Received: 5 August 2021 | Revised: 12 October 2021 | Accepted: 15 November 2021

DOI: 10.1111/mec.16293

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Novel genome characteristics contribute to the invasiveness of *Phragmites australis* (common reed)

Dong-Ha Oh[1] | Kurt P. Kowalski[2] | Quynh N. Quach[3] | Chathura Wijesinghege[1] | Philippa Tanford[3,4] | Maheshi Dassanayake[1] | Keith Clay[3,5]

[1]Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana, USA

[2]U.S. Geological Survey, Great Lakes Science Center, Ann Arbor, Michigan, USA

[3]Department of Ecology & Evolutionary Biology, Tulane University, New Orleans, Louisiana, USA

[4]Department of Biology, Washington University in St. Louis, St. Louis, Missouri, USA

[5]Department of Biology, Indiana University, Bloomington, Indiana, USA

Correspondence
Dong-Ha Oh and Maheshi Dassanayake, Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana, USA
Emails: ohdongha@gmail.com; maheshid@lsu.edu

Kurt P. Kowalski, U.S. Geological Survey, Great Lakes Science Center, Ann Arbor, Michigan, USA
Email: kkowalski@usgs.gov

Keith Clay, Department of Ecology & Evolutionary Biology, Tulane University, New Orleans, Louisiana, USA.
Email: clay@tulane.edu

Funding information
U.S. Geological Survey, Grant/Award Number: G18AC00373; National Science Foundation, Grant/Award Number: MCB-1616827 and NSF-IOS-EDGE-1923589

Abstract

The rapid invasion of the non-native *Phragmites australis* (Poaceae, subfamily Arundinoideae) is a major threat to native wetland ecosystems in North America and elsewhere. We describe the first reference genome for *P. australis* and compare invasive (ssp. *australis*) and native (ssp. *americanus*) genotypes collected from replicated populations across the Laurentian Great Lakes to deduce genomic bases driving its invasive success. Here, we report novel genomic features including a *Phragmites* lineage-specific whole genome duplication, followed by gene loss and preferential retention of genes associated with transcription factors and regulatory functions in the remaining duplicates. Comparative transcriptomic analyses revealed that genes associated with biotic stress and defence responses were expressed at a higher basal level in invasive genotypes, but native genotypes showed a stronger induction of defence responses when challenged by a fungal endophyte. The reference genome and transcriptomes, combined with previous ecological and environmental data, add to our understanding of mechanisms leading to invasiveness and support the development of novel, genomics-assisted management approaches for invasive *Phragmites*.

KEYWORDS
Arundinoideae, fungal inoculation, gene expression, invasive, *Phragmites australis*, reference genome, whole genome duplication

1 | INTRODUCTION

Invasion of native ecosystems by non-native species is a worldwide problem that damages both ecosystems and economies. Invasive plants can negatively affect agricultural production (Dogra et al., 2010; Pimentel et al., 2005) and displace native species through multiple mechanisms including enemy release (Keane & Crawley, 2002; Mitchell & Power, 2003), allelopathy (Chengxu et al., 2011; Kalisz et al., 2020), and novel traits (Divíšek et al., 2018; Kolar & Lodge, 2001). They are key drivers of global environmental change and

require substantial economic resources for management (Diagne et al., 2021). Given the biological impacts of invasive species, there is an urgent need for studies across diverse systems to evaluate the long-term consequences of invasive species and to better understand the underlying biological mechanisms of invasion. Dozens of hypotheses have been proposed to explain why some species become invasive (Catford et al., 2009) and why some habitats are vulnerable to invasion (Stohlgren et al., 2002), which can contribute to better management of invasive species and a better understanding of general mechanisms of invasiveness (Hierro et al., 2005). The genetic and molecular bases for evolution of invasive lineages have been of longstanding interest, but there is still limited information addressing this problem (Baker & Stebbins, 1965; Bock et al., 2015). Previous studies have examined the genetics of invasive and weedy species independent of direct comparisons with noninvasive genotypes or populations (Liu, Yan, et al., 2020; Peng et al., 2014), but better insights can be gained from comparative studies where invasive and noninvasive genotypes or populations of the same species are investigated in both the native and invasive ranges (Hodgins et al., 2013; Lavergne & Molofsky, 2007; Lockwood & Somero, 2011; Stern & Lee, 2020) or where invasive and noninvasive populations occur intermixed in the same habitats (Mounger et al., 2021; Sherman et al., 2016). Understanding the evolutionary bases of invasiveness will contribute to the development of more effective management and control strategies.

Phragmites australis (Cav.) Trin. ex Steud. (Common Reed, Poaceae) is globally distributed (Figure 1a) and provides multiple ecosystem services in its native range (Kiviat, 2013; Rooth & Stevenson, 2000; Whitaker et al., 2015). A native subspecies (*P. australis* ssp. *americanus*) has been present in North American wetlands for thousands of years (Saltonstall et al., 2004). However, the non-native, invasive subspecies (*P. australis* ssp. *australis*) (Martin & Blossey, 2013; Saltonstall, 2002) was introduced to North America from Europe prior to 1900 and has been aggressively disrupting and displacing native plant communities (Mozdzer et al., 2013; Saltonstall, 2002) and altering wildlife habitat and ecosystem properties (Perez et al., 2013; Rogalski & Skelly, 2012). The invasive subspecies occurs throughout the contiguous United States (U.S.) and the entire Laurentian Great Lakes basin (Bourgeau-Chavez et al., 2013; Saltonstall, 2002; Tulbure & Johnston, 2010) (Figure 1b) and is one of the most problematic invasive plant species in wetland habitats in eastern North America, with millions of dollars per year invested in control efforts (Kowalski et al., 2015; Meyerson et al., 2016). It co-occurs with the native subspecies in many areas (Figure 1c) but exhibits more robust growth (Figure 1d,e) with larger inflorescences, leaves, and height (Figure 1f,g). Both native and invasive subspecies reproduce by seed and clonally via rhizomes (Figure 1h). A variety of mechanisms promoting *P. australis* invasions have been proposed, including efficient resource utilization and lower construction costs (Caplan et al., 2014), genetic diversity and mode of reproduction (Kettenring et al., 2011; Kettenring & Mock, 2012), and escape from natural enemies relative to native *Phragmites* genotypes (Allen et al., 2015; Cronin et al., 2015; Lambert et al., 2007), but effective control strategies are lacking (Hazelton et al., 2014). The Great Lakes Restoration Initiative has identified invasive species (including *P. australis)* as one of its five most urgent issues since 2010 (Great Lake Restoration Initiative, 2019). Invasive *P. australis* has also been recognized as the leading plant model for studying genetic mechanisms underlying plant invasions (Cesarino et al., 2020; Meyerson et al., 2016). *P. australis* therefore provides an excellent system to test genetic adaptations and control measures in plant invasions, given that both native and invasive populations coexist over a large geographic range (Figure 1a, b).

Phragmites australis exhibits a range of ploidy levels from 2 to 12× and higher, with tetraploids reported as dominant in temperate Europe and North America and octoploids dominant in Asia (Clevering & Lissner, 1999; Saltonstall, 2003; te Beest et al., 2012). *P. australis* genotypes, ploidy levels, and genome size have been assessed for traits that may favour invasiveness such as photosynthetic rate and nitrogen use efficiency, rhizome size, shoot emergence rate, and herbivory resistance (Guo et al., 2014; Park & Blossey, 2008; Pyšek et al., 2020). However, *P. australis* has not been investigated from a genomic perspective and lacks a reference genome that can serve as a foundational resource to investigate genomic traits underlying plant invasions and to identify genetic targets for biocontrol. Considering the diverse range of ploidy levels (Keller, 2000; Liu, Yin, et al., 2020; Pyšek et al., 2020) and that the genetic mechanisms and natural selection underlying diploidization of polyploids are largely unknown for angiosperms (Li et al., 2021), *P. australis* provides an ideal subject to study such evolutionary mechanisms. Further, *Phragmites* belongs to the grass subfamily Arundinoideae, which has been poorly explored at the genomic level compared to other grass subfamilies, even though multiple species are ecologically dominant or invasive on a global scale (Grass Phylogeny Working Group, 2001).

We report here the first reference genome for *P. australis* using a representative genotype from the invasive subspecies *P. australis* ssp. *australis*, as well as comparative transcriptomic analyses of invasive and native genotypes coexisting in the Great Lakes region of North America. Our results identify variation in gene expression correlated with invasiveness and provide a key genomic resource for grasses and the subfamily Arundinoideae, novel insights into evolution in an underexplored grass clade, and a genomic foundation for development of new management approaches.

2 | MATERIALS AND METHODS

2.1 | Plant materials

For genome sequencing, tillers and associated rhizome tissues were collected from a single *P. australis* clump of chloroplast haplotype M (Martin & Blossey, 2013; Saltonstall, 2002) at the Ottawa National Wildlife Refuge near Toledo, Ohio (Figure 1c, marked with "G") and propagated in a walk-in growth chamber as detailed in Supporting Information Methods. For transcriptome analyses, we collected three additional invasive and three native genotypes from four sites around the Great Lakes in Michigan and Ohio, U.S. (Figure 1c). Native

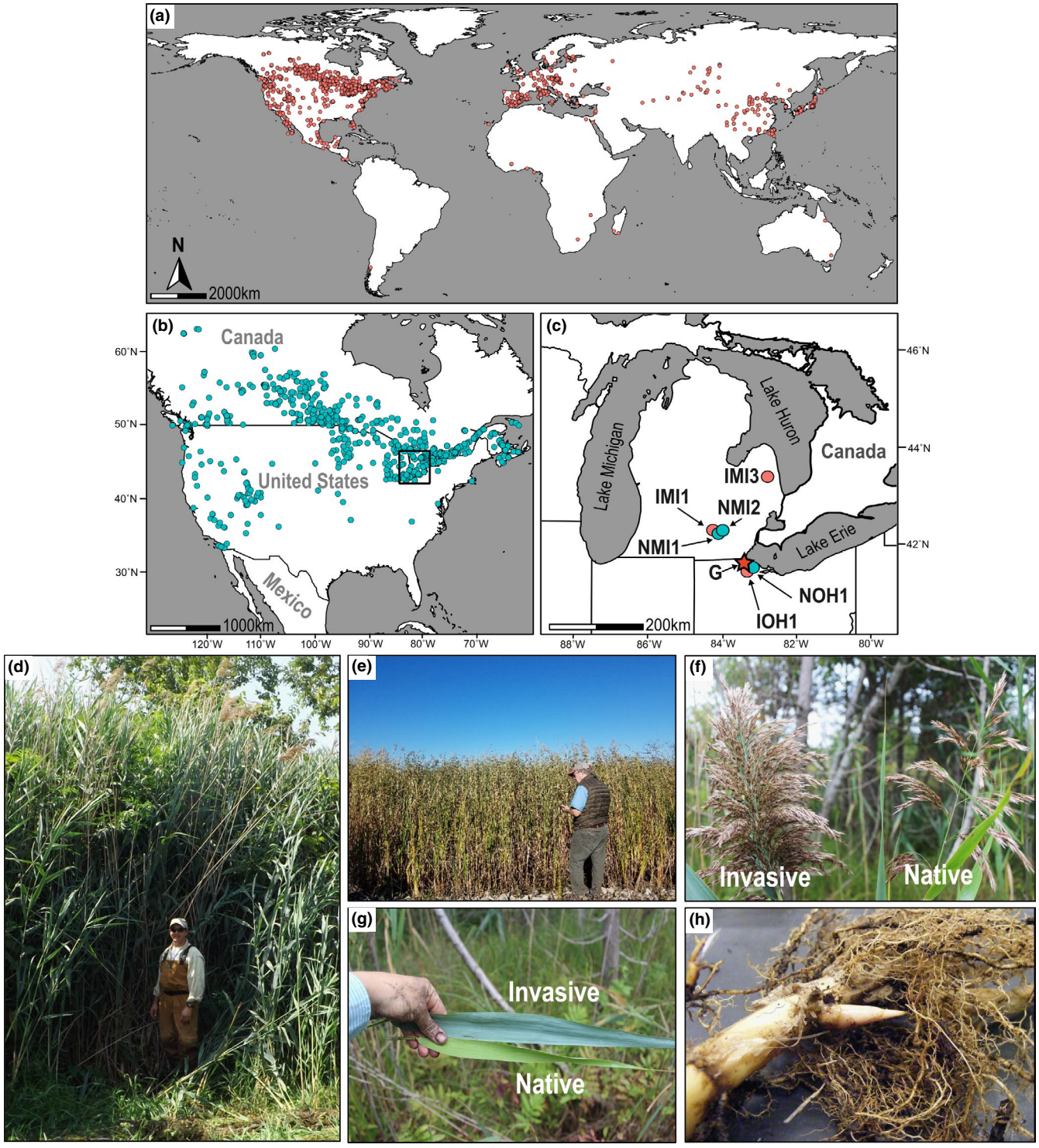


FIGURE 1 The invasive and native subspecies of common reed (*Phragmites australis*). (a) Reported global distribution of *Phragmites australis* ssp. *australis* ("Invasive") and (b) Reported distribution of *P. australis* ssp. *americanus* ("Native") in the U.S.A. and Canada based on the Global Biodiversity Information Facility (GBIF.org, 2020a, 2020b). (c) A magnified view of the box in panel b showing sample site locations in Michigan (MI) and Ohio (OH) in the U.S.A. Invasive (I) and Native (N) *Phragmites* plants were sampled from three Michigan sites (MI1, MI2, and MI3) and one Ohio site (OH1) in the Great Lakes region. Clonal fragments for the reference genome (G) were collected near the OH1 site. See Figure 5a for more details. (d) Invasive *P. australis* ssp. *australis* stand growing in a western Lake Erie coastal wetland. (e) Native *P. australis* ssp. *americanus* stand located in a western Lake Erie wetland. (f) Seed heads from the invasive and native *P. australis* growing in Michigan. (g) Invasive and native *P. australis* leaves collected in Michigan. (h) Rhizome and dense fibrous roots from an invasive *P. australis* plant growing in Michigan

