Global-scale drivers of ploidy state in aquatic macrophytes

Tatiana Lobato-de Magalhães[1], Kevin Murphy[2], Andrey Efremov[3], Victor Chepinoga[4], Thomas Davidson[5], and Eugenio Molina-Navarro[6]

[1]Autonomous University of Queretaro
[2]University of Glasgow
[3]Ul'anovskij gosudarstvennyj pedagogiceskij universitet imeni I N Ul'anova
[4]Central Siberian Botanical Garden SB RAS
[5]Aarhus Universitet
[6]University of Alcalá Research and Learning Resources Center

November 24, 2020

Abstract

To determine potential drivers of the global distribution of ploidy in aquatic macrophyte species we allocated ploidy state to 1572 species occurring in 238 10 × 10° gridcells worldwide. Analysis of the relationship of 16 global-scale spatial, landscape, environmental, and biotic variables with ploidy state using Boosted Regression Trees revealed temperature variables and evapotranspiration as the strongest predictors. There were contrasting latitudinal patterns between haploid/diploid and polyploid species, while species richness measures also influenced ploidy state. Polyploid species occupied larger geographical ranges than haploid/diploid species. Mixed ploidy species showed the highest latitudinal range size and maximum latitude of species occurrence. Our findings suggest that increased chromosome number is associated with tolerance of a wider range of environmental conditions in macrophyte species. Mixed ploidy could reflect adaptability to expand geographical occurrence via chromosome number change, with such species predominantly occupying a latitude range intermediate between haploid/diploid and polyploid macrophyte dominance.

Introduction

Chromosome number change (e.g., polyploidy or whole-genome duplication; dysploidy or change in single chromosome numbers) is a common genetic mutation in plant species with complex effects on plant ecological responses and several evolutionary implications (Segraves, 2017). Ploidy state ("cytotype") has been determined, initially by chromosome counts and more recently by techniques such as flow cytometry (Kron et al., 2007; Rice et al., 2015), though variation in DNA content among cytotypes is still unknown for many aquatic plant species, especially in tropical regions (Ramsey & Ramsey, 2014).

The existence of different distribution patterns both between and within species across geographical gradients such as latitude and altitude, in vascular plant populations showing different cytotypes has long been known (e.g., Tischler, 1935; Löve & Löve, 1949; Blackburn & Morton, 1957; Johnson & Packer, 1965; Hardy et al., 2000; Johnson et al., 2003; Kubátová et al., 2008; Martin & Husband, 2013). Recent work offers strong evidence for the existence of a latitude-related gradient of occurrence of polyploidy in angiosperms (Rice et al., 2019), which the authors suggested may be related to increasing climatic or other environmental stress and indirect effects of environment on life form and species richness occurring at high latitudes. In addition, there is previous evidence that ploidy state is a factor linked to range-size in angiosperms (Petit & Thompson, 1999; Lowry & Lester, 2006; Martin & Husband, 2009). Whether this applies to those vascular plant species that live in freshwater and brackish inland waterbodies ("aquatic macrophytes"), has, however, not been specifically investigated at a global-scale, and was the subject of our study.

As a group, aquatic macrophytes are generally considered to occupy a "stressful environment", in terms of factors that may reduce their productivity, in all or most of the habitats where they are found (Grime, 1979; Santamaría, 2002). These factors include light, nutrient and carbon availability, temperature regime, sediment conditions, and alkalinity (e.g., Bornette & Puijalon, 2011; Iversen et al., 2019; Sun et al, 2019). However, the intensity of environmental stress and disturbance tolerated by individual species, and sometimes among individual populations of a given species across its range varies considerably. Aquatic macrophytes consequently show a wide range of resulting life-strategies for survival and growth (e.g., Garbey et al., 2004; Lui et al., 2005; Beck & Alahuhta, 2017) and changes in number of chromosomes could be one factor that contributes to the success of such survival strategies.

Given the assumption that the latitudinal distribution pattern seen in angiosperm ploidy may be particularly related to environmental stress conditions (e.g., Stebbins, 1984; Rice et al., 2019), it is clearly of interest to explore further how ploidy state varies in plants, such as aquatic macrophytes which both occur widely across the planet and are inherently adapted to cope with potentially-high environmental stress conditions. These stressors may be associated with latitude and other spatial, environmental, and landscape-level drivers, both natural and human-related, known to influence the macroecology of these plants (e.g., Murphy et al., 2019, 2020; Alahuhta et al., 2020 in press). If, as has been asserted, polyploidy can be a response to habitat loss and isolation, then high levels of ploidy may be a strategy of plant species which enhances their survival in human-impacted landscapes (Plue et al., 2018), as well as in natural habitats experiencing strong environmental stress or disturbance pressures (e.g., Chambers et al., 1999; Ulum et al. 2020). Despite the increasing evidence that ploidy is a factor of importance for environmental adaptation, relatively few studies have to date addressed the ecological drivers of change in ploidy state for aquatic plants (Šmarda et al., 2013; Soltis et al., 2016; Segraves, 2017).

Our objectives were to address hypotheses related to four principal questions: (1) Is latitude correlated with aquatic macrophyte ploidy, and in particular is there an opposing latitudinal pattern between haploid/diploid and polyploid species? (2) Can latitudinal ploidy patterns in aquatic macrophytes be alternatively explained by environmental and landscape variables? (3) Are there any signs of species interactions with ploidy patterns (e.g., effects associated with species richness)? (4) Do polyploids occupy larger geographical ranges than haploid/diploid aquatic macrophyte species?

Material and Methods

Global aquatic macrophyte ploidy database

Collation of ploidy information was undertaken from a total of 468 sources, commencing with the database underpinning the study (Rice et al., 2019; who included >800 macrophyte species in their world dataset), and complemented by information derived from Floras, scientific papers, and unpublished compiled databases that review in detail data for chromosome number and ploidy level of each species. A substantial part of the relevant ploidy information was found in local journals, dissertations, and books, as well as in non-English language publications (source data are available: see Data Accessibility section below). Ploidy state was characterized in three classes, following Wani et al. (2018) and Dar et al. (2020):

1. **Haploid/Diploid (D)** , species having solely haploid or diploid (or both) populations;
2. **Polyploid (P)** , species exhibiting various levels and forms of polyploidy; and
3. **Mixed ploidy (D&P)** , species showing "other" ploidy, i.e., with both haploid/ diploid, and polyploid populations (or variants showing, for example, agmatoploidy or dysploidy) occurring in different parts of their range.

In allocating percentage occurrence for species showing each ploidy state per global gridcell and ecozone: (i) we used the definition of "aquatic macrophyte" originally proposed by Chambers et al. (2008) and subsequently extended by Murphy et al. (2019), to set up a global pool of 3496 vascular macrophyte species; (ii) we note that polyploidy is an ongoing process and most polyploids are allopolyploids (Tippery et al., 2018; Levin, 2019), so we did not separate our database into autopolyploid and allopolyploid species; (iii) we also note that complete agmatoploidy, complete symploidy, and polyagmatoploidy are considered a

"misinterpretation of polyploidy" by many authors (e.g., Guerra, 2016); and (iv) there is a general lack of appropriate data for species of tropical regions (Ramsey & Ramsey, 2014).

Spatial, environmental and landscape variables

The 16 predictor variables included in the study were (see Appendix A1):

1. **Spatial** : LAT: mid-gridcell latitude (°absolute); ALT: area of high altitude land per gridcell as % gridcell area >1000 m above sea level, a.s.l.;
2. **Environmental** : ET0 (mm): potential evapotranspiration; AI (ratio between precipitation and ET0 × 10,000): aridity index; TYR (°C): average annual temperature; TMX (°C): maximum temperature of warmest month; TRG (°C): maximum temperature of warmest month – minimum temperature of coldest month; TDRY (°C): average temperature of driest quarter; PCP (mm): annual precipitation; PCPDR (mm): precipitation of driest quarter; PCPS: precipitation seasonality (coefficient of variation of monthly precipitation); CCV (m year^{-1}): historic (Late Quaternary) climate change velocity;
3. **Landscape** : grAH (km^2): area of aquatic habitat present per gridcell; CROP (% agricultural land cover per gridcell);
4. **Biotic** : S_{tot}: total macrophyte species richness; S_{end}: species richness of ecozone-endemic macrophytes, both as number of species per gridcell.

