Small genome size and variation in ploidy levels support the naturalization of vascular plants but constrain their invasive spread

Petr Pyšek[1,2] (iD), Magdalena Lučanová[3,4] (iD), Wayne Dawson[5] (iD), Franz Essl[6] (iD), Holger Kreft[7,8,9] (iD), Ilia J. Leitch[10] (iD), Bernd Lenzner[6] (iD), Laura A. Meyerson[11] (iD), Jan Pergl[1] (iD), Mark van Kleunen[12,13] (iD), Patrick Weigelt[7,8,9] (iD), Marten Winter[14] (iD) and Wen-Yong Guo[15,16] (iD)

[1]Department of Invasion Ecology, Institute of Botany, Czech Academy of Sciences, Průhonice, CZ-252 43, Czech Republic; [2]Department of Ecology, Faculty of Science, Charles University, Viničná 7, Prague, CZ-128 44, Czech Republic; [3]Department of Evolutionary Biology of Plants, Institute of Botany, Czech Academy of Sciences, Průhonice, CZ-252 43, Czech Republic; [4]Department of Botany, Faculty of Science, University of South Bohemia, Branišovská 1760, České Budějovice, CZ-370 05, Czech Republic; [5]Department of Biosciences, Durham University, South Road, Durham, DH1 3LE, UK; [6]Division of Bioinvasions, Global Change & Macroecology, Department of Botany and Biodiversity Research, University of Vienna, Wien, 1030, Austria; [7]Biodiversity, Macroecology & Biogeography, University of Göttingen, Büsgenweg 1, Göttingen, 37077, Germany; [8]Centre of Biodiversity and Sustainable Land Use (CBL), University of Goettingen, Büsgenweg 1, Göttingen, D-37077, Germany; [9]Campus-Institute Data Science, Goldschmidtstraße 1, Göttingen, 37077, Germany; [10]Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK; [11]University of Rhode Island, Natural Resources Science, 9 East Alumni Avenue, Kingston 02881 RI, USA; [12]Ecology, Department of Biology, University of Konstanz, Universitätsstrasse 10, Constance, D-78464, Germany; [13]Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, Taizhou University, Taizhou, 318000, China; [14]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstraße 4, Leipzig, 04103, Germany; [15]Research Centre for Global Change and Complex Ecosystems, School of Ecological and Environmental Sciences, East China Normal University, Shanghai, 200241, China; [16]Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, School of Ecological and Environmental Sciences, East China Normal University, Shanghai, 200241, China

Author for correspondence:
Petr Pyšek
Email: *pysek@ibot.cas.cz*

Summary

- Karyological characteristics are among the traits underpinning the invasion success of vascular plants.
- Using 11 049 species, we tested the effects of genome size and ploidy levels on plant naturalization (species forming self-sustaining populations where they are not native) and invasion (naturalized species spreading rapidly and having environmental impact).
- The probability that a species naturalized anywhere in the world decreased with increasing monoploid genome size (DNA content of a single chromosome set). Naturalized or invasive species with intermediate monoploid genomes were reported from many regions, but those with either small or large genomes occurred in fewer regions. By contrast, large holoploid genome sizes (DNA content of the unreplicated gametic nucleus) constrained naturalization but favoured invasion.
- We suggest that a small genome is an advantage during naturalization, being linked to traits favouring adaptation to local conditions, but for invasive spread, traits associated with a large holoploid genome, where the impact of polyploidy may act, facilitate long-distance dispersal and competition with other species.

Introduction

The rapidly increasing availability of data on the global distributions of alien plant species (van Kleunen *et al.*, 2015, 2019; Dawson *et al.*, 2017; Pyšek *et al.*, 2017; Pagad *et al.*, 2018; Essl *et al.*, 2019) and plant traits (Kattge *et al.*, 2020; Weigelt *et al.*, 2020) has created an opportunity to explore the mechanisms underpinning biological invasions across large geographical scales, and more robustly than ever before (Pyšek *et al.*, 2020b). Previous research has assessed how plant invasions are affected by

species traits such as reproductive mode (Dellinger *et al.*, 2016; Razanajatovo *et al.*, 2016, 2019), life strategy (Guo *et al.*, 2018) and mutualistic interactions (Delavaux *et al.*, 2019; Pyšek *et al.*, 2019) (see Pyšek & Richardson, 2007; van Kleunen *et al.*, 2010 for overviews). However, the quality of data available for particular traits varies among taxa, regions and traits. Although some traits are thought to determine a plant's invasion success, these suggestions were often based on anecdotal evidence without proper statistical testing due to a lack of data. Plant genome size, that is the amount of nuclear DNA (Bennett & Leitch, 2005a;

Greilhuber *et al.*, 2005), is one such trait. The role of genome size in plant invasions was first suggested > 20 yr ago based on studies of individual genera (Rejmánek, 1996; Grotkopp *et al.*, 2002; Garcia *et al.*, 2008) and has only recently been addressed by comparative analyses involving large multispecies data sets. Indeed, there are now several papers demonstrating that naturalized or invasive species tend to have smaller genomes than those that have not successfully naturalized or invaded (Kubešová *et al.*, 2010; Pandit *et al.*, 2014; Pyšek *et al.*, 2015), which may explain why species with small genomes are significantly over-represented among invasive taxa (Suda *et al.*, 2015).

From the 'large genome constraint' hypothesis (Knight *et al.*, 2005), it follows that species with large genomes are less likely to be invasive than species with small genomes. In part, this is because species with large genomes have a narrower range of other trait states since trait options become more limited as genome size increases. For example, while plant species with large genomes are restricted to being obligate perennials, those with small genomes have a wider diversity of life-cycle strategies they may adopt (e.g. they may be ephemeral, annual or perennial; Bennett, 1972). Furthermore, species with small genomes have a wide range of seed sizes, whereas very small seeds that are easily dispersed over long distances are rarely associated with large-genome species, an observation attributed to developmental constraints associated with big genomes (Beaulieu *et al.*, 2007). Stomata size (which increases with increasing genome size), stomatal density (which decreases with increasing genome size), and the rate of stomatal opening and closing can also play a role as these traits influence carbon fixation and water-use efficiency (Beaulieu *et al.*, 2008). Thus, species with larger genomes are more restricted in the range of suitable environmental conditions they can grow in compared to species with smaller genomes. In addition, nutrient-poor soils can select against species with large genomes because these species have higher demands for phosphorus and nitrogen to build DNA and RNA and maintain their cells (Šmarda *et al.*, 2013; Guignard *et al.*, 2016). Therefore, many traits associated with large genomes do not overlap with the characteristics found in species that are successful invaders (Suda *et al.*, 2015).

Recently, the emerging picture has been extended by considering intraspecific variation in genome size, demonstrating its effect at the level of individual genotypes within a species. For example, in *Phragmites australis* (Cav.) Trin. ex Steud., the invasive populations in North America differ from native European populations in several growth, physiological and reproductive traits (Meyerson *et al.*, 2016, 2020; Pyšek *et al.*, 2019), which were related to their smaller genome sizes compared with the native North American populations (Pyšek *et al.*, 2018).

Genome size may also interact with ploidy level to influence the invasion success of alien species (te Beest *et al.*, 2012; Pandit *et al.*, 2014; Meyerson *et al.*, 2016). For example, ploidy level has been shown to affect various attributes such as breeding system, plant–herbivore interactions or photosynthetic rate (Naiki & Nagamasu, 2004; Guggisberg *et al.*, 2006; Meyerson *et al.*, 2016). As well as the nucleotypic effect of doubling the number of chromosomes, in polyploids the impact of intra- or interspecific hybridization has been shown to play a role in influencing the success of

polyploids together with genotype and the level of heterozygosity (Levin, 1983). In general, polyploidy (or whole genome duplication) is common among invasive plants (Amsellem *et al.*, 2001; te Beest *et al.*, 2012). Polyploidization does not only increase the DNA content (Soltis & Soltis, 1999) but also often genetic variability, such that some polyploid species have been shown to possess a better ability to colonize new habitats than their diploid relatives (Ehrendorfer, 1980; Stebbins, 1985; Thompson & Lumaret, 1992; DeWalt & Hamrick, 2004; Schierenbeck & Ainouche, 2006). On the contrary, there is often a strong drive for genome size increases to be followed by genome downsizing, especially in angiosperms (Wang *et al.*, 2021), because the accumulation of extra DNA (with or without genome duplication) means higher nutritional costs for maintaining a larger genome. For example, studies have shown positive correlations between phosphorus and/or nitrogen availability in the soil and the genome sizes of the plants growing in the community. From long-term fertilizer-addition experiments, it is obvious that plants with large genomes are not able to grow and compete successfully in nutrient-poor conditions (Šmarda *et al.*, 2013; Guignard *et al.*, 2016). This cost is considered to be one of the major evolutionary forces driving genome downsizing and the reason why most plants have small genome sizes, although all flowering plants have undergone repeated polyploidization events throughout their evolutionary history (Wendel, 2015; Faizullah *et al.*, 2021; Wang *et al.*, 2021). However, the mechanisms of the combined effects of these two karyological characteristics, genome size and ploidy level, on plant invasion success are not well-understood.

Despite these observations and studies, no comprehensive global-level synthesis exists that sheds light on possible interactions between genome size and other karyological traits across vascular plants and tests the effects of genome size in a context-dependent framework. It is well-established that for biological invasions, the effect of biological traits is context-dependent (Thuiller *et al.*, 2006; Blanchet *et al.*, 2009; Moodley *et al.*, 2013; Pyšek *et al.*, 2015, 2020a) with the stage of the invasion process being particularly important (Williamson, 2006; van Kleunen *et al.*, 2007; Gravuer *et al.*, 2008; Pyšek *et al.*, 2009). The invasion process starts with a human-mediated introduction, followed by a casual stage (when a species' persistence in the wild depends on repeated input of propagules), then naturalization (forming self-reproducing populations beyond cultivation), and finally invasion (naturalized species spreading rapidly into new environments; Richardson *et al.*, 2000; Blackburn *et al.*, 2011).

In this study, we address this stage dependence by employing several measures that quantify the global success of plant species which allowed us to disentangle the effects of genome size and ploidy on species invasions. Specifically, we explore the role that genome size plays in plant naturalization (i.e. the establishment of a species in its new range) and invasion (i.e. its further spread often associated with impact on invaded ecosystems), how it interacts with ploidy level, and whether the effects of these two karyological characteristics change from naturalization to invasion. We hypothesize that small genomes provide an advantage that successfully distinguishes species that are naturalized or invasive anywhere in the world from those that are not. However, we

also hypothesize that ploidy level has an opposing effect – and hence that the naturalization and invasion success of species increases with increasing ploidy levels, that is with a higher amount of DNA in the nucleus and greater genetic diversity.