genotypes were readily distinguished from invasive genotypes by their smaller stature and thinner tillers with distinctive reddish coloration at the nodes (Figure 1e, f, g). Specimens of both subspecies are in many Midwestern herbaria (https://midwestherbaria.org/portal/), and voucher specimens generated from the genomic reference have been deposited in the Indiana University, Louisiana State University, and University of Michigan herbaria. Both native and invasive genotypes were confirmed by chloroplast haplotype sequences as previously described (Saltonstall, 2002, 2016). *Phragmites* plants were grown in the growth chamber and subjected to endophyte inoculation treatments before RNA isolation for RNA-seq analyses, as detailed in Supporting Information Methods.

2.2 | Genome sequencing, assembly, and annotation

For the genome sequencing, genomic DNA was isolated from leaf tissue of the individual sample collected ("G" in Figure 1c) using a 2% CTAB extraction protocol (Doyle & Dickson, 1987), converted to SMRTbell libraries (Pacific Bioscience) with 20-Kb target insert size, and sequenced for continuous long reads (CLR) using a PacBio RSII single-molecule real-time sequencing platform. After post-processing (SMRT Analysis v2.3, Pacific Bioscience), 4.79 million reads (mean length 8.80 Kbp, N50 13.11 Kbp, and total 42.19 Gbp) were assembled into contigs using Canu assembler (v. 1.4) (Koren et al., 2017), with the target genome size set to 1 Gbp, generating the reference genome assembly (Supporting Information Methods). In addition, the same genomic DNA sample was converted to TruSeq DNA paired-ends libraries (Illumina) for whole-genome shotgun sequencing on a NextSeq platform (Illumina) and used to test for functional diploidy based on sequence diversity sampled (Supporting Information Methods and Figure S1).

For genome annotation, we first used RepeatModeler v. 1.0.8 and RepeatMasker v. 4.0.9 (http://www.repeatmasker.org/) to detect transposable elements and repetitive sequences in the assembled *P. australis* contigs. De novo-detected repetitive sequences were combined with known monocot repeat sequences in RepBase (v. 23.07; http://www.girinst.org/) to mask repeats in *P. australis* contigs. Protein-coding gene models on the repeat-masked genome were predicted using the MAKER (v. 2.31.10) (Campbell et al., 2014) with *P. australis* benchmarking universal single-copy orthologues (BUSCO)-trained parameters for ab initio gene model prediction (Waterhouse et al., 2018) and hints from *P. australis* transcriptomes de novo assembled using Trinity (v. 2.1.1) (Haas et al., 2013) as well as from homologues from *Sorghum* (Phytozome ID:454) and *Brachypodium* (Phytozome ID:314). In addition, we performed a reference-guided transcriptome assembly using StringTie (v. 2.0.1) (Pertea et al., 2015) to report putative isoform models associated with gene models predicted in the reference genome. Transcriptomes used to assist the protein-coding gene model prediction of the reference genome were assembled based on paired-end RNA-seq reads derived from the leaf, shoot, and rhizome tissues of the same sample used for the reference genome assembly. The gene model encoding the longest open reading frame (ORF) was selected as the representative for each of the observed protein-coding gene loci.

2.3 | Comparative and functional analyses

We used BUSCO (v. 3) (Waterhouse et al., 2018) to assess the completeness of the representative *P. australis* protein-coding gene models in comparison with grass genomes available in Monocots PLAZA (v. 4.5) (Van Bel et al., 2018) (Table S1). Maximum-likelihood species trees were created using OrthoFinder (v. 2.2.7) (Emms & Kelly, 2019) and RAxML (v. 8.2) (Stamatakis, 2006) based on a concatenated alignment of 782 protein-coding sequences. We selected five monocot genomes for in-depth comparative analyses, based on their availability of recently updated gene models with >90% BUSCO scores. These genomes lack additional whole genome duplication (WGD) events more recent than the ρ WGD event documented for grasses (McKain et al., 2016). MCscan (as implemented in JCVI v. 1.1.7) (Tang et al., 2008) was used to compare syntenic depths between *P. australis* and other monocot genomes. SynMap (Lyons & Freeling, 2008) and CLfinder pipeline (Oh & Dassanayake, 2019) were used to detect colinear paralogue and orthologue pairs within the *P. australis* genome and between *P. australis* and monocot genomes, respectively. We estimated synonymous substitution rates at four-fold degenerate sites using codeml (Yang, 1997) as described previously (Oh & Dassanayake, 2019). We detected orthologue groups among representative gene models of *P. australis* and other monocot genomes using OrthoFinder (v. 2.2.7) (Emms & Kelly, 2019) and MMseqs2 (Steinegger & Söding, 2017) and subsequently identified *P. australis* orthologue groups that are "conserved" and "duplicated" in orthologue copy numbers compared to other monocot species as detailed in Supporting Information Methods.

Gene ontology (GO) annotations were transferred to the *Phragmites* representative gene models based on sequence similarities to plant proteins with GO annotations as of 1 January, 2020 in the GO consortium (http://geneontology.org/). In short, *Phragmites* protein sequences were compared with reference sequences with a GO annotation using MMseqs2 with the maximum sensitivity (-s 7.5). If the protein alignment covers minimum 30% of both the query and subject sequences, the GO annotation was transferred to the *Phragmites* protein. We used BiNGO (Maere et al., 2005) to detect GO terms enriched in *Phragmites* gene models. GO terms were further clustered and summarized using GOMCL (Wang et al., 2020).

2.4 | Endophyte treatment, transcriptome assembly, and RNA-seq analyses

Invasive and native genotypes of *P. australis* were propagated from rhizome cuttings, grown in a growth chamber for 60 days, and subjected to *Alternaria alternata* (accession KT923239) (Clay et al., 2016) fungal endophyte inoculation. *A. alternata* was used for inoculations

because it was widespread across all regions sampled and was the second most common endophyte isolated from invasive *Phragmites* leaf tissues (Clay et al., 2016). The endophyte has also been shown to enhance allelopathic effects of host plants (Aschehoug et al., 2014). RNA for control and endophyte-treated samples were extracted from mature leaf and rhizome tissue separately as eight biological replicates (four plants, sampled twice per plant) per sample for six genotypes (three native and three invasive genotypes) and used to generate 192 RNAseq libraries following Illumina TruSeq library preparation guidelines as described in Supporting Information Methods. RNA-seq reads were filtered for adapter sequences using FASTP (Chen et al., 2018) and aligned to the reference genome using HISAT2 (v. 2.2). Expression counts for each gene model were estimated using StringTie (v. 2.0.1) with default parameters (Pertea et al., 2016). RNA-Seq reads were mapped to protein-coding gene models in the reference genome using bowtie2 (v. 2.4.4) with the "--very-sensitive" option to assess the rate of mapped reads among transcriptome samples. Differentially expressed genes (DEGs) were identified based on a minimum 2-fold difference in expression levels, with adjusted p-values < .05 estimated by DESeq2 (Love et al., 2014), between pairs of genotypes or between endophyte-inoculated and control samples within each genotype. In addition, putative protein-coding transcript sequences were obtained from filtered RNA-Seq reads by the Trinity (v. 2.1.1) and TransDecoder (v. 5.5.0) pipeline (Haas et al., 2013) (Table S2) and used for estimating the phylogenetic relationship of *P. australis* subspecies and genotypes using the Agalma pipeline (v. 2), which was designed to work with orthologous gene alignments derived from both genomes and transcriptomes (Dunn et al., 2013).

3 | RESULTS

3.1 | Assembly and annotation of the common reed *Phragmites australis*

We sequenced the genome of an invasive genotype of *Phragmites australis* ssp. *australis* collected in the U.S. Fish and Wildlife Service, Ottawa National Wildlife Refuge, Ohio (Figure 1c, star-marked "G"). We obtained ~42.19 Gbp of high confidence sequence data from leaf genomic DNA, representing a 37-fold genome coverage, using PacBio SMRT sequencing technology. Assembled using Canu (Koren et al., 2017) (version 1.4; see Methods), the reference genome provides 1.14 Gbp of 13,411 gap-free contigs with more than half of the assembled genome captured in 1370 contigs (N50) larger than 194.6 kbp (L50). The largest contig was 3.22 Mbp (Table 1). Illumina short reads mapped to the primary assembly confirmed the genotype used as the reference to be functionally diploid based on single nucleotide polymorphisms that represented a nonreference allele frequency distribution with a peak at 0.5, as expected with a functionally diploid genome (Figure S1). In total, 56.19% of the genome consisted of repetitive sequences, with sequences derived from long terminal repeat (LTR) retrotransposons constituting 36.42% of the genome (Table S3). Based on the repeat-masked genome sequence, we annotated 64,857 protein-coding gene models with a total length of 72.35 Mbp, which accounted for 6.35% of the genome (Table 1). The genomic location of these gene models and their best orthologues in rice and *Arabidopsis thaliana* are provided in Supporting Information data set 1. BUSCO analysis (Waterhouse et al., 2018) found 93.3% of single-copy gene models expected for land plants in the *Phragmites* reference genome (Table S1).

TABLE 1 The *Phragmites australis* draft genome

Assembly	
Assembled genome size	1,139,927,050 bps
Largest contig size	3,219,705 bps
N50 contig length (L50)	194,574 bps
Total number of contigs	13,411
Number of contigs >1 Mbps	67
Number of contigs >N50	1370
Annotation	
Number of protein-coding gene loci	64,857
Total length of predicted ORFs	72,347,598 bps
Longest ORF	14,790 bps
Median length of ORFs	927 bps
Proportion of ORFs in the genome	6.35%
Proportion of repeats in the genome	56.19%
SINEs	0.14%
LINEs	1.74%
LTR elements	36.42%
DNA elements	11.43%
Unclassified repeats	6.46%

We constructed a species tree using 6404 gene models representing 782 gene families from *P. australis* and 13 other published grass genomes (PLAZA monocot database v. 4.5) (Van Bel et al., 2018), with pineapple (*Ananas comosus*, Bromeliaceae) as the outgroup. *Phragmites australis*, representing the first genome from the Arundinoideae subfamily, was placed sister to subfamily Chloridoideae, consistent with the PACMAD clade species tree (Figure 2) (Burke et al., 2016; Hardion et al., 2017).

