Population Structure and Phylogeny of Some U.S. Peach Cultivars

Chunxian Chen and William R. Okie

U.S. Department of Agriculture, Agricultural Research Service, Southeastern Fruit and Tree Nut Research Lab, 21 Dunbar Road, Byron, GA 31008

ADDITIONAL INDEX WORDS. microsatellite, phylogenetic analysis, *Prunus persica*, simple sequence repeat

ABSTRACT. Peach (*Prunus persica*) cultivars maintained at the U.S. Department of Agriculture program at Byron, GA, have never been characterized with any molecular markers. In this study, 20 microsatellite markers were used to genotype 112 cultivars and the data were analyzed to discern their population structure and phylogenetic relationships. STRUCTURE simulations revealed four K clusters and broad genetic admixture among the cultivars. Principal coordinate analysis (PCoA) showed the cultivar groups from western, northeastern, and southeastern U.S. states were adjacent to each other except cultivars from Michigan (close to most southeastern state groups) and Florida (most distant from the other groups). Principal component analysis (PCA) showed that these cultivars had no obvious PCA partitioning boundaries. The intertwined distribution in both PCoA and PCA partitions suggested many of them were genetically closely related to each other largely because most shared same ancestral parentages. Most pairwise distance means within and between the cultivar groups were relatively low, suggesting close phylogenetic relations among those cultivars, as were demonstrated in the phylogenetic tree. Limiting factors and perspectives relevant to peach breeding are discussed.

Peach (*Prunus persica*) originated in northwest China, was introduced to Persia and the Mediterranean region via the Silk Route, and spread to colonies in America and other regions around the world (Faust and Timon, 1995; Li et al., 2013). Although peach appears to favor its native temperate cool dry climate and well-drained soil, diverse cultivars developed around the world have expanded the adaptability to a wide range of climatic and edaphic conditions from subtropical to subarctic zones (Meland et al., 2014; Okie, 1998; Rouse et al., 1985), thanks to natural selection over several thousand years and conventional breeding over the past century (Correa et al., 2019; Li and Wang, 2020; Okie, 1998; Okie et al., 1985; Reig et al., 2013). Largely due to chill requirement and cold hardiness, individual peach cultivars generally are adapted to restricted climate conditions and zones. As a result, breeding programs exist in different climatic regions around the world to meet localized cultivar needs. Relatively large collections of peach accessions have been maintained in stone fruit breeding programs around the world, as demonstrated in reviews, cultivar releases, and/or population genetics studies using those accessions; for example, from China (Cao et al., 2016; Li and Wang, 2020; Li et al., 2013; Shen et al., 2015), Spain (Aranzana et al., 2003, 2010; Carrillo-Navarro et al., 2015; Reig et al., 2013), Brazil (Correa et al., 2019; Thurow et al., 2020), and the United States (Chaparro et al., 2014; Chen and Okie, 2017, 2020a, 2020b, 2020c; Okie et al., 1985).

Knowledge of population structure, genetic relatedness, phylogenies, and pedigrees gained from genotyping or fingerprinting of peach accessions maintained in breeding programs would be useful in improving breeding efficiency through optimized parental combination, trait heritability prediction, and/or marker-assisted selection (MAS).

Microsatellite or simple sequence repeat (SSR) markers are commonly used in peach genotyping and/or population studies (Aranzana et al., 2003, 2010; Chen and Okie, 2017, 2021; Forcada et al., 2013; Li et al., 2013; Shen et al., 2015), followed by single nucleotide polymorphism (SNP) markers (Cao et al., 2016; Thurow et al., 2020). For example, genetic distance analysis of 48 SSRs and 653 peach accessions in a Chinese breeding collection clustered the accessions into two main groups: one with only four wild peach species and the other with most cultivars and landraces. STRUCTURE analysis assigned 469 accessions to three subpopulations: Oriental, Occidental, and Landraces. Linkage disequilibrium (LD) analysis of each subpopulation showed that LD decayed faster in the Oriental subgroup than in the Occidental subgroup. Intriguingly, inclusion of Chinese landraces greatly increased the genetic diversity in Occidental breeding programs (Li et al., 2013). STRUCTURE analysis of data from 49 SSRs with 195 peach accessions in another Chinese breeding program (i.e., 158 local cultivars of different Chinese ecological regions, 27 modern cultivars, and 10 wild accessions) divided them into eight subpopulations, which coincided with the genetic distance derived clusters and the conventionally descriptive subgroups: juicy honey peach, southwestern peach I, wild peach, Buddha peach + southwestern peach II, northern peach, southern crisp peach, ornamental peach, and *Prunus davidiana* + *Prunus kansuensis*. Most modern cultivars were related to the juicy honey peach and southwestern peach I subgroups, whereas the others were genetically diverse (Shen et al., 2015). Genome-wide association study of 12 traits using whole-genome sequencing data and 129 peach accessions collected in another Chinese breeding program revealed that most of the qualitative traits had more precise association signals compared with previous linkage analysis;

Received for publication 9 Aug. 2021. Accepted for publication 14 Sept. 2021. Published online 11 November 2021.

The research was partly supported by the U.S. Department of Agriculture (USDA) National Program of Plant Genetic Resources, Genomics and Genetic Improvement (Project number: 6042-21000-005-00D). We thank Luke Quick for sampling and Minling Zhang for genotyping. This article reports the results of research only. Mention of a trademark or proprietary product is solely for the purpose of providing specific information and does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products that may also be suitable.

