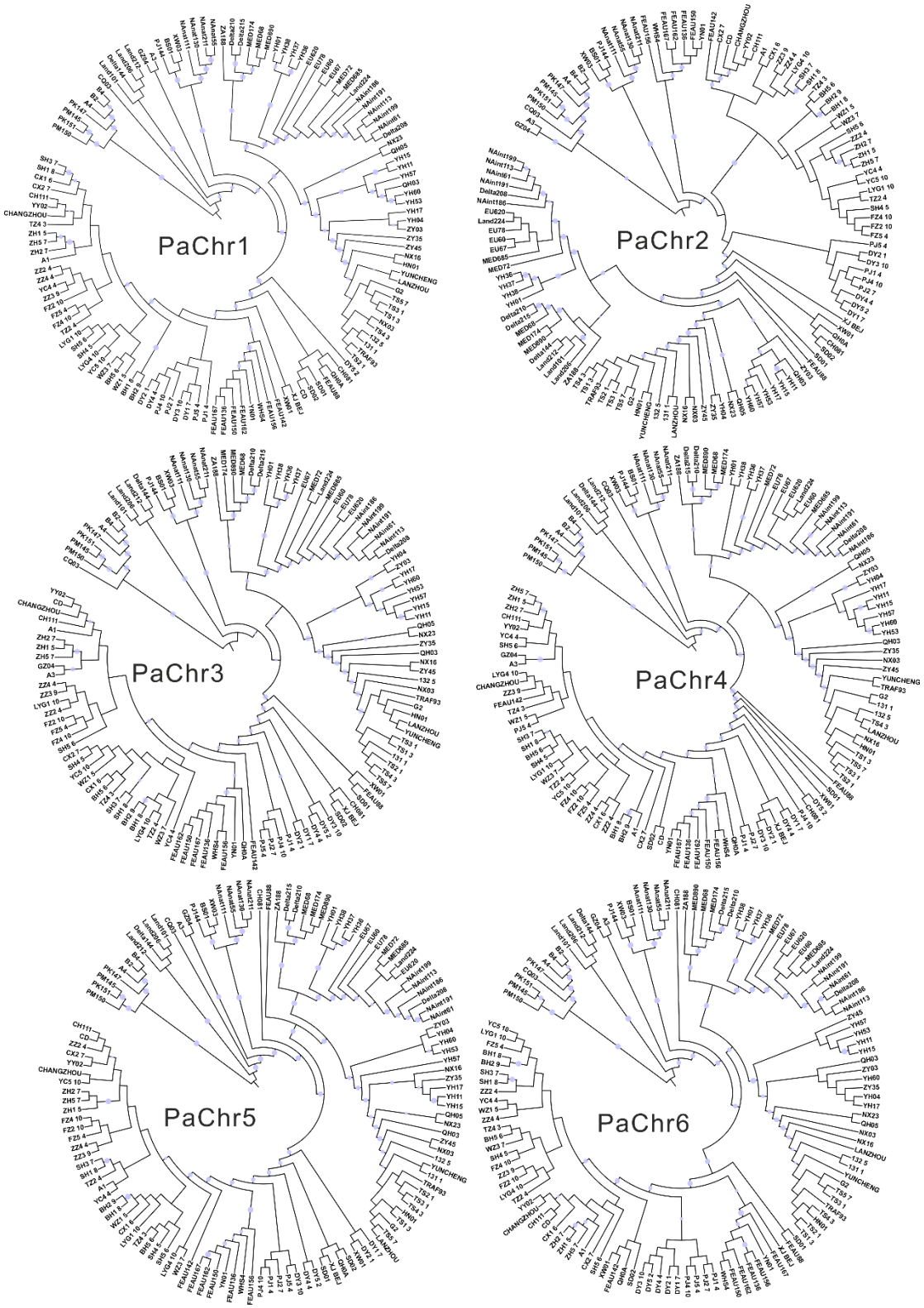
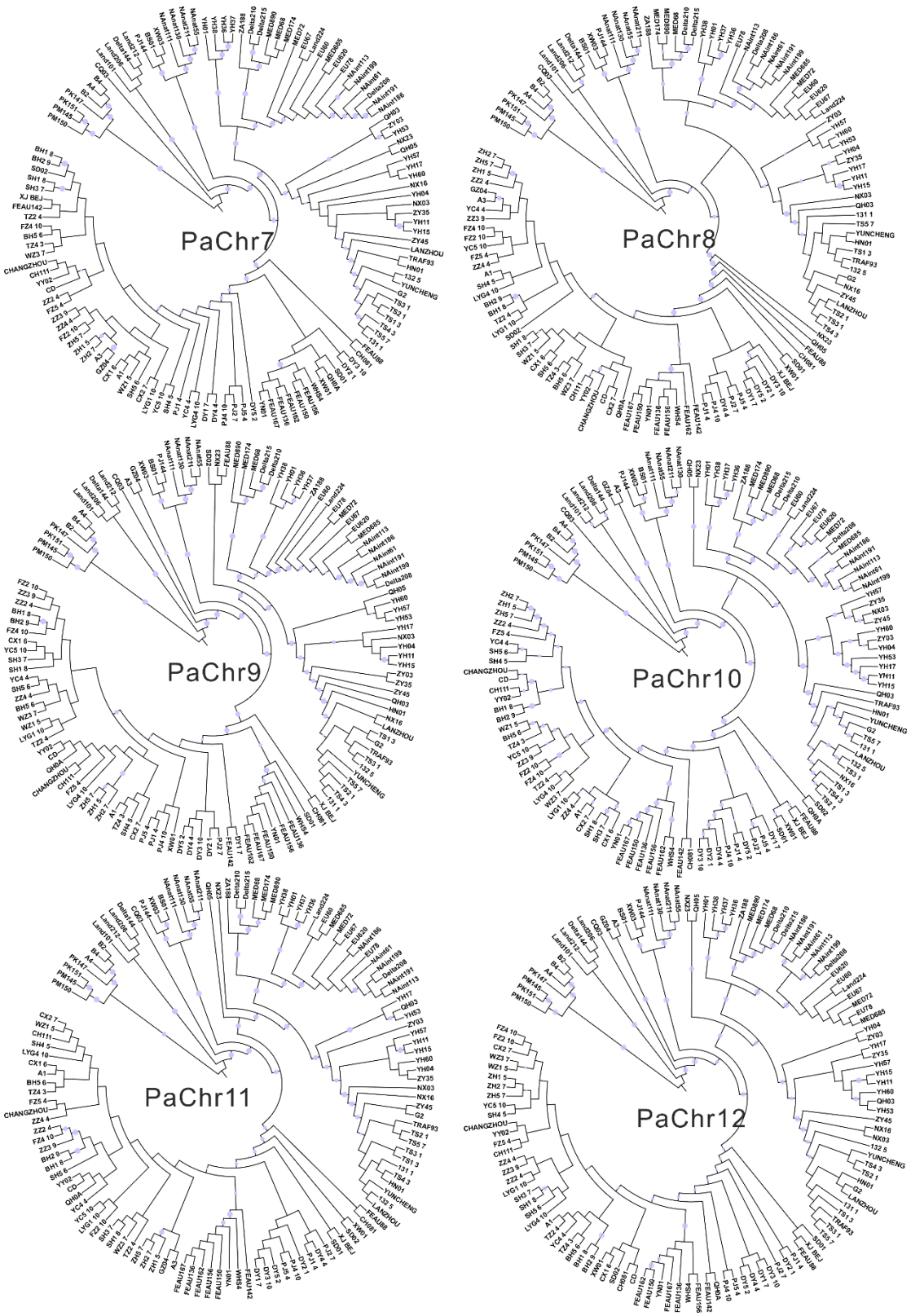
