

# Introgression and chloroplast capture from octoploids confer stress tolerance to allotetraploid common reed (*Phragmites australis*)

Cui Wang<sup>1</sup>, Lele Liu<sup>2</sup>, Wenyi Sheng<sup>2</sup>, Meiqi Yin<sup>2</sup>, Carla Lambertini<sup>3</sup>, Xiao Guo<sup>4</sup>, Chunyan Li<sup>5</sup>, and Weihua Guo<sup>2</sup>

<sup>1</sup>Affiliation not available

<sup>2</sup>Shandong University

<sup>3</sup>Università degli Studi di Milano

<sup>4</sup>Qingdao Agricultural University

<sup>5</sup>Qingdao baomaide Biotechnology Co.,Ltd

September 04, 2024

## Abstract

Hybridization is a potent mechanism for generating genetic diversity and transferring adaptive genetic loci across species or populations (lineages), enabling organisms to explore broader ecological niches. However, the effects of hybridization between species/lineages with different ploidy levels remain underexplored. In this study, we used transcriptomic approaches to investigate the abiotic stress tolerance of common reed (*Phragmites australis*) and in relation to polyploidy and hybridization. Our findings revealed that a brackish water population of a tetraploid lineage acquired salinity tolerance through adaptive introgression from an octoploid lineage. Among 46 adaptive genes with high  $F_{ST}$  values between populations, nine were significantly enriched in response to salicylic acid. In a common garden experiment, we chose two hybrid genotypes with similar genetic backgrounds to assess the consistency of adaptability from introgression under varying cadmium (Cd) concentrations. The Heze genotype exhibited a significant decline in net photosynthesis rate and chlorophyll content with increasing Cd concentrations, whereas the Panjin genotype exhibited greater tolerance. Correspondingly, the Heze genotype also had a higher number of differentially expressed genes under both low and high Cd concentrations. These findings suggest that introgressed loci may have varying functions in hybrid populations. This study enhances our understanding of how hybridization between lineages of different ploidy levels can lead to improved stress tolerance and its role in adaptive evolution.

## Introduction

Plant species inhabiting diverse environmental conditions often undergo a series of local adaptations, leading to adaptive phenotypic traits and genetic changes that can result in incipient speciation or population divergence. Genetic elements, including the nuclear genome, coevolve synergistically with organelles, resulting in coordinated adaptations (Roux et al., 2016). However, this dynamic can shift in secondary contact zones or adjacent distribution areas where hybridization frequently occurs. In these regions, genetic introgression—where foreign genetic material from one species is introduced into another—can significantly impact the evolutionary trajectory of the recipient species. Additionally, organelle capture where a species retains and gradually replaces its original organelles (e.g., chloroplasts and mitochondria) with those from another species can occur. On one hand, organelle capture may introduce cytonuclear incompatibility and potentially lead to cytonuclear conflicts in genetically admixed individuals (Case, Finseth, Barr, & Fishman, 2016; Ellison & Burton, 2010; Sloan et al., 2018). On the other hand, this process can also interactively influence the reorganization of nuclear genome from the parental species, selectively retaining coevolved adaptive genes in the nuclear genome, and ultimately altering the local adaptivity of the recipient species (H. Wang et al., 2024).

Polyploidization is a widespread phenomenon in nature, with multiple polyploidization events occurring in closely related species or even within the same species (Brysting, Fay, Leitch, & Aiken, 2004; Cordeiro & Felix, 2018; Figueiredo et al., 2014; Wagner et al., 2019). Polyploids are known for their advantageous adaptation to various environments, often outperforming their diploid counterparts in terms of tolerance to ambient stresses such as drought, high salinity, cold, heat and nutrient deficits (Del Pozo & Ramirez-Parra, 2015). Hybridization between different ploidy levels is also common in nature. *Adaptive introgression, where adaptive loci are transferred from one species or population to another, has been observed in various plant species, including Arabidopsis (Baduel, Hunter, Yeola, & Bomblies, 2018). This process can introduce beneficial phenotypic traits from one population to another. However, whether these adaptive loci can function stably in hybrid populations remains unclear.*

*Phragmites* are a group of aquatic plants that have been playing important roles in the wetland ecosystems. They are commonly seen in streams, lowland rivers, estuaries, and creeks, and have high adaptability to different salinity regimes (Srivastava, Kalra, & Naraian, 2014). Genetically, the common reed has diverged into eight phylogeographic lineages worldwide (C Lambertini et al., 2006; L. L. Liu et al., 2022; Saltonstall, 2002; C. Wang, Liu, Yin, Eller, Brix, Wang, Salojärvi, et al., 2021). Using non-coding chloroplast regions trnT-trnL and rbcL-psaI, researchers have identified up to 57 haplotypes that based on their phylogenetic relationships can be used to distinguish these phylogeographic lineages (Eller et al., 2017; Saltonstall, 2016). In China, two major haplotypes were discovered: haplotype O (denoted as CN) primarily distributed in northern China, and haplotype P (denoted as AU) found in Korea and middle to southern China (L. L. Liu et al., 2022; Saltonstall, 2002; Tanaka, Irbis, & Inamura, 2017). These two distinct haplotypes are distantly related with haplotype O associated with the Eurasian lineage, and haplotype P belonging to a lineage primarily distributed in East Asia and Australia. They also differ in ploidy levels: haplotype O is tetraploid while Haplotype P is octoploid (Carla Lambertini et al., 2020; L. L. Liu et al., 2022; C. Wang, Liu, Yin, Eller, Brix, Wang, Salojärvi, et al., 2021). Ploidy level, along with sexual (seed production) and asexual (rhizome growth) reproductive modes, are key traits contributing to the strong environmental tolerance of *Phragmites* species (Guan et al., 2023; Meyerson et al., 2016; Te Beest et al., 2012). Population with higher genetic diversity, which may correlate with higher ploidy levels, have been found to be more tolerant to high salinity (Sun et al., 2021). The distribution of ploidy levels appears geographically specific, and individuals with the same ploidy level may have formed from different polyploidization events (C. Wang, Liu, Yin, Eller, Brix, Wang, Salojärvi, et al., 2021).

Ambient heavy metal pollution is another significant source of stress for plants. Cadmium (Cd), a nonessential and hazardous element, often contaminates aquatic ecosystems due to its water-soluble compounds, which facilitate rapid spread in soil and water sediments (Wright & Welbourn, 1994). Cadmium exposure can severely disrupt plant growth and metabolism, leading to reduced root elongation and biomass, water and nutrient deficiencies, and decreased enzyme activity (Kahle, 1993). In response, plants activate detoxification processes, such as increased synthesis of chelating peptides to sequester cadmium ions and restricted ion transport to compartmentalize cadmium in root vacuoles (Zhu, Li, Duan, Liu, & Chen, 2021). Aquatic plant species can extend their resistance to heavy metal stress through reduced water flow and interactions with rhizosphere microorganisms involved in phytoremediation processes (Javed et al., 2019).

The bioaccumulation of heavy metal has been well studied in *P. australis* (Bonanno & Giudice, 2010; Březinová & Vymazal, 2022). Common reed demonstrates high tolerance and adaptation to heavy metals by increasing rhizosphere surface area when exposed to polluted sediments (Cicero-Fernández, Peña-Fernández, Expósito-Camargo, & Antizar-Ladislao, 2016). Given its wide distribution range, and significant intraspecific phylogeographic variation, the effect of genetic divergence should be considered when assessing bioaccumulation and stress tolerance. Indeed, different lineages of common reed may exhibit diversified plasticity in their functional and physiological traits, showcasing high adaptability to local environmental conditions. These variations in plasticity may result in distinct responses to environmental stresses, such as temperature fluctuations or changes in CO<sub>2</sub> levels (Eller et al., 2017). Although high levels of cadmium can be lethal to plants, lower levels (e.g., 50 µM) may not significantly perturb gene expression in some species, such as rice (Oono et al., 2016). Therefore, including a gradient of cadmium levels in experiments can be helpful for

testing stress tolerance, plant responses and phytoremediation potential.