Climate variables were obtained from the Bioclim project source (www.worldclim.org/data/bioclim.html) and downloaded at 30 arc-seconds resolution, except for ET0 and AI, which were obtained from Trabucco and Zomer (2019), who derived them (also 30 arc-seconds resolution) from the Worldclim dataset; CCV spatial dataset was obtained from Sandel et al. (2019); water (grAH) and agricultural (CROP) cover data from the USGS Global Land Cover Characterization (GLCC, https://doi.org/10.5066/F7GB230D), and altitude from the USGS Global Multi-Resolution Terrain Elevation Data 2010 (https://topotools.cr.usgs.gov/GMTED_-viewer/). Mean, median and standard deviation for climate variables per gridcell were obtained using a 10° gridcell shapefile and the Zonal statistics tool in QGIS (version 3.4.9-Madeira). The area occupied by each agricultural land use category (and subsequently total agricultural land cover) per gridcell was obtained using the above-mentioned shapefile and the Zonal Histogram tool in the ArcMap Spatial Analyst toolbox using ArcMap v. 9.3.1, see Murphy et al. (2019, 2020) for further details. All-species richness and ecozone-endemic species richness values were obtained from Murphy et al. (2019).

Mapping of spatial, environmental, landscape and biotic variables

Records were compiled for 238 10 x 10deg latitude x longitude gridcells covering the six world ecozones (Palaearctic, Nearctic, Neotropics, Afrotropics, Orient, and Australasia: see Appendix A7 of Murphy et al. (2019) for full information on gridcell locations and ecozone boundaries) that primarily contain the freshwater and brackish inland waterbodies in which aquatic macrophytes occur, using ESRI(r) ArcMap v. 9.3.1. and following procedures described in detail by Murphy et al. (2019, 2020).

Regression procedures to predict ploidy occurrence

Boosted regression trees (BRTs), which are largely unaffected by the distribution of the data (De'ath, 2007), were used to predict the percentage of occurrence of each of the three ploidy states for macrophyte species present per gridcell, using 16 variables (LAT, ET0, AI, TYR, TMX, TRG, TDRY, PCP, PCPDR, PCPS, CCV, grAH, ALT, CROP, S_{tot} and S_{end}). The tree analysis was conducted following the guidelines of Elith et al. (2008). All regression trees overfit the model and it is necessary to apply some cost-complexity based simplification of the initial tree by iteratively dropping each variable and assessing the loss in predictive power. The predictors that are retained in the simplified model can therefore be thought of as significant predictors. Tree complexity was set at three with a learning rate of 0.0005, and with the bag fraction set at 0.75 and only results from the simplified trees are presented. Partial dependence plots of fitted diversity function versus observed values for variables significantly predicting the response variables were prepared; these seek to present the influence uniquely attributable to a single predictor. To support interpretation of the outcomes two approaches are used, BRT and simple linear regression biplots. In addition, 'standard'

3

regression trees were constructed using the same response and set of predictor variables as in the BRT analysis. These provide a better visualisation of relationship of % ploidy status and the key predictor variables along with key cut off values of these environmental drivers in the data set. BRT analysis was carried out using gbm package (Greenwell et al.; 2018) and standard regression trees were constructed using the rpart package (Therneau et al., 2019). These analyses were performed in R v. 4.0.2 (R Core Team, 2020). Linear regression plots were constructed using Excel.

Geographical distribution of ploidy

One-way analysis of variance (ANOVA) with Tukey's mean-separation test procedure was used to determine the significance of differences in total range distribution ("extent of occupancy of aquatic habitat" ahEOO km^2: Murphy et al., 2019) for each of the three ploidy states. The same approach was used to examine differences in endemic and invasive species ploidy between gridcells and ecozones across world latitude bands. We measured variation in latitudinal range in degabsolute (maximum latitude, minimum latitude, median latitude, and total latitudinal range) for a subset of 256 species from 11 aquatic plant families (Alismataceae, Araceae, Eriocaulaceae, Haloragaceae, Hydrocharitaceae, Isoetaceae, Lentibulariaceae, Nymphaeaceae, Potamogetonaceae, Ranunculaceae, and Typhaceae) for which both latitudinal range and ploidy information were available, see procedures in Murphy et al. (2020). We performed all analyses using R v. 4.0.2 (R Core Team, 2020).

Results

Global macrophyte ploidy

Ploidy state information was collected for 1572 species, representing 45.0% of the known total global pool of vascular aquatic macrophyte species (ranging between 50.7 - 100% of the total number of species present per gridcell) (Appendix A2). The macrophyte ploidy database showed that 48.4% of the species are haploid/diploid (D), 34.6% polyploid (P), and 17.0% had mixed ploidy (D&P) (Fig. 1a). We observed that the incidence of percentage of P in macrophyte assemblages increased with increasing distance from the Equator, the percentage occurrence of D decreased at high latitudes, and D&P species showed an intermediate peak of occurrence around the 50deg latitude band (Fig. 1b; Fig. 2; Appendix 3).

Overall, the Northern hemisphere has more available macrophyte ploidy information than is available in the literature for Southern hemisphere species, and species of non-tropical regions are much better represented than tropical species (Appendix A2). Two mainly tropical ecozones, Afrotropics (AFR) and Neotropics (NEO), had the lowest proportion of species with ploidy information available, reflecting the general lack of karyological effort involving low-latitude macrophyte species (for tropical endemics in particular). The other partly tropical ecozones, Orient (OR) and Australasia (AUS), are not as affected by this since they have a higher proportion of species from temperate areas. However, our findings indicated that there is generally little difference between ecozones in the percentage of occurrence of species from the three ploidy states. All six ecozones have between 40–50% D species present (with the highest value in NEO); between 30–40% for P species; and 20–30% for D&P species (except the NEO ecozone, which shows about 18% D&P species present).

Regarding endemic species, the observed global pattern across latitude bands was similar for both haploid/diplod and polyploid species, with significantly higher numbers of endemic species of both ploidy states in the tropics compared to those at the highest latitudes, >60degabsolute (p<0.05). However, the actual mean numbers of polyploid endemic species per gridcell were only about half those for diploid species (reaching 10–11 polyploid species in the high latitude bands, compared to 18–19 haploid/diploid species in the same latitudes). A subsidiary peak was also seen for both endemic species ploidy states at 40–50degabsolute, significantly higher than for the highest latitude bands but lower than values for the tropical latitudes. For mixed ploidy endemic species both the pattern and numbers were different from those seen for haploid/diplod and polyploid species. In this case there was a single peak of mixed ploidy species occurrence in the 40-60deg latitude band, tailing away towards the Poles and also towards the Equator, while mean numbers of mixed ploidy endemic species per gridcell were also much lower, reaching no more than three – four species per

4

gridcell in the peak latitude band.

BRT results and supporting simple regression biplots

Despite the observed patterns of increase, decline, or unimodal response to latitude for P, D, and D&P respectively (Fig. 1b) this variable was not found to be a significant predictor in any of the BRT analyses (Fig. 3–5; Appendix 3). Instead, for both haploid/diplod and polyploid species, temperature (either annual average or maximum of the warmest month) had strong predictive power as did evapotranspiration. The simplified models, using only the significant predictors for haploidy/diplody and polyploidy, explained 92% and 81% of the variance respectively. The standard regression trees also explained a high proportion of variance (Fig. 3–5) and they indicated that macrophyte species richness had a positive influence in the case of haploidy/diploidy and negative for polyploidy. Species number explained some variation in percentage of haploidy/diploidy at average temperatures above 17 degC. Conversely the percentage of polyploidy was influenced by species number below 15.7 degC with the highest incidence of polyploidy at low temperature and low species number (Fig. 3 and 4).

