

**CHARACTERISTICS OF THE NATURALIZED FLORA OF SOUTHERN AFRICA
LARGELY REFLECT THE NON-RANDOM INTRODUCTION OF ALIEN SPECIES
FOR CULTIVATION**

Ali Omer^{1,2}, Trevor Fristoe¹, Qiang Yang¹, Noëlie Maurel¹, Patrick Weigelt³, Holger Kreft^{3,4},
Jonas Bleilevens¹, Wayne Dawson⁵, Franz Essl⁶, Jan Pergl⁷, Petr Pyšek^{7,8}, & Mark van
Kleunen^{1,9}

¹Ecology, Department of Biology, University of Konstanz, Universitätsstraße 10, D-78464
Konstanz, Germany

²Department of Forest Management, Faculty of Forestry, University of Khartoum, North
Khartoum, 13314, Sudan

³Biodiversity, Macroecology & Biogeography, University of Goettingen, Büsgenweg 1, D-
37077 Göttingen, Germany

⁴Centre of Biodiversity and Sustainable Land Use (CBL), University of Goettingen,
Büsgenweg 1, D-37077 Göttingen, Germany

⁵Department of Biosciences, Durham University, South Road, Durham DH1 3LE, United
Kingdom

⁶Division of Conservation Biology, Vegetation- and Landscape Ecology, Department of
Botany and Biodiversity Research, University of Vienna, Rennweg 14, Vienna 1030, Austria

⁷Czech Academy of Sciences, Institute of Botany, Department of Invasion Ecology, CZ-252
43 Průhonice, Czech Republic

22 ⁸Department of Ecology, Faculty of Science, Charles University, Viničná 7, CZ-128 44 Prague,
23 Czech Republic

24 ⁹Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation,
25 Taizhou University, Taizhou 318000, China

26

27 Correspondence: ali.omer@uni-konstanz.de

28

ABSTRACT

Most studies on biological invasions focus on the later stages of the invasion process, i.e. after species have already become naturalized. It is frequently overlooked, however, that patterns in origin, phylogeny and traits of naturalized alien species might largely reflect which species have been introduced in the first place. Here, we quantify and account for such introduction biases by analyzing 5,317 plant species introduced for cultivation in Southern Africa. We show that this cultivated alien flora represents a non-random subset of the global flora, and that this bias at the introduction stage largely drives patterns in origin, growth form and phylogenetic composition of the naturalized flora. For example, while species from Australasia are, compared to the global flora, disproportionally overrepresented in the naturalized cultivated flora of Southern Africa, this pattern is solely driven by their higher likelihood of having been introduced for cultivation. We also show that among cultivated aliens, naturalization success was correlated with intermediate seed mass and height, as well as high specific leaf area. Our quantification of introduction biases demonstrates that they are huge, and that accounting for it is essential to avoid over- or under-estimation of the characteristics of successfully naturalized alien plants.

Keywords

biological traits, garden plants, geographical origin, invasive plants, non-native plants, naturalization success, naturalization extent.

INTRODUCTION

Anthropogenic activities are increasingly changing abiotic and biotic components of the biosphere (Lewis & Maslin 2015). The exchange of biota among different parts of the globe is one of the most distinct of these man-made changes. As a result, over 13,000 vascular plant species have now invaded regions outside their known native geographic ranges (van Kleunen *et al.* 2015a). While the invasion process consists of four stages —transport, introduction, establishment (i.e. naturalization) and spread (i.e. advancing to become invasive) (Richardson *et al.* 2000; Blackburn *et al.* 2011)— most research focuses on the last two stages (but see Hulme *et al.* 2008; Faulkner *et al.* 2020). As a consequence, still little is known about what characterizes the introduced species.

As naturalized species (and ultimately invasive species) emerge from the pool of introduced species, it is important to know whether species with specific characteristics are more likely to have been introduced than what is expected by chance (i.e. whether there is a so-called introduction bias) (Chrobock *et al.* 2011; Maurel *et al.* 2016). For example, it has been shown that species from some origins (e.g. Europe) and certain plant lineages (e.g. species in the Pinaceae family) have been more successful at becoming naturalized than others. However, with a few exceptions (McGregor *et al.* 2012; Moodley *et al.* 2013), it is not known whether species from certain continents and families are over-represented among naturalized species because they were more likely to be transported and introduced elsewhere (i.e. they show an introduction bias) or whether they have a higher inherent ability to naturalize once introduced. Accounting for biases resulting from the characteristics of species that have been introduced is therefore an essential, but rarely considered, step for understanding geographic and phylogenetic patterns in naturalization success.

Investigations into plant characteristics associated with invasion success started with Herbert Baker's publication of a list of characteristics of the "ideal weed" (Baker 1965). Since then many studies have searched for functional traits (e.g. growth-form, height, seed mass, specific leaf area) or geographic, phylogenetic and ecological characteristics that promote invasion success (Hamilton *et al.* 2005; Pyšek & Richardson 2007; Bucharova & van Kleunen 2009; Pyšek *et al.* 2009; Ordonez *et al.* 2010; Kueffer *et al.* 2013; van Kleunen *et al.* 2015b; Haeuser *et al.* 2018). While some traits appear to be frequently associated with naturalized or invasive species, literature reviews have also revealed a lack of predictive power and inconsistent associations between plant traits and invasion success (Pyšek & Richardson 2007; van Kleunen *et al.* 2015b; van Kleunen *et al.* 2018a). This could be because the effects of traits on invasion success are context dependent (i.e. vary among regions or environments) (Kueffer *et al.* 2013) or vary among invasion stages (Dawson *et al.* 2009; Moodley *et al.* 2013). Furthermore, most studies considered only linear relationships between invasion success and traits (but see (Küster *et al.* 2008; Haeuser *et al.* 2018), although many relationships may be non-linear (van Kleunen *et al.* 2015b). So, to improve our understanding of the role of plant characteristics in naturalization success, we need to quantify and account for biases associated with the introduction stage (Maurel *et al.* 2016) and to test for non-linear effects (van Kleunen *et al.* 2018a).

Many alien organisms have been unintentionally introduced (Hulme *et al.* 2008), and comprehensive data for such introductions are lacking. Hence it is not possible to know all species that have been introduced to a region. However, in the case of vascular plants, the majority of alien species have been intentionally introduced for cultivation (Reichard & White 2001; Lambdon *et al.* 2008; Dodd *et al.* 2015; Faulkner *et al.* 2016). At least 75% of the ~13,000 naturalized alien plants worldwide are known to be grown in domestic gardens somewhere around the globe (van Kleunen *et al.* 2018b), and plants with economic use are 18

times more likely to naturalize than those not used in such a way (van Kleunen *et al.* 2020). So, introduced cultivated floras offer valuable, yet underutilized, data sources for quantifying biases in the types of plants humans have selected from the global flora for introduction into cultivation, and the consequences of such introduction biases for patterns in the origins, phylogenetic composition and traits of the plants that naturalized (Maurel *et al.* 2016; Pergl *et al.* 2016). Without such information, we might draw incorrect conclusions about what drives naturalization success.

Here, we used a unique checklist of the cultivated flora of Southern Africa (Glen 2002), to test whether these species represent a more biased subset of the global flora, outside of Southern Africa, than what would be expected by chance. This was done in terms of the geographical origins, phylogenetic composition and traits of the cultivated alien species. In addition, we tested whether those that have become naturalized deviate from the ones that were brought into cultivation and did not escape and naturalize. Specifically, we asked the following questions: (1) Are there biases in the continent of origin and the native range size of the introduced cultivated flora of Southern Africa, and if so, can they explain origin and native-range-size patterns in the naturalized flora? (2) Are there phylogenetic and growth-form biases in the cultivated flora of Southern Africa, and if so, can they explain the phylogenetic and growth-form composition of the naturalized flora? (3) How do plant functional traits (i.e. seed mass, height and specific leaf area) mediate the naturalization success of cultivated plants in Southern Africa?

118 **METHODS**

119 **The cultivated and naturalized floras of Southern Africa**

120 A list of taxa (including species and infraspecific taxa; hereafter jointly referred to as ‘species’)
121 that are known to be cultivated in at least one of the 10 countries in Southern Africa (Fig. 1a),
122 was extracted from the ‘Cultivated Plants of Southern Africa’ (Glen 2002). This checklist
123 contains over 8,000 plant names, with information on species’ geographic origins. Although
124 Glen (2002) does not claim to be exhaustive, it is the most comprehensive list of cultivated
125 plants available for this region.

