

ARTICLE

Received 23 Mar 2016 | Accepted 21 Sep 2016 | Published 31 Oct 2016

DOI: 10.1038/ncomms13313

OPEN

# Plants capable of selfing are more likely to become naturalized

Mialy Razanajatovo<sup>1</sup>, Noëlie Maurel<sup>1</sup>, Wayne Dawson<sup>2</sup>, Franz Essl<sup>3,4</sup>, Holger Kreft<sup>5</sup>, Jan Pergl<sup>6</sup>, Petr Pyšek<sup>6,7</sup>, Patrick Weigelt<sup>5</sup>, Marten Winter<sup>8</sup> & Mark van Kleunen<sup>1</sup>

Many plant species have established self-sustaining populations outside their natural range because of human activities. Plants with selfing ability should be more likely to establish outside their historical range because they can reproduce from a single individual when mates or pollinators are not available. Here, we compile a global breeding-system database of 1,752 angiosperm species and use phylogenetic generalized linear models and path analyses to test relationships between selfing ability, life history, native range size and global naturalization status. Selfing ability is associated with annual or biennial life history and a large native range, which both positively correlate with the probability of naturalization. Path analysis suggests that a high selfing ability directly increases the number of regions where a species is naturalized. Our results provide robust evidence across flowering plants at the global scale that high selfing ability fosters alien plant naturalization both directly and indirectly.

Konstanzer Online-Publikations-System (KOPS)  
URL: <http://nbn-resolving.de/urn:nbn:de:bsz:352-0-374616>

<sup>1</sup>Ecology, Department of Biology, University of Konstanz, Universitätsstrasse 10, Konstanz D-78457, Germany. <sup>2</sup>Conservation Ecology Group, Department of Biosciences, Durham University, South Road, Durham DH1 3LE, UK. <sup>3</sup>Division of Conservation, Vegetation and Landscape Ecology, University of Vienna, Wien 1030, Austria. <sup>4</sup>Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Matieland 7602, South Africa. <sup>5</sup>Biodiversity, Macroecology and Biogeography, University of Göttingen, Büsgenweg 1, Göttingen D-37077, Germany. <sup>6</sup>Institute of Botany, Department of Invasion Ecology, The Czech Academy of Sciences, Průhonice, CZ-25243, Czech Republic. <sup>7</sup>Department of Ecology, Faculty of Science, Charles University, Viničná 7, Prague 2, CZ-12844, Czech Republic. <sup>8</sup>German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, Leipzig D-04103, Germany. Correspondence and requests for materials should be addressed to M.R. (email: [mialy.razanajatovo@uni-konstanz.de](mailto:mialy.razanajatovo@uni-konstanz.de)).

The human-caused introduction and subsequent spread of species into new regions has become a defining feature of global environmental change in the Anthropocene<sup>1</sup>. At least 3.9% of species in the global flora has now established self-sustaining populations in regions where they did not occur naturally<sup>2</sup>. Most such naturalized plants have been introduced for horticulture<sup>3</sup>, and subsequently escaped from gardens, nurseries and arboreta<sup>4</sup>. Other plant naturalizations have resulted from accidental introductions, for example as seed contaminant of imported crops<sup>5</sup>. Some naturalized plants have become invasive in many regions of the world, where they often have detrimental impacts on the environment and human societies<sup>6</sup>. Naturalization is a key stage of the invasion process when species overcome the barriers that prevent them from establishing self-sustaining populations in the wild<sup>7,8</sup>. Thus, understanding what drives naturalization is a key biological question of global interest<sup>9</sup>.

The ability to reproduce is of particular importance for understanding naturalization, because propagule supply is required for the founding and maintenance of populations<sup>10,11</sup>. Baker<sup>12</sup> posed that species capable of uniparental reproduction are more likely to establish after long-distance dispersal, because they can reproduce from a single individual. This hypothesis, known as Baker's Law or Baker's Rule<sup>13</sup>, may apply to natural long-distance dispersal events as well as to species introduced by humans to new regions, because suitable mates and/or pollinators may be scarce there<sup>14</sup>. When mates are scarce, self-compatible species should have an advantage over self-incompatible ones. Furthermore, when pollinators are scarce, those self-compatible species that can self-pollinate—and thus are autofertile—should especially have a major advantage. The role of the breeding system could thus be most relevant during naturalization<sup>14</sup>.

Previous tests of Baker's Law with alien plants have yielded contradicting findings. While many naturalized and invasive plants have high selfing ability in their alien range<sup>15–18</sup>, some globally invasive plants do not possess this ability<sup>19,20</sup>. Some studies even suggest that the breeding system only plays a minor role in naturalization<sup>21</sup>. However, previous works were limited to specific taxonomic or functional groups of plants, to the invasion rather than the naturalization stage and/or to a restricted part of the globe. Thus, we still lack a comprehensive test of the role of breeding systems in explaining naturalization of alien plant species at the global scale.

A species' breeding system has commonly been defined as a qualitative trait<sup>22,23</sup> (for example, self-compatible versus self-incompatible). However, species classified as self-compatible might not be fully self-compatible, and those classified as self-incompatible might be partly self-compatible<sup>24</sup>. Furthermore, as shown in some species, even self-compatible ones, selfing may not occur at all while in others selfing may occur rarely, sometimes, often or nearly always<sup>22</sup>. This continuity in the degree of self-compatibility (SC) and autofertility (AF) may be important for naturalization, and thus, quantitative metrics of selfing ability are preferable. Despite the extensive number of studies on the breeding systems of flowering plants around the globe, only few efforts have been made to compile the individual quantitative findings into a global database<sup>22,25</sup>. Such an approach, covering as many regions and plant species as possible, is necessary to address general scientific questions related to plant reproductive strategies.

Several plant characteristics are often reported as being associated with naturalization as well as with selfing ability. First, short-lived species such as annuals and biennials predominantly occur in habitats subjected to frequent disturbances, and therefore often colonize anthropogenic habitats (that is, they become weedy<sup>26</sup>), and are also often selfers<sup>14,27,28</sup>. Second, species with selfing ability are more likely to have larger native

geographic ranges than those lacking this ability<sup>29,30</sup>, and species with large geographic ranges or a wide habitat breadth are more likely to naturalize elsewhere<sup>31–33</sup>. It is thus important to unravel the complex causal associations among these different factors to assess whether relationships between selfing ability and naturalization, if any, are direct or indirect. Furthermore, species with different degrees of evolutionary relatedness might have different degrees of similarity in plant traits associated with naturalization and in traits associated with selfing ability. Therefore, it is necessary to account for other traits and phylogeny<sup>34</sup> when testing the relationship between naturalization and breeding system.

