Distinctive macroinvertebrate communities in a subtropical river network

Rebecca Zawalski\textsuperscript{a,b}, Weston H. Nowlin\textsuperscript{a}, Karl Cottenie\textsuperscript{c}, Archis Grubh\textsuperscript{d} and Astrid N. Schwalb\textsuperscript{a}

\textsuperscript{a}Biology Department, Texas State University, San Marcos, TX, USA; \textsuperscript{b}Oklahoma Conservation Commission, Oklahoma City, OK, USA; \textsuperscript{c}Department of Integrative Biology, University of Guelph, Guelph, ON, N1G 2W1, Canada; \textsuperscript{d}Texas Parks and Wildlife Department, Rivers Studies Program, San Marcos, TX, USA

ABSTRACT

Macroinvertebrates are widely used as bio-indicators in streams and rivers, and it is usually assumed that their community composition is primarily controlled by local environmental conditions. We examined the distribution of macroinvertebrates within the Guadalupe River basin (3256 km\textsuperscript{2}) in Central Texas across physiographic gradients. Spatial analysis with variables that considers flow direction, connectivity and distances between sites (asymmetric eigenvector maps, AEM) detected distinctive communities in the lower reaches of the mainstem, in spring-influenced reaches, and in a tributary with intermittent reaches. Variation partitioning with redundancy analysis showed that large-scale factors, i.e. riverine network patterns (large-scale AEM variables), climatic variation and ecoregion explained a significant proportion (28\%) of the variation in community composition within a river basin. The riverine network patterns were the most important factor, explaining 12\% alone. Local environmental factors were significant, but completely confounded within these spatial patterns. We propose that there are distinctive macroinvertebrate communities depending on the location in the river network and this may apply to other (subtropical) rivers, which should be tested by future studies. We recommend spatial analysis that considers distances and connectivity within a river network as a powerful tool to recognize multiscale riverine network patterns, which can help to identify priority areas for conservation and to develop sound monitoring programs.

ARTICLE HISTORY

Received 3 October 2018
Revised 14 January 2019
Accepted 17 January 2019

KEYWORDS

metacommunity; AEM spatial variables; biomonitoring; biogeography; aquatic insects; springs; intermittent flow

Introduction

Biological monitoring is a widely accepted survey methodology to evaluate the ecological health of rivers and streams (Barbour et al. 1999). Although it is well known that both
local and regional factors may affect the distribution of communities, one of the principle assumptions of macroinvertebrate biomonitoring is that local communities are primarily controlled by local environmental conditions. Therefore, when high dispersal rates among spatially connected communities (i.e. so-called mass effects) override the importance of local factors, biomonitoring may lead to inaccurate information about the environmental health of a local aquatic system (Vilmi et al. 2016). The relative importance of local environmental conditions versus dispersal for metacommunity structure may depend on the location within the river network. For example, an analysis of three river basins in Maryland, USA found that local environmental factors were most important for macroinvertebrate community structure in headwater sites, but that dispersal-driven processes were more important in riverine mainstem sites (Brown and Swan 2010, but see Schmera et al. 2018). In addition, the spatial location of sites within a riverine network influences composition and diversity across a drainage, with diversity typically being greater at the confluence points in the mainstem network and in the lower reaches of a drainage (Altermatt et al. 2013).

Numerous studies have examined the relative importance of local environmental versus spatial factors for the structuring of metacommunities (e.g. Cottenie 2005; Logue et al. 2011). Most studies conducted in streams and rivers found that local environmental factors are generally more important than spatial factors (Heino et al. 2015b; but see Heino et al. 2015a), their relative importance, however, may vary with distances between sites and the spatial extent surveyed (Heino et al. 2015b). In contrast, dispersal processes over larger time scale such as historical colonization events are often ignored, although they may play a major role for metacommunity structure (Castillo-Escrivà et al. 2017). It is also important to differentiate between the role of dispersal processes at the spatial scale of metacommunities (local communities linked by dispersal, Leibold et al. 2004) and a biogeographical scale (e.g. along a macroclimatic gradient, Gonçalves-Souza et al. 2014). In particular, studies examining patterns in community composition and diversity over larger spatial extents, e.g. across drainage basins, should also consider biogeographic patterns which are results of long-term dispersal effects (e.g. historical colonization events, historical dispersal barriers) and large-scale environmental differences (e.g. climatic gradient; Leibold et al. 2010; Heino et al. 2015b, 2017).

The purpose of this study was to examine multiscale spatial patterns of benthic macroinvertebrates within a riverine network (hereafter referred to as ‘riverine network patterns’) in the Guadalupe River basin in central Texas, USA (Figure 1). The spatial extent of the basin examined by this study allowed us to also examine larger scale spatial patterns of macroinvertebrate communities that are likely the result of biogeographic processes. An increasing number of studies have used complex spatial analyses such as asymmetric eigenvector map (AEM) analysis for variation partitioning in metacommunity analyses of macroinvertebrates (e.g. Göthe et al. 2013; Zhang et al. 2014; Cauvy-Fraunié et al. 2015). However, AEM analysis can also be used to identify multiscale riverine network patterns for macroinvertebrates at a larger spatial scale to identify locations, reaches, and segments of the river with a distinct macroinvertebrate community composition. This could serve as a useful starting point to discuss and further investigate potential factors that drive these patterns, instead of attributing any differences between monitoring sites to local environmental factors, and could ultimately lead to a more effective design of monitoring programs. Therefore, we addressed the following main questions: (1) What is the spatial pattern of macroinvertebrates communities in the Guadalupe Basin? Which multiscale riverine network patterns do they show? (2) Which groups of
macroinvertebrates and environmental factors (local and larger scale) are associated with these patterns? (3) What is the relative importance of climatic variation, ecoregion and larger scale riverine network patterns? (4) To what extent are climatic variation, ecoregion, and riverine network patterns correlated with local environmental conditions and land use-land cover?

