OPEN

Niche partitioning of microbial communities in riverine floodplains

Marc Peipoch [1*], Scott R. Miller[2], Tiago R. Antao[2] & H. Maurice Valett[2]

Riverine floodplains exhibit high floral and faunal diversity as a consequence of their biophysical complexity. Extension of such niche partitioning processes to microbial communities is far less resolved or supported. Here, we evaluated the responses of aquatic biofilms diversity to environmental gradients across ten riverine floodplains with differing degrees of flow alteration and habitat diversity to assess whether complex floodplains support biofilm communities with greater biodiversity and species interactions. No significant evidence was found to support a central role for habitat diversity in promoting microbial diversity across 116 samples derived from 62 aquatic habitats, as neither α (H': 2.8–4.1) nor β (Sørensen: 0.3–0.39) diversity were positively related to floodplain complexity across the ten floodplains. In contrast, our results documented the sensitivity of biofilm communities to regional templates manifested as gradients of carbon, nitrogen, and phosphorous availability. Large-scale conditions reflecting nitrogen limitation increased the relative abundance of N-fixing cyanobacteria (up to 0.34 as fraction of total reads), constrained the total number of interactions among bacterial taxa, and reinforced negative over positive interactions, generating unique microbial communities and networks that reflect large-scale species sorting in response to regional geochemical gradients.

Unraveling the biogeochemistry[1–3], trophic significance[4–6], and biodiversity[7,8] of microbial assemblages attached to hard substrata of riverbeds–aka epilithic biofilms–has been a dominant focus of aquatic ecology, leading to recognition of epilithic biofilms as key control agents of ecosystem processes. The structure of river biofilms have been examined at several spatial scales, from species segregation within a few nanometers[9,10] to differentiation of biofilm functions along river sections[11,12], and landscape-scale responses of biotic constituents to changes in ecosystem integrity[13]. Inasmuch as biofilm assemblages are becoming increasingly scrutinized at multiple scales and scenarios, little is known about the assembly of biofilm communities among and within riverine floodplains, one of the most iconic elements of fluvial landscapes.

Adjacent to streams and rivers of different sizes, well-preserved floodplains are the paradigm of landscape heterogeneity, blending terrestrial and aquatic habitats as a mosaic of patches characterized by high turnover rates[14–16]. Habitat complexity of this kind gives rise to remarkable gradients of environmental conditions and resource availability across multiple spatial scales[17,18]. Complex floodplains are carved by cut-and-fill alluviation that sustains a variety of ecological niches and biological species, making them one of the world's most biologically diverse ecosystems in relation to their area[15,19]. Floodplain complexity has been shown to enhance biodiversity of plants[20], invertebrates[21,22], fish[23,24] and birds[25]. Some have argued that river floodplains are the "primary arena" for species interactions across vast regional landscapes[26]. This perspective relies on a direct association between niche heterogeneity and species diversity as a fundamental mechanism explaining enhanced species richness and distribution across riverine floodplains, a relationship that has been conceptualized and tested only for organisms of much larger size than microbes.

Whether it is appropriate to transfer ecological theory developed from and for macro-organisms (*i.e.*, those visible to the naked eye) to the ecology of microbial communities is an ongoing debate[27,28]. Differences in dispersal capabilities and rates of local adaptation are undoubtedly to be given weight when rationalizing how micro- and macro-organisms may respond to habitat complexity and environmental gradients presented by riverine floodplains (*i.e.*, niche partitioning). For instance, hydrologic regimes are known to exert powerful controls on invertebrate[29,30] and fish[31] communities by favoring specific lifecycles, but large-scale interactions between hydrology and biofilm communities are far less predictable and less well understood[32]. As the ecotone between terrestrial and aquatic landscape elements, floodplains include different 'zones' of organization. Main channel habitats (*e.g.*, riffles, pools, shallow shorelines) occur within a zone of transport and connectivity linking upstream and downstream segments[16,33]. Off-channel habitats include parafluvial elements residing near and frequently disturbed by flowing

[1]Stroud Water Research Center, Avondale, PA, USA. [2]Division of Biological Sciences, University of Montana, Missoula, MT, USA. *email: mpeipoch@stroudcenter.org
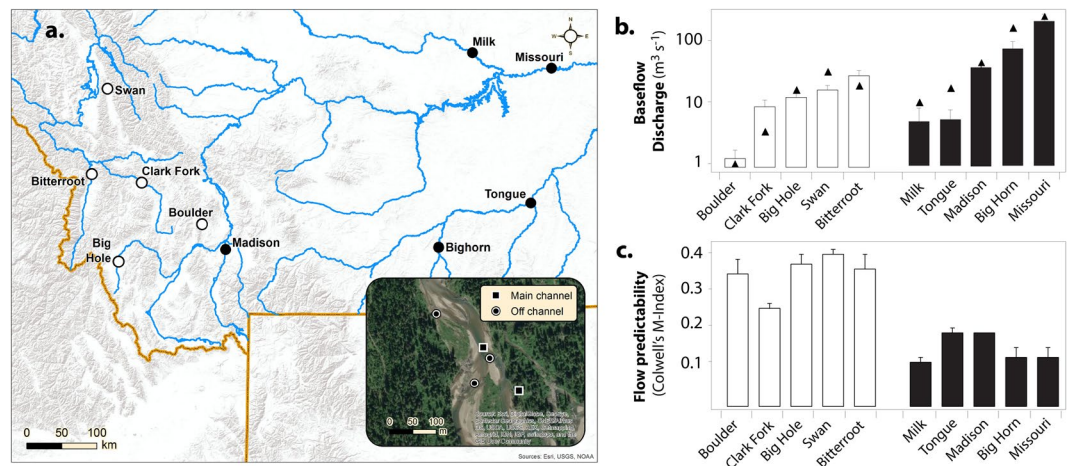
Figure 1. River location and flow. Map of Montana, USA, showing floodplain locations and pictorial representation of main- and off-channel sampling locations on the Swan River (**a**). Average baseflow discharge (m3/s) for the 1970-2017 period; triangles indicate river discharge during sampling; white bars correspond to floodplain locations in relatively free-flowing rivers, black bars correspond to river floodplains located downstream of one or more impoundments (**b**). Index of the predictability of flooding events in each river floodplain, color coding as in panel b, error bars were generated from 999 bootstrapped samples (**c**). Satellite imagery were obtained from World Imagery (http://goto.arcgisonline.com/maps/World_Imagery) and modified using ArcGIS 10.3 (https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview). Credits attribution for both satellite and terrain imagery are Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community.

water and orthofluvial patches with vegetation development characteristic of later stages of succession due to infrequent exposure to flooding disturbance. Springbrooks, side channels, and ponds represent a diverse array of niches in the off-channel zone. While these floodplain habitats include a plethora of environmental opportunities[34,35], such habitat heterogeneity occurs over hundreds (if not thousands) of meters and may last several days or weeks. Considering the great dispersal capabilities and rapid rates of adaptation characteristic of microbial communities like epilithic biofilms, the extent to which biofilm assemblages reflect and respond to environmental heterogeneity occurring at these spatial and temporal scales is unclear. Hence, it remains to be seen how variation in the degree of habitat and environmental heterogeneity among complex floodplains translates into more 'complex' benthic biofilm communities with greater numbers of microbial taxa and species interactions.

Despite uncertainties surrounding the general sensitivity of microbial diversity to environmental gradients[36], biogeochemical controls on river biofilms have been identified at multiple levels of ecological organization including microhabitats[37], stream reaches[38], and watersheds[13], all with significant effects on the microbial diversity of biofilm communities. More recently, a growing body of research has begun to incorporate network analysis based on microbial co-occurrence patterns (*i.e.*, microbial association networks) to better understand interactions among microbial taxa, community-level responses to environmental conditions, and to elucidate keystone taxa in microbial communities[13,28,39]. Network analysis is a powerful approach for identifying spatial associations among microbial taxa, especially when communities are assessed across pronounced environmental gradients to which different taxa may respond in similar or opposite manners. Some authors have argued that network analysis is more suitable than diversity metrics (*e.g.*, alpha or beta diversity) when dealing with complex microbial datasets[40]. Interpreting microbial interactions and identifying keystone taxa by exploring microbial association networks may be particularly well suited for riverine floodplains given the degree of habitat heterogeneity that floodplains convey within small areas. Concomitantly, these approaches allow investigation of hierarchical controls over spatial patterns of diversity presented by the occurrence of a variety of floodplains within river networks draining diverse elements of regional landscapes.

Here, we compare environmental heterogeneity, microbial diversity, and co-occurrence microbial networks across riverine floodplains with differing flow regimes and habitat diversity. Ten riverine floodplains were sampled using a hierarchical approach to assess how biofilm assemblages respond to environmental gradients at biome, floodplain, zone, and habitat scales. Our goals were to 1) determine the extent to which complex riverine floodplains with greater habitat heterogeneity and more pronounced environmental gradients support biofilm communities with greater biodiversity and more extensive biological interaction (*i.e.*, well connected biofilm networks), 2) document the occurrence of keystone taxa among epilithic assemblages, and 3) identify and characterize critical scales or organization for biofilm communities in riverine floodplains.

