Freshwater Biology

Freshwater Biology (2016) **61,** 646–657 doi:10.1111/fwb.12732

Detecting signatures of competition from observational data: a combined approach using DNA barcoding, diversity partitioning and checkerboards at small spatial scales

TREVOR T. BRINGLOE*[†], SARAH J. ADAMOWICZ*[†], VIVIAN F. I. HARVEY*[†], JOHN K. JACKSON[‡] AND KARL COTTENIE*

*Department of Integrative Biology, University of Guelph, Guelph, ON, Canada
[†]Biodiversity Institute of Ontario, University of Guelph, Guelph, ON, Canada
[‡]Stroud Water Research Center, Avondale, PA, U.S.A.

SUMMARY

1 Competitive interactions are assumed to play a role in governing species distributions but are difficult to infer using observational data. In addition, morphological identification may overlook species, obscuring patterns of competitive exclusion. To address these limitations, we showcase a molecular (DNA barcoding) approach to species-unit delineation and sample stream insect communities at small spatial scales where environmental filtering and dispersal limitations are minimised.
2 Rocks from riffles were sampled for stream insects, in particular caddisfly (Trichoptera), beetle (Coleoptera) and blackfly (Simuliidae) larvae from filterer and grazer guilds at White Clay Creek, Pennsylvania, U.S.A. (20 June, 2013). Additive partitioning of species diversity was used to establish at which spatial scale(s) variation in the distribution of species occurs; this informed us of the spatial scale at which to conduct checkerboard analyses, which tested for signatures of species segregations (e.g. competitive exclusion).
3 Only a small portion of the total species diversity occurred at the smallest spatial level (mean = 30% on individual rocks). Distributional variation in species was greatest at the kilometre scale (where beta-diversity was 37% and 41% for filterers and grazers respectively). At the scale of White Clay Creek sampled (*c.* 3 km extent), species of filterers showed a strong tendency to segregate on individual rocks, while grazers showed random structure approaching aggregation.
4 This study demonstrates a potential approach to assessing the role of competition in structuring communities using observational data and highlights the importance of accurate species units for analysis (e.g. DNA barcoding). Using a combination of analyses, we were able to link patterns of segregation to competitive interactions among stream insect species in a filterer guild, while determining these interactions were not important in a grazer guild.

Keywords: community structure, competition, molecular operational taxonomic units, spatial scale, stream insects

Introduction

The principle of competitive exclusion predicts that when two sympatric species populations occupy the same ecological niche, one will inevitably drive the other to local extinction (Gause, 1932; Hardin, 1960). Despite the logic and general interest behind competitive exclusion, modern ecology continues to struggle with attributing observational patterns in species distributions to competition, given that multiple ecological and evolutionary processes can yield the same distributional patterns (Ulrich, 2004; Mayfield & Levine, 2010; Kunstler *et al.*, 2012; Warren *et al.*, 2014). One commonly used test for competition, checkerboard analysis, infers the

Correspondence: Trevor Bringloe, Biology Department, University of New Brunswick, PO Box 4400, Fredericton, NB E3B 5A3, Canada.
E-mail: trevor.bringloe@unb.ca

 © 2016 John Wiley & Sons Ltd

importance of competition in species assemblages based on the prediction that certain combinations of species are 'forbidden' in a site due to interspecific competition (Diamond, 1975; Gotelli & McCabe, 2002). If species tend to segregate (as per Diamond, 1975), this segregation should be reflected as a mutual tendency for two species not to co-occur, or a 'checkerboard pattern', in a site-by-species matrix when sampling multiple sites (Stone & Roberts, 1990). Diamond's (1975) community assembly rule, however, has been contrasted with other explanations for species segregations such as random chance (Connor & Simberloff, 1979; Ulrich, 2004), neutral processes (Bell, 2001, 2005) and habitat heterogeneity (Schoener & Adler, 1991). While researchers recognise checkerboards can be used to detect non-random assemblages of species (Gotelli & McCabe, 2002), the interpretation of non-random C-scores remains compromised by the possibility of species segregating due to environmental filtering and dispersal limitations; consequently, uncertainty remains in the interpretation of checkerboards (Boschilia, Oliveira & Thomaz, 2008; Kamilar & Ledogar, 2011). We therefore propose a combined approach to reduce the potential confounding effects of environmental filtering due to habitat heterogeneity and dispersal limitation when analysing observational data. Because the influence of both these processes increases with spatial scale, sampling over a small spatial extent will reduce the likelihood of detecting their confounding effects (Willis & Whittaker, 2002; Cavender-Bares *et al.*, 2009). With these carefully considered sampling designs, environmental and dispersal factors will become relatively weak in driving species distributions compared to local competitive interactions, which facilitates the interpretation of the checkerboard results. The scale at which competitive interactions govern the assemblage of species will depend on the system (i.e. habitat, taxa) under investigation.

To determine the optimal 'small spatial scale' for non-random distributions of species, we propose using a diversity components analysis as a starting point. Additive diversity partitioning breaks down regional community richness (γ) into constituent local diversity (α) and distributional heterogeneity coefficient (β) components ($\gamma = \alpha + \beta$); additive β is defined as the amount of diversity (typically number of species) by which the regional data set (γ) exceeds the average amount of diversity in a single sampling unit (α) (Tuomisto, 2010). The classical additive approach to measuring β diversity, although not without limitations (Veech & Crist, 2010; Wilsey, 2010), has undergone a recent surge in use with the introduction of null models (Veech *et al.*, 2002; Crist *et al.*, 2003; Veech, 2005). Here, we employ the intuitive and accessible concept of additive diversity partitioning to test for structure in the distribution of species diversity at a small spatial scale (e.g. structure in residual species diversity after environmental filtering and dispersal limitation mechanisms have taken effect). While diversity components are often used to assess variation in the distribution of species at regional scales, such localised assessments are far less common. Of the few studies that additively partition diversity over a very localised spatial extent (<5 km), β diversity is often greater than expected by chance, meaning diversity remains non-randomly distributed even at small spatial extents (De Troch *et al.*, 2008; Gheerardyn *et al.*, 2010; Ligeiro, Melo & Callisto, 2010; Negro, Rolando & Palestrini, 2011; Rodríguez-Zaragoza *et al.*, 2011). If species diversity remains structured at small local spatial extents where regional environmental variation and dispersal limitations are minimised, then we propose here that the potential for detecting biotic competition interactions becomes a viable hypothesis, which can be tested observationally with checkerboard analysis.

