Development of genomic resources for cattails (*Typha*), a globally important macrophyte genus

Running title: Development of genomic resources for *Typha*

Alberto Aleman[1*], Marcel E. Dorken[1,2], Aaron B. A. Shafer[1,3], Tulsi Patel[2], Polina A. Volkova[4], Joanna R. Freeland[1,2*]

[1]Environmental and Life Sciences Graduate Program, Trent University, Peterborough, Ontario, Canada

[2]Department of Biology, Trent University, Peterborough, Ontario, Canada

[3]Department of Forensic Sciences, Trent University, Peterborough, Ontario, Canada

[4]Papanin Institute for Biology of Inland Waters, Russian Academy of Sciences, Borok, Nekouz District, Yaroslavl Region, Russia

***Corresponding authors:**

Joanna Freeland: joannafreeland@trentu.ca

Alberto Aleman: albertolopezaleman@trentu.ca

Abstract

1. Background: A critical knowledge gap in freshwater plants research is the lack of genetic tools necessary to answer fundamental questions about their demographic histories, adaptation, and phylogenetic relationships. One example of this is *Typha*, a global genus of freshwater plants foundational to wetlands that is also becoming an increasingly problematic biological invader in numerous regions worldwide; while important insights have been discovered for this genus, existing markers are insufficient to answer fundamental questions about their demographic histories, adaptation, and phylogenetic relationships, to identify introduced and hybrid lineages, and to examine patterns of hybridisation and introgression.

2. Methods: We optimised a library preparation and data processing protocol to develop genome–wide nuclear and plastid resources for studying the evolutionary history, genetic structure and diversity, hybridisation, local adaptation, invasiveness, and geographic expansion dynamics of *Typha*.

3. Main results: We sequenced 140 *Typha* samples and identified ~120K nuclear SNPs that differentiate *T. angustifolia*, *T. domingensis* and *T. latifolia* and retrieved their plastome sequences. We observed genetic introgression among the three species.

4. Conclusions: Following a fast, straightforward, and cost–efficient genomic library preparation protocol, we produced a suite of genome–wide resources to facilitate investigations into the taxonomy and population genetics of *Typha* and to advance the genomic understanding of wetland plants.

5. Contributions: The protocol described, the updated chromosome–level genome assembly of *T. latifolia*, the catalogue of species-specific SNPs, and the chloroplast sequences produced in this study comprise permanent resources that can be applied to study the genetic

composition of multiple populations and hybrid zones and will be incorporated into future studies of *Typha*, an ecologically important and globally invasive macrophyte.

Keywords: Chloroplast–genome assembly, diagnostic markers, hybridisation, introgression, *Typha* phylogenomics.

Introduction

Freshwater plants are essential to aquatic ecosystems, shaping their habitats' structure and ecological functions (Chambers et al., 2008; Christie et al., 2009; Rejmankova, 2011). Although freshwater plants have been increasingly incorporated into applications that include habitat restoration and invasive species management, they remain highly understudied compared to terrestrial plants (Evangelista et al., 2014; Iversen et al., 2022). One key knowledge gap in freshwater plants research is the lack of genetic tools necessary to answer fundamental questions about their demographic histories, adaptation, and phylogenetic relationships (Fay et al., 2019; Maréchal, 2019; O'Hare et al., 2018; Yannelli et al., 2022).

Genetic characterisation of freshwater plants has been hampered by biological and technical challenges, plus biases in scientific research (Matheson & McGaughran, 2022; Troudet et al., 2017). Hundreds to thousands of molecular markers are often required for genomic–based research on topics such as gene flow and adaptation of non–model organisms (da Fonseca et al., 2016; Stapley et al., 2010), but the *de novo* development of genomic resources can be both time–consuming and expensive (Hu et al., 2020; Ortega et al., 2020; Prieto et al., 2021). Overcoming these challenges is now feasible using novel, rapid and cost–effective methods that capture genome–wide genetic variation, allowing researchers to address questions related to taxonomy, evolution, and conservation (Andrews et al., 2016; Goodwin et al., 2016).

Typha L. (cattails) is a global genus of rhizomatous perennial, monoecious, self-compatible, and wind-pollinated freshwater plants foundational to wetlands (reviewed in Bansal et al., 2019). Cattails are a valuable ecosystem resource and play a fundamental ecological role by cycling nutrients, preventing erosion, maintaining stable water levels and providing food and shelter for wildlife (Andrews & Pratt, 1978; Bonanno & Cirelli, 2017;

Dieye et al., 2017; Kimmerer, 2013; Svedarsky et al., 2019). One major challenge in *Typha* research has been taxonomic identification, which cannot be fully accomplished using morphological characters due to their high intraspecific variability and interspecific hybridisation. Consequently, the richness of cattail species, their taxonomy, provenance (i.e., alien or native lineages), and phylogenetic relationships remain unsolved (Ciotir & Freeland, 2016; Volkova & Bobrov, 2022; Zhou et al., 2018). A refined *Typha* taxonomy along with species-specific genetic markers are necessary to identify introduced and hybrid lineages, which are increasingly documented as invasive, e.g., *T. domingensis* Pers. in Central America, *T.* × *glauca* Godr. (*T. angustifolia* L. × *T. latifolia* L.) in North America, and *T. latifolia* L. in Oceania and Western Europe, (Bansal et al., 2019; *GISD*, n.d.; Govaerts, 2004; Hall, 2009; Maldonado, 2019; Xu et al., 2013).

High–throughput sequencing technologies, novel, cost and time–accessible genome library preparations, and the recent assembly of the *T. latifolia* genome (287.19 Mb) (Goodwin et al., 2016; Rowan et al., 2019; Widanagama et al., 2022) collectively present an opportunity to develop a suite of genomic resources for *Typha*. In addition to taxonomic resolution, these resources will facilitate investigations of the evolutionary history, genetic structure and diversity, hybridisation and introgression, local adaptation, invasiveness and geographical expansion of this genus. We applied a high–throughput sequencing protocol for enzymatic fragmentation, library preparation, and data processing and produced genome–wide resources for three *Typha* spp. By optimising the method from Rowan et al. (2019), we generated a catalogue of nuclear SNPs that characterise *T. angustifolia*, *T. domingensis*, and *T. latifolia* that include chloroplast–genome sequences in a fast, straightforward, and cost–efficient manner.

Materials and methods

Reference genome

We used the reference–based scaffolder Chromosomer 0.1.4a (Tamazian et al., 2016) to align the 1158 *T. latifolia* scaffolds from Widanagama et al. (2022) (Genbank accession: JAIOKV000000000.1) with the *T. latifolia* isolate L0001 (15 chromosomes, GenBank accession: JAAWWQ000000000.1) and produce a local chromosome–level *T. latifolia* genome. Widanagama et al. (2022) had a significantly higher mapping success of unrelated re–sequenced *Typha* spp. compared to the isolate L0001, suggesting it is a more representative *Typha* genome assembly. The scaffolds were aligned as chromosomes using BLAST+ 2.12.0 (Camacho et al., 2009) with the software default settings, and the alignments were anchored in Chromosomer, establishing a gap length = 0 and a ratio threshold = 1.

Sampling, DNA extraction, and sequencing

Samples were either obtained from previous studies or collected across Eurasia, and DNA was extracted at Trent University following published protocols (Bhargav et al., 2022; Ciotir et al., 2017; Pieper et al., 2020, 2017; Tangen et al., 2022; Tisshaw et al., 2020). Briefly, leaf tissue was dried in desiccant silica beads and stored at -20°C. Dried leaf material was ground with a Retsch® MM300 mixer mill (Haan, Germany). DNA was extracted from 25 - 30 mg of semi-fine powder of each sample using the EZNA Plant DNA kit (Omega-Bio Tek) or the Fastpure plant DNA isolation mini kit (Nanjing Vazyme Biotech-China) protocols for dried material, with a final elution of 100 μL. We obtained DNA from 38 *T. angustifolia*, 25 *T. domingensis*, and 77 *T. latifolia* samples (n = 140) (Figure 1; Supplementary Table S1) previously identified to taxon using a combination of genetic analyses of microsatellite loci (Kirk et al., 2011; Snow et al., 2010) and morphological characteristics (Grace & Harrison, 1986; Smith, 1967). Extracted DNA was quantified using a Qubit fluorometer (Thermofisher

Scientific) and calculated as the mean of three independent readings for each sample. All samples were either standardised to 2 ng/μL by dilution with nuclease–free water or left undiluted if at concentrations less than 2 ng/μL (0.4 – 1.9 ng/μL).

