1 **Assessing genetic and phenotypic diversity in pepper (*Capsicum annuum* L.)**
2 **landraces from North-West Spain**

3 Antonio Rivera ^a, Ana B. Monteagudo ^b, Ernesto Igartua ^c, Alfredo Taboada ^a, Alba García-Ulloa ^b,
4 Federico Pomar ^b, Manuel Riveiro-Leira ^d, Cristina Silvar ^{a,*}

5

6 ^a Centro de Investigaciones Agrarias de Mabegondo, 15318 A Coruña, Spain

7 ^b Grupo de Investigación en Bioloxía Evolutiva, Departamento de Bioloxía Animal, Bioloxía
8 Vexetal e Ecoloxía, Centro de Investigaciones Científicas Avanzadas (CICA), Universidade da
9 Coruna, 15071 A Coruña, Spain

10 ^c Departamento de Genética y Producción Vegetal, Estación Experimental Aula Dei, CSIC, 50059
11 Zaragoza, Spain

12 ^d Estación Experimental Agrícola do Baixo Miño, 36471 Pontevedra, Spain

13

14 *Corresponding author: Cristina Silvar

15 Tel: +34 981167000

16 E-mail address: c.silvar@udc.es

17

18

19

20

21

22

23

24

25 **Abstract**

26 Pepper (*Capsicum annuum* L.) is one of the most valuable vegetables in the world. Over the last
27 decades, highly performing cultivars have progressively replaced the diversified and heterogeneous
28 landraces worldwide, causing wide genetic erosion in this crop. The recovery of these ancient
29 landraces, which might preserve alleles of agricultural interest and local adaptations, results of
30 pivotal importance for the development of new varieties and the maintenance of a sustainable
31 agriculture. In the present work, a collection of twenty-six landrace-derived inbred lines and three
32 landraces from North-West Spain were evaluated for their agronomic performance and genetic
33 diversity based on a set of twenty-seven morphological descriptors and twenty microsatellite
34 markers. The collection featured phenotypic variability for all the studied traits, which were
35 influenced by the location, except for the yield. The principal component analysis divided the
36 landraces in well-defined groups, with only Arnoia, Punxin and Blanco Rosal showing some degree
37 of overlapping. The greater part of the variance was accounted for traits such as fruit weight,
38 pericarp thickness and fruit shape and colour. The molecular analyses suggested a high level of
39 genetic diversity within the collection and the presence of specific alleles, which were not
40 previously detected in other Spanish pepper landraces. Multivariate and Bayesian clustering showed
41 that landraces were primarily grouped according to their geographical origin and secondarily in
42 agreement with the characteristics of their fruits. Six groups of landraces, with a great genetic
43 differentiation, were clearly identified. Only the landraces Mougan and Arnoia possessed an allele
44 associated to the pungency character.

45

46

47 Keywords: *Capsicum annuum*, pepper, genetic diversity, morphological characterization, SSR

48

49 **1. Introduction**

50 The genus *Capsicum* (family *Solanaceae*) originated in the tropical South American region centered
51 in what is now Bolivia (Eshbaugh, 1993). Nowadays, the number of recognized species in the genus
52 is twenty-seven, five of which were domesticated from distinct events at different primary
53 diversification centers (Andrews, 1995; Baral and Bosland, 2002). Among these five, *Capsicum*
54 *annuum* L. is the most widespread and economically important *Capsicum* species worldwide as well
55 as the most used in commercial cultivar breeding programs (Bosland and Votava, 2000). *C. annum*
56 was domesticated in Mexico from the wild bird pepper or ‘Chiltepin’ (*C. annum* var.
57 *glabriusculum*) and subsequently introduced to Europe by Columbus in the 15th century (Andrews,
58 1995; Perry et al., 2007; Kraft et al., 2014). Afterwards, it was rapidly distributed to Africa, India
59 and China, where it came into wide cultivation giving way to a current crop of immense cultural
60 and economic importance because of its multiple uses (Bosland and Votava, 2000; Kumar et al.,
61 2006). Thousands of years of human selections in multiple environments and cultural contexts led
62 to the impressive phenotypic diversity of contemporary *C. annum* fruits (Nuez et al., 1996; Djian-
63 Caporalino et al., 2007; Nicolai et al., 2013). In general, continued selection was driven to obtain
64 lines with non-deciduous, pendant, larger and non-pungent fruits with greater shape variation and
65 increased fruit mass (Paran and van der Knaap, 2007). The main negative effect of migrations and
66 consequent artificial selections in the secondary diversification centers was the dramatic reduction
67 in the genetic basis of pepper, as occurred for the majority of cultivated species (Tang et al., 2010).
68 Such genetic erosion has become more evident during the past century, when modern breeding
69 promoted the development and worldwide utilization of highly performing commercial cultivars
70 and hybrids, which are genetically uniform and therefore more vulnerable to any biotic or abiotic
71 stress (Votaba et al., 2005). At the present time, with the imminent climate change and the
72 worldwide increasing demand of food, to recover the lost genetic variability and made it available
73 to plant breeders become a mandatory task. In this context, pepper landraces, i.e. native varieties

74 empirically selected by farmers over time and well adapted to the agro-environments in which they
75 have been cultivated for long time, represent valuable reservoirs of genetic diversity that has not
76 been fully exploited (Zeven, 1998; Pacheco-Olvera et al., 2012; Liu et al., 2013). Additionally,
77 landraces are commonly associated to better sensory traits, which would serve to explore new niche
78 markets oriented to worldwide consumers willing to pay higher prices for new pepper varieties
79 combining improved taste, healthy benefits and nutritional properties (Casals et al., 2011; Patil et
80 al., 2014).

81 In Spain, as in other parts of Southern Europe, the versatility of agro-climatic regions and the
82 heterogeneity of the land favoured the survival in cultivation of a large number of specifically
83 adapted pepper landraces very diverse phenotypically (Carravedo et al., 2005; González-Pérez et
84 al., 2014). Part of these landraces was confined to North-West Spain (region of Galicia), where they
85 suffered from selection by farmers, generally directed at fruit attributes, under restricted areas. As a
86 consequence, groups of morphologically recognizable landraces have been defined based on
87 ethnobotanical characteristics in the North (Couto, Couto Grande and Piñeira), Center (Padron and
88 Mougán) and South (Blanco Rosal, Arnoia, Punxín and Oimbra) of Galicia (unpublished data).
89 These landraces are still broadly cultivated in NW Spain, where they sustain the local economy of
90 small horticultural cooperatives. Seeds from different populations of each local variety were
91 collected by the Centro de Investigaciones Agrarias de Mabegondo (CIAM) (A Coruña, Spain),
92 where landrace-derived inbred lines have been obtained during the last decade. This germplasm
93 collection has been partially investigated and data revealed that some of these landraces contain
94 remarkable amounts of vitamin C and antioxidant compounds (unpublished data) and might result
95 exceptionally valuable for the processing industries (Rodríguez-Bao et al., 2004). Such peculiarities
96 make this collection an interesting resource that might be promptly employed for the breeding and
97 scientific communities, as long as a meticulous characterization is being performed. For that
98 purpose, the combination of a morphological approach together with a molecular evaluation will

99 invariably result in more reliable and accurate conclusions (Geleta et al., 2005; Zeinalabedini et al.,
100 2012; Mercati et al., 2014). Among many types of molecular markers that have been developed
101 during the past decades, including the lately released Affymetrix GeneChip array (Ashrafi et al.,
102 2012; Hill et al., 2013), microsatellites or Simple Sequence Repeats (SSRs) are still the most
103 attractive ones for genetic diversity analysis and breeding in plants (Varshney et al., 2005). SSRs
104 are co-dominant, multi-allelic, highly variable and they can be assayed in any laboratory with
105 minimum facilities as well as automated with capillary sequencers for moderate throughput
106 (Schuelke, 2000). In addition, microsatellites performed superiorly than SNPs in resolving
107 population structure (Hamblin et al., 2007; Singh et al., 2013).

108 The main goal of the present work was to evaluate a collection of twenty-six inbred lines derived
109 from nine different pepper landraces and three landraces still cultivated in NW Spain. The genetic
110 relationships and the level of genetic diversity within and among landraces were investigated in a
111 set of twenty publicly available SSR markers. Likewise, the agronomic performance of lines was
112 assessed with twenty-seven morphological traits (IPGR descriptors) in two different locations.

113 **2. Materials and methods**

114 *2.1. Plant material*

115 Twenty-six pepper (*C. annuum*) inbred lines derived from nine different landraces, collected as
116 seeds from farmers in the main growing regions of Galicia (NW Spain) and three landraces (AR45,
117 AR55 and AR65), were used in this work (Table 1, Fig. S1).

