Open Science principles for accelerating trait-based science across the Tree of Life

Rachael V. Gallagher [1]*, Daniel S. Falster [2], Brian S. Maitner [3], Roberto Salguero-Gómez[4,5,6], Vigdis Vandvik[7,8], William D. Pearse [9], Florian D. Schneider [51], Jens Kattge [10,11], Jorrit H. Poelen [51], Joshua S. Madin [12], Markus J. Ankenbrand [13,14,15], Caterina Penone [16], Xiao Feng [3], Vanessa M. Adams [17], John Alroy [1], Samuel C. Andrew[18], Meghan A. Balk[19], Lucie M. Bland[20], Brad L. Boyle[3], Catherine H. Bravo-Avila [21,22], Ian Brennan [23], Alexandra J. R. Carthey [1], Renee Catullo[23], Brittany R. Cavazos[24], Dalia A. Conde [25,26,27], Steven L. Chown [28], Belen Fadrique[21], Heloise Gibb[29], Aud H. Halbritter[7,8], Jennifer Hammock[30], J. Aaron Hogan[31], Hamish Holewa [18], Michael Hope[18], Colleen M. Iversen[32], Malte Jochum [11,16,33], Michael Kearney[34], Alexander Keller [13,14], Paula Mabee[35], Peter Manning [36], Luke McCormack[37], Sean T. Michaletz [38], Daniel S. Park [39], Timothy M. Perez[21,22], Silvia Pineda-Munoz[40], Courtenay A. Ray[41], Maurizio Rossetto[42,43], Hervé Sauquet [2,42,44], Benjamin Sparrow[45], Marko J. Spasojevic[46], Richard J. Telford [7,8], Joseph A. Tobias[47], Cyrille Violle[48], Ramona Walls[49], Katherine C. B. Weiss [41], Mark Westoby [1], Ian J. Wright [1] and Brian J. Enquist[3,50]

Synthesizing trait observations and knowledge across the Tree of Life remains a grand challenge for biodiversity science. Species traits are widely used in ecological and evolutionary science, and new data and methods have proliferated rapidly. Yet accessing and integrating disparate data sources remains a considerable challenge, slowing progress toward a global synthesis to integrate trait data across organisms. Trait science needs a vision for achieving global integration across all organisms. Here, we outline how the adoption of key Open Science principles—open data, open source and open methods—is transforming trait science, increasing transparency, democratizing access and accelerating global synthesis. To enhance widespread adoption of these principles, we introduce the Open Traits Network (OTN), a global, decentralized community welcoming all researchers and institutions pursuing the collaborative goal of standardizing and integrating trait data across organisms. We demonstrate how adherence to Open Science principles is key to the OTN community and outline five activities that can accelerate the synthesis of trait data across the Tree of Life, thereby facilitating rapid advances to address scientific inquiries and environmental issues. Lessons learned along the path to a global synthesis of trait data will provide a framework for addressing similarly complex data science and informatics challenges.

Traits, broadly speaking, are measurable attributes or characteristics of organisms. Traits related to function (for example, leaf size, body mass, tooth size or growth form) are often used to understand how organisms interact with their environment and other species via key vital rates such as survival, development and reproduction[1–5].

Trait-based approaches have long been used in systematics and macroevolution to delineate taxa and reconstruct ancestral morphology and function[6–8] and to link candidate genes to phentoypes[9–11]. The broad appeal of the trait concept is its ability to facilitate quantitative comparisons of biological form and function. Traits also allow us to mechanistically link organismal responses to abiotic and biotic factors with measurements that are, in principle, relatively easy to capture across large numbers of individuals. For example, appropriately chosen and defined traits can help identify lineages that share similar life-history strategies for a given environmental regime[12,13]. Documenting and understanding the diversity and composition of traits in ecosystems directly contributes to our understanding of organismal and ecosystem processes, functionality, productivity and resilience in the face of environmental change[14–19].

In light of the multiple applications of trait data to address challenges of global significance (Box 1), a central question remains: How can we most effectively advance the synthesis of trait data within and across disciplines? In recent decades, the collection, compilation and availability of trait data for a variety of organisms has accelerated rapidly. Substantial trait databases now exist for plants[20–23], reptiles[24,25], invertebrates[23,26–29], fish[30,31], corals[32], birds[23,33,34], amphibians[35], mammals[23,36–38] and fungi[23,39], and parallel efforts are no doubt underway for other taxa. Though considerable effort has been made to quantify traits for some groups (for example, Fig. 1), substantial work remains. To develop and test theory in biodiversity science, much greater effort is needed to fill in trait data across the Tree of Life by combining and integrating data and trait collection efforts.

A full list of affiliations appears at the end of the paper.

Box 1 | Research programs dependent on comprehensive trait data across the Tree of Life

Access to open data on the traits of all organisms will allow the pursuit of long-standing questions in ecological science, including:

Defining major axes of strategy variation across the Tree of Life. Measurements of organismal traits can be used to identify the position of species along trade-off spectrums that shape fitness. For instance, traits (for example, leaf nitrogen content, leaf mass per area, seed mass and maximum height) have been used to capture trade-offs between plant species at global scales[91]. With access to open data on the traits of a wide breadth of organisms, we will be able to identify major axes of functional specialization across the Tree of Life. Traits such as adult body mass, offspring mass at independence, mass-specific metabolic rate, and body temperature capture core differences in the ecological strategies within and across groups[92–94]. Though there is tremendous potential in the comparison of traits across the Tree of Life via the OTN, consideration of scale and biology are needed to identify where it is appropriate to make such comparisons[95].

Conservation of functional diversity in protected areas. Reserve-selection procedures[96] seek to identify representative networks of potential protected areas at the lowest financial cost. Species are often the biological unit being targeted regardless of the functional diversity they collectively represent. However, there is an increased focus on ecosystem services and protecting ecosystem function by better representing phylogenetic or functional diversity[97]. With trait data from the Tree of Life, systematic conservation planning[98] could: (1) optimize reserve designs to maximize the conservation of ecosystem function predicted from species traits; or (2) assess the adequacy of current protected areas in representing ecosystem function.

Reserve design typically targets species or populations, not ecosystem functions, and is based on three fundamental principles: comprehensiveness, adequacy and representativeness[98]. Using theoretical predictions about the relationship between traits and ecosystem functions we could predict which traits, or range of trait values[99], should be preserved when designing protected area networks. In this context, comprehensiveness would allow for the inclusion of the full range of ecosystem functions (as captured by traits of species present) recognized at an appropriate scale within and across relevant ecosystem units (for example, bioregions and biomes); adequacy would seek to maintain the viability and integrity of ecosystem function and model how functional redundancy (that is, different species performing the same functional role[100]) may scale across landscapes to identify the minimum viable reserve size for maintaining ecosystem functions; and representativeness would seek to capture the diversity of functions and the gradients across which they occur, including the level of intraspecific variation and plasticity inherent to the trait being examined.

Strengthening predictions of the effects of global change on biodiversity. Traits already provide valuable information for models predicting global change impacts on the biosphere. Access to large-scale data on plant traits allows the distribution of traits to be captured in Earth system models rather than only modelling competition among broad plant functional types[101–103]. Species-based predictions from correlative niche models may also benefit from the integration of data on species biology to capture mechanistic links between function and the environment[104–107].

