Scientia Horticulturae 237 (2018) 28–35

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Species identification, genetic diversity and population structure of sweet cherry commercial cultivars assessed by SSRs and the gametophytic self-incompatibility locus

Congli Liu, Xiliang Qi, Lulu Song, Yuhong Li, Ming Li*

Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou, 450009, China

ARTICLE INFO

Keywords:
Prunus avium L.
Genetic diversity
Population structure
SSR markers
S-genotype

ABSTRACT

Assessing the genetic diversity and population structure of sweet cherry is essential for the efficient preservation of germplasms and exploitation of genetic resources. In this study, 10 highly polymorphic simple sequence repeat markers, combined with the gametophytic self-incompatibility locus, were used to evaluate the genetic diversity levels, phylogenetic relationships and population structure among 95 sweet cherry commercial cultivars. A total of 67 alleles were detected, with a mean of 6.09 alleles per locus. In total, five out of 10 simple sequence repeat primers chose according to their polymorphism information content values were able to completely differentiate 95 accessions based on *S*-genotype information. The un-weighted pair group method with arithmetic mean dendrogram and model-based structural analysis clearly divided the accessions into three distinct populations (Pop 1, 2 and 3) that were consistent with their pedigree-based relationships. Pop 1 and Pop 2 contained 82% of the analyzed accessions and some admixed accessions dispersed between the populations. Pop 3 mainly included sweet cherry cultivars from the Ukraine and their descendants. This study provides an efficient method for species identification and demonstrates its potential application in analyzing genetic fingerprints and diversity levels, as well as verifying parentage, in sweet cherry.

1. Introduction

Sweet cherry (*Prunus avium* L.) is an economically important diploid species in the Rosaceae family. The species originated from the area surrounding the Caspian and Black Seas (Webster et al., 1996), and is now widespread through most of the world's temperate regions. China is becoming an important producer of sweet cherry and had a total production area of more than 134,800 ha in 2013 (Huang et al., 2014).

Sweet cherries are self-incompatible determined by the Gametophytic Self-Incompatibility system (De Nettancourt, 2001) and cross-breeding is the most effective way of developing new cultivars. It appears that modern breeding has decreased the levels of genetic diversity in sweet cherry cultivars, as seen in grape (Emanuelli et al., 2013), peach (Li et al., 2013), apple (Urrestarazu et al., 2012) and other fruit trees (Bourguiba et al., 2012). The genetic diversity has been estimated in 141 landraces and 66 modern varieties using 26 simple sequence repeats (SSRs), which showed that 40% of the alleles present in wild cherry have been lost in modern cherry varieties, because only small numbers of individuals are used as parents in cherry breeding programs (Mariette et al., 2010). A comparative assessment of chloroplast DNA diversity levels in the wild and sweet cherry showed that the level was greater in wild *P. avium* than in sweet cherry cultivars (Panda et al., 2003). Genetic diversity in sweet cherry has been minimized owing to repeated use of a few common varieties as parents in breeding programs (Choi and Kappel, 2004; Mariette et al., 2010; Wunsch and Hormaza, 2002). For example, more than 20 commercial cultivars, including well-known 'Rainier', 'Lapins', 'New Star', 'Summit', 'Sunburst', 'Sylvia' and 'Sweetheart', were developed directly from a common parent 'Van' (Schuster, 2012). Only five founding clones and one self-incompatible genetic source were found in 66 sweet cherry selections released from four breeding programs in North America (Choi and Kappel, 2004). Thus, exploring and introducing sweet cherry germplasms with new genetic backgrounds in breeding programs is crucial to overcome the present narrow genetic background. Therefore, the assessment of genetic diversity, as well as the relationships and structure of the germplasms, is the basis for efficient germplasm resources conservation and for selecting more elite germplasms in sweet cherry breeding programs.

Molecular identification based on molecular markers has become a popular tool for evaluating genetic relationships between and within

* Corresponding author.
E-mail address: liming06@caas.cn (M. Li).

https://doi.org/10.1016/j.scienta.2018.03.063
Received 25 January 2018; Received in revised form 27 March 2018; Accepted 29 March 2018

individuals or populations, and for assessing the purity and accurate identification of *Prunus* species (Bourguiba et al., 2012; Emanuelli et al., 2013; Li et al., 2013). Because of their advantages, including high degree of polymorphism, reproducibility, effectiveness and extensive genomic coverage, SSR markers were frequently used for germplasm characterization, species identification, genetic fingerprinting, parentage verification, genetic diversity analyses and genetic linkage map construction in sweet cherry (Campoy et al., 2016; Fernandez et al., 2012; Guajardo et al., 2015; Lacis et al., 2009; Marchese et al., 2017; Mariette et al., 2010; Schueler et al., 2003; Stanys et al., 2012; Wunsch and Hormaza, 2002). Although most of these SSRs have been cloned and sequenced in peach, these markers are transferable among *Prunus* species, including cherry (Fernandez et al., 2012), apricot (Zhebentyayeva et al., 2003) and plum (Mnejja et al., 2004). With the release of the sweet cherry genome sequence (Shirasawa et al., 2017), more microsatellites are available from the sweet cherry genome database (http://cherry.kazusa.or.jp/index.html).

