Received: 16 January 2020 | Revised: 7 August 2020 | Accepted: 30 August 2020

DOI: 10.1111/gcb.15344

TECHNICAL ADVANCE

Global Change Biology WILEY

Trait-based ecology at large scales: Assessing functional trait correlations, phylogenetic constraints and spatial variability using open data

Martin A. Wilkes[1] | François Edwards[2] | J. Iwan Jones[3] | John F. Murphy[3] | Judy England[4] | Nikolai Friberg[5,6] | Daniel Hering[7] | N. LeRoy Poff[8] | Philippe Usseglio-Polatera[9] | Wilco C. E. P. Verberk[10] | Jon Webb[11] | Lee E. Brown[12]

[1]Centre for Agroecology, Water and Resilience, Coventry University, Ryton-on-Dunsmore, UK

[2]Centre for Ecology and Hydrology, Wallingford, UK

[3]Queen Mary University of London, London, UK

[4]Environment Agency, Wallingford, UK

[5]Norwegian Institute for Water Research, Oslo, Norway

[6]University of Copenhagen, Copenhagen, Denmark

[7]University of Duisburg-Essen, Duisburg, Germany

[8]Colorado State University, Fort Collins, CO, USA

[9]Université de Lorraine, Nancy, France

[10]Radboud University, Nijmegen, The Netherlands

[11]Natural England, York, UK

[12]School of Geography/water@leeds, University of Leeds, Leeds, UK

Correspondence
Martin A. Wilkes, Centre for Agroecology, Water and Resilience, Coventry University, Ryton-on-Dunsmore, UK.
Email: martin.wilkes@coventry.ac.uk

Abstract

The growing use of functional traits in ecological research has brought new insights into biodiversity responses to global environmental change. However, further progress depends on overcoming three major challenges involving (a) statistical correlations between traits, (b) phylogenetic constraints on the combination of traits possessed by any single species, and (c) spatial effects on trait structure and trait–environment relationships. Here, we introduce a new framework for quantifying trait correlations, phylogenetic constraints and spatial variability at large scales by combining openly available species' trait, occurrence and phylogenetic data with gridded, high-resolution environmental layers and computational modelling. Our approach is suitable for use among a wide range of taxonomic groups inhabiting terrestrial, marine and freshwater habitats. We demonstrate its application using freshwater macroinvertebrate data from 35 countries in Europe. We identified a subset of available macroinvertebrate traits, corresponding to a life-history model with axes of resistance, resilience and resource use, as relatively unaffected by correlations and phylogenetic constraints. Trait structure responded more consistently to environmental variation than taxonomic structure, regardless of location. A re-analysis of existing data on macroinvertebrate communities of European alpine streams supported this conclusion, and demonstrated that occurrence-based functional diversity indices are highly sensitive to the traits included in their calculation. Overall, our findings suggest that the search for quantitative trait–environment relationships using single traits or simple combinations of multiple traits is unlikely to be productive. Instead, there is a need to embrace the value of conceptual frameworks linking community responses to environmental change via traits which correspond to the axes of life-history models. Through a novel integration of tools and databases, our flexible framework can address this need.

KEYWORDS

ecological niche modelling, functional traits, life history, macroecology, phylogenetics, spatial ecology, trait-based ecology

1 | INTRODUCTION

Trait-based ecology uses the phenotypic characteristics of organisms to study biodiversity responses to environmental change. We define 'functional traits' as characteristics strictly inherent to the organism, requiring no reference to external conditions (Violle et al., 2007), which contribute to fitness (Cadotte & Tucker, 2017). Via their expression within species assemblages, functional traits indirectly control ecosystem functioning (Moore & Olden, 2017; Wilkes et al., 2019). Two potential advantages of trait-based ecology over traditional taxonomy-based approaches are improved mechanistic understanding of species–environment relationships and greater generality of those relationships over large (e.g. continental-global) extents (Verberk, Van Noordwijk, & Hildrew, 2013). There is a long history of ecological research on the role of traits in predicting shifts in community composition along environmental gradients (Grime, 1977; Poff, 1997; Southwood, 1977; Townsend & Hildrew, 1994; Winemiller, Fitzgerald, Bower, & Pianka, 2015). Several initiatives have collated trait information for different groups of organisms, and multiple traits are routinely measured on individual organisms or referenced from databases after taxonomic identification (BirdLife International, 2019; FishBase Consortium, 2018; Kattge et al., 2011; Schmidt-Kloiber & Hering, 2015; Vieira et al., 2006).

Applications of trait-based ecology have driven important new insights into biodiversity responses to global environmental change (Brown et al., 2018; Díaz et al., 2016; Newbold et al., 2012; Stuart-Smith et al., 2013). However, to make further progress, there are at least three major challenges that need to be overcome when working at the largest scales. These challenges involve (a) statistical correlations between trait, (b) phylogenetic constraints on the combination of traits possessed by any single species, and (c) spatial effects on trait structure (occurrence probability- or abundance-weighted means of traits in a community) and trait–environment relationships (statistical links between trait structure and environmental variables).

1.1 | Statistical correlations and phylogenetic constraints

The vast majority of trait-based studies have searched for single trait–environment relationships, with few results supporting a priori predictions (Hamilton et al., 2019; Peres-Neto, Dray, & Braak, 2017). There is now ample evidence to suggest that such negative results are due to two related issues affecting trait independence. First, statistical correlations between traits (challenge 1) arise because a combination of traits is more adaptive in a given environment (e.g. multivoltinism and small body size) and, therefore, species possessing one trait are more likely to also possess the other (Resh et al., 1988; Usseglio-Polatera, Bournaud, Richoux, & Tachet, 2000). Second, trait non-independence arises due to phylogenetic constraints (challenge 2) whereby closely related species possess a similar combination of traits because of shared ancestry (de Bello et al., 2017; Verberk et al., 2013). Traits do not evolve independently in response to a single selection pressure but as specific combinations of traits linked through evolutionary processes such as trade-offs (investment in one trait reduces the resources available for another) and spin-offs (investment in one trait increases fitness advantages conferred by another trait; Verberk et al., 2013). Consequently, a causal mechanism by which a trait appears to influence a species' persistence may be related to another, correlated trait. Spurious causal relationships may seriously confound trait-based applications (Moor et al., 2017; Pilière et al., 2016; Poff et al., 2006; Webb, Hoeting, Ames, Pyne, & Poff, 2010; Weemstra et al., 2016). This problem of disentangling cause and correlation can be circumvented by delineating relatively independent sets of traits that respond to key environmental gradients (Verberk et al., 2013). In generalized life-history models, these axes relate to resource availability and resistance and resilience to stress and disturbance (Van Looy et al., 2019; Winemiller et al., 2015).

1.2 | Spatial effects on trait structure and trait–environment relationships

A major purported benefit of using a trait-based approach in large-scale ecology is that it offers additional information over taxonomy-based analyses because trait structure is less confounded by biogeographical processes limiting species' distributions (Dolédec, Statzner, & Bournaud, 1999; Menezes, Baird, & Soares, 2010). Ideally, trait structure and trait–environment relationships would be consistent across whole continents, or even globally, enabling the development of generalized predictive frameworks (challenge 3). Implicit in the definition of this ideal scenario is a general use of the term 'spatial', which includes both the spatialized environment and 'pure' spatial effects (Clappe, Dray, & Peres-Neto, 2018). Hence, under the current paradigm, if communities are represented using traits, responses to a given environmental change would be similar in any location, regardless of the spatial structures underlying species' distributions.

To take an example from the freshwater realm, previous global (Brown et al., 2018), continental (Blanck & Lamouroux, 2006; Statzner, Bis, Dolédec, & Usseglio-Polatera, 2001; Statzner, Dolédec, & Hugueny, 2004) and river basin scale (Dolédec et al., 1999; Heino, Schmera, & Erős, 2013) meta-analyses have shown that while trait–environment relationships are generally consistent across those scales, systematic spatial effects on trait structure are often clearly evident. For instance, working in alpine streams, Brown et al. (2018) reported a significant contraction of functional richness and a shift in trait structure towards taxa with smaller body sizes and shorter life cycles, as well as shifts in diets, at higher latitudes. In contrast, Statzner et al. (2001) reported low variability in trait structure and trait response to disturbance among a collection of datasets from various stream types in Europe. As yet, however, there have been no explicit, spatially continuous assessments of the variability of trait structure at continental scales, nor of how this spatial variability can confound the results of trait-based studies. Our ability to deliver such a comprehensive analysis has traditionally been hampered by a lack of data at the relevant scales.

