Adaptation in Unstable Environments and Global Gene Losses: Small but Stable Gene Networks by the May–Wigner Theory

Shaohua Xu,[1,2,†] Shao Shao,[1,†] Xiao Feng,[1] Sen Li,[1] Lingjie Zhang,[1] Weihong Wu,[1] Min Liu,[1] Miles E. Tracy,[1] Cairong Zhong,[3] Zixiao Guo,[1] Chung-I Wu,[1] Suhua Shi,[1,*] and Ziwen He[1,*]

[1]State Key Laboratory of Biocontrol and Guangdong Provincial Key Laboratory of Plant Resources, School of Life Sciences, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-sen University, Guangzhou, China

[2]School of Ecology, Sun Yat-sen University, Shenzhen, China

[3]Institute of Wetland Research, Hainan Academy of Forestry (Hainan Academy of Mangrove), Haikou, China

[†]These authors contributed equally to this work.

***Corresponding authors:** E-mails: heziwen@mail.sysu.edu.cn; lsssh@mail.sysu.edu.cn

Associate editor: Michael Purugganan

Abstract

Although gene loss is common in evolution, it remains unclear whether it is an adaptive process. In a survey of seven major mangrove clades that are woody plants in the intertidal zones of daily environmental perturbations, we noticed that they generally evolved reduced gene numbers. We then focused on the largest clade of Rhizophoreae and observed the continual gene set reduction in each of the eight species. A great majority of gene losses are concentrated on environmental interaction processes, presumably to cope with the constant fluctuations in the tidal environments. Genes of the general processes for woody plants are largely retained. In particular, fewer gene losses are found in physiological traits such as viviparous seeds, high salinity, and high tannin content. Given the broad and continual genome reductions, we propose the May–Wigner theory (MWT) of system stability as a possible mechanism. In MWT, the most effective solution for buffering continual perturbations is to reduce the size of the system (or to weaken the total genic interactions). Mangroves are unique as immovable inhabitants of the compound environments in the land–sea interface, where environmental gradients (such as salinity) fluctuate constantly, often drastically. Extending MWT to gene regulatory network (GRN), computer simulations and transcriptome analyses support the stabilizing effects of smaller gene sets in mangroves vis-à-vis inland plants. In summary, we show the adaptive significance of gene losses in mangrove plants, including the specific role of promoting phenotype innovation and a general role in stabilizing GRN in unstable environments as predicted by MWT.

***Key words:* gene loss, mangrove, unstable environment, adaptive evolution, genome, network stability.**

Introduction

Identifying genetic changes underlying environmental adaptation is one of the central topics of evolution. A widely acknowledged genetic modification is the increase of genetic materials via gene duplication or even whole-genome duplication (WGD), which may promote adaptation by increasing gene expression or evolving new functions (Van de Peer et al. 2017; Ren et al. 2018; Wu et al. 2019; Xu et al. 2023; Feng et al. 2024). In contrast, gene losses are perceived as a process of devolution, the disuse of specific genes and traits that are costly to maintain. The examples include parasitic and domesticated plants, which experienced extensive gene number reductions involved in noncarnivorous nutrient acquisition and photosynthesis (Ibarra-Laclette et al. 2013; Carretero-Paulet et al. 2015; Palfalvi et al. 2020).

In recent years, the frequently occurring gene losses in evolution have been increasingly paid attention. Many cases have shown that gene loss, as a type of gene copy number variation that changes genic dosage, has great potential in adaptive evolution, known as the hypothesis of "less is more" (Albalat and Cañestro 2016; Xu et al. 2019). For example, losses of specific genes are associated with adaptation to pollinator changes, agronomic trait evolution, environmental adaptation, and the transfer from outcrossing to selfing (Hoballah et al. 2007; Sas et al. 2016; Wu et al. 2017; Helsen et al. 2020; Monroe et al. 2021).

Increasing studies have found much more gene loss than has been documented as adaptive. Whether the

Received: January 12, 2024. **Revised:** March 07, 2024. **Accepted:** March 15, 2024

large-scale gene losses during evolution are neutral or adaptive is still debated. What may be difficult to explain is that the generalized gene losses do not appear likely to be associated with environments. The examples include the thousands of gene losses in metazoan and bacteria and the extensive gene losses in plants after experiencing WGDs (Wolf et al. 2012; Albalat and Cañestro 2016; Van de Peer et al. 2017). Then, what might explain the adaptive potential of extensive gene loss?

Here, we note a hint of general gene loss of mangroves. Mangroves are woody plants that inhabit intertidal environments. As the dynamic land–sea interface, the intertidal zone is one of the most unstable environments exhibiting fluctuating salinity, water flooding, etc. (Lin et al. 1993; Tomlinson 2016). While unstable environments are fairly common in nature, intertidal zones, being compound environments between the land and sea, are almost unique. Environmental variables in compound environments can fluctuate more widely than in other habitats, particularly in the variance and kurtosis (tailedness) of many environmental gradients.

Since mangrove plants have tolerated a wide range of environmental stresses, it is tempting to assume that dealing with multiple stresses and innovation of phenotypic traits are accomplished by increasing genomic complexity (Duke 1992; Shi et al. 2005; Liang et al. 2008). However, previous studies indicated both genome sizes and gene sets have been simplified (Lyu et al. 2018; He et al. 2020; Xu et al. 2020). Using mangroves as models, we plan to investigate the possible driving forces explaining large-scale gene losses. Based on previous studies, the number of gene losses associated with phenotypic trait evolution should be small. Therefore, we hypothesize that the most unstable environments inhabited by mangroves could be a primary driving force. Maintaining gene regulatory network (GRN) stability under continuous perturbation may be the primary challenge in highly fluctuating intertidal habitats.

With multiple stressors and high fluctuation in intertidal zones, the perturbation of mangroves' GRN is more significant than in inland plants (Liang et al. 2012; Cruz et al. 2020; Feng et al. 2020; Su et al. 2022; Xu et al. 2023; see Discussion). We explore whether the extensive gene loss has contributed to GRN stabilization, defined as every node's ability to return to equilibrium after perturbation. A classical theory describing ecological network stability is the May–Wigner theory (MWT), which suggests large and highly connected networks such as the food webs are unstable (May 1972; Allesina and Tang 2012). As network connectivity interaction strength or the number of variables increases, instability also tends to increase. Interestingly, although food webs are the first biological systems treated with MWT, it has been suggested that GRNs, as well as many nonbiological systems like electric grids and international banking, fit the theory much more closely (Chen et al. 2019). The gene loss, corresponding to smaller networks, can potentially increase GRN stability.

In this study, we conduct a comparative genomic analysis in Rhizophoreae mangroves, as part of a large worldwide mangrove genome sequencing project (He et al. 2022). We describe four de novo assembled mangrove genomes and one inland species. Rhizophoreae is a typical mangrove clade that comprises 18 or ~26% of the total mangrove species (Shi et al. 2005; Tomlinson 2016; Xu et al. 2017). Having this many species allows for robust inference of gene number evolution and enhanced power to identify adaptive gene losses. We also found a common trend of gene number reduction in seven independent mangrove clades. The phenomenon enables us to explore the adaptive role of gene loss, first identifying gene losses that appear to be associated with key adaptive traits, then testing whether the extensive gene loss could be explained by increased GRN stability.

Results

The Genome and Phylogeny of Rhizophoreae Mangroves

We sequenced the genomes of four Rhizophoreae mangrove species (*Ceriops tagal*, *Ceriops zippeliana*, *Kandelia candel*, and *Rhizophora mangle*) and a closely related inland species *Carallia pectinifolia* using a combination of PacBio Single-Molecule Real-Time (SMRT) sequencing, 10x Genomics, and high-throughput chromosome conformation capture (Hi-C) technology (supplementary table S1, Supplementary Material online; Materials and Methods). The genomes are of high quality and complete and were conducted with functional annotation (Table 1; supplementary tables S2 to S6 and figs. S1 to S8, Supplementary Material online).

We then examined the species divergence of the Rhizophoreae clade into four genera: *Rhizophora* (six species), *Bruguiera* (five species), *Ceriops* (five species), and *Kandelia* (two species; Rivera-Monroy et al. 2017). The phylogenetic tree was reconstructed using sequences of single-copy gene families, and the divergence time was calibrated with fossils (supplementary fig. S9, Supplementary Material online; Materials and Methods; Yang 2007). The invasion of the intertidal zone is inferred to have occurred 50.8 to 47.8 million years ago (Ma), in agreement with previous conclusions (Fig. 1; Xu et al. 2017; He et al. 2020). This period coincides with early Eocene hyperthermal events of 55 to 51 Ma (Zachos et al. 2008), when the rising sea levels may force the Rhizophoreae inland ancestors to invade intertidal zones. Then species divergence of Rhizophoreae takes place in two hot spots both correlated with climatic optimum events: the among-genus divergence coincides with the Mid-Eocene Climatic Optimum (MECO: ~42 Ma) and the speciation within the genus happened during a Mid-Miocene Climatic Optimum (MMCO: ~15 Ma). Sea-level elevation during these events may have promoted Rhizophoreae mangrove dispersal and expanded their habitats. Subsequent temperature drops lowered sea levels and isolated the surviving populations, promoting their evolution into distinct species. The results also indicated the

Table 1 Genome assembly statistics of six Rhizophoraceae species

	Carallia pectinifolia	*Kandelia candel*	*Ceriops zippeliana*	*Ceriops tagal*	*Rhizophora mangle*	*Rhizophora apiculata*
Sequencing strategy	SMRT + Hi-C	SMRT	SMRT	10x Genomics + Hi-C	SMRT	Hi-C
Estimated genome size (*k*-mer; Mb)	355	278	324	283	380	274
Total assembly length (Mb)	324.51	213.51	320.06	229.80	281.00	232.00
Number of scaffolds	111	474	2,436	5,177	2,379	244
N50 (Mb)	29.47	9.45	3.46	11.73	1.93	13.31
BUSCO completeness (%)	96.3	97.5	95.2	95.3	96.5	97.8
GC content (%)	35.41	35.57	34.15	37.00	33.50	35.03
Repeats (%)	39.69	31.86	56.23	32.77	46.62	34.63
Number of genes	31,896	24,295	23,158	21,450	25,466	23,565
Annotated genes (%)	89.85	97.58	96.12	99.28	93.52	98.71

high fluctuation of intertidal zones and their extensive influence on mangroves' survival and divergence.

Accelerated Gene Loss and Depressed Gene Duplication in Rhizophoreae Mangroves

We found Rhizophoreae mangrove genomes contained 21,450 to 25,466 protein-coding genes, many fewer than the 31,221 to 42,950 in their inland relatives (Fig. 1; supplementary tables S7 and S8, Supplementary Material online). In a wide-ranging comparison with 82 core eudicot plants, the eight Rhizophoreae mangroves are among the most gene poor (supplementary fig. S10, Supplementary Material online). Compared to their inland relative *C. pectinifolia* that shares the exact WGD event, Rhizophoreae mangroves have about 8,400 fewer genes on average.

A decrease in gene number can result from accelerated gene loss or slowed origination. We assessed the gene duplication rate by examining tandem gene duplications. On average, 5,101 (3,550 to 6,394) tandem duplications were identified in four inland species of Malpighiales, but only 1,603 (919 to 2,326) in the eight Rhizophoreae mangroves. Since tandem gene duplications usually occurred recently, their rareness indicated a lower gene duplication rate. We roughly assessed gene loss by identifying gene loss after WGDs. Rhizophoreae mangroves (an average of 9,919) retained fewer gene duplications than *C. pectinifolia* (13,357) that shared the same WGD event, indicating a higher gene loss rate. Therefore, accelerated gene loss and depressed gene duplication reduced the gene number in Rhizophoreae mangroves.

The Common Trend of Extensive Gene Number Reduction among Multiple Mangrove Clades

We surveyed the gene content of six other independent mangrove clades in the Rosids in addition to Rhizophoreae (supplementary table S7, Supplementary Material online). In the analyzed 20 mangrove genomes from the seven clades, 19 have fewer gene numbers than inland relatives. The only exception is the mangrove *Pemphis acidula*, which experienced a recent species-specific WGD event. In total, we found a significant trend of gene number decrease in these independent mangrove clades (Fig. 2; Wilcoxon signed-rank test; $P = 1.04 \times 10^{-7}$). The trend is conserved among mangroves in the same clade and is unaffected by the different evolutionary scenarios among mangrove clades. Rhizophoreae and *Sonneratia* have experienced the same WGD events with their inland relatives; *Laguncularia racemosa* and *Lumnitzera* genus have no recent WGDs. They all have fewer genes than their inland relatives. Furthermore, the six mangrove clades have lower tandem gene duplication rates (supplementary fig. S11, Supplementary Material online; Wilcoxon signed-rank test; $P = 1.47 \times 10^{-4}$). We found all the genomes surveyed have similarly high gene annotation completeness, and the examination of expressed genes in the leaf again showed smaller gene numbers in mangroves (supplementary figs. S12 and S13, Supplementary Material online). It confirmed the common trend of gene number reduction in mangrove clades. The common trend makes it reasonable to hypothesize that gene number reduction is an adaptive strategy in environmental evolution.

