communications earth & environment

Article

https://doi.org/10.1038/s43247-024-01478-5

Geodiversity of a European river network controls algal biodiversity and function

Thomas Fuß [1,2] ✉, Lukas Thuile Bistarelli [1], Franziska Walther[3], Simon Vitecek [4,5], Lauren Talluto [1] & Gabriel Singer [1,2] ✉

Biodiversity and functioning often follow spatial gradients, yet with unclear causal linkage. In spatially complex rivers, regional-scale factors associated with hydrological connections and catchment properties control downstream transport of material and dispersal of organisms, both being crucial for ecosystem functioning. In a single snapshot study, we here show how a river's network structure interacts with its terrestrial matrix to control key environmental conditions and periphyton community composition at the local habitat scale, which in turn drive primary production. We found the high geodiversity of the Vjosa River network to promote high periphyton beta-biodiversity through regional (dispersal) and local (species sorting) processes. Community turnover driven by species sorting rather than purely by dispersal was identified as relevant for production rates, suggesting a match between environment and community composition to be conditional for functioning. Hence, anthropogenic perturbation of regional mechanisms by river modification may affect ecosystem functions through interfering with metacommunity structure.

In rivers, biogeochemical processes of global importance[1,2] are driven by a remarkably concentrated fraction of Earth´s biodiversity[3,4]. Alarmingly, we are losing riverine biodiversity at an unprecedented rate[5] without having properly understood what organizes or maintains it[6] nor being able to assess functional implications of its dramatic loss[7,8]. One reason for this knowledge gap is the need to consider regional complexity when it comes to studying the linkage of community structure and functioning in natural river networks. This is particularly concerning as human activities result in changes to land cover and river hydrology at regional scales.

In rivers, local ecosystems are connected in an asymmetric and hierarchical dendritic network, producing a strong, directional influence of upstream on downstream ecosystems[9]. Indeed, rivers are better conceptualized as fluvial meta-ecosystems, where flowing water transports materials and energy from upstream to downstream and dispersal links local communities to metacommunities at the regional river network scale[10,11]. Hence, a river network´s spatial configuration in its terrestrial matrix (i.e. geology and land cover) and downstream water flow therein may be the ultimate controls for biogeochemical processes, as they shape environmental conditions at any location in the network and structure its metacommunity from bacteria to fish[12–14]. A biogeochemical process like primary production is the consequence of species interacting with resources in a specific environment—understanding a river network´s primary production thus becomes a question of metacommunity structure-function coupling.

Metacommunity ecology provides paradigms (e.g. species sorting and mass effects) for understanding the relative importance of local environmental conditions, regional dispersal, and their interactions[15,16]. However, observational studies on primary producer metacommunity structure in river networks are scarce[12], and assessments of its implications for functioning are non-existent. Conceptually, efficient species sorting along environmental gradients depends on dispersal, with most efficient species sorting (i.e., the best match between community and environmental conditions) expected at intermediate levels of dispersal[16–18]. Low dispersal may prevent potentially productive species from reaching suitable local environments, while high dispersal may overwhelm local environmental selection and cause species persistence at environmentally unsuitable sites (i.e. mass effects)[19,20]. The degree of match between environmental conditions and a community´s traits may determine functional performance[21,22]. For example, the productivity of periphyton species may depend on how well they exploit resources, which in turn is determined by the match between their morphology, a key species trait, and environmental conditions such as flow velocity. Species with a tall growth form have an advantage over smaller species, as their increased exposure to water flow facilitates nutrient and light acquisition, however, the tall growth form also makes them highly

[1]Department of Ecology, University of Innsbruck, 6020 Innsbruck, Austria. [2]Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), 12587 Berlin, Germany. [3]Chair of Planning of Landscape and Urban Systems (PLUS), Institute for Spatial and Landscape Planning, Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, 8093 Zurich, Switzerland. [4]WasserCluster Lunz—Biologische Station GmbH, 3293 Lunz am See, Austria. [5]University of Natural Resources and Life Sciences, 1180 Vienna, Austria. ✉e-mail: thomas.fuss@uibk.ac.at; gabriel.singer@uibk.ac.at

susceptible to abrasion[12,23]. Contrary, smaller species are less exposed to water flow and resist higher flow velocities, which makes them better adapted to high flow velocity environments[12,23]. The occurrence of beneficial traits of communities in prevailing local environmental conditions is defined by community assembly processes, and these act at both local and regional scales. Thus, progress toward understanding of primary production in river networks will benefit from studying local as well as regional controls, including their interactions in a spatially explicit manner[2,10,24].

Questions of spatial scale in studies of riverine primary production have hitherto been investigated as effects of distal (e.g. climate, land cover, geology) and proximal controls (e.g. resources, hydraulics) on a strictly local description of production[25–27]. Ultimately, local nutrient conditions and light are pivotal for primary production[26,28]. However, rivers gain nutrients from their terrestrial matrix, and at the same time receive sediments that affect turbidity and thus light availability. Therefore geology and land cover act as regional modulators of primary production[25,29,30]. In-stream light may vary from small to large spatial scales as it depends on light attenuation by riparian vegetation, valley topography and turbidity derived from suspended sediments transported over long distances[30,31]. Typically, primary production responds to light variation along a saturation curve, reflecting limitation and inhibition at low and high light, respectively. Production-Irradiance (PI) curves are shaped by average light, nutrients, temperature, and the autotrophic community[27,32,33]. Indeed, compositional turnover at the community-level is an often observed adaptation to local light conditions, in addition to physiological reactions at the cell-level via upregulating chlorophyll-*a* content[33–36]. Parameters describing the shape of a PI-curve may thus be directly linked to the functional capacity determined by the composition of the local community.

In a single-snapshot study, we here investigate periphyton primary production across an entire river network in a spatially explicit manner that transcends the use of scale-specific predictors for localized production measurements as well as unidimensional river size gradients. We recognize the river´s dendritic network structure and its terrestrial matrix as regional controls shaping local environmental conditions and mediating community composition through dispersal (Fig. 1). We hypothesize that environmental conditions, community composition and their interaction through species sorting finally (co-)define the functional performance of periphyton in a real river setting. Following our causal framework (Fig. 1), we used multiple variation partitioning analyses with periphyton primary production as the final response to a cascade of drivers from regional to local scale. Our multi-scale approach is timely, as human activities increasingly influence regional mechanisms driving fluvial meta-ecosystems. For instance, large-scale land cover conversion changes a river network´s terrestrial matrix, and fragmentation interferes with a river network´s sediment and flow regimes[37,38], thereby altering transport and connectivity dynamics. Here, we study the hydromorphologically unaltered Vjosa (Aoos) River network draining a catchment of 6,704 km² ranging from the Greek Pindos Mountains to the Adriatic Sea in Southern Albania (Fig. 2). The catchment is partly a United Nations Educational, Scientific and Cultural Organization (UNESCO) geopark with high geological diversity. Perhaps uniquely for its size in Europe, the Vjosa River network is still characterized by unobstructed water flow and sediment transport along a near-natural river corridor consisting of gorges and both braided and meandering river sections[39,40], making it a well-suited river system for our large-scale meta-ecosystem functioning study.

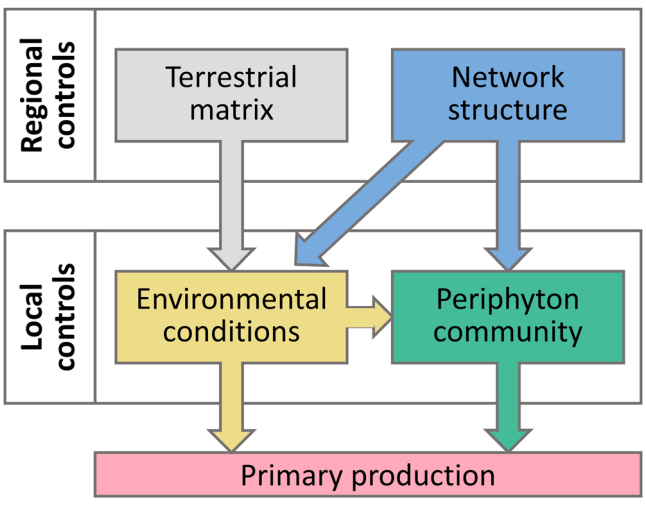


Fig. 1 | Conceptual framework. We hypothesized that (i) a river´s network structure and the terrestrial matrix (i.e. land cover and geology) control environmental conditions, (ii) periphyton community composition is sorted by environmental conditions and its dispersal is constrained by the network structure, and (iii) environmental conditions and periphyton community affect primary production. Arrows illustrate hypotheses.

Results

The multi-scale analysis of riverine meta-ecosystem functioning following our causal framework (Fig. 1) indicated the importance of the Vjosa River network's terrestrial matrix and network structure as regional controls for primary production because of their influence on both local controls, environmental conditions and the community composition. River network structure has a special role in our analysis as it constrains the transport of nutrients and sediments and (passive) algal dispersal, which are all considered to occur mainly in a downstream direction[12,41]. 15 spatial variables arising from eigenvector analysis constitute potentially detectable directional spatial gradients at various spatial scales among our 46 study sites (see asymmetric eigenvector map analysis in "Methods"). We used those spatial variables for the hypothesis tests as descriptors of network structure (Fig. 1), each potentially acting as a proxy for a spatial process at a distinct spatial scale. Forward-selected spatial variables included in final models (Table 1) were spatial variable 1, which represented an upstream-downstream gradient across the entire river network and strongly correlated with log-transformed subcatchment area (Pearson's $r = -0.90$, $df = 44$, p value < 0.001), and—at an intermediate spatial scale—spatial variable 2, which differentiated the Drinos subcatchment from the main stem, and spatial variable 4, which differentiated the Sarantaporos and Aoos subcatchments (Fig. 3b). Further included spatial variables 5, 6, 13 and 14 described spatial relationships among sites at smaller scales, which cannot be easily interpreted (see Supplementary Fig. S1). Remaining spatial variables did not describe environmental gradients or shifts in community composition as they were not included in any final model.

