Synergistic approach of PCR-based fragment length analysis and amplicon deep sequencing reveals rich diversity of S-alleles in sweet cherries from the Caucasian region of origin

Susan Schröpfer[1]*, Mirko Schuster[1], José Quero-Garcia[2], Gregorio López-Ortega[3] and Henryk Flachowsky[1]

[1]Julius Kühn-Institut (JKI), Federal Research Centre for Cultivated Plants, Institute for Breeding Research on Fruit Crops, Dresden-Pillnitz, Germany, [2]UMR BFP, INRAE, University of Bordeaux, Villenave'Ornon, France, [3]Atlantic Green S.L, Ctra. Almonte-El Rocío, Huelva, Spain

Introduction: The self-incompatibility system in sweet cherry (*Prunus avium* L.) prevents fertilization with own or genetically related pollen, and is genetically determined by the multi-allelic S-locus. Therefore, determining S-alleles is crucial for plant breeding and fruit production, as it enables the selection of compatible combinations of S-genotypes for successful pollination.

Methods: In this study, S-alleles were identified in a total of 260 genotypes from the Caucasian region, the species' center of origin. S-allele genotyping was conducted using PCR fragment length analysis with the standard marker PaConsI-F/R2 and reference genotypes, complemented by sequence analysis through amplicon deep sequencing.

Results and discussion: The genotypes collected from Azerbaijan and Turkey exhibit a high allelic richness at the S-locus, particularly compared to modern sweet cherry cultivars worldwide. Nine previously undescribed S-alleles were identified and designated as S_{45}, S_{46}, S_{47}, S_{48}, S_{49}, S_{50}, S_{51}, S_{52} and S_{53}. Given the expected high diversity for other traits, this plant material represents a valuable resource for further breeding research and introgression of new traits in future breeding programs. Furthermore, our results underscore that fragment length alone may not be sufficient for unambiguous assignment of S-alleles due to minimal length differences between different alleles. To address this issue, an S-allele reference ladder was developed using the rich diversity for precise

assignment of the *S*-alleles. This tool can be applied in future experiments as a robust and cost-effective method for accurate *S*-genotyping across different runs and laboratories. Additionally, several selected *S*-genotypes were planted in a trial field and will be maintained as an S-allele reference collection.

1 Introduction

Sweet cherry (*Prunus avium* L.) belongs to the most popular temperate fruit crops. Total global production of sweet cherry in 2018 was 2.6 million tons with a steady upward trend over the past 20 years (FAOSTAT; http://www.fao.org/faostat/en/#data/QC/visualize), highlighting its increasing agronomic and economic importance worldwide. In addition to its role as a cultivated fruit tree, wild *P. avium* trees coexist in the same geographic areas as an important component of natural forest ecosystems in Europe, Asia, and Africa (Pinosio et al., 2020).

The origin of modern sweet cherry cultivars is believed to be in countries south of the Caucasus around the Caspian and Black Seas. Birds probably drove the subsequent spread to Europe, which is considered the native territory of sweet cherry (Iezzoni et al., 2017; Dondini et al., 2018). It is known from archaeological findings that early inhabitants in Europe consumed *P. avium* as a wild fruit around 4,000 – 5,000 B.C. Domestication of sweet cherry probably occurred in the Danube Valley in the Neolithic period 4,000 years ago (Faust and Surányi, 1996). In the Roman period, sweet cherry cultivation was introduced to Europe, documented e.g. by Roman mosaics from the 3rd century found in Germany (Faust and Surányi, 1996) and written references dating to 300 B.C (Brown et al., 1996). The importance of sweet cherry cultivation increased in the 16th century, with the most intensity in Central Europe (Watkins, 1976). At that time, various cultivars were already cultivated (Burger et al., 2011). In the 19th century, early settlers from Europe introduced sweet cherries to the American continent (Brown et al., 1996; Faust and Surányi, 1996).

Little is known about the development of cherry cultivars in the Middle Ages. Cherries were likely cultivated in different regions of Europe, and thus a large number of landraces have been developed and adapted to their local conditions (Iezzoni et al., 1990; Hanke and Flachowsky, 2017). Many of these landraces were used in breeding programs and some are still used for local production. Targeted breeding of sweet cherry cultivars, nowadays practiced in many countries, is still a very young discipline and started at the beginning of the 19th century (Iezzoni et al., 2017). From a genetic point of view, modern cultivars are only a few generations distant from their early ancestors (Iezzoni et al., 1990). Breeding of new sweet cherry cultivars with better properties is difficult because the genetic basis of this species is very limited (Hanke and Flachowsky,

2017; Iezzoni et al., 2017). Studies using single nucleotide polymorphism (SNP) markers to analyze the genetic diversity of wild accessions, landraces, and modern cultivars revealed a significant genome-wide loss of variation among the three groups (Pinosio et al., 2020). Two different, successive bottlenecks have led to less diversity. The first and largest decline in diversity occurred during domestication, and a later, minor bottleneck occurred because of breeding.

In order to expand the scope of breeding material in terms of diversity, collecting, characterizing, and evaluating genetic resources from areas expected to have high diversity is beneficial to achieve future breeding goals. In 1926, the Russian botanist Nikolai I. Vavilov described the Caucasus region as one of the centers of origin of cultivated plants (Vavilov, 1926). Since the diversity of the species of interest in such centers is very high, they are ideally suited as places for collecting trips to expand the breeding gene pool. In this work, seedlings of sweet cherry landraces and wild *P. avium* trees from two countries south of the Caucasus (Azerbaijan and Turkey) were genotyped. In Turkey, the collection area was located in the southern part of the Kaçkar Mountains. The vegetation there is restricted to the river and stream valleys, and vegetation-free mountain ridges form a natural barrier between different plant populations. In Azerbaijan, which is an important fruit-growing center in the Caucasus region with diverse agro-environmental conditions, the main fruit production areas are located in the Northwest part near Guba and Khachmaz. The natural distribution of sweet cherry in Azerbaijan was described along the tributaries of the Kur River in the center and in the Talish Mountains near Lankaran in the south (Grossgejm, 1952).

One important feature with implications for the cultivation and breeding of sweet cherry is self-incompatibility (SI), which prevents fertilization with own or genetically related pollen and ensures a constant genetic exchange within and between populations (De Nettancourt, 1984). In commercial sweet cherry production, a significant percentage of the cultivated area is required for genetically compatible pollinator genotypes (Eeraerts et al., 2019), and parameters such as e.g. flowering time, picking time, and distance to the main cultivar need to be considered for optimal orchard management. The mechanism of gametophytic self-incompatibility (GSI) in *Prunus* species is determined by a single highly polymorphic genetic locus, the so-called *S*-locus. This locus

consists of two closely coupled *S*-genes responsible for self-recognition and self-rejection after pollination (Tao and Iezzoni, 2010). Genetic variants of the *S*-locus are referred to as '*S*-haplotypes' or '*S*-alleles' and typically, two different *S*-alleles are present in diploid sweet cherry genotypes. The pistil *S*-gene is specifically expressed in the diploid pistil tissue and encodes for an enzyme with ribonuclease activity, the S-RNase, which acts as cytotoxin in self-pollen tubes (Tao and Iezzoni, 2010). The pollen *S*-gene encodes for an S-haplotype-specific F-box protein (SFB) and is expressed in the haploid male gamete (Tao and Iezzoni, 2010). According to an actual model of *S-RNase*-based self-incompatibility, S-RNases are transmitted from the style to the cytoplasm of the growing pollen tube and there, a general undefined RNase inhibitor detoxifies them. It is assumed that SFB proteins from the same *S*-allele type protect S-RNases from degradation, resulting in pollen tube growth inhibition through RNA degradation by the S-RNase activity (Hua et al., 2008).

Cultivars with the same *S*-genotype are cross-incompatible and belong to the same incompatibility group (IG). PCR methods based on the detection of the length polymorphism of the *S-RNase* gene (Tao et al., 1999; Wiersma et al., 2001; Sonneveld et al., 2003; Sonneveld et al., 2006) and the *SFB* gene (Vaughan et al., 2008) have been developed for the determination of the different *S*-alleles. The improved consensus primer pair flanking the first intron of the *S-RNase* gene, namely PaConsI-F/R2, was suggested for high throughput *S*-allele genotyping (Sonneveld et al., 2006). This system has become well-established and has been used extensively for *S*-allele genotyping of sweet cherry (Vaughan et al., 2008; Ganopoulos et al., 2012; Sharma et al., 2016; Cachi et al., 2017; Marchese et al., 2017; Azizi-Gannouni et al., 2018) and has also be transferred to other *Prunus* species (Halász et al., 2021a, Halász et al., 2021b; Nicolás-Almansa et al., 2023). It further served as a basis for the further development of the genotyping method, such as the one-tube reaction assay described by Cmejlova et al. (2023). Different *S*-alleles of sweet cherry were numbered consecutively and to date, 23 *S*-alleles were described in commercial sweet cherry cultivars (Schuster, 2020; Cmejlova et al., 2023). Further *S*-alleles have been identified in wild sweet cherries accessions (Cachi et al., 2017), which have not yet been found in cultivated cherries.

