Plant Molecular Biology Reporter (2018) 36:23–35
https://doi.org/10.1007/s11105-017-1058-7

ORIGINAL PAPER

Comparative Analysis of SSR Markers Developed in Exon, Intron, and Intergenic Regions and Distributed in Regions Controlling Fruit Quality Traits in *Prunus* Species: Genetic Diversity and Association Studies

Beatriz García-Gómez[1] · Mitra Razi[2] · Juan A. Salazar[3] · Angela S. Prudencio[1] · David Ruiz[1] · Luca Dondini[4] · Pedro Martínez-Gómez[1]

Published online: 24 November 2017

Abstract

Simple sequence repeats (SSRs) are genome domains located in both coding and non-coding regions in eukaryotic genomes. Although SSRs are often characterized by low polymorphism, their DNA-flanking sequences could be a useful source of DNA markers, which could help in genetic studies and breeding because they are associated with genes that control traits of interest. In this study, 56 genotypes from different *Prunus* species were used, including peach, apricot, plum, and almond (already phenotyped for several agronomical traits, including self-compatibility, flowering and ripening time, fruit type, skin and flesh color, and shell hardness). These *Prunus* genotypes were molecularly characterized using 28 SSR markers developed in exons, introns, and intergenic regions. All these genes were located in specific regions where quantitative trait loci (QTLs) for certain fruit quality traits were also located, including flowering and ripening times and fruit flesh and skin color. A sum of 309 SSR alleles were identified in the whole panel of analyzed cultivars, with expected heterozygosity values of 0.61 (upstream SSRs), 0.17 (exonic SSRs), 0.65 (intronic SSRs), and 0.58 (downstream SSRs). These values prove the low level of polymorphism of the exonic (gene-coding regions) markers. Cluster and structural analysis based on SSR data clearly differentiated the genotypes according to either specie (for the four species) and pedigree (apricot) or geographic origin (Japanese plum). In addition, some SSR markers mainly developed in intergenic regions could be associated with genes that control traits of interest in breeding and could therefore help in marker-assisted breeding. These findings highlight the importance of using molecular markers able to discriminate between the functional roles of the gene allelic variants.

Keywords SSR mining · Genome structure · *Prunus* · *Agronomical traits* · Breeding

Introduction

The plant family *Rosaceae* from the order *Rosales* comprises over 100 different genera and 3000 species (Shulaev et al. 2008) with a range of ornamental and agricultural uses. The *Prunus* genus, a member of this family, comprises 230 species divided into three main subgenera (*Amygdalus*, *Cerasus*, and *Prunus*) and a fourth subgenus of lesser interest, *Eplectocladus*, which includes desert almond species (Potter 2012). The following production data from 2015 demonstrate the importance of *Prunus* production in the world: (i) peach and nectarine [*P. persica* (L.) Batsch] ($2n = 2\times = 16$) fruits (21.63 million tons) and almond kernels (2.91 million tons) [*P. amygdalus* (Batsch) syn. *P. dulcis* (Miller) Webb] ($2n = 2\times = 16$) in the subgenus *Amygdalus*; (ii) sweet (*P. avium* L.) ($2n = 2\times = 16$), sour (*P. cerasus* L.) ($2n = 4\times = 32$), and

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11105-017-1058-7) contains supplementary material, which is available to authorized users.

✉ Pedro Martínez-Gómez
pmartinez@cebas.csic.es

[1] Department of Plant Breeding, CEBAS-CSIC, PO Box 164, E-30100 Espinardo, Murcia, Spain

[2] Department of Horticulture, Faculty of Agriculture, University of Zajan, Zajan, Iran

[3] Departamento de Producción Agrícola, Universidad de Chile, Santiago, Chile

[4] Dipartimento di Scienze Agrarie, Università degli Studi di Bologna, Via Fanin 46, 40127 Bologna, Italy

ground (*P. fruticosa* Pall.) ($2n = 4\times = 32$) cherry fruits (49 million tons) in the subgenus *Cerasus*; (iii) prune (*P. domestica* L.) ($2n = 6\times = 48$), Japanese plum (*P. salicina* Lindl) ($2n = 2\times = 16$), sloe (*P. spinosa* L.) ($2n = 4\times = 32$), and cherry plum (myrobalan) (*P. cerasifera* Ehrh.) ($2n = 2\times = 16$) fruits (11.52 million tons) in the subgenus *Cerasus*, section Prunus; and (iv) apricot (*P. armeniaca* L.) ($2n = 2\times = 16$) and mei (or Japanese apricot) (*P. mume* von Siebold & Zuccarini) ($2n = 2\times = 16$) fruits (4.11 million tons) in the subgenus *Cerasus*, section Armeniaca (http://faostat.fao.org).

Simple sequence repeats (SSRs, microsatellites) are small-scale DNA variations (from 1 bp to 1 kb) consisting of short repeat motifs present in both coding and non-coding regions of DNA sequences (Tautz and Renz 1984). SSRs show a high level of length polymorphism due to mutations of one or more repeats. In addition, they are the marker of choice for the assessment of genetic diversity within plant species because of their high polymorphism, abundance, and codominant inheritance (Gupta et al. 1996), and also because of their multi-allelic nature, reproducibility, and extensive genome coverage (Kantety et al. 2002). From the evolutionary point of view, SSRs are interesting since they are present in all eukaryotic genomes, including in plants, although they are distributed in low frequency in the coding regions. Furthermore, SSRs present in the non-coding regions are neutral, while those present in the coding regions are not (Powell et al. 1996; Varshney et al. 2005). Genic SSRs are also widely used to estimate population structure and in association analysis, because they can identify specific allelic variants of genes responsible for phenotype variability for specific traits of interest. The identification of specific allelic variants is a fundamental pre-requisite for the development of molecular tools able to support breeding programs selecting for specific traits by marker-assisted selection. Furthermore, association analysis eliminates the main drawbacks of classical linkage analysis, such as the cumbersome and expensive development of specific mapping populations, and this technique has significant potential in assessing a larger number of alleles with higher mapping resolution (Yu et al. 2006; Gupta et al. 2014).

SSR markers are extensively used in genetic diversity analysis studies in *Prunus* species. The identification of linked SSR markers underlying important agronomic traits such as disease resistance, stress tolerance, and fruit quality is of great importance for apricot breeding (Salazar et al. 2013, 2014, 2016). To give an example, the identified SSR marker PGS1.21 was recently used in the selection for *Plum pox virus* (PPV) resistance in apricot (Rubio et al. 2014). Nowadays, it is much more feasible to detect genomic variations thanks to the efficiency and sensitivity of the low-cost next generation sequencing (NGS) technologies (Martínez-Gómez et al. 2011, 2012; Saxena et al. 2014). Moreover, DNA databases, which are being added to on a daily basis, have become particularly attractive resources for the in silico mining of expressed sequence tag (EST)–SSRs or genic SSRs. In silico approaches can also be used in transferability studies because they contain conserved genic regions (Sorkheh et al. 2016).