C.C. is the corresponding author. E-mail: chunxian.chen@usda.gov.

This is an open access article distributed under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

candidate genes controlling the traits could be useful for characterization of functions and utilization in MAS (Cao et al., 2016).

Similarly, 16 SSR markers were used to fingerprint a collection of 212 peach and nectarine (*P. persica* var. *nectarina*) cultivars in a Spanish breeding program. According to the data, nonmelting canning peach cultivars (eaten fresh in Spain) were grouped together and separated distantly from the other cultivars tested. The level of homozygosity seemed to be a separator of conventionally bred cultivars from chance seedling-derived cultivars (Aranzana et al., 2003). In another study by the same authors, 224 peach and nectarine cultivars were genotyped using 50 SSRs evenly distributed on the peach reference map. The dataset divided the cultivars into three main groups primarily in line with melting peaches, melting nectarines, and nonmelting others. The nonmelting group had more homozygous alleles than melting counterparts, with exceptions of melting ‘Chinese Cling’, ‘Early Crawford’, and ‘Admiral Dewey’. The three cultivars, heavily used in the early peach breeding in the United States, were clustered together with commercial melting peaches, somewhat reflecting their foundational or representative impact on modern American and European commercial cultivars. Population structure analysis revealed subpopulations, and LD analysis suggested high LD conservation up to 13 to 15 cM in peach; future use of whole-genome scanning of traits could improve breeding efficiency (Aranzana et al., 2010). In another Spanish station at Aula Dei, population structure analysis of 92 peach cultivars genotyped with 40 SSRs revealed three subpopulations: Spanish local cultivars, international modern cultivars, and cultivars of admixture from the two subgroups. The local subgroup was somewhat less diverse than the modern subgroup. LD analysis revealed a high level of LD up to 20 cM, and suggested association mapping could be used to identify markers for MAS in peach breeding (Forcada et al., 2013). In a study of a Brazilian low-chill peach breeding germplasm using SNPs, analysis of population structure, genetic variability, and LD patterns of 220 peach genotypes revealed three main distinct subpopulations, suggesting SNP-trait associations in previously mapped regions could facilitate better MAS (Thurrow et al., 2020).

Recently we genotyped peach accessions in a U.S. Department of Agriculture (USDA), Agricultural Research Service (ARS) stone fruit breeding program using eight polymorphic chloroplast SSRs discovered among *Prunus* species. The data revealed eight unique maternal lineage groups (Chen and Okie, 2017). The collection includes many elite U.S. cultivars that are used in peach programs in the United States and around the world. In this study, 20 nuclear SSR markers were used to genotype 112 U.S. peach cultivars in the collection to describe their population structure and phylogenetic relationships.

Materials and Methods

MATERIALS USED. The 112 peach cultivars used in this study were from public and private breeding programs in the United States and maintained at variety blocks in the USDA-ARS Southeastern Fruit and Tree Nut Research Laboratory stone fruit breeding program at Byron, GA. Most cultivars were developed by U.S. state, federal, or private breeding programs; many were also used as breeding parents and/or were in commercial production at some time. To facilitate subsequent statistical analysis and descriptive comparison, the cultivars were divided into 11 sources (groups) (Table 1, Supplemental Table 1), where they

Table 1. Sources (groups), descriptions, and numbers of U.S. peach cultivars used in this study.

Source ^z	Source description	No.
AR	University of Arkansas, Fayetteville	5
CA	U.S. Department of Agriculture (USDA), Parlier, CA	3
CZ	Zaiger Genetics, Modesto, CA	15
CB	Other programs in California except the sources of CA and CZ	7
FL	University of Florida, Gainesville	4
GA	USDA, Byron, GA	36
SE	Other southeastern cultivar source states except Florida and Georgia	12
MI	Paul Friday Farm, Coloma, MI and other cultivars from Michigan	17
NJ	Rutgers University, New Brunswick, NJ	3
WV	USDA, Kearneysville, WV	4
NE	Other northeastern cultivar source states except New Jersey and West Virginia	6
Total		112

^zCultivar sources (groups) mostly are named by abbreviated states where cultivars were developed, except CZ, CB, SE, and NE. CZ and CB are two additional sources in California. SE and NE include other southeastern and northeastern cultivar source U.S. states not listed here, respectively.

were developed. The 11 cultivar sources (groups) are as follows: University of Arkansas, Fayetteville, AR (AR); USDA, Parlier, CA (CA); Zaiger Genetics, Modesto, CA (CZ); other cultivar sources in California, except CA and CZ (CB); University of Florida, Gainesville (FL); USDA, Byron, GA (GA); other southeastern cultivar source states except Florida and Georgia (SE); Paul Friday Farm, Coloma, MI, and other cultivar sources in Michigan [MI (all the cultivar names start with PF, the initials of Paul Friday)]; Rutgers University, New Brunswick, NJ (NJ); USDA, Kearneysville, WV (WV); and other northeastern cultivar source states except New Jersey and West Virginia (NE). Genomic DNA was isolated from 200 mg of young leaves using a cetyltrimethylammonium bromide (CTAB) protocol (Doyle and Doyle, 1987), with in-house modifications to scale down to use 2-mL Eppendorf tubes for the isolation.

SSR MARKERS. Twenty SSR markers (Supplemental Table 2), previously characterized (Chen et al., 2014), were selected to genotype the 112 cultivars in this study. These SSR markers were selected primarily for their proved amplification performance, known chromosomal positions, and potentially high polymorphisms, as demonstrated previously (Chen et al., 2014).