The next generation of Earth system models may also benefit from integration of open trait data for (at least) two types of organisms: (1) terrestrial vertebrates; and (2) microbes. Terrestrial vertebrates provide key ecological functions (for example, dispersal and disturbance[108]) and their loss may result in significant changes to ecosystem function[109]. However, capturing the influence of terrestrial megafauna on forest structure, function and biogeochemical cycles will be improved by access to data on traits (for example, size and diet) to parameterize process-based models. Access to knowledge about the traits of soil-borne microbes, whose activities increase atmospheric warming via the decomposition of soil carbon, may help predict planetary responses to climate change[110–112]. Molecular-level traits in microbes (for example, enzyme activity) allow the estimation of decomposition rates but are difficult to measure, amplifying the need to share these data openly.
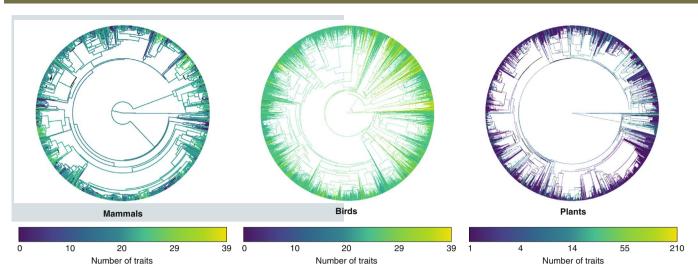

Current barriers to global trait-based science

Despite the recognized importance of traits, several common research practices limit our capacity for meaningful synthesis across the Tree of Life. These practices include failure to publish usable datasets alongside new findings[40], missing or inadequate metadata[41], minimal descriptions of methods used to collate, clean and analyse trait datasets in published works[42], and inadequate coordination between researchers and institutions with common goals, such as filling strategic spatial or taxonomic gaps in trait knowledge[43,44]. Our limited ability to access and redistribute trait data contributes to the widespread reproducibility crisis within science[45]. Any study relying on data that cannot easily be re-used introduces barriers to verifying the claims made by those studies and thereby questions the reproducibility of the science[46], which is becoming of prime importance to many scientific journals. Such limitations have been common within trait-based science.

Access to data is not the only impediment to a global synthesis of trait knowledge. Barriers to synthesis exist because researchers and institutions are apprehensive that the time and resources they spend to create new observations or share legacy data (for example, observations from field guides, specimens, or publications without data supplements) will not be recognized. Identifying who should receive credit for contributing trait observations (whether via co-authorship or other formal recognition) is a complex issue, particularly where data involve a chain of expertise (for example, when trait data are extracted from taxonomic treatments involving specimen collectors, digitizers, taxonomists and curators). Funding bodies are often reluctant to support data management, limiting recognition of the sizeable effort expended on creating bespoke solutions to curating and harmonizing trait data from different sources[46].

Opportunities exist for expanding the spatial and taxonomic coverage of trait observations, particularly by strengthening interdisciplinary connections across single organismic groups. Despite certain plant traits (for example, growth form, height and leaf size) being carefully catalogued in taxonomic species descriptions[47], these data have only recently been exchanged with large-scale databases such as TRY[21] or BIEN (http://bien.nceas.ucsb.edu/bien/). Although several informatics challenges in biodiversity science have now been overcome (for example, synthesizing global species occurrence information (https://www.gbif.org/) and sharing genetic data on individuals (https://www.ncbi.nlm.nih.gov/genbank/)), trait science lacks a vision for achieving global integration across all organisms. We argue that this is not simply a failure of the traits community to learn from existing successful networks. Instead, cataloguing traits is a more complex task that is highly context-dependent and therefore needs a more refined network model than that offered by a centralized repository.

Fig. 1 | Mammal, bird and plant phylogenies coloured according to the number of traits for which we have data for each species and lineage. The plant phylogeny is sparsely populated for traits but contains more taxa ($n = 10{,}596$) than the mammal and bird phylogenies ($n = 5{,}747$ and $9{,}993$, respectively). Trait data were downloaded from refs. [25,34,87]. We counted the number of traits present across these datasets for each species and mapped those onto phylogenies using posteriors[37,88] and a random subset of plant species within a single phylogeny[89]. Terminal branches (representing species) and ancestral lineages (using ancestral state reconstruction[90]) were coloured according to the number of reconstructed traits. Note that this is an exploratory analysis conducted purely to show variation in the availability of trait data across taxonomic groups.

We propose that widespread adoption of key Open Science principles (Box 2) could be transformative for trait science in achieving a global synthesis. These principles would lay a strong foundation for transparency, reproducibility and recognition, and encourage a culture of data sharing and collaboration beyond established networks. Openness reinforces the scientific process by allowing increased scrutiny of methods and results, resulting in the deeper exploration of findings and their significance[42,48–51]. The scope of trait science would increase if researchers and institutions: (1) made datasets available in machine-accessible formats under clear licensing arrangements; (2) created and adopted standardized protocols, handbooks or metadata formats for data collection, documentation and management (see refs. [48,49]); and (3) created human-centred networks to reduce the complexity of integrating existing data from disparate sources (for example, specimens, published literature, citizen-science initiatives[50,51] and large-scale digitization efforts). These different sources exhibit systematic differences in error rates, validation, context, reproducibility and objectivity relative to field-collected trait observations. Without a model of recognition that embraces transparency and fairness, much trait data will remain hidden from science.

Introducing the Open Traits Network

The Open Traits Network (OTN) is a collaborative initiative for accelerating trait data synthesis. Specifically, it is a global, decentralized community welcoming all researchers and institutions pursuing the collaborative goal of standardizing and integrating trait data across all organisms. We promote five main objectives built upon Open Science ideals that could transform trait science:

(1) Openly sharing data, methods, protocols, codes and workflows.
(2) Citing original data collectors and providing scholarly credit.
(3) Providing appropriate metadata together with trait observations.
(4) Collecting trait data following reproducible, standardized methods and protocols (when available), or committing to their development.
(5) Providing training resources in trait collection and database construction using Open Science principles.

We envision a future for trait research where protocols for data exchange and re-use are transparent, research findings are reproducible, and all trait data (either newly collected or from legacy sources) are openly available to the research community and broader public. While several network models exist in trait research (Fig. 2), the OTN adopts a decentralized but connected structure with an emphasis on bringing people together through data and expertise.

Often, groups building smaller-scale databases do so in isolation, using their own tools and workflows tailored to their research question; they are decentralized and disconnected (Fig. 2a). Decentralization has certain advantages, including retaining the power to determine which traits are most useful in a study system and how they should be compiled. There is little formal support or interaction across this style of network, so researchers often collect redundant data and develop similar tools for data collection, cleaning and integration, which can lead to duplication of effort. There are many small, isolated and heterogeneous data sources of this sort, increasing the disconnect between pools of trait data[52].

For some organisms, centralized hubs exist to aggregate and standardize trait data across disparate sources (see refs. [21,32,53–57]) (Fig. 2b). These trait repositories have become the main access point for trait data on well-studied taxa such as plants and corals, but they remain mostly isolated, limiting the sharing of expertise and information across taxa. As these repositories continue to grow, difficulties with data integration and synthesis will also increase due to the momentum of entrenched workflows and exchange protocols that may not be interoperable.

