Article

S-Locus Genotyping in Japanese Plum by High Throughput Sequencing Using a Synthetic *S*-Loci Reference Sequence

Afif Hedhly [1,*], María Engracia Guerra [2], Jerome Grimplet [1,3] and Javier Rodrigo [1,3,*]

1   Departamento de Ciencia Vegetal, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Avda Montañana 930, 50059 Zaragoza, Spain
2   Área de Fruticultura Mediterránea, CICYTEX-Centro de Investigación 'Finca La Orden-Valdesequera', A-V, KM 372, Guadajira, 06187 Badajoz, Spain
3   Instituto Agroalimentario de Aragón-IA2, CITA-Universidad de Zaragoza, 50013 Zaragoza, Spain
*   Correspondence: ahedhli@cita-aragon.es (A.H.); jrodrigo@aragon.es (J.R.)

Abstract: Self-incompatibility in *Prunus* species is governed by a single locus consisting of two highly multi-allelic and tightly linked genes, one coding for an F-box protein—i.e., SFB in *Prunus*-controlling the pollen specificity and one coding for an *S-RNase* gene controlling the pistil specificity. Genotyping the allelic combination in a fruit tree species is an essential procedure both for cross-based breeding and for establishing pollination requirements. Gel-based PCR techniques using primer pairs designed from conserved regions and spanning polymorphic intronic regions are traditionally used for this task. However, with the great advance of massive sequencing techniques and the lowering of sequencing costs, new genotyping-by-sequencing procedures are emerging. The alignment of resequenced individuals to reference genomes, commonly used for polymorphism detection, yields little or no coverage in the *S*-locus region due to high polymorphism between different alleles within the same species, and cannot be used for this purpose. Using the available sequences of Japanese plum *S*-loci concatenated in a rosary-like structure as synthetic reference sequence, we describe a procedure to accurately genotype resequenced individuals that allowed the analysis of the *S*-genotype in 88 Japanese plum cultivars, 74 of them are reported for the first time. In addition to unraveling two new *S*-alleles from published reference genomes, we identified at least two *S*-alleles in 74 cultivars. According to their *S*-allele composition, they were assigned to 22 incompatibility groups, including nine new incompatibility groups reported here for the first time (XXVII-XXXV).

Keywords: self-incompatibility; *S*-allele; *S*-genotyping-by-sequencing; *Prunus salicina*; Japanese plum

1. Introduction

In flowering plants, Gametophytic Self-Incompatibility (GSI) is the most widespread and ancient post-pollination mechanism of non-self and self-recognition that inhibits the growth of self-pollen tubes in the style, preventing self-fertilization [1,2]. Self-incompatibility allows maintaining genetic diversity and avoiding the potentially negative effects of inbreeding [3]. It is controlled by a single genetic locus subjected to strong negative frequency selection, i.e., rare alleles have a mating advantage, [4] that would lead to an increase in the number and diversity of *S*-alleles [5]. The *S*-locus consists of two tightly linked genes that determine both pollen and pistil specificity [6]. In *Prunus* sp., as in other species of the Rosaceae, pollen allele specificity is determined by an F-box protein (SFB) that is expressed in the pollen grain [7], and pistil allele specificity is determined by a ribonuclease (S-RNase) that is expressed in styles [8,9], and inhibits the growth of self-pollen tubes [10]. The genotyping of *S*-locus in both natural and domesticated plant species has confirmed the high allelic diversity and sequence polymorphism [11,12].

In fruit trees, *S*-allele genotyping is of great practical and scientific importance, since it allows the adequate design of crosses in breeding programs and the correct choice of pollinators for self-incompatible cultivars in new orchards [13]. In Japanese plum

(*Prunus salicina* and hybrids of *P. salicina* with other *Prunus species*), most cultivars are self-incompatible and require cross-pollination to set fruit. Therefore, knowledge of the *S*-genotype of each cultivar is essential for the design of new orchards with an adequate proportion of inter-compatible cultivars and to solve problems of low fruit set related to lack of pollination [14,15]. It also allows the characterization of allelic series in natural populations to advance our understanding of gene flow and the evolution of mating barriers in plants, and their ecological and evolutionary implications [16,17].

In fruit tree species, self-incompatibility and inter-incompatibility relationships between cultivars were initially determined on the basis of controlled crosses followed by characterization of fruit set under orchard conditions [18]. Later, incompatibility was determined more precisely by microscopic observation of pollen tube growth through the style in self- and cross-pollinated flowers [19], first in the field and later in semi-in vivo cut flowers in the laboratory (e.g., [20–22]). These procedures allowed the characterization of cultivar self-(in)compatibility and the establishment of incompatibility groups by using an arbitrary nomenclature system of *S*-alleles. Towards the end of the 20th century, the identification of proteins (S-RNases and SFBs) associated with *S*-alleles allowed the identification of *S*-genotypes by biochemical methods [23,24]. Likewise, the cloning and sequencing of the genes encoded by the *S*-locus allowed obtaining the first *S*-allele sequences upon which different PCR techniques evolved (reviewed in [25,26]). In *Prunus* and other genera belonging to the Rosaceae, degenerate primers designed from conserved exonic regions within the *S-RNase* and *SFB* genes and spanning polymorphic intronic regions allowed trans-specific *S*-genotyping using only a few primer pairs [27]. In the absence of intronic length polymorphism, other PCR-based techniques were developed, such as the single strand conformation polymorphism (SSCP) technique [28]. Nowadays, the tremendous advances in whole genome sequencing are opening new avenues for less time- and resource-consuming *S*-genotyping procedures. However, reference genomes, which are currently available for many economically important fruit tree species, cannot be used to accurately align and resequence the *S*-genotype of individuals due to the extreme polymorphism that characterizes the *S*-locus. To circumvent this pitfall, several approaches have recently emerged, including methods based on bar-coded amplicons [29–31], separate alignment on several reference genomes of the same species [32], and *de novo* assemblies of either the whole genome, followed by BLAST and phylogenetic analyses [31], or the *S*-allele itself [33,34]. In apple (*Malus domestica*), de Franceschi et al. managed to *S*-genotype 63 resequenced cultivars using 32 apple *S-RNase* sequences concatenated in a string-like structure as reference sequence [33].

In this work, building on the previously described procedure used in *M. domestica* [33], we present an improved methodology to use information from the entire range of known *S*-loci (genomic regions including *S-RNases* and *SFBs*), rather than just one of the two genes that form the locus, for successful *S*-genotyping 88 individual shotgun sequences of Japanese plum cultivars. To circumvent the problem of unspecific high depth of coverage in intronic regions, which can lead to misidentification of alleles, we developed a simple but quantitative three-step validation that relies not only on the depth but also on the breadth of coverage. This methodology allowed us to accurately ascertain the *S*-genotype of most cultivars, enabling accurate *S*-genotyping of any individual (i.e., the four alleles) as long as it has an already known *S*-allele, and to identify individuals with supernumerary *S*-alleles, or with likely novel and unknown alleles.

2. Results and Discussion

2.1. An S-Locus Based Synthetic Sequence

Since the first two Japanese plum *S-RNase* alleles were cloned in 1999 from the Japanese cultivar 'Sordum' [35], a further 53 *S-RNases* and 36 *SFBs* sequences—corresponding to 39 different loci—have been published or submitted directly to the NCBI database (Table 1 and references therein) representing Chinese, Japanese, North American, Israeli, Hungarian, and Tunisian germplasms. As this is a locus involved in mate recognition, a high degree of

diversification and high number of alleles was suggested and confirmed in both wild and domesticated plant species [5,36,37]. Further genotyping of germplasm from China—where the species is believed to have originated—will undoubtedly reveal more polymorphism. To confirm the identity of all *S*-alleles and detect possible duplicates, i.e., redundant or inconsistent-, we performed phylogenetic and pairwise similarity analyses at the nucleotide and amino acid levels for all alleles (data not shown). We detected some redundancies, such as that the S_1, S_3, S_4, S_5, and S_6 alleles were actually duplicates of the S_a, S_k, S_c, S_e, and S_f alleles, respectively (Table 1). There was one case of two different *S*-alleles receiving the same name (S_{23}-RNase), and both sequences were conserved. Despite the high similarity between S_{10} and S_{32} (99.3% identity at the nucleotide level), we could not fully determine whether they are the same allele. There was only 96.0% identity at the amino acid level, differing at the 3 amino acids immediately downstream of the conserved RC4 region. In addition, both arrangements are shared with other *S*-alleles. Both alleles were also conserved to build the synthetic sequence.

To complete our search for all available information and, if possible, to retrieve full *S*-loci sequences, we performed a BLASTN search against the three available reference genomes with either short and conserved regions or full *S-RNase* and *SFB* sequences. In addition, we searched for highly similar candidate genes that fulfill a tail-to-tail orientation as in all *Prunus S*-loci, being able to retrieve three full *S*-loci. The *S-RNase* and *SFB* from cultivar 'Zhongli n°6' turned out to be 100% identical to *Sb-RNase* and *SFBb* of *P. salicina*, respectively (Figure 1A,B). All other alleles did not appear to match any sequenced allele and were therefore considered novel alleles (named S_{sany} and S_{zgl}). The alleles recovered from the 'Sanyueli' reference genome showed lower than expected trans-specific similarity within *Prunus*, whereas the *S-RNase* and *SFB* recovered from the 'ZGL' reference genome were highly similar to *S06 S-RNase* from *P. speciosa* and *SFB5* from *P. tenella*, respectively. High similarity between alleles of different species compared to those of the same species—called trans-specific polymorphism—is a common phenomenon within GSI [12,38–40]. Sequence analysis of both genes showed typical structural arrangements and size ranges as in other *P. salicina* alleles [35,41–46]. Alignment of the three *S-RNases* revealed that they have the hypothetical signal peptide sequence (and the first intron at the junction between the signal peptide and the mature protein), the three conserved regions, and the single hyper-variable region (Figure 1C) typical of ribonucleases of the T2 family [24,39]. Alignment of the three *SFB* alleles indicated that they have the single *F-box* domain, the two variable regions, and the two hypervariable regions (Figure 1D) typical of *Prunus SFB* genes [47].