Limestone was identified as the dominant geological class in the entire catchment. Additional common geological substrates were sandstone, flysch and ultramafic igneous rock (Fig. 3a, see Supplementary Table S1). Dominating land-cover classes were forest and natural vegetation. Agricultural areas, bare rock, urban infrastructure and water bodies contributed only little to the catchment's land cover. Principal component analysis (PCA) revealed main gradients of geology (geo-PCs) and land cover (land-PCs). Important geological gradients distinguished limestone from ultramafic igneous rock and sandstone (geo-PC1), ultramafic igneous rock from sandstone (geo-PC2), limestone from flysch (geo-PC3), and conglomerate and sandstone (geo-PC4) (see factor loadings of all PCAs in Supplementary Fig. S2).

Sites across the Vjosa River network revealed strong variation in environmental conditions. For example, in-stream light ranged from 124.5 to 554.8 µmol $m^{-2}s^{-1}$ and water temperature from 6.5 to 22.7 °C (see Supplementary Table S2). Variation partitioning showed that a large fraction of variation in nutrients, temperature and conductivity could be explained by geology alone (55%) and only 3% by the network structure alone (Fig. 3c). While land cover was not significant, geology was identified as important along three main gradients (geo-PCs 1–3). Specifically, subcatchments dominated by limestone exhibited higher concentrations of nutrients, except for concentrations of magnesium and silica, which were highest in subcatchments with high relative amounts of ultramafic igneous rock and

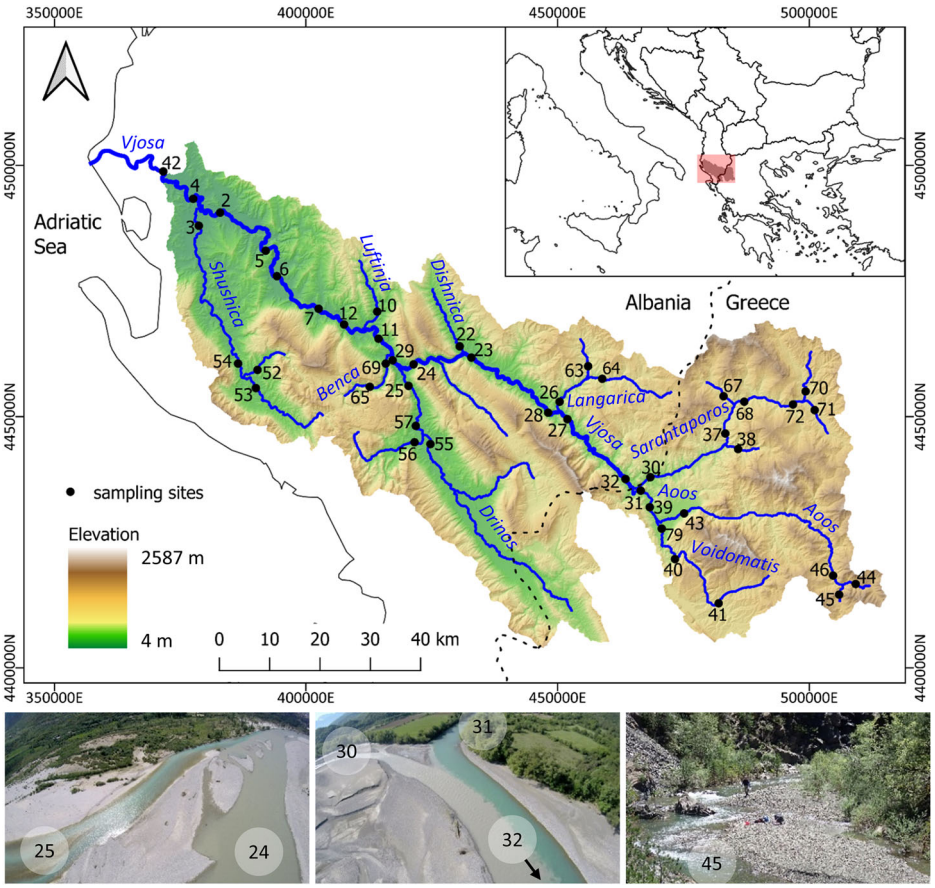


Fig. 2 | Map of the Vjosa river network with sampling sites. Pictures show sites at the two confluences Drinos-Vjosa, and Aoos-Sarantaporos and at the upstream section of the Aoos (from left to right). At each of the confluences, the two tributaries show strong differentiation in sediment loads due to diverging subcatchment geology. The presence/absence of easily transported fine sediment drives strong spatial autocorrelation of in-stream light on any of the tributaries (Fig. 6). The dominant source of fine sediment is the Sarantaporos catchment, where abundant highly erodible sandstone creates turbid water. In contrast, limestone dominated the rather clear southern tributaries Voidomatis, Drinos, Benca and Shushica (see also Fig. 3a).

flysch (see results from the redundancy analysis in Supplementary Fig. S3). Spatial variables 1 and 14 (from asymmetric eigenvector map analysis) were selected as main descriptors of network structure driving environmental conditions. Variation partitioning revealed that network structure alone played a more important (25%) role than geology alone (14%) as driver of in-stream light (Fig. 3c). A substantial fraction of the in-stream light variation (33%) was shared by network structure and geology.

The periphyton metacommunity consisted of 489 taxa, obtained as amplicon sequence variants (ASVs), and was mainly constituted of Bacillariophyta (diatoms, relative abundance 80%) and Chrysophyta (19%). The most abundant genera were the chrysophyte *Hydrurus* (35%) and the diatoms *Asterionella* (12%), *Cymbella* (12%), *Diatoma* (9%), *Nitzschia* (5%), *Gomphonema* (5%), *Navicula* (4%) and *Achnanthidium* (4%). Additive partitioning of diversity, which estimates how much of the regional periphyton richness (number of ASVs in the Vjosa River network) was composed of local richness (number of ASVs at each site) and the turnover across local communities (β-diversity, see Methods)[42], revealed that mean local richness (53 ASVs) was lower than expected under the null model (164 ASVs, p-value < 0.001), whereas ASV turnover across sites comprised the majority of the regional periphyton diversity and was higher than expected (436 > 325 ASVs, p-value < 0.001). Not considering the 236 singletons, we classified 140 ASVs as satellite and 10 as core taxa, across sites, on average 40% (±10 SD) of present ASVs were classified in one of these two groups. Further, we could assign 16 and 24 ASVs to low-profile and high-profile groups, respectively, across sites, on average 18% (±6 SD) of present ASVs were classified in one of these two groups. Results of variation partitioning suggest that network structure and environmental conditions explained the variation in periphyton community composition similarly well (15% and 13%, respectively), with 9% of variation commonly explained by the network structure and environmental conditions (Fig. 3d, Table 1). Forward-selected environmental predictors were in stream-light and two main environmental gradients derived from PCA, which ordered sites along a gradient of low to high temperature, conductivity and concentrations of nutrients (Ca, K, Na, SO_4, Cl, NO_3) (env-PC1) and along a gradient of low to high Mg and Si concentrations (env-PC2) (Fig. 4).

Finally, we targeted the controls of periphyton functioning (potential primary production), which we measured through (local) community-level Production-Irradiance (PI) curves described by two parameters: the biomass-specific maximum gross production rate (P_{max}) and the light intensity at P_{max} (I_{opt}). The final model to describe variation in P_{max} included one selected environmental gradient and two selected gradients of community turnover derived from PCA (Table 1): P_{max} was positively related to Mg and Si concentrations (env-PC2) (Fig. 4) and to cc-PC1 and cc-PC3, (Fig. 5). Along both selected cc-PCs we found a shift in relative abundance from high-profile (e.g. *Diatoma vulgare* var *linearis*, *Encyonema minutum, Melosira varians*) to low-profile taxa (e.g. *Achnanthidium minutissimum*, *Encyonopsis* sp.) (see factor loadings of ASVs in Supplementary Fig. S4). Satellite taxa dominated especially the higher end of cc-PC1 while core taxa prevailed in the center of the gradient. Also, richness increased and biomass decreased along cc-PC1; biomass also decreased along cc-PC3 (Fig. 5). Cc-PC3 decreased with log-transformed subcatchment area (Pearson's $r = -0.49$, $df = 37$, p-value < 0.01). Partial regression revealed that the bulk explanatory power lay in the commonly explained fraction (32%) and variation solely explained by environmental conditions or community composition was not significant (Fig. 3e). In contrast, we found no direct effect of community composition on the light intensity at P_{max} (I_{opt}), for which only in-stream light could be identified as a driver (Fig. 3e).

Discussion

Our study reveals how the essence of a river—a dendritically structured network embedded in a terrestrial matrix – controls its primary production via shaping environmental conditions and the autotrophic community

Table 1 | Results of tests to identify causal links between the Vjosa river network's terrestrial matrix, its network structure, environmental conditions, periphyton community composition and periphyton functioning (see hypotheses Fig. 1)

Response variable(s)		Predictor variables		Global test					Selected predictors[a]	Final model			
				Type	n	df	adjR²	P		Type	n	df	adjR²	P
Environmental conditions	Nutrients, temperature, conductivity[b]	Terrestrial matrix	land-PC1-3; geo-PC1-5	RDA	46	37	0.63	***	geo-PC1-3 (all ***)	RDA	46	40	0.64	***
		Network structure	SV1-15	RDA	46	30	0.2	*	SV1*, SV14*					
	light	Terrestrial matrix	land-PC1-3; geo-PC1-5	MLR	31	24	0.43	**	geo-PC1,2,4 (*, **, *, resp.); all neg	MLR	31	24	0.71	***
		Network structure	SV1-15	MLR	31	15	0.52	*	SV2**pos, SV4**pos, SV14**neg					
Periphyton community	site-by-ASV matrix[c]	Env. conditions	env-PC1-5; light	RDA	27	20	0.27	***	env-PC1***, env-PC2***, light**	RDA	27	17	0.37	***
		Network structure	SV1 – 15	RDA	27	15	0.31	**	SV1**, SV2*, SV4*, SV5*, SV6*, SV13*					
Periphyton functioning	log(P_{max})	Env. conditions	env-PC1-5; light	MLR	26	19	0.29	*	env-PC2**pos	MLR	24	20	0.36	**
		Periphyton community	cc-PC1-8	MLR	24	15	0.42	*	cc-PC1*pos, cc-PC3*pos					
	log(I_{opt})	Env. conditions	env-PC1-5; light	MLR	26	19	0.34	*	light**pos	LM	26	24	0.33	**
		Periphyton community	cc-PC1-8	MLR	24	15	0	n.s.	–	–	–	–	–	–

land-PC, geo-PC, env-PC, cc-PC principal components derived by principal component analysis (PCA) on the relative coverage of 7 land cover classes, 13 geological classes, *z*-standardized log(nutrients) and log(conductivity) and temperature, and on Hellinger-transformed rarefied site-by-ASV table, respectively. *SV* spatial variable, *RDA* redundancy analysis, *MLR* multiple linear regression, *LM* linear model, *n* sample size, *df* degrees of freedom, *adjR²*, adjusted R^2; *light* in-stream light availability; P_{max} biomass-specific maximum production rate, I_{opt} light intensity at P_{max}; *pos/neg* positive or negative relationship of response variable and selected predictor, respectively.