The models for mixed ploidy were not as strong, though 75% of the variance was explained, by two variables – endemic species richness of macrophytes and average temperature (Fig. 5). The standard regression tree included other variables (evapotranspiration and temperature range) but the association with endemic species richness was marked, with mixed ploidy only occurring once endemic richness in a gridcell fell below a value of 150, with a negative association thereafter, and the percentage of mixed ploidy rising as endemic richness declined.

Climatic and biotic variables were stronger drivers of ploidy state than location in either human-impacted landscapes (CROP) or highly-stressed areas of the planet (CCV and ALT). Three current climate variables (potential evapotranspiration: ET0; temperature range: TRG; annual precipitation: PCP) were strong predictors of ploidy state, though varying in importance between the three models.

Geographical distribution: ploidy state ranges

We observed that polyploid macrophyte species occupy larger geographical ranges than haploid/diploid species at global scale (Fig. 6). Overall, mixed ploidy had the biggest geographical range size ($p<0.001$) on a gridcell basis, followed by polyploids and then haploid/diploid species, which also differ from each other significantly ($p < 0.05$). Regarding the species latitudinal range for the subset of 256 species of 11 aquatic plant families used in this analysis, mixed ploidy plants again showed significantly greater latitudinal range, as well as higher maximum and mean latitudes of occurrence than either haploid/diploid or polyploid species ($p<0.001$) (Fig. 6). Latitudinal range diagrams are shown in Appendix A5 for families with >5 species per ploidy category which have available ploidy state information per ploidy category, in each of 11 macrophyte families for which latitudinal range data were also available (see Murphy et al., 2020).

Discussion

Latitudinal and environmental patterns of ploidy

In plants, abiotic factors such as temperature, light regime (especially photoperiod), drought, or salinity, as well as isolation, may influence the incidence of different ploidy states (e.g., Hagerup, 1932; Comai et al., 2005; te Beest et al., 2012; Husband et al., 2013; Šmarda et al., 2013; Ramsey & Ramsey, 2014; Lobato-de Magalhães et al., 2019). Polyploids often have broader ecological tolerances than their diploid antecedents (Soltis et al., 2016). Notwithstanding the strong spatial latitudinal patterns observed here, we found that individual environmental and biotic variables (though often following latitudinal gradients) are strongly predictive of, or associated with ploidy patterns in aquatic macrophytes. These primarily include temperature (as annual average, maximum of the warmest month, or annual temperature range), potential evapotranspiration, annual precipitation, total macrophyte species richness, and ecozone-endemic macrophyte species richness. Polyploid macrophyte species were massively more prevalent in harsh environments, particularly those with low temperatures and reduced growing season length, typical of sub-Arctic to Arctic latitudes.

Overall, our results provide evidence from macrophytes to support the general assertion that increased chromosome number may be associated with a wide range of environmental conditions, but especially those typifying the environmentally-stressed high latitudes (e.g., drastic changes in temperature and photoperiod between seasons). It should be noted that Cyperaceae (the most diverse family in floras of the tundra and taiga zone), Poaceae, and (in part) Juncaceae contribute strongly to the regional floras of high latitudes, and all are well represented in the macrophyte floras of such regions. In Cyperaceae it has been stated that "chromosome instability is almost the rule" (Tena-Flores et al., 2014). Agmatoploidy, dysploidy and symploidy have been reported in several aquatic and wetland genera of these three families, including *Carex*, *Eleocharis*, *Luzula*, and *Rhynchospora*(Tena-Flores et al., 2013, 2014; Luceño et al., 1998; Guerra, 2016).

Changes in ploidy perhaps cause changes in plant reproduction (e.g., apomixis or clonal recruitment), which can affect the evolution of species' geographic distribution range (Eckert, 2002; Ulum et al. 2020). For example, aquatic mosses, as with their terrestrial cousins, developed various adaptations, such as the alternation of gametophyte (haploid) and sporophyte (diploid) generations, to aid survival in the often environmentally-stressed and highly-disturbed conditions (for example, fast-flowing or torrential upland streams), in which they typically occur (Grime, 1979; Lang & Murphy, 2011; Goga et al. 2018). Polyploids often show a tendency towards clonality, apomixis, and self-compatibility, all of which are adaptations that generally enhance reproductive assurance (Herben et al., 2017).

Recently, Ulum et al. (2020) observed that different cytotypes of*Ranunculus auricomus* L. (Ranunculaceae) (diploid, tetraploid, and hexaploid races, typically occurring in moist woodland conditions) vary their reproduction strategy (e.g., apomixis frequency) in response to extended photoperiod. These authors mention that the polyploid cytotypes of this species may better buffer environmental stress, thereby facilitating the establishment of the species, which is congruent with our findings. A wider range of environmental conditions (e.g., temperature and photoperiod) related to high latitudes could be also be playing a role as a driver of polyploidy in aquatic plants. Middle latitude occurrence of mixed ploidy species could reflect an adaptation by the ancestral diploid races of these species to expand the species range into higher latitudes, via an increase in chromosome number to form new races better suited to the higher-latitude conditions, while also retaining the original diploid cytotypes in their lower-latitude home areas. This could well contribute to an explanation of the observed higher occurrence of mixed ploidy macrophyte species in latitudes intermediate between the peaks of haploidy/diploidy occurrence at low latitudes, and polyploidy at very high latitudes (see schematic summary diagram, Fig. 1b).

Species interactions – species richness and endemism effects upon ploidy state

The set and number of macrophyte species present within the area of a given gridcell integrate are measures of the assemblage which integrate all the various environmental stress and disturbance factors (Grime, 1979) that individually appear to affect ploidy state occurrence. Diversity is one such measure, and our findings suggest that macrophyte species richness (both for total species and endemics) has some influence on, or at least an association with ploidy state in macrophytes. Low interspecific competition (suggested by low number of species present per unit area, though competition is also strongly influenced by other characteristics of species assemblage in plants, especially production: Grime, 1979) probably acts to increase the incidence of polyploidy (Rice et al., 2019). The number of macrophyte species present per unit area of the planet's surface decreases with increasing distance from the Equator, and particularly so at high latitudes (Murphy et al., 2019, 2020), and it is notable that the incidence of polyploidy in macrophyte assemblages increases within this reducing alpha-diversity of macrophytes at high latitudes. However, if we look at the percentage of polyploid species vs. latitude there is no strong change until higher than 40 or even 50°absolute latitudes, i.e. once cool-temperate conditions are reached, and temperature variation at a given latitude is hence also likely to be driving polyploid occurrence. This fits with our findings for latitude-associated climatic variables like temperature and evapotranspiration which are good predictors of ploidy state in our dataset.

As expected, there are differences in the mean ploidy state of gridcells occurring in the six target ecozones, primarily relating to differences in climate between the temperate and more tropical ecozones. These observations on geographical variations in ploidy state occurrence for ecozone-endemic macrophyte species lend

6

support to the evidence that we found regarding the value of endemic species richness in predicting mixed ploidy state using BRT analyses.

The predictive power of the number of endemic taxa in a grid cell on the percentage of mixed ploidy species was remarkable, explaining 49% of the total variation in the percentage D&P. However, it is hard to determine whether endemic species richness has a direct effect on ploidy state or whether endemic species richness is more likely just a proxy summary of multiple factors that actually drive this pattern (see schematic summary diagram, Fig. 1b): notably climatic (annual temperature range, evapotranspiration) and landscape factors (area of water body and anthropic disturbance).

