3 Biotech (2018) 8:16
https://doi.org/10.1007/s13205-017-1036-7

ORIGINAL ARTICLE

Molecular identification and genetic analysis of cherry cultivars using capillary electrophoresis with fluorescence-labeled SSR markers

Chenglin Liang[1] · Tian Wan[1] · Shiyan Xu[2] · Binbin Li[1] · Xiangnan Li[1] · Ying Feng[1] · Yuliang Cai[1]

Received: 21 May 2017 / Accepted: 4 December 2017 / Published online: 11 December 2017

Abstract

Molecular identification and genetic analysis of cherry are necessary for solving the problem of synonyms and homonyms that occur in cherry production. In this study, capillary electrophoresis with fluorescent-labeled simple sequence repeat (SSR) primers was used to identify 63 cherry cultivars (varieties and rootstocks) planted in Shaanxi province, China. A total of 146 alleles were amplified by 10 SSR primer pairs, ranging from 10 to 20 per locus (mean: 14); among the SSR primer pairs, genotype number ranged from 12 to 26 (mean: 18). The mean values of gene diversity, heterozygosity, and polymorphism information content were 0.7549 (range 0.4011–0.8782), 0.5952 (range 0.3810–0.9683), and 0.7355 (range 0.3937–0.8697), respectively. An unweighted pair-group method with arithmetic average cluster analysis was used to separate the cherry cultivars. A model-based structure analysis separated the cultivars into three populations, which was consistent with the results of a phylogenic and principal component analysis. Based on Bayes' rule, the cultivars were further subdivided into seven populations. Some of the 63 cherry cultivars that are often confused in production were distinguished, and DNA fingerprinting of cherry cultivars was established. This research will significantly assist in the identification of cherry cultivars at the molecular level.

Keywords Cherry · Genetic diversity · Simple sequence repeat · Fluorescent capillary electrophoresis · DNA fingerprinting

Introduction

Cherries are thought to have originated around the Caspian and Black Seas, and are also found across mainland Europe and in western Asia (Webster 1996). Sweet cherry (*Prunus avium* L.) is an important fruit in temperate regions of the world (Wünsch and Hormaza 2002) and its production has rapidly developed, because the numerous cultivated varieties exhibit early maturity with good quality, highly nutritious large and brightly colored fruit with moderate sweet and sour flavors. In contemporary tree fruit production, the selection of rootstocks is an important long-term management decision that may influence fruit production and quality (Turkoglu et al. 2012; Ognjanov et al. 2015). Due to their effective stress resistance, early fruiting, and dwarfing, *Prunus mahaleb*, *P. cerasus*, and *P. pseudocerasus* are currently widely used in cherry production as rootstocks. Therefore, genetic analysis and identification of sweet cherry and its stock are of great significance for the production industry. It is difficult to accurately morphologically identify the clones of sweet cherry varieties during the seedling period (Struss et al. 2001), and due to a lack of reliable early identification methods for sweet cherry varieties, both synonyms and homonyms are present in significant numbers (Turet-Sayar et al. 2012) causing large losses in cherry production. In addition, the genetic relationships among cultivars are unclear. In the past, genetic resources in cherries were evaluated according to phenotype characterization, but variability effects of environmental factors limited trial stability and predictive accuracy. Currently, the identification period is too long, and the fruits of some cherry varieties are very similar in appearance; therefore, rapid

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13205-017-1036-7) contains supplementary material, which is available to authorized users.

✉ Yuliang Cai
yuanyicyl@nwsuaf.edu.cn

[1] College of Horticulture, Northwest A&F University, Yangling, Shaanxi, China

[2] Shaanxi Provincial Fruit Trees Propagation Center, Tongchuan, Shaanxi, China

and accurate identification of sweet cherry varieties has become problematic and needs to be resolved to ensure effective cherry production.

Many studies have shown that there are significant differences among varieties of sweet cherry at the molecular level, and these differences can be detected by DNA analysis (Cipriani et al. 1999). Among the various marker types, simple sequence repeats (SSRs), or microsatellite markers, are a desirable tool used to construct DNA fingerprinting and analyze genetic diversity (Cipriani et al. 1999). SSRs, consisting of 1–6 nucleotides as repeating units, are widely distributed on chromosomes within the eukaryotic genomes (Smith 1994). SSR markers are the markers of choice in genetic diversity assessment, fingerprinting, and genotyping, due to their codominant Mendelian inheritance, high levels of polymorphism, and rapid and convenient detection (Turet-Sayar et al. 2012). SSR markers have been applied to many kinds of plants, especially fruit tree species such as sweet cherry, peach, and apple (Wünsch and Hormaza 2004; Ercisli et al. 2011). In *Prunus* species, most of the available SSR sequences have been developed from peach (Cipriani et al. 1999; Dirlewanger et al. 2002), sweet cherry (Dirlewanger et al. 2002), and sour cherry (Downey and Iezzoni 2000; Lacis et al. 2009) species, and have been widely used over the last 10 years to assess the genetic diversity among genotypes (Guarino et al. 2010) and rootstocks (Turkoglu et al. 2010). Gene polymorphism analysis of conventional SSR molecular markers has been conducted through polyacrylamide gel electrophoresis (PAGE) combined with manual reading; however, this method is time-consuming, labor intensive, and non-automated. Moreover, there are still considerable difficulties in the collection and analysis of massive and multi-batch data that are mainly reflected by the difficulty in allele identification and the unmanageability of data obtained from different batches (Chandra et al. 2014). Therefore, in this study, we used capillary electrophoresis with fluorescent-labeled SSR primers because of its high efficiency and automation. The technique has been widely used in studies of many plants involving molecular markers (Hayden et al. 2008; Liang et al. 2010; Chandra et al. 2014).

Here, we used 10 pairs of SSR fluorescent-labeled primers to amplify 63 cherry cultivars planted in Shaanxi province, China. Genotyping was performed by capillary electrophoresis, enabling the DNA fingerprinting of 63 cherry cultivars to be established, and we explored the relationships among them. In China, Shaanxi province has developed into a cherry-growing region, and it is, therefore, important for the cherry industry to identify the cherry cultivars at the molecular level.

Materials and methods

Plant materials

Tender leaves of 63 cherry cultivars consisting of 39 sweet cherry cultivars, 6 sour cherry cultivars, and 18 cherry rootstocks were collected in six regions of Shaanxi province from April to May 2016. Young leaves collected from these populations were immediately placed in zip-locked plastic bags containing silica gel for drying. The sample information is given in Table S1.