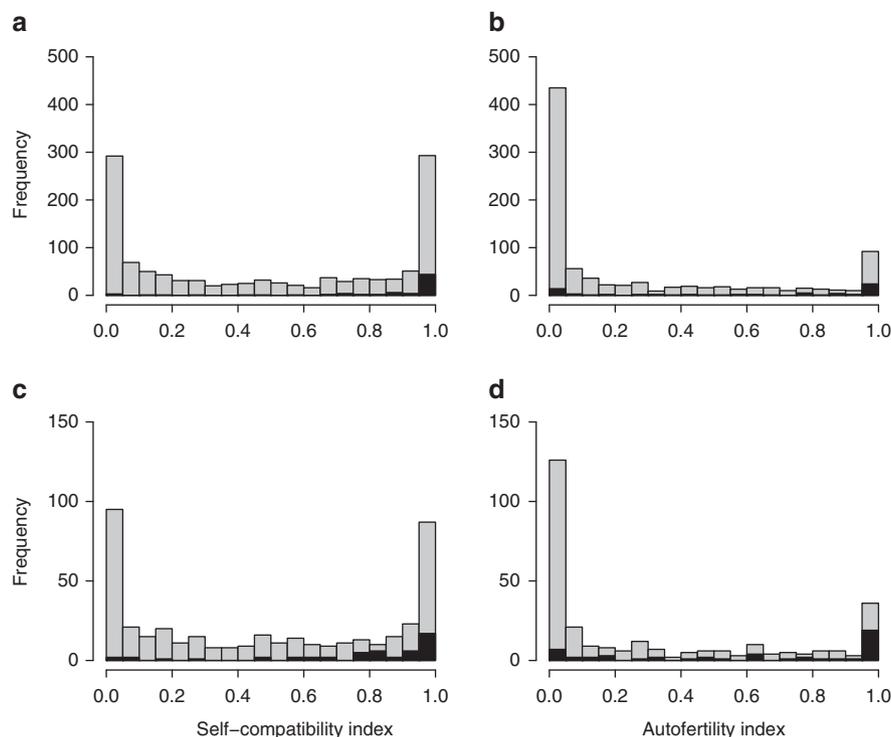
To test the effects of SC and AF on naturalization at the global scale, we assembled a global breeding-system database of flowering plants. We quantified species' selfing ability using indices of SC and AF<sup>22,35</sup>. We combined these data with information on life history, native range size and global information on the incidence and the extent of naturalization (expressed as being naturalized or not outside the native range and the number of regions where naturalized, respectively) using the Global Naturalized Alien Flora database (GloNAF)<sup>2</sup>, which is the most comprehensive data source on naturalized alien plant species distributions. We used a phylogenetically informed generalized linear model approach to account for the effect of relatedness among species. Further, we used path analysis to separate the potential direct effects of SC and AF on naturalization from the indirect associations caused by correlations with life history and native range size. Our questions were: (i) are the incidence (whether or not a species has naturalized somewhere) and extent of naturalization (the number of regions where naturalized) of alien plant species associated with SC and AF, as suggested by Baker's Law? (ii) On the basis of a path analysis, what is the strength of potential direct and indirect effects of SC and AF on the incidence and extent of naturalization when accounting for life history and native range size?

Here we show that SC and AF are positively related to whether a species has naturalized somewhere and the number of regions where a species is naturalized, both directly and indirectly. Furthermore, path analysis suggests that SC and AF are associated with annual or biennial life history and a large native range, which both positively correlate with the probability of naturalization. High SC and AF also increase with the number of regions where a species is naturalized. Our results provide evidence of the validity of Baker's Law across flowering plants at the global scale, and of its significance for alien plant naturalization.

## Results

**Global patterns in self-compatibility and autofertility.** Based on a comprehensive literature search, we found breeding-system data for a total of 1,752 species (Supplementary Data 1) covering all major clades of angiosperms (Supplementary Fig. 1), from 161 families (Supplementary Fig. 2), and from all over the world (Supplementary Fig. 3). The indices of SC and AF had a continuous bimodal frequency distribution from zero (indicating absence of selfing ability) to one (indicating high selfing ability), and were higher for short-lived (annual and biennial) species than for perennials (Fig. 1).

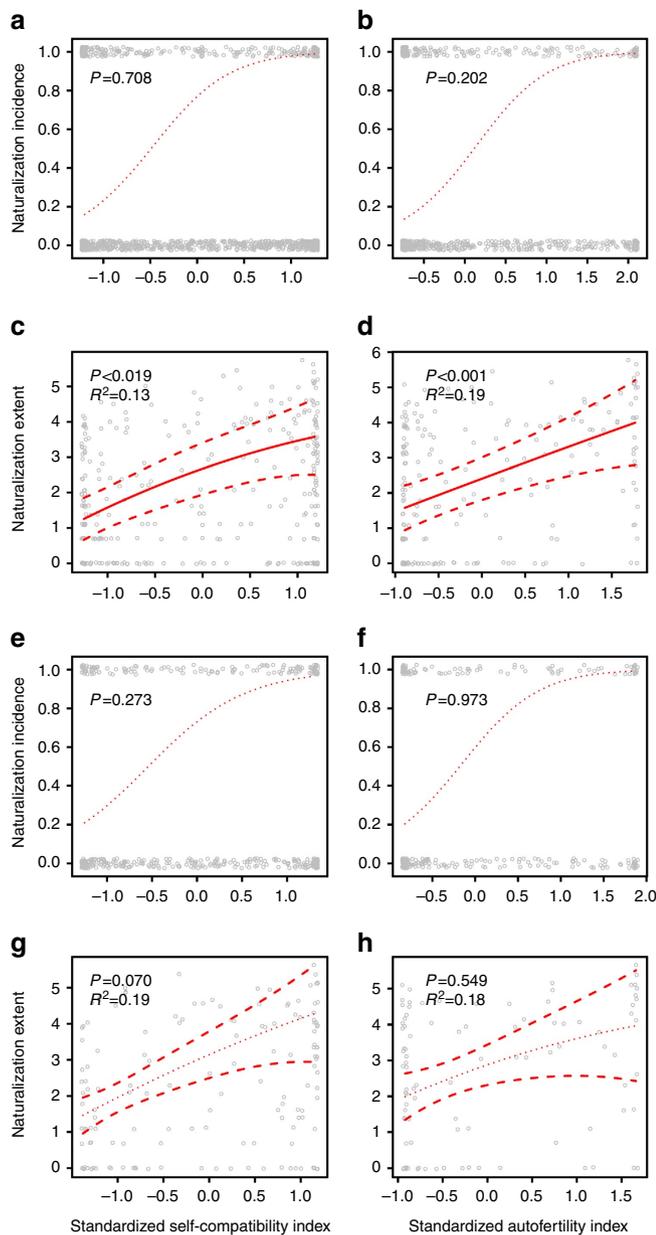
**Naturalization incidence associated with selfing ability.** Of the 1,752 species in our database, 498 have become naturalized somewhere. Although the naturalization incidence appeared to be positively associated with SC and AF indices, calculated for fruit set, that is, fruit/flower ratio (Fig. 2a,b), these effects were not significant in the phylogenetic logistic regressions (Table 1).



**Figure 1 | Global patterns in self-compatibility and autofertility.** Frequency distributions of four indices of selfing ability compiled in our global database on breeding systems of angiosperms. **(a)** SC index calculated based on fruit set ( $n=1,192$ ). **(b)** AF index calculated based on fruit set ( $n=873$ ). **(c)** SC index calculated based on the number of seeds per flower ( $n=422$ ). **(d)** AF index calculated based on the number of seeds per flower ( $n=286$ ). Sample sizes refer to the total number of species from individual studies in the breeding-system database. Annual and biennial species are shown in black and perennial ones in grey.

However, a species was significantly more likely to naturalize somewhere if its native range was large, as an increase of one standard deviation (0.854 in the analysis with SC, and 0.848 in the analysis with AF) in the log-transformed number of regions where a species is native increased its odds of being naturalized somewhere by a factor of  $e^{0.710}=2.034$  and  $e^{0.720}=2.054$  (Table 1). Moreover, annuals and biennials had a higher naturalization incidence than perennials, as an annual or a biennial life history increased a species' odds of being naturalized somewhere by a factor of  $e^{0.731}=2.077$  compared with a perennial life history, at least so in the analysis with AF (Table 1). Although the logistic regressions that accounted for phylogeny had a better fit than the ones without phylogeny, the results of the latter were qualitatively similar (Supplementary Table 1). Path analysis, in which we could not account for phylogeny, but which allowed us to explicitly differentiate between direct and indirect relationships, confirmed that the degrees of SC and AF of species had no significant direct effects on naturalization incidence (Fig. 3a,b). However, species with a high SC and AF were more likely to naturalize, as they were likely to have a large native range, which in turn can be inferred to increase their ability to naturalize (Fig. 3a,b). Moreover, SC and AF were highest for species with an annual or biennial life history, which were more likely to naturalize than perennials (Fig. 3a,b). Although the results were slightly different for SC and AF indices based on seed production per flower instead of fruit set, these differences were small. Then, naturalization incidence was also not significantly correlated with SC and AF in the phylogenetic logistic regressions (Table 1; Fig. 2e,f). However, in the path analysis, the indirect relationship between SC, native range size and naturalization incidence was significant (Fig. 3e), but such an indirect relationship was not significant in the analysis with AF (Fig. 3f).

