J. Amer. Soc. Hort. Sci. 148(5):221–228. 2023. https://doi.org/10.21273/JASHS05292-23

Genetic Diversity of New Almond Accessions from Central Asian and Cold-adapted North American Germplasm

Per McCord
Department of Horticulture, Washington State University, Irrigated Agriculture Research and Extension Center, 24106 N. Bunn Road, Prosser, WA 99350, USA

Vishal Singh, Amita Kaundal, and Teryl Roper
Department of Plants, Soils, and Climate, Utah State University, 4820 Old Main Hill, Logan, UT 84322, USA

Keywords. discriminant analysis, hierarchical clustering, plant genetic resources, principal components analysis, *Prunus dulcis*, SSR

Abstract. We evaluated the genetic diversity of a newly available collection of 94 almond [*Prunus dulcis* (Mill.) D.A. Webb] accessions from the former Improving Perennial Plants for Food and Bioenergy (IPPFBE) Foundation. Most of the collection (87 accessions) were collected as seeds from trees growing in the central Asian nations of Kyrgyzstan, Tajikistan, and Uzbekistan, and included several examples of *Prunus bucharica* (Korsh.) Hand.-Mazz, and related wild species. Of the remaining accessions, six were sourced from a nursery in northern Utah in the United States, and one was a seedling of 'Nonpareil', a major commercial cultivar. DNA fingerprints were generated from 10 simple sequence repeat markers. To evaluate the comparative diversity of these new accessions, 66 accessions from the US Department of Agriculture, National Plant Germplasm System (NPGS) almond germplasm collection near Davis, CA, USA, were also included. These NPGS accessions were chosen to represent those collected in similar regions of Central Asia and the Caucasus. The fingerprints were analyzed via hierarchical clustering, principal components analysis (PCA), and discriminant analysis of principal components (DAPC). Hierarchical clustering suggested that half of the Utah-sourced accessions are closely related to each other and to the 'Nonpareil' seedling. Additional close relationships were detected (including at least one duplication or mislabeling), and two *P. bucharica* accessions from the IPPFBE collection were separated from the rest of the collection. A plot of the first two principal components clearly separated wild almond relatives (*P. bucharica* and *Prunus fenzliana* Fritsch) from the remaining accessions. PCA after removal of the wild species separated the 'Nonpareil' seedling, the Utah-sourced accessions, and many of the IPPFBE accessions (mostly from Uzbekistan) from nearly all other individuals. The third principal component identified an additional population structure that separated groups of predominantly IPPFBE or NPGS accessions. DAPC showed a considerable admixture of accessions from Azerbaijan, and a little to no admixture of accessions from Georgia and Tajikistan. These results suggest that central Asian/Caucasian almond germplasm is generally distinct from 'Nonpareil' and its relatives, and that although there is overlap between the NPGS and IPPFBE collections from this region, the IPPFBE collection does enhance the diversity of available almond germplasm.

Almonds [*Prunus dulcis* (Mill.) D.A. Webb] are perhaps the most economically important member of the genus *Prunus* L. The edible kernels of almonds have been consumed by humans since at least 11,000 BCE (Gradziel 2017). Current world production of almonds is estimated at 4.1 million tonnes and is increasing, having grown 60% since 2010 (Food and Agriculture Organization of the United Nations 2021). The United States (exclusively California) is the largest producer, and is responsible for more than half (58%) of world production. Other significant almond-growing nations include Spain, Australia, Iran, Turkey, Morocco, Syria, Italy, Tunisia, and Algeria (Food and Agriculture Organization of the United Nations 2021; US Department of Agriculture, National Agricultural Statistics Service 2021).

Almonds are consumed alone, raw or roasted; ground in water and strained to produce almond milk; or as an ingredient in confectioneries, desserts, and savory dishes (nonsweet). Although energy dense, almonds are low in saturated fat, and their consumption has been linked to lower levels of low-density lipoprotein, or "bad" cholesterol; a reduction in total cholesterol; and, most importantly, a reduction in cardiovascular disease [reviewed in Richardson et al. (2009)]. In addition, almonds are a source of minerals, including calcium, magnesium, phosphorus, manganese, and copper; and vitamins, including riboflavin and vitamin E (Kodad 2017).

As the demand for almonds grows, production is likely to expand beyond traditional areas where current cultivars are adapted. Limited water resources and climate instability are additional drivers of the need for new almond cultivars, including rootstocks adapted to saline conditions (Acharya et al. 2022). Plant genetic resources, in the form of germplasm collections, are a valuable component of breeding and cultivar development efforts because they can contain novel superior alleles for existing traits of interest (Bhullar et al. 2010; Sharma et al. 2021), or to address changes in abiotic stressors and pathogens (Batlle et al. 2017).

Received for publication 19 Jan 2023. Accepted for publication 21 Aug 2023. Published online 4 Oct 2023.
P.M. is the corresponding author. E-mail: phmccord@wsu.edu.

Almond germplasm collections are held in several countries, such as Italy (Rigoldi et al. 2015), Morocco (Hamzaoui et al. 2014), Portugal (Cabrita et al. 2014), Spain (Espiau et al. 2002), and Uzbekistan (Zaurov et al. 2015). In the United States, a public almond germplasm collection is maintained near Davis, CA, by the US Department of Agriculture, Agricultural Research Service, National Plant Germplasm System (NPGS). It currently contains 178 accessions of *P. dulcis*, including some that exist as small populations (with subaccession numbers) collected or donated originally as seed.

Almonds are believed to have originated in Central Asia (Gradziel 2017), and indeed the majority (n = 131) of the NPGS almond accessions are listed as coming from this part of the world. A number of wild almond species also exist, several of which are hypothesized to be progenitors of *P. dulcis* (Kester et al. 1990; Ladizinsky 1999; Zeinalabedini et al. 2010). Several of these species have been used in modern breeding programs to introgress desirable traits [reviewed in Socias i Company et al. (2017)]. The NPGS collection contains accessions (including seed-derived subaccessions) of *Prunus arabica* (Olivier) Meikle (n = 1 accession), *Prunus argentea* (Lam.) Reider (n = 1), *Prunus bucharica* (Korsh.) Hand.-Mazz (n = 19), *Prunus fenzliana* Fritsch (n = 38), *Prunus kuramica* (Korsh.) Kitam. (n = 4), *Prunus tangutica* (Baalin) Koehne (n = 2), *Prunus tenella* Batsch (n = 1), *Prunus triloba* Lindl. (n = 2), and *Prunus webbii* (Spach) Verh. (n = 6), as well as a small number of hybrids between *P. dulcis* and other species, including some of the wild ones just mentioned. The germplasm collections in Spain, Portugal, Italy, and Morocco have been evaluated for their genetic diversity (Cabrita et al. 2014; Kodad et al. 2008; Martí et al. 2009; Rigoldi et al. 2015), as have collections in a number of countries, such as Iran (Fathi et al. 2008) and Tunisia (Gouta et al. 2010). The majority of the accessions present in these collections represent cultivated (nonferal) *P. dulcis*, although a small number of wild accessions have been included, as well as wild almond species as outgroups. The genetic diversity of the NPGS collection was analyzed by Preece and Aradhya (2013), but their study included all *Prunus* species in the Davis collection, and did not focus specifically on almond (i.e., evaluate its intraspecific diversity).

