The conservation and rehabilitation of ecosystem structure and functioning in the Neotropics requires of a deep knowledge on the assembly process of communities. In this thesis, I used the metacommunity framework to dissect the relative influences of dispersal (in ecological and evolutionary timeframes), selection (driven by abiotic factors) and ecological drift on the assembly process of benthic communities in the Orinoco basin. Here, you will find evidence pointing that one or more events of dispersal limitation in an evolutionary timeframe shaped distinct pools of taxa in the basin. The extent and composition of these pools, which partially matches the distribution of the ecoregions, determine the main drivers of the community assembly. Within each ecoregion, dispersal, selection and drift interact to constrain the structure and dynamics of communities and metacommunities among and within streams. Depending on the taxa belonging to each pool of species, the role of one of these processes may prevail over the others.

2020 - DISSECTING THE ASSEMBLY PROCESS OF BENTHIC COMMUNITIES FROM NEOTROPICAL STREAMS

DISSECTING THE ASSEMBLY PROCESS OF BENTHIC COMMUNITIES FROM NEOTROPICAL STREAMS

DOCTORAL THESIS

2020

JUAN DAVID GONZÁLEZ-TRUJILLO

Universitat de Girona

ICRA Catalan Institute for Water Research

UNIVERSIDAD NACIONAL DE COLOMBIA

Dissecting the assembly process of benthic communities from Neotropical streams

Juan David González-Trujillo

Juan David G

Universidad Nacional de Colombia, Bogotá, Colombia.
Universitat de Girona, Girona, España.
Institut Català de Recerca de l'Aigua, Girona, España.
2020

Dissecting the assembly process of benthic communities from Neotropical streams

Juan David González-Trujillo

Tesis presentada como requisito parcial para optar al título de:
Doctor en Ciencias – Biología
Universidad Nacional de Colombia
Ph.D. in Water Science and Technology
Universidad de Girona

Director por la Universidad de Girona:
Dr. Sergi Sabater
Director por la Universidad Nacional de Colombia:
Dr. Jhon Ch. Donato-R

Codirectora por la Universidad de Barcelona:
Dra. Isabel Muñoz

Tesis con fines de doble titulación bajo el Convenio de cotutela entre la Universidad de Girona y la Universidad Nacional de Colombia.

Universidad Nacional de Colombia, Bogotá, Colombia.
Universitat de Girona, Girona, España.
Institut Català de Recerca de l’Aigua, Girona, España.
2020

“The most unique feature of earth is the existence of life, and the most extraordinary feature of life is its diversity.”

Cardinale et al., 2012

A todas aquellas personas que
de una u otra manera contribuyeron a
que tenga hoy una historia que contar.
En particular, aquellas de mi familia
biológica y no biológica.

JD, 2019

Agradecimientos

Esta es una de las secciones más difíciles de escribir, en sólo cuatro años pueden pasar tantas cosas… que… no es extraño sentirse en un berenjenal mientras recuerdas y escribes acerca de todas aquellas personas que pusieron su granito de arena no sólo en el ámbito académico, sino también en lo personal. Amigos y amigas dispuestas a darme sus dos manos e infinidades de sopapos para corregir el rumbo (y de vez en cuando una que otra palmadita en la espalda… haha).

A todas y todos, en serio, ¡Muchas gracias!

Para no desentonar con una de las ideas centrales de la tesis (‘*History matters!*’), me tomo la libertad de agradecer cronológicamente a todas aquellas personas que contribuyeron a construir estas páginas que ven frente a ustedes…

Primero que todo, por supuesto, quiero agradecerle a mi familia. En especial a mi papá, mamá y hermano. Muchas de estas líneas jamás se hubieran escrito si ustedes no hubieran sido mis principales alcahuetas. Siempre apoyándome y soportándome, inclusive desde la distancia, me dieron el ánimo, comida y recursos más que de sobra para sacar adelante las diferentes fases de este proyecto.

Hablando de alcahuetas, siempre estaré muy agradecido contigo **Liz**. Gracias por acompañarme, soportarme y aconsejarme durante el final y el comienzo de un nuevo viaje. Que haya tenido los datos para hacer esta tesis es en gran parte culpa tuya. Gracias por la compañía durante todos los viajes, las horas que dedicaste a ayudarme a buscar bichos en sopas de sedimento, y como no, a las risas e historias que amenizaron las jornadas en carro.

A **Juan Pablo**, **Gabriela** y **Katterin**, quiénes también acompañaron en las aventuras buscando y muestreando los ríos de la tesis, días cargados de trabajo y noches cortas con una que otra cervecita para romper con la monotonía. Sin ustedes el campo hubiera tardado un tiempo considerable más y, por tanto, no tendría el conjunto de datos con el que pude jugar durante dos años. A **Gabriela** y **Katterin**, también quiero agradecerles por permitirme crecer junto a ustedes y sus proyectos de tesis de maestría. Puedo decir que una buena parte de las grandes enseñanzas que me dejaron estos cuatro años vienen de nuestro trabajo en conjunto.

En el comienzo de esta aventura, dos mujeres muy berracas ayudaron a darle mucha forma a la propuesta de tesis. **Argenis Bonilla** y **Dolors Armenteras**, muchas gracias por el tiempo que sacaron y las ganas que le pusieron a la revisión de la propuesta de tesis. De la primera presentación a la versión final hubo una diferencia abismal, y eso se los debo a las dos. ¡Muchas Gracias! También me gustaría agradecerle a **Jhon** e **Isabel** por su amistad, los buenos consejos y la gran compañía que representaron durante estos años. A **Jhon,** y especialmente a **Elizabeth,** les doy mil gracias por la mano que me dieron para sortear los mil y un trámites burocráticos que iban apareciendo como matas de monte en claros amazónicos.

Para terminar de agradecer la alcahuetería, solo me falta nombrar al "jefote". **Sergi**, de todo corazón, ¡Muchas gracias! Si no te hubieras tomado el tiempo de sentarte conmigo por allá en una casa llena de cocodrilos en los Llanos Orientales y proponerme hacer una tesis "de ríos colombianos", quién sabe que estaría haciendo ahora… Con toda seguridad, no hubiera tenido la oportunidad de hacer una tesis en mi país. A pesar de nuestras discusiones y diferencias intermitentes (como esos ríos que tanto te gustan), siempre te estaré agradecido por la (infinita) paciencia que me tuviste estos años y, sobretodo, por tu amistad y por abrirme la puerta de tu casa y presentarme esa familia tan peculiar que tienes. Dale mis gracias a Araceli, Laia, Pau, Nabil, Núria, María y Sara por esas tardes familiares que me regalaron. Por supuesto, no puedo olvidarme de agradecerte por tener siempre dos minutos para hablar de filosofía barata, y por supuesto, por el centenar de galletitas que me robé de tu despacho.

Las salidas de campo no hubieran sido tan amenas y divertidas sino hubiera sido por los anfitriones que abrieron las puertas de sus casas o reservas. A Doña **Miriam**, Don **Javier**, **Marcos y Diana**, gracias por abrirme la puerta del Guaviare y recibirme siempre como con una gran sonrisa, un cafecito y un buen pescaito con patacón. A **Miguel Rodríguez** y su familia por abrinos su cabaña: puerta al Páramo y la alta montaña. Finalmente, al equipo del parque **La Chorrera** por permitirnos entrar y disfrutar de las hermosas quebraditas que por allí corren.

Mi tiempo en Girona fue muy divertido en parte por el grupo de amigos y amigas que se fue aglutinando poco a poco como crema Catalana. **Albert**, parcerito, muchas gracias por ser el primer gran anfitrión a mi llegada a la comarca. Eres un crack como corredor, como persona y como bailarín, y también, ¡como ingeniero! Gracias por tomarte el tiempo para enseñarme acerca de los caudales y de cómo estimarlos a partir de datos secundarios. El método de Témez… como golazo en el descuento. Junto a **Dídac** y **Julio**, las conversaciones durante la comida no volverán a ser iguales sin la intervención del consejo de sabíos.

Quisiera también agradecer a todos y todas en **SomRIE**. A Vicenç, Rafa, Pep, Núria, Carmencita, Núria P., Anna, Duygu, Carmè, Dani, Dídac, Oriana, Arianna, Jordi, Miriam, Ferran, Elisabet, Nina, Albert, Sergi, gracias por compartir tiempo, conocimientos, malos chistes, som-risas y, como no, alguna que otra cerveza y con fuet y olivas. A la **Dra. Catalán**, a **Dani** y a **Dídac** quisiera darles un agradecimiento muy especial. **Núria**, ya te lo dije antes: "@NCatalanGarcia is one of the women I admire the most! Leading collaborative projects, untangling the carbon cycle, writing cool stories, promoting #WomenInScience or giving good-but-old-fashion advice (and sometimes partying), she's always inspiring!" ¡Gracias por las buenas charlas, cuentos y risas! A **Dani** – no se me moja la canoa diciéndolo: ¡me hace muy feliz que se cruzaran nuestros caminos en un país tan lejano! Gracias por tus enseñanzas, paciencia, buena vibra, comida, y por supuesto, bromas costeñitas. Todo pasa por algo, ya tu saé… ⇑ A **Dídac**, gracias por ser tan original… Aunque digas lo contrario, he aprendido mucho de ti y de tu forma de ser ¡NO cambies nunca porfa!

También me gustaría mencionar a dos miembres del trío Calavera: ¡gracias **Ferri**, gracias **Miriam,** por hacer tantos momentos memorables! Mai oblidaré les nostres aventures al Duero i a la tomatina, banderes del trío calavera. Gràcies per alegrar els congressos i nits a Girona. Aunque siempre creeré que no debimos marcharnos de Porto… Un ron colita i a dormir! **Moltes mercès**

El **ICRA**, aunque no lo quieras, se vuelve parte de tu día a día y allí es donde conectas con personitas muy especiales. De sus historias y trabajos aprendes día a día, y entre tantos planes, cuando menos te das cuenta se convierten en algo así como una "familia institucional". Gracias **Giannis, Esther, Fede, Luca, Nick, Duygu, Natalia, Niko** por brindarme su amistad y tiempo para aprender un poquito de cada uno. Entre chiste y chanza, entre conciertos y bicicrucis, entre zampanzar y Salvador Catà, siempre quedan muchas muchas enseñanzas y buenos recuerdos.

Unas gracias totales a mi familia no sanguínea en Girona. **María**, **Fede**, **Luca**, **Rami** y **Giannis**. Gracias por adoptarme durante estos años. Sé que nos faltó compartir uno que otro invierno (haha), pero creo que tenían que haber momentos de distancia para poder extrañarnos y así juntarnos de nuevo alrededor de vermú toreros hasta las tantas… Dirán que poco o nada aportaron a esta tesis, pero no. Algunas de las ideas de este documento vienen de nuestras conversaciones y de una que otra anécdota. Además, el final de una tesis no es documento como tal, sino tu crecimiento personal. En este sentido, ustedes aportaron muchísimo. Siempre ayudándome a ver las cosas desde otra perspectiva, sacando un poco de tiempo para escucharme, dándome muchos consejos de cómo proceder y, de vez en cuando, jalándome las orejas merecidamente.

Hablando de consejos, jalones de oreja y confesionarios, quisiera agradecerles a todos aquellos que en el algún momento pausaron un momento de su vida para escucharme. A esas personas especiales que abrieron un poco de su mundo para que me sintiera en confianza de compartir un poco del mío. En especial, a **Sebastián**, hermano, muchas gracias por estar siempre ahí sin importar las diferencias horarias cambiantes. Siempre alentando, siempre escuchando y nunca juzgando. ¡Gracias totales! A **Marisol** y **Pacho**, esas cuatro manos siempre llegaron cuando más las necesitaba. Muchas gracias por acompañarme incondicionalmente durante estos años, gracias por todos esos cafés y pausas para el *brunch*… ¡mil gracias por sacarme a pasear en el momento que más lo necesitaba! Finalmente, a **Nata**, gracias por aparecer siempre para animarme con un poco de Bullying, un café y un buen libro. **¡Un gran abrazo para todos!**

A todos aquellos amigos que conocí alrededor del Mundo y que me dejaron más de una enseñanza en la ecología o la biogeografía: **Juan Escamilla**, **Boni**, **Dylan Craven**, siempre ando expectante de que en algún momento podamos repetir esos buenos momentos compartimos juntos. A **Dani Petsch** y a **Víctor Saito**, dos brasileros que con gusto aceptaron discutir ideas conmigo y me impregnaron de su alegría brasileiiiira. To the "**MetaCom group**" – **Miguel, Romain, Julie, Núria, George, Sophie, Alberto** – thanks for the advice and the good vibe you have when we meet. **I hope to meet you soon!**

No puedo olvidar agradecerle a **Lina Corredor** por tantos años de amistad y tantas charlas acerca del pasado, presente y futuro. Y por supuesto, por ponerle tanto cariño y dedicación a la portada que envuelve este documento ¡Me encantó!

También debo agradecer a la Universidad Nacional de Colombia, a la Universidad de Girona y el Institut Català de Reserca de l'Aigua (ICRA) por acogerme y prestarme sus diferentes facilidades durante estos años. También al Departamento Administrativo de Ciencia, Tecnología e Innovación "Colciencias" y al Secretariado de universidades e investigación de la Generalitat de Catalunya y el fondo europeo social por los fondos recibidos para adelantar los estudios doctorales. Por ultimo, a la dirección de Investigación Sede Bogotá y su Convocatoria nacional de proyectos para el fortalecimiento de la investigación, creación e innovación de la universidad nacional de Colombia 2016-2018 (DIB proyecto No. 34856).

Finalmente, quisiera agradecer especialmente a los evaluadores del proyecto y del documento final, y a aquellos revisores anónimos que espero algún día conocer. A su vez a todas aquellas y aquellos líderes sociales y ambientales que han contribuido a que después de 500 años aún tengamos ríos prístinos que estudiar **¡A su memoria y bienestar!**

Resumen

- La conservación y rehabilitación de la estructura y funcionamiento de los ecosistemas requiere de un conocimiento profundo de las causas y consecuencias de su biodiversidad. Sin embargo, este conocimiento es aún escaso en regiones Neotropicales.
- En esta tesis, utilizo el marco conceptual de las metacomunidades para caracterizar los posibles efectos que tienen los procesos de *dispersión* (en tiempos ecológicos y evolutivos), *selección* (impulsada por factores abióticos) y *deriva ecológica* en la diversidad y distribución de las comunidades bentónicas fluviales.
- La zona de estudio, de un área de aproximadamente 40,000 km^2, abarcó entre 26 y 32 segmentos de ríos prístinos del Orinoco colombiano. Los puntos de muestreo abarcaron un gradiente de elevación de 300 a 3400 m.s.n.m. que incluyó un conjunto heterogéneo de ecorregiones y paisajes.
- Mediante una aproximación de ligar patrones y posibles mecanismos, esta tesis proporciona evidencias de que la *dispersión*, la *selección* y la *deriva* están directamente involucradas en el proceso de ensamblaje de las comunidades bentónicas fluviales.
- Mis hallazgos indican que uno o más eventos de limitación de la dispersión en un marco de tiempo evolutivo (eventos de aislamiento alopátrico) formaron diferentes *pools* de especies dentro de la cuenca del Orinoco. La extensión de estos *pools* coincide parcialmente con la distribución de las ecorregiones, lo que sugiere que los eventos que moldearon los paisajes fluviales y la estructura de la vegetación afectaron de manera similar la diversidad y distribución de las especies bentónicas en ecosistemas fluviales.
- Adicionalmente, dentro de cada ecorregión, la *dispersión*, la *selección* y la *deriva* están interactuando para restringir la estructura y la dinámica de las comunidades y metacomunidades entre y dentro de los ríos. Dependiendo de la comunidad (p.e. diatomeas o insectos), el papel de cada uno de estos procesos puede prevalecer sobre el de los demás.
- Estos hallazgos tienen implicaciones tanto para la investigación básica como para la aplicada (p.e. biomonitoreo) en las disciplinas de la ecología de metacomunidades y de agua dulce, así como en la conservación y la biogeografía.

Palabras clave: metacomunidad, paisaje fluvial, legado histórico, pool de especies, Colombia.

Abstract

- The conservation and rehabilitation of ecosystem structure and functioning requires of a deep knowledge on the causes and consequences of its biodiversity. The assembly of Neotropical communities, particularly in riverine ecosystems, remains to be dissected.
- I used the metacommunity framework to dissect the relative influences of *dispersal* (in ecological and evolutionary timeframes), *selection* (driven by abiotic factors) and *ecological drift* on the assembly process of freshwater benthic communities.
- The study was carried out at 26-32 different stream segments within an area of about 40,000km^2, in the Colombian Orinoco. The area encompasses an elevation gradient from 3400 to 300m a.s.l. and includes a heterogeneous assembly of ecoregions and landscapes.
- By using a pattern-matching approach, I provide evidences supporting that *dispersal*, *selection* and *drift* are directly involved in the assembly of freshwater benthic communities.
- My findings indicate that one or more events of dispersal limitation (i.e. allopatric isolation) in an evolutionary timeframe shaped distinct pools of taxa in the Orinoco basin. The extent of these pools partially matches the distribution of the ecoregions, suggesting that those events molding the riverscapes and the vegetation structure similarly affect the diversity and distribution of benthic species.
- Within each ecoregion, *dispersal, selection* and *drift* interact to constrain the structure and dynamics of communities and metacommunities among and within streams. Depending on the taxa belonging to each pool of species, the role of one of these processes may prevail over the others.
- These findings have implications for both basic and applied research in the disciplines of metacommunity and freshwater ecology as well as of conservation and biogeography.

- **Keywords:** metacommunity, riverscape, historical legacy, species pool, Colombia.

Resum

- La conservació i rehabilitació de l'estructura i funcionament dels ecosistemes requereix d'un coneixement profund de les causes i conseqüències de la seva biodiversitat. No obstant això, aquest coneixement és encara incipient en regions Neotropicales.
- En aquesta tesi, utilitzo el marc conceptual de les metacomunidades per caracteritzar els possibles efectes que tenen els processos de dispersió (en temps ecològics i evolutius), selecció (impulsada per factors abiòtics) i deriva ecològica en la diversitat i distribució de les comunitats bentòniques.
- La zona d'estudi, d'una àrea d'aproximadament 40,000 km2, va abastar entre 26 i 32 segments de rius prístina de l'Orinoco colombià. Els punts de mostreig van abastar un gradient d'elevació de 300 a 3400 m.s.n.m. que va incloure un conjunt heterogeni de ecoregions i paisatges.
- Mitjançant una aproximació de lligar patrons i possibles mecanismes, aquesta tesi proporciona evidències que la dispersió, la selecció i la deriva estan directament involucrades en el procés d'encaix de les comunitats bentòniques.
- Els meus troballes indiquen que un o més esdeveniments de limitació de la dispersió en un marc de temps evolutiu (esdeveniments d'aïllament alopátrico) van formar diferents pools d'espècies dins de la conca de l'Orinoco. L'extensió d'aquests pools coincideix parcialment amb la distribució de les ecoregions, el que suggereix que els esdeveniments que van modelar els paisatges fluvials i l'estructura de la vegetació van afectar de manera similar la diversitat i distribució de les espècies bentòniques en ecosistemes fluvials.
- Addicionalment, dins de cada ecorregión, la dispersió, la selecció i la deriva estan interactuant per a restringir l'estructura i la dinàmica de les comunitats i metacomunidades entre i dins dels rius. Depenent de la comunitat (p.e. diatomees o insectes), el paper de cada un d'aquests processos pot prevaler sobre el dels altres.
- Aquestes troballes tenen implicacions tant per a la investigació bàsica com per l'aplicada (p.e. biomonitoreo) en les disciplines de l'ecologia de metacomunidades i d'aigua dolça, així com en la conservació i la biogeografia.

Paraules clau: metacomunitat, paisatge fluvial, llegat històric, pool d'espècies, Colòmbia.

Table of Contents

List of boxes

List of figures

List of tables

List of publications

I. **González-Trujillo, J. D.,** Donato-Rondon J. Ch, Muñoz, I., Sabater, S. (2019) Historical processes constrain metacommunity structure by shaping different pools of invertebrate taxa within the Orinoco basin. **Diversity and Distributions.** https://doi.org/10.1111/ddi.12996

II. **González-Trujillo, J. D.,** Pedraza-Garzón, E., Donato-Rondon J. Ch, Sabater, S. (*accepted*) Ecoregional characteristics drive the distribution patterns of Neotropical stream diatoms. **Journal of Phycology.**

III. **González-Trujillo, J. D.**, Petsch, D., Córdoba-Ariza, G., Rincón-Palau, K., Donato-Rondon J. Ch, Castro-Rebolledo, M., Sabater, S. (2019) Upstream refugia and dispersal ability may override benthic-community responses to high-Andean streams deforestation. **Biodiversity and Conservation 28: 1513**. https://doi.org/10.1007/s10531-019-01739-2

IV. **González-Trujillo, J. D.,** Saito, V., Petsch, D., Muñoz, I., Sabater, S. (*submitted*) Historical legacies and contemporary processes shape beta diversity in Neotropical montane streams.

List of Scientific conferences

I. **González-Trujillo, J. D.**, Muñoz, I., Sabater, S. 2019. Biogeographical legacies constrain invertebrate and diatom metacommunity structure within a heterogeneous basin. *11th Symposium for European Freshwater Sciences. Zagreb, Croacia.* Award: 1st Colyn Reynolds Price for the Best oral presentation by a young scientist.

II. **González-Trujillo, J. D.** 2019. Historical legacies and contemporary constraints explain invertebrate and diatom metacommunity structure within a heterogeneous basin. *2nd International Biogeography Society Conference – Humboldt 250. Quito, Ecuador.*

III. **González-Trujillo, J. D.** 2019. The role of Historical events in the assembly process of benthic communities in the Orinoco basin. *6th Fresh Blood for Fresh Water Conference. Tihany, Hungría.*

IV. **González-Trujillo, J. D.** Diversidad funcional de invertebrados aplicada a la evaluación y mitigación de perturbaciones en ríos de altoandinos. 2018. *1er simposio en diversidad funcional en zoologia - V Congreso colombiano de Zoología. Bogotá, Colombia.*

V. **González-Trujillo, J. D.,** Donato-Rondon J. Ch, Muñoz, I., Sabater, S. 2018. Ecoregional features explain invertebrate diversity at local and regional scales in rivers from the Colombian Orinoco basin. *XIX Conferencia de la Asociación Ibérica de Limnología. Coimbra, Portugal.*

List of funding projects

I. **Proyecto:** evaluación de los impactos ligados a la pérdida del bosque de ribera a través de las redes tróficas en ecosistemas fluviales en la Serranía La Lindosa (San José de Guaviare, Colombia).
Código:39731
Código QUIPU: 201010028802
Duración: 12 meses (Finalizo en marzo de 2020).
Vinculación: investigador principal.
Tareas: muestreo de las comunidades, procesamientos de muestras, análisis de datos, redacción de informes.

II. **Proyecto:** Linking functional diversity patterns of algae and invertebrates to scale-dependent constrains of rivers from the Orinoco basin
Código:40172
Código QUIPU: 201010027596
Duración: 24 meses (Finalizo en febrero de 2020).
Vinculación: co-investigador.
Tareas: muestreo de las comunidades, procesamientos de muestras, análisis de datos, redacción de informes.

III. **Proyecto:** Diversidad de algas en un contexto multi-escalar: implicaciones para el biomonitoreo de ecosistemas fluviales colombianos.
Código:34856
Código QUIPU: 201010025963
Duración: 14 meses (Finalizó en octubre de 2018).
Vinculación: co-investigador.
Tareas: muestreo de las comunidades, procesamientos de muestras, análisis de datos, redacción de informes.

General Introduction

The conservation and rehabilitation of ecosystem structure and functioning requires a deep knowledge on the causes and consequences of its biodiversity (McGill *et al.*, 2006; Pavoine & Bonsall, 2011) or, in other words, on the understanding of the major drivers of biodiversity patterns and how such patterns influence community assembly and ecosystem functioning**.** In light of this, **dissecting the assembly process (Box 1.1)** of biological communities has become one of the aims - and major challenges - of ecologists nowadays (Lessard *et al.*, 2012; Violle *et al.*, 2014; Mori *et al.*, 2018)**.** However, this is not a completely straightforward process, as **the assembly process** vary according to the type of biological community, the type of ecosystem (i.e. aquatic or terrestrial) or the spatio-temporal scale considered in the context of each study (Heino *et al.*, 2012; Heino, 2013; Datry *et al.*, 2017).

Historically, research on the **community assembly process** has focused on answering two pivotal questions: ***Why are there so many species and which are the major determinants of species distribution and abundance in biological communities?*** Starting from Darwin and Wallace's Natural Selection and passing through Kimura and Hubbell's Neutral Theory of Diversity (Kimura, 1991; Hubbell, 2006), multiple theories and hypotheses have been proposed (and tested) in attempting to provide an ultimate answer. Hutchinson (1959), for instance, presented one of the first attempts to answer both questions from an ecological perspective. In his famous '**Homage to Santa Rosalia'**, Hutchinson gathered several lines of theoretical and empirical evidence to conclude that the reason behind species diversity lies in the fact that complex trophic organization of a community is more stable than a simple one.

Hutchinson’s seminal work represented a cornerstone on which ecological research was built in America. Research was particularly focused on those controls limiting the possible diversity proposed by Hutchison: namely, space, niche requirements and the mosaic nature of the environment. For example, MacArthur and Wilson (1967) explored the role of space as a major driver of diversity in their Theory of Island Biogeography. Hardin (1960), Diamond (1975) and Connor and Simberloff (1979), on the other hand, focused on testing the role of competence, limiting similarity and character displacement in shaping biodiversity patterns. Meanwhile, other researchers focused on disentangling the effect of environmental heterogeneity in biodiversity patterns (see the recent review of Cadotte & Tucker, 2017 and references therein).

Box 1 The community assembly process

The **community assembly process** portrays all those processes triggering diversity at different scales (Pavoine & Bonsall, 2011).

The ‘assembly process’ was firstly coined by Jared Diamond (1975) to explain the diversity and distribution of Pidgeon species in an insular system. By using species life history, dispersal and competitive ability, Diamond proposed a set of ‘**assembly rules**’ that determines which species will coexist and be part of the same community in a specific island of the insular system.

Nowadays, the **assembly** of communities is viewed as the **process** where the successful colonization and coexistence of species at the local scale is determined by the dispersal of species from a regional pool and the subsequent selection by biotic and abiotic constraints (Vellend, 2010; Lessard *et al.*, 2012; Mori *et al.*, 2018).

Parallel to the developments in America, alternative explanations to the two pivotal questions surrounding the community assembly process were developed in Europe. In general, these explanations were more holistic in nature. Margalef (1951), for instance, claimed that ecology, genetics and evolution should be considered at the same time to explain the relationship between the species richness and the number of individuals. He was also one of the first researchers who proposed the historical biogeography as a major factor to be considered for understanding diversity patterns at large scales. As would be later stated by Ricklefs (1987), Margalef sugested that patterns of diversity at the local scale cannot be explained without considering the diversity at the regional scale, which has been majorly determined by past geological and climatic events.

The introduction of the hierarchical theory (O'Neill et al. 1986) and the recognition of the scale-dependency of the observed patterns (Wiens, 1989) brought to light the importance of setting the **community assembly process** in a multi-scalar framework. In this regard, following studies tried to account for the relative influence exerted by ecological processes constraining species occurrence at the local scale, by processes constraining species dispersal at larger scales, and by evolutionary and biogeographic processes that lead to variation in the diversity and composition of the **regional pools of species** (Buschke *et al.*, 2014; Cornell & Harrison, 2014; Mittelbach & Schemske, 2015). However, the fact that multiple processes may be simultaneously invoked, led to the perception that community ecology was a “mess” or a “black box”, on which too much contingency made difficult to state useful generalizations (Lawton, 1999).

Recently, Vellend (2010) proposed that **community-level patterns** are shaped by the interaction of **‘big four’ processes (Box 1.2): *speciation***, ***dispersal***, ***selection*** and ***ecological drift***. By doing this, Vellend’s theory aimed at synthetizing all the previous research and proposing a framework to study the assembly of biological communities. These ‘big four’ processes encapsulate all the ‘minor processes’ that have been found to trigger biodiversity at different scales. For instance, the competitive exclusion and the environmental filtering, which are *a priori* two distinct processes, are ultimately ***selection*** processes, since they are selecting the fittest species.

Hence, by embracing this framework, community ecologists have a relatively simple framework to dissect the assembly process and assess the relative influence of the main drivers.

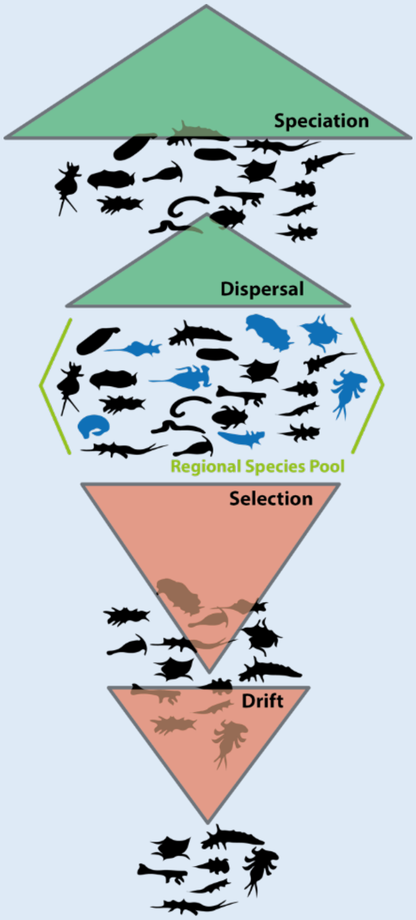
As first ascertained by Wiens, **scale** cannot be ignored when studying the **assembly process** (Wiens, 1989; Leibold & Chase, 2018; Morton *et al.*, 2018). Indeed, one of the reasons behind the context dependence in community-level patterns lies on the pervasive influence of scale is not adequately included (Chave, 2013). In this regard, the **metacommunity framework (Box 1.3)** offers a complementary framework to study the assembly process, since it explicitly considers more than one scale simultaneously (Leibold *et al.*, 2004). Moreover, by explicitly including ***dispersal*** in assembly models, the metacommunity framework binds together the effects of local-scale processes such as ***selection*** and ***drift*** and those of shaping diversity at regional-scale processes such as ***speciation***.

In freshwaters, research on the assembly process, and particularly, on ***dispersal-driven processes*** using the metacommunity framework have gained attention during the last decade (Brown *et al.*, 2011; Altermatt, 2013; Heino, 2013). Indeed, the use of this framework has provided new insights into the effect of the unique features of riverine landscapes – such as dispersal ability and connectivity-related constraints - in the assembly of freshwater communities (recently reviewed by Tonkin *et al.,* 2018).

Box 2 The big four processes and the role of the regional species pool

Aiming to synthetize the processes underlying the assembly of biological communities, Vellend borrowed some ideas from genetics and proposed that, as with (micro)evolutionary change, community dynamics is underlain by just four high- level processes: ***selection***, ***drift***, ***dispersal***, and ***speciation*** (Vellend, 2010).

The framework resulting from this perspective considers that there are only **four fundamental** ways to change the **number of species (species richness)**:



1. Speciation: The creation of new species. It occurs within evolutionary (or microevolutionary) timeframes, when subpopulations diverge from others, forming new species.

2. Dispersal: The movement of organisms across space. It can occur in ecological or evolutionary timeframes when an individual of a species arrives from somewhere else and joins the local community.

These two processes are the major determinants of the **regional pool of species**.

A "**pool**" or "**regional pool**" of species refers to the set of all species available to colonize local communities (Srivastava, 1999). Since the introduction of this concept in the Theory of Island Biogeography, species pools have helped to disentangle the relative contributions of local, regional and historical factors in the assembly process (Cornell & Harrison, 2014).

3. Selection: A deterministic fitness difference between individuals of different species. If a species has a fitness advantage over other species, it can exclude them from the community.

4. Drift: Random changes in species relative abundances. Even if all individuals have the same chances of survival and reproduction, there is a nonzero probability that, some individuals die before reproducing. If this happens, the number of species will decrease.

Metacommunity-based studies have brought to light that freshwater communities are mainly shaped by the interplay between ***selection*** and ***dispersal*** as well as ***ecological drift*** (Thompson & Townsend, 2006; Heino *et al.*, 2015b; Siqueira *et al.*, 2020). Regarding to ***selection***, habitat heterogeneity and disturbance regimes are the main determinants of community structure at the local scale (Poff *et al.*, 1997; Heino *et al.*, 2003; Lepori & Hjerdt, 2006). Spates, floods and droughts are the main factors of disturbance, and major determinants of the habitat heterogeneity in riverine ecosystems (Townsend & Hildrew, 1994; Poff, 1997; Ward, 1998; Lake, 2000). However, these determinants usually fail to account for the variability in community species composition, since their relative influence is highly dependent on the biological community under study: algae and invertebrates, for instance, have different life-history strategies to overcome the constraints imposed by local-scale stressor, many of them reliant on their dispersal ability (Townsend & Hildrew, 1994; Kärnä *et al.*, 2015; Burdon *et al.*, 2016; Morton *et al.*, 2018).

