The phylogeny of *Salix* revealed by whole genome re-sequencing suggests different sex-determination systems in major groups of the genus

Sergey Gulyaev[1,#], Xin-Jie Cai[1,#], Fei-Yi Guo[2], Satoshi Kikuchi[3], Wendy L. Applequist[4], Zhi-Xiang Zhang[2], Elvira Hörandl[5], Li He[1,6,*]

[1]College of Forestry, Fujian Agriculture and Forestry University, Fuzhou, 350002, China; [2]College of Ecology and Nature Conservation, Beijing Forestry University, Beijing, 100083, China; [3]Hokkaido Research Center, Forestry and Forest Products Research Institute, Forest Research and Management Organization. Hitsujigaoka 7, Toyohira, Sapporo City, Hokkaido, 062-8516 Japan; [4]William L. Brown Center, Missouri Botanical Garden, St. Louis, Missouri 63110, USA; [5]Department of Systematics, Biodiversity and Evolution of Plants (with Herbarium), University of Goettingen, Göttingen, 37073, Germany; [6]Eastern China Conservation Centre for Wild Endangered Plant Resources, Shanghai Chenshan Botanical Garden, Shanghai 201602, China

**For correspondence. E-mail heli198724@163.com*

[#]These authors contributed equally to this paper.

- **Background and Aims** The biggest genus of Salicaceae *sensu lato* (*s.l.*) *Salix* L. has been shown to consist of two main clades: clade *Salix*, in which species have XY sex determination systems (SDSs) on chromosome 7, and clade *Vetrix* including species with ZW SDSs on chromosome 15. Here, we test the utility of whole genome re-sequencing (WGR) for phylogenomic reconstructions of willows to infer changes between different SDSs.
- **Methods** We used more than 1 TB of whole genome re-sequencing (WGR) data from 70 *Salix* taxa to ascertain SNPs on the autosomes, the sex-linked regions (SLRs), and the chloroplast genomes, for phylogenetic and species tree analyses. To avoid bias, we chose reference genomes from both groups, *Salix dunnii* from clade *Salix* and *S. purpurea* from clade *Vetrix*.
- **Key Results** Two main largely congruent groups were recovered: the paraphyletic *Salix* grade and the *Vetrix* clade. The autosome dataset trees resolved four subclades (C1-C4) in *Vetrix*. C1 and C2 comprise species from the Hengduan Mountains and adjacent areas and Eurasia, respectively. Section *Longifoliae* (C3) grouped within the *Vetrix* clade but fell into the *Salix* clade in trees based on the chloroplast dataset analysis. *Salix triandra* from Eurasia (C4) was revealed as sister to the remaining species of clade *Vetrix*. In *Salix*, polyploid group C5 is paraphyletic to clade *Vetrix* and subclade C6 is consistent with Argus's subgenus *Protitea*. Chloroplast datasets separated both *Vetrix* and *Salix* as monophyletic, and yielded C5 embedded in *Salix*. Using only diploid species, the SLR and autosomal datasets both yielded trees with *Vetrix* and *Salix* as well supported clades.
- **Conclusion** WGR data are useful for phylogenomic analyses of willows. The different sex determining systems may contribute to the isolation of the two major groups, but the reproductive barrier between them needs to be studied.

Key words: *Salix*, phylogeny, sex determination, whole genome re-sequencing, chloroplast dataset, willows.

INTRODUCTION

Salicaceae *sensu lato* (*s.l.*) includes over 50 genera, with approximately 1000 species of woody trees and shrubs (Chase *et al*., 2002; Li *et al*., 2019). S*alix* L. (willows) is the largest genus (Skvortsov, 1999; Fang *et al.*, 1999; Ohashi *et al.*, 2006), and includes ~450 species mainly distributed in the Northern Hemisphere (reviewed in Argus, 1997; Skvortsov, 1999; He *et al*., 2021a). However, species with valuable biological features have been widely introduced and cultivated in a variety of places around the world (Isebrands and Richardson, 2014).

Reproduction by separate sexes (dioecy), reduced flowers, polyploidization, and frequent natural hybridization, with often wide ranges of intraspecific phenotypic and genotypic variation, all create problems for willow classification (Argus, 1997; Skvortsov, 1999; Fang *et al*., 1999). Molecular evidence in the first decade of the 21st century proved *Salix* to be a monophyletic group, and significant progress in genus delimitation has been achieved (Supplementary data Table S1).

Subgeneric classification of *Salix* has, however, remained in a process of endless revision (Supplementary data Table S1). Skvortsov (1968, translated into English in Skvortsov 1999), recognized three subgenera in Eurasia, *Salix*, *Chamaetia* and *Vetrix.* However, he made no taxonomic decision on species from other continents. He admitted that the separation between the subgenera *Chamaetia* and *Vetrix* was not clear. These subgenera are more closely related to each other than to *Salix s.l.* (sensu Skvortsov's), which shows "primitive" morphological features (Barkalov and Kozyrenko, 2014; Cronk *et al*., 2015; Wu *et al.*, 2015) typically found in *Populus*. Dorn (1976), based on morphological characteristics of American willows, accepted only the subgenera *Salix* and *Vetrix.* Argus (1997) conducted a morphological cladistic analysis and classified the North American species into four

subgenera: *Salix*, *Longifoliae*, *Chamaetia,* and *Vetrix*. Later on, Argus (2010) accepted five subgenera in Flora of North America: *Protitea*, *Longifoliae*, *Salix*, *Chamaetia*, and *Vetrix*.

Hardig *et al.* (2010) used *matK* chloroplast markers and ribosomal DNA ITS sequences, generally supporting Argus's subgenera. However, Chen *et al.* (2010) revealed the two subgenera *Chamaetia* and *Vetrix* to be monophyletic based on three plastid markers. Abdollahzadeh *et al.* (2011) considered that all of these subgenera were non-monophyletic except *Longifoliae*. Using ETS and ITS sequences of nuclear ribosomal DNA and four plastid markers, Wu *et al.* (2015) supported the merging of the subgenera *Chamaetia* and *Vetrix*, with sect. *Amygdalinae* as sister to it. Lauron-Moreau *et al.* (2015a) proposed to subdivide *Salix* into two subgenera (*Salix* and *Vetrix*), but in a corrected version (Lauron-Moreau *et al.*, 2015b) four clades were recognised, basically consistent with subgenera *Protitea*, *Salix*, *Longifoliae*, and *Vetrix* (including *Chamaetia*), respectively. Other molecular studies around the same time mostly recognized two major clades within *Salix*: one which is composed of species from subgenera *Salix, Longifoliae* and *Protitea*, and the other including species of subgenera *Chamaetia* and *Vetrix*, along with sect. *Amygdalinae*, and representatives of the formerly recognized segregate genera *Chosenia* (*Salix arbutifolia*) and *Toisusu* (*Salix cardiophylla*) (Wu *et al.*, 2015; Zhang *et al.*, 2018b). Restriction-site associated DNA (RAD) sequencing has been applied to estimate the phylogeny of willows, whereby particularly Wagner *et al.* (2018; 2020; 2021a) confirmed the monophyly of *Chamaetia* and *Vetrix*, suggesting they be treated as the *Chamaetia/Vetrix* clade, but excluding sect. *Amygdalinae*. He *et al.* (2021a) discussed adaptive evolution patterns of some *Chamaetia/Vetrix* species and their radiation in the Hengduan Mountains, showing subdivision within the clade into the Hengduan and Eurasian subclades. Recently, chloroplast genomes of 32 species confirmed the monophyly of three well-supported clades that are each separated on long branches:

Chamaetia/Vetrix, subg. *Salix*, and in between the "*Amygdalinae*" clade with *Salix triandra* (Wagner *et al.*, 2021b).

The emergence and development of high-throughput sequencing technologies has provided new approaches (Morey *et al.,* 2013). In particular, it has become possible to sequence large genomes at low cost with relative confidence in the data quality (Shendure *et al.,* 2017), and whole genome sequencing (WGS) has become widely used (Unamba *et al.*, 2015), and further extended to whole genome resequencing (WGR) (Shendure *et al.*, 2017). This method can now be used for phylogenomic studies (e.g., Malmstrøm *et al*., 2017; Ma *et al.*, 2018; Olofsson *et al.*, 2019).

Compared to the nuclear genome, the chloroplast genome is considerably smaller, varying among plants in size and coding genes (Daniell *et al.*, 2016). Moreover, the substitution rate of the chloroplast genome is low, which makes it potentially a useful resource for molecular phylogenetic studies (Raubeson and Jansen, 2005). The disadvantage is the low resolution of plastid phylogenies and the predominant maternal inheritance of chloroplast genome. Phylogenetic incongruence with nuclear data may appear due to maternal or biparental mode of transmission (Stull *et al.*, 2020). Offspring of interspecific hybridization inherit nuclear genes from both parents. However, as shown by Zhang and Liu (2003), since *Salix* has maternal inheritance, the chloroplast genome is expected to remain largely identical with that of the female parent. Several studies applied chloroplast genomes to willows species for resolving some phylogenetic issues. Zhang *et al*. (2018b) used whole chloroplast genome sequences of 42 members of Salicaceae *s.l.* mainly for divergence time estimation; nevertheless, two major clades within *Salix*, mainly comprised of species of *Chamaetia/Vetrix* and *Salix*, respectively, were recognized. Li *et al*. (2019) used chloroplast genomes of *Salix interior* along with chloroplast genomes of 23 species from Salicaceae *s.l.* to reconstruct the intrageneric relationships of the family. The study by Wagner *et al.* (2021b)

confirmed that chloroplast genomes can separate large clades, but chloroplast genome evolution at the species level is shaped by low divergence, reticulate evolution, and homoplasy. Thus, despite low resolution and low support values within clades, chloroplast genome data can be successfully used for investigation of the major clades of *Salix*.

Hybridization and sex determination

Salix is interesting for the study of sex determination systems (SDRs), and these may relate to the phylogenetic relationships in the following way, making it important to clarify relationships among willows. *Populus* and *Salix* are sister genera in Salicaceae, and both are dioecious. It has therefore been suggested that this state was present in the common ancestor of these genera (Dai *et al.*, 2014) before they diverged from each other about 40-45 Mya (Boucher *et al.*, 2003; Wu *et al.*, 2015). Despite the long period during which dioecy could have existed, the chromosomes carrying the sex determination region (SDR) are homomorphic (van Buijtenen and Einspahr, 1959). In both genera the SDR systems are based on a single factor located either on the female (W) chromosome, or the male (Y) chromosome (Renner and Müller, 2021). This factor is either female-specific expressed on the W chromosome or dominantly repressed by the male Y chromosome. Changes in the position of the SDR region can happen relatively easily by translocation, female heterogamety, or new mutation (Renner and Müller, 2021). It is therefore possible that the SDR has evolved to some extent independently in these two genera, and that turnover events could have occurred, in which an established sex-determining system was replaced by a different one. Indeed, both male and female heterogamety are now known within the genus *Salix*. Studies of sex determination systems (SDSs) were mainly based on members of subg. *Vetrix*, including *Salix polyclona*, *S. suchowensis*, *S. viminalis*, *S. purpurea*, and *S. triandra,* which have female heterogamety (ZW) with physically extensive sex-linked regions (SLRs) on

chromosome 15 in all species (reviewed in He *et al*., 2021b), Table 1). However, recent studies have demonstrated male heterogamety (XY) and SLRs on chromosome 7 in two species of subg. *Salix*, *S. dunnii* (He *et al*., 2021b) and *S. nigra* (Sanderson *et al.*, 2021). Thus, the different SDSs may support a biologically important subdivision of *Salix* that could be used as a character in subgeneric classification (although sex determination has so far been studied in only a few species of *Salix*).