Recently, a private collection of almond germplasm was made available for public research and development. This collection is also focused on germplasm of central Asian origin. To determine whether this private collection would enhance the diversity of available germplasm, we explored its genetic diversity and compared it to the diversity of the NPGS collection.

Materials and Methods

Private collection. The private collection consisted of 94 accessions (Supplemental Table 1) maintained as a research orchard in Thatcher, UT, USA. The majority of these accessions (n = 87) were collected in 2010 and 2011 as true seed (not budwood) from Central Asia, including Kyrgyzstan (n = 9 accessions), Tajikistan (n = 15), and Uzbekistan (n = 63). One of the accessions from Uzbekistan and three from Tajikistan were identified as *P. bucharica*; the rest were considered *P. dulcis*. Of the remaining seven accessions in the collection, one was a seedling of 'Nonpareil', a major commercial almond cultivar in the United States. The others were almond seedlings from unknown mother trees from a now-defunct nursery in Weber County, UT, USA. All accessions were grown as own-rooted seedlings. This germplasm collection was originally assembled by the Improving Perennial Plants for Food and Bioenergy (IPPFBE) Foundation, an organization founded to promote the development of perennial crops on the grazing lands of southern Idaho, USA, and northern Utah, USA. When the IPPFBE Foundation ceased operations in 2015, the germplasm collection (which included other tree fruit and nut species as well) was donated to Utah State University (Logan, UT, USA), which has maintained and evaluated the germplasm since then.

The NPGS collection. Sixty-six accessions from the NPGS almond germplasm collection (part of the National Clonal Germplasm Repository) near Davis, CA, USA, were also sampled by requesting leaf tissue through the Germplasm Resources Information Network (GRIN) (US Department of Agriculture, Agricultural Research Service 2022). These accessions represent 57 unique inventory numbers in the federal collection; six of the accessions sampled were represented by multiple (originally) seed-propagated individuals, and two to three subsamples were taken from each. They were chosen to represent germplasm from a similar geographic region as the foreign portion of the IPPFBE collection (Supplemental Table 1). In addition to the central Asian nations of Kazakhstan (n = 1 accession), Kyrgyzstan (n = 10), Turkmenistan (n = 3), and Uzbekistan (n = 15), accessions from Iran (n = 1), Pakistan (n = 3), and the Caucasus [Armenia (n = 2), Azerbaijan (n = 14), and Georgia (n = 17)] were also included. Although most of the NPGS accessions were identified as *P. dulcis*, two were identified as *P. fenzliana*, and five simply as *Prunus* sp.

DNA extraction and simple sequence repeat fingerprinting. DNA was extracted from leaf tissue via cetyltrimethylammonium bromide–based extraction (Murray and Thompson 1980) followed by phenol–chloroform purification. Fingerprints were generated using 10 simple sequence repeat (SSR) markers, a subset of 15 markers used by Preece and Aradhya (2013) in their overall analysis of *Prunus* diversity in the NPGS collection. All 15 markers were assayed, but two failed to amplify, one was monomorphic, and two produced patterns that could not be scored reliably. The forward primer of each primer pair was labeled at the 5′ end with 6-carboxyfluorescein fluorescent dye. Polymerase chain reaction (PCR) was performed in 10-μL volumes and consisted of 20 ng template DNA, 0.25 μM (combined) forward and reverse primers, and 1× Platinum II Hot-Start PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA). PCR conditions consisted of 94 °C for 2 min, followed by 35 cycles of denaturation at 94 °C for 15 s, primer annealing at 60 °C for 15 s, and product extension at 68 °C for 15 s, with a final extension at 68 °C for 10 min and an indefinite hold at 4 °C. Two of the markers (BPPCT002 and pchgms1) were amplified using a two-step amplification protocol as follows: initial denaturation at 94 °C for 2 min, followed by 35 cycles of denaturation at 98 °C for 5 s, and a combination annealing/extension at 60 °C for 15 to 20 s, with no final extension step. Post-PCR, the samples were analyzed via capillary electrophoresis using a capillary analyzer (Applied Biosystems 3730, Thermo Fisher Scientific) and the LIZ500 size standard (Thermo Fisher Scientific). The electropherograms were scored using the *Fragman* package (Covarrubias-Pazaran et al. 2016) in R software ver. 4.x (R Foundation for Statistical Computing, Vienna, Austria).

Fingerprint analysis. Missing genotypes (1.6%) were first imputed as the mean of the population using the *tab* function in the R package *adegenet* (Jombart 2008). The imputed dataset of

SSR genotypes was converted to a Euclidean distance matrix, and a dendrogram was generated via hierarchical clustering (complete linkage method) using *hclust* in base R. Principal component analysis (PCA) and discriminant analysis of principal components (DAPC) were performed via *adegenet*. DAPC is a relatively new technique (Jombart et al. 2010) that performs discriminant analysis using principal components (PCs) as input variables. Traditional PCA seeks to maximize overall variance; DAPC focuses on maximizing between-group variation, which is ideal for exploring the structure of populations. Before these analyses, variables were scaled via the *scaleGen* function. Because DAPC is a classification procedure, groups must be specified before running the analysis. The countries/states of origin for each accession were chosen as the original groups. The admixture of individuals was explored using bar plots generated by the *compoplot* function, which generates bar plots comparable to those produced by STRUCTURE software (Pritchard et al. 2000). K-means clustering was also used to choose the optimal number of groups for DAPC, with the choice of K governed by examining a scree plot of within-cluster sums of squares. In this scenario, the loadings of the original input variables (SSR alleles) on the discriminant functions were also examined. The Genome Database for Rosaceae (GDR) (Jung et al. 2019) was used to search for agronomic traits or candidate genes near SSRs with high loading.

Results and Discussion

Marker performance. A total of 182 alleles was generated by 10 SSR markers. The number of alleles per marker ranged from 2 to 30 (Table 1). A slightly larger number of alleles was found in the NPGS germplasm vs. the IPPFBE collection (152 vs. 149), which seems reasonable given the larger geographic area represented by the NPGS collection. Observed heterozygosity was significantly lower than expected heterozygosity for all markers as a group ($P < 0.001$, paired t test). Markers UDP98-409 and BPPCT040 were particularly less heterozygous than expected.