[a]Forward selected.[92]

[b]Z-standardized log(Ca, K, Mg, Si, Na, SO_4, Cl, NO_3, conductivity) and temperature.

[c]Rarefied and Hellinger transformed.

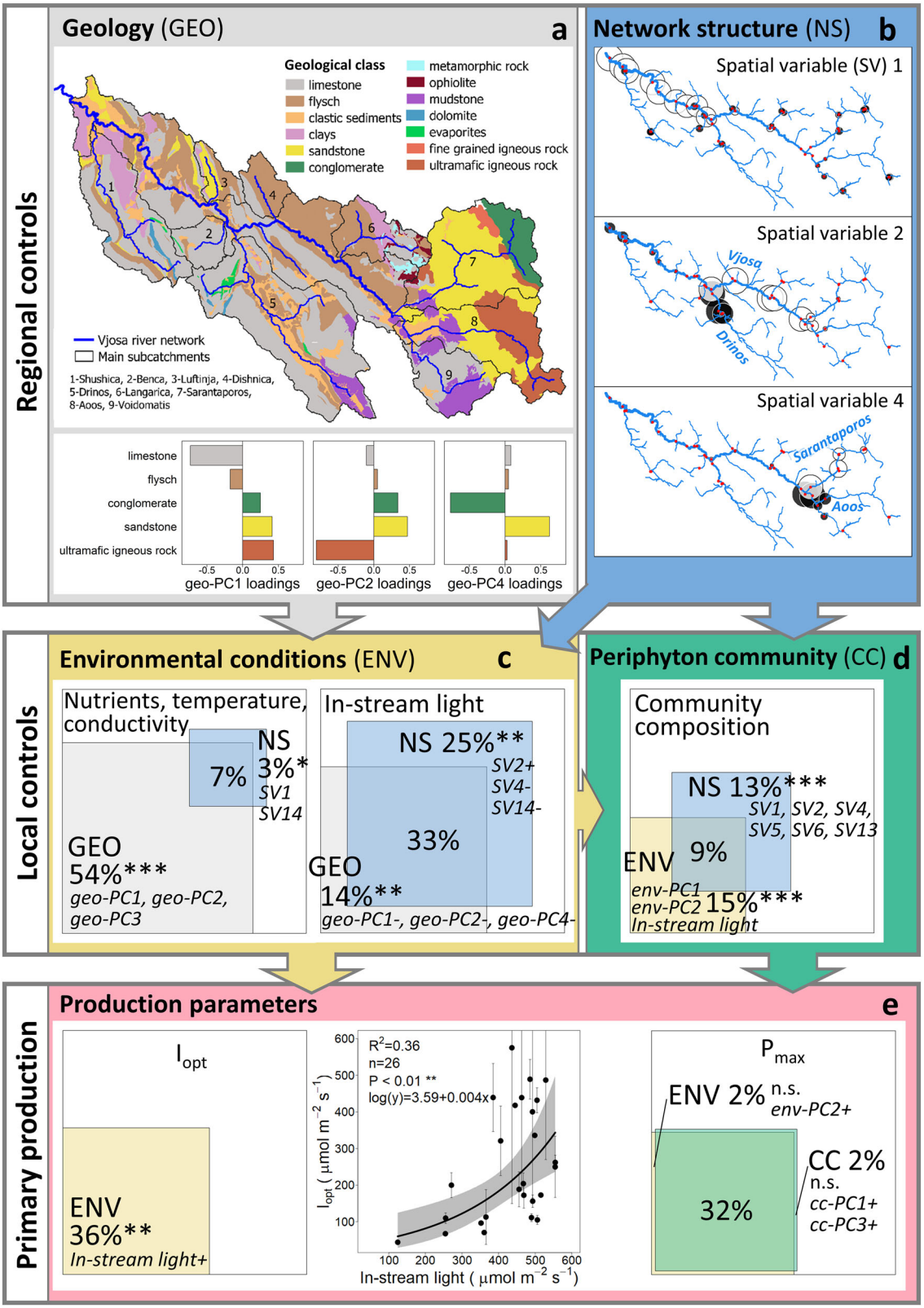
An asterisks show degrees of significance (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

through various mechanisms acting and interacting across spatial scales. Largely, our analysis based on variation partitioning supported our hypothesized causal framework (Fig. 1). Environmental conditions strongly reflected geological characteristics of subcatchments (Fig. 3c), yet network structure alone explained substantial additional fractions of in-stream light heterogeneity (Fig. 3c right), presumably by shaping the transport of suspended sediments (Fig. 6). Both environmental conditions and the network structure independently predicted periphyton community composition, indicating species sorting along environmental gradients and dispersal to shape metacommunity patterns (Fig. 3d). Our results suggest the regional scale as the main spatial scale promoting biodiversity in the Vjosa River network. Environmental gradients and community composition, both strongly shaped by regional controls, ultimately influenced production parameters (Fig. 3e). Importantly, the influence of environmental conditions and community composition on the maximum gross production rate (P_{max}) largely overlapped (Fig. 3e right). This suggests that especially species sorting along environmental gradients, which leads to a good match between environment and community, enables high functioning.

The catchment's remarkably high geological diversity clearly drove environmental patterns (Fig. 3c). Surprisingly, network structure could only explain 10% of the variation in nutrients, temperature and conductivity across the Vjosa River network (of which 7% was shared with geology). We attribute this to natural nutrient concentration and temperature discontinuities among sampled sites which diminished spatial autocorrelation of downstream to upstream sites. We mainly sampled tributaries just upstream of their confluences and sites immediately below where water bodies were evidently mixed. This sampling design led to long watercourse distances between sites on the same river reach (i.e., between confluences), and in-stream nutrient uptake and transformation potentially caused nutrient turnover at smaller scales than this distance between sites[43]. To overcome this limitation, future studies should include sites in close proximity on the same river reach[44]. In addition, human point sources and abundant thermal and karstic springs might have created discontinuities in environmental gradients. Further, the often strong environmental differentiation of tributaries due to variable geology in their respective subcatchments may have caused discontinuities at downstream to upstream sites of confluences[9].

In contrast to the weak spatial structure of environmental gradients, network structure was an important driver for in-stream light presumably by constraining the transport of suspended sediments (Fig. 3c). This spatial dependency is linked to the substantial load of suspended sediments sourced from sandstone-dominated subcatchments and their unobstructed transport through the Vjosa River network (Fig. 2, Fig. 6)[39,40]. Negative effects of selected geological gradients can all be summarized as reduced in-stream light in subcatchments with large fractions of sandstone, as sandstone was found at the positive end of all selected geo-PCs (i.e. main geological gradients, Fig. 3a). Hence, effects of the network structure and geology are difficult to disentangle, which resulted in a large commonly explained fraction of in-stream light variation.

Variation partitioning showed that the periphyton metacommunity was shaped by environmental conditions and the network structure (Fig. 3d), which is corroborated by the higher β-diversity among local habitats than expected by chance. This suggests that dispersal and environmental differentiation among local habitats were sufficient to enable efficient species sorting, which caused a large fraction of the Vjosa River network's regional diversity to lie in differentiated local habitats[16,45]. Taxa were sorted along gradients of light availability, nutrient concentrations, temperature and conductivity (integrated in env-PC1 and env-PC2), the dominant factors for periphyton community composition also identified in other studies[35,36,46,47]. For example, the cold-water stenotherm *Hydrurus foetidus* was most abundant in cold-water mountain streams in the Aoos, Voidomatis and Sarantaporos subcatchments (see results from the redundancy analysis in Supplementary Fig. S5). The relevant environmental gradients were to some degree spatially structured as revealed by the overlapping fraction in variation partitioning of environmental and spatial



effects (Fig. 3d, Fig. 6). However, shifts in community composition attributed to the network structure alone suggest that dispersal dynamics in the Vjosa River network also created periphyton metacommunity patterns independent of environmental conditions. We note that this result may also derive from unknown spatially structured environmental or biotic controls[48]. Describing the manifold controls on periphyton communities, including physical stressors such as shear stress and grazing by herbivores, which may have a very localized imprint, is a challenge, which commonly leads to poorly explained metacommunity patterns[12,49]. This may also reveal limitations of snap-shot surveys, which cannot capture temporal variability in environmental conditions and frequent disturbance during community assembly[50]. Notably, discharge dynamics modulate the residence time and

Fig. 3 | Main results that confirm the hypotheses presented in Fig. 1. Maps of the Vjosa river network representing geological classes (**a**) and important spatial variables (SVs) (**b**, remaining spatial variables are shown in Supplementary Fig. S1); colors of bubbles reflect opposite sides of the spatial gradient and sizes represent absolute score values, hence, importance of sites for the respective spatial variable. Squared Venn diagrams show results of variation partitioning analysis for the respective response variables, i.e. nutrients (Ca, K, Mg, Si, Na, SO_4, Cl, NO_3), temperature and conductivity (**c**, left), in-stream light (**c**, right), community compositions (**d**), biomass-specific maximum production rate (P_{max}, **e**, right), and light intensity at P_{max} (I_{opt}, **e**, left). Gray, blue, yellow and green squares show relative fractions of variation of response variables explained by geology (GEO), network structure (NS), environmental conditions (ENV) and community compositions (CC), respectively. We tested statistical significance for relative fractions based on Monte Carlo permutation tests with 999 permutations[94,97]. Asterisks show significance (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$) of relative fractions. Shared fractions cannot be tested for significance. Overlaps of squares show commonly explained variation. Significant predictors of respective relative fractions are shown in italics; ± signs denote positive or negative effects, respectively, of single predictors on single responses modeled by linear or multiple regressions. When used as predictors, geological classes, environmental conditions (nutrients, temperature and conductivity) and community composition were reduced to the principal components geo-PCs, env-PCs and cc-PCs, respectively. Factor loadings indicate the strength and direction of the relationship between three important geo-PCs and a subset of geological classes (**a**, factor loadings of env-PCs, land-PCs and remaining geo-PCs are shown in Supplementary Fig. S2). Regression plot (**e**, center), including regression line and 95% confidence band, shows effect of in-stream light on I_{opt}; error bars represent standard errors. Geology shaped environmental conditions (**c**), particularly the effect on in-stream light (**c**, right) can be attributed to the relative fraction of sandstone in the subcatchment (see panel **a** for factor loadings of selected geo-PCs). In-stream light also showed spatial patterns detected by spatial variables (**b**), presumably derived from sediment transport along the river network. Effects of environmental conditions and the network structure on community composition point to species sorting and dispersal processes (**d**). I_{opt} tracked in-stream light conditions (**e**, left and center) and environmental conditions and community composition both shaped P_{max} (**e**, right), whereas the strong overlap suggests the importance of a good match between environment and community.