We test these related hypotheses by combining the most comprehensive databases on the distribution of naturalized alien plant species worldwide (Global Naturalized Alien Flora database – GloNAF; van Kleunen *et al.*, 2015, 2019), and plant nuclear DNA content (Plant DNA *C*-values database; https://cvalues.science.kew.org). From the latter, we extracted genome size data (*C*-values) for 11 049 species (Fig. 1; Supporting Information Table S1). We considered both holoploid (*C*-value, i.e. the total amount of DNA in an unreplicated gametic nucleus of a cell) and monoploid genome sizes (*C*$_x$-value, i.e. the genome size of one unreplicated chromosome set; *sensu* Greilhuber *et al.*, 2005) in our analyses. While analyses of the holoploid genome size provide us with insights into the overall effect of nuclear DNA content, analyses of monoploid genome size allow comparison of genome size across different ploidy levels. We also compiled information on all ploidy levels reported for each species, and determined whether the species was diploid or polyploid. We used GloNAF, which includes data on the regional distribution of 13 939 naturalized plant taxa worldwide and was used to characterize the naturalization success of those species with genome size and ploidy data. For a subset of regions with available data, we also collected information on invasive species following the definition of the Convention on Biological Diversity (CBD, 2000). We used the GloNAF database and invasion data set to characterize the success of each species as an alien by expressing its naturalization incidence (whether it is reported as naturalized from anywhere in the world), naturalization extent (from how many regions it is reported as naturalized), and invasion extent, using the number of regions in GloNAF where the species is reported as invasive.

Materials and Methods

Data on karyological characteristics

We used the Plant DNA *C*-values Database (release 7.1, April 2019, Leitch IJ, Johnston E, Pellicer J, Hidalgo O, Bennett MD; https://cvalues.science.kew.org), which contains genome size data (i.e. *C*-values) for 12 273 species, comprising data for land plants and several algal lineages; for our study, we focussed on data for vascular plants. To have a taxonomically unified data set, we ran the species through the R package TAXONSTAND (Cayuela & Oksanen, 2014) and worked only with species with an accepted name. In the case of multiple genome size values for one species, we considered if these represented individuals with more than one ploidy level (i.e. multiples of genome size and/or
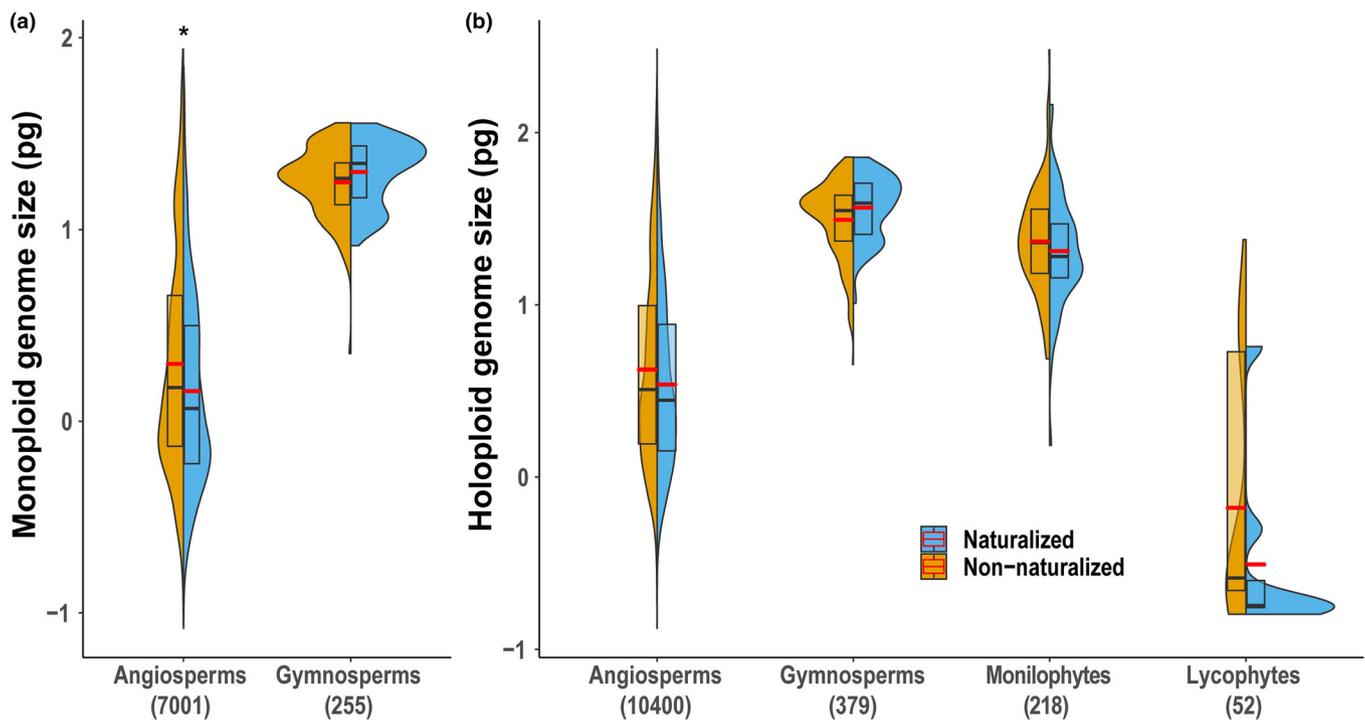
Fig. 1 Split-violin plots and boxplots showing the distribution of (a) the monoploid genome size (*C*$_x$-value) for vascular plants: angiosperms and gymnosperms. Missing information on ploidy levels of monilophytes and lycophytes meant it was not possible to calculate their monoploid genome sizes. (b) Holoploid genome size (*C*-value) for vascular plants: angiosperms, gymnosperms, monilophytes, and lycophytes. The different colours correspond to naturalized and non-naturalized species, with the red and black lines in the boxplot indicating the mean and median values for each group. The number of species in each taxonomic group is indicated below its name. After accounting for species relatedness, the only significant difference between naturalized and non-naturalized species was found in the monoploid genome size for angiosperms (*, $P < 0.05$), all other groups had relatively similar genome size values between naturalized and non-naturalized species ($P > 0.05$). The data were $\log_{10}$-transformed to improve visualization. See Supporting Information Table S1 for a detailed summary.

chromosome numbers). If there were multiple values for the same ploidy level, we applied the following criteria to choose the most reliable value. We preferred: values made using flow cytometry with the intercalating dye propidium iodide, a combination of Feulgen densitometry and flow cytometry, or computerized image analysis (Greilhuber *et al.*, 2005), as the only methods of genome size estimation following best practice recommendations (Sliwinska *et al.*, 2022); values reported by authors from well-established laboratories following best practice (see Galbraith *et al.*, 2021); the newer and repetitive values estimated using best practice methods; prime estimates – that is, the values marked by the authors of the Plant DNA *C*-values database, which are considered to represent the most reliable value obtained under best-practice methods when multiple estimates have been reported (as originally defined by Bennett & Smith, 1976). This data set was complemented by genome size values obtained from a further 40 studies from the literature that have not yet been incorporated into the database. They were identified by a dedicated search in Web of Science using the search terms 'genome size' and '*C*-value' (Methods S1). In total, our genome size data set comprised *C*-values for 11 049 species (10 400 angiosperms, 379 gymnosperms, 218 monilophytes and 52 lycophytes; Table S1). While the Plant DNA *C*-values database reports genome sizes as $1C$, $2C$ or $4C$-values (Greilhuber *et al.*, 2005), $1C$-values (i.e. the amount of DNA in the unreplicated gametic nucleus) given in picograms (pg) were extracted for all the analyses presented here.

We compiled information on all ploidy levels reported for each species, based also on chromosome numbers, reported in the Plant DNA *C*-values database (if available). If only chromosome numbers were present, we calculated the ploidy level from the basic chromosome number reported for a genus and the multiples of the genome size estimate. If no chromosome number was given for a species, we did not estimate the ploidy level (to avoid assigning the genome size to the wrong ploidy level). In such cases, we only analysed the holoploid genome size (*C*-value). For monilophytes and lycophytes, where data for chromosome numbers are mostly missing, we did not work with the ploidy level at all. Based on the estimated ploidy levels, we determined whether the species was diploid or polyploid. In cases where multiple ploidy levels were reported for a species (only 4.8%, i.e. 534 out of the 11 049 species, are ploidy-variable in our data set), the mean *C*-value was calculated. While we recognize that this removed some of the variation in our dataset, since our full data set comprises such a large range of genome sizes, genome size variation associated with intraspecific variation in ploidy levels was considered to be insignificant. Altogether we obtained holoploid genome sizes for 11 049 species, of which 5089 were diploids, 2196 polyploids and 3764 species lacked ploidy level information. Monoploid genome size (C_x-value, that is the $2C$ holoploid genome size divided by ploidy level, corresponding to the genome size of one unreplicated chromosome set; *sensu* Greilhuber *et al.*, 2005) was calculated for 7256 species (see Table S1 for sample sizes of each vascular plant lineage). For statistical analysis, we used both holoploid and monoploid genome sizes.

Holoploid genome size and monoploid genome size are closely related (Fig. S1; $R^2 = 0.90$), yet we report both analyses in the

paper. While the monoploid values inform about the effect of the genome size alone, the results obtained for holoploid values reflect the joint effects of both genome size and ploidy, that is the coarse overall pattern that will then be disentangled by other analyses.

Measures of naturalization and invasion success

We used the Global Naturalized Alien Flora (GloNAF) database (van Kleunen *et al.*, 2019; accessed on 7 April 2021) to characterize naturalization success. This database includes data on the distribution of 13 939 naturalized plant taxa in 1029 regions (i.e. countries, states, provinces and districts, including 381 islands) world-wide (van Kleunen *et al.*, 2019). For a subset of regions with available data ($n = 349$), we collected information on which species are classified as invasive. Here, we followed the widely accepted approach in environmental policy, based on the standard definition of the Convention on Biological Diversity (2000), in which the term 'invasive' refers to the subset of naturalized alien species, which exert negative impacts on the environment. To avoid the influence of different interpretations of the term invasive by individual researchers and to ensure a geographically balanced sampling (see Pyšek *et al.*, 2017; Essl *et al.*, 2019 for details), we based our list of invasive alien plant species on three global data sources which contain standardized information on invasiveness: (1) the CABI Invasive Species Compendium (http://www.cabi.org/isc); (2) the ISSG Global Invasive Species Database (http://www.iucngisd.org/gisd); and (3) the invasive plant species database (Weber, 2017).

