Global Ecology and Biogeography

WILEY

Global Ecology and Biogeography — A Journal of Macroecology

RESEARCH ARTICLE OPEN ACCESS

Gaps in Global Alien Plant Trait Data and How to Fill Them

Matthias Grenié[1,2,3] | Helge Bruelheide[3,4] | Wayne Dawson[5] | Franz Essl[6] | Mark van Kleunen[7,8] | Ingolf Kühn[3,4,9] | Holger Kreft[10,11,12] | Petr Pyšek[13,14] | Patrick Weigelt[10,15] | Marten Winter[2,3]

[1]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, Grenoble, France | [2]Leipzig University, Leipzig, Germany | [3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 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, Uiversity of Liverpool, Liverpool, 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, Konstanz, Germany | [8]Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, Taizhou University, Taizhou, China | [9]Dept. Community Ecology, Helmholtz Centre for Environmental Resaerch – UFZ, 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, Průhonice, Czech Republic | [14]Department of Ecology, Faculty of Science, Charles University, Prague, Czech Republic | [15]Department of Environmental Science, Radboud Institute for Biological and Environmental Sciences, Radboud University, Nijmegen, the Netherlands

Correspondence: Matthias Grenié (matthias.grenie@univ-grenoble-alpes.fr)

Received: 17 October 2024 | **Revised:** 27 August 2025 | **Accepted:** 15 September 2025

Handling Editor: Phoebe Zarnetske

Funding: This work was supported by Akademie V d esk Republiky, RVO 67985939; Deutsche Forschungsgemeinschaft, 264740629, DFG- FZT 118, 202548816, Research unit FOR2716 DynaCom; Grantov Agentura esk Republiky, EXPRO grant no. 1928807X; Austrian Science Fund, pr. no. I 5825B.

Keywords: functional traits | missing trait | naturalised plants | non-native plants | raunkiaerian shortfall | trait database

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 naturalised 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 harmonisation. 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 (2764 traits for 14,539 species) was only 1.5% complete, with most traits measured for very few species. Only ten traits were available for more than 50% of all alien plants. Four percent of the species lacked all trait data, while 27% of species had data for the three key plant traits: leaf mass per area, seed mass, and plant height. We observed a strong latitudinal gradient in trait knowledge, with tropical regions showing lower trait knowledge than higher latitudes, particularly in the Northern Hemisphere. Growth form, range size, and invasion status were the strongest predictors of trait knowledge, with widespread, invasive tree species being better recorded than other alien species.

Main Conclusions: We identified large trait data gaps at a global scale for alien plants, which limits our ability to study functional invasion ecology at large spatial scales. These gaps are partly driven by uneven sampling and a lack of trait data integration across sources. We recommend prioritising the most invasion-relevant traits and coordinating community efforts of plant and invasion scientists to sample them in a standardised way, which could help close these gaps.

1 | Introduction

Biological invasions are an important component of global changes (IPBES 2023); we need a better understanding of their underlying ecological processes for effective management (Díaz et al. 2019; IPBES 2023). Plant invasions are known to cause many long-lasting impacts on terrestrial and aquatic ecosystems, above and below ground. Consequences include, for example, severely changing ecosystem processes, leading to biotic homogenisation of formerly distinct biotas, changing local abundance and species richness of outcompeted native biota (Pyšek et al. 2012; D'Antonio and Flory 2017; Kumar Rai and Singh 2020; Lázaro-Lobo et al. 2023; Dostál 2024). Global invasion of plants led to approximately 4% of the global flora being established outside of its native range (van Kleunen et al. 2015), becoming alien (i.e., non-native, exotic) somewhere on this planet. Identifying plant traits promoting successful species introductions, establishment and dispersal is one of the major aims of plant invasion ecology (e.g., Pyšek et al. 2008; Drenovsky et al. 2012; Knapp and Kühn 2012; van Kleunen et al. 2015; Gallien and 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-fertilise are more likely to become naturalised than outcrossing species (Razanajatovo et al. 2016).

Despite massive collaborative mobilisation 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 embrace the context-dependency of traits in biological invasions (Catford et al. 2022; Milanović et al. 2025), traits should be considered in conjunction to use them to manage alien species (Küster et al. 2008; Pyšek et al. 2020). Traits have been used to assess the risk of invasion of species, as well as the economic costs of invasion (Keller and Drake 2009; Palma et al. 2021; Cuthbert et al. 2025). For example, species with more acquisitive strategies, with small seeds and high SLA, have been shown to be more invasive (Palma et al. 2021). Thus, the trait availability of alien species is key to using traits to explain idiosyncrasies associated with different invasion syndromes (Küster et al. 2008; Novoa et al. 2020) and across environmental gradients (Golivets et al. 2024). Although one of the holy grails of invasion ecology, there has been no universal set of functional traits identified as being significantly associated with successful plant invasions across different invasion stages, habitats, regions, climates, propagule pressure, and residence time (Gioria et al. 2023). Traits and trait states associated with invasion success differ across studies and are always dependent on the availability of trait measures, which is limited as we show here.

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 naturalised range) have more trait data present in databases than species in poorer countries. We also expect invasive (i.e., species with ecological and/or negative economic impacts) species to have data available for more traits than naturalised 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 the 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 a 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.

2 | Methods

2.1 | Alien Plants List

We extracted a list of alien vascular plants from the Global Naturalised 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 labelled as 'naturalised' or 'invasive'. Our list of alien species comprised 16,044 taxonomic names of vascular plants before taxonomic harmonisation. We also extracted the naturalised geographic distribution (at TDWG4 resolution level) of each taxon from the GloNAF database.

2.2 | Taxonomic Harmonisation

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 standardise species names, we performed a full taxonomic harmonisation workflow (Grenié et al. 2022). For the four trait databases as well as GloNAF, we referred to the raw available names with authors and subspecific epithets if available (i.e., the name from the original source). We leveraged the speed and reliability of the Taxonomic Name Resolution Service (TNRS, v.5.1 Boyle et al. 2021, Boyle et al. 2013) with its R package TNRS v.0.3.6 (Maitner 2024) to match all of these names against the World Checklist of Vascular Plants (Govaerts 2023). We set TNRS to only return the best match. To merge all datasets, we only retained the accepted binomial names for all species. In the end, we had 14,073 matched species names between GloNAF and all trait databases (see Data S2 and Figure S1).

2.3 | 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. We didn't consider the geographical provenance of our data, though available for all of AusTraits and BIEN data, because it is only available for 42% of TRY data (Kattge et al. 2020), and isn't easily tractable in GIFT as the trait measurements are coming from floras.

2.3.1 | 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.

2.3.2 | 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.

2.3.3 | 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.

2.3.4 | TRY

We queried all publicly available traits in TRY v6.0 (Kattge et al. 2020). We obtained data for 301,799 species and 2460 traits.

A list of all of the used original data sources is found in Data S1.

2.3.5 | Aligning Common Trait Definitions

We created a single species-by-trait matrix from all trait databases after harmonising the traits across them (see details in Data S2) 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.

2.3.6 | 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 with the TRY SLA trait with petiole excluded (TRY:3115), closely matching with the TRY SLA trait with petiole, midrib, and rachis excluded (TRY:3086), and related to two other TRY SLA traits where the petiole was included (TRY:3116) and where it is undefined if the petiole was or was 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 the data supplements.

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 3351 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 2764 unique traits. In the end, in the 'full' trait network, our combined trait dataset contained 14,063 species (after taxonomic harmonisation) and

2250 observed traits as 514 traits were never observed across our set of target species.

2.4 | 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 Data S2 for a list of traits and extended justification.