SSR GENOTYPING AND SCORING PROCEDURE. The SSR genotyping and scoring procedure was previously described (Chen et al., 2014). Genotyping was performed on an eight-capillary sequencing instrument (Applied Biosystems 3500 Genetic Analyzer; Thermo Fisher Scientific, Waltham, MA) to generate fluorescent dye-based chromatographic files that were used to score SSR alleles using GeneMarker (SoftGenetics, State College, PA) and form an allele table for subsequent analysis.

SSR ALLELE DATA ANALYSIS. The SSR allele table was converted to the format required by STRUCTURE 2.3.4 (Pritchard et al., 2000), PowerMarker 3.25 (Liu and Muse, 2005), GenAIEx 6.5 (Peakall and Smouse, 2006, 2012), and the R package adegenet (Jombart, 2008; Jombart et al., 2010), respectively.

STRUCTURE 2.3.4 was used to determine the population structure and estimate K (K-means clustering) of the collection using simulation run under the following set of parameters: 100,000 as the length of Burnin period, 500,000 as the number of Markov Chain Monte Carlo repetitions after Burnin, admixture as the ancestry model, allele frequencies correlated, no prior defined population information, K = 1 to 15, and 10 iterations each K. The results generated by STRUCTURE were entered into the website program Structure Harvester (Earl and Vonholdt, 2012) for visualizing likelihood values of multiple K and detecting the number of genetic groups that best fit the data. The estimated cluster membership coefficient matrices of the preceding multiple runs of STRUCTURE were used as input for CPLUMPP 1.1.2 (Jakobsson and Rosenberg, 2007) to find optimal alignments of replicate cluster analyses of the same data. The output from CLUMPP was used directly as input into DISTRUCT (Rosenberg, 2004), a cluster visualization program, to graphically display the population structure. PowerMarker was used to summarize the genotyping data, including the major allele frequency (MAF), genotype number detected (GN), allele number detected (AN), gene diversity (GD), heterozygosity (H), and polymorphism information content (PIC); calculate the shared allele and Euclidean genetic distances; and draw the distances based on phylogenetic trees with robustness of branches tested with 1000 bootstraps. The trees were displayed using TreeView X (Page, 1996). GenAlEx 6.5 was used to perform PCoA with the standardized pairwise genetic distance matrixes and draw the PCoA chart to show the coordinated partitions of the groups and individual genotypes, respectively. The data formatted for GenAlEx were imported to the R package adegenet to perform discriminant analysis of principal components, a method equivalent to PCA of populations with molecular marker data, as previously described (Jombart, 2008; Jombart et al., 2010).

Results

SUMMARY OF SSR GENOTYPING DATA. The genotyping results of the 20 SSR markers were summarized (Supplemental Table 3), varying in MAF, GN, AN, GD, H, and PIC values. Given the primers detecting GN and AN in the samples, CX2A01, CX1E06, and CX2E02 had both high numbers (31 and 18, 27 and 16, and 24 and 14, respectively), whereas CX2B12, CX6H11, and CX3H06 had both low numbers (6 and 5, 6 and 5, and 8 and 7, respectively). The PIC values of CX2F12, CX1E06, and CX2E08 (0.63, 0.62, and 0.56, respectively) were among the highest, and those of CX2B12, CX3H06, and CX1G04 (0.03, 0.07, and 0.09, respectively), among the lowest.

POPULATION STRUCTURE. Four K clusters (K = 4) were determined by STRUCTURE simulations. Most accessions across the four clusters had broad genetic admixture among them (Fig. 1). Given the cultivars with a source state, almost all of them were elite scion cultivars generated by public or private breeding programs in these states, including USDA stone fruit breeding programs in California, Georgia, and West Virginia. Most of the scion cultivars shared similar composite patterns predominant in red or yellow varying among genotypes from different groups. The genotypical composites of the 17 MI cultivars looked more uniform than any other groups, suggesting they were genetically very close to each other. They collectively appeared to have the least allelic admixture in this collection. In contrast, the cultivars

Fig. 1. Population structure of the 112 peach cultivars in the U.S. Department of Agriculture, Byron, GA, collection. Four K-means clusters were represented by red, yellow, green, and blue colors. The 11 cultivar sources (groups) are described in Table 1.

developed in the USDA scion breeding program at Byron, GA, looked most diverse.

PCoA partitioning showed most groups (sources) were adjacent to each other except MI and FL. MI was close to the groups from the southeast, whereas the FL group, including only four old cultivars and apparently not reflecting the current breeding there, was most distant from the other groups. However, cultivars developed by these states except FL were closely adjacent in two coordinate partitions (Fig. 2), suggesting the close relationships among them. Likewise, the close relationships among the cultivars and groups were also demonstrated by the largely intertwined distribution in the PCA chart with a few exceptions (Fig. 3), which mostly were rootstock or ornamental cultivars (i.e., Candy Cane).

PAIRWISE GENETIC DISTANCES AND PHYLOGENY. The means of pairwise genetic distances within the same groups and between different groups were calculated to indicate the trend of differences among them (Table 2). The SDs of the means were between 0.03 and 0.08 (data not shown). A larger mean implied that individual accessions more genetically distant within the same group or between the different groups. Within the groups, AR had the highest mean (0.51), whereas MI had the lowest (0.21), which corroborated the results of STRUCTURE simulations. The within-group mean was 0.40. Means between the groups ranged from 0.29 (MI-GA, the lowest) to 0.47 (AR-WV, the highest), with an average of 0.39 (Table 2).