Results

Floodplain locations and river flow.   We selected ten riverine floodplains distributed across montane (n = 6) and grassland (n = 4) biomes within the Rocky Mountains and Great Plains of Montana, USA (Fig. 1a). Average baseflow discharge ranged from 1 to > 100 m³s⁻¹ across rivers and biomes (Fig. 1b). Half of the floodplains, including all grassland systems, are located downstream of large scale impoundments, while the other half,

Parameter	Biome	Floodplain	Zone	Habitat	Residual
Dissolved Organic Carbon	0	63	4	13	20
Ammonium	0	0	13	0	87
Nitrate + Nitrite	0	60	4	2	34
Dissolved Organic Nitrogen	0	64	3	3	30
Total Nitrogen	11	63	9	0	17
Soluble Reactive Phosphorous	0	73	0	21	6
Total Phosphorous	0	87	1	9	4
molar N:P ratio	0	65	7	3	25
Benthic chlorophyll-a	0	8	0	13	79
Benthic organic matter	0	0	8	19	73

Table 1. Variance Component Analysis. Values represent the estimated percentage of variance explained by each of the spatial scales and for each biogeochemical variable. Residual variance represents unexplained variation.
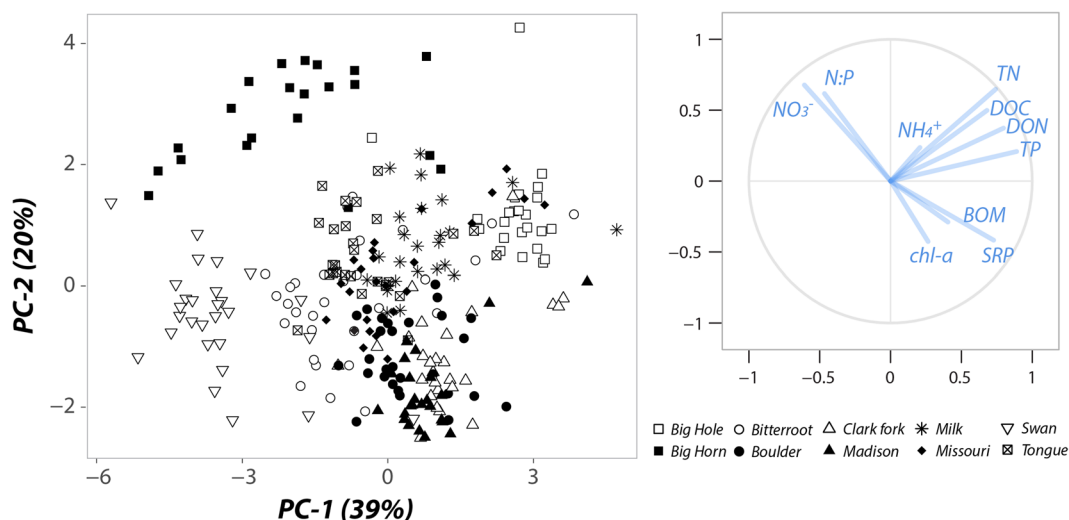

Figure 2. Environmental variation in riverine floodplains. Principal Components Analysis (PCA) of ten biogeochemical variables characterizing environmental heterogeneity for all habitats within each of the ten riverine floodplains. Left panel shows all PCA-scores (n = 265 habitats) along PCA-1 and -2 components. Unique symbol shapes represent individual floodplains. Right panel indicates the PCA loadings (biogeochemical variables) corresponding to each of the ten biogeochemical parameters measured in each sampled habitat.

mostly associated with montane rivers, are in relatively free-flowing waters (Fig. 1a,b). While extensive examination of flow regimes is beyond the scope of this study, Colwell's M-Index[41] was used to describe the effects of impoundments on flow seasonality. The index is minimal when the probabilities of occurrence for river flows of a given magnitude are independent of the season, *e.g.*, less predictable flooding seasonality. Flow predictability was significantly lower with river dams present (0.11 ± 0.01) than when absent (0.27 ± 0.02; Fig. 1c). Long-term discharge data in free-flowing rivers show recurrent spring, snowmelt-driven floods characteristic of western USA, contrasting with more irregular seasonality in river flow for those floodplains under the influence of upstream dams (Fig. S1).

Environmental variation across riverine floodplains. Most of the observed variation in water chemistry was attributable to differences among individual floodplains rather than at scales above or within them (Table 1). Average concentrations of total nitrogen (TN) were greatest in the Bighorn, Missouri, and Milk Rivers (Table S1), where flow regime is also more heavily altered (Fig. 1c). The Big Hole River had the greatest concentrations of dissolved organic forms of carbon (C) and nitrogen (N), while the Swan River showed minimum concentrations for all chemical constituents except nitrate (N-NO$_3^-$; Table S1). Similar conditions of high phosphorus (P) availability and low N-NO$_3^-$ concentrations in the Clark Fork, Boulder, and Madison Rivers, suggest potential N-limitation of biological activity in aquatic habitats of these floodplains (Table S1). Variation in both biofilm biomass and chlorophyll-a abundance was mostly unexplained (*i.e.*, residual variance 73 and 79%, respectively, Table 1), but variance component analysis revealed meaningful influence at the habitat-scale (Table 1). Ordination analysis separated riverine floodplains along its first component (PC-1) following a gradient of dissolved organic carbon (DOC) and total nutrient concentrations (Fig. 2). In contrast, PC-2 captured variation in bioavailable N,
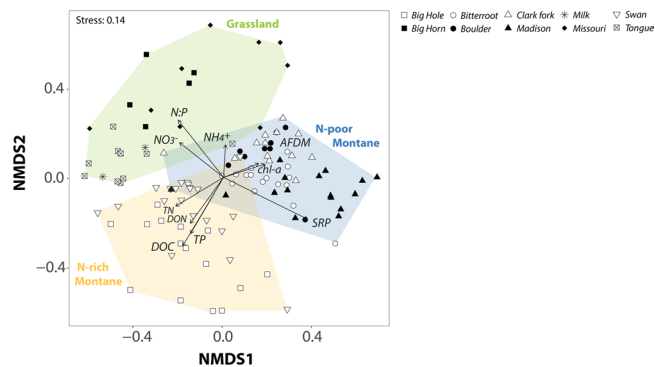
Figure 3. Bacterial community composition among floodplains. Non-metric Multidimensional Scaling (NMDS) based on Bray-Curtis dissimilarity. Data points correspond to biofilm samples (shape coding as in Fig. 2). Vectors correspond to projected environmental variables by finding directions in NMDS space with maximal correlation to each environmental variable. Polygons illustrate the grassland biome (green) and the two montane sub-biomes, N-rich (orange) and N-limited floodplains (blue). Statistical differences among clusters are detailed in Table S2.
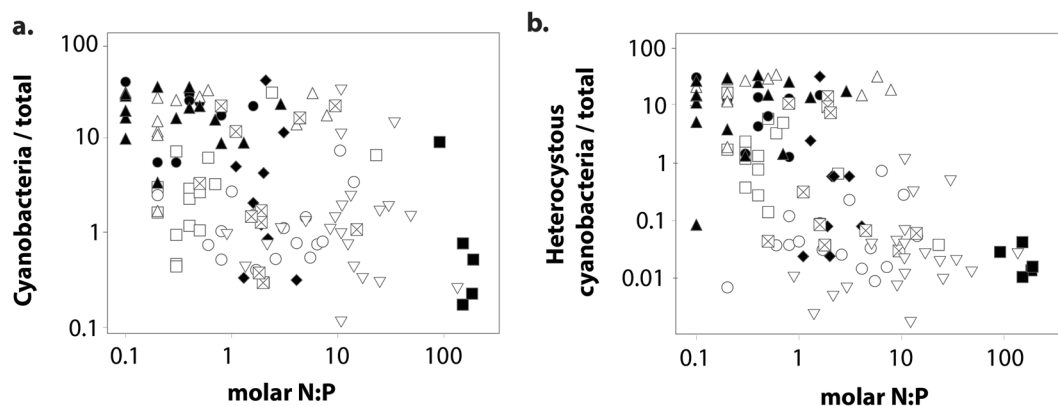

Figure 4. Relative abundance of cyanobacteria versus N:P. Sequence reads of (**a**) cyanobacteria and (**b**) heterocystous cyanobacteria are expressed as a fraction of total sequence reads. Data points correspond to biofilm samples (shape coding as in Fig. 2).

from high-NO_3 waters (Bighorn, Big Hole, or Swan) to more N-limited floodplains with high P concentrations and greater chlorophyll-a abundance (Clark Fork, Boulder, Madison, Bitterroot; Fig. 2 and Table S1).

Biogeochemical variation within each floodplain (*i.e.*, at zone and habitat scales) was widely spread across PC-1 and PC2 components depending on the floodplain (Fig. 2), which is in agreement with a substantial percentage of variation in DOC, P, chlorophyll-a abundance, and BOM attributed to the habitat scale (Table 1). Among habitats, highest concentrations of inorganic forms of N and P were observed in slow-moving waters (*e.g.*, springbrooks) whereas TN and DON concentrations were significantly greater in floodplain ponds (Table S1). Overall, flow predictability (Colwell's M-Index) was positively related to both habitat diversity (Shannon's H, R = 0.74, Fig. S2-a) and environmental heterogeneity (CV for habitat centroids, R = 0.88, Fig. S2-b) across the ten riverine floodplains, indicating that floodplains in relatively free-flowing rivers should provide greater niche diversity for biofilm communities.