Recent studies have tended to decrease the number of subgenera, predominantly agreeing on the recognition of only two major clades, the *Salix* clade consisting of subgenera *Salix, Longifoliae,* and *Protitea,* and the *Chamaetia/Vetrix* clade which includes subg. *Chamaetia* and *Vetrix* along with *Salix arbutifolia* (*Chosenia*), *S. cardiophylla* (*Toisusu*), and sect. *Amygdalinae*. Furthermore, interspecific hybridisation occurs mostly within the *Vetrix* or *Salix* clades (Wagner *et al.*, 2021a), raising the question whether the two main clades with different SDSs have reproductive barriers.

Thus, in the present study we aimed: (i) to use WGR data to reconstruct the phylogeny of sampled specimens (this approach has never been applied to the reconstruction of willow phylogeny across the whole genus), (ii) to test whether phylogenies based on autosomes, sex-linked regions and chloroplast sequences are consistent with one another, and (iii) to find out whether differences in sex determination systems between species of the two clades offer a reliable character for their subgeneric division, or not.

MATERIALS AND METHODS

Taxon sampling

We included 90 *Salix* samples (23 sections 62 species and 8 varieties) in our analysis (Supplementary data Table S2); 59 samples representing 48 taxa were newly collected from China, Japan and North America for this study. The species represent the 5 previously

recognized subgenera *Salix*, *Protitea*, *Vetrix*, *Longifoliae*, and *Chamaetia.* The plant material was frozen in liquid nitrogen and stored at -80°C until total genomic DNA was extracted or dried by silica gel. To cover more taxa and sections of the genus, resequencing genome data were included from 26 samples of He *et al.* (in preparation) and Guo *et al.* (2021), representing 23 taxa. All these taxa were identified using relevant floras and taxonomic papers (Fang *et al.,* 1999; Ohashi, 2006; Argus, 2010; He *et al.,* 2014; He *et al.*, 2015 ; Liu *et al*., 2016; He and Chen, 2017; He, 2018; Liu *et al.,* 2020; Zeng and He, 2020). We also downloaded whole genome sequencing data of five samples: *Salix brachista* (SRR7341535), *S. dunnii* (SRR12893418), *S. purpurea* (SRR3927002), *S. suchowensis* (SRR10197854), and *S. viminalis* (ERR1558612) from NCBI, which have available assembled genome(s) (reviewed in He *et al*., 2021b), and chose *Populus euphratica* (I_233/ SRR13324572) as outgroup.

Ploidy determination

The ploidy level of 49 individuals representing 42 taxa was measured by flow cytometry (FCM). *Salix polyclona* (2x = 2n = 38, He *et al*. in preparation) was used as an external standard for ploidy determination. The FCM protocol of Doležel *et al*. (2007) was used. About 20–50 mg of silica gel-dried leaf tissue was incubated for 80 min in 1 ml LB01 buffer, and then chopped with a razor blade. The cell culture was then collected by gentle pipetting and filtered through a 38 µm nylon mesh. Before analysis, the samples were stained with 80µg ml−1 PI simultaneously with 80µg ml−1 RNase in an ice bath for 30 min. A total of about 5000 nuclei were measured for each sample.

Ploidy levels were estimated using a MoFlo-XDP flow cytometer (Beckman Coulter, Inc., Indianapolis, United States), and FloMax V2.0 (Sysmex Partec GmbH, Münster, Germany) was used to evaluate the histograms for each sample. The ploidy level was

obtained based on the following equation: Sample ploidy (integer) = reference ploidy * mean position of the G1 sample peak / mean position of the G1 reference peak. A maximum coefficients of variation (CV) value of 8 was accepted for each sample peak (G0/G1 peak), to control the quality of ploidy level measurements.

Sequencing, reads mapping, and variant calling

Total genomic DNA for 59 samples was extracted from fresh leaves frozen in liquid nitrogen and stored at -80°C or silica gel-dried leaves using the Qiagen DNeasy Plant Mini Kit (Qiagen, Valencia, CA) following the manufacturer's instructions. Whole-genome re-sequencing using paired-end libraries was performed on Illumina NovaSeq 6000 with 150-bp read length on each end by NovoGene (Beijing, China) and Majorbio (Shanghai, China). The high-quality genome assemblies of *Salix dunnii* (female, including chr07X, clade *Salix*) (He *et al*., 2021b) and *S. purpurea* v5.1 (female, including phased Chr15Z and Chr15W, clade *Vetrix*) (Zhou *et al.,* 2020), and chloroplast genomes of *S. dunnii* (He *et al.,* 2021b) and *S. purpurea* (GenBank: KP019639.1) were used as reference genomes for all 90 samples of *Salix* and outgroup in reads mapping and variant calling.

The sequenced reads of all samples were filtered and trimmed using fastp, and reads with length < 60 bp were discarded (Chen *et al*., 2018). Then clean reads were aligned to the genomes and chloroplast genome sequences of *Salix dunnii* and *S. purpurea* using the BWA-MEM algorithm from BWA 0.7.12 (Li and Durbin, 2009; Li, 2013). SAMTOOLS 0.1.19 (Li *et al*., 2009) was used to extract primary alignments, sort, and merge the mapped data. We used Sambamba 0.7.1 (Tarasov *et al*., 2015) to mark potential duplications from the PCR amplification step of library preparation for all bam files.

We called variants for all the bam files using GATK's 'HaplotypeCaller' and 'GenotypeGVCFs' (Genome Analysis Toolkit v. 4.1.8.1). For each bam file,

'HaplotypeCaller' for chromosome regions were run with '--sample-ploidy 2', and for chloroplast regions with '--sample-ploidy 1'. Genomic VCFs (GVCFs) of chromosome and chloroplast regions were obtained for each sample. Before joint genotyping with 'GenotypeGVCFs', 'GenomicsDBImport' and 'CombineGVCFs' were used to merge the GVCFs of chromosome regions and chloroplast regions from all samples, respectively. Hard filtering of the SNPs was carried out using the best practice quality recommendations of the GATK group (QD < 2.0, FS > 60.0, MQ < 40.0, MQRankSum < -12.5, ReadPosRankSum < -8.0, SOR > 3.0). Biallelic sites were extracted for subsequent filtering for the chromosome regions. For the sites in the chloroplast regions, we kept all polymorphisms. We then excluded sites with extremely high coverage across all samples (twice the average coverage) and treated the sample-level genotype depth '< 4' as no call, and included sites with at most 10% of no-call genotypes in all sample (for sites in chloroplast sequences, we allowed 50% max missing). We also removed sites with a minor allele frequency less than 0.05. This process yielded four high-quality SNP datasets: 416 SNPs using the chloroplast genome of *Salix dunnii* as reference (CP-dun), 3,036,086 SNPs using the *S. dunnii* nuclear genome as the reference, 988 SNPs using the chloroplast genome of *S. purpurea* as reference (CP-pur), and 3,350,756 with the *S. purpurea* nuclear genome reference.

Phylogenetic analysis

Since different SDSs were identified in the two main diploid clades, we extracted the SNPs in autosomal regions from both nuclear datasets (excluding the entire chromosomes 7 and 15). A python script (https://github.com/zhangrengang/degeneracy) was used to generate all four-fold degenerate sites in the genomes of *Salix dunnii* and *S. purpurea*, based on their gene annotation (GFF3 files, Zhou *et al.*, 2020, He *et al.*, 2021b). We extracted the SNPs at these four-fold degenerate (FF-D) sites from the two nuclear autosomal datasets. Polyploidy

can confuse phylogenetic analysis, especially if interspecific hybridisation was involved into duplication of the genome, which can lead to incongruent reconstructions as well as reticulate evolutionary patterns (Alix *et al*., 2017). We therefore obtained 233,684 FF-D SNPs for all individuals, and 207,155 FF-D SNPs for diploids, from the autosome regions AR-dun dataset; similarly, we obtained 258,908 FF-D SNPs for all individuals and 230,084 for diploids from the AR-pur dataset.

Phylogenetic relationships were inferred by a maximum likelihood approach using RAXML v.8.2.4 (Stamatakis, 2014) based on the concatenated sequences of six datasets: AR-dun (all individuals), AR-Di-dun (diploids only), CP-dun, AR-pur (all individuals), AR-Di-pur (diploids only), and CP-pur. Support values were calculated using 100 rapid bootstrap replicates (-f a option) based on the GTR+GAMMA nucleotide substitution model as Wagner *et al*. (2020) and He *et al*. (2021a) used for the phylogenomics of willows.

The sex determination systems are not known for any polyploid willows. We therefore extracted SNPs from the sex-linked regions only for the diploids. For the dataset using the *S. dunnii* reference, this region is X-LR, chr07:5675000–8880000 of chromosome 7, and for that using the *S. purpurea* reference it is the Z-linked region of chromosome 15 (Z-LR, Chr15Z: 2341099–6715814) (Zhou *et al*., 2020; He *et al*., 2021b). Because recombination suppression of sex-linked regions may lead to gene duplications, we used only single-copy genes (SCGs) in the chromosome 7X and 15Z regions, identified using OrthoFinder (Emms and Kelly, 2019). Because of assembly quality and availability of phased sex-linked regions, we used only the genomes of *S. brachista*, *S. dunnii*, and *S. purpurea* to identify the chromosome 7X SCGs, and only *S. dunnii* and *S. purpurea* (excluding Chr15W) to identify the chromosome 15Z SCGs. Then we used the two SCG datasets to extract the SNPs in the X-LR and Z-LR. This yielded 1,165 SNPs in 32 SCGs in X-LR, and 663 SNPs in 15 of 74 SCGs in Z-LR. We developed a custom script to obtain the sequences of each SCG in these

datasets, and used ModelFinder (Kalyaanamoorthy *et al.*, 2017) to select the best model under the Bayesian information criterion and IQ-tree based on best model (Minh *et al.,* 2020) to construct individual gene trees. Species trees of the two datasets were estimated using ASTRAL (Zhang *et al.,* 2018a).

RESULTS

Ploidy determination

Among the 42 measured taxa, the ploidy levels of 17 species and 5 varieties were measured for the first time (Supplementary data Table S2, Figure S1). *Salix annulifera*, *S. baileyi*, *S. caroliniana*, *S. cheilophila*, *S. fargesii* var. *hypotricha*, *S. fargesii* var. *kansuensis*, *S. hypoleuca* var. *kansuensis*, *S. hypoleuca* var. *platyphylla*, *S. luctuosa*, *S. permollis*, *S. spathulifolia*, and *S. tangii* were revealed to be diploid. *Salix austrotibetica*, *S. balansaei*, *S. chienii*, *S. paraplesia*, *S. sclerophylloides*, *S. spathulifolia* var. *glabra*, and *S. staintoniana* were found to be tetraploid. *Salix taipaiensis, S. wangiana*, and *S. yuhuangshanensis* are hexaploid. Of the other 20 taxa, 17 had ploidy levels congruent with previous reports (Supplementary data Table S2). However, the ploidy levels of tetraploid *Salix daltoniana*, diploid *S. cf. scouleriana*, and hexaploid *S. sinica* differed from those previously reported (*S. daltoniana* diploid (Fang *et al*., 1999); *S. scouleriana* tetraploid (Argus, 2010); and *S. sinica* tetraploid, (He *et al*., in preparation)).

Whole genome re-sequencing

After filtering, we obtained more than 7 (mean 11) Gb of clean reads per new sequenced sample (Supplementary data Table S3), and on average 11.4 Gb of all sample. The average depths of the new sequences ranged from 14.5× to 44.1× and 11.9× to 44.2×, using *Salix dunnii* and *S. purpurea* as reference genomes, respectively, while all the 91 samples had

average depths of 14.5× to 62.9× and 11.9× to 63.5×, respectively (Supplementary data Table S4 and S5). Based on more than 1 TB (0.65 TB new sequenced) clean reads, we obtained 416 chloroplast and 3,036,086 nuclear high-quality SNPs using *S. dunnii* as reference, 988 and 3,350,756 high-quality SNPs with the *S. purpurea* as reference.