We examined whether the theories explaining previously reported large-scale gene losses in plants could also explain the phenomenon in mangroves. The simplest way to make such distinct differences in gene numbers is large-scale gene duplication, such as WGDs. However, mangroves did not experience fewer WGD events than their inland relatives (He et al. 2020; Hu et al. 2023). The most dramatic gene loss reported is in parasitic plants and aquatic plants, among which the extensive gene losses, including the essential ones, are explained by regressive evolution that refers to the loss of useless characteristics over time. However, the phenomenon in mangroves does not fit the theory since the essential genes are usually retained in mangroves (see sections The Genome and Phylogeny of Rhizophoreae Mangroves and The Patterns of Gene Number Reduction). Another possible benefit of decreasing gene numbers is a lower energy cost, leading to smaller genome sizes (Kang et al. 2015). However, the mangroves of Combretaceae have fewer genes but comparable genome sizes to inland relatives (Xie et al. 2023; Zhu et al. 2023). It indicated that the constraint of energy cost might not be the primary driving force, though the intertidal zones are usually nutritionally deficient (Reef et al. 2010). Since none of the above theories could explain the phenomenon in mangroves, we take advantage of the

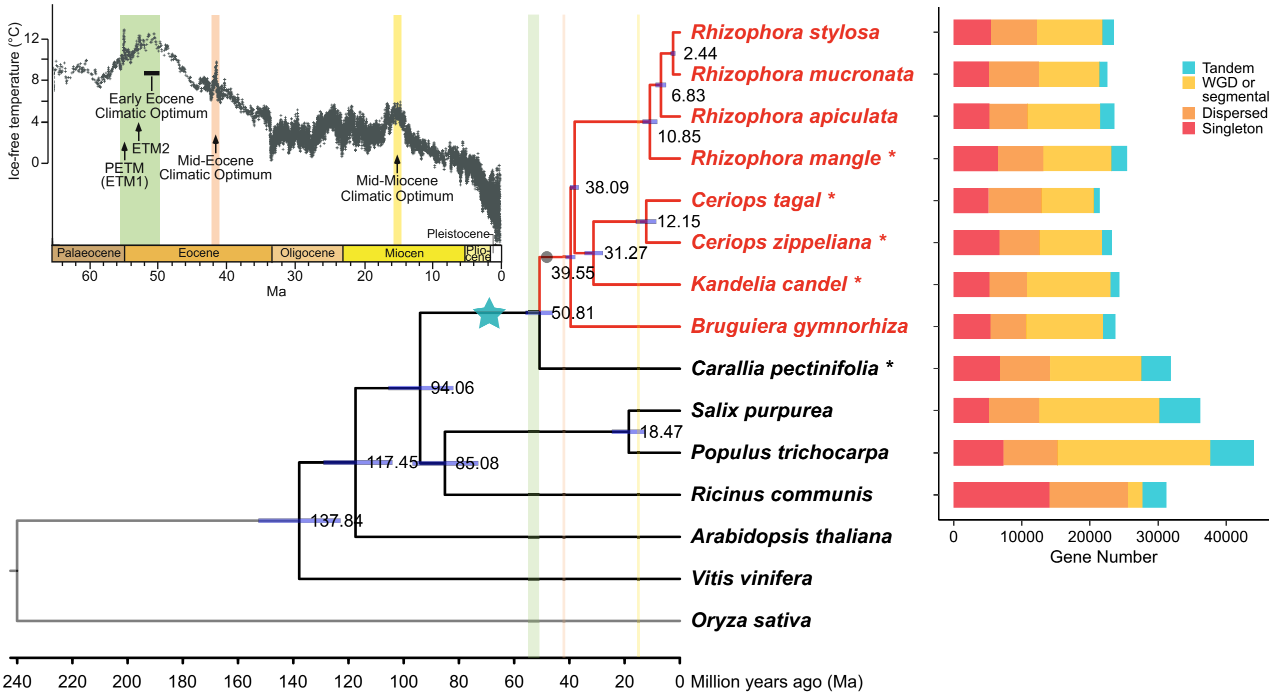


Fig. 1. Origination and gene family reduction of Rhizophoreae mangroves. The origination, among-genera divergence, and within-genus divergence of Rhizophoreae mangroves coincided with three climatic optimum events in the upper left panel. The estimated divergence times are displayed near the nodes, with bars representing 95% confidence intervals. The WGD event shared by Rhizophoreae is marked with a star. The gray dot represents the lower time boundary of the earliest fossil of the common ancestor of Rhizophoreae mangroves (47.8 Ma). Rhizophoreae mangroves are indicated with red font. De novo assembled genomes in this study are marked with asterisks. The upper left panel shows historical temperature changes over 65 Myr (modified from Zachos et al. 2008). Three climate warming events are highlighted. The right panel shows fewer duplicated genes in Rhizophoreae mangroves than in nonmangroves.

largest mangrove clade, Rhizophoreae, to explore the evolutionary forces driving the extensive gene losses.

Gene Loss and Trait Innovation

We first tested whether the extensive gene losses could be explained by adaptive trait evolution, known as the “less is more” hypothesis (Albalat and Cañestro 2016; Xu et al. 2019). Rhizophoreae mangroves have acquired specialized phenotypic traits such as viviparous seeds, high salinity tolerance, and high tannin content, greatly improving the survivability in the intertidal zone (Tomlinson 2016).

Gene Loss and Vivipary

The viviparous seed that germinates before abscission from mother plants is one of the most significant trait innovations of Rhizophoreae mangroves (Fig. 3a). It increases seed survival by isolating them from stressful environments and promotes long-distance dispersal by prolonging seed flooding time (Shi et al. 2005; Liang et al. 2008). We found fewer gene copies reported to participate in seed dormancy and maturation in Rhizophoreae mangroves than in their inland relatives (supplementary fig. S14, Supplementary Material online; Berardini et al. 2015). Overall, we found 66 reduced, but only nine expanded gene families in Rhizophoreae (supplementary table S9, Supplementary Material online).

The most significant reduction occurred in families that participated in seed reserve accumulation and desiccation tolerance, which are essential for seed maturation and dormancy. Four gene families that play roles in storage were almost completely lost or significantly reduced in Rhizophoreae mangroves (Fig. 3b and c; supplementary figs. S15 to S17, Supplementary Material online; Wan et al. 2007; López-Ribera et al. 2014; Miquel et al. 2014; Lunn et al. 2018). The genes coding LATE EMBRYOGENESIS ABUNDANT (LEA) proteins that are implicated in desiccation tolerance, such as *SEED MATURATION PROTEIN 1* (*SMP1*), have significantly fewer genes in Rhizophoreae mangroves (Fig. 3b; supplementary fig. S18 and table S9, Supplementary Material online; Hoekstra et al. 2001; Kushwaha et al. 2012). Other reduced gene families involved in seed maturation are summarized in the Supplementary Material (supplementary table S9 and fig. S19A, Supplementary Material online).

Pathways that regulate seed maturation and dormancy have also experienced gene loss in Rhizophoreae mangroves. The LAFL network is the critical component in the seed maturity regulation process and comprises LEAFY COTYLEDON 1 (LEC1)/LEC1-LIKE (L1L), LEC2, FUS3, and ABA-insensitive 3 (ABI3; Jia et al. 2013; Chen et al. 2020). Reduced *LEC1/2* in Rhizophoreae mangroves may have altered the network, causing seeds to stop accumulating storage materials and germination earlier

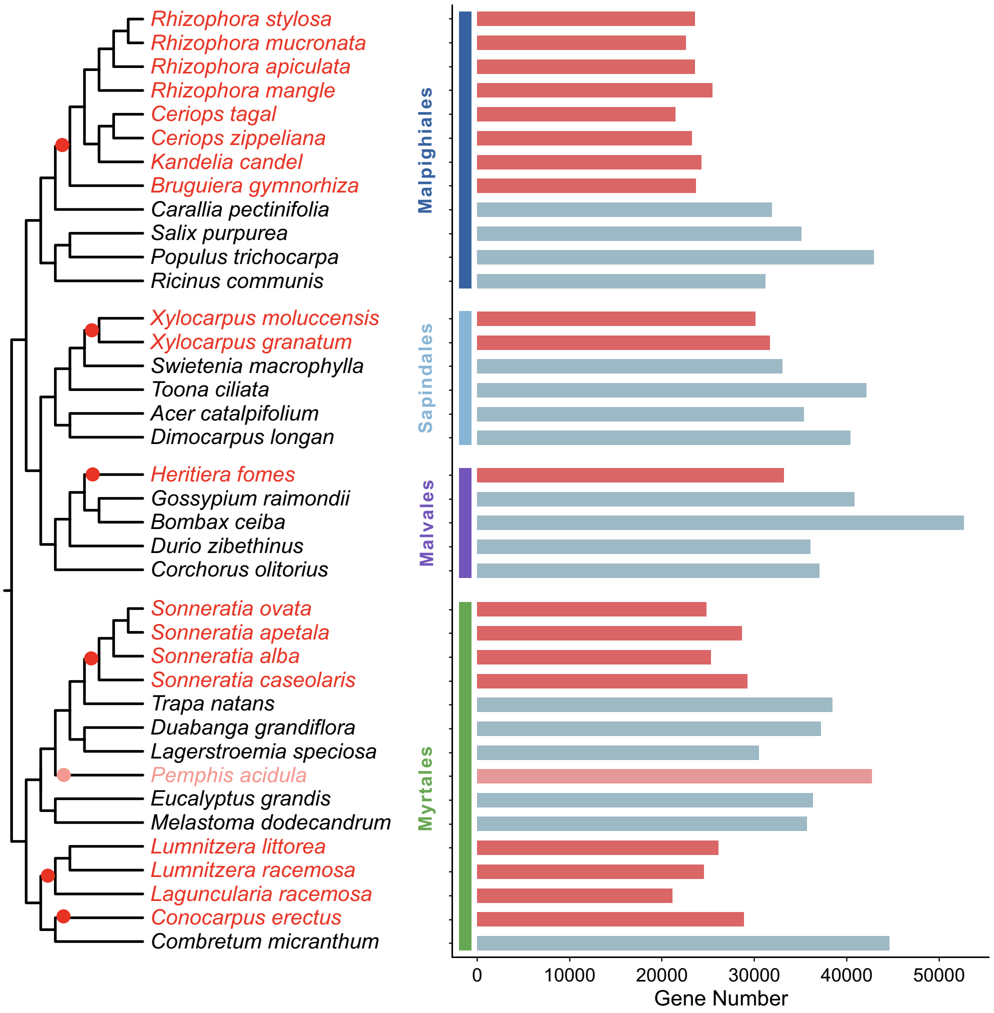


Fig. 2. Extensive gene reduction in independent mangrove clades. The red circles on the phylogenetic tree represent independent origin events of seven mangrove clades, with mangrove plants marked in red. The right panel showed the total gene number of each species. The mangrove clade *P. acidula* experienced a recent species-specific WGD event and was marked in pink.

(Chen et al. 2020). Abscisic acid (ABA) is the major phytohormone controlling seed dormancy. And we found genes that regulate ABA signaling, such as *DELAY OF GERMINATION 1* (*DOG1*) and *EARLY METHIONINE-LABELLED 6* (*GEA6*), are absent in Rhizophoreae mangroves (Fig. 3b and c; supplementary figs. S20 and S21 and table S10, Supplementary Material online). The DOG1 protein binds to AHG1 and releases SnRK2 to positive regulate ABA signaling and seed dormancy (Fig. 3c; Nishimura et al. 2018). The GEA6 protein protects the embryo at the drying stage (Carles et al. 2002). The gene loss in these pathways may block seed dormancy and result in viviparous seeds. Although a few genes such as *CYP707A1* that mediates ABA metabolism have expanded, gene family size reduction has generally played a more critical role than expansion in viviparous seed evolution of Rhizophoreae mangroves (Okamoto et al. 2006; supplementary fig. S19B, Supplementary Material online).

Gene Loss and High Salinity Tolerance

The gene loss also contributes to the tolerance of high salinity, one of the main challenges in intertidal zones. Among 540 gene families reported to respond to high salinity (Berardini et al. 2015), we found only six expanded such as the *HIGH-AFFINITY K+ TRANSPORTER 1* (*HKT1*) gene, and 43 shrunk families in Rhizophoreae mangroves (supplementary table S11, Supplementary Material online). The most significant reduction occurred in gene families that negatively regulate the ABA signaling pathway and hypertonic salinity response. For example, the *RESPONSE TO ABA AND SALT 1* (*RAS1*) gene, an ABA sensitivity enhancing negative regulator of salt tolerance during seed germination and early seedling growth (Ren et al. 2010), was lost in six Rhizophoreae mangroves.

Gene Loss and High Tannin Content

Tannins are common flavonoid polymers of plants. Its antioxidant activity is a major mechanism for dealing with oxidative stress in intertidal environments (Xu et al. 2017). We found that the leaves and stems of Rhizophoreae mangroves are significantly richer in tannins than inland relative: 168.8 and 302.4 mg/g in mangroves, and 82.5 and 185.3 mg/g in *C. pectinifolia* (Fig. 4b). Compared with inland relatives, we did not find significant copy number change of genes leading to tannin biosynthesis. At the expression level, we found that the expression of most key genes in the tannin synthesis pathway is much higher in Rhizophoreae mangroves, such as the *LAR* and *ANR* genes

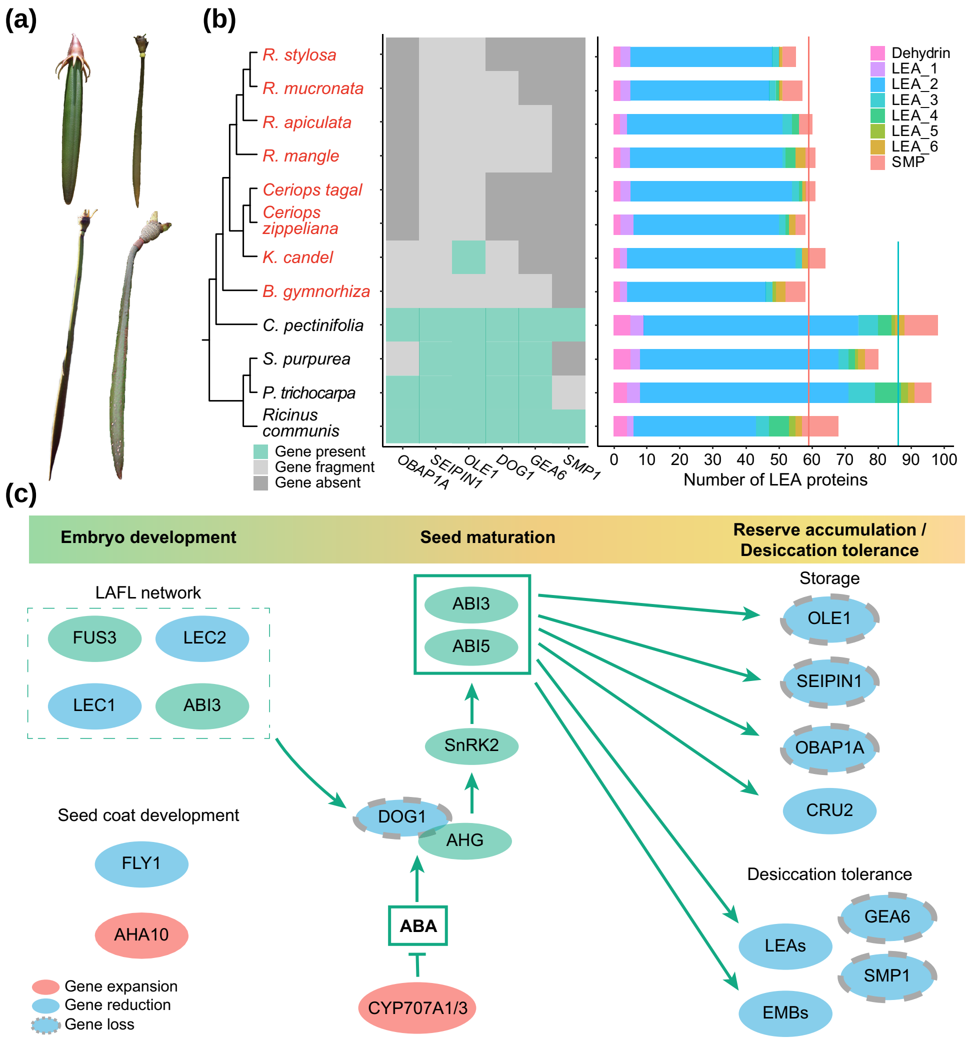


Fig. 3. Gene loss and viviparous seed evolution of Rhizophoreae mangroves. a) Viviparous propagules of Rhizophoreae mangroves. From upper left to bottom right: *B. gymnorhiza*, *Ceriops tagal*, *K. candel*, and *R. apiculata*. The photo source is Mangrove iD e-book (Duke 2017). b) Loss of genes related to seed development in Rhizophoreae mangroves. Six genes that were lost in Rhizophoreae mangroves are shown. The right panel shows the gene copy number of *LEA*. c) Diagram showing the possible mechanisms of vivipary in Rhizophoreae mangroves. Expansion and contraction of genes involved in embryo development, seed maturation, and reserve accumulation. Genes circled by elliptical dotted lines have been lost in Rhizophoreae mangroves.