As checkerboard analysis is based on species presence and absence, it is crucial to start with correct species delineation. In the above outlined approach, we advocate to sample environmentally similar sites at small spatial scales. If, however, cryptic species exist in these situations, false co-occurrences will potentially swamp the competitive signal. Biologists are increasingly aware of severe limitations in our understanding of how many species occupy this earth and our ability to detect them (Hebert, Ratnasingham & deWaard, 2003b; Hebert *et al.*, 2003a; Smith *et al.*, 2008; Hamilton *et al.*, 2010; Mora *et al.*, 2011). DNA sequence data tend to provide a more accurate measurement of existing species diversity than does morphology for many invertebrate taxa, due to the prevalence of undescribed species (Smith *et al.*, 2008), sexually dimorphic organisms (Blagoev *et al.*, 2013), evolutionarily distinct yet morphologically cryptic species (Hebert *et al.*, 2004; Witt, Threloff & Hebert, 2006; Zhou *et al.*, 2010; Jackson *et al.*, 2014) and difficulty of identifying juvenile and/or damaged individuals (Sweeney *et al.*, 2011; Jackson *et al.*, 2014). While there are a multitude of molecular operational taxonomic unit (MOTU) delineation approaches available (e.g. Jones, Ghoorah & Blaxter, 2011; Puillandre *et al.*, 2012; Fujisawa & Barraclough, 2013; Ratnasingham & Hebert, 2013), many studies seek to establish a % divergence threshold among species, a value that may vary depending on the taxa of interest. DNA barcoding, in particular, can be used to assign specimens to known species as well as

for rapid delimitation of species-like MOTUs based on DNA sequence data, using a marker that is standardised across most animal taxa, a portion of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene. DNA barcoding may use a fluid divergence threshold or a variety of MOTU delineation methods, which are typically tuned for a given taxonomic group in order to be as close as possible to either classical taxonomic units or evolutionarily distinct species. As such, DNA barcoding is becoming increasingly used in ecological studies to increase the frequency and accuracy of species-level identifications (Valentini, Pompanon & Taberlet, 2008; Pfrender *et al.*, 2010) and has been used in studies of diversity components (Bringloe, 2014), checkerboard analyses (Pfenninger *et al.*, 2007), food webs (Kaartinen *et al.*, 2010), patterns in species distributions (Gill *et al.*, 2014) and diet studies (Clare *et al.*, 2011). Here, we further showcase the utility of DNA barcoding for ecological studies, particularly in the context of diversity partitioning and checkerboard analyses.

We focused our study on two guilds within stream systems, filterers and grazers, specifically targeting larval caddisflies (Trichoptera), beetles (larval and adult; Coleoptera) and larval blackflies (Simuliidae). Filterers are organisms that sieve moving water for particulates (including animals), while grazers feed on organic material such as algae on the surface of rocks. We were particularly interested in investigating the filterer guild, given extensive reports and experimental studies demonstrating strong competitive interactions among Hydropsychidae and subordinate Simuliidae (Jansson & Vuoristo, 1979; Hemphill & Cooper, 1983; Thorp, 1983; Hemphill, 1988, 1991; Georgian & Thorp, 1992). In addition, Martin, Adamowicz & Cottenie (2016), studying the same taxonomic groups in similar habitat as investigated here (wadeable temperate streams), demonstrated that spatial distance and environmental variation play a non-significant or minor role in structuring communities at the scale of the stream reach across a relatively large spatial extent (<15% community variation explained across *c.* 100 km). Providing consistent microhabitats, rocks from riffles were sampled across a small spatial extent (scale of *c.* 3 km), thus we are confident spatial and environmental effects are negligible in explaining the non-random distribution patterns detected. Note, however, that finding limited environmental variation within the stream is contingent upon the parameters measured; while we include parameters important for biomonitoring programs and which are known to exert an influence upon stream communities, additional parameters (e.g. extremely local-scale flow) could play a role. See supporting information, appendix S1 (including Tables S1, S2 and Figure S1) for additional support on this important assumption.

Our research question has two components: (i) can known competitive interactions among filterer species (e.g. Hydropsychidae and Simuliidae) be detected using a purely observational approach at White Clay Creek? And (ii) if our observational approach is accurate, what is the relative role of competitive interactions in structuring grazers at White Clay Creek? Our first objective was to use additive diversity partitioning to determine (i) if stream insect species are non-randomly distributed, and if so (ii) at what spatial scale(s) does important species turnover occur while maintaining environmental consistency? The scale at which this variation in species distributions occurs (metres versus kilometres) would then be used as the spatial extent for checkerboard analyses testing for signatures of taxon-specific interactions. Our second objective was to test specifically for signatures of competition in explaining the non-random community structure detected using additive partitioning of species diversity. We hypothesise that members of the same guild compete with one another at White Clay Creek, possibly for food and space resources; if so, we will detect significantly large checkerboard scores within these guilds.

Methods

Sampling of White Clay Creek

In order to assess stream insect communities for signatures of competition, rocks were sampled throughout the White Clay Creek catchment in Chester County, Pennsylvania (U.S.A., 20 June 2013; Fig. 1). Five stations (e.g. within-stream sites) were sampled spanning two arms of White Clay Creek (Fig. 1). At each station, three riffles were sampled running perpendicular to the stream flow (n = 15 total riffles). At each riffle, two rocks were randomly sampled (n = 30 total rocks); a rock sample sometimes consisted of two to three rocks if the selected rock was part of a tight cluster (given Hydropsychidae build their net case in crevices; max rock size = 197 cm^2, min rock size = 55 cm^2, mean ± SE rock size = 104.9 ± 6.6 cm^2). Sampling consisted of gently moving rocks from the stream into a water-filled bucket where the rock(s) were hand scrubbed to remove insects. Once cleaned and inspected, each rock was wrapped in tinfoil before being returned to the stream; the weight of the tinfoil needed to cover the rock(s) was converted into a measurement of the area sampled

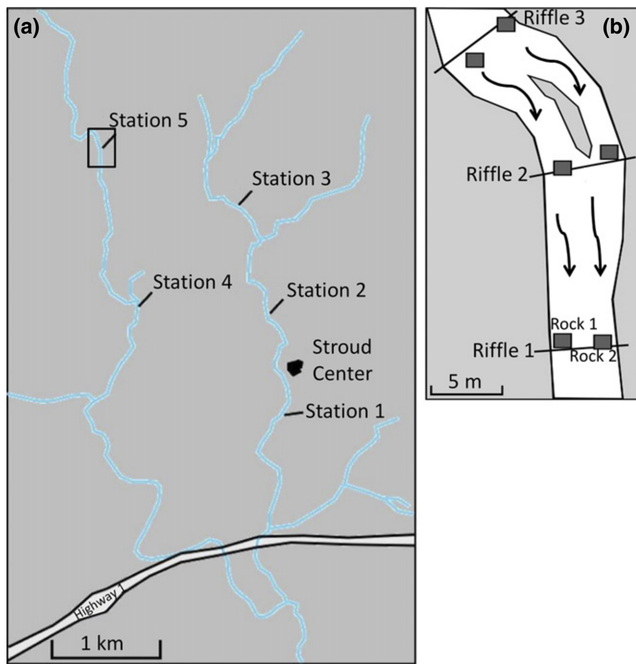

Fig. 1 Field sampling design implemented at White Clay Creek, PA, U.S.A. (20 June, 2013). (a) White Clay Creek catchment depicting the stations sampled and location of the Stroud Water Research Center. (b) Sampling scheme at a single station where rocks were sampled from several riffles.