The GloNAF database and the invasion data set were used to characterize the success of each species as an alien by using the following measures: (1) naturalization incidence – whether the species is reported as naturalized from anywhere in the world, that is included in GloNAF based on its occurrence as naturalized in at least one region; (2) naturalization extent (for the subset of those species that are naturalized and included in GloNAF), defined as how many regions it is reported from. Analogously, we defined (3) invasion extent, expressed as the number of regions in GloNAF from which the species is reported as invasive. While the fact that information on invasiveness is only available for a subset of regions does not prevent comparisons between invasive species in terms of how widespread they are (it can be assumed that widely distributed invasive species will be distinguished from rare ones even in the subset of data), it might provide biased data on whether or not a species is invasive somewhere. Therefore, we did not consider the invasion incidence analogously to naturalization incidence.

Phylogeny

To consider the relatedness of species in further analysis, we constructed phylogenetic trees for the species analysed here using the largest dated vascular plant phylogeny presently available, the extended GBOTB mega-tree of Smith & Brown (2018). To build this dated tree, Smith & Brown (2018) combined molecular data from GenBank with data from the Open Tree of Life

project, and Jin & Qian (2019) further checked and extended this mega-tree, including 74 533 taxa and all families of extant vascular plants. We pruned the phylogeny to the set of species used in this study based on the supertree using the default setting of the V. PHYLOMAKER package (Jin & Qian, 2019) in R v.4.1.0 (R Core Team, 2021). The final time-calibrated phylogenetic tree used here is shown in Fig. S2, plotted using iTOL (Letunic & Bork, 2019).

Statistical analyses

Given the uneven distribution of genome sizes across taxonomic groups (Leitch & Leitch, 2012; Suda et al., 2015; Pellicer et al., 2018), below we give the descriptive statistics on naturalization incidence and extent for data pooled across all taxa, as well as for angiosperms, gymnosperms, monilophytes and lycophytes separately. We ran a phyloANOVA in the PHYTOOLS package (Revell, 2012) for naturalized and non-naturalized species in each group to assess the mean difference in genome size. Because the taxonomic group was not a significant factor in almost all of the analyses (Fig. 1), all statistical analyses presented below are for pooled data.

Even though most previous studies found a negative relationship between genome size and species naturalization and invasiveness (Kubešová et al., 2010; Pyšek et al., 2018), and although a small genome size is linked with many traits associated with successful invasion (Suda et al., 2015), a minimum genome size is necessary to perform basic biological activities such as cell division (Fridley & Craddock, 2015). In addition, the relationship between genome size and species success is not uniform across the huge range of genome sizes encountered in the species analysed. The monoploid genome size of naturalized species ranged from 0.08 to 87.16 pg (i.e. 1089-fold), with similarly large ranges of values for holoploid genome size and among invasive species (Table S1). Therefore, we used both linear and quadratic phylogenetic logistic regressions (Ives & Garland, 2010) to analyse the relationships between naturalization incidence and genome size. Similarly, we tested for associations between naturalization and invasion extent with genome size using phylogenetic generalized linear regression. To check whether there was any variation between major angiosperm groups, we also repeated the global naturalization incidence analyses for monocots and eudicots separately. For both groups, we found similar patterns as in the combined data set (Table S2; Fig. S3); thus, we did not separate them in further analyses. The Akaike information criterion (AIC) was used to select the most parsimonious model between linear and quadratic models for each analysis. As holoploid genome size is affected by ploidy level, we further checked whether the global relationships between holoploid genome size and each of the three response variables (i.e. naturalization incidence and extent, invasion extent) were consistent for the most common ploidy levels in our data set, specifically diploids, tetraploids and hexaploids. We then used both phylogenetic logistic regression and phylogenetic linear regression to analyse the effects of the number of ploidy levels of a species, and its interactions with genome size, on

naturalization incidence and naturalization and invasion extent. The monoploid and holoploid genome sizes and naturalization and invasion extents were log_{10}-transformed to improve model normality of the residuals and then further standardized to zero mean and standard deviations of 1 to facilitate comparisons and visualization (Schielzeth, 2010).

To test for differences between diploids and polyploids in their naturalization success, we ran a phylogenetic logistic regression and phylogenetic linear regression for naturalization incidence and extent, respectively, using the following groups: diploids, polyploids and diploids + polyploids. We further summarized the number of ploidy levels reported for each species (Table S3). To test for the effect of the number of ploidy levels reported for a species on naturalization incidence, we ran generalized linear mixed effects models (GLMMs) with a binomial error distribution with genus nested within family as a random factor to account for species relatedness. We ran similar GLMMs with a negative binomial error distribution to test the effect on naturalization extent and invasion extent. We then ran Bonferroni's multiple post hoc tests to test the significance of differences between the GLMM model estimates of ploidy levels using the EMMEANS package (Lenth, 2020). We further tested the effects of monoploid genome size and several ploidy levels with sufficient species numbers on all three dependent variables.

The phylogenetic logistic and phylogenetic generalized linear regressions were implemented in the PHYLOLM package (Ho & Ané, 2014). Generalized linear mixed effects models were applied in the LME4 package (Bates et al., 2015). We used R^2 to quantify the variance explained by the variables in each model. For phylogenetic logistic models, the RR2 package (Ives & Li, 2018) was applied to obtain the pseudo-R^2 (Ives, 2019).

Results

Overview of the global data on genome size for taxonomic groups

Among angiosperms and gymnosperms with available monoploid genome size data (i.e. genome size of one unreplicated chromosome set), the percentages of invasive species were similar, but gymnosperms had a higher percentage of naturalized species than angiosperms (Table S1; Fig. 1). Pooled across all the assessed plant taxonomic groups and out of the 11 049 plant species analysed, 2548 (35.1%) and 680 (9.4%) were naturalized and invasive, respectively. Even though naturalized and invasive angiosperms had a wider range of monoploid genome sizes (Table S1), gymnosperms had larger genomes than angiosperms, regardless of species naturalization or invasion success (Fig. 1).

For angiosperms and gymnosperms, the distributions of the holoploid genome sizes (i.e. the total amount of DNA in an unreplicated gametic nucleus of a cell) were generally similar to those shown for monoploid genome sizes (Table S1). Across all vascular plants, the holoploid genome sizes of gymnosperms and monilophytes had larger genomes than angiosperms and lycophytes (Fig. 1).

Effect of genome size on naturalization and invasion

The probability of a species being naturalized somewhere in the world (i.e. naturalization incidence) decreased with increasing monoploid genome size, and the linear and quadratic regressions showed similar model fits in terms of AIC (Fig. 2a; Table 1); the pattern also held for eudicots and monocots when analysed separately (Fig. S3; Table S2). Although monoploid genome size was also negatively related to naturalization and invasion extent, quadratic regressions showed better model fits, as indicated by smaller AICs (Table 1), and hence revealed non-linear relationships between monoploid genome size and these two extent measures (Fig. 2b,c). Naturalized or invasive species with intermediate monoploid genome sizes tended to occur in a large number of regions (i.e. their naturalization extent and invasion extent are greater) compared to species with either small or large genomes. It was notable that the constraint imposed by

possessing a small monoploid genome size was much higher for invasion extent than for naturalization extent as can be seen from the shapes of the quadratic fits (Fig. 2b vs 2c).

As for monoploid genome size, the naturalization incidence was negatively associated with holoploid genome size. However, the quadratic regression, which showed similar model performance in terms of AIC (Table 1), revealed that species with intermediate holoploid genome size have a greater probability of becoming naturalized than species with smaller or bigger genomes (Fig. 3a). An analysis based on separated ploidy levels revealed (Table S4; Fig. S4) that this overall pattern was shaped by diploid and tetraploid species, whereas hexaploid species showed the opposite relationship, that is, hexaploid species with large holoploid genome sizes were more likely to be naturalized than hexaploids with small genome sizes. Nevertheless, the quadratic regression which had a smaller AIC value than the linear one revealed that hexaploids generated a pattern similar to the
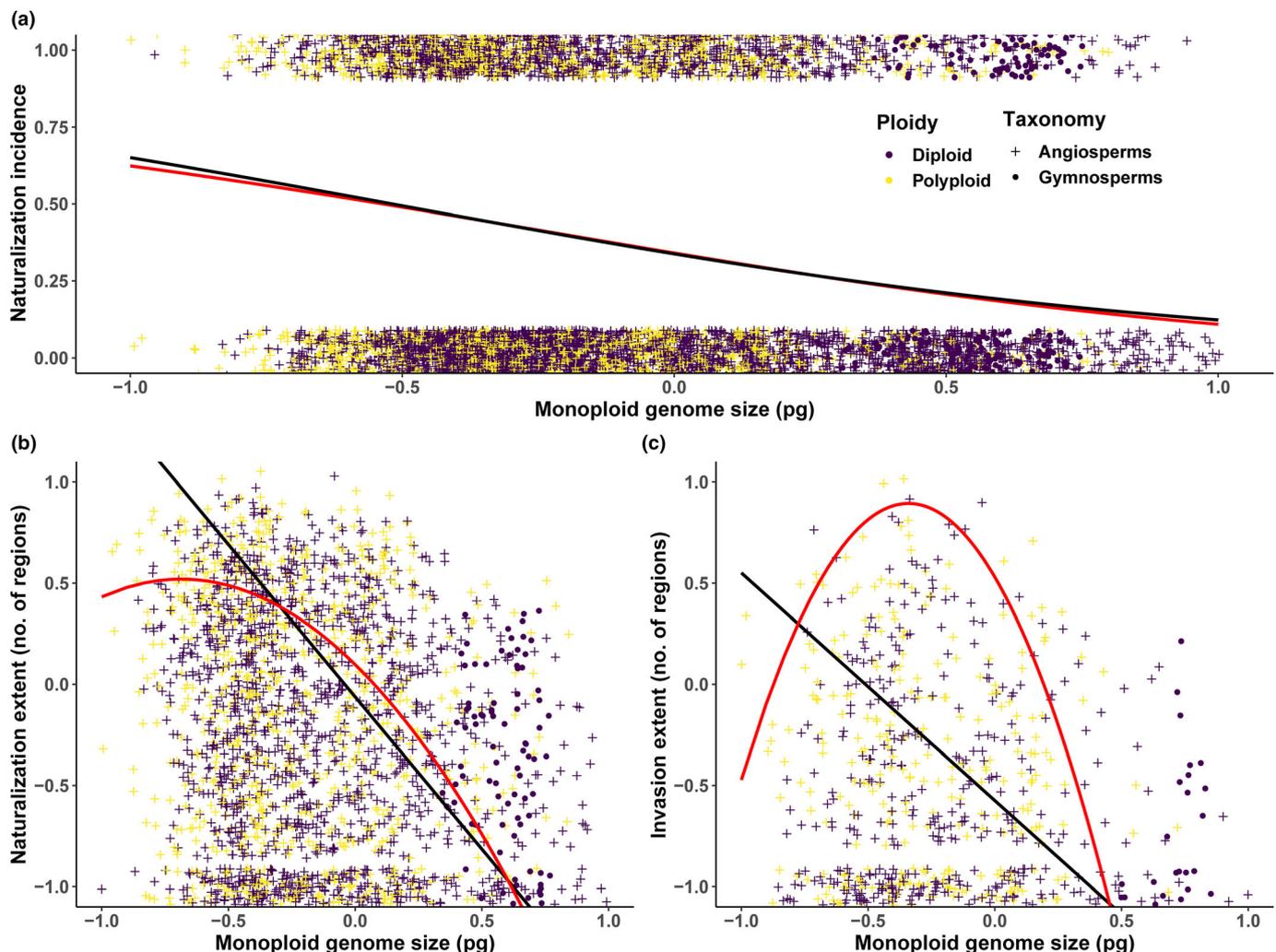
Fig. 2 Effects of the monoploid genome size on (a) naturalization incidence (i.e. whether or not the species is reported as naturalized from at least one region in the GloNAF database), (b) naturalization extent (i.e. the number of GloNAF regions) and (c) invasion extent. Black and red solid lines correspond to linear and quadratic fits, respectively. Colours and shapes represent the ploidy status and taxonomic groups, respectively. The detailed statistical summary is shown in Table 1. The naturalization extent and invasion extent were log$_{10}$-transformed and then, as was also the monoploid genome size, were standardized to zero mean and standard deviations of one. To improve the visibility, points were jittered.