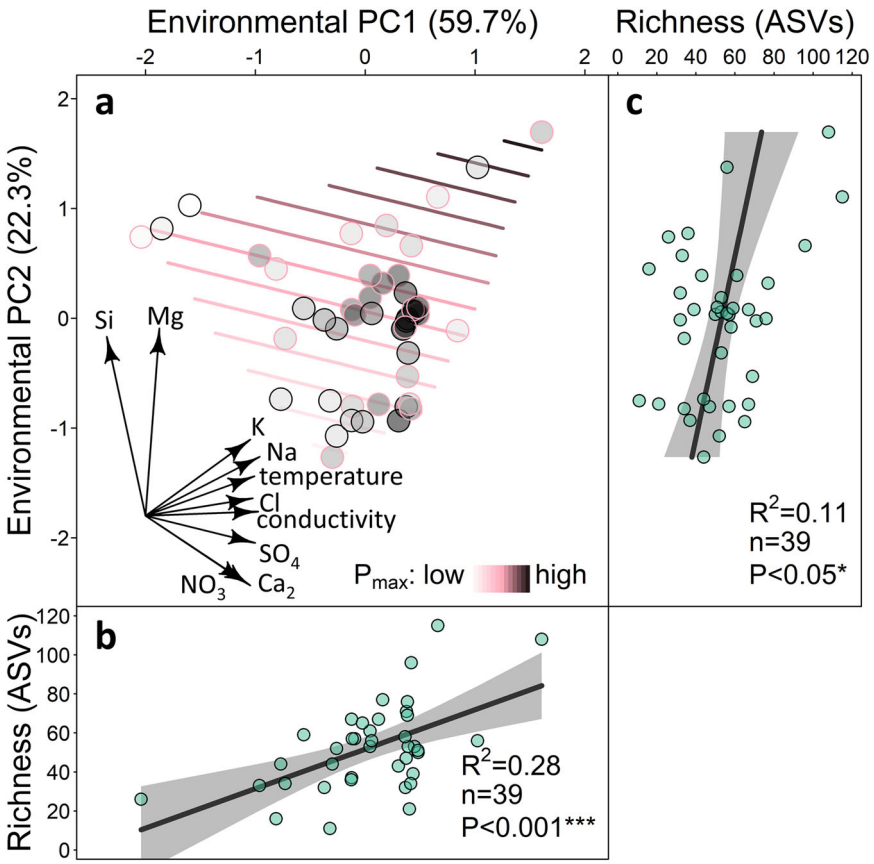


Fig. 4 | Main environmental gradients related to periphyton richness and the highest maximum production rate (P_{max}). Principal component (PC) analysis for environmental conditions (based on nutrients, temperature and conductivity for 46 sites) (**a**). Env-PC1 ordered sites along a gradient of low to high temperature, conductivity and concentrations of nutrients (Ca, K, Na, SO_4, Cl, NO_3) and env-PC2 was a gradient from low to high Mg and Si concentrations. Contour lines illustrate effects of env-PC gradients on the biomass-specific maximum production rate (P_{max}), fitted as smooth surface on the ordination diagram. P_{max} was measured only at sites with red circled dots (see Supplementary Table S5). Darkness of dots represents larger sized catchment areas. Richness, as amplicon sequence variant (ASVs) counts, increased along env-PC1 and env-PC2; the lines and the shaded areas represent predictions and the 95% confidence intervals of a linear regression model, respectively (**b**, **c**).

degree of connectivity between local habitats, thereby potentially shifting the balance between community assembly processes during a year[51].

Environmental conditions controlled primary production performance in concert with community composition. A strong predictor of P_{max} was env-PC2, a proxy for magnesium and silica (Fig. 4), which was also found to be an important niche dimension (i.e., predictor for community composition (Fig. 3d) and richness (Fig. 4)) that parallels a main environmental gradient related to the Vjosa River network's geodiversity (see Supplementary Fig. S3). The dominant algal class in the Vjosa River network are diatoms with silica-rich frustules, and magnesium is an essential component of chlorophyll. Both may constrain a community's capability of responding to environmental conditions by compositional changes and regulation of cellular chlorophyll-*a* concentration and thus affect observable P_{max}[33–36].

Importantly, the environmental influence on P_{max} was almost entirely shared with two gradients of community turnover (cc-PC1 and cc-PC3) (Fig. 3e), which identifies shifts in community composition along environmental gradients as functionally relevant for primary production. Indeed, these results point to the importance of a good match between community and environment achieved by species sorting to enable high P_{max}. Vice versa, communities driven by mass effects necessarily suffer from a poor match to the environment and have only limited P_{max}. This reasoning is supported by the concomitant shift from core to satellite taxa particularly towards the higher end of cc-PC1 paralleling increasing richness along cc-PC1 (Fig. 5). From a niche perspective, core and satellite taxa translate to generalists and specialists; then this finding is consistent with another study, which found generalists to form species-poor periphyton communities while specialists dominated highly diverse

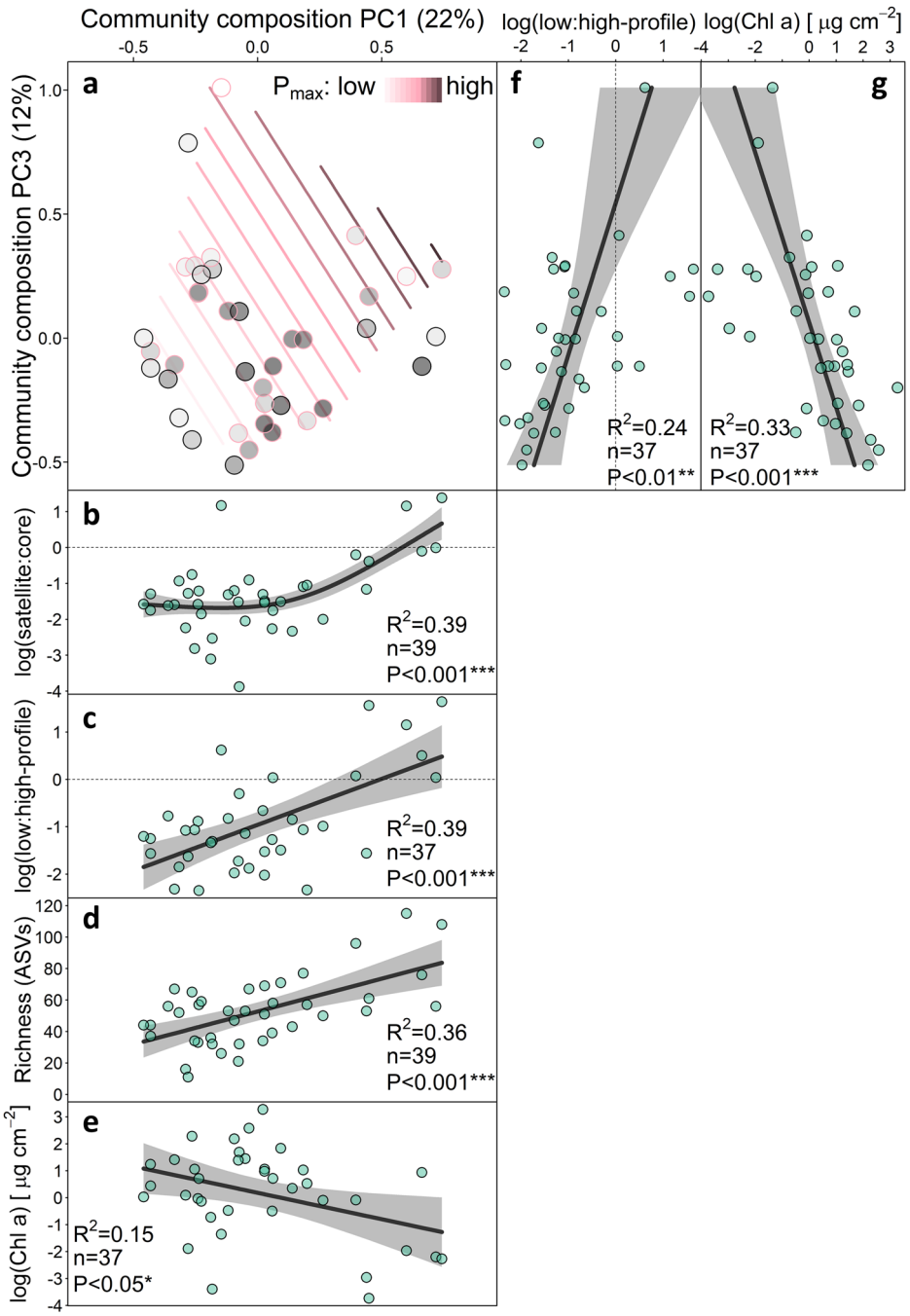


Fig. 5 | Compositional shifts of the periphyton community related to community descriptors and the highest maximum production rate (P_{max}). Principal component (PC) analysis based on a Hellinger-transformed rarefied site-by-ASV table for 39 sites (**a**, see factor loadings of ASVs in Supplementary Fig. S4). Contour lines illustrate effects of PC gradients on the biomass-specific maximum production rate (P_{max}), fitted as a smooth surface on the ordination diagram. P_{max} was measured only at sites with pink circles. Darkness of dots represents larger sized catchment area. The solid line in (**b**) illustrates a non-linear increase of the satellite to core-taxa ratio (i.e. a gradient from communities dominated by common taxa to communities dominated by rare taxa) along cc-PC1 derived by a generalised additive model; the shaded area represents the standard error. Solid lines in panels (**c**–**g**) were derived by linear regression and show shifts from high-profile to low-profile taxa along cc-PC1 (**c**), as well as an increase of richness, as amplicon sequence variant (ASVs) counts (**d**) and a decrease of biomass, as log(chlorophyll-*a* cm^{-2}, **e**), along cc-PC1, respectively. Along cc-PC3, the community composition shifted from high to low-profile taxa (**f**) and biomass increased (**g**). Dashed lines in panels (**b**, **c**, **f**) indicate where satellite vs core taxa, and high-profile vs low-profile taxa are equally abundant. Shaded areas in panels c-g are 95% confidence bands.