Fruit quality traits are key traits from the breeding point of view in all the *Prunus* species (Infante et al. 2008, 2011). Nevertheless, the polygenic nature of most of the traits related to fruit quality, with genes distributed throughout the entire genome, makes it very difficult to develop linked markers. As a result, several researchers have focused on the study and characterization of such polygenic traits in different *Prunus* species and the identification of linked quantitative trait loci (QTLs) (Salazar et al. 2014). Many of the QTLs linked to fruit quality traits have been identified in scaffolds 3 and 4 in several *Prunus* species, including peach (Eduardo et al. 2011; Fresnedo-Ramírez et al. 2015), apricot (Salazar et al. 2013), and Japanese plum (Salazar et al. 2017).

The objective of this work was to perform a comparative analysis of genetic diversity and trait association studies in apricot, peach, Japanese plum, and almond genotypes by using 28 SSR markers isolated from either genomic sequences and transcripts in regions controlling fruit quality traits in *Prunus* that have been identified in exons, introns, and intergenic regions.

Materials and Methods

Plant Material

The plant material assayed in the validation assay analysis consisted of 56 *Prunus* genotypes from different species including apricot (Table 1), Japanese plum (Table 2), peach (Table 3), and almond (Table 4). The 56 *Prunus* genotypes used had previously been phenotyped for the following agronomical traits: (i) self-compatibility, evaluated by bagging as either self-compatible when there was fruit set or self-incompatible when there was no fruit set; (ii) flowering time, evaluated every 1 to 2 days until 50% of the flowers were completely opened (F_{50}), ranging from very early (before February 16) to very late (after March 12); (iii) ripening time, considered when the fruit had suitable firmness and color at the commercial maturity stage, ranging from early (before May 10) to very late (after June 20); (iv) fruit type, skin and flesh color, determined by a Minolta Chroma Meter (CR-300, Minolta, Ramsey, USA); and (v) shell hardness in the case of almond.

SSR Identification

The SSR markers identified in the introns and intergenic regions (Fig. 1) were identified by mapping the transcriptome reads obtained by NGS [reads in BAM format, available after

Table 1 Apricot cultivars assayed including the pedigree, origin, and main agronomic characteristics (self-compatibility, flowering and ripening time, and skin color)

Cultivar	Pedigree	Origin	Self-compatibility	Flowering	Ripening	Skin color
Altera	Harcot × (San Castrese × R. Imola)	Italy	Self-compatible	Medium	Early	Orange/red
Bergeron	Unknown	France	Self-compatible	Late	Very late	Light orange
Bergarouge	Bergeron × Orange Red	France	Self-incompatible	Late	Late	Orange/red
Búlida	Unknown	Spain	Self-compatible	Medium	Medium	Light orange
Canino	Unknown	Spain	Self-compatible	Early	Medium	Yellow
Cornia	NJA1 × Bella di Imola	Italy	Self-incompatible	Medium	Medium	Orange/red
Currot	Unknown	Spain	Self-compatible	Very early	Very early	Light yellow
Dorada	Bergeron × Moniquí	Spain	Self-compatible	Late	Late	Yellow
Estrella	Orange Red × Z211–18	Spain	Self-incompatible	Medium	Medium	Orange/red
Goldrich	Sunglo x Perfection	USA	Self-incompatible	Late	Late	Orange
Harcot	Modem 604 × NJA1	Canada	Self-incompatible	Late	Medium	Orange/red
Kioto	Unknown	France	Self-compatible	Late	Medium	Orange
Lito	Stark Early Orange × Tyrithos	Greece	Self-compatible	Late	Late	Orange
Mauricio	Unknown	Spain	Self-compatible	Early	Early	Yellow
Micaelo	Orange Red × Tardif de Bordaneil	Spain	Self-compatible	Medium	Medium	Orange/red
Mirlo Blanco	Rojo Pasión × Búlida Precoz	Spain	Self-compatible	Early	Very early	Light orange/-red
Mirlo Anaranj-ado	Rojo Pasión × Búlida Precoz	Spain	Self-compatible	Early	Very early	Light orange/-red
Tardorange	Orange Red × Tardif de Bordaneil	Spain	Self-compatible	Late	Late	Orange/red
Orange Red	Lasgerdi Mashad × NJA2	USA	Self-incompatible	Late	Medium	Orange/red
Palsteyn	Blenhein × Canino	South Afri-ca	Self-compatible	Medium	Medium	Orange
Petra	Goldrich × Pelese di Giovanniello	Italy	Self-compatible	Medium	Late	Orange
Ninfa	Ouardi × Tyrinthos	Italy	Self-compatible	Early	Early	Yellow
Pieve	Harcot × Reale di Imola	Italy	Self-compatible	Medium	Late	Yellow
Pieve Tardiva	Harcot × Reale di Imola	Italy	Self-compatible	Medium	Late	Orange/red
Reale di Imola	Unknown	Italy	Self-compatible	Late	Late	Light orange
Rojo Pasión	Orange Red × Currot	Spain	Self-compatible	Medium	Early	Light orange/-red
Tardif de Valence	Unknown	France	Self-compatible	Late	Late	Orange
Valorange	Orange Red × Currot	Spain	Self-compatible	Medium	Medium	Orange/red

an RNA-Seq approach on five different cDNA libraries from 'Rojo Pasión' and two from 'Z506-7' (Salazar et al. 2015), were processed by SAMtools (samtools.sourceforge.net) and displayed in alignment with the peach genome by Tablet (Milne et al. 2013)] with the *Prunus* reference genome [*Prunus persica* genome v2.0 (http://www.rosaceae.org/

Table 2 Japanese plum cultivars assayed including the pedigree, origin, and main agronomic characteristics (self-compatibility, flowering and ripening time, skin color and flesh color)

Cultivar	Pedigree	Origin	Self-compatibility	Flowering	Ripening	Skin color	Flesh color
Angeleno	Queen Ann × OP	USA	Self-incompatible	Medium	Very late	Violet-Black	Yellow
Black Diamond	Angeleno × OP	USA	Self-incompatible	Early	Medium	Black	Red
Black Splendor	Black Amber × OP	USA	Self-incompatible	Very early	Early	Black	Red
Honey Dawn	Unknown	South Africa	Self-incompatible	Early	Early	Yellow	Yellow
Pioneer	Unknown	South Africa	Self-compatible	Very early	Early	Red	Yellow
Red Beauty	Eldorado × Burmosa	USA	Self-incompatible	Early	Early	Red	Yellow
Santa Rosa	Unknown	USA	Self-compatible	Medium	Medium	Red	Yellow
Songold	Golden King × Wickson	South Africa	Self-incompatible	Late	Late	Yellow	Yellow

peach/genome) (Verde et al. 2017). Sequences were uploaded to the Phytozome website, and mask options were used to identify and eliminate repetitive domains. Most of these SSRs have been identified in scaffolds 3 and 4 of the *Prunus* reference genome (http://www.rosaceae.org/peach/genome; Verde et al. 2017) (Table S1), where different QTLs linked to fruit quality traits have been described in peach (Eduardo et al. 2011; Fresnedo-Ramírez et al. 2015), apricot (Salazar et al. 2013), and Japanese plum (Salazar et al. 2017). Regarding the SSRs developed in exon regions, EST sequences of different *Prunus* species, including peach, apricot, sweet cherry, mei, almond, sour cherry, and prune, were downloaded in FASTA format from Genbank (ftp://ncbi.nlm.nih.gov/genbank/genomes/) (Fig. 1). The genomic SSRs were detected using GMATo software, available for free at http://sourceforge.net/p/GMATo (Sorkheh et al. 2016) (Table S1).