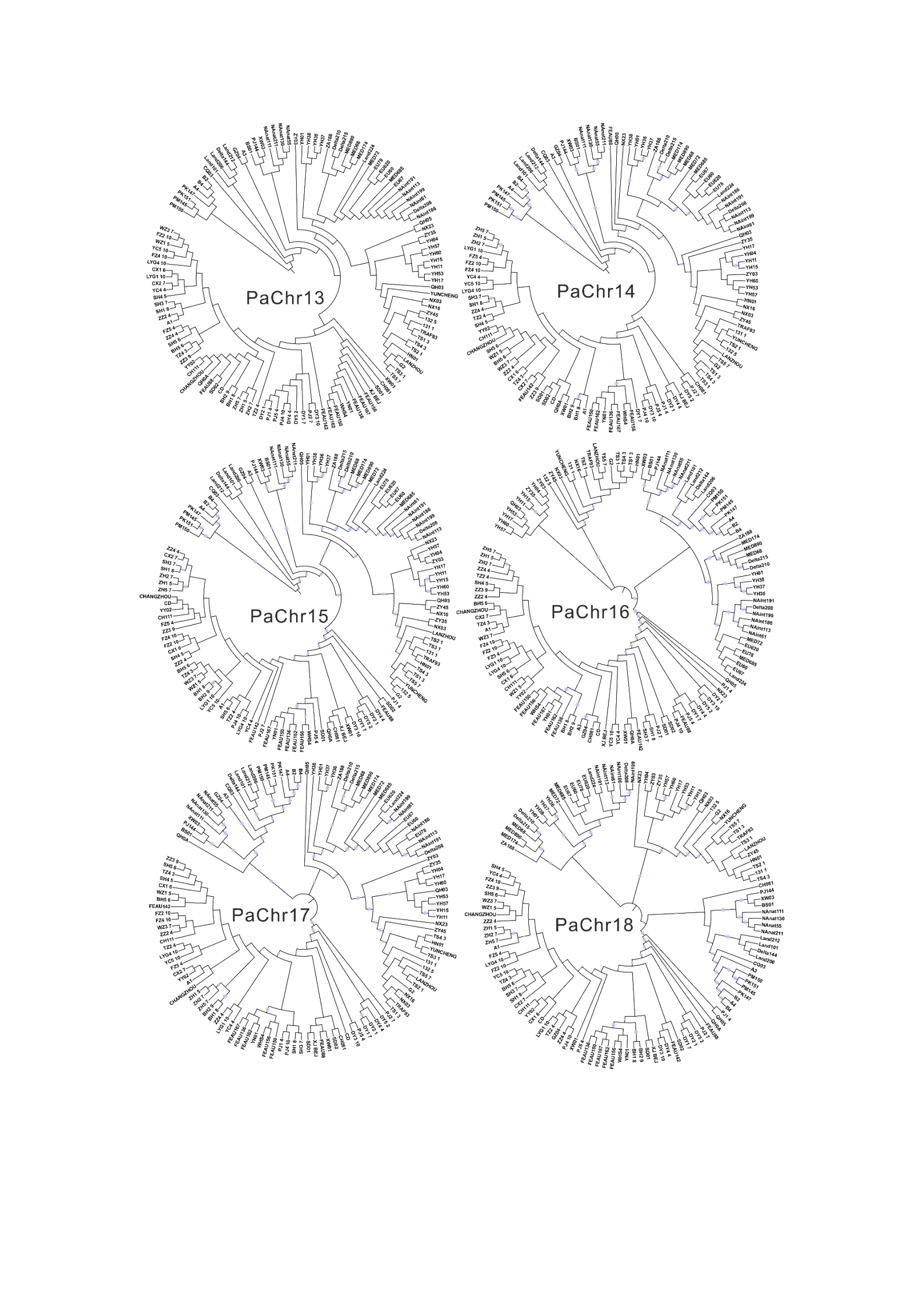
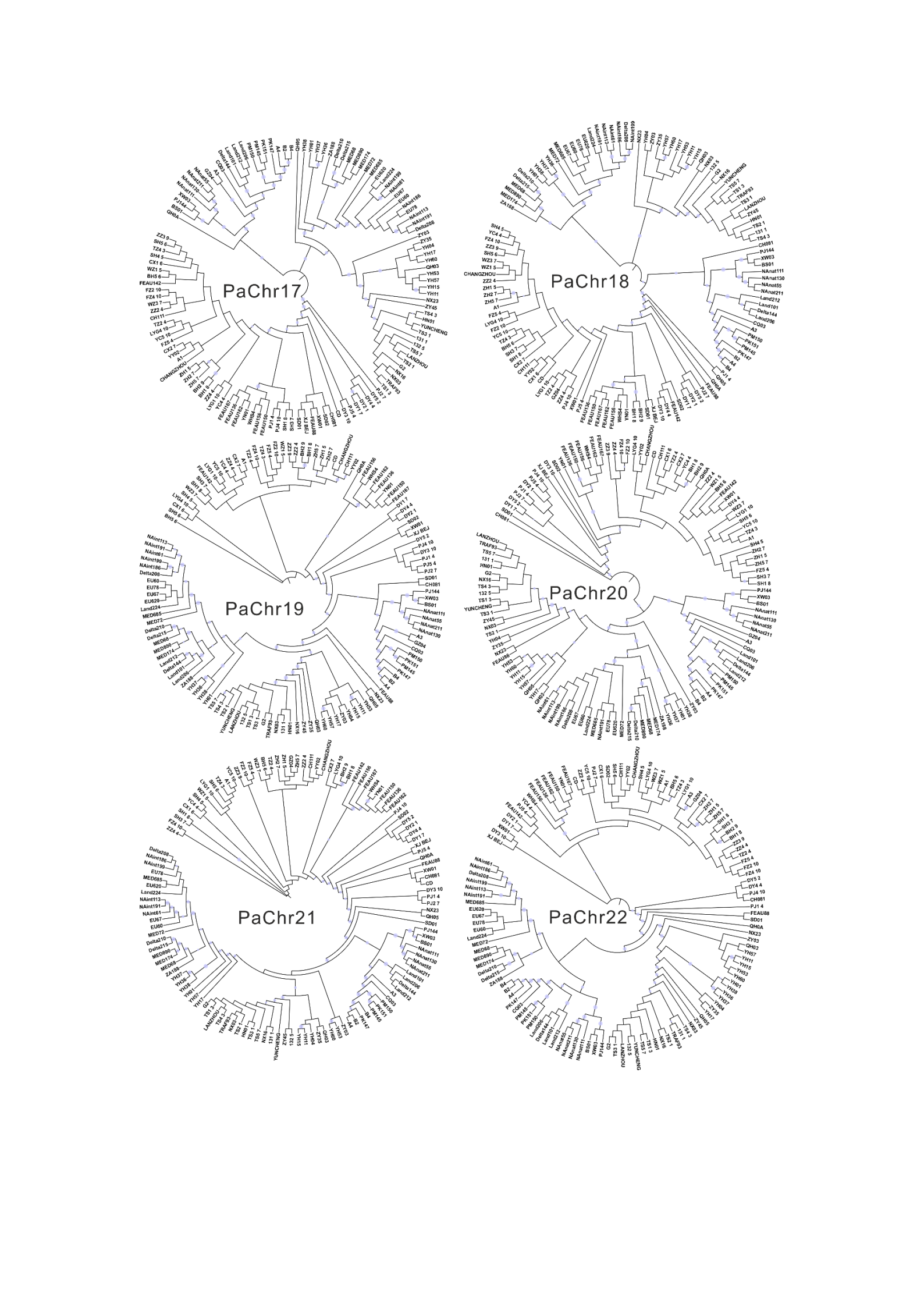
Supplementary Materials




