

# Genetic diversity and core collection of *Malus × domestica* in northwestern Spain, Portugal and the Canary Islands by SSRs

Santiago Pereira-Lorenzo<sup>a,\*</sup>, Ana María Ramos-Cabrer<sup>a</sup>, Vanessa Ferreira<sup>b</sup>,  
María Belén Díaz-Hernández<sup>a</sup>, Valdemar Carnide<sup>b</sup>, Olinda Pinto-Carnide<sup>b</sup>, Raúl Rodrigues<sup>c</sup>,  
María Encarnación Velázquez-Barrera<sup>d</sup>, Domingo Rios-Mesa<sup>d</sup>, Javier Ascasíbar-Errasti<sup>e</sup>,  
Isaura Castro<sup>b</sup>

<sup>a</sup> Universidad de Santiago de Compostela, Escola Politécnica Superior, Departamento de Producción Vegetal e Ingeniería de Proyectos, Campus de Lugo, 27002 Lugo, Spain

<sup>b</sup> Centro de Investigação e de Tecnologias Agro-Ambientais e Biológicas (CITAB), Universidade de Trás-os-Montes e Alto Douro, 5000-801 Vila Real, Portugal

<sup>c</sup> Centro de Investigação de Montanha (CIMO), Escola Superior Agrária de Ponte de Lima, Instituto Politécnico de Viana do Castelo (ESAPL/IPVC), 4990-706, Refóios, Portugal

<sup>d</sup> Cabildo Insular de Tenerife, Área de Agricultura, Pza España 1, 38001 Santa Cruz de Tenerife, Canary Islands, Spain

<sup>e</sup> Centro de Investigacións Agrarias de Mabegondo, Xunta de Galicia, Apdo. 10, 15080 A Coruña, Spain

## ARTICLE INFO

### Keywords:

Apple  
Local cultivars  
Molecular markers  
Structure  
Variability  
Germplasm banks

## ABSTRACT

The domesticated apple (*Malus × domestica* Borkh.) is one of the most relevant cultivated fruit crops worldwide, in particular in temperate zones, and the fourth most economically important. In order to avoid loss of genetic diversity, germplasm banks have been established to conserve proper genetic diversity, among them we have evaluated five that covers western Iberian Peninsula (the one from Galicia and three from northern to southern Portugal) and Canary Islands. A total of 634 accessions composed of old and local autochthonous apple germplasm, as well as commercial and reference varieties, plus 16 accessions that were collected as wild apples, were studied. Of the total, 484 local accessions were obtained from Spain and 67 from Portugal. The objective of the present study was to analyse the genetic diversity and genetic structure, and define core collection (a small group of cultivars that preserve the allele diversity of the set), by using 13 microsatellite markers (SSRs).

Of the total of 634 accessions studied, corresponding to 370 unique genotypes, were grouped into two reconstructed panmictic populations or RPPs; the first RPP included 506 accessions (304 unique genotypes), which belonged to local and autochthonous varieties; and the second RPP clustered the remaining 128 (66 distinct genotypes), which were commercial and reference varieties. In total, 29 new synonyms (74 accessions) were identified between Galicia and the Canary Islands, which indicated the close historical relationship between these two regions located 2000 km apart.

A core collection to preserve the whole allelic diversity of the west of the Iberian Peninsula (Galicia and Portugal) and the Canary Islands was defined keeping 10% of the genotypes (63) with respect to total accessions of the complete study, 14% of them (54) from Galicia with wild accessions or 15% (55) from Galicia without wild accessions, 40% (30) from the Canary Islands and 36% (24) from Portugal.

The results obtained will allow update of the germplasm banks with correct identification of cultivars, differentiating the autochthonous ones from those derived from commercial cultivars, with the elimination of repetitions (42%), and proposal of the core collection necessary to preserve allele diversity.

## 1. Introduction

Relatively recently (more than 50 million years ago), genome-wide duplication (GWD) gave rise to the 17 chromosomes in the Pyrae tribe

(Velasco et al., 2010), confirming the origin of the cultivated apple in *Malus sieversii*, closely related to *M. orientalis* and *M. × asiatica* (a form of Chinese cultivated apple), and less closely related to *M. sylvestris*, *M. baccata*, *M. micromalus* and *M. prunifolia*. The wild apple *M. sieversii*

\* Corresponding author.

E-mail addresses: [santiago.pereira.lorenzo@usc.es](mailto:santiago.pereira.lorenzo@usc.es) (S. Pereira-Lorenzo), [ana.amos@usc.es](mailto:ana.amos@usc.es) (A.M. Ramos-Cabrer), [vanessa.cm.ferreira29@gmail.com](mailto:vanessa.cm.ferreira29@gmail.com) (V. Ferreira), [belen.diaz@usc.es](mailto:belen.diaz@usc.es) (M.B. Díaz-Hernández), [vcarnide@utad.pt](mailto:vcarnide@utad.pt) (V. Carnide), [olindac@utad.pt](mailto:olindac@utad.pt) (O. Pinto-Carnide), [raulrodrigues@esa.ipvvc.pt](mailto:raulrodrigues@esa.ipvvc.pt) (R. Rodrigues), [mariavb@tenerife.es](mailto:mariavb@tenerife.es) (M.E. Velázquez-Barrera), [domingor@tenerife.es](mailto:domingor@tenerife.es) (D. Rios-Mesa), [javier.ascasibar.errasti@xunta.gal](mailto:javier.ascasibar.errasti@xunta.gal) (J. Ascasíbar-Errasti), [icastro@utad.pt](mailto:icastro@utad.pt) (I. Castro).

<https://doi.org/10.1016/j.scienta.2018.05.053>

Received 20 March 2018; Received in revised form 8 May 2018; Accepted 31 May 2018  
0304-4238/ © 2018 Elsevier B.V. All rights reserved.

(Ldb.) Roem. was found in the Tian Shan mountains of Central Asia, and it is believed that its origin dates from about 4000–10,000 years ago. The apples were taken to the west by the Silk Road, where they hybridized with other wild apples such as *M. baccata* (L.) Borkh. in Siberia, *M. orientalis* Uglitz. in the Caucasus, and *M. sylvestris* Mill. in Europe (Cornille et al., 2014). The process of domestication retained most of the genetic variation compared to its close wild relative, *Malus sieversii* M. Roem. (Gross et al., 2014, 2018; Lassois et al., 2016; Urrestarazu et al., 2016), with wild-to-crop introgressants, particularly from the European crabapple *M. sylvestris* (Cornille et al., 2012; Leforestier et al., 2015).

In the apple tree, grafting produced ‘instant domestication’ as in other perennial crops, from spontaneous hybridizations, but without significant genetic introgressants from other species such as *M. sylvestris* as had been believed (Velasco et al., 2010). The specific fruit of the Pyreae, the pome, is formed by the growth of the lower part of the sepals, which is related to the MADS-box genes, involved in the development of flowers and fruits and in the metabolism of sorbitol (Velasco et al., 2010).

So far in Spain, a total of 1453 accessions from seven collections have been referenced, 174 of which correspond to foreign cultivars (93 cultivars) (Pereira Lorenzo et al., 2017). The accessions are conserved in different germplasm banks: 321 in the Mabegondo Centre for Agrarian Research (CIAM) in A Coruña, Galicia; 122 in the Centre for Conservation of Agricultural Biodiversity (CCBAT) in Cabildos de Tenerife, La Palma and Gran Canaria, Canary Islands; 309 curated at the Public University of Navarre (UPNa) in Pamplona, Navarra; 247 in the Asturias Regional Agrifood Research and Development Service (SERIDA) in Villaviciosa, Asturias; 207 in the Aragón Agrifood Research and Technology Centre (CITA) in Zaragoza, Aragón; 143 in the Spanish National Research Council Aula Dei Experimental Station (EEAD-CSIC) in Zaragoza, Aragón; and 104 curated at the University of Lleida (UdL) in Gimenells, Catalonia.