118 *2.2. Agro-morphological characterization*

119 The twenty-nine lines were evaluated over three years (2005-2007) in two experimental fields
120 located in Northern and Southern Galicia. The first one was at the Centro de Investigaciones
121 Agrarias de Mabegondo (Mabegondo, A Coruña, Spain) (43° 15'N, 8° 18'W). The second field was
122 located at Estación Experimental Agrícola de Baixo Miño (Salceda de Caselas, Pontevedra, Spain)

(42° 6'N, 8° 33'W), about 10 km far from the border with Portugal. Both areas, approximately 200 Km distant, belong to the same climate region (Mediterranean maritime) according to the classification of Papadakis (1975) and they exhibit little differences in their agro-climatic conditions. During the time of experiments (2005-2007) slightly higher mean temperatures were observed in Salceda (13.5°C) with respect to Mabegondo (11°C), with averages per year of 14°C, 13°C and 13.5°C for the first and 11°C, 12°C and 11.5°C for the second site. On the contrary, the rainfall accomplished better at the latter location (mean values of 85.4 L/m² against 71.2 L/m²). Thus, average values per year (2005, 2006 and 2007) for Mabegondo were 70.1, 105.8 and 80.2 L/m², while 62.8, 82.4 and 68.5 L/m² for Salceda. Seeds were sown under greenhouse conditions and the seedlings were transplanted to the field in the month of April. The experimental design was a complete randomized block with three replications for a total of 84 plants per plot. Agro-morphological data were collected from 10 plants and 25 fruits per replicate. Twenty-seven traits associated to both the plant and the fruit were evaluated on the basis of the *Capsicum* descriptors developed by the International Plant Genetic Resources Institute (IPGRI, 1995): plant height (cm), plant growth habit, plant canopy width (cm), stem length (cm), stem diameter (cm), stem colour, stem shape, stem pubescence, nodal anthocyanin, fruit length (cm), fruit width (cm), fruit weight (g), fruit pedicel length (cm), fruit wall thickness (mm), placenta length (mm), number of locules, fruit set, fruit colour at intermediate stage, fruit colour at mature stage, presence of anthocyanin stripes, fruit shape, fruit shape at pedicel attachment, neck at base of fruit, fruit shape at blossom end, fruit blossom end appendage, fruit cross-sectional corrugation and fruit surface. Harvesting was performed from middle July to middle September and the yield of each line was recorded in kilograms per square meter (kg/m²) and number of fruits per plant (n° fruits/plant).

. Quantitative data were examined with an analysis of variance (ANOVA) using a 99% confidence interval followed by means comparison with Waller-Duncan's test or Least Significant Differences (LSD) test ($\alpha=0.05$). Qualitative characters were expressed as relative frequencies for each landrace

148 at any specific location (Mabegondo or Salceda). In addition, both sets of variables were subjected
149 to principal component analysis (PCA) to obtain a general overview of the structure of variation
150 within and among landraces.. All statistical analyses were performed with SPSS software version
151 17.0 (SPSS, 2008).

152 2.3. *Molecular marker analysis*

153 Genomic DNA was isolated from young leaves of each line following the CTAB method (Doyle
154 and Doyle, 1987). DNA quality was evaluated on agarose gels and DNA concentrations were
155 determined with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington
156 DE). Final DNA concentrations were adjusted to 25 ng μl^{-1} and they were stored at -20°C until used.
157 Twenty publicly available microsatellite markers (thirteen genomic and seven EST-based SSRs),
158 distributed across the twelve pepper chromosomes, were selected according to their observed
159 polymorphism in Spanish pepper landraces (González-Pérez et al., 2014) (Table 2). The sequences
160 of primer pairs were obtained from Lee et al. (2004), Minamiyama et al. (2006), Yi et al. (2006),
161 Ben-Chaim et al. (2006); Nagy et al. (2007). All 20 SSRs were unlinked, according to the genetic
162 maps of Barchi et al. (2007), Wu et al. (2009), Mimura et al. (2012) and Sugita et al. (2013). PCR
163 amplification and detection of microsatellite markers were performed according to the methodology
164 reported by Schuelke (2000) based on the fluorescently labelling of universal primers. Briefly, an
165 M13 tail (5'-CACGACGTTGTAAAACGAC-3') was added to the 5' end of each publicly available
166 forward primer cited above. A universal primer with a complementary sequence to the M13 tail was
167 labelled with different fluorophores (6-Fam, Hex or Ned). PCR was carried out in a final volume of
168 15 μl , which contained 50-75 ng of genomic DNA, 1X PCR Buffer (NZYTech), 2.5 mM MgCl_2
169 (NZYTech), 0.02 μM of the forward primer, 0.2 μM of reverse primer, 0.18 μM of the universal
170 primer, dNTPs (NZYTech) at 0.2 mM each, and 0.4 U of Taq DNA Polymerase (NZYTech). All
171 fragments were amplified using the following touchdown PCR profile: an initial denaturing step of
172 5 min at 94°C was followed by 35 cycles with denaturation at 94°C for 30 s and extension at 72°C

for 30 s, respectively. The annealing temperature was decreased in 0.5°C increments from 62°C in the first cycle to 56°C and was then kept constant for the remaining 35 cycles (always 30–50 s). A final extension step was performed at 72°C for 10 min. DNA fragments were resolved in an Applied Biosystems 3130xl Genetic Analyzer. The allele sizes were assigned with the GeneMapper 3.7 software (Applied Biosystems). The putative presence of the pungency trait in the *C. annuum* lines was assessed by using the molecular marker MAP1, which was developed based on a DNA sequence with high similarity to *Pun1* locus (Rodriguez-Maza et al., 2012).

The number of observed alleles per locus, the observed heterozygosity (H_o), the Nei's unbiased gene diversity index (uH_e) (Nei, 1978) and the Wright's fixation index (F_{is}) were calculated using GenAlex software v6.5 (Peakall and Smouse, 2012). PIC (Polymorphic Information Content) values for each marker were computed according to the formula of Botstein et al. (1980) implemented in the Excel Microsatellite Toolkit (Park, 2001). To determine the genetic uniqueness of each inbred line, the multilocus DNA profile of all the lines was compared by using the software Identity 1.0 (Wagner and Sefc, 1999). Redundancy was determined as the proportion of distinguishable genotypes (Ellstrand and Roose, 1987) and the percentage of genotypes that could be related by mutation was calculated by considering such as those that only differed in one allele (Boccacci et al., 2006; Pereira-Lorenzo et al., 2011).

The marker data were used to generate a 0/1 matrix (presence/absence of allele at the marker locus), which was employed to estimate the genotypic distances between lines. Pairwise similarities were calculated using the Dice coefficient (Nei and Li, 1979). A tree depicting the genetic similarities of the different lines was built from the similarity matrix using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) (1000 bootstrapped). Analyses were carried out using R software (v3.0.2) (R Development Core Team, 2014) and the graphical representation of the tree was performed with MEGA software v6 (Tamura et al., 2013). Genetic relationships among landraces

197 were further assessed with the software NTSYSpc v. 2.1 (Rohlf, 2000) by using a Nei's distance
198 matrix based on allelic frequencies (Nei, 1978).

199 In order to investigate the structure of the collection and to assign individuals to populations based
200 on the SSR genotypes, a Bayesian model-based clustering procedure implemented in the software
201 STRUCTURE 2.3.4 was used (Pritchard et al., 2000). The tests were performed using an admixture
202 model with correlated allele frequencies and by setting the number of populations (K) from 1 to 12,
203 with 10 independent simulations for each K . Each run consisted of a burn-in period of 500,000 steps
204 and 2×10^6 Markov Chain Monte Carlo (MCMC) repetitions. In order to assess the best K value
205 supported by the data, the ΔK method described by Evanno et al. (2005) was used through Structure
206 Harvester v. 6.93 (Earl and vonHoldt, 2012) to examine the rate of change in successive posterior
207 probabilities over the range of K values. Genotypes were assigned to the group for whom they had
208 the highest membership coefficient, considering strong affinity when the membership coefficient
209 (qI) was $\geq 80\%$ (Breton et al., 2008; Ferreira dos Santos et al., 2011). CLUMPP (Jakobsson and
210 Rosenberg, 2007) was employed to generate Q values for groups from the STRUCTURE data using
211 the Greedy K algorithm. Genetic variation between groups and sub-groups defined by UPGMA and
212 STRUCTURE was assessed with an analysis of molecular variance (AMOVA) by using the
213 software Arlequin v. 3.5 (Excoffier et al., 2005).

214 3. Results

215 3.1. Morphological characterization

216 A broad phenotypic variation was found for most of the parameters studied in the collection of
217 pepper lines. The distribution of the different qualitative morphological characters was represented
218 as the percentage of relative frequencies for each group of landraces at any different location (Fig.
219 S2). Little discordance was observed between the two experimental sites. Among the agronomic
220 traits, all the pepper plants showed a growth habit erect, the stem colour was mostly green with
221 sparse pubescence, except for Mougan, which showed an intermediate-dense level of pubescence.
222 Stem shape was cylindrical and anthocyanin was visible in the internodes, varying from light to
223 dark purple (Fig. S2). The qualitative fruit characters showed broader variation. Fruit set ranged
224 from intermediate to high and fruit shape varied from triangular (Blanco Rosal, Oimbra and Punxin)
225 to blocky (Couto Grande, Piñeira, Mougan and Padron) (Fig. S2). Fruit colour at intermediate stage
226 was white for Blanco Rosal, Oimbra and Punxin, and green for the others, whereas fruit colour at
227 mature stage mostly varied between red and dark red. All fruits had a smooth surface and only few
228 showed anthocyanin stripes (Fig. S2).