Hierarchical clustering. Hierarchical clustering generated a dendrogram with seven main branches (Fig. 1). The first branch contained only two *P. bucharica* accessions. UT-PD-64 was sourced from the Forestry Institute in the Gissar Valley of Tajikistan, whereas UT-PD-68 was collected in the wild from Surkhandarya Province in Uzbekistan. None of the other accessions identified as wild almond species (*P. bucharica*, *P. fenzliana*) was separated from *P. dulcis* material. The second branch, with 16 accessions, included seven from Uzbekistan (six from the IPPFBE collection), three each from Kyrgyzstan and Tajikistan, two from Georgia, and one from Azerbaijan. The third branch of the dendrogram was the smallest branch, with only *P. dulcis* accessions (n = 7) and consisted solely of individuals from the IPPFBE collection, including the seedling of 'Nonpareil' (UT-PD-04) and four of the Utah, USA–sourced accessions. A closer inspection of this branch revealed that UT-PD-06 and UT-PD-09 had identical SSR fingerprints. Given that UT-PD-06 is from Utah and has bitter kernels and UT-PD-09 is supposedly from Uzbekistan and has sweet kernels, this is likely the result of a labeling error during sample collection, with the same sample being collected twice. In addition, UT-PD-57 and UT-PD-58, both from Surkhandarya Province in Uzbekistan, differed by only one allele.

Table 1. Simple sequence repeat markers used to fingerprint the almond germplasm collections.

Marker	*Prunus* chromosome no.	Alleles (n)	Expected heterozygosity	Observed heterozygosity
BPPCT002	2	18	0.85	0.64
BPPCT004	2	22	0.76	0.61
BPPCT014	5	12	0.74	0.5
BPPCT039	3	30	0.93	0.73
BPPCT040	4	14	0.84	0.52
CPSCT026	7	2	0.05	0.03
CPSCT042	7	16	0.86	0.61
pchgms1	2	18	0.86	0.73
UDP96-003	4	27	0.92	0.69
UDP98-409	8	23	0.91	0.45

The fourth branch, with 15 accessions, contained two NPGS accessions from Armenia. One (Cal-PD-24/DPRU 2941) is described as *P. fenzliana*, whereas the other (Cal-PD-23/DPRU 2773-002A), although listed in GRIN as *P. dulcis*, includes the notation "looks like *fenzliana*" in its passport documentation. In addition, the fourth branch included eight accessions from Uzbekistan (five from IPPFBE), two from Kyrgyzstan, and a single accession each from Azerbaijan, Georgia, and Turkmenistan. The fifth branch of the tree was the third largest, with 31 accessions. Sixteen of these were collected in Uzbekistan (n = 11 from the IPPFBE collection). Nearly half (n = 7) of the IPPFBE accessions from Tajikistan were present, as well as three from Kyrgyzstan (n = 2 from IPPFBE), two each from Azerbaijan and Georgia, and one accession from Pakistan. The sixth and largest branch, with 49 accessions, was the only branch with a majority of individuals from the NPGS collection (n = 26). A significant number of these were from Georgia (n = 11), followed by Azerbaijan (n = 6), Kyrgyzstan (n = 5), Pakistan (n = 2), and one each from Iran and Turkmenistan. The IPPFBE accessions in this branch were mostly from Uzbekistan (n = 19), with an additional four from Kyrgyzstan.

The seventh branch of the tree was the second largest and contained 40 accessions. Three nursery-sourced accessions from northern Utah (UT-PD-01, -05, and -07) were in this branch, although they do not appear to be closely related to each other. Of the remaining accessions, 25 were from Uzbekistan (n = 19 from IPPFBE). Four were from Azerbaijan (NPGS), including one described as *P. fenzliana*. Three were from Tajikistan (all IPPFBE), two from Kyrgyzstan (NPGS), and a single accession each was from Georgia, Kazakhstan, and Turkmenistan (all NPGS).

Principal components analysis. Plotting the first two PCs (Fig. 2) separated the wild species *P. bucharica* and *P. fenzliana* from the remaining accessions. Cal-PD-52 (DPRU 3012) is described as *P. dulcis*, including in its passport information, but clustered with the *P. fenzliana* accessions. In contrast, the *P. fenzliana* accession from Azerbaijan (Cal-PD-31/DPRU 3206-002A) clustered with the *P. dulcis* accessions. When these wild species were removed and the PCA was rerun, plotting the first two PCs separated the 'Nonpareil' seedling, all Utah, USA–sourced accessions, and 13 additional IPPFBE accessions (all but one from Uzbekistan), from nearly all other accessions with a first PC > 3 (Fig. 3). In addition, the accession from Iran (Cal-PD-07/DPRU 2374) and one of the accessions from Turkmenistan (Cal-PD-06/DPRU 2306-0004A) were separated from all others by the second PC (> 10), whereas an NPGS accession from

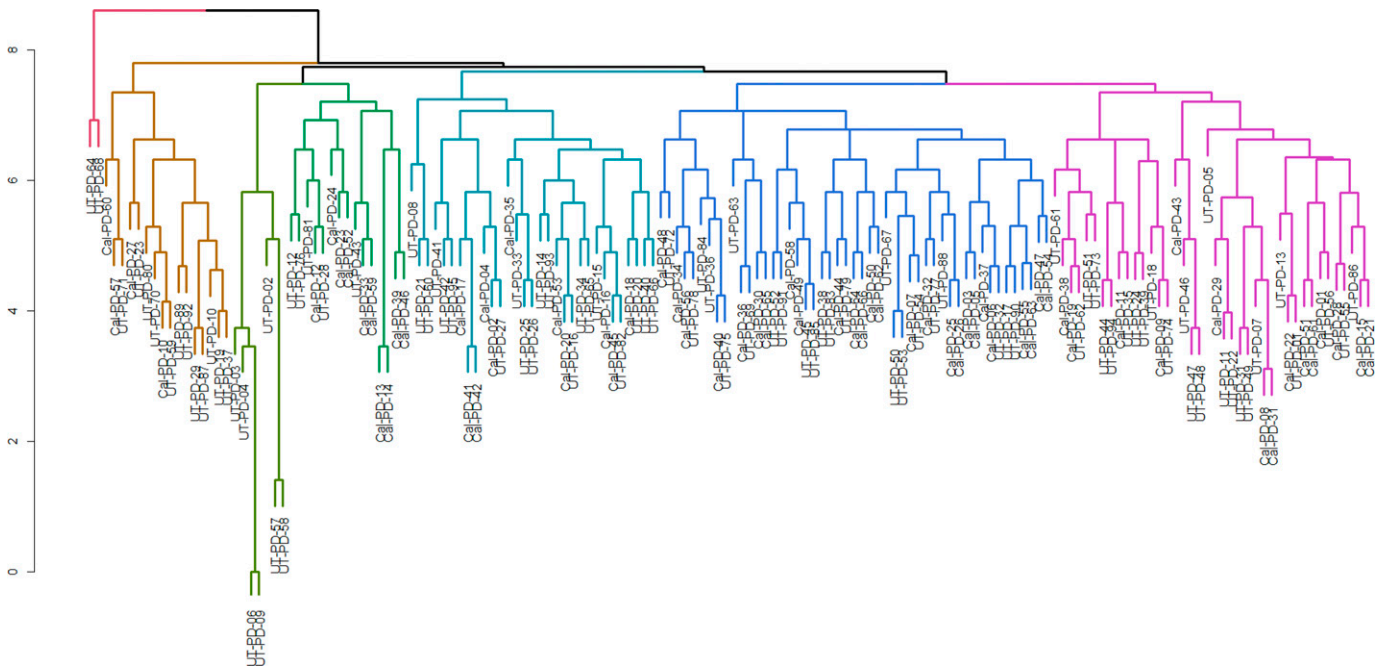


Fig. 1. Hierarchical clustering of Improving Perennial Plants for Food and Bioenergy (IPPFBE) Foundation and National Plant Germplasm System (NPGS) almond accessions. IPPFBE accessions have the prefix UT-; NPGS accessions have the prefix Cal-.

