Allopatric speciation in cattails: Genomics reveal bottlenecks, balancing selection, and adaptive introgressions in *Typha*, a wetland ecosystem engineer

Alberto Aleman[1], Aaron Shafer[1,2], Joanna Freeland[1], Marcel Dorken[1,3]

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

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

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

Corresponding authors:

Marcel E. Dorken: marceldorken@trentu.ca

Alberto Aleman: albertolopezaleman@trentu.ca

Word count for main text: 4248

Keywords: climate driven barriers; divergence; drift; evolutionary history; genomic landscapes; hybridisation; secondary contact

Abstract

Speciation can be broadly understood within two non-mutually exclusive frameworks: genetic drift under isolation and natural selection under ecological divergence. Here, we examined the genomic diversity and differentiation of five *Typha* species, a group of plants foundational to freshwaters with widespread, partially sympatric distributions and at least one widespread hybrid zone. Using genome-wide data from 207 individuals, we examined the contributions of demographic fluctuations, selection, and hybridisation in driving their speciation history. Demographic reconstructions revealed sequential bottlenecks and expansions coincident with lineage splits, and indicated a drift-driven scenario with no migration events for all five species. The genomic landscapes showed balancing selection, sparse divergent selection, and low net divergence. Introgressions from *T. latifolia* to *T. angustifolia* and *T. domingensis* were found. Our findings suggest histories of allopatric speciation followed by range expansions and secondary contacts, leading to contemporary hybridisation between some species. Our results also emphasise the roles of balancing selection and introgression as sources of standing genetic variation. Allopatric speciation in *T. latifolia* and *T. angustifolia* could explain their ability to hybridise, highlighting the need to stop the human-mediated dispersal of *Typha* (e.g., the intercontinental sourcing via garden centres).

Introduction

Speciation can follow two non-mutually exclusive pathways: geographic isolation ("allopatric speciation") and natural selection ("ecological speciation") (Stankowski and Ravinet, 2021). In isolation, divergence is primarily driven by genetic drift (Mayr *et al.*, 1963), and although local adaptations occur, they are not assumed to be the driving force underlying species formation (Hernández-Hernández *et al.*, 2021). Conversely, divergent selection is mandatory during ecological speciation (Pinho and Hey, 2010; Nosil, 2012). Understanding how these pathways lead to lineages diverging is key to speciation genomics and can help us characterise how biological diversity originates (Seehausen *et al.*, 2014).

Global climatic oscillations have repeatedly shaped species' ranges and population sizes, triggering isolation and providing opportunities for allopatric divergence (Hewitt, 2000). These fluctuations include the rapid transformations of the Neogene (Andersson, 2009) and the glacial cycles of the Quaternary, which promoted speciation through drift events (e.g., Chacón *et al.*, 2019; Kong *et al.*, 2022; Dagallier *et al.*, 2024). Drift-driven speciation can be linked to sharp changes in species' past demographic sizes (Bock *et al.*, 2023) and is the most common cause of divergence in plants (Hernández-Hernández *et al.*, 2021). In contrast, natural selection promotes ecological divergence by fixing beneficial alleles and purging deleterious ones (Schluter, 2001). Under this model, speciation leads to the emergence

of “genomic landscapes” (Wolf and Ellegren, 2017). These landscapes are highly differentiated (“islands”) or conserved (“valleys”) genomic regions along (or linked to) loci associated with natural selection and reproductive isolation (Ravinet *et al.*, 2017). Reconstructing these landscapes can be used to identify the contributions of natural selection to speciation (Andrew and Rieseberg, 2013; Ravinet *et al.*, 2017; Han *et al.*, 2017; Irwin *et al.*, 2018; Ottenburghs *et al.*, 2020; Kessler *et al.*, 2023; Shang *et al.*, 2023).

Gene flow plays an important role in speciation (Feder *et al.*, 2012; Tigano and Friesen, 2016). The restriction of gene flow, through allopatry or natural selection, permits genetic differentiation and the formation of distinct lineages (Morjan and Rieseberg, 2004), whereas gene flow generally plays a homogenising role (Woodruff, 2001). However, gene flow between genetically distinct species (hybridisation) has diverse outcomes (Harrison and Larson, 2014; Taylor and Larson, 2019). Hybridisation can cause the formation of new species, the extinction of parental taxa through genetic swamping, or the reinforcement of species’ boundaries (Noor, 1999; Garner *et al.*, 2018; Runemark *et al.*, 2019). In stable hybrid zones, hybridisation can cause introgressions (the movement of genes from one species to another), promoting adaptation when introgressed loci provide higher fitness (Whitney *et al.*, 2010; Suarez-Gonzalez, Lexer, *et al.*, 2018; Leroy *et al.*, 2020; Rendón-Anaya *et al.*, 2021).

Cattails (*Typha*) are rhizomatous perennial, monoecious, self-compatible, and wind-pollinated plants crucial to freshwater and brackish ecosystems; they play a vital ecological role in cycling nutrients, preventing erosion, maintaining water levels, and providing food and shelter for amphibians, birds, fish, insects, and mammals (Bonanno and Cirelli, 2017; Bansal *et al.*, 2019; Svedarsky *et al.*, 2019). The three most widespread *Typha* spp. (*T. angustifolia* L., *T. domingensis* Pers., and *T. latifolia* L.) have extensive areas of sympatry and allopatry (Figure 1) and are capable of hybridising (Smith, 1967). *Typha angustifolia* and *T. latifolia* share a widespread hybrid zone in North America involving the highly impactful invasive *T.* × *glauca*, which forms dense stands, altering habitats and outcompeting and displacing native plants (reviewed in Bansal *et al.*, 2019).

Here, we examined the genomic diversity and differentiation of five cattail species to evaluate the role of drift, selection, and hybridisation in their speciation history. These comprised *T. angustifolia*, *T. domingensis*, and *T. latifolia*; we also included two closely related *Typha*, *T. laxmannii* Lepech. and *T. shuttleworthii* W. D. J. Koch and Sond (Zhou *et al.*, 2018), which have more restricted ranges (Figure 1). The genomes of 207 plants sampled across multiple continents were sequenced to address the following questions: (1) What are the demographic histories and (2) levels of genetic diversity and divergence of these species? (3) How has selection shaped their genomic landscapes? (4) Has hybridisation led to (adaptive or non-

adaptive) introgressions between *T. angustifolia* and *T. latifolia*? Understanding how *Typha* species have diverged could provide important insights for wetland conservation. Speciation in allopatry (without developing reproductive genetic barriers) could help explain cattails' ability to create invasive hybrids; this, in turn, could highlight potential risks associated with human-mediated dispersal, for example, intercontinental *Typha* sourcing by garden centres (Ciotir and Freeland, 2016).

Materials and Methods

Data preparation

Sampling, DNA extraction, and sequencing

Samples were obtained from a previous study (Aleman *et al.*, 2024) and supplemented with additional collections to increase taxonomic diversity and sampling range (Figure 1; Supplementary Table S1). Plants were identified beforehand using known morphological characteristics (Smith, 1967; Grace and Harrison, 1986) and/or a combination of three to four microsatellite loci (Ciotir *et al.*, 2013, 2017; Ciotir and Freeland, 2016; Tisshaw *et al.*, 2020; Pieper *et al.*, 2020; Bhargav *et al.*, 2022). DNA was extracted following published protocols (Pieper *et al.*, 2017, 2020) and then converted into Nextera XT libraries for reduced-representation genotyping-by-sequencing as per Aleman *et al.* (2024) (Supplementary information – Methods). Paired-end sequencing was conducted on a Miseq (151 bp) and a Novaseq 6000 (126 bp) at The Centre for Applied Genomics (Toronto, Ontario, Canada) for 64 *T. angustifolia*, 25 *T. domingensis*, 104 *T. latifolia*, 11 *T. laxmannii*, and 3 *T. shuttleworthii*.

Raw data processing

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 cleaned reads shorter than 100 bp. All reads

were mapped to the *T. latifolia* nuclear (15 chromosomes, 285.11 Mb, GenBank accession JAIOKV000000000.2) and chloroplast (161.57 kb, GenBank accession NC_013823) genomes with BWA 0.7.17 (Li and Durbin, 2009). Mapping statistics were evaluated using SAMtools 1.15.1 (Li *et al.*, 2009).

Genotype-calling

Genotype calling was performed through ANGSD 0.93 (Korneliussen *et al.*, 2014). Two datasets were created using all samples. SNPs were retrieved requiring a minimum mapping and Phred scores of 20 and a minimum p-value of $1e^{-6}$ (referred to as the *SNP dataset*).

