Check for updates

New Phytologist

Research

A change in sex chromosomes caused by allopolyploidy in a willow lineage triggered new evolutionary changes in descendant species

Zhi-Qing Xue[1], Xin-Yu Liu[1], Yuan Wang[1], Ren-Gang Zhang[2], Zhi-Ying Zhu[1], Jie Ding[1], Yi Wang[1], Elvira Hörandl[3], Deborah Charlesworth[4] and Li He[1]

[1]Key Laboratory of National Forestry and Grassland Administration on East China Plant Conservation and Utilization, Shanghai Chenshan National Botanical Garden, Shanghai, 201602, China; [2]Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Populations, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201, Yunnan, China; [3]Department of Systematics, Biodiversity and Evolution of Plants (with Herbarium), University of Goettingen, 37073, Göttingen, Germany; [4]Institute of Ecology and Evolution, School of Biological Sciences, University of Edinburgh, Edinburgh, EH9 3FL, UK

Authors for correspondence:
Deborah Charlesworth
Email: deborah.charlesworth@ed.ac.uk

Li He
Email: lhe@cemps.ac.cn

Received: *6 July 2026*
Accepted: *1 November 2026*

New Phytologist (2026)
doi: 10.1111/nph.71771

Key words: allopolyploid, neofunctionalization, pericentromeric regions, recombination suppression, *Salix*, sex chromosome turnover.

Summary

- Polyploidization through hybridization between species with different sex chromosomes is evolutionarily surprising. To understand stable sex-determining systems in such hybrids, we studied a dioecious allotetraploid (4*n*) willow lineage that arose by hybridization between *Salix* species from the *Salix* and *Vetrix* clades.
- We generated chromosome-level assemblies for three allotetraploids and one related diploid, characterized their sex-linked regions (SLRs) to infer their evolutionary history, and analyzed mRNA and small RNA (sRNA) expression profiles of candidate sex-determining genes.
- Female genome assemblies confirmed that the same two subgenomes contributed to all allotetraploids studied. Both progenitor species have male heterogamety (XY systems), with SLRs on different chromosomes (7 and 15), whereas the allotetraploids have female heterogamety (ZW systems).
- Three 4*n* species, *Salix babylonica*, *Salix matsudana*, and *Salix chienii*, retained Z-SLRs in the former chromosome 15Y pericentromeric location, which carry male-determining partial duplications of the *ARR17* gene, and an inversion between the SLRs arose adjacent to these duplications in their common ancestor. Another 4*n* species, *Salix pentandra*, split before these 4*n* species diversified, and the *ARR17* partial duplications are in the chromosome's pseudo-autosomal (recombining) region. We identify a W-linked *ARR22*-like gene and propose it as a candidate sex-determining factor in this species, pending functional testing.

Introduction

Polyploidization has repeatedly evolved in flowering plants, despite several factors acting against such events. As reviewed by Twyford *et al.* (2025), polyploids mostly evolve through hybridization (resulting in 'allopolyploids') involving unreduced gametes carrying the diploid chromosome number. However, interploidy crosses often fail to produce viable seed, due to gene dosage imbalance in the endosperm. Hybridization between dioecious species is further impeded because sterility is especially likely in hybrids between species with different sex chromosomes or sex-determining systems. However, as reviewed by Westergaard (1958), there may be a dominant sex-determining factor that can act in individuals with multiple sets of autosomes. This may explain how some species maintain dioecy after polyploidization.

Indeed, in the autotetraploid *Salix polyclona*-W2 (ZZZW/ZZZZ), the single W can determine femaleness, as in its diploid progenitor, *Salix polyclona* (He *et al.*, 2023; Wang *et al.*, 2026). The tetraploid *Rumex acetosella* retains dioecy and a single pair of sex chromosomes (Cuñado *et al.*, 2007). However, the hexaploid *Diospyros kaki* has lost the dioecy of its diploid progenitors (Akagi *et al.*, 2014; Masuda *et al.*, 2022).

Allopolyploids derived from hybridization between dioecious lineages with different sex-determining systems are less frequently reported (He & Hörandl, 2022; He *et al.*, 2024), though the allopolyploid *Mercurialis canariensis* shares male heterogamety and the same sex-linked region (SLR) with one of its diploid progenitor species (Gerchen *et al.*, 2022), whereas the other progenitor remains unknown. Among the *c.* 450 *Salix* species, *c.* 40% are polyploid (Wagner *et al.*, 2021), with polyploids occurring in both the *Salix* and *Vetrix* clades (Gulyaev *et al.*, 2022; Ogutcen *et al.*, 2024), and at least one case of maintenance of dioecy in an allotetraploid is well documented in *Salix babylonica* (He *et al.*, 2024). How stable sex-determining systems evolve in

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.
All rights reserved, including rights for text and data mining and training of artificial intelligence technologies or similar technologies.

hybrids from species with different sex chromosomes is still an open question.

We studied three further dioecious allotetraploids, *S. pentandra*, *S. chienii*, and *S. matsudana*, which all probably originated from hybridization between species from the same two clades, *Salix* and *Vetrix*, as in *S. babylonica* (Zhang *et al.*, 2020; Gulyaev *et al.*, 2022). Fig. 1 and Supporting Information Fig. S1 summarise the proposed origin of the *S. babylonica* sex-determining system, to help describe our new findings. The *Vetrix* clade of the genus *Salix* includes species with male or female heterogamety, with SLRs on chromosome 15 in both cases; for brevity, we term these 15XY and 15ZW species (Fig. 2). Diploid species, including *S. arbutifolia*, *S. exigua*, *S. interior*, and *S. triandra*, have 15XY systems (Hu *et al.*, 2023; Xue *et al.*, 2024; Wang *et al.*, 2024a,b), the ancestral state in the genus *Salix* (Xue *et al.*, 2024; Wang *et al.*, 2024a). In the *Salix* clade, including *S. chaenomeloides*, *S. dunnii*, and *S. mesnyi*, the SLR is on chromosome 7 (7XY; He *et al.*, 2021a; Wang *et al.*, 2022; Wang *et al.*, 2024b). Allotetraploidy in *S. babylonica* has been shown to have arisen following hybridization between two XY species, involving a cross between a 7X7X female from the *Salix* clade and a 15X15Y male from the *Vetrix* clade (He *et al.*, 2024) (Figs 1, S1), yet it shows female heterogamety, as do the species studied here; we refer to these as the 15ZW clade II, distinguishing them from late-branching *Vetrix* '15ZW clade I' species that independently changed to female heterogamety, which include *S. purpurea*, *S. viminalis*, *S. baileyi*, *S. polyclona*, *S. integra*, and *S. gordejevii* (Pucholt *et al.*, 2017; Zhou *et al.*, 2020; Liu *et al.*, 2026; Wang *et al.*, 2026).

Like *S. babylonica*, the other allopolyploid species in 15ZW clade II probably inherited homeologous SLRs and sex-determining factors from their *Salix* and *Vetrix* clade XY ancestors (15XY and 7XY; Fig. 1a). The 15XY and 7XY ancestors probably (Wang *et al.*, 2022, 2024b) have a single-gene male-determining system, similar to that in *Populus* species (Müller *et al.*, 2020; Xue *et al.*, 2020). As reviewed by Wang *et al.* (2026) and Leite Montalvão *et al.* (2022), the SLRs carry partial duplicates of a gene, *ARR17*, that is essential for female development. These partial duplicate sequences generate small RNAs (sRNAs) that promote male development by either silencing the transcription of intact autosomal or pseudo-autosomal *ARR17* copies by DNA methylation or mediating cleavage of *ARR17* mRNA (Müller *et al.*, 2020; Xue *et al.*, 2020); they are thus Y-linked. In these XY willows, each homolog of the autosome 19 carries two intact *ARR17* copies.

The allotetraploid species carry a chromosome descended from a V_Y, and also one descended from a V_X, consistent with the involvement of unreduced $15V_X15V_Y$ gametes early in the change to a tetraploid genotype, as proposed for *S. babylonica* (He *et al.*, 2024). Polyploids in the 15ZW clade II might have originated directly through hybridization involving unreduced gametes from both progenitor species (Fig. S1). Alternatively, diploid hybrids may have formed first, followed by aneuploid gamete formation (due to poor segregation of homeologs during meiosis), which could have favored genome doubling (Fig. S2), restoring fertility by allowing the production of diploid male and female gametes (De Storme & Geelen, 2013). Under either mode of polyploid formation, the ancestral tetraploid would have produced several diploid gamete types carrying the *Vetrix*-clade chromosome 15, including $15V_X15V_Y$ and $15V_X15V_X$ (and, as there is no extensive, fully Y-linked region showing genetic degeneration or major losses of Y-specific genes, possibly also viable $15V_Y15V_Y$ pollen), generating the zygote genotypes shown in Fig. S2. The tetraploids also have two copies of *Salix*-clade chromosomes, which provide essential chromosome 15 genes and could allow a viable $15V_Y15V_Y$ genotype. Furthermore, the diploid ZZ genotype in *Vetrix* 15ZW clade I willows (in which the Z is derived from a Y chromosome) is also viable (Wang *et al.*, 2026).

It is also likely that the $15V_X15V_Y7S_X7S_X$ genotype was female in the ancestral allotetraploid, as in the extant ones, whose $15V_X$ now acts as a W chromosome (He *et al.*, 2024). In this genotype,

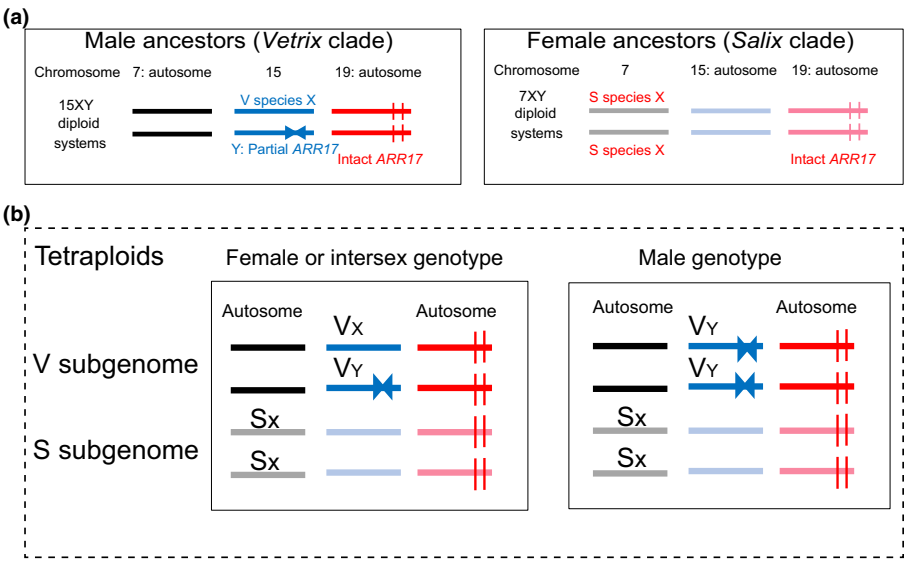


Fig. 1 Hypothesis for the origin of the 15ZW clade II polyploids inferred from *Salix babylonica* (He *et al.*, 2024). (a) The male *Vetrix* (V) ancestor had the 15X15Y genotype, whereas the *Salix* (S) female ancestor was 7X7X. (b) The initial tetraploid hybrids. The $15V_X15V_Y$ genotype is female or a sex-reversed intersex, due to having a higher ratio of intact *ARR17* copies on chromosome 19 to partial *ARR17* copies on chromosome 15 than in $15V_X15V_Y$ diploids. Different colors represent the different chromosomes, and different symbols indicate the locations of known sex-determining factors.

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

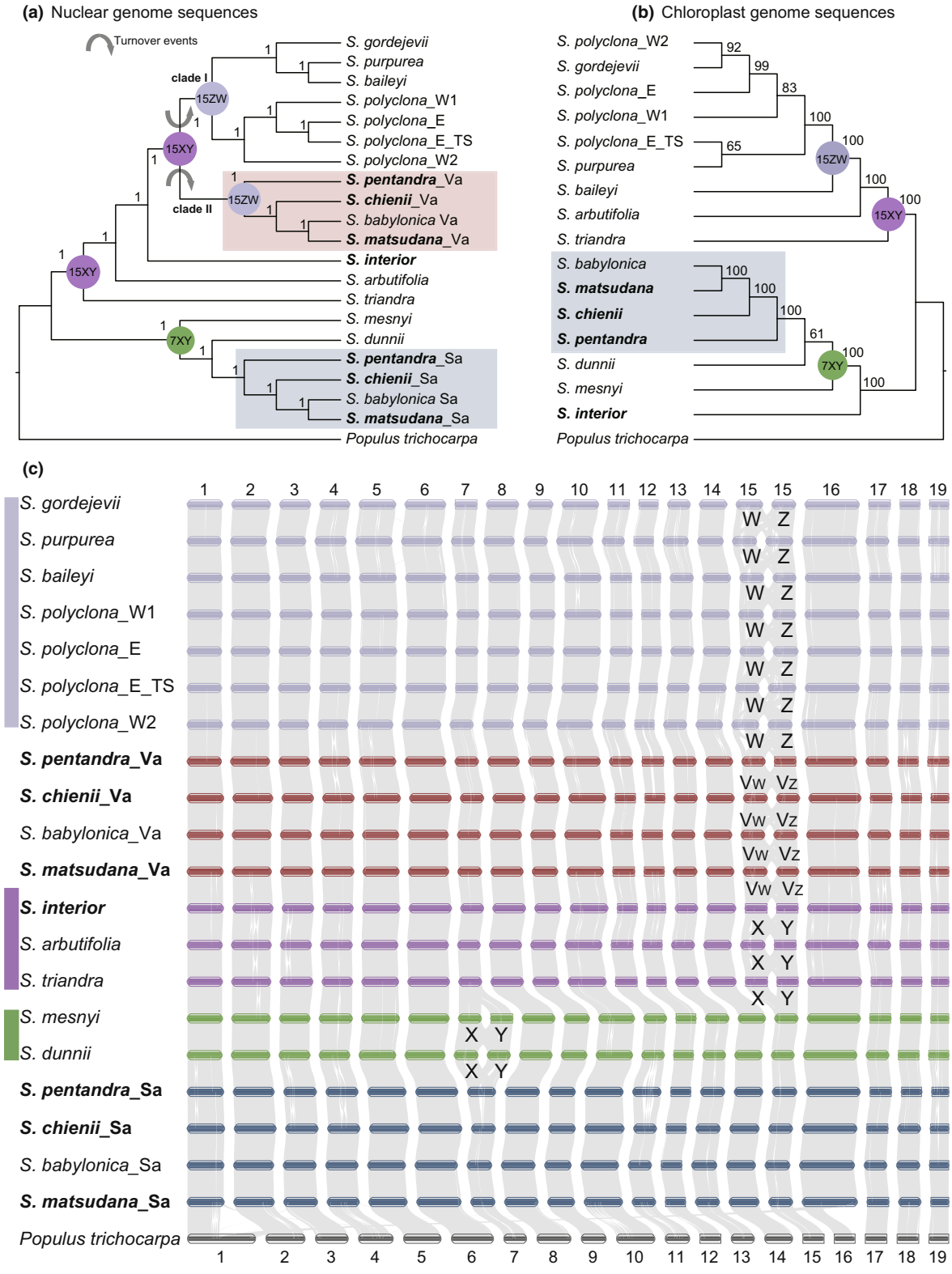


Fig. 2 The origins of the *Salix* allopolyploid lineage and comparisons between the two subgenomes and diploid genomes of other species. (a) Phylogenetic tree estimates based on 5039 single-copy autosomal nuclear genes present in the genome sequences of all 13 willow species analyzed, and in the outgroup *Populus trichocarpa*. The two independent 15XY to 15ZW turnovers are indicated in the diploid/autotetraploid 15ZW clade I and the allotetraploid 15ZW clade II. Numbers at the nodes show the local posterior probabilities. Bold text indicates the two subgenomes of the allopolyploid 15ZW clade II. (b) Phylogenetic tree based on chloroplast genome sequences of the 13 willow species and one poplar outgroup. Numbers marked on the tree represent bootstrap values. (c) Gene synteny between chromosomes, with the subgenomes of the allotetraploids in bold font.