Table 1 Summary of the linear and quadratic phylogenetic regressions of the effect of monoploid/holoploid genome size on naturalization incidence (whether or not the species is reported as naturalized from at least one region in the GloNAF database), and naturalization and invasion extent (the number of the GloNAF regions where the species is naturalized or invasive).

Models		Estimate	SE	Z-value	P-value	R^2	Phylogenetic signal	AIC
Naturalization incidence								
Linear	Monoploid genome size	−1.295	0.103	−12.539	< 0.001	0.089	0.097	8927
Quadratic	Monoploid genome size	−1.305	0.058	−11.290	< 0.001	0.081	0.100	8927
	Monoploid genome size^2	−0.139	0.239	−0.583	0.56			
Linear	Holoploid genome size	−1.008	0.083	−12.178	< 0.001	0.067	0.117	13 723
Quadratic	Holoploid genome size	−0.973	0.088	−11.097	< 0.001	0.068	0.124	13 722
	Holoploid genome size^2	−0.988	0.205	−4.822	< 0.001			
Naturalization extent								
Linear	Monoploid genome size	−1.506	0.121	−12.409	< 0.001	0.061	0.272	9444
Quadratic	Monoploid genome size	−1.230	0.128	−9.627	< 0.001	0.077	0.267	9405
	Monoploid genome size^2	−0.895	0.139	−6.422	< 0.001			
Linear	Holoploid genome size	−0.632	0.072	−8.770	< 0.001	0.021	0.220	13 535
Quadratic	Holoploid genome size	−0.368	0.073	−5.005	< 0.001	0.064	0.210	13 377
	Holoploid genome size^2	1.427	0.111	12.808	< 0.001			
Invasion extent								
Linear	Monoploid genome size	−1.125	0.179	−6.290	< 0.001	0.055	0.110	2406
Quadratic	Monoploid genome size	−2.140	0.211	−10.127	< 0.001	0.139	0.100	2344
	Monoploid genome size^2	−3.142	0.385	−8.158	< 0.001			
Linear	Holoploid genome size	0.430	0.102	4.222	< 0.001	0.018	0.094	3174
Quadratic	Holoploid genome size	2.003	0.114	17.622	< 0.001	0.325	0.065	2822
	Holoploid genome size^2	3.005	0.145	29.683	< 0.001			

quadratic result based on analysing all species (Fig. S4A), even though the intermediate genome size value of hexaploids with the highest probability of naturalization was larger than the value obtained by analysing across all species (Fig. S4A; Table S4).

Interestingly, the linear and quadratic regressions yielded different results for naturalization and invasion extent: the linear regressions revealed negative or positive associations, respectively, but for both naturalization and invasion, the quadratic relationships showed that species with an intermediate holoploid genome size were less widespread (i.e. having the smallest extent) compared with species with either small or large holoploid genomes (Fig. 3b,c). This intermediate genome size pattern for naturalization extent was generally consistent for diploids and tetraploids, except for differences between the amplitudes of the relationships across the range of genome sizes – diploids had lower and tetraploids had higher amplitudes than the model based on all the data (Fig. S4B). This indicates that for species with intermediate holoploid genome sizes, tetraploids are predicted to have a much smaller, and diploids slightly greater naturalization extents compared to the average across all species. However, compared with the model based on all the data, diploids showed the opposite trend for invasion extent, that is, diploid species with intermediate holoploid genome sizes invaded more regions than those with either small or large genomes (Fig. S4C; Table S4).

Effect of ploidy level and ploidy diversity on naturalization and invasion

We found a positive relationship between the number of different ploidy levels reported for a species (Table S3) and its naturalization

incidence (Fig. 4a; Table 2; $P < 0.001$). Furthermore, while polyploids were more likely to become naturalized than diploids (Fig. 4b; $P < 0.001$), the highest naturalization incidence was recorded for species that are reported to include both diploid and polyploid individuals (Fig. 4b; Table S5; $P < 0.05$). By contrast, the number of ploidy levels of a species was negatively related to its naturalization extent (Fig. 4c; Table 2; $P < 0.001$).

The impact of polyploidy on naturalization success (incidence and extent) and invasion extent was explored in more detail for specific ploidy levels (Table S6). The specific ploidy level reported for a species had a significant effect on naturalization incidence ($P < 0.001$) and naturalization extent ($P < 0.01$) but not on invasion extent (Table S7). In all models that yielded significant results, diploids were less likely to naturalize than species with higher ploidy levels, and the likelihood of naturalization success (i.e. naturalization incidence and extent) was generally higher for species with intermediate ploidy levels from triploids to hexaploids, compared with octoploids and higher ploidy levels (i.e. ≥ 10×; Fig. 4d).

Interactions between genome size and ploidy level affecting species naturalization and invasion success

While we found significant main effects of genome size and the number of ploidy levels on naturalization incidence only, the interactions between these two karyological variables were highly significant in their effect on naturalization and invasion extent ($P < 0.001$). Naturalization extent decreased with increasing monoploid genome size, but the decrease was faster for species reported to have just one ploidy level (Table S8). In addition, the
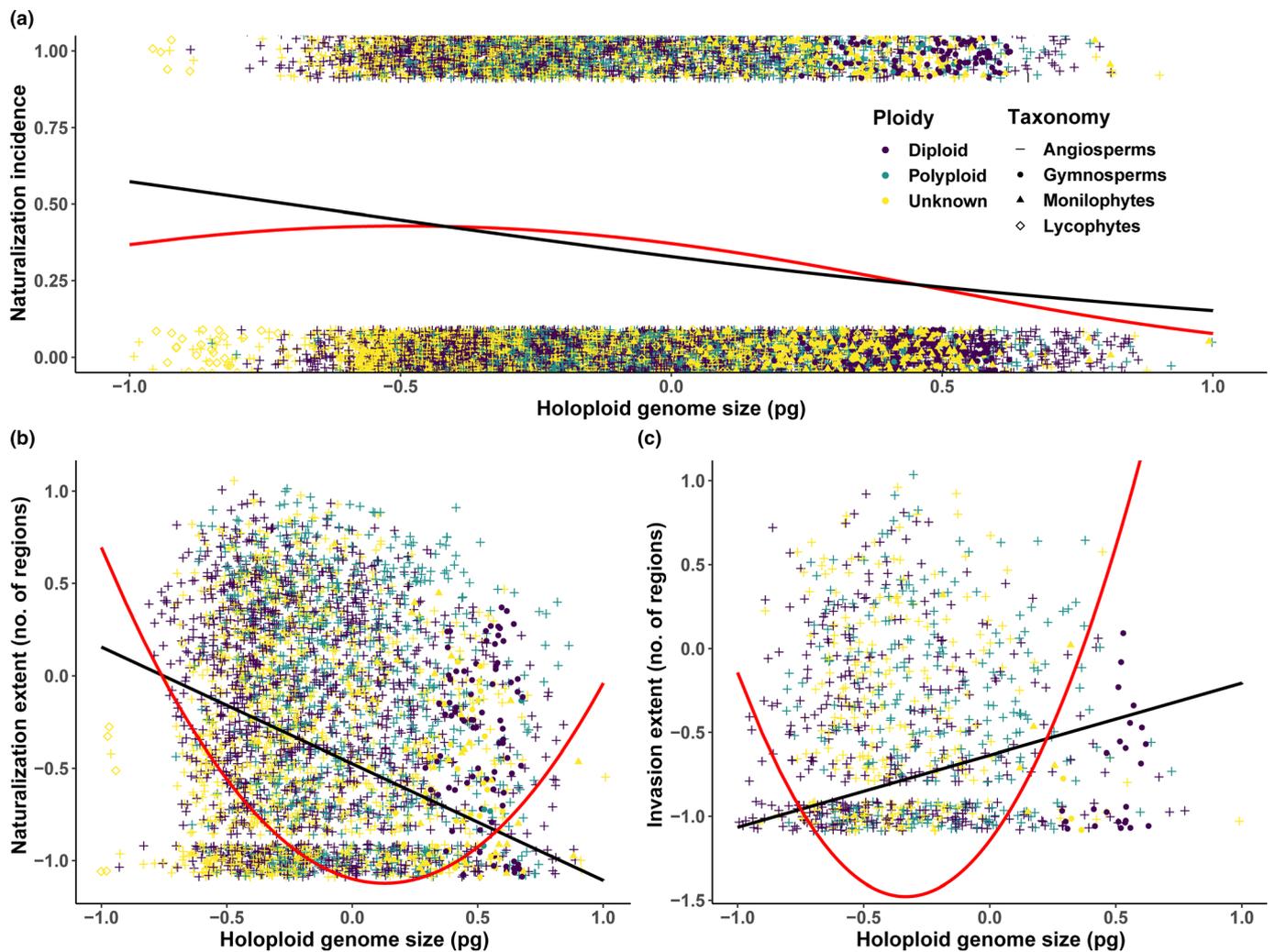
Fig. 3 Effects of the holoploid genome size (*C*-values) on (a) naturalization incidence (i.e. whether or not the species is reported as naturalized from at least one region in the GloNAF database), (b) naturalization extent (i.e. the number of GloNAF regions) and (c) invasion extent. Black and red solid lines correspond to a linear and quadratic fits, respectively. Colours and shapes represent the ploidy status and taxonomy groups, respectively. The detailed statistical summary is shown in Table 1. The naturalization extent and invasion extent were $\log_{10}$-transformed and then were standardized, as was the holoploid genome size, to zero mean and standard deviations of one. To improve the visibility, points were jittered.

monoploid genome size had a greater effect than the ploidy level (diploid, tetraploid and hexaploid) on naturalization extent but not on invasion extent, as here the ploidy level had a positive effect, while the monoploid genome size had no significant effect (Table S9).