Lastly, under this heading, we need to emphasize the possible effects of information bias, since endemics, especially those of tropical regions, are under-represented in the macrophyte karyological literature (Ramsey & Ramsey, 2014), and we do not know the possible impact of this upon our results relating to macrophyte endemism. Future improvement in knowledge of ploidy state in such plants may help rectify this situation.

Geographical distribution: ploidy state ranges

Confirming the known evidence that plant range distribution is associated with ploidy state (Petit & Thompson, 1999; Lowry & Lester, 2006; Martin & Husband, 2009), our results demonstrated that polyploid species generally occupy larger geographical ranges than diploid macrophyte species. Additionally, we observed higher geographical and latitudinal range size of mixed ploidy species compared with diploid and polyploid species. Plant species that have both types of ploidy (D&P) are likely to be able to occupy a wider range of environmental conditions than species with more limited ploidy. This may occur because their component polyploid populations are likely to be better suited to more stressed high latitudes (or possibly high mountain environments) while their haploid/diploid populations can occupy low-latitude habitats experiencing more benign conditions, such as higher temperatures, thereby leading to broader overall distribution. An example of a macrophyte species showing geographically-demarcated populations of mixed ploidy is *Rotala ramosior* (L.) Koehne (Lythraceae). In this New World species Neotropical populations, from Mexico southwards, are diploid, whilst Nearctic populations, occurring as far north as Canada, are polyploid (Les, 2017). In other cases, different chromosome races of a species may occur in the same area, as, for example, in *Phragmites australis* (Cav.) Trin. ex Steud. (Poaceae), which has more than a dozen chromosome races varying from 2n=28 to 2n=120, most frequent being 2n=48, 72, 96. Other examples are *Ranunculus gmelinii* DC. (Ranunculaceae) with three chromosome races (2n=16 (2x); 2n=32 (4x); 2n=64 (8x)) occurring (with different frequency) throughout its distribution area, and *Comarum palustre* L. (Rosaceae) which has two main races (2n=28 (4x); 2n=42 (6x)) not correlated with either morphology or geography (e.g. Rice et al., 2015). Although there is some information in the literature, such as that for *Rotala ramosior,*on distributions of the different cytotypes of some macrophyte species, very little work has been done hitherto on the detail of geographical cytotype occurrence in macrophyte species, let alone the landscape-scale environmental and other factors which may drive ploidy state distribution in these plants.

Ploidy of invasive aquatic macrophyte species

Change in ploidy is generally thought to be associated with the success of species invasion and invasiveness, as well as rapid environmental and climatic adaptation over varying spatial and time scales (te Beest et al., 2012; Wani et al., 2018; Levin, 2019). Regarding invasiveness, of the 3496 known macrophyte species 52 are seriously invasive into at least one other world ecozone outwith their native ecozone(s) (e.g., Pieterse & Murphy, 1993; Hussner et al., 2017; Hill et al., 2020), and we have ploidy information for 49 of those species (Appendix A5). The proportion of invasive species exhibiting each ploidy state is similar: haploid/diploid (17 species), polyploidy (15), and mixed ploidy (17). There are differences in the incidence of invasive macrophytes with different ploidy between ecozones, but there is little or no consistent pattern to suggest that ploidy is important in determining invasiveness in aquatic macrophyte species. Our findings are supported by those of Kubátová et al. (2008), who looked at ploidy state in native (Old World) and invasive (North American) populations of *Lythrum salicaria* L. (Lythraceae), and found that invasive spread was not associated with differences in ploidy level (2x, 3x, 4x, 6x).

It is also interesting to examine in some detail the results of a recent study of invasive plant ploidy at a local scale, in this case the Kashmir Himalaya region, straddling the boundary between Orient and Palaearctic, and mostly located at high elevation (Dar et al., 2020). The authors provide evidence to support their assertion that polyploidy is strongly associated with invasiveness in the alien flora of this region. However, we doubt that polyploid species are actually predominant in the alien macrophyte flora of Kashmir: mixed ploidy species are more likely to play this role in the region. The reason for the difference between our and their conclusion is that in their study the authors cite only a single reference for ploidy state of each species, often quite old and in many cases geographically very limited, although some of the species are widely distributed across the world. In contrast we used multiple additional sources in allocating ploidy state, often more recent, and usually covering additional areas of the planet – the net consequence being a higher probability of finding races of the target species which show both types of ploidy, should these exist, resulting in more likely allocation of mixed ploidy status to the species. Of 390 alien plant species included in their study 55 (14.1%) are macrophytes, which were categorised by Dar et al. (2020) as 18.2% D, 72.7% P and 9.1% D&P species. However, the outcome using ploidy information for the same species from our dataset suggests a breakdown of 9.1% D, 32.7% P and 58.2% D&P for these species, which does not support the assertion that polyploidy predominates in this set of species invasive in Kashmir (though it remains possible that polyploid races of the species concerned are more invasive than their diploid cytotypes: no data are however available on this issue). This example demonstrates the relevance of global ploidy datasets on seeing the "whole picture", though it will be interesting to see the outcomes of more studies along these lines, at local scale, from other parts of the world.

Conclusions

The total number of species recorded in our ploidy database is probably close to the total for which ploidy has so-far been assessed in macrophyte species. This ploidy database provides a useful resource to support further studies about the effects of varying environmental conditions on ploidy state in aquatic plants, at both regional and local scales, particularly in the context of scenarios of future climate change which suggest increased temperature and drought worldwide.

The findings of our study, utilising this database, improve knowledge of patterns of global distribution and diversity relating to ploidy state within a group of freshwater organisms, namely aquatic macrophytes, which both occur widely across the planet and are inherently adapted to a substantially-varied set of environmental stress conditions worldwide. The study also provides new information about drivers of ploidy at global-scale for this group of plants, including the use of plant assemblage diversity parameters to predict ploidy. In particular, our results provide evidence that climatic factors, especially temperature and evapotranspiration, play a strong role in driving distributions of macrophyte ploidy state, across the planet. Overall, our data strongly support the assertion that increased chromosome number may be an adaptation assisting vascular plant survival in a range of harsh environmental conditions, especially those typifying environmentally-stressed high latitudes.

Acknowledgment

References

Alahuhta, J., Lindholm, M., Båstrup-Spohr, L., Garcia-Giron, J., Toivanen, M., Heino, J., Murphy, K. (2020 in press). Macroecology of macrophytes in the freshwater realm: Patterns, mechanisms and implications. *Aquatic Botany*

Beck, M.W. & Alahuhta, J. (2017). Ecological determinants of *Potamogeton* taxa in glacial lakes: assemblage composition, species richness, and species-level approach. *Aquatic Sciences 79* : 427-441. https://doi.org/10.1007/s00027-016-0508-x

Blackburn, K. B., & Morton, J. K. (1957). The incidence of polyploidy in the Caryophyllaceae of Britain and of Portugal. *New Phytologist 56* , 344-351. https://www.jstor.org/stable/2429613

Bornette, G. & Puijalon, S. (2011). Response of aquatic plants to abiotic factors: A review. *Aquatic Sciences 73* , 1–14. https://doi.org/10.1007/s00027-010-0162-7

Chambers, R., Meyerson, L. & Saltonstall, K. (1999). Expansion of *Phragmites australis* into tidal wetlands of North America. *Aquatic Botany 64,* 261-273. https://doi.org/10.1016/S0304-3770(99)00055-8

Chambers, P.A., Lacoul, P., Murphy, K.J. & Thomaz, S.M. (2008). Global diversity of aquatic macrophytes in freshwater. *Hydrobiologia 595* , 9-26.