Reconstructing the genomic landscape requires variant and invariant loci; hence, a dataset including SNPs and invariant sites (referred to as the *all-sites dataset*) was created using the same mapping and Phred filters. SNPs with more than 20% missing data and sites mapped to the plastome were removed from the *SNP dataset* with VCFtools 0.1.16 (Danecek *et al.*, 2011). The plastome data were also removed from the *all-sites dataset*—for the coverage requirements in the *all-sites dataset*, see *Role of selection and introgressive hybridisation on species' divergence*. No individuals were removed from any dataset. The absence of clones (multiple ramets from the same genet) was verified by calculating kinship coefficients with Plink 2.0 (Chang *et al.*, 2015) and the *SNP dataset*.

Genetic structure and species' demographic histories

Two approaches were used to investigate the genetic structure and relationships among samples: (1) a principal component analysis (PCA; PC1 to PC3) and (2) a neighbour-joining (NJ) tree based on the samples' pairwise distances (expressed as allele counts, transformed with the R 4.3.1 package ape 5.7-1 (Paradis and Schliep, 2019; R Core Team, 2022) were both produced with Plink 1.9 (Purcell *et al.*, 2007) and the *SNP dataset*.

Species splits and migration events were inferred with TreeMix 1.13 (Pickrell and Pritchard, 2012) and the *SNP dataset*. This analysis tested migration events between 0 and 10, and the optimal number of migrations was chosen based on the data variance. Demographic fluctuations were reconstructed using a Stairway Plot 2.1.2 analysis (Liu and Fu, 2015, 2020), which generates estimates of N_e over generations in the past. Using the *SNP dataset*, a folded site frequency spectrum (SFS) per species was produced in easysfs 0.0.1 (Gutenkunst *et al.*, 2009); given that intraspecific structure was observed in *T. latifolia*, distinguishing two lineages ("Eastern" and "Western", see *Results*), independent SFSs were produced for each of these lineages. The resultant SFSs and the mutation rate of *Arabidopsis thaliana* (7×10^{-9} mutations per site per generation (Weng *et al.*, 2019)) were used as inputs for Stairway Plot.

Role of selection and introgressive hybridisation on species' divergence

To estimate the genetic diversity within and divergence between species, F_{ST} (Weir and Cockerham, 1984), d_{XY} (Nei and Miller, 1990), and π (Nei and Li, 1979) were computed in 5 kb windows for all species pairs using pixy 1.2.7 (Korunes and Samuk, 2021) and the *all-sites dataset.* To ensure data reliability, windows with coverage below 50% (in each pairwise comparison) were discarded. The means for each statistic and the net divergence between species (d_a; Nei and Li (1979)) were calculated.

To test the role of selection on species' genetic differentiation, islands and valleys of divergence consistent with the alternative types of selection proposed by Irwin *et al.* (2018) were identified using the retained windows in each pairwise comparison (see *above*). Individual F_{ST}, d_{XY}, and π values were standardised into Z-scores (e.g., $ZF_{ST} = [(\text{window } F_{ST} - \text{genome-wide median } F_{ST}) / \text{genome-wide } F_{ST} \text{ standard deviation}]$), and windows with exceptionally high or low diversity and divergence were characterised by detecting outliers where the absolute value of a Z-score $\geq$ 1.96; intermediate d_{XY} values were also documented (Z-score between –0.1257 and 0.1257). Following Kessler *et al.* (2023) and Shang *et al.* (2023), islands and valleys of divergence (i.e., positive and negative F_{ST} outliers) were classified into four models of selection using joint combinations of statistics as follows: (i) divergent with gene flow (high F_{ST} and d_{XY}, low π); (ii) divergent without gene flow (high F_{ST}, intermediate d_{XY}, and low π);

(iii) background (high F_{ST}, low d_{XY} and π); and (iv) balancing (low F_{ST}, high d_{XY} and π).

Additionally, Tajima's D (Tajima, 1989) was estimated in 5 kb windows for each species using VCFtools and the *all-sites dataset*. Within each species, only those windows with more than 50 SNPs were kept, and loci with Tajima's D higher than +2 or lower than –2 were interpreted as experiencing balancing selection and selective sweeps, respectively.

ABBA BABA was used to identify introgressions; *T. laxmannii* was designated as outgroup, and all possible combinations of donor and recipient species were examined (Supplementary Table S2). The script ABBABABAwindows.py (Martin and Jiggins, 2017) was used to compute f_d (Martin *et al*., 2015) in 5 kb windows. Two or more loci with a positive f_d, ZF_{ST} and Zd_{XY} $\leq$–1.96, and at least within 50 kb of each other, were classified as introgressions. Introgressions from *T. latifolia* to *T. angustifolia* were observed (see *Results*); thus, ABBA BABA was replicated, splitting data into "North America" and "Europe" to test whether introgressions occur on both continents.

Results

Genetic structure and species' demographic histories

We assembled 21,759,123 nuclear SNPs across 207 samples and kept 77,207 with missing data below 20%. The PCA (PC1 to PC3, Supplementary Figure S1) and NJ tree (Figure 2) distinguished all five species; the NJ tree also revealed intraspecific structure in *T. latifolia* distinguishing two lineages, named "Western" (n = 77; North America, Western and Central Europe) and "Eastern" (n = 27; Eastern Europe and the Iturup Island in the Russian Kuril Chain).

TreeMix indicated *T. angustifolia* – *T. domingensis* and *T. latifolia* – *T. shuttleworthii* as sister lineages, and *T. laxmannii* as outgroup species (Figure 1); this topology was consistent with the NJ tree. The most likely number of gene flow events was determined to be zero, explaining over 99.9% of the data variance. *Typha laxmannii*, inferred by Treemix to be the oldest species, exhibited a stable N_e from ~18 million to ~5 million generations ago, when a sharp decline appeared to coincide with sharp expansions for *T. angustifolia* and *T. domingensis*. Another sharp decline for *T. laxmannii* appeared to coincide with a sharp expansion ~3 million generations ago for Eastern *T. latifolia*. Western *T. latifolia* appeared to experience a cycle of expansion-decline-expansion between ~2 million and ~400,000 generations ago; finally, *T. shuttleworthii* appeared to undergo a sharp expansion ~1 million generations ago. The series of sharp declines

and expansions inferred by Stairway Plot overlapped with the order of species splits revealed by Treemix; moreover, assuming a ~2-year generation time, this series of events coincides with the divergence times of *T. angustifolia – T. domingensis* and *T. latifolia – T. shuttleworthii* in the literature (Figure 3*; see Discussion*).

Role of selection and introgressive hybridisation on species' divergence

Covering 65% of the genome (variant and invariant sites, depth = 4×), mean F_{ST} varied from 0.317 (*T. angustifolia–T. domingensis*) to 0.617 (*T. latifolia–T. laxmannii*), d_{XY} from 0.016 (*T. latifolia–T. shuttleworthii*) to 0.030 (*T. domingensis–T. shuttleworthii*), π from 0.005 (*T. latifolia*) to 0.024 (*T. domingensis*), and d_a from 0.007 (*T. latifolia–T. shuttleworthii*) to 0.015 (*T. latifolia–T. angustifolia*) (Table 1; Figure 2). *Typha latifolia* was the only species with low diversity ($\pi < 0.01$). The net divergence among the five species was low (d_a <0.02).

The genomic landscapes between species pairs showed a heterogeneous role of selection on species' differentiation (Supplementary Figures S3 to S12). Most loci (5 kb) with exceptionally high or low F_{ST} could not be associated with any of the types of selection tested (on average, 1398 out of 1839; Table 2). On average, from the 441 loci that could be associated with selection, balancing selection was the most frequently detected in pairwise comparisons (375 loci, 0.66% of the genome); background selection was the second (46 loci, 0.08% of the genome),

followed by selection without and with gene flow (15 and 6 loci – 0.03%, and 0.01% of the genome, respectively). A prominent island of background selection was observed between *T. laxmannii* and most species in our study (~350 kb, chromosome 14); no genes could be identified in this island.

Except for *T. latifolia*, all species had mean positive Tajima's D per species, from 0.05 (*T. angustifolia*) to 1.55 (*T. laxmannii*) (Figure 4) and more loci with values higher than +2 (~8,129) vs lower than –2 (~43), consistent with the widespread balancing selection observed across the genomic landscapes; *T. shuttleworthii* had no loci with values higher than 2 a mean Tajima's D of 0.66, *T. domingensis* mean was 1.14; *T. latifolia* had a mean negative Tajima's D (–1.27) and a large number of windows lower than –2 (12,252), in line with the most recent expansion observed in its demographic history.