Rowan et al. (2019) reported a relatively rapid and cost-effective library preparation technique for genomic sequencing by enzymatic fragmentation followed by ligation of short adapter sequences (i.e., tagmentation) using transposases (Nextera XT) and relatively low yields of DNA. Our protocol was based on this method with a few modifications. First, each DNA sample was tagmented with the Illumina Tagment DNA enzyme (TD) and buffer kit (small kit, #20034210). As the ratio of TD enzyme to DNA is crucial for the reaction, we initially followed Rowan et al. (2019) and subsequently optimised the reagent volumes for our library preparation as 5.5 μL of 5× TD buffer, 0.5 μL of 1× TD enzyme and 4 μL of DNA (standardised or undiluted), keeping all reagents on ice during the preparation. Samples were incubated at 55°C for 10 minutes and left at room temperature for five minutes, and 5 μL of each sample were run on an agarose gel to confirm the efficacy of the tagmentation reaction, evidenced by visible smears. The tagmented DNA was then amplified using unique dual indexing based on combinations from a total of 24 N7 (47 bases) and 8 S5 (51 bases) adapters (Alpha DNA, Canada). The PCR cocktail included 0.2 μM of each index, 0.5U of KAPA HiFi HotStart DNA polymerase (Roche), 12.5 μL of 5× KAPA reagent, 5 μL of tagmented DNA, and 6.5 μL of nuclease–free water to a final volume of 25 μL. The PCR cycle comprised 72°C (3 minutes); 95°C (30 seconds); and 14 cycles of 95°C (10 seconds), 55°C (30 seconds), and 72°C (30 seconds). Once again, visible smears confirmed amplification success after running 5 μL of the PCR product on an agarose gel; then, 10 μL of each sample were pooled, and the remaining PCR products were stored at -20°C. The pooled library was purified with a QIAquick PCR purification kit (QIAGEN) following the manufacturer's protocol, with a final

elution in 50 μL of elution buffer. The library was quantified using a D1000 Tapestation assay (Agilent Technologies, USA) and a Qubit fluorometer (Thermo Fisher Scientific). A quality–control paired–end sequencing was executed using a Miseq (151 bp) to ensure the genomic library was compiled successfully. Finally, paired–end sequencing was performed on a Novaseq 6000 (126 bp) at The Centre for Applied Genomics (Toronto, Ontario).

Raw data processing, filtering, and SNP–calling

The quality of the demultiplexed raw sequences was evaluated using FastQC 0.11.9 (Andrews, 2017) and MultiQC 1.14 (Ewels et al., 2016). Read pairing and adapter pruning were carried out with trimmomatic 0.39 (Bolger et al., 2014), removing any cleaned reads shorter than 100 bp. Paired and remaining unpaired reads were mapped to our chromosome–level *T. latifolia* nuclear genome plus the *T. latifolia* plastome reference (Genbank accession: NC_013823.1) using the mem module of BWA 0.7.17 (Li & Durbin, 2009). Mapped reads from Miseq and Novaseq 6000 sequencers were merged, and mapping statistics were evaluated with the flagstat and coverage modules of SAMtools 1.15.1 (Li et al., 2009).

Genotype-calling was performed with ANGSD 0.93 (Korneliussen et al., 2014) following the SAMtools model, retrieving hard-called SNPs with a minimum *p*–value of $1e^{-6}$, minimum mapping and sequencing qualities of 20, discarding any indels and triallelic sites, and outputting a binary Variant Call Format file (*-doGeno 4 -gl 1 -skipTriallelic 1 -SNP_pval 1e-6 -minMapQ 20 -minQ 20 -doMajorMinor 1 -domaf 1 -doPost 1* - doBcf 1).

For the nuclear analyses, SNPs with > 50% missing data across all samples and sites mapped to the plastome were removed in VCFtools 0.1.16 (Danecek et al., 2011). We did not apply any additional filters to SNP identification: our samples represent a broad geographical sampling (Figure 1; Supplementary Table S1) and thus were not expected to be in Hardy–Weinberg equilibrium; additionally, as allele frequencies were unlikely to be representative of

regional allele frequencies, we did not apply a minor allele frequency filter; neither did we filter for linkage equilibrium, as eliminating alleles that are in linkage disequilibrium is likely to decrease the resolution to detect hybridisation and introgression (Alexander, 2020; Pearman et al., 2022).

Genetic structure and diagnostic markers

We used nuclear SNPs to assess the most likely number of genetic clusters across all samples and the membership of each plant to these clusters using three complementary approaches: i) ADMIXTURE 1.3.0 (Alexander & Lange, 2011) was run with $K = 1 – 10$, and the optimal number of clusters was chosen via the cross–validation procedure, ii) a neighbour–joining tree from the samples' pairwise genetic distance matrix (expressed as allele counts, transformed on the R 4.2.1 package ape 5.7-1 (Paradis et al., 2004; R Core Team, 2022)), and iii) a Principal Component Analysis (PCA) were performed with Plink 1.90 (Purcell et al., 2007). To avoid over- or under-estimating genetic structure (Janes et al., 2017), we verified that the assignment of samples to genetic clusters (corresponding to three species, see *Results*) was consistent for each approach.

Potential introgression was tested by running ADMIXTURE ($K = 1 – 5$) on three datasets, each comprising a combination of two genetic clusters, using only those SNPs that remained variable based on the two species being compared. We confirmed that the cross–validation procedure for the runs of each species' pair chose the optimal number of clusters as two ($K = 2$) and used the admixture proportion (*Q score*) as an index of potential introgression of each sample. Applying Senn & Pemberton (2009) and Smith et al. (2018) thresholds, individuals whose Q score was $0.05 \leq Q \leq 0.95$, were considered as genetically introgressed.

To compare the levels of differentiation between clusters, values of Weir and Cockerham's genetic differentiation (F_{ST}; Weir & Cockerham, 1984) and genetic divergence

(d_{xy}; Nei & Miller, 1990) for every variable site in 10 Kbp windows between species pairs were computed with pixy 1.2.7 (Korunes & Samuk, 2021), and the means were calculated. Species–specific SNPs were identified i) for each species' pair and ii) by running three paired comparisons of one species versus the other two on each run, using DiagnoSNPs 1.0 (Arce-Valdés, 2022). We removed genetically introgressed individuals before estimating levels of genetic differentiation and identifying species–specific SNPs.

Chloroplast genome reconstruction and phylogenetic analysis

We implemented a reference–guided workflow to reconstruct whole–chloroplast–genome sequences. Nucleotide calling was performed individually for each of the 140 samples in ANGSD, using the reads that mapped to the plastome reference, requiring minimum mapping and base qualities of 20, and using N's for missing data (*-dofasta 2 -minMapQ 20 -minQ 20 -doCounts 1*). The sequences were aligned to the chloroplast genomes of *T. przewalskii* Skvortsov, *T. lugdunensis* P. Chabert, *T. orientalis* C. Presl, and *Sparganium natans* (GenBank accessions: NC_061354.1, NC_061353.1, NC_050678.1, and NC_058577.1), following Smith et al. (2021) by applying MAFFT 7.0 default settings (i.e., with the flag *--nzero*) (Katoh et al., 2019). Snp–sites 2.5.1 (Page et al., 2016) was used to remove regions of the genome with ambiguous positions, gaps, and missing data, such that if any of those was found in a sequence, that position was removed for all sequences. Nucleotide diversity (π) was calculated in the R package pegas (Paradis, 2010).