**Materials and methods**

**Study area**

The Guadalupe River basin is relatively large (3256 km²) and encompasses a pronounced regional physiographic gradient. Like the other Gulf coast rivers in Texas, the Guadalupe River flows from northwest to southeast, experiencing a climatic gradient with increased precipitation from west to east, and increasing temperatures from north to south (see below). Both high disturbance intensity and stable conditions occur in this basin, where flashfloods are common and where groundwater-fed tributaries with relatively consistent flow occur next to tributaries with intermittent reaches and higher variation in seasonal flow. The basin was never glaciated and also includes major springs, which are considered hotspots for endemic macroinvertebrates (Bowles and Arsurf 1993).

The Guadalupe River basin contains portions of four level III ecoregions of North America (https://www.epa.gov/eco-research/ecoregions-north-america, Figure 1): the Edwards Plateau, the Texas Blackland Prairies, the East Central Texas Plains and the Western Gulf Coastal Plain. All four ecoregions can be characterized by homogeneity...
associated with both abiotic – soils, vegetation, geology, climate and physiography (Omernik 1987; Griffith et al. 2004), and biotic factors, including algal coverage. The ecoregions do not follow a strict up to downstream pattern but two of the four ecoregions alternate in the middle and lower reaches of the Guadalupe (Figure 1).

The Edwards Plateau is dominated by karst limestone geology and many headwaters and stream reaches are strongly spring-influenced, containing clear water with high physicochemical stability. The Blackland Prairies is dominated by clays and silty soils and contains larger fraction of cropland and urban space. The East Central Texas Plains is largely composed of savanna and mostly used for pasture. The Western Gulf Coastal Plain is a low gradient plain that ends at the Gulf of Mexico. Rainfall varies across the basin, from a minimum of 406 mm per year in the north and western portions of the basin (Edwards Plateau) to a maximum of 1473 mm per year in the southern and eastern regions (Western Gulf Coast Plain).

Two of the major tributaries, the Comal and San Marcos rivers, are strongly groundwater influenced (i.e. these rivers are fed by large spring complexes at their headwaters) and thus exhibit stable physicochemical conditions and relatively more consistent seasonal flows. In contrast, the other tributary, the Blanco River and the upper portion of the Guadalupe River mainstem (sites 14–16 on the Guadalupe, Figure 1) exhibit much higher variation in seasonal flows with some sections of these rivers going dry during dry years or experiencing large-scale flooding during wet periods. Indeed, the Blanco River experienced large-scale and historic flooding in early May 2015 several months before we started sampling for this study.

Field data collection

We collected macroinvertebrates and local environmental data at 28 sites across the basin between 15 July and 5 October 2015 (each site sampled once). Sites were located in the main tributaries of the Guadalupe River including the Comal River (2 sites), the San Marcos River (4 sites), and the Blanco River (7 sites); the remaining 15 sites were distributed along the mainstem of the Guadalupe River stretching over 492 river kilometers from the most upstream (site 14, Figure 1) to the most downstream site (site 28, Figure 1, Table S1). Each site was sampled once across the mid- and late-summer growing season. In order to minimize the effects of temporally declining flows affecting local conditions across the basin during the sampling interval, lower discharge and lower stream order sites in the drainage were sampled at the beginning of the study period. Over the course of the study, flows across the basin generally declined, with discharge at the furthest downstream site (Guadalupe River at Victoria, Texas; USGS Station ID 08176500) receding from 52.9 to 18.5 m$^3$/L ($\bar{x} \pm 1$ SD = 29.1 ± 9.2 m$^3$/L) over the 82-d sampling interval (https://waterdata.usgs.gov/tx/nwis/rt).

Prior to collecting macroinvertebrate samples at each site, we measured pH, water temperature, dissolved oxygen (DO; mg/L), and conductivity (µS/cm) using a multiparameter probe (YSI 556). Water velocity immediately upstream from each sample point in a riffle was measured with a Hach flow meter (FH950). The percent sediment size composition at each sampling point was estimated using a modified Wentworth scale (Wentworth 1922) and percent algae cover was estimated using an underwater viewing window. Duplicate water samples were taken at each sampling location using 2-L brown Nalgene bottles which were rinsed three times with site water before sample collection. Water samples were placed in a cooler on ice and transported to the lab at the Texas State University, where samples were filtered and preserved within 48 hours of collection.
Water samples were filtered to determine the concentration of \( \text{NH}_4^+ \), \( \text{NO}_3^- \), soluble reactive phosphorus (SRP, assumed to be \( \text{PO}_4^{3-} \)), total suspended solids (TSS), nonvolatile suspended solids (NVSS), and suspended chlorophyll-a (Chla). Nutrients and suspended materials were determined through lab-specific standardized methods (Caston et al. 2009).

Local environmental conditions such as substrate type and composition and water velocity are known to affect macroinvertebrate community composition within a given sampling site, (Allen 1995). We collected invertebrate samples from haphazardly placed locations within riffles at each sampling site, which are the most ideal mesohabitat to sample when evaluating macroinvertebrates since it consistently contains higher diversity (Brown and Brussock 1991, Barbour et al. 1999) and many environmental monitoring programs focus on macroinvertebrates from riffles (Carter and Resh 2001).

