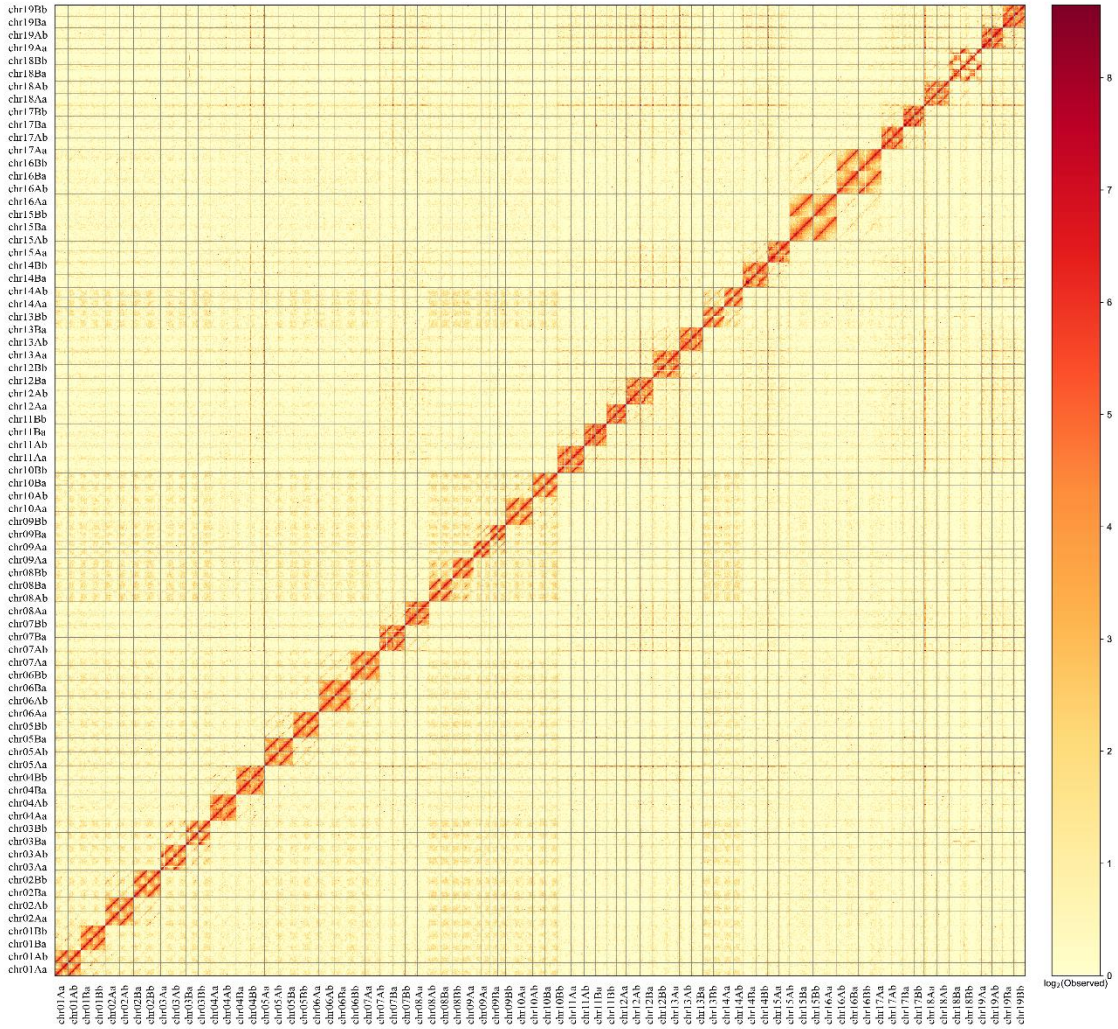
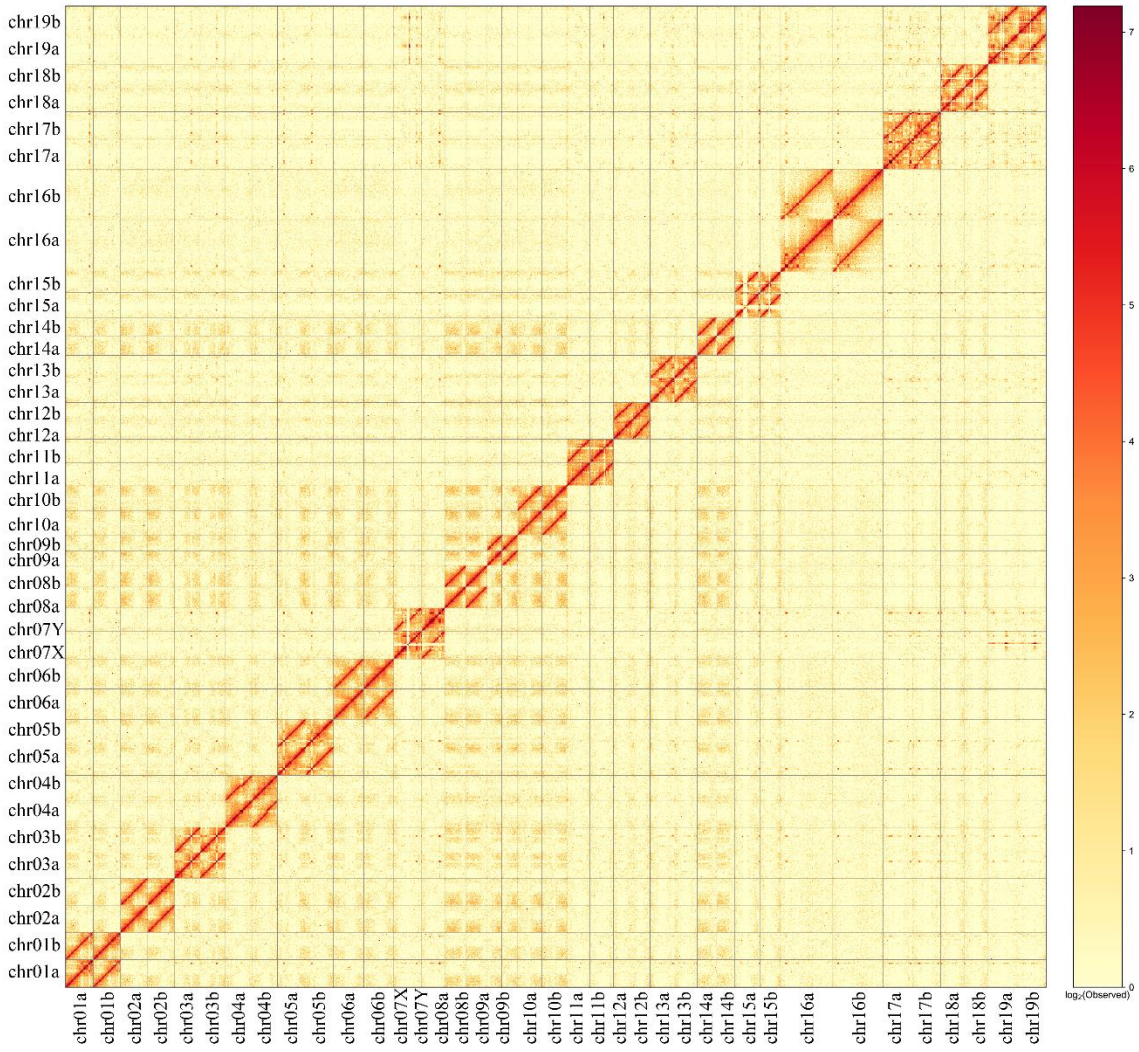
Allopolyploidization from two dioecious ancestors leads to recurrent evolution of sex chromosomes