The dispersal ability is a key life-history trait modulating the effect of selection on species occurrence, and consequently, on the structure of riverine metacommunities (Tonkin *et al.*, 2018). Hence, by including dispersal ability or mode in statistical models, researchers have been able to unravel the major drivers of metacommunity structure and dynamics at larger scales (i.e. Cañedo-Argüelles *et al.*, 2015a; Sarremejane *et al.*, 2017). For example, community structure is expected to reflect the environmental conditions if species dispersal rates are moderate, because species can track the most suitable conditions (Holyoak *et al.*, 2005). Meanwhile, if species dispersal rates are low or very high, the structure of local communities may be partly decoupled from that expected from the effect of local environmental conditions alone (Kärnä *et al.*, 2015; González-Trujillo *et al.*, 2019).

Box 3 The Metacommunity Framework

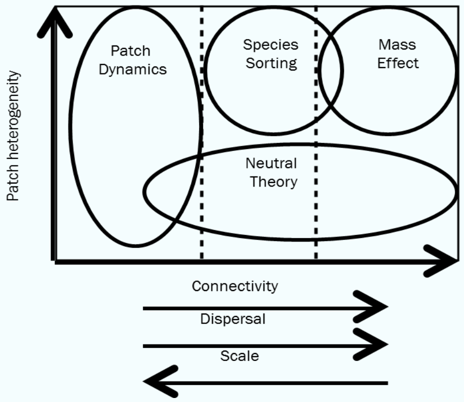
- **Metacommunity (definition):** a set of **local communities** connected through the ***dispersal*** of multiple interacting species (Leibold *et al.*, 2004).

The central tenet of the **metacommunity framework** is that the diversity of biological communities cannot be understood by focusing exclusively in local-scale processes (e.g. competition). This is because regional-scale processes, such as the dispersal of organisms among local communities, can influence the structure and dynamics of biological communities.

The **Metacommunity framework** was firstly stated as a set of **four major hypotheses/paradigms** (Leibold *et al.*, 2004; Holyoak *et al.*, 2005):

- **Species sorting:** analogous to the niche paradigm, it emphasizes in the sorting of species into the most appropriate habitat based on biotic and abiotic filters.
- **Mass effects:** under this paradigm, high rates of dispersal can swamp local effects and allow species to persist in unfavorable habitats.
- **Patch dynamics:** this paradigm states that landscapes are mosaics of patches that can be occupied by every species of the regional pool
- **Neutral assembly:** under this paradigm, species in communities are considered ecologically equivalent. Thus, the dynamics and structure of local communities are dictated by the random loss and gain of individuals.

However, mounting evidence has demonstrated that these hypotheses are not mutually exclusive, and therefore, are not statistically testable (see Leibold and Chase 2018 and references therein). In this regard, Leibold and Chase (2018) claimed that researchers should better focus on the causes and consequences of these paradigms rather than on describing which one fixed better to empirical patterns. As major causes, they proposed the **patch heterogeneity**, **connectivity**, **dispersal** and **scale**.



The interaction of these processes determines the metacommunity dynamics and shape the structures previously described in the literature (Modified from Leibold and Chase 2018).

However, not all freshwater species can disperse in the three spatial dimensions which can be recognized in fluvial ecosystems (e.g. longitudinal, lateral and vertical, Kärnä *et al.*, 2015). Recent research brought to light that species dispersal ability is constrained to a greater degree by the structure of the riverine landscapes; and particularly, by the dendritic-like structure of fluvial networks (Brown & Swan, 2010; Altermatt *et al.*, 2013). In this regard, mounting evidence has demonstrated that the branching organization of fluvial networks can exert strong controls on metacommunity dynamics – and ultimately shaping biodiversity patterns, through regulating the extent and rates of ***dispersal*** (Altermatt, 2013; Datry *et al.*, 2016; Finn *et al.*, 2016; Altermatt & Fronhofer, 2018). Specially, in dendritic networks and for obligate aquatic organisms, such as fishes (Kärnä *et al.*, 2015; Tonkin *et al.*, 2018), or in fluvial network possessed of intermittent stems (Cañedo-Arguelles *et al.,* 2015).

So forth, recent research on riverine metacommunities has particularly highlighted the preponderant role that ***dispersal*** can play in the assembly process. However, few studies have yet considered that dispersal occurs in **ecological** and in **evolutionary timeframes** (but see Bonada *et al.*, 2005). When dispersal constraints occur within an **ecological timeframe**, the distributional patterns may reflect both the spatially structured local environment and the structure of dispersal pathways. Conversely, when strong and longstanding constraints take place, species dispersal may be impeded within an **evolutionary timeframe**, which could create distributional patterns that are footprints of the **historical contingencies** of the region under study (Brown, 1995).

The influence of **evolutionary events** on the species diversity and distribution has been usually studied by Biogeography. However, since the contemporary patterns of community structure can be shaped by the effect of past climatic and geological events (Vyverman *et al.*, 2007b; Dias *et al.*, 2014), a conceptual unification between biogeography and community ecology would provide a better understanding of the causes and consequences of biodiversity across scales (Jenkins & Ricklefs, 2011; Ricklefs & Jenkins, 2011).

In this context, the **metacommunity framework** represents an ideal model to study the intersection between both disciplines. Previous research has demonstrated that including proxies of the biogeographical effects, such as biogeographical provinces or basin identity, can increase the explanatory power of metacommunity-based analyses (Declerck *et al.*, 2011; Heino *et al.*, 2017). Thus, applying a ***biogeography-oriented metacommunity framewor*k** may be of interest to understand the interplay between ecological and evolutionary processes in shaping the contemporary patterns of biodiversity distribution; particularly, in the Neotropical region.

In the Neotropics, anthropogenic pressures are still relatively small (though widening; Sabater *et al.*, 2017). Thus, the effect of anthropogenic stressors is not confounding diversity patterns, as it can occur elsewhere (i.e. Gutiérrez-Cánovas *et al.*, 2013). Moreover, past geological and climatic events have left imprints of the contemporary distribution of Neotropical species in a wide number of ecosystems and landscape: being the Andean uplift events strong drivers of the current distribution of ecoregions (Hoorn *et al.*, 2010; Hazzi *et al.*, 2018). This diverse ecoregional mosaic is drained by several stream networks that are interconnected within large basins, and therefore, local communities that had once been isolated by past events may currently be re-connected through dispersal, leading to the onset of metacommunity dynamics within an ecological timeframe. Yet, the interplay between ecological and evolutionary factors in shaping the assembly process remains to be explored.

Thesis aims and outline

This thesis uses *a* ***biogeography-oriented metacommunity framework*** to dissect the relative influences of ***dispersal*** (in ecological and evolutionary timeframes), ***selection*** (driven by abiotic factors) and **ecological drift** on the assembly process of benthic communities. This means that the relative influence of contemporary constraints (environmental- and dispersal-related) and the possible legacies of historical events on the assembly process of benthic communities is evaluated by merging the toolbox of the metacommunity ecology (regression models on distance matrices, unsupervised classification analyses and constrained and unconstrained ordination methods) into a biogeographical framework (see next chapter for further details).

In this context, the aims of this thesis are:

(i) Evaluate the effect of ecoregional features on the different facets of diversity of algae and invertebrates at different spatial scales.

(ii) Determine the main drivers of the structure of algal and invertebrates metacommunities.

(iii) Establish the degree of convergence between the biological responses of algae and invertebrates to environmental factors that act in each type of ecoregion.

(iv) Analyze the influence of dispersal barriers on the different forms of diversity of algae and invertebrates at different spatial scales.

(v) Distinguish the roles of historical and contemporary processes on the benthic communities of diatoms and insects in tropical montane streams.

To distinguish the role of ***dispersal*** (and **selection** in some extent) in historical and contemporary timeframes, I selected two communities with disparate life histories in terms of dispersal mode and ability: stream ***diatoms*** and ***invertebrates***. ***Diatoms***, on one hand, are sessile microorganisms whose small body size is usually related to a high dispersal capability (Padial *et al.*, 2014). ***Invertebrates***, on the other hand, are organisms with different dispersal capabilities. Their dispersal mode and ability are mainly linked to their life histories and evolutionary background (Shurin *et al.*, 2009; Saito *et al.*, 2015). Therefore, while diatom communities may mainly respond to the contemporary environmental conditions rather than dispersal limitation factors (Passy, 2012), invertebrate communities may respond to both environmental and dispersal constrains.

Additionally, to account for the effect of historical imprints on contemporary patterns, I created an ecoregional classification scheme which allows to reconstruct a hypothetical scenario for the geological history of the ecoregions in the basin (see chapter I).

Outline

The thesis is structured in seven chapters. In the **Chapter I,** I outline a primer on the study zone, encompassing the evolutionary history of the Orinoco basin and the legacy of landscape features from geological and climatic events. These events have generated a wide gradient of biogeographical provinces and environmental conditions in the Orinoco basin, which I included on disentangling the main drivers of the community assembly. Specifically, an ecoregional classification system has been developed to include the effect of historical legacies in the analyses.

After a brief description of the Material and Methods in the **Chapter II,** the subsequent chapters used this ecoregional classification as a basis for exploring the role of evolutionary events on the present-day patterns of metacommunity structure. If evolutionary history influences the structure of freshwater communities in parallel with ecoregions' physiography and geomorphological structure, it is expected that a nearly discrete species pool would characterize each ecoregion. This was tested in the **Chapter III** by characterizing the distribution of invertebrate species pools in the Orinoco basin, and testing how their distribution affects the structure of the metacommunity across the stream network of the basin. As much as it was expected, the historical and evolutionary processes have resulted in species pools differing between the ecoregions of the Orinoco basin. This finding highlights the relevance of historical (= dispersal limitation in evolutionary timeframes) in shaping the metacommunity structure at the basin scale.

The role of historical events shaping the shaping microorganism biogeographical distribution patterns is still strongly debated (Fenchel & Finlay, 2005). Their small size and passive dispersal strategies brought forth the hypothesis that they could colonize anywhere, overcoming the possible contingencies of historical events on their distribution (Martiny *et al.*, 2006). In the **Chapter IV**, the role of historical events on microorganism distribution was tested using diatom communities. It was found that species composition was mainly explained by the ecoregion identity, followed by environmental conditions and spatial structure. Indeed, diatom distribution was highly constrained to the ecoregions, each of which harbored a distinctive pool of diatom taxa.

Chapters II and **III** provide evidence on the strength of historical contingencies, and its effect on contemporary environmental conditions, in affecting the assembly process of benthic communities in the Neotropics. All evidences up to now point that the historical events have mainly constrained dispersal in an evolutionary timeframe. However, dispersal could be constrained in an ecological timeframe by topographical barriers (e.g. mountains, Finn *et al.*, 2016), by the connectedness degree within the dendritic network (Sarremejane *et al.*, 2017), or by the effect of some anthropic stressors (Gutiérrez-Cánovas *et al.*, 2013). Thus, in the **Chapter V**, a set of beta-diversity metrics is used to assess the effect of deforestation on the distribution patterns of benthic communities, and to determine how it is modulated by the occurrence of upstream refugia and the dispersal abilities of the different communities.

The final question to be answered in this thesis is whether or not, patterns of micro- and macroorganisms are shaped by the same drivers. In the **Chapter VI,** this question is addressed by comparing patterns of beta diversity. Beta diversity provides fundamental insights into mechanisms of community assembly (Anderson *et al.*, 2011). Thus, I characterized and compared beta diversity patterns of benthic communites among streams of the Orinoco basin. Since biodiversiy is multi-faceted, I estimated it by using the taxonomic-, functional- and phylogenetic-based approaches. A series of models including environmental, spatial and historical descriptors were fit to determine the major drivers of beta diversity and its components (turnover and nestedness).

Finally, the **Chapter VII** includes the concluding remarks and some perspectives for future studies in the Neotropical region.

I. The Orinoco basin: a geographical and physiographical primer

An evolutionary overview of the Orinoco basin

Biodiversity patterns emerging in riverine ecosystems cannot be addressed without considering the geological and climatic history of their riverscapes (*sensu* Wiens, 2002). This is because present-day features of riverine landscapes are imprints of past evolutionary events, slightly to severely modified by anthropogenic impacts (Dollar *et al.*, 2007; Thorp *et al.*, 2010). In pristine or near-pristine conditions, river's physical structure and behavior overtime are determined by the top-down effect of geology and climate on catchment- and channel-level modulators such as vegetation or sediment load, respectively (Schumm, 1977; Frissell *et al.*, 1986) (Figure I-1).

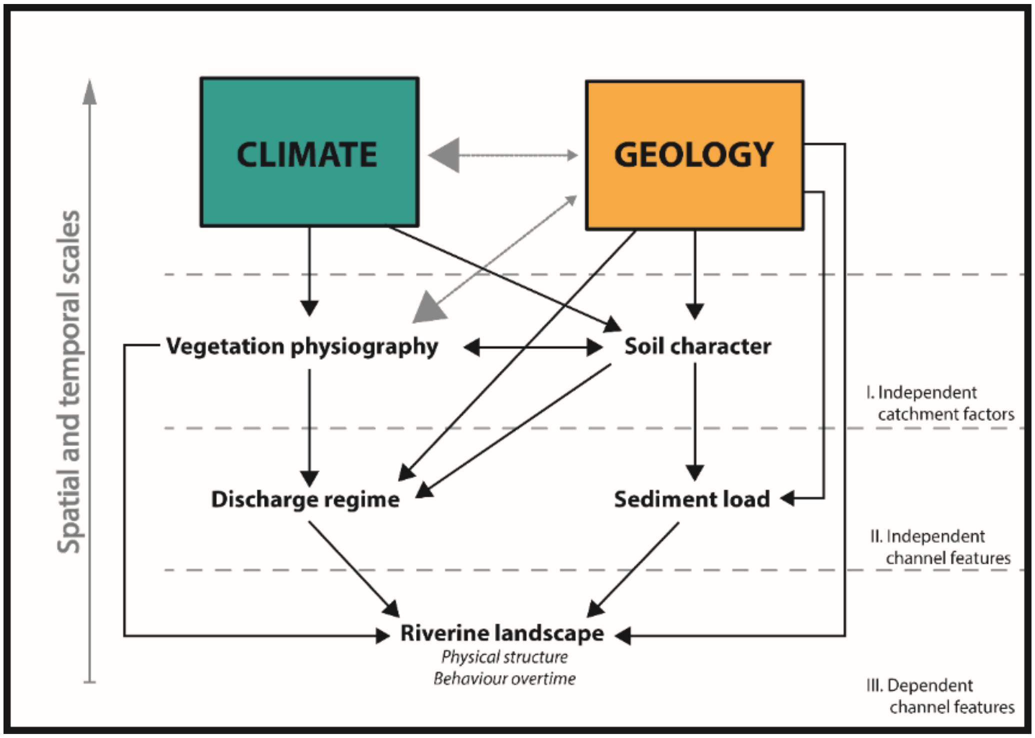


Figure I-1 Major factors and variables shaping the structure of riverine landscapes. After Schumm 1977.

This is the case of the **Orinoco basin**: the third largest basin in South America, covering about 990,000 km^2 in Venezuela and eastern part of Colombia (Romero Ruíz *et al.*, 2004) (Figure 0.2). In this basin, the physiographic features - and contemporary distribution of species – cannot be explained without considering past geological and climatic events (Goosen, 1971; van der Hammen, 1974; Flórez, 2003). Particularly, those historical legacies from events that occurred during the tertiary and quaternary periods.

Figure I-2 Geographical location and extent of the Orinoco basin. The green shaded area represents the study zone.

Historical legacies on species distribution can be appreciated in all the basins influenced in some degree by the Andean Orogeny. At least five great uplifts along the Andean chains affected the regional climate of South America and reconfigured the continent's drainage patterns, giving rise to distinct fluvial systems as well as to distinct aquatic species (e.g. Goosen, 1971; van der Hammen, 1974; Hughes & Eastwood, 2006; Hoorn et al., 2010). In the Amazon basin, for instance, Andean uplifts conformed a big lake during a long period of time ('the Pebas system'), causing Amazon streams to change their flow directionality (from W-E to E-W) and connectivity patterns within and between sub-basins (Albert et al. 2018). These events, together with changes in climate, have been later related to dispersal constraints and species radiation in reptile and fish communities (Hoorn *et al.*, 2010; Albert *et al.*, 2018). Although similar links have not been established in the Orinoco yet, the parallel evolutionary stories (Bicudo *et al.*, 2019) suggest a similar influence of historical contingencies on contemporary patterns.

The major events which shaped the contemporary landscapes in the Amazon and Orinoco basins occurred during the **Tertiary** and **Quaternary** periods (Bicudo *et al.*, 2019). During the **Tertiary**, the Andean orogeny caused uplift of the source areas which molded in a great extent the basin geographical and physiographical features (Goosen, 1971; Romero Ruíz *et al.*, 2004) and determined the changes in climatic conditions (e.g. temperature) as well as in the vegetation structure overtime (van der Hammen, 1958). On the other hand, during the **Quaternary**, the Pleistocene glaciers and inter-glaciers changed repeatedly the face of the riverine landscapes (Goosen, 1971; van der Hammen, 1974; Flórez, 2003).

The legacies from **Tertiary** and **Quaternary** changes are currently observable in stream's geomorphology and riverine vegetation physiography as well as in the climatic conditions. Among the basin, for instance, it is possible to observe a wide diversity of plant assemblages and geomorphological patterns along a longitudinal gradient spanning through meandered streams, braided rivers, and constrained and armored channels (see Figures I-3 to I-9). Although there is a wide number of vegetation types (at least 22 according to Rangel-ch, 2016), three major vegetation formations can be distinguished in the basin: the **montane**, **tropical (rain) forest** and **tropical savanna** formations (van der Hammen, 1974), each one differing in evolutionary origin and history. The **savanna** formation is restricted to the center of the basin, where the confluence of winds creates a semi-humid climate (<2000mm of rainfall per year) with some water deficiency during the dry season. Surrounding the savanna, there is the **tropical (rain) forest** – an Amazon-like forest with an extremely-humid climate (about 4000mm of rainfall per year). Finally, increasing in altitude from the piedmonts to the mountain peaks, there is the **montane vegetation formation**.

The **montane formations** host a large environmental heterogeneity throughout an elevational gradient. Indeed, montane systems have been objects of fascination for earlier naturalists such as von Humboldt and Bonpland (1807). From the bottom to top, mountains host a series of nearly discrete belts of vegetation changing altitudinally in a stepwise manner. Each belt is a footprint of past evolutionary events (van der Hammen, 1957, 1974) and is characterized by distinctive geomorphological and physiographical features. Because of their discreteness in eco-evolutionary features, altitudinal belts have served as the foundation for a growing body of knowledge documenting the mechanisms associated to the high diversity in montane systems (e.g. Nottingham *et al.*, 2018; Fattorini, Di Biase & Chiarucci, 2019; Encalada *et al.*, 2019; Muellner-Riehl *et al.*, 2019).

Among these altitudinal and longitudinal gradients, there is also an intricate network of rivers and tributaries which drains down three large forms of relief that grow in altitude from 100 to 3500m a.s.l.: the Orinoco's **ancient massifs and shields**, its recently **raised ridges**, and its **tectonic depressions** and **sedimentary plains** (Stallard, 1985). These reliefs were molded by the Andean orogenesis and, they are one of the major drivers of the in-stream habitat heterogeneity and water chemical characteristics (Stallard, 1985). For example, the **ancient massifs** have passed through at least three weathering cycles (Goosen 1970). Thus, rivers draining this relief have metamorphic rocks in their riverbeds and lower values of electric conductivity, and in general, lower concentrations of ions (see for instance the Guiana Shield streams of this study). On the other hand, the **raised ridges** shaped strong lateral influences which drive increases in salinity and chloride at the lowlands of the basin (Gonzalez-Penagos *et al.*, 2014). However, these increases are not uniform, since **sedimentary plains** are devoid of components associated with the meteoric water signature (Gonzalez-Penagos *et al.*, 2014).

The reconstruction of the Orinoco's geological history by means of path-length matrix

Rationale

As earlier discussed, geological and climatic events have shaped unique combinations of river forms and riparian ecosystems at the ecoregional scale (Hoorn et al., 2010; Rull, 2008); being ecoregions, therefore, contemporary imprints of those past historical events. To account for the historical relationships among ecoregions, we built a path-length matrix (Legendre *et al.*, 1994). This matrix represents a hypothetical evolutionary scenario of the Orinoco basin, since it models the order of appearance of each ecoregion in the basin and the most plausible relationships among ecoregions regarding their origin.

The matrix was constructed based on available geological, biological, historical and palynological evidences (van der Hammen, 1958, 1974; Goosen, 1971; Stallard, 1985; Flórez, 2003; Hubert & Renno, 2006; Hoorn *et al.*, 2010; Jaramillo *et al.*, 2010; Rangel-ch, 2016; Flantua & Hooghiemstra, 2018; Bicudo *et al.*, 2019; Flantua *et al.*, 2019). The process to build this path-length matrix involved three major steps:

(i) **Create an ecoregional classification scheme to classify the rivers.** This scheme is grounded on the type of ecoregion and the type of Functional Process Zone (FPZ, *sensu* Thorp et al. 2008), as both can be considered as contemporary imprints of past geological and climatic events.

(ii) **Reconstruct a hypothetical scenario of basin's evolutionary history by sorting them based on their time of appearance in the basin.** We defined "time of appearance" as the age at which appeared the most similar to contemporary vegetation and land surface forms. To distinguish between two ecoregions with similar time of appearance but different origin, we also included in this reconstruction the major geological and/or climatic events which triggered changes in their landforms or vegetation forms.

(iii) **Translate the hypothetical scenario to a cladogram, which is then converted into a path-length matrix.** We used a rooted tree where the first step represents the distinction between the Guiana shield and the other ecoregions because of the last Andean uplift.

Next, we describe each step in detail:

i. Create an ecoregional classification scheme to classify the rivers.

River characterization schemes have been traditionally pursued as a scientific way to understand their functioning as well as a basis to guide their management (Kondolf & Piégay, 2005). Many schemes have been proposed among the years (see Thorp *et al.*, 2010 and references therin); however, each one is closely linked to the study aim, and usually, they are defined based on the local environmental features (Thorp *et al.*, 2010). For this study, we followed a large-scale **ecoregional classification scheme**, which is grounded on the type of ecoregion and the type of Functional Process Zone (FPZ, *sensu* Thorp et al. 2008), as both can be considered as contemporary imprints of past geological and climatic events.

The starting point was to define the **Functional Process Zones** as the elementary unit of the scheme (‘**FPZ**’; *sensu* Thorp, Thoms & Delong, 2006). **Functional Process Zones** (Thorp *et al.*, 2010), or Functional sectors (Petts & Amoros, 1996), represent hierarchical units between a valley and a reach. These units or patches are based on a hydrogeomorphic perspective with function being related to hydrologic and geomorphic processes. These explicitly incorporate the effects that geology and climate have on river’s physiography and geomorphology overtime, but do not explicitly consider other aspects of the evolutionary history which can affect species distribution, such as climatic fluctuations or barriers to dispersal. Thus, if the same type of **FPZ** occurred in two distinct biogeographical provinces, the explanatory power of the unit will be lower since the two **FPZ** can host different pool of species.

In light of this, I included the biogeographical history of the basin into the **FPZ framework** by considering the Ecoregion type where each **FPZ** occurs. Ecoregions were delimited based on vegetation type, climate, altitude and age of major landform (Table I-1). Similar classification schemes (i.e. altitudinal belts of vegetation) have served as the foundation for a growing body of knowledge documenting the mechanisms associated to the high diversity in Tropical systems (e.g. Nottingham *et al.*, 2018; Fattorini, Di Biase & Chiarucci, 2019; Encalada *et al.*, 2019; Muellner-Riehl *et al.*, 2019). Moreover, I demonstrated that each ecoregion host a nearly unique pool of diatom (Chapter IV) and invertebrate species (Chapter III), indicating that riverine species distribution may be a legacy of past historical events.

Table I-1 Physiographic features of the seven ecoregions assessed within the Orinoco basin. The included features are recognizable at large spatial scale. The number of streams sampled in each ecoregion is reported within parenthesis.

	Páramo (n=3)	**high-Andean (n=5)**	**Piedmont (n=5)**	**Alluvial fans (n=4)**	**Alluvial plains (n=3)**	**High Plains (n=3)**	**Guiana shield (n=3)**	***References***
Age of major landform	early Pleistocene to Holocene	early Pleistocene to Holocene	Pliocene to early Holocene	Pliocene to Holocene	Pliocene to Holocene	early Holocene	Miocene	(van der Hammen, 1958, 1974; Goosen, 1971; Restrepo & Toussaint, 1988; Flórez, 2003)
Altitude (m asl)	>3.000	2000-3000	300-1000	<300	<300	<400	<500	(van der Hammen, 1974)
Vegetation type	Páramo and subPáramo	Andean forest	Lower tropical forest	Lower tropical forest	Gallery forest	Savanna and Gallery forest	Dry tropical woodland	(van der Hammen, 1974)
Climate	Tundra (ET)	Oceanic (Cfb)	Rainforest (Af)	Rainforest (Af)	Rainforest (Af)	Monsoon (Am)	Monsoon (Am)	Köppen-Geiger classification (Kottek *et al.*, 2006)
Channel form	gorge	mountain	mountain	Braided	meandering	meandering	meandering	(*sensu* Petts & Amoros, 1996)

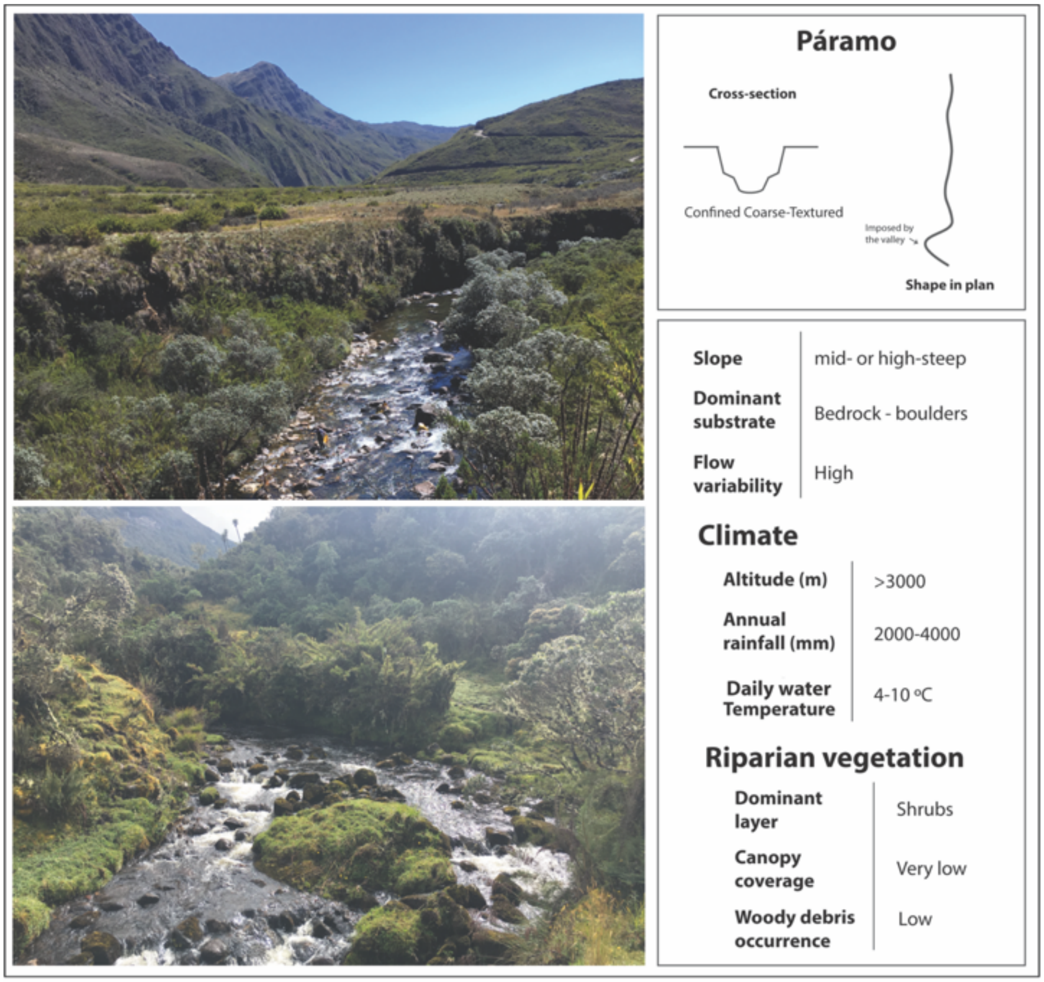
The ecoregional classification scheme

The ecoregional classification scheme is hierarchical (mesohabitats within riffles, riffles within stream segments, and segments within ecoregions) and consists of seven major categories (‘Eco-FPZs’):

1. The Páramo; endemic to the Neotropics, it is the most recent ecoregion in the Andean hills. It is globally recognized as a diversity hotspot as well as a “natural water fabric” (Buytaert *et al.*, 2006).
2. The high-Andean ecoregion; located immediately below the Páramo, is a belt of cloud forest abundant in fertile zones, with comparatively high precipitation and low temperatures.
3. The Piedmont ecoregion; reaching down the Andean foothills, it encompasses a belt of rainforest growing on steep slope zones with a warm and humid climate.
4. The Alluvial fans ecoregion; where the streams leave the Andean hills and join the greater mainstreams, such as the Meta or Guaviare rivers.

5, 6, **7**. The remaining three ecoregions are distributed throughout the lowlands (locally known as “Llanos Orientales”). During the early Holocene, the Meta fall divided the lowlands into an elevated zone to the east, which is characterized by a savanna-like climate and vegetation (the high Plains ecoregion), and a lower zone to the west that includes the alluvial plains and the Guiana shield ecoregions. Both ecoregions are elevated zones with a comparatively flat morphology and distinct vegetation. They were both shaped during the Andean uplifts, although the former is more ancient, and has a geology that is largely influenced by the Guiana shield reliefs.

A detailed description of each ecoregion is provided below (in order of decreasing altitude):



Slope	mid- or high-steep
Dominant substrate	Bedrock - boulders
Flow variability	High

Climate

Altitude (m)	>3000
Annual rainfall (mm)	2000-4000
Daily water Temperature	4-10 ºC

Riparian vegetation

Dominant layer	Shrubs
Canopy coverage	Very low
Woody debris occurrence	Low

Figure I-3 The Páramo ecoregion. It has its structural origin during the **Quaternary** (van der Hammen, 1974; Flantua & Hooghiemstra, 2018). During this period, glacial cover the Andes cordilleras from 3000m above sea level and their slow movement was flattening the terrain until deglaciation period (Flantua *et al.*, 2019). Since the last glaciation event, streams have been running throughout glacial-origin depressions, dissecting the terrain, and forming terraces along the flat terrain above (Flórez, 2003). As a result, these rivers have the form of Confined Coarse-Textured zones with deep and near-vertical valleys constraining their horizontal displacement. The moderate-to-high slopes give an idea of high energy rivers with high erosion potential. Stream segments are composed by several pool-riffle sequences. Bedrock and boulders are bed substrata in riffles, and pools are composed mainly by pebbles and gravel. The vegetation is quite characteristic; shrubs and herbaceous strata are the dominant forms. These forms do not represent neither a great coverage from the main channel nor an input of allochthonous resources.