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the number of intact *ARR17* copies on chromosome 19, relative to the number of male-determining partial copies on chromosome 15, is higher than in the diploid state (Figs 1b, S1, S2). This may have allowed this genotype to be female, despite the presence of the maleness-determining factors derived from the $15V_Y$ of its male *Vetrix* clade progenitor (He *et al.*, 2024). Crosses between $15V_X15V_Y$ female and $15V_Y15V_Y$ male genotypes can segregate the same two genotypes in a 1 : 1 ratio (Fig. S2). However, if the $15V_Y$ female-suppressing sequences in the $V_X/V_Y/S_X/S_X$ genotype impair female function, selection could favor stronger female-determining factors on chromosome 19, or weakening of the chromosome 15 male-determining factor(s), potentially explaining the puzzling change from an XY to a ZW system.

To better understand the evolution of 15ZW clade II allotetraploids and test the ideas outlined above, we studied three further members of this lineage, including *S. pentandra*, which branched before the split of *S. babylonica* and two other species. We describe new phased chromosome-level genome assemblies, including one from a male of the diploid *S. interior*, a 15XY basal species in the *Vetrix* clade, as an outgroup to polarize changes. We show that this allotetraploid lineage shares a common hybrid origin of the kind described above (between a *Salix* clade 7X7X female and a *Vetrix* clade 15X15Y male), and analyze changes in the SLRs and sex-determining factors of the different species.

Materials and Methods

Plant materials

Our samples used for new whole-genome sequencing included three females from each of the three allotetraploid *Salix* species with female heterogamety (*Salix chienii* W.C. Cheng, *Salix matsudana* Koidz., and *Salix pentandra* L.) and one male diploid, *Salix interior* Rowlee, a basal species of the *Vetrix* clade based on previous studies (Gulyaev *et al.*, 2022; He *et al.*, 2024; Xue *et al.*, 2024). For each species, we collected young leaves for genome sequencing, and young leaves, catkins, stems, and roots for transcriptome sequencing, plus flower buds from both male and female individuals of the three tetraploid species (with three biological replicates per sex) for mRNA and sRNA sequencing. Additionally, we collected fresh leaves from 15 to 16 flowering individuals of each sex per species and immediately preserved them in silica gel for subsequent population resequencing and analyses to determine SLRs. Voucher specimens are deposited in the Shanghai Chenshan National Botanical Garden herbarium. Detailed sample information is in Table S1. We also analyzed publicly available population-level next-generation short-read resequencing data from 20 females and 18 males *S. interior* in Xue *et al.* (2024) and mRNA datasets of flower buds of *S. babylonica* L. from He *et al.* (2024).

Genome sequencing and RNA sequencing

We performed whole-genome DNA sequencing for *S. matsudana*, *S. chienii*, *S. interior*, and *S. pentandra*, integrating data from PacBio HiFi, Hi-C, and Illumina PCR-free technologies. Additional long-read data were obtained for *S. matsudana* using Oxford Nanopore Technologies (ONT). All sequencing was performed by Novogene (Beijing, China).

Genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method for PacBio HiFi sequencing. HiFi libraries were prepared using the SMRTbell Express Template Prep Kit 2.0 and sequenced on the PacBio Revio platform at > 100× coverage, except for *S. matsudana*, which was sequenced on the PacBio Sequel II platform at 30× coverage. Hi-C libraries were prepared following standard procedures (Belton *et al.*, 2012) and sequenced on an Illumina NovaSeq 6000. PCR-free libraries were prepared using the Illumina TruSeq DNA PCR-Free Library Preparation Kit with genomic DNA extracted using the DNeasy Plant Mini Kit (Qiagen). For *S. matsudana*, ONT sequencing on the PromethION platform generated *c.* 68 Gb of ultra-long reads (> 80× coverage). Illumina paired-end short reads were also generated for whole-genome sequencing of all flowering individuals.

RNA sequencing included both mRNA and sRNA. For mRNA sequencing, total RNA was extracted using the CTAB method, assessed using a Fragment Analyzer 5400, and used for library preparation. For sRNA sequencing, total RNA was extracted using the RNAprep Pure Plant Plus Kit. sRNA libraries were prepared through adapter ligation, reverse transcription, PCR amplification, purification, and size selection. All libraries were sequenced on an Illumina NovaSeq 6000 platform. Detailed sequencing information is provided in Methods S1.

Genome assembly and annotation

The assembly process and software for *S. matsudana* followed the approach in He *et al.* (2024), with an updated pipeline subsequently applied to the remaining three species: *S. chienii*, *S. interior*, and *S. pentandra*. Haplotype-resolved contig assemblies were generated using Hifiasm (Cheng *et al.*, 2021), followed by Hi-C-based chromosome scaffolding with Juicer and 3D-DNA (Durand *et al.*, 2016b; Dudchenko *et al.*, 2017) and manual correction using Juicebox (Durand *et al.*, 2016a). Assembly gaps were filled using quartet (Lin *et al.*, 2023). For chromosomes with incomplete telomeric sequences, targeted reassembly of telomere-associated HiFi reads was performed to extend chromosome ends. Chromosomes were numbered according to their homologous relationships with *S. babylonica* or *S. arbutifolia* Pall. Chloroplast and mitochondrial genomes were assembled using GetOrganelle (Jin *et al.*, 2020), and redundant contigs were removed using Redundans (Pryszcz & Gabaldón, 2016). Assembly quality was assessed using k-mer-based genome size estimates, contig N50, read mapping, and BUSCO completeness scores (Simão *et al.*, 2015).

Repetitive sequences were identified using EDTA (Ou *et al.*, 2019) and annotated with RepeatMasker (http://www.repeatmasker.org/RepeatMasker/). Telomeric repeats were identified using TANDEM REPEATS FINDER (TRF, Benson, 1999).

Genes were annotated using MAKER2 (Holt & Yandell, 2011), integrating *ab initio* predictions from AUGUSTUS (Stanke

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al., 2008) and SNAP (Korf, 2004), transcript evidence from Trinity (Grabherr *et al.*, 2011), HISAT2 (Kim *et al.*, 2015), STRINGTIE (Pertea *et al.*, 2015), and PASA (Haas *et al.*, 2003), and homologous protein evidence from 17 *Salix* and *Populus* species and *Arabidopsis thaliana* (L.) Heynh. (Table S2). All annotations were integrated with EVidenceModeler (EVM) to generate consensus gene models (Haas *et al.*, 2008). Transposable element (TE)-derived coding regions were identified using TEsorter (Zhang *et al.*, 2022) and excluded. PASA was used to upgrade the EVM annotations. Genes with abnormal coding sequences or lengths below 50 amino acids were removed. Noncoding RNAs were annotated using tRNAscan-SE (Lowe & Eddy, 1997), BARRNAP (https://github.com/tseemann/barrnap), and RFAMSCAN. Annotation completeness was evaluated using BUSCO. Gene functions were assigned using EGGNOG-MAPPER (Huerta-Cepas *et al.*, 2017), DIAMOND searches against protein databases (Buchfink *et al.*, 2015), and InterProScan for conserved domains and motifs (Jones *et al.*, 2014). The detailed procedures are described in Methods S1.

Phylogenetic analysis

To identify the evolutionary origin of the allotetraploid lineage, we conducted comprehensive phylogenetic analyses using 13 willow species with haplotype-resolved assemblies (Table S2), with *Populus trichocarpa* Torr. & A. Gray as the outgroup. For the four allotetraploid species, we analyzed both Sa and Va haplotype sequences of autosomes, and the haplotypes a for the diploid and autotetraploid species. Single-copy orthologs (SCOs) were extracted from the genomes using ORTHOFINDER (Emms & Kelly, 2019). The resulting protein sequences were aligned using MAGUS (Smirnov & Warnow, 2020) with default parameters. We reconstructed gene trees with IQ-TREE (Minh *et al.*, 2020). The species tree was then reconstructed using ASTRAL based on the inferred gene trees. Chloroplast genomes of all 13 species were also obtained from available assemblies. After aligning using MAFFT (Katoh & Standley, 2013) and trimming using GBLOCKS (Castresana, 2000) with default parameters, IQ-TREE was used to construct a phylogenetic tree. To test the origins of the Sa, Sb, Va, and Vb haplotypes, we also reconstructed chromosome-specific phylogenies based on SCOs from all 19 chromosomes.

Syntenic patterns between assembly haplotypes

To compare gene arrangements in our reference genomes, we performed analyses of all chromosome sequences from each haplotype used in the phylogenetic analyses with MCSCAN (Python version) (Tang *et al.*, 2008).

Identifying SLRs

Short-read resequencing data were filtered with FASTP (Chen *et al.*, 2018) following the parameters in Wang *et al.* (2024b), and the clean data were used to identify the sex-linked region by identifying sex-specific *k-mers,* chromosome quotient (CQ), and F_{ST} (for diploid *S. interior* only) analyses in windows across the likely sex chromosomes, as detailed in the Results, figure legends, and Methods S1. Changepoint analysis (Killick & Eckley, 2014) was used to detect SLR boundaries objectively within the chromosomes analyzed.

SLR features and evolution

We calculated the densities of protein-coding genes and repetitive sequences (total TEs, LTR-Gypsy, LTR-Copia, tandem repeat, and putative centromeric retrotransposon in maize ('CRM') element types) in 20-kb windows across the genome assemblies. We also estimated the numbers of genes in the SLRs and pseudoautosomal regions (PARs). Candidate pericentromeric regions were identified based on high TE densities and low gene densities, both of which are characteristic of pericentromeric regions and have been observed in other willows (Wang *et al.*, 2025).

To assess gene losses from the SLRs and PARs, we used the *S. dunnii* C.K. Schneid. haplotype *a* assembly of chromosome 15 (an autosome) as a reference to represent the ancestral state, following He *et al.* (2024). Integrating our results with Wang *et al.* (2024b, 2026), we identified genes shared between W- and Z-SLRs in each species and between the X-SLR and Y-SLR in *S. interior*. Genes present in the Z-SLR but lacking BLAST hits in the W-SLR were classified as potentially Z-SLR-specific, and vice versa for W-SLR-specific genes. We further analyzed the sharing of such candidate W-SLR-specific genes among *S. babylonica*, *S. chienii*, and *S. matsudana*.

To investigate (1) subgenome divergence between the two subgenomes of the tetraploid lineage and their diploid *Salix* relatives (using the *a* haplotypes of the V and S subgenomes, the *S. baileyi* C.K. Schneid. *a* haplotype, and *S. dunnii a* haplotype), (2) divergence within the tetraploid lineage (*S. pentandra* subgenomes vs those of the other three tetraploid species), and (3) the timing of SLR formation (Z-W SLRs compared with their orthologous regions in the *a* haplotype assembly of autosome 15 of *S. dunnii*), we quantified sequence divergence by calculating synonymous site divergence (Ks) values. Specifically, we first identified homologous gene pairs by BLAST searches, and excluded pairs with < 70% amino acid sequence identity as potentially nonorthologous. *Ks* values were estimated using yn00 in PAML (Yang, 2007). Divergence times were estimated as $T = \mathrm{Ks}/(2r)$, assuming a neutral mutation rate of 1.0×10^{-9} per year in Salicaceae (Berlin *et al.*, 2011).

Identifying ancestrally sex-linked genes

Previous studies (He *et al.*, 2024; Wang *et al.*, 2024a) suggested that Saarb15bG0044000 and its orthologs were sex-linked before the divergence of *S. arbutifolia*, *S. purpurea* L., and *S. babylonica*. We identified homologs in the species studied here and the published *Salix* genomes of 15ZW clade I and 15XY species. We then reconstructed the phylogeny tree for this gene, and for further genes identified by OrthoFinder as conserved SCOs in SLRs of W and Z chromosomes from *S. babylonica*, *S. matsudana*, and *S. chienii*, chromosome 15 of *S. pentandra*, and the autosome 15a of the 7XY diploid species *S. dunnii*, to test

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

whether other sex-linked gametologs pairs also show evidence of divergence between the ancestral Z and W chromosomes before the 4*n* clade evolved.

Identifying candidate sex-determining factors

To find *ARR17* sequences in the 4*n* species, we used the *Populus trichocarpa ARR17* homolog (Potri.019G133600 of Müller *et al.*, 2020) for BLASTN searches. We classified hits as complete *ARR17* homologs only when they contained all five exons and others as partial *ARR17*-like sequences.

In *S. pentandra*, *ARR17*-like sequences are not in the SLRs (see the Results section). We functionally annotated this species' candidate W-specific genes using the TAIR database (https://www.arabidopsis.org/), and estimated their expression levels by RNA-seq. For candidate sex-determining genes, defined as being in the SLR and also having female-biased expression, we used BLASTP to identify their homologs and any autosomal homologs in whole-genome sequences of this and other species.

We analyzed the distribution of *ARR17* and candidate femaleness-promoting *ARR22* (see the Results section) sequences in all currently published *Salix* genome assemblies, as well as those of the previously reported male-determining candidates *MSF* (Wang *et al.*, 2024b) and *GATA15* (Hyden *et al.*, 2023), and female-determining candidate partial *PISTILLATA* (*PI*) sequences and their intact *PI* progenitors (Wang *et al.*, 2026).