The phylogenetic tree split U.S. cultivars into approximately six clusters (I–VI) (Fig. 4). Each cluster was generally mixed with cultivars developed by different programs but with some clustered together from the same programs. Cluster I included several of our recent releases that were grouped together with the cultivar O’Henry, the source of desired red blush and firmness in these cultivars (Augustprince, Early Augustprince, and their parent Rich Joy, plus their other parent Sunprince).

Fig. 2. Principal coordinate analysis of the 11 peach cultivar sources (groups). Cultivar sources are described in Table 1.

Fig. 3. Principal component analysis (PCA) of the 11 peach cultivar sources (groups). The legend for the groups, PCA eigenvalues, and discriminant analysis (DA) eigenvalues were shown at the three corners, respectively. Cultivar sources are described in Table 1.

‘Blazeprince’ was also in the group, but its open-pollinated (OP) seedling ‘Scarletprince’ was grouped closely with ‘PF 7A’, which might be its pollen parent. Cluster I had most cultivars from our program, in which the cultivar O’Henry from CB was used to improve peach fruit blush and eating quality. A trio of recent “Joy” releases, ‘Rich Joy’, ‘Crimson Joy’, and ‘Liberty Joy’, also fell in this cluster. Cluster II included ‘Julyprince’, ‘Flameprince’, ‘Redglobe’, and ‘Harvester’, which are some of the more widely used commercial cultivars in the southeast United States. ‘Redglobe’ is in the parentage of ‘Flameprince’, and ‘Julyprince’ descended from ‘Harvester’. ‘Harvester’ was descended from a possible mutation of ‘Fireglow’, which was a parent of ‘Redglobe’. ‘GaLa’, an OP of ‘Harvester’, was not in this group, but close to its pollen parent, ‘Jefferson’. The MI cultivars tended to clump together in Cluster II. These were generally derived from crosses of cold-hardy MI cultivars with a limited number of superior cultivars from the three groups (CA, CZ, and CB) from California. ‘Springcrest’ may be related to the parents from

Table 2. Mean of pairwise genetic distances of individual genotypes within and between cultivar sources (groups). The 11 cultivar sources are described in Table 1.

	AR	CA	CB	CZ	FL	GA	MI	NE	NJ	SE	WV
AR	0.51										
CA	0.45	0.32									
CB	0.47	0.38	0.42								
CZ	0.43	0.34	0.38	0.33							
FL	0.45	0.39	0.41	0.37	0.31						
GA	0.45	0.35	0.38	0.34	0.36	0.33					
MI	0.43	0.31	0.35	0.30	0.33	0.29	0.21				
NE	0.44	0.38	0.41	0.38	0.39	0.37	0.34	0.40			
NJ	0.44	0.40	0.42	0.38	0.40	0.39	0.35	0.40	0.40		
SE	0.44	0.37	0.40	0.36	0.37	0.36	0.31	0.38	0.40	0.37	
WV	0.47	0.39	0.42	0.40	0.39	0.38	0.33	0.42	0.42	0.41	0.44

Fig. 4. Neighbor joining tree of U.S. peach cultivars based on Euclidean distance matrices. Clusters I to VI were marked to facilitate illustration of the tree. The bar represents the distance of 0.1.

California used in the MI cultivars. ‘Big Red’ was a grower-named selection from USDA-ARS, Parlier, CA, which likely had similar parentage. Cluster III included nectarines (‘Arctic Sweet’, ‘Arctic Belle’). Cluster IV appeared to be broadly mixed with cultivars of different sources, including low-chill cultivars in the FL and TX groups. They tended to group together probably due to use of Florida accessions in the Texas breeding program. For example, ‘Texprince’ is a seedling of ‘Flordaking’. Although ‘Flordaguard’ was a rootstock, it came out of the fruiting cultivar breeding work. Cluster V was mixed mostly with cultivars of

different sources. Cluster VI included North Carolina (NC) peaches ('Contender', 'Clayton', and 'Carolina Gold') that were grouped closely together, which were resistant to bacterial spot and often used as a source of the resistance. With them was 'Redhaven' from MI, which was in the pedigrees of NC peaches, and a parent of 'TruGold'. Most flat or saucer peaches ('Sauzee King', 'Sweet Cap', and 'Flat Wonderful') grouped near each other, along with many of the white-fleshed peaches from Arkansas.

Discussion

Peach is an introduced crop with a relatively short cultivation and variety development history in the United States (Okie, 1998; Okie et al., 1985). To meet local industry needs and climatic requirements, over the past century germplasm was exchanged, and cultivars released from dozens of public and private breeding programs, of which fewer than 10 are still active. This study revealed the microsatellite-based population structure, genetic relatedness, diversity and phylogeny of the U.S. peach cultivars maintained at the USDA, Byron, GA, stone fruit breeding program (Chen, 2021; Chen and Okie, 2017), which is ≈ 80 years old. The results provide additional molecular marker insight into these U.S. cultivars developed by different state, federal, and private peach breeding programs. The admixture nature and close relationship of most U.S. cultivars should be expected and are apparently due to exchange and shared use of breeding accessions among those breeding programs.