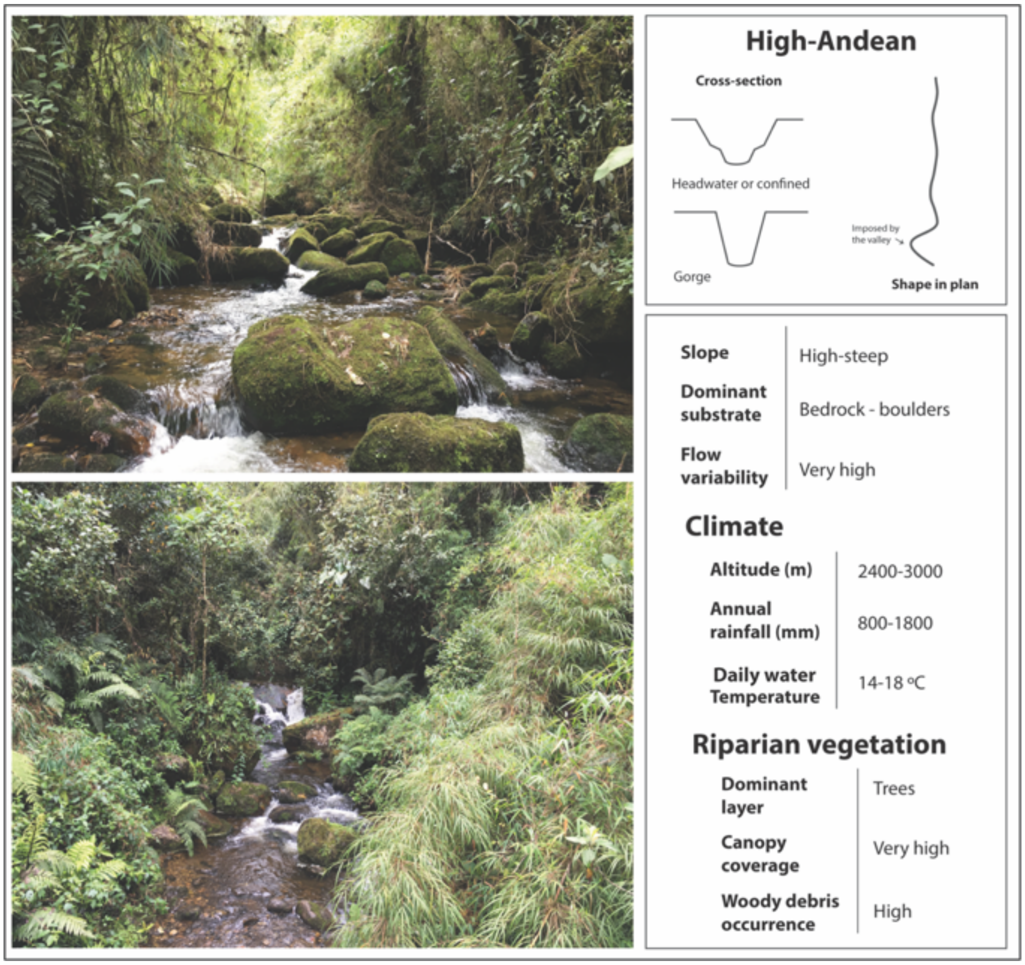


Figure I-4 The high-Andean ecoregion. The origin of these streams is closely linked to the Andean orogeny, and more recently, to deglaciation events during the **Quaternary** (van der Hammen, 1974; Flantua & Hooghiemstra, 2018). Located below glacial boundary (between 2100-3000m), high-Andean rivers preserve the dissected valleys and steep slopes originated during Andean orogeny. In some rivers, these characteristics are more accentuated by high water flow discharges linked to deglaciation periods. Thus, water and sediment regimes and cross-sections are similar to those of Gorge, Confined Coarse-Textured and Headwater FPZ's. Due to the action of high energy and unpredictable spate events, reaches have steep-pool and riffle-pool sequences with in-stream beds composed mainly by bedrock and boulders. Nevertheless, the constant erosion of margins also incorporates smaller substrata, such as cobbles or pebbles, which can be found in pools. A dense forest, with leafy trees, cover the major part of the main channel and provides an input of litter and organic matter (OM) to the system. The high retention of this input could determine the bed stability and several functional aspects of high-Andean rivers (Petts & Amoros,1996).

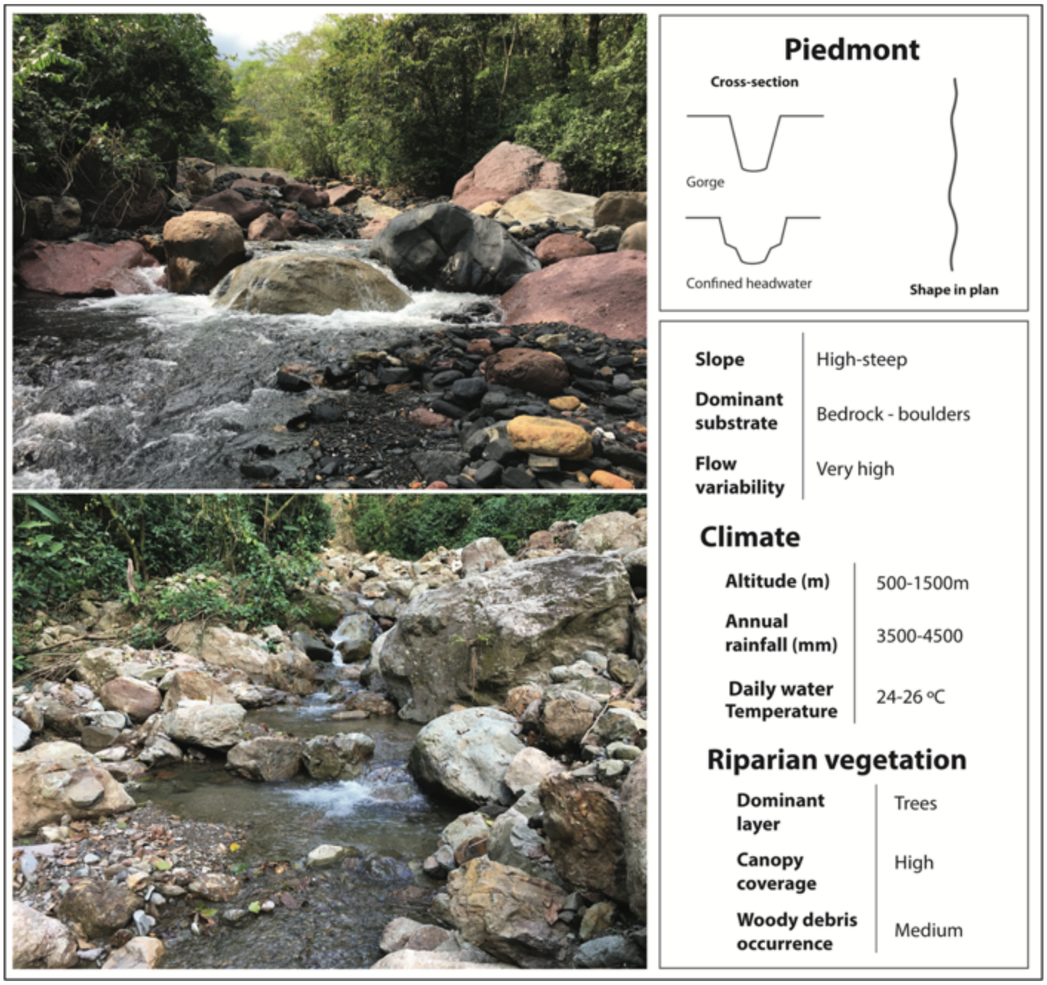
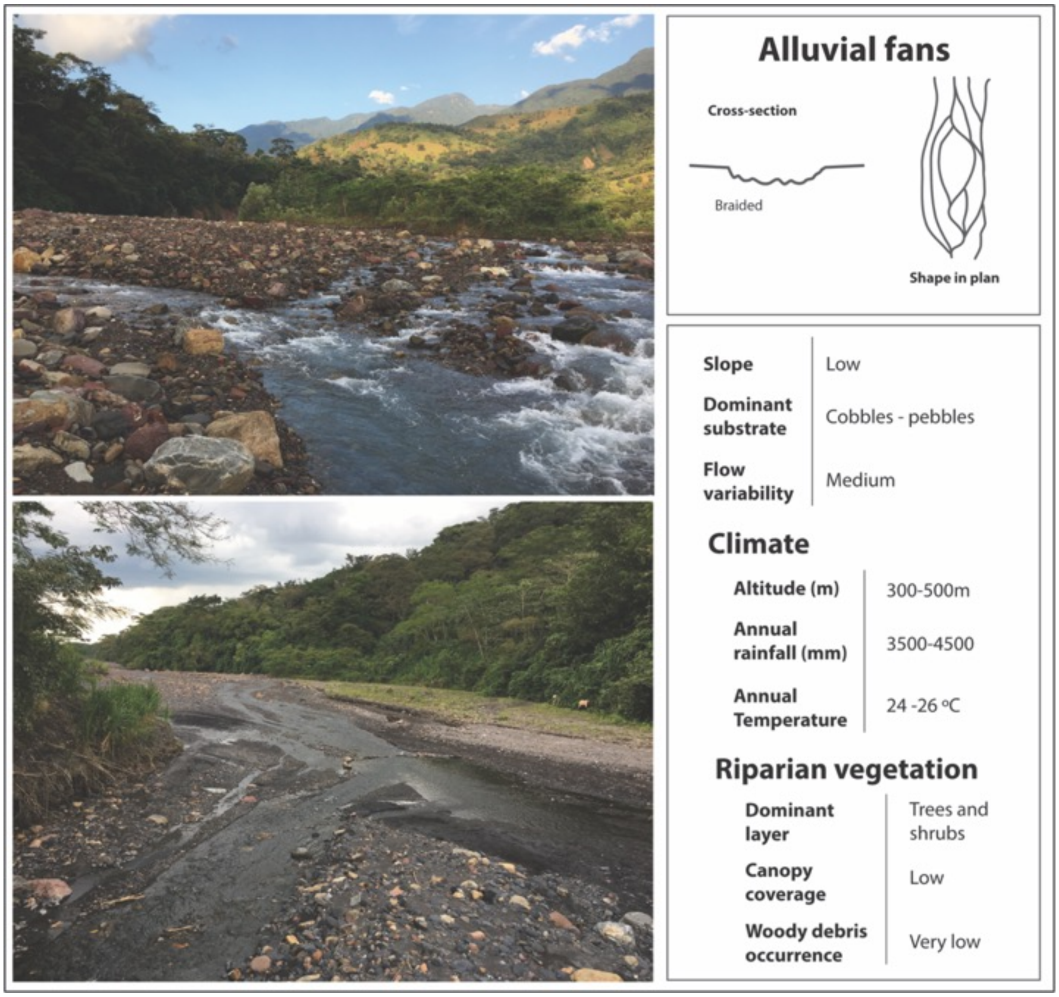


Figure I-5 The Piedmont ecoregion. The Andean uplifts caused the folding of the Piedmont over the flat terrain between elevations from 400 to 1000m. This event created a concave terrain of low-stability over which was formed several draining systems along the Andes (Goosen, 1971). Low-Andean river have features like that of transfer zones; e.g. highly dissected and confined cross-sections, and a nearly straight shape in plan (Schumm, 1977). As a result, Gorge and Headwater are the main FPZ's of these rivers. As with high-Andean rivers, reaches of these rivers are characterized by step-pool sequences with coarser substrates and leafy trees covering the channel. Nevertheless, they have higher temperatures among the year because their lower elevation. Moreover, their steeper slopes of channels and greater valley instability caused a constant removal of accumulations of wood and smaller substrates. In this regard, woody debris seems to be less important for bed stability and system functioning in low-Andean zones.



Slope	Low
Dominant substrate	Cobbles - pebbles
Flow variability	Medium

Climate

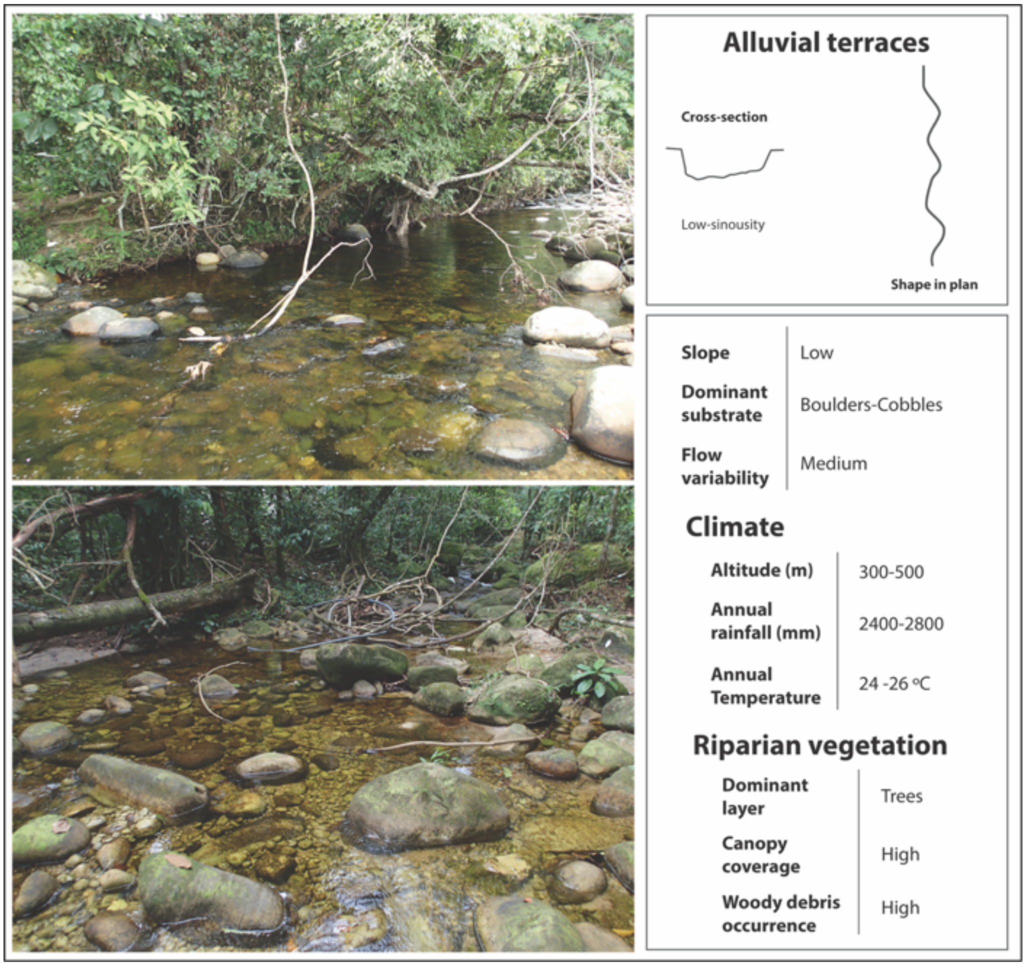
Altitude (m)	300-500m
Annual rainfall (mm)	3500-4500
Annual Temperature	24 -26 ºC

Riparian vegetation

Dominant layer	Trees and shrubs
Canopy coverage	Low
Woody debris occurrence	Very low

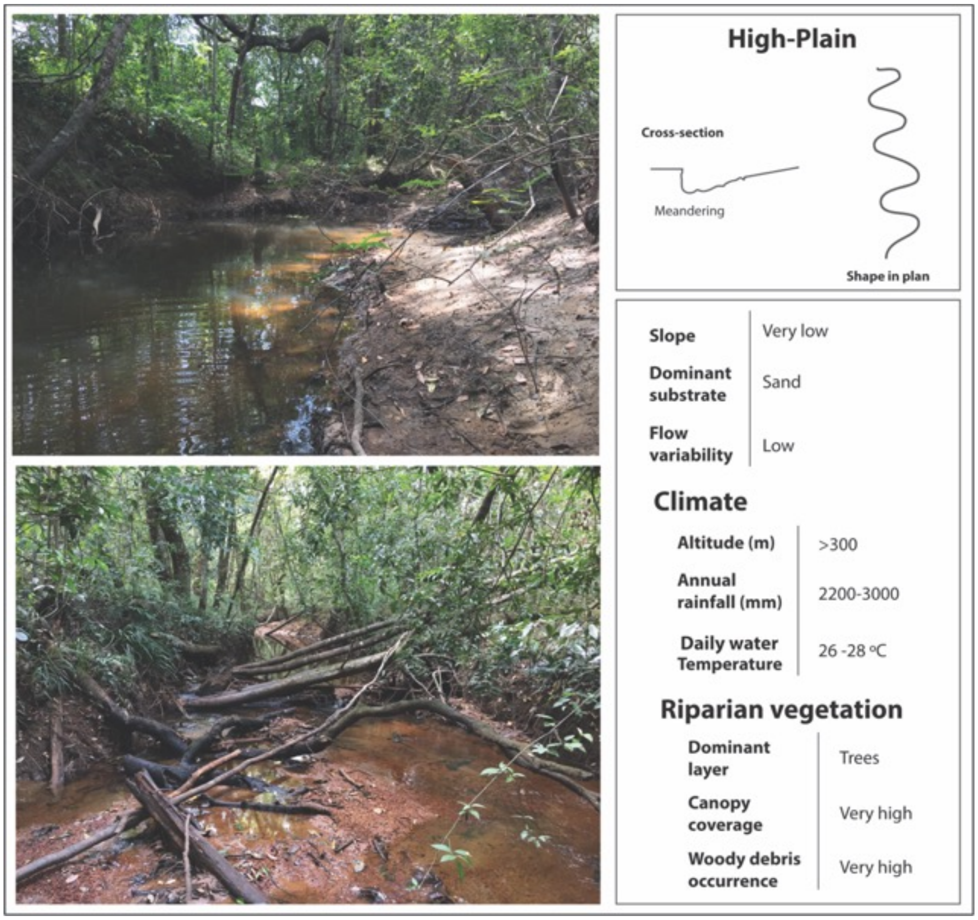
Figure I-6 The Alluvial fans ecoregion. Streams reach a plain terrain immediately after they out from the constriction characteristic of Piedmont ecoregions. The unconfined valley together with the abruptly change of slope caused streams take a braided or ‘alluvial-fan’ shape. Most of these streams have this shape until they meet the Meta river. Coarser substrates are retained in the low-Andean zones; the smaller ones form the alluvium surface over which these streams are running. These streams have frequent variations in stream discharge that promote a constant rearrangement of river channels. Channels are composed of bars of different forms and substrate types; however, cobbles and pebbles are the most abundant substrates overall. The riparian vegetation has a great diversity of species and strata; however, the taller trees only cover a small percentage of channel margins. Moreover, the high-magnitude discharges and the instability of riverbed do not allow a high retention of OM or litter.

The Andean orogeny was followed by periods of high sedimentation. Sediments of the upper cordillera were transported to the low-elevation zones by the action of rivers (Flórez, 2002). The valley physiography of next river categories had its origin after these sedimentation phases filled up the internal sea of the Orinoco region. Once the sea water was depleted, events such as the sediment discharge of Andean rivers, the uplift of the Meta fault, and several wind currents, shaped the river forms that we can see at the present day.



Slope	Low
Dominant substrate	Boulders-Cobbles
Flow variability	Medium
Climate	
Altitude (m)	300-500
Annual rainfall (mm)	2400-2800
Annual Temperature	24 -26 ºC
Riparian vegetation	
Dominant layer	Trees
Canopy coverage	High
Woody debris occurrence	High

Figure I-7 The Alluvial plains ecoregion. A large number of streams run throughout the alluvial terraces formed during the sedimentation periods of the Orinoco basin. Some of these, western to the Meta river uplift, are straight and low-energy rivers. They run throughout smaller faults with moderate-to-shallow sloped; therefore, the main course is partially confined. The low sinuosity is the most representative feature of these streams. Their low energy is linked to cross-sections with a uniform shape and low rates of bed degradation. Nevertheless, it is common to find the typical riffle-pool sequences created mainly by cobble bars. Pebbles and gravels are the most frequent type of substratum in channel beds; however, boulder could be abundant in streams near the uplift of “proto-cordilleras”. A dense riparian vegetation with taller trees is characteristic and, it covers the channel bed in a wider degree. As a result, several “litter packages” could be found along the channel. Some woody debris formations could be also observed frequently along the channel.



Slope	Very low
Dominant substrate	Sand
Flow variability	Low
Climate	
Altitude (m)	>300
Annual rainfall (mm)	2200-3000
Daily water Temperature	26 -28 ºC
Riparian vegetation	
Dominant layer	Trees
Canopy coverage	Very high
Woody debris occurrence	Very high

Figure I-8 The high Plains ecoregion. The movement of the Meta fault caused an uplift of the platform located south-eastern from the Meta river. The change in altitude (approx. 50m; Goosen 1970) blocked the discharge of water and sediments of rivers coming from the mountains. This event followed by a desertification phase characterized by strong winds during the **Pleniglacial** that caused a general loss of superficial water and forest (van der Hammen, 1974). Only rivers running throughout low-to-moderate channel dissections surrounded by gallery forest can be observed today. These rivers have shallow slopes, low flow velocities, and therefore, low energy. Features also contributing to their meandering shape in plan. The bed of high plain streams is composed mainly by the sand carried out during the **Pleniglacial** and extensive layers of allochthonous litter. Some riffles could be found in zones of woody debris, but low-velocity or pool zones are the most common along the river. The base-level of these rivers is closely linked to the gallery forest that surround them. Gallery forest provides canopy coverage and a great input of litter. Moreover, the forest provides food and refugium for invertebrates and fishes when the channel is laterally expanded in wet seasons.

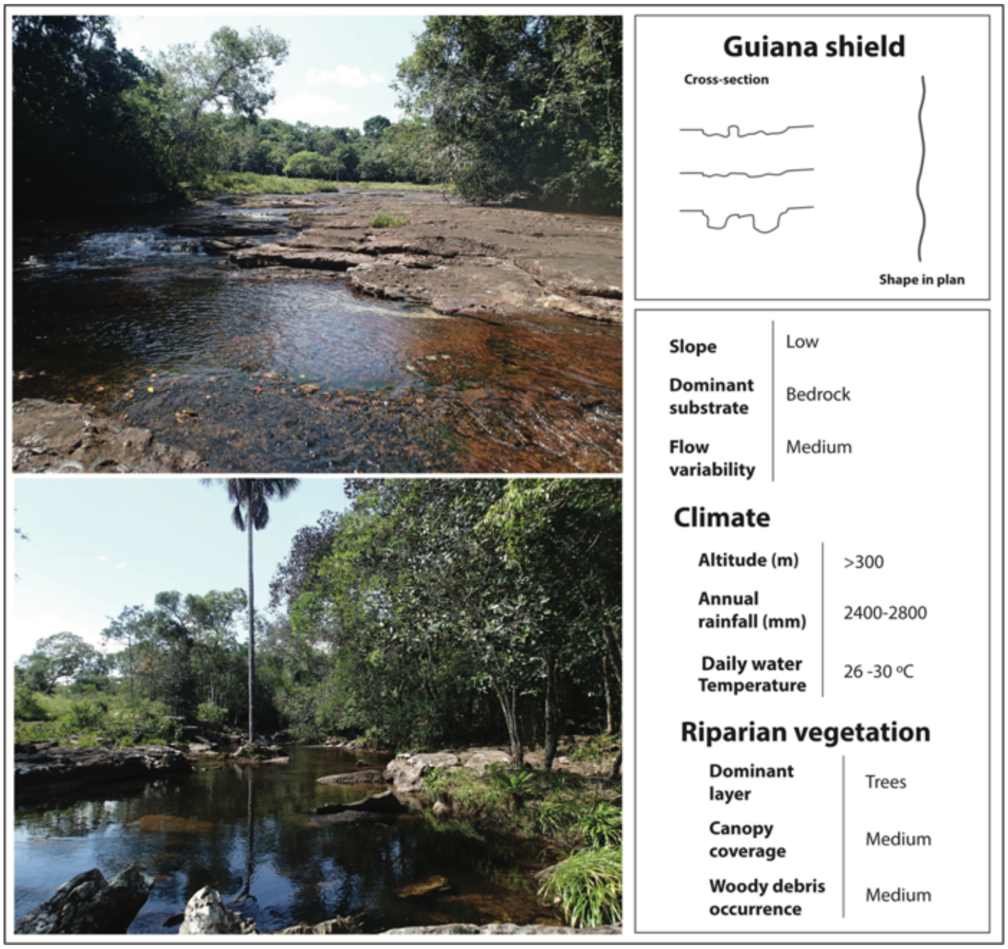


Figure I-9 The Guiana shield ecoregion. Some events related to the Andean uplift caused the emergence of "proto-cordilleras" and "Serranías" beside the mountain chain. These uplifts were composed mainly of bedrock, which was exposed to the surface by the action of strong winds. The streams running through these formations have the form and characteristics of Bedrock zones. Channel bed is basically the exposed bedrock, which at the same time dictates the morphology of the channel overall. Shallow slopes are characteristic of these rivers and the typical riffle-pool sequences could be observed in zones were depressions occur. In years of low rainfall, pools serve as refugia for organisms. The riparian vegetation is quite characteristic of savannas: shrubs and herbaceous strata are the most abundant. However, taller trees occur at the margins of the river channel providing some coverage. Litter packages only occur in zones were macrophytes are present, given that the armored and “flattened” bed surface are not ideal for retention. Bedrock and Macrophytes (particularly, *Macarenia clavigera*) are therefore the most common habitats to invertebrates and fishes.

ii. Reconstruct a hypothetical scenario of basin’s evolutionary history by sorting them based on their time of appearance in the basin.

I reconstruct a hypothetical scenario of basin’s evolutionary history by sorting ecoregions based on their time of appearance in the basin (figure SM1.10). The “time of appearance” was defined as the age at which appeared the most similar to contemporary vegetation and land surface forms. To distinguish between two ecoregions with similar time of appearance but different origin, I also included in this reconstruction the major geological and/or climatic events which triggered changes in their landforms or vegetation forms.

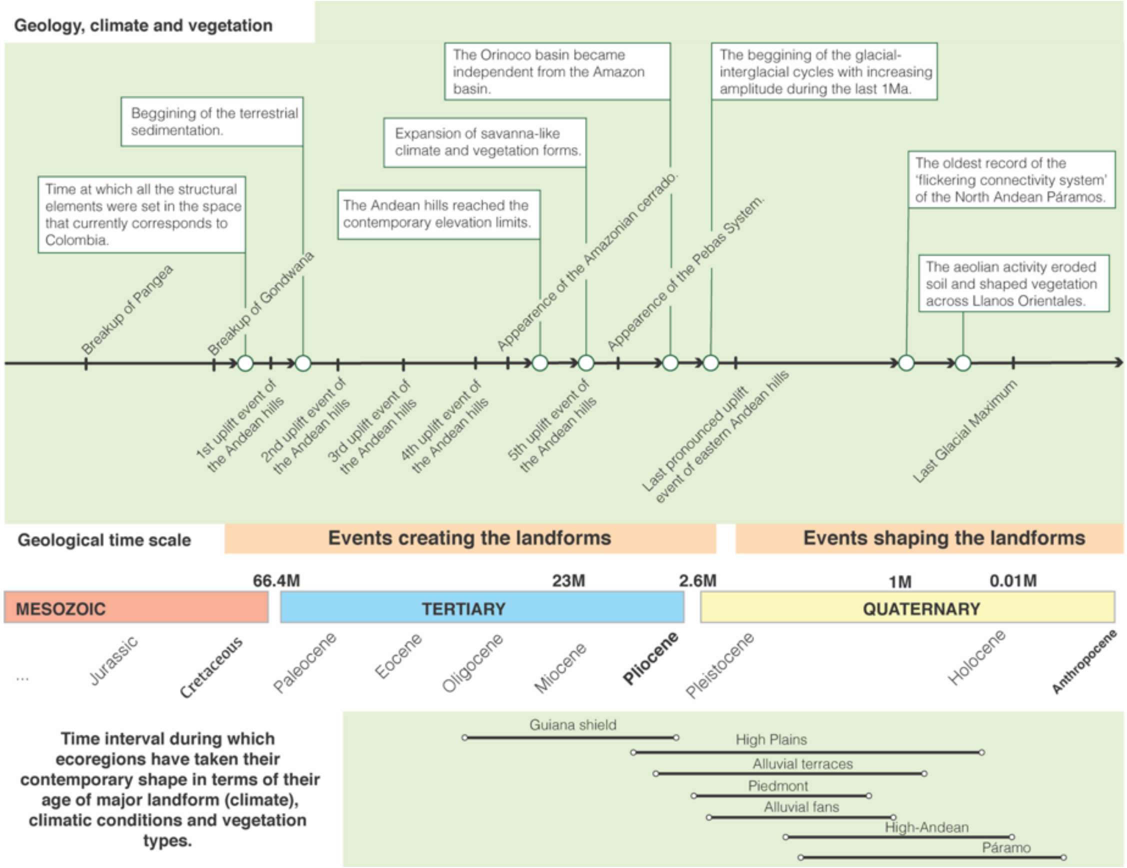


Figure I-10 Major geological and climatic events and vegetational changes occurring during the tertiary and quaternary in the Orinoco basin. The hypothetical intervals of time during which ecoregions appeared are drawn in the bottom.

iii. Translate the hypothetical scenario to a cladogram, which is then converted into a path-length matrix.

The final step consisted in numerically establishing the historical relationships among rivers. In my case, the scarcity of data did not allow to build a high-resolution model, where the relationships among rivers were included. Instead, I modeled historical relationships among ecoregions (without within-group variance) by means of a path-length matrix. While simple, including this matrix allows to establish possible relationships between Orinoco’s evolutionary history and beta diversity (dissimilarity) patterns among the ecoregions of this basin. This approach represents a starting point to dig deeper on the role of historical legacies on present-day patterns of biodiversity.

To build the path-length matrix, I followed Legendre *et al.*'s (1994) instructions. In this regard, I firstly transformed the historical relationships among ecoregions into a tree, and then, into a distance matrix; this implied rooting the tree, finding the sequence in which the ecoregions appeared during evolution (stepii, see above), and finally, writing a full distance matrix for the rivers under study (the one used in the regression models).

As starting point, I used the age of the origin of the major landforms (Table I-1, see the step ii). The age, however, did not account for a possible similar origin of two or more ecoregions. Thus, to reconstruct the sequence of apparition, I included those geological of climatic events (and the period where they occurred) linked to the origin of each ecoregion and those that triggered the differentiation in terms of landforms and those that probably triggered isolation between populations inhabiting two or more ecoregions (i.e. the last uplift event which separated the Guiana shield from the other ecoregions).

The order in which the ecoregions descend from one another along the tree was noted and transcribed in the form of a 'character-state' tree. As there is not evidence of common ancestors or 'intermediate states' between two ecoregions, all ecoregions were considered as terminal nodes of the three. The unspecified intermediate states (represented by '∇'), which represents points of divergence/separation between two or more ecoregions, were considered as evolutionary step and mapped into the tree. Each unspecified intermediate state was linked to one or more geological and/or climatic events which presumably triggered the origin of a new ecoregion.

The code of the constructed tree is the following:

> **Root → (Guiana shield, ∇ → (High Plains, ∇ → (Alluvial fans, ∇ → (∇ → (Piedmont, Alluvial fans), ∇ → (High-Andean, Páramo))))**

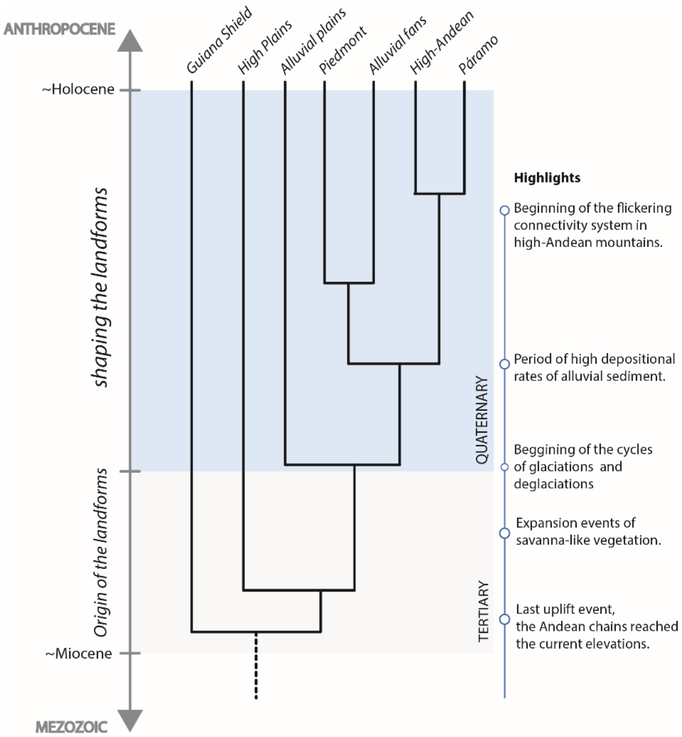
This tree reads as follows: the **ROOT** leads to the **Guiana shield** and a group of Andean-origin ecoregions, which were separated by the last uplift event (Flórez, 2003; Bicudo *et al.*, 2019). Since their separation, the vegetation and geomorphology of Guiana shield were not directly affected by Andean orogeny, probably leading to distinct evolutionary histories in this ecoregion.

The Andean-origin group leads to **High-Plains** and to a group of ecoregions directly affected by the dynamics of mountains (i.e. sediment deposition or orographic rain): the separation was mainly driven by a high sediment deposition, which was triggered by the last uplift events; and changes in the climate, which triggered the expansion of savannas (Goosen, 1971; Flórez, 2003). The isolation of this ecoregion was strengthened by the occurrence of aeolian activity which eroded soils and constrained vegetation structure to savanna-like species (Goosen, 1971; van der Hammen, 1974; Gonzalez-Penagos *et al.*, 2014).

