CHAPTER 20

TraitNet: furthering biodiversity research through the curation, discovery, and sharing of species trait data

Shahid Naeem and Daniel E. Bunker

20.1 Introduction: traits and ecosystem functioning

Trait-based ecological and evolutionary research has undergone an extraordinary expansion in the last thirty years (Fig. 20.1), and species traits, broadly defined, have been key to advances in many fields of natural science (Table 20.1). For example, the evolutionary ecology of species' niches involves fundamental trade-offs in seed size (Moles *et al.* 2005), leaf economic traits (Ackerly 2004, Wright *et al.* 2004), and allometric constraints (West *et al.* 1997). Traits have been used to predict the risk of species invasions (Veltman *et al.* 1996, Kolar and Lodge 2002, Lloret *et al.* 2005, Ruesink *et al.* 2005), the risk of species extinction (Gittleman and Purvis 1998, Foufopoulos and Ives 1999, Purvis *et al.* 2000a), and crop responses to climate change (Lynch *et al.* 2004). Table 20.1 summarizes these and other examples of basic and applied research that are dependent on species trait data.

While early efforts to understand the effects of biodiversity on ecosystem functioning focused on species richness, recent efforts have recognized the role of functional diversity as the driver of these effects (Petchey *et al.* (Chapter 4), Diaz and Cabido 2001, Petchey *et al.* 2004b). The functional traits of species are the means by which species interact with their environment, and thus are directly responsible for the effects of species on ecosystem processes (e.g. Gordon 1998, Eviner and Chapin 2003, Kelso *et al.* 2003, Diaz *et al.* 2004, Eviner 2004), as well as the response of species to environmental change (Grime 1979, Tilman 1982, Huston and Smith 1987, Brown 1995, Enquist *et al.* 2003, Brown *et al.* 2004, Navas and Moreau-Richard 2005, McGill *et al.* 2006).

The BioMERGE Research Coordination Network strove to expand biodiversity–ecosystem functioning research to larger scales by developing models to predict functioning from species traits (Naeem *et al.* (Chapter 1), Duffy *et al.* (Chapter 5), Naeem *et al.* 2007, Naeem 2008). For example, Solan *et al.* (2004) used traits to estimate changes in estuarine sediment turnover in the face of biodiversity loss and Bunker *et al.* (2005) forecasted changes in forest carbon sequestration under different management practices. Another example is that of McIntyre *et al.* (2007) who estimated the influence of fish biodiversity on freshwater ecosystem functions. These early efforts primarily addressed selection effects (SE; see Chapter 7) by utilizing traits that predict *per capita* effects on functioning in combination with various traits associated with extinction risk. These efforts ignored complementarity effects (CE; see Chapter 7), in part because we still know relatively little about which traits may lead to complementarity, but also because species trait data are, at best, dispersed throughout the literature, and at worst lacking altogether. Indeed, even as the development of comprehensive vegetation databases (e.g. VegBank, Center for Tropical Forest Science, and SALVIAS) and phylogenetic databases

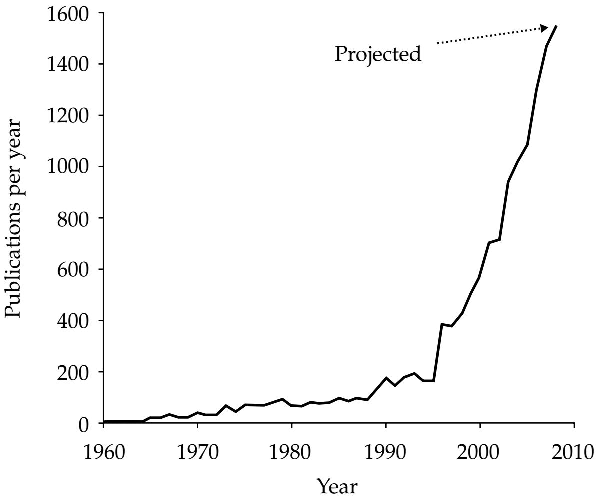

Figure 20.1 Exponential increase in ecological and evolutionary research using traits. Results from Scopus search for (("ecology" or "evolution") and "traits").

