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

A critical knowledge gap in freshwater plants research is the lack of genetic resources necessary to answer fundamental questions about their demographic histories, adaptation, and taxonomy. One example of this is *Typha*, a macrophyte genus essential to wetlands that is also becoming an increasingly problematic biological invader in numerous regions worldwide; while important insights have been discovered for this genus, currently available genetic markers are insufficient to resolve its phylogenetic relationships, population structure, and hybridization dynamics. We performed a cost and time–accessible library preparation for high–throughput sequencing to develop a suite of genomic resources for *Typha*. Genome–wide nuclear SNPs of 140 *Typha* samples from North America, temperate Eurasia and Africa revealed three independent genetic clusters corresponding to *T. angustifolia*, *T. domingensis* and *T. latifolia*. Data from ~40% of the nuclear genome was obtained, permitting the characterization of 119,324 nuclear diagnostic markers (SNPs that differentiate the three species). A reference–guided workflow to reconstruct whole–chloroplast–genome sequences was implemented, recovering ~60% of the genome per sample. Three genetic lineages were identified from the cpDNA phylogenetic analysis, agreeing with those determined from nrDNA. With a cost below 15 USD per sample and a processing time of two hours for the library preparation, this is a rapid and cost–effective protocol for population genomic studies. The genomic locations of diagnostic markers and the chloroplast sequences produced in this study will be permanent resources that can be incorporated into future studies of *Typha,* a globally important and evolutionarily dynamic genus.

Keywords: Hybridization, diagnostic markers, chloroplast–genome assembly, high–throughput sequencing, *Typha* phylogenomics, introgression.

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 plant 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).

Genomic characterization 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 feasible with 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 freshwater plants foundational to wetlands (reviewed in Bansal et al., 2019). Cattails are a valuable ecosystemic and human resource that 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). While

62 their importance is undisputed, one major challenge in *Typha* research has been taxonomic
63 identification, which cannot be fully accomplished using morphological characters due to
64 their high intraspecific variability and interspecific hybridization. Consequently, the richness
65 of cattail species, their taxonomy, their provenance (e.g., alien versus native lineages), and
66 their phylogenetic relationships remain incompletely resolved (Ciotir & Freeland, 2016;
67 Volkova & Bobrov, 2022; Zhou et al., 2018). A refined *Typha* taxonomy accompanied by
68 diagnostic markers, i.e., those loci that have alternative fixed alleles between two species, is
69 necessary to identify lineages of both introduced and hybrid taxa, which are increasingly
70 documented as invasive, e.g., *T. domingensis* Pers. in Western Europe and Central America,
71 *T. latifolia* L. in Oceania, and *T.* × *glauca* Godr. (*T. angustifolia* L. × *T. latifolia* L.) in North
72 America (Bansal et al., 2019; *GISD*; Govaerts, 2004; Hall, 2009; Maldonado, 2019).
73 High–throughput sequencing technologies, novel, cost and time–accessible genome
74 library preparations, and the recent assembly of the *T. latifolia* genome (Goodwin et al., 2016;
75 Rowan et al., 2019; Widanagama et al., 2022) collectively present an opportunity to develop a
76 suite of genomic resources for *Typha*. In addition to taxonomic resolution, these resources can
77 facilitate investigations of the genetic variation, local adaptation, population structure,
78 evolutionary history, and hybridization dynamics in this genus. We applied a high–throughput
79 sequencing protocol for enzymatic library preparation and data processing to produce
80 genome–wide resources in *Typha* spp. By optimizing the method from Rowan et al. (2019),
81 we generated a catalogue of nuclear SNPs that differentiate *T. angustifolia*, *T. domingensis*,
82 and *T. latifolia*, plus extensive chloroplast genome sequences in a fast, straightforward, and
83 cost–efficient manner. These permanent resources will be incorporated into future
84 evolutionary and population genetics research in *Typha*.