**Naturalization extent associated with selfing ability.** Among the 498 species in the breeding-system database that have become naturalized somewhere, the naturalization extent (measured as the number of regions) ranged from 1 to 409 regions (out of 843) of the GloNAF database. In the phylogenetic linear regression, the naturalization extent was significantly positively related to the SC and AF indices, calculated for fruit set (Table 1 and Fig. 2c,d). An increase of 1 s.d. in the SC index (0.410) and AF index (0.374) of a species increased the log-transformed number of regions where it is naturalized by 0.256 and 0.472, respectively (Table 1). The extent of naturalization was not significantly correlated with native range size, but annuals and biennials had a larger naturalization extent than perennials (Table 1). An annual or a biennial life history increased the log-transformed number of regions where a species is naturalized by 1.463 and 1.048 compared with a perennial life history in the analysis of SC and AF, respectively (Table 1). The results were qualitatively similar when not correcting for phylogeny (Supplementary Table 1). Path analysis confirmed that the degrees of SC and AF of species as well as life history had significant direct relationships with naturalization extent (Fig. 3c,d). An increase of one standard deviation in the SC index (0.410) and AF index (0.374) of a species increased the log-transformed number of regions where it is naturalized by 0.128 and 0.244 s.d., respectively. As SC and AF were highest for species with an annual or biennial life history, they were also indirectly correlated with naturalization extent (Fig. 3c,d). The paths from selfing ability to native range size and from native range size to naturalization extent were not significant (Fig. 3c,d). Although the results were slightly different for SC indices based on seed production per flower instead of fruit set, these differences were small. Then, the correlation between SC and naturalization extent was marginally significant (Table 1). However, in the path analysis, the degree of SC had a



**Figure 2 | Global naturalization of alien plant species in relation to selfing ability.** (a–d) Analysis with indices calculated using fruit set.

(a) A phylogenetic logistic regression testing how naturalization incidence (expressed as being naturalized somewhere or not) depends on SC index ( $n = 1,181$ ). (b) A phylogenetic logistic regression testing how naturalization incidence depends on AF index ( $n = 866$ ). (c) A phylogenetic linear model testing how naturalization extent (natural log-transformed number of regions where the species is naturalized) depends on SC index ( $n = 295$ ). (d) A phylogenetic linear model testing how naturalization extent depends on AF index ( $n = 203$ ). (e–h) Analysis with indices calculated using seed production per flower. (e) A phylogenetic logistic regression testing how naturalization incidence depends on SC index ( $n = 417$ ). (f) A phylogenetic logistic regression testing how naturalization incidence depends on AF index ( $n = 282$ ). (g) A phylogenetic linear model testing how naturalization extent depends on SC index ( $n = 144$ ). (h) A phylogenetic linear model testing how naturalization extent depends on AF index ( $n = 95$ ). Selfing-ability indices were standardized to a mean of zero and an s.d. of 1. Sample sizes refer to the total number of species from individual studies in the breeding-system database. Solid lines and dotted lines indicate significant and nonsignificant relationships, respectively. Dashed lines represent standard errors of the phylogenetic linear regression. Denoted  $R^2$  are total explained variance by the variables in the model.

significant direct relationship with the naturalization extent (Fig. 3g). The results were different for AF indices based on seed production per flower as then the effect of AF on naturalization extent was not significant (Table 1 and Figs 2h and 3h).

## Discussion

We combined the largest quantitative database on breeding systems of angiosperms collected to date with the most comprehensive database on alien plant species distributions (GloNAF<sup>2</sup>) to assess potential relationships between selfing ability and naturalization. Our models explained 13–24% of the variation in naturalization incidence and extent of species (Table 1 and Figs 2 and 3), which is comparable with the explanatory power of models in previous ecological studies testing the importance of different traits on the naturalization or invasion of alien species (for example, ref. 36). We showed that both incidence and extent of naturalization were positively associated with SC and AF. However, the relationship between selfing ability and whether or not a species has naturalized somewhere was only indirect through a correlation with life history, and a positive relationship with native range size. In contrast, selfing ability was positively associated with the number of regions where a species has naturalized, and selfing ability significantly accounted for 36–59% of the explained variation in naturalization relative to the other variables in the models (Table 1). Similar direct effects have been found at local to regional scales. For example, many South African Iridaceae species that have become naturalized elsewhere have, in the native range, a higher ability for autonomous selfing than their non-naturalized counterparts<sup>37</sup>. Many successful naturalized alien plants also demonstrated high selfing ability in the introduced range<sup>16,18,33</sup>. Our results show for the first time that the association between selfing ability and naturalization holds at a global scale across many families.

SC provides plants with the ability to reproduce even when their populations are small—as is usually the case during the initial establishment—and mating is likely to involve few and closely related individuals (Baker's Law<sup>12,14</sup>). Indeed, SC confers a fitness advantage when mates are limited (reproductive assurance<sup>28</sup>). Selfing may increase seed production<sup>38</sup> and therefore propagule pressure, which is one of the major drivers of naturalization and invasion<sup>39</sup>. Modes of self-pollination are diverse<sup>22,40</sup>; some self-compatible plants still require pollinators whereas others can do autonomous self-pollination, resulting in AF. Geitonogamy, that is, pollen transfer from a flower to another flower of the same plant, and also facilitated autogamy only occur when the introduced plants encounter suitable pollinators in the new environment. However, when pollinators are limited, AF is required to assure reproduction<sup>40,41</sup>. As the indices of SC and AF are not independent, and both are associated with naturalization, it is not possible to infer from our study whether selfing ability is more important for overcoming mate or pollinator limitation. Recently, Pannell *et al.*<sup>14</sup> proposed a narrower circumscription of Baker's Law focusing on the consequences of mate rather than pollinator limitation. However, in a recent study, it was shown that naturalized species might attract more floral visitors than cultivated, but non-naturalized alien species<sup>42</sup>, suggesting that pollinator limitation may be important and should not be ignored.

While our study suggests that high selfing ability may contribute to the naturalization of alien plant species, questions regarding the consequences of selfing for the demography of those species may arise. A major disadvantage of selfing is inbreeding depression<sup>43</sup>. Although many naturalized plants in our database have a high degree of selfing ability, some have also been shown to have substantial inbreeding depression, for