Phylogenetic trees based on nuclear and chloroplast data

As we hypothesised that the species of the two main clades of *Salix* have different sex determination systems (see Introduction), we chose *Salix dunnii* and *S. purpurea* from the two clades as reference genomes in order to analyse sex-linked regions separately from autosomal ones, and used a total of eight datasets to reconstruct phylogenies and species trees of *Salix* species (see Methods, Table 2, Figures 1-4, and Supplementary data Figs S2-5).

Five major clades and one paraphyletic group were found by analyses of putatively autosomal sequences, using both the AR-dun and AR-pur datasets, which include both diploid and polyploid species (Fig. 1, Supplementary data Fig S2). Clade *Vetrix* includes four subclades: C1, C2, C3, and C4. *Salix* includes one subclade (C6) and a small polyploid group, C5, consisting of *Salix shihtsuanensis, S. chienii* and *S. matsudana*, plus three branches with *S. cf. fragilis, S. pentandra, S. paraplesia*, and *S. lucida*, paraphyletic to the *Vetrix* clade. Individuals of the same species are well supported as monophyletic in the two trees, except a few species or complex with multiple ploidy levels (*S. polyclona*, *S. shihtsuanensis*), allopolyploid origin (*S. opsimantha*), or unclear morphological boundaries (*S. luctuosa* and *S. hypoleuca*; *S. fargesii* and *S. moupinensis*).

Compared to the complete autosome datasets, the data from just the diploid species (AR-Di-dun, AR-Di-pur, and sex-linked region datasets (SCG-X-LR and SCG-Z-LR)) are expected to be less affected by the problems outlined above for polyploid species, such as paraphyly and reticulate evolution. All analyses of the diploid data support separate

monophyletic groups of *Vetrix* and *Salix* species (Figs. 2-3, Supplementary Data Fig. S3-4). Due to lower numbers of SNPs, the SCG-X-LR and SCG-Z-LR species trees did not distinguish C1 and C2, but revealed C3 as sister to C1 and C2 group. The AR-Di-dun and AR-Di-pur analyses revealed a similar topology of C1, C2, C3, C4 of clade *Vetrix* as the whole autosome datasets.

Both trees based on chloroplast genomes, CP-dun and CP-pur again include two major well-supported clades, *Salix* and *Vetrix* (Fig. 4, Supplementary Data Fig. S5). Bootstrap support values for most of the internal nodes within the major clades is low, and no subclades were clearly defined within *Vetrix*, with the notable exception of subclade P1, *Salix triandra.* Representatives of subgenera *Chamaetia* and *Vetrix* were intermixed in one highly supported clade. In contrast, the *Salix* clade was fully supported as distinct; all species previously recognized as members of subg. *Salix* (including *S. exigua* and *S. interior*) fell into it, in two well-supported subclades.

DISCUSSION

Phylogenomic analyses of Salix

Five major clades and one paraphyletic group were formed in autosome data-based trees (Fig. 1 and Supplementary Data Fig. S2). The subclade C1 is mainly comprised of Asian species, especially endemic species of the Qinghai–Tibetan Plateau (QTP, including Hengduan Mountains and Himalaya), which is basically supporting the finding of a radiation by He *et al*. (2021a) in their Hengduan Mountains (HDM) clade, whereas C2 includes Eurasian and North American willows (Fig. 1; Supplementary data Fig. S2, Table S2). The species of both subclades are present in the *Vetrix* clade of chloroplast genome-based trees (Figure 3; Supplementary Data Fig. S4). The American species *Salix exigua* and *S. interior* (C3) have been considered to belong to subgenus *Longifoliae* (Argus, 2010), but their position within

the *Vetrix* is highly supported in our analysis. This is incongruent with the chloroplast genome trees and may be caused by interspecific hybridization or incomplete lineage sorting of nuclear gene copies (see *"Placement of section Longifoliae*"). *Salix triandra* is the only species in C4, and appears as a sister to C3 and the other *Vetrix* representatives, which is consistent with previous studies (Chen *et al*., 2010; Hardig *et al*., 2010; Wu *et al*., 2015; Lauron-Moreau *et al*., 2015a; Wagner *et al*. 2021b; see also "Placement of *Salix triandra*").

Members of subg. *Salix* as circumscribed by Argus (2010) formed the C5 grade, which is paraphyletic to the *Vetrix* clade. Hence, subgenus *Salix sensu* Argus is a paraphyletic grade when the polyploid C5 group is included. In contrast, in the chloroplast genome trees the species of C5 were embedded into *Salix*, subclades P2 and P3 (Fig. 1, 4; Supplementary Data Fig. S2, S5). This C5 grade is comprised of representatives of sections *Pentandrae*, *Salix*, and *Salix shihtsuanensis* (and its varieties). *Salix shihtsuanensis* were wrongly placed into section *Sieboldianae* based on morphological characters (Fang *et al*., 1999). The position of C5 is similar to the ITS tree of Lauron-Moreau *et al*. (2015b, correction 'Fig 3'), in which the clade mainly consisting of subg. *Salix* species formed a sister group to clades *Longifoliae* and *Chamaetia/Vetrix*. The sampled taxa of C5 (Fig. 1; Supplementary data Fig. S2, Table S2) were all identified as polyploids which may have affected their positions in the two nuclear-based trees. Polyploidy is further a major cause for paraphyly because parental diploid species and polyploid derivatives coexist, whereas reciprocal monophyly establishes only after extinction of ancestors (Hörandl, 2007; Hörandl and Stuessy, 2010). *Salix matsudana* of C5 appeared ~4 million years ago and was predicted to be allotetraploid (Zhang *et al*., 2020). Considering the conflict of the nuclear and chloroplast trees (Figs 1, 4; Supplementary data Figs S2, S5), it might be possible that the members of C5 subclade arose from allopolyploid offspring of crosses between species from clades *Salix* and *Vetrix*; admixture analysis of nuclear data support the hypothesis of hybrid origin (He L, Shanghai, China, unpubl. res.).

However, at present no final taxonomic conclusion can be drawn on the C5 group, and the way how sex is determined in polyploid systems also need to be clarified, as well as SDSs of more species.

Although chloroplast genome-based analysis supports two major robust clades within *Salix*, nuclear-based alternative topologies suggest a more complex subgeneric evolution, as expected. Nevertheless, the four trees based on diploid species only supported both *Vetrix* and *Salix* as monophyletic (Figs 2, 3; Supplementary Data Figs S3, S4). Taken together, representatives of *Chamaetia/Vetrix* grouped together in all our trees, suggesting their close affinity and supporting their merging into one clade *Vetrix*, despite the uncertain placements of *Salix triandra* and section *Longifoliae* (see below). It is also worth to note that since willows are widely distributed across the Northern Hemisphere (Skvortsov, 1999) for further investigation it is important to extend the geographical range of sampling in order to obtain a phylogeny that would more profoundly reflect subgeneric relationships within the genus on a worldwide scale.

Placement of Salix triandra

Although sex determination systems of a few species have been identified so far, representatives of two major groups (*Salix* and *Vetrix*) already exhibit different heterogamety. Species in clade *Salix* have a XX/XY system, whereas species in clade *Vetrix* have a ZW/ZZ system, which may act as a barrier to gene exchange (Stöck *et al*., 2021).

Salix triandra formed a sister branch to the rest of the *Vetrix* clade in all trees. For the past decade *S. triandra* has attracted great interest from willow taxonomists. Affiliation of this species to subg. *Salix* was put in doubt by Trybush *et al*., (2008), in whose study it fell out of *Salix* and *Vetrix* and formed a third cluster with approximately equal genetic similarity to both subgenera. Chen *et al*. (2010) and Wu *et al*. (2015) distinguished *Triandrae* clade (=

sect. *Amygdalinae*), including *S. triandra*, as sister to the whole *Chamaetia/Vetrix* clade. *Salix triandra* has a female heterogamety (ZW) sex determination system (Li *et al*., 2020), the same as all the tested species of *Vetrix* (Table 1). Moreover, hybrids of *Salix triandra* and *S. viminalis* can produce viable seeds; however, no fertile cross was recorded with species of subgenus *Salix* (*S. alba, S. pentandra*) (Karp *et al*., 2011). Nevertheless, diploid *Salix triandra* can hybridize with tetraploid *S. fragilis* from subgenus *Salix*, resulting in a triploid hybrid (*S.* × *alopecuroides*), which is often found in Europe, because the two parental species frequently co-occur along rivulets (Rechinger, 1954; Neumann and Polatschek, 1972; Neumann, 1981; Dobes *et al*., 1997; Wagner *et al*., 2021a). Interestingly, Neumann (1981) reported that these hybrids often have catkins with both male and female flowers, which highlights the complexity of sex determination in polyploid willows as a result of interspecific crosses. The homomorphic sex chromosomes of willows may not act as a complete reproductive barrier (He *et al*., 2021b; Stöck *et al*., 2021). However, the sex determination system of *S. fragilis* is not yet characterized, and it is unknown how SDRs actually work in polyploids. Fertility and abundance of *S.* × *alopecuroides* is unknown, but it does not form populations (pers. obs. E. Hörandl).

The separated phylogenetic position and the high genetic divergence of *Salix triandra* from the other *Vetrix* species speaks against an inclusion of the species into subgenus *Vetrix*. However, its inclusion in subg. *Salix* also seems unwarranted. Although we assigned it under the *Vetrix* clade in this study according to its SDS, further studies with more samples of the "*Amygdalinae*" clade are needed to decide on a final taxonomic placement.

Placement of section Longifoliae

Skvortsov (1968) treated *Longifoliae* as a section of subgenus *Salix*. He was the first to propose to raise it to a subgenus, but he did not classify species outside Eurasia. He also

claimed that species from this section possess a hypodermis lacking chloroplast on both sides of the leaves, which is similar to *Chosenia* (*S. arbutifolia*), which, in turn, belongs to the *Chamaetia/Vetrix* clade (Chen *et al.*, 2010; Hardig *et al.*, 2010). However, Azuma *et al.* (2000) considered the possibility of independent evolution of this morphological feature in different lineages.

In hybridization experiments Mosseler (1989; 1990) used a selection of species from the *Salix* and *Vetrix* clades, including *S. exigua/interior* (Argus 2010). The latter exhibited pollen-pistil incongruity in crosses with members of clade *Salix*, resulting in seed abortion. In contrast, *S. exigua/interior* did not show a pollination barrier in crosses with diploid representatives of clade *Vetrix* (*S. eriocephala* and *S. petiolaris*), and produced viable F1 progeny. These studies suggest a closer affinity of *S. exigua/interior* with clade *Vetrix* than with *Salix*.

No previous molecular phylogenetic studies of *Salix* have proposed to treat *Longifoliae* as a part of *Vetrix*. Chong *et al*. (1995), examining allozyme variation in order to estimate genetic distance between *S. exigua* and other North American willows, found out that the species was equally distant from both *Vetrix* and *Salix*, suggesting the revision of its taxonomical placement. In the result of Leskinen and Alström-Rapaport (1999), *S. exigua* fell out of the main group of *Salix* species, so the authors suggested its earlier divergence. In the ITS tree of Lauron-Moreau *et al*. (2015b) four major clades were formed. One of them consisted of most *Longifoliae* species and appeared as a sister clade to *Chamaetia/Vetrix,* which, however, was not consistent with the plastid-based tree, in which *Longifoliae* perfectly grouped with members of clade *Salix.*

If we assume that differences in the SDSs lead to reproductive barriers between two clades (clade *Salix* and clade *Vetrix*), *S. exigua/interior*, which was found to inter-cross successfully with clade *Vetrix* species producing viable and vigorous F1 progeny (Mosseler,

1989; 1990), may probably have a female heterogamety as do species in the *Vetrix* clade investigated so far. However, due to the lack of research on SDSs of any species of *Longifoliae*, this hypothesis needs to be tested. It is highly likely that phylogenetic incongruence of *S. exigua* and *S. interior* in chloroplast and nuclear trees appeared because of early interspecific hybridization or incomplete lineage sorting of nuclear gene copies (Figs 1, 4; Supplementary Data Figs S2, S5). Therefore, whether *S. exigua* and *S. interior* are the only two unique species of *Longifoliae*, or all species of this subgenus (section) have the same pattern of sex incapability when used in interspecific hybridizations, remains in question. Further identification of the sex determination systems will help to investigate our hypothesis, especially of species of sections *Longifoliae* and *Amygdalinae.*

Reproductive barriers via sex determination systems in willows

In recent years, both male and female heterogamety have been found within *Salix* (see Table 1). Based on our analyses, we hypothesise that the two major groups, *Salix* and *Vetrix*, might have different male and female heterogamety, respectively. One exception is *Salix triandra* (female heterogamety; Li *et al.,* 2020), whose taxonomic position remains uncertain (see "Placement of *Salix triandra*" above). We next discuss whether the sex-determination difference between the two major clades in *Salix* might act as a reproductive barrier.