Some successful large-scale initiatives have followed the centralized and connected network model (for example, the Global Biodiversity Information Facility (GBIF; https://www.gbif.org/) and GenBank (https://www.ncbi.nlm.nih.gov/genbank/)). These platforms mandate strict data exchange protocols to facilitate synthesis using standardized vocabularies (for example, the Darwin Core[58] and Humboldt Core[59]). These protocols have been central to the explosive growth of biodiversity data as they facilitate the exchange of information using common data formats[58–60]. Ontologies that provide unified terms and concepts necessary to represent traits have been developed (for example, Uberon, the multispecies anatomy ontology for animals[61], and TOP, the Thesaurus of Plant characteristics[62]).
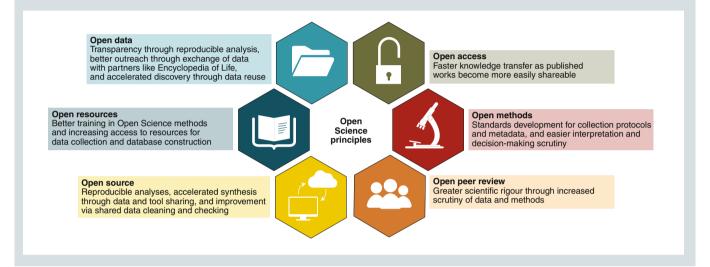
Box 2 | Using Open Science principles in trait research

Open Science principles outline a movement towards making all aspects of the scientific process transparent and accessible to a wide audience[51]. Open Science principles are rapidly being adopted across the sciences.

The Box 2 figure shows the six core principles of Open Science and their potential benefits to trait science. Three Open Science principles are particularly relevant to the Open Traits Network and trait-based research more broadly: Open Data, Open Source and Open Methods.

In this context, knowledge is considered open if anyone can freely access, use, modify and share it, subject at most to measures that preserve provenance and openness (http://opendefinition.org/). Several pronouncements about Open

Science principles have already been made, including the Berlin Declaration (https://openaccess.mpg.de/Berlin-Declaration), the Bouchout Declaration (http://www.bouchoutdeclaration.org/declaration/), and the Denton Declaration (https://openaccess.unt.edu/denton-declaration) on open access to science data. Other initiatives champion open practices such as the Bari Manifesto on interoperability[113] and the FORCE 11 network, which developed the 'Joint Declaration of Data Citation Principles' (https://www.force11.org/datacitationprinciples) and the 'FAIR' principles (https://www.force11.org/group/fairgroup/fairprinciples). The FAIR principles address several of the challenges facing trait-based research, namely making data findable, accessible, interoperable and reusable.

These provide integration with other data types (for example, genetic and environmental) and their corresponding ontologies (for example, Gene Ontology[63] and Environmental Ontology[64]).

Despite these successes, we argue that a centralized and connected network structure will not facilitate trait data synthesis. Trait observations are highly nuanced and hierarchical. Describing multiple aspects of a phenotype for any organism with traits is not amenable to a simplified set of exchange fields that apply across the Tree of Life. While the centralized and connected model (Fig. 2b) does have benefits, it lacks the necessary flexibility to connect trait data where ontologies and exchange formats do not exist. The likely result is that established trait networks will remain isolated and disconnected.
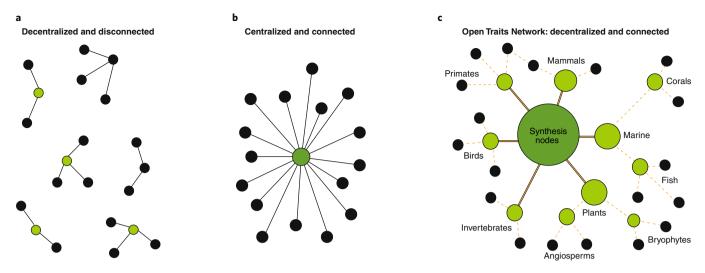
The decentralized but connected model (orange connections in Fig. 2c) adopted by the OTN maintains the key advantages of a decentralized network (for example, taxon- or discipline-specific decision making) while enhancing the level of connectivity among groups, allowing for easier sharing of expertise, tools and data. These network characteristics also buffer against node loss (for example, due to lack of funding). Decentralized and connected networks are characterized by socially mediated improvements in learning[65] as they capitalize on the aggregated judgement of many experts rather than singular opinions[66]. The OTN model capitalizes on existing connections within disciplines and links domains across the Tree of Life to disseminate knowledge about traits. By recognizing the importance of specialist taxon groups (light-green nodes in Fig. 2c) and accommodating their needs into the development of cross-domain tools for synthesis (dark-green nodes in Fig. 2c), the OTN model will be particularly beneficial for low-profile taxa

that may not be accommodated by a centralized effort to synthesize data. The OTN's open, decentralized network structure will allow researchers to retain agency and independence while also creating a collaborative effort to minimize the duplication of effort.

How (and why) to participate in the OTN

The OTN seeks to broaden its membership by lowering barriers to inclusion and advocating for approaches to trait science that benefit data custodians. New members can join the OTN via our website (www.opentraits.org) through two mechanisms: (1) adding a member profile (for example, name, location, expertise and collaboration statement); and/or (2) registering their open-source (or embargoed) trait datasets in the OTN Trait Dataset Registry (see Activity 1). The registry contains metadata for trait datasets and links users to the open dataset. New entries to the registry will be reviewed by OTN members before being added. This step will facilitate interaction between new and established OTN members and encourage deeper collaboration. Once registered, members will receive regular updates about the OTN, including newly registered trait datasets, notifications about upcoming chances for face-to-face meetings, and funding opportunities. Members will also benefit from the OTN through the sharing of resources, funding calls and workshops where appropriate.

OTN membership spans scientists (and institutions) with high-level expertise in trait data science and synthesis activities through to those with strong motivations to work with traits but little expertise. The OTN has already conducted an international workshop facilitated by an open call for participants, with more workshops planned. Following this initial communication process, we are

Fig. 2 | Architectures of three alternative networks in which research groups (nodes) interact in collecting and organizing trait data. Black nodes are individuals, groups or institutions conducting projects. Light-green nodes are those harmonizing data and developing protocols, where node size is proportional to available resources. Dark-green nodes are synthesis nodes that collect standardized trait data and knowledge. **a**, Groups are disconnected and decentralized, risking duplication of effort (often the status quo). **b**, Groups are linked to a centralized repository, potentially limiting innovation. **c**, The Open Traits Network, represented by orange lines. Nodes are linked within biological domains (for example, plants or marine) and include expertise from diverse disciplines (for example, systematics, palaeobiology, ecology and biomechanics) allowing for more efficient and specialized decisions about trait collection. Data synthesis across domains or disciplines is facilitated by joining nodes based on common workflows, theoretical frameworks and data-sharing protocols that adhere to the principles of the Open Traits Network.

currently sharing ideas and act upon them within subgroups. Being a decentralized network, the OTN does not need to rely on funding and dedicated personnel to complete tasks, though larger goals will benefit from financial support. Instead, we will communicate the joint aims and gaps between network nodes (Fig. 2) and arrange workshops and activities where necessary.