229 The analysis of variance for quantitative descriptors was firstly performed independently at each
230 experimental location, in order to check the potential effect of variable climatic conditions over the
231 evaluated periods. Significant differences among years were found at each site for all traits, except
232 for fruit length (Mabegondo), fruit wall thickness and number of locules (Salceda). Differences in
233 yield traits were detected only in Mabegondo (Table S1). Nevertheless, the percentage of variance
234 explained by the 'year' together with the interaction 'landrace $\times$ year' was small compared to that
235 arising from the genotypic effects (Table S1), and it mostly resulted from divergent means observed
236 in 2006 (data not shown). ANOVA analysis considering the two different locations showed
237 significant ($p < 0.001$) differences for all of the quantitative characters under study. In general,
238 differences were detected among landraces, environments and for the interaction 'landrace $\times$

location' (Table 3). The yield, expressed as n° fruits/plant and kg/m² was slightly influenced by the experimental site. Thus, 'landrace × location' interaction was not significant for the trait kg/m², but the location showed a significant effect on the n° fruits/plant (Table 3). Couto and Padron exhibited the highest yield for both places (141 and 146.6 kg/m², respectively), whereas Couto Grande and Piñeira produced the lowest ones (12.6 and 14 kg/m², respectively), although these were not statistically different from those obtained with Oimbra, Punxin, Arnoia and Blanco Rosal (Table S2). A similar trend was observed for the fruit weight trait, being much lower in Couto and Padron (27.4 and 28.1g, respectively), compared to Couto Grande (193.9g). In particular, the effect of 'landrace' was predominant for the majority of fruit and plant characters, except for stem diameter and fruit pedicel length, whose variability was better explained by the location effect (Table 3). Location also displayed a notable influence on characters such as plant height and width. In general, for those traits, the region of Mabegondo exhibited higher values than Salceda (Table S3). The landraces that displayed a greater differentiation depending of the environment were Blanco Rosal, Mougan and Punxin, especially for the attributes plant height, plant width and fruit weight (Table S3). Only the total number of harvested kilograms per square meter were not significantly different at both locations, the mean value for this character at Mabegondo being of 53.2 kg/m², while 58.7 kg/m² for Salceda (Table S3).

A principal component analysis was performed separately with both sets of descriptors in order to more accurately determine the most effective attributes in discriminating among landraces, and to avoid a potential reduction in the variance due to excessive number of variables or the presence of scale-based traits. Those traits related to yield were considered as quantitative characters. PCA based on quantitative data resulted in three principal components with eigenvalues >1, cumulatively accounting for 85.61% of the total variance (Table S4). The first component explained 47.21% of the total variance and it was positively and strongly correlated with fruit weight, width and wall thickness, and negatively correlated in a robust manner with kg/m² (Table S4). The second

264 component, which explained 29.98% of the total variance, had plant height, width and stem length
265 with the highest coefficients. Finally, the third component contributed only for 8.42% of the total
266 variance and the n° fruits/plant was the main trait responsible for the observed variability (Table
267 S4). These three components were used to obtain the diagram of dispersion for all the twenty-nine
268 lines, giving the picture of the differences among the nine groups of landraces (Fig. 1). Couto
269 Grande, Piñeira, Oimbra, Padron, Couto and Mougan, were represented as well-defined clusters,
270 while Arnoia and Punxin, and Blanco Rosal to a lesser extent, showed some degree of overlapping
271 (Fig. 1). The first axis, mainly related to variation in fruit characters, clearly separated the landraces
272 with large fruits (e.g. Couto Grande, Oimbra and Piñeira) from those with the smallest ones
273 (Padron, Couto and Mougan) (Fig. 1, Table S2). The second component, mostly concerned with
274 plant traits, differentiated Padron (tallest plants with highest width and stem length) from Mougan
275 (shortest plants with lowest width and stem length) (Fig. 1, Table S2). The third axis did not allow
276 resolving the overlay among Arnoia, Punxin and Blanco Rosal, as these landraces displayed similar
277 values for the n°fruits/plant trait (Fig. 1, Table S2). Regarding the qualitative characters, those
278 descriptors that showed little polymorphism or varied in only one landrace were not considered for
279 the principal component analysis. In that case, the three first components of the PCA accounted,
280 respectively, for 43.41, 16.13 and 14.61% of the total variance (Table S4). The first component was
281 positively correlated with the fruit shape at blossom end, the fruit colour and fruit shape, whereas
282 the second and third components were mainly correlated with the corrugation and the fruit shape at
283 pedicel attachment, respectively (Table S4). The projection of the genotypes on a two-dimensional
284 plot revealed that the accessions of each landrace presented a higher degree of dispersion in the
285 graph than that observed in the above PCA (Fig. 2). Nevertheless, a better resolution was observed
286 for Arnoia, Punxin and Blanco Rosal, which showed decreasing values for the first component,
287 supported by descriptors such as fruit colour and shape (Fig. 2).

288 3.2. *Molecular analysis*

289 All amplified loci showed polymorphism in the analysed collections. The 20 SSR markers produced
 290 a total of 60 alleles in the 29 pepper landrace-derived inbred lines (Table 2), ranging from 2
 291 (Hpms1-148, HmpsE075, Hpms2-2, Hpms2-24, CA523558, HpmsE145, EPMS342, HpmsE082) to
 292 7 (GPMS197), with an average of 3 distinct alleles per locus. Allele sizes varied between 118 bp
 293 (EMPS426) and 360 bp (EMPS342). The observed heterozygosity (H_o) was zero for 15 out of 20
 294 markers, and for the others ranged from 0.034 (HpmsE075) to 0.103 (GMPS197) (Table 2). The
 295 lowest Nei's unbiased gene diversity (uH_e) was 0.19 (HpmsE082) and the maximum was 0.84
 296 (GMPS197). F_{is} values were positive and close to one for the twenty tested SSRs (Table 2). The
 297 polymorphic information content (PIC), which depicts the number of alleles and their distribution,
 298 was calculated to determine the informativeness of each marker. The average PIC value for the
 299 whole collection of pepper landraces was of 0.44. The most informative marker was GMPS197,
 300 with a PIC value of 0.82, while the minimum PIC was recorded for HpmsE082 (PIC=0.17) (Table
 301 2).

302 The H_o and Nei's unbiased gene diversity index (uH_e), which summarize the fundamental genetic
 303 variation of a group, were calculated in the different pepper landraces, excluding those with only
 304 one line (Table 4). The observed heterozygosity was very low, with a mean value of 0.014 for the
 305 whole collection, and only Couto, Mougán, Oimbra and Piñeira exhibited few heterozygous
 306 genotypes. The Nei's index displayed an average of 0.51 for the combination of lines, the most
 307 diverse landrace being Arnoia (uH_e =0.21), whereas Blanco Rosal, Padron and Punxin did not
 308 display any differentiation among their derived lines. The average value for the inbreeding
 309 coefficient (F_{is}) in the entire group of landraces was high (0.97), confirming the low level of
 310 heterozygosity among the genotypes analyzed (Table 4). Twelve out of sixty alleles (20%) were
 311 considered as unique alleles, i.e., alleles present in only one landrace and absent in the others. Five
 312 of them (8.3%) in five different loci (HpmsE016, HpmsE031, Hpms1-143, HpmsE082 and Hpms1-
 313 139) were recorded in inbred lines derived from Couto (Table 4). Analysis with the software

314 Identity revealed that four groups of landraces (Blanco Rosal, Punxin, Padron and Mougan)
315 possessed lines with the same SSR profile. The size of duplicate genotypes varied from two in
316 Blanco Rosal and Punxin, to six in Padron. Within the remaining nineteen unique genotypes,
317 differences involving a single allele were detected between Co12B and Co5A. Similarly, few lines
318 from distinct landraces differ with others in only two alleles and they might be also related by
319 mutation. This is the case for Blanco Rosal, which diverge from AR330-05 at two loci or Couto
320 Grande, which shared 90% of its marker profile with Piñeira. The potential pungency of the pepper
321 accessions was investigated with the specific marker MAP1. Five inbred lines derived from
322 Mougan (MG14-04, MG77-04, MG279-05, MG318-05 and MG335-05) and two from Arnoia
323 (AR45 and AR55) carried the allele of 494 bp associated to pungency. The remaining lines
324 displayed the non-pungent allele of 479 bp (data not shown).

325 Those genotypes previously identified as redundant were not considered in cluster and
326 STRUCTURE analyses, although they were maintained in Fig. 3 for the sake of clarity. Pairwise
327 similarities among lines were computed using the Dice coefficient and the resulting matrix was
328 employed to generate a dendrogram using the UPGMA clustering algorithm. Based on the
329 similarity of their alleles, all inbred lines derived from Couto, Couto Grande and Piñeira were
330 clustered in a well-defined group with a bootstrap of 93.1 %, and they differed from the other
331 landraces at a similarity level of ca. 0.34 (Fig. 3a). At a coefficient of similarity of 0.5, the
332 collection of pepper landraces from Galicia was divided into three major clusters: cluster A (11
333 lines), cluster B (12 lines) and cluster C (6 lines), supported by bootstrap values of 100%, 70.9%
334 and 93.1%, respectively (Fig. 3a). Clustering pattern was primarily based upon geographical origin:
335 cluster A corresponds with landraces original from Southern Galicia (Arnoia, Blanco Rosal, Oimbra
336 and Punxin), while cluster B includes landraces from the central part (Padron and Mougan), and
337 cluster C comprises landraces derived from the Northernmost region of Galicia (Fig. 3a). All three
338 clusters could be further split into smaller groups, most of them also supported by good confidence