To trace the evolutionary history of *ARR22* genes, we used a *S. babylonica* homolog sequence (Sabab15Bg0046500) as the query. Homologous sequences are present in the W-SLRs of all four 4*n* species, and paralogous sequences are detected on their autosomes. We reconstructed a phylogenetic tree using all the chromosome 15 sequences (including the 15ZW clade I), and integrated gene presence or absence character states with the species tree for ancestral state reconstruction with MESQUITE v.3.81 by both character tracing and likelihood methods (Maddison & Maddison, 2007).

Gene expression analyses

To check that the developmental stages of the flower buds used for transcriptome studies correspond to the onset of sex differentiation, when genes involved in sex differentiation are expressed, we used cryo-section transmission electron microscopy in *S. pentandra* and *S. chienii*, with the stages used for this analysis sampled in the annual flowering season at times close to those used for our RNA-seq and sRNA-seq analyses. For *S. pentandra*, RNA-seq and microscopy samples were collected on October 8, 2024, and October 9, 2025, respectively. For *S. chienii*, the respective dates were November 3, 2024 and October 16, 2025. Pistillate and staminate buds were sectioned using a Leica CM30508 cryostat microtome, and sections were examined and photographed using an LEICA M205 FA optical microscope.

To examine gene expression and identify candidate sex-determining genes, we used RNA-seq data from both female and male flower buds. Raw reads were filtered using FASTP and mapped to the reference genome assemblies using HISAT2. Gene expression levels were quantified using featureCounts and transcripts per million. Differential expression analyses were performed using raw read counts with DESEQ2 v.1.42.1 (Love *et al.*, 2014). Genes were classified as differentially expressed genes based on an absolute $\log_2$ fold change ($\log_2$FC > 1 for female-biased expression and $\log_2$FC < −1 for male-biased expression).

For sRNA analyses, adaptor sequences, noncoding RNAs (rRNAs, tRNAs, snoRNAs, and snRNAs), and plasmid contaminants were removed using the SRNAanno database (Chen *et al.*, 2021). Cleaned reads were aligned to their genomes with SRNAMINER v1.1 (Li *et al.*, 2024), and IGV-sRNA was used to visualize and quantify average nucleotide read coverage across replicates and to examine sRNA distributions around each *ARR17*-like duplicate.

Results

Chromosome-level genome assemblies and annotation

We obtained genome assemblies of allotetraploid *Salix chienii*, *S. matsudana*, and *S. pentandra* (Table 1). As the allotetraploids were suspected to have female heterogamety, based on findings in *S. babylonica* (He *et al.*, 2024; Fig. 2), we sequenced female individuals of these three 4*n* species to include haplotypes carrying the 15Z- and 15W-SLRs. As an outgroup to aid inference of evolutionary changes, a male of the diploid *S. interior*, a 15XY species, was also sequenced (Table S3). The assembly sizes of the newly sequenced allotetraploids (Table S2) ranged from 1.35 to 1.60 Gb, roughly twice that of the diploid, as expected. All four assemblies had high quality (Table 1), with over 99% of Illumina and HiFi reads mapping to the assemblies and few assembly gaps, none within the sex-linked regions (Tables S5, S6); they included 98.7–99.2% of complete BUSCO genes found in the OrthoDB database, with high continuity, coverage, and accuracy (Tables S4, S7).

The estimated numbers of protein-coding genes identified in each haploid genome range from just under 30 000 to almost 35 000 (Table S4). We functionally annotated > 97% of the predicted genes in all species (Tables S8–S10). In these assemblies, repetitive sequence densities exceed 39% (Table S4), predominantly long terminal repeats (LTRs; Table S11), which are no higher than in diploid willows (Table S12).

Phylogenetic analyses

Our species tree resolves *Salix* and *Vetrix* as distinct clades (Fig. 2a), in agreement with previous studies (He *et al.*, 2021b; Gulyaev *et al.*, 2022; Wang *et al.*, 2026). Within the allotetraploid 15ZW clade II, *S. chienii* is sister to a monophyletic clade including *S. babylonica* and *S. matsudana*, and *S. pentandra* is sister to this three-species clade (Fig. 2a,b).

The phylogenetic trees generated from datasets for each of the 19 chromosomes support the conclusion that each current allotetraploid possesses two *Salix* clade haplotypes (Sa and Sb) and two *Vetrix* clade haplotypes (Va and Vb, Figs 2a, S3). A phylogenetic

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 Statistics of the four genome assemblies.

Category	*Salix chienii*				*Salix matsudana*				*Salix pentandra*				*Salix interior*	
Gender	Female				Female				Female				Male	
Ploidy	4*n*				4*n*				4*n*				2*n*	
Total number of contigs	102				100				84				40	
Total number of scaffolds	89				78				78				-	
Number of gaps	13				22				6				0	
Busco (%)	99				98.9				99.2				98.7	
Repeat content (%)	49.7				44.44				50.43				38.78	
Subgenome	Va	Vb	Sa	Sb	Va	Vb	Sa	Sb	Va	Vb	Sa	Sb	a	b
Total assembly size (Mb)	387.33	364.06	396.09	359.72	336.44	323.52	358.51	335.28	418.79	391.11	410.83	375.83	294.51	286.68
Number of distinct protein-coding genes	28 532	28 077	28 892	28 498	27 904	27 674	28 268	28 157	28 608	28 802	28 934	28 561	27 947	27 722
Number of genes	31 104	30 781	31 045	31 663	29 192	28 967	29 843	29 593	38 325	36 752	31 871	32 528	29 275	29 031

All subgenomes yielded 19 pseudochromosome assemblies.

tree based on chloroplast genome sequences confirmed a common origin of all four tetraploid clade II 15ZW lineage species, involving a female progenitor from the *Salix* clade, which has a 7X7X system, and a male progenitor from the *Vetrix* clade, which has a 15X15Y system (Fig. 2b), as in *S. babylonica* (He *et al.*, 2024). No interchromosomal rearrangements were detected between the *Salix* and *Vetrix* clade species, though a few chromosomes had intrachromosomal rearrangements, including inversions (Fig. 2c).

Identifying the sex chromosomes and sex-linked regions in the allotetraploids and the diploid *S. interior*

To determine the heterogametic sex and identify SLRs, we used two approaches: analysis of sex-specific *k*-mers and chromosome quotients (CQ) in our Illumina short reads (see details in the Methods S1; Table S13). For the CQ values, we used the cq-calculate.pl script to calculate the ratio of read coverage between the two sexes for sequences matching two or four reference haplotypes, normalized against autosomes (Hall *et al.*, 2013). The results confirmed male heterogamety in *S. interior* and female heterogamety in the allotetraploids (Figs 3, S4–S11). *S. chienii* and *S. matsudana* showed patterns similar to *S. babylonica* (He *et al.*, 2024); results for *S. pentandra* differ and are described separately below. The middle regions of chromosome 15Va had densely concentrated female-specific *k-mers*, as expected under female heterogamety, consistent with the SLR suggested by CQ, which showed male : female ratios close to 0 (Figs S6–S9). These results suggest that 15Va is the W-linked haplotype, while 15Vb with CQ values close to 2, represents the $15V_Z$ haplotype (Fig. 3a,b).

SLR boundaries in the Va and Vb haplotypes were identified using the larger regions suggested by changepoint analyses of *k-mers* and CQ signals (Table 2 shows the estimated sizes of the SLRs identified in the newly sequenced 4*n* species; regions outside the SLRs are candidate PARs). The SLRs in *S. chienii* and *S. matsudana* have low gene densities and share an inversion in a similar location to that previously found in *S. babylonica* (He *et al.*, 2024) (Table 2; Figs 3a,b, 4, S12). These results support the regions in the middle parts of both chromosome 15 haplotypes in all three species being pericentromeric.

In *S. pentandra*, both analyses identified a physically much smaller SLR near the middle of 15Va and Vb haplotypes, but otherwise similar to those of the other species' SLRs (Figs S11, S12), lacking their shared inversion. The W-SLR overlaps part of the W-SLRs in the other species. Its Z-SLR is even smaller and corresponds to only a small portion of the left-hand part of the Z-SLRs in the other species (Figs S12, S13; Tables S14, S15). These SLRs are adjacent to large, exclusively repetitive regions, which are considerably larger in *S. pentandra* than the central repetitive regions of the other species; it includes 10.46 Mb of the Z haplotype and 4.90 Mb of the W haplotype (Figs 3c, S12). Consistent with the hypothesis in Fig. 1, none of the three newly studied allotetraploids showed *k-mer* or CQ signals suggesting sex-linkage on the *Salix* clade 15S *a* and *b* haplotypes (Figs S6, S8, S10).

In the diploid, *S. interior*, part of haplotype 15a was enriched for male-specific *k-mers*, suggesting a Y haplotype, and 15b showed CQ values consistent with being the X haplotype. High F_{ST} values between the sexes, indicating sex-linkage, were found exclusively on chromosome 15, in repeat-rich middle regions similar to those in the allotetraploid species (Figs 3d, S4, S5).

Evidence that $15V_Z$ in 15ZW Clade II lineage is derived from a $15V_Y$

We confirmed that, during the evolution of female heterogamety in the 15ZW Clade II lineage, the ancestral *Vetrix*-clade 15Y

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

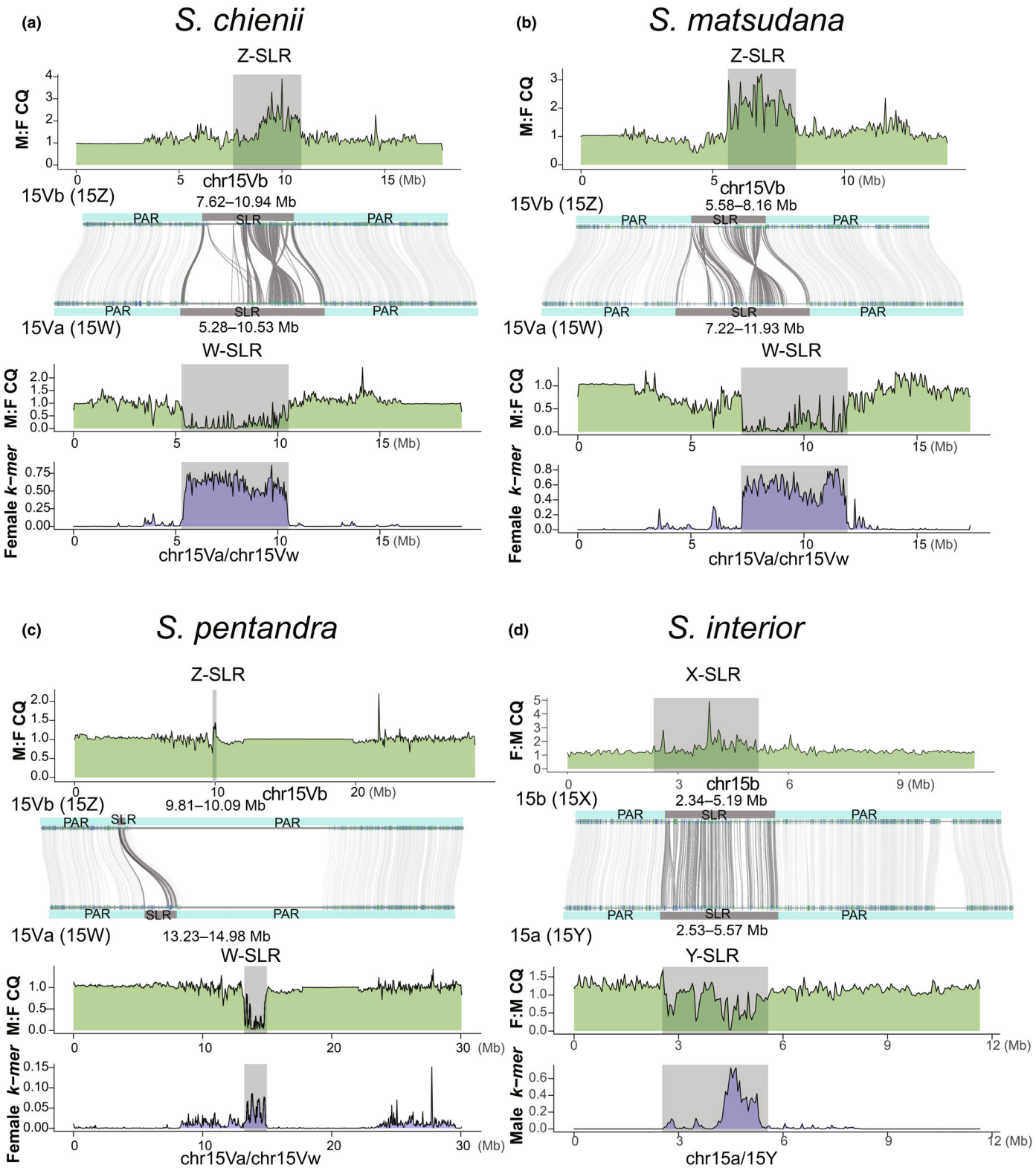


Fig. 3 Inference of sex-linked regions (SLRs) in three allotetraploid 15ZW clade II species and the diploid *Salix interior*. For all three polyploid species with female heterogamety (a, b, c) and *S. interior* with male heterogamety (d), *k-mer* sex-specific abundances and CQ values (see the Materials and Methods section) were used to detect SLRs on the respective chromosome 15 homologs. The *k-mer* analysis results are shown only for the W-linked regions for polyploids, to show results for female-specific sequences, while male-specific sequences are shown for *S. interior* Y-linked regions (see the Materials and Methods section; as shown in Supporting Information Figs S4–S11, very few sequences did not correspond to the stated heterogamety). The inferred SLRs are shown as gray rectangles; blue and green vertical lines on the chromosome diagrams in the synteny plots indicate forward- and reverse-oriented genes, respectively.

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2 Sex-linked regions (SLRs) in the newly sequenced *Salix* species, including *Salix interior*, the diploid outgroup to the 4*n* species, using three types of windowed analyses.