In Portugal, regional apple genetic resources are mainly conserved in the collection of the Regional Directorate of Agriculture and Fisheries of the North (DRAPN) in Felgueiras (Quinta de Sergude) (39 accessions), the Viseu Agrarian Station of the Regional Directorate of Agriculture and Fisheries of the Centre (161 accessions), the Tavira Agricultural Experiment Centre of the Regional Directorate of Agriculture and Fisheries of the Algarve (DRAPA; 26 accessions) and recently also in the germplasm collection of the Agrarian School of the Viana do Castelo Polytechnic Institute which contains more than 100 accessions.

On this way, the general objective of this study was to evaluate the genetic diversity and relationships of local apple accessions from western Iberian Peninsula, including Galicia and Portugal, and Canary Islands by SSR markers proposed by the ECPGR for apple, in order (1) to determine duplicates, synonyms, and homonym for a better management of apple collections; (2) to better understand the genetic diversity and structure in the continuous area of western Iberian Peninsula and Canary Islands, with historical relationships with Galicia and Portugal.

## 2. Material and methods

### 2.1. Plant material

We studied a total of 634 apple accessions, comprising 484 and 67 local and native from Spain and Portugal, respectively, and 16 accessions collected as alleged wild apple trees (Table A.1 in Supplementary material). Moreover, were also analysed 66 reference varieties (commercial cultivars and foreign genotypes) including six references recommended by the ECPGR as standards between laboratories, supplied by the Research Institute of Horticulture and Seeds, INRA, Agrocampus Ouest, University of Angers, France (FR), including other *Malus* species. In the case of multiple accessions with the same denomination, sampled

in one or more germplasm collections, different codes were granted.

The plant material (native and local apple cultivars) came from germplasm banks located in Portugal (PO) and Spain (SP). In Spain, the plant material came from two germplasm collections, one at CIAM of the Xunta de Galicia in Galicia (GAL) and the other from the Canary Island Councils of Tenerife, La Palma and Gran Canaria in the Canary Islands (CAN). The wild material (WILD) was collected in the field in the province of Lugo. The Portuguese plant material was sampled in three germplasm collections, those of the Agrarian School of the Viana do Castelo Polytechnic Institute in Ponte de Lima (ESAPL), the ‘Quinta de Sergude’ from the Regional Directorate of Agriculture and Fisheries of the North’ (DRAPN) and the Tavira Agricultural Experiment Centre of the Regional Directorate of Agriculture and Fisheries of the Algarve (DRAPA), located in the Portuguese provinces of Minho, Douro Litoral and Algarve, respectively.

### 2.2. SSR analysis

DNA extraction, amplification and data analysis were carried out according to Ferreira et al. (2016). Each sample was amplified using a set of 13 microsatellites proposed by the ECPGR for apple with a high degree of polymorphism determined in previous studies, belonging to different linkage groups to ensure their independence: CH04e05, CH01h10, CH02d08, CH01f02, CH02c11, CH04c07, CH02c09, CH01h01, CH05f06 and CH03d07 (Liebhard et al., 2002); GD12 and GD147 (Hokanson et al., 1998); and Hi02c07 (Silfverberg-Dilworth et al., 2006) (Table A.2 in Supplementary material). Among this set of markers, eight (CH04e05, CH02c11, CH02c09, CH02d08, CH04c07, CH01h01, Hi02c07 and CH01h10) have been considered as the most relevant for the evaluation of European apple genetic resources (Lateur et al., 2013) (Table A.2 in Supplementary material). The forward primers were labelled with 6-FAM, VIC, NED or PET fluorochrome, and three different multiplex PCRs were designed.

### 2.3. Data analysis

Only unique genotypes were used for the analyses. SPAGeDi ver. 1.3 software (Hardy and Vekemans, 2002) was used to estimate the observed heterozygosity ( $H_o$ ) and expected heterozygosity ( $H_e$ ) (Nei, 1978). GenoDive ver. 2.0b23 software (Meirmans and Van Tienderen, 2004) was used to estimate Molecular variance (AMOVA) (Excoffier et al., 1992; Michalakis and Excoffier, 1996), the inbreeding coefficient ( $F_{IS}$ ) and the goodness-of-fit test ( $F_{IT}$ ) (Weir and Cockerham, 1984). The absence of a defined allele was codified as ‘0’ and the presence as ‘1’ to build up a binary data matrix, to compute the Jaccard coefficient ( $J_C$ ) which does not consider the shared absence of a character as a similarity (Lo et al., 2009).

We also performed a Bayesian analysis implemented by STRUCTURE software (Pritchard et al., 2000) by using the admixture model with unlinked loci and correlated allele frequencies, as defined by Pereira-Lorenzo et al. (2010) and Porras-Hurtado et al. (2013) who recommended a minimum of 20 iterations to estimate the ancestry membership proportions of a population. Diploid and triploid genotypes were analysed separately. We computed  $K = 1$ –15 unknown RPPs (reconstructed panmictic populations) of genotypes with the options popinfo = 0 and popflag = 0, which considered that the sampled genotypes were of unidentified origin, assigning them probabilistically to RPPs based on a  $q_i$  (probability of membership) of 80%, while a lower probability meant an admixed genotype. The best  $K$  value supported by the data (Evanno et al., 2005) was obtained by using Structure Harvester (Earl and vonHoldt, 2012).

RPPs identified by the Bayesian model-based clustering method were characterized by descriptive statistics, such as the number of genotypes retained in each RPP, the number of genotypes strongly assigned to each RPP (i.e.  $q_i > 80\%$ ), the average  $H_e$ , the number of total alleles retained per RPP, the number of rare alleles ( $P < 0.05$ ) and the

number of unique alleles (i.e. those only found in one division level). A factorial correspondence analysis (FCA) was performed considering the RPPs, using Genetix 4.05 (Belkhir et al., 2004).

A core collection, i.e. the minimum number of samples which represents the greatest diversity of the initial collection, was defined using PowerCore (ver. 1.0) (Kim et al., 2007) with the heuristic method for the total number of genotypes, and for each set of samples: Portugal, Galicia and the Canary Islands.

### 3. Results

#### 3.1. SSR polymorphism

The total 634 accessions studied were grouped into 370 genotypes (Table A.3 in Supplementary material). Of these 634 total accessions, 506 (grouped into 304 genotypes) belonged to local and autochthonous varieties and the remaining 128 accessions were reference varieties (grouped into 66 genotypes). Excluding references, 41 (11.1%) unique genotypes were from Portugal and 264 (71.3%) from Spain, which were mostly from Galicia (209 genotypes, 56.5%) and less from Canary Islands (55 genotypes, 14.9%). In Portugal, unique genotypes were 17 (4.6%) from DRAPN, 11 (3.0%) from ESAPL, and 13 (3.5%) from DRAPA. Among the 559 accessions studied from Spain (including local, wild and reference accessions), 244 presented unique genotypes, of which 193 were diploids and 51 putative triploids (they presented three alleles for some loci). Among the 316 remaining accessions were found 78 genotypes, 48 diploids and 30 putative triploids.

