PRIMARY RESEARCH PAPER

Multifaceted assessment of stream fish alpha and beta diversity using spatial models

Zachery D. Zbinden [ID] **· Aaron D. Geheber ·**
Richard J. Lehrter II · William J. Matthews

Abstract Our understanding of the factors determining stream community assembly has mainly been built upon taxonomic diversity. Additionally, most investigations of factors governing assembly are limited in scope and use of spatial model testing. Therefore, we combine a sizeable environmental dataset (local and landscape level) with robust spatial analysis to model the factors determining multiple facets of stream fish diversity—across a region and within four drainages. We found that total dissimilarity was more explained by replacement than nestedness. Fish diversity was governed by spatial and environmental factors, although the degree depended on the diversity facet and drainage. Adjusted R^2 for combined models (i.e., spatial plus environment) of fish beta diversity was as high as 0.65. Local instream habitat variables were the most explanatory factors. Spatial factors based on overland Euclidean distance and asymmetric (directional) distance revealed more explanatory power than symmetric hydrologic distance. The purely spatial variation was often as high as 33% explained. Multiple facets were strongly correlated and provided similar results—albeit some differences arose, especially within different drainages; these differences highlight how difficult it is to generalize the multifaceted approach. However, the importance of exploring spatial models for determining the role of spatial processes on community assembly cannot be overstated.

Keywords Asymmetric eigenvector maps · Dispersal limitation · Environmental filtering · Species replacement · Upstream dispersal · Variation partitioning

Z. D. Zbinden (✉) · A. D. Geheber · R. J. Lehrter II · W. J. Matthews
Department of Biology, University of Oklahoma, Norman, OK, USA
e-mail: zdzbinde@uark.edu

Present Address:
Z. D. Zbinden
Department of Biological Sciences, University of Arkansas, Fayetteville, AR, USA

Present Address:
A. D. Geheber
School of Natural Sciences, University of Central Missouri, Warrensburg, MO, USA

Present Address:
R. J. Lehrter II
National Ecological Observatory Network, Boulder, CO, USA

Introduction

Conservation and management of biodiversity rely on understanding the factors and processes responsible for creating and maintaining assemblages of species (López-Delgado et al., 2020). Additional facets—or dimensions—of biodiversity can provide complementary information about the factors that structure communities (Pavoine & Bonsall, 2011; Nakamura et al., 2020). Nevertheless, much of our understanding of community assembly has relied solely on taxon diversity (TD), based on species abundance or incidence, while largely neglecting other facets (Pool et al., 2014; Mori et al., 2018). Taxon diversity does not account for evolutionary and ecological species relationships, and therefore, it fails to encompass essential pieces of the natural puzzle (Cardoso et al., 2014; Heino & Tolonen, 2017; Roa-Fuentes et al., 2019). Although additional means of quantifying diversity based on functional traits (McNaughton, 1977; Grossman et al., 1982) or phylogenetic relationships (Faith, 1992) have existed for some time, examination of more than one facet of diversity within a single study has only recently increased in popularity (Webb et al., 2002; Meynard et al., 2011; Branco et al., 2020). Still, most community assembly studies in freshwater systems exclusively focus on TD (Perez Rocha et al., 2018; Hill et al., 2019).

Functional diversity (FD) represents the unique morphological, physiological, and ecological traits present within a community (Pool et al., 2014). FD differs less among localities than TD due to functional redundancy (Villéger et al., 2012, 2013). However, because species traits often mediate environmental filtering, FD is predicted to be more closely related to environmental factors and, therefore, can better illuminate abiotic relationships with community structure (Poff, 1997; Mori et al., 2018). Phylogenetic diversity (PD) represents the unique evolutionary history present within a community (Faith, 1992). PD can provide additional insights into how evolutionary processes structure contemporary communities (Webb et al., 2002). For example, environmental filtering may drive co-occurrence among closely related species with similar characteristics, tolerances, and preferences. However, competitive interactions among those same species place constraints on coexistence (Mayfield & Levine, 2010; Geheber & Geheber, 2016; Geheber, 2019).

Compared to TD, FD and PD measures show fewer limitations when quantifying and comparing diversity across large spatial scales (Perez Rocha et al., 2018). The reason is that both FD and PD are not as sensitive as TD to dissimilarity saturation—when communities have maximum dissimilarity due to a lack of shared species (Tuomisto et al., 2012). Dissimilarity saturation occurs when sampling over long ecological or spatial gradients. However, even in the absence of shared species, communities can still overlap in evolutionary history (e.g., shared families) and functionality (e.g., piscivores and detritivores) (Brown et al., 2017). Consequently, the utility of these additional facets of biodiversity depends on the spatial extent of investigation (Legendre, 2014). TD is least practical at large spatial scales at which species pools totally differ. At such scales, PD and FD may prove more useful. Evidence suggests that multiple facets of biodiversity may be more complementary at lower than higher latitudes (Stevens & Tello, 2018); this may suggest that TD alone—at appropriate scales—is a suitable metric for the study and conservation of temperate stream systems.

Identifying extrinsic factors that correlate with patterns of community structure and diversity is a significant ecological pursuit and a primary step towards a complete understanding of assembly mechanisms and ecological dynamics (Vellend, 2010; Leibold & Chase, 2018). Diversity measures are often coupled with environmental and spatial characteristics to parse the effects of local and landscape-level abiotic factors, spatially structured abiotic factors, and purely spatial factors on community structure (Legendre et al., 2005). These associations—or lack thereof—are presumed to support different underlying mechanisms: environmental filtering, dispersal limitation, or stochasticity (Cottenie, 2005; Leibold & Chase, 2018). Although diversity variation explained by purely spatial factors—after accounting for the environment—is typically used to infer dispersal limitation, it should be noted that unmeasured environmental variation can also drive such a pattern (Landeiro et al., 2011). In addition, the strength of the relationships among geographic variation of community structure and environmental and spatial factors is often used to infer metacommunity-level processes (Leibold et al., 2004), although the predictions for the fundamental archetypes (e.g., Neutral Theory; Species Sorting; Patch Dynamics; Mass Effects), are

not often well defined or mutually exclusive (Leibold & Chase, 2018). More studies across taxa, systems, and spatial scales are needed to determine the value of studying additional facets of diversity (Stevens & Tello, 2018).

Investigations that aim to partition variation among spatial and environmental predictors typically use spatial eigenfunction analysis (Griffith & Peres-Neto, 2006) to decompose a matrix of spatial sampling configuration (e.g., pairwise Euclidean distance among sites) into eigenvectors which can be used to model latent spatial processes such as dispersal limitation (Borcard & Legendre, 2002). However, Euclidean distance is usually deemed inappropriate for modeling spatial processes for stream obligate species; therefore, hydrologic network distance (i.e., river distance or watercourse distance) is typically used instead (Fetzner & Crandall, 2003; Landeiro et al., 2011). However, other methods may better portray the unique spatial structure of stream networks because they consider both the network position and directional processes such as flow direction (Blanchet et al., 2008a) and might more robustly infer the effects of dispersal limitation (Landeiro et al., 2011).

The present study aimed to determine if multiple facets of diversity provide complementary information regarding stream fish assemblage structure. To better understand species assembly, beta diversity facets were partitioned into species replacement (i.e., turnover) and nestedness components (Baselga, 2010). Replacement is indicated when species substitutions occur between different sites, presumably due to environmental filtering (Nunes et al., 2016; Heino & Tolonen, 2017), and nestedness occurs when less speciose sites are subsets of more speciose ones (Baselga, 2010); it is related to stochastic dispersal or colonization–extinction dynamics along gradients (Taylor & Warren, 2001; Si et al., 2016). The relative importance of replacement and nestedness also relies on the degree of environmental heterogeneity and dispersal (Gianuca et al., 2017; Zeni et al., 2020). For example, no nestedness is expected in an entirely homogenous system, as no directional species loss should be found in the absence of environmental gradients (Gianuca et al., 2017). However, in environmentally heterogeneous systems, nestedness is expected to depend on the degree of dispersal between environments, with higher dispersal leading to more homogenization

and less nestedness. Therefore, the contribution of each component to total dissimilarity can aid in identification of the underlying mechanisms driving assembly. Generally, species replacement contributes much more than nestedness to overall dissimilarity, as shown across taxa and regions (Soininen et al., 2018; Branco et al., 2020).

To further explore potential assembly mechanisms, we also examined the relative influences of spatial and environmental factors on observed fish diversity. Specifically, this approach allows further elucidation of the strengths of habitat filtering and dispersal limitation mechanisms. To this end, we employ a robust set of environmental predictors in combination with tests of 10 unique spatial models that go beyond pairwise hydrologic distance, some of which include asymmetric network processes (i.e., upstream/downstream dispersal). These models employ directional processes (i.e., flow direction) that affect spatial patterns of community organization and can better estimate the role of spatial processes in structuring lotic communities (Horváth et al., 2016).

Our overarching goal was to examine relationships among abiotic factors (i.e., spatial and environmental factors) and facets of diversity to understand observed patterns of community structure and processes governing assembly. To approach this goal, we aimed to test the following hypotheses: (i) dissimilarity among fish assemblages (i.e., beta diversity) would be driven more by species replacement than nestedness due to environmental and spatial filtering; (ii) spatial and environmental factors should explain facets of stream fish diversity; however, environmental factors should have a relatively stronger influence due to the predominance of local environmental filtering at the spatial scale sampled; (iii) environmental factors should be most influential for describing variation in FD and PD based on the assumption that species functionality and evolutionary histories are directly impacted by environmental conditions. In addition to our hypotheses, we set out to demonstrate the efficacy of using more appropriate spatial models for fluvial networks to capture greater signal of significant spatial processes than has been typically observed in past studies of this kind. Tests of the hypotheses above will provide a more complete understanding of how fish assemblages vary across streams and watersheds needed to inform conservation and management.

Materials and methods

Study area

This study was conducted in four temperate river drainages in the Red River Basin, Oklahoma, USA: Clear Boggy River, Muddy Boggy River, Kiamichi River, and Little River (Fig. 1). The sampling extent was approximately 100 km (latitudinally) by 200 km (longitudinally) and covered an area of approximately 16,500 km². The area encompassed three distinct ecoregions, including the cross timbers region (west), the south-central plains (south), and the highland region of the Ouachita Mountains (east). The western and southern parts of the study region are lower elevation (min. = 101 m a.s.l.). In comparison, the eastern and northern parts are at higher elevation (max. = 753 m a.s.l.) and more topographically rugged. The region is rural and dominated by forests and cattle pastures. Human impact is relatively low, but parts of the eastern drainages (Kiamichi and Little Rivers) are intensely managed for commercial timber (Taylor & Lienesch, 1996).

The sampled stream reaches were concentrated in the highland regions and primarily small, shallow (mean max depth = 1.2 m; mean max width = 10 m) headwater streams (stream order = 1–3). These warm-water streams had moderate summer temperatures (mean = 25.6 °C) and dissolved oxygen (mean = 9.0 ppm). Streams varied from being clear, high gradient, and shaded tributaries in highlands to turbid, low gradient mainstems in the lowlands.
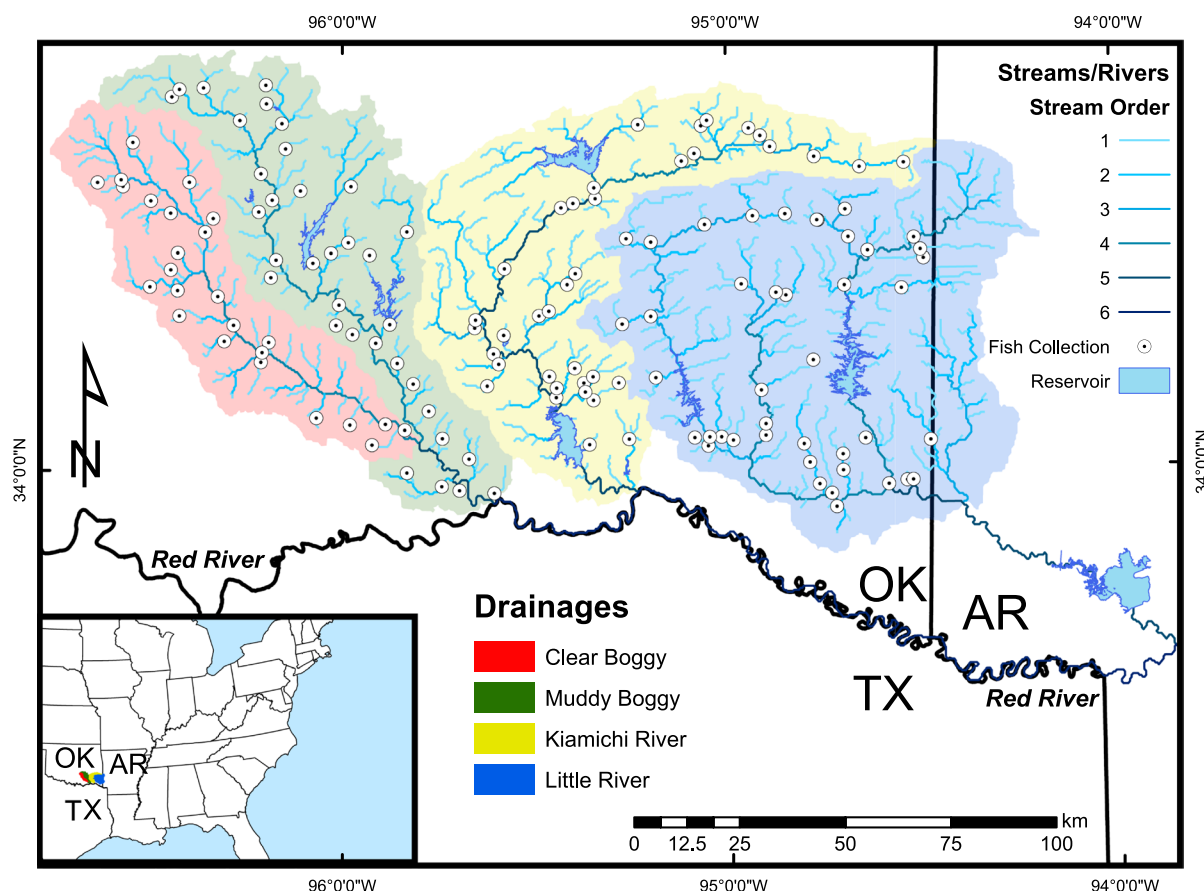
Fig. 1 Map of 138 freshwater fish collections analyzed in this study. Locations were chosen haphazardly and with the intent of providing maximum spatial coverage across each drainage. The Muddy Boggy River, Clear Boggy River, Kiamichi River, and Little River are tributaries of the Red River located in southeastern Oklahoma, USA. Neighboring states Oklahoma (OK), Arkansas (AR), and Texas (TX) are shown for reference

Generally, stream reaches had riffle-pool-run habitats, and substrates were dominated by mud, sand, or gravel. Riparian areas were commonly covered in water willow and spike rushes and usually highly vegetated with large trees. Floating and submerged vegetation was present but less common.