1.3 | Integrating open data sources to test assumptions of trait-based ecology

There are now >1 billion species' occurrence records in the Global Biodiversity Information Facility (GBIF), and >1.7 million sequences publicly available in the Barcode of Life Database. The Open Tree of Life now has >2.6 million tips in its synthetic phylogenetic tree. High-resolution climate (e.g. WorldClim) and elevation (e.g. SRTM) data are openly available and readily integrated into large-scale statistical models (Fick & Hijmans, 2017). This creates new opportunities for trait-based research at large scales that have yet to be fully realized (Culina, Crowther, Ramakers, Gienapp, & Visser, 2018; Violle, Reich, Pacala, Enquist, & Kattge, 2014).

We stipulated three basic criteria for trait independence: (a) there should be minimal statistical correlations between traits; (b) traits should be minimally constrained by phylogeny; and (c) trait structure, and its response to environmental change, should be minimally variable with respect to spatial coordinates at the scale considered. These criteria correspond to the three challenges introduced above. Due to the confounding effects of trait correlations and phylogenetic constraints acting on the full set of trait data, we hypothesized that trait–environment relationships would be more strongly evident when trait categories violating these criteria were excluded from the analysis (Van Looy et al., 2019).

By combining openly available environmental data and species' occurrence, trait and phylogenetic records with computational modelling, we establish a new, generalized analytical framework for quantifying trait correlations, phylogenetic constraints and spatial variability at large scales. We demonstrate its application with a case study on freshwater macroinvertebrates and test the implications of our findings for trait-based applications by re-analysing published macroinvertebrate community data from stream sites spanning the major alpine regions of Europe. Subsequently, we discuss present capabilities and recommend future directions in trait-based ecology.

2 | MATERIALS AND METHODS

2.1 | Framework development

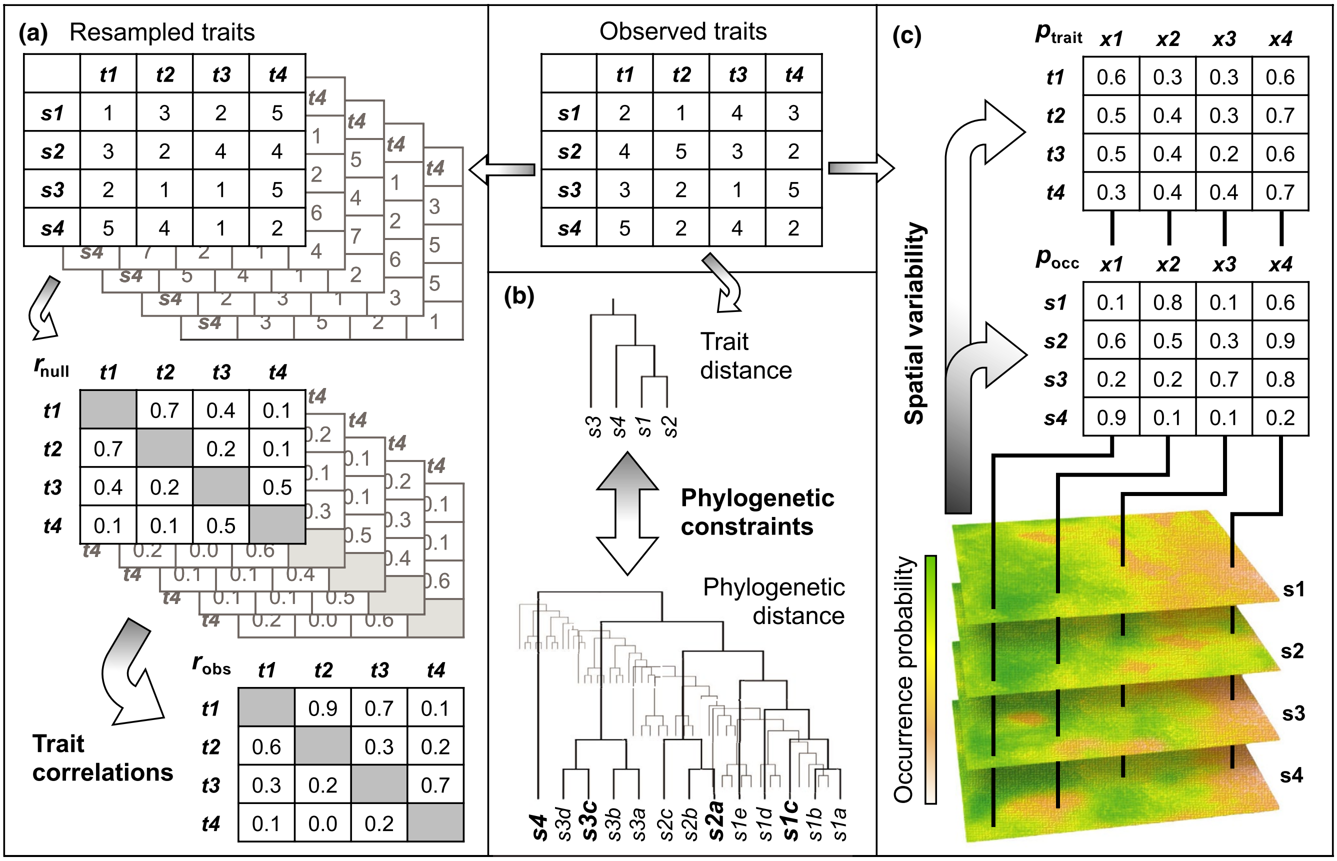
Our framework applies three separate analyses corresponding to the criteria stipulated above (Figure 1). The steps involved in our analyses are directly applicable to fuzzy coded traits assigned at mixed taxonomic levels but may be adapted for application to any trait types and taxonomic resolutions. Trait correlations are assessed by resampling the observed species × traits matrix a large number of times, respecting the rules of the fuzzy scoring system used to quantify traits (Figure 1a). This is necessary because the rules used to assign fuzzy scores can vary between trait categories and taxa, potentially introducing correlations which are artefacts of the scoring system. For each sample of the species × traits matrix, a correlation matrix is produced, generating null distributions of correlation coefficients (Spearman's ρ) for each pair of traits (r_{null}). Observed pairwise correlations (r_{obs}) are then ranked among the null distributions to derive two-tailed p values, Bonferroni adjusted for multiple pairwise comparisons. For applications involving exclusively binary or continuously measured traits, this step may be simplified to a standard correlation test. In cases where binary or continuously measured traits are mixed with fuzzy coded traits then our framework may be applied directly.

To enable phylogenetic constraints to be quantified, a phylogenetic tree is constructed with tips corresponding to species within the taxa included in the trait database at mixed taxonomic levels (Figure 1b). A trait distance matrix is produced from the observed traits and the constraint quantified as the cumulative correlation between traits and phylogenetic distance, compared to a null hypothesis of no phylogenetic autocorrelation. Separate trait distance matrices may be produced to quantify phylogenetic constraints for different groups of traits. The analysis is repeated by sampling species-level branches in the phylogenetic tree a large number of times (represented by grey phylogenetic trees in the background of Figure 1b) to quantify the sensitivity of results to the mixed taxonomic resolution of the trait database. For instance, in the simplified example presented in Figure 1b, taxon *s2* in the trait database corresponds to three possible tips (species) in the phylogenetic tree (*s2a*, *s2b*, *s2c*), one of which would be selected in each sample. For studies using traits assigned at the species level, sampling from the phylogenetic tree is not necessary.

Spatial variability is assessed by first fitting species distribution models for each taxon in the trait database, then randomly sampling a large number of grid cells (x) within the study area to generate predicted probability occurrences (p_{occ}) in each sampled cell (Figure 1c). Community (occurrence probability)-weighted means of each trait (p_{trait}) are then calculated and gradient analysis used to quantify the turnover of taxonomic and trait structure in space. This procedure is repeated a large number of times to assess the sensitivity of results to spatial sampling. If sufficient abundance data exist at the requisite scale, abundance-based species distribution models (including joint species distribution models) and abundance-weighted mean traits can be used. Further description of the steps involved in applying the framework is given below in the context of our case study.