The f_d identified introgressions from *T. latifolia* to *T. angustifolia* on chromosomes 11 and 14 (~145 kb); no genes could be identified in these introgressions. The introgressed regions exhibited values of Tajima's D from –0.77 to –1.96, suggesting positive selection. Splitting the data by continent confirmed that these introgressions are found in North America and Europe (Supplementary Figures S13 and S14). Additional introgressions from *T. latifolia* to *T. domingensis* were identified in chromosomes 8, 14, and 15; no genes were found in these introgressions, and their Tajima's D values ranged from –0.004 to –1.73.

Discussion

During allopatric speciation, genetic divergence primarily accumulates through drift (Abbott *et al.*, 2013); in contrast, ecological (sympatric) speciation requires divergent selection (Rundle and Nosil, 2005; Nosil, 2012). This means that ecological speciation should produce stronger reproductive barriers than geographic isolation, and that speciation in isolation could enable hybridisation if species experience secondary contact (Hewitt, 2000; Rheindt and Edwards, 2011; Sobel, 2016)—although the later development of genetic barriers between species (e.g., through reinforcement (Noor, 1999)) would depend on their history after secondary contact and hybridisation occur (Runemark *et al.*, 2019; Moran *et al.*, 2021). Here, we examined the genomic diversity and differentiation of five cattail species to evaluate the role of drift, selection, and hybridisation in their speciation history. We revealed (1) species splits (without migration events) coincident with past sharp demographic declines, (2) low net genomic divergence between species (d_a from 0.007 to 0.015), (3) predominant valleys of balancing selection between species, and (4) introgressions from *T. latifolia* to *T. angustifolia*—both in North America and Europe, despite the absence of hybrids in the latter—and *T. domingensis*. Our results suggest that demographic contractions are the primary factor responsible for the speciation of these *Typha*, possibly followed by range expansions and secondary contact, enabling hybridisation between some species.

Demographic history and scarcity of divergent selection in Typha

Treemix reported a topology with no migration events, indicating that species split events without gene flow events can account for all the observed genetic variation among the species in our study. While it does not rule out the possibility of hybridisation after secondary contact, this result suggests that genetic drift—without gene flow—was the primary driver of differentiation among the species in this study. Stairway Plot revealed a series of sharp N_e declines and expansions that coincide chronologically with the Treemix topology and the divergence times of Zhou *et al.* (2018). Geoclimatic events, including the Miocene aridification and the Quaternary glacial cycles (Hewitt, 2000; Herbert *et al.*, 2016; Butiseacă *et al.*, 2021), could have led to these bottlenecks, both periods potentially fragmenting the extent and connectivity of wetland habitats. These reductions in N_e likely reflect episodes of geographic isolation that triggered species divergence via genetic drift. Stairway Plot also indicated that *T. latifolia* possibly underwent an expansion-decline-expansion cycle between ~2 million and ~400,000 generations ago, which is in line with the genome-wide negative Tajima's D (–1.765) in this species.

The islands and valleys of divergence associated with divergent selection were scarce, occurring on average in 0.12% of the genome and 2.88% of the loci with exceptionally high or low F_{ST} in pairwise comparisons. This result suggests a negligible role of ecological divergence

in the speciation of the *Typha* in this study. Some signatures of divergent due to selection could have been eroded by time and recombination, which is consistent with the late stages of speciation (Burri *et al.*, 2015)— the age of *Typha* is estimated between 20 and 70 Ma (Zhou *et al.*, 2018; Widanagama *et al.*, 2022)—however, the net divergence (d_a) among these *Typha* species is low (<0.02; 'grey zone of speciation' (Roux *et al.*, 2016)), which suggests that reproductive isolation remains incomplete despite *Typha* being and old genus and the high relative divergence ($F_{ST} > 0.25$ (Wright, 1984)) between species.

Balancing selection and introgressions as sources of genetic variation

Standing genetic variation can be as important as mutations to adaptation (Barrett and Schluter, 2008; Matuszewski *et al.*, 2015). Several mechanisms maintain standing genetic variation, including balancing selection (Llaurens *et al.*, 2017; Fijarczyk *et al.*, 2018). Across the genomic landscapes of the species in our study, most regions associated with adaptive divergence were valleys of balancing selection—as confirmed by Tajima's D results—in which the maintenance of standing polymorphisms causes an excessively reduced divergence (Guerrero and Hahn, 2017). Signatures of balancing selection are congruent with *Typha* spp. ability to self-fertilise; selfing leads to increased homozygosity, so balancing selection can be crucial in preserving polymorphism at important loci (Glémin, 2021). Genes experiencing balancing selection have been shown to underlie a wide

range of phenotypes (Isildak *et al.*, 2021; Promy *et al.*, 2023) and promote adaptation to divergent habitats (Delph and Kelly, 2014; Wu *et al.*, 2017). A predominant role of balancing selection has been detected among other widespread plants, including *Arabidopsis*, *Arbutus*, *Capsella*, *Populus*, and *Quercus* (Santiso *et al.*, 2016; Meireles *et al.*, 2017; Wu *et al.*, 2017; Bachmann *et al.*, 2018; Koenig *et al.*, 2019; Wang *et al.*, 2019, 2020; Rendón-Anaya *et al.*, 2019; Le Veve *et al.*, 2023; Shang *et al.*, 2023). The growing evidence supporting a prominent role of balancing selection (Delph and Kelly, 2014; Fijarczyk and Babik, 2015; Guerrero and Hahn, 2017; Llaurens *et al.*, 2017; Thorburn *et al.*, 2022; Kessler *et al.*, 2023) challenges the view of background selection being the leading force shaping species' genetic diversity (Comeron, 2017).

Introgressive hybridisation is another important source of standing genetic variation (Tigano and Friesen, 2016; Suarez-Gonzalez, Lexer, *et al.*, 2018). Research on the *T.* × *glauca* hybrid swarm in North America has demonstrated the fertility of F1 hybrids in natural populations (Grace and Harrison, 1986; Snow *et al.*, 2010; Travis *et al.*, 2010; Kirk *et al.*, 2011) and their ability to backcross (Pieper *et al.*, 2017; Bhargav *et al.*, 2022). Under these circumstances, introgressions between *T. latifolia* and *T. angustifolia* could be expected, and those identified might be under positive selection. Introgressions were also present in Europe, where these species do not currently hybridise (Ciotir *et al.*, 2017)—suggesting that hybridisation

between these species could have happened in the past. Introgressions have been documented in other widespread plants, including *Helianthus* (Kim and Rieseberg, 1999; Whitney *et al.*, 2006, 2010), *Populus* (Suarez-Gonzalez, Hefer, *et al.*, 2018; Rendón-Anaya *et al.*, 2021), and *Quercus* (Goicoechea *et al.*, 2019; Leroy *et al.*, 2020; Li *et al.*, 2021), and are recognised as essential sources of genetic variation for adaptation.

Genetic diversity in Typha

Genetic diversity reflects the reservoir of traits and potential responses to environmental changes that species encounter, thereby influencing their resilience in different ecosystems (Gregorius, 1987; Frankham *et al.*, 2002; Hartl and Clark, 2006). Excluding *T. latifolia*, species' genetic diversity was high ($\pi > 0.01$ (Begun *et al.*, 2007)), consistent with large population sizes and substantial gene flow. Low genetic diversity is often considered detrimental, as it can reflect a reduced adaptive potential (Teixeira and Huber, 2021; Kardos *et al.*, 2021). However, some species—including invasive species—thrive despite having low genetic diversity (Tsutsui *et al.*, 2000; Roman and Darling, 2007; Charlesworth and Jensen, 2022). *Typha* tolerates wide-ranging climates, nutrients, pHs, pollutants, and water levels (Sojda and Solberg, 1993; Kadlec and Wallace, 2008; Sesin *et al.*, 2021), and the most widespread and commonly recognised cattail species is *T. latifolia*, presumed to have large census sizes, and is considered invasive in Oceania (Xu *et al.*, 2013). Low

genetic diversity in *T. latifolia* suggests that alternative mechanisms, such as epigenetic modifications and associated phenotypic plasticity, may be responsible for the success of this species (Mounger *et al.*, 2021). In *Arabidopsis thaliana*, epigenetic diversity underlies morphological variation and phenotypic plasticity when genetic diversity is low (Zhang *et al.*, 2013; Kooke *et al.*, 2015; Schmid *et al.*, 2018). An inquiry for future research is testing if epigenetic modifications and phenotypic plasticity are more extensive in *T. latifolia* than in other *Typha* species.