(e.g. Flora of North America, Tree of Life project) has proceeded rapidly, trait databases, where they exist at all, have remained specialized to particular regions, taxa, or sets of traits. This chapter describes TraitNet, a recently established research coordination network designed to facilitate trait-based research in general, and this new direction in biodiversity and ecosystem functioning research in particular.

20.2 TraitNet: enabling trait-based research

TraitNet aims to facilitate collaboration between ecologists and evolutionary biologists who utilize species trait data. Traits are used across a broad spectrum of disciplines, including ecology, evolution, and conservation biology. While each discipline has developed its own operational definitions, protocols, and databases, there is little coordination across disciplines. Because of these diverse uses and definitions of species traits, TraitNet takes an expansive view of what may constitute a 'trait', including most any character that can be quantified at the individual level. We take this approach because, while our efforts are motivated by the need to understand and quantify the traits that drive ecosystem functioning, we intend TraitNet to support the broad range of research that utilizes species trait data.

In order to facilitate this new, cross-disciplinary, collaborative, trait-based research, TraitNet will pursue five primary goals: (1) identify key questions and core hypotheses in trait-based research, (2) identify data gaps that hinder the advancement of intra- and inter-disciplinary trait-based research, (3) coordinate the standardization of collection and curation of trait data, (4) build a model database to address core hypotheses, and (5) facilitate the development of cross-disciplinary computational tools for merging, disseminating, and sharing trait data. There are several core activities, including workshops, online collaboratories, and database development (Fig. 20.2).

TraitNet aims to bring together species trait data from a variety of taxa across different trophic levels and from a variety of habitats and locations to address specific interdisciplinary hypotheses (see examples below). Our initial coverage will be greatest among terrestrial plants because several plant trait networks are well developed and will serve as useful starting points (Table 20.2). The interdisciplinary hypotheses chosen for study by the network of participants will likely require traits of herbivores, predators, detritivores, and other trophic groups in addition to traits of primary producers. Additionally, TraitNet will not limit itself to terrestrial systems. It will also examine aquatic ecosystems or transition habitats such as wetlands. To that end, we have assembled a group of core participants that is weighted towards terrestrial plant ecologists due to current advances in

Table 20.1 A broad sampling of trait-based research where trait is broadly defined. Note that additional research areas include trait-based taxonomy, trait-based phylogenetics, and morphometrics.