Predictor data

Environmental predictor data

At each sampling location, 23 variables were measured along the stream reach (instream variables hereafter). These variables included: the size of the stream (depth (m) and width (m) recorded at the maximum); the proportion of habitat types present (pool, riffle, channel, and backwater); dissolved oxygen (mg/l; measured using a Horiba Water Quality Monitor model U-5000, Alvin, TX.); relative current velocity (none; slow < 15 cm/s; medium < 30 cm/s; fast > 30 cm/s); the proportion of substrate types present (mud, sand, gravel, cobble, bedrock) (Simonson et al., 1993); incidence of structure types (filamentous algae, macrophytes, submerged structure, coarse woody debris, boulders); and riparian characteristics (bank incision, incidence of bank stability, pasture, and woodland) (Marsh-Matthews & Matthews, 2000).

Landscape-level environmental attributes were assigned to each sampling location using data made available by HydroATLAS version 1.0 at 15 arc-second (~500 m) resolution (Linke et al., 2019). The total data consisted of 281 individual attributes for 56 variable types across five categories, including hydrology–physiography, climate, land cover, geology, and anthropogenic influences (See Supplementary S1 & S2 for more information). The associated RiverATLAS v.10 shapefile for North America was imported into ArcMap (ESRI v.10.8), trimmed, and then exported as a shapefile. Sampling location data were also imported into ArcMap and exported as a shapefile. Both files were projected using NAD 1983 (2011) Contiguous USA Albers projection. These shapefiles were imported into R to join the river database attributes to the sampling locations and instream data using the R package 'sf' (Pebesma, 2018).

The instream and landscape-level variables were joined for a total of six categories characterizing each site: instream, hydrology/physiography, climate, land cover, geology, and anthropogenic variables. Invariant

factors were removed. Principal Component Analysis (PCA) was computed for each of the six environmental sets separately. Prior to PCA, the number of variables for each category were: instream = 23, hydro-physio = 31, climate = 92, landcover = 49, geology = 27, anthropogenic = 14. First, highly collinear variables within a set were removed by calculating variation inflation factors (VIF; Fox & Monette, 1992). In a stepwise manner, variables with $VIF \geq 10$ were removed using the R package 'usdm' (Naimi, 2013). The appropriate number of PCs retained for each group was determined via the randomization method based on significant pseudo-F ratios, *Rnd-F* (Peres-Neto et al., 2005) implemented in the R package 'PCDimension' (Coombes et al., 2019). Principal components were computed for each of the six environmental sets using the R function 'prcomp' with standardization, scaling, and default settings. Six sets of environmental PCs were retained. After preliminary analyses, two sites were outliers (> 3 standard deviations than PC mean) and removed before repeating the analyses.

Spatial predictor data

We compared ten different models of spatial autocorrelation (Fig. 2). All models resulted in spatial eigenvectors either via distance-based Moran's eigenvector map analysis (dbMEM; Dray et al., 2006) or asymmetric eigenvector map analysis (AEMs) (Blanchet et al., 2008a). MEMs and AEMs represent spatial structures in the data ranging from broad to fine scale.

Standard symmetric spatial models 1 and 2

The first spatial model (MEM_LAND) was simply the Euclidean distance among sampling locations which was decomposed into spatial eigenvectors via function 'dbmem' from the R package 'adespatial' (Dray et al., 2018). Only positive spatially autocorrelated eigenvectors were retained for further analysis. Model 2 (MEM_HYDRO) differed from Model 1 only because it was based on hydrologic network (river) distance instead. The spatial join between the RiverATLAS and Site shapefiles was used to measure the fluvial distance between all sites using the 'riverdist' R package (Tyers, 2017). This distance matrix was then transformed into a rectangular matrix of spatial eigenvectors as with Model 1 (Fig. 2). The
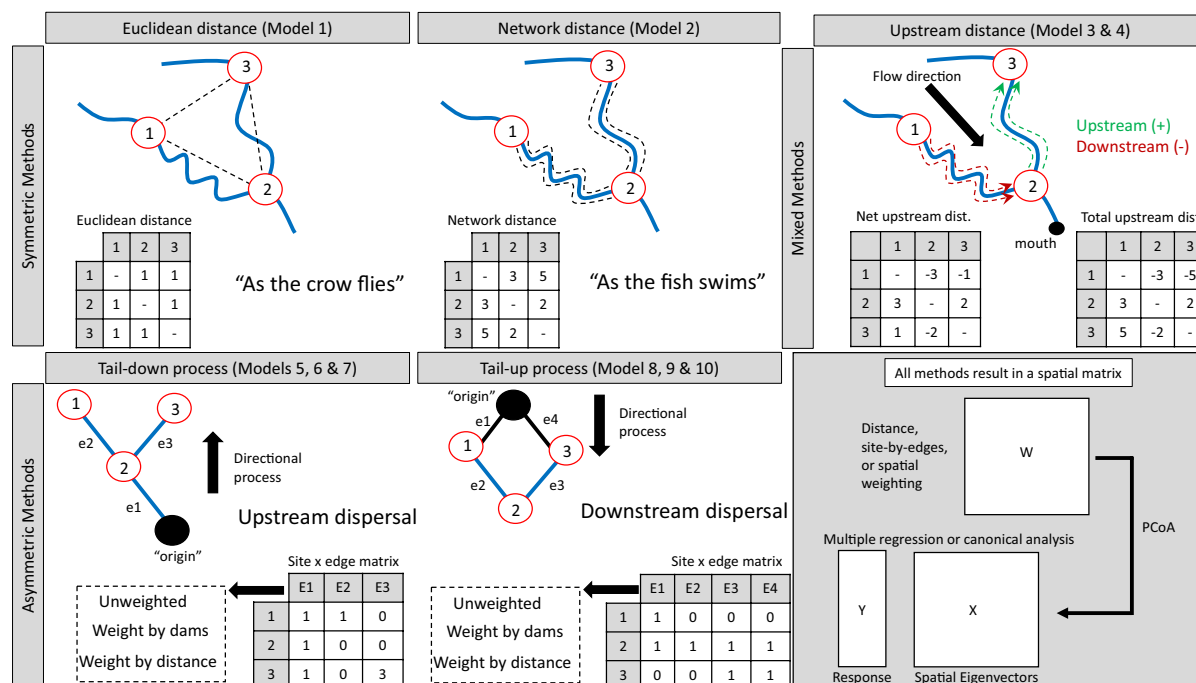
Fig. 2 Spatial autocorrelation of fish diversity was examined across a riverscape in southeastern Oklahoma. Ten models of spatial connectivity were compared. These models are illustrated with a simple three-collection site example. Sites are denoted as red circles with numbers, and the stream network is denoted with blue lines, and distances are denoted with dashed lines, where appropriate. Each method results in a spatial matrix which is decomposed to a series of spatial eigenvectors

threshold for truncating the distance matrix was the longest edge of the minimum spanning tree among sites (as was the case for all MEM models).

Mixed spatial models 3 and 4

The remaining models are all asymmetric with respect to flow direction so that upstream and downstream distances are not treated equally. This can help determine if asymmetric dispersal occurs (e.g., dispersal biased downstream with the current). Essentially, the flow direction is used to weight distances as a means of modeling resistance to dispersal (e.g., Mozzaquattro et al., 2020). The following two models differed from Model 2 in calculating the distance along the river network. These models incorporated flow direction along the stream network to differentiate between downstream and upstream distances (Fig. 2). This distance was calculated using the 'upstream' functions from the R package 'riverdist' (Tyers, 2017). Downstream distances were weighted negative (i.e., reducing distance), while upstream distances were weighted positive (i.e., increasing distance). Models 3 and 4 differed only slightly because Model 3 measured 'net' upstream distance or the difference between up and downstream totals. Model 4 measured 'total' upstream distance. Total network distance (Model 4) was similar to basic hydrologic (Model 2) distance, except the distances had a sign depending on the net direction ($\pm$; upstream/downstream). The 'upstream' distance matrices for Models 3 and 4 were then decomposed into spatial eigenvectors using the same procedure as with Models 1 and 2 (Fig. 2).

Asymmetric models 5 to 10

The remaining models were all based on Asymmetric Eigenvector Map analysis (AEM; Blanchet et al., 2008a, b). AEM was similar to dbMEM in that it generated a rectangular set of spatial eigenvectors that represented the spatial autocorrelation among sampling locations. AEM, however, included directionality in the model so that processes like dispersal from

mainstems to tributaries (tail-down) or directionally biased dispersal mediated by flow (tail-up) were more accurately modeled. AEM analysis required the geographic coordinates of sampled locations and a file containing information about which sites were directly linked by the river network (Blanchet et al., 2011; López-Delgado et al., 2019). We used the AEM-related functions in the R package 'adespatial' (Dray et al., 2018) to build the spatial matrices and decompose them into spatial eigenvectors.

Tail-down models are related to processes that point upstream (Ver Hoef & Peterson, 2010), such as dispersal from mainstem rivers to tributary streams, and this includes our spatial Models 5, 6, and 7. The 'adespatial' function aem.build.binary was used with site coordinates and links with a 'rot.angle'$=0$ degrees because our network flows from north to south and the function infers process from the bottom to the top of the plot area where the network is visualized. The model adds a site to the connection diagram; it serves as the origin of the spatial process that flows through the network. For the tail-down models, we added two additional nodes to the network, one at the confluence of the Muddy Boggy and Red River and another at the confluence of the Kiamichi River and Red River (Fig. 2). These nodes served to represent the spatial network better and were removed after calculating the eigenvectors. Tail-down asymmetric eigenvectors were calculated as described above and without weighting the spatial site by edge matrix for Model 5 (AEM_TDN_UW). Two additional and similar models were calculated that were weighted based on the presence of dams for Model 6 (AEM_TDN_DAM) or distance for Model 7 (AEM_TDN_DIST). For all AEMs, only positive eigenvectors were retained; this was determined by calculating Moran's I statistic and selecting only the eigenvectors with I greater than that expected by chance using the function 'moran.randtest' from the R package 'adespatial' (Dray et al., 2018).

Tail-up models point downstream and indicate downstream biased dispersal, presumably due to the flow of stream current (Ver Hoef & Peterson, 2010). Models 8, 9, and 10 are all tail-up models and were calculated using the same procedure as the tail-down models described above, but with minor differences. First, the 'rot.angle' was set to 180 degrees so the spatial process would be modeled upstream to downstream. An additional site was added to the connection diagram at an arbitrarily chosen location upstream of all other sites and centrally located; this site was then connected to all terminal branches in the network. The origin site was then attached to this single most upstream site so that the flow of current could be modeled through the network. An unweighted Model 8 (AEM_TUP_UW), a dam-weighted Model 9 (AEM_TUP_DAM), and a distance-weighted Model 10 (AEM_TUP_DIST) were produced. Again, only positively autocorrelated eigenvectors were retained.

Fish species data

The sampling approach used here is detailed in Matthews (1986) and Matthews & Marsh-Matthews (2017). The goal of community sampling is to take all species present in proportion to their relative abundances. Fish collections were made between May and July in 2014 and 2015 when flows were low and most amenable to wading. One hundred fifty-one locations were sampled and chosen haphazardly based on access and to provide maximum spatial coverage of the study drainages. Sampling consisted of thoroughly seining all available microhabitats present (e.g., pool, riffle, channel, edge, etc.) within a stream reach approximately 100 m in length. Individual microhabitats were repeatedly seined until an adequate sample representing the species and their relative abundances were obtained (i.e., additional hauls produced 'more of the same species') (Matthews & Marsh-Matthews, 2017). Channels and pools were sampled by pulling seines downstream, while riffle and edge habitats were sampled by kick seining. Depending on the stream's width (min$=2$ m; max$=40$ m; mean$=10$ m), either one or both of two seine nets were used (4.57 m$\times$1.22 m$\times$4.88 mm mesh and/or 2.44 m$\times$1.22 m$\times$4.88 mm mesh). Sampling times varied depending on the size of the stream (mean sampling time was 46 min) and consisted of 25–50 seine hauls. Our sampling approach had yielded highly similar numbers of individuals and species when the same stream reaches were repeatedly sampled over time (Matthews & Marsh-Matthews, 2017; Zbinden et al., 2022). Specimens were preserved in 10% formalin and identified to species in the laboratory. Specimens were archived in the Sam Noble Oklahoma Museum of Natural History, Norman, Oklahoma. All collections were made under

protocols approved by the IACUC of the University of Oklahoma (R14-002).