3.2 | Signatures of a whole genome duplication event in the lineage of *P. australis*

We found signatures of a previously unreported whole genome duplication (WGD) in the *P. australis* genome that occurred after its divergence from the subfamily Panicoideae (Figure 3a–d). This was detected as a more recent lineage specific event following the ρ WGD event identified for multiple lineages in Poales (McKain et al., 2016). We compared the *P. australis* genome with five representative monocot reference genomes, including pineapple (*A. comosus*) and four grass species that did not experience an additional WGD event

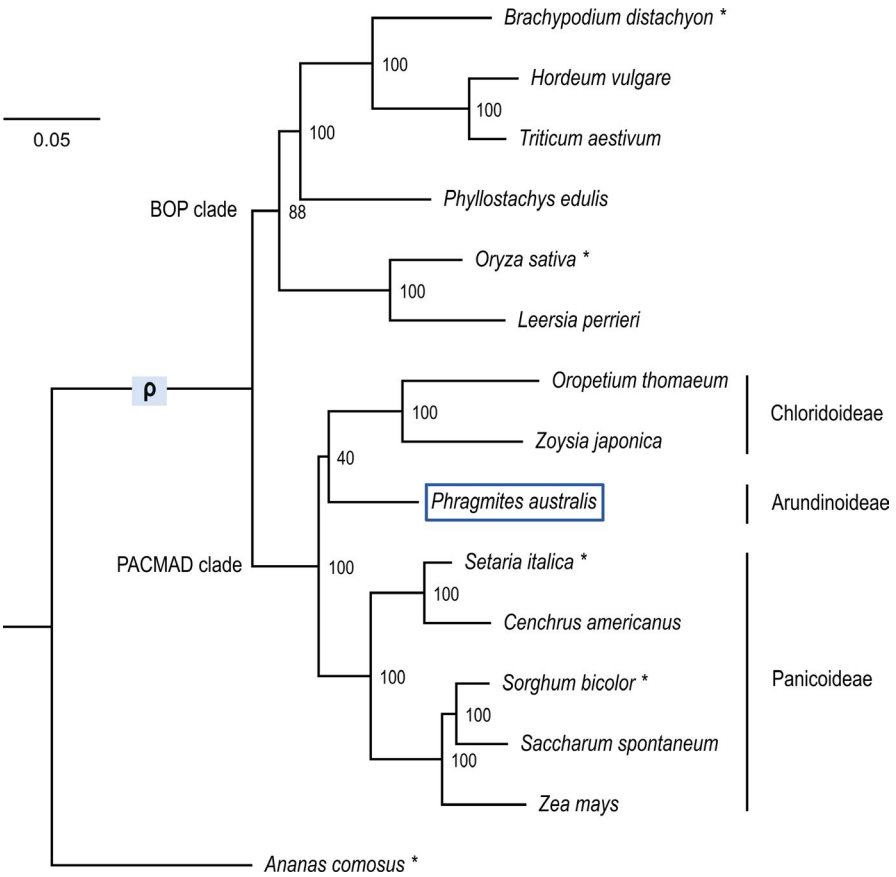


FIGURE 2 Phylogenetic position of *Phragmites australis* in the underinvestigated Arundinoideae subfamily, based on the draft genome. For the *P. australis* draft genome and 14 monocot genomes publicly available in Monocots PLAZA (v. 4.5) (Van Bel et al., 2018), a maximum likelihood species tree was constructed based on 26,878 amino acid alignments from 782 orthologue groups, selected based on the criteria that at least seven species among the set had a single orthologue. All sites that included gaps in more than 20% of taxa were excluded. The number in each branch shows percent support from 1000 bootstrap replicates. The branch with the ρ genome duplication (McKain et al., 2016) is marked, and species with asterisks were used for comparative analyses

following the ρ duplication (Figures 2 and 3d marked with asterisks). The *P. australis* genome presented 36.7% BUSCO genes as duplicated, while only 1%–5% BUSCOs were duplicated in the five comparator genomes (Figure 3a and Table S1). A genome-wide alignment between *Setaria italica* and *P. australis* found that 50.2% of *S. italica* gene models are represented twice in the *P. australis* genome as colinear orthologues in synteny blocks (Figure 3b, and Figure S2a). Similar patterns were observed in comparisons between *P. australis* and the other four monocot genomes (Figure S2a). A genome-wide self-alignment using SynMap identified 14,005 gene loci, comprising 21.6% of all *P. australis* protein-coding genes, that were organized into 1501 paralogous synteny blocks consisted of at least five colinear paralogue pairs. Synteny blocks were widespread across the *P. australis* genome and found in 66.8% of all contigs with 10 or more protein-coding gene loci (Figures S2b,c), further supportive of a WGD event in the lineage leading to *P. australis*.

To further assess the timing of genome duplications that resulted in the observed paralogous synteny blocks, we used neutral evolutionary substitutions calculated for four-fold degenerate (4D) sites in codons that allows positioning of timing of duplicated events within a clade. We plotted the distribution of synonymous substitution rate (*Ks*) at 4D sites for *P. australis* colinear paralogue pairs and for colinear orthologue pairs between *P. australis* and comparator species (Figure 3c). The peak *Ks* for *P. australis* colinear paralogue pairs (Figure 3c, dark grey) was smaller than those observed for any pairwise comparisons between species, while larger than the value found between invasive *P. australis* ssp. *australis* and native *P. australis* ssp. *americanus* (Figure 3c, bottom row). We also found a small proportion of synteny blocks within the *P. australis* genome that probably represent the ρ duplication at the root of the Poaceae (Figure S3). Our analysis suggests that the majority of colinear paralogues in synteny blocks are derived from a single WGD event that occurred after the divergence of *Phragmites* from the Panicoideae, represented here by *S. italica* and *S. bicolor*, but before the divergence between the subspecies *australis* and *americanus* (Figure 3d).

3.3 | Biased retention of transcription factors and signalling-related genes following the lineage-specific genome duplication in *P. australis*

While the synteny blocks are widespread across the *P. australis* genome, the relatively short length of each synteny block (Figure S2c) and the fact that only 21.6% of duplicated genes were retained in synteny blocks indicates a substantial fractionation following genome duplication, probably due to the selective retention of duplicates that enhance plant fitness while reducing nonessential duplicates or extra copies of genes under strong dosage-dependent selection. We therefore compared the functions of duplicated *P. australis* genes relative to those genes with conserved copy numbers based on orthologue groups including other grass species (Figure 4).

For all orthologue groups (OGs) detected among *P. australis* and the five comparator monocot species (Methods and Supporting

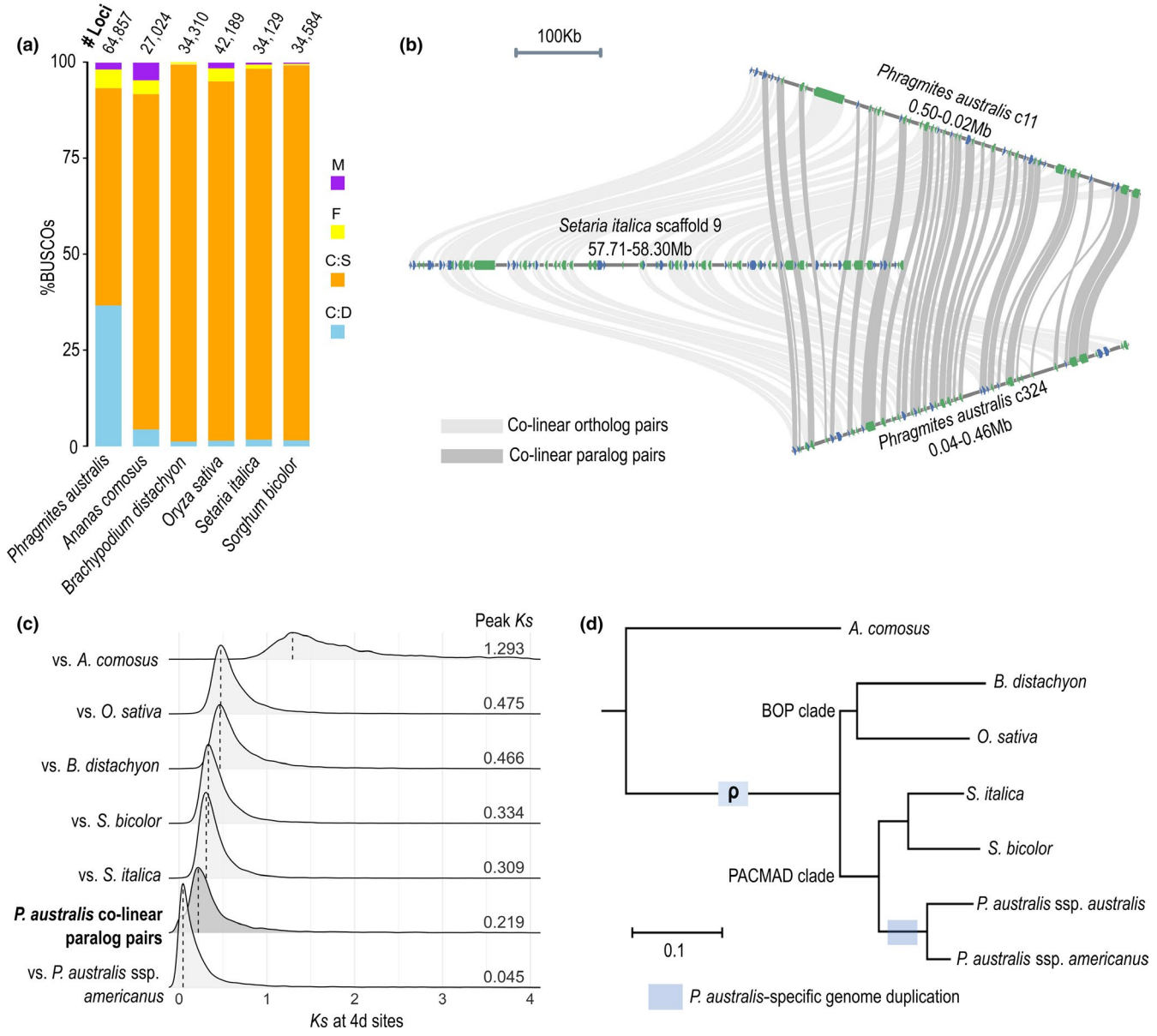
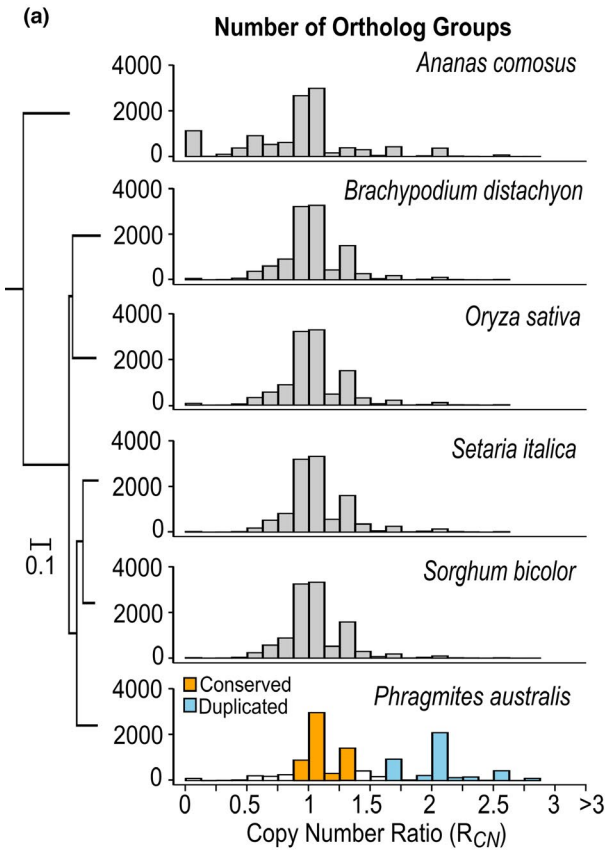
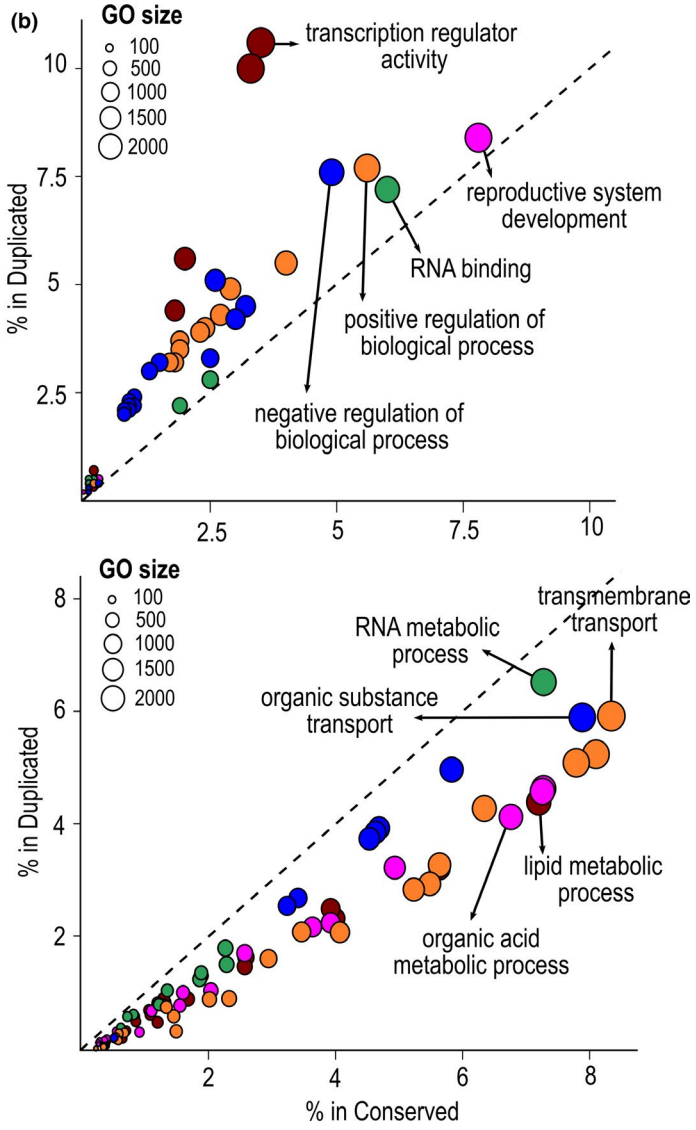