SSR Analysis

Total genomic DNA was isolated using a procedure described by Doyle and Doyle (1987). Extracted apricot genomic DNA was PCR-amplified using 28 different primer pairs flanking SSR sequences distributed among the whole genome in exon (EST-SSRs), intron, and intergenic regions (Table S1). SSR-PCRs were performed according to the multiplex PCR protocol, as described by Campoy et al. (2010), using primers labeled with FAM, VIC, NED, or PET fluorescent dyes. SSR amplifications were analyzed in an ABI Prism 3730 DNA Analyzer (Applied Biosystems, MA, USA), and SSR peaks were visualized with Peak Scanner 1.0 software.

Molecular Diversity Evaluation and Cluster Analysis

The SSR allelic data obtained were used to estimate diversity parameters, including the number of alleles, size range, and the observed heterozygosity (Ho) as the number of heterozygous genotypes divided by the total number of genotypes. In addition, the power of discrimination (PD) of each SSR marker was calculated as $PD = 1 - \sum g^2_i$ where g_i is the frequency of the i^{th} genotype. A neighbor-joining (NJ) dendrogram was produced using MEGA 7 software (Kumar et al. 2016). Relative support for the branches in each dendogram was

Table 3 Peach cultivars assayed including the pedigree, origin, type, and main agronomic characteristics (self-compatibility, flowering and ripening time, and flesh color)

Cultivar	Pedigree	Origin	Type	Self-compatibility	Flowering	Ripening	Flesh color
Alisio 10	Sevilla 2 × Flor da Prince	Spain	Peach	Self-compatible	Very early	Very early	Yellow
Astoria	Unknown	Spain	Peach	Self-compatible	Very early	Very early	Yellow
Carioca	Unknown	Spain	Flat Peach	Self-compatible	Early	Early	White
Flariba 117	Unknown	Spain	Nectarine	Self-compatible	Very early	Very early	Yellow
Flat July	Unknown	France	Flat Peach	Self-compatible	Late	Late	White
Honey prima	Unknown	USA	Nectarine	Self-compatible	Early	Early	Yellow
Levante 10	Precocinho × OP	Spain	Peach	Self-compatible	Very early	Very early	Yellow
Mesembrina	Fantasia × (Jalousia × Summergrand)	France	Flat nectarine	Self-compatible	Medium	Medium	Yellow
Mistral 30	Big Bang × Messembrina	Spain	Flat nectarine	Self-compatible	Medium	Medium	Yellow
Precocinho	Diamante × OP	Brazil	Peach	Self-compatible	Very early	Early	Yellow
Romea	Catherina × OP	Italy	Peach	Self-compatible	Late	Late	Yellow
Siroco 10	UFO 3 × IC22	Spain	Flat peach	Self-compatible	Early	Early	White

Table 4 Almond cultivars assayed including the pedigree, origin, and main agronomic characteristics (self-compatibility, shell hardness, and flowering time)

Cultivar	Pedigree	Origin	Self-compatibility	Flowering	Shell hardness
Antoñeta	Ferragnès × Tuono	Spain	Self-compatible	Late	Hard
Desmayo	Unknown	Spain	Self-incompatible	Early	Hard
Garrigues	Unknown	Spain	Self-incompatible	Medium	Hard
Marta	Ferragnès × Tuono	Spain	Self-compatible	Late	Hard
Nonpareil	Unknown	USA	Self-incompatible	Early	Soft
Penta	S5133 × Lauranne	Spain	Self-compatible	Very Late	Hard
R1000	Tardif Nonpareil × Tuono	France	Self-compatible	Late	Semi-hard
Tardona	S5133 × R1000	Spain	Self-compatible	Extra Late	Hard

assessed with 2000 replicates of NJ bootstrap. Finally, genetic distance estimation was performed using the maximum composite likelihood (MCL) method.

Trait Association Studies

Marker-trait association analyses between SSRs and agronomic traits were performed using TASSEL v5 software (Bradbury et al. 2007). A general lineal model (GLM) using numeric data joined to genotype data and principal component analysis (PCA) was used for generating Manhattan plots for each trait and year. In addition, a mixed lineal model (MLM) was applied using kinship data to define the relationship between individuals, because MLM sometimes has higher statistical power than GLM and may detect more true associations.

Results and Discussion

Molecular Diversity of Assayed SSRs and *Prunus* Genotypes

SSR amplifications were successful with all the 28 tested primers (Table S2). Polymorphic bands were generated for all the primers used, with a total of 309 scored polymorphic