Species	Chromosome	Based on CQ results	Based on F_{ST} values	Based on *k-mer* results	Positions of inversions	SLR based on synteny of genes	Median synonymous site W-Z divergence SLR/PAR (Ks, %)[1]
S. chienii	15Va (W)	5.38–10.53	/	5.38–10.53	5.28–5.38; 8.46–9.79	5.28–10.53	2.2/0.6
	15Vb (Z)	8.88–10.88	/	/	7.62–7.68; 8.93–10.45	7.62–10.94	
S. matsudana	15Va (W)	7.28–11.93	/	7.28–11.88	9.79–11.32	7.22–11.93	2.2/0.3
	15Vb (Z)	5.58–8.13	/	/	6.52–8.00	5.58–8.16	
S. pentandra	15Va (W)	13.23–14.88	/	13.43–14.98	/	13.25–14.97	1.7/0.8
	15Vb (Z)	No clear change point signal	/	/	/	9.81–10.09	
S. interior	15a (Y)	2.53–5.57	4.10–5.47	2.57–5.38	/	2.53–5.57	1.4/0.7
	15b (X)	No clear change point signal	3.74–5.10	/	/	2.34–5.19	

The SLR locations and sizes inferred by changepoint analyses using the CQ, *K-mer*, and F_{ST} results (see the Materials and Methods) are shown in megabases (Mb) in the relevant assemblies. The gene syntenic region between X-SLR and Y-SLR or between Z-SLR and W-SLR was identified by synteny analysis within these ranges. '/' indicates regions that could not be determined by the method in the column.
[1]The value is 1.2% in *S. babylonica* (see Supporting Information Table S20). The higher value, compared with the other species, partly reflects this species' larger SLR, so the genes used for the Ks estimate are not the same as for the other species.

became 15Z (with 15X becoming 15W). Previous analyses (He *et al.*, 2024; Wang *et al.*, 2024a) tested whether the gene trees of SCOs from candidate SLR sequences clustered X and W sequences together, separately from clusters including Y and Z sequences (termed 'sex-patterned'), following an approach used in studies of other species (García-Moreno & Mindell, 2000; Handley *et al.*, 2004; Kirkpatrick *et al.*, 2021). These analyses found only one SCO, Saarb15bG0044000, whose topology suggested that it was already sex-linked in an ancestor of the 15ZW clade II species. In the additional 4*n* species analyzed here, the 15Vw and 15Vz sequences of *S. matsudana* (a close relative of *S. babylonica*) again cluster with the *Vetrix*-clade 15X and 15Y sequences, respectively, supporting the origin of 15Vw from 15X, and 15Vz from 15Y. In *S. pentandra* and *S. chienii*, multiple duplications of this gene (Fig. S14) prevented inference of the origin of their V_Z and V_W chromosomes.

Previous studies found two male-specific sequences (a set of *ARR17* partial duplicates and *MSF*) in a physically small region of the 15Y-SLR of the diploid XY species *S. arbutifolia* and *S. triandra*. Both are also present in the $15V_Z$-SLR of *S. babylonica* (a 15ZW allotetraploid) and *S. polyclona*-TS (a 15ZW clade I diploid), but not in the 15X- or 15W-SLRs of these species (He *et al.*, 2024; Wang *et al.*, 2024b, 2026). Both factors are Z-linked in allotetraploids *S. chienii* and *S. matsudana*, and in the Y-SLR of *S. interior*, supporting the conclusion from *S. babylonica* that the Z evolved from a $15V_Y$ chromosome (Table 3; Fig. S12). In *S. pentandra*, however, these sequences are present in both Z and W haplotypes within the PARs, consistent with this species lacking the inversion present in the other species, and with evidence described below that recombination has occurred between regions corresponding to the Z- and W-linked regions of the other allotetraploids.

Evolution of the SLRs in the 15ZW clade II species

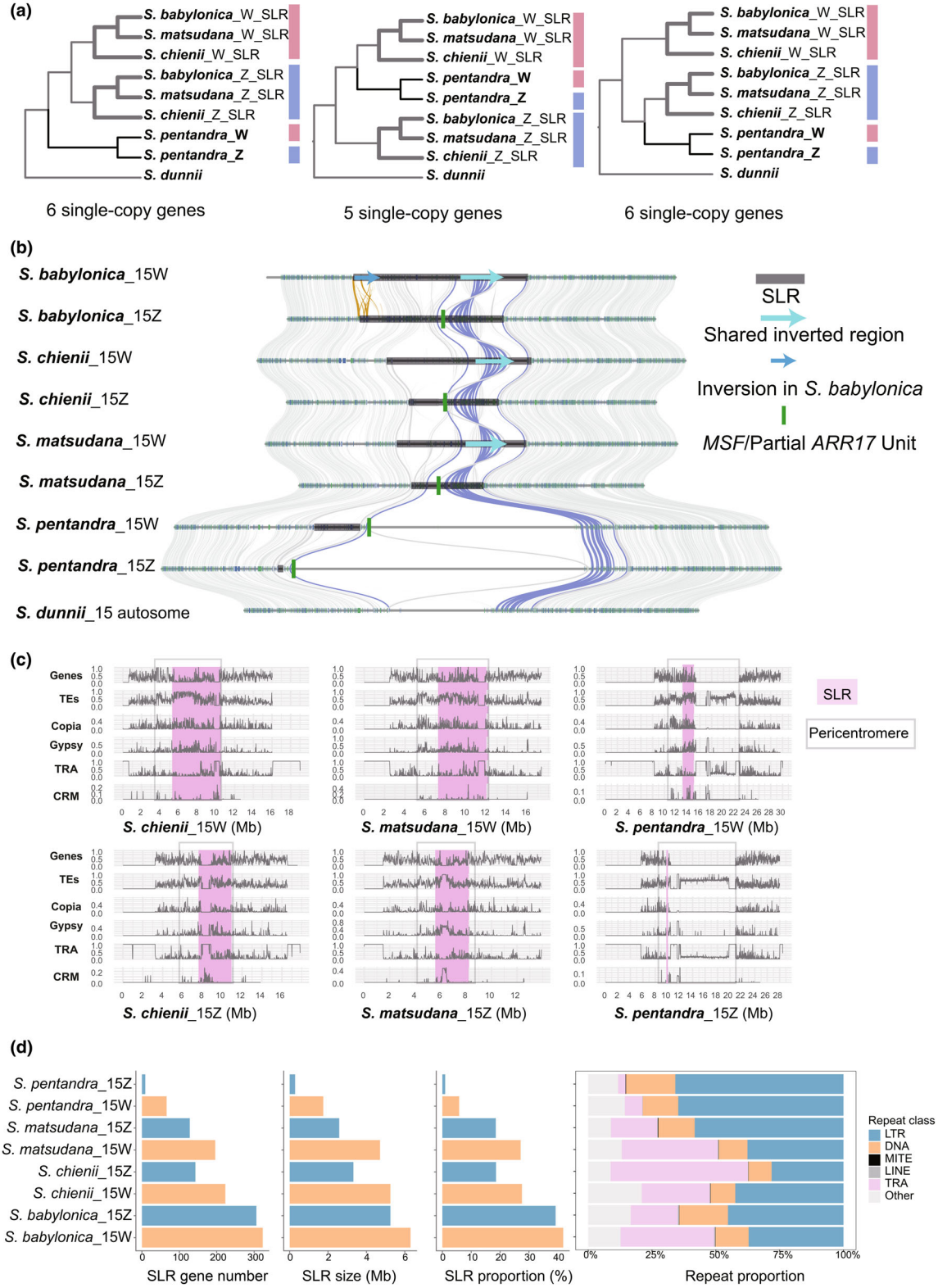
We identified 51 SCO protein-coding genes in the SLRs of *S. chienii*, *S. matsudana*, and *S. babylonica*, the sex chromosomes of *S. pentandra*, and autosome 15a of the outgroup 7XY diploid *S. dunnii* (Table S16). Of these, 17 showed 'sex-patterned' trees suggesting that they were already in a region in which recombination was infrequent in a common ancestor of *S. chienii*, *S. matsudana*, and *S. babylonica* (Fig. 4a), and possibly even in their XY ancestor (Table S15); 15 of these genes are within the inversion described above (Fig. 4b; Tables S16, 17), consistent with the inversion having stopped recombination. The remaining 34 genes showed evidence of recombination after species divergence. Consistent with the absence of the inversion, the 17 'sex-patterned' SCO sequences cluster differently in *S. pentandra* (Fig. 4a).

In the species with the inversion, 15 gametolog pairs have *Ks* values similar to or slightly larger than those between each species and its *S. pentandra* ortholog (Table S15). The male-determining partial *ARR17* sequences are located within or near the left end of the inversion shared by *S. babylonica*, *S. chienii*, and *S. matsudana* (Fig. 4b; Tables S16, S17), supporting the conclusion that an inversion that includes *ARR17* partial duplicates caused sex-linkage in their common ancestor, before divergence from the ancestor of *S. pentandra*. Sixteen W-SLR-specific genes are shared by these three species (Table S18), consistent with their origin in a common ancestor. One of these, orthogroup 0000018 (an unknown function gene located near the right-hand end of the inversion and barely expressed), is absent from both the 15W and 15Z haplotypes of *S. pentandra*.

The SLRs of the allotetraploids are located in the middle of both the Z and W haplotypes (Fig. 3). As their sex chromosomes are metacentric (Fig. 4c; He *et al.*, 2024), this supports the suggestion above that the SLRs are largely pericentromeric. Wang *et al.* (2025) identified the core centromeres of *Populus* and *Salix* chromosomes by CENH3 occupancy and showed that the surrounding regions have properties characteristic of pericentromeric regions, with higher TE densities than other chromosome regions (and some sequences resembling centromere-specific elements termed 'CRM', Naish & Henderson, 2024;

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

Zhong *et al.*, 2002), and consequently lower gene densities. The W-SLRs are consistently larger than the Z-SLRs and carry more duplicated genes and repetitive sequences (Table S14), corresponding to their descent inferred above from X-SLRs, which are consistently larger than Y-SLRs reported in other *Salix*. However, both haplotypes in the 4*n* species share these properties



© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Fig. 4 Gene trees, synteny, and features of sex chromosomes (chromosome 15) in the allotetraploid 15ZW clade II. (a) Phylogenetic tree analyses of the single-copy ortholog (SCO) genes shared by the inferred sex-linked regions (SLRs) of *Salix babylonica*, *Salix matsudana*, *Salix chienii*, and the sex chromosomes of *Salix pentandra* in the 15ZW tetraploid clade II lineage. (b) Collinearity analysis of genes in the assemblies of the two haplotypes of these species' chromosome 15. The horizontal bars represent the chromosomes, and the black regions indicate the SLRs inferred in Fig. 3, showing the different lengths. The 17 SCOs shown in part (a) and described in the text are colored blue; these genes are largely syntenic, included in the regions indicated by cyan arrows, which are inverted between the W and Z in the species other than *S. pentandra*. (c) Protein-coding gene and repeat densities in the W and Z haplotypes of the sex chromosomes in the sequenced females of each species. Pink areas indicate the SLRs inferred in Fig. 3, and the regions of low gene and high repeat density are indicated by the boxes. TRA indicates tandem repeat arrays. Centromeric retrotransposon in maize (CRM) indicates centromere-associated repeats, consistent with the SLRs being largely pericentromeric regions (see the text); they include large regions apparently devoid of protein-coding genes due to very high repetitive content, precluding gene annotation, which probably includes the centromeres. As explained in the text, *S. pentandra* differs from the other three allotetraploid species, and we detected recombination between the regions that are sex-linked in the other species. The figure shows that this species' small SLRs are to the left of the region where they are found in the other species. It lacks the inverted region found in the other species. The large gap in the figure for this species indicates a gene-depleted region, which includes only repetitive sequences. (d) Summary of the numbers of genes and densities of repeat types (and the composition of repeats) in the eight SLRs of the four tetraploid *Salix* lineage species studied, showing the distinctive characteristics in *S. pentandra*. TE, transposable element.

(Figs 4c, S15–S18), indicating that they do not merely reflect the history of the Z region as a formerly Y-linked region, or their current W linkage. The arrangement of the SLRs has also changed since the allopolyploids speciated, while the pericentromeric regions of the autosomes have undergone fewer rearrangements, despite also being repeat-rich (Figs 4b, S19). In *S. chienii* and *S. matsudana*, the 15Z- and W-SLRs appear to include most of the pericentromeric regions, whereas those of *S. pentandra* overlap only small parts of its pericentromeric regions, with its Z-SLR spanning only 0.28 Mb and containing only nine protein-coding genes (Fig. 4c). Consistent with an independent evolutionary history and different PAR boundary inferred above, *c.* 20 Mb of the 28.3 Mb syntenic PAR in *S. pentandra* is highly repetitive, unlike the other species' PARs (the density is 71% in the 15W haplotype, and 68% in the 15Z one); it also has substantial terminal sub-telomeric repetitive regions (Figs 4b,c, S11; Table S14).

The numbers of protein-coding genes annotated within the SLRs varied widely, from 9 to 323 (Table S14). The largest SLR among the four 15ZW clade II species is 6.3 Mb (the *S. babylonica* 15W-SLR; Fig. 4d), a size that mainly reflects an additional inversion in *S. babylonica* (Fig. 4b), which creates sex linkage of a region that is pseudoautosomal in its two closely related species.

The tetraploidy event occurred after the divergence of the *Vetrix* and *Salix* clades, as that inter-clade split corresponds to a median *Ks* of 7.8%, very similar to the inter-subgenome *Ks* of 7.0–7.6% (Fig. S20; Table S15). This divergence is about three times greater than the split between *S. pentandra* Sa or Va sequences and their S or V subgenome counterparts in each of the other three species, for which genome-wide median *Ks* values ranged from 2.4% to 2.7%. The four intraspecies Z-W *Ks* medians are even lower, between 1.2% and 2.3% (Tables 2, S15), suggesting that many Z-W gametologs stopped recombining after species divergence. Although these inter-gametolog values are small, they are higher in the SLRs than in the PARs in *S. babylonica* ($P < 0.001$, Wilcoxon test), *S. chienii* ($P < 0.001$), and *S. matsudana* ($P < 0.001$), but not in *S. pentandra* ($P = 0.251$). This supports the inference above that the SLR in *S. pentandra* originated more recently (Figs S20, S21; Tables 2, S19, S20).

Despite incomplete diploidization, we could investigate specific gene losses and gains in the 15Z- and 15W-SLR using the autosome 15a sequence of *S. dunnii* as an outgroup to infer changes. In all 4*n* species, the proportions of genes lost are higher in the SLRs than in the PARs. Surprisingly, however, the numbers are almost always smaller for the W- than the Z-SLR. By contrast, W-SLRs gained more genes than Z-SLRs, except in *S. pentandra*, whose W-SLRs have evolved independently and more recently (column F of Table S21). A similar asymmetry occurs between the 15X- and 15Y-SLRs of *S. interior* and other diploid *Salix* species (He *et al.*, 2024; Wang *et al.*, 2024b, 2026). The consistently greater gene losses from Z- than W-linked regions (inferred, respectively, to have evolved from Y- and X-linked regions) may reflect the greater ancestral losses expected for Y- than X-linked regions.