Before analyses, a species by site matrix was modified to remove rare species (Ross et al., 1985). First, nine species were removed from the matrix that were large-bodied species not sampled consistently with seine nets, for example, gar (*Lepisosteus* spp. Lacepède, 1803), freshwater drum (*Aplodinotus grunniens* Rafinesque, 1819), and smallmouth buffalo [*Ictiobus bubalus* (Rafinesque, 1818)]. Second, 13 species were removed that were collected at < 3 sites. Third, sites were removed from the matrix if they contained < 5 total collected species ($N = 13$). These sites mainly were isolated springs or backwater sloughs. The justification for treating the data in this way was to ensure all sites included were part of the river network and that the species included were those best sampled by our collecting methods (Ross et al., 1985; Poos & Jackson, 2012). Therefore, these modifications should reduce variation due to sampling error.

Functional data

A set of 15 functional fish characteristics was compiled using the Fishtraits database (Frimpong & Angermeier, 2009) and additional literature (e.g., Miller & Robison, 2004; Robison & Buchanan, 2020). The set included seven trophic ecology traits based on feeding: benthic, surface, algae, macrophyte, detritus, piscivore, and eggs. A species received a value of 0 ('no') for each trait unless it has been reported otherwise in the literature for adult individuals at any time over the species' range (Frimpong & Angermeier, 2009). Otherwise, the value was 1 ('yes'). The remaining eight variables were maximum total length, age at maturity, longevity, fecundity, spawning frequency, spawning season length, and minimum and maximum temperature tolerance (Frimpong & Angermeier, 2009 and references therein). These variables were chosen following others who relied on similar size, life history, and trophic-related characteristics to quantify fish functional diversity (Pool et al., 2014; Burgad et al., 2019).

The dimensionality of functional trait data was reduced using non-metric multidimensional scaling on Gower distances among species based on trait data standardized to a mean of zero and one standard deviation. NMDS was performed with function 'metaMDS' (without transformation) in the R package 'vegan' with four axes and 1,000 starts. The solution was constrained to four dimensions because subsequent functional diversity analyses are limited to n-1 functional dimensions where n is the minimum number of species in any site (as mentioned above, we removed sites with less than five species). This constraint led us to choose NMDS over principal coordinates analysis (PCoA = metric multidimensional scaling), typically used for functional diversity analyses of this kind (Villéger et al., 2008). NMDS is designed to provide the most reliable solution relative to the actual distances in a given number of dimensions (Legendre & Legendre, 2012). Stress plots of observed dissimilarity vs. ordination distance were compared for both NMDS and PCoA, and NMDS resulted in a better fit.

Phylogenetic data

A phylogenetic hypothesis of the study species was obtained using a previously assembled time-calibrated tree of ray-finned fishes based on a 27-gene multi-locus alignment (Rabosky et al., 2018). The backbone tree was subset to a phylogram representing character change limited to the species collected in this study using the R package 'fishtree' (Chang et al., 2019). The phylogeny is provided in the Supplementary Material (S3). Derivation of phylogenetic alpha and beta diversity are explained below.

Response data

Alpha diversity

Three facets of stream fish *alpha diversity* were calculated, used to model the contribution of spatial and environmental predictors, and finally used to partition diversity variation among spatial and environmental sets. This procedure was done for each drainage independently and for the whole region. Taxon alpha diversity (TDα) was the number of species found at a given site. Functional alpha diversity (FDα) was calculated based on the convex hull volume of functional space occupied by species assembled at each location. The metric quantified for functional alpha diversity was functional richness, or the proportion of total functional space occupied by each assemblage (Villéger et al., 2008). The function multidimFD

(http://villeger.sebastien.free.fr/Rscripts.html) was used to calculate the functional richness of each assemblage. Phylogenetic alpha diversity (PDα) was calculated based on the sum of the total phylogenetic branch length of the species present within a given site (Faith, 1992). We used the function 'pd' in the R package 'picante' (Kembel et al., 2020), together with the fish phylogeny and fish incidence data described above.

Beta diversity

Beta diversity was based on compositional differences. The same three facets of stream fish *beta diversity* were calculated, used to model spatial and environmental predictors, and finally used to partition diversity variation among spatial and environmental sets. This procedure was done for each drainage independently and for the whole region. Taxon beta diversity was examined using separate analyses on species abundance and incidence data. In addition, each diversity facet was partitioned into overall dissimilarity, replacement, and nestedness components for a total of 12 separate analyses of stream fish beta diversity across the region and within each drainage.

Pairwise beta diversity was calculated and analyzed independently among all sites across the region and within four drainages. Twelve beta diversity matrices were produced using the R package 'betapart' (Baselga & Orme, 2012). Abundance-based Taxon Beta diversity (TDβ) was calculated using function 'beta.pair.abund,' which creates three beta matrices: Bray–Curtis dissimilarity (Bray) and its additive components Bray Balanced (Bray.Bal) and Bray Gradient (Bray.Gra) (Baselga, 2013). Incidence-based Taxon Beta diversity (TDβ) was calculated using function 'beta.pair' which creates three beta matrices: Sørensen dissimilarity (β_{sor}) and its additive components Simpson dissimilarity (β_{sim}) and nestedness (β_{sne}). The same was done for both Functional Beta diversity (FDβ) and Phylogenetic Beta diversity (PDβ) but only using incidence data (abundance based not available in 'betapart') with functions 'functional.beta.pair' and 'phylo.beta.pair', respectively. Finally, a correlation matrix was constructed by calculating the pairwise correlations among each of the 12 resultant beta matrices using Procrustes analysis, and significance was tested using the Procrustean randomization test (Jackson, 1995).

Statistical analyses

Forward selection of predictors and variation partitioning of response

For each vector of alpha diversity (TDα, FDα, and PDα) and each matrix of beta diversity (TDβ Abund, TDβ Incidence, FDβ, and PDβ), we wanted to quantify variation explained by different environmental sets and spatial models. A multiple regression framework was used for modeling the alpha diversity vectors. A distance-based redundancy analysis framework was used for the beta diversity matrices via function 'dbrda' (McArdle & Anderson, 2001) in the R package 'vegan' (Oksanen et al., 2016). Otherwise, the forward selection and variation partitioning procedures are the same for all response data.

First, for a given diversity response, factors from a predictor set (i.e., environmental or spatial) were selected using forward selection (Blanchet et al., 2008b). The selection proceeded only if the overall model, including the response and all factors from a set, was significant ($\alpha = 0.1$). Variables were selected if they increased the adjusted R^2 of the models, were significant ($\alpha = 0.05$), and the adjusted R^2 did not exceed that of the overall model. This method has been suggested over Akaike information criterion (AIC) selection when the objective is to describe patterns as accurately as possible (Bauman et al., 2018; Pollice et al., 2020). Selected environmental variables from the different categories were combined. Collinear variables were removed using the same stepwise VIF procedure detailed above before variation partitioning. The same procedure was employed for the significant spatial variables from different models.

Variation for a given diversity response was then partitioned among the selected spatial and environmental predictors using partial multiple regression or partial distance-based redundancy analysis (Anderson & Legendre, 1999). In all cases, this was done with the 'vegan' function 'varpart.' Each fraction of explained variation was then tested for significance using ANOVA.

Results

The analyzed data set contained 138 fish collections from 4 river drainages in southeastern Oklahoma

(Fig. 1). We collected 58 freshwater fish species and 31,743 individuals. Approximately half (47%) of the species were collected in all drainages, another 41% collected in 2–3, and only 12% were restricted to one. Despite differences driven by unique species, the four river drainages showed much overlap in fish assemblage structure (Supplement S4 to S6).

Functional traits for the 58 fish species were dimensionally reduced to 4 NMDS axes (Supplement S7). The final stress of the solution was 0.091, and the ordination distances were highly correlated with the observed Gower distances (non-metric $R^2 = 0.992$ and linear $R^2 = 0.914$). Phylogenetic relationships of the 58 species of fish were all well resolved (Supplement S3). This tree was not monophyletic and was used only to calculate phylogenetic diversities.

Spatio-environmental predictors

The six environmental sets were reduced to principal components numbering: instream = 3, hydro-physio = 3, climatic = 4, land coverage = 6, geologic = 2, and anthropogenic = 3 PCs. These factors varied among the four river drainages with climate, landcover, and geologic PCs showing a longitudinal trend—eastern and western drainages having different positions along PC1 (Fig. 3). Additional information regarding components and loadings can be found in the Supplementary Material (S8–S13). The ten spatial models resulted in separate matrices of eigenvectors representing the spatial structure among sampling locations. The number of eigenvectors associated with positive eigenvalues analyzed from each of the ten spatial models were as follows: MEM_LAND = 44, MEM_HYDRO = 8, MEM_UPSTR_NET = 32, MEM_UPSTR_TOTAL = 30, AEM_TDN_UW = 14, AEM_TDN_DAM = 15, AEM_TDN_DIST = 15, AEM_TUP_UW = 24, AEM_TUP_DAM = 16, AEM_TUP_DIST = 19.

Alpha diversity across the region

Total species richness across drainages ranged from 40 to 46 species; total functional richness ranged from

Fig. 3 Ordinations of principal components analysis (PCA) of environmental factors characterizing the fish sampling locations in southeastern Oklahoma. There are six major groups of environmental factors, including instream, hydrologic/physiographic, climatic, geologic, landcover/use, and anthropogenic
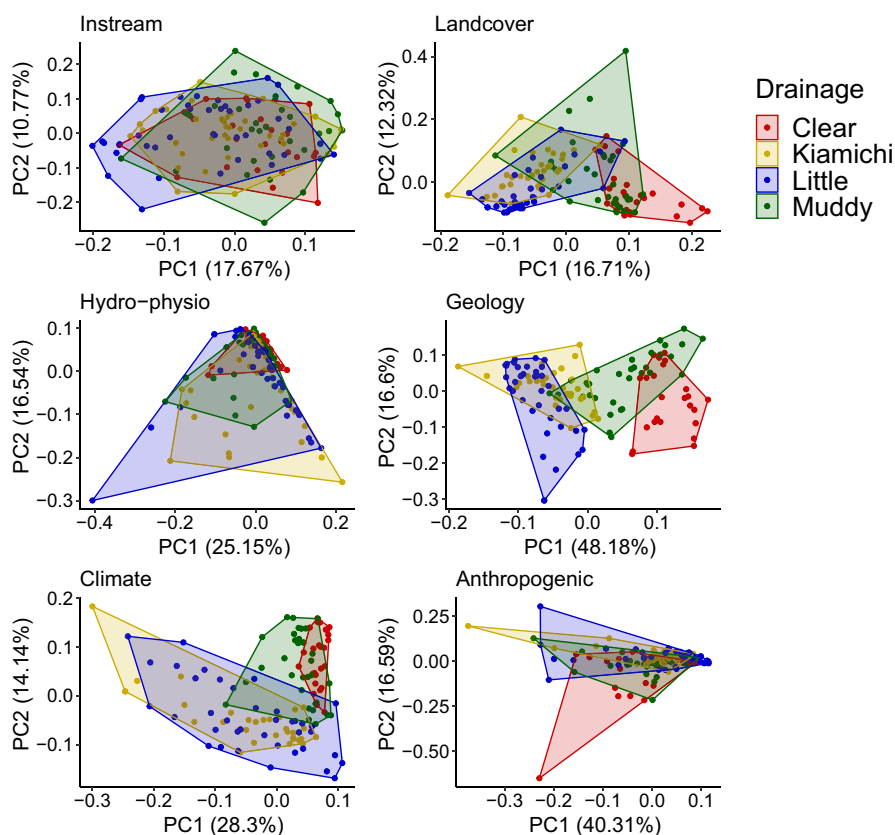
0.75 to 0.80 of the total volume; total phylogenetic diversity ranged from 10.9 to 13.4 branch lengths. The differences in mean alpha diversity among drainages were minor: TDα ranged from 10.5 to 11.8; FDα, as a percentage of total volume, ranged from 0.13 to 0.20; and PDα ranged from 5.9 to 6.3 branch lengths (Fig. 4; Supplement 14). For the entire set of 138 sites analyzed, mean alpha diversities per site were as follows: TDα = 11.1 species, FDα = 0.15 of total functional volume, and PDα = 6 branch lengths.

Across all 138 sites, variation of both TDα and FDα was explained by spatial and environmental predictors. PDα was not significantly explained by either (Table 1). TDα was better explained by environmental compared to spatial factors. These environmental factors were related to instream and landcover variables. Although a small proportion of the variation in TDα was attributed to purely spatial processes, the essential spatial predictors were derived from the weighted, tail-up, asymmetric models at a broad spatial scale (AEM_TU_DIST 1). The overall spatial and environmental model explained more variation in FDα than TDα, and partitioning suggested a predominance of purely spatial processes—opposite of TDα.