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature Reviews Genetics 6* , 836–846. https://doi.org/10.1038/nrg1711

Dar, M.A., Wani, G. A., Reshi, Z.A., Al-Qarawi, A.A., Abd Allah, E.F. & Shah, M.A. (2020). Stage-specific ploidy level variations in invasive species in comparison to rare endemics in Kashmir Himalaya. *Flora 262* , 151525.

De'ath, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology* , *88(1)* , 243–251. https://doi.org/10.1890/0012-9658(2007)88[243:BTFEMA]2.0.CO;2

Eckert C.G. (2002) The loss of sex in clonal plants. In: Stuefer J.F., Erschbamer B., Huber H., Suzuki JI. (eds) Ecology and Evolutionary Biology of Clonal Plants. Springer, \soutDordrecht. https://doi.org/10.1007/978-94-017-1345-0_15

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology* , *77* , 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Garbey, C., Murphy, K.J., Thiébaut, G. & Muller, S. (2004). Variation in P-content in aquatic plant tissues offers an efficient tool for determining plant growth strategies along a resource gradient. *Freshwater Biology 49* , 346-356. https://doi.org/10.1111/j.1365-2427.2004.01188.x

Goga, M., Ručová, D., Kolarčik, V., Sabovljević, M., Bačkor, M., & Lang, I. (2018). Usnic acid, as a biotic factor, changes the ploidy level in mosses. *Ecology and Evolution 8(5)* , 2781–2787. https://doi.org/10.1002/ece3.3908

Greenwell, B., Boehmke, B., Cunningham, J., & GBM, D. (2018). gbm: generalized boosted regression models. R package version 2.1.5.

Grime, J. P. (1979). *Plant strategies and vegetation processes* . Chichester: Wiley.

Guerra, M. (2016). Agmatoploidy and symploidy: a critical review. *Genetics and Molecular Biology 39(4),* 492-496 (2016). https://doi.org/10.1590/1678-4685-GMB-2016-0103

Hagerup, O. (1932). Über polyploidie in beziehung zu klima, ökologie und phylogenie. *Hereditas 16* , 19–40. Doi:10.1111/j.1601-5223.1932.tb02560.x

Hardy, O.J., Vanderhoeven, S., De Loose, M. & Meerts, P. (2000). Ecological, morphological and allozymic differentiation between diploid and tetraploid knapweeds (*Centaurea jacea*) from a contact zone in the Belgian Ardennes. *New Phytologist 146* , 281–290. https://doi.org/10.1046/j.1469-8137.2000.00631.x

Herben, T., Suda, J. & Klimesova, J. (2017). Polyploid species rely on vegetative reproduction more than diploids, a re-examination of the old hypothesis. *Annals of Botany, 120 (2)* , 341-349. https://doi.org/10.1093/aob/mcx009.

Hill, M.P., Coetzee, J.A., Martin, G.D., Smith, R. & Strange, E.F. (2020). Invasive alien aquatic plants in South African freshwater ecosystems. In: van Wilgen B., Measey J., Richardson *Biological Invasions in South Africa.* Invading Nature D., Wilson J., Zengeya T. (eds) - Springer Series in Invasion Ecology, vol 14. Springer, Cham. https://doi.org/10.1007/978-3-030-32394-3_4

9

Husband, B.C., Baldwin, S.J. & Suda, J. (2013). The incidence of polyploidy in natural plant populations: major patterns and evolutionary processes. In: Greilhuber, J., Dolezel, J., & Wendel, J. (eds.)*Plant genome diversity 2,* 255-276, Springer. http://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-3-7091-1160-4_16

Hussner, A., Stiers, I., Verhofstad, M.J.J.M. et al. (2017). Management and control methods of invasive alien freshwater aquatic plants: a review. *Aquatic Botany 136* ,112-137. https://doi.org/10.1016/j.aquabot.2016.08.002

Iversen, L. L., Winkel, A., Baastrup-Spohr, L., Hinke, A. B., Alahuhta, J., Baattrup-Pedersen, A., Birk, S., Brodersen, P., Chambers, P A., Ecke, F., Feldmann, T., Gebler, D., Heino, J., Jespersen, T S., Moe, S J., Riis, T., Sass, L., Vestergaard, O., Maberly, S C., Sand-Jensen, K., Pedersen, O. (2019). Catchment properties and the photosynthetic trait composition of freshwater plant communities. *Science 366* , 878-881. https://doi.org/10.1126/science.aay5945

Johnson, A.W. & Packer, J.G. (1965). Polyploidy and environment in Arctic Alaska. *Science 148* , 237–239. https://doi.org/10.1126/science.148.3667.237

Johnson, M.T.J., Husband, B.C. & Burton, T.L. (2003). Habitat differentiation between diploid and tetraploid *Galax urceolata*(Diapensiaceae). *International Journal of Plant Science* 164, 703–710. https://doi.org/10.1086/376813

Kron, P., Suda, J. & Husband, B.C. (2007). Applications of flow cytometry to evolutionary and population biology. *Annual Review of Ecology, Evolution and Systematics 38* , 847–876. https://doi.org/10.1146/annurev.ecolsys.38.091206.095504

Kubátová, B., Trávníček, P., Bastlová, D., Curn, V., Jarolímová, V. & Suda, J. (2008). DNA ploidy-level variation in native and invasive populations of *Lythrum salicaria* at a large geographical scale.*Journal of Biogeography 35* , 167–176. https://doi.org/10.1111/j.1365-2699.2007.01781.x

Lang, P. & Murphy, K.J. (2012) Environmental drivers, life strategies and bioindicator capacity of aquatic bryophyte communities in high-latitude upland streams. *Hydrobiologia 679* , 1-17. https://doi.org/10.1007/s10750-011-0838-6

Les, D.H. (2017). *Aquatic dicotyledons of North America: Ecology, life history, and systematics* . CRC Press, Taylor & Francis. https://doi.org/10.1201/9781315118116

Levin, D.A. (2019). Plant speciation in the age of climate change.*Annals of Botany, 124* , 769-775. https://doi.org/10.1093/aob/mcz108

Lobato-de Magalhaes, T., Cabrera-Toledo, D. &*Mahinda Martinez, M. (2019). Microsatellite loci transferability and genetic diversity of the aquatic plant *Nymphoides fallax* Ornduff (Menyanthaceae), endemic to the Mexican and Guatemalan highlands. *Limnology 20* , 233–241. https://doi.org/10.1007/s10201-019-00571

Love, A & Love, D. (1949). The geobotanical significance of polyploidy. I. Polyploidy and latitude. *Portugaliae Acta Biologica A (R. B. Goldschmidt Jubilee Volume),* 273–352.

Lowry, E. & Lester, S.E. (2006). The biogeography of plant reproduction: potential determinants of species'range sizes.*Journal of Biogeography 33* , 1975–1982. https://doi.org/10.1111/j.1365-2699.2006.01562.x

Luceno, M., Mendes, A.P., Venzela, A.L.L & Alves M.V. (1998). Agmatoploidy and symploidy in *Rhynchospora cephalotes* (L.) Vahl (Cyperaceae). *Cytologia 63* , 79-81.