The phylogenetic relationships of the chloroplast genome sequences were reconstructed in RAxML-NG 1.1 (Kozlov et al., 2019). Model selection was based on jmodeltest 2.1.10 results (Darriba et al., 2012) using the Akaike Information Criterion. RAxML was run under a TVM + G4 + I model with the automatic and thorough bootstrap

options, starting from 100 random trees and employing *Sparganium natans* as the outgroup. The best–scoring ML tree was visualized.

Results

Genome scaffolding, mapping statistics, and genotyping

Approximately 99.81% of the scaffold sequences were aligned to the template genome. The scaffolds were anchored to 15 chromosomes producing a genome of 285.11 Mb (GenBank accession: JAIOKV000000000.2). The total assembled size was comparable to the *T. latifolia* genome sizes of Widanagama et al. (2022) (287.19 Mb) and the isolate L0001 (214.13 Mb). Updating the chromosome–level genome assembly simplified our downstream analyses while keeping the highest mapping success of unrelated re–sequenced *Typha* spp., and facilitating an accurate recombination map for future studies of speciation, hybridisation, and the genomic landscape of introgression in *Typha*.

After quality control, 982 M clean paired–end reads were retained, and ~98% mapped to the reference genome. With minimum mapping and sequencing qualities = 20, the average depth and breadth of coverage for the nuclear sequences were 4× and 42%, respectively. Over 60% of the plastome breadth was covered across all samples (mean depth = 711×), enabling us to use 96,591 bp for the phylogenetic reconstruction. We assembled 12,177,703 bi–allelic nuclear SNPs across the 140 *Typha* samples (7,122,151 with a MAF > 0.05). The total genotyping rate, i.e., the mean proportion of samples with data for each SNP, was 0.68.

Genetic structure and diagnostic markers

The admixture analysis, the PCA, and the neighbour–joining tree each established the most likely number of genetic clusters as three (K = 3) (Figure 1): in line with previous taxonomic identifications, 38 samples were identified within the *T. angustifolia* cluster (15 of which had *T. latifolia* introgression and 5 of which had both *T. domingensis* and *T. latifolia*

introgression); 25 samples were in the *T. domingensis* cluster (one with *T. angustifolia* introgression, 12 with *T. latifolia* introgression, and 3 with both *T. angustifolia* and *T. latifolia* introgression); and 77 samples were in the *T. latifolia* cluster (one with both *T. angustifolia* and *T. domingensis* introgression, and one with *T. angustifolia* introgression). Using only the 18 *T. angustifolia*, 9 *T. domingensis*, and 75 *T. latifolia* non-introgressed samples, the mean pairwise interspecific F_{ST} and d_{xy} values ranged from 0.25 to 0.49 and 0.28 to 0.35 (Table 1), respectively, with *T. latifolia* showing the highest differentiation with both *T. angustifolia* and *T. domingensis.* We identified 119,324 nuclear species-specific SNPs by pairwise comparisons between the three species and 16,856 SNPs when one species was compared to the other two (*T. angustifolia* = 10,537; *T. domingensis* = 3,838; *T. latifolia* = 3,044).

Chloroplast genome reconstruction and phylogenetic relationships

After removing all ambiguities and missing data from the chloroplast genomes, we were left with an alignment of 96,591 bp across all 140 sequences and the four references, with 4,916 segregating sites and π = 0.003. The phylogenetic reconstruction was congruent with the nuclear genetic structure results (i.e., individuals were consistently assigned to the same nuclear and plastid lineages), grouping the 143 *Typha* samples into three lineages, with *T. angustifolia* in one clade, *T. domingensis* and *T. orientalis* sharing another, and *T. latifolia* and *T. przewalskii* in a third one (Figure 2). All interspecific nodes were strongly supported.

Discussion

We aimed to produce a suite of genome–wide resources to facilitate investigations into the taxonomy and population genetics of *Typha* and to advance the genomic understanding of wetland plants. Following a fast, straightforward, and cost–efficient genomic library preparation protocol (Rowan et al., 2019), we sequenced 140 *Typha* samples, obtaining an average breadth of 42% of the nuclear genome, characterising 119,324 nuclear SNPs that collectively differentiate three *Typha* spp., and producing chloroplast sequences with a breadth of coverage > 60% per sample. With a cost below 15 USD per sample and a processing time of two hours for the library preparation, our workflow is a rapid and cost–effective protocol that can be applied in population genomic studies for investigating levels of genetic diversity and differentiation, identifying conservation units and alien taxa, and investigating hybrid zones, among other purposes. Additionally, our results reveal the feasibility of reconstructing whole-chloroplast-genome sequences as a by-product of the enzymatic fragmentation for high-throughput sequencing libraries, making plastome research simpler and inexpensive for species with an available reference.

Three genetic clusters were identified from both nuclear and chloroplast genomes, corresponding to *T. angustifolia*, *T. domingensis*, and *T. latifolia*, and genetic introgression was detected among the three species. While there are some reports of hybridisation between *T. domingensis* and either *T. angustifolia* or *T. latifolia* (Ciotir et al., 2017; Govaerts, 2004; Smith, 1967), range-wide surveys are lacking. Future research should address the extent to which hybridisation is shaping the genetic differentiation and diversity of these three species. Furthermore, *T. domingensis* is increasingly invading regions in Nigeria (Ringim et al., 2016), Costa Rica (Trama et al., 2017) and North America, potentially expanding its range across the latter (Spencer & Vincent, 2013; Zhang et al., 2008). However, the taxonomic identity of

these plants is unclear –are they hybrids, non–native lineages, native lineages responding to environmental change, or misidentified *T. angustifolia* (Bansal et al., 2019)? By characterising SNPs that differentiate *T. domingensis*, we provide a valuable resource to answer this and other questions across the evolutionary and hybridisation history of *Typha*.

The markers that differentiate *T. angustifolia* from *T. latifolia* will have important applications in North America, where the two species interbreed across a large area and produce an invasive interspecific hybrid (*T.* × *glauca*) that dominates wetlands, alters nutrient cycling, and reduces biodiversity across the Great Lakes Region (Bansal et al., 2019); additionally, this hybrid is expanding throughout the Prairie Pothole Region, causing native plant diversity to decrease in invaded potholes (Jones et al., 2023), and may impact essential habitat for millions of breeding and migratory waterfowl species (Tangen et al., 2022). Until now, molecular resources to characterise *T. angustifolia*, *T. latifolia*, and *T.* × *glauca* were limited to sets of relatively few individual markers that have produced important insights: RAPDs, chloroplast DNA sequences, and codominant SSR loci have contributed to exposing the sexual fertility of first–generation hybrids (F1s) (Snow et al., 2010), asymmetric hybridisation, with *T. angustifolia* being mainly the maternal parent (Ball & Freeland, 2013; Kuehn et al., 1999; Pieper et al., 2017), overall comparable levels of sexual and clonal reproduction in parents and F1s (Pieper et al., 2020; Travis et al., 2011), heterosis in F1s (Bunbury-Blanchette et al., 2015; Travis et al., 2010; Zapfe & Freeland, 2015), a high frequency of F1s in natural populations (Kirk et al., 2011; Travis et al., 2010), the capability of F1s to backcross, plus partial sterility in F1s coupled with hybrid breakdown of F2s and advanced–generation hybrids (Bhargav et al., 2022; Pieper et al., 2017). However, critical inquiries remain unresolved because existing markers are insufficient to expose the prevalence of advanced–generation hybrids and backcrosses in wild populations. The

expansive suite of SNPs identified in this study will facilitate investigations into the extent of hybridisation, hybrid breakdown dynamics, and adaptive introgression across the *T.* × *glauca* hybrid zones, allowing researchers to understand the processes shaping this genus speciation and species boundaries, and to inform conservation and management strategies.