Subject area	Description	Example traits	References
Bioremediation	Using species to remediate pollution	Heavy metal resistance, specific root length, root surface area, root volume and average root diameter	Von Canstein *et al.* 2002, Pulford and Watson 2003, Merkl *et al.* 2005
Biodiversity and ecosystem functioning	Mechanisms by which changes in biodiversity change ecosystem functioning	Species response and effect traits	Lavorel and Garnier 2002, Solan *et al.* 2004, Bunker *et al.* 2005, Thompson *et al.* 2005
Comparative method	Using phylogenies and traits to understand evolutionary adaptation	Traits such as leaf mass per area, seed mass, genome size	Ackerly 2004, Moles *et al.* 2005
Community ecology	How trait filtering governs community composition and structure, including assembly rules, competition, facilitation and limiting similarity	Body size, height, leaf traits, trophic position, light requirements, clonality	Gaudet and Keddy 1988, Weiher and Keddy 1995a, Weiher and Keddy 1995b, Weiher and Keddy 1999, Ackerly *et al.* 2002, Suding *et al.* 2003, Cavender-Bares *et al.* 2004, Suding *et al.* 2005, Grime 2006, McGill *et al.* 2006
Conservation biology	Estimate threat levels for species or likelihood of extinction	Gestation period, range size, number of offspring, trophic position	Gittleman and Purvis 1998, Foufopoulos and Ives 1999, Purvis *et al.* 2000a
Ecosystem ecology	Trait-specific influences of organisms on ecosystem processes and biogeochemistry	Woody caudices, multi-branched rhizomes, N-fixing symbiotic associations, C3 or C4 pathway	Gordon 1998, Eviner and Chapin 2003, Kelso *et al.* 2003, Diaz *et al.* 2004, Eviner 2004
Gradient analysis	Mechanisms and patterns of biodiversity along ecological gradients	R^*, dispersal mode, reproductive structures, elements of leaf design	Tilman and Wedin 1991, Thuiller *et al.* 2004
Endemism	Determining what traits are associated with endemism	Stature, dispersal, pollen/ovule ratios, number of flowers	Lavergne *et al.* 2003, Lavergne *et al.* 2004
Fire ecology	Predicting fires based on plant traits	Resprouting capability, seed bank	Saha and Howe 2003, Pausas *et al.* 2004
Food webs	Structure and dynamics of communities governed by trophic interactions	Dietary or energy transfer linkages and trophic position, body size, morphological traits	Layman *et al.* 2005
Functional diversity	Identification and quantification of functional diversity	All functional traits	Petchey and Gaston 2002a, Mason *et al.* 2003, Botta-Dukat 2005, Mouillot *et al.* 2005a, Mouillot *et al.* 2005c
Guild analysis	Grouping species by environmental resource exploitation irrespective of taxonomy	C3, C4, annuals and biennial forbs, ephemeral spring forbs, spring forbs, summer/fall forbs, legumes, woody shrubs	Simberloff and Dayan 1991, Kindscher and Wells 1995, Blondel 2003
Heritability	Quantifying the heritability of species traits	Various traits	Iyengar *et al.* 2002, Caruso *et al.* 2005, Garant *et al.* 2005
Macroecology	Patterns of species adaptations at geographic scales	Body size, photosynthetic pathway, dispersal syndrome	Brandle *et al.* 2002, Burns 2004, Morin and Chuine 2006

(*Continues*)

Table 20.1 (*continued*)

Subject area	Description	Example traits	References
Metabolic theory of ecology	Metabolism as a basis for linking individual organisms to population, community, and ecosystem ecology.	Body size, physiological traits, and correlates such as growth, range	Brown 1995, West *et al.* 1997, Enquist *et al.* 2003, Brown *et al.* 2004
Natural selection	How species evolve in response to selective pressure	Body size, fledging weight, dispersal syndrome, palatability, host specificity	Boughman 2001, Etterson and Shaw 2001, Merila *et al.* 2001, Nosil *et al.* 2002, Beatty *et al.* 2004, Hoskin *et al.* 2005
Palaeobiology	Using leaf physiognomy to estimate past climate	Leaf size, leaf morphology	Royer *et al.* 2005, Royer and Wilf 2006
Plant ecological strategies	How plant species 'secure carbon profit during vegetative growth and ensure gene transmission into the future'	Leaf-span-per-area, leaf-lifespan	Hodgson *et al.* 1999, Westoby *et al.* 2002
Population ecology	Predicting properties of dynamics (e.g. probability of extinction)	Body size, age at first reproduction, or number of offspring	Fagan *et al.* 2001, McGill *et al.* 2006
Species invasions	Predicting invasiveness of species based on traits	Body size, endemism, reproductive rate	Veltman *et al.* 1996, Kolar and Lodge 2002, Lloret *et al.* 2004, Hamilton *et al.* 2005, Lloret *et al.* 2005, Ruesink *et al.* 2005
Succession	Temporal change in communities predicted by traits	Respiration rate, seed number, growth rate, maximum life span, induced dormancy, R^*, stress tolerance	Grime 1979, Tilman 1982, Huston and Smith 1987, Navas and Moreau-Richard 2005
Unified neutral theory of biodiversity and biogeography	Trait neutral core patterns in distribution and abundance forms contrasting hypotheses to trait-based patterns	Dispersal, growth rates	Hubbell 2001, Nee and Stone 2003, He and Hu 2005, Hubbell 2005, Ostling 2005, Wootton 2005