Biofilm diversity and composition across riverine floodplains. Bacterial diversity (α-diversity as Shannon's H index) varied significantly among floodplains ($F_{(8,107)} = 5.07$; $P < 0.001$) but not among biomes, zones, or habitat types (Table S2). Shannon's index ranged from 2.8 in the Missouri River (503 bacterial families) to 4.1 in the Tongue River (660 bacterial families [Table S2]). Significant differences in β-diversity were also observed among floodplains ($R^2 = 0.2$; $F_{(8,686)} = 18.67$; $P < 0.001$) but not at any other spatial scale (Table S2). At the floodplain level, most of the β-diversity reflected species replacement among habitats ($\beta_{turnover}$) rather than a nested loss of bacterial taxa from rich to poor habitats (*i.e.*, $\beta_{nestedness}$ [Table S2]). Species turnover in biofilm communities across floodplains was dominated by increased relative abundance of *cyanobacteria* within select floodplains (up to 34% of total reads), with a 3-fold increase in cyanobacteria contributions to β-diversity in N-limited floodplains (average total of 33 ± 5%) compared to in N-rich floodplains (9 ± 4%). Differences in chlorophyll-a abundance among floodplains (Fig. 2) reflected variation in the relative abundances of both cyanobacteria and eukaryotic algae (Fig. S3-a). Eukaryotic algal composition varied among rivers (Fig. S3-b), but was dominated by photosynthetic stramenopiles (*i.e.*, heterokont algae). Diatoms comprised 87% of the stramenopile sequence

| Semi-biome	Nodes (no.)	Edges (no.)	Node Degree	Closeness Centrality	Betwenness Centrality	Keystone Taxa (no.)	Positive edges (%)	Negative Edges (%)	Modularity
Montane (N-limited)	178	781	4 (2, 13)	0.32 (0.27, 0.38)	0.0008 (0.000, 0.019)	23	72 (335, 229)	28 (91, 124)	0.374
Montane (N-rich)	273	3669	12 (3, 45)	0.36 (0.29, 0.41)	0.003 (0.000, 0.011)	32	99 (2773, 597)	1 (6, 33)	0.212
Grassland	299	3438	15 (4, 39)	0.36 (0.31, 0.42)	0.004 (0.000, 0.009)	44	98 (2773, 597)	2 (51, 17)	0.286

Table 2. Biofilm network parameters. Central tendencies and variation of the computed network parameters to quantify network complexity in the grassland biome (n = 27), montane N-rich sub-biome (n = 37), and montane N-limited sub-biome (n = 54). Values for average degree, closeness centrality, and betwenness centrality are expressed as median (25th, 75th percentile). For the percentage of positive and negative interactions (edges), the amount of edges connecting OTUs associated with the same floodplain zone and the amount of edges connecting OTUs associated with different zones (main- and off-channel) are also indicated within parenthesis.

reads for the full data set and were the most abundant members of this group in most rivers (Fig. S3-c), although other stramenopiles (*e.g.*, brown algae) were important contributors to the diversity of specific floodplains.

Bacterial community composition varied significantly between grassland and montane biomes (ANOSIM-R: 0.279) and more so across floodplains (ANOSIM-R: 0.679). Unconstrained ordination revealed distinct communities at more N-limited floodplains (Clark Fork, Madison, and Boulder), significantly different from those of other montane rivers or in grassland floodplains (Fig. 3; Table S3). As such, bacterial communities in montane floodplains separated in two distinct groups along a gradient of DOC, N, and P availability, from high NO_3 and DOC concentrations in Swan and Big Hole Rivers (hereafter referred to as N-rich conditions) to more N-limited floodplains with greater P availability and chlorophyll-a abundance (Fig. 3; Table S1). The second NMDS axis separated bacterial communities between grassland floodplains and the two groups in the montane biome (Fig. 3). Together, these results suggest the occurrence of three sub-biomes experienced by epilithic biofilms of the rivers that flow within these regional geochemical gradients. Significant differences in both environmental conditions (ANOSIM-R: 0.215–0.415, *p*-value < 0.01) and bacterial community structure (ANOSIM-R: 0.215-0.687, Table S3) were observed among the three sub-biomes supporting the patterns observed in NMDS ordination. Community differentiation of a N-poor montane sub-biome from N-rich montane floodplains and the grassland biome is likely related to the increasing relative abundance of cyanobacteria in biofilm communities as the N:P ratio of the waters in which they reside declines (R = −0.46, *p*-value = 0.018, Fig. 4a). On average, N:P ratios in the N-poor sub-biome were an order of magnitude lower (1.8 ± 0.4) than in the N-rich sub-biome (12.7 ± 3.9) and grassland biome (30.7 ± 11.8). Variation in the relative abundance of a clade of multicellular cyanobacteria that produce specialized nitrogen-fixing cells (heterocysts) when N is limiting was also largely related to the floodplain in which they occurred ($R^2 = 0.71$; $F_{(8,105)} = 31.92$; $P < 0.0001$). This pattern principally reflects differences in N:P ratios among rivers, since the proportion of heterocystous cyanobacterial reads decreased with increasing N:P ratios (R = −0.63, *p*-value = 0.009, Fig. 4b;). Only marginal differentiation of biofilm communities was observed between zones (ANOSIM-R: 0.051) and not at all among habitats (P > 0.05). Low variation in bacterial community similarity within floodplains concur with greater variation in water chemistry among floodplains than within them (Table 1), including the most significant predictors of variation in bacterial community similarity.

Environmental controls on biofilm communities. Constrained ordination showed that concentrations of DOC and P (as soluble reactive phosphorous [SRP]) were the most significant predictors of variation in bacterial community similarity, with each accounting for a similar portion of the total variation explained by the model (Table S4-a). Phosphorous concentration was significantly correlated to the relative abundance of cyanobacteria across floodplain habitats (R = 0.58, P < 0.05), and this may be an avenue by which P abundance influenced overall differences in bacterial community composition. Total N and N:P ratios were also identified by the model as significant predictors, but with much lower proportions of variation in community structure attributed to them (Table S4-a). Once explanatory variables were selected, results of variance partitioning showed that geographic location was as good a predictor of bacterial community structure as the four selected environmental predictors combined, and the largest portion of the explained variation in community structure was ascribed to shared variation between geographic and environmental components (Table S4-b). Together, unconstrained and constrained ordination indicated that a combination of geographic location and biogeochemical condition accounted for the greatest amount of the variation in bacterial community structure at the habitat level (Table S4-b), although most variation remained unexplained (77.6%, Table S4-b).

Co-occurrence bacterial networks. As guided by results from unconstrained ordination, we computed a total of three co-occurrence bacterial networks congruent with recognition of the sub-biomes characterized by flow regimes and biogeochemical conditions. Generating independent networks for the combined rivers within the grassland biome and for the two biogeochemical settings in the montane biome allowed sufficient sample size and reliability for network analysis (Table 2). Total number of nodes (*i.e.*, OTUs) in the co-occurrence network of N-limited floodplains was lower than for bacterial networks of grassland or montane N-rich floodplains (Table 2; Fig. 5a–c). Even greater differences were found for the total number of interactions (*i.e.*, edges) among nodes, which tripled in grassland and N-rich communities compared to in N-limited bacterial communities (Table 2;
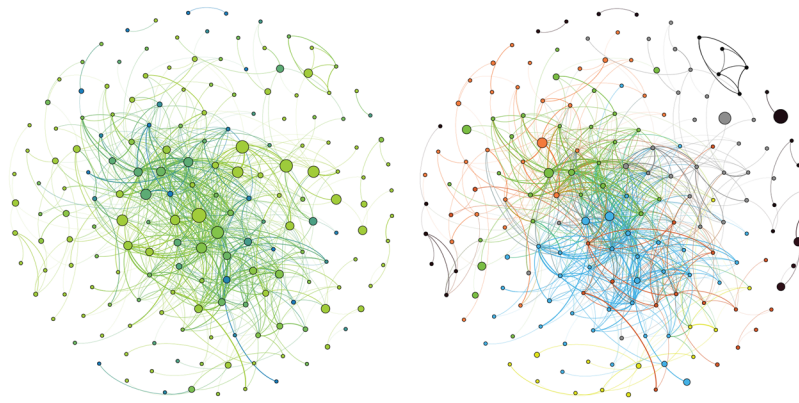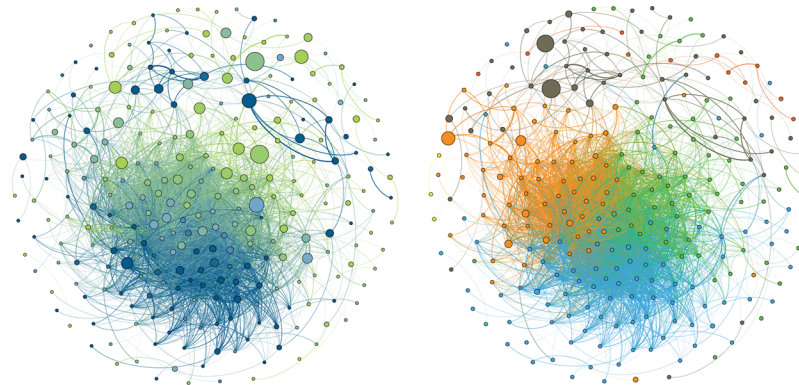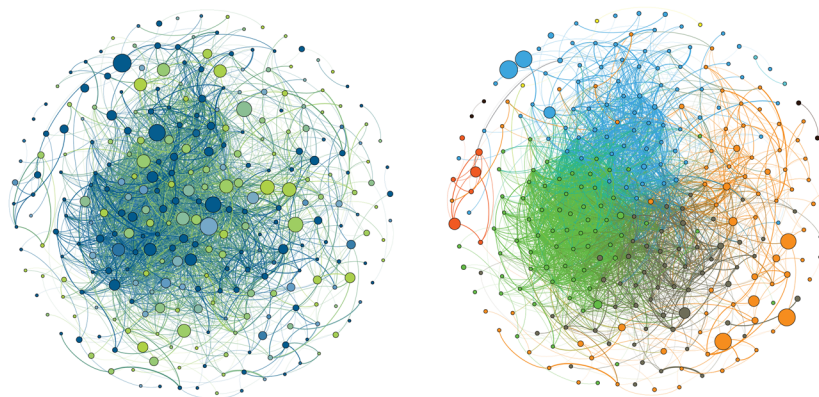
Figure 5. Co-occurrence microbial networks. Computed co-occurrence networks for river-floodplain systems in the N-limited (**a**), N-rich (**b**), and grassland (**c**) sub-biomes. Nodes in networks on the left are color coded to reflect relative abundance in main- and off-channel zones, while node sizes are proportional to values of betweenness centrality. Nodes in networks on the right are color coded to reflect the different bacterial modules to which they belong, with large (i.e., higher number of nodes) and small (representing less than 2.5% of the network nodes) modules in blue and black, respectively. Node size is proportional to OTU relative abundance. An interaction (edge) reflects a significant Pearson correlation ($-0.6 > R > 0.6$ and p-value ≤ 0.01) with edge thickness proportional to correlation coefficients.

