

Gaps in global alien plant trait data and how to fill them

Authors

Matthias Grenié ^{1,2,3} ORCID: [0000-0002-4659-7522](https://orcid.org/0000-0002-4659-7522)

email: matthias.grenie@univ-grenoble-alpes.fr

Helge Bruehlheide ^{3,4} ORCID: [0000-0003-3135-0356](https://orcid.org/0000-0003-3135-0356)

Wayne Dawson ⁵ ORCID: [0000-0003-3402-0774](https://orcid.org/0000-0003-3402-0774)

Franz Essl ⁶ ORCID: [0000-0001-8253-2112](https://orcid.org/0000-0001-8253-2112)

Mark van Kleunen ^{7,8} ORCID: [0000-0002-2861-3701](https://orcid.org/0000-0002-2861-3701)

Ingolf Kühn ^{9,4,3} ORCID: [0000-0003-1691-8249](https://orcid.org/0000-0003-1691-8249)

Holger Kreft ^{10,11,12} ORCID: [0000-0003-4471-8236](https://orcid.org/0000-0003-4471-8236)

Petr Pyšek ^{13,14} ORCID: [0000-0001-8500-442X](https://orcid.org/0000-0001-8500-442X)

Patrick Weigelt ^{10,15} ORCID: [0000-0002-2485-3708](https://orcid.org/0000-0002-2485-3708)

Marten Winter ^{2,3} ORCID: [0000-0002-9593-7300](https://orcid.org/0000-0002-9593-7300)

Affiliations

¹ Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, 38000 Grenoble, France

² Leipzig University, Ritterstraße 26, 04109 Leipzig, Germany

³ German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig,
Puschstraße 4, 04103 Leipzig, Germany

⁴ Institute of Biology/Geobotany and Botanical Garden, Martin Luther University
Halle-Wittenberg, Halle, Germany

⁵ Department of Evolution, Ecology and Behaviour, Institute of Infection, Veterinary and
Ecological Sciences, University of Liverpool, Liverpool, L69 7ZB, UK

⁶ Division of BioInvasions, Global Change & Macroecology, Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria

⁷ Ecology, Department of Biology, University of Konstanz, Universitätsstrasse 10, D-78457 Konstanz, Germany

⁸ Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, Taizhou University, Taizhou 318000, China.

⁹ Dept. Community Ecology, Helmholtz Centre for Environmental Research - UFZ, Theodor-Lieser-Str. 4, 06120 Halle, Germany

¹⁰ Biodiversity, Macroecology & Biogeography, University of Göttingen, Göttingen, Germany

¹¹ Centre of Biodiversity and Sustainable Land Use, University of Göttingen, Göttingen, Germany

¹² Campus Institute Data Science, Göttingen, Germany

¹³ Czech Academy of Sciences, Institute of Botany, Department of Invasion Ecology, CZ-25243 Průhonice, Czech Republic

¹⁴ Department of Ecology, Faculty of Science, Charles University, Viničná 7, CZ-12844 Prague, Czech Republic

¹⁵ Department of Environmental Science, Radboud Institute for Biological and Environmental Sciences, Radboud University, Heyendaalseweg 135, 6525AJ Nijmegen, The Netherlands

Acknowledgments

M.G., H.B., IK and M.W. gratefully acknowledge the support of iDiv funded by the German Research Foundation (DFG– FZT 118, 202548816), specifically funding through sDiv, the Synthesis Centre of iDiv. PP was supported by EXPRO grant no. 19-28807X (Czech Science Foundation) and long-term research development project RVO 67985939 (Czech Academy of Sciences). MvK acknowledges support of the German Research Foundation DFG (grant

264740629). FE appreciates funding by the Austrian Science Fund (FWF; pr.no. I 5825-B). HK acknowledges funding by German Research Foundation DFG (Research unit FOR2716 DynaCom).

Conflict of Interest

There is no conflict of interest to declare.

Author Contributions

M.G. and M.W. conceived the ideas and designed methodology; M.G. collected and analyzed the data; M.G. and M.W. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Title

Gaps in alien plant trait data and how to fill them

Running title

Global gaps in alien plant trait data

Abstract

Aim

Functional traits help to understand the ecological processes underlying biological invasions. The extent to which trait data are available for alien plants at the global scale is unknown. In this study, we assess the availability of trait data and identify global gaps and biases

Location

Global

Time Period

Present

Major taxa studied

Vascular plants

Methods

We used the GloNAF database to get a global list of plants naturalized outside their native range and their distributions. We combined data from the four largest trait databases: AusTraits, BIEN, GIFT, and TRY, on which we performed taxonomic and trait harmonization. We studied the availability of trait data. Then, based on the distribution data, we tested to what extent trait knowledge was driven by ecological and socioeconomic variables.

Results

We found that the species-by-trait matrix (16,044 species for 2,764 traits) was only 1.4% filled, with most traits measured on very few species. Only ten traits were available for more than 50% of all alien plants. Four percent of the species had not a single trait measured, while 27% of species had data for the three key plant traits of leaf mass per area, seed mass, and plant height. We observed a strong latitudinal gradient in trait knowledge with areas in the Tropics showing lower trait knowledge than higher latitudes. Growth form, range size, and invasion status were the strongest predictors of trait knowledge.

Main conclusions

We identified large trait data gaps at a global scale for alien plants. These gaps are partly driven by uneven sampling and a lack of trait data integration across sources. We recommend that prioritizing the most invasion-relevant traits and distributing community efforts of plant and invasion scientists to sample them in a standardized way can help close these gaps.

Keywords

functional traits; naturalized plants; non-native plants; missing trait; raunkiaerian shortfall; trait database

Introduction

Biological invasions are one of the most important drivers of biodiversity loss and we need a better understanding of the underlying ecological processes for effective management (Díaz *et al.*, 2019; Roy *et al.*, 2023). Global invasion of plants led to approximately four percent of the global flora being established outside of its native range (Küster *et al.*, 2008; e.g., Pyšek *et al.*, 2008; Drenovsky *et al.*, 2012; van Kleunen *et al.*, 2015; Gallien & Carboni, 2017), becoming alien (i.e. non-native) somewhere on this planet. Identifying plant traits promoting successful species introductions, establishment and dispersal is one of the major aims of plant invasion ecology (Küster *et al.*, 2008; e.g., Pyšek *et al.*, 2008; Drenovsky *et al.*, 2012; van Kleunen *et al.*, 2015; Gallien & Carboni, 2017). With the knowledge of species traits the management and prediction of population dynamics of established or future alien species can be much more effective. This includes biological invasions: for example, species that are able to self-fertilize are more likely to become naturalized than outcrossing species (Razanajatovo *et al.*, 2016).

Despite massive collaborative mobilization efforts, large trait data gaps remain for the global vascular plant flora (Kattge *et al.*, 2020; Visscher *et al.*, 2022). Recent efforts have identified and systematically described such gaps (Hortal *et al.*, 2015; Cornwell *et al.*, 2019; Rudbeck *et al.*, 2022; Maitner *et al.*, 2023). For example, we observe a systematic lack of trait data for species occurring in less studied regions such as in the Tropics or Africa, while species in temperate regions have been studied more intensely (Hortal *et al.*, 2015). Geographical accessibility is also known to drive data availability (Oliveira *et al.*, 2016), which means that we have greater trait knowledge of accessible regions than less accessible ones. In addition, species in economically wealthier areas show greater occurrence data coverage than species in comparatively poorer areas (Meyer *et al.*, 2016).

To fully embrace context-dependency of traits in biological invasions, traits should be considered in conjunctions to use them to manage alien species (Küster *et al.*, 2008; Pyšek *et*

al., 2020). Thus, trait availability of alien species is key to using traits to explain idiosyncrasies associated with different invasion syndromes (Novoa *et al.*, 2020) and across environmental gradients (Golivets *et al.*, 2024).

There is no assessment of the availability of trait data for alien plant species. To better understand potential biases or vice versa to identify collection priorities in trait data of alien plants, we need to identify the factors determining the trait knowledge of alien plants, spatially and taxonomically. Given the previous studies on biases in plant trait data (Cornwell *et al.*, 2019; Gallagher *et al.*, 2020; Kattge *et al.*, 2020; Maitner *et al.*, 2023), we know that species with larger ranges have greater data availability than species with smaller ranges, because they are more likely to be sampled. In addition, we expect that species occurring in wealthier countries (in either their native or naturalized range) have more trait data present in databases, than species in poorer countries. We also expect invasive (i.e. species with ecological and/or economic negative impacts) species to have data available for more traits than naturalized species that were never recorded as invasive, because the impacts of invasiveness should incentivise trait research on invasive plants.

Scientists rarely leverage multiple trait databases, which would likely improve trait coverage (Feng *et al.*, 2022). This is because plant trait databases are often not directly interoperable and do not follow commonly defined standards. As a result, most trait studies do not use the full range of available trait data (Feng *et al.*, 2022). The lack of shared trait definitions impedes the reuse of plant traits across databases (Garnier *et al.*, 2017), while combination of traits coming from different databases requires care and specific expertise (Keller *et al.*, 2023), but ensures a greater trait coverage and reduces the potential taxonomic and spatial biases (Maitner *et al.*, 2023).

Here we map and quantify the trait knowledge for alien plants at global scale, using the largest trait dataset for non-native species assembled to date, by combining four major plant trait databases. We then assess the main ecological and socioeconomic factors driving the data

availability of plant traits.

Methods

Alien plants list

We extracted a list of alien vascular plants from the Global Naturalized Alien Flora (GloNAF) database v.2.0 (van Kleunen *et al.*, 2019) that is being continuously updated by the database creators (accessed 2024-05-30). We extracted the list of species in GloNAF that were labeled as “naturalized” or “invasive”. Our list of alien species comprised 16,044 taxonomic names of vascular plants before taxonomic harmonization. We also extracted the naturalized geographic distribution (at TDWG4 resolution level) of each taxon from the GloNAF database.

Taxonomic harmonization

We gathered trait data from four different plant trait databases: AusTraits (Falster *et al.*, 2021), BIEN (Enquist *et al.*, 2016), GIFT (Weigelt *et al.*, 2020), and TRY (Kattge *et al.*, 2020). We used these databases because they represent the largest and most accessible trait databases for plants globally (Feng *et al.*, 2022).

As the different databases used different taxonomic backbones to standardize species names, we performed taxonomic harmonization (Grenié *et al.*, 2022) by merging the accepted binomial species names between GloNAF and each trait database. In the end, we had 13,705 matched species names between GloNAF and all trait databases (see [Supplementary Information 1](#) for details).

Trait data

For each trait database, we counted the number of measured traits for each species. Some traits may have more than one observation. We however decided to be the least conservative for our analyses: we considered as “measured” a trait that was at least measured once across all databases. We did so as any single measured trait gives greater alien species trait knowledge than no observation.

AusTraits. We extracted all traits available for species referenced in GloNAF from AusTraits version 6.0 (Falster *et al.*, 2021). We obtained data for 33,494 taxa (including infraspecific ones) and 497 traits.

BIEN. We queried all traits available in BIEN through the BIEN R package (Maitner *et al.*, 2018). We used BIEN version 4.2.6 (released 2022-08-09, Enquist *et al.*, 2016). We obtained data for 109,394 species and 52 traits.

GIFT. We used the GIFT database (Weigelt *et al.*, 2020) as it offers complimentary traits from global databases and notably contains the growth form for most plant species. We used GIFT version 3.1, including both public and private records through the GIFT R package (Denelle *et al.*, 2023). We obtained data for 287,229 species and 106 traits.

TRY. We queried all publicly available traits in TRY v6.0 (Kattge *et al.*, 2020). We obtained data for 301,799 species and 2,460 traits.

Aligning common trait definitions. We created a single species-by-trait matrix from all trait databases after harmonizing the traits across them (see details in [Supplementary Information S1](#)), to make correspondence tables for all possible pairwise database combinations we leveraged the Australian Plant Trait Dictionary (APD) v2.0.0 (Wenk *et al.*, 2024), which provides trait correspondence between AusTraits and all three other databases we used.

Final trait dataset. We created three distinct trait datasets based on how stringent we were to consider traits similar in their definition across our correspondence tables. In the first option (“full” trait network), we considered all traits that were exactly matching, close, or related as being the same. The second option (“close” trait network) considered only traits that were exactly or closely matching. The final option, the most stringent one (“exact” trait network), considered two traits the same only if they were exactly matching. For example, in our network Austraits “Leaf lamina mass per area” trait ([APD:0011231](#)), was considered exactly matching TRY SLA trait with petiole excluded (TRY:3115), closely matching with TRY SLA trait with petiole, midrib and rachis excluded (TRY:3086), and related to two other TRY SLA traits where

petiole included (TRY:3116) and where it is undefined if petiole was or not included (TRY:3117). In the “full” network, all these traits would be lumped together, while in the “close” network the trait from AusTraits would be connected to TRY:3115 and TRY:3086, in the more stringent “exact” network, only TRY:3115 would be connected to the leaf lamina mass per area trait from AusTraits. We provide the “full” network in [Table S1](#).

We performed our analyses with all three versions of the trait networks but present only the “full” option hereafter as the results were quantitatively and qualitatively similar across all versions. Our trait name network initially contained 3,351 unique trait names across databases and 804 links between exact, close, and related matches of traits. Using our correspondence tables, considering the “full” trait network, we obtained 2,764 unique traits. In the end, in the “full” trait network, our combined trait dataset contained 13,705 species (after taxonomic harmonization) and 2,243 observed traits as 521 traits were never observed across our set of target species.

Trait combinations

Because one can't measure all traits for all species to fully describe phenotypes, researchers identified generic trait combinations reflecting major ecological trade-offs to compare as many species as possible (Westoby, 1998; Díaz *et al.*, 2016; Bergmann *et al.*, 2020). We focused on three ecological trait trade-offs: the Leaf-Height-Seed Mass (Westoby, 1998), the global spectrum of plant form and function (Díaz *et al.*, 2016; aboveground spectrum traits hereafter), and the root economics space (Bergmann *et al.*, 2020). See [Supplementary Information S2](#) for list of traits and extended justification.

Modeling trait knowledge

We tested to what extent the determinants of other shortfalls of biodiversity (Hortal *et al.*, 2015; Rudbeck *et al.*, 2022) correlated with the number of measured traits per species (our response variable). We extracted for each GloNAF region of alien plant species occurrence

several predictors provided in the GIFT database: the average gross domestic product per capita (GDPpc) from 2015 (Kummu *et al.*, 2018), the mean access time from major cities (Weiss *et al.*, 2018), and the Human Influence Index (WCS & CIESIN, 2005), which aggregates and averages disparate sources of anthropogenization (density of roads, density of population, land-use, etc.) per region. We computed the average of all predictor variables across the entire range for each species. For GDPpc, because we hypothesized that species occurring in wealthier countries in their non-native range and/or their native range would have more traits measured, we computed two GDPpc, one across the native range of the species, the other across its non-native range. As species with larger ranges are more likely to have more traits measured, especially larger native ranges, we considered separately the number of regions where a species is native and the number of regions where it is non-native. We assumed that species occurring in more diverse habitats have a higher chance of being sampled, as they are more likely of occurring in a well sampled environment, we thus counted the number of biomes a species occurs in from Dinerstein *et al.* (2017). We also included the simplified growth form of the species (tree, shrub, herb, or other) extracted from GIFT as a predictor variable, available for all species.

Final data subset. We only kept species for which all predictors were known for at least 80% of the regions they occur in as naturalized species. This led to a total of 13,253 species being included in the analyses presented here.

Statistical Model. Our response variable was the number of traits measured per species out of our theoretical maximum of 2,764 traits; we thus performed a negative binomial generalized linear model that we fit using the `glm.nb()` function in the MASS package (Venables & Ripley, 2002). We used the nine above-mentioned predictor variables: species growth form, species total range size, species non-native range size, the number of biomes a species occurs in, the average Human Influence Index across its entire range, the standard deviation of Human Influence Index across its entire range, the average GDP per capita across its native range, the

average GDP per capita across its non-native range, and the average accessibility across its entire range. All predictor variables were centered to 0 mean and scaled to 1 standard deviation prior to the analysis. All predictors showed low multicollinearity with variance inflation factors all < 5 . The checks using the `check_model()` function of the performance package (Lüdecke *et al.*, 2021) showed normal residuals and no evidence for overdispersion nor zero-inflation. We used Nagelkerke's pseudo-R-square for GLMs through the `r2_nagelkerke()` function from the performance package. All data extraction and analyses were done using R 4.2.2 (R Core Team, 2022).