The last group of ecoregions leads to **Alluvial plains** and to a group of **montane ecoregions**; the former being originated from the deposition of alluvial sediments coming from the glaciation-deglaciation dynamics occurred within the quaternary (Goosen, 1971; van der Hammen, 1974; Flórez, 2003). Species from this ecoregion were isolated from those in the high plains because of their differences in terms of climate, soils and hydrology. Differences driven by the aeolian activity and the partial uplift of the Meta fault (Goosen, 1971; van der Hammen, 1974). This last uplift, which was progressive, elevated high Plains 10-to-30m with respect to alluvial terraces; and also, changed the hydrology of the zone. Currently, riverscapes located northern to the Meta fault tend to flood in wet seasons; meanwhile, in those located southern to the Meta fault, water is scarce and tend to flow in rivers with a well-protected riparian zone.

The group of montane ecoregions leads to the last four ecoregions, which were differentiated mainly by changes in the regime of sediments, changes in altitude and changes in climatic conditions (van der Hammen, 1974). It is worth to notice that two of the most recent ecoregions (high-Andean and Páramo) have a similar origin and they were differentiated by changes in connectivity triggered by glaciation dynamics during the last 2My (resembling ‘a flickering connectivity system’ *sensu* Flantua *et al.*, 2019).

The graphical representation of the tree is:



As a final step, the tree was transformed into a path-length matrix. Path lengths were measured by counting the number of steps separating ecoregions in the hypothetical reconstruction. In this way, the final path-length matrix describing historical relationships among ecoregions looks like:

	ROOT	*GUIANA SHIELD*	*HIGH PLAINS*	*ALLUVIAL PLAINS*	*PIEDMONT*	*ALLUVIAL FANS*	*HIGH ANDEAN*	*PARAMO*
ROOT	0	1	2	3	4	4	4	4
GUIANA SHIELD		0	3	4	6	6	6	6
HIGH PLAINS			0	3	5	5	5	5
ALLUVIAL PLAINS				0	3	4	4	4
PIEDMONT					0	2	4	4
ALLUVIAL FANS						0	4	4
HIGH ANDEAN							0	2
PARAMO								0

Study Zone

The study was carried out at 26-32 different stream segments within an area of about 40,000km^2, in the Colombian Orinoco. The area encompasses an elevation gradient from 3400 to 300m a.s.l. and includes a heterogeneous assembly of ecoregions and landscapes. From three to six pristine or near-pristine streams were selected and sampled in each one of the seven ecoregions (Figure I.12).

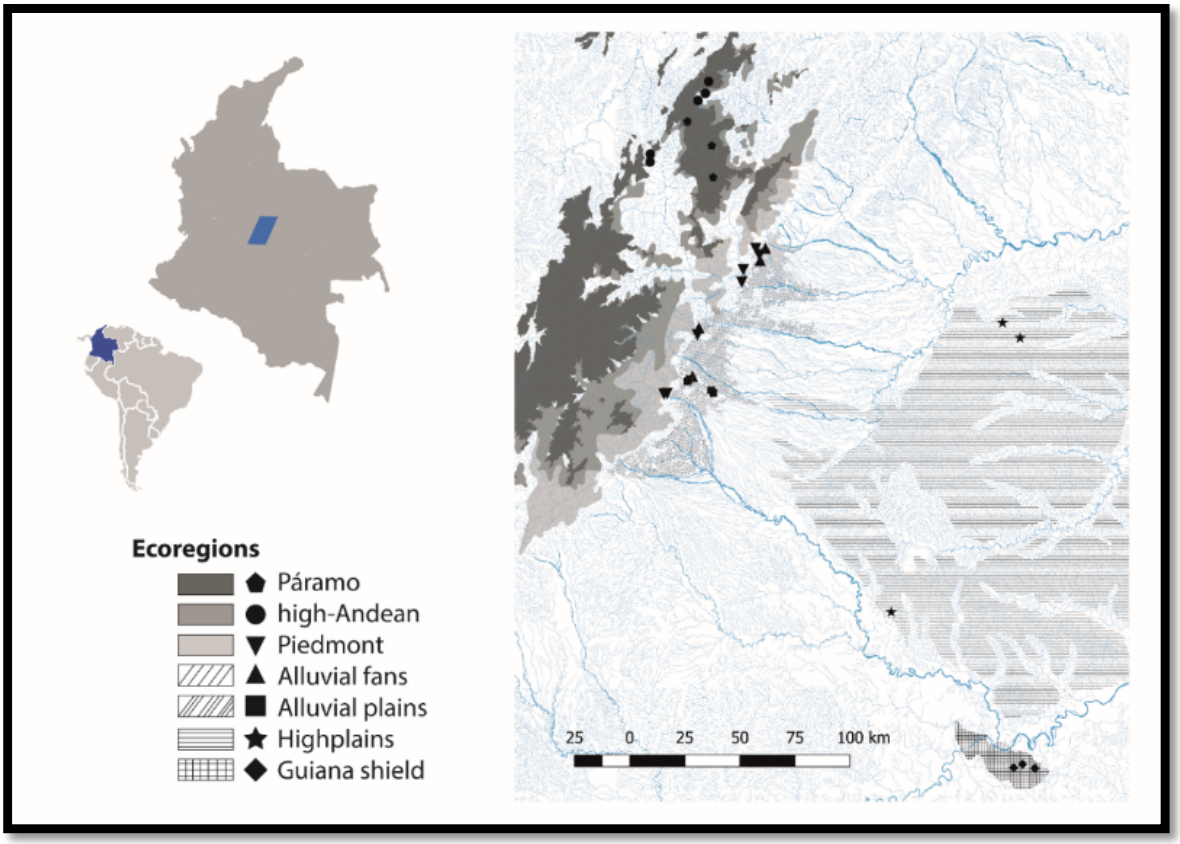


Figure I-11 Geographical extent of ecoregions and position of the sampled streams in the Orinoco basin.

The Table I-2 contains the water quality and physiographic features of the rivers sampled in the Orinoco basin.

Table I-2 Mean and coefficients of variation for the physiographic and water chemistry variables within ecoregions.

		Guiana shield		**Paramo**		**high-Andean**		**highPlains**		**Piedmont**		**Alluvial.fans**		**Alluvial plains**	
		Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)
Water Chemistry	*Temperature (ºC)*	26.3	4	11.4	12	11.4	12	25.6	5	22.3	4	23.5	7	23.4	3.8
	Conductivity (µS/cm²)	4.6	47	39.6	66	16.7	57	7.7	67	151.9	196	200.7	200	6.5	8.2
	O2 (mg/L)	6.2	24	7.2	5	7.8	5	5.1	15	7.4	8	7.2	10	7.3	7
	% Sat. O_2	78.5	23	96.5	2	98.1	3	63.3	14	92.3	8	90.7	10	92.1	2.9
	pH	5.0	18	7.5	8	6.8	16	4.8	13	6.2	22	6.3	15	4.9	9.3
	PRS (ug/L)	8.3	55	12.9	64	22.0	84	9.8	65	12.7	52	10.4	43	10.6	18.3
	N-NH_4 *(ug/L)*	14.3	42	10.7	39	9.3	39	19.8	70	10.1	54	11.0	56	6.8	23.5
	N-NO_3 *(mg N/L)*	0.03	32	0.02	0	0.05	87	0.03	72	0.10	182	0.09	74	46.6	44.2
	TOC (mg C)	4.6	24	3.0	27	4.1	45	3.5	33	4.8	178	2.7	88	4.48	19.2
Physiography	*Elevation (m a.s.l.)*	221.00	4	2136.20	62	2729.64	9	227.00	2	685.85	19	552.00	6	548.5	16
	Slope (%)	2.67	2	13.67	31	12.56	17	0.67	20	11.57	32	5.83	35	6.38	39.2
	Discharge (m³/s)	0.28	32	0.54	75	0.36	63	0.07	58	0.37	143	1.46	88	0.34	88.2
	Channel width (m)	9.89	22	6.14	34	5.23	37	4.13	39	5.49	37	15.50	56	6.9	26.1
	TSS (mg/L)	0.002	123	0.007	143	0.001	100	0.004	36	0.005	139	0.002	60	0.001	0.0001
	% Bedrock	89	33	5	150	19	83	0	-	10	144	4	192	0	-
	% Boulder	2	165	47	31	43	26	0	-	43	41	36	50	27	114
	% Cobble	7	248	30	37	20	37	0	-	18	59	27	46	45	45
	% Pebble	4	224	12	50	15	69	0	-	21	59	24	45	20	51
	% Gravel	2	150	4	171	1	264	13	210	7	108	8	59	8	100
	% Sand and Lime	2	150	0	-	0	-	87	32	3	178	1	255	0	0

II. Material and Methods

Here I provide an overview of the methods employed in this thesis to dissect the assembly process of benthic communities in the Orinoco basin (Table II.1). A detailed description of each one of analyses and techniques can be found in next chapters.

Table II-1 Summary of the sampling and data-analysis methods employed to study the assembly process of benthic communities in the Orinoco basin.

		Chapter III	**Chapter IV**	**Chapter V**	**Chapter VI**
Community data	*Diatoms*		X	X	X
	Invertebrates	X		X	X
	Traits				X
	Phylogenies				X
Environmental characterization	*Hydrology*	X	X		X
	Water Quality	X	X	X	X
	Canopy shading	X	X	X	X
Spatial modelling	*MEMs*	X	X		X
	Ecoregional structure	X	X		
Data analyses	*Spp. accumulation curves*	X	X	X	
	Spp. Indicator Value (IndVal)	X	X		
	Additive partitioning	X			
	CAP	X	X		
	PERMDISP	X	X		
	EMS	X			
	RDA + Variance partitioning	X	X		
	Modelling Beta diversity			X	X

** **MEM** – Moran Eigenvector Maps; **CAP** - Canonical analysis of principal coordinates; **PERMDISP** – Test for Homogeneity of Multivariate Group dispersions; **EMS** – Analysis of the Elements of Metacommunity Structure; **RDA** – Redundancy analysis.*

Community-level data

Community sampling

In total, I sampled 26 (for diatoms) and 32 (for invertebrates) stream segments during the dry season of 2016 (from January to February). Sites were in separated branches of different tributaries and scattered throughout the three large forms of relief (see Chapter I).

Diatom sampling. Diatom communities were sampled at three different riffle sections of each stream segment, each section spanning from 20 to 60m long. The selected riffles were representative of the stream's range of substrata types, flow velocities, channel widths and depths, and canopy cover. At each riffle section, we collected 8 cm^2 of surface, brush-scraped algal material from 30 boulders of similar size. Algal material was pooled by riffle section (= 3 samples per stream segment) and subsequently preserved in a *Transeau* solution.

Invertebrate sampling. Invertebrates were collected using a multi-habitat sampling procedure, with 5 Surber (mesh size: 350um; area: 0.09 m^2) samples collected in stream substrata that were selected according to their corresponding habitat coverage. For instance, if a riffle was composed of 60% of boulders, 30% gravel, and 10% cobbles, 3 Surber samples of the first, 1 of the second, and 1 of the third substratum type were collected. The substratum distribution in each riffle was evaluated visually using the Wentworth scale (mm, diameter-based) as a reference (Wentworth, 1922). We only sampled boulders (diameter >250mm) smaller than the sample frame.

Taxonomical identification

Diatoms. In the laboratory, samples' organic material was cleaned using hydrogen peroxide. Clean diatom frustules were mounted on permanent slides using a Naphrax® medium, the slides were then observed under a 1000x light microscope and identified at the finest level possible using specialized monographs. At least 400 valves were counted in each slide, and counts processed as relative abundances.
Diatom monographs used for determination:

- Bellinger E.G. & Sigee D.C. (2015). *Freshwater algae: identification and use as bioindicators*. John Wiley & Sons.
- Krammer K. & Lange-Bertalot H. (1991). Bacillariophyceae. 2/3: Centrales, Fragilariaceae, Eunotiaceae. In: *Süsswasserflora von Mitteleuropa*.
- Krammer K. & Lange-Bertalot H. (1986). Bacillariophyceae. *Susswasserflora von Mitteleuropa.*
- Metzeltin D. & Lange-Bertalot H. (2007). *Tropical Diatoms of South America II. Special remarks on biogeography disjunction. Iconographia Diatomologica 18: 1-877, incl. 296 pl.*

Invertebrates. Individuals were sorted and identified to the genus level, following specialized monographs. Chironomidae and Ephemeroptera were dissected and mounted in Euparal® following the protocol of Domínguez (2006) and Andersen et al. (2013). The pupae of Chironomidae were mounted to confirm some taxonomical identities (Prat et al. 2014).

Monographs used for determination:

- Andersen P., Cranston P.S. & Epler J.H. (2013). *Chironomidae of the Holarctic Region: Keys and Diagnoses, Part 1: Larvae*. Entomological Society of Lund, Sweden.
- Domínguez E. & Fernández H.R. (2009). *Macroinvertebrados bentónicos sudamericanos: sistemática y biología*. Fundación Miguel Lillo Tucumán.
- González-Córdoba M., del Carmen Zúñiga M. & Manzo V. (2015). Riqueza genérica y distribución de Elmidae (Insecta: Coleoptera, Byrrhoidea) en el departamento del Valle del Cauca, Colombia. *Biota Colombiana* **16**, 51–74.
- Merritt R.W. & Cummins K.W. (2008). *An introduction to the Aquatic insects of North America*. Kendall/Hunt Pub. Co.
- Prat N., González-Trujillo J.D. & Ospina-Torres R. (2014). Key to chironomid pupal exuviae (Diptera: Chironomidae) of tropical high Andean streams. *Revista de Biologia Tropical* **62**
- Trivinho-Strixino S. & Strixino G. (1995). *Larvas de Chironomidae (Diptera) do Estado de São Paulo: identificação e diagnose dos - 50 - géneros*.

Trait-based databases

The functional diversity of diatom communities was described in terms of four traits (Table II.2). These traits were selected because of their documented response to in-stream environmental conditions and dispersal-related constraints (see Passy, 2007 and Lange, Townsend & Matthaei, 2016). Each trait was resolved into different discrete states (see the Supplementary Material). Trait states were described based on microscope observations and on the available literature (Lange *et al.*, 2016) and online databases (www.algaebase.com).

Table II-2 Diatoms traits. Biological traits, and their categories used to estimate the functional diversity of diatom communities.

	Trait	States
A.	Cell volume	Nano (5 < 100 um^3), Micro (100 < 300 um^3), Meso (300 < 600 um^3), Large (600 < 1500 um^3), Very Large (>1500 um^3)
B.	Growth form	Solitary, filamentous, colonial.
C.	Attachment to substrata	Postrate, Epyphite, non-attached, Pad-stalks, Mucilage tubes
D.	Motility	Drift, flag-glid, motile, non-motile

The trait composition of insect communities was described in terms of seven traits (Table II.3). These traits were selected because of their documented response to in-stream environmental conditions and dispersal-related constraints. Particularly, we focused on traits linked to taxon dispersal and on those that respond to changes in habitat heterogeneity, flow-related disturbance, resource availability. To describe each trait, we used field and laboratory observations as well as secondary information gathered from previous studies from the Neotropical zone (Tomanova & Usseglio-Polatera, 2007; González-Trujillo, 2016). Each trait was resolved into different states and a fuzzing code procedure (Chevene *et al.*, 1994) was used to describe the affinity of each genus to each trait state. A value between 0 (no affinity) and 3 (high affinity) was assigned to each trait category (see the Supplementary Material). The affinity values of each trait for each genus were standardized so that their sum for a given genus and for a given trait was 1.

Table II-3 Invertebrate traits. Biological traits, and their categories used to estimate the functional diversity of invertebrate communities.

	Trait	States
A.	Respiration mode*	Spiracles – 'spi'; Gas gills – 'ggill'; Plastron and Spiracular gills – 'pla:sgil'; Tegument – 'teg'; Tracheal gills – 'tgil'; blood-based respiration – 'blo'
B.	Morphological adaptations to flow constraints	Case – 'case'; Suckers – 'Suc'; Silt glands – 'silt'; abdominal claws – a.clw; tarsal claws – t.clw; no adaptation – 'n.flow'
C.	Body size (At adult or last instars phases)	BS1 (<2.5mm); BS2 (2.5-5mm); BS3 (5-10mm); BS4 (10-20mm); BS5 (20-40mm); BS6 (>40mm)
D.	Feeding strategy	Collector-Gatherer – 'CG'; Shredder – 'SH'; Scrapper – 'SC'; Filterer – 'F'; Predator – 'P'
E.	Locomotion	Surface swimmer – 'SW'; Full water swimmer – 'FW'; Crawler – 'Cr'; Burrower/Interstitial – 'Bu'; Temporary attached – 'TA'; Permanently attached – 'PA'
F.	Aquatic stages	Egg – 'egg'; Larva/Nymph – 'inmat'; pupa – 'pup'; Adult – 'adu'
G.	Dispersal type	Aquatic passive – 'AcP'; Aquatic active – 'AcA'; Aerial passive – 'AeP'; Aerial active – 'AeA'

*** Trait states are ordered according to their efficiency in oxygen uptake.**

Phylogeny construction

I used the time-calibrated phylogeny of diatoms published by Nakov *et al*. (Nakov *et al.*, 2018), which is the most inclusive diatom species phylogeny to date. This phylogenetic approach was reconstructed by maximum likelihood using 11 genes (18 S rRNA, 28 S rRNA, 16 S rRNA, *atp*B, *psa*A, *psa*B, *psb*A, *psb*C, *rbc*L, *cob* and *cox*I) and 38 calibration points (Nakov *et al.*, 2018). The phylogeny included almost all taxa of our community dataset. The missing taxa were grafted to the most recent common ancestor of all members sharing the lowest taxonomic level available in the tree. This way, each inserted taxon was separated from the other on the basis of its taxonomic name and replaced on an independent branch in the tree.

I used the supertree published in Saito et al. (2015) as a basis for constructing our hypothetical phylogenetic tree for aquatic insects. Saito et al.'s tree contains a great number of the genera of our community dataset. Similar to the procedure for constructing the diatom supertree, the missing insect taxa were grafted to the most recent common ancestor of all members sharing the lowest taxonomic level available in the tree. We used information of ages nodes compiled from several molecular phylogenies (see Saito et al., 2015 for details). When this information was not available, the topology was solved using taxonomy because it is a reasonable surrogate to phylogeny. In this regard, the Linnean classification system was used to establish relationships between sister groups. Analogous to Saito's et al. (2015), we compiled the ages of 28 nodes and assigned the supertree branch lengths using the BLADJ algorithm in Phylocom. BLADJ spaces undated nodes evenly between dated nodes using an adjuster algorithm (Webb *et al.*, 2008).

I used the Mesquite software (Maddison & Maddison, 2018) for editing the diatom and insect supertrees.

Environmental characterization

Environmental descriptors for each stream segment included hydrology, substratum, and water quality.

Canopy shading (%) was estimated analyzing a series of upward photographs taken using a fisheye lens. The luminosity and white balance of each photo were manually adjusted to color leaves in black and sky in white. The proportion between black and white pixels was used to estimate the percentage of canopy shading.

Water quality. Physical and chemical variables were measured during invertebrate sampling (January-February 2017) and on two further occasions (November 2016 and, January-February 2018), that corresponded to high and low water flows, respectively.

Conductivity, **pH**, **oxygen**, and **temperature** were recorded using a HANNA HI98194 water quality meter upon arrival (early morning) and departure (dawn) from the site. **Instantaneous discharge** was estimated in the three riffles by measuring of water depth and flow velocity at 15cm intervals along three cross-sections. At each interval, we also recorded the dominant substrate. **Flow velocity** was measured with a digital flow meter (SCHILTKNECHT – MiniAir 20).

On each occasion, one liter of water was collected for physico-chemical analyses, filtered through 0.7µm glass fiber filters (Whatman GF/F, Kent, UK) and stored frozen until analysis. In the laboratory, **ammonium** and **nitrate** concentrations were determined on a Dionex ICS-5000 ion chromatography system (Dionex Corporation, Sunnyvale, U.S.A.). **Reactive**

phosphorus (PRS) concentrations were determined colorimetrically using the fully automated discrete analyzer Smartchem 140 (AMS Allaince, Frépillon, France). **Total suspended solids (TSS)** were analyzed by filtering 500ml of water through a pre-weighed GFF and drying the filtrate for 1 hour at 105ºC.

Hydrological variables. Long-term hydrological variables were estimated using the rational method modified by Témez (2003). I used this method because of the absence of historical series of the stream's water flow discharge. Benthic community patterns cannot be properly explained without incorporating the hydrological variability of streams. Therefore, being aware of the limitation of the method in terms of its accuracy, I used the simplest descriptors of stream's hydrological variability of one year before samplings.

The rational method estimates a stream's water flow as function of the total precipitation, the basin area and associated land uses, the time of concentration, and the runoff coefficient. Once the daily water flow had been determined, we estimated the threshold at which the stream's basal flow was surpassed, as a unit of disturbance for benthic communities. We then calculated: (i) the number of days elapsed since the last flood event (defined as the one doubling the basal flow discharge); (ii) the number of flood events; and (iii) the ratio between the maximum and basal flow discharges.

The Rational Method modified by Témez (2003). Témez proposed that the instantaneous discharge can be estimated by using the following formula [1]:

$$\mathrm{Q} = \frac{C \cdot I \cdot A}{3{,}6} \cdot K \; [1]$$

Where, the instantaneous discharge (*Q*) is estimated as a function of a runoff coefficient (*C*), the average precipitation intensity (I), sub-basin area (*A*), the

Uniformity coefficient (*K*), and a constant value. The sub-basin area was estimated from a Digital Elevation Model in QGIS 2.18 (http://qgis.osgeo.org).

- *Runoff coefficient – C:*

The runoff coefficient is estimated as a function of the Maximum daily precipitation (*Pd*) and the runoff threshold value (P_0). When the daily precipitation does not exceed the threshold, the runoff coefficient is equal to 0. The following formula is used [2]:

$$\mathrm{C} = \frac{(Pd - P) \cdot (Pd + 23 \cdot P_0)}{(Pd + 11 \cdot P_0)^2} \; [2]$$

The runoff threshold (P_0) is a parameter whose value is estimated from a table in Témez (2003). Depending on the sub-basin land use, slope, hydrologic features, and soil type, the value will be higher or slower. Therefore, we obtained a value per stream. All these features were assessed in QGIS 2.18 using the Shapefiles provided by governmental entities. Shapefiles were downloaded from SIAC webpage (http://www.siac.gov.co/catalogo-de-mapas).

- *Average precipitation intensity – I:*

The average precipitation intensity is estimated as a function of the average daily precipitation (I_d = punctual precipitation / 24h), the ratio I_1/I_d, and the concentration time – *t* (eq. 3).

$$\mathrm{I} = I_d \cdot \left(\frac{I_1}{I_d}\right)^{\frac{28^{0,1} - t^{0,1}}{28^{0,1} - 1}} \; [3]$$

The ratio I_1/I_d is taken from a map of isobars, and the concentration time (*t*) is estimated as a function of stream length (*L*) and slope (*J*) (Eq. 4). Stream length and slope were measured from a Digital Elevation Model in QGIS 2.18.

$$\mathrm{t} = 0{,}3 \cdot \left(\frac{L}{J^{1/4}}\right)^{0,76} \; [4]$$

- *Uniformity coefficient – K:*

The uniformity is coefficient is a value for correcting the concentration time. It is calculated as a function of the concentration time (*t*) (eq. 5).

$$\mathrm{K} = 1 + \cdot \left(\frac{0{,}61^{1{,}25}}{0{,}61^{1{,}25} + 14}\right) [5]$$

Spatial modelling

Moran's eigenvector maps (MEMs) were used to model the spatial structure of the streams within each ecoregion (Legendre & Legendre, 2012). Following the method of Declerck et al. (2011), the resulting eigenvector maps were arranged in blocks in a staggered matrix, each block corresponding to a distinct ecoregion. MEMs model the spatial dependence in a set of locations as an orthogonal combination of MEM variables derived from geographic coordinates, the number of ecoregions ("blocks"), and the stream within each ecoregion (Declerck *et al.*, 2011). I used the "create.dbMEM.model" function of the "adespatial" package (Dray *et al.*, 2018) to create the MEM model.

Ecoregional structure. I created a matrix of seven dummy variables, representing each ecoregion type, to model the dependence of community structure on the ecoregional characteristics.

Statistical analyses

Species accumulation curves are graphs which record the cumulative number of species or taxa in a sample, stream or basin as a function of the cumulative effort expended searching for them. In this thesis, the cumulative effort was measured at the riffle level.

Analysis of Species Indicator Value. This analysis (IndVal, Cáceres and Legendre 2009) look for the preference of species to each ecoregion. The IndVal analysis is based on the concepts of *specificity* (highest when the species is present in just one ecoregion) and *fidelity* (highest when the species is present in all streams of one ecoregion). A high indicator value is obtained by a combination of high specificity and fidelity. I performed IndVal analysis using the function “multipatt” from the indicspecies package (Cáceres & Legendre, 2009).

Additive partitioning of diversity. This approach accounts for the compositional differentiation of groups (e.g. patches) by partitioning the regional or gamma (γ) diversity into within- and between- group components, representing the alpha (α) and beta (β) diversities, respectively (Jost, 2007). We used the approach of Crist et al. (2003) to include distinct spatial scales in a hierarchical sampling design, where the gamma (γ) diversity can be expressed as the proportional contributions of diversity due to each level of the hierarchy (Crist et al. 2003). The highest level corresponded to the Orinoco basin, with its gamma diversity partitioned within (α) and among riffles of each stream (β1), among streams of each ecoregion (β2), and among ecoregions within the basin (β3).

Canonical analysis of principal coordinates (CAP, Anderson & Willis, 2003). This analysis identifies the axes through the multivariate space that best discriminate among *a priori* groups. In this thesis, I allocated riffles to correct ecoregion types (leave-one-out procedure) and test for among-ecoregion differences in community composition using random permutations (9 999 runs). I performed CAP calculations with the 'CAPdiscrim' function of the BiodiversityR package (Kindt & Kindt, 2018).

Test for homogeneity of dispersion (PERMDISP, Anderson, 2006). This analysis was used to estimate the within-ecoregion dissimilarity and to compare among-ecoregion differences in the distance between the observations and the group centroid. The significance of the among-group differences was tested with permutation tests for least-squares residuals. The null hypothesis of no differences the among-ecoregions in terms of environmental and community dissimilarity was tested using a permutation test with 9,999 runs. We performed PERMIDISP tests with the 'betadisper' function of the VEGAN package (Oksanen *et al.*, 2013).

Elements of metacommunity structure (EMS). This analysis determines which idealized metacommunity structure (Leibold & Mikkelson, 2002) or quasi-structure (Presley *et al.*, 2010) bests fits the observations obtained from the community matrix. Based on a species-by-site incidence matrix, in which sites are rearranged by similarities in species composition and species are rearranged by similarities in distribution (Leibold & Mikkelson, 2002), EMS analysis assesses the coherence, turnover, and boundary clumping of species distributions, looking for the best-fit model. I used the 'Metacommunity' function of the METACOM package (Dallas, 2014) to perform all EMS analyses.

Redundancy Analysis (RDA). The relative contribution of the environmental variables, spatial structure and ecoregional identity on metacommunity composition was quantified with the distance-based redundancy analysis (Legendre & Anderson, 1999), followed by variation partitioning analysis (Borcard *et al.*, 1992). RDA and variance partitioning analysis were performed using the “rda” and “varpart” functions of the VEGAN package (Oksanen *et al.*, 2013), respectively.

Modelling beta diversity. To model beta diversity patterns, I used linear models and generalized linear models depending on the study question and the type of descriptors (e.g. continuous, discrete or distance-based). The patterns of beta-diversity within streams were quantified using a simple linear regression of the pairwise distances between reaches of each stream. The coefficients of these regressions were used to evaluate if the community-environment relationship change at each stream and depending on the community.

I used two approaches to model beta diversity patterns among streams. **The first approach** included pooling the pairwise distances between reaches of each stream and performed generalized least squares (GLS) regressions. The coefficients of these regressions allow to evaluate if the responses to deforestation (Chapter V) varied depending on the community. I used GLS since these models can deal with data violating the heteroskedasticity assumption - heterogeneous variances among groups (streams). The heterogeneity of each stream was included in the random term of the model as suggested by Zuur et al. (2009).

In **the second approach**, I modelled the pairwise similarity against the geographic distance, the environmental dissimilarity matrices and a path-length matrix that indicated the evolutionary hypothesis of the streams in the Orinoco basin. Models were built using a GLM approach with a Quasi-Binomial error distribution with a logit link function. As described by Dias *et al.* (2014), using this approach has three advantages: (1) it accounts for values found between 0 and 1, (2) it handles similarity values of 0 (minimal similarity) and 1 (maximal similarity) without previous trans- formation and (3) it accounts for the usually curvilinear shape of the distance–decay plots. To test the significance of partial regression coefficients, I performed a permutation of the response variables (i.e. the rows and columns of dissimilarity matrices) while the original order of the descriptors was held constant. Therefore, the statistical significance was assessed using the frequency at which the null simulated coefficients were higher and/or lower than the observed coefficients.

III. Historical processes constrain metacommunity structure by shaping different pools of invertebrate taxa within the Orinoco basin

ABSTRACT

The aim was to identify and characterize the distribution of invertebrate taxa in the Orinoco basin and how their distribution affects the metacommunity structure along the river network. We characterized the invertebrate communities and environmental characteristics of 25 streams from six ecoregions in the Orinoco basin. The ecoregions present different historical evolution, altitude, slope, climate, and vegetation features. We used multiplicative diversity partitioning to compare the contributions of riffle (α and β1), stream (β2) and ecoregional (β3) scales to the overall gamma diversity of the basin. We also applied the Elements of Metacommunity Structure (EMS) framework to delineate metacommunity types and a distance-based redundancy analysis (dbRDA) to assess the relative relevance of environmental, spatial, and ecoregional drivers in species compositions of invertebrate communities. Streams showed significant differences in community composition among the ecoregions. Several discrete pools of invertebrate taxa occurred in the basin that largely matched the distribution of the ecoregions. Consequently, the metacommunity in the basin resembled a Clementsian idealized structure. The species composition of invertebrate communities was mostly explained by ecoregion type and its interaction with the local environment, particularly its physiographic features. Historical and evolutionary processes have resulted in species pools differing between the ecoregions of the Orinoco basin. At the basin scale, the metacommunity structure seems to be constrained by ecoregional features rather than by spatial structure or the local environment. Hence, using the basin as a unit for biodiversity conservation and river management would not adequately reflect the diversity and distribution patterns in highly heterogeneous basins such as the Orinoco basin.

-o-

An edited version of this chapter has been published in the journal of ***Diversity and Distributions****. Please, cite the respective paper as:*

- *González-Trujillo, J. D., Donato-Rondon J. Ch, Muñoz, I., Sabater, S. Diversity and Distributions (2019).* [*https://doi.org/10.1111/ddi.12996*](https://doi.org/10.1111/ddi.12996)

INTRODUCTION

Basins have been considered an appropriate scale to unravel the metacommunity assembly process in riverine ecosystems (Brown & Swan, 2010; Grönroos et al., 2013; Cañedo-Argüelles et al., 2015). In a metacommunity context, tributary communities are viewed as subsets of a regional species pool whose species composition is shaped by environmental sorting, dispersal-driven processes and stochastic events. As they partly or completely encompass a river network within a delimited area, a species pool common to all the tributary communities is usually expected (Brown & Swan, 2010). However, a unique species pool cannot always be assumed to occur across the basin (Higgins *et al.*, 2005). Historical events, such as those that occurred during the Tertiary and Quaternary periods, could lead to the development of several species pools that are patchily distributed within basins or sub-basins (Hubert & Renno, 2006; Hoorn *et al.*, 2010).

The metacommunity framework allows assessment of the role of historical, regional, spatial, and environmental factors in the assembly process across scales (Mittelbach & Schemske, 2015; Leibold & Chase, 2018). In recent decades, the use of pattern-based statistical methods, such as the analysis of the Elements of Metacommunity Structure (“EMS”, Leibold & Mikkelson, 2002) and direct ordination approaches (e.g., “dbRDA”, Legendre & Anderson, 1999), have helped to disentangle the main drivers of metacommunity structure and dynamics (Leibold & Chase, 2018). The major constraints for metacommunities at the basin scale, for instance, include environmental sorting, the dendritic structure of the river network and its effects on habitat connectivity, and the dispersal ability of the species (e.g. Brown & Swan, 2010; Grönroos et al., 2013; Cañedo-Argüelles et al., 2015; Sarremejane et al., 2017). However, the influence of these processes on metacommunity structure may be affected by the occurrence of distinct pools of species in the basin (Benito *et al.*, 2018).