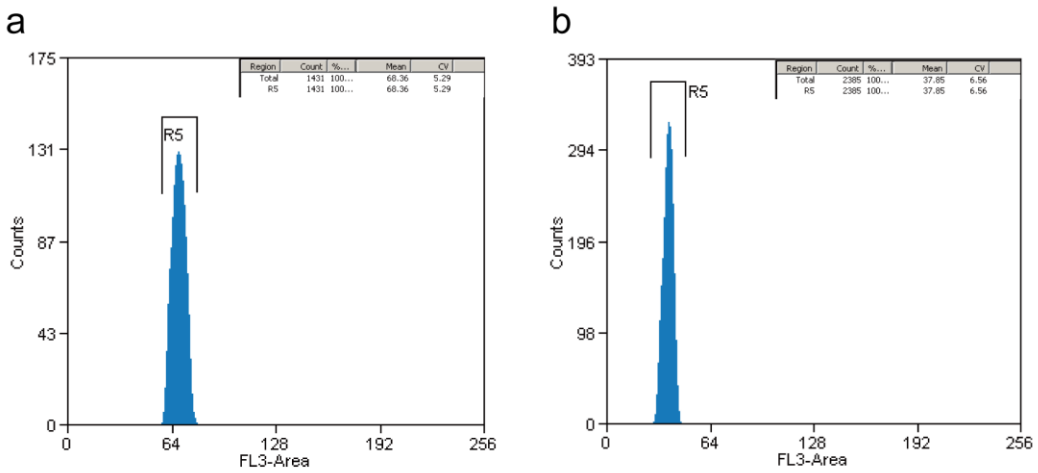
He *et al.*



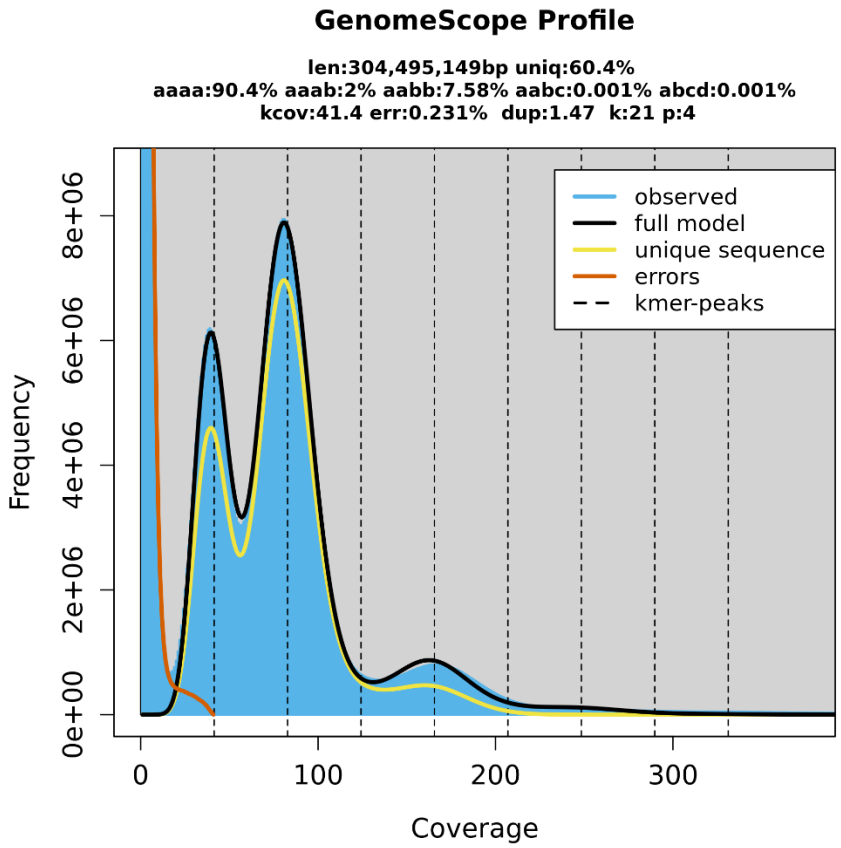
Supplementary Fig. 1. Genome-wide analysis of chromatin interactions in the *S. babylonica* genome based on Hi-C data. Aa=Sa, Ab=Sb, Ba=Va, Bb=Vb



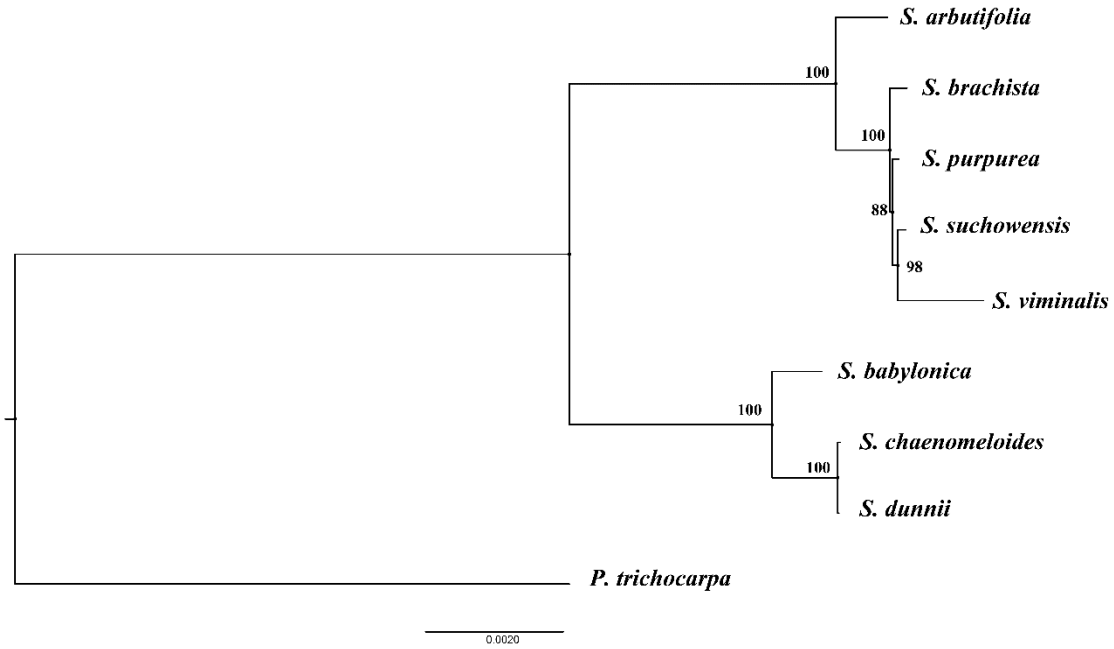
Supplementary Fig. 2. Genome-wide analysis of chromatin interactions in the *S. dunnii* genome based on Hi-C data.



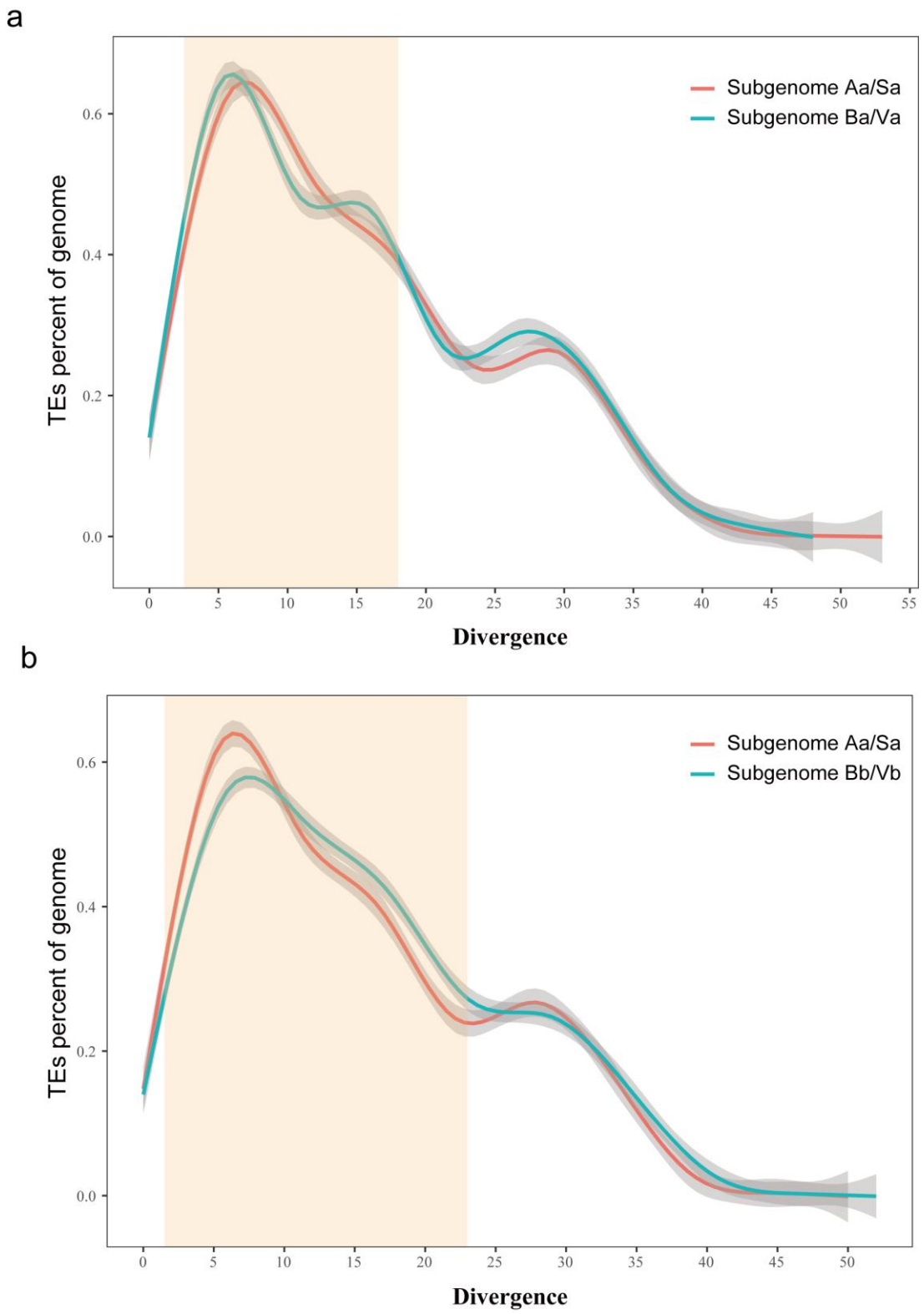
Supplementary Fig. 3. Flow cytometry histograms of willows. **a** *Salix babylonica* (Saba01F). **b** The Saba01F's external standard *S. dunnii* (FNU-M-1).