Results

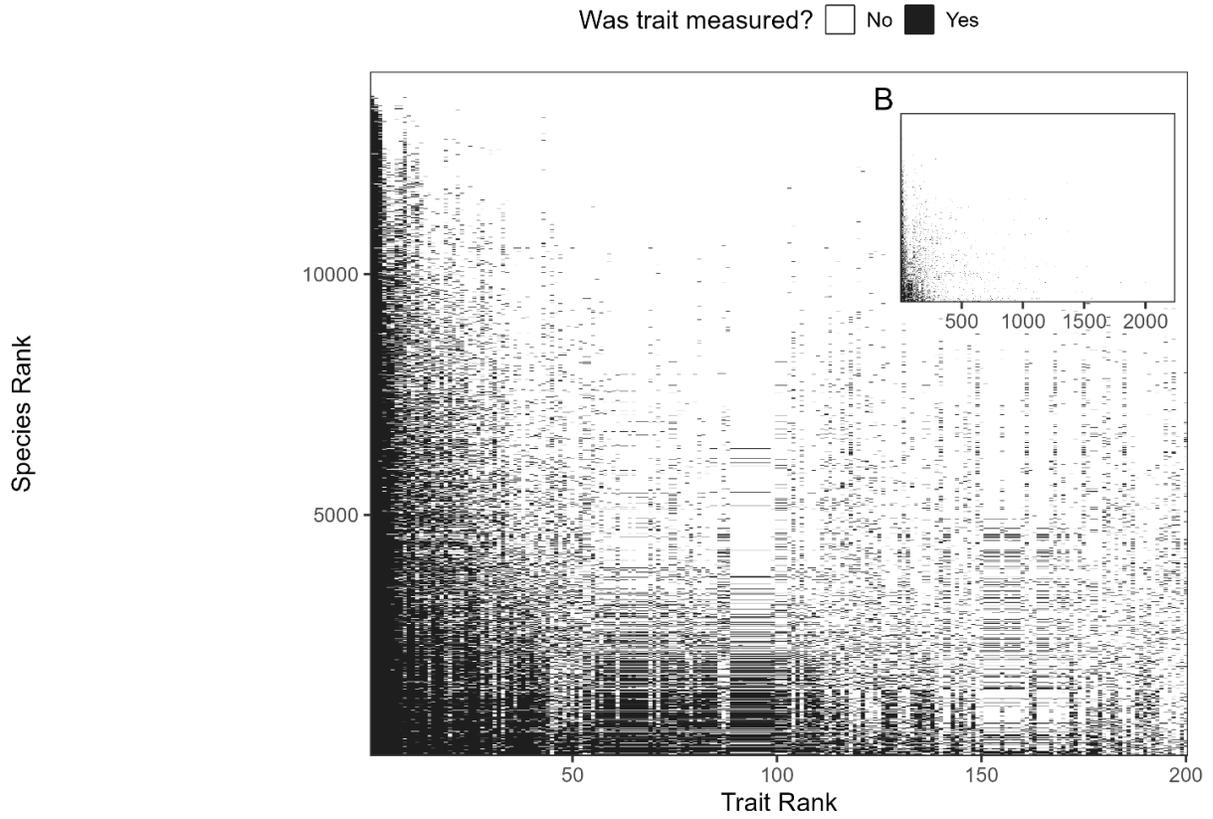
A sparse species-traits matrix

We obtained a species-traits table containing observations for 13,705 alien vascular plant species (out of 16,044 known at global scale) and 2,243 different traits (out of 2,764) ([Figure 1A](#)). This gives a theoretical maximum of 48,646,400 possible species by trait combinations, while our observed matrix only contained 691,653 (1.4%) of them. Narrowing on the 200 most frequently measured traits ([Figure 1B](#)) gives 3,208,800 combinations of which 501,736 (15.6%) are observed.

We listed 25 most frequently measured traits. About 90% of alien plant species had data available for each of the following three traits: growth form, woodiness, and life history ([Figure 1C](#)). Around 70% of the alien species had two other traits measured at least once: plant height, which is fundamental to understanding the ecology of species; and leaf type, distinguishing between a broad leaf or a needle. Each of four other traits were available for more than half of the species: leaf compoundness (compound or simple leaf), dispersal syndrome (anemochorous, etc.), seed mass, and photosynthetic pathway (C3, C4, or other). The remaining 16 traits are available for less than half of the species. Most of those traits describe fundamental ecological characteristics: flowering phenology, nitrogen fixation, fruit type, leaflet number per leaf, leaf length, leaf phenology, species tolerance to frost, leaf phyllotaxis, leaf width, pollination syndrome, fruit length, plant sex type, flower color, leaf mass per area, leaf area, seed germination rate. Among the 25 most frequently measured traits, the different databases had various unique contributions ([Figure S1](#)). For some traits, like leaf type (available for 9,528 out of 10,162 species), TRY covered 6,121 species, while for plant height. For leaf type, GIFT was the greatest provider of unique species-trait observations unavailable in other databases (2,564 species out of 10,467). The contribution of AusTraits for leaf type, though smaller, mostly brought unique observations (398 species unobserved in other databases).

Species-traits observations in BIEN were generally also available in TRY, but BIEN had some unique contributions for certain traits, for example for pollination syndrome (113 species uniquely in BIEN versus 525 in AusTraits, 989 in GIFT, and 1080 in TRY for a total of 5,460 species).

A



C

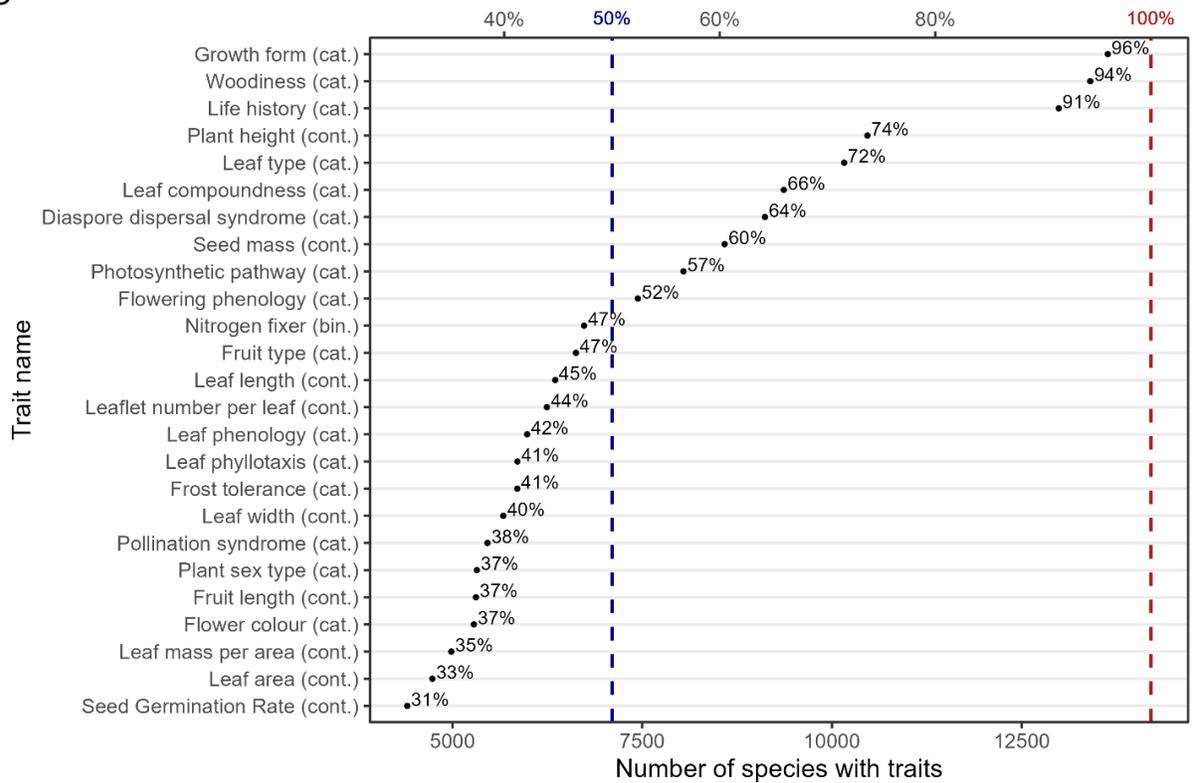


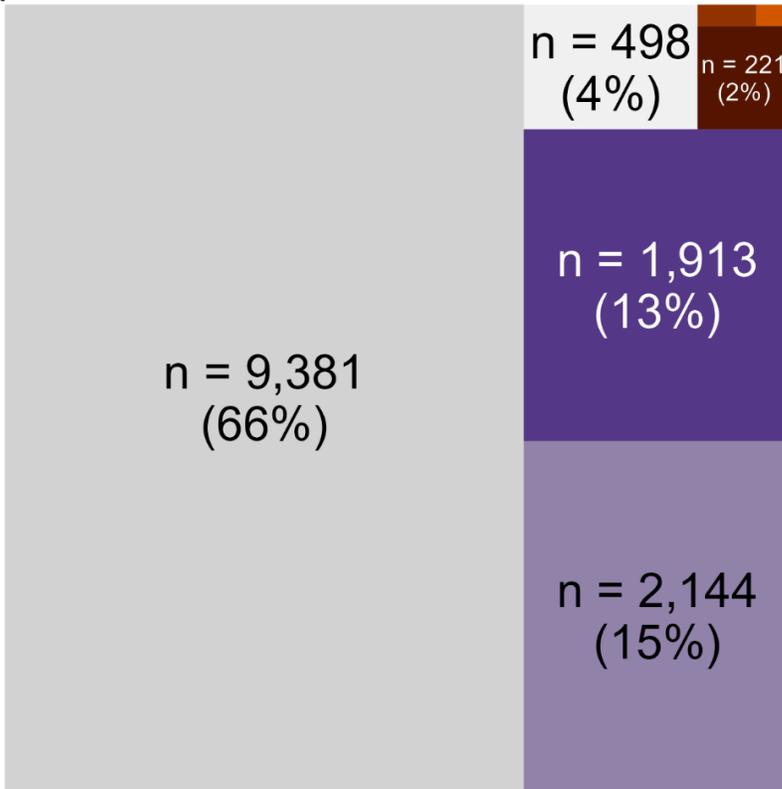
Figure 1. (A) Species-by-trait matrix for alien vascular plants of the 200 most measured traits. Each pixel represents the measurement of one trait for one species. Traits are ordered on the x-axis from most to least measured. Species are ordered in columns ordered from most to least measured (bottom to top). The color of the pixel shows whether the trait was ever measured for this species (dark gray) or never (white). **(B)** shows the same figure with all 2,273 measured traits. **(C)** Proportion of species covered by the 25 most frequently measured traits. The x-axis shows the number of species with the given trait measured (the top x-axis gives the corresponding proportion). The y-axis shows the different traits with their names as labels, the labels indicate whether the traits are continuous (cont.), categorical (cat.), or binary (bin.). The points are the proportions of alien species with at least one trait value for the trait indicated on the y-axis. The proportions are displayed above the points. The red dashed line represents 100% cover of the alien species (16,044 in our dataset), while the blue dashed line represents 50%.

Uneven proportion of measured combination of traits

We mapped species for which we had the measured combinations of traits (see details in [Table S2-1](#)): Leaf-Height-Seed Mass (LHS), aboveground spectrum traits, and root traits. For only 4% of alien species not a single trait measurement exists (498 species, [Figure 2A](#)), which means that 96% of alien plant species have at least one trait measured in our consolidated dataset. However, to compare species or compute functional diversity metrics, the traits measured across species need to be the same. Only 26.8% of the species with non-native occurrences have a measured LHS combination (4,307 species), only 13.3% have the aboveground spectrum traits fully described (2,134 species), and root traits are measured for only 1.7% (267 species). When jointly considering multiple combinations, the numbers drop further: 1.6% of species (250 species) have measured LHS and root traits, while 1.4% of species (221 species) have both the full aboveground spectrum and root combinations measured.

The available trait combinations show strong taxonomic biases ([Figure 2B](#)). While we expect families with a larger number of alien species to show a higher proportion of trait combinations available, smaller families like Myrtaceae have trait data for most of their alien species, while the bigger families like Poaceae show lacking trait combinations for most alien species.

A



Trait combination



B

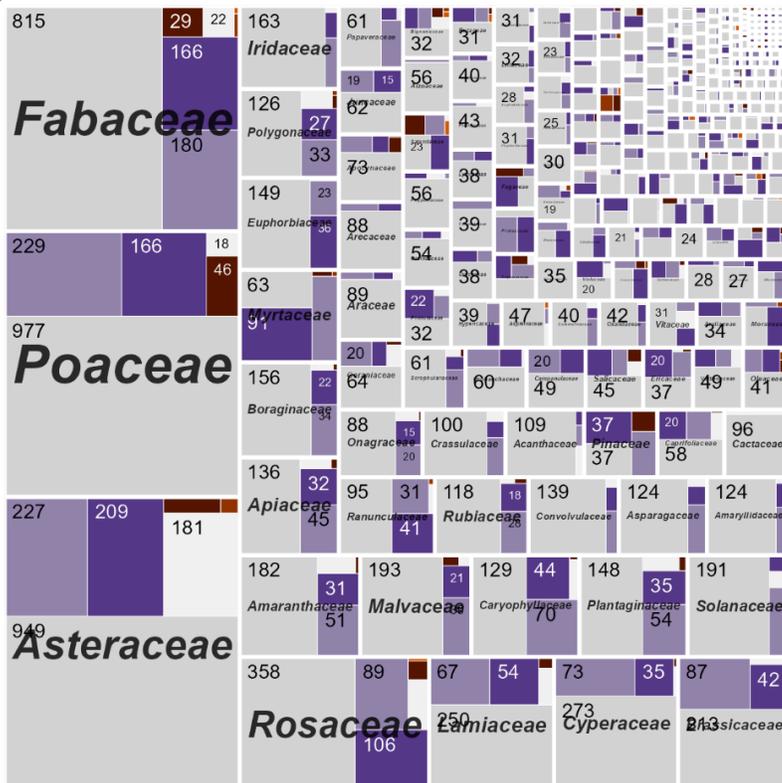
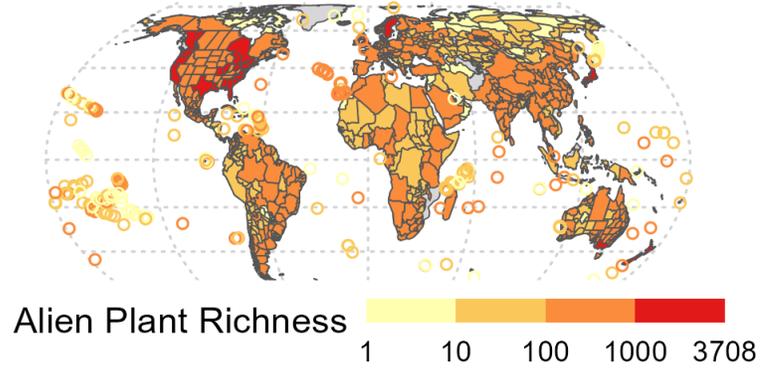


Figure 2. Shares of alien vascular plant species covered by different trait combinations. The area of rectangles is directly proportional to the number of species in each category. **(A)** Organized by trait combinations. The numbers show the number of species and respective proportion of species compared to the global number of alien species (16,044). **(B)** Organized by families and trait combinations. The blocks contained within thick white borders are botanical families. Families are ordered by decreasing number of alien species (e.g., Asteraceae has the most aliens). The numbers depicted are the number of species of the family with the given trait combination. Areas are colored depending on measured trait combinations. The colors correspond to different trait combinations: light gray, no measured trait; gray, at least one measured trait; light purple, LHS; dark purple, aboveground spectrum traits; bright orange; root traits; darker orange, LHS and root traits; brown, aboveground spectrum and root traits.