DNA extraction

Genomic DNA was extracted from leaves using a modified cetyl trimethylammonium bromide (CTAB) method (Doyle and Doyle 1987; Hormaza 1999). Dry leaves were ground to a powder by tissue grinding apparatus and then placed in a centrifuge tube. The tissue powders were exposed to 700 µL extraction buffer (3% CTAB, 5 M NaCl, 0.5 M EDTA, 1 M Tris–HCl pH 8.0) with 2 µL β-mercaptoethanol and incubated at 65 °C for 45 min. Following incubation, equal volumes of a chloroform and isoamylalcohol (24:1) solution were added before centrifugation at 15,000×*g* for 10 min. Genomic DNA from the aqueous phase was precipitated at − 20 °C for 30 min by twice adding ice-cold isopropanol and centrifuging at 15,000×*g* for 6 min. The sediment was gently washed with 70% ethanol and centrifuged at 15,000×*g* for 30 s. This procedure was repeated twice. The precipitate was dried naturally. Genomic DNA was dissolved in 100 µL ddH_2O (double distilled water). DNA was checked by 1.0% (w/v) agarose gel electrophoresis and its concentration was determined using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Finally, the DNA was diluted to 100 ng μL^{-1}.

SSR primers and polymerase chain reaction (PCR) amplification

Among the 104 specific primers for *P. avium*, *P. persica*, *P. salicina*, *P. dulcis*, *P. cerasus*, and *P. pseudocerasus*, 10 microsatellite primers (Dirlewanger et al. 2002; Vaughan and Russell 2004; Ai et al. 2007), which can produce clear, simple, and repeatable bands, were selected to analyze the 63 cultivars (Table 1).

PCR reactions were performed in 25 µL volumes containing 10 × PCR Buffer (100 mM Tris–HCl pH 8.8 at 25 °C, 500 mM KCl, 0.8% (v/v) Nonidet), 10 mM dNTP, 25 mM $MgCl_2$, 10 µM each primer, 100 ng genomic DNA, 5U μL^{-1} Taq polymerase, and ddH_2O. The PCR amplification procedure was conducted at 95 °C for 3 min,

Table 1 Primers used in simple sequence repeat (SSR) analysis of cherry cultivars with fluorescent capillary electrophoresis

Locus	Repeat motif	Forward and reverse primer sequences (5′–3′)	Size (bp)	References
SC2	$(AC/CA)_8$	ATTCGGGTCGAACTCCCT ACGAGCACTAGAGTAACCCTCTC	136–177	Ai et al. (2007)
SC3	$(AC)_8TT(TA)_5$	ACCCACAAATCAAGCATATCC AGCTTCAGCCACCAAGC	140–172	Ai et al. (2007)
BPPCT013	$(AG)_{28}$	ACCCACAAATCAAGCATATCC AGCTTCAGCCACCAAGC	140–172	Dirlewanger et al. (2002)
BPPCT026	$(AG)_8GG(AG)_6$	ATACCTTTGCCACTTGCG TGAGTTGGAAGAAAACGTAACA	134–182	Dirlewanger et al. (2002)
EMPaS01	$(GA)_9(GA)_{11}$	CAAAATCAACAAAATCTAAACC CAAGAATCTTCTAGCTCAAACC	215–266	Vaughan and Russell (2004)
EMPaS02	$(TTG)_7CTGC(TG)_{10}(AG)_8$	CTACTTCCATGATTGCCTCAC AACATCCAGAACATCAACACAC	109–159	Vaughan and Russell (2004)
CPSCT038	$(GA)_{18}$	CAGGAACCCTATTCCCACAA TCAATGGCACCCATTTTACA	182–209	Mnejja et al. (2004)
EPDCU5060	$(CAT)_8$	ACCAAATTGGACATGCAACC CGGTCGAGAAGACTGAGGAG	98–148	GDR database
EPDCU5183	$(CT)_{20}$	AGCAGTCTTTGCCAAATCAA TACAGGGTCCACATGATCCA	95–175	GDR database
EPPCU4092	$(AAAG)_6(AAG)_4$	AAGAAGAAGACGACGACGAC TCTGTATCCACCACGAGACC	124–262	GDR database

followed by 10 cycles of 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s and then 20 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s, and a final extension step at 72 °C for 6 min. PCR products were separated by 9% PAGE in 1 × TBE buffer at 180 V for 1.5 h. The gels were stained with 0.5% (w/v) silver nitrate for 10 min and colored with 1.5% (w/v) NaOH and 0.4% (v/v) formaldehyde for 5 min.

Capillary electrophoresis detection

Fluorescent primers were obtained from Sagon Biotech Co., Ltd (Shanghai, China) and the fluorescent dyes were FAM, HEX, and PAMRA. An electronic version of a test table was made, and the machine table was generated automatically. A mixture of 990 µL HIDI and 10 µL ROX500 or LIZ500 was placed in a 96-well reaction plate with a continuous pipette, with each well having a volume of 10 µL. The well plate was sealed with sealing plate film, placed in a flat plate centrifuge, and exposed to a relative centrifugal force (RCF) of 500×*g*. In the PCR instrument, the denaturation process was conducted at 98 °C for 5 min, without heating the hot cover and at the end of the procedure, the 96-well plate was placed immediately on iced water. Once cooled, the well plate was placed in a flat plate centrifuge and exposed to an RCF of 2000×*g*. Finally, the samples were analyzed using an ABI3730XL sequence analyzer (ABI Corporation, Foster City, CA, USA).

Data analysis

The results of the peak patterns produced by the sequence analyzer were analyzed by GeneMapper v5.0 (Hayden et al. 2008). The data were counted by peak feature and fragment size of the corresponding peaks.

PowerMarker ver. 3.25 (Liu and Muse 2005) software was used to calculate allele frequency, genotype number, allele number, gene diversity, heterozygosity, and the polymorphism information content (PIC).

The population structure of the 63 cherry cultivars was analyzed using 10 SSR primer pairs by the model-based software Structure v2.3.4 (Pritchard et al. 2010). The model choice criterion implemented in structure to detect the true *K* is an estimate of the posterior probability of the data for a given *K*, Pr(*X*|*K*) (Pritchard et al. 2000). This value, called 'Ln*P*(*D*)' in structure output, is obtained by first computing the log likelihood of the data at each step of the Markov Chain Monte Carlo (MCMC) replication. The number of populations, *K*, was set a priori from 1 to 10, and calculated in 20 independent simulations. For each simulation, with the selection of admixture and related frequency models, 10,000 iterations were performed before a burn-in length of 10,000 MCMC replications (Pritchard et al. 2000; Falush et al. 2007). The other parameters were set to default values. The optimal *K* value was determined by the posterior probability [Ln*P*(*D*)] and an ad hoc statistic ΔK based on the rate of change in [Ln*P*(*D*)] between successive *K* values (Evanno et al. 2005). The cultivars were assigned to corresponding populations based on *K* values.