339 values. Thus, cluster A was sub-divided in two sub-groups: A.1 (Oimbra and Punxin) and A.2
 340 (Arnoia and Blanco Rosal), cluster B consisted of two other sub-groups: B.1 (Padron, MG14-04 and
 341 MG101-04) and B.2 (Mougan), and finally cluster C was divided into sub-group C.1 (Couto) and
 342 C.2 (Couto Grande and Piñeira) (Fig. 3a). The UPGMA dendrogram based on Nei's genetic
 343 distances defined by SSR markers confirmed that landraces were clearly grouped in three clusters
 344 following a geographical pattern (Fig. S4). The narrowest genetic distance (0.21) was found
 345 between Arnoia and Blanco Rosal, suggesting that these two materials share a common genetic
 346 background. Couto Grande was closely related to Piñeira (0.33) and these two peppers, together
 347 with Couto, constituted the Northern group and positioned rather distant (0.93) from the other
 348 clusters (Fig. S3). Genetic relationships among landrace-derived pepper lines were further
 349 investigated with the software STRUCTURE, which allowed assigning individuals to specific
 350 populations. The log probability of data ($\ln[\Pr(X/K)]$), after the Evanno et al. (2005) correction,
 351 showed a clear maximum for $\Delta K=6$. The majority of lines had membership coefficients $qI > 80\%$
 352 for each specific group (Fig. 3b), and only MG14-04 and MG101-04 showed admixture ($qI < 80\%$),
 353 possessing a genomic composition between Padron and Mougan. The partitioning level observed
 354 after the Bayesian analysis was clearly in agreement with the six groups (A.1, A.2, B.1, B.2, C.1,
 355 C.2) previously defined by the cluster division (Fig. 3b). As stated above, the lines distributed
 356 among the six inferred populations on the basis of their geographical origin. The further exploration
 357 of increasing values of K , allowed a greater differentiation of groups of landraces (Fig. 3c). Thus, at
 358 $K=8$, Arnoia and Blanco Rosal (A.1) were split into two different populations, although inbred lines
 359 derived from Arnoia still maintained some degree of ancestry (qI between 6.8% and 63%) with
 360 Blanco Rosal lines (Fig. 3c). Similarly, the group A.2 became divided into two smaller groups
 361 composed by Oimbra and Punxin, respectively. The other groups (B.1, B.2, C.1, and C.2) remained
 362 inseparable (Fig. 3c). A detailed analysis of results suggested that apart from their origin, the
 363 genetic relationships among all sub-groups showed rather good coherence with pepper plant and

364 fruit traits, such as it was deduced by the output obtained after morphological assessment. Thus,
 365 cluster A was characterized by plants with large fruits (weight>80 g, width>5 cm, pericarp
 366 thickness>3mm) mostly triangular and white at the intermediate stage of ripening. Cluster B
 367 basically include landraces with smaller fruits (weight<40 g, width<5 cm, pericarp
 368 thickness<3mm), which are blocky in shape and green colour in the plant. Interestingly, cluster C
 369 comprises lines with both small and large blocky fruits with green colour; Couto possessed the
 370 characteristics associated to small peppers, while Couto Grande and Piñeira displayed traits typical
 371 of larger berries (Fig. 3c, Table S2). Subsequent subdivisions might be explained by other
 372 qualitative and quantitative characters, which clearly differ among landraces within each cluster
 373 (Table S2).

374 AMOVA was performed between all lines, grouped according to cluster and STRUCTURE outputs.
 375 Analysis of clusters A, B and C indicated that 61.65% of the variation was due to differences within
 376 groups, while 38.35% was due to differences among groups (Table 5). The overall F_{ST} value of 0.38
 377 suggested a large genetic differentiation between groups of landraces. Pairwise estimates of F_{ST}
 378 using AMOVA revealed a significant degree of differentiation (p value between 0.0001 and 0.05)
 379 among the three groups with values ranging from 0.23 to 0.44 (between A and B: 0.44 ($p<0.0001$),
 380 A and C: 0.23 ($p<0.05$), B and C: 0.40 ($p<0.0001$)). The analysis of molecular variance at the six
 381 hierarchical levels showed that most of the variation occurred among groups (71.46%), although a
 382 considerable proportion was also found within groups (28.54%) (Table 5). In this case, the greatest
 383 genetic differentiation was found between A.1 and B.2. ($F_{ST}=0.46$, $p<0.05$), while the lowest was
 384 observed between B.1 and C.1 ($F_{ST}=0.88$, $p<0.0001$).

385 4. Discussion

386 A collection comprising twenty-nine inbred lines, mostly derived from nine pepper landraces
387 collected at different regions in NW Spain was evaluated with twenty-seven agro-morphological
388 descriptors and twenty SSR markers. These old landraces, selected locally by farmers across many
389 years on the basis of a recognizable performance, are still cultivated nowadays because of their high
390 quality and good acceptance in the national market. On a global scale, this collection represents a
391 valuable source of unexploited variability in a cultivated background that might be easily employed
392 with breeding purposes. For that goal, the accurate assessment and description of trait variation is of
393 crucial significance in the success of any program aimed at the selection of genotypes harboring
394 those features mainly demanded by consumers. Similarly, the evaluation of genetic variability
395 among genotypes results essential for the conservation and protection of any genetic resource, but
396 also for broadening the genetic basis of cultivated varieties, promoting a sustainable agriculture
397 (FAO 2010; Ortiz et al., 2010; Sparato and Negri, 2013). Both approaches are addressed in the
398 present work in order to gain a better knowledge on Galician pepper landraces and design suitable
399 strategies for their future exploitation.

400 Agronomic results of field trials pointed out a notable phenotypic variability among landraces.
401 Qualitative traits were not highly affected by the location and only a few discrepancies arose among
402 and within landraces, which could be also attributed to the ambiguity of any visual classification
403 based on pre-established morphological categories. On the contrary, the year and location of field
404 trials exerted a certain influence on the majority of quantitative descriptors, although the latter
405 possessed a more remarkable effect. Productivity, expressed as kg/m^2 , was significantly different
406 over the three assayed years in Mabegondo, but it was not significantly affected by the regional
407 conditions, performances of all distinct landraces being comparable between both experimental
408 sites. Indeed, differences in yield among lines within each landrace were neither detected for this
409 parameter, suggesting a high adaptability of the different genotypes. Other parameters associate to

410 yield, such as the number of fruits per plant and fruit weight were significantly affected by the
411 location, although in those cases the ‘landrace $\times$ location’ interaction represented a small percentage
412 of the total variance. Indeed, a low correlation (20%) was found between the number of fruits per
413 plant and the kg/m^2 . In opposition to productivity, environmental conditions showed a certain effect
414 on other landraces traits mostly associated to plant and fruit growth, such as plant width and height,
415 stem diameter and fruit pedicel length. Such effect might be attributed to slightly variable climatic
416 parameters among years and locations. For instance, rainfalls were superior in 2006 and the average
417 water availability was greater in the Northern site, which might explain in part the better
418 performance of the majority of landraces at this location. Nevertheless, these environmental
419 variations were not sufficient to modify the productivity of landraces, suggesting that the different
420 genotypes could easily adapt to distinct agro-climatic areas.

421 The most representative traits describing the phenotypic diversity of the genotypes were defined by
422 PCA analysis. Thus, nine groups of landraces were primarily separated on the basis of fruit
423 dimension, weight, pericarp thickness and productivity (kg/m^2). Only Arnoia, Punxin and Blanco
424 Rosal showed some degree of overlapping for quantitative traits, which was resolved when the
425 qualitative descriptors were investigated. Among qualitative traits, fruit shape, colour and blossom
426 end were the ones that most contributed to the organization of genotypes in the space defined by the
427 first component. However, in this case, the distinctive genotypes from specific landraces were
428 plotted in a more disperse manner. Such organization might be attributed to the influence of other
429 characters but also to the presence among the different landraces of certain intermediate
430 morphotypes hard to classify based on a visual inspection (Portis et al., 2006; Bozokalfa et al.,
431 2009; Yumnam et al., 2012). Plant characters such as plant height, plant width and stem length were
432 the most important attributes in the second component, featuring 32.8% of total variation, and
433 drastically separating Padron and Mougán. Earlier works also pointed to plant traits as less
434 significant contributors to the classification of accessions in clusters (Bozokalfa et al., 2009;

Yumnam et al., 2012). Groupings primarily based on fruit traits agreed with those previously observed in pepper landrace collections from Italy and Turkey, although in those reports plant descriptors explained a variable percentage of the variance ranging from 13.9% (Bozokalfa et al., 2009) to 62.4% (Portis et al., 2006). Similar trends were also reported for other *Capsicum* species and closely related horticultural crops, like tomato (Mazzucato et al., 2008; Carvalho et al., 2014; Mercati et al., 2014). Results suggest that each landrace have been strongly selected at every specific region using basic agro-morphological traits founded on farmers' preferences. Globally, this agro-morphological characterization provided useful information for farmers and breeders in order to set up well-defined horticultural types and easily select those which better adapt to specific uses. The uniformity of Galician landraces, but also the occurrence in the collection of uncommon types, such as landraces with large, fleshy and pungent fruits (AR45 and AR55) and landraces with small, non-pungent berries (Padron and MG101-04), together with the presence in several landraces of remarkable amounts of nutraceutical compounds (unpublished data) will serve to develop pepper varieties with increased added value and to explore innovative market niches in the short term.

