


# Small genome separates native and invasive populations in an ecologically important cosmopolitan grass

PETR PYŠEK <sup>1,2,8</sup>, HANA SKÁLOVÁ,<sup>1</sup> JAN ČUDA,<sup>1,2</sup> WEN-YONG GUO,<sup>1</sup> JAN SUDA,† JAN DOLEŽAL,<sup>1,4</sup> ONDŘEJ KAUZÁL,<sup>1,2</sup> CARLA LAMBERTINI,<sup>5</sup> MAGDALENA LUČANOVÁ,<sup>1,3</sup> TEREZIE MANDÁKOVÁ,<sup>6</sup> LENKA MORAVCOVÁ,<sup>1</sup> KLÁRA PYŠKOVÁ,<sup>1,2</sup> HANS BRIX,<sup>5</sup> AND LAURA A. MEYERSON<sup>7</sup>

<sup>1</sup>Institute of Botany, The Czech Academy of Sciences, Zámek 1, CZ-252 43 Průhonice, Czech Republic

<sup>2</sup>Department of Ecology, Faculty of Science, Charles University, Viničná 7, CZ-128 44 Prague, Czech Republic

<sup>3</sup>Department of Botany, Faculty of Science, Charles University, Benátská 2, CZ-128 00 Prague, Czech Republic

<sup>4</sup>Museum and Gallery of the Orlické hory Mts, Jiráskova 2, CZ-516 01 Rychnov nad Kněžnou, Czech Republic

<sup>5</sup>Department of Bioscience, Faculty of Science, Aarhus University, Ole Worms Alle 1, DK-8000 Aarhus C, Denmark

<sup>6</sup>Plant Cytogenomics Research Group, CEITEC – Central European Institute of Technology, Masaryk University, 625 00 Brno, Czech Republic

<sup>7</sup>Department of Natural Resources Science, The University of Rhode Island, Kingston, Rhode Island, 02881 USA

**Abstract.** The literature suggests that small genomes promote invasion in plants, but little is known about the interaction of genome size with other traits or about the role of genome size during different phases of the invasion process. By intercontinental comparison of native and invasive populations of the common reed *Phragmites australis*, we revealed a distinct relationship between genome size and invasiveness at the intraspecific level. Monoploid genome size was the only significant variable that clearly separated the North American native plants from those of European origin. The mean Cx value (the amount of DNA in one chromosome set) for source European native populations was  $0.490 \pm 0.007$  (mean  $\pm$  SD), for North American invasive  $0.506 \pm 0.020$ , and for North American native  $0.543 \pm 0.021$ . Relative to native populations, the European populations that successfully invaded North America had a smaller genome that was associated with plant traits favoring invasiveness (long rhizomes, early emerging abundant shoots, resistance to aphid attack, and low C:N ratio). The knowledge that invasive populations within species can be identified based on genome size can be applied to screen potentially invasive populations of *Phragmites* in other parts of the world where they could grow in mixed stands with native plants, as well as to other plant species with intraspecific variation in invasion potential. Moreover, as small genomes are better equipped to respond to extreme environmental conditions such as drought, the mechanism reported here may represent an emerging driver for future invasions and range expansions.

**Key words:** biogeography; climate; common reed; plant invasion; source populations; species traits.

## INTRODUCTION

Finding factors that promote invasiveness of species introduced outside their native range has been a key issue in invasion biology since the beginning of the field, and predicting which species will become invasive still represents an ultimate goal of invasion ecologists (Kolar and Lodge 2001, Richardson and Pyšek 2006). A large body of studies addressing this topic have identified a number of species' biological and ecological traits that foster invasive behavior in plants (see Pyšek and Richardson [2007] and van Kleunen et al. [2010] for overviews of such traits). The rapid development of new technologies in recent decades now makes routine screenings of traits possible. As a result, karyological

characteristics can be quantified, including genome size, a trait long suspected to play a role in invasiveness (Rejmánek 1996).

The relationship between plant genome size (i.e., the nuclear DNA amount; Greilhuber et al. 2005) and invasiveness was first recognized for individual genera, such as *Pinus* and *Artemisia* (Rejmánek 1996, Grotkopp et al. 2002, Garcia et al. 2008), and it was shown that many life-history patterns are indirectly, but consistently, associated with genome size (Grotkopp et al. 2004, Meyerson et al. 2016). Another line of evidence comes from analyses of multispecies data sets; several papers demonstrated that naturalized or invasive species tend to have smaller genomes than those that have not successfully naturalized or invaded (Kubešová et al. 2010, Kuester et al. 2014, Pandit et al. 2014, Pyšek et al. 2015).

Kubešová et al. (2010) compared 1C values (holoploid genome size *sensu* Greilhuber et al. 2005) of nearly 100 naturalized alien species in the Czech Republic with their congeners and confamilials not reported to be naturalized or invasive anywhere in the world and found that

Manuscript received 28 February 2017; revised 8 October 2017; accepted 16 October 2017. Corresponding Editor: Brenda J. Grewell.

<sup>8</sup> E-mail: pysek@ibot.cas.cz

† Deceased

naturalized species (i.e., those that established in the Czech Republic, forming self-sustaining populations) were characterized by smaller genomes than their congeners and confamilials that had not successfully naturalized. For 890 species across the angiosperm phylogeny, Pandit et al. (2014) demonstrated that invasiveness was negatively related to genome size and positively to ploidy level and that including both traits improved the explanatory power of the models. Although the two karyological traits are seemingly in conflict, as polyploidy at least initially leads to an increase in the 1C value (Suda et al. 2015), it is their interaction that underlies their actual effects on plant phenotype and physiology, and ultimately on invasion success (Pandit et al. 2014). Another comprehensive analysis comparing the frequency distributions of genome sizes in 242 of the globally most invasive taxa in natural environments with those for all angiosperms revealed that noninvasive plants are less strongly skewed toward small genomes, and that small genomes are significantly over-represented among invasive taxa (Suda et al. 2015). Nonetheless, a study using *Acacia* as a model genus did not find the relationship between genome size and invasiveness (Gallagher et al. 2011).

While small genomes are often typical of species that are invasive, the strongest mechanistic signal emerging from these analyses is that species with large genomes tend to be excluded from being invasive (Suda et al. 2015). According to the “large genome constraint” hypothesis (Knight et al. 2005), species with small genomes can attain a much wider array of trait states compared to species with large genomes, and many traits associated with large genomes are not compatible with the characteristics of successful invaders (Suda et al. 2015).

While the above literature collectively provides robust evidence for this phenomenon, the majority of these studies tested the effect of genome size on invasiveness in isolation (but see Grotkopp et al. 2002, Gallagher et al. 2011, Pyšek et al. 2015, Meyerson et al. 2016). Most did not consider the effects of biological species traits through which the karyological characteristics manifest (Suda et al. 2015) or the confounding factors known to affect the likelihood of a successful invasion, such as propagule pressure, habitat legacy, and residence time (Richardson and Pyšek 2006). However, when genome size was tested together with other traits known to promote invasiveness (Pyšek and Richardson 2007, van Kleunen et al. 2010) in a model that also took into account the potentially confounding factors mentioned above, genome size was one of the variables that explained the naturalization success of central European plant species in North America (Pyšek et al. 2015).