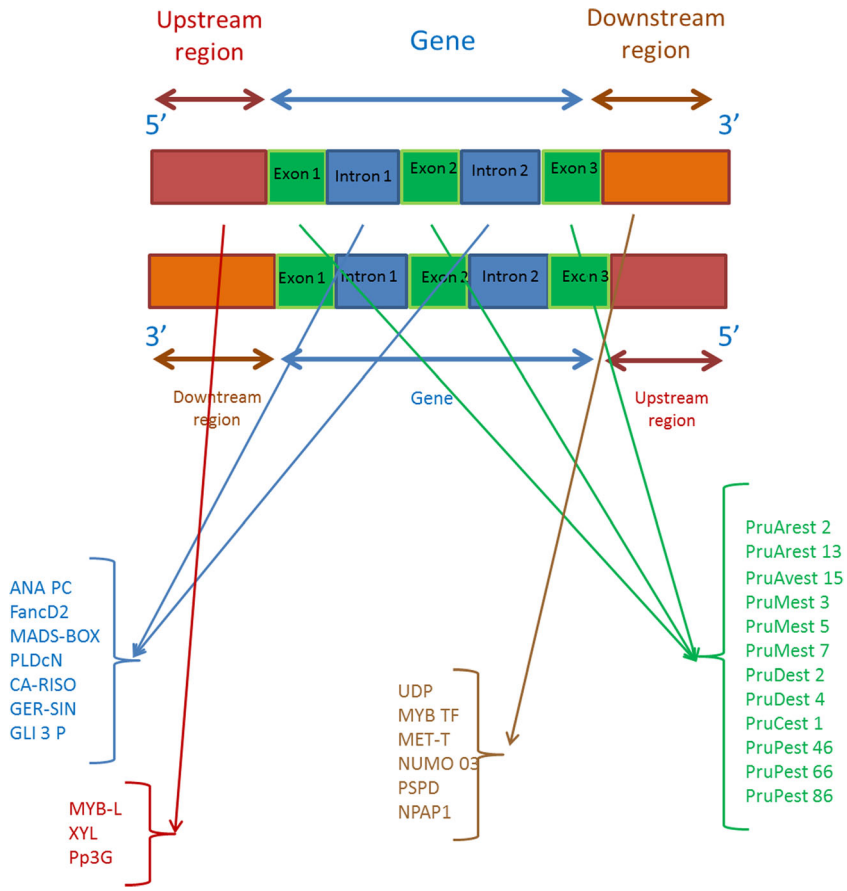


Fig. 1 SSRs identified in exon (EST-SSRs), intron, and intergenic regions

Table 5 Simple sequence repeat (SSR) markers assayed and polymorphism obtained in the *Prunus* cultivars assayed

SSR marker	No. of Alleles	Size range	Heterozygosity	Power of discrimination
Upstream				
MYB-L	17	164–205	0.49	0.59
XYL-1	17	237–265	0.60	0.50
Pp3G	25	154–209	0.74	0.47
Mean	*19.6*		*0.61*	*0.52*
Exon				
PruArest 2	2	187–195	0.00	0.21
PruArest 13	2	201–211	0.01	0.04
PruAvest 15	3	175–193	0.24	0.03
PruMest 3	1	183	0.00	0.03
PruMest 5	9	212–277	0.32	0.49
PruMest 7	5	278–296	0.90	0.50
PruDest 2	3	231–297	0.52	0.09
PruDest 4	1	169	0.00	0.03
PruCest 1	2	452–460	0.00	0.23
PruPest 46	4	138–155	0.00	0.32
PruPest 66	4	140–152	0.00	0.60
PruPest 86	2	189–218	0.00	0.24
Mean	*3.2*		*0.17*	*0.23*
Intron				
ANA PC	20	205–237	0.73	0.70
MADS BOX	11	186–213	0.76	0.43
FancD2	18	115–153	0.78	0.61
PLDcN	10	144–183	0.69	0.65
CA-RISO	22	51–181	0.54	0.23
GER-SIN	19	193–237	0.80	0.55
GLI 3 P	8	211–238	0.28	0.65
Mean	*15.4*		*0.65*	*0.54*
Downstream				
UDP	22	206–284	0.64	0.43
MYB TF	15	232–260	0.34	0.73
MET-T	22	128–212	0.65	0.71
NUMO D3	13	191–216	0.62	0.67
PSP D1	14	203–258	0.58	0.61
NPAP1	18	132–167	0.66	0.70
Mean	*17.3*		*0.58*	*0.64*

bands. Most of the identified markers were polymorphic, and only two (Prumest 3 and PruDest 4) were monomorphic. The highest number of presumed alleles revealed by the 28 SSRs was 25 (Pp3G), with a mean of 10.4 alleles per locus (Table 5). In addition, the 28 SSR markers assayed for genetic diversity studies generated a total of 150 genotypes with an average of 6 genotypes per marker. SSR alleles were identified when genotypes from all the species were considered with mean expected heterozygosity values of 0.61 (upstream SSRs), 0.17 (exonic SSRs), 0.65 (intronic SSRs), and 0.58 (downstream SSRs), proving the lower level of polymorphism in the exonic SSR markers. The maximum observed heterozygosity (Ho) values were recorded by SSR marker GER-SIN (0.80), and minimum Ho values were recorded by SSR markers PruArest 2, PruDest 4, PruCest 1, PruPest 46, PruPest 66, and PruPest 86 (0) (Table 5).

SSR markers in gene-coding (exonic) regions showed lower polymorphism in comparison to SSR markers developed in intron (intronic) and intergenic (upstream SSRs and downstream SSRs) regions. This lower polymorphism of SSRs developed in exons has also been described in maize (Holland et al. 2001). In this specie, 67% of promoter

Table 6 Heterozygosity levels of *Prunus* cultivars assayed

Cultivars	Heterozygosity
Apricot	
Bergeron	0.371
Búlida	0.400
Currot	0.371
Estrella	0.428
Goldrich	0.457
Lito	0.400
Mauricio	0.342
Mirlo Blanco	0.342
Mirlo Anaranjado	0.342
Orange Red	0.342
Palsteyn	0.485
Rojo Pasión	0.457
Valorange	0.342
Canino	0.200
Micaelo	0.400
Tardorange	0.457
Bergerouge	0.300
Altera	0.314
Ninfa	0.485
Cornia	0.400
Harcot	0.400
Kioto	0.257
Petra	0.428
Pieve	0.285
Pieve Tardiva	0.342
Dorada	0.228
Reale di Imola	0.314
Tardif Valence	0.428
Mean	*0.368*
Plum	
Santa Rosa	0.342
Red Beauty	0.371
Cultivars	Heterozygosity
Pioneer	0.314
Black Splendor	0.371
Black Diamond	0.342
Honey Dawn	0.371
Angeleno	0.400
Sungold	0.342
Mean	*0.356*
Peach	
Alisio 10	0.342
Astoria	0.200
Romea	0.257
Flariba 117	0.257
Honey prima	0.228
Carioca	0.257
Siroco 10	0.257

Table 6 (continued)

Cultivars	Heterozygosity
Mistral 30	0.285
Mesembrina	0.285
Levante 10	0.171
Flat July	0.228
Precocinho	0.228
Mean	*0.249*
Almond	
Desmayo	0.342
Marta	0.400
Penta	0.457
Tardona	0.285
Antoñeta	0.400
R1000	0.342
Nonpareil	0.457
Garrigues	0.371
Mean	*0.381*

(upstream) markers, 58% of intron markers, and 13% of exon markers exhibited amplified product-length polymorphism. These results agree with our results. Furthermore, Huang et al. (2016) also found that the intergenic regions exhibited the highest relative abundance and diversity in six species of birds.