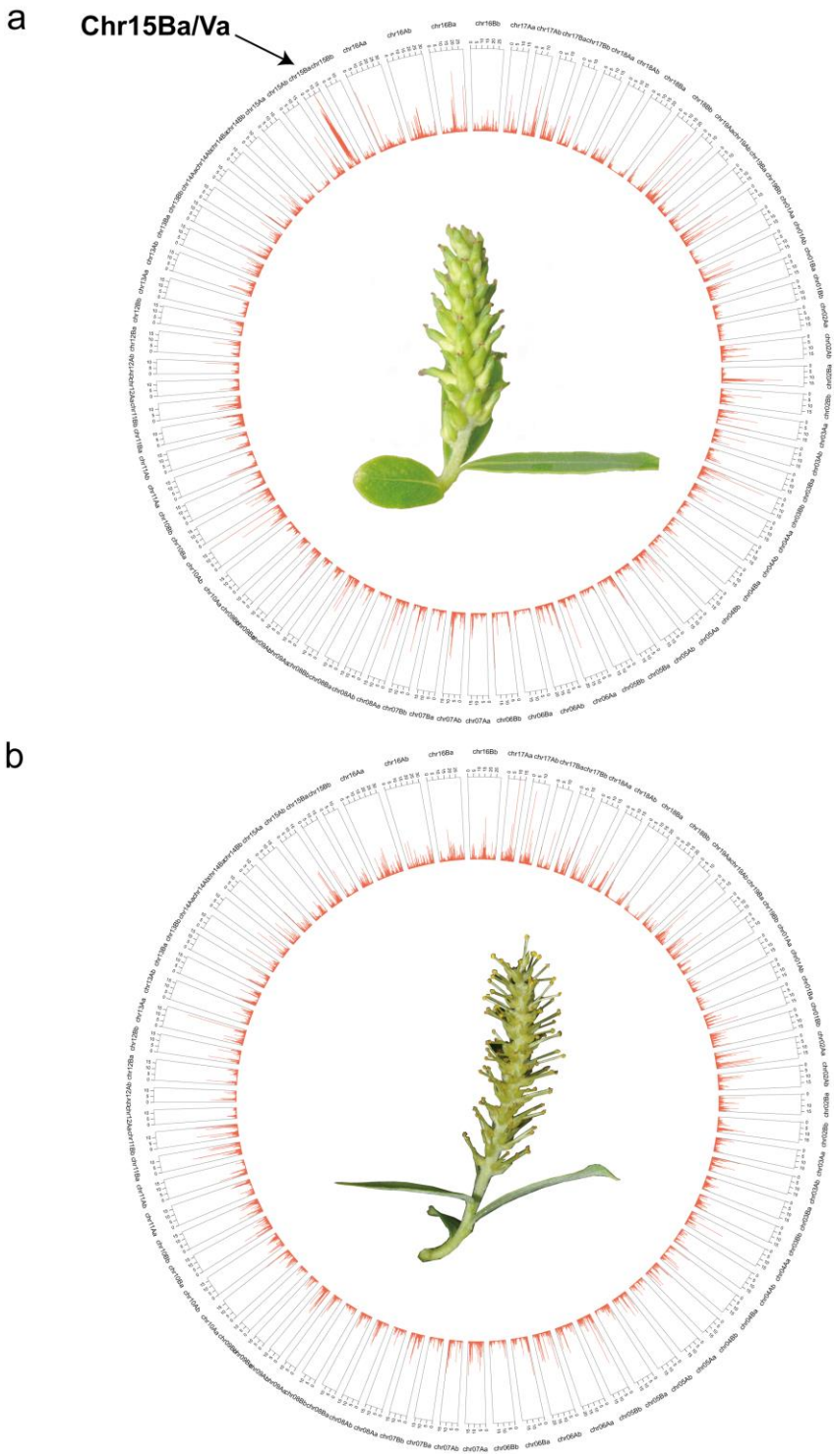
Supplementary Fig. 4. GenomeScope plots for *Salix babylonica* (Saba01F). The higher proportion of aabb than aaab (i.e., aaab < aabb) indicates that *S. babylonica* is allotetraploid.



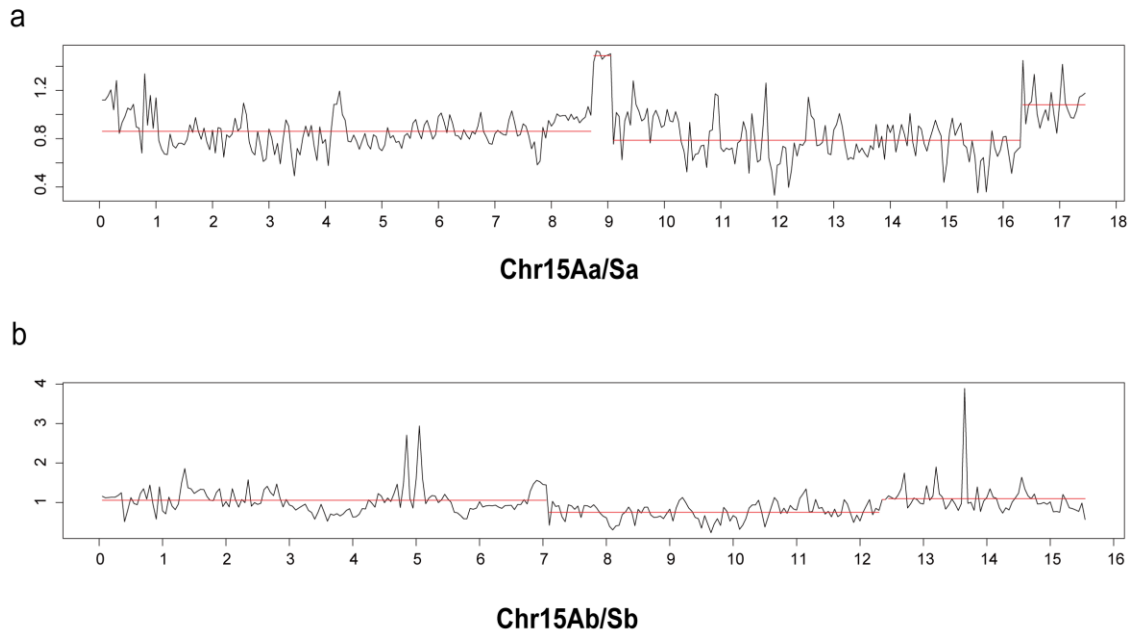
Supplementary Fig. 5. Phylogenetic tree based on chloroplast genome data. Numbers marked on the tree represent bootstrap values.