Sex-linked gene expression and the evolution of sex-determining factors in willows

We next searched the SLRs for candidate femaleness-promoting genes by comparing gene expression between the two sexes, using flower buds at an early developmental stage, when sex-determining genes are expected to be expressed, based on evidence shown in Fig. S22. Many 15W-SLR genes have female-biased expression, but some are probably W-SLR-specific. Even after excluding such genes, female-biased expression was significantly more common in SLR than PAR genes (Fisher's exact tests), and possibly reflects feminization of genes in the W-SLR (panel E of Table S22; Fig. S23).

Because a new SLR appears to have evolved in *S. pentandra*, we examined all female-biased genes (including those specific to the W-SLR) shared by *S. babylonica*, *S. chienii*, and *S. matsudana* (Table S23). Functional annotations of their *A. thaliana* orthologs identified only three genes likely to be associated with flowering regulation: *SCD1*, *BRC1*, and *ARP4* (Falbel *et al.*, 2003; Kandasamy *et al.*, 2005; Feng *et al.*, 2022). However, their functions in these willows remain unknown.

Note, however, that feminization could be promoted by changes in other genomic regions, and that the polyploids might not have evolved new female-determining factors. We explained

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

New Phytologist (2026)
www.newphytologist.com

Table 3 The chromosomal distribution of sex-determining candidate sequences in *Salix* species.

	Genes and copy numbers on the chromosomes indicated. For species with male heterogamety, the table indicates that the Y-linked region carries the gene(s)																	
	ARR22			Intact *ARR17* duplicates			Partial *ARR17* duplicates			*MSF*: male-specific factor			Intact *PI*		Partial *PI*		*GATA15*	
Taxon and chromosome	15W	15Z	3	15W	15Z	19	7	15W	15Z	15W	15Z	7	2	5	15 W	15Z	15W	15Z
Salix gordejevii	2	0	10	0	0	1	0	2	0	1	0	0	2	2	5	0	0	0
Salix purpurea	1	1	10	4	0	2	0	0	7	1	0	0	1	1	2	0	1	0
Salix baileyi	3	0	10	1	0	2	0	0	3	1	0	0	2	2	2	0	0	0
Salix polyclona – W1	6	1	10	2	2	4	0	7	7	2	2	0	2	2	4	0	0	0
S. polyclona – E	5	1	10	2	2	4	0	6	4	2	1	0	3	2	9	0	0	0
S. polyclona – E – TS	7	2	10	14	2	4	0	0	4	0	1	0	2	2	9	0	0	0
S. polyclona – W2 (autotetraploid)	4	1	20	15	5	6	0	0	16	0	6	0	5	4	3	0	0	0
4*n* allopolyploid species																		
S. babylonica	**3 (SLR)**	**0**	**19**	**0**	**0**	**8**	**0**	**0**	**4 (SLR)**	**0**	**2 (SLR)**	**0**	**4**	**4**	**0**	**0**	**0**	**0**
S. chienii	**2 (SLR)**	**0**	**20**	**0**	**0**	**12**	**0**	**0**	**4 (SLR)**	**0**	**2 (SLR)**	**0**	**4**	**4**	**0**	**0**	**0**	**0**
S. matsudana	**4 (SLR)**	**0**	**20**	**0**	**0**	**8**	**0**	**0**	**4 (SLR)**	**0**	**2 (SLR)**	**0**	**4**	**4**	**0**	**0**	**0**	**0**
S. pentandra	**3 (SLR)**	**0**	**20**	**0**	**0**	**6**	**0**	**4 (PAR)**	**4 (PAR)**	**2 (PAR)**	**2 (PAR)**	**0**	**4**	**4**	**0**	**0**	**0**	**0**
The above four 4*n* ZW allopolyploid species are shown separately from other species																		
S. interior	0	0	10	0	0	4	0	–	**4 (Y)**	–	2 (Y)	0	**2**	**2**	0	0	**0**	**0**
S. arbutifolia	0	0	10	0	0	4	0	–	7 (Y)	–	2 (Y)	0	2	2	0	0	0	0
S. triandra	0	0	8	0	0	4	0	–	7 (Y)	–	4 (Y)	0	2	2	0	0	0	0
S. mesnyi	0	0	10	0	0	4	4 (Y)	–	0	–	0	2 (Y)	2	2	0	0	0	0
S. dunnii	0	0	10	0	0	4	6 (Y)	–	0	–	0	2 (Y)	2	2	0	0	0	0

Note: Bold font indicates the species analysed in this study.

New Phytologist © 2026 New Phytologist Foundation. © 2026 The Author(s).

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

above that, compared with the XY ancestor, polyploids have a change in the dosage of intact *ARR17* genes on chromosome 19, relative to the partial copies on chromosome 15, because, in the ZW genotype, the partial duplicates are present only on the Z (He *et al.*, 2024). In the *S. babylonica* genome, for example, there are eight intact *ARR17*-like genes on the four copies of autosome 19, and four Z-linked partial duplicates arranged as two inverted repeats, derived from the ancestral V_Y (Tables S24–S26). *S. matsudana* is similar (Fig. 5a), whereas the *S. chienii* female genotype has a total of 12 intact copies, possibly further promoting femaleness. A phylogenetic tree based on exon 1 sequences (136 bp) of the partial and intact *ARR17* sequences, including those from *S. interior* (Fig. 5b; Table S27), indicated that the 15Z- and 15Y-linked partial duplicates all evolved from an ancestral *Salix* clade chromosome 19 sequence.

We confirmed that mRNA-seq reads from exon 1 of the putatively female-promoting intact *ARR17* genes on chromosome 19 in *S. chienii* and *S. matsudana* are indeed expressed at much lower levels in flower buds of ZZ males than in those of ZW females, with expression sometimes undetectable in males (Figs 5c,d, S24a, S25a). In male buds, sRNA abundance is higher for exon 1 of the partial duplicates on 15Z and from their upstream regions than in female buds (Figs S24b, S25b). sRNA reads can be mapped to the intact *ARR17* gene regions of both haplotypes of both homeologs of chromosome 19 in males of all three newly studied species, but not in females of the two species tested, *S. chienii* and *S. matsudana* (Figs 5c–e, S26).

Although the relationship between *ARR17* sequences and sex determination probably differs in *S. pentandra*, as explained below, all the newly studied 15ZW clade II species are similar to *S. babylonica* in showing the expected male-specific expression of orthologs of the downstream male-function *PISTILLATA* (*PI*) gene, with higher expression in male flower buds (Fig. 6a; Tables S28, S29). *PI* is important for the development of male floral organs in *A. thaliana*, and its suppression can convert male flowers into female ones in spinach (Sather *et al.*, 2010). In 15ZW clade I species, partial *PI* orthologs can also determine femaleness (Wang *et al.*, 2026).

The evidence for evolution of a new female-determining factor in *S. pentandra* is much stronger. First, it has only six intact *ARR17* copies, presumably reflecting the loss of a copy since the initial hybrid formed (Fig. 5a; Tables S24, S26). Second, the partial *ARR17* duplicates are found in both the Z and W haplotypes (Fig. 5a). They form a Y-Z clade with the *S. interior* 15Y-SLR (Fig. 5b), consistent with findings from the other species suggesting a shared ancestral origin (Wang *et al.*, 2024b), and with a proposed crossover event in the pericentromeric region of *S. pentandra* that transferred these sequences from the Z haplotype (the former Y) to the W haplotype (whose ancestral X carried no such sequences). The presence of partial *ARR17* duplicates in both sexes (in both the V_Y-derived V_Z and the V_W, see Fig. 1) suggests that the V_Z copies no longer function in sex determination. Indeed, although we observed sRNA accumulation derived from the partial *ARR17* duplicates in both sexes of this species, we detected almost no *ARR17* mRNA in flower buds of either sex (Figs 5e, S26, 27). To maintain a 1 : 1 sex ratio, *S. pentandra* would therefore be expected to have evolved a new female-determining factor on the $15V_W$ chromosome. We therefore searched for V_W-specific genes that were either expressed only in females or showed female-biased expression. One such candidate femaleness-promoting gene is homologous to a different *A. thaliana* RR gene, *ARR22* (Fig. 6a; Table S30).

All the 15ZW clade II *Salix* species studied here have *ARR22* copies on chromosome 3 (usually five copies on each homeolog), and all also have duplicates in their 15W-SLRs (Fig. 6b), No such copies are found on the sex chromosomes of the extant diploid 15XY and 7XY species, but duplicates are again present on the 15W chromosome in the diploid 15ZW clade I species (Table 3, Table S31). All the W-linked copies originated from the same chromosome 3 copy (chr03-*ARR22*_05 in Figs 6c, S28). Ancestral state reconstruction suggests that the 15W copies in both clades I and II were derived from their ancestral 15X *ARR22* copy (Fig. S29), and that the two clades changed independently from 15X to 15W. In the other 4*n* species, the W-linked *ARR22* copies were not detectably expressed (except for one copy in *S. matsudana*, Table S32), nor were most autosomal copies. We discuss the female-specific expression of the W-linked duplicate *ARR22* copy in *S. pentandra* below.

Discussion

The hybrid origin of 15ZW clade II allotetraploid willows and the change in heterogamety

The molecular basis of sex determination in *Salix* has mostly focused on diploid species (Wang *et al.*, 2022, 2024b; Hyden *et al.*, 2023), while among polyploids, the only allotetraploid studied so far is *S. babylonica* (He *et al.*, 2024). We confirmed that three further 15ZW clade II species also originated from hybridization between a *Salix* clade female and a *Vetrix* clade male ancestor more than 13 Mya (Table S15; Fig. 7), and unreduced gamete formation (Ramsey & Schemske, 1998) seems to have been involved. These species all have sex chromosomes derived from the $15V_X$ and $15V_Y$ chromosomes of the male ancestor, together with $7S_X$ chromosomes from the female ancestor (Figs 1, S30).

Our results suggest that the change to female heterogamety may have been a direct consequence of hybridization and polyploidization. Specifically, combining different sex-determining systems might initially have created a monoecious phenotype, as recently found in an F1 *S. alba* × *babylonica* hybrid (Alarcón-Bolaños *et al.*, 2025), and been followed by re-establishment of dioecy. During these events, the 15XY system from the *Vetrix* clade ancestor contributed the present sex chromosome pair, with only two sex chromosome types: V_Z derived from the former Y, and V_W derived from the former X. Thus, as previously proposed for *S. babylonica* (He *et al.*, 2024), the change to female heterogamety after hybridization between two species with male heterogamety may have been triggered by the lower dosage of maleness-promoting partial *ARR17* duplicates on the $15V_Z$-SLR relative to the at least eight intact *ARR17* copies on chromosome 19, which must presumably be suppressed in order for the male phenotype

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

New Phytologist

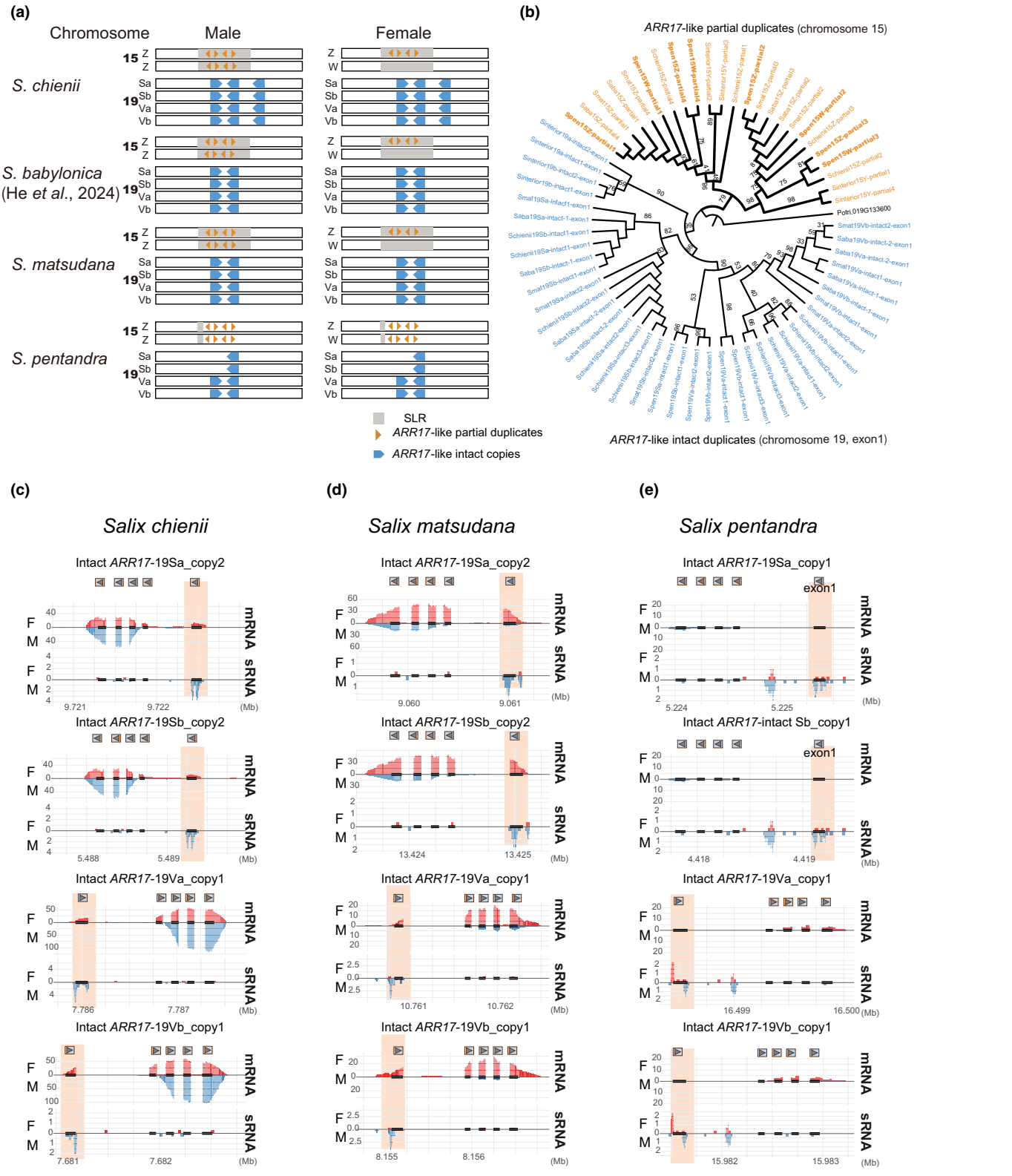


Fig. 5 Numbers, locations, and expression patterns of *ARR17*-like sequences in the tetraploid *Salix* 15ZW clade II lineage. (a) Locations and numbers of intact and partial copies. (b) Phylogenetic tree based on exon 1 of the *ARR17* copies, including the close outgroup diploid *Salix interior*. The tree was rooted using the intact exon 1 of the *Populus trichocarpa ARR17* gene (Potri.019G133600). The two colors show the chromosome 15 and 19 copies. Numbers marked on the tree represent bootstrap values. (c–e) mRNA-seq read estimates and sRNA enrichment pattern of intact *ARR17* on Chr19Sa, Chr19Sb, Chr19Va and Chr19Vb copies in female (red) and male flower (blue) buds of *S. chienii*, *S. matsudana* and *S. pentandra*. The shaded region indicates the location of the first exon. Results for other copies are shown in Supporting Information Figs S24–S27.