Although this system works well, it has two weaknesses. Firstly, the fragment lengths detected by different electrophoresis systems can vary slightly due to the differences in instrumentation and chemistry used. This makes it difficult to assign alleles accurately, especially if there are no standards with a known *S*-allele genotype for comparison. Secondly, since the assignment of *S*-alleles is based on the fragment length of amplicons, it is always uncertain whether two PCR fragments of the same length from two different genotypes also undoubtedly originate from the same *S*-allele.

In this study, seedlings of landraces and wild sweet cherry trees collected in two countries in the center of origin were investigated to (i) determine the diversity of *S*-alleles using a newly introduced *S*-allele reference ladder and amplicon deep sequencing and (ii) develop a reference set of genotypes representing the currently known *S*-alleles of sweet cherry.

TABLE 1 *S*-allele genotyping by fragment length analysis and amplicon deep sequencing.

S-allele	length (bp)			reference genotype[1]	sequence analysis			
	ABI	Beckman	Diff.		Seq-ID	best BLAST hit	ident. [%]	short description
S_{31}	203	208	5		OQ511575	JQ280529.1	100.0	*P. avium*, S-RNase-S_{31}, partial cds
S_3	229	234	5	'Nordwunder' (S_3S_{12}), 'Fertard' (S_3S_6), 'Alex' ($S_3S_{3'}$), 'Sir Tom' (S_3S_{13})	OQ511563	AJ635286	100.0	*P. avium*, S-RNase S_3, exons 1-2
$S_{3'}$	229	234	5	'Alex' ($S_3S_{3'}$)				
S_{14}	325	331	6	'Ferminia' (S_1S_{14}), 'Adriana' (S_1S_{14})	OQ511568	DQ266445	100.0	*P. avium*, S-RNase S_{14}, partial cds
S_{45}	334	340	6		OQ511581	KF975455	91.8	*P. armeniaca*, S-RNase S_{53}, complete cds
S_{29}	335	340	5		OQ511573	DQ266441	99.7	*P. avium*, S-Rnase S_{29}, partial cds
S_{18}	337	342	5		OQ511570	DQ983363 [b]	97.6	*P. tenella*, S-RNase S_4, partial cds
S_2	338	344	6	'Feria' (S_2S_4), 'Narana' (S_2S_9)	OQ511562	AJ635284	100.0	*P. avium*, S-RNase S_2 precursor, exons 1-2
S_{12}	339	345	344[a]	6(5)	'Nordwunder' (S_3S_{12}), '0900 Ziraat' (S_3S_{12})				
S_7	341	n.d.		'Benjaminler' (S_1S_7), 'Pflugwirtle' (S_7S_{19})				

(Continued)

TABLE 1 Continued

S-allele	length (bp)			reference genotype[1]	sequence analysis			
	ABI	Beckman	Diff.		Seq-ID	best BLAST hit	ident. [%]	short description
S_{46}	344	350		6		OQ511582	GU968641	95.6	*P. dulcis*, S-RNase S_{22}, partial cds
S_{47}	345	350		5		OQ511583	EU448294	95.9	*P. dulcis*, S-RNase S_{31}, partial cds
S_9	351	357		6	'Ferobi' (S_6S_9), 'Grace Star' (S_4'S_9), 'Narana' (S_2S_9)	OQ511567	AJ635271	100.0	*P. avium*, S-RNase S_9, exons 1-2
S_{48}	360	365		5		OQ511584	MG712735	75.5	*P. armeniaca*, S-RNase S_{52}, partial cds
S_{36b3}	363	368		5		OQ511579	EU042130	100.0	*P. cerasus*, S-RNase S_{36b3}, complete cds
S_{36b2}	365	370	369[a]	5(4)		OQ511578	EU042129	100.0	*P. cerasus*, S-RNase S_{36b2} complete cds
S_{21}	369	n.d.			'Talaguera Brilante' ($S_{21}S_{22}$), 'Maergelikirsche' (S_5S_{21})				
S_{34}	372	377	378[a]	5(6)		OQ511577	JQ280525	100.0	*P. avium, S-RNase S_{34}*, partial cds
S_1	374	380		6	'Ferminia' (S_1S_{14}), 'Adriana' (S_1S_{14}), 'Basler Adlerkirsche' (S_1S_5), 'Benjaminler' (S_1S_7), 'Lapins' (S_1S_4')	OQ511561	AJ635282	100.0	*P. avium*, S-RNase S_1, exons 1-2
S_{49}	376	381	380[a]	5(4)		OQ511585	AJ635280	82.5	*P. avium*, S-RNase S_{16}, exons 1-2
S_{30}	379	385		6		OQ511574	DQ266442	100.0	*P. avium*, S-RNase, S_{30} allele, partial cds
S_{50}	385	390		5		OQ511586	DQ058402	81.7	*P. dulcis*, S-RNase S_i, partial cds
S_5'	386	392		6		OQ511565	EU077235	100.0	*P. avium*, S-RNase S_5, cultivar 'Kronio', partial cds
S_5	387	393		6	'Rita' (S_5S_{22}), 'Basler Adlerkirsche' (S_1S_5), 'Maergelikirsche' (S_5S_{21})				
S_{51}	388	394	393[a]	6(5)		OQ511587	AB364469	91.6	*P. mume*, S-RNase S_8, partial cds
S_{17}	391	396		5		OQ511569	AM690354[b]	94.6	*P. webbii*, partial S-RNase S_{n3}, exons 1-3
S_{52}	394	399		5		OQ511588	DQ790375	74.2	*P. salicina*, S-Rnase S_b, partial cds
S_{42}	406	411		5		OQ511580	FN429354[b]	91.2	*P. dulcis*, S-Rnase S_{23}, complete cds
S_{32}	410	416		6		OQ511576	EU516388[b]	100.0	*P. armeniaca*, S-RNase S_{17}, partial cds
S_{22}	417	423	422[a]	6(5)	'Talaguera Brilante' ($S_{21}S_{22}$), 'Pico Colorado' (S_6S_{22}), 'Rita' (S_5S_{22})	OQ511572	EF429142	100.0	*P. avium*, S-RNase S_{22}, complete cds
S_{53}	418	424		6		OQ511589	AB364468	80.0	*P. mume*, S-RNase S7, complete cds
S_{19}	421	427		6	'Pflugwirtle B 12' (S_7S_{19})	OQ511571	GU968641[b]	95.7[b]	*P. pseudocerasus*, S-RNase S_{11}, complete cds

(Continued)

TABLE 1 Continued

S-allele	length (bp)			reference genotype[1]	sequence analysis			
	ABI	Beckman	Diff.		Seq-ID	best BLAST hit	ident. [%]	short description
S_6	436	443	7	'Ferobi' (S_6S_9), 'Fertard' (S_3S_6), 'Pico Colorado' (S_6S_{22})	OQ511566	EU077236	100.0	*P. avium*, S-RNase S_6, partial cds
S_4	445	451	6	'Feria' (S_2S_4), 'Sir Don' (S_4S_{13})	OQ511564	AJ635288	100.0	*P. avium*, S-RNase S_4, exons 1-2
$S_{4'}$	445	451	6	'Grace Star' ($S_{4'}S_9$), 'Lapins' ($S_1S_{4'}$)				
S_{13}	n.a.	n.a.		'Sir Don' (S_4S_{13}), 'Sir Tom' (S_3S_{13})				

[1]cultivars with known S-alleles used as reference for genotyping with ABI system.
[a]additional observed allele lengths with less frequency in the data set.
[b]no reference sequence (complete cds, exon 1-2) of S_{17}, S_{18}, S_{19}, S_{32} and S_{42} in NCBI database available.
n.a., no amplification; n.d., not determined.
For each S-allele, the fragment length of the PCR product amplified with the PaConsI-F/R2 primer combination and determined with the ABI and Beckman systems, is given in ascending order. Cultivars with known S-alleles that were used as reference genotypes and analyzed on the ABI system are shown. In addition, amplicon deep sequencing of PCR products from selected individuals (indicated in Supplementary Table S1) was performed and the sequence obtained (Seq-ID) was compared with the PaConsI-F/R2 amplicon length. The best hit from the BLASTn analysis of the sequence is given, including the percentage of sequence identity and a short description. Red numbers indicate allele lengths that may be critical for accurate S-allele genotyping.