Of the 67 local Portuguese accessions, 34 unique genotypes were identified, 29 diploids and five putative triploids. The remaining 33 accessions were grouped into 14 genotypes, 11 diploids and three putative triploids.

In the case of Portugal, of the 92 accessions previously described by Ferreira et al. (2016), 16 (12 genotypes) are present in the Galician collection with other names (synonyms). Twenty-nine synonyms were found between the Galician and Canary Islands collections, representing 74 accessions, 43 from Galicia, 28 from the Canary Islands and three references (Table 1). In this work, no additional synonymy was found between Portugal and the Canary Islands.

The 13 SSR loci amplified a total of 230 alleles, ranging from a minimum of 12 (Hi02c07) to a maximum of 24 different alleles (CH03d07), with an average number of alleles per locus of 17.7 (Table 2). Of the total of 230 alleles, 142 were rare alleles ( $P < 0.05$ ).

The average  $H_e$  of the 13 loci for the 370 genotypes (diploids and triploids) was 0.815, varying between 0.530 (Hi02c07) and 0.890 (CH01f02).

On the other hand, the maximum value of  $F_{IS}$  was 0.126 ( $P < 0.01$ )

**Table 2**

Genetic diversity for diploid and triploid genotypes ( $n = 370$ ) and for only diploid genotypes ( $n = 281$ ). Expected Heterozygosity ( $H_e$ ), Observed Heterozygosity ( $H_o$ ) and inbreeding coefficient ( $F_{IS}$ ).

Locus	Diploid and triploid genotypes				Diploid genotypes	
	(n = 370)				(n = 281)	
	No. Alleles	No. Rare alleles	$H_e$	$F_{IS}$	$H_o$	$H_e$
CH02c09	14	6	0.856	−0.026	0.882	0.857
CH02d08	19	13	0.836	−0.063**	0.849	0.840
CH05f06	13	5	0.855	0.024	0.811	0.852
GD147	18	12	0.849	−0.025	0.823	0.849
CH04e05	21	16	0.726	0.017	0.685	0.716
CH02c11	14	5	0.879	−0.016	0.883	0.873
Hi02c07	12	7	0.530	0.126**	0.392	0.475
CH01h10	15	8	0.780	−0.013	0.801	0.789
CH03d07	24	17	0.876	−0.004	0.853	0.874
CH04c07	22	15	0.864	0.001	0.840	0.859
CH01f02	23	15	0.890	−0.020	0.878	0.893
CH01h01	19	13	0.847	0.043**	0.779	0.842
GD12	16	10	0.802	0.057***	0.759	0.798
Total	230	142				
Average	17.7	11	0.815	0.008	0.731	0.752

\*\*  $P < 0.01$ .

\*\*\*  $P < 0.05$ .

for Hi02c07, and the minimum was −0.063 ( $P < 0.01$ ) for CH02d08. The mean value of  $F_{IS}$  was 0.008, with a slight deficit of heterozygotes, which is also evident for the 281 diploid genotypes, with  $H_o = 0.731$  and  $H_e = 0.752$  (Table 2).

#### 3.2. Structure and origin of the cultivars

Bayesian analysis (STRUCTURE) was performed for diploid (281) and triploid (89) unique genotypes independently. For diploids,  $\ln [Pr(X/K)]$  reached a maximum value at  $K = 2$  (Fig. A.1 in Supplementary material) (no substructure was detected) estimated with STRUCTURE HARVESTER (Earl and vonHoldt, 2012), which corresponds to strong differentiation of two main groups of genotypes with a  $qI > 80\%$  (including 202 genotypes of the total 281, 71.8%). The first group (RPP1) included 119 local genotypes and 130 accessions (56.25%), and the second group (RPP2) had 83 genotypes and 100 accessions (43.48%) including commercial and local cultivars, all with  $qI > 80\%$ . The third group presented the remaining 79 genotypes with a  $qI < 80\%$ , considered 'admixed' (Figs. 1 and A.1 in Supplementary material, Tables 3 and A.3 in Supplementary material). RPP1 clustered 16 genotypes from

**Table 1**

Synonyms between Canary Islands and Galicia.

Accessions from Canary Islands	Accessions from Galicia	Commercial cultivars
'Cencia'	'Alfoz_01', 'De Septiembre'	
'Manzano'	'Magazos_02', 'Manzana De Invierno', 'Pontellas'	
'Acida', 'Antiguo' ('Antigua'), 'Francés' ('Francesa'), 'Francesa Antigua', 'Llagada', 'Raneta', 'Reineta', 'Reineta Parda',	'Mingán Almuña', 'Tabardilla Parda', 'Chantada_02', 'Reina De Reinetas', 'Tabardilla Antigua'	'Reineta Gris'
'Similar A Reineta'		
'Rayada'	'Breixa_02', 'Facha Fonte', 'Parecida A Carapanón', 'Senra_01', 'Tabardilla Francesa', 'Tabardilla Romana'	
'Del País', 'Garafia_01', 'Tenerife_01'	'Burricios_04'	'Red Spur', 'Top Red'
'Francés', 'Manzano', 'Pajarita'	'Herbón_07', 'De Agosto', 'De Cedo', 'De Puebla', 'Do Apóstol', 'Grande Plana', 'Nogueirosa_01', 'Roja Plana', 'Torres Agrelo_03'	
'Roja Rayada'	'Masma_01'	
'Manzano'	'Partovia_04', 'Tabardilla Roja', 'Tabardilla Tardía'	
'Antiguo', 'Rojo Temprano'	'De Julio', 'De San Juan', 'Roja De Julio', 'Roja De Verano'	
'Manzano'	'Pero'	
'Manzano'	'Blanca', 'Blanca Plana', 'De Pera', 'De San Roque', 'Maza Fresca', 'Parecida A De Cera'	
'Francés', 'Jugosa', 'Perotolo'	'E2', 'Pequeña Verde'	



**Fig. 1.** Analysis of structure for 281 diploid genotypes and 89 triploids of apple in the western Iberian Peninsula (Galicia and Portugal) and Canary Islands analysed by STRUCTURE (Pritchard et al., 2000) using 13 loci SSR in  $K = 2$  reconstructed populations, RPP1 (in light grey) and RPP2 (in dark green).

$qI$ , coefficient of coancestry (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Canary Islands, 71 from Galicia, nine from Portugal and 23 were mostly Spanish cider cultivars used as references. RPP2 clustered 11 genotypes from Canary Islands, 35 from Galicia, 11 from Portugal and 26 references.

For triploid individuals,  $\ln [Pr(X/K)]$  reached a maximum value at  $K = 2$  (Fig. A.1 in Supplementary material, Tables 3 and A.3 in Supplementary material), but other secondary submaxima for  $K = 3$  and  $K = 4$  were found; according to different authors, this may indicate the presence of substructure (Ferreira et al., 2016). Nevertheless, we have considered the highest probability for  $K = 2$ , observing clear differentiation of the two main groups, 29 genotypes (32.58%) and 33 accessions (21.15%) in RPP1, and 50 genotypes (56.18%) and 123 accessions (78.85%) in RPP2, both groups with  $qI > 80\%$ , and a third group (admixed) with 10 genotypes and  $qI < 80\%$ . The high number of accessions in RPP2 is due to the high number of duplicates for some cultivars such as ‘Reineta’ which only had 49 accessions (Tables 3 and A.3 in Supplementary material). RPP1 grouped six genotypes from Canary Islands, 19 from Galicia, two from Portugal and two Spanish cider cultivars used as references. RPP2 clustered seven genotypes from Canary Islands, 34 from Galicia, four from Portugal and five Spanish cultivars used as references.