2.5 | Modelling 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 and 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 hypothesised 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 to occur 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.

2.5.1 | Final Data Subset

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

2.5.2 | Statistical Model

Our response variable was the number of traits measured per species out of our theoretical maximum of 2764 traits; we thus performed a negative binomial generalised linear model that we fit using the glm.nb() function in the MASS package (Venables and 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 the 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 a 0 mean and scaled to a 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 or zero-inflation. We used Nagelkerke's pseudo-R-square for GLMs through the r2_nagelkerke() function from the performance package.

2.5.3 | Phylogenetic Model

Because our trait knowledge model was species-based, we wanted to test the effect of adding phylogenetic correction to the model. We fitted a Poisson phylogenetic regression model, based on the same predictors as our non-phylogenetic model, using the phyloglm() function in the phylolm package (Ho and Ane 2014). We provided a phylogenetic tree of all of our non-native species, assembled through the rtrees package (Li 2023) using a reference global plant phylogeny (Smith and Brown 2018).

All data extraction and analyses were done using R 4.2.2 (R Core Team 2022).

3 | Results

3.1 | A Sparse Species × Traits Matrix

We obtained a species-traits table containing observations for 14,063 alien vascular plant species (out of 14,539 known at a global scale) and 2250 different traits (out of 2764) (Figure 1A). This gives a theoretical maximum of 40,185,796 possible species by trait combinations, while our observed matrix only contained 622,513 (1.5%) of them. Narrowing on the 200 most frequently measured traits (Figure 1B) gives 2,907,800 combinations, of which 504,234 (17.3%) are observed.

We listed 25 most frequently measured traits. More than 90% of alien plant species had data available for each of the following three traits: growth form, woodiness, and life history (Figure 1C). Close to 75% 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. Five other traits were available for more than half of the species: leaf compoundness (compound or simple leaf), dispersal syndrome (anemochorous, etc.), seed mass, photosynthetic pathway (C3, C4, or other), and flowering phenology. The remaining 15 traits are available for less than half of the species. Most of those traits describe

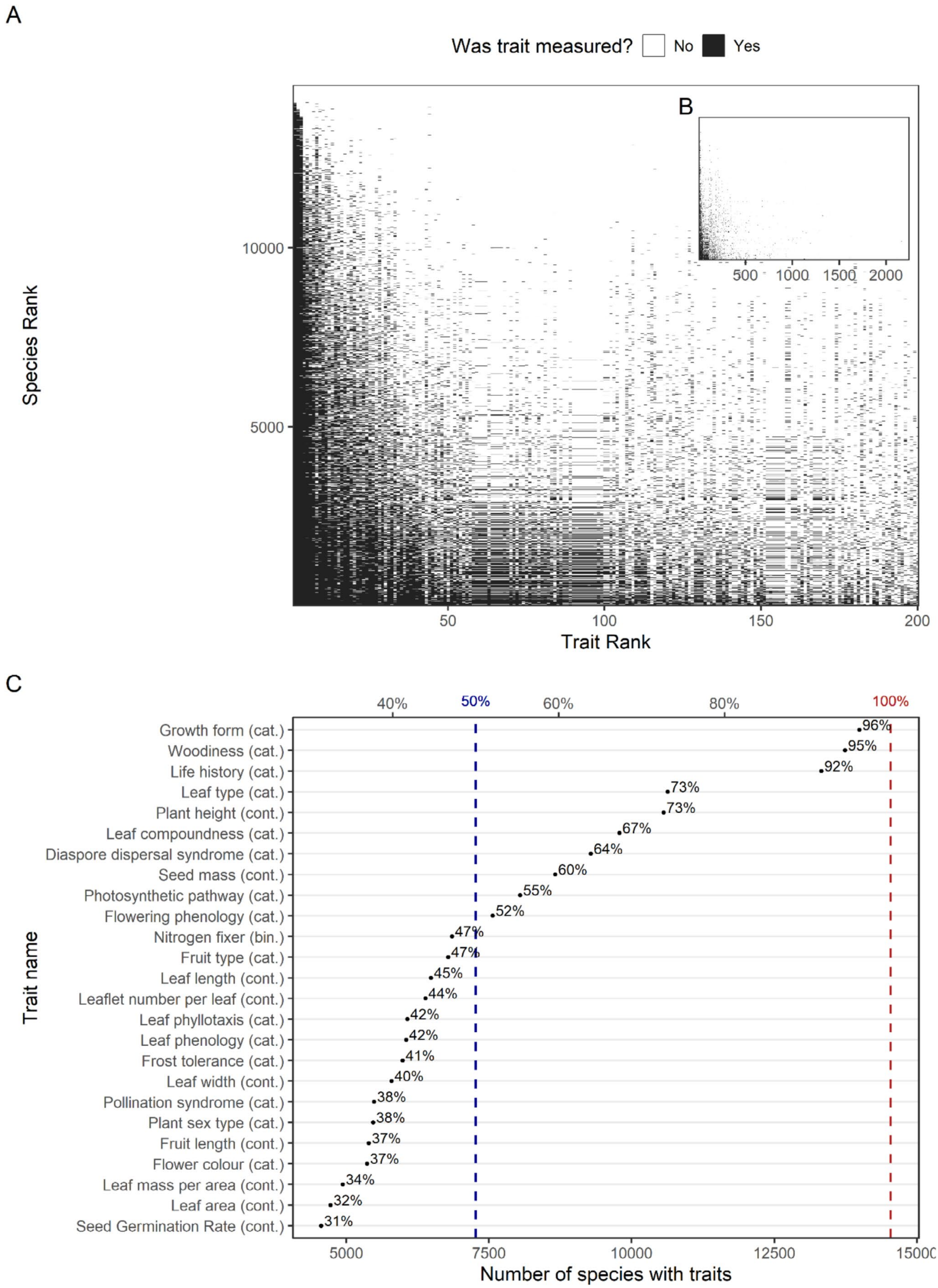

FIGURE 1 | Legend on next page.

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 colour of the pixel shows whether the trait was ever measured for this species (dark grey) or never (white). (B) shows the same figure with all 2250 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 (14,539 in our dataset), while the blue dashed line represents 50%.

fundamental ecological characteristics, 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 colour, leaf mass per area, leaf area, and seed germination rate. Among the 25 most frequently measured traits, the different databases had various unique contributions (Figure S2). For plant growth form (available for 13,991 species), TRY covered 13,622 species, and 3309 species were found in common in all four databases; but GIFT was the greatest provider of unique species-trait observations unavailable in other databases (290 species out of 13,991). The contribution of AusTraits for plant growth form, though smaller, mostly brought unique observations (39 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 (114 species uniquely in BIEN versus 559 in AusTraits, 948 in GIFT, and 985 in TRY for a total of 5487 species). These contributions only consider unique species-trait combinations across databases and do not consider the availability of different trait measurements for the same species.

3.2 | Uneven Proportion of Measured Combination of Traits

We mapped species for which we had the measured combinations of traits (see details in Table S1): Leaf-Height-Seed Mass (LHS), aboveground spectrum traits, and root traits. For only 3.3% of alien species, not a single trait measurement exists (476 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. 29.1% of the species with non-native occurrences have a measured LHS combination (4227 species), 14.3% have the aboveground spectrum traits fully described (2079 species), and root traits are measured for only 1.8% (266 species). When jointly considering multiple combinations, the numbers drop further: 1.7% of species (244 species) have measured LHS and root traits, while 1.5% of species (216 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 a lack of trait combinations for most alien species.

3.3 | 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.