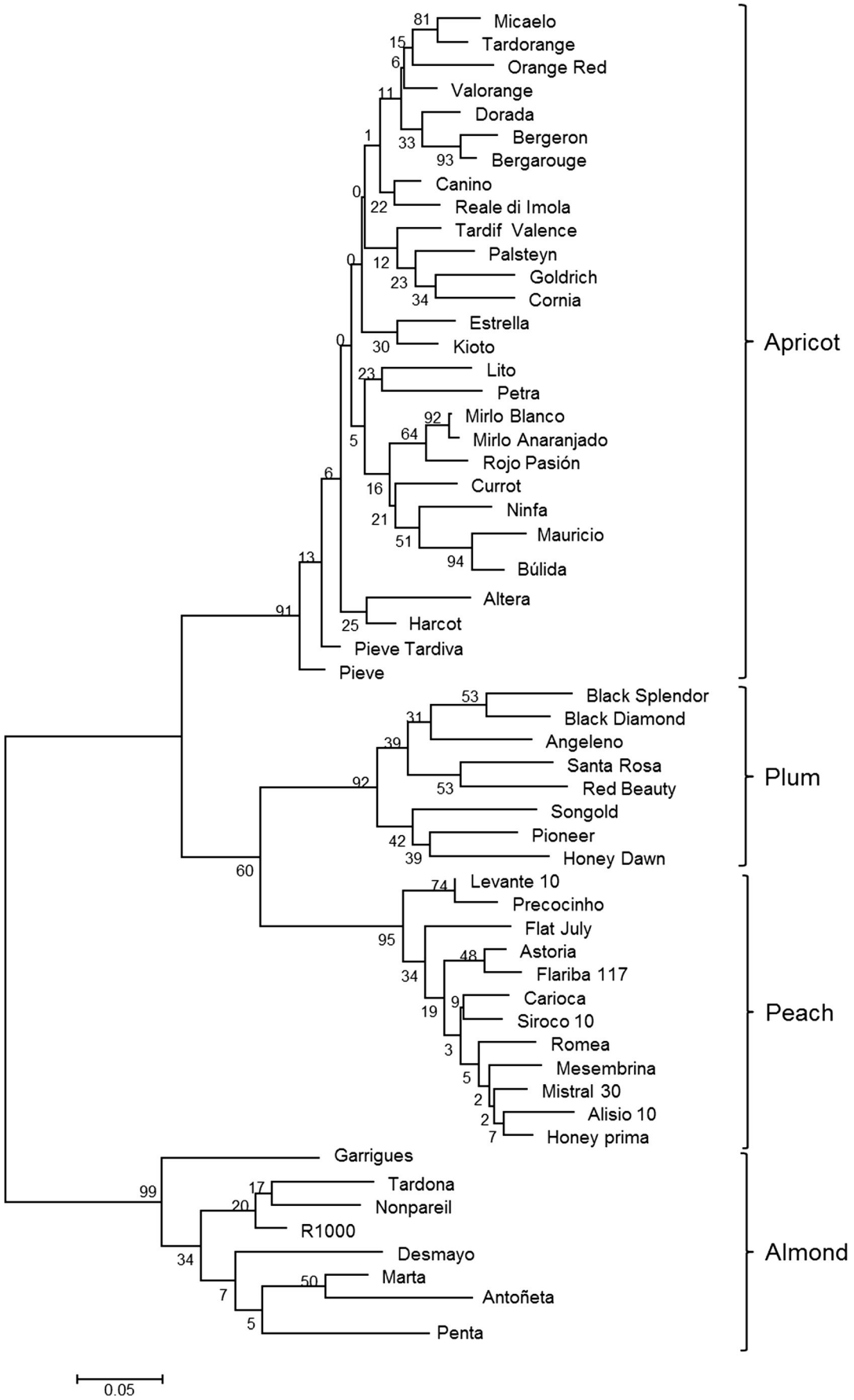
The genetic heterozygosity of *Prunus* genotypes evaluated using SSR markers ranged from 0.171 ('Levante 10' peach) to 0.485 ('Plasteyn' apricot), with an average value of 0.33 (Table 6). The mean observed heterozygosity value in almond (0.381) was higher than the observed heterozygosity in apricot (0.368) and plum (0.356), proving the lower level of polymorphism in the peach genotype markers (0.249). The greater heterozygosity for almond, apricot, and plum relative to peach can be attributed to the mating system differences between these species. Peach is self-fertilizing and inbreeding, whereas almond, apricot, and plum are normally self-incompatible and therefore outcrossing. These differences in the heterozygosity of *Prunus* species identified using SSRs and attributed to the mating system have also been described by other authors using intronic and intergenic SSRs (Martínez-Gómez et al. 2003; Sánchez-Pérez et al. 2006).

SSR markers are ideally selectively neutral, but researchers often have concerns that these markers are prone to selective pressures when they occur in or near coding regions (Morgante et al. 2002). For example, SSR loci in *Glycine* are three times more likely to occur in translated regions when derived from transcriptomic data than genomic data, and the majority of the loci found in translated regions are trinucleotide repeats (Hodel et al. 2016).

Clustering of *Prunus* Genotypes

Figure 2 shows the phenetic relationships among the different *Prunus* genotypes assayed. The NJ dendogram showed four main clusters that group together the genotypes of the four species assayed: apricot, plum, peach, and almond. Inside each specie cluster, the different genotypes were grouped together according to pedigree (apricot) or geographical origin (plum).

In the first cluster, apricot genotypes, there is very strong support for clustering cultivars with related pedigrees. We can see one cluster with the traditional Spanish cultivars, including 'Mauricio', 'Búlida', 'Currot', and the Italian 'Ninfa'. 'Búlida Precoz' seems to be a mutation of 'Búlida'. Other clearly identified clusters include descendants of the cultivar 'Orange Red', 'Micaelo', 'Tardorange', and 'Valorange' and the descendants of 'Rojo Pasión', 'Mirlo Anaranjado', and 'Mirlo Blanco'. The French cultivar 'Bergeron' and its descendant 'Bergerouge' are also very close. The same trend seems to be shown by 'Pieve', 'Pieve Tardiva', and 'Altera', which are clustered closer to 'Harcot' than to the cultivar in their pedigree, 'Reale di Imola'. In the cluster grouping plum genotypes together, however, there is very strong support for clustering cultivars with related geographical origins. We can see two main clusters with the plums from either USA ('Black Splendor', 'Black Diamond', 'Angeleno', 'Santa Rosa', and 'Red Beauty') or from South Africa ('Songold', 'Pioner', and 'Honey Dawn'). On the other hand, in the case of the peach and almond clusters, no significant groups were detected according to origin or pedigree, although there was a cluster


Micaelo
Tardorange
Orange Red
Valorange
Dorada
Bergeron
Bergarouge
Canino
Reale di Imola
Tardif Valence
Palsteyn
Goldrich
Cornia
Estrella
Kioto
Lito
Petra
Mirlo Blanco
Mirlo Anaranjado
Rojo Pasión
Currot
Ninfa
Mauricio
Búlida
Altera
Harcot
Pieve Tardiva
Pieve
Apricot
Black Splendor
Black Diamond
Angeleno
Santa Rosa
Red Beauty
Songold
Pioneer
Honey Dawn
Plum
Levante 10
Precocinho
Flat July
Astoria
Flariba 117
Carioca
Siroco 10
Romea
Mesembrina
Mistral 30
Alisio 10
Honey prima
Peach
Garrigues
Tardona
Nonpareil
R1000
Desmayo
Marta
Antoñeta
Penta
Almond
0.05

◀ **Fig. 2** Dendogram obtained by NJ cluster analysis based on the mean character difference distances among the peach, almond, plum, and apricot cultivars evaluated with the 28 EST-SSRs assayed. Numbers below branches represent bootstrapping values. The scale bar represents simple matching distance

between 'Mesmbrina' peach and its descendant 'Mistral 30', and a cluster was formed between 'Marta' and 'Antoñeta' almond, which are from the same 'Ferragnès' × 'Tuono' cross (Fig. 2).