PaChr7
PaChr8
PaChr9
PaChr10
PaChr11
PaChr12


PaChr13
PaChr14
PaChr15
PaChr16
PaChr17
PaChr18


PaChr17
PaChr18
PaChr19
PaChr20
PaChr21
PaChr22

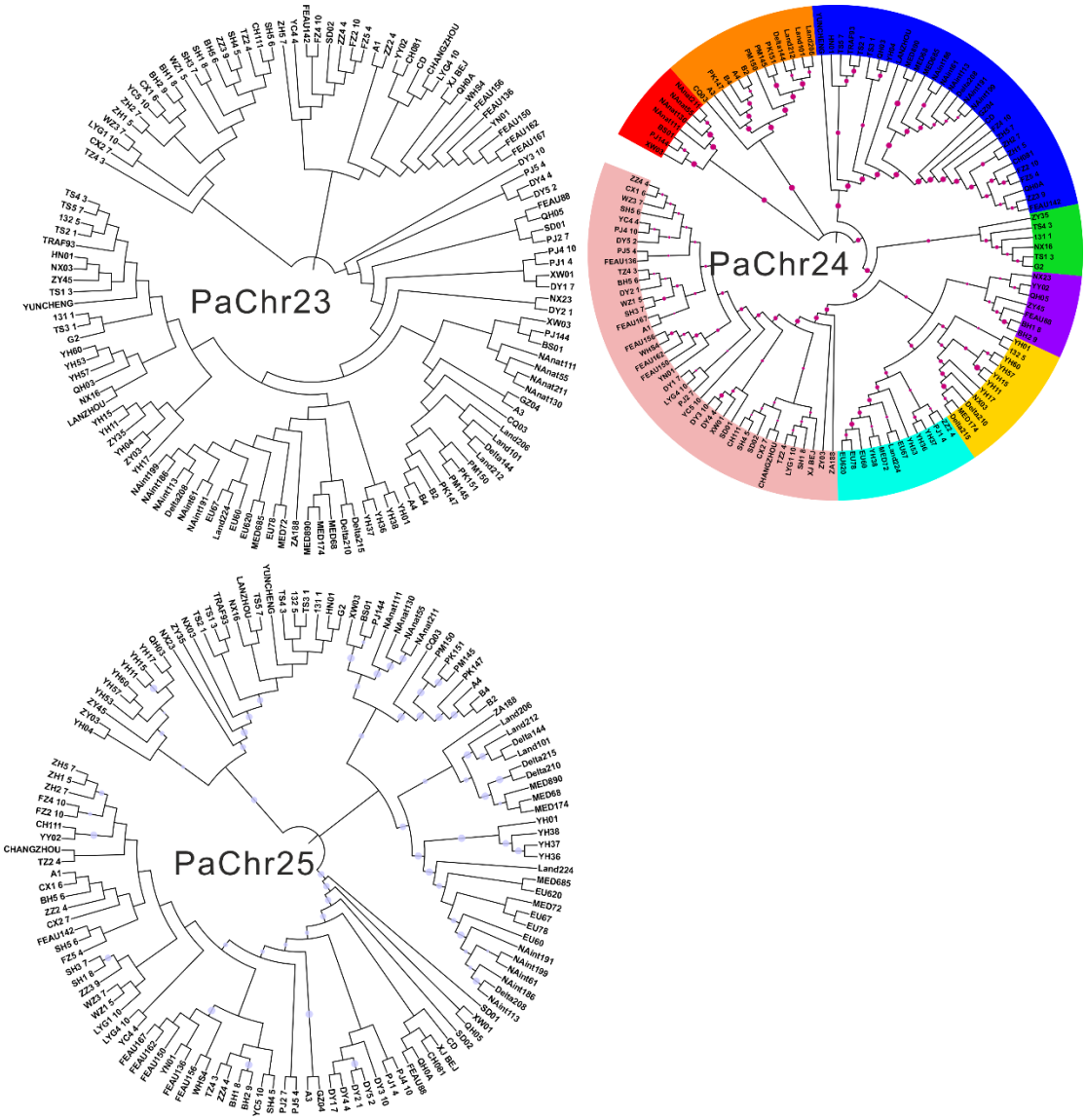


Figure S1. Phylogeny of *Phragmites* species for each chromosome. Circular phylogenetic trees (cladograms) constructed independently for each of the 25 chromosomes of *Phragmites australis* using RAXML. Tip labels denote individual samples. Nodes are marked with filled squares/circles indicating high bootstrap support, sized/coloured by support value]. Discordant topologies among chromosomes are consistent with their own evolutionary history.

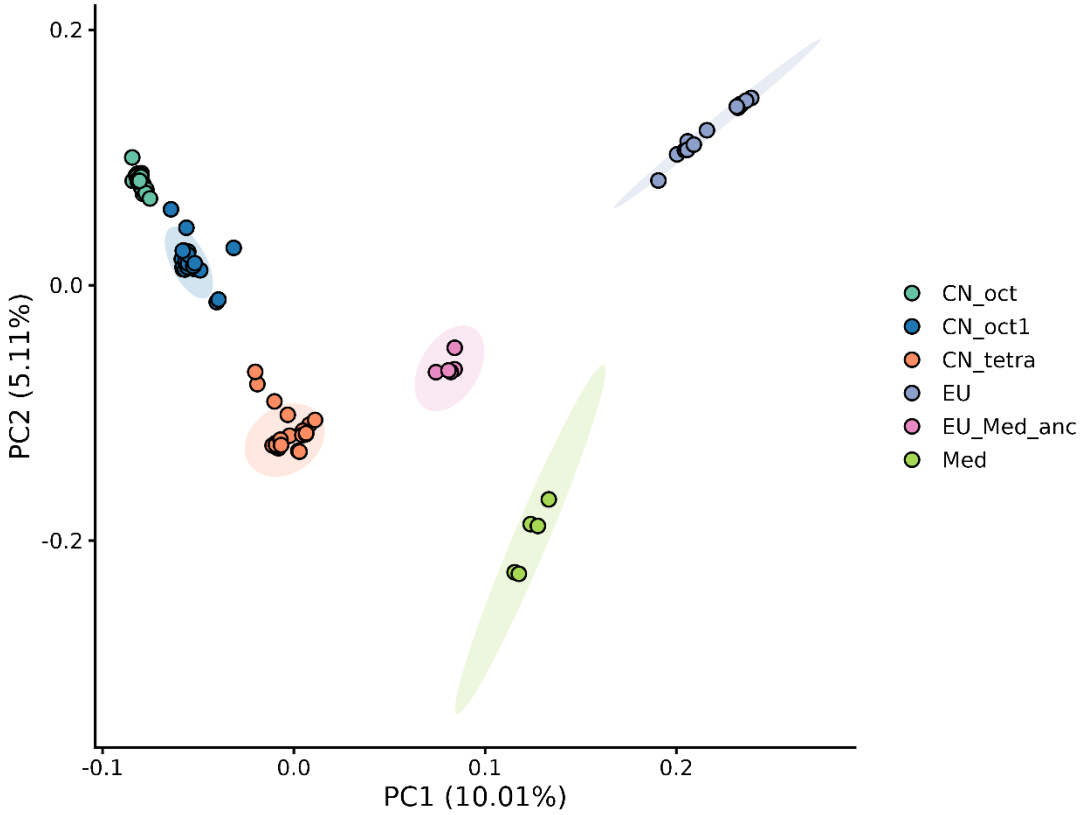


Figure S2. Principal component analysis (PCA) within *Phragmites australis*. PCA based on genome-wide SNPs across all sampled individuals. PC1 accounts for 10.01% of the total variance, and PC2 accounts for 5.11% of the variance. Individuals separate into five major groups: CN_tetra (China tetraploid), CN octoploid, EU (Europe), EU_Med_anc (Europe–Mediterranean contacts), and Med (Mediterranean). The CN octoploid lineage further separates into two distinct subclusters (CN_oct and CN_oct1) along PC2. Points are coloured by group, with shaded ellipses indicating the 95% confidence interval for each group.