3.4 | Traits of Widespread Invasives Are Measured More Often

The species' invasion status influenced the knowledge of trait combinations (Figure 4A). Non-native species referred to as invasive in at least one region in GloNAF ('invasive' in this section) had a greater coverage in all trait combinations than non-native species never reported as invasive ('non-invasive' in this section), which themselves showed higher coverage than species never referenced as non-natives ('natives' in this section). We found strong evidence that LHS traits are more frequently measured for invasive species (48.2% of species) than for non-invasive species (24.8%; $\chi^2 = 325$, df = 1, $p < 0.001$). 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%). The number of traits available per species followed the same pattern; invasive species had 74.8 traits available on average, while non-invasives had 34.7 and native ones 7.44 (All pairwise *t*-tests showed $p < 0.001$).

We also observed a difference in trait knowledge depending on the geographical spread of species. The 100 most widespread species in GloNAF consistently showed higher trait-combination

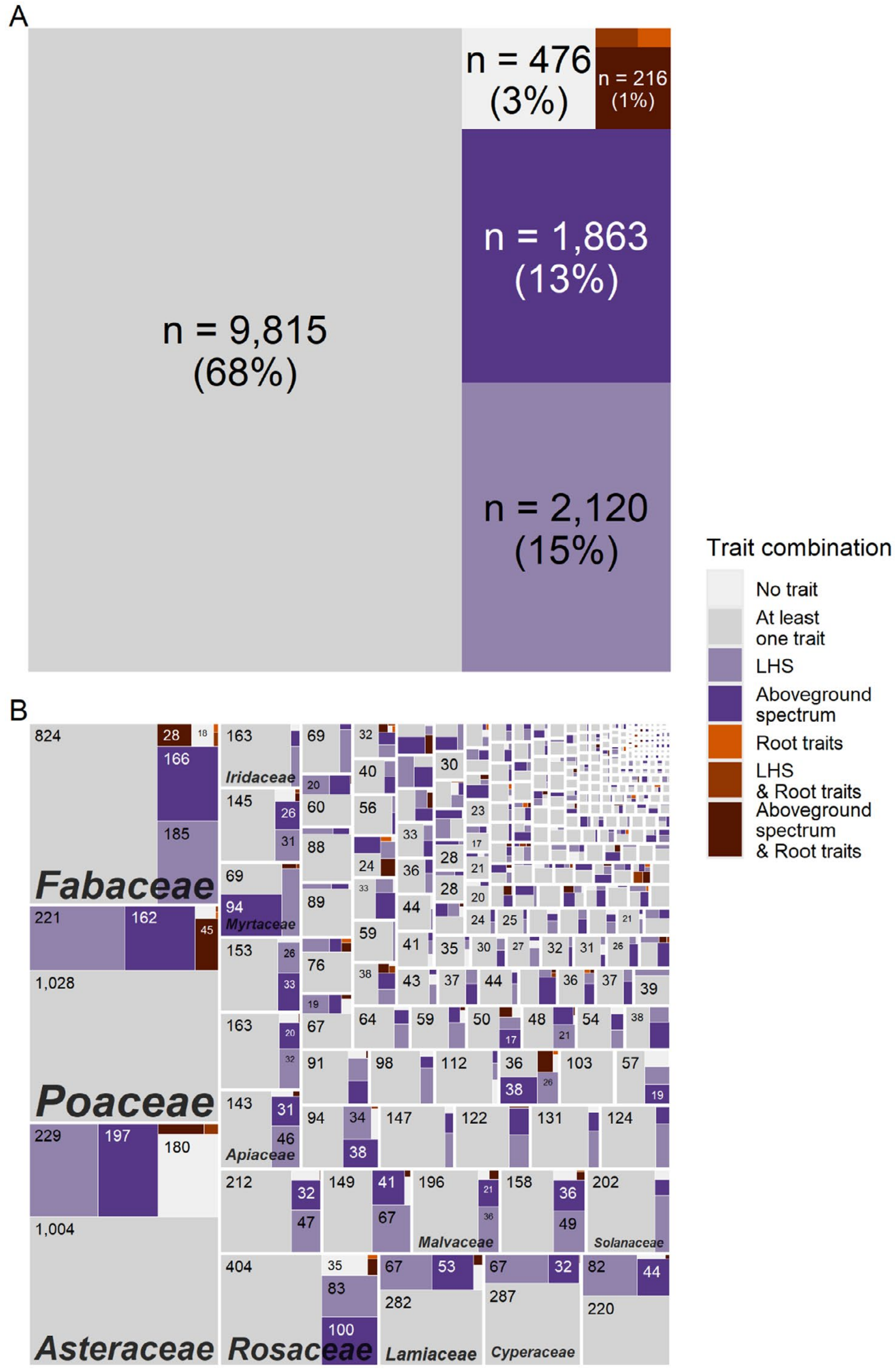

FIGURE 2 | Shares of alien vascular plant species covered by different trait combinations for all alien species (A) and per plant family (B). The area of rectangles is directly proportional to the number of species in each category. The numbers show the number of species and in (A) additionally respective proportion of species compared to the global number of alien species (14,539). In (B), the blocks contained within thick white borders represent botanical families 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 coloured depending on measured trait combinations. The colours correspond to different trait combinations: light grey, no measured trait; grey, 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. For ease of navigation and reading, an interactive online version of this figure is available at: https://rekyt.github.io/alientraitgaps/, archived for long-term on Zenodo doi: 10.5281/zenodo.13940200.

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, $p = 0.45$).

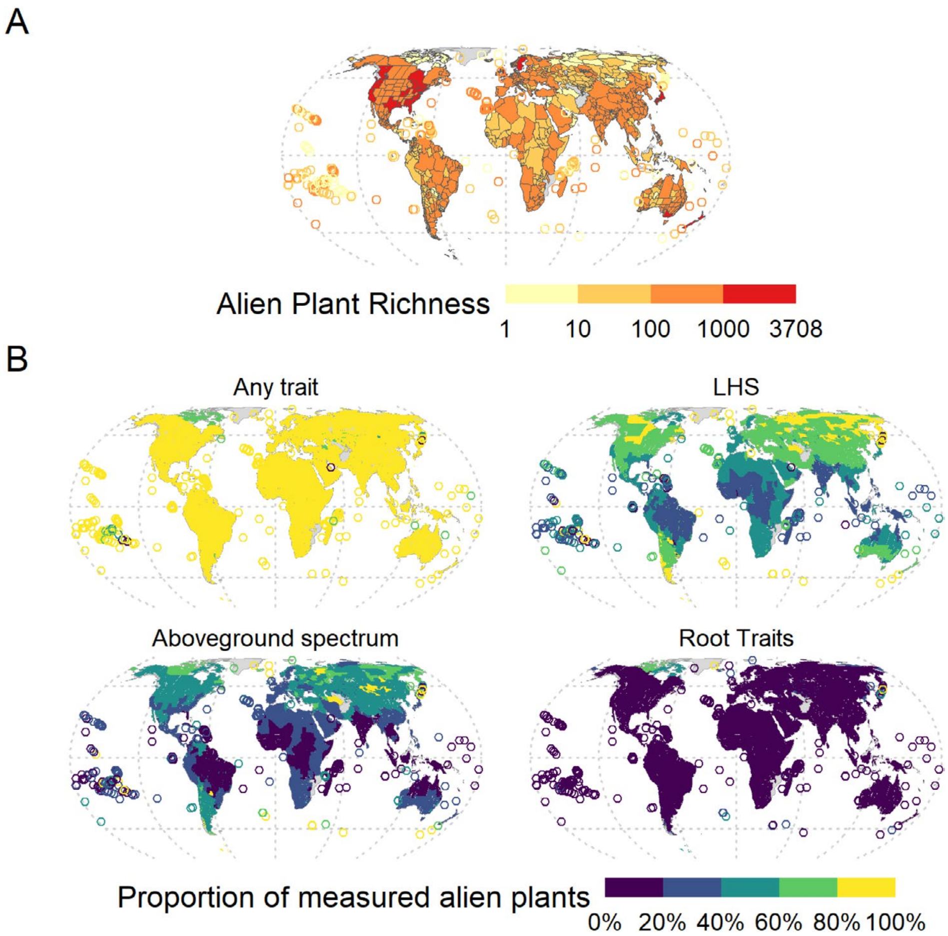

FIGURE 3 | Maps of alien vascular plant species richness and proportion of aliens with measured trait combinations. (A) Alien species richness based on GloNAF. For readability reasons, the richness scale has been discretised in four colours following a log10 scale. Grey areas show where no data were 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).

