

Germplasm characterization and genetic structure of Ukrainian sweet cherry (*Prunus avium* L.) assessed by microsatellite and S-locus markers

Yaroslav I. Ivanovych

yaroslav.ivanovych@slu.se

Swedish University of Agricultural Sciences

Research Article

Keywords: cultivar identification, microsatellite markers, ex situ conservation, *Prunus avium*, self-incompatibility, Ukraine

Posted Date: October 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11236353/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Ukrainian sweet cherry (*Prunus avium* L.) germplasm is among the larger national collections in Europe, yet its cultivar identities and recorded pedigrees have not been tested against molecular evidence, and the collections holding it are at risk. Ninety-four accessions were genotyped at 18 microsatellite (SSR) loci and three self-incompatibility (S) locus markers, with allele sizes at six shared loci calibrated to the European reference scale to permit direct cross-country comparison — the first such calibration applied to this germplasm. Identity testing resolved five duplicate cultivar names into three homonym pairs, two clonal pairs, and identified two additional accessions as mislabeled propagations, leaving 90 unique genotypes. Comparison with European reference collections revealed a reciprocal exchange of identity between the founder cultivars 'Drogans gelbe Knorpelkirsche' and 'Dönissens gelbe Knorpelkirsche' in both Ukrainian and European holdings. S-genotypes were determined for the 94 accessions and 17 duplicates, yielding 19 S-alleles, resolving six of eleven conflicts in the published record, and providing first records for 33 Ukrainian cultivars and selections. Parentage testing confirmed three pedigrees in full and one parent in 25 further pedigrees, rejected 13, and proposed four new assignments; 'Drogans gelbe Knorpel' was confirmed as the parent of 17 pedigrees, and two distinct genotypes were shown to have been propagated under the name 'Valerii Chkalov'. Genetic structure reflected a narrow founder base. Cultivar labels must not be assumed correct without verification, and *ex situ* conservation necessitates documentation at the individual-tree level, supported by convergent evidence from SSR, S-locus, and parentage.

1. Introduction

Sweet cherry (*Prunus avium* L.) is among the most economically significant stone fruit crops in southern Ukraine and has one of the country's longest breeding histories. Systematic breeding at Melitopol began in 1933, yielding more than 90 cultivars, 45 of which are listed in the State Register (Turovtsev, Turovtseva et al. 2011). The Bakhmut (Artemivsk) program developed more than 50 stone-fruit cultivars (Taranenko 2009, Taranenko 2011), 27 of which were sweet cherry. Additional cultivars originated from the Mliiv Research Station of Pomology, the Institute of Horticulture, and the Nikita Botanical Garden in Crimea. A compilation of the State Register, breeders' publications, Ukrainian pomological literature (Symyrenko 1963, Turovtsev, Taranenko et al. 2004), and genebank passport data reports 282 named sweet cherry cultivars, selections, and landraces of Ukrainian origin. Of these, 180 are linked to the Institute of Horticulture NAAS network, while 77 are associated with the Nikita Botanical Garden. Crimean landraces form a small but unique group, consisting of about twenty locally selected forms, mostly first recorded in the late nineteenth century (Symyrenko 1963, Lukicheva 2010).

No single national source documents this resource comprehensively. The State Register of Plant Varieties Suitable for Dissemination in Ukraine functions mainly as a commercial tool rather than a conservation list: it includes varieties approved for commercial planting, with entries maintained through an annual fee paid by rights holders (UIPVE 2026). Combining current and lapsed entries, the register has historically listed about one-third of Ukrainian sweet cherry cultivars. Similarly, genebank records are incomplete: EURISCO's Ukrainian inventory records 486 *P. avium* accessions, with 99 originally entered under the obsolete name *Cerasus avium* and not retrievable via the accepted binomial. Of these, 355 are Ukrainian: 180 as advanced cultivars and 143 as hybrid material, with eight semi-natural and one traditional cultivar or landrace. As a result, Crimean landraces mentioned in pomological literature are virtually absent from the inventory.

This material is now at critical risk. The Nikita Botanical Garden collection, which includes over 400 accessions from 19 countries and is primarily composed of the Crimean and central Ukrainian eco-geographical groups — accounting for just over half—was lost to Ukraine when Crimea was annexed in 2014 (Lukicheva 2010). The Melitopol Research

Station of Horticulture, home to 123 sweet cherry and 114 sour cherry accessions from at least 35 countries, has been under occupation since February 2022, and scientific activities have halted. The Bakhmut (Artemivsk) Nursery Research Station, which maintained 280 cherry accessions as part of Ukraine's National Bank of Plant Genetic Resources, has been destroyed and cannot be revived. Some genotypes have likely been lost permanently and will need to be recovered from private collections or reimported from abroad (Udovychenko, Grynyk et al. 2024). Approximately one-third of Ukraine's sweet cherry orchard territory is currently under occupation (Udovychenko, Grynyk et al. 2024), including the "Melitopol cherry" production region, which was officially registered as a Ukrainian geographical indication in September 2020 (Association 2020). However, it was subsequently appropriated by the Russian patent office in July 2024 (UANPIO 2024). As a consequence, material stored outside Ukraine has sometimes become the more secure copy — an issue foreseen in 1999 when L.I. Taranenko personally transported budwood from Donetsk to Kyiv for export under a signed Memorandum of Understanding, doubting the survival of her work in place (Iezzoni 2005).

Recovery of that kind depends entirely on the names attached to the material. For Ukrainian sweet cherry, those names carry at least four independent classes of error: incorrect or missing country of origin; Soviet-era Russian-language names retained as accepted names, compounded by transliteration and, in some cases, translation into third languages; propagation and labeling errors that produce homonyms, duplicates, and mislabeled trees; and recorded pedigrees that cannot be verified, either because the stated parent was never registered or because the parental tree no longer exists. These classes compound, and molecular datasets inherit all of them: a genotype is only as reliable as the identity of the tree from which it was sampled. These issues relate to documentation flaws, not conservation. The material might last, but essential information for use could be missing.

Self-incompatibility data alone tends to amplify rather than resolve this issue. The S-genotype determines the cross-incompatibility group and therefore the choice of pollinizer, so an S-genotype assigned to a misidentified tree yields an incorrect orchard recommendation. For 'Valerii Chkalov', the published record reports S_1S_9 , S_1S_6 and S_3S_9 — three genotypes that fall into three different incompatibility groups. When such determinations derive from accessions whose identity was never independently verified, the S-genotype literature propagates labeling error under the guise of molecular precision.

A European reference framework exists (Campoy, Lerigoleur-Balsemin et al. 2016, Barreneche, Cárcamo de la Concepción et al. 2021, Ordidge, Litthauer et al. 2021), based on the SSR panel recommended by the ECPGR Working Group on *Prunus* (Clarke and Tobutt 2009). Ukrainian material has been minimally integrated into this framework, mainly through third-country holdings, which raises questions about provenance and nomenclatural accuracy.

The molecular characterization of Ukrainian fruit and grapevine germplasm has only recently begun and remains limited (Bocharova, Kovaliova et al. 2009, Risovannaia, Gorislavets et al. 2014, Udovychenko, Ivanovych et al. 2016, Karastan, Muliukina et al. 2018). Cross-country SSR comparisons between Ukrainian germplasm and foreign reference collections, initially applied to apples, have only recently been undertaken (Udovychenko, Skytte af Sætra et al. 2026). Regarding sweet cherry, the only genotyping research on Ukrainian cultivars from local collections was conducted (Ivanovych, Udovychenko et al. 2017a, Ivanovych and Volkov 2018).

Neither 'Drogans gelbe Knorpelkirsche' nor 'Valerii Chkalov', both recorded as the parent stock for a significant portion of the documented Ukrainian assortment (Turovtsev, Turovtseva et al. 2009, Turovtsev, Turovtseva et al. 2011), have been subjected to molecular comparison with European reference material. Such verification cannot be accomplished using Ukrainian material alone. This study aims to (i) verify the true-to-type identity of 94 Ukrainian

and reference sweet cherry accessions, addressing duplicates and homonyms; (ii) curate S-genotypes for Ukrainian cultivars against the published record and assign cross-incompatibility groups; (iii) compare Ukrainian accessions with European reference collections through allele-size calibration; and (iv) describe the genetic structure of the collection as a reflection of its breeding history and founder base.

2. Materials and methods

2.1 Plant material and DNA extraction

The plant material of sweet cherry varieties (Table S1) was collected from the stock plantations of IH NAAS, Melitopol RSH IH NAAS, Bakhmut (Artemivsk) NRS IH NAAS, Symyrenko RSP IH NAAS, and natural populations during 2013–2016. In total, 62 Ukrainian cultivars, 17 non-Ukrainian cultivars, and 10 wild accessions of Ukrainian origin were included in the study. DNA of 16 sweet cherry reference cultivars was provided from the Centre de Ressources Génétiques (CRG) *Prunus* (France) and INRAE-Bordeaux; reference accessions were selected to overlap with the earlier dataset (Campoy, Lerigoleur-Balsemin et al. 2016).

Total genomic DNA was extracted from leaves using the CTAB method (Doyle and Doyle 1987) with minor modifications. DNA quality and quantity were estimated by electrophoresis on 1% agarose. Provided DNA samples were extracted using DNeasy® Plant Mini Kit (Qiagen, Germany) following the manufacturer's protocol.

Accessions are maintained as living specimens in the collections listed in Table S1 and are identified by their accession numbers, which serve as voucher identifiers; no herbarium specimens have been prepared. Material from the Melitopol, Bakhmut, and Nikita collections is no longer available for resampling.

2.2 Microsatellite fingerprinting

The PCR reactions with M13-tailed fluorescent-labeled SSR primers were conducted in 15 µL of reaction mixture containing 2–10 ng genomic DNA, 1× PCR buffer, 2.5 mM MgCl₂, 0.2 mM of each dNTP, 0.2 µM of each primer (Table S2), and 0.25 U JumpStart™ *Taq* DNA Polymerase (Sigma, USA). The amplification was carried out in Mastercycler® 5341 (Eppendorf, Germany) and Veriti® (Applied Biosystems®, USA). The conditions for amplification were as follows: 94°C for 15 min followed by 15 cycles of 94°C for 45 s, 56–58°C for 45 s, 72°C for 45 s and then 25 cycles of 94°C for 45 s, 53°C for 45 s, 72°C for 45 s with the final elongation step of 72°C for 10 min (Schuelke 2000, Clarke and Tobutt 2003).

Each PCR was run on separate plates. After amplification, the presence, estimated concentration, and dilution level of PCR products were assessed. Capillary electrophoresis was conducted at INRAE-Bordeaux using the Beckman Coulter CEQ™ 8000 Genetic Analysis System. Initial peak detection was performed by M. Brault using the preinstalled software for allele identification and size measurement. SSR genotyping was performed at INRAE UMR 1332 BFP during an STSM in 2016.

Overall, we used primer pairs to amplify microsatellite genomic sequences at 18 previously published SSR loci in our experiments (Testolin, Marrazzo et al. 2000, Aranzana, Garcia-Mas et al. 2002, Dirlwanger, Cosson et al. 2002, Yamamoto, Mochida et al. 2002, Clarke and Tobutt 2003, Mnejja, Garcia-Mas et al. 2004, Vaughan and Russell 2004, Howad, Yamamoto et al. 2005) and three primer pairs for the S-locus (Sonneveld, Tobutt et al. 2003, Sonneveld, Robbins et al. 2006, Vaughan, Russell et al. 2006) (Table S2).

To harmonize fingerprinting of sweet cherry varieties across laboratories, the ECPGR Working Group on *Prunus* previously recommended a standard set of 16 SSR loci (Clarke, Sargent et al. 2009), and a set of 14 aligned SSR loci was more recently reported (Ordidge, Litthauer et al. 2021). The used panel shares six loci with each list (BPPCT037, EMPA002, EMPaS02, EMPaS14, EMPA026, and EMPaS10 with the former; BPPCT037, EMPA002, EMPaS02, EMPaS14, BPPCT034, and CPSCT038 with the latter), i.e., eight distinct loci in total, of which the six shared were used to harmonize to the ECPGR allele scale (Ordidge, Litthauer et al. 2021).