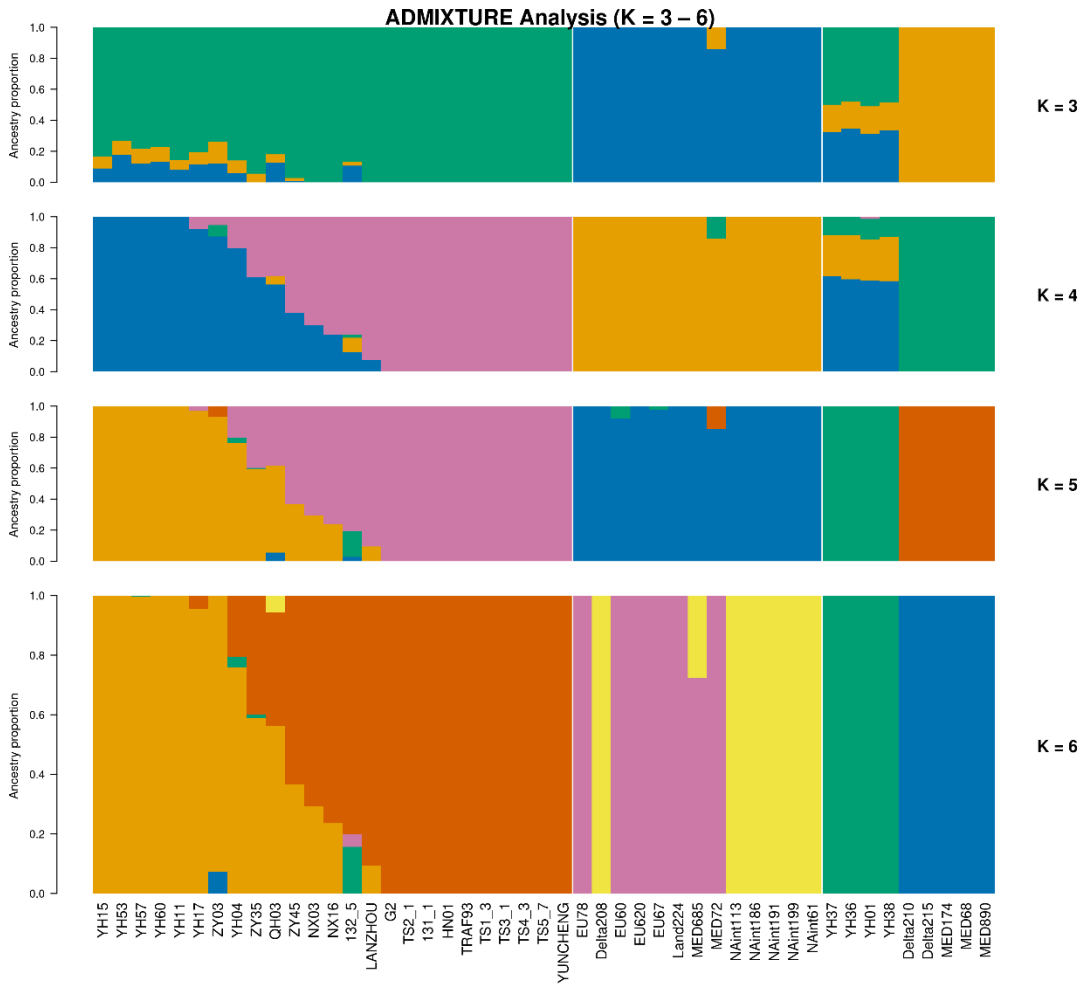
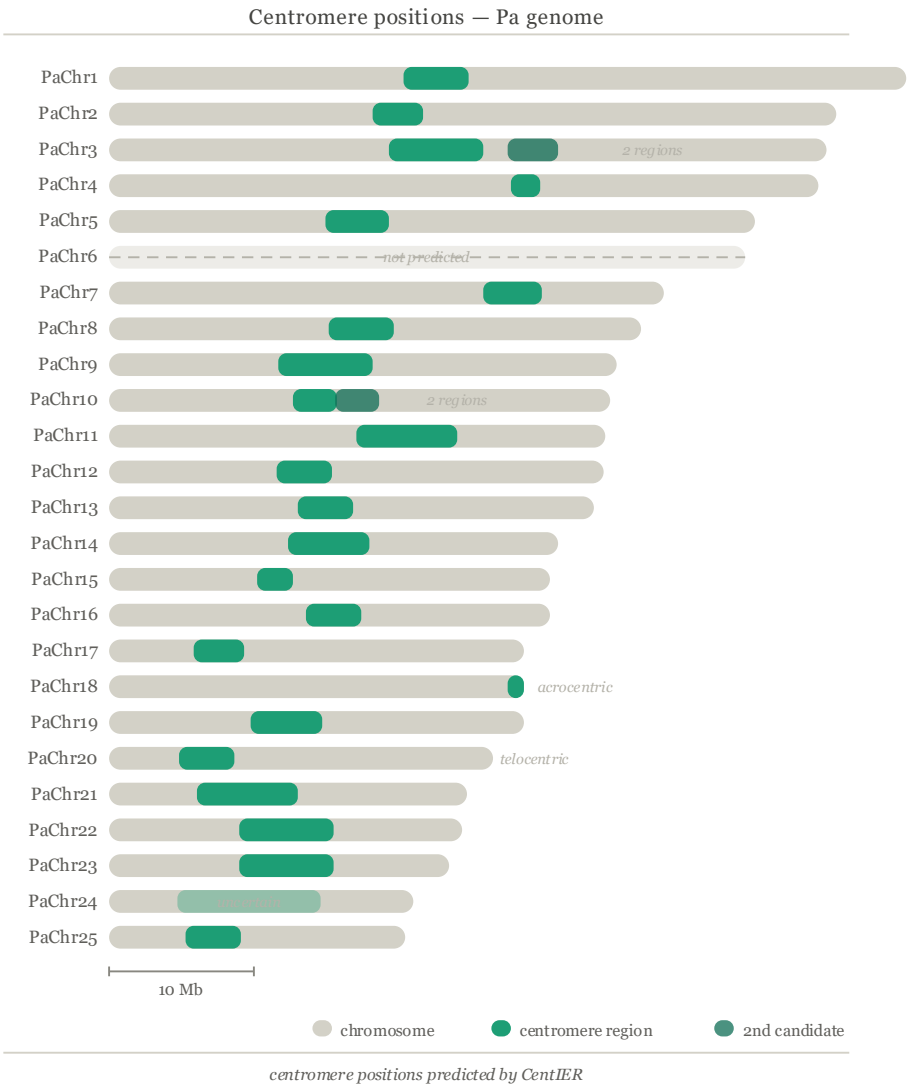
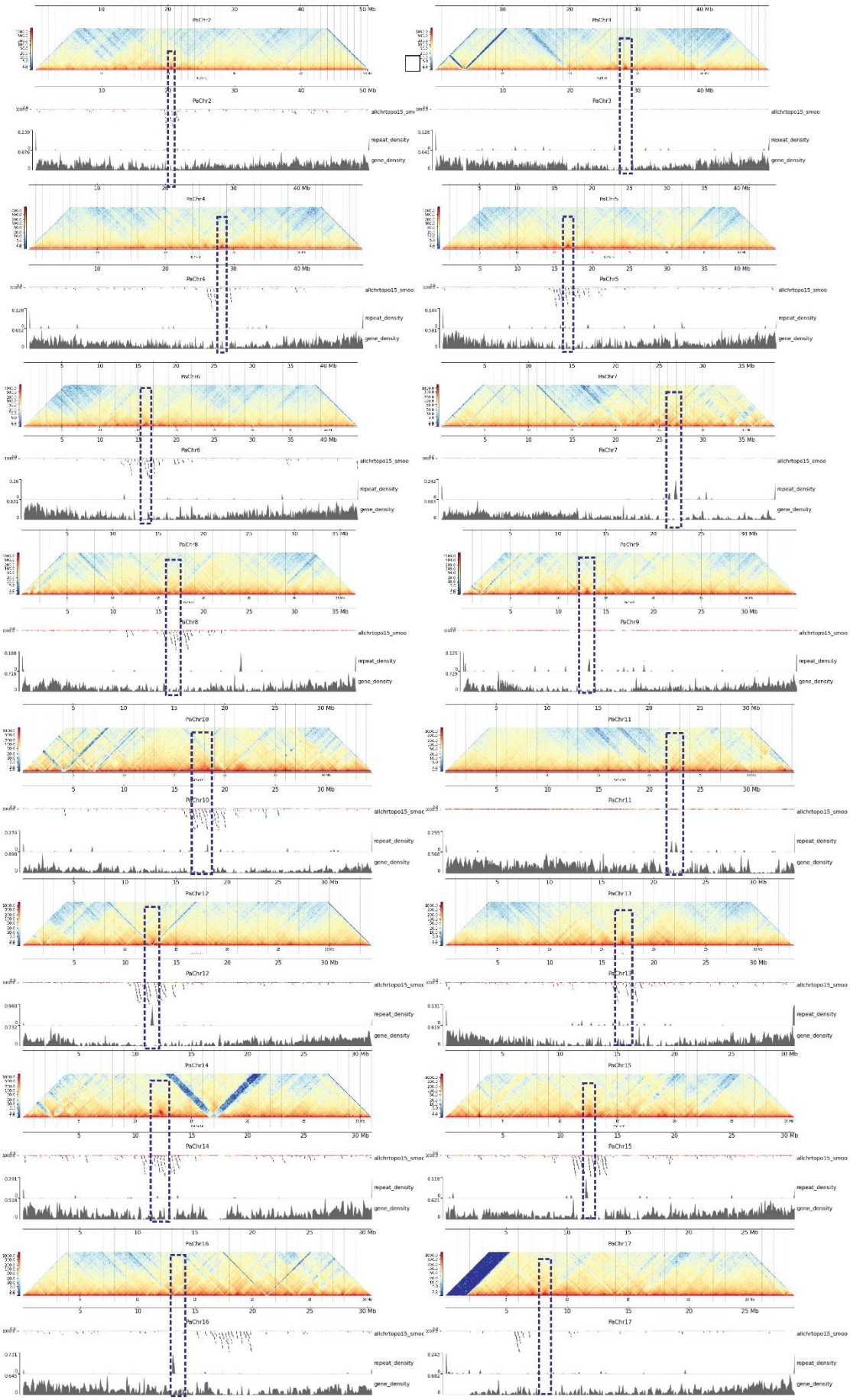
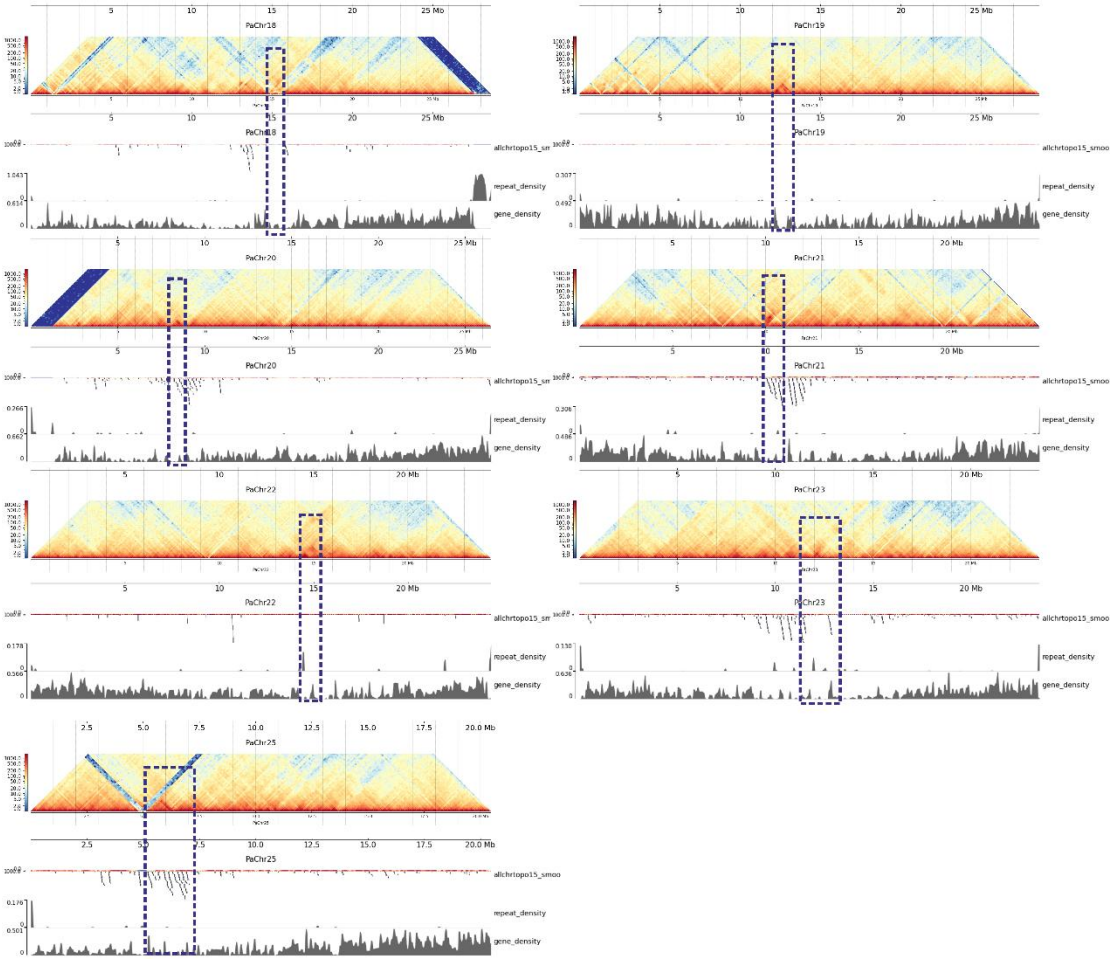