**Table 1 | Global naturalization of alien plant species in relation to selfing ability.**

Response variables Explanatory variables	Naturalization incidence*				Naturalization extent†				R <sup>2‡</sup>
	Estimates	s.e.	z	P	Estimates	s.e.	t	P	
<i>Selfing-ability indices calculated using fruit set</i>									
Analysis with self-compatibility <span style="float:right">n = 1,181</span>									
Intercept	– 1.146	0.126	– 9.07	<0.0001	1.924	0.264	7.28	<0.0001	0.13
Native range size§	0.710	0.082	8.67	<0.0001	0.101	0.096	1.05	0.293	7.14
Annual/biennial§	0.568	0.418	1.36	0.174	1.463	0.475	3.08	0.002	50.00
Self-compatibility index§	0.026	0.070	0.37	0.708	0.256	0.108	2.36	0.019	35.71
Annual/biennial§ × self-compatibility§	0.243	0.362	0.67	0.502	– 0.373	0.444	– 0.84	0.402	7.14
Analysis with autofertility <span style="float:right">n = 866</span>									
Intercept	– 1.276	0.171	– 7.47	<0.0001	2.241	0.294	7.69	<0.0001	0.19
Native range size§	0.720	0.097	7.42	<0.0001	0.033	0.120	0.27	0.785	0.10
Annual/biennial§	0.731	0.352	2.08	0.038	1.048	0.387	2.70	0.007	40.46
Autofertility§	0.104	0.081	1.28	0.202	0.472	0.132	3.58	<0.001	59.15
Annual/biennial§ × autofertility§	0.319	0.241	1.32	0.186	– 0.030	0.302	– 0.10	0.920	0.05
<i>Selfing-ability indices calculated using seed production per flower</i>									
Analysis with self-compatibility <span style="float:right">n = 417</span>									
Intercept	– 0.660	0.152	– 4.35	<0.0001	2.327	0.147	15.85	<0.0001	0.19
Native range size§	0.618	0.126	4.89	<0.0001	0.155	0.148	1.05	0.296	6.44
Annual/biennial§	0.502	0.42	1.19	0.233	1.480	0.455	3.25	0.001	66.48
Self-compatibility index§	0.121	0.110	1.09	0.273	0.275	0.150	1.82	0.070	25.62
Annual/biennial§ × self-compatibility§	– 0.058	0.381	– 0.15	0.880	– 0.237	0.475	– 0.50	0.619	1.46
Analysis with autofertility <span style="float:right">n = 282</span>									
Intercept	– 0.688	0.186	– 3.69	<0.001	2.707	0.250	10.81	<0.0001	0.18
Native range size§	0.900	0.176	5.11	<0.0001	0.068	0.189	0.36	0.720	1.10
Annual/biennial§	0.298	0.407	0.72	0.473	1.624	0.488	3.33	0.001	86.94
Autofertility§	0.004	0.145	0.03	0.973	0.121	0.201	0.60	0.549	7.04
Annual/biennial§ × autofertility§	0.276	0.316	0.87	0.382	– 0.311	0.410	– 0.76	0.449	4.91

Results of four phylogenetic logistic regressions and four phylogenetic linear regressions testing the global naturalization of alien plant species in relation to selfing ability (measured as self-compatibility and autofertility indices), native range size, annual/biennial versus perennial life history and the interaction between life history and selfing ability. Global naturalization was measured as naturalization incidence outside the native range (expressed as being naturalized somewhere or not) and naturalization extent (natural log-transformed number of regions where the species is naturalized). Native range size was measured as the number of TDWG level-2 regions. Self-compatibility and autofertility indices were calculated using fruit set and seed production per flower from our global breeding-system database. Sample sizes refer to the total number of species from individual studies in the breeding-system database.

\*Phylogenetic logistic regression:  $\alpha = 0.035$  in the analysis with self-compatibility and 0.026 in the analysis with autofertility (indices calculated using fruit set); and  $\alpha = 0.033$  in the analysis with self-compatibility and 0.029 in the analysis with autofertility (indices calculated using seed production per flower).

†Phylogenetic linear model:  $\lambda = 0.141$  in the analysis with self-compatibility and 0.142 in the analysis with autofertility (indices calculated using fruit set); and  $\lambda < 0.0001$  in the analysis with self-compatibility and 0.048 in the analysis with autofertility (indices calculated using seed production per flower).

‡Total explained variance by the variables in the model, and the relative importance of each variable calculated as the difference in deviance between the full model and a model without the variable of interest. R<sup>2</sup> could not be calculated for the phylogenetic logistic regressions.

§Variables were rescaled to have a mean of zero and an s.d. of 1.

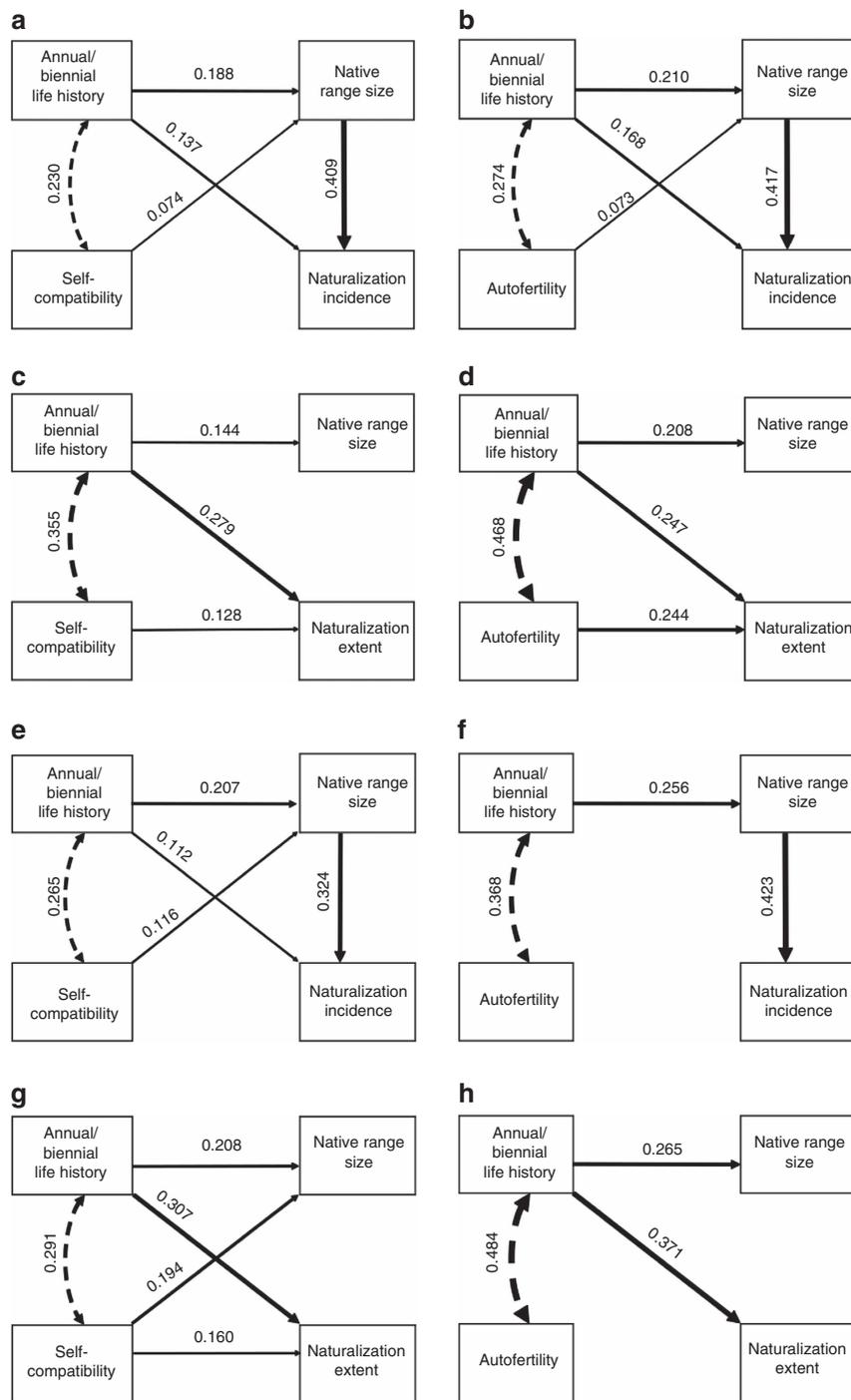
example the tree *Acacia dealbata* Link in the introduced range in South Africa<sup>44</sup>. On the other hand, some self-compatible and autofertile alien plants have negligible inbreeding depression, such as the geophyte *Lilium formosanum* Wallace, also naturalized in South Africa<sup>45</sup>. These two examples suggest that inbreeding depression costs vary among species. Furthermore, selfing ability can vary within species (for example, in different locations<sup>46</sup>). The roles of intraspecific variation in selfing ability and potential short- and long-term disadvantages of selfing for naturalized alien plants require further tests.