Various mechanisms can limit plant hybridisation, and both pre- and postzygotic barriers are known (Baack *et al.,* 2015), and prezygotic reproductive barriers like different elevation preferences (Wagner *et al*., 2021a) have been documented between willow species in the subgenera *Vetrix* and *Salix.* Mosseler and Papadopol (1989) concluded that common prezygotic barriers, such as ecological or spatial isolation did not act, since a wide range of hybrids were found in natural populations within a 100-km area around Toronto. On the other hand, the same study recorded the flowering times of seven species from the subgenera *Vetrix*

and *Salix*, and revealed two phenological groups: early flowering *Vetrix* species, and late flowering willows belonging to *Salix*, suggesting that this could be a prezygotic barrier preventing interspecific hybridisation between these two groups of species. Moreover, Mosseler (1989) demonstrated pollen-pistil incongruity between species of the *Vetrix* and *Salix*, since successful pollinations were rare between representatives of subgenera.

Considering the impact of polyploidy on hybridisation, in general, different ploidy levels produce strong crossing barriers in willows (Wagner *et al*., 2021a). On the other hand, it has been recorded that homoploid crosses of species within the same subgenus usually yield viable seeds (Argus, 1974; Mosseler, 1990; Choudhary *et al*., 2013; Gramlich and Hörandl, 2016). For postzygotic barriers it is difficult to disentangle crossing barriers by different ploidy levels from those by putative different SDS, and hence we focus here on crosses on the same ploidy level between the groups.

Data concerning postzygotic barriers and viability of seeds and hybrids obtained from interspecific crosses are scarce. Argus (1974) carried out pollinations between four *Salix* species. Among them, crosses between tetraploid *S. discolor* from subg. *Vetrix* and *S. lucida* from subg. *Salix* (4*x*; Dorn, 1976); no seeds were produced. Diploid *Salix exigua* (sect. *Longifoliae*, subg. *Salix*, Argus, 2010) expressed neither pre- nor postzygotic barriers with diploid representatives of subg. *Vetrix* (Mosseler and Papadopol, 1989; Mosseler, 1989; 1990), suggesting that it may belong to this subgenus (which is consistent with the trees based on nuclear sequences in Figs 1-3, Supplementary Data Figs S2-S4, but not with the chloroplast result in Fig. 4 and Supplementary Data Fig. S5). *Salix humboldtiana,* a diploid South American species from subg. *Salix* (Argus, 1997), not studied here, was used in breeding experiments for developing multipurpose willows. When used as a pistillate parent, it produced progeny, though weak, when pollinated by diploid *S. viminalis*, *S. purpurea* and *S. daphnoides* (also from *Vetrix* clade, not studied here) (Argus, 2010; Bubner *et al*., 2018;

Förster *et al*., 2021). Finally, the allopolyploid origin of the C5 grade in our phylogeny could have resulted from parental lineages of the diploid *Vetrix* and *Salix* clades.

Salix fits a single factor-model of the SDS (He *et al*., 2021b; Renner and Müller 2021; Sanderson *et al*. 2021). According to Renner and Müller (2021), females in the ZW system are heterozygous for a dominant W-linked femaleness factor. In the XY system, a Y-linked specific factor in the heterozygous males dominantly suppresses female functions. A hypothetical scheme in Supplementary Figure S6 shows the four offspring genotype classes when a zW female is inter-crossed with a xY male. Only two of them, with W (15, dominant) and x (7, recessive), and z (15, recessive) and Y (7, dominant) have no conflict between the dominance of the SD factors on the different chromosomes, and (assuming that presence of a chromosome 7 has no effect on sex in the species whose SD locus is on chromosome 15 unless it carries a dominant factor, and vice versa) might be expected to be female and male, respectively. The zx lacks the W femaleness-specific factor, but also lacks the Y-specific factor, so it could be female. Although, WY were observed in fish (*Xiphophorus maculatus*) (Kallman 1984), its YW individuals can be female or male. However, SDS systems in plants and animals evolved in different ways (Mank, 2022). The outcome cannot be predicted with certainty for the WY of willows too. In the reciprocal cross (zz male × xx female), all offspring are xz. Hence, even if these hybrids were viable, fertility could be very low, especially in competition with the parental population. Even if fertile, there would be probably an excess of female genotypes. Eventually, polyploidy could overcome expression bias of the diploid hybrid by gene duplication, or two locus-systems could evolve; such mechanisms could have helped the allopolyploid C5 group to establish and to evolve fertile species, but this needs to be studied.

The rate of sex chromosome divergence is neither unidirectional nor correlated with time (reviewed by Mank 2022). In willows, homomorphism of sex chromosomes does not

mean that they have recently evolved (He *et al*., 2021b; Mank 2022). Although, turnover events are possible within each clade of *Salix* (Renner and Müller, 2021). It is reasonable to assume that species within each clade share the same ancestral SDS on same chromosome (Almeida *et al.*, 2020). SDSs of willows may have contributed to isolation of the two major groups. A comprehensive dated tree of willows and divergence time estimation of X (Z) and Y (W) with a broader sampling should further aid our understanding of the correlation of the SDSs and *Salix* diversification.

CONCLUSIONS

Overall, within the genus there are the two clearly defined groups, *Vetrix* and *Salix*. The former one comprises four subclades: endemic Asian species of (C1); Eurasian and North American species (C2); two species of sect. *Longifoliae* (C3); *Salix triandra* (C4). The *Salix* group becomes paraphyletic by inclusion of the mainly polyploid species (C5), and includes the group with mainly species of subg. *Protitae* (C6). Our analysis suggested that species expressing female or male heterogamety belong to different clades. However, it is not clear whether the difference in heterogamety is a barrier to hybridisation. This uncertainty is partly due to the fact that the SDSs of only seven species have been identified, while the type of heterogamety remains unknown for many species used in breeding experiments, and partly because it remains unclear how ploidy affects interspecific mating (see "Placement of *Salix triandra*"). Furthermore, the C5 polyploids (Fig. 1, Supplementary data Fig. S2, Table S2) could have originated from hybridization of species of the *Salix* and *Vetrix*. If so, this would support the incompleteness of the postzygotic reproductive barriers due to different SDS systems between willows of *Salix* and *Vetrix* clades.

SUPPLEMENTARY DATA

Supplementary data are available online at https://academic.oup.com/aob and consist of the following. **Figure S1**: Selected flow cytometry histograms of the estimated samples. **Figure S2**: Phylogeny inferred for 70 taxa and 91 samples of the genus *Salix* and the outgroup *Populus euphratica* based on maximum likelihood analyses of the AR-pur dataset using RAxML. **Figure S3**: Phylogeny inferred for 38 diploid taxa and 49 samples of the genus *Salix* and the outgroup *Populus euphratica* based on maximum likelihood analyses of the AR-Di-pur dataset using RAxML. **Figure S4**: Phylogeny inferred for 38 taxa and 49 diploid samples of the genus *Salix* and the outgroup *Populus euphratica* based on the SCG-Z-LR dataset analysed with ASTRAL species tree methods. **Figure S5**: Phylogeny inferred for 70 taxa and 91 samples of the genus *Salix* and the outgroup *Populus euphratica* based on maximum likelihood analyses of the CP-pur dataset using RAxML **Figure S6**. Hypothetical scheme of crosses between the ZW/ZZ and XX/XY systems. **Table S1**: The recent significant treatments of *Salix* with focus on subgeneric classification. **Table S2**: Details of plant materials used in this study. **Table S3:** Statistics of quality control results of whole genome sequencing datasets of 91 samples. **Table S4**: Summary of mapping results of 91 samples using *Salix dunnii* as reference genome.**Table S5**: Summary of mapping results of 91 samples using *Salix purpurea* as reference genome.

ACKNOWLEDGMENTS

We thank Deborah Charlesworth for her useful comments on our paper. We thank Shuai Liao, Dong Liu, Robbie E. Hart, Sheng-Jie Fu, Fang He, Jing Ma, Bin Chen, Andrew Townesmith, Chun-Lan Lian, Wan-Cheng Hou, Dan-Ni Zhao, Gongga Expedition, PE Xizang Expedition, Hao Zhang, Jun Zhao, Si-Wen Zeng, Yi Lei, Ce Shang, and Tian Liang for sampling. We are indebted to Ren-Gang Zhang for his useful python script (https://github.com/zhangrengang/degeneracy). Sequence data presented in this article can be downloaded from the NCBI database under BioProject accession PRJNA765215. The developed custom script is available athttps://github.com/oasiswho/vcfbox.

FUNDING

This study was financially supported by the National Natural Science Foundation of China (grant No. 31800466, 32171813) and the Natural Science Foundation of Fujian Province of China (grant No. 2018J01613).

LITERATURE CITED

Abdollahzadeh A, Osaloo KS, Maassoumi AA. 2011. Molecular phylogeny of the genus *Salix* (Salicaceae) with an emphasize to its species in Iran. *The Iranian Journal of Botany* **17**: 244–253.

Alix K, Gerard PR, Schwarzacher T, Heslop-Harrison JSP. 2017. Polyploidy and interspecific hybridization: partners for adaptation, speciation and evolution in plants. *Annals of Botany* **120**: 183–194.

Almeida P, Proux-Wera E, Churcher A, *et al.* 2020. Genome assembly of the basket willow, *Salix viminalis*, reveals earliest stages of sex chromosome expansion. *BMC Biology* **18**: 1–18.

Argus GW. 1974. An experimental study of hybridization and pollination in *Salix* (willow). *Canadian Journal of Botany* **52**: 1613–1619.

Argus GW. 1997. Infrageneric Classification of *Salix* (Salicaceae) in the New World. *Systematic Botany Monographs* **52**: 1–121.

Argus GW. 2010. Salix. In: Flora of North America Editorial Committee, eds. *Flora of North America North of Mexico*. New York: Oxford University Press, **7**: 23–51.

Azuma T, Kajita, T, Yokoyama J, Ohashi H. 2000. Phylogenetic relationships of *Salix* (Salicaceae) based on rbcL sequence data. *American Journal of Botany* **87**: 67–75.

Baack E, Melo MC, Rieseberg LH, Ortiz-Barrientos D. 2015. The origins of reproductive isolation in plants. *New Phytologist* **207**: 968–984.

Barkalov VY, Kozyrenko MM. 2014. Phylogenetic analysis of the Far Eastern *Salix* (Salicaceae) based on sequence data from chloroplast DNA regions and ITS of nuclear ribosomal DNA. *Journal of Plant Science and Conservation* **3**: 3–19.

Boucher LD, Manchester SR, Judd WS. 2003. An extinct genus of Salicaceae based on twigs with attached flowers, fruits, and foliage from the Eocene Green River Formation of Utah and Colorado, USA. *American Journal of Botany* **90**: 1389–1399.