FIGURE 3 Signatures of *Phragmites australis*-specific whole genome duplication (WGD). (a) Percentages of complete duplicated (C:D), complete single-copy (C:S), fragmented (F), and missing (M) orthologues among 1375 Benchmarking universal single-copy orthologues (BUSCOs) and the number of protein-coding gene loci (# loci), in the genomes of *P. australis* and other monocot species. (b) An example microsynteny between a 500-kb *Setaria italica* genomic block and two duplicated *P. australis* genomic blocks. Ribbons connect colinear orthologue (light grey) and paralogue (dark grey) pairs identified by MCscan (Tang et al., 2008) as described in Methods. (c) Synonymous substitution rates (*Ks*) at four-fold degenerate (4d) sites were estimated for colinear orthologue pairs between *P. australis* and *A. comosus* (total 2720 orthologue pairs), *B. distachyon* (10,559 pairs), *O. sativa* (11,176 pairs), *S. italica* (12,878 pairs), and *S. bicolor* (12,322 pairs), respectively, as well as 6600 colinear paralogue pairs detected within the *P. australis* genome. For comparison with the native *P. australis* ssp. *americanus*, we used 11,445 reciprocal best homologue pairs between *P. australis* reference genome and a de novo transcriptome assembly from *P. australis* ssp. *americanus*. Probability distributions and peak values of *Ks* are shown. (d) A maximum- likelihood species tree shows the branches associated with the *P. australis*-specific WGD, as well as the ρ WGD event shared among grasses (McKain et al., 2016)

Information data set 2), we calculated the orthologue copy number ratio (R_{CN}) for each species by dividing the copy number in the species with the average of the other five species (Figure 4a and Table S4). All species except *Phragmites* showed a peak at $R_{CN} \approx 1$. By contrast, the R_{CN} distribution in *P. australis* showed two peaks, one with R_{CN} at 1 (Figure 4a, "conserved") and a second peak at $R_{CN} \approx 2$ (Figure 4a, "duplicated"), indicating that a substantial number of *P. australis* OGs have their copy numbers doubled compared to other species (see Methods for details). We searched for enriched gene functions in the duplicated group (11,002 *P. australis* genes in 4113 OGs) and in the Conserved group (6981 *P. australis* genes in 5600 OGs). Figure 4b, c present the five largest functional clusters, detected using GOMCL



$$R_{CN}(i) = \frac{\text{Copy number (species } i\text{)}}{\text{Mean copy number in all species except } i}$$



(c)

Cluster	Representative GO term [a]	Duplicated		Conserved		Background
		Genes (%) [b]	P [c]	Genes (%) [b]	P [c]	Genes (%) [b]
D1	transcription regulator activity	834 (10.6)	**1.9E-16**	182 (3.5)	-	2963 (8.1)
D2	reproductive system development	665 (8.4)	**1.9E-2**	408 (7.8)	6.6E -1	2772 (7.5)
D3	RNA binding	567 (7.2)	**7.8E-5**	309 (6.0)	7.3E -1	2197 (6.0)
D4	positive regulation of biological process	616 (7.7)	**6.5E-3**	293 (5.6)	-	2524 (6.8)
D5	negative regulation of biological process	608 (7.6)	**1.9E-8**	259 (4.9)	-	2244 (6.1)
C1	lipid metabolic process	349 (4.4)	-	378 (7.2)	**1.5E-6**	2017 (5.5)
C2	organic acid metabolic process	368 (4.6)	-	382 (7.3)	**1.4E-6**	2040 (5.5)
C3	RNA metabolic process	519 (6.5)	5.4E-2	382 (7.3)	**6.4E-4**	2172 (5.9)
C4	transmembrane transport	471 (5.9)	-	438 (8.3)	**3.3E-2**	2693 (7.3)
C5	organic substance transport	469 (5.9)	-	414 (7.9)	**8.3E-3**	2467 (6.7)

FIGURE 4 Legend on next page

FIGURE 4 Functions enriched among *Phragmites australis* genes that remained duplicated after the WGD event. (a) Comparison of orthologue copy numbers between *P. australis* and five other monocot species. For each orthologue group identified by OrthoFinder as described in Methods, copy number ratio (R_{CN}) was calculated for each species by dividing the orthologue number in the species with the mean orthologue copy number in the other five species. Histograms show a shift towards increased R_{CN} uniquely in *P. australis*. We identified "conserved" (orange) and "duplicated" (sky-blue) groups among *P. australis* genes whose copy numbers remain unchanged and uniquely increased, respectively, compared to other monocot species. (b, c) GO terms enriched in either duplicated or conserved *P. australis* gene groups. The proportion of genes annotated with each GO term in duplicated and conserved groups is plotted as circles in (b). GO terms at least 80% overlapping with a bigger GO term are clustered and the largest five GO clusters enriched in either group were shown with the same colour in (b) and (c). [a]The largest GO term in each GO cluster; MF, molecular function; BP, biological process. [b]Percentages are the total number of genes with a GO annotation in each group. [c]*p*-Values of enrichment compared to the background, after Benjamini-Hochberg correction for multiple testing. Values <0.05 are in bold

(Wang et al., 2020), considering GO terms significantly enriched exclusively in either conserved or duplicated groups. The largest clusters included 72% and 64% of all *Phragmites* genes represented by 123 and 146 GO terms enriched in the duplicated and the conserved groups, respectively (Supporting Information data set 3). Functions enriched exclusively in the duplicated group were largely related to regulation of gene expression (Figure 4b,c). For example, "transcription regulator activity" is the most enriched functional cluster among duplicated genes (Figure 4b,c, cluster D1), representing 10.6% of all genes in the duplicated group with a GO annotation, compared to only 3.5% and 8.1% in the conserved group and all genes (used as the background for the enrichment analysis), respectively (Supporting Information data set 3). By contrast, the conserved group was enriched in functions associated with primary metabolic processes and transport (C1–5 in Figure 4b, c and Supporting Information data set 3). Selective gene retention after WGD events are often linked to adaptive traits or traits leading to lineage diversification. The biased retention of transcription factors and regulatory processes in *P. australis* spp. *australis* following its most recent WGD event is indicative of greater genomic plasticity conducive to invasive lifestyles (Clark & Donoghue, 2018).

3.4 | Divergence in basal transcriptome profiles between invasive and native *P. australis* subspecies

To gain further insight into genetic factors promoting invasiveness, we compared transcriptomes of invasive and native *Phragmites* subspecies collected from the Great Lakes region (Figures 1c and 5a). We generated eight replicates of RNA-seq samples from leaf and rhizome tissues, respectively (see Methods), followed by independent de novo transcriptome assembly of the six genotypes (Table S2). The maximum likelihood tree based on concatenated alignments of 5394 homologue groups separated the six genotypes into invasive and native subspecies as expected, and the three invasive genotypes were placed together with the invasive genotype used as the reference genome (Figure 5b).

We aligned RNA-seq reads to the reference genome and gene model sequences to explore the differences in basal transcriptome profiles between invasive and native genotypes and estimate the relative abundance of reference genes, as detailed in Methods (Supporting Information data set 4). Proportions of RNA-seq reads aligned to protein-coding gene model sequences did not show significant differences between invasive and native genotypes, while the native genotypes showed overall slightly fewer reads aligned to the reference genome sequences (Table S5). Further, when normalized to the total number of reads aligned to protein-coding gene models and expressed gene compositions as implemented in DESeq2, the distribution of estimated expression values and number of significantly differentially expressed genes (DEGs) did not show a bias towards either invasive or native genotypes (Figure S4). The leaf- and rhizome-derived transcriptomes were distinct from each other in a principal component analysis (PCA) (on PC1), while both showed additional separation between native and invasive genotypes (on PC2) (Figure 5c). In agreement with the PCA, the number of DEGs between invasive and native genotypes was much greater than DEGs detected between any other genotype comparisons (Figure 5d). On average, 11.5 ± 1.0% of the 64,857 reference gene models showed higher basal expression in the invasive genotypes compared to the native genotypes (Figure 5d, upper right sections of the diagonal plots). Similarly, 12.1 ± 2.1% of gene models were more highly expressed in the native genotypes than in the invasive genotypes (Figure 5d, lower left sections).

In the leaf samples, "response to stimulus" and "response to stress" were the largest representative GO terms showing significant enrichment among DEGs with higher basal expression in the invasive genotypes compared to the native genotypes in all cross-genotype comparisons (Supporting Information data set 5). Interestingly, among child GO terms of these two GO terms, only those categorized as "response to biotic stimulus" and "defence response" showed significant enrichment in the invasive genotypes, but not "response to abiotic stress" (Figure S5). On average, 9.2% of all DEGs were annotated with "defence response" and showed higher basal expression in the invasive genotypes compared to 6.4% and 6.2% in the native genotypes and the background, respectively (Figure 5e, upper panel). This bias towards a higher basal expression of genes associated with biotic stress and defence in the invasive genotypes was less clear in the rhizome tissue (Figure 5e, lower panel). By contrast with the functional enrichment in defence responses observed for the invasive genotypes, the native genotypes were biased towards genes associated with "transmembrane transport" and its child GO terms, including "ammonium transport", are represented by DEGs with higher basal expression in all cross-genotype comparisons (Figure 5f and Supporting Information data set 5). This

transcriptomic signal was more prominent in leaf samples compared to rhizomes as similarly observed for the defence response.

3.5 | Divergence in invasive and native *P. australis* transcriptomic responses to biotic stress induced by fungal endophyte inoculation

We inoculated target genotypes with the fungal endophyte commonly isolated from *P. australis*, *Alternaria alternata* (Clay et al., 2016) (see Methods), to investigate further the transcriptomic signal established at basal expression contrasting the invasive genotypes from native genotypes biased towards biotic stress responses (Figures 1c, 5a and 6). We compared total 24 RNA-seq profiles (i.e., two tissue types, two treatments, and six genotypes) in eight replicates (192 samples) generated for leaf and rhizome tissue separately harvested at preinoculation (basal expression) and post-inoculation (response to biotic stress) to deduce DEGs and their representative enriched functions in six individual genotypes (see Methods, Supporting Information data set 4). In general, more DEGs were significantly induced (Figure 6a) than repressed (Figure 6b) in response to the endophyte inoculation. The two invasive genotypes IOH1 and IMI1 showed more attenuated responses to endophyte inoculation compared to the three native genotypes. However, the invasive genotype IMI3 showed a response similar in magnitude to the native genotype NMI2 (Figure 6a,b and Supporting Information data set 6). Further, IMI3 and the three native genotypes shared a substantial number of endophyte-induced DEGs (Figure 6a red boxes).

We searched for enriched functional associations between endophyte-induced and repressed genes in the invasive and native genotypes (Figure 6c and Supporting Information data set 6). In the native genotypes and IMI3 genotype, endophyte-induced genes were enriched in "ATP binding" and "defence response", while endophyte-repressed genes were enriched in processes associated with photosynthesis (Figure 6c). A closer inspection into these induced genes annotated under GO molecular function "ATP-binding" and GO biological process "defence response" revealed that many encode membrane receptor kinases known for roles in defence signalling (Supporting Information data set 6).

Despite genotypic variation in response to the endophyte inoculation, the native genotypes showed an overall stronger response in induced genes associated with biotic stress and defence, while these functions had higher basal expression in the invasive genotypes. Genotypes IMI3 and NMI2, while showing the highest number of induced DEGs annotated under defence response, also showed the highest number of repressed DEGs associated with photosynthesis compared to the rest of the genotypes, suggesting that the strong defence response concurrently downregulated photosynthesis in these two genotypes (Figure 6c–e). This is less prominent in NOH1 where both the number and enrichment of photosynthesis-related endophyte-repressed DEGs were among the smallest compared to other genotypes.