Whether a plant species has naturalized outside the native range and how widely it has naturalized might depend on characteristics of the mode of introduction such as introduction effort or pathway, and the spatial distribution of propagules<sup>47</sup>. The species in our study might not have the same probability of being introduced to all potentially suitable regions outside their native ranges<sup>48,49</sup>. Particularly for species that have not become naturalized at all, one cannot exclude that they have not naturalized simply because they were not introduced. This could have obscured a direct effect of selfing ability on naturalization incidence. Nevertheless, because we accounted for native range size, which is a good predictor of the probability of being introduced elsewhere<sup>50</sup>, variation in introduction effort is at least partly accounted for.

Species with higher selfing ability in our database cover larger native ranges (Fig. 3a,b). This finding is in line with the results of

a recent study showing that selfing plants have larger native ranges than their outcrossing close relatives<sup>31</sup>. These results support Baker's Law, in that the ability of plants to overcome mate limitation has contributed to a greater native range expansion<sup>12</sup>. Native range size itself can be a strong predictor of naturalization as shown in previous studies. For example, introduced Poaceae and Fabaceae that had larger native ranges in Asia and Africa had a larger naturalized range in North America<sup>51</sup>. Similarly, temperate shrubs and trees that have larger native ranges had greater naturalization success globally<sup>29</sup>. The probability of central European plants becoming naturalized outside their native range was also determined by native range size<sup>52</sup>. So, overall, species widespread at home tend to be more likely to naturalize.

There are several potential reasons why naturalization may be associated with a large native range. First, if a species is widespread and common, it is more likely to be transported and introduced elsewhere<sup>52</sup>. Second, a large native range could also be associated with traits conferring higher dispersal ability, such as small and numerous seeds<sup>53</sup> and selfing ability<sup>31</sup>. Third, as a large native range tends to contain a broad variety of environmental conditions, species occurring in a large geographic area may be ecologically versatile<sup>52</sup>. It is difficult to say which of these potential mechanisms is most important. However, the strong relationship between native range size and naturalization incidence and the absence of such relationship with naturalization



**Figure 3 | Direct and indirect association of global naturalization with selfing ability.** (a–d) Selfing-ability indices calculated using fruit set from our global breeding-system database. (a) Effects of SC on naturalization incidence (expressed as being naturalized somewhere or not);  $n = 1,181$ , no. of parameters = 7, CFI = 0.993, RMSEA = 0.048 and  $R^2_{\text{naturalization incidence}} = 0.20$ . (b) Effects of AF on naturalization incidence;  $n = 866$ , no. parameters = 7, CFI = 0.988, RMSEA = 0.061 and  $R^2_{\text{naturalization incidence}} = 0.24$ . (c) Effects of SC on the extent of naturalization (number of regions where the species is naturalized);  $n = 295$ , no. parameters = 8, CFI = 1.000, RMSEA = 0.000 and  $R^2_{\text{naturalization extent}} = 0.19$ . (d) Effects of AF on the extent of naturalization;  $n = 203$ , no. parameters = 8, CFI = 1.000, RMSEA = 0.000 and  $R^2_{\text{naturalization extent}} = 0.21$ . (e–h) Selfing-ability indices calculated using seed production per flower from our global breeding-system database. (e) Effects of SC on naturalization incidence;  $n = 417$ , no. parameters = 7, CFI = 0.982, RMSEA = 0.062 and  $R^2_{\text{naturalization incidence}} = 0.14$ . (f) Effects of AF on naturalization incidence;  $n = 282$ , no. parameters = 7, CFI = 0.974, RMSEA = 0.090 and  $R^2_{\text{naturalization incidence}} = 0.22$ . (g) Effects of SC on the extent of naturalization;  $n = 144$ , no. parameters = 8, CFI = 0.990, RMSEA = 0.061 and  $R^2_{\text{naturalization extent}} = 0.18$ . (h) Effects of AF on the extent of naturalization;  $n = 95$ , no. parameters = 8, CFI = 0.991, RMSEA = 0.059 and  $R^2_{\text{naturalization extent}} = 0.18$ . Sample sizes refer to the total number of species from individual studies in the breeding-system database. Curved dashed lines with two-headed arrows indicate bivariate correlations between variables; solid lines with one-headed arrow indicate directed paths connecting the variables. The thickness of each path is proportional to the value of the path coefficient. Only path coefficients that are significant ( $z$ -test,  $P < 0.05$ ) are shown. CFI, Comparative Fit Index; RMSEA, root mean square error of approximation.

extent suggest that at least the increased likelihood of being introduced elsewhere plays a role.

Annual and biennial species were more likely to be naturalized elsewhere, and in larger numbers of regions. Many widely naturalized species are associated with habitats prone to frequent disturbances and thus have weedy characteristics in their native range, and many weedy species are annuals or short-lived plants<sup>20</sup>. Annual and biennial species have an advantage over perennials by reaching maturity earlier. Such characteristics of r-strategists could help alien species to rapidly colonize vacant space—due to widespread and frequent man-made disturbances—before other species do (priority effect<sup>49</sup>). Furthermore, annual and biennial species may depend more on reproductive assurance than perennials, as the latter have more opportunities to reproduce, and frequently can multiply vegetatively<sup>13,14</sup>. Thus, annuals and biennials are likely to have high selfing ability. Our results on a large set of angiosperm species suggest that, along with selfing ability, annual and biennial life histories contribute to alien plant naturalization.

We provide the first comprehensive evidence that, across a large number of angiosperms and at the global scale, selfing ability plays a significant role in the naturalization of alien plant species. Species with high selfing ability are more likely to naturalize outside their native ranges, as such species are also more likely to occupy larger native ranges, which in turn increases their ability to naturalize in new regions to which they were introduced. These results therefore provide robust evidence that high selfing ability fosters naturalization outside the native range, thereby supporting Baker's Law<sup>12</sup>.

## Methods

**Compilation of a global database on breeding systems.** To assemble as much as possible of the published quantitative data on breeding systems of angiosperm species into a global database, we searched for plant breeding-system literature in Web of Science (<http://apps.webofknowledge.com>; Supplementary Fig. 4). We used the search keyword string "TS = (('breeding system' OR 'mating system' OR 'self-compatible' OR 'self-fertile' OR autogam' OR 'auto-fertile' OR outcross' OR apomixis) AND plant)", and searched for all indexed documents from the year 1900 onwards. The last search was done in February 2015. We scanned all 6,678 resulting titles, and excluded the ones that did not present breeding-system data for angiosperms (for example, studies on animals, cryptogams and gymnosperms), and studies that only assessed the realized outcrossing rates based on molecular markers without assessing selfing ability. Subsequently, we screened the abstracts of the remaining 1,675 articles. Studies that explicitly included all or a subset of the following treatments were selected: bagging of flowers to exclude pollinators (testing for AF), bagging of flowers in combination with self-pollination (testing for SC), bagging of flowers in combination with cross-pollination (testing for maximal fruit and seed production when bagged), bagging of flowers in combination with emasculation (testing for apomixis), supplemental hand-pollination (testing for pollen limitation) and unmanipulated control flowers (testing for fruit and seed production under natural conditions). Out of the 705 studies that fulfilled our criteria, we examined 696 full-text documents, as nine studies were not accessible. We included in our database all studies that provided measures of fruit set (fruit/flower ratio) and/or seed production (number of seeds per plant or per flower or per fruit, or seed/ovule ratio) for all or a subset of the above-mentioned treatments. Of the 696 articles, 516 yielded suitable data. We additionally included data from 244 documents that we encountered independently of the literature search (including 113 references used by Raduski *et al.*<sup>25</sup>).

For each species reported in a study, we documented all available fruit and seed production data in each of the different pollination treatments. If data were provided in the text or in tables, we extracted them directly. If the data were provided in graphs only, we extracted the data using ImageJ<sup>54</sup>. According to the availability of the data, we obtained for each species and treatment, the percentage of treated flowers that produced fruits (fruit set) and the number of seeds produced per flower (Supplementary Table 2). When the number of seeds per flower was not provided, we calculated it by multiplying fruit set by the number of seeds per fruit (the percentages of values that were calculated this way are given in Supplementary Table 2).