To build a synthetic reference sequence, we retained the three full *S*-loci sequences, plus 1000 bp upstream and downstream the locus to increase specificity, and the longest genomic sequence of each *S*-locus gene pair and concatenated them into a rosary-like structure with 1000 bp N spacers (Figure 1E). To have a reference sequence in single strand-like structure and mimic the structure of the *S*-locus, reverse complements of *S-RNase* sequences were concatenated in tail-to-tail orientation with *SFB* alleles within each locus. In the absence of *SFB* alleles, the *S-RNases* sequences were also reverse complemented. Alignment of shotgun sequences on available reference genomes, as used for other applications, does not produce a useful alignment at the *S*-locus region due to the extreme polymorphism that characterizes this locus [33,34]. The technical advance of the method is based on the use of the entire *S*-locus allelic series of the species—i.e., all S-RNases and SFBs—to construct a synthetic reference sequence on which the resequenced genotypes are aligned. Compared to the previously described methodology based on the use of only the *S-RNase* sequence [33], our genotyping procedure is based on the parallel calling of the four *SFB* and *S-RNase* alleles that can be found in a diploid individual, and should therefore reduce allele misidentification and result in more robust genotyping.

Table 1. *Prunus salicina SFB* and *S-RNase* allele sequences deposited at GenBank. Sequences used to build the concatenated artificial sequence are indicated.

Allele	Cultivar (*S*-locus Genotype)	Nucleotide Sequence		Protein Sequence		Coding Sequence (CDS)	Reference
		ID	Length	ID	Length		
SFBa	Burmosa (S_aS_b)	AM746961.1	992 bp	CAN90151.1	331 aa	partial	[40]
SFBa	Ozarkpremier (S_aS_f)	DQ849087.1	977 bp	ABI15333.1	325 aa	partial	[45]
SFBa [1]	Sordum (S_aS_b)	AB252410.1	1131 bp	BAF42763.1	376 aa	complete	[48]
Sa-RNase [1]	Sordum (S_aS_b)	AB252411.1	1277 bp	BAF42764.1	226 aa	complete	[48]
S1-RNase	Red-beaut (S_aS_b)	AF433649.1	576 bp	AAP97311.1	95 aa	partial	[49]
SFBb	Black Golden	KJ396620.1	1131 bp	AHX39360.1	376 aa	complete	[50]
SFBb	Gaixiandali (S_bS_d)	DQ849088.1	978 bp	ABI15334.1	326 aa	partial	[45]
SFBb	Hamra Bedri	KJ396618.1	1131 bp	AHX39358.1	376 aa	complete	[50]
SFBb	Santa Rosa (S_cS_e)	KJ396607.1 [2]	1131 bp	AHX39347.1	376 aa	complete	[50]
SFBb [1]	Sordum (S_aS_b)	AB252412.1	1131 bp	BAF42765.1	376 aa	complete	[48]
Sb-RNase [1]	Sordum (S_aS_b)	AB252413.1	2332 bp	BAF42766.1	221 aa	complete	[48]
Sb-RNase	Unknown	DQ646488.1	2406 bp	ABG36934.1	159 aa	partial	[51]
SFBc	Ain Torkia	KJ396613.1	1128 bp	AHX39353.1	375 aa	complete	[50]
SFBc	Bedri1 (S_eS_h)	KJ396608.1	1128 bp	AHX39348.1	375 aa	complete	[43,50,52]
SFBc	Meiguili (S_cS_e)	DQ849084.1	1128 bp	ABI15330.1	375 aa	complete	[47]
SFBc [1]	Santa Rosa (S_cS_e)	AB280792.1	2124 bp	BAF91847.1	375 aa	complete	[53]
Sc-RNase	Oishiwasesumomo (S_cS_d)	AB084144.1	1781 bp	BAC20940.1	177 aa	partial	[41]
Sc-RNase	Santa Rosa (S_cS_e)	AB280791.1	2630 bp	BAF91846.1	230 aa	complete	[53]
Sc-RNase	Unknown	DQ646489.1	1920 bp	ABG36936.1	172 aa	partial	[51]
S4-RNase	Royal Zee (S_cS_e)	AF432418.1	1283 bp	AAP97308.1	95 aa	partial	[49]
SFBd [1]	Formosa (S_bS_d)	AM746962.1	992 bp	CAN90152.1	331 aa	partial	[40]
Sd-RNase [1]	Oishiwasesumomo (S_cS_d)	AB084145.1	1976 bp	BAC20941.1	169 aa	partial	[41]
SFBe	Aouina Hamra Bedria (S_eS?)	KJ396612.1	1131 bp	AHX39352.1	376 aa	complete	[50]
SFBe	Black Diamant/ Black Diamond (S_eS_h)	KJ396605.1	1128 bp	AHX39345.1	375 aa	complete	[43,50]
SFBe	Cidre (S_aS_e)	KJ396611.1	1128 bp	AHX39351.1	375 aa	complete	[41,50]
SFBe	Meiguili (S_cS_e)	DQ849086.1	1036 bp	ABI15332.1	345 aa	partial	[45]
SFBe	Santa Rosa (S_cS_e)	AB280794.1	1248 bp	BAF91849.1	375 aa	partial	[53]
SFBe	Stanley	KJ396606.1 [3]	1128 bp	AHX39346.1	375 aa	complete	[50]
SFBe [1]	Unknown	DQ646490.1	1749 bp	ABG36937.1	373 aa	complete	[51]
Se-RNase [1]	Santa Rosa (S_cS_e)	AB280793.1	2622 bp	BAF91848.1	239 aa	complete	[53]
S5-RNase	Royal Zee (S_cS_e)	AF433647.1	1553 bp	AAP97309.1	95 aa	partial	[49]
SFBf [1]	Huangpili (S_7S_f)	DQ849089.1	972 bp	ABI15335.1	324 aa	partial	[45]
SFBf	Janha	KJ396610.1	996 bp	AHX39350.1	331 aa	complete	[50]
SFBf	Meski Hamra	KJ396614.1	1086 bp	AHX39354.1	361 aa	complete	[50]

Table 1. *Cont.*

Allele	Cultivar (*S*-locus Genotype)	Nucleotide Sequence		Protein Sequence		Coding Sequence (CDS)	Reference
		ID	Length	ID	Length		
SFBf	Meski Kahla	KJ396617.1	1086 bp	AHX39357.1	361 aa	complete	[50]
SFBf	Unknown	DQ989578.1	1081 bp	ABM54900.1	360 aa	partial	[49]
SFBf	Zaghwenia (S_cS_f)	KJ396615.1	993 bp	AHX39355.1	330 aa	complete	[50,52]
Sf-RNase [1]	White Plum (S_fS_g)	AB084147.1	1554 bp	BAC20943.1	132 aa	partial	[41]
Sf-RNase mRNA	Huangpili (S_7S_f)	DQ512911.1	762 bp	ABF61820.1	215 aa	partial	[54]
S6-RNase	Wikson	AF433648.1	1212 bp	AAP97310.1	95 aa	partial	[49]
SFBg	Bonnie (S_gS_h)	AM746963.1	992 bp	CAN90153.1	331 aa	partial	[40]
SFBg [1]	Unknown	DQ989579.1	1084 bp	ABM54901.1	361 aa	partial	[51]
Sg-RNase [1]	Bonnie (S_gS_h)	AM746950.1	1536 bp	CAN90140.1	169 aa	partial	[40]
Sg-RNase	Honey Rosa (S_bS_g)	AB093131.1	1266 bp	BAC75456.1	79 aa	partial	[41]
SFBh	Ain Tasstouria	KJ396616.1	1125 bp	AHX39356.1	374 aa	complete	[50]
SFBh	Bedri2	KJ396619.1	1125 bp	AHX39359.1	374 aa	complete	[50]
SFBh	Jabounia Safra	KJ396609.1	1125 bp	AHX39349.1	374 aa	complete	[50]
SFBh [1]	Nvgelei (S_cS_h)	DQ849118.1	1131 bp	ABI15337.1	376 aa	complete	[45]
SFBh	Unspecified	DQ646491.1	1253 bp	ABG36938.1	374 aa	partial	[51]
Sh-RNase [1]	Kelsey (S_fS_h)	AB084148.1	1172 bp	BAC20944.1	175 aa	partial	[41]
Si-RNase [1]	Bakemonosumomo (S_bS_i)	AB084149.1	968 bp	BAC20945.1	170 aa	partial	[41]
Sj-RNase [1]	Tecumseh (S_fS_j)	AB093132.1	2346 bp	BAC75457.1	173 aa	partial	[41]
SFBk [1]	Wickson (S_fS_k)	DQ992485.1	1083 bp	ABM54902.1	361 aa	partial	[44,51]
Sk-RNase	Friar (S_hS_k)	DQ790372.1	374 bp	ABH07013.1	63 aa	partial	[55]
Sk-RNase [1]	Starkgold (S_gS_k)	AB093133.1	1035 bp	BAC75458.1	187 aa	partial	[41,56]
Sk-RNase	Wickson (S_fS_k)	EU113311.1	632 bp	ABW86860.1	149 aa	partial	[44,57]
S3-RNase	Unspecified	AF432417.1	467 bp	AAP97307.1	95 aa	partial	[58]
Sl-RNase [1]	Combination (S_gS_l)	AB093134	1533 bp	BAC75459.1	188 aa	partial	[41]
Sm-RNase [1]	Botan (S_aS_m)	AB093135.1	930 bp	BAC75460.1	171 aa	partial	[41]
Sn-RNase [1]	Superior (S_aS_n)	AB093136	843 bp	BAC75461.1	121 aa	partial	[41,56]
St-RNase [1]	Karari (S_bS_t)	AB573636.1	525 bp	BAJ13374.1	175 aa	partial	[59]
SFB7 [1]	Huangpili (S_7S_f)	DQ849085	975 bp	ABI15331.1	325 aa	partial	[45]
S7-RNase [1]	Huangpili (S_7S_f)	AY781290.1	369 bp	AAV34703.1	94 aa	partial	[60]
S7-RNase mRNA	Huangpili (S_7S_f)	DQ512912.1	761 bp	ABF61821.1	217 aa	partial	[51]
S8-RNase [1]	82-1-109	AY902455.1	450 bp	AAW83158.1	94 aa	partial	[60]
S8-RNase mRNA	82-1-109	DQ512913.1	750 bp	ABF61822.1	213 aa	partial	[51]
S9-RNase [1]	Tai Ping Guo	AY996051.1	815 bp	AAY15124.1	94 aa	partial	[60]
SFB10 [1]	Mali ($S_{10}S_{14}$)	DQ849090.1	1041 bp	ABI15336.1	347 aa	partial	[45]