A



B

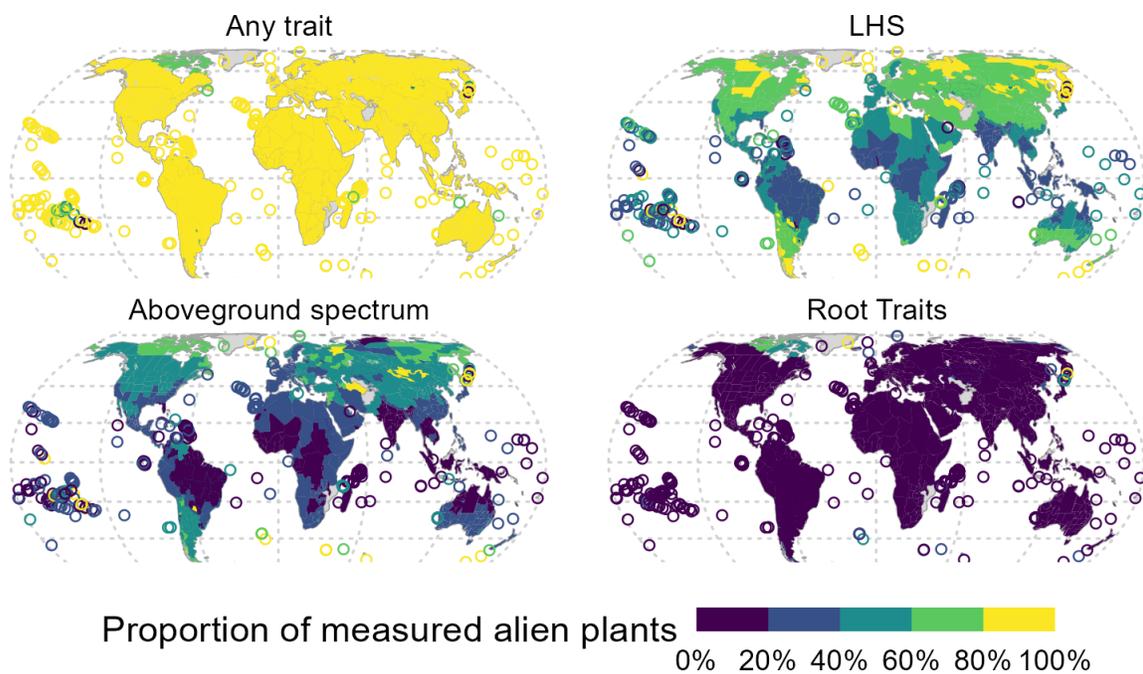


Figure 3. Maps of alien vascular plant species richness and proportion of aliens with measured trait combinations. **(A)** Alien species richness based on GloNAF. For readability reasons the richness scale has been discretized in four colors following a log₁₀ scale. Gray areas show where no data were available. **(B)** Proportions of alien species in each region with measured trait combination (at least one trait; Leaf-Height-Seed Mass, LHS; aboveground spectrum; and root traits). Empty circles represent islands and are not scaled for readability reasons. All maps are projected in Equal Earth projection (EPSG:8857).

Combinations of traits are poorly represented across the world

We mapped the alien species richness and the proportion of alien species with measured trait combinations ([Figure 3](#)). The latter measure varied strongly depending on which particular traits we combined ([Figure 3B](#)). Most regions showed trait coverage over 80% when considering single traits. The LHS traits show the highest degrees of coverage variation of all trait combinations considered. Some regions with higher alien species richness also showed high LHS coverage like the North of North America or different regions in Russia. Regions with very low alien species richness (fewer than 10 alien species) showed a high LHS coverage (over 80%). We observed a strong latitudinal gradient in trait coverage. Temperate regions showed an LHS trait coverage over 60%, while most tropical regions showed LHS trait coverage below 60% (with the notable exception of Brazil with several regions over 60% coverage). For both aboveground and especially root traits, most regions showed coverage below 20% of the species, except for regions with low alien species richness. Some regions of North America, South America, and Central Asia, however, showed coverage between 20% and 40% for aboveground spectrum traits.

Traits of widespread and invasive species are more measured

The species' invasion status influenced the knowledge of trait combinations ([Figure 4A](#)). Species referenced as invasive in at least one region in GloNAF ("invasive" in this section) had a greater coverage in all trait combinations than species never referenced as invasive ("non-invasive" in this section). We found strong evidence that LHS traits are better measured for invasive species (48.2% of species) than for non-invasive species (24.8%; $\chi^2 = 325$, $df = 1$, $p\text{-value} < 0.001$). We observed similar differences for aboveground traits and root traits (26.9% vs. 11.3% for aboveground traits; and 3.9% vs. 1.3% for root traits). Even when considering any trait, invasive species were better measured than non-invasives (respectively 99.1% had at least one measured trait vs. 96.1%).

We also observed a difference in trait knowledge depending on geographical spread of species. The 100 most widespread species in GloNAF consistently showed higher trait-combination knowledge than less widespread species (81% vs. 27.5% for LHS traits; 59% vs. 13.5% for aboveground traits; and 18% vs. 1.7% for root traits). The only case where we found no difference between the most widespread and other species was when considering whether they had data on at least one trait (96% vs. 85.3%, respectively, $\chi^2 = 0.57$, $df = 1$, $p\text{-value} = 0.45$).

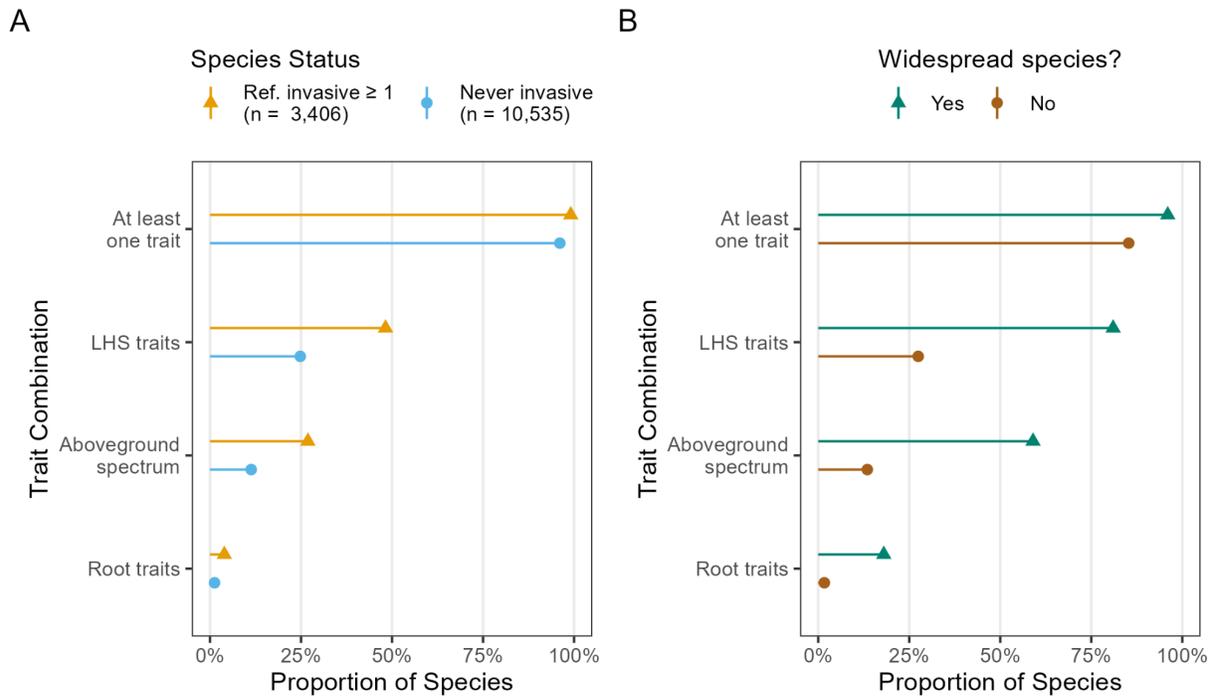


Figure 4. (A) Effect of invasiveness on measured trait combinations across species. The x-axis shows the proportion of species with the given trait combination. The y-axis shows the different trait combinations. Shapes and colors distinguish species based on their invasion status: orange triangles are for species mentioned as invasive at least once and blue circles for the ones never mentioned as invasive. **(B)** Effect of range size (over 100 GloNAF regions from which the species is reported) on trait combination knowledge across species. The x-axis shows the proportion of species with the given trait combination, the y-axis shows different trait combinations. The points and lines of different shapes and colors distinguish species based on their widespreadness: whether they are part of the 100 most widespread species (expressed as the number of GloNAF regions) or not.

Plant growth form and range size predict trait knowledge

We modeled the number of traits measured per species as a function of predictors averaged over the entire range size of the species as well as the growth form of species. We found evidence for effects of all our tested variables ($p < 0.001$, [Figure 5](#), see partial residual plots in [Figure S2](#)). Our model had a Nagelkerke's pseudo- R^2 of 59.4%. The strongest variable explaining the number of measured traits was growth form: measured as trees, shrubs, herbs, and others. Trees had on average more trait information than shrubs (56.5 on average versus 37.4 traits), which had more than herbs (33.3), which had more than species of other growth forms (19.7). The next predictor with the strongest effect was the species total range size, with a positive effect meaning that for every factor 10 increase in the range (in km^2), there was a 85% increase in the number of traits for a species. The number of biomes a species occurs in and the human influence index averaged across its range also had a positive effect on the number of measured traits per species. The other variables all decreased the number of measured traits per species, with GDPpc in the native range having a stronger negative effect (decreasing the number of traits by 20%) than the non-native range size (19%), GDPpc in the non-native range (9%) and, as well as the accessibility of the range (9%). We performed similar analyses with thresholds of 70% and 90% and obtained similar results ([Figure S3](#))

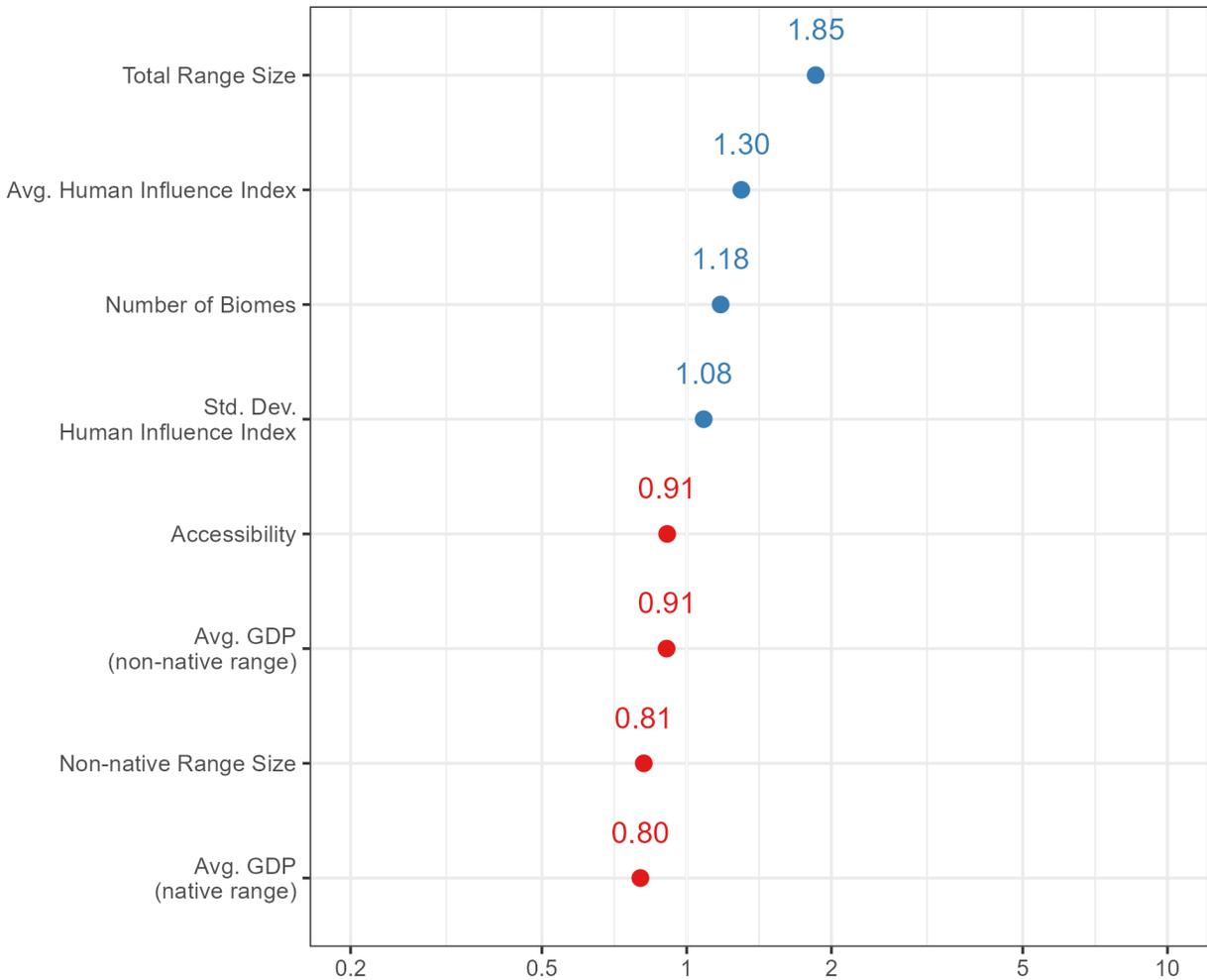


Figure 5. Summary plot showing the incidence rate ratios for variables explaining number of traits measured per alien vascular plant species. Variables with blue dots increase the number of traits measured per species, while variables in red decrease it. The coefficients are interpreted as follows: an increase in 1 of the total range increases the number of measured traits by 85%, while an increase in 1 in accessibility decreases the number of measured traits by 9% ($1 - 0.91$). All tested variables showed p-values < 0.0001 .

Discussion

We assembled the largest collection of traits for alien vascular plant species worldwide from the four biggest global plant trait databases and systematically assessed the completeness of available information. We showed that the global alien species-by-trait matrix was mostly empty and that the most well measured traits were categorical. When considering multiple traits together, we found that only a fraction of species had “classical” trait combinations measured. The knowledge of traits was mostly driven by plant growth form, invasiveness status, and range size. Furthermore, when considering these combinations across space, we identified that most regions in the world showed consistently low trait coverage for alien plant species. The huge gaps in global trait availability of plants in general and, in particular, of alien plants might prevent us from identifying potentially important traits for invasion processes at large scales.

Biases in trait knowledge

Several dimensions of biodiversity are known for showing strong geographic, taxonomic, and trait biases (Hortal *et al.*, 2015; Cornwell *et al.*, 2019; Hughes *et al.*, 2021; Rudbeck *et al.*, 2022). Traits of alien plants are no exception. The lack of trait data, the so-called Raunkiaerian shortfall, is characterized by several biases. First comes the trait bias; although many traits have been measured, and around 70% of species have at least 10 traits measured ([Figure S4](#)), only few traits are consistently measured across species. Second, the taxonomic bias: when traits are measured, we found that even the common ones are not measured uniformly across species within botanical families ([Figure 2](#), [Figure S5](#)). Third, regarding the biogeographic bias, we identified a strong latitudinal gradient in trait knowledge ([Figure 3](#), [Figure S6](#)), with greater trait knowledge for species occurring in temperate regions than for species occurring in tropical regions. Finally, for the invasion and range related biases, we identified that the invasion status and widespreadness of species strongly correlated with the knowledge on their traits. Accounting for these trait, taxonomic, biogeographic, and spatial biases requires careful

analyses. They call for greater attention to data collection, mobilization, and integration to compensate for biases (See [“How to fill the trait data gap” section](#)).

Trait relevance

We decided to focus on commonly used and clearly defined trait combinations, namely LHS traits from Westoby (1998), the global spectrum of plants form and functions (GSPFF) from Díaz *et al.* (2016), and root traits from Bergmann *et al.* (2020) and Weigelt *et al.* (2021). These three trait-combination frameworks are easy to interpret, allowing us to compare species at a global scale. LHS was one of the first clearly defined combinations compared across species (1998), while later the GSPFF traits (2016), extended the LHS set of traits with the two main dimensions being size and the leaf economics spectrum; finally, the root traits (2020; 2021) add information regarding resource acquisition versus conservation, and collaboration with soil microbes (mycorrhizal fungi). They all measure aspects of the strategies of plant species globally, and position them across ecological gradients. Invasion ecology relies heavily on these general ecological trait frameworks, and given the sparsity of data available for these trait combinations, our understanding of the roles of traits in plant invasions can only advance if we identify the most ecologically relevant traits and fill the data gaps.

Although we know that the provenance of the traits, i.e. where the measurements come from (e.g. native or non-native range), is relevant in invasion ecology (e.g. Parker *et al.*, 2013), we could not consider this factor. The main reason is the lack of georeferenced measures (e.g., only 42% of trait observations are georeferenced in TRY v.5.0; Kattge *et al.*, 2020). Additionally, we wanted to include as much trait data as possible in our gap analysis. Identifying if trait measurements are from native or non-native ranges is challenging. Species can show very different trait values between their native and non-native ranges (Leishman *et al.*, 2014). Those differences can potentially point to underlying ecological plasticity, evolutionary processes, or non-random selection of phenotypes at introduction, which are important to understand when

managing invasions. Representing naturalized species trait variability requires measuring them in both the non-native and the native range.