Macroinvertebrate samples were collected using a 500-µm Hess sampler (35-cm diameter). This mesh size is commonly used in biomonitoring programs, although it misses smaller benthic organisms, especially early stages of many macroinvertebrates. At each sampling site, four Hess samples were collected from within a riffle area, except for 5 of the 28 sites (3 Hess samples at sites 21, 24, 25 and 28, and 2 Hess samples at site 7). To account for differences in the number of Hess samples, macroinvertebrate densities for each taxon were expressed as number of individuals/m². During sampling, substrate was agitated for a 2-minute interval and samples were preserved in 90% ethanol (EtOH) for processing in the laboratory. Macroinvertebrate samples were identified under a stereo-microscope (Nikon SMZ745T) to the lowest practical taxonomic level (typically genus) using relevant several taxonomic keys (Merritt et al. 2008, Diaz 2014). A total of 59 macroinvertebrate taxa were identified, including six non-insect taxa (Table S2). Non-insect taxa were identified to order and all other taxa were identified to genus, except Diptera, which were identified to family.

**Land cover data**

Land cover data were downloaded from the United States Geological Survey and overlaid on sample site locations in ArcGIS v10.4 using the National Land Cover Database (NLCD 2011 version). Land use-land cover (LULC) was determined as percent composition among 20 categories: developed open space, developed low intensity, developed medium intensity, developed high intensity, open water, perennial ice/snow, barren land (rock/sand/clay), deciduous forest, evergreen forest, mixed forest, dwarf scrub, shrub/scrub, herbaceous grassland, herbaceous sedge, lichens, moss, pasture/hay, cultivated, woody wetlands and emergent herbaceous wetlands (NLCD 2011 Product Legend; https://www.mrlc.gov/nlcd2011.php). Three spatial scales of LULC for each sampling site were examined based on Allan (2004) and Becker et al. (2014): (1) a reach scale with land cover in a 100-m buffer on either side of the river with a 2 km buffer upstream from each site; (2) a riparian scale with land cover in a 100-m buffer for total distance upstream for each site; and (3) a catchment scale with land cover for the whole watershed upstream of the site. We followed the procedure outlined in Becker et al. (2014) to combine and reduce LULC into eight categories: urban, cultivated, evergreen forest, deciduous forest, mixed forest, rangeland, wetlands and open water. Barren land was removed from any analyses because it made up <1% of the coverage area (Dodds and Oakes 2008, Becker et al. 2014). Ecoregions for each site were based upon USEPA Level-III Ecoregions, downloaded from the EPA (Griffith et al. 2004), and overlaid across the Guadalupe River Basin in ArcGIS. Estimates of river slope were generated
using a digital elevation model (DEM), and river distances between sites were evaluated by using a river network map in ArcGIS. Mean annual precipitation data for each site were obtained from Texas Parks and Wildlife Department and reported as the annual mean during the 2000–2010 period.

### Data analysis

Prior to analysis, values obtained from duplicate water samples for each analyte from each site were averaged. To avoid issues with multicollinearity in analyses, variables which were highly correlated \((r > 0.70)\) were removed from the dataset. Mean, maximum, minimum and point slope estimates for each site were highly correlated, so only site mean slope was used in further analyses. TSS and NVSS were also highly correlated, thus TSS was used in further analyses. A Pearson correlation matrix for each group of predictor variable data set revealed that the riparian and catchment scales for LULC percent coverage were highly correlated for nearly all variables and the riparian LULC scale were removed from further analyses (Becker et al. 2014).

Twenty taxa were excluded from analysis because they contained <5% of taxa at all sites (Zhao et al. 2017). To address question (1) and to evaluate potentially complex multiscale spatial patterns within the river network, we used an AEM analysis, a spatial modeling technique that considers autocorrelation at different spatial scales. AEM analysis was developed for ecosystems such as rivers in which directional physical processes (water currents) can affect the distribution of organism (Blanchet et al. 2008a). We computed Eigenfunction-based spatial variables (eigenvectors) from a directional downstream distance matrix accounting for connectivity between sites (Blanchet et al. 2008a), and assigned weights to the edge matrix based on watercourse distances in kilometers between sites. A forward selection procedure was used to reduce the number of spatial eigenvectors to predict the variation in community composition (Blanchet et al. 2008b). We also used Moran’s eigenvector map (MEM) analysis (Dray et al. 2006) to model hydrological connections between sites without considering the direction of the flow. However, the results were very similar to the AEM-analysis and were therefore not included here.

To evaluate the association of differences in community composition with higher densities of certain groups and with environmental factors (local environmental factors, LULC and climatic variables) (question 2, Table S3) we used redundancy analyses (RDA). To examine the relative importance (question 3) of climatic variation (water temperature, precipitation), ecoregion and large-scale riverine network patterns (AEM variables), we used variation decomposition based on redundancy analysis (RDA, Cottenie 2005). To focus on larger scale patterns, we only used the first three large-scale AEM variables selected by the forward selection (see above). The computed percentage of explained variation was adjusted for the number of explanatory variables (i.e. adjusted \(R^2\), Peres-Neto et al. 2006). The dependent abundances of macroinvertebrates (genera or family) were Hellinger transformed to minimize the disproportional influence of rare species on the redundancy analysis (Legendre and Gallagher 2001). To determine to what extent ecoregion, climatic variation, and large-scale AEM variables were correlated (and how much variation they shared) with local environmental conditions and land use-land cover (question 4), we ran pairwise variation decomposition based on RDA (see above), i.e. ecoregion vs. local environmental conditions, and ecoregion vs. land use-land cover variables, and the same for climatic variation and large-scale AEM variables. Comparisons among more than two groups of variables (e.g. variation decomposition with AEM variables, local environmental factors and land use-land cover) were not possible, because of high
correlation between the variables. All analyses were done in R (R Development Core Team, 2017, version 3.4.0) using the package vegan (Oksanen et al. 2017).