Discussion

Taxonomic patterns in naturalization and invasion

Our analysis of naturalization and invasion success of angiosperms, gymnosperms, monilophytes and lycophytes (Table S1) reflected that in particular vascular plant groups, relationships between genome size and alien species success differed, depending on genomic differences between the major land plant lineages (Wood *et al.*, 2009; Leitch & Leitch, 2012) including, among other factors, the frequency of polyploid speciation (Wood

et al., 2009; Van de Peer *et al.*, 2017). Gymnosperms stand out as the group with the largest mean genome size of all major land plant lineages (Pellicer *et al.*, 2018). They are only rarely polyploid, have a much narrower range of genome sizes than the other taxonomic groups (Pellicer *et al.*, 2018), and woodiness is the prevailing life form. This may contribute to explaining why the values we obtained for gymnosperms contrast with the overall trend we found in other groups, that is, a small genome is an advantage for naturalization and a large genome for being invasive. It should be noted, however, that monilophytes and lycophytes (2.4%) and gymnosperms (3.4%) comprise only a small proportion of the total data set, which is dominated by angiosperms (94.2%). As the analyses of the effects of genome size and ploidy levels did not reveal significant differences between taxonomic groups in how these karyological characteristics shape global patterns of naturalization and invasion, the results discussed below refer to trends driven by angiosperms.
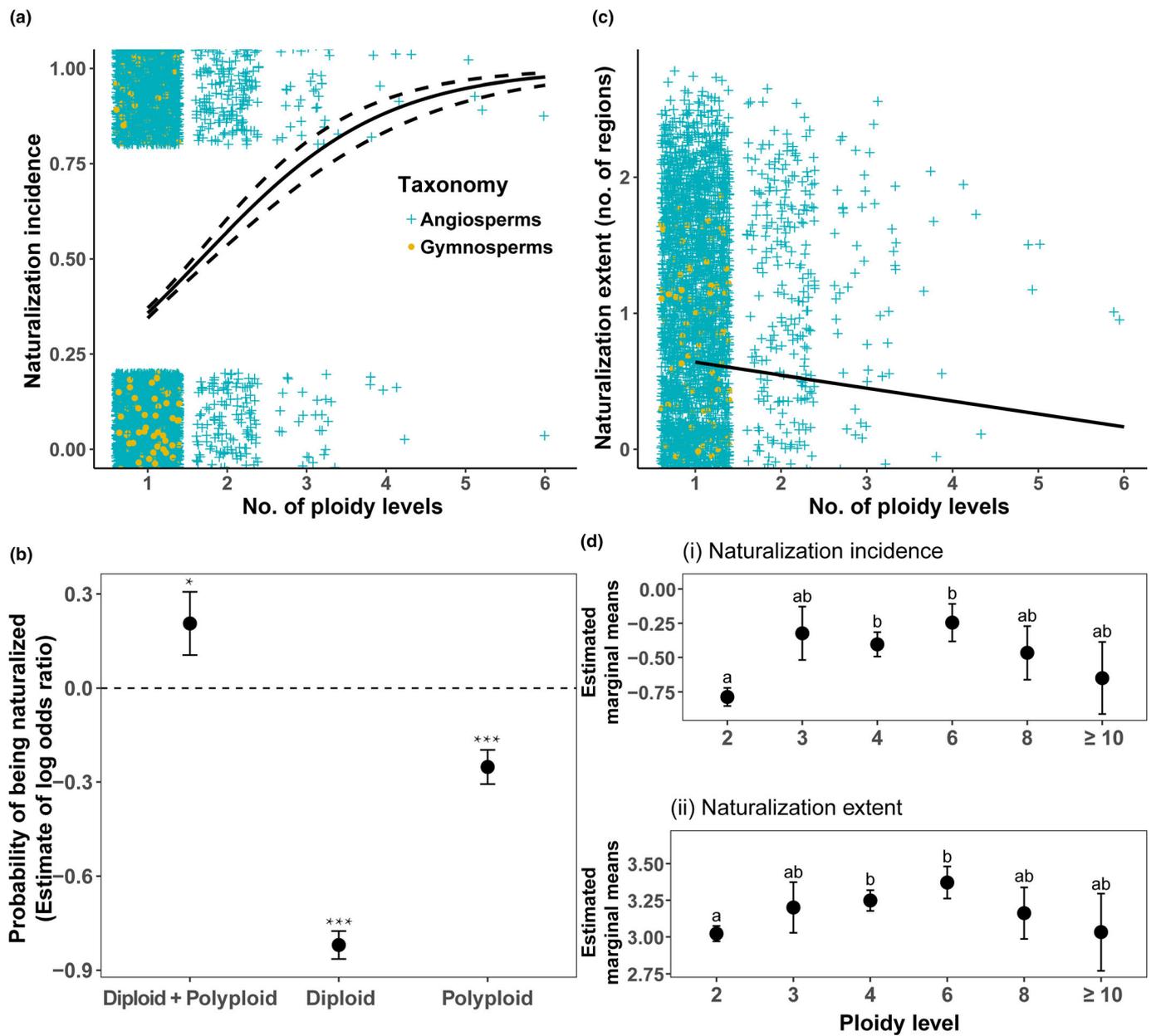
Fig. 4 Effects of ploidy level on species (a) naturalization incidence and (c) naturalization extent. The regression lines in (a) were from a phylogenetic logistic regression model, and in (c) from a phylogenetic generalized linear regression model (Table 2). In (a), the dashed lines are 95% confidence intervals. (b) Effects of ploidy types on naturalization incidence. Estimates (log odds ratios of the probability of being naturalized) of the effects and the standard errors (bars) were obtained from phylogenetic logistic regression models (Supporting Information Table S5). *, $P < 0.05$; ***, $P < 0.001$. (d) Multiple comparisons between different ploidy levels for naturalization success: naturalization incidence; naturalization extent. As the relationship between ploidy level and invasion extent was not significant, the plot is not shown here (Table S7). Y-axis represents the estimated marginal means and standard errors (bars) for each ploidy level, the different letters in each subplot indicate significant differences between ploidy levels ($P < 0.05$).

Table 2 Effects of the number of ploidy levels on species' naturalization incidence and naturalization extent.

Models	Estimate	SE	Z-value	P-value	R^2	Phylogenetic signal
Naturalization incidence						
No. of ploidy levels	0.872	0.071	12.297	< 0.001	0.048	0.159
Naturalization extent						
No. of ploidy levels	−0.095	0.028	−3.377	< 0.001	0.003	0.394

We did not run a similar analysis for invasion extent because only a few species with more than one ploidy level also had available invasion extent data (Supporting Information Table S3).

A small genome supports the naturalization of plants but constrains invasive spread

Presenting the effects of both monoploid and holoploid genome size on our measures of species' success at different invasion stages makes it possible to disentangle, to some extent, the contrasting roles played by genome size and ploidy. Nevertheless, we recognise the difficulty of fully disentangling these two karyological features given that a robust definition of what constitutes a polyploid is challenging as all seed plants are considered to have undergone at least one round of whole genome duplication in their evolutionary history (Liu *et al.*, 2022). For monoploid genome size, all three analysed measures (naturalization incidence; naturalization extent; invasion extent) decrease with increasing genome size, which, in general, confirms previous findings reporting a negative relationship between an alien plant's success and its genome size (Kubešová *et al.*, 2010; Pandit *et al.*, 2014; Pyšek *et al.*, 2015). Yet, the comparison of quadratic fits of the relationship between naturalization extent vs invasion extent and monoploid genome size (Fig. 2b,c, respectively) indicates that while possessing a small monoploid genome can be a constraint for successful invasion, it is much less so, if at all, for naturalization extent. Therefore, our results suggest that a negative relationship between genome size and naturalization, previously only reported for measures corresponding to our 'naturalization incidence' (as in Kubešová *et al.*, 2010), also holds for naturalization extent but remains more complex when it comes to invasion extent.

The analyses of holoploid genome size, where both genome size and polyploidy are combined, highlight the opposite effects of these two karyological characters on naturalization and invasion, and suggest that different mechanisms are operating at these two stages of the invasion process. For naturalization success, measured as either incidence or extent, the effect of holoploid genome size is the same as that of monoploid – that is, the success of naturalization decreases with increasing holoploid genome size; thus, species with small genomes are more successful than those with larger genomes. However, the shapes of the quadratic relationships (Fig. 3b,c) are very different from the analyses based on monoploid genome size (i.e. compare with Fig. 2b,c). At the lower end of the range of holoploid genome sizes, the extent of both naturalization and invasion decreases with increasing size of the holoploid genome, reflecting the negative impacts of traits associated with increasing genome sizes (i.e. 'the large genome constraint' hypothesis; Knight *et al.*, 2005). However, with yet further increases in genome size, the positive effect of higher ploidies (as documented by the separate analyses of ploidy levels; Table 2; see also te Beest *et al.*, 2012) prevails, leading to species with larger holoploid genomes increasing in their extent of naturalization and invasion. Moreover, the advantages of polyploidy appear much more pronounced for invasion extent than naturalization extent, as indicated by the steeper curve of the quadratic fit (Fig. 3c vs 3b, respectively). This corresponds to our finding that the relationship observed for the holoploid genome size as a predictor is largely driven by polyploids (Fig. S4C).

One reason why a small genome size might be advantageous during the naturalization stage could be its potential association with short generation times, an evolutionary feature that plays a key role in enabling adaptation to local conditions following introduction into a new area. In addition, according to the 'large genome constraint' hypothesis (Knight *et al.*, 2005), a small genome can attain a broader range of trait states that increase the potential for adaptation under novel conditions (Suda *et al.*, 2015). Furthermore, species with smaller genomes benefit from increased phenotypic plasticity (Zenni *et al.*, 2014; Meyerson *et al.*, 2020), which is especially advantageous if plants need to cope with stressful conditions during establishment – plants with large genomes may be less tolerant of environmental stress and less plastic (Knight *et al.*, 2005). For invasive spread, however, the association of a large genome with certain traits may represent an advantage. In *Phragmites australis*, populations with larger genomes had traits favouring spread, such as increased biomass allocation to the production of seeds and above-ground runners, potentially facilitating invasion into new areas and hence becoming more widespread (Pyšek *et al.*, 2018, 2019). The latter trait, representing vegetative regeneration and spread based on the dispersal of vegetative plant parts, has been repeatedly shown to be associated with the successful invasion not only of *P. australis* (Čuda *et al.*, 2021) but of vascular plants in general (Pyšek, 1997; Pyšek & Richardson, 2007).