126 To align the list of cultivated plants with other datasets used in this study, we
127 harmonized the taxonomic names according to The Plant List (version 1.1;
128 <http://www.theplantlist.org>) using the R package ‘Taxonstand’ (Cayuela *et al.* 2019). We
129 removed all non-seed plants (i.e. ferns and their allies, 114 species) from the final list, because
130 data on their naturalization success are less complete. The final list included 6,912 cultivated
131 species of which 1,595 are native to Southern Africa and 5,317 are alien. As no species was
132 introduced from Antarctica, we did not consider this continent in our analyses.

133 To identify which of the cultivated plants are currently naturalized in Southern Africa,
134 we used the Global Naturalized Alien Flora (GloNAF) database (van Kleunen *et al.* 2019;
135 accessed February 2020). GloNAF includes data on more than 13,000 vascular plant taxa and
136 their naturalization status in over 1,000 regions around the globe (van Kleunen *et al.* 2019).
137 From GloNAF, we extracted the naturalized species of all 10 Southern African countries (see
138 Fig 1a).

139 **Data on geographic origin and species characteristics**

140 For the cultivated alien flora, Gen (2002) provides already some information on the regions of
141 origin. However, as this information may not be complete, we collected additional native range

data from the Germplasm Resources Information Network (GRIN; <https://ars-grin.gov>), the World Checklist of Selected Plant Families (WCSP; <http://apps.kew.org/wcsp>) and the Plant of the World Online database (POWO 2019); <http://www.plantsoftheworldonline.org/>). As continents of origin, we used the nine level-1 regions of the World Geographical Scheme for Recording Plant Distributions of the Taxonomic Databases Working Group (TDWG) (Brummitt 2001) (Fig 1a). As an estimate of native range size, we used the number of TDWG level-2 regions (total number is 52) each introduced cultivated species is native to (see also (Maurel *et al.* 2016; Razanajatovo *et al.* 2016). This information was available for 5,067 of the 5,317 introduced cultivated species in our dataset.

To test whether the introduced and naturalized cultivated floras of Southern Africa are biased with regard to growth-form (i.e. whether some growth-forms are more or less frequent than expected by chance), we collected growth-form data from multiple sources (Supplementary data S1). As these sources use different growth-form schemes, we harmonized them to seven standard categories: short-lived (i.e. annual or biennial) free-standing herb, long-lived free-standing herb, free-standing woody, aquatic, climber, epiphyte and parasite (see Supplementary data S1 for more details). We found such data for 218,429 of the 326,101 species in the extant global flora, and for 4,708 of the 5,317 introduced cultivated species in our dataset.

For functional traits, there is less complete data for the global flora than there is for native range size and growth-form. Therefore, we only analyzed whether these traits are associated with naturalization success of the introduced cultivated species (i.e. we did not compare them to the global flora). We chose specific leaf area, height and seed mass, because they are part of the leaf-height-seed (LHS) plant-strategy scheme (Westoby 1998) and capture major variation in plant strategies (e.g. the acquisitive-conservative continuum) (Diaz *et al.*

2016). Data on these traits were extracted from the TRY database (Kattge *et al.* 2020), and complemented with data from other sources (Supplementary data S2). For the cultivated plants ($n = 5,317$), data on specific leaf area was available for 1,389 (26%) species, data on seed mass for 2,816 (52%) species, and data on height for 3,265 (61%) species.

Phylogenetic tree construction

To assess phylogenetic biases among cultivated and naturalized plants in Southern Africa (i.e. whether species from certain clades are more likely to be cultivated and naturalized than expected by chance), we constructed a phylogenetic tree of the global seed-plant flora (for more details see Supplementary Data S2), including all species with accepted names in The Plant List ($n = 326,101$). In addition, for the analyses of the introduced cultivated flora of Southern Africa, we also constructed a phylogenetic tree for those species by pruning the global flora tree (5,317 species, Supplementary Figure S2).

Statistical analysis

All statistical analyses were done in R, version 3.6.1 (R Core Team 2019).

Tests for biases in origin, native range size, growth form and phylogenetic composition

To test for biases in geographic origins, native range size and growth-forms of the introduced and naturalized cultivated floras of Southern Africa, we used randomization tests (Supplementary Fig. S3). We first created global source pools of species that are not native to Southern Africa. We did this separately for continent of origin, native range size and growth-form (see Supplementary Data S4).

Using the global source pools, and the introduced and naturalized cultivated species lists, we did three types of comparisons (Supplementary Fig. S3). First, we tested whether the introduced cultivated species in Southern Africa are more or less likely to originate from a particular continent, to have a bigger or smaller native range size, and to be more or less likely to belong to a particular growth-form than expected by chance (i.e. when compared to the global flora outside of Southern Africa). Second, we tested this also for all naturalized cultivated species in Southern Africa, to assess which patterns one would find if one would not account for introduction bias. Third, to account for introduction bias, we tested whether the naturalized cultivated species are a random subsample of the introduced cultivated flora of Southern Africa. For the first and second comparisons, we randomly drew 9,999 times from the global source pools the number of introduced cultivated species with available data (5,317, 5,067 and 4,708 for the continent-of-origin, native-range-size and growth-form tests, respectively), and the number of naturalized cultivated species (571, 565 and 541 for the continent-of-origin, native-range-size and growth-form tests, respectively), separately. For the third comparison, we used the introduced cultivated species of Southern Africa as species source pool, and randomly drew from it 9,999 times the number of species that have become naturalized and had available data (571, 565 and 541 for the continent-of-origin, native-range-size and growth-form tests, respectively). The observed proportions of introduced and naturalized cultivated species from each specific continent of origin, native range size or growth-form category were then compared to the distributions of the corresponding random draws. We considered the observed number to be smaller or greater than expected if it was within the lower 2.5% or upper 2.5% of the distributions of random draws, respectively.

To test if there is a phylogenetic bias in the composition of the introduced and naturalized cultivated floras, we again used a randomization test. We assessed whether the introduced and naturalized cultivated species are more closely related (i.e. clustered), less

213 closely related (i.e. overdispersed) or equally related to what would be expected by chance. To
214 quantify the phylogenetic clustering of the species, we calculated Faith's phylogenetic diversity
215 (PD), which is a measure of the evolutionary history contained within a set of species and is
216 calculated as the cumulative length of the branches connecting the taxa (Faith 1992). First, we
217 calculated the observed PD of introduced and naturalized cultivated species. Then, we
218 randomly drew 999 subsets with the same number as introduced cultivated species (5,317) and
219 naturalized cultivated species (571) from the global flora outside of Southern Africa to
220 calculate the expected PD values. For the third comparison, we compared the observed PD of
221 naturalized cultivated species to a distribution of expected PD values calculated from 999
222 randomly drawn subsets of species from the phylogeny of cultivated species. As an index of
223 phylogenetic signal, we calculated the standardized effect size (SES) as the difference between
224 the observed phylogenetic diversity and the mean value of the distribution of expected values
225 divided by the standard deviation of the expected values.

226 *Association of naturalization extent with geographic origins and growth forms*

227 To test whether, among the naturalized species, the naturalization extent (i.e. the number of
228 occupied regions in Southern Africa; 41 in total as some countries have multiple regions)
229 depended on the geographic origin or growth-forms of the species, we did again a
230 randomization test. We created a pool of the number of regions in Southern Africa that each of
231 the naturalized species occur in. Then, separately for each continent of origin or growth-form,
232 we randomly drew a number of values equal to the number of naturalized species that belong
233 to the respective continent or growth-form, and calculated the mean number of regions. This
234 was repeated 9,999 times.

Association of naturalization success with functional traits

To test whether the naturalization success of introduced cultivated species in Southern Africa is associated with plant height, seed mass or specific leaf area, we used generalized linear models (GLMs) with a binomial error distribution and a logit link function. Correlation between these variables were weak (all $|r| \leq 0.17$; see Supplementary Fig. S4), but, as there were many missing data, we ran separate GLMs for each predictor variable. Each predictor was standardized to a mean of zero and a standard deviation of one. To test for potential non-linear relationships between naturalization success and predictor variables, we also included a quadratic term for each predictor variable, but removed it when it was not significant. We also implemented a phylogenetically corrected GLM using the “phyloglm” function of the R package ‘phylolm’, version 2.6 (Tung Ho & Ané 2014). However, because the results were very similar (Supplementary Table S3) and the standard binomial GLMs had the lowest AIC values, we only present results of the latter.

RESULTS

Of the entire cultivated flora, 5,317 species (77%) have been introduced, and of those 571 (10.7%) are naturalized in at least one region of Southern Africa.