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) to 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 was 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

DNA extractions were obtained either from previous studies (Bhargav et al., 2022; Ciotir et al., 2017; Pieper et al., 2020, 2017; Tangen et al., 2022; Tisshaw et al., 2020) or from samples collected in Eurasia and extracted at Trent University following the published protocols. We obtained DNA from 140 *T. angustifolia*, *T. domingensis*, and *T. latifolia* samples (Figure 1), previously identified using a combination of morphological characteristics (Grace & Harrison, 1986; Smith, 1967) and genetic analyses of microsatellite loci (Bansal et al., 2019; Kirk et al., 2011; Pieper et al., 2020; Snow et al., 2010). Extracted DNA was quantified using a Qubit fluorometer (Thermofisher Scientific) and calculated as the mean of three independent readings from each sample. All samples were either standardized to 2 ng/µL by dilution with nuclease–free water or, if at concentrations less than 2 ng/µL (0.4 – 1.9 ng/µL), left undiluted.

108 Each DNA sample was prepared for Nextera XT sequencing libraries by enzymatic
109 fragmentation followed by ligation of short adapter sequences with the Illumina Tagment
110 DNA enzyme (TD) and buffer kit (small kit, #20034210). As the ratio of TD enzyme to DNA
111 is crucial for the reaction, we initially followed the recommendations of Rowan et al. (2019)
112 and subsequently optimized the reagent volumes for our library preparation as 5.5 µL of 5×
113 TD buffer, 0.5 µL of 1× TD enzyme, and 4 µL of DNA (standardized or undiluted), keeping
114 all reagents on ice during the preparation. Samples were incubated at 55°C for 10 minutes and
115 left at room temperature for five minutes. 5 µL of each sample were run on an agarose gel to
116 confirm the efficacy of the tagmentation reaction, evidenced by visible smears. The
117 fragmented DNA was then amplified using unique dual indexing based on combinations from
118 a total of 24 N7 (47 bases) and 8 S5 (51 bases) adapters (Alpha DNA, Canada). The PCR
119 cocktail included 0.2 µM of each index, 0.5U of KAPA HiFi HotStart DNA polymerase
120 (Roche), 12.5 µL of 5× KAPA reagent, 5 µL of tagmented DNA, and 6.5 µL of nuclease–free
121 water to a final volume of 25 µL. The PCR cycle comprised 72°C (3 minutes); 95°C (30
122 seconds); and 14 cycles of 95°C (10 seconds), 55°C (30 seconds), and 72°C (30 seconds).
123 Visible smears confirmed amplification success after running 5 µL of the PCR product on an
124 agarose gel; then, 10 µL from each sample were pooled, and the remaining PCR products
125 were stored at -20°C. The pooled library was purified with a QIAquick PCR purification kit
126 (QIAGEN) following the manufacturer's protocol, with a final elution in 50 µL of buffer. The
127 library was quantified using a D1000 Tapestation assay (Agilent Technologies, USA) and a
128 Qubit fluorometer (Thermo Fisher Scientific). A quality–control paired–end sequencing was
129 executed using a Miseq (151 bp) to ensure the library was compiled successfully. Finally,
130 paired–end sequencing was performed on a Novaseq 6000 (126 bp) at The Centre for Applied
131 Genomics (Toronto, Ontario).

132 *Raw data processing, filtering, and SNP–calling*

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

141 SNP–calling was performed with ANGSD 0.93 (Korneliussen et al., 2014) following
142 the SAMtools model, exclusively retrieving SNPs with a minimum *P*–value of $1e^{-6}$, minimum
143 mapping and sequencing qualities of 20, omitting any triallelic sites, and outputting a variant
144 call format (VCF) file (*-doGeno 4 -gl 1 -skipTriallelic 1 -SNP_pval 1e-6 -minMapQ 20 -minQ*
145 *20 -doMajorMinor 1 -domaf 1 -doPost 1 -doCounts 1 -dovcf* 1). For the nuclear analyses,
146 SNPs with > 50% missing data across all samples and sites that mapped to the chloroplast
147 genome were removed with vcftools 0.1.16 (Danecek et al., 2011). We did not apply
148 additional filters to SNP identification: our samples represented broad geographical sampling
149 and thus were not expected to be in Hardy–Weinberg equilibrium; additionally, as allele
150 frequencies were unlikely to be representative of regional allele frequencies we did not apply
151 a minor allele frequency filter. Neither did we filter for linkage equilibrium, as eliminating
152 alleles that are in linkage disequilibrium is likely to decrease the power of analyses to detect
153 hybridization and introgression (Alexander, 2020; Pearman et al., 2022).