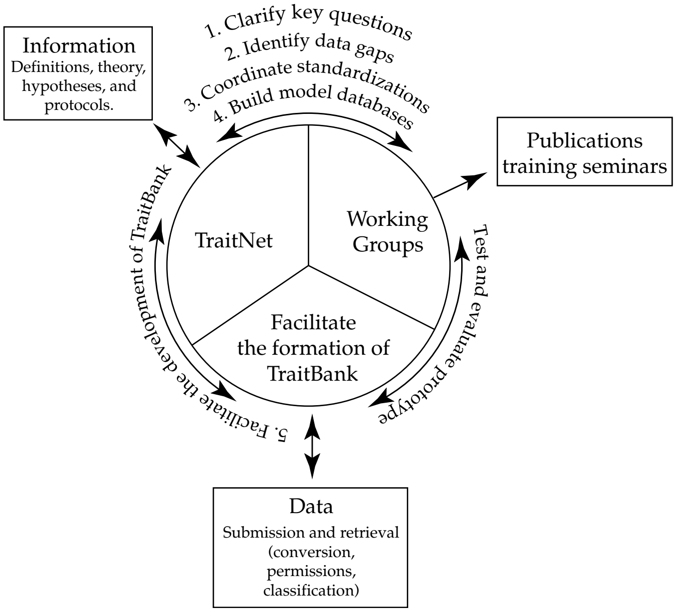

Figure 20.2 Architecture of TraitNet. Five objectives structure TraitNet, an NSF-funded Research Coordination Network that coordinates the development of cross-disciplinary trait-based research. Working groups consist of members across trait-based disciplines that will collectively clarify key questions within and across disciplines, identify data gaps, coordinate standardizations, and build model databases to test hypotheses that emerge from syntheses. Working groups will provide workshops and training sessions to broadly disseminate TraitNet results. Finally, TraitNet will coordinate the design of a prototype universal database entitled TraitBank.

the field, but which also includes researchers who specialize in insect, mammalian, microbial, aquatic, and disease ecology. These core participants are actively recruiting additional investigators within their respective areas of expertise. TraitNet participation is expected to grow substantially as we identify additional researchers who focus on other habitats, taxonomic groups, and trophic groups.

Hypotheses derived from cross-disciplinary research are often characterized by a scope that requires multiple traits collected uniformly from a diversity of species across widely dispersed localities. The TraitNet working groups will explore and identify such hypotheses, determine data needs and gaps, develop and test model datasets for addressing these hypotheses, and coordinate the establishment of trait databases, guidelines, and training for their use, and enable a variety of multidisciplinary research activities. As Table 20.1 illustrates, the potential number of cross-disciplinary approaches is very large. Here we provide three of the many possible hypotheses TraitNet will explore:

- *Dimensionality of life-history trade-offs.* While an endless number of traits can be measured on individuals and species, many traits are highly correlated with one another, and it has been suggested that a small number trade-off axes can explain the majority of variation in plant form and function (Grime 1979, Coley *et al.* 1985, Charnov 1997, Reich *et al.* 1999, Hubbell 2001, Westoby *et al.* 2002, Wright *et al.* 2004). How are species life histories constrained by these fundamental trade-offs, how many axes of differentiation exist, and how does the extent of these trade-offs vary across environmental gradients and among biomes? These key questions require data on multiple traits, collected from multiple species, from multiple sites, and standardized when different protocols were used.
- *Mechanisms of exotic species invasions*. The success of invasive species has often been attributed to an escape from natural enemies, whereby one would predict successful invaders to have ‘better’ traits than the native species they displace, such as greater height, lower R^*, or lower construction costs (Miller and Werner 1987, Gaudet and Keddy 1988, Nagel and Griffin 2001, Seabloom *et al.* 2003, Bunker 2004). Alternatively, the success of some invaders has been attributed to novel traits, such as nitrogen fixation (Vitousek and Walker 1989) or allelopathic effects (Bais *et al.* 2003), that allow them to dominate new habitats. While both mechanisms certainly play a strong role, the relative importance of each in driving species invasions is not clear. An effective