[where weight(g)/0.004 = area (cm^2), conversion calibrated in the lab], which was then used in the conversion of species abundances per rock into density per m^2 of surface area of substrate. The density data were needed for the calculation of null distributions in the partition analysis (see below). Insects washed into the bucket were filtered through a 200 μm hand net and moved into a 118–237 mL jar and preserved with 95% ethanol. Samples were stored in a freezer (−20 °C), and ethanol was changed at least once within 24 h. Sample processing consisted of removing all the invertebrates from the bulk sample using a 4 × dissecting microscope and sorting insects according to order or family (for Trichoptera, Coleoptera and Simuliidae) into 20 mL vials (CABIN; McDermott, Paull & Strachan, 2010).

DNA barcoding of specimens

DNA barcoding was used to make species-level delineations of stream insects. When available, 20 individuals/family/rock were randomly selected for DNA barcoding. Specimens were sequenced for the barcode region of the mitochondrial COI gene following protocols established by the Canadian Centre for DNA Barcoding at the Biodiversity Institute of Ontario (BIO), University of Guelph (see DNA barcoding pipeline and Table S3 in supporting information). Specimen information, photographs and density data are available as an online data set through the Barcode of Life Data Systems (BOLD; Ratnasingham & Hebert, 2007; project White Clay Creek Microhabitat Sampling, WCCMS: dx.doi.org/10.5883/DS-WCCMS; GenBank accession numbers KJ449339–KJ450830). Species were defined according to MOTUs based on the sequence data. Delineations were done on the basis of barcode index numbers (BINs), which delineates species units based on a seed threshold of 2.2% divergence, but subsequently refines these delineations based on patterns of intra and interspecific genetic variation (Ratnasingham & Hebert, 2013). Although Linnaean names were not required for our analyses here, species names were assigned to BINs within our data set whenever possible, to contribute to knowledge of species inventories in White Clay Creek, a model site for ecological study. When available, species names were preferentially assigned to BINs using records from a previous study specifically from White Clay Creek (Sweeney *et al*., 2011). If a sampled BIN was not assigned a name through this previous work, then available names on BOLD were used, providing a given BIN met metadata requirements, was assigned to a single species name, and did not share a Linnaean name with another BIN. If multiple BINs were linked to the same species name, then specimens were recorded as the BIN number (*sp*.BOLD#) rather than being assigned a Linnaean name (see Table S4). In some cases, published literature linked to the BIN page was used to confirm the Linnaean name, or a given species belonged to a species complex (in which case, we linked the species name according to the respective complex, that is, *Simulium tuberosum* belongs to the *Simulium tuberosum* complex; see Table S4). Sampling effort was assessed using Chao2 estimation curves (Colwell *et al*., 2012); observed richness was close to mean estimated richness (see supporting information Figure S2).

Morphological IDs were used in a small number of cases (<5% of specimens) where sequencing success was <85% for a given rock sample (based on Merritt, Cummins & Berg, 2008). Density estimates were calculated based on available data if sequencing success was >85%. For cases of excessive failures (≤85% success), failed specimens were identified according to the taxonomic diversity of the higher taxon to which that specimen belonged. Failed specimens belonging to single-species families were morphologically examined to confirm familial membership and recorded as that single species. Low-diversity families with distinguishable single-species genera were assigned to species

according to genus membership. Barcoding was attempted a second time using the same procedure as stated earlier for cases of failures in species-rich families (more than one species/genus); this generally brought sequence success up to acceptable levels (>85%).

Additive diversity components

Additive diversity partitioning was used to assess whether species were non-randomly distributed at White Clay Creek, and, if so, at what spatial scale variation in species distributions occurs. Additive partition analyses were performed on species richness using PARTITION 3.0 (Crist *et al.*, 2003). Total diversity for White Clay Creek was defined as:

$$\text{White Clay Creek insect diversity}(\gamma) = \alpha_{\text{rock}} + \beta_{\text{rock}} + \beta_{\text{riffle}} + \beta_{\text{station}}$$

Partitions were performed for two guilds, the filterers (families Simuliidae, Hydropsychidae, Philopotamidae, and Polycentropodidae) and grazers (families Elmidae, Psephenidae, Glossosomatidae, Georidae, Hydroptilidae and Uenoidae) (Merritt *et al.*, 2008). In order to assess departures from null distributions, expected values of diversity components were determined using individual-based randomisation procedures (simulating diversity components if individual specimens are randomly allocated to each rock). Significance of departures (greater or less) in the observed values was determined using two-tailed *P*-values (>0.975 if less than expected value, <0.025 if greater than expected value). Five thousand iterations were run for each partition in the calculation of the null distributions. Standard errors for observed α component estimations were also calculated from the raw data. The spatial design was balanced most of the time, except for one rock with no recorded species of grazers; this sample was necessarily removed from the analysis given it lacked an index of diversity.

Checkerboard analyses

Partition analysis was used to determine which spatial scale(s) show non-random variation in the distribution of stream insect species, which in turn would be indicative of a sorting mechanism (i.e. this analysis informed at which spatial scale checkerboards should be performed). Checkerboards were calculated using the co-occurrence module in ECOSIM7.0 (Acquired Intelligence Inc., Kelsey-Bear & Pinyon Publishing; Montrose, CO, U.S.A.). The number of checkerboard units for a species pair is calculated as $CU = (r_i - S) \times (r_j - S)$, where r_i and r_j are the number of cases where species i and j occur without the other, and S is the number of sites inhabited by both species (Stone & Roberts, 1990). The C-score is the average of checkerboard units across all species pairs. To test for significant departures from random chance, observed C-scores were compared to the null distribution of C-scores based on random iterations of the input matrix; significance was determined by the two-tail probability that values greater than the average null C-score were due to chance (e.g. 0.025 for significantly greater observed values, 0.975 for significantly lesser values). The C-score was evaluated for species presence/absence matrices for both filterer and grazer guilds separately and for all species sampled for rocks spanning the entirety of White Clay Creek (n = 30; scale of <3 km; station level from partition analyses). The analysis for all species sampled included two predatory species (*Rhyacophila carolina* and *Psychomyia flavida*).