Kyrgyzstan (Cal-PD-47/DPRU 3003) and an IPPFBE accession from Uzbekistan (UT-PD-81) were separated from the main body in the opposite direction (PC < –5). When plotting the first and third PCs (Fig. 4), accessions with a third PC < –2 were mostly (15 of 20) from the NPGS collection from multiple countries, whereas accessions with a third PC > 2 were primarily from the IPPFBE collection (16 of 19).

Discriminant analysis of principal components. A biplot of the first two discriminant functions using all *P. dulcis* accessions identified the single Iranian accession (Cal-PD-07/DPRU 2374-0012A) as an extreme outlier (data not shown). Reanalysis without this accession (or the *P. bucharica* and *P. fenzliana* accessions) assigned nearly all accessions (89.6%) to their original classification (country of origin). The biplot of the first two

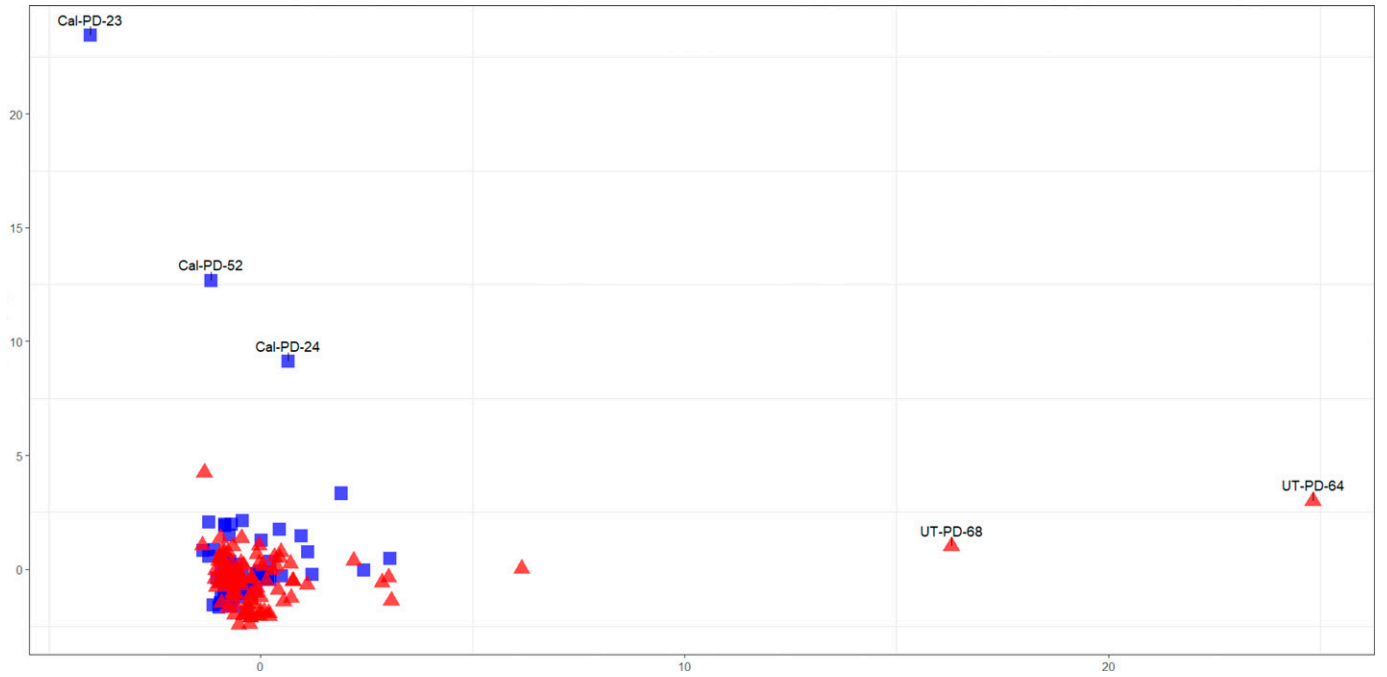


Fig. 2. Plot of the first two principal components (PCs) from a principal components analysis, with the first PC on the *x*-axis. Individuals from the National Plant Germplasm System collection are denoted by blue squares, and Improving Perennial Plants for Food and Bioenergy accessions, red triangles. UT-PD-64 and -68 are *Prunus bucharica*; Cal-PD-23, -24, and -52 are *Prunus fenzliana*.