communities[52]. At the higher end of cc-PC1, the observed high P_{max} in species-rich communities of specialists can then be interpreted as reflecting complementarity enhancing functioning (Fig. 5). In contrast, species-poor communities dominated by widespread generalists are incapable of achieving efficient primary production. Revealingly, the dominant compositional shift of the community (cc-PC1) also correlates strongly with the ratio of low-profile to high-profile taxa and weakly with biomass (Fig. 5). This suggests the turnover in community composition to also reflect a transition from early to late successional stages (high to low end of cc-PC1). Low-profile taxa (e.g. *Achnanthidium minutissimum*, *Encyonopsis* sp.) are characteristic for early successional stages because they show higher resistance to flow events and therefore have a colonization advantage[23,53]. Functionally, the greater surface to volume ratio of low-profile taxa allows more efficient nutrient uptake and, in turn, higher biomass-specific P_{max}[53]. In contrast, succession increases self-shading in thicker periphytic mats, drives senescence, species loss and nutrient depletion, and thereby reduces biomass-specific P_{max}[33,54]. A weaker change of the ratio of low-profile to high-profile taxa and an even more notable change of biomass, both likely driven by succession, also happens along cc-PC3 paralleled by a relationship to subcatchment area. Higher disturbance frequency in smaller streams may have led to earlier successional stages, while larger rivers with stable flows allowed substantial periphyton growth[55], with implications for periphyton functioning. Altogether, diverse communities of magnesium- and silica-supplied low-profile specialists in early successional stages produced higher P_{max} compared to less diverse communities of magnesium- and silica-deficient high-profile generalists in later successional stages. It remains unclear why in-stream light did not drive P_{max} despite its important role for community composition. A possible cause for this could be that measured P_{max} is a potential and not an actually in-situ realized production rate.

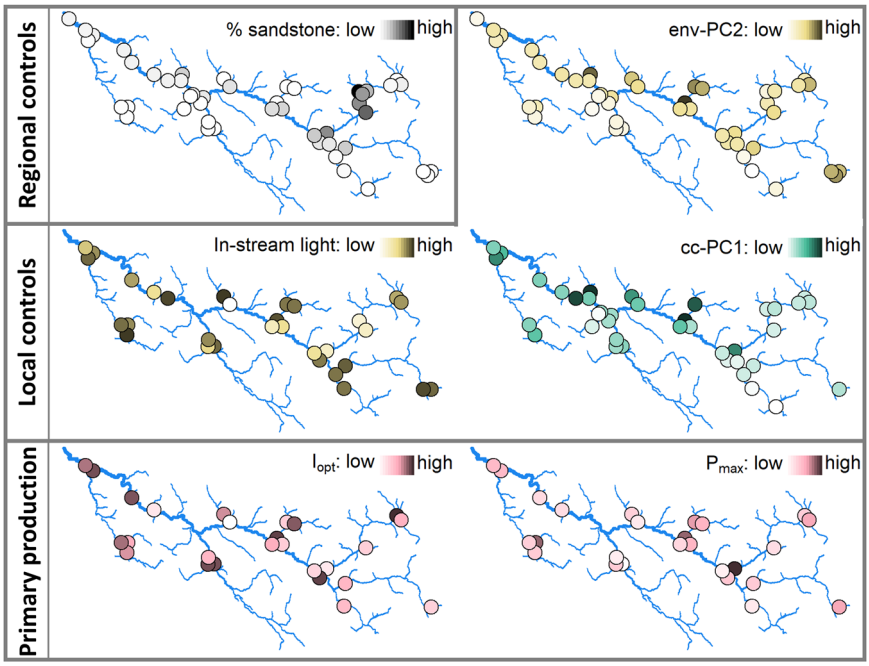


Fig. 6 | Spatial patterns of important drivers and responses across the Vjosa River network. Note that the representation of the majority of the network is maintained for all parameters, even though site count for different parameters decreases from a total of 46 (% sandstone) to a subset of 26 (I_{opt} and P_{max}) sites. Env-PC and cc-PC are principal components derived by principal component analysis (PCA) on z-standardized log(nutrients), log(conductivity) and temperature, and on Hellinger-transformed rarefied site-by-ASV table, respectively; P_{max}, biomass-specific maximum production rate; I_{opt}, light intensity at P_{max}.

Also, undetected microscale heterogeneity of light conditions across the vertical extension of periphytic biofilms may blur the in-stream light-P_{max} relationship[54].

The strong effect of in-stream light on the light intensity at the maximum gross production rate (I_{opt}) is additional evidence for adaptation of the community to environmental conditions. However, compositional changes of the community could not explain variation in I_{opt}. This is surprising, as in-stream light was found to shape community composition. Presumably, communities responded to varying light availability rather on the cellular level via upregulating chlorophyll-*a* than by compositional changes[35].

Overall, match/mismatch scenarios between the environment and community composition driven by metacommunity mechanisms and variation in successional stages influenced by river size-related disturbance frequency emerged as drivers of periphyton light physiology in the Vjosa River network. Hence, it is ultimately the Vjosa River network's drainage pattern on its geologically diverse landscape that shapes periphyton metacommunity structure-function coupling via various intermediate mechanisms. Consequently, our understanding of the role of the community and its adaptations to in-stream light in a fluvial meta-ecosystem would benefit from assessing temporal dynamics of match/mismatch scenarios between the community and in-stream light and its effects on ecosystem function.

Our study emphasizes the importance of a holistic, spatially explicit approach to understanding fluvial meta-ecosystems, primary production and ecosystem functioning in river networks. The great biodiversity and efficient biogeochemical performance of a river network depend on its spatial structure and its terrestrial matrix. The very same structure also makes them heavily vulnerable to human activities[39]. Securing future river ecosystem functioning will depend on improved conceptualization and parameterization of river networks as fluvial meta-ecosystems. To this aim, natural heritage sites like the Vjosa provide strong reference points for the fragmented, dammed and dewatered rivers dominating human landscapes. Consequently, protecting near-natural river networks can also guide regional-scale conservation and restoration of altered rivers.

Methods

The study river network and its terrestrial matrix

We investigated 46 sites mainly clustered around confluences (i.e. 2 sites shortly upstream of their confluence and one site below after water bodies were mixed) and distributed across the Vjosa River network to represent the majority of the river network and enable time-efficient fieldwork in the period 18 April-4 May 2018. We specifically chose a sampling season with expectedly intermediate discharge (according to daily records for the time period 1958–1990[39]) where it is most likely to observe the interplay between species sorting and dispersal dynamics as drivers of community composition. The river network and catchment were delineated from a 25 m resolution digital elevation model[56]. Fractions of land cover classes and geological classes were calculated for the subcatchment of each site using the Coordination of Information on the Environment (CORINE) land cover data[57] and geological data (Greece: European Geological Data Infrastructure[58]; Albania: digitalized geological map from the Geological Research and the Oil and Gas Institute of Albania, 2002), respectively. Original CORINE land cover data were grouped into seven dominating land cover classes (see Supplementary Table S3 and Supplementary Fig. S6). A total of 40 geological classes occurring in both geological databases were grouped into 13 classes (Fig. 3a, see Supplementary Table S1). Geographical data was processed in QGIS 2.18[59].

River network structure

Spatial variables derived from spatial eigenfunction analysis represent spatial patterns unfolding across the entire study area or parts thereof and are traditionally used as proxies for dispersal processes[12,17,48]. Asymmetric eigenvector map analysis models spatial variables describing spatial patterns produced by directional spatial processes, which are inherent to river networks[60]. Asymmetric eigenvector map analysis was based on hydrological distances between sites weighted by water travel time (tt) and performed in R version 3.6.1[61] using packages *aem*[62] and *watershed*[63]. To capture site connections in the river network, we created a binary site-by-edge matrix, where ones represented links (edges) in the flow-connected upstream direction of a respective site and zeros were assigned to all other links[64]. We computed the tt of a water parcel (or transported matter) based on the length of the links and the mean velocity derived from a fitted discharge-velocity power model[65,66] (see Supplementary Methods) and applied a decay function to compute weights as $1-(tt/tt_{max})^{0.5}$[12,67,68]. Maximum travel time (tt_{max}) of a single link was 1336.4 min. We then obtained the weighted site-by-edge matrix as the Hadamard product of the site-by-edge-matrix and the weights. Finally, we obtained 45 principal components of the weighted site-by-edge matrix as spatial variables. From these 45 spatial

variables, we selected only those with significant positive spatial autocorrelation using Moran's I statistic[60]. The resulting set of 15 spatial variables describing potentially detectable downstream processes were used as predictors in the global models (see Methods section "Data analysis"). When variation partitioning reveals a significant fraction of variation in community composition or environmental conditions to be explained by spatial variables, then this suggests dispersal processes or transport of matter, respectively, to generate the identified spatial structure. The detection of specific important spatial variables through forward selection (see Methods section "Data analysis") allowed us to draw conclusions about the location and extent of spatial processes.