It is worth noting that the cultivars in the study are fruit scions except two rootstocks ('Guardian' and 'Flordaguard') and one ornamental peach ('Candy Cane'). Population structure analysis showed the four K clusters (represented by red, yellow, green, and blue colors) and admixture nature among most of the U.S. cultivars. They shared similar composite patterns predominant in red or yellow and varied slightly among different groups. Based on early use of Chinese and European peach introductions in most U.S. scion breeding programs and the allelic composites used by STRUCTURE simulations to determine the clusters, the four K clusters might be well corresponding to the following ancestral gene pools and represented by red, yellow, green, and blue in the cluster visualization (Fig. 1), respectively: 1) Chinese scion introductions; 2) European scion introductions; 3) Chinese wild species and ornamental introductions; and 4) other feral or rare genotypes. In other words, allelic composites (red) might be directly from 'Chinese Cling' and its offspring, and those (yellow) from the early European introductions and their progeny. They appeared predominant because of the early heavy use of these accessions in both pools as parents in most U.S. scion breeding programs (Myers et al., 1989; Okie et al., 1985; Scorza et al., 1985). The allelic composites in the green cluster apparently were derived from Chinese wild species and ornamental peaches that led to selections of modern rootstocks and ornamental peaches, for instance, 'Guardian' and 'Candy Cane' from the USDA, Byron, GA, stone fruit breeding program (Okie, 2013; Okie et al., 1994). The allelic composites in the blue cluster appeared to be rare and scattered in only a handful of cultivars. Although the results were gained from limited SSR markers, the K-means clustering information gained by STRUCTURE simulations should be useful for further categorization of these accessions and assistance in choice of parents. In addition, although reportedly a doubled haploid, 'TruGold' was heterozygous for some alleles, which agreed with a previous report (Sithther et al., 2012).

Similar breeding and research efforts have also been taken by other programs around the world (Correa et al., 2019; Li and Wang, 2020; Li et al., 2013; Reig et al., 2013; Thurrow et al., 2020). These breeding activities have led to the release of numerous localized cultivars, and the genetic researchers have gained knowledge of those collections that could facilitate breeding in several ways. Although a large number of peach cultivars have been released around the world, most of them apparently offer only incremental improvement and thus become obsolete rapidly (Okie, 1998), as we have observed at the USDA, Byron, GA, program (Chen, 2021). The situation might be related to several facts or conjectures. Peach is self-compatible, and thousands of years of natural introgression and selection have led to a high rate of homozygosity and inbreeding tendency (Chen et al., 2014; Scorza et al., 1985). In addition, use of a limited array of parents in conventional hybridization, primarily based on organoleptic characteristics of fruit, and breeders' experience and bias, might have accelerated and enhanced the trend. As a result, the natural and conventional selection might have resulted in loss of many agriculturally important genes with traits that were not easily visible or constitutively expressed, and in reduction of genetic diversity among peach cultivars. For example, compared with that of native wild peach trees, lifespan of commercial peach trees is substantially reduced, primarily due to heartwood decay in the scaffolds (Chen et al., 2015). The decay does not occur in several rootstock and wild *Prunus* accessions, suggesting the reduction probably is due to loss of defense-related genes because organoleptic fruit characteristics and yield often are among top factors evaluated during selection of scion cultivars. Therefore, understanding and knowledge of the genetic and phylogenetic relationship among potential parents should be considered. Parents with greater diversity and genetic distances could be prioritized, which might increase the likelihood of heterosis or outbreeding enhancement in peach breeding. Such benefits have been demonstrated in many staple crops, using rice (*Oryza sativa*) as an example, in which whole-genome knowledge of many genes and traits are well characterized and applied to breed super hybrids with many desired traits (Qian et al., 2016; Yu et al., 2020). That should be the direction of breeding pursuits, although peach and other woody perennial fruit crops are much more difficult and need more time to decipher the genes and traits in various potential germplasms. The trend toward fewer peach breeding programs and more impediments on germplasm exchanges will continue to reduce the range and choice of desired germplasms available for use in crossing and thus make well-informed parental selection even more critical.

Literature Cited

- Aranzana, M.J., E.K. Abbassi, W. Howad, and P. Arus. 2010. Genetic variation, population structure and linkage disequilibrium in peach commercial varieties. *BMC Genet.* 11:69, <https://doi.org/10.1186/1471-2156-11-69>.
- Aranzana, M.J., J. Carbo, and P. Arus. 2003. Microsatellite variability in peach [*Prunus persica* (L.) Batsch]: Cultivar identification, marker mutation, pedigree inferences and population structure. *Theor. Appl. Genet.* 106:1341–1352, <https://doi.org/10.1007/s00122-002-1128-5>.
- Cao, K., Z.K. Zhou, Q. Wang, J. Guo, P. Zhao, G.R. Zhu, W.C. Fang, C.W. Chen, X.W. Wang, X.L. Wang, Z.X. Tian, and L.R. Wang. 2016. Genome-wide association study of 12 agronomic traits in peach. *Nat. Commun.* 7:13246, <https://doi.org/10.1038/ncomms13246>.
- Carrillo-Navarro, A., A. Guevara-Gazquez, M. Perez-Jimenez, L. Ruiz-Garcia, and J. Cos-Terrer. 2015. 'Alisio 15 (R)': An early-maturing peach cultivar for the fresh fruit market. *HortScience* 50:312–314, <https://doi.org/10.21273/HORTSCI.50.2.312>.