Fig. 5). In grassland and N-rich montane networks, nearly every (98 & 99%, respectively) bacterial interaction was positive, whereas more than a quarter of the correlations found in the N-limited network were negative (Table 2). Moreover, at the zone level (*i.e.*, main-channel vs. off-channel) most of the interactions in grassland and N-rich montane floodplains involved taxa whose relative abundance co-varied positively across floodplain zones (Fig. 5b,c). In contrast, approximately half the bacterial interactions in N-limited floodplains involved pairs of OTUs with opposite zonal patterns in relative abundance (Fig. 5a).

Network connectivity–Four parameters were assessed to estimate node connectivity with the wider network including node degree, closeness centrality (CC), betweenness centrality (BC), and keystone species designation. Node degree represents the average amount of edges per node in each network, a measure much lower and less variable in N-limited biofilms than in the other networks (Table 2). Similarly, CC (measured as the inverse of total distance to other nodes) was lowest in N-limited communities indicating a larger fraction of OTUs poorly connected to the larger network (Fig. 5a–c). Median betweenness centrality (BC; fraction of times a node is found on the shortest path between two nodes) was also lowest in the N-limited network (Table 2), but a few nodes in this sub-biome showed a disproportionate influence on the interactions between other taxa (note high 75th percentile BC values in Table 2; Fig. 5a). The proportional abundance of keystone taxa (*i.e.*, nodes exceeding the 75th percentile of node degree, CC, and BC) was similar among networks (12–15% of all nodes; Table 2) without propensity to occur within a specific floodplain zone, and thus indicating that keystone taxa are equally likely to be found in main channel and floodplain habitats (Fig. 5a–c; Table S5). The relative abundance of keystone taxa in each habitat (as fraction of total reads) was positively correlated with sample taxon richness in grassland (R = 0.83; *p*-value < 0.001) and N-rich (R = 0.87; *p*-value < 0.001) communities, but not in N-limited floodplains (Fig. S4).

Network modularity–The degree to which a network is composed of sets of nodes highly connected to each other but poorly connected to other sets describes network modularity. A total of 7 modules were identified in each network encompassing more than 95% of the total number of nodes. Greater modularity occurred in N-limited biofilms compared to N-rich and grassland communities (Table 2). High network modularity in N-limited conditions is illustrated by fewer connections among taxa belonging to different modules (Fig. 5d), contrasting with the numerous linkages across modules in N-rich and grassland assemblages (Fig. 5e,f). With few exceptions, more abundant community members were included in one or more of the larger network modules (*i.e.*, those that include a greater number of nodes). However, in grassland and N-rich communities, abundant taxa were rather poorly connected (low node degree and BC) and located at the edge of their respective modules (Fig. 5d–f), while most abundant taxa in N-limited communities were generally more strongly connected within each module (Fig. 5d). For the N-limited floodplains, two abundant Cyanobacteria families–unnamed families but one mostly representing *Leptolyngbya* and *Microcoleus*, and the other *Calothrix* and *Nostoc* genera– were the most well-connected taxa in the second largest module (orange in Fig. 5d). One of these cyanobacteria families was also identified as a keystone taxon (Table S5) and its abundance was strongly correlated (R = 0.73; P-value < 0.001) with the relative abundance of heterocystous cyanobacteria across all biofilm samples. Together, these results suggest a central role of N-fixing cyanobacteria in microbial networks of river-floodplain systems under N limiting conditions (Figs 5d and S5).

Discussion

As human pressures on riverine landscapes continue to increase, understanding the consequences of floodplain simplification for freshwater biodiversity is of high priority[42]. Hypothesized losses of species diversity due to degradation of floodplain habitat diversity are supported by extensive evidence from a variety of taxa[43–45]. But the validity of this hypothesis when applied to microbial communities is far less resolved and was the main rationale for this study. We gathered little evidence for niche differentiation at the habitat scale that leads to unique microbial communities within floodplains. Biophysically complex floodplains supported alpha- and beta-diversity similar to that in grassland floodplains with limited habitat heterogeneity. The lack of significant differentiation of biofilm communities among the array of aquatic habitats surveyed contrasts with the occurrence of distinct biogeochemical mixtures of C, P, and N as central determinants for biofilm community differentiation within and across riverine floodplains. Hence, our results do not generally support the idea that increased habitat diversity and environmental variation in riverine floodplains lead to more diverse biofilm communities. Instead microbial composition was organized at the scale of sub-biomes and among the floodplains within them, reflecting biogeochemical and geographic character of entire river-floodplain systems and across large, regional scales that contain them.

Partitioning biofilm communities into universal groups may be impossible to determine objectively, but practical divisions can prove informative. Identifiable communities based on assemblage structure and interactions across landscapes have been used, for instance, to describe how catchment properties relate to changes in bacterial communities in a context of microbial biogeography[13,46]. Our results suggest a similar approach may be taken to assess regional influences on microbial communities of floodplain landscapes. Biofilm community composition in riverine floodplains is driven by regional gradients in geochemical conditions more than endogenous environmental gradients provided by floodplain habitat diversity. Of central importance to that geochemical template is the relative and absolute availability of N and P, fundamental resource pools that may act as master variables for species sorting. Regional flow regimes and gradients of nitrogen limitation exerted strong influences on species sorting and interactions of biofilm assemblages, emphasizing spatial niche differentiation at large scales.

River flow, floodplain complexity, and biofilm diversity.
River impoundments alter natural flow regimes and reduce lateral hydrologic connectivity while decreasing the abundance and diversity of floodplain habitats such as backwaters, springbrooks, and ponds[18,42,47]. By selecting floodplains in relatively free-flowing rivers and others located downstream of dams, we intended to capture a gradient of habitat complexity and environmental heterogeneity with subsequent implications for biofilm communities. Our simple analysis of flow

alteration showed that river dams reduced flow predictability, a major organizing factor for lotic communities[30]. As expected, declining flow predictability was strongly correlated to reduction in the diversity of aquatic habitats found in each floodplain, and to biogeochemical variation among habitats where epilithic biofilms develop. Together these patterns are indicative of river and floodplain simplification associated with river dams[18]. Results support our initial premise that more natural flow regimes translate into greater floodplain complexity and stronger environmental gradients, but increased environmental heterogeneity alone did not strongly translate to greater bacterial co-existence and enhanced microbial biodiversity.