Besides morphological traits, molecular markers are a powerful tool to discriminate for genetic diversity in crops, especially the microsatellites, which are able to identify distinctive alleles. The collection of pepper landraces from NW Spain was evaluated with a set of twenty microsatellite loci. In total, 60 alleles were detected, with a mean of 3 alleles per locus. These values were lower than those previously reported by González-Pérez et al. (2014) (117 alleles/average 4) in a collection of pepper landraces from Spain. However, if we consider that the current collection consist of a smaller number of landraces coming from a restricted area in Spain, data suggests that the Galician set still maintain a noteworthy level of genetic diversity. Indeed, the average number of alleles per locus was similar or even higher than those reported for larger and more diverse panels of *C. annuum* genetic resources (Kwon et al., 2005; Minamiyama et al., 2006; Hanacek et al., 2009; Pacheco-Olvera et al., 2012; Rai et al., 2013; Nimmakayala et al., 2014). Similarly, PIC values,

460 which provide an estimate of the discriminatory power of each SSR locus, showed a comparable
461 range and average to those reported above and, most importantly, to the Spanish group (0.17-0.82
462 Galician vs 0.13-0.87 Spanish; 0.44 Galician vs 0.45 Spanish). Twelve private alleles were detected
463 in seven landraces and six of them are fixed in Blanco Rosal, Couto, Padron and Punxin (allele
464 frequencies of 100%). Four of those also appeared in the Spanish collection as specific alleles not
465 detected in non-Spanish *C. annuum* accessions (González-Pérez et al., 2014) and one allele of 254
466 bp associated to marker HpmsE082 was exclusively identified in all Couto-derived inbred lines.
467 The presence of private alleles has been reported for other *Capsicum* collections (Hanacek et al.,
468 2009; Ibiza et al., 2012; Nicolai et al., 2013) and it might be of adaptive significance, as a result of
469 selection for adaptation to local constraints. Capture and assessment of such specific alleles and
470 genotypes should be an important objective of any conservation and exploitation strategies (Paran
471 and Fallik, 2011).

472 The observed heterozygosity (H_o) was very low while the coefficient of inbreeding (F_{is}) was close
473 to one in the entire collection, as expected considering the highly inbred nature of *C. annuum* and
474 the fact that we are dealing with a set of landrace-derived inbred lines (Onus and Pickersgill, 2004).
475 The genetic diversity, expressed as a measure of the Nei's unbiased gene diversity index, was
476 similar to that observed in Spanish and French *C. annuum* collections (Nicolai et al., 2013;
477 González-Pérez et al., 2014) and higher than those reported in studies with larger diverse sets of *C.*
478 *annuum* landraces from Mexico, Colombia, Italy or Turkey (Lanteri et al., 2003; Toquica et al.,
479 2003; Oyama et al., 2006; Aktas et al., 2009; Pacheco-Olvera et al., 2012). The genetic diversity
480 among the Arnoia lines was the highest ($uHe=0.21$), because this group included the landraces
481 AR45, AR55 and AR65, which are not inbred lines.

482 Multivariate and population structure approaches clearly resolved the *C. annuum* lines into greatly
483 distinctive clusters, primarily based on their geographical origin and secondarily founded on fruit
484 characters. Similar clustering according to a geographical pattern was detected by Portis et al.,

(2006), who observed a division of Italian pepper landraces in Southern and Northern groups, and also reported in other *Capsicum* spp (Albrecht et al., 2012; Moses et al., 2014). Subsequent Bayesian analysis reinforced the presence of six minor clusters ($\Delta K=6$) within the three main geographical groups and pointed to the assignment of genotypes to each population according to fruit characteristics. This clustering pattern attending to fruit traits is in agreement with the overall trend reported in *C. annuum* during domestication and selection. . Thus, ancestral small red hot peppers, would suffer from continuous selection to give larger pepper fruits with increased mass, variable colour and absence of pungency (Paran et al., 1998; Portis et al., 2006; Paran and van der Knaap, 2007; Tam et al., 2009; Hill et al., 2013). So far, numerous studies have found notable distinctions among groups of lines with small and pungent peppers, lines with elongated and pungent fruits and lines with huge and blocky or triangular sweet types (Lefebvre et al., 1993; Nagy et al., 2007; Geleta et al., 2005). Plants with elongated hot peppers ('Cayenne' type) are not represented in the Galician collection. However, large and fleshy fruited landraces carrying a non-pungent allele in the MAP1 locus could be recognized mainly in clusters A.1 and A.2 and partially (Piñeira and Couto Grande) in cluster C. The separation of large sweet peppers into two well-differentiate clusters suggest that these landraces might originated from two distinct genetic pools. This hypothesis is supported by the fact that Piñeira and Couto Grande are located at the northernmost part of Galicia and their berry morphology was more blocky ('Bell' type) than triangular, bearing an explicit sunken blossom end. Cluster B comprised Padron and Mougan, which are characterized by small and thin-walled fruits. All inbred lines derived from Mougan possessed an allele for pungency, while this was not detected in Padron. This might be explained because Padron lines have been selected for the absence of pungency from a former landrace, which was originally pungent. The remaining agro-morphological traits were maintained highly similar to the ancestral genetic material. The F_{ST} values observed after the AMOVA confirmed a greater level of differentiation among groups, suggesting that the different landraces have likely been under

510 confined selection long enough for specific genes to be gained or lost. Interestingly, such estimates
511 became smaller between groups of landraces belonging to neighboring areas, pointing to some gene
512 flow in the past (Pressoir and Berthaud, 2004).

513 The landrace Couto outlined as the most distinctive one, being positioned at the highest Nei's
514 genetic distance according to the phylogenetic tree. In a previous work, this landrace together with
515 Padron constituted a specific cluster with a characteristic genomic constitution closely related to
516 elongated peppers from the South of Spain and with various alleles in common to the group of
517 Spanish *C. annuum* 'Serrano' types (González-Pérez et al., 2014). Interestingly, this landrace
518 retained one of the key traits that was left behind after domestication, as it is the erect peduncle of
519 their flowers and fruits (Paran and van der Knaap, 2007) but, on the contrary, it does not possess the
520 pungency character, commonly associated to small pepper types (Ortiz et al., 2010; Qin et al.,
521 2014). The first peppers travelled from Central America to Spain in post-Columbian times and
522 quickly spread throughout the Iberian Peninsula and the rest of Europe (Andrews, 1995). Although
523 further studies are needed it cannot be ruled out that the Couto landrace was originated from these
524 ancient peppers, part of which promptly migrated to the North Western part of Spain, suffering
525 from isolation and independent selection pressure in confined geographical areas.

526 Just as phenotypic characterization, molecular assessment pointed at well-defined groups of
527 Galician pepper landraces possessing noteworthy levels of genetic diversity, comparable to those
528 observed in larger *C. annuum* sets. Marker data will guaranty a more efficient conservation of this
529 genetic resource and it will help breeders straightaway to speed up the selection process of
530 improved varieties. Farmers might also benefit directly from the genotypic information generated in
531 this work. Thus, a database holding the genetic profiles expected for each landrace was constructed
532 and molecular data have been already used in elucidating seed admixture at some horticultural
533 cooperatives that grow Padron and Couto, two landraces very similar in gross morphology.
534 Similarly, this genetic information, complemented with additional work, will be extremely useful

535 for the future typification and protection of these pepper landraces in extensive commercial
536 exploitations.

537 **5. Conclusions**

538 In the present work, a multidisciplinary approach was carried out with the goal of assessing the
539 variability within a collection of pepper landraces from NW Spain. The results revealed differences
540 among all studied landraces at the phenotypic and genotypic levels. Well-defined and distinctive
541 groups of landraces were observed, supporting the existence of the nine groups established *a priori*
542 based on ethnobotanical criteria. Furthermore, data pointed to the collection as a valuable reservoir
543 of genetic and phenotypic diversity that might be further exploited in breeding programs for the
544 achievement of pepper varieties with interesting properties, which could enhance farmer's
545 opportunities for entering new markets.

546

547

548

549

550

551

552

553

554

555

556

557

558 **Acknowledgements**

559 This work was supported by Xunta de Galicia (Project EM2014/048) and by the Spanish Institute
560 for Agricultural and Food Research and Technology (INIA), co-financed by the European Regional
561 Development Fund (FEDER) (Project RTA2011–00118). The authors are grateful to Dr. Fernanda
562 Rodriguez (SAI, University of Coruña) for great technical support.

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581 **References**

- 582 Andrews, J., 1993. Diffusion of mesoamerican food complex to southeastern Europe. *Geograph.*
583 *Rev.* 83, 194–204.
- 584 Andrews, J., 1995. *Peppers: The domesticated Capsicums*. Austin: University of Texas Press.
- 585 Aktas, H., Abak, K., Sensoy, S., 2009. Genetic diversity in some Turkish pepper (*Capsicum*
586 *annuum* L.) genotypes revealed by AFLP analyses. *African J. Biotech.* 8, 4378–4386.
- 587 Ashrafi, H., Hill, T., Stoffel, K., Kozik, A., Yao, J., et al., 2012. De novo assembly of the pepper
588 transcriptome (*Capsicum annuum*): a benchmark for *in silico* discovery of SNPs, SSRs and
589 candidate genes. *BMC Genom.* 13, 571–586.
- 590 Baral, J.B., Bosland, P., 2002. An updated synthesis of the *Capsicum* genus. *Capsicum Eggplant*
591 *News.* 21, 11–21.
- 592 Barchi, L., Bonnet, J., Boudet, C., Signoret, P., Nagy, I., et al., 2007. A high-resolution,
593 intraspecific linkage map of pepper (*Capsicum annuum* L.) and selection of reduced recombinant
594 inbred line subsets for fast mapping. *Genome* 50, 51–60.
- 595 Ben-Chaim, A., Borovsky, Y., Falise, M., Mazourek, M., Kang, B.C., Paran, I., Jahn, M., 2006.
596 QTL analysis for capsaicinoid content in *Capsicum*. *Theor. Appl. Genet.* 113, 1481–1490.
- 597 Boccacci, P., Akkak, A., Botta, R., 2006. DNA typing and genetic relations among European
598 hazelnut (*Corylus avellana* L.) cultivars using microsatellite markers. *Genome* 49, 598–611.
- 599 Bosland, P.W., Votava, E.J., 2000. *Peppers: vegetable and spice Capsicum*. CABI Publishing,
600 Oxon. 204 p.
- 601 Bozokalfa, M.K., Esiyok, D., Turhan, K., 2009. Patterns of phenotypic variation in a germplasm
602 collection of pepper (*Capsicum annuum* L.) from Turkey. *Spanish J. Agri. Res.* 7, 83–95.
- 603 Botstein, D., White, R., Skolnick, M., Davis, R., 1980. Construction of genetic linkage map in man
604 using restriction fragment length polymorphism. *Am. J. Hum. Genet.* 32, 314–331.
- 605 Breton, C., Pinatel, C., Medail, F., Bonhomme, F., Berville, A., 2008. Comparison between
606 classical and Bayesian methods to investigate the history of olive cultivars using SSR–
607 polymorphisms. *Plant Sci.* 175, 524–532.
- 608 Carravedo, M., Ochoa, M.J., Gil, R., 2005. Catálogo genético de pimientos autóctonos conservados
609 en el Banco de Germoplasma de Hortícolas de Zaragoza. Gobierno de Aragón, España.
- 610 Carvalho, S.I.C., Ragassi, C.F., Bianchetti, L.B., Reifschneider, F.J.B., Buso, G.S.C., Faleiro, F.G.,
611 2014. Morphological and genetic relationships between wild and domesticated forms of peppers
612 (*Capsicum frutescens* L. and *C. chinense* Jacquin). *Genet. Mol. Res.* 13, 7447–7464