- Chaparro, J.X., P.J. Conner, and T.G. Beckman. 2014. 'GulfAtlas' peach. *HortScience* 49:1093–1094, <https://doi.org/10.21273/HORTSCI.49.8.1093>.
- Chen, C. 2021. Peach cultivar releases and fruit trait distribution in the USDA-ARS Byron program. *Acta Hort.* (1304):29–36, <https://doi.org/10.17660/ActaHortic.2021.1304.4>.
- Chen, C., C.H. Bock, M.H. Hotchkiss, M.M. Garbelotto, and T.E. Cottrell. 2015. Observation and identification of wood decay fungi from the heartwood of peach tree limbs in central Georgia, USA. *Eur. J. Plant Pathol.* 143:11–23, <https://doi.org/10.1007/s10658-015-0661-4>.
- Chen, C., C.H. Bock, W.R. Okie, F.G. Gmitter, S. Jung, D. Main, T.G. Beckman, and B.W. Wood. 2014. Genome-wide characterization and selection of expressed sequence tag simple sequence repeat primers for optimized marker distribution and reliability in peach. *Tree Genet. Genomes* 10:1271–1279, <https://doi.org/10.1007/s11295-014-0759-4>.
- Chen, C. and W.R. Okie. 2017. Characterization of polymorphic chloroplast microsatellites in *Prunus* species and maternal lineages in peach genotypes. *J. Amer. Soc. Hort. Sci.* 142:217–224, <https://doi.org/10.21273/Jashs04070-17>.
- Chen, C. and W.R. Okie. 2020a. 'Crimson Joy' peach. *HortScience* 55:972–973, <https://doi.org/10.21273/Hortsci14983-20>.
- Chen, C. and W.R. Okie. 2020b. 'Liberty Joy' peach. *HortScience* 55:951–952, <https://doi.org/10.21273/Hortsci14907-20>.
- Chen, C. and W.R. Okie. 2020c. 'Rich Joy' peach. *HortScience* 55:591–592, <https://doi.org/10.21273/Hortsci14720-19>.
- Chen, C. and W.R. Okie. 2021. Genetic relationship and parentages of historical peaches revealed by microsatellite markers. *Tree Genet. Genomes* 17:35, <https://doi.org/10.1007/s11295-021-01517-8>.
- Correa, E.R., M. Nardino, W.S. Barros, and M.D.B. Raseira. 2019. Genetic progress of the peach breeding program of Embrapa over 16 years. *Crop Breed. Appl. Biotechnol.* 19:319–328, <https://doi.org/10.1590/1984-70332019v19n3a44>.
- Doyle, J.J. and J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19:11–15.
- Earl, D.A. and B.M. Vonholdt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4:359–361, <https://doi.org/10.1007/s12686-011-9548-7>.
- Faust, M. and B. Timon. 1995. Origin and dissemination of peach. *Hort. Rev.* 17:331–379, <https://doi.org/10.1002/9780470650585.ch10>.
- Forcada, C.F.I., N. Oraguzie, E. Igarua, M.A. Moreno, and Y. Gogorcena. 2013. Population structure and marker-trait associations for pomological traits in peach and nectarine cultivars. *Tree Genet. Genomes* 9:331–349, <https://doi.org/10.1007/s11295-012-0553-0>.
- Jakobsson, M. and N.A. Rosenberg. 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806, <https://doi.org/10.1093/bioinformatics/btm233>.
- Jombart, T. 2008. adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405, <https://doi.org/10.1093/bioinformatics/btn129>.
- Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genet.* 11:94, <https://doi.org/10.1186/1471-2156-11-94>.
- Li, X.W., X.Q. Meng, H.J. Jia, M.L. Yu, R.J. Ma, L.R. Wang, K. Cao, Z.J. Shen, L. Niu, J.B. Tian, M.J. Chen, M. Xie, P. Arus, Z.S. Gao, and M.J. Aranzana. 2013. Peach genetic resources: Diversity, population structure and linkage disequilibrium. *BMC Genet.* 14:84, <https://doi.org/10.1186/1471-2156-14-84>.
- Li, Y. and L.R. Wang. 2020. Genetic resources, breeding programs in China, and gene mining of peach: A review. *Hort. Plant J.* 6:205–215, <https://doi.org/10.1016/j.hpj.2020.06.001>.
- Liu, K. and S.V. Muse. 2005. PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics* 21:2128–2129, <https://doi.org/10.1093/bioinformatics/bti282>.
- Meland, M., O. Frøynes, and C. Kaiser. 2014. Evaluation of peach cultivars in cool, mesic Ullensvang, Norway. *HortTechnology* 24:618–622, <https://doi.org/10.21273/HORTTECH.24.5.618>.
- Myers, S.C., W.R. Okie, and G. Lightner. 1989. The Elberta peach. *Fruit Var. J.* 43:130–138.
- Okie, W.R. 1998. Handbook of peach and nectarine varieties: Performance in the southeastern United States and index of names. U.S. Dept. Agr. Hdbk No. 714.
- Okie, W.R. 2013. Peach tree named 'Candy Cane'. U.S. Plant Patent No. 22,443. Filed 29 July 2011. Issued 5 Mar. 2013.
- Okie, W.R., T.G. Beckman, A.P. Nyczepir, G.L. Reighard, W.C. Newall, and E.I. Zehr. 1994. BY520-9, a peach rootstock for the southeastern United States that increases scion longevity. *HortScience* 29:705–706, <https://doi.org/10.21273/HORTSCI.29.6.705>.
- Okie, W.R., D.W. Ramming, and R. Scorza. 1985. Peach, nectarine, and other stone fruit breeding by the USDA in the last 2 decades. *HortScience* 20:633–641.
- Page, R.D.M. 1996. TreeView: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12:357–358.
- Peakall, R. and P.E. Smouse. 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* 6:288–295, <https://doi.org/10.1111/j.1471-8286.2005.01155.x>.
- Peakall, R. and P.E. Smouse. 2012. GenAEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539, <https://doi.org/10.1093/bioinformatics/bts460>.
- Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Qian, Q., L.B. Guo, S.M. Smith, and J.Y. Li. 2016. Breeding high-yield superior quality hybrid super rice by rational design. *Natl. Sci. Rev.* 3:283–294, <https://doi.org/10.1093/nsr/nww006>.
- Reig, G., S. Alegre, F. Gatiús, and I. Iglesias. 2013. Agronomical performance under Mediterranean climatic conditions among peach [*Prunus persica* L. (Batsch)] cultivars originated from different breeding programmes. *Scientia Hort.* 150:267–277, <https://doi.org/10.1016/j.scienta.2012.11.006>.
- Rosenberg, N.A. 2004. DISTRUCT: A program for the graphical display of population structure. *Mol. Ecol. Notes* 4:137–138, <https://doi.org/10.1046/j.1471-8286.2003.00566.x>.
- Rouse, R.E., W.B. Sherman, and R.H. Sharpe. 1985. Flordagrande - A peach for sub-tropical climates. *HortScience* 20:304–305.
- Scorza, R., S.A. Mehlenbacher, and G.W. Lightner. 1985. Inbreeding and coancestry of freestone peach cultivars of the eastern United States and implications for peach germplasm improvement. *J. Amer. Soc. Hort. Sci.* 110:547–552.
- Shen, Z.J., R.J. Ma, Z.X. Cai, M.L. Yu, and Z. Zhang. 2015. Diversity, population structure, and evolution of local peach cultivars in China identified by simple sequence repeats. *Genet. Mol. Res.* 14:101–117, <https://doi.org/10.4238/2015.January.15.13>.
- Sitther, V., D.P. Zhang, S.A. Dhekney, D.L. Harris, A.K. Yadav, and W.R. Okie. 2012. Cultivar identification, pedigree verification, and diversity analysis among peach cultivars based on simple sequence repeat markers. *J. Amer. Soc. Hort. Sci.* 137:114–121, <https://doi.org/10.21273/JASHS.137.2.114>.
- Thurrow, L.B., K. Gasic, M.D.B. Raseira, S. Bonow, and C.M. Castro. 2020. Genome-wide SNP discovery through genotyping by sequencing, population structure, and linkage disequilibrium in Brazilian peach breeding germplasm. *Tree Genet. Genomes* 16:10, <https://doi.org/10.1007/s11295-019-1406-x>.
- Yu, S.B., J. Ali, C.P. Zhang, Z.K. Li, and Q.F. Zhang. 2020. Genomic breeding of green super rice varieties and their deployment in Asia and Africa. *Theor. Appl. Genet.* 133:1427–1442, <https://doi.org/10.1007/s00122-019-03516-9>.