Transcriptomics has proven to be an effective tool for constructing phylogenetic relationships among individuals and detecting local adaptations in diverging species (Pieri et al., 2024). In this study, we aim to use transcriptomic approaches to explore the stress tolerance of individuals resulting from hybridizations between octoploids and tetraploids in contact zones. We address the following questions: 1) Are octoploids more tolerant to abiotic stresses such as salinity than tetraploids? 2) Is there adaptive introgression between ploidy levels that enhances genetic tolerance in hybrids? 3) If adaptive introgression related to abiotic stress tolerance occurs, do the introgressed loci function stably within the hybrid population? To answer these questions, we first constructed a phylogenetic tree encompassing populations of different ecotypes and genetic lineages to determine their evolutionary relationships. We then conducted a genome scan to identify polymorphic loci between the AU and CN lineages that may be under natural selection. Finally, we examined the gene expression profiles of hybrid individuals with similar genetic backgrounds to assess whether stress tolerance remains stable in hybrid populations.

## Materials and Methods

### RNAseq reads obtained from NCBI database

We obtained RNAseq data of 32 individuals and whole genome sequencing reads for two samples from the NCBI database, representing different geographic origins (**Supplementary Table S1**). Identifications of brackish water and freshwater ecotypes were justified in the original publications (**Supplementary Table S1**). Of these, RNAseq data from 13 individuals were generated from a common garden in Shandong University in July 2020, allowing for direct comparison of gene expression profiles. Ploidy level information was sourced from previous studies, determined either through flow cytometry or predicted using alternative allele frequency after genome mapping (C. Wang, Liu, Yin, Eller, Brix, Wang, Salojärvi, et al., 2021). Specifically, individuals in the AU lineage are octoploid, while those in the USnat, EU, Med, and CN lineages are tetraploid. The reference genome sequence of *P. australis* was obtained from our previous work (NCBI accession: ASM4037322v1) (C. Wang et al., 2024). For phylogenetic analysis, we included the 34 samples as well as the six RNAseq samples of the Heze and Panjin genotype control samples.

### Read mapping and variant identification

The quality of the RNA-seq reads was checked with fastqc (Andrews 2010). Reads and bases with a Phred score higher than 20 were retained (Martin, 2011). For transcriptomic data analysis, clean reads were aligned to the *P. australis* reference genome of a Chinese individual, using the STAR aligner two-pass procedure (Dobin et al., 2013), and the resulting bam files were sorted with Samtools (Li et al., 2009). To ensure data accuracy, mapped reads from the transcriptomic data were screened for duplicates using the ‘MarkDuplicates’ function in the GATK pipeline (v 4.5.0.0), and any artificial duplicates resulting from the sequencing process were removed. Then we used the “SplitNCigarReads” tool on the RNA-seq samples to split reads at intronic regions. Genomic variants were called using DeepVariant (1.1.0), with the argument `-model_type WES`. Joint variant calling across all samples was performed using GLnexus, including an additional joint call for each population with all invariant sites, configured for DeepVariantWES. Although DeepVariant is optimized for SNP calling in diploids, we used this approach due to its high accuracy. To filter out false variants, only biallelic variants with less than 50% missingness were retained, and variants with a minor allele frequency (MAF) below 0.05 were filtered out using VCFtools. To avoid the effects of partially called SNPs, we filtered SNPs using the `-vcf-half-call missing` parameter. Finally, lineages were grouped according to their phylogenetic classification.

### Genetic diversity, $F_{ST}$ and genome scan

We calculated genetic diversity ( $\pi$  and  $\theta$ ) following the methods outlined in Pieri et al. (2024) and C. Wang, Liu, Yin, Eller, Brix, Wang, Salojärvi, et al. (2021). Briefly, ancestral alleles were first annotated by mapping the fastq files of six Poaceae species to the reference genome. The ancestral state of the alleles was determined by selecting the most common bases using ANGSD (Korneliussen, Albrechtsen, & Nielsen, 2014). Based on

this ancestral information, we calculated  $\pi$  and  $\vartheta$  from the unfolded site frequency spectra using the bam files. Additionally, we estimated  $F_{ST}$  between the AU and CN populations, as well as between the brackish water and freshwater populations within the CN lineage, using ANGSD with a window size of 10kb. We selected SNPs with  $F_{ST}$  values in the top 1%, considering these as potential sites under selection.

### Plant treatment

The plants used for the cadmium stress experiments were grown in the greenhouse of Shandong University, Qingdao, China (36°23'10" N, 120deg36'44" E). The greenhouse equipped with glass ceilings receives natural light, with light intensity similar with natural conditions and a photoperiod of 12-14 hours of daylight from May to October. The average daily temperature was maintained at 30.2 and the air humidity around 85% by ventilating fans. The same greenhouse settings were maintained throughout the experimental period to guarantee controlled and constant growth conditions. First of all, rhizomes of two genotypes Heze and Panjin of *P. australis* were collected from the common garden in May, 2022. After one week in the water, the rhizomes were cut into pieces of similar lengths with buds and placed in the pots. Each rhizome piece was planted in a pot with mixed vermiculite and nutrient soil (in a volumn ratio of 2:1) as the substrate. The pots had a top diameter of 16.0 cm, a bottom diameter of 13.0 cm, and a height of 17.5 cm. The water saturated soil weight was about 1.25 kg in each pot, and the dry weight was about 0.35 kg. Three clones grown from the rhizomes of the same genotype were used as independent biological replicates (in separate pots) for each treatment. The concentration of the Cd solution was made of 1.61g CdCl<sub>2</sub> per liter of water, i.e., 1mg mL<sup>-1</sup> Cd. A survey of spatial and temporal distribution of Cd pollution in farmland soil showed that the farmland soil Cd concentrations ranged from 0.012 to 23.33 mg\*kg<sup>-1</sup> in China, with the geometric mean of 0.473 mg\*kg<sup>-1</sup> (J. Wang, Wei, & Pan, 2023). According to the control standards of pollutants in sludge for agricultural use in China issued in 2018 (GB4284-2018), the total cadmium content below 3 mg kg<sup>-1</sup> can be classified as Class A pollutants, while the total cadmium content below 15 mg kg<sup>-1</sup> can be classified as Class B pollutants. Class A sludge products can be used for farmland, gardens, and pasture, whereas Class B sludge products can be used for gardens, pasture, and non-food crop farmland. Therefore, we set three Cd treatment levels in this study: 0mg kg<sup>-1</sup>(the control, C), 4mg kg<sup>-1</sup> (the low Cd treatment, L), and 20mg kg<sup>-1</sup>(the high Cd treatment, H). After two-month cultivation of the buds, we will treat these plants with CdCl<sub>2</sub> solution. After one-month treatment, the fresh leaves of each individual were collected for RNA extraction, and the cultivars were used for functional traits measurement.

### Phenotypic traits measurement

Phenotypic traits were measured for each clone. Plants with apparently abnormal values were checked immediately and the values corrected whenever necessary. To investigate the influence of Cd on photosynthesis, we measured net photosynthetic rate (A) in mornings between 8-11 am on sunny days, using a Li-6800 portable InfraRed Gas Analyser (IRGA). Chlorophyll content was extracted using the acetone method (Su et al., 2010) and measured using a spectrophotometer. Actual quantum efficiency of PSII (YII) shows the actual light energy conversion efficiency, and maximum quantum efficiency of PSII (Fv/Fm) shows the maximum efficiency of PSII which converts absorbed light energy into chemical energy. We first placed the plants in darkness for 30 minutes as a pretreatment, and then measured YII, Fv/Fm and electron transfer efficiency (ETR) in photosynthesis of healthy leaves using PAM-2500, DualPam-100. Ion leakage rate (ILR) as an indicator of cell membrane damage, was measured using conductivity meter (FiveEasy Plus, Mettler Toledo, Switzerland). Leaf cadmium after high Cd treatment was separated by boiling with concentrated nitric acid, and measured using an ICP-MS,Thermo Fisher iCAP RQ,America. Multiple comparisons based on Turkey's post-hoc HSD test was performed with ANOVA by the R package *agricolae* (Mendiburu, 2019) to test the differences among groups in physiological traits due to Cd treatments.