2.3 Allele-size calibration against reference cultivars

Since allele sizes measured on the Beckman Coulter CEQ™ 8000 were not calibrated against the ECPGR reference genotypes, the six loci common to the aligned ECPGR panel (BPPCT034, BPPCT037, CPSCT038, EMPa002, EMPaS02, and EMPaS14) were adjusted to the European allele scale prior to any comparison across collections.

Calibration used 13 cultivars genotyped in both this study and European collections, including 'Napoleon', 'Bigarreau Hâtif Burlat', 'Regina', 'Garnet', 'Cristobalina', 'Saint Georges', 'Xapata', 'Jaboulay', 'Folfer', 'Lapins', 'Adriana', 'Margit', and the 'Drogans gelbe Knorpel.' and 'Dönissens gelbe Knorpel.' pair. Their published profiles are identical across all six loci (Barreneche, Cárcamo de la Concepción et al. 2021, Ordidge, Litthauer et al. 2021), making the correspondence independent of the reference dataset. For each locus, Ukrainian allele calls were ranked and paired with the corresponding reference alleles across these cultivars, yielding a consensus per-allele correspondence table (Table S3).

Alleles positioned between two calibration anchors were determined by linear interpolation, while those outside the calibrated range were extrapolated using the nearest anchors. In both cases, the resulting value was rounded to the nearest allele in the combined ECPGR ladder, which includes 2,394 genotype profiles (Barreneche, Cárcamo de la Concepción et al. 2021, Ordidge, Litthauer et al. 2021), allowing a ± 1 bp tolerance to account for residual rounding and the odd/even scoring effects observed in these datasets (Ordidge, Litthauer et al. 2021).

To harmonize allele calls at six microsatellite loci with the ECPGR reference scale, a custom script was developed that derives the locus-specific correspondence table and implements it across all 94 accessions (Table S3). All applied rules are reproducible from the correspondence table and the indicators provided in Table S3; no allele call was based on undocumented judgment.

Offsets were locus-specific, ranging from about + 16 to + 21 bp. Calibration verification involved re-harmonizing the reference cultivars, thereby recovering the published European profiles (e.g., 6/6 loci for 'Napoleon', 'Xapata', and the 'Drogans gelbe Knorpel.' and 'Dönissens gelbe Knorpel.' pair). All 94 accessions were calibrated directly from raw allele calls; harmonization with the combined ECPGR ladder differed from calibration based solely on the narrower allele set (Barreneche, Cárcamo de la Concepción et al. 2021) in four accessions (eight locus-genotypes), all of which had alleles outside the calibrated range. Alleles requiring extrapolation or located midway between two reference alleles are flagged per accession and locus in Table S3; although retained in all analyses, these calls are less certain.

2.4 S-locus genotyping

Self-incompatibility (S) genotypes for all accessions were identified for 94 accessions using three primer combinations that target the polymorphic regions of the S-locus: the first and second introns of the S-RNase gene and the single intron of the S-haplotype-specific F-box (SFB) gene (Sonneveld, Tobutt et al. 2003, Sonneveld, Robbins et al. 2006, Vaughan, Russell et al. 2006). Second-intron S-RNase products were sized by agarose gel electrophoresis, whereas the fluorescently labeled first-intron S-RNase and SFB fragments were sized by capillary

electrophoresis on the CEQ™ 8000 analyzer. S-alleles were assigned based on the consensus fragment sizes reported for S-haplotypes and verified for the reference cultivars (Schuster 2012), following the approach used for Ukrainian germplasm in a pilot study (Ivanovych, Triapitsyna et al. 2017). The identity of some accessions was re-examined in complex with SSR genetic profiles and, where necessary, S-genotypes corrected.

The curated S-genotypes were utilized both to assign cross-incompatibility groups and, as an independent verification of candidate parent–offspring relationships, since a genuine parent and offspring are required to share at least one S-allele.

An additional 17 SSR-uncharacterized accessions were incorporated into the analysis as duplicates to confirm the S-genotypes of cultivars, using the polymorphic regions of the first and second introns of S-RNase solely. The comprehensive list of 111 accessions with assigned S-genotypes has been compiled (Table S4).

2.5 Marker informativeness

The discriminatory power of the 18 SSR loci was characterized by the number of different alleles (N_a), number of effective alleles (N_e), Shannon's information index (I), observed (H_o) and expected (H_e) heterozygosity, and Wright's fixation index (F), computed in GenAlEx 6.503 (Peakall and Smouse 2012). The same software was used to estimate the probability of identity for unrelated individuals (PI) and for full siblings ($PIsibs$). Polymorphic information content (PIC) was calculated in Cervus 3.0.7 (Kalinowski, Taper et al. 2007). Allele frequencies were computed in GenAlEx separately for the 94-accession Ukrainian panel (18 loci) and for the extended ECPGR reference set (14 loci; Ordidge et al., 2021; Barreneche et al., 2021), each serving as the internal reference for the corresponding analysis.

2.6 Identity, relationship, and parentage inference

Identity, relationship, and parentage were inferred across three data streams, each utilizing allele frequencies derived from its respective reference population. (i) Within-Ukrainian comparisons of the 94 accessions at the 18 unharmonized SSR loci were conducted using Familias 3.4.1 (Kling, Tillmar et al. 2014), with internal Ukrainian-panel allele frequencies. A mutation rate of 0.005 was applied at all loci in all three streams. Direct-match tests were employed to identify clones and synonyms (Table S5), while parent–offspring tests evaluated recorded pedigrees and identified probable ancestors (Table S6). (ii) Comparisons within the ECPGR dataset of 2,394 accessions across 14 standard SSR loci (Barreneche, Cárcamo de la Concepción et al. 2021, Ordidge, Litthauer et al. 2021). Direct-match and parent–offspring test results on cultivars of Ukrainian origin and related parental forms were extracted (Table S7). (iii) Comparisons of 90 deduplicated, harmonized Ukrainian accessions with the ECPGR dataset were conducted using six overlapping loci from the 14-locus SSR panel. Only test results with a direct match, no inconsistent loci, at least five overlapping markers, and an LR of 1.9×10^4 or higher, the value at which no false-positive match was observed within the reference set, were included (Table S8).

Comparisons (ii and iii) were conducted with a custom R script because Familias was too impractical at this scale. The script reproduces the Familias blind-search workflow and output columns (Kling, Tillmar et al. 2014). For each pair, per-locus genotype likelihoods conditional on the identity-by-descent state, $P(\text{genotype} \mid \text{IBD} = 0, 1, 2)$, were combined across loci, and the realized IBD coefficients ($\kappa_0, \kappa_1, \kappa_2$) were obtained by maximum-likelihood estimation of relatedness (Milligan 2003). Kinship was then $\phi = \kappa_1/4 + \kappa_2/2$, and the coefficient of relationship $r = 2\phi = \kappa_1/2 + \kappa_2$. Pairwise likelihood ratios for the tested relationship versus the unrelated hypothesis, plus a triage note, complete the output.

Each pair was summarised by the likelihood ratio (LR), Identity-by-State (IBS), ϕ , κ_0 – κ_2 and r . In the within-Ukrainian stream (i), where allele frequencies come from a small, founder-heavy panel and are non-independent, LR, ϕ , κ and r

were treated as indicative only, and relationship conclusions rested on the count of exclusionary ($IBS = 0$) loci; for the reference-based streams (ii) and (iii), the large ECPGR frequency base (2,394 accessions) makes these statistics reliable, and they were interpreted quantitatively. In all streams, each discordant locus was classified from the genotypes as a scoring wobble or single-repeat mutation (both alleles typed, ± 2 bp), a probable allele dropout (a homozygous call masking a shared allele), or a genuine exclusion (opposite homozygotes, or heterozygotes sharing no allele and > 2 bp apart); a single genuine exclusion rejected a first-degree relationship.

The codominant *S*-locus functioned as an independent verification: the parent and offspring must share a single *S*-allele; otherwise, they are excluded from the LR. Corrected parental forms were assigned solely when there was concordance between SSR and *S*-locus evidence (Table S1).

The standard 14-locus genetic profiles of Ukrainian accessions from the combined ECPGR dataset (Barreneche, Cárcamo de la Concepción et al. 2021, Ordidge, Litthauer et al. 2021) were extracted (Table S9). For accessions genotyped in this study, the data were supplemented with the corresponding harmonized six-locus calls (Table S3), enabling direct comparisons across collections.

Pedigree relationships were represented as a directed graph in Graphviz (Ellson, Gansner et al. 2004), based on the corrected parental assignments in Table S1. Each edge was annotated to indicate whether the connection was molecularly confirmed, a candidate assignment, or a literature record for a putative ancestor unavailable for genotyping. Nodes display the curated *S*-genotype (Section 2.4), enabling the presentation of SSR-based parentage and *S*-locus compatibility in a single diagram.

2.7 Genetic structure and relationships

The genetic structure was determined using only the 18 neutral SSR loci. The three *S*-locus assays (PaConsl, PaConslI, F-box) were excluded because they target a single non-recombining haplotype block maintained by gametophytic balancing selection, which violates the assumptions of neutrality and linkage equilibrium required by the admixture model. All analyses were conducted on a clone-corrected dataset of 90 accessions.

The Discriminant Analysis of Principal Components (DAPC) was performed on the 18 neutral SSR loci in *adegenet* (Jombart, Devillard et al. 2010). Clusters were inferred *de novo* with *find.clusters* ($K = 1-10$; 1000 replicate BIC runs), and $K = 3$ was retained as both the BIC minimum and the elbow. The number of retained principal components was set by the α -score ($n.pca = 6$); cross-validation (*xvalDapc*, 100 replicates, group-mean success) indicated a higher but broadly equivalent optimum ($n.pca = 15$), and reassignment was 99% concordant between the two (1/90 individuals, an admixed accession of low posterior). The parsimonious α -score value was therefore retained, with two discriminant functions ($n.da = 2$).

Bayesian clustering was performed with STRUCTURE 2.3.4 (Pritchard, Stephens et al. 2000) through the command line. An admixture model with correlated allele frequencies was employed. No prior information on population, sampling location, or phenotypes was provided; the germplasm status and breeding origin categories used to order accessions in the figures are purely for display and did not influence the inference. The number of clusters, K , was varied from 1 to 11, with 20 independent runs for each K , each consisting of 500,000 burn-in iterations followed by 500,000 Markov chain Monte Carlo iterations for a total of 220 runs.

The most likely number of clusters was assessed using the mean estimated log-likelihood of the data $L(K)$ and the ΔK statistic (Evanno, Regnaut et al. 2005), with K ranging from 1 to 11. This was done using STRUCTURE HARVESTER (Earl and vonHoldt 2012), which was re-implemented in a custom script. Since independent runs can

produce different cluster-label arrangements and, at higher K , different modes, the 20 replicate Q-matrices at each K were aligned and averaged with CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007). All subsequent analyses and bar plots are based on these CLUMPP consensus matrices, not on individual runs. An accession was assigned to a specific cluster if its membership coefficient was $Q \geq 0.80$; otherwise, it was considered admixed.

Genetic relationships among the accessions were summarized using DARwin 6.0.21 (Perrier, Flori et al. 2003). Pairwise dissimilarities were calculated using the simple matching index for allelic data (ploidy = 2), accounting for pairwise allele deletion and a minimum of 50% valid data for each unit pair. Node support was assessed via 10,000 bootstrap resamplings across loci, with a minimum of 60% of data considered valid; three accessions (MEL/2, TEM/1–2, FRA 77) did not meet this criterion and were excluded, leaving 87 tips. A dendrogram was constructed from the original dissimilarity matrix using the Unweighted Neighbour-Joining method (Gascuel 1997, Gascuel and Steel 2006) and midpoint-rooted in iTOL ver. 7 (Letunic and Bork 2024).