For the null distribution in C-scores, species incidences were randomly shuffled using the sequential swap algorithm into 50 000 iterations of the input matrix, as per Fayle & Manica (2010) recommendation for reducing type I errors. A fixed row and column null model was used in the analysis, as opposed to fully randomised null matrices. When the number of times a species occurred (column totals) is fixed in the analysis, the rarity or commonness of each species is retained in the null matrices; similarly, fixed row totals retain the species richness/site. The fixed model is most appropriate for 'island lists' where the full species list is known for well-defined habitat patches, as in our comprehensive sampling of rocks at White Clay Creek (Gotelli, 2000). Fixed column approaches are also recommended because they are robust to type I errors (Gotelli, 2000). C-scores were standardised for comparisons between analyses, where the standardised C-score = $C_{obs} - C_{sim}$/standard deviation$_{sim}$. Given the C-score is standardised by the number of standard deviations above or below the expected C-score, values >2 or <−2 represent non-random values.

Results

DNA barcoding success

Sequencing success was high for White Clay Creek specimens (1492/1562 = 95.5%); sequencing success was 218/234 for Coleoptera specimens (93.2%; 30 specimens were adults), 320/330 for Simuliidae (97.0%) and 954/998 for Trichoptera (95.6%). In total, DNA barcode data

returned 38 BINs; with the addition of two morphologically identified species of Trichoptera, 40 species were identified. Of these species, 23 were members of the filterer guild, while 15 were grazers and two were predatory. Of the species (BINs) with sequence data, 19/38 (50%) could not be assigned a single Linnaean name (see Methods). Only three cases occurred where two BINs were closely related (i.e. <2% average divergence).

Additive diversity components and checkerboards

α diversity at the smallest scale (rock) was considerably smaller than expected by chance for both guilds, accounting for 35% and 25% of the total diversity for filterers and grazers respectively (Fig. 2, Table 1). β rock was not different from expected values for both guilds, accounting for 9% and 12% in filterers and grazers respectively (Fig. 2, Table 1). β riffle and β station, on the other hand, were always greater than expected by chance; β riffle was 18% and 22%, while β station was 37% and 41% for filterers and grazers respectively (Fig. 2, Table 1).

Because partition analyses indicated that non-random β diversity was greatest at the larger within-stream spatial scale (e.g. β station; km scale), checkerboards were performed for rocks at the full scale of White Clay Creek sampled. A significantly high C-score was detected in the analysis of filterers, indicating a high degree of segregation of species (Fig. 2, Table 2). Grazers, on the other hand, showed random structure, but did approach significant co-occurrence (i.e. aggregation) of species ($P = 0.087$; Fig. 2, Table 2). When all species were analysed, the C-score indicated significant segregation of species, although the standardised value was intermediary between the analyses of filterers and grazers (Table 2).

Discussion

Competition is assumed to play a role in governing species distributions, but detecting signatures of competition remains difficult with observational data. Here, DNA barcoding was used to make accurate species-level delineations of stream insects, addressing concerns that invertebrate diversity is often overlooked with morphological identifications (Hebert *et al.*, 2004; Smith *et al.*, 2008; Bringloe, 2014). Despite the incomplete reference libraries and at times conflicting identifications in public databases (19/38 BINs could not be confidently assigned a Linnaean name, Table S4), DNA

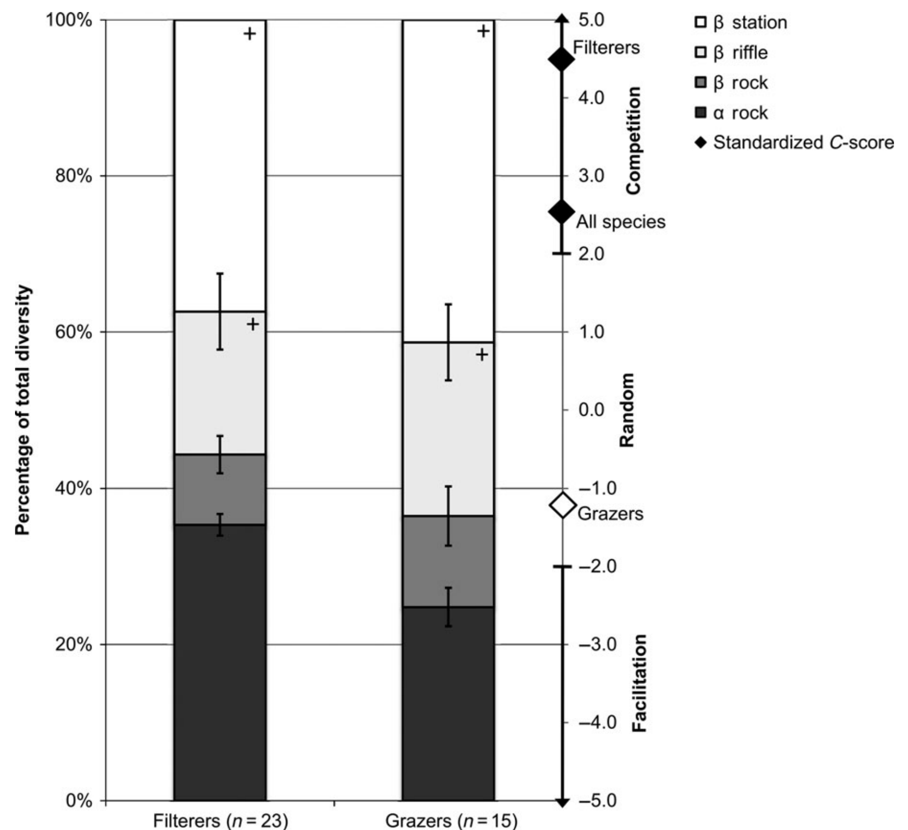

Fig. 2 Additive partitions of sampled filterer and grazer species diversity at White Clay Creek, with accompanying standardised C-scores (PA, U.S.A.; 20 June 2013). Error bars for observed values represent 1 SE and apply to the α component at a given spatial level. For additive partitions, + indicates a beta value greater than expected by chance. For the standardised C-scores, a filled marker indicates a value significantly different than expected by chance (e.g. not random).

Table 1 Additive richness components and checkerboard values for stream filterer and grazer insect species at White Clay Creek (PA, U.S.A.; 20 June 2013). Upper and lower expected values of diversity components were calculated based on 5000 individual-based randomisation iterations of the input data. For diversity partitioning results, *P*-values <0.025 indicate an observed value significantly greater than the expected value; *P*-values >0.975 indicate an observed value significantly smaller than the expected value (significant values are in bold).

