP027/ESP026	ESP	Castilla La Mancha, Ciudad Real, Villanueva de los Infantes		38º 44' 18'' N	3º 0' 50'' W
Shs130	BGV003709	CM-C-41	ESP026	ESP	Castilla La Mancha, Cuenca, Sotorribas, Ribagorda		40º 20' 3'' N	2º 13' 50'' W
Shs228		CGN20512	NLD037	NLD				
Stp6	BGV011928	CF82	ESP027/ESP026	ESP	Andalusia, Jaén, Villacarrillo	largo	38º 6' 37'' N	3º 5' 6'' W
Stp19	BGV010636	2437	ESP027/ESP026	ESP	Castile and Leon, Soria, Almazán	de la tierra	41º 29' 17'' N	2º 31' 54'' W
Stp20	BGV002491	CA-C-19	ESP026	ESP	Canary Islands, Santa Cruz de Tenerife (La Palma), Barlovento, Topaciegas		28º 49' 46'' N	17º 49' 25'' W
Stp23	BGV010608	BGHZ2337	ESP027/ESP026	ESP	Castile and Leon, Segovia, Ayllón	de la tierra	41º 25' 19'' N	3º 22' 32'' W
Stp34	BGV004892	V-C-58	ESP026	ESP	Valencian Community, Valencia, Ademuz		40º 3' 45'' N	1º 17' 8'' W
Stp35	BGV004895	V-C-61	ESP026	ESP	Valencian Community, Valencia, Ademuz		40º 3' 45'' N	1º 17' 8'' W
Stp51	BGV010632	2422	ESP027/ESP026	ESP	Castile and Leon, Valladolid, Peñafiel	del país	41º 35' 51'' N	4º 7' 2'' W
Stp56	BGV000042	A-C-10	ESP026	ESP	Aragon, Huesca, Quicena	gordo	42º 8' 57'' N	0º 21' 34'' W
Stp58	BGV000437	AN-C-83	ESP026	ESP	Andalusia, Jaén, Úbeda		38º 0' 46'' N	3º 22' 12'' W
Stp59	BGV010314	426	ESP027/ESP026	ESP	Castilla La Mancha, Guadalajara, Yunquera de Henares	largo	40º 45' 19'' N	3º 9' 52'' W
Stp62	BGV000452	AN-C-98	ESP026	ESP	Andalusia, Jaén, Cazorla		37º 54' 48'' N	3º 0' 0'' W
Stp65	BGV008299	IVIA-044	ESP026	ESP	Valencian Community	del país		
Stp69	BGV000039	A-C-7	ESP026	ESP	Aragon, Teruel, Tramacastilla		40º 25' 55'' N	1º 34' 23'' W
Stp73	BGV004961	V-C-129	ESP026	ESP	Valencian Community, Alicante, Rojales	corto verde	38º 5' 25'' N	0º 43' 19'' W
Stp89	BGV000050	A-C-18	ESP026	ESP	Aragon, Zaragoza, Rueda de Jalón		41º 38' 3'' N	1º 16' 25'' W
Stp97	BGV000467	AN-C-113	ESP026	ESP	Andalusia, Jaén, Martos	jaenero	37º 43' 29'' N	3º 57' 58'' W
Stp109	BGV000513	AN-C-159	ESP026	ESP	Andalusia, Huelva, Cortegana		37º 54' 43'' N	6º 49' 6'' W
Stp111	BGV010322	NC019957	ESP027/ESP026	ESP	Castilla La Mancha, Guadalajara, Yélamos de Arriba		40º 38' 22'' N	2º 50' 35'' W

(1) CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

(2) CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan, LKA: Sri Lanka, NLD: The Netherlands, USA: United States of America.

Table S1 (cont.). Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers).