3. Results

3.1 Marker informativeness and allele-size calibration

All 18 SSR loci were polymorphic across the 94 accessions, yielding 151 alleles with a mean of 8.4 per locus, ranging from 2 (M19A) to 16 (EPPB4230). The number of effective alleles averaged 2.9 (1.0 at M19A to 5.7 at BPPCT037), PIC was 0.554 (0.021–0.801), observed heterozygosity was 0.583 (0.022–0.840), and expected heterozygosity was 0.600 (0.021–0.825), with the same two loci defining both extremes throughout (Table 1). M19A was effectively monomorphic in this material and contributed little to discrimination, whereas BPPCT037, EPDCU3392, BPPCT040, and CPSCT034 were the most informative.

Table 1

Indices of genetic diversity and discriminatory power of microsatellite loci in the genetic fingerprinting of 94 sweet cherry accessions

Locus	LG	Size, bp [†]	Na	Ne	Ho	He	PIC	F	HW	P_{ID}	P_{ID-SIB}
BPPCT034	2	232– 290	13	2.297	0.585	0.565	0.536	-0.036	NS	0.218	0.522
BPPCT037	5	144– 180	10	5.715	0.840	0.825	0.801	-0.019	NS	0.054	0.351
BPPCT040	4	138– 166	11	4.151	0.798	0.759	0.726	-0.051	NS	0.091	0.393
CPSCT022	7	265– 279	8	2.351	0.578	0.575	0.510	-0.006	NS	0.246	0.524
CPSCT034	2	192– 210	9	3.997	0.779	0.750	0.713	-0.039	NS	0.099	0.400
CPSCT038	2	202– 246	8	1.865	0.453	0.464	0.430	0.022	NS	0.321	0.598
EMPA002	1	124– 150	5	2.058	0.598	0.514	0.408	-0.163	NS	0.342	0.579
EMPA026	8	224– 252	6	2.139	0.261	0.532	0.440	0.510	***	0.311	0.562
EMPaS02	3	152– 168	8	3.048	0.693	0.672	0.632	-0.032	NS	0.148	0.451
EMPaS10	4	168– 205	10	2.598	0.582	0.615	0.583	0.053	NS	0.180	0.487
EMPaS14	5	218– 232	3	2.099	0.648	0.524	0.413	-0.237	NS	0.337	0.573
EPDCU3392	7	128– 150	9	4.654	0.793	0.785	0.755	-0.011	NS	0.076	0.377
EPPB4230	5	276– 314	16	3.294	0.677	0.696	0.673	0.027	NS	0.115	0.431
EPPCU3090	6	190– 206	8	2.876	0.553	0.652	0.596	0.152	NS	0.177	0.468
M19A	6	189– 192	2	1.022	0.022	0.021	0.021	-0.011	NS	0.958	0.979
pchgms49	8	160– 204	9	1.828	0.467	0.453	0.427	-0.032	NS	0.325	0.605
pchgms55	4	144– 204	9	3.791	0.602	0.736	0.695	0.182	NS	0.111	0.410

Note: Na, number of different alleles; Ne, number of effective alleles; Ho, observed heterozygosity; He, expected heterozygosity; F: Wright's fixation index (inbreeding coefficient); PIC, polymorphic information content; HW: significance of deviation from Hardy–Weinberg equilibrium: NS = not significant, ***: significant at the 0.1% level; P_{ID} , probability of identity of unrelated cultivars; P_{ID-SIB} , probability of siblings identity. [†] – The indicated fragment sizes are larger due to the use of M13-tailed primers and are not harmonized with ECPGR.

Locus	LG	Size, bp [†]	Na	Ne	Ho	He	PIC	F	HW	P _{ID}	P _{ID-SIB}
UDP98-022	1	102– 118	7	2.899	0.559	0.655	0.620	0.147	NS	0.154	0.461
Mean			8.389	2.927	0.583	0.600	0.554	0.025		0.237	0.510
Total			151							6.705×e- 14	3.14×e- 6
Note: Na, number of different alleles; Ne, number of effective alleles; Ho, observed heterozygosity; He, expected heterozygosity; F: Wright's fixation index (inbreeding coefficient); PIC, polymorphic information content; HW: significance of deviation from Hardy–Weinberg equilibrium: NS = not significant, ***: significant at the 0.1% level; P _{ID} , probability of identity of unrelated cultivars; P _{ID-SIB} , probability of siblings identity. [†] – The indicated fragment sizes are larger due to the use of M13-tailed primers and are not harmonized with ECPGR.											

The fixation index averaged 0.025, ranging from – 0.237 at EMPaS14 to 0.510 at EMPA026. Eleven loci exhibited a slight excess of heterozygotes, with only EMPA026 significantly deviating from Hardy–Weinberg equilibrium. These accessions comprise a curated collection of cultivars, breeding selections, and wild forms connected through numerous parent–offspring relationships, including duplicates, rather than a randomly mating population. Therefore, fixation indices and Hardy–Weinberg deviations are presented as markers of allele behavior rather than indicators of demographic processes. The anomaly at EMPA026 is likely technical, as this locus showed an excess of homozygous calls and contributed disproportionately to apparent exclusions in parentage testing (Section 3.5). This pattern suggests a segregating null allele rather than inbreeding.

Out of the 151 alleles examined, 38 were found at a frequency below 0.6% (meaning only one copy among 188 haploid genomes), 36 ranged from 0.6% to 2.5%, and 77 exceeded 2.5%. Rare alleles were mainly found in wild and landrace materials rather than in cultivars, especially in the ornamental wild form 'Plena', where 24 of 30 typed alleles were rare or private. Similar patterns appeared in wild accessions AVCV3–AVCV8, and among landraces in 'Saint Georges' (9) and 'Xapata' (4). Ukrainian cultivars contained few rare alleles, with ACC_GEN/1 having 7 and 'Nektarna' 4. The loci EPPB4230 and BPPCT034 possessed the highest numbers of rare and private alleles, with 10 and 9, respectively.

The current panel was assembled around loci with a well-established utility in *Prunus* fingerprinting and predates the synchronized 14-locus reference list; six loci overlap with it and were adjusted (Section 2.3) to the European allele scale (Ordidge, Litthauer et al. 2021). In addition to recovering the published profiles of the 13 shared reference cultivars, binning quality can be evaluated through the frequency spectrum of the harmonized calls, since misclassification would tend to assign Ukrainian alleles to European allele sizes that are rare or absent in the reference. Such misclassification was not observed: every harmonized Ukrainian allele occurring at a frequency above 5% corresponds to a common allele within the reference ladder, and no call fell outside this range (Tables S3, S8). Two alleles are more prevalent in Ukrainian than in European material—BPPCT037 154 (0.149 vs 0.031) and 134 (0.085 vs 0.028)—but both are established reference alleles rather than novel variants, and their higher frequency is consistent with the largely separate breeding histories of the Ukrainian pool.

The mean allele number per locus (8.4) is lower than the 14.5 reported for the standard 14-locus panel across 376–1,275 typed individuals (Ordidge, Litthauer et al. 2021) and the 9.8 among 220 unique genotypes (Barreneche, Cárcamo de la Concepción et al. 2021). Expected and observed heterozygosity values (0.600 and 0.583) are similarly somewhat below the corresponding values of 0.670/0.670 and 0.770/0.780. These differences correlate

with sample size, the narrower geographic scope of the current study, and the different loci employed, and should not be interpreted as indicating reduced diversity within Ukrainian germplasm. The values fall within the range reported for comparable single-country panels (Guarino, Santoro et al. 2009, Ganopoulos, Kazantzis et al. 2011).

3.2 Duplicates, homonyms, clone correction

Identity testing was conducted on sample codes rather than on cultivar labels, with each label treated as a hypothesis to be tested. Published records, nursery documentation, and breeders' unpublished communications provided the a priori expectation, while conclusions were drawn from the molecular data. Because allele frequencies estimated within this panel are not independent (Section 2.6), identity was assessed based on the number of loci at which two accessions differed and on their *S*-genotypes rather than on likelihood ratios; the complete blind-search output is provided in Table S5.

Each of five cultivar names was represented by two accessions (Table 2). Identical genotypes were obtained for the two 'Oleh' accessions (15 loci compared, no differences) and the two 'Otrada' accessions (18 loci, one difference). The pairs shared *S*-genotypes S_1S_9 and S_3S_5 , respectively. The single discrepancy in 'Otrada' at udp98-022 (102/106 vs 102/102) is a homozygous call at an otherwise heterozygous locus, i.e., a probable allele dropout, and does not argue against clonal identity. The remaining three pairs were not identical: 'Siestronka' differed at 12 of 18 loci, 'Yaroslavna' at 11 of 17, and 'Donetska krasavytsia' at 12 of 17. These are homonyms rather than duplicates. The two 'Siestronka' accessions also carry different *S*-genotypes (S_1S_6 and S_3S_5) and thus belong to different cross-incompatibility groups, whereas the 'Yaroslavna' (S_3S_5) and 'Donetska krasavytsia' (S_3S_9) pairs share *S*-genotypes, which is consistent with a close relationship — plausibly full sibs propagated under one name — but not with identity.

Table 2

Tested accession pairs showing comparison details: loci analyzed, loci with differences, S-genotype, cross-incompatibility group (CIG), recording status, and conclusions

Accession 1	Accession 2	Loci compared	Loci differing	S-genotype	CIG	Recorded	Conclusion
Oleh OLE/1	Oleh OLE/2	15	0	S_1S_9 / S_1S_9	XVIII	duplicate	clone
Otrada OTR/1–3	Otrada OTR/2	18	1	S_3S_5 / S_3S_5	VII	duplicate	clone (dropout at udp98-022)
Annushka ANN/1	Taina TAY/1	17	0	S_3S_4 / S_3S_4	III	distinct cultivars	clone; TAY/1 mislabeled
Lehenda Mliieva LEG/1	Zoriana ZOR/1	16	0	S_6S_9 / S_6S_9	X	distinct cultivars	clone; ZOR/1 mislabeled
ACC_SES/1	Biriuza BIR/1	18	2	S_1S_6 / S_1S_6	XX	distinct cultivars	identity not resolved
ACC_VAL/3	Shans SAN/1	14	2	S_3S_6 / S_3S_6	VI	distinct cultivars	close relative, not clone
ACC_YAR/1	Yaroslavna YAR/2	17	11	S_3S_5 / S_3S_5	VII	duplicate	homonyms, distinct
ACC_DON/1	Donetska krasavytsia DON/2	17	12	S_3S_9 / S_3S_9	XVI	duplicate	homonyms, sibs
ACC_SES/1	Siestronka SES/2	18	12	S_1S_6 / S_3S_5	XX / VII	duplicate	homonyms, distinct

The blind search also identified two unexpected clonal pairs from the records: 'Annushka' (ANN/1) and 'Taina' (TAY/1) were identical across all 17 loci examined and shared S_3S_4 , while 'Lehenda Mliieva' (LEG/1) and 'Zoriana' (ZOR/1) were identical across all 16 loci tested and shared S_6S_9 . In both instances, the concern is with the accessions rather than the cultivars: according to the documented lineage, TAY/1 and ZOR/1 are likely mislabeled propagations of 'Annushka' and 'Lehenda Mliieva'.

Two additional pairs were almost identical but not quite. ACC_SES/1 differed from 'Biriuza' at only two of 18 loci and shared both S_1S_6 and the cross-incompatibility group XX. Since the differences (EPPB4230, pchgms55) involve homozygous calls and are thus dropout-compatible, the data cannot confirm or exclude clonal identity, and the accession remains provisionally distinct pending re-genotyping. Meanwhile, ACC_VAL/3 differed from 'Shans' at two of 14 loci while sharing S_3S_6 . In this case, one difference indicates a true exclusion, suggesting they are close relatives rather than clones.

For analyses assuming independent genotypes, one representative from each confirmed clonal pair was retained, thereby excluding OTR/2, TAY/1, ZOR/1, and OLE/2 and reducing the dataset from 94 to 90 accessions. Mislabeled but genetically distinct accessions were kept with provisional identifiers (ACC_DON/1, ACC_YAR/1, ACC_SES/1, ACC_VAL/3) because discarding them would eliminate critical evidence of labeling errors in the collections.