2 Materials and methods

2.1 Plant material

Stones from local sweet cherry cultivars and wild sweet cherries were collected in 2008 in Turkey and 2014 in Azerbaijan, separated from the fruit flesh, air-dried and stored under dry conditions. The collection area in Turkey was located in the northeast at the southern part of the Kaçkar Mountains. In Azerbaijan, sampling was performed in the center of the country along the tributaries of the Kur river and in the Talish Mountains in the south. In both countries, stones from local sweet cherries and wild sweet cherries were collected. The stones were transferred to the Julius Kühn-Institute (JKI), Institute for Breeding Research on Fruit Crops (Dresden, Germany) and sown in soil in a greenhouse after four-month of stratification at 4°C in perlite. In the year following sowing, the trees obtained (one to ten seedlings per mother tree) were raised in a nursery and then planted on the experimental field in Dresden-Pillnitz (in 2010: collection of Turkish origin; in 2017 collection of Azerbaijan origin). The cultivars used as reference genotypes for S-allele genotyping (Table 1) were grown in the same experimental field except of 'Talaguera Brilante' and 'Pico Colorado'. Leaf material of these cultivars was kindly provided by Ana Wünsch (Agrifood Research and Technology Centre of Aragón) for reference. Selected genotypes of reference set for S-allele genotyping (Table 2) were grafted on the root stock *Prunus avium* 'Alkavo' and planted in the experimental field in 2023.

2.2 DNA extraction and S-allele genotyping

The total genomic DNA was extracted from peeled flower or leaf buds (65 – 85 mg) or leaf material (60 - 70 mg) using the DNeasy Plant Mini Kit (QIAGEN GmbH, Hilden, Germany) according to the manufactures protocol. After DNA quantification using the NanoDrop 2000c (ThermoFisher Scientific, Germany), DNA concentration was adjusted to 10 ng/µL in ddH$_2$O. The primer set PaConsI-F/PaConsI-R2 (Sonneveld et al., 2003, Sonneveld et al., 2006) was used for amplification of the first intron of the S-RNase gene. The primer PaConsI-R2 was labeled at 5'-end with 6-FAM (for analysis with the ABI system) or with BMN-5 (for measurements with the Beckman system), both supplied by Biomers.net GmbH (Ulm, Germany). The PCR was performed in a final volume of 6 to 10 µL using the Type-it Microsatellite PCR Kit (QIAGEN GmbH, Hilden, Germany) adding the Q-solution with following PCR conditions: initial denaturation at 94°C for 2 min; 35 cycles of 94°C for 1 min, 54°C for 1 min and 72°C for 1 min; and a final elongation step at 60°C for 30 min.

For fragment length analysis with the ABI system, the PCR reaction was diluted with ddH$_2$O in the ratio 1:100 or 1:200 in final volume of 200 µL, supplemented with 8.95 µL Applied Biosystems Hi-Di Formamid (Thermo Fisher Scientific Inc., USA) and 0.05 µL Applied Biosystems GeneScan 600 LIZ dye Size Standard v2.0 (Thermo Fisher Scientific Inc., USA). PCR fragment length determination was performed after denaturation at 95°C for 5 min on the ABI Genetic Analyzer 3500xl system (Applied Biosystems® ThermoFisher Scientific, Germany) using the POP-7 Polymer (Thermo Fisher Scientific Inc., USA) and analysis was done with the Applied Biosystems GeneMapper Software 6.

For analysis with the Beckman system, the PCR reaction was diluted 1:20 with ddH$_2$O and 2 µL was added to a mixture consisting of 24.9 µL sample loading solution and 0.1 µL DNA size standard 600 (Beckman Coulter, Krefeld, Germany). To prevent evaporation, a drop of mineral oil was added to the sample. Fragment length analysis was performed on the Beckman Coulter CEQ™ 8000 Genetic Analysis system (Beckman Coulter, Krefeld, Germany) equipped with a GenomeLab 33–75 separation capillary array (Beckman Coulter) and sizes were scored with the fragment analysis tool of the GenomeLab GeXP software (Beckman Coulter).

Allelic richness and private allelic richness were calculated for the S-locus using the Allelic Diversity Analyzer Version 1.0 (Szpiech et al., 2008). The maximal standardized sample size (MAX_G) was set to 50.

TABLE 2 Reference collection for *S*-allele genotyping established at JKI Dresden-Pillnitz.

Genotype	*S*-alleles	length [bp]		Reference collection	
				Ladder	Accession nr.
'Adriana'	S_1S_{14}	374	325	x	KIZ-KIA0174
'Alex'	$S_3S_{3'}$	229	229	–	KIZA04-19
'Benjaminler'	S_1S_7	374	341	x	KIZA06-5
'Feria'	S_2S_4	338	445	x	KIZA02-5
'Ferobi'	S_6S_9	436	351	x	KIZA02-1
K18-200	$S_{14}S_{29}$	325	335	x	PiSue09-08.1
K18-204	S_4S_{31}	445	203	x	PiSue09-13.4
K18-218	$S_{48}S_{49}$	360	376	x	PiSue09-14.8
K18-233	$S_{42}S_{52}$	406	394	x	PiSue09-16.3
K18-235	$S_{18}S_{42}$	337	406	x	PiSue09-16.5
K18-245	$S_{14}S_{50}$	325	385	x	PiSue09-17.5
K18-302	$S_{42}S_{53}$	406	418	x	PiSue09-23.3
K18-339	$S_{19}S_{32}$	421	410	x	PiSue09-31.10
K18-343	$S_{30}S_{51}$	379	388	x	PiSue09-32.4
K18-346	$S_{51}S_{17}$	388	391	x	PiSue09-32.8
K18-361	S_4S_{46}	445	344	x	PiSue09-11.3
K18-362	S_1S_{45}	374	334	x	PiSue09-11.4
K18-402	S_1S_{36b2}	374	365	x	PiSue15-68.1
K18-415	$S_{5'}S_{36b3}$	386	363	x	PiSue15-73.5
K18-447	S_4S_{47}	445	345	x	PiSue15-82.4
K18-469	$S_{14}S_{34}$	325	372	x	PiSue15-87.2
'Lapins'	$S_1S_{4'}$	374	445	–	KIZA07-5
'Maergelikirsche'	S_5S_{21}	369	387	–	KIA1071
'Nordwunder'	S_3S_{12}	229	339	x	KIZ-KIA0075
'Pflugwirtle B 12'	S_7S_{19}	341	421	x	KIZ-KIA1051
'Rita'	S_5S_{22}	387	417	x	KIZA04-9
'Talaguera Brilante'	$S_{21}S_{22}$	369	417	x	–

The genotypes used for the *S*-allele ladder are indicated. The indicated fragment length of the PaConsI-F/R2 PCR amplicon refers to the analysis on the ABI instrument.

2.3 Ultra-deep sequencing of *S*-allele amplicons

The first intron of *S*-alleles was amplified using the unlabeled primer set PaConsI-F/PaConsI-R2 with the Thermo Scientific™ Phusion High-Fidelity PCR Kit (Thermo Fisher Scientific Inc., USA) according to the manufactures protocol. The PCR was performed in a final volume of 50 µL using 20 ng genomic template DNA and 1x HF buffer with following PCR conditions: initial denaturation at 98°C for

30 sec; 30 cycles of 98°C for 10 sec, annealing 58°C (optimized by gradient PCR) for 30 sec and 72°C for 30 sec; and a final elongation step at 72°C for 7 min. The PCR products were purified using the Thermo Scientific™ GeneJET PCR Purification Kit (Thermo Fisher Scientific Inc., USA) or the PCR Kleen™ Purification Spin Columns (Bio-Rad Laboratories GmbH, Germany) according to the manufacturer protocols, eluted with ddH₂O. DNA quantification was done using the NanoDrop 2000c (Thermo Fisher Scientific, Germany). GENEWIZ Germany GmbH (Leipzig, Germany) provided amplicon ultra-deep sequencing on the Illumina platform (2x250 bp configuration) as well as unique sequence identification. The Basic Local Alignment Search Tool (BLAST) was used to find high similar sequences in the nucleotide collection of the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/, last access 22.09.2020) for each identified unique sequence with a portion ≥ 1% of total reads. The Qiagen CLC Main Workbench 21.0.1 (Qiagen GmbH, Hilden, Germany) was used to create the alignment and the phylogenetic trees.

3 Results

3.1 Collection of sweet cherry genotypes from the region south of the Caucasus

In June 2008, stones from 18 trees of local sweet cherry cultivars or wild sweet cherries (Supplementary Table S1) were collected in Turkey at an altitude of 800 – 1,600 m above sea level in three side valleys of the Çoruh River near the town of Yusufeli (Figure 1, sites A-J).

In 2014, stones from local cultivars and wild sweet cherries were collected in Azerbaijan, respectively. Stones from 11 genotypes were collected in the central region in Akdash, Ganja, and Tovus. Further 13 samples were taken in Lankaran and the Talish Mountains in the southern region near the border with Iran. Each sample consisted of several stones of each genotype. The stones were stratified and grown as described, and one to ten individual seedlings per mother tree were obtained. A total of 97 genotypes originating from Azerbaijan and 163 genotypes with an origin from Turkey were received and cultivated in the experimental field of the Julius Kühn-Institut in Dresden, Germany.