Finally, pairwise genetic distances among the 56 *Prunus* genotypes calculated by the MCL method presented an overall value of 0.413 (Table S3). The lowest values (0.006) were reported for related cultivars, such as the apricots 'Mirlo Anaranjado' and 'Mirlo Blanco', which are both from 'Rojo Pasión' × 'Búlida Precoz'. A high value was also found for the apricot genotype 'Tardorange' and the almond genotype 'R-1000'. These genetic distances in general, including the great distance between 'Tardorange' and 'R-1000', confirmed the phenetic relationships shown in Fig. 2. Genetic distance analysis of the SSRs assayed also confirmed the completely different origins of the *Prunus* genotypes assayed.

Trait Association Genetics

Marker-trait association analyses between SSRs and agronomic traits were performed using TASSEL v5 software. A mixed linear model (MLM) using numeric data joined the genotype (SSR marker) and phenotype (fruit quality traits). In this analysis, some SSR markers mainly developed in intergenic regions could be associated with genes that control traits of interest in breeding and therefore could help in marker-assisted breeding. The most significant (p value lower than 0.005) marker associated with the traits of interest showed R^2 values ranging from 0.588 (NPAP1 linked to ripening) and 0.097 (PruPest 46 linked to skin color) (Table 7). In general, SSRs located in intron regions showed better association with phenotypic traits. However, some associations have also been found in coding regions (exonic), as recently described in mango by Lal et al. (2017).

Table 8 shows the associations identified at the single species level. Several markers showed good linkage with flesh and skin color and flowering time in plum. When the genotypes were separated into two groups according to their flesh color (red and yellow), it was possible to observe that the allele was only present in all genotypes with red-fleshed fruits. The alleles 135 bp (CA-RISO) and 174 bp (PLDcN) were present in all the genotypes with black-skinned fruits. Finally, the presence of the alleles 207 bp from the MADBOX marker correlated with late flowering. In the case of peach, only one marker showed linkage with flowering time (Table 8). The presence of the alleles 210 bp from the MAD BOX marker correlated with very early flowering in

Table 7 Marker trait association by mixed linear model (MLM) using TASSEL v5 software for different fruit quality traits in apricot, Japanese plum, and peach cultivars

Marker	Trait	F marker	Df marker	Df error	p value	R^2 marker
Intron						
ANA PC	Skin color	3.74	7	18	0.011	0.402
	Flesh color	3.74	7	18	0.011	0.402
CA-RISO	Flowering	3.02	4	25	0.036	0.306
	Skin color	2.68	4	22	0.050	0.282
	Flesh color	2.68	4	22	0.050	0.282
PLDcN	Skin color	2.23	4	40	0.050	0.190
MADS BOX	Flowering	2.29	6	35	0.050	0.250
GLI 3 P	Skin color	2.82	5	20	0.043	0.281
	Flesh color	2.75	6	31	0.029	0.147
GER-SIN	Flesh color	2.27	9	30	0.044	0.196
Downstream						
MYB TF	Ripening	2.58	6	29	0.039	0.277
NPAP1	Ripening	5.02	8	20	0.001	0.588
Exon						
PruPest 66	Skin color	4.39	1	32	0.044	0.097
PruPest 46	Skin color	4.39	1	32	0.041	0.097
Upstream						
Pp3G	Skin color	9.67	5	23	0.045	0.438
	Flesh color	5.34	5	23	0.002	0.399

F Fisher test, *Df* degrees of freedom

Table 8 SSR alleles obtained with the highest level of co-segregation within phenotypes (fruit quality traits) of the Japanese plum and peach cultivars assayed

Marker	SSR allele (bp)	Percentage of a given phenotype[a]	
Japanese plum			
Flesh color			
		Red	Yellow
CA-RISO	135	100 (2)	0 (0)
Skin color			
		Black	Other colors
CA-RISO	135	100 (3)	0 (0)
PLDcN	174	100 (3)	0 (0)
Flowering time			
		Late	Early-medium
MADS BOX	207	100 (1)	0 (0)
Peach			
Flowering time			
		Medium-late	Very early
MADS BOX	210	0 (0)	100 (3)
	212	0 (0)	100 (2)

[a] Between parenthesis is indicated the number of observations of each phenotype

genotypes 'Astoria', 'Alisio 10', and 'Flariba', although this was not the case in the very early Brazilian genotype 'Precocinho'. Finally, in almond, no marker showed linkage with any trait.

In agreement with our results, several authors have identified a QTL linked to fruit skin color (Eduardo et al. 2011) and flesh color (Yamamoto et al. 2001) in peach in the position where the CA-RISO marker (Scaffold_4: 16074538..16074688) is located. In the position where the Linkage Group 3 CA-RISO markers (Scaffold_3: 15343738..15343905) are located, several authors have also identified several QTLs linked to fruit skin color in apricot (Ruiz et al. 2010), peach (Eduardo et al. 2011), and sweet cherry (Sooriyapathirana et al. 2010), and QTLs linked to flesh color in peach (Abbott et al. 1998; Quilot et al. 2004; Illa et al. 2011) and sweet cherry (Sooriyapathirana et al. 2010).

In *Prunus* species, several members of the MADS-BOX Transcription Factor family (Bianchi et al. 2015) have been associated with the control of genes responsible for bud dormancy and flowering (Bielenberg et al. 2008; Jiménez et al. 2009, 2010). In addition, in the position where the SSR-MAD-BOX marker is located in Linkage group 3 (Scaffold_3: 18608201..18608410), several authors have identified a QTL linked to flowering time in apricot (Olukolu et al. 2009), peach (Romeu et al. 2014), and almond (Sánchez-Pérez et al. 2007). The involvement of MADS box genes in flowering induction has already been described in other Rosaceae species such as apple. In fact, the use of a silver birch MADS Box gene in the genetic transformation of apples has produced early flowering phenotypes that have been used for fast-track breeding approaches (Flachowsky et al. 2007).