2.2 | Study area

For the purposes of demonstrating our framework, we defined our study area as the extent of European Union Member States, excluding the Outermost Regions held by France, Spain and Portugal (European Union, 2019). To avoid large gaps within this geopolitical region that may have confounded the analyses, we also included



Resampled traits

	t1	t2	t3	t4
s1	1	3	2	5
s2	3	2	4	4
s3	2	1	1	5
s4	5	4	1	2

Observed traits

	t1	t2	t3	t4
s1	2	1	4	3
s2	4	5	3	2
s3	3	2	1	5
s4	5	2	4	2

r_{null}	t1	t2	t3	t4
t1		0.7	0.4	0.1
t2	0.7		0.2	0.1
t3	0.4	0.2		0.5
t4	0.1	0.1	0.5	

r_{obs}	t1	t2	t3	t4
t1		0.9	0.7	0.1
t2	0.6		0.3	0.2
t3	0.3	0.2		0.7
t4	0.1	0.0	0.2	

p_{trait}	x1	x2	x3	x4
t1	0.6	0.3	0.3	0.6
t2	0.5	0.4	0.3	0.7
t3	0.5	0.4	0.2	0.6
t4	0.3	0.4	0.4	0.7

p_{occ}	x1	x2	x3	x4
s1	0.1	0.8	0.1	0.6
s2	0.6	0.5	0.3	0.9
s3	0.2	0.2	0.7	0.8
s4	0.9	0.1	0.1	0.2

FIGURE 1 Overview of the framework, using a hypothetical example of four taxa (*s*) and four fuzzy coded trait modalities (*t*) assigned at mixed taxonomic resolutions. (a) Trait correlations are assessed by resampling the observed $s \times t$ matrix a large number of times, respecting the rules of the fuzzy scoring system used to assign traits (Figure S1), to produce null distributions of correlation coefficients for each pair of traits (r_{null}). Observed pairwise correlations (r_{obs}) are then ranked among the null distributions to derive *p* values. (b) Phylogenetic constraints are quantified by constructing a phylogenetic tree with tips corresponding to species represented in the trait database. A trait distance matrix is produced from the observed traits and the constraint quantified as the cumulative correlation between traits and phylogenetic distance. The analysis is repeated by sampling species-level branches (e.g. *s2a*, *s2b*, *s2c*) in the phylogenetic tree a large number of times to quantify the sensitivity of results to the mixed taxonomic resolution of the trait database. (c) Spatial variability is assessed by first fitting species distribution models for each genus represented in the trait database, then randomly sampling a large number of grid cells (*x*) within the study area to generate predicted probability occurrences (p_{occ}) in each sampled cell ($x = 4$ shown here for demonstration purposes). Community (occurrence probability)-weighted means of each trait (p_{trait}) are then calculated and gradient analysis used to quantify the turnover of taxonomic and trait structure in space. The analysis may be simplified where binary or continuously measured traits are available at the species level

the United Kingdom, Norway, Switzerland, Albania, Bosnia and Herzegovina, Montenegro, Macedonia and Serbia. In total, the study area extended to 5,110,076 km^2.

2.3 | Model organism group

In Europe, trait-based freshwater ecology is set to play an increasingly important role in biomonitoring (Reyjol et al., 2014). All macroscopic (macro)invertebrates, including insects, crustaceans, molluscs and other major groups are considered within the scope of monitoring. This role for trait-based ecology using macroinvertebrates is supported by well-established trait databases (Schmidt-Kloiber & Hering, 2015). The most comprehensive of these databases, compiled by Tachet, Richoux, Bournaud, and Usseglio-Polatera (2010), contains a set of 63 functional trait modalities in 11 categories, covering 484 taxa classified at mixed taxonomic levels and coded using a fuzzy scoring system (Chevenet, Dolédec, & Chessel, 1994; Table S1). For clarity, we use the term 'trait category' (e.g. food) to refer to a group of 'trait modalities' (e.g. detritus). Despite the popularity of this reference database, to date there has been no comprehensive assessment of the extent of statistical correlations, phylogenetic constraints and spatial variability among the traits it contains. However, applications of classification and ordination techniques have previously indicated partial phylogenetic constraints among the traits and taxa included in the database (Usseglio-Polatera et al., 2000; Usseglio-Polatera, Richoux, Bournaud, & Tachet, 2001).

Previous work in North America found that many of the available traits for freshwater insects were highly collinear and tightly linked to phylogeny among the 311 taxa considered (Poff et al., 2006). A set of traits describing feeding mode, dispersal mode, body size and voltinism (number of generations per year) were identified as relatively unaffected, or 'labile' (i.e. more readily altered through evolutionary processes, independently of other traits). Others have suggested that strong interrelationships among some macroinvertebrate traits mean that traits should not be analysed as independent variables but rather as linked sets or 'syndromes' (Verberk et al., 2008a, 2008b, 2013). To date, however, there is no consensus on the extent of non-independence among macroinvertebrate traits, despite a decade of debate (Menezes et al., 2010; Statzner & Bêche, 2010).

Based on the full set of data published in Tachet et al. (2010), traits encompassed categories of body size (maximum body length), lifespan, voltinism, aquatic stages (eggs, larva, pupa, adult), reproduction mode, dispersal mode, resistance forms, respiration mode, locomotion, food and feeding mode (Table S1). After removing taxa with incomplete trait information, a total of 443 taxa in 23 orders and 152 families remained. Of these, most had traits assigned at the genus (52%) or species (39%) level. The remainder (mostly Diptera) were assigned at the tribe, subfamily or family levels (9%). This trait database was used as the basis for three sets of analyses designed to assess trait correlations (Table 1), phylogenetic constraints (Table 2) and spatial variability (Table 3). More details on the steps involved are found in the corresponding tables and the Methods S1. The analyses were conducted in R version 4.0.0 (R Core Team, 2020) unless otherwise specified. Key R scripts for estimating trait correlations, phylogenetic constraints and spatial gradients in taxonomic and trait structure are available in the macroTraits GitHub repository (https://github.com/wilkesma/macroTraits).

TABLE 1 Steps involved in quantifying trait correlations

Step	Description	Source data	Method (package)
a1	Assess the fuzzy scoring system used to quantify traits	Tachet et al. (2010)	Quantify minimum and maximum possible fuzzy scores for trait modalities within each trait category (see Figure S1). See function *fuzzy_trait_correlations* in GitHub repository[a]
a2	Generate an ensemble of possible species × trait matrices	Step a1	Resample the species × trait matrix 1×10^6 times, respecting the rules of the fuzzy scoring system. See function *fuzzy_trait_correlations* in GitHub repository[a]
a3	Create null distributions of correlation coefficients	Step a2	Calculate Spearman's ρ between each pair of trait modalities in the ensemble of possible species × trait matrices
a4	Assess the significance of observed correlations	Tachet et al. (2010); Step a1	Rank the observed correlations among the null distributions to obtain two-tailed, Bonferroni adjusted p values

[a] https://github.com/wilkesma/macroTraits.

TABLE 2 Steps involved in quantifying phylogenetic constraints

Step	Description	Source data	Method (package)
b1	Obtain genetic data	Barcode of Life Database (BOLD)	Search BOLD for cytochrome *c* oxidase subunit I (COI) sequences using taxa names from the trait database (bold v0.8.6 in R)
b2	Clean genetic data	Step b1	Retain the longest sequences with least missing base pairs for each species
b3	Align sequences	Step b2	Perform multiple sequence alignment using Clustal Omega (msa v1.14.0 in R) followed by removal of gaps and poorly aligned sequences using Gblocks (ips v0.0-7 in R)
b4	Retrieve synthetic phylogenetic tree	Open Tree of Life (OTL)	Search OTL using species names from the multiple sequence alignment (rotl v3.0.7 in R)
b5	Select nucleotide substitution model	Steps b3–b4	Subset the multiple sequence alignment to retain only those species represented in the synthetic tree. Select the nucleotide substitution model using jModelTest (phangorn v2.4.0 in R), providing the settings for branch length estimation (step b6)
b6	Prepare final phylogenetic tree	Steps b3–b5	Estimate branch lengths using aligned sequences of species represented in the synthetic tree (MrBayes v3.2.6). Settings are generated using the selected nucleotide substitution model (step b5) and by constraining tree topology using the synthetic phylogenetic tree (paleotree v3.1.3 in R)
b7	Quantify phylogenetic constraints	Trait database; Step b6	Calculate the cumulative correlation between phylogenetic distance and trait distance for each trait category (phylosignal v1.2.1 in R). See function *phylo_constraints* in GitHub repository[a]

[a] https://github.com/wilkesma/macroTraits.