Here, focusing on common reed, *Phragmites australis*, we use conditional inference trees (Hothorn et al. 2006) and Multiple Indicators Multiple Causes models (MIMIC; Hancock 2001, Ríos-Bedoya et al. 2009) to test if genome size plays a role in invasiveness, and if so, whether the effects of genome size persist in interaction with other traits. We included traits covering a range of

plant functions relating to growth, reproduction, physiology, karyology, tissue chemistry, and herbivory and tested their associations with genome size because previous research has shown that traits representing the above functional groups play a role in plant invasiveness (Pyšek and Richardson 2007, van Kleunen et al. 2010). Our paper provides novel insights into this topic because (1) by focusing on differences between intraspecific genotypes of different phylogeographic populations rather than on different species, we investigate the concept of genome-size-related invasiveness in a phylogenetically controlled framework that eliminates biases associated with variation in the mode of reproduction, geographic distribution, and human preferences; (2) we assess the role of genome size in invasiveness in concert with a wide array of experimentally measured functional traits to parse the role of genome size, plant traits, and their interactions on invasiveness (Küster et al. 2008, Pyšek et al. 2009, 2015). Though one previous study using the grass species *Phalaris arundinacea* examined the role between genome size and invasiveness at the intraspecific level (Lavergne et al. 2010), it used many fewer populations. Moreover, (3) by studying three *P. australis* groups of differing phylogeographic origin, we simultaneously test the ecological and biogeographical implications of the role of genome size; the ecological comparison is that of native North American populations growing sympatrically with invasive populations introduced from Europe, and the biogeographical insights are gained by comparing the populations from the native range in Europe with those populations invading North America. This approach allows for testing the changes in genome size and plant traits that occurred over at least 200 years of invasion (Burk 1880, Saltonstall 2002). Finally, (4) we discuss whether genome size can be used as both an indicator of invasiveness and for screening potentially invasive populations (Pyšek et al. 2013, Suda et al. 2015).

## METHODS

### *Study species*

*Phragmites australis* (Cav.) Trin. ex Steud. (common reed, Poaceae) is a tall, helophytic, wind-pollinated perennial grass with shoots up to 4 m tall, forming an extensive system of rhizomes and stolons (runners; Haslam 1972), with a single inflorescence developing on each fertile stem, producing 500–2,000 seeds (McKee and Richards 1996). The species exhibits great genetic, karyological, and morphological variation. It belongs to one of the most ploidy-variable species known, with published cytotypes from  $4x$  to  $12x$ , based on  $x = 12$  (te Beest et al. 2012; see also Clevering and Lissner 1999, Lambertini et al. 2006), and there is marked intraspecific variation in genome size, as well as phylogeographic genetic diversity within the species and the whole genus (Saltonstall 2011, Lambertini 2016).

*Phragmites australis* colonizes a wide range of environmental conditions (Meyerson et al. 2000, 2010) and extends from the tropics to cold temperate regions in both hemispheres, which places it among the world's most cosmopolitan and globally important wild plants providing ecosystem services (Tucker 1990). In its confirmed introduced range, which for the European native *P. australis* subsp. *australis* is North America (Saltonstall 2002), it is a noxious invader that has converted botanically diverse wetlands into low-diversity ecosystems and it outcompetes the North American native *P. australis* subsp. *americanus* (Meyerson et al. 2000). Since the seminal paper by Saltonstall (2002) that provided DNA markers to distinguish the native North American populations and invasive populations introduced from Europe, a significant body of research has emerged to reveal that the patterns observed are results of multiple introductions (Meyerson and Cronin 2013, Lambertini 2016). Genetically, the introduced North American populations today differ from their ancestral European populations in nuclear DNA amount (Lambertini et al. 2012), corroborating the fact that the introduced populations have evolved since their separation (Guo et al. 2014). Because of multiple introductions, long-distance dispersal of seed by wind and birds, and intensive use and transport by humans, there are high intrapopulation genetic distances in both ranges (Lambertini et al. 2006, 2008), which makes the exact source populations of the introduced genotypes to North America very unlikely to be traced. While a detailed discussion is beyond the scope of this paper, several papers provide a comprehensive overview (Lambertini et al. 2006, Lambertini 2016).

#### Population sampling

We used *P. australis* clones representing distinct populations obtained from the collections of the University of Rhode Island, Kingston, Rhode Island, USA and University of Aarhus, Denmark, and also included some field-collected clones. The collection was established in 2011 and rhizomes were first propagated in sand in plastic pots of dimensions 30 × 30 × 20 cm. The clones for the experiment were chosen based on chloroplast DNA following Saltonstall (2002) with the aim of achieving representative coverage of native European populations ( $n = 21$ ), those that are invasive in North America ( $n = 17$ ) and North American native ( $n = 19$ ; see Appendix S1: Table S1 for details on clones used in the experiment).

#### Experimental set up

The experiment was performed in the Experimental Garden of the Institute of Botany CAS in Průhonice, Czech Republic (49°59'38.972" N, 14°33'57.637" E), 320 m above sea level in the temperate climate zone, with a mean annual temperature of 8.6°C and precipitation of 610 mm. The clones were grown in round pots 60 cm in diameter at the top, 36 cm in height (effective pot size

80 L), filled with sand, and mixed with 480 g of slow-release fertilizer Osmocote Pro (release time 12–14 months; ICL Specialty Fertilizers, Geldermalsen, The Netherlands). A piece of rhizome (standard size of 20–30 cm) with an emerging stem was planted in each pot on 7–8 July 2012. Two to six replicates per clone, depending on the availability of the plant material and early survival were used, giving a total of 273 pots with experimental plants. The plants were regularly watered using tap water delivered by an automatic watering system (Hunter Industries, San Marcos, Texas, USA). To ensure comparable water supply to all plants, three holes were drilled in each pot 25 cm from the bottom to allow drainage of excessive water and achieve the same water level in each pot. If plants started to exhibit signs of iron-deficiency (yellowing), 0.2 g Fe as iron in chelation complex of DTPA dissolved in 300 mL of tap water was added to each pot. Plants grew until full senescence (November) and, after the aboveground biomass was harvested, the pot surface was covered with spruce brushwood and the pot sides wrapped with bubble foil to protect the plants from frost. In early April, the frost protection was removed, and 200 g of Osmocote and 0.2 g Fe added to each pot, and an addition of the same Fe dose was repeated in May/June. The experiment was terminated after 2.5 years, in autumn 2014. For harvest, the aboveground biomass was cut about 3 cm above the sand surface, when shoots were senescent (tan) in late November/early December of 2012 and 2013. During the final harvest in October 2014, belowground biomass was also harvested, excavated from the substrate, rinsed and separated into roots and rhizomes, and the length of each was measured. The root and rhizome biomass was oven-dried at 60°C and weighed in the same way as for aboveground biomass.

#### Traits measured

During the experiment, we recorded information about a number of traits that can be organized into several functional groups related to *growth* (dynamics of shoot number and their maximum height over the growing seasons; presence and total length of creeping shoots/runners; aboveground dry biomass at the end of each growing season and belowground biomass at final harvest, separated into rhizomes and roots; total length of rhizomes in the pot), *reproduction* (flowering intensity expressed as the percentage of flowering shoots over the growing season; the biomass of panicles and proportional allocation of biomass to generative reproduction; total seed number per pot, sorted under microscope into potentially viable and damaged seed, with derived variable seed production referring to whether or not the clone produced seed at all), *physiology* (leaf water content; specific leaf area; leaf area per shoot; leaf area per pot calculated as a product of average leaf area per shoot and shoot number; leaf toughness measured as the force necessary to penetrate the leaf; maximum photosynthesis measured on fully developed top leaf; frost damage scored by using a four degree scale

following an unexpectedly late severe frost in the spring of 2014), *herbivory* (ambient aphid attack recorded as an index derived from the estimate of the percentage of infested leaves and average number of aphids per leaf; stem galls attack estimated as a percentage of infested shoots; leaf damage by chewers derived from the percentage of shoots in the pot with chewing damage and percentage of leaves damaged per shoot), and *tissue chemistry* (C, N, P, and total phenolic content determined separately in leaves, roots, and rhizomes, and used to calculate C:N ratio for each tissue). For a complete list of individual variables, detailed descriptions of the traits measured, and timing of sampling see Appendix S2: Table S1.