Biofilm diversity (as average α-diversity) was highest in more regulated rivers (*e.g.*, Big Horn River) with lower habitat diversity and biogeochemical variation, and the degree of community differentiation (β-diversity) was unrelated to the range of abiotic conditions observed within each floodplain. These results suggest similar degrees of community differentiation among bacterial taxa with varying environmental tolerances and among riverine floodplains of contrasting habitat diversity. Hence, processes of niche partitioning promoting species co-existence for communities of macro-organisms of floodplain landscapes[20–25] may not equally apply to biofilm communities. One may argue that niche partitioning in biofilm assemblages may be obscured by scale effects on ecological phenomena. Previous work at small scales (10^{-1} m) has demonstrated how increased flow and substrate heterogeneity can enhance both β-diversity[37] and functional diversity[48,49]. Heterogeneous habitat across floodplain landscapes manifest at spatial scales in the order of 10^2–10^3 m, likely larger than the spatial dimension evoking microbial responses[50]. Although our experimental designed allowed us to compare habitat and bacterial diversity across biome, floodplain, zone, and habitat scales, it is possible that our finest level of resolution was too coarse to discern niche differentiation in microbial assemblages[36,51]. However, the argument of a large-scale decoupling between environmental variation and microbial responses contrasts with regional abiotic controls on aquatic biofilms found in this and previous studies on similar ecosystems and scales[13,52,53].

Environmental controls on biofilm assemblages of riverine floodplains.   Differences in community structure were clearly greater among floodplains than at any other spatial scale. Similar patterns were found for the strongest predictors of biofilm community structure (*i.e.*, DOC and SRP) that differed more strongly across floodplains than within them. Riverine floodplains represented distinct spatial units distributed across regional environmental gradients, leading to ubiquitous occurrence of bacterial taxa with similar environmental tolerances[52] despite quantifiable heterogeneity and complexity at the habitat scale. Recognizing it as the fundamental energy source to heterotrophic microbes, a great number of studies have demonstrated how quality and quantity of DOC can significantly affect biofilm community composition[54], as well as microbial contribution to ecosystem processes[55,56]. Biogeochemical distinctions were also tied to SRP and N availability manifested as regional gradients of N to P ratios which led to changes in the abundance of cyanobacteria and eukaryotes, and properties of microbial co-occurrence networks. Our results illustrated striking dissimilarity in community structure between montane sub-biomes of contrasting N and P availability. Regional variation in P concentrations can be attributed to the existence of P-rich geologic formations in southwest Montana[57], leading to riverine systems perennially enriched in SRP, promoting N-limitation of biological processes when N demand increases[58]. This was the case for the Boulder, Clark Fork, and Madison Rivers, and to a lesser extent, the Bitterroot River. Under conditions of high water temperature, high P availability and low N concentrations, heterocystous cyanobacteria with the ability to fix atmospheric-N heavily influence the structure and function of lotic ecosystems[59,60]. Significant increases in N-fixing cyanobacteria, distinct community structure, and unique network topology in SRP enriched (*i.e.*, low-N) systems suggest a critical role for N-limitation in structuring composition, function, and negative species interactions of biofilm communities.

Nitrogen limitation and microbial networks in river floodplains.   Microbial communities in floodplain landscapes are proposed to function as meta-communities[61,62], in which local extinctions due to direct competition can be rapidly overcome by recolonization from nearby habitats[63]. Hence, niche differentiation in the form of environmental control on community assembly should have strong influences on local community structure[64]. Co-occurrence of microbial taxa across floodplain habitats is likely to capture those unique responses of microbial taxa to local environmental conditions. Positive co-occurrence is thought to reflect interactions between ecologically similar taxa[65] that require similar environmental settings and resources[66], and likely perform similar or complementary functions. By contrast, negative interactions are indicative of direct competition for resources among microbial taxa, or the result of resource partitioning[67]. Agler, *et al.*[68] provided compelling results supporting these interpretations by showing how positive co-occurrence accounted for 86% of the interactions among microbial taxa of the same kingdom (fungi or bacteria), while only 23% of cross-kingdom interactions were positive. A dominance of positive interactions in microbial networks has been commonly observed in benthic biofilms[53,69], plant microbiomes[68], soil communities[70,71], or marine bacterioplankton[67]. In our study, positive interactions dominated each of the three computed networks, but more notably when N was not limiting biological activity. We relate the increased proportion of negative interactions in N-limited floodplains to increased resource competition along extreme gradients of N-limiting conditions as described by N:P ratios that differed by orders of magnitude within and among floodplains. Nitrogen fixation is highly sensitive to changes in abiotic controls with rapid, nonlinear responses when thermal or nutrient thresholds are exceeed[59,72,73], providing a significant competitive advantage to N-fixing bacteria when favorable conditions (high temperature, P, and low N) are met[52]. Further effects of increased resource competition on N-limited biofilms may be the lack of positive relationships between keystone taxa abundance and bacterial richness observed in other N-rich rivers of grassland and montane biomes. Positive co-occurrence exists among keystone species and other taxa directly or indirectly affected by them, likely taxa requiring similar environmental settings or performing complementary functions[65,68], resulting in overall increases in taxa richness[74]. But in riverine floodplains affected by regional constraints of low N availability and concomitant P enrichment, patterns of bacterial richness were

more likely influenced by local increases in the relative abundance of N-fixing cyanobacteria that possessed a clear competitive advantage in the use of the limiting nutrient resources. This interpretation is further supported by increased network modularity in N-limited floodplains and the central role of cyanobacteria in the largest modules (sub-networks). Two of the most abundant cyanobacteria families (representing 10% of total reads) were found in the second largest module and included many genera with the ability to fix N (*e.g.*, *Nostoc*). Networks with more modular structures result from the emergence of distinct network units (*i.e.*, modules) coping with resource constraints and unique niche opportunities[75,76], which can lead to greater abundance of negative relationships in the network[77] as we observed in N-limited communities. Our results indicate that effects of niche partitioning associated with N-limiting conditions can lead to increased modularity of epilithic biofilm communities[78], in which N-fixing cyanobacteria may play a central role in determining assembly and network structure for biofilm communities when regional constraints on N and P availability increase resource competition among and within floodplain habitats of N-poor sub-biomes.

Methods

Field sampling.   A nested sampling design was employed to capture spatial organization of river-floodplain systems at different scales. First, a *ca.* 10 km of floodplain in each river system was selected and equally divided as upper, middle, and lower sections that were sequentially surveyed over a three-day period. Within each section, lateral affiliation was characterized by 'main-channel' and 'off-channel' zones. Five aquatic habitats within each zone were sampled when present: *riffles*, *runs*, *pools*, *channel confluences*, and *shoreline* within the boundaries of main-channel zone; and *backwaters*, *side channels*, *parafluvial springbrooks*, *orthofluvial springbrooks*, and *ponds* in the off-channel zone with varying distances to the river's main channel. Detailed definitions of each habitat type can be found in Peipoch, *et al.*[52]. A total of 265 aquatic habitats were surveyed for biogeochemical characterization (details described below) during baseflow conditions (Fig. 1b). Epilithic biofilms were present in approximately half of the surveyed habitats (n = 126), from which composite samples for 16S- rRNA sequencing were collected.

Four out of ten rivers [*Bitterroot*, *Clark Fork*, *Boulder*, and *Madison]* were sampled during July-August of 2013 and six [*Swan*, *Big Hole*, *Bighorn*, *Tongue*, *Milk*, and *Missouri*] were collected during June-August of 2014. Water samples for inorganic and total dissolved nutrients were collected in triplicate at each habitat and immediately filtered (0.7-μm pore size, glass fiber filter, Whatman International, Kent, UK), kept on ice during transport, and frozen until further analysis in the laboratory. Samples for the analysis of chlorophyll and organic matter (i.e., ash-free dry mass, AFDM) were collected as composite scrapes from three to five different cobbles of known area within each habitat and stored on ice under dark conditions until further processing in the laboratory. Biofilms from 15–20 different locations within sampled habitats were carefully scrapped from rock surfaces using sterile disposable spatulas to generate a composite sample that was stored in 2-mL Safe-Lock microcentrifuge tubes (Eppendorf AG, Germany). Samples were kept on ice during transport centrifuged at 13000 rpm for 5 minutes to remove excess water, and frozen at −20 °C until DNA extraction processing.

Analytical methods.   Dissolved organic carbon (DOC) concentrations were determined using an Aurora 1030 W TOC Analyzer (Oceanographic Int., College Station, Texas, USA) and chemical oxidation methods[79]. The concentrations of ammonium-nitrogen (N-NH$_4$), nitrate-nitrogen (N-NO$_3$), and Soluble Reactive Phosphorous (SRP) were determined by micro-segmented flow analysis using an Astoria 2 Analyzer (Astoria-Pacific Inc., Clackamas, Oregon, USA) following standard procedures[80]. Determination of total nitrogen (TN) and total phosphorous (TP) was conducted by performing an alkaline persulfate digestion to each sample followed by analysis of N-NO$_3$ and SRP using the same instrumentation and methods detailed above[81]. Biofilm samples for BOM as AFDM were dried at 100 °C for 24 h, weighed as dry mass, subsampled if necessary, combusted at 550 °C for 6 h and weighed as ash. Benthic chlorophyll-*a* was determined by macerating samples in 90% acetone for 24 h and measuring extract absorbance in a V-550 UV/VIS spectrophotometer (Jasco Inc., Easton, Maryland, USA) before and after acidification to allow proper corrections for phaeopigments[82,83].