Conclusions

Understanding the causes of speciation is a central aspect of evolutionary biology. Here, we tested the roles of drift, selection, and hybridisation in driving species divergence in *Typha*, an old and widespread plant genus that is foundational to freshwater ecosystems. Our results support bottlenecks and geographic isolation as the primary causes of speciation, as well as widespread balancing selection across *Typha*, and introgressions from *T. latifolia* into *T. angustifolia* and *T. domingensis*. The absence of divergent selection does not rule out the role of ecological divergence as a promoter of speciation for the *Typha* in our study—this genic view of speciation requires only a handful of genes (Lexer and Widmer, 2008)—rather, it indicates a primary role of genetic drift. Our results add *Typha* to the body of evidence supporting an important role of balancing selection in preserving standing genetic variation despite the accumulation of divergence (e.g., Selechnik *et al.*, 2019).

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, to whom we show our respect. We thank Polina A. Volkova and Tulsi Patel for their invaluable contributions to the laboratory and the field. The Natural Sciences and Engineering Research Council of Canada financially supported this work, and Alberto Aleman is funded by the Environmental and Life Sciences Graduate Program at Trent University. SHARCNET and Compute Canada provided computational resources for this study. We are grateful to Camille Kessler, Marie-Laurence Cossette, Beibei Zhou, Xinwei Xu, the Associate Editor, and the anonymous reviewers for their valuable feedback, which significantly improved the quality of this manuscript. Finally, we thank Enrique Ruiz for his work in Figure 1.

Author contribution statement

Conceptualisation, developing methods, conducting research, data interpretation, writing, data analysis, and preparation of figures and tables: Aaron Shafer, Alberto Aleman, Joanna Freeland, and Marcel Dorken. All authors contributed to the manuscript and approved its final version.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data archiving

Data and scripts for this study can be found at https://gitlab.com/WiDGeT_TrentU/graduate_theses/-/tree/master/aleman/hdy and https://github.com/al-aleman/totoras_hdy.

Raw sequencing data are available from the authors upon request.

Research Ethics Statement

This research did not require any ethical permissions.

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, *et al.* (2013). Hybridization and speciation. *Journal of Evolutionary Biology* **26**: 229–246.

Aleman A, Dorken ME, Shafer ABA, Patel T, Volkova PA, Freeland JR (2024). Development of genomic resources for cattails (Typha), a globally important macrophyte genus. *Freshwater Biology* **69**: 74–83.

Andersson C (2009). Neogene Climates. In: Gornitz V (ed) *Encyclopedia of Paleoclimatology and Ancient Environments*, Springer Netherlands: Dordrecht, pp 609–612.

Andrew RL, Rieseberg LH (2013). Divergence is focused on few genomic regions early in speciation: incipient speciation of sunflower ecotypes. *Evolution* **67**: 2468–2482.

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

Bachmann JA, Tedder A, Laenen B, Steige KA, Slotte T (2018). Targeted Long-Read Sequencing of a Locus Under Long-Term Balancing Selection in Capsella. *G3 Genes|Genomes|Genetics* **8**: 1327–1333.

Bansal S, Lishawa SC, Newman S, Tangen BA, Wilcox D, Albert D, *et al.* (2019). Typha (Cattail) Invasion in North American Wetlands: Biology, Regional Problems, Impacts, Ecosystem Services, and Management. *Wetlands* **39**: 645–684.

Barrett RDH, Schluter D (2008). Adaptation from standing genetic variation. *Trends Ecol Evol* **23**: 38–44.

Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh Y-P, Hahn MW, *et al.* (2007). Population Genomics: Whole-Genome Analysis of Polymorphism and Divergence in Drosophila simulans. *PLOS Biology* **5**: e310.

Bhargav VV, Freeland JR, Dorken ME (2022). Evidence of hybrid breakdown among invasive hybrid cattails (Typha × glauca). *Heredity* **129**: 195–201.

Bock DG, Cai Z, Elphinstone C, González-Segovia E, Hirabayashi K, Huang K, *et al.* (2023). Genomics of plant speciation. *Plant Communications* **4**: 100599.

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bonanno G, Cirelli GL (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.

Burri R, Nater A, Kawakami T, Mugal CF, Olason PI, Smeds L, *et al.* (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Res* **25**: 1656–1665.

Butiseacă GA, Vasiliev I, van der Meer MTJ, Krijgsman W, Palcu DV, Feurdean A, *et al.* (2021). Severe late Miocene droughts affected western Eurasia. *Global and Planetary Change* **206**: 103644.

Chacón J, Luebert F, Selvi F, Cecchi L, Weigend M (2019). Phylogeny and historical biogeography of Lithospermeae (Boraginaceae): Disentangling the possible causes of Miocene diversifications. *Molecular Phylogenetics and Evolution* **141**: 106626.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**: 7.

Charlesworth B, Jensen JD (2022). How Can We Resolve Lewontin's Paradox? *Genome Biology and Evolution* **14**: evac096.

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**: 137–150.

Ciotir C, Kirk H, Row JR, Freeland JR (2013). Intercontinental dispersal of Typha angustifolia and T. latifolia between Europe and North America has implications for Typha invasions. *Biol Invasions* **15**: 1377–1390.

Ciotir C, Szabo J, Freeland J (2017). Genetic characterization of cattail species and hybrids (Typha spp.) in Europe. *Aquatic Botany* **141**: 51–59.

Comeron JM (2017). Background selection as null hypothesis in population genomics: insights and challenges from Drosophila studies. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**: 20160471.

Dagallier L-PMJ, Condamine FL, Couvreur TLP (2024). Sequential diversification with Miocene extinction and Pliocene speciation linked to mountain uplift explains the diversity of the African rain forest clade Monodoreae (Annonaceae). *Annals of Botany* **133**: 677–696.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, *et al.* (2011). The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158.

Delph LF, Kelly JK (2014). On the importance of balancing selection in plants. *New Phytologist* **201**: 45–56.

Ewels P, Magnusson M, Lundin S, Käller M (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**: 3047–3048.

Feder JL, Egan SP, Nosil P (2012). The genomics of speciation-with-gene-flow. *Trends Genet* **28**: 342–350.

Fijarczyk A, Babik W (2015). Detecting balancing selection in genomes: limits and prospects. *Molecular Ecology* **24**: 3529–3545.

Fijarczyk A, Dudek K, Niedzicka M, Babik W (2018). Balancing selection and introgression of newt immune-response genes. *Proceedings of the Royal Society B: Biological Sciences* **285**: 20180819.

Frankham R, Briscoe DA, Ballou JD (2002). *Introduction to conservation genetics*. Cambridge University Press: Cambridge, UK□; New York.

Garner AG, Goulet BE, Farnitano MC, Molina-Henao YF, Hopkins R (2018). Genomic Signatures of Reinforcement. *Genes (Basel)* **9**: 191.

Glémin S (2021). Balancing selection in self-fertilizing populations. *Evolution* **75**: 1011–1029.

Goicoechea PG, Guillardín L, Fernández-Ibarrodo L, Valbuena-Carabaña M, González-Martínez SC, Alía R, *et al*. (2019). Adaptive Introgression Promotes Fast Adaptation In Oaks Marginal Populations. : 731919.

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.

Gregorius H-R (1987). The relationship between the concepts of genetic diversity and differentiation. *Theoret Appl Genetics* **74**: 397–401.

Guerrero RF, Hahn MW (2017). Speciation as a sieve for ancestral polymorphism. *Molecular Ecology* **26**: 5362–5368.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009). Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *PLOS Genetics* **5**: e1000695.

Han F, Lamichhaney S, Grant BR, Grant PR, Andersson L, Webster MT (2017). Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Res* **27**: 1004–1015.

Harrison RG, Larson EL (2014). Hybridization, Introgression, and the Nature of Species Boundaries. *Journal of Heredity* **105**: 795–809.

Hartl DL, Clark AG (2006). *Principles of Population Genetics*, Fourth Edition, Fourth Edition. Oxford University Press: Oxford, New York.

Herbert TD, Lawrence KT, Tzanova A, Peterson LC, Caballero-Gill R, Kelly CS (2016). Late Miocene global cooling and the rise of modern ecosystems. *Nature Geosci* **9**: 843–847.