TABLE 3 Steps involved in quantifying spatial variability

Step	Description	Source data	Method (package)
c1	Prepare environmental data	WorldClim (BIO1, BIO4, BIO12, BIO15); Shuttle Radar Topography Mission (elevation); SoilGrids (soil pH)	Download climate and elevation data at 10-arc-minute resolution (getData function, raster v2.8-19 in R). Compute slope from elevation data using a 3 × 3 cell moving window. Download the 'PHIHOX' variable (soil pH) manually from SoilGrids (https://www.isric.org/explore/soilgrids). Crop layers to the study area and aggregate to a common resolution of 2.5 km (aggregate function, raster v2.8-19 in R)
c2	Download species' occurrence data	Global Biodiversity Information Facility (GBIF)	Search GBIF using generic names for taxa assigned at species or genus level in the trait database, retaining only georeferenced records within the study area with coordinate uncertainties ≤ 2.5 km (gbif function, dismo v1.1-4 in R)
c3	Augment GBIF data	*Réseau de Contrôle de Surveillance* (RCS)	Fill gaps in GBIF data coverage using national or regional datasets (e.g. RCS monitoring data from France), retaining only genus and species-level occurrence records
c4	Fit species distribution models	Steps c1–c3	Given the environmental layers, use MaxEnt to predict the occurrence probability in every 2.5 km^2 grid cell for each genus represented in the augmented GBIF data (enmtools.maxent function, ENMTools v0.2 in R)
c5	Sample spatial gradients	Step c4	Generate an ensemble of 1,000 random samples of 1,000 grid cells (without replacement) from within the study area. Retrieve the geographical coordinates, elevation and predicted occurrence probabilities for every genus in each sampled grid cell. See function *spatial_gradient* in GitHub repository[a]
c6	Calculate community-weighted means	Step c5	For each sampled grid cell in the ensemble, multiply the predicted occurrence probability by the corresponding trait score for each trait modality and genus. See function *spatial_gradient* in GitHub repository[a]
c7	Quantify spatial variability (turnover)	Steps c4–c6	For each of the 1,000 spatial gradient samples, fit gradient forest models to occurrence probabilities and community-weighted mean traits (gradientForest v0.1-17 in R) to quantify the turnover of taxonomic and trait structure along spatial gradients (northing, easting, elevation). See function *spatial_gradient* in GitHub repository[a]

[a] https://github.com/wilkesma/macroTraits.

2.4 | Re-analysis of alpine macroinvertebrate community data

To test our hypothesis that trait correlations and phylogenetic constraints confound trait-based analyses, we re-analysed macroinvertebrate community (abundance) data from the European subset of alpine sites studied by Brown et al. (2018). The data were drawn from 261 alpine stream sites located across the Scandinavia Mountains, the Pyrenees and the European Alps, with varying proportions (0%–80%) of glacial ice cover in their upstream catchments. The original paper used a more focused trait database which was harmonized across alpine regions globally, whereas the comparative analyses presented here are based upon the European database of Tachet et al. (2010). For all analyses, we compared the results obtained using all 11 trait categories to those obtained after excluding trait categories that most strongly violated our criteria for trait independence. We calculated commonly used functional diversity (FD) indices for each site (Villéger, Mason, & Mouillot, 2008) and estimated the importance of deterministic community assembly processes (dispersal- and niche-based) along the gradient of glacier cover using the same approach as the original paper (Brown et al., 2018). We also computed turnover (mean Bray–Curtis dissimilarity) in taxonomic and trait structure within and between regions. Finally, we tested for relationships between linear combinations of traits and glacial ice cover using the fourth corner (individual trait–glacial ice cover relationships) and RLQ (overall trait–glacial ice cover relationship) methods (Dray et al., 2014). It should be noted that, while the rest of the analyses reported here are based upon trait occurrence probability, this analysis uses abundance-weighted trait structure. This difference is justified due to the lack of abundance data at the continental scale and the fact that we do not make direct comparisons between inferences generated using occurrence probability- and abundance-weighted data.

3 | RESULTS AND DISCUSSION

3.1 | Trait correlations

The maximum possible number of unique trait combinations, expressed as the product of unique combinations of resampled scores within each trait category (Figure S4), was >2 × 10^{36}. Thus, if all traits were truly independent, we would expect all 443 taxa to exhibit their own unique trait profile, yet only 408 unique trait combinations existed in the trait database. This finding indicates trait non-independence, as previously found for stream insects in North America (Poff et al., 2006). Statistically significant correlations between trait modalities were widespread (Figure 2). Negative

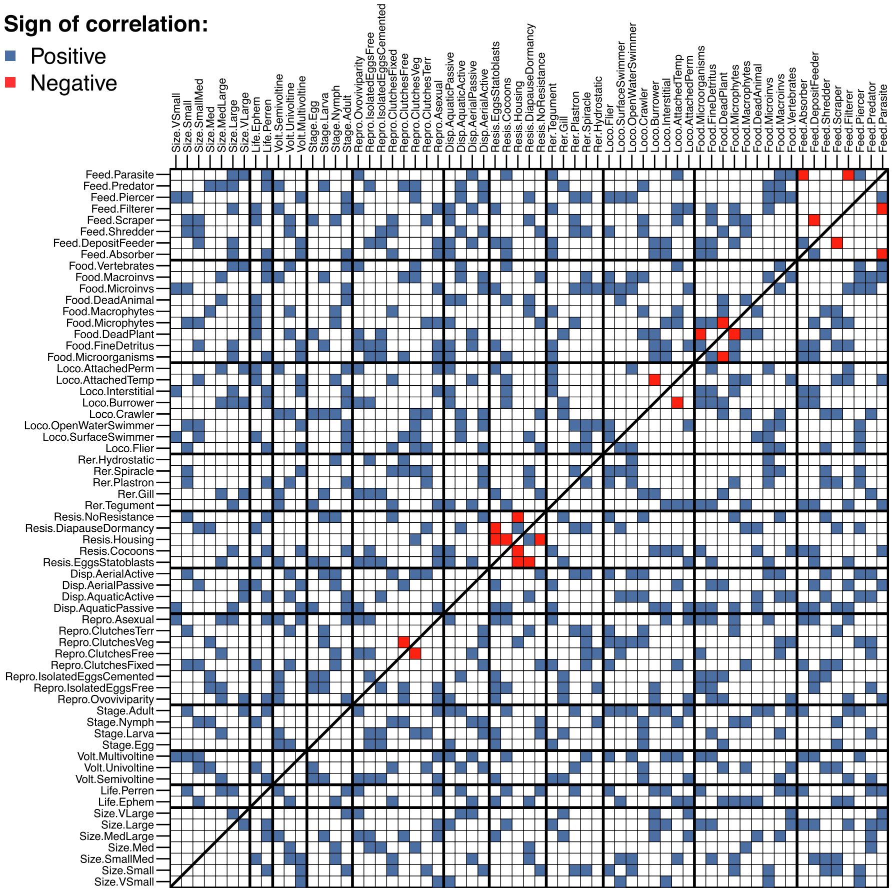


FIGURE 2 Statistically significant (Bonferroni adjusted $p < .05$) correlations between trait modalities, with colour representing the sign of the correlation. See Table S1 for full names of trait modalities and trait categories

correlations were only found between trait modalities within the same trait categories, whereas correlations between different trait categories were always positive. This is because the fuzzy scoring system typically assigns each taxon nonzero values in a limited number of trait modalities within each trait category, damping any signal from negative trait correlations.

Trait modalities for food (describing diet) and feeding mode (how an organism obtains food) were highly intercorrelated, as expected (Cummins & Klug, 1979). The mean number of significant correlations ($p_{adj} < .05$) per trait modality within each trait category was highest for lifespan (37 correlations), dispersal mode (34) and voltinism (32; Figure S5). Many of these highly intercorrelated traits were linked in intuitive combinations whereby, for example, multivoltine taxa also have ephemeral lives (≤1 year), are present in the aquatic environment at juvenile life stages and disperse passively. These same taxa are also more likely to have smaller bodies. However, body size, along with resistance forms, was the least correlated with other traits (24 correlations each), and also minimally correlated with one another.

3.2 | Phylogenetic constraints

All trait categories were significantly constrained by phylogeny but to varying degrees (Figure 3). Aquatic stages was the most severely constrained category, exhibiting strong and significant correlations up to phylogenetic distances of >1.5 (Figure 3d). These extreme distances correspond approximately to the class level (Figure S6), reflecting the fundamental difference between insects with non-aquatic life stages and obligate aquatic classes, as well as the presence or absence of a pupal (aquatic) stage distinguishing between holometabolous and hemimetabolous insect orders. Reproduction mode was positively correlated with phylogenetic distance for both closely and distantly related species (Figure 3e). This indicates convergent evolution of reproductive strategies among lineages, particularly the occurrence of ovoviviparity among diverse Annelida, Porifera, Bryozoa, Bivalvia, Gastropoda, Crustacea and Insecta taxa. Body size, voltinism and resistance forms were the least constrained trait categories (Figure 3a,c,g) with significant correlations persisting up to phylogenetic distances of <0.6 (corresponding to congeneric taxa; Figure S6). Body size and voltinism even vary within a species across latitudinal clines (Bonada & Dolédec, 2018; Horne, Hirst, & Atkinson, 2015). Because of their phylogenetic independence, these trait categories could offer most additional information over taxonomy-based analyses. Food was marginally less constrained by phylogeny than feeding mode, supporting the notion that macroinvertebrates obtain a more generalist diet than expected based on mouthpart morphology and feeding behaviour alone (Dangles, 2002; Tomanova, Goitia, & Helešic, 2006).