Table 20.2 Examples of databases that include trait data. Current collaborators in bold.

Database name	Description
BiolFlor	Focuses on the German flora and includes > 60 traits and > 3600 plant species
BioPop	Database of plant traits of the Mid-European flora including 60 traits and > 4,700 species
Center for Tropical Forest Science Trait Database	A newly initiated effort to collect functional trait data for 6,200 tree species found in 18 large forest dynamics plots located in 14 tropical countries
Ecological flora of the British Isles	Database of plant traits of the flora of the British Isles including > 130 traits and > 1700 plant species
Ecological flora of California	A database of ecological characteristics, including life history, phenology, morphology and other traits for the California flora. Under development by David Ackerly
FishBase	Worldwide fish species database with more than 29,000 species
Glopnet	Global compilation of leaf economic traits from > 2500 plant species. Initiated by Peter Reich, Ian Wright, and Mark Westoby
Hawaii Plant Trait Database	A database of systematic, biogeographic, functional, physiological, and ecological data for Hawaii's native and alien flora. Initiated by Rebecca Montgomery, Lawren Sack, Becky Ostertag, Susan Cordell, and Jon Price
LEDA Traitbase	Focuses on the Northwest European flora and plant traits that describe three key features of plant dynamics: persistence, regeneration, and dispersability
NatureServe Explorer	Conservation data on more than 50,000 plants, animals, and ecological communities of the USA and Canada
NatureServe InfoNatura	Conservation information on the more than 5,500 birds, mammals, and amphibians of Latin America and the Caribbean
Seed Information Database, Kew Botanic Garden	Database of seed characteristics, including storage behavior, weight, dispersal, germination, oil content, protein content, morphology, for several thousand plant species, with plans to include > 24,000 species
TRY	Global plant trait database to develop new plant functional classifications for earth system modelling
USDA PLANTS	The PLANTS Database provides standardized information about the vascular plants, mosses, liverworts, hornworts, and lichens of the USA and its territories. It includes names, distributional data, characteristics, images, and crop information

test would require species trait data on plant invader species, on the native species they may displace, on palatability to native herbivores, and data on traits of potential natural enemies such as body size, diet, and growth rate.

- *Predicting species, community and ecosystem responses to global change.* Predicting the response of species to climate change, pollution, and land use change is a key challenge to ecologists. These predictions could be developed by correlating species traits with either observed responses to global drivers or across natural environmental gradients. In either case trait data from a wide variety of species, across multiple trophic levels, from a variety of habitats would be required. Similarly, predicting the effects of these global drivers on ecosystem function will require additional trait data that mechanistically link species with their *per capita* effects on ecosystem functioning (Etterson and Shaw 2001, Solan *et al.* 2004, Bunker *et al.* 2005).

20.3 The challenges of data integration

Much of what TraitNet aims to accomplish will rely on integrating data from disparate sources. Integrating disparate data is a complex process with many challenges. These challenges are not trivial. Here we outline the main issues and describe our approach to addressing them. In order to meet these challenges, we have included several informatics experts in our group of core participants, representing several organizations including the Science Environment for Ecological Knowledge project (SEEK), the Pacific Ecoinformatics and Computational Ecology Lab, the National Center for Ecological Analysis and Synthesis (NCEAS), the National Evolutionary Synthesis Center, and the Microsoft European Science Initiative. To address our Core Hypotheses, we will build a model database that also will serve as a training ground for building a fully accessible and open source trait data archive termed TraitBank.