Self-incompatibility in sweet cherry is controlled by a multi-allelic gametophytic self-incompatibility (*S*-) locus, which includes a style-specific ribonuclease (*S-RNase*) and a pollen-specific F-box protein gene (Yamane et al., 2003). In total, 18 different *S-RNase* alleles were detected in 47 incompatibility groups in sweet cherry after compiling the *S*-genotypes of 734 sweet cherry accessions (Schuster, 2012). The high level of genetic diversity allows it to be used as a genetic marker for variety identification and genetic diversity analyses when combined with SSR markers. Additionally, the diversity of *S*-alleles is directly linked to varietal diversity in cherry (Mariette et al., 2010). The lowest number of *S*-alleles was detected in cultivars from the new cherry growing regions of East Asia, North America and Australia (Schuster, 2012). A large number of different *S*-alleles and new incompatibility groups were found in an Italian sweet cherry germplasm collection that conserved abundant cherry resources (Marchese et al., 2017). Therefore, *S*-alleles and incompatibility groups of sweet cherry are not only especially important when choosing pollinator cultivars and designing specific combinations in breeding programs, but they are also essential for species identification and the construction of DNA fingerprints as genetic markers.

Here, 95 sweet cherry commercial cultivars, including 40 Chinese breeding varieties from different research institutes and 55 international commercial or old cultivars imported from different parts of the world, were analyzed with two types of markers: the *S*-locus and SSR markers. The objectives of this study were: (1) to select the most effective method for distinguishing different germplasms using the fewest primers; (2) to evaluate the genetic diversity levels and population structures within the collections; and (3) to confirm the pedigree-based relationships of the accessions.

2. Materials and methods

2.1. Plant materials and DNA extraction

A total of 95 sweet cherry commercial varieties maintained by the Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences (Zhengzhou, China) were investigated in this study. Of these, 40 cultivars were released in China, while the rest were introduced from other countries. Detailed information, including the cultivar name, parentage information, origin, *S*-genotype and self-incompatibility group was shown in Table 1. Total genomic DNA was isolated from young leaves or buds using a Plant Genomic DNA Kit (TIANGEN, China) and quantified by running on a 1.5% gel. The samples were then diluted to 10 ng/μl.

2.2. SSR analysis and S-allele identification

A set of 50 SSR primers previously developed in peach and sweet cherry was used to screen 10 representative sweet cherry cultivars. Finally, 10 pairs of primers were selected to test all of the accessions according to their locations on *Prunus* maps and the polymorphism of each marker. A list of primers is provided in Table 2. All forward primers were labeled at the 5′-end with one of the fluorescent dyes 6-FAM, ROX or HEX. SSR amplification was carried out in a volume of 20 μl, containing 10 ng DNA, 1.2 μl 2.5 mmol/L dNTPs, 2 μl 10 × PCR buffer (Mg^{2+}), 1 U *Taq* DNA polymerase (Takara Biotechnology Company, Dalian, China) and 0.8 μl each of 10 μmol/l forward and reverse primers. PCR amplification was carried out according to the following temperature profile: an initial step of 3 min at 94 ℃, 35 cycles of 45 s at 94℃, 45 s at 57 ℃ and 45 s at 72 ℃, with a final extension of 10 min at 72 ℃. PCR products were separated by capillary electrophoresis on an ABI PRISM 310 Genetic Analyzer using Genescan LIZ 500 as the size standard. Two reference cultivars ‘Summit’ and ‘Hongdeng’ were included in each run to monitor the repeatability of different amplifications.

Another consensus primer, Pru-T2/SI32 (Wang et al., 2010), which was designed to amplify the first intron of the *S-RNase* gene, was used to detect the *S*-genotypes of the 95 accessions. 60 of these had been analyzed for the *S-RNase* alleles in previous reports (Schuster, 2012; Wang et al., 2010; Wünsch and Hormaz, 2004), and the remaining 35 accessions were analyzed for the first time in this work. The forward primer was labeled with 6-FAM. PCR assays were performed as in the SSR analysis, and the primer information is provided in Table 2.

2.3. Data collection and transformation

The original SSR data were organized in a square matrix, with “?” representing the code for missing data. Then, the software Data Formater (Fan et al., 2016) was used for transforming the original size-based data into the input files for different population genetics software programs, including NTSYSpc, Powermarker and STRUCTURE.

2.4. Genetic diversity and phylogenetic tree analyses

To evaluate the genetic diversity within cultivars, the following parameters were estimated with Powermarker version 3.25 (Liu and Muse, 2005): major allele frequency (*MAF*), polymorphism information content (*PIC*), number of observed alleles per locus (*Ao*), expected heterozygosity (*He*), gene diversity (G_D) and genotype number (*Gn*). The genetic similarity coefficients between 95 individuals according to Dice were calculated by NTSYS-pc version 2.10e (Rohlf, 2000). A dendrogram was constructed using un-weighted pair group method with arithmetic mean (UPGMA) with a SHAN clustering program based on the genetic similarity coefficients between individuals according to SSR markers and the *S*-locus. Moreover, primers were selected according to their *PIC* values to analyze their efficiency using NTSYS-pc version 2.10e.

2.5. Analysis of population structure

The population structure of the sweet cherry germplasms was analyzed using STRUCTURE v.2.3.4 software, a model-based clustering method that can sort individuals into a number of clusters (K) based on their genetic similarity levels (Pritchard et al., 2000). Five independent runs for K values ranging from 2 to 7 were performed with a burn-in length of 100,000 followed by 100,000 Markov chain Monte Carlo. The STRUCTURE output was submitted to a website, and the program STRUCTURE HARVESTER (http://taylor0.biology.ucla.edu/struct_harvest/) was used as previously described (Earl and vonHoldt, 2012; Evanno et al., 2005) to estimate the most likely number of clusters.