(Fig. 4c; supplementary table S12, Supplementary Material online).

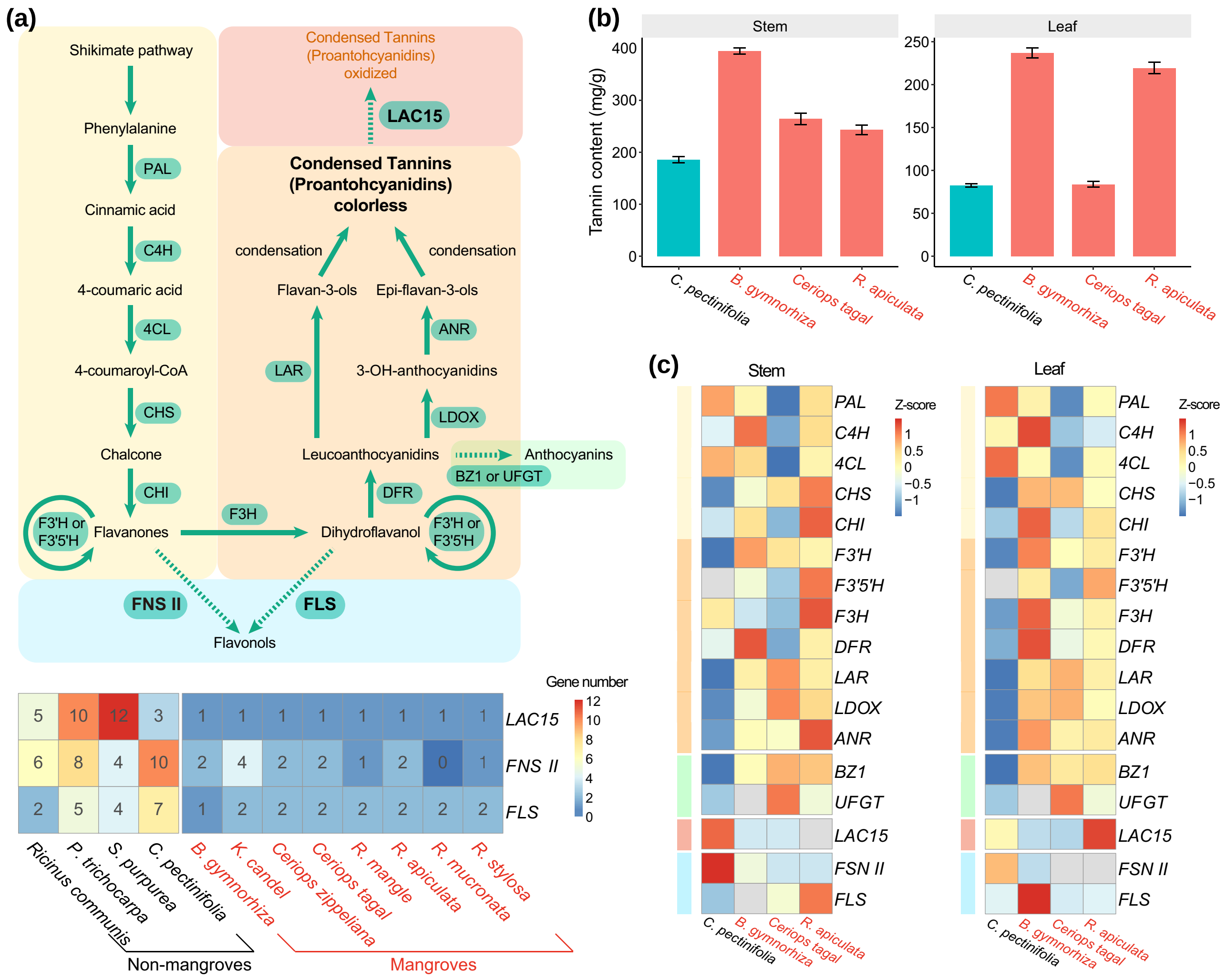
We identified significantly fewer copies of genes consuming tannins in Rhizophoreae mangroves (Fig. 4a; supplementary table S13 and fig. S22A and B, Supplementary Material online). The copy number reductions include genes encoding FLAVONE SYNTHASE II (FNS II) and FLAVONOL SYNTHASE (FLS), which consume precursors of condensed tannins to synthesize flavone and flavonol. Another significant reduction occurred in the *TRANSPARENT TESTA 10* (*TT10* or *LAC15*) that promotes oxidative polymerization of flavonoids. It amplified to 3 to 12 copies mainly by tandem duplication in inland relatives and retained only a single copy in Rhizophoreae mangroves (Fig. 4a; supplementary fig. S22C, Supplementary Material online). Reducing the *TT10* family may decelerate the oxidization of colorless tannins and contribute to accumulating more epicatechin monomers and soluble tannins (Pourcel et al. 2005). At the same time, expression levels of the *FNS II* gene in both leaves and stems and *TT10* in stems are significantly lower in Rhizophoreae mangroves than in *C. pectinifolia* (Fig. 4c). Therefore, high expression of tannin synthesis genes, low expression, and copy number reduction of genes that consume tannins or their precursors contribute to the high tannin content of Rhizophoreae mangroves.

The Patterns of Gene Number Reduction

Although the previous section confirmed the crucial role of gene loss in the innovation of phenotypic traits, most gene losses are not associated with these processes. To explain the extensive gene loss in mangroves, we examined what these gene losses are and when they occurred.

Gene Loss Minimizes Influence on Essential Processes

Dramatic gene number reductions have been reported in parasitic and carnivorous plants and seagrasses due to trait degeneration, including core loci conserved across most inland plants (Lee et al. 2018; Sun et al. 2018; Palfalvi et al. 2020; Cai et al. 2021). However, in Rhizophoreae mangroves, the majority (95.2% to 97.8%) of core eudicot genes can still be found with complete structure despite the small number of genes. Constrained by dosage balance, gene families involved in central processes and housekeeping functions are highly conserved in copy number and sequence, while large families are more likely to harbor redundant genes (Wang et al. 2018; Xu et al. 2019). We measured copy



	Ricinus communis	P. trichocarpa	S. purpurea	C. pectinifolia	B. gymnorhiza	K. candel	Ceriops zippeliana	Ceriops tagal	R. mangle	R. apiculata	R. mucronata	R. stylosa
LAC15	5	10	12	3	1	1	1	1	1	1	1	1
FNS II	6	8	4	10	2	4	2	2	1	2	0	1
FLS	2	5	4	7	1	2	2	2	2	2	2	2

Fig. 4. Molecular mechanisms of high tannin content in Rhizophoreae mangroves. a) Gene copy number evolution in the flavonoid biosynthesis pathway. The upper panel is a sketch map of the pathway. Genes that oxidize tannin or consume their precursors have fewer copies in Rhizophoreae mangroves and are highlighted in bold font. The bottom panel shows the copy number of *LAC15*, *FNS II*, and *FLS* genes in Rhizophoreae mangroves and nonmangroves. b) There is a higher tannin content in the stems and leaves of three Rhizophoreae mangroves than in *C. pectinifolia*. c) Expression of genes in the flavonoid biosynthesis pathway. The gene expression was first measured by TPM. TPM values of each species were then normalized to make a between-species comparison (Materials and Methods). Heatmap showing the normalized TPM values transformed into a Z-score. The gene grouping is consistent with the pathway in a).

number variability across species using the coefficient of variation (CV) among four inland plants. Duplicate gene loss was estimated using the fold change of mean copy number between mangroves and their inland relatives. The gene families with the most variable copy numbers among inland plants were likelier to lose duplicate genes in Rhizophoreae mangroves (Fig. 5a). Large gene families were more likely to experience gene loss (supplementary figs. S23 and S24, Supplementary Material online). We conclude that the Rhizophoreae mangroves are likelier to lose nonessential genes but maintain essential processes.

Genes Involved in Environmental Interaction Are Disproportionately Lost

We then checked the molecular functions and biological pathways involved in the gene losses. We identified 2,670 gene families that contain more than twofold fewer genes in Rhizophoreae mangroves than the four inland relatives. Six biological processes are found among the reduced gene families more often than expected from the genome average: biotic stimulus response, hypoxia response, hormone response, secondary metabolic process, cell wall organization, and cell death (Fig. 5b; supplementary table S14, Supplementary Material online). If we focus on the 165 gene families with the most significant reduction in tandem duplication rates on the mangrove lineages, we still see enrichment in genes involved in biotic stress (supplementary table S15 and fig. S25, Supplementary Material online). Gene families lost in Rhizophoreae but retained in *C. pectinifolia* are disproportionally involved in hypoxia response (supplementary table S16 and fig. S26, Supplementary Material online).

The most significant reduction occurs in genes in response to biotic stimulus, which includes "response to

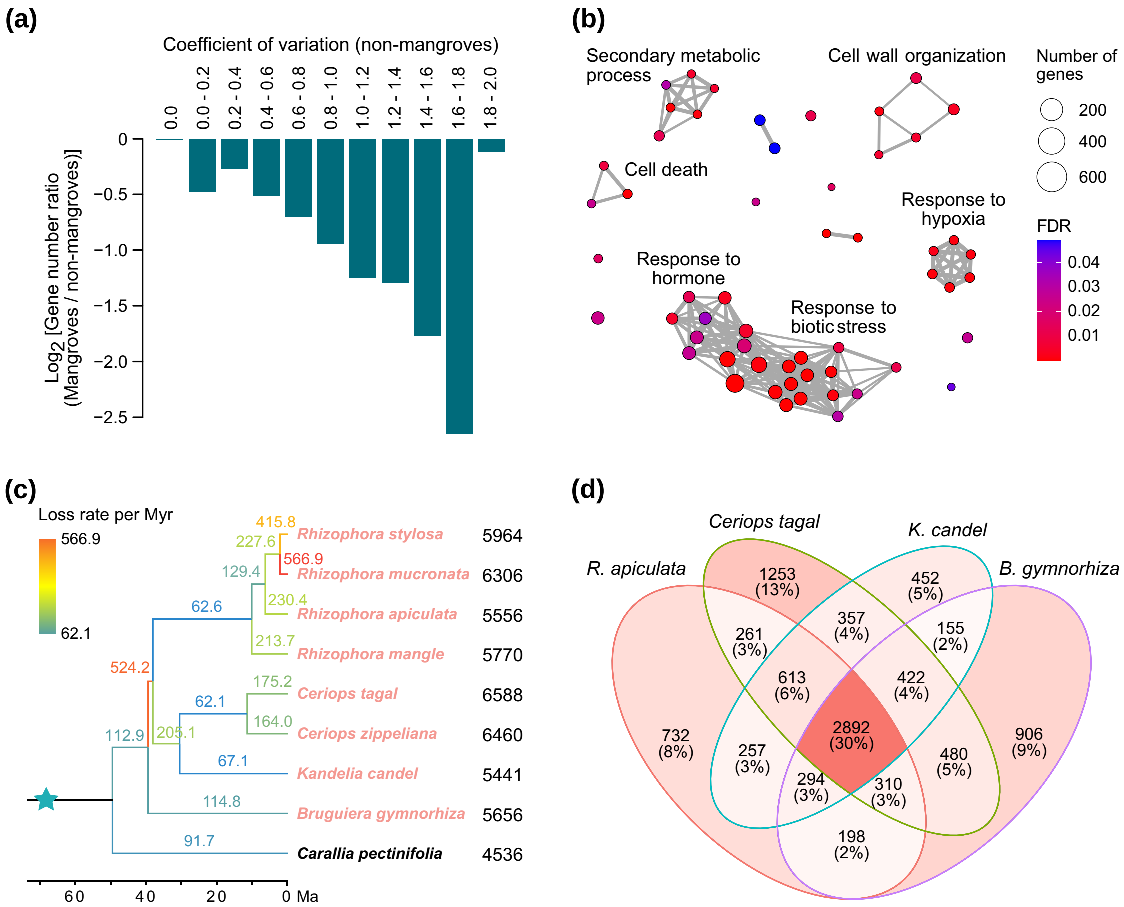


Fig. 5. Patterns of gene family reduction in Rhizophoreae mangroves. a) A decrease in gene numbers in Rhizophoreae mangroves positively correlates with gene number variability among inland plants. Copy number variability is measured using the CV (standard deviation/mean). b) GO term enrichments (in the Biological Process category) for gene families with fewer genes in Rhizophoreae mangroves. The GO terms with false discovery rates (FDRs) less than 0.05 are displayed. FDRs are calibrated using the Benjamini and Hochberg method. The Jaccard coefficient measures similarity between gene sets and is defined as the size of the intersection divided by the size of the union of gene sets. If the Jaccard similarity coefficient > 0.2 between two GO terms, an edge would be drawn to connect the GO terms. Thicker edges indicate stronger similarities. c) Comparison of gene loss between Rhizophoreae mangroves and *C. pectinifolia*. The numbers and colors of the branches represent the inferred loss rates per million years. The total number of gene loss events from the ancestor to present in a species is marked on the right side. The WGD events shared among Rhizophoreae are marked with stars on the branches. d) The Venn diagram represents the intersection of loss events for the four genera of Rhizophoreae mangroves. Colors from white to red indicate an increasing number of lost events.

bacterium," "response to oomycetes," and "defense response." In addition, disease resistance genes (*R* genes) containing nucleotide-binding site (NBS) domains are disproportionately lost in Rhizophoreae mangroves: 56 (30 to 74) remain compared to 317 (156 to 544) retained in non-mangroves (supplementary table S17, Supplementary Material online). Such loss of genes underlying the responses to biotic stress may be beneficial because there are relatively few pathogens in these stressful and fluctuating environments (Wang et al. 2012). Furthermore, abiotic stresses often induce genes in biotic stress response, resulting in unnecessary expression fluctuation (Huot et al. 2014). Indeed, disease gene loss has been observed in diverse plants inhabiting extreme environments (Zhang et al. 2019; Liu et al. 2021). Although mangroves suffer varying abiotic stresses in intertidal habitats, many genes involved in stress response have been lost. ABA and salicylic acid are involved in multiple stress processes, and genes responding to these hormones have also been lost (Huot et al. 2014). Furthermore, genes involved in programmed cell death, cell wall organization, and secondary metabolic processes are also easily influenced by abiotic stresses (Le Gall et al. 2015; Isah 2019). These results indicated that gene loss concentrates on processes closely related to environmental interaction.