4 | DISCUSSION

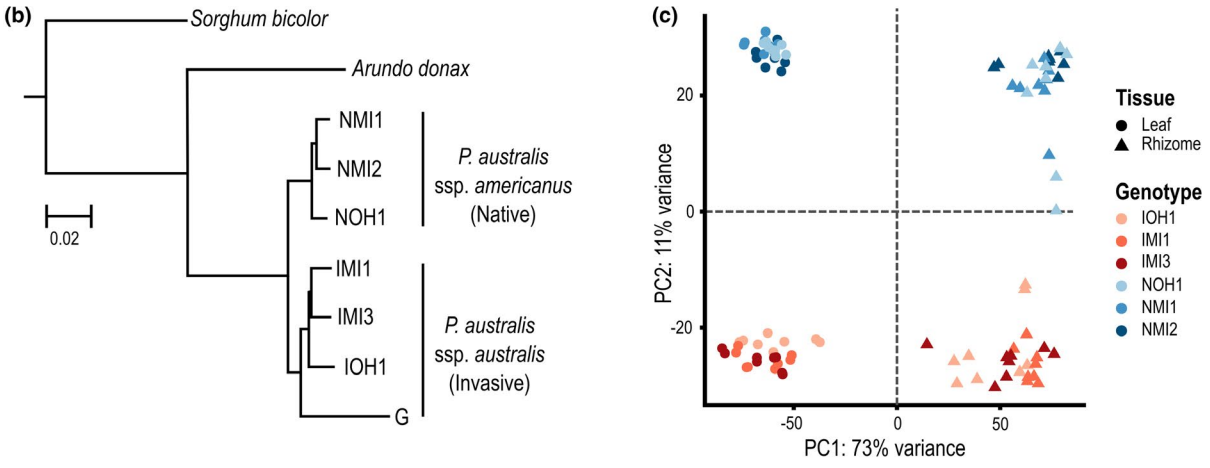
The genome of common reed (*Phragmites australis*) reported here provides the first reference genome for this ecologically important species, a genomic model for the model invasive grass *P. australis* spp. *australis* (Martin & Blossey, 2013; Saltonstall, 2002), and the first reference genome for the grass subfamily Arundinoideae (Figures 1 and 2, Table 1). Our genomic analyses, including nonreference allele frequency distributions and comparative genomics with other grasses, indicate that the reference genome was derived from a functionally diploid plant although *P. australis* has been generally reported as a tetraploid in North America (Clevering & Lissner, 1999). The reference genome provides a core set of reference genes within unique genomic backgrounds independent of chromosome-level ploidy of contemporary polyploidy. However, it can be used as a fundamental genetic resource to tag unique genomic loci fluorescently and monitor chromosomal-level variation, identify multivalent chromosomes, or study whether higher-ploidy level populations can result from population-specific additional successive WGD events, endoreplication found in clonal propagation, and/or introgression between populations.

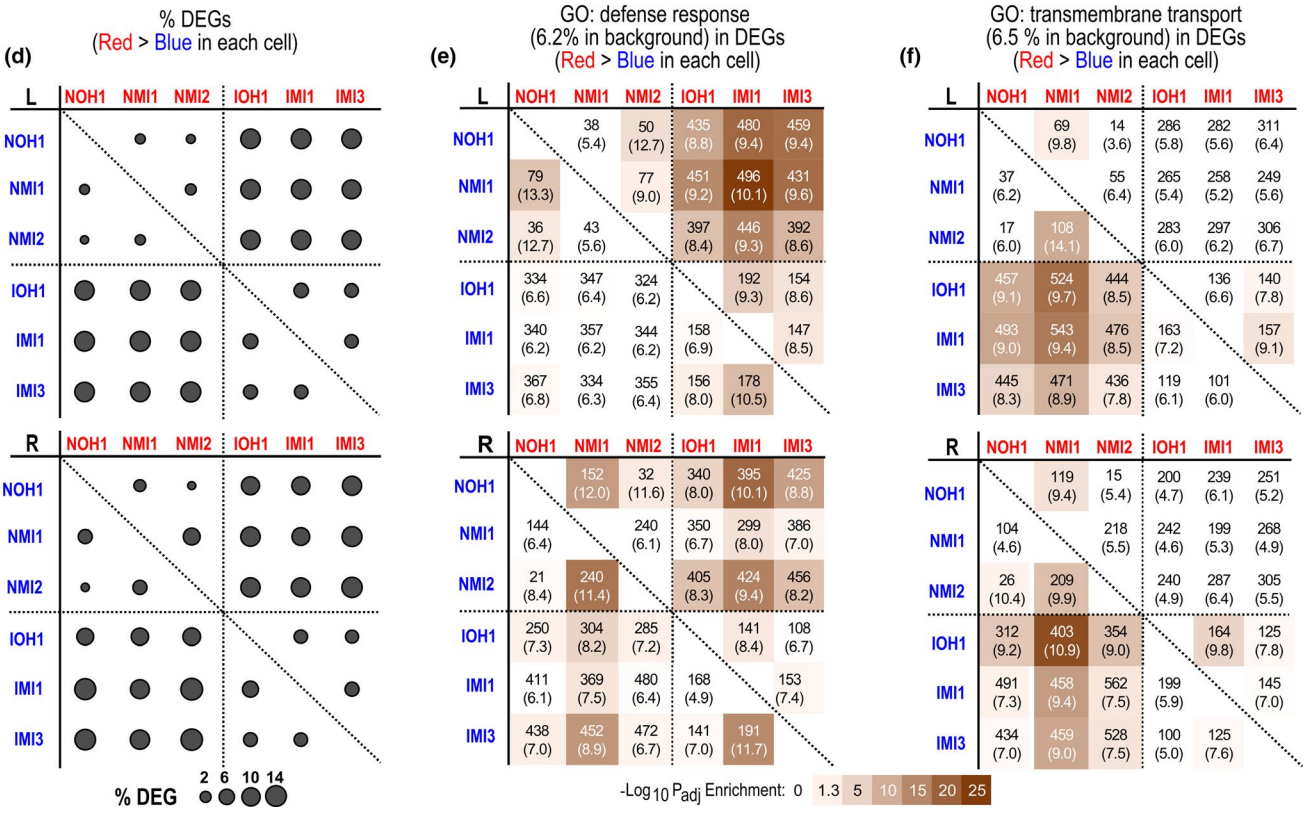
Contemporary polyploids can show cytological diploidization (chromosomal changes leading to bivalent chromosome pairing as a cellular mechanism) or genic diploidization (a WGD event often

FIGURE 5 Comparison of transcriptomes between invasive and native *Phragmites australis* genotypes. (a) Details on genotype collections of invasive (I) and native (N) subspecies from locations marked on Figure 1c,b A maximum likelihood tree based on deduced protein sequences of 5394 homologous gene groups. Protein sequences were deduced from transcriptomes de novo assembled using RNA-seq reads derived from the six *P. australis* genotypes shown in (a) as well as the reference gene models (marked with "G"). Publicly available transcriptome and genome sequences were used for *Arundo donax* (Barrero et al., 2015) and *Sorghum bicolor*, respectively. All branches were 100% supported by 100 bootstrap tests. (c) RNA-seq reads from leaf and rhizome tissues of the six genotypes were aligned to the *P. australis* reference genome. Principal component analysis separated different tissues and genotypes. (d) Differentially expressed genes (DEGs) showing significant (adjusted p-value <.05 and fold-difference ≥2) changes in basal-level expression in leaf (L) and rhizome (R) were identified between all pairs of genotypes. In the diagonal plot, the circle in each cell represents the proportion of DEGs among 64,857 *P. australis* reference gene models in which the genotype in the column (red) shows higher basal-level expression than the clone in the row (blue). (e, f) GO terms "defence response" (e) and "transmembrane transport" (f) showed enrichment among DEGs in which invasive and native genotypes showed higher basal-level expression in pairwise comparisons, respectively. Each cell shows the number and percentage of DEGs annotated with the GO term among all DEGs showing higher basal-level expression in the genotype specified by the column (red) compared to the genotype by the row (blue). Adjusted p-values of enrichment were calculated compared to a background of 41,595 reference gene models annotated with any GO term and represented as a colour heatmap

(a)

ID	Genotype	Haplotype	Description	Coordinates
IMI1	Invasive	M	Railroad, Garvey Rd, Chelsea, MI, USA	42.307211, -84.059662
NMI1	Native	E	Private, Garvey Rd, 50 meters SW from IMI1 site	42.307135, -84.060249
NMI2	Native	E	Private, McKinley Rd, Chelsea, MI, USA	42.337980, -84.00040
IMI3	Invasive	M	Sandusky State Game Area, Michigan Dept. Natural Resources, Sandusky, MI, USA	43.394026, -82.790088
IOH1	Invasive	M	Show Pool, U.S. Fish and Wildlife Service, Ottawa National Wildlife Refuge, Oak Harbor, OH, USA	41.614505, -83.195807
NOH1	Native	E	Show Pool, 110 meters NW from IOH1 site	41.614755, -83.197046
G	Invasive	M	Show Pool, 3 kilometers from IOH1 site	41.631281, -83.228892





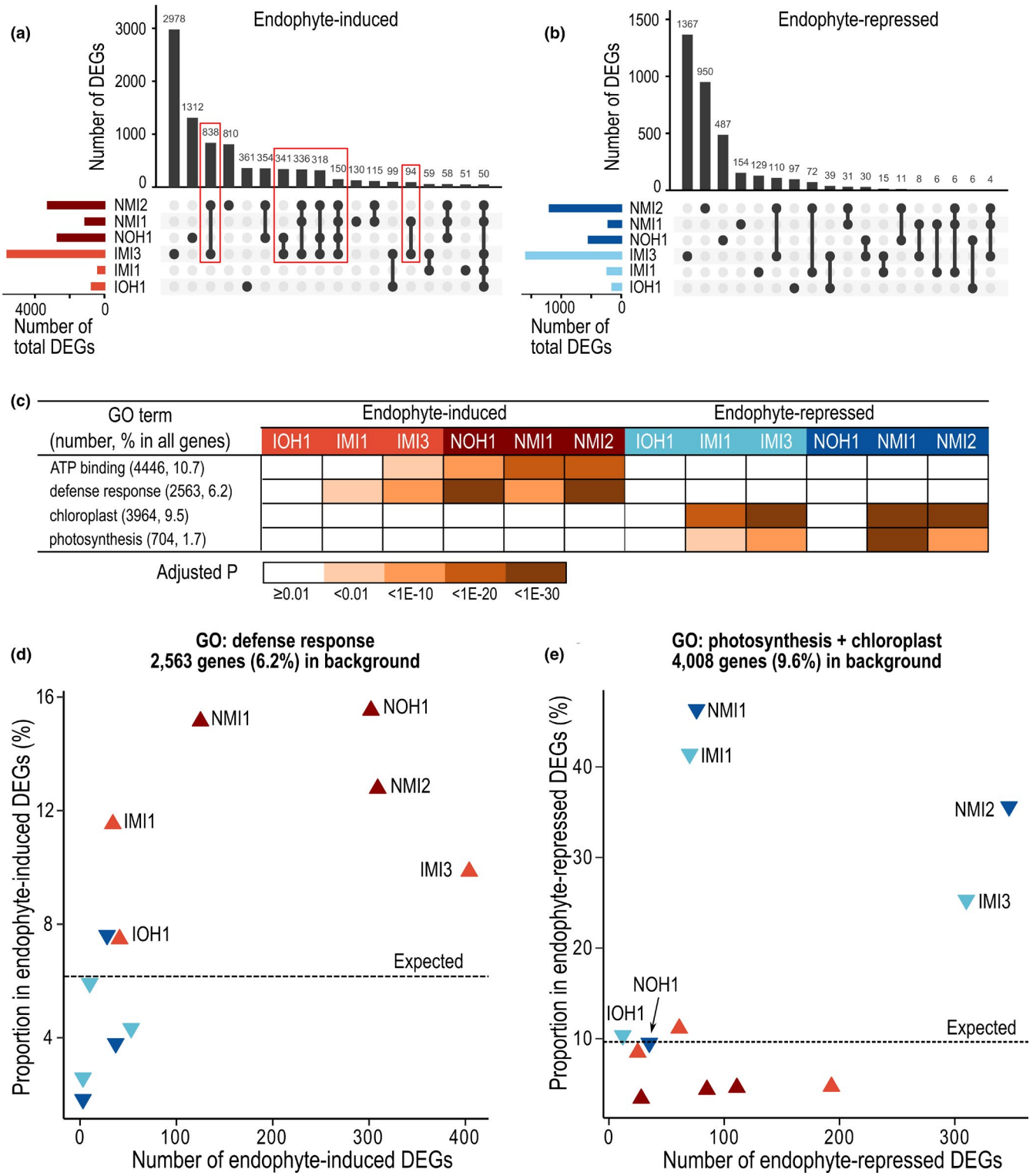


FIGURE 6 Transcriptome responses to *Alternaria alternata* fungal endophyte inoculation of invasive and native *Phragmites australis* genotypes. (a, b) Upset plots showing the number of shared and unique endophyte-induced (a) and repressed (b) DEGs among the six invasive and native genotypes. Red boxes show endophyte-induced DEGs shared between IMI3 genotype and a native genotype. (c) The pattern of GO enrichment in six clones involving the largest number of endophyte-induced DEGs (represented by GO terms "ATP binding" and "defence response") and endophyte-repressed DEGs (represented by GO terms "chloroplast" and "photosynthesis"). (d, e) Number and percent proportion of endophyte-induced (upward triangles) or repressed (downward triangles) DEGs annotated with the GO term "defence response" (d) or either of the GO terms "photosynthesis" and "chloroplast" (e). The percent proportions in the entire *P. australis* gene models are marked with dashed lines as the expected value when there is no enrichment