The traits related to karyology included holoploid (1C value; the amount of DNA in the whole chromosome complement of the nucleus) and monoploid (1Cx value; the amount of DNA in one chromosome set) genome size *sensu* Greilhuber et al. (2005). Nuclear genome size was determined by DNA flow cytometry using a Sysmex-Partec CyFlow SL instrument equipped with green (532 nm, 100 mW output power) solid state laser (Sysmex Partec GmbH, Görlitz, Germany). Sample preparation followed the simplified two-step procedure using Otto buffers as detailed in Doležal et al. (2007). *Bellis perennis* (2C = 3.38 pg; Schönswetter et al. 2007) was chosen as an appropriate internal reference standard (with close but not overlapping genome size for *Phragmites* accessions). A fully developed intact leaf of the analysed *Phragmites* plant (~0.5 cm<sup>2</sup>) and an appropriate volume of the internal standard were chopped with a sharp razor blade in a Petri dish containing 1 mL of ice-cold Otto I buffer (0.1 mol/L citric acid, 0.5% Tween 20). The suspension was filtered through a 42-µm nylon mesh and incubated for 20 min at room temperature. Samples were stained with 1 mL of Otto II buffer (0.4 mol/L Na<sub>2</sub>HPO<sub>4</sub> · 12 H<sub>2</sub>O) supplemented with β-mercaptoethanol (2 µL/mL) and propidium iodide + RNase IIA (both at concentrations 50 µL/mL). Samples were stained for 10 min at room temperature and analysed with the flow cytometer. Fluorescence intensity of 5,000 particles was recorded by the same operator. Only histograms with both peaks of similar height (the smaller peak must have been at least 70% as high as the higher peak) and coefficients of variation of G<sub>0</sub>/G<sub>1</sub> (quiescence/post mitotic gap phase of the cell cycle) peaks of both sample and standard below 3.0% were considered. Each plant was re-estimated at least three times on different days to minimize potential random instrumental drift. If a between-day variation exceeded 2%, the most outlying value was discarded and the sample reanalyzed. Accessions with intraspecific genome size variation above 4.0% were run simultaneously to confirm the divergence. The genome size measurements were complemented by counting chromosomes in a subset of the clones (see Appendix S3).

#### *Climatic variables*

To express climate in the geographic area of the clone origin we used 19 bioclimatic variables available through

the WorldClim database to create three principal component axes (PCAs) that explain nearly 85% of the variation present in the original 19 variables, following the procedure described by Dupin et al. (2011; database *available online*).<sup>9</sup> The original variables included minimum, maximum, and mean temperature and precipitation data for the past 50 years broken out in biologically relevant ways (e.g., precipitation seasonality or mean temperature during the wettest or driest quarters; see, Hijmans et al. 2005). The resulting PCAs represent three uncorrelated linear combinations of the original data, with the first one (PCA1) attributed mainly to mean annual temperature, second (PCA2) mainly to precipitation during wet or warm periods, and third (PCA3) to precipitation during drought (Dupin et al. 2011), see Appendix S1: Table S1.

#### *Statistical analysis*

To evaluate the role that the plant traits played in determining the invasiveness of *P. australis* populations, including interactions among the traits, we first imputed missing data via a multiple imputation process in the Amelia II package (Honaker et al. 2011), and averaged the data per clone. To identify the important predictors for classifying a clone in the three a priori defined groups (European native, North American invasive, and North American native), we used a conditional inference tree, CIT (Hothorn et al. 2006). This classification tree model is more robust than the linear models with nonlinearity and interactions between explanatory variables, is easy to interpret (Breiman et al. 1984, De'ath and Fabricius 2000), and includes statistical correction to avoid overfitting and bias compared to a traditional decision tree (Hothorn et al. 2006, see Appendix S4 for details).

Because we were interested in revealing how the variables identified by the classification tree as significant interact with other variables that do not improve the predictive power of the CIT but may express their influence on invasiveness indirectly through other traits, we employed Structural Equation Modelling (SEM). SEM is a method of representing causal relations between variables via a set of equations (Grace et al. 2010, 2012). Among the 44 traits, some were highly correlated (Appendix S2: Table S1, Fig. S1), and some may not be important for the classification of the groups. Hence, we identified the trait variables to be used in SEM via variable selection. Variable or feature selection is a common procedure in classification and regression to select the most relevant variables in high-dimensional models, especially when the number of predictor variables is higher than sample size, which was the case of our analysis (Guyon 2003, Wasserman and Roeder 2009). Several variable-selection models were carried out to detect the variables most relevant for explaining the classification

<sup>9</sup><http://www.worldclim.org>

of the clones into the three groups, including random forest (RF; Breiman 2001), Boruta (Kursa and Rudnicki 2010), conditional inference forests (CIF; Strobl et al. 2007, 2008, 2009), and gradient boosting models (GBM; De'ath 2007, Elith et al. 2008). These four algorithms performed quite well (Appendix S2: Figs S2, S3), and provided remarkably similar results (Appendix S2: Table S2). The correct classification rate of the RF model increased with the number of input variables and remained steady after including the first 15 variables (Appendix S2: Fig. S4, S5). The traits measured in the experiment were selected so as to cover a wide range of plant functions relating to climate, growth, reproduction, physiology, karyology, tissue chemistry, and herbivory (see Appendix S2: Table S1 for more information), and variables representing some of these trait groups appeared less relevant based on the exploratory variable selection procedure. Therefore, to reflect the scope of the study in its entirety, some variables representative of the above trait groups were included, especially those previously reported as having an effect on plant invasiveness, such as photosynthetic capacity or specific leaf area (SLA; Pyšek and Richardson 2007, Mozdzer and Zieman 2010, van Kleunen et al. 2010, Guo et al. 2014, Pyšek et al. 2015). For traits that partly overlapped (such as biomass or height sampled at different times of the year), we followed the selection methods. These screening procedures yielded a reduced data set with 15 variables for further analysis: monoploid genome size (representing *karyology*); spring shoot height, below/aboveground biomass, rhizome length, presence of runners (*growth*); reproductive allocation (*reproduction*), SLA, frost damage, photosynthetic capacity (*physiology*); aphid attack, gall attack (*herbivory*); rhizome C/N, rhizome P (*tissue chemistry*); and climate PCA3 (see Appendix S2: Table S1 for their description and sampling details).

Previous studies suggested that the leaf or whole-plant traits in general showed a multidimensional correlation or a correlation network (Osnas et al. 2013, Poorter et al. 2014). Our pilot analysis based on the data presented here also revealed that the causal-effect relationships between the traits were impossible to find and that the small sample size of each group made the use of the standard SEM inappropriate. Instead, a MIMIC model, a special case of SEM, was used (Gallo et al. 1999, Giles 1999, Hancock 2001, Ríos-Bedoya et al. 2009). The MIMIC model consists of two parts: the multiple indicators part is a measurement model that defines the correlations between a latent variable and its indicators, and the multiple-causes part is a structural model that specifies the causal relationships between a latent variable and covariates (observed predictors; Jöreskog and Goldberger 1975). The MIMIC model captures the advantages of the standard SEM model, i.e., the modelling of the measurement error, but also allows for the simultaneous detection of associations between the latent variables and among the indicators and the covariates; it also works well with a small sample size (Hancock 2001).

Here we used the MIMIC model to investigate both the effect of climate in the region of the clone origin to capture the possible effect of local adaptation and the effect of genome size on the “performance” of the clones. In our study the latent variable “performance” integrates the indirect effects of the measured traits on invasiveness of the *P. australis* clones because it describes how these traits are related to genome size. Genome size can be considered, based on the CIT analysis, as a proxy for invasiveness because it separates the invasive clones from the non-invasive. Several criteria were used to test how well the model fit the data, i.e., the  $\chi^2$ -test, the root mean squared error of approximation (RMSEA), *P* value of close fit ( $p_{\text{close}}$ ), and the comparative fit index (CFI). All of the analyses were performed in R version 3.2.4 (R Core Team 2016). For further details, see Appendix S4.