3.3 Cross-country validation against ECPGR and EU.CHERRY

The reference resource integrates the aligned ECPGR dataset with the EU.CHERRY dataset (Barreneche, Cárcamo de la Concepción et al. 2021, Ordidge, Litthauer et al. 2021). The datasets are linked: EU.CHERRY (324 accessions from 19 countries) serves as the core dataset of the ECPGR EU.CHERRY project and was integrated into the aligned database, with only eight accessions differing between them. When a genotype appears under both an ECPGR code and its EU.CHERRY name, it counts as a single record, not as independent confirmation. The 2-bp EMPa017 offset between datasets was applied during the assembly of the combined set of 2,394 accessions (Ordidge, Litthauer et al. 2021).

Two comparisons were made. The combined reference set was first searched against itself to determine which synonym groups it contains and how Ukrainian material is represented within it (Table S7). Then, the 90 deduplicated accessions from this study were compared against that set using six harmonized loci (Table S8), with profiles listed together in Table S9. Since only five or six loci overlap between the two panels, a zero-mismatch match does not necessarily confirm identity; at six loci, the chance of unrelated genotypes matching remains significant. Therefore, matches were considered consistent with identity only if they were unique or aligned with documentary evidence or the S-genotype.

The five reference cultivars genotyped in both frameworks — 'Xapata', 'Saint Georges', 'Cristobalina', 'Napoleon', and 'Garnet' — all showed exact matches to their European counterparts across the harmonized loci. This confirms that the allele-size calibration transfers accurately between both scales.

The key finding concerns 'Drogans gelbe Knorpel.' and 'Dönissens gelbe Knorpel.', which are clearly separate synonym groups within the reference set. The first group includes 'Drogans Gelbe Knorpel.' (DEU451-005), 'Grosse blanche de Verchocq', 'Dzintars' (LVA001011), 'Bigarreau Jaune Tardif', 'Kirsche Kierling', 'Branca', and Pa_48. The second group encompasses 'Donnisens Gelbe Knorpel.' (VF Bonn, 1979-006), Hungarian accessions 'Fehér Cseresznye', 'Fehér Ropogós Pollenadó', 'Pákozdi Fehér', and Pa_45/Pa_46. Accession #110 [FRA], labeled as 'Guigne jaune donissen' from France, matches the 'Drogans gelbe Knorpel.' group at all six loci, while DRO/1 [UKR], labeled 'Drohana zhovta', aligns with the 'Dönissens gelbe Knorpel.' group. Both are misclassified, yet their S-genotypes agree separately: DRO/1 [UKR] has S_3S_6 , and accession #110 [FRA] has S_1S_5 . The inversion is not limited to the material examined here: within the reference set, accessions called 'Guigne jaune donissen' (V3862) fall within the 'Drogans gelbe Knorpel.' group. The aligned analysis further clarifies this by grouping them into two distinct clusters of over ten members each — one associated with 'Dönissens Gelbe Knorpel.', and the other including eight names that are not repeated across countries, such as 'Drogans Gelbe Knorpel.' in Germany and 'Grosse blanche de Verchocq' in France (Ordidge, Litthauer et al. 2021).

Finally, several queries yielded multiple equally good matches — AVCV7 matched eight reference accessions from five countries; the accession labeled 'Tunisian/Barg' aligned with the 'Bigarreau Hâtif Burlat' group; and 'Jaboulay' corresponded to several unrelated accessions. These results reflect the limited resolution of six loci rather than true synonymy. They are marked accordingly in Table S8 and should not be considered definitive identities. The same caution applies to the reference groups: as noted (Ordidge, Litthauer et al. 2021), resolving issues of synonymy versus mislabeling is currently impossible, and some indistinguishable groups may contain sports and clones of independent value. To clarify the Ukrainian cases, genotyping at the full 14-locus standard set would be necessary.

3.4 S-genotypes: published vs curated Ukrainian records

S-genotypes were determined for 111 accessions, including 94 accessions characterized using SSR markers and an additional 17 duplicate accessions that were typed solely at the S-locus. Among these, 95 accessions were of Ukrainian origin. The data were subsequently compared with all existing published records pertaining to the same cultivars (Tables S4, S10). A total of nineteen distinct S-alleles were identified, representing an increase from twelve alleles documented in the previous survey of this germplasm (Ivanovych, Triapitsyna et al. 2017b). The newly identified alleles were found predominantly in wild and landrace materials.

The allele profile of the enlarged Ukrainian sample replicates that previously reported: S_9 (20.3%), S_5 (18.8%), S_3 (14.6%), S_6 (14.1%), S_1 (10.9%), and S_4 (9.9%) collectively constitute 88.6% of the alleles analyzed, with each allele's frequency differing by no more than 1.5 percentage points from the values obtained in the earlier sample comprising 35 accessions. Consequently, the higher frequency of S_5 and particularly S_9 relative to other European cherry germplasm (Schuster 2012) remains evident in a nearly three times larger—and the distributions of alleles and cross-incompatibility groups are not reiterated herein (Ivanovych, Triapitsyna et al. 2017).

Among the 78 published records encompassing 66 cultivars, 22 are corroborated by the present data, 13 are contradicted, and 31 pertain to cultivars not represented in this study. Eleven cultivars exhibit conflicting S-genotypes in the literature. Of these, six are resolved herein in favor of a specific variant: 'Donetskyi uholok' S_3S_5 (Sharma, Cachi et al. 2016, Schuster 2017) rather than S_5S_9 (Lisek, Rozpara et al. 2015); 'Krupnoplidna' S_5S_9 (Békefi 2006, Schuster, Flachowsky et al. 2007) rather than S_4S_5 (Hegedűs, Taller et al. 2013); 'Kytaivska chorna' S_6S_9 rather than S_2S_5 (Lisek, Rozpara et al. 2015); 'Prysadybna' S_6S_9 (Lisek, Rozpara et al. 2015, Schuster 2017) rather than S_5S_9 (Liu, Qi et al. 2018); and 'Donchanka' S_4S_5 and 'Vasylysa prekrasna' S_5S_9 , in both cases against records of S_3S_6 and S_3S_9 , respectively (Lisek, Rozpara et al. 2015, Sharma, Cachi et al. 2016, Schuster 2017). For 'Skorospilka', the S_5S_9 record (Liu, Qi et al. 2018) is excluded, whereas the S_2S_3 genotype (Paprštein, Patzak et al. 2018, Patzak, Henychová et al. 2020) aligns with the genotypes of its two documented progeny (Section 3.5). The case of 'Valerii Chkalov' is discussed separately in Section 3.5. Records pertaining to 'Kutuzovka', 'Perspektyvna', and 'Transportabelna' remain unresolved, as these cultivars are not included in this study.

3.5 Corrected pedigrees

The recorded parentages were examined for each accession with documented ancestry by employing the count of exclusionary loci and the S-locus as independent verification methods (Section 2.6). Likelihood ratios and relatedness coefficients were not utilized to substantiate conclusions within the Ukrainian panel. The results are summarized for each accession in **Table S1** and presented pairwise in Table S6.

Three pedigrees have been fully confirmed. 'Otrada' traces from 'Annushka' × 'Anons', with neither parent showing exclusions, and the S-genotypes forming a consistent trio ($S_3S_4 \times S_5S_9 \rightarrow S_3S_5$). This also dismisses the alternative record attributing it to 'Drogans gelbe Knorpel.' × 'Valerii Chkalov'. 'Studentka' descends from 'Siestronka' × 'Donetska krasavytsia'; when the correct accessions (SES/2 and DON/2) are used instead of the mislabeled ones (ACC_SES/1 and ACC_DON/1), the only difference at EMPaS02 is a dropout. 'Prysadybna' descends from 'Dönissens gelbe Knorpel.' × 'Jaboulay', with a single discrepancy of approximately ± 2 base pairs against 'Jaboulay', and an S-genotype consistent with the cross (S_6S_9 from $S_3S_6 \times S_6S_9$).

In an additional 25 accessions, one recorded parent was confirmed, while the other could not be tested. The ongoing limitation is the absence of an authentic 'Valerii Chkalov': 19 of the 50 recorded pedigrees list it as the pollen parent, yet the only accession bearing that name (ACC_VAL/3) is mislabeled (Section 3.2). Consequently, in each case, the second parent is documented as not tested rather than rejected. This affects 'Annushka', 'Ruksandra', and 'Valeriia'

(all confirmed as progeny of 'Donchanka'), as well as 'Donetskyi uholok', 'Kazka', 'Melitopolska myrna', 'Mirazh', 'Talisman', 'Prestyzhna', and others. Similarly, 'Lasunia' and 'Anshlah' are confirmed progeny of 'Prestyzhna'; 'Amazonka' is progeny of 'Siestronka'; 'Lehenda Mliieva' is progeny of 'Krupnoplidna'; and 'Sato Nishiki' is progeny of 'Napoleon', with the second parent unavailable in each case.

'Drogans gelbe Knorpel.' emerges as the predominant founder, appearing as a confirmed progenitor in 17 accessions. Part of this reflects nomenclature rather than additional crosses: 'Bigarreau Napoléon blanc' is a somatic sport of 'Drogans gelbe Knorpel.' (Symyrenko 1963) and is therefore genetically indistinguishable from it. Pedigrees that list 'Napoléon blanc' as a parent—'Anons', 'Dachnytsia', 'Krupnoplidna', and 'Prestyzhna'—therefore pertain to the same genotype. This equivalence does not apply to all such records, as 'Mliivska chorna' and 'Nektarna', both documented as 'Napoléon blanc' open-pollinated varieties, are excluded at EMPA026 because they carry opposite homozygotes.

The exclusion pattern indicates that two distinct trees were propagated under the name 'Valerii Chkalov' (Fig. 1). The authentic cultivar is S_1S_9 , yet to be confirmed; however, three Bakhmut cultivars recorded as its progeny — 'Annushka' and 'Ruksandra' (S_3S_4), and 'Donetskyi uholok' (S_3S_5) — lack both S_1 and S_9 and are instead compatible with the S_3S_9 genotype documented for the cultivar by independent sources (Liu, Qi et al. 2018, Paprštejn, Patzak et al. 2018, Patzak, Henychová et al. 2020). All Melitopol progeny ('Kazka', 'Mirazh', 'Talisman', 'Prestyzhna', 'Elektra', and others) remain compatible with S_1S_9 . The sole exception is 'Valeriiia' (S_1S_4), the only Bakhmut cultivar requiring S_1 and therefore incompatible with the S_3S_9 form. Therefore, the pollen donor used at Bakhmut is regarded as an unidentified form propagated under the designation 'Valerii Chkalov' (Bakhmut), rather than as a specific cultivar.

Recorded parentage was rejected outright in 13 accessions, which are entered as Unknown. These comprise 'Chervneva', 'Dzherelo', 'Elektra', 'Etyka', 'Mliivska chorna', 'Nektarna', 'Shans', 'Temporion', 'Vasylysa prekrasna', 'Veselka', both 'Donetska krasavytsia' accessions, and ACC_SES/1. In most cases, a single locus with opposite homozygous genotypes—most often EMPA026 or EMPaS10—excludes first-degree descent, and when the S -genotypes also share no allele, the exclusion is corroborated independently. An additional four accessions, including the breeding selections 'D 44 – 15' and 'D 58 – 52', lack records that could be tested; both behave as highly connected nodes with numerous first-degree links and cannot be resolved to a unique cross.

Four parentage assignments are proposed for the first time. 'Pokazkova' and 'Shchedrist' each show a clear first-degree relationship to 'Dönissens gelbe Knorpel.' with no exclusions, and are most plausibly co-offspring of 'Dönissens gelbe Knorpel.' rather than parent and offspring of each other. The unidentified white-fruited accession AVCV1 is fully trio-compatible with 'Drogans gelbe Knorpel.' × 'Donchanka', with its S -genotype S_1S_4 comprising S_1 from 'Drogans gelbe Knorpel.' and S_4 from 'Donchanka'. 'Mykhailivska chorna' (S_1S_4) shows a first-degree connection to 'Lapins' (S_1S_4), which is suggested to be its progeny, or to either clone of 'Halka' (S_1S_4), a Czech cultivar (Schuster 2020). All four are presented as candidate assignments awaiting validation through independent testing.