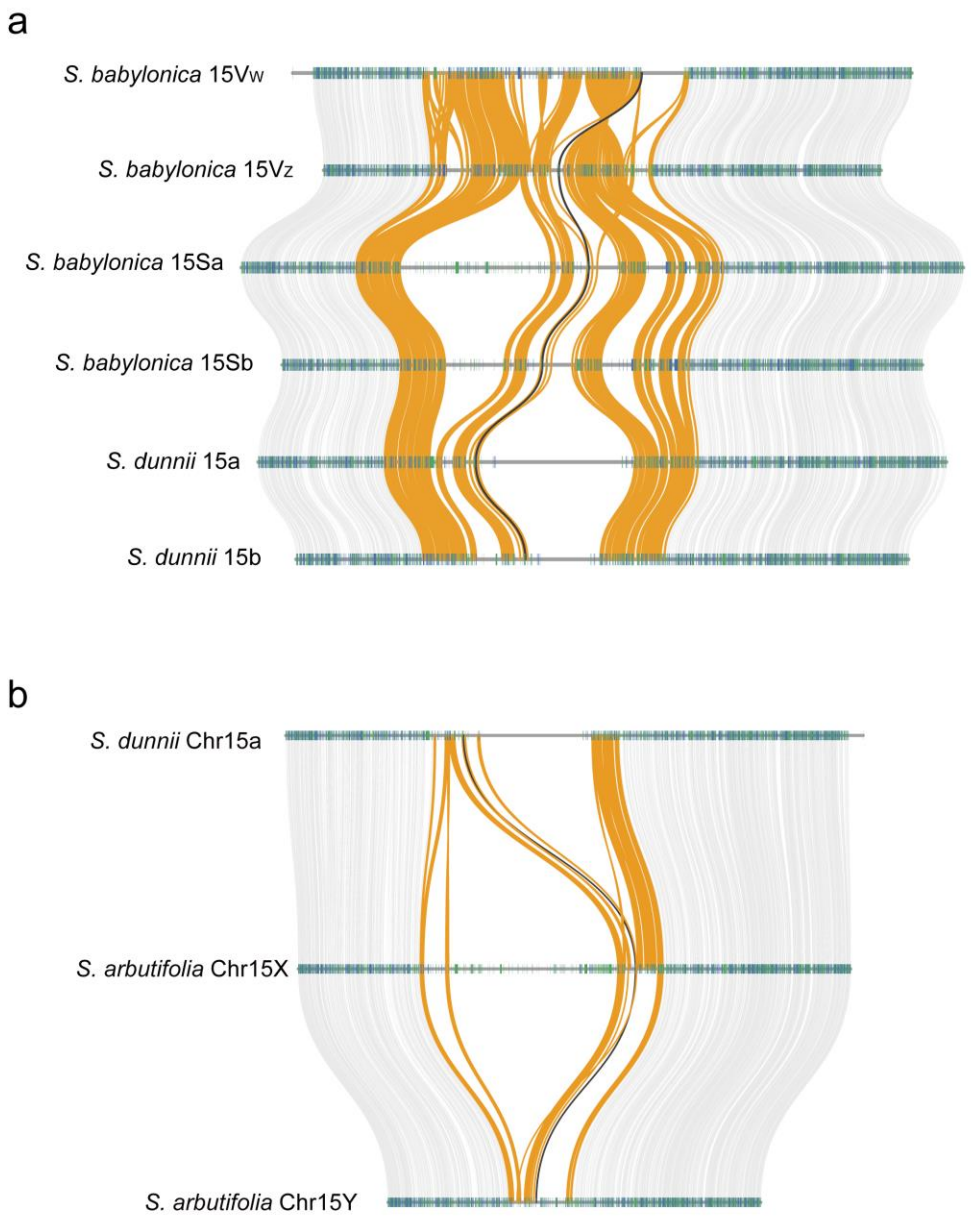
Supplementary Fig. 6. The distribution of sequence divergence rates of TEs between two subgenomes of *S. babylonica*. a The divergence landscape of TEs between Aa/Sa and Ba/Va. **b** The divergence landscape of TEs between Aa/Sa and Bb/Vb. Source data are provided as a Source Data file.



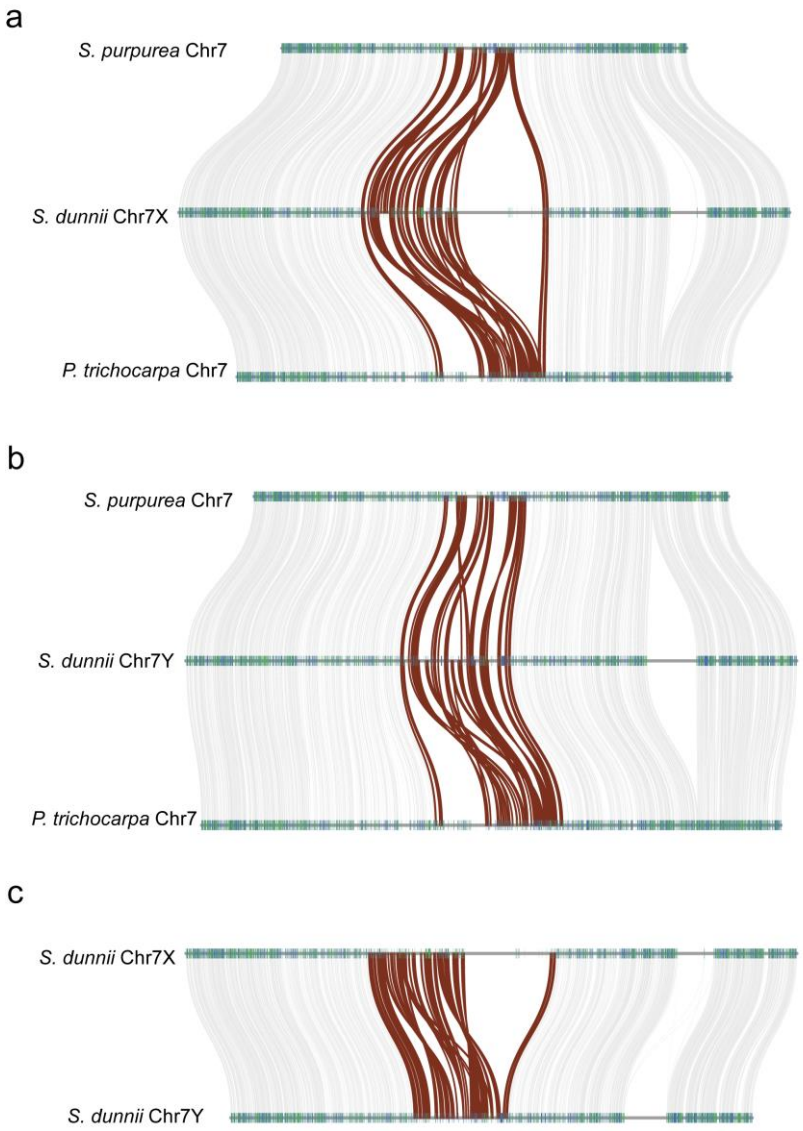
Supplementary Fig. 7. The female and male specific *k-mer* distribution in *S. babylonica*. a The female specific *k-mer* distribution in *S. babylonica*. Inside the circle is a female catkin. There is a significant signal on chromosome 15Ba/Va. **b** The male specific *k-mer* distribution in *S. babylonica*. Inside the circle is a male catkin. Aa=Sa, Ab=Sb, Ba=Va, and Bb=Vb. Source data are provided as a Source Data file.