3.5 | Plant Growth Form and Range Size Predict Trait Knowledge

We modelled 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 S3). Our model had a Nagelkerke's pseudo-R^2 of 62.7%. 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 of 10 increase in the range (in km^2), there was an 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%). The analyses were performed considering species for which the predictors were available for at least 80% of their total range; we obtained similar results when performing the same analyses with a threshold of 70% and 90% (Figure S4). Considering the phylogeny in the model didn't affect the direction of the effect of all of the variables, which all remained with $p < 0.001$ (Figure S5).

4 | 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

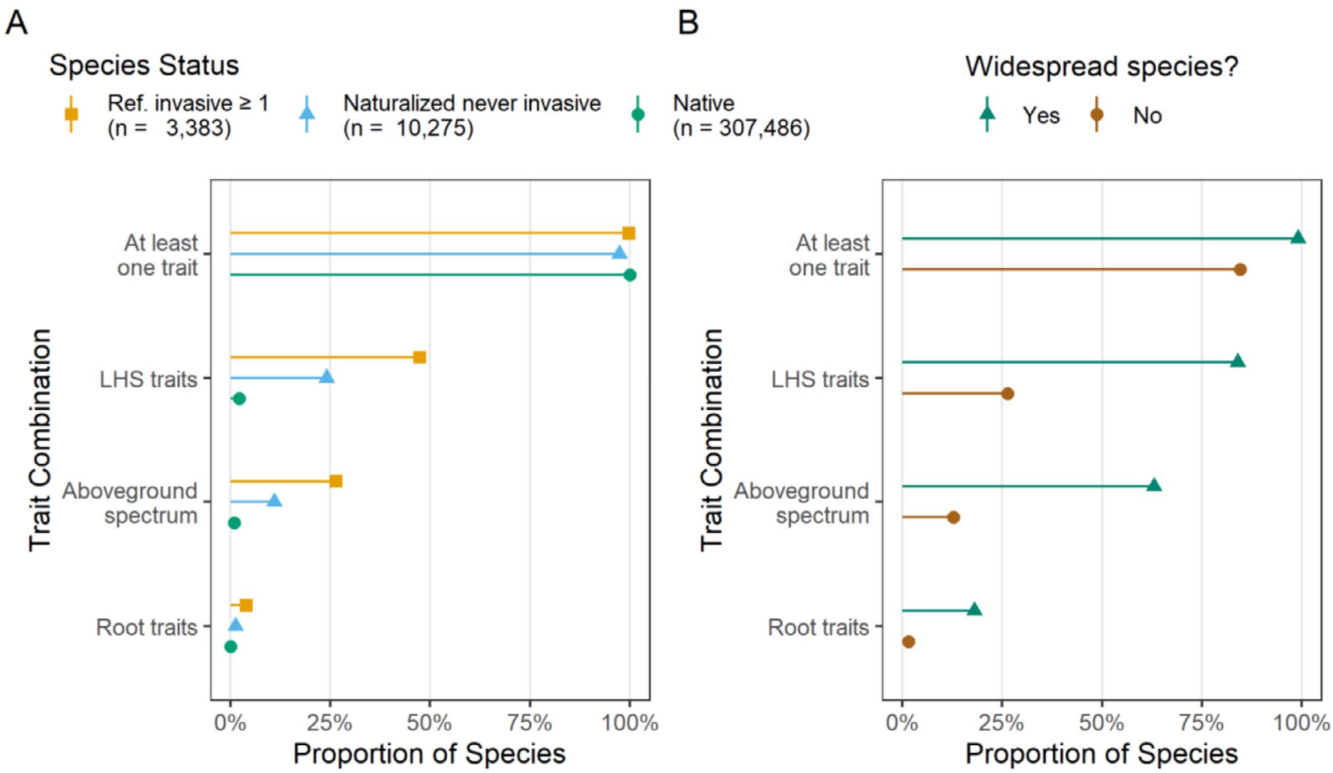

FIGURE 4 | (A) Effect of invasiveness on measured trait combinations across species. The *x*-axis shows the proportion of species with the given trait combination. The *y*-axis shows the different trait combinations. Shapes and colours distinguish species based on their invasion status: orange squares are for species mentioned as invasive at least once, blue triangles for the ones never mentioned as invasive, and green circles for the natives ones. (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 colours 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.

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.

4.1 | Biases in Trait Knowledge

Several dimensions of biodiversity are known for showing strong geographic, taxonomic, and trait biases (Tyler et al. 2012; Hortal et al. 2015; Cornwell et al. 2019; Webb and Vanhoorne 2020; Hughes et al. 2021; Rudbeck et al. 2022). Traits of alien plants are no exception. The lack of trait data, the so-called Raunkiærian shortfall, is characterised 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 S6), only a 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 S7). Third, regarding the biogeographic bias, we identified a strong latitudinal gradient in trait knowledge (Figure 3, Figure S8), 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 of their traits. Accounting for these trait, taxonomic, biogeographic, and spatial biases requires careful analyses. They call for greater attention to data collection, mobilisation, and integration to compensate for biases (See 'How to fill the trait data gap' section).

As we expected, we found that non-native plants with larger ranges and occurring in more biomes had more traits measured across the databases. Opposite to our expectations, we found negative relationships between the number of traits measured and the average GDP in countries in both their native and non-native ranges. While average GDP should correlate with research effort, and as such collection effort, this negative relationship could be due to the relationships between average GDP over species ranges and the area of their ranges. Large-range species, which tend to have more traits measured, will show lower average GDP over their ranges. Small-range species may occur over higher GDP areas, but show a lower number of traits measured because of their overall smaller range. These findings call for additional studies on the determinants of trait knowledge for both native and non-native plants.