Hernández-Hernández T, Miller EC, Román-Palacios C, Wiens JJ (2021). Speciation across the Tree of Life. *Biological Reviews* **96**: 1205–1242.

Hewitt G (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.

Irwin DE, Milá B, Toews DPL, Brelsford A, Kenyon HL, Porter AN, *et al.* (2018). A comparison of genomic islands of differentiation across three young avian species pairs. *Molecular Ecology* **27**: 4839–4855.

Isildak U, Stella A, Fumagalli M (2021). Distinguishing between recent balancing selection and incomplete sweep using deep neural networks. *Molecular Ecology Resources* **21**: 2706–2718.

Kadlec RH, Wallace S (2008). *Treatment Wetlands*. CRC Press.

Kardos M, Armstrong EE, Fitzpatrick SW, Hauser S, Hedrick PW, Miller JM, *et al.* (2021). The crucial role of genome-wide genetic variation in conservation. *Proceedings of the National Academy of Sciences* **118**: e2104642118.

Kessler C, Wootton E, Shafer ABA (2023). Speciation without gene-flow in hybridizing deer. *Molecular Ecology* **32**: 1117–1132.

Kim S-C, Rieseberg LH (1999). Genetic Architecture of Species Differences in Annual Sunflowers: Implications for Adaptive Trait Introgression. *Genetics* **153**: 965–977.

Kirk H, Connolly C, Freeland JR (2011). Molecular genetic data reveal hybridization between Typha angustifolia and Typha latifolia across a broad spatial scale in eastern North America. *Aquatic Botany* **95**: 189–193.

Koenig D, Hagmann J, Li R, Bemm F, Slotte T, Neuffer B, *et al.* (2019). Long-term balancing selection drives evolution of immunity genes in Capsella (M Przeworski, IT Baldwin, and Z Gao, Eds.). *eLife* **8**: e43606.

Kong H, Condamine FL, Yang L, Harris AJ, Feng C, Wen F, *et al.* (2022). Phylogenomic and Macroevolutionary Evidence for an Explosive Radiation of a Plant Genus in the Miocene. *Systematic Biology* **71**: 589–609.

Kooke R, Johannes F, Wardenaar R, Becker F, Etcheverry M, Colot V, *et al.* (2015). Epigenetic Basis of Morphological Variation and Phenotypic Plasticity in Arabidopsis thaliana. *Plant Cell* **27**: 337–348.

Korneliussen TS, Albrechtsen A, Nielsen R (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**: 356.

Korunes KL, Samuk K (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources* **21**: 1359–1368.

Le Veve A, Burghgraeve N, Genete M, Lepers-Blassiau C, Takou M, De Meaux J, *et al.* (2023). Long-Term Balancing Selection and the Genetic Load Linked to the Self-Incompatibility Locus in Arabidopsis halleri and A. lyrata. *Mol Biol Evol* **40**: msad120.

Leroy T, Louvet J-M, Lalanne C, Le Provost G, Labadie K, Aury J-M, *et al.* (2020). Adaptive introgression as a driver of local adaptation to climate in European white oaks. *New Phytologist* **226**: 1171–1182.

Lexer C, Widmer A (2008). The genic view of plant speciation: recent progress and emerging questions. *Philos Trans R Soc Lond B Biol Sci* **363**: 3023–3036.

Li H, Durbin R (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al.* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Li X, Wei G, El-Kassaby YA, Fang Y (2021). Hybridization and introgression in sympatric and allopatric populations of four oak species. *BMC Plant Biology* **21**: 266.

Liu X, Fu Y-X (2015). Exploring population size changes using SNP frequency spectra. *Nat Genet* **47**: 555–559.

Liu X, Fu Y-X (2020). Stairway Plot 2: demographic history inference with folded SNP frequency spectra. *Genome Biology* **21**: 280.

Llaurens V, Whibley A, Joron M (2017). Genetic architecture and balancing selection: the life and death of differentiated variants. *Molecular Ecology* **26**: 2430–2448.

Martin SH, Davey JW, Jiggins CD (2015). Evaluating the Use of ABBA–BABA Statistics to Locate Introgressed Loci. *Molecular Biology and Evolution* **32**: 244–257.

Martin SH, Jiggins CD (2017). Interpreting the genomic landscape of introgression. *Curr Opin Genet Dev* **47**: 69–74.

Matuszewski S, Hermisson J, Kopp M (2015). Catch Me if You Can: Adaptation from Standing Genetic Variation to a Moving Phenotypic Optimum. *Genetics* **200**: 1255–1274.

Mayr E, Mayr AAP of Z (Emeritus) M of CZE, Mayr EW (1963). *Animal Species and Evolution*. Belknap Press of Harvard University Press.

Meireles JE, Beulke A, Borkowski DS, Romero-Severson J, Cavender-Bares J (2017). Balancing selection maintains diversity in a cold tolerance gene in broadly distributed live oaks. *Genome* **60**: 762–769.

Moran BM, Payne C, Langdon Q, Powell DL, Brandvain Y, Schumer M (2021). The genomic consequences of hybridization (PJ Wittkopp, Ed.). *eLife* **10**: e69016.

Morjan CL, Rieseberg LH (2004). How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology* **13**: 1341–1356.

Mounger J, Ainouche ML, Bossdorf O, Cavé-Radet A, Li B, Parepa M, *et al*. (2021). Epigenetics and the success of invasive plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**: 20200117.

Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci U S A* **76**: 5269–5273.

Nei M, Miller JC (1990). A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. *Genetics* **125**: 873–879.

Noor MAF (1999). Reinforcement and other consequences of sympatry. *Heredity* **83**: 503–508.

Nosil P (2012). *Ecological Speciation*. Oxford University Press.

Ottenburghs J, Honka J, Müskens GJDM, Ellegren H (2020). Recent introgression between Taiga Bean Goose and Tundra Bean Goose results in a largely homogeneous landscape of genetic differentiation. *Heredity* **125**: 73–84.

Paradis E, Schliep K (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**: 526–528.

Pickrell JK, Pritchard JK (2012). Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLOS Genetics* **8**: e1002967.

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**: 714–725.

Pieper SJ, Nicholls AA, Freeland JR, Dorken ME (2017). Asymmetric Hybridization in Cattails (Typha spp.) and Its Implications for the Evolutionary Maintenance of Native Typha latifolia. *Journal of Heredity* **108**: 479–487.

Pinho C, Hey J (2010). Divergence with Gene Flow: Models and Data. *Annual Review of Ecology, Evolution, and Systematics* **41**: 215–230.

Promy NT, Newberry M, Gulisija D (2023). Rapid evolution of phenotypic plasticity in patchy habitats. *Sci Rep* **13**: 19158.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, *et al.* (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics* **81**: 559–575.

R Core Team (2022). R: The R Project for Statistical Computing.

Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, Rafajlović M, *et al.* (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology* **30**: 1450–1477.

Rendón-Anaya M, Ibarra-Laclette E, Méndez-Bravo A, Lan T, Zheng C, Carretero-Paulet L, *et al.* (2019). The avocado genome informs deep angiosperm phylogeny, highlights introgressive hybridization, and reveals pathogen-influenced gene space adaptation. *Proceedings of the National Academy of Sciences* **116**: 17081–17089.

Rendón-Anaya M, Wilson J, Sveinsson S, Fedorkov A, Cottrell J, Bailey MES, *et al.* (2021). Adaptive Introgression Facilitates Adaptation to High

Latitudes in European Aspen (Populus tremula L.). *Molecular Biology and Evolution* **38**: 5034–5050.

Rheindt FE, Edwards SV (2011). Genetic Introgression: An Integral but Neglected Component of Speciation in Birds. *tauk* **128**: 620–632.

Roman J, Darling J (2007). Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution* **22**: 454–464.

Roux C, Fraïsse C, Romiguier J, Anciaux Y, Galtier N, Bierne N (2016). Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence. *PLOS Biology* **14**: e2000234.

Rundle H, Nosil P (2005). Ecological speciation - Rundle - 2005 - Ecology Letters - Wiley Online Library. *Ecology letters*.

Runemark A, Vallejo-Marin M, Meier JI (2019). Eukaryote hybrid genomes. *PLOS Genetics* **15**: e1008404.

Santiso X, Lopez L, Retuerto R, Barreiro R (2016). Population Structure of a Widespread Species under Balancing Selection: The Case of Arbutus unedo L. *Front Plant Sci* **6**.