A “pool” or “regional pool” of species refers to the set of all species available to colonize local communities (Srivastava, 1999). Since the introduction of this concept in the Theory of Island Biogeography, species pools have helped to disentangle the relative contributions of local, regional and historical factors to shaping metacommunity structure (Cornell & Harrison, 2014). An accurate characterization of species pools can help to determine how large-scale forces interact with local constraints to shape the species composition of local communities and metacommunities across different scales (Ricklefs, 1987; Cornell & Harrison, 2014). Indeed, almost all analyses to address the assembly process (e.g. null models or EMS) depend on how the species pool is defined in its composition and spatial extent (Lessard *et al.*, 2012; Buschke *et al.*, 2014). Hence, the appropriate characterization of the species pool(s) should be the first step towards understanding the assembly process (Cornell & Harrison, 2014).

In the Neotropical region, geological and climatic events such as Andean uplifts and glacier retreats have shaped unique combinations of river forms and riparian ecosystems at the ecoregional scale (Rull, 2008; Hoorn *et al.*, 2010). For example, when descending from the Andes to the Amazon, there is a matrix of streams with constrained channels that are surrounded by shrubs in the Páramo, followed by rivers with gorge channels and steep slopes in the Andean-cloud forest and Piedmont regions, and then meandering rivers surrounded by well-developed rainforest in the Amazonian region. Since these ecoregional features do not always match the extension of the stream networks, this scenario provides the opportunity to test the influence of the potential occurrence of different pools of species metacommunity structure. The few comparable studies carried out in the Neotropics have indicated the occurrence of several pools of species across the region (Brasil *et al.*, 2017; Datry *et al.*, 2017; Benito *et al.*, 2018). However, to our knowledge only one has addressed how the occurrence of different pools of species may shape the structure of phytoplankton metacommunities

(Benito *et al.*, 2018). This study, however, included a community of poor dispersers in disconnected systems (Andean lakes); which lead to clumped dispersion patterns in species distribution (Vilmi *et al.*, 2017).

In this study we characterized the distribution of invertebrate species pools in the Orinoco basin, and how their distribution affects the structure of the metacommunity across the stream network of the basin. In the absence of major disturbances, these patterns may provide an appropriate comparator to shed light on the effects of deforestation, mining, damming, and intensive agriculture on the biodiversity in the region (Sabater *et al.*, 2017). We hypothesized that if evolutionary history influenced the structure of freshwater communities in parallel with ecoregions' physiography and geomorphological structure, a nearly discrete species pool would characterize each ecoregion. If this were the case, beta diversity would be greater between ecoregions than between streams of the same ecoregion. To test this hypothesis, we characterized and compared invertebrate communities from six different ecoregions which differ in their historical evolution, current climates, vegetation cover, and riverine channel form and slope (Table I-1).

MATERIAL AND METHODS

Study area

The Orinoco basin is the third largest basin in South America, covering an area of about 990,000 km^2 that is in most of Venezuela and in the eastern part of Colombia (Romero Ruíz *et al.*, 2004). The complex geological and climatic history of the basin has shaped a broad range of ecosystems across heterogeneous landscapes (Romero Ruíz *et al.*, 2004). The geological and climatic events of the Late Miocene and Pliocene, in particular, have shaped the current physiographic features and ecoregions (e.g. Goosen, 1971; van der Hammen, 1974; Hughes & Eastwood, 2006; Hoorn et al., 2010). Within the basin, there is an intricate network of rivers and tributaries, spanning from 0 to 3500m a.s.l, and three constraining relief forms: ancient massifs and shields, recently raised ridges, and tectonic depressions or accumulation plains (Stallard, 1985). The water chemistry of the streams and rivers is primarily determined by the geological features (Stallard, 1985), while historical events have contributed to shaping riverscape biodiversity (Říčan, 2011).

We sampled aquatic invertebrates at 25 stream segments within an area of about 40,000km^2 in the Colombian Orinoco (Figure III-1). The area encompassed a broad altitudinal gradient (from 3,400 to 300 m a.s.l) to include a largely heterogeneous area. Our sampling sites were distributed among six ecoregions of the Orinoco basin, with 3 to 7 streams selected in each ecoregion according to their geomorphological/physiographical comparability (see Table I-1) and with no signs of anthropogenic impact. We selected only pristine or near-pristine streams. A larger number of streams were sampled in the ecoregions with greater variability (i.e. Alluvial fans n=4, high-Andean n=5, Piedmont n=7), based on pre-survey data (González-Trujillo *unpublished data*).

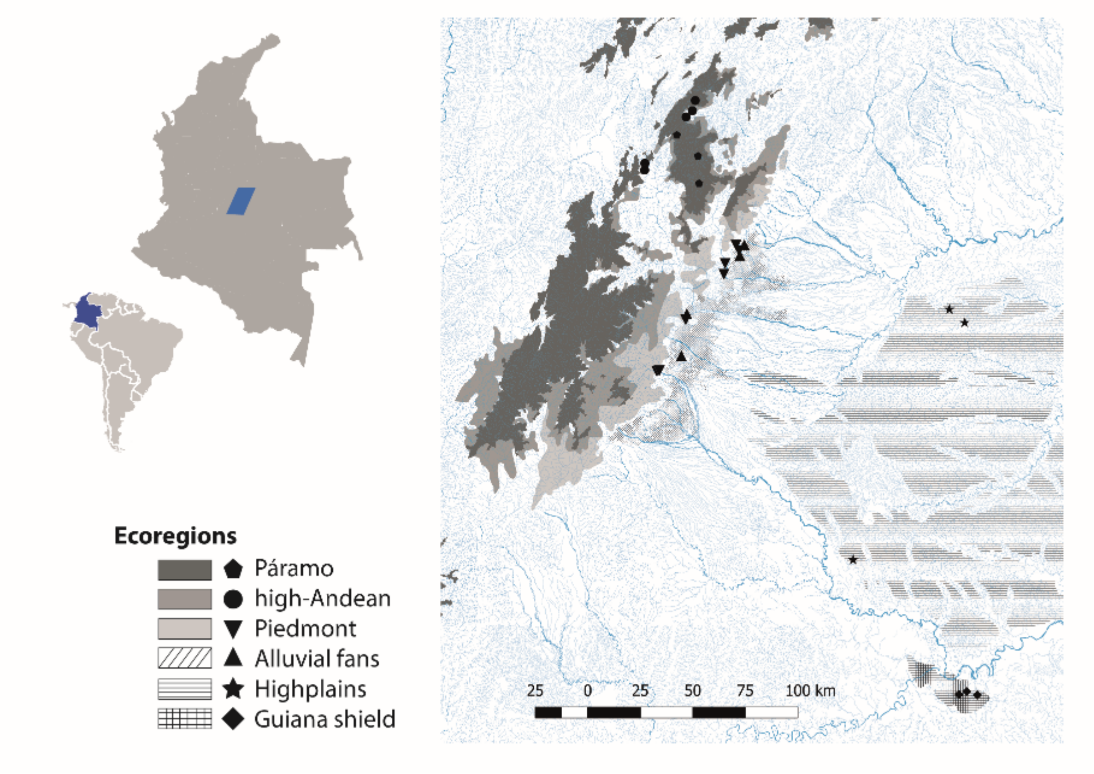


Figure III-1 Geographical area of the ecoregions and the positions of the sampled streams in the Orinoco basin.

Environmental characterization of the streams

Environmental descriptors for each stream segment included hydrology, substratum, and water quality (Table III-1). Longer-term hydrological variables were estimated using the rational method modified by Témez (2003). This method estimates a stream's water flow as function of the total precipitation, the basin area and associated land uses, the time of concentration, and the runoff coefficient (see the Material and Methods chapter). Once the daily water flow had been determined, we estimated the threshold at which the stream's basal flow was surpassed, as a unit of disturbance for the invertebrate communities. We then calculated: (i) the number of days elapsed since the last flood event (defined as the one doubling the basal flow discharge); (ii) the number of flood events; and (iii) the ratio between the maximum and basal flow discharges.

Table III-1 Water quality and physiographic variables used for modeling the environmental component in dbRDA models.

Data set	***scale***	***variable***
Water quality	*riffle-level*	*Conductivity*
		pH
		Saturated Oxygen
		Reactive phosphorus soluble - PRS
		Total organic carbon – TOC
		Canopy shading
		Total suspended solids - TSS
		Ammonium – N-NH$_4$
		Nitrate – N-NO$_3$
Physiography	*Stream-level*	*Elevation*
		Slope
		Temperature - T
		Cross-sectional channel width - W
		Average depth
		Average flow velocity
		Instantaneous discharge - Q
		Max. discharge / Inst. Discharge - Qmax.Q
		number of days after the last spate - nEvents2tQ
		number of spate events - n2tQdays
		Percentage of low mobile substrate (> Boulders)
		Percentage of mid mobile substrate (Pebble-Boulders)
		Percentage of high mobile substrate (< Pebble)

In each stream segment (100 to 200m long), we selected three riffle areas that were representative of the range of substratum types, flow velocities, channel widths and depths, and canopy cover occurring along the stream. Physical and chemical variables were measured during invertebrate sampling (January-February 2017) and on two further occasions (November 2016 and, January-February 2018), that corresponded to high and low water flows, respectively. Instantaneous discharge was estimated in the three riffles by measuring of water depth and flow velocity at 15cm intervals along three cross-sections. At each interval, we also recorded the dominant substrate. Flow velocity was measured with a digital flow meter (SCHILTKNECHT – MiniAir 20). Canopy shading (%) was estimated from vertical photographs using a fisheye lens and subsequent image analysis. Conductivity, pH, oxygen, and temperature were recorded using a HANNA HI98194 water quality meter upon arrival (early morning) and departure (dawn) from the site.

On each occasion, 1 L of water was collected for physico-chemical analyses, filtered through 0.7µm glass fiber filters (Whatman GF/F, Kent, UK) and stored frozen until analysis. In the laboratory, ammonium and nitrate concentrations were determined on a Dionex ICS-5000 ion chromatography system (Dionex Corporation, Sunnyvale, U.S.A.). Reactive phosphorus (PRS) concentrations were determined colorimetrically using the fully automated discrete analyzer Smartchem 140 (AMS Allaince, Frépillon, France). Total suspended solids (TSS) were analyzed by filtering 500ml of water through a pre-weighed GFF and drying the filtrate for 1 hour at 105ºC. The mean and coefficient of variation of all the variables per ecoregion are summarized in Table 0.2 (Chapter I).

Invertebrate sampling

Stream invertebrates were sampled during the dry season (January - February 2017). Invertebrates were collected using a multi-habitat sampling procedure, with 5 Surber (mesh size: 350um; area: 0.09 m^2) samples collected in stream substrata that were selected according to their corresponding habitat coverage. For instance, if a riffle was composed of 60% of boulders, 30% gravel, and 10% cobbles, 3 Surber samples of the first, 1 of the second, and 1 of the third substratum type were collected. The substratum distribution in each riffle was evaluated visually using the Wentworth scale (mm, diameter-based) as a reference (Wentworth, 1922). We only sampled boulders (diameter >250mm) smaller than the sample frame. In six of the 25 streams sampled, only two riffle sections (10 Surber samples) were sampled because of problems with access. Our final data set comprised a total of 343 Surber samples (6 ecoregions, 25 streams, and 69 riffle sections).

In the laboratory, invertebrates were sorted and identified to the genus level, following Trivinho-Strixino & Strixino (1995), Merritt & Cummins (2008), Domínguez & Fernández (2009) and González-Córdoba et al. (2015). Chironomidae and Ephemeroptera were dissected and mounted in Euparal®

following the protocol of Domínguez (2006) and Andersen et al. (2013). The pupae of Chironomidae were mounted to confirm some taxonomical identities (Prat et al. 2014).

Data analyses

We used a four-step approach to: (i) explore the relative contribution of different spatial scales to the overall basin (gamma) diversity; (ii) test for significant differences in environmental and community composition variability among and within ecoregions; (iii) assess the elements of metacommunity structure (EMS, Leibold & Mikkelson, 2002) by matching the observed patterns to an idealized model; and (iv) quantify the relative contribution of environmental conditions, spatial structure and ecoregional identity to explaining metacommunity structure.

i. Additive partitioning of diversity

The additive partitioning of diversity (Whittaker 1972) has been used to test the uniformity of diversity across spatial scales or across patches within a matrix (Lande, 1996). This approach accounts for the compositional differentiation of groups (e.g. patches) by partitioning the regional or gamma (γ) diversity into within- and between- group components, representing the alpha (α) and beta (β) diversities, respectively (Jost, 2007). Thus, if diversity is uniformly distributed, the ratio between the alpha and gamma components will have a value of 1 (Lande, 1996). We used the approach of Crist et al. (2003) to include distinct spatial scales in a hierarchical sampling design, the lower sampling levels being nested within higher levels. By doing this, the gamma (γ) diversity can be expressed as the proportional contributions of diversity due to each level of the hierarchy (Crist et al. 2003). In our hierarchical sampling design, the highest level corresponded to the Orinoco basin, with its gamma diversity partitioned within (α) and among riffles of each stream (β1), among streams of each ecoregion (β2), and among ecoregions within the basin (β3). The sum of α + β1 + β2 + β3 is equivalent to γ.

Additive partitioning can be conducted on the most widely used diversity metrics: species richness, the Shannon entropy index, and the Gini-Simpson index (Lande 1996, Jost 2007). We used only the species richness and the Shannon entropy because these are the two metrics that increase in line with the compositional differentiation in the additive partitioning framework (Jost et al. 2010). Under this framework, compositional differentiation estimates how much diversity is added (by the occurrence of different species and the increasing in their respective abundances) at each hierarchical level. For instance, the additive species richness β3 is a measure of the average number of species from the Orinoco basin (γ) absent from a "sample" taken at the ecoregional level. The expected diversity at each level was calculated 9,999 times by individual based randomization of the community matrix (see Crist et al., 2003 for further details on the formulas and calculations). We performed all the calculations using the 'adipart' function of the VEGAN package (Oksanen *et al.*, 2013).

ii. Environmental and community composition variability at the ecoregional scale

Following Heino et al. (2013), we used both the canonical analysis of principal coordinates (CAP, Anderson & Willis, 2003) and the test for homogeneity of dispersion (PERMDISP, Anderson, 2006) to assess variability in community composition and environmental conditions among and within ecoregions. Community data was pooled at the riffle level (summing up the counts of five Surber samples) in order to match the scale at which environmental variables were measured at each stream. CAP identifies the axes through the multivariate space that best discriminate among *a priori* groups. In our case, we allocated riffles to correct ecoregion types (leave-one-out procedure) and test for among-ecoregion differences in community composition using random permutations. A total of 9,999 permutations were run to test the null hypothesis that ecoregion centroids do not differ. We used the Euclidean distance for standardized environmental variables and the Hellinger distance (Legendre & Gallagher, 2001) with the invertebrate incidence- and abundance-based data. We calculated the Hellinger distance for the incidence- and abundance-based data to maintain comparability with the redundancy analysis. Before analysis, we ran diagnostic tests to guarantee the best discrimination among the ecoregions. We used PERMDISP to estimate the within-ecoregion dissimilarity and to compare among-ecoregion differences in the distance between the observations and the group centroid. The significance of the among-group differences was tested with permutation tests for least-squares residuals. The null hypothesis of no differences the among-ecoregions in terms of environmental and community dissimilarity was tested using a permutation test with 9,999 runs. We performed CAP calculations with the 'CAPdiscrim' function of the BiodiversityR package (Kindt & Kindt, 2018), and PERMIDISP tests with the 'betadisper' function of the VEGAN package (Oksanen *et al.*, 2013).

iii. Elements of metacommunity structure

The analysis of the elements of metacommunity structure (EMS) determines which idealized metacommunity structure (Leibold & Mikkelson, 2002) or quasi-structure (Presley *et al.*, 2010) bests fits the observations obtained from the community matrix. Based on a species-by-site incidence matrix, in which sites are rearranged by similarities in species composition and species are rearranged by similarities in distribution (Leibold & Mikkelson, 2002), EMS analysis assesses the coherence, turnover, and boundary clumping of species distributions, looking for the best-fit model. By ordering the matrix, the sites with similar species composition are located closer to one another, and, in a similar way, the species with similar occurrence among the sites are closer to one another (Gauch, 1982). Coherence is assessed by counting the number of embedded absences in the ordinated matrix and by comparing that value to a null distribution of embedded absences (i.e. a checkerboard distribution with gaps in the species range). If the number of absences is significantly less than that occurring at random, then the turnover is evaluated. Turnover is assessed by counting the number of species replacements along the matrix and comparing that value to the null distribution. A significant negative turnover suggests a nested distribution, whereas a significant positive turnover suggests an evenly spaced, Clementsian or Gleasonian structure. Finally, boundary clumping was evaluated using Morisita's dispersion index and subsequently tested against expected distributions using a chi-squared test. Values significantly greater than 1 indicate clumped range boundaries, values significantly less than 1 indicate hyperdispersed range boundaries, while equaling 1 indicate randomly distributed range boundaries.

We used the 'Metacommunity' function of the METACOM package (Dallas, 2014) to perform all EMS analyses. EMS can be sensitive to the grain size (Marcilio-Silva *et al.*, 2017); therefore, we created two datasets with different grain sizes (corresponding to the riffle and stream levels). The selected null model constrained species the richness of each riffle or stream to equal empirical richness, with equiprobable occurrences for each species. It corresponded to the 'r1' option in the function. This null model has a more desirable combination of type I and type II error properties than Random 0 or Random 4 (Gotelli & Graves, 1996; Gotelli, 2000), and has been successfully used in analyzing coherence (e.g. Presley & Willig, 2010). All null models were based on 9,999 permutations.

iv. dbRDA and variance partitioning.

The relative contribution of the environmental variables, spatial structure and ecoregional identity on metacommunity composition was quantified with the distance-based redundancy analysis ('dbRDA', Legendre & Anderson, 1999), followed by variation partitioning analysis (Borcard *et al.*, 1992). Before analysis, community data was pooled at the riffle level (summing up the counts of five Surber samples) in order to match the scale at which environmental variables were measured at each stream. Community data was Hellinger-transformed and environmental variables were standardized to have a mean 0 and variance 1, as recommended by Legendre & Gallagher (2001) and Legendre and Legendre (2012). The significance of all the models was tested using 9,999 permutations. dbRDA and variance partitioning analysis were performed using the “rda” and “varpart” functions of the VEGAN package (Oksanen *et al.*, 2013), respectively.

We aimed to identify the influence of local environmental variables by using two datasets, one of which included nine water quality variables (including canopy cover) while the other included 13 physiographic variables (Table III-1). The most significant variables were retained through a forward selection procedure using the “ordiR2step” function of the VEGAN package (Oksanen *et al.*, 2013). Moran’s eigenvector maps (MEMs) were used to model the spatial structure of the streams within each ecoregion (Legendre & Legendre, 2012). Following the method of Declerck et al. (2011), the resulting eigenvector maps were arranged in blocks in a staggered matrix, each block corresponding to a distinct ecoregion. MEMs model the spatial dependence in a set of locations as an orthogonal combination of MEM variables derived from geographic coordinates, the number of ecoregions (“blocks”), and the stream within each ecoregion (Declerck *et al.*, 2011). We used the “create.dbMEM.model” function of the “adespatial” package (Dray *et al.*,

2018) to create the MEM model. Finally, to identify possible biogeographic effects, we created a matrix of six dummy variables representing the respective ecoregions.

All analyses and graphical outputs were performed using the R statistical software version 3.4.1 (R Core Team, 2018).

RESULTS

We identified a total of 69,206 individuals from 133 genera belonging to 52 families and 14 orders. Rarefaction analysis indicated that the ecoregions were well sampled, as the curves approached the asymptote (Figure III-2). The most diverse ecoregion in terms of taxon richness was the high-Andean (74 genera among streams and 43 genera on average per stream), followed by the high plains (73 genera among streams and 33 genera on average per stream), the piedmont (65 genera among streams and 24 genera on average per stream), the alluvial fans (63 genera among streams and 25 genera on average per stream), the Guiana shield (62 genera among streams and 33 genera on average per stream) and finally, the Páramo (47 genera among streams and 27 genera on average per stream).

IndVal analysis revealed 67 genera as potential indicators of the ecoregions: 8 for the Páramo streams, 8 for the high-Andean streams, 5 for the Piedmont streams, 6 for the Alluvial fan streams, 23 for the highPlain streams, and 17 for the Guiana shield streams. Additionally, 28 and 7 genera were potential indicators of combinations of two or three ecoregions, respectively (Supplementary Material).

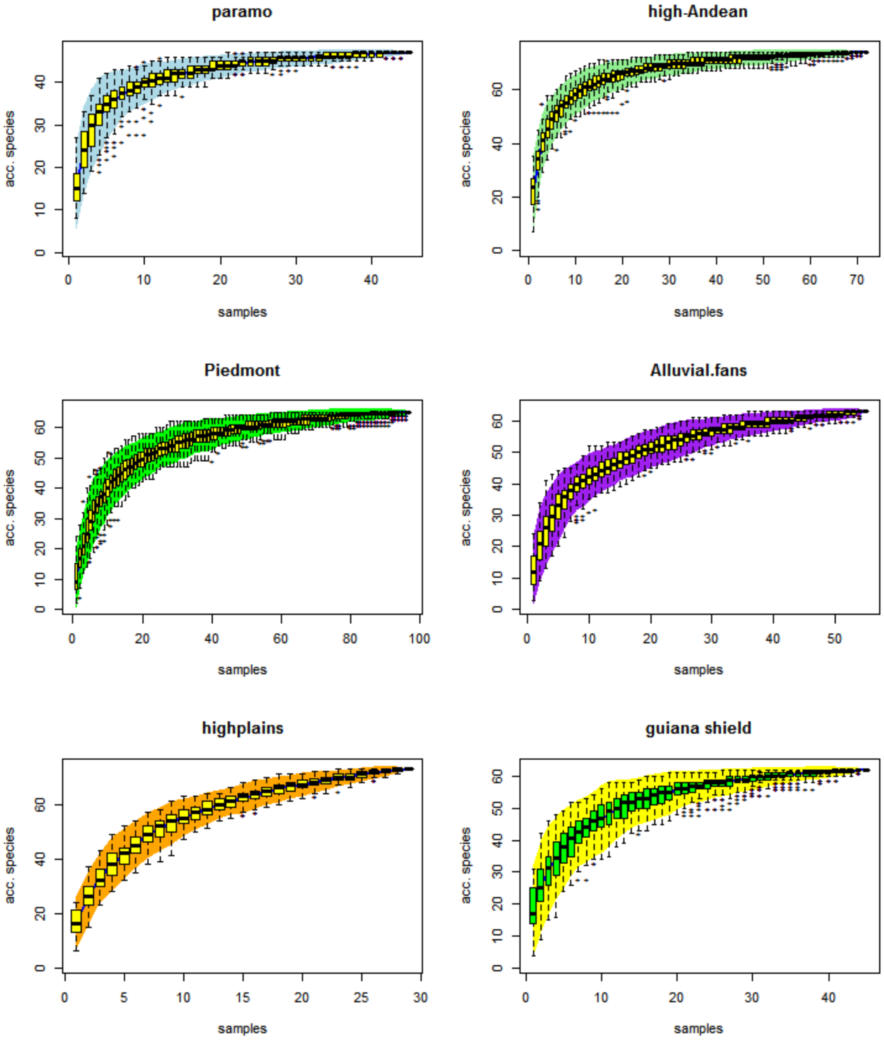


Figure III-2 Species accumulation curves per ecoregions. Data of Surber samples were pooled by Ecoregion. We used the "random" accumulation method, which adds sites in random order.

i. Diversity partitioning

Additive partition of diversity based on taxon richness and Shannon's entropy index showed that samples at each riffle (α - *within riffles*) presented lower diversity, on average, than that expected with the null hypothesis (p-value < 0.001, Table III-2), indicating that the taxa are not homogeneously distributed across the spatial scales. Instead, the significantly greater contribution of the upper hierarchical levels (β2 - among streams, β3 - among ecoregions) indicated that the non-random clusters of taxa could be separated at larger scales. However, the relative contribution of streams and ecoregions diminished as the more abundant species gained greater weight in Shannon's index estimations. This decreasing trend suggested that compositional differentiation at the upper levels was mainly due to the presence of unique or low-abundant species in streams and ecoregions.

Table III-2 Additive partitioning of invertebrate diversity in the Orinoco basin, Colombia. The expected values for both taxon richness and Shannon's index correspond to the average values of the null distribution.

	ORINOCO BASIN (n=25)			
Taxa richness - S	**Observed**	**Expected**	**(%)**	**p-value**
α - within riffles	30.33	100.34	22.8	<0.001
β- among riffles	8.22	15.08	6.2	<0.001
β- among streams	25.44	12.99	19.1	<0.001
β- among ecoregions	69	4.61	51.9	<0.001
γ	133			
Shannon Index -H'				
α- within riffles	2.26	3.33	67.1	<0.001
β- among riffles	0.07	0.02	2.1	<0.001
β- among streams	0.49	0.01	14.5	<0.001
β- among ecoregions	0.55	0.004	16.3	<0.001
γ	3.37			

ii. Environmental and community composition variability at the ecoregional scale

We found that invertebrate community composition (Figure III-3A-B), water quality (Figure III-3C) and stream physiography differed significantly among the ecoregions (Figure III-3D); however, classification success rates varied appreciably depending on the dataset. Water quality variables had the lowest discriminant power (66.2%, p-value = 0.001), followed by physiography (91.2%, p-value = 0.001), invertebrate abundance (91.2%, p-value = 0.001), and invertebrate incidence (95.6%, p-value = 0.001). In all the datasets, the Piedmont and alluvial fan ecoregions had the lowest percentages of correct classification (Table III-3). When the invertebrate data were considered, these were the only ecoregions with a percentage of correct classification below 100%.

The PERMDISP tests indicated high variability in community composition and environmental conditions within the ecoregions (Table III-4). Environmental and community dissimilarities differed significantly among the ecoregions. In terms of water quality, streams in the Piedmont and alluvial fan ecoregions were the most variable. Streams from the alluvial fans were the most variable in terms of physiography, while streams in the other ecoregions showed similar variability. Regarding community composition, the within-ecoregion variability was similar for both types of data (incidence and abundance). In both datasets, streams in the alluvial fans exhibited the highest within-ecoregion variability.

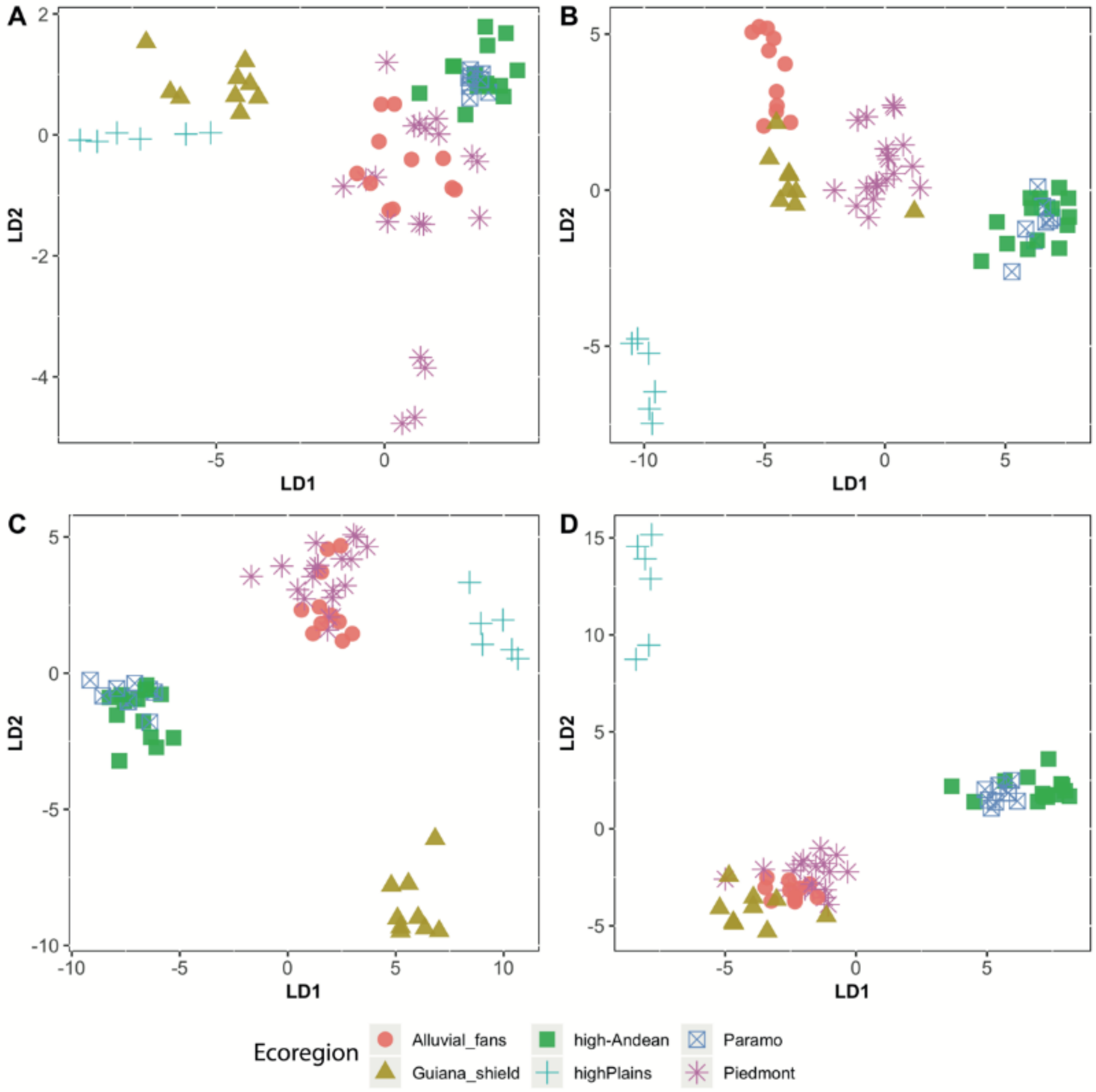


Figure III-3 Ordination plots of the canonical analysis of principal coordinates (CAP) for (A) water quality data, (B) physiographic variables data, (C) invertebrate incidence-based data and (D) abundance-based data.

Table III-3 Percentage of correct classification and mean distance from group centroids for the environmental and community composition data obtained in CAP and PERMDISP analyses, respectively. The F- and p-values were obtained from tests for differences in multivariate dispersions among the streams (999 permutations).

	Water quality		**Physiography**		**Invertebrate incidence data**		**Invertebrate abundance data**	
Ecoregion	% of correct classification	Mean distance from centroid	% of correct classification	Mean distance from centroid	% of correct classification	Mean distance from centroid	% of correct classification	Mean distance from centroid
Páramo	56%	0.6994	67%	1.846	100%	0.5036	100%	0.5415
High-Andean	57%	1.5641	93%	1.281	100%	0.4727	100%	0.4839
Piedmont	74%	2.6298	100%	2.144	89%	0.6535	89%	0.6379
Alluvial fans	64%	2.0423	100%	3.502	64%	0.6145	91%	0.5218
high-Plains	67%	0.8624	100%	1.545	100%	0.6044	100%	0.5678
Guiana shield	78%	1.3404	78%	2.123	100%	0.5116	100%	0.4956
	F	2.9186	F	5.7144	F	16.227	F	3.7336
	p-value	0.018	p-value	0.0002	p-value	0.001	p-value	0.006

iii. Elements of metacommunity structure

The invertebrate metacommunity of the Colombian Orinoco basin exhibited a Clementsian structure (Table III-4). Notably, the EMS analysis results were not affected by data taken at the riffle level or weighted at the stream level. Considering the lower contribution of riffles in the additive diversity partitioning, this result suggests that differences in community structure were negligible within streams. Overall, the invertebrate metacommunity was characterized by a positive coherence, positive turnover, and smaller values of boundary clumping. The positive coherence suggests that taxa are exposed to similar environmental gradients, while the positive turnover indicates that community composition along the gradient changes more than would be expected for equiprobable distributions. Finally, the boundary clumping pattern indicates that the compositional changes occur by clusters of taxa.

Table III-4 Elements of metacommunity structure (EMS) analysis of the Orinoco basin using data at the riffle- and stream-level. These metrics indicated a Clementsian metacommunity structure.

	Coherence				Species turnover				Boundary clumping		
	EAbs	*p*	Mean	SD	Rep	*P*	Mean	SD	Morisita's Index	*p*	Interpretation
riffle-level	3552	<0.001	5778	125	971995	<0.001	577537	41615	1.86	<0.001	*Clementsian*
stream-level	1161	<0.001	1910	71	141731	<0.001	89381	7114	1.48	<0.001	*Clementsian*

iv. dbRDA and variance partitioning

The variance partitioning revealed that the ecoregional identity and environmental variables were significant predictors of community structure (Figure III-4), whereas the effects of spatial structure were not significant. The full models, including water quality, physiographic descriptors and ecoregions, accounted for 48% and 59% of the variation in community composition when using incidence- and abundance-based data, respectively. When taxon abundances were considered, the overall performance of the model increased, but did not affect all the environmental descriptors retained by the forward selection procedure. For both types of data, the most parsimonious models for water quality descriptors included pH, oxygen saturation, TSS, conductivity, PRS, and canopy cover. The variables included in the parsimonious model for physiography using the incidence-based data were elevation, slope, substrata with high mobility, channel width and depth, the number of events doubling the basal flow, and the number of days elapsed since the last flood event. In the case of the abundance-based data, stream depth was not selected by the forward selection procedure, while flow velocity was selected as a significant variable.