followed by gene loss as a molecular mechanism), and these two mechanisms are independent from each other (Li et al., 2021; Ma & Gustafson, 2005; Tayalé & Parisod, 2013). The random shotgun sequencing of genomic DNA performed using Illumina reads suggest that the source DNA of the *P. australis* reference genome is derived from a functional diploid organism (Figure S1). Functional diploidy together with disomic inheritance is common in sexually reproducing plants even when they represent true polyploids, as shown for nearly half of polyploid species demonstrating bivalent chromosome pairing (Li et al., 2021). The status of *P. australis* as a functional diploid is in line with previous studies reporting cytological evidence of bivalent pairing of *P. australis* chromosomes between metaphase I and II during pollen development (Gorenflot, 1976) and karyotyping of cells from root tissues (Raicu et al., 1972).

WGD events are ubiquitous in land plant evolution at large, but less common within lineages, and mark important divergence points in lineage evolution while serving as a significant source for novel adaptations (Soltis & Soltis, 2016). We found a previously unreported whole genome duplication (WGD) event leading to the *P. australis* lineage, independent of the three ancient WGD events known in the grasses (McKain et al., 2016). The WGD event detected in the *P. australis* genome predates the divergence between spp. *australis* and *americanus* but postdates the divergence from the Panicoideae (Figures 3 and 4). As expected with substantial gene fractionation following WGD events, the *P. australis* genome lost up to 48% of its duplicated genes (Figure S2a) but retained over 14,005 duplicated genes. The high level of gene fractionation observed in the current reference genome suggests that it is a mesopolyploid (Li et al., 2021), given that *Phragmites* has undergone a more recent WGD that is still in the process of genic diploidization (Wang et al., 2011). The retained duplicated genes in the reference genome were strongly enriched with transcription factors and other genes associated with regulatory processes. Genes encoding transcription factors have been preferentially retained over other gene groups following paleo WGD events in other pan-global species, such as *Arabidopsis thaliana* (Blanc & Wolfe, 2004; De Bodt et al., 2005; Freeling et al., 2007), and are thought to be a hallmark feature of WGD events underlying rapid diversification and global distribution of angiosperms (Birchler, 2019; Cheng et al., 2018). They may also provide a selective advantage in invasive species when introduced to new environments (Moura et al., 2021; te Beest et al., 2012). Therefore, the duplicated gene space detected in the *Phragmites* lineage enriched in regulatory genes provides novel genetic material for selection to act facilitating its potential as an invasive species.

Adaptive innovations initiated at the genomic level and further diversified at the transcriptome level can lead to distinct ecological fates. Our comparative analysis using native and invasive genotypes from the Great Lakes region of North America indicates that gene expression associated with defence against biotic stress is primed in the invasive genotypes compared to native genotypes. When a biotic stress response was induced by inoculation with the fungal endophyte *A. alternata*, the native genotypes were consistently more responsive (Figures 5 and 6). Altered interactions with pathogens and native plant species by invasive *Phragmites* could contribute to its success (Mangla & Callaway, 2008; Schroeder et al., 2020). A primed transcriptome with elevated transcript abundance for stress-responsive genes has been recognized as an adaptive strategy exhibited by plants adapted to various abiotic or biotic stresses (Baccelli et al., 2020; Dräger et al., 2004; Wang, DiTusa, et al., 2021). A primed transcriptome for biotic stresses allows faster defence responses to an array of biotic stresses compared to an induced response from those species/populations that are not primed for these biotic stresses. Therefore, populations with primed transcriptomes to biotic stresses are expected to have a selective advantage over nonprimed populations if introduced to environments with a high biotic stress. It will be important to determine if the expression differences reported here are geographically variable and extend beyond the specific samples used in this study.

Phragmites australis is known for its intraspecific cytological ploidy polymorphism (Clevering & Lissner, 1999; Lambertini et al., 2006) coincident with its widespread success as a globally recognized invasive species. A higher gene content present in polyploids may predispose a species for invasive lifestyles adapted to a broad range of habitats potentially defined by both biotic and abiotic stresses (te Beest et al., 2012). However, whether ploidy level plays a deterministic role in facilitating invasiveness in *P. australis* remains debatable when categorical higher ploidy levels exclusively assigned to invasive genotypes are absent. For example, Gulf Coast *P. australis* subsp. *berlandieri* is cytologically hexaploid yet is not considered invasive, and cytologically octoploid *P. australis* has been documented along the Charles River in Massachusetts (Keller, 2000). A recent study (Wang, Wang, et al., 2021) showed that there were gene expression-based functional enrichment differences related to environmental stress tolerance between tetraploid and octoploid *P. australis* populations. However, our comparative transcriptomics analysis does not limit the identification of traits associated with enriched transcripts in certain genotypes that may additionally have distinct ploidy levels because expression of genes is determined based on a unique genomic locus in the reference genome. This approach does not differ between transcriptomes analysed from diploid or polyploid organisms when the reference genome is created as a "haploid reference" assigned to the species. Further, the transcript expression observed from bulk tissues in plants generally include a mixture of cells at different ploidy levels regardless of the species ploidy level (De Rocher et al., 1990). Even diploid plants such as *Arabidopsis thaliana* tend to have higher ploidy cells exceeding the diploid cell fraction with developmental age (Galbraith et al., 1991). Expression differences related to tissue-level polyploidy with plant age or induced by environmental stress, as well as species-level polyploidy, also provide potential mechanisms for invasive species adaptation in new environments (De Rocher et al., 1990; Kettenring & Mock, 2012; te Beest et al., 2012). Exploring *P. australis* pangenomes in future studies could provide further insight into how genic or cytological ploidy polymorphism generates selective advantages for invasive genotypes of *P. australis* relative to noninvasive genotypes.

Our results provide a foundation for the development of novel species- and subspecies-specific genetic approaches for control of

invasive *Phragmites* (Harvey-Samuel et al., 2017) and potentially other invasive plant species. There is widespread interest in controlling invasive *Phragmites* in many parts of its invasive range. Our genomic data help to identify unique and essential genes that could be targeted in genetic control approaches using RNAi with higher species specificity than attainable using chemical or mechanical control. Our transcriptomic data also point to potential target genes highly expressed in the invasive versus native genotypes. However, our results also provide some caution in particular approaches if targeted genes are duplicated or shared among both native and invasive groups. While RNAi approaches have been widely explored for controlling invasive insect pests and plant pathogens (Cagliari et al., 2019; Mamta & Rajam, 2017), to our knowledge, no RNAi-based treatments have been developed to control problematic plants given the dearth of genomic data from invasive plant species. Future attempts at genetic control for invasive *Phragmites* should take into account that genetic variation exists in invasive *Phragmites* genotypes and that native genotypes can co-occur in the same habitats.

ACKNOWLEDGEMENTS

The authors would like to thank The U.S. Fish and Wildlife Service and Michigan Department of Environment, Great Lakes and Energy for access to study sites and the Great Lakes Restoration Initiative for financial support. We thank Doug Rusch and Ram Podicheti and the staff of the Indiana University Center for Genomics and Bioinformatics for their help in initial genome and transcriptome analyses, Matt Filipek for laboratory assistance, and the Indiana University Greenhouse staff for their assistance with growing and propagating *Phragmites*. We thank Dr. Wesley Bickford for data collection and analysis support and Drs Ping Gong, Kathleen Ferris, and Simon Barak for their thoughtful comments on a previous version of this paper. This research was supported by the United States Geological Survey Cooperative Agreement G18AC00373 to KC. MD and DHO acknowledge the support of National Science Foundation awards MCB-1616827, NSF-IOS-EDGE-1923589, and the Next-Generation BioGreen21 Programme of the Republic of Korea (PJ01317301). CW was supported by an Economic Development Assistantship award from Louisiana State University. The authors also acknowledge the LSU High Performance Computing services for providing computational resources needed for data analyses.

CONFLICT OF INTEREST

The authors declare no competing interests. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

AUTHOR CONTRIBUTION

Keith Clay, Maheshi Dassanayake and Kurt P. Kowalski conceived and designed the experiments and obtained the funding. Philippa Tanford., Keith Clay, and Kurt P. Kowalski conducted data collections and related experiments. Dong-Ha Oh, Chathura Wijesinghege, and Maheshi Dassanayake designed the bioinformatic work flows and performed computational analyses. Dong-Ha Oh and Quynh N. Quach generated figures. All authors contributed to drafting and finalizing the manuscript.

OPEN RESEARCH BADGES

This article has earned an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at https://doi.org/10.6084/m9.figshare.14036756, https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA705976 and https://genomevolution.org/coge/GenomeInfo.pl?gid=59768.

DATA AVAILABILITY STATEMENT

All raw and assembled sequence data have been deposited to NCBI under BioProject PRJNA705976. In addition, the reference genome sequence, gene models, and representative RNA-Seq tracks are available in the CoGe database (https://genomevolution.org/coge/) with the Genome ID 59768, for browsing and other CoGe-embedded comparative analyses (Nelson et al., 2018). Large Supporting Information data sets have been deposited in Figshare (https://doi.org/10.6084/m9.figshare.14036756) and also in the USGS ScienceBase repository (https://doi.org/10.5066/P9NLU6Q4).

ORCID

Dong-Ha Oh https://orcid.org/0000-0003-1526-9814
Kurt P. Kowalski https://orcid.org/0000-0002-8424-4701
Quynh N. Quach https://orcid.org/0000-0001-6504-6358
Chathura Wijesinghege https://orcid.org/0000-0002-7688-4418
Maheshi Dassanayake https://orcid.org/0000-0003-3123-3731
Keith Clay https://orcid.org/0000-0002-3956-0887

REFERENCES

Allen, W. J., Young, R. E., Bhattarai, G. P., Croy, J. R., Lambert, A. M., Meyerson, L. A., & Cronin, J. T. (2015). Multitrophic enemy escape of invasive *Phragmites australis* and its introduced herbivores in North America. *Biological Invasions*, *17*(12), 3419–3432. https://doi.org/10.1007/s10530-015-0968-2

Aschehoug, E. T., Callaway, R. M., Newcombe, G., Tharayil, N., & Chen, S. (2014). Fungal endophyte increases the allelopathic effects of an invasive forb. *Oecologia*, *175*(1), 285–291. https://doi.org/10.1007/s00442-014-2891-0

Baccelli, I., Benny, J., Caruso, T., & Martinelli, F. (2020). The priming fingerprint on the plant transcriptome investigated through meta-analysis of RNA-Seq data. *European Journal of Plant Pathology*, *156*(3), 779–797. https://doi.org/10.1007/s10658-019-01928-3

Baker, H. G., & Stebbins, G. L. (1965). Genetics of colonizing species, proceedings. *International Union of Biological Sciences Symposia on General Biology 1964: Asilomar, Calif.*

Barrero, R. A., Guerrero, F. D., Moolhuijzen, P., Goolsby, J. A., Tidwell, J., Bellgard, S. E., & Bellgard, M. I. (2015). Shoot transcriptome of the giant reed, *Arundo donax*. *Data in Brief*, *3*, 1–6. https://doi.org/10.1016/j.dib.2014.12.007

Birchler, J. A. (2019). Genomic balance plays out in evolution. *The Plant Cell*, *31*(6), 1186–1187. https://doi.org/10.1105/tpc.19.00329

Blanc, G., & Wolfe, K. H. (2004). Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *The Plant Cell*, *16*(7), 1679–1691. https://doi.org/10.1105/tpc.021410

Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., Turner, K. G., Whitney, K. D., & Rieseberg, L. H. (2015). What we still don't know about invasion genetics. *Molecular Ecology*, *24*(9), 2277–2297. https://doi.org/10.1111/mec.13032