Figure S3. ADMIXTURE analysis of population structure in *Phragmites australis* (K = 3–6). Ancestry proportions estimated by ADMIXTURE for K = 3 to K = 6 clusters. Each vertical bar represents an individual, with colours indicating the proportion of ancestry assigned to each inferred cluster; individuals are grouped along the x-axis by geographic locations, with sample IDs shown at the bottom.

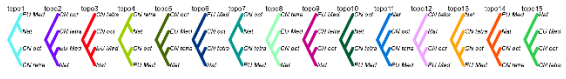


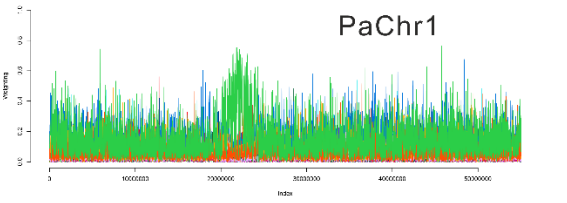
Supplementary Figure S4. Predicted centromere positions across the *P. australis* genome. Predicted centromere regions for the 25 pseudochromosomes identified using CentIER. Gray bars represent pseudochromosomes, and green blocks indicate the predicted centromere regions. Dark green blocks denote secondary candidate centromere regions where multiple predictions were obtained. Dashed bars indicate chromosomes for which no centromere could be confidently predicted. The labels acrocentric and telocentric indicate chromosome morphology inferred from the predicted centromere positions. The scale bar represents 10 Mb.

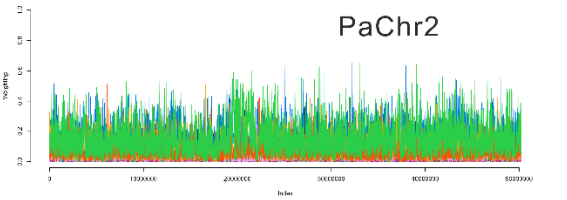

PaChr2
PaChr3
PaChr4
PaChr5
PaChr6
PaChr7
PaChr8
PaChr9
PaChr10
PaChr11
PaChr12
PaChr13
PaChr14
PaChr15
PaChr16
PaChr17
allchrtopo15_smoo
repeat_density
gene_density

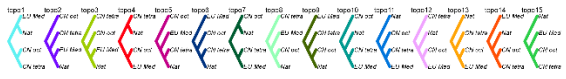


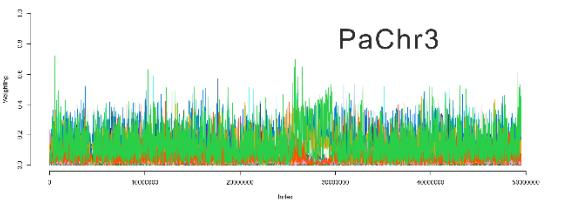
Supplementary Figure S5. Hi-C contact maps supporting centromere identification on representative pseudochromosomes. Hi-C contact maps of representative pseudochromosomes are shown together with repeat density and gene density tracks. Blue dashed boxes indicate the centromere regions predicted by CentIER. The predicted centromeres coincide with local Hi-C interaction patterns and are characterized by elevated repeat density and reduced gene density, providing independent support for centromere localization.