### RNAseq library preparation

The plant total RNA was extracted using the RNAPrep Pure Plant Kit (Tiangen, Beijing, China) according to the instructions provided by the manufacturer. RNA concentration and purity was measured using NanoDrop 2000(Thermo Fisher Scientific, Wilmington, DE). RNA integrity were assessed using the RNA

Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). A total amount of 1  $\mu$ g RNA per sample was used as input material for the preparations of RNA samples. Sequencing libraries were generated using Hieff NGS Ultima Dual-mode mRNA Library Prep Kit for Illumina (Yeasen Biotechnology (Shanghai) Co., Ltd.) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. First strand cDNA was synthesized, and second strand cDNA synthesis was subsequently performed. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptors with hairpin loop structure were ligated to prepare for hybridization. The library fragments were purified with the AMPure XP system (Beckman Coulter, Beverly, USA). Then 3  $\mu$ l USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37°C for 15 min followed by 5 min at 95°C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. The libraries were sequenced on an Illumina NovaSeq 6000 platform to generate 150 bp paired-end reads, according to the manufacturer's instructions.

### Read preprocessing and DEG analyses

Quality of the reads were checked using fastqc (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Adapters were clipped and low quality bases (phred score < 20) were removed from the reads in 4 bp sliding windows using Trimmomatic (Bolger, Lohse, & Usadel, 2014). Genome assembly and annotation of *P. australis* was obtained from Clean reads were mapped to the reference genome using STAR aligner (Dobin et al., 2013). After sorting using Samtools (Danecek et al., 2021), the mapped reads were processed using StringTie (Pertea et al., 2015) to obtain the gene and transcript count matrix. Principal component analysis (PCA) was conducted on differentially expressed genes among the treatment groups using DESeq2 following rlog transformation of gene counts, which represent the number of reads mapped to the reference genome (Love, Huber, & Anders, 2014).

DESeq2 provides statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting P values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted P-value < 0.01 & Fold Change[?]<sup>2</sup> found by DESeq2 were assigned as differentially expressed. Gene Ontology (GO) enrichment assessment analyses sets of genes in the context of GO categories to determine whether certain functional categories are overrepresented or enriched within the background whole genome gene sets. Here we tested enriched GO terms for the differentially expressed genes (DEGs) using goatools (Klopfenstein et al., 2018), using a criterion of P < 0.05 with Bonferroni correction. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was performed using KAAS (Moriya, Itoh, Okuda, Yoshizawa, & Kanehisa, 2007) and the enriched pathway of differentially expressed genes were analyzed using clusterProfiler (Wu et al., 2021). The Venn diagram was performed using ggvenn (Yan & Yan, 2021).

Coefficient of variation (CVa) for gene expression data was calculated by first normalizing the raw read counts using the variance stabilizing transformation (VST) method, which stabilizing variance across expression levels. Samples were then grouped by population and treatment. The CVa, defined as the ratio of the standard deviation to the mean expression level, was computed for each gene, with genes exhibiting zero variance excluded from further analysis. Finally, we visualized the distribution of CVa values using density plots, highlighting peaks and medians.

### Statistical analyses

For the hypothesis testing, the Wald test was used in DESeq2 to evaluate genes that are significantly differently expressed among the haplotypes and treatments with Bonferroni correction. Then, the differentially expressed genes were used to infer the associated gene ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms using Fisher's exact test with Bonferroni multiple correction. ANOVA was conducted to compare the physiological traits including net photosynthetic rate, chlorophyll content, actual

quantum efficiency of PSII, maximum quantum efficiency of PSII, electron transfer efficiency, ion leakage rate across haplotypes and treatments, followed by Turkey’s post-hoc HSD test for multiple corrections.

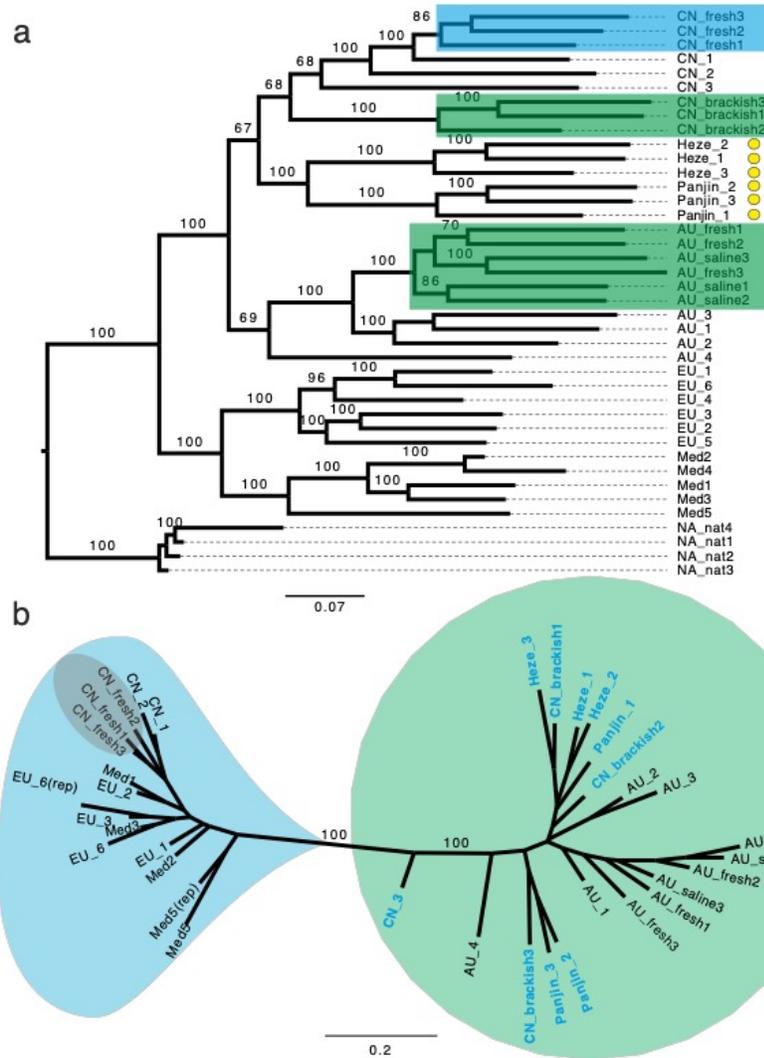
## Results

### Variant calling and mapping accuracies

A total of 40 samples were mapped to the reference genome, resulting in 1,505,327 SNPs. After filtering out SNPs with a minor allele frequency of less than 0.05 and more than 10% of missingness and removing the linkage disequilibrium within a 50kb window size and 10kb step size, we obtained 81,566 biallelic variant sites. These sites were used to construct a nuclear phylogenetic tree, and to perform PCA analysis for *P. australis* lineages. For the chloroplast phylogenetic trees, 211 sites were retained after filtering out those with a minor allele frequency of less than 0.05, and a missing rate higher than 50%. To assess the transcriptomic response to heavy metal exposure among the hybrid lineages, we prepared 18 RNAseq libraries, including control, low cadmium treatment, and high cadmium treatment groups. A total of 128.92 Gb of clean RNA-seq data was obtained from the 18 samples, with each sample producing 5.88Gb of clean data. The percentage of Q30 bases exceeded 94.31% across all samples. All reads from the 18 libraries were mapped to the reference genome with high accuracy, with unique mapping rates ranging from 85.94 to 93.96% (**Supplementary Table S2**). After filtering out genes with a count number of less than 1, a total of 40,089 genes were retained.