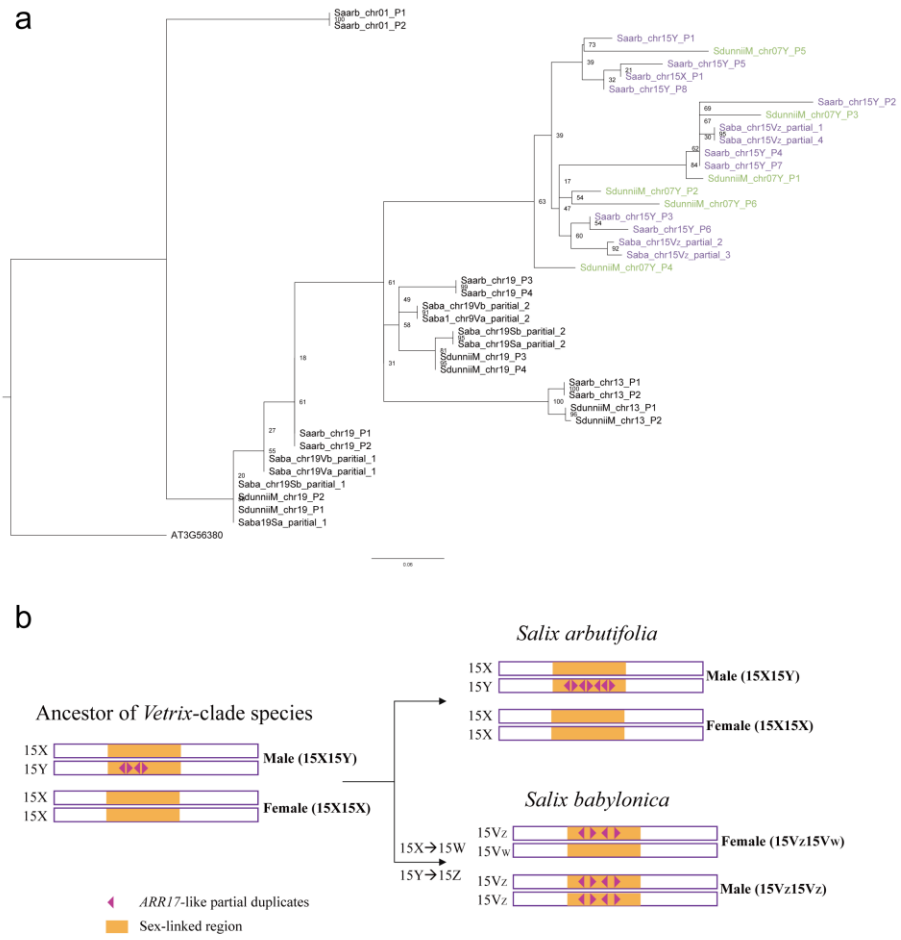
Supplementary Fig. 8. The CQ results (male vs. female alignments) of *S. babylonica*. a The CQ results on chromosomes 15Aa/Sa. **b** The CQ results on 15Ab/Sb. Source data are provided as a Source Data file.



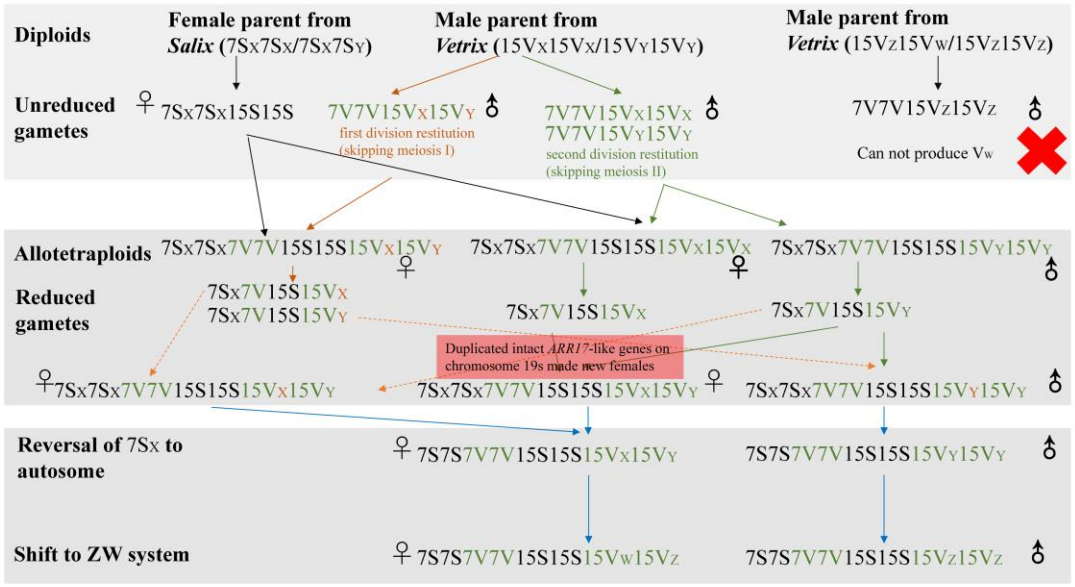
Supplementary Fig. 9. Synteny between sex chromosomes and their homologous chromosome(s). a The collinearity connection between sex chromosomes of *S. babylonica* (15Vw and 15Vz) and homologous autosomes of *S. babylonica* (15Sa and 15Sb) and *S. dunnii* (15a, 15b). **b** The collinearity connection between sex chromosomes of *S. arbutifolia* (15X and 15Y) and homologous autosomes of *S. dunnii* (15a). The yellow regions represent sex-linked regions, and the black lines connect ancient sex-linked homologous genes. Source data are provided as a Source Data file.



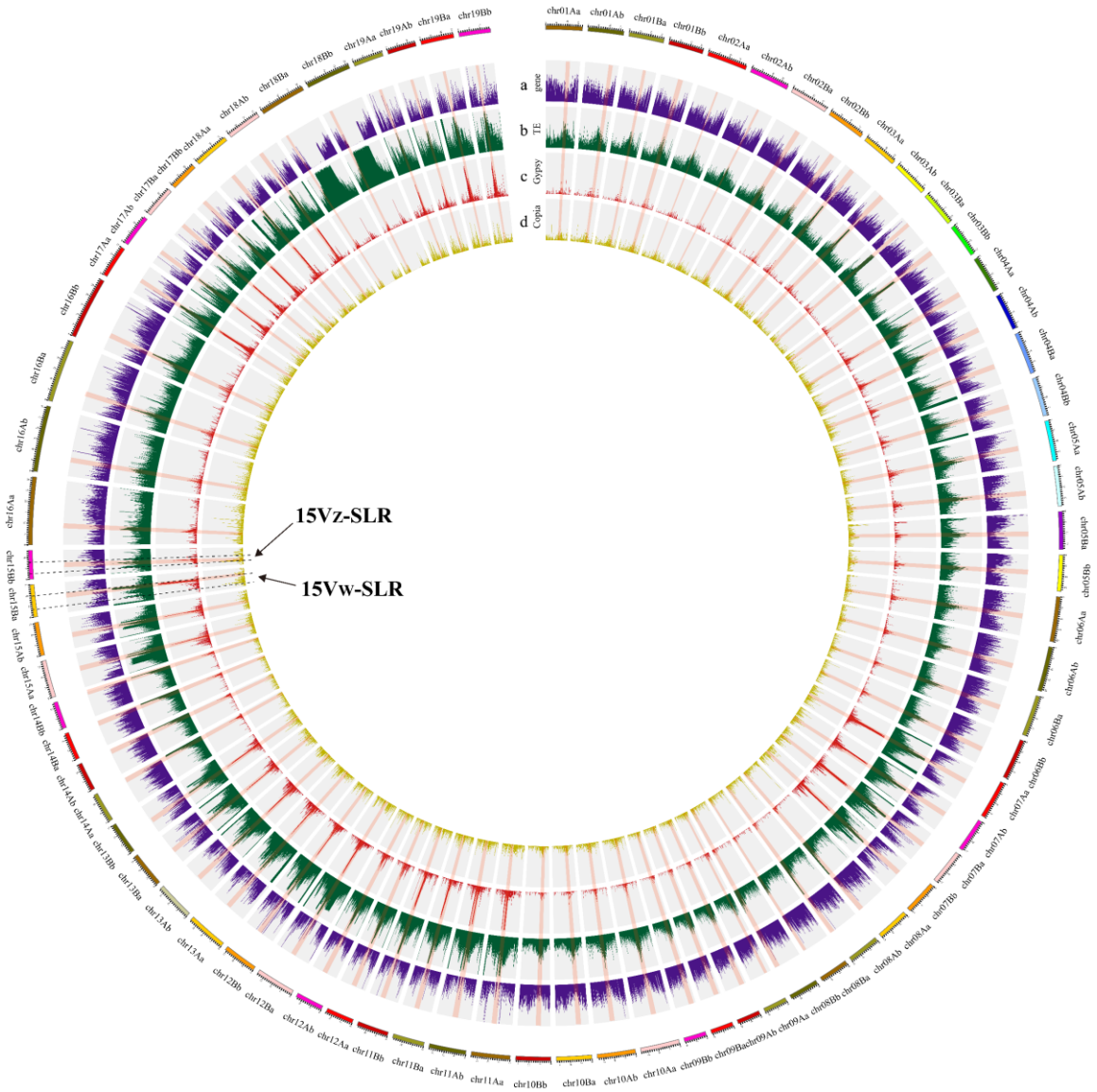
Supplementary Fig. 10. The collinearity results between sex chromosomes of *S. dunnii* and autosomes of *S. purpurea* and *P. trichocarpa*. The brown regions represent sex-linked region. **a** The collinearity connection between sex chromosomes of *S. dunnii* 7X and homologous autosomes of *S. purpurea* Chr7 and *P. trichocarpa* Chr 7. **b** The collinearity connection between sex chromosomes of *S. dunnii* 7Y and homologous autosomes of *S. purpurea* Chr7 and *P. trichocarpa* Chr7. **c** The collinearity connection between sex chromosomes of *S. dunnii* (7X and 7Y). Source data are provided as a Source Data file.