Conclusions

SSR markers in gene-coding (exonic) regions showed lower polymorphism than SSR markers developed in intron (intronic) and intergenic (upstream SSRs and downstream SSRs) regions. Furthermore, cluster and structural analysis based on SSR data clearly differentiated the genotypes according to either specie (for the four species) and pedigree (apricot) or geographic origin (Japanese plum). In addition, some SSR markers mainly developed in intergenic regions could be linked to genes that control traits of interest in breeding and could therefore help in marker-assisted breeding.

Acknowledgements The authors offer grateful thanks to the University of Zajan and the Chilean government (FONDECYT Postdoctoral fellowship No. 3160080) for financial assistance in the stay of Mitra Razi and Juan A. Salazar in Murcia. This study was also supported by the "Apricot breeding" project of the Spanish Ministry of Economy and Competiveness (AGL2013-41452-R) and the project "Breeding stone fruit species assisted by molecular tools" of the Seneca Foundation of the Region of Murcia (19879/GERM/15).

References

Abbott AG, Rajapakse S, Sosinski B, Lu ZX, Sossey-Alaoui K, Gannavarapu M, Reighard G, Ballard RE, Baird WV, Callahan A (1998) Construction of saturated linkage maps of peach crosses segregating for characters controlling fruit quality, tree architecture and pest resistance. Acta Hortic 465:41–49

Bianchi V, Rubio M, Trainotti L, Verde I, Bonghi C, Martínez-Gómez P (2015) *Prunus* transcription factors: breeding perspectives. Front Plant Sci 6:443. https://doi.org/10.3389/fpls.2015.00443

Bielenberg DG, Wang Y, Li ZG, Zhebentyayeva T, Fan SH, Reighard GL, Scorza R, Abbott AG (2008) Sequencing and annotation of the

evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-BOX transcription factors as candidate genes for regulation of terminal bud formation. Tree Genet Genomes 4(3): 495–507. https://doi.org/10.1007/s11295-007-0126-9

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23(19):2633–2665. https://doi.org/10.1093/bioinformatics/btm308

Campoy JA, Martínez-Gómez P, Ruiz D, Rees J, Celton JM (2010) Developing microsatellite multiplex and megaplex PCR systems for high throughput characterization of breeding progenies and linkage maps spanning the apricot genome. Plant Mol Biol Report 28(4):560–568. https://doi.org/10.1007/s11105-010-0186-0

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Eduardo I, Pacheco I, Chietera G, Bassi D, Pozzi C, Vecchietti A, Rossini L (2011) QTL analysis of fruit quality traits in two peach intraspecific populations and importance of maturity date pleiotropic effect. Tree Genet Genomes 7(2):323–335. https://doi.org/10.1007/s11295-010-0334-6

Flachowsky H, Peil A, Sopanen T, Elo A, Hanke V (2007) Overexpression of BpMADS4 from silver birch (*Betula pendula* Roth.) induces early-flowering in apple (*Malus domestica* Borkh.) Plant Breed 126(2):137–145. https://doi.org/10.1111/j.1439-0523.2007.01344.x

Fresnedo-Ramírez J, Bink MC, van de Weg E, Famula TR, Crisosto CH, Frett TJ, Gradziel TM (2015) QTL mapping of pomological traits in peach and related species breeding germplasm. Mol Breed 35:1–19

Gupta PK, Balyan HS, Sharma PC, Ramesh B (1996) Microsatellites in plants: a new class of molecular markers. Curr Sci 70:45–54

Gupta S, Kumari K, Muthamilarasan M, Parida SK, Prasad M (2014) Population structure 586 and association mapping of yield contributing agronomic traits in foxtail millet. Plant Cell Rep 33(6):881–893. https://doi.org/10.1007/s00299-014-1564-0

Hodel RG, Gitzendanner MA, Germain-Aubrey CC, Liu X, Crowl AA, Sun M, Soltis DE (2016) A new resource for the development of SSR markers: millions of loci from a thousand plant transcriptomes. Appl Plant Sci 4(6):1600024. https://doi.org/10.3732/apps.1600024

Holland JB, Helland SJ, Sharopova N, Rhyne DC (2001) Polymorphism of PCR-based markers targeting exons, introns, promoter regions, and SSRs in maize and introns and repeat sequences in oat. Genome 44(6):1065–1076. https://doi.org/10.1139/g01-110

Huang J, Li W, Jian Z, Yue B, Yan YF (2016) Genome-wide distribution and organization of microsatellites in six species of birds. Biochem Syst Ecol 67:95–102. https://doi.org/10.1016/j.bse.2016.05.023

Illa I, Eduardo I, Audergon JM, Barale F, Dirlewanger E, Li X, Moing A, Lambert P, Le Dantec L, Gao Z, Poëssel JL, Pozzi C, Rossini L, Vecchietti A, Arús P, Howad W (2011) Saturating the *Prunus* (stone fruits) genome with candidate genes for fruit quality. Mol Breed 28(4):667–682. https://doi.org/10.1007/s11032-010-9518-x

Infante R, Martínez-Gómez P, Predieri S (2008) Quality oriented fruit breeding: peach [*Prunus persica* (L.) Batsch]. J Food Agric Environ 6:342–356

Infante R, Martínez-Gómez P, Predieri S (2011) Breeding for fruit quality in *Prunus*. In: Jenks MA, Bebeli PJ (eds) Breeding for fruit quality. Editorial Wiley & Blackwel, New York (EEUU), pp 201–229

Jiménez S, Lawton-Rauh AL, Reighard GL, Abbott AG, Bielenberg DG (2009) Phylogenetic analysis and molecular evolution of the dormancy associated MADS-BOX genes from peach. BMC Plant Biol 9(1):81. https://doi.org/10.1186/1471-2229-9-81

Jiménez S, Reighard GL, Bielenberg DG (2010) Gene expression of DAM5 and DAM6 is suppressed by chilling temperatures and inversely correlated with bud break rate. Plant Mol Biol 73(1-2):157–167. https://doi.org/10.1007/s11103-010-9608-5

Kantety RV, La Rota M, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. Plant Mol Biol 48(5/6):501–510. https://doi.org/10.1023/A:1014875206165

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 bigger datasets. Mol Biol Evol 33(7): 1870–1874. https://doi.org/10.1093/molbev/msw054

Lal S, Singh AK, Sing SK, Srivastav M, Sing BP, Sharma N, Sing NM (2017) Association analysis for pomological traits in mango (*Mangifera indica* L.) by genic-SSR markers. Trees 31:1391–1409

Martínez-Gómez P, Arulsekar S, Potter D, Gradziel TM (2003) An extended interspecific gene pool available to peach and almond breeding as characterized using simple sequence repeat (SSR) markers. Euphytica 131(3):313–322. https://doi.org/10.1023/A:1024028518263

Martínez-Gómez P, Crisosto CH, Bonghi C, Rubio M (2011) New approaches to *Prunus* transcriptome analysis. Genetica 139(6):755–769. https://doi.org/10.1007/s10709-011-9580-2