Schluter D (2001). Ecology and the origin of species. *Trends Ecol Evol* **16**: 372–380.

Schmid MW, Heichinger C, Coman Schmid D, Guthörl D, Gagliardini V, Bruggmann R, *et al.* (2018). Contribution of epigenetic variation to adaptation in Arabidopsis. *Nat Commun* **9**: 4446.

Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, *et al.* (2014). Genomics and the origin of species. *Nat Rev Genet* **15**: 176–192.

Selechnik D, Richardson MF, Shine R, DeVore JL, Ducatez S, Rollins LA (2019). Increased Adaptive Variation Despite Reduced Overall Genetic Diversity in a Rapidly Adapting Invader. *Front Genet* **10**: 1221.

Sesin V, Davy CM, Freeland JR (2021). Review of *Typha* spp. (cattails) as toxicity test species for the risk assessment of environmental contaminants on emergent macrophytes. *Environmental Pollution* **284**: 117105.

Shang H, Field DL, Paun O, Rendón-Anaya M, Hess J, Vogl C, *et al.* (2023). Drivers of genomic landscapes of differentiation across a Populus divergence gradient. *Molecular Ecology* **32**: 4348–4361.

Smith SG (1967). Experimental and Natural Hybrids in North American Typha (Typhaceae). *The American Midland Naturalist* **78**: 257–287.

Snow AA, Travis SE, Wildová R, Fér T, Sweeney PM, Marburger JE, *et al.* (2010). Species-specific SSR alleles for studies of hybrid cattails (Typha latifolia × T. angustifolia; Typhaceae) in North America. *American Journal of Botany* **97**: 2061–2067.

Sobel JM (2016). Speciation, Geography of. In: Kliman RM (ed) *Encyclopedia of Evolutionary Biology*, Academic Press: Oxford, pp 183–191.

Sojda RS, Solberg KL (1993). 13.4.13. Management and Control of Cattails.

Stankowski S, Ravinet M (2021). Defining the speciation continuum. *Evolution* **75**: 1256–1273.

Suarez-Gonzalez A, Hefer CA, Lexer C, Douglas CJ, Cronk QCB (2018). Introgression from Populus balsamifera underlies adaptively significant variation and range boundaries in P. trichocarpa. *New Phytol* **217**: 416–427.

Suarez-Gonzalez A, Lexer C, Cronk QCB (2018). Adaptive introgression: a plant perspective. *Biology Letters* **14**: 20170688.

Svedarsky D, Grosshans R, Venema H, Ellis-Felege S, Bruggman J, Ostlund A, *et al.* (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*: 1–14.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.

Taylor SA, Larson EL (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nat Ecol Evol* **3**: 170–177.

Teixeira JC, Huber CD (2021). The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences* **118**: e2015096118.

Thorburn D-M, Sagonas K, Lenz T, Chain F, Feulner P, Bornberg-Bauer E, *et al.* (2022). Systematic genome scans reveal common population-specific signals of balancing selection.

Tigano A, Friesen VL (2016). Genomics of local adaptation with gene flow. *Molecular Ecology* **25**: 2144–2164.

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**: 12091–12103.

Travis SE, Marburger JE, 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**: 7–16.

Tsutsui ND, Suarez AV, Holway DA, Case TJ (2000). Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences* **97**: 5948–5953.

Wang B, Mojica JP, Perera N, Lee C-R, Lovell JT, Sharma A, *et al.* (2019). Ancient polymorphisms contribute to genome-wide variation by long-term balancing selection and divergent sorting in Boechera stricta. *Genome Biology* **20**: 126.

Wang M, Zhang L, Zhang Z, Li M, Wang D, Zhang X, *et al.* (2020). Phylogenomics of the genus Populus reveals extensive interspecific gene flow and balancing selection. *New Phytologist* **225**: 1370–1382.

Weir BS, Cockerham CC (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**: 1358–1370.

Weng M-L, Becker C, Hildebrandt J, Neumann M, Rutter MT, Shaw RG, *et al.* (2019). Fine-Grained Analysis of Spontaneous Mutation Spectrum and Frequency in Arabidopsis thaliana. *Genetics* **211**: 703–714.

Whitney KD, Randell RA, Rieseberg LH (2006). Adaptive Introgression of Herbivore Resistance Traits in the Weedy Sunflower Helianthus annuus. *The American Naturalist* **167**: 794–807.

Whitney KD, Randell RA, Rieseberg LH (2010). Adaptive introgression of abiotic tolerance traits in the sunflower Helianthus annuus. *New Phytologist* **187**: 230–239.

Widanagama SD, Freeland JR, Xu X, Shafer ABA (2022). Genome assembly, annotation, and comparative analysis of the cattail Typha latifolia. *G3 Genes|Genomes|Genetics* **12**: jkab401.

Wolf JBW, Ellegren H (2017). Making sense of genomic islands of differentiation in light of speciation. *Nat Rev Genet* **18**: 87–100.

Woodruff DS (2001). Populations, Species, and Conservation Genetics. In: Levin SA (ed) *Encyclopedia of Biodiversity*, Elsevier: New York, pp 811–829.

Wright S (1984). *Evolution and the Genetics of Populations, Volume 4: Variability Within and Among Natural Populations*. University of Chicago Press.

Wu Q, Han T-S, Chen X, Chen J-F, Zou Y-P, Li Z-W, *et al.* (2017). Long-term balancing selection contributes to adaptation in Arabidopsis and its relatives. *Genome Biology* **18**: 217.

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**: e70728.

Zhang Y-Y, Fischer M, Colot V, Bossdorf O (2013). Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytologist* **197**: 314–322.

Zhou B, Tu T, Kong F, Wen J, Xu X (2018). Revised phylogeny and historical biogeography of the cosmopolitan aquatic plant genus Typha (Typhaceae). *Sci Rep* **8**: 8813.