Fundamental genetic tools are essential for investigating freshwater plants' biology, management, and conservation (O'Hare et al., 2018). The protocol described in this paper, the updated chromosome–level genome assembly of *T. latifolia*, the catalogue of species-specific SNPs, and the chloroplast sequences produced for each sample comprise permanent resources that can be applied to study the genetic composition of multiple populations and hybrid zones. Genome–wide sequencing techniques and reference–based chloroplast genome assemblies are promising tools to clarify the demographic histories, dispersal, adaptation, and taxonomy of multiple congeneric macrophyte species (Russello et al., 2015; Straub et al., 2012), and substantial genome–wide research will allow us to tackle these and other knowledge gaps in *Typha* and other freshwater taxa.

Acknowledgements

We acknowledge that the laboratory procedures and data analyses were conducted at Trent University, which is on the traditional territory of the Mississauga Anishinaabeg. The Natural Sciences and Engineering Research Council of Canada (NSERC) financially supported this work, and Alberto Aleman is funded by the Environmental & Life Sciences Graduate Program at Trent University. The work of Polina A. Volkova was supported by the Russian Science Foundation grant no. 23-14-00115. We thank V. Bhargav, N. Tikhomirov, and M. Ivanova for providing plant tissue samples; T. Pimenov, M. Aksyonova and the staff of the Dagestansky Nature Reserve, in particular, G. S. Dzhamirzoyev, for their help in the field, AO "IEPI" for organising fieldwork in Krasnodar Region, and SHARCNET and Compute Canada for providing computational resources. Finally, we thank Camille Kessler for her comments on the manuscript and Enrique Ruiz for his work in Figure 2 (top).

Author Contributions

Conceptualisation: ABAS, JRF, and MED. Developing methods, conducting the research, data interpretation and writing: AA, ABAS, JRF, MED, PAV, and TP. Data analysis and preparation figures & tables: AA.

Data Accessibility Statement

The updated nuclear genome assembly will be submitted to GenBank (JAIOKV000000000.2). Code, species-specific SNP locations (chromosome, position) and chloroplast-genome sequences produced in this study will be available at https://gitlab.com/WiDGeT_TrentU/graduate_theses/-/tree/master/aleman/fwb and https://github.com/al-aleman/totoras_fwb. High-throughput sequencing raw data and Variant Call Format files are available from the authors upon reasonable request.

Conflict of Interest

The authors declare no conflict of interest.

References

Alexander, A. (2020). *GBS_SNP_filter v1.17* [R]. https://github.com/laninsky/GBS_SNP_filter

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, *12*(1), 246. https://doi.org/10.1186/1471-2105-12-246

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), Article 2. https://doi.org/10.1038/nrg.2015.28

Andrews, N., & Pratt, D. (1978). Energy Potential of Cattails (Typha spp.) and Productivity in Managed Stands. *Journal of the Minnesota Academy of Science*, *44*(2), 5–8.

Andrews, S. (2017). *FastQC: A quality control tool for high throughput sequence data. 2010*.

Arce-Valdés, L. R. (2022). *LuisRodrigoArce-Valdes/DiagnoSNPs* [Shell]. https://github.com/LuisRodrigoArce-Valdes/DiagnoSNPs (Original work published 2021)

Ball, D., & Freeland, J. R. (2013). Synchronous flowering times and asymmetrical hybridization in Typha latifolia and T. angustifolia in northeastern North America. *Aquatic Botany*, *104*, 224–227. https://doi.org/10.1016/j.aquabot.2012.08.006

Bansal, S., Lishawa, S. C., Newman, S., Tangen, B. A., Wilcox, D., Albert, D., Anteau, M. J., Chimney, M. J., Cressey, R. L., DeKeyser, E., Elgersma, K. J., Finkelstein, S. A., Freeland, J., Grosshans, R., Klug, P. E., Larkin, D. J., Lawrence, B. A., Linz, G., Marburger, J., … Windham-Myers, L. (2019). Typha (Cattail) Invasion in North American Wetlands: Biology, Regional Problems, Impacts, Ecosystem Services, and Management. *Wetlands*, *39*(4), 645–684. https://doi.org/10.1007/s13157-019-01174-7

Bhargav, V. V., Freeland, J. R., & Dorken, M. E. (2022). Evidence of hybrid breakdown among invasive hybrid cattails (Typha × glauca). *Heredity*, *129*(3), Article 3. https://doi.org/10.1038/s41437-022-00557-7

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bonanno, G., & Cirelli, G. L. (2017). Comparative analysis of element concentrations and translocation in three wetland congener plants: Typha domingensis, Typha latifolia and Typha angustifolia. *Ecotoxicology and Environmental Safety*, *143*, 92–101. https://doi.org/10.1016/j.ecoenv.2017.05.021

Bunbury-Blanchette, A. L., Freeland, J. R., & Dorken, M. E. (2015). Hybrid Typha×glauca outperforms native T. latifolia under contrasting water depths in a common garden. *Basic and Applied Ecology*, *16*(5), 394–402. https://doi.org/10.1016/j.baae.2015.04.006

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Chambers, P. A., Lacoul, P., Murphy, K. J., & Thomaz, S. M. (2008). Global diversity of aquatic macrophytes in freshwater. In E. V. Balian, C. Lévêque, H. Segers, & K. Martens (Eds.), *Freshwater Animal Diversity Assessment* (pp. 9–26). Springer Netherlands. https://doi.org/10.1007/978-1-4020-8259-7_2

Christie, H., Norderhaug, K. M., & Fredriksen, S. (2009). Macrophytes as habitat for fauna. *Marine Ecology Progress Series*, *396*, 221–233. https://doi.org/10.3354/meps08351

Ciotir, C., & Freeland, J. (2016). Cryptic intercontinental dispersal, commercial retailers, and the genetic diversity of native and non-native cattails (Typha spp.) in North America. *Hydrobiologia*, *768*(1), 137–150. https://doi.org/10.1007/s10750-015-2538-0

Ciotir, C., Szabo, J., & Freeland, J. (2017). Genetic characterization of cattail species and hybrids (Typha spp.) in Europe. *Aquatic Botany*, *141*, 51–59. https://doi.org/10.1016/j.aquabot.2017.03.005

da Fonseca, R. R., Albrechtsen, A., Themudo, G. E., Ramos-Madrigal, J., Sibbesen, J. A., Maretty, L., Zepeda-Mendoza, M. L., Campos, P. F., Heller, R., & Pereira, R. J. (2016). Next-generation biology: Sequencing and data analysis approaches for non-model organisms. *Marine Genomics*, *30*, 3–13. https://doi.org/10.1016/j.margen.2016.04.012

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and high-performance computing. *Nature Methods*, *9*(8), 772. https://doi.org/10.1038/nmeth.2109

Dieye, Y., Sambou, V., Faye, M., Thiam, A., Adj, M., & Azilinon, D. (2017). Thermo-mechanical characterization of a building material based on Typha Australis. *Journal of Building Engineering*, *9*, 142–146. https://doi.org/10.1016/j.jobe.2016.12.007

Evangelista, H. B. A., Thomaz, S. M., & Umetsu, C. A. [UNESP. (2014). An analysis of publications on invasive macrophytes in aquatic ecosystems. *Aquatic Invasions*, 521. https://doi.org/10.3391/ai.2014.9.4.10

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Fay, M. F., Gargiulo, R., & Viruel, J. (2019). The present and future for population genetics, species boundaries, biogeography and conservation. *Botanical Journal of the Linnean Society*, *191*(3), 299–304. https://doi.org/10.1093/botlinnean/boz076

GISD. (n.d.). Retrieved March 13, 2023, from http://www.iucngisd.org/gisd/speciesname/Typha+latifolia

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), Article 6. https://doi.org/10.1038/nrg.2016.49

Govaerts, R. (2004). The Monocot Checklist Project. *TAXON*, *53*(1), 144–146. https://doi.org/10.2307/4135499