Environmental conditions

To describe environmental conditions, we assessed concentrations of nutrients, temperature, conductivity and in-stream light conditions. We filtered stream water on site through sterile 0.2 µm membrane filters (Sartorius, Göttingen, Germany) and stored samples at 4 °C in the dark pending analysis within 4 weeks. Water samples for Ca, K, Mg, Si and Na analysis were acidified (ultrapure HCl to <0.2 pH) and analysed by an inductively coupled plasma optical emission spectrometer (ICP-OES, Thermo iCAP 6500, Thermo Scientific Fisher, Waltham, USA). SO_4, Cl and NO_3 were measured by ion chromatography (Dionex ICS-2000, Thermo Scientific Fisher). Soluble reactive phosphorus concentrations were below detection (<0.01 mg l^{-1}) at all sites and therefore not considered. Conductivity and water temperature were measured in-situ (WTW 340i portable conductivity meter, Xylem, Weilheim, Germany) at the time of sampling. Additionally, water temperature was continuously recorded at 10-min intervals for on average 7 days (minimum of 1 day) per site with MiniDOT loggers (Precision Measurement Engineering, Vista, USA) and averaged over 24 h.

Average daily solar irradiance (I, W m^{-2}) across the Vjosa River network was modeled from solar angle, time of day, date, latitude and topographical shading at each site at 10-min intervals for an assumed 2-week growth period before sampling with *r.sun* in GRASS[69,70]. Then, we used light measurements (HOBO Pendant temperature/light data loggers, UA-002-64; Onset, Bourne, USA) at the water surface (I_0) and bottom (I_z) recorded during primary production measurements in the water-filled incubation container (see Methods section "Periphyton functioning") to compute site-specific light attenuation coefficients k [m^{-1}] = (log(I_0) - log(I_z))/z. Finally, in-stream light conditions (as photon flux density) were computed from I by multiplication with 2.02[71] and using Lambert-Beer's law for a depth of 0.17 m (z), which reflects both the experimental conditions during primary production measurements and the average sampling depth of periphyton-covered stones. Our estimate of in-stream light conditions accounts for topographical shading and turbidity, but ignores local short-term weather effects and riparian shading given the open landscape character and the sampled sites' wide channels (12.9–80.7 m).

Periphyton community

The large-scale focus of our study required sampling of periphyton reflecting metacommunity structure and network-wide environmental variation rather than microhabitat differences. We thus targeted ubiquitous shallow-water periphyton communities and randomly collected 9–13 periphyton-covered stones at a depth of 10–30 cm along the shoreline of an approximately 100 m long reach, in the Vjosa River network. We assessed taxonomic composition, diversity and biomass of periphyton communities. Following incubation (see Methods section "Periphyton functioning"), we scraped the sun-facing side of the stones with a sterile scalpel and pooled the resulting slurry in a Falcon tube. The slurry was diluted with sterile-filtered stream water and homogenized before taking 1 ml aliquots for DNA extraction and pigment analysis, respectively. The first was conserved in 1 ml 99% EtOH, both were kept frozen (−20 °C) pending analysis. Sampled biofilm area was measured by image analysis from high-angle photographs of the scraped stones.

Following DNA extraction[72,73] and amplification of the 18S rRNA (primers DIV4for: 5′-GCGGTAATTCCAGCTCCAATAG-3′ and DIV4rev3:5′-CTCTGACAATGGAATACGAATA- 3′[74,75]), library preparation (2 × 300 bp) and sequencing on a MiSeq Illumina platform were performed by LGC Genomics (Berlin, Germany). We used the *DADA2* R package[76] to denoise, merge and remove chimeras from paired-end amplicons and to infer amplicon sequence variants (ASVs)[77]. We blasted ASVs against the Protist Ribosomal and the silva v132 reference databases[78,79] to assign taxonomy by using a naive Bayesian classifier[80]. We prioritised the results of the Protist Ribosomal database and only used assignments derived from the silva v132 database for ASVs which were not already assigned to the species level. Subsequently, we only retained ASVs which were assignable to phototrophic taxa at the rank of division or lower. We cannot guarantee that all non-phototrophic taxa were excluded from the study, as not all ASVs could be assigned to species and, additionally, not all species have a clear definition of lifestyles (autotrophic, mixotrophic, heterotrophic). However, we expect a stronger bias by diminishing the resolution of the community by a more rigorous exclusion of potentially non-phototrophic ASVs than by wrongly retaining the expected low number of these ASVs. Cyanobacteria, an important taxonomic group of the autotrophic biofilm community, could not be considered. ASVs were rarefied to the dataset-wide minimum of 1949 reads (site 38). We had to dismiss data from 7 sites due to bad extraction or poor sequencing results and retained a final set of 39 sites.

To describe periphyton diversity, we calculated local richness as the number of ASVs at each site. We applied additive partitioning using the function *adipart* in the R package *vegan*[81] and we tested the deviation of the observed diversity components from those expected under a null model that randomly permuted the ASV occurrence matrix 999 times by keeping original local richness and (regional) ASV occurrences fixed.

For analyses which involve community composition, we (i) excluded singletons, i.e. ASVs which occurred at only one site, as they provide only stochastic information about their environmental niche and spatial occurrence[17,82], and (ii) Hellinger transformed[83] the rarefied site-by-ASV matrix. We calculated Levins' niche width index $B_j = 1/\sum_{i=1}^{N} p_{ij}^2$, where N is the total number of sites and p_{ij} the proportion of ASV j found at site i[82,84,85]. Low B indicates unevenly distributed ASVs across only a few sites, whereas high B describes evenly distributed ASVs across many sites. B may describe gradients from specialists with narrow niche width to generalists, and also from low to high dispersers. We adopted a less interpretative classification and assigned ASVs at arbitrary cutoffs $B < 4$ and $B > 20$ as satellite taxa and core taxa, respectively. Further, we pooled assigned species to two functional groups, low-profile and high-profile growth forms, a classification found to describe resistance to physical disturbance and correlate with ability to take up nutrients[23,86]. ASVs not assigned to the low-profile or high-profile group comprise functionally ambiguous motile taxa (35), planktonic taxa (21) and taxa which are either not in the used database[86] or which could not be identified to species level, a prerequisite for an unambiguous classification. Finally, we calculated the ratio of Hellinger-transformed relative fractions of satellite to core taxa and of the low-profile to high-profile group for each site, respectively.

Aliquots for chlorophyll-*a* analysis were freeze-dried and extracted with dimethylformamide[87] before analysis by high-performance liquid chromatography (HPLC; Waters, Millford, USA). Chlorophyll-*a* was standardized by scraped stone area and used as a proxy for periphyton biomass in subsequent analysis.

Periphyton functioning

We assessed local respiration and production rates due to time constraints at a subset of 31 sites, which were chosen in a way to maintain the representation of the entire Vjosa River network (Fig. 6). We incubated periphyton-covered stones in 4 gas-proof glass chambers (23x17x9 cm, IKEA, Delft, Netherlands) filled with stream water (see Supplementary Fig. S7 for pictures of the experimental setup). The chambers were put in a lightproof plastic container filled with stream water to 17 cm depth, which itself was placed in the stream to maintain ambient water temperature. The lid of the container held a battery-powered warm-white LED light (LED

Fluter 50 W 12 V, Westech-Solar Energy GmbH, Planegg, Germany) illuminating the chambers. We ensured water movement inside the chambers and the container with battery-powered underwater toy propellers (Brandstätter Group, Zirndorf, Germany) to prevent fine sediment deposition and temperature or oxygen gradients.

Dissolved oxygen (DO) concentrations were measured non-invasively at 7 s-intervals with optode sensor spots glued to the inside of the chambers (PreSens Precision Sensing GmbH, Regensburg, Germany). Community respiration rate (CR) was calculated based on DO change during an incubation of >15 min in darkness. Net primary production (NPP) was calculated from DO changes over consecutive incubations (>10 min each) at three different light intensities achieved with neutral-density filters placed just below an artificial light source. We measured light with HOBO Pendant temperature/light data loggers at the water surface (I_0) and bottom (I_z) at a depth of 0.17 m. This setting resulted in I_0 of 30.0, 72.1 and 108.5 µmol $m^{-2}s^{-1}$. To obtain three biomass-standardized gross primary production rates (P, mg O_2 h^{-1} μg^{-1} chlorophyll-*a*) for the three light levels at each site, we (i) summed up replicated DO change measurements, (ii) multiplied by water volume in the chambers, (iii) normalized by chlorophyll-*a* per cm^2 (see below), and (iv) subtracted CR from NPP. Finally, we fitted a PI-curve using Steele's equation, $P = P_{max}\ I_z/I_{opt}\ \exp(1 - I_z/I_{opt})$[88], in an iterative inverse modeling approach with the R package *FME*[89] and using the sum of squared residuals as objective criterion. The fitted parameters P_{max} and I_{opt} mean the biomass-specific maximum gross production rate and the light intensity at P_{max}, respectively. Due to occasional malfunction of the equipment or a poor PI fit, we had to dismiss data from 5 sites and eventually retained P_{max} and I_{opt} values of 26 sites. For our study, we deemed PI-curves to study the link between community features and primary production as more appropriate than reach-scale measures of primary production (e.g. as open-channel metabolism). The latter are inherently confounded by temporally and spatially variable abiotic factors, e.g. diurnal changes in light intensity and heterogeneous light distribution, and thus may not allow to detect community-primary production relationships.