Supplemental Table 1. U.S. peach cultivars used in this study and their designated sources (groups).

Cultivar	Source ^z
White Cloud	AR
White County	AR
White Diamond	AR
White River	AR
White Rock	AR
Big Red	CA
Fairtime	CA
Galaxy	CA
Arctic Belle	CZ
Arctic Gold	CZ
Arctic Sweet	CZ
Crimson Snow	CZ
Flavorich	CZ
Glacier White	CZ
Honey Royale	CZ
Joanna Sweet	CZ
Sweet Cap	CZ
Sauzee King	CZ
Spice Zee	CZ
Tri Lite	CZ
Yumm	CZ
Strawberry Free	CZ
Zephyr	CZ
Eastern Glo	CB
June Gold	CB
Karla Rose	CB
O'Henry	CB
Springtime	CB
Snow Queen	CB
Summerset	CB
Flordadawn	FL
Flordaguard	FL
Flordaking	FL
Flordaprince	FL
Augustprince	GA
Early Augustprince	GA
Autumnprince	GA
Blazeprince	GA
Crimson Joy	GA
Liberty Joy	GA
Rich Joy	GA
Coronet	GA
Dixiland	GA
Durbin	GA
Empress	GA
Fireprince	GA
Flameprince	GA
Goldprince	GA
Julyprince	GA
Juneprince	GA
Juneprincess	GA
Lobo	GA
Roseprincess	GA
Rubyprince	GA
Scarletpearl	GA
Scarletprince	GA

(Continued on next page)

Supplemental Table 1. (Continued)

Cultivar	Source ^z
Sentinel	GA
Southern Pearl	GA
Springcrest	GA
Springprince	GA
Starlite	GA
Summergold	GA
Summerprince	GA
Sunbrite	GA
Sunland	GA
Sunprince	GA
Sureprince	GA
Suwanee	GA
Candy Cane	GA
Guardian	GA
Carolina Gold	SE
Clayton	SE
Contender	SE
Carolina Red	SE
Winblo	SE
Carored	SE
Camden	SE
Majestic	SE
GaLa	SE
Harvester	SE
Texking	SE
Texprince	SE
PF10 Ball	MI
PF11 Ball	MI
PF17	MI
PF19-007	MI
PF23	MI
PF24C	MI
PF27A	MI
PF36-007	MI
PF5D Big	MI
PF7A	MI
PF8 Ball	MI
PF9-007	MI
PF-Late Large 23	MI
PF-Late 24-007	MI
Cresthaven	MI
Halehaven	MI
Redhaven	MI
Flat Wonderful	NJ
Jerseyqueen	NJ
Saturn	NJ
Flavrburst	WV
Sentry	WV
SummerFest	WV
TruGold	WV
Redskin	NE
Redglobe	NE
Loring	NE
Cary Mac	NE
Reliance	NE
Jefferson	NE

^zCultivar sources are described in Table 1.