Grace, J., & Harrison, J. (1986). The biology of Canadian weeds. 73. Typha latifolia L., Typha angustifolia L. and Typha xglauca Godr. *Canadian Journal of Plant Science - CAN J PLANT SCI*, *66*, 361–379. https://doi.org/10.4141/cjps86-051

Hall, S. J. (2009). Cultural Disturbances and Local Ecological Knowledge Mediate Cattail (Typha domingensis) Invasion in Lake Pátzcuaro, México. *Human Ecology*, *37*(2), 241–249. https://doi.org/10.1007/s10745-009-9228-3

Hu, Z.-M., Zhong, K.-L., Weinberger, F., Duan, D.-L., Draisma, S. G. A., & Serrão, E. A. (2020). Linking Ecology to Genetics to Better Understand Adaptation and Evolution: A Review in Marine Macrophytes. *Frontiers in Marine Science*, *7*. https://www.frontiersin.org/articles/10.3389/fmars.2020.545102

Iversen, L. L., Girón, J. G., & Pan, Y. (2022). Towards linking freshwater plants and ecosystems via functional biogeography. *Aquatic Botany*, *176*, 103454. https://doi.org/10.1016/j.aquabot.2021.103454

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K = 2 conundrum. *Molecular Ecology*, *26*(14), 3594–3602. https://doi.org/10.1111/mec.14187

Jones, S. A., DeKeyser, E. S., Dixon, C., & Kobiela, B. (2023). Invasive Species Change Plant Community Composition of Preserved Prairie Pothole Wetlands. *Plants*, *12*(6), Article 6. https://doi.org/10.3390/plants12061281

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, *20*(4), 1160–1166. https://doi.org/10.1093/bib/bbx108

Kimmerer, R. (2013). *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge and the Teachings of Plants*. Milkweed Editions.

Kirk, H., Connolly, C., & Freeland, J. R. (2011). Molecular genetic data reveal hybridization between Typha angustifolia and Typha latifolia across a broad spatial scale in eastern North America. *Aquatic Botany*, *95*(3), 189–193. https://doi.org/10.1016/j.aquabot.2011.05.007

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, *21*(4), 1359–1368. https://doi.org/10.1111/1755-0998.13326

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

Kuehn, M. M., Minor, J. E., & White, B. N. (1999). An examination of hybridization between the cattail species Typha latifolia and Typha angustifolia using random amplified polymorphic DNA and chloroplast DNA markers. *Molecular Ecology*, *8*(12), 1981–1990. https://doi.org/10.1046/j.1365-294x.1999.00792.x

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Maldonado, G. (2019). *The Paradox of Culturally Useful Invasive Species: Chuspatel (Typha Domingensis) Crafts of Lake Patzcuaro, Mexico*.

Maréchal, E. (2019). Marine and Freshwater Plants: Challenges and Expectations. *Frontiers in Plant Science*, *10*, 1545. https://doi.org/10.3389/fpls.2019.01545

Matheson, P., & McGaughran, A. (2022). Genomic data is missing for many highly invasive species, restricting our preparedness for escalating incursion rates. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-17937-y

Nei, M., & Miller, J. C. (1990). A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. *Genetics*, *125*(4), 873–879. https://doi.org/10.1093/genetics/125.4.873

O'Hare, M. T., Aguiar, F. C., Asaeda, T., Bakker, E. S., Chambers, P. A., Clayton, J. S., Elger, A., Ferreira, T. M., Gross, E. M., Gunn, I. D. M., Gurnell, A. M., Hellsten, S., Hofstra, D. E., Li, W., Mohr, S., Puijalon, S., Szoszkiewicz, K., Willby, N. J., & Wood, K. A. (2018). Plants in aquatic ecosystems: Current trends and future directions. *Hydrobiologia*, *812*(1), 1–11. https://doi.org/10.1007/s10750-017-3190-7

Ortega, A., Geraldi, N. R., Díaz-Rúa, R., Ørberg, S. B., Wesselmann, M., Krause-Jensen, D., & Duarte, C. M. (2020). A DNA mini-barcode for marine macrophytes. *Molecular Ecology Resources*, *20*(4), 920–935. https://doi.org/10.1111/1755-0998.13164

Page, A. J., Taylor, B., Delaney, A. J., Soares, J., Seemann, T., Keane, J. A., & Harris, S. R. (2016). SNP-sites: Rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial Genomics*, *2*(4), e000056. https://doi.org/10.1099/mgen.0.000056

Paradis, E. (2010). pegas: An R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, *20*(2), 289–290. https://doi.org/10.1093/bioinformatics/btg412

Pearman, W. S., Urban, L., & Alexander, A. (2022). Commonly used Hardy–Weinberg equilibrium filtering schemes impact population structure inferences using RADseq data. *Molecular Ecology Resources*, *22*(7), 2599–2613. https://doi.org/10.1111/1755-0998.13646

Pieper, S., Dorken, M., & Freeland, J. (2020). Genetic structure in hybrids and progenitors provides insight into processes underlying an invasive cattail (Typha × glauca) hybrid zone. *Heredity*, *124*(6), Article 6. https://doi.org/10.1038/s41437-020-0307-y

Pieper, S. J., Nicholls, A. A., Freeland, J. R., & Dorken, M. E. (2017). Asymmetric Hybridization in Cattails (Typha spp.) and Its Implications for the Evolutionary Maintenance of Native Typha latifolia. *Journal of Heredity*, *108*(5), 479–487. https://doi.org/10.1093/jhered/esx036

Prieto, A. E., Hardion, L., & Beisel, J.-N. (2021). Designing robust DNA barcode libraries for metabarcoding of freshwater plants by integrating herbarium collections and contemporary floristic inventories. *ARPHA Conference Abstracts*, *4*, e64713. https://doi.org/10.3897/aca.4.e64713

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

R Core Team. (2022). *R: The R Project for Statistical Computing*. https://www.r-project.org/

Rejmankova, E. (2011). The role of macrophytes in wetland ecosystems. *Journal of Ecology and Environment*, *34*(4), 333–345. https://doi.org/10.5141/JEFB.2011.044

Ringim, A., Babura, B., & Harry. (2016). Implication of Invasive Plant Typha domingensis on Biodiversity: An Ecological Study of the Hadejia-Nguru Wetlands, Nigeria. *Scholarly Journal of Biological Science*, *4*, 40–46.

Rowan, B. A., Heavens, D., Feuerborn, T. R., Tock, A. J., Henderson, I. R., & Weigel, D. (2019). An Ultra High-Density Arabidopsis thaliana Crossover Map That Refines the Influences of Structural Variation and Epigenetic Features. *Genetics*, *213*(3), 771–787. https://doi.org/10.1534/genetics.119.302406

Russello, M. A., Waterhouse, M. D., Etter, P. D., & Johnson, E. A. (2015). From promise to practice: Pairing non-invasive sampling with genomics in conservation. *PeerJ*, *3*, e1106. https://doi.org/10.7717/peerj.1106

Senn, H. V., & Pemberton, J. M. (2009). Variable extent of hybridization between invasive sika (Cervus nippon) and native red deer (C. elaphus) in a small geographical area. *Molecular Ecology*, *18*(5), 862–876. https://doi.org/10.1111/j.1365-294X.2008.04051.x

Smith, C. I., McKain, M. R., Guimond, A., & Flatz, R. (2021). Genome-scale data resolves the timing of divergence in Joshua trees. *American Journal of Botany*, *108*(4), 647–663. https://doi.org/10.1002/ajb2.1633

Smith, S. G. (1967). Experimental and Natural Hybrids in North American Typha (Typhaceae). *The American Midland Naturalist*, *78*(2), 257–287. https://doi.org/10.2307/2485231

Smith, S. L., Senn, H. V., Pérez-Espona, S., Wyman, M. T., Heap, E., & Pemberton, J. M. (2018). Introgression of exotic Cervus (nippon and canadensis) into red deer (Cervus elaphus) populations in Scotland and the English Lake District. *Ecology and Evolution*, *8*(4), 2122–2134. https://doi.org/10.1002/ece3.3767

Snow, A. A., Travis, S. E., Wildová, R., Fér, T., Sweeney, P. M., Marburger, J. E., Windels, S., Kubátová, B., Goldberg, D. E., & Mutegi, E. (2010). Species-specific SSR alleles for studies of hybrid cattails (Typha latifolia × T. angustifolia; Typhaceae) in North America. *American Journal of Botany*, *97*(12), 2061–2067.