Principal component analysis (PCA) was operated by NTsys2.10. The unweighted pair-group method with arithmetic average (UPGMA) cluster analysis was performed based on Nei's genetic distance matrix with MEGA6 (Tamura et al. 2013). Both were used to identify the relationship among populations and species.

Results

Using capillary electrophoresis with fluorescent-labeled SSR markers, GeneMapper could read the exact size of target fragments according to the location of the target peak compared with the internal standard (GeneScan™ 500 LIZ®) in the same capillary lane. The target fragments were recorded based on the highest peak position, with a single peak representing a homozygote, while a double peak represented a heterozygote (Fig. 1). Thus, compared with the conventional native PAGE, capillary electrophoresis with fluorescent primers read the fragment size exactly, and therefore, the test data were more accurate.

A total of 146 polymorphic bands obtained from 10 pairs of SSR primers (Table 1) were used for the DNA fingerprinting, where the pairs marked with A, B, C, D, E, F, G, H, I, and J were grouped from front to back and separated by "–". The number of alleles amplified by 10 SSR primer pairs, ranged from 10 (EPDCU5060) to 20 (EPDCU5183) and the genotype number ranged from 12 (EPPCU4092) to 26 (EPDCU5183), with an average of 18. The mean values of gene diversity, heterozygosity, and PIC were 0.7549, 0.5952, and 0.7355, respectively (Table 2).

In each primer, allele sizes from small to large were assigned a sequence number starting at 01 (Table 3) and these strings were arranged in a digital fingerprint of the 63 tested cherry cultivars that were all uniquely identified.

Ln$P(D)$ was referred to as $L(K)$ afterwards. The distribution of $L(K)$ did not show a clear mode for the true K, but an ad hoc quantity based on the second-order rate of

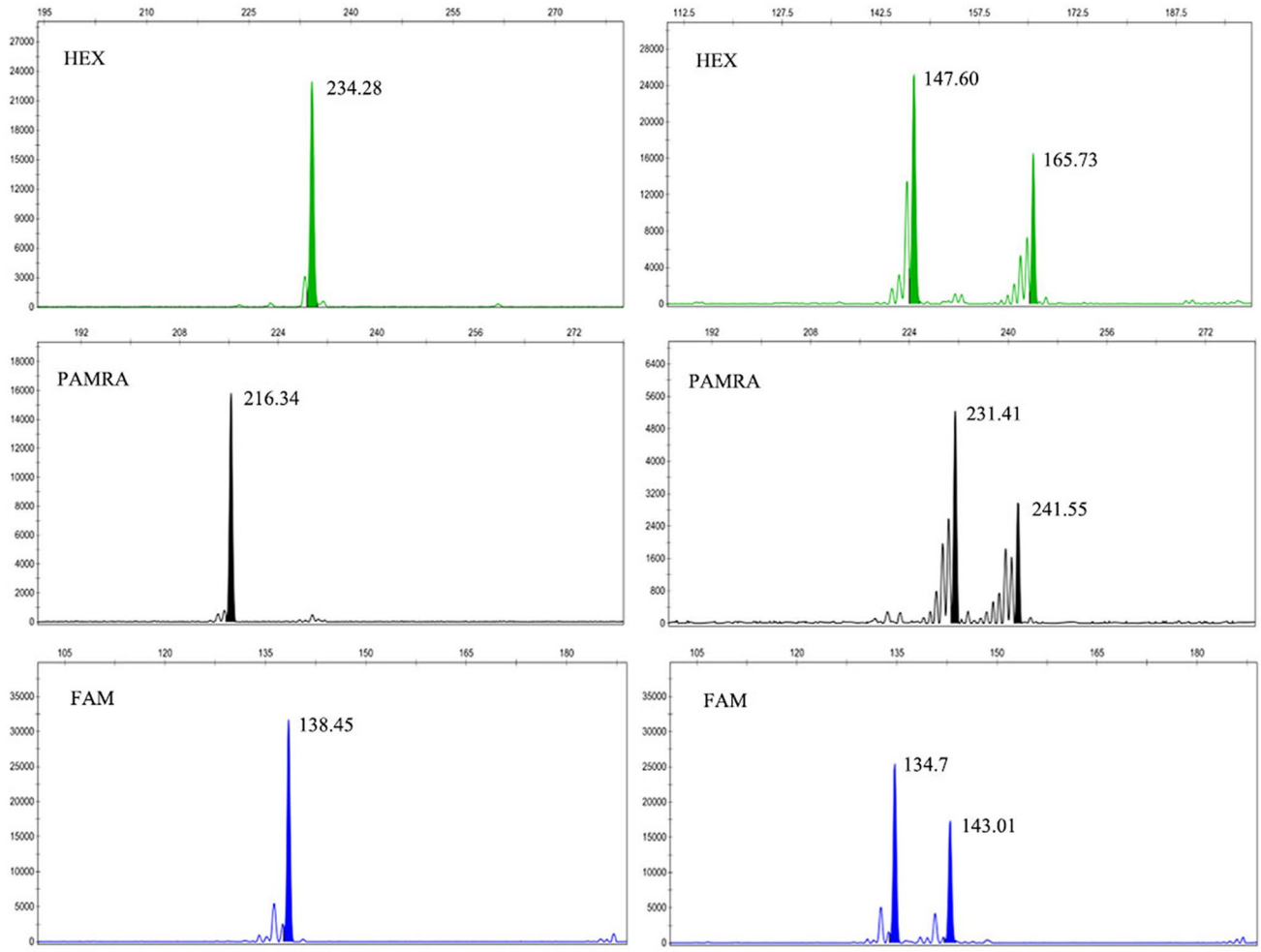

Fig. 1 Capillary electrophoresis peaks detected using a sequence analyzer, with the fluorescent dyes FAM, HEX, and PAMRA. The *x*-axis represents fragment size of an amplified microsatellite and the *y*-axis represents the fluorescence intensity of amplified products