(Pritchard et al., 2000) using 13 loci SSR in  $K = 2$  reconstructed populations (RPPs). Of the total of 230 alleles, 91 unique alleles (alleles that appear only in an RPP) were found. If we consider only the 281 diploid genotypes, we found nine unique alleles (9.89%) for RPP1 and 26 (28.57%) for RPP2. In the case of triploid genotypes (89), we found a higher number of unique alleles: 32 unique alleles (35.16%) belonging to RPP1 and 33 (36.26%) to RPP2 (Table A.4 in Supplementary material). The mean values of heterozygosity, both observed ( $H_o$ ) and expected ( $H_e$ ), for diploid individuals (281 genotypes) were similar in the two RPPs:  $H_o$  of 0.765 (RPP1) and 0.797 (RPP2) and  $H_e$  of 0.765 (RPP1) and 0.799 (RPP2) (Table A.5 in Supplementary material). In the case of triploid individuals (89 genotypes), in RPP1, the lowest  $H_e$  was 0.408 (locus Hi02c07) and the highest 0.895 (locus CH03d07), with an average value in this RPP of 0.784 (Table A.6 in Supplementary material). In the case of RPP2,  $H_e$  varied between 0.682 (locus CH01h10) and 0.903 (locus CH02c11), with an average value of 0.805. (Table A.6 in Supplementary material). Allelic richness was estimated using FSTAT software, based on a minimum size of 74 diploid individuals. For RPP1, the maximum value obtained was 15.6 (locus CH03d07), the minimum value was 4.6 (locus Hi02c07) and the average value was 10.8. In the case of RPP2, the highest value was 16.4 (CH01f02) and the lowest 6.9 (locus Hi02c07), with an average of 11.5. In the case of the admixed genotypes, the maximum value was 17.7 (locus CH03d07), the minimum was 8.9 (locus Hi02c07) and the

average was 12.3 (Table A.5 in Supplementary material). The mean value of FIS for RPP1 was 0.011. In RPP2, we observed an average value for FIS of 0.003. The mean value of FIT was very similar to that for RPP1, with a maximum value of 0.181 ( $P < 0.01$ ) (locus Hi02c07) and a minimum of  $-0.061$  (CH02c09) (Table A.5 in Supplementary material).

Regarding triploids, the mean value of FIS for RPP1 was 0.004, with a maximum value of 0.173 ( $P < 0.01$ ) (locus CH01h01) and a minimum value of  $-0.215$  ( $P < 0.01$ ) (locus CH01f02). In RPP2, we observed a mean value of  $-0.124$ , with a maximum of  $-0.020$  (locus CH01h10) and a minimum of  $-0.213$  ( $P < 0.01$ ) (locus CH02d08) (Table A.6 in Supplementary material). The mean value of FIT was very similar to that for RPP1, with a maximum value of 0.173 (locus CH01h01) and a minimum of  $-0.215$  ( $P < 0.01$ ) (locus CH01f02).

### 3.3. FCA and main groups of cultivars

FCA applied to diploid individuals corroborated the results obtained by STRUCTURE, with two groups of well-differentiated cultivars, one with local apple cultivars from Galicia, Portugal and the Canary Islands (RPP1) and the other with commercial cultivars and local cultivars (RPP2). These groups were distributed mainly on axis 1, with 83% of the total variation (Fig. 2). The result of AMOVA analysis for  $K = 2$ , when carried out on accessions with  $qI > 80$ , indicated that the proportion of genetic differentiation within the two main groups accounted for most of the molecular variance (92.8%) and only 7.2% ( $P < 0.001$ ) accounted for variation among the groups.

In this work, seven main groups of cultivars were relevant, ‘Pero’, ‘Camoesa’, ‘Repinaldo’, ‘Reineta’, ‘Tres en Cunca’, ‘Rayada’ and ‘Torres Agrelo’, because they showed a high number of genotypes (59) for the west of the Iberian Peninsula (Galicia and Portugal) and the Canary Islands (Tables A.3 and A.7 in Supplementary material). Diploid and triploid genotypes appear in all these groups of cultivars, except in the case of ‘Tres en Cunca’ which includes only triploid individuals, and ‘Repinaldo’ which has only diploid individuals (Table A.7 in Supplementary material). The identified genotypes are present in one or two of the RPPs, as well as in the group of admixed individuals. Nevertheless, most of the 59 genotypes of these groups are present in RPP2, specifically 22 of them (37.29%), 20 (33.90%) in the admixed group and 17 (28.81%) in RPP1 (Table A.7 in Supplementary material).

The groups with the largest number of accessions are ‘Pero’, with 18 genotypes (30.51%) in total (13 of which belong to RPP1), and ‘Reineta’ with 13 genotypes (22.03%) (nine of them belonging to RPP2). They are followed by the ‘Camoesa’ group with nine genotypes (15.25%), the majority (eight genotypes) belonging to the admixed group, and the

**Table 3**

Classification of 370 genotypes of *Malus domestica* by STRUCTURE (Pritchard et al., 2000) using 13 loci SSR in  $K = 2$  reconstructed populations (RPPs).

Ploidy	RPP	$qI$ (coefficient of coancestry)	N° (%) of genotypes	N° (%) de accessions	The most representative cultivars
Diploids	RPP1	$> 0.8$	119 (42.35)	155 (36.30)	Spanish and Portuguese local cvs. (‘Peros’) Reference cvs. (‘Golden Delicious’)
	RPP2	$> 0.8$	83 (29.54)	158 (37.00)	
	Admixed	$< 0.8$	79 (28.11)	114 (26.70)	
Triploids	RPP1	$> 0.8$	29 (32.58)	49 (23.67)	Spanish and Portuguese local cvs. (‘Peros’) ‘Reineta’ Group
	RPP2	$> 0.8$	50 (56.18)	141 (68.12)	
	Admixed	$< 0.8$	10 (11.24)	17 (8.21)	



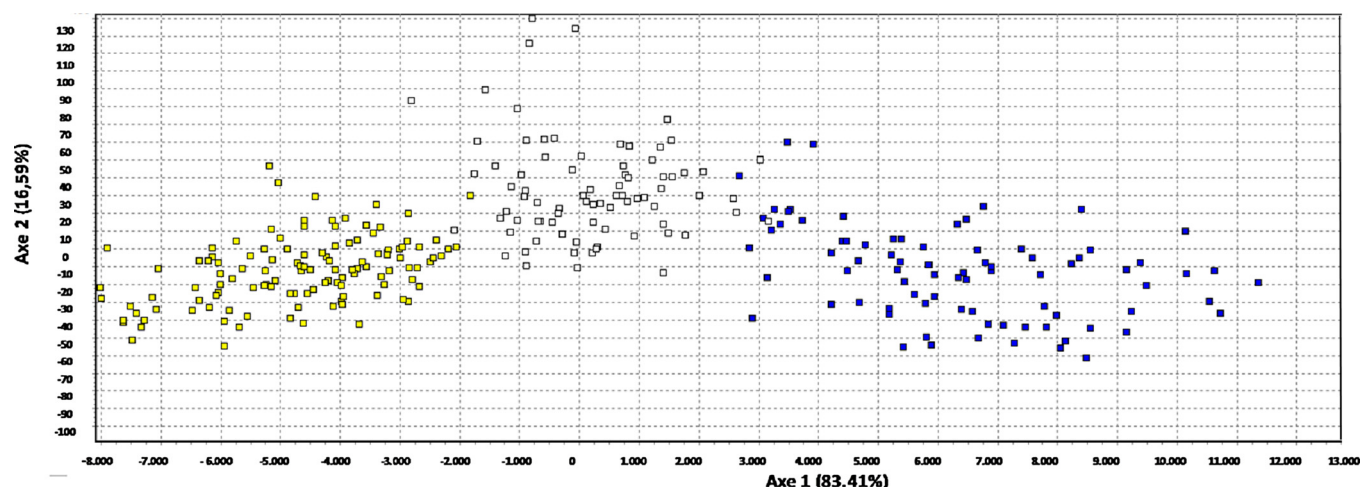


Fig. 2. FCA for 281 genotypes diploids. In light grey local cvs. from Galicia, Portugal and Canary Islands (RPP1); in dark grey reference and local cvs. (RPP2); and in white genotypes classified as admixed (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