Fig. 4 Distributions of three facets of stream fish alpha diversity from streams in southeastern Oklahoma. The four major drainages are shown separately for comparative purposes. The distributions are plotted as density functions or 'ridges.' Points for each alpha value are shown and have been 'jittered' to create visual separation on the y-axis. Black lines represent the medians
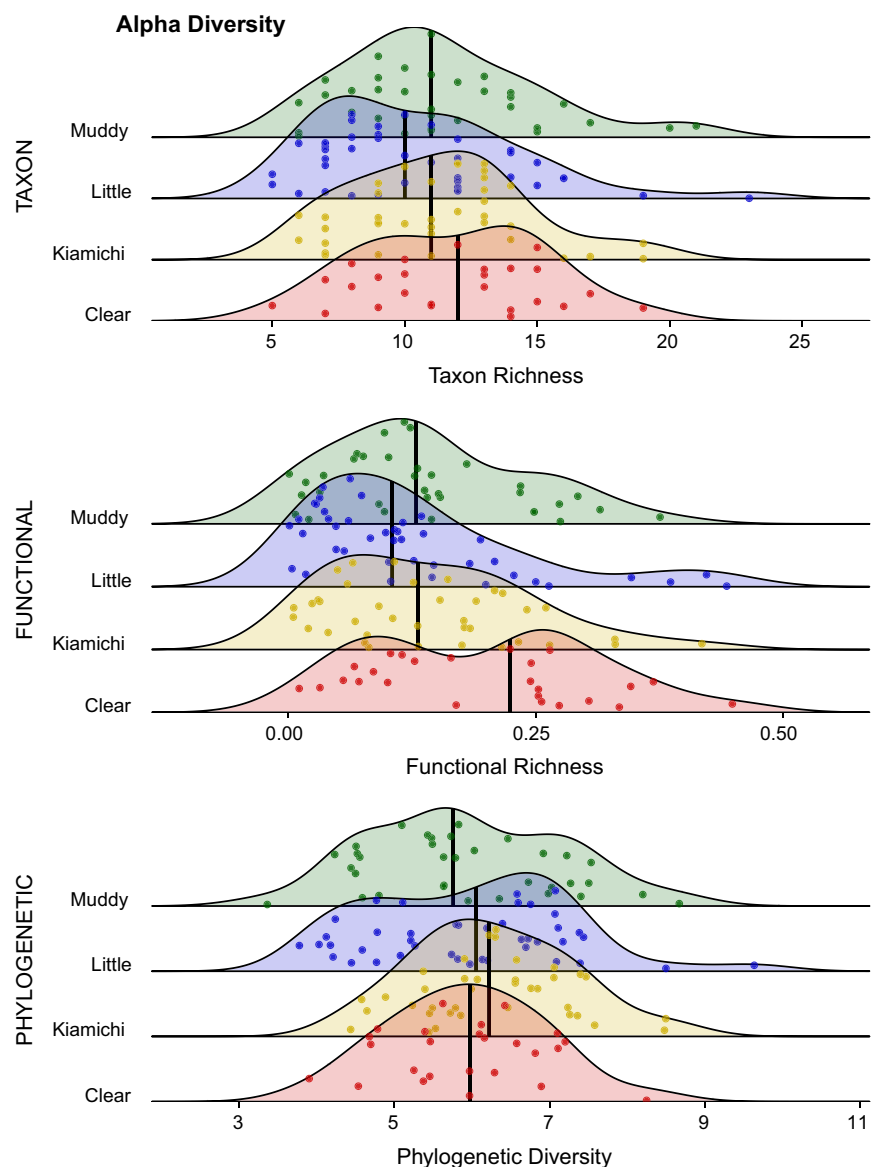
Table 1 Combined results of variation partitioning and forward model selection (FWD) of *alpha diversity* of stream fish collected in southeastern Oklahoma

	TDα	FDα	PDα
Variation partitioning			
Total (A + B + C)	0.14	0.24	–
Spatial (A + B)	0.09	0.13	–
Environ. (B + C)	0.12	0.11	–
Pure spatial (A)	0.02	0.13	–
Pure environ. (C)	0.05	0.11	–
Shared (B)	0.06	0.00	–
Environmental (FWD)			
Instream	0.08	0.07	–
Hydro-physio	–	–	–
Climate	–	0.04	–
Landcover	0.05	0.06	–
Geology	–	–	–
Anthropogenic	–	–	–
Spatial models (FWD)			
MEM_LAND	–	–	–
MEM_HYDRO	–	–	–
MEM_UPSTR_TOTAL	–	0.07	–
MEM_UPSTR_NET	–	0.11	–
AEM_TDN_UW	–	–	–
AEM_TDN_DAM	–	–	–
AEM_TDN_DIST	–	–	–
AEM_TUP_UW	–	–	–
AEM_TUP_DAM	0.07	–	–
AEM_TUP_DIST	0.09	–	–

Each facet of alpha diversity (taxon = TDα, functional = FDα, and phylogenetic = PDα) was analyzed separately and presented in each column. Column rows are divided into three sections: *Variation partitioning* results showing the variance explained by spatial and environmental components, forward model selection results comparing variance explained by the *six major environmental variable groups* (analyzed as principal components), and results of multiple linear regression of diversity data and ten *spatial models*. All table values are adjusted R^2, and only significant relationships are shown ($\alpha = 0.05$), except for variation partitioning row Shared (*B*), which is not a testable fraction

Again, instream and landcover variables were significant, but no significant spatial structure of environmental factors was involved (Table 1). The contributing spatial factors were derived from the upstream distance models (MEM_UPSTR_TOTAL) at broad and fine spatial scales (e.g., MEM3 to 32). All three alpha diversity facets were strongly correlated according to the Spearman's rank correlation (TDα v. FDα $\rho = 0.94$; TDα v. PDα $\rho = 0.84$; FDα v. PDα $\rho = 0.76$).

Alpha diversity within drainages

Models of alpha diversity within drainages yielded differing results (Table 2). For example, neither of the three facets (taxon, functional, phylogenetic) was significantly related to environmental or spatial factors in the Clear Boggy Drainage. Environmental factors accounted for some variation across all three facets within the Kiamichi Drainage, but the lack of significant spatial factors prevented variation partitioning. The opposite was true for TDα in the Muddy Boggy Drainage and PDα in the Little River Drainage where spatial factors could explain some variation, but not environmental ones. Both FDα and PDα were explained by environmental and spatial factors within the Muddy Boggy Drainage. For both, most of the variation was due to the combination of pure and spatially structured environments. A quarter of the variance was attributed to purely spatial factors derived from Euclidean distance models (MEM_LAND). Both TDα and FDα were explained by environmental and spatial factors within the Little River Drainage. However, a much more significant proportion of variance was explained purely by space (~80%). The contributing spatial models here were based on hydrologic distance (MEM_HYDRO) or asymmetric tail-down models, suggesting an influence of downstream diversity on upstream assemblages.

Beta diversity across the region

The four drainages showed similar levels of both pairwise (Fig. 5) and multi-site (Supplement S14) beta diversity across all facets. Replacement contributed more to overall dissimilarity than nestedness for all beta diversity facets, but the magnitude of the difference varied by facet (Fig. 5). Abundance-based TDβ had the highest pairwise dissimilarity overall, and PDβ had the least (Fig. 5). Distributions of incidence-based TDβ and FDβ were more similar overall, although that of FDβ was more skewed towards higher dissimilarity. Compared to incidence-based TDβ, FDβ showed lower levels of replacement and higher levels of nestedness (Fig. 5).

Table 2 Within drainage results of variation partitioning and forward model selection (FWD) of *alpha diversity* of stream fish collected in southeastern Oklahoma
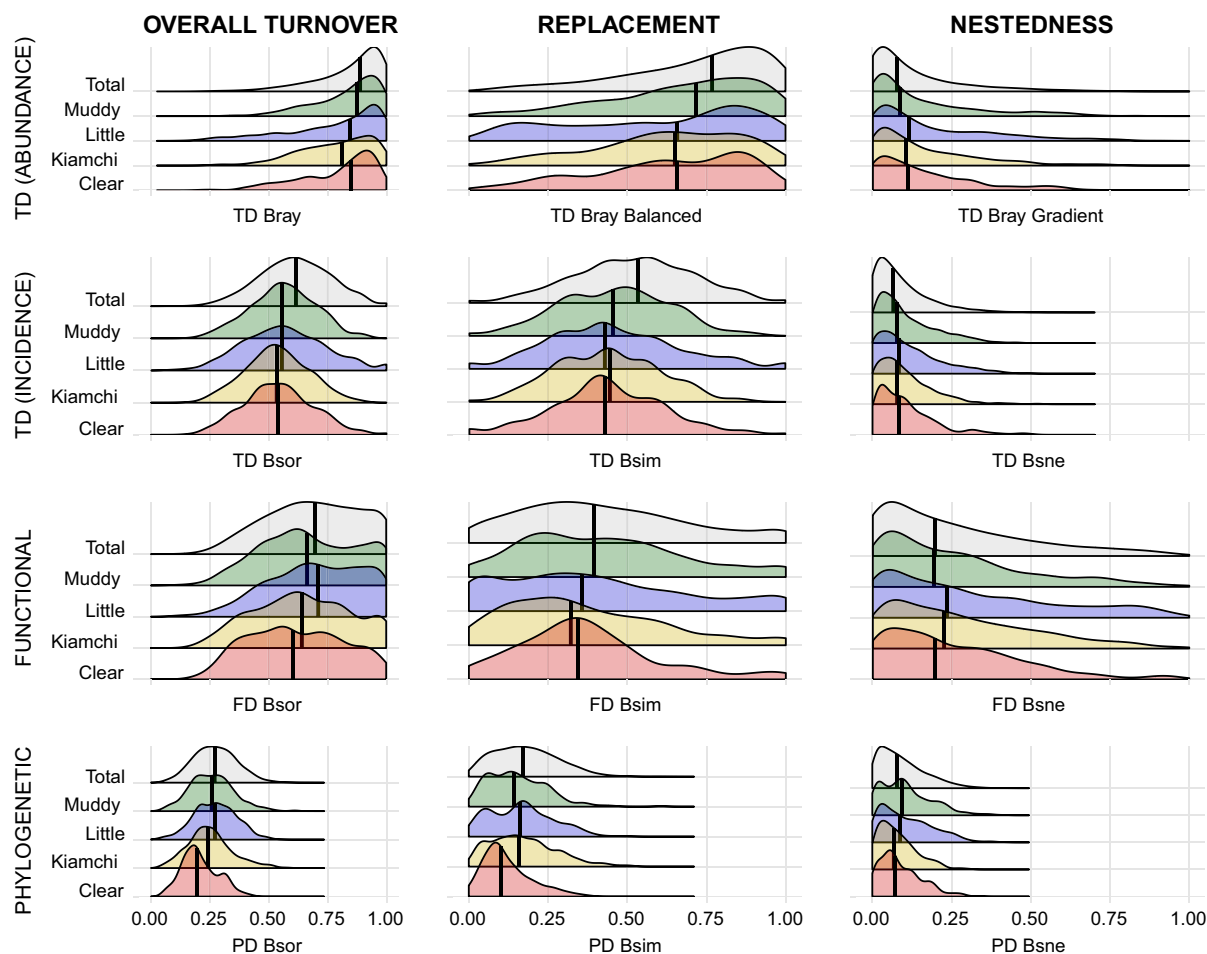
Alpha diversity

	Clear Boggy			Muddy Boggy			Kiamichi R			Little R		
	TDα	FDα	PDα	TDα	FDα	PDα	TDα	FDα	PDα	TDα	FDα	PDα
Variation partitioning												
Total (A + B + C)	–	–	–	–	0.48	0.43	–	–	–	0.28	0.37	–
Spatial (A + B)	–	–	–	–	0.41	0.28	–	–	–	0.24	0.34	–
Environ. (B + C)	–	–	–	–	0.36	0.36	–	–	–	0.06	0.05	–
Pure spatial (A)	–	–	–	–	0.12	0.10	–	–	–	0.22	0.32	–
Pure environ. (C)	–	–	–	–	0.08	0.15	–	–	–	0.04	0.03	–
Shared (B)	–	–	–	–	0.28	0.19	–	–	–	0.02	0.02	–
Environmental (FWD)												
Instream	–	–	–	–	–	–	0.08	–	–	0.06	0.05	–
Hydro-physio	–	–	–	–	–	–	0.18	0.27	0.09	–	–	–
Climate	–	–	–	–	–	–	–	–	–	–	–	–
Landcover	–	–	–	–	0.36	0.21	–	–	–	–	–	–
Geology	–	–	–	–	–	0.09	–	–	–	–	–	–
Anthropogenic	–	–	–	–	–	–	–	–	–	–	–	–
Spatial models (FWD)												
MEM_LAND	–	–	–	0.29	0.41	0.28	–	–	–	–	–	–
MEM_HYDRO	–	–	–	–	–	–	–	–	–	0.08	0.21	–
MEM_UPSTR_TOTAL	–	–	–	–	–	–	–	–	–	–	–	–
MEM_UPSTR_NET	–	–	–	–	–	–	–	–	–	–	–	–
AEM_TDN_UW	–	–	–	–	–	–	–	–	–	0.14	–	0.11
AEM_TDN_DAM	–	–	–	–	–	–	–	–	–	–	0.07	–
AEM_TDN_DIST	–	–	–	–	–	–	–	–	–	0.16	0.15	0.07
AEM_TUP_UW	–	–	–	0.13	–	–	–	–	–	–	–	–
AEM_TUP_DAM	–	–	–	–	–	–	–	–	–	–	–	–
AEM_TUP_DIST	–	–	–	–	–	–	–	–	–	–	–	–

Each facet of alpha diversity (taxon = TDα, functional = FDα, and phylogenetic = PDα) was analyzed separately and presented in each column. Rows are divided into three sections: *Variation partitioning* results showing the variance explained by spatial and environmental components, forward model selection results comparing variance explained by the *six major environmental variable groups* (analyzed as principal components), and results of multiple linear regression of diversity data and ten *spatial models*. All table values are adjusted R^2, and only significant relationships are shown ($\alpha = 0.05$)

Across the region and all facets, spatial and environmental predictors better explained replacement than overall dissimilarity, based on adjusted R^2, while unable to explain nestedness (Table 3). While models could account for variance in facets differently, the proportion of total explained variance for each facet was attributed to largely the same environmental factors and specific spatial models. Across all facets, variance explained by spatial factors (A + B; ~80% total variance explained) was higher than environmental factors (B + C; ~70% total variance explained). Similarly, the variance explained purely by space (A; ~33% total variance) was invariably higher than that explained purely by environment (C; ~20% total variance explained). Although spatially structured environment (B; ~45% total variance explained) was greater than pure space. Generally, around half of the variance explained by the total model (A + B + C) was shared between space and environment (B; Table 3), suggesting that much of the environmental factors related to beta diversity were spatially structured. One exception was the model of FDβ$_{sim}$, where the total model explained 40% of the variance and the shared variance explained between space and environment

Fig. 5 Distributions of beta diversity facets (taxon abundance based, taxon incidence based, functional, and phylogenetic) of stream fish assemblages from streams in southeastern Oklahoma. The overall dissimilarity of each facet is further partitioned into additive components of replacement and nest-edness. The four major drainages are shown separately for comparative purposes. The distributions are plotted as density functions or 'ridges.' Black lines represent the medians. All plots have the same *x*-axis range (0 = most similar to 1 = least similar)

(B) was 9%. Models of incidence-based TDβ and PDβ explained the most variance, followed by those of abundance-based TDβ and FDβ.