Bubner B, Köhler A, Zaspel I, *et al.* 2018. Breeding of multipurpose willows on the basis of *Salix daphnoides* Vill., *Salix purpurea* L. and *Salix viminalis* L. *Landbauforschung* **68**: 53–66.

van Buijtenen JP, Einspahr DW. 1959. Note on the Presence of Sex Chromosomes in *Populus* tremuloides. *Botanical Gazette* **121**: 60–61.

Chase MW, Zmarzty S, Lledo MD, Wurdack KJ, Swensen SM, Fay MF. 2002. When in doubt, put it in Flacourtiaceae: a molecular phylogenetic analysis based on plastid rbcL DNA sequences. *KEW bulletin* **57**: 141–181.

Chen J-H, Sun H, Wen J, Yang Y-P. 2010. Molecular phylogeny of *Salix* L. (Salicaceae) inferred from three chloroplast datasets and its systematic implications. *Taxon* **59**: 29–37.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884–i890.

Chong DKX, Zsuffa L, Aravanopoulos FA. 1995. Genetic relationship between *Salix exigua* and other North American willows (*Salix* L.): Evidence from allozyme variation. *Biochemical Systematics and Ecology* **23**: 767–771.

Cronk QC, Needham I, Rudall PJ. 2015. Evolution of Catkins: Inflorescence Morphology of Selected Salicaceae in an Evolutionary and Developmental Context. *Front Plant Science* **6**: 1030.

Dai X-G, Hu Q-J, Cai Q-L, *et al.* 2014. The willow genome and divergent evolution from poplar after the common genome duplication. *Cell Research* **24**: 1274–1277.

Daniell H, Lin C-S, Yu M, Chang W-J. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biology* **17**: 1–29.

Dobes C, Hahn B, Morawetz W. 1997. Chromosomenzahlen zur gefässpflanzen-flora Österreichs. *Linzer Biologische Beiträge* **29**: 5–43.

Doležel J, Greilhuber J, Suda J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2**: 2233–2244.

Dorn RD. 1976. A synopsis of American Salix. *Canadian Journal of Botany* **54**: 2769–2789.

Emms D, Kelly S. 2018. OrthoFinder2: fast and accurate phylogenomic orthology analysis from gene sequences. *Genome Biology* **20**: 238**.**

Fang C-F, Zhao S-D, Skvortsov AK. 1999. Salicaceae. In: Wu Z, Y., Raven, P.,H., eds. *Flora of China.* Beijing and St. Louis: Science Press and Missouri Botanical Garden Press **4**: 139–274.

Förster N, Antoniadou K, Zander M, *et al.* 2021. Chemoprofiling as Breeding Tool for Pharmaceutical Use of *Salix*. *Frontiers in Plant Science* **12**: 579820.

Guo F-Y, Liu K-J, Wang Y-C, Li E-Z, Zhan Z-F, Zhang Z-X. 2021. Complete chloroplast genome sequence of *Salix sinopurpurea* (Salicaceae). *Mitochondrial DNA Part B* **6**: 718–719.

Gramlich S, Hörandl E. 2016. Fitness of natural willow hybrids in a pioneer mosaic hybrid zone. *Ecology and Evolution* **6**: 7645–7655

Hardig TM, Anttila CK, Brunsfeld SJ. 2010. A Phylogenetic Analysis of *Salix* (Salicaceae) Based on matK and Ribosomal DNA Sequence Data. *Journal of Botany* **2010**: 1–12.

He L. 2018. Taxonomy and nomenclature of *Salix baileyi*, *S. rehderiana*, and *S. disperma*. *Phytotaxa* **349**: 54–60.

He L, Chen S-P. 2017. Identity of *Salix chingiana* (Salicaceae) from China. *Phytotaxa* **323**: 97–100.

He L, Jia K-H, Zhang R-G, *et al.* 2021b. Chromosome-scale assembly of the genome of *Salix dunnii* reveals a male-heterogametic sex determination system on chromosome 7. *Molecular Ecology Resources* **21**: 1966–1982.

He L, Liao S, Zhang Z-X. 2014. Taxonomic revision of *Salix opsimantha* Schneid. and *Salix oreinoma* Schneid. *Acta Botanica Boreali-Occidentalia Sinica* **34**: 1904–1908.

He L, Liao S, Zhang Z-X. 2015. Nomenclature and taxonomy of *Salix floccosa*, *S. opsimantha* and *S. austrotibetica*. *Phytotaxa* **201**: 158–164.

He L, Wagner ND, Hörandl E. 2021a. Restriction‐site associated DNA sequencing data reveal a radiation of willow species (*Salix* L., Salicaceae) in the Hengduan Mountains and adjacent areas. *Journal of Systematics and Evolution* **59**: 44–57.

Hörandl E. 2007. Neglecting evolution is bad taxonomy. *Taxon* **56**: 1–5.

Hörandl E, Stuessy TF. 2010. Paraphyletic groups as natural units of biological classification. *Taxon* **59**: 1641–1653.

Hou J, Ye N, Zhang D-F, *et al.* 2015. Different autosomes evolved into sex chromosomes in the sister genera of *Salix* and *Populus*. *Scientific Reports* **5**: 1–6.

Isebrands JG, Richardson J. 2014. *Poplars and willows: trees for society and the environment*. Oxford: CABI.

Kallman KD. 1984. A new look at sex determination in poeciliid fishes. In Turner BJ, eds. *Evolutionary Genetics of Fishes*. New York: Plenum Publishing, 95–171.

Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods* **14**: 587–589.

Karp A, Hanley SJ, Trybush SO, Macalpine W, Pei M, Shield I. 2011. Genetic improvement of willow for bioenergy and biofuels. *Journal of Iintegrative Plant Biology* **53**: 151–65.

Lauron-Moreau A, Pitre FE, Argus GW, Labrecque M, Brouillet L. 2015a. Phylogenetic relationships of American willows (*Salix* L., Salicaceae). *PLoS One* **10**: e0121965.

Lauron-Moreau A, Pitre FE, Argus GW, Labrecque M, Brouillet L. 2015b. Correction: Phylogenetic Relationships of American Willows (*Salix* L., Salicaceae). *PLoS One* **10**: e0138963.

Leskinen E, Alström-Rapaport C. 1999. Molecular phylogeny of Salicaceae and closely related Flacourtiaceae: Evidence from 5.8 S, ITS 1 and ITS 2 of the rDNA. *Plant Systematics and Evolution* **215**: 209–227.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Li H, Handsaker B, Wysoker A, *et al.* 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Li M-M, Wang D-Y, Zhang L, *et al.* 2019. Intergeneric Relationships within the Family Salicaceae s.l. based on Plastid Phylogenomics. *International Journal of Molecular Sciences* **20**: 3788.

Li W, Wu H-T, Li X-P, Chen Y-N, Yin T-M. 2020. Fine mapping of the sex locus in *Salix triandra* confirms a consistent sex determination mechanism in genus *Salix*. *Horticulture Research* **7**: 1-12.

Liu L-J, He L, Applequist WL. 2020. Untangling two Chinese *Salix* species (Salicaceae) published by CK Schneider, with lectotypification of four names. *Willdenowia* **50**: 159-163.

Liu L-J, He L, Zhang Z-X. 2016. Two new records of *Salix* L. from China. *Acta Botanica Boreali-Occidentalia Sinica* **36**: 426–428.

Ma Z-Y, Wen J, Ickert-Bond SM, Nie Z-L, Chen L-Q, Liu X-Q. 2018. Phylogenomics, biogeography, and adaptive radiation of grapes. *Molecular Phylogenetics and Evolution* **129**: 258–267.

Malmstrøm M, Matschiner M, Tørresen OK, Jakobsen KS, Jentoft S. 2017. Whole genome sequencing data and de novo draft assemblies for 66 teleost species. *Scientific Data* **4**: 1–13.

Mank JE. 2022. Are plant and animal sex chromosomes really all that different? *Philosophical Transactions of the Royal Society, B,* in press.

Minh BQ, Schmidt HA, Chernomor O, *et al.* 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**: 1530–1534.

Mosseler A, Papadopol CS. 1989. Seasonal isolation as a reproductive barrier among sympatric *Salix* species. *Canadian Journal of Botany* **67**: 2563–2570.

Mosseler A. 1989. Interspecific pollen–pistil incongruity in *Salix*. *Canadian Journal of Forest Research* **19**: 1161–1168.

Mosseler A. 1990. Hybrid performance and species crossability relationships in willows (*Salix*). *Canadian Journal of Botany* **68**: 2329–2338.

Neumann A, Polatschek A. 1972. Cytotaxonomischer beitrag zur gattung *Salix*. *Annalen des Naturhistorischen Museums in Wien* **76**: 619–633.

Neumann A. 1981. Die mitteleuropäischen *Salix*-Arten. *Mitteilungen der forstlichen Bundesversuchsanstalt Wien* **134**, 1–152.

Ohashi H. 2006. Salicaceae. In: Boufford D, Ohba H., eds. *Flora of Japan*. Tokyo: Kodansha, IIa: 7–25.

Olofsson JK, Cantera I, Van de Paer C, *et al.* 2019. Phylogenomics using low-depth whole genome sequencing: a case study with the olive tribe. *Molecular Ecology Resources* **19**: 877–892.

Raubeson LA, Jansen RK. 2005. Chloroplast genomes of plants. *Plant diversity and evolution: genotypic and phenotypic variation in higher plants*: 45–68. Wallingford: CABI Publishing.

Rechinger KH fil. 1957. *Salix* L. *Hegi, G.: Illustrierte Flora von Mitteleuropa* **3**: 1.

Renner SS, Muller NA. 2021. Plant sex chromosomes defy evolutionary models of expanding recombination suppression and genetic degeneration. *Nature Plants* **7**: 392–402.

Sanderson BJ, Feng G, Hu N, *et al.* 2021. Sex determination through X-Y heterogamety in *Salix nigra*. *Heredity (Edinb)* **126**: 630–639.

Shendure J, Balasubramanian S, Church GM, *et al.* 2017. DNA sequencing at 40: past, present and future. *Nature* **550**: 345–353.

Skvortsov AK. 1968. *Willows of the U.S.S.R. A Taxonomic and Geographic Revision.* Moscow: Publishing Office "Nauka".

Skvortsov AK. 1999. *Willows of Russia and Adjacent Countries.* Joensuu: University of Joensuu.

Sohma K. 1993. Pollen diversity in *Salix* (Salicaceae). *Science Reports of Tohoku University. Series 4 (Biology)* **40**: 77–178.

Stamatakis A. 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* **30**: 1312–1313.

Stöck M, Dedukh D, Reifová R, Lamatsch DK, Starostová Z, Janko K. 2021. Sex chromosomes in meiotic, hemiclonal, clonal and polyploid hybrid vertebrates: along

the ‘extended speciation continuum’. *Philosophical Transactions of the Royal Society B* **376**: 20200103.

Stull GW, Soltis PS, Soltis DE, Gitzendanner MA, Smith SA. 2020. Nuclear phylogenomic analyses of asterids conflict with plastome trees and support novel relationships among major lineages. *American Journal of Botany* **107**: 790–805.

Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**: 2032–2034.

Trybush S, Jahodová Š, Macalpine W, Karp A. 2008. A Genetic Study of a Salix Germplasm Resource Reveals New Insights into Relationships Among Subgenera, Sections and Species. *BioEnergy Research,* **1**: 67–79.

Unamba CI, Nag A, Sharma RK. 2015. Next Generation Sequencing Technologies: The Doorway to the Unexplored Genomics of Non-Model Plants. *Frontiers in Plant Science* **6**: 1074.

Wagner ND, Gramlich S, Hörandl E. 2018. RAD sequencing resolved phylogenetic relationships in European shrub willows (*Salix* L. subg. *Chamaetia* and subg. *Vetrix*) and revealed multiple evolution of dwarf shrubs. *Ecology and Evolution* **8**: 8243–8255.