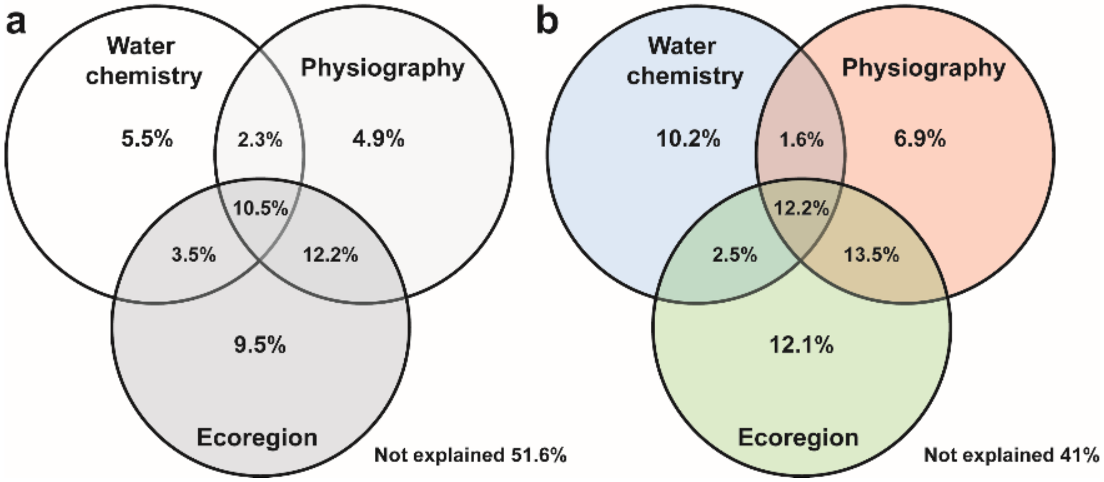


Figure III-4 Venn diagram showing the variation partitioning results for aquatic invertebrates based on (a) incidence and (b) abundance. The values indicate the proportion of variance (R^2-adjusted) explained by each component and its respective interactions. The spatial component was not significant and was not included in the representation.

In both models, the greatest amount of variance in the community structure was explained by the ecoregion and its joint effects with physiographic features (Figure III-4). When abundances were included, the variance due to water quality increased, while that for the other components decreased. This suggested that water quality had a stronger effect on the distributions of taxon abundance, but its influence on taxon occurrence was weak and depended on physiographic and ecoregional features. These two components exhibited a large amount of shared variance (incidence-based data: 25.1%; abundance-based data: 22.9%), which was even higher than the variance explained by the components individually in both models.

DISCUSSION

Our study used small grain-size samples from across a broad biogeographical area to identify the main drivers of metacommunity structure in the Orinoco basin. At the basin scale, metacommunity structure was shaped by the occurrence of different pools of taxa rather than by spatial structure or the local environment. We found evidence that the distribution of the species pools was substantially shaped by the action of dispersal-limiting processes within historical and ecological timeframes. Our results suggest that events from the Tertiary and Quaternary periods constrained dispersal and shaped discrete pools of invertebrate taxa, while some physiographic features of each ecoregion exerted some control on invertebrate dispersal within the ecological timeframe. Consequently, these results have important implications for the study and conservation of invertebrate metacommunities in a region with high biodiversity (Myers *et al.*, 2000).

Diversity, distribution and metacommunity structure at the basin scale

The diversity partitioning analysis indicated that the ecoregional level is contributing the greatest to the gamma diversity of the Orinoco basin. CAP and PERMDISP analyses suggested that the dissimilarity in taxon composition is greater among than within ecoregions. These findings demonstrate a patchy distribution within the basin, with each ecoregion having a distinct pool of invertebrate taxa. Previous studies in tropical and subtropical basins found distributions similarly constrained by ecoregion, both in diatom (Benito *et al.*, 2018) and invertebrate (Dedieu *et al.*, 2015; Pero *et al.*, 2019) communities. However, finding this pattern was unexpected in our study given that all the tributaries belong to a single river network.

Previous studies in similar settings have shown that the overall connectedness of the network, as well as the position of each tributary and its particular environmental features, are major determinants of metacommunity structure (e.g. Brown & Swan, 2010; Finn et al., 2011; Altermatt, 2013; Grönroos et al., 2013). However, in our study, the lower or no explanatory power of the environment and spatial structure in the dbRDA models and subsequent variance partitioning indicates that other factors may act as the underlying drivers. RDA models usually do not explain a considerable amount of the total variance in community composition, the variance explained by 'pure' spatial or environmental components rarely exceeds the 15% (to compare with different assemblages, see Heino, Grönroos, Soininen, Virtanen, & Muotka, 2012). Hence, although the variance was small (R^2_{adj} = 0.1-0.12, percentage of model's explained variance between 19 and 21%, Figure III-4), the comparatively higher variance explained by the pure ecoregional component points to ecoregional features as major drivers of community composition. Besides, the higher variance shared with the other components indicates the possibility of a joint effect of ecoregional features and stream environmental conditions on community composition (expanded below).

Taken collectively, the abovementioned findings suggest that past historical events shaped taxon distribution in the way they shaped the distribution and physiographic features of the ecoregional types across the Orinoco basin (e.g. Goosen, 1971; van der Hammen, 1974; Hughes & Eastwood, 2006; Hoorn et al., 2010). Several studies have indicated that a concordance between species pools and ecoregion (or biogeographic region) distributions provides indirect evidence of the effects of historical events on present-day biodiversity patterns (Presley & Willig, 2010; Brasil *et al.*, 2017; Hazzi *et al.*, 2018).

Our EMS analysis results further supported the role of historical events as major drivers. We identified a metacommunity resembling a Clementsian idealized structure (Leibold & Mikkelson, 2002), which suggests that the turnover between the ecoregions is due to changes in clumped sets of associated taxa rather than the responses of individual species. Although the Clementsian idealized structure is rather common in the freshwater realm (Heino *et al.*, 2015b), we found no evidence to link the occurrence of this structure with some of the underlying drivers described elsewhere, such as an occurrence a patchy distribution of resources (Willig *et al.*, 2011), habitat types (Presley *et al.*, 2012), and environmental conditions (Tonkin *et al.*, 2015) among the landscapes. Instead, our findings are analogous to those described by Presley and Willig (2010) and Brasil et al. (2017) for Caribbean bat and Amazonian invertebrate communities, respectively. Those studies found that Clementsian idealized structures may stem from the area of origin and the associated specialization of species. Similar clumped patterns in species distributions have been already identified in the tropical Andes and Amazonia (e.g. Hubert & Renno, 2006; Hughes & Eastwood, 2006; Hoorn et al., 2010; Hazzi et al., 2018), with most of these attributable to allopatric speciation by vicariance. For instance, Hughes and Eastwood (2006) found that the clumped distribution of *Lupinus* plants was driven by ecological factors afforded by the emergence of island-like habitats after the Andean uplift. However, further understanding of the historical origin and evolution of the clumped distribution observed in the basin requires defining the phylogenetic relationships between the invertebrate taxa.

The higher turnover among the ecoregions may be attributed not only to historical and evolutionary processes, but also to processes limiting dispersal within an ecological timeframe. Contemporary features of ecoregions, such as climate, landscape topology, and stream forms, are a result of past historical events (van der Hammen, 1958; Goosen, 1971). All these features can be viewed as physical boundaries to species dispersal in the present-day, helping to explain the considerable amount of variance explained by stream physiography and ecoregional identity in RDA models. Stream discharge, temperature and substrata are examples of the variables conditioned by ecoregion that can influence species dispersal. In the Andean streams, for instance, temperature and the frequency of spate episodes may shape larger scale trends of invertebrate diversity and distribution by limiting species dispersal either physically or physiologically (Gill *et al.*, 2016). Future studies could investigate how species (or genus-level) dispersal ability modulates the effect of the different ecoregional constraints on species occurrence and distribution among river networks. This knowledge will be essential for unraveling the role of evolutionary and ecological processes in shaping the structure of metacommunities in the present day.

The variance partitioning and CAP analyses revealed that the effect of water quality on metacommunity structure is not influenced by the ecoregional context, even when geological differences between the ecoregions are considered. Nevertheless, water quality could be one of the strongest environmental filters explaining beta diversity across streams of some ecoregions. In Piedmont or the alluvial fans, for instance, we observed large variability in water conductivity, as well as in nitrate and total organic carbon (TOC) concentrations. Conductivity may be a strong environmental filter in streams (Cañedo-Argüelles *et al.*, 2012), determining invertebrate occurrence as well as the higher beta diversity observed in streams of these ecoregions. Future studies, including assessing a larger number of streams per ecoregion, could help to determine the role of water quality in the environmental filtering at the ecoregional scale.

Implications for metacommunity ecology in riverine ecosystems

The understanding of metacommunity dynamics in riverine ecosystems has been greatly enhanced by considering rivers as dendritic networks (Tonkin et al., 2018). As mentioned above, several studies have demonstrated that the branching organization of river networks can strongly affect metacommunity dynamics by regulating dispersal within the river network (Brown & Swan, 2010; Altermatt *et al.*, 2013). However, as far as we know, few studies have addressed how the occurrence of several pools of species can modulate the control exerted by the riverine network in a basin (e.g. Heino et al., 2015b; Tonkin et al., 2015; Brasil et al., 2017).

Our study indicated that historical events have an important role in the assembly of metacommunities at the basin scale. Particularly, we observed that the occurrence of different regional pools can override the effects of other factors previously described as determinants of metacommunity structure (e.g. local environment or the spatial structure of tributaries). Similar to previous findings in algal communities (Benito et al. 2018), neotropical species distribution seemed to be significantly restricted to each ecoregion or biogeographic district. Hence, the communities that we studied do not constitute a metacommunity in an ecological timeframe or in the basin extent. Conversely, the Orinoco basin can be viewed as a mosaic of functional metacommunity units, with the structure of each unit varying according to the taxa belonging to each species pool and the environmental context at each ecoregion. Brasil et al. (2017) observed a similar context-dependent effect of historical events on the structure of damselfly communities in near-pristine basins from the Amazonas region. However, further studies are needed to determine if these results can be extended to all Tropical basins.

Implications for conservation

One common aim of ecologists and conservationists is to determine how diversity varies over space and time (Jost *et al.*, 2010). In riverine ecosystems, the basin has been used as an appropriate scale to study the causes and consequences of biological diversity (Brown & Swan, 2010; Grönroos et al., 2013; Cañedo-Argüelles et al., 2015). However, as indicated by our results, using the basin may not be appropriate in highly heterogeneous basins such as the Orinoco basin. Before establishing a conservation unit in tropical basins, a first step should be determining the extent of the pool of species, and their relative influence on metacommunity structure. This agrees with the observation of Poiani et al., (2000), who reported that the units for conservation should be delineated according to the context of each basin or sub-basin. Our results indicate that such delineation should consider the type and extent of the different ecoregions in the basin, given that Ecoregions harbor distinct pools of species because of historical contingencies. Therefore, they can be considered as functional metacommunity units for assessing biodiversity patterns and designing effective conservation actions.

IV. Ecoregional characteristics drive the distribution patterns of Neotropical stream diatoms

ABSTRACT

We assessed the relative influence of ecoregional features in explaining diatom distribution in the Orinoco river basin. Ecoregions in the Colombian Orinoco can be seen as imprints of the evolutionary history of the basin, for their current biodiversity and physiographic features are the result of the geological and climatic shifts that have occurred since the Tertiary. Thus, they represent an ideal testing ground for studying the interplay between ecological and evolutionary processes shaping diversity patterns of microorganisms, such as diatoms, in the present day. To attempt to study this interplay, we compared diatom community composition variance within and among seven ecoregions and assessed the explanatory power of environmental, spatial and historical drivers**.** This was done by a combination of correlation analyses, multivariate methods and constrained ordinations. We also deconstructed the whole community dataset into ecological guilds (low and high-profile, and motile) to explore their individual response to the contemporary and historical drivers. Taken together, these analyses indicated that contemporary constraints to species occurrence and dispersal as well as the legacies of historical events can provide an explanation for the contemporary distribution of diatoms in the Colombian Orinoco. Particularly, we provided evidence showing that both historical legacies and contemporary environmental conditions (temperature, pH, and phosphorus concentration) are interacting to determine diatoms' distribution. Our results suggest the need to consider ecoregional gradients for unraveling the mechanisms shaping tropical diversity as well as for designing conservation plans.

-o-

An edited version of this chapter has been accepted in the *journal of Phycology*. Please, cite the respective paper as:

- **González-Trujillo, J. D., Donato-Rondon J. Ch, Pedraza-Garzón, E., Sabater, S. Journal of Phycology (2020).**

INTRODUCTION

In the Neotropics, the available literature indicates that microorganism distribution is mostly limited by altitude-driven shifts in the local environment (Nottingham *et al.*, 2018), such as it was formerly described by Humboldt and Bonpland in plant communities (von Humboldt & Bonpland, 1807). However, a recent study brought to light the existence of clusters or ‘provinces’ of freshwater lakes with unique combinations of diatom species (Benito *et al.*, 2018). Such a pattern hinted that historical events could have constrained species’ dispersal in an evolutionary timeframe, particularly through isolations caused by the orogenesis of the Andean uplifts. Similar distribution patterns associated to these historical events have previously been described in reptile, fish and plant communities (Hughes & Eastwood, 2006; Hoorn *et al.*, 2010). Nevertheless, further evidence is needed to determine if there is a general pattern in Neotropical freshwaters, particularly in river networks, where the greater connectivity could erase the clustered pattern because of the greater propensity to disperse in these landscapes.

The effect of dispersal limitation on species distribution can occur in ecological and evolutionary timeframes. When dispersal constraints are weak, or occur within an ecological timeframe, the distributional patterns may reflect both the spatially-structured local environment and the structure of dispersal pathways. Conversely, when strong and longstanding constraints take place, species dispersal may be impeded within an evolutionary timeframe, which could create distributional patterns that are footprints of the historical contingencies (Brown, 1995). Stream networks are an example of structural constraints that can shape the distribution of microorganisms within ecological timeframes (Liu *et al.*, 2013; Widder *et al.*, 2014), but which are themselves also be affected by historical contingencies. At the same time, streams are also affected by contemporary environmental factors (e.g. changing hydrology, water chemistry, altered temperature regimes), which comprise additional potential modulators of microorganisms' distribution (Martiny *et al.*, 2006; Soininen, 2012).

Few studies have yet addressed the relative influence of all of these constraints on diatom distributional patterns (but see Heino et al., 2017). In this regard, Neotropical streams constitute an appropriate setting to disentangle the interplay between contemporary and historical factors. Firstly, because compared to Europe and other areas with large and long-lasting anthropogenic impact, the pressures received by this region are still relatively small (though widening Sabater *et al.*, 2017). Secondly, because Neotropical basins have complex evolutionary histories that have been determined by major geological and climatic events, of which the Andean uplift played a particularly strong role shaping the current distribution of the ecoregions (Hoorn *et al.*, 2010; Hazzi *et al.*, 2018). Thirdly, this diverse ecoregional mosaic is drained by several stream networks that are interconnected within large basins, and therefore, local communities that had once been isolated by past events may currently be re-connected through dispersal, leading to the onset of metacommunity dynamics within an ecological timeframe.

In order to evaluate the major drivers of stream diatom diversity, we compared the variance of community composition among and within ecoregions and assessed the explanatory power of habitat (stream)- and ecoregion-level variables in diatoms' compositional patterns. Ecoregions in the Colombian Orinoco can be seen as imprints of the evolutionary history of the basin, for their current biodiversity and physiographic features are the result of the geological and climatic shifts that have occurred since the Tertiary (Goosen, 1971; van der Hammen, 1974; Hughes & Eastwood, 2006; Hoorn *et al.*, 2010): for example, events such as Andean uplifts and glacier retreats shaped unique combinations of river forms and riparian ecosystems at the ecoregional scale (Rull, 2008; Hoorn *et al.*, 2010). Therefore, by including the ecoregion type and characteristics, we are able to account for the potential legacies of past geological and climatic events.

We used constrained and unconstrained ordination methods to evaluate the relative influence of contemporary constraints (environmental- and dispersal-related) and possible legacies of historical events on the assembly process of diatom communities. Our predictions about the interplay between these drivers were based on the framework proposed by Martiny et al. (2006). According to this framework, a random assembly is expected if we find a non-spatially-structured pattern and lower explanatory power in the three sets of variables (Fig. IV-1A). Constraints to species dispersal are expected be the main underlying factors if diatom communities are more similar between neighboring streams or within ecoregions, and if diatom communities' species compositions are significantly explained by the spatial or ecoregional structure (Fig. IV-1B). Environmental filtering is expected to occur if diatom communities' species compositions are mostly explained by the environment dataset (Fig. IV-1C). Finally, it is also recognized that distribution patterns could result from a combination of environmental filtering and dispersal-driven effects. Under this scenario, communities' species compositions are expected to be explained by an interaction between the environmental and spatial and/or ecoregional components (Fig. IV-1D).

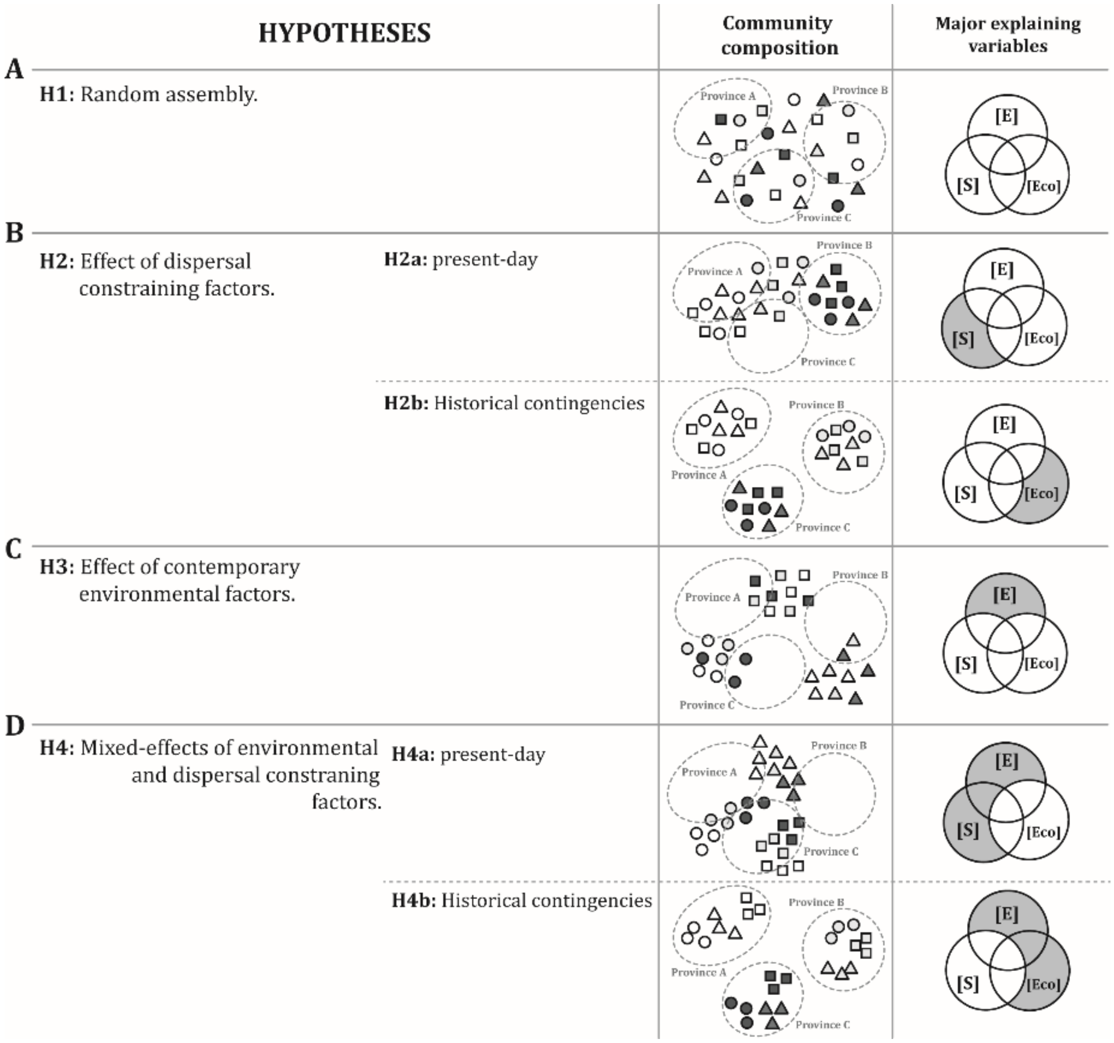


Figure IV- 1 Conceptual scheme representing the framework stated by Martiny et al. (2006). For each hypothesis the main drivers are represented, as well as the expected patterns in community similarity among- and within-ecoregions, and the set of explanatory variables that would account for the major part of the constrained variance in RDA models (shaded in grey). The shape of the points represents the environmental similarity (e.g. streams represented by triangles are supposed to be more similar). The color of the points represents the spatial proximity between streams; streams of similar color are supposed to be geographically closer.

MATERIAL AND METHODS

Study sites

The Orinoco basin is the third largest basin in South America, and covers an area of about 990,000 km^2 in Venezuela and eastern part of Colombia (Romero Ruíz *et al.*, 2004). The complex geological and climatic history of the basin has created a heterogeneous landscape containing a broad range of ecosystems (Romero Ruíz *et al.*, 2004). An intricate network of rivers and tributaries drains down three large forms of relief that grow in altitude from 100 to 3500m a.s.l.: the Orinoco's ancient massifs and shields, its recently raised ridges, and its tectonic depressions and accumulation plains (Stallard, 1985). These geological features are the main determinants of the streams' geomorphological chemical characteristics.

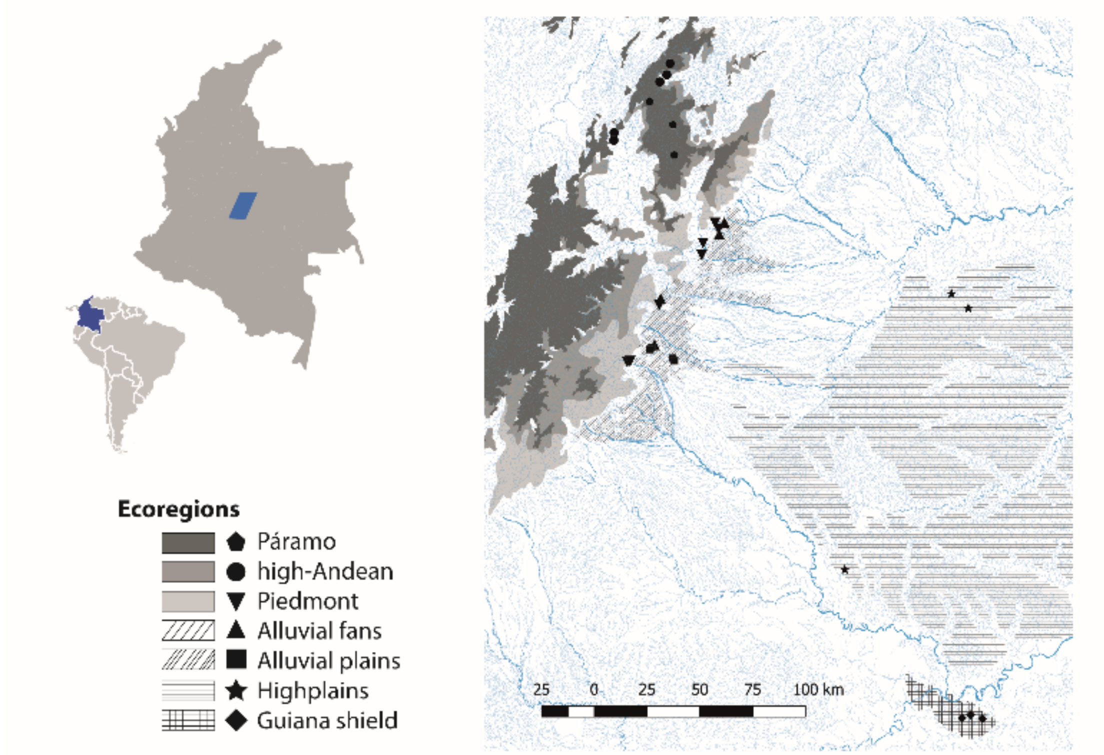


Figure IV-2 Geographical extent of ecoregions and position of the sampled streams in the Orinoco basin.

Our sampling was performed at 26 stream segments within an area of about 40,000km^2, in the Colombian Orinoco (geographical center N 3º34'- W 73º31', Figure IV-2). The area encompasses an elevation gradient from 300 to 3400 m a.s.l. and includes a heterogeneous assembly of ecoregions and landscapes. From three to five pristine or near-pristine streams were selected and sampled in seven of these ecoregions, which were *a priori* identified based on their physiographic features (i.e. the age of major landforms, altitude, vegetation and climate, channel shape and slope, and the main streambed substratum; Table I-1).

The larger number of streams corresponded to ecoregions on which variability was expected to be higher (González-Trujillo *personal observations*). The ecoregions included in the study were (in order of decreasing altitude): (i) The Páramo; endemic to the Neotropics, it is the most recent ecoregion in the Andean hills. It is globally recognized as a diversity hotspot as well as a "natural water fabric" (Buytaert *et al.*, 2006). (ii) The high-Andean ecoregion; located immediately below the Páramo, is a belt of cloud forest abundant in fertile zones, with comparatively high precipitation and low temperatures (iii) The Piedmont ecoregion; reaching down the Andean foothills, it encompasses a belt of rainforest growing on steep slope zones with a warm and humid climate. (iv) The Alluvial fans ecoregion; where the streams leave the Andean hills and join the greater mainstreams, such as the Meta or Guaviare rivers.

(v-vii) The remaining three ecoregions are distributed throughout the lowlands (locally known as "Llanos Orientales"). During the early Holocene, the Meta fall divided the lowlands into an elevated zone to the east, which is characterized by a savanna-like climate and vegetation (the high Plains ecoregion), and a lower zone to the west that includes the alluvial plains and the Guiana shield ecoregions. These two ecoregions are elevated zones with a comparatively flat morphology and distinct vegetation. They were both shaped during the Andean uplifts, although the former is more ancient, and has a geology that is largely influenced by the Guiana shield reliefs.

Diatom sampling

We sampled diatom communities at three different riffle sections of each stream segment, each section spanning from 20 to 60m long. The selected riffles were representative of the stream's range of substrata types, flow velocities, channel widths and depths, and canopy cover. At each riffle section, we collected 8 cm^2 of surface, brush-scraped algal material from 30 boulders and cobbles. In the case of streams from Guiana shield and high-Plains, where boulders and cobbles were scarce, we also took samples from bedrock, pebbles and sand. Algal material was pooled by riffle section (= 3 samples per stream segment) and subsequently preserved in a *Transeau* solution. In the laboratory, the organic material from samples was cleaned using hydrogen peroxide. Clean diatom frustules were mounted on permanent slides using a Naphrax® medium, the slides were then observed under a 1000x light microscope and identified at the finest level possible using specialized monographs (Krammer & Lange-Bertalot, 1986, 1991; Metzeltin & Lange-Bertalot, 2007; Bellinger & Sigee, 2015). At least 400 valves were counted in each slide, and counts processed as relative abundances.

We identified a total of 297 diatom taxa (Table S.IV-1). In several cases, the identity of the taxa did not match the monograph descriptions (presumably new species); these were classified as *affine to* (aff.) and were recorded as a separate taxonomic group. To alleviate taxonomic bias between samples, we assembled the 297 morphospecies into 72 taxonomic groups (Supplementary Material). Finally, we assigned each identified taxa to their corresponding ecological guild (low-profile, high-profile, motile) (Passy, 2007; Rimet & Bouchez, 2011).

Environmental characterization

Each stream segment was characterized by its hydrology, substratum type and distribution, physical water conditions, and water chemistry (see the introductory chapter). Long-term hydrological variables were estimated following the modified rational method (Témez, 2003). The variables calculated were i) the number of days elapsed after the last flood event (defined as twice the basal flow discharge), ii) the number of flood events ("nEvents2t"), and iii) the ratio between the maximum and basal flow discharges ("Q.max/Q"). These variables were derived from the daily discharge estimation, which was a function of the total precipitation, the basin area and associated land uses, the concentration time, and the runoff coefficient (see the Chapter II for an extended explanation on the computation).

Physical and chemical variables were measured before (November 2016), during (January-February 2017) and after (January-February 2018) the sampling campaign. Samplings were performed during the low water period, as diversity is maximized during this period. Instantaneous water flow was estimated in the three riffles using water depth and flow velocity, which were measured by means of three cross sections with measurement intervals of 15 cm. Flow velocity was measured using a digital flow meter (Schiltknecht – MiniAir 20). Conductivity, pH, oxygen, and temperature were recorded using a Hanna HI98194 water quality meter. Canopy shading (%) was estimated analyzing a series of upward photographs taken using a fisheye lens. The luminosity and white balance of each photo were manually adjusted to color leaves in black and sky in white. The proportion between black and white pixels was used to estimate the percentage of canopy shading. Finally, considering the high heterogeneity among ecoregions in terms of substratum type, we reclassified them as low- (bedrock and boulders), mid- (cobbles and pebbles) and high mobility (gravel and sand) substratum before including them in RDA models.

One liter of water was collected, filtered through 0.7µm glass fiber filters (Whatman GF/F, Kent, UK) stored frozen to be used for chemical analyses. Ammonium and nitrate concentrations of were determined in a Dionex ICS-5000 ion chromatograph (Dionex Corporation, Sunnyvale, U.S.A.). PRS concentration was determined colorimetrically using a fully automated discrete analyzer Alliance Instruments Smartchem 140 (AMS, Frépillon, France). The total suspended solids (TSS) were determined by filtering 500ml of water in a pre-weighted GFF filter, drying the filter for 1 hour at 105ºC, and weighing the resulting contents. The environmental variables' means and coefficient of variation in each ecoregion are summarized in Table I-2 (see Chapter I).

Data analysis

We used a Hellinger-transformed community dataset for all the analyses, and standardized all environmental variables to mean=0 and variance=1, as recommended by Legendre and Legendre (2012). To avoid the possible influence of singletons, we summed up sample counts and used a matrix of relative abundances per stream in the subsequent analyses (abundance-data matrix of 26 rows/streams and 72 columns/taxonomical entities). We calculated beta diversity using the Hellinger distance for abundance-based data to maintain comparability with the redundancy analysis. Additionally, we performed all analyses for the whole community dataset, as well as in the separated datasets of the low-profile, high-profile, and motile guilds. All analyses and graphical outputs were done using the R Statistical software v 3.5.1 (R Core Team, 2018) and the "ggplot2" package (Wickham, 2016).

i. Elevation-driven diversity and local environment patterns

In mountain systems, changes in species richness and community composition are generally linked to elevation, following a monotonical function (i.e. Nottingham et al. 2018). We therefore tested the existing monotonic relationships between elevation and diatom alpha and beta diversity, as well as between elevation and water chemistry and physiographical characteristics. We used the Spearman´s rank tests for taxa richness, and the Shannon index (alpha diversity) and Mantel tests (9,999 permutations) for environmental and community dissimilarity (beta diversity).

ii. Ecoregional diversity patterns

We used two analyses to evaluate community composition dissimilarity among and within ecoregions. First, we performed a non-Metrical Multidimensional scaling (nMDS) ordination, followed by an ANOSIM analysis (Clarke, 1993) to explore differences in diatom communities between ecoregions. Secondly, we used a canonical analysis of principal coordinates (CAP, Anderson & Willis, 2003) to test for differences among ecoregions. The CAP test finds axes along the multivariate space that best discriminate *a priori* groups. In our case, stream ecoregion types were tested in order to find among-ecoregion differences in community composition, using 9,999 permutations. Finally, we used an analysis of indicator species (IndVal, Cáceres and Legendre 2009) to look for the preferences of species in each ecoregion. The IndVal analysis is based on the concepts of *specificity* (highest values when the species is present in only one ecoregion) and *fidelity* (highest when the species is present in all streams of a given ecoregion). A high indicator value, which ranges from 0 to 1, is obtained by the combination of high specificity and fidelity. We performed IndVal analysis using the function “multipatt” from the indicspecies package (Cáceres & Legendre, 2009). CAP calculations were performed using the function ‘CAPdiscrim’ from the BiodiversityR package (Kindt & Kindt, 2018) and the ANOSIM test was done using the function ‘anosim’ from the VEGAN package (Oksanen *et al.*, 2013).

iii. Drivers of diversity and distribution patterns

We used a Redundancy Analysis ('RDA', Legendre & Anderson, 1999) to explore the main drivers (environmental, spatial and biogeographical) of community composition. Average values of all environmental variables described in Table I-2 were used in the analysis. We used Moran’s eigenvector maps (MEM, Legendre & Legendre, 2012) to model the spatial structure, using the function “create.dbMEM.model” from the “adespatial” package (Dray *et al.*, 2018). We also created a matrix of seven dummy variables, representing each ecoregion type, to model the dependence of community structure on the ecoregional characteristics.