DNA extraction and sequencing.   Deoxyribonucleic acid (DNA) was extracted from biofilms using the PowerBiofilm® DNA Isolation Kit (MO BIO Lab. Inc., Carlsbad, California, USA) following manufacturer instructions. The V4 region of the16S rRNA gene was amplified in triplicate using primers F515/R806 tagged with 12-base Golay codes[84]. Samples from rivers surveyed in 2013 and 2014 were sequenced using the Illumina MiSeq platform at Argonne National Laboratory and Genomics Core Laboratory of the University of Montana, respectively. Approximately 250 base-pair reads were generated from each direction, resulting in nearly the entire targeted region having double coverage. Reads were assembled, Organizational Taxonomic Units (OTUs) were generated based on 97% sequence similarity, and chimeras were removed using USEARCH v7.0. For data analysis we used both MOTHUR [https://aem.asm.org/content/75/23/7537.short] and QIIME2 [https://peerj.com/preprints/27295/]. For each application we did 3 analyses: for the samples in 2013, for the samples in 2014, and for a mix of both years. We did this because the amount of information available on the two different sampling events was quite different, with orders of magnitude more data being available for the sampling of 2014. Also the protocol followed was slightly different – the data were available demultiplexed from 2014 whereas 2013 required demultiplexing. The analysis with MOTHUR provided more consistent results than QIIME2 in that the relative proportions of the mixed analysis were in line with the observations from the analysis of the individual years. In addition, we also validated that the relative proportions estimated by MOTHUR and QIIME2 analysis were consistent for each of the years when taken in isolation. Overall, we followed a slightly adapted version of the MiSEQ SOP available at https://mothur.org/wiki/MiS. Prior to data analysis of the bacterial community, sequences identified as chloroplasts were removed from the dataset (but see further analysis of chloroplasts sequences below).

All OTUs present in less than a quarter of the total samples and/or with less than 20 sequences were also removed. Each sample was then rarefied to contain 10,000 sequences.

Then, and because MOTHUR does not provide taxonomic assignments for chloroplast sequences, we used QIIME to compare the relative abundances of different photosynthetic microorganisms. Following the identification of different algal groups with QIIME, we conducted downstream analyses of streptophyte and stramenopile algal OTUs, respectively. For streptophyte OTUs, we used local nucleotide BLAST (blastn) searches to distinguish plant versus algal OTUs; approximately 82% of streptophyte sequence reads (268,623/326,888) were derived from plants and were not included in further analyses. We also used blastn searches against the NCBI GenBank nr database to assign stramenopile OTUs to different classes (*e.g.*, Bacillariophyceae (diatoms), Phaeophyceae (brown algae) etc.).

Data analyses. Variables were log-transformed to meet normality and homodestacity requirements before any of the analysis performed. Variance component analysis for each environmental variable was performed using restricted maximum likelihood estimations with the 'remlVCA' of the *VCA* package in order to identify the specific spatial scale at which each biotic and abiotic variable was varying most. Principal Component Analysis (PCA) was performed using the 'nipals' function from the *ade4* package to characterize environmental variation within and among riverine floodplains. Environmental heterogeneity of each floodplain was calculated as the mean and coefficient of variation (CV) of distance from each habitat site to the floodplain group centroid in the PCA ordination using the 'betadisper' function of the *vegan* package. We explored biofilm community similarity at the Family level using Non-metric Multidimensional Scaling (NMDS) based on Bray-Curtis dissimilarity calculated by the 'metaMDS' function from the *vegan* package. Statistical differences in community similarity among groups were tested using the 'anosim' function, and environmental variables were fitted into the NMDS plot to facilitate interpretation and guide further analyses without incorporating environmental constraints using the 'envfit' function. Beta diversity (dissimilarity between communities of different sites) was calculated for each floodplain as the sum of its two components or phenomena, *nestedness* of species assemblages (*i.e.*, species in specific habitats being subsets of communities from richer habitats) and *spatial turnover* (species replacement among habitats), following procedures and methods detailed in the 'beta.pair.abund' function of the *betapart* package[85]. Contributions of specific taxa to floodplain's beta diversity were estimated using the 'beta.div' function of the *vegan* package.

We used partial Redundancy Analysis (pRDA) to assess environmental controls on biofilm community composition while accounting for spatial effects on community similarity. We first generated distance-based Moran's Eigenvector Maps (db-MEMs) following procedures described by Dray, *et al.*[86]. In brief, db-MEMs characterize spatial variation among sites by generating eigenvector-based variables (similar to PCA eigenvalues) that can be later used as predictors for community patterns. Geographic distances used as input for db-MEMs were calculated using Euclidean distances and UTM coordinates. Once spatial variables were obtained, we performed an RDA to exclude those that were not significantly explaining patterns of community similarity using the 'rda' and 'ordiR2step' functions from *vegan* package. Then, a pRDA was done using the 'varpart' function of vegan to estimate individual and shared explanatory power of environmental and geographic distances on biofilm community structure. Adjusted R^2 values were used to evaluate the specific contribution of environmental and spatial predictors of biofilm assemblages. Relationships among environmental heterogeneity, microbial diversity, and network properties at multiple scales were explored using regression analysis. All statistical analyses were performed using the R software version 3.5.0[87].

To ensure network analysis was performed out of nonrandom co-occurrence interactions we first calculated a C-score index[88] and compared it to a null model with fixed OTU/site frequencies using the 'ecospat.Cscore' function from *ecospat* package[89]. Then, co-occurrence interactions among microbial taxa were examined at a 90% identity threshold corresponding roughly to a taxonomic level of Family[90] for the sake of network visualization[28,77]. Valid co-occurrence associations were considered when pairwise correlations were equal or in excess of 0.6 or −0.6, and statistically significant (*p*-value < 0.01). All the necessary steps for network inference were performed in the R environment (http://www.r-project.org) using *igraph* package[91]. Networks were then visualized and examined using the gephi platform[92]. To assess network structure we calculated average node degree and network's modularity following Newman[93] procedures. Node degree corresponds to the average number of edges per node and modularity quantifies the abundance of modules (set of nodes highly connected to each other while poorly connected to nodes of other groups) in the network. We identified keystone taxa (aka hubs) in each network when their values of node degree, closeness centrality, and betweenness centrality exceeded the 75th percentile of each of those parameters. Closeness Centrality is inversely proportional to the sum of distances to each other node of the network and Betweenness Centrality is equal to the number of times a node is in between the shortest path between any two nodes in the network. Together, node degree, CC, and BC values are indicative of the propensity of an OTU to act as keystone taxa[94].

Sequencing data derived from this study are available as BioProject PRJNA551494at in https://www.ncbi.nlm.nih.gov/.

References

1. Battin, T. J. *et al*. Biophysical controls on organic carbon fluxes in fluvial networks. *Nature Geoscience* **1**, 95–100, https://doi.org/10.1038/ngeo101 (2008).
2. Romani, A. M. & Sabater, S. Structure and activity of rock and sand biofilms in a Mediterranean stream. *Ecology* **82**, 3232–3245, https://doi.org/10.1890/0012-9658(2001)082[3232:saaora]2.0.co;2 (2001).
3. Battin, T. J., Besemer, K., Bengtsson, M. M., Romani, A. M. & Packmann, A. I. The ecology and biogeochemistry of stream biofilms. *Nature Reviews Microbiology* **14**, 251–263, https://doi.org/10.1038/nrmicro.2016.15 (2016).

4. Romani, A. M. & Sabater, S. Effect of primary producers on the heterotrophic metabolism of a stream biofilm. *Freshwater biology* **41**, 729–736, https://doi.org/10.1046/j.1365-2427.1999.00413.x (1999).

5. Hill, W. R., Mulholland, P. J. & Marzolf, E. R. Stream ecosystem responses to forest leaf emergence in spring. *Ecology* **82**, 2306–2319, https://doi.org/10.2307/2680233 (2001).

6. Hall, E. K. *et al.* Looking inside the box: using Raman microspectroscopy to deconstruct microbial biomass stoichiometry one cell at a time. *Isme Journal* **5**, 196–208, https://doi.org/10.1038/ismej.2010.115 (2011).

7. Besemer, K., Singer, G., Hodl, I. & Battin, T. J. Bacterial Community Composition of Stream Biofilms in Spatially Variable-Flow Environments. *Applied and Environmental Microbiology* **75**, 7189–7195, https://doi.org/10.1128/aem.01284-09 (2009).

8. Besemer, K. *et al.* Biophysical controls on community succession in stream biofilms. *Applied and Environmental Microbiology* **73**, 4966–4974, https://doi.org/10.1128/aem.00588-07 (2007).

9. Battin, T. J. *et al.* Microbial landscapes: new paths to biofilm research. *Nature Reviews Microbiology* **5**, 76–81, https://doi.org/10.1038/nrmicro1556 (2007).

10. Gloag, E. S. *et al.* Self-organization of bacterial biofilms is facilitated by extracellular DNA. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11541–11546, https://doi.org/10.1073/pnas.1218898110 (2013).

11. Merbt, S. N., Auguet, J. C., Blesa, A., Marti, E. & Casamayor, E. O. Wastewater Treatment Plant Effluents Change Abundance and Composition of Ammonia-Oxidizing Microorganisms in Mediterranean Urban Stream Biofilms. *Microbial Ecology* **69**, 66–74, https://doi.org/10.1007/s00248-014-0464-8 (2015).