We recognize that altruism is unlikely to offer enough motivation to ensure widespread participation in the OTN. The sharing of trait datasets is not merely a technical problem to be solved; it relies on custodians having the skills, incentives and motivation to contribute. The key incentives for individuals to join the OTN include increasing the findability of their data and expertise and having access to a ready-made network of trait scientists and institutions engaging in relevant initiatives. Data are a powerful asset for researchers, and release under open-license schemes accompanied by well-defined metadata offers great potential for new collaborations and increased visibility. A persistent concern is that scientists will lose control of their hard-earned data under open licensing, though this underestimates the potential for new collaborations and may unnecessarily increase distrust within the scientific community[67]. Access to scientific networks can provide valuable exposure and connection[49], particularly for early-career researchers and those in developing nations, although it is important to understand the risks involved. By emphasizing the importance of community engagement and support, the OTN seeks to make trait-data sharing and synthesis an opportunity for all involved rather than simply a technical challenge to be solved.

Milestones toward an open approach to trait-based science
We highlight five OTN activities (several of which are already operational) that demonstrate the power of a decentralized and connected network to increase knowledge transfer in trait science. Trait scientists have made significant achievements in key areas, such as the synthesis of large numbers of observations within taxonomic groups[20–39] and the development of theory and frameworks to use these data when testing ideas and large-scale empirical studies[1–19]. However, basic foundations are still lacking to quantify how and why traits vary across organisms.

Activity 1: Maintaining a global registry of trait-based initiatives. Several data gaps impede synthetic analyses across taxa, geographical locations and ontogeny. The heterogeneous ways in which trait data have been collected to date have resulted in a patchy and unrepresentative data landscape across trait types, taxa, regions and times of the year[68,69]. The OTN bridges these gaps by maintaining a Trait Dataset Registry that can be accessed at http://opentraits.org/datasets.html.

The OTN Registry contains information on existing open (or embargoed) datasets so that gaps can be identified and ultimately filled through collective effort. Core information for the registry includes Digital Object Identifier (DOI), taxonomic coverage, curator and format. The OTN Registry also provides the opportunity for contributors to identify if and where code to process and manipulate raw data is located (see Activity 2). As it develops, the OTN Registry will relate trait concepts to ontologies provided through the Open Biomedical Ontologies Foundry (http://www.obofoundry.org). The OTN Registry maps to several Open Science principles (for example, Open Source, Open Data and Open Access; Box 2) and is designed to support data retrieval and integration.

The OTN does not place restrictions on what members may consider traits of importance to a taxonomic group. Most traits can be measured from individuals and fit into existing definitions, though this may not be appropriate for organisms where individual or taxonomic boundaries are unclear (for example, microbes[70] and fungi[71]). It can be argued that traits encompass emergent properties of populations (for example, abundance and geographic range size) or represent interactions among species (for example, diet type). Within the OTN, we believe that more important than imposing strict definitions around traits is engaging the community in discussion about the utility of available data for answering novel ecological and evolutionary questions.

Activity 2: Sharing reproducible workflows and tools for aggregating trait data. The OTN leverages collaborative software development via platforms like GitHub (https://github.com/) to create modular, open-source software to access, harmonize and re-use

data with seamless piping of data from one software tool to the next. OTN contributors have already developed several open-source tools such as the traitdataform package, which assists R users to format their data and harmonize units (http://ecologicaltraitdata.github. io/traitdataform). The code for the Coral Traits database[32] (https:// github.com/jmadin/traits) could be modified to guide the creation of databases on other organisms. The FENNEC project provides a tool for accessing and viewing community trait data as a self-hosted website service[72] (https://github.com/molbiodiv/fennec). The OTN can act as a connector between developers and the broader community seeking to synthesize trait data, facilitating the training of scientists in all aspects of reproducible data management.

Activity 3: Advocating for a free flow of data and appropriate credit. One goal of the OTN is to increase the use of open datasets and to ensure due credit is given to researchers who collect or synthesize primary data. Without effective reward or motivation for collecting new trait observations or sharing legacy data, a trait synthesis across the Tree of Life will remain unattainable. Currently, motivation for collecting and sharing new primary data is not strong and direct funding for trait data management is scarce.

The OTN can strengthen the attribution of credit to data providers and promote new data collection via two paths. Firstly, the OTN will encourage citation back to primary source via a permissive license model that secures authorship attribution (for example, Creative Commons Attribution 4.0 Int; CC BY 4.0) and the use of DOIs and Open Researcher and Contributor ID (ORCID) identifiers. Open-access datasets with a DOI can be tracked to understand patterns of re-use and to assess the impact of the author's decision to share.

There is an important distinction between sharing data within a network and making data publicly available under an open license. Clear license arrangements increase visibility and promote fair attribution and citation (for example, using Creative Commons licenses such as CC-BY or CC0). CC-BY requires attribution (that is, citation) to the original creator whereas CC0 does not legally require users of the data to cite the source, though this does not affect ethical norms for attribution in research communities (https://creative-commons.org/share-your-work/public-domain/cc0/). Identifying who should be credited for prior work on legacy data is complicated by the involvement of many individuals. This issue could be solved, in part, by inviting organizations to be named as contributors or co-authors on outputs using their data or (looking forward) implementing new ways of documenting who should be credited for making specimens or datasets usable in trait science.

Incentives to collect new trait data can be linked to the Open Science practice of pre-registration. In pre-registration, authors archive a public proposal for research activities (for example, via the Centre for Open Science; https://cos.io/prereg/) which, if approved, may receive in-principle acceptance from participating journals. As of March 2019, 168 journals are willing to give in-principle acceptance following pre-review of the study design prior to conducting field or experimental work. Ten of these participating journals regularly feature papers on trait-based science (for example, *BMC Ecology* and *Ecology and Evolution*). We envision a situation where the OTN Trait Registry (Activity 1) could be used to identify spatial or taxonomic gaps in trait data that could be coupled to pre-registered hypotheses. Together, pre-registration and in-principle acceptance of findings could incentivize the collection of new data, circumventing the growing reliance on available data with known gaps.

Activity 4: Creating a trait core to facilitate synthesis and standardization. Trait science requires its own 'core' terminology or data standard that is flexible enough to capture the complexity of trait data. Building on efforts to standardize occurrence data (that is, Darwin Core[58]) and biological inventories (that is, Humboldt

Core[43,59]), the OTN envisions a trait core offering a set of cross-domain metadata standards and controlled vocabularies that are (ideally) connected to trait ontologies via unambiguous identifiers. This standard terminology would be implemented across trait-data publications, unifying data in decentralized repositories as well as centralized data portals.

A trait core would allow trait data to be: (1) interpreted accurately within the context of their collection (that is, including information on associated data on factors such as environmental conditions at collection sites, taxa covered, data custodians or collection methods); and (2) known by compatible terms so that observations of similar phenomena can be grouped and compared (that is, what is meant by 'generation time' or 'establishment' across taxonomic groups[73,74]). Existing initiatives may provide logical cornerstones for referencing terms and concepts, including Ecological Metadata Language[41]. Several initiatives implement the Ecological Metadata Language (for example, The Knowledge Network for Biocomplexity[75], Darwin Core[58] and Humboldt Core[59]) and the use of referencing terms from anatomy or phenotype ontologies (for example, the Plant Ontology[66] and the Vertebrate Trait Ontology[67]) to relate traits to publicly defined terms, allowing annotated data to be processed computationally (http://www.obofoundry.org).