Table 1. *Cont.*

| Allele	Cultivar (*S*-locus Genotype)	Nucleotide Sequence		Protein Sequence		Coding Sequence (CDS)	Reference
		ID	Length	ID	Length		
S10-RNase [1]	Hong Men Ling	DQ003310.1	880 bp	AAX98270.1	94 aa	partial	[60]
S11-RNase [1]	Zhushali	DQ512908.1	670 bp	ABF61817.1	119 aa	partial	[51]
S12-RNase [1]	Kongquedan ($S_{12}S_{13}$)	DQ512909.1	197 bp	ABF61818.1	65 aa	partial	[51]
S13-RNase [1]	Kongquedan ($S_{12}S_{13}$)	DQ512910.1	191 bp	ABF61819.1	63 aa	partial	[51]
S14-RNase [1]	Mali ($S_{10}S_{14}$)	EF177345.1	194 bp	ABM53475.1	37 aa	partial	[45]
S15-RNase [1]	Kuandiandali ($S_{15}S_{16}$)	EF177346.1	571 bp	ABM53476.1	39 aa	partial	[61]
S16-RNase [1]	Kuandiandali ($S_{15}S_{16}$)	EF177347.1	460 bp	ABM53477.1	39 aa	partial	[61]
S17-RNase [1]	Xinjiangsanhao ($S_{17}S_{19}$)	EF177348.1	261 bp	ABM53478.1	39 aa	partial	[61]
S18-RNase [1]	Danchengxingmei ($S_{18}S_{20}$)	EF177349.1	413 bp	ABM53479.1	38 aa	partial	[61]
S19-RNase [1]	Xinjiangsanhao ($S_{17}S_{19}$)	EU113259.1	690 bp	ABW88922.1	55 aa	partial	[62]
S20-RNase [1]	Danchengxingmei ($S_{18}S_{20}$)	EU113260.1	1925 bp	ABW88923.1	55 aa	partial	[62]
S21-RNase [1]	Zhenzhuli ($S_{21}S_{22}$)	EU113261.1	1601 bp	ABW88924.1	116 aa	partial	[62]
S22-RNase [1]	Zhenzhuli ($S_{21}S_{22}$)	EU113262.1	853 bp	ABW88925.1	117 aa	partial	[62]
S23-RNase [1]	Dameigui ($S_{23}S_{24}$)	EU113263.1	750 bp	ABW88926.1	54 aa	partial	[62]
S23-RNase [1]	Guangdonghongli	FJ377732.1 [4]	597 bp	ACI88749.1	64 aa	partial	[63]
S24-RNase [1]	Dameigui ($S_{23}S_{24}$)	EU113264.1	1155 bp	ABW88927.1	117 aa	partial	[62]
S25-RNase [1]	Unspecified	EU113265.1	920 bp	ABW88928.1	116 aa	partial	[62]
S26-RNase [1]	Unspecified	EU113266.1	1492 bp	ABW88929.1	117 aa	partial	[62]
S27-RNase [1]	Unspecified	EU113267.1	657 bp	ABW88930.1	54 aa	partial	[62]
S32-RNase [1]	Guali	GU574195.1 [5]	898 bp	ADD20972.1	100 aa	partial	[64]
S41-RNase [1]	Qingzhouli	GU968758.1	719 bp	ADF42685.1	121 aa	partial	[65]

[1] Used for genotyping. [2] BLASTN sequence is 100% identical to SFBb but the cultivar is likely wrong; 'Santa Rosa' was genotyped elsewhere and here as *ScSe*. [3] BLASTN sequence is 100% identical to *SFBe* but the cultivar name is reported in the literature as *Prunus domestica*. [4] A duplicate case: same name given to different alleles. [5] Highly similar to *S10-Rnases* at the nucleotide level.

2.2. A Powerful S-Allele Calling Procedure

Raw individual shotgun sequences from 88 Japanese plum cultivar samples (83 genomic, four exon targeted-capture, and one transcriptome libraries (Table S1 and references therein) were aligned to the synthetic sequence using Bowtie2, following the end-to-end default sensitive mode (Figure 2). To increase the stringency of the alignment for higher *S*-allele specificity, reads with a mapping quality lower than 20, reads shorter than 20bp in length, and reads with more than two mismatches with the reference sequence were filtered out.

To accurately genotype the *S*-locus of all samples, we followed a strategy that combined a visual inspection of the alignment with the implementation of a quantitative analysis. First, we visually inspected all alignments and established a preliminary *S*-genotype based on gene coverage. Second, we used different tools that report per-base depth and per-gene breadth of the coverage. The per-base depth was plotted as raw data (area graph) and fitted as probability density distribution. The per-gene breadth of the coverage was plotted as bar plots of the number of reads per gene and the proportion of the gene covered by at least one read base (see Figure S1 for all genotypes analyzed). Whereas each plot showed a bit of information about the likely *S*-genotype, careful examination of all figures followed by quantitative analysis was necessary to contrast the preliminary visual inspection and to call a positive *S*-allele. The probability density plot was especially informative about the actual likelihood to observe a given peak. The size of the area below the curve that a given gene covers is in fact and approximation of the probability of observing that peak considering that the whole area under the curve should sum up to one (see Table S2 for all genotypes analyzed). The proportion of the gene covered by at least one read base was especially relevant for discarding those unspecific alignments with high coverage but limited breadth, which are characteristic of intronic and intergenic regions with repetitive and transposable elements that are frequently found throughout the genome. We considered adding the rest of the genome to the synthetic sequence, but this approach significantly increased the duration of the alignment procedure with similar results to our three-steps validation procedure. Positive *S*-allele calls were based on the highest gene breadth of coverage and high likelihood in congruence with the visual inspection of the alignment.

Of the 88 shotgun samples, we were able to unambiguously genotype two *S*-alleles in 74 cultivars, only one *S*-allele in 11 cultivars, and none of the *S*-alleles in four cultivars (Table 2). The low depth and breadth of coverage in the *S*-locus regions extracted from the reference genomes (S_b, S_{sany} and S_{zgl}, see Figure S1) for most cultivars not carrying those three alleles illustrate the high degree of polymorphism at the *S*-locus, and how reference genomes cannot be used for genotyping this locus. Our methodology clearly improved the approach used in *M. domestica* based solely on per base depth of coverage, which is highly sensitive to non-specific coverage of intronic regions [33]. Although this limitation was successfully circumvented in the previous work by setting the per-base depth of coverage only for exonic regions, this might lead to misidentification of alleles in species with lower polymorphism than apple in transcribed regions. Our approach, which includes intronic regions while solving the problem of non-specific alignment, allowed us to reduce these identifications errors.
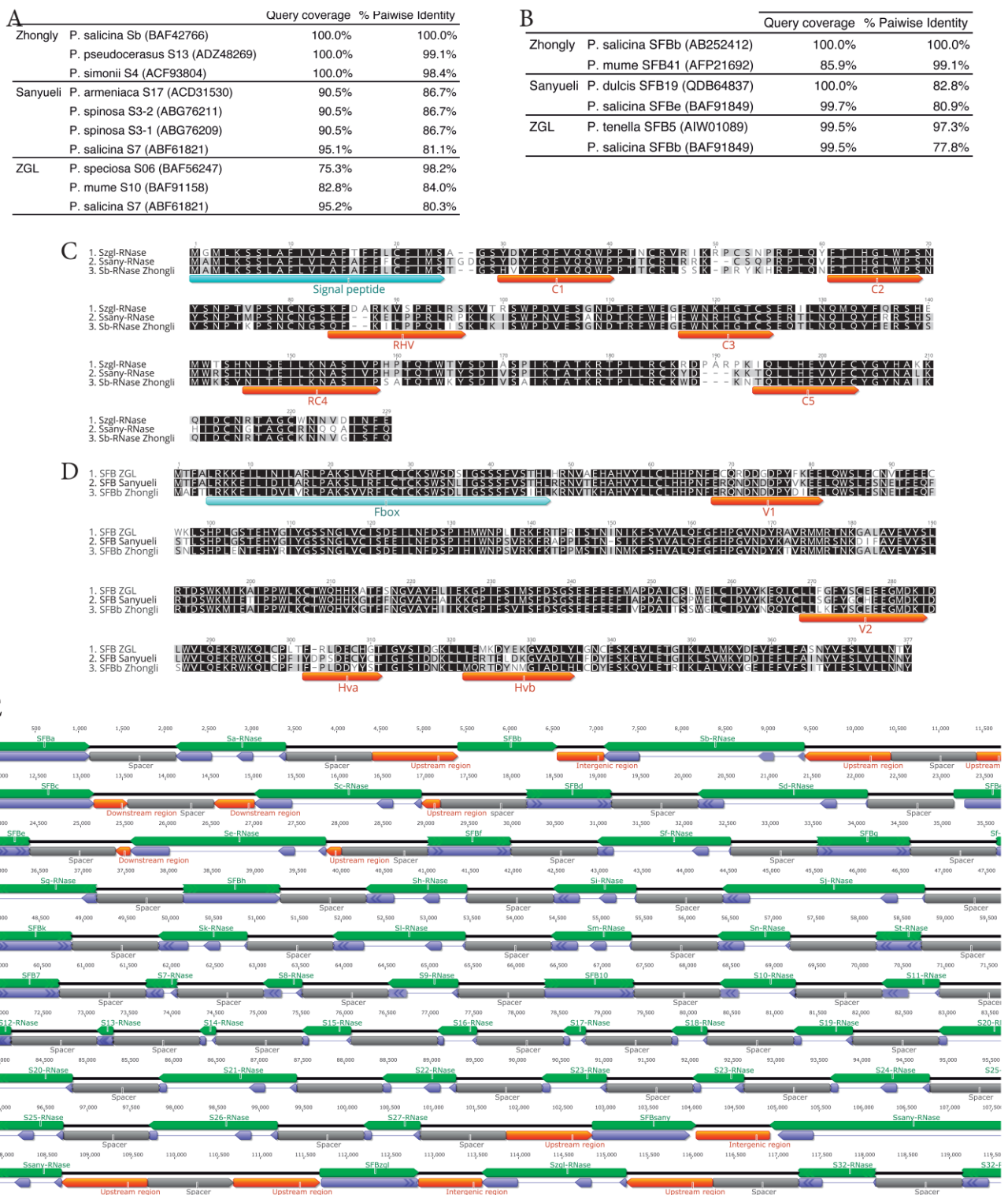
Figure 1. Full *S*-loci sequences retrieved from published reference genomes and synthetic reference sequence structure. (**A**,**B**) Pairwise trans-specific identities of the 6 retrieved *S-RNase* and *SFB* alleles. (**C**) Comparison of the amino-acids predicted from the coding regions of the three S-RNases. Signal peptide, conserved regions, and hypervariable regions (as defined by [66]) are highlighted. (**D**) Comparison of the amino acids predicted from the coding regions of the three SFBs. F-box, variable, and hypervariable regions (as defined by [47]) are highlighted. The amino acids of the S-RNases and SFBs were aligned using Muscle5 [67]. (**E**) Structure of the synthetic reference sequence used for aligning resequenced individuals. Genes (green), CDSs (blue), upstream, intergenic, and downstream