3.2 *S*-allele genotyping combining on PCR fragment length analysis and amplicon deep sequencing

The first intron of the *S-RNase* gene was amplified from each analyzed genotype using the primer pair PaConsI-F/R2 (Sonneveld et al., 2006). The lengths of the obtained fragments were determined and compared with two different capillary gel electrophoresis devices from the manufacturers ABI and Beckman (Table 1). For the assignment of the measured fragment lengths to already described *S*-alleles, several cultivars with known *S*-alleles
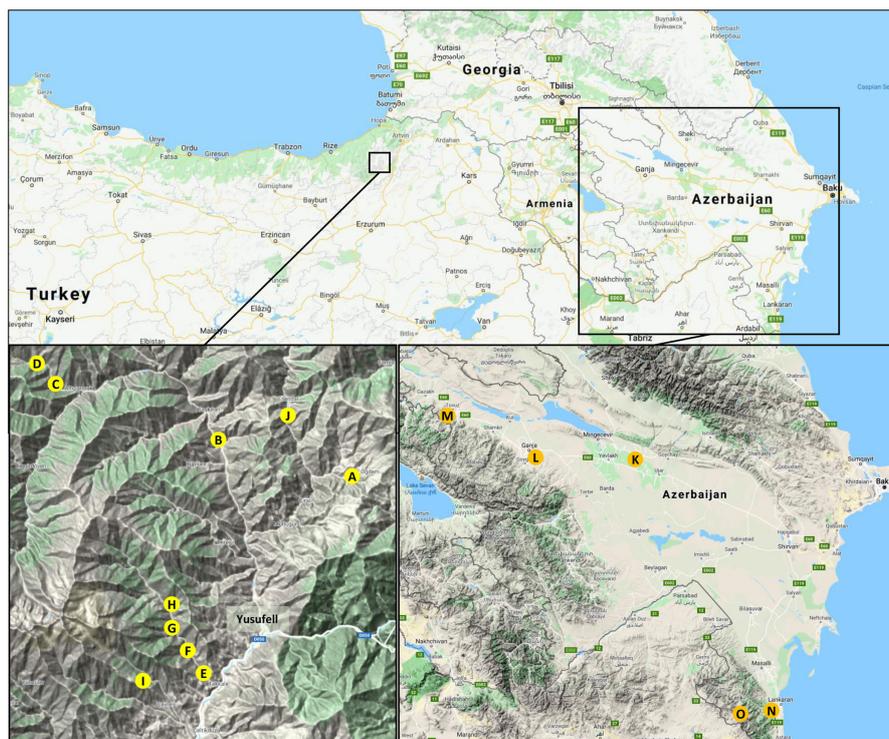
FIGURE 1
Collection areas of sweet cherry stones south of the Caucasus. The upper map gives an overview of the Caucasus region between the Black Sea and the Caspian Sea. The collection areas in Turkey and Azerbaijan, both located in the South Caucasus (Lesser Caucasus Mountains), are indicated by the squares. In the lower panels, the different collection sites are indicated by letters (left: Turkey, A-J; right: Azerbaijan, K-O). In the following, the number of genotypes analyzed is given in the form "collection site (number of mother trees, number of seedlings)". A (1, 1); B (4, 40); C (1, 10); D (2, 20), E (1, 10); F (3, 30); G (2, 19); H (1, 10); I (2, 19); J (1, 4); K (3, 5), L (2, 6); M (6, 23); N (4, 19); O (9, 44). Detailed information on the plant material collected is available in Supplementary Table S1. (map material: Google Maps, 2020).

(Schuster, 2017) were selected, analyzed on the ABI system and used as reference. For each of these *S*-alleles, one to five different cultivars served as independent references for genotyping. The *S*-allele genotyping results from each tree originating from stones collected in Azerbaijan and Turkey are presented in Supplementary Table S1. Comparing both analytical instruments, the determined fragment lengths differed in most cases by 5 to 6 bp. In addition to the fragment lengths assignable to *S*-alleles S_1, S_2, S_3/S_3', S_4/S_4', S_5, S_6, S_7, S_9, S_{12}, S_{14}, S_{19}, S_{21}, and S_{22}, further fragment lengths were observed when analyzing the collected material from the Caucasian region. For these fragments no clear *S*-allele assignment using the reference genotypes was possible. Therefore, amplicon deep sequencing of the PaConsI-F/R2 PCR products from selected individuals (Supplementary Table S1) was performed. The lengths of the sequences obtained from the sequencing approach correlated strongly with the values measured by fragment length analysis of the PCR products. BLASTn analysis of some sequences resulted in alignments with high levels of sequence identity (nearly 100%) with known *S*-allele sequences deposited in the NCBI database (Table 1). Using these sequence analyses, the *S*-alleles S_5', S_{29}, S_{30}, S_{31}, S_{34}, S_{36b2}, S_{36b3} could be linked to observed PCR fragment lengths. Further assignments to previously known *S*-alleles S_{17}, S_{18}, S_{32}, and S_{42} were made by comparing fragment lengths described in the

literature (Vaughan et al., 2008; Cachi et al., 2017). For these alleles, no reference sequence of the corresponding region of the *S-RNase* gene was available in the NCBI database.

3.3 Newly identified *S*-alleles

Nine different PCR fragments were amplified that could not be assigned to any of the previously described *S*-alleles either based on their fragment length or on the basis of BLAST analyses using the amplicon sequence. According to the *S*-allele nomenclature, these newly identified *S*-alleles have been named, in ascending order, S_{45} (334 bp), S_{46} (344 bp), S_{47} (345 bp), S_{48} (360 bp), S_{49} (376 bp), S_{50} (385 bp), S_{51} (388 bp), S_{52} (394 bp) and S_{53} (418 bp). The given fragment lengths refer to the measurements performed on the ABI instrument. Sequence alignments (Supplementary Figure S1) and a phylogram (Supplementary Figure S2) were generated with all the sequences obtained and a distinctness between the various *S*-allele sequences could be demonstrated. Particular attention should be paid to the *S*-allele sequences that are very similar in fragment length (Table 1). The newly defined alleles S_{46} and S_{47} differ by only one base pair in their fragment length using the ABI instrument. Using the Beckman system, an identical fragment length of 350 bp
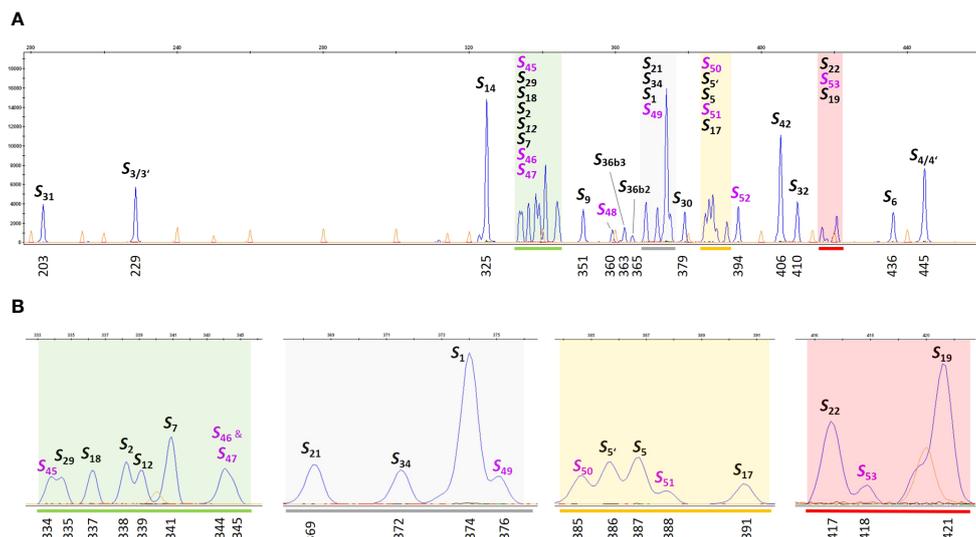
FIGURE 2
S-allele reference ladder for genotyping. *S*-allele specific fragments were amplified by PCR from reference genotypes (Table 2) using a FAM-labeled primer pair PaConsI-F/PaConsI-R2 flanking the first intron of the *S-RNase* gene. An *S*-allele reference ladder sample combining each PCR reaction was analyzed on the ABI Genetic Analyzer 3500xl system. **(A)** The diagram displays all *S*-allele specific peaks (blue) in relation to the size standard (orange). The calculated size of the fragment and the assigned *S*-allele is labeled (*S*-alleles, which were newly defined in this work are lettered in violet). **(B)** For better resolution, the colored areas in **(A)**, are displayed with a higher zoom factor.

was determined for both alleles. However, the sequence alignment of both *S*-alleles revealed only 54% sequence identity and significant differences including large sequence gaps/insertions and substitutions (Supplementary Figure S1). A high degree of genetic divergence is also shown in the phylogram, as both sequences appear in distant branches (Supplementary Figure S2). This finding is also evident for other *S*-alleles (Table 1). Sequence differences between *S*-alleles with similar fragment lengths were also found for the newly identified *S*-alleles S_{45} and S_{29}, S_{49} and S_1,

S_{50} and $S_{5'}$, S_{51} and S_5 (NCBI sequence accession AJ635290), as well as S_{53} and S_{22}.