Wagner ND, He L, Hörandl E. 2020. Phylogenomic Relationships and Evolution of Polyploid *Salix* Species Revealed by RAD Sequencing Data. *Frontiers in Plant Science,* **11.** 1077.

Wagner ND, He L, Hörandl E. 2021a. The Evolutionary History, Diversity, and Ecology of Willows (*Salix* L.) in the European Alps. *Diversity* **13**: 146–161.

Wagner ND, Volf M, Hörandl E. 2021b. Highly diverse shrub willows (*Salix* L.) share highly similar plastomes. *Frontiers in Plant Science*: 1740.

Westergaard M. 1958. The mechanism of sex determination in dioecious flowering plants. *Advances in genetics* **9**: 217–281.

Wu J, Nyman T, Wang DC, Argus GW, Yang YP, Chen JH. 2015. Phylogeny of *Salix* subgenus *Salix* s.l. (Salicaceae): delimitation, biogeography, and reticulate evolution. *BMC Evolutionary Biology* **15**: 1–13.

Yoshida K, Makino T, Yamaguchi K, *et al.* 2014. Sex chromosome turnover contributes to genomic divergence between incipient stickleback species. *PLoS Genetics* **10**: e1004223.

Zeng S-W, He L. 2020. Lectotypification and Morphological Notes of *Salix annulifera* C. Marquand & Airy Shaw (Salicaceae). *Journal of Tropical and Subtropical Botany* **28**: 236–240.

Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018a. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC bioinformatics* **19**: 15–30.

Zhang J, Yuan H, Li Y, *et al.* 2020. Genome sequencing and phylogenetic analysis of allotetraploid *Salix matsudana* Koidz. *Horticulture Research* **7**: 201.

Zhang L, Xi Z, Wang M, Guo X, Ma T. 2018b. Plastome phylogeny and lineage diversification of Salicaceae with focus on poplars and willows. *Ecology and Evolution* **8**: 7817–7823.

Zhang Q, Liu Y. 2003. Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. *Plant and Cell Physiology* **44**: 941–951.

Zhou R, Macaya-Sanz D, Carlson CH, *et al.* 2020. A willow sex chromosome reveals convergent evolution of complex palindromic repeats. *Genome Biology* **21**: 1–9.

FIGURE LEGENDS

Fig. 1. Phylogeny inferred for 70 taxa and 91 samples of the genus *Salix* and the outgroup *Populus euphratica* based on maximum likelihood analyses of the AR-dun (using the genome of *Salix dunnii* as reference, excluding chromosomes 7 and 15) dataset (233,684 four-fold degenerate SNPs) using RAxML. The subgenera were according to Skvortsov (1999), Ohashi (2006), and Argus (2010). Fang *et al*. (1999) divided *Salix* species only on sectional level, however not all of those sections are monophyletic groups (He *et al*., 2021a). Thus, only species that were confirmed to belong to particular subgenera, based on Skvortsov (1999), Ohashi (2006), and Argus (2010)'s works, were coloured, accordingly.

Fig. 2. Phylogeny inferred for 38 diploid taxa and 49 samples of the genus *Salix* and the outgroup *Populus euphratica* based on maximum likelihood analyses of the AR-Di-dun (using the genome of *Salix dunnii* as reference, excluding chromosomes 7 and 15) dataset (207,155 four-fold degenerate SNPs) using RAxML.

Fig. 3. Phylogeny inferred for 38 taxa and 49 diploid samples of the genus *Salix* and the outgroup *Populus euphratica* based on the SCG-X-LR dataset (using the genome of *Salix dunnii* as reference, 32 single-copy genes 1,165 SNPs) analysed with ASTRAL species tree methods.

Fig. 4. Phylogeny inferred for 70 taxa and 91 samples of the genus *Salix* and the outgroup *Populus euphratica* based on maximum likelihood analyses of the CP-dun (using the chloroplast genome of *Salix dunnii* as reference) dataset (416 SNPs) using RAxML.

TABLES

Table 1. Summary of current information about sex determination systems in *Salix*, adapted from He *et al.* (2021b)

Species	Male or female heterogamety	Chromosome carrying the sex-determining locus	References
***Salix* clade**			
S. dunnii	Male (XX/XY)	7	He *et al*. (2021b)
S. nigra	Male (XX/XY)	7	Sanderson *et al*. (2021)
***Vetrix* clade**			
S. triandra	Female (ZW/ZZ)	15	Li *et al.* (2020)
S. purpurea	Female (ZW/ZZ)	15	Zhou *et al*. (2020)
S. suchowensis	Female (ZW/ZZ)	15	Hou *et al*. (2015)
S. viminalis	Female (ZW/ZZ)	15	Almeida *et al*. (2020)
S. polyclona	Female (ZW/ZZ)	15	He *et a*l., in prep.

Table 2. Total eight datasets used to conduct phylogenetic analyses

Dataset	
Nuclear genome	
	four-fold degenerate SNPs of autosomal region
AR-dun	233,684
AR-pur	258,908
	four-fold degenerate SNPs of autosomal region (diploids only)
AR-Di-dun	207,155
AR-Di-pur	230,084
	SNPs of single copy genes in sex-linked region (diploid only)
SCG-X-LR	1,165
SCG-Z-LR	663
Chloroplast genome	
CP-dun	416
CP-pur	988

NOTE.—Dun and X-LR used *Salix dunnii* genome as reference, X-LR represents the X-linked region on the chromosome 7 of *Salix dunnii*. Pur and Z-LR used *Salix purpurea* genome as reference; Z-LR represents the Z-linked region on the chromosome 15 of *Salix purpurea*. AR, CP, and SCG representautosomal region, chloroplast genome, and single copy gene, respectively.