Gene Losses Occurred Continuously during Mangrove Evolution

The Rhizophoreae mangroves originated ~50 Ma and subsequently diverged into 18 species. We explored whether the extensive gene losses are associated with the inhabitation of intertidal zones or continuously occur during mangrove evolution. Compared with inland relatives, gene loss in Rhizophoreae mangroves could be the repressed gene duplication or the accelerated gene nonfunctionalization. The gene duplication rate was inferred in the tandem gene duplications. Phylogenetic inference suggests that tandem duplications mainly occurred after species divergence (supplementary fig. S27, Supplementary Material online). Therefore, the small number of tandem duplications in Rhizophoreae mangroves resulted from decreased duplication rates in this clade. We infer gene loss events based on gene presence/absence information on the gene family trees (Materials and Methods). Among 17,352 gene families, we see an average of 5,968 loss events in Rhizophoreae mangrove plants, higher than the 4,536 in *C. pectinifolia* (Fig. 5c). The inferred gene loss rate of the common ancestor branch of Rhizophoreae mangrove is 112.9 per million years (Myr), only slightly higher than the 91.7/Myr of *C. pectinifolia*. At the stage of differentiation into four genera, the loss rate increased sharply, reaching a maximum of 524.2/Myr. The gene losses continuously occurred at high rates during the following species divergence. Correspondingly, we found the gene losses are different among Rhizophoreae mangroves. Only 30% of the total loss events are shared by species from four genera (Fig. 5d). Therefore, the large-scale gene loss of Rhizophoreae mangrove was not a single event that occurred in the common ancestor, but has continued during mangrove evolution.

Gene Losses and GRN Stability in the May–Wigner Framework

The results above suggest that the continuous gene losses are concentrated on environmental interactions but avoid essential genes. Since no previous theories could explain the broad gene losses across multiple taxa of mangroves, we postulate a mechanism associated with the specific mangrove habitats. In other words, constantly fluctuating intertidal environments may demand the stability of GRN. In the following sections, we will use the largest mangrove clade, Rhizophoreae, to suggest that GRN stability could be explained by the MWT.

The Basic Theory and the Observations

If a random network contains n nodes and with probability C the nodes are connected with an average interaction strength α, May (1972) suggested the network would be almost certainly unstable if $\alpha(nC)^{1/2} > 1$. That is, the central point of the MWT is that a large and complex system would be unstable. We applied the theory to GRNs based on the work of Chen et al. (2019). Briefly, in a GRN with n genes, the perturbations of gene i with mRNA concentration x_i could be denoted as $F_i(x_1, x_2, \dots x_n) = \frac{dx_i}{dt} = b_i + \sum_{j=1, j\neq i}^{n} a_{ij}x_j - D_i x_i$. The b_i is the basal transcription rate, D_i is the decay rate, and a_{ij} is the regulatory strength of gene j on gene i. Then the Jacobian matrix $\boldsymbol{A}$ describing interactions among genes could be calculated as $\frac{\partial F_n}{\partial x_n}(X)$. As a result, the matrix elements are $\boldsymbol{A}_{ii} = -D_i$ and $\boldsymbol{A}_{ij} = a_{ij}$, where $\boldsymbol{A}_{ii}$ and $\boldsymbol{A}_{ij}$ represent the diagonal and off-diagonal elements, respectively (Materials and Methods). Then the GRN stability could be evaluated by the leading eigenvalue (the eigenvalue with the largest real part) of the matrix $\boldsymbol{A}$. The leading eigenvalue with a negative real part indicates the stability of the network (May 1972).

We test the prediction that small networks are more stable using simulation. The topology of GRNs was represented by scale-free networks, whereby a small number of nodes have a high level of connectivity with the rest having very few connections (Guelzim et al. 2002). The connectivity, mean, and variation of interaction strength of the real GRN were set according to the estimation by Chen et al. (2019). The diagonal element $\boldsymbol{A}_{ii}$ was set as -1, according to May (1972). In the simulation, we changed the gene number and complexity of the network. We simulated 1,000 times for each condition to calculate the probability of being stable. Consisting with our hypothesis, we found a trend that the GRN stability decreases with gene number or network complexity increase (Fig. 6a). We also set a nonlinear model and calculated gene mRNA concentration fluctuations using iterative simulation (Materials and Methods). We again found that the GRN stability decreases with the increase in gene numbers (supplementary figs. S28 and S29, Supplementary Material online).

Indeed, the observed patterns of gene loss in Rhizophoreae mangroves are consistent with the model predictions. The concentration of environmental interactions of gene losses may moderate the influence of environmental fluctuations on GRN stability. The gene losses avoid essential genes while simplifying GRNs. When mapping the mangrove gene losses to *Arabidopsis thaliana*, we found that 2,624 of the 10,480 loss events could be assigned to the reliably connected gene network that contains 6,411 of the 27,416 total genes (Burks et al. 2022; Fisher's exact test; $P = 2.1 \times 10^{-7}$). The gene losses are also enriched on large gene families. More than half of the multicopy gene families (copy number ≥ 3) experienced gene loss (926 of 1,803). Constrained by the criteria that all the family's genes are reliably linked, the concentration increased to 52 of 71 (Fisher's exact test; $P = 1.0 \times 10^{-4}$). Therefore, it may minimize the effects on growth and development and cause subnetwork simplification instead of a complete loss. More importantly, gene losses continuously occurred from the inhabitation of intertidal zones to recent species divergence. It suggested that gene loss may be mainly driven by fluctuating environments instead of associated with phenotypic traits.

Reduced GRN Sizes and the Enhanced Stability under Salinity Changes

We then directly test whether the GRN of mangroves is more stable after gene loss. Pervasive transposable element (TE) loss was also observed (Lyu et al. 2018). TE activation by short-term environmental insults induces stress-responsive transcripts and instability in GRN (Pfeiffer et al. 2000; Schuermann et al. 2005; Argueso et al. 2008). Therefore, TE and gene loss may have promoted GRN stability under environmental change.

To test this hypothesis, we examined TE and gene expression under various salinity conditions. Rhizophoreae

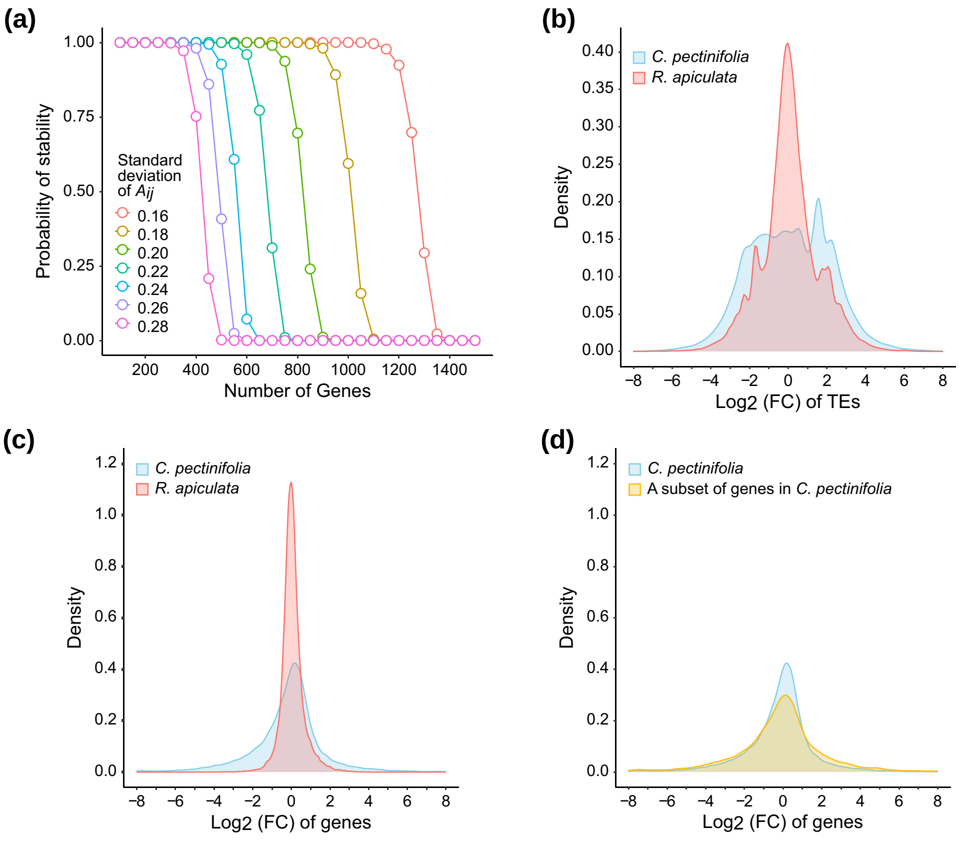


Fig. 6. The role of gene family reduction in increasing GRN stability. a) The GRN stability is affected by gene number and interaction strength. The probability of GRN stability was calculated in 1,000 simulations. The lines represent GRN with different standard deviations of $\boldsymbol{A}_{ij}$, the off-diagonal element of the matrix. b to d) Gene and TE loss and GRN stability across saline environments. Expression of both TEs and genes is more stable in *R. apiculata* than in *C. pectinifolia* across salinity levels. The subset of genes in *C. pectinifolia* in d) represents the gene families with fewer members in Rhizophoreae mangroves, and salinity levels influence them more. The significance level of the differences in expression fold changes between mangrove and nonmangrove was tested using the Kolmogorov–Smirnov test ($P < 10^{-15}$).

mangroves have smaller genomes that contain fewer TEs (supplementary tables S18 and S19, Supplementary Material online) than most plant taxa. The numbers of intact long terminal repeat retrotransposons (LTR-RTs) inserted into Rhizophoreae genomes range from 96 to 665, significantly fewer than 1,025 in *C. pectinifolia* (supplementary fig. S30, Supplementary Material online; Wilcoxon signed-rank test; $P = 0.004$). We treated *C. pectinifolia* and *Rhizophora apiculata* with 0 and 250 mM NaCl, examined gene expression alteration between the two salt concentrations (supplementary table S20, Supplementary Material online; Materials and Methods). Relatively stable GRNs in *R. apiculata* are expected to change the expression of fewer genes than the inland species. TE expression in *C. pectinifolia* significantly changed after 250 mM NaCl treatment (as seen in other plants), with 69.05% of TE copies' expression changing more than twofold (Fig. 6b; supplementary fig. S31, Supplementary Material online). In contrast, the expression of only 41.39% of *R. apiculata* TEs changed more than twofold (Kolmogorov–Smirnov test; $P < 10^{-15}$). Differences in expression plasticity of protein-coding genes are even more pronounced: 38.47% *C. pectinifolia* transcripts changed at least twofold, in sharp contrast to only 7.79% in *R. apiculata* (Fig. 6c; supplementary fig. S32, Supplementary Material online; Kolmogorov–Smirnov test; $P < 10^{-15}$). We further examined the expression levels of 2,670 gene families with an average gene number of mangroves less than half that of four inland relatives. In these most significant gene reductions in Rhizophoreae mangroves, their homologs in *C. pectinifolia* are the most labile and show a significantly higher expression alteration than other genes (Fig. 6d; Kolmogorov–Smirnov test; $P < 10^{-15}$). These results indicate that Rhizophoreae mangroves maintain GRN stability under salinity change, probably due to the simplification of GRNs.

Discussion

Whether the frequently occurring large-scale gene loss could be a significant player in adaptative evolution is unknown. Although the prevalence of gene loss events has been uncovered during evolution, only a few specific cases have shown that gene losses are essential sources of adaptive genetic variation (Albalat and Cañestro 2016; Monroe et al. 2018; Sharma et al. 2018; Huelsmann et al. 2019; Helsen et al. 2020). This study identified a shared trend of extensive gene loss in independent mangrove clades that inhabited unstable intertidal environments. We then deeply investigate the gene loss events in the largest mangrove clade Rhizophoreae. A small subset of gene losses has dominated the innovation of adaptive traits, while most of the gene loss may have contributed to the GRN stability in highly fluctuating environments.

Small gene losses have participated in the evolution of three examined phenotypic traits. The loss of *DOG1* essential in seed dormancy and a set of genes related to seed desiccation tolerance and reserve accumulation may contribute to the formation of viviparous seeds (Feng et al. 2021). Similar gene losses were identified in high salinity response and tannin biosynthesis pathways, in which genes negatively regulate the processes that are extensively lost. It is consistent with the "less is more" hypothesis, which suggests gene loss altered dosage balance of macromolecular complexes or biological pathways could ultimately affect the phenotype and evolutionary fitness (Edger and Pires 2009; Birchler and Veitia 2012).

However, gene losses related to phenotype evolution account for only a small percentage. Instead, the gene loss is concentrated in large and highly variable gene families, which are mainly involved in stress response and metabolic processes easily induced by environmental fluctuations. We thus speculated that the extensive gene losses in mangroves may be driven by the unique habitats of extremely unstable intertidal environments. Multiple explanations exist for the associations between gene set reduction and environmental instability. However, few of these explanations seem to have the power to account for the broad and consistent reduction in mangrove genomes. For example, intertidal environments seem to have fewer pathogens than inland habitats, which may allow mangroves to shed some genes from their genomes. The extensive gene losses may also decrease energy consumption, benefiting mangroves' growth in nutrient-deficient intertidal zones (Wagner 2005; Strickler et al. 2006; Reef et al. 2010). However, the gene number of carnivorous plants inhabiting nutrient-deficient environments is larger than that of mangroves, so the benefits should be weak (Fukushima et al. 2017).

The MWT is a hypothesis that is more generally applicable and testable, which suggests large and highly connected GRNs are unstable (May 1972; Allesina and Tang 2012; Chen et al. 2019). If any of other explanations is true, MWT may still contribute to the GRN stability for small genomes that evolved for other reasons. When invading the most variable intertidal environments from inland ancestors, the instability should be more pronounced in mangroves (Lyu et al. 2018; Xu et al. 2023). Consistent with the theory, we found significantly fewer genes in Rhizophoreae mangroves due to gene duplication repression and an accelerated gene loss rate. Both simulation and experimental data suggest that gene loss has promoted the GRN stability of mangroves.

The hypothesis relies on the fact that the highly fluctuating intertidal environments have greatly perturbed mangroves' GRNs. The intertidal zone is the area above water level at low tide and underwater at high tide, the interface between terrestrial and marine habitats. Being sandwiched between two suitable environments, it is more susceptible to extreme conditions than a single environment due to its unique composite nature (supplementary text and fig. S33, Supplementary Material online). The most striking feature of intertidal zones distinguishing from other extreme environments is the high fluctuations (Tomlinson 2016; Rivera-Monroy et al. 2017; Duke et al. 2022). Some other extreme environments also stress plants, such as hot springs, deep seas, or polar environments. And the environmental factors are usually much more stable compared with intertidal zones.