Overall, genome size has been shown to be related to a wide range of biological traits at cellular, anatomical, morphological and physiological levels, and with ecological and evolutionary consequences in plants and in animals (Bennett & Leitch, 2005b; Gregory, 2005). According to the so-called nucleotypic theory (Bennett, 1972), genome size is known to influence, for example, the duration of cell division, the size of the nucleus, and thus the whole cell and its minimum generation time (Bennett, 1972, 1973; Edwards & Endrizzi, 1975; Cavalier-Smith, 1985). Many correlative studies have suggested that genome size can be viewed as an adaptive trait influencing, for example minimum generation time or time to flowering (Leitch & Bennett, 2007; Jian *et al.*, 2017), seed characteristics (Grotkopp *et al.*, 2004; Beaulieu *et al.*, 2007; Krahulcová *et al.*, 2017), relative growth rates of seedlings (Grotkopp *et al.*, 2004), specific leaf area (Morgan & Westoby, 2005; Beaulieu *et al.*, 2007), stomatal size and density (Beaulieu *et al.*, 2008; Hodgson *et al.*, 2010), pollen size (Knight *et al.*, 2010) and flower size (Meagher & Vassiliadis, 2005), as well as underpinning relationships with environmental gradients (Knight & Ackerly, 2002), invasiveness (Suda *et al.*, 2015), the evolution of specific phylogenetic groups (Chrtek *et al.*, 2009; Luo *et al.*, 2017) and phenotypic plasticity (Meyerson *et al.*, 2020; Faizullah *et al.*, 2021). Exploring the relationship between genome size and physiological processes influencing water use efficiency, photosynthesis and nutrient demands have provided strong evidence suggesting environmental selection on genome size which, in turn, is likely to play a role in influencing the evolutionary trajectory of plants.

However, it needs to be noted that associations between genome size and traits are often very complex, resulting in trade-offs that ultimately determine the extent of the impact of genome size variation on different aspects of a plant's biology (Faizullah *et al.*, 2021). Therefore, for some traits, we cannot say with

certainty whether relationships reported between genome size and specific traits are underpinned by the nucleotypic effects of genome size or are just correlations arising from genome size co-varying with other traits that are driving the reported relationships. In many cases, all we can say is that the genome sizes of plants are frequently associated with these traits. One example of a trait that is directly impacted by genome size itself, and in turn could constrain invasion success, is the duration of meiosis and mitosis, which are both positively correlated with genome size. For a species with a big genome, the duration of meiosis and mitosis cannot be short because of the time needed to replicate all the DNA in a large genome compared with a small genome. Similarly, a species with a large genome cannot have small cells (one cannot fit a big genome in a small cell); hence, for example guard cells are restricted to being large, which in turn can impact gas exchange and hence photosynthesis and water use efficiency.

With regard to the above-discussed results, it needs to be noted that while the relationships between genome size metrics and naturalization/invasion extent were highly significant, the genome size mostly explained less than 10% of the variation in the measured variables (with the exception of quadratic relationships with invasion extent; Table 1). This suggests that many other factors are at play in determining the success of alien species (Pyšek *et al.*, 2020a).

Polyploidy and variation in ploidy levels favour naturalization but limit invasive spread

Pandit *et al.* (2014), using a data set of 890 plant species, showed that invasiveness was negatively related to genome size and positively to chromosome number (which they used to infer ploidy levels) and that including both these karyological traits in models increased their explanatory power. However, they only tested the effect of holoploid genome size (1*C*-value) and found a positive relationship between genome size and chromosome numbers, similar to the positive associations between genome size and polyploidy revealed in this study (Fig. S5). Here, by exploring both monoploid and holoploid genome size effects, we were able to begin disentangling the contribution of both karyological characteristics to species' invasion success and interpret the results within the context of the different stages of the invasion process.

Like genome size, polyploidy also had contrasting effects on naturalization and invasion in our study. The number of different ploidy levels reported for a species, as a measure of variation, was positively related to naturalization incidence but negatively related to naturalization extent. However, the pattern is more complex. Species reported to be diploid only were less likely to naturalize than species with a higher number of different ploidy levels. This is probably caused by the smaller monoploid genome size of polyploids compared to diploids (Fig. S6), which favours naturalization success. However, naturalization incidence was highest for species that occurred both as diploids and polyploids (Fig. 4b). The explanation here could be that polyploidization brings anatomical, physiological and morphological changes, which may lead to improved fitness and a shift in ecological requirements in novel cytotypes, enhancing their ability to

survive and adapt to novel environmental conditions (Levin, 2002; Otto, 2007). Moreover, heteroploidy systems are even more variable because of gene flow between the ploidies, which enables the establishment of newly formed polyploids (Burton & Husband, 2001; Čertner *et al.*, 2017; Kolář *et al.*, 2017). We suggest that diploid–polyploid combinations maximize the variability of many species' characteristics, providing a greater genetic basis for adaptation under new conditions that facilitates naturalization.

A key assumption in the study of polyploidy is that species with a greater number of ploidy levels are likely to be more widespread because the newly originated polyploids need to escape from the competition with their ancestors. To achieve this, polyploids may shift their ecological niche and/or geographic distribution, and/or differentiate from their diploid ancestors in a process resulting in the evolution of novel traits and avoidance of cytotype exclusion. Moreover, the unique traits of individual polyploid cytotypes may enhance opportunities to colonize novel environments (Kolář *et al.*, 2017; Rice *et al.*, 2019). So far, among angiosperms, *c.* 16% of species analysed extensively at the population level have been reported to have several ploidy levels. The existence of multiple ploidies increases the intraspecific diversity of a species providing a potential advantage for survival (Kolář *et al.*, 2017). However, our results suggest that the explanations for the better performance of species with a greater variation in ploidy levels hold more for the population establishment stage linked with species naturalization than for the invasion stage – an observation that is similar to the contrasting effects of genome size on these two stages.

Whether the higher success rate of naturalization in species with higher numbers of ploidy levels is also underpinned by their higher diversification (Landis *et al.*, 2018) and speciation rates compared with diploids (Puttik *et al.*, 2019; Soltis & Soltis, 2020) is unclear. One interesting additional analysis providing insights into this issue could be to distinguish the effect of allopolyploids, resulting from hybridization and potentially benefitting from fixed heterosis, from autopolyploids that are of nonhybrid origin. Unfortunately, testing this factor was not possible due to the lack of information to separate the two types of polyploidy that is only available for *c.* 20% of species in our data set. Nevertheless, if the introduction of a species to a novel environment is sufficiently stressful to trigger the formation of polyploid variants and results in a higher species diversification rate and given that higher diversification rates have recently been shown to lead to higher naturalization success (Lenzner *et al.*, 2021), it is possible that the interplay between polyploidy and increased diversification rates may also play a role in supporting the naturalization success.

Competing interests

None declared.

Author contributions

PP, ML and W-YG conceived the ideas and designed the methodology. WD, FE, HK, IJL, BL, ML, JP, PP, MvK, PW and MW prepared the data. W-YG analysed the data, PP, W-YG and ML led the writing of the manuscript, and all authors contributed to the writing.

ORCID

Wayne Dawson https://orcid.org/0000-0003-3402-0774
Franz Essl https://orcid.org/0000-0001-8253-2112
Wen-Yong Guo https://orcid.org/0000-0002-4737-2042
Mark van Kleunen https://orcid.org/0000-0002-2861-3701
Holger Kreft https://orcid.org/0000-0003-4471-8236
Ilia J. Leitch https://orcid.org/0000-0002-3837-8186
Bernd Lenzner https://orcid.org/0000-0002-2616-3479
Magdalena Lučanová https://orcid.org/0000-0002-4612-3693
Laura A. Meyerson https://orcid.org/0000-0002-1283-3865
Jan Pergl https://orcid.org/0000-0002-0045-1974
Petr Pyšek https://orcid.org/0000-0001-8500-442X
Patrick Weigelt https://orcid.org/0000-0002-2485-3708
Marten Winter https://orcid.org/0000-0002-9593-7300

Data availability

The data used in the paper, GloNAF (2) and Plant DNA *C*-values Database (release 7.1, April 2019, Leitch IJ, Johnston E, Pellicer J, Hidalgo O, Bennett MD; https://cvalues.science.kew.org) are publicly available. All data and R codes used in the study are also available on GitHub (https://github.com/wyeco/GS_glonaf).

References

Amsellem L, Chevallier MH, Hossaert-McKey M. 2001. Ploidy level of the invasive weed *Rubus alceifolius* (Rosaceae) in its native range and in areas of introduction. *Plant Systematics and Evolution* **228**: 171–179.

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using LME4. *Journal of Statistical Software* **67**: 1–48.

Beaulieu JM, Leitch IJ, Patel S, Pendharkar A, Knight CA. 2008. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist* **179**: 975–986.

Beaulieu JM, Moles AT, Leitch IJ, Bennett MD, Dickie JB, Knight CA. 2007. Correlated evolution of genome size and seed mass. *New Phytologist* **173**: 422–437.

te Beest M, Le Roux JJ, Richardson DM, Brysting AK, Suda J, Kubešová M, Pyšek P. 2012. The more the better? The role of polyploidy in facilitating plant invasions. *Annals of Botany* **109**: 19–45.

Bennett MD. 1972. Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society B: Biological Sciences* **181**: 109–135.

Bennett MD. 1973. Nuclear characters in plants. *Brookhaven Symposium in Biology* **25**: 344–366.

Bennett MD, Leitch IJ. 2005a. Nuclear DNA amounts in angiosperms: progress, problems and prospects. *Annals of Botany* **95**: 45–90.

Bennett MD, Leitch IJ. 2005b. Genome size evolution in plants. In: Gregory TR, ed. *The evolution of the genome.* San Diego, CA, USA: Elsevier, 89–162.

Bennett MD, Smith JB. 1976. Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society London B: Biological Sciences* **274**: 227–274.

Blackburn TM, Pyšek P, Bacher S, Carlton JT, Duncan RP, Jarošík V, Wilson JRU, Richardson DM. 2011. A proposed unified framework for biological invasions. *Trends in Ecology & Evolution* **26**: 333–339.

Blanchet S, Leprieur F, Beauchard O, Staes J, Oberdorff T, Brosse S. 2009. Broad-scale determinants of non-native fish species richness are context-dependent. *Proceedings of the Royal Society B: Biological Sciences* **276**: 2385–2394.