Code	BGV	Accession name	Active Collection[a]	CTY[b]	Origin	Local name	Latitude	Longitude
Stp112	BGV003696	CM-C-27	ESP026	ESP	Castilla La Mancha, Albacete, Tarazona de la Mancha		39º 16' 0'' N	1º 54' 42'' W
Stp124	BGV002490	CA-C-18	ESP026	ESP	Canary Islands, Santa Cruz de Tenerife (La Palma), Barlovento, Topaciegas		28º 49' 46'' N	17º 49' 25'' W
Stp127	BGV003688	CM-C-19	ESP026	ESP	Castilla La Mancha, Albacete, Alcalá del Júcar		39º 11' 40'' N	1º 25' 41'' W
Stp128	BGV000512	AN-C-158	ESP026	ESP	Andalusia, Huelva, Cortegana		37º 54' 43'' N	6º 49' 6'' W
Stp144	BGV004026	E-C-56	ESP026	ESP	Extremadura, Cáceres, Madrigal de la Vera		40º 8' 53'' N	5º 22' 8'' W
Stp150	BGV003679	CM-C-10	ESP026	ESP	Castilla La Mancha, Guadalajara, Mondéjar	tronquero	40º 19' 25'' N	3º 6' 30'' W
Stp155	BGV002473	CA-C-1	ESP026	ESP	Canary Islands, Santa Cruz de Tenerife, Hermigua, Lomo San Pedro		28º 8' 59'' N	17º 12' 0'' W
Stp156	BGV010324	NC019964	ESP027/ESP026	ESP	Castilla La Mancha, Guadalajara, Yélamos de Arriba		40º 38' 22'' N	2º 50' 35'' W
Stp161	BGV004938	V-C-106	ESP026	ESP	Valencian Community, Valencia, Bocairent		38º 46' 3'' N	0º 36' 36'' W
Stp172	BGV015701	C-C-50	ESP026	ESP	Catalonia, Girona, Sant Jaume de Llierca		42º 12' 49'' N	2º 36' 22'' E
Stp178	BGV014958	BGHZ2359	ESP027/ESP026	ESP	Aragon, Huesca, Valle de Bardají, Llert	amarillo	42º 24' 36'' N	0º 26' 3'' E
Stp179	BGV014959	BGHZ2360	ESP027/ESP026	ESP	Aragon, Huesca, Valle de Bardají, Llert	amarillo	42º 24' 36'' N	0º 26' 3'' E
Stp187	BGV014968	BGHZ2388	ESP027/ESP026	ESP	Castilla La Mancha, Albacete, Riópar, El Lugar Nuevo	largo	38º 30' '' N	2º 23' '' W
Stp211		NC044016	ESP027	ESP	Navarra, Navarra, Mendigorría		42º 37' 41'' N	1º 50' 7'' W
Stp222		NC076258	ESP027	ESP	Castile and Leon, Palencia, Dueñas	del terreno	41º 52' 34'' N	4º 32' 44'' W
Sty2	BGV011736	BGHZ2315	ESP027/ESP026	ESP	Castile and Leon, Ávila, Mombeltrán	de la tierra	40º 15' 41'' N	5º 1' 0'' W
Sty3	BGV011742	F4	ESP027/ESP026	ESP	Castilla La Mancha, Albacete, Ayna		38º 33' 12'' N	2º 4' 7'' W
Sty9	BGV004304	MU-C-50	ESP026	ESP	Region of Murcia, Murcia, Murcia, Monteagudo	pequeño verde	38º 1' 15'' N	1º 6' 6'' W
Sty11	BGV010683	546/1	ESP027/ESP026	ESP	Castile and Leon, Ávila, Cebreros	pequeño amarillo	40º 27' 21'' N	4º 27' 47'' W

(1) CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

(2) CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan, LKA: Sri Lanka, NLD: The Netherlands, USA: United States of America.

Table S1 (cont.). Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers).