The pattern is obvious in salinity and waterlogging, the most primary stresses in intertidal zones. Many aquatic plants, such as seagrasses, are completely submerged in water throughout their life cycle. The submerge degree would not change greatly. While halophytes usually adapted to a range of salinity, the ranges are narrow and their changes require long time spans (Gil et al. 2014). In contrast, the salinity and waterlogging of mangrove habitats vary greatly in time and space. The average daily tidal range of the intertidal zone is 0.5 to 2 m; in some areas, it is as high as 4 m (Wells 1983; Harrison et al. 1997). Correspondingly, the environmental salinity could change from $< 10‰$ to $> 30‰$ in a single day and fluctuate from freshwater to higher than seawater biweekly or monthly, superimposed on the influence of climatic factors (Lin et al. 1993; Studer and Poulin 2013; Yang et al. 2013; Nguyen et al. 2015). It means that a mangrove tree would be shocked by salinity and waterlogging frequently. In space, different individuals from one population and different populations from one species experience varying salinity and waterlogging under the influence of tidal inundation, rainfall, soil type, topography, and many other factors (Ball 1998; Duke et al. 1998; Nguyen et al. 2015; Tomlinson 2016). Individuals of *Kandelia obovata* in one river system experience annual mean salinity ranging from 10‰ to 30‰ (Yang et al. 2013). Two populations of *R. apiculata* in Thailand grow in conditions with 1‰ and 35‰ salinity, respectively (Rivera-Monroy et al. 2017). Furthermore, the complex interaction of multiple environmental factors makes environments more unpredictable. Although there are reports that mangrove habitats have fewer pathogens, it also reflects the extreme environments of intertidal zones for all organisms (Cheng et al. 2009; Wang et al. 2012).

With such highly fluctuating environments, many studies have found the altered gene expression and thus perturbed GRN in mangroves (Liang et al. 2012; Cruz et al. 2020; Feng et al. 2020; Su et al. 2022; Xu et al. 2023). The environmental fluctuations, such as salinity changes, have greatly affected mangroves' physiology and morphology (Reef and Lovelock 2015; Lechthaler et al. 2016; Méndez-Alonzo et al. 2016; Garcia et al. 2017). Correspondingly, the environmental fluctuations have altered many genes' expression (Wang et al. 2016; Xiao et al. 2016). This study found that GRNs of inland plants are highly perturbed under these environmental fluctuations, but mangrove GRNs have high stability. We then found that the GRN stability and the extensive gene loss events uniquely found in mangroves could be linked by the MWT. Since MWT has been proven responsible for the stability of other kinds of networks, we suggested that the theory may better explain the general trend of gene family reduction among mangroves than many other explanations. The gene number reduction may be the simplest way to maintain GRN stability.

Stability is a fundamental property of all networks. In the case of GRNs, which are among the most important networks governing complex biological processes, their stability ensures the growth and reproduction of organisms. Indeed, GRN stability has been investigated recently. miRNA-mediated mRNA degradation and network topology were suggested to influence GRN stability (Guo and Amir 2021). However, size reduction may be the most effective way to simplify and stabilize a network. In true GRNs, biological function constrained the size reduction process, and essential genes could not be lost. For example, although the loss of transcription factors (TFs) that are usually highly connected contributes more to GRN stability, it may cause functional deficiency. Therefore, only 10% of TFs are lost compared with inland relatives, while the whole gene set has lost 25%. However, it does not mean the large gene family reduction observed in mangroves simply results are useless. Many genes that could be mapped in the *A. thaliana* key subnetwork were unexpectedly lost. The biased loss in large gene families may maximize the effect on GRN stability while minimizing the influence on biological functions.

Indeed, the MWT could explain more general gene loss than observed in mangroves. A similar pattern in plant genomes is the extensive gene loss after WGDs. The losses did not affect essential processes but limited the total amount of gene numbers after multiple rounds of gene number doubling. Therefore, the gene losses in mangroves and after WGDs may be the constraint of GRN stability explained by the MWT. The gene number reductions have often been observed in unicellular organisms, probably because of their cells' direct contact with environments. The streamlined and compact genomes of many marine bacteria, such as *Prochlorococcus* and Pelagibacterales, and the eukaryotic *Fabrea salina* were proposed to promote the adaptation to a wider range of stresses (Giovannoni et al. 2014; Zhang et al. 2022). Furthermore, genomes streamline as the optimal temperature increase in bacteria and archaea (Sabath et al. 2013).

We note, however, that genome shrinkage may also reduce the adaptive potential of mangroves to the forthcoming climate change and human activities (Guo et al. 2018). The tremendous loss of disease and drought tolerance genes has dramatically decreased the potential of mangrove readaptation to terrestrial environments. In a long period of evolution, monotone gene loss and GRN simplification should not be the best solution for adaptation to unstable environments. Indeed, a biphasic model of punctuated genome evolution has been proposed. The short, explosive phase with extensive gene duplications such as WGD leads to a rapid increase in genome complexity. It is followed by a much longer phase of genome complexity decrease (Wolf and Koonin 2013). Mangroves inhabiting the most fluctuating environments may have accelerated genome reduction rates. It also indicated that habitat protection is essential for mangrove conservation.

Materials and Methods

Plant Material and Sequencing

Mature individual tissues of the six Rhizophoraceae species were collected from southern China, Singapore, and Malaysia (supplementary table S1, Supplementary Material online). Genomic DNA was extracted from leaves using a modified CTAB method (Doyle and Doyle 1987). Total RNA was extracted from leaves and roots using the TRIzol reagent (Invitrogen). We have selected the most suitable strategy for genome sequencing, assembly, and annotation for each species (supplementary table S1, Supplementary Material online). SMRT (Pacific Biosciences, Menlo Park, CA, United States) long reads were sequenced using PacBio Sequel II platforms. 10x Genomics linked reads were sequenced using the BGISEQ-500 platform. DNA short-insert paired-end reads were sequenced using Illumina NovaSeq 6000, Illumina X-TEN, and BGISEQ-500 platforms. RNA-seq was performed using the same platform as DNA short reads.

De Novo Genome Assembly

PacBio long reads were used to de novo assemble the genome using two programs: Falcon (Chin et al. 2016) and MECAT2 (Xiao et al. 2017). The assemblies were polished using Quiver (Chin et al. 2013) or Racon (v.1.3.1; Vaser et al. 2017). DNA short reads were mapped to the polished genomes to correct sequencing errors using Pilon (v1.22; Walker et al. 2014). Assembly redundancy was removed using Purge Haplotigs (Roach et al. 2018).

The *Ceriops tagal* genome was assembled using a combination of 10x Genomics sequencing data and Hi-C data. First, the 10X Genomics linked reads were used to assemble the genome with Supernova (v2.0.1; Weisenfeld et al. 2017). Next, DNA short reads were used for gap filling using GapCloser (v1.12; Luo et al. 2012). Assembly redundancy was then removed using Redundans (v0.14a; Pryszcz and Gabaldón 2016). Finally, the assemblies were mapped to chromosomes using HiC-Pro (v3.2; Servant et al. 2015), Juicer (Durand et al. 2016), and 3D de novo assembly (3D-DNA) pipeline (Dudchenko et al. 2017). The same process was used to the chromosome assembly of *C. pectinifolia* and *R. apiculata*. The accuracy and completeness of the assembled genomes were evaluated using the 2,326 eudicot universal single-copy orthologs from BUSCO (v5.0.0; Seppey et al. 2019; Manni et al. 2021).

Repetitive Sequence and Protein-Coding Gene Annotation

Homology and de novo prediction approaches were used for repetitive sequence annotation. Homology prediction was performed by aligning genome sequences to the Repbase library (http://www.girinst.org/repbase; Bao et al. 2015) using RepeatMasker and RepeatProteinMask (Tarailo-Graovac and Chen 2009). De novo prediction was conducted based on genome sequences and recognition of repetitive sequence characteristics using RepeatModeler (Flynn et al. 2020), TRF (Benson 1999), and LTR_FINDER (Xu and Wang 2007). LTR-RTs were identified using LTRharvest (v1.6.1; Ellinghaus et al. 2008) and LTR_FINDER_parallel (v1.1; Ou and Jiang 2019). Intact LTR-RTs were then identified using LTR_retriever (v2.8.7; Ou and Jiang 2018).

Based on repeat-masked genomes, we predicted protein-coding genes using a combination of three methods: homologous sequence alignment, de novo prediction, and gene expression. First, high-quality published genomes were used for homology prediction with GeneWise (Birney et al. 2004). Then, by identifying conservative transcription sites, Augustus (v3.3.1; Stanke et al. 2006), GlimmerHMM (Majoros et al. 2004), and GeneMark-ET (Lomsadze et al. 2014) were used for de novo predictions. TopHat2 (v2.1.1; Kim et al. 2013), Cufflinks (v2.2.1; Trapnell et al. 2012), HISAT2 (Kim et al. 2019), and StringTie (v1.3.5; Kovaka et al. 2019) were used based on transcript prediction. Finally, the genes identified by the three methods were integrated into the final nonredundant (NR) gene set by MAKER (v3.00; Cantarel et al. 2008) and EVidenceModeler (Haas et al. 2008).

Gene functions were annotated by searching for the best match in a series of function databases, including NCBI NR, Swiss-Prot, TrEMBL, and Kyoto Encyclopedia of Genes and Genomes (KEGG) database, using BLAST (v2.2.6; -e 1e-5). In addition, protein domains and Gene Ontology (GO) classifications were annotated using InterProScan (v5.30-69.0; Jones et al. 2014).

Gene Type Analysis and Genome Feature Visualization

To identify gene duplications, all proteins within a genome were aligned by BLASTp (v2.11.0; $e < 10^{-10}$; identity ≥ 40%). Gene types and collinear blocks were then identified by *MCScanX* (Wang et al. 2012). Collinear blocks containing at least five gene pairs were retained for subsequent analyses. Genomic features (collinear blocks, GC content, gene density, TE content, and scaffold length) were visualized using Circos (v0.69-9; Krzywinski et al. 2009). GC content, gene density, and TE content were calculated using a sliding window method with 200 kb.

Gene Family Clustering and Divergence Time Estimation

Gene family analysis was performed in 14 core eudicot plants, with *Vitis vinifera* as an outgroup. First, protein sequences of the 14 species were aligned with DIAMOND (v0.9.24.125; -e 1e-10; identity ≥ 40%; Buchfink et al. 2014). Gene families were clustered using OrthoFinder (v2.3.11; -I 1.5; Emms and Kelly 2019). Next, amino acid sequences of single-copy orthologous were aligned with MAFFT (v7.429; --maxiterate 1000 -localpair; Katoh and Standley 2013) and converted into nucleotide sequences using PAL2NAL (v14; Suyama et al. 2006). After removing poorly aligned regions with Gblocks (v0.91; -b5 = n; Castresana 2000), genes from each species were synthesized into supergenes, which were then used to draw the species tree. Finally, the species tree was constructed using RAxML-NG (v1.0.2; Kozlov et al. 2019) and replicated 1,000 times using the GTR + G model.

To ensure the accuracy of divergence time estimation, we used *Oryza sativa* (Ouyang et al. 2007) as an outgroup of the 14 core eudicot plants. Reliable single-copy orthologous genes were found and aligned as before. Divergence time was inferred using the MCMCTree program from PAML package (v4.9j; Yang 2007) with approximate likelihood calculation and repeated 10,000,000 times. The program was run twice to ensure time convergence. Several reliable fossils were selected to calibrate: (i) The root nodes of monocots and eudicots were limited to less than 247 Ma; (ii) the root node of core eudicots was limited to less than 125 Ma; (iii) recent ancestors of Rhizophoreae have appeared in early Eocene (56 to 47.8 Ma); and (iv) recent ancestor of *Rhizophora* has appeared in late Eocene (37.8 to 33.9 Ma; Muller 1981; Graham 2006; Morris et al. 2018).

Candidate Genes Involved in Seed Dormancy, ABA Response, and Salt Tolerance

We first searched the TAIR database (https://www.arabidopsis.org/; Berardini et al. 2015) for genes involved in seed germination and dormancy, salt tolerance, ABA biosynthesis, and signaling. Gene families containing these genes from *A. thaliana* were collected as candidates. Gene families with more than twice the number of genes in mangroves compared to nonmangroves were marked as expanded and those with half the number in mangroves as reduced. The copy number expansions were further validated by checking protein sequence alignments and phylogenetic tree reconstructions. Searching homologous sequences in genome assemblies validated gene loss candidates. If homologous sequences were found, gene structure completeness was examined by searching for premature termination and frameshift mutations.

High Tannin Content

Materials

Three Rhizophoreae mangrove species (*Bruguiera gymnorhiza*, *Ceriops tagal*, and *R. apiculata*) were sampled from Dongzhai Harbor, Hainan, and *C. pectinifolia* was sampled from the South China Botanical Garden. Fresh leaves and stems were sampled and stored at −80 °C for tannin content determination and RNA sequencing. For each species, three individuals were sampled as biological replicates.

Tannin Content Determination

Tannin from leaves and stems was extracted by water and treated with phosphomolybdic acid in an alkaline environment to form a blue compound (Kloster 1974). We used a microplate reader to measure the content at 760 nm and adjust it to zero with distilled water. The regression equation determined under standard conditions is:

$$y = ax + b, \tag{1}$$

where x is the standard concentration (mg/mL) and y is ΔA (A determination − A control).

$$\text{Tannin content (mg/g)} = \frac{(\Delta A - b) / a \times V1}{W \times V1 / V2} \times D = \frac{(\Delta A - b) \times V2 \times D}{a \times W} \tag{2}$$

(V1, sample volume, mL; V2, extract volume, mL; D, sample dilution factor, mg/mL; W, sample fresh weight, g).

Gene Expression Measurements

Total RNA from leaves and stems of the four species was extracted using Plant RNA Kit (Omega Bio-tek). The Illumina platform was used for sequencing. Raw sequence reads were first processed to remove adapters, and then fastp (v0.20.0; -q 20 -u 10 -n 5; Chen et al. 2018) was employed to remove low-quality reads. Clean sequences from each sample were aligned to the relevant genomes using HISAT2 (v2.2.1). SAM files were sorted and converted to BAM files using SAMtools (v1.9; Li et al. 2009). Transcriptome assembly was conducted using the simplified StringTie workflow (v2.1.4; -e), which directly estimates and analyzes the expression of a given transcript in the reference genome annotation file. Transcript expression is estimated using transcripts per kilobase of exon model per million mapped reads (TPM). For cross-species comparison, gene expression levels are normalized according to the method of Brawand et al. (2011). For genes whose expression levels are within the interquartile range, we identified 1,000 genes with the most conservative ranking among all samples. Next, the median of the expression levels of these genes in each sample was estimated, and these medians were adjusted to a common value through a scaling factor. Finally, these factors were used to correct the expression levels of all genes in the sample. We further used the sum of the expression levels of all isoenzymes as the expression level of the enzyme and used the R package "pheatmap" to draw the expression heatmap.

Gene Family Dynamics and Tandem Duplication Analysis

Gene families in mangrove species with no more than half the genes in nonmangroves are defined as significantly reduced in mangroves. GO annotations of these genes were mapped using *A. thaliana* genome information for corresponding gene families. GO enrichment analysis was performed using the R package "clusterProfiler" (Yu et al. 2012).