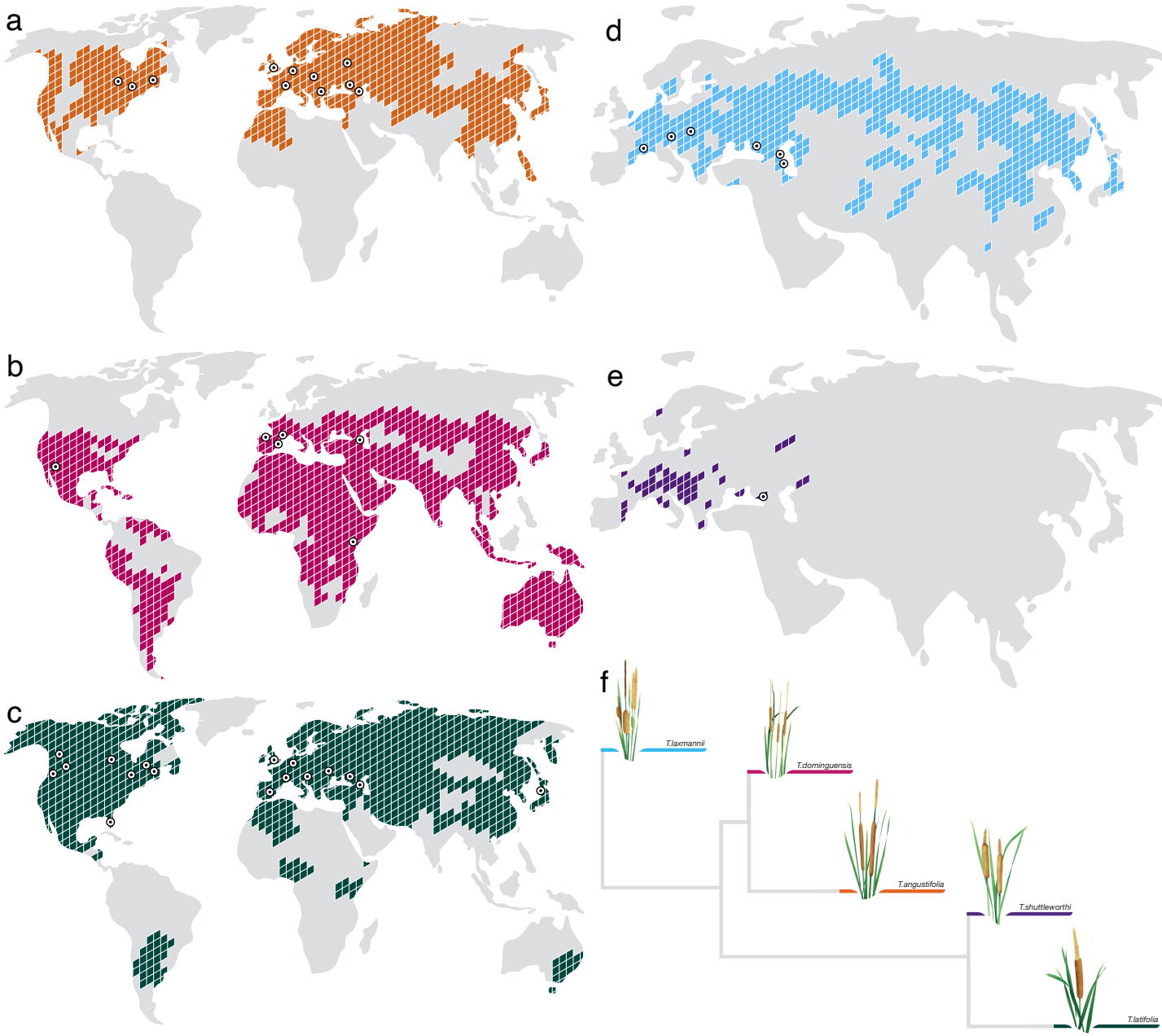
TABLES AND FIGURES

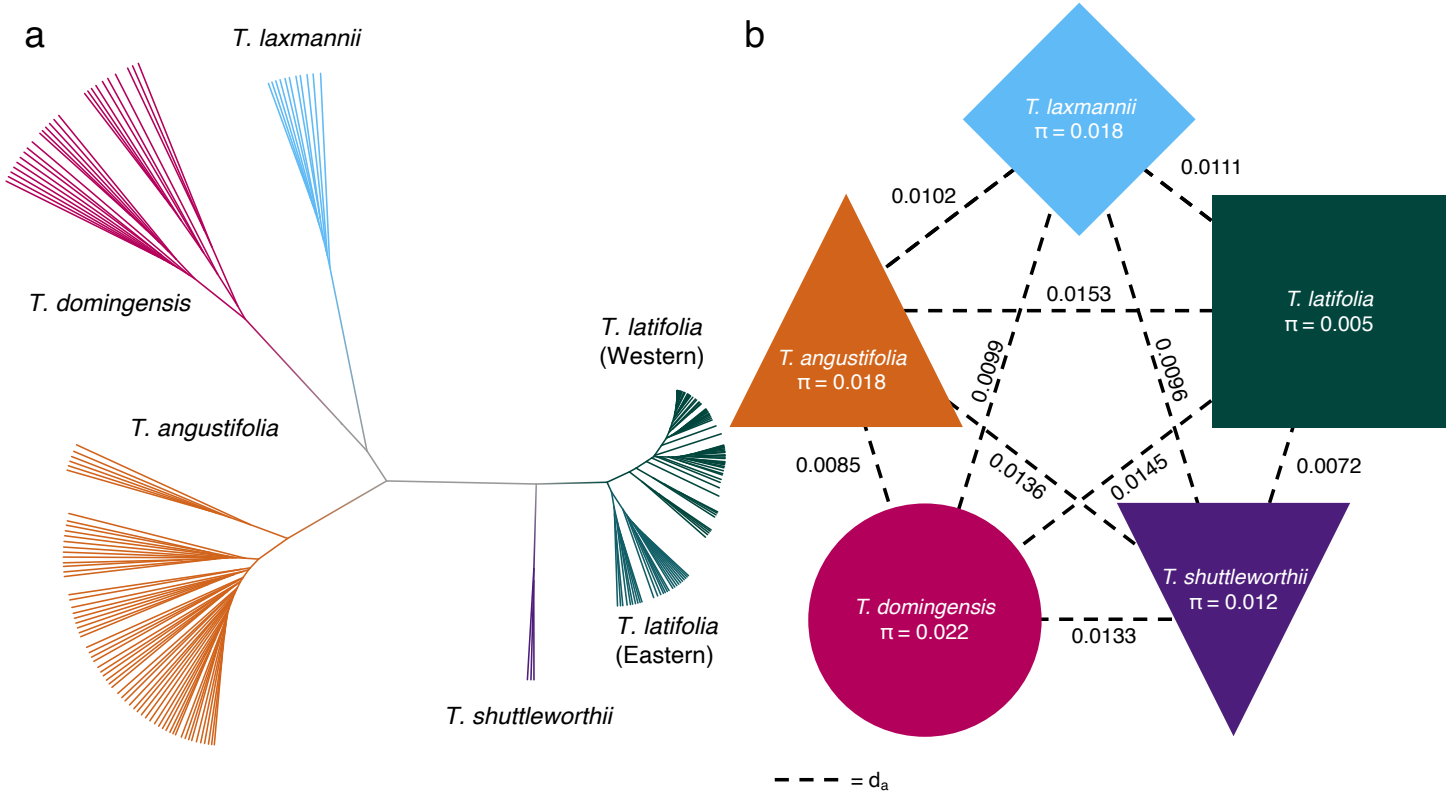
Figure 1. Distributions (coloured grids) and sample sites (black and white circles) in this study for *T. angustifolia* (a), *T. domingensis* (b), *T. latifolia* (c), *T. laxmannii* (d), and *T. shuttleworthii* (e). Distributions from Ciotir and Freeland (2016) and GBIF. Sample sizes are not represented graphically. (f) Species relationships inferred by Treemix; the maximum likelihood tree produced indicated zero migration events. The colour scheme represents species, as in the tree in (f). Note: Shading is based on maps that represent geopolitical boundaries of areas from which species have been recorded.

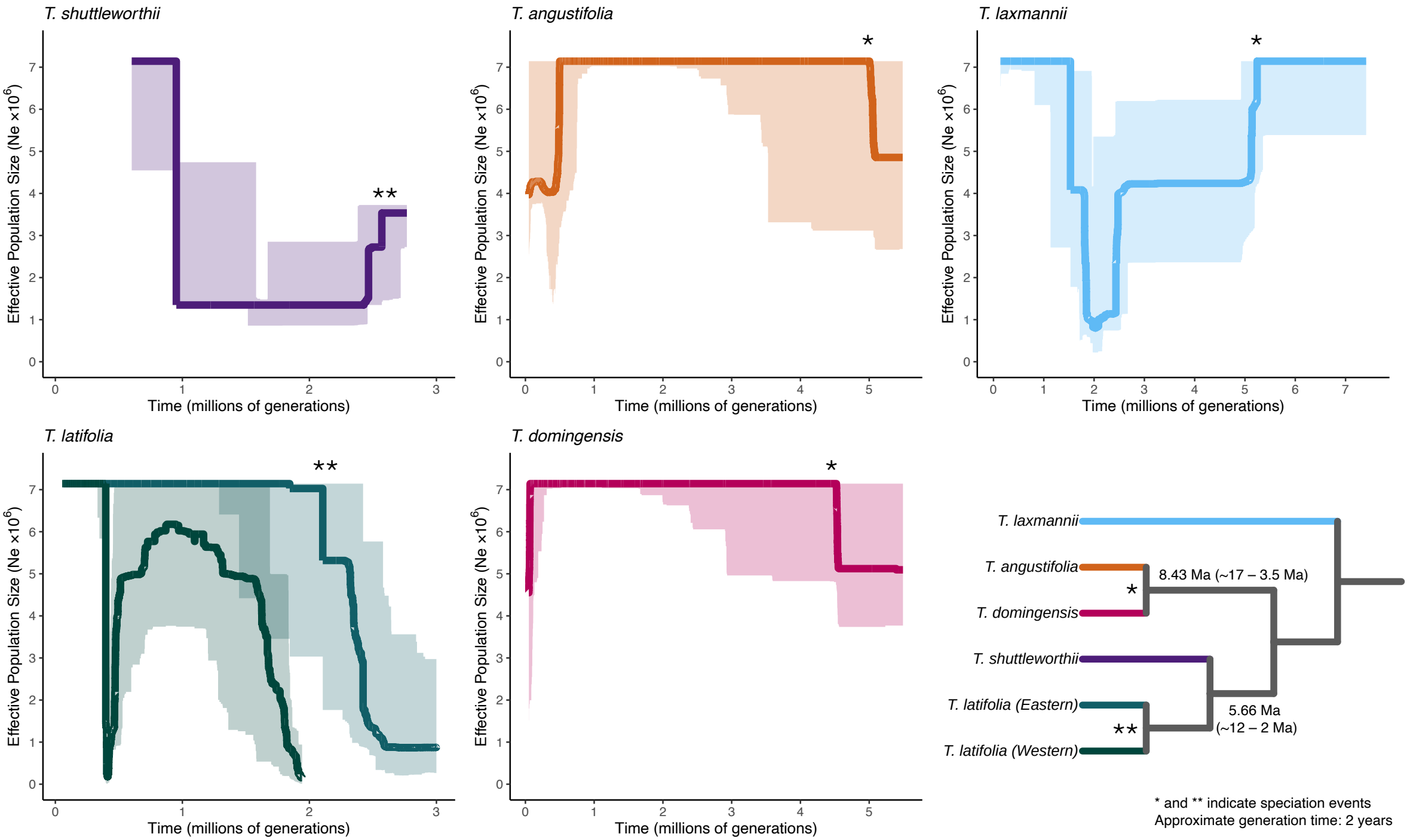
Figure 2. Genetic structure, species relationships, and diversity and divergence for five *Typha* species (a) Neighbour-joining tree. Branches represent individuals; different colours indicate species, as in the labels. (b) Nucleotide diversity (π, inside shapes) for each species and net divergence (d_a, dashed lines) between species. Shapes and colours represent species.