Bourgeau-Chavez, L. L., Kowalski, K. P., Carlson Mazur, M. L., Scarbrough, K. A., Powell, R. B., Brooks, C. N., Huberty, B., Jenkins, L. K., Banda, E. C., Galbraith, D. M., Laubach, Z. M., & Riordan, K. (2013). Mapping invasive *Phragmites australis* in the coastal Great Lakes with ALOS PALSAR satellite imagery for decision support. *Journal of Great Lakes Research*, *39*, 65–77. https://doi.org/10.1016/j.jglr.2012.11.001

Burke, S. V., Wysocki, W. P., Zuloaga, F. O., Craine, J. M., Pires, J. C., Edger, P. P., Mayfield-Jones, D., Clark, L. G., Kelchner, S. A., & Duvall, M. R. (2016). Evolutionary relationships in Panicoid grasses based on plastome phylogenomics (Panicoideae; Poaceae). *BMC Plant Biology*, *16*(1), 140. https://doi.org/10.1186/s12870-016-0823-3

Cagliari, D., Dias, N. P., Galdeano, D. M., Dos Santos, E. Á., Smagghe, G., & Zotti, M. J. (2019). Management of pest insects and plant diseases by non-transformative RNAi. *Frontiers in Plant Science*, *10*, 1319. https://doi.org/10.3389/fpls.2019.01319

Campbell, M. S., Holt, C., Moore, B., & Yandell, M. (2014). Genome annotation and curation using MAKER and MAKER-P. *Current Protocols in Bioinformatics*, *48*, 4.11.1–39. https://doi.org/10.1002/0471250953.bi0411s48

Caplan, J. S., Wheaton, C. N., & Mozdzer, T. J. (2014). Belowground advantages in construction cost facilitate a cryptic plant invasion. *AoB Plants*, *6*, plu020. https://doi.org/10.1093/aobpla/plu020

Catford, J. A., Jansson, R., & Nilsson, C. (2009). Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. *Diversity & Distributions*, *15*(1), 22–40. https://doi.org/10.1111/j.1472-4642.2008.00521.x

Cesarino, I., Dello Ioio, R., Kirschner, G. K., Ogden, M. S., Picard, K. L., Rast-Somssich, M. I., & Somssich, M. (2020). Plant science's next top models. *Annals of Botany*, *126*(1), 1–23. https://doi.org/10.1093/aob/mcaa063

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

Cheng, F., Wu, J., Cai, X., Liang, J., Freeling, M., & Wang, X. (2018). Gene retention, fractionation and subgenome differences in polyploid plants. *Nature Plants*, *4*(5), 258–268. https://doi.org/10.1038/s41477-018-0136-7

Chengxu, W., Mingxing, Z., Xuhui, C., & Bo, Q. (2011). Review on allelopathy of exotic invasive plants. *Procedia Engineering*, *18*, 240–246. https://doi.org/10.1016/j.proeng.2011.11.038

Clark, J. W., & Donoghue, P. C. J. (2018). Whole-genome duplication and plant macroevolution. *Trends in Plant Science*, *23*(10), 933–945. https://doi.org/10.1016/j.tplants.2018.07.006

Clay, K., Shearin, Z. R. C., Bourke, K. A., Bickford, W. A., & Kowalski, K. P. (2016). Diversity of fungal endophytes in non-native *Phragmites australis* in the Great Lakes. *Biological Invasions*, *18*(9), 2703–2716. https://doi.org/10.1007/s10530-016-1137-y

Clevering, O. A., & Lissner, J. (1999). Taxonomy, chromosome numbers, clonal diversity and population dynamics of *Phragmites australis*. *Aquatic Botany*, *64*(3), 185–208. https://doi.org/10.1016/S0304-3770(99)00059-5

Cronin, J. T., Bhattarai, G. P., Allen, W. J., & Meyerson, L. A. (2015). Biogeography of a plant invasion: plant-herbivore interactions. *Ecology*, *96*(4), 1115–1127. https://doi.org/10.1890/14-1091.1

De Bodt, S., Maere, S., & Van de Peer, Y. (2005). Genome duplication and the origin of angiosperms. *Trends in Ecology & Evolution*, *20*(11), 591–597. https://doi.org/10.1016/j.tree.2005.07.008

De Rocher, E. J., Harkins, K. R., Galbraith, D. W., & Bohnert, H. J. (1990). Developmentally regulated systemic endopolyploid in succulents with small genomes. *Science*, *250*(4977), 99–101. https://doi.org/10.1126/science.250.4977.99

Diagne, C., Leroy, B., Vaissière, A.-C., Gozlan, R. E., Roiz, D., Jarić, I., Salles, J.-M., Bradshaw, C. J. A., & Courchamp, F. (2021). High and rising economic costs of biological invasions worldwide. *Nature*, *592*(7855), 571–576. https://doi.org/10.1038/s41586-021-03405-6

Divíšek, J., Chytrý, M., Beckage, B., Gotelli, N. J., Lososová, Z., Pyšek, P., Richardson, D. M., & Molofsky, J. (2018). Similarity of introduced plant species to native ones facilitates naturalization, but differences enhance invasion success. *Nature Communications*, *9*(1), 4631. https://doi.org/10.1038/s41467-018-06995-4

Dogra, K. S., Sood, S. K., Dobhal, P. K., & Sharma, S. (2010). Alien plant invasion and their impact on indigenous species diversity at global scale: A review. *Journal of Ecology and the Natural Environment*, *2*(9), 175–186.

Doyle, J. J., & Dickson, E. E. (1987). Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon*, *36*(4), 715–722. https://doi.org/10.2307/1221122

Dräger, D. B., Desbrosses-Fonrouge, A.-G., Krach, C., Chardonnens, A. N., Meyer, R. C., Saumitou-Laprade, P., & Krämer, U. (2004). Two genes encoding *Arabidopsis halleri* MTP1 metal transport proteins co-segregate with zinc tolerance and account for high *MTP1* transcript levels. *The Plant Journal*, *39*(3), 425–439. https://doi.org/10.1111/j.1365-313x.2004.02143.x

Dunn, C. W., Howison, M., & Zapata, F. (2013). Agalma: an automated phylogenomics workflow. *BMC Bioinformatics*, *14*, 330. https://doi.org/10.1186/1471-2105-14-330

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 238. https://doi.org/10.1186/s13059-019-1832-y

Freeling, M., Rapaka, L., Lyons, E., Pedersen, B., & Thomas, B. C. (2007). G-boxes, bigfoot genes, and environmental response: characterization of intragenomic conserved noncoding sequences in *Arabidopsis*. *The Plant Cell*, *19*(5), 1441–1457. https://doi.org/10.1105/tpc.107.050419

Galbraith, D. W., Harkins, K. R., & Knapp, S. (1991). Systemic endopolyploidy in *Arabidopsis thaliana*. *Plant Physiology*, *96*(3), 985–989. https://doi.org/10.1104/pp.96.3.985

GBIF.org. (2020a). *Phragmites australis* subsp. *americanus*. GBIF Occurence Download. https://doi.org/10.15468/dl.xbuvdm

GBIF.org. (2020b). *Phragmites australis* subsp. *australis*. GBIF Occurence Download. https://doi.org/10.15468/DL.Q8HC34

Gorenflot, R. (1976). Le complexe polyploïde du *Phragmites australis* (Cav.) Trin. ex Steud. (= *P. communis* Trin.). *Bulletin De La Société Botanique De France*, *123*(5–6), 261–271. https://doi.org/10.1080/00378941.1976.10835694

Grass Phylogeny Working Group (2001). Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden*, *88*(3), 373–457. https://doi.org/10.2307/3298585

Great Lake Restoration Initiative (2019). Great Lake Restoration Initiative (GLRI) Action Plans. https://www.glri.us/documents#actionplan

Guo, W.-Y., Lambertini, C., Nguyen, L. X., Li, X.-Z., & Brix, H. (2014). Preadaptation and post-introduction evolution facilitate the invasion of *Phragmites australis* in North America. *Ecology and Evolution*, *4*(24), 4567–4577. https://doi.org/10.1002/ece3.1286

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B. O., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., … Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hardion, L., Verlaque, R., Haan-Archipoff, G., Cahen, D., Hoff, M., & Vila, B. (2017). Cleaning up the grasses dustbin: systematics of the Arundinoideae subfamily (Poaceae). *Plant Systematics and Evolution*, *303*(10), 1331–1339. https://doi.org/10.1007/s00606-017-1451-6

Harvey-Samuel, T., Ant, T., & Alphey, L. (2017). Towards the genetic control of invasive species. *Biological Invasions*, *19*(6), 1683–1703. https://doi.org/10.1007/s10530-017-1384-6

Hazelton, E. L. G., Mozdzer, T. J., Burdick, D. M., Kettenring, K. M., & Whigham, D. F. (2014). *Phragmites australis* management in the United States: 40 years of methods and outcomes. *AoB Plants*, *6*. https://doi.org/10.1093/aobpla/plu001

Hierro, J. L., Maron, J. L., & Callaway, R. M. (2005). A biogeographical approach to plant invasions: the importance of studying exotics in their introduced and native range. *Journal of Ecology*, *93*(1), 5–15. https://doi.org/10.1111/j.0022-0477.2004.00953.x

Hodgins, K. A., Lai, Z., Nurkowski, K., Huang, J., & Rieseberg, L. H. (2013). The molecular basis of invasiveness: differences in gene expression of native and introduced common ragweed (*Ambrosia artemisiifolia*) in stressful and benign environments. *Molecular Ecology*, *22*(9), 2496–2510. https://doi.org/10.1111/mec.12179

Kalisz, S., Kivlin, S. N., & Bialic-Murphy, L. (2020). Allelopathy is pervasive in invasive plants. *Biological Invasions*, https://doi.org/10.1007/s10530-020-02383-6

Keane, R. M., & Crawley, M. J. (2002). Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution*, *17*(4), 164–170. https://doi.org/10.1016/S0169-5347(02)02499-0

Keller, B. E. M. (2000). Genetic variation among and within populations of *Phragmites australis* in the Charles River watershed. *Aquatic Botany*, *66*(3), 195–208. https://doi.org/10.1016/S0304-3770(99)00077-7

Kettenring, K. M., McCormick, M. K., Baron, H. M., & Whigham, D. F. (2011). Mechanisms of *Phragmites australis* invasion: feedbacks among genetic diversity, nutrients, and sexual reproduction. *Journal of Applied Ecology*, *48*(5), 1305–1313. https://doi.org/10.1111/j.1365-2664.2011.02024.x

Kettenring, K. M., & Mock, K. E. (2012). Genetic diversity, reproductive mode, and dispersal differ between the cryptic invader, *Phragmites australis*, and its native conspecific. *Biological Invasions*, *14*(12), 2489–2504. https://doi.org/10.1007/s10530-012-0246-5

Kiviat, E. (2013). Ecosystem services of *Phragmites* in North America with emphasis on habitat functions. *AoB Plants*, *5*, plt008.https://doi.org/10.1093/aobpla/plt008

Kolar, C. S., & Lodge, D. M. (2001). Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution*, *16*(4), 199–204. https://doi.org/10.1016/s0169-5347(01)02101-2

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

Kowalski, K. P., Bacon, C., Bickford, W., Braun, H., Clay, K., Leduc-Lapierre, M., Lillard, E., McCormick, M. K., Nelson, E., Torres, M., White, J., & Wilcox, D. A. (2015). Advancing the science of microbial symbiosis to support invasive species management: a case study on *Phragmites* in the Great Lakes. *Frontiers in Microbiology*, *6*, 95. https://doi.org/10.3389/fmicb.2015.00095

Lambert, A. M., Winiarski, K., & Casagrande, R. A. (2007). Distribution and impact of exotic gall flies (*Lipara* sp.) on native and exotic *Phragmites australis*. *Aquatic Botany*, *86*(2), 163–170. https://doi.org/10.1016/j.aquabot.2006.09.017

Lambertini, C., Gustafsson, M. H. G., 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*(3), 161–182. https://doi.org/10.1007/s00606-006-0412-2

Lavergne, S., & Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(10), 3883–3888. https://doi.org/10.1073/pnas.0607324104

Li, Z., McKibben, M. T. W., Finch, G. S., Blischak, P. D., Sutherland, B. L., & Barker, M. S. (2021). Patterns and processes of diploidization in land plants. *Annual Review of Plant Biology*, *72*, 387–410. https://doi.org/10.1146/annurev-arplant-050718-100344