Homologous genes separated by no more than five genes were defined as tandem duplications. The synonymous substitution rate (Ks) between tandem duplicates was calculated using KaKs_Calculator (v2.0; -m YN; Wang et al. 2010). Amino acid usage bias in these duplicates was calculated, and the results were displayed using the R package "pheatmap" (https://CRAN.R-project.org/package=pheatmap). Fisher's exact test was used to test whether a gene family contains significantly more tandem duplications in eight mangroves than the four Malpighiales nonmangroves, or vice versa ($P < 0.05$). These gene families were used for GO enrichment analysis.

Gene Tree Construction, Gene Tree Species Tree Reconciliation, and Gene Loss Identification

We reconstructed gene trees of homologous groups produced by OrthoFinder to identify gene duplication and loss events. First, gene families with fewer than four loci or containing genes from only one species were removed. Then, for each of the remaining gene families, sequence alignment and conversion of protein sequences to CDS were performed using the method outlined above. Next, poorly aligned regions were removed using Gblocks (-b5 = h). Finally, gene family maximum likelihood (ML) trees were reconstructed using RAxML-NG under the GTR + G model with 100 bootstraps.

Notung (v2.9.1.5; Stolzer et al. 2012) was then used to reconcile gene trees with the species tree. The software outputs reconstructed gene trees with the most negligible duplication/loss cost based on the principle of maximum parsimony. Briefly, the scenario with the fewest duplication/loss events was chosen as the inferred gene evolution model. For example, the simultaneous gene absence in sister species A, B, and C would be inferred as the single gene loss event in their ancestor. The analysis was based on the duplication/loss model. First, the "rooting" mode was used to establish roots for rootless gene trees. Then "reconcile" mode was used to reconcile each family's gene and species tree. Finally, the "rearrange" mode (–threshold 50%) was used to rearrange weakly supported edges (bootstrap < 50%). We carefully checked the rearranged trees and further restricted the criteria for a duplication event "at least one species exists in both subbranches," thus removing some untrustworthy duplication events.

Gene loss events were identified in the eight mangrove species and *C. pectinifolia*. In each phylogenetic tree of a gene family, we first identified subtrees that contain only nine species. In each subtree, we checked whether Notung had determined loss events. If we further found no genes in its children nodes, the node was considered as having experienced a gene loss event. Gene families have more than twice the loss events in mangroves than *C. pectinifolia* were retained for GO enrichment analyses.

Protein Domain Identification

Hidden Markov model (HMM) profiles in the Pfam database (34.0; Mistry et al. 2021) were prepared as an HMM database by the hmmpress program in the HMMER suite (3.2.1; Eddy 2011). The HMM profiles were used to identify NBS (PF00931) and LEA (dehydrin [PF00257], LEA_1 [PF03760], LEA_2 [PF03168], LEA_3 [PF03242], LEA_4

[PF02987], LEA_5 [PF00477], LEA_6 [PF10714], and SMP [PF04927]) domains in the genomes of 14 species using PfamScan (1.6; Mistry et al. 2007).

Simulation of the Effect of Gene Loss on GRN Stability

For a GRN containing n genes, the change in the mRNA concentration x_i of gene i with time t is dx_i/dt. Perturbations near equilibrium can be approximated by ordinary differential equations, although the system should be nonlinear:

$$F_i(x_1, x_2, \ldots x_n) = \frac{dx_i}{dt} = b_i + \sum_{j=1,\, j\neq i}^{n} a_{ij}x_j - D_i x_i, \qquad (3)$$

where b_i is the basal transcription rate, a_{ij} is the regulatory intensity of gene j on gene i, and D_i is the degradation rate. The interaction between genes can construct a Jacobian matrix $\boldsymbol{A}$ and take the partial derivative of (3) and bring it into $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(X) & \frac{\partial F_1}{\partial x_2}(X) & \cdots & \frac{\partial F_1}{\partial x_n}(X) \\ \frac{\partial F_2}{\partial x_1}(X) & \frac{\partial F_1}{\partial x_2}(X) & & \frac{\partial F_2}{\partial x_n}(X) \\ \vdots & & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1}(X) & \frac{\partial F_n}{\partial x_2}(X) & \cdots & \frac{\partial F_n}{\partial x_n}(X) \end{pmatrix}, \qquad (4)$$

where $\boldsymbol{A}_{ii} = -D_i$ and $\boldsymbol{A}_{ij} = a_{ij}$. The diagonal element $\boldsymbol{A}_{ii}$ represents the effect of gene i on itself or the gene decay rate, and the off-diagonal element $\boldsymbol{A}_{ij}$ represents the regulation strength of gene j on gene i. When the real part of the maximum eigenvalue of matrix $\boldsymbol{A} < 0$, the GRN reaches a stable state. The GRN with power-law distribution is generated by the "igraph" package (no. of edges $= 0.031 \times n^2$), $\boldsymbol{A}_{ii}$ is set to -1, and $\boldsymbol{A}_{ij}$ is randomly drawn from a normal distribution with mean 0 and standard deviation from 0.16 to 0.28 (Chen et al. 2019). Each condition was repeated 1,000 to statistically stabilize the probability.

A sigmoidal transfer function is added to Equation (3), making the new equation have the necessary nonlinearity:

$$\frac{dx_i}{dt} = S\left(b_i + \sum_{j=1,\, j\neq i}^{n} a_{ij}x_j\right) - D_i x_i \qquad (5)$$

where $S(x) = \frac{1}{1+e^{-x}}$. The initial mRNA concentration of gene x_i is set to $1/n$, b_i is set to 1, and other parameters are consistent with linear simulation (D'haeseleer 2000). When the change amplitude of the mRNA concentration of all genes $< 1 \times 10^{-6}$, the GRN is stable.

Salt Treatment Experiment

Materials

We measured transcriptome profiles of mangrove species *R. apiculata* and its inland relative *C. pectinifolia* under salinity changes. Seedlings of each species were divided into two groups: one was grown in pure culture solution (0 mM NaCl) and the other with 250 mM NaCl added. After treatment for 24 h, fresh leaves and roots were sampled and stored at −80 °C for RNA sequencing. For each condition, three or two individuals were sampled as biological replicates.

Measurement of Fold Change in Gene and TE Expression

Total RNA extraction, sequencing, and transcriptome assembly methods were the same as in the High Tannin Content section. The RNA-seq was aligned to the corresponding reference genomes to measure TE expression, with the HISAT2 parameter set to --no-mixed. TE expression was measured using TEtranscripts (v2.2.1; Jin et al. 2015). Fold change of gene and TE expression were estimated using the R package "DESeq2" (Love et al. 2014).

Supplementary Material

Supplementary material is available at *Molecular Biology and Evolution* online.

Acknowledgments

We thank Yuxin Chen from Stanford University and Yongsen Ruan from Sun Yat-sen University for the helpful discussion and Wei Lun Ng from Xiamen University Malaysia for assistance with field sampling.

Author Contributions

S.Shi. and Z.H. conceived and designed the research. S.X., S.Sha., X.F., S.L., W.W., M.L., M.E.T., C.Z., and Z.G. collected materials. S.X., S.Sha., and X.F. performed the experiments. S.X., S.Sha., L.Z., and Z.H. carried out data processing and analysis. S.X., S.Sha., Z.H., C.-I.W., and S.Shi. wrote the first draft of the manuscript. All authors read and approved the final version.

Funding

This study was supported by the National Natural Science Foundation of China (32170230, 32100176, and 32330005), the Guangdong Basic and Applied Basic Research Foundation (2022A1515010962 and 2023B1515020083), and the Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai; 311021006).

Conflict of Interest

None declared.

Data Availability

The genome assembly sequences and raw Illumina reads have been deposited in the National Center for

Biotechnology Information (NCBI) with BioProject number PRJNA996495. The genome annotations are available at Figshare (https://doi.org/10.6084/m9.figshare.25323121).

References

Albalat R, Cañestro C. Evolution by gene loss. *Nat Rev Genet.* 2016:**17**(7):379–391. https://doi.org/10.1038/nrg.2016.39.

Allesina S, Tang S. Stability criteria for complex ecosystems. *Nature* 2012:**483**(7388):205–208. https://doi.org/10.1038/nature10832.

Argueso JL, Westmoreland J, Mieczkowski PA, Gawel M, Petes TD, Resnick MA. Double-strand breaks associated with repetitive DNA can reshape the genome. *Proc Natl Acad Sci U S A.* 2008:**105**(33):11845–11850. https://doi.org/10.1073/pnas.0804529105.

Ball MC. Mangrove species richness in relation to salinity and waterlogging: a case study along the Adelaide River floodplain, northern Australia. *Glob Ecol Biogeogr Lett.* 1998:**7**(1):73–82. https://doi.org/10.2307/2997699.

Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA.* 2015:**6**(1):1–6. https://doi.org/10.1186/s13100-015-0041-9.

Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 1999:**27**(2):573–580. https://doi.org/10.1093/nar/27.2.573.

Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E. The Arabidopsis information resource: making and mining the "gold standard" annotated reference plant genome. *Genesis* 2015:**53**(8):474–485. https://doi.org/10.1002/dvg.22877.

Birchler JA, Veitia RA. Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. *Proc Natl Acad Sci U S A.* 2012:**109**(37):14746–14753. https://doi.org/10.1073/pnas.1207726109.

Birney E, Clamp M, Durbin R. GeneWise and genomewise. *Genome Res.* 2004:**14**(5):988–995. https://doi.org/10.1101/gr.1865504.

Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M, *et al.* The evolution of gene expression levels in mammalian organs. *Nature* 2011:**478**(7369):343–348. https://doi.org/10.1038/nature10532.

Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods.* 2014:**12**(1):59–60. https://doi.org/10.1038/nmeth.3176.

Burks DJ, Sengupta S, De R, Mittler R, Azad RK. The *Arabidopsis* gene co-expression network. *Plant Direct* 2022:**6**(4):1–22. https://doi.org/10.1002/pld3.396.

Cai L, Arnold BJ, Xi Z, Khost DE, Patel N, Hartmann CB, Manickam S, Sasirat S, Nikolov LA, Mathews S, *et al.* Deeply altered genome architecture in the endoparasitic flowering plant *Sapria himalayana* Griff. (Rafflesiaceae). *Curr Biol.* 2021:**31**(5):1002–1011.e9. https://doi.org/10.1016/j.cub.2020.12.045.

Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Alvarado AS, Yandell M. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 2008:**18**(1):188–196. https://doi.org/10.1101/gr.6743907.

Carles C, Bies-Etheve N, Aspart L, Léon-Kloosterziel KM, Koornneef M, Echeverria M, Delseny M. Regulation of *Arabidopsis thaliana Em* genes: role of ABI5. *Plant J.* 2002:**30**(3):373–383. https://doi.org/10.1046/j.1365-313X.2002.01295.x.

Carretero-Paulet L, Librado P, Chang TH, Ibarra-Laclette E, Herrera-Estrella L, Rozas J, Albert VA. High gene family turnover rates and gene space adaptation in the compact genome of the carnivorous plant *Utricularia gibba*. *Mol Biol Evol.* 2015:**32**(5):1284–1295. https://doi.org/10.1093/molbev/msv020.

Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 2000:**17**(4):540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334.

Chen K, Li GJ, Bressan RA, Song CP, Zhu JK, Zhao Y. Abscisic acid dynamics, signaling, and functions in plants. *J Integr Plant Biol.* 2020:**62**(1):25–54. https://doi.org/10.1111/jipb.12899.

Chen Y, Shen Y, Lin P, Tong D, Zhao Y, Allesina S, Shen X, Wu CI. Gene regulatory network stabilized by pervasive weak repressions: microRNA functions revealed by the May-Wigner theory. *Natl Sci Rev.* 2019:**6**(6):1176–1188. https://doi.org/10.1093/nsr/nwz076.

Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018:**34**(17):i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Cheng ZS, Pan JH, Tang WC, Chen QJ, Lin YC. Biodiversity and biotechnological potential of mangrove-associated fungi. *J For Res.* 2009:**20**(1):63–72. https://doi.org/10.1007/s11676-009-0012-4.

Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, *et al.* Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods.* 2013:**10**(6):563–569. https://doi.org/10.1038/nmeth.2474.

Chin CS, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, Dunn C, O'Malley R, Figueroa-Balderas R, Morales-Cruz A, *et al.* Phased diploid genome assembly with single-molecule real-time sequencing. *Nat Methods.* 2016:**13**(12):1050–1054. https://doi.org/10.1038/nmeth.4035.

Cruz MV, Mori GM, Oh DH, Dassanayake M, Zucchi MI, Oliveira RS, de Souza AP. Molecular responses to freshwater limitation in the mangrove tree *Avicennia germinans* (Acanthaceae). *Mol Ecol.* 2020:**29**(2):344–362. https://doi.org/10.1111/mec.15330.

D'haeseleer P. Reconstructing gene networks from large scale gene expression data [PhD thesis]. The University of New Mexico; 2000.

Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 1987:**19**:11–15. https://worldveg.tind.io/record/33886.

Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP, *et al. De novo* assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science* 2017:**356**(6333):92–95. https://doi.org/10.1126/science.aal3327.

Duke NC. Mangrove floristics and biogeography. In: *Tropical mangrove ecosystems*. Washington (DC): American Geophysical Union; 1992. p. 63–100.

Duke NC. *World Mangrove iD: expert information at your fingertips*. Australia: MangroveWatch Publication; 2017.

Duke NC, Ball MC, Ellison JC. Factors influencing biodiversity and distributional gradients in mangroves. *Glob Ecol Biogeogr Lett.* 1998:**7**(1):27–47. https://doi.org/10.2307/2997695.

Duke NC, Mackenzie JR, Canning AD, Hutley LB, Bourke AJ, Kovacs JM, Cormier R, Staben G, Lymburner L, Ai E. ENSO-driven extreme oscillations in mean sea level destabilise critical shoreline mangroves—an emerging threat. *PLOS Clim.* 2022:**1**(8):e0000037. https://doi.org/10.1371/journal.pclm.0000037.

Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, Aiden EL. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* 2016:**3**(1):95–98. https://doi.org/10.1016/j.cels.2016.07.002.

Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol.* 2011:**7**(10):e1002195. https://doi.org/10.1371/journal.pcbi.1002195.

Edger PP, Pires JC. Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes. *Chromosome Res.* 2009:**17**(5):699–717. https://doi.org/10.1007/s10577-009-9055-9.

Ellinghaus D, Kurtz S, Willhoeft U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* 2008:**9**(1):1–14. https://doi.org/10.1186/1471-2105-9-18.

Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 2019:**20**(1):1–14. https://doi.org/10.1186/s13059-019-1832-y.

Feng X, Chen Q, Wu W, Wang J, Li G, Xu S, Shao S, Liu M, Zhong C, Wu C, *et al.* Genomic evidence for rediploidization and adaptive evolution following the whole-genome triplication. *Nat Commun*. 2024:**15**(1):1635. https://doi.org/10.1038/s41467-024-46080-7.