### Nuclear and chloroplast phylogeny of *P. australis* lineages

Given that the native North American population diverged early from other populations (C. Wang, Liu, Yin, Eller, Brix, Wang, Salojarvi, et al., 2021), we rooted the nuclear phylogeny on the USnat lineage. The lineages then diverged in the following order: Med, EU, AU and CN, with AU clustering closely with CN (**Fig. 1a**). In contrast, the unrooted chloroplast phylogenetic tree grouped the CN lineage with other lineages (**Fig. 1b**), while the AU lineage formed its own distinct clade. Discrepancies between the nuclear and chloroplast phylogenies were observed for some individuals, where they clustered with the CN lineage in the nuclear phylogenetic tree but grouped with the AU lineage in the chloroplast phylogenetic tree. These include the populations labeled as CN\_brackish, Heze, and Panjin. These mismatches suggest substantial gene flow between the two lineages, indicating ancient hybridization events. In contrast, the Australian populations, including both brackish water and freshwater ecotypes, all clustered within the same genetic group, suggesting that the octoploid AU lineage possesses genes conferring salinity tolerance (**Fig. 1a,b**).



## Genome scan reveals selective genes with high $F_{ST}$ values between populations

Consistent with expectations, principal component analysis (PCA) based on the nuclear gene dataset revealed that the tetraploid USnat lineage is genetically distinct from the other lineages and accounts for the largest proportion of variations (**Fig. 2a**). The octoploid AU lineage is separated from the rest of the tetraploid lineages (**Supplementary Figure S1**). We selected individuals representing pure CN and AU lineages and estimated  $F_{ST}$  between these populations. Furthermore, we conducted  $F_{ST}$  estimation between the CN brackishwater and freshwater populations. To identify regions potentially under selection, we applied a threshold corresponding to the top 1% of  $F_{ST}$  values across the genome. The analysis identified a total of 811 regions with high  $F_{ST}$  values between the AU and CN lineages, predominantly located on chromosome 1 (79 windows) and chromosome 18 (74 windows) (**Fig. 2b**). Among these, 111 windows overlapped with high  $F_{ST}$  regions identified between the CN Brackish water and freshwater populations (**Fig. 2d**). Most of these windows were found on chromosome 18, 10 and 1, suggesting strong adaptive introgression from the AU lineage into the CN lineage (**Fig. 2c**). Within these overlapping regions, we identified 46 genes and regard them as candidate selective genes. Of these, nine genes were enriched in Gene Ontology (GO) terms related to responses to salicylic acid (**Supplementary Table S3**). These genes were proved to involve in response to environmental cues, gene expression regulation, plant defense, and see longevity (**Supplementary Table**

S4 ). Interestingly, five of these nine genes are tandem duplicates.

#### Hosted file

image2.emf available at <https://authorea.com/users/827986/articles/1222430-introgression-and-chloroplast-capture-from-octoploids-confer-stress-tolerance-to-allotetraploid-common-reed-phragmites-australis>

#### Diverse gene regulatory strategies in response to heavy metal stress

To investigate whether the selective genes function stably in response to abiotic stress across different hybrid populations, we treated two genotypes with a series of cadmium and conducted a comprehensive transcriptomic analysis. PCA of the RNAseq dataset showed the first principal component (PC1) accounted for 72% of the variance, while the second principal component (PC2) accounted for 20% of the variations. The transcriptomes of the Heze and Panjin genotypes were clearly separated along PC1. Minimal variation was observed among the Panjin individuals, regardless of Cd treatments, suggesting this genotype is less sensitive to cadmium exposure. In contrast, PC2 separated the gene expression profiles of the Heze control group from those of the treatment groups. One sample treated with a high cadmium concentration grouped with the control genotypes. Although most samples had more than 92% of reads mapped to the reference genome, this particular sample showed a lower mapping rate of 86%, indicating that the individual may not have been in optimal condition during sampling (**Fig. 3a**). As a result, we excluded this genotype from subsequent analyses. In the Panjin genotype, compared to the Heze genotype, 2,078 genes were upregulated and 1,492 genes were downregulated (see **Supplementary Fig. S2**). Among these, 14 selective genes were found to be upregulated, while nine selective genes were downregulated (**Supplementary Table S5**).

The upregulated genes are predominantly involved in biological processes such as responses to external biotic stimuli, defense against bacteria and oomycetes, cell surface receptor signaling pathways, responses to salicylic acid, intracellular water homeostasis, and phosphorylation. Regarding cellular components, these genes mainly impact the plasma membrane and contribute to protein serine/threonine kinase activity, ADP binding, phosphotransferase activity involving alcohol group acceptors, cyclase activity, and transmembrane signaling receptor activity. In contrast, the downregulated genes did not show any significantly enriched Gene Ontology (GO) terms.

The coefficient of variation (CVa) analysis is essential for assessing the variability in gene expression across different treatment conditions relative to the mean expression levels. The Heze genotype under treatment conditions exhibited the highest median CVa among all groups, with a value of 1.53 (**Fig. 3b**; **Supplementary Table S6**). In contrast, the control groups of the Heze and Panjin genotypes displayed similar CVa values, with 0.983 and 0.920, respectively. The adaptive genes that were upregulated in the Panjin C vs Heze C comparison, showed the highest CVa values in Panjin HL treatment group. Conversely, downregulated genes in the Panjin C vs Heze C comparison exhibited the highest CVa values in the Heze HL cadmium groups (**Supplementary Figure S3, S4**). Nonetheless, all the selective genes showed higher CVa values than the median CVa of all genes (**Supplementary Table S7**). This suggests that Heze genotype showed higher response to the cadmium stress, and the adaptive genes have played a special role with independent regulatory pathways in each genotype.

#### Hosted file

image3.emf available at <https://authorea.com/users/827986/articles/1222430-introgression-and-chloroplast-capture-from-octoploids-confer-stress-tolerance-to-allotetraploid-common-reed-phragmites-australis>

javascript:void(0) In the Heze genotype, 1,075 genes exhibited differential expression when comparing low cadmium-treated samples to the control group (**Heze L vs Heze C**), with the majority of these genes being upregulated (956 genes). When exposed to a high concentration of cadmium, the Heze genotype showed 1,841 DEGs, with 1,371 genes upregulated and 470 genes downregulated relative to the control (**Heze H vs Heze C**), reflecting a more robust response with increased stress (**Fig. 4a**). In total, 712 genes

were upregulated in both low and high cadmium treatment groups (**Heze L vs Heze C**), constituting 74.5% of the upregulated genes in the low cadmium treatment (**Fig. 4b**). Conversely, only 28 genes were downregulated in both treatment groups (**Fig. 4b**). There are no genes that are simultaneously upregulated and downregulated across any treatment combinations.

In the Panjin genotype, only 170 DEGs were identified between low cadmium-treated samples and the control group (**Panjin L vs Panjin C**), with a higher number of downregulated genes (101 genes;**Fig. 4a**) compared to the upregulated ones. In contrast, samples treated with high cadmium levels (**Panjin H vs Panjin C**) showed 300 DEGs, of which more than half were downregulated (182 genes).

### Hosted file

image4.emf available at <https://authorea.com/users/827986/articles/1222430-introgression-and-chloroplast-capture-from-octoploids-confer-stress-tolerance-to-allotetraploid-common-reed-phragmites-australis>

### GO enrichment

#### *Biological process*

In the Heze genotype, the upregulated DEGs in the **Heze L vs Heze C** group were mostly involved in biological processes such as intracellular water homeostasis, response to salicylic acid, cell surface receptor signaling pathway, cell volume homeostasis, biological process involved in interspecies interaction between organisms and response to external biotic stimulus (**Fig. 4c**). For the **Heze H vs Heze C** group, upregulated DEGs were mainly associated with cell surface receptor signaling pathways and intracellular water homeostasis. For the Panjin genotypes, upregulated DEGs in the **Panjin L vs Panjin C** comparison were enriched in several GO terms related to responses to various light conditions (including light intensity, red/far-red light, UV-A, UV-B, and blue light), as well as developmental process, chlorophyll biosynthesis and metabolic processes (**Supplementary Table S8**). In contrast, downregulated DEGs in the **Panjin H vs Panjin C** comparison were associated with response to endogenous stimulus, organic substances, and hormones.