After constructing the first models, we conducted a forward selection procedure to reduce the number of explanatory variables and obtain more parsimonious models. We included variables with a variance inflation factors (VIF) inferior to 3 to prevent multicollinearity. The selection procedure was performed separately for the variables of the environmental and the spatial components, as recommended by Borcard et al. (2018), using the “ordiR2step”. The significance of all models was tested using a permutation-based ANOVA (999 permutations). Finally, we conducted a variation partitioning analysis (Borcard *et al.*, 1992) using the most parsimonious models.

RESULTS

Our sampling captured a representative fraction of the Orinoco diatom flora, as evidenced by the species accumulation curves performed when streams were grouped in their corresponding ecoregions (Figure IV-3), which approached asymptotes in most of the ecoregions.

i. Elevation-driven diversity and local environment patterns

The diatom community did not exhibit any increasing or decreasing alpha diversity trends along the altitudinal gradient (see the Supplementary Material). Neither their taxa richness (Spearman test: rho = -0.144, S = 3347.1, p = 0.482) nor their Shannon diversity (Spearman test: rho = -0.101, S = 3221.1, p-value = 0.423) were significantly correlated with elevation. The dissimilarity among diatom communities (beta diversity) increased as differences in elevation increased (Mantel test: $\rho = 0.468$, p-value < 0.0001). The dissimilarity of low-profile (Mantel test: $\rho = 0.468$, p-value < 0.0001) and high-profile guilds (Mantel test: $\rho = 0.425$, p-value < 0.0001) exposed a similar trend. Motile diatom group was the group which varied the less as differences in elevation increased (Mantel test: $\rho = 0.215$, p-value < 0.01).

Contrary to the observed patterns in community dissimilarity, environmental dissimilarity between streams was not significantly linked to elevation differences (Mantel test: $\rho = 0.0034$, p-value = 0.48) (see the Supplementary Material). Streams from different ecoregions at similar elevation ranges can exhibit very distinct environmental characteristics. This was consistent in the separate chemical (Mantel test: $\rho = 0.0093$, p-value = 0.284) and physiographic variables ($\rho = 0.0085$, p-value = 0.414) evaluations. By analyzing each variable, we found that only five variables were significantly correlated with elevation. Temperature significantly decreased (Spearman test: rho = -0.855, S = 5425.4, p-value < 0.0001), while saturated oxygen (Spearman test: rho = 0.849, S = 442.58, p-value < 0.0001) and pH

(Spearman test: rho = 0.531, S = 1370.7, p-value < 0.01) increased significantly. A significant increase was also observed in water conductivity (Spearman test: rho = 0.429, S = 1668.8, p-value = 0.028) and PRS (Spearman test: rho = 0.416, S = 1708.9, p-value = 0.035).

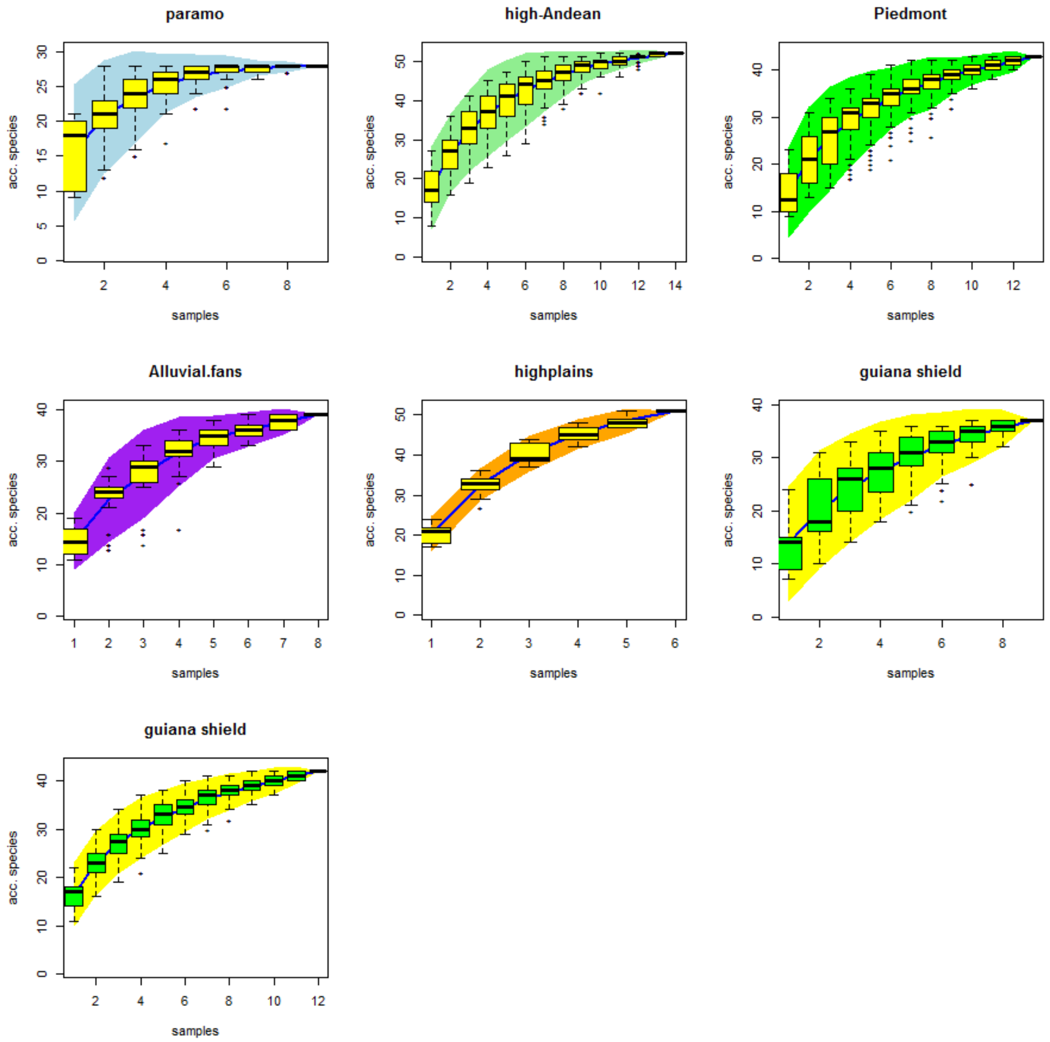


Figure IV-3 Species accumulation curves per ecoregions. Sample counts were pooled by Ecoregion. We used the "random" accumulation method, which adds sites in random order.

ii. Ecoregional diversity patterns

The occurrence of distinct regional diatom pools within the basin was evidenced by the ANOSIM and CAP results, which partially supported hypotheses H2b and H4b (Figure IV-1). Most ecoregions included in the analysis hosted a distinctive set of diatom taxa (Figures IV-4 and IV-5B). Differences were consistent when considering the complete dataset (ANOSIM R = 0.6323, p-value <0.001, Figure IV-4A), as well in the low-profile (ANOSIM R = 0.4675, p-value <0.001, Figure IV-4B), high-profile (ANOSIM R = 0.5962, p-value <0.001, Figure IV-4C) or motile (ANOSIM R = 0.2505, p-value <0.001, Figure IV-4D) guild datasets.

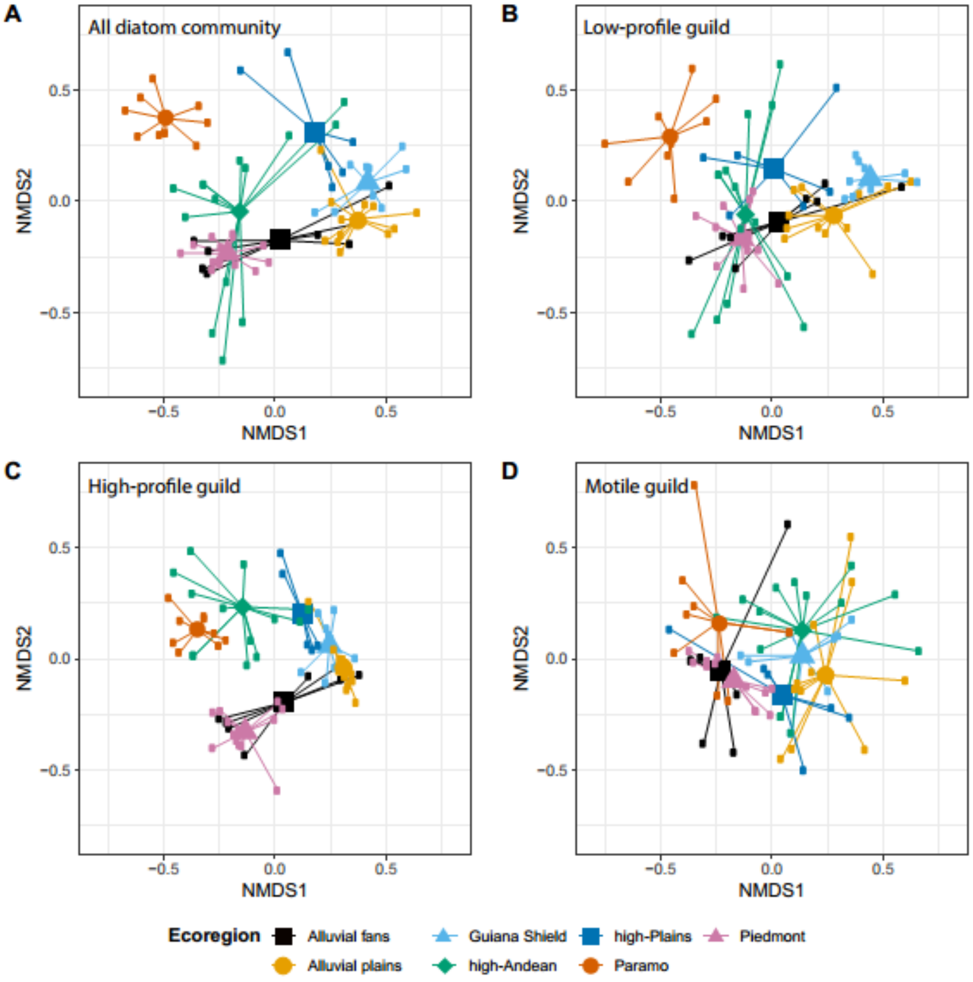


Figure IV-4 Non-metrical multidimensional scaling (nMDS) ordinations of diatom community. Community data was Hellinger-transformed. Distances are represented in a Euclidean space. Stress: 0.17 (A), 0.19 (B), 0.13 (C), 0.19 (D).

Figure IV-5 summarizes the results of the CAP analyses. The percentages of correct classification were similar (about 80%) for the environmental and the community datasets, suggesting a partial match between the community composition and the environmental conditions. The differences were observed when the contemporary environmental conditions in the ecoregions (Figure IV-5A) were compared to their community composition (Figure IV-5B). Some ecoregions can be confidently classified according to their diatom communities but not according to their physiography and/or water chemistry (e.g. Páramo, high-Andes or Piedmont). While others (high Plains and Alluvial fans) can be better classified according to their environmental conditions.

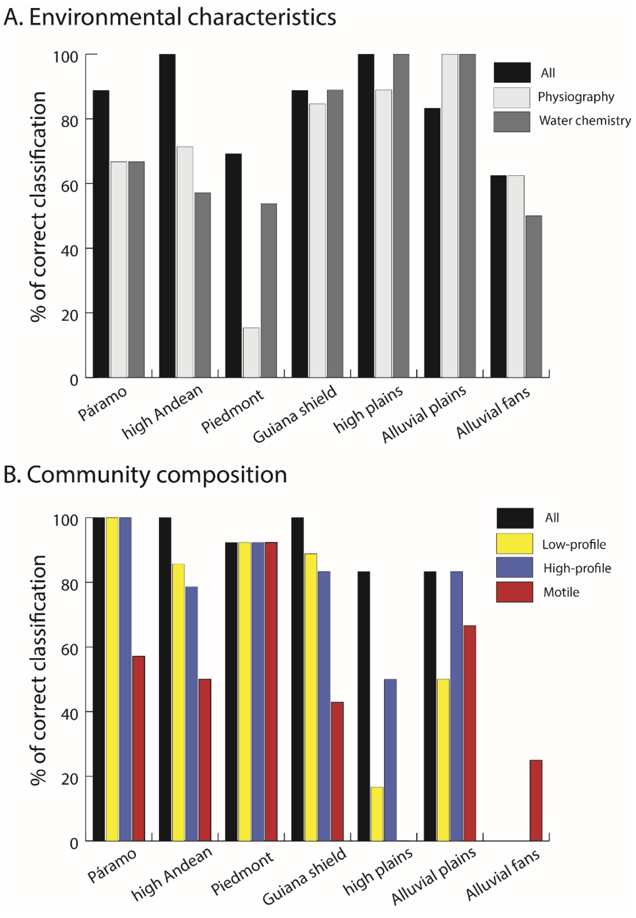


Figure IV-5 CAP analysis results showing the percentage of correct classification of each ecoregion based on their (A) local environment and (B) community composition.

The IndVal analysis revealed 52 morphospecies to be potential indicators of ecoregional distribution (Table IV-1), being *Hannaea arcus* and *Fragilaria capucina* var. *vaucheriae* the most indicative for the Páramos, *Planothidium* sp. pl. and *Cocconeis* sp. pl. for the high-Andeans, *Synedra* sp. pl. and *Gomphonema micropumilum* for the Piedmonts, *Frustulia rostrata* and *Navicula cryptocephala* for the Alluvial terraces, several *Eunotia* and *Gomphonema gracile* for the high-Plains, and *Cymbellopsis* sp. pl for the Guiana shield. There were not species assigned as indicators for the Alluvial fans. Additionally, 17 and 2 genera were potential indicators of combinations of two or three ecoregions, respectively.

iii. Drivers of diversity and distribution patterns

The diatom community structure was partially modeled by the three sets of explanatory variables (environmental, spatial structure and ecoregional identity) (Supplementary Material). The RDA model that included all diatom taxa had the largest variance ($R_{adj} = 0.526$, p-value < 0.0001), followed by the models of the high-profile ($R_{adj} = 0.526$, p-value < 0.0001), low-profile ($R_{adj} = 0.451$, p-value < 0.0001) and motile ($R_{adj} = 0.309$, p-value < 0.0001) taxa. In all four models, the largest constrained variance was explained by the ecoregional component [Eco], and by its joint effect with the environmental component [E + Eco]. The pure effects of the environmental [E] and spatial [S] components - and its respective interaction [E + S] - were relatively low compared to the former components (Figure IV-6).

Table IV-1 List of diatom morphospecies selected as indicator taxa of every ecoregion and group of ecoregions (IndVal Analysis). *Specificity* is the highest (=1) when the species is present in just one ecoregion and *fidelity* is the highest (=1) when the species is present in all streams of one ecoregion.

Ecoregion	Morphospecies	Specificity	Fidelity	IndVal	*p value*
Páramo	*Hannaea arcus* Ehrenberg	0.9057	1	0.952	0.001
	Fragilaria capucina var. Vaucheriae Desmazières 1830	0.8703	1	0.933	0.001
	Cymbella sp. Pl.	0.9412	0.8889	0.915	0.001
	Tabellaria floculosa (Roth) Kützing 1844	0.6466	1	0.804	0.001
	Reimeria uniseriata Sala, Guerrero & Ferrario 1993	0.8262	0.7778	0.802	0.001
	Gomphonema sp. Pl.	0.7185	0.8889	0.799	0.009
	Encyonema minutum (Hilse) D.G.Mann	0.6718	0.8889	0.773	0.001
	Achnanthidium affine (Grunow) Czarnecki 1994	0.7362	0.6667	0.701	0.032
	Fragilaria sp. Pl.	0.942	0.4444	0.647	0.007
	Epithemia cf. turgida	0.9655	0.3333	0.567	0.01
High-Andean	*Planothidium* sp. Pl.	0.8263	0.6429	0.729	0.002
	Cocconeis sp. Pl.	0.8198	0.5714	0.684	0.004
Piedmont	*Synedra* sp. Pl.	0.8797	0.4615	0.637	0.006
	Gomphonema micropumilum Reichardt 1997	0.7712	0.4615	0.597	0.011
	Caloneis sp. Pl.	1	0.3077	0.555	0.005
Alluvial terraces	*Frustulia rostrata* Hustedt	0.8719	0.8333	0.852	0.001
	Navicula cryptocephala Kützing 1844	0.9761	0.25	0.494	0.03
High-Plains	*Eunotia curvula* Hustedt 1952	1	0.6667	0.816	0.001
	Eunotia ventricosa Patrick 1940	1	0.6667	0.816	0.001
	Gomphonema gracile Ehrenberg 1838	0.8739	0.6667	0.763	0.001
	Aulacoseira sp. Pl.	1	0.5	0.707	0.001
	Stenopterobia densestriata (Hustedt) Krammer 1987	0.9091	0.5	0.674	0.001
	Aulacoseira granulata (Ehrenberg) Simonsen 1979	1	0.3333	0.577	0.008
	Eunotia rabenhorstiana (Grunow) Hustedt 1949	1	0.3333	0.577	0.01
Guiana Shield	*Cymbellopsis* sp. Pl.	0.8566	0.5556	0.69	0.001
Páramo & high--Andean	*Encyonema silesiacum* (Bleisch) D.G.Mann 1990	0.9214	0.8261	0.872	0.001
	Fragilaria capucina Desmazières 1830	0.9952	0.6522	0.806	0.001
	Diatoma mesodon (Ehrenberg) Kützing 1844	1	0.3043	0.552	0.007
High-Andean & high Plains	*Eunotia exigua* (Brébisson ex Kützing) Rabenhorst 1864	1	0.35	0.592	0.01

Table IV-1. *Continued.*

Piedmont & Alluvial fans	*Gomphonema lagenula* Kützing 1844	0.9446	0.9048	0.924	0.001
	Nitzschia palea (Kützing) W.Smith 1856	0.8795	0.9524	0.915	0.001
	Gomphonema parvulum (Kützing) Kützing 1849	0.8401	0.5714	0.693	0.004
	Amphora cf. *montana*	0.8846	0.4286	0.616	0.012
Alluvial terraces & Guiana Shield	*Encyonopsis frequentis* Krammer, 1997	0.8747	0.9524	0.913	0.001
	Encyonema sp. Pl.	0.8247	0.7619	0.793	0.002
	Eunotia rhomboidea Hustedt 1950	1	0.2381	0.488	0.045
Alluvial terraces & high Plains	*Frustulia* sp. Pl.	0.8093	0.9444	0.874	0.001
	Chamaepinnularia sp. Pl.	0.8066	0.6111	0.702	0.001
Guiana Shield – high Plains	*Eunotia bilunaris* (Ehrenberg) Schaarschmidt 1880	0.8615	0.3333	0.536	0.018
Piedmont & high Plains	*Adlaphia* sp. Pl	0.8257	0.3684	0.552	0.014
Alluvial fans & Piedmont & high-Andean	*Achnanthidium eutrophilum* (Lange-Bertalot) Lange-Bertalot 1999	0.9893	0.5714	0.752	0.002
	Pinnularia sp. Pl.	0.9398	0.4839	0.674	0.007

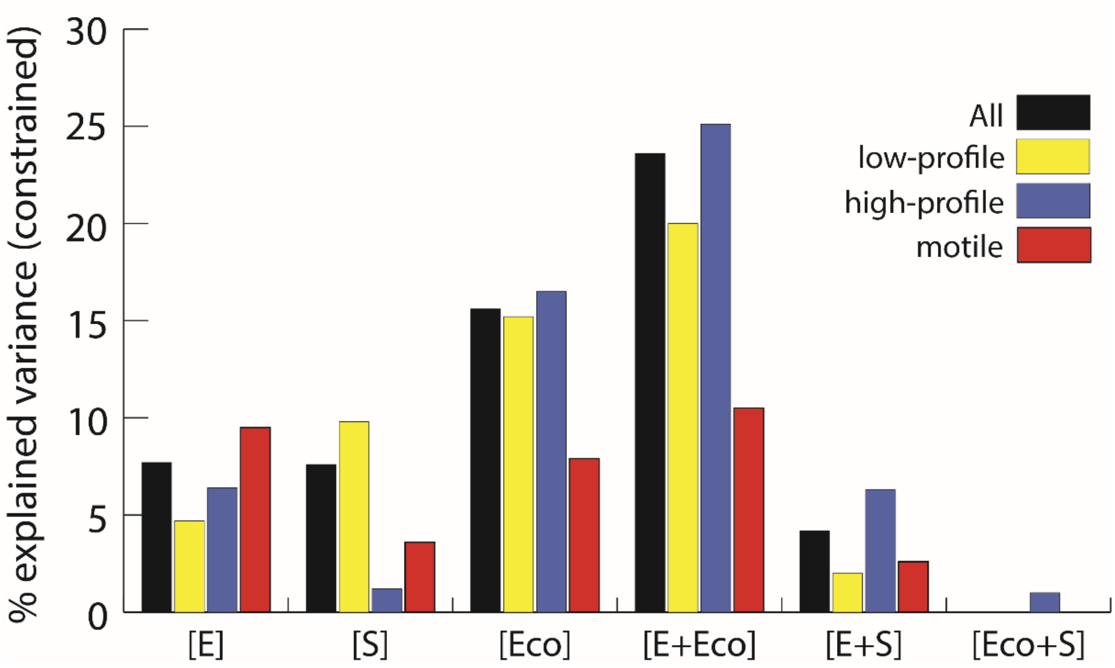


Figure IV-6 Variation partitioning results for the whole community (black), low-profile guild (green), high-profile guild (blue) and motile guild (red) datasets. Variance was partitioned for the effect of the environment [E], spatial structure [S] and ecoregion [Eco], and its respective joint effects: environment and spatial [E+S], environment and ecoregion [E+Eco], and spatial and ecoregion [S+Eco]. The joint effect of the three components [E+S+Eco] was non-significant in the four models.

The forward selection procedure retained the first and last MEM eigenvectors, indicating that community structure was explained to a similar degree by the streams' spatial structure [S] at large and local scales (Table IV-2). Regarding the environmental component [E], the forward selection procedure retained pH and temperature in all the datasets (Table IV-2). The first axis represented a gradient of pH, while the second and third axes were distinctly correlated to temperature. The hydro-morphological variables and nutrients were more highly correlated with the second and third axes respectively (Table IV-3). The other environmental variables were highly correlated with the second and the subsequent axes, depending on the dataset used.

Table IV- 2 Spatial and environmental variables retained in each model for the forward selection procedure. In bold, the common variables retained in the four models.

	Whole community	**Low-profile**	**High-Profile**	**Motile**
Spatial structure	**dbMEM – 1,**2,5,6	**dbMEM1,**2,4,7	**dbMEM1,**5	**dbMEM1,**6
Environmental conditions	**pH, Temperature** Qmax/Q Low mobility substrate PRS N2tQ Conductivity TOC	**pH,Temperature** N-NH$_4$ Low mobility substrate PRS N-NO$_3$ N2tQ Conductivity	**pH, Temperature** Qmax/Q NH$_4$ N-NO$_3$ TOC	**pH, Temperature** Q Low mobility substrate nEvents2t PRS

Table IV- 3 Correlation coefficients of the environmental variables with each axis of the distinct RDA models. In red are shown the variables with higher correlation coefficients. In red the variables with the highest correlation to the first two axes of the RDA.

Dataset	Variable	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6
Whole	**pH**	**-0.925**	0.012	0.004	-0.187	-0.166	-0.175
	T	**0.524**	**0.594**	-0.539	0.127	0.109	-0.072
	N.NH4	-0.053	-0.281	-0.796	0.309	0.338	0.119
	Qmax.Q	0.214	0.440	0.140	-0.252	0.140	0.473
	low_mob.subs	-0.044	-0.259	0.167	0.094	-0.826	0.314
	PRS	-0.353	0.139	0.643	0.354	0.071	-0.066
	n2tQdays	0.163	-0.522	-0.092	0.013	0.242	-0.692
	Cond	-0.269	0.404	-0.453	-0.198	-0.156	-0.178
	TOC	0.026	0.218	0.080	0.631	-0.078	-0.173
low-profile	**pH**	**-0.871**	0.077	-0.024	-0.257	-0.272	0.015
	T	**0.602**	**-0.593**	-0.297	-0.343	0.048	-0.192
	N.NH4	-0.081	0.210	-0.622	-0.360	0.256	-0.145
	low_mob.subs	0.120	0.529	-0.257	0.217	-0.001	-0.268
	PRS	-0.410	-0.163	-0.050	0.744	-0.419	-0.267
	n2tQdays	0.021	0.314	0.447	-0.172	-0.017	-0.628
	Cond	-0.167	-0.156	-0.384	-0.521	-0.659	0.305
high-profile	**pH**	**0.971**	-0.059	0.182	-0.009	0.090	0.113
	T	**-0.536**	**-0.765**	-0.236	0.137	-0.178	-0.148
	N.NH4	0.086	-0.119	-0.855	0.316	-0.259	0.284
	TOC	-0.013	-0.180	-0.308	-0.924	0.074	-0.114
	Qmax.Q	-0.308	-0.264	0.417	-0.071	-0.295	0.755
	PRS	0.322	0.316	0.427	-0.159	-0.706	-0.301
motile	**pH**	**0.977**	-0.029	0.065	-0.035	0.116	0.154
	T	**-0.534**	**-0.783**	0.030	0.079	-0.279	-0.132
	low_mob.subs	0.090	0.352	0.219	0.070	0.193	-0.212
	PRS	0.326	0.342	0.299	-0.719	-0.321	-0.257
	TOC	-0.015	-0.192	-0.885	-0.380	0.074	-0.100
	Q	0.073	-0.232	0.465	-0.007	0.643	-0.485
	nEvents2tQ	-0.176	-0.501	0.334	-0.486	0.438	0.373

DISCUSSION

Contrarily to previous findings, changes in diatom diversity and distribution were partially uncoupled from the Andean elevational gradient, but greatly explained by the basin ecoregional gradient. Overall, our study provides several evidences in support of Martiny et al.'s (2006) fourth hypothesis (Figure IV-1, H4B) as the primary explanation for the observed patterns in community composition. It means that the interplay between contemporary processes and past historical events is driving the distribution of diatoms in the Colombian Orinoco basin. Consequently, the species composition of a given community is shaped by the effect of dispersal limitation within an evolutionary timeframe and the effect of environmental filtering within an ecological timeframe. Particularly, our findings suggest that the historical legacies associated with the Andean uplifts would have shaped distinct regional pools of diatom taxa, promoting diversity at the regional scale.

Since von Humboldt & Bonpland (1807), elevation-driven changes in temperature have been stated as one of the major drivers of biodiversity patterns (Jackson, 2009). Nottingham et al. (2018), for instance, recently showed that alpha and beta diversity patterns of microbes, fungi and plant communities coincide with an altitudinal gradient in the Peruvian Andes and concluded that temperature was the main driver of microorganism diversity and distribution. Our results partially conformed to these findings. Contrarily to the Nottingham et al.'s study, stream environmental features were spatially-constrained according to the ecoregion type, rather than changing monotonically with elevation.

Spatially-constrained environments can obscure longitudinal, latitudinal- and/or elevational patterns of diversity when the environmental variability does not change linearly in an spatial dimension (Passy *et al.*, 2018). Our results are in line with this statement, since we found that ecoregional, rather than elevational drivers, are underlying the distribution of diatoms in Neotropical streams. Ecoregion type outperformed elevation in explaining the environmental features of the streams (i.e. CAP 83.1% of correct stream classification based on ecoregions). We observed that, regardless of their spatial proximity, diatom communities from streams within the same ecoregions were more similar to each other than to those from different ecoregions.

Uncovering this pattern provides a potential mechanism for a better understanding of diatom distribution, and possibly diatom biogeography in the Neotropics. The ecoregionaly-constrained distribution of diatom taxa suggests that historical constraints of dispersal may have been the major determinants of their present-day distribution. Although a similar distribution has already been observed in Neotropical lakes diatoms (Benito *et al.*, 2018), it is still surprising to observe this level of discreteness in stream networks, since their dispersal is limited to a lower degree than lakes (Vilmi *et al.*, 2017).

The observed concordance between the distributions of diatoms and ecoregions can be explained by the evolutionary species pool hypothesis (SPH) (Pither & Aarssen, 2005), which states that the longer a set of environmental conditions persists in time and space, the greater the opportunity for the evolution of adapted species (Pither & Aarssen, 2005). This theory predicts that several pools of species might result from the conjoint action of events limiting dispersal and the filtering of species according to their tolerances. This interplay has already shown to be the main driver of biogeographic provincialism of *Lupinus* plants (Hughes & Eastwood, 2006). Although further evidences on diatom phylogenetic relationships and radiation events would be needed to confirm this mechanism, the higher variance of the [Eco + E] interaction (Figure IV-6), and the occurrence of ecoregionaly-constrained pools of diatom taxa, provide initial supporting evidences. We have recently observed a similar pattern in the Orinoco basin assessing invertebrate communities (González-Trujillo *et al.*, 2020). In this regard, the concordance in distributional patterns between two biological communities (diatoms and invertebrates) with distinct dispersal abilities may imply that historical legacies play a key role determining the present-day distribution of freshwater biota in Neotropical basins.

Historical legacies alone cannot fully explain diatom distribution. This is evident when comparing the models for the different diatom guilds, which suggests that different drivers may be underlying the distribution of each guild. For example, the non-spatially-structured distribution of motile diatoms suggests that other underlying processes, such as mass effects or stochastic forces (Jamoneau *et al.*, 2018), must have important roles. The motile guild collects species able to use the available resources and support disturbances (Dong *et al.*, 2016; Passy, 2016; Falasco *et al.*, 2019), which makes them potentially successful under a multiplicity of circumstances, and rather independent of any spatial structure. On the other hand, the distribution of low profile species is expected to be weakly structured in the space because of their high dispersal ability, which is mostly relevant at large scales (Passy, 2016). Still, these species have preferences for high mountain sites because of their tight attachment abilities and preferences for poor nutrients and capacity to resist fast water velocities (Bottin *et al.*, 2014; Dong *et al.*, 2016), and this may support their preference for high-altitude ecoregions. The high-profile species, which assembles species remaining erect or colonial, spanned mostly through low altitude sites, as much as it has been observed under other situations (Passy 2006, Bottin et al. 2014). Overall, there is a certain degree of discreteness in the distribution of the different guilds which may support the occurrence of biogeographical boundaries on the contemporary distribution of diatoms in the Neotropics. Nonetheless, further research is needed to confirm if the drivers of diatom biogeography can change as a function of other functional traits (e.g. growth forms).

Ecoregional characteristics can also exert a high influence on ecological timeframes, acting as large-scale drivers of local environmental conditions (Varanka & Luoto, 2012; Neff & Jackson, 2013; Vilmi *et al.*, 2017). In this study, the stream's relevant physical and chemical characteristics, such as water pH, temperature, conductivity and soluble reactive phosphorus (Table IV-2) were related to some of the ecoregions. Of these, pH provided the higher degree of differentiation, and separated the High Andean and Páramo regions from the alluvial terraces and the Guaiana shield (Supplementary Material). The RDAs identified diatom taxa such as *Hannaea arcus*, *Fragilaria capucina*, or *Tabellaria* *flocculosa*, which are common in moderately low pH conditions, and others, such as *Frustulia rostrata* or *Navicula cryptocephala,* which are more common in circumneutral pHs (Whitmore, 1989; Poulíčková *et al.*, 2010). Result that is also supported by the IndVal analysis, which attributed these sets of species as the best indicator taxa of the Páramos and the latter to be good indicators of the alluvial terraces. These findings are in consonance with the fact that algae living under low pH conditions have developed adaptative systems, such as the use of active pumps to prevent H^+ ions entering the cell (Seckbach & Oren, 2007).

Concluding remarks

Here, we have provided evidence that ecoregional characteristics drive the community composition of Neotropical stream diatoms. Historical legacies, particularly, might explain the observed ecoregionaly-constrained distribution at the regional scale. Analogous findings have already been found for microbial biogeography in high-latitude streams (Vyverman *et al.*, 2007a; Verleyen *et al.*, 2009; Liu *et al.*, 2013; Widder *et al.*, 2014). However, to our knowledge, our study is one of the firsts to assess the main drivers underlying microorganism distribution in Neotropical streams. Understanding the main drivers for a basal-diversity pattern may help define the subsequent conservation steps not only for the algal and diatom flora but also for the ecosystems hosting them. Our overall results suggest that, not only future studies, but also the delimitation of zones of conservation priority should consider the type and the extent of the different ecoregions.