	Spatial scale	α or β	Obs	SE	Exp	Lower limit	Upper limit	*P*(obs>exp)
Filterers (*n* = 23)	Rock (*n* = 30)	α	8.13	0.32	19.80	19.10	20.37	**>0.999**
		β	2.07		2.00	1.33	2.70	0.323
	Riffle (*n* = 15)	α	10.20	0.55	21.79	21.07	22.53	**>0.999**
		β	4.20		1.15	0.40	1.87	**<0.001**
	Station (*n* = 5)	α	14.40	1.12	22.94	22.20	23.00	**>0.999**
		β	8.60		0.06	0.00	0.80	**<0.001**
Grazers (*n* = 15)	Rock (*n* = 30)	α	3.72	0.37	10.78	10.07	11.38	**>0.999**
		β	1.75		1.80	1.14	2.46	0.579
	Riffle (*n* = 15)	α	5.47	0.57	12.58	11.80	13.47	**>0.999**
		β	3.33		2.09	1.27	3.07	**<0.001**
	Station (*n* = 5)	α	8.80	0.73	14.67	13.80	15.00	**>0.999**
		β	6.20		0.33	0.00	1.20	**<0.001**

Table 2 Checkerboard results for stream insect species at White Clay Creek (PA, U.S.A.; 20 June 2013). SD refers to standard deviation of expected C-score values. Significant values are bolded. The total species checkerboard analysis is slightly more than the number of species for filterers and grazers combined, given the inclusion of a third predatory guild (with two species).

Analysis (*n* = species)	Obs. C-score	Exp. C-score	SD	Standardised C-score	*P*(obs < exp)	*P*(obs > exp)
Filterers (*n* = 23)	9.73	9.26	0.10	4.52	>0.999	**<0.001**
Grazers (*n* = 15)	3.50	3.60	0.08	−1.19	0.087	0.920
All (*n* = 40)	28.63	28.11	0.21	2.49	0.986	**0.015**

barcoding successfully delineated analysis units for our study, which correspond closely to biological species in stream insects (Zhou *et al.*, 2010; Sweeney *et al.*, 2011; Jackson *et al.*, 2014). Note, the aims of our study required the delineation of species (i.e. units of study), not necessarily the identification of species (assigning a species name; Collins & Cruickshank, 2013), and as such, the unidentified species do not impact our results. Our molecular approach to species delineation was then coupled with additive diversity partitioning and checkerboard analyses at a small spatial scale, where environmental filtering and dispersal processes become less important in governing species distributions (Willis & Whittaker, 2002; Cavender-Bares *et al.*, 2009); these analyses in turn assessed if signatures of competition could be detected using our observational data. Our study represents a combined approach as a means to detect meaningful signals of competition from observational data.

Non-random distributions in species diversity at White Clay Creek

Species diversity for both filterer and grazer guilds was non-randomly distributed within White Clay Creek (Fig. 2). Using an analysis method that considers the density of individuals, we found a significant shortage of species at the smallest spatial level (e.g. 8.1 observed versus 19.8 expected for filterers, α rock), where on average only 30% of the total species pool occurred (Fig. 2). Similarly, Ligeiro *et al.* (2010) analysed stream insect diversity at the family level using a small-scale design (<5 km) and also found α diversity at the level of centimetres was only 30% of the total family pool. These results indicate that within-stream processes significantly affect the distribution of species within these guilds, suggesting that competition in the form of inter- and intraspecific interactions could lead to these non-random patterns.

We also have to consider, however, the dispersive capabilities of the insect taxa analysed, which could account for high β diversity at the scale of kilometres. Some species of Trichoptera are considered ‘weak dispersers’, with some individuals travelling short distances (roughly a kilometre) between headwaters (Bilton, Freeland & Okamura, 2001), while individuals within some families such as Hydropsychidae may travel much longer distances (up to 16 km; Coutant, 1982). Simuliidae individuals are known to travel large distances (15 km) in search of a blood meal (Merritt *et al.*, 2008).

Presumably, dispersal limitations are therefore less important in Simuliidae for the extent of White Clay Creek sampled (*c.* 3 km). Of the two families of Coleoptera sampled, Elmidae (riffle beetles) is capable of flight as adults (Elliot, 2008), while some evidence suggests that the Psephenidae are poor dispersers, hardly travelling more than 5 m from resident streams during a brief terrestrial adult stage (Miller, Blinn & Keim, 2002). Given the filterers analysed were largely comprised of the more dispersive Hydrosychidae and Simuliidae families, dispersal limitation likely does not explain the significantly high β diversity at the scale of kilometres. In addition, Martin, Adamowicz & Cottenie (2016) demonstrate spatial distances play at most a minor role in structuring community composition of stream invertebrates at an *c.* 100 km scale (up to 3% of variation explained in Diptera). It seems very unlikely spatial distances played a role in structuring insects in the 3 km extent of White Clay Creek we sampled. We cannot discount, however, that grazers may have exhibited some level of dispersal limitation given we sampled poorly dispersive families of grazing Trichoptera and the Coleoptera family Psephenidae. Because the checkerboard analyses showed grazers did not segregate, the significantly high β diversity at the scale of kilometres for grazers may be more readily explained by dispersal limitations rather than competitive interactions.

Competitive interactions at White Clay Creek

Once it was established that variation in species diversity within the filterer and grazer guilds at White Clay Creek was structured at the scale of kilometres (e.g. most of the departure from random expectations in β diversity occurred at this within-stream scale), checkerboard analyses were used to investigate if species within these guilds segregate; patterns of segregation would, in turn, be indicative of important food or space resource competition. Note that the environment was assumed to have not impacted species distributions at White Clay Creek due to the small spatial scale sampled. Based on larger scale sampling of nine streams in southern Pennsylvania, the relative importance of environmental filtering in driving β diversity does appear to be minimal at the scale of White Clay Creek (see Figure S1). Most environmental variation when sampling at a regional level (*c.* 100 km extent) occurred at the stream level, with comparatively less environmental variation within a given stream (e.g. transect level, with most of that expected to be between pools and riffles, whereas only riffles were included in the main study here). Additive β diversity in Pennsylvanian streams increased or decreased accordingly with this environmental variation (Table S2, Figure S1), showing that, on average, only 30% of the regional diversity will remain in a given stream (presumably due to the environmental filtering mechanism). As such, we believe our assumption of environmental variation playing a minor role in structuring species distributions at the scale of White Clay Creek is valid.