#### *Cell components*

In the Heze genotype, upregulated DEGs in the **Heze H vs Heze C** comparison were enriched in functions related to the photosystem II (PSII) antenna complex, PSII associated light-harvesting, complex II, light-harvesting complex and thylakoid light-harvesting complex. Conversely, the downregulated DEGs in the **Heze H vs Heze C** comparison were predominantly associated with the extracellular region. For the Panjin genotype, upregulated DEGs in the **Panjin L vs Panjin C** comparison were enriched in functions related to the plastid membrane, bounding membrane of organelle, chloroplast thylakoid membrane, plastid thylakoid membrane, thylakoid membrane, and photosynthetic membrane.

#### *Molecular functions*

For the Heze genotype, upregulated DEGs in **Heze L vs Heze C** treatment were enriched in ADP binding and transmembrane signaling receptor activity. Conversely, downregulated DEGs in the same treatment were enriched in polysaccharide binding functions. In the **Heze H vs Heze C** treatment group, upregulated DEGs were associated with calcium ion binding, and hydrolase activity, hydrolyzing O-glycosyl compounds. For the Panjin genotype, upregulated DEGs in the **Panjin L vs Panjin C** comparison were enriched in flavonol synthase activity and naringenin 3-dioxygenase activity.

#### javascript:void(0)KEGG pathway

For the Heze genotype, DEGs in the **Heze L vs Heze C** comparison were associated with nucleotide excision repair and starch and sucrose metabolism pathways. In the **Heze H vs Heze C** comparison, both upregulated and downregulated DEGs were involved in several pathways, including phenylpropanoid biosynthesis, starch and sucrose metabolism, cysteine and methionine metabolism, MAPK signaling pathway-plant, photosynthesis, carotenoid biosynthesis, plant-pathogen interaction, and amino sugar and nucleotide

sugar metabolism pathways (**Fig. 4d**). For the Panjin genotype, DEGs in **Panjin L vs Panjin C** comparison were enriched in plant hormone signal transduction pathway. In the **Panjin H vs Panjin C comparison**, DEGs were associated with starch and sucrose metabolism, glycosphingolipid biosynthesis, and ATP-dependent chromatin remodeling pathways.

#### javascript:void(0)Functional traits measurements

Four functional traits showed significant differences among cadmium treatments, including A, Chl, Fv/Fm and ILR, regardless of lineage effect (**Fig. 5 ; Table 1**). Among all the six traits, Fv/Fm was significantly differed between the two lineages ( $P < 0.01$ ), and affected by the interaction of cadmium and lineage (**Fig. 5 ; Table 1**). For both genotypes, ILR increased rapidly with the increasing of cadmium and reached an exceptional high level at the treatment with high cadmium concentration (**Fig. 5**). Traits such as A, Chl and Fv/Fm showed significant decreases as the concentration of cadmium raised in Heze genotype but not in Panjin genotype. There were not any significant differences between two genotypes in leaf Cd content under high Cd treatment (Heze,  $1.165 \pm 0.849$  mg/kg; Panjin,  $1.875 \pm 1.56$  mg/kg; t-test,  $t = -0.687, p = 0.525$ ).

**Table1 Functional traits of the Heze and Panjin population**

Population	Name	Treatment	Y(II)	ETR	Fv/Fm
Heze	Heze_1	Control	0,155	42,2	0,804
Heze	Heze_2	Control	0,134	42,6	0,803
Heze	Heze_3	Control	0,148	40,2	0,811
Heze	Heze_L_1	L	0,147	38,2	0,806
Heze	Heze_L_2	L	0,152	41,6	0,81
Heze	Heze_L_3	L	0,138	40,3	0,796
Heze	Heze_H_1	H	0,135	38,7	0,773
Heze	Heze_H_2	H	0,139	39,6	0,752
Heze	Heze_H_3	H	0,147	38,8	0,762
Panjin	Panjin_1	Control	0,162	40,4	0,807
Panjin	Panjin_2	Control	0,146	41,6	0,805
Panjin	Panjin_3	Control	0,143	41,2	0,811
Panjin	Panjin_L_1	L	0,131	40,5	0,804
Panjin	Panjin_L_2	L	0,158	38,5	0,807
Panjin	Panjin_L_3	L	0,162	39,9	0,803
Panjin	Panjin_H_1	H	0,133	39,6	0,788
Panjin	Panjin_H_2	H	0,142	39,7	0,786
Panjin	Panjin_H_3	H	0,145	38,4	0,796

#### Hosted file

image5.emf available at <https://authorea.com/users/827986/articles/1222430-introgression-and-chloroplast-capture-from-octoploids-confer-stress-tolerance-to-allotetraploid-common-reed-phragmites-australis>

#### Discussion

##### Adaptive introgression conferred salt tolerance to allotetraploids

The phylogenetic pattern revealed in this study aligns with our previous population genetic study based on RADseq, revealing that USnat is more distant from the rest of the lineages, with AU and CN formed one clade, and EU and Med in another clade (C. Wang, Liu, Yin, Eller, Brix, Wang, Salojärvi, et al., 2021). Thirteen individuals representing all lineages from the RADseq study were reanalyzed here using RNAseq data (including AU1, AU2, AU3, Med1, Med2, Med3, Med5, EU2, EU3, EU4, CN1, CN2), confirming the

AU lineage is octoploid, and the other lineages are tetraploids. Contrasting phylogenetic patterns between nuclear and chloroplast genes were observed between the AU and CN lineages, suggesting extensive ancient hybridization has happened between the two lineages, therefore the individuals with discordant phylogenetic trees can be visualized as ancient hybrids. This is plausible because AU lineages are primarily distributed in regions further south than southern China, extending to Australia, while CN lineages are found along the Yellow River and further north (C. Wang, Liu, Yin, Eller, Brix, Wang, Salojärvi, et al., 2021). As a result, much of eastern and central China lies within a secondary contact zone. In the octoploid AU lineage, both freshwater and salt-tolerant populations clustered together with little genetic differentiation, which is indication that they have strong tolerance to the saline levels in the ambient environment. In contrast, the brackish water and freshwater populations of CN lineage were distinctly separated, with freshwater population embedded with the pure tetraploids and brackish water individuals placed between tetraploids and octoploids. The CN brackish water population retains chloroplast DNA from the AU lineage, suggesting this lineage resulted from hybridization between AU and CN, and chloroplast capture happened. This suggests that adaptive loci associated with salt tolerance may have been conferred from the octoploid AU lineage to tetraploid CN lineage through introgression, leaving the hybrid brackish water population more resistant to high salinity.

This finding aligns with conclusions drawn from a previous gene expression study on untreated common reed samples, wherein individuals from lineages with higher ploidy levels were more likely to express genes enriched in stress tolerance (C. Wang, Wang, et al., 2021). The enhanced tolerance to abiotic stress in polyploids is a phenomenon observed in many plant species. Plants with higher ploidy levels often exhibit optimized morphological and anatomical structures in their leaves, reducing water loss and maintaining cellular homeostasis during salt or drought stress, as evidenced in common reed, *Arabidopsis* and birches (Sheng et al., 2024; Van de Peer, Ashman, Soltis, & Soltis, 2021). In rice, diploids tend to synthesize higher levels of glucose, fructose, and chlorophyll compared with autotetraploid under Cd stress (Ghouri et al., 2023).