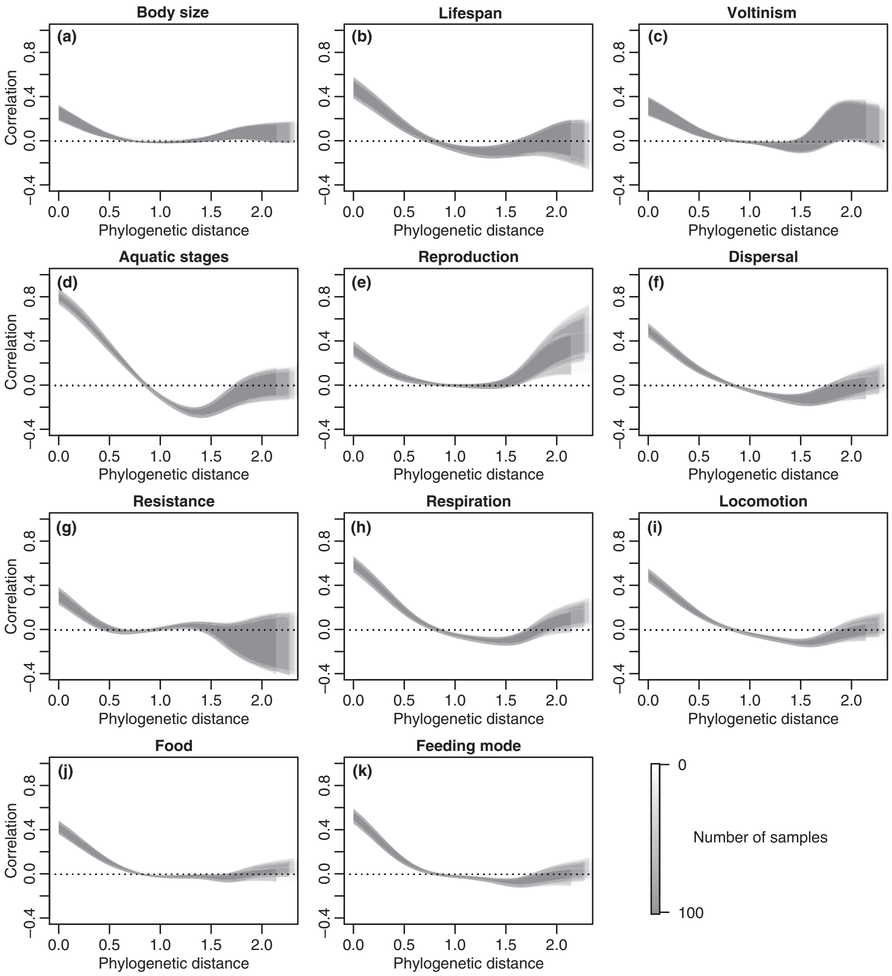


FIGURE 3 Phylogenetic constraints on traits. Cumulative correlation of trait modalities within 11 trait categories (a–k) with phylogenetic distance (sum of branch lengths) sampled from the species-level phylogenetic tree (Figure S2). Polygons shaded to reflect the distribution of 100 samples in data space. Dashed lines indicate the expected value of Moran's *I* under the null hypothesis of no autocorrelation. See Figure S6 for taxonomic interpretation of phylogenetic distances

3.3 | Spatial variability

Taxonomic and trait structure were approximately equal in terms of their total variability (maximum cumulative importance) along spatial gradients (Figure 4c,e,g). However, trait turnover along the latitudinal gradient ('northing' in the reprojection) was much steeper than taxonomic turnover, with a clearer delineation of Mediterranean regions. Across the continent, the environmental correlates of elevation were associated with a more consistent effect on trait structure regardless of location than was the case with taxonomic structure. This can be seen by comparing maps of taxonomic (Figure 4a) and trait (Figure 4b) turnover. With trait turnover the major mountain regions consistently occupy a similar part of ordination space (green colours in Figure 4b) regardless of location (e.g. compare the Alps and Scandinavian Mountains). This tendency for similar trait structure among communities from anywhere within Europe indicates that trait responses to environmental variation are less dependent on location than taxonomic responses. Northing was the single most important gradient (Figure 4c) followed by easting (Figure 4e). These findings reflect large-scale climatic drivers of macroinvertebrate community assembly acting on both taxonomic and trait structure (Brown et al., 2018). The turnover of individual traits along spatial gradients varied strongly by trait modality but no single trait category stood out as particularly invariant (Figure 4d,f,h).

3.4 | Implications for trait-based ecology

While our findings support the notion that trait non-independence (statistical correlations and phylogenetic constraints) is widespread, not all traits were equally affected. Trait categories describing body size, resistance forms and, to a lesser extent, food were the most labile, a similar result to previous work on stream insects in North America (Poff et al., 2006). These three trait categories correspond to the typical axes of life-history models, namely resilience (smaller bodies are associated with r-selected species), resistance to disturbance (resistance forms) and resource utilization (food), respectively (e.g. Townsend & Hildrew, 1994; Van Looy et al., 2019). Although these same traits were clearly associated with systematic spatial variation, the gradient forest analysis suggested that patterns in trait structure are more consistently related to strong environmental gradients, such as those associated with elevation, when compared to patterns in taxonomic structure. Low spatial variability in trait structure was evident for major mountain regions. Thus, large-scale trait-based analyses across distant mountain regions are less