The challenges of integrating trait databases

Even though there are efforts in unifying the format of plant trait databases, they are far from being interoperable or even automatically integrable. One of the challenges we faced in our study was to combine data from heterogeneous trait sources. The four databases we used are complimentary in terms of species coverage and using all four increased data coverage ([Figure S1](#)) but posed serious conceptual and analytical challenges. First, trait data are increasingly shared openly, which means a greater trait coverage for species but scattering into multiple data sources (Gallagher *et al.*, 2020). This problem calls for more attention for data integration and better data sharing practices (Feng *et al.*, 2022). Second, even though the databases are open, they may not be easily accessible. We focused on four trait databases (AusTraits, BIEN, GIFT, and TRY) because all of them offer an R package to access and/or clean their data (Maitner *et al.*, 2018; Falster *et al.*, 2021; Lam *et al.*, 2022; Weigelt & Denelle, 2022). Third, the trait and distribution databases all used different taxonomies, which is a known issue when combining data (Grenié *et al.*, 2022); this emphasizes the importance of preserving original species names when aggregating data. In principle, one would even need to know the taxonomic concept used (Berendsohn, 1995). Fourth, we had to align the trait definitions across databases. Thankfully, standard vocabularies, thesauri, and ontologies facilitate this integration (Garnier *et al.*, 2017; Wenk *et al.*, 2024) but only AusTraits and TRY provided links to trait ontologies. For all other comparisons, the first author manually paired the traits from all the databases. Our proposed correspondence method and cross-database table is available as Supplementary Information in the hope that it would be useful for other studies ([Table S1](#)). Community-developed trait correspondence schemes, for example through the Open Trait Networks (Gallagher *et al.*, 2020), would help enforce interoperability of trait databases so that

definitions would be more consensus-driven and openly discussed (Wenk *et al.*, 2024). In a world of increasing automatic algorithms matching data, looking for patterns (e.g. via AI) an expert driven unifying global plant trait correspondence scheme is the only way to minimize errors in those automatic processes.

How to fill the trait data gap?

Prioritizing trait acquisition

It seems unrealistic to expect all trait gaps to be filled with in-situ measurements in the near future. Adopting a prioritization framework similar to the one used in conservation biology would be more realistic (Arponen, 2012). Prioritization schemes use well-defined criteria on species, traits, or regions, to target data sampling or data integration and increase their trait coverage. The prioritization depends on the aims and purpose of the sampling.

Any prioritization approach would have to make a decision on the origin of the respective trait measure (native or non-native range). Traits of alien species can be measured anywhere in their range but this would limit their ecological applicability as discussed before. In an ideal world, traits are measured in both the native and alien range equally (which is frequently not the case, see Parker *et al.*, 2013). New trait measures should come with a clear georeferenced locality information, including habitat characteristics and note on the invasion status of the species. For example, we could prioritize species to be sampled based on their impact through their (potential) invasiveness (e.g., with their Environmental Impact Classification for Alien Taxa –EICAT– score; Blackburn *et al.*, 2014). Such a prioritization, however, risks reinforcing the gap in trait knowledge between invasive and non-invasive species that we have identified in this study.

We showed greater gaps in trait knowledge of alien plant species in the Tropics than in temperate regions, which suggest a need for a geographic prioritization scheme. Areas richer in alien species could be targeted, as these are more likely to harbor many invasive species

(Chytrý *et al.*, 2012) and suffer from the impacts of invasion. Another region-based approach would prioritize regions with the highest potential increase in projected new alien species in relation to the existing trait knowledge (e.g., Seebens *et al.*, 2021). Finally, because it is likely that many of the trait gaps will not be filled soon, we could rely on methods to prioritize species/traits/locations that would minimize the error from trait imputation methods (Penone *et al.*, 2014; Schrodt *et al.*, 2015; Joswig *et al.*, 2023). Then species and traits would be prioritized to reduce the uncertainty of the imputation the most. For example, we could prioritize species from families where only a few species have been sampled.

Closing the trait gaps

Once species, traits and locations have been prioritized, we need to find ways to close the trait gaps. In this section, we list potential solutions to do so. They fall into two categories: mobilization of existing data and collecting new data. Major gaps in trait data that we identified do not necessarily mean that the traits have never been sampled. Potentially, these traits were measured but never contributed and aggregated into databases. There may be solutions to get these data from previously acquired sources.

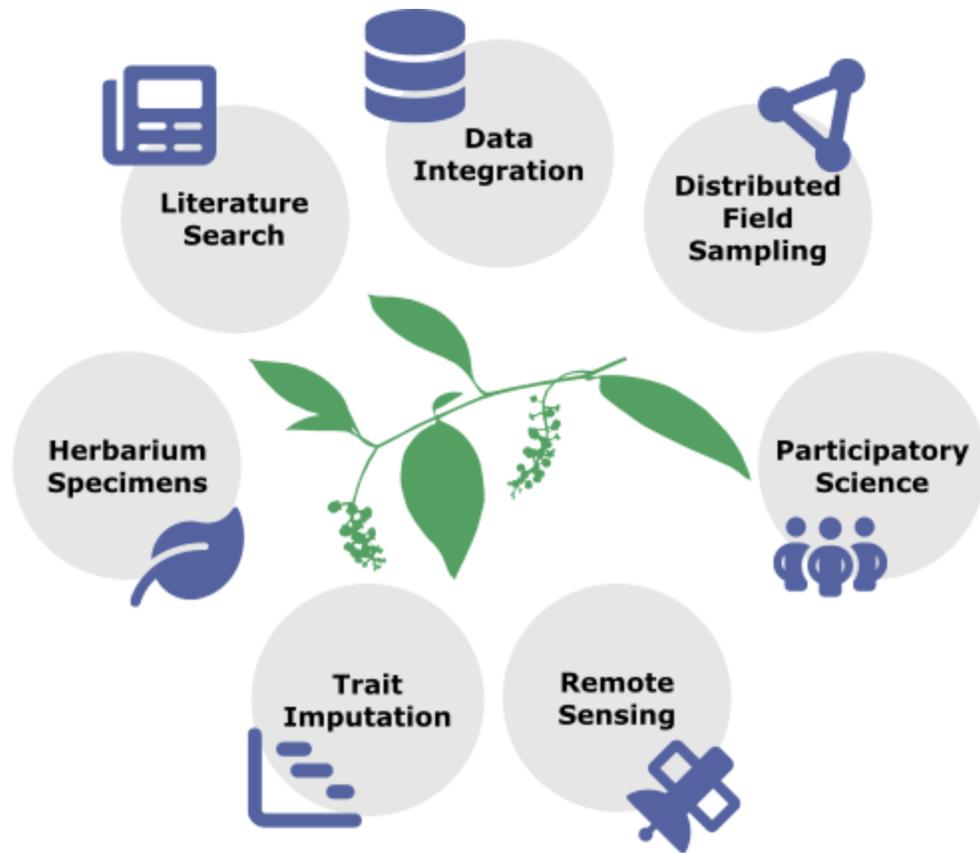


Figure 6. Strategies to increase alien species trait knowledge.

Trait data are increasingly shared openly in the literature. The four trait databases we used do not continuously monitor the published articles for trait data (pers. comm. from database managers). Targeted literature searches, for specific species and traits, could give access to more trait data than available in databases. LT-Brazil is a recent successful example of this strategy (Mariano *et al.*, 2021), where researchers more than doubled the coverage for leaf traits of Brazilian vascular plant species in TRY (i.e. LT-Brazil is now included in TRY) through a well crafted literature search. Recent advancements in natural language processing might in addition reduce the manual effort needed for mobilizing traits from the literature (Domazetoski *et al.*, 2023).

If the traits are not available in databases nor directly from the literature, they may well be privately available from researchers. A targeted call for data can help increase data coverage of some areas and species (Newbold *et al.*, 2012; Kattge *et al.*, 2020). For example, the manager of the PREDICTS database issued a call for data in *Frontiers of Biogeography* that successfully increased data coverage in under-represented regions (Newbold *et al.*, 2012). The calls could be publicly made or through direct contacts with researchers who mobilized the data, like GloNAF did (van Kleunen *et al.*, 2019). These calls should always be accompanied by incentives for data providers like specific citation requirements.

Distributed field campaigns could help to acquire a few traits of alien species. After trait and species prioritization, a call for a global measurement campaign could be issued. The campaign would require standard protocols distributed to partner labs across the world and then pooling their data, also to avoid the definition of new trait states, which is not necessary given >2000 existing plant trait names already. This approach has been used successfully to perform experiments at global scale on nutrient addition through the NutNet network for example (Borer *et al.*, 2014), but it has not been used to acquire trait data to our knowledge.

Participatory science has been rising across many fields in ecology (Silvertown, 2009), empowering large communities to take part in and help science. With the rise of AI-driven plant

identification smartphone applications (Hart *et al.*, 2023), it would be possible to acquire trait data from these applications, though limited on the type of traits that could be (easily) acquired. It would require determining which data can be confidently and accurately acquired by participants, through which tools (e.g., photographs, manual measurements, apps such as BioLeaf (Machado *et al.*, 2016) or LeafByte (Getman-Pickering *et al.*, 2020)) with an appropriate protocol.

With the increasing coverage of satellite imagery, trait ecologists leverage remotely sensed data (Homolová *et al.*, 2013; Feilhauer *et al.*, 2018; Cherif *et al.*, 2023). Recent studies extend their approach to the traits of alien species, though at the community rather than the species level (Huang & Asner, 2009; Niphadkar & Nagendra, 2016). This approach is limited to traits that can be remotely sensed for species occurring in open areas (Niphadkar & Nagendra, 2016). Those approaches also need robust ground truthing data for model calibration (Dechant *et al.*, 2023). Remote-sensing trait distribution forms a dynamic field with strong ongoing efforts to leverage its high resolution capabilities (Torresani *et al.*, 2024).

Several studies show that useful trait data can be extracted from herbarium specimens (Davis, 2022). Herbaria are globally underused resources and can help access useful trait data, especially from difficult to acquire or rare species. While herbarium specimens have been used to reconstruct the spread history of alien species (Mandák *et al.*, 2004; Williamson *et al.*, 2005; Fuentes *et al.*, 2008), they have not been systematically mobilized for trait data. In particular, because many herbaria provide digitized specimens, it would be possible to acquire trait data semi-automatically from these images (Davis, 2022).

Most solutions stated above require collective work from plant, invasion, and remote sensing scientists as well funding schemes which focus on pure data collection campaigns, which rarely exist. We want to emphasize the importance of community building in this regard to tackle the issue of trait data through community efforts. Potential routes to close the gaps in trait knowledge rely on the good will of individual past or present contributors (people who acquired

the data, collected the species for herbaria, citizen scientists, participating labs, etc.) and research funders. We want to underline that any of these scientific contributions should be valued and recognized as they create a basis for progress in research.

Conclusion

We identified large trait gaps for alien plant species at global scale. These gaps are partly driven by uneven sampling and missing integration of data. With distributed efforts of the global community of plant and invasion scientists these gaps can be reduced. Our suggestions should encourage efforts to harmonize plant trait information to be able to unify plant trait databases. Such developments should result in FAIR and open data, increasing incentives for people to deposit their trait data in databases (Wilkinson *et al.*, 2016; Islam *et al.*, 2022). The advent of large-scale trait-based invasion ecology will improve the understanding of biological invasions.

Data Availability

All the code and data are available online through Zenodo as a permanent archive (<https://doi.org/10.5281/zenodo.13940200>) and on GitHub for the development version (<https://github.com/Rekyt/alientraitgaps>).

References

- Arponen, A. (2012) Prioritizing species for conservation planning. *Biodiversity and Conservation*, **21**, 875–893.
- Berendsohn, W.G. (1995) The concept of “potential taxa” in databases. *TAXON*, **44**, 207–212.
- Bergmann, J., Weigelt, A., van der Plas, F., Laughlin, D.C., Kuyper, T.W., Guerrero-Ramirez, N., Valverde-Barrantes, O.J., Bruehlheide, H., Freschet, G.T., Iversen, C.M., Kattge, J., McCormack, M.L., Meier, I.C., Rillig, M.C., Roumet, C., Semchenko, M., Sweeney, C.J., van Ruijven, J., York, L.M. & Mommer, L. (2020) The fungal collaboration gradient dominates the root economics space in plants. *Science Advances*, **6**, eaba3756.
- Blackburn, T.M., Essl, F., Evans, T., Hulme, P.E., Jeschke, J.M., Kühn, I., Kumschick, S., Marková, Z., Mrugała, A., Nentwig, W., Pergl, J., Pyšek, P., Rabitsch, W., Ricciardi, A., Richardson, D.M., Sendek, A., Vilà, M., Wilson, J.R.U., Winter, M., Genovesi, P. & Bacher, S. (2014) A Unified Classification of Alien Species Based on the Magnitude of their Environmental Impacts. *PLOS Biology*, **12**, e1001850.
- Borer, E.T., Harpole, W.S., Adler, P.B., Lind, E.M., Orrock, J.L., Seabloom, E.W. & Smith, M.D.

- (2014) Finding generality in ecology: a model for globally distributed experiments. *Methods in Ecology and Evolution*, **5**, 65–73.
- Cherif, E., Feilhauer, H., Berger, K., Dao, P.D., Ewald, M., Hank, T.B., He, Y., Kovach, K.R., Lu, B., Townsend, P.A. & Kattenborn, T. (2023) From spectra to plant functional traits: Transferable multi-trait models from heterogeneous and sparse data. *Remote Sensing of Environment*, **292**, 113580.
- Chytrý, M., Wild, J., Pyšek, P., Jarošík, V., Dendoncker, N., Reginster, I., Pino, J., Maskell, L.C., Vilà, M., Pergl, J., Kühn, I., Spangenberg, J.H. & Settele, J. (2012) Projecting trends in plant invasions in Europe under different scenarios of future land-use change. *Global Ecology and Biogeography*, **21**, 75–87.
- Cornwell, W.K., Pearse, W.D., Dalrymple, R.L. & Zanne, A.E. (2019) What we (don't) know about global plant diversity. *Ecography*, **42**, 1819–1831.
- Davis, C.C. (2022) The herbarium of the future. *Trends in Ecology & Evolution*, S0169534722002956.
- Dechant, B., Kattge, J., Pavlick, R., Schneider, F., Sabatini, F., Moreno-Martinez, A., Butler, E., Bodegom, P. van, Vallicrosa, H., Kattenborn, T., Boonman, C., Madani, N., Wright, I., Dong, N., Feilhauer, H., Penuelas, J., Sardans, J., Aguirre-Gutierrez, J., Reich, P., Leitao, P., Cavender-Bares, J., Myers-Smith, I.H., Duran, S., Croft, H., Prentice, I.C., Huth, A., Rebel, K., Zaehle, S., Simova, I., Diaz, S., Reichstein, M., Schiller, C., Bruehlheide, H., Mahecha, M., Wirth, C., Malhi, Y. & Townsend, P. (2023) Intercomparison of global foliar trait maps reveals fundamental differences and limitations of upscaling approaches.
- Denelle, P., Weigelt, P. & Kreft, H. (2023) GIFT – an R package to access the Global Inventory of Floras and Traits. 2023.06.27.546704.
- Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Prentice, I.C., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P.B., Moles, A.T., Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Wright, S.J., Sheremet'ev, S.N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M.D. & Gorné, L.D. (2016) The global spectrum of plant form and function. *Nature*, **529**, 167–171.
- Díaz, S., Settele, J., Brondízio, E.S., Ngo, H.T., Agard, J., Arneith, A., Balvanera, P., Brauman, K.A., Butchart, S.H.M., Chan, K.M.A., Garibaldi, L.A., Ichii, K., Liu, J., Subramanian, S.M., Midgley, G.F., Miloslavich, P., Molnár, Z., Obura, D., Pfaff, A., Polasky, S., Purvis, A., Razaque, J., Reyers, B., Chowdhury, R.R., Shin, Y.-J., Visseren-Hamakers, I., Willis, K.J. & Zayas, C.N. (2019) Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science*, **366**, eaax3100.
- Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., Hansen, M., Locke, H., Ellis, E.C., Jones, B., Barber, C.V., Hayes, R., Kormos, C., Martin, V., Crist, E., Sechrest, W., Price, L., Baillie, J.E.M., Weeden, D., Suckling, K., Davis, C., Sizer, N., Moore, R., Thau, D., Birch, T., Potapov, P., Turubanova, S., Tyukavina, A., de Souza, N., Pinteá, L., Brito, J.C., Llewellyn, O.A., Miller, A.G., Patzelt, A., Ghazanfar, S.A., Timberlake, J., Klöser, H., Shennan-Farpón, Y., Kindt, R., Lillesø, J.-P.B., van Breugel, P., Graudal, L., Voge, M., Al-Shammari, K.F. & Saleem, M. (2017) An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *Bioscience*, **67**, 534–545.
- Domazetoski, V., Kreft, H., Bestova, H., Wieder, P., Koynov, R., Zarei, A. & Weigelt, P. (2023) Using natural language processing to extract plant functional traits from unstructured text.
- Drenovsky, R.E., Grewell, B.J., D'Antonio, C.M., Funk, J.L., James, J.J., Molinari, N., Parker, I.M. & Richards, C.L. (2012) A functional trait perspective on plant invasion. *Annals of*

Botany, **110**, 141–153.