Lui, K., Thompson, F.L. & Eckert, C.G. (2005). Causes and consequences of extreme variation in reproductive strategy and vegetative growth among invasive populations of a clonal aquatic plant, *Butomus umbellatus* L. (Butomaceae). *Biological Invasions 7* , 427–444. https://doi.org/10.1007/s10530-004-4063-3

Martin, S.L. & Husband, B.C. (2009). Influence of phylogeny and ploidy on species ranges of North American angiosperms. *Journal of Ecology 97* , 913–922. https://doi.org/10.1111/j.1365-2745.2009.01543.x

Martin, S.L. & Husband, B.C. (2013). Adaptation of diploid and tetraploid *Chamerion angustifolium* to elevation but not local environment. *Evolution 67* , 1780–1791. https://doi.org/10.1111/evo.12065

Murphy, K., Efremov, A., Davidson, T.A., Molina-Navarro, E., Fidanza, K., Crivelari Betiol, T.C., Chambers, P., Tapia-Grimaldo, J., Varandas Martins, S., Springuel, I., Kennedy, M., Mormul, R., Dibble, E., Hofstra, D., Lukacs, B.A., Gebler, D., Bastrup-Spohr, L. & Urrutia Estrada, J. (2019). World distribution, diversity and endemism of aquatic macrophytes. *Aquatic Botany 158* . https://doi.org/10.1016/j.aquabot.2019.06.006

Murphy, K., Carvalho, P., Efremov, A., Tapia Grimaldo, J., Molina-Navarro, E., Davidson, T.A. & Thomaz, S.M. (2020). Latitudinal variation in global range-size of aquatic macrophyte species shows evidence for a Rapoport Effect. *Freshwater Biology 65* , 1622 - 1640. https://doi.org/10.1111/fwb.13528

Petit, C. & Thompson, J.D. (1999). Species diversity and ecological range in relation to ploidy level in the flora of the Pyrenees.*Evolution and Ecology 13* , 45–66. https://doi.org/10.1023/A:1006534130327

Pieterse A.H. & Murphy K.J. (1993). *Aquatic Weeds: 2nd Edition* . Oxford University Press, Oxford, UK.

Plue, J., Kimberley, A. & Slotte, T. (2018). Interspecific variation in ploidy as a key plant trait outlining local extinction risks and community patterns in fragmented landscapes. *Functional Ecology32(8)* , 2095–2106. https://doi.org/10.1111/1365-2435.13127

R Core Team (2020). R*: a Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. R-project.org/

Ramsey, J. & Ramsey, T.S. (2014). Ecological studies of polyploidy in the 100 years following its discovery. *Philosophical Transactions of the Royal Society B 369* : 20130352. https://doi.org/10.1098/rstb.2013.0352

Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N. M., Salman-Minkov, A., . . . & Mayrose, I. (2015). The Chromosome Counts Database (CCDB)–a community resource of plant chromosome numbers.*New Phytologist* , *206(1)* , 19-26. https://doi.org/10.1111/nph.13191

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. https://doi.org/10.1038/s41559-018-0787-9

Sandel, B., Arge, L., Dalsgaard, B., Davies, R. G., Gaston, K. J., Sutherland, W. J., & Svenning, J. C. (2011). The influence of Late Quaternary climate-change velocity on species endemism. *Science* , *334(6056)* , 660-664. https://doi.org/10.5061/dryad.b13j1

Santamaría, L. (2002). Why are most aquatic plants widely distributed? Dispersal, clonal growth and small-scale heterogeneity in a stressful environment. *Acta Oecologica 23* , 137–154. https://doi.org/10.1016/S1146-609X(02)01146-3

Segraves, K. A. (2017). The effects of genome duplications in a community context. *New Phytologist 215* , 57-69. https://doi.org/10.1111/nph.14564

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. https://doi.org/10.1111/nph.12399

Soltis, D.E., Visger, C.J., Marchant, D.B. & Soltis, P.S. (2016). Polyploidy: pitfalls and paths to a paradigm. *American Journal of Botany 103* , 1146–1166. https://doi.org/10.3732/ajb.1500501

Stebbins, G. L. (1984). Polyploidy and the distribution of the Arctic-Alpine flora - new evidence and a new approach. *Botanica Helvetica 94* , 1–13.

Sun, J., Hunter, P.D., Tyler, A.N., & Willby, N.J. (2019). Lake and catchment-scale determinants of aquatic vegetation across almost 1,000 lakes and the contrasts between lake types. *Journal of Biogeography 46* , 1066–1082. https://doi.org/10.1111/jbi.13557

te Beest, M., Le Roux, J.J., Richardson, D.M., Brysting, A.K., Suda, J., Kubesova, M. & Pysek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. *Annals of Botany109* , 19–45. https://doi.org/10.1093/aob/mcr277

Tena-Flores, J.A., González-Elizondo, M.S., Herrera-Arrieta, Y., Almaraz-Abarca, N., Mayek-Pérez, N., da Silva, C.R.N. & Vanzela, A.L.L. (2013). Karyotype characterization of eight Mexican species of*Eleocharis* (Cyperaceae). *Botanical Sciences 91* , 119-128. http://rdcb.cbg.ipn.mx/handle/20.500.12273/66

Tena-Flores, J.A., González-Elizondo, M.S., Herrera-Arrieta, Y., Almaraz-Abarca, N., Mayek-Pérez, N. & Vanzela, A.L.L. (2014). Karyotype characterization of four Mexican species of *Schoenoplectus*(Cyperaceae) and first report of polyploid mixoploidy for the family.*Caryologia* , *67(2)* , 124-134. https://doi.org/10.1080/00087114.2014.931633

Therneau, T., Atkinson, B., & Ripley, B. (2019). rpart: Recursive Partitioning and Regression Trees. R package version 4.1–15.

Tippery, N.P., Sears, N.L., Zentner, A.B. & Sivadas, V. (2018). Evidence for allopolyploid speciation in *Nymphoides*(Menyanthaceae). *Systematic Botany 43* , 117–129. https://doi.org/10.1600/036364418X696950

Tischler, G. (1935). Die Bedeutung der Polyploide für die Verbreitung der Angiospermen. *Biologische Jahrbuch* 47, 1.

Trabucco, A., & Zomer, R. J. (2018). Global aridity index and potential evapotranspiration (ET0) climate database v2. *CGIAR Consort Spat Inf* . https://doi.org/10.6084/m9.figshare.7504448.v3

Ulum, F. B., Costa Castro, C., & Hörandl, E. (2020). Ploidy-dependent effects of light stress on the mode of reproduction in the*Ranunculus auricomus* complex (Ranunculaceae). *Frontiers in Plant Science* , *11* , 104. https://doi.org/10.3389/fpls.2020.00104

Wani, G. A., Shah, M. A., Reshi, Z. A., & Dar, M. A. (2018). Polyploidy determines the stage of invasion: clues from Kashmir Himalayan aquatic flora. *Acta Physiologiae Plantarum* , *40(3)*.https://doi.org/10.1007/s11738-018-2629-4