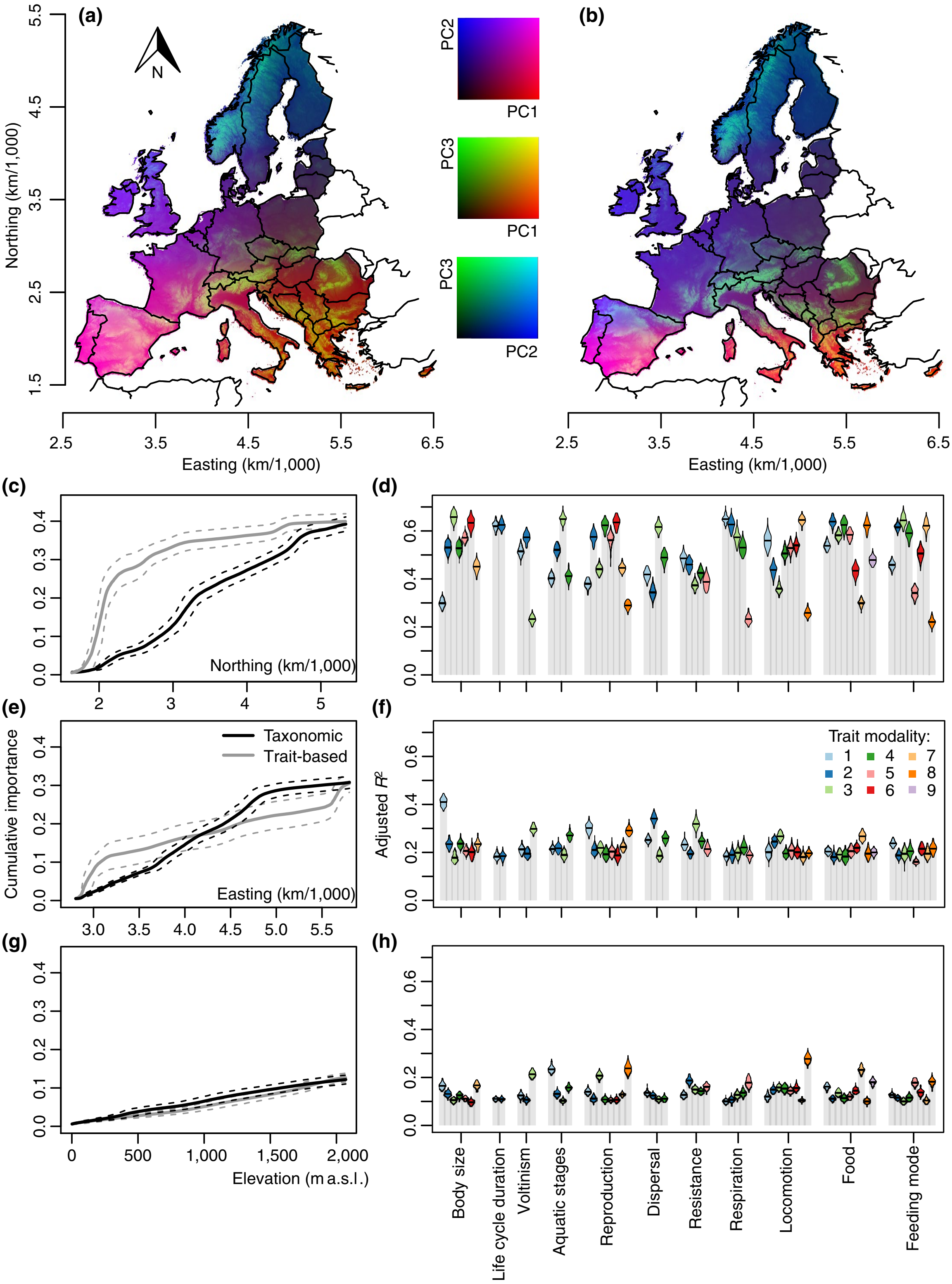


FIGURE 4 Spatial instability of taxonomic and trait occurrence probability structure. Summary of spatial turnover in taxonomic (a) and trait (b) composition based on predicted occurrence probabilities from multiple species distribution models, with colour bands scaled to axis scores from a principal components (PC) analysis of transformed data from gradient forest models including northings, eastings and elevation as predictors. As colour changes from one point on the map to another, the predicted (a) taxonomic or (b) trait composition changes proportionally. (c, e, g) Turnover functions (cumulative importance) from gradient forest models predicting taxonomic and trait structure. Dashed lines indicate the 95% confidence intervals from sampling of grid cells in (a) and (b). (d, f, h) Bean plots showing the mean (black horizontal line) and distribution (coloured by trait modality) of the importance of individual trait modalities within 11 trait categories in the gradient forest models. See Table S1 for names of trait modalities

confounded by spatial structure than taxonomy-based analyses, but they may still be affected by trait non-independence. We therefore re-examined the alpine dataset from Brown et al. (2018) to test the hypothesis that trait non-independence confounds trait-based applications at large scales. We compared trait-based indicators calculated using the full set of traits to those calculated using only body size, resistance forms and food.

Functional richness (related to the number of traits represented in the community) and functional evenness (related to both the number of traits and the abundance distribution) were highly sensitive to the set of traits used (Figure 5a,b). In contrast, there was a close relationship between functional dispersion (related to the abundance distribution only) values calculated using the two alternative sets of traits (Figure 5c), showing that purely abundance-based FD indices are robust to variation in the traits included. The trait space occupied by macroinvertebrate communities using each set of traits was strikingly similar among the three regions (Figure 5d–f). Analysis of turnover within and between regions provided clear evidence for the benefits of working with traits at large scales, with drastically lower mean Bray–Curtis dissimilarities for trait structure compared to taxonomic structure (Figure 5g). There was little difference in trait turnover between the scenarios including all traits and labile traits only, although labile trait turnover was less in all cases (Figure 5g). Using the labile subset of traits did not result in a stronger deterministic community assembly response to increasing glacier cover (Figure 5h,i), a gradient of habitat harshness associated with decreasing channel stability, water temperature and organic matter (Jacobsen & Dangles, 2012). This is because the community assembly model is based on reduced dimensionality trait space (Brown et al., 2018), meaning that inclusion of additional, highly intercorrelated traits had a negligible impact on the result, and may have led to overfitting (Figure 5h). Including region and/or stream catchment as random effects failed to improve the fit of generalized additive models of deterministic community assembly processes relative to the global model (Table S2). This indicates the presence of a consistent trait–environment relationship across the major European alpine regions (Table S3). However, the fourth corner (Table S4) and RLQ (Table S5) analyses reported no significant trait–glacial ice cover relationships, re-emphasizing the need to go beyond the search for trait–environment relationships using single traits or linear combinations of multiple traits (Hamilton et al., 2019; Peres-Neto et al., 2017). Overall, our re-analysis of these data suggests that trait-based ecologists should think carefully about which traits to include in large-scale analyses, especially when occurrence-based FD indices are of interest.

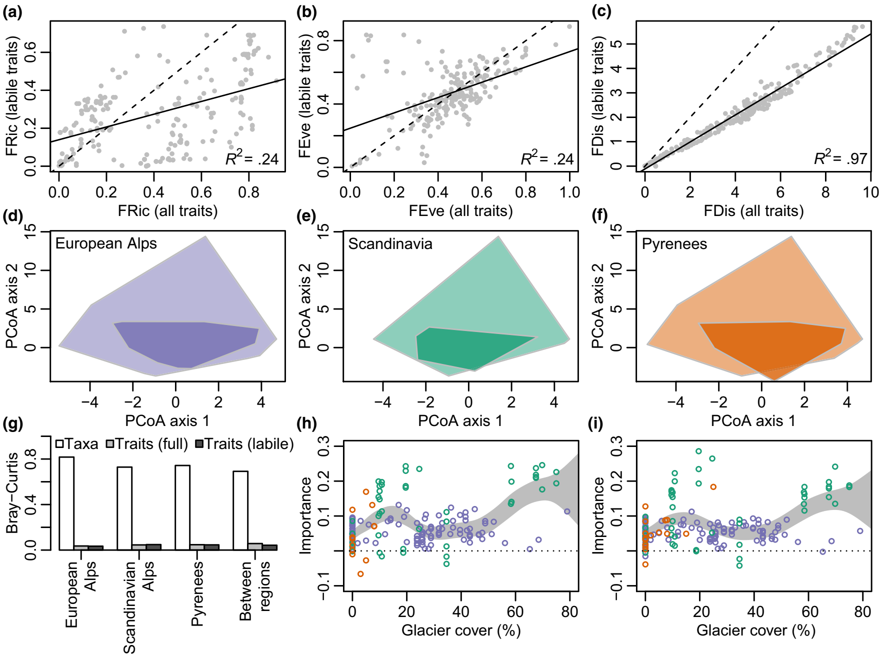


FIGURE 5 Sensitivity of trait-based analyses on macroinvertebrate community data from alpine regions of Europe using two alternative sets of traits: all 11 available trait categories and three relatively labile trait categories. Comparison of (a) functional richness (FRic), (b) functional evenness (FEve) and (c) functional dispersion (FDis). Dashed lines in (a, b, c) show 1:1 relationships between FD indices calculated using each set of traits, whereas the solid lines indicate the best fit of general linear models with goodness-of-fit noted within each panel. First two axes of a principal coordinates analysis (PCoA) showing the trait space occupied by macroinvertebrate taxa observed in (d) the European Alps, (e) the Scandinavian Mountains and (f) the Pyrenees. Lighter and darker coloured polygons in (d–f) represent convex hulls calculated using all trait categories and the relatively labile trait categories, respectively. Comparison of turnover (mean Bray–Curtis dissimilarity) in taxonomic and trait structure within and between regions (g). Note that the between regions scenario in (g) is based on mean abundances of taxa within each region. The importance of deterministic community assembly processes estimated using (h) all trait categories and (i) labile trait categories only. Shaded areas in (h, i) indicate 95% confidence intervals around the mean fit of generalized additive models. Importance values greater than zero indicate a significant role for deterministic community assembly processes (combined effect of dispersal- and niche-based processes)

4 | CONCLUSIONS

We have shown how three major challenges (statistical independence, phylogenetic independence and spatial variability) in large-scale trait-based ecology can be better understood using openly available ecological, phylogenetic and environmental data. In the case of freshwater macroinvertebrates, traits were strongly intercorrelated and constrained by phylogeny, although certain traits were less affected (body size, resistance forms, food). Our findings support the applicability of a life-history model for this diverse group comprising axes (and corresponding traits) of resilience (body size), resistance (resistance forms) and resource utilization (food). These traits map directly onto a generalized model for river ecosystems recently proposed by Van Looy et al. (2019). However, our findings suggest that while these traits may be useful to ecologists working at continental scales, the search for trait–environment relationships through simple combinations of trait modalities and environmental variables is unlikely to be productive. Instead, ecologists should embrace the heuristic value of trait-based conceptual models for understanding how communities respond to environmental change.

We have harnessed advances in data availability and computational analyses to establish a novel, generalized framework for large-scale trait-based ecology. The framework can be applied to any taxonomic group and habitat type, although the specific approach to assessing statistical correlations, phylogenetic constraints and spatial variability should be adjusted to reflect the type of trait data available (continuous, binary or fuzzy; taxonomic resolution), as well as the availability and coverage of phylogenetic and occurrence or abundance data (Figure 1). While the framework provides quantitative information on the performance of each trait category, background knowledge specific to the taxonomic group and habitats of interest must be used judiciously in future applications. For example, in addition to results gained by applying the framework, reference to existing life-history models for the group of interest will be useful to guide the selection of traits.