- Enquist, B.J., Condit, R., Peet, R.K., Schildhauer, M. & Thiers, B.M. (2016) *Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity*, PeerJ Inc.
- Falster, D., Gallagher, R., Wenk, E.H., Wright, I.J., Indiarto, D., Andrew, S.C., Baxter, C., Lawson, J., Allen, S., Fuchs, A., Monro, A., Kar, F., Adams, M.A., Ahrens, C.W., Alfonzetti, M., Angevin, T., Apgaua, D.M.G., Arndt, S., Atkin, O.K., Atkinson, J., Auld, T., Baker, A., von Balthazar, M., Bean, A., Blackman, C.J., Bloomfield, K., Bowman, D.M.J.S., Bragg, J., Brodribb, T.J., Buckton, G., Burrows, G., Caldwell, E., Camac, J., Carpenter, R., Cattford, J.A., Cawthray, G.R., Cernusak, L.A., Chandler, G., Chapman, A.R., Cheal, D., Cheesman, A.W., Chen, S.-C., Choat, B., Clinton, B., Clode, P.L., Coleman, H., Cornwell, W.K., Cosgrove, M., Crisp, M., Cross, E., Crous, K.Y., Cunningham, S., Curran, T., Curtis, E., Daws, M.I., DeGabriel, J.L., Denton, M.D., Dong, N., Du, P., Duan, H., Duncan, D.H., Duncan, R.P., Duretto, M., Dwyer, J.M., Edwards, C., Esperon-Rodriguez, M., Evans, J.R., Everingham, S.E., Farrell, C., Firn, J., Fonseca, C.R., French, B.J., Froud, D., Funk, J.L., Geange, S.R., Ghannoum, O., Gleason, S.M., Gosper, C.R., Gray, E., Groom, P.K., Grootemaat, S., Gross, C., Guerin, G., Guja, L., Hahs, A.K., Harrison, M.T., Hayes, P.E., Henery, M., Hochuli, D., Howell, J., Huang, G., Hughes, L., Huisman, J., Ilic, J., Jagdish, A., Jin, D., Jordan, G., Jurado, E., Kanowski, J., Kasel, S., Kellermann, J., Kenny, B., Kohout, M., Kooyman, R.M., Kotowska, M.M., Lai, H.R., Laliberté, E., Lambers, H., Lamont, B.B., Lanfear, R., van Langevelde, F., Laughlin, D.C., Laugier-Kitchener, B.-A., Laurance, S., Lehmann, C.E.R., Leigh, A., Leishman, M.R., Lenz, T., Lepschi, B., Lewis, J.D., Lim, F., Liu, U., Lord, J., Lusk, C.H., Macinnis-Ng, C., McPherson, H., Magallón, S., Manea, A., López-Martinez, A., Mayfield, M., McCarthy, J.K., Meers, T., van der Merwe, M., Metcalfe, D.J., Milberg, P., Mokany, K., Moles, A.T., Moore, B.D., Moore, N., Morgan, J.W., Morris, W., Muir, A., Munroe, S., Nicholson, Á., Nicolle, D., Nicotra, A.B., Niinemets, Ü., North, T., O'Reilly-Nugent, A., O'Sullivan, O.S., Oberle, B., Onoda, Y., Ooi, M.K.J., Osborne, C.P., Paczkowska, G., Pekin, B., Guilherme Pereira, C., Pickering, C., Pickup, M., Pollock, L.J., Poot, P., Powell, J.R., Power, S.A., Prentice, I.C., Prior, L., Prober, S.M., Read, J., Reynolds, V., Richards, A.E., Richardson, B., Roderick, M.L., Rosell, J.A., Rossetto, M., Rye, B., Rymer, P.D., Sams, M.A., Sanson, G., Sauquet, H., Schmidt, S., Schönenberger, J., Schulze, E.-D., Sendall, K., Sinclair, S., Smith, B., Smith, R., Soper, F., Sparrow, B., Standish, R.J., Staples, T.L., Stephens, R., Szota, C., Taseski, G., Tasker, E., Thomas, F., Tissue, D.T., Tjoelker, M.G., Tng, D.Y.P., de Tombeur, F., Tomlinson, K., Turner, N.C., Veneklaas, E.J., Venn, S., Vesk, P., Vlasveld, C., Vorontsova, M.S., Warren, C.A., Warwick, N., Weerasinghe, L.K., Wells, J., Westoby, M., White, M., Williams, N.S.G., Wills, J., Wilson, P.G., Yates, C., Zanne, A.E., Zemunik, G. & Ziemnińska, K. (2021) AusTraits, a curated plant trait database for the Australian flora. *Scientific Data*, **8**, 254.
- Feilhauer, H., Schmid, T., Faude, U., Sánchez-Carrillo, S. & Cirujano, S. (2018) Are remotely sensed traits suitable for ecological analysis? A case study of long-term drought effects on leaf mass per area of wetland vegetation. *Ecological Indicators*, **88**, 232–240.
- Feng, X., Enquist, B.J., Park, D.S., Boyle, B., Breshears, D.D., Gallagher, R.V., Lien, A., Newman, E.A., Burger, J.R., Maitner, B.S., Merow, C., Li, Y., Huynh, K.M., Ernst, K., Baldwin, E., Foden, W., Hannah, L., Jørgensen, P.M., Kraft, N.J.B., Lovett, J.C., Marquet, P.A., McGill, B.J., Morueta-Holme, N., Neves, D.M., Núñez-Regueiro, M.M., Oliveira-Filho, A.T., Peet, R.K., Pillet, M., Roehrdanz, P.R., Sandel, B., Serra-Diaz, J.M., Šímová, I., Svenning, J.-C., Violle, C., Weitemier, T.D., Wiser, S. & López-Hoffman, L. (2022) A review of the heterogeneous landscape of biodiversity databases: Opportunities and challenges for a synthesized biodiversity knowledge base. *Global Ecology and Biogeography*, **31**, 1242–1260.

- Fuentes, N., Ugarte, E., Kühn, I. & Klotz, S. (2008) Alien plants in Chile: inferring invasion periods from herbarium records. *Biological Invasions*, **10**, 649–657.
- Gallagher, R.V., Falster, D.S., Maitner, B.S., Salguero-Gómez, R., Vandvik, V., Pearse, W.D., Schneider, F.D., Kattge, J., Poelen, J.H., Madin, J.S., Ankenbrand, M.J., Penone, C., Feng, X., Adams, V.M., Alroy, J., Andrew, S.C., Balk, M.A., Bland, L.M., Boyle, B.L., Bravo-Avila, C.H., Brennan, I., Carthey, A.J.R., Catullo, R., Cavazos, B.R., Conde, D.A., Chown, S.L., Fadrique, B., Gibb, H., Halbritter, A.H., Hammock, J., Hogan, J.A., Holewa, H., Hope, M., Iversen, C.M., Jochum, M., Kearney, M., Keller, A., Mabee, P., Manning, P., McCormack, L., Michaletz, S.T., Park, D.S., Perez, T.M., Pineda-Munoz, S., Ray, C.A., Rossetto, M., Sauquet, H., Sparrow, B., Spasojevic, M.J., Telford, R.J., Tobias, J.A., Violle, C., Walls, R., Weiss, K.C.B., Westoby, M., Wright, I.J. & Enquist, B.J. (2020) Open Science principles for accelerating trait-based science across the Tree of Life. *Nature Ecology & Evolution*, 1–10.
- Gallien, L. & Carboni, M. (2017) The community ecology of invasive species: where are we and what's next? *Ecography*, **40**, 335–352.
- Garnier, E., Stahl, U., Laporte, M.-A., Kattge, J., Mougnot, I., Kühn, I., Laporte, B., Amiaud, B., Ahrestani, F.S., Bönnisch, G., Bunker, D.E., Cornelissen, J.H.C., Díaz, S., Enquist, B.J., Gachet, S., Jaureguiberry, P., Kleyer, M., Lavorel, S., Maicher, L., Pérez-Harguindeguy, N., Poorter, H., Schildhauer, M., Shipley, B., Violle, C., Weiher, E., Wirth, C., Wright, I.J. & Klotz, S. (2017) Towards a thesaurus of plant characteristics: an ecological contribution. *Journal of Ecology*, **105**, 298–309.
- Getman-Pickering, Z.L., Campbell, A., Aflitto, N., Grele, A., Davis, J.K. & Ugine, T.A. (2020) LeafByte: A mobile application that measures leaf area and herbivory quickly and accurately. *Methods in Ecology and Evolution*, **11**, 215–221.
- Golivets, M., Knapp, S., Essl, F., Lenzner, B., Latombe, G., Leung, B. & Kühn, I. (2024) Future changes in key plant traits across Central Europe vary with biogeographical status, woodiness, and habitat type. *Science of The Total Environment*, **907**, 167954.
- Grenié, M., Berti, E., Carvajal-Quintero, J., Dädlow, G.M.L., Sagouis, A. & Winter, M. (2022) Harmonizing taxon names in biodiversity data: A review of tools, databases and best practices. *Methods in Ecology and Evolution*, 2041–210X.13802.
- Hart, A.G., Bosley, H., Hooper, C., Perry, J., Sellors-Moore, J., Moore, O. & Goodenough, A.E. (2023) Assessing the accuracy of free automated plant identification applications. *People and Nature*, **5**, 929–937.
- Homolová, L., Malenovský, Z., Clevers, J.G.P.W., García-Santos, G. & Schaepman, M.E. (2013) Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity*, **15**, 1–16.
- Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015) Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, **46**, 523–549.
- Huang, C. & Asner, G.P. (2009) Applications of Remote Sensing to Alien Invasive Plant Studies. *Sensors*, **9**, 4869–4889.
- Hughes, A.C., Orr, M.C., Ma, K., Costello, M.J., Waller, J., Provoost, P., Yang, Q., Zhu, C. & Qiao, H. (2021) Sampling biases shape our view of the natural world. *Ecography*, **44**, 1259–1269.
- Islam, S., Weiland, C. & Addink, W. (2022) From data pipelines to FAIR data infrastructures: A vision for the new horizons of bio-and geodiversity data for scientific research. *Research Ideas and Outcomes*, **8**, e93816.
- Joswig, J.S., Kattge, J., Kraemer, G., Mahecha, M.D., Rüger, N., Schaepman, M.E., Schrod, F. & Schuman, M.C. (2023) Imputing missing data in plant traits: A guide to improve gap-filling. *Global Ecology and Biogeography*, **32**, 1395–1408.
- Kattge, J., Bönnisch, G., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Tautenhahn, S., Werner,

G.D.A., Aakala, T., Abedi, M., Acosta, A.T.R., Adamidis, G.C., Adamson, K., Aiba, M., Albert, C.H., Alcántara, J.M., Alcázar C, C., Aleixo, I., Ali, H., Amiaud, B., Ammer, C., Amoroso, M.M., Anand, M., Anderson, C., Anten, N., Antos, J., Apgaua, D.M.G., Ashman, T.-L., Asmara, D.H., Asner, G.P., Aspinwall, M., Atkin, O., Aubin, I., Baastrup-Spohr, L., Bahalkeh, K., Bahn, M., Baker, T., Baker, W.J., Bakker, J.P., Baldocchi, D., Baltzer, J., Banerjee, A., Baranger, A., Barlow, J., Barneche, D.R., Baruch, Z., Bastianelli, D., Battles, J., Bauerle, W., Bauters, M., Bazzato, E., Beckmann, M., Beeckman, H., Beierkuhnlein, C., Bekker, R., Belfry, G., Belluau, M., Beloiu, M., Benavides, R., Benomar, L., Berdugo-Lattke, M.L., Berenguer, E., Bergamin, R., Bergmann, J., Bergmann Carlucci, M., Berner, L., Bernhardt-Römermann, M., Bigler, C., Bjorkman, A.D., Blackman, C., Blanco, C., Blonder, B., Blumenthal, D., Bocanegra-González, K.T., Boeckx, P., Bohlman, S., Böhning-Gaese, K., Boisvert-Marsh, L., Bond, W., Bond-Lamberty, B., Boom, A., Boonman, C.C.F., Bordin, K., Boughton, E.H., Boukili, V., Bowman, D.M.J.S., Bravo, S., Brendel, M.R., Broadley, M.R., Brown, K.A., Bruelheide, H., Brumnich, F., Bruun, H.H., Bruy, D., Buchanan, S.W., Bucher, S.F., Buchmann, N., Buitenwerf, R., Bunker, D.E., Bürger, J., Burrascano, S., Burslem, D.F.R.P., Butterfield, B.J., Byun, C., Marques, M., Scalon, M.C., Caccianiga, M., Cadotte, M., Cailleret, M., Camac, J., Camarero, J.J., Company, C., Campetella, G., Campos, J.A., Cano-Arboleda, L., Canullo, R., Carbognani, M., Carvalho, F., Casanoves, F., Castagneyrol, B., Catford, J.A., Cavender-Bares, J., Cerabolini, B.E.L., Cervellini, M., Chacón-Madrigal, E., Chapin, K., Chapin, F.S., Chelli, S., Chen, S.-C., Chen, A., Cherubini, P., Chianucci, F., Choat, B., Chung, K.-S., Chytrý, M., Ciccarelli, D., Coll, L., Collins, C.G., Conti, L., Coomes, D., Cornelissen, J.H.C., Cornwell, W.K., Corona, P., Coyea, M., Craine, J., Craven, D., Cromsigt, J.P.G.M., Csecserits, A., Cufar, K., Cuntz, M., da Silva, A.C., Dahlin, K.M., Dainese, M., Dalke, I., Dalle Fratte, M., Dang-Le, A.T., Danihelka, J., Dannoura, M., Dawson, S., de Beer, A.J., De Frutos, A., De Long, J.R., Dechant, B., Delagrangé, S., Delpierre, N., Derroire, G., Dias, A.S., Diaz-Toribio, M.H., Dimitrakopoulos, P.G., Dobrowolski, M., Doktor, D., Dřevojan, P., Dong, N., Dransfield, J., Dressler, S., Duarte, L., Ducouret, E., Dullinger, S., Durka, W., Duursma, R., Dymova, O., E-Vojtkó, A., Eckstein, R.L., Ejtehadi, H., Elser, J., Emilio, T., Engemann, K., Erfanian, M.B., Erfmeier, A., Esquivel-Muelbert, A., Esser, G., Estiarte, M., Domingues, T.F., Fagan, W.F., Fagúndez, J., Falster, D.S., Fan, Y., Fang, J., Farris, E., Fazlioglu, F., Feng, Y., Fernandez-Mendez, F., Ferrara, C., Ferreira, J., Fidelis, A., Finegan, B., Firn, J., Flowers, T.J., Flynn, D.F.B., Fontana, V., Forey, E., Forgiarini, C., François, L., Frangipani, M., Frank, D., Frenette-Dussault, C., Freschet, G.T., Fry, E.L., Fyllas, N.M., Mazzochini, G.G., Gachet, S., Gallagher, R., Ganade, G., Ganga, F., García-Palacios, P., Gargaglione, V., Garnier, E., Garrido, J.L., de Gasper, A.L., Gea-Izquierdo, G., Gibson, D., Gillison, A.N., Giroldo, A., Glasenhardt, M.-C., Gleason, S., Gliesch, M., Goldberg, E., Göldel, B., Gonzalez-Akre, E., Gonzalez-Andujar, J.L., González-Melo, A., González-Robles, A., Graae, B.J., Granda, E., Graves, S., Green, W.A., Gregor, T., Gross, N., Guerin, G.R., Günther, A., Gutiérrez, A.G., Haddock, L., Haines, A., Hall, J., Hambuckers, A., Han, W., Harrison, S.P., Hattingh, W., Hawes, J.E., He, T., He, P., Heberling, J.M., Helm, A., Hempel, S., Hentschel, J., Hérault, B., Hereş, A.-M., Herz, K., Heuertz, M., Hickler, T., Hietz, P., Higuchi, P., Hipp, A.L., Hirons, A., Hock, M., Hogan, J.A., Holl, K., Honnay, O., Hornstein, D., Hou, E., Hough-Snee, N., Hovstad, K.A., Ichie, T., Igić, B., Illa, E., Isaac, M., Ishihara, M., Ivanov, L., Ivanova, L., Iversen, C.M., Izquierdo, J., Jackson, R.B., Jackson, B., Jactel, H., Jagodzinski, A.M., Jandt, U., Jansen, S., Jenkins, T., Jentsch, A., Jespersen, J.R.P., Jiang, G.-F., Johansen, J.L., Johnson, D., Jokela, E.J., Joly, C.A., Jordan, G.J., Joseph, G.S., Junaedi, D., Junker, R.R., Justes, E., Kabzems, R., Kane, J., Kaplan, Z., Kattenborn, T., Kavelenova, L., Kearsley, E., Kempel, A., Kenzo, T., Kerkhoff, A., Khalil, M.I., Kinlock, N.L., Kissling,