Across the region, all six environmental variable types explained significant amounts of variation of beta diversity for all facets and beta components (except nestedness). There were slight differences in the proportions of variance explained for each environmental factor category and spatial model across facets. Instream variables were consistently ranked highest (by adjusted R^2) in 7 of 8 models—excluding nestedness—accounting for an average of 43% of total variance explained (Table 3). Landcover and

geology-related variables were generally the next best at explaining variation in beta diversity, followed by climate, hydrologic and physiographic, and anthropogenic factors (Table 3).

The proportion of variance attributed to different spatial models was consistent across facets for the regional models. Across the region, all ten spatial models explained significant amounts of variation of beta diversity for all facets and components (except nestedness). The spatial model based on overland Euclidean distance among sites (MEM_LAND) was consistently ranked highest (7/8 models) on average, accounting for about 76% of the total variance

Table 3 Steam fish assemblages from southeastern Oklahoma were studied by partitioning multiple facets of beta diversity (taxon, functional, and phylogenetic) into components of dissimilarity (overall, replacement, and nestedness)

	TDβ Abundance based			TDβ Incidence based			FDβ			PDβ		
	Bray	BrayBal	BrayGra	βsor	βsim	βsne	βsor	βsim	βsne	βsor	βsim	βsne
	R^2 (adj.)			R^2 (adj.)			R^2 (adj.)			R^2 (adj.)		
Variation partitioning												
Total (A + B + C)	0.28	0.50	–	0.47	0.59	–	0.17	0.40	–	0.41	0.65	–
Spatial (A + B)	0.22	0.39	–	0.37	0.47	–	0.14	0.25	–	0.32	0.51	–
Environ. (B + C)	0.19	0.35	–	0.33	0.43	–	0.11	0.24	–	0.27	0.45	–
Pure spatial (A)	0.09	0.15	–	0.14	0.16	–	0.06	0.16	–	0.13	0.20	–
Pure environ. (C)	0.06	0.11	–	0.10	0.12	–	0.03	0.15	–	0.08	0.14	–
Shared (B)	0.12	0.24	–	0.23	0.31	–	0.08	0.09	–	0.19	0.31	–
Environmental (FWD)												
Instream	0.09	0.20	–	0.20	0.28	–	0.08	0.23	–	0.16	0.27	–
Hydro-physio	0.06	0.13	–	0.10	0.13	–	0.02	0.09	–	0.06	0.10	–
Climate	0.08	0.14	–	0.15	0.20	–	0.06	0.14	–	0.11	0.17	–
Landcover	0.10	0.19	–	0.15	0.20	–	0.06	0.11	–	0.14	0.23	–
Geology	0.08	0.15	–	0.16	0.22	–	0.07	0.17	–	0.13	0.21	–
Anthropogenic	0.04	0.08	–	0.06	0.09	–	0.03	0.06	–	0.06	0.10	–
Spatial models (FWD)												
MEM_LAND	0.20	0.41	–	0.40	0.59	–	0.10	0.27	–	0.30	0.48	–
MEM_HYDRO	0.08	0.15	–	0.14	0.19	–	0.04	0.10	–	0.12	0.17	–
MEM_UPSTR_TOTAL	0.12	0.32	–	0.24	0.32	–	0.11	0.20	–	0.23	0.38	–
MEM_UPSTR_NET	0.07	0.15	–	0.17	0.23	–	0.05	0.17	–	0.15	0.27	–
AEM_TDN_UW	0.11	0.24	–	0.22	0.31	–	0.07	0.16	–	0.18	0.35	–
AEM_TDN_DAM	0.12	0.24	–	0.26	0.38	–	0.08	0.21	–	0.18	0.34	–
AEM_TDN_DIST	0.13	0.27	–	0.25	0.36	–	0.09	0.26	–	0.19	0.34	–
AEM_TUP_UW	0.09	0.18	–	0.14	0.20	–	–	–	–	0.11	0.19	–
AEM_TUP_DAM	0.08	0.13	–	0.11	0.12	–	0.02	–	–	0.09	0.13	–
AEM_TUP_DIST	0.07	0.14	–	0.09	0.11	–	–	–	–	0.08	0.11	–

For each facet (×2 for taxon because both abundance-based and incidence-based data were analyzed), results for the variation partitioning analysis, forward selection of environmental variables, and forward selection of spatial variables are shown. All values in the table are adjusted R^2. Bray = Bray–Curtis dissimilarity. BrayBal = balanced variation in abundance and BrayGra = abundance gradients (akin to replacement and nestedness for abundance data). Bsor = overall dissimilarity, βsim = replacement component, and βsne = nestedness component

explained. The following best models depended on facet and component, but generally was either Model 3 (MEM_UPSTR_TOTAL; ~55% total variance explained) or the weighted tail-down asymmetric models (Model 6, AEM_TDN_DAM, ~50% total variance explained; or Model 7, AEM_TDN_DIST, ~54%). Although tail-up models consistently explained significant amounts of variation in beta diversity, they always explained less than the tail-down counterparts. Strikingly, the commonly applied Model 2 (MEM_HYDRO) was consistently among the worst-performing models (based on adjusted R^2).

A complete list of significant spatial and environmental factors is available in the Supplementary Material (S15).

Beta diversity within drainages

Analyses of each drainage separately also produced more mixed results (Table 4). Models explained more variation in beta diversity for the Muddy Boggy and Little River drainages than the Clear Boggy and Kiamichi. The replacement component of overall dissimilarity was always better explained (as with the

Table 4 Steam fish assemblages from southeastern Oklahoma were studied by partitioning multiple facets of beta diversity (taxon, functional, and phylogenetic) into components of dissimilarity (overall, replacement, and nestedness)

		TDβ-Abund			TDβ-Incid			FDβ			PDβ		
		Br	Br.B	Br.G	βsor	βsim	βsne	βsor	βsim	βsne	βsor	βsim	βsne
Clear Boggy													
VarPart	Total (A+B+C)	0.20	0.34	–	0.17	0.21	–	–	–	–	–	–	–
	Spatial (A+B)	0.11	0.19	–	0.05	0.06	–	–	–	–	–	–	–
	Environ. (B+C)	0.14	0.27	–	0.16	0.19	–	–	–	–	–	–	–
	Pure spatial (A)	0.06	0.07	–	0.01	0.02	–	–	–	–	–	–	–
	Pure environ. (C)	0.09	0.15	–	0.12	0.15	–	–	–	–	–	–	–
	Shared (B)	0.05	0.12	–	0.04	0.04	–	–	–	–	–	–	–
ENV FWD	Instream	–	0.10	–	0.10	0.13	–	–	–	–	–	–	–
	Hydro-physio	0.13	0.17	–	0.15	0.21	–	–	–	–	–	–	–
	Climate	0.10	0.12	–	–	0.09	–	–	–	–	–	–	–
	Landcover	0.05	–	–	–	–	–	–	–	–	–	–	–
	Geology	0.08	0.18	–	–	–	–	–	0.13	–	–	–	–
	Anthropogenic	–	–	–	–	–	–	–	0.15	–	0.12	0.09	–
Spatail FWD	MEM_LAND	0.07	0.20	–	–	–	–	–	–	–	–	–	–
	MEM_HYDRO	0.05	0.19	–	–	–	–	–	–	–	–	–	–
	MEM_UPSTR_TOTAL	0.04	0.09	–	0.05	0.09	–	–	–	–	–	–	–
	MEM_UPSTR_NET	0.04	0.08	–	–	–	–	–	–	–	–	–	–
	AEM_TDN_UW	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TDN_DAM	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TDN_DIST	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TUP_UW	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TUP_DAM	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TUP_DIST	–	–	0.38	–	–	–	–	–	–	–	–	–
Kiamichi River													
VarPart	Total (A+B+C)	0.13	0.24	–	0.17	0.22	–	0.04	–	–	0.21	0.27	–
	Spatial (A+B)	0.11	0.18	–	0.12	0.18	–	0.02	–	–	0.19	0.26	–
	Environ. (B+C)	0.09	0.18	–	0.09	0.13	–	0.02	–	–	0.11	0.15	–
	Pure spatial (A)	0.05	0.06	–	0.08	0.09	–	0.02	–	–	0.10	0.12	–
	Pure environ. (C)	0.02	0.06	–	0.05	0.04	–	0.02	–	–	0.02	0.01	–
	Shared (B)	0.07	0.12	–	0.05	0.09	–	0.00	–	–	0.09	0.14	–
ENV FWD	Instream	0.05	0.08	–	0.06	0.09	–	–	–	–	0.09	0.16	–
	Hydro-physio	0.06	–	0.31	0.04	0.06	–	0.02	–	–	0.07	–	–
	Climate	–	–	0.33	0.05	–	–	–	–	–	–	–	–
	Landcover	0.09	0.07	–	0.04	–	–	–	–	–	–	–	–
	Geology	–	0.06	–	0.07	0.10	–	–	0.08	–	–	0.07	–
	Anthropogenic	–	0.05	–	–	–	–	–	–	–	–	–	0.17

Table 4 (continued)

		TDβ-Abund			TDβ-Incid			FDβ			PDβ		
		Br	Br.B	Br.G	βsor	βsim	βsne	βsor	βsim	βsne	βsor	βsim	βsne
Spatail FWD	MEM_LAND	0.07	–	–	0.04	–	–	–	–	–	–	0.08	–
	MEM_HYDRO	0.11	0.21	–	0.13	0.21	–	0.03	–	–	0.16	0.35	–
	MEM_UPSTR_TOTAL	–	–	–	–	–	–	–	–	–	–	–	–
	MEM_UPSTR_NET	0.02	–	–	–	–	–	–	–	–	–	–	–
	AEM_TDN_UW	–	–	–	–	–	–	–	–	–	0.05	0.09	–
	AEM_TDN_DAM	–	–	–	–	–	–	–	–	–	0.05	0.09	–
	AEM_TDN_DIST	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TUP_UW	–	–	–	–	–	–	–	–	–	0.09	0.12	–
	AEM_TUP_DAM	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TUP_DIST	–	–	–	–	–	–	–	–	–	–	–	–
		TDβ-Abund			TDβ-Incid			FDβ			PDβ		
		Br	Br.B	Br.G	βsor	βsim	βsne	βsor	βsim	βsne	βsor	βsim	βsne
Muddy Boggy													
VarPart	Total (A + B + C)	0.17	0.41	–	0.39	0.42	0.53	0.23	0.40	–	0.39	0.55	0.30
	Spatial (A + B)	0.08	0.33	–	0.28	0.30	0.27	0.11	0.21	–	0.16	0.12	0.12
	Environ. (B + C)	0.15	0.28	–	0.30	0.31	0.53	0.18	0.40	–	0.34	0.43	0.23
	Pure spatial (A)	0.03	0.13	–	0.09	0.10	0.26	0.05	–	–	0.05	0.12	0.06
	Pure environ. (C)	0.10	0.08	–	0.11	0.11	0.09	0.13	0.19	–	0.23	0.43	0.17
	Shared (B)	0.05	0.08	–	0.19	0.20	0.18	0.05	0.21	–	0.11	0.00	0.06
ENV FWD	Instream	0.10	0.21	–	0.17	0.20	–	0.05	–	–	0.19	0.27	–
	Hydro-physio	0.04	0.10	–	0.07	0.10	–	0.02	0.09	–	0.05	0.21	–
	Climate	0.02	0.10	–	0.05	0.06	–	–	–	–	0.07	0.13	–
	Landcover	0.06	0.12	–	0.05	–	0.37	0.07	0.41	0.29	0.18	0.25	0.25
	Geology	0.02	0.05	–	0.06	0.09	–	0.05	0.12	–	0.05	–	–
	Anthropogenic	–	–	–	–	–	–	–	–	–	–	0.22	–
Spatail FWD	MEM_LAND	0.05	0.20	–	0.05	0.15	0.51	0.06	–	–	0.11	–	0.13
	MEM_HYDRO	–	0.19	–	0.23	0.22	–	–	–	–	0.05	–	–
	MEM_UPSTR_TOTAL	–	–	–	–	–	–	–	–	–	–	–	–
	MEM_UPSTR_NET	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TDN_UW	0.03	0.08	–	0.06	0.09	–	–	–	–	0.06	0.12	–
	AEM_TDN_DAM	0.03	0.08	–	0.06	0.09	–	–	–	–	0.05	0.12	–
	AEM_TDN_DIST	–	–	–	–	–	–	–	–	–	-	–	–
	AEM_TUP_UW	–	0.11	–	0.10	0.13	–	–	–	–	–	–	–
	AEM_TUP_DAM	–	0.06	–	–	–	–	–	–	–	–	–	–
	AEM_TUP_DIST	0.02	–	–	–	–	–	0.04	0.26	–	0.07	–	–

regional analysis). Generally, the nestedness component was either not explained by space or environment (9/16) or only explained by one: space (2/16) or environment (3/16). However, incidence-based TDβ and PDβ nestedness components (2/16) were explained by both environment and space for the Muddy Boggy Drainage. Here, the nestedness component $TD\beta_{sne}$ was primarily explained by purely spatial factors and, to a lesser extent, by spatially structured environment and pure environment. However, the nestedness component of PDβ was explained mainly by the pure and spatially structured environment.