4.2 | Trait Relevance

We decided to focus on commonly used and clearly defined trait combinations, namely LHS traits from Westoby (1998),

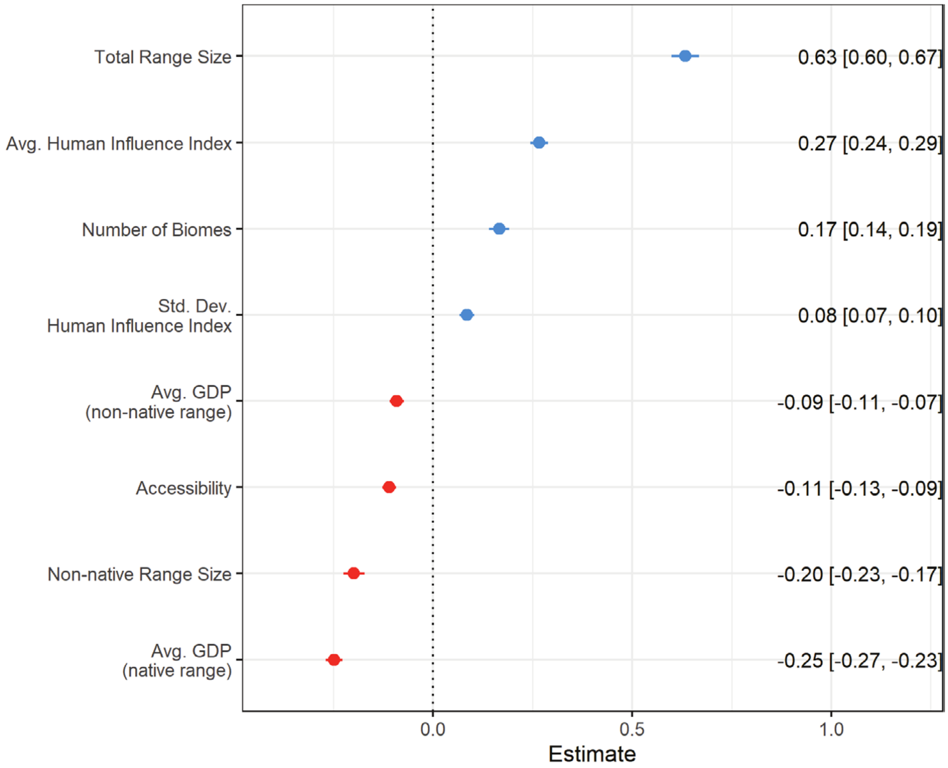

FIGURE 5 | Summary plot showing the incidence rate ratios for variables explaining number of traits measured per alien vascular plant species. Variables with blue dots increase the number of traits measured per species, while variables in red decrease it. All tested variables showed $p < 0.0001$.

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, that is, 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 many 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 naturalised species' trait variability requires measuring them in both the non-native and the native range.

We here made the simplifying assumption that all trait measurements were perfectly recorded, with no measurement nor reporting errors. Considering these errors would certainly reduce even further our trait knowledge. It was recently shown for the TRY database that only 23% of the original SLA measurements from TRY were actually original, representative, logical, comparable, and traceable (Augustine et al. 2024). While we know the ecological importance of intraspecific trait variation for plants (Westerband et al. 2021), we also simplified our trait matrix by considering any single trait measurement for a single species enough to know the trait value for the species. Our study could be further extended by studying the number of trait measurements known for each trait and each species to estimate how well we know the intraspecific variation for each species.

4.3 | 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, both of which are criteria to follow FAIR principles (i.e., Findability, Accessibility, Interoperability, and Reusability) in data stewardship (Wilkinson et al. 2016; Keller et al. 2023). One of the challenges we faced in our study was to combine data from heterogeneous trait sources. The four databases we used are complementary in terms of species coverage, and using all four increased data coverage (Figure S2) 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 and 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 emphasises 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 are available as Supporting Information in the hope that it would be useful for other studies. Community-developed trait correspondence schemes, for example through the OpenTrait Network (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 or looking for patterns, an expert-driven unifying global plant trait correspondence scheme is the only way to minimise errors in those automatic processes.

4.4 | How to Fill the Trait Data Gap?

4.4.1 | Prioritising Trait Acquisition

It seems unrealistic to expect all trait gaps to be filled with in situ measurements in the near future. Given the immense diversity of the plant traits reported here (more than 2764 different traits), prioritising the most commonly studied traits would seem more tractable. Adopting a prioritisation framework similar to the one used in conservation biology would be more realistic (Arponen 2012). Prioritisation schemes use well-defined criteria on species, traits, or regions to target data sampling or data integration and increase their trait coverage. The prioritisation depends on the aims and purpose of the sampling.

Any prioritisation 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 a note on the invasion status of the species. For example, we could prioritise 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 prioritisation, 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 suggests a need for a geographic prioritisation scheme. Areas richer in alien species could be targeted, as these are more likely to harbour many invasive species (Chytrý et al. 2012) and suffer from the impacts of invasion. Another region-based approach would prioritise 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 prioritise species/traits/locations that would minimise the error from trait imputation methods (Penone et al. 2014; Schrodt et al. 2015; Joswig et al. 2023). Then species and traits would be prioritised to reduce the uncertainty of the imputation the most. For example, we could prioritise species from families where only a few species have been sampled.

4.4.2 | Closing the Trait Gaps

Once species, traits and locations have been prioritised, 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: mobilisation 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).

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 mobilising 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 mobilised 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.

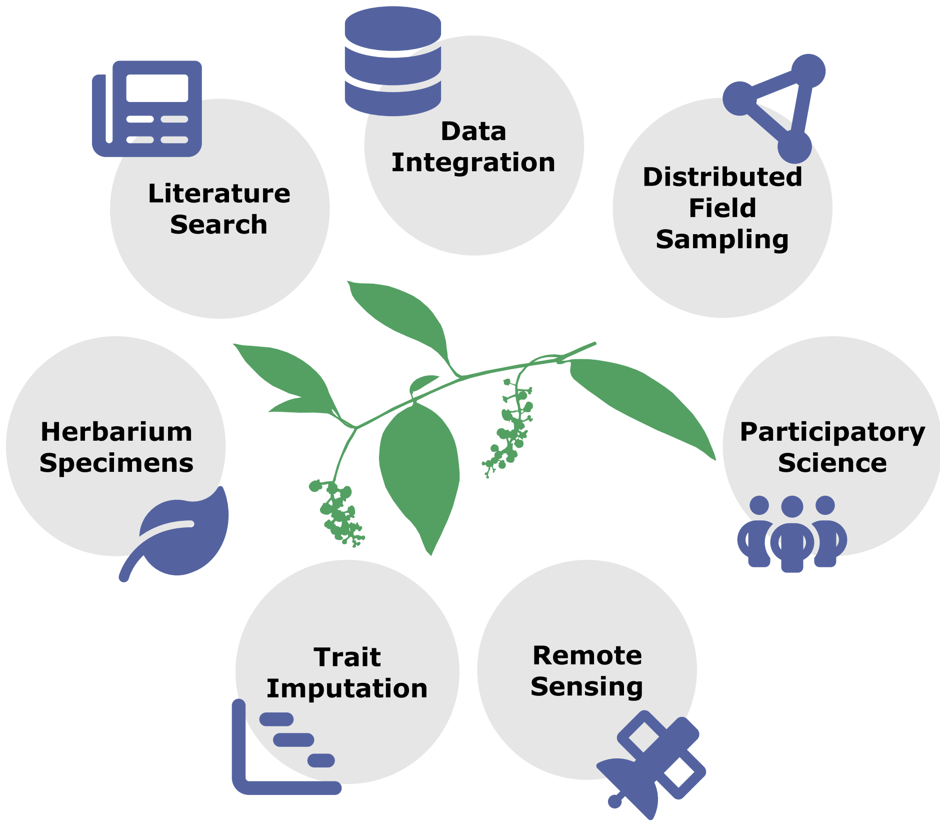

FIGURE 6 | Strategies to increase alien species trait knowledge.

Distributed field campaigns could help to acquire a few traits of alien species. After trait and species prioritisation, 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 a 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. There is, though, a series of 'Functional Plant Trait Courses' organised by V. Vandvik and B. Enquist who organised several campaigns to acquire additional trait data (https://plantfunctionaltraitscourses.w.uib.no/).

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 and Asner 2009; Niphadkar and Nagendra 2016). This approach is limited to traits that can be remotely sensed for species occurring in open areas (Niphadkar and 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 mobilised for trait data. In particular, because many herbaria provide digitised specimens, it would be possible to acquire trait data semi-automatically from these images (Davis 2022).