Table 2 Results of the ten simple sequence repeat (SSR) primers

Primer	Allele frequency	Genotype number	Allele number	Gene diversity	Heterozygosity	PIC
SC2	0.3254	18	14	0.8181	0.5238	0.7982
SC3	0.2619	22	18	0.8486	0.9048	0.8337
BPPCT013	0.3492	17	15	0.7778	0.9683	0.7507
BPPCT026	0.4444	21	13	0.7572	0.6508	0.7381
EMPaS01	0.3889	16	13	0.7458	0.5556	0.7104
EMPaS02	0.5000	18	12	0.7168	0.5714	0.6993
CPSCT038	0.2778	20	18	0.8337	0.4127	0.8150
EPDCU5060	0.3968	15	10	0.7720	0.5556	0.7464
EPDCU5183	0.2778	26	20	0.8782	0.4286	0.8697
EPPCU4092	0.7698	12	13	0.4011	0.3810	0.3937
Mean	0.3992	18	14	0.7549	0.5952	0.7355

change of the likelihood function with respect to K (ΔK) did show a clear peak at the true value of K (Evanno et al., 2005). $L'(K) = L(K) - L(K-1)$, $L''(K) = |L'(K+1) - L'(K)|$. Finally, ΔK was estimated as the mean of the absolute values of $L''(K)$ averaged over 20 runs divided by the standard deviation of $L(K)$, $\Delta K = m(|L''(K)|)/s[L(K)]$ (Pritchard et al., 2000). According to this formula, ΔK reached a peak at $K = 3$ (Fig. 2b). The cultivars were classified into three populations, P1, P2, and P3 (Fig. 3a). From Bayes' rule, when several values of K give similar estimates of Ln$P(D)$, the smallest is often "correct" (Pritchard et al. 2010). Based on this principle, that $K = 7$ qualified (Fig. 2a). Accordingly, 63 cultivars were divided into seven populations, P1, P2, P3, P4, P5, P6, and P7 (Fig. 3d).

PCA discriminated the 63 cultivars in two dimensions and grouped all the cultivars into three clusters (P1, P2, and P3) labeled with three ellipses (Fig. 3b), which was consistent with the result obtained for the population structure, as shown in Fig. 3a. The UPGMA cluster analysis separated all the cultivars; however, there was intermixing of cultivars from P1, P2, and P3 in the phylogenetic tree (Fig. 3c).

Discussion

In this study, plant materials covering 63 cherry genotypes were characterized with 10 SSR markers, which have been reported to be polymorphic in a wide sample of *Prunus* species, and their cross-species transferability has been demonstrated in the previous studies (Dirlewanger et al. 2002; Vaughan and Russell 2004; Ai et al. 2007). The 10 SSR primer pairs amplified a total of 146 alleles, ranging from 12 to 20 alleles per primer pair, with an average of 14 alleles per primer pair. In our study, the SC2 and SC3 primer pairs created 14 and 18 alleles per locus, with 18 and 22 genotypes, respectively; however, this contrasts with Ai et al. (2007) who found that these primers produced only 4 and 2 alleles per locus, respectively. Dirlewanger et al. (2002) reported that allele number and heterozygosity of BPPCT026 in cherry cultivars were 6 and 0.67, respectively; this heterozygosity was consistent with our results (0.6508), but we observed almost twice the amount of alleles (13) for BPPCT026. It was assumed that some close bands were easily avoided when the PAGE method was applied, and capillary electrophoresis with fluorescent primers provided more precise genotyping. For the EMPaS01 and EMPaS02 primer pairs, the number of alleles was 13 and 12, while the heterozygosity was 0.5556 and 0.5714, respectively, which differed from the results of Vaughan and Russell (2004) where the number of alleles was 4 and 5, and heterozygosity was 0.75 and 0.81, respectively. This conflict in results could be due to differences in number of genotypes that were used and level of polyploidy (Turet-Sayar et al. 2012). For EPPCU4092, allele frequency as high (0.7698), but genotype number (12), allele number (13), gene diversity (0.4011), heterozygosity (0.3810), and PIC (0.3937) were generally low. In contrast, allele frequency (0.2778) of the EPDCU5183 primer pairs was low and heterozygosity (0.4286) was relatively low, while genotype number (26), allele number (20), gene diversity (0.8782), and PIC (0.8697) were all high. PIC is an important index to assess the fitness of SSR primers, reflecting the amount of polymorphism information that primers can produce (Song et al. 2016). In our study, the average gene diversity (0.7549) in accordance with the average PIC value (0.7355) was high, which was consistent with the result of Lacis et al. (2009), but higher than that of Sharma et al. (2015). It has been suggested that the chances of scoring undesired alleles are minimized when capillary electrophoresis with fluorescent primers is used (Pan et al. 2003).

According to Evanno et al. (2005), ΔK reaches a peak at $K = 3$; thus, the 63 cultivars were classified into three populations that were consistent with the PCA analysis. We found similarity with the UPGMA analysis, but there was

Table 3 DNA fingerprinting of 63 cherry accessions by simple sequence repeat (SSR) markers