Progress towards a universal framework for making robust predictions of ecological responses to environmental change across major habitat types and taxonomic groups depends on identifying traits directly related to niche dimensions (Winemiller et al., 2015). The identification of labile traits, such as that we have demonstrated here, may lead to better indicators of community resilience and resistance to disturbances (Van Looy et al., 2019). Such important work will improve our understanding of the role of traits in controlling ecological stability following disturbance (Donohue et al., 2013; Pennekamp et al., 2018; Radchuk et al., 2019), across time (Yang, Fowler, Jackson, & Donohue, 2019) and space (Zelnik, Arnoldi, & Loreau, 2019). Improved knowledge of these controls will be central to our ability to forecast future ecosystem dysfunction, and therefore inform efforts to prevent, mitigate and adapt to global environmental change.

DATA AVAILABILITY STATEMENT

The R scripts implemented in this study are archived at https://github.com/wilkesma/macroTraits. The existing datasets that support the findings of the case study can be obtained from the original sources cited in the manuscript.

ORCID

Martin A. Wilkes https://orcid.org/0000-0002-2377-3124
Wilco C. E. P. Verberk https://orcid.org/0000-0002-0691-583X
Lee E. Brown https://orcid.org/0000-0002-2420-0088

REFERENCES

BirdLife International. (2019). *BirdLife data zone*. Retrieved from http://datazone.birdlife.org

Blanck, A., & Lamouroux, N. (2006). Large-scale intraspecific variation in life-history traits of European freshwater fish. *Journal of Biogeography*, *34*(5), 862–875. https://doi.org/10.1111/j.1365-2699.2006.01654.x

Bonada, N., & Dolédec, S. (2018). Does the Tachet trait database report voltinism variability of aquatic insects between Mediterranean and Scandinavian regions? *Aquatic Sciences*, *80*(1), 7. https://doi.org/10.1007/s00027-017-0554-z

Brown, L. E., Khamis, K., Wilkes, M., Blaen, P., Brittain, J. E., Carrivick, J. L., … Milner, A. M. (2018). Functional diversity and community assembly of river invertebrates show globally consistent responses to decreasing glacier cover. *Nature Ecology & Evolution*, *2*(2), 325–333. https://doi.org/10.1038/s41559-017-0426-x

Cadotte, M. W., & Tucker, C. M. (2017). Should environmental filtering be abandoned? *Trends in Ecology & Evolution*, *32*(6), 429–437. https://doi.org/10.1016/J.TREE.2017.03.004

Chevenet, F., Dolédec, S., & Chessel, D. (1994). A fuzzy coding approach for the analysis of long-term ecological data. *Freshwater Biology*, *31*(3), 295–309. https://doi.org/10.1111/j.1365-2427.1994.tb01742.x

Clappe, S., Dray, S., & Peres-Neto, P. R. (2018). Beyond neutrality: Disentangling the effects of species sorting and spurious correlations in community analysis. *Ecology*, *99*(8), 1737–1747. https://doi.org/10.1002/ecy.2376

Culina, A., Crowther, T. W., Ramakers, J. J. C., Gienapp, P., & Visser, M. E. (2018). How to do meta-analysis of open datasets. *Nature Ecology & Evolution*, *2*(7), 1053–1056. https://doi.org/10.1038/s41559-018-0579-2

Cummins, K. W., & Klug, M. J. (1979). Feeding ecology of stream invertebrates. *Annual Review of Ecology and Systematics*, *10*(1), 147–172. https://doi.org/10.1146/annurev.es.10.110179.001051

Dangles, O. (2002). Functional plasticity of benthic macroinvertebrates: Implications for trophic dynamics in acid streams. *Canadian Journal of Fisheries and Aquatic Sciences*, *59*(9), 1563–1573. https://doi.org/10.1139/f02-122

de Bello, F., Šmilauer, P., Diniz-Filho, J. A. F., Carmona, C. P., Lososová, Z., Herben, T., & Götzenberger, L. (2017). Decoupling phylogenetic and functional diversity to reveal hidden signals in community assembly. *Methods in Ecology and Evolution*, *8*(10), 1200–1211. https://doi.org/10.1111/2041-210X.12735

Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., Dray, S., … Gorné, L. D. (2016). The global spectrum of plant form and function. *Nature*, *529*(7585), 167–171. https://doi.org/10.1038/nature16489

Dolédec, S., Statzner, B., & Bournaud, M. (1999). Species traits for future biomonitoring across ecoregions: Patterns along a human-impacted river. *Freshwater Biology*, *42*(4), 737–758. https://doi.org/10.1046/j.1365-2427.1999.00509.x

Donohue, I., Petchey, O. L., Montoya, J. M., Jackson, A. L., McNally, L., Viana, M., … Emmerson, M. C. (2013). On the dimensionality of ecological stability. *Ecology Letters*, *16*(4), 421–429. https://doi.org/10.1111/ele.12086

Dray, S., Choler, P., Dolédec, S., Peres-Neto, P. R., Thuiller, W., Pavoine, S., & ter Braak, C. J. F. (2014). Combining the fourth-corner and the

RLQ methods for assessing trait responses to environmental variation. *Ecology*, *95*(1), 14–21. https://doi.org/10.1890/13-0196.1

European Union. (2019). *Countries*. Retrieved from https://europa.eu/european-union/about-eu/countries_en

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

FishBase Consortium. (2018). *FishBase*. Retrieved from https://www.fishbase.org

Grime, J. P. (1977). Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *The American Naturalist*, *111*(982), 1169–1194. https://doi.org/10.1086/283244

Hamilton, A. T., Schäfer, R. B., Pyne, M. I., Chessman, B., Kakouei, K., Boersma, K. S., … Stamp, J. (2019). Limitations of trait-based approaches for stressor assessment: The case of freshwater invertebrates and climate drivers. *Global Change Biology*, *26*(2), 364–379. https://doi.org/10.1111/gcb.14846

Heino, J., Schmera, D., & Erős, T. (2013). A macroecological perspective of trait patterns in stream communities. *Freshwater Biology*, *58*(8), 1539–1555. https://doi.org/10.1111/fwb.12164

Horne, C. R., Hirst, A. G., & Atkinson, D. (2015). Temperature-size responses match latitudinal-size clines in arthropods, revealing critical differences between aquatic and terrestrial species. *Ecology Letters*, *18*(4), 327–335. https://doi.org/10.1111/ele.12413

Jacobsen, D., & Dangles, O. (2012). Environmental harshness and global richness patterns in glacier-fed streams. *Global Ecology and Biogeography*, *21*(6), 647–656. https://doi.org/10.1111/j.1466-8238.2011.00699.x

Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., … Wirth, C. (2011). TRY – A global database of plant traits. *Global Change Biology*, *17*(9), 2905–2935. https://doi.org/10.1111/j.1365-2486.2011.02451.x

Menezes, S., Baird, D. J., & Soares, A. M. V. M. (2010). Beyond taxonomy: A review of macroinvertebrate trait-based community descriptors as tools for freshwater biomonitoring. *Journal of Applied Ecology*, *47*(4), 711–719. https://doi.org/10.1111/j.1365-2664.2010.01819.x

Moor, H., Rydin, H., Hylander, K., Nilsson, M. B., Lindborg, R., & Norberg, J. (2017). Towards a trait-based ecology of wetland vegetation. *Journal of Ecology*, *105*(6), 1623–1635. https://doi.org/10.1111/1365-2745.12734

Moore, J. W., & Olden, J. D. (2017). Response diversity, nonnative species, and disassembly rules buffer freshwater ecosystem processes from anthropogenic change. *Global Change Biology*, *23*(5), 1871–1880. https://doi.org/10.1111/gcb.13536

Newbold, T., Scharlemann, J. P. W., Butchart, S. H. M., Sekercioglu, C. H., Alkemade, R., Booth, H., & Purves, D. W. (2012). Ecological traits affect the response of tropical forest bird species to land-use intensity. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1750), 20122131. https://doi.org/10.1098/rspb.2012.2131

Pennekamp, F., Pontarp, M., Tabi, A., Altermatt, F., Alther, R., Choffat, Y., … Petchey, O. L. (2018). Biodiversity increases and decreases ecosystem stability. *Nature*, *563*(7729), 109–112. https://doi.org/10.1038/s41586-018-0627-8

Peres-Neto, P. R., Dray, S., & te Braak, C. J. F. (2017). Linking trait variation to the environment: Critical issues with community-weighted mean correlation resolved by the fourth-corner approach. *Ecography*, *40*(7), 806–816. https://doi.org/10.1111/ecog.02302