We found breeding-system data for 1,829 plant taxa. Because some of these taxa may be synonyms, and to facilitate alignment of the species list with other databases (see below), we standardized the scientific names of the species in our database following The Plant List (<http://www.theplantlist.org>), using the package 'Taxonstand'<sup>55</sup> in R<sup>56</sup>. For 14 taxa that did not occur in The Plant List, we kept the names given in the original study from which breeding-system data were extracted.

Because a species' breeding system is often related to its life history, we included for each species, information on its life history (annual, biennial and perennial). When this was not provided in the publication from which the breeding-system data were extracted, we consulted other sources: Encyclopedia of Life (<http://eol.org>), Royal Botanic Gardens, Kew (<http://epic.kew.org>), the United States Department of Agriculture database (<http://plants.usda.gov/java/>), Tropicos (<http://www.tropicos.org>), Instituto de Botanica Darwinio ([http://www.darwin.edu.ar/Proyectos/Flora Argentina/fa.htm](http://www.darwin.edu.ar/Proyectos/FloraArgentina/fa.htm)); and did extensive internet searches.

As a broad native distribution may influence the global naturalization of an alien species, and might be affected by a plant's breeding system and life history, we compiled information on the native range size of each species in the database using the number of TDWG level-2 regions the species is native to (52 regions in total; adapted from the scheme of the Biodiversity Information Standards or Taxonomic Databases Working Group<sup>57</sup>). We extracted data from the World Checklist of Selected Plant Families (WCSP; <http://apps.kew.org/wcsp/>), the Germplasm Resources Information Network (<http://www.ars-grin.gov/cgi-bin/npgs/html/index.pl>), the Global Biodiversity Information Facility (<http://www.gbif.org/>), Tropicos (<http://www.tropicos.org>), the United States Department of Agriculture database (<http://plants.usda.gov/java/>), Encyclopedia of Life (<http://eol.org>), Australian National Botanic Gardens (<https://www.anbg.gov.au/index.html>), CIRAD (<http://arbres-reunion.cirad.fr/accueil>), and IUCN (<http://www.iucnredlist.org/>). For 21 species, we obtained information on the native range from the original study from which breeding-system data had been extracted.

To test whether global naturalization of species is related to their breeding systems, we used the Global Naturalized Alien Flora (GloNAF) database version 1.1<sup>2</sup>, which is the most comprehensive database of naturalized plants. It includes 13,168 alien plant species that have become naturalized in a total of 843 regions (481 mainland regions and 362 island regions) on the globe (~83% of the Earth's land area). For each species in our database, we added data on presence/absence in the GloNAF database (that is, whether or not a species is listed in GloNAF) and the number of regions where the species has become naturalized outside its native range.

We removed 30 taxa that were not identified to the species level from the breeding-system database, because we could not find information on native and naturalized ranges for these taxa. In total, after taxonomic standardization, we had quantitative breeding-system data for 1,752 angiosperm species (Supplementary Data 1) from 161 families (Supplementary Fig. 2), collected in 116 regions (covering tropics, subtropics and temperate zones, Supplementary Fig. 3) using 763 studies. Our database covers experiments performed between 1868 and 2015 in natural populations (1,380 species), plantings (61 species), common gardens (143 species) and greenhouses (269 species). The breeding-system database was not compiled with any *a priori* bias towards the alien species. For 127 species, the breeding system was assessed outside their native range.

**Phylogeny.** To account for potential biases due to different degrees of evolutionary relatedness among species in the analyses<sup>34</sup>, we constructed a phylogenetic tree of the 1,752 species in our database. We first pruned an existing megatree of 31,749 plant species<sup>58</sup> using the 'ape' package<sup>59</sup> in R. The pruned tree returned 821 species in our database that matched with the tips of the megatree. We then added all species belonging to identical genera as polytomies on the pruned tree and obtained a new tree with 1,266 species. We manually added the 486 remaining missing species by supplementing nodes that represented the latest common ancestor of the added taxon. For families that did not occur in the megatree, we decided on the last common ancestor based on taxonomic information on the Angiosperm Phylogeny Website version 13 (<http://www.mobot.org/MOBOT/research/APweb/>). For genera that did not occur in the megatree, we created polytomies under the families. We resolved resulting polytomies using information from online and published phylogenies (Supplementary Table 3). We then obtained the topology of our phylogenetic tree, and estimated the branch length of the tree using *phylocom*<sup>60</sup> considering node ages given by Wikström<sup>61</sup>. For each species that was reported in more than one study ( $n = 125$ ), we added new nodes leading to multiple tips in the tree. To ensure phylogenetic equality, we adjusted the within-species branch lengths to the equivalent of 100 years<sup>62</sup>.

**Data analysis.** To quantify the degree of SC and AF of each species, we calculated an index of SC<sup>22</sup> and an index of AF<sup>22,35</sup>, respectively:

$$SC = \frac{\text{Outcome from self-pollination}}{\text{Outcome from cross-pollination}} \quad (1)$$

$$AF = \frac{\text{Outcome from pollinator exclusion}}{\text{Outcome from cross-pollination}} \quad (2)$$

AF, thus, did not exclude apomixis, which is an asexual mode of reproduction by seed. Each index was calculated for fruit set and for seed production per flower.

An index of zero indicates a low degree of SC or AF; and a value of one indicates a high degree of SC or AF<sup>22,35</sup>. Values larger than one occurred when the outcome from self-pollination or pollinator exclusion was larger than that from cross-pollination<sup>22</sup>. We set such values larger than one to one (values set to one were up to 16% of observations; also see ref. 25). We averaged the indices for each

species within a study. This means that if a study reports selfing-ability data for a species in multiple locations, we used an averaged index. However, if a species was reported in more than one study, we included multiple indices for this species. The maximum number of indices for a species was three (Supplementary Data 1).

To assess potential relationships between SC and AF on global naturalization, we used two approaches. First, we used a generalized linear model approach that allowed us to correct for phylogenetic non-independence of species by including a variance-covariance matrix that contains the phylogenetic distance among species combined with a model of evolution. Second, we used path analysis, in which we could not account for phylogeny, but which allowed us to unravel direct and indirect effects of SC and AF on naturalization.

To test whether naturalization is associated with the degree of SC and AF of species, respectively, we fitted a phylogenetic logistic regression<sup>63</sup> using the package 'phylolm'<sup>64</sup> in R. As a response variable, we used naturalization incidence, that is, whether a species has or has not been recorded as naturalized outside its native ranges, as indicated by its presence/absence in the GloNAF database. As predictor variables, we included native range size (natural log-transformed number of TDWG level-2 regions), annual/biennial life history (yes/no, no if perennial), index of SC or AF, and its interaction with life history. We had such data for a total of 1,115 species from 481 studies for the analysis of SC, and 828 species from 373 studies for that of AF, based on fruit set. As some species were reported in more than one study, we considered the unique combination species-study source as the unit of observation. Therefore, we used more than one index (maximum of three) of SC or AF for 5% and 4% of the species, respectively. Nevertheless, the results remained qualitatively similar when we used one average index per species (Supplementary Table 4). To facilitate comparisons among estimates, we standardized the covariates to a mean of zero and an s.d. of 1 (ref. 65). To quantify the variance explained by the variables in our models, we reported the  $R^2$  for the phylogenetic linear regressions<sup>66</sup>, and we calculated the relative importance of each variable as the difference in deviance between the full model and a model without the variable of interest. Although  $R^2$  can be calculated for non-phylogenetic logistic regression<sup>66</sup> (see values in Supplementary Table 1), such an approach quantifying the  $R^2$  is not available for the phylogenetic logistic regressions.