## RESULTS

A conditional inference tree identified monoploid genome size as the only important, highly significant variable (Fig. 1;  $P < 0.001$ ), and most remarkably, clearly separated the North American native group from the European native and European invasive in North America at a value of 0.53 pg (Fig. 1). The exceptional importance of monoploid genome size was also evident from all four variable-selection models (Appendix S2: Table S2). The mean Cx value for European native was  $0.490 \pm 0.007$  pg (mean  $\pm$  SD,  $n = 21$ ), for European invasive in North America  $0.506 \pm 0.020$  pg ( $n = 17$ ) and for North American native  $0.543 \pm 0.021$  pg of DNA ( $n = 19$ ), and pairwise comparisons revealed significant differences between all three pairs ( $P < 0.05$  for the European native and North American invasive pair,  $P < 0.001$  for the other two pairs; Fig. 2).

The MIMIC model for the three groups had a good fit:  $\chi^2 = 58.76$ ,  $df = 62$ ,  $P = 0.593$ , RMSEA = 0.000 (90% CI: 0.000–0.072),  $p_{\text{close}} = 0.83$ , CFI = 1.000, indicating that it is a very good approximation to the data. The covariate monoploid genome size had a significant negative effect on the performance, meaning that the larger the genome, the lower the performance demonstrated by the groups. The third climate axis, representing precipitation in dry season, had no effect on performance (Fig. 3). The results of an additional pairwise MIMIC model comparing North American native and invasive clones were similar to those of the complete three-group model, with genome size driving the differences in performance (Appendix S2: Fig. S6a).

Among the traits representing indicators in the MIMIC models, some were closely correlated with the latent variable “performance,” that integrates the indirect effects of the measured traits. The clones producing long rhizomes, tall shoots in spring that are abundant at the end of the growing season, are, via the correlation with performance, most strongly associated with small genomes, while those heavily attacked by aphids, producing more runners and showing a high rhizome C:N ratio

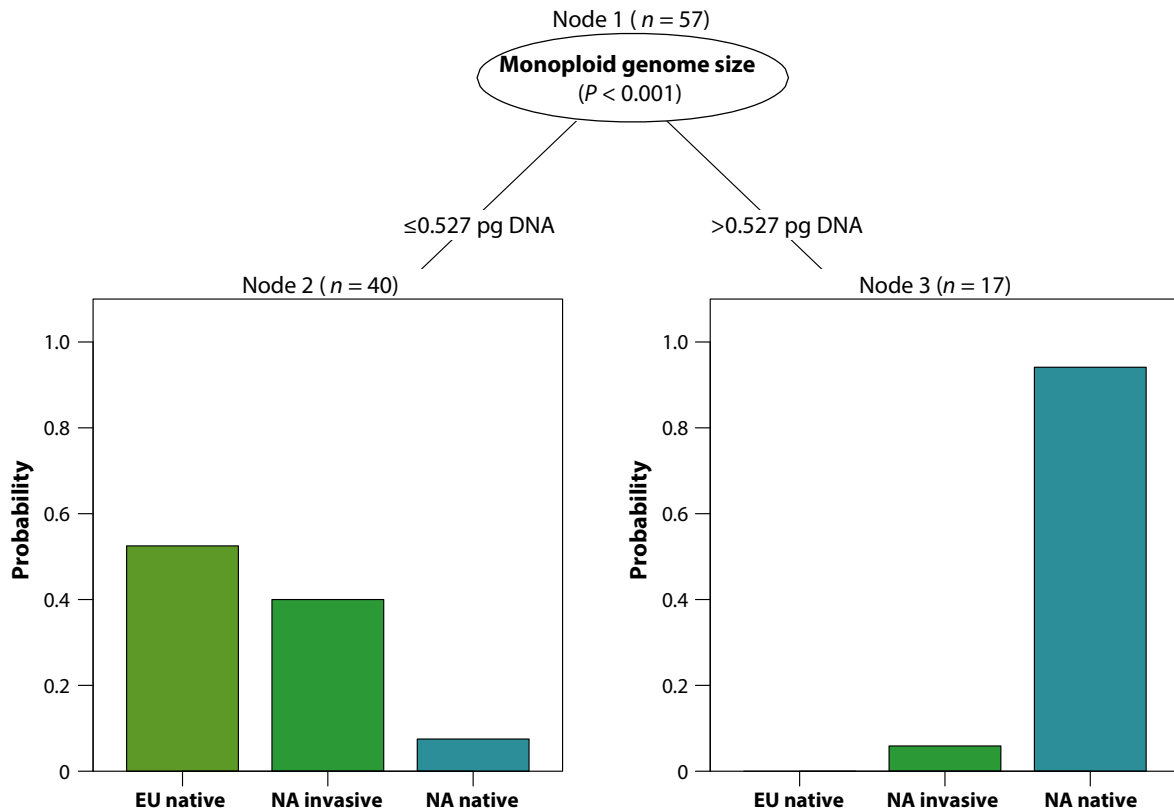


FIG. 1. Conditional inference tree to predict the groups. The histograms under each leaf show the proportion of observations that fall into each group at the terminal nodes. EU, European; NA, North American.

are associated with large genomes (Fig. 3, Appendix S2: Figs. S6a, S7).

Based on boxplots and a density plot, the clones invasive in North America had significantly bigger genomes, higher rhizome C:N, allocated more biomass to reproduction, and produced more runners than European natives, but also had lower belowground:aboveground biomass ratios (Appendix S2: Fig. S7). The MIMIC model for these two groups also fit quite well (Appendix S2: Fig. S6b), but neither monoploid genome size nor the third climate axis had significant effects on performance, probably due to the small difference and great overlap of genome sizes between the European native and North American invasive groups (Fig. 2). However, there was a significant positive effect of the third climate axis on genome size indicating that smaller genomes originated from drier areas (Appendix S2: Fig. S6b).

#### DISCUSSION

Genome size constrains many functional traits related to growth, reproductive success, and dispersal. Due to its effects on cell size parameters and cell division rates, genome size also affects size- and rate-dependent traits

(Suda et al. 2015). Of particular importance to invasion potential are the relationships between genome size and minimum generation time (Bennett 1972, Leitch and Bennett 2007), seed characteristics (Grotkopp et al. 2004, Beaulieu et al. 2007b), seedling relative growth rate (Grotkopp et al. 2004), specific leaf area (Morgan and Westoby 2005, Beaulieu et al. 2007a), and stomatal size and density regulating water use and photosynthetic efficiency (Beaulieu et al. 2008, Hodgson et al. 2010).

Our paper experimentally confirmed the relationship between genome size and invasiveness at the intraspecific level by comparing multiple native and invasive populations of *Phragmites australis*. We strengthened prior studies that investigated the relationship between genome size and invasiveness, by testing individual genotypes of a single species (allowing phylogenetic control) instead of comparing multi-species data sets or species within genera (Rejmánek 1996, Grotkopp et al. 2002, 2004, Garcia et al. 2008, Kubešová et al. 2010, Gallagher et al. 2011, Pandit et al. 2014). This is the most appropriate approach to make progress in understanding the mechanisms of invasion since populations, not entire species, invade (Richardson and Pyšek 2012). Only one previous study has examined whether there is a population effect of genome size on invasiveness.