Table 4 (continued)

		TDβ-Abund			TDβ-Incid			FDβ			PDβ		
		Br	Br.B	Br.G	βsor	βsim	βsne	βsor	βsim	βsne	βsor	βsim	βsne
Little River													
VarPart	Total (A + B + C)	0.38	0.66	–	0.46	0.64	–	0.17	0.47	–	0.24	0.40	–
	Spatial (A + B)	0.20	0.41	–	0.25	0.31	–	0.10	0.28	–	0.24	0.26	–
	Environ. (B + C)	0.28	0.53	–	0.40	0.57	–	0.09	0.30	–	0.24	0.38	–
	Pure spatial (A)	0.09	0.13	–	0.06	0.07	–	0.08	0.17	–	0.00	0.01	–
	Pure environ. (C)	0.17	0.24	–	0.21	0.33	–	0.07	0.20	–	0.06	0.14	–
	Shared (B)	0.11	0.28	–	0.20	0.24	–	0.02	0.10	–	0.18	0.24	–
ENV FWD	Instream	0.15	0.32	–	0.26	0.38	–	0.07	0.34	–	0.09	0.14	–
	Hydro-physio	0.12	0.30	–	0.21	0.38	–	0.06	0.21	–	0.07	0.14	–
	Climate	0.12	0.30	–	0.22	0.32	–	0.04	0.15	–	0.11	0.15	–
	Landcover	0.14	0.31	–	0.28	0.40	–	0.06	0.24	–	0.11	0.21	–
	Geology	0.15	0.30	–	0.26	0.36	–	0.08	0.28	–	0.14	0.21	–
	Anthropogenic	0.08	0.24	–	0.15	0.23	–	0.02	0.08	–	0.10	0.16	–
Spatail FWD	MEM_LAND	0.17	0.33	–	0.29	0.43	–	0.09	–	–	0.22	0.31	–
	MEM_HYDRO	0.08	0.32	–	0.31	0.41	–	0.11	0.39	–	0.15	0.12	–
	MEM_UPSTR_TOTAL	–	–	–	–	–	–	–	–	–	–	–	–
	MEM_UPSTR_NET	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TDN_UW	0.07	0.17	–	0.14	0.20	–	0.06	0.19	–	0.15	0.17	–
	AEM_TDN_DAM	0.05	0.14	–	0.11	0.16	–	0.05	0.15	–	0.08	0.13	–
	AEM_TDN_DIST	0.08	0.19	-	0.14	0.21	–	0.07	0.19	–	0.11	0.16	–
	AEM_TUP_UW	0.03	0.06	0.28	–	–	–	–	–	–	–	–	–
	AEM_TUP_DAM	–	–	–	–	–	–	–	–	–	–	–	–
	AEM_TUP_DIST	–	0.05	–	–	–	–	–	–	–	–	–	–

Results are presented here from each of four drainages analyzed separately. For each facet (×2 for taxon because both abundance-based and incidence-based data were analyzed) results for the variation partitioning analysis, forward selection of environmental variables, and forward selection of spatial variables are shown. All values in the table are adjusted R^2. Br. = Bray–Curtis dissimilarity. Br.B = balanced variation in abundance and Br.G = abundance gradients (akin to replacement and nestedness for abundance data). βsor = Overall dissimilarity, βsim = replacement component, and βsne = nestedness component

Both facets based on taxon diversity (abundance and incidence based) were explained by environmental and spatial factors for all drainages. Analyses of individual drainages showed that purely environmental factors generally explained more variation in assemblage composition than purely spatial factors (opposite of that found above across the region). However, this was not true for the Kiamichi Drainage, where pure space was more important than pure environment. FDβ was not successfully modeled in either the Clear Boggy or Kiamichi drainages, and the same was true for PDβ in the Clear Boggy.

Among drainages, there was variation in which environmental factors explained the most variation in assemblage composition (Table 4). Generally, instream and landscape-scale landcover factors were among the most important (adjusted R^2). Differences in the important environmental attributes were more apparent among drainages than among facets. Fewer differences were observed among drainages regarding important spatial models. Euclidean distance (MEM_LAND) often explained a high amount of variation (as with the regional scale). The hydrologic distance model (MEM_HYDRO) did relatively better at explaining variation at this spatial scale compared to the regional scale. Asymmetric tail-down models often explained about half as much variance as Euclidean distance. Depending on the drainage, asymmetric tail-up models also explained significant compositional variance.

All 12 pairwise beta diversity matrices were significantly correlated, some very highly so (Supplement

S16). The highest correlations tended to be among components within facets (e.g., overall, replacement, nestedness). This is expected, as replacement and nestedness are components of overall dissimilarity.

Discussion

Despite correlation among diversity facets, they did provide complementary information. For example, variation partitioning models of FDβ and FDα showed uniquely low contribution of spatially autocorrelated environmental factors, suggesting higher levels of distinctly spatial (e.g., dispersal limitation) and environmental processes (e.g., related to stream hierarchy) may govern these facets. Furthermore, despite general patterns emerging related to which specific environmental and spatial factors explained the highest variance, there were differences in the ability of models to explain different facets. Moreover, across all facets, the combined spatial and environmental model of PDβ$_{sim}$ (replacement) explained the highest amount of variation (adjusted $R^2 = 0.65$). The total variance explained by the combined model varied substantially among facets and components.

The degree to which additional facets of diversity provide complementary information may depend on several factors, including location (Burgad et al., 2019), latitude (Stevens & Tello, 2018), spatial scale (Legendre, 2014), and phylogenetic scale (Branco et al., 2020), all of which are naturally correlated with expected variation of the facets. Functional diversity, for example, relies on both the phylogenetic scale of the study and the traits used to measure those functions (Mammola et al., 2021). Functional and phylogenetic diversity might provide more in-depth insight into whole community data with greater functional and phylogenetic scale (i.e., not constrained only to one group—fish). For a particular study system, it likely cannot be known how valuable additional insights gleaned from multiple facets will be until they are quantified and modeled (Heino & Tolonen, 2017). For example, the explanatory power of abiotic factors for different diversity facets varied considerably for each drainage despite being adjacent and in the same basin; this supports the notion that general patterns related to different diversity facets may not be generalizable.

Alpha diversity

Across the entire region and within drainages, environmental and spatial predictors showed a disparity in explaining variation among facets of alpha diversity. At the region scale, PDα could not be explained by either space or environment, suggesting that other factors such as stochasticity, historical contingency, or biotic interactions may be of greater importance for determining PDα (Li et al., 2020; Morales-Castilla et al., 2020). At the drainage scale, we recovered this same pattern except for the Muddy Boggy, where spatial and environmental factors explained PDα. The Muddy Boggy drainage encompasses both upstream highland and lowland areas known to differ in habitat and, consequently, fish assemblage structure (Pigg, 1977; Zbinden & Matthews, 2017). The same pattern is apparent with FDα in the Muddy Boggy drainage but not with TDα, which was not explained by either space or environment. These findings suggest that the upland and lowland areas differ concerning functional and phylogenetic diversity (different kinds of species present yet similar richness) due to longitudinal zonation (McGarvey & Hughes, 2008; Stegmann et al., 2019) and that phylogenetic and functional diversity can provide additional information beyond TD (Saito et al., 2015).

Across the region, models accounted for more variation in FDα compared to TDα. Also, these models revealed a predominance of broad- and fine-scale spatial factors based on the degree of upstream distance between sites that governs variability in functional diversity. This suggests communities become less similar in terms of FDα when greater upstream separation presumably limits dispersal from functionally richer downstream locations to more isolated headwater communities (Carvalho & Tejerina-Garro, 2015; Caetano et al., 2021). While environmental factors that account for spatial and temporal heterogeneity also drive variation in FDα among communities, these factors do not appear to be spatially structured. We would expect them to be spatially structured if diversity differences were due primarily to longitudinal changes in the environment (upstream–downstream gradient) rather than dispersal limitation.

Variation in TDα was not well explained within drainages either. For the Muddy Boggy and Little drainages, FDα was the best explained facet of diversity, but FDα was not related to spatial and

environmental factors in the other drainages studied. This difference among drainages is likely responsible for the lower correlation coefficients observed for FDα across the entire region. Functional diversity (like PDα) in the Muddy Boggy seems partly governed by environmental differences between upland and lowland areas. For the Little River, we observed a more predominant role of purely spatial factors, whereas spatially structured environmental factors had minimal influence (similar to the regional model). Hydrologic distance and tail-down processes were the essential spatial factors here; this suggests FDα in the Little River is determined by dispersal limitation, particularly dispersal from downstream to upstream assemblages. If this were the case, we would expect to see greater amounts of nestedness in the Little River, which we found (Fig. 5).

Beta diversity

Overall, dissimilarity among assemblages was generally high. This highlights the importance of inter-site differentiation in maintaining total diversity (i.e., gamma) within drainages (Whittaker, 1960; Branco et al., 2020), and particularly for headwater streams (stream order ≤ 3), which make up most reaches studied here (115/138 sites, Supplement S17). Headwater stream reaches are generally less speciose than downstream reaches, yet headwater streams are more variable among reaches in terms of species composition (Finn et al., 2011).

We found replacement was the more prominent component of overall dissimilarity between assemblages across all drainages and facets of diversity. Note, however that levels of nestedness were relatively higher for FDβ and PDβ than TDβ, as has been reported elsewhere (Heino & Tolonen, 2017; Branco et al., 2020). For PDβ, levels of replacement were still greater but more similar to levels of nestedness. A dominant role of replacement—sometimes referred to as spatial turnover—has been observed for many different systems across the globe (Soininen et al., 2018; Branco et al., 2020). This suggests that compositional differences between assemblages result from species replacing one another rather than simple gain and loss between sites (Baselga, 2010). Therefore, we expect to find evidence for environmental and spatial filtering driving that replacement (Boschilia et al., 2016). Indeed, spatio-environmental models accounted for

significant variation in diversity, with up to 65 percent of variation explained. This suggests that community composition is explained by environmental and spatial filtering. Spatially structured environmental variables at the landscape and instream scale played a part in explaining beta diversity as expected (Marsh-Matthews & Matthews, 2000); however, 'purely' environmental factors (i.e., not spatially structured) were predominately associated with instream and hydrophysiological factors. Additionally, we found consistently high levels of replacement at both spatial scales. López-Delgado et al. (2020) found the same pattern across scales. They suggested this is consistent with dispersal being limited enough between assemblages to not overwhelm the process of habitat filtering (i.e., species sorting). Indeed, most sampling locations in this study were far enough apart to likely limit dispersal for the primarily small-bodied species collected here.

There was high variability in explanatory power at the drainage scale across different facets of beta diversity (and across drainages). As with alpha diversity above, it appeared that fish diversity in the Muddy Boggy and Little River drainages was more predictable than that in the Clear Boggy and Little River drainages (i.e., less stochastic). The best explained facet of beta diversity differed among drainages. We observed that environmental factors explained more variation in facets of diversity within Clear and Muddy Boggy drainages. Environmental factors were also important in the Little River, but space and environment were more similar for some facets. For the Kiamichi drainage, however, spatial factors were more important explainers of variation in diversity. Again, instream and hydrological/physiographic factors played a consistent explanatory role, but variability was detected among facets and drainages regarding which sets of spatial factors explained the most variation. The overland distance was consistently able to explain variation at the within-drainage scale but less so than at the regional scale. Essential contributions to variation explained within drainages were also attributed to hydrologic distance and asymmetric tail-down models.

Importance of spatial modeling

Explicitly considering different spatial scales and testing different spatial models may be more valuable

than investigating multiple facets of diversity. Realistic spatial models may show that dispersal limitation is more important than previously thought. For example, our models attributed more diversity variation to purely spatial factors at the largest spatial scale assessed (total region) than the smaller spatial scale (within basins). Studies of stream community variation often demonstrate a small or non-significant relationship with purely spatial factors (e.g., Zbinden & Matthews, 2017; Huang et al., 2019; López-Delgado et al., 2019; Benone et al., 2020). This may be partly due to the small spatial scales analyzed and the use of only river network distance for spatial modeling. We predict minor dispersal limitation and spatially structured environmental heterogeneity at smaller spatial scales (Jackson et al., 2001; Cottenie, 2005).