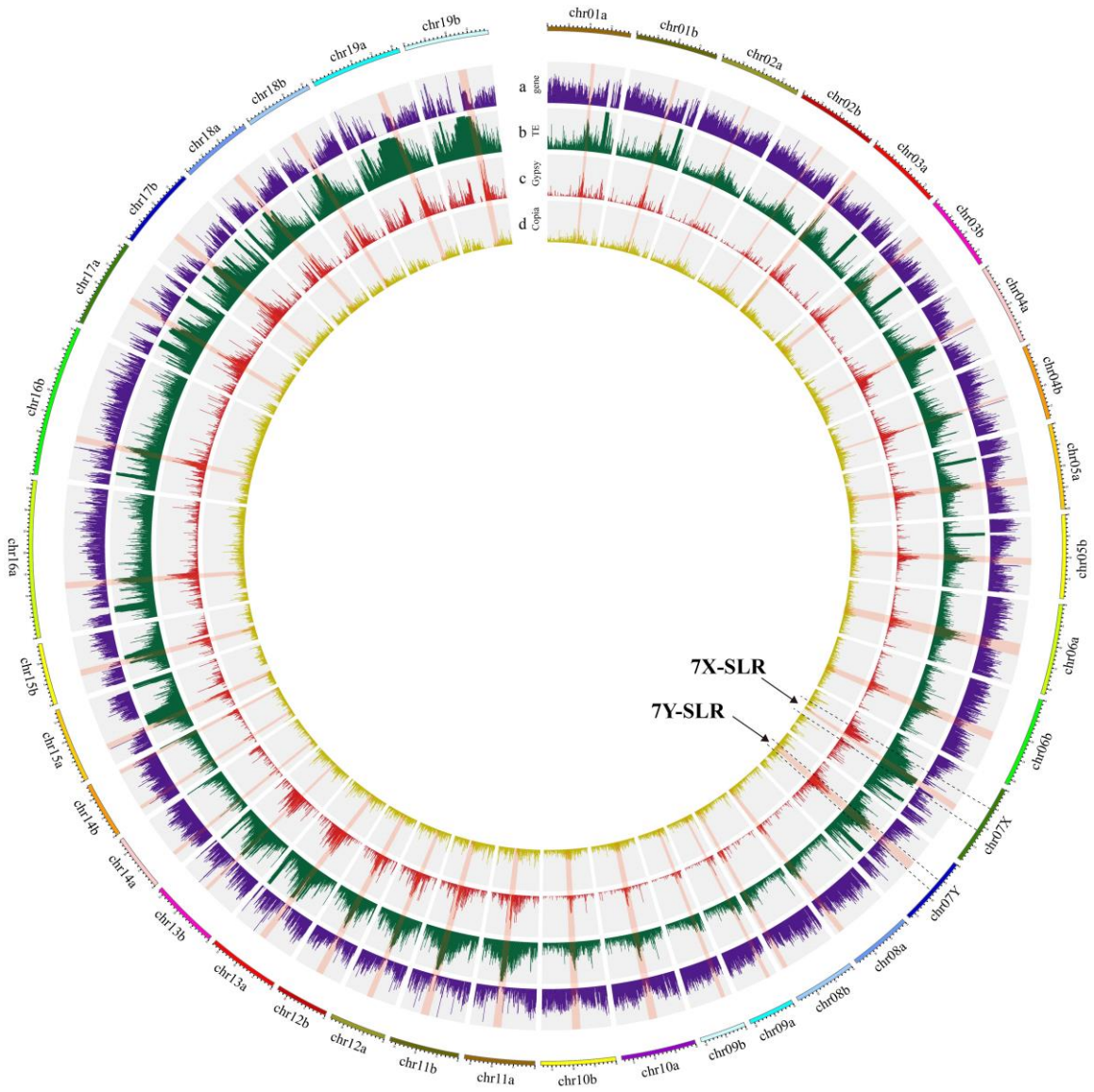
Supplementary Fig. 11. Evolutionary history of *ARR17*-like partial duplicates. a The phylogeny of *ARR17*-like partial duplicates. Sarrb: *S. arbutifolia*, SdunniiM: *S. dunnii*, Saba: *S. babylonica*, AT: *Arabidopsis thaliala*. The purple and green represent genomic contents from ancestor of the *Vetrix*-clade and *Salix*-clade. **b** The evolutionary trajectory of *ARR17*-like partial duplicates in ancestral species, *S. arbutifolia* and *S. babylonica*.



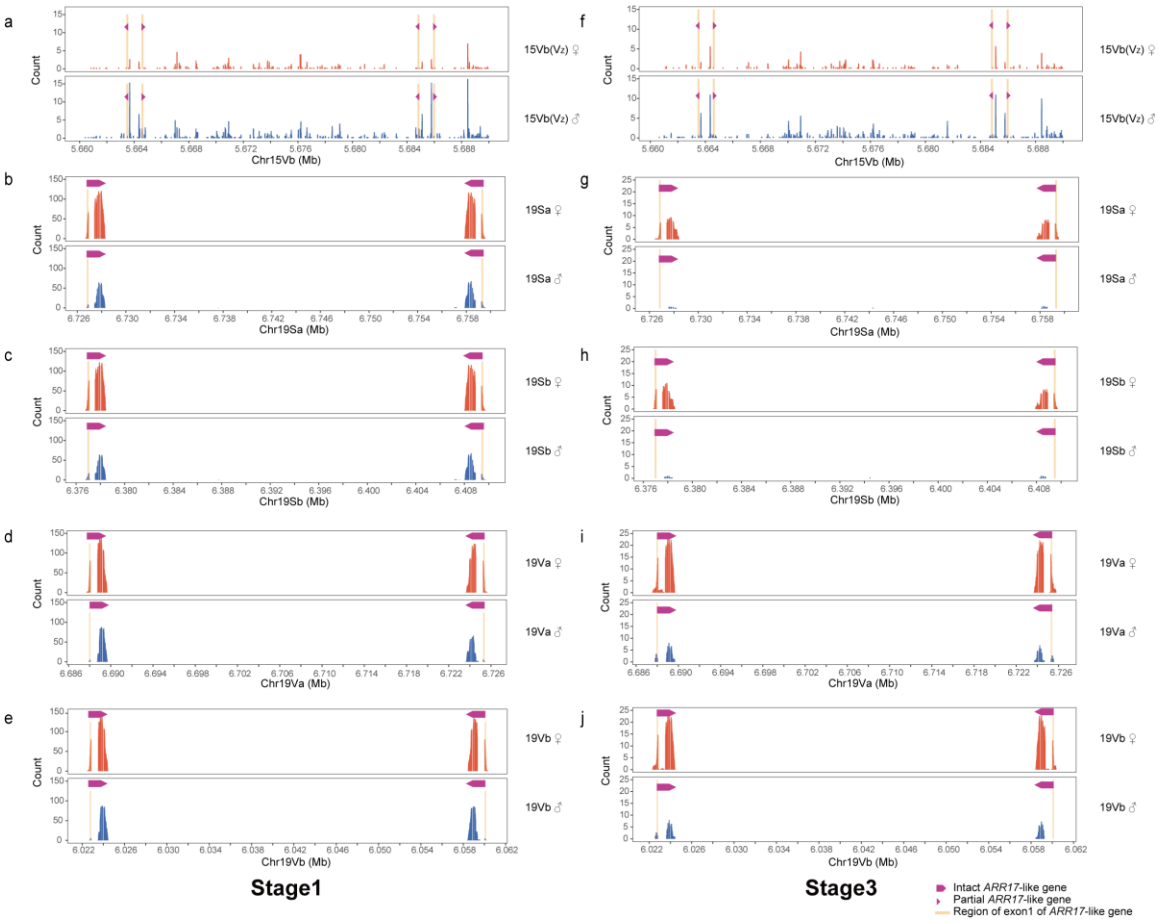
Supplementary Fig. 12. Hypothetical evolutionary origin of *Salix babylonica*. Black arrows represent gametes from female progenitor of *Salix* clade. Orange and green arrows represent gametes from male progenitor of *Vetrix* clade, and subsequent evolution routes. Blue arrows represent the X to autosome and XY to ZW transition.