3.6 Genetic structure and breeding history of the collection

The genetic structure was analyzed as background for the identity results, using Discriminant Analysis of Principal Components (DAPC) and STRUCTURE independently, without incorporating prior information (Figure S1). The analysis used the same 18 neutral SSR loci and included 90 clone-corrected accessions.

DAPC identified the optimal number of clusters ($K = 3$) using both the Bayesian Information Criterion (BIC) and the minimum and elbow methods. With a membership threshold of $Q \geq 0.80$, 83 of 90 accessions were successfully

classified (Fig. 2A). Cluster 1 ($n = 48$) consisted predominantly of modern Ukrainian cultivars, along with four accessions of unresolved identity and a small number of introduced reference samples. Cluster 2 ($n = 22$) included wild accessions and most foreign reference samples. Cluster 3 ($n = 13$) was composed largely of early selection cultivars. By contrast, STRUCTURE identified $K = 2$ using the ΔK criterion and assigned 82 of 90 accessions (Fig. 2B). In this analysis, Cluster 1 ($n = 63$) included Ukrainian-bred material from both temporal periods, while Cluster 2 ($n = 19$) again comprised wild and introduced accessions.

The two methods agree on the primary division. Collapsing the DAPC clusters into bred (C1 + C3) and wild-plus-introduced (C2) partitions produces concordant assignments for 76 of 90 accessions, with a single reciprocal disagreement — ACC_KYT/3, an accession whose identity could not be verified. The remaining discrepancies occur when one method assigns an accession while the other classifies it as admixed. Consequently, DAPC resolves three cohorts with essentially the same confidence as STRUCTURE resolves two.

The attempt to replicate the third division within the ancestry model proved unsuccessful. As the value of K increased, the proportion of accessions assigned at $Q \geq 0.80$ decreased from 91% at $K = 2$ to 34% at $K = 4$ and further to 16% at $K = 7$. Notably, only the wild-plus-introduced cluster remained coherent throughout this process (**Figure S2**). Given that ΔK predominantly detects the highest level of genetic structure (Evanno, Regnaut et al. 2005), the secondary peaks of ΔK were found to be approximately an order of magnitude below the signal at $K = 2$ and were therefore not considered for interpretation. This outcome aligns with expectations for a breeding collection characterized by a limited number of founders recurring within the pedigrees: the ancestry is genuinely continuous and is represented as pervasive admixture by STRUCTURE, while DAPC—which maximizes separation between groups and does not rely on any population-genetic model—still facilitates the delineation of interpretable boundaries. Consequently, the DAPC groups are regarded as breeding cohorts with porous boundaries rather than as discrete populations, and the germplasm status and origin categories used to order Fig. 2 were not included in either prior analysis.

Several individual assignments pertain to the determination of the identity results. 'Melitopolska chorna' was admixed according to STRUCTURE ($Q = 0.58$), aligning with its documented offspring of suggested French cultivar 'Frantsuzka chorna', although DAPC classified it with high confidence ($Q = 0.96$) — exemplifying the divergence in the estimations of the two methodologies. A sample of 'Drogans gelbe Knorpel.' (#110 [FRA]) and 'Dönissens gelbe Knorpel.' (DRO/1 [UKR]) were both classified based on their genetic background, regardless of their exchanged labels. ACC_NIJ/2 was identified within the wild-plus-introduced group by both methods ($Q = 0.99$ and 0.90); as its definitive identity remains unknown, it is regarded as an unverified field accession rather than a cultivar. AVCV7 was the sole wild-collected accession to cluster with cultivated germplasm ($Q = 0.91$ and 0.96), which is consistent with a cultivar-derived seedling at the collection site.

The neighbor-joining dendrogram (Fig. 3) is presented as a summary of pairwise distances rather than as an independent estimate of grouping. Only eleven nodes reached 70% bootstrap support, all of which subtend terminal pairs or small groups of two to five accessions, with the strongest support observed for ACC_SES/1 paired with 'Biriuzza' (99%). No nodes supporting the major groupings received such support. Consequently, the dendrogram provides informative insights primarily for the true-to-type comparisons detailed in Section 3.2. However, grouping analyses are derived from the DAPC rather than from the tree topology.

4. Discussion

This study considers the cultivar name as a hypothesis subject to testing rather than as a mere identifier. The findings demonstrate that this distinction is of significant importance and not merely procedural. When a name is evaluated against genotype, S-locus constitution, and documented parentage concurrently, three lines of evidence provide mutual validation: SSR profiles alone establish distinctness but do not confirm identity; S-genotypes alone assign an accession to an incompatibility group without differentiating among cultivars within that group; and recorded pedigrees, as illustrated below, are inherently reflective of the period during which they were documented. Therefore, convergence among all three lines of evidence constitutes the minimum criterion for asserting identity, while divergence among them indicates the presence of error rather than simply signaling it. Consequently, verification cannot be delegated solely to a single external reference. The reciprocal exchange of identity documentation presented here is also present in European holdings; therefore, a name confirmed through comparison with one collection may merely replicate that collection's error. Under this approach, what remains invariant is not the name but the individual tree, and it is at this level that records must be maintained.

Ukrainian cultivars are stored in numerous collections (Table S10), with the largest in Poland and Russia, and additional accessions in Germany, the Czech Republic, Hungary, the Baltic states, Romania, the UK, the USA, and China (Schuster 2017, EURISCO 2026). Much of this material is difficult to locate via name searches. Names have been transliterated or translated into various languages, including Polish, Czech, Romanian, German, and French. Russian-to-English transliteration is most common, often with typographical errors. Old landraces are known by many synonyms, such as 'Black Tartarian' (Symyrenko 1963). Breeding forms are often labeled with unregistered hybrid numbers such as hybrids '19-21B', '2115', 'D 48–52', 'D 24–37', and 'D 49 – 1', or numbered in ways that do not correspond to known identities. The exception is 'D 44 – 1', which corresponds to 'Otrada' (Lisek, Rozpara et al. 2015). Some carry undocumented or lost Soviet-era names, such as 'Debanka' (Schuster, Flachowsky et al. 2007), or forms that are no longer interpretable yet are still claimed as Ukrainian, such as 'Amicitia', 'Ranna Krupnoplodna', and 'Shorlry' (Liu, Qi et al. 2018).

Ukrainian germplasm is represented more broadly in European collections than just by accession names, mostly through Latvian and German holdings and often under transliterated names: 'Valerii Chkalov' (FRA057-001, LVA001023, Pa_118, 14-007-2537, all part of a single synonym group), 'Krupnoplidna' (LVA001025, Pa_136), 'Aelita' (LVA001027), 'Donchanka' (LVA001026), 'D 24–37' (LVA001024), and 'Donetskyi uholok' (Ugoljok, LVA001028, similar to EMCherry 7). Several samples were directly recovered from the Ukrainian panel: DOU/1 matched Ugoljok and EMCherry 7, KRU/4 matched 'Krupnoplidna', AEL/1 matched 'Aelita', MEL/2 matched the Pillnitz accession Pa_224 ('Melitopolskaja tscherkaja'), and DOR/1 matched EMCherry 8. In contrast, ACC_VAL/3 did not match any European 'Valerii Chkalov' accessions, supporting its reclassification (Section 3.2).

The passport data show the same discontinuity. The German national inventory in EURISCO records only four *P. avium* accessions of Ukrainian origin, yet Ukrainian cultivars are demonstrably held at Dresden-Pillnitz under German accession codes: 'Valerij Tschkalov' (Pa_118), 'Krupnoplodnaja' (Pa_136), and 'Melitopolskaja tscherkaja' (Pa_224) were identified here genotypically. The last one matches genuine 'Melitopolska chorna', while a second accession (Pa_88) carries a different genotype and thus represents a homonym within the German collection itself. Further accessions are held there under tentative Ukrainian names that were not resolved, namely 'Debanka-?' (Pa_74) and 'Rekordiska-?' (Pa_91), both annotated as cultivars of Ukrainian origin while leaving the accession itself unverified (Ordidge, Litthauer et al. 2021). Retrieval is further impeded by taxonomy (EURISCO 2026), since part of this material is filed under *Cerasus avium* (L.) Moench rather than *Prunus avium* L. and is missed by a query on the currently accepted name.

Provenance can be ascertained solely through genotypic analysis. This has implications extending beyond curation, as several of the sampled collections may no longer be accessible for reinspection (Udovychenko, Grynyk et al. 2024). Additionally, for certain germplasm, the remaining material is dispersed among various institutional holdings, private collections, and international genebanks. When the original tree is unavailable for resampling, the material can only be re-identified through genotype matching across collections. This process, in turn, requires profiles recorded on a standardized allele scale (Section 2.3) and stored alongside the passport data, thereby enabling recognition of the accessions reported herein in other contexts.

Several S-genotypes documented for this germplasm by the present author necessitate reattribution rather than correction: the genotypes remain valid, but the cultivar names assigned to them are incorrect. Accessions GEN/1 and GEN/2 ('Heneralska'), VAL/3 ('Valerii Chkalov'), TAY/1 ('Taina'), and ZOR/1 ('Zoriana') are demonstrated to be mislabeled (Section 3.2). Due to the unavailability of additional accessions for these cultivars, none of the four cultivars are represented in this study; their reported S-genotypes—including the three-allele profile of GEN/1, previously interpreted as indicative of potential triploidy—pertain to unidentified trees. For 'Yaroslavna', 'Donetska krasavytsia', and 'Siestronka', the mislabeled accession (YAR/1, DON/1, SES/1) is accompanied by a verified accession (YAR/2, DON/2, SES/2), thereby supporting the cultivar records based on the latter. The accession maintained as 'Drogans gelbe Knorpel' (DRO/1) bears the S_3S_6 genotype characteristic of 'Dönissens gelbe Knorpel' (Section 3.3). Cross-incompatibility group assignments were based on the last systematic review (Schuster 2020), and the revised attributions are accordingly updated in Table S4.

The corrected identification of the 'Drogans gelbe Knorpel.' / 'Dönissens gelbe Knorpel.' reference accessions affects the historical record in a specific way. Cultivars documented as open-pollinated progeny of 'Drogans gelbe Knorpel.'—namely 'Liubava', 'Yaroslavna', 'Donchanka', 'Aelita', and others—are hereby confirmed to be descendants of 'Drogans gelbe Knorpel.', based on accession #110 [FRA], which the reference comparison designates as authentic (Section 3.3). Consequently, the breeding records are accurate, and the materials used in the original crosses were correctly identified at the time of their creation. The discrepancy lies instead with the accessions of 'Drogans gelbe Knorpel.' (DRO/1 [UKR] and DRO/2 [UKR]) that were maintained and exhibit the genotype and S-genotype characteristic of 'Dönissens gelbe Knorpel.'. Therefore, the error is related to recent propagation or sampling within the living collection rather than to the documented pedigrees. The confusion between these two yellow-fruited cultivars appears to be common in European collections rather than specific to Ukrainian ones, and any pedigree claiming either as a parent should be considered provisional until the accession is genotyped.

This also constrains the interpretation of earlier molecular research on this germplasm. Of the 24 accessions fingerprinted using ISSR (Ivanovych, Udovychenko et al. 2017), 16 share sample identifiers with the present study, four of which are shown here to be mislabeled (VAL/3, DON/1, YAR/1, and DRO/1); the remaining eight are different trees bearing the same cultivar names and were not cross-verified. Consequently, the groupings reported in those studies that involve the affected accessions accurately reflect the sampled individual trees rather than the cultivars to which they were attributed, and are therefore not considered independent lines of support. The same considerations apply to the microsatellite analysis of 27 cultivars (Ivanovych and Volkov 2018). Three accessions—YAR/1, TAY/1, and DON/1—are identified herein as mislabeled, with one of these explicitly observable in the published dendrogram: 'Taina' (TAY/1) clusters immediately adjacent to 'Annushka' (ANN/1) within subcluster B1. This clustering is not attributable to a close pedigree relationship but rather to the two accessions being clones (Section 3.2). More notably, the assignment of F_1 or F_2 progeny of 'Valerii Chkalov' relied on documented pedigrees rather than a genotyped reference, as no accession of that cultivar was incorporated into the analysis. Furthermore,

since no authentic 'Valerii Chkalov' was available in the current study (Section 3.2), such inferences remain unverified.