regions (orange), and 1000 bp spacers (grey) and their orientation are illustrated. Sizes of the boxes are proportional to sequence length. Accession numbers of the sequences of each allele are given in Table 1. Full *S*-loci regions for S_b, S_{sany}, and S_{zgl}, including 1000 bp upstream and downstream the locus and the intergenic regions, were retrieved from the threes published reference genomes of 'Zhongly' (Chr6:14567566:14573591), 'Sanyueli' (Chr6:33823425:33831243) and ZGL (Chr2:2766997:2772560), respectively.

All individuals with a putative S_{10} or S_{32} allele did not show an exclusive alignment to any of the two loci. Although this result was expected due to the highly similarity between both alleles at the nucleotide level, we finally decided to assign a mixed genotype $S_{10/32}$ to all individuals with this genotype. De novo assembly or full sequencing of both *S*-alleles in their respective cultivars will clarify whether they represent the same or different alleles. All unidentified *S*-alleles were assumed to be novel unsequenced alleles. Furthermore, since the *S-RNase* and *SFB* genes are known to be expressed in flower tissues during the anthesis period [7–10], we demonstrated that the *S*-genotype of the cultivar 'Sanyueli', as established from genomic reads ($S_{10/32}S_{sany}$) could also be inferred from flower transcriptome reads (Figure S1, Table S2). The lower likelihood of positive calls and shorter breadths of gene coverage can be explained by the absence of coverage within the intronic regions. In addition, we extended our analysis to test whether shotgun sequences from exome-targeted-capture libraries could be useful for *S*-genotyping. We chose four libraries from cultivars of previously known genotypes [35,41,42] and aligned them to the synthetic sequence. Although we successfully confirmed the *S*-genotype of all 4 cultivars (Table 2), lower values of likelihood and breadth of coverage should also reflect the absence of intronic coverage among others (Figure S1, Table S2). In summary, our approach represents a relevant addition to the methodologies previously described in *M. domestica* [33] and *Arabidopsis halleri* [34], as it combined the strategy of using the entire *S*-locus, the high stringency of the alignment of shotgun sequences from different types of libraries, and the quantitative analysis based on the determination of both the probability of positive calls and the percentage of gene coverage.

2.3. S-Allele Genotypes, Methodology Validation, and Incompatibility Groups

The approach developed here allowed for the identification of the *S*-genotype in 88 Japanese plum cultivars, 74 of them reported for the first time (Table 2). At least two *S*-alleles were identified in 74 cultivars, which, according to their *S*-allele composition, were assigned to 22 incompatibility groups, including nine new incompatibility groups reported for the first time (XXVII-XXXV, Table 2). In addition, 18 cultivars were included in group 0, since no other cultivars with the same *S*-genotype have been reported so far and, therefore, they can be considered as potential universal pollinators [14]. This information could be highly valuable for growers and breeders, since the knowledge of incompatibility relationships is intended to help fruit growers to select compatible pollinizers and breeders to choose parental genotypes.

To validate our methodology, we compared our results with already known *S*-genotypes of those cultivars for which *S*-locus sequences are available in the GenBank (Tables 1 and 2). The results confirmed the *S*-genotype of 13 cultivars: 'Sordum' [35], 'Oishiwase' and 'Taiyo' [41], 'Angeleno' [43], 'Fortune' and 'Friar' [68], 'Jinshali' and 'Qiuji' [45], 'Pingguoli' and 'Zuili' [46], and the two self-compatible cultivars 'Honeyrosa' [42] and 'Santa Rosa' [68]. However, the *S*-genotype of 'Beauty' (S_bS_d) and 'Friar' (S_bS_h) differed from previously reported genotypes. 'Beauty' was previously genotyped as S_cS_e [41] and 'Friar' as S_hS_k [55]. Based on a previous report on the *S*-allele inheritance in 'Friar' [68], its most likely *S*-genotype is S_bS_h. On the other hand, further research using additional plant material from 'Beauty' would be needed to clarify its *S*-genotype. Although further validation of the remaining cultivars -as more *S*-locus sequences become available- and clarification of those conflicting results would provide definitive support for the validity of this methodology,

confirmation of the *S*-genotype in 14 cultivars is a strong indication that our methodology is reliable.

In 11 cultivars, only one *S*-allele could be identified and, therefore, they were not assigned to any incompatibility group and were considered unclassified (Table 2). In four cultivars, no *S*-allele could be identified, suggesting that at least two novel *S*-alleles were detected. *S*-allele de novo assembly complemented with PCR analysis based on the amplification of *S-RNase* and *SFB* genes in these 15 unclassified cultivars would be needed to identify the new *S*-alleles and clarify their *S*-genotype. Three pairs of cultivars have a highly similar names and the same *S*-genotype ('Changlixiangjiaoli' and 'Changlixiangbian', 'Guofeng2' and 'N2 Guofeng', 'Cuihongli' and 'Cuihong') so they could be considered synonyms. On the other hand, 'Li he' and 'Lihe', with highly similar names, showed different *S*-genotypes, so they are distinct genotypes and could be considered homonyms. The presence of three and four alleles in the cultivars 'Guiyang' and 'Xingyikongxinli', respectively, means that either they have undergone some *S*-locus duplication events, or, more likely, they are triploid and tetraploid individuals, respectively.

No information on self-(in)compatibility is available for most of the cultivars tested (all except two self-compatible cultivars 'Honeyrosa' [42] and 'Santa Rosa' [68]). Controlled pollinations and subsequent microscopic analysis of pollen tube growth in self-pollinated flowers would be necessary, since the identification of the *S*-alleles of a cultivar does not accurately predict its self-(in)compatibility [14].