‘Repinaldo’ group with seven genotypes (11.86%) (four of them in the admixed group, two in RPP2 and one in RPP1). The groups with the lowest number of genotypes are ‘Torres Agrelo’ with five genotypes (8.47%), ‘Rayada’ with four genotypes (6.78%) and ‘Tres en Cunca’ with three (5.08%), with their genotypes mainly belonging to RPP2 (Table A.7 in Supplementary material). All this reinforces the aforementioned, that local cultivars, for example ‘Peros’, are included in the RPP1 autochthonous group, and those related to commercial cultivars, such as ‘Reineta’, are included in RPP2.

It should be noted that the eight triploid genotypes of ‘Reineta’, belonging to RPP2, were grouped in 49 accessions, which indicates a considerable number of repetitions or a high percentage of clonality for this varietal group. Despite this, it is important to point out the great intracultural variability existing in these varietal groups; in the case of the ‘Reineta’ group, those eight triploid genotypes presented a minimum JC of 0.255 and a maximum of 0.882, with a mean value of 0.431 (Table A.7 in Supplementary material), and the eight different diploid genotypes of ‘Pero’ had a minimum JC of 0.091 and a maximum of 0.452, with a mean value of 0.215 (Table A.7 in Supplementary material).

‘Manzano’ from Canary Islands was found to be synonym to five Galician genotypes; eight more genotypes with this name were singular and genetically diverse, five of them diploids, and considered admixed, and three triploids, two clustering in RPP1 and one in RPP2.

### 3.4. Core collection

The application of a heuristic method using the PowerCore program (Kim et al., 2007) to the total of 370 diploid and triploid genotypes selected a total of 63 genotypes (17.03%) that conserves all the 230 alleles of the study. Excluding reference genotypes, and performing the same analysis only for Galicia (209 genotypes), then 54 genotypes (25.83%) and 200 alleles were conserved. In further analysis for Galicia, without those accessions that were collected as wild (200 genotypes), 55 genotypes (27.50%) and 189 alleles were conserved. In the case of the Canary Islands, of 55 genotypes, 30 (54.54%) and 150 alleles were selected; finally, for Portugal, of 41 genotypes, 24 (58.53%) and 140 alleles were selected (Tables 4 and A.3 in Supplementary material).

Taking into account the number of starting genotypes/accessions, in the case of the total core collection, 17.03/9.94% of the genotypes/accessions were retained to preserve the allelic variability. These percentages were slightly higher in the core collection of Galicia, 25.83/14.21% (total Galicia) and 27.50/15.07% (Galicia without wild), increasing considerably in the cases of the Canary Islands and Portugal,

with percentages of 54.54/39.47% and 58.53/35.82%, respectively.

Considering the maximum values of JC, the total core collection and those of Galicia, with and without wild accessions, presented a value of 0.750, while for the Canary Islands and Portugal, this value was reduced to 0.517 and 0.452, respectively (Table 4).

He was high (> 0.800) and almost equal in all core collections, since these conserve the greatest allelic diversity in the minimum number of individuals (Table A.8 in Supplementary material). The average He value was 0.856 for the total core collection, 0.845 for Galicia without wild individuals, 0.834 for Canary Islands and 0.828 for Portugal (Tables A.3 and A.8 in Supplementary material). FIS values varied significantly among the different core collections: values of 0.019 (total core collection), 0.023 (Portugal core collection) and 0.043 (Canary Islands core collection) indicated a slight deficit of heterozygotes, and markedly reduced values of −0.012 and −0.020 (Galicia core collections with and without wild accessions, respectively) seem to indicate a slightly higher degree of heterozygosity.

## 4. Discussion

### 4.1. Genetic diversity

Our study has allowed expansion of the study area with respect to the work of Ferreira et al. (2016), by incorporating the complete collection of Galicia, the collection of the Canary Islands (not compared previously) and a greater number of reference cultivars, in a geographical continuum that covers the entire north-west Iberian Peninsula and continues with the Canary Islands.

Regarding synonyms, genotypes of Portuguese and Galician collections were coincident, as previously described by Ferreira et al. (2016), indicating that the exchange of clonal cultivars has occurred between Galicia and Portugal and Canary Islands, but not between Portugal and the Canary Islands, surely due to a reduced interchange between those two geographical areas. On the other hand, it should be noted that an equal number of genotypes have been found in both Galicia and the Canary Islands. In the case of Portugal, synonyms include 16 accessions corresponding to 12 common genotypes, with 3.24% of the 370 total genotypes; in the case of Galicia and the Canary Islands, 74 accessions, corresponding to 13 common genotypes, represent 3.51% of the total genotypes.

The percentage of duplicates (clonality) is high at 41.64%, which denotes the great difficulty of morphologically identifying accessions, which must necessarily be complemented with molecular markers. This percentage of clonality was higher than that obtained in other studies in which it was around 30% (Ferreira et al., 2016; Liang et al., 2015), but

**Table 4**  
Comparison of parameters among Core Collections.

	General Core Collection	Core Collection Galicia	Core Collection Galicia without wild genotypes	Core Collection Canary Islands	Core Collection Portugal
No. Genotypes	370	209	200	55	41
No. of genotypes selected for the “Core Collection” (%)	63 (17.03)	54 (25.83)	55 (27.50)	30 (54.54)	24 (58.53)
No. of accessions	634	380	365	76	67
Selected genotypes respecting the total (%)	9.94	14.21	15.07	39.47	35.82
No. Alleles	230	200	189	150	140
He	0.856	0.853	0.845	0.834	0.828
$F_{IS}$	0.019	−0.012	−0.020	0.043	0.023
Minimum JC	0.000	0.000	0.000	0.000	0.044
Maximum JC	0.750	0.750	0.750	0.517	0.452
Average JC	0.143	0.152	0.166	0.158	0.165

Expected Heterozygosity (He), inbreeding coefficient (FIS), Jaccard Coefficient (JC).

practically identical to that obtained by Urrestarazu et al. (2012) for cultivars from the northeast of the peninsula (43.5%). The duplicates present in the study were expected due to the synonyms between Spanish and Portuguese cultivars (Ferreira et al., 2016); however, there are other causes such as labelling or misidentification, as has occurred in other collections of Spanish and Italian germplasm (Liang et al., 2015; Urrestarazu et al., 2012).