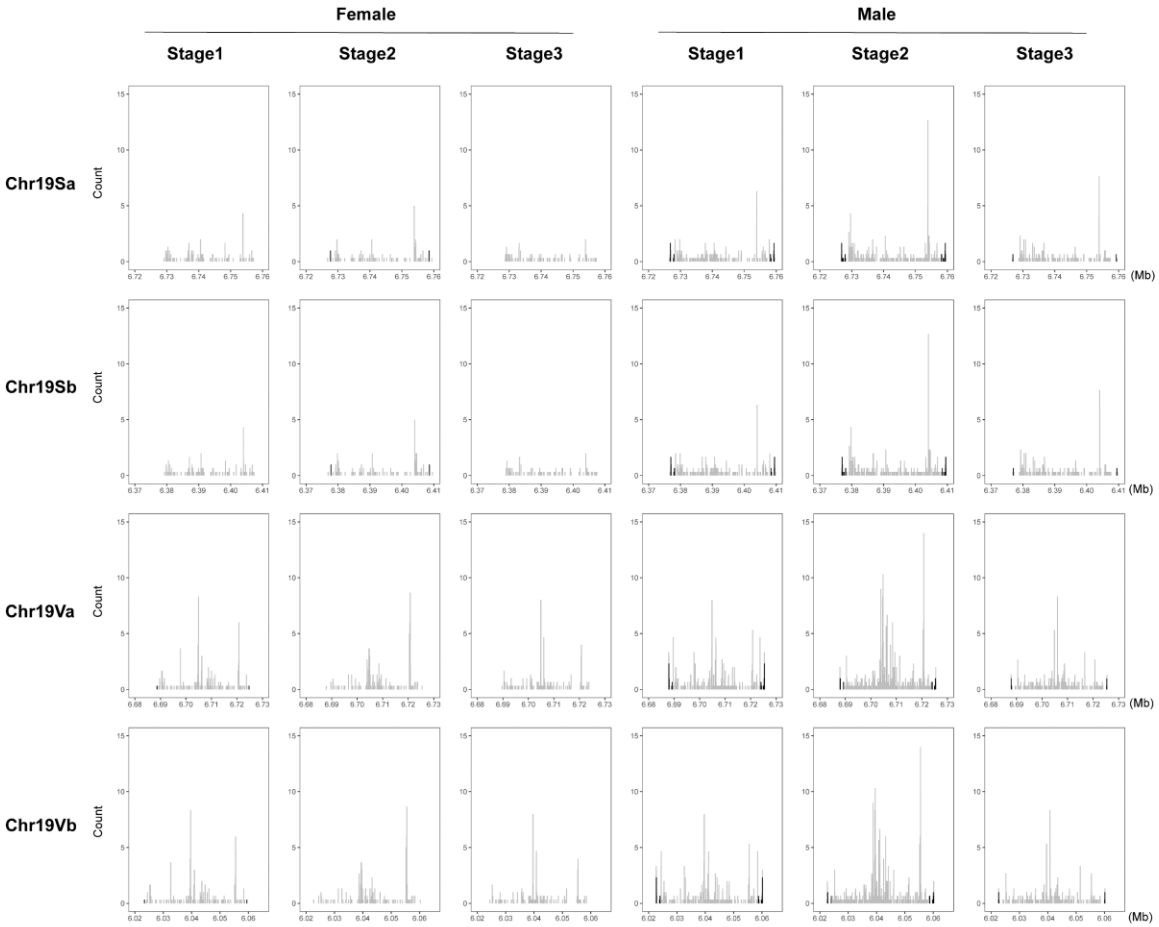
Supplementary Fig. 13. Genome structure of *S. babylonica*. (a) gene density, (b) total TE density, (c) LTR-Gypsy density, (d) LTR-Copia density. The orange area indicates the location of the predicted centromere. Aa=Sa, Ab=Sb, Ba=Va, and Bb=Vb. SLR: sex-linked region. Source data are provided as a Source Data file.



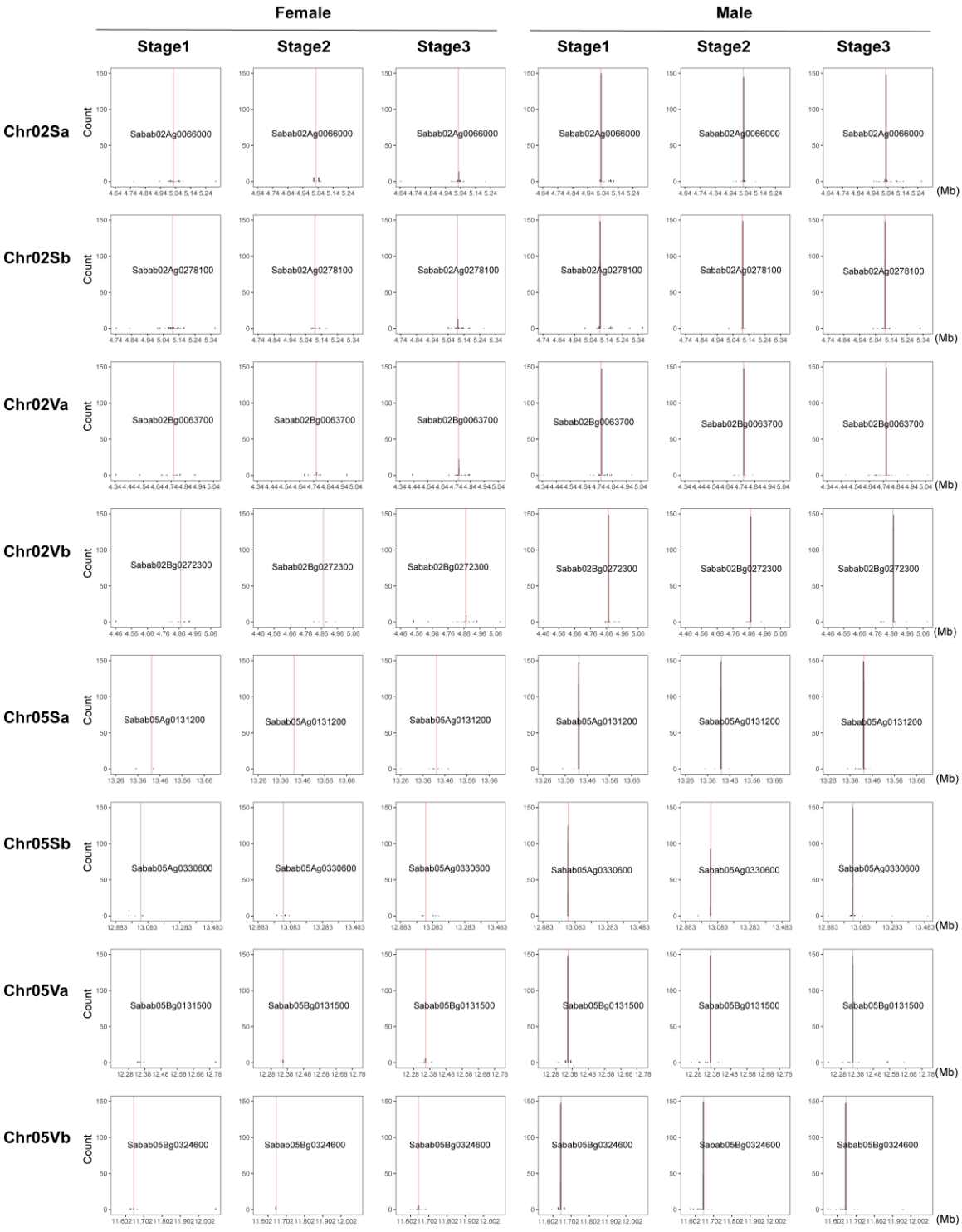
Supplementary Fig. 14. Genome structure of *S. dunnii*. (a) gene density, (b) total TE density, (c) LTR-Gypsy density, (d) LTR-Copia density. The orange area indicates the location of the predicted centromere. SLR: sex-linked region. Source data are provided as a Source Data file.