Japanese plum, which include pure *P. salicina* cultivars and hybrids of *P. salicina* with other *Prunus* species [14,69], is an ideal system to undertake such *S*-genotyping for several scientific and practical reasons. *P. salicina* is considered to have originated in south-western China from where it spread to other Chinese territories, then to Japan, from Japan to California, and from there to other parts of the world such as Europe, New Zealand, South America and South Africa [70–72]. The highest diversity of *S*-alleles was observed in cultivars putatively originating from the southern provinces of China (Table 2, Table S1), further supporting the putative origin of the species in this region. Furthermore, no cultivar of putative Chinese origin had the S_a and S_l alleles, suggesting a probable introgression of these alleles from other *Prunus* species through hybridization. Interestingly, the cultivars 'Changlixiangjiaoli' and 'Changlixiangbian' from Hebei province in China, genotyped here as $S_e S_k$ (Table 2 and Table S1), were claimed to belong to *P. simonii* in a previous study aimed at characterizing the origin and diversity of this species [72]. Whether this *S*-genotype reflects trans-specific polymorphism or an introgression of one or both alleles in *Japanese plum* hybrids is unknown. Undoubtedly, further studies on Chinese germplasm aimed at distinguishing pure *P. salicina* cultivars from interspecific hybrids will reveal more introgressed *S*-alleles.
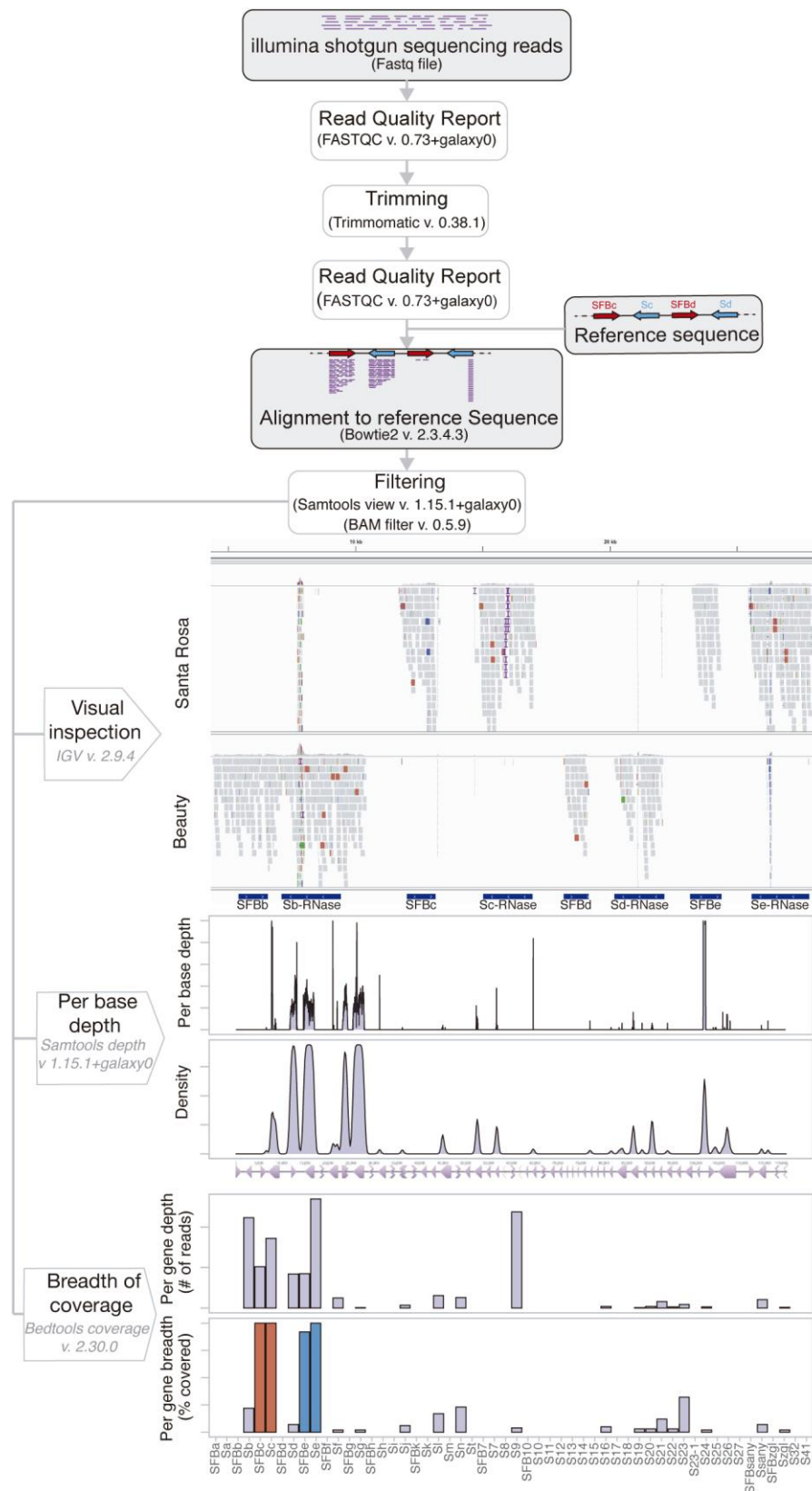
Figure 2. Bioinformatics pipeline for *S*-genotyping by high throughput sequencing. The visual inspection illustrates cultivars 'Santa Rosa' and 'Beauty; the per base depth and per gene depth of the coverage illustrate the cultivar 'Santa Rosa' (see Figure S1 for all genotypes analyzed).

Table 2. Incompatibility groups (I. G.), *S*-genotype and library type (E: exon-targeted capture, G: genomic, T: transcriptome) of 88 Japanese plum cultivars.

I. G.	S-alleles	Cultivar	Origin [1]	Library	Reference
I	S_aS_b	Sordum	Japan	E	[35]
II	S_bS_c	Fortune	California, USA	G	[68]
		Zaoshengyueguang	Japan	G	
		Dashizhongsheng	Japan	G	
		Taiyo	Japan	E	[41]
III	S_bS_f	Zhenyuanqiyuexiang	Guizhou China	G	
IV	S_bS_h	Friar	California, USA	G	[68]
		Friar [2]	California, USA	G	[55] (S_hS_k)
		Qiuji	Japan	G	[45]
		Guofeng7	Liaoning, China	G	
		C20	Spain	G	
		Zhongly nº 6	Northern China	G	
V	S_bS_i	Fujianfurongli	Fujian, China	G	
VI	S_fS_h	Satsuma	California, USA	G	
VII	S_cS_h	Fulihong	Shaanxi, China	G	
		Angeleno	California, USA	G	[43]
IX	S_fS_g	Jinshali	Yunnan, China	G	[45]
		Cehengjixueli	Guizhou, China	G	
XIV	S_aS_c	C46	Spain	G	
XIX	S_bS_d	Abazhoumeiguili	Sichuan, China	G	
		Beauty [2]	California, USA	G	[41] (S_cS_e)
XXI	S_eS_k	Wuxiangli	Hebei, China	G	
		Changlixiangjiaoli	Hebei, China	G	
		Changlixiangbian	Hebei, China	G	
XXIII	S_cS_d	Dashizaosheng	Japan	G	
		Jingshang	Japan	G	
		Oishiwase	Japan	G	[41]
		Oishiwase	Japan	E	[41]
XXIV	S_eS_{11}	Jinxiqiuli	Liaoning, China	G	
		Qiyuexiang	Liaoning, China	G	
XXV	$S_{15}S_{16}$	Pingguoli	Liaoning, China	G	[46]
XXVII [3]	S_bS_7	Heyuansanhuali	Guangdong, China	G	
		Qingpihongxin	Guangdong, China	G	
		Zaohuangli	Anhui, China	G	
		Jinganzhushali	Jiangxi, China	G	
		Jaiqingzi	Jiangsu, China	G	
XXVII [3]	S_bS_8	Huahongli	Yunnan, China	G	
		Brace	California, USA	G	
XXIX [3]	$S_eS_{10/32}$	Shuili	Guangxi, China	G	
		Dahuili	Henan, China	G	
		Xiangjiaoli	Liaoning, China	G	
		Qiuxiang	Liaoning, China	G	
XXX [3]	S_gS_k	Yuanshuai	Japan	G	
XXXI [3]	$S_hS_{10/32}$	Guofeng2	Liaoning, China	G	
		N2 Guofeng	Liaoning, China	G	
XXXII [3]	S_iS_9	Changkuangsanhuali	Guangdong, China	G	
		Sanhua plum	Guangdong, China	G	
XXXIII [3]	$S_{10/32}S_{11}$	Tiankanmali	Sichuan, China	G	
		Li He	China	G	
		Wushan plum	Chongging, China	G	
		Wanshuang plum	Chongging, China	G	
XXXIV [3]	$S_{11}S_{20}$	Yuhuangli	Shaanxi, China	G	
		Huangli	Shandong, China	G	
		Lihe	Hebei, China	G	
XXXV [3]	$S_bS_cS_f$	Guiyang	Japan	G	
Group 0	S_aS_l	Hongshou	Japan	G	
	S_bS_{11}	Yanzhili	Fujian, China	G	

Table 2. *Cont.*

I. G.	S-alleles	Cultivar	Origin [1]	Library	Reference
	S_dS_g	Fenghuali	Zhejiang, China	G	
	S_eS_{12}	Huangguli	Zhejiang, China	G	
	S_eS_{20}	Daqingke	Shandong, China	G	
	S_hS_{15}	Guofeng17	Liaoning, China	G	
	S_hS_7	Zuili	Zhejiang, China	G	[46]
	S_kS_8	Bullbank	California, USA	G	
	S_7S_{20}	Pingdingxiang	Shandong, China	G	
	S_8S_9	Wanshuhuanai	Fujian, China	G	
	$S_{10/32}S_{15}$	Zhengzhouzaoli	Henan, China	G	
	$S_{10/32}S_{sany}$	Sanyueli	Guangdong, China	G	
		Sanyueli	Guangdong, China	T	
	$S_{10/32}S_{zgl}$	Cuihongli	Sichuan, China	G	
		Cuihong plum	Sichuan, China	G	
	$S_{11}S_{16}$	Saozouli	Guangxi, China	G	
	$S_eS_kS_{10/32}S_{11}$	Xingyikongxinli	Guizhou, China	G	
	Cultivars in which only one *S* allele could be identified				
	S_b	Cuipinwannai	Fujian, China	G	
	S_c	Guoli	Liaoning, China	G	
	S_e	Tianmumili	Zhejiang, China	G	
	S_e	Hongxinli	Anhui, China	G	
	S_h	Guomei	Liaoning, China	G	
	$S_{10/32}$	Yinhong plum	Sichuan, China	G	
	$S_{10/32}$	Fengtang plum	Guizhou, China	G	
	S_{11}	Tongkeli	Guangxi, China	G	
	S_{19}	Non-specified	Tibet, China	G	
	S_{21}	Damili	Guangdong, China	G	
	S_{22}	Zhushali	Yunnan, China	G	
	Cultivars in which no *S* allele could be identified				
		Kuaishili	Guangdong, China	G	
		Lushanli	Jiangxi, China	G	
		Ni Ma Qu Ji Li	China	G	
		Wuyuecui	Sichuan, China	G	
	Self-compatible cultivars				
SC	S_bS_g	Honeyrosa	Japan	E	[42]
SC	S_cS_e	Santa Rosa	California, USA	G	[68]

[1] Presumed origin, and in its absence the sampling location is given. [2] Cultivars described with different *S*-genotype. [3] New Incompatibility Groups first reported herein.