Pilière, A. F. H., Verberk, W. C. E. P., Gräwe, M., Breure, A. M., Dyer, S. D., Posthuma, L., … Schipper, A. M. (2016). On the importance of trait interrelationships for understanding environmental responses of stream macroinvertebrates. *Freshwater Biology*, *61*(2), 181–194. https://doi.org/10.1111/fwb.12690

Poff, N. L. (1997). Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society*, *16*(2), 391–409. https://doi.org/10.2307/1468026

Poff, N. L., Olden, J. D., Vieira, N. K. M., Finn, D. S., Simmons, M. P., & Kondratieff, B. C. (2006). Functional trait niches of North American lotic insects: Traits-based ecological applications in light of phylogenetic relationships. *Journal of the North American Benthological Society*, *25*(4), 730–755. https://doi.org/10.1899/0887-3593(2006)025[0730:FTNONA]2.0.CO;2

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Radchuk, V., Laender, F. D., Cabral, J. S., Boulangeat, I., Crawford, M., Bohn, F., … Kramer-Schadt, S. (2019). The dimensionality of stability depends on disturbance type. *Ecology Letters*, *22*(4), 674–684. https://doi.org/10.1111/ele.13226

Resh, V. H., Brown, A. V., Covich, A. P., Gurtz, M. E., Li, H. W., Minshall, G. W., … Wissmar, R. C. (1988). The role of disturbance in stream ecology. *Journal of the North American Benthological Society*, *7*(4), 433–455. https://doi.org/10.2307/1467300

Reyjol, Y., Argillier, C., Bonne, W., Borja, A., Buijse, A. D., Cardoso, A. C., … van de Bund, W. (2014). Assessing the ecological status in the context of the European Water Framework Directive: Where do we go now? *Science of the Total Environment*, *497–498*, 332–344. https://doi.org/10.1016/j.scitotenv.2014.07.119

Schmidt-Kloiber, A., & Hering, D. (2015). www.freshwaterecology.info – An online tool that unifies, standardises and codifies more than 20,000 European freshwater organisms and their ecological preferences. *Ecological Indicators*, *53*, 271–282. https://doi.org/10.1016/J.ECOLIND.2015.02.007

Southwood, T. R. E. (1977). Habitat, the templet for ecological strategies? *The Journal of Animal Ecology*, *46*(2), 337–365. https://doi.org/10.2307/3817

Statzner, B., & Bêche, L. A. (2010). Can biological invertebrate traits resolve effects of multiple stressors on running water ecosystems? *Freshwater Biology*, *55*, 80–119. https://doi.org/10.1111/j.1365-2427.2009.02369.x

Statzner, B., Bis, B., Dolédec, S., & Usseglio-Polatera, P. (2001). Perspectives for biomonitoring at large spatial scales: A unified measure for the functional composition of invertebrate communities in European running waters. *Basic and Applied Ecology*, *2*(1), 73–85. https://doi.org/10.1078/1439-1791-00039

Statzner, B., Dolédec, S., & Hugueny, B. (2004). Biological trait composition of European stream invertebrate communities: Assessing the effects of various trait filter types. *Ecography*, *27*(4), 470–488. https://doi.org/10.1111/j.0906-7590.2004.03836.x

Stuart-Smith, R. D., Bates, A. E., Lefcheck, J. S., Duffy, J. E., Baker, S. C., Thomson, R. J., … Edgar, G. J. (2013). Integrating abundance and functional traits reveals new global hotspots of fish diversity. *Nature*, *501*(7468), 539–542. https://doi.org/10.1038/nature12529

Tachet, H., Richoux, P., Bournaud, M., & Usseglio-Polatera, P. (2010). *Invertébrés d'eau douce: Systématique, biologie, écologie* (2nd ed.). Paris, France: Centre National de la Recherche Scientifique Press.

Tomanova, S., Goitia, E., & Helešic, J. (2006). Trophic levels and functional feeding groups of macroinvertebrates in neotropical streams. *Hydrobiologia*, *556*(1), 251–264. https://doi.org/10.1007/s10750-005-1255-5

Townsend, C. R., & Hildrew, A. G. (1994). Species traits in relation to a habitat templet for river systems. *Freshwater Biology*, *31*(3), 265–275. https://doi.org/10.1111/j.1365-2427.1994.tb01740.x

Usseglio-Polatera, P., Bournaud, M., Richoux, P., & Tachet, H. (2000). Biological and ecological traits of benthic freshwater macroinvertebrates: Relationships and definition of groups with similar traits. *Freshwater Biology*, *43*(2), 175–205. https://doi.org/10.1046/j.1365-2427.2000.00535.x

Usseglio-Polatera, P., Richoux, P., Bournaud, M., & Tachet, H. (2001). A functional classification of benthic macroinvertebrates based on biological and ecological traits: Application to river condition assessment and stream management. *Archiv Für Hydrobiologie. Supplementband*, *139*(1), 53–83.

Van Looy, K., Tonkin, J. D., Floury, M., Leigh, C., Soininen, J., Larsen, S., … Wolter, C. (2019). The three Rs of river ecosystem resilience: Resources, recruitment, and refugia. *River Research and Applications*, *35*(2), 107–120. https://doi.org/10.1002/rra.3396

Verberk, W. C. E. P., Siepel, H., & Esselink, H. (2008a). Applying life-history strategies for freshwater macroinvertebrates to lentic waters. *Freshwater Biology*, *53*(9), 1739–1753. https://doi.org/10.1111/j.1365-2427.2008.02036.x

Verberk, W. C. E. P., Siepel, H., & Esselink, H. (2008b). Life-history strategies in freshwater macroinvertebrates. *Freshwater Biology*, *53*(9), 1722–1738. https://doi.org/10.1111/j.1365-2427.2008.02035.x

Verberk, W. C. E. P., Van Noordwijk, C. G. E., & Hildrew, A. G. (2013). Delivering on a promise: Integrating species traits to transform descriptive community ecology into a predictive science. *Freshwater Science*, *32*(2), 531–547. https://doi.org/10.1899/12-092.1

Vieira, N. K. M., Poff, N. L., Carlisle, D. M., Moulton II, S. R., Koski, M. L., & Kondratieff, B. C. (2006). *A database of lotic invertebrate traits for North America*. Retrieved from https://pubs.usgs.gov/ds/ds187/

Villéger, S., Mason, N. W. H., & Mouillot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, *89*(8), 2290–2301. https://doi.org/10.1890/07-1206.1

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional! *Oikos*, *116*(5), 882–892. https://doi.org/10.1111/j.0030-1299.2007.15559.x

Violle, C., Reich, P. B., Pacala, S. W., Enquist, B. J., & Kattge, J. (2014). The emergence and promise of functional biogeography. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(38), 13690–13696. https://doi.org/10.1073/pnas.1415442111

Webb, C. T., Hoeting, J. A., Ames, G. M., Pyne, M. I., & Poff, N. L. (2010). A structured and dynamic framework to advance traits-based theory and prediction in ecology. *Ecology Letters*, *13*(3), 267–283. https://doi.org/10.1111/j.1461-0248.2010.01444.x

Weemstra, M., Mommer, L., Visser, E. J. W., van Ruijven, J., Kuyper, T. W., Mohren, G. M. J., & Sterck, F. J. (2016). Towards a multidimensional root trait framework: A tree root review. *New Phytologist*, *211*(4), 1159–1169. https://doi.org/10.1111/nph.14003

Wilkes, M. A., Gittins, J. R., Mathers, K. L., Mason, R., Casas-Mulet, R., Vanzo, D., … Jones, J. I. (2019). Physical and biological controls on fine sediment transport and storage in rivers. *Wiley Interdisciplinary Reviews: Water*, *6*(2), e1331. https://doi.org/10.1002/wat2.1331

Winemiller, K. O., Fitzgerald, D. B., Bower, L. M., & Pianka, E. R. (2015). Functional traits, convergent evolution, and periodic tables of niches. *Ecology Letters*, *18*(8), 737–751. https://doi.org/10.1111/ele.12462

Yang, Q., Fowler, M. S., Jackson, A. L., & Donohue, I. (2019). The predictability of ecological stability in a noisy world. *Nature Ecology & Evolution*, *3*(2), 251–259. https://doi.org/10.1038/s41559-018-0794-x

Zelnik, Y. R., Arnoldi, J., & Loreau, M. (2019). The three regimes of spatial recovery. *Ecology*, *100*(2), e02586. https://doi.org/10.1002/ecy.2586

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Wilkes MA, Edwards F, Jones JI, et al. Trait-based ecology at large scales: Assessing functional trait correlations, phylogenetic constraints and spatial variability using open data. *Glob Change Biol*. 2020;26:7255–7267. https://doi.org/10.1111/gcb.15344