Supplementary Fig. 15. The expression patterns of *ARR17*-like duplicates in female and male buds. a (stage 1), **f** (stage 3) Transcript level of small RNAs of partial *ARR17*-like duplicates and around regions on sex chromosome 15Vz. Orange area indicates exon1 of *ARR17*-like copies. **b**, **c, d,** and **e** (stage 1); **g**, **h**, **i**, and **j** (stage 3) Transcript level of intact *ARR17*-like genes on Chr19Sa, Chr19Sb, Chr19Va and Chr19Vb. The purple triangles indicate partial *ARR17*-like duplicates (the tip is the end of the duplicate). The purple arrows show the intact *ARR17*-like duplicates (the tip is the end of the gene). Source data are provided as a Source Data file.



Supplementary Fig. 16. Transcript level of small RNAs around intact *ARR17*-like duplicates of *S. babylonica*. Black represents the positions of intact *ARR17*-like duplicates, while grey represents the other regions. Source data are provided as a Source Data file.



Supplementary Fig. 17. Transcript level of intact *PI*-like genes of *S. babylonica* across three stages.

Red area indicates exon1 of *PI*-like genes. Source data are provided as a Source Data file.

Supplementary Table 1. Sequencing information of *S. dunnii* and *S. babylonica.*

	S. dunnii			***S. babylonica***		
Sequencing method	HiFi	Illumina short reads	HiC	HiFi	Illumina short reads	HiC
Total Number of reads	2613804	268886000	247657000	2473021	526734000	1092986000
Total Number of sequenced bases (Gb)	39.482	40.333	37.149	45.659	79.010	163.601
N50 (bp)	15650	150	150	19307	150	150
Coverage (X)	112	108	107	64	94	254

Supplementary Table 2. Gap distribution in genome assembly of *S. babylonica*.

Chr	Start	End	Gap length
chr03Sa/Aa	4726713	4726813	100
chr03Sb/Ab	4671630	4671728	98
chr04Va/Ba	18248904	18249004	100
chr04Va/Ba	18479352	18479452	100
chr04Va/Ba	19215784	19215883	99
chr04Va/Ba	19539379	19539479	100
chr04Vb/Bb	9162330	9162430	100
chr04Vb/Bb	17052173	17052273	100
chr04Vb/Bb	17164068	17164165	97
chr04Vb/Bb	17846690	17846790	100
chr07Sa/Aa	8402332	8402432	100
chr07Sa/Aa	17121795	17121895	100
chr07Sb/Ab	15644231	15644331	100
chr07Vb/Bb	7344360	7344458	98
chr11Sb/Ab	9941265	9941365	100
chr11Va/Ba	8334071	8334171	100
chr12Sa/Aa	6483302	6483402	100
chr17Sa/Aa	64135	64176	41
chr17Sb/Ab	124864	124964	100
chr17Sb/Ab	376658	376758	100
chr17Sb/Ab	2734570	2734670	100
chr18Vb/Bb	12649511	12649611	100
chr19Sa/Aa	2339349	2339449	100
chr19Sa/Aa	7419513	7419613	100

Supplementary Table 3. The BUSCO results of *S. dunnii* and *S. babylonica* (protein).

	S. dunnii	***S. babylonica***
Complete BUSCOs	1416 (98.4%)	1422 (98.8%)
Complete and single-copy BUSCOs	24 (1.7%)	14 (1.0%)
Complete and duplicated BUSCOs	1392 (96.7%)	1408 (97.8%)
Fragmented BUSCOs	8 (0.6%)	6 (0.4%)
Missing BUSCOs	16 (1.0%)	12 (0.8%)

Supplementary Table 4. Statistics of protein-coding genes of the genome of *Salix dunnii* and *S. babylonica.*

	S. dunnii	***S. babylonica***
Gene number	65485	123231
Transcript number	119374	167939
Average gene region length (bp)	3683.4	3421.6
Average transcript length (bp)	2212.4	1764.2
Average CDS length (bp)	1352	1311.9
Average exons per transcript number	6.4	5.9
Average exon length (bp)	346.5	298.3
Average intron length (bp)	421.7	428.3

Supplementary Table 5. Genome datasets used in the paper.

Taxa	Version	URL	Reference
Salix suchowensis	v4.1	http://popgenie.org/	1
Salix purpurea	v5.1	https://genome.jgi.doe.gov/portal/pages/dynamicOrganismDownload.jsf?organism=Spurpurea	2
Salix viminalis	used in	requested from the first author Pedro Almeida	3
Salix brachista	used in	https://www.ncbi.nlm.nih.gov/assembly/GCA_009078335.1	4
Salix babylonica	saba01F	This paper	
Salix arbutifolia	B649AF		5
Salix dunnii	FNU-M-1	This paper	
Salix chaenomeloides	GWHBDOA00000000	https://ngdc.cncb.ac.cn/gwh/Assembly/21464/show	6
Populus trichocarpa	v4.1	https://genome.jgi.doe.gov/portal/pages/dynamicOrganismDownload.jsf?organism=Ptrichocarpa	
Salix suchowensis	SRR10197854	https://www.ncbi.nlm.nih.gov	
Salix purpurea	SRR3927002	https://www.ncbi.nlm.nih.gov	
Salix viminalis	ERR1558612	https://www.ncbi.nlm.nih.gov	
Salix chaenomeloides	SRR16018669	https://www.ncbi.nlm.nih.gov	

Supplementary Table 6. The statistics of gene loss in SLR and PARs.

Species	Region	Type	Loss gene number	Total gene number	Degradation rate (%)
S. babylonica	SLR	15Vw Loss	8	282	2.84
S. babylonica	SLR	15Vz Loss	11	282	3.90
S. babylonica	PARs	15Vw Loss	11	886	1.24
S. babylonica	PARs	15Vz Loss	14	886	1.58
S. purpurea	SLR	15W Loss	8	320	2.50
S. purpurea	SLR	15Z Loss	27	320	8.44
S. purpurea	PARs	15W Loss	2	1042	0.19
S. purpurea	PARs	15Z Loss	3	1042	0.29
S. arbutifolia	SLR	15X Loss	8	87	9.20
S. arbutifolia	SLR	15Y Loss	10	87	11.49
S. arbutifolia	PARs	15X Loss	6	1027	0.58
S. arbutifolia	PARs	15Y Loss	6	1027	0.58
S. dunnii	SLR	7X Loss	1	82	1.22
S. dunnii	SLR	7Y Loss	1	82	1.22
S. dunnii	PARs	7X Loss	6	1135	0.53
S. dunnii	PARs	7Y Loss	7	1135	0.62

Supplementary references

1. Dai, X. *et al.* The willow genome and divergent evolution from poplar after the common genome duplication. *Cell Res.* **24**, 1274–1277 (2014).
2. Zhou, R. *et al.* A willow sex chromosome reveals convergent evolution of complex palindromic repeats. *Genome Biol.* **21**, 1–19 (2020).
3. Almeida, P. *et al.* Genome assembly of the basket willow, *Salix viminalis*, reveals earliest stages of sex chromosome expansion. *BMC Biol.* **18**, 1–18 (2020).
4. Chen, J. *et al.* Genome-wide analysis of Cushion willow provides insights into alpine plant divergence in a biodiversity hotspot. *Nat. Commun.* **10**, 5230 (2019).
5. Wang, Y. *et al.* Gap-free X and Y chromosomes of *Salix arbutifolia* reveal an evolutionary change from male to female heterogamety in willows, without a change in the sex-determining region. *New Phytol.* **242**, 2872–2887 (2024).
6. Wang, D. *et al.* Repeated turnovers keep sex chromosomes young in willows. *Genome Biol.* **23**, 1–23 (2022).