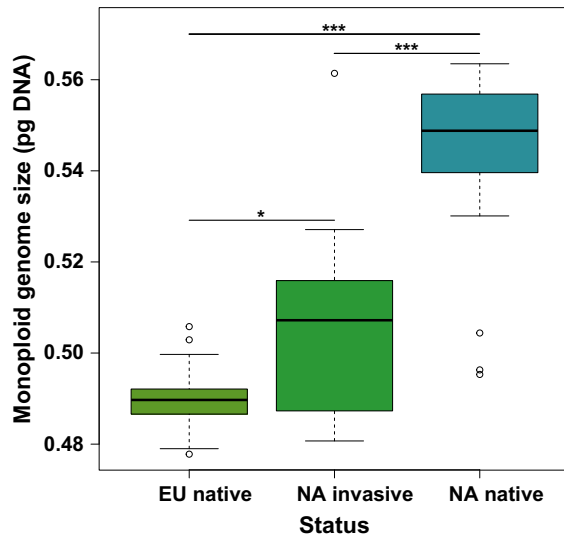


FIG. 2. Boxplot representing the monoploid genome size (1Cx value; the amount of DNA in one chromosome set) of different groups. Data are shown as mean (heavy line)  $\pm$  the interquartile range (boxes),  $\pm 1.5$  times the interquartile range (whiskers), and outliers (circles).  $N = 21$  for native European populations (EU native), 17 for invasive North American population (NA invasive), and 19 for native North American population (NA native), respectively. One-way ANOVA;  $F_{2,54} = 48.26$  and  $P < 0.001$ . A Bonferroni post-hoc test revealed significant differences between groups, represented by asterisks: \*  $P < 0.05$ ; \*\*\*  $P < 0.001$ .

Lavergne et al. (2010) investigated whether North American invasive genotypes of the grass *Phalaris arundinacea* had smaller genomes than the European native genotypes from which they were derived. However, the differences in genome size between native and invasive *Phalaris* genotypes were much less pronounced (below 2%) than for *Phragmites australis* in our study.

Moreover, the natural experiment created by the historical intercontinental introductions of *P. australis* provided an opportunity to address not only the ecological differences between native North American populations and invasive populations introduced from Europe, but also to explore post-introduction evolutionary change in the latter. We accomplished this by comparing North American invasive populations originating from Europe with their ancestors that still occur as native in Europe. It has been suggested that *P. australis* populations invading North America differ from European populations in some traits, potentially indicating preadaptation and post-introduction evolution (Guo et al. 2014, Cronin et al. 2015). The naturally occurring distribution of three lineages in Europe and North America therefore facilitates a biogeographic, ecological, and presumed evolutionary assessment of the association of genome size with invasiveness (Cronin et al. 2015).

Populations of European origin that are currently invasive in North America have slightly (by 3.2% on average) but significantly ( $P < 0.05$ ) bigger genomes

than their ancestors native to Europe. This could point to a non-random filtering of bigger genomes from the European source pool during the introduction process or post-introduction interbreeding. Interestingly, this is the opposite of the genome reduction by natural selection during the invasion process suggested for *Phalaris arundinacea* by Lavergne et al. (2010). However, these authors based their study on six populations in each range, fewer than the *Phragmites australis* study reported here. One explanation for the discrepancy between the two studies may be the different genome sizes of these two grass species. While the mean 1C value of *Phalaris arundinacea* was 4.63 pg (i.e., the intermediate genome size category as defined by Leitch et al. [1998]), the mean 1C value of the most common (4x) cytotype of *Phragmites australis* in our study was 1.01 pg (i.e., a very small genome). Because of its low DNA content, constraints imposed by large genomes (Knight et al. 2005) are likely to play less important role in *P. australis* than in *P. arundinacea*, providing less opportunity for reduction of genome size during the process of invasion.

Therefore, within *P. australis* where invasive genotypes in North America have bigger genomes than their European ancestors, there is no indication of the selection for smaller genomes during the invasion; the advantage of a small genome and associated traits could have been manifest without such selection processes because populations with small genomes coming from Europe directly competed with native North American populations with larger genomes. Nevertheless, our data allow for outlining the hypothetical sequence of changes in genome size, and associated plant traits, during the historically, biogeographically, and evolutionarily framed invasion process. The story unfolds with European populations having small genomes. The European populations were introduced to North America, but as the results of our study on genome size variation indicate, those populations that established and spread had, on average, slightly bigger genomes (0.506 pg) than those that were presumably filtered out following introduction from the native European range (0.490 pg). As we show, bigger genomes are associated with traits favoring spread, such as increased allocation to generative reproduction and production of runners (Fig. 3, Appendix S2: Fig. S6b), and might have thus been advantageous at the initial phase of invasion. Yet, relative to the native North American *P. australis* populations, the genome of the European populations that became invasive in North America was comparatively small enough (smaller on average by 6.9%) to generate trait differences that provided the invading populations with competitive superiority.

A question can be raised, with respect to the above-described biogeographical perspective, why the native European populations whose genome is the smallest of the three groups compared in our study, are not invasive. We suggest that this is paradoxical rather than contradictory because the populations of *P. australis* in Europe

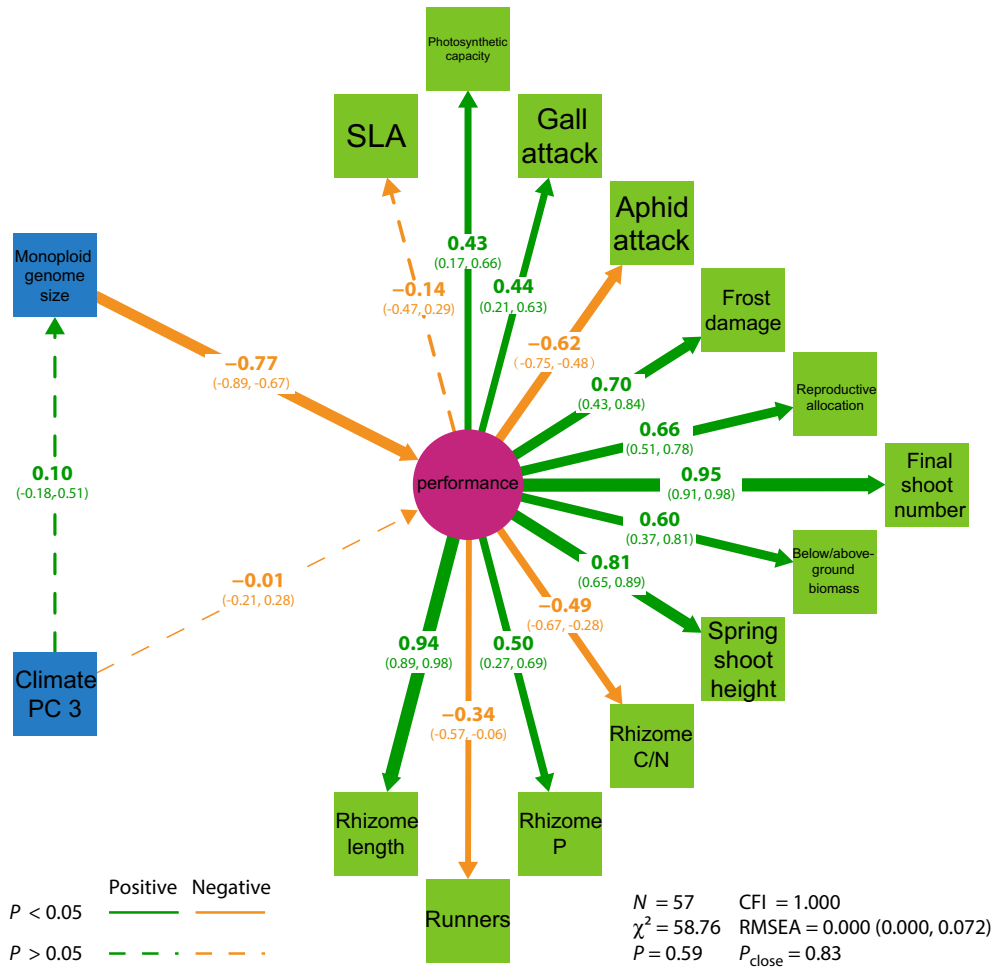


FIG. 3. Multiple Indicators Multiple Causes (MIMIC) models examining the impact of covariates on the latent variable, performance, for the three groups. The direction of the arrow in the structural model (from genome size and climate PC 3 to performance) indicates causal relationship; the direction of arrow in the measurement model (from performance to measured traits) indicates reflective relationship, i.e., measured traits are reflective effects of the latent variable, performance. Solid orange arrows represent negative paths ( $P < 0.05$ ), solid green arrows represent positive paths ( $P < 0.05$ ) and dashed arrows represent non-significant paths ( $P > 0.05$ ). The numbers on the arrows are the standardized path coefficients, and the numbers in brackets are the CI of the coefficients. The width of the arrow is proportional to the value of the path coefficient.  $N$ , number of samples; CFI, comparative fit index; RMSEA, root mean square error of approximation (90% CI);  $p_{close}$ ,  $P$  value of close fit. To simplify, the correlations between the indicators are not shown (Appendix S2: Table S3).