Progress towards a trait core is already being made through the development of a prototypal Ecological Trait Standard[76] (Box 3). However, the development and adoption of a trait core requires consultation and coordination within the broader scientific community, a goal which the OTN is ideally placed to advance. The OTN can mobilize expertise for cross-domain workshops and advocate for funding, which allows not only meetings of experts but also the creation of cyber-infrastructure for synthesis nodes (dark-green nodes in Fig. 2c). Links to emerging initiatives for biodiversity data standardization (for example, Species Index of Knowledge[57]) will also be vital for success, as will ratification of the core through the Biodiversity Information Standards (TDWG, www.tdwg.org).

Activity 5: Facilitating consistent approaches to measuring traits within major groups. The OTN will share new developments towards protocols and handbooks for major clades that standardize approaches to capture trait observations. Protocols are necessary because downstream activities such as developing metadata standards (Activity 4) will be impossible if trait measurement protocols do not exist. Some research communities have adopted standardized terms[56,62] and data collection protocols (for example, plants[20,77–80], invertebrates[29,81–83], mammals[36] and aquatic life[30,32,84]), though these may not always fit the requirements of some studies (for example, where trait variability rather than the average trait of species is targeted[85]). Protocols and handbooks may not emerge rapidly and should have the flexibility to be open to innovation through a commitment to version control and updates as techniques evolve. Two versions of the plant trait measurement handbook have been published[77,86] and several online resources exist that can be updated regularly (see http://prometheuswiki.org/tiki-custom_home.php).

Standardizing approaches to trait measurement across research communities will reduce ambiguity when aggregating data and improve the quality of resulting datasets. Integrating trait standardization and databasing into taxonomic workflows constitutes a challenge and an opportunity[7] that holds the promise of bridging the long disconnect between structural and functional traits. The presence of a range of biodiversity collections personnel in the OTN and an open invitation for more to join is expected to catalyse the adoption of trait-based thinking into taxonomic practices.

Concluding remarks

This is the opportune time to push towards a new approach to sharing and synthesizing trait data across all organisms. Trait science has great potential to increase its taxonomic, phylogenetic and spatial

Box 3 | A pipeline for harmonizing trait data from disparate sources

Preparation and re-use of open trait data requires the following action steps[76]:

1. **Data collection and data handling within the project context**. Applying project-specific methodology and storing data in tables suitable for data analysis, applying project-specific terms for taxa, traits and column labels. Adopting standard methodologies and terminologies from the start will simplify steps 2 and 3 and facilitate data publication. OTN Activity 5 aims to build consensus for common trait definitions and measurement methodologies for major organismal groups.

2. **Harmonization of taxa and traits into standard terms using ontologies**. Prior to publication, all taxa should be harmonized into accepted names linked to ontologies using uniform resource identifiers (URIs). Those can be provided in metadata or secondary data tables. Ontologies for traits are scarce, but if available should also be referred to via URIs to deliver unambiguous trait definitions. OTN Activity 2 will foster the development of ontologies linking trait data to publicly available resources.

3. **Standardization of table descriptors and metadata using a standard vocabulary**. Data should be published in tables using standard terms for column names, such as those provided by the Ecological Trait-data Standard vocabulary (ETS; DOI: 10.5281/zenodo.2605377). This ETS implements a minimal terminology

that can be adapted to include traits from a variety of organisms and uses of uniform resource identifiers for taxa, traits, methods and units, thereby following the standards for a semantic web of scientific data. Metadata should point to the applied standard to facilitate interpretation by humans and machines. OTN Activity 4 will engage in the cross-domain, community-based development for standard vocabularies for trait data.

4. **Publication of data and upload to a public repository**. Open Access file hosting services offer permanent hosting and findability of data by assigning a DOI to data publications and stating authorship and conditions for re-use under Creative Commons licenses. OTN Activity 3 will support recognition for data publications and thereby mitigate the investment into data standardization for smaller research projects.

5. **Synthesis of trait data and re-use in downstream products**. Open-access data publications with high-quality metadata are a valuable complement for meta-analysis or the functional analysis of abundance data. By keeping original author terms and values, the quality of the derived datasets can be assured and controlled for a better integration of multiple datasets. Availability of such high-quality data will facilitate reproducibility and enable computer-aided analysis of large databases. OTN Activity 1 supports the findability of trait data by creating a public registry, and OTN Activity 2 will develop tools to aid the compilation of databases.

scopes by leveraging data-science tools, embracing Open Science principles, and creating stronger connections between researchers, institutions, publishers and funding bodies. We hope that trait enthusiasts, regardless of field and research stage, will engage with the OTN via our website (www.opentraits.org) and help build new connections between disciplines, institutions and taxonomic domains. By adding metadata profiles for datasets to the OTN Trait Dataset Registry, trait collection efforts become more findable, as do the researchers who have compiled them. We envision that by connecting people with common goals, we can work collectively towards a synthesis of global trait data to preserve the nuances of taxon-specific expertise while also facilitating collaboration across domains. We urge scientists and institutions keen to commit to Open Science principles to make use of existing resources, including those offered by the Centre for Open Science (https://cos.io/), the Open Science Training Handbook (https://open-science-training-handbook.gitbook.io/book/), the Open Science Training Initiative (http://www.opensciencetraining.com/index.php) and FOSTER (https://www.fosteropenscience.eu/toolkit).

To support and expand the activities of the OTN, we will grow membership and develop communities around synthesis nodes to undertake key activities and secure funding support, in particular for the development of a trait core. Funding for international workshops, technical support and implementation meetings could drive a new era of trait-based synthesis that mirrors the achievement of similar initiatives such as GBIF, which now houses >1 billion occurrence records.

By supporting a reciprocal exchange of expertise and outputs using Open Science principles between researchers and institutions, we can mobilize data for a cross-taxa, worldwide, trait-based data resource to examine, understand and predict nature's responses to global change. As a better-connected OTN emerges, data streams and coordination will improve, allowing us to deliver information to support globally important research agendas (Box 1) as well as specific data and knowledge to the public through integration with third-party portals. Lessons learned along the path to a global synthesis of trait data across all organisms will provide a framework

for addressing similarly complex, context-dependent challenges in biodiversity informatics and beyond.