W.D., Kitajima, K., Kitzberger, T., Kjølner, R., Klein, T., Kleyer, M., Klimešová, J., Klipel, J., Kloeppel, B., Klotz, S., Knops, J.M.H., Kohyama, T., Koike, F., Kollmann, J., Komac, B., Komatsu, K., König, C., Kraft, N.J.B., Kramer, K., Kreft, H., Kühn, I., Kumarathunge, D., Kuppler, J., Kurokawa, H., Kurosawa, Y., Kuyah, S., Laclau, J.-P., Lafleur, B., Lallai, E., Lamb, E., Lamprecht, A., Larkin, D.J., Laughlin, D., Le Bagousse-Pinguet, Y., le Maire, G., le Roux, P.C., le Roux, E., Lee, T., Lens, F., Lewis, S.L., Lhotsky, B., Li, Y., Li, X., Lichstein, J.W., Liebergesell, M., Lim, J.Y., Lin, Y.-S., Linares, J.C., Liu, C., Liu, D., Liu, U., Livingstone, S., Llusià, J., Lohbeck, M., López-García, Á., Lopez-Gonzalez, G., Lososová, Z., Louault, F., Lukács, B.A., Lukeš, P., Luo, Y., Lussu, M., Ma, S., Maciel Rabelo Pereira, C., Mack, M., Maire, V., Mäkelä, A., Mäkinen, H., Malhado, A.C.M., Mallik, A., Manning, P., Manzoni, S., Marchetti, Z., Marchino, L., Marcilio-Silva, V., Marcon, E., Marignani, M., Markesteijn, L., Martin, A., Martínez-Garza, C., Martínez-Vilalta, J., Mašková, T., Mason, K., Mason, N., Massad, T.J., Masse, J., Mayrose, I., McCarthy, J., McCormack, M.L., McCulloh, K., McFadden, I.R., McGill, B.J., McPartland, M.Y., Medeiros, J.S., Medlyn, B., Meerts, P., Mehrabi, Z., Meir, P., Melo, F.P.L., Mencuccini, M., Meredieu, C., Messier, J., Mészáros, I., Metsaranta, J., Michaletz, S.T., Michelaki, C., Migalina, S., Milla, R., Miller, J.E.D., Minden, V., Ming, R., Mokany, K., Moles, A.T., Molnár, V., Molofsky, J., Molz, M., Montgomery, R.A., Monty, A., Moravcová, L., Moreno-Martínez, A., Moretti, M., Mori, A.S., Mori, S., Morris, D., Morrison, J., Mucina, L., Mueller, S., Muir, C.D., Müller, S.C., Munoz, F., Myers-Smith, I.H., Myster, R.W., Nagano, M., Naidu, S., Narayanan, A., Natesan, B., Negoita, L., Nelson, A.S., Neuschulz, E.L., Ni, J., Niedrist, G., Nieto, J., Niinemets, Ü., Nolan, R., Nottebrock, H., Nouvellon, Y., Novakovskiy, A., Network, T.N., Nystuen, K.O., O'Grady, A., O'Hara, K., O'Reilly-Nugent, A., Oakley, S., Oberhuber, W., Ohtsuka, T., Oliveira, R., Öllerer, K., Olson, M.E., Onipchenko, V., Onoda, Y., Onstein, R.E., Ordonez, J.C., Osada, N., Ostonen, I., Ottaviani, G., Otto, S., Overbeck, G.E., Ozinga, W.A., Pahl, A.T., Paine, C.E.T., Pakeman, R.J., Papageorgiou, A.C., Parfionova, E., Pärtel, M., Patacca, M., Paula, S., Paule, J., Pauli, H., Pausas, J.G., Peco, B., Penuelas, J., Perea, A., Peri, P.L., Petisco-Souza, A.C., Petraglia, A., Petritan, A.M., Phillips, O.L., Pierce, S., Pillar, V.D., Pisek, J., Pomogaybin, A., Poorter, H., Portsmouth, A., Poschlod, P., Potvin, C., Pounds, D., Powell, A.S., Power, S.A., Prinzing, A., Puglielli, G., Pyšek, P., Raavel, V., Rammig, A., Ransijn, J., Ray, C.A., Reich, P.B., Reichstein, M., Reid, D.E.B., Réjou-Méchain, M., de Dios, V.R., Ribeiro, S., Richardson, S., Riibak, K., Rillig, M.C., Riviera, F., Robert, E.M.R., Roberts, S., Robroek, B., Roddy, A., Rodrigues, A.V., Rogers, A., Rollinson, E., Rolo, V., Römermann, C., Ronzhina, D., Roscher, C., Rosell, J.A., Rosenfield, M.F., Rossi, C., Roy, D.B., Royer-Tardif, S., Rüger, N., Ruiz-Peinado, R., Rumpf, S.B., Rusch, G.M., Ryo, M., Sack, L., Saldaña, A., Salgado-Negret, B., Salguero-Gomez, R., Santa-Regina, I., Santacruz-García, A.C., Santos, J., Sardans, J., Schamp, B., Scherer-Lorenzen, M., Schleuning, M., Schmid, B., Schmidt, M., Schmitt, S., Schneider, J.V., Schowaneck, S.D., Schrader, J., Schrod, F., Schuldt, B., Schurr, F., Selaya Garvizu, G., Semchenko, M., Seymour, C., Sfair, J.C., Sharpe, J.M., Sheppard, C.S., Sheremetiev, S., Shiodera, S., Shipley, B., Shovon, T.A., Siebenkäs, A., Sierra, C., Silva, V., Silva, M., Sitzia, T., Sjöman, H., Slot, M., Smith, N.G., Sodhi, D., Soltis, P., Soltis, D., Somers, B., Sonnier, G., Sørensen, M.V., Sosinski Jr, E.E., Soudzilovskaia, N.A., Souza, A.F., Spasojevic, M., Sperandii, M.G., Stan, A.B., Stegen, J., Steinbauer, K., Stephan, J.G., Sterck, F., Stojanovic, D.B., Strydom, T., Suarez, M.L., Svenning, J.-C., Svitková, I., Svitok, M., Svoboda, M., Swaine, E., Swenson, N., Tabarelli, M., Takagi, K., Tappeiner, U., Tarifa, R., Tauougourdeau, S., Tavsanoğlu, C., te Beest, M., Tedersoo, L., Thiffault, N., Thom, D., Thomas, E., Thompson, K., Thornton, P.E., Thuiller, W., Tichý, L., Tissue, D., Tjoelker, M.G., Tng, D.Y.P., Tobias, J., Török, P., Tarin, T., Torres-Ruiz, J.M., Tóthmérész, B., Treurnicht, M., Trivellone, V., Trolliet, F., Trotsiuk, V.,

- Tsakalos, J.L., Tsiripidis, I., Tysklind, N., Umehara, T., Usoltsev, V., Vadeboncoeur, M., Vaezi, J., Valladares, F., Vamosi, J., van Bodegom, P.M., van Breugel, M., Van Cleemput, E., van de Weg, M., van der Merwe, S., van der Plas, F., van der Sande, M.T., van Kleunen, M., Van Meerbeek, K., Vanderwel, M., Vanselow, K.A., Vårhammar, A., Varone, L., Vasquez Valderrama, M.Y., Vassilev, K., Vellend, M., Veneklaas, E.J., Verbeeck, H., Verheyen, K., Vibrans, A., Vieira, I., Villacís, J., Violle, C., Vivek, P., Wagner, K., Waldram, M., Waldron, A., Walker, A.P., Waller, M., Walther, G., Wang, H., Wang, F., Wang, W., Watkins, H., Watkins, J., Weber, U., Weedon, J.T., Wei, L., Weigelt, P., Weiher, E., Wells, A.W., Wellstein, C., Wenk, E., Westoby, M., Westwood, A., White, P.J., Whitten, M., Williams, M., Winkler, D.E., Winter, K., Womack, C., Wright, I.J., Wright, S.J., Wright, J., Pinho, B.X., Ximenes, F., Yamada, T., Yamaji, K., Yanai, R., Yankov, N., Yguel, B., Zanini, K.J., Zanne, A.E., Zelený, D., Zhao, Y.-P., Zheng, J., Zheng, J., Ziemińska, K., Zirbel, C.R., Zizka, G., Zo-Bi, I.C., Zotz, G. & Wirth, C. (2020) TRY plant trait database – enhanced coverage and open access. *Global Change Biology*, **26**, 119–188.
- Keller, A., Ankenbrand, M.J., Bruelheide, H., Dekeyser, S., Enquist, B.J., Erfanian, M.B., Falster, D.S., Gallagher, R.V., Hammock, J., Kattge, J., Leonhardt, S.D., Madin, J.S., Maitner, B., Neyret, M., Onstein, R.E., Pearse, W.D., Poelen, J.H., Salguero-Gomez, R., Schneider, F.D., Tóth, A.B. & Penone, C. (2023) Ten (mostly) simple rules to future-proof trait data in ecological and evolutionary sciences. *Methods in Ecology and Evolution*, **14**, 444–458.
- van Kleunen, M., Dawson, W. & Maurel, N. (2015) Characteristics of successful alien plants. *Molecular Ecology*, **24**, 1954–1968.
- van Kleunen, M., Pyšek, P., Dawson, W., Essl, F., Kreft, H., Pergl, J., Weigelt, P., Stein, A., Dullinger, S., König, C., Lenzner, B., Maurel, N., Moser, D., Seebens, H., Kartesz, J., Nishino, M., Aleksanyan, A., Ansong, M., Antonova, L.A., Barcelona, J.F., Breckle, S.W., Brundu, G., Cabezas, F.J., Cárdenas, D., Cárdenas-Toro, J., Castaño, N., Chacón, E., Chatelain, C., Conn, B., de Sá Dechoum, M., Dufour-Dror, J.-M., Ebel, A.L., Figueiredo, E., Fragman-Sapir, O., Fuentes, N., Groom, Q.J., Henderson, L., Inderjit, Jogan, N., Krestov, P., Kupriyanov, A., Masciadri, S., Meerman, J., Morozova, O., Nickrent, D., Nowak, A., Patzelt, A., Pelsner, P.B., Shu, W., Thomas, J., Uludag, A., Velayos, M., Verkhosina, A., Villaseñor, J.L., Weber, E., Wieringa, J.J., Yazlık, A., Zeddner, A., Zykova, E. & Winter, M. (2019) The Global Naturalized Alien Flora (GloNAF) database. *Ecology*, **100**, e02542.
- Kummu, M., Taka, M. & Guillaume, J.H.A. (2018) Gridded global datasets for Gross Domestic Product and Human Development Index over 1990–2015. *Scientific Data*, **5**, 180004.
- Küster, E.C., Kühn, I., Bruelheide, H. & Klotz, S. (2008) Trait interactions help explain plant invasion success in the German flora. *Journal of Ecology*, **96**, 860–868.
- Lam, O.H.Y., Tautenhahn, S., Walther, G., Boenisch, G., Baddam, P. & Kattge, J. (2022) The “rtry” R package for preprocessing plant trait data. EGU22-13251.
- Leishman, M.R., Cooke, J. & Richardson, D.M. (2014) Evidence for shifts to faster growth strategies in the new ranges of invasive alien plants. *Journal of Ecology*, **102**, 1451–1461.
- Lüdecke, D., Ben-Shachar, M.S., Patil, I., Waggoner, P. & Makowski, D. (2021) performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, **6**, 3139.
- Machado, B.B., Orue, J.P., Arruda, M.S., Santos, C.V., Sarath, D.S., Goncalves, W.N., Silva, G.G., Pistori, H., Roel, A.R. & Rodrigues-Jr, J.F. (2016) BioLeaf: A professional mobile application to measure foliar damage caused by insect herbivory. *Computers and electronics in agriculture*, **129**, 44–55.
- Maitner, B., Gallagher, R., Svenning, J.-C., Tietje, M., Wenk, E.H. & Eiserhardt, W.L. (2023) A global assessment of the Raunkiaeran shortfall in plants: geographic biases in our

- knowledge of plant traits. *New Phytologist*, **240**, 1345–1354.
- Maitner, B.S., Boyle, B., Casler, N., Condit, R., Donoghue, J., Durán, S.M., Guaderrama, D., Hinchliff, C.E., Jørgensen, P.M., Kraft, N.J.B., McGill, B., Merow, C., Morueta-Holme, N., Peet, R.K., Sandel, B., Schildhauer, M., Smith, S.A., Svenning, J.-C., Thiers, B., Violle, C., Wiser, S. & Enquist, B.J. (2018) The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods in Ecology and Evolution*, **9**, 373–379.
- Mandák, B., Pyšek, P. & Bímová, K. (2004) History of the invasion and distribution of *Reynoutria* taxa in the Czech Republic: a hybrid spreading faster than its parents.
- Mariano, E., Gomes, T.F., Lins, S.R.M., Abdalla-Filho, A.L., Soltangheisi, A., Araújo, M.G.S., Almeida, R.F., Augusto, F.G., Canisares, L.P., Chaves, S.S.F., Costa, C.F.G., Diniz-Reis, T.R., Galera, L.A., Martinez, M.G., Morais, M.C., Perez, E.B., Reis, L.C., da P. Simon, C., Mardegan, S.F., Domingues, T.F., Miatto, R.C., Oliveira, R.S., Reis, C.R.G., Nardoto, G.B., Kattge, J. & Martinelli, L.A. (2021) LT-Brazil: A database of leaf traits across biomes and vegetation types in Brazil. *Global Ecology and Biogeography*, **30**, 2136–2146.
- Meyer, C., Weigelt, P. & Kreft, H. (2016) Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecology Letters*, **19**, 992–1006.
- Newbold, T., Hudson, L., Purves, D.W., Scharlemann, J.P.W., Mace, G. & Purvis, A. (2012) call for data: PREDICTS: Projecting Responses of Ecological Diversity in Changing Terrestrial Systems. *Frontiers of Biogeography*, **4**.
- Niphadkar, M. & Nagendra, H. (2016) Remote sensing of invasive plants: incorporating functional traits into the picture. *International Journal of Remote Sensing*, **37**, 3074–3085.
- Novoa, A., Richardson, D.M., Pyšek, P., Meyerson, L.A., Bacher, S., Canavan, S., Catford, J.A., Čuda, J., Essl, F., Foxcroft, L.C., Genovesi, P., Hirsch, H., Hui, C., Jackson, M.C., Kueffer, C., Le Roux, J.J., Measey, J., Mohanty, N.P., Moodley, D., Müller-Schärer, H., Packer, J.G., Pergl, J., Robinson, T.B., Saul, W.-C., Shackleton, R.T., Visser, V., Weyl, O.L.F., Yannelli, F.A. & Wilson, J.R.U. (2020) Invasion syndromes: a systematic approach for predicting biological invasions and facilitating effective management. *Biological Invasions*, **22**, 1801–1820.
- Oliveira, U., Paglia, A.P., Brescovit, A.D., de Carvalho, C.J.B., Silva, D.P., Rezende, D.T., Leite, F.S.F., Batista, J.A.N., Barbosa, J.P.P.P., Stehmann, J.R., Ascher, J.S., de Vasconcelos, M.F., De Marco Jr, P., Löwenberg-Neto, P., Dias, P.G., Ferro, V.G. & Santos, A.J. (2016) The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. *Diversity and Distributions*, **22**, 1232–1244.
- Parker, J.D., Torchin, M.E., Hufbauer, R.A., Lemoine, N.P., Alba, C., Blumenthal, D.M., Bossdorf, O., Byers, J.E., Dunn, A.M., Heckman, R.W., Hejda, M., Jarošík, V., Kanarek, A.R., Martin, L.B., Perkins, S.E., Pyšek, P., Schierenbeck, K., Schlöder, C., van Klinken, R., Vaughn, K.J., Williams, W. & Wolfe, L.M. (2013) Do invasive species perform better in their new ranges? *Ecology*, **94**, 985–994.
- Penone, C., Davidson, A.D., Shoemaker, K.T., Di Marco, M., Rondinini, C., Brooks, T.M., Young, B.E., Graham, C.H. & Costa, G.C. (2014) Imputation of missing data in life-history trait datasets: which approach performs the best? *Methods in Ecology and Evolution*, **5**, 961–970.
- Pyšek, P., Bacher, S., Kühn, I., Novoa, A., Catford, J.A., Hulme, P.E., Pergl, J., Richardson, D.M., Wilson, J.R.U. & Blackburn, T.M. (2020) MAcroecological Framework for Invasive Aliens (MAFIA): disentangling large-scale context dependence in biological invasions. *NeoBiota*, **62**, 407–461.
- Pyšek, P., Richardson, D.M., Pergl, J., Jarošík, V., Sixtová, Z. & Weber, E. (2008) Geographical and taxonomic biases in invasion ecology. *Trends in Ecology & Evolution*, **23**, 237–244.