Checkerboard analyses detected significant segregation of species at White Clay Creek within the filterers, but random structure within the grazers (Fig. 2, Table 2). Most of the species of filterers analysed within White Clay Creek belonged to the families Simuliidae and Hydropsychidae (net-spinning caddisflies). Both of these families attach to stream rocks as aquatic larvae in fast-flowing streams and are known to undergo intense competition for filter-feeding locations (Thorp, 1983; Georgian & Thorp, 1992). Hydropsychidae are particularly territorial with each other and with other taxa for these feeding locations, exhibiting a range of aggressive behaviours ranging from stridulation as a warning mechanism to sometimes fighting to the death for retreats (Jansson & Vuoristo, 1979). Simuliidae, on the other hand, will actively avoid filter-feeding sites occupied by Hydropsychidae (based on *in situ* experiments; Hemphill & Cooper, 1983; Hemphill, 1988, 1991). These experimental and behavioural studies are consistent with our finding that filterers at White Clay Creek are competitively structured, supporting our observational approach to detecting competitive interactions in communities. Given Simuliidae are subordinate to the larger, more aggressive Hydropsychidae, these competitive interactions should result in exclusion of simuliid species within the filter-feeding locations (i.e. on rocks) when hydropsychids are present. Note that, given the highly asymmetric competitive relationship between Hydropsychidae and Simuliidae, a different process must allow the inferior competitor to persist within White Clay Creek to fully explain the high C-score. Stream systems are known to undergo regular disturbance, which allows inferior competitors to persist in the presence of dominant ones; Simuliidae species in particular are early colonisers and will taper in density as other species settle into a microhabitat (Hemphill & Cooper, 1983; Hemphill, 1991; Effenberger *et al.*, 2008). It is likely the community structure at White Clay Creek undergoes cycles of segregation and random structure; monitoring these patterns over time would shed light on this issue. Another possibility is that habitat preference may occur at a scale smaller than the rocks we used as

sample units, such as with subtle differences in water flow across the rocks, thereby promoting some degree of co-occurrence of taxa. This may explain why some species of Hydropsychidae and Simuliidae occur in the same sample, but clearly not all species of Hydropsychidae and Simuliidae (19 total in our study) co-occur consistently on the same rock (Fig. 2).

Given competitive interactions were detected in the filterer guild, we next sought to determine the relative role of competition in the grazer guild. Our analysis of grazers showed random structure approaching aggregation of species at the rock level (Fig. 2). Patterns of aggregation may reflect patchy resources; in this case, rocks growing sufficient food resources such as algae may be patchy where sunlight is reaching the streambed. In this case, the patchy resource needs to be consistently targeted for sampling to infer if competition is occurring where the resource is available; we may not have achieved this in our own sampling. Another possible explanation for the random structure is that the grazers are still recovering from a disturbance episode; over time, we might expect the grazers to settle into a non-random assemblage of species, in particular aggregation as the checkerboards suggest the community structure is approaching this. Lastly, a competitive signal in community structure could be obscured if grazers are relatively mobile (i.e. do not have same constraints as filterers in occupying 'ideal sites'). Tentatively, however, our analysis appears to rule out competition as driving the structure of species occurrences among grazers; it is worth noting the densities of grazer species was low compared to densities of filterers (in particular Hydropsychidae and Simuliidae), which would reduce the pressure of competition and permit co-occurrence providing resources are not limited for this guild. The significant C-score for all species analysed together likely reflects the tendency to segregate in the more species-rich filterer guild (Fig. 2).

Several limitations with regard to our study must be addressed. The sample size of 30 rocks is a possible concern, particularly two rocks per riffle. Despite this, the partition analysis indicated that the species pool does not increase significantly when moving from the rock to riffle level (Table 1, Fig. 1), suggesting heterogeneity in species distributions at this scale is negligible. The rarefaction curve for the sampling of 30 rocks also indicates the species pool (over the 3 km extent) was adequately sampled (see Figure S2). These results – in tandem with the small spatial scale sampled, the fact that all rocks were sampled within a few hours, and the robustness of checkerboard analyses to type I errors (Gotelli, 2000) – strongly suggest our sample size was appropriate for the filterer guild; whether grazers truly showed random structure or if a type II error occurred is less clear. Note that other analyses including other abundant insect orders not investigated in this study (i.e. Ephemeroptera, Plecoptera, Chironomidae) could reveal important segregation of species within grazers. With regard to the checkerboard analysis, competitively superior species present an inherent problem; if certain species occur at every site and drive other species to exclusion, this will not produce a checkerboard as expected of competitive interactions. This occurs because there is no way to assess if the excluded species would exist without the presence of the common one. In this case, important competitive interactions occur within the community but will go undetected by the checkerboard analysis. Similarly, checkerboards rely solely on presence/absence data, or competitive exclusion, as a proxy for competitive interactions. If competitive exclusion is a rare outcome in competitive interactions, then density data are needed to investigate the effect two species have on one another; this issue may have masked important interactions within the grazer guild, resulting in random structure. While diversity partitioning cannot infer the importance of competition between specific taxa, it can be used to determine if diversity is non-randomly structured according to different orders of diversity (weighting the diversity index to emphasise rare or common species), and thus could be used to further elucidate density-dependent patterns. Finally, while checkerboards can determine if species tend to segregate within specific taxa, they cannot elucidate if specific species pairs segregate more than expected by chance. A null model for interpreting departures in the number of checkerboards for species pairs would be a welcome addition to currently available methods for studying species-specific interactions.

Competitive interactions have long been assumed to play a role in governing species distributions (Gause, 1932; Hardin, 1960; Gurevitch *et al.*, 1992), but their importance has remained difficult to infer using observational data. By sampling over a small spatial extent, however, we minimised the importance of other processes such as environmental filtering and dispersal limitations, which allowed us to attribute non-random distributions of species to competitive interactions. We coupled this principle with a molecular approach, DNA barcoding, to species delineation in order to increase accuracy of our data set when searching for signatures of competition in communities. We verified the utility of the above approach by first detecting non-random

structure in filterer species distributions at White Clay Creek, a result consistent with experimental evidence suggesting these non-random patterns are attributable to competitive interactions. We then demonstrated that competitive interactions did not appear to play an important role in structuring grazers at White Clay Creek. Researchers seeking to investigate the role of competitive interactions in community structure might employ a similar approach; if researchers can supplement sampling of species diversity at regional spatial extents with more localised efforts, then we will achieve a clearer picture of the full scale of processes governing the distribution of biodiversity.

Acknowledgments

We thank everyone who helped during field collections: Jan Battle, Mike Broomall and the Stroud Water Research Center staff in Pennsylvania. We also thank Hafiz Maherali and Bernard Sweeney for input early in the project, and those who helped with sample processing: Emma Garden, Krisselle Rondolo and Susannah Ripley. We thank the staff of the Canadian Centre for DNA Barcoding at the Biodiversity Institute of Ontario, University of Guelph, for molecular analysis. This project was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) through a Discovery Grant (386591-2010) and an NSERC Post-Graduate Scholarship, the Ontario Ministry of Research and Innovation through the Global Leadership Round in Genomics and Life Sciences program (ORF-RE GL2-01-015), a grant from the Government of Canada through Genome Canada and the Ontario Genomics Institute in support of the International Barcode of Life project (2008-OGI-ICI-03), an Ontario Graduate Scholarship, the O'Brien Fellowship, a Hart Aquatic Biodiversity Scholarship, and the University of Guelph through the Work-Study program and an International Collaborative Research travel grant from the College of Biological Sciences. Support for John Jackson and the staff at the Stroud Water Research Center was provided by the Pennswood Endowment Fund, the Stroud Water Research Center Endowment Funds, and a National Science Foundation Grant DEB-1052716.