**Results**

**Multiscale riverine network patterns and associated macroinvertebrates and environmental factors (Questions 1 and 2)**

The analysis of AEM variables resulted in nine significant variables (Figure 2), which explained 52% ($p = 0.001$) of the variation in macroinvertebrate community composition across the Guadalupe Basin. Most of the AEM variables represented large-scale spatial patterns (variables V1 to V6; Figure 2). Overall, the analysis revealed four notable patterns which were associated with different genera and environmental factors. First, there was a unique macroinvertebrate community composition in the lower portion of the Guadalupe River after the confluence with the San Marcos River (V1; Figure 2), characterized by higher densities of the mayfly genus *Traverella*, the predatory stonefly *Neoperla*, and the riffle beetle *Hexacyleleopus* (Figure 3a). This spatial pattern largely corresponded with the climatic gradient in the basin, with greater precipitation (and higher water temperatures, Figure 3b) in the lower portions of the basin (Figure 3b). These changes in community structure in the lower portion of the Guadalupe River also correlated with several local environmental factors: higher TSS and Chla concentrations, slightly higher pH, and higher proportion of sand in benthic substrates (Figure 3c). With respect to LULC patterns in the basin, these taxonomic changes were also correlated with an increase in the percent coverage of wetlands and agriculture at the catchment scale (Figure 3b).

The second riverine network pattern was associated with the spring-influenced reaches along the Comal and San Marcos rivers (V21, Figure 3a both spring sites have the highest negative values for second RDA axis). These sites exhibited higher densities of macroinvertebrates with lower dispersal abilities or those lacking desiccation-resistant resting stages (i.e. *Hyalella*, the water penny *Psephenus*), and the riffle beetle *Microcyleleopus*) (Figure 3a). In these reaches, the LULC patterns had higher percentages of urban, mixed forest, and open water (presumably associated with the headwater spring complexes).

A third distinctive community type in the basin was found along the Blanco River characterized by high abundances of the net spinning caddisfly *Chimarra* (variables V2, V6 and partly V4; Figures 2 and 3(a)). Local environmental conditions at these sites tended to have greater water depths and a higher proportion of cobble in benthic substrates (Figure 3(c)). Blanco sites also had slightly a higher proportion of ranchland use (up to 72%, compared to up to 60% elsewhere in the basin, Figure 3(b)).

The last major community type was associated with sites located within the Texas Blackland Prairies ecoregion (several sites in the lower San Marcos, the upper Guadalupe, and the lower Guadalupe rivers, compare large open squares just before and after the bigger confluence in V5, Figure 2 with blue Texas Blackland Prairies ecoregion in Figure 1). Although there were no obvious associations with environmental factors (Figure 3), there were increased densities of the mayflies *Baetis* and *Leptohyphes* at the sites located in this region.

Other riverine network patterns that were detected included a different pattern around the confluence of the San Marcos with the Guadalupe River (V3; Figure 2), and another smaller scale pattern with the most upstream reaches of the upper Guadalupe being different (V11; Figure 2).
Figure 2. Significant AEM variables. White squares symbolize negative AEM scores, black squares symbolize positive AEM scores, the bigger the size of the square the higher the AEM score (but differs for each panel). Sites are placed according to their geographic coordinates, and the lines are connections between sites.
Figure 3. Biplots of redundancy analysis with sites in the upper Guadalupe (white triangles), the lower Guadaluper (black squares), the Comal River (black triangles), the Blanco (black circles) and the San Marcos River (white circles). (a) The arrows indicate the significant AEM variables, and the letters show genera that distinguished between different spatial patterns. Chim: *Chimarra*, Trav: *Traverella*, Hexa: *Hexacylleopus*, Neop: *Neoperla*, Hyal: *Hyallella*, Psep: *Psephenus*, Micr: *Mycrocylleopus*, (b) Arrows indicate significant climatic variables and catchment land-use, land-cover variables. Ag: cultivated, DecFor: deciduous forest, Ran: rangeland, Wet: wetlands and (c) Arrows indicate significant local environmental variables.
Relative importance of climatic variation, ecoregion, and larger scale riverine network patterns (Question 3)

Large-scale AEM variables (the first three variables) explained 12% of the variation in community composition across the Guadalupe Basin after eliminating the shared effects of the other factors climate and ecoregion (i.e. pure effects). The pure effects of climate variation and ecoregion were 4% and 3%, respectively (Figure 4). In addition, ecoregion and AEM variables shared 7% of the variation, and all variables shared 5% of the variation. Due to multicollinearity between the factors, the total amount of variation explained was 28% instead of 31%.

Correlation with local environmental and LULC factors (Question 4)

Pairwise variation partitioning in RDA indicated that that ecoregion shared more variation with LULC factors (15%) than with local environmental factors (7%). Similarly, climatic factors shared a small proportion of the variation with LULC factors (7%), but none with local environmental factors. Not surprisingly, both LULC factors and local environmental factors were spatially structured. Local environmental factors shared 28% (all variation) with (large-scale) AEM variables, and LULC factors 22%.