154 *Genetic structure and diagnostic markers*

155 We first used nuclear SNPs to assess the most likely number of genetic clusters across
156 all samples and the membership of each plant to these clusters using three complementary
157 approaches. ADMIXTURE 1.3.0 (Alexander & Lange, 2011) was run with K = 1 – 10, and
158 the optimal number of clusters was chosen via the cross–validation procedure. A neighbour–
159 joining tree from the samples' pairwise genetic distance matrix (transformed on the R 4.2.2
160 package ape (Paradis et al., 2004; R Core Team, 2022)), and a Principal Component Analysis
161 (PCA) without an *a priori* assumption were produced in Plink 1.90 (Purcell et al., 2007). We
162 verified that the assignment of samples to genetic clusters (corresponding to three species, see
163 *Results*) was consistent for each approach.

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

171 To compare the levels of differentiation between clusters, values of Weir and
172 Cockerham's genetic differentiation (F_{ST}; Weir & Cockerham, 1984) for every variable site
173 and genetic divergence (d_{xy}; Nei & Miller, 1990) in 10 Kbp windows between species pairs
174 were computed using pixy 1.2.7 (Korunes & Samuk, 2021), and the means were calculated.
175 Species–specific SNPs for each pair of genetic clusters were identified with DiagnoSNPs 1.0
176 (Arce-Valdés, 2022). We removed introgressed individuals before estimating genetic
177 differentiation levels and identifying species–specific SNPs.

178 *Chloroplast genome reconstruction and phylogenetic analysis*

179 We implemented a reference–guided workflow to reconstruct whole–chloroplast–
180 genome sequences. Nucleotide calling was performed individually for each of the 140
181 samples in ANGSD using the reads that mapped to the chloroplast genome reference,
182 requiring minimum mapping and base qualities of 20, and using Ns as missing data (*-dofasta*
183 *2 -minMapQ 20 -minQ 20 -doCounts 1*). Sequences were aligned to the chloroplast genomes
184 of *T. przewalskii* Skvortsov, *T. lugdunensis* P. Chabert, *T. orientalis* C. Presl, and
185 *Sparganium natans* (GenBank accessions: NC_061354.1, NC_061353.1, NC_050678.1, and
186 NC_058577.1), following Smith et al. (2021) by applying MAFFT 7.0 default settings (plus
187 the flag *--nwildcard*) (Katoh et al., 2019). Snp–sites 2.5.1 (Page et al., 2016) was used to
188 remove sections of the genome with ambiguous positions, gaps, and missing data, such that if
189 any of those was found in a sequence, that position was removed for all sequences.
190 Nucleotide diversity (π) was calculated in the R package pegas (Paradis, 2010).

191 The phylogenetic relationships of the chloroplast genome sequences were
192 reconstructed through RAxML-NG 1.1 (Kozlov et al., 2019). Model selection was based on
193 jmodeltest 2.1.10 results (Darriba et al., 2012) using the AIC. A transversion model was
194 established, assuming gamma–distributed rates with four substitution categories and
195 estimating the proportion of invariable sites (TVM + G4 + I). RAxML was run with a
196 thorough bootstrap starting from 100 random trees, with *Sparganium natans* as the outgroup,
197 and letting automatic bootstrapping (--autoMRE). The best–scoring ML tree was plotted.

198 **Results**

199 *Genome scaffolding, mapping statistics, and genotyping*

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

208 After quality control, 982 M clean paired–end reads were retained, and ~98% mapped
209 to the reference genome. With minimum mapping and sequencing qualities = 20, the average
210 depth and breadth of coverage of the nuclear sequences were 4× and 42%, respectively. Over
211 60% of the chloroplast genome breadth was covered across all samples (mean depth = 711×),
212 enabling us to use 96,591 bp for the phylogenetic reconstruction. We assembled 12,177,703
213 bi–allelic nuclear SNPs across the 140 *Typha* samples (7,122,151 with a MAF > 0.05). The
214 total genotyping rate, i.e., the mean proportion of samples with data for each SNP, was 0.68.