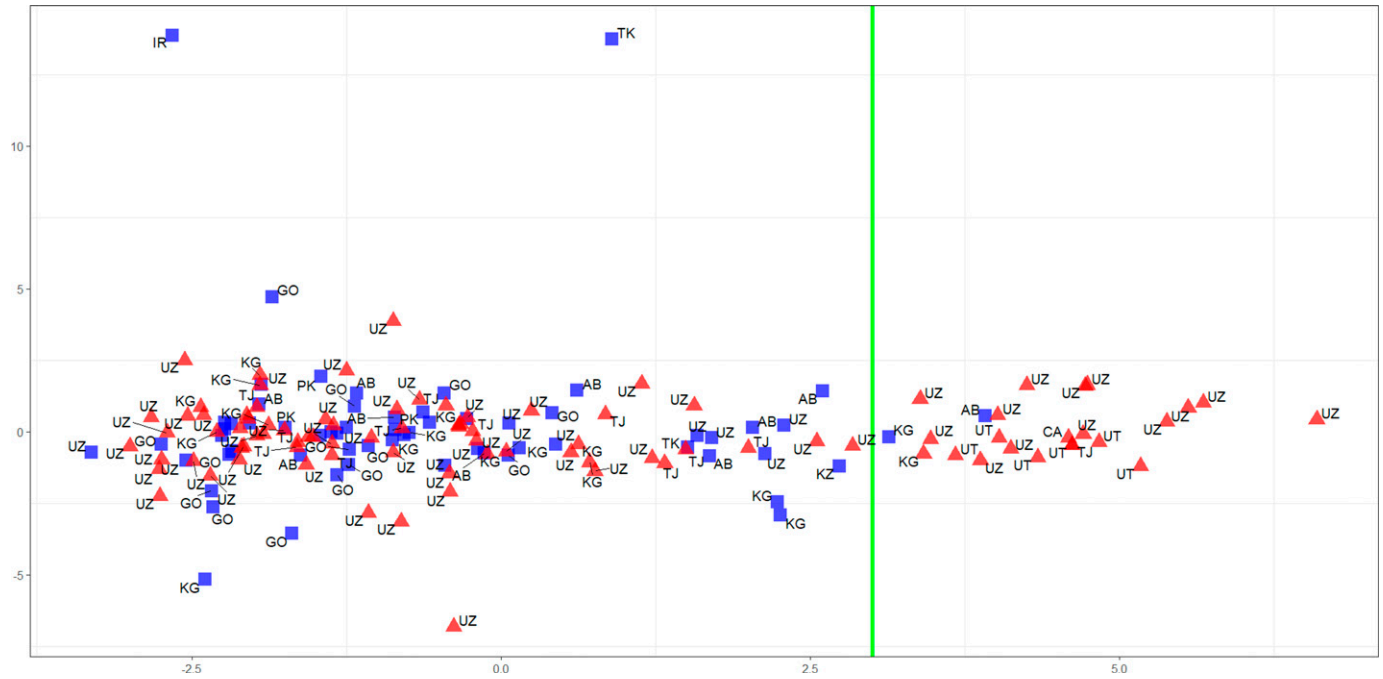


Fig. 3. Plot of first two principal components, with *Prunus bucharica* and *Prunus fenzliana* accessions removed. Individuals from the National Plant Germplasm System collection are denoted by blue squares, and Improving Perennial Plants for Food and Bioenergy accessions, red triangles. Two-letter codes are used for the country (or state) of origin: AB = Azerbaijan; CA = California, USA; GO = Georgia; IR = Iran; KG = Kyrgyzstan; KZ = Kazakhstan; PK = Pakistan; TJ = Tajikistan; TK = Turkmenistan; UT = Utah, USA; UZ = Uzbekistan.

discriminant functions (Fig. 5) suggested that the NPGS accessions from Georgia, Kazakhstan, Pakistan, and Turkmenistan were the most distinct. The bar plots of group membership probabilities (Figs. 6 and 7) showed that accessions from Kyrgyzstan and Azerbaijan were the most admixed, frequently with each other, as well as with populations from Uzbekistan. The connection with Uzbekistan is not surprising, given that the names of the mother trees of some of the IPPFBE accessions from Kyrgyzstan (UT-PD-87 and UT-PD-90, for example) suggest these trees originated in neighboring Uzbekistan. Given that Azerbaijan and Kyrgyzstan are separated by a distance of more than 1200 km, including the Caspian Sea, the admixture is striking and suggests intentional

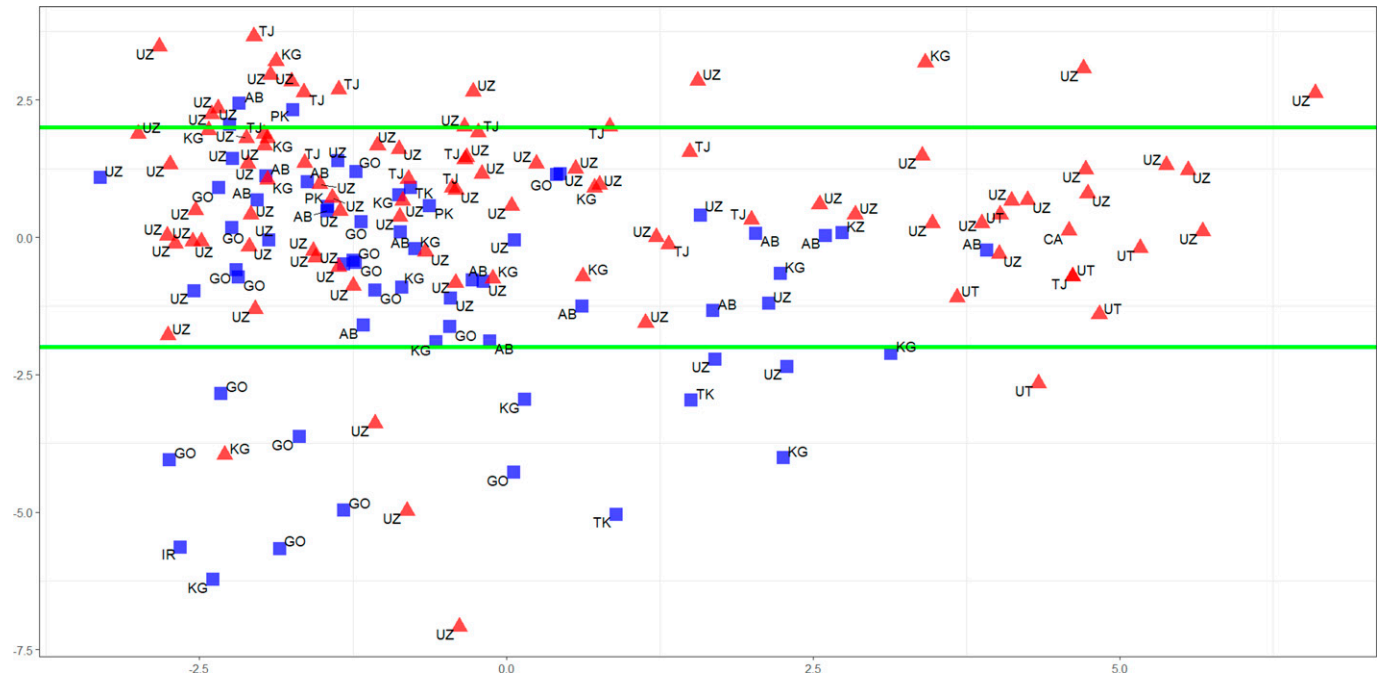


Fig. 4. Plot of first and third principal components of the National Plant Germplasm System (blue squares) and Improving Perennial Plants for Food and Bioenergy (red triangles) accessions, with *Prunus bucharica* and outliers removed. Two-letter codes are used for the country (or state) of origin: AB = Azerbaijan; CA = California, USA; GO = Georgia; IR = Iran; KG = Kyrgyzstan; KZ = Kazakhstan; PK = Pakistan; TJ = Tajikistan; TK = Turkmenistan; UT = Utah, USA; UZ = Uzbekistan.

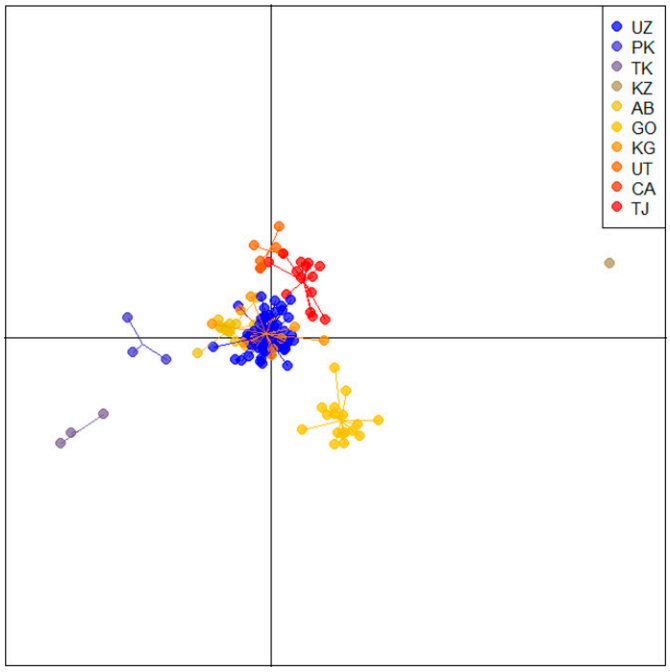


Fig. 5. Scatterplot of the first two discriminant functions based on discriminant analysis of principal components. *Prunus bucharica*, *Prunus fenzliana*, and the single Iranian accession were removed. Two-letter codes are used for the country (or state) of origin: AB = Azerbaijan; CA = California, USA; GO = Georgia; KG = Kyrgyzstan; KZ = Kazakhstan; PK = Pakistan; TJ = Tajikistan; TK = Turkmenistan; UT = Utah, USA; UZ = Uzbekistan.