Data analysis

Our statistical approach followed classical examples of metacommunity analysis which partition effects of geographical distance and environmental conditions based on proportions of explained variation[49,90]. Geological classes, land cover classes, environmental variables (nutrients, conductivity, temperature) and the Hellinger-transformed rarefied site-by-ASV table were reduced to principle components (geo-PCs, land-PCs, env-PCs and cc-PCs) derived by four independent principle component analyses (PCA), respectively, when used as predictors. Due to its pivotal role in our study, in-stream light was not included in the PCA on environmental variables but treated separately. We followed the same strategy in five independent analyses with three sequential steps. Each analysis involved two variables (or sets thereof) as drivers (see Table 1 column "Predictor variables") and one response (or set of responses, see Table 1 column "Response variable(s)") according to our hypotheses (Fig. 1). In the first step, two separate global tests, either redundancy analysis (RDA) in case of multivariate responses, or multiple linear regression (MLR) in case of only one response, assessed the relevance of each driver for the response. For all global models including PCs as predictors, we only considered those PCs which explained at least 3% of the total variation of the respective data (see Table 1 column "Predictor variables"). This rather low threshold gives high numbers of retained PCs compared to other methods[91], yet it is necessary to obtain an overall correct Type I error in the global model forming the basis for predictor selection in the second step[92]. Predictors in global models were checked for multicollinearity using variance inflation factors[93]. In the second step, on each significant global model we applied forward selection based on two stopping criteria (significance level alpha and adjusted R^2-value) to retain only strong predictors in each driver matrix[92]. In the third step, drivers reduced to selected predictors were used to estimate relative fractions of the total variation of the response by means of variation partitioning using partial RDA or partial regression[64,90,94]. The last step resulted in three adjusted R^2-fractions, of which two were exclusively attributable to the two drivers, while one shared component was the result of their covariation and cannot be attributed to a single driver. This sequence of tests resulted, for instance, in an assessment of variation of in-stream light driven solely by the terrestrial matrix, solely by the network structure, or commonly by both. Similarly, we assessed the importance of environmental conditions and the network structure or both as drivers of community composition. Ultimately, we attributed variation of production parameters (P_{max} and I_{opt}) to local environmental conditions, community composition or both.

To better interpret effects of community composition on production parameters, we finally also tested whether functionally important gradients of community composition, i.e. forward selected cc-PCs, determined the ratio of satellite to core taxa, the ratio of low-profile to high-profile taxa, biomass and richness (see Supplementary Table S4). In case of a non-linear relationship, we applied generalized additive models (GAM). We used the R packages *packfor*[95] for forward selection, *mgcv* for running GAMs[96] and *vegan*[81] for running RDA, PCA and variation partitioning.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data that supports the findings of this study are openly available in the FigShare online repository at https://doi.org/10.6084/m9.figshare.25713882.

Received: 13 August 2023; Accepted: 3 June 2024;
Published online: 12 June 2024

References

1. Battin, T. J. et al. The boundless carbon cycle. *Nat. Geosci.* **2**, 598–600 (2009).
2. Segatto, P. L., Battin, T. J. & Bertuzzo, E. The metabolic regimes at the scale of an entire stream network unveiled through sensor data and machine learning. *Ecosystems* **24**, 1792–1809 (2021).
3. Strayer, D. L. & Dudgeon, D. Freshwater biodiversity conservation: Recent progress and future challenges. *J. N. Am. Benthol. Soc* **29**, 344–358 (2010).
4. Román-Palacios, C., Moraga-López, D. & Wiens, J. J. The origins of global biodiversity on land, sea and freshwater. *Ecol. Lett.* **25**, 1376–1386 (2022).
5. Almond, R. E. A., Grooten, M. & Petersen, T. *WWF (2020) Living Planet Report 2020 - Bending the curve of biodiversity loss*. https://f.hubspotusercontent20.net/hubfs/4783129/LPR/PDFs/ENGLISH-FULL.pdf (2020).
6. Altermatt, F. Diversity in riverine metacommunities: A network perspective. *Aquat. Ecol.* **47**, 365–377 (2013).
7. Hooper, D. U. et al. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature* **486**, 105–108 (2012).
8. Cardinale, B. J. et al. Biodiversity loss and its impact on humanity. *Nature* **489**, 326–326 (2012).
9. Benda, L. et al. The network dynamics hypothesis: How channel networks structure riverine habitats. *Bioscience* **54**, 413 (2004).
10. Battin, T. J. et al. Biophysical controls on organic carbon fluxes in fluvial networks. *Nat. Geosci.* **1**, 95–100 (2008).
11. Loreau, M., Mouquet, N. & Holt, R. D. Meta-ecosystems: a theoretical framework for a spatial ecosystem ecology. *Ecol. Lett.* **6**, 673–679 (2003).
12. Liu, J., Soininen, J., Han, B.-P. & Declerck, S. A. J. Effects of connectivity, dispersal directionality and functional traits on the metacommunity structure of river benthic diatoms. *J. Biogeogr.* **40**, 2238–2248 (2013).

13. Tonkin, J. D. et al. The role of dispersal in river network metacommunities: Patterns, processes, and pathways. *Freshw. Biol.* **63**, 141–163 (2018).
14. Besemer, K. et al. Headwaters are critical reservoirs of microbial diversity for fluvial networks. *Proc. Biol. Sci.* **280**, 20131760 (2013).
15. Holyoak, M., Leibold, M. A., Mouquet, N. & Holt, R. D. Metacommunities: spatial dynamics and ecological communities. (The University of Chicago Press, Chicago, Illinois, USA, 2005).
16. Leibold, M. A. et al. The metacommunity concept: A framework for multi-scale community ecology. *Ecol. Lett.* **7**, 601–613 (2004).
17. Horváth, Z., Vad, C. F. & Ptacnik, R. Wind dispersal results in a gradient of dispersal limitation and environmental match among discrete aquatic habitats. *Ecography (Cop.).* **39**, 726–732 (2016).
18. Declerck, S. A. J., Winter, C., Shurin, J. B., Suttle, C. A. & Matthews, B. Effects of patch connectivity and heterogeneity on metacommunity structure of planktonic bacteria and viruses. *ISME J.* **7**, 533–542 (2013).
19. Shmida, A. & Wilson, M. V. Biological determinants of species diversity. *J. Biogeogr.* **12**, 1–20 (1985).
20. Mouquet, N. & Loreau, M. Community patterns in source-sink metacommunities. *Am. Nat.* **162**, 544–557 (2003).
21. Besemer, K. Biodiversity, community structure and function of biofilms in stream ecosystems. *Res. Microbiol.* **166**, 774–781 (2015).
22. Bannar-Martin, K. H. et al. Integrating community assembly and biodiversity to better understand ecosystem function: the Community Assembly and the Functioning of Ecosystems (CAFE) approach. *Ecol. Lett.* **21**, 167–180 (2018).
23. Passy, S. I. Diatom ecological guilds display distinct and predictable behavior along nutrient and disturbance gradients in running waters. *Aquat. Bot.* **86**, 171–178 (2007).
24. Leibold, M. A., Chase, J. M. & Ernest, S. K. M. Community assembly and the functioning of ecosystems: how metacommunity processes alter ecosystems attributes. *Ecology* **98**, 909–919 (2017).
25. Fuß, T., Behounek, B., Ulseth, A. J. & Singer, G. A. Land use controls stream ecosystem metabolism by shifting dissolved organic matter and nutrient regimes. *Freshw. Biol.* **62**, 582–599 (2017).
26. Bernot, M. J. et al. Inter-regional comparison of land-use effects on stream metabolism. *Freshw. Biol.* **55**, 1874–1890 (2010).
27. Uehlinger, U., König, C. & Reichert, P. Variability of photosynthesis-irradiance curves and ecosystem respiration in a small river. *Freshw. Biol.* **44**, 493–507 (2000).
28. Mulholland, P. J. et al. Inter-biome comparison of factors controlling stream metabolism. *Freshw. Biol.* **46**, 1503–1517 (2001).
29. McGroddy, M. E., Baisden, W. T. & Hedin, L. O. Stoichiometry of hydrological C, N, and P losses across climate and geology: An envionmental matrix approach across New Zealand primary forests. *Global Biogeochem. Cycles* **22**, 1–14 (2008).
30. Allan, D. J., Erickson, D. L. & Fay, J. The influence of catchment land use on stream integrity across multiple spatial scales. *Freshw. Biol.* **37**, 149–161 (1997).
31. Young, R. G. & Huryn, A. D. Effects of land use on stream metabolism and organic matter turnover. *Ecol. Appl.* **9**, 1359–1376 (1999).
32. Hill, W. R., Boston, H. L. & Steinman, A. D. Grazers and nutrients simultaneously limit lotic primary productivity. *Can. J. Fish. Aquatic Sci.* **49**, 504–512 (1992).
33. Guasch, H. & Sabater, S. Seasonal variations in photosynthesis-irradiance responses by biofilms in Mediterranean streams. *J. Phycol.* **31**, 727–735 (1995).
34. Falkowski, P. G. Light-Shade Adaptation in Marine Phytoplankton. in *Primary Productivity in the Sea* (ed. Falkowski, P. G.) 99–119 (Springer US, 1980). https://doi.org/10.1007/978-1-4684-3890-1_6.
35. Bengtsson, M. M., Battin, T. J., Wagner, K., Schwab, C. & Urich, T. Light availability impacts structure and function of phototrophic stream biofilms across domains and trophic levels. *Mol. Ecol*. 2913–2925 https://doi.org/10.1111/mec.14696 (2018).
36. Lange, K., Liess, A., Piggott, J. J., Townsend, C. R. & Matthaei, C. D. Light, nutrients and grazing interact to determine stream diatom community composition and functional group structure. *Freshw. Biol.* **56**, 264–278 (2011).
37. Quinton, J. N., Govers, G., Van Oost, K. & Bardgett, R. D. The impact of agricultural soil erosion on biogeochemical cycling. *Nat. Publ. Gr.* **3**, 311–314 (2010).
38. Hall, R. O. et al. Turbidity, light, temperature, and hydropeaking control primary productivity in Colorado River, Grand Canyon. *Limnol. Oceanogr.* **60**, 512–526 (2015).
39. Schiemer, F. et al. The Vjosa River corridor: A model of natural hydro-morphodynamics and a hotspot of highly threatened ecosystems of European significance. *Landsc. Ecol.* **35**, 953–968 (2020).
40. Hauer, C. et al. Measuring of sediment transport and morphodynamics at the Vjosa river / Albania. *Rep. to Riverwatch* 85 (2019).
41. Kristiansen, J. 16. Dispersal of freshwater algae — a review. *Hydrobiologia* **336**, 151–157 (1996).
42. Lande, R. Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* **76**, 5–13 (1996).
43. McGuire, K. J. et al. Network analysis reveals multiscale controls on streamwater chemistry. *Proc. Natl. Acad. Sci.* **111**, 7030–7035 (2014).
44. Som, N. A., Monestiez, P., Ver Hoef, J. M., Zimmerman, D. L. & Peterson, E. E. Spatial sampling on streams: Principles for inference on aquatic networks. *Environmetrics* **25**, 306–323 (2014).
45. Heino, J., Melo, A. S. & Bini, L. M. Reconceptualising the beta diversity-environmental heterogeneity relationship in running water systems. *Freshw. Biol.* **60**, 223–235 (2015).
46. Soininen, J. Determinants of benthic diatom community structure in boreal streams: The role of environmental and spatial factors at different scales. *Int. Rev. Hydrobiol.* **89**, 139–150 (2004).
47. DeNicola, D. M. Periphyton responses to temperature at different ecological levels. in *Algal Ecology: Freshwater Benthic Ecosystems* (eds. Stevenson, R. J., Bothwell, M. L. & Lowe, R. L.) 149–181 (Academic Press, 1996).
48. Heino, J. et al. Metacommunity organisation, spatial extent and dispersal in aquatic systems: Patterns, processes and prospects. *Freshwater Biol.* **60**, 845–869 (2015).
49. Cottenie, K. Integrating environmental and spatial processes in ecological community dynamics. *Ecol. Lett.* **8**, 1175–1182 (2005).
50. Korhonen, J. J., Soininen, J. & Hillebrand, H. A quantitative analysis of temporal turnover in aquatic species assemblages across ecosystems. *Ecology* **91**, 508–517 (2010).
51. Stadler, M. & del Giorgio, P. A. Terrestrial connectivity, upstream aquatic history and seasonality shape bacterial community assembly within a large boreal aquatic network. *ISME J.* **16**, 937–947 (2022).
52. Soininen, J. & Heino, J. Variation in niche parameters along the diversity gradient of unicellular eukaryote assemblages. *Protist* **158**, 181–191 (2007).
53. Tapolczai, K., Bouchez, A., Stenger-Kovács, C., Padisák, J. & Rimet, F. Trait-based ecological classifications for benthic algae: review and perspectives. *Hydrobiologia* 1–17 https://doi.org/10.1007/s10750-016-2736-4 (2016).
54. Dodds, W. K., Biggs, B. J. F. & Lowe, R. L. Photosynthesis-irradiance patterns in benthic microalgae: Variations as a function of assemblage thickness and community structure. *J. Phycol.* **35**, 42–53 (1999).
55. Widder, S. et al. Fluvial network organization imprints on microbial co-occurrence networks. *Proc. Natl. Acad. Sci.* **111**, 12799–12804 (2014).
56. Copernicus. Digitel elevation model v1.1. https://land.copernicus.eu/pan-european/satellite-derived-products/eu-dem/eu-dem-v1.1 (2018).
57. Copernicus. CORINE land cover data. https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012 (2012).
58. European Geological Data Infrastructure. EGDI 1:1 Million OneGeology pan-european Surface Geology. http://www.europe-geology.eu (2018).
59. Quantum GIS Development Team. Quantum GIS Geographic Information System. *Open Source Geospatial Foundation Project* http://qgis.osgeo.org (2015).