Figures

Figure 1 . Ploidy state of 1572 aquatic macrophyte species at a global-scale: A. Venn diagram of number of species per ploidy category; B. Schematic summary diagram of latitudinal patterns of ploidy, and climatic and biotic variables. (D = haploid/diploid, P = polyploid, D&P = mixed ploidy, S_{end} = species richness of ecozone-endemic macrophytes; S_{tot} = all-species richness; see Methods text for full explanation other abbreviations)
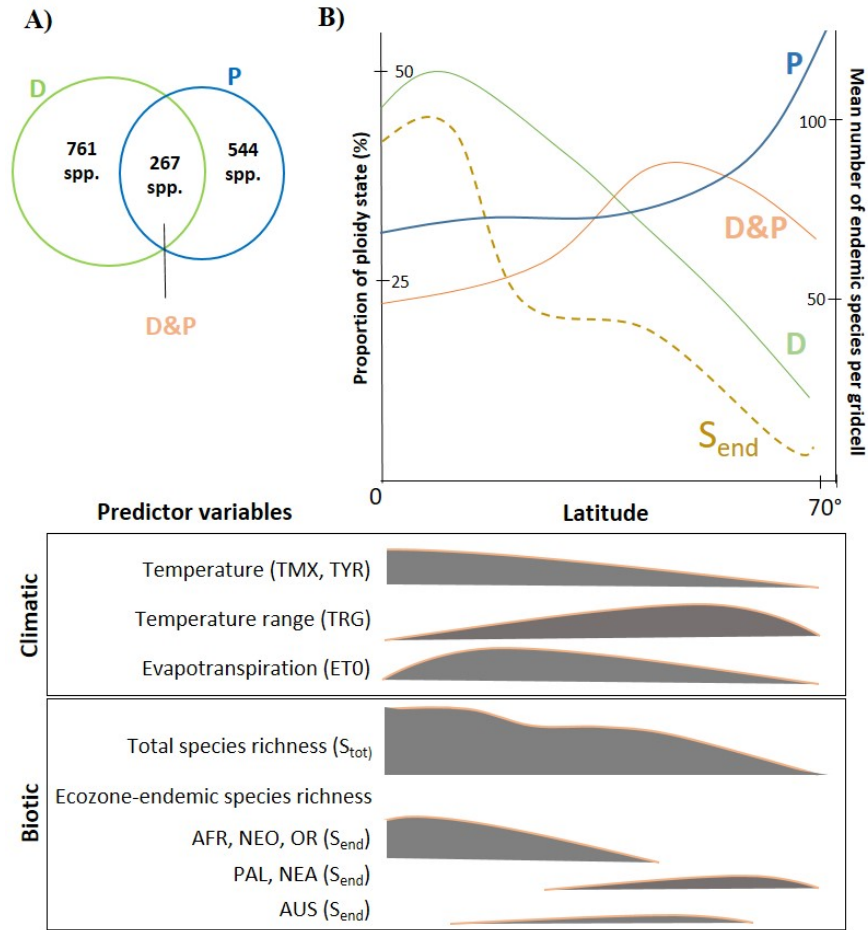
Figure 2 . Global map of aquatic macrophyte species ploidy state (a) haploid/diploid, (b) polyploid, (c) mixed ploidy.
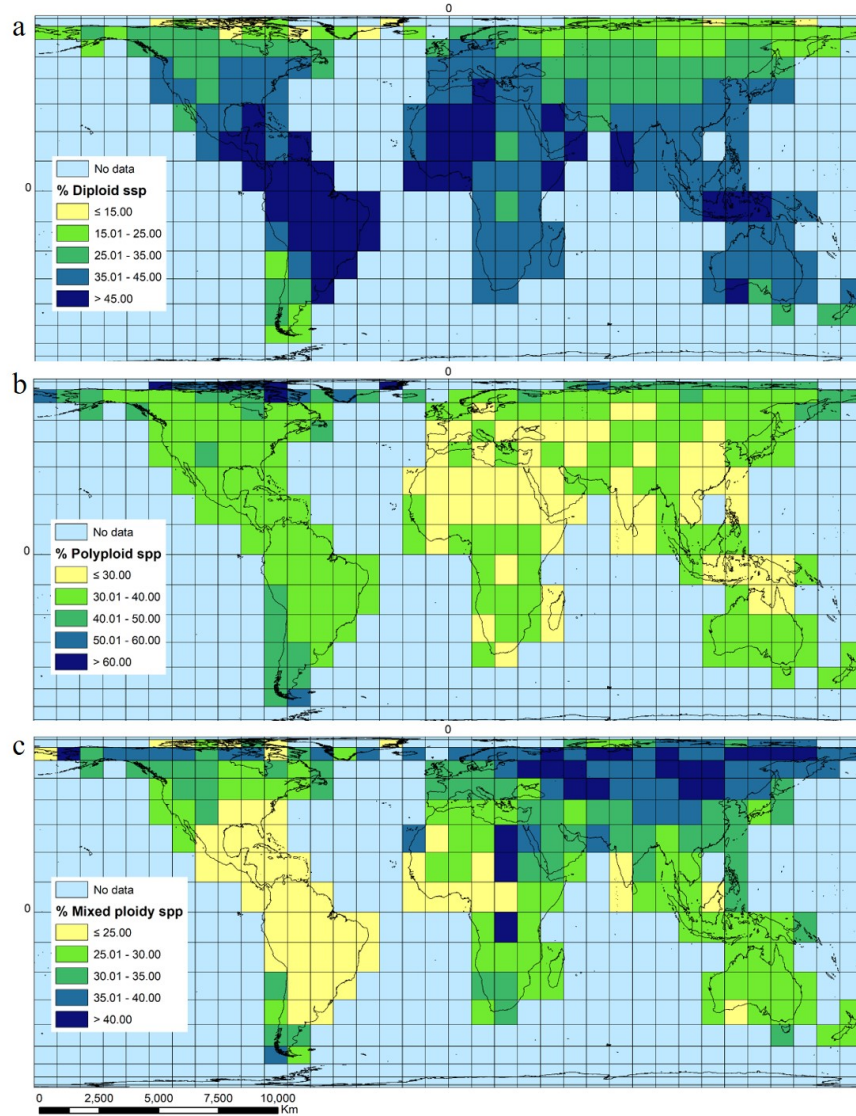
Figure 3 . a) Percentage of haploidy/diploidy versus absolute latitude; b) Pruned regression tree results based on percentage of haploidy/diploidy, the total deviance explained is variance remaining in the sum of the leaves as a proportion of the variance in the entire data set. Cut off values for the most important environmental variable are given at each split, average annual temperature (TYR), potential evapotranspiration (ET0) and number of endemic species (End). The numbers at each leaf is the average percentage of haploidy/diploidy of the grid square in that group, c) The partial dependency plots showing the shape of the relationship between percentage of haploidy/diploidy and its best predictor variables; average annual temperature (TYR), potential evapotranspiration (ET0), total species richness (S_{tot}) and maximum temperature of the warmest month (TMX).
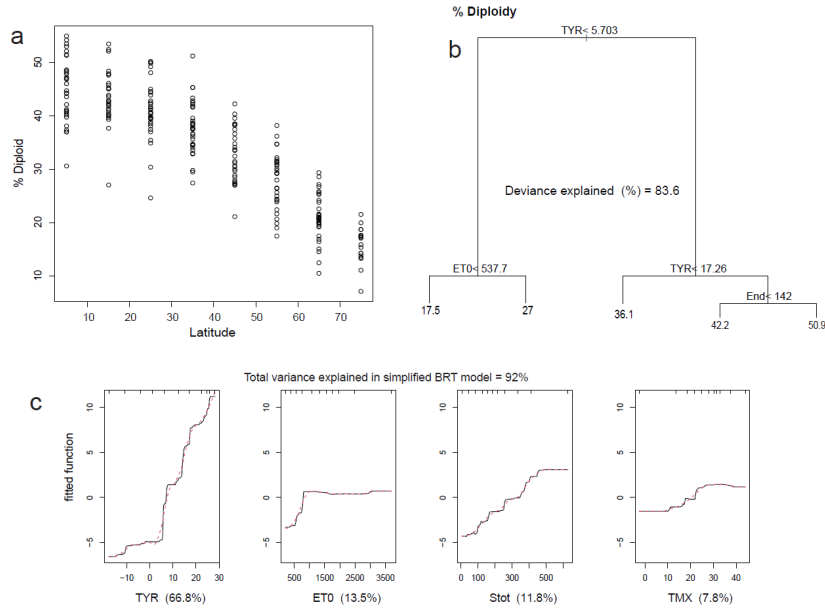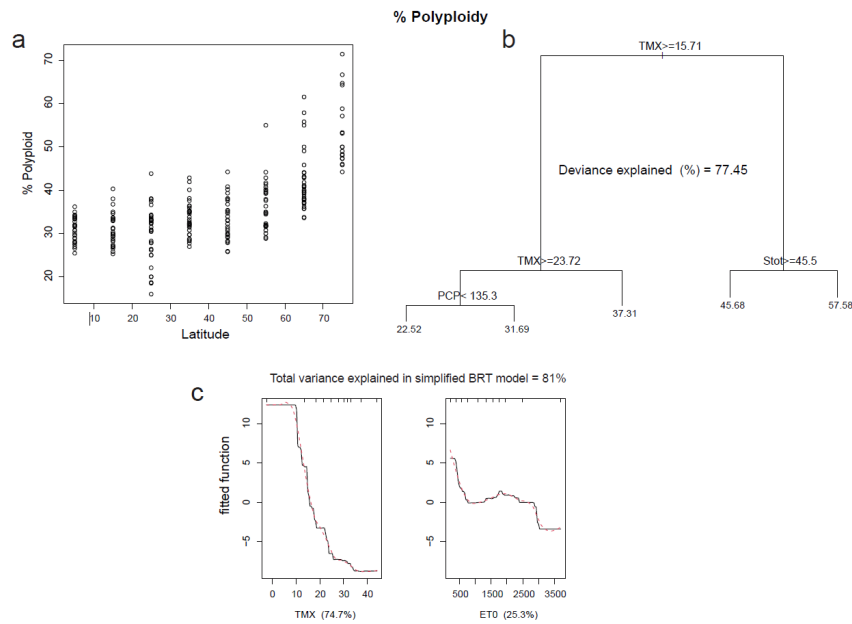
14