Figure 1

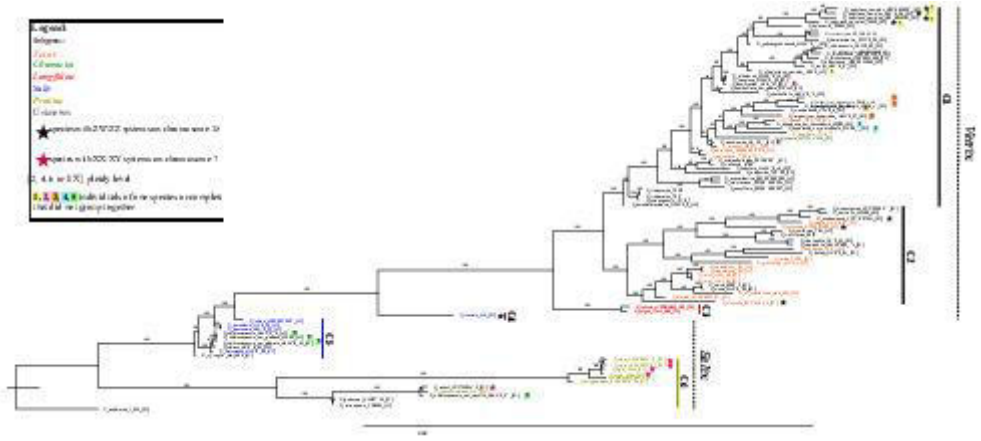


Figure 2

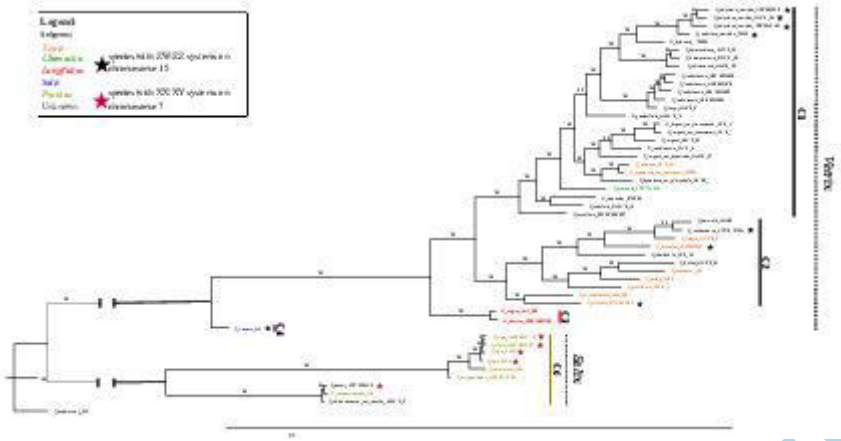


Figure 3

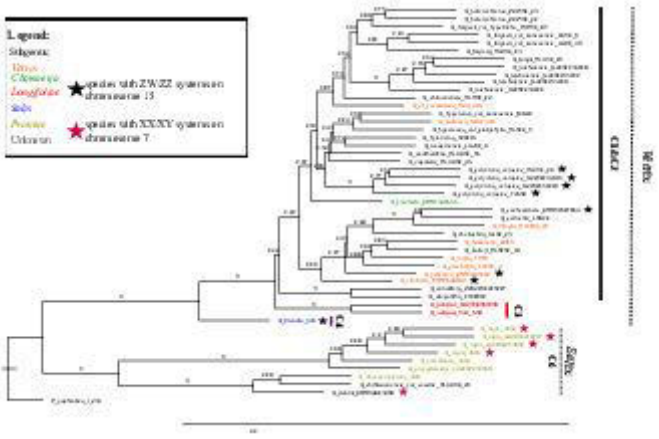
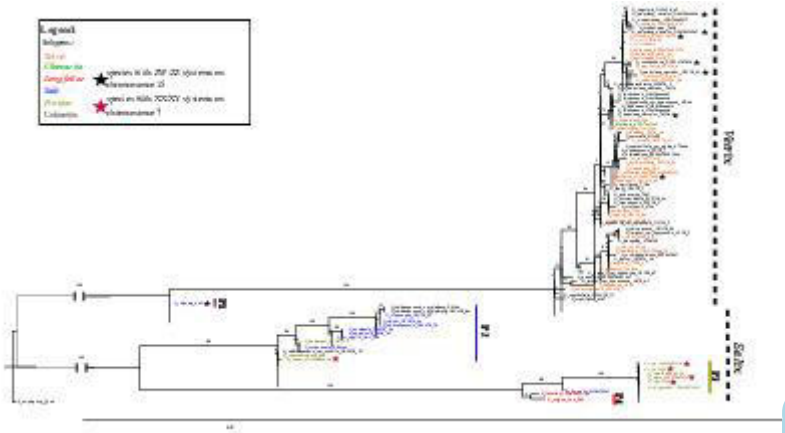
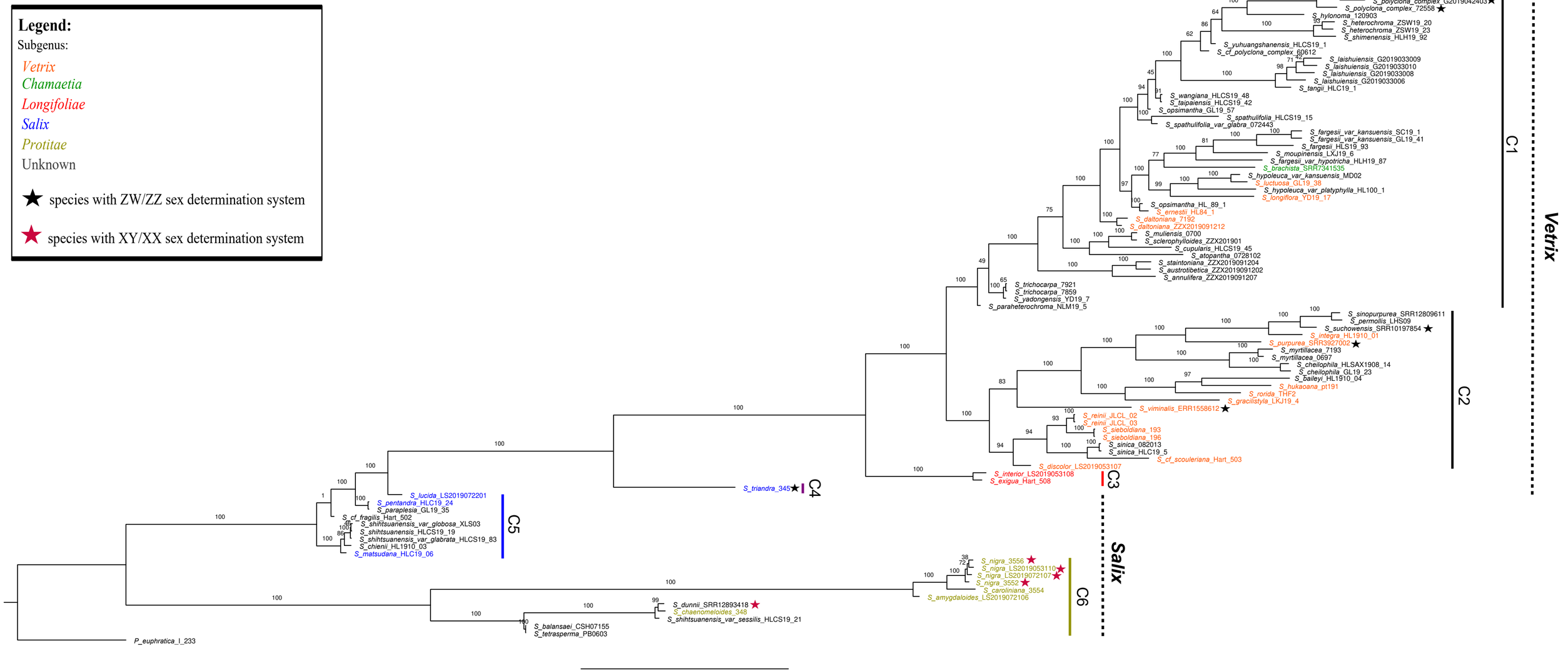
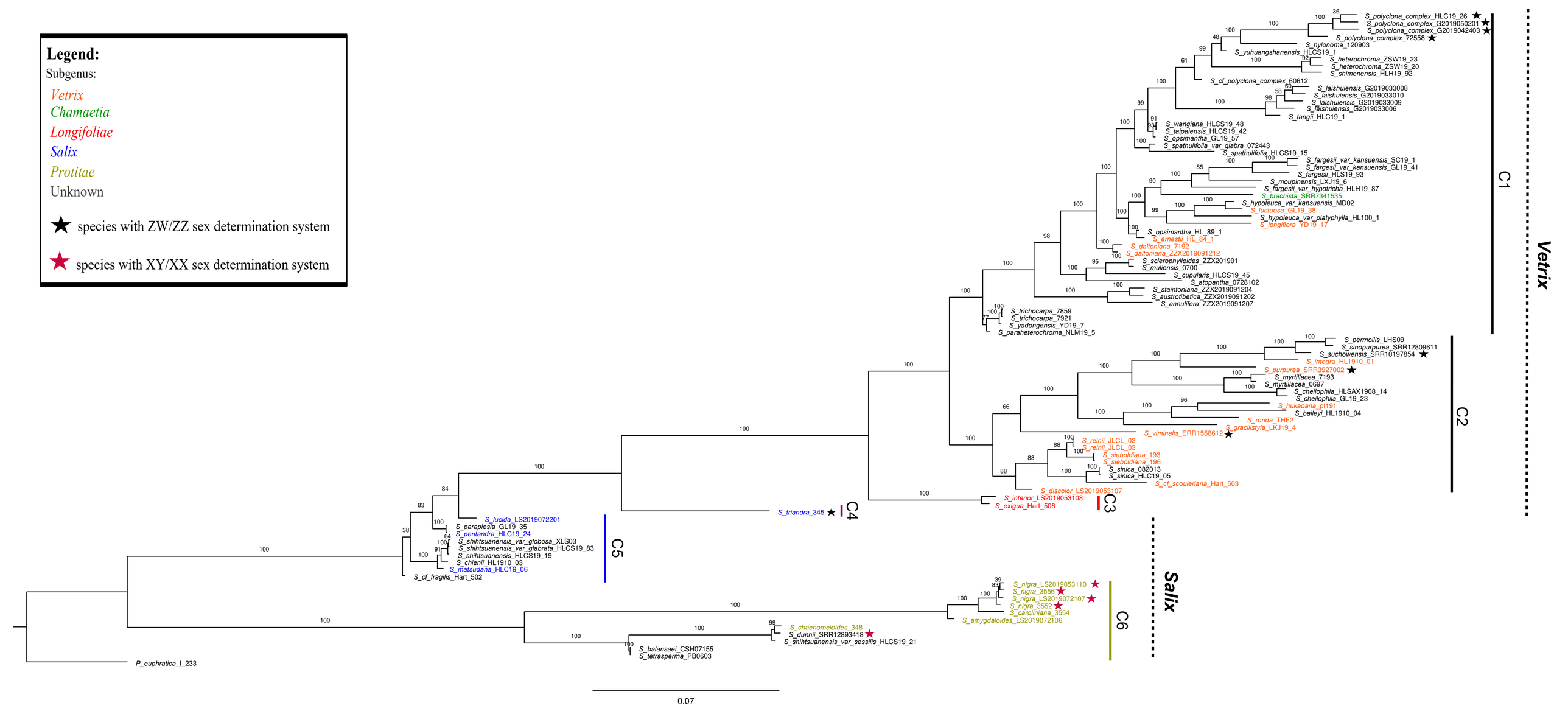
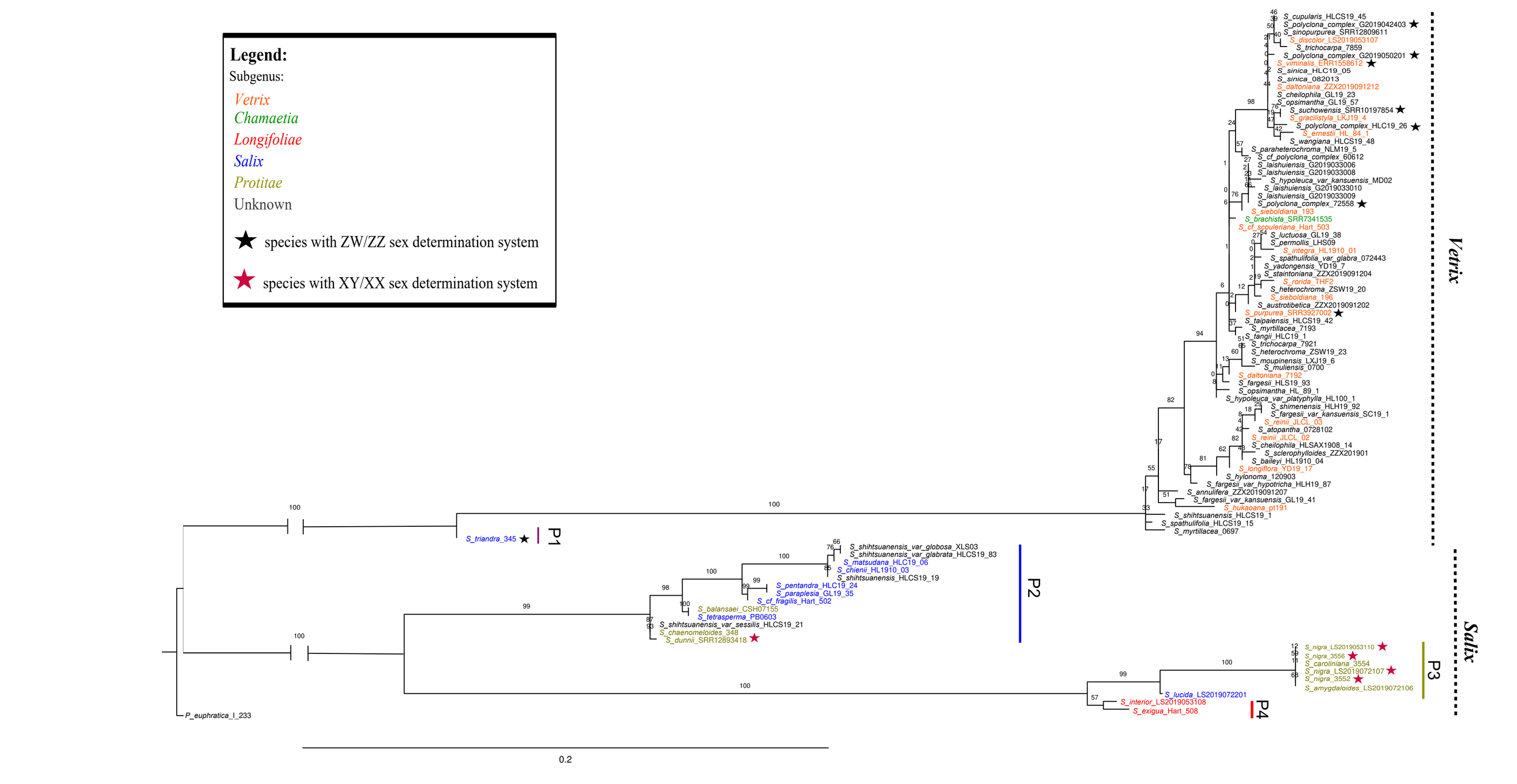


Figure 4




Legend:
Subgenus:
Vetrix
Chamaetia
Longifoliae
Salix
Protitae
Unknown
★ species with ZW/ZZ sex determination system
★ species with XY/XX sex determination system
S_polyclona_complex_HLC19_26 ★
S_polyclona_complex_G2019050201 ★
S_polyclona_complex_G2019042403 ★
S_polyclona_complex_72558 ★
S_hylonoma_120903
S_heterochroma_ZSW19_20
S_heterochroma_ZSW19_23
S_shimenensis_HLH19_92
S_yuhuangshanensis_HLCS19_1
S_cf_polyclona_complex_60612
S_laishuiensis_G2019033009
S_laishuiensis_G2019033010
S_laishuiensis_G2019033008
S_laishuiensis_G2019033006
S_tangii_HLC19_1
S_wangiana_HLCS19_48
S_taipaiensis_HLCS19_42
S_opsimantha_GL19_57
S_spathulifolia_HLCS19_15
S_spathulifolia_var_glabra_072443
S_fargesii_var_kansuensis_SC19_1
S_fargesii_var_kansuensis_GL19_41
S_fargesii_HLS19_93
S_moupinensis_LXJ19_6
S_fargesii_var_hypotricha_HLH19_87
S_brachista_SRR7341535
S_hypoleuca_var_kansuensis_MD02
S_luctuosa_GL19_38
S_hypoleuca_var_platyphylla_HL100_1
S_longiflora_YD19_17
S_opsimantha_HL_89_1
S_ernestii_HL84_1
S_daltoniana_7192
S_daltoniana_ZZX2019091212
S_muliensis_0700
S_sclerophylloides_ZZX201901
S_cupularis_HLCS19_45
S_atopantha_0728102
S_staintoniana_ZZX2019091204
S_austrotibetica_ZZX2019091202
S_annulifera_ZZX2019091207
S_trichocarpa_7921
S_trichocarpa_7859
S_yadongensis_YD19_7
S_paraheterochroma_NLM19_5
C1
Vetrix
S_sinopurpurea_SRR12809611
S_permollis_LHS09
S_suchowensis_SRR10197854 ★
S_integra_HL1910_01
S_purpurea_SRR3927002 ★
S_myrtillacea_7193
S_myrtillacea_0697
S_cheilophila_HLSAX1908_14
S_cheilophila_GL19_23
S_baileyi_HL1910_04
S_hukaoana_pt191
S_rorida_THF2
S_gracilistyla_LKJ19_4
S_viminalis_ERR1558612 ★
S_reinii_JLCL_02
S_reinii_JLCL_03
S_sieboldiana_193
S_sieboldiana_196
S_sinica_082013
S_sinica_HLC19_5
S_cf_scouleriana_Hart_503
C2
S_discolor_LS2019053107
S_interior_LS2019053108
S_exigua_Hart_508
C3
S_triandra_345 ★
C4
S_lucida_LS2019072201
S_pentandra_HLC19_24
S_paraplesia_GL19_35
S_cf_fragilis_Hart_502
S_shihtsuanensis_var_globosa_XLS03
S_shihtsuanensis_HLCS19_19
S_shihtsuanensis_var_glabrata_HLCS19_83
S_chienii_HL1910_03
S_matsudana_HLC19_06
C5
S_nigra_3556 ★
S_nigra_LS2019053110 ★
S_nigra_LS2019072107 ★
S_nigra_3552 ★
S_caroliniana_3554
S_amygdaloides_LS2019072106
C6
Salix
S_dunnii_SRR12893418 ★
S_chaenomeloides_348
S_shihtsuanensis_var_sessilis_HLCS19_21
S_balansaei_CSH07155
S_tetrasperma_PB0603
P_euphratica_I_233
0.08