3. Materials and Methods

3.1. Building the Reference Synthetic S-Loci Sequence

Japanese plum sequences for *S-RNase* and *SFB* were retrieved from NCBI based on the literature and on direct database search. In addition, we conducted a BLASTN search on the three available Japanese plum reference genomes of Chinese origin (cultivars 'Sanyueli' assembly V2.0 at rosaceae.org [73], 'Zhongli' assembly V1.0 at rosaceae.org [74] and 'ZGL' assembly GCA_020226455.1 at NCBI [75]) and recovered three full length *S*-locus regions. The synthetic sequence was built by concatenating individual sequences in a string-like structure. Depending on the sequences available per locus, it contained sequences from three full *S*-locus regions, individual *S*-locus sequences in tail-to-tail orientation (SFB is transcribed on the forward strand and S-RNase on the reverse), and individual *S-RNase* sequences for loci with unsequenced *SFB* alleles. BLAST searches were performed at NCBI [76] or locally on the online platform useglaxy.eu using the Galaxy BLAST+ blastn v 2.10.1+galaxy0 tool [77,78]. Sequence editing, alignment, and concatenation were performed with Geneious prime 202.2.2.

3.2. Alignment to Reference Sequence

Individual shotgun sequencing reads from Japanese plum genomic, transcriptomic and exon targeted-capture libraries were aligned to the synthetic sequence. Alignment was performed on the online platform usegalaxy.eu [78] following the pipeline outlined in Figure 2. The raw reads were first subjected to quality control (FastQC, [79]), then trimmed (Trimmomatic [80]), and finally aligned to the reference sequence using Bowtie2 [81] with the end-to-end default sensitive mode. Samtools View [82] was used to filter out reads with mapping quality below 20, and Bam Filter [83] was used to filter out reads shorter than 20 bp or having more than 2 mismatches with the reference sequence.

3.3. S-Allele Typing Methodology

To accurately type *S*-allele genotypes, we used a combination of visual inspection of the alignment using IGV [84] and quantitative analyses using SAMtools [82] and BEDtools coverage [85], followed by graphical analyses using the R package ggplot2 [86]. The likelihood of a positive call was inferred by integrating the probability density distribution of read depth at the gene level, and by determining the proportion of the gene covered by reads. Final *S*-allele call was based on full congruence between the three approaches. A positive call should have a high probability, the highest proportion of gene coverage, and should be congruent with visual inspection of the alignment.

4. Conclusions

Building on previous *S*-genotyping procedures [33,34], we developed a modified and robust methodology to *S*-genotype Japanese plum cultivars using whole genome shotgun sequences. The technical advance of the method is based on the use of the entire *S*-locus allelic series of the species to construct a synthetic reference sequence on which resequenced genotypes are aligned. *S*-genotyping is based on parallel calling of the four alleles that form the *S*-locus, combined with a simple but powerful quantitative methodology. Further work is underway to de-novo-assemble Japanese plum *S*-alleles and obtain longer sequence for many alleles for which only short sequences are available or discover novel *SFB* and *S-RNases* alleles.

S-genotyping cultivars of Chinese origin would ultimately provide valuable information on the original allelic series that can be found in pure *P. salicina* cultivars and would allow us to identify all those *S*-alleles, and their respective species, that are the result of the frequent hybridization events in modern Japanese plum hybrid cultivars. It would also enable us to build a comprehensive database of inter-compatibility relationships for all available germplasm, which would be of great value for breeders and growers. The approach developed here could be applied to any other plant species—with a GSI system—provided that several *S*-loci have been sequenced. A study is underway to apply this *S*-genotyping-by-sequencing procedure to several economically important *Prunus* fruit tree species such as *Prunus armeniaca*, *Prunus avium* and *Prunus dulcis*.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.   de Nettancourt, D. *Incompatibility in Angiosperms*; Springer: Berlin/Heidelberg, Germany, 1977.
2.   Steinbachs, J.E.; Holsinger, K.E. S-RNase-medited Gametophytic Self-Incompatibility is Ancestral in Eudicots. *Mol. Biol. Evol.* **2002**, *19*, 825–829. [CrossRef] [PubMed]
3.   de Nettancourt, D. *Incompatibility and Incongruity in Wild and Cultivated Plants*; Springer: Berlin/Heidelberg, Germany, 2001.
4.   Uyenoyama, M.K. Evolutionary Dynamics of Self-Incompatibility Alleles in Brassica. *Genetics* **2000**, *156*, 351–359. [CrossRef] [PubMed]
5.   Wright, S. The distribution of self-sterility alleles in populations. *Genetics* **1939**, *24*, 538. [CrossRef]
6.   Wu, J.; Gu, C.; Khan, M.A.; Gao, Y.; Wang, C.; Korban, S.S.; Zhang, S. Molecular Determinants and Mechanisms of Gametophytic Self-Incompatibility in Fruit Trees of Rosaceae. *CRC Crit. Rev. Plant Sci.* **2013**, *32*, 53–68. [CrossRef]
7.   Ushijima, K.; Sassa, H.; Dandekar, A.M.; Gradziel, T.M.; Tao, R.; Hirano, H. Structural and Transcriptional Analysis of the Self-Incompatibility Locus of Almond: Identification of a Pollen-Expressed F-Box Gene with Haplotype-Specific Polymorphism. *Plant Cell* **2003**, *15*, 771–781. [CrossRef] [PubMed]
8.   Tao, R.; Yamane, H.; Sassa, H.; Mori, H.; Gradziel, T.M.; Dandekar, A.M.; Sugiura, A. Identification of Stylar RNases Associated with Gametophytic Self-Incompatibility in Almond (*Prunus dulcis*). *Plant Cell Physiol.* **1997**, *38*, 304–311. [CrossRef] [PubMed]
9.   Tao, R.; Yamane, H.; Sugiura, A.; Murayama, H.; Sassa, H.; Mori, H. Molecular Typing of S-alleles through Identification, Characterization and cDNA Cloning for S-RNases in Sweet Cherry. *J. Am. Soc. Hortic. Sci.* **1999**, *124*, 224–233. [CrossRef]
10. McCubbin, A.G.; Kao, T. Molecular Recognition and Response in Pollen and Pistil Interactions. *Annu. Rev. Cell Dev. Biol.* **2000**, *16*, 333–364. [CrossRef]
11. Castric, V.; Vekemans, X. Invited Review: Plant self-incompatibility in natural populations: A critical assessment of recent theoretical and empirical advances. *Mol. Ecol.* **2004**, *13*, 2873–2889. [CrossRef]
12. Halász, J.; Molnár, A.B.; Ilhan, G.; Ercisli, S.; Hegedüs, A. Identification and Molecular Analysis of Putative Self-Incompatibility Ribonuclease Alleles in an Extreme Polyploid Species, *Prunus laurocerasus* L. *Front. Plant Sci.* **2021**, *12*, 715414. [CrossRef]
13. Herrera, S.; Lora, J.; Hormaza, J.I.; Rodrigo, J. Pollination Management in Stone Fruit Crops. In *Production Technology of Stone Fruits*; Mir, M.M., Iqbal, U., Mir, S.A., Eds.; Springer: Singapore, 2021; pp. 75–102.
14. Guerra, M.E.; Rodrigo, J. Japanese plum pollination: A review. *Sci. Hortic.* **2015**, *197*, 674–686. [CrossRef]
15. Guerra, M.E.; Guerrero, B.I.; Casadomet, C.; Rodrigo, J. Self-(in)compatibility, S-RNase allele identification, and selection of pollinizers in new Japanese plum-type cultivars. *Sci. Hortic.* **2020**, *261*, 109022. [CrossRef]
16. Busch, J.W.; Schoen, D.J. The evolution of self-incompatibility when mates are limiting. *Trends Plant Sci.* **2008**, *1s*, 128–136. [CrossRef] [PubMed]
17. Pickup, M.; Brandvain, Y.; Fraïsse, C.; Yakimowski, S.; Barton, N.H.; Dixit, T.; Lexer, C.; Cereghetti, E.; Field, D.L. Mating system variation in hybrid zones: Facilitation, barriers and asymmetries to gene flow. *New Phytol.* **2019**, *224*, 1035–1047. [CrossRef] [PubMed]
18. Williams, R.R. Techniques used in fruit-set experiments. In *Towards Regulated Cropping*; Williams, R.R., Wilson, D., Eds.; Grower Books: London, UK, 1970; pp. 57–61.
19. Herrero, M.; Dickinson, H.G. Pollen tube growth following compatible and incompatible intraspecific pollinations in *Petunia hybrida*. *Planta* **1980**, *148*, 217–221. [CrossRef] [PubMed]
20. Rodrigo, J.; Herrero, M. Evaluation of pollination as the cause of erratic fruit set in apricot 'Moniqui'. *J. Hortic. Sci.* **1996**, *71*, 801–805. [CrossRef]
21. Sanzol, J.; Herrero, M. Identification of self-incompatibility alleles in pear cultivars (*Pyrus communis* L.). *Euphytica* **2002**, *128*, 325–331. [CrossRef]
22. Guerra, M.E.; Wünsch, A.; López-Corrales, M.; Rodrigo, J. Lack of Fruit Set Caused by Ovule Degeneration in Japanese plum. *J. Am. Soc. Hortic. Sci.* **2011**, *136*, 375–381. [CrossRef]
23. Bošković, R.; Tobutt, K.R. Correlation of Stylar Ribonuclease Zymograms with Incompatibility Alleles in Sweet Cherry. *Euphytica* **1996**, *90*, 245–250. [CrossRef]
24. Sassa, H.; Nishio, T.; Kowyama, Y.; Hirano, H.; Koba, T.; Ikehashi, H. Self-incompatibility (S) alleles of the Rosaceae encode members of a distinct class of the T2/S ribonuclease superfamily. *Mol. Gen. Genet.* **1996**, *250*, 547–557.
25. Halász, J.; Hegedüs, A. A critical evaluation of methods used for S-genotyping: From trees to DNA level. *Int. J. Hortic. Sci.* **2006**, *12*, 19–29. [CrossRef]
26. Yamane, H.; Tao, R. Molecular Basis of Self-(in)compatibility and Current Status of S-genotyping in Rosaceous Fruit Trees. *J. Jpn. Soc. Hortic. Sci.* **2009**, *78*, 137–157. [CrossRef]
27. Sutherland, B.G.; Robbins, T.P.; Tobutt, K.R. Primers amplifying a range of *Prunus* S-alleles. *Plant Breed.* **2004**, *123*, 582–584. [CrossRef]
28. Delorme, V.; Gaude, T.; Heizmann, P.; Dumas, C. Use of Inmunochemical and SSCP analyses to test homozygosity at the S Locus of *Brassica oleracea* genotypes. *Mol. Breed.* **1995**, *1*, 237–244. [CrossRef]
29. Jorgensen, R.A. *Plant Genetics and Genomics: Crops and Models*; Springer: Berlin/Heidelberg, Germany, 2014.