Furthermore, river network distance alone does not always adequately account for the complex spatial structure of river networks. More realistic models may be necessary to adequately capture the effects of spatial processes (Mozzaquattro et al., 2020). The spatial models used here are not mutually exclusive (as environmental variables are not), but the contribution of each can help make inferences about the type of spatial processes underlying variation in assemblage structure. Spatially structured environmental variables often contribute heavily to the total variance in diversity explained (Leprieur et al., 2009). This is expected across most spatially structured riverscapes because nearest-neighboring locations are usually similar in environmental conditions. This often results in substantial diversity overlap (between 'close' sites) due to environmental filtering/species sorting processes.

It should be noted that diversity variation explained by purely spatial factors is typically used to infer the relative influence of dispersal limitation. However, any unaccounted-for environmental variation that happens to be spatially structured can also create a pattern of what appears to be purely spatial variation (Landeiro et al., 2011). This potential issue was drastically minimized in the present study because of the extensive set of environmental factors considered (Heino & Tolonen, 2017).

Spatio-environmental factors that govern the diversity of stream communities are spatially represented across both the riverscape and across the landscape (Peterson et al., 2013). Due to this dual representation, studies examining the spatial autocorrelation

of stream assemblages should consider geographic and network distances to better understand processes that may occur at different spatial scales (i.e., across drainages or catchments and along the stream network). While spatial eigenvectors based on watercourse distance are expected to depict dispersal limitation more accurately, the overland distance may also explain significant amounts of compositional variation (Landeiro et al., 2011). In our study region, we expected overland distance (MEM_LAND) to explain assemblage variation due to autocorrelation between diversity and drainage-specific environmental variables (i.e., spatially structured environmental filtering). Indeed, this model consistently recovered the strongest signal; this suggests that streams in close proximity (i.e., short overland Euclidean distance) tend to have similar environments and similar fish assemblages even if separated by large network distances. This signal was strongest at the largest spatial scale assessed (entire region), and it became less apparent when drainages were analyzed individually. Thus, overland distance is likely more indicative of spatial autocorrelation due to spatially structured environmental variation, while watercourse distances indicate dispersal limitation.

Additionally, our findings demonstrate that downstream and upstream dispersal limitation are not equal, as evidenced by the significant variation in beta diversity and alpha diversity explained by the total upstream distance (Model 3, MEM_UPSTR_TOTAL) compared to that of symmetric network distance (Model 2, MEM_HYDRO). Furthermore, our asymmetric spatial models indicate neighboring assemblages can affect one another via migration and immigration in more of a tail-down manner—immigrants moving upstream possibly due to mass effects or source-sink dynamics (Leibold et al., 2004). However, tail-up models were often significant as well, suggesting that downstream migration is also important. On a related note, Hydro-physio PC1 was associated with network position, and one of the strongest variable loadings was the river volume in the catchment upstream of a site. Increased volume is associated with the size and number of tributary streams in the catchment above a reach and is presumably associated with greater potential for colonizers. Understandably, the presence of dams along the stream network can alter dispersal in both directions, as evidenced by the consistent significant signal of both

tail-up and tail-down models weighted by the presence of dams. Moreover, the dam-weighted models explained as much, or more, variance in fish assemblages as the distance-weighted models, thus, providing further evidence of how barriers can impact connectivity among metacommunities (Shao et al., 2019; Mozzaquattro et al., 2020).

Conclusion

Determining the extent of additional insight gained by including novel diversity facets may not be straightforward. We found differences in the utility of additional diversity facets even among neighboring drainages within the same latitude. This may be related to differences in environmental variation and ecosystem complexity among the drainages assessed here (Maasri et al., 2021; Terui et al., 2021). The consensus across facets was that environmental and spatial factors play a fundamental role in governing fish assemblage structure, reflecting the predominance of environmental filtering and dispersal limitation. Assemblage structure is mainly the result of species replacement, particularly so for headwater streams. This suggests that tributary streams might be managed individually based on unique habitat and assemblages (Lowe & Likens, 2005; Pracheil et al., 2013). However, we did find evidence of upstream dispersal affecting assemblage variation. Therefore, managing high diversity downstream sites can help maintain diversity across a metacommunity (Kanno et al., 2012; Zbinden et al., 2022).

Local, instream factors were critical for determining reach-scale fish diversity. There may be no good landscape-scale substitute for the variables that describe heterogeneity in micro-habitat (e.g., substrate, riffles, pools). These data are more challenging to gather than landscape-level data (i.e., available on the web), but they are indispensable (Montaña & Winemiller, 2010; Blanchet et al., 2014). Furthermore, evidence that instream heterogeneity governs species assembly suggests managers should focus on protecting areas that encompass diverse habitats (López‑Delgado et al., 2020). Ideally, protected areas would encompass the full gradient of instream heterogeneity found within a region (Roa-Fuentes et al., 2019).

In riverine networks, testing spatial autocorrelation models based on multiple distance-based approaches can reveal how processes like dispersal limitation determine assemblage structure. Directional models are likely to estimate these effects better than symmetric models (i.e., simple network distance). Disentangling the factors that control community assembly is vital for successfully conserving and managing aquatic biodiversity (Jackson et al., 2001; Chen & Olden, 2020; Maasri et al., 2021). Our capacity to conserve biodiversity relies on understanding how and why it changes across space and time.

Data, material and code availability All data and R code available at https://github.com/zdzbinden/OK_fish_diversity. Fish collections are housed at the Sam Noble Oklahoma Museum of Natural History, Norman, OK.

Declarations

Conflict of interest The authors have no conflicts of interest or competing interests to disclose.

Ethical approval Fish collections were approved by the Institutional Animal Care and Use Committee at the University of Oklahoma and permitted by the Oklahoma Department of Wildlife Conservation.

References

Anderson, M. J. & P. Legendre, 1999. An empirical comparison of permutation methods for tests of partial

regression coefficients in a linear model. Journal of Statistical Computation and Simulation 62(3): 271–303.

Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography 19(1): 134–143.

Baselga, A., 2013. Separating the two components of abundance-based dissimilarity: balanced changes in abundance vs. abundance gradients. Methods in Ecology and Evolution 4(6): 552–557.

Baselga, A. & C. D. L. Orme, 2012. betapart: an R package for the study of beta diversity. Methods in Ecology and Evolution 3(5): 808–812.

Bauman, D., T. Drouet, S. Dray & J. Vleminckx, 2018. Disentangling good from bad practices in the selection of spatial or phylogenetic eigenvectors. Ecography 41(10): 1638–1649.

Benone, N. L., C. G. Leal, L. L. dos Santos, T. P. Mendes, J. Heino & L. F. de Assis Montag, 2020. Unravelling patterns of taxonomic and functional diversity of Amazon stream fish. Aquatic Sciences 82(4): 1–11.

Blanchet, F. G., P. Legendre & D. Borcard, 2008a. Modelling directional spatial processes in ecological data. Ecological Modelling 215(4): 325–336.

Blanchet, F. G., P. Legendre & D. Borcard, 2008b. Forward selection of explanatory variables. Ecology 89(9): 2623–2632.

Blanchet, F. G., P. Legendre, R. Maranger, D. Monti & P. Pepin, 2011. Modelling the effect of directional spatial ecological processes at different scales. Oecologia 166(2): 357–368.

Blanchet, S., M. R. Helmus, S. Brosse & G. Grenouillet, 2014. Regional vs local drivers of phylogenetic and species diversity in stream fish communities. Freshwater Biology 59(3): 450–462.

Borcard, D. & P. Legendre, 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153(1–2): 51–68.

Boschilia, S. M., E. F. de Oliveira & A. Schwarzbold, 2016. Partitioning beta diversity of aquatic macrophyte assemblages in a large subtropical reservoir: prevalence of turnover or nestedness? Aquatic Sciences 78(3): 615–625.

Branco, C. C., P. C. Bispo, C. K. Peres, A. F. Tonetto, R. A. Krupek, M. Barfield & R. D. Holt, 2020. Partitioning multiple facets of beta diversity in a tropical stream macroalgal metacommunity. Journal of Biogeography 47(8): 1765–1780.

Brown, B. L., E. R. Sokol, J. Skelton & B. Tornwall, 2017. Making sense of metacommunities: dispelling the mythology of a metacommunity typology. Oecologia 183(3): 643–652.

Burgad, A. A., G. L. Adams & R. Adams, 2019. Temporal and spatial dynamics of fish community structure during watershed alteration in two Ouachita River systems. Ecology of Freshwater Fish 28(3): 459–472.

Caetano, V., M. Camana, R. B. Dala-Corte & A. S. Melo, 2021. Scale-sensitive stream slope drives nested fish trait-based diversity. Aquatic Ecology 55(3): 1051–1063.

Cardoso, P., F. Rigal, J. C. Carvalho, M. Fortelius, P. A. Borges, J. Podani & D. Schmera, 2014. Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. Journal of Biogeography 41(4): 749–761.

Carvalho, R. A. & F. L. Tejerina-Garro, 2015. Environmental and spatial processes: what controls the functional structure of fish assemblages in tropical rivers and headwater streams? Ecology of Freshwater Fish 24(2): 317–328.

Chang, J., D. L. Rabosky, S. A. Smith & M. E. Alfaro, 2019. An R package and online resource for macroevolutionary studies using the ray-finned fish tree of life. Methods in Ecology and Evolution 10(7): 1118–1124.

Chen, K. & J. D. Olden, 2020. Threshold responses of riverine fish communities to land use conversion across regions of the world. Global Change Biology 26(9): 4952–4965.

Coombes, K. R., M. Wang & M. K. R. Coombes, 2019. Package 'PCDimension'. https://cran.r-project.org/package=PCDimension. Accessed 6 Dec 2021.

Cottenie, K., 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecology Letters 8(11): 1175–1182.

Dray, S., P. Legendre & P. R. Peres-Neto, 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling 196(3–4): 483–493.

Dray, S., G. Blanchet, D. Borcard, G. Guenard, T. Jombart, P. Legendre & H. Wagner, 2018. Package 'adespatial': multivariate multiscale spatial analysis. https://cran.r-project.org/package=adespatial. Accessed 6 Dec 2021.

Faith, D. P., 1992. Conservation evaluation and phylogenetic diversity. Biological Conservation 61(1): 1–10. https://doi.org/10.1016/0006-3207(92)91201-3.

Fetzner, J. W., Jr. & K. A. Crandall, 2003. Linear habitats and the nested clade analysis: an empirical evaluation of geographic versus river distances using an Ozark crayfish (Decapoda: Cambaridae). Evolution 57(9): 2101–2118.

Finn, D. S., N. Bonada, C. Múrria & J. M. Hughes, 2011. Small but mighty: headwaters are vital to stream network biodiversity at two levels of organization. Journal of the North American Benthological Society 30(4): 963–980.

Fox, J. & G. Monette, 1992. Generalized collinearity diagnostics. Journal of the American Statistical Association 87(417): 178–183.

Frimpong, E. A. & P. L. Angermeier, 2009. Fish traits: a database of ecological and life-history traits of freshwater fishes of the United States. Fisheries 34(10): 487–495.

Geheber, A. D., 2019. Contemporary and historical species relationships reveal assembly mechanism intricacies among co-occurring darters (Percidae: Etheostomatinae). Copeia 107(3): 464–474.

Geheber, A. D. & P. K. Geheber, 2016. The effect of spatial scale on relative influences of assembly processes in temperate stream fish assemblages. Ecology 97(10): 2691–2704.

Gianuca, A. T., S. A. Declerck, P. Lemmens & L. De Meester, 2017. Effects of dispersal and environmental heterogeneity on the replacement and nestedness components of β-diversity. Ecology 98(2): 525–533.

Griffith, D. A. & P. R. Peres-Neto, 2006. Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. Ecology 87(10): 2603–2613.

Grossman, G. D., P. B. Moyle & J. O. Whitaker Jr., 1982. Stochasticity in structural and functional characteristics of an Indiana stream fish assemblage: a test of community theory. The American Naturalist 120(4): 423–454.

Heino, J. & K. T. Tolonen, 2017. Ecological drivers of multiple facets of beta diversity in a lentic macroinvertebrate metacommunity. Limnology and Oceanography 62(6): 2431–2444.

Hill, M. J., J. Heino, J. C. White, D. B. Ryves & P. J. Wood, 2019. Environmental factors are primary determinants of different facets of pond macroinvertebrate alpha and beta diversity in a human-modified landscape. Biological Conservation 237: 348–357.

Horváth, Z., C. F. Vad & R. Ptacnik, 2016. Wind dispersal results in a gradient of dispersal limitation and environmental match among discrete aquatic habitats. Ecography 39: 726–732.

Huang, L., J. Huang, Z. Wu, Y. Mo, Q. Zou, E. Jeppesen & N. Wu, 2019. Beta diversity partitioning and drivers of variations in fish assemblages in a headwater stream: Lijiang River, China. Water 11(4): 680.

Jackson, D. A., 1995. PROTEST: a PROcrustean randomization TEST of community environment concordance. Ecoscience 2(3): 297–303.

Jackson, D. A., P. R. Peres-Neto & J. D. Olden, 2001. What controls who is where in freshwater fish communities the roles of biotic, abiotic, and spatial factors. Canadian Journal of Fisheries and Aquatic Sciences 58(1): 157–170.