The above-mentioned strategies help fill the trait gaps by acquiring new data. Trait imputation (also known as trait gap filling) is a complementary strategy that leverages trait correlations as well as additional data (whether spatial and/or phylogenetic depends on the exact method) to infer the trait values for species with missing values (Schrodt et al. 2015; Joswig et al. 2023). Trait imputation should be performed carefully, considering the strengths and weaknesses of the different imputation methods as well as the ecological context of the original trait measurements used to fit the imputation models (Penone et al. 2014; Johnson et al. 2021; Blomberg and Todorov 2025; Gorné et al. 2025).

Most solutions stated above require collective work from plant, invasion, and remote sensing scientists, as well as funding schemes which focus on pure data collection campaigns, which rarely exist. We want to emphasise 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 goodwill 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 recognised as they create a basis for progress in research.

5 | Conclusion

We identified large trait gaps for alien plant species at a global scale. These gaps are partly driven by uneven sampling and missing integration of data. With the distributed efforts of the global community of plant and invasion scientists, these gaps can be reduced. Our suggestions should encourage efforts to harmonise 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.

Author Contributions

M.G. and M.W. conceived the ideas and designed the methodology; M.G. compiled and analysed the data; M.G. and M.W. led the writing of the manuscript. All authors discussed approaches and intermediate results and contributed critically to the drafts and gave final approval for publication.

Acknowledgements

M.G., H.B., I.K. 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. P.P. was supported by EXPRO grant no. 19-28807X (Czech Science Foundation) and long-term research development project RVO 67985939 (Czech Academy of Sciences). M.K. acknowledges support of the German Research Foundation DFG (grant 264740629). F.E. appreciates funding by the Austrian Science Fund (FWF; pr. no. I 5825-B). H.K. acknowledges funding by German Research Foundation DFG (Research unit FOR2716 DynaCom). We appreciate the constructive comments of two anonymous reviewers and the Handling Editor (Phoebe Zarnetske) to a previous version of this manuscript. Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

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). Our interactive online Supporting Information figures are available at https://rekyt.github.io/alientraitgaps/.

References

Arponen, A. 2012. “Prioritizing Species for Conservation Planning.” *Biodiversity and Conservation* 21: 875–893.

Augustine, S. P., I. Bailey-Marren, K. T. Charton, N. G. Kiel, and M. S. Peyton. 2024. “Improper Data Practices Erode the Quality of Global Ecological Databases and Impede the Progress of Ecological Research.” *Global Change Biology* 30: e17116.

Berendsohn, W. G. 1995. “The Concept of ‘Potential Taxa’ in Databases.” *Taxon* 44: 207–212.

Bergmann, J., A. Weigelt, F. van der Plas, et al. 2020. “The Fungal Collaboration Gradient Dominates the Root Economics Space in Plants. *Science*.” *Advances* 6: eaba3756.

Blackburn, T. M., F. Essl, T. Evans, et al. 2014. “A Unified Classification of Alien Species Based on the Magnitude of Their Environmental Impacts.” *PLoS Biology* 12: e1001850.

Blomberg, S. P., and O. S. Todorov. 2025. “The Fallacy of Single Imputation for Trait Databases: Use Multiple Imputation Instead.” *Methods in Ecology and Evolution* 16: 658–667.

Borer, E. T., W. S. Harpole, P. B. Adler, et al. 2014. “Finding Generality in Ecology: A Model for Globally Distributed Experiments.” *Methods in Ecology and Evolution* 5: 65–73.

Boyle, B., N. Hopkins, Z. Lu, et al. 2013. “The Taxonomic Name Resolution Service: An Online Tool for Automated Standardization of Plant Names.” *BMC Bioinformatics* 14: 16.

Boyle, B. L., N. Matasci, D. Mozzherin, et al. 2021. “Taxonomic Name Resolution Service, Version 5.1.”

Catford, J. A., J. R. U. Wilson, P. Pyšek, P. E. Hulme, and R. P. Duncan. 2022. “Addressing Context Dependence in Ecology.” *Trends in Ecology & Evolution* 37: 158–170.

Cherif, E., H. Feilhauer, K. Berger, et al. 2023. “From Spectra to Plant Functional Traits: Transferable Multi-Trait Models From Heterogeneous and Sparse Data.” *Remote Sensing of Environment* 292: 113580.

Chytrý, M., J. Wild, P. Pyšek, et al. 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., W. D. Pearse, R. L. Dalrymple, and A. E. Zanne. 2019. “What We (Don't) Know About Global Plant Diversity.” *Ecography* 42: 1819–1831.

Cuthbert, R. N., T. W. Bodey, E. Briski, et al. 2025. “Harnessing Traits to Predict Economic Impacts From Biological Invasions.” *Trends in Ecology & Evolution* 40: P639–P650.

D'Antonio, C., and S. L. Flory. 2017. “Long-Term Dynamics and Impacts of Plant Invasions.” *Journal of Ecology* 105: 1459–1461.

Davis, C. C. 2022. “The Herbarium of the Future.” *Trends in Ecology & Evolution* 38: S0169534722002956.

Dechant, B., J. Kattge, R. Pavlick, et al. 2023. “Intercomparison of Global Foliar Trait Maps Reveals Fundamental Differences and Limitations of Upscaling Approaches.”

Denelle, P., P. Weigelt, and H. Kreft. 2023. “GIFT—An R Package to Access the Global Inventory of Floras and Traits.” *Methods in Ecology and Evolution* 14: 2738–2748.

Díaz, S., J. Kattge, J. H. C. Cornelissen, et al. 2016. “The Global Spectrum of Plant Form and Function.” *Nature* 529: 167–171.

Díaz, S., J. Settele, E. S. Brondízio, et al. 2019. “Pervasive Human-Driven Decline of Life on Earth Points to the Need for Transformative Change.” *Science* 366: eaax3100.

Dinerstein, E., D. Olson, A. Joshi, et al. 2017. "An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm." *Bioscience* 67: 534–545.

Domazetoski, V., H. Kreft, H. Bestova, et al. 2023. "Using Natural Language Processing to Extract Plant Functional Traits From Unstructured Text."

Dostál, P. 2024. "Temporal Development in the Impacts of Plant Invasions: Search for the Underlying Mechanisms." *Journal of Evolutionary Biology* 37: 588–604.

Drenovsky, R. E., B. J. Grewell, C. M. D'Antonio, et al. 2012. "A Functional Trait Perspective on Plant Invasion." *Annals of Botany* 110: 141–153.

Enquist, B. J., R. Condit, R. K. Peet, M. Schildhauer, and B. M. Thiers. 2016. *Cyberinfrastructure for an Integrated Botanical Information Network to Investigate the Ecological Impacts of Global Climate Change on Plant Biodiversity*. PeerJ Inc.

Falster, D., R. Gallagher, E. H. Wenk, et al. 2021. "AusTraits, a Curated Plant Trait Database for the Australian Flora." *Scientific Data* 8: 254.

Feilhauer, H., T. Schmid, U. Faude, S. Sánchez-Carrillo, and S. Cirujano. 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., B. J. Enquist, D. S. Park, et al. 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., E. Ugarte, I. Kühn, and S. Klotz. 2008. "Alien Plants in Chile: Inferring Invasion Periods From Herbarium Records." *Biological Invasions* 10: 649–657.

Gallagher, R. V., D. S. Falster, B. S. Maitner, et al. 2020. "Open Science Principles for Accelerating Trait-Based Science Across the Tree of Life." *Nature Ecology & Evolution* 4: 1–10.