Burton T, Husband B. 2001. Fecundity and offspring ploidy in matings among diploid, triploid and tetraploid *Chamerion angustifolium* (Onagraceae): consequences for tetraploid establishment. *Heredity* **87**: 573–582.

Cavalier-Smith T. 1985. *The evolution of genome size.* Chichester, UK: J. Wiley.

Cayuela L, Oksanen J. 2014. *Taxonstand: taxonomic standardization of plant species names, v. 1.6.* Vienna, Austria: R Foundation for Statistical Computing.

Čertner M, Fenclová E, Kúr P, Kolář F, Koutecký P, Krahulcová A, Suda J. 2017. Evolutionary dynamics of mixed-ploidy populations in an annual herb: dispersal, local persistence and recurrent origins of polyploids. *Annals of Botany* **120**: 303–315.

Chrtek J, Zahradníček J, Krak K, Fehrer J. 2009. Genome size in *Hieracium* subgenus *Hieracium* (Asteraceae) is strongly correlated with major phylogenetic groups. *Annals of Botany* **104**: 161–178.

Convention on Biological Diversity. 2000. *Alien species that threaten ecosystems, habitats or species.* Nairobi, Kenya: UNEP/CBD/COP/5/8, Secretariat of the Convention on Biological Diversity.

Čuda J, Skálová H, Meyerson LA, Pyšek P. 2021. Regeneration of *Phragmites australis* from rhizome and culm fragments: an experimental test of environmental effects, population origin and invasion status. *Preslia* **93**: 237–254.

Dawson W, Moser D, van Kleunen M, Kreft H, Pergl J, Pyšek P, Weigelt P, Winter M, Lenzner B, Blackburn TM et al. 2017. Global hotspots and correlates of alien species richness across taxonomic groups. *Nature Ecology & Evolution* **1**: 186.

Delavaux CS, Weigelt P, Dawson W, Duchicela J, Essl F, van Kleunen M, König C, Pergl J, Pyšek P, Stein A et al. 2019. Mycorrhizal fungi influence global plant geography. *Nature Ecology & Evolution* **3**: 424–429.

Dellinger AS, Essl F, Hojsgaard D, Kirchheimer B, Klatt S, Dawson W, Pergl J, Pyšek P, van Kleunen M, Weber E et al. 2016. Niche dynamics of alien species do not differ among sexual and apomictic flowering plants. *New Phytologist* **209**: 1313–1323.

DeWalt SJ, Hamrick JL. 2004. Genetic variation of introduced Hawaiian and native Costa Rican populations of an invasive tropical shrub, *Clidemia hirta* (Melastomataceae). *American Journal of Botany* **91**: 1155–1162.

Edwards GA, Endrizzi JE. 1975. Cell size, nuclear size, and DNA content relationship in *Gossypium*. *Canadian Journal of Genetics and Cytology* **17**: 181–186.

Ehrendorfer F. 1980. Polyploidy and distribution. In: Lewis WH, ed. *Polyploidy: biological relevance.* New York, NY, USA; London, UK: Plenum Press, 45–60.

Essl F, Dawson W, Kreft H, Pergl J, Pyšek P, van Kleunen M, Weigelt MT, Dullinger S, Lenzner B et al. 2019. Drivers of the relative richness of naturalized and invasive plant species on the Earth. *AoB Plants* **11**: plz051.

Faizullah L, Morton JA, Hersch-Green EI, Walczyk AM, Leitch AR, Leitch IJ. 2021. Exploring environmental selection on genome size in angiosperms. *Trends in Plant Science* **26**: 1039–1049.

Fridley JD, Craddock A. 2015. Contrasting growth phenology of native and invasive forest shrubs mediated by genome size. *New Phytologist* **207**: 659–668.

Galbraith D, Loureiro J, Antoniadi I, Bainard J, Bureš P, Cápal P, Castro M, Castro S, Čertner M, Čertnerová D et al. 2021. Best practices in plant flow cytometry. *Cytometry. Part A* **99**: 311–317.

Garcia S, Canela MÁ, Garnatje T, McArthur ED, Pellicer J, Sanderson SC, Vallès J. 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.

Gravuer K, Sullivan JJ, Williams PA, Duncan RP. 2008. Strong human association with plant invasion success for *Trifolium* introductions to New Zealand. *Proceedings of the National Academy of Sciences, USA* 105: 6344–6349.

Gregory TR. 2005. The *C*-value enigma in plants and animals: a review of parallels and an appeal for partnership. *Annals of Botany* 95: 133–146.

Greilhuber J, Doležel J, Lysak MA, Bennett MD. 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, Rejmánek M, Rost TL. 2002. Toward a causal explanation of plant invasiveness: seedling growth and life-history strategies of 29 pine (*Pinus*) species. *The American Naturalist* 159: 396–419.

Grotkopp E, Rejmánek M, Sanderson MJ, Rost TL. 2004. Evolution of genome size in pines (*Pinus*) and its life history correlates: supertree analysis. *Evolution* 58: 1705–1729.

Guggisberg A, Mansion G, Kelso S, Conti E. 2006. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid–polyploid species complex of *Primula*. *New Phytologist* 171: 617–632.

Guignard MS, Nichols RA, Knell RJ, Macdonald A, Romila C-A, Trimmer M, Leitch IJ, Leitch AR. 2016. Genome size and ploidy influence angiosperm species biomass under nitrogen and phosphorus limitation. *New Phytologist* 210: 1195–1206.

Guo W-Y, van Kleunen M, Winter M, Weigelt P, Stein A, Pierce S, Pergl J, Moser D, Maurel N, Lenzner B *et al*. 2018. The role of adaptive strategies in plant naturalization. *Ecology Letters* 21: 1380–1389.

Ho LST, Ané C. 2014. *Package 'PHYLOLM': phylogenetic linear regression, v. 2.2.* [WWW document] URL https://cran.r-project.org/web/packages/phylolm/index.html [accessed 8 June 2022].

Hodgson JG, Sharafi M, Jalili A, Díaz S, Montserrat-Martí G, Palmer C, Cerabolini B, Pierce S, Hamzehee B, Asri Y *et al*. 2010. Stomatal vs genome size in angiosperms: the somatic tail wagging the genomic dog? *Annals of Botany* 105: 573–584.

Ives AR. 2019. *R*²s for correlated data: phylogenetic models, LMMs, and GLMMs. *Systematic Biology* 68: 234–251.

Ives AR, Garland T. 2010. Phylogenetic logistic regression for binary dependent variables. *Systematic Biology* 59: 9–26.

Ives AR, Li D. 2018. RR2: an R package to calculate *R*²s for regression models. *Journal of Open Source Software* 3: 1028.

Jian Y, Xu C, Guo Z, Wang S, Xu Y, Zou C. 2017. Maize (*Zea mays* L.) genome size indicated by 180-bp knob abundance is associated with flowering time. *Scientific Reports* 7: 5954.

Jin Y, Qian HV. 2019. PHYLOMAKER: an R package that can generate very large phylogenies for vascular plants. *Ecography* 42: 1353–1359.

Kattge J, Bönisch G, Díaz S, Lavorel S, Prentice IC, Leadley P, Tautenhahn S, Werner GDA, Aakala T, Abedi M *et al*. 2020. TRY plant trait database: enhanced coverage and open access. *Global Change Biology* 26: 119–188.

van Kleunen M, Dawson W, Essl F, Pergl J, Winter M, Weber E, Kreft H, Weigelt P, Kartesz J, Nishino M *et al*. 2015. Global exchange and accumulation of non-native plants. *Nature* 525: 100–103.

van Kleunen M, Johnson SD, Fischer M. 2007. Predicting naturalization of southern African Iridaceae in other regions. *Journal of Applied Ecology* 44: 594–603.

van Kleunen M, Pyšek P, Dawson W, Essl F, Kreft H, Pergl J, Weigelt P, Stein A, Dullinger S, König C *et al*. 2019. The Global Naturalized Alien Flora (GloNAF) database. *Ecology* 100: e02542.

van Kleunen M, Weber E, Fischer M. 2010. A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters* 13: 235–245.

Knight CA, Ackerly DD. 2002. Genome size variation across environmental gradients in the California flora. *Ecology Letters* 5: 66–76.

Knight CA, Clancy RB, Götzenberger L, Dann L, Beaulieu JM. 2010. On the relationship between pollen size and genome size. *Journal of Botany* 2010: 612017.

Knight CA, Molinari NA, Petrov DA. 2005. The large genome constraint hypothesis: evolution, ecology and phenotype. *Annals of Botany* 95: 177–190.

Kolář F, Čertner M, Suda J, Schönswetter P, Husband BC. 2017. Mixed-ploidy species: progress and opportunities in polyploid research. *Trends in Plant Science* 22: 1041–1055.

Krahulcová A, Trávníček P, Krahulec F, Rejmánek M. 2017. Small genomes and large seeds: chromosome numbers, genome size and seed mass in diploid *Aesculus* species (Sapindaceae). *Annals of Botany* 119: 957–964.

Kubešová M, Moravcová L, Suda J, Jarošík V, Pyšek P. 2010. Naturalized plants have smaller genomes than their non-invading relatives: a flow cytometric analysis of the Czech alien flora. *Preslia* 82: 81–96.

Landis JB, Soltis DE, Li Z, Marx HE, Barker MS, Tank DC, Soltis PS. 2018. Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany* 105: 348–363.

Leitch AR, Leitch IJ. 2012. Ecological and genetic factors linked to contrasting genome dynamics in seed plants. *New Phytologist* 194: 629–646.

Leitch IJ, Bennett MD. 2007. Genome size and its uses: the impact of flow cytometry. In: Doležel J, Greilhuber J, Suda J, eds. *Flow cytometry with plant cells. Analysis of genes, chromosomes and genomes*. Weinheim, Germany: Wiley-VCH, 153–176.

Lenth R. 2020. *Emmeans: estimated marginal means, aka least-squares means, v. 1.4.4.* R package version. [WWW document] URL https://CRAN.R-project.org/package=emmeans [accessed 8 June 2022].

Lenzner B, Magallón S, Dawson W, Kreft H, König C, Pergl J, Pyšek P, Weigelt P, van Kleunen M, Winter M *et al*. 2021. Role of diversification rates and evolutionary history as a driver of plant naturalization success. *New Phytologist* 229: 2998–3008.

Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v.4: recent updates and new developments. *Nucleic Acids Research* 47: W256–W259.

Levin DA. 1983. Polyploidy and novelty in flowering plants. *The American Naturalist* 122: 1–25.

Levin DA. 2002. *The role of chromosomal change in plant evolution*. New York, NY, USA: Oxford University.

Liu Y, Wang S, Li L, Yang T, Dong S, Wei T, Wu S, Liu Y, Gong Y, Feng X *et al*. 2022. The *Cycas* genome and the early evolution of seed plants. *Nature Plants* 8: 389–401.