613 Casals, J., Pascual, L., Cañizares, J., Cebolla–Cornejo, J., Casañas, F., Nuez, F., 2011. The risks of
614 success in quality vegetable markets, possible genetic erosion in Marmande tomatoes (*Solanum*
615 *lycopersicum* L.) and consumer dissatisfaction. *Sci. Hort.* 130, 78–84.

616 Djian–Caporalino, C., Lefebvre, V., Sage–Daubeze, A.M., Palloix, A., 2007. Capsicum, in: Singh,
617 R.J., (Ed.), Genetic resources, chromosome engineering and crop improvement series, Vol 3,
618 Vegetable crops. CRC Press, Florida (USA). pp. 185–243.

619 Doyle, J.J., Doyle, J.J., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf
620 tissue. *Phytochem. Bull* 19, 11–19.

621 Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for
622 visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4,
623 359–361.

624 Ellstrand, N.C., Roose, M.L., 1987. Patterns of genotypic diversity in clonal plants species. *Am. J.*
625 *Bot.* 74, 123–131.

626 Eshbaugh, W.H., 1993. Peppers: History and exploitation of a serendipitous new crop discovery, in:
627 Janick, J., Simon, J.E., (Eds). *New crops*. Wiley, New York (USA). pp. 132–139.

628 Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the
629 software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620.

630 Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: An integrated software package for
631 population genetics data analysis. *Evol. Bioinform. Online* 1, 47–50.

632 FAO., 2010. The second report on the state of the world’s plant genetic resources for food and
633 agriculture. Rome, 370 p. Available: <http://www.fao.org/docrep/013/i1500e/i1500e00.htm>.
634 Accessed 2013 December.

635 Ferreira dos Santos, A.R., Ramos–Cabrer, A.M., Díaz–Hernandez, M.B., Pereira–Lorenzo, S.,
636 2011. Genetic variability and diversification process in local pear cultivars from Northwestern
637 Spain using microsatellites. *Tree Genet. Genomes* 7, 1041–1056.

638 Geleta, L.F., Labuschagne, M.T., Viljoen, C.D., 2005. Genetic variability in pepper (*Capsicum*
639 *annuum* L.) estimated by morphological data and amplified fragment length polymorphism
640 markers. *Biodiversity Conserv.* 14, 2361–2375.

641 González–Pérez, S., Garcés–Claver, A., Mallor, C., Sáenz de Miera, L.E., Fayos O, Pomar F,
642 Merino F, Silvar C., 2014. New insights into *Capsicum* spp relatedness and the diversification
643 process of *C. annum* in Spain. *Plos One* 9, 1–23.

644 Greenleaf, W.H., 1986. Pepper breeding, in: Bassett, M.J., (Ed.) *Breeding vegetable crops*. AVI
645 Pub. Co., Westport (USA), pp. 67–134.

646 Hamblin, M.T., Warburton, M.L., Buckler, E.S., 2007. Empirical comparison of simple sequence
647 repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. *PloS*
648 *One* 2, e1367.

649 Hanacek, P., Vyhnánek, T., Rohrer, M., Cieslarová J., Staveliková, H., 2009. DNA polymorphism
650 in genetic resources of red pepper using microsatellite markers. *Hortic Sci.* 36, 127–132.

651 Hill, T.A., Ashrafi, H., Reyes–Chin–Wo, S., Yao, J., Stoffel, K., et al., 2013. Characterization of
652 *Capsicum annuum* genetic diversity and population structure based on parallel polymorphism
653 discovery with a 30k unigene pepper gene chip. *Plos One* 8, e 56200.

654 Ibiza, V.P., Blanca, J., Cañizares, J., Nuez, F., 2012. Taxonomy and genetic diversity of
655 domesticated *Capsicum* species in the Andean region. *Genet. Resour. Crop Evol.* 59, 1077–1088.

656 Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for
657 dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*
658 23, 1801–1806.

659 Kraft, K.H., Brown, C.H., Nabhan, G.P., Luedeling, E., Luna Ruiz, J.J., et al., 2014. Multiple lines
660 of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico. *Proc. Natl.*
661 *Acad. Sci. USA* 111, 6165–6170.

662 Kumar, S., Kumar, R., Singh, J., 2006. Cayenne/American pepper (*Capsicum* species), in: Peter,
663 K.V., (Ed.) *Handbook of herbs and spices*. Woodhead, Cambridge, pp. 299–312.

664 Kwon, Y.L., Sham, J.M., Yi, G.B., Yi, S.I., Kim, K.M., Soh, E.H. Bae, K.M. Park, E.K., Song, I.H.,
665 Kim, B.D., 2005. Use of SSR markers to complement tests of distinctiveness, uniformity and
666 stability (DUS) of pepper (*Capsicum annuum* L.) varieties. *Mol. Cells* 19, 428–435

667 Lanteri, S., Acquadro, A., Quagliotti, L., Portis, E., 2003. RAPD and AFLP assessment of variation
668 in a landrace of pepper (*Capsicum annuum* L.) grown in North–West Italy. *Genet. Resour. Crop*
669 *Evol.* 50, 723–735.

670 Lee, J.M., Nahm, S.H., Kim, Y.M., Kim, B.D., 2004. Characterization and molecular genetic
671 mapping of microsatellite loci in pepper. *Theor. Appl. Genet.* 108, 619–627.

672 Lefebvre, V., Palloix, A., Rives, M., 1993. Nuclear RFLP between pepper cultivars (*Capsicum*
673 *annuum* L.). *Euphytica* 71, 189–199.

674 Liu, W.Y., Yang, H.B., Jo, Y.D., Jeong, H.J., Kang, B.C., 2013. Classical genetics and traditional
675 breeding in peppers, in: Kang, B.C., Kole, C. (Eds.), *Genetics, genomics and breeding of peppers*
676 *and eggplants*. CRC Press, London, pp. 16–39.

677 Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer*
678 *Res.* 27, 209–220.

679 Mazzucato, A., Papa, R., Bitocchi, R., Pietro, P., Nanni, L., Negri, V., Picarella, M.E., Siligato, F.,
680 Soressi, G.P., Tiranti, B., Veronesi, F., 2008. Genetic diversity, structure and marker–trait
681 associations in a collection of Italian tomato (*Solanum lycopersicum* L.) landraces. *Theor. Appl.*
682 *Genet.* 116, 657–669

683 Mercati, F., Longo, C., Poma, D., Araniti, F., Lupini, A., Mammano, M.M., Fiore, M.C.,
684 Abenavoli, M.R., Sunseri, F., 2014. Genetic variation of an Italian long shelf–life tomato (*Solanum*

685 *lycopersicum* L.) collection by using SSR and morphological fruit traits. Genet. Resour. Crop Evol.
686 DOI: 10.1007/s10722-014-0191-5.

687 Mimura, Y., Inoue, T., Minamiyama, Y., Kubo, N., 2012. An SSR based genetic map of pepper
688 (*Capsicum annuum* L.) serves as an anchor for the alignment of major pepper maps. Breeding Sci.
689 62, 93–98.

690 Minamiyama, Y., Tsuro, M., Hirai, M., 2006. An SSR-based linkage map of *Capsicum annuum*.
691 Mol. Breeding 18, 157–169.

692 Moses, M., Umaharan, P., Dayanandan, S., 2014. Microsatellite based analysis of the genetic
693 structure and diversity of *Capsicum chinense* in the Neotropics. Genet. Resour. Crop Evol. 61, 741–
694 755.

695 Nagy, I., Stägel, A., Sasvari, Z., Roder, M., Ganai, M., 2007. Development, characterization, and
696 transferability to other Solanaceae of microsatellite markers in pepper (*Capsicum annuum* L.).
697 Genome 50, 668–688.

698 Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of
699 individuals. Genetics 89, 583–590.

700 Nei, M., Li, W.H., 1979. Mathematical model for studying genetic variation in terms of restriction
701 endonucleases. Proc. Natl. Acad. Sci. USA 76, 5269–5273.

702 Nicolai, M., Cantet, M., Lefebvre, V., Sage-Palloix, A.M., Palloix, A., 2013. Genotyping a large
703 collection of pepper (*Capsicum* spp.) with SSR loci brings new evidence for the wild origin of
704 cultivated *C. annuum* and the structuring of genetic diversity by human selection of cultivar types.
705 Genet. Resour. Crop Evol. 60, 2375–2390.