3.4 *S*-allele reference ladder as a robust tool for genotyping

The large number of *S*-alleles studied in this work shows that the fragment lengths of different alleles are sometimes very close to
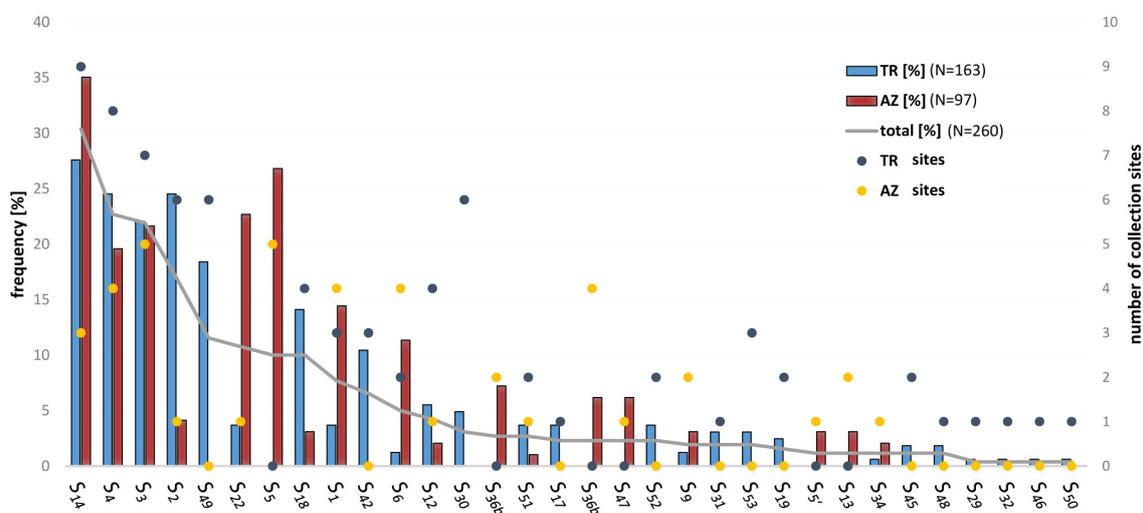

FIGURE 3
Frequency and distribution of single *S*-alleles in the collection area south of the Caucasus. The percentage frequency of *S*-alleles in the collected sweet cherries is shown as a grey line in descending order. The percentage proportion of *S*-alleles in the progeny with origin in Turkey (TR) and Azerbaijan (AZ) is represented by columns. The colored dots represent the total number of collection sites where this *S*-allele was detected (in total 10 sites in TR and 5 in AZ).

TABLE 3 Occurrence and distribution of S-alleles in genetic resources originating from Caucasian sweet cherries.

Collected in country	Turkey		Azerbaijan	
Sampled donor trees	17		24	
Analyzed individuals of the progeny	163		97	
Detected S-alleles	26		18	
common for both countries	12			
	S_1, S_2, S_3, S_4, S_6, S_9, S_{12}, S_{14}, S_{18}, S_{22}, S_{34}, S_{51}			
Unique for country	14	S_{17}, S_{19}, S_{29}, S_{30}, S_{31}, S_{32}, S_{42}, S_{45}, S_{46}, S_{48}, S_{49}, S_{50}, S_{52}, S_{53}	6	S_5, $S_{5'}$, S_{13}, S_{36b2}, S_{36b3}, S_{47}
Frequent S-alleles[1]	7	S_2, S_3, S_4, S_{14}, S_{18}, S_{42}, S_{49}	7	S_1, S_3, S_4, S_5, S_6, S_{14}, S_{22}
Occurrence at single collection sites	9	S_{17} (I), S_{22} (B), S_{29} (A), S_{31} (B), S_{32} (H), S_{34} (D), S_{46} (J), S_{48} (B), S_{50} (D)	8	S_2 (N), $S_{5'}$ (M), S_{12} (L), S_{18} (N), S_{22} (O), S_{34} (O), S_{47} (N), S_{51} (M)
Stray find	5	S_{29} (A), S_{32} (H), S_{34} (D), S_{46} (J), S_{50} (D)	1	S_{51} (M)

[1]found in more than 10% of the analyzed individuals.
The total number of different S-alleles is given and these are specified for the respective category. The indicated letters refer to the collection sites defined in Figure 1.

each other. The high density of marker alleles requires the use of suitable references for matching. To create a reference collection for S-allele genotyping, different genotypes with known S-alleles were selected to represent each S-allele at least once. These genotypes were planted in the experimental field of the Julius Kühn-Institute in Dresden, Germany (Table 2), where they will be maintained for future studies. To make the use of such a large number of references in capillary gel electrophoresis simple and practical, an S-allele reference ladder was developed, consisting of a mixture of PCR products amplified from selected reference genotypes (Table 2). The resulting chromatogram of this reference ladder is shown in Figure 2. Delineated peaks allow accurate genotyping of 31 different S-alleles. In addition, the alleles S_{46} and S_{47} form a common broad peak in the pooled reference sample. In the respective individual samples, the peaks at 344 bp and 345 bp are clearly distinguishable from each other (not shown). Differentiation between S_3 and $S_{3'}$ as well as S_4 and $S_{4'}$ is not possible with this analytical method. The S-allele reference ladder can be used as a single standard sample to calibrate run-to-run and lab-to-lab differences on any capillary electrophoresis instrument. This increases the robustness of S-allele genotyping in future studies.

3.5 Distribution of S-alleles among the genetic resources originated from the Caucasus region

Among the genotypes originated from sweet cherry stones collected in the Caucasus region, 32 different S-alleles were

observed. An overview of the distribution and frequency of these alleles in Turkey and Azerbaijan is given in Figure 3; Table 3. Twelve S-alleles were detected in progeny from both countries. Fourteen S-alleles were found only in sweet cherries originating from Turkey, whereas six S-alleles were found only in cherries from Azerbaijan. Most of the newly identified S-alleles, namely S_{45}, S_{46}, S_{48}, S_{49}, S_{50}, S_{52}, and S_{53} were found only in cherries from Turkey. Among them, the S_{49} allele was frequently observed, with a proportion of 18% in all progeny and a wide distribution in the Turkish collection region (present in six out of ten collection sites) (Supplementary Table S2). The remaining alleles were found at much lower frequencies, in most cases only at one or two collection sites (Table 3; Supplementary Table S2). S_{47} was found only in cherries from Azerbaijan with a frequency of 6% and S_{51} was detectable in plants from both countries with a low abundance (4% in the Turkish collection, 1% in the Azerbaijan collection). The most frequent allele in cherries from both countries was S_{14}, followed by S_4 and S_3, which were also found in cherries from both countries (Figure 3). In addition, the alleles S_2, S_{49}, S_{22}, and S_5 also show high frequencies, but with a clear correlation with the collection region. S_2 and S_{49} are characteristic of the Turkish material and S_{22} and S_5 occur mainly in the material collected in Azerbaijan.

One hundred and six S-genotypes were be detected in total. The frequency of each S-allele genotype and its distribution in the collection material according to its origin is shown in Supplementary Table S3. There are 74 different S-allele combinations present in cherries from Azerbaijan and 39 in cherries from Turkey. Only seven S-allele combinations (7%) were found in plants from both countries, while the other genotypes were found in only one country.

3.6 High S-allele diversity in sweet cherries collected in the Caucasus

To study the S-allele diversity of the collected material, the number of distinct alleles (allelic richness) and the number of private alleles (private allelic richness) were calculated and normalized to a standardized sample size (Figure 4). The cherries from Turkey exhibited a higher allelic richness compared to cherries from Azerbaijan (Figure 4A). This was also the case for the number of private alleles (Figure 4B). Comparing the allelic richness of the distinct collection sites (Figure 4A), a high degree of variability was observed, with sites B and F (both in Turkey) showing the highest allelic richness.