Feng X, Li G, Xu S, Wu W, Chen Q, Shao S, Liu M, Wang N, Zhong C, He Z, *et al.* Genomic insights into molecular adaptation to intertidal environments in the mangrove *Aegiceras corniculatum*. *New Phytol*. 2021:**231**(6):2346–2358. https://doi.org/10.1111/nph.17551.

Feng X, Xu S, Li J, Yang Y, Chen Q, Lyu H, Zhong C, He Z, Shi S. Molecular adaptation to salinity fluctuation in tropical intertidal environments of a mangrove tree *Sonneratia alba*. *BMC Plant Biol*. 2020:**20**(1):1–14. https://doi.org/10.1186/s12870-020-02395-3.

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci U S A*. 2020:**117**(17):9451–9457. https://doi.org/10.1073/pnas.1921046117.

Fukushima K, Fang X, Alvarez-Ponce D, Cai H, Carretero-Paulet L, Chen C, Chang TH, Farr KM, Fujita T, Hiwatashi Y, *et al.* Genome of the pitcher plant *Cephalotus* reveals genetic changes associated with carnivory. *Nat Ecol Evol*. 2017:**1**(3):1–9. https://doi.org/10.1038/s41559-016-0059.

Garcia JDS, Dalmolin ÂC, França MGC, Mangabeira PAO. Different salt concentrations induce alterations both in photosynthetic parameters and salt gland activity in leaves of the mangrove *Avicennia schaueriana*. *Ecotoxicol Environ Saf*. 2017:**141**:70–74. https://doi.org/10.1016/j.ecoenv.2017.03.016.

Gil R, Bautista I, Boscaiu M, Lidon A, Wankhade S, Sanchez H, Llinares J, Vicente O. Responses of five Mediterranean halophytes to seasonal changes in environmental conditions. *AoB PLANTS*. 2014:**6**(0):plu049. https://doi.org/10.1093/aobpla/plu049.

Giovannoni SJ, Cameron Thrash J, Temperton B. Implications of streamlining theory for microbial ecology. *ISME J*. 2014:**8**(8):1553–1565. https://doi.org/10.1038/ismej.2014.60.

Graham A. Paleobotanical evidence and molecular data in reconstructing the historical phytogeography of Rhizophoraceae. *Ann Missouri Bot Gard*. 2006:**93**(2):325–334. https://doi.org/10.3417/0026-6493(2006)93[325:PEAMDI]2.0.CO;2.

Guelzim N, Bottani S, Bourgine P, Képès F. Topological and causal structure of the yeast transcriptional regulatory network. *Nat Genet*. 2002:**31**(1):60–63. https://doi.org/10.1038/ng873.

Guo Y, Amir A. Exploring the effect of network topology, mRNA and protein dynamics on gene regulatory network stability. *Nat Commun*. 2021:**12**(1):1–10. https://doi.org/10.1038/s41467-020-20314-w.

Guo Z, Li X, He Z, Yang Y, Wang W, Zhong C, Greenberg AJ, Wu CI, Duke NC, Shi S. Extremely low genetic diversity across mangrove taxa reflects past sea level changes and hints at poor future responses. *Glob Chang Biol*. 2018:**24**(4):1741–1748. https://doi.org/10.1111/gcb.13968.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Robin CR, Wortman JR. Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. *Genome Biol*. 2008:**9**(1):1–22. https://doi.org/10.1186/gb-2008-9-1-r7.

Harrison PJ, Khan N, Yin K, Saleem M, Bano N, Nisa M, Ahmed SI, Rizvi N, Azam F. Nutrient and phytoplankton dynamics in two mangrove tidal creeks of the Indus River delta, Pakistan. *Mar Ecol Prog Ser*. 1997:**157**:13–19. https://doi.org/10.3354/meps157013.

He Z, Feng X, Chen Q, Li L, Li S, Han K, Guo Z, Wang J, Liu M, Shi C, *et al.* Evolution of coastal forests based on a full set of mangrove genomes. *Nat Ecol Evol*. 2022:**6**(6):738–749. https://doi.org/10.1038/s41559-022-01744-9.

He Z, Xu S, Zhang Z, Guo W, Lyu H, Zhong C, Boufford DE, Duke NC, Shi S. Convergent adaptation of the genomes of woody plants at the land-sea interface. *Natl Sci Rev*. 2020:**7**(6):978–993. https://doi.org/10.1093/nsr/nwaa027.

Helsen J, Voordeckers K, Vanderwaeren L, Santermans T, Tsontaki M, Verstrepen KJ, Jelier R. Gene loss predictably drives evolutionary adaptation. *Mol Biol Evol*. 2020:**37**(10):2989–3002. https://doi.org/10.1093/molbev/msaa172.

Hoballah ME, Gübitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo A, Arnold M, Kuhlemeier C. Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell*. 2007:**19**(3):779–790. https://doi.org/10.1105/tpc.106.048694.

Hoekstra FA, Golovina EA, Buitink J. Mechanism of plant desiccation tolerance. *Trends Plant Sci*. 2001:**6**(9):431–438. https://doi.org/10.1016/S1360-1385(01)02052-0.

Hu Y, Wang X, Xu Y, Yang H, Tong Z, Tian R, Xu S, Yu L, Guo Y, Shi P, *et al.* Molecular mechanisms of adaptive evolution in wild animals and plants. *Sci China Life Sci*. 2023:**66**(3):453–495. https://doi.org/10.1007/s11427-022-2233-x.

Huelsmann M, Hecker N, Springer MS, Gatesy J, Sharma V, Hiller M. Genes lost during the transition from land to water in cetaceans highlight genomic changes associated with aquatic adaptations. *Sci Adv*. 2019:**5**(9):1–12. https://doi.org/10.1126/sciadv.aaw6671.

Huot B, Yao J, Montgomery BL, He SY. Growth-defense tradeoffs in plants: a balancing act to optimize fitness. *Mol Plant*. 2014:**7**(8):1267–1287. https://doi.org/10.1093/mp/ssu049.

Ibarra-Laclette E, Lyons E, Hernández-Guzmán G, Pérez-Torres CA, Carretero-Paulet L, Chang TH, Lan T, Welch AJ, Juárez MJA, Simpson J, *et al.* Architecture and evolution of a minute plant genome. *Nature* 2013:**498**(7452):94–98. https://doi.org/10.1038/nature12132.

Isah T. Stress and defense responses in plant secondary metabolites production. *Biol Res*. 2019:**52**(1):39. https://doi.org/10.1186/s40659-019-0246-3.

Jia H, McCarty DR, Suzuki M. Distinct roles of LAFL network genes in promoting the embryonic seedling fate in the absence of VAL repression. *Plant Physiol*. 2013:**163**(3):1293–1305. https://doi.org/10.1104/pp.113.220988.

Jin Y, Tam OH, Paniagua E, Hammell M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* 2015:**31**(22):3593–3599. https://doi.org/10.1093/bioinformatics/btv422.

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, *et al.* InterProScan 5: genome-scale protein function classification. *Bioinformatics* 2014:**30**(9):1236–1240. https://doi.org/10.1093/bioinformatics/btu031.

Kang M, Wang J, Huang H. Nitrogen limitation as a driver of genome size evolution in a group of karst plants. *Sci Rep*. 2015:**5**:11636. https://doi.org/10.1038/srep11636

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013:**30**(4):772–780. https://doi.org/10.1093/molbev/mst010.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019:**37**(8):907–915. https://doi.org/10.1038/s41587-019-0201-4.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013:**14**(4):1–13. https://doi.org/10.1186/gb-2013-14-4-r36.

Kloster MB. The determination of tannin and lignin. *J Am Water Works Assoc*. 1974:**66**(1):44–46. https://doi.org/10.1002/j.1551-8833.1974.tb01970.x.

Kovaka S, Zimin AV, Pertea GM, Razaghi R, Salzberg SL, Pertea M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol*. 2019:**20**(1):1–13. https://doi.org/10.1186/s13059-019-1910-1.

Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 2019:**35**(21):4453–4455. https://doi.org/10.1093/bioinformatics/btz305.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009:**19**(9):1639–1645. https://doi.org/10.1101/gr.092759.109.

Kushwaha R, Lloyd TD, Schäfermeyer KR, Kumar S, Downie AB. Identification of Late Embryogenesis Abundant (LEA) protein putative interactors using phage display. *Int J Mol Sci*. 2012:**13**(6):6582–6603. https://doi.org/10.3390/ijms13066582.

Lechthaler S, Robert EMR, Tonné N, Prusova A, Gerkema E, Van As H, Koedam N, Windt CW. Rhizophoraceae mangrove saplings use hypocotyl and leaf water storage capacity to cope with soil water salinity changes. *Front Plant Sci*. 2016:**7**:1–13. https://doi.org/10.3389/fpls.2016.00895.

Lee H, Golicz AA, Bayer PE, Severn-Ellis AA, Chan CKK, Batley J, Kendrick GA, Edwards D. Genomic comparison of two independent seagrass lineages reveals habitat-driven convergent evolution. *J Exp Bot*. 2018:**69**(15):3689–3702. https://doi.org/10.1093/jxb/ery147.

Le Gall H, Philippe F, Domon JM, Gillet F, Pelloux J, Rayon C. Cell wall metabolism in response to abiotic stress. *Plants* 2015:**4**(1): 112–166. https://doi.org/10.3390/plants4010112.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics* 2009:**25**(16):2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Liang S, Fang L, Zhou R, Tang T, Deng S, Dong S, Huang Y, Zhong C, Shi S. Transcriptional homeostasis of a mangrove species, *Ceriops tagal*, in saline environments, as revealed by microarray analysis. *PLoS One* 2012:**7**:e36499. https://doi.org/10.1371/journal.pone.0036499

Liang S, Zhou RC, Dong SS, Shi SH. Adaptation to salinity in mangroves: implication on the evolution of salt-tolerance. *Chinese Sci Bull*. 2008:**53**(11):1708–1715. https://doi.org/10.1007/s11434-008-0221-9.

Lin G, Da SL, Sternberg L. Effects of salinity fluctuation on photosynthetic gas exchange and plant growth of the red mangrove (*Rhizophora mangle* L.). *J Exp Bot*. 1993:**44**(1):9–16. https://doi.org/10.1093/jxb/44.1.9.

Liu Y, Zeng Z, Zhang YM, Li Q, Jiang XM, Jiang Z, Tang JH, Chen D, Wang Q, Chen JQ, *et al.* An angiosperm NLR Atlas reveals that NLR gene reduction is associated with ecological specialization and signal transduction component deletion. *Mol Plant*. 2021:**14**(12):2015–2031. https://doi.org/10.1016/j.molp.2021.08.001.

Lomsadze A, Burns PD, Borodovsky M. Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Res*. 2014:**42**(15):e119–e119. https://doi.org/10.1093/nar/gku557.

López-Ribera I, La Paz JL, Repiso C, García N, Miquel M, Hernández ML, Martínez-Rivas JM, Vicient CM. The evolutionary conserved oil body associated protein OBAP1 participates in the regulation of oil body size. *Plant Physiol*. 2014:**164**(3):1237–1249. https://doi.org/10.1104/pp.113.233221.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014:**15**(12):1–21. https://doi.org/10.1186/s13059-014-0550-8.

Lunn D, Wallis JG, Browse J. Overexpression of Seipin1 increases oil in hydroxy fatty acid-accumulating seeds. *Plant Cell Physiol*. 2018:**59**(1):205–214. https://doi.org/10.1093/pcp/pcx177.

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Yunjie L, *et al.* SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience* 2012:**1**(1):18. https://doi.org/10.1186/2047-217X-1-18.

Lyu H, He Z, Wu CI, Shi S. Convergent adaptive evolution in marginal environments: unloading transposable elements as a common strategy among mangrove genomes. *New Phytol*. 2018:**217**(1): 428–438. https://doi.org/10.1111/nph.14784.

Majoros WH, Pertea M, Salzberg SL. TigrScan and GlimmerHMM: two open source *ab initio* eukaryotic gene-finders. *Bioinformatics* 2004:**20**(16):2878–2879. https://doi.org/10.1093/bioinformatics/bth315.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021:**38**(10): 4647–4654. https://doi.org/10.1093/molbev/msab199.

May RM. Will a large complex system be stable? *Nature* 1972:**238**(5364):413–414. https://doi.org/10.1038/238413a0.

Méndez-Alonzo R, López-Portillo J, Moctezuma C, Bartlett MK, Sack L. Osmotic and hydraulic adjustment of mangrove saplings to extreme salinity. *Tree Physiol*. 2016:**36**(12):1562–1572. https://doi.org/10.1093/treephys/tpw073.

Miquel M, Trigui G, D'Andréa S, Kelemen Z, Baud S, Berger A, Deruyffelaere C, Trubuil A, Lepiniec L, Dubreucq B. Specialization of oleosins in oil body dynamics during seed development in arabidopsis seeds. *Plant Physiol*. 2014:**164**(4): 1866–1878. https://doi.org/10.1104/pp.113.233262.

Mistry J, Bateman A, Finn RD. Predicting active site residue annotations in the Pfam database. *BMC Bioinformatics* 2007:**8**(1):1–14. https://doi.org/10.1186/1471-2105-8-298.

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, *et al.* Pfam: the protein families database in 2021. *Nucleic Acids Res*. 2021:**49**(D1):D412–D419. https://doi.org/10.1093/nar/gkaa913.

Monroe JG, McKay JK, Weigel D, Flood PJ. The population genomics of adaptive loss of function. *Heredity (Edinb)*. 2021:**126**(3): 383–395. https://doi.org/10.1038/s41437-021-00403-2.

Monroe JG, Powell T, Price N, Mullen JL, Howard A, Evans K, Lovell JT, McKay JK. Drought adaptation in *Arabidopsis thaliana* by extensive genetic loss-of-function. *Elife* 2018:**7**:1–18. https://doi.org/10.7554/eLife.41038.

Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ. The timescale of early land plant evolution. *Proc Natl Acad Sci U S A*. 2018:**115**(10):E2274-E2283. https://doi.org/10.1073/pnas.1719588115.

Muller J. Fossil pollen records of extant angiosperms. *Bot Rev*. 1981:**47**(1):1–142. https://doi.org/10.1007/BF02860537.

Nguyen HT, Stanton DE, Schmitz N, Farquhar GD, Ball MC. Growth responses of the mangrove *Avicennia marina* to salinity: development and function of shoot hydraulic systems require saline conditions. *Ann Bot*. 2015:**115**(3):397–407. https://doi.org/10.1093/aob/mcu257.

Nishimura N, Tsuchiya W, Moresco JJ, Hayashi Y, Satoh K, Kaiwa N, Irisa T, Kinoshita T, Schroeder JI, Yates JR, *et al.* Control of seed dormancy and germination by DOG1-AHG1 PP2C phosphatase complex via binding to heme. *Nat Commun*. 2018:**9**(1):1–14. https://doi.org/10.1038/s41467-018-04437-9.