References

Bell G. (2001) Neutral macroecology. *Science*, **293**, 2413–2418.

Bell G. (2005) The co-distribution of species in relation to the neutral theory of community ecology. *Ecology*, **86**, 1757–1770.

Bilton D.T., Freeland J.R. & Okamura B. (2001) Dispersal in freshwater invertebrates. *Annual Review of Ecology and Systematics*, **32**, 159–181.

Blagoev G.A., Nikolova N.I., Sobel C.N., Hebert P.D.N. & Adamowicz S.J. (2013) Spiders (Araneae) of Churchill, Manitoba: DNA barcodes and morphology reveal high species diversity and new Canadian records. *BMC Ecology*, **44**, 1–17.

Boschilia S.M., Oliveira E.F. & Thomaz S.M. (2008) Do aquatic macrophytes co-occur randomly? An analysis of null models in a tropical floodplain. *Oecologia*, **156**, 203–214.

Bringloe T.T. (2014) *Heirarchical Spatial Structuring of Stream Insect Diversity through DNA Barcoding*. MSc Thesis, University of Guelph, Ontario.

Cavender-Bares J., Kozak K.H., Fine P.V.A. & Kembel S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.

Clare E.L., Barber B.R., Sweeney B.W., Hebert P.D.N. & Fenton M.B. (2011) Eating local: influences of habitat on the diet of little brown bats (*Myotis lucifugus*). *Molecular Ecology*, **20**, 1772–1780.

Collins R.A. & Cruickshank R.H. (2013) The seven deadly sins of DNA barcoding. *Molecular Ecology Resources*, **13**, 969–975.

Colwell R.K., Chao A., Gotelli N.J., Lin S.Y., Mao C.X., Chazdon R.L. *et al.* (2012) Models and estimators linking individual-based and sample-based rarefaction, extrapolation, and comparison of assemblages. *Journal of Plant Ecology*, **5**, 3–21.

Connor E.F. & Simberloff D. (1979) The assembly of species communities: chance or competition? *Ecology*, **60**, 1132–1140.

Coutant C.C. (1982) Evidence for upstream dispersion of adult caddisflies (Trichoptera, Hydropsychidae) in the Columbia River. *Aquatic Insects*, **4**, 61–66.

Crist T.O., Veech J.A., Gering J.C. & Summerville K.S. (2003) Partitioning species diversity across landscapes and regions: a hierarchical analysis of α ß and γ diversity. *The American Naturalist*, **162**, 734–743.

De Troch M., Melgo-Ebarle J.L., Angsinco-Jimenez L., Gheerardyn H. & Vincx M. (2008) Diversity and habitat selectivity of harpacticoid copepods from sea grass beds in Pujada Bay, the Philippines. *Journal of the Marine Biological Association of the United Kingdom*, **88**, 515–526.

Diamond J.M. (1975) Assembly of species communities. In: *Ecology and Evolution of Communities* (Eds M.L. Cody & J.M. Diamond), pp. 342–444. Harvard University, Belknap Press, Cambridge.

Effenberger M., Engel J., Sabastian D. & Matthaei C.D. (2008) Disturbance history influences the distribution of stream invertebrates by altering microhabitat parameters: a field experiment. *Freshwater Biology*, **53**, 996–1011.

Elliot J.M. (2008) The ecology of riffle beetles (Coleoptera: Elmidae). *Freshwater Reviews*, **1**, 189–203.

Fayle T.M. & Manica A. (2010) Reducing over-reporting of deterministic co-occurrence patterns in biotic communities. *Ecological Modelling*, **221**, 2237–2242.

Fujisawa T. & Barraclough T.G. (2013) Delimiting species using single-locus data and the generalized mixed yule coalescent (GMYC) approach: a revised method and evaluation on simulated data sets. *Systematic Biology*, **62**, 707–724.

Gause G.F. (1932) Experimental studies on the struggle for existence. *The Journal of Experimental Biology*, **9**, 389–402.

Georgian T. & Thorp J.H. (1992) Effects of microhabitat selection on feeding rates of net-spinning caddisfly. *Ecology*, **73**, 229–240.

Gheerardyn H., De Troch M., Vincx M. & Vanreusel A. (2010) Diversity and community structure of harpacticoid copepods associated with cold-water coral substrates in the Porcupine Seabight (North-East Atlantic). *Helgoland and Marine Research*, **64**, 53–62.

Gill B.A., Harrington R.A., Kondratieff B.C., Zamidio K.R., Poff N.L. & Funk W.C. (2014) Morphology taxonomy, DNA barcoding, and species diversity in southern Rocky Mountain headwater streams. *Freshwater Science*, **33**, 288–301.

Gotelli N.J. (2000) Null model analysis of species co-occurrence patterns. *Ecology*, **81**, 2606–2621.

Gotelli N.J. & McCabe D.J. (2002) Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model.. *Ecology*, **83**, 2091–2096.

Gurevitch J., Morrow L.L., Wallace A. & Walsh J.S. (1992) A meta-analysis of competition in field experiments. *The American Naturalist*, **140**, 539–572.

Hamilton A.J., Basset Y., Benke K.K., Grimbacher P.S., Miller S.E., Novotny V. *et al.* (2010) Quantifying uncertainty in estimation of tropical arthropod species richness. *The American Naturalist*, **176**, 90–95.

Hardin G. (1960) The competitive exclusion principle. *Science*, **131**, 1292–1297.

Hebert P.D.N., Cywinska A., Ball S.L. & deWaard J.R. (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 313–321.

Hebert P.D.N., Penton E.H., Burns J.M., Janzen D.H. & Hallwachs W. (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States*, **101**, 14812–14817.

Hebert P.D.N., Ratnasingham S. & deWaard J.R. (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 96–99.

Hemphill N. (1988) Competition between two stream dwelling filter-feeders *Hydrophyche oslari* and *Simulium virgatum*. *Oecologia*, **77**, 73–80.

Hemphill N. (1991) Disturbance and variation in competition between 2 stream insects. *Ecology*, **72**, 864–872.

Hemphill N. & Cooper S.D. (1983) The effect of physical disturbance on the relative abundances of two filter-feeding insects in a small stream. *Oecologia*, **58**, 378–382.

Jackson J.K., Battle J.M., White B.P., Pilgrim E.M., Stein E.D., Miller P.E. *et al.* (2014) Cryptic biodiversity in streams: a comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. *Freshwater Science*, **33**, 312–324.

Jansson A. & Vuoristo T. (1979) Significance of stridulation in larval Hydropsychidae (Trichoptera). *Behaviour*, **71**, 167–186.

Jones M., Ghoorah A. & Blaxter M. (2011) jMOTU and Taxonerator: turning DNA Barcode sequences into annotated operational taxonomic units. *PLoS ONE*, **6**, e19259.