Two circumstances of the period in which these records were compiled bear on their interpretation. Breeding by open pollination was administratively regarded as a low standard of work, and breeders were at times obliged to record an inferred rather than an observed origin when reporting results (V.M. Mezhenyskyi, pers. comm., 2020). The practice was not uniform: open pollination is explicitly recorded for 'Donchanka', 'Aelita', 'Liubava' and 'Naiada', so where a controlled cross is stated, it cannot be assumed to be a substitution. It does, however, provide a mechanism for records that cannot be reconciled with any tested genotype, and for cases where one recorded parent is not otherwise documented at all: the maternal parent of 'Donetska krasavytsia' is given as 'Drohana rozheva', a name that appears in this pedigree statement alone and in no register or pomological source. Both accessions bearing the cultivar name were rejected here, and 'Drogans gelbe Knorpel.' and 'Dönissens gelbe Knorpel.' were excluded as parents of either. Six of the 13 rejected pedigrees originate from the Bakhmut programme, where the reporting conditions described above applied.

In summary, the corrected assignments suggest a breeding history rooted in a narrow founder base: a single cultivar, 'Drogans gelbe Knorpel.' is confirmed as the parent of 17 examined accessions, four of which were documented under the name of its sport, 'Bigarreau Napoléon blanc' (Symyrenko 1963), and are genetically indistinguishable from it. Concurrently, 17 of the 50 accessions with revised assignments could not be aligned with their recorded parentage and are designated as unknown. Pedigree records for this germplasm should be regarded as hypotheses requiring verification rather than as established facts when selecting parentage for new crosses.

These findings possess practical significance for collection documentation. In cases where a cultivar name includes two distinct genotypes, such as 'Valerii Chkalov', both should be recorded under separate accession identifiers instead of one being omitted, as neither can presently be confirmed as the authentic clone. When published S-genotypes originate from trees of unverified identity, the record should be reassigned to the corresponding accession rather than corrected under the cultivar name, since the genotype itself is not inaccurate. For international compilations, an S-genotype is interpretable solely if it is reported against a traceable accession, considering that the same cultivar name may denote different trees across various collections. Material distributed as budwood or scion should include its verified genotype along with the name, and pedigrees that cannot be validated against available parental material should be marked as unverified in passport data rather than silently inherited by subsequent records.

Several constraints bound these conclusions. While the marker panel effectively resolves clonal identity and reliably supports parentage exclusion, eighteen loci do not confidently distinguish among close relatives beyond the first degree; therefore, sibling and half-sibling assignments are reported as consistent with the marker data rather than definitively demonstrated. S-genotyping relies on fragment length at three markers, and alleles that could not be matched to reference cultivars were not confirmed by sequencing due to the unavailability of remaining material; since fragments of identical length may originate from unrelated alleles, such variants are reported without designation and are excluded from all identity and parentage inferences. The two loci most effective in distinguishing the yellow-fruited founder cultivars within the European reference set, CPPCT022 and CPPCT006, are absent from this panel. Consequently, the reclassification of those accessions depends on S-locus constitution, calibrated shared loci, and pedigree concordance rather than on diagnostic markers; genotyping either locus would independently verify such classifications. All genotyping was conducted within a single laboratory without inter-laboratory replication; calibration to the reference scale mitigates but does not eliminate potential discrepancies.

Lastly, the accessions constitute a curated collection rather than a population sample, and the structure analysis is therefore descriptive.

5. Conclusions

Molecular identification of 94 sweet cherry accessions from Ukrainian collections, anchored to European reference material through locus-specific allele-size calibration, demonstrates that cultivar names within this germplasm cannot be regarded as dependable identifiers. Of the five cultivar names represented by two accessions each, three were identified as homonyms rather than duplicates, and an additional two accessions were found to be mislabeled propagations of other cultivars. Four accessions were eliminated as clones, and six were retained under provisional identifiers as mislabeled yet genetically distinct material.

The most significant outcome pertains to the two yellow-fruited founder cultivars. Accessions identified as 'Drogans gelbe Knorpel.' and 'Dönissens gelbe Knorpel.' each incorporate the genotype of the other, and this reciprocal exchange is also observed within the European reference set. The two differentiate at the *S*-locus (S_1S_5 and S_3S_6), thereby placing them in distinct cross-incompatibility groups. Although both are early-selection cultivars, they are extensively held within collections and recur throughout the documented pedigrees of the modern assortment. Consequently, this exchange influences parentage records and any *S*-genotype designation made by name rather than by genotype. It is important to note that this error impacts the accessions preserved within the living collection rather than the historical breeding records; cultivars recorded as progeny of 'Drogans gelbe Knorpel.' are validated against the authentic genotype, confirming that the parental material was accurately identified at the time of crossing and that the error originated during subsequent propagation.

The curation of *S*-genotypes in comparison with the published record has resolved six of the eleven cultivars exhibiting conflicting genotypes. Moreover, it has identified four cultivar names whose previously published *S*-genotypes correspond to unidentified trees, thereby necessitating reattribution rather than correction. *S*-genotypes are reported here for the first time for 33 named Ukrainian cultivars and breeding selections, thereby expanding the *S*-genotyped Ukrainian assortment by half relative to the 66 cultivars for which published records already existed. A total of nineteen *S*-alleles were documented, with additional alleles predominantly found in wild and landrace material.

Parentage testing confirmed three recorded pedigrees in full and one parent in an additional 25 accessions, rejected the recorded parentage of 13, and proposed four new assignments. 'Drogans gelbe Knorpel.' was confirmed as a parent of 17 accessions, four of which were recorded under the name of its somatic sport, 'Bigarreau Napoléon blanc,' and are therefore genetically indistinguishable from it. Testing of a further 19 recorded pedigrees is blocked by the absence of an authenticated 'Valerii Chkalov', and the exclusion pattern indicates that two distinct trees were propagated under that name in the Melitopol and Bakhmut programs. Recorded pedigrees for this germplasm should therefore be treated as hypotheses requiring verification before use in parent selection.

When analyzed as a background component, the genetic structure distinguishes cultivated, wild, and introduced materials using both methodologies. At a threshold of $Q \geq 0.80$, DAPC correctly classified 83 of 90 accessions across three cohorts, while STRUCTURE correctly assigned 82 of 90 accessions across two cohorts. The two methods concurred on the classification of bred versus wild and introduced groups for 76 of 90 accessions, with only a single reciprocal disagreement. These findings, with $K = 3$ as determined by DAPC and $K = 2$ as identified by STRUCTURE, are consistent with a breeding collection in which a limited number of founders recur throughout the pedigrees.

In germplasm management, claims of identity must be substantiated by the convergence of independent evidence—namely SSR profiles, S-genotypes, and parentage exclusions. Furthermore, the individual tree code, rather than the cultivar label, must serve as the primary record until such convergence is confirmed. Applying this protocol to Ukrainian collections requires genotyping with the comprehensive 14-locus ECPGR standard set, including the S-locus; depositing the resulting reference profiles alongside passport data; and permanently affixing machine-readable labels to individual trees, linked to a unique accession identifier. This approach ensures that identity information remains with the propagule rather than its planting location. Routine verification at each propagation cycle, along with re-verification of any accession exchanged with other collections, will mitigate the recurrence of labeling errors documented herein. Establishing a verified reference collection through the Institute of Horticulture NAAS network would facilitate the identification and consolidation of dispersed Ukrainian germplasm across collections.

Declarations

Author contributions. Yaroslav Ivanovych conceived and designed the study, carried out the laboratory work and data analysis, prepared all figures and tables, and wrote the manuscript. The author read and approved the final manuscript.

Funding. This work was supported by the COST Association through COST Action FA1104, "Sustainable production of high-quality cherries for the European market," which funded the author's Short-Term Scientific Mission (COST-STSM-ECOST-STSM-FA1104-180116-070455) to INRA Bordeaux in 2016.

Competing interests. The author declares that he has no known competing financial interests or personal relationships that could appear to influence the work reported in this paper.

Data availability. All data generated or analyzed in this study are included in this article and its supplementary information files. No nucleotide sequence data were generated in this study; all genotyping was based on PCR fragment-length analysis. The complete SSR and S-locus genotype data, accession passport information, and the allele-size calibration to the ECPGR reference scale are provided in Supplementary Files 1–2. No public sequence database accessions are applicable, as no sequence data were produced.

References

1. Aranzana MJ, Garcia-Mas J, Carbó J, Arús P (2002) Development and variability analysis of microsatellite markers in peach. *Plant Breed* 121:87–92. <https://doi.org/10.1046/j.1439-0523.2002.00656.x>
2. Association of Melitopol Cherry Producers (2020) Specification for the geographical indication “Melitopol Cherry” [Spetsyfikatsiia geografichnoho zaznachennia “Melitopolska chereshnia”]. Industrial Property [Promyslova vlasnist] 8 – 2:1185–1190. Ukrainian National Office for Intellectual Property and Innovations (UANIPIO), Kyiv
3. Barreneche T, Cárcamo de la Concepción M, Blouin-Delmas M, Ordidge M, Nybom H, Lacis G, Feldmane D, Sedlak J, Meland M, Kaldmäe H, Kahu K, Békefi Z, Stanivuković S, Đurić G, Höfer M, Galik M, Schüller E, Spornberger A, Sirbu S, Drogoudi P, Agulheiro-Santos AC, Kodad O, Vokurka A, Lateur M, Fernández Fernández F, Giovannini D, Quero-García J (2021) SSR-based analysis of genetic diversity and structure of sweet cherry (*Prunus avium* L.) from 19 countries in Europe. *Plants* 10:1983. <https://doi.org/10.3390/plants10101983>

4. Bocharova BR, Kovaliova IA, Mazurenko LS (2009) Identification of grapevine clone genotypes by use of microsatellite markers. *Cytol Genet* 43:371–378. <https://doi.org/10.3103/S0095452709060024>
5. Békefi Z (2006) Review of sweet and sour cherry incompatibility. *Int J Hortic Sci* 12:111–116. <https://doi.org/10.31421/IJHS/12/2/642>
6. Campoy JA, Lerigoleur-Balsemin E, Christmann H, Beauvieux R, Girollet N, Quero-García J, Dirlewanger E, Barreneche T (2016) Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of *Prunus avium* L. landraces and bred cultivars. *BMC Plant Biol* 16:49. <https://doi.org/10.1186/s12870-016-0712-9>
7. Clarke JB, Sargent DJ, Bošković RI, Belaj A, Tobutt KR (2009) A cherry map from the inter-specific cross *Prunus avium* ‘Napoleon’ × *P. nipponica* based on microsatellite, gene-specific and isoenzyme markers. *Tree Genet Genomes* 5:41–51. <https://doi.org/10.1007/s11295-008-0166-9>
8. Clarke JB, Tobutt KR (2003) Development and characterization of polymorphic microsatellites from *Prunus avium* ‘Napoleon’. *Mol Ecol Notes* 3:578–580. <https://doi.org/10.1046/j.1471-8286.2003.00517.x>
9. Clarke JB, Tobutt KR (2009) A standard set of accessions, microsatellites and genotypes for harmonizing the fingerprinting of cherry collections for the ECPGR. *Acta Hortic* 814:615–618. <https://doi.org/10.17660/ActaHortic.2009.814.104>
10. Dirlewanger E, Cosson P, Tavaud M, Aranzana J, Poizat C, Zanetto A, Arús 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* 105:127–138. <https://doi.org/10.1007/s00122-002-0867-7>
11. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15
12. Earl DA, vonHoldt BM (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>
13. Ellson J, Gansner ER, Koutsofios E, North SC, Woodhull G (2004) Graphviz and Dynagraph – static and dynamic graph drawing tools. In: Jünger M, Mutzel P (eds) *Graph drawing software*. Springer, Berlin, Heidelberg, pp 127–148. https://doi.org/10.1007/978-3-642-18638-7_6
14. EURISCO (2026) EURISCO Catalogue: *Prunus avium* L. accessions (Ukraine, Poland, Germany, and other European national inventories). European Search Catalogue for Plant Genetic Resources, IPK Gatersleben. <https://eurisco.ipk-gatersleben.de>. Accessed 28 Jul 2026
15. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
16. Ganopoulos IV, Kazantzis K, Chatzicharisis I, Karayiannis I, Tsaftaris AS (2011) Genetic diversity, structure and fruit trait associations in Greek sweet cherry cultivars using microsatellite based (SSR/ISSR) and morpho-physiological markers. *Euphytica* 181:237–251. <https://doi.org/10.1007/s10681-011-0416-z>
17. Gascuel O (1997) Concerning the NJ algorithm and its unweighted version, UNJ. In: Mirkin B, McMorris FR, Roberts FS, Rzhetsky A (eds) *Mathematical hierarchies and biology*, vol 37. American Mathematical Society, Providence, pp 149–170
18. Gascuel O, Steel M (2006) Neighbor-Joining revealed. *Mol Biol Evol* 23:1997–2000. <https://doi.org/10.1093/molbev/msl072>