Code	BGV	Accession name	Active Collection[a]	CTY[b]	Origin	Local name	Latitude	Longitude
Sty15	BGV011872	546/2	ESP027/ESP026	ESP	Castile and Leon, Ávila, Cebreros	pequeño blanco	40º 27' 21'' N	4º 27' 47'' W
Sty17	BGV011582	2735	ESP027/ESP026	ESP	Aragon, Teruel, Alcañiz	amarillo con raya	41º 3' 9'' N	0º 7' 54'' W
Sty24	BGV011556	2503	ESP027/ESP026	ESP	Extremadura, Cáceres, Logrosán	rubio corto	39º 20' 22'' N	5º 29' 32'' W
Sty29	BGV000523	AN-C-170	ESP026	ESP	Andalusia, Granada, La Peza		37º 16' 35'' N	3º 18' 58'' W
Sty37	BGV000520	AN-C-167	ESP026	ESP	Andalusia, Granada, Juviles		36º 57' 4'' N	3º 13' 30'' W
Sty40	BGV003711	CM-C-43	ESP026	ESP	Castilla La Mancha, Cuenca, Cuenca		40º 4' 40'' N	2º 7' 49'' W
Sty41	BGV011544	2538	ESP027/ESP026	ESP	Extremadura, Badajoz, Talarrubias	verde	39º 2' 21'' N	5º 13' 57'' W
Sty53	BGV011913	HU049	ESP027/ESP026	ESP	Aragon, Huesca, Sabiñánigo, Molino de Villobas	amarillo	42º 23' 34'' N	0º 17' 14'' W
Sty55	BGV003705	CM-C-37	ESP026	ESP	Castilla La Mancha, Cuenca, Sotorribas, Sotos		40º 11' 47'' N	2º 9' 45'' W
Sty67	BGV004981	V-C-149	ESP026	ESP	Valencian Community, Alicante, Beneixama		42º 9' 19'' N	3º 4' 15'' E
Sty70	BGV000460	AN-C-106	ESP026	ESP	Andalusia, Jaén, Alcaudete	del terreno	37º 35' 29'' N	4º 5' 10'' W
Sty77	BGV003370	CL-C-10	ESP026	ESP	Castile and Leon, Palencia, Palencia		42º 0' 32'' N	4º 32' 0'' W
Sty80	BGV000047	A-C-15	ESP026	ESP	Aragon, Zaragoza, Lumpiaque		41º 37' 56'' N	1º 17' 55'' W
Sty86	BGV000451	AN-C-97	ESP026	ESP	Andalusia, Jaén, Cazorla		37º 54' 48'' N	3º 0' 0'' W
Sty92	BGV000375	AN-C-21	ESP026	ESP	Andalusia, Granada, Vegas del Genil, Purchil		37º 10' 25'' N	3º 39' 52'' W
Sty94	**BGV010301**	**360**	**ESP027/ESP026**	**ESP**	**Castilla La Mancha, Guadalajara, Corduente, Torete**		**40º 48' 57'' N**	**2º 3' 21'' W**
Sty98	BGV003371	CL-C-11	ESP026	ESP	Castile and Leon, Palencia, Venta de Baños		41º 55' 16'' N	4º 29' 38'' W
Sty102	**BGV010296**	**340**	**ESP027/ESP026**	**ESP**	**Castilla La Mancha, Guadalajara, Algar de Mesa**		**41º 8' 6'' N**	**1º 57' 28'' W**
Sty103	BGV000035	A-C-3	ESP026	ESP	Aragon, Teruel, Torres de Albarracín		40º 25' 41'' N	1º 31' 51'' W
Sty105	BGV000473	AN-C-119	ESP026	ESP	Andalusia, Granada, Santa Fe	castellano pinchudo	37º 11' 26'' N	3º 43' 4'' W
Sty115	BGV000522	AN-C-169	ESP026	ESP	Andalusia, Granada, La Peza		37º 16' 58'' N	3º 16' 52'' W
Sty116	BGV004925	V-C-93	ESP026	ESP	Valencian Community, Castellón, Fanzara	de vinagre	40º 1' 14'' N	0º 18' 55'' W

(1) CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

(2) CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan, LKA: Sri Lanka, NLD: The Netherlands, USA: United States of America.

Table S1 (cont.). Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers).