Kanno, Y., W. T. Russ, C. J. Sutherland & S. B. Cook, 2012. Prioritizing aquatic conservation areas using spatial patterns and partitioning of fish community diversity in a near-natural temperate basin. Aquatic Conservation: Marine and Freshwater Ecosystems 22(6): 799–812.

Kembel, S. W., D. D. Ackerly, S. Blomberg, W. Cornwell, P. Cowan, M. Helmus, H. Morlon & C. Webb, 2010. R tools for integrating phylogenies and ecology. R Package' picante. https://cran.r-project.org/package=picante. Accessed 6 Dec 2021

Landeiro, V. L., W. E. Magnusson, A. S. Melo, H. M. Espirito-Santo & L. M. Bini, 2011. Spatial eigenfunction analyses in stream networks: do watercourse and overland distances produce different results? Freshwater Biology 56(6): 1184–1192.

Legendre, P., 2014. Interpreting the replacement and richness difference components of beta diversity. Global Ecology and Biogeography 23(11): 1324–1334.

Legendre, P. & L. Legendre, 2012. Numerical Ecology, Elsevier, Amsterdam:

Legendre, P., D. Borcard & P. R. Peres-Neto, 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. Ecological Monographs 75(4): 435–450.

Leibold, M. A. & J. M. Chase, 2018. Metacommunity Ecology, Princeton University Press, Princeton:

Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau & A. Gonzalez, 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7(7): 601–613.

Leprieur, F., J. D. Olden, S. Lek & S. Brosse, 2009. Contrasting patterns and mechanisms of spatial turnover for native and exotic freshwater fish in Europe. Journal of Biogeography 36(10): 1899–1912.

Li, M., J. Liu, J. D. Tonkin, J. Shen, N. Xiao & J. Wang, 2020. The effects of abiotic and biotic factors on taxonomic and phylogenetic diversity of stream epilithic bacteria around Qiandao Lake. Aquatic Sciences 82(4): 1–12.

Linke, S., B. Lehner, C. O. Dallaire, J. Ariwi, G. Grill, M. Anand, P. Beames, V. Burchard-Levine, S. Maxwell, H. Moidu, F. Tan & M. Thieme, 2019. Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. Scientific Data 6(1): 1–15.

López-Delgado, E. O., K. O. Winemiller & F. A. Villa-Navarro, 2019. Do metacommunity theories explain spatial variation in fish assemblage structure in a pristine tropical river? Freshwater Biology 64(2): 367–379.

López-Delgado, E. O., K. O. Winemiller & F. A. Villa-Navarro, 2020. Local environmental factors influence beta-diversity patterns of tropical fish assemblages more than spatial factors. Ecology 101(2): e02940.

Lowe, W. H. & G. E. Likens, 2005. Moving headwater streams to the head of the class. BioScience 55(3): 196–197.

Maasri, A., M. Pyron, E. R. Arsenault, J. H. Thorp, B. Mendsaikhan, F. Tromboni, et al., 2021. Valley-scale hydrogeomorphology drives river fish assemblage variation in Mongolia. Ecology and Evolution 11: 6527–6535.

Mammola, S., C. P. Carmona, T. Guillerme & P. Cardoso, 2021. Concepts and applications in functional diversity. Functional Ecology 35: 1869–1885.

Marsh-Matthews, E. & W. J. Matthews, 2000. Geographic, terrestrial and aquatic factors: which most influence the structure of stream fish assemblages in the midwestern United States? Ecology of Freshwater Fish 9(1–2): 9–21.

Matthews, W. J., 1986. Fish faunal structure in an Ozark stream: stability, persistence and a catastrophic flood. Copeia. https://doi.org/10.2307/1444997.

Matthews, W. J. & E. Marsh-Matthews, 2017. Stream fish community dynamics, Johns Hopkins University Press, Baltimore:

Mayfield, M. M. & J. M. Levine, 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. Ecology Letters 13(9): 1085–1093.

McArdle, B. H. & M. J. Anderson, 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology 82(1): 290–297.

McGarvey, D. J. & R. M. Hughes, 2008. Longitudinal zonation of Pacific Northwest (USA) fish assemblages and the species-discharge relationship. Copeia 2008(2): 311–321.

McNaughton, S. J., 1977. Diversity and stability of ecological communities: a comment on the role of empiricism in ecology. The American Naturalist 111(979): 515–525.

Meynard, C. N., V. Devictor, D. Mouillot, W. Thuiller, F. Jiguet & N. Mouquet, 2011. Beyond taxonomic diversity patterns: how do α, β and γ components of bird functional and phylogenetic diversity respond to environmental gradients across France? Global Ecology and Biogeography 20(6): 893–903.

Miller, R. J. & H. W. Robison, 2004. Fishes of Oklahoma, University of Oklahoma Press, Norman:

Montaña, C. G. & K. O. Winemiller, 2010. Local-scale habitat influences morphological diversity of species assemblages of cichlid fishes in a tropical floodplain river. Ecology of Freshwater Fish 19(2): 216–227.

Morales-Castilla, I., T. J. Davies & M. Á. Rodríguez, 2020. Historical contingency, niche conservatism and the tendency for some taxa to be more diverse towards the poles. Journal of Biogeography 47(4): 783–794.

Mori, A. S., F. Isbell & R. Seidl, 2018. β-diversity, community assembly, and ecosystem functioning. Trends in Ecology & Evolution 33(7): 549–564.

Mozzaquattro, L. B., R. B. Dala-Corte, F. G. Becker & A. S. Melo, 2020. Effects of spatial distance, physical barriers, and habitat on a stream fish metacommunity. Hydrobiologia 847: 3039–3054.

Naimi, B., 2013. usdm: uncertainty analysis for species distribution models. https://cran.r-project.org/package=usdm. Accessed 6 Dec 2021

Nakamura, G., L. O. Gonçalves & L. D. S. Duarte, 2020. Revisiting the dimensionality of biological diversity. Ecography 43(4): 539–548.

Nunes, C. A., R. F. Braga, J. E. C. Figueira, F. D. S. Neves & G. W. Fernandes, 2016. Dung beetles along a tropical altitudinal gradient: environmental filtering on taxonomic and functional diversity. PLoS One 11(6): e0157442.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'hara et al., 2016. Package 'vegan' Community ecology package. https://cran.r-project.org/package=vegan. Accessed 6 Dec 2021

Pavoine, S. & M. B. Bonsall, 2011. Measuring biodiversity to explain community assembly: a unified approach. Biological Reviews 86(4): 792–812.

Pebesma, E., 2018. Package "sf" - Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal 10: 439–446.

Peres-Neto, P. R., D. A. Jackson & K. M. Somers, 2005. How many principal components? Stopping rules for determining the number of non-trivial axes revisited. Computational Statistics & Data Analysis 49(4): 974–997.

Perez Rocha, M., L. M. Bini, S. Domisch, K. T. Tolonen, J. Jyrkänkallio-Mikkola, J. Soininen, et al., 2018. Local environment and space drive multiple facets of stream macroinvertebrate beta diversity. Journal of Biogeography 45(12): 2744–2754.

Peterson, E. E., J. M. Ver Hoef, D. J. Isaak, J. A. Falke, M. J. Fortin, C. E. Jordan, et al., 2013. Modelling dendritic ecological networks in space: an integrated network perspective. Ecology Letters 16(5): 707–719.

Pigg, J., 1977. A survey of the fishes of the Muddy Boggy River in south central Oklahoma. Proceedings of the Oklahoma Academy of Science 57: 68–82.

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.

Pollice, A., G. Jona-Lasinio, M. Gaglio, F. G. Blanchet & E. A. Fano, 2020. Modelling the effect of directional spatial ecological processes for a river network in Northern Italy. Ecological Indicators 112: 106144.

Pool, T. K., G. Grenouillet & S. Villéger, 2014. Species contribute differently to the taxonomic, functional, and phylogenetic alpha and beta diversity of freshwater fish communities. Diversity and Distributions 20(11): 1235–1244.

Poos, M. S. & D. A. Jackson, 2012. Addressing the removal of rare species in multivariate bioassessments: the impact of methodological choices. Ecological Indicators 18: 82–90.

Pracheil, B. M., P. B. McIntyre & J. D. Lyons, 2013. Enhancing conservation of large-river biodiversity by accounting for tributaries. Frontiers in Ecology and the Environment 11(3): 124–128.

Rabosky, D. L., J. Chang, P. O. Title, P. F. Cowman, L. Sallan, M. Friedman, et al., 2018. An inverse latitudinal gradient in speciation rate for marine fishes. Nature 559(7714): 392–395.

Roa-Fuentes, C. A., J. Heino, M. V. Cianciaruso, S. Ferraz, J. O. Zeni & L. Casatti, 2019. Taxonomic, functional, and phylogenetic β-diversity patterns of stream fish assemblages in tropical agroecosystems. Freshwater Biology 64(3): 447–460. https://doi.org/10.1111/fwb.13233.

Robison, H. W. & T. M. Buchanan, 2020. Fishes of Arkansas, University of Arkansas Press, Fayetteville:

Ross, S. T., W. J. Matthews & A. A. Echelle, 1985. Persistence of stream fish assemblages: effects of environmental change. The American Naturalist 126(1): 24–40.

Saito, V. S., T. Siqueira & A. A. Fonseca-Gessner, 2015. Should phylogenetic and functional diversity metrics compose macroinvertebrate multimetric indices for stream biomonitoring? Hydrobiologia 745(1): 167–179.

Shao, X., Y. Fang, J. W. Jawitz, J. Yan & B. Cui, 2019. River network connectivity and fish diversity. Science of the Total Environment 689: 21–30.

Si, X., A. Baselga, F. Leprieur, X. Song & P. Ding, 2016. Selective extinction drives taxonomic and functional alpha and beta diversities in island bird assemblages. Journal of Animal Ecology 85(2): 409–418.

Simonson, T. D., J. Lyons & P. D. Kanehl, 1993. Guidelines for evaluating fish habitat in Wisconsin streams. US Forest Service Technical Report NC-164. https://doi.org/10.2737/NC-GTR-164.

Soininen, J., J. Heino & J. Wang, 2018. A meta-analysis of nestedness and turnover components of beta diversity across organisms and ecosystems. Global Ecology and Biogeography 27: 96–109.

Stegmann, L. F., R. P. Leitão, J. Zuanon & W. E. Magnusson, 2019. Distance to large rivers affects fish diversity patterns in highly dynamic streams of Central Amazonia. PloS One 14(10): e0223880.

Stevens, R. D. & J. S. Tello, 2018. A latitudinal gradient in dimensionality of biodiversity. Ecography 41(12): 2016–2026.

Taylor, C. M. & P. W. Lienesch, 1996. Regional parapatry of the congeneric cyprinids *Lythrurus snelsoni* and *L. umbratilis*: species replacement along a complex environmental gradient. Copeia 1996(2): 493–497.

Taylor, C. M. & M. L. Warren Jr., 2001. Dynamics in species composition of stream fish assemblages: environmental variability and nested subsets. Ecology 82(8): 2320–2330.

Terui, A., S. Kim, C. L. Dolph, T. Kadoya & Y. Miyazaki, 2021. Emergent dual scaling of riverine biodiversity.

Proceedings of the National Academy of Sciences 118(47): e2105574118. https://doi.org/10.1073/pnas.2105574118.

Tuomisto, H., L. Ruokolainen & K. Ruokolainen, 2012. Modelling niche and neutral dynamics: on the ecological interpretation of variation partitioning results. Ecography 35(11): 961–971.

Tyers, M., 2017. Riverdist: River Network Distance Computation and Applications. https://cran.r-project.org/package=riverdist. Accessed 6 Dec 2021.

Vellend, M., 2010. Conceptual synthesis in community ecology. The Quarterly Review of Biology 85(2): 183–206. https://doi.org/10.1086/652373.

Ver Hoef, J. M. & E. E. Peterson, 2010. A moving average approach for spatial statistical models of stream networks. Journal of the American Statistical Association 105(489): 6–18.

Villéger, S., N. W. Mason & D. Mouillot, 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. Ecology 89(8): 2290–2301.

Villéger, S., J. R. Miranda, D. F. Hernandez & D. Mouillot, 2012. Low functional β-diversity despite high taxonomic β-diversity among tropical estuarine fish communities. PloS One 7(7): e40679.

Villéger, S., G. Grenouillet & S. Brosse, 2013. Decomposing functional β-diversity reveals that low functional β-diversity is driven by low functional turnover in E uropean fish assemblages. Global Ecology and Biogeography 22(6): 671–681.

Webb, C. O., D. D. Ackerly, M. A. McPeek & M. J. Donoghue, 2002. Phylogenies and community ecology. Annual Review of Ecology and Systematics 33(1): 475–505.

Whittaker, R. H., 1960. Vegetation of the Siskiyou mountains, Oregon and California. Ecological Monographs 30(3): 279–338.

Zbinden, Z. D. & W. J. Matthews, 2017. Beta diversity of stream fish assemblages: partitioning variation between spatial and environmental factors. Freshwater Biology 62(8): 1460–1471.

Zbinden, Z. D., A. D. Geheber, W. J. Matthews & E. Marsh-Matthews, 2022. Fish communities, species of greatest conservation need, and protected areas in southeastern Oklahoma, 2014–2016. Proceedings of the Oklahoma Academy of Science 101: 14–32.

Zeni, J. O., D. J. Hoeinghaus, C. A. Roa-Fuentes & L. Casatti, 2020. Stochastic species loss and dispersal limitation drive patterns of spatial and temporal beta diversity of fish assemblages in tropical agroecosystem streams. Hydrobiologia 847(18): 3829–3843.