Kaartinen R., Stone G.N., Hearn J., Lohse K. & Roslin T. (2010) Revealing secret liaisons: DNA barcoding changes our understanding of food webs. *Ecological Entomology*, **35**, 623–638.

Kamilar J.M. & Ledogar J.A. (2011) Species co-occurrence patterns and dietary resource competition in primates. *American Journal of Physical Anthropology*, **144**, 131–139.

Kunstler G., Lavergne S., Courbaud B., Thuiller W., Vieilledent G., Zimmermann N.E. *et al.* (2012) Competitive interactions between forest trees are driven by species' trait hierarchy, not phylogenetic or functional similarity: implications for forest community assembly. *Ecology Letters*, **15**, 831–840.

Ligeiro R., Melo A.S. & Callisto M. (2010) Spatial scale and the diversity of stream macroinvertebrates in a Neotropical catchment. *Freshwater Biology*, **55**, 424–435.

Martin G., Adamowicz S.J. & Cottenie K. (2016) Taxonomic resolution based on DNA barcoding affects environmental signal in metacommunity structure. *Freshwater Science*, **35**.

Mayfield M.M. & Levine J.M. (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters*, **13**, 1085–1093.

McDermott H., Paull T. & Strachan S. (2010) *Canadian Aquatic Biomonitoring Network: Laboratory Manuals, Processing, Taxonomy, and Quality Control of Benthic Macroinvertebrate Samples*. Environment Canada.

Merritt R.W., Cummins K.W. & Berg M.B. (2008) *An Introduction to the Aquatic Insects of North America*, 4th edn. Kendall Hunt, Dubuque.

Miller M.P., Blinn D.W. & Keim P. (2002) Correlations between observed dispersal capabilities and patterns of genetic differentiation in populations of four aquatic insect species from the Arizona White Mountains, U.S.A. *Freshwater Biology*, **47**, 1660–1673.

Mora C., Tittensor D.P., Adl S., Simpson A.G.B. & Worm B. (2011) How many species are there on earth and in the ocean? *PLoS ONE*, **9**, e1001127.

Negro M., Rolando A. & Palestrini C. (2011) The impact of overgrazing on dung beetle diversity in the Italian maritime alps. *Environmental Entomology*, **104**, 1081–1092.

Pfenninger M., Nowak C., Kley C., Steinke D. & Streit B. (2007) Utility of DNA taxonomy and barcoding for the inference of larval community structure in morphologically cryptic *Chironomus* (Diptera) species. *Molecular Ecology*, **16**, 1957–1968.

Pfrender M.E., Hawkins C.P., Bagley M., Courtney G.W., Creutzburg B.R., Epler J.H. *et al.* (2010) Assessing macroinvertebrate biodiversity in freshwater ecosystems: advances and challenges in DNA-based approaches. *The Quarterly Review of Biology*, **85**, 319–340.

Puillandre N., Lambert A., Brouillet S. & Achaz G. (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, **21**, 1864–1877.

Ratnasingham S. & Hebert P.D.N. (2007) BOLD: the barcode of life data system. *Molecular Ecology Notes*, **7**, 355–364.

Ratnasingham S. & Hebert P.D.N. (2013) A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS ONE*, **8**, e66213.

Rodríguez-Zaragoza F.A., Cupul-Magaña A.L., Galván-Villa C.M., Ríos-Jara E., Ortiz M., Robles-Jarero E.G. *et al.* (2011) Additive partitioning of reef fish diversity variation: a promising marine biodiversity research tool. *Biodiversity Conservation*, **20**, 1655–1675.

Schoener T.W. & Adler G.H. (1991) Greater resolution of distributional complementarities by controlling for habitat affinities: a study with Bahaman lizards and birds. *The American Naturalist*, **137**, 669–692.

Smith A.M., Rodriguez J.R., Whitfield J.B., Deans A.R. & Janzen D.H. (2008) Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proceedings of the National Academy of Sciences of the United States*, **105**, 12359–12364.

Stone L. & Roberts A. (1990) The checkerboard score and species distributions. *Oecologia*, **85**, 74–79.

Sweeney B.W., Battle J.M., Jackson J.K. & Dapkey T. (2011) Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *Journal of the North American Benthological Society*, **30**, 195–216.

Thorp J.H. (1983) An evaluation of hypotheses on the evolutionary differentiation of catchnets in net-spinning caddisflies (Hydropsychidae). *Oikos*, **40**, 308–312.

Tuomisto H. (2010) A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, **33**, 2–22.

Ulrich W. (2004) Species co-occurrences and neutral models: reassessing J. M. Diamond's assembly rules. *Oikos*, **107**, 603–609.

Valentini A., Pompanon F. & Taberlet P. (2008) DNA barcoding for ecologists. *Trends in Ecology and Evolution*, **24**, 110–117.

Veech J.A. (2005) Analyzing patterns of species diversity as departures from random expectations. *Oikos*, **108**, 149–155.

Veech J.A. & Crist T.O. (2010) Toward a unified view of diversity partitioning. *Ecology*, **91**, 1988–1992.

Veech J.A., Summerville K.S., Crist T.O. & Gering J.C. (2002) The additive partitioning of species diversity: recent revival of an old idea. *Oikos*, **99**, 3–9.

Warren D.L., Cardillo M., Rosauer D.F. & Bolnick D.I. (2014) Mistaking geography for biology: inferring processes from species distributions. *Trends in Ecology and Evolution*, **29**, 572–580.

Willis K.J. & Whittaker R.J. (2002) Species diversity-scale matters. *Science*, **295**, 1245–1248.

Wilsey B.J. (2010) An empirical comparison of beta diversity indices in establishing prairies. *Ecology*, **91**, 1984–1988.

Witt J.D.S., Threloff D.L. & Hebert P.D.N. (2006) DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. *Molecular Ecology*, **15**, 3073–3082.

Zhou X., Jacobus L.M., DeWalt R.E., Adamowicz S.J. & Hebert P.D.N. (2010) Ephemeroptera, Plecoptera, and Trichoptera fauna of Churchill (Manitoba, Canada): insights into biodiversity patterns from DNA barcoding. *Journal of the North American Benthological Society*, **29**, 814–837.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Environmental variation analysis, including Tables S1, S2 and Figure S1.

Appendix S2. DNA barcoding pipeline, including Table S3. Primers and regime times used to sequence the COI barcode regions of stream insects.

Table S4. Sampled insect species list for White Clay Creek (PA, U.S.A.; 20 June 2013).

Figure S2. Rarefaction curve for of stream insect species at White Clay Creek (PA, U.S.A.; 20 June 2013).

(Manuscript accepted 12 January 2016)

© 2016 John Wiley & Sons Ltd, *Freshwater Biology*, **61,** 646–657