Accessions	DNA fingerprinting
Katalin	A07A08-B05B14-C05C14-D10D10-E05E05-F06F06-G04G04-H06H08-I06I06-J08J08
Brooks	A14A14-B14B14-C14C14-D10D12-E03E08-F06F09-G05G13-H06H06-I04I04-J08J08
Tieton	A05A05-B14B18-C14C15-D10D10-E05E08-F05F06-G04G04-H04H06-I15I16-J08J10
Kpynhonnouhax	A06A06-B05B14-C05C14-D10D13-E05E05-F06F08-G04G06-H04H10-I05I05-J08J08
HongDeng	A06A06-B05B14-C05C14-D11D13-E06E10-F06F08-G05G05-H06H10-I15I16-J08J08
Qinying III	A07A08-B05B14-C05C14-D10D10-E05E08-F06F06-G04G04-H06H06-I06I14-J08J08
Summit	A06A06-B05B14-C05C14-D11D11-E05E05-F06F06-G04G04-H05H09-I04I04-J08J08
Lapins	A08A08-B14B18-C14C15-D11D11-E05E05-F06F06-G12G12-H05H05-I04I04-J08J08
Rainier	A06A06-B14B18-C14C15-D10D10-E08E08-F06F06-G04G04-H06H06-I04I04-J08J08
Chelan	A06A08-B05B14-C05C14-D11D11-E05E08-F06F06-G13G13-H06H10-I04I04-J08J08
Starkrimson	A06A08-B05B14-C05C14-D10D10-E05E08-F06F06-G05G13-H06H08-I14I14-J08J08
Van	A05A05-B04B13-C05C14-D10D12-E08E08-F06F06-G04G04-H06H06-I04I04-J08J08
Qinying I	A05A05-B05B14-C05C14-D10D10-E08E10-F05F08-G05G05-H04H06-I04I04-J08J08
Sunburst	A06A08-B14B18-C14C15-D11D11-E05E05-F06F06-G12G12-H06H06-I05I05-J08J08
Hungary-A	A06A08-B05B14-C05C14-D10D10-E05E08-F06F06-G04G12-H06H08-I14I14-J08J08
Kocmhqecka	A06A08-B05B14-C04C13-D09D09-E05E08-F06F06-G04G12-H05H07-I14I14-J08J08
Burlat	A05A05-B05B05-C05C14-D10D10-E08E10-F05F08-G05G05-H04H06-I04I16-J08J08
Russia VIII	A06A08-B04B13-C04C13-D10D10-E01E03-F06F06-G05G05-H05H09-I15I16-J01J03
Sylvia-1	A05A05-B05B14-C05C14-D10D12-E08E08-F06F06-G04G04-H06H06-I07I07-J08J13
Santina-1	A06A09-B05B14-C05C14-D10D11-E05E05-F06F06-G05G13-H06H06-I04I04-J08J08
S7	A07A09-B05B10-C05C10-D03D10-E05E08-F05F06-G03G05-H04H06-I06I09-J08J08
S9	A07A09-B05B14-C05C14-D03D10-E05E08-F05F06-G05G13-H04H06-I06I09-J08J10
Maoyingtao	A04A06-B02B05-C02C05-D10D10-E03E03-F09F09-G11G18-H03H06-I04I14-J08J09
Meili	A07A10-B03B10-C03C14-D01D10-E05E08-F06F10-G05G07-H06H08-I04I06-J08J08
Aode	A06A06-B03B10-C03C14-D02D10-E05E08-F06F07-G04G04-H06H08-I04I06-J08J08
Qinling wild cherry	A12A12-B03B04-C03C04-D05D07-E04E06-F11F12-G10G15-H02H03-I03I06-J08J09
Meilei	A07A13-B03B09-C05C09-D03D13-E05E05-F05F09-G07G07-H04H06-I05I05-J08J08
Aojie	A07A09-B05B10-C05C14-D03D11-E05E08-F05F06-G03G05-H04H06-I04I08-J08J08
Mahaleb-Y	A01A05-B08B08-C06C08-D06D06-E02E13-F04F04-G15G15-H03H03-I08I10-J03J04
Mahaleb-R	A01A05-B08B11-C08C11-D06D08-E02E13-F04F04-G15G15-H03H03-I09I12-J08J11
Ouli-R	A07A07-B02B06-C02C06-D10D10-E11E11-F01F02-G16G16-H01H01-I13I13-J08J12
Ouli-Y	A07A07-B01B06-C01C06-D04D10-E12E12-F01F01-G16G16-H01H01-I01I01-J08J12
Carmen-1	A06A06-B05B14-C05C14-D10D13-E05E05-F06F08-G04G06-H04H10-I05I05-J08J13
Spur type Mahaleb	A01A06-B08B12-C08C12-D07D10-E02E02-F04F07-G15G15-H03H03-I12I12-J02J03
Mahaleb CDR-2	A02A05-B08B11-C08C11-D07D08-E02E02-F04F11-G15G15-H03H03-I13I13-J08J11
Valerij Cskalov	A07A07-B04B13-C05C14-D10D13-E05E05-F06F08-G05G07-H04H10-I04I04-J08J08
Rita	A05A05-B14B18-C14C15-D10D10-E05E10-F05F06-G04G04-H06H06-I15I16-J08J08
Regina	A06A09-B05B14-C05C14-D10D10-E05E08-F06F08-G05G13-H04H08-I16I19-J08J08
Skeena	A06A09-B13B17-C14C15-D10D11-E05E08-F06F10-G05G13-H06H06-I20I20-J08J08
Sweetheart	A06A06-B14B18-C14C15-D11D11-E05E08-F06F06-G05G13-H06H06-I04I04-J08J08
Techlovan	A06A06-B05B14-C05C14-D11D13-E05E08-F06F10-G05G05-H06H06-I05I05-J08J08
Kordia	A05A05-B05B14-C05C14-D10D13-E05E08-F07F10-G04G04-H06H08-I15I17-J08J13
Carmen-2	A07A07-B05B14-C05C14-D10D11-E05E05-F06F06-G05G05-H06H08-I04I04-J08J13
Sylvia-2	A06A09-B04B13-C05C14-D10D13-E05E08-F06F08-G05G13-H04H08-I15I16-J08J08
Carina	A06A06-B04B14-C05C14-D10D13-E05E08-F06F10-G05G05-H06H08-I07I07-J08J08
Snelders	A06A09-B05B14-C05C14-D10D13-E05E08-F06F08-G05G13-H04H08-I16I19-J08J09
13-33	A06A06-B04B13-C05C14-D10D12-E05E08-F06F08-G05G05-H04H10-I04I04-J08J08
8-129	A07A07-B04B13-C05C14-D10D12-E08E10-F06F08-G07G13-H06H06-I04I04-J08J08
Jiahong	A06A06-B05B14-C05C14-D10D11-E05E08-F06F08-G05G05-H06H06-I05I05-J06J08
Shamidou	A06A06-B05B14-C05C14-D11D11-E05E05-F06F06-G05G05-H06H10-I05I05-J08J08

Table 3 (continued)

Accessions	DNA fingerprinting
2-82	A07A07-B05B14-C05C14-D10D12-E05E08-F06F06-G05G05-H06H10-I07I07-J08J08
Mingzhu	A06A06-B14B14-C14C14-D10D12-E06E08-F06F06-G04G04-H10H10-I18I18-J08J08
Santina-2	A06A06-B04B13-C05C14-D10D12-E08E08-F06F06-G04G04-H06H06-I05I05-J08J08
Ouli	A10A10-B01B07-C01C07-D02D10-E09E09-F02F03-G08G09-H01H01-I04I19-J07J07
Daqingye-1	A04A06-B02B05-C02C05-D10D10-E03E03-F09F09-G01G17-H03H06-I04I14-J08J09
Mahaleb 2-10	A02A05-B08B16-C06C08-D04D08-E02E02-F04F07-G15G15-H03H03-I13I13-J05J08
Mahaleb 2-70	A02A05-B13B13-C11C14-D06D07-E02E02-F07F07-G02G14-H03H03-I08I13-J08J11
Mahaleb 1-162	A02A05-B08B11-C06C08-D07D08-E02E02-F04F07-G15G15-H03H03-I09I12-J08J11
Mahaleb 1-199	A02A05-B08B08-C06C08-D06D08-E02E02-F07F07-G15G15-H03H03-I09I12-J08J11
Mahaleb CDR-1	A02A05-B08B11-C08C11-D08D10-E02E02-F04F11-G15G15-H03H03-I09I12-J08J11
ZY-1	A07A13-B05B09-C05C09-D03D13-E05E05-F05F09-G07G07-H04H06-I04I04-J08J08
Gisela V	A04A11-B02B05-C03C05-D01D10-E05E07-F06F09-G03G07-H03H04-I02I06-J08J08
Gisela VI	A04A11-B02B03-C03C05-D01D10-E05E07-F06F09-G03G07-H03H04-I02I04-J08J08