215 *Genetic structure and diagnostic markers*

216 The admixture analysis, the PCA, and the neighbour–joining tree each established the
217 most likely number of genetic clusters as three (K = 3) (Figure 1): in line with previous
218 taxonomic identifications, 38 samples were identified within the *T. angustifolia* cluster (15 of
219 which had *T. latifolia* introgression and 5 of which had both *T. domingensis* and *T. latifolia*
220 introgression); 25 samples were in the *T. domingensis* cluster (one with *T. angustifolia*
221 introgression, 12 with *T. latifolia* introgression, and 3 with both *T. angustifolia* and *T. latifolia*

222 introgression); and 77 samples were in the *T. latifolia* cluster (one with both *T. angustifolia*
223 and *T. domingensis* introgression, and one with *T. angustifolia* introgression). Using only the
224 18 *T. angustifolia*, 9 *T. domingensis*, and 75 *T. latifolia* not introgressed individuals, the mean
225 pairwise interspecific F_{ST} and d_{xy} values ranged from 0.25 to 0.49 and 0.28 to 0.35,
226 respectively, with *T. latifolia* showing the highest differentiation with both *T. angustifolia* and
227 *T. domingensis.* We identified 119,324 nuclear species-specific SNPs (Table 1) based on the
228 comparisons between species pairs.

229 *Chloroplast genome reconstruction and phylogenetic relationships*

230 After removing all ambiguities and missing data from the chloroplast genomes, we
231 were left with an alignment of 96,591 bp across all 140 sequences and the four reference
232 genomes, with 4,916 segregating sites and $\pi = 0.003$. The phylogenetic reconstruction was
233 consistent with the nuclear genetic structure results, grouping the 143 *Typha* samples into
234 three lineages, with *T. angustifolia* in one clade, *T. domingensis* and *T. orientalis* sharing
235 another, and *T. latifolia* and *T. przewalskii* in a third clade (Figure 2). All interspecific nodes
236 were strongly supported (100%, Felsenstein bootstrap proportion).

237 **Discussion**

238 We aimed to produce a suite of genome–wide resources to facilitate investigations into
239 the taxonomy and population genetics of *Typha* and to advance the general understanding of
240 wetland plants. Following a fast, straightforward, and cost–efficient genomic library
241 preparation protocol (Rowan et al., 2019), we sequenced 140 *Typha* samples, obtaining an
242 average breadth of 42% of the nuclear genome, characterizing 119,324 nuclear SNPs that
243 collectively differentiate three *Typha* spp., and producing chloroplast sequences with a
244 breadth of > 60% per sample. With a cost below 15 USD per sample and a processing time of
245 two hours for the library preparation, our workflow is a rapid and cost–effective protocol that

246 can be applied in population genomic studies for investigating levels of genetic diversity and
247 differentiation, identifying conservation units and alien taxa, characterizing hybrid classes,
248 and examining hybridization dynamics, among other purposes.

249 Three genetic clusters were identified from both nuclear and chloroplast genomes,
250 corresponding to *T. angustifolia*, *T. domingensis*, and *T. latifolia*; and potential introgression
251 was detected among the three species. Hybridization of *T. domingensis* with *T. angustifolia*
252 and *T. latifolia* has been historically documented (Govaerts, 2004; Smith, 1967); however,
253 those inferences were based on experimental and taxonomical observations and recently
254 exposed with microsatellite loci by Ciotir et al. (2017). Future investigations should address
255 the extent to which hybridization is shaping the genetic variation among these three species.
256 Furthermore, *T. domingensis* is increasingly invading Costa Rica (Trama et al., 2017), Nigeria
257 (Ringim et al., 2016), and North America, potentially expanding its range across the latter
258 (Spencer & Vincent, 2013; Zhang et al., 2008). However, the taxonomic identity of these
259 plants is unclear –are they hybrids, non–native lineages, or native lineages responding to
260 environmental change? By identifying SNPs that characterize *T. domingensis*, we provide a
261 valuable resource to answer this and other questions across the evolutionary and hybridization
262 history of *Typha*.