30. Mable, B.K.; Hagmann, J.; Kim, S.T.; Adam, A.; Kilbride, E.; Weigel, D.; Stift, M. What causes mating system shifts in plants? *Arabidopsis lyrata as a case study. Heredity* **2017**, *118*, 52–63.

31. Mable, B.K.; Brysting, A.K.; Jørgensen, M.H.; Carbonell, A.K.Z.; Kiefer, C.; Ruiz-Duarte, P.; Lagesen, K.; Koch, M.A. Adding Complexity to Complexity: Gene Family Evolution in Polyploids. *Front. Ecol. Evol.* **2018**, *6*, 114. [CrossRef]

32. Tsuchimatsu, T.; Goubet, P.M.; Gallina, S.; Holl, A.C.; Fobis-Loisy, I.; Bergès, H.; Marande, W.; Prat, E.; Meng, D.; Long, Q.; et al. Patterns of Polymorphism at the Self-Incompatibility Locus in 1083 *Arabidopsis thaliana* Genomes. *Mol. Biol. Evol.* **2017**, *34*, 1878–1889. [CrossRef]

33. De Franceschi, P.; Bianco, L.; Cestaro, A.; Dondini, L.; Velasco, R. Characterization of 25 full-length S-RNase alleles, including flanking regions, from a pool of resequenced apple cultivars. *Plant Mol. Biol.* **2018**, *97*, 279–296. [CrossRef]

34. Genete, M.; Castric, V.; Vekemans, X. Genotyping and De Novo Discovery of Allelic Variants at the Brassicaceae Self-Incompatibility Locus from Short-Read Sequencing Data. *Mol. Biol. Evol.* **2020**, *37*, 1193–1201. [CrossRef]

35. Yamane, H.; Tao, R.; Sigiura, A. Identification an cDNA Cloning for *S*-RNases in Self-incompatible Japanese Plum (*Prunus salicina* Lindl. cv. Sordum). *Plant Biotechnol.* **1999**, *16*, 389–396. [CrossRef]

36. Lawrence, M.J. Population Genetics of the Homomorphic Self-incompatibility Polymorphisms in Flowering Plants. *Ann. Bot.* **2000**, *85*, 221–226. [CrossRef]

37. Igic, B.; Bohs, L.; Kohn, J.R. Historical inferences from the self-incompatibility locus. *New Phytol.* **2004**, *161*, 97–105. [CrossRef]

38. Richman, A.D.; Kohn, J.R. Learning from rejection: The evolutionary biology of single-locus incompatibility. *Trends Ecol. Evol.* **1996**, *11*, 497–502. [CrossRef] [PubMed]

39. Ioerger, T.R.; Clark, A.G.; Kao, T. Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9732–9735. [CrossRef]

40. Sutherland, B.G.; Tobutt, K.R.; Robbins, T.P. Trans-specific S-RNase and SFB alleles in *Prunus* self-incompatibility haplotypes. *Mol. Genet. Genom.* **2008**, *279*, 95–106. [CrossRef]

41. Beppu, K.; Yamane, H.; Yaegaki, H.; Yamaguchi, M.; Kataoka, I.; Ryutaro, T. Diversity of S-RNase genes and S-haplotypes in Japanese plum (*Prunus salicina* Lindl.). *J. Hortic. Sci. Biotechnol.* **2002**, *77*, 658–664. [CrossRef]

42. Beppu, K.; Takemoto, Y.; Yamane, H.; Yaegaki, H.; Yamaguchi, M.; Kataoka, I.; Tao, R. Determination of *S*-haplotypes of Japansese plum (*Prunus salicina* Lindl.) cultivars by PCR and cross-pollination test. *J. Hortic. Sci. Biotechnol.* **2003**, *78*, 315–318. [CrossRef]

43. Sapir, G.; Stern, R.A.; Eisikowitch, D.; Goldway, M. Cloning of four new Japanese plum *S*-alleles and determination of the compatibility between cultivars by PCR analysis. *J. Hortic. Sci. Biotechnol.* **2004**, *79*, 223–227. [CrossRef]

44. Sapir, G.; Stern, R.A.; Goldway, M.; Shafir, S. SFBs of Japanese Plum (*Prunus salicina*): Cloning Seven Alleles and Determining Their Linkage to the S-RNase Gene. *HortScience* **2007**, *42*, 1509–1512. [CrossRef]

45. Zhang, S.L.; Huang, S.X.; Kitashiba, H.; Nishio, T. Identification of S-haplotype-specific F-box gene in Japanese plum (*Prunus salicina* Lindl.). *Sex. Plant Reprod.* **2007**, *20*, 1–8. [CrossRef]

46. Zhang, S.J.; Huang, S.X.; Heng, W.; Wu, H.Q.; Wu, J.; Zhang, S.L. Identification of *S*-genotypes in 17 Chinese cultivars of Japanese plum (*Prunus salicina* Lindl.) and molecular characterisation of 13 novel *S*-alleles. *J. Hortic. Sci. Biotechnol.* **2008**, *83*, 635–640. [CrossRef]

47. Ikeda, K.; Igic, B.; Ushijima, K.; Yamane, H.; Hauck, N.R.; Nakano, R.; Sassa, H.; Iezzoni, A.F.; Kohn, J.; Tao, R. Primary structural features of the *S* haplotype-specific F-box protein, SFB, in *Prunus*. *Sex. Plant Reprod.* **2004**, *15*, 235–243. [CrossRef]

48. Tao, R.; Watari, A.; Hanada, T.; Habu, T.; Yaegaki, H.; Yamaguchi, M.; Yamane, H. Self-compatible peach (*Prunus persica*) has mutant versions of the S-haplotypes found in self-incompatible Prunus species. *Plant Mol. Biol.* **2007**, *63*, 109–123. [CrossRef] [PubMed]

49. Sapir, G.; Goldway, M. Prunus Salicina S-RNases. Nucleotide Sequences AF432418.1, AF433647.1, AF433648.1, AF433649.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

50. Abdallah, D.; Baraket, G.; Ben Tamarzizt, H.; Ben Mustapha, S.; Salhi Hannachi, A. Identification, Evolutionary Patterns and Intragenic Recombination of the Gametophytic Self Incompatibility Pollen Gene (SFB) in Tunisian Prunus Species (Rosaceae). *Plant Mol. Biol. Report.* **2016**, *34*, 339–352.

51. Sapir, G.; Goldway, M. Identification of F-box Gene in Prunus Salicina. Nucleotide Sequences AF432417.1, AF433649.1, DQ646488.1, DQ646489.1, DQ646490.1, DQ646491.1, DQ989578.1, DQ989579.1, DQ992485.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

52. Abdallah, D.; Baraket, G.; Perez, V.; Ben Mustapha, S.; Salhi-Hannachi, A.; Hormaza, J.I. Analysis of self-incompatibility and genetic diversity in diploid and hexaploid plum genotypes. *Front. Plant Sci.* **2019**, *10*, 896. [CrossRef] [PubMed]

53. Watari, A.; Hanada, T.; Yamane, H.; Esumi, T.; Tao, R.; Yaegaki, H.; Yamaguchi, M.; Beppu, K.; Kataoka, I. A Low Transcriptional Level of se-RNase in the Se-haplotype Confers Self-compatibility in Japanese Plum. *J. Am. Soc. Hortic. Sci.* **2007**, *132*, 396–406. [CrossRef]

54. Zhang, S.L.; Huang, S.X.; Zhang, S.J.; Wu, J.; Wu, H.Q.; Heng, W.; Zhang, Y.Y. Prunus Salicina S-RNase Genes. Nucleotide Sequences DQ512908.1, DQ512909.1, DQ512910.1, DQ512911.1, DQ512912.1, DQ512913.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

55. Halász, J.; Hegedus, A.; Szabó, Z.; Nyéki, J.; Pedryc, A. DNA-based S-genotyping of Japanese Plum and Pluot Cultivars to Clarify Incompatibility Relationships. *HortScience* **2007**, *42*, 46–50. [CrossRef]

56. Beppu, K.; Syogase, K.; Yamane, H.; Tao, R.; Kataoka, I. Inheritance of self-compatibility conferred by the Se-haplotype of Japanese plum and development of Se-RNase gene-specific PCR primers. *J. Hortic. Sci. Biotechnol.* **2010**, *85*, 215–218. [CrossRef]