Discussion

We found that a considerable portion of the variation in community composition of macroinvertebrates in the Guadalupe River basin was explained by variation in larger scale factors (i.e. climatic variation, ecoregion and riverine network patterns). Such factors are not routinely included in monitoring programs, despite findings that such large-scale patterns are important (e.g. Feminella 2000; Mykrä et al. 2004). Not surprisingly, part of the variation in climate and ecoregion in the Guadalupe basin were spatially structured (i.e. shared variation with large-scale AEM variables), but both climate and ecoregion also explained a small amount of the variation in community structure on its own. The pure effects of large-scale spatial variables were most important and all significant AEM variables combined explained ~50% of the variation in macroinvertebrate community structure. This result suggests that the location of a site within a river network (e.g. the
presence of spring influenced reaches, confluence points in the network) and historical and current connectivity were more important in determining macroinvertebrate community composition than the physiographic gradients across the basin.

The importance of local environmental factors for metacommunity structuring has been shown by many metacommunity studies in rivers (Heino et al. 2015b; but see Heino et al. 2015a), suggesting that species sorting is the prevalent dynamic in rivers. It is important, however, to consider the spatial extent encompassed by a study and the distances between sites in relation to the dispersal abilities of the study organisms as different processes may act on different scales. For example, a study on neotropical lepidopterans and spiders found that environmental effects dominated at the metacommunity scale, whereas at the biogeographical scale dispersal-based processes were more important (Gonçalves-Souza et al. 2014). The biogeographic scale was also considered in a study on aquatic organisms across drainage basins in Finland, which showed that basin identity and local environmental variables were both important for community structure, whereas the spatial effects within a basin were usually negligible (Heino et al. 2017, area of three drainage basins: 63,609 km²). In contrast to this, our study found that large-scale spatial patterns can also play a role within a basin (area: 3256 km²). Unfortunately, we could not determine the relative importance of local environmental factors, as all local environmental factors were spatially structured (local environmental factors shared all explained variation with (large scale) AEM variables, see above). In order to avoid the local environmental factors to be completely confounded with large-scale spatial patterns, a different sampling design would have been needed with additional sites being placed closer to each other. It is also possible that local factors would explain additional variation in the distribution of macroinvertebrates, when more specific requirements of organisms were included in analyses, such as food availability and prey presence. For freshwater mussels (technically also macroinvertebrates, but usually not included in routine biomonitoring), the inclusion of a moving niche component, that is, the presence/absence of their host fish, they need for successful recruitment, explained substantially more of their distribution across river basins in Ontario, Canada (Schwalb et al. 2013). Maybe the biggest obstacle is the scarce information available about life history of many macroinvertebrates, hence more life history studies of macroinvertebrates are needed.

The AEM-variables revealed interesting multiscale riverine network patterns in macroinvertebrate community composition, which indicated a distinct community composition for the lower portion of the basin (lower Guadalupe mainstem), the spring-influenced reaches, the Blanco River, and the different ecoregions. The different community composition in the lower Guadalupe mainstem was associated with local environmental factors, but also larger scale climatic factors, which are usually associated with a biogeographical spatial scale. In particular, the plecopterans (mostly the genus Neoperla) were present in the lower mainstem Guadalupe sites, but were largely absent from sites extending up into the Edwards Plateau. The large-scale southeast-northwest spatial occurrence pattern of plecopterans in the study watershed is consistent with previous descriptions of plecopterans distributions in the region. This is likely related to precipitation patterns and its influence on stream permanence and the past biogeographic dispersal patterns of plecopterans in the region (Stewart et al. 1973; Szytko and Stewart 1977). In addition, larger scale differences in water quality and food availability likely influence the occurrence of species in the lower Guadalupe mainstem. The lower mainstem has higher suspended Chl-a and higher TSS concentrations, indicating greater food availability for filter feeders such as Traverella, which also occurred in higher abundances in the lower Guadalupe.
The spring influenced reaches of the basin were characterized by macroinvertebrate communities containing species with lower dispersal abilities or an inability to tolerate periods of desiccation, such as amphipods (*Hyalella azteca*) and smaller bodied riffle beetles (*Microhyalella pusillus*). It is also noteworthy that the proportion of shredders was considerably higher near Spring Lake in the San Marcos River and near Comal Springs compared to the rest of the basin (0–3%, see Zawalski 2017 for more details). Diversity is typically higher in lower reaches of river networks (Altermatt et al. 2013) and that holds true for freshwater fishes and unionid mussels in Texas (Dascher et al. 2018). However, we found the highest diversity of macroinvertebrates near springs (Zawalski 2017). Spring-influenced reaches of riverine networks, especially in arid areas, play an important role as ecological and evolutionary refugia (Davis et al. 2013) in that some species found in these reaches may only be able to persist in these refugia during extended or severe drought periods. As such, protection of these segments from anthropogenic impacts, including restriction of groundwater pumping from regional aquifers provides crucial refuge and protection for these spring-associated and dispersal-limited taxa (Bowles and Arsuflu 1993) as well as maintaining locations of higher species diversity in the landscape.