263 The markers that differentiate *T. angustifolia* from *T. latifolia* will have important
264 applications in North America, where the two species interbreed across a large area and
265 produce an invasive interspecific hybrid (*T.* × *glauca*) that dominates wetlands, alters nutrient
266 cycling, and reduces biodiversity across the Great Lakes Region (Bansal et al., 2019);
267 additionally, this hybrid is expanding throughout the Prairie Pothole Region, causing native
268 plant diversity to decrease in invaded potholes (Jones et al., 2023), and may impact essential
269 habitat for millions of breeding and migratory waterfowl species (Tangen et al., 2022). Until

270 now, molecular resources to differentiate *T. angustifolia*, *T. latifolia*, and *T.* × *glauca* were
271 limited to sets of relatively few individual markers that have produced important insights:
272 RAPDs, chloroplast DNA sequences, and codominant SSR loci have contributed to exposing
273 the sexual fertility of first–generation hybrids (F1) (Snow et al., 2010), asymmetric
274 hybridization (with *T. angustifolia* being mainly the maternal parent) (Ball & Freeland, 2013;
275 Kuehn et al., 1999; Pieper et al., 2017), overall comparable levels of sexual and clonal
276 reproduction in parents and F1s (Pieper et al., 2020; Travis et al., 2011), heterosis in F1s
277 (Bunbury-Blanchette et al., 2015; Travis et al., 2010; Zapfe & Freeland, 2015), a high
278 frequency of F1s in natural populations (Kirk et al., 2011; Travis et al., 2010), and partial
279 sterility of F1s coupled with hybrid breakdown of F2s and advanced–generation hybrids
280 (Bhargav et al., 2022; Pieper et al., 2017). However, critical questions remain unresolved
281 because existent markers are insufficient to expose the prevalence of advanced–generation
282 hybrids and backcrosses in wild populations. The expansive suite of SNPs identified in this
283 study will facilitate investigations into the extent of hybridization, hybrid breakdown
284 dynamics, and adaptive introgression across the *T.* × *glauca* hybrid zone, allowing
285 researchers to understand the evolutionary processes shaping this genus speciation and
286 species boundaries, and to inform conservation and management strategies.
287 Fundamental genetic tools are essential for understanding macrophytes biology,
288 management, and conservation (O'Hare et al., 2018). The protocol described in this paper, the
289 updated chromosome–level assembly genome of *T. latifolia*, the SNPs catalogue, and the
290 chloroplast sequences produced for each sample comprise permanent resources that can be
291 applied to study the genetic composition of *Typha* populations and hybrid zones. Genome–
292 wide sequencing techniques and reference–based chloroplast genome assemblies are
293 promising tools to clarify the demographic histories, dispersal, adaptation, and taxonomy of

294 multiple congeneric macrophyte species (Russello et al., 2015; Straub et al., 2012), and
295 substantial genome–wide research will allow us to tackle these and other knowledge gaps in
296 *Typha* and other taxa.

297 **Acknowledgements**

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

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

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

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, *49*(W1), W293–W296. https://doi.org/10.1093/nar/gkab301

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

Wickham, H. (2016). *Ggplot2*. Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4

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

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

585 **Data Accessibility Statement**

586 The updated genome assembly, raw reads supporting this study's findings, SNP
587 information and chloroplast sequences produced in the present study will be deposited into
588 online databases upon publication. Upon publication, related metadata and all the scripts used
589 will be available (https://gitlab.com/WiDGeT_TrentU). We will provide unique sample
590 identifier tags that can be matched to the deposited genetic data, metadata, and scripts.

591 **Benefit–Sharing Statement**

592 Benefits from this research accrue from sharing our data and results on public
593 databases, as described above.

594 **Author Contributions**

595 Joanna R. Freeland, Aaron B. A. Shafer and Marcel E. Dorken conceived this study.
596 Joanna R. Freeland and Polina A. Volkova collected and characterized the plant tissue
597 samples. Tulsi Patel performed the lab work. Alberto Aleman executed the bioinformatics and
598 wrote the manuscript. All authors contributed to the manuscript and approved this version.