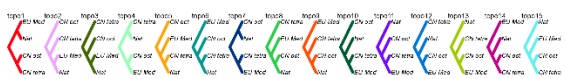


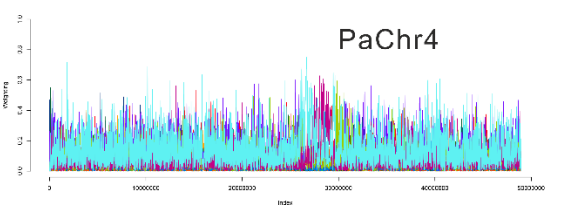

PaChr1
Weighting
Index

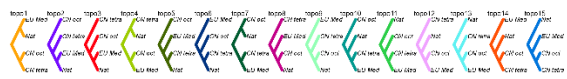

PaChr2
Weighting
Index

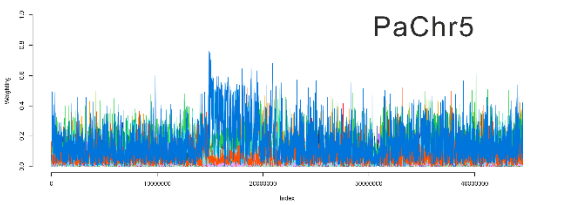


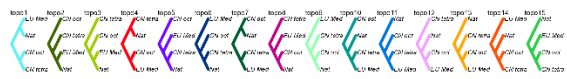

PaChr3
Weighting
Index

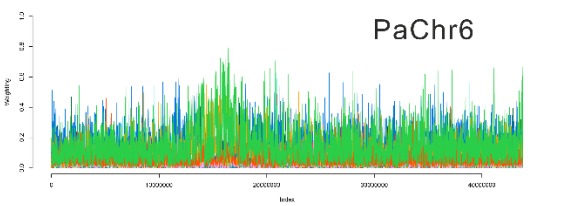


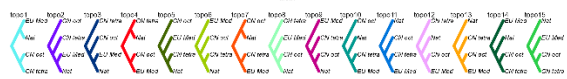

PaChr4
Weighting
Index

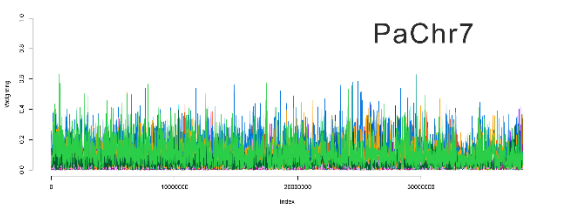


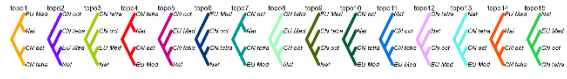

PaChr5
Weighting
Index

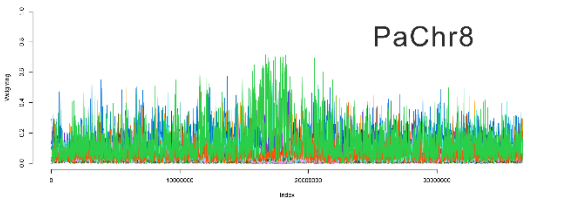


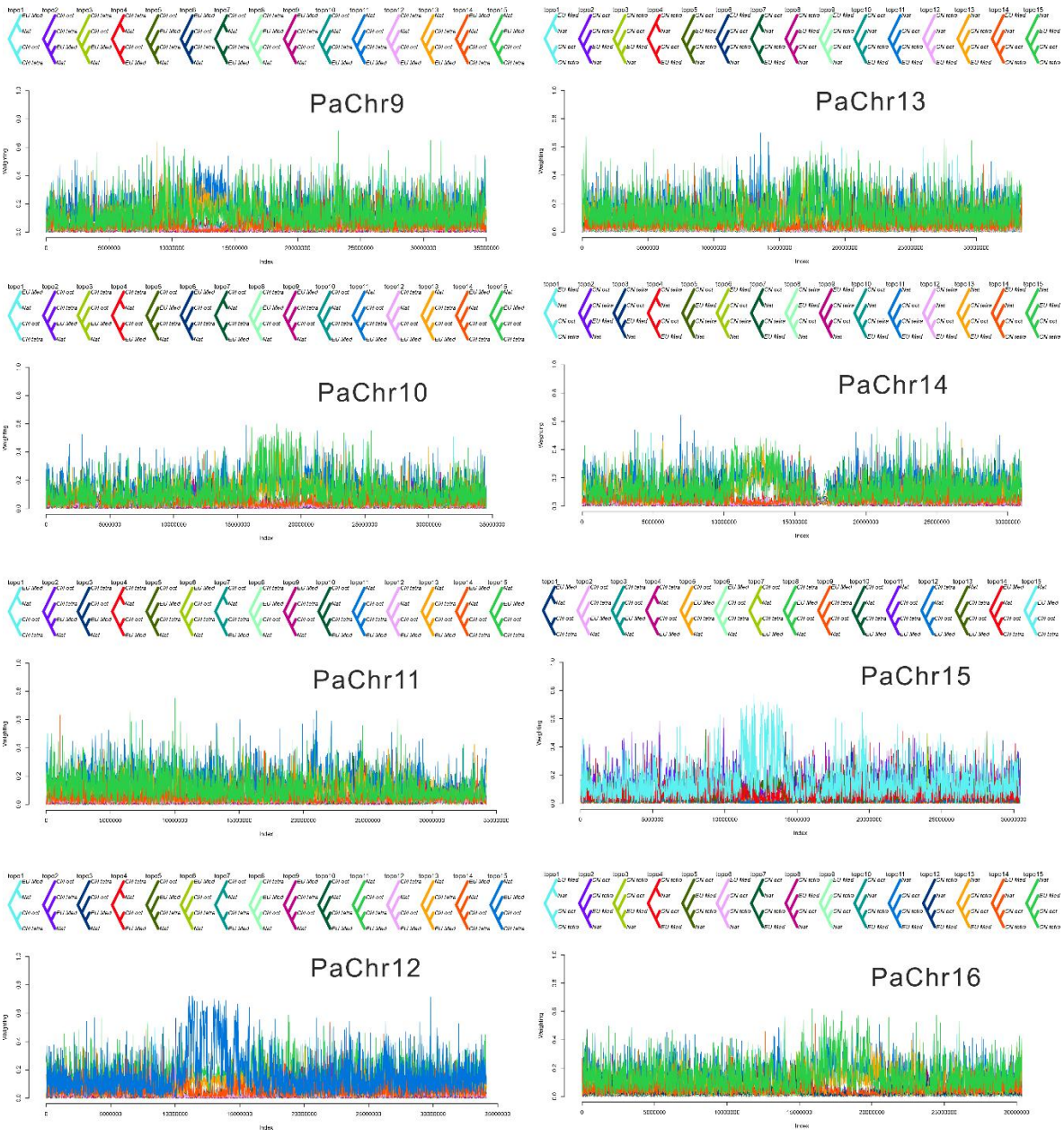

PaChr6
Weighting
Index

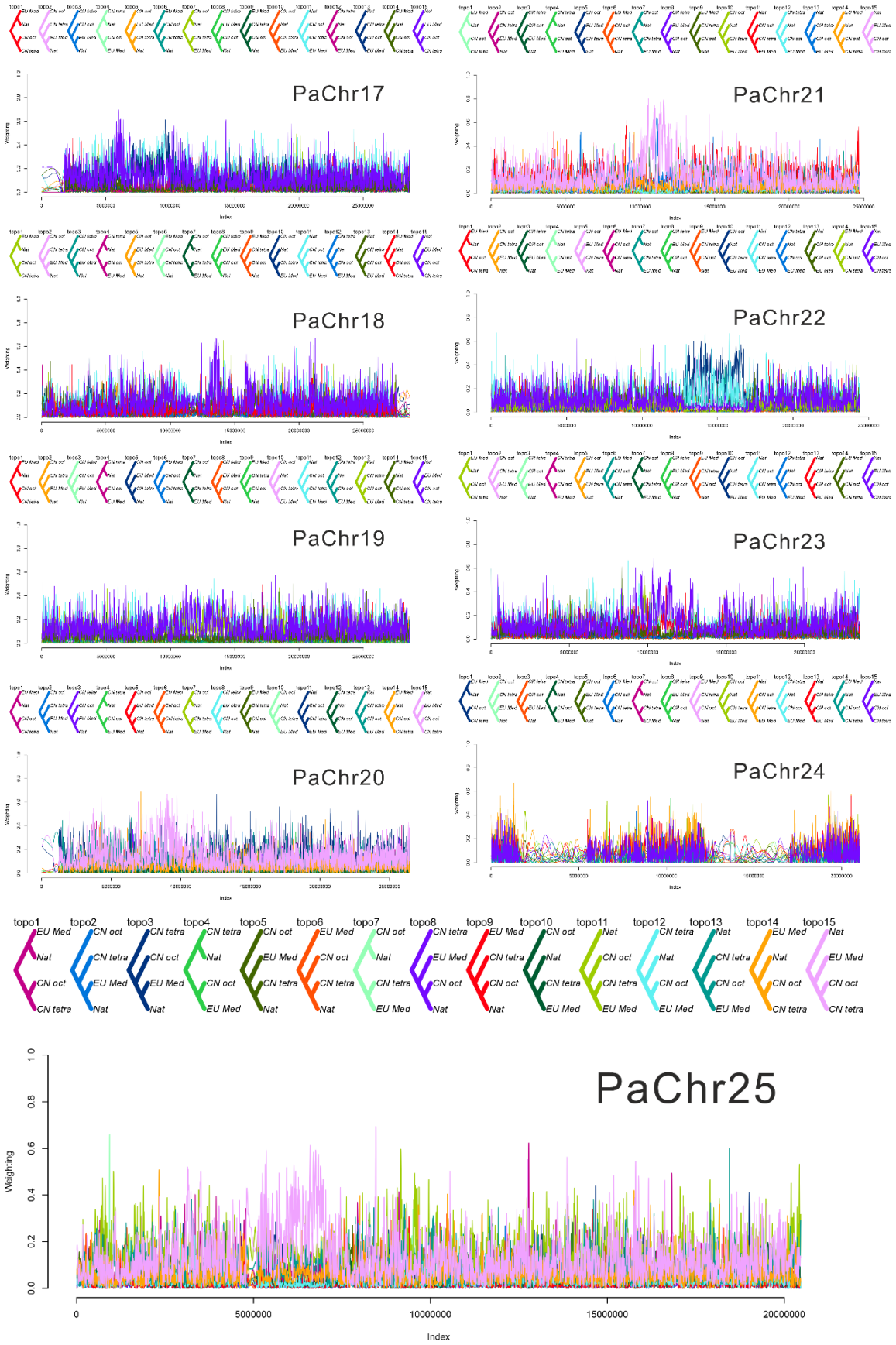



PaChr7
Weighting