Spencer, J. M., & Vincent, M. A. (2013). Southern cat-tail (Typha domingensis, Typhaceae) discovered in Ohio. *Phytoneuron*, *2013–22*, 1–5.

Stapley, J., Reger, J., Feulner, P. G. D., Smadja, C., Galindo, J., Ekblom, R., Bennison, C., Ball, A. D., Beckerman, A. P., & Slate, J. (2010). Adaptation genomics: The next generation. *Trends in Ecology & Evolution*, *25*(12), 705–712. https://doi.org/10.1016/j.tree.2010.09.002

Straub, S. C. K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., & Liston, A. (2012). Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American Journal of Botany*, *99*(2), 349–364. https://doi.org/10.3732/ajb.1100335

Svedarsky, D., Grosshans, R., Venema, H., Ellis-Felege, S., Bruggman, J., Ostlund, A., & Lewis, J. (2019). Integrated management of invasive cattails (Typha spp.) for wetland habitat and biofuel in the Northern Great Plains of the United States and Canada: A review. *Mires and Peat*, *25*, 1–14. https://doi.org/10.19189/MaP.2018.APG.367

Tamazian, G., Dobrynin, P., Krasheninnikova, K., Komissarov, A., Koepfli, K.-P., & O'Brien, S. J. (2016). Chromosomer: A reference-based genome arrangement tool for producing draft chromosome sequences. *GigaScience*, *5*(1), 38. https://doi.org/10.1186/s13742-016-0141-6

Tangen, B. A., Bansal, S., Freeland, J. R., Travis, S. E., Wasko, J. D., McGonigle, T. P., Goldsborough, L. G., Gow, K., Marburger, J. E., & Meier, J. A. (2022). Distributions of native and invasive Typha (cattail) throughout the Prairie Pothole Region of North America. *Wetlands Ecology and Management*, *30*(1), 1–17. https://doi.org/10.1007/s11273-021-09823-7

Tisshaw, K., Freeland, J., & Dorken, M. (2020). Salinity, not genetic incompatibilities, limits the establishment of the invasive hybrid cattail Typha × glauca in coastal wetlands. *Ecology and Evolution*, *10*(21), 12091–12103. https://doi.org/10.1002/ece3.6831

Trama, F. A., Viale, F. L. S. R. P., Kumar, A. S., Stynoski, J. L., Colton, M. B. M., & Springer, M. C. (2017). The Management of Typha domingensis (Typhaceae) affects Macroinvertebrate Assemblages in the Palo Verde Wetland, Guanacaste, Costa Rica. *Ecological Restoration*, *35*(2), 175–189. https://doi.org/10.3368/er.35.2.175

Travis, S. E., Marburger, J. E., Windels, S. K., & Kubátová, B. (2011). Clonal Structure of Invasive Cattail (Typhaceae) Stands in the Upper Midwest Region of the US. *Wetlands*, *31*(2), 221–228. https://doi.org/10.1007/s13157-010-0142-7

Travis, S. E., Marburger, J. E., Windels, S., & Kubátová, B. (2010). Hybridization dynamics of invasive cattail (Typhaceae) stands in the Western Great Lakes Region of North America: A molecular analysis. *Journal of Ecology*, *98*(1), 7–16. https://doi.org/10.1111/j.1365-2745.2009.01596.x

Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F. (2017). Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-09084-6

Volkova, P. A., & Bobrov, A. A. (2022). Easier than it looks: Notes on the taxonomy of Typha L. (Typhaceae) in East Europe. *Aquatic Botany*, *176*, 103453. https://doi.org/10.1016/j.aquabot.2021.103453

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, *38*(6), 1358–1370. https://doi.org/10.2307/2408641

Widanagama, S. D., Freeland, J. R., Xu, X., & Shafer, A. B. A. (2022). Genome assembly, annotation, and comparative analysis of the cattail Typha latifolia. *G3 Genes|Genomes|Genetics*, *12*(2), jkab401. https://doi.org/10.1093/g3journal/jkab401

Xu, Z., Feng, Z., Yang, J., Zheng, J., & Zhang, F. (2013). Nowhere to Invade: Rumex crispus and Typha latifolia Projected to Disappear under Future Climate Scenarios. *PLOS ONE*, *8*(7), e70728. https://doi.org/10.1371/journal.pone.0070728

Yannelli, F. A., Bazzichetto, M., Conradi, T., Pattison, Z., Andrade, B. O., Anibaba, Q. A., Bonari, G., Chelli, S., Ćuk, M., Damasceno, G., Fantinato, E., Geange, S. R., Guuroh, R. T., Holle, M. J. M., Küzmič, F., Lembrechts, J. J., Mosyaftiani, A., Šikuljak, T., Teixeira, J., … Sperandii, M. G. (2022). Fifteen emerging challenges and opportunities for vegetation science: A horizon scan by early career researchers. *Journal of Vegetation Science*, *33*(1), e13119. https://doi.org/10.1111/jvs.13119

Zapfe, L., & Freeland, J. R. (2015). Heterosis in invasive F1 cattail hybrids (Typha × glauca). *Aquatic Botany*, *125*, 44–47. https://doi.org/10.1016/j.aquabot.2015.05.004

Zhang, X.-H., Tapia, M., Webb, J. B., Huang, Y.-H., & Miao, S. (2008). Molecular signatures of two cattail species, Typha domingensis and Typha latifolia (Typhaceae), in South Florida. *Molecular Phylogenetics and Evolution*, *49*(1), 368–376. https://doi.org/10.1016/j.ympev.2008.03.032

Zhou, B., Tu, T., Kong, F., Wen, J., & Xu, X. (2018). Revised phylogeny and historical biogeography of the cosmopolitan aquatic plant genus Typha (Typhaceae). *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-27279-3

Tables and Figures

Table 1. Mean Weir and Cockerham's pairwise genetic differentiation (F_{ST}) and genetic divergence (d_{XY}) measured in 10 Kbp windows for every variable site, number of SNPs, and diagnostic markers (SNPs with fixed opposite alleles) between three *Typha* spp.

Pairwise comparison	F_{ST}	d_{XY}	SNPs	Diagnostic SNPs
T. angustifolia – T. domingensis	0.25	0.28	10,358,977	33,436
T. angustifolia – T. latifolia	0.44	0.35	8,786,870	30,113
T. domingensis – T. latifolia	0.49	0.34	9,380,607	55,775