599 **Conflict of Interest**

600 The authors declare no conflict of interest.

601 **Supporting Information**

602 **Tables and Figures**

603 **Table 1.** Mean Weir and Cockerham's pairwise genetic differentiation (F_{ST}) for every variable

604 site, mean genetic divergence (d_{xy}) measured in 10 Kbp windows, number of SNPs, and

605 diagnostic markers (SNPs with fixed opposite alleles) between three *Typha* spp.

606

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

607

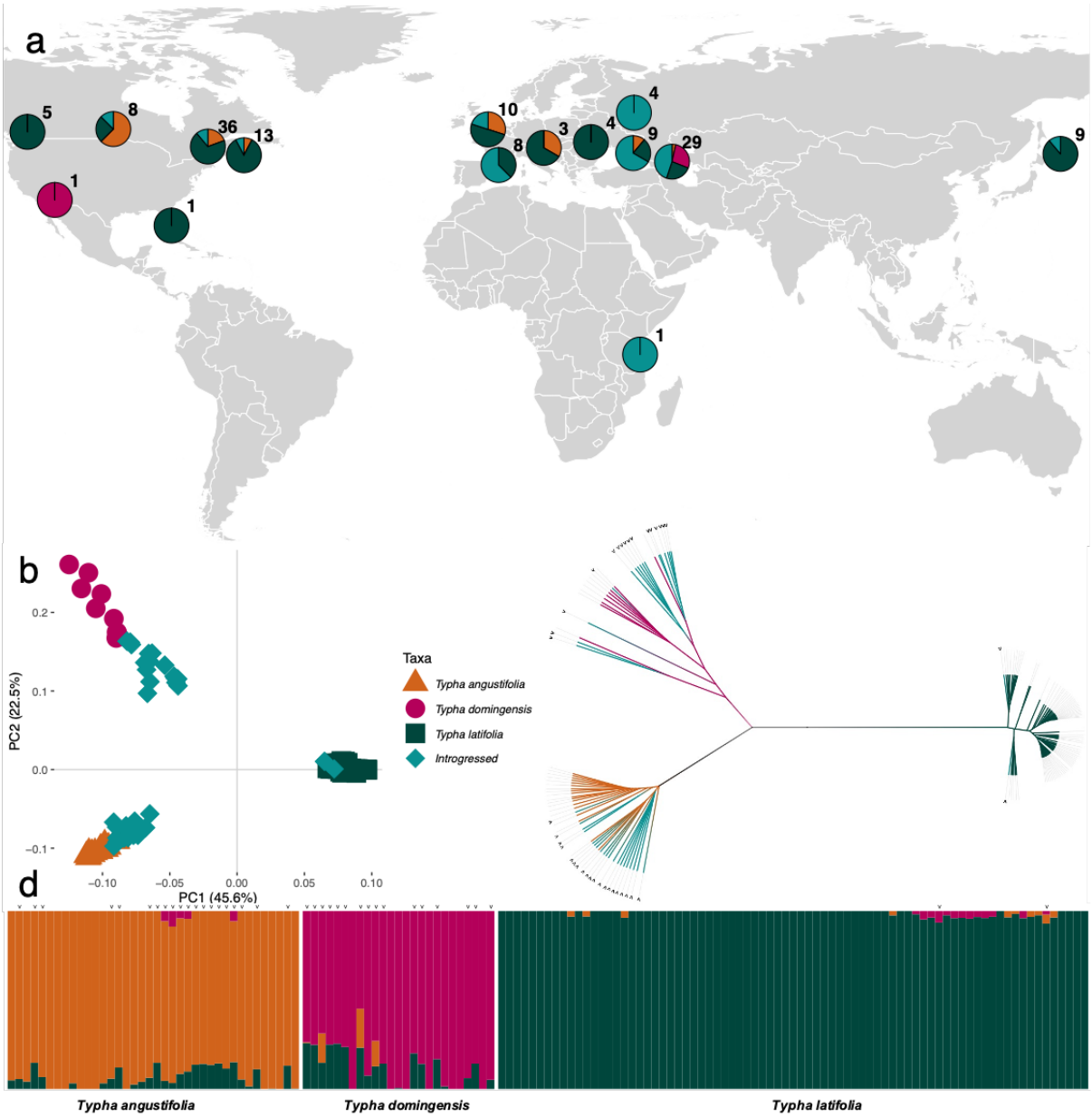

Figure 1. Genetic structure results of 12,177,703 nuclear SNPs obtained for three *Typha* spp. (a) Sampling locations in this study. Coloured areas of the pie charts reflect the proportion of samples from each taxon, and the number of samples is written top–right of each pie chart. (b) Principal Component Analysis for PC1 and 2. Shapes represent individuals, colours represent taxa, (c) Neighbour–joining tree. Tips represent individuals; colours represent clusters as a and b, (d) ADMIXTURE analysis for $K = 3$. Vertical bars