A genome scan between the AU and CN population, as well as between the CN brackish and freshwater population, revealed many outliers with high  $F_{ST}$  values. Since these differentiated loci could be indicative of either speciation islands or adaptive loci, we focused on regions only where  $F_{ST}$  values overlapped between lineages and ecotypes to identify loci that likely represent adaptive traits. Nine of these genes are enriched in the biological process of responding to salicylic acid, which is known to enhance plant tolerance to various abiotic stresses, including salinity, temperature extremes and heavy metals (Guozhang, Guchou, & Zhengxun, 2004). Among these, the gene LOC133925052 codes for a S40-7-like protein playing a role in response to environmental cues and leaf senescence in rice; the gene LOC133925058 is known to positively mediate salt stress tolerance by binding to an E-box like motif to regulate gene expression; the protein OEL21155 is a receptor-like kinases (RLKs) in plant defense; the protein XP\_062206816.1 can positively regulate seed longevity in *Arabidopsis* (**Supplementary Table S4**). These are all relevant to the actual situation of the ecotypes involved in the study and the CVa of these genes in response to Cd are higher than neutral genes, indicating they are indeed responsive to abiotic stress and that our screening method is effective.

### Differed response to cadmium-induced stress in ancient hybrid populations

The principal component analysis (PCA) plot of the gene expression variations revealed that a substantial proportion of gene expression variation on PC1 separates the genotypes. This suggests that genetic background is still the most crucial factor in determining the genetic response to heavy metal stress in common reed. Both genotypes have genetic background as ancient hybrids. Hybridization can lead to species divergence and the formation of stable phylogenetic lineages, as illustrated by the phylogenetic trees (Gardner et al., 2023). However, individuals within the same genetic lineage exhibited markedly different responses to cadmium-induced stress, indicating potential genetic incompatibilities. However, similar fluctuations in gene expression have also been observed in non-hybrid *Brassica rapa* populations sampled over different time periods, suggesting that plants may undergo rapid evolutionary changes in gene expression in response to climate change (Hamann et al., 2021). Gene expression patterns were observed to change along with environmental

factors, such as elevations (Ye et al., 2023), latitude (Chen et al., 2020) and light (Mural, 1991), showing their flexibilities with local adaptation. In this study, differences in gene expression and photosynthetic capacity between the two populations may be attributed to local adaptations to their respective geographic origins. The Panjin population in Liaoning Province is geographically further north than the Heze population in Shandong Province.

Indeed, a comparison between control samples (**Panjin control vs Heze control**) showed that a large number of upregulated genes in Panjin are enriched in biological processes such as response to external biotic stimulus, defense response to bacteria and oomycetes, cell surface receptor signaling, response to salicylic acid, intracellular water homeostasis, and phosphorylation. This indicates that the Panjin genotype may be preadapted and obtained enhanced defense responses to biotic stimuli, potentially signifying a higher level of resistance to certain pathogens or environmental stresses. The involvement of signaling pathways and phosphorylation suggests complex regulatory mechanisms underpinning these responses. Physiologically, Panjin demonstrated significantly higher maximum quantum efficiency of photosystem II (Fv/Fm) in physiological reactions, signifying better performance under the same conditions. Under high Cd levels, photosynthesis parameters such as A and chlorophyll (Chl) showed a significant decrease in Heze but not in Panjin, suggesting that Heze is more adversely affected by cadmium stress than Panjin. Since cadmium content in leaf tissues did not significantly differ between the two genotypes, and considering the transcriptomic data were obtained from leaves, we can conclude that the stronger response in Heze is caused by the cadmium treatment itself rather than the heavy metal transportation and accumulation process.

That elucidates why, upon exposure to varying cadmium levels, Panjin genotype exhibited only a few DEGs between the treatment and control groups. In contrast, Heze genotype showed a high number of upregulated DEGs, indicating its greater sensitivity and lower resistance to Cd stress compared to Panjin. Similarly, more pathways related to structural defense, energy production, and stress signaling and number of DEGs related to pathways responsible for starch and sucrose metabolism were detected in Heze than in Panjin, suggesting the Heze genotype initiated comprehensive defense against the detrimental effects caused by cadmium. Panjin showed more pathways that prioritize maintaining cell integrity and regulating gene expression to adapt to stress, suggesting it has a more efficient or resilient strategy. The cellular responses are also different between the two populations. In Heze genotype, DEGs were predominantly enriched in biological processes such as intracellular water homeostasis, cell volume homeostasis, and cell surface receptor signaling pathway, regardless of exposure to low or high Cd levels. The increase in ion leakage rate (ILR) with rising Cd levels indicates cell membrane impairment under stress, a phenomenon supported by observations of undeveloped ultrastructure in vacuoles of common reed under electron microscopy (Hakmaoui, Ater, Boka, & Baron, 2007). The plant may endeavor to protect itself against the cell membrane damage through the activation of cell wall remodeling (Loix et al., 2017) to maintain the cellular water osmotic pressure.

Photosynthesis is a conserved life activity from Cyanobacteria to green plants. The deleterious cadmium was found to affect photosynthesis in many plant species by ways of decreasing the chlorophyll and carbon fixation, resulting in disruptions in development and growth (Haider et al., 2021). In common reed, exposure to cadmium results in significant damage to four functional traits associated with the photosynthetic apparatus in Heze genotype and the cell membrane in both haplotypes. Similarly, in lettuce, the treatment with high levels of Cd demonstrates a substantial decline in chlorophyll contents, net photosynthetic rate (A) and maximum quantum efficiency of PSII (Fv/Fm) when compared to control plants (Dias et al., 2013).

When treated with low levels of Cd content, both populations showed upregulated DEGs associated with organelle membrane and thylakoid membrane. It is noteworthy, however, that the number of DEGs is relatively small in Panjin genotype. The photosynthetic apparatus in plants is particularly susceptible to the adverse effects of cadmium stress. The photosystems, consist of light-harvesting complexes and reaction-center complexes interconnected by electron transport chains, involve numerous protein complexes and pigmentations. Consequently, any disruption in these essential components due to Cd stress will result in reduced efficiency photosynthesis. A high level of Cd stress will lead to “a disturbed shape, wavy appearance of grana and stroma thylakoids and swollen intrathylakoidal space” for the chloroplast ultrastructure in

common reed (Hakmaoui et al., 2007; Parmar, Kumari, & Sharma, 2013).

The DEGs in the comparison of **Heze H vs Heze C** treatments are enriched in the phenylpropanoid biosynthesis pathway, essential for lignin biosynthesis in secondary cell wall modification and the clearance of harmful reactive oxygen species. The metabolic products may also sequester cadmium to the extracellular region, impeding its mobility and enhancing plant resistance. In addition, these DEGs may promote the production of flavonoids, which act as plant hormones (Ge, Xin, & Tian, 2023). Unlike the passive protection mechanisms observed in Heze genotype, Panjin genotype upregulates genes to capture a broader spectrum of light types and intensities, coupled with chlorophyll biosynthesis, to mitigate the negative impact of the low cadmium level.

A large number of genes were found to be involved in species interaction and plant-pathogen interactions in Heze genotype, suggesting that the plant's response to cadmium stress may share pathways with pathogen defense particularly against oomycetes. The accumulation of cadmium can be intertwined with pathogen defense and potentially trigger Systemic acquired resistance (SAR) via salicylic acid (SA) and jasmonic acid (JA) mediated pathway. This synergistic effect, observed extensively in Brassicaceae complies with the elemental defense hypothesis (Z. Liu et al., 2022). Moreover, Heze genotype upregulates genes associated with calcium ion binding. This strategic upregulation may involve the occupation of known ion channels to compete with Cd, thereby exerting an antagonistic effect on cadmium absorption (Huang et al., 2017).

## Conclusions

Through extensive phylogenetic analysis using genetic variants from transcriptomic data, combined with gene expression and functional traits measured in common garden experiments, we tested the hypothesis regarding differences in abiotic stress tolerance between octoploids and tetraploids, as well as the functional stability of hybrid populations in common reed. We conclude that octoploids are more tolerant to high-salinity environments than tetraploids. Evidence suggests that adaptive introgression has occurred, with adaptive genes being transferred from the octoploid lineage to the tetraploid lineage. However, the introgressed loci do not function stably in hybrid populations, likely due to local adaptations within the divergent hybrid populations or incompatible regulatory networks between the parental genomes.