All SSR loci analysed showed a high degree of polymorphism, presenting between 12 and 24 alleles per locus. For the total of 370 individuals and 230 alleles, the average number of alleles per locus was 17.7, slightly lower than the values found in the American germplasm bank *Malus × domestica* and in accessions from all over Spain (Gross et al., 2014; Pereira Lorenzo et al., 2017), studies comprising a wider geographic area of apple cultivation. However, it was higher than the means obtained by other authors in restricted geographic regions of Portugal, Spain and Italy (Ferreira et al., 2016; Foroni et al., 2012; Liang et al., 2015; Pereira Lorenzo et al., 2007; Pina et al., 2014; Urrestarazu et al., 2012).

Nevertheless, the expected heterozygosity did not seem to be affected by the magnitude of the geographic region analysed in this and other studies of apple germplasm diversity. The average He of the 13 loci ( $He = 0.81$ ) was comparable to that found not only in collections of local cultivars from the northeast of the Iberian Peninsula and Portuguese Azores Islands (Foroni et al., 2012; Pina et al., 2014; Urrestarazu et al., 2012) but also of all Spain, Italy and USA (Gross et al., 2014; Liang et al., 2015; Pereira Lorenzo et al., 2017).

The mean value of FIS was 0.008, the same as for the 281 diploid genotypes, which shows a slight deficit of heterozygotes, although lower than the 0.067 shown for the American bank of cultivated apple (Gross et al., 2014) in which only diploid genotypes were considered.

#### 4.2. Structural and genetic similarity

The Bayesian analysis with STRUCTURE defined two reconstructed populations or RPPs, the first with local and indigenous accessions (RPP1) and the second with local and reference accessions (RPP2), in agreement with the two large groups defined for Spain (Pereira Lorenzo et al., 2017) and Portugal (Ferreira et al., 2016). In addition to the important gene flow found at European level (Urrestarazu et al., 2016), it seems that this group of Spanish local cultivars could differ from those that have originated other gene pools in Europe (Pereira Lorenzo et al., 2017). Additionally, an admixed group, which represents a group halfway between RPP1 and RPP2, included both reference and local accessions from Spain and Portugal.

The presence of local cultivars in RPP2 could be explained by the possible introduction of foreign material into local material, giving rise to new hybridizations that could have produced lineages that continued to propagate, as reported by Pereira Lorenzo et al. (2017). Local Portuguese cultivars have been related to the group ‘Reineta’ (Ferreira

et al., 2016), a group that in our study, as well in that of Pereira Lorenzo et al. (2017), is related to the commercial group. The Portuguese group of cultivars related to the group ‘Reineta’ could be explained by an early hybridization that occurred in Portugal. Moreover, the high number of triploid accessions for RPP2 (123 genotypes), in comparison with that for RPP1 (33 genotypes), reinforces the importance of the ‘Reineta’ group in this region, known in Galicia as ‘Tabardilla’ and in Portugal as ‘Espriega’ (Ferreira et al., 2016).

For the diploid individuals, with 281 genotypes, we observed heterozygosity values, both observed and expected, similar to those for the two RPPs, and lower than those for the third, admixed, group; the same was observed for allelic richness, which suggests that this is the result of cultivars not assigned to a specific RPP which have wide genetic diversity, because presumably they accumulate allelic variables of both subgroups.

In both diploid and triploid genotypes, the lowest He was found for the Hi02c07 locus, which may be related to the presence of null alleles. These He values were also reflected in the FIS values for the two RPPs, with a maximum for the Hi02c07 locus.

For RPP2 for the triploid individuals, it is worth mentioning the presence of significantly negative FIS values, in contrast to what was observed for RPP1. This could be due to the fact that the ‘Reineta’ group dominant in RPP2 (triploids) presents a high degree of heterozygosity, indicating the selection made by farmers in favour of larger fruits and subsequent propagation by grafting, giving rise to what is known as instantaneous domestication in perennial crops (Gross et al., 2014).

The genetic differentiation found in our study ( $F_{ST} = 0.072$ ) suggests low or no differentiation between groups according to Weir and Cockerham (1984), and is in line with previous studies of genetic diversity and structure of local apple cultivars such as those of Portugal ( $F_{ST} = 0.074$ ) (Ferreira et al., 2016) and of the northeast of Spain, for which values of 0.076 and 0.070 were obtained by Urrestarazu et al. (2012) and Pina et al. (2014), respectively. Liang et al. (2015) observed a total  $F_{ST}$  value of 0.056 among the four subgroups obtained in a core collection of apple cultivars from Italy, suggesting the existence of weaker population differentiation between the subgroups than that obtained in the present study.

#### 4.3. Varietal groups

In this study, we have been able to assign seven varietal groups, ‘Pero’, ‘Camoesa’, ‘Repinaldo’, ‘Reineta’, ‘Tres en Cunca’, ‘Rayada’ and ‘Torres Agrelo’, highlighting in RPP1 ‘Peros’, with a greater number of genotypes, and in RPP2 ‘Reineta’ (synonymous with ‘Tabardilla’ in Galicia and ‘Espriega’ in Portugal). All the varietal groups present a high degree of intravarietal diversity, although it is slightly lower in ‘Reineta’, as shown by the different genotypes found for each of these groups and the low JC among them.

In Portuguese germplasm, ‘Pero’ has been assigned to the main

group of local cultivars, RPP1, and the groups ‘Camoesa’ and ‘Reineta’ to RPP2 (Ferreira et al., 2016). In the present study encompassing a much higher number of accessions and a wider geographic region, the ‘Camoesa’ group happened to be admixed, as reported for a broad germplasm set in Spain by Pereira Lorenzo et al. (2017). Nevertheless, the designations ‘Camoesa’, ‘Pero’, ‘Tres en Cunca’, ‘Tres en Rama’, ‘Repinaldo’ and ‘Tabardilla’ (synonymous in Galicia with ‘Reineta’) appear referenced in the 18th century in Galicia between 1746 and 1775, with indication that the name was taken from the Castle of Camoes in Galicia (Sarmiento, 1986). Therefore, ‘Camoesa’ and ‘Reineta’ groups, being related by hybridization (Ferreira et al., 2016), had to have their origin in the 18th century or earlier.

#### 4.4. Core collections

By applying the heuristic method using the PowerCore program (Kim et al., 2007), we have been able to define the core collections for the west of the Iberian Peninsula (Galicia and Portugal) and the Canary Islands, that is, small groups of cultivars that preserve all the genetic variability of the whole. This same methodology has been applied for the germplasm banks of Galicia, the Canary Islands and Portugal. The size of the core collections depends, obviously, on the initial size of the sample and its allelic diversity, varying from a maximum value for the total core collection, up to a minimum value in the Portugal core collection.

A core collection has been defined to preserve the whole allelic diversity with the minimum redundancy to be useful in further comparative studies as a set of reference, substantially smaller than the one from which it is formed (van Hintum et al., 2000), including between 5% and 20% of the accessions of the complete collections, but usually no more than 10% of the entries. In our study, taking into account the number of starting accessions, in the case of the total core collection, 9.94% of genotypes were conserved to preserve the total allelic variability, and 14.21% and 15.07% were conserved for Galicia with and without including wild accessions, respectively, under 20% of the complete collections. Taking into account the genotypes, core collections would have 17% of the total set, surpassing 20% in the subgroups for the Canary Islands and Portugal. These values are in the range of 13% reported for the total collection for fig cultivars in Spain (Balas et al., 2014), and well above values reported by Le Cunff et al. (2008) for grapevine, in which only 4% of an entire collection represented the whole genetic diversity of the cultivated vines.

In the case of the Canary Islands and Portugal, even considering the starting accessions, it is necessary to conserve 39.47% and 35.82%, respectively, well above 20%, which indicates the high allelic diversity of these subgroups spread over a large number of genotypes.