Legend:
Subgenus:
Vetrix
Chamaetia
Longifoliae
Salix
Protitae
Unknown
★ species with ZW/ZZ sex determination system
★ species with XY/XX sex determination system
S_polyclona_complex_HLC19_26 ★
S_polyclona_complex_G2019050201 ★
S_polyclona_complex_G2019042403 ★
S_polyclona_complex_72558 ★
S_hylonoma_120903
S_yuhuangshanensis_HLCS19_1
S_heterochroma_ZSW19_23
S_heterochroma_ZSW19_20
S_shimenensis_HLH19_92
S_cf_polyclona_complex_60612
S_laishuiensis_G2019033008
S_laishuiensis_G2019033010
S_laishuiensis_G2019033009
S_laishuiensis_G2019033006
S_tangii_HLC19_1
S_wangiana_HLCS19_48
S_taipaiensis_HLCS19_42
S_opsimantha_GL19_57
S_spathulifolia_var_glabra_072443
S_spathulifolia_HLCS19_15
S_fargesii_var_kansuensis_SC19_1
S_fargesii_var_kansuensis_GL19_41
S_fargesii_HLS19_93
S_moupinensis_LXJ19_6
S_fargesii_var_hypotricha_HLH19_87
S_brachista_SRR7341535
S_hypoleuca_var_kansuensis_MD02
S_luctuosa_GL19_38
S_hypoleuca_var_platyphylla_HL100_1
S_longiflora_YD19_17
S_opsimantha_HL_89_1
S_ernestii_HL_84_1
S_daltoniana_7192
S_daltoniana_ZZX2019091212
S_sclerophylloides_ZZX201901
S_muliensis_0700
S_cupularis_HLCS19_45
S_atopantha_0728102
S_staintoniana_ZZX2019091204
S_austrotibetica_ZZX2019091202
S_annulifera_ZZX2019091207
S_trichocarpa_7859
S_trichocarpa_7921
S_yadongensis_YD19_7
S_paraheterochroma_NLM19_5
C1
Vetrix
S_permollis_LHS09
S_sinopurpurea_SRR12809611
S_suchowensis_SRR10197854 ★
S_integra_HL1910_01
S_purpurea_SRR3927002 ★
S_myrtillacea_7193
S_myrtillacea_0697
S_cheilophila_HLSAX1908_14
S_cheilophila_GL19_23
S_hukaoana_pt191
S_baileyi_HL1910_04
S_rorida_THF2
S_gracilistyla_LKJ19_4
S_viminalis_ERR1558612 ★
C2
S_reinii_JLCL_02
S_reinii_JLCL_03
S_sieboldiana_193
S_sieboldiana_196
S_sinica_082013
S_sinica_HLC19_05
S_cf_scouleriana_Hart_503
S_discolor_LS2019053107
S_interior_LS2019053108
S_exigua_Hart_508
C3
S_triandra_345 ★
C4
S_lucida_LS2019072201
S_paraplesia_GL19_35
S_pentandra_HLC19_24
S_shihtsuanensis_var_globosa_XLS03
S_shihtsuanensis_var_glabrata_HLCS19_83
S_shihtsuanensis_HLCS19_19
S_chienii_HL1910_03
S_matsudana_HLC19_06
S_cf_fragilis_Hart_502
C5
Salix
S_nigra_LS2019053110 ★
S_nigra_3556 ★
S_nigra_LS2019072107 ★
S_nigra_3552 ★
S_caroliniana_3554
S_amygdaloides_LS2019072106
C6
S_chaenomeloides_348
S_dunnii_SRR12893418 ★
S_shihtsuanensis_var_sessilis_HLCS19_21
S_balansaei_CSH07155
S_tetrasperma_PB0603
P_euphratica_I_233
0.07


Legend:
Subgenus:
Vetrix
Chamaetia
Longifoliae
Salix
Protitae
Unknown
★ species with ZW/ZZ sex determination system
★ species with XY/XX sex determination system
S_cupularis_HLCS19_45
S_polyclona_complex_G2019042403 ★
S_sinopurpurea_SRR12809611
S_discolor_LS2019053107
S_trichocarpa_7859
S_polyclona_complex_G2019050201 ★
S_viminalis_ERR1558612 ★
S_sinica_HLC19_05
S_sinica_082013
S_daltoniana_ZZX2019091212
S_cheilophila_GL19_23
S_opsimantha_GL19_57
S_suchowensis_SRR10197854 ★
S_gracilistyla_LKJ19_4
S_polyclona_complex_HLC19_26 ★
S_ernestii_HL_84_1
S_wangiana_HLCS19_48
S_paraheterochroma_NLM19_5
S_cf_polyclona_complex_60612
S_laishuiensis_G2019033006
S_laishuiensis_G2019033008
S_hypoleuca_var_kansuensis_MD02
S_laishuiensis_G2019033010
S_laishuiensis_G2019033009
S_polyclona_complex_72558 ★
S_sieboldiana_193
S_brachista_SRR7341535
S_cf_scouleriana_Hart_503
S_luctuosa_GL19_38
S_permollis_LHS09
S_integra_HL1910_01
S_spathulifolia_var_glabra_072443
S_yadongensis_YD19_7
S_staintoniana_ZZX2019091204
S_rorida_THF2
S_heterochroma_ZSW19_20
S_sieboldiana_196
S_austrotibetica_ZZX2019091202
S_purpurea_SRR3927002 ★
S_taipaiensis_HLCS19_42
S_myrtillacea_7193
S_tangii_HLC19_1
S_trichocarpa_7921
S_heterochroma_ZSW19_23
S_moupinensis_LXJ19_6
S_muliensis_0700
S_daltoniana_7192
S_fargesii_HLS19_93
S_opsimantha_HL_89_1
S_hypoleuca_var_platyphylla_HL100_1
S_shimenensis_HLH19_92
S_fargesii_var_kansuensis_SC19_1
S_reinii_JLCL_03
S_atopantha_0728102
S_reinii_JLCL_02
S_cheilophila_HLSAX1908_14
S_sclerophylloides_ZZX201901
S_baileyi_HL1910_04
S_longiflora_YD19_17
S_hylonoma_120903
S_fargesii_var_hypotricha_HLH19_87
S_annulifera_ZZX2019091207
S_fargesii_var_kansuensis_GL19_41
S_hukaoana_pt191
S_shihtsuanensis_HLCS19_1
S_spathulifolia_HLCS19_15
S_myrtillacea_0697
Vetrix
S_triandra_345 ★
P1
S_shihtsuanensis_var_globosa_XLS03
S_shihtsuanensis_var_glabrata_HLCS19_83
S_matsudana_HLC19_06
S_chienii_HL1910_03
S_shihtsuanensis_HLCS19_19
S_pentandra_HLC19_24
S_paraplesia_GL19_35
S_cf_fragilis_Hart_502
S_balansaei_CSH07155
S_tetrasperma_PB0603
S_shihtsuanensis_var_sessilis_HLCS19_21
S_chaenomeloides_348
S_dunnii_SRR12893418 ★
P2
S_nigra_LS2019053110 ★
S_nigra_3556 ★
S_caroliniana_3554
S_nigra_LS2019072107 ★
S_nigra_3552 ★
S_amygdaloides_LS2019072106
P3
S_lucida_LS2019072201
S_interior_LS2019053108
S_exigua_Hart_508
P4
Salix
P_euphratica_I_233
0.2


Legend:
Subgenus:
Vetrix
Chamaetia
Longifoliae
Salix
Protitae
Unknown
species with ZW/ZZ sex determination system
species with XY/XX sex determination system
S_wangiana_HLCS19_48
S_ernestii_HL_84_1
S_polyclona_complex_HLC19_26
S_suchowensis_SRR10197854
S_gracilistyla_LKJ19_4
S_sinica_HLC19_05
S_daltoniana_ZZX2019091212
S_discolor_LS2019053107
S_trichocarpa_7859
S_sinopurpurea_SRR12809611
S_polyclona_complex_G2019042403
S_cupularis_HLCS19_45
S_sinica_082013
S_polyclona_complex_G2019050201
S_viminalis_ERR1558612
S_opsimantha_GL19_57
S_cheilophila_GL19_23
S_laishuiensis_G2019033009
S_laishuiensis_G2019033006
S_laishuiensis_G2019033008
S_laishuiensis_G2019033010
S_polyclona_complex_72558
S_hypoleuca_var_kansuensis_MD02
S_sieboldiana_193
S_cf_scouleriana_Hart_503
S_purpurea_SRR3927002
S_trichocarpa_7921
S_luctuosa_GL19_38
S_sieboldiana_196
S_yadongensis_YD19_7
S_muliensis_0700
S_integra_HL1910_01
S_permollis_LHS09
S_heterochroma_ZSW19_20
S_moupinensis_LXJ19_6
S_daltoniana_7192
S_heterochroma_ZSW19_23
S_spathulifolia_var_glabra_072443
S_staintoniana_ZZX2019091204
S_rorida_THF2
S_austrotibetica_ZZX2019091202
S_opsimantha_HL89_1
S_paraheterochroma_NLM19_5
S_cf_polyclona_complex_60612
S_fargesii_HLS19_93
S_hypoleuca_var_platyphylla_HL100_1
S_tangii_HLC19_01
S_myrtillacea_7193
S_brachista_SRR7341535
S_taipaiensis_HLCS19_42
S_reinii_JLCL_02
S_reinii_JLCL_03
S_atopantha_0728102
S_cheilophila_HLSAX1908_14
S_baileyi_HL1910_04
S_shimenensis_HLH19_92
S_fargesii_var_kansuensis_SC19_1
S_sclerophylloides_ZZX201901
S_hylonoma_120903
S_longiflora_YD19_17
S_fargesii_var_hypotricha_HLH19_87
S_hukaoana_pt191
S_shihtsuanensis_HLCS19_1
S_myrtillacea_0697
S_spathulifolia_HLCS19_15
S_annulifera_ZZX2019091207
S_fargesii_var_kansuensis_GL19_41
Vetrix
S_triandra_345
P1
S_shihtsuanensis_var_glabrata_HLCS19_83
S_shihtsuanensis_var_globosa_XLS03
S_shihtsuanensis_HLCS19_19
S_chienii_HL1910_03
S_matsudana_HLC19_06
S_pentandra_HLC19_24
S_paraplesia_GL19_35
S_cf_fragilis_Hart_502
S_balansaei_CSH07155
S_tetrasperma_PB0603
S_shihtsuanensis_var_sessilis_HLCS19_21
S_dunnii_SRR12893418
S_chaenomeloides_348
P2
S_nigra_LS2019072107
S_nigra_3556
S_amygdaloides_LS2019072106
S_nigra_LS2019053110
S_caroliniana_3554
S_nigra_3552
P3
S_lucida_LS2019072201
P4
S_interior_LS2019053108
S_exigua_Hart_508
Salix
P_euphratica_I_233
0.2