20.3.1 Intellectual property rights

Intellectual property rights are a critical issue for any research network and even more so when data are aggregated from multiple sources. Trait-based research progresses best when data sharing is maximal, but currently the sharing of raw data is not common except within groups. Workshops, collaboratories, training sessions, and the TraitNet website will provide a forum for discussion of the many issues surrounding intellectual property rights and how they would affect database tools, resources, and the design and implementation of TraitBank in the future.

To that end, we propose that data owners will retain full rights and full control over their data. TraitNet will facilitate collaborations between participants that would otherwise be less likely to occur. Our model database will be fully searchable, whereby one could search for all available traits for a particular species, or all species with a particular trait, or for a set of traits among a set of species. If participants so choose, we can set up the search system to return only whether the data exist, and who owns the data. It would then be up to the participant to contact the data owner and propose collaboration.

20.3.2 Taxonomic standardization

Definitions of biological taxa change with taxonomic revisions over time. For instance, a single species may be split by one revision into several species, and then lumped back into a single species in subsequent revisions. A trait value measured on the lumped species cannot be assigned to any one of the split species, and a trait value measured on one of the split species cannot be assumed to represent the entirety of the lumped species. In addition, species are often cited with only the name authority, but not the underlying taxon concept reference. For these reasons, taxonomic names by themselves cannot be considered a unique index for TraitNet datasets. This obstacle applies to all data that are specific to individual species, such as GenBank and VegBank.

Fortunately, efforts are under way to address this complex issue. The SEEK Taxon project has created an internationally accepted standard for taxonomic data, the Taxonomic Concept Schema (TCS), and work is under way to implement the Taxonomic Object Service (TOS), a repository and web service allowing for translation between taxonomic concept authorities (Graham and Kennedy 2007). TraitNet will collaborate with SEEK Taxon to ensure that our database schema will function with the TOS when it is fully functional and populated with taxon concepts.

To ensure that TraitNet data will be compatible with TOS, TraitNet will require that participants include name authorities and taxon concept references, as well as subspecies when appropriate, in their data submissions. Only in this way can we specify, for each record, the taxonomic concept upon which the measurement was taken. For instance, a full taxon concept reference might be: "*Aus beus* Sarg. 1893 Smith 1989," where Smith 1989 is a link to the reference in which the concept is described or defined (Fig. 20.3).

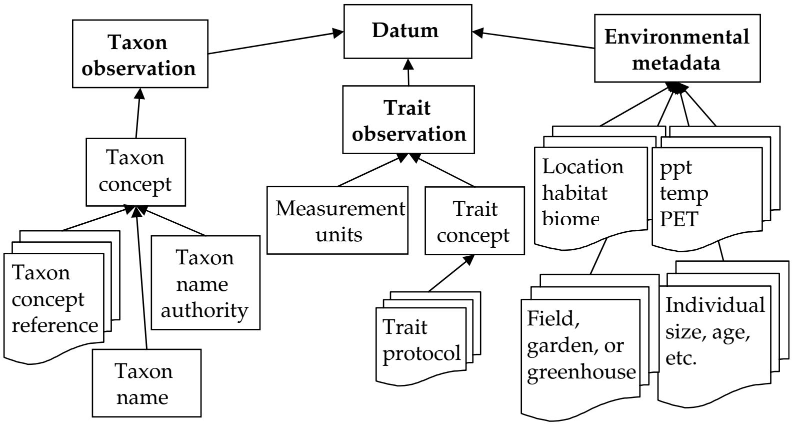

Figure 20.3 TraitNet trait observation schema. The ideal trait observation will include critical metadata, such as trait protocols, taxon concepts and environmental data. These metadata will make trait observations robust to various future uses of the trait data and thus ensure the longevity and usefulness of trait data collections.