(and other parts of the species' native range) exhibited the same, or even more pronounced expansive behavior in the past as the invasive populations have more recently in North America. In Europe, *P. australis* is considered to be the most competitive of all wetland plants, with a great ecological amplitude that extends from lime- and acid-oligotrophic to eutrophic waters (Ellenberg 1988, Rodwell 1998, Packer et al. 2017). The species is monodominant in wetlands over extensive areas of land, commonly covering several square kilometers (Ellenberg 1988), indicating that its expansive behavior in the native range was, during post-glacial colonization (e.g., in UK; Pigott and Pigott 1963, Ingrouille 1995), even more pronounced than that of the invasive populations in North America (Packer et al. 2017).

Given this behavior, the fact that the native European populations have small genome does not contradict our overall finding of the role that genome size plays in *Phragmites* success. It should further be noted that the species is still a dominant of extensive wetland areas despite the decline of its populations that occurred in Europe and elsewhere several decades ago. The causes of this dieback were attributed to interactions of hydrology, herbivory by beetles, and eutrophication reducing the resilience of reeds to exposure from waves and erosion (Brix 1999, Roberts 2000, Ostendorp et al. 2003).

The competitive superiority of the European populations that invade North America results from the fact that relative to native North American populations, invaders invest in long rhizomes rather than aboveground



runners, emerge early in the spring (shoot height on first measurement date is a proxy for early emergence because the earlier the shoots emerged, the taller they were in the first weeks of growth), reach a higher shoot density at the peak of the growing season, do not suffer heavy aphid attack, and allocate more nitrogen and less carbon to their rhizome tissues. How genome size selects for these traits, however, requires further study. Between these two groups, comparing traits that obviously function to maximize growth in invasive populations is highly relevant because invasive and native populations often grow in mixed stands in North America where they directly compete (Cronin et al. 2015).

The biogeographic comparison of traits, sensu Colautti et al. (2014), yielded different results. After at least 150 years since its first introduction to North America, invasive *P. australis* exhibited increased C:N ratios in rhizome tissues and lower root/shoot ratios, greater allocation to generative reproduction, and produced more runners than their European ancestors. Collectively, these traits facilitate dispersal and invasive spread rather than just local growth and effective site preemption, supporting the notion that different traits confer advantages at different stages of the invasion process (Pyšek et al. 2009, 2015, Richardson and Pyšek 2012). A similar stage-dependent context can be inferred at the species level by comparing genome sizes of naturalized species in the Czech flora with their non-invasive congeners; small genomes were important for promoting successful establishment outside of the native range (i.e., naturalization), but less important during the transition from naturalized to invasive species, which is characterized by rapid and massive spread (Kubešová et al. 2010).

It is important to realize that genome size primarily affects plant traits (Suda et al. 2015, Meyerson et al. 2016) rather than invasiveness per se; even the variation not explained by our models could represent additional effects of monopleid genome size mediated through an unmeasured variable influencing the physiology of experimental plants, e.g., stomatal size and density and adjustment of the water economy (Clevering and Lissner 1999, Saltonstall et al. 2007, Mozdzer and Zieman 2010).

However, it should be noted that an alternative explanation for the patterns we observed cannot be excluded, i.e., that the European populations possessing the smallest genomes were not the sources of this invasive genotype introduction to North America. To test this hypothesis, genetic data is needed to relate the populations currently invading North America to their European ancestors thus allowing an inference about the post-introduction evolution of genome size. Therefore, employing phylogeographic data to explore whether or not the introduction history may confound the putative effect of genome size represents a challenge for future research. Nonetheless, our results strongly suggest that, at least for our study system, genome size is the most suitable proxy for synthesizing information on traits associated with invasiveness. This is particularly relevant from the perspective of the

global *P. australis* invasion; there are a number of phylogeographic groups delimited with an increasing knowledge of the worldwide distribution of haplotypes (Saltonstall 2002, Lambertini et al. 2006, Lambertini 2016) but as yet, the North American continent is the most thoroughly researched and has the best available information (Meyerson et al. 2009). Data are much more scarce for other regions, such as Australia or China, where the presence of introduced populations has been suggested but not substantiated. The distinct separation of invasive and non-invasive populations based on genome size alone would allow for fast and effective identification and screening of invasive genotypes and create an early warning system for areas where invasive populations threaten native wetlands (Pyšek et al. 2013, Suda et al. 2015).

Another key question resulting from the relationship of genome size to invasiveness is how population level intraspecific variation in genome size is likely to drive species responses to global change. Populations possessing small genomes may be better equipped to handle extreme environments (Knight et al. 2005, Suda et al. 2015, Meyerson et al. 2016). Our findings demonstrate this for *P. australis*: small genomes tend to occur in seasonally dry areas, a significant stressor for a plant with such a high water demand. But as we show, small-genome populations simultaneously manifest greater potential for invasion. This highlights an invasion paradox since invasions in extreme environments are less common than in productive, resource-rich environments (Chytrý et al. 2008). However, if those environmental constraints are overcome, the phenomenon suggested here may represent an emerging driver for future invasions that are not yet fully realized. We indicate this at the population level within one species, but suggest that this phenomenon may be generalizable and valid across more plant species.

#### ACKNOWLEDGMENTS

P. Pyšek and L. A. Meyerson conceived the idea and designed the experiment, L. A. Meyerson, H. Brix, and C. Lambertini collected the plant material, J. Čuda, J. Doležal, O. Kauzál, M. Lučanová, T. Mandáková, K. Pyšková, L. Moravcová, H. Skálová, J. Suda, L. A. Meyerson, and P. Pyšek collected the data, W. Y. Guo and J. Čuda analyzed the data, P. Pyšek, L. A. Meyerson, and J. Suda wrote the paper with contributions from W. Y. Guo, J. Čuda, and H. Skálová, and all authors commented on the manuscript. The study was supported by project no. 14-15414S (Czech Science Foundation, to P. Pyšek, H. Skálová, J. Čuda, J. Suda, M. Lučanová, L. A. Meyerson), and long-term research development project RVO 67985939 (The Czech Academy of Sciences). P. Pyšek acknowledges support by Praemium Academiae award from The Czech Academy of Sciences. L. A. Meyerson was also funded by US NSF DEB research grant 1049914 and the U.S. Fulbright Commission and the University of Rhode Island College of Environment and Life Sciences Agricultural Experiment Station Project R100H-332, 311000-6044. We thank Michal Pyšek for logistic support, Jan Pergl for computing climatic variables and consultation on preliminary statistics, Tomáš Hájek for photosynthesis measurement, and Marie Brůnová, David Cmíral, Dominika Prajzlerová, Ivana Rajznoverová, Zuzana Sixtová, Melissa Burger, Gillian Baird, and Xiao Guo for technical assistance.