The genetic diversity of the S-locus was also analyzed for cultivated sweet cherries of different countries. Therefore, data previously published and regularly updated by (Schuster, 2020) were used. This data set contains the S-genotypes of 2,966 sweet cherry cultivars including information to the respective country of origin. The sweet cherries collected in the Caucasus exhibited a higher allelic richness (Figure 4D) compared to the sweet cherry cultivars listed by (Schuster, 2020). This was also the case for the private allelic richness (Figure 4E). With increasing sample size g, the number of private alleles is steadily increasing calculated for the
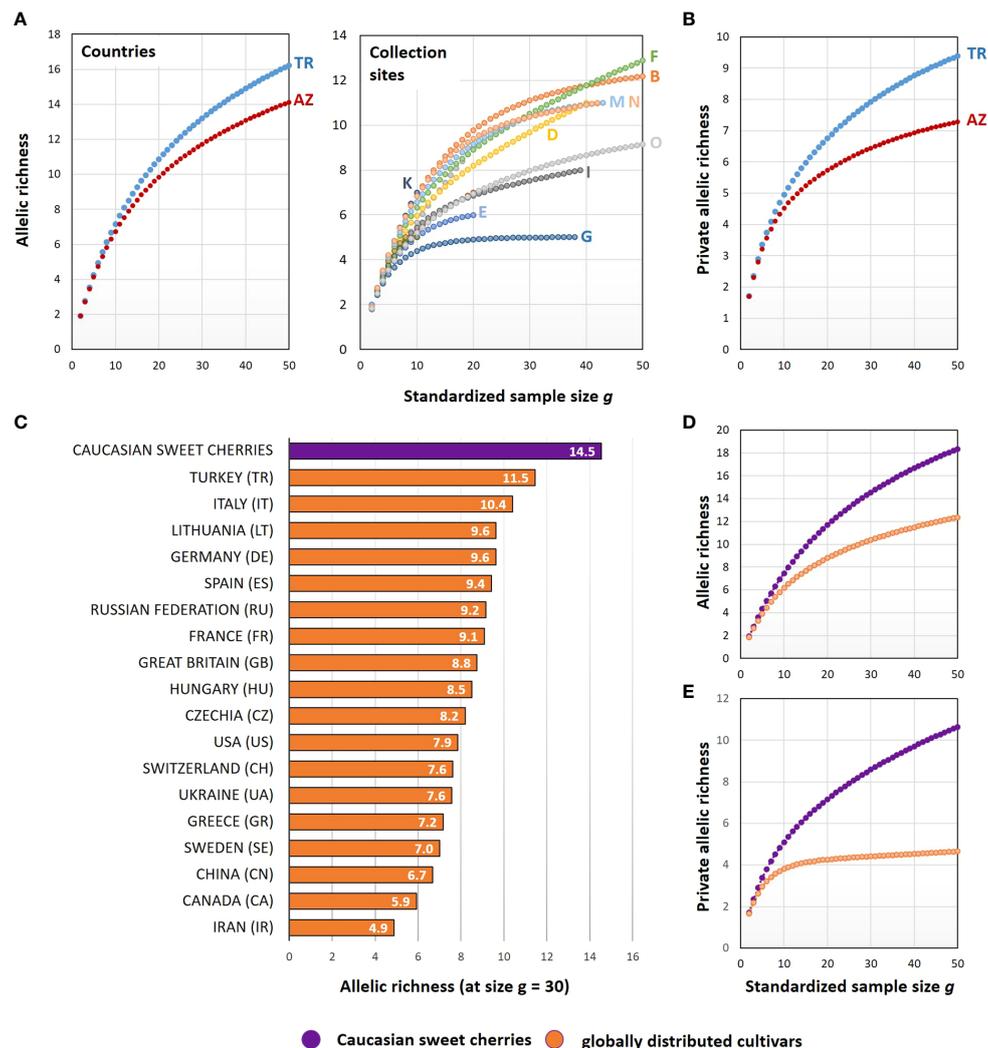
FIGURE 4

Comparison of the *S*-allele diversity of sweet cherries collected in the Caucasus and sweet cherry cultivars distributed worldwide. Allelic richness and the private allelic richness were calculated for the sweet cherries collected in Turkey and Azerbaijan and for sweet cherry cultivars originating from different countries. For this comparison, only countries with at least 30 cultivars for which the *S*-alleles are known were considered. The information on the *S*-alleles of these cultivars comes from (Schuster, 2020) **(A, B)** The grouping of the sweet cherries collected in the Caucasus was done for different countries (TR: Turkey; AZ: Azerbaijan) and collection sites (A−O, see Figure 1). **(C)** The allelic richness of the *S*-locus is presented for the sweet cherries collected in the Caucasus and for sweet cherry cultivars distributed worldwide, grouped by their country of origin. **(D, E)** Comparison of all cultivated sweet cherries and the sweet cherries collected in the Caucasus.

sweet cherries from the Caucasus, reaching a value of about 11 private alleles at sample size g=50. In contrast, the number of private alleles stagnates for the group of sweet cherry cultivars and reaches only a value of about five.

A more detailed view to the genetic diversity of the *S*-locus of cultivated sweet cherries grouped to their country of origin and compared to the Caucasian sweet cherries is represented in Figure 4C. The highest allelic richness was found in the Caucasian cherries. They showed a clear distance to the remaining sweet cherry cultivars. The Caucasian cherries were followed by cultivars from Turkey, Italy, Lithuania, Germany and Spain. The lowest allelic richness was found in cultivars from Iran, followed by Canada and China.

In summary, a high degree of allelic richness of the *S*-locus was observed for the plants originating from seeds of wild sweet cherries

and local sweet cherry cultivars, which were collected in only a limited area in the Caucasian center of origin. In this regard, the comparison with the large group of worldwide-distributed sweet cherry cultivars reflected the reduction of genetic diversity of this fruit crop by selective breeding.

4 Discussion

4.1 Reference genotypes and amplicon-deep sequencing for *S*-allele genotyping

Determining *S*-alleles in sweet cherries is important for plant breeding and fruit production, as it allows for the selection of compatible combinations of *S*-genotypes successful pollination. The

development of lab methods for predicting the compatibility of genotypes, which eliminates the need for elaborate test crosses, gained significant importance about 30 years ago. Initially, biochemical methods were developed using protein extracts obtained from style tissue to determine S-alleles based on the S-RNase polymorphism (Bošković and Tobutt, 1996; Bošković et al., 1997). A major simplification of S-genotyping was achieved by using PCR-based methods relying on the amplification of DNA length polymorphisms within the S-locus and the development of allele-specific primer combinations (Sonneveld et al., 2001, Sonneveld et al., 2003, Sonneveld et al., 2006).

Based on the results obtained in these initial studies, a consensus primer pair PaConsI-F/PaConsI-R2 was developed (Sonneveld et al., 2006) that binds specifically to conserved regions flanking the first intron of the S-RNase gene. This primer pair has since been used as a standard for the determination of S-alleles. Due to the large literature reference database developed with the PaConsI-F/R2 primer combination, this marker was also used in this work. However, the exact assignment of S-alleles remained difficult due to variations in analysis parameters such as different methods for separating amplicons (e.g. agarose gel versus capillary gel electrophoresis), types of devices for capillary gel electrophoresis (e.g. Beckman Coulter CEQ 8000, Thermofisher 3500xL Genetic Analyzer), and fluorescent dyes used for labeling primers. These differences can influence measured values in fragment length determination and cause system-related size shifts. This can also be proven by the data measured in this work. A size shift of 5 to 6 bp between the two different capillary electrophoresis devices used could be detected (Table 1, Supplementary Table S1). Further, the steadily increasing number of identified S-alleles, some of which differ in fragment length by only one or a few bases, exacerbates this problem. Examples are the S_2, S_7 and S_{12} alleles (Sonneveld et al., 2006; Vaughan et al., 2006; Cachi et al., 2017). The use of reference genotypes with known S-alleles is essential for the exact assignment of amplicons of almost the same size and the normalization of system-related size shifts.

In this study, one to five reference genotypes were analyzed for each of the 16 previously described S-alleles S_1, S_2, S_3, $S_{3'}$, S_4, $S_{4'}$, S_5, S_6, S_7, S_9, S_{12}, S_{13}, S_{14}, S_{19}, S_{21}, and S_{22} to allow a reliable assignment of the amplified fragment lengths to the respective S-allele. Unfortunately, reference material for S-genotyping was not available for all the S-alleles already described. These include the $S_{5'}$, S_{17}, S_{18}, S_{29}, S_{30}, S_{31}, S_{32}, and S_{34} alleles. Some of them have only been reported in wild cherry (Vaughan et al., 2008), whereas they have not yet been detected in sweet cherry cultivars. Examples are the alleles S_{29}, S_{31}, S_{32}, and S_{34} (Schuster, 2020).

Analysis of the cherry genotypes collected in Turkey and Azerbaijan revealed a large number of S-alleles whose amplicons could not be assigned to the already known S-alleles of the available reference genotypes due to their different fragment size. To clarify whether these amplicons originated from previously unknown S-alleles, these fragments were sequenced using Illumina ultra-deep sequencing. Sequence alignments of the obtained sequences with sequences available in public databases allowed their assignment to the S-alleles $S_{5'}$, S_{29}, S_{30}, S_{31}, S_{32}, and S_{34}. In this way, the assignment of several amplicons to previously described S-alleles could be verified and confirmed (Table 1). For S_{17} and S_{18}, no reference

sequences covering the region of the first intron could be found in the database. Here, an assignment was made by comparison with literature values (Cachi et al., 2017), considering possible size shifts.

Together with the S-alleles newly described in this work, 36 different S-haplotypes were analyzed in this study. In order to facilitate future S-allele genotyping studies and to make them comparable, a reference set of genotypes was compiled. This reference set of genotypes has been planted in the JKI's trial field so that material from these plants can be provided on request for future genotyping studies. The reference set is structured in such a way that each S-haplotype is included at least once (Table 2). Further genotypes can be continuously added to this reference set to fill in missing or newly identified S-alleles, such as the recently published S_{54} allele from the variety 'Techlovicka' (Cmejlova et al., 2023).