Patterns in origin of introduced and naturalized cultivated plants

The continent (i.e. TDWG level-1 region; Fig. 1a) that has donated most species to the introduced cultivated flora of Southern Africa is Temperate Asia (29.6%), followed by Southern America (28.6%) and Northern America (22.8%) (Fig. 1b). Species originating from Temperate Asia (41.1%) and Southern America (29.7%) also compose the largest numbers of naturalized species, closely followed by other parts of Africa, Europe, and Northern America (Fig. 1b). The smallest donor of introduced and naturalized cultivated species is the Pacific

Islands region (134 and 19 species, respectively). There are also 32 species of unknown origin, only known from cultivation, of which 2 are naturalized.

Compared to the global flora, the numbers of cultivated species in Southern Africa that are native to Temperate Asia, Africa (excluding Southern Africa), Northern America, Australasia and Europe are significantly higher than expected (Fig. 2a; Supplementary Fig. S5). The numbers of cultivated species from Southern America and the Pacific Islands are significantly lower than expected (Fig. 2a; Supplementary Fig. S5). The pattern for the naturalized cultivated flora of Southern Africa is overall very similar (Fig. 2a; Supplementary Fig. S5). However, the numbers of naturalized species from Southern America and the Pacific Islands do not deviate significantly from expectation, and the number of species from Tropical Asia is significantly higher than expected (Fig. 2a; Supplementary Fig. S5).

When the geographic origins of naturalized species are compared to those of the cultivated flora of Southern Africa, species from Temperate Asia, Tropical Asia, Europe, Northern America, and Africa are significantly over-represented (Fig. 2a; Supplementary Fig. S5). In other words, once introduced, species from those continents are more likely to naturalize. The naturalization extents of the species from the Americas were significantly higher than expected, whereas the ones from Australasia were significantly less widespread than expected (Fig. 2b).

Patterns in native range size of introduced and naturalized cultivated plants

Compared to the global flora, the native range sizes of the introduced and naturalized cultivated floras are significantly larger than expected (introduced: standardized effect size [SES] = 79.9, $P < 0.001$; naturalized: SES = 56.5, $P < 0.001$; Fig. 3). When the average native range size of the naturalized species is compared to that of the introduced cultivated flora of Southern Africa,

it is still bigger than expected ($SES = 16.7$, $P < 0.001$, Fig.3), but the standardized effect size is lower than when compared to the global flora.

Phylogenetic patterns among introduced and naturalized cultivated plants

Analyses of Faith's phylogenetic diversity revealed that both the introduced and naturalized cultivated species in Southern Africa are phylogenetically clustered subsets of the global seed-plant flora (introduced: standardized effect size [SES] = -22.5 $P < 0.001$; naturalized: SES = -10.0 , $P < 0.001$; Fig. 4a&b, Supplementary Fig. S1). The naturalized species are also a phylogenetically clustered subset of the cultivated species (SES = -6.6 , $P < 0.001$), but less so than when compared to the global flora (Fig. 4a, Supplementary Fig. S2).

Patterns in growth forms of introduced and naturalized cultivated plants

Compared to the relative representation of different growth-forms in the global flora, more epiphytes, short-lived herbs, long-lived herbs and woody species have been introduced for cultivation in Southern Africa than expected by chance (Fig. 5a, Supplementary Fig. S6). Aquatics, climbers and parasitic plants, on the other hand, are under-represented (Fig. 5a, Supplementary Fig. S6). Likewise, in the naturalized cultivated flora, short-lived herbs and woody species are over-represented, and parasitic plants are under-represented (Fig. 5a, Supplementary Fig. S6). In addition, epiphytes are also significantly under-represented, but aquatics, climbers and long-lived herbs, do not deviate from the expected numbers (Fig. 5a, Supplementary Fig. S6).

When comparing proportions of the different growth-forms in the naturalized flora with those of the cultivated flora, short-lived herbs are still significantly over-represented, but woody species are not (Fig. 5a, Supplementary Fig. S6). In addition, climbers are over-

represented (Fig. 5a, Supplementary Fig. S5). The naturalization extents of the aquatic, short-lived herbs and long-lived herbs were significantly higher than expected, whereas the other growth-forms did not deviate from expectations (Fig. 5b).

Associations of functional traits with naturalization of cultivated plants

Among the cultivated flora of Southern Africa, the probability of naturalization increased with specific leaf area of the species (Fig. 6a, Supplementary Table S3). Furthermore, the probability of naturalization was non-linearly related to seed mass and height of the plants, with the highest naturalization probability at intermediate trait values (Fig. 6b&c, Supplementary Table S3).

DISCUSSION

By comparing the cultivated flora of Southern Africa to the global flora, our study is—to the best of our knowledge—the first one that quantified introduction biases. We showed that the species that humans have introduced for cultivation in Southern Africa do not represent a random subset of the global flora. They are biased with regard to their native origin, growth form and phylogenetic composition. In some cases, the characteristics favored in cultivated species also increased the probability of naturalization of those species. In other cases, patterns in characteristics of naturalized species were entirely determined by introduction biases. However, we were also able to identify correlates of naturalization success that were otherwise masked by introduction biases. So, patterns in origin, growth form, phylogenetic composition and functional traits of naturalized species may largely reflect biases in the species that humans chose to introduce for cultivation, and this needs to be accounted for.

Patterns in geographic origin and native range size

We showed that continents vary in their contributions to the introduced and naturalized cultivated floras of Southern Africa. In a previous global analysis, it was shown that in the naturalized flora of the African continent, as well as in the naturalized floras of most other continents, alien species from Africa, Europe, Northern America and Temperate Asia are over-represented (van Kleunen *et al.* 2015a). Similarly, we found that these continents are also over-represented as donors of the naturalized cultivated flora, and that in addition plants from Australasia and Tropical Asia are over-represented. Interestingly, the over-representation of Australasian species in the naturalized flora is completely accounted for by a higher than expected introduction. In other words, Australasian species do not have higher naturalization probabilities once introduced. On the other hand, the Tropical Asian species were not more likely to be introduced but have a higher naturalization probability once introduced. This is in line with the results of a recent global analysis of naturalization success of economic plants (van Kleunen *et al.* 2020), and suggests that species from Tropical Asia have an innate higher naturalization potential.

A recent study showed that Africa, Australasia, Europe, Northern America and Temperate Asia are disproportionally over-represented as donors of economic plants globally (van Kleunen *et al.* 2020). The same continents are also over-represented in the cultivated flora of Southern Africa (Fig. 2a). The over-representation of European plants is likely a consequence of the colonial history of Southern Africa, as European settlers frequently brought plants with them (Crosby 1986; Faulkner *et al.* 2020; van Wilgen *et al.* 2020). The high representation of Temperate Asian and African plants could reflect that many species native to Europe are also native to parts of those continents. This is particularly the case for species in the Mediterranean, which is one of the world's centers of origin of cultivated plants (Vavilov 1992). The over-representation of Northern American and Australasian plants could potentially

be explained by the fact that the European colonial empires exchanged many economic plants between their colonies (Heywood 1989).

The high naturalization success of plants from Africa, Europe and Temperate Asia could reflect that many of them are from the Mediterranean, and thus find climatically suitable areas in Southern Africa (Peel *et al.* 2007). Indeed, climatic suitability might also be one of the determinants of naturalization success. However, climatic similarity cannot fully explain naturalization success, as species from other continents with climates similar to Southern Africa, such as Australasia, are under-represented among naturalized species. Possibly the higher naturalization probability of plants from certain continents could also be explained by higher planting frequencies (i.e. propagule pressure) (Lockwood *et al.* 2005; Colautti *et al.* 2006) or an earlier introduction. However, as data on planting frequencies and year of introduction are not available for the species in our dataset, we could not test this.

We found that naturalization success in Southern Africa was positively associated with native range size. This was also evident from the observation that the sum of the proportions of species native to the different continents increased from the global flora, via the introduced cultivated flora, to the naturalized cultivated flora (Fig. 2a). This indicates that species that are native to multiple continents, and thus have large native ranges, have a higher chance to be introduced for cultivation and naturalize. This provides further evidence that native range size is positively associated with invasion success (Scott & Panetta 1993; Goodwin *et al.* 1999; Prinzing *et al.* 2002; Maurel *et al.* 2016; Razanajatovo *et al.* 2016), but see (Milbau & Stout 2008) for an exception. However, it cannot fully explain the positive association between native range size and the probability of naturalization among the cultivated plants. It could be though that widespread species were introduced earlier, as has been found for garden plants introduced to Central Europe (Maurel *et al.* 2016), and that they therefore had more time to naturalize.

Furthermore, species with large native ranges could be ecologically and genetically more versatile and have traits that allow them to easily establish (Pysek *et al.* 2009).