## LITERATURE CITED

- Beaulieu, J. M., I. J. Leitch, and C. A. Knight. 2007a. Genome size evolution in relation to leaf strategy and metabolic rates revisited. *Annals of Botany* 99:495–505.
- Beaulieu, J. M., A. T. Moles, I. J. Leitch, M. D. Bennett, J. B. Dickie, and C. A. Knight. 2007b. Correlated evolution of genome size and seed mass. *New Phytologist* 173:422–437.
- Beaulieu, J. M., I. J. Leitch, S. Patel, A. Pendharkar, and C. A. Knight. 2008. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist* 179:975–986.
- Bennett, M. D. 1972. Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society B* 181:109–135.
- Breiman, L. 2001. Random forests. *Machine Learning* 45:5–32.
- Breiman, L., J. H. Friedman, R. A. Olshen, and C. G. Stone. 1984. *Classification and regression trees*. Chapman and Hall, Orange, California, USA.
- Brix, H. 1999. The European research project on reed die-back and progression (EUREED). *Limnologia* 29:5–10.
- Burk, I. 1880. List of plants recently collected on ship's ballast in the neighborhood of Philadelphia. *Proceedings of the National Academy of Sciences USA* 29:105–109.
- Chytrý, M., L. C. Maskell, P. Pyšek, M. Vilà, X. Font, and S. M. Smart. 2008. Habitat invasions by alien plants: a quantitative comparison among Mediterranean, subcontinental and oceanic regions of Europe. *Journal of Applied Ecology* 45:448–458.
- Clevering, O. A., and J. Lissner. 1999. Taxonomy, chromosome numbers, clonal diversity and population dynamics of *Phragmites australis*. *Aquatic Botany* 64:185–208.
- Colautti, R. I., J. D. Parker, M. W. Cadotte, P. Pyšek, C. S. Brown, D. Sax, and D. M. Richardson. 2014. Quantifying the invasiveness of species. *NeoBiota* 21:7–27.
- Cronin, J. T., G. P. Bhattarai, W. J. Allen, and L. A. Meyerson. 2015. Biogeography of a plant invasion: plant-herbivore interactions. *Ecology* 96:1115–1127.
- De'ath, G. 2007. Boosted trees for ecological modeling and prediction. *Ecology* 88:243–251.
- De'ath, G., and K. E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81:3178–3192.
- Doležel, J., J. Greilhuber, and J. Suda. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* 2:2233–2244.
- Dupin, M., P. Reynaud, V. Jarošík, R. Baker, S. Brunel, D. Eyre, J. Pergl, and D. Makowski. 2011. Effects of the training dataset characteristics on the performance of nine species distribution models: application to *Diabrotica virgifera virgifera*. *PLoS ONE* 6:e20957.
- Elith, J., J. R. Leathwic, and T. Hastie. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology* 77:802–813.
- Ellenberg, H. 1988. *Vegetation ecology of central Europe*. Cambridge University Press, Cambridge, UK.
- Gallagher, R. V., M. R. Leishman, J. T. Miller, C. Hui, D. M. Richardson, J. Suda, and B. Trávníček. 2011. Invasiveness in introduced Australian acacias: the role of species' traits and genome size. *Diversity and Distributions* 17:884–897.
- Gallo, J. J., P. V. Rabins, and J. C. Anthony. 1999. Sadness in older persons: 13-year follow-up of a community sample in Baltimore, Maryland. *Psychological Medicine* 29:341–350.
- Garcia, S., M. A. Canela, T. Garnatje, E. D. McArthur, J. Pellicer, S. C. Sanderson, and J. Valles. 2008. Evolutionary and ecological implications of genome size in the North American endemic sagebrushes and allies (*Artemisia*, Asteraceae). *Biological Journal of the Linnean Society* 94:631–649.
- Giles, D. 1999. Measuring the hidden economy: implications for econometric modelling. *Economic Journal* 109:370–380.
- Grace, J. B., T. M. Anderson, H. Olf, and S. M. Scheiner. 2010. On the specification of structural equation models for ecological systems. *Ecological Monographs* 80:67–87.
- Grace, J. B., D. R. Schoolmaster, G. R. Guntenspergen, A. M. Little, B. R. Mitchell, K. M. Miller, and E. W. Schweiger. 2012. Guidelines for a graph: theoretic implementation of structural equation modeling. *Ecosphere* 3:1–44.
- Greilhuber, J., J. Doležel, M. A. Lysak, and M. D. Bennett. 2005. The origin, evolution and proposed stabilization of the terms 'genome size' and 'C-value' to describe nuclear DNA contents. *Annals of Botany* 95:255–260.
- Grotkopp, E., M. Rejmánek, and T. L. Rost. 2002. Toward a causal explanation of plant invasiveness: seedling growth and life-history strategies of 29 pine (*Pinus*) species. *American Naturalist* 159:396–419.
- Grotkopp, E., M. Rejmánek, M. J. Sanderson, and T. L. Rost. 2004. Evolution of genome size in pines (*Pinus*) and its life-history correlates: supertree analyses. *Evolution* 58:1705–1729.
- Guo, W. Y., C. Lambertini, L. X. Nguyen, X. Z. Li, and H. Brix. 2014. Preadaptation and post-introduction evolution facilitate the invasion of *Phragmites australis* in North America. *Ecology and Evolution* 4:4567–4577.
- Guyon, I. 2003. An introduction to variable and feature selection. *Journal of Machine Learning Research* 3:1157–1182.
- Hancock, G. R. 2001. Effect size, power, and sample size determination for structured means modeling and MIMIC approaches to between-groups hypothesis testing of means on a single latent construct. *Psychometrika* 66:373–388.
- Haslam, S. M. 1972. *Phragmites communis* Trin. *Journal of Ecology* 60:585–610.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978.
- Hodgson, J. G., et al. 2010. Stomatal vs. genome size in angiosperms: the somatic tail wagging the genomic dog? *Annals of Botany* 105:573–584.
- Honaker, J., G. King, and M. Blackwell. 2011. Amelia II: a program for missing data. *Journal of Statistical Software* 45:1–47.
- Hothorn, T., K. Hornik, and A. Zeileis. 2006. Unbiased recursive partitioning: a conditional inference framework. *Journal of Computational and Graphical Statistics* 15:651–674.
- Ingrouille, N. 1995. *Historical ecology of British Flora*. Springer, London, UK.
- Jöreskog, K., and S. Goldberger. 1975. Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association* 70:631–639.
- Knight, C. A., N. A. Molinari, and D. A. Petrov. 2005. The large genome constraint hypothesis: evolution, ecology and phenotype. *Annals of Botany* 95:177–190.
- Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution* 16:199–204.
- Kubešová, M., L. Moravcová, J. Suda, V. Jarošík, and P. Pyšek. 2010. Naturalized plants have smaller genomes than their non-invading relatives: a flow cytometric analysis of the Czech alien flora. *Preslia* 82:81–96.
- Kuester, A., J. K. Conner, T. Culley, and R. S. Baucom. 2014. How weeds emerge: a taxonomic and trait-based examination using United States data. *New Phytologist* 202:1055–1068.
- Kursa, M. B., and W. R. Rudnicki. 2010. Feature selection with the Boruta package. *Journal of Statistical Software* 36:1–13.