4.2 High resolution of S-genotyping by S-allele reference ladder

To improve the comparability and assignment of amplicons to known S-alleles based on their fragment length, an S-allele reference ladder was developed (Figure 2). This reference ladder contains pooled PCR products amplified from all S-haplotypes previously distinguished by fragment length analysis. This ladder can be added as a single control sample to all future studies. This allows for rapid normalization of instrument-dependent size shifts and calibration of run-to-run differences. The ladder also allows discrimination between S-alleles with amplified fragment lengths that are very close to each other. The resolution has been shown to be very high. By using the S-allele reference ladder, even very close peaks could be clearly assigned to different S-alleles. For example, alleles S_{45} (334 bp), S_{29} (335 bp), S_{18} (337 bp), S_2 (338 bp), S_{12} (339 bp) and S_7 (341 bp), which are only one or two base pairs apart, form clearly distinguishable peaks in the S-allele reference ladder (Figure 2). Despite the high accuracy of the fragment length determination, a doubtless assignment to known S-alleles by a simple comparison with published fragment lengths is not possible due to the instrument-related size shifts of several base pairs. This is only possible by using suitable reference samples in each experiment.

However, this work also showed that fragment length alone is a weak criterion for the detection of a specific S-allele. Based on sequence comparisons of the newly described S_{46} and S_{47} alleles, amplicons of identical size are expected. However, both S-allele sequences have a low degree of sequence identity (Supplementary Figure S1) and are phylogenetically distant from each other (Supplementary Figure S2). While the Beckman system detected an identical fragment length for both S-alleles, the ABI system detected a minimal size shift between the two alleles. In case of doubt, sequencing or allele-specific PCRs should be performed.

4.3 Limitations of S-allele genotyping using the universal marker PaConsI-F/R2

For some alleles, the PaConsI-F/R2 primer pair does not allow reliable determination of the S-allele. Examples of this are the alleles

$S_{3'}$ and $S_{4'}$. Both alleles derived from their ancestral alleles S_3 and S_4 by induced mutagenesis. In this process, irradiation generated deletions at the S-locus that lie outside the region flanked by the primers PaConsI-F/R2 (Lewis and Crowe, 1954; Sonneveld et al., 2005). Consequently, PCR fragment sizes amplified from S_3 and $S_{3'}$ as well as S_4 and $S_{4'}$ with PaConsI-F/R2 are identical to each other. Thus, no distinction can be made between the ancestral and the mutated allele using PaConsI-F/R2. This is only possible by a subsequent analysis with allele-specific primer pairs. This is not the case for another mutant S-allele, called $S_{5'}$. This allele is of natural origin and occurs in Sicilian varieties such as 'Kronio' (Marchese et al., 2007). The PaConsI-F/R2 fragment amplified from the $S_{5'}$ allele differs from the ancestral S_5 allele by a deletion of one base pair. This difference can be detected by fragment length analysis and the use of the reference ladder.

Furthermore, the marker PaConsI-F/R2 is not suitable for the amplification from the S_{13} allele as no PCR product or eventually a SSR like trace peak is produced, which has been already reported (Sonneveld et al., 2006; Marchese et al., 2010; Lisek et al., 2015) and confirmed in this work. Although the primer binding sites are present in S_{13}, a short highly variable tandem repeat, presenting up to 25 (AT) repeats in *P. avium* and up to 38 in *P. cerasus* (Marchese et al., 2010) in the amplified region probably inhibits the PCR, as already discussed by Cmejlova et al. (2023). For three genotypes of the Caucasian collection, the allele S_{13} was assigned on the basis that in these only one PCR product was detected and a diploid, self-incompatible genotype was assumed. It cannot be completely excluded that other S alleles, which cannot be amplified with PaConsIF/R2, are present in these few genotypes. For confirmation of the S_{13} allele, fluorescently labelled allele-specific primers could be included in future studies, that amplify across the two SSR regions present in the first and second intron of the S_{13} S-RNase gene (Marchese et al., 2010).

The use of S-allele-specific primer pairs and other markers within the S-locus is an alternative to overcome the problems mentioned above. Recently, a new one-tube PCR assay with subsequent fragment analysis for S genotyping of *P. avium* has been developed (Cmejlova et al., 2023), based on multiplex PCR with 27 different primers, 8 of which are labeled with a fluorescent dye. This complex assay contains sequence-optimized primers for the universal amplification of the first intron of the S-RNase gene based on the PaConsI-F/R2 binding sites, a set of several S-allele specific primers, and a marker for the detection of *MGST* alleles.

4.4 Basis material with high level of diversity for breeding

In this work, nine previously undescribed S-alleles were identified in the collected genetic material from the Causcasus. Although, the collection area was limited to a small area in Turkey and Azerbaijan. It can be assumed that the genetic diversity in the area of origin of the sweet cherry is much greater and that the number of distinguishable S-alleles will increase with further investigations. This goes hand in hand with the increasing complexity in the application of S-allele genotyping. On the other hand, only a fraction of the S-allele diversity has been incorporated into varieties cultivated today. This is well demonstrated in this study by the comparison of allelic richness and private allelic richness between the analyzed material from the Caucasus and the globally distributed cultivars performed in this study (Figure 4). For the applicability of S-allele genotyping in the field of breeding and fruit production, the question arises to what extent the nomenclature of alleles by ascending numbering should also be maintained for original and wild material of the species *P. avium*. However, the newly established S-allele reference ladder in combination with amplicon deep sequencing considerably facilitate the determination of S-alleles. Both tools can now be used to determine new S-alleles and to identify preferred genotypes as part of marker-assisted selection.

Genetic diversity in cultivated sweet cherries is severely limited. Domestication has been described as a major bottleneck and breeding has further reduced diversity (Pinosio et al., 2020). For the breeding of sweet cherry cultivars with new traits adapted to changing environmental conditions, the low diversity in cultivated elite breeding material is a severe limitation. Collecting genetic resources in the region of origin of the species aims to obtain basic material characterized by a high diversity of traits of great interest for future breeding programs. In this study, it could be shown by means of S-allele genotyping that the genetic diversity in the collected material from the Caucasus is significantly higher than the diversity of globally distributed sweet cherries. It can be assumed that other traits with significance for breeding new, adapted varieties also show higher diversity. The use of these genetic resources in future breeding programs for the introduction of new traits is aimed at. Thus, the S-alleles, which have so far only been determined in basic material for breeding, should also be considered for future genotyping projects.

5 Conclusions

Based on the results of this study, the following conclusions can be drawn. The genetic diversity of the collection material from the center of origin appears to be significantly higher (at least on the basis of the S-alleles) than the diversity in the current spectrum of traditional and commercially grown sweet cherry cultivars. This seems to be due to the fact that originally only a few genotypes contributed to the development of the genepool of cultivated sweet cherries. Major progress and genetic gains through breeding can only be expected in the future if the diversity in the breeders' gene pool is expanded. This can be achieved by collecting material in the center of origin and the continuous introgression of new traits into the breeding pool as part of pre-breeding programs. In addition, the areas of origin should be safeguarded through suitable protective measures.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

SS: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. MS: Conceptualization, Resources, Writing – review & editing. JQ-G: Resources, Writing – review & editing. GL-O: Resources, Writing – review & editing. HF: Conceptualization, Funding acquisition, Writing – original draft, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1355977/full#supplementary-material

References

Azizi-Gannouni, T., Barreneche, T., Quero-Garcia, J., Campoy, J. A., Ouergui, I., Albouchi, A., et al. (2018). *S*-locus diversity of sweet cherry cultivars in Tunisia. *J. New Sci.* 58, 3738–3742.

Bošković, R., Russell, K., and Tobutt, K. R. (1997). Inheritance of stylar ribonucleases in cherry progenies, and reassignment of incompatibility alleles to two incompatibility groups. *Euphytica* 95, 221–228. doi: 10.1023/A:1002945529157

Bošković, R., and Tobutt, K. R. (1996). Correlation of stylar ribonuclease zymograms with incompatibility alleles in sweet cherry. *Euphytica* 90, 245–250. doi: 10.1007/BF00023865

Brown, S. K., Iezzoni, A. F., and Fogle, H. F. (1996). "Cherries," in *Fruit Breeding, Volume 1, Tree and Tropical Fruits*. Eds. J. Janick and J. N. Moore (John Wiley & Sons, Inc, NewYork), 213–255.

Burger, P., Terral, J., Ruas, M., Ivorra, S., and Picq, S. (2011). Assessing past agrobiodiversity of *Prunus avium* L. (Rosaceae): a morphometric approach focussed on the stones from the archaeological site Hôtel-Dieu (16th century, Tours, France). *Veget Hist Archaeobot* 20, 447–458. doi: 10.1007/s00334-011-0310-6

Cachi, A. M., Wunsch, A., Vilanova, A., Guardia, M., Ciordia, M., and Aleta, N. (2017). *S*-locus diversity and cross-compatibility of wild *Prunus avium* for timber breeding. *Plant Breed.* 136, 126–131. doi: 10.1111/pbr.12450

Cmejlova, J., Paprstein, F., Suran, P., Zeleny, L., and Cmejla, R. (2023). A new one-tube reaction assay for the universal determination of sweet cherry (*Prunus avium* L.) self-(In)Compatible *MGST*- and *S*-alleles using capillary fragment analysis. *Int. J. Mol. Sci.* 24 (8), 6931. doi: 10.3390/ijms24086931

De Nettancourt, D. (1984). "Incompatibility,". *Cell. Interactions. (Berlin Heidelberg: Springer)* 624–639. doi: 10.1007/978-3-642-69299-4_26

Dondini, L., Lugli, S., and Sansavini, S. (2018). "Cherry Breeding: Sweet Cherry (*Prunus avium* L.) and Sour Cherry (*Prunus cerasus* L.)," in *Advances in Plant Breeding Strategies: Fruits*. Eds. . Al-Khayri, S. Jain and D. Johnson (Springer, Cham).