Luo C, Chen D, Cheng X, Zhao H, Huang C. 2017. Genome size estimations in *Chrysanthemum* and correlations with molecular phylogenies. *Genetic Resources and Crop Evolution* 64: 1451–1463.

Meagher TR, Vassiliadis C. 2005. Phenotypic impacts of repetitive DNA in flowering plants. *New Phytologist* 168: 71–80.

Meyerson LA, Cronin JT, Bhattarai GP, Brix H, Lambertini C, Lučanová M, Rinehart S, Suda J, Pyšek P. 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.

Meyerson LA, Pyšek P, Lučanová M, Wigginton S, Tran C-T, Cronin JT. 2020. Plant genome size influences stress tolerance of invasive and native plants via plasticity. *Ecosphere* 11: e03145.

Moodley D, Geerts S, Richardson DM, Wilson JRU. 2013. Different traits determine introduction, naturalization and invasion success in woody plants: Proteaceae as a test case. *PLoS ONE* 8: e75078.

Morgan HD, Westoby M. 2005. The relationship between nuclear DNA content and leaf strategy in seed plants. *Annals of Botany* 96: 1321–1330.

Naiki A, Nagamasu H. 2004. Correlation between distyly and ploidy level in *Damnacanthus* (Rubiaceae). *American Journal of Botany* 91: 664–671.

Otto SP. 2007. The evolutionary consequences of polyploidy. *Cell* 131: 452–462.

Pagad S, Genovesi P, Carnevali L, Schigel D, McGeoch MA. 2018. Introducing the global register of introduced and invasive species. *Scientific Data* 5: 170202.

Pandit MK, White SM, Pocock MJO. 2014. The contrasting effects of genome size, chromosome number and ploidy level on plant invasiveness: a global analysis. *New Phytologist* 203: 697–703.

Pellicer J, Hidalgo O, Dodsworth S, Leitch IJ. 2018. Genome size diversity and its impact on the evolution of land plants. *Genes (Basel)* 9: 88.

Puttik MN, Clark C, Donoghue PCJ. 2019. Size is not everything: rates of genome size evolution, not *C*-value, correlate with speciation in angiosperms. *Proceedings of the Royal Society B: Biological Sciences* 282: 20152289.

Pyšek P. 1997. Clonality and plant invasions: can a trait make a difference? In: de Kroon H, van Groenendael J, eds. *The ecology and evolution of clonal plants.* Leiden, the Netherlands: Backhuys Publishers, 405–427.

Pyšek P, Bacher S, Kühn I, Novoa A, Catford JA, Hulme PE, Pergl J, Richardson DM, Wilson JRU, Blackburn TM. 2020a. MAcroecological Framework for Invasive Aliens (MAFIA): disentangling large-scale context-dependence in biological invasions. *NeoBiota* 62: 407–461.

Pyšek P, Hulme PE, Simberloff D, Bacher S, Blackburn TM, Carlton JT, Dawson W, Essl F, Foxcroft LC, Genovesi P *et al.* 2020b. Scientists' warning on invasive alien species. *Biological Reviews* 95: 1511–1534.

Pyšek P, Jarošík V, Pergl J, Randall R, Chytrý M, Kühn I, Tichý L, Danihelka J, Chrtek J, Sádlo J. 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, Manceur AM, Alba C, McGregor KF, Pergl J, Štajerová K, Chytrý M, Danihelka J, Kartesz J, Klimešová J *et al.* 2015. Naturalization of central European plants in North America: species traits, habitats, propagule pressure, residence time. *Ecology* 96: 762–774.

Pyšek P, Pergl J, Essl F, Lenzner B, Dawson W, Kreft H, Weigelt P, Winter M, Kartesz J, Nishino M *et al.* 2017. Naturalized alien flora of the world: species diversity, taxonomic and phylogenetic patterns, geographic distribution and global hotspots of plant invasion. *Preslia* 89: 203–274.

Pyšek P, Richardson DM. 2007. Traits associated with invasiveness in alien plants: where do we stand? In: Nentwig W, ed. *Biological invasions. Ecological studies, vol. 193.* Berlin, Heidelberg, Germany: Springer, 97–125.

Pyšek P, Skálová H, Čuda J, Guo W-Y, Doležal J, Kauzál O, Lambertini C, Pyšková K, Brix H, Meyerson LA. 2019. Physiology of a plant invasion: biomass production, growth and tissue chemistry of invasive and native *Phragmites australis* populations. *Preslia* 91: 51–75.

Pyšek P, Skálová H, Čuda J, Guo W-Y, Suda J, Doležal J, Kauzál O, Lambertini C, Lučanová M, Mandáková T *et al.* 2018. Small genome separates native and invasive populations in an ecologically important cosmopolitan grass. *Ecology* 99: 79–90.

R Core Team. 2021. *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. [WWW document] URL https://www.R-project.org/ [accessed 8 June 2022].

Razanajatovo M, Maurel N, Dawson W, Essl F, Kreft H, Pergl J, Pyšek P, Weigelt P, Winter M, van Kleunen M. 2016. Plants capable of selfing are more likely to become naturalized. *Nature Communications* 7: 13313.

Razanajatovo M, van Kleunen M, Kreft H, Dawson W, Essl F, Pergl J, Pyšek P, Winter M, Weigelt P. 2019. Autofertility and self-compatibility moderately benefit Island colonization of plants. *Global Ecology and Biogeography* 28: 341–352.

Rejmánek M. 1996. A theory of seed plant invasiveness: the first sketch. *Biological Conservation* 78: 171–181.

Revell LJ. 2012. PHYTOOLS: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.

Rice A, Šmarda P, Novosolov M, Drori M, Glick L, Sabath N, Meiri S, Belmaker J, Mayrose I. 2019. The global biogeography of polyploid plants. *Nature Ecology & Evolution* 3: 265–273.

Richardson DM, Pyšek P, Rejmánek M, Barbour MG, Panetta FD, West CJ. 2000. Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions* 6: 93–107.

Schielzeth H. 2010. Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* 1: 103–113.

Schierenbeck KA, Ainouche ML. 2006. The role of evolutionary genetics in studies of plant invasions. In: Cadotte MW, McMahon SM, Fukami F, eds. *Conceptual ecology and invasion biology: reciprocal approaches to nature.* Dordrecht, the Netherlands: Springer, 193–221.

Sliwinska E, Loureiro J, Leitch IJ, Šmarda P, Bainard J, Bureš P, Chumová Z, Horová L, Koutecký P, Lučanová M *et al.* 2022. Application-based guidelines for best practices in plant flow cytometry. *Cytometry. Part A* 101: 749–781.

Šmarda P, Hejcman M, Březinová A, Horová L, Steigerová H, Zedek F, Bureš P, Hejcmanová P, Schellberg J. 2013. Effect of phosphorus availability on the selection of species with different ploidy levels and genome sizes in a long-term grassland fertilization experiment. *New Phytologist* 200: 911–921.

Smith SA, Brown JW. 2018. Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany* 105: 302–314.

Soltis DE, Soltis PS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology & Evolution* 14: 348–351.

Soltis PS, Soltis DE. 2020. Plant genomes: markers of evolutionary history and drivers of evolutionary change. *Plants People Planet* 3: 74–82.

Stebbins GL. 1985. Polyploidy, hybridization and the invasion of new habitats. *Annals of the Missouri Botanical Garden* 72: 824–832.

Suda J, Meyerson LA, Leitch I, Pyšek P. 2015. The hidden side of plant invasions: the role of genome size. *New Phytologist* 205: 994–1007.

Thompson JD, Lumaret R. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. *Trends in Ecology & Evolution* 7: 302–307.

Thuiller W, Richardson DM, Rouget M, Proches Ş, Wilson JRU. 2006. Interactions between environment, species traits and human uses describe patterns of plant invasions. *Ecology* 87: 1755–1769.

Van de Peer Y, Mizrachi E, Marchal K. 2017. The evolutionary significance of polyploidy. *Nature Reviews Genetics* 18: 411–424.

Wang X, Morton JA, Pellicer J, Leitch IJ, Leitch AR. 2021. Genome downsizing after polyploidy: mechanisms, rates and selection pressures. *The Plant Journal* 107: 1003–1015.

Weber E. 2017. *Invasive plant species of the world: reference guide to environmental weeds, 2nd edn.* Wallingford, UK: CABI Publishing.

Weigelt P, König C, Kreft H. 2020. GIFT – A Global Inventory of Floras and Traits for macroecology and biogeography. *Journal of Biogeography* 47: 16–43.

Wendel JF. 2015. The wondrous cycles of polyploidy in plants. *American Journal of Botany* 102: 1753–1756.

Williamson M. 2006. Explaining and predicting the success of invading species at different stages of invasion. *Biological Invasions* 8: 1561–1568.

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences, USA* 106: 13875–13879.

Zenni RD, Lamy J-B, Lamarque LJ, Porté AJ. 2014. Adaptive evolution and phenotypic plasticity during naturalization and spread of invasive species: implications for tree invasion biology. *Biological Invasions* 16: 635–644.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Relationship between the monoploid genome size and the holoploid genome size.

Fig. S2 Phylogenetic tree of the species included in the study.

Fig. S3 Effects of the monoploid genome size on naturalization incidence for eudicots and monocots.

Fig. S4 Effects of holoploid genome size (*C*-value) on naturalization incidence, naturalization extent and invasion extent.

Fig. S5 Relationships between ploidy levels and holoploid and monoploid genome size.

Fig. S6 Frequency densities of monoploid genome sizes for diploid and polyploid species.

Methods S1 Studies used to complement genome size.

Table S1 Summary of the distribution of monoploid (C_x-value), and holoploid genome sizes (*C*-value) for vascular plants together with data on the number of naturalized and invasive species.

Table S2 Summary of linear and quadratic phylogenetic regressions of the effect of monoploid genome size on naturalization incidence.

Table S3 Numbers and percentages of species by the number of ploidy levels in which they are reported to occur.

Table S4 Summary of the linear and quadratic phylogenetic regression models of the effect of holoploid genome size on naturalization incidence and naturalization extent.

Table S5 Effects of the diploids and polyploidy on species naturalization incidence, naturalization extent and invasion extent.

Table S6 Summary of the number and percentage of naturalized and invasive species for each ploidy level.

Table S7 Summary of models testing the effect of different ploidy levels on naturalization incidence, naturalization extent and invasion extent.

Table S8 Effects of the number of ploidy levels, the monoploid genome size and their interaction on species naturalization incidence, naturalization extent and invasive extent.

Table S9 Summary of the phylogenetic regression models of the effect of monoploid genome size and ploidy level on naturalization incidence and naturalization and invasion extent.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.