Okamoto M, Kuwahara A, Seo M, Kushiro T, Asami T, Hirai N, Kamiya Y, Koshiba T, Nambara E. CYP707A1 and CYP707A2, which encode abscisic acid 8′-hydroxylases, are indispensable for proper control of seed dormancy and germination in Arabidopsis. *Plant Physiol*. 2006:**141**(1):97–107. https://doi.org/10.1104/pp.106.079475.

Ou S, Jiang N. LTR_retriever: a highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiol*. 2018:**176**(2):1410–1422. https://doi.org/10.1104/pp.17.01310.

Ou S, Jiang N. LTR_FINDER_parallel: parallelization of LTR_FINDER enabling rapid identification of long terminal repeat retrotransposons. *Mob DNA*. 2019:**10**(1):1–3. https://doi.org/10.1186/s13100-019-0193-0.

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, *et al.* The TIGR rice genome annotation resource: improvements and new features. *Nucleic Acids Res*. 2007:**35**(Database):D883–D887. https://doi.org/10.1093/nar/gkl976.

Palfalvi G, Hackl T, Terhoeven N, Shibata TF, Nishiyama T, Ankenbrand M, Becker D, Förster F, Freund M, Iosip A, *et al.* Genomes of the *Venus flytrap* and close relatives unveil the roots of plant carnivory. *Curr Biol*. 2020:**30**(12):2312–2320.e5. https://doi.org/10.1016/j.cub.2020.04.051.

Pfeiffer P, Goedecke W, Obe G. Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. *Mutagenesis* 2000:**15**(4):289–302. https://doi.org/10.1093/mutage/15.4.289.

Pourcel L, Routaboul JM, Kerhoas L, Caboche M, Lepiniec L, Debeaujon I. *TRANSPARENT TESTA10* encodes a laccase-like enzyme involved in oxidative polymerization of flavonoids in Arabidopsis seed coat. *Plant Cell*. 2005:**17**(11):2966–2980. https://doi.org/10.1105/tpc.105.035154.

Pryszcz LP, Gabaldón T. Redundans: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Res*. 2016:**44**(12):e113. https://doi.org/10.1093/nar/gkw294.

Reef R, Feller IC, Lovelock CE. Nutrition of mangroves. *Tree Physiol*. 2010:**30**(9):1148–1160. https://doi.org/10.1093/treephys/tpq048.

Reef R, Lovelock CE. Regulation of water balance in mangroves. *Ann Bot*. 2015:**115**(3):385–395. https://doi.org/10.1093/aob/mcu174.

Ren R, Wang H, Guo C, Zhang N, Zeng L, Chen Y, Ma H, Qi J. Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Mol Plant*. 2018:**11**(3):414–428. https://doi.org/10.1016/j.molp.2018.01.002.

Ren Z, Zheng Z, Chinnusamy V, Zhu J, Cui X, Iida K, Zhu JK. RAS1, a quantitative trait locus for salt tolerance and ABA sensitivity in *Arabidopsis*. *Proc Natl Acad Sci U S A*. 2010:**107**(12):5669–5674. https://doi.org/10.1073/pnas.0910798107.

Rivera-Monroy VH, Lee SY, Kristensen E, Twilley RR. *Mangrove ecosystems: a global biogeographic perspective: structure, function, and services*. Cham: Springer; 2017.

Roach MJ, Schmidt SA, Borneman AR. Purge haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics* 2018:**19**(1):1–10. https://doi.org/10.1186/s12859-018-2485-7.

Sabath N, Ferrada E, Barve A, Wagner A. Growth temperature and genome size in bacteria are negatively correlated, suggesting genomic streamlining during thermal adaptation. *Genome Biol Evol*. 2013:**5**(5):966–977. https://doi.org/10.1093/gbe/evt050.

Sas C, Müller F, Kappel C, Kent TV, Wright SI, Hilker M, Lenhard M. Repeated inactivation of the first committed enzyme underlies the loss of benzaldehyde emission after the selfing transition in *Capsella*. *Curr Biol*. 2016:**26**(24):3313–3319. https://doi.org/10.1016/j.cub.2016.10.026.

Schuermann D, Molinier J, Fritsch O, Hohn B. The dual nature of homologous recombination in plants. *Trends Genet*. 2005:**21**(3):172–181. https://doi.org/10.1016/j.tig.2005.01.002.

Seppey M, Manni M, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol Biol*. 2019:**1962**:227–245. https://doi.org/10.1007/978-1-4939-9173-0_14.

Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, Heard E, Dekker J, Barillot E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. 2015:**16**(1):1–11. https://doi.org/10.1186/s13059-015-0831-x.

Sharma V, Hecker N, Roscito JG, Foerster L, Langer BE, Hiller M. A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. *Nat Commun*. 2018:**9**(1):1–9. https://doi.org/10.1038/s41467-018-03667-1.

Shi S, Huang Y, Zeng K, Tan F, He H, Huang J, Fu Y. Molecular phylogenetic analysis of mangroves: independent evolutionary origins of vivipary and salt secretion. *Mol Phylogenet Evol*. 2005:**34**(1):159–166. https://doi.org/10.1016/j.ympev.2004.09.002.

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: *ab initio* prediction of alternative transcripts. *Nucleic Acids Res*. 2006:**34**(Web Server):W435–W439. https://doi.org/10.1093/nar/gkl200.

Stolzer M, Lai H, Xu M, Sathaye D, Vernot B, Durand D. Inferring duplications, losses, transfers and incomplete lineage sorting with nonbinary species trees. *Bioinformatics* 2012:**28**(18):i409–i415. https://doi.org/10.1093/bioinformatics/bts386.

Strickler SS, Gribenko Alexey V, Gribenko Alexander V, Keiffer TR, Tomlinson J, Reihle T, Loladze VV, Makhatadze GI. Protein stability and surface electrostatics: a charged relationship. *Biochemistry* 2006:**45**(9):2761–2766. https://doi.org/10.1021/bi0600143.

Studer A, Poulin R. Cercarial survival in an intertidal trematode: a multifactorial experiment with temperature, salinity and ultraviolet radiation. *Parasitol Res*. 2013:**112**(1):243–249. https://doi.org/10.1007/s00436-012-3131-3.

Su BY, Wang YS, Sun CC. Response adaptive mechanisms of three mangrove (*Avicennia marina*, *Aegiceras corniculatum*, and *Bruguiera gymnorrhiza*) plants to waterlogging stress revealed by transcriptome analysis. *Front Mar Sci*. 2022:**9**:1–15. https://doi.org/10.3389/fmars.2022.929649.

Sun G, Xu Y, Liu H, Sun T, Zhang J, Hettenhausen C, Shen G, Qi J, Qin Y, Li J, *et al.* Large-scale gene losses underlie the genome evolution of parasitic plant *Cuscuta australis*. *Nat Commun*. 2018:**9**(1):4–11. https://doi.org/10.1038/s41467-018-04721-8.

Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res*. 2006:**34**(Web Server):W609–W612. https://doi.org/10.1093/nar/gkl315.

Tarailo-Graovac M, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinforma*. 2009:**25**(1):4.10.1–4.10.14. https://doi.org/10.1002/0471250953.bi0410s25.

Tomlinson PB. *The botany of mangroves*. Cambridge: Cambridge University Press; 2016.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012:**7**(3):562–578. https://doi.org/10.1038/nprot.2012.016.

Van de Peer Y, Mizrachi E, Marchal K. The evolutionary significance of polyploidy. *Nat Rev Genet*. 2017:**18**(7):411–424. https://doi.org/10.1038/nrg.2017.26.

Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate *de novo* genome assembly from long uncorrected reads. *Genome Res*. 2017:**27**(5):737–746. https://doi.org/10.1101/gr.214270.116.

Wagner A. Energy constraints on the evolution of gene expression. *Mol Biol Evol*. 2005:**22**(6):1365–1374. https://doi.org/10.1093/molbev/msi126.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, *et al.* Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014:**9**(11):e112963. https://doi.org/10.1371/journal.pone.0112963.

Wan L, Ross ARS, Yang J, Hegedus DD, Kermode AR. Phosphorylation of the 12 S globulin cruciferin in wild-type and *abi1-1* mutant *Arabidopsis thaliana* (thale cress) seeds. *Biochem J*. 2007:**404**(2):247–256. https://doi.org/10.1042/BJ20061569.

Wang P, Moore BM, Panchy NL, Meng F, Lehti-Shiu MD, Shiu SH. Factors influencing gene family size variation among related species in a plant family, Solanaceae. *Genome Biol Evol*. 2018:**10**(10):2596–2613. https://doi.org/10.1093/gbe/evy193.

Wang L, Pan D, Lv X, Cheng CL, Li J, Liang W, Xing J, Chen W. A multilevel investigation to discover why *Kandelia candel* thrives in high salinity. *Plant Cell Environ*. 2016:**39**(11):2486–2497. https://doi.org/10.1111/pce.12804.

Wang Y, Sheng HF, He Y, Wu JY, Jiang YX, Tam NFY, Zhou HW. Comparison of the levels of bacterial diversity in freshwater, intertidal wetland, and marine sediments by using millions of Illumina tags. *Appl Environ Microbiol*. 2012:**78**(23):8264–8271. https://doi.org/10.1128/AEM.01821-12.

Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, *et al.* *MCScanx*: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*. 2012:**40**(7):e49-e49. https://doi.org/10.1093/nar/gkr1293.

Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. Kaks_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics Proteomics Bioinforma*. 2010:**8**(1):77–80. https://doi.org/10.1016/S1672-0229(10)60008-3.

Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. Direct determination of diploid genome sequences. *Genome Res*. 2017:**27**(5):757–767. https://doi.org/10.1101/gr.214874.116.

Wells JT. Dynamics of coastal fluid muds in low-, moderate-, and high-tide-range environments. *Can J Fish Aquat Sci*. 1983:**40**(S1):s130–s142. https://doi.org/10.1139/f83-276.

Wolf YI, Koonin EV. Genome reduction as the dominant mode of evolution. *BioEssays*. 2013:**35**(9):829–837. https://doi.org/10.1002/bies.201300037.

Wolf YI, Makarova KS, Yutin N, Koonin EV. Updated clusters of orthologous genes for archaea: a complex ancestor of the archaea and the byways of horizontal gene transfer. *Biol Direct*. 2012:**7**(1):1–15. https://doi.org/10.1186/1745-6150-7-46.

Wu S, Han B, Jiao Y. Genetic contribution of paleopolyploidy to adaptive evolution in angiosperms. *Mol Plant*. 2019:**13**(1):59–71. https://doi.org/10.1016/j.molp.2019.10.012.

Wu W, Liu X, Wang M, Meyer RS, Luo X, Ndjiondjop MN, Tan L, Zhang J, Wu J, Cai H, *et al.* A single-nucleotide polymorphism causes smaller grain size and loss of seed shattering during African rice domestication. *Nat Plants*. 2017:**3**(6):1–7. https://doi.org/10.1038/nplants.2017.64.

Xiao CL, Chen Y, Xie SQ, Chen KN, Wang Y, Han Y, Luo F, Xie Z. MECAT: fast mapping, error correction, and *de novo* assembly for single-molecule sequencing reads. *Nat Methods*. 2017:**14**(11):1072–1074. https://doi.org/10.1038/nmeth.4432.

Xiao X, Hong Y, Xia W, Feng S, Zhou X, Fu X, Zang J, Xiao Y, Niu X, Li C, *et al.* Transcriptome analysis of *Ceriops tagal* in saline environments using RNA-sequencing. *PLoS One*. 2016:**11**(12):1–19. https://doi.org/10.1371/journal.pone.0167551.

Xie W, Guo Z, Wang J, He Z, Li Y, Feng X, Zhong C, Shi S. Evolution of woody plants to the land-sea interface—the atypical genomic features of mangroves with atypical phenotypic adaptation. *Mol Ecol*. 2023:**32**(6):1351–1365. https://doi.org/10.1111/mec.16587.

Xu S, Guo Z, Feng X, Shao S, Yang Y, Li J, Zhong C, He Z, Shi S. Where whole-genome duplication is most beneficial: adaptation of mangroves to a wide salinity range between land and sea. *Mol Ecol*. 2023:**32**(2):460–475. https://doi.org/10.1111/mec.16320.

Xu S, He Z, Zhang Z, Guo Z, Guo W, Lyu H, Li J, Yang M, Du Z, Huang Y, *et al.* The origin, diversification and adaptation of a major mangrove clade (Rhizophoreae) revealed by whole-genome sequencing. *Natl Sci Rev*. 2017:**4**(5):721–734. https://doi.org/10.1093/nsr/nwx065.

Xu YC, Niu XM, Li XX, He W, Chen JF, Zou YP, Wu Q, Zhang YE, Busch W, Guoa YL. Adaptation and phenotypic diversification in Arabidopsis through loss-of-function mutations in protein-coding genes. *Plant Cell*. 2019:**31**(5):1012–1025. https://doi.org/10.1105/tpc.18.00791.

Xu Z, Wang H. LTR-FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res*. 2007:**35**(Web Server):W265–W268. https://doi.org/10.1093/nar/gkm286.

Xu S, Wang J, Guo Z, He Z, Shi S. Genomic convergence in the adaptation to extreme environments. *Plant Commun*. 2020:**1**(6):100117. https://doi.org/10.1016/j.xplc.2020.100117.

Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007:**24**(8):1586–1591. https://doi.org/10.1093/molbev/msm088.

Yang SC, Shih SS, Hwang GW, Adams JB, Lee HY, Chen CP. The salinity gradient influences on the inundation tolerance thresholds of mangrove forests. *Ecol Eng*. 2013:**51**:59–65. https://doi.org/10.1016/j.ecoleng.2012.12.049.

Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012:**16**(5):284–287. https://doi.org/10.1089/omi.2011.0118.

Zachos JC, Dickens GR, Zeebe RE. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 2008:**451**(7176):279–283. https://doi.org/10.1038/nature06588.

Zhang B, Hou L, Qi H, Hou L, Zhang T, Zhao F, Miao M. An extremely streamlined macronuclear genome in the free-living protozoan *Fabrea salina*. *Mol Biol Evol*. 2022:**39**(4):msac062. https://doi.org/10.1093/molbev/msac062.

Zhang T, Qiao Q, Novikova PY, Wang Q, Yue J, Guan Y, Ming S, Liu T, De J, Liu Y, *et al.* Genome of *Crucihimalaya himalaica*, a close relative of Arabidopsis, shows ecological adaptation to high altitude. *Proc Natl Acad Sci U S A*. 2019:**116**(14):7137–7146. https://doi.org/10.1073/pnas.1817580116.

Zhu R, Shao S, Xie W, Guo Z, He Z, Li Y, Wang W, Zhong C, Shi S, Xu S. High-quality genome of a pioneer mangrove *Laguncularia racemosa* explains its advantages for intertidal zone reforestation. *Mol Ecol Resour*. 2023:**00**:1–16. https://doi.org/10.1111/1755-0998.13863.