represent individuals, and the probability of belonging to different taxa is coded with different colours. Arrows in b and d indicate the location of potentially introgressed individuals. Colour and shape code as the box in b: *T. angustifolia* (orange, triangles), *T. domingensis* (pink, circles), *T. latifolia* (green, squares), and potentially introgressed individuals (cyan, diamonds). a-b plotted in ggplot2 (Wickham, 2016), c plotted with iTOL v6 (Letunic & Bork, 2021), and d plotted in R.

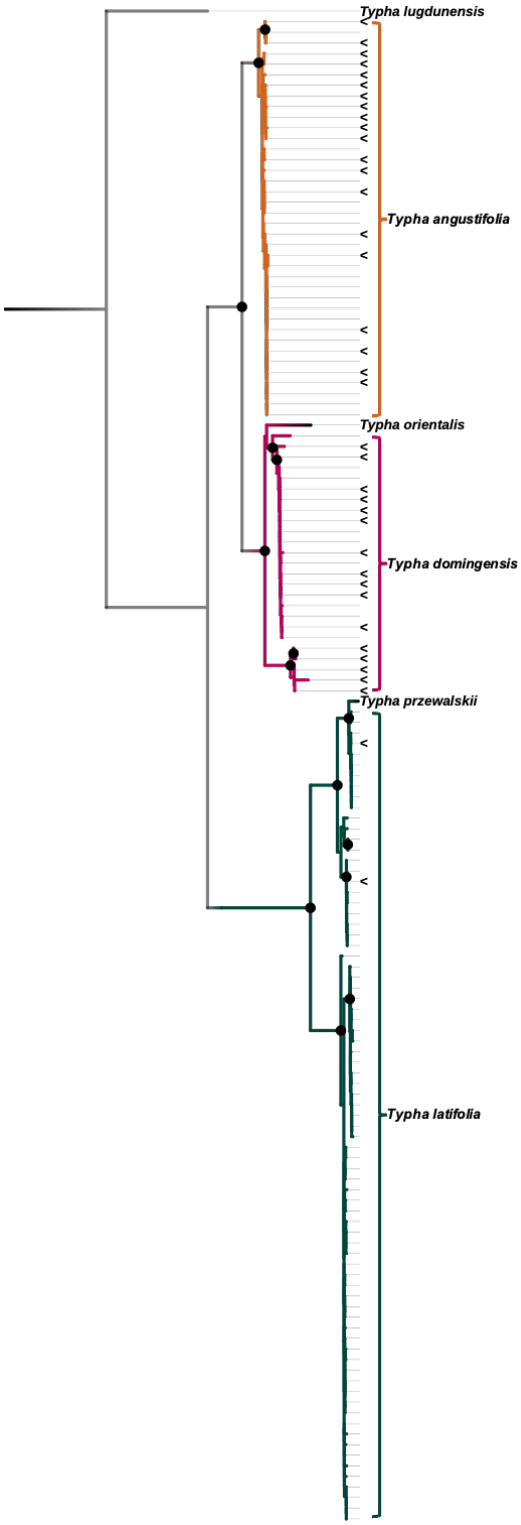

Figure 2. Chloroplast phylogeny of 140 *Typha* samples and four NCBI references, based on 96,591 bp. The tree was produced with *Sparganium natans* as the outgroup and drawn without it. Coloured lines recreate the genetic structure results in Figure 1, i.e., *T. latifolia* (green), *T. angustifolia* (orange), and *T. domingensis* (pink). Black circles indicate branch support = 100% (Felsenstein bootstrap proportion). Arrows indicate the location of potentially introgressed individuals (genetic structure results).