## Data availability

The RNAseq data for the 18 samples were deposited in NCBI SRA database, with accession number PRJNA1066047.

## javascript:void(0) Acknowledgments

The authors thank Miss Dongzi Jiang for her contribution to the greenhouse experiment. This work is supported by the National Natural Science Foundation of China (No. 32100304; 32470388; U22A20558; 32271588) and the Natural Science Foundation of Shandong Province (ZR2021QC119). We thank CSC-IT Center for Science, Finland for computation resources.

## Author contributions:

LL. Liu, C. W and WH. Guo conceived the idea; C. W. and LL. Liu analyzed the transcriptomic and phenotypic data; LL. Liu, WY. Sheng, MQ.Yin, C. Lambertini and X. Guo maintained the plant and carried out the cadmium treatment experiments; C. W. led the writing of the paper with contribution from LL. Liu, WH. Guo, C. Lambertini, X. Guo, MQ. Yin and WY. Sheng.

## References

Baduel, P., Hunter, B., Yeola, S., & Bomblies, K. (2018). Genetic basis and evolution of rapid cycling in railway populations of tetraploid *Arabidopsis arenosa*. *PLoS genetics*, *14* (7), e1007510. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30* (15), 2114-2120. Bonanno, G., & Giudice, R. L. (2010). Heavy metal bioaccumulation by the

organs of *Phragmites australis* (common reed) and their potential use as contamination indicators. *Ecological indicators*, 10 (3), 639-645. Březinová, T. D., & Vymazal, J. (2022). Distribution of heavy metals in *Phragmites australis* growing in constructed treatment wetlands and comparison with natural unpolluted sites. *Ecological Engineering*, 175, 106505. Brysting, A. K., Fay, M. F., Leitch, I. J., & Aiken, S. G. (2004). One or more species in the arctic grass genus *Dupontia*?—a contribution to the Panarctic Flora project. *Taxon*, 53 (2), 365-382. Case, A. L., Finseth, F. R., Barr, C. M., & Fishman, L. (2016). Selfish evolution of cytonuclear hybrid incompatibility in *Mimulus*. *Proceedings of the Royal Society B: Biological Sciences*, 283 (1838), 20161493. Chen, L., Cai, Y., Qu, M., Wang, L., Sun, H., Jiang, B., . . . Wu, C. (2020). Soybean adaptation to high-latitude regions is associated with natural variations of GmFT2b, an ortholog of FLOWERING LOCUS T. *Plant, Cell & Environment*, 43 (4), 934-944. Cicero-Fernandez, D., Pena-Fernandez, M., Exposito-Camargo, J. A., & Antizar-Ladislao, B. (2016). Role of *Phragmites australis* (common reed) for heavy metals phytoremediation of estuarine sediments. *International journal of phytoremediation*, 18 (6), 575-582. Cordeiro, J. M. P., & Felix, L. P. (2018). Intra- and interspecific karyotypic variations of the genus *Senna* Mill. (Fabaceae, Caesalpinioideae). *Acta Botanica Brasílica*, 32 (1), 128-134. Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., . . . Davies, R. M. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10 (2), giab008. Del Pozo, J. C., & Ramirez-Parra, E. (2015). Whole genome duplications in plants: an overview from *Arabidopsis*. *Journal of Experimental Botany*, 66 (22), 6991-7003. Dias, M. C., Monteiro, C., Moutinho-Pereira, J., Correia, C., Goncalves, B., & Santos, C. (2013). Cadmium toxicity affects photosynthesis and plant growth at different levels. *Acta physiologiae plantarum*, 35, 1281-1289. Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., . . . Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29 (1), 15-21. Eller, F., Skalova, H., Caplan, J. S., Bhattarai, G. P., Burger, M. K., Cronin, J. T., . . . Kettenring, K. M. (2017). Cosmopolitan species as models for ecophysiological responses to global change: the common reed *Phragmites australis*. *Frontiers in Plant Science*, 8, 1833. Ellison, C., & Burton, R. (2010). Cytonuclear conflict in interpopulation hybrids: the role of RNA polymerase in mtDNA transcription and replication. *Journal of evolutionary biology*, 23 (3), 528-538. Figueiredo, M., Bruno, R., Barros e Silva, A., Nascimento, S., Oliveira, I., & Felix, L. (2014). Intraspecific and interspecific polyploidy of Brazilian species of the genus *Inga* (Leguminosae: Mimosoideae). *Genetics and Molecular Research*, 13 (2), 3395-3403. Gardner, E. M., Bruun-Lund, S., Niissalo, M., Chantarasuwan, B., Clement, W. L., Geri, C., . . . Khew, G. (2023). Echoes of ancient introgression punctuate stable genomic lineages in the evolution of figs. *Proceedings of the National Academy of Sciences*, 120 (28), e2222035120. Ge, W., Xin, J., & Tian, R. (2023). Phenylpropanoid pathway in plants and its role in response to heavy metal stress: a review. *Sheng wu Gong Cheng xue bao = Chinese Journal of Biotechnology*, 39 (2), 425-445. Ghouri, F., Shahid, M. J., Liu, J., Lai, M., Sun, L., Wu, J., . . . Shahid, M. Q. (2023). Polyploidy and zinc oxide nanoparticles alleviated Cd toxicity in rice by modulating oxidative stress and expression levels of sucrose and metal-transporter genes. *Journal of Hazardous Materials*, 448, 130991. Guan, B., Yu, J., Wu, M., Liu, X., Wang, X., Yang, J., . . . Zhang, X. (2023). Clonal integration promotes the growth of *Phragmites australis* populations in saline wetlands of the Yellow River Delta. *Frontiers in Plant Science*, 14, 1162923. Guozhang, K., Guchou, S., & Zhengxun, W. (2004). Salicylic acid and its environmental stress tolerance in plants. *Guangxi Zhiwu*, 24 (2), 178-183. Haider, F. U., Liqun, C., Coulter, J. A., Cheema, S. A., Wu, J., Zhang, R., . . . Farooq, M. (2021). Cadmium toxicity in plants: Impacts and remediation strategies. *Ecotoxicology and Environmental Safety*, 211, 111887. Hakmaoui, A., Ater, M., Boka, K., & Baron, M. (2007). Copper and cadmium tolerance, uptake and effect on chloroplast ultrastructure. Studies on *Salix purpurea* and *Phragmites australis*. *Zeitschrift für Naturforschung C*, 62 (5-6), 417-426. Hamann, E., Pauli, C. S., Joly-Lopez, Z., Groen, S. C., Rest, J. S., Kane, N. C., . . . Franks, S. J. (2021). Rapid evolutionary changes in gene expression in response to climate fluctuations. *Molecular Ecology*, 30 (1), 193-206. Huang, D., Gong, X., Liu, Y., Zeng, G., Lai, C., Bashir, H., . . . Cheng, M. (2017). Effects of calcium at toxic concentrations of cadmium in plants. *Planta*, 245, 863-873. Javed, M. T., Tanwir, K., Akram, M. S., Shahid, M., Niazi, N. K., & Lindberg, S. (2019). Phytoremediation of cadmium-polluted water/sediment by aquatic macrophytes: role of plant-induced pH changes. In *Cadmium toxicity and tolerance in plants* (pp. 495-529): Elsevier. Kahle, H. (1993). Response of roots of trees to heavy metals. *Environmental and experimental botany*, 33 (1), 99-119. Klopfenstein, D., Zhang, L., Pedersen, B. S., Ramirez, F., Warwick Vesztrocy, A.,