Code	BGV	Accession name	Active Collection[a]	CTY[b]	Origin	Local name	Latitude	Longitude
Sty122	BGV000408	AN-C-54	ESP026	ESP	Andalusia, Cádiz, Los Barrios	enano	36º 11' 12'' N	5º 29' 29'' W
Sty134	BGV003366	CL-C-6	ESP026	ESP	Castile and Leon, Segovia, Hontalbilla	de conserva	41º 20' 49'' N	4º 7' 13'' W
Sty136	BGV010350	563	ESP027/ESP026	ESP	Castilla La Mancha, Ciudad Real, Calzada de Calatrava		38º 42' 21'' N	3º 46' 33'' W
Sty146	BGV004920	V-C-88	ESP026	ESP	Valencian Community, Castellón, Figueroles	de vinagre	40º 7' 8'' N	0º 14' 12'' W
Sty148	BGV014552	07-A16-06	ESP026	ESP	Valencian Community , Castellón, Ribesalbes		40º 1' 24'' N	0º 16' 35'' W
Sty157	BGV000479	AN-C-125	ESP026	ESP	Andalusia, Córdoba, Lucena, Jauja	de la tierra	37º 18' 23'' N	4º 39' 15'' W
Sty165	BGV015469	BGHZ4943	ESP027/ESP026	ESP	Extremadura, Cáceres, Alía, La Calera		39º 30' 4'' N	5º 15' 1'' W
Sty169	BGV011886	2899/3	ESP027/ESP026	ESP	Aragon, Huesca, Campo	del terreno	42º 24' 40'' N	0º 23' 54'' E
Sty180	BGV014960	BGHZ2361	ESP027/ESP026	ESP	Aragon, Huesca, Labuerda	amarillo	42º 27' 4'' N	0º 8' 9'' E
Sty190	BGV014970	BGHZ2391	ESP027/ESP026	ESP	Castilla La Mancha, Cuenca, Alcohujate		40º 25' 6'' N	2º 36' 53'' W
Sty191	BGV015122	BGHZ2392	ESP027/ESP026	ESP	Extremadura, Cáceres, Carrascalejo		39º 38' 48'' N	5º 13' 0'' W
Sty192	BGV014971	BGHZ2393	ESP027/ESP026	ESP	Extremadura, Cáceres, Logrosán	extremeño	39º 20' 22'' N	5º 29' 32'' W
Sty193	BGV014972	BGHZ2394	ESP027/ESP026	ESP	Andalusia, Huelva, Santa Olalla del Cala	largo	37º 54' 23'' N	6º 14' 0'' W
Sty209		NC026203	ESP027	ESP	Castilla La Mancha, Toledo, Noblejas	corto	39º 58' 49'' N	3º 26' 22'' W
Sty215		NC051904	ESP027	ESP	La Rioja, La Rioja, San Roman de Cameros	amarillo del Duque de Altamira	42º 13' 59'' N	2º 28' 26'' W
Sty223		NC076278	ESP027	ESP	Madrid, Madrid, Tielmes		40º 14' 40'' N	3º 18' 52'' W
Sty224		NC076454	ESP027	ESP	Madrid, Madrid, Arganda del Rey		40º 18' 3'' N	3º 26' 17'' W
Sty239	BGV015704	CM-C-67	ESP026	ESP	Castilla La Mancha, Cuenca, Villarejo-Periesteban	de Huete	39º 52' 17'' N	2º 26' 15'' W
VL7	BGV004305	MU-C-51	ESP026	ESP	Region of Murcia, Murcia, Murcia, San Benito	largo verde	37º 57' 52'' N	1º 7' 35'' W
VL151	BGV000489	AN-C-135	ESP026	ESP	Andalusia, Huelva, Calañas	del país	37º 39' 21'' N	6º 52' 39'' W

(1) CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

(2) CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan, LKA: Sri Lanka, NLD: The Netherlands, USA: United States of America.

Table S1 (cont.). Accessions characterized in this work including passport information (In bold accessions selected for genotyping with molecular markers).

Code	BGV	Accession name	Active Collection[a]	CTY[b]	Origin	Local name	Latitude	Longitude
VL163	BGV001779	C-C-9	ESP026	ESP	Catalonia, Tarragona, Gratallops		41º 11' 39'' N	0º 46' 41'' E
VL219		NC055480	ESP027	ESP	Basque Country, Vizcaya, Zaratamo	de encurtir	43º 12' 41'' N	2º 52' 25'' W
VL227		CGN20853	NLD037	JPN		Sagami Hanpaku Fushinari Kyuri		
VL234	BGV015107	V05A0754	CHN122	CHN		Hei Wu She		
VL235	BGV015113	V05A0781	CHN122	CHN		Shou Guang Qiu Gua		
VL236	BGV015115	V05A0921	CHN122	CHN		Long Quan Qing Huang Gua		
VL237	BGV015116	V05A0922	CHN122	CHN		De Hui Huang Gua		
VL238	BGV015118	V05A0926	CHN122	CHN		San Ye Zao		
W72	BGV010629	2413	ESP027/ESP026	ESP	Castile and Leon, Valladolid, Peñafiel	del terreno	41º 35' 51'' N	4º 7' 2'' W
W166	BGV015698	NC094962(b)	ESP026	ESP	Valencian Community, Castellón, Sierra Engarcerán, Els Rosildos	del terreno	40º 17' 34'' N	0º 2' 48'' W
W210		NC043980	ESP027	ESP	Navarra, Navarra, Mendigorría	blanco	42º 37' 41'' N	1º 50' 7'' W
W233		CGN20517	NLD037	LKA		Yellow 1		

a CHN122: Chinese Academy of Agricultural Science (CAAS), ESP026: Institute of Conservation and Improvement of Valencian Agrodiversity, (COMAV), ESP027: Vegetables Genebank BGHZ, NLD37: Center for Genetic Resources (CGN).