Gallien, L., and M. Carboni. 2017. "The Community Ecology of Invasive Species: Where Are We and What's Next?" *Ecography* 40: 335–352.

Garnier, E., U. Stahl, M.-A. Laporte, et al. 2017. "Towards a Thesaurus of Plant Characteristics: An Ecological Contribution." *Journal of Ecology* 105: 298–309.

Getman-Pickering, Z. L., A. Campbell, N. Aflitto, A. Grele, J. K. Davis, and T. A. Ugine. 2020. "LeafByte: A Mobile Application That Measures Leaf Area and Herbivory Quickly and Accurately." *Methods in Ecology and Evolution* 11: 215–221.

Gioria, M., P. E. Hulme, D. M. Richardson, and P. Pyšek. 2023. "Why Are Invasive Plants Successful?" *Annual Review of Plant Biology* 74: 635–670.

Golivets, M., S. Knapp, F. Essl, et al. 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.

Gorné, L. D., J. Aguirre-Gutiérrez, F. C. Souza, et al. 2025. "Use and Misuse of Trait Imputation in Ecology: The Problem of Using Out-Of-Context Imputed Values." *Ecography* n/a: e07520.

Govaerts, R., ed. 2023. "WCVP: World Checklist of Vascular Plants, Version 12."

Grenié, M., E. Berti, J. Carvajal-Quintero, G. M. L. Dädlow, A. Sagouis, and M. Winter. 2022. "Harmonizing Taxon Names in Biodiversity Data: A Review of Tools, Databases and Best Practices." *Methods in Ecology and Evolution* 14: 12–25.

Hart, A. G., H. Bosley, C. Hooper, et al. 2023. "Assessing the Accuracy of Free Automated Plant Identification Applications." *People and Nature* 5: 929–937.

Ho, L. S. T., and C. Ane. 2014. "A Linear-Time Algorithm for Gaussian and Non-Gaussian Trait Evolution Models." *Systematic Biology* 63: 397–408.

Homolová, L., Z. Malenovský, J. G. P. W. Clevers, G. García-Santos, and M. E. Schaepman. 2013. "Review of Optical-Based Remote Sensing for Plant Trait Mapping." *Ecological Complexity* 15: 1–16.

Hortal, J., F. de Bello, J. A. F. Diniz-Filho, T. M. Lewinsohn, J. M. Lobo, and R. J. Ladle. 2015. "Seven Shortfalls That Beset Large-Scale Knowledge of Biodiversity." *Annual Review of Ecology, Evolution, and Systematics* 46: 523–549.

Huang, C., and G. P. Asner. 2009. "Applications of Remote Sensing to Alien Invasive Plant Studies." *Sensors* 9: 4869–4889.

Hughes, A. C., M. C. Orr, K. Ma, et al. 2021. "Sampling Biases Shape Our View of the Natural World." *Ecography* 44: 1259–1269.

IPBES. 2023. *Summary for Policymakers of the Thematic Assessment Report on Invasive Alien Species and their Control of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*, edited by H. E. Roy, A. Pauchard, P. Stoett, et al. IPBES secretariat. https://doi.org/10.5281/zenodo.7430692.

Islam, S., C. Weiland, and W. Addink. 2022. "From Data Pipelines to FAIR Data Infrastructures: A Vision for the New Horizons of Bio- and Geodiversity Data for Scientific Research." *Research Ideas & Outcomes* 8: e93816.

Johnson, T. F., N. J. B. Isaac, A. Paviolo, and M. González-Suárez. 2021. "Handling Missing Values in Trait Data." *Global Ecology and Biogeography* 30: 51–62.

Joswig, J. S., J. Kattge, G. Kraemer, et al. 2023. "Imputing Missing Data in Plant Traits: A Guide to Improve Gap-Filling." *Global Ecology and Biogeography* 32: 1395–1408.

Kattge, J., G. Bönisch, S. Díaz, et al. 2020. "TRY Plant Trait Database—Enhanced Coverage and Open Access." *Global Change Biology* 26: 119–188.

Keller, A., M. J. Ankenbrand, H. Bruelheide, et al. 2023. "Ten (Mostly) Simple Rules to Future-Proof Trait Data in Ecological and Evolutionary Sciences." *Methods in Ecology and Evolution* 14: 444–458.

Keller, R. P., and J. M. Drake. 2009. *Trait-Based Risk Assessment for Invasive Species. Bioeconomics of Invasive Species: Integrating Ecology, Economics, Policy, and Management*, edited by R. P. Keller, D. M. Lodge, M. A. Lewis, and J. F. Shogren. Oxford University Press.

Knapp, S., and I. Kühn. 2012. "Origin Matters: Widely Distributed Native and Non-Native Species Benefit From Different Functional Traits." *Ecology Letters* 15: 696–703.

Kumar Rai, P., and J. S. Singh. 2020. "Invasive Alien Plant Species: Their Impact on Environment, Ecosystem Services and Human Health." *Ecological Indicators* 111: 106020.

Kummu, M., M. Taka, and J. H. A. Guillaume. 2018. "Gridded Global Datasets for Gross Domestic Product and Human Development Index Over 1990–2015." *Scientific Data* 5: 180004.

Küster, E. C., I. Kühn, H. Bruelheide, and S. Klotz. 2008. "Trait Interactions Help Explain Plant Invasion Success in the German Flora." *Journal of Ecology* 96: 860–868.

Lam, O. H. Y., S. Tautenhahn, G. Walther, G. Boenisch, P. Baddam, and J. Kattge. 2022. "The 'Rtry' R Package for Preprocessing Plant Trait Data. EGU22-13251."

Lázaro-Lobo, A., Á. Alonso, R. D. Fernández, et al. 2023. *Impacts of Plant Invasions on Ecosystem Functionality: A Perspective for Ecosystem Health and Ecosystem Services. Plant Invasions and Global Climate Change*, edited by S. Tripathi, R. Bhadouria, P. Srivastava, R. Singh, and D. R. Batish, 31–56. Springer Nature.

Leishman, M. R., J. Cooke, and D. M. Richardson. 2014. "Evidence for Shifts to Faster Growth Strategies in the New Ranges of Invasive Alien Plants." *Journal of Ecology* 102: 1451–1461.

Li, D. 2023. "Rtrees: An R Package to Assemble Phylogenetic Trees From Megatrees." *Ecography* 2023, no. 7: e06643.

Lüdecke, D., M. S. Ben-Shachar, I. Patil, P. Waggoner, and D. Makowski. 2021. "Performance: An R Package for Assessment, Comparison and Testing of Statistical Models." *Journal of Open Source Software* 6: 3139.

Machado, B. B., J. P. Orue, M. S. Arruda, et al. 2016. "BioLeaf: A Professional Mobile Application to Measure Foliar Damage Caused by Insect Herbivory." *Computers and Electronics in Agriculture* 129: 44–55.

Maitner, B. 2024. "EnquistLab/RTNRS: CRAN Release 0.3.6."

Maitner, B., R. Gallagher, J.-C. Svenning, M. Tietje, E. H. Wenk, and W. L. Eiserhardt. 2023. "A Global Assessment of the Raunkiæran Shortfall in Plants: Geographic Biases in Our Knowledge of Plant Traits." *New Phytologist* 240: 1345–1354.

Maitner, B. S., B. Boyle, N. Casler, et al. 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., P. Pyšek, and K. Bímová. 2004. "History of the Invasion and Distribution of Reynoutria Taxa in The Czech Republic: A Hybrid Spreading Faster Than Its Parents." *Preslia* 76, no. 1: 15–64.