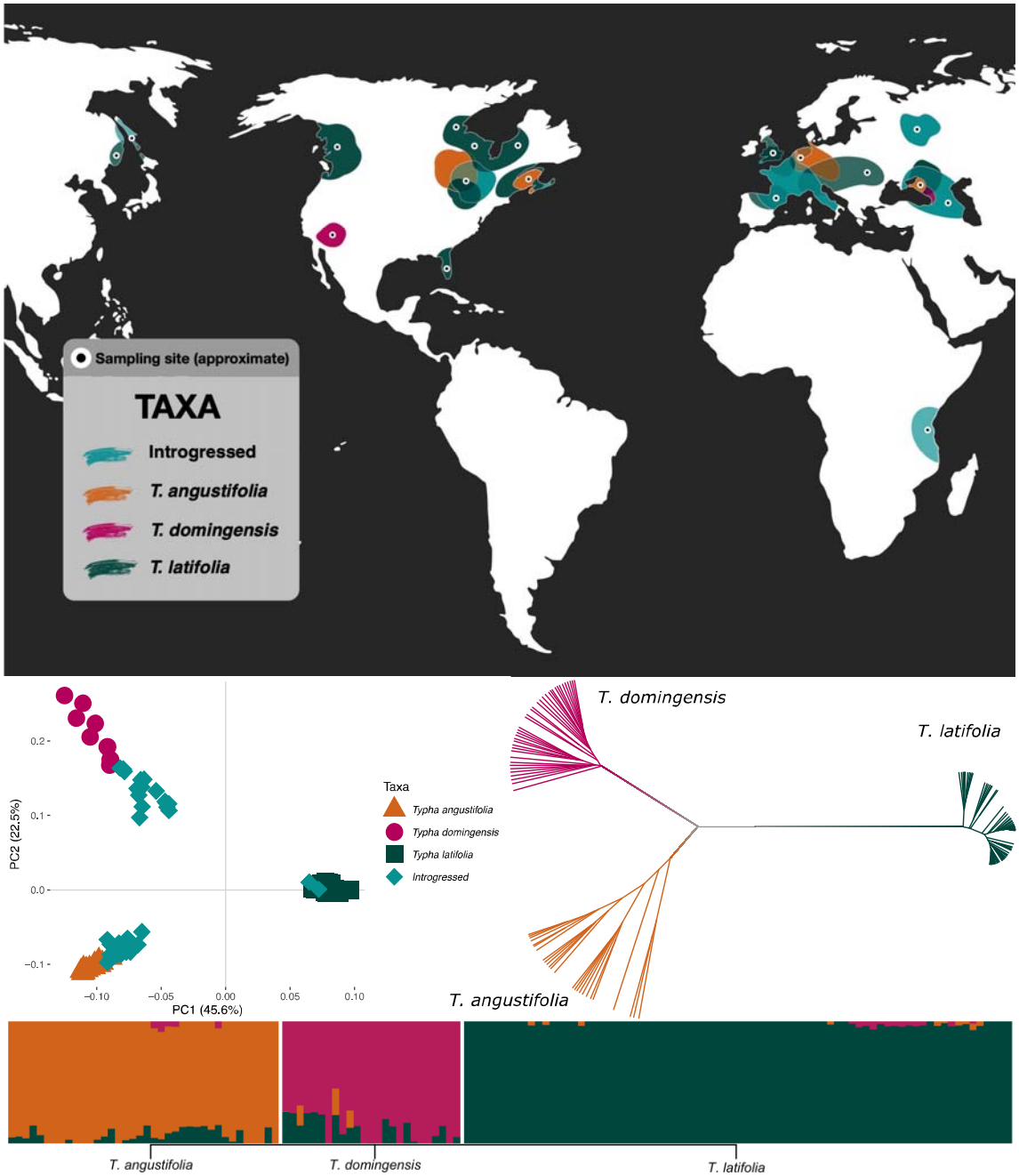

Figure 1. Genetic structure results of 12,177,703 nuclear SNPs obtained for three *Typha* spp. Top: Sampling locations in this study. Black points indicate approximate sampling sites, and coloured areas indicate the taxa identified. The number of samples is not shown. Left: Principal Component Analysis for PC1 and PC2. Shapes represent individuals, and colours represent taxa, as in the box. Right: Neighbour–joining tree. Branches represent individuals, and colours indicate species belonging as labelled. Bottom: ADMIXTURE (K = 3). Vertical bars represent individuals, and the admixture proportion is shown with different colours.

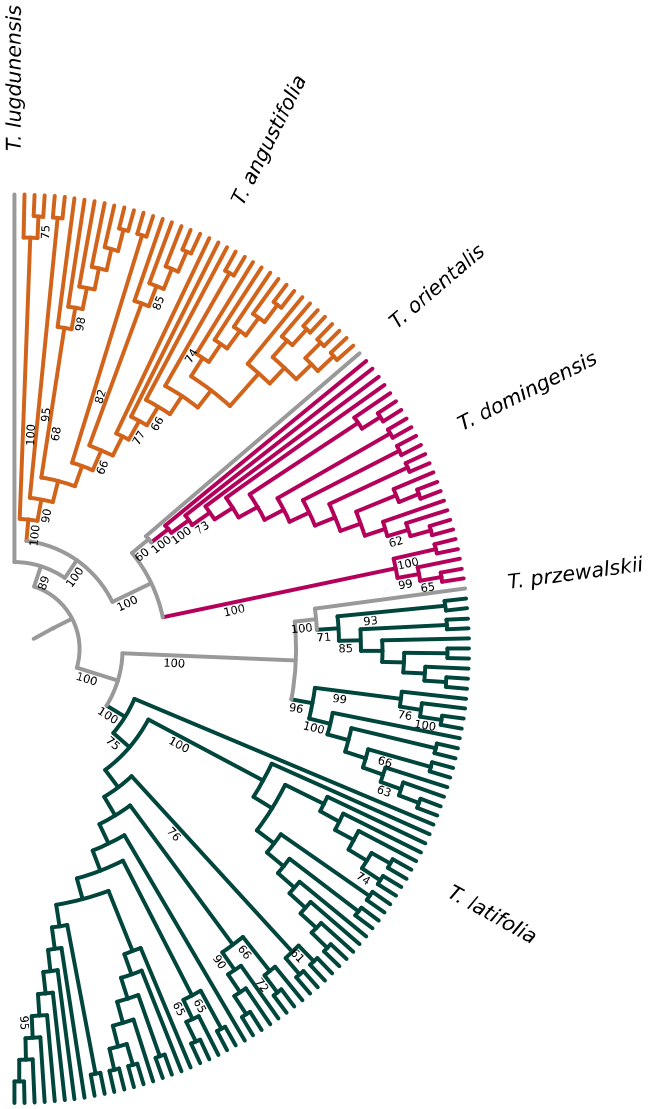

Figure 2. Chloroplast phylogeny of 140 samples from three *Typha* spp. and four NCBI references, based on 96,591 bp. The tree was produced with *Sparganium natans* as the outgroup and drawn without it. Branches represent individuals, and colours indicate species belonging as labelled. Numbers indicate branch support ≥ 60. Branch lengths are not shown.

Supplementary Table S1. Sampling identifiers, geographic coordinates, species membership, and genetic introgression status of the 140 *Typha* collected for this study.

Sampling ID	Latitude (approximate)	Longitude (approximate)	Species	Genetic introgression
BHP1	50.79	-119.21	*T. latifolia*	No
BJL2	50.57	-119.61	*T. latifolia*	No
BJL3_1	50.57	-119.61	*T. latifolia*	No
BRT2B	44.74	-63.24	*T. angustifolia*	No
BUN1	53.89	-122.81	*T. latifolia*	No
C05TL	52.72	-1.37	*T. latifolia*	No
CA13_TL	44.85	7.71	*T. latifolia*	No
CA3TL	47.68	22.46	*T. latifolia*	No
CC1_TL	50.77	2.31	*T. latifolia*	No
CL10TA	50.77	2.31	*T. angustifolia*	Yes
COTL6	52.72	-1.37	*T. latifolia*	No
CY6ATA	51.43	-0.11	*T. angustifolia*	No
CY7ATA	51.43	-0.11	*T. angustifolia*	Yes
CYTZ_4ATD	-6.79	39.21	*T. domingensis*	Yes
E74B	45.03	-63.50	*T. angustifolia*	Yes
EDD_44_1	47.37	-68.33	*T. latifolia*	No
EDW_2_15	47.37	-68.33	*T. latifolia*	No
EDW_3_6	47.37	-68.33	*T. latifolia*	No
EL14_TA	52.44	5.85	*T. latifolia*	No
ELTA_01	52.44	5.85	*T. angustifolia*	No
ELTA_08	52.44	5.85	*T. angustifolia*	No
FRD47_2	45.96	-66.64	*T. latifolia*	No
FRTD_1	41.64	13.34	*T. domingensis*	Yes
GORE_3	43.68	-70.44	*T. latifolia*	No
HAW_03_12	44.60	-63.55	*T. latifolia*	No