Unmeasured factors such as flow permanence may be an important driver for the distinctive community in the Blanco River. Several reaches (at least 3 sites) were dry in 2014 and all of the sites in the Blanco River experienced an especially large flooding event in 2015, shortly before the sampling. The communities in this river were characterized by high abundances of *Bifemurum*, which could be especially resilient to such disturbance as it was found in higher abundances in unstable substrate in the Ardèche river, France (when compared to other net-spinning caddisflies, Dolédec and Tachet 1989). A higher concentration of suspended food sources could also be a factor. Algal growth is enhanced in pools that become isolated or when flow is decreased. In addition, when intermittent reaches go dry, terrestrial organic matter can accumulate and aquatic plant material be decomposed and providing a high input of organic matter when these reaches are re-flooded (Williams 2006), which could be a good food source for collectors such as net building caddisflies. Indeed, the majority of collector-filterers in the Guadalupe basin can be found in the Blanco River as well as the headwaters of the Guadalupe and Comal Rivers (Zawalski 2017). The intermittent reaches may also be a good habitat for macroinvertebrates with high colonization potential. Interestingly, 99% of the genera present in the intermittent reaches in the Blanco River (i.e. dry in 2014) had winged adult stages. The average for all sites in the Blanco River was 96% compared to the rest of the basin with 82% not including the spring sites. At spring sites the percentage of winged results were the lowest (13% and 37% for San Marcos and Comal springs, respectively).

It is well known that differences in regional species pools must be considered for biomonitoring and the development of biotic and multimetric indices, especially if they are based on biological attributes of species instead of functional metrics (Pont et al. 2006). Our data suggest that biogeographic differences can play a role not just between basins (e.g. Heino et al. 2017), but also within river basins. Furthermore, our results indicate that the distribution of macroinvertebrates may also depend on the location in the river network (near springs, lower reaches, tributary with intermittent reaches). The importance of location in the river network for the distribution of macroinvertebrates and metacommunity structuring has been previously shown, e.g. with focus on the arrangement of tributaries (Rice et al. 2001) and river network properties (Altermatt et al. 2013), and comparing headwaters and mainstem (Brown and Swan 2010). Based on our results, we predict that metacommunity structure and dynamics in a subtropical river network will vary because of different disturbance levels found in different parts of the river network (Table 1). The environmental
conditions are most stable close to the springs, which allows macroinvertebrates with low dispersal abilities and those unable to tolerate periods of desiccation to become abundant. As these springs are evolutionary refugia species sorting should still occur despite limited dispersal over a sufficiently large time-scale. Such a species sorting with limited dispersal (sensu Winegardner et al. 2012) assumes a tradeoff between competition and dispersal, and the abundant low dispersal macroinvertebrates will co-exist with high dispersal species (Table 1). In contrast, in intermittent reaches with high disturbance levels due to drying and flash flooding, only macroinvertebrates with high dispersal abilities become abundant. The community composition is determined by colonization after disturbance events, followed by succession driven by local environmental condition and biotic interactions (i.e. species sorting with high dispersal, Table 1). However, the relative importance of dispersal vs. environmental filtering may vary with hydrological phases of dry, flowing, and non-flowing conditions in intermittent systems (Datry et al. 2016). Finally, the distinct community in the lower mainstem of the Guadalupe is subject to an intermediate disturbance level, where the community composition is mainly determined by abiotic conditions (species sorting with efficient dispersal, Table 1), but probably also by biogeographic patterns. The frequency of high- vs. low-dispersal groups of macroinvertebrates (e.g. percentage of winged adults) reflects the relative frequency found in the entire basin (Table 1). We propose that the changes we observed in macroinvertebrate communities are predictable based on the location in the river network and may apply to other (subtropical) rivers, which should be tested by future studies.

It has been postulated that ecologists should also consider distances along the river network (Heino et al. 2015b), and newly developed statistical methods allow for such a spatial analysis in river networks (Legendre and Legendre 2012). Spatial variables considering watercourse distances (i.e. AEM variables) have been used increasingly in ecological studies, but have not yet gained popularity in more applied studies. We recommend spatial analyses that consider distances and connectivity as a powerful tool to recognize multiscale riverine network patterns, and which may otherwise go undetected. For instance, a

| River network feature | Disturbance level | Metacommunity dynamics | Macroinvertebrate community | Associated genera |
|-----------------------|-------------------|------------------------|-----------------------------|-------------------|
| Close to springs      | Stable environment| Species sorting with limited dispersal, assumes tradeoff between competition and dispersal | Macroinvertebrates abundant with low dispersal abilities and/or inability to tolerate periods of desiccation, coexisting with species with high dispersal abilities (13% and 37% with winged adult stages) | Amphipod: Hyalella, water penny: Psephenus, riffle beetle: Myrcoclypeus |
| Tributary with intermittent reaches | High disturbance due to drying and flash flooding | Species sorting with high dispersal, community composition determined by colonization after disturbance, followed by succession. | Macroinvertebrates with high dispersal abilities most abundant. (99% with winged adult stages in intermittent reaches) | Net spinning cadisfly: Chimarra |
| Lower basin mainstem | Intermediate disturbance with occasional flooding but permanent flow | Species sorting with efficient dispersal, Community composition determined by abiotic conditions and biogeographic patterns | Frequency of high- and low-dispersal abilities reflective of entire basin. Community distinct from rest of basin | Filter feeding mayfly: Traverella, predatory stonefly: Neoperla |

Table 1. Proposed community structure and dynamics (sensu Winegardner et al. 2012), and associated genera in relation to disturbance levels in different parts of the river network for which a distinctive community was detected in the Guadalupe basin.
survey throughout the basin will be necessary to identify distinct communities, and those differences will then need to be considered when monitoring human impact. In addition, such an analysis can help to identify priority areas for conservation and management plans (such as spring-influenced reaches, see above). Our example uses macroinvertebrates, but it could be easily applied to fish or other groups of organisms. Using multiscale spatial analysis also helps to identify the relative importance of processes at different spatial scales, and may indicate mechanisms responsible for these patterns. Thus, it would be an important step for designing an effective monitoring program, detecting human impact, and developing mitigation plans.