References

1. Adler, P. B. et al. Functional traits explain variation in plant life history strategies. *Proc. Natl Acad. Sci. USA* **111**, 740–745 (2014).
2. Chapin, F. S. III, Autumn, K. & Pugnaire, F. Evolution of suites of traits in response to environmental stress. *Am. Nat.* **142**, S78–S92 (1993).
3. Chown, S. L. & Gaston, K. J. Macrophysiology–progress and prospects. *Funct. Ecol.* **30**, 330–344 (2016).
4. Kooijman, S. A. L. M. *Dynamic Energy and Mass Budgets in Biological Systems* (Cambridge Univ. Press, 2000).
5. Diaz, S., Cabido, M. & Casanoves, F. Plant functional traits and environmental filters at a regional scale. *J. Veg. Sci.* **9**, 113–122 (1998).
6. Harmon, L. J. et al. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* **64**, 2385–2396 (2010).
7. Sauquet, H. & Magallón, S. Key questions and challenges in angiosperm macroevolution. *New Phytol.* **219**, 1170–1187 (2018).
8. Sneath, P. H. & Sokal, R. R. *Numerical Taxonomy: The Principles and Practice of Numerical Classification* (W. H. Freeman & Co, 1973).
9. Edmunds, R. C. et al. Phenoscape: identifying candidate genes for evolutionary phenotypes. *Mol. Biol. Evol.* **33**, 13–24 (2015).
10. Mungall, C. J. et al. The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species. *Nucleic Acids Res.* **45**, D712–D722 (2016).
11. Gkoutos, G. V., Schofield, P. N. & Hoehndorf, R. The anatomy of phenotype ontologies: principles, properties and applications. *Brief. Bioinform.* **19**, 1008–1021 (2017).
12. Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A. & Wright, I. J. Plant ecological strategies: some leading dimensions of variation between species. *Annu. Rev. Ecol. Syst.* **33**, 125–159 (2002).
13. Kiørboe, T., Visser, A. & Andersen, K. H. A trait-based approach to ocean ecology. *ICES J. Mar. Sci.* **75**, 1849–1863 (2018).
14. Kunstler, G. et al. Plant functional traits have globally consistent effects on competition. *Nature* **529**, 204–207 (2016).
15. Laughlin, D. C. Nitrification is linked to dominant leaf traits rather than functional diversity. *J. Ecol.* **99**, 1091–1099 (2011).
16. Finegan, B. et al. Does functional trait diversity predict above-ground biomass and productivity of tropical forests? Testing three alternative hypotheses. *J. Ecol.* **103**, 191–201 (2015).

17. Laigle, I. et al. Species traits as drivers of food web structure. *Oikos* **127**, 316–326 (2018).
18. Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. Toward a metabolic theory of ecology. *Ecology* **85**, 1771–1789 (2004).
19. West, G. B., Brown, J. H. & Enquist, B. J. A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122–126 (1997).
20. Iversen, C. M. et al. A global Fine-Root Ecology Database to address below-ground challenges in plant ecology. *New Phytol.* **215**, 15–26 (2017).
21. Kattge, J. et al. TRY–a global database of plant traits. *Glob. Change Biol.* **17**, 2905–2935 (2011).
22. Bernhardt-Römermann, M., Poschlod, P. & Hentschel, J. BryForTrait–A life-history trait database of forest bryophytes. *J. Veg. Sci.* **29**, 798–800 (2018).
23. Bennett, J. M. et al. GlobTherm, a global database on thermal tolerances for aquatic and terrestrial organisms. *Sci. Data* **5**, 180022 (2018).
24. Meiri, S. Traits of lizards of the world: variation around a successful evolutionary design. *Glob. Ecol. Biogeogr.* **27**, 1168–1172 (2018).
25. Myhrvold, N. P. et al. An amniote life-history database to perform comparative analyses with birds, mammals, and reptiles. *Ecology* **96**, 3109–3109 (2015).
26. Schäfer, R. B. et al. A trait database of stream invertebrates for the ecological risk assessment of single and combined effects of salinity and pesticides in South-East Australia. *Sci. Total Environ.* **409**, 2055–2063 (2011).
27. Bland, L. Global correlates of extinction risk in freshwater crayfish. *Animal Conserv.* **20**, 532–542 (2017).
28. Brun, P., Payne, M. R. & Kiørboe, T. A trait database for marine copepods. *Earth Syst. Sci. Data* **9**, 99–113 (2017).
29. Parr, C. L. et al. GlobalAnts: a new database on the geography of ant traits (Hymenoptera: Formicidae). *Insect Conserv. Divers.* **10**, 5–20 (2017).
30. Froese, R. & Pauly, D. *Progress Report on FishBase* (Fisheries Centre, University of British Columbia, 2010).
31. Frimpong, E. A. & Angermeier, P. L. Fish traits: a database of ecological and life-history traits of freshwater fishes of the United States. *Fisheries* **34**, 487–495 (2009).
32. Madin, J. S. et al. The Coral Trait Database, a curated database of trait information for coral species from the global oceans. *Sci. Data* **3**, 160017 (2016).
33. Garnett, S. T. et al. Biological, ecological, conservation and legal information for all species and subspecies of Australian bird. *Sci. Data* **2**, 150061 (2015).
34. Wilman, H. et al. EltonTraits 1.0: species-level foraging attributes of the world's birds and mammals: Ecological Archives E095-178. *Ecology* **95**, 2027 (2014).
35. Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C. & Costa, G. C. AmphiBIO, a global database for amphibian ecological traits. *Sci. Data* **4**, 170123 (2017).
36. Jones, K. E. et al. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**, 2648–2648 (2009).
37. Faurby, S. et al. PHYLACINE 1.2: the phylogenetic atlas of mammal macroecology. *Ecology* **99**, 2626 (2018).
38. Galán-Acedo, C., Arroyo-Rodríguez, V., Andresen, E. & Arasa-Gisbert, R. Ecological traits of the world's primates. *Sci. Data* **6**, 55 (2019).
39. Flores-Moreno, H. et al. *fungaltraits aka fun^fun: a dynamic functional trait database for the world's fungi* (GitHub, 2019); https://doi.org/10.5281/zenodo.1216257.
40. Sholler, D., Ram, K., Boettiger, C. & Katz, D. S. Enforcing public data archiving policies in academic publishing: A study of ecology journals. *Big Data Soc.* **6**, 2053951719836258 (2019).
41. Fegraus, E. H., Andelman, S., Jones, M. B. & Schildhauer, M. Maximizing the value of ecological data with structured metadata: an introduction to Ecological Metadata Language (EML) and principles for metadata creation. *Bull. Ecol. Soc. Am.* **86**, 158–168 (2005).
42. Parker, T. H. et al. Transparency in ecology and evolution: real problems, real solutions. *Trends Ecol. Evol.* **31**, 711–719 (2016).
43. Hortal, J. et al. Seven shortfalls that beset large-scale knowledge of biodiversity. *Annu. Rev. Ecol. Syst.* **46**, 523–549 (2015).
44. Cornwell, W. K., Pearse, W. D., Dalrymple, R. L. & Zanne, A. E. What we (don't) know about global plant diversity. *Ecography* **42**, 1819–1831 (2019).
45. Stodden, V., Seiler, J. & Ma, Z. An empirical analysis of journal policy effectiveness for computational reproducibility. *Proc. Natl Acad. Sci. USA* **115**, 2584–2589 (2018).
46. Lowndes, J. S. S. et al. Our path to better science in less time using open data science tools. *Nat. Ecol. Evol.* **1**, 0160 (2017).
47. Weigelt, P., König, C. & Kreft, H. GIFT–a global inventory of floras and traits for macroecology and biogeography. *J. Biogeogr.* https://doi.org/10.1111/jbi.13623 (2019).
48. Parker, T., Nakagawa, S. & Gurevitch, J., IIEE workshop participants. Promoting transparency in evolutionary biology and ecology. *Ecol. Lett.* **19**, 726–728 (2016).
49. McKiernan, E. C. et al. Point of view: How open science helps researchers succeed. *eLife* **5**, e16800 (2016).
50. Munafò, M. R. et al. A manifesto for reproducible science. *Nat. Hum. Behav.* **1**, 0021 (2017).
51. Nosek, B. A. et al. Promoting an open research culture. *Science* **348**, 1422–1425 (2015).
52. Farley, S. S., Dawson, A., Goring, S. J. & Williams, J. W. Situating ecology as a big-data science: current advances, challenges, and solutions. *BioScience* **68**, 563–576 (2018).
53. Falster, D. S. et al. BAAD: a Biomass And Allometry Database for woody plants. *Ecology* **96**, 1445–1445 (2015).
54. Salguero-Gómez, R. et al. COMADRE: a global data base of animal demography. *J. Anim. Ecol.* **85**, 371–384 (2016).
55. Salguero-Gómez, R. et al. The COMPADRE Plant Matrix Database: an open online repository for plant demography. *J. Ecol.* **103**, 202–218 (2015).
56. Marques, G. M. et al. The AmP project: comparing species on the basis of dynamic energy budget parameters. *PLOS Comput. Biol.* **14**, e1006100 (2018).
57. Conde, D. A. et al. Data gaps and opportunities for comparative and conservation biology. *Proc. Natl Acad. Sci. USA* **116**, 9658–9664 (2019).
58. Wieczorek, J. et al. Darwin Core: an evolving community-developed biodiversity data standard. *PLOS ONE* **7**, e29715 (2012).
59. Guralnick, R., Walls, R. & Jetz, W. Humboldt Core–toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. *Ecography* **41**, 713–725 (2018).
60. Deans, A. R. et al. Finding our way through phenotypes. *PLOS Biol.* **13**, e1002033 (2015).
61. Haendel, M. A. et al. Unification of multi-species vertebrate anatomy ontologies for comparative biology in Uberon. *J. Biomed. Semant.* **5**, 21 (2014).
62. Garnier, E. et al. Towards a thesaurus of plant characteristics: an ecological contribution. *J. Ecol.* **105**, 298–309 (2017).
63. The Gene Ontology Consortium. The Gene Ontology resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2018).
64. Buttigieg, P. L., Morrison, N., Smith, B., Mungall, C. J. & Lewis, S. E. The environment ontology: contextualising biological and biomedical entities. *J. Biomed. Semant.* **4**, 43 (2013).
65. Becker, J., Brackbill, D. & Centola, D. Network dynamics of social influence in the wisdom of crowds. *Proc. Natl Acad. Sci. USA* **114**, E5070–E5076 (2017).
66. Page, S. E. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies - New Edition* (Princeton Univ. Press, 2008).
67. Tenopir, C. et al. Data sharing by scientists: practices and perceptions. *PLOS ONE* **6**, e21101 (2011).
68. Tyler, E. H. et al. Extensive gaps and biases in our knowledge of a well-known fauna: implications for integrating biological traits into macroecology. *Glob. Ecol. Biogeogr.* **21**, 922–934 (2012).
69. Kissling, W. D. et al. Towards global data products of Essential Biodiversity Variables on species traits. *Nat. Ecol. Evol.* **2**, 1531–1540 (2018).
70. Lajoie, G. & Kembel, S. W. Making the most of trait-based approaches for microbial ecology. *Trends Microbiol.* **27**, 814–823 (2019).
71. Dawson, S. K. et al. Handbook for the measurement of macrofungal functional traits: a start with basidiomycete wood fungi. *Funct. Ecol.* **33**, 372–387 (2019).
72. Ankenbrand, M. J., Hohlfeld, S. C., Weber, L., Förster, F. & Keller, A. Functional exploration of natural networks and ecological communities. *Methods Ecol. Evol.* **9**, 2028–2033 (2018).
73. Gaillard, J.-M. et al. Generation time: a reliable metric to measure life-history variation among mammalian populations. *Am. Nat.* **166**, 119–123 (2005).
74. Steiner, U. K., Tuljapurkar, S. & Coulson, T. Generation time, net reproductive rate, and growth in stage-age-structured populations. *Am. Nat.* **183**, 771–783 (2014).
75. Andelman, S. J., Bowles, C. M., Willig, M. R. & Waide, R. B. Understanding environmental complexity through a distributed knowledge network. *BioScience* **54**, 240–246 (2004).
76. Schneider, F. D. et al. Towards an ecological trait-data standard. *Methods Ecol. Evol.* **10**, 2006–2019 (2019).
77. Perez-Harguindeguy, N. et al. A new handbook for standardised measurement of plant functional traits worldwide. *Aust. J. Bot.* **64**, 715–716 (2013).
78. Fang, J. et al. Methods and protocols for plant community inventory. *Biodivers. Sci.* **17**, 533–548 (2009).
79. Sack, L. et al. A unique web resource for physiology, ecology and the environmental sciences: PrometheusWiki. *Funct. Plant Biol.* **37**, 687–693 (2010).
80. Bjorkman, A. D. et al. Tundra Trait Team: a database of plant traits spanning the tundra biome. *Glob. Ecol. Biogeogr.* **27**, 1402–1411 (2018).