germplasm exchange, likely during the Soviet period. Although no IPPFBE accessions were collected from Azerbaijan, several accessions, such as UT-PD-92 (collected in Kyrgyzstan) and UT-PD-43 (collected in Uzbekistan) appear to be Azerbaijani in origin. Similarly, although Cal-PD-39/DPRU 2992-0007A and Cal-PD-40/DPRU 2993-0007A were collected in Azerbaijan, both appear to be largely derived from Uzbekistani populations. The least admixed of the larger populations (represented by more than three accessions) were the Georgian accessions from the NPGS collection, and the Tajikistani accessions from the IPPFBE collection. The Georgian accessions showed no evidence of admixture. Most of the accessions from Tajikistan were also pure, with two accessions (UT-PD-19 and UT-PD-21) showing some admixture with populations from Kyrgyzstan and Uzbekistan. Similar to the hierarchical clustering results, which suggested a close relationship, the Utah, USA–sourced accessions and the 'Nonpareil' seedling were admixed with each other. That the Weber County nursery was a source of diverse material (Roper T, unpublished) is supported by the fact that at least one of the Utah accessions, UT-PD-07, shows significant admixture with the Tajikistani population.

Using K-means clustering to identify groups before DAPC suggested five optimal clusters (data not shown). A plot of the loadings of the original SSR alleles on the discriminant functions suggested that alleles from BPPCT002 and pchgms1 had the strongest effect on clustering (Supplemental Fig. 1). A marker search of GDR showed that BPPCT002 lies very close (within 0.1 cM) to the *SH* shell-hardness locus (Arús et al. 1998). Although germplasm in both the IPPFBE and NPGS collections show considerable diversity for this trait, the data are incomplete (Supplemental Table 1), and further phenotypic characterization is needed to confirm trait associations with particular DNA markers.

Conclusion

Almonds are an economically important and healthful crop with acreage that is expanding. This expansion encourages the development of new cultivars adapted to new growing areas, and novel germplasm may be the source of alleles needed for adaptation in these locations. The NPGS and IPPFBE germplasm collections both contain numerous accessions from the almond center of origin. Analysis of the IPPFBE collection indicates that

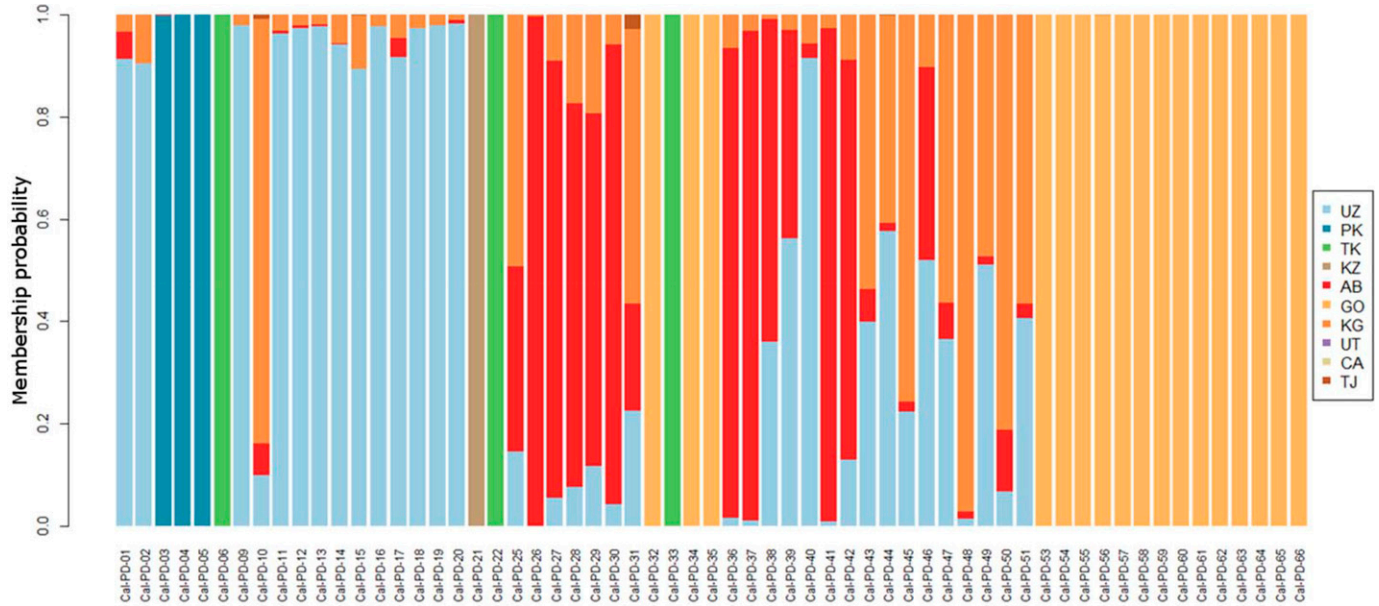


Fig. 6. Bar plot of group (country of origin) membership probabilities for each accession from the National Plant Germplasm System collection (not including *Prunus fenzliana* or the Iranian accessions). Two-letter codes are used for the country of origin: AB = Azerbaijan; CA = California, USA; GO = Georgia; KG = Kyrgyzstan; KZ = Kazakhstan; PK = Pakistan; TJ = Tajikistan; TK = Turkmenistan; UT = Utah, USA; UZ = Uzbekistan.