To test whether the extent of naturalization for species naturalized in at least one region is positively related to its degree of SC and AF, respectively, we analysed the subset of species with breeding-system data that are listed in the GloNAF database. This subset comprised 274 species from 168 studies for SC indices, and 195 species from 114 studies for AF indices, based on fruit set. We fitted phylogenetic linear regressions<sup>67</sup> using the package 'phylolm' in R. As a response variable, we used naturalization extent measured as the natural log-transformed number of regions where the species is naturalized. As predictor variables, we used the same covariates as in the phylogenetic logistic regression described above.

We additionally analysed the complex association of SC and AF of plant species with their life history (annual/biennial or perennial), their native range size and their naturalization outside the native range using path analysis<sup>68</sup>. To specify the path structure, we assumed the following relationships: (i) naturalization outside the native range is driven by the degree of SC and AF of a species, respectively, its native range size and life history (annual/biennial vs perennial); (ii) native range size of a species is driven by its degree of SC or AF, and life history; and (iii) the degrees of SC or AF of a plant species are correlated with an annual or a biennial life history. This path analytical approach allowed us to quantify the potential direct and indirect effects of SC or AF of a plant species on its naturalization outside the native range.

As in the previous models, we separately analysed naturalization incidence (whether the species have been recorded as naturalized outside their native ranges), and for the subset of only the naturalized species, their naturalization extent (the number of regions where they are naturalized). For the analysis of the binary variable naturalization incidence, we used a diagonally weighted least square estimation method with robust standard errors. For the analysis of naturalization extent, we log-transformed the number of regions where the species is native to and the number of regions, where the species is naturalized, and used a maximum likelihood estimation method with robust standard errors (MLR). To facilitate comparisons among the different path coefficients, we standardized all continuous variables to a mean of zero and s.d. of 1. We assessed model fit using a relative measure, the Comparative Fit Index, and an absolute measure, the root mean square error of approximation. A Comparative Fit Index  $>0.95$  and a root mean square error of approximation  $<0.06$  are considered good<sup>69</sup>. In addition, we reported  $R^2$  for the endogenous variables naturalization incidence and naturalization extent. We did the path analysis using the 'lavaan' package<sup>70</sup> in R. The results remained qualitatively similar when we used one average index per species (Supplementary Fig. 5).

**Data availability.** Data for this article, including four selfing-ability indices and information on data sources, species name, life history, native range and global naturalization as assessed based on the Global Naturalized Alien Flora database may be found in the Supplementary Information (Supplementary Data 1). Fruit set and seed production data in pollinator exclusion, self-pollination and cross-pollination treatments are available from the corresponding author upon request.

## References

- Lewis, S. L. & Maslin, M. A. Defining the anthropocene. *Nature* **519**, 171–180 (2015).
- van Kleunen, M. *et al.* Global exchange and accumulation of non-native plants. *Nature* **525**, 100–103 (2015).
- Reichard, S. H. & White, P. Horticulture as a pathway of invasive plant introductions in the United States. *Bioscience* **51**, 103–113 (2001).
- Hulme, P. E. Addressing the threat to biodiversity from botanic gardens. *Trends Ecol. Evol.* **26**, 168–174 (2011).
- Hulme, P. E. *et al.* Grasping at the routes of biological invasions: a framework for integrating pathways into policy. *J. Appl. Ecol.* **45**, 403–414 (2008).
- Blackburn, T. M. *et al.* A unified classification of alien species based on the magnitude of their environmental impacts. *PLoS Biol.* **12**, e1001850 (2014).
- Richardson, D. M. *et al.* Naturalization and invasion of alien plants: concepts and definitions. *Divers. Distrib.* **6**, 93–107 (2000).
- Blackburn, T. M. *et al.* A proposed unified framework for biological invasions. *Trends Ecol. Evol.* **26**, 333–339 (2011).
- Richardson, D. M. & Pyšek, P. Naturalization of introduced plants: ecological drivers of biogeographical patterns. *New Phytol.* **196**, 383–396 (2012).
- Barrett, S. C. Why reproductive systems matter for the invasion biology of plants? in *Fifty Years of Invasion Ecology: the Legacy of Charles Elton* (ed. Richardson, D. M.) 195–210 (John Wiley & Sons, 2011).
- Burns, J. H. *et al.* Greater sexual reproduction contributes to differences in demography of invasive plants and their noninvasive relatives. *Ecology* **94**, 995–1004 (2013).
- Baker, H. G. Self-compatibility and establishment after long distance dispersal. *Evolution* **9**, 347–349 (1955).
- Stebbins, G. L. Self fertilization and population variability in the higher plants. *Am. Nat.* **91**, 337–354 (1957).
- Pannell, J. R. *et al.* The scope of Baker's law. *New Phytol.* **208**, 656–667 (2015).
- Burns, J. H., Ashman, T.-L., Steets, J. A., Harmon-Threatt, A. & Knight, T. M. A phylogenetically controlled analysis of the roles of reproductive traits in plant invasions. *Oecologia* **166**, 1009–1017 (2011).
- Rambuda, T. D. & Johnson, S. D. Breeding systems of invasive alien plants in South Africa: does Baker's rule apply? *Divers. Distrib.* **10**, 409–416 (2004).
- van Kleunen, M. & Johnson, S. D. Effects of self-compatibility on the distribution range of invasive European plants in North America. *Conserv. Biol.* **21**, 1537–1544 (2007).
- Hao, J. H., Qiang, S., Chrobok, T., van Kleunen, M. & Liu, Q. Q. A test of Baker's law: breeding systems of invasive species of Asteraceae in China. *Biol. Invasions* **13**, 571–580 (2011).
- Lafuma, L. & Maurice, S. Increase in mate availability without loss of self-incompatibility in the invasive species *Senecio inaequidens* (Asteraceae). *Oikos* **116**, 201–208 (2007).
- Sutherland, S. What makes a weed a weed: life history traits of native and exotic plants in the USA. *Oecologia* **141**, 24–39 (2004).
- Moodley, D., Geerts, S., Richardson, D. M. & Wilson, J. R. The importance of pollinators and autonomous self-fertilisation in the early stages of plant invasions: *Banksia* and *Hakea* (Proteaceae) as case studies. *Plant Biol.* **18**, 124–131 (2015).
- Lloyd, D. G. & Schoen, D. J. Self-and cross-fertilization in plants. I. Functional dimensions. *Int. J. Plant Sci.* **153**, 358–369 (1992).
- Bawa, K. S. Breeding systems of tree species of a lowland tropical community. *Evolution* **28**, 85–92 (1974).
- Mena-Ali, J. I. & Stephenson, A. G. Segregation analyses of partial self-incompatibility in self and cross progeny of *Solanum carolinense* reveal a leaky S-allele. *Genetics* **177**, 501–510 (2007).
- Raduski, A. R., Haney, E. B. & Igić, B. The expression of self-incompatibility in angiosperms is bimodal. *Evolution* **66**, 1275–1283 (2012).
- Baker, H. G. *Characteristics and Modes of Origin of Weeds in The Genetics of Colonizing Species*. (eds Baker, H. G. & Stebbins, G. L.) 147–168 (Academic Press, 1965).
- Baker, H. G. The evolution of weeds. *Annu. Rev. Ecol. Syst.* **5**, 1–24 (1974).
- Pannell, J. R. & Barrett, S. C. Baker's law revisited: reproductive assurance in a metapopulation. *Evolution* **52**, 657–668 (1998).
- Pyšek, P. *et al.* Temperate trees and shrubs as global invaders: the relationship between invasiveness and native distribution depends on biological traits. *Biol. Invasions* **16**, 577–589 (2014).
- Pyšek, P. *et al.* Naturalization of central European plants in North America: species traits, habitats, propagule pressure, residence time. *Ecology* **96**, 762–774 (2015).
- Grossenbacher, D., Briscoe Runquist, R., Goldberg, E. E. & Brandvain, Y. Geographic range size is predicted by plant mating system. *Ecol. Lett.* **18**, 706–713 (2015).
- Randle, A. M., Slyder, J. B. & Kalisz, S. Can differences in autonomous selfing ability explain differences in range size among sister-taxa pairs of *Collinsia* (Plantaginaceae)? An extension of Baker's Law. *New Phytol.* **183**, 618–629 (2009).