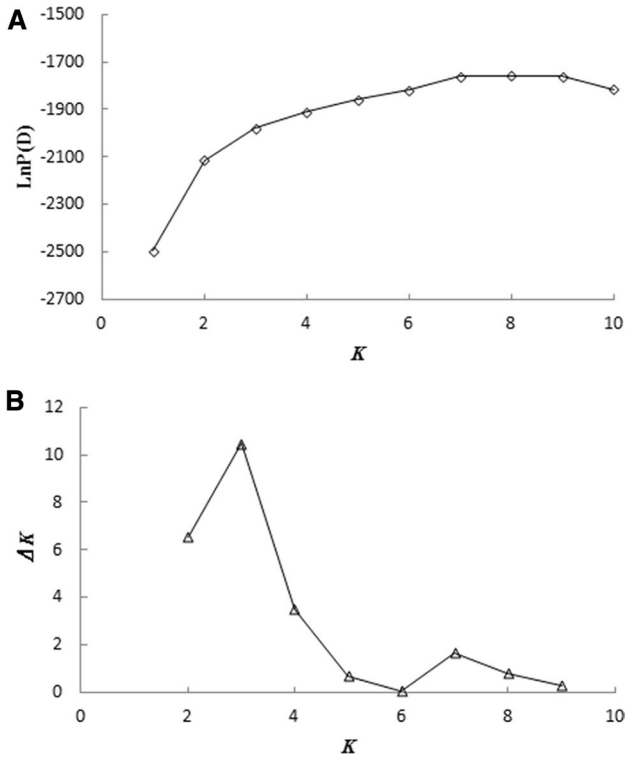

Fig. 2 Estimation of Ln$P(D)$ and ΔK in the 63 cherry accessions. K is the number of populations

some within-population intermixing of cultivars and may be a result of anthropogenic-mediated gene exchange between different regions. The results of our analyses broadly compare well with the species identified among the 63 cultivars (Table S1). Using Bayes' rule, which states that the smallest value of K is often correct when similar estimates of Ln$P(D)$ are given (Pritchard et al. 2010), the 63 cultivars were further subdivided into seven populations. In terms of the genetic structure, various colors were intertwined, which may be due to hybridization or transferability (Gasic et al. 2009).

Prunus humilis was chosen as the reference cultivar, and when $K = 3$, *P. humilis* plants and other rootstocks (except *P. mahaleb*) were bracketed, while when $K = 7$, the 63 cultivars were further subdivided, and the three *P. humilis* plants were grouped together from the other rootstocks, suggesting that the classifications were reliable. When $K = 3$, the 63 cultivars were classified into three groups, and *P. avium* was completely assigned to a large group P2, but 'Russia VIII' was not. After being further subdivided, *P. avium* was mainly classified into two groups P2 and P6, but again, 'Russia VIII' was not. Originating from Russia, 'Russia VIII' is a hybrid progeny of 'Iuliia' and 'Valerij Cskalov', and was noteworthy (as indicated by an asterisk in Fig. 3b), because its K value differed from other varieties, but it was always classified together with rootstocks. In the UPGMA cluster analysis, 'Russia VIII' and 'Burlat' were in the same clade. No studies of the molecular identification of 'Russia VIII' have been reported.

Both 'Qinying I', a natural mutation of 'Burlat', and 'Shamidou', a natural mutation of 'Summit', could be differentiated from their parent cultivars. However, our results are inconsistent with the findings of Wünsch and Hormaza (2002) who reported that 'Burlat C1' (a compact mutation of 'Burlat') or 'Van Spur' and 'Early Van Compact' (both mutations of 'Van') could not be differentiated from their parent cultivars. This further demonstrates that for the genetic fingerprinting of cherry, the utility of capillary electrophoresis with fluorescent SSR primers was more accurate than PAGE.

In sweet cherry varieties, both synonyms and homonyms are abundant (Turet-Sayar et al. 2012) and cause large losses in cherry production. Some test varieties

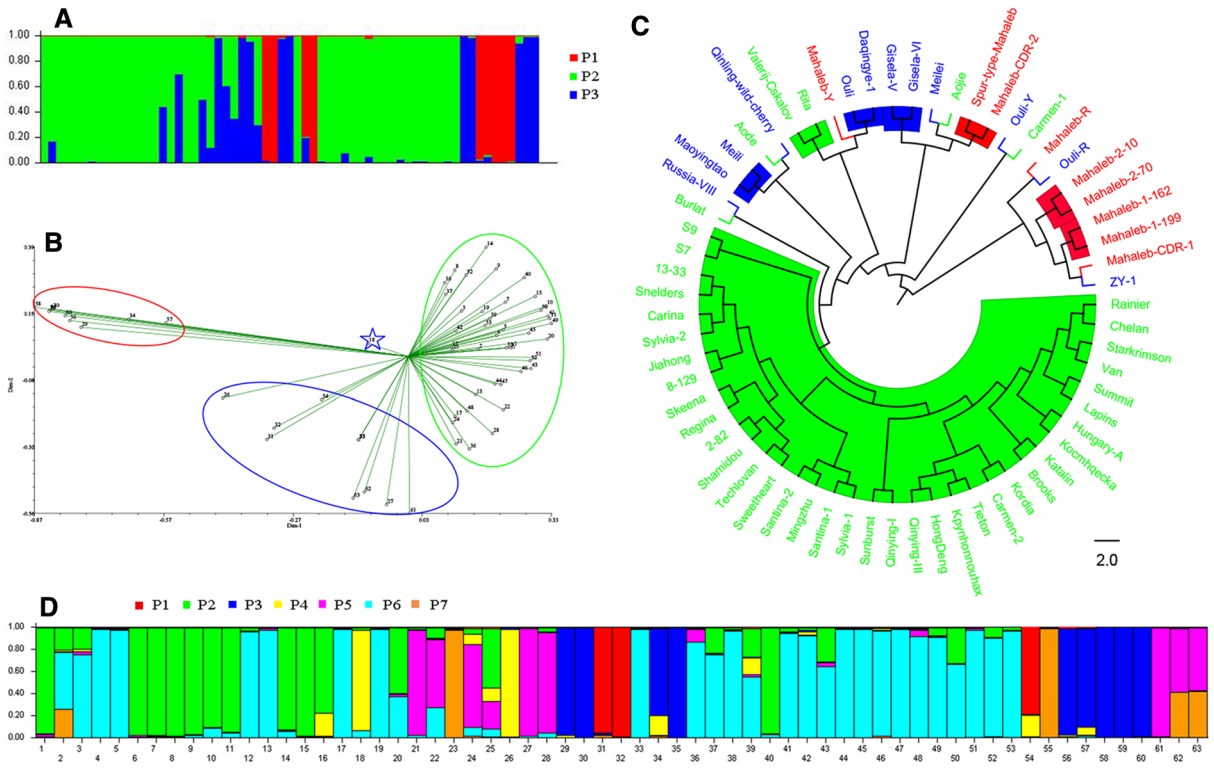

Fig. 3 Characterization of genetic structure in the 63 cherry accessions. **a**, **d** population structure. Colored bars represent accessions grouped into the corresponding inferred population. **b** Principal component analysis (PCA) plot of 63 accessions encircled by different colors corresponding to the model-based structure. **c** Phylogenic tree of 63 accessions based on unweighted pair-group method with arithmetic average (UPGMA). Colored clades correspond to inferred populations