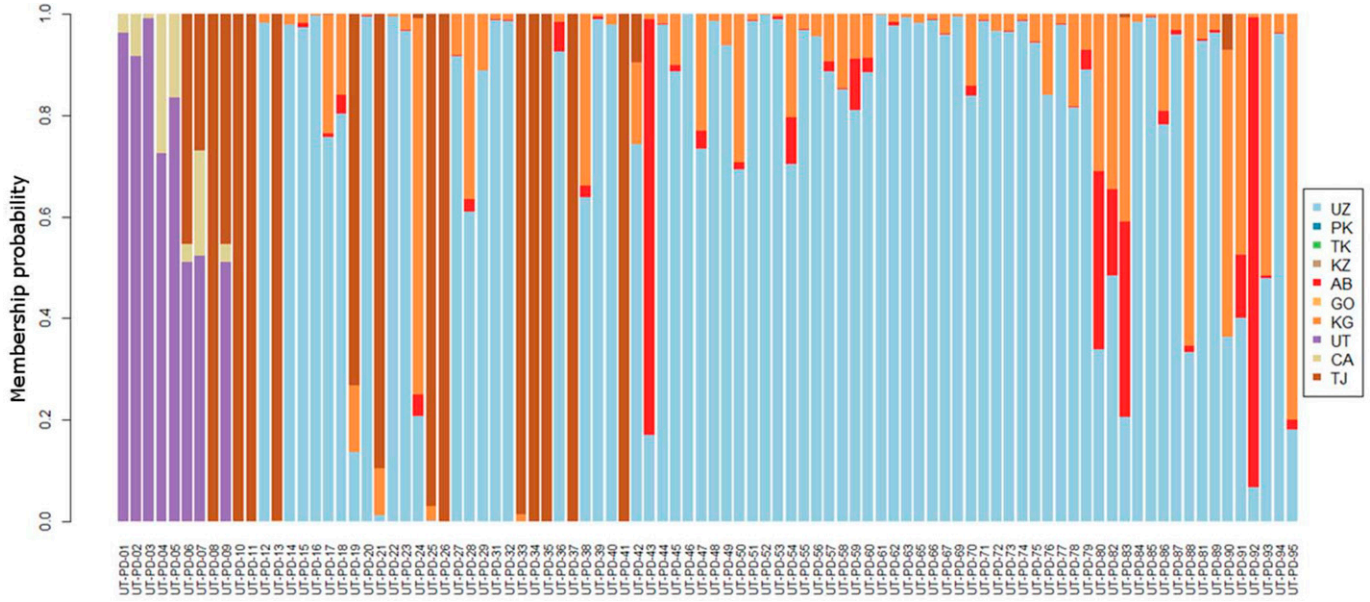


Fig. 7. Bar plot of group (country of origin) membership probabilities for each accession from the Improving Perennial Plants for Food and Bioenergy collection (not including *Prunus bucharica* accessions). Two-letter codes are used for the country (or state) of origin: AB = Azerbaijan; CA = California, USA; GO = Georgia; KG = Kyrgyzstan; KZ = Kazakhstan; PK = Pakistan; TJ = Tajikistan; TK = Turkmenistan; UT = Utah, USA; UZ = Uzbekistan.

although it and the NPGS collection are broadly similar, the new material contributes positively to the genetic diversity of almond germplasm. This is particularly true for material from Tajikistan and Uzbekistan. Although there are a number of accessions from Uzbekistan in the NPGS collection, IPPFBE accessions expand the available diversity from this country. With respect to Tajikistan, only one *Prunus* accession (a seedling family with subaccessions) from this country is currently listed in the NPGS collection, and its taxonomy is unclear (listed as *Prunus* sp.). In addition to these two countries, Iran also appears to contain genetically distinct almond germplasm, and efforts should be undertaken to acquire additional Iranian accessions. The 'Nonpareil'-derived germplasm is distinct from most accessions collected from Central Asia and the Caucasus. Its survival and fruiting in the high elevation (~1200 m) of northern Utah, USA, suggests it also contains valuable alleles for successful almond cultivation in severe climates. Phenotypic and genotypic analysis of germplasm in both collections should be expanded to identify novel superior alleles for breeding high-quality, resilient almonds.

References Cited

Acharya BR, Sandhu D, Dueñas C, Pudussery M, Kaundal A, Ferreira JFS, Suarez DL, Skaggs TH. 2022. Morphological, physiological, biochemical, and transcriptome studies reveal the importance of transporters and stress signaling pathways during salinity stress in *Prunus*. Sci Rep. 12:1274. https://doi.org/10.1038/s41598-022-05202-1.

Arús P, Ballester J, Jáuregui B, Joobeur T, Truco M, de Vicente MC. 1998. The European *Prunus* mapping project: Update on marker development in almond. Acta Hortic. 484:331–338. https://doi.org/10.17660/ActaHortic.1998.484.57.

Batlle I, Dicenta F, Socias i Company R, Gradziel TM, Wirthensohn M, Duval H, Vargas F. 2017. Classical genetics and breeding, p 111–148. In: Socias i Company R, Gradziel TM (eds). Almonds: Botany, production, and uses. CABI, Wallingford, Oxfordshire, UK. https://doi.org/10.1079/9781780643540.0000.

Bhullar NK, Zhang Z, Wicker T, Keller B. 2010. Wheat gene bank accessions as a source of new alleles of the powdery mildew resistance gene *Pm3*: A large scale allele mining project. BMC Plant Biol. 10:88. https://doi.org/10.1186/1471-2229-10-88.

Cabrita L, Apostolova E, Neves A, Marreiros A, Leitão J. 2014. Genetic diversity assessment of the almond [*Prunus dulcis* (Mill) D.A. Webb] traditional germplasm of Algarve, Portugal, using molecular markers. Plant Genet Resour. 12:S164–S167. https://doi.org/10.1017/S1479262114000471.

Covarrubias-Pazaran G, Díaz-García L, Schautman B, Salazar W, Zalapa J. 2016. Fragman: An R package for fragment analysis. BMC Genet. 17:62. https://doi.org/10.1186/s12863-016-0365-6.

Espiau, MT, Ansón JM, Socias i Company R. 2002. The almond germplasm bank of Zaragoza. Acta Hortic. 591:275–278. https://doi.org/10.17660/ActaHortic.2002.591.42.

Fathi A, Ghareyazi B, Haghnarazi A, Ghaffari MR, Pirseyedi SM, Kadkhodei S, Naghavi MR, Mardi M. 2008. Assessment of the genetic diversity of almond (*Prunus dulcis*) using microsatellite markers and morphological traits. Iran J Biotechnol. 6(2):98–106.

Food and Agriculture Organization of the United Nations. 2021. Crops and livestock products: Almonds, in shell [production quantity, world + (total)]. https://www.fao.org/faostat/en/#data/QCL. [accessed 2 Dec 2022].

Gouta H, Ksia E, Buhner T, Moreno MÁ, Zarrouk M, Mliki A, Gogorcena Y. 2010. Assessment of genetic diversity and relatedness among Tunisian almond germplasm using SSR markers. Hereditas. 147(6):283–292. https://doi.org/10.1111/j.1601-5223.2009.02147.x.

Gradziel TM. 2017. History of cultivation, p 43–69. In: Socias i Company R, Gradziel TM (eds). Almonds: Botany, production, and uses. CABI, Wallingford, Oxfordshire, UK. https://doi.org/10.1079/9781780643540.0000.

Hamzaoui AE, Oukabli A, Moumni M. 2014. Morphological and molecular diversity and genetic structure of Moroccan cultivated almond (*Prunus dulcis* Mill.) beside some foreign varieties. Plant Genet Resour. 12(3):308–316. https://doi.org/10.1017/S1479262114000094.

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, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: A new method for the analysis of structured populations. BMC Genet. https://doi.org/10.1186/1471-2156-11-94.

Jung S, Lee T, Cheng CH, Buble K, Zheng P, Yu J, Humann J, Ficklin SP, Gasic K, Scott K, Frank M, Ru S, Hough H, Evans K, Peace C, Olmstead M, DeVetter LW, McFerson J, Coe M, Wegrzyn JL, Staton ME, Abbott AG, Main D. 2019. 15 Years of GDR: New data and functionality in the Genome Database for Rosaceae. Nucl Acid Res. 47:D1137–D1145. https://doi.org/10.1093/nar/gky1000.