Phylogenetic patterns

In a global analysis, Pysek *et al.* (2017) showed that some plant families have more or fewer naturalized species than expected based on their species richness. Therefore, it is not surprising that we found a strong phylogenetic structure for the naturalized plants in Southern Africa. Importantly, however, we also found that there was a strong phylogenetic pattern in the introduced cultivated flora of Southern Africa, and that the phylogenetic signal of naturalization was weaker when compared to the cultivated flora instead of the global flora. This shows that a large proportion of the phylogenetic pattern in naturalization success is due to an introduction bias. A similar result was recently found in an analysis of naturalization success of economic plants in the extant global seed flora (van Kleunen *et al.* 2020). Nevertheless, the phylogenetic signal of naturalization among cultivated plants was still highly significant. This indicates that naturalization ability is phylogenetically clustered, possibly because some phylogenetically conserved characteristics that are shared among closely related species promote naturalization success (Pyšek & Richardson 2007; van Kleunen *et al.* 2010; van Kleunen *et al.* 2015b).

Patterns in species traits

We found evidence that some growth forms were over-represented and others were under-represented in the introduced and naturalized cultivated floras of Southern Africa. However, it is important to note that these biases that we observed are contingent on the assumption that the global flora is well represented by the 204,100 species for which we found growth-form

data. As was the case for geographic origin and phylogeny, this pattern in the naturalized flora was partly driven by introduction biases. South Africa has been called the “world capital of tree invasions” (Richardson *et al.* 2020), and our study confirms that free-standing woody species form the largest growth-form category among the naturalized cultivated species. Importantly, however, our study also shows that woody species were only over-represented among naturalized plants because they were more likely to be introduced. Due to a paucity of native trees suitable for timber production, many woody species have been introduced to Southern Africa for forestry plantations, and also for environmental purposes such as the provisioning of shade and stabilization of sand dunes (Richardson *et al.* 2003; Richardson *et al.* 2020; van Wilgen *et al.* 2020). The short-lived herbs, on the other hand, were both more likely to be introduced and to naturalize following introduction. This may partly reflect that many annual crop species are short-lived herbs, and that most crop species in Africa have been introduced (Wood 1988).

Although most studies only consider linear effects of species traits, we found clear evidence that naturalization success was highest for cultivated plants with intermediate values of seed mass and plant height. While very small seeds might not contain sufficient resources to warrant seedling survival in stressful environments, very large seeds might have limited dispersal ability and are produced in low numbers (Westoby *et al.* 1992; Kempel *et al.* 2013). With regard to plant height, although a short stature might be beneficial in stressful low-resource environments, tall stature increases competitive ability as well as seed dispersal (Moles *et al.* 2009). However, very tall species may due to their higher water requirements not be successful in Southern Africa. In contrast to seed mass and height, specific leaf area showed a positive association with naturalization success. While the benefits of low and high specific leaf area are also likely to depend on the habitat, many studies have found that invasion success is associated with high specific leaf area (Pyšek & Richardson 2007; van Kleunen *et al.* 2010).

This might be because specific leaf area is indicative of fast growth (Lambers & Poorter 1992), which might be advantageous particularly in the resource-rich environments where invasions typically start.

Conclusions

We showed that there are strong patterns in the geographic origin, phylogeny and species characteristics of the introduced cultivated flora of Southern Africa, and that these biases largely underlie patterns in the naturalized flora. For example, while Australasian species are disproportionately over-represented among naturalized cultivated plants, this is solely because a disproportional number of Australasian species has been introduced. Similarly, the large number of naturalized free-standing woody species is due to a disproportional number of woody species that has been introduced. On the other hand, certain groups of species, such as the ones from Tropical Asia and climbers, might appear as having no specifically high naturalization probabilities when comparing them to the global flora, although they have disproportionately high naturalization probabilities once introduced. We further showed that among the introduced cultivated plants of Southern Africa, naturalization success is related to functional traits (i.e. seed mass, height and specific leaf area). Importantly, we showed that such relationships can be non-linear, with the highest naturalization probabilities for species with intermediate seed masses and heights. Overall, our case study for Southern Africa does not only show that it is important to account for introduction biases, but also to quantify the biases. Therefore, we emphasize the need to collate data not only on those species that have become naturalized but also on the ones that have been introduced for cultivation and have not yet become naturalized.

ACKNOWLEDGEMENTS

AO thanks the Ministry of Higher Education of Sudan and the University of Konstanz for funding and support from the International Max Planck Research School for Organismal Biology. Mvk thanks the German Research Foundation DFG for funding (grants 264740629 and 432253815). We thank C. Gommel, K. Mamonova, V. Pasqualetto and B. Rüter for help with data extraction, and Lesley Henderson for providing lists of naturalized species for South Africa. FE acknowledges funding by the Austrian Science Foundation FWF (grant I 3757-B29). PP and JP were supported by EXPRO grant no. 19-28807X (Czech Science Foundation) and long-term research development project RVO 67985939 (Czech Academy of Sciences).

REFERENCES

- Baker, H.G. (1965). Characteristics and modes of origin of weeds. In: *The Genetics of Colonizing Species*. (eds. Baker, HG & Stebbins, G). Academic Press Inc., N.Y. New York, pp. 147-172.
- Blackburn, T.M., Pysek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarosik, V. *et al.* (2011). A proposed unified framework for biological invasions. *Trends Ecol Evol*, 26, 333-339.
- Brummitt, R.K. (2001). *World Geographical Scheme for Recording Plant Distributions*. 2 edn. Hunt Institute for Botanical Documentation, Pittsburgh.
- Bucharova, A. & van Kleunen, M. (2009). Introduction history and species characteristics partly explain naturalization success of North American woody species in Europe. *Journal of Ecology*, 97, 230-238.
- Cayuela, L., macarro, I., Stein, A. & Oksanen, J. (2019). Taxonstand: Taxonomic Standardization of Plant Species Names. R package version 2.2. <https://CRAN.Rproject.org/package=Taxonstand>.
- Chrobock, T., Kempel, A., Fischer, M. & van Kleunen, M. (2011). Introduction bias: Cultivated alien plant species germinate faster and more abundantly than native species in Switzerland. *Basic and Applied Ecology*, 12, 244-250.
- Colautti, R.I., Grigorovich, I.A. & MacIsaac, H.J. (2006). Propagule pressure: A null model for biological invasions. *Biological Invasions*, 8, 1023-1037.

- Crosby, A. (1986). *Ecological Imperialism: The Biological Expansion of Europe*. Cambridge University Press., Cambridge.
- Dawson, W., Burslem, D.F.R.P. & Hulme, P.E. (2009). Factors explaining alien plant invasion success in a tropical ecosystem differ at each stage of invasion. *Journal of Ecology*, 97, 657-665.
- Diaz, S., Kattge, J., Cornelissen, J.H., Wright, I.J., Lavorel, S., Dray, S. *et al.* (2016). The global spectrum of plant form and function. *Nature*, 529, 167-171.
- Dodd, A.J., Burgman, M.A., McCarthy, M.A. & Ainsworth, N. (2015). The changing patterns of plant naturalization in Australia. *Diversity and Distributions*, 21, 1038-1050.
- Faith, D.P. (1992). Conservation Evaluation and Phylogenetic Diversity. *Biological Conservation*, 61, 1-10.
- Faulkner, K.T., Burness, A., Byrne, M.J., Kumschick, S., Peters, K., Robertson, M.P. *et al.* (2020). South Africa's Pathways of Introduction and Dispersal and How They Have Changed Over Time. In: *Biological Invasions in South Africa* (eds. van Wilgen, BW, Measey, J, Richardson, DM, Wilson, JR & Zengeya, TA). Springer International Publishing Cham, pp. 313-354.
- Faulkner, K.T., Robertson, M.P., Rouget, M. & Wilson, J.R.U. (2016). Understanding and managing the introduction pathways of alien taxa: South Africa as a case study. *Biological Invasions*, 18, 73-87.
- Glen, H.F. (2002). *Cultivated Plants of Southern Africa: Botanical Names, Common Names, Origins, Literature*. Jacana: National Botanical Institute, Johannesburg.
- Goodwin, B.J., McAllister, A.J. & Fahrig, L. (1999). Predicting invasiveness of plant species based on biological information. *Conservation Biology*, 13, 422-426.
- Haeuser, E., Dawson, W., Thuiller, W., Dullinger, S., Block, S., Bossdorf, O. *et al.* (2018). European ornamental garden flora as an invasion debt under climate change. *J Appl Ecol*, 55, 2386-2395.
- Hamilton, M.A., Murray, B.R., Cadotte, M.W., Hose, G.C., Baker, A.C., Harris, C.J. *et al.* (2005). Life-history correlates of plant invasiveness at regional and continental scales. *Ecology Letters*, 8, 1066-1074.
- Heywood, V. (1989). Patterns, extents and modes of invasions by terrestrial plants. In: *Biological Invasions : A Global Perspective* (eds. Darke, JA, Mooney, HA, di Castri, F, Groves, RH, Kruger, FJ, Rejmanek, M *et al.*). John Wiley & Sons, pp. 31-60.
- Hulme, P.E., Bacher, S., Kenis, M., Klotz, S., Kuhn, I., Minchin, D. *et al.* (2008). Grasping at the routes of biological invasions: a framework for integrating pathways into policy. *J Appl Ecol*, 45, 403-414.