Mariano, E., T. F. Gomes, S. R. M. Lins, et al. 2021. "LT-Brazil: A Database of Leaf Traits Across Biomes and Vegetation Types in Brazil." *Global Ecology and Biogeography* 30: 2136–2146.

Meyer, C., P. Weigelt, and H. Kreft. 2016. "Multidimensional Biases, Gaps and Uncertainties in Global Plant Occurrence Information." *Ecology Letters* 19: 992–1006.

Milanović, M., J. D. Bakker, L. Biederman, et al. 2025. "Successful Alien Plant Species Exhibit Functional Dissimilarity From Natives Under Varied Climatic Conditions but Not Under Increased Nutrient Availability." *Journal of Vegetation Science* 36: e70032.

Newbold, T., L. Hudson, D. W. Purves, J. P. W. Scharlemann, G. Mace, and A. Purvis. 2012. "Call for Data: PREDICTS: Projecting Responses of Ecological Diversity in Changing Terrestrial Systems." *Frontiers of Biogeography* 4: 155–156.

Niphadkar, M., and H. Nagendra. 2016. "Remote Sensing of Invasive Plants: Incorporating Functional Traits Into the Picture." *International Journal of Remote Sensing* 37: 3074–3085.

Novoa, A., D. M. Richardson, P. Pyšek, et al. 2020. "Invasion Syndromes: A Systematic Approach for Predicting Biological Invasions and Facilitating Effective Management." *Biological Invasions* 22: 1801–1820.

Oliveira, U., A. P. Paglia, A. D. Brescovit, et al. 2016. "The Strong Influence of Collection Bias on Biodiversity Knowledge Shortfalls of Brazilian Terrestrial Biodiversity." *Diversity and Distributions* 22: 1232–1244.

Palma, E., P. A. Vesk, M. White, J. B. Baumgartner, and J. A. Catford. 2021. "Plant Functional Traits Reflect Different Dimensions of Species Invasiveness." *Ecology* n/a: e03317.

Parker, J. D., M. E. Torchin, R. A. Hufbauer, et al. 2013. "Do Invasive Species Perform Better in Their New Ranges?" *Ecology* 94: 985–994.

Penone, C., A. D. Davidson, K. T. Shoemaker, et al. 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., S. Bacher, I. Kühn, et al. 2020. "Macroecological Framework for Invasive Aliens (MAFIA): Disentangling Large-Scale Context Dependence in Biological Invasions." *NeoBiota* 62: 407–461.

Pyšek, P., V. Jarošík, P. E. Hulme, et al. 2012. "A Global Assessment of Invasive Plant Impacts on Resident Species, Communities and Ecosystems: The Interaction of Impact Measures, Invading Species' Traits and Environment." *Global Change Biology* 18: 1725–1737.

Pyšek, P., D. M. Richardson, J. Pergl, V. Jarošík, Z. Sixtová, and E. Weber. 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*. Austria.

Razanajatovo, M., N. Maurel, W. Dawson, et al. 2016. "Plants Capable of Selfing Are More Likely to Become Naturalized." *Nature Communications* 7: 13313.

Rudbeck, A. V., M. Sun, M. Tietje, et al. 2022. "The Darwinian Shortfall in Plants: Phylogenetic Knowledge Is Driven by Range Size." *Ecography* 2022: e06142.

Schrodt, F., J. Kattge, H. Shan, et al. 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., S. Bacher, T. M. Blackburn, et al. 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.

Smith, S. A., and J. W. Brown. 2018. "Constructing a Broadly Inclusive Seed Plant Phylogeny." *American Journal of Botany* 105: 302–314.

Torresani, M., C. Rossi, M. Perrone, et al. 2024. "Reviewing the Spectral Variation Hypothesis: Twenty Years in the Tumultuous Sea of Biodiversity Estimation by Remote Sensing." *Ecological Informatics* 82: 102702.

Tyler, E. H. M., P. J. Somerfield, E. V. Berghe, et al. 2012. "Extensive Gaps and Biases in Our Knowledge of a Well-Known Fauna: Implications for Integrating Biological Traits Into Macroecology." *Global Ecology and Biogeography* 21: 922–934.

van Kleunen, M., W. Dawson, and N. Maurel. 2015. "Characteristics of Successful Alien Plants." *Molecular Ecology* 24: 1954–1968.

van Kleunen, M., P. Pyšek, W. Dawson, et al. 2019. "The Global Naturalized Alien Flora (GloNAF) Database." *Ecology* 100: e02542.

Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics With S*. 4th ed. Springer.

Visscher, A. M., F. Vandelook, E. Fernández-Pascual, et al. 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)*. https://data.nasa.gov/dataset/last-of-the-wild-project-version-2-2005-lwp-2-global-human-influence-index-hii-dataset-igh.

Webb, T. J., and B. Vanhoorne. 2020. "Linking Dimensions of Data on Global Marine Animal Diversity." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 375: 20190445.

Weigelt, A., L. Mommer, K. Andraczek, et al. 2021. "An Integrated Framework of Plant Form and Function: The Belowground Perspective." *New Phytologist* 232: 42–59.

Weigelt, P., and P. Denelle. 2022. "GIFT: Access to the Global Inventory of Floras and Traits (GIFT)."

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

Weiss, D. J., A. Nelson, H. S. Gibson, et al. 2018. "A Global Map of Travel Time to Cities to Assess Inequalities in Accessibility in 2015." *Nature* 553: 333–336.

Wenk, E. H., H. Sauquet, R. V. Gallagher, et al. 2024. "The AusTraits Plant Dictionary." *Scientific Data* 11: 537.

Westerband, A. C., J. L. Funk, and K. E. Barton. 2021. "Intraspecific Trait Variation in Plants: A Renewed Focus on Its Role in Ecological Processes." *Annals of Botany* 127: 397–410.

Westoby, M. 1998. "A Leaf-Height-Seed (LHS) Plant Ecology Strategy Scheme." *Plant and Soil* 199: 213–227.

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3: 1–9.

Williamson, M., P. Pyšek, V. Jarošík, and K. Prach. 2005. "On the Rates and Patterns of Spread of Alien Plants in The Czech Republic, Britain, and Ireland." *Écoscience* 12: 424–433.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data Sources S1:** Additional references for original data sources. **Data Sources S2:** The index for all of the material available in the "DataS2" file, should be the following: **Supplementary Information S1:** Details on the taxonomic harmonization process (i.e., aligning the taxonomies of the different databases used). **Supplementary Information S2:** Details on the trait harmonization process (i.e., aligning the trait definitions across trait databases). **Supplementary Information S3:** Ecological justification on the used trait sets. **Supplementary Table S3-1:** Name, references and traits contained in the used trait sets. **Supplementary Figure S1:** Venn diagram of the number of shared species across the used databases (GloNAF, AusTraits, BIEN, GIFT, and TRY). **Supplementary Figure S2:** Euler diagrams showing the contribution in number of unique and shared species of trait databases for the twenty traits available for most species. **Supplementary Figure S3:** Partial residual plots of variables explaining the number of traits available per species. **Supplementary Figure S4:** Estimated coefficients plots for linear models explaining the number of available traits, with explaining variables available across at least 70%, 80% or 90% of the range of the species. **Supplementary Figure S5:** Comparison of estimated coefficients across phylogenetic vs. non-phylogenetic models. **Supplementary Figure S6:** Cumulative number of species per number of available traits. **Supplementary Figure S7:** Treemaps of number of available traits per species and per family. **Supplementary Figure S8:** Global map of the median and standard deviation of the number of traits available per species.