Naldi, A., . . . Weigel, M. (2018). GOATOOLS: A Python library for Gene Ontology analyses. *Scientific reports*, *8* (1), 10872. Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. *BMC bioinformatics*, *15*, 1-13. Lambertini, C., Guo, W.-Y., Ye, S., Eller, F., Guo, X., Li, X.-Z., . . . Brix, H. (2020). Phylogenetic diversity shapes salt tolerance in *Phragmites australis* estuarine populations in East China. *Scientific reports*, *10* (1), 17645. Lambertini, C., Gustafsson, M., Frydenberg, J., Lissner, J., Speranza, M., & Brix, H. (2006). A phylogeographic study of the cosmopolitan genus *Phragmites* (Poaceae) based on AFLPs. *Plant Systematics and Evolution*, *258*, 161-182. Liu, L. L., Yin, M. Q., Guo, X., Wang, J. W., Cai, Y. F., Wang, C., . . . Eller, F. (2022). Cryptic lineages and potential introgression in a mixed-ploidy species (*Phragmites australis*) across temperate China. *Journal of Systematics and Evolution*, *60* (2), 398-410. Liu, Z., Sun, Z., Zeng, C., Dong, X., Li, M., Liu, Z., & Yan, M. (2022). The elemental defense effect of cadmium on *Alternaria brassicicola* in *Brassica juncea*. *BMC Plant Biology*, *22* (1), 1-14. Loix, C., Huybrechts, M., Vangronsveld, J., Gielen, M., Keunen, E., & Cuypers, A. (2017). Reciprocal interactions between cadmium-induced cell wall responses and oxidative stress in plants. *Frontiers in Plant Science*, *8*, 1867. Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, *15* (12), 1-21. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, *17* (1), 10-12. Meyerson, L. A., Cronin, J. T., Bhattarai, G. P., Brix, H., Lambertini, C., Lučanová, M., . . . Pyšek, P. (2016). Do ploidy level and nuclear genome size and latitude of origin modify the expression of *Phragmites australis* traits and interactions with herbivores? *Biological Invasions*, *18*, 2531-2549. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic acids research*, *35* (suppl\_2), W182-W185. Mural, R. J. (1991). Fundamentals of light-regulated gene expression in plants. *Plant Genetic Engineering*, 191-211. Oono, Y., Yazawa, T., Kanamori, H., Sasaki, H., Mori, S., Handa, H., & Matsumoto, T. (2016). Genome-wide transcriptome analysis of cadmium stress in rice. *BioMed Research International*, *2016*. Parmar, P., Kumari, N., & Sharma, V. (2013). Structural and functional alterations in photosynthetic apparatus of plants under cadmium stress. *Botanical Studies*, *54* (1), 1-6. Perteua, M., Perteua, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature biotechnology*, *33* (3), 290-295. Pieri, A., Beleggia, R., Gioia, T., Tong, H., Di Vittori, V., Frascarelli, G., . . . Fiorani, F. (2024). Transcriptomic response to nitrogen availability reveals signatures of adaptive plasticity during tetraploid wheat domestication. *The Plant Cell*, koae202. Roux, F., Mary-Huard, T., Barillot, E., Wenes, E., Botran, L., Durand, S., . . . Budar, F. (2016). Cytonuclear interactions affect adaptive traits of the annual plant *Arabidopsis thaliana* in the field. *Proceedings of the National Academy of Sciences*, *113* (13), 3687-3692. Saltonstall, K. (2002). Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proceedings of the National Academy of Sciences*, *99* (4), 2445-2449. Saltonstall, K. (2016). The naming of *Phragmites* haplotypes. *Biological Invasions*, *18* (9), 2433-2441. Sheng, W., Liu, L., Wu, Y., Yin, M., Yu, Q., Guo, X., . . . Guo, W. (2024). Exploring salt tolerance and indicator traits across four temperate lineages of the common wetland plant, *Phragmites australis*. *Science of The Total Environment*, *912*, 169100. Sloan, D. B., Warren, J. M., Williams, A. M., Wu, Z., Abdel-Ghany, S. E., Chicco, A. J., & Havird, J. C. (2018). Cytonuclear integration and co-evolution. *Nature Reviews Genetics*, *19* (10), 635-648. Srivastava, J., Kalra, S. J., & Naraian, R. (2014). Environmental perspectives of *Phragmites australis* (Cav.) Trin. Ex. Steudel. *Applied Water Science*, *4* (3), 193-202. Sun, X.-S., Chen, Y.-H., Zhuo, N., Cui, Y., Luo, F.-L., & Zhang, M.-X. (2021). Effects of salinity and concomitant species on growth of *Phragmites australis* populations at different levels of genetic diversity. *Science of The Total Environment*, *780*, 146516. Tanaka, T. S., Irbis, C., & Inamura, T. (2017). Phylogenetic analyses of *Phragmites* spp. in southwest China identified two lineages and their hybrids. *Plant Systematics and Evolution*, *303* (6), 699-707. Te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubešová, M., & Pyšek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. *Annals of botany*, *109* (1), 19-45. Van de Peer, Y., Ashman, T.-L., Soltis, P. S., & Soltis, D. E. (2021). Polyploidy: an evolutionary and ecological force in stressful times. *The Plant Cell*, *33* (1), 11-26. Wagner, F., Ott, T., Zimmer, C., Reichhart, V., Vogt, R., & Oberprieler, C. (2019). 'At the crossroads towards polyploidy': genomic divergence and extent of homoploid hybridization are drivers

for the formation of the ox-eye daisy polyploid complex (*Leucanthemum*, Compositae-Anthemideae). *New Phytologist*, 223 (4), 2039-2053. Wang, C., Liu, L., Yin, M., Eller, F., Brix, H., Wang, T., . . . Guo, W. (2021). Genome-wide analysis illustrates the effect of polyploidization in *Phragmites australis*. *bioRxiv* . Wang, C., Liu, L., Yin, M., Eller, F., Brix, H., Wang, T., . . . Guo, W. (2021). Genome-wide analysis tracks the emergence of intraspecific polyploids in *Phragmites australis*. *bioRxiv* , 2021.2009. 2005.458733. Wang, C., Liu, L., Yin, M., Liu, B., Wu, Y., Eller, F., . . . Salojärvi, J. (2024). Chromosome-level genome assemblies reveal genome evolution of an invasive plant *Phragmites australis*. *Communications Biology*, 7 (1), 1007. doi:10.1038/s42003-024-06660-1Wang, C., Wang, T., Yin, M., Eller, F., Liu, L., Brix, H., & Guo, W. (2021). Transcriptome analysis of tetraploid and octoploid common reed (*Phragmites australis*). *Frontiers in Plant Science*, 12 , 653183. Wang, H., Zhang, W., Yu, Y., Fang, X., Zhang, T., Xu, L., . . . Xiao, H. (2024). Biased gene introgression and adaptation in the face of chloroplast capture in *Aquilegia amurensis*. *Systematic Biology* , syae039. Wang, J., Wei, H., & Pan, B. (2023). Accumulation characteristics and probabilistic risk assessment of Cd in agricultural soils across China. *Huan Jing ke Xue= Huanjing Kexue*, 44 (7), 4006-4016. Wright, D., & Welbourn, P. (1994). Cadmium in the aquatic environment: a review of ecological, physiological, and toxicological effects on biota. *Environmental Reviews*, 2 (2), 187-214. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., . . . Zhan, L. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The innovation*, 2 (3). Yan, L., & Yan, M. L. (2021). Package ‘ggvenn’. In: Google Scholar. Ye, L.-J., Möller, M., Luo, Y.-H., Zou, J.-Y., Zheng, W., Liu, J., . . . Gao, L.-M. (2023). Variation in gene expression along an elevation gradient of *Rhododendron sanguineum* var. *haemaleum* assessed in a comparative transcriptomic analysis. *Frontiers in Plant Science*, 14 , 1133065. Zhu, T., Li, L., Duan, Q., Liu, X., & Chen, M. (2021). Progress in our understanding of plant responses to the stress of heavy metal cadmium. *Plant Signaling & Behavior*, 16 (1), 1836884.