Index



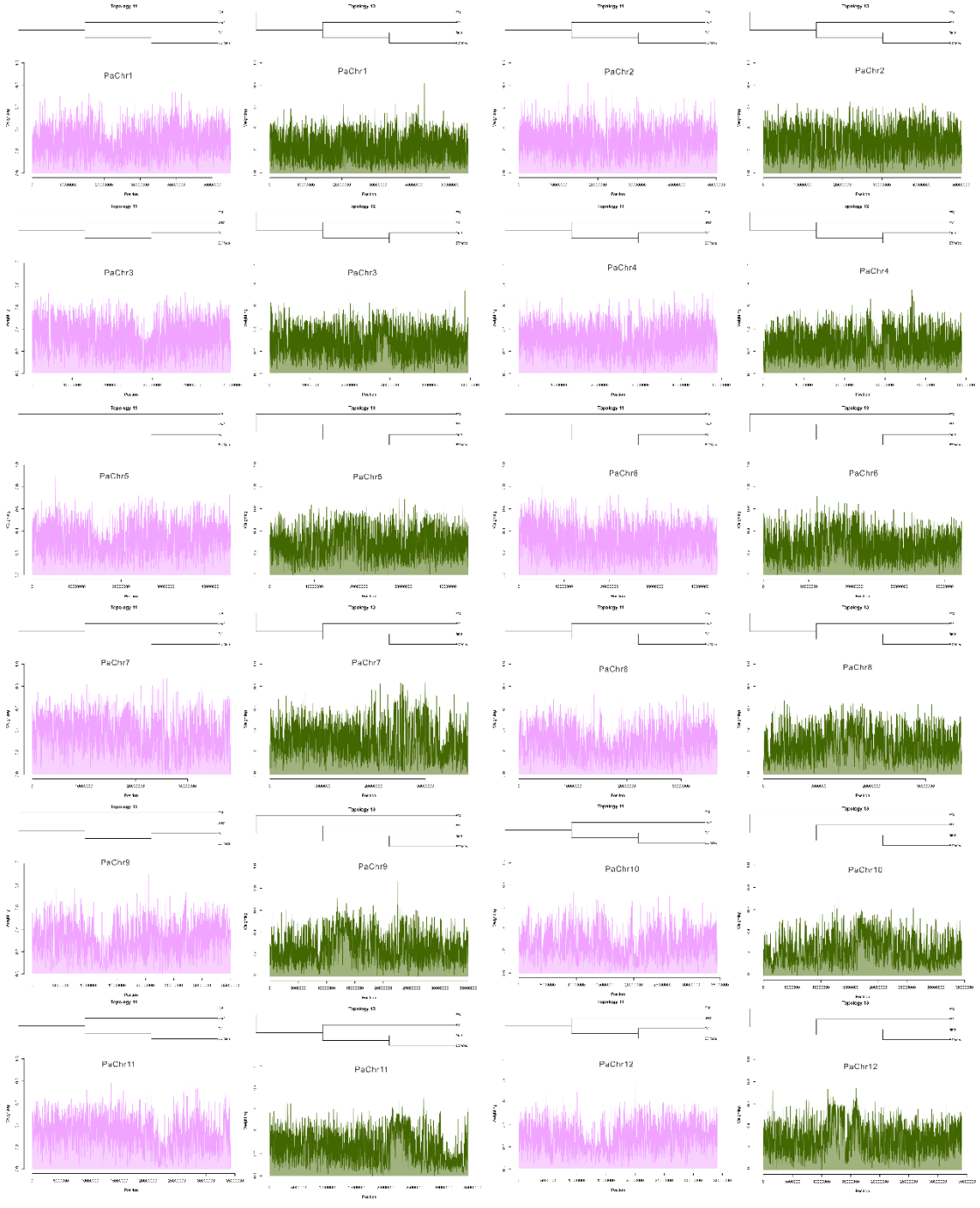

PaChr8
Weighting
Index

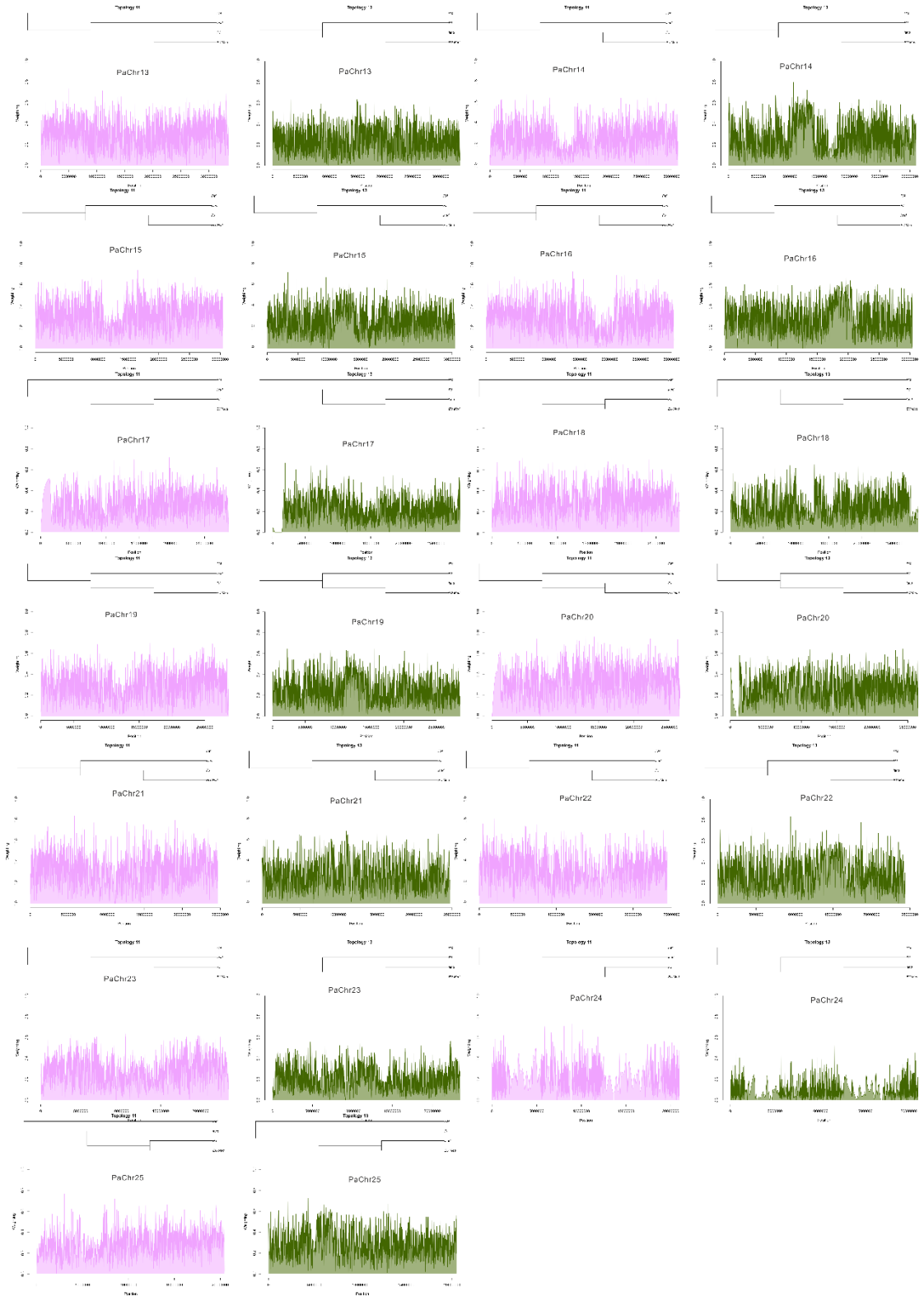

PaChr9
PaChr13
PaChr10
PaChr14
PaChr11
PaChr15
PaChr12
PaChr16



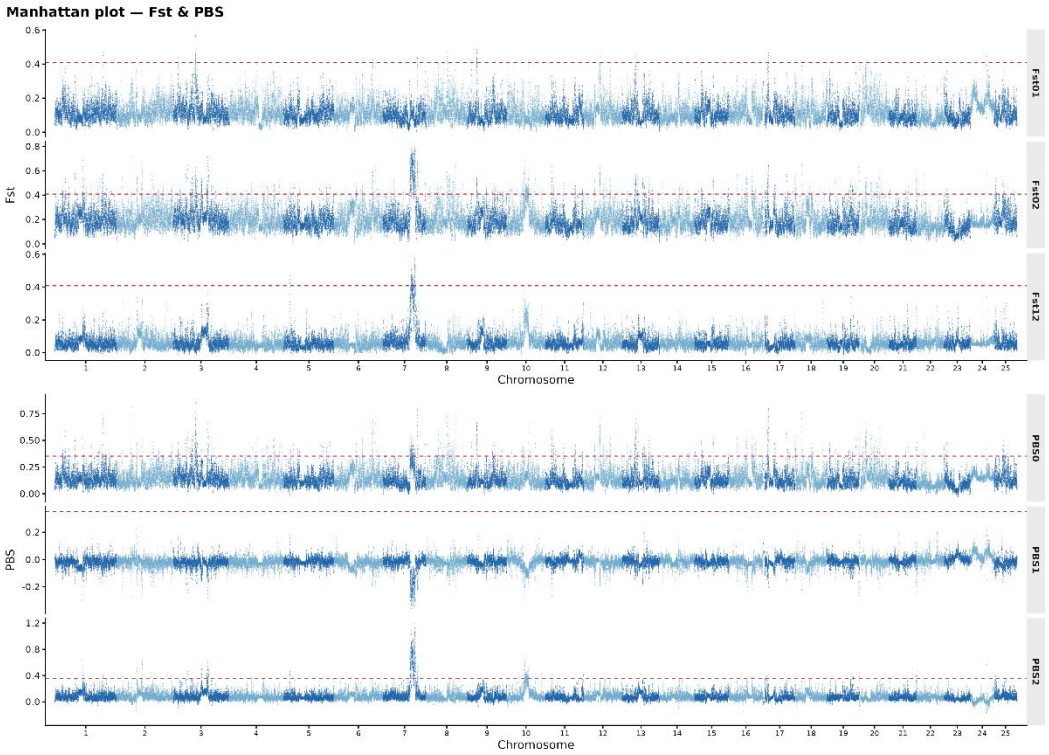
Supplementary Figure S6. Evidence of introgression from the tetraploid CN lineage into the octoploid AU lineage. Topology weighting analysis across PaChr1-25. The upper panel

shows the 15 alternative topologies evaluated, and the lower panel shows topology weighting along the chromosome. Topology 15, corresponding to (NAnat,(EUMed,(CNoct,CNtetra))), is consistent with introgression from the tetraploid CN lineage into the octoploid CN lineage.