Liu, B. O., Yan, J., Li, W., Yin, L., Li, P., Yu, H., Xing, L., Cai, M., Wang, H., Zhao, M., Zheng, J., Sun, F., Wang, Z., Jiang, Z., Ou, Q., Li, S., Qu, L. U., Zhang, Q., Zheng, Y., … Wan, F. (2020). *Mikania micrantha* genome provides insights into the molecular mechanism of rapid growth. *Nature Communications*, *11*(1), 340. https://doi.org/10.1038/s41467-019-13926-4

Liu, L.-L., Yin, M.-Q., Guo, X., Wang, J.-W., Cai, Y.-F., Wang, C., Yu, X.-N., Du, N., Brix, H., Eller, F., Lambertini, C., & Guo, W.-H. (2020). Cryptic lineages and potential introgression in a mixed-ploidy species (*Phragmites australis*) across temperate China. *Journal of Systematics and Evolution*. https://doi.org/10.1111/jse.12672

Lockwood, B. L., & Somero, G. N. (2011). Transcriptomic responses to salinity stress in invasive and native blue mussels (genus *Mytilus*). *Molecular Ecology*, *20*(3), 517–529. https://doi.org/10.1111/j.1365-294X.2010.04973.x

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), 550. https://doi.org/10.1186/s13059-014-0550-8

Lyons, E., & Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. *The Plant Journal*, *53*(4), 661–673. https://doi.org/10.1111/j.1365-313X.2007.03326.x

Ma, X.-F., & Gustafson, J. P. (2005). Genome evolution of allopolyploids: a process of cytological and genetic diploidization. *Cytogenetic and Genome Research*, *109*(1–3), 236–249. https://doi.org/10.1159/000082406

Maere, S., Heymans, K., & Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, *21*(16), 3448–3449. https://doi.org/10.1093/bioinformatics/bti551

Mamta, B., & Rajam, M. V. (2017). RNAi technology: a new platform for crop pest control. *Physiology and Molecular Biology of Plants*, *23*(3), 487–501. https://doi.org/10.1007/s12298-017-0443-x

Mangla, S., & Callaway, R. M. (2008). Exotic invasive plant accumulates native soil pathogens which inhibit native plants. *Journal of Ecology*, *96*, 58–67. https://doi.org/10.1111/j.1365-2745.2007.01312.x

Martin, L. J., & Blossey, B. (2013). The runaway weed: Costs and failures of *Phragmites australis* management in the USA. *Estuaries and Coasts*, *36*(3), 626–632. https://doi.org/10.1007/s12237-013-9593-4

McKain, M. R., Tang, H., McNeal, J. R., Ayyampalayam, S., Davis, J. I., dePamphilis, C. W., Givnish, T. J., Pires, J. C., Stevenson, D. W., & Leebens-Mack, J. H. (2016). A phylogenomic assessment of ancient polyploidy and genome evolution across the Poales. *Genome Biology and Evolution*, *8*(4), 1150–1164. https://doi.org/10.1093/gbe/evw060

Meyerson, L. A., Cronin, J. T., & Pyšek, P. (2016). *Phragmites australis* as a model organism for studying plant invasions. *Biological Invasions*, *18*(9), 2421–2431. https://doi.org/10.1007/s10530-016-1132-3

Mitchell, C. E., & Power, A. G. (2003). Release of invasive plants from fungal and viral pathogens. *Nature*, *421*(6923), 625–627. https://doi.org/10.1038/nature01317

Mounger, J., Ainouche, M. L., Bossdorf, O., Cavé-Radet, A., Li, B., Parepa, M., Salmon, A., Yang, J., & Richards, C. L. (2021). Epigenetics and the success of invasive plants. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *376*(1826), 20200117. https://doi.org/10.1098/rstb.2020.0117

Moura, R. F., Queiroga, D., Vilela, E., & Moraes, A. P. (2021). Polyploidy and high environmental tolerance increase the invasive success of plants. *Journal of Plant Research*, *134*(1), 105–114. https://doi.org/10.1007/s10265-020-01236-6

Mozdzer, T. J., Brisson, J., & Hazelton, E. L. G. (2013). Physiological ecology and functional traits of North American native and Eurasian

introduced *Phragmites australis* lineages. *AoB Plants*, *5*, https://doi.org/10.1093/aobpla/plt048

Nelson, A. D. L., Haug-Baltzell, A. K., Davey, S., Gregory, B. D., & Lyons, E. (2018). EPIC-CoGe: managing and analyzing genomic data. *Bioinformatics*, *34*(15), 2651–2653. https://doi.org/10.1093/bioinformatics/bty106

Oh, D.-H., & Dassanayake, M. (2019). Landscape of gene transposition-duplication within the Brassicaceae family. *DNA Research*, *26*(1), 21–36. https://doi.org/10.1093/dnares/dsy035

Park, M. G., & Blossey, B. (2008). Importance of plant traits and herbivory for invasiveness of *Phragmites australis* (Poaceae). *American Journal of Botany*, *95*(12), 1557–1568. https://doi.org/10.3732/ajb.0800023

Peng, Y., Lai, Z., Lane, T., Nageswara-Rao, M., Okada, M., Jasieniuk, M., O'Geen, H., Kim, R. W., Sammons, R. D., Rieseberg, L. H., & Stewart, C. N. Jr (2014). De novo genome assembly of the economically important weed horseweed using integrated data from multiple sequencing platforms. *Plant Physiology*, *166*(3), 1241–1254. https://doi.org/10.1104/pp.114.247668

Perez, A., Mazerolle, M. J., & Brisson, J. (2013). Effects of exotic common reed (*Phragmites australis*) on wood frog (*Lithobates sylvaticus*) tadpole development and food availability. *Journal of Freshwater Ecology*, *28*(2), 165–177. https://doi.org/10.1080/02705060.2012.750629

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., & Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols*, *11*(9), 1650–1667. https://doi.org/10.1038/nprot.2016.095

Pertea, M., Pertea, 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. https://doi.org/10.1038/nbt.3122

Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, *52*(3), 273–288. https://doi.org/10.1016/j.ecolecon.2004.10.002

Pyšek, P., Čuda, J., Šmilauer, P., Skálová, H., Chumová, Z., Lambertini, C., Lučanová, M., Ryšavá, H., Trávníček, P., Šemberová, K., & Meyerson, L. A. (2020). Competition among native and invasive *Phragmites australis* populations: An experimental test of the effects of invasion status, genome size, and ploidy level. *Ecology and Evolution*, *10*(3), 1106–1118. https://doi.org/10.1002/ece3.5907

Raicu, P., Staicu, S., Stoian, V., & Roman, T. (1972). The *Phragmites communis* Trin. chromosome complement in the Danube Delta. *Hydrobiologia*, *39*(1), 83–89. https://doi.org/10.1007/BF00047596

Rogalski, M. A., & Skelly, D. K. (2012). Positive effects of nonnative invasive *Phragmites australis* on larval bullfrogs. *PLoS One*, *7*(8), e44420. https://doi.org/10.1371/journal.pone.0044420

Rooth, J. E., & Stevenson, J. C. (2000). Sediment deposition patterns in *Phragmites australis* communities: Implications for coastal areas threatened by rising sea-level. *Wetlands Ecology and Management*, *8*(2), 173–183. https://doi.org/10.1023/A:1008444502859

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 of the United States of America*, *99*(4), 2445–2449. https://doi.org/10.1073/pnas.032477999

Saltonstall, K. (2003). Microsatellite variation within and among North American lineages of *Phragmites australis*. *Molecular Ecology*, *12*(7), 1689–1702. https://doi.org/10.1046/j.1365-294x.2003.01849.x

Saltonstall, K. (2016). The naming of *Phragmites* haplotypes. *Biological Invasions*, *18*(9), 2433–2441. https://doi.org/10.1007/s10530-016-1192-4

Saltonstall, K., Peterson, P. M., & Soreng, R. J. (2004). Recognition of *Phragmites australis* subsp. *americanus* (Poaceae: Arundinoideae) in North America: Evidence from morphological and genetic analyses. *SIDA, Contributions to Botany*, *21*(2), 683–692.

Schroeder, C. S., Halbrook, S., Birnbaum, C., Waryszak, P., Wilber, W., & Farrer, E. C. (2020). *Phragmites australis* associates with belowground fungal communities characterized by high diversity and pathogen abundance. *Diversity*, *12*(9), 363. https://doi.org/10.3390/d12090363

Sherman, C. D. H., Lotterhos, K. E., Richardson, M. F., Tepolt, C. K., Rollins, L. A., Palumbi, S. R., & Miller, A. D. (2016). What are we missing about marine invasions? Filling in the gaps with evolutionary genomics. *Marine Biology*, *163*(10), 198. https://doi.org/10.1007/s00227-016-2961-4

Soltis, P. S., & Soltis, D. E. (2016). Ancient WGD events as drivers of key innovations in angiosperms. *Current Opinion in Plant Biology*, *30*, 159–165. https://doi.org/10.1016/j.pbi.2016.03.015

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, *22*(21), 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Steinegger, M., & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, *35*(11), 1026–1028. https://doi.org/10.1038/nbt.3988

Stern, D. B., & Lee, C. E. (2020). Evolutionary origins of genomic adaptations in an invasive copepod. *Nature Ecology & Evolution*, *4*(8), 1084–1094. https://doi.org/10.1038/s41559-020-1201-y

Stohlgren, T. J., Chong, G. W., Schell, L. D., Rimar, K. A., Otsuki, Y., Lee, M., Kalkhan, M. A., & Villa, C. A. (2002). Assessing vulnerability to invasion by nonnative plant species at multiple spatial scales. *Environmental Management*, *29*(4), 566–577. https://doi.org/10.1007/s00267-001-0006-2

Tang, H., Bowers, J. E., Wang, X., Ming, R., Alam, M., & Paterson, A. H. (2008). Synteny and collinearity in plant genomes. *Science*, *320*(5875), 486–488. https://doi.org/10.1126/science.1153917

Tayalé, A., & Parisod, C. (2013). Natural pathways to polyploidy in plants and consequences for genome reorganization. *Cytogenetic and Genome Research*, *140*(2–4), 79–96. https://doi.org/10.1159/000351318

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. https://doi.org/10.1093/aob/mcr277

Tulbure, M. G., & Johnston, C. A. (2010). Environmental conditions promoting non-native *Phragmites australis* expansion in Great Lakes coastal wetlands. *Wetlands*, *30*(3), 577–587. https://doi.org/10.1007/s13157-010-0054-6

Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A., Van de Peer, Y., Coppens, F., & Vandepoele, K. (2018). PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Research*, *46*(D1), D1190–D1196. https://doi.org/10.1093/nar/gkx1002

Wang, 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. https://doi.org/10.3389/fpls.2021.653183

Wang, G., DiTusa, S. F., Oh, D.-H., Herrmann, A. D., Mendoza-Cozatl, D. G., O'Neill, M. A., Smith, A. P., & Dassanayake, M. (2021). Cross species multi-omics reveals cell wall sequestration and elevated global transcript abundance as mechanisms of boron tolerance in plants. *New Phytologist*, *230*(5), 1985–2000. https://doi.org/10.1111/nph.17295

Wang, G., Oh, D.-H., & Dassanayake, M. (2020). GOMCL: a toolkit to cluster, evaluate, and extract non-redundant associations of Gene Ontology-based functions. *BMC Bioinformatics*, *21*(1), 139. https://doi.org/10.1186/s12859-020-3447-4

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.-H., Bancroft, I., Cheng, F., Huang, S., Li, X., Hua, W., Wang, J., Wang, X., Freeling, M., Pires, J. C., Paterson, A. H., Chalhoub, B. … Brassica

rapa Genome Sequencing Project Consortium (2011). The genome of the mesopolyploid crop species Brassica rapa. *Nature Genetics*, *43*(10), 1035–1039. https://doi.org/10.1038/ng.919

Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E. V., & Zdobnov, E. M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*, *35*(3), 543–548. https://doi.org/10.1093/molbev/msx319

Whitaker, K., Rogers, K., Saintilan, N., Mazumder, D., Wen, L., & Morrison, R. J. (2015). Vegetation persistence and carbon storage: Implications for environmental water management for *Phragmites australis*. *Water Resources Research*, *51*(7), 5284–5300. https://doi.org/10.1002/2014WR016253

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences: CABIOS*, *13*(5), 555–556. https://doi.org/10.1093/bioinformatics/13.5.555

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Oh, D.-H., Kowalski, K. P., Quach, Q. N., Wijesinghege, C., Tanford, P., Dassanayake, M., & Clay, K. (2022). Novel genome characteristics contribute to the invasiveness of *Phragmites australis* (common reed). *Molecular Ecology*, 31, 1142–1159. https://doi.org/10.1111/mec.16293