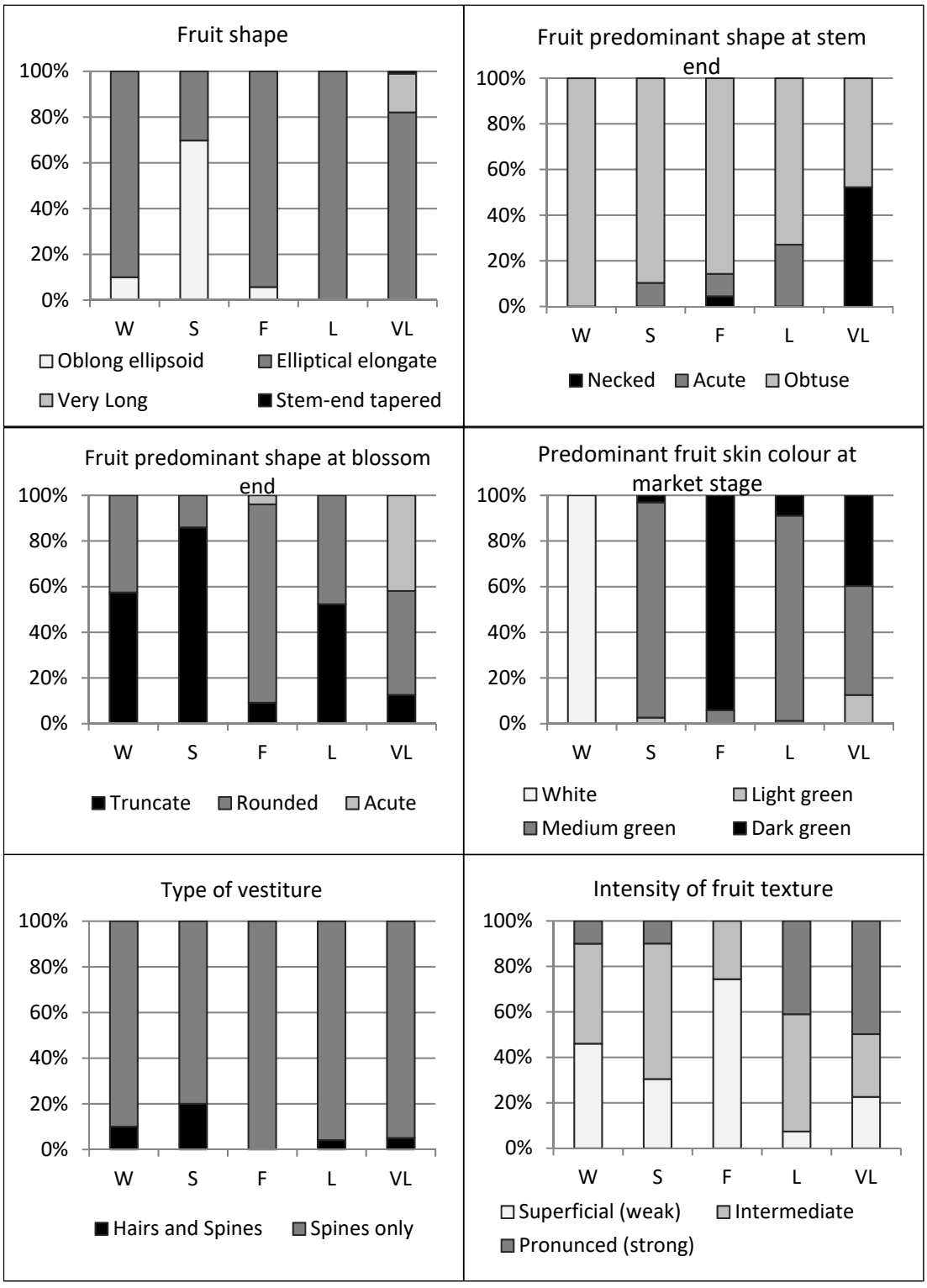
b CHN: China, COG: Congo, ESP: Spain, IRN: Iran, JPN: Japan, LKA: Sri Lanka, NLD: The Netherlands, USA: United States of America.