- R Core Team (2022) *R: A language and environment for statistical computing*, Vienna, Austria.
- Razanajatovo, M., Maurel, N., Dawson, W., Essl, F., Kreft, H., Pergl, J., Pyšek, P., Weigelt, P., Winter, M. & van Kleunen, M. (2016) Plants capable of selfing are more likely to become naturalized. *Nature Communications*, **7**, 13313.
- Roy, H., Pauchard, A., Stoett, P., Renard Truong, T., Bacher, S., Galil, B., Hulme, P., Ikeda, T., Sankaran, K.V. & McGeoch, M. (2023) Summary for Policymakers of the Thematic Assessment Report on Invasive Alien Species and their Control. *IPBES Invasive Alien Species Assessment*, 1–56.
- Rudbeck, A.V., Sun, M., Tietje, M., Gallagher, R.V., Govaerts, R., Smith, S.A., Svenning, J.-C. & Eiserhardt, W.L. (2022) The Darwinian shortfall in plants: phylogenetic knowledge is driven by range size. *Ecography*, **2022**, e06142.
- Schrodt, F., Kattge, J., Shan, H., Fazayeli, F., Joswig, J., Banerjee, A., Reichstein, M., Bönisch, G., Díaz, S., Dickie, J., Gillison, A., Karpátne, A., Lavorel, S., Leadley, P., Wirth, C.B., Wright, I.J., Wright, S.J. & Reich, P.B. (2015) BHPMF – a hierarchical Bayesian approach to gap-filling and trait prediction for macroecology and functional biogeography. *Global Ecology and Biogeography*, **24**, 1510–1521.
- Seebens, H., Bacher, S., Blackburn, T.M., Capinha, C., Dawson, W., Dullinger, S., Genovesi, P., Hulme, P.E., van Kleunen, M., Kühn, I., Jeschke, J.M., Lenzner, B., Liebhold, A.M., Pattison, Z., Pergl, J., Pyšek, P., Winter, M. & Essl, F. (2021) Projecting the continental accumulation of alien species through to 2050. *Global Change Biology*, **27**, 970–982.
- Silvertown, J. (2009) A new dawn for citizen science. *Trends in ecology & evolution*, **24**, 467–471.
- Torresani, M., Rossi, C., Perrone, M., Hauser, L.T., Féret, J.-B., Moudrý, V., Simova, P., Ricotta, C., Foody, G.M., Kacic, P., Feilhauer, H., Malavasi, M., Tognetti, R. & Rocchini, D. (2024) Reviewing the Spectral Variation Hypothesis: Twenty years in the tumultuous sea of biodiversity estimation by remote sensing. *Ecological Informatics*, **82**, 102702.
- Venables, W.N. & Ripley, B.D. (2002) *Modern applied statistics with S*, 4th edn. Springer, New York.
- Visscher, A.M., Vandeloos, F., Fernández-Pascual, E., Pérez-Martínez, L.V., Ulian, T., Diazgranados, M. & Mattana, E. (2022) Low availability of functional seed trait data from the tropics could negatively affect global macroecological studies, predictive models and plant conservation. *Annals of Botany*, **130**, 773–784.
- WCS & CIESIN (2005) Last of the Wild Project, Version 2, 2005 (LWP-2): Global Human Influence Index (HII) Dataset (Geographic).
- Weigelt, A., Mommer, L., Andrzejek, K., Iversen, C.M., Bergmann, J., Bruehlheide, H., Fan, Y., Freschet, G.T., Guerrero-Ramírez, N.R., Kattge, J., Kuyper, T.W., Laughlin, D.C., Meier, I.C., van der Plas, F., Poorter, H., Roumet, C., van Ruijven, J., Sabatini, F.M., Semchenko, M., Sweeney, C.J., Valverde-Barrantes, O.J., York, L.M. & McCormack, M.L. (2021) An integrated framework of plant form and function: the belowground perspective. *New Phytologist*, **232**, 42–59.
- Weigelt, P. & Denelle, P. (2022) *GIFT: Access to the global inventory of floras and traits (GIFT)*.
- Weigelt, P., König, C. & Kreft, H. (2020) GIFT – A Global Inventory of Floras and Traits for macroecology and biogeography. *Journal of Biogeography*, **47**, 16–43.
- Weiss, D.J., Nelson, A., Gibson, H.S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T.C.D., Howes, R.E., Tusting, L.S., Kang, S.Y., Cameron, E., Bisanzio, D., Battle, K.E., Bhatt, S. & Gething, P.W. (2018) A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, **553**, 333–336.
- Wenk, E.H., Sauquet, H., Gallagher, R.V., Brownlee, R., Boettiger, C., Coleman, D., Yang, S., Auld, T., Barrett, R., Brodribb, T., Choat, B., Dun, L., Ellsworth, D., Gosper, C., Guja, L., Jordan, G.J., Le Breton, T., Leigh, A., Lu-Irving, P., Medlyn, B., Nolan, R., Ooi, M.,

- Sommerville, K.D., Vesk, P., White, M., Wright, I.J. & Falster, D.S. (2024) The AusTraits plant dictionary. *Scientific Data*, **11**, 537.
- Westoby, M. (1998) A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, **199**, 213–227.
- Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B. & Bourne, P.E. (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, **3**, 1–9.
- Williamson, M., Pyšek, P., Jarošík, V. & Prach, K. (2005) On the rates and patterns of spread of alien plants in the Czech Republic, Britain, and Ireland. *Écoscience*, **12**, 424–433.

Supplementary Information

Supplementary Information S1. Details on trait harmonization

Only AusTraits offered direct correspondence between its trait definitions and TRY, through direct mention of TRY Trait ID and, for certain traits, links to standard definitions provided by the TOP trait Thesaurus (Garnier *et al.*, 2017; Falster *et al.*, 2021; Wenk *et al.*, 2024). For all other databases, we manually matched the names of traits of each pair of databases (giving a total of six correspondence tables). For each trait of an input database, we tried to match trait(s) in the target database based on their associated metadata. We then noted if traits were either identical or related, the latter meaning that they are not exactly measuring the same entity but could be considered similar in a certain context. For example, BIEN reports “leaf area per leaf dry mass”, i.e. LMA, while AusTraits uses a trait named “specific_leaf_area”, which corresponds to SLA. Even if both traits are named differently, they correspond to the same biological measurement as $LMA = 1/SLA$, they can be considered fully identical. TRY reports six different types of SLA/LMA depending if it includes the petiole and/or is done at the leaflet level for compound leaves or if it’s undetermined. At the date of the analyses, no other database differentiated SLA values as much, thus we considered them identical to the “undetermined” value (TraitID in TRY: 3117) of TRY and considered the five other types of SLA as similar but not identical. Another example is that GIFT references “maximum plant height” while BIEN reports “measured plant height”; these two traits can be considered related while not identical. We built a trait name network where nodes are trait names and edges are links between traits, with three categories of links, as in the Austraits Plant Dictionary ontology: exact, if the traits were exactly matching ; close, if the traits were close but not exactly matching ; related, if the traits are related more lightly. The network allowed us to jointly consider all links across all databases. We considered traits the same across databases if they were part of the same connected component in the

network. Each connected component then got assigned a new consolidated name representing the final trait name. We obtained three different trait name networks based on how stringent we were to consider two traits connected:

- The “full” network considered all links (exact, close, and related) to obtain connected components.
- The “close” network considered only exact and close links.
- The “exact” network considered only exact links.

We present only the result of the full network in the article as the results for the two other networks were qualitatively and quantitatively similar. We also focused on the “full” network as it allowed us to cast the widest net possible and show the most optimistic picture for trait data of non-native species.

Supplementary Information S2. Ecological justification of used trait combinations.

Westoby suggested the Leaf Mass per Area-Height-Seed Mass trio to quickly classify the ecology of many species with the fewest traits. We considered three ecologically meaningful trait combinations and assessed how often they were measured across non-native species, that are detailed in the table below:

Table S2-1. Description of the three considered trait sets

Name of trait combination	Source of trait combination	Number of traits	Traits in the set	Justification
Leaf-Height-Seed Mass (LHS)	Westoby, 1998	3	<ul style="list-style-type: none"> ● Leaf mass per area, ● Plant height, ● Seed mass 	Compares the general life history strategy of species at global scale without being context-dependent
Global Spectrum of Plant Form and Function (GSPFF)	Díaz et al., 2016	6	<ul style="list-style-type: none"> ● Leaf mass per area, ● Plant height, ● Seed mass ● Leaf area, ● Stem Specific Density, ● Leaf Nitrogen Content 	Captures the main axes of variation in aboveground plant strategies worldwide
Root traits	Bergmann et al., 2020	4	<ul style="list-style-type: none"> ● Specific root length, ● Root diameter, ● Root tissue density, ● Root nitrogen content 	Positions species along the root economics spectrum and contrasts species along two axes: a “conservation” gradient corresponding to the fast-slow continuum and a “collaboration” gradient corresponding to reliance on mycorrhizal fungi to extract nutrients.

Supplementary Tables and Figures

Table S1. Final trait alignment table for “full” network. See companion CSV file:

<https://drive.google.com/file/d/1IXabqGXkxhdQJCiZBzg1ZyXBCAASP-qO/view?usp=sharing>

Description of fields:

- **component**, number of the connected component the trait name network ordered by size.
- **consolidated_name**, automatically assigned name of the component for understanding.
- **component_size**, size of the connected component in number of nodes (= trait names)
- **austraits_trait_name**, trait name provided as in AusTraits.
- **bien_trait_name**, trait name provided as in BIEN
- **gift_trait_name**, trait name provided as in GIFT corresponding to the “Trait2” level.
- **try_trait_id**, trait id as provided in TRY TraitID column

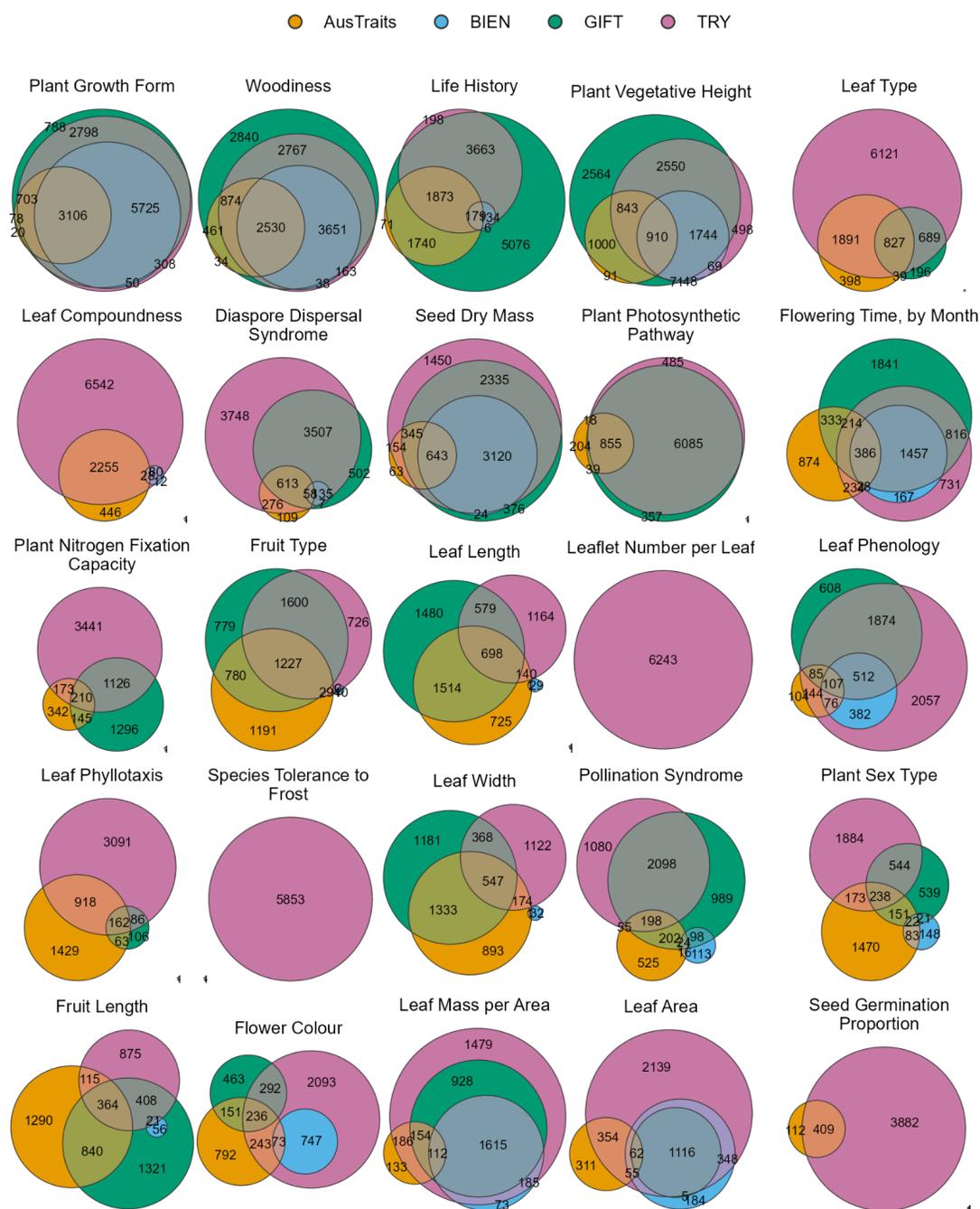


Figure S1. Complementarity of trait databases for the twenty-five most measured traits. Each circle has an area proportional to the number of species covered by a database for this trait, the intersections between circles is also proportional to the number of shared species. The numbers give the number of species in each section. Each database is always represented using the same color across traits: red for AusTraits, blue for BIEN, green for GIFT, and purple for TRY.

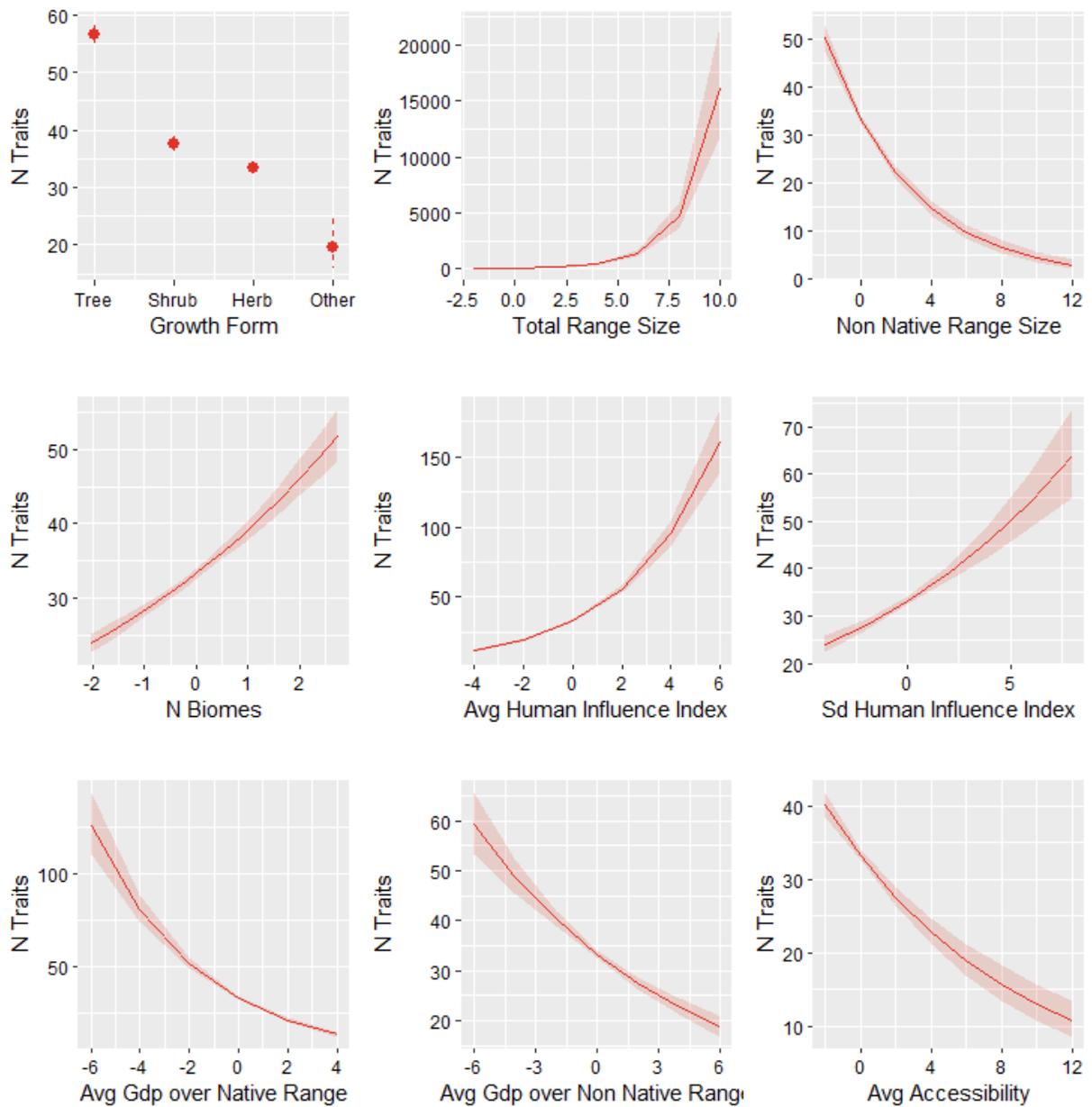


Figure S2. Partial residual plots of the trait knowledge model. These follow the negative-binomial model of the number of traits measured predicted with the 9 variables indicated as the x-axis of each panel (see Methods section for details). It shows the predicted variation of the number of traits in function of the variation of one predictor keeping all the other predictors equal to their mean value (here 0 as they were scaled).