706 Nimmakayala, P., Abburi, V.L., Abburi, L., Alaparathi, S.B., Cantrell, R., et al., 2014. Linkage
707 disequilibrium and population-structure analysis among *Capsicum annuum* L. cultivars for use in
708 association mapping. Mol. Genet. Genomics 289, 513–521.

709 Nuez, F., Gil, R., Costa, J., 1996. El cultivo de pimientos, chiles y ajies. Mundi-Prensa, Madrid.
710 607 p.

711 Onus, A.N., Pickersgill, B., 2004. Unilateral incompatibility in *Capsicum* (Solanaceae): occurrence
712 and taxonomic distribution. Ann. Bot. 94, 289–295.

713 Ortiz, R., Delgado de la Flora, F., Alvarado, G., Crossa, J., 2010. Classifying vegetable genetic
714 resources—A case study with domesticated *Capsicum* spp. Sci. Hort. 126, 186–191.

715 Oyama, K., Hernández-Verdugo, S., Sánchez, C., González-Rodríguez, A., Sánchez-Peña, P., et
716 al., 2006. Genetic structure of wild and domesticated populations of *Capsicum annuum*
717 (Solanaceae) from northwestern Mexico analyzed by RAPDs. Genet. Resour. Crop Evol. 53, 553–
718 562.

719 Pacheco–Olvera, A., Hernández–Verdugo, S., Rocha–Ramírez, V., González–Rodríguez, A.,
720 Oyama, K., 2012. Genetic diversity and structure of pepper (*Capsicum annuum* L.) from
721 Northwestern Mexico analyzed by microsatellite markers. *Crop Sci.* 52, 231–241

722 Papadakis, J., 1975. *Climates of the world and their agricultural potentialities*. Edición Argentina,
723 Buenos Aires, 200 pp

724 Paran, I., Fallik, E., 2011. Breeding for fruit quality in pepper (*Capsicum* spp.), in: Bebeli, P., Jenks,
725 M. (Eds.), *Breeding for fruit quality*. Wiley–Blackwell Publishing, San Francisco (USA), pp. 307–
726 318.

727 Paran, I., van der Knaap, E., 2007. Genetic and molecular regulation of fruit and plant
728 domestication traits in tomato and pepper. *J. Exp. Bot.* 58, 3841–3852.

729 Paran, I., Aftergoot, E., Shifriss, C., 1998. Variation in *Capsicum annuum* revealed by RAPD and
730 AFLP markers. *Euphytica* 99, 167–173.

731 Park, S.D.E., 2001. The Excel microsatellite toolkit. Available:
732 <http://animalgenomics.ucd.ie/sdepark/mstoolkit/>. Accessed 2014 December.

733 Patil, B.S., Crosby, K., Byrne, D., Hirschi, K., 2014. The intersection of plant breeding, human
734 health, and nutritional security: lessons learned and future perspectives. *HortScience* 49, 116–127

735 Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic
736 software for teaching and research—an update. *Bioinformatics* 28, 2537–2539.

737 Pereira–Lorenzo, S., Costa, R.M.L., Ramos–Cabrer, A.M., Ciordia–Ara, M., Ribeiro, C.A.M.,
738 Borges, O., Barreneche, T., 2011. Chestnut cultivar diversification process in the Iberian Peninsula,
739 Canary Islands, and Azores. *Genome* 54, 301–315

740 Perry, L., Dickau, R., Zarrillo, S., Holst, I., Pearsall, D.M., et al., 2007. Starch fossils and the
741 domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. *Science* 315, 986–
742 988.

743 PGRI, AVRDC, CATIE, 1995. *Descriptors for Capsicum (Capsicum spp.)*. International Plant
744 Genetic Resources Institute, Rome, Italy. 49 p.

745 Portis, E., Nervo, G., Cavallanti, F., Barehi, L., Lanteri, S., 2006. Multivariate analysis of genetic
746 relationships between Italian pepper landraces. *Crop Sci.* 46, 2517–2525.

747 Pritchard, J.K., Stephens, M., Rosenberg, N.A., Donnelly, P., 2000. Association mapping in
748 structured populations. *Am. J. Hum. Genet.* 67, 170–181.

749 Pressoir, G., Berthaud, J., 2004. Patterns of population structure in maize landraces from the Central
750 Valleys of Oaxaca in Mexico. *Heredity* 92, 88–94.

751 Qin, C., Changshui, Y., Shen, Y., Fang, X., Chen, L., et al., 2014. Whole genome sequencing of
752 cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. *Proc.*
753 *Natl. Acad. Sci. USA* 111, 5135–5140.

754 R Development Core Team, 2014. R: a language and environment for statistical computing. R
755 Foundation for Statistical Computing, Vienna, Austria. Website <http://www.r-project.org/>

756 Rai, V.P., Kumar, R., Kumar, S., Rai, A., Kumar, S., et al., 2013. Genetic diversity in *Capsicum*
757 germplasm based on microsatellite and random amplified microsatellite polymorphism markers.
758 *Physiol. Mol. Biol. Plants* 19, 575–586.

759 Rodríguez–Bao, J.M., Terren–Poves, L., Rivera–Martínez, A., Andres–Ares, J.L., Fernandez Paz,
760 J., 2004. Pimientos autóctonos de Galicia. *Horticultura Internacional* 43, 34–40.

761 Rodríguez–Maza, M.J., Garcés–Claver, A., Park, S.W., Kang, B.C., Arnedo–Andrés, M.S., 2012. A
762 versatile PCR marker for pungency in *Capsicum* spp. *Mol. Breeding* 30, 889–898.

763 Rohlf, F.J., 2000. NTSYS–pc Numerical taxonomy and multivariate analysis system, version 2.1.
764 Exeter Software, Setauket.

765 Schuelke, M., 2000. An economic method for the fluorescent labeling PCR fragments. *Nature*
766 *Biotech.* 18, 233–234.

767 Singh, N., Choudhury, D.R., Singh, A.K., Kumar, S., Srinivasan, K., Tyagi, R.K., Singh, N.K.,
768 Singh, R., 2013. Comparison of SSR and SNP markers in estimation of genetic diversity and
769 population structure of indian rice varieties. *Plos One* 8, e84136.

770 Sparato, G., Negri, V., 2013. The European seed legislation on conservation varieties: Focus,
771 implementation, present and future impact n landrace on farm conservation. *Genet. Resour. Crop*
772 *Evol.* 60, 2421–2430.

773 SPSS, 2008. SPSS Statistics for Windows, Version 17.0. Chicago: SPSS Inc.

774 Sugita, T., Semi, Y., Sawada, H., Utoyama, Y., Hosomi, Y., et al., 2013. Development of simple
775 sequence repeat markers and construction of a high–density linkage map of *Capsicum annuum*.
776 *Mol. Breeding* 31, 909–920.

777 Tam, S.M., Lefebvre, V., Palloix, A., Sage–Palloix, A.M., Grand–bastien, M.A., 2009. LTR–
778 retrotransposons Tnt1 and T135 markers reveal genetic diversity and evolutionary rela– tionships of
779 domesticated peppers. *Theor. Appl. Genet.* 119, 973–989.

780 Tang, H., Sezen, U., Paterson, A.H., 2010. Domestication and plant genomes. *Curr. Opinion Plant*
781 *Biol.* 13, 160–166

782 Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular
783 Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* 30, 2725–2729.

784 Toquica, S.P., Rodriguez, F., Martinez, E., Duque, M.C., Tohme, J., 2003. Molecular
785 characterization by AFLPs of *Capsicum* germplasm from the Amazon department in Colombia.
786 *Genet Resour. Crop Evol.* 5, :639–647

787 Votava, E.J., Baral, J.B., Bosland, P.W., 2005. Genetic diversity of chile (*Capsicum annuum* var.
788 *annuum* L.) landraces from northern New Mexico, Colorado, and Mexico. *Econ. Bot.* 59: 8–17.

Wagner, H.W., Sefc, K.M., 1999. IDENTITY 1.0. Center for Applied Genetics, University of Agricultural Sciences Vienna.

Yi, G., Lee, J.M., Lee, S., Choi, D., Kim, B.D., 2006. Exploitation of pepper EST–SSRs and an SSR–based linkage map. *Theor. Appl. Genet.* 114, 113–130.

Yumnam, J.S., Tyagi, W., Pandey, A., Meetei, N.T., Rai, M., 2012. Evaluation of genetic diversity of chilli landraces from North Eastern India based on morphology, SSR markers and the Pun1 locus. *Plant Mol. Biol. Rep.* 30, 1470–1479.

Zeinalabedini, M., Sohrabi, S., Nikoumanesh, K., Imani, A., Mardi, M., 2012. Phenotypic and molecular variability and genetic structure of Iranian almond cultivars. *Plant Syst. Evol.* 298, 1917–1929.

Zeven, A.C., 1998. Landraces, a review of definitions and classifications. *Euphytica* 104, 117–139.

Wu, F., Eannetta, N.T., Xu, Y., Durrett, R., Mazourek, M., et al., 2009. A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus *Capsicum*. *Theor. Appl. Genet.* 118, 1279–1293.