12. Merbt, S. N., Auguet, J. C., Casamayor, E. O. & Marti, E. Biofilm recovery in a wastewater treatment plant-influenced stream and spatial segregation of ammonia-oxidizing microbial populations. *Limnology and Oceanography* **56**, 1054–1064, https://doi.org/10.4319/lo.2011.56.3.1054 (2011).

13. Hosen, J. D., Febria, C. M., Crump, B. C. & Palmer, M. A. Watershed Urbanization Linked to Differences in Stream Bacterial Community Composition. *Frontiers in Microbiology* **8**, https://doi.org/10.3389/fmicb.2017.01452 (2017).

14. Stanford, J. A., Gayeski, N. J., Pavlov, D. S., Savvaitova, K. A. & Kuzishchin, K. V. Biophysical complexity of the Krutogorova River (Kamchatka, Russia). *Verhandlungen Internationale Vereinigung für theoretische und angewandte Limnologie* **28**, 1354–1361 (2003).

15. Stanford, J. A., Lorang, M. S. & Hauer, F. R. The shifting habitat mosaic of river ecosystems. *Verhandlungen Internationale Vereinigung für theoretische und angewandte Limnologie* **29**, 123–136 (2005).

16. Ward, J. V., Tockner, K. & Schiemer, F. Biodiversity of floodplain river ecosystems: Ecotones and connectivity. *Regulated Rivers-Research & Management* **15**, 125–139, https://doi.org/10.1002/(sici)1099-1646 (1999).

17. Green, D. G. & Sadedin, S. Interactions matter - complexity in landscapes and ecosystems. *Ecological Complexity* **2**, 117–130, https://doi.org/10.1016/j.ecocom.2004.11.006 (2005).

18. Peipoch, M., Brauns, M., Hauer, F. R., Weitere, M. & Valett, H. M. Ecological Simplification: Human Influences on Riverscape Complexity. *Bioscience* **65**, 1057–1065, https://doi.org/10.1093/biosci/biv120 (2015).

19. Tockner, K., Lorang, M. S. & Stanford, J. A. River flood plains are model ecosystems to test general hydrogeomorphic and ecological concepts. *River Research and Applications* **26**, 76–86, https://doi.org/10.1002/rra.1328 (2010).

20. Bornette, G., Amoros, C. & Lamouroux, N. L. Aquatic plant diversity in riverine wetlands: The role of connectivity. *Freshwater Biology* **39**, 267–283, https://doi.org/10.1046/j.1365-2427.1998.00273.x (1998).

21. Robinson, C. T., Tockner, K. & Ward, J. V. The fauna of dynamic riverine landscapes. *Freshwater Biology* **47**, 661–677, https://doi.org/10.1046/j.1365-2427.2002.00921.x (2002).

22. Stanford, J. A. & Ward, J. V. An ecosystem perspective of alluvial rivers - connectivity and the hyporheic corridor. *Journal of the North American Benthological Society* **12**, 48–60, https://doi.org/10.2307/1467685 (1993).

23. Aarts, B. G. W., Van den Brink, F. W. B. & Nienhuis, P. H. Habitat loss as the main cause of the slow recovery of fish faunas of regulated large rivers in Europe: The transversal floodplain gradient. *River Research and Applications* **20**, 3–23, https://doi.org/10.1002/rra.720 (2004).

24. Copp, G. H. The habitat diversity and fish reproductive function of floodplain ecosystems. *Environmental Biology of Fishes* **26**, 1–27, https://doi.org/10.1007/bf00002472 (1989).

25. Kingsford, R. T. Ecological impacts of dams, water diversions and river management on floodplain wetlands in Australia. *Austral Ecology* **25**, 109–127, https://doi.org/10.1046/j.1442-9993.2000.01036.x (2000).

26. Hauer, F. R. *et al.* Gravel-bed river floodplains are the ecological nexus of glaciated mountain landscapes. *Science Advances* **2**, https://doi.org/10.1126/sciadv.1600026 (2016).

27. Prosser, J. I. *et al.* Essay - The role of ecological theory in microbial ecology. *Nature Reviews Microbiology* **5**, 384–392, https://doi.org/10.1038/nrmicro1643 (2007).

28. Barberan, A., Casamayor, E. O. & Fierer, N. The microbial contribution to macroecology. *Frontiers in Microbiology* **5**, https://doi.org/10.3389/fmicb.2014.00203 (2014).

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

30. Poff, N. L. *et al.* The natural flow regime. *Bioscience* **47**, 769–784, https://doi.org/10.2307/1313099 (1997).

31. Olden, J. D., Poff, N. L., Douglas, M. R., Douglas, M. E. & Fausch, K. D. Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology & Evolution* **19**, 18–24, https://doi.org/10.1016/j.tree.2003.09.010 (2004).

32. Widder, S. *et al.* Fluvial network organization imprints on microbial co-occurrence networks. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 12799–12804, https://doi.org/10.1073/pnas.1411723111 (2014).

33. Ward, J. V. & Stanford, J. A. Ecological connectivity in alluvial river ecosystems and its disruption by flow regulation. *Regulated Rivers-Research & Management* **11**, 105–119, https://doi.org/10.1002/rrr.3450110109 (1995).

34. Arscott, D. B., Tockner, K. & Ward, J. V. Thermal heterogeneity along a braided floodplain river (Tagliamento River, northeastern Italy). *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 2359–2373, https://doi.org/10.1139/cjfas-58-12-2359 (2001).

35. Ward, J. V. & Tockner, K. Biodiversity: towards a unifying theme for river ecology. *Freshwater Biology* **46**, 807–819 (2001).

36. Zeglin, L. H. Stream microbial diversity in response to environmental changes: review and synthesis of existing research. *Frontiers in Microbiology* **6**, https://doi.org/10.3389/fmicb.2015.00454 (2015).

37. Singer, G., Besemer, K., Schmitt-Kopplin, P., Hodl, I. & Battin, T. J. Physical Heterogeneity Increases Biofilm Resource Use and Its Molecular Diversity in Stream Mesocosms. *Plos One* **5**, https://doi.org/10.1371/journal.pone.0009988 (2010).

38. Battin, T. J., Wille, A., Sattler, B. & Psenner, R. Phylogenetic and functional heterogeneity of sediment biofilms along environmental gradients in a glacial stream. *Applied and Environmental Microbiology* **67**, 799–807, https://doi.org/10.1128/aem.67.2.799-807.2001 (2001).

39. Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nature Reviews Microbiology* **10**, 538–550, https://doi.org/10.1038/nrmicro2832 (2012).

40. Proulx, S. R., Promislow, D. E. L. & Phillips, P. C. Network thinking in ecology and evolution. *Trends in Ecology & Evolution* **20**, 345–353, https://doi.org/10.1016/j.tree.2005.04.004 (2005).

41. Colwell, R. K. Predictability, constancy, and contingency of periodic phenomena. *Ecology* **55**, 1148–1153, https://doi.org/10.2307/1940366 (1974).

42. Wohl, E., Hall, R. O., Lininger, K. B., Sutfin, N. A. & Walters, D. M. Carbon dynamics of river corridors and the effects of human alterations. *Ecological Monographs* **87**, 379–409, https://doi.org/10.1002/ecm.1261 (2017).

43. Strayer, D. L. & Dudgeon, D. Freshwater biodiversity conservation: recent progress and future challenges. *Journal of the North American Benthological Society* **29**, 344–358, https://doi.org/10.1899/08-171.1 (2010).

44. Dudgeon, D. *et al*. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews* **81**, 163–182, https://doi.org/10.1017/s1464793105006950 (2006).

45. Brook, B. W., Sodhi, N. S. & Bradshaw, C. J. A. Synergies among extinction drivers under global change. *Trends in Ecology & Evolution* **23**, 453–460, https://doi.org/10.1016/j.tree.2008.03.011 (2008).

46. Crump, B. C., Adams, H. E., Hobbie, J. E. & Kling, G. W. Biogeography of bacterioplankton in lakes and streams of an arctic tundra catchment. *Ecology* **88**, 1365–1378, https://doi.org/10.1890/06-0387 (2007).

47. Tockner, K., Pennetzdorfer, D., Reiner, N., Schiemer, F. & Ward, J. V. Hydrological connectivity, and the exchange of organic matter and nutrients in a dynamic river-floodplain system (Danube, Austria). *Freshwater Biology* **41**, 521–535, https://doi.org/10.1046/j.1365-2427.1999.00399.x (1999).

48. Cardinale, B. J., Palmer, M. A., Swan, C. M., Brooks, S. & Poff, N. L. The influence of substrate heterogeneity on biofilm metabolism in a stream ecosystem. *Ecology* **83**, 412–422, https://doi.org/10.2307/2680024 (2002).

49. Peipoch, M. *et al*. Small-scale heterogeneity of microbial N uptake in streams and its implications at the ecosystem level. *Ecology* **97**, 1329–1344, https://doi.org/10.1890/15-1210.1 (2016).

50. Green, J. L. *et al*. Spatial scaling of microbial eukaryote diversity. *Nature* **432**, 747–750, https://doi.org/10.1038/nature03034 (2004).