Figure 3. Evolutionary history of the *Typha* in this study. Demographic fluctuations over generations in the past reconstructed with Stairway Plot. The tree (bottom right) is based on the TreeMix results in this study; the dates correspond to Zhou *et al.* (2018). * and ** indicate speciation events.

Figure 4. Genome-wide Tajima's D for five *Typha* species. Points represent 5 kb windows, solid lines represent means, and dashed lines represent 95% confidence intervals. The introgressions from *T. latifolia* to *T. angustifolia* and *T. domingensis* detected by ABBA BABA are highlighted in green.


a
b
c
d
e
f
T.laxmannii
T.dominguensis
T.angustifolia
T.shuttleworthi
T.latifolia


a
T. laxmannii
T. domingensis
T. angustifolia
T. latifolia
(Western)
T. latifolia
(Eastern)
T. shuttleworthii
b
T. laxmannii
π = 0.018
T. angustifolia
π = 0.018
T. latifolia
π = 0.005
T. domingensis
π = 0.022
T. shuttleworthii
π = 0.012
0.0102
0.0111
0.0153
0.0099
0.0096
0.0085
0.0136
0.0145
0.0072
0.0133
= d_a


T. shuttleworthii
T. angustifolia
T. laxmannii
T. latifolia
T. domingensis
Effective Population Size (Ne ×10^6)
Time (millions of generations)
*
**
T. laxmannii
T. angustifolia
T. domingensis
T. shuttleworthii
T. latifolia (Eastern)
T. latifolia (Western)
8.43 Ma (~17 – 3.5 Ma)
5.66 Ma
(~12 – 2 Ma)
* and ** indicate speciation events
Approximate generation time: 2 years

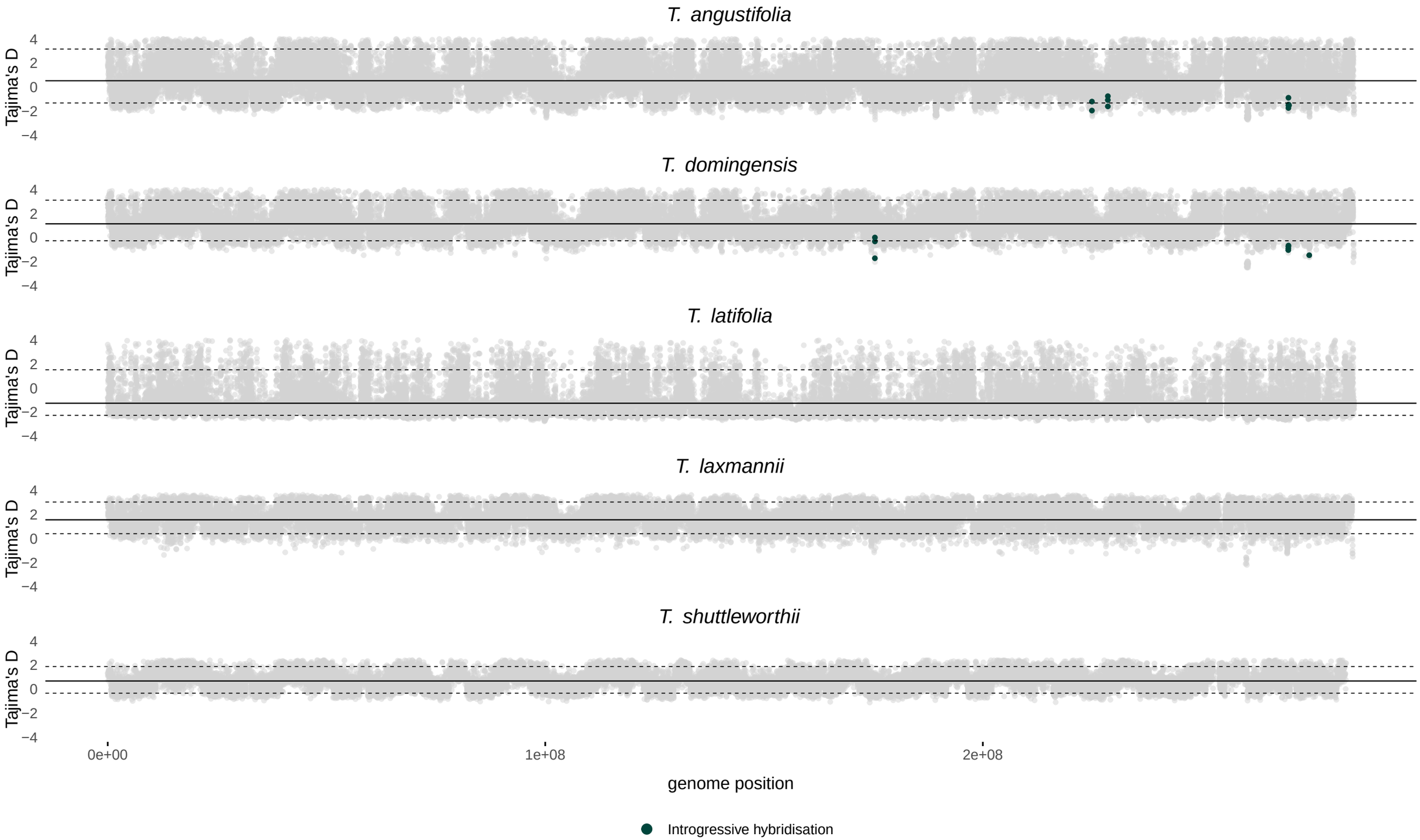

T. angustifolia
T. domingensis
T. latifolia
T. laxmannii
T. shuttleworthii
Tajima's D
4
2
0
−2
−4
0e+00
1e+08
2e+08
genome position
Introgressive hybridisation

Table 1. Mean relative (F_{ST}, below diagonal) and absolute (d_{XY}, above diagonal) genetic divergence between five *Typha* species, measured in 5 kb windows.

	T. laxmannii	*T. domingensis*	*T. angustifolia*	*T. shuttleworthii*	*T. latifolia*
T. laxmannii		0.029	0.028	0.024	0.023
T. domingensis	0.330		0.029	0.030	0.029
T. angustifolia	0.384	0.317		0.028	0.027
T. shuttleworthii	0.380	0.398	0.447		0.016
T. latifolia	0.617	0.614	0.575	0.544	

Table 2. Number of genetic islands and valleys of divergence (5 kb) consistent with alternative types of selection between five *Typha* species: (i) divergent with gene flow, (ii) divergent without gene flow, (iii) background selection, (iv) balancing selection. The number of islands and valleys of divergence that could not be assigned to the alternative types of selection tested is shown in (v).

	(i)	(ii)	(iii)	(iv)	(v)
T. angustifolia - *T. domingensis*	0	23	46	551	1384
T. angustifolia - *T. latifolia*	0	1	10	45	2169
T. angustifolia - *T. laxmannii*	0	18	82	585	996
T. angustifolia - *T. shuttleworthii*	0	15	99	468	1255
T. domingensis - *T. latifolia*	1	5	9	123	1932
T. domingensis - *T. laxmannii*	0	17	79	655	879
T. domingensis - *T. shuttleworthii*	0	5	81	574	1096
T. latifolia - *T. laxmannii*	15	19	7	124	1552
T. latifolia - *T. shuttleworthii*	42	27	1	29	1825
T. laxmannii - *T. shuttleworthii*	0	17	45	591	893