Figure 4 . a) Percentage of polyploidy versus absolute latitude; b) Pruned regression tree results based on percentage of polyploidy, the total deviance explained is variance remaining in the sum of the leaves as a proportion of the variance in the entire data set. Cut off values for the most important environmental variable are given at each split: maximum temperature of the warmest month (TMX), Total species richness (S_{tot}) and annual precipitation (PCP). The numbers at each leaf is the average percentage of polyploidy of the grid square in that group, c) The partial dependency plots showing the shape of the relationship between percentage of polyploidy and its best predictor variables: Maximum temperature of the warmest month (TMX) and potential evapotranspiration (ET0).
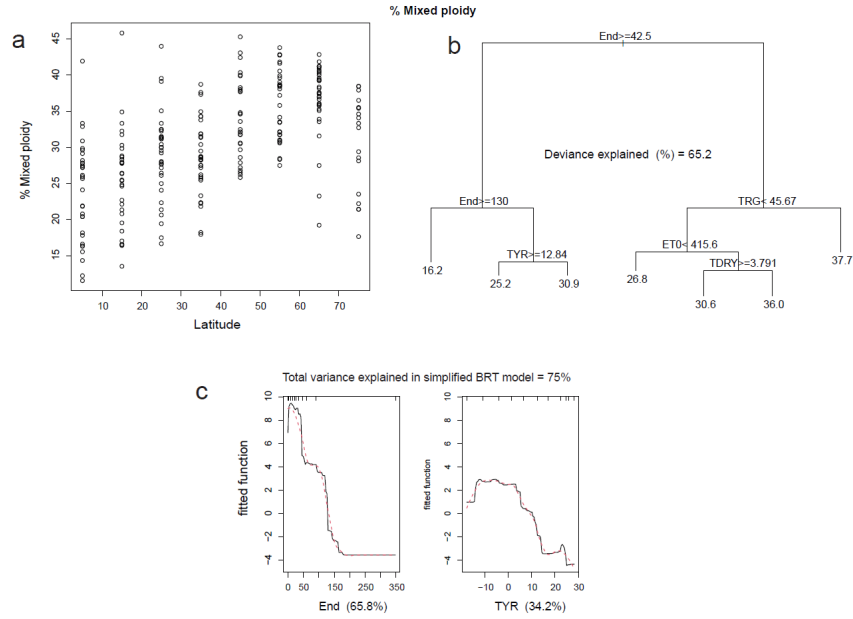
Figure 5 . Percentage of mixed ploidy versus absolute latitude; b) Pruned regression tree results based on percentage of mixed ploidy, the total deviance explained is variance remaining in the sum of the leaves as a proportion of the variance in the entire data set. Cut off values for the most important environmental variable are given at each split: Endemic species richness (End) maximum temperature of the warmest month (TMX), Maximum of the warmest month – minimum of the coldest month (TRG), average annual temperature (TYR), potential evapotranspiration (ET0), average temperature of the direst quarter (TDRY). The numbers at each leaf is the average percentage of mixed ploidy of the grid square in that group, c) The partial dependency plots showing the shape of the relationship between percentage of mixed ploidy and its best predictor variables: Endemic species richness (End) and average annual temperature (TYR).

Figure 6 . Geographical range size per ploidy states (D = haploid/diploid, P = polyploid, D&P = mixed ploidy) (p<0.01).
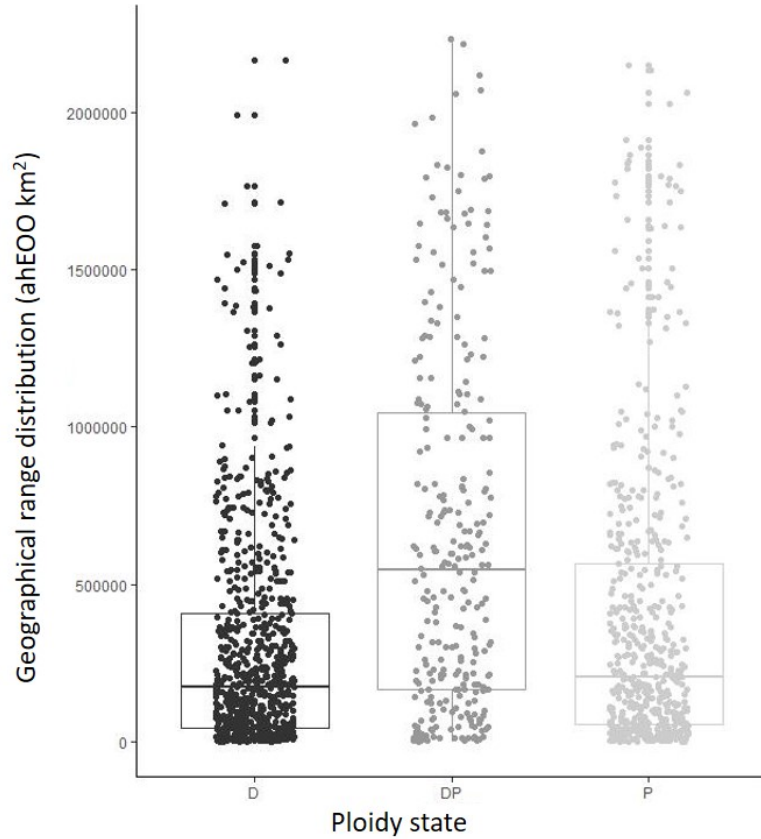
All species -- Geographical range distribution: D&P > P > D
Subset of 11 families:
Latitudinal range: D&P > P = D
Maximum latitude: D&P > P = D
Mean latitude: D&P > P = D
Minimum latitude: D&P = P = D

Supplementary material

A1. Maps of environmental and landscape variables.

A2. Map of ploidy information available per gridcell.

A3. Simple regression biplots for each of the main explanatory variables.

A4. Latitudinal range diagrams for 11 families.

A5. List of invasive species and their ploidy state.

Hosted file

A1. Environmental and landscape var maps.pdf available at https://authorea.com/users/378432/articles/494933-global-scale-drivers-of-ploidy-state-in-aquatic-macrophytes

Hosted file

A2. `Ploidy map.pdf` available at https://authorea.com/users/378432/articles/494933-global-scale-drivers-of-ploidy-state-in-aquatic-macrophytes

Hosted file

A3. `Regression biplots.pdf` available at https://authorea.com/users/378432/articles/494933-global-scale-drivers-of-ploidy-state-in-aquatic-macrophytes

Hosted file

A4. `Lat range diags.pdf` available at https://authorea.com/users/378432/articles/494933-global-scale-drivers-of-ploidy-state-in-aquatic-macrophytes

Hosted file

A5. `Invasive species ploidy.pdf` available at https://authorea.com/users/378432/articles/494933-global-scale-drivers-of-ploidy-state-in-aquatic-macrophytes