- Kattge, J., Bonisch, G., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P. *et al.* (2020). TRY plant trait database - enhanced coverage and open access. *Glob Chang Biol*, 26, 119-188.
- Kempel, A., Chrobock, T., Fischer, M., Rohr, R.P. & van Kleunen, M. (2013). Determinants of plant establishment success in a multispecies introduction experiment with native and alien species. *Proc Natl Acad Sci U S A*, 110, 12727-12732.
- Kueffer, C., Pyšek, P. & Richardson, D.M. (2013). Integrative invasion science: model systems, multi-site studies, focused meta-analysis and invasion syndromes. *New Phytol*, 200, 615-633.
- Küster, E.C., Kühn, I., Bruehlheide, H. & Klotz, S. (2008). Trait interactions help explain plant invasion success in the German flora. *Journal of Ecology*, 96, 860-868.
- Lambdon, P.-W., Pyšek, P., Basnou, C., Hejda, M., Arianoutsou, M., Essl, F. *et al.* (2008). Alien flora of Europe: Species diversity, temporal trends, geographical patterns and research needs. *Preslia -Praha*, 80, 101-149.
- Lambers, H. & Poorter, H. (1992). Inherent variation in growth-rate between higher-plants - a search for physiological causes and ecological consequences. *Advances in Ecological Research*, 23, 187-261.
- Lewis, S.L. & Maslin, M.A. (2015). Defining the Anthropocene. *Nature*, 519, 171-180.
- Lockwood, J.L., Cassey, P. & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends Ecol Evol*, 20, 223-228.
- Maurel, N., Hanspach, J., Kuhn, I., Pyšek, P. & van Kleunen, M. (2016). Introduction bias affects relationships between the characteristics of ornamental alien plants and their naturalization success. *Global Ecology and Biogeography*, 25, 1500-1509.
- McGregor, K.F., Watt, M.S., Hulme, P.E. & Duncan, R.P. (2012). What determines pine naturalization: species traits, climate suitability or forestry use? *Diversity and Distributions*, 18, 1013-1023.
- Milbau, A. & Stout, J.C. (2008). Factors associated with alien plants transitioning from casual, to naturalized, to invasive. *Conserv Biol*, 22, 308-317.
- Moles, A.T., Warton, D.I., Warman, L., Swenson, N.G., Laffan, S.W., Zanne, A.E. *et al.* (2009). Global patterns in plant height. *Journal of Ecology*, 97, 923-932.
- Moodley, D., Geerts, S., Richardson, D.M. & Wilson, J.R. (2013). Different traits determine introduction, naturalization and invasion success in woody plants: Proteaceae as a test case. *PLoS One*, 8, e75078.
- Ordóñez, A., Wright, I.J. & Olff, H. (2010). Functional differences between native and alien species: a global-scale comparison. *Functional Ecology*, 24, 1353-1361.

- Peel, M.C., Finlayson, B.L. & McMahon, T.A. (2007). Updated world map of the Koppen-Geiger climate classification. *Hydrology and Earth System Sciences*, 11, 1633-1644.
- Pergl, J., Sadlo, J., Petrik, P., Danihelka, J., Chrtek, J., Hejda, M. *et al.* (2016). Dark side of the fence: ornamental plants as a source of wild-growing flora in the Czech Republic. *Preslia*, 88, 163-184.
- POWO (2019). Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; <http://www.plantsoftheworldonline.org/> Retrieved 10-26 August 2020.
- Prinzing, A., Durka, W., Klotz, S. & Brandl, R. (2002). Which species become aliens? *Evolutionary Ecology Research*, 4, 385-405.
- Pyšek, P., Jarosik, V., Pergl, J., Randall, R., Chytrý, M., Kuhn, I. *et al.* (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., Jarošík, V., Pergl, J., Randall, R., Chytrý, M., Kühn, I. *et al.* (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., Pergl, J., Essl, F., Lenzner, B., Dawson, W., Kreft, H. *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, D.M. (2007). Traits associated with invasiveness in alien plants: where do we stand? In: *Biological Invasions* (ed. Nentwig, W). Springer Berlin Heidelberg Berlin, pp. 97-125.
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria.
- Razanajatovo, M., Maurel, N., Dawson, W., Essl, F., Kreft, H., Pergl, J. *et al.* (2016). Plants capable of selfing are more likely to become naturalized. *Nat Commun*, 7, 13313.
- Reichard, S.H. & White, P. (2001). Horticulture as a pathway of invasive plant introductions in the United States. *Bioscience*, 51, 103-113.
- Richardson, D.M., Cambray, J.A., Chapman, R.A., Dean, W.R.J., Griffiths, C.L., Le Maitre, D.C. *et al.* (2003). Vectors and pathways of biological invasions in South Africa - past, present, and future. In: *Invasive Species: Vectors And Management Strategies* (eds. G, R & J, C), pp. 292-349.
- Richardson, D.M., Foxcroft, L.C., Latombe, G., Le Maitre, D.C., Rouget, M. & Wilson, J.R. (2020). The Biogeography of South African Terrestrial Plant Invasions. In: *Biological Invasions in South Africa* (eds. van Wilgen, BW, Measey, J, Richardson, DM, Wilson, JR & Zengeya, TA). Springer International Publishing Cham, pp. 67-96.

- Richardson, D.M., Pysek, P., Rejmanek, M., Barbour, M.G., Panetta, F.D. & West, C.J. (2000). Naturalization and invasion of alien plants: concepts and definitions. *Diversity* 6, 93-107.
- Scott, J.K. & Panetta, F.D. (1993). Predicting the Australian Weed Status of Southern African Plants. *J Biogeogr*, 20, 87-93.
- Tung Ho, L.s. & Ané, C. (2014). A Linear-Time Algorithm for Gaussian and Non-Gaussian Trait Evolution Models. *Systematic Biology*, 63, 397-408.
- van Kleunen, M., Bossdorf, O. & Dawson, W. (2018a). The Ecology and Evolution of Alien Plants. *Annual Review of Ecology, Evolution, and Systematics*, 49, 25-47.
- van Kleunen, M., Dawson, W., Essl, F., Pergl, J., Winter, M., Weber, E. *et al.* (2015a). Global exchange and accumulation of non-native plants. *Nature*, 525, 100-103.
- van Kleunen, M., Dawson, W. & Maurel, N. (2015b). Characteristics of successful alien plants. *Mol Ecol*, 24, 1954-1968.
- van Kleunen, M., Essl, F., Pergl, J., Brundu, G., Carboni, M., Dullinger, S. *et al.* (2018b). The changing role of ornamental horticulture in alien plant invasions. *Biol Rev*, 93, 1421-1437.
- van Kleunen, M., Pysek, P., Dawson, W., Essl, F., Kreft, H., Pergl, J. *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. *Ecol Lett*, 13, 235-245.
- van Kleunen, M., Xu, X., Yang, Q., Maurel, N., Zhang, Z., Dawson, W. *et al.* (2020). Economic use of plants is key to their naturalization success. *Nat Commun*, 11, 3201.
- van Wilgen, B.W., Measey, J., Richardson, D.M., Wilson, J.R. & Zengeya, T.A. (2020). Biological invasions in South Africa: an overview. In: *Biological Invasions in South Africa* (eds. van Wilgen, BW, Measey, J, Richardson, DM, Wilson, JR & Zengeya, TA). Springer International Publishing, pp. 3-31.
- Vavilov, N.I. (1992). *Origin and Geography of Cultivated Plants* (transl. D Lo'Ve). Cambridge University Press, Cambridge.
- Westoby, M. (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, 199, 213-227.
- Westoby, M., Jurado, E. & Leishman, M. (1992). Comparative evolutionary ecology of seed size. *Trends Ecol Evol*, 7, 368-372.
- Wood, D. (1988). Introduced crops in developing countries. *Food Policy*, 13, 167-177.