60. Blanchet, F. G., Legendre, P. & Borcard, D. Modelling directional spatial processes in ecological data. *Ecol. Modell.* **215**, 325–336 (2008).
61. R-Development-Core-Team. R: a language and environment for statistical computing. (2014).
62. Blanchet, F. Guillaume Legendre, P. & Gauthier, O. AEM: Tools to construct Asymmetric eigenvector maps (AEM) spatial variables. (2015).
63. Talluto, L. watershed: Tools for Watershed Delineation. R package version 0.4.9. https://github.com/flee-group/watershed (2023).
64. Blanchet, F. G., Legendre, P., Maranger, R., Monti, D. & Pepin, P. Modelling the effect of directional spatial ecological processes at different scales. *Oecologia* **166**, 357–368 (2011).
65. Burgers, H. E., Schipper, A. M. & Hendriks, A. J. Size relationships of water discharge in rivers: Scaling of discharge with catchment area, main-stem length and precipitation. *Hydrol. Process.* **28**, 5769–5775 (2014).
66. Leopold, L. B. & Maddock, T. Jr. The hydraulic geometry of stream channels and some physiographic implications. *USGS Prof. Pap.* **242**, 1–57 (1953).
67. Dray, S., Legendre, P. & Peres-Neto, P. R. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecol. Modell.* **196**, 483–493 (2006).
68. Borcard, D., Gillet, F. & Legendre, P. *Spatial Analysis of Ecological Data. In: Numerical Ecology with R. Use R!* (Springer, Cham., 2018). https://doi.org/10.1007/978-3-319-71404-2_7.
69. GRASS Development Team. Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.2. (2017).
70. Hofierka, J. & Šúri, M. The solar radiation model for Open source GIS: implementation and applications. in *Proceedings of the Open source GIS - GRASS users conference 2002 - Trento, Italy, 11−13 September 2002* 11–13 (2002).
71. Mavi, H. S. & Tupper, G. *Agrometeorology: principles and applications of climate studies in agriculture*. (New York: Food Products Press, 2004).
72. Kolmakova, O. V. et al. Effects of zooplankton carcasses degradation on freshwater bacterial community composition and implications for carbon cycling. *Environ. Microbiol.* **21**, 34–49 (2019).
73. Nercessian, O., Noyes, E., Kalyuzhnaya, M. G., Lidstrom, M. E. & Chistoserdova, L. Bacterial populations active in metabolism of C1 compounds in the sediment of Lake Washington, a freshwater lake. *Appl. Environ. Microbiol.* **71**, 6885–6899 (2005).
74. Zimmermann, J., Glöckner, G., Jahn, R., Enke, N. & Gemeinholzer, B. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Mol. Ecol. Resour.* **15**, 526–542 (2015).
75. Visco, J. A. Environmental monitoring: Inferring the diatom index from next- generation sequencing data. *Environ. Sci. Technol.* **49**, 7597–7605 (2015).
76. Callahan, B. J., Mcmurdie, P. J., Rosen, M. J., Han, A. W. & Johnson, A. J. A. DADA2: High resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
77. Callahan, B. J., Mcmurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).
78. Morien, E. & Parfrey, L. W. SILVA v128 and v132 dada2 formatted 18s 'train sets'. https://doi.org/10.5281/ZENODO.1447330 (2018).
79. Guillou, L. et al. The Protist Ribosomal Reference database (PR^2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**, 597–604 (2013).
80. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Navïe Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
81. Oksanen, J. et al. vegan: Community Ecology Package. *Community Ecology Package* (2020).
82. Székely, A. J. & Langenheder, S. The importance of species sorting differs between habitat generalists and specialists in bacterial communities. *FEMS Microbiol. Ecol.* **87**, 102–112 (2013).
83. Legendre, P. & Gallagher, E. D. Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**, 271–280 (2001).
84. Levins, R. *Evolution in changing environments*. *Princeton University Press* (1968).
85. Pandit, S. N., Kolasa, J. & Correnie, K. Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology* **90**, 2253–2262 (2009).
86. Rimet, F. & Bouchez, A. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. *Knowl. Manag. Aquat. Ecosyst*. 1–12 https://doi.org/10.1051/kmae/2012018 (2012).
87. Wright, S. W. & Jeffrey, S. W. High-resolution HPLC system for chlorophylls and carotenoids of marine phytoplankton. in *Phytoplankton pigments in oceanography: guidlines to modern methods* (eds. Jeffrey, S. W., Mantoura, R. F. C. & Wright, S. W.) 327–341 (1997).
88. Steele, J. H. Environmental control of photosynthesis in the sea. *Limnol. Oceanogr.* **7**, 137–150 (1962).
89. Soetaert, K., Petzoldt, T. & Setzer, R. Inverse modelling, sensitivity and Monte Carlo Analysis in R using package FME. *J. Stat. Softw.* **33**, 1–28 (2010).
90. Borcard, D., Legendre, P. & Drapeau, P. Partialling out the spatial component of ecological variation. *Ecology* **73**, 1045–1055 (1992).
91. Cangelosi, R. & Goriely, A. Component retention in principal component analysis with application to cDNA microarray data. *Biol. Direct* **2**, (2007).
92. Blanchet, F. G., Legendre, P. & Borcard, D. Forward selection of explanatory variables. *Ecology* **89**, 2623–2632 (2008).
93. Groß, J. Variance inflation factors. *R News* **3**, 13–15 (2003).
94. Legendre, P. & Legendre, L. Numerical Ecology - Second English Edition. *Dev. Environ. Model*. 1–839 https://doi.org/10.1017/CBO9781107415324.004 (1998).
95. Dray, S., Legendre, P. & Blanchet, G. packfor: Forward Selection with permutation (Canoco p.46). (2016).
96. Wood, S. N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B (Statistical Methodol)* **73**, 3–36 (2011).
97. Peres-Neto, P. R., Legendre, P., Dray, S. & Borcard, D. Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* **87**, 2614–2625 (2006).

Acknowledgements

We thank Sajmir Beqiraj, Emil Birnstiel, Sonia Herrero Ortega, Olivia Wilfling, Remo Wüthrich and the staff of the water chemical laboratory at the Leibniz Institute of Freshwater Ecology and Inland Fisheries for support in the field and in the lab. Financial support came from the project FLUFLUX (ERC-STG 716196). The Vice Rector for Research at the University of Innsbruck provided support for publication fees.

Author contributions

TF, LTB and GS designed the research. TF, LTB, FW, LT, SV and GS organized the field campaign and contributed to sampling. TF and LTB performed respiration and production measurements, biofilm sampling and community analysis. TF performed statistical analysis with the help of LTB, FW, LT and GS. TF wrote the first draft of the manuscript, and LTB, FW, LT, SV and GS contributed substantially to revisions.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-024-01478-5.

Correspondence and requests for materials should be addressed to Thomas Fuß or Gabriel Singer.

Peer review information *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Martina Grecequet. A peer review file is available.