803

804

805

806

807

808

809

810

811

812

813

814

815

Table 1 Common name and origin of the landraces from NW Spain used in the present work

Landrace	Lines*	Origin	
		Region	Province

Arnoia	AR330-05, AR55, AR65, AR45, AR274-05	Arnoia	Orense
Blanco Rosal	BR97-04, BR113-05	O Rosal	Pontevedra
Couto	Co5A, Co10A, Co12	Narón	A Coruña
Couto Grande	CG-10-117	Narón	A Coruña
Mougan	MG279-05, MG77-04, MG101-04, MG335-05, MG14-04, MG318-05	Guntin	Lugo
Oimbra	OI27-05, OI30-04	Oimbra	Orense
Padron	ETH80-05, ETH134-05, ETH209-05, PA145, ETH211-05, ETH166-05	Herbón	A Coruña
Piñeira	PI309-05, PI275-05	Ribadeo	Lugo
Punxin	PX74-05, PX118-04	Punxin	Orense

* All lines, except AR45, AR55 and AR65, are landraces-derived inbred lines

831 **Table 2** Microsatellite markers evaluated for the genotyping of the *C. annuum* collection. Name, chromosomal position, repeat sequence and
832 source are indicated. For each locus: number of alleles observed (Na), range of alleles sizes, observed heterozygosity (H_o), Nei's unbiased gene
833 diversity (uH_e), Wright's fixation index (F_{is}) and polymorphic information content (PIC)

Marker	Chrom	Repeat	Source	Na	Range (bp)	H_o	uH_e	F_{is}	PIC
Hpms1-148	1	(GA) ₁₄	Lee et al. (2004)	2	211-213	0.000	0.508	1.000	0.375
Hpms1-139	1	(CT) ₂ (AG) ₁₅	Lee et al. (2004)	3	304-320	0.069	0.550	0.872	0.440
Hpms1-143	2	(AG) ₁₂	Lee et al. (2004)	4	240-246	0.000	0.498	1.000	0.456
HpmsE016	3	(TACA) ₃	Yi et al. (2006)	4	180-210	0.000	0.709	1.000	0.641
Hpms1-173	3	(GA) ₁₆ (TG) ₂	Lee et al. (2004)	4	176-186	0.000	0.745	1.000	0.684
CAMS351	4	(TG) ₃ ...(AG) ₂₆	Minamiyama et al. (2006)	3	223-227	0.000	0.547	1.000	0.476
CA523558	6	T ₂₀	Ben-Chaim et al. (2006)	2	194-196	0.000	0.373	1.000	0.299
Hpms1-5	6	(AT) ₁₁ (GT) ₁₇	Lee et al. (2004)	3	334-338	0.000	0.653	1.000	0.563
EPMS426	7	(AT) ₁₅	Nagy et al. (2007)	3	118-130	0.000	0.547	1.000	0.476
HpmsCaSIG19	7	(CT) ₆ (AT) ₈	Lee et al. (2004)	3	237-241	0.037	0.662	0.943	0.575
HpmsE145	8	(TTA) ₄	Yi et al.(2006)	2	233-245	0.000	0.382	1.000	0.305
EPMS342	8	(CTT) ₇	Yi et al. (2006)	2	349-360	0.069	0.460	0.847	0.350
Hpms2-24	9	(CT) ₁₇ (CA) ₅ A ₂₁	Lee et al. (2004)	2	204-220	0.000	0.373	1.000	0.299
HpmsE082	9	(ATC) ₅	Yi et al. (2006)	2	250-254	0.000	0.195	1.000	0.173
CAMS336	10	(TC) ₁₆	Minamiyama et al. (2006)	3	178-192	0.000	0.668	1.000	0.582
HpmsE031	10	(TC) ₆	Yi et al. (2006)	3	185-189	0.000	0.249	1.000	0.227
Hpms2-2	11	(GT) ₉	Lee et al. (2004)	2	167-169	0.000	0.334	1.000	0.274
CAMS405	11	(TC) ₁₈	Minamiyama et al.(2006)	4	250-256	0.000	0.587	1.000	0.523
HpmsE075	12	(ACC) ₆	Yi et al. (2006)	2	218-221	0.034	0.354	0.901	0.287
GPMS197	12	(GA) ₃ (TAT) ₁₆	Nagy et al. (2007)	7	293-351	0.103	0.849	0.876	0.814

834

Table 3 Analysis of variance (sum of squares) for test of significance of differences among landraces, locations and their interaction for thirteen quantitative traits in a collection of twenty-nine Galician lines

Source of variation	df	Plant width		Stem length		Stem diameter		Plant height		Fruit lenght		Fruit width		Fruit weight	
Landrace (L)	8	174043.95	***	90959.47	***	78.60	***	854375.17	***	14384.94	***	5666.77	***	7.70E+06	***
Location (Lo)	1	54212.37	***	450.07	***	127.34	***	213559.21	***	85.89	***	45.39	***	5.22E+04	***
L×Lo	8	22096.56	***	1172.62	***	17.47	***	81402.63	***	354.20	***	22.02	***	5.33E+04	***
Error	2807	418739.13		54655.31		252.48		1120461.95		6992.66		1655.45		2.62E+06	
Source of variation	df	Fruit pedicel length		Fruit wall thickness		N° of locules		Placenta lenght		N° fruits/plant		Kg/m ²			
Landrace (L)	8	593.33	***	3938.91	***	262.41	***	18712.88	***	34.54	***	1.14E+06	***	-	
Location (Lo)	1	652.42	***	82.65	***	16.32	***	170.93	**	15.40	***	1.76E+03	ns	-	
L×Lo	8	60.19	***	53.79	***	24.48	***	1009.33	***	45.95	***	8.09E+03	ns	-	
Error	2807	1787.46		2281.45		1128.51		86345.38		587.51		2.16E+05		-	

*, **, *** Significant at $P=0.05$, 0.01 and 0.001, respectively.

841

842

843

844

845

846

847

848

849

850 **Table 4** Diversity parameters calculated for the different groups of landraces. N=number of
851 genotypes, H_o =observed heterozygosity, uH_e =Nei's unbiased gene diversity (uH_e), F_{is} =Wright's
852 fixation index. Numbers in brackets in the last column indicate the frequencies for each private
853 allele.

Landrace	N	H_o	uH_e	F_{is}	N° private alleles
Arnoia	5	0.000	0.21	1.00	1 (60%)
Blanco Rosal	2	0.000	0.00	na	1 (100%)
Couto	3	0.033	0.05	0.20	4 (100%), 1 (50%)
Couto Grande	1	na	na	na	1 (na)
Padron	6	0.000	0.00	na	1 (na)
Mougan	6	0.042	0.16	0.72	1 (83%), 1 (75%)
Oimbra	2	0.025	0.09	0.56	-
Piñeira	2	0.025	0.13	0.67	-
Puxin	2	0.000	0.00	na	1 (na)
Total	29	0.016	0.51	0.97	12

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869
870
871
872

873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888

Table 5 AMOVA partition of molecular variance for two and six groups suggested by the UPGMA derived dendrogram and the STRUCTURE analysis

Source of variations	<i>df</i>	Sum of squares	Variance components	% Total	<i>p</i> value
<i>3 groups</i>					
Among groups	2	63.37	1.56	38.35	0.0001
Within groups	55	138.52	2.52	61.65	-
Total	57	201.89	4.08	100	-
<i>6 groups</i>					
Among groups	5	141.58	2.9	71.46	0.0001
Within groups	52	60.32	1.16	28.54	-
Total	57	201.89	4.06	100	-

889 **Figure captions**

890

891 **Fig. 1.** Scatter plot of the PCA analysis based on thirteen quantitative characters. A) PC1 vs PC2,
892 B) PC3 vs PC2

893

894 **Fig. 2.** Scatter plot of the PCA analysis based on eight qualitative traits. A) PC1 vs PC2, B) PC3 vs
895 PC2

896

897 **Fig. 3.** A) Dendrogram of twenty-nine landraces-derived inbred lines revealed by UPGMA cluster
898 analysis based on Dice distance derived from twenty SSRs. Numbers in brackets indicate bootstrap
899 values expressed as percentages. B) Structure bar plots at $K=6$ and $K=8$ (C). The twenty-nine
900 genotypes are ordered according to the UPGMA tree. Each colour represents a different group and
901 the proportion of a given genotype's colour represents the proportion that genotype belongs to the
902 corresponding group.

903

904

905

906

907

908

909

910

911

912

913

914 **Supplementary material**

915

916 **Table S1** Analysis of variance (sum of squares) for test of significance of differences among
917 landraces, years and their interaction for thirteen quantitative traits in a collection of twenty-nine
918 Galician lines assayed at two different locations (Mabegondo and Salceda)

919

920 **Table S2** Means $\pm$ SE for thirteen quantitative traits for nine landraces. Values with different letter
921 within the same column indicate significant differences at $P=0.001$

922

923 **Table S3** Means $\pm$ SE for thirteen quantitative traits for nine landraces grouped according to the
924 two experimental locations

925

926 **Table S4** Estimates of variance, accumulated variance and weighting coefficients (autovectors) of
927 the first three principal components for thirteen quantitative characters (A) and eight qualitative
928 traits (B) evaluated on twenty-nine pepper lines

929

930 **Fig. S1.** Map of the region of Galicia (NW Spain) showing the geographical origin of the nine
931 landraces used

932

933 **Fig. S2.** Distribution of the qualitative morphological characters throughout the groups of landraces
934 according to the location (Mabegondo or Salceda). Each chart corresponds to a different character
935 and each bar to a different landrace. The percentages are calculated over the number of plants found
936 to have every type of the character.

937

938 **Fig. S3.** UPGMA tree based on Nei's genetic distances obtained from SSR markers showing
939 relationships among the pepper landraces

940

941

942

943

944

945