FIGURES

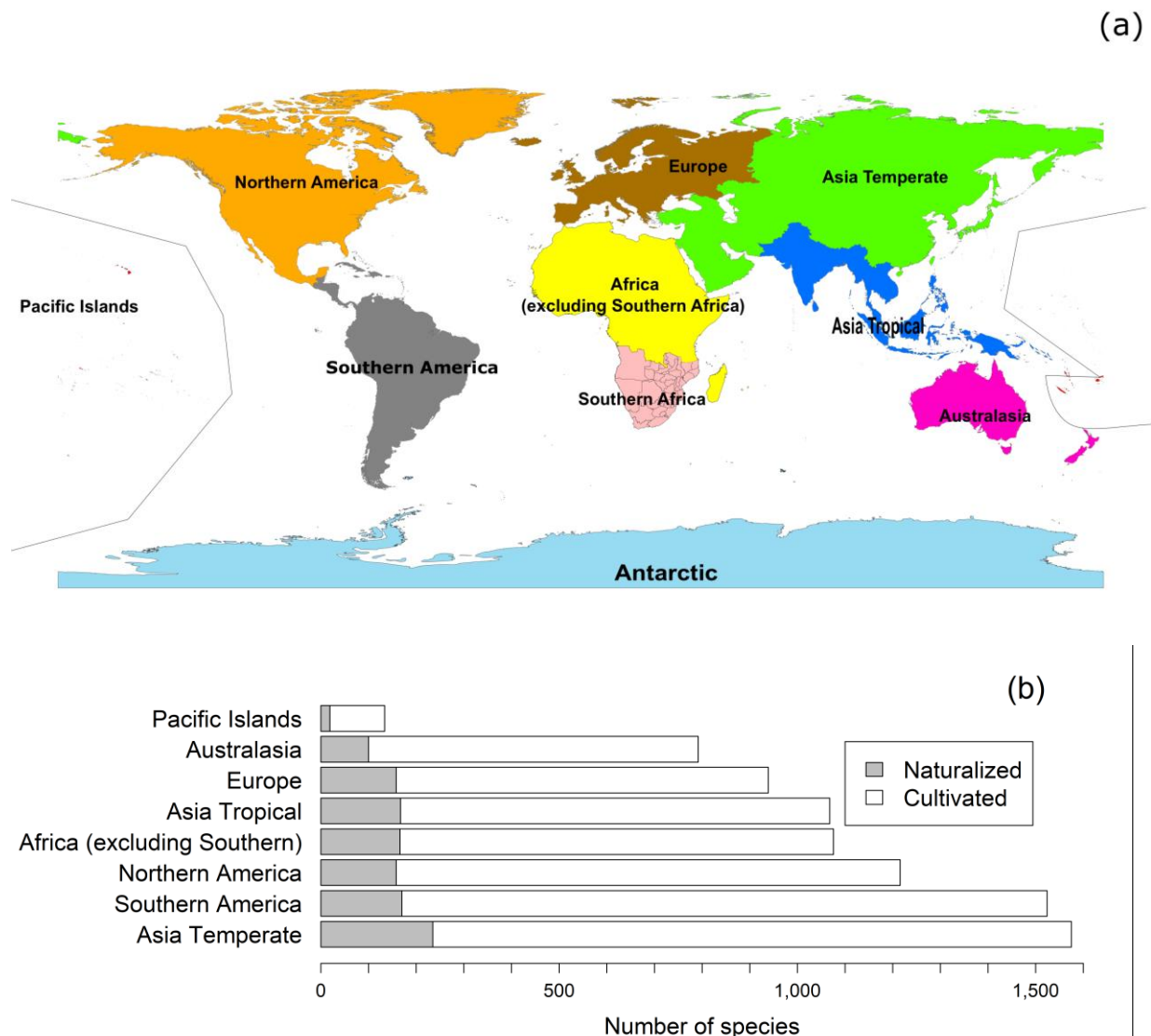


Figure 1 Donor continents of the introduced and naturalized cultivated floras of Southern Africa. (a) The nine TDWG continents and the ten Southern African countries: Angola, Botswana, Eswatini, Lesotho, Malawi, Mozambique, Namibia, South Africa, Zambia and Zimbabwe (our study area). (b) Numbers of introduced cultivated plant species in Southern Africa according to their continents of origin. The cultivated species that have become naturalized are indicated in grey. As none of the introduced cultivated species is native to Antarctica, this continent is not included in b.

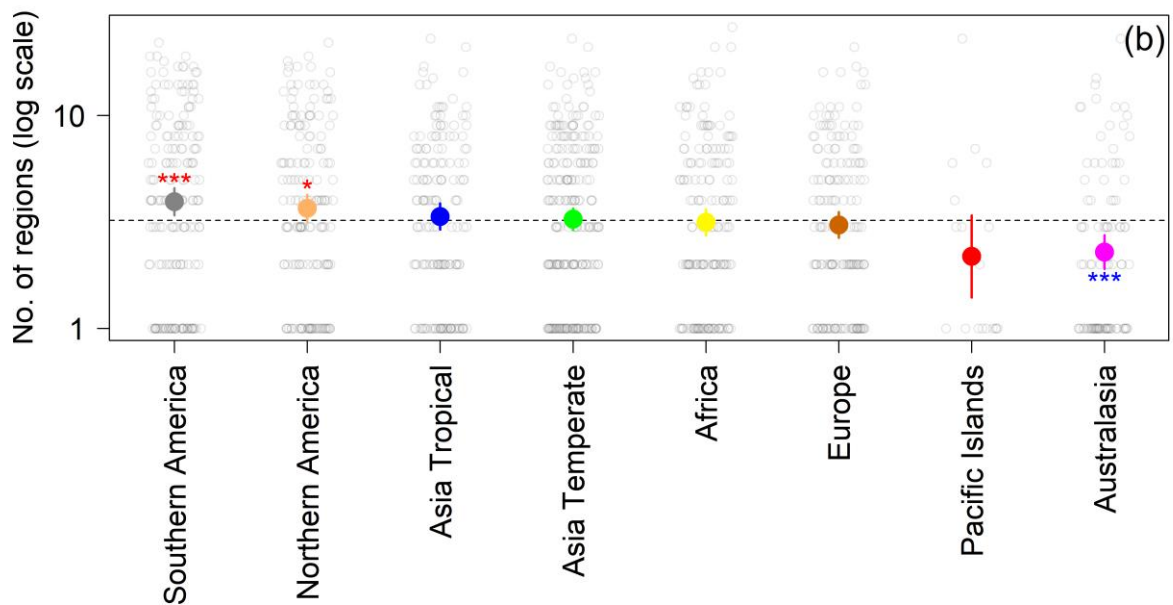
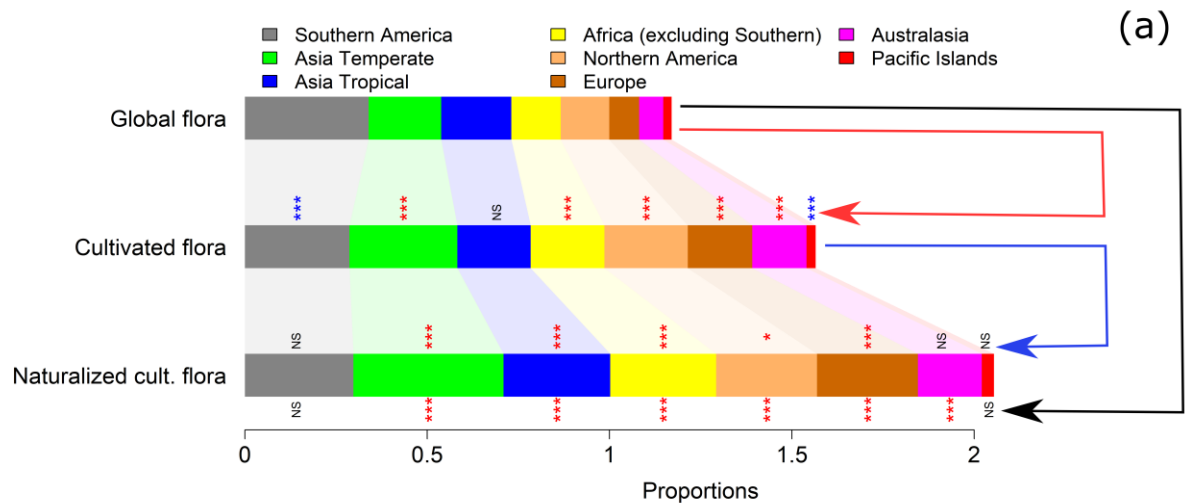


Figure 2 Patterns in geographic origin of introduced and naturalized cultivated plants. (a) Proportions of species native to different continents in the global flora, and the introduced and naturalized cultivated floras of Southern Africa. The lighter colored connections between the stacked horizontal bars visualize which proportions increase, remain the same or decrease if one goes from the global flora, via the introduced cultivated flora, to the naturalized cultivated flora. Headed arrows represent the flows of species from the global flora to the introduced cultivated flora (red) and naturalized flora (black), and from the introduced cultivated flora of southern Africa to the naturalized flora (blue). Asterisks below the stacked horizontal bars of the naturalized cultivated flora indicate whether the species flows from the respective continents are significantly over-represented (red) or under-represented (blue) among the naturalized cultivated flora relative to the proportions in the global flora. Similarly, asterisks on the connections between the stacked bars indicate whether species from the respective continent are over- or under-represented among the introduced cultivated flora relative to the global flora, and among the naturalized cultivated flora relative to the introduced cultivated flora. ***: $P < 0.001$, NS: not significant (see Supplementary Figure S5). Note, that the

cumulative proportions exceed one, because species can be native to multiple continents. As none of the introduced cultivated species is native to Antarctica, this continent is not included in the figure. (b) The mean (\pm 95% CI) of the log₁₀ of the number of Southern African regions in which the naturalized cultivated species have become naturalized per continent of origin. The data for each individual species are shown in grey. The horizontal dashed line indicates the mean number of regions in which the naturalized cultivated species managed to naturalize. Red and blue asterisks indicate that the observed extents of the naturalized ranges are significantly larger and smaller, respectively, than expected (***: $P < 0.001$, **: $P < 0.01$, *: $P < 0.05$).