Acknowledgements

We thank Gaby Timmins, Pete Diaz, and Nathaniel Dede-Bamfo for technical support, Rebecca Tucker for her help with sampling, and Don Apodaca for his help in the field and in the laboratory.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Texas Parks and Wildlife Department through Texas Water Development Board under grant no. 1349311646.

Notes on contributors

Rebecca Zawalski was an MSc student co-advised by Drs. Nowlin and Schwalb during this study, but is now a Blue Thumb Field Educator at the Oklahoma Conservation Commission.

Weston Nowlin is a professor of ecosystem ecology in the Biology Department at the Texas State University.

Karl Cottenie is an associate professor of metacommunity ecology in the Department of Integrative Biology at the University of Guelph.

Dr. Archis Grubh is an aquatic biologist with research interests in macroinvertebrate community dynamics, and currently works for Texas Parks and Wildlife Division.

Astrid Schwalb is an assistant professor of stream ecology at the Texas State University.

References

Allen JD. 1995. *Stream ecology. Structure and function of running waters*. Dordrecht/ Boston/ London: Kluwer Academic Publishers.

Allan JD. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annu Rev Ecol Evol Syst. 35(1):257–284.

Altermatt F, Seymour M, Martinez N. 2013. River network properties shape $\alpha$-diversity and community similarity patterns of aquatic insect communities across major drainage basins. J Biogeogr. 40(12):2249–2260.

Barbour MT, Gerritsen J, Snyder BD, Stribling JB. 1999. Rapid bioassessment protocols for use in wadeable streams and rivers: periphyton, benthic macroinvertebrates and fish, 2nd ed. EPA 841-B-99-002. Washington, DC: Environmental Protection Agency; Office of Water.
Becker JC, Rodibaugh KJ, Labay BJ, Bonner TH, Zhang Y, Nowlin WH. 2014. Physiographic gradients determine nutrient concentrations more than land use in a Gulf Slope (USA) river system. Freshw Sci. 33(3):731–744.

Blanchet FG, Legendre P, Borcard D. 2008a. Modelling directional spatial processes in ecological data. Ecol Model. 21:325–336.

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

Bowles DE, Arsuffi TL. 1993. Karst aquatic ecosystems of the Edwards Plateau region of central Texas, USA: a consideration of their importance, threats to their existence, and efforts for their conservation. Aquatic Conserv: Mar Freshw Ecosyst. 3(4):317–329.

Brown AV, Brussock PP. 1991. Comparisons of benthic invertebrates between riffles and pools. Hydrobiologia. 220(2):99–108.

Brown BL, Swan CM. 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. J Anim Ecol. 79(3):571–580.

Carter JL, Resh VH. 2001. After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. J North Amer Benthol Soc. 20(4):658–682.

Castillo-Escrivà A, Valls L, Rochera C, Camacho A, Rochera L, Mesquita-Joanes F. 2017. Disentangling environmental, spatial, and historical effects on ostracod communities in shallow lakes. Hydrobiologia. 787(1):61–72.

Caston, CB, Nowlin, WH, Gaulke, A, Vanni, MJ. 2009. The relative importance of heterotrophic bacteria to pelagic ecosystem dynamics varies with reservoir trophic state. Limnology and Oceanography. 54: 2143–2156.

Cauvy-Fraunie S, Espinosa R, Andino P, Jacobsen D, Dangles O. 2015. Invertebrate metacommunity structure and dynamics in an Andean glacial stream network facing climate change. PloS One. 10(8):e0136793.

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

Dascher ED, Burlakova LE, Karatayev AY, Ford DF, Schwalb AN. 2018. Distribution of unionid freshwater mussels and host fishes in Texas. A study of broad-scale spatial patterns across basins and a strong climate gradient. Hydrobiologia. 810(1):315–331.

Datry T, Bonada N, Heino J. 2016. Towards understanding the organisation of metacommunities in highly dynamic ecological systems. Oikos. 125(2):149–159.

Davis J, Pavlova A, Thompson R, Sunnucks P. 2013. Evolutionary refugia and ecological refuges: key concepts for conserving Australian arid zone freshwater biodiversity under climate change. Glob Change Biol. 19(7):1970–1984.

Diaz PH. 2014. *Photographic guide to the aquatic invertebrates of the upper San Marcos River Hays County, San Marcos, Texas*. 1st ed. ISBN-13 978-1500260170.

Diaz PH. 2016a. *Key to Ephemeroptera of Texas*. San Marcos: U. S. Fish and Wildlife Service.

Diaz PH. 2016b. *Key to Trichoptera families of Texas*. San Marcos: U. S. Fish and Wildlife Service.

Dodds WK, Oakes RM. 2008. Headwater influences on downstream water quality. Environ Manage. 41(3):367–377.

Dolédec S, Tachet H. 1989. Ecological observations and life histories of five net-spinning caddisflies (Trichoptera) of the lower ardèche river. Aquatic Insects. 11(2):89–99.

Dray SP, Legendre P, Peres-Neto PR. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol Model. 196(3/4):483–493.

Feminella JW. 2000. Correspondence between stream macroinvertebrate assemblages and 4 ecoregions of the southeastern USA. J North Amer Benthol Soc. 19(3):442–461.