19. Guarino C, Santoro S, De Simone L, Cipriani G (2009) *Prunus avium*: nuclear DNA study in wild populations and sweet cherry cultivars. *Genome* 52:320–337. <https://doi.org/10.1139/g09-007>
20. Hegedűs A, Taller D, Papp N, Szikriszt B, Ercisli S, Halász J, Stefanovits-Bányai É (2013) Fruit antioxidant capacity and self-incompatibility genotype of Ukrainian sweet cherry (*Prunus avium* L.) cultivars highlight their breeding prospects. *Euphytica* 191:153–164. <https://doi.org/10.1007/s10681-013-0919-x>
21. Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P (2005) Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map. *Genetics* 171:1305–1309. <https://doi.org/10.1534/genetics.105.043661>
22. Iezzoni AF (2005) Acquiring cherry germplasm from Central and Eastern Europe. *HortScience* 40:304–308. <https://doi.org/10.21273/HORTSCI.40.2.304>
23. Ivanovych Y, Volkov R (2018) Genetic relatedness of sweet cherry (*Prunus avium* L.) cultivars from Ukraine determined by microsatellite markers. *J Hortic Sci Biotechnol* 93:64–72. <https://doi.org/10.1080/14620316.2017.1342568>
24. Ivanovych YI, Udovychenko KM, Bublyk MO, Volkov RA (2017a) ISSR-PCR fingerprinting of Ukrainian sweet cherry (*Prunus avium* L.) cultivars. *Cytol Genet* 51:40–47. <https://doi.org/10.3103/S0095452717010066>
25. Ivanovych YI, Triapitsyna N, Udovychenko K, Volkov R (2017b) Self-incompatibility allele identification in Ukrainian sweet cherry (*Prunus avium* L.) cultivars. *Visn ukr tov genet sel* 15:150–158. <https://doi.org/10.7124/visnyk.utgis.15.2.873>
26. Jakobsson M, Rosenberg NA (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>
27. Jombart T, Devillard S, Balloux F (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>
28. Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* 16:1099–1106. <https://doi.org/10.1111/j.1365-294X.2007.03089.x>
29. Karastan OM, Muliukina NA, Papina OS (2018) Verification of grape pedigree by microsatellite analysis. *Cytol Genet* 52:331–342. <https://doi.org/10.3103/S0095452718050031>
30. Kling D, Tillmar AO, Egeland T (2014) Familias 3 – extensions and new functionality. *Forensic Sci Int Genet* 13:121–127. <https://doi.org/10.1016/j.fsigen.2014.07.004>
31. Letunic I, Bork P (2024) Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res* 52:W78–W82. <https://doi.org/10.1093/nar/gkae268>
32. Lisek A, Rozpara E, Głowacka A, Kucharska D, Zawadzka M (2015) Identification of S-genotypes of sweet cherry cultivars from central and Eastern Europe. *Hortic Sci* 42:13–21. <https://doi.org/10.17221/103/2014-HORTSCI>
33. Liu C, Qi X, Song L, Li Y, Li M (2018) Species identification, genetic diversity and population structure of sweet cherry commercial cultivars assessed by SSRs and the gametophytic self-incompatibility locus. *Sci Hortic* 237:28–35. <https://doi.org/10.1016/j.scienta.2018.03.063>
34. Lukicheva LA (2010) Sweet cherry gene pool collection of the Nikitsky Botanical Garden [Genofondovaia kolleksiia chereschni Nikitskogo botanicheskogo sada]. In: Gene pool of southern fruit crops and its use [Genofond yuzhnykh plodovykh kultur i ego ispolzovanie], vol 132. Nikitsky Botanical Gardens – National Scientific Centre, Yalta, pp 115–129

35. Milligan BG (2003) Maximum-likelihood estimation of relatedness. *Genetics* 163:1153–1167.
<https://doi.org/10.1093/genetics/163.3.1153>
36. 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* 4:163–166. <https://doi.org/10.1111/j.1471-8286.2004.00603.x>
37. Ordidge M, Litthauer S, Venison E, Blouin-Delmas M, Fernandez-Fernandez F, Höfer M, Kägi C, Kellerhals M, Marchese A, Mariette S, Nybom H, Giovannini D (2021) Towards a joint international database: alignment of SSR marker data for European collections of cherry germplasm. *Plants* 10:1243.
<https://doi.org/10.3390/plants10061243>
38. Paprštejn F, Patzak J, Sedlák J, Henychová A (2018) Certified methodology for the evaluation of sweet cherry germplasm using molecular genetic methods [Certifikovaná metodika hodnocení genofondu třešně s využitím molekulárně genetických metod]. Research and Breeding Institute of Pomology Holovousy, Holovousy. ISBN 978-80-87030-62-2
39. Patzak J, Henychová A, Paprštejn F, Sedlák J (2020) Evaluation of S-incompatibility locus, genetic diversity and structure of sweet cherry (*Prunus avium* L.) genetic resources by molecular methods and phenotypic characteristics. *J Hortic Sci Biotechnol* 95:84–92. <https://doi.org/10.1080/14620316.2019.1647798>
40. Peakall R, Smouse PE (2012) GenAlEx 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>
41. Perrier X, Flori A, Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) *Genetic diversity of cultivated tropical plants*. Science Publishers, Enfield, pp 43–76
42. 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>
43. Risovannaia VI, Gorislavets SM, Memetova ES, Volodin VA (2014) The study of the genetic resources of Ukrainian grapes in the framework of international projects [Izuchenie geneticheskikh resursov vinograda Ukrainy v ramkakh mezhdunarodnykh proektov]. *Vinogradarstvo i Vinodelie = Viticulture and Winemaking* 44:10–13
44. Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 18:233–234. <https://doi.org/10.1038/72708>
45. Schuster M (2012) Incompatible (S-) genotypes of sweet cherry cultivars (*Prunus avium* L.). *Sci Hortic* 148:59–73. <https://doi.org/10.1016/j.scienta.2012.09.012>
46. Schuster M (2017) Self-incompatibility (S) genotypes of cultivated sweet cherries – an overview 2017. OpenAgrar Repository. <https://doi.org/10.5073/20171213-111734>
47. Schuster M (2020) Self-incompatibility (S) genotypes of cultivated sweet cherries – an overview update 2020. OpenAgrar Repository. <https://doi.org/10.5073/20201016-141600>
48. Schuster M, Flachowsky H, Köhler D (2007) Determination of self-incompatible genotypes in sweet cherry (*Prunus avium* L.) accessions and cultivars of the German Fruit Gene Bank and from private collections. *Plant Breed* 126:533–540. <https://doi.org/10.1111/j.1439-0523.2007.01401.x>
49. Sharma K, Cachi AM, Sedlák P, Skřivanová A, Wünsch A (2016) S-genotyping of 25 sweet cherry (*Prunus avium* L.) cultivars from the Czech Republic. *J Hortic Sci Biotechnol* 91:117–121.
<https://doi.org/10.1080/14620316.2015.1110997>
50. Sonneveld T, Robbins TP, Tobutt KR (2006) Improved discrimination of self-incompatibility S-RNase alleles in cherry and high-throughput genotyping by automated sizing of first intron polymerase chain reaction products.

51. Sonneveld T, Tobutt KR, Robbins TP (2003) Allele-specific PCR detection of sweet cherry self-incompatibility (S) alleles S1 to S16 using consensus and allele-specific primers. *Theor Appl Genet* 107:1059–1070. <https://doi.org/10.1007/s00122-003-1274-4>
52. Symyrenko LP (1963) Pomology, vol 3: stone fruits, quince, southern rowan, dogwood, medlar, cultivated hazelnuts and forest hazel [Pomologiya T. 3. Kostochkovye porody, aiva, riabina iuzhnaia, kizil, mushmula, funduki sadovye i lesnoi leshchinnyi orekh]. Malakhova PF (ed). State Publishing House of Agricultural Literature of the Ukrainian SSR [Gos. izd-vo selskokhoz. lit. USSR], Kyiv
53. Taranenko LI (2009) Stone fruit crops cultivars for Donbass and northern regions of Ukraine [Sorta kostochkovykh kultur dlia Donbassa i severnykh regionov Ukrainy]. *Plant Var Stud Prot* 2:85–95. [https://doi.org/10.21498/2518-1017.2\(10\).2009.59582](https://doi.org/10.21498/2518-1017.2(10).2009.59582)
54. Taranenko LI (2011) Isolation of varieties of stone fruit crops that are promising for cultivation in the northern regions of Ukraine [Vydelenie sortov kostochkovykh kultur, perspektivnykh dlia vzdelyvaniia v severnykh raionakh Ukrainy]. *Plod Vinograd Yuga Rossii* 10:49–56
55. Testolin R, Marrazzo T, Cipriani G, Quarta R, Verde I, Dettori MT, Pancaldi M, Sansavini S (2000) Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. *Genome* 43:512–520. <https://doi.org/10.1139/g00-010>
56. Turovtsev MI, Turovtseva VO, Turovtseva NM (2009) Sweet cherry (*Cerasus avium* Moench.) breeding in the Institute of Irrigating Horticulture named after M. Sydorenko of UAAS. *Sci Bull Nat Univ Live Environ Sci Ukraine* 13:51–58
57. Turovtsev MI, Turovtseva VO, Turovtseva NM (2011) Sweet cherries breeding work results in the Institute of Irrigating Horticulture named after M.F. Sidorenko NAAS [Rezultaty selektsiinoi roboty z chereshneiu v Instytuti zroshuvanoho sadivnytstva im. M.F. Sydorenka NAAN]. *Ukr J Ecol* 1:81–86
58. Turovtsev NI, Taranenko LI, Pavliuk VV (2004) Plum, sour cherry, sweet cherry [Sliva, vishnia, chershnia], vol 4. Urozhai, Kyiv
59. UANIPIO (2024) The Ukrainian government delegation appeals to the WIPO General Assembly for theft by Russia of the geographical indication “Melitopol Cherry”. Ukrainian National Office for Intellectual Property and Innovations, Kyiv. <https://nipo.gov.ua>. Accessed 30 Jul 2026
60. Udovychenko K, Grynyk I, Budzanivska I, Garkava-Gustavsson L, Yareshchenko O (2024) Maintaining continuity in the fruit growing sector in the context of the war in Ukraine. *CABI Rev* 19:1–7. <https://doi.org/10.1079/cabireviews.2024.0004>
61. Udovychenko K, Skytte af Sättra J, Moks M, Denancé C, Kondratenko T, Garkava-Gustavsson L (2026) Genetic diversity in Ukrainian, Swedish, and Estonian apple germplasm revealed with SSR markers. *Tree Genet Genomes* 22:17. <https://doi.org/10.1007/s11295-026-01744-x>
62. Udovychenko KM, Ivanovych YI, Triapitsyna NV (2016) Association of allele variants of gene Rf with red fruit color of apple cultivars of Ukrainian breeding. *Agric Sci Pract* 3:54–59. <https://doi.org/10.15407/agrisp3.03.054>
63. UIPVE (2026) State Register of Plant Varieties Suitable for Dissemination in Ukraine [Derzhavnyi reiestr sortiv roslyn, prydatnykh dlia poshyrennia v Ukraini]. Ukrainian Institute for Plant Variety Examination, Kyiv. <https://sops.gov.ua>. Accessed 30 Jul 2026
64. 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* 4:429–431. <https://doi.org/10.1111/j.1471-8286.2004.00673.x>