57. Sapir, G.; Goldway, M.; Stern, R.A.; Shafir, S. S-RNase Based S-genotyping of Japanese Plum (*Prunus salicina* Lindl.) Cultivars. Nucleotide Sequence EU113311.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

58. Sapir, G.; Goldway, M. *Prunus salicina* S-RNases. Nucleotide Sequence AF432417.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

59. Beppu, K.; Kumai, M.; Yamane, H.; Tao, R.; Kataoka, I. Molecular and genetic analyses of the *S*-haplotype of the self-compatible Japanese plum (*Prunus salicina* Lindl.) 'Methley'. *J. Hortic. Sci. Biotechnol.* **2012**, *87*, 493–498. [CrossRef]

60. Huang, S.X.; Zhang, S.L.; Zhang, Y.Y.; Wu, J.; Heng, W.; Zhao, C.P. Identification of New S Genotype in Prunus salicina. Nucleotide Sequences AY781290.1, AY902455.1, AY996051.1, DQ003310.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

61. Zhang, S.L.; Zhang, S.J.; Huang, S.X.; Wu, H.Q.; Wu, J. Identification of Self-Incompatibility Genotypes of Japanese Plum. Nucleotide Sequences EF177346.1, EF177347.1, EF177348.1, EF177349.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

62. Zhang, S.L.; Zhang, S.J.; Huang, S.X.; Heng, W.; Wu, H.Q.; Wu, J. Identification of Self-Incompatibility Genotypes of Japanese plum by S-Allele-Specific PCR Analysis. Nucleotide sequences EU113259.1, EU113260.1, EU113261.1, EU113262.1, EU113263.1, EU113264.1, EU113265.1, EU113266.1, EU113267.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

63. Wuyun, T.N.; Li, H.G. Isolation of Six SFB New Genes in the Prunus armeniaca (China Apricot) Nucleotide Sequence FJ377732.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

64. Wu Yun, T.N.; Li, H.G. Isolation of Six New S-RNase Genes in the *Prunus armeniaca*. Nucleotide Sequence GU574195.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

65. Na, T.; Yun, W.; Li, H.G. Isolation of Six New S-RNase Genes in the *Prunus salicina*. Nucleotide Sequence GU968758.1. Available online: https://www.ncbi.nlm.nih.gov/nuccore (accessed on 28 December 2022).

66. Ushijima, K.; Sassa, H.; Tao, R.; Yamane, H.; Dandekar, A.M.; Gradziel, T.M.; Hirano, H. Cloning and characterization of cDNAs encoding S-RNases from almond (*Prunus dulcis*): Primary structural features and sequence diversity of the S-RNases in Rosaceae. *Mol. Gen. Genet.* **1998**, *260*, 261–268. [CrossRef]

67. Edgar, R.C. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat. Commun.* **2022**, *13*, 6968. [CrossRef] [PubMed]

68. Guerra, M.E.; Rodrigo, J.; López-Corrales, M.; Wünsch, A. *S*-RNase genotyping and incompatibility group assignment by PCR and pollination experiments in Japanese plum. *Plant Breed.* **2009**, *128*, 304–311. [CrossRef]

69. Okie, W.R.; Hancock, J.F. Plums. In *Temperate Fruit Crop Breeding*; Hancock, J.F., Ed.; Springer: Dordrecht, Germany, 2008; pp. 337–357.

70. Faust, M.; Surányi, D. Origin and Dissemination of Plums. In *Horticultural Reviews*; Janick, J., Ed.; John Wiley & Sons Inc.: New York, NY, USA, 1999; pp. 179–231.

71. Okie, W.R.; Weinberger, J.H. Plums. In *Fruit Breeding, Volume I: Tree and Tropical Fruits*; Janick, J., Moore, J.N., Eds.; John Wiley and Sons, Inc.: New York, NY, USA, 1996; Volume 1, pp. 559–608.

72. Wei, X.W.; Shen, F.; Zhang, Q.; Liu, N.; Zhang, Y.; Xu, M.; Liu, S.; Zhang, Y.; Ma, X.; Liu, W. Genetic diversity analysis of Chinese plum (*Prunus salicina* L.) based on whole-genome resequencing. *Tree Genet. Genomes* **2021**, *17*, 26. [CrossRef]

73. Liu, C.; Feng, C.; Peng, W.; Hao, J.; Wang, J.; Pan, J.; He, Y. Chromosome-level draft genome of a diploid plum (*Prunus salicina*). *GigaScience* **2020**, *9*, giaa13. [CrossRef] [PubMed]

74. Huang, Z.; Shen, F.; Chen, Y.; Cao, K.; Wang, L. Chromosome-scale genome assembly and population genomics provide insights into the adaptation, domestication, and flavonoid metabolism of Chinese plum. *Plant J.* **2021**, *108*, 1174–1192. [CrossRef] [PubMed]

75. Tan, Q.; Li, S.; Zhang, Y.; Chen, M.; Wen, B.; Jiang, S.; Chen, X.; Fu, X.; Li, D.; Wu, H.; et al. Chromosome-level genome assemblies of five *Prunus* species and genome-wide association studies for key agronomic traits in peach. *Hortic. Res.* **2021**, *8*, 213. [CrossRef]

76. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]

77. Cock, P.J.A.; Chilton, J.M.; GrÃ¼ning, B.; Johnson, J.E.; Soranzo, N. NCBI BLAST+ integrated into Galaxy. *GigaScience* **2015**, *4*, 39. [CrossRef]

78. Afgan, E.; Baker, D.; Batut, B.; van den Beek, M.; Bouvier, D.; Cech, M.; Chilton, J.; Clements, D.; Coraor, N.; Grüning, B.A.; et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses. *Nucleic Acids Res.* **2018**, *48*, W395–W402.

79. Andrews, S. FastQC: A Quality Control Tool for High Througput Sequence Data. Available online: http://www.bioinformataics.babraham.ac.uk/proyects/fastqc (accessed on 28 December 2022).

80. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef]

81. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef] [PubMed]

82. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M.; et al. Twelve years of SAMtools and BCFtools. *GigaScience* **2021**, *10*, giab008. [CrossRef]

83. Breese, M.R.; Liu, Y. NGSUtils: A software suite for analyzing and manipulating next-generation sequencing datasets. *Bioinformatics* **2013**, *29*, 494–496. [CrossRef] [PubMed]

84. Robinson, J.T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E.S.; Getz, G.; Merisov, J.P. Integrative genomics viewer. *Nat. Biotech.* **2011**, *29*, 24–26. [CrossRef]

85. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [CrossRef] [PubMed]

86. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016.

87. Cao, K.; Li, Y.; Deng, C.H.; Gardiner, S.E.; Zhu, G.; Fang, W.; Chen, C.; Wang, X.; Wang, L. Comparative population genomics identified genomic regions and candidate genes associated with fruit domestication traits in peach. *Plant Biotechnol. J.* **2019**, *17*, 1954–1970. [CrossRef]

88. Fang, B.; Xu, Y.; Yu, J. The complete chloroplast genome sequence of *Prunus salicina* 'Wushan plum'. *Mitochondrial DNA B.* **2021**, *26*, 1243–1244. [CrossRef]

89. Wang, X.; Liu, S.; Zuo, H.; Zheng, W.; Zhang, S.; Huang, Y.; Pingcuo, G.; Ying, H.; Zhao, F.; Li, Y.; et al. Genomic basis of high-altitude adaptation in Tibetan Prunus fruit trees. *Curr. Biol.* **2021**, *31*, 3848–3860. [CrossRef]

90. Cao, K.; Peng, Z.; Zhao, X.; Li, Y.; Liu, K.; Arús, P.; Fang, W.; Chen, C.; Wang, X.; Wu, J.; et al. Chromosome level genome assemblies of four wild peach species provide insights into genome evolution and genetic basis of stress resistance. *BMC Biol.* **2022**, 139. [CrossRef]

91. Xu, Y.; Fang, B.; Wang, Y.; Liu, J.; Liu, C.; Yu, J. Phylogenomic analysis and development of molecular markers for the determination of twelve plum cultivars (Prunus, Rosaceae). *BMC Genomics* **2022**, *23*, 745. [CrossRef]

92. Fiol, A.; Garcí a-Gómez, B.E.; Jurado-Ruiz, F.; Alexiou, K.; Howard, W.; Aranzana, M.J. Characterization of Japanese Plum (*Prunus salicina*) PsMYB10 Alleles Reveals Structural Variation and Polymorphisms Correlating with Fruit Skin Color. *Frontiers Plant Sci.* **2021**, *12*, 655267. [CrossRef] [PubMed]

93. Fiol, A.; Jurado-Ruiz, F.; López-Girona, E.; Aranzana, M.J. An efficient CRISPR-Cas9 enrichment sequencing strategy for characterizing complex and highly duplicated genomic regions. A case study in the *Prunus salicina* LG3 MYB10 genes cluster. *Plant Methods* **2022**, *18*, 105. [CrossRef] [PubMed]

94. Fang, Z.-Z.; Wang, K.-L.; Dai, H.; Zhou, D.-R.; Jiang, C.-C.; Espley, R.V.; Deng, C.; Lin, Y.-J.; Pan, S.-L.; Ye, X.-F. The genome of low-chill Chinese plum "Sanyueli" (*Prunus salicina* Lindl.) provides insights into the regulation of the chilling requirement of flower buds. *Mol. Ecol. Resour.* **2022**, *22*, 1919–1938. [CrossRef] [PubMed]

95. Numaguchi, K.; Akagi, T.; Kitamura, Y.; Ishikawa, R.; Ishii, T. Interspecific introgression and natural selection in the evolution of Japanese apricot (*Prunus mume*). *The Plant J.* **2020**, *104*, 1551–1567. [CrossRef]