Kester DE, Gradziel TM, Grasselly C. 1990. Almonds (*Prunus*). Acta Hortic. 290:701–760. https://doi.org/10.17660/ActaHortic.1991.290.16.

Kodad O. 2017. Chemical composition of almond nuts, p 428–448. In: Socias i Company R, Gradziel TM (eds). Almonds: Botany, production, and uses. CABI, Wallingford, Oxfordshire, UK. https://doi.org/10.1079/9781780643540.0000.

Kodad O, Alonso JM, Sánchez A, Oliveira MM, Socias I Company R. 2008. Evaluation of genetic diversity of *S*-alleles in an almond germplasm collection. J Hortic Sci Biotechnol. 83(5):603–608. https://doi.org/10.1080/14620316.2008.11512430.

Ladizinsky G. 1999. On the origin of almond. Genet Resources Crop Evol. 46:143–147. https://doi.org/10.1023/A:1008690409554.

Martí ÀF, Alonso JM, Espiau MT, Rubio-Cabetas MJ, Socias I Company R. 2009. Genetic diversity in Spanish and foreign almond germplasm assessed by molecular characterization with simple sequence repeats. J Am Soc Hortic Sci. 134:535–542. https://doi.org/10.21273/JASHS.134.5.535.

Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. Nucl Acids Res. 8(19):4321–4326. https://doi.org/10.1093/nar/8.19.4321.

Preece JE, Aradhya M. 2013. The *Prunus* collection at the National Clonal Germplasm Repository in Davis, California. Acta Hortic. 985:47–54. https://doi.org/10.17660/ActaHortic.2013.985.5.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics. 155:945–959. https://doi.org/10.1093/genetics/155.2.945.

Richardson DP, Astrup A, Cocaul A, Ellis P. 2009. The nutritional and health benefits of almonds: A healthy food choice. Food Sci Technol Bull. 6:41–50. https://doi.org/10.1616/1476-2137.15765.

Rigoldi MP, Rapposelli E, De Giorgio D, Resta P, Porceddu A. 2015. Genetic diversity in two Italian almond collections. Electron J Biotechnol. 18(1):40–45. https://doi.org/10.1016/j.ejbt.2014.11.006.

Sharma S, Schulthess AW, Bassi FM, Badaeva ED, Neumann K, Graner A, Özkan H, Werner P, Knüpffer H, Kilian B. 2021. Introducing beneficial alleles from plant genetic resources into the wheat germplasm. Biology (Basel). 10(10):982. https://doi.org/10.3390/biology10100982.

Socias i Company R, Ansón JM, Espiau MT. 2017. Taxonomy, botany and physiology, p 1–42. In: Socias i Company R, Gradziel TM (eds). Almonds: Botany, production, and uses. CABI, Wallingford, Oxfordshire, UK. https://doi.org/10.1079/9781780643540.0000.

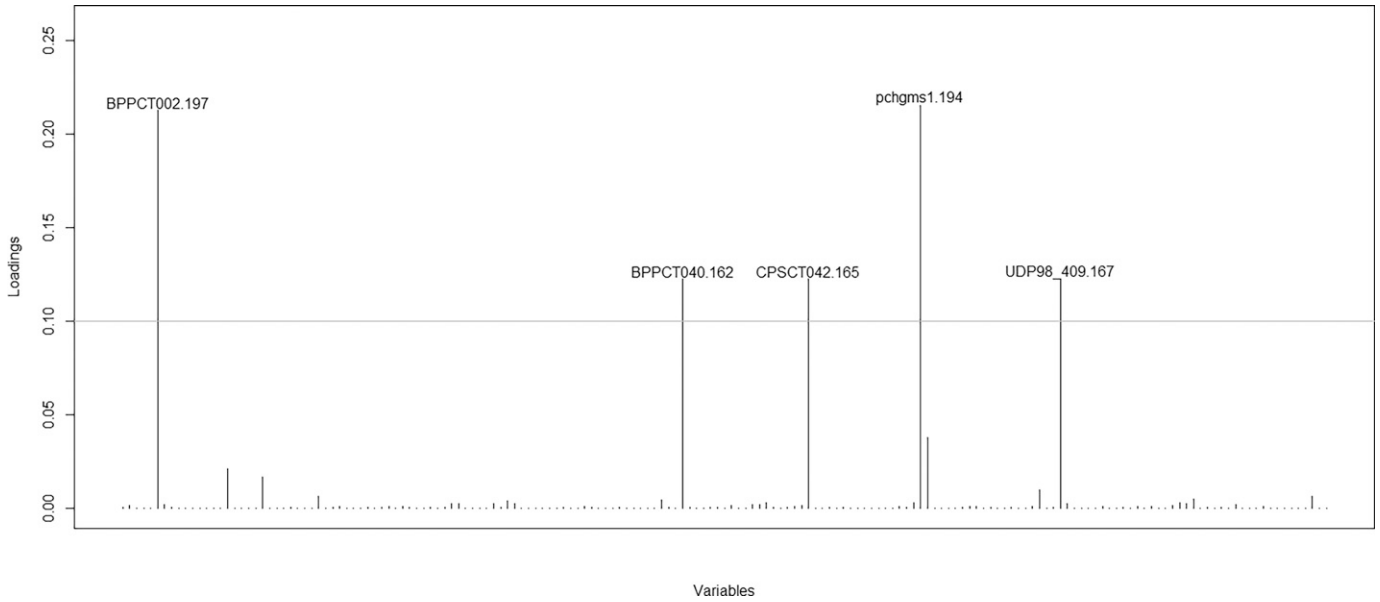
US Department of Agriculture, Agricultural Research Service. 2022. Germplasm Resources Information Network (GRIN). https://npgsweb.ars-grin.gov/gringlobal/search. [accessed 2 Dec 2022].

US Department of Agriculture, National Agricultural Statistics Service. 2021. Noncitrus fruits and nuts 2021 summary. https://usda.library.cornell.edu/concern/publications/zs25x846c. [accessed 2 Dec 2022].

Zaurov DE, Eisenman SW, Ford T, Khokhlov S, Kenjebaev S, Shalpykov KT, Funk CR. 2015. Genetic resources of almond species in the former USSR. HortScience. 50(1):18–29. https://doi.org/10.21273/HORTSCI.50.1.18.

Zeinalabedini M, Khayam-Nekoui M, Grigorian V, Gradziel TM, Martínez-Gómez P. 2010. The origin and dissemination of the cultivated almond as determined by nuclear and chloroplast SSR marker analysis. Sci Hortic. 125(4):593–601. https://doi.org/10.1016/j.scienta.2010.05.007.

Supplemental Information



Supplemental Fig. 1. Plot displaying the loading of the original simple sequence repeat (SSR) alleles on the discriminant functions of the discriminant analysis of principal components. *Prunus bucharica* and *Prunus fenzliana* accessions were excluded. Alleles are labeled by the name of the SSR marker, a period, and the size of the allele in bases.