#### 5. Conclusions

The high genetic diversity found in this studied was reflected in the values of expected heterozygosity, comparable to that of wider geographic regions, suggesting Galicia Portugal and Canary Islands as hotspots of apple germplasm diversity. The large-scale comparison of plant material conserved in the collections of the western Iberian Peninsula and the Canary Islands has allowed us to differentiate two main groups, the native and local cultivars and the reference cultivars and some cvs. considered as local, with moderate genetic differentiation according to the FST values, confirming that the local groups of cultivars is extended in the west of the Iberian Peninsula and Canary Islands. The presence of specific alleles for each of the two large groups will allow, in the future, traceability for Protected Geographical Indications and Denominations of Origin. Seven groups of relevant cultivars have been detected within the germplasm of the collections of the western Iberian Peninsula and the Canary Islands: ‘Pero’, ‘Reineta’ (local name ‘Tabardilla’), ‘Torres Agrelo’, ‘Tres en Cunca’, ‘Rayada’, ‘Camoesa’ and ‘Repinaldo’. ‘Camoesa’, ‘Pero Camoesa’ and ‘Tabardilla’ were already

referenced in the 18th century in Galicia (Sarmiento, 1986). Exchanges of local clonal cultivars between distant geographical areas, such as Galicia and the Canary Islands, were of the same order as those between neighbouring geographical areas, such as Galicia and Portugal, which may be explained by the importance of the historical relationship of Galicia with both areas.

One core collection has been defined to preserve the whole allelic diversity of the west of the Iberian Peninsula (Galicia and Portugal) and the Canary Islands, as well as two for Galicia (with and without wild accessions), one for the Canary Islands and another one for Portugal, which implies a considerable reduction in conservation efforts and allows further comparison studies involving these core collections and others from different regions towards preserving worldwide apple allelic diversity.

#### Funding sources

This work was supported by European Investment Funds by FEDER/COMPETE/POCI–Competitiveness and Internationalization Operational Programme, under the Project POCI-01-0145-FEDER-006958, National Funds by FCT–Portuguese Foundation for Science and Technology, under the project UID/AGR/04033/2013 and the scholarship SFRH/BD/96400/2013 and Galicia Norte Portugal European Grouping of Territorial Cooperation (GNP-EGTC) funds supporting an internship at University of Santiago de Compostela, under the IACOBUS program.

#### Acknowledgements

The authors acknowledge the “Research Institute of Horticulture and Seeds” INRA, Agrocampus Ouest, University of Angers, for providing some of the reference DNA samples. Moreover, we thank to A. Assunção (of DRAPN) and A. Marreiros (of DRAPA) for some Portuguese samples.

#### Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.scienta.2018.05.053>.

#### References

- Balas, F.C., Osuna, M.D., Domínguez, G., Pérez-Gragera, F., López-Corrales, M., 2014. Ex situ conservation of underutilised fruit tree species: establishment of a core collection for *Ficus carica* L. using microsatellite markers (SSRs). *Tree Genet. Genomes* 10, 703–710. <http://dx.doi.org/10.1007/s11295-014-0715-3>.
- Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F., 2004. GENETIX 4.05, logiciel Sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions. CNRS UMR 5000. Université de Montpellier II, Montpellier, France.
- Cornille, A., Gladieux, P., Smulders, M.J.M., Roldán-Ruiz, I., Laurens, F., Le Cam, B., Nersisyan, A., Clavel, J., Olonova, M., Feugey, L., Gabrielyan, I., Zhang, X.G., Tenaillon, M.I., Giraud, T., 2012. New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. *PLoS Genet.* 8 (5), e1002703. <http://dx.doi.org/10.1371/journal.pgen.1002703>.
- Cornille, A., Giraud, T., Smulders, M.J.M., Roldán-Ruiz, I., Gladieux, P., 2014. The domestication and evolutionary ecology of apples. *Trends Genet.* 30, 57–65. <http://dx.doi.org/10.1016/j.tig.2013.10.002>.
- Earl, D.A., vonHoldt, B.M., 2012. Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. <http://dx.doi.org/10.1007/s12686-011-9548-7>.
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software structure a simulation study. *Mol. Ecol.* 14, 2611–2620. <http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x>.
- Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.
- Ferreira, V., Ramos-Cabrera, A.M., Carnide, V., Pinto-Carnide, O., Assunção, A., Marreiros, A., Rodrigues, R., Pereira-Lorenzo, S., Castro, I., 2016. Genetic pool structure of local apple cultivars from Portugal assessed by microsatellites. *Tree Genet. Genomes* 12, 1–15. <http://dx.doi.org/10.1007/s11295-016-0997-8>.
- Foroni, I., Baptista, C., Monteiro, L., Lopes, M.S., Mendonça, D., Melo, M., Carvalho, C., Monjardino, P., Lopes, D.J.H., da Câmara Machado, A., 2012. The use of