Martínez-Gómez P, Sánchez-Pérez R, Rubio M (2012) Clarifying omics concepts, challenges, and opportunities for *Prunus* breeding in the postgenomic era. OMICS 16(5):268–283. https://doi.org/10.1089/omi.2011.0133

Milne I, Stephen G, Bayer M, Cock PJ, Pritchard L et al (2013) Using tablet for visual exploration of second-generation sequencing data. Brief Bioinform 14(2):193–202. https://doi.org/10.1093/bib/bbs012

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nature Genet 30:194–200

Olukolu B, Trainin T, Fan S, Kole C, Bielenberg D, Reighard G, Abbott A, Holland D (2009) Genetic linkage mapping for molecular dissection of chilling requirement and budbreak in apricot (*Prunus armeniaca* L.) Genome 52(10):819–828. https://doi.org/10.1139/G09-050

Potter D (2012) Basic information on the stone fruit crops. In: Kole C, Abbott AG (eds) Genetics, genomics and breeding of stone fruits. CRC Press, New York, pp 1–21. https://doi.org/10.1201/b13104-2

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1(7):215–222. https://doi.org/10.1016/S1360-1385(96)86898-0

Quilot B, BH W, Kervella J, Génard M, Foulongne M, Moreau K (2004) QTL analysis of quality traits in an advanced backcross between *Prunus persica* cultivars and the wild relative species *P. davidiana*. Theor Appl Genet 109(4):884–897. https://doi.org/10.1007/s00122-004-1703-z

Romeu JF, Monforte AJ, Sánchez G, Granell A, García-Brunton J, Badenes ML, Ríos G (2014) Quantitative trait loci affecting reproductive phenology in peach. BMC Plant Biol 14(1):52. https://doi.org/10.1186/1471-2229-14-52

Rubio M, Ruiz D, Egea J, Martínez-Gómez P, Dicenta F (2014) Opportunities of marker assisted selection for plum pox virus resistance in apricot breeding. Tree Genet Genomes 10(3):513–525. https://doi.org/10.1007/s11295-014-0700-x

Ruiz D, Lambert P, Audergon JM, Gouble B, Bureau S, Reich M, Dondini L, Tartarini S, Adami M, Bassi D, Testolin R (2010) Identification of QTLs for fruit quality traits in apricot. Acta Hortic 862:587–592

Salazar JA, Ruiz D, Egea J, Martínez-Gómez P (2013) Transmission of fruit quality traits in apricot (*Prunus armeniaca* L.) and analysis of linked quantitative trait loci (QTLs) using simple sequence repeat (SSR) markers. Plant Mol Biol Report 31(6):1506–1517. https://doi.org/10.1007/s11105-013-0625-9

Salazar JA, Ruiz D, Campoy JA, Sánchez-Pérez R, Crisosto CH, Martínez-García PJ, Blenda A, Jung S, Main D, Martínez-Gómez P, Rubio M (2014) Quantitative trait loci (QTL) and Mendelian trait loci (MTL) analysis in *Prunus* a breeding perspective and beyond. Plant Mol Biol Report 32(1):1–18. https://doi.org/10.1007/s11105-013-0643-7

Salazar JA, Rubio M, Ruiz D, Tartarini S, Martínez-Gómez P, Dondini L (2015) SNP development for genetic diversity analysis in apricot. Tree Genet Genomes 11(1):15. https://doi.org/10.1007/s11295-015-0845-2

Salazar JA, Ruiz D, Campoy JA, Tartarini S, Dondini L, Martínez-Gómez P (2016) Inheritance of reproductive phenology traits and related QTL identification in apricot. Tree Genet Genomes 12(4):71. https://doi.org/10.1007/s11295-016-1027-6

Salazar JA, Pacheco I, Shinya P, Zapata P, Silva C, Ruiz D, Martínez-Gómez P, Infante R (2017) Genotyping by sequencing for SNP-based linkage analysis and identification of QTLs linked to fruit quality traits in Japanese plum (*Prunus salicina* Lindl.) Front Plant Sci 8:476. https://doi.org/10.3389/fpls.2017.00476

Sánchez-Pérez R, Martínez-Gómez P, Dicenta F, Egea J, Ruiz D (2006) Level and transmission of genetic heterozygosity in apricot, explored by simple sequence repeat markers. Genet Resour Crop Evol 53(4):763–770. https://doi.org/10.1007/s10722-004-4636-0

Sánchez-Pérez R, Howad D, Dicenta F, Arús P, Martínez-Gómez P (2007) Mapping major genes and quantitative trait loci controlling agronomic traits in almond. Plant Breed 126(3):310–318. https://doi.org/10.1111/j.1439-0523.2007.01329.x

Saxena RK, Edwards D, Varshney RK (2014) Structural variations in plant genomes. Brief Funct Genomics 13(4):296–307. https://doi.org/10.1093/bfgp/elu016

Shulaev V, Korban SS, Sosinski B, Abbott AG, Aldwinckle HS, Folta KM, Iezzoni A, Main D, Arús P, Dandekar AM, Lewers K, Gardiner SE, Potter D, Veilleux E (2008) Multiple models for Rosaceae genomics. Plant Physiol 147(3):985–1003. https://doi.org/10.1104/pp.107.115618

Sooriyapathirana SS, Khan A, Sebolt AM, Wang D, Bushakra JM, Wang KL, Allan AC, Gardiner SE, Chagné H, Iezzoni AF (2010) QTL analysis and candidate gene mapping for skin and flesh color in sweet cherry fruit (*Prunus avium* L.) Tree Genet Genomes 6(6):821–832. https://doi.org/10.1007/s11295-010-0294-x

Sorkheh K, Prudencio AS, Ghebinejad A, Dehkordi MK, Erogul D, Rubio M, Martínez-Gómez P (2016) In silico search, characterization and validation of new EST-SSR markers in the genus *Prunus*. BMC Res Notes 9(1):336. https://doi.org/10.1186/s13104-016-2143-y

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukayotic genomes. Nucleic Acid Res 12(10):4127–4138. https://doi.org/10.1093/nar/12.10.4127

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23(1):48–55. https://doi.org/10.1016/j.tibtech.2004.11.005

Verde I, Jenkins J, Dondini L, Micali S, Pagliarani G, Vendramin E, Schmutz J (2017) The peach v2.0 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. BMC Genomics 18(1):225. https://doi.org/10.1186/s12864-017-3606-9

Yamamoto T, Shimada T, Imai T, Yaegaki H, Haji T, Matsuta N, Yamaguchi M, Hayashi T (2001) Characterization of morphological traits based on a genetic linkage map in peach. Breed Sci 51(4):271–278. https://doi.org/10.1270/jsbbs.51.271

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38(2):203–208. https://doi.org/10.1038/ng1702