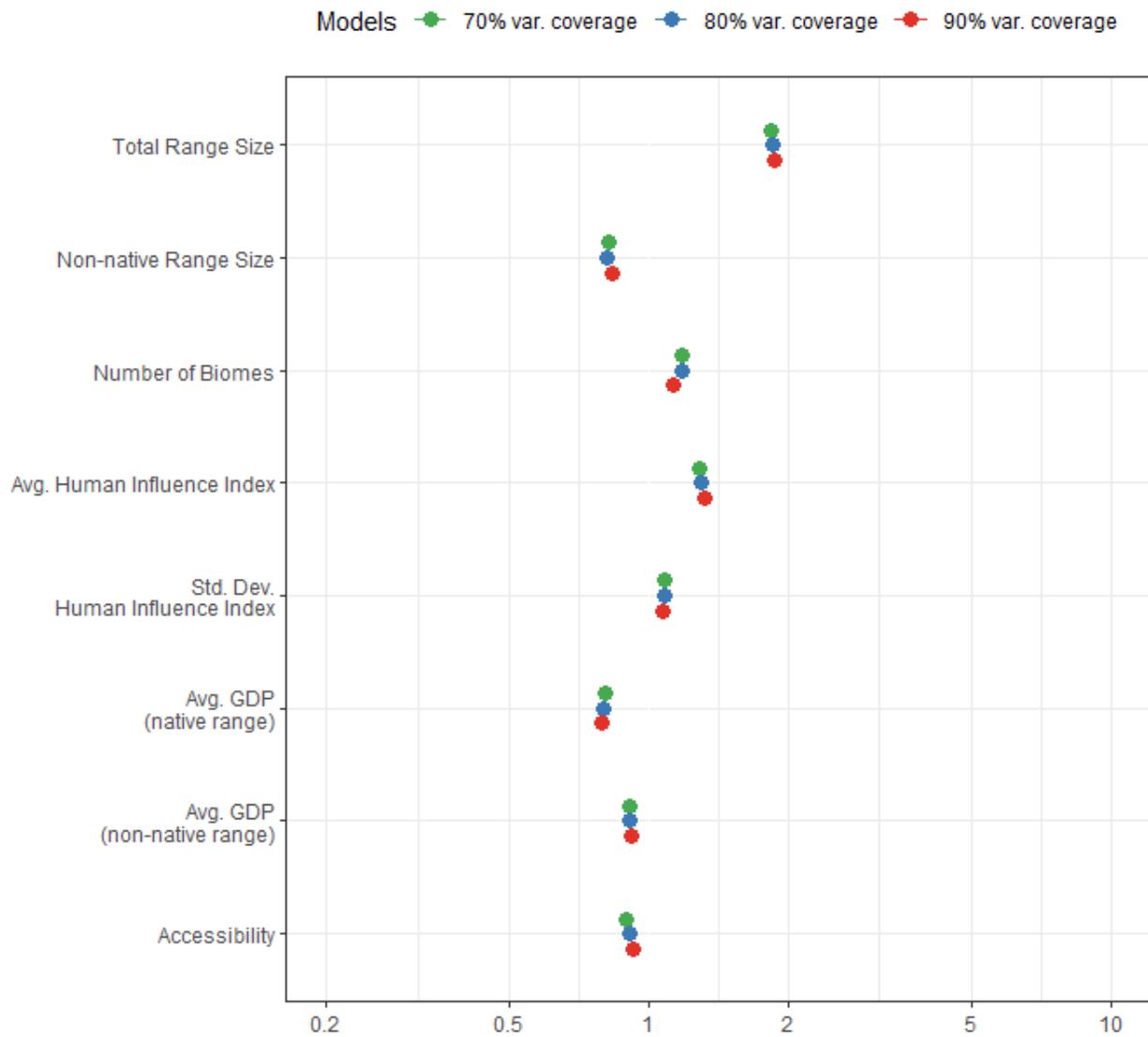


Figure S3. Coefficients of alternative models of number of traits measured by species where variables were available for 70%, 80%, or 90% of the total range of species. The different models are thus fitted on a different number of species.

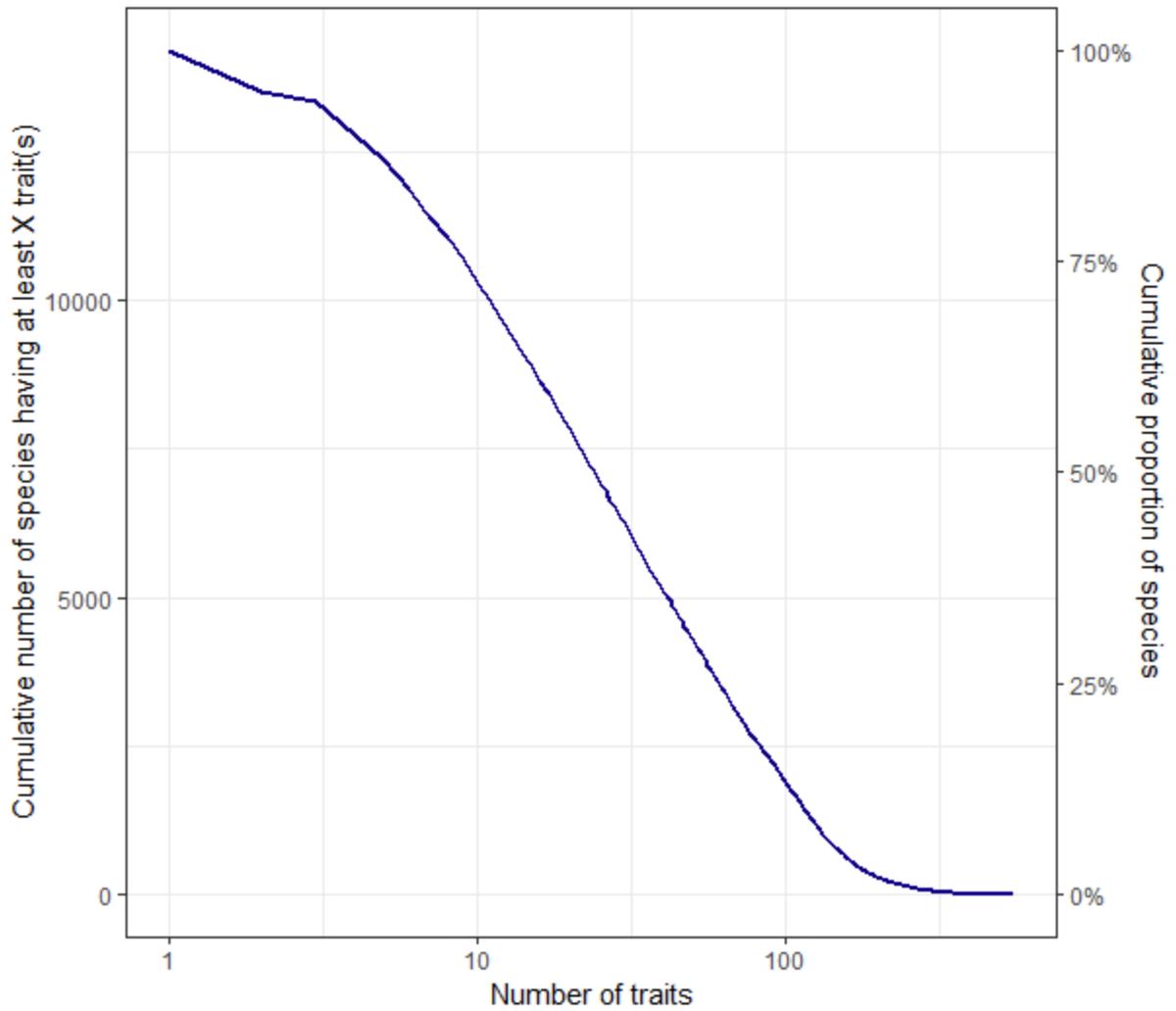


Figure S4. Cumulative number and proportion of alien vascular plant species with at least a given number of traits indicated by the x-axis. The x-axis represents the number of traits. The way to read the graph is the following: “Close to 100% of species have at least 1 trait. More than 10,000 species have at least 10 measured traits”.

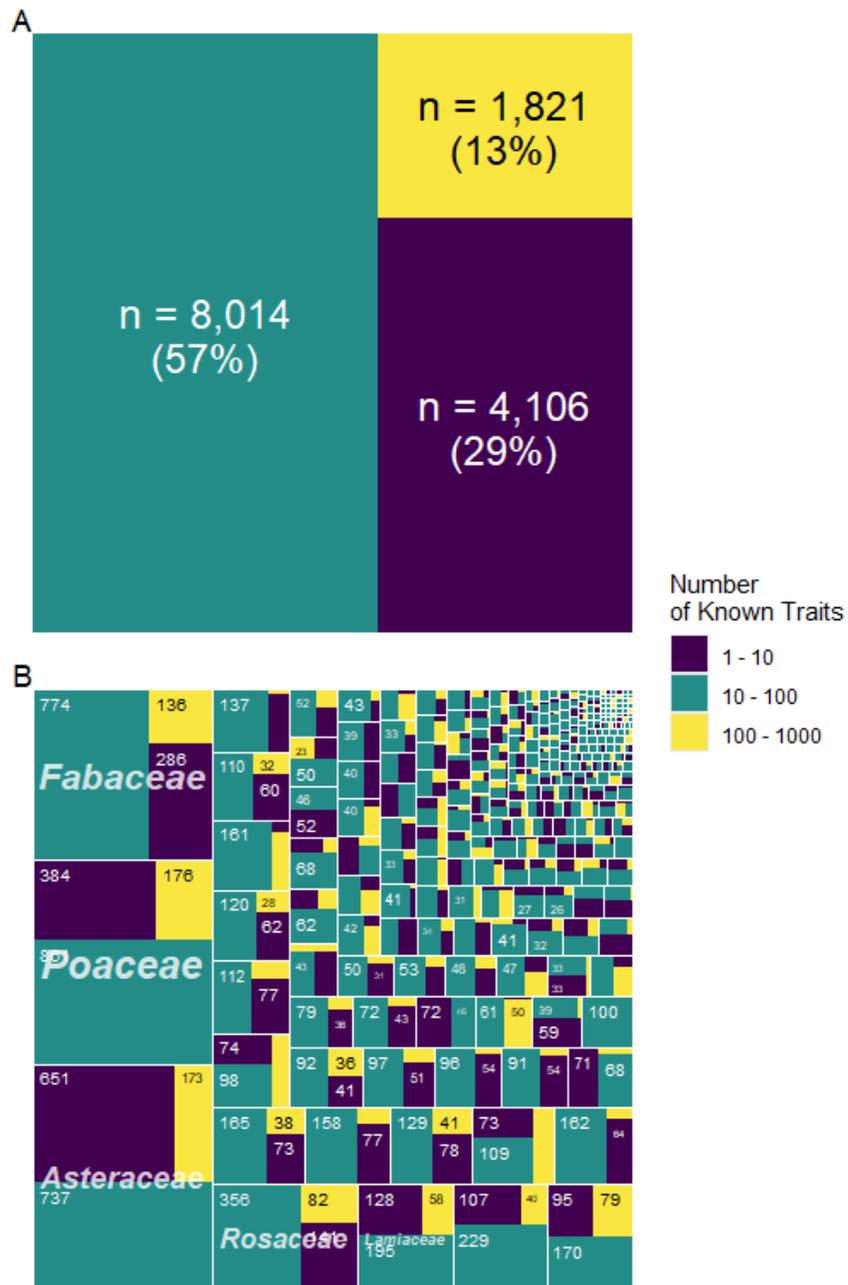


Figure S5. Treemap with number of traits measured per family. **(A)** All species ordered by number of traits with the respective numbers and proportions of species per amount of measured traits. **(B)** Ordered by family. The numbers correspond to the number of species within a family with the given trait knowledge. Way to read the graph: “There are 384 Poaceae species which have 1 to 10 measured traits”.

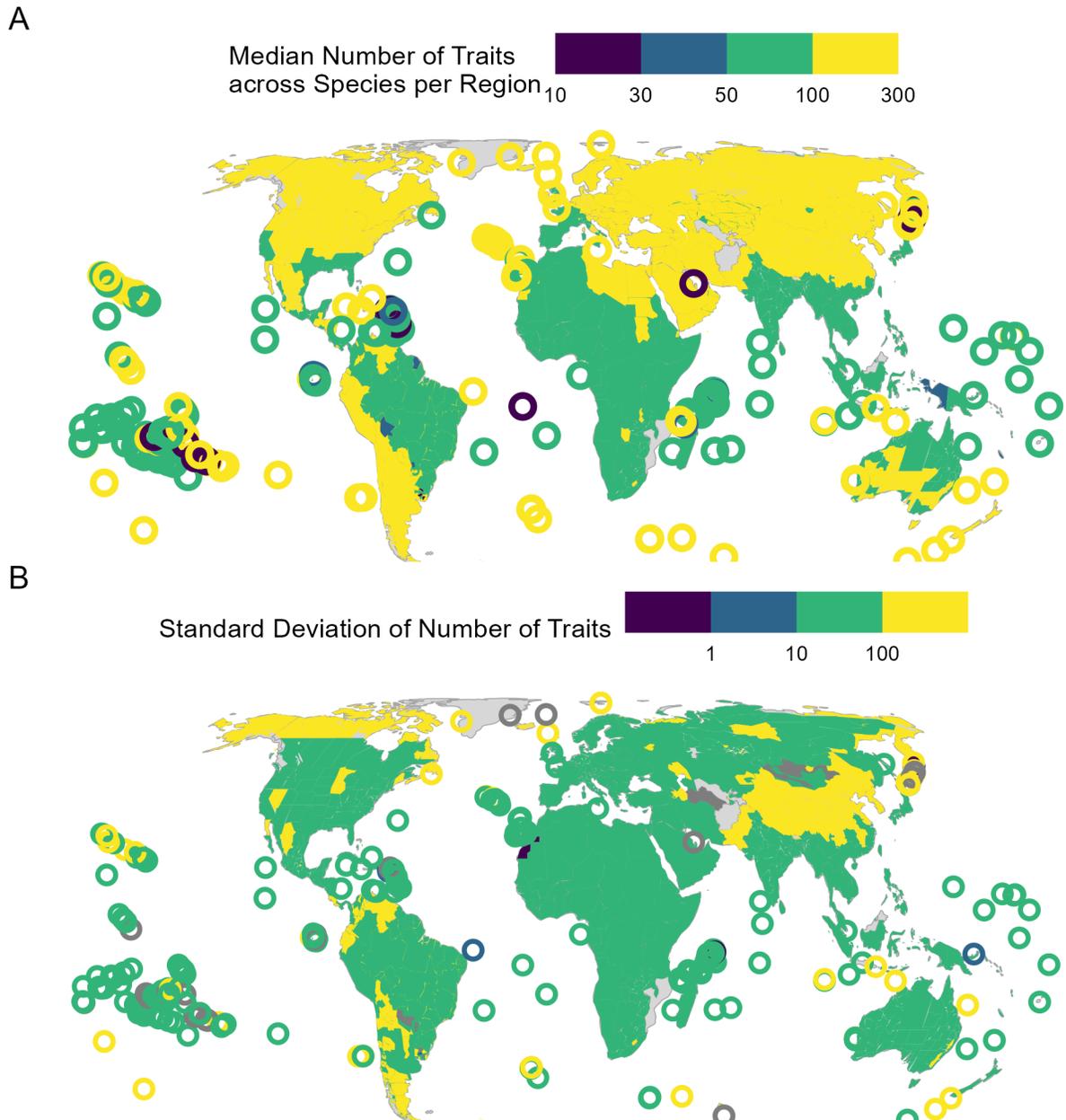


Figure S6. Map showing the median number of measured traits across species per region (**A**) and the standard deviation of the number of traits across species per region (**B**). Note that the scale has been discretized for ease of reading. Empty circles represent islands and are not to scale for readability reasons. Light gray areas show areas with no information on alien vascular plant species, dark gray regions are where there were no variations in number of measured traits per species. The map is projected using the Equal-Earth projection.