33. Pyšek, P. *et al.* Successful invaders co-opt pollinators of native flora and accumulate insect pollinators with increasing residence time. *Ecol. Monogr.* **81**, 277–293 (2011).
34. Felsenstein, J. Phylogenies and the comparative method. *Am. Nat.* **125**, 1–15 (1985).
35. Eckert, C. G. *et al.* Plant mating systems in a changing world. *Trends Ecol. Evol.* **25**, 35–43 (2010).
36. Dehnen-Schmutz, K., Touza, J., Perrings, C. & Williamson, M. The horticultural trade and ornamental plant invasions in Britain. *Conserv. Biol.* **21**, 224–231 (2007).
37. van Kleunen, M., Manning, J. C., Pasqualetto, V. & Johnson, S. D. Phylogenetically independent associations between autonomous self-fertilization and plant invasiveness. *Am. Nat.* **171**, 195–201 (2008).
38. Busch, J. W. & Delph, L. F. The relative importance of reproductive assurance and automatic selection as hypotheses for the evolution of self-fertilization. *Ann. Bot.* **109**, 553–562 (2012).
39. Colautti, R. I., Grigorovich, I. A. & MacIsaac, H. J. Propagule pressure: a null model for biological invasions. *Biol. Invasions* **9**, 885–885 (2007).
40. Goodwillie, C., Kalisz, S. & Eckert, C. G. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annu. Rev. Ecol. Syst.* **36**, 47–79 (2005).
41. Kalisz, S., Vogler, D. W. & Hanley, K. M. Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. *Nature* **430**, 884–887 (2004).
42. Razanajatovo, M., Föhr, C., Fischer, M., Prati, D. & van Kleunen, M. Non-naturalized alien plants receive fewer flower visits than naturalized and native plants in a Swiss botanical garden. *Biol. Conserv.* **182**, 109–116 (2015).
43. Charlesworth, D. & Charlesworth, B. Inbreeding depression and its evolutionary consequences. *Annu. Rev. Ecol. Syst.* **18**, 237–268 (1987).
44. Rodger, J. G. & Johnson, S. D. Self-pollination and inbreeding depression in *Acacia dealbata*: can selfing promote invasion in trees? *S. Afr. J. Bot.* **88**, 252–259 (2013).
45. Rodger, J. G., van Kleunen, M. & Johnson, S. D. Does specialized pollination impede plant invasions? *Int. J. Plant Sci.* **171**, 382–391 (2010).
46. Petanidou, T. *et al.* Self-compatibility and plant invasiveness: comparing species in native and invasive ranges. *Perspect. Plant Ecol. Evol. Syst.* **14**, 3–12 (2012).
47. Lambdon, P. W., Lloret, F. & Hulme, P. E. How do introduction characteristics influence the invasion success of Mediterranean alien plants? *Perspect. Plant Ecol. Evol. Syst.* **10**, 143–159 (2008).
48. Pyšek, P., Jarošík, V. & Pergl, J. Alien plants introduced by different pathways differ in invasion success: unintentional introductions as a threat to natural areas. *PLoS ONE* **6**, e24890 (2011).
49. van Kleunen, M., Dawson, W. & Maurel, N. Characteristics of successful alien plants. *Mol. Ecol.* **24**, 1954–1968 (2015).
50. Bucharová, A. & van Kleunen, M. Introduction history and species characteristics partly explain naturalization success of North American woody species in Europe. *J. Ecol.* **97**, 230–238 (2009).
51. Rejmánek, M. A theory of seed plant invasiveness: the first sketch. *Biol. Conserv.* **78**, 171–181 (1996).
52. Pyšek, P. *et al.* The global invasion success of Central European plants is related to distribution characteristics in their native range and species traits. *Divers. Distrib.* **15**, 891–903 (2009).
53. Moles, A. T. & Westoby, M. Seed size and plant strategy across the whole life cycle. *Oikos* **113**, 91–105 (2006).
54. Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with ImageJ. *Biophoton. Int.* **11**, 36–42 (2004).
55. Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F. S. & Golicher, D. J. Taxonstand: an R package for species names standardisation in vegetation databases. *Methods Ecol. Evol.* **3**, 1078–1083 (2012).
56. R Core Team. *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2012).
57. Brummitt, R. K., Pando, F., Hollis, S. & Brummitt, N. *World Geographical Scheme for Recording Plant Distributions* (International Working Group on Taxonomic Databases for Plant Sciences (TDWG), 2001).
58. Zanne, A. E. *et al.* Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).
59. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
60. Webb, C. O., Ackerly, D. D. & Kembel, S. W. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* **24**, 2098–2100 (2008).
61. Wikström, N., Savolainen, V. & Chase, M. W. Evolution of the angiosperms: calibrating the family tree. *Proc. R. Soc. Lond. B Biol. Sci.* **268**, 2211–2220 (2001).
62. Felsenstein, J. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. *Am. Nat.* **171**, 713–725 (2008).
63. Ives, A. R. & Garland, T. Phylogenetic logistic regression for binary dependent variables. *Syst. Biol.* **59**, 9–26 (2010).
64. Ho, L. S. T. & Ane, C. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.* **63**, 397–408 (2014).
65. Schielzeth, H. Simple means to improve the interpretability of regression coefficients. *Methods Ecol. Evol.* **1**, 103–113 (2010).
66. Nakagawa, S. & Schielzeth, H. A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4**, 133–142 (2013).
67. Freckleton, R., Harvey, P. & Pagel, M. Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* **160**, 712–726 (2002).
68. Grace, J. B. *Structural Equation Modeling and Natural Systems* (Cambridge University Press, 2006).
69. Hu, L. t. & Bentler, P. M. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Modeling* **6**, 1–55 (1999).
70. Rosseel, Y. lavaan: an R package for structural equation modeling. *J. Stat. Softw.* **48**, 1–36 (2012).

### Acknowledgements

We thank A. Raduski and collaborators for sharing their database on self-incompatibility; E. Rehn for acquiring PDF files of many studies; B. Hodapp and D. Wagner for their help in data extraction; the International Max Planck Research School for Organismal Biology and the Equal Opportunity Council of the University of Konstanz for supporting M.R.; and the DFG (Projects KL1866/3-1 and KL1866/9-1) for funding. P.P. and J.P. were supported by long-term research development project RVO 67985939 (The Czech Academy of Sciences), 504/11/1028 (GA CR) and project no. 14-36079G, Centre of Excellence PLADIAS (Czech Science Foundation). P.P. appreciates the support by Praemium Academiae award from The Czech Academy of Sciences. H.K. and P.W. acknowledge funding in the scope of the BEFmate project from the Ministry of Science and Culture of Lower Saxony. M.W. acknowledges the support of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, funded by the German Research Foundation (FZT 118).

### Author contributions

M.R. and M.v.K. compiled the global breeding-system database. W.D., F.E., H.K., J.P., P.P., P.W., M.W. and M.v.K. compiled the Global Naturalized Alien Flora database. M.R. and M.v.K. developed the research questions and analysed the data. N.M. contributed to the data analyses. M.R. wrote the manuscript with input from M.v.K., N.M., W.D., F.E., H.K., J.P., P.P., P.W. and M.W.

### Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Razanajatovo, M. *et al.* Plants capable of selfing are more likely to become naturalized. *Nat. Commun.* **7**, 13313 doi: 10.1038/ncomms13313 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016