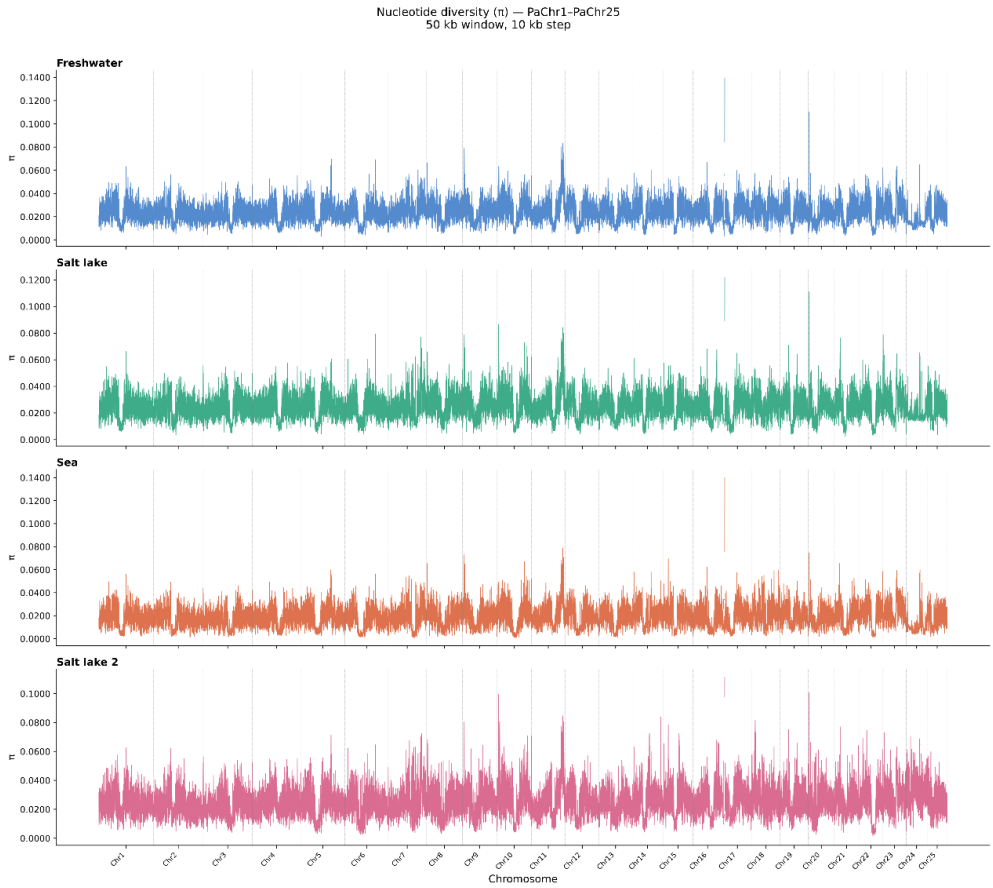




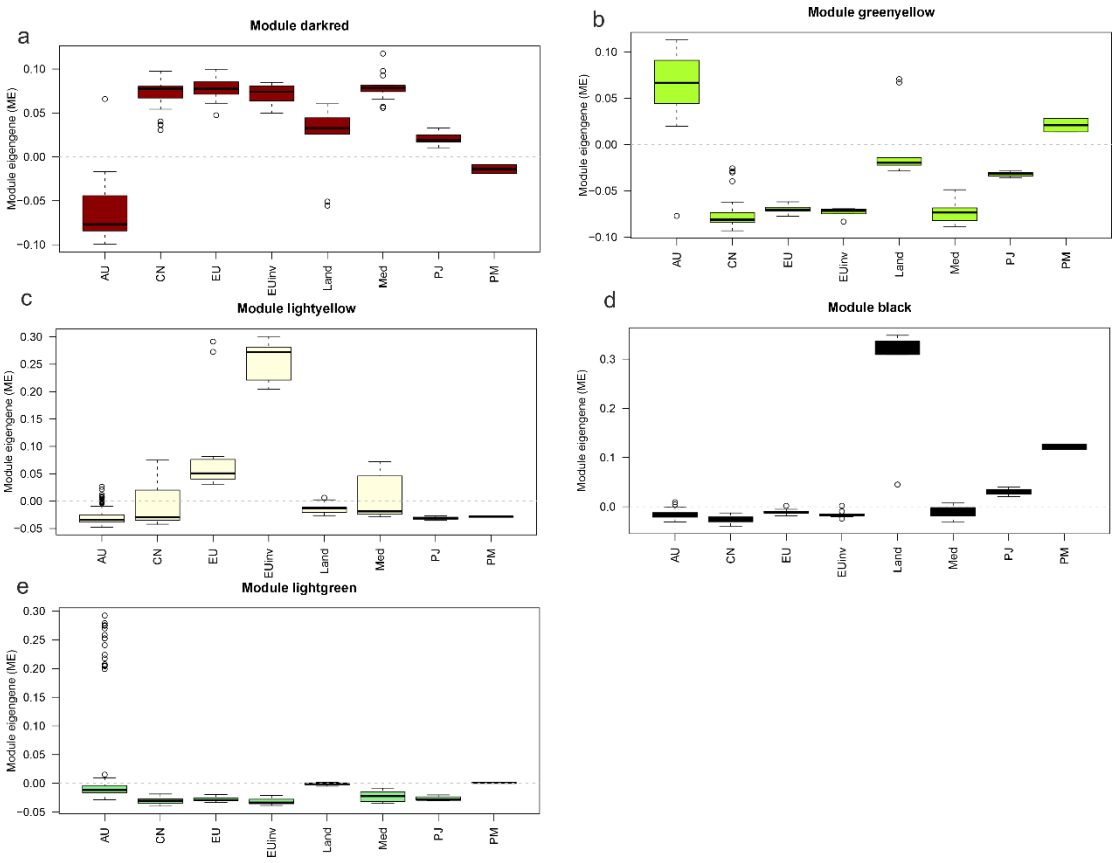
Supplementary Figure S7. Chromosome-wide Twisst analysis supporting the hybrid origin of the USland lineage. Topology weighting (Twisst) was performed for each pseudochromosome. The two topologies with the highest genome-wide support are shown above each weighting profile. Variation in topology weighting along the chromosomes indicates mosaic ancestry resulting from introgression between the Med lineage and *P. mauritianus*, leading to the formation of the USland lineage.



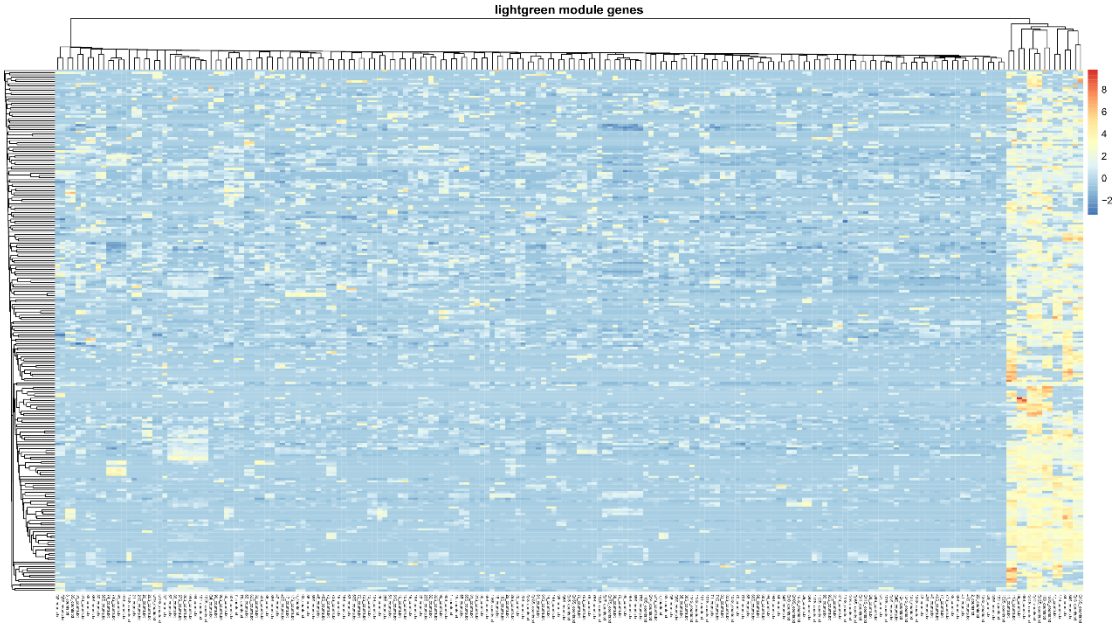
Supplementary Figure S8. Genome-wide FST and PBS among the CN freshwater, CN salt-lake, and CN seawater populations. Manhattan plots of pairwise FST (upper panels) and population branch statistic (PBS; lower panels) calculated across the genome. Population indices are defined as follows: **0**, CN salt-lake; **1**, CN freshwater; and **2**, CN seawater. Pairwise FST comparisons are shown for FST01, FST02, and FST12, and PBS values are shown for PBS0, PBS1, and PBS2. Red dashed lines indicate the top 1% empirical threshold.



Supplementary Figure S9. Genome-wide nucleotide diversity (π) of the four ecotypes. Genome-wide nucleotide diversity (π) was estimated for the freshwater, salt-lake, seawater, and salt-lake 2 populations using a 50-kb sliding window with a 10-kb step size. Values are plotted across all 25 pseudochromosomes.



Supplementary Figure S10. Module eigengene (ME) values of representative co-expression modules across populations. Module eigengene (ME) values of the **darkred** (a), **greenyellow** (b), **lightyellow** (c), **black** (d), and **lightgreen** (e) co-expression modules across populations (AU, CN, EU, EUinv, Land, Med, PJ, and PM). Positive and negative ME values indicate relatively higher and lower module expression, respectively. Boxes represent the median and interquartile range (IQR), whiskers extend to 1.5 × IQR, and points beyond the whiskers represent outliers. Elevated ME values in the lightgreen module are observed in a subset of Australian (AU) samples and are independent of salt treatment.



Supplementary Figure S11. Heatmap of gene expression in the lightgreen co-expression module. Heatmap showing the standardized (Z-score) expression levels of genes assigned to the lightgreen co-expression module across all samples. Genes and samples were hierarchically clustered according to their expression profiles. Warmer colors indicate higher relative expression, whereas cooler colors indicate lower relative expression.