81. Moretti, M. et al. Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Funct. Ecol.* **31**, 558–567 (2017).

82. Ferris, H. *NEMAPLEX: The Nematode-Plant Expert Information System* (Univ. California Davis, 2005); http://nemaplex.ucdavis.edu/

83. Tennessen, J. M., Barry, W. E., Cox, J. & Thummel, C. S. Methods for studying metabolism in *Drosophila*. *Methods* **68**, 105–115 (2014).

84. Palomares, M. L. D. & Pauly, D. SeaLifeBase v.12/2010 (2010); www.sealifebase.org

85. Le Bagousse-Pinguet, Y. et al. Traits of neighbouring plants and space limitation determine intraspecific trait variability in semi-arid shrublands. *J. Ecol.* **103**, 1647–1657 (2015).

86. Cornelissen, J. et al. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Aust. J. Bot.* **51**, 335–380 (2003).

87. Maitner, B. S. et al. The bien r package: a tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods Ecol. Evol.* **9**, 373–379 (2018).

88. Jetz, W., Thomas, G., Joy, J., Hartmann, K. & Mooers, A. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).

89. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302–314 (2018).

90. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).

91. Díaz, S. et al. The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).

92. Andersen, K. H. et al. Characteristic sizes of life in the oceans, from bacteria to whales. *Annu. Rev. Mar. Sci.* **8**, 217–241 (2016).

93. Neuheimer, A. B. et al. Adult and offspring size in the ocean over 17 orders of magnitude follows two life history strategies. *Ecology* **96**, 3303–3311 (2015).

94. Ernest, S. M. et al. Thermodynamic and metabolic effects on the scaling of production and population energy use. *Ecol. Lett.* **6**, 990–995 (2003).

95. Weiss, K. C. & Ray, C. A. Unifying functional trait approaches to understand the assemblage of ecological communities: synthesizing taxonomic divides. *Ecography* **42**, 2012–2020 (2019).

96. Ball, I. R., Possingham, H. P. & Watts, M. in *Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools* (eds Moilanen, A. et al.) 185–195 (Oxford Univ. Press, 2009).

97. Pollock, L. J., Thuiller, W. & Jetz, W. Large conservation gains possible for global biodiversity facets. *Nature* **546**, 141–144 (2017).

98. Margules, C. R. & Pressey, R. L. Systematic conservation planning. *Nature* **405**, 243–253 (2000).

99. Gross, N. et al. Functional trait diversity maximizes ecosystem multifunctionality. *Nat. Ecol. Evol.* **1**, 0132 (2017).

100. Loreau, M. Does functional redundancy exist? *Oikos* **104**, 606–611 (2004).

101. van Bodegom, P. M., Douma, J. C. & Verheijen, L. M. A fully traits-based approach to modeling global vegetation distribution. *Proc. Natl Acad. Sci. USA* **111**, 13733–13738 (2014).