- microsatellites to analyze relationships and to decipher homonyms and synonyms in Azorean apples (*Malus × domestica* Borkh.). *Plant Syst. Evol.* 298, 1297–1313. <http://dx.doi.org/10.1007/s00606-012-0637-1>.
- Gross, B.L., Henk, A.D., Richards, C.M., Fazio, G., Volk, G.M., 2014. Genetic diversity in *Malus × domestica* (Rosaceae) through time in response to domestication. *Am. J. Bot.* 101 (10), 1770–1779. <http://dx.doi.org/10.3732/ajb.1400297>.
- Gross, B.L., Wedger, M.J., Martinez, M., Gayle, M.V., Hale, C., 2018. Identification of unknown apple (*Malus × domestica*) cultivars demonstrates the impact of local breeding program on cultivar diversity. *Genet. Resour. Crop Evol.* 1–11. <http://dx.doi.org/10.1007/s10722-018-0625-6>.
- Hardy, O.J., Vekemans, X., 2002. Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* 2, 618–620. <http://dx.doi.org/10.1046/j.1471-8286.2002.00305.x>.
- Hokanson, S.C., Szewc-McFadden, A.K., Lamboy, W.F., McFerson, J.R., 1998. Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus × domestica* Borkh. core subset collection. *Theor. Appl. Genet.* 97, 671–683. <http://dx.doi.org/10.1007/s001220050943>.
- Kim, K.W., Chung, H.K., Cho, G.T., Ma, K.H., Chandrabalan, D., Gwag, J.G., Kim, T.S., Cho, E.G., Park, Y.J., 2007. PowerCore: a program applying the advanced M strategy with a heuristic search for establishing allele mining sets. *Bioinformatics* 23, 2155–2162.
- Lassois, L., Denancé, C., Ravon, E., Guyader, A., Guisnel, R., Hibrand-Saint-Oyant, L., Poncet, C., Lasserre-Zuber, P., Feugey, L., Durel, C.E., 2016. Genetic diversity, population structure, parentage analysis, and construction of core collections in the French apple germplasm based on SSR markers. *Plant Mol. Biol. Rep.* 34, 827–844. <http://dx.doi.org/10.1007/s11105-015-0966-7>.
- Lateur, M., Ordridge, M., Engels, J., Lipman, E., 2013. Report of a working group on *Malus/Pyrus*. In: Fourth Meeting, 7–9 March 2012, Weggis, Switzerland. Bioversity International, Rome, Italy. pp. 36.
- Le Cunff, L., Fournier-Leve, L.A., Laucou, V., Vezzulli, S., Lacombe, T., Adam-Blondon, A.F., Boursiquot, J.M., This, P., 2008. Construction of nested core collections to optimize the exploitation of natural diversity in *Vitis vinifera* L. subsp. sativa. *BMC Plant Biol.* 8 (31). <http://dx.doi.org/10.1186/1471-2229-8-31>.
- Leforestier, D., Ravon, E., Muranty, E., Cornille, A., Lemaire, C., Giraud, T., Durel, C.E., Branca, A., 2015. Genomic basis of the differences between cider and dessert apple varieties. 2015. *Evol. Appl.* 8 (7), 650–661. <http://dx.doi.org/10.1111/eva.12270>.
- Liang, W., Dondini, L., De Franceschi, P., Paris, R., Sansavini, S., Tartarini, S., 2015. Genetic diversity, population structure and construction of a core collection of apple cultivars from Italian germplasm. *Plant Mol. Biol. Rep.* 33, 458–473. <http://dx.doi.org/10.1007/s11105-014-0754-9>.
- Liebhart, R., Gianfranceschi, L., Koller, B., Ryder, C.D., Tarchini, R., Van De Weg, E., Gessler, C., 2002. Development and characterisation of 140 new microsatellites in apple (*Malus × domestica* Borkh.). *Mol. Breed.* 10, 217–241. <http://dx.doi.org/10.1023/A:1020525906332>.
- Lo, E.Y., Stefanovic, S., Dickinson, T.A., 2009. Population genetic structure of diploid sexual and polyploid apomictic hawthorns (Crataegus; Rosaceae) in the Pacific Northwest. *Mol. Ecol.* 18, 1145–1160.
- Meirmans, P.G., Van Tienderen, P.H., 2004. GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* 4, 792–794. <http://dx.doi.org/10.1111/j.1471-8286.2004.00770.x>.
- Michalakis, Y., Excoffier, L., 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142, 1061–1064.
- Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583–590.
- Pereira Lorenzo, S., Ramos Cabrer, A.M., Díaz Hernández, M.B., 2007. Evaluation of genetic identity and variation of local apple cultivars (*Malus × domestica* Borkh.) from Spain using microsatellite markers. *Genet. Resour. Crop Evol.* 54, 405–420.
- Pereira Lorenzo, S., Urrestarazu, J., Ramos-Cabrer, A.M., Miranda, C., Pina, A., Dapena, E., Moreno, M.A., Errea, P., Llamero, N., Díaz-Hernández, M.B., Santesteban, G., Laquidain, M.J., Gogorcena, Y., Urbina, V., Dalmases, J., Ascasibar Errasti, J., Royo, J.B., 2017. Analysis of the genetic diversity and structure of the Spanish apple genetic resources suggests the existence of an Iberian gene pool. *Ann. Appl. Biol.* 171 (3), 424–440. <http://dx.doi.org/10.1111/aab.12385>.
- Pereira-Lorenzo, S., Costa, R.M.L., Ramos-Cabrer, A.M., Marques Ribeiro, C.A., Serra da Silva, M.F., Manzano, G., Barreneche, T., 2010. Variation in grafted European chestnut and hybrids by microsatellites reveals two main origins in the Iberian Peninsula. *Tree Genet. Genomes* 6, 701–715. <http://dx.doi.org/10.1007/s11295-010-0285-y>.
- Pina, A., Urrestarazu, J., Errea, P., 2014. Analysis of the genetic diversity of local apple cultivars from mountainous areas from Aragon (Northeastern Spain). *Sci. Hortic.* 174, 1–9. <http://dx.doi.org/10.1016/j.scienta.2014.04.037>.
- Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, A., Carracedo, A., Lareus, M.V., 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. *Front. Genet.* 4, 98. <http://dx.doi.org/10.3389/fgene.2013.00098>.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Sarmiento, F.M., 1986. In: Pensado, J.L. (Ed.), Catálogo de voces vulgares y en especial de voces gallegas de diferentes vegetales. Ediciones Universidad Salamanca, Salamanca.
- Silfverberg-Dilworth, E., Matasci, C.L., Van De Weg, W.E., Van Kaauiwen, M.P.W., Walser, M., Kodde, L.P., Soglio, V., Gianfranceschi, L., Durel, C.E., Costa, F., Yamamoto, T., Koller, B., Gessler, C., Patocchi, A., 2006. Microsatellite markers spanning the apple (*Malus × domestica* Borkh.) genome. *Tree Genet. Genomes* 2, 202–224. <http://dx.doi.org/10.1007/s11295-006-0045-1>.
- Urrestarazu, J., Miranda, C., Santesteban, L.G., Royo, J.B., 2012. Genetic diversity and structure of local apple cultivars from Northeastern Spain assessed by microsatellite markers. *Tree Genet. Genomes* 8, 1163–1180. <http://dx.doi.org/10.1007/s11295-012-0502-y>.
- Urrestarazu, J., Denancé, C., Ravon, E., Guyader, A., Guisnel, R., Feugey, L., Poncet, C., Lateur, M., Houben, P., Ordridge, M., Fernandez-Fernandez, F., Evans, K.M., Papstein, F., Sedlak, J., Nybom, H., Garkava-Gustavsson, L., Miranda, C., Gassmann, J., Kellerhals, M., Suprun, I., Pikunova, A.V., Krasova, N.G., Torutaeva, E., Dondini, L., Tartarini, S., Laurens, F., Durel, C., 2016. Analysis of the genetic diversity and structure across a wide range of germplasm reveals prominent gene flow in apple at the European level. *BMC Plant Biol.* 16, 130. <http://dx.doi.org/10.1186/s12870-016-0818-0>.
- van Hintum, Th.J.L., Brown, A.H.D., Spillane, C., Hodgkin, T., 2000. Core Collections of Plant Genetic Resources. IPGRI Technical Bulletin No. 3. International Plant Genetic Resources Institute, Rome, Italy.
- Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S.K., Troggio, M., Pruss, D., Salvi, S., Pindo, M., Baldi, P., Castelletti, S., Cavauiolo, M., Coppola, G., Costa, F., Cova, V., Dal Ri, A., Goremykin, V., Komjanc, M., Longhi, S., Magnago, P., Malacarne, G., Malnoy, M., Micheletti, D., Moretto, M., Perazzolli, M., Si-Ammour, A., Vezzulli, S., Zini, E., Eldredge, G., Fitzgerald, L.M., Gutin, N., Lanchbury, J., Macalma, T., Mitchell, J.T., Reid, J., Wardell, B., Kodira, C., Chen, Z., Desany, B., Niaz, F., Palmer, M., Koepke, T., Jiwani, D., Schaeffer, S., Krishnan, V., Wu, C., Chu, V.T., King, S.T., Vick, J., Tao, Q., Mraz, A., Stormo, A., Stormo, K., Bogden, R., Ederle, D., Stella, A., Vecchietti, A., Kater, M.M., Masiero, S., Lasserre, P., Lepinasse, Y., Allan, A.C., Bus, V., Chagné, D., Crowhurst, R.N., Gleave, A.P., Lavezzo, E., Fawcett, J.A., Proost, S., Rouzé, P., Sterck, L., Toppo, S., Lazzari, B., Hellens, R.P., Durel, C.E., Gutin, A., Bumgarner, R.E., Gardiner, S.E., Skolnick, M., Egholm, M., Van de Peer, Y., Salamini, F., Viola, R., 2010. The genome of the domesticated apple (*Malus × domestica* Borkh.). *Nat. Genet.* 42 (10), 833–839. <http://dx.doi.org/10.1038/ng.654>.
- Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. <http://dx.doi.org/10.1111/j.1558-5646.1984.tb05657.x>.