20.3.3 Trait protocol standardization

Variations in collection protocols introduce challenges that are similar to those introduced by taxonomic revisions. A trait concept may remain fixed, but the protocol used to quantify the trait may change as new protocols are introduced. A trait database must be able to incorporate revised trait collection protocols as they are developed to ensure that data produced though all protocols for a given trait concept are quantitatively comparable. For example, wood density is the trait concept of mass per unit volume. However the protocol to collect wood density varies. Wood mass may be measured on oven-dry wood samples or on air-dried samples with 12–15 per cent moisture content. Both metrics quantify wood density, but data from air-dried samples must be corrected to account for the moisture content. TraitNet will define trait concepts and associated trait collection protocols. Each trait protocol for a given trait concept must be quantitatively comparable (Fig. 20.3).

20.4 Integrating TraitNet into ongoing ecoinformatics frameworks

TraitNet will build on current ecoinformatic efforts to address these issues. Ecological Metadata Language (EML) has been developed by NCEAS' Knowledge Network for Biocomplexity project and is widely considered the standard for documenting metadata for ecological datasets. The SEEK project has extended and formalized critical aspects of EML in the Observation Ontology (OBOE), a formal model of scientific observations that includes trait measurements (Madin *et al.* 2007, Madin *et al.* 2008). Thus TraitNet will use and extend EML to specify the Taxon Concepts, Trait Concepts, and associated environmental data outlined in Fig. 20.3. Eventually, these trait concepts will be included in SEEK's formal ontologies such as OBOE.

Because species trait data are used for such diverse research ends, the data must be collected and archived with sufficient metadata to ensure wide applicability to potentially unforeseen research questions. For instance, an investigator may collect wood density data with the intention of calculating above-ground biomass at a given site. However, future investigators may ask how wood density varies within species at a given site, within species across environmental gradients, among species, among size classes, or even throughout the year. To ensure that a given trait observation contains maximum scientific value, the collector will want to document explicitly the conditions under which it is collected, including latitude and longitude, habitat, individual age or body size, time and date of collection, etc. Many trait observations, such as wood density, require substantial effort and/or expense to collect. Only minimal additional effort is required to collect substantial metadata that will ensure that a trait observation has lasting scientific value (Fig. 20.3).

Finally, an ideal network environment would also: (1) allow for automated integration of trait data contributions; (2) include a web-enabled search engine that would allow user-friendly access to the general public, including students, educators, and policymakers; and (3) enable seamless access to related community, phylogenetic, and environmental databases.

As the understanding and appreciation of functional biodiversity grows among the public, TraitBank has the potential to be much more than a resource for natural scientists. For example, a birder may wish to learn more about the functional traits of a group of species found at a particular site, or a farmer or natural resource manager might want to gain additional insights into the types of weeds or invasive species found on site. A teacher on a field trip may wish to design an exercise based on local species, or a student may require trait data for their paper or science project. In this way, TraitBank will offer the public detailed trait data for individual species, and thus augment existing resources such as the Global Biodiversity Information Facility (http://www.gbif.org/).

20.5 Species traits, functional diversity, and the future of biodiversity research

TraitNet is designed to coordinate a wide array of scientific disciplines that are centred on a specific research theme but would benefit enormously from cross-disciplinary coordination. TraitNet will facilitate cross-disciplinary research among ecological and evolutionary fields centered on trait-based research. All of these diverse activities rely strongly on understanding the functional characteristics, or traits, of species. It is only through these traits that scientists can understand and predict the responses of species to their environment as well as species effects on ecosystem functioning and the ecosystem services that humans increasingly demand from both natural and managed ecosystems (Naeem *et al*., Chapter 5; Jackson *et al*., Chapter 13). Indeed, this shift towards trait-based ecology and functional diversity will bring a new perspective on biological diversity to a wide array of fundamental and applied scientists, students, and educators in ecology, evolution, and environmental biology.