51. Leff, L., Van Gray, J. B., Marti, E., Merbt, S. N. & Romani, A. M. *Aquatic Biofilms and Biogeochemical Processes*. (2016).

52. Peipoch, M., Jones, R. & Valett, H. M. Spatial Patterns in Biofilm Diversity across Hierarchical Levels of River-Floodplain Landscapes. *Plos One* **10**, https://doi.org/10.1371/journal.pone.0144303 (2015).

53. Niederdorfer, R., Peter, H. & Battin, T. J. Attached biofilms and suspended aggregates are distinct microbial lifestyles emanating from differing hydraulics. *Nature Microbiology* **1**, https://doi.org/10.1038/nmicrobiol.2016.178 (2016).

54. Olapade, O. A. & Leff, L. G. Influence of dissolved organic matter and inorganic nutrients on the biofilm bacterial community on artificial substrates in a northeastern Ohio, USA, stream. *Canadian Journal of Microbiology* **52**, 540–549, https://doi.org/10.1139/w06-003 (2006).

55. Ardon, M. & Pringle, C. M. The quality of organic matter mediates the response of heterotrophic biofilms to phosphorus enrichment of the water column and substratum. *Freshwater Biology* **52**, 1762–1772, https://doi.org/10.1111/j.1365-2427.2007.01807.x (2007).

56. Johnson, L. T., Royer, T. V., Edgerton, J. M. & Leff, L. G. Manipulation of the Dissolved Organic Carbon Pool in an Agricultural Stream: Responses in Microbial Community Structure, Denitrification, and Assimilatory Nitrogen Uptake. *Ecosystems* **15**, 1027–1038, https://doi.org/10.1007/s10021-012-9563-x (2012).

57. Ingman, G. L. (ed Department of Health and Enviroamental Sciences) (State of Montana, 1992).

58. Lohman, K. & Priscu, J. C. Physiological indicators of nutrient deficiency in cladophora (chlorophyta) in the clark fork of the Columbia River, Montana. *Journal of Phycology* **28**, 443–448, https://doi.org/10.1111/j.0022-3646.1992.00443.x (1992).

59. Welter, J. R. *et al*. Does N-2 fixation amplify the temperature dependence of ecosystem metabolism? *Ecology* **96**, 603–610, https://doi.org/10.1890/14-1667.1 (2015).

60. Galmes, J., Kapralov, M. V., Copolovici, L. O., Hermida-Carrera, C. & Niinemets, U. Temperature responses of the Rubisco maximum carboxylase activity across domains of life: phylogenetic signals, trade-offs, and importance for carbon gain. *Photosynthesis Research* **123**, 183–201, https://doi.org/10.1007/s11120-014-0067-8 (2015).

61. Leibold, M. A. *et al*. The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* **7**, 601–613, https://doi.org/10.1111/j.1461-0248.2004.00608.x (2004).

62. Leibold, M. A. & Loeuille, N. Species sorting and patch dynamics in harlequin metacommunities affect the relative importance of environment and space. *Ecology* **96**, 3227–3233, https://doi.org/10.1890/14-2354.1 (2015).

63. Langenheder, S. *et al*. Bacterial metacommunity organization in a highly connected aquatic system. *Fems Microbiology Ecology* **93**, https://doi.org/10.1093/femsec/fiw225 (2017).

64. Leibold, M. A. & McPeek, M. A. Coexistence of the niche and neutral perspectives in community ecology. *Ecology* **87**, 1399–1410, https://doi.org/10.1890/0012-9658(2006)87[1399:cotnan]2.0.co;2 (2006).

65. Barberan, A., Bates, S. T., Casamayor, E. O. & Fierer, N. Using network analysis to explore co-occurrence patterns in soil microbial communities. *Isme Journal* **6**, 343–351, https://doi.org/10.1038/ismej.2011.119 (2012).

66. Eiler, A., Heinrich, F. & Bertilsson, S. Coherent dynamics and association networks among lake bacterioplankton taxa. *Isme Journal* **6**, 330–342, https://doi.org/10.1038/ismej.2011.113 (2012).

67. Fuhrman, J. A. & Steele, J. A. Community structure of marine bacterioplankton: patterns, networks, and relationships to function. *Aquatic Microbial Ecology* **53**, 69–81, https://doi.org/10.3354/ame01222 (2008).

68. Agler, M. T. *et al*. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *Plos Biology* **14**, https://doi.org/10.1371/journal.pbio.1002352 (2016).

69. Qu, X. D. *et al*. Influences of anthropogenic land use on microbial community structure and functional potentials of stream benthic biofilms. *Scientific Reports* **7**, https://doi.org/10.1038/s41598-017-15624-x (2017).

70. Faust, K. *et al*. Cross-biome comparison of microbial association networks. *Frontiers in Microbiology* **6**, https://doi.org/10.3389/fmicb.2015.01200 (2015).

71. Ma, B. *et al*. Genetic correlation network prediction of forest soil microbial functional organization. *The ISME Journal* **12**, 2492–2505, https://doi.org/10.1038/s41396-018-0232-8 (2018).

72. Grimm, N. B. & Fisher, S. G. Nitrogen limitation in a sonoran desert arizona usa stream. *Journal of the North American Benthological Society* **5**, 2–15, https://doi.org/10.2307/1467743 (1986).

73. Berrendero, E. *et al*. Nitrogen fixation in a non-heterocystous cyanobacterial mat from a mountain river. *Scientific Reports* **6**, https://doi.org/10.1038/srep30920 (2016).

74. Berry, D. & Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in Microbiology* **5**, https://doi.org/10.3389/fmicb.2014.00219 (2014).

75. Parter, M., Kashtan, N. & Alon, U. Environmental variability and modularity of bacterial metabolic networks. *Bmc Evolutionary Biology* **7**, https://doi.org/10.1186/1471-2148-7-169 (2007).

76. Wu, J. G. & David, J. L. A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. *Ecological Modelling* **153**, 7–26, https://doi.org/10.1016/s0304-3800(01)00499-9 (2002).

77. Lupatini, M. *et al*. Network topology reveals high connectance levels and few key microbial genera within soils. *Frontiers in Environmental Science* **2**, https://doi.org/10.3389/fenvs.2014.00010 (2014).

78. Freedman, Z. B. & Zak, D. R. Atmospheric N deposition alters connectance, but not functional potential among saprotrophic bacterial communities. *Molecular Ecology* **24**, 3170–3180, https://doi.org/10.1111/mec.13224 (2015).

79. Menzel, D. W. & Vaccaro, R. F. The measurement of dissolved organic and particulate carbon in seawater. *Limnology and Oceanography* **9**, 138–142 (1964).

80. APHA. *Standard methods for examination of water and wastewater*. Twenty-first edition edn, (American Public Health Association, 2006).

81. Patton, C. J. & Kryskalla, J. R. Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory: evaluation of alkaline persulfate digestion as an alternative to Kjeldahl digestion for determination of total and dissolved nitrogen and phosphorus in water. Water-Resources Investigations Report 2003-4174, https://doi.org/10.3133/wri034174.

82. Sartory, D. P. & Grobbelaar, J. U. Extraction of chlorophyll a from freshwater phytoplankton for spectrophotometric analysis. *Hydrobiologia* **114**, 177–187, https://doi.org/10.1007/bf00031869 (1984).

83. Steinman, A. D., Lamberti, G. A. & Leavitt, P. R. In *Methods in stream ecology* (eds Hauer, F. R. & Lamberti, G. A.) Ch. 17, 357–380 (Elsevier Academic Press, 2006).

84. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America.* **108**, 4516–22, https://doi.org/10.1073/pnas.1000080107 (2011).

85. Baselga, A. Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* **19**, 134–143, https://doi.org/10.1111/j.1466-8238.2009.00490.x (2010).

86. Dray, S., Legendre, P. & Peres-Neto, P. R. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* **196**, 483–493, https://doi.org/10.1016/j.ecolmodel.2006.02.015 (2006).

87. R: A language and environment for statistical computing (Vienna, Austria, 2008).

88. Stone, L. & Roberts, A. The checkerboard score and species distributions. *Oecologia* **85**, 74–79, https://doi.org/10.1007/bf00317345 (1990).

89. Package 'ecospat': Spatial Ecology Miscellaneous Methods (2018).

90. Konstantinidis, K. T. & Tiedje, J. M. Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. *Current Opinion in Microbiology* **10**, 504–509, https://doi.org/10.1016/j.mib.2007.08.006 (2007).

91. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal*, 1695 (2006).

92. Bastian, M., Heymann, S. & Jacomy, M. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the Third International ICWSM Conference* (2009).

93. Newman, M. E. J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 8577–8582, https://doi.org/10.1073/pnas.0601602103 (2006).

94. Berry, D. & Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in Microbiology* **5**, https://doi.org/10.3389/fmicb.2014.00219 (2014).

Acknowledgements

Author contributions

Conceived and designed the experiments: M.P. and H.M.V. Performed the experiments: M.P. and H.M.V. Analyzed the data: M.P., S.R.M. and T.R.A. Contributed reagents/materials/analysis tools: S.R.M. and T.R.A. Wrote the paper: M.P., S.R.M. and H.M.V.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-52865-4.

Correspondence and requests for materials should be addressed to M.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.