HAW_10_9	44.60	-63.55	*T. latifolia*	No
HM3TL	47.51	19.04	*T. latifolia*	No
HOR_6	48.14	26.51	*T. latifolia*	No
HUF_13TA	47.69	17.65	*T. angustifolia*	No
HUS_TL1	47.51	19.04	*T. latifolia*	No
IP2	44.46	-64.32	*T. latifolia*	No
IR_36	44.97	-64.06	*T. latifolia*	No
IR_39	44.97	-64.06	*T. latifolia*	No
IR_41	44.97	-64.06	*T. latifolia*	No
KL14_TL	46.64	14.31	*T. latifolia*	No
KL1TL	46.64	14.31	*T. latifolia*	No
LIETD_01	42.21	2.61	*T. domingensis*	Yes
LIETD_12	42.21	2.61	*T. domingensis*	Yes
LiETL_11	42.21	2.61	*T. latifolia*	No
M23	45.08	-64.49	*T. latifolia*	No
M30	45.08	-64.49	*T. latifolia*	No
M37	45.08	-64.49	*T. latifolia*	No
M85	45.08	-64.49	*T. latifolia*	No
MBNK_2_8	51.51	-97.00	*T. latifolia*	No
MOD_01_0	46.98	-70.55	*T. latifolia*	No
MOD_02_1	46.98	-70.55	*T. latifolia*	No
MOD_04_1	46.98	-70.55	*T. latifolia*	No
MOD_05_2	46.98	-70.55	*T. latifolia*	No
MOD_07_0	46.98	-70.55	*T. angustifolia*	No
MOD_08_1	46.98	-70.55	*T. latifolia*	No
MOD_14_0	46.98	-70.55	*T. latifolia*	No
MOD_16_1	46.98	-70.55	*T. latifolia*	No
MOD_17_1	46.98	-70.55	*T. latifolia*	No

MOD_18_0	46.98	-70.55	*T. latifolia*	No
MOD_25_2	46.98	-70.55	*T. latifolia*	No
MOD_26_2	46.98	-70.55	*T. latifolia*	No
MOD_29_1	46.98	-70.55	*T. latifolia*	No
MOD_35_3	46.98	-70.55	*T. latifolia*	No
MOD_45_2	46.98	-70.55	*T. latifolia*	No
MOD_54_0	46.98	-70.55	*T. latifolia*	No
NBU_6	42.53	2.83	*T. domingensis*	Yes
OR_TA4	43.28	-2.13	*T. domingensis*	Yes
OR_TL5	43.28	-2.13	*T. latifolia*	No
P6	45.41	-64.33	*T. latifolia*	No
P9	45.41	-64.33	*T. latifolia*	No
PA2	33.19	-112.15	*T. domingensis*	No
PIW_04_6	43.84	-79.10	*T. angustifolia*	No
PIW_09_1	43.84	-79.10	*T. angustifolia*	No
PIW_13_0	43.84	-79.10	*T. angustifolia*	No
PIW_16_2	43.84	-79.10	*T. angustifolia*	No
PIW_24_3	43.84	-79.10	*T. angustifolia*	Yes
PIW_25_3	43.84	-79.10	*T. angustifolia*	Yes
PWE19TL	51.62	-3.94	*T. latifolia*	No
RGT2	45.48	-74.30	*T. angustifolia*	No
SAD_17_3	45.27	-66.09	*T. latifolia*	No
SP2_TA	52.41	4.68	*T. angustifolia*	No
TAKT_1	45.27	37.38	*T. angustifolia*	No
TAKT_2	45.27	37.38	*T. angustifolia*	Yes
TALD_1	55.73	37.61	*T. angustifolia*	Yes
TAVG_1	55.32	40.37	*T. angustifolia*	Yes
TAVG_2	55.32	40.37	*T. angustifolia*	Yes

TA_01_F	26.56	-81.95	*T. latifolia*	No
TB7	44.39	-64.25	*T. latifolia*	No
TDD_04	43.09	47.46	*T. domingensis*	Yes
TDD_08	43.09	47.46	*T. domingensis*	Yes
TDD_09	43.09	47.46	*T. domingensis*	Yes
TDD_15	43.09	47.46	*T. domingensis*	Yes
TDD_16	43.09	47.46	*T. latifolia*	No
TDD_17	43.09	47.46	*T. latifolia*	No
TDD_18	43.09	47.46	*T. domingensis*	No
TDD_22	43.09	47.46	*T. domingensis*	No
TDD_30	43.09	47.46	*T. domingensis*	No
TDD_32	43.09	47.46	*T. domingensis*	Yes
TDD_33	43.09	47.46	*T. domingensis*	Yes
TDD_35	43.09	47.46	*T. latifolia*	No
TDD_36	43.09	47.46	*T. latifolia*	Yes
TDD_37	43.09	47.46	*T. latifolia*	No
TDD_39	43.09	47.46	*T. latifolia*	No
TDD_40	43.09	47.46	*T. domingensis*	No
TDD_41	43.09	47.46	*T. latifolia*	No
TDD_42	43.09	47.46	*T. latifolia*	No
TDD_43	43.09	47.46	*T. angustifolia*	Yes
TDD_44	43.09	47.46	*T. domingensis*	Yes
TDD_46	43.09	47.46	*T. domingensis*	No
TDD_47	43.09	47.46	*T. domingensis*	No
TDD_49	43.09	47.46	*T. angustifolia*	Yes
TDD_50	43.09	47.46	*T. domingensis*	No
TDD_51	43.09	47.46	*T. domingensis*	Yes
TDD_52	43.09	47.46	*T. domingensis*	Yes

TDD_53	43.09	47.46	*T. domingensis*	No
TDD_54	43.09	47.46	*T. domingensis*	Yes
TLIK_01	45.09	147.89	*T. latifolia*	No
TLIK_03	45.09	147.89	*T. latifolia*	No
TLIK_06	45.09	147.89	*T. latifolia*	No
TLIK_15	45.09	147.89	*T. latifolia*	No
TLIK_24	45.09	147.89	*T. latifolia*	No
TLMK_1	45.21	38.89	*T. latifolia*	No
TLPK_1	45.21	38.89	*T. latifolia*	No
TLSK_02	47.06	142.53	*T. latifolia*	No
TLSK_04	47.06	142.53	*T. latifolia*	No
TLSK_11	47.06	142.53	*T. latifolia*	Yes
TLSK_12	47.06	142.53	*T. latifolia*	No
TMO2_7	46.09	-64.78	*T. latifolia*	No
TXKT_1	45.29	37.41	*T. angustifolia*	Yes
TXKT_2	45.29	37.41	*T. angustifolia*	Yes
TXKT_3	45.29	37.41	*T. angustifolia*	Yes
TXKT_4	45.29	37.41	*T. angustifolia*	Yes
TXKT_6	45.29	37.41	*T. angustifolia*	Yes
TXVG_1	55.32	40.37	*T. angustifolia*	Yes
VFW6	49.63	-125.40	*T. latifolia*	No
Vikram_25	44.30	-78.32	*T. angustifolia*	Yes
Vikram_33	44.30	-78.32	*T. angustifolia*	No
Vikram_76	44.30	-78.32	*T. latifolia*	No
Vikram_94	44.30	-78.32	*T. angustifolia*	Yes
WBD20_1	44.28	-84.23	*T. angustifolia*	No
WBD_27_4	44.28	-84.23	*T. angustifolia*	No
WBD_28_3	44.28	-84.23	*T. angustifolia*	No

WBD_32_1	44.28	-84.23	*T. angustifolia*	Yes
WBD_33_1	44.28	-84.23	*T. angustifolia*	No
WBW5_4	44.28	-84.23	*T. latifolia*	No

634