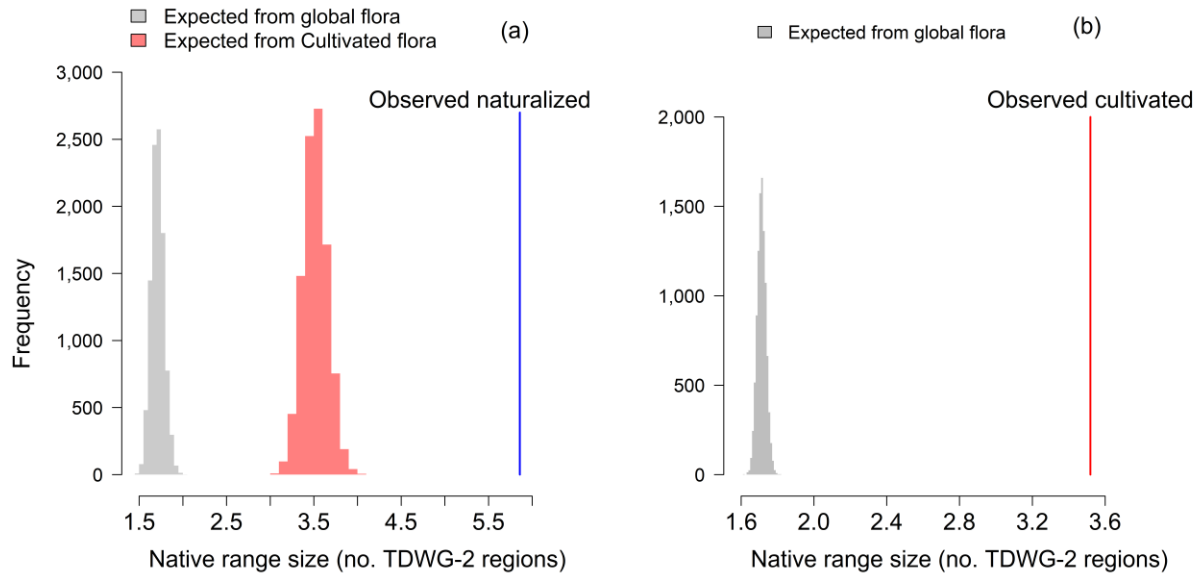


Figure 3 Results of tests for biases in native range size of the introduced and naturalized cultivated floras of Southern Africa. (a) The observed average native range size of a species in the naturalized cultivated flora of Southern Africa (blue vertical line) relative to expectations based on the global flora (grey histogram) and the introduced cultivated flora (red histogram). (b) The observed average native range size of a species in the introduced cultivated flora of Southern Africa (red line) relative to expectations based on the global flora (grey histogram). The histograms of the expected average native range size are based on 9,999 random draws from their respective reference floras. The observed native range size (blue and red vertical lines) are bigger than expected, indicating significant positive associations between native range size and introduction and naturalization success.

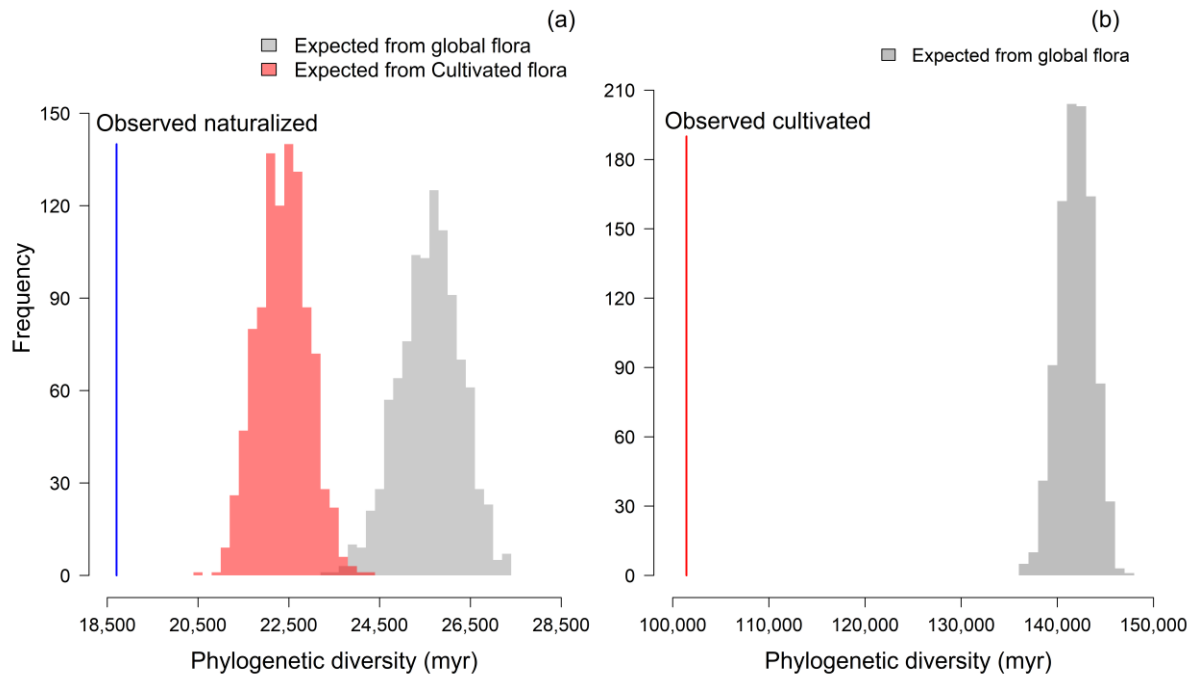


Figure 4 Results of tests for phylogenetic clustering of the introduced and naturalized cultivated floras of Southern Africa. (a) Phylogenetic diversity of the naturalized cultivated flora of Southern Africa (blue vertical line) relative to expectations based on the global flora (grey histogram) and the introduced cultivated flora (red histogram). (b) Phylogenetic diversity of the introduced cultivated flora of Southern Africa (red line) relative to expectations based on the global flora (grey histogram). The histograms of expected phylogenetic diversities are based on 999 random draws from the respective reference floras. The observed phylogenetic diversities (blue and red vertical lines) are smaller than expected, indicating significant phylogenetic clustering.