65. Vaughan SP, Russell K, Sargent DJ, Tobutt KR (2006) Isolation of S-locus F-box alleles in *Prunus avium* and their application in a novel method to determine self-incompatibility genotype. *Theor Appl Genet* 112:856–866.
<https://doi.org/10.1007/s00122-005-0187-9>
66. Yamamoto T, Mochida K, Imai T, Shi YZ, Ogiwara I, Hayashi T (2002) Microsatellite markers in peach [*Prunus persica* (L.) Batsch] derived from an enriched genomic and cDNA libraries. *Mol Ecol Notes* 2:298–301.
<https://doi.org/10.1046/j.1471-8286.2002.00242.x>

Figures

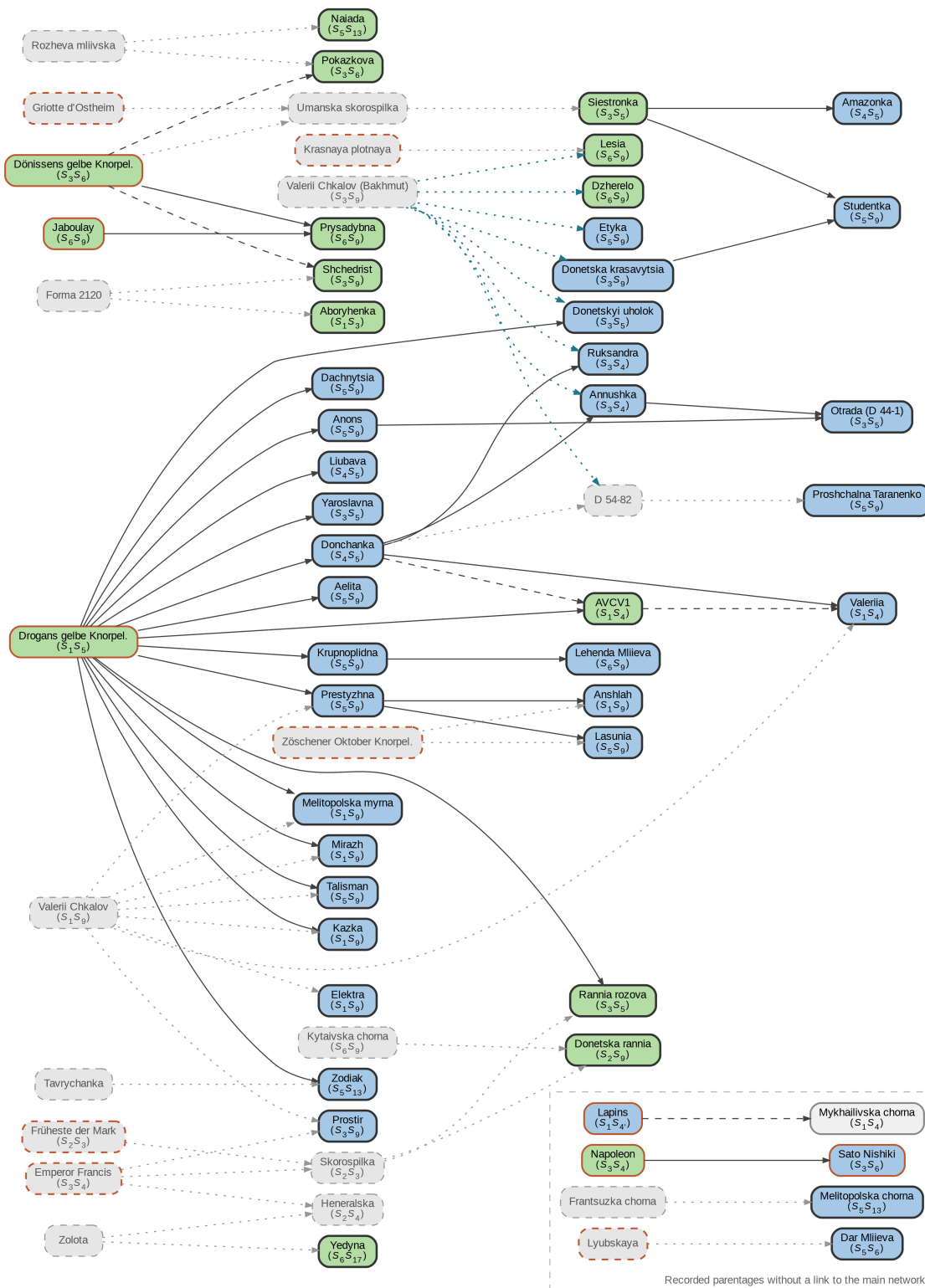


Figure 1

Pedigree network of Ukrainian sweet cherry germplasm, reconstructed from SSR parentage testing and S-genotypes. Nodes represent cultivars and breeding forms; the S-genotype is shown in parentheses beneath each name, with the accession code omitted for clarity (see Table S1 for codes). Node fill color indicates germplasm status – blue for modern cultivars, green for early selection cultivars, while gray indicates an ancestor documented in the literature but not genotyped in this study. Node borders denote origin: black for Ukrainian accessions, orange for non-Ukrainian reference cultivars, and gray for unresolved origins. Edges represent parentage: solid lines are

supported by SSR and S-genotype data; dashed lines indicate candidate assignments compatible with but not directly supported by the data; gray dotted lines signify parentage recorded in the literature when the putative ancestor could not be genotyped; teal dotted lines show descent from a putative second cultivar propagated under the name 'Valerii Chkalov' (see the text). When two dotted edges converge on a single node, they represent mutually exclusive alternatives rather than the two parents of a single cross. Records of open-pollination involving pollen mixtures are omitted, as a mixture component cannot be designated as a specific parent. The enclosed group at the lower right contains recorded parentages without a connection to the main network. Supporting data are provided in Tables S1 (assignments), S6 (pairwise results), and S4 and S10 (S-genotypes).

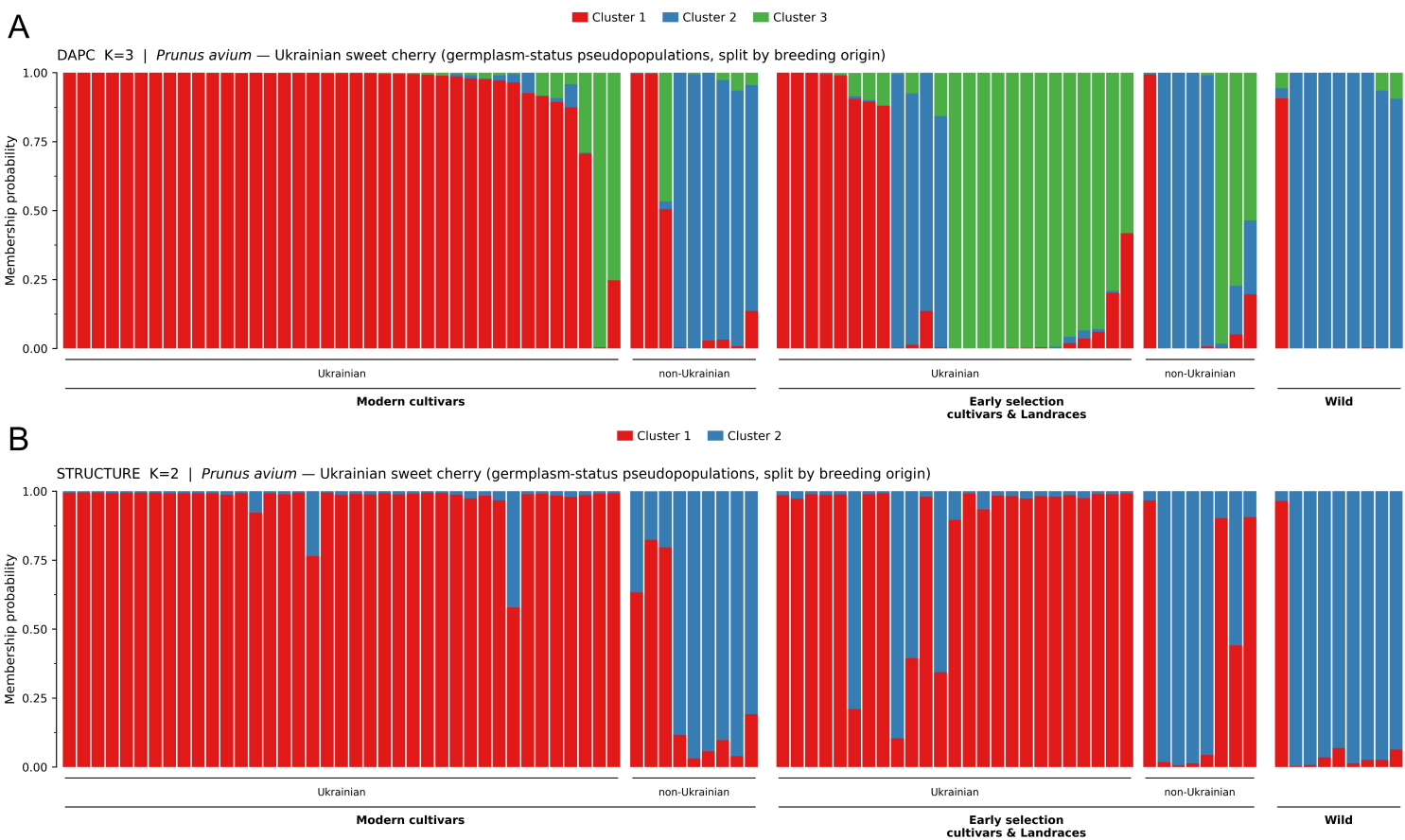


Figure 2

Genetic structure of 90 clone-corrected sweet cherry (*Prunus avium* L.) accessions based on 18 neutral SSR loci. **(A)** DAPC ($K = 3$, selected by Bayesian Information Criterion); **(B)** STRUCTURE ($K = 2$, selected by ΔK), illustrating CLUMPP-averaged ancestry proportions over twenty iterations. Each vertical bar denotes one accession, segmented into membership coefficients: Cluster 1 (red), Cluster 2 (blue), Cluster 3 (green). Cluster numbering is arbitrary and independent across panels. The accessions are arranged in the same sequence in both panels, ordered by germplasm status (Modern cultivars; Early selection cultivars & Landraces; Wild), and within each status, by breeding origin (Ukrainian / non-Ukrainian), then sorted according to membership in Panel A (and as follows in **Figure S2**). These classifications were not provided as prior information to either analysis. At a threshold of $Q \geq 0.80$, a total of 83 of 90 accessions were assigned to a cluster by DAPC, and 82 of 90 by STRUCTURE.



Figure 3

Neighbor-joining dendrogram of 87 sweet cherry (*Prunus avium* L.) accessions based on 18 neutral SSR loci. The tree was constructed from simple-matching dissimilarities by the unweighted neighbor-joining method and is shown midpoint-rooted; bootstrap values from 10,000 resamplings are given at nodes with $\geq 70\%$ support. Leaf labels give the resolved cultivar name followed by country and sample code; accessions whose identity could not be confirmed are labeled "Unknown cv." with the original name in brackets. Label color indicates germplasm status: modern cultivars in dark blue, early selection cultivars in light blue, landraces in orange, wild accessions in green, and

unresolved in grey. Grey circles denote accessions that remained admixed ($Q < 0.80$) in the DAPC. The inner strip displays individual DAPC membership coefficients ($K = 3$, scale 50–100%); the outer strip illustrates breeding origin (Ukrainian in blue, non-Ukrainian in yellow, unknown in grey).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [IvanovychSupplementaryFile1.xlsx](#)
- [IvanovychSupplementaryFile2.xlsx](#)
- [IvanovychSupplementaryFile3.docx](#)
- [UACHerryIDIvanovychtables.docx](#)