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

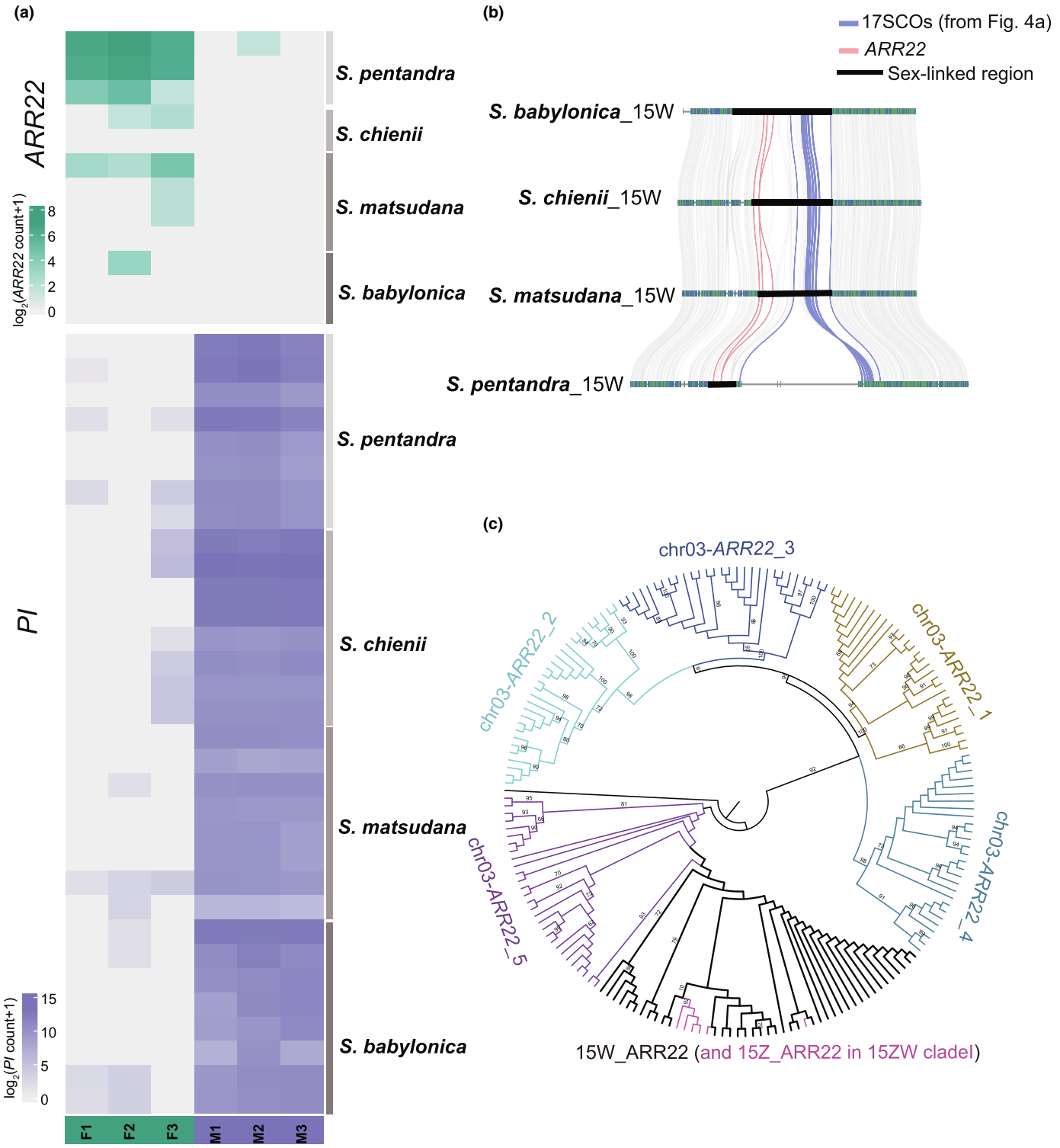


Fig. 6 Expression of the *ARR22* and *PISTILLATA* homologs in the tetraploid *Salix* 15ZW clade II lineage species. (a) Expression patterns of the *ARR22* (a candidate femaleness-determining factor) and *PISTILLATA* (a downstream gene involved in male development) in female and male flower buds (the columns labeled F1 to F3 and M1 to M3; see also Supporting Information Tables S29 and S32). (b) The positions of the *ARR22* genes on chromosome 15 relative to the single-copy ortholog genes shown in Fig. 4. (c) Phylogenetic tree of *ARR22* duplicates, showing the close relationship of the copies on chromosome 15 with one set of chromosome three sequences. Numbers on the tree represent bootstrap values.

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

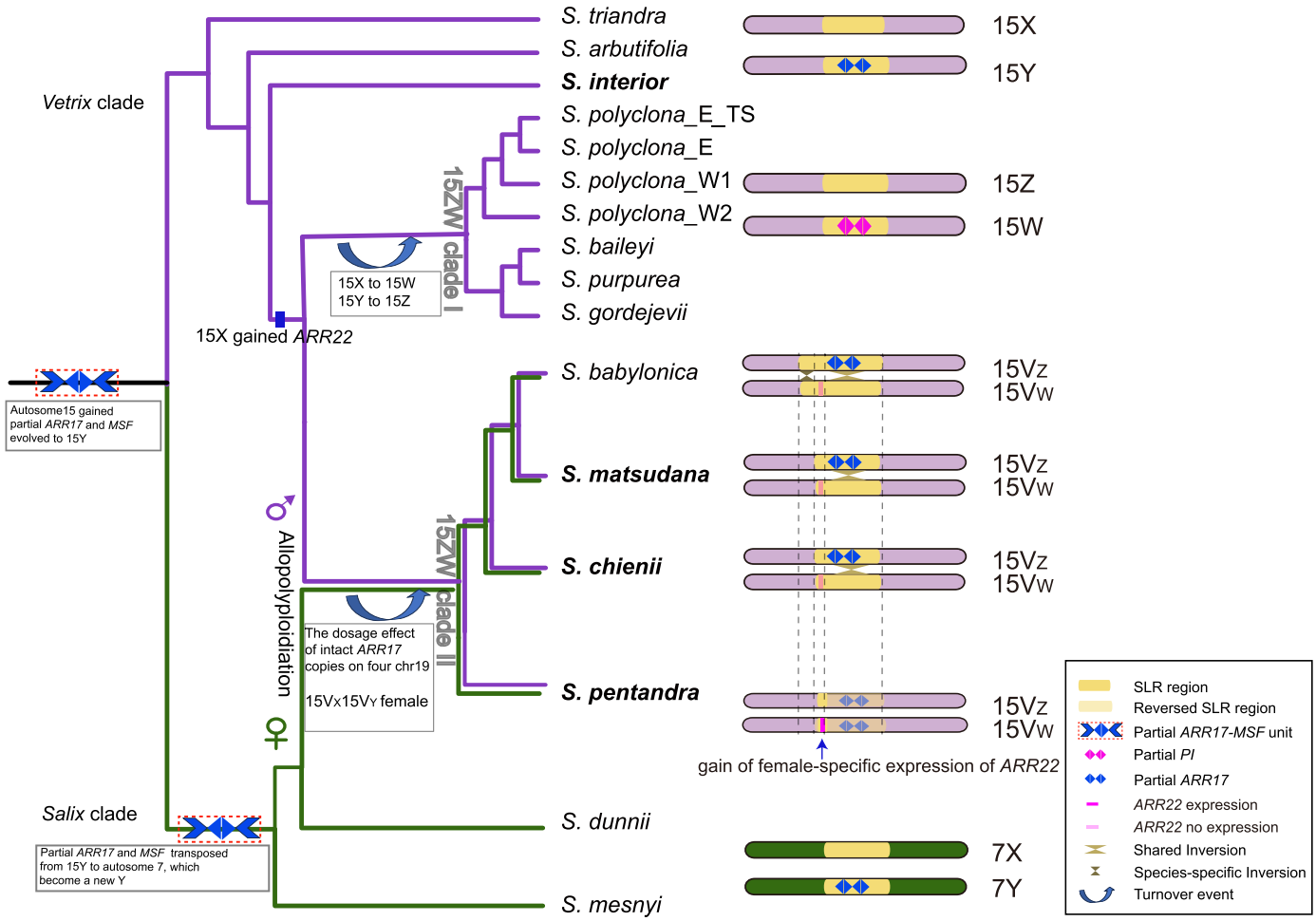


Fig. 7 Hypothesis for the sex determination systems in the *Salix* genus. The left side shows the phylogenetic relationships and the hybrid origin of the allotetraploid lineage. The right side shows the hypothesized female genotypes (males are $15V_ZV_Z$ homozygotes for the allotetraploid lineage) and sex-determining genes of the different clades. In 7XY (bottom) and 15XY (top) species, Y-linked partial *ARR17*-like duplicates are present in the Y-linked regions and can inhibit the intact *ARR17*-like copies on chromosome 19. In the 15ZW clade I species (diploids and autopolyploids), partial *PI*-like duplicates on chromosome 15 determine femaleness, and the system based on *ARR17*-like genes has been lost (Wang *et al*., 2026). The 15ZW allotetraploid lineage: (1) sex chromosome evolution in the 15ZW clade II, with sex-linked region (SLR) boundary changes driven by shared inversions in the ancestral lineage and species-specific inversions that expanded the SLR in *Salix babylonica*. The reversion restoring recombination leads to re-evolution of SLR in *Salix pentandra*. (2) *ARR22* expressed in *S. pentandra*: partial *ARR17*-like duplicates on both Z and W could suppress all intact *ARR17*-like copies on chromosome 19 in both males and females. *ARR22* is a strong candidate for the female-determining factor.

to develop. Such sensitivity to dosage differences resembles that suggested in swordtail fish (Volff & Schartl, 2001). Under this model, the 4*n* species' females have the $15V_Z/V_W$ genotype (derived from the $15V_X15V_Y$ genotype), whereas the former $7S_X$ chromosome has no effect on the sex phenotype and has become an autosome (Figs 1, S1, S30). Thus, a simpler model than that of He *et al*. (2024) is that the ancestral Y became converted into a Z, while the former X, which carried no female-suppressing partial *ARR17* copies and thus would not have required their loss or inactivation, automatically became a W. Dominance would also not need to evolve, as femaleness would automatically be dominant due to the presence of many intact *ARR17* copies on the chromosome 19 homeologs. An ancestral 4*n* species might thus have had a changed heterogametic sex without selection favoring the new system as proposed by Bull (1983). The further change, to a new female-determining system that is inferred in *S. pentandra*, is discussed below.

Recombination in the willow chromosome 15 SLRs

The SLRs of these species may not have evolved by recombination suppression, as they are at least partly within pericentromeric regions (Fig. 4c). These are (sometimes very extensive) regions with very low but nonzero recombination rates that flank centromeres (Mahtani & Willard, 1988; Barra & Fachinetti, 2018; Langley *et al*., 2019; Altemose *et al*., 2022). Like SLRs, they are repeat-rich and heterochromatic. Pericentromeric sex-determining regions (Charlesworth, 2019) have recently been found in various plants (She *et al*., 2025; Akagi *et al*., 2025a,b), including poplars (Xue *et al*., 2020; Li *et al*., 2023) and willows

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(He *et al.*, 2024; Wang *et al.*, 2024b, 2026), and the SLRs of the 4*n* *Salix* species are also repeat-rich and gene-poor. As rearrangements are frequent in pericentromeric regions (Barra & Fachinetti, 2018), the rearrangements, including inversions between these SLRs, may be either neutral consequences of being in a pericentromeric SLR or causes of low recombination rates. The very low W–Z *Ks* values in these regions are consistent with occasional recombination.

The inversion shared by *S. babylonica*, *S. matsudana*, and *S. chienii*, but not by *S. pentandra*, may, however, have suppressed exchanges, converting a rarely recombining pericentromeric region into one that is now completely W-linked. The few genes with sex-patterned gene trees in these three species, indicating sex-linkage that predates their divergence, are mostly within this inversion, and both likely male-promoting factors, the partial *ARR17* and *MSF* sequences, are also within or near it (Fig. 4a,b; Tables S16, S17). Consistent with the 'neutral' model above, the two W-specific genes within the inversion are not strong candidates for femaleness factors, as both are expressed at low levels (Table S18).

In *S. babylonica*, the further inversion extending the SLR into the adjacent syntenic region (Fig. 4) seems especially likely to have involved selection. The initial female $15V_Z/V_W$ genotype was derived from a $15V_X/15V_Y$ genotype (Fig. 1), which, as mentioned above, might have had poor female function or exhibited some intersex/male function. Under the dosage model explained above, a partially W-linked allele that increased female function at the expense of maleness might have appeared only in this species (though no candidate has been identified so far). Such an allele would have favored an inversion that restricted it to females.