- Küster, E. C., I. Kühn, H. Bruelheide, and S. Klotz. 2008. Trait interactions help explain plant invasion success in the German flora. *Journal of Ecology* 96:860–868.
- Lambertini, C. 2016. Heteroplasmy due to chloroplast paternal leakage: another insight into *Phragmites* haplotypic diversity in North America. *Biological Invasions* 18:2443–2455.
- Lambertini, C., M. H. G. Gustafsson, J. Frydenberg, J. Lissner, M. Speranza, and H. Brix. 2006. A phylogeographic study of the cosmopolitan genus *Phragmites* (Poaceae) based on AFLPs. *Plant Systematics and Evolution* 258:161–182.
- Lambertini, C., M. H. G. Gustafsson, J. Frydenberg, M. Speranza, and H. Brix. 2008. Genetic diversity patterns in *Phragmites australis* at the population, regional and continental scales. *Aquatic Botany* 88:160–170.
- Lambertini, C., B. K. Sorrell, T. Riis, B. Olesen, and H. Brix. 2012. Exploring the borders of European *Phragmites* within a cosmopolitan genus. *AoB Plants* 2012:pls020.
- Lavergne, S., N. J. Muenke, and J. Molofsky. 2010. Genome size reduction can trigger rapid phenotypic evolution in invasive plants. *Annals of Botany* 105:109–116.
- Leitch, I. J., and M. D. Bennett. 2007. Genome size and its uses: the impact of flow cytometry. Pages 153–176 in J. Doležal, J. Greilhuber, and J. Suda, editors. *Flow cytometry with plant cells*. Wiley-VCH Verlag, Weinheim, Germany.
- Leitch, I. J., M. W. Chase, and M. D. Bennett. 1998. Phylogenetic analysis of DNA C-values provides evidence for a small ancestral genome size in flowering plants. *Annals of Botany* 82(Suppl. A):85–94.
- McKee, J., and A. J. Richards. 1996. Variation in seed production and germinability in common reed (*Phragmites australis*) in Britain and France with respect to climate. *New Phytologist* 133:233–243.
- Meyerson, L. A., and J. T. Cronin. 2013. Evidence for multiple introductions of *Phragmites australis* to North America: detection of a new non-native haplotype. *Biological Invasions* 15:2605–2608.
- Meyerson, L. A., K. Saltonstall, L. Windham, E. Kiviat, and S. E. G. Findlay. 2000. A comparison of *Phragmites australis* in freshwater and brackish marsh environments in North America. *Wetland Ecology and Management* 8: 89–103.
- Meyerson, L. A., K. Saltonstall, and R. M. Chambers. 2009. *Phragmites australis* in eastern North America: a historical and ecological perspective. Pages 57–82 in B. R. Silliman, E. D. Grosholz, and M. D. Bertness, editors. *Human impacts on salt marshes: a global perspective*. University of California Press, Berkeley, California, USA.
- Meyerson, L. A., A. M. Lambert, and K. Saltonstall. 2010. A tale of three lineages: expansion of common reed (*Phragmites australis*) in the U.S. *Invasive Plant Science and Management* 3:515–520.
- Meyerson, L. A., J. T. Cronin, G. P. Bhattarai, H. Brix, C. Lambertini, M. Lučanová, S. Rinehart, J. Suda, and P. Pyšek. 2016. Do ploidy level and nuclear genome size and latitude of origin modify the expression of *Phragmites australis* traits and interactions with herbivores? *Biological Invasions* 18: 2531–2549.
- Morgan, H. D., and M. Westoby. 2005. The relationship between nuclear DNA content and leaf strategy in seed plants. *Annals of Botany* 96:1321–1330.
- Mozdzer, T. J., and J. C. Zieman. 2010. Ecophysiological differences between genetic lineages facilitate the invasion of non-native *Phragmites australis* in North American Atlantic coast wetlands. *Journal of Ecology* 98:451–458.
- Osnas, J. L., J. W. Lichstein, P. B. Reich, and S. W. Pacala. 2013. Global leaf trait relationships: mass, area, and the leaf economics spectrum. *Science* 340:741–744.
- Ostendorp, W., M. Dienst, and K. Schmieder. 2003. Disturbance and rehabilitation of lakeside *Phragmites* reeds following an extreme flood in Lake Constance (Germany). *Hydrobiologia* 506:687–695.
- Packer, J., L. A. Meyerson, H. Skálová, P. Pyšek, and C. Kueffer. 2017. Biological flora of British Isles: *Phragmites australis*. *Journal of Ecology* 105:1123–1162.
- Pandit, M. K., S. M. White, and M. J. O. Pocock. 2014. The contrasting effects of genome size, chromosome number and ploidy level on plant invasiveness: a global analysis. *New Phytologist* 203:697–703.
- Pigott, C. D., and M. E. Pigott. 1963. Late-glacial and post-glacial deposits at Malham, Yorkshire. *New Phytologist* 62:317–334.
- Poorter, H., H. Lambers, and J. R. Evans. 2014. Trait correlation networks: a whole-plant perspective on the recently criticized leaf economic spectrum. *New Phytologist* 201:378–382.
- Pyšek, P., and D. M. Richardson. 2007. Traits associated with invasiveness in alien plants: where do we stand? Pages 97–125 in W. Nentwig, editor. *Biological invasions*. Springer, Berlin, Germany.
- Pyšek, P., V. Jarošík, J. Pergl, R. Randall, M. Chytrý, I. Kühn, L. Tichý, J. Danihelka, J. Chrtek Jun, and J. Sádlo. 2009. The global invasion success of Central European plants is related to distribution characteristics in their native range and species traits. *Diversity and Distributions* 15:891–903.
- Pyšek, P., et al. 2013. Hitting the right target: taxonomic challenges of, and for, biological invasions. *AoB Plants* 5:plt042.
- Pyšek, P., et al. 2015. Naturalization of central European plants in North America: species traits, habitats, propagule pressure, residence time. *Ecology* 96:762–774.
- R Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [www.r-project.org](http://www.r-project.org)
- Rejmánek, M. 1996. A theory of seed plant invasiveness: the first sketch. *Biological Conservation* 78:171–181.
- Richardson, D. M., and P. Pyšek. 2006. Plant invasions: merging the concepts of species invasiveness and community invasibility. *Progress in Physical Geography* 30:409–431.
- Richardson, D. M., and P. Pyšek. 2012. Naturalization of introduced plants: ecological drivers of biogeographic patterns. *New Phytologist* 196:383–396.
- Ríos-Bedoya, C., C. Pomerleau, R. Neuman, and O. Pomerleau. 2009. Using MIMIC models to examine the relationship between current smoking and early smoking experiences. *Nicotine and Tobacco Research* 11:1035–1041.
- Roberts, J. 2000. Changes in *Phragmites australis* in southeastern Australia: a habitat assessment. *Folia Geobotanica* 35:353–362.
- Rodwell, J. S. (ed.) 1998. *British plant communities*. 4. Aquatic communities, swamps and tall-herb fens. Cambridge University Press, Cambridge, UK.
- Saltonstall, K. 2002. Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proceedings of the National Academy of Sciences USA* 99:2445–2449.
- Saltonstall, K. 2011. Remnant native *Phragmites australis* maintains genetic diversity despite multiple threats. *Conservation Genetics* 12:1027–1033.
- Saltonstall, K., K. Glennon, A. Burnett, R. B. Hunter, and K. L. Hunter. 2007. Comparison of morphological variation indicative of ploidy level in *Phragmites australis* (Poaceae) from eastern North America. *Rhodora* 109:415–429.
- Schönswetter, P., J. Suda, M. Popp, H. Weiss-Schneeweiss, and C. Brochmann. 2007. Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. *Molecular Phylogenetics and Evolution* 42:92–103.

- Strobl, C., A. L. Boulesteix, A. Zeileis, and T. Hothorn. 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 8:25.
- Strobl, C., A. L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9:307.
- Strobl, C., T. Hothorn, and A. Zeileis. 2009. Party on! A new, conditional variable-importance measure for random forests available in the party package. *R Journal* 1/2:14–17.
- Suda, J., L. A. Meyerson, I. Leitch, and P. Pyšek. 2015. The hidden side of plant invasions: the role of genome size. *New Phytologist* 205:994–1007.
- te Beest, M., J. J. Le Roux, D. M. Richardson, A. K. Brysting, J. Suda, M. Kubešová, and P. Pyšek. 2012. The more the better? The role of polyploidy in facilitating plant invasions. *Annals of Botany* 109:19–45.
- Tucker, G. C. 1990. The genera of Arundinoideae (Gramineae) in the southeastern United States. *Journal of Arnold Arboretum* 71:145–177.
- van Kleunen, M., E. Weber, and M. Fischer. 2010. A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters* 13:235–245.
- Wasserman, L., and K. Roeder. 2009. High-dimensional variable selection. *Annals of Statistics* 37:2178–2201.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1002/ecy.2068/supinfo>