Göthe E, Angeler DG, Sandin L. 2013. Metacommunity structure in a small boreal stream network. J Anim Ecol. 82(2):449–458.

Gonçalves-Souza T, Romero GQ, Cottenie K. 2014. Metacommunity versus biogeography: a case study of two groups of Neotropical vegetation-dwelling arthropods. Plos One.

Griffith GE, Bryce SA, Omernik JM, Comstock JA, Rogers AC, Harrison B, Hatch SL, Bezanson D. 2004. Ecoregions of Texas (color poster with map, descriptive text, and photographs): Reston, VA, U. S. Geological Survey (map scale 1:2,500,000).

Heino J, Melo AS, Bini LM, Altermatt F, Al-Shami SA, Angeler DG, Bonada N, Brand C, Callisto M, Cottenie K, et al. 2015. A comparative analysis reveals weak relationships between ecological factors and beta diversity of stream insect metacommunities at two spatial levels. Ecol Evol. 5(6):1235–1248.
Heino J, Melo AS, Siqueira T, Soininen J, Valanko S, Bini LM. 2015. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. Freshw Biol. 60(5):845–869.

Heino J, Soininen J, Alahuhta J, Lappalainen J, Virtanen R. 2017. Metacommunity ecology meets biogeography: effects of geographical region, spatial dynamics and environmental filtering on community structure in aquatic organisms. Oecologia. 183:12.

Legendre P, Gallagher ED. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia. 129(2):271–280.

Legendre P, Legendre L. 2012. **Numerical Ecology**, 3rd edn. Amsterdam: Elsevier.

Leibold MA, Ecomno EP, Peres-Neto P. 2010. Metacommunity phylogenetics: separating the roles of environmental filters and historical biogeography. Ecol Lett. 13(10):1290–1299.

Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, Holt RD, Shurin JB, Law R, Tilman D, Loreau M. 2004. The metacommunity concept: A framework for multi-scale community ecology. Ecol Lett. 7:601–613.

Logue JB, Mouquet N, Peter H, Hillebrand H. Metacommunity Working Group 2011. Empirical approaches to metacommunities: a review and comparison with theory. Trends Ecol. Evolut. 26(9): 482–491.

Merritt RW, Cummins KW, Berg MB. 2008. An introduction to the aquatic insects of North America. 4th ed. Dubuque, IA: Kendall-Hunt Publishers.

Mykrä H, Heino J, Muotka T. 2004. Variability of lotic macroinvertebrate assemblages and stream habitat characteristics across hierarchical landscape classifications. Environ Manager. 34(3):341–352.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MH, Wagner H. 2017. Community Ecology Package. R package version 2.4-3. https://CRAN.R-project.org/package=vegan

Omernik JM. 1987. Map Supplement Ecoregions of the Conterminous United States. Ann Assoc Amer Geogr. 77(1):118–125.

Peres-Neto PR, Legendre P, Dray S, Borcard D. 2006. Variation partitioning of species data matrices: estimation and comparison of fractions. Ecology. 87(10):2614–2625.

Pont D, Hugueny B, Beier U, Goffaux D, Melcher A, Noble R, Rogers C, Roset N, Schmutz S. 2006. Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages. J Appl Ecol. 43(1):70–80.

Core Team R. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rice SP, Greenwood MT, Joyce CB. 2001. Tributaries, sediment sources, and the longitudinal organization of macroinvertebrate fauna along river systems. Can J Fish Aquat Sci. 58(4):824–840.

Schmera D, Árva D, Boda P, Bódis E, Bolgovics A, Borics G, Csercsa A, Deák C, Krasznai EÁ, Lukács BA, Mauchart P. 2018. Does isolation influence the relative role of environmental and dispersal-related processes in stream networks? An empirical test of the network position hypothesis using multiple taxa. Freshw Biol. 2018:74–85.

Schwalb AN, Morris TJ, Mandrak NE, Cottenie K. 2013. Distribution of unionid freshwater mussels depends on the distribution of host fishes on a regional scale. Diversity Distrib. 19(4):446–454.

Stewart KW, Baumann RW, Stark BP. 1973. The distribution and past dispersal of southwestern United States Plecoptera. Trans Amer Entomol Soc. 99:507–546.

Szytko SW, Stewart KW. 1977. The stoneflies (Plecoptera) of Texas. Trans Amer Entomol Soc. 103: 327–378.

Vilmi A, Karjalainen SM, Hellsten S, Heino J. 2016. Bioassessment in a metacommunity context: Are dia-tom communities structured solely by species sorting? Ecol Indic. 62:86–94.

Wentworth CK. 1922. A scale of grade and class terms for clastic sediments. J Geol. 30(5):377–392.

Williams DD. 2006. **The biology of temporary waters**. Oxford: Oxford University Press.

Winegardner AK, Jones BK, Ng IS, Siqueira T, Cottenie K. 2012. The terminology of metacommunity ecology. Trends Ecol Evol (Amst). 27(5):253–254.

Zawalski R. 2017. Benthic macroinvertebrate community structure of the Guadalupe River Basin, San Marcos, TX. Texas State University.

Zhang Y, Zhang J, Wang L, Lu D, Cai D, Wang B. 2014. Influences of dispersal and local environmental factors on stream macroinvertebrate communities in Qinjiang River, Guangxi, China. Aquat Biol. 20(3):185–194.

Zhao K, Song K, Pan Y, Wang L, Da L, Wang Q. 2017. Metacommunity structure of zooplankton in river networks: roles of environmental and spatial factors. Ecol Indic. 73:96–104.