A different female-determining system in *S. pentandra*

The *S. pentandra* W and Z SLRs appear to have evolved independently of those in the other 4*n* species. This species lacks inversions, suggesting that its pericentromeric region may still recombine occasionally, allowing the exchanges between the Z and W haplotypes detected by gene tree analysis (Fig. 4a), and explaining the presence of the usually Z- (and Y) specific partial *ARR17* and *MSF* sequences in both the V_W and V_Z haplotypes. The *S. pentandra* SLRs overlap the left-hand portions of its chromosome 15 inferred pericentromeric regions and are both repeat-rich (Fig. 4c; Table S14), but genes are still syntenic with those in the SLRs of the other three tetraploid species, as in *S. triandra*, whose 15XY SLRs are also only partly pericentromeric (Wang *et al.*, 2024b). The especially low divergence between W- and Z-linked genes in *S. pentandra* (Table 2; Fig. S20) suggests a recent loss of recombination, not an ancestral state from which the other species evolved, even though this is an early-branching species. Intriguingly, *S. pentandra*'s new, very small, fully sex-linked region (Table S14) carries three W-specific copies of a sequence with 91–97% protein sequence identity to the autosome 3 gene, *ARR22*. This gene encodes a type-C cytokinin response regulator with no known involvement in *A. thaliana* flower development: its homolog regulates expression of multiple genes involved in abiotic stress responses, including drought and freezing tolerance (Kang *et al.*, 2013), and probably also participates in cytokinin-responsive phosphorelay signal transduction (Kiba *et al.*, 2004). However, its ortholog in kiwifruit (Fig. S31) is the Y-linked Shy Girl (*SyGl*) gene, which acts as a *SuF* factor suppressing female and promoting male functions (Akagi *et al.*, 2018). Under the model that *SuF* factors are sexually antagonistic, with alternative alleles promoting female functions at the expense of male ones, as part of a 'trade-off' system in flower development (Charlesworth & Charlesworth, 1978), such W-linked alleles in *S. pentandra* might therefore promote female functions. The coding regions of most *ARR22* copies in *Salix* are longer than those of *SyGl*, and *ARR22* may have gained new domains, potentially increasing female-promoting activity. The situation may resemble the evolution of the African clawed frog *DM-W* from *DMRT1*; these genes have high sequence identity in their DNA-binding domains, but dissimilar *trans*-activation domains, and *DM-W* has novel functions (Yoshimoto *et al.*, 2008; Bewick *et al.*, 2011).

Conclusions

Taken together, our results suggest that a 4*n* ancestor may have had a male-biased sex ratio because of the presence of a Z chromosome in females, which may still have properties inherited from its Y ancestor, and they suggest that this situation probably led to different evolutionary changes in the different 4*n* species. In *S. pentandra*, W-specific *ARR22* duplications creating female-biased expression may have restored an equal sex ratio in *S. pentandra*, while in the other species the dosage of maleness-promoting partial *ARR17* duplicates on the former Y (which evolved into the 4*n* species' Z), relative to that of intact autosomal copies, has allowed maintenance of both sexes, but may also have changed in response to a dosage of autosomal copies insufficient for reliable female development (Fig. S30). The W haplotypes of the latter species have evolved independently, with one shared inversion that is lacking in *S. pentandra*, and one inversion unique to *S. babylonica*. These inversions might have been favored because they prevent maleness-promoting alleles on the Z (inherited from its former haplotype, the V_Y) from recombining onto the W, or might simply reflect the tendency of pericentromeric regions to rearrange. We also cannot exclude the possibility that femaleness-promoting alleles may have evolved in genes in the *S. babylonica* W-SLR, and even in those of the other species.

Acknowledgements

We are grateful to Wendy L. Applequist, Wan-Cheng Hou, Yi-Jiao Wang, Li Tu, Gang Liu, Guang-Nan Gong, Wen Shao, Lu-Yao Li, and Xi Chen for their kind help during the preparation of this manuscript. This study received financial support from the National Natural Science Foundation of China (grant no. 32500303 to Z-QX, 32171813 to LH) and the Special Fund for Scientific Research of the Shanghai Landscaping & City Appearance Administrative Bureau (grant nos. G232403, G242417, G252403 to LH and G252411 to Z-QX).

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

Competing interests

None declared.

Author contributions

Conceptualization: LH and DC. Project administration: LH and Z-QX. Methodology: LH, Z-QX, X-YL, Yuan Wang and R-GZ. Investigation: LH and Z-QX. Lab work: JD, Z-QX and Z-YZ. Formal analysis: LH, Z-QX and X-YL, and Yuan W. Writing – original draft: Z-QX and L.H. Writing – review and editing: LH, DC, EH, Z-QX and Yi Wang.

ORCID

Deborah Charlesworth https://orcid.org/0000-0002-3939-9122
Jie Ding https://orcid.org/0009-0006-2282-5902
Elvira Hörandl https://orcid.org/0000-0002-7600-1128
Li He https://orcid.org/0000-0002-4591-8056
Xin-Yu Liu https://orcid.org/0009-0007-7954-5159
Yi Wang https://orcid.org/0000-0001-6611-5319
Yuan Wang https://orcid.org/0009-0008-3300-9366
Zhi-Qing Xue https://orcid.org/0000-0001-9219-9894
Ren-Gang Zhang https://orcid.org/0000-0002-8028-9208
Zhi-Ying Zhu https://orcid.org/0000-0003-2824-7139

Data availability

The genome assembly sequences and resequencing datasets have been deposited in the National Genomics Data Center (NGDC, https://ngdc.cncb.ac.cn/bioproject/) with the BioProject number: PRJCA055340 (*Salix chienii*, *S. matsudana*, *S. pentandra*, and *S. interior*).

References

Akagi T, Fujita N, Shirasawa K, Tanaka H, Nagaki K, Masuda K, Horiuchi A, Kuwada E, Kawai K, Kunou R *et al.* 2025a. Rapid and dynamic evolution of a giant Y chromosome in *Silene latifolia*. *Science* **387**: 637–643.

Akagi T, Henry IM, Ohtani H, Morimoto T, Beppu K, Kataoka I, Tao R. 2018. A Y-encoded suppressor of feminization arose via lineage-specific duplication of a cytokinin response regulator in kiwifruit. *Plant Cell* **30**: 780–795.

Akagi T, Henry IM, Tao R, Comai L. 2014. A Y-chromosome–encoded small RNA acts as a sex determinant in persimmons. *Science* **346**: 646–650.

Akagi T, Segawa T, Uchida R, Tanaka H, Shirasawa K, Yamagishi N, Yaegashi H, Natsume S, Takagi H, Abe A *et al.* 2025b. Evolution and functioning of an X-A balance sex-determining system in hops. *Nature Plants* **11**: 1339–1352.

Alarcón-Bolaños P, Pittet L, He L, Hörandl E. 2025. Hybridization resulted in shifts from dioecy to monoecy in weeping willows (*Salix* L.). *Genes* **16**: 958.

Altemose N, Logsdon GA, Bzikadze AV, Sidhwani P, Langley SA, Caldas GV, Hoyt SJ, Uralsky L, Ryabov FD, Shew CJ *et al.* 2022. Complete genomic and epigenetic maps of human centromeres. *Science* **376**: eabl4178.

Barra V, Fachinetti D. 2018. The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nature Communications* **9**: 4340.

Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. 2012. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* **58**: 268–276.

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27**: 573–580.

Berlin S, Fogelqvist J, Lascoux M, Lagercrantz U, Rönnberg-Wästljung AC. 2011. Polymorphism and divergence in two willow species, *Salix viminalis* L. and *Salix schwerinii* E. Wolf. *G3: Genes, Genomes, Genetics* **1**: 387–400.

Bewick AJ, Anderson DW, Evans BJ. 2011. Evolution of the closely related, sex-related genes *DM-W* and *DMRT1* in African clawed frogs (*Xenopus*). *Evolution* **65**: 698–712.

Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nature Methods* **12**: 59–60.

Bull JJ. 1983. *Evolution of sex determining mechanisms*. Menlo Park, CA, USA: The Benjamin/Cummings Publishing Co.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**: 540–552.

Charlesworth B, Charlesworth D. 1978. A model for the evolution of dioecy and gynodioecy. *The American Naturalist* **112**: 975–997.

Charlesworth D. 2019. Young sex chromosomes in plants and animals. *New Phytologist* **224**: 1095–1107.

Chen C, Li J, Feng J, Liu B, Feng L, Yu X, Li G, Zhai J, Meyers BC, Xia R. 2021. sRNAanno – a database repository of uniformly annotated small RNAs in plants. *Horticulture Research* **8**: 45.

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

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. 2021. Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nature Methods* **18**: 170–175.

Cuñado N, Navajas-Pérez R, de la Herrán R, Ruiz Rejón C, Ruiz Rejón M, Santos JL, Garrido-Ramos MA. 2007. The evolution of sex chromosomes in the genus *Rumex* (Polygonaceae): identification of a new species with heteromorphic sex chromosomes. *Chromosome Research* **15**: 825–833.

De Storme N, Geelen D. 2013. Sexual polyploidization in plants – cytological mechanisms and molecular regulation. *New Phytologist* **198**: 670–684.

Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP *et al.* 2017. *De novo* assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **356**: 92–95.

Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. 2016a. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Systems* **3**: 99–101.

Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, Aiden EL. 2016b. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Systems* **3**: 95–98.

Emms DM, Kelly S. 2019. ORTHOFINDER: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**: 238.

Falbel TG, Koch LM, Nadeau JA, Segui-Simarro JM, Sack FD, Bednarek SY. 2003. *SCD1* is required for cytokinesis and polarized cell expansion in *Arabidopsis thaliana* [corrected]. *Development* **130**: 4011–4024.

Feng J, Deng Q, Lu H, Wei D, Wang Z, Tang Q. 2022. *Brassica juncea BRC1*-1 induced by SD negatively regulates flowering by directly interacting with BjuFT and *BjuFUL* promoter. *Frontiers in Plant Science* **13**: 986811.

García-Moreno J, Mindell DP. 2000. Rooting a phylogeny with homologous genes on opposite sex chromosomes (Gametologs): a case study using Avian CHD. *Molecular Biology and Evolution* **17**: 1826–1832.

Gerchen JF, Veltsos P, Pannell JR. 2022. Recurrent allopolyploidization, Y-chromosome introgression and the evolution of sexual systems in the plant genus *Mercurialis*. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **377**: 20210224.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**: 644–652.

Gulyaev S, Cai X-J, Guo F-Y, Kikuchi S, Applequist WL, Zhang Z-X, Hörandl E, He L. 2022. The phylogeny of *Salix* revealed by whole genome re-sequencing suggests different sex-determination systems in major groups of the genus. *Annals of Botany* **129**: 485–498.

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK Jr, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD *et al.* 2003. Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Research* **31**: 5654–5666.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. *Genome Biology* **9**: R7.

Hall AB, Qi Y, Timoshevskiy V, Sharakhova MV, Sharakhov IV, Tu Z. 2013. Six novel Y chromosome genes in *Anopheles mosquitoes* discovered by independently sequencing males and females. *BMC Genomics* **14**: 273.

Handley LJ, Ceplitis H, Ellegren H. 2004. Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. *Genetics* **167**: 367–376.

He L, Guo FY, Cai XJ, Chen HP, Lian CL, Wang Y, Shang C, Zhang Y, Wagner ND, Zhang ZX *et al.* 2023. Evolutionary origin and establishment of a dioecious diploid-tetraploid complex. *Molecular Ecology* **32**: 2732–2749.

He L, Hörandl E. 2022. Does polyploidy inhibit sex chromosome evolution in angiosperms? *Frontiers in Plant Science* **13**: 976765.

He L, Jia KH, Zhang RG, Wang Y, Shi TL, Li ZC, Zeng SW, Cai XJ, Wagner ND, Hörandl E *et al.* 2021a. 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, Wagner ND, Hörandl E. 2021b. 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.

He L, Wang Y, Wang Y, Zhang R-G, Wang Y, Hörandl E, Ma T, Mao YF, Mank JE, Ming R. 2024. Allopolyploidization from two dioecious ancestors leads to recurrent evolution of sex chromosomes. *Nature Communications* **15**: 6893.

Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**: 491.

Hu N, Sanderson BJ, Guo M, Feng G, Gambhir D, Hale H, Wang D, Hyden B, Liu J, Smart LB *et al.* 2023. Evolution of a ZW sex chromosome system in willows. *Nature Communications* **14**: 7144.

Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P. 2017. Fast genome-wide functional annotation through orthology assignment by EGGNOG-MAPPER. *Molecular Biology and Evolution* **34**: 2115–2122.

Hyden B, Zou J, Wilkerson DG, Carlson CH, Robles AR, DiFazio SP, Smart LB. 2023. Structural variation of a sex-linked region confers monoecy and implicates *GATA15* as a master regulator of sex in *Salix purpurea*. *New Phytologist* **238**: 2512–2523.

Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 2020. GETORGANELLE: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biology* **21**: 241.

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G *et al.* 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**: 1236–1240.

Kandasamy MK, Deal RB, McKinney EC, Meagher RB. 2005. Silencing the nuclear actin-related protein *AtARP4* in Arabidopsis has multiple effects on plant development, including early flowering and delayed floral senescence. *The Plant Journal* **41**: 845–858.

Kang NY, Cho C, Kim J. 2013. Inducible expression of *Arabidopsis response regulator 22* (*ARR22*), a type-C ARR, in transgenic *Arabidopsis* enhances drought and freezing tolerance. *PLoS ONE* **8**: e79248.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Kiba T, Aoki K, Sakakibara H, Mizuno T. 2004. *Arabidopsis* response regulator, *ARR22*, ectopic expression of which results in phenotypes similar to the *wol* cytokinin-receptor mutant. *Plant & Cell Physiology* **45**: 1063–1077.

Killick R, Eckley IA. 2014. CHANGEPOINT: an R package for changepoint analysis. *Journal of Statistical Software* **58**: 1–19.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**: 357–360.

Kirkpatrick M, Sardell JM, Pinto BJ, Dixon G, Peichel CL, Schartl M. 2021. Evolution of the canonical sex chromosomes of the guppy and its relatives. *G3: Genes, Genomes, Genetics* **12**: jkab435.

Korf I. 2004. Gene finding in novel genomes. *BMC Bioinformatics* **5**: 59.

Langley SA, Miga KH, Karpen GH, Langley CH. 2019. Haplotypes spanning centromeric regions reveal persistence of large blocks of archaic DNA. *eLife* **8**: e42989.

Leite Montalvão AP, Kersten B, Kim G, Fladung M, Müller NA. 2022. *ARR17* controls dioecy in *Populus* by repressing B-class MADS-box gene expression. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **377**: 20210217.

Li G, Chen C, Chen P, Meyers BC, Xia R. 2024. sRNAminer: a multifunctional toolkit for next-generation sequencing small RNA data mining in plants. *Science Bulletin* **69**: 784–791.

Li W, Fu W, Hou J, Yang Y, Yin T. 2023. Evolution of plant sex and molecular mechanisms underlying plants sex separation. *Forest Research* **3**: 1.

Lin Y, Ye C, Li X, Chen Q, Wu Y, Zhang F, Pan R, Zhang S, Chen S, Wang X *et al.* 2023. QUARTET: a telomere-to-telomere toolkit for gap-free genome assembly and centromeric repeat identification. *Horticulture Research* **10**: uhad12.

Liu X-Y, Zhang Z-X, Cheng Y-J, Xue Z-Q, He L. 2026. Chromosome-scale assembly of *Salix integra* reveals a conserved sex-determination mechanism in 15ZW clade I willows. *Journal of Systematics and Evolution* **n/a**: jse.70081.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESEQ2. *Genome Biology* **15**: 550.