Eeraerts, M., Smagghe, G., and Meeus, I. (2019). Pollinator diversity, floral resources and semi-natural habitat, instead of honey bees and intensive agriculture, enhance pollination service to sweet cherry. *Agriculture Ecosyst. Environ.* 284, 106586. doi: 10.1016/j.agee.2019.106586

Faust, M., and Surányi, D. (1996). "Origin and dissemination of cherry," in *Horticultural Reviews*. Ed. J Jules. (Hort Rev) 19, 263–317.

Ganopoulos, I., Aravanopoulos, F., Argiriou, A., and Tsaftaris, A. (2012). Genome and population dynamics under selection and neutrality: an example of *S*-allele diversity in wild cherry (*Prunus avium* L.). *Tree Genet. Genomes* 8, 1181–1190. doi: 10.1007/s11295-012-0504-9

Grossgejm, A. A. (1952). *Flora Kavkaza Tom V Rosaceae – Leguminosae* (Moskow/Leningrad: Isdatelstvo Akademij Nauk SSSR).

Halász, J., Makovics-Zsohár, N., Szőke, F., Ercisli, S., and Hegedűs, A. (2021a). Simple Sequence Repeat and *S*-Locus Genotyping to Assist the Genetic Characterization and Breeding of Polyploid *Prunus* Species, *P. spinosa* and *P. domestica subsp. insititia*. *Biochem. Genet.* 59, 1065–1087. doi: 10.1007/s10528-021-10090-7

Halász, J., Molnár, A. B., Ilhan, G., Ercisli, S., and Hegedűs, A. (2021b). Identification and molecular analysis of putative self-incompatibility ribonuclease alleles in an extreme polyploid species, *Prunus laurocerasus* L. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.715414

Hanke, M. V., and Flachowsky, H. (2017). "Süßkirsche (*Prunus avium*) und Sauerkirsche (*Prunus cerasus*)," in *Obstzüchtung und wissenschaftliche Grundlagen* (Berlin, Heidelberg: Springer Spektrum), 247–280. doi: 10.1007/978-3-662-54085-5

Hua, Z. H., Fields, A., and Kao, T. H. (2008). Biochemical models for *S-RNase*-based self-incompatibility. *Mol. Plant* 1, 575–585. doi: 10.1093/mp/ssn032

Iezzoni, A., Schmidt, H., and Albertini, A. (1990). Cherries (*Prunus*). *Acta Hortic.* 290, 111–176. doi: 10.17660/ActaHortic.1991.290.4

Iezzoni, A., Wünsch, A., Höfer, M., Giovannini, D., Jensen, M., Quero-Garcia, J., et al. (2017). "Biodiversity, germplasm resources and breeding methods," in *Cherries: Botany, Production and Uses*. Eds. I. A. Quero-Garcia J., J. Pulawska and G. Lang (Wallingford: CABI).

Lewis, D., and Crowe, L. K. (1954). Structure of the incompatibility gene. *Heredity* 8, 357–363. doi: 10.1038/hdy.1954.38

Lisek, A., Rozpara, E., Glowacka, A., Kucharska, D., and Zawadzka, M. (2015). Identification of S-genotypes of sweet cherry cultivars from Central and Eastern Europe. *Hortic. Sci.* 42, 13–21. doi: 10.17221/103/2014-HORTSCI

Marchese, A., Boskovic, R. I., Caruso, T., Raimondo, A., Cutuli, M., and Tobutt, K. R. (2007). A new self-compatibility haplotype in the sweet cherry 'Kronio', *S5*, attributable to a pollen-part mutation in the *SFB* gene. *J. Exp. Bot.* 58, 4347–4356. doi: 10.1093/jxb/erm322

Marchese, A., Bošković, R. I., Caruso, T., and Tobutt, K. R. (2010). Intra-allelic variation in introns of the *S13*-RNase allele distinguishes sweet, wild and sour cherries. *Tree Genet. Genomes* 6, 963–972. doi: 10.1007/s11295-010-0305-y

Marchese, A., Giovannini, D., Leone, A., Mafrica, R., Palasciano, M., Cantini, C., et al. (2017). *S*-genotype identification, genetic diversity and structure analysis of Italian sweet cherry germplasm. *Tree Genet. Genomes* 13, 93. doi: 10.1007/s11295-017-1176-2

Nicolás-Almansa, M., Ruiz, D., Salazar, J. A., Guevara, A., Cos, J., Martínez-Gómez, P., et al. (2023). Phenotypic and molecular characterization of new interspecific Japanese plum × apricot hybrids (plumcots). *Scientia Hortic.* 318, 112131. doi: 10.1016/j.scienta.2023.112131

Pinosio, S., Marroni, F., Zuccolo, A., Vitulo, N., Mariette, S., Sonnante, dG., et al. (2020). A draft genome of sweet cherry (*Prunus avium* L.) reveals genome-wide and local effects of domestication. *Plant J.* 103, 1420–1432. doi: 10.1111/tpj.14809

Schuster, M. (2017). *Self-incompatibility (S) genotypes of cultivated sweet cherries – An overview 2017* (OpenAgrar-Repositorium).

Schuster, M. (2020). *Self-incompatibility (S) genotypes of cultivated sweet cherries – An overview update 2020*. [Data set]. OpenAgrar Repository (Wiley & Sons Ltd). doi: 10.5073/20201016-141600

Sharma, K., Cachi, A. M., Sedlák, P., Skřivanová, A., and Wünsch, A. (2016). *S*-genotyping of 25 sweet cherry (*Prunus avium* L.) cultivars from the Czech Republic. *J. Hortic. Sci. Biotechnol.* 91, 117–121. doi: 10.1080/14620316.2015.1110997

Sonneveld, T., Robbins, T. P., Bošković, R., and Tobutt, K. R. (2001). Cloning of six cherry self-incompatibility alleles and development of allele-specific PCR detection. *Theor. Appl. Genet.* 102, 1046–1055. doi: 10.1007/s001220000525

Sonneveld, T., Robbins, T. P., and Tobutt, K. R. (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. *Plant Breed.* 125, 305–307. doi: 10.1111/j.1439-0523.2006.01205.x

Sonneveld, T., Tobutt, K. R., and Robbins, T. P. (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. doi: 10.1007/s00122-003-1274-4

Sonneveld, T., Tobutt, K. R., Vaughan, S. P., and Robbins, T. P. (2005). Loss of pollen-S function in two self-compatible selections of *Prunus avium* is associated with deletion/mutation of an *S Haplotype–Specific F-Box Gene*. *Plant Cell* 17, 37–51. doi: 10.1105/tpc.104.026963

Szpiech, Z. A., Jakobsson, M., and Rosenberg, N. A. (2008). ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* 24, 2498–2504. doi: 10.1093/bioinformatics/btn478

Tao, R., and Iezzoni, A. F. (2010). The S-RNase-based gametophytic self-incompatibility system in *Prunus* exhibits distinct genetic and molecular features. *Scientia Hortic.* 124, 423–433. doi: 10.1016/j.scienta.2010.01.025

Tao, R., Yamane, H., Sugiura, A., Murayama, H., Sassa, H., and Mori, H. (1999). Molecular typing of *S*-alleles through identification, characterization and cDNA cloning for S-RNases in sweet cherry. *J. Am. Soc. Hortic. Sci.* 124, 224–233. doi: 10.21273/JASHS.124.3.224

Vaughan, S. P., Boskovic, R. I., Gisbert-Climent, A., Russell, K., and Tobutt, K. R. (2008). Characterisation of novel *S*-alleles from cherry (*Prunus avium* L.). *Tree Genet. Genomes* 4, 531–541. doi: 10.1007/s11295-007-0129-6

Vaughan, S. P., Russell, K., Sargent, D. J., and Tobutt, K. R. (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. doi: 10.1007/s00122-005-0187-9

Vavilov, N. I. (1926). Centres of origin of cultivated plants. *Bull. Appl. Bot. Genet. Plant Breed* 16, 1–248.

Watkins, R. (1976). "Cherry, plum, peach, apricot and almond: *Prunus spp*. (Rosaceae)," in *Evolution of crop plants*. Ed. N. W. Simmonds (Longman Group Ltd, New York), 242–247.

Wiersma, P. A., Wu, Z., Zhou, L., Hampson, C., and Kappel, F. (2001). Identification of new self-incompatibility alleles in sweet cherry (*Prunus avium* L.) and clarification of incompatibility groups by PCR and sequencing analysis. *Theor. Appl. Genet.* 102, 700–708. doi: 10.1007/s001220051700