have the same names in production, but the appearance of their fruit is very different, such as for 'Sylvia-1' and 'Sylvia-2', 'Santina-1' and 'Santina-2', and 'Carmen-1' and 'Carmen-2'. From their appearance, i.e., fruit shape, size, and color (Yamamoto et al. 2015), 'Sylvia-1', 'Santina-1', and 'Carmen-2' could be identified based on their official description. Their source of introduction was also clear, and therefore, it was confirmed that 'Sylvia-1', 'Santina-1', and 'Carmen-2' were the correct varieties. To further determine whether 'Sylvia-1' and 'Sylvia-2', 'Santina-1' and 'Santina-2', and 'Carmen-1' and 'Carmen-2' were the same varieties, we analyzed them at the molecular level. Since they were collected from different sampling sites, it suggests that they may have different origins from the other Canadian or Hungarian cultivars in our experiment. This may also be due to cross-regional migration or breeding and cultivation in different regions, indicating the complex nature of the history of cherry domestication (Yang et al. 2015). The results of the UPGMA cluster analysis indicated that there was a significant genetic difference between cultivars, and therefore, those possibilities above were ruled out. In this study, no other tested varieties had alleles that were identical to 'Sylvia-2', 'Santina-2', or 'Carmen-1'. There is a need to collect more cherry cultivar resources and detect and analyze them to further confirm the classification 'Sylvia-2', 'Santina-2', and 'Carmen-1'.

Most rootstocks are seed propagated (sexual propagation). *P. pseudocerasus* and *P. tomentosa* have received much attention recently because of their high resistance and vigorous growth, and many studies of these species have selected varieties to provide superior germplasm resources for breeding. The seeds used for seedling stock production are typically produced in fruit processing plants or are obtained from wild-grown trees (Ercisli et al. 2006; Mratinic et al. 2012). In recent years, many rootstocks have been asexually reproduced by cutting and layering. In this study, the significance of SSR molecular identification for rootstocks was that superior rootstocks reproduced by seed were selected to asexually propagate, and were then widely promoted in production.

Prunus mahaleb, a cherry breeding resource plant, was introduced from Hungary to China by the Northwest Agriculture and Forestry University, because of its dwarfing and strong resistance to crown gall and salt, among many other excellent biological characteristics (Hrotkó 2016). As a cosmopolitan sweet cherry rootstock, *P. mahaleb* has become one of the main sweet cherry rootstocks in northwest China,

and in our study, whether $K = 3$ or 7, the mahaleb series (comprising cultivars 'Mahaleb-Y', 'Mahaleb-R', 'Spur type Mahaleb', 'Mahaleb CDR-2', 'Mahaleb 2-10', 'Mahaleb 2-70', 'Mahaleb 1-162', 'Mahaleb 1-199', and 'Mahaleb CDR-1') was classified as a single category. The mahaleb series was selected from seedlings with excellent characteristics, and is widely propagated by cutting in cherry production.

The genetic relationship reflects the difference in the genetic background between cultivars; thus, it is possible to breed elite varieties through the selection of genetically distant cultivars as hybrid parents (Yang et al. 2015). The genetic relationship and the genetic distance were conducive to grafting optimal varieties and rootstocks to become superior varieties.

Conclusion

With the utility of fluorescent capillary electrophoresis for genetic fingerprinting in cherry, SSRs revealed a high mean number of alleles per locus as well as high heterozygosity, gene diversity, and PIC values. The cultivars were divided into three different populations. After subdivision, they were grouped into seven populations. Some cherry varieties that are often confused in production were distinguished. The establishment of DNA fingerprinting for cherry cultivars planted in Shaanxi province, China, will be useful in cherry cultivar selection, planting, and production.

Acknowledgements This study was supported by the international science and technology cooperation project between China and Hungary: "Cooperative research to determine high quality and resistance in the selection of new cherry varieties, and the development of intensive culture and processing techniques (2016YFE0130900)". The authors thank Dr. Wei Liang for her guidance regarding data analysis and Dr. Jianlong Liu for his editing of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest in the publication.

References

Ai CX, Yu XM, Zhang LS, Wei HR, Xin L, Yuan KJ, Jin SN, Sun QR, Liu QZ (2007) Development of SSR markers in sweet cherry using selectively amplified microsatellite (in Chinese). Acta Hortic Sinic 34(2):311–316. https://doi.org/10.16420/j.issn.0513-353x.2007.02.009

Chandra A, Grisham MP, Pan YB (2014) Allelic divergence and cultivar-specific SSR alleles revealed by capillary electrophoresis using fluorescence-labeled SSR markers in sugarcane. Genome 57(6):363–372. https://doi.org/10.1139/gen-2014-0072

Cipriani G, Lot G, Huang WG, Marrazzo MT, Peterlunger E, Testolin R (1999) Ac/gt and ag/ct microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterisation and cross-species amplification in prunus. Theor Appl Genet 99(1–2):65–72. https://doi.org/10.1007/s001220051209

Dirlewanger E, Cosson P, Tavaud M, Aranzana J, Poizat C, Zanetto A, Arus P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theor Appl Genet 105(1):127–138. https://doi.org/10.1007/s00122-002-0867-7

Downey SL, Iezzoni AF (2000) Polymorphic DNA markers in black cherry (*Prunus serotina*) are identified using sequences from sweet cherry, peach, and sour cherry. J Am Soc Hortic Sci 125(1):76–80

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

Ercisli S, Esitken A, Orhan E, Ozdemir O (2006) Rootstocks used for temperate fruit trees in Turkey: an overview. Sodininkystė Ir DaržIninkystė 25(3):27–33

Ercisli S, Agar G, Yildirim N, Duralija B, Vokurka A, Karlidag H (2011) Genetic diversity in wild sweet cherries (*Prunus avium*) in Turkey revealed by SSR markers. Genet Mol Res 10(2):1211–1219. https://doi.org/10.4238/vol10-2gmr1196