Lowe TM, Eddy SR. 1997. TRNASCAN-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**: 955–964.

Maddison WP, Maddison DR. 2007. *MESQUITE: a modular system for evolutionary analysis*. Version 2.3. [WWW document] URL https://www.mesquiteproject.org.

Mahtani MM, Willard HF. 1988. A primary genetic map of the pericentromeric region of the human X chromosome. *Genomics* **2**: 294–301.

Masuda K, Ikeda Y, Matsuura T, Kawakatsu T, Tao R, Kubo Y, Ushijima K, Henry IM, Akagi T. 2022. Reinvention of hermaphroditism via activation of a *RADIALIS*-like gene in hexaploid persimmon. *Nature Plants* **8**: 217–224.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**: 1530–1534.

Müller NA, Kersten B, Leite Montalvão AP, Mähler N, Bernhardsson C, Bräutigam K, Carracedo Lorenzo Z, Hoenicka H, Kumar V, Mader M *et al.* 2020. A single gene underlies the dynamic evolution of poplar sex determination. *Nature Plants* **6**: 630–637.

Naish M, Henderson IR. 2024. The structure, function, and evolution of plant centromeres. *Genome Research* **34**: 161–178.

Ogutcen E, de Lima Ferreira P, Wagner ND, Marinček P, Vir Leong J, Aubona G, Cavender-Bares J, Michálek J, Schroeder L, Sedio BE *et al.* 2024. Phylogenetic insights into the Salicaceae: the evolution of willows and beyond. *Molecular Phylogenetics and Evolution* **199**: 108161.

Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, Lugo CSB, Elliott TA, Ware D, Peterson T *et al.* 2019. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biology* **20**: 275.

Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. 2015. STRINGTIE enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology* **33**: 290–295.

Pryszcz LP, Gabaldón T. 2016. REDUNDANS: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Research* **44**: e113.

Pucholt P, Hallingbäck HR, Berlin S. 2017. Allelic incompatibility can explain female biased sex ratios in dioecious plants. *BMC Genomics* **18**: 251.

Ramsey J, Schemske DW. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* **29**: 467–501.

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Sather DN, Jovanovic M, Golenberg EM. 2010. Functional analysis of B and C class floral organ genes in spinach demonstrates their role in sexual dimorphism. *BMC Plant Biology* **10**: 46.

She H, Liu Z, Xu Z, Zhang H, Wu J, Cheng F, Wang X, Qian W. 2025. A complete telomere-to-telomere assembly of the *Spinacia oleracea* genome reveals Y chromosome evolution and centromere landscape. *Plant Communications* **6**: 101410.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210–3212.

Smirnov V, Warnow T. 2020. MAGUS: multiple sequence alignment using graph clUStering. *Bioinformatics* **37**: 1666–1672.

Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics* **24**: 637–644.

Tang H, Wang X, Bowers JE, Ming R, Alam M, Paterson AH. 2008. Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. *Genome Research* **18**: 1944–1954.

Twyford AD, Conover JL, Doyle JJ, Mason AS, Soltis DE, Soltis PS, Wendel JF. 2025. The polyploid continuum and the landscape of polyploid genomic variation. *American Journal of Botany* **112**: e70121.

Volff JN, Schartl M. 2001. Variability of genetic sex determination in poeciliid fishes. *Genetica* **111**: 101–110.

Wagner ND, He L, Hörandl E. 2021. The evolutionary history, diversity, and ecology of Willows (*Salix* L.) in the European Alps. *Diversity* **13**: 146.

Wang D, Li Y, Li M, Yang W, Ma X, Zhang L, Wang Y, Feng Y, Zhang Y, Zhou R *et al.* 2022. Repeated turnovers keep sex chromosomes young in willows. *Genome Biology* **23**: 200.

Wang Y, Gong G-N, Wang Y, Zhang R-G, Hörandl E, Zhang Z-X, Charlesworth D, He L. 2024a. Gap-free X and Y chromosome assemblies of *Salix arbutifolia* reveal an evolutionary change from male to female heterogamety in willows, without a change in the position of the sex-determining locus. *New Phytologist* **242**: 2872–2887.

Wang Y, Xue Z-Q, Zhang R-G, Zhu Z-Y, Hörandl E, Wang X-R, Mao Y-F, Charlesworth D, He L. 2026. Recurrent sex chromosome turnover mediated by distinct *ARR17* and *PISTILLATA* duplications in willows. *Genome Biology* **27**: 137.

Wang Y, Zhang R-G, Hörandl E, Zhang Z-X, Charlesworth D, He L. 2024b. Evolution of sex-linked genes and the role of pericentromeric regions in sex chromosomes: insights from diploid willows. *Molecular Biology and Evolution* **41**: msae235.

Wang Y, Zhao L, Wang D, Chen K, Luo T, Luo J, Jiang C, He Z, Huang H, Xie J *et al.* 2025. Four near-complete genome assemblies reveal the landscape and evolution of centromeres in Salicaceae. *Genome Biology* **26**: 111.

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

Xue L, Wu H, Chen Y, Li X, Hou J, Lu J, Wei S, Dai X, Olson MS, Liu J *et al.* 2020. Evidences for a role of two Y-specific genes in sex determination in *Populus deltoides*. *Nature Communications* **11**: 5893.

Xue Z-Q, Applequist WL, Hörandl E, He L. 2024. Sex chromosome turnover plays an important role in the maintenance of barriers to post-speciation introgression in willows. *Evolution Letters* **8**: 467–477.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**: 1586–1591.

Yoshimoto S, Okada E, Umemoto H, Tamura K, Uno Y, Nishida-Umehara C, Matsuda Y, Takamatsu N, Shiba T, Ito M. 2008. A W-linked DM-domain gene, *DM-W*, participates in primary ovary development in *Xenopus laevis*. *Proceedings of the National Academy of Sciences, USA* **105**: 2469–2474.

Zhang J, Yuan H, Li Y, Chen Y, Liu G, Ye M, Yu C, Lian B, Zhong F, Jiang Y *et al.* 2020. Genome sequencing and phylogenetic analysis of allotetraploid *Salix matsudana* Koidz. *Horticulture Research* **7**: 201.

Zhang R-G, Li G-Y, Wang X-L, Dainat J, Wang Z-X, Ou S, Ma Y. 2022. TESORTER: an accurate and fast method to classify LTR-retrotransposons in plant genomes. *Horticulture Research* **9**: uhac017.

Zhong CX, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, Birchler JA, Jiang J, Dawe RK. 2002. Centromeric retroelements and satellites interact with Maize kinetochore protein CENH3. *Plant Cell* **14**: 2825–2836.

Zhou R, Macaya-Sanz D, Carlson CH, Schmutz J, Jenkins JW, Kudrna D, Sharma A, Sandor L, Shu S, Barry K *et al.* 2020. A willow sex chromosome reveals convergent evolution of complex palindromic repeats. *Genome Biology* **21**: 38.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Possible evolutionary origin of the clade II 15ZW polyploids, adapted from the interpretation of the *Salix babylonica* case in supplementary fig. 12 of He *et al.* (2024).

Fig. S2 One possible evolutionary pathway for the evolution of the clade II 15ZW polyploids, with diploid hybrids having formed first.

Fig. S3 Species tree based on single-copy orthologs present on all the chromosome 19 homeologs.

Fig. S4 Identification of the sex chromosome in the mapping population samples of *Salix interior*.

Fig. S5 The change-point signals of chromosome quotient, F_{ST}, and *k-mer*, gene, and repeat densities (including centromeric retrotransposon in maize) on the 15 chromosomes of *Salix interior*.

Fig. S6 Identification of the sex chromosome in the mapping population samples of *Salix chienii*.

Fig. S7 The change-point signals of CQ and *k-mers* on the 15 chromosomes of *Salix chienii*.

Fig. S8 Identification of the sex chromosome in the mapping population samples of *Salix matsudana*.

Fig. S9 The change-point signals of CQ and *k-mer* on the 15 chromosomes of *Salix matsudana*.

Fig. S10 Identification of the sex chromosome in the mapping population samples of *Salix pentandra*.

Fig. S11 The change-point signals of CQ and *k-mer* on the 15 chromosomes of *Salix pentandra*.

Fig. S12 Gene synteny between the sex chromosomes in the four different 15ZW clade II tetraploid species studied here and a close diploid outgroup, *Salix interior*.

Fig. S13 Gene synteny between the sex-linked regions in the four 15ZW clade II tetraploid species studied here.

Fig. S14 The phylogenetic relationships among *Salix* species reconstructed from ancestral sex-linked gene (Saarb15b0044000) proposed by He *et al.* (2024) and Wang *et al.* (2024a,b).

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Fig. S15 Circos plot of all 76 *Salix chienii* chromosomes.

Fig. S16 Circos plot of 76 *Salix matsudana* chromosomes.

Fig. S17 Circos plot of *Salix pentandra* chromosomes.

Fig. S18 The gene and repeat densities (including centromeric retrotransposon in maize-like elements) on the sex chromosomes of *Salix babylonica*.

Fig. S19 Collinearity and genomic composition of 15 chromosomes (15Va/15W, 15Vb/15Z, 15Sa/autosome and 15Sb/autosome) in *Salix chienii*, *Salix matsudana*, and *Salix pentandra*.

Fig. S20 Divergence between Z- and W-linked coding sequences of 15ZW clade II 4*n* species, estimated as synonymous site divergence (*Ks*) values.

Fig. S21 Estimated *Ks* values between gene pairs in the haplotypes in the chromosome 15 assemblies, for genes present in the W haplotype assemblies in the four 4*n* *Salix* species studied.

Fig. S22 Pistillate and staminate flower development in *Salix pentandra* and *Salix chienii* in longitudinal sections of the stages used for our samples for RNA-seq.

Fig. S23 Gene-biased expression estimates ($\log_2$ F/M ratios) among the sex chromosome.

Fig. S24 Expression patterns of *ARR17*-like duplicates in female and male flower buds of *Salix chienii*.

Fig. S25 Expression patterns of *ARR17*-like duplicates in female and male flower buds of *Salix matsudana*.

Fig. S26 Transcript level of partial *ARR17*-like duplicate sRNA and surrounding regions on autosome 19.

Fig. S27 Expression patterns of *ARR17*-like duplicates in female and male flower buds of *Salix pentandra*.

Fig. S28 The phylogenetic relationship of *ARR22* in *Salix* genus.

Fig. S29 Reconstructions of ancestral character states of *ARR22* on 15Z and 15W in willow species.

Fig. S30 Hypothesis for the sex determination systems and associated sex determining-genes in species of the 15ZW clade II.

Fig. S31 Protein sequences and exon base sequence alignments between the *SyGl* gene in kiwifruit (*Actinidia chinensis*) and *ARR22* in 15W-SLR *Salix* species in this study.

Methods S1 Genome sequencing, RNA sequencing, genome assembly and annotation, and identification of the sex-linked regions.

Table S1 Details of plant materials used for this study.

Table S2 Available genome datasets used in the paper.

Table S3 DNA sequencing information of four *Salix* species.

Table S4 Statistics of the four genome assemblies.

Table S5 Gap distribution in genome assemblies of *Salix* species in this study.

Table S6 Proportion of second- and third-generation sequencing reads mapped to the four studied genomes and their coverage percentages.

Table S7 The BUSCO results of *Salix* species in this study (protein).

Table S8 Distribution of protein-coding genes and RNAs in each region of the genome of four *Salix* species.

Table S9 Summary of annotated protein-coding genes in four *Salix* genomes.

Table S10 Function annotation of the predicted protein-coding genes of four *Salix* species.

Table S11 Summary of repeat content in the four *Salix* genomes.

Table S12 Transposable element content across the entire genomes of *Salix* species.

Table S13 Summary of whole-genome re-sequencing data of *Salix chienii*, *Salix matsudana*, and *Salix pentandra*.

Table S14 Statistics on the size and the proportion of repeated sequences of sex chromosomes in tetraploid willows.

Table S15 Synonymous-site divergence among genes in the four tetraploid *Salix* species and comparisons with divergence from the diploid outgroups, *Salix baileyi* and *Salix dunnii*.

Table S16 The locations of 51 single-copy orthologs shared across the 4*n* clade and *ARR17*-like partial duplicates.

Table S17 The locations of inversions, partial *ARR17*-like duplicates, and ancient sex-linked genes in the allotetraploid lineage.

Table S18 Shared W-SLR-specific genes (absent from the Z-SLR) across *Salix babylonica*, *Salix chienii*, and *Salix matsudana*, with their *Arabidopsis thaliana* homologs and functional annotations.

Table S19 Synonymous-site divergence estimates between gametologous gene pairs in 4*n* *Salix* species and diploid *S. interior*.

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

Table S20 Synonymous-site divergence between genes in the two haplotypes assembled in the four tetraploid *Salix* species.

Table S21 Numbers of sex chromosome genes, estimates of degeneration in terms of numbers of genes lost, and the proportions of gene losses from the regions of sex chromosomes in the *Salix* species studied.

Table S22 Estimated expression levels in flower buds of genes in sex-linked regions and PARs in the four tetraploid species.

Table S23 Genes exhibiting female-biased expression within the W-linked regions shared among *Salix babylonica*, *Salix chienii*, and *Salix matsudana*, with their *Arabidopsis thaliana* homologs and functional annotations.

Table S24 Identification of intact *ARR17*-like genes and *ARR17*-like partial duplicates in the *Salix chienii* genome.

Table S25 Identification of intact *ARR17*-like genes and *ARR17*-like partial duplicates in the *Salix matsudana* genome.

Table S26 Identification of intact *ARR17*-like genes and *ARR17*-like partial duplicates in the *Salix pentandra* genome.

Table S27 Identification of intact *ARR17*-like genes and *ARR17*-like partial duplicates in the *Salix interior* genome.

Table S28 Identification of *PI* genes in the studied *Salix* genomes.

Table S29 Expression patterns of the PI-like genes on chromosomes 2 and 5 in male and female flowering bud individuals of three 4*n* *Salix* species.

Table S30 Expression patterns of 17 W-specific genes and their homologous genes in male and female flowering bud individuals of *Salix pentandr*a.

Table S31 Identification of *ARR22* genes in the studied *Salix* genomes.

Table S32 Expression patterns of *ARR22*-like genes in male and female flowering bud individuals of three 4*n* 15ZW clade II species.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

Disclaimer: The New Phytologist Foundation remains neutral with regard to jurisdictional claims in maps and in any institutional affiliations.

© 2026 The Author(s).
New Phytologist © 2026 New Phytologist Foundation.

14698137, 0, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.71771 by Georg-August-Universitaet, Wiley Online Library on [20/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License