Supplemental Table 2. Simple sequence repeat markers used for genotyping of U.S. peach cultivars.

Marker	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Size (bp) ^z	Accession ID	Chr ^y
CX1G04	ATTCTTCAGATTGTTCTGCTCTTCT	ACACAGTACAATCACGCTACCATT	173	CV045920	1
CX1E06	TTTGAACCAACACAAAGTCTGGT	GAATGTCACCAGAGATATAAGGGG	124	DY636511	1
CX1G06	CGCCTTTTCATCTTTTCTAACTC	AGATTGTTCTTTTGAAAGCCTGAG	208	DY651077	1
CX1B09	AAAAAGCAACAAACCCAATTACC	AATTAATTGAGCTCCTTTTGGTG	260	AM291698	2
CX2A01	CATCTTCTTAACCTGGTCTTTGGG	ATGTTTTCCTTGGGGTATACTGTC	219	CV052011	2
CX2E02	CCAAAATTACACACCCGAAAT	TACAGTGGTGAGAGCAACAGAGAC	121	EE488456	3
CX2B03	CACATCAGAATCTTACCATCTTC	ATTCTTCCTGAGATCAAATCCAAC	232	DY634703	3
CX2F06	TTTTTATCACACACAGCCAAAGT	AACCACCAGCTGCTCTCTAGTAAT	161	HE645523	4
CX2E08	AACAAAGAGACAGAGAGCACACAG	AGAATAGGACCTGGCGGAACT	300	DY637160	4
CX2D10	GACCGCATCAATAATACAATACGA	TGCTGTTTGGATACGAATTAAGAA	189	DW345359	5
CX2B12	AATTTTATTGGTCTGTGAGTGCAA	GTGGTGATGAGTCTGTAAACTTGG	247	FC864765	5
CX2F12	AAGGGTACCAATACCAAATTTCAA	GTCGCATTCTTTTGTTTTCTTT	266	HE640488	5
CX3A02	ATATCATATCATCCCATATCCCCA	TTTTCTTGTCAGACAGCTTTTC	229	HE643816	6
CX3F03	GAAGAGCAGAGCAAAATACACAGA	TCTCTGTTAGTAGACTGTCCACGC	108	AM291395	6
CX3B06	ATTTCCAAACCCTCTCTATCGC	GTTTTGGTCTTGAGATGAATGTTG	293	DY644129	7
CX3H06	ACCGAAAGAGAGATGGATGTTAAG	GATATAAATGGCTCAAAAGCTCGT	293	GW868981	7
CX6H11	GCTCTCTCTGCATTCTAGACTAT	TGCTAGAAATACATTCAATTCCGA	169	AM290588	7
CX3H08	AGAGAAAGAAAATGACCAATCCAG	AATATTTTCATCATCTGCATCAGGA	145	BU574017	8
CX3H09	AAGGTTATGTGATTTTGCTCTTCC	AGCGCAATTGTTACTCTTTGAAAT	144	DW345069	8
CX3F11	GGATTGAAACTGAATACAAAACCC	GCCATCATCACCATCAAGTAAATA	155	DY637734	8

^zThe expected amplicon size calculated from the expressed sequence tag sequences (their accession codes in the next column). bp = base pairs.^yNumbers refer to which chromosome (scaffold) the markers locate in the peach reference genome.

Supplemental Table 3. Summary of simple sequence repeat marker genotyping data from U.S. peach cultivars.

Marker	MAF ^z	GN	AN	GD	H	PIC
CX1G04	0.98	2	2	0.04	0.00	0.03
CX1E06	0.55	11	7	0.61	0.57	0.55
CX1G06	0.54	7	5	0.53	0.46	0.42
CX1B09	0.51	6	3	0.58	0.50	0.49
CX2A01	0.66	8	6	0.47	0.42	0.39
CX2E02	0.65	11	7	0.52	0.44	0.47
CX2B03	0.55	7	6	0.53	0.50	0.42
CX2F06	0.87	2	2	0.23	0.26	0.20
CX2E08	0.42	7	4	0.66	0.54	0.59
CX2D10	0.90	8	6	0.18	0.15	0.18
CX2B12	0.99	2	2	0.02	0.00	0.02
CX2F12	0.48	12	7	0.68	0.60	0.63
CX3A02	0.96	4	4	0.09	0.04	0.09
CX3F03	0.67	8	5	0.46	0.46	0.38
CX3B06	0.64	4	3	0.46	0.46	0.36
CX3H06	0.99	2	2	0.03	0.03	0.03
CX6H11	0.96	3	3	0.09	0.00	0.08
CX3H08	0.61	6	3	0.55	0.52	0.48
CX3H09	0.86	4	3	0.25	0.25	0.23
CX3F11	0.74	10	5	0.42	0.38	0.38
Mean	0.73	6	4	0.37	0.33	0.32

^zMAF = major allele frequency; GN = genotype number detected; AN = allele number detected; GD = gene diversity, H = heterozygosity; PIC = polymorphism information content.