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14(8):2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7(4):574–578. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Gasic K, Han Y, Kertbundit S, Shulaev V, Iezzoni AF, Stover EW, Bell RL, Wisniewski ME, Korban SS (2009) Characteristics and transferability of new apple EST-derived SSRs to other Rosaceae species. Mol Breed 23(3):397–411. https://doi.org/10.1007/s11032-008-9243-x

Guarino C, Santoro S, Simone LD, Cipriani G (2010) Molecular characterisation of ancient *Prunus avium* L. germplasm using sweet cherry SSR markers. J Hortic Sci Biotech 85(4):295–304

Hayden MJ, Nguyen TM, Waterman A, Mcmichael GL, Chalmers KJ (2008) Application of multiplex-ready PCR for fluorescence-based SSR genotyping in barley and wheat. Mol Breed 21(3):271–281. https://doi.org/10.1007/s11032-007-9127-5

Hormaza JI (1999) Early selection in cherry combining RAPDs with embryo culture. Sci Hortic (Amsterdam) 79(1–2):121–126. https://doi.org/10.1016/S0304-4238(98)00204-0

Hrotkó K (2016) Potentials in *Prunus mahaleb* L. for cherry rootstock breeding. Sci Hortic (Amsterdam) 205(410):70–78. https://doi.org/10.1016/j.scienta.2016.04.015

Lacis G, Rashal I, Ruisa S, Trajkovski V, Iezzoni AF (2009) Assessment of genetic diversity of Latvian and Swedish sweet cherry (*Prunus avium* L.) genetic resources collections by using SSR (microsatellite) markers. Sci Hortic (Amsterdam) 121(4):451–457. https://doi.org/10.1016/j.scienta.2009.03.016

Liang J, Pan YB, Li YR, Fang FX (2010) Assessment of genetic diversity in *Saccharum* using SSR markers and capillary electrophoresis. Guihaia 30(1):106–111

Liu K, Muse SV (2005) Powermarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21(9):2128–2129. https://doi.org/10.1093/bioinformatics/bti282

Mnejja M, Garcia-Mas J, Howad W, Badenes ML, Arús P (2004) Simple-sequence repeat (SSR) markers of Japanese plum (*Prunus salicina Lindl.*) are highly polymorphic and transferable to peach and almond. Mol Ecol Notes 4(2):163–166. http://doi.org/10.1111/j.1471-8286.2004.00603.x

Mratinic E, Fotiric Aksic M, Jovkovic R (2012) Analysis of wild sweet cherry (*Prunus avium* L.) germplasm diversity in

south-east Serbia. Genetik 44(2):259–268. https://doi.org/10.2298/GENSR1202259M

Ognjanov V, Ljubojević M, Ninićtodorović J, Bošnjaković D, Barać G, Čukanović J, Mladenovića E (2015) Morphometric diversity in dwarf sour cherry germplasm in Serbia. J Hortic Sci Biotech 87(2):117–122. https://doi.org/10.1080/14620316.2012.11512841

Pan YB, Cordeiro GM, Richard EP, Henry RJ (2003) Molecular genotyping of sugarcane clones with microsatellite DNA markers. Maydica 48(4):319–329

Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2):945–959

Pritchard JK, Wen X, Falush D (2010) Documentation for structure software: version 2.3. J Pediatr Surg 41(10):55–63

Sharma K, Xuan H, Sedlák P (2015) Assessment of genetic diversity of Czech sweet cherry cultivars using microsatellite markers. Biochem Syst Ecol 63:6–12. https://doi.org/10.1016/j.bse.2015.09.013

Smith D (1994) Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome 37(6):977–983

Song X, Zhang C, Li Y, Feng S, Yang Q, Huang S (2016) SSR analysis of genetic diversity among 192 diploid potato cultivars. Hortic Plant J 2(3):163–171. https://doi.org/10.1016/j.hpj.2016.08.006

Struss D, Boritzki M, Glozer K, Southwick SM (2001) Detection of genetic diversity among populations of sweet cherry (*Prunus avium* L.) by AFLPs. J Hortic Sci Biotech 76(3):362–367. https://doi.org/10.1080/14620316.2001.11511378

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) Mega6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12):2725–2729. https://doi.org/10.1093/molbev/mst197

Turet-Sayar M, Turkec A, Demir T (2012) Identification of sweet cherry cultivars (*Prunus avium* L.) and analysis of their genetic relationship using microsatellite DNA fingerprinting. J Agric Sci 4(8):134–140. https://doi.org/10.5539/jas.v4n8p134

Turkoglu Z, Bïlgener S, Ercïslï S, Bakïr M, Koc A, Akbulut M, Gercekcioglu R, Gunes M, Esitken A (2010) Simple sequence repeat-based assessment of genetic relationships among *Prunus* rootstocks. Genet Mol Res 9(4):2156–2165. https://doi.org/10.4238/vol9-4gmr957

Turkoglu Z, Koc A, Ercisli S (2012) Genetic relationships among *Prunus* rootstocks for sweet cherry (*Prunus avium* L.) cultivars. Plant Genet Resour 10(2):101–107. https://doi.org/10.1017/S147926211200007X

Vaughan SP, Russell K (2004) Characterization of novel microsatellites and development of multiplex PCR for large-scale population studies in wild cherry, *Prunus avium*. Mol Ecol Notes 4(3):429–431. https://doi.org/10.1111/j.1471-8286.2004.00673.x

Webster AD (1996) The taxonomic classification of sweet and sour cherries and a brief history of their cultivation. In: Webster AD, Looney NE (eds) Cherries: crop physiology, production and uses. CAB International, Wallingford, Oxon, pp 3–24

Wünsch A, Hormaza JI (2002) Molecular characterisation of sweet cherry (*Prunus avium* L.) genotypes using peach [*Prunus persica* (L.) Batsch] SSR sequences. Heredity 89:56–63. https://doi.org/10.1038/sj.hdy.6800101

Wünsch A, Hormaza JI (2004) Molecular evaluation of genetic diversity and *S*-allele composition of local Spanish sweet cherry (*Prunus avium* L.) cultivars. Genet Resour Crop Evol 51(6):635–641. https://doi.org/10.1023/B:GRES.0000024649.06681.43

Yamamoto K, Ninomiya S, Kimura Y, Hashimoto A, Yoshioka Y, Kameoka T (2015) Strawberry cultivar identification and quality evaluation on the basis of multiple fruit appearance features. Comput Electron Agric 110:233–240. https://doi.org/10.1016/j.compag.2014.11.018

Yang XS, Wen-Jin SU, Wang LJ, Lei J, Chai SS, Liu QC (2015) Molecular diversity and genetic structure of 380 sweet potato accessions as revealed by SSR markers. J Integr Agric 14(4):633–641. https://doi.org/10.1016/S2095-3119(14)60794-2