102. Sakschewski, B. et al. Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model. *Glob. Change Biol.* **21**, 2711–2725 (2015).

103. Butler, E. E. et al. Mapping local and global variability in plant trait distributions. *Proc. Natl Acad. Sci. USA* **114**, E10937–E10946 (2017).

104. Kearney, M. & Porter, W. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecol. Lett.* **12**, 334–350 (2009).

105. Fordham, D. A. et al. How complex should models be? Comparing correlative and mechanistic range dynamics models. *Glob. Change Biol.* **24**, 1357–1370 (2018).

106. Enriquez-Urzelai, U., Kearney, M. R., Nicieza, A. G. & Tingley, R. Integrating mechanistic and correlative niche models to unravel range-limiting processes in a temperate amphibian. *Glob. Change Biol.* **25**, 2633–2647 (2019).

107. Benito Garzón, M., Robson, T. M. & Hampe, A. ΔTrait SDMs: species distribution models that account for local adaptation and phenotypic plasticity. *New Phytol.* **222**, 1757–1765 (2019).

108. Berzaghi, F. et al. Assessing the role of megafauna in tropical forest ecosystems and biogeochemical cycles–the potential of vegetation models. *Ecography* **41**, 1934–1954 (2018).

109. Galetti, M. & Dirzo, R. Ecological and evolutionary consequences of living in a defaunated world. *Biol. Conserv.* **163**, 1–6 (2013).

110. Huang, Y. et al. Orchimic (v1. 0), a microbe-mediated model for soil organic matter decomposition. *Geosci. Model Dev.* **11**, 2111–2138 (2018).

111. McGuire, K. L. & Treseder, K. K. Microbial communities and their relevance for ecosystem models: decomposition as a case study. *Soil Biol. Biochem.* **42**, 529–535 (2010).

112. Todd-Brown, K. E., Hopkins, F. M., Kivlin, S. N., Talbot, J. M. & Allison, S. D. A framework for representing microbial decomposition in coupled climate models. *Biogeochemistry* **109**, 19–33 (2012).

113. Hardisty, A. R. et al. The Bari Manifesto: an interoperability framework for essential biodiversity variables. *Ecol. Inform.* **49**, 22–31 (2019).

Author contributions

R.V.G. wrote the manuscript with contributions from D.S.F., B.S.M., R.S.-G., V.V., W.D.P., F.D.S., J.K., J.H.P., J.S.M., M.J.A., C.P., X.F., V.M.A., J.A., S.C.A., M.A.B., L.M.B., B.L.B., C.H.B.-A., I.B., A.J.R.C., R.C., B.R.C., D.A.C., S.L.C., B.F., H.G., A.H.H., J.H., J.A.H., H.H., M.H., C.M.I., M.J., M.K., A.K., P.Mabee, P.Manning, L.M., S.T.M., D.S.P., T.M.P., S.P.-M., C.A.R., M.R., H.S., B.S., M.J.S., R.J.T., J.A.T., C.V., R.W., K.C.B.W., M.W., I.J.W. and B.J.E.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to R.V.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[1]Department of Biological Sciences, Macquarie University, Sydney, New South Wales, Australia. [2]Evolution and Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales, Australia. [3]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA. [4]Department of Zoology, Oxford University, Oxford, UK. [5]Centre for Biodiversity and Conservation Science, University of Queensland, Brisbane, Queensland, Australia. [6]Evolutionary Demography Laboratory, Max Plank Institute for Demographic Research, Rostock, Germany. [7]Department of Biological Sciences, University of Bergen, Bergen, Norway. [8]Bjerknes Centre for Climate Research, University of Bergen, Bergen, Norway. [9]Ecology Center and Department of Biology, Utah State University, Logan, UT, USA. [10]Max Planck Institute for Biogeochemistry, Jena, Germany. [11]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. [12]Hawai'i Institute of Marine Biology, University of Hawai'i at Manoa, Manoa, HI, USA. [13]Department of Bioinformatics, Biocenter, University of Wuerzburg, Wuerzburg, Germany. [14]Center for Computational and Theoretical Biology, Biocenter, University of Wuerzburg, Wuerzburg, Germany. [15]Comprehensive Heart Failure Center, University Hospital Wuerzburg, Wuerzburg, Germany. [16]Institute of Plant Sciences, University of Bern, Bern, Switzerland. [17]Discipline of Geography and Spatial Sciences, University of Tasmania, Hobart, Tasmania, Australia. [18]Commonwealth Scientific and Industrial Research Organisation (CSIRO), Canberra, Australian Capital Territory, Australia. [19]Bio5 Institute, University of Arizona, Tucson, AZ, USA. [20]School of Life and Environmental Sciences, Centre for Integrative Ecology, Deakin University, Geelong, Victoria, Australia. [21]Department of Biology, University of Miami, Miami, FL, USA. [22]Fairchild Tropical Botanic Garden, Coral Gables, FL, USA. [23]Research School of Biology, Australian National University, Canberra, Australian Capital Territory, Australia. [24]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA, USA. [25]Species360 Conservation Science Alliance,

Bloomington, MN, USA. [26]Interdisciplinary Center on Population Dynamics, University of Southern Denmark, Odense, Denmark. [27]Department of Biology, University of Southern Denmark, Odense, Denmark. [28]School of Biological Sciences, Monash University, Melbourne, Victoria, Australia. [29]Department of Ecology, Environment and Evolution and Centre for Future Landscapes, La Trobe University, Melbourne, Victoria, Australia. [30]National Museum of Natural History, Smithsonian Institution, Washington, DC, USA. [31]International Center for Tropical Botany, Department of Biological Sciences, Florida International University, Miami, FL, USA. [32]Climate Change Science Institute and Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA. [33]Institute of Biology, Leipzig University, Leipzig, Germany. [34]School of BioSciences, The University of Melbourne, Melbourne, Victoria, Australia. [35]Department of Biology, University of South Dakota, Vermillion, SD, USA. [36]Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Frankfurt, Germany. [37]Center for Tree Science, The Morton Arboretum, Lisle, IL, USA. [38]Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada. [39]Department of Organismic and Evolutionary Biology and Harvard University Herbaria, Harvard University, Cambridge, MA, USA. [40]School of Biological Sciences and School of Earth & Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA. [41]School of Life Sciences, Arizona State University, Tempe, AZ, USA. [42]National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Sydney, New South Wales, Australia. [43]Queensland Alliance of Agriculture and Food Innovation, University of Queensland, Brisbane, Queensland, Australia. [44]Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Universite Paris-Saclay, Orsay, France. [45]TERN / School of Biological Sciences, Faculty of Science, The University of Adelaide, Adelaide, South Australia, Australia. [46]Department of Evolution, Ecology, and Organismal Biology, University of California Riverside, Riverside, CA, USA. [47]Department of Life Sciences, Imperial College London, London, UK. [48]CEFE, CNRS, Univ Montpellier, Université Paul Valéry Montpellier, Montpellier, France. [49]CyVerse, University of Arizona, Tucson, AZ, USA. [50]Santa Fe Institute, Santa Fe, NM, USA. [51]Unaffiliated: Florian D. Schneider, Jorrit H. Poelen. *e-mail: rachael.gallagher@mq.edu.au