547

548

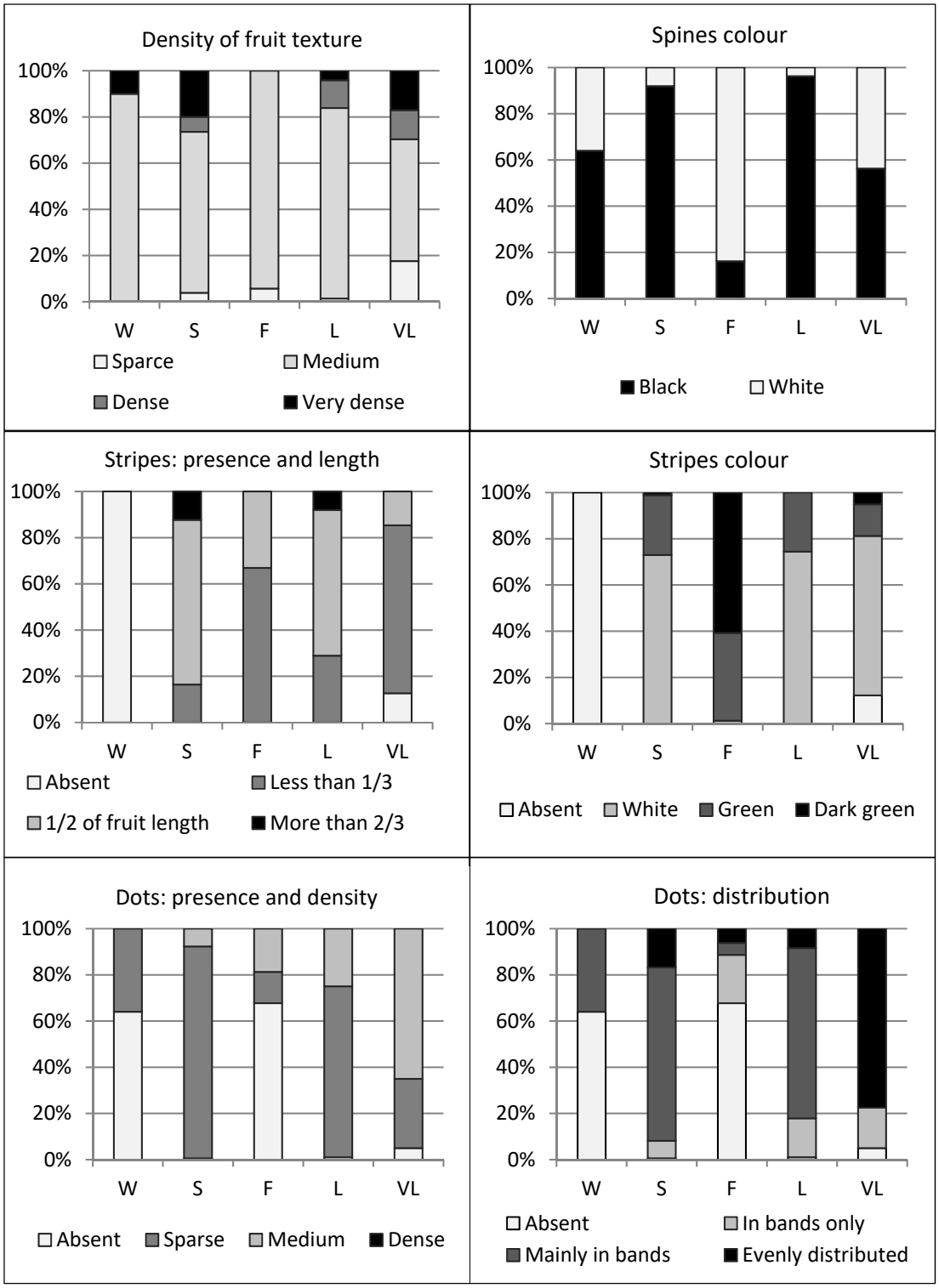
549 Online Resource **S2** Histograms constructed using the qualitative traits in each of the established groups
550 (W='White' type; S='Short' type; F='French' type; L='Long' type; VL='Very long' type)

551



552

553


Density of fruit texture
100%
80%
60%
40%
20%
0%
W
S
F
L
VL
Sparce
Medium
Dense
Very dense
Spines colour
Black
White
Stripes: presence and length
Absent
Less than 1/3
1/2 of fruit length
More than 2/3
Stripes colour
Absent
White
Green
Dark green
Dots: presence and density
Absent
Sparse
Medium
Dense
Dots: distribution
Absent
In bands only
Mainly in bands
Evenly distributed

554

555

556

557

558