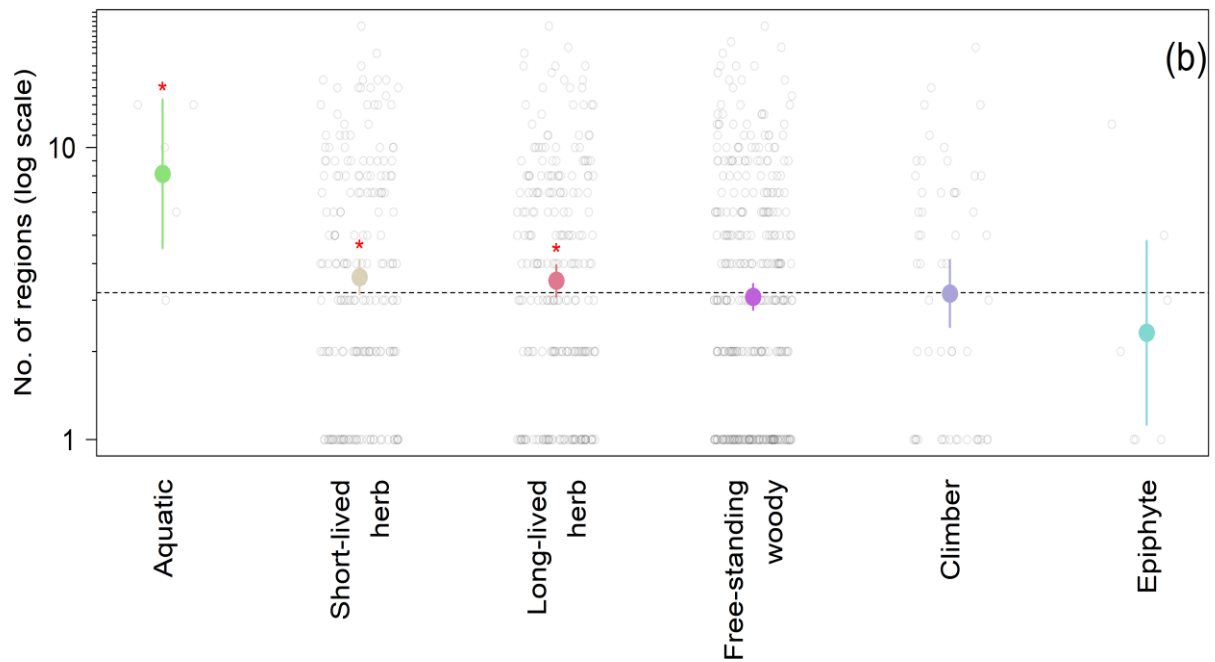
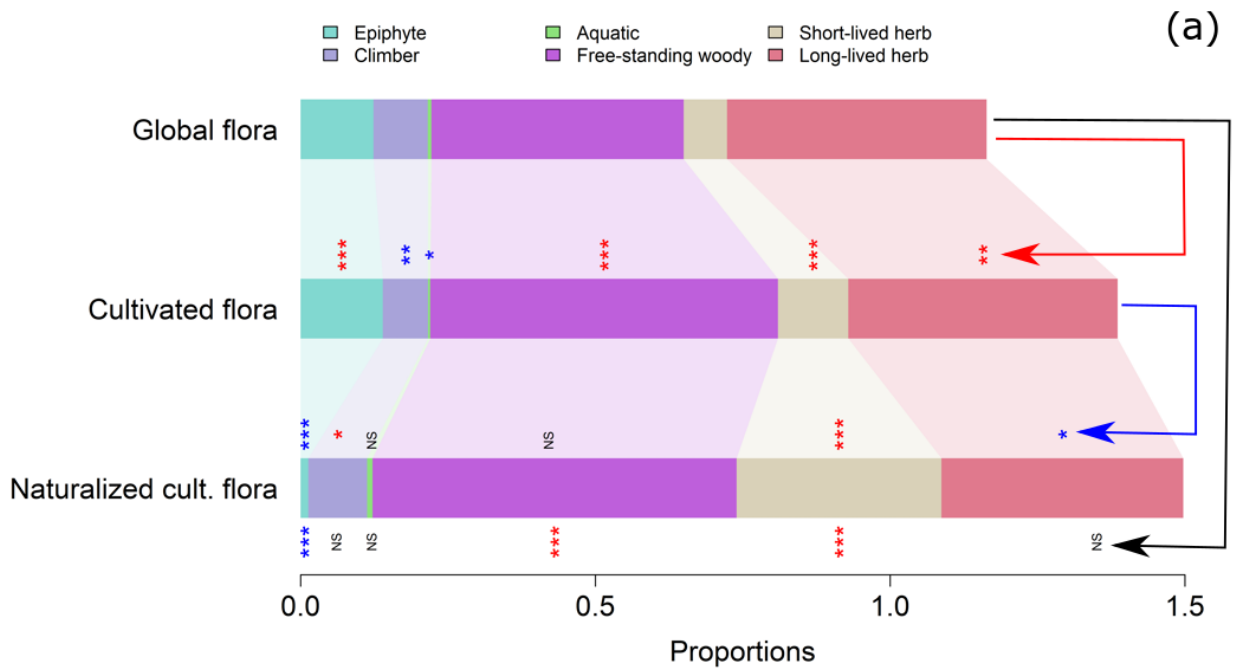


Figure 5 Patterns in growth-forms of introduced and naturalized cultivated plants. (a) Proportions of species with different growth forms for the global flora, and the introduced and naturalized cultivated floras of Southern Africa. The lighter colored connections between the stacked horizontal bars indicate whether the proportions increase, remain the same or decrease if one goes from the global flora, via the introduced cultivated flora, to the naturalized cultivated flora. Headed arrows represent the flows of species from the global flora to the introduced cultivated flora (red) and naturalized flora (black), and from the introduced

cultivated flora of southern Africa to the naturalized flora (blue). Asterisks below the stacked
 horizontal bars of the naturalized cultivated flora indicate whether species with different
 growth-forms are significantly over-represented (red) or under-represented (blue) among the
 naturalized flora relative to the proportions in the global flora. Similarly, asterisks on the
 connections between the stacked bars indicate whether species with the respective growth-
 forms are over- or under-represented among the introduced cultivated flora relative to the
 global flora, and among the naturalized cultivated flora relative to the introduced cultivated
 flora. ***: $P < 0.001$, *: $P < 0.05$, NS: not significant (see Supplementary Figure S6). Note,
 that the cumulative proportions exceed one, because some species can take on more than one
 growth-form. The parasitic plants are not shown as they are a very small fraction of the global
 flora. (b) The mean (\pm 95% CI) of the log10 of the number of Southern African regions in
 which the naturalized cultivated species have become naturalized per growth form category.
 The data for each individual species are shown in grey. The horizontal dashed line indicates
 the mean number of regions in which the naturalized cultivated species managed to naturalize.
 Red and blue asterisks indicate that the observed extents of the naturalized ranges are
 significantly larger and smaller, respectively, than expected (*: $P < 0.05$).

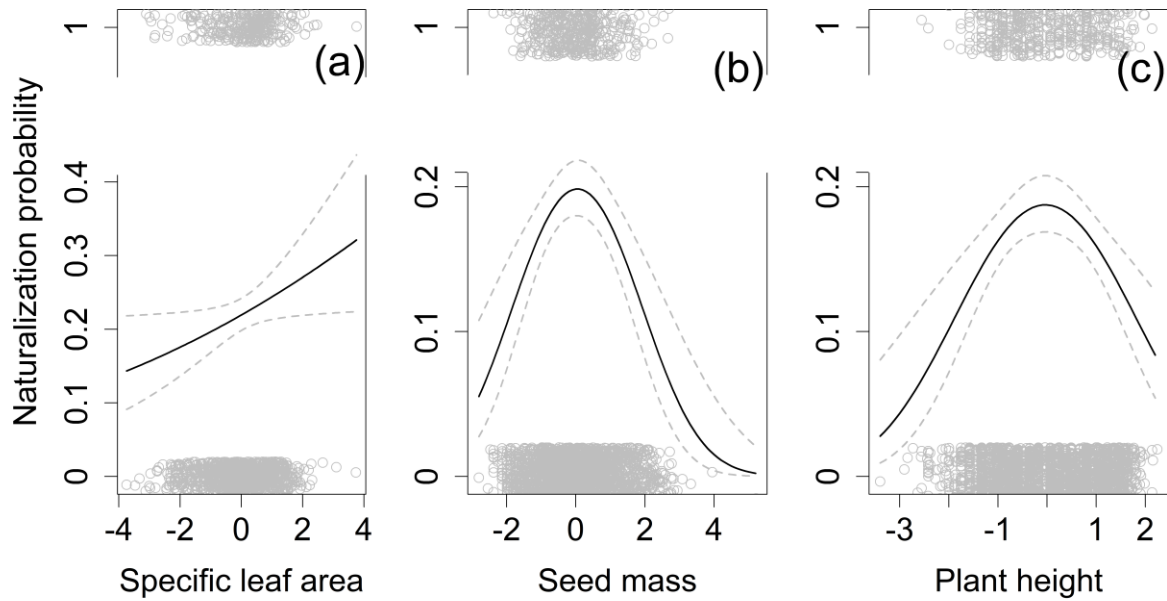


Figure 6 Naturalization of introduced cultivated alien plants in Southern Africa in relation to (a) specific leaf area, (b) individual seed mass, and (c) plant height. The solid lines are the predicted relationships from binomial GLMs (Supplementary Table S3), and the dashed lines indicate the 95% CIs. The results of phylogenetic binomial GLMs were very similar to those of the standard binomial GLMs (Supplementary Table S3). The values of the functional traits of the naturalized and non-naturalized species are shown as points at one and zero respectively (jittered to increase visibility). The values of the traits were standardized to a mean of zero and a standard deviation of one.