V. Upstream refugia and dispersal ability may override benthic-community responses to high-Andean streams deforestation

ABSTRACT

Deforestation is a major driver of biodiversity loss in the Tropical region, but the role of upstream refugia and dispersal ability on the community response to this disturbance is unknown. We assessed the relevance of undisturbed upstream patches ("refugia") on the responses of benthic communities to forest cover loss. We selected four Andean rivers with a well-protected forest in their upstream section and different degree of forest cover loss downstream and evaluated the dissimilarity patterns of three benthic communities (diatoms, Chironomidae, and the assemblage of Ephemeroptera, Plecoptera and Trichoptera – 'EPT') along their water courses. We evaluated the dissimilarity by using Euclidean (environment), Sørensen (incidence data) and Morisita-Horn (abundance data) pairwise distances. We found that diatom beta-diversity, as organisms with passive but higher dispersal ability, significantly tracked the environmental changes caused by forest loss. However, insect communities, whose *a priori* are active dispersers and can track for suitable conditions, were weakly affected by deforestation. These results provide evidences that the existence of well-preserved upstream reaches along patched corridors may allow non-tolerant species to remain extant throughout dispersal-driven feedbacks. This being the case in the Andean streams, where effects of deforestation on benthic communities were small, and not uniform. Our results reinforce the strategy of preserving upstream sections in order to achieve successful restoration or rehabilitation goals.

-o-

An edited version of this chapter has been published in the journal of *Biodiversity and Conservation*. Please, cite the respective paper as:

- **González-Trujillo, J. D., Petsch, D., Córdoba-Ariza, G., Rincón-Palau, K., Donato-Rondon J. Ch, Castro-Rebolledo, M., Sabater, S. Biodiversity and Conservation (2019) 28: 1513.** https://doi.org/10.1007/s10531-019-01739-2

INTRODUCTION

Although deforestation is one of the major drivers of biodiversity loss in the Tropical region (Barlow *et al.*, 2016), its effects on biodiversity structural and functional components have been scarcely evaluated in the Tropics (see on this regard the meta-analysis evidences by Richardson and Béraud 2014). Deforestation effect studies have mostly focused on comparing changes on alpha-diversity (i.e. local richness) and bioassessment-derived metrics (e.g. the Biological Monitoring Working Party, 'BMWP'; see Hawkes 1998) between forested and non-forested streams (e.g. Bojsen and Jacobsen 2003; Lorion and Kennedy 2009; Iñiguez-Armijos et al. 2014). As these alpha-diversity and bioassessment-derived metrics do not take into account the potential influence of dispersal (Hitt & Angermeier, 2008), they cannot quantify the relevance of refugia occurrence (i.e. sites where the absence of an impairment have enabled the native communities to survive) among the stream network in structuring local communities and basin-level metacommunities (i.e. a set of communities potentially linked by dispersal; Leibold et al. 2004). Addressing this knowledge gap is crucial to improve bioassessment protocols as well as the performance of mitigation plans and conservation practices.

The use of alpha-diversity and bioassessment-derived metrics assumes that community composition will reflect the environmental change linked to one or more impairments (Siqueira *et al.*, 2014a; Buss *et al.*, 2015). Under this assumption, species at impaired sites will be those with greater resistance to the new environmental conditions (e.g. species more tolerant to increased input of sediments covering the streambed, reduced inputs of allochthonous organic material or input of fertilizers in deforest segments). In this way, environmental sorting would be the main process underlying community and metacommunity structure (i.e. species composition and distribution across local communities). However, streams are dendritic and highly connected ecosystems (Campbell Grant *et al.*, 2007; Altermatt *et al.*, 2013), where dispersal-driven processes may override the effect of environmental sorting on the structure of local communities and metacommunities. This is mostly related to two situations (e.g. Tonkin et al. 2014, 2018; Kitto et al. 2015; Sarremejane et al. 2017): when species have high dispersal ability, through which they can virtually establish everywhere; or when there exists a high connectivity between source (e.g. refugia) and sink patches (e.g. impaired sites). Considering that undisturbed headwater segments can serve as refugia for species under human-related disturbances (Sedell *et al.*, 1990), both species dispersal ability and the distance to the refugia cannot be ignored when assessing the effects of impairment.

One approach to explore the influence of dispersal and connectivity on metacommunity structure is to focus on characterizing the beta diversity patterns of communities possessed of species with distinct dispersal abilities (Kärnä *et al.*, 2015; Tonkin *et al.*, 2018). The beta diversity summarizes the variability in species composition between sampling units or 'local communities' (Anderson *et al.*, 2006). Therefore, analyzing the beta diversity patterns along stress gradients may help disentangling the effects of environmental heterogeneity, network spatial structure, and/or dispersal-driven processes on metacommunity structure (Gutiérrez-Cánovas *et al.*, 2013; Heino *et al.*, 2015a). By decomposing beta diversity in its turnover and nestedness components (Baselga 2010), it is also possible to infer the main effects of a particular stressor on metacommunity structure (Gutiérrez-Cánovas *et al.*, 2013; Socolar *et al.*, 2016). That is, if nestedness is the major driver, the stressor may promote assemblage disaggregation and leads communities at impaired zones to be subsets of the undisturbed ones (Gaston *et al.*, 2000; Tornés & Ruhí, 2013). On the contrary, if turnover is the main driver, the stressor may promote the replacement of some species by others according to their stress tolerance (Qian *et al.*, 2005).

We therefore used beta-diversity metrics to assess the effect of deforestation on the distribution patterns of benthic communities, and to determine how it is modulated by the occurrence of upstream refugia and dispersal ability. So forth, we evaluated beta-diversity patterns of high-elevation Tropical benthic communities within four Andean streams, all having an upstream section with well-protected forest and two downstream sections with varying degrees of forest cover loss (Figure V-1A). As populations of high-dispersal species may persist at impaired reaches through colonization from upstream refugia – in a source-sink dynamics (Leibold *et al.*, 2004; Leibold & Chase, 2018), we expected communities possessed of species with low dispersal ability to track better the local environment changes linked to deforestation. Thus, we hypothesized the relationship between beta-diversity and environmental dissimilarity between reaches of each stream to be weaker as dispersal ability increases (Figure V-1B). Besides, if local environmental conditions change in line with forest cover, we also hypothesized that within-stream beta-diversity will change accordingly to the degree of forest cover loss along each stream: a lower beta-diversity should be expected when the forest cover is similar along the stream, as local environment is similar between reaches (Figure V-1B, stream “S1”); a gradual increase in beta-diversity had to be expected when the stream had a gradual loss of forest cover (Figure V-1B, streams “S2” & “S3”); and, finally, when the forest cover would abruptly decrease (Figure V-1B, stream “S4”), beta-diversity should be expected to increase abruptly as well.

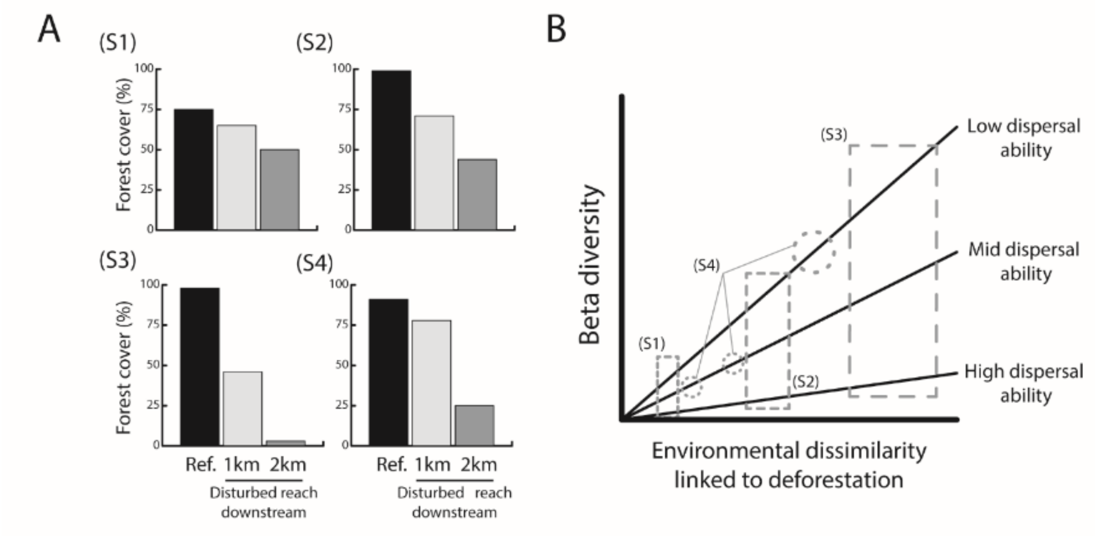


Figure V-1 Percentage of forest cover at each stream reach (A) and conceptual scheme representing the expected biotic and environmental dissimilarity among reference and impact stream reaches (B). The dotted rectangles represent the expected position of each stream in the relationship between community and environmental dissimilarity. The dotted circles represent S4 stream reaches.

We tested these hypotheses by using three benthic communities possessed of taxa with specific dispersal abilities, since they show differences in their body size and dispersal strategy (De Bie *et al.*, 2012; Padial *et al.*, 2014). The first community is made up by benthic diatoms, which are passively dispersing organisms. Diatoms are sessile microorganisms whose dispersal is mostly passive and confined to the stream network (Liu *et al.*, 2013). However, the small body size of diatoms is usually related to a high dispersal capacity (Padial et al. 2014). The second community selected was the non-biting midges (Chironomidae), which hardly disperse actively due to their small body size but can passively reach longer distances in both aerial and aquatic pathways (Serra et al. 2016). And the third community is that made up by Ephemeroptera, Plecoptera and Trichoptera (EPT), which are taxa with active dispersal through aquatic and aerial pathways (Malmqvist, 2000; Graham *et al.*, 2017). Given their high dispersal ability, EPT taxa tend to seek for the most suitable places for inhabiting or laying eggs (Thompson & Townsend, 2006; Lancaster & Downes, 2017). Overall, these three communities represent a great set to assess for the role of dispersal on structuring benthic metacommunities. All three communities are frequently used in bioassessments due to their sensitivity to environmental changes (e.g. Barbour et al. 1999; Leira and Sabater 2005; Buss and Vitorino 2010; Lencioni et al. 2012; Milošević et al. 2013; Tornés et al. 2018); and therefore, can be accurately used for assessing the effects of deforestation on river integrity and health.

METHODS

Study design and site selection

The Andean mountains are one of the most diverse regions in the world, but also highly impaired by historical land-use changes (Myers *et al.*, 2000; Etter *et al.*, 2008). About 62% of the natural ecosystems had been cleared by 2000, and causes are multiple and include human settlement expansion, cattle rising, illegal cropping, and the expansion of agricultural frontiers (Rodríguez Eraso *et al.*, 2012). Although Andean countries have laws for protecting the forest, the minimum width - 30m in general – is usually not maintained along stream corridors. This has caused that Andean streams exhibit a patched pattern of forest along their corridors: the most cleared zones are located near human settlements, and the most preserved are confined to protected or inaccessible areas, such as headwaters or stepper valleys (Armenteras *et al.*, 2011). We used this pattern as an opportunity to study the relevance of these patches on the response of three different benthic communities to deforestation.

We selected four second-order streams from the high-Andean region which differed to each other on the patterns of riparian forest cover loss. All four streams drain to the Meta River basin and are located between 2300 and 2800 m a.s.l (department of Cundinamarca, Colombia; Figure V-2A). Their waters drain areas mainly cover by high-Andean forest ecosystems but some of them (S1 and S4) flow partially through Páramo ecosystems. Páramo regions are endemic to the Neotropical Andean Mountains and are recognizable by their extreme temperature variations (intraday variations of more than 20ºC), soils with a high water retention capacity (such as Andisols), and vegetation chiefly comprised of giant rosette plants (*Espeletia*), shrubs, and grasses (Buytaert *et al.*, 2006).

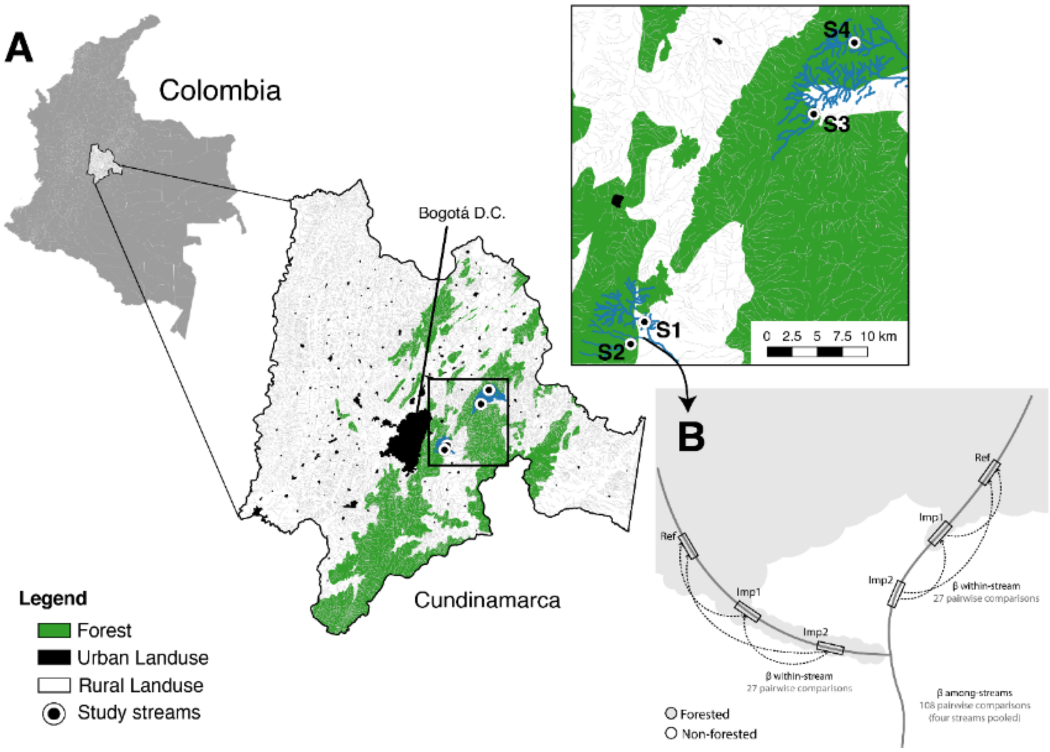


Figure V-2 Study zone map (A) and the scheme representing the sampling design and the procedure to estimate pairwise distances within and among streams (B). The rectangles represent stream reaches.

Using aerial photographs (Google Maps Pro 2017, version 7.1), we estimated the forest cover along the stream channel within buffers of 30m- (representing the minimum width established by the Colombian Ministry of Environment: decree 2245 of 2017). After scanning photographs, we selected one segment of minimum 2km-long per stream and established three sampling reaches of 100m-long. Stream reaches, which consisted of at least three riffles, were separated to each other by ca. 1 km (Figure V-2B). Although some studies have demonstrated that insect dispersal may be as high as 5 km (reviewed by Bohonak and Jenkins 2003), recent evidence points out that dispersal of Diptera, Plecoptera, Ephemeroptera and Trichoptera individuals may be lesser than 1 km within a generation (Macneale *et al.*, 2005; Poff *et al.*, 2006; Winking *et al.*, 2014; Graham *et al.*, 2017). Therefore, this distance was considered enough to reliably assess the influence of dispersal ability on community responses to deforestation. Particularly, in our study streams, where we found at least one waterfall between the sampling reaches of each stream.

Environmental and biological data collection

The sampling was conducted during the dry season (10th to 22nd February 2017), when biological communities (insects and diatoms) were well developed. Environmental and biological samples were collected in the riffle sections of each stream segment. We measured 11 environmental variables in order to characterize the environmental conditions of each of the reaches (Table V.1). Conductivity, pH, oxygen, and temperature were recorded using a HANNA HI98194 water quality meter. Canopy cover (%) was estimated from an upward photo taken through a fisheye lens. Channel width, depth and current velocity were measured at two cross-channel transects located at the beginning and ending sections of each riffle. Depth and water velocity were measured at 0.25m intervals, and the average value per riffle section was used in the analyses. The total suspended solids (TSS) were determined by filtering 500ml of stream water in a pre-weighted GFF filter and drying the filter for 1 hour at 105ºC. Organic matter was collected from each Surber sample after sorting all invertebrates (see below), dried for 48-72 hours at 70ºC, and oven-dried for 4 hours at 500ºC to obtain the ash-free dry weight (Elosegi & Sabater, 2008). The ash-free dry weight of each sample was used as a proxy of the Benthic Organic Matter (BOM) collected at each sample.

We collected brush-scraped material from a 0.008m^2 surface of 30 boulders of similar size at each riffle for the diatom samples. Samples were preserved in a Transeau solution (1:1 ratio) and taken to the laboratory where diatom frustules were cleaned from the organic material using hydrogen peroxide. We mounted frustules on permanent slides using Naphrax® medium. Slides were observed with immersion oil at a 1000x magnification under a light microscope and identified at genus-level using specialized monographs (Krammer and Lange-Bertalot 1986, 1988, 1991a, 1991b; Metzeltin and Lange-Bertalot 2007; Bellinger and Sigee 2015). At least 400 valves were counted for each slide.

The invertebrates were collected following a multi-habitat sampling procedure. This was performed by collecting 5 Surber (mesh size 350 mm, area 0.09 m^2) samples in stream substrata selected according to the corresponding habitat coverage. For instance, if a riffle was composed by 60% of boulders, 30% gravels, and 10% cobbles, 3 Surber samples of the first, 1 of the second, and 1 of the third substrate type were collected. A total of 60 samples were taken for this study (4 streams/segments x 3 riffles x 5 samples), each sample being preserved in a plastic bottle with 96% alcohol. In the laboratory, all the individuals were sorted and identified at genus-level following (Trivinho-Strixino & Strixino, 1995; Merritt & Cummins, 2008; Domínguez & Fernández, 2009; Prat *et al.*, 2011, 2014).

Data preparation

The grain sizes for insect and diatom communities were a Surber sample and a composed sample per riffle, respectively. To create comparable datasets, Chironomidae and EPT abundances were weighed according to the substrate coverage at each riffle. Thus, all three datasets had a grain size equivalent to a riffle section; and therefore, a total of 36 rows (4 stream segments x 3 reaches - 1 reference, 2 downstream impairments – x 3 riffles).

Data analysis

We tested the abovementioned hypotheses by measuring the relationship strength between community and environmental dissimilarities within each stream. Community dissimilarity, estimated by using Sørensen (Sørensen, 1948) and Morisita-Horn (Horn, 1966) distances, tell us how much community composition changed between reaches with different degrees of forest loss. Meanwhile, environmental dissimilarity, estimated by using a Euclidean distance, tell us how much the local environment changed due to the loss of forest cover. Thus, if environmental sorting is determining community composition, we expected this relationship to be stronger in streams with a greater degree of forest cover loss along their course. The greater the loss, the greater should be the impact on the local environment; and therefore, the greater should be the change in species composition.

To evaluate and compare the strength of this relationship between streams and communities, we used the regression coefficients obtained from linear regressions of the Sørensen and Morisita-Horn pairwise distances against the Euclidean pairwise distances. Pairwise-Euclidean distances were estimated by using uncorrelated variables (Pearson correlation coefficient < 0.6; Booth et al. 1994) to minimize collinearity: BOM, pH, conductivity, temperature, TSS, Canopy cover, oxygen saturation, depth, current velocity, and width. On the other hand, a matrix of incidence data (presence-absence) was used for estimating Sørensen distances, and a matrix of relative abundances was used for estimating Morisita-Horn distances. These two distances allow us to evaluate how much of the dissimilarity was driven by differences in species composition (Sørensen) or by changes in species relative abundance (Morisita-Horn). Besides, the Morisita-Horn distances shows a high resistance to under-sampling by giving higher weight to the most abundant species. Since the Morisita-Horn is related to the Sørensen index, two communities having the same species and analogous abundances distribution, will provide the same values in the two indices (Jost *et al.*, 2011). This property is useful for disentangling in which way riparian loss affects species composition or their abundances. Additionally, we decomposed the Sørensen index (β_{SOR}) in their Turnover (β_{SIM}) and Nestedness (β_{NES}) components (Baselga, 2010).

As our aim was to evaluate how the occurrence of upstream refugia can modulate community-level responses to deforestation, we only estimated pairwise distances within-streams but not among them (Figure V-2B). Moreover, as it has been found in other Tropical zones (Múrria *et al.*, 2015), our streams were significantly different in terms of their community composition (see Supplementary Material). Thus, we considered that estimating distances between streams could add unnecessary noise to the analysis, while not providing additional evidences for testing our hypotheses

We estimated regression coefficients for each stream (“within-stream”) and for all streams (“among-streams”) (Figure V-2B). Within-streams, we quantified the relationship using a simple linear regression of the pairwise distances between reaches of each stream (n= 27 pairs for stream). We used the coefficients of these regressions to evaluate if the community-environment relationship changes at each stream and depending on the community. For quantifying the among-streams relationship, we pooled the pairwise distances between reaches of each stream (n= 108 pairs) and performed generalized least squares (GLS) regressions. The coefficients of these regressions allow us to evaluate if the responses to deforestation varied depending on the community. We used GLS since these models can deal with data violating the heteroskedasticity assumption - heterogeneous variances among groups (streams). The heterogeneity of each stream was included in the random term of the model as suggested by Zuur et al. (2009). All analyses and graphical outputs were performed with the packages *vegan* (Oksanen *et al.*, 2013)*, ade4* (Chessel *et al.*, 2017), *betapart* (Baselga &

Orme, 2012b), and *ggplot2* (Wickham, 2010) of R Statistical software v 3.4.1 (R Core Team, 2018). All the assumptions were previously tested by following Zuur et al. (2009) suggestions.

RESULTS

We identified 26 genera of diatoms, 28 genera of Chironomidae and 26 genera of EPT (9 genera of Ephemeroptera, 1 genus of Plecoptera, and 16 genera of Trichoptera) in the whole set of stream reaches (Supplementary Material). The accumulation curves indicated that sampling was adequate, since curves tended to approach an asymptote (Figure V-3).

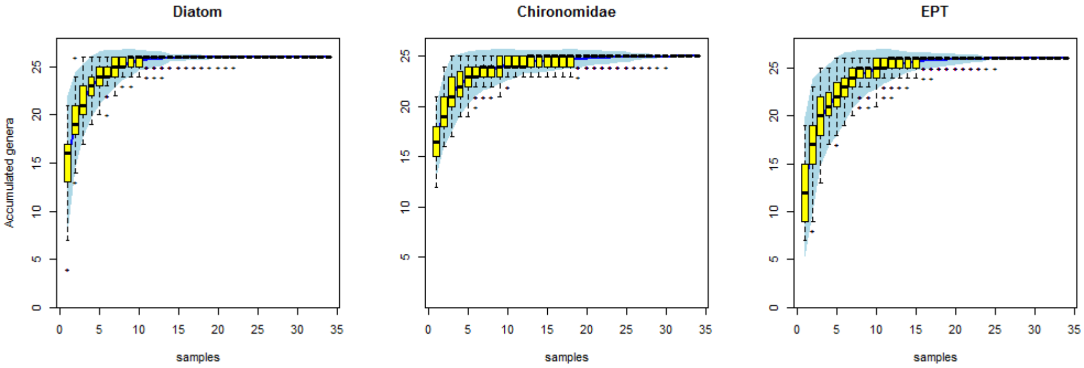


Figure V-3 Species accumulation curves of the three assemblages. We used the "random" accumulation method, which adds samples randomly.

Deforestation effects on stream's local environmental were evidenced 1km and 2km downstream from undisturbed reaches (Figure V-4). The physical and chemical characteristics tended to change in line with the percentage of forest cover loss. However, the magnitude of change varied depending on the stream. We observed a low dissimilarity between reference and impaired reaches at streams S1 and S2, but a high dissimilarity within reaches of the stream S3 (the most impaired stream). At the stream S4, the environmental conditions did not reflect the abrupt loss of forest coverage starting 2km downstream of the undisturbed reach (Figure V-4C and V-4D).

Canopy cover, temperature, TSS, and conductivity were the variables that significantly changed within each stream (Table V-1). Canopy cover diminished in disturbed reaches, while temperature, TSS and conductivity increased at undisturbed reaches.

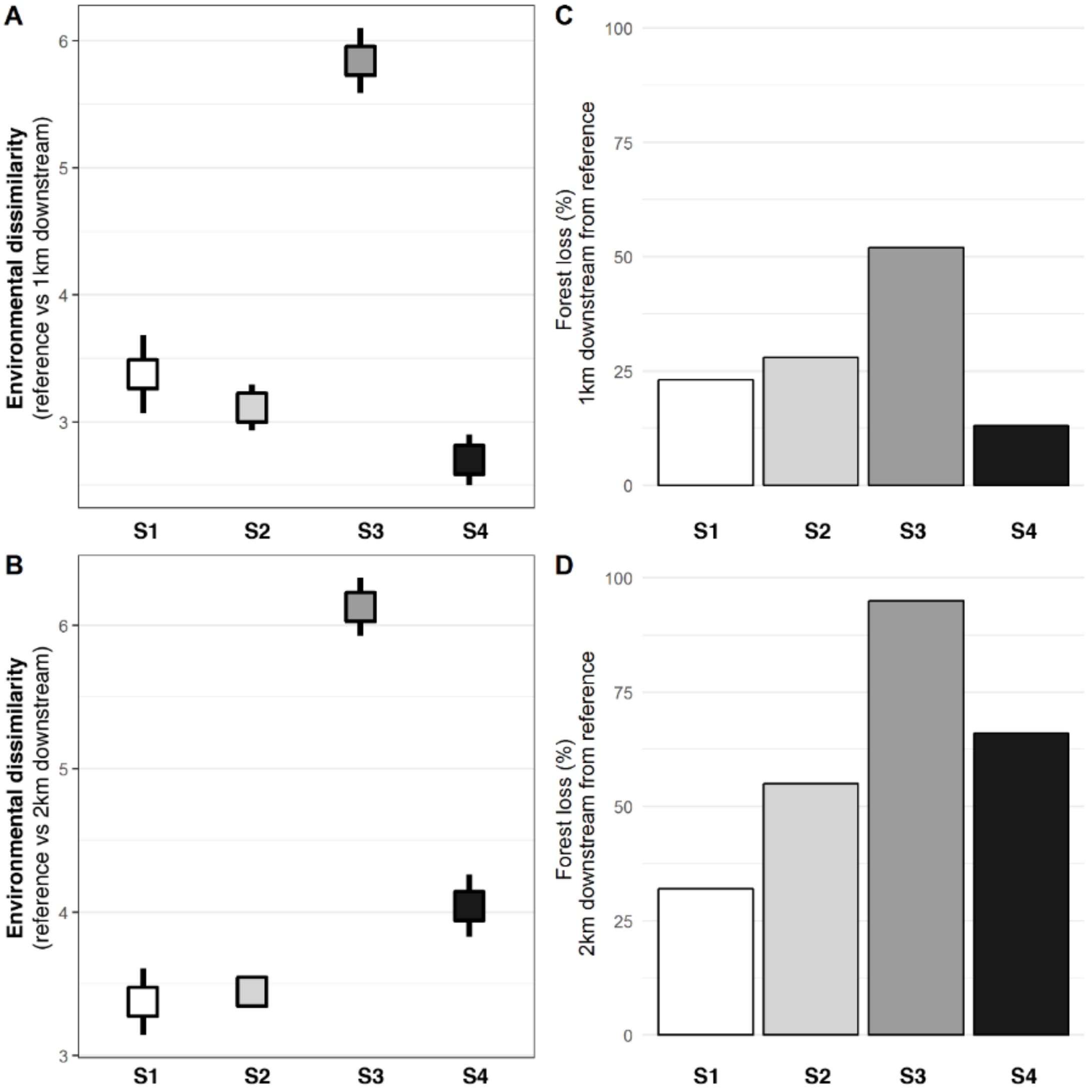


Figure V-4 On the left column, the mean and standard error (SE) of the Euclidean pairwise distances (representing environmental dissimilarity) between the undisturbed and the impaired reaches located 1km (A) and 2km (B) downstream. On the right column, the percentage of forest cover the undisturbed to the impaired reaches (1km - C, 2km - D) of each stream (S1-S4).

Table V-1 Range of values of the physical and chemical variables measured in the study streams.

	(S1) Chuscal			(S2) Chiflon			(S3) Tunjo			(S4) Zaque		
Reach	**Reference (Upstream)**	**Impact 1 (1km down)**	**Impact 2 (2km down)**	**Reference (Upstream)**	**Impact 1 (1km down)**	**Impact 2 (2km down)**	**Reference (Upstream)**	**Impact 1 (1km down)**	**Impact 2 (2km down)**	**Reference (Upstream)**	**Impact 1 (1km down)**	**Impact 2 (2km down)**
Altitude	2364	2271	2227	2528	2349	2368	2746	2544	2453	2953	2895	2861
Riparian forest (%) (buffer width -30m)	75.7	53.4	44.1	99.5	70.7	44.3	97.5	45.7	44.1	91.4	78.3	25.5
Canopy (%)	58.7-65.2^{a}	16.1-34.1^{b}	34.9-42.1^{a}	68.1-69.9^{a}	18.5-46.4^{b}	22.6-25.5^{c}	63.2-72.5^{a}	0.4-1.2^{b}	0-0.1^{c}	25.5-34.8^{a}	24.1-33.9^{a}	0.2-7.8^{a}
T (ºC)	13.4-13.5^{a}	15.8-16.4^{b}	15.7-16.4^{c}	11.5-11.9^{a}	11.4-12.4^{a}	14.2-15.5^{b}	9.3-9.8^{a}	13.8-14.2^{b}	15.8-16.4^{c}	11.1-12.1^{a}	11.1-12^{a}	10.2-10.4^{b}
pH	7.41-7.46^{a}	7.34-7.4^{a}	7.43-7.49^{a}	7.49-7.54^{a}	7.4-7.41^{a}	7.61-7.67^{a}	7.53-7.54^{a}	7.44-7.46^{a}	7.17-7.25^{a}	7.16-7.18^{a}	6.91-6.98^{a}	6.98-7.19^{a}
Sat.Ox (%)	96.2-96.9^{a}	96.2-97.6^{a}	96-96-8^{a}	95.3-95.8^{a}	94-95.5^{a}	95.6-96.6^{a}	94-96.6^{a}	95.1-95.5^{a}	94.9-96.2^{a}	95-95.8^{a}	95.2-95.8^{a}	92.8-94^{a}
Cond (uS/cm)	25.8-27.8^{a}	27.8-29.3^{a}	27.2-28.8^{b}	17.85-18-12^{a}	18.3-21.29^{a}	22.1-22.7^{b}	8.41-8.45^{a}	9.69-9.72^{a}	10.85-10.9^{b}	3.83-3.87^{a}	4.63-4.79^{a}	5.09-5.2^{a}
TSS (g/L)	0.0002^{a}	0.0032^{b}	0.0031^{c}	0.0001^{a}	0.0013^{b}	0.0026^{c}	0.0002^{a}	0.0251^{b}	0.0221^{c}	0.0019^{a}	0.00036^{b}	0.0024^{c}
BOM (mg.m2)	17-24.5^{a}	7.1-37.4^{a}	47.7-75.2^{a}	20.5-88.1^{a}	15.1-89.5^{a}	17-40^{a}	23.3-123.4^{a}	2.6-13.2^{a}	9.8-56.9^{a}	20.4-45.4^{a}	8.6-13.1^{a}	52.6-94.5^{a}
Channel depth (m)	0.22-0.39^{a}	0.17-0.26^{a}	0.17-0.26^{a}	0.14-0.18^{a}	0.09-0.13^{a}	0.13-0.19^{a}	0.18-0.22^{a}	0.18-0.34^{a}	0.18-0.26^{a}	0.17-0.23^{a}	0.26-0.31^{a}	0.2-0.25^{a}
Channel width (m)	3.9-5.7^{a}	2.7-5.8^{a}	4.5-6.2^{a}	2.9-4.5^{a}	2.5-3^{a}	2.5-4.8^{a}	2.5-5.6^{a}	3.7-4.3^{a}	4.4-7.2^{a}	6.6-8.4^{a}	3.2-5.3^{a}	3.55-5.1^{a}
Current velocity (m/s)	0.29-0.36^{a}	0.26-0.39^{a}	0.26-0.47^{a}	0.14-0.53^{a}	0.06-0.32^{a}	0.22-0.41^{a}	0.31-0.56^{a}	0.58-0.62^{a}	0.39-0.55^{a}	0.26-0.47^{a}	0.32-0.42^{a}	0.61-0.68^{a}

Lettering denotes significant differences between the undisturbed reach and the undisturbed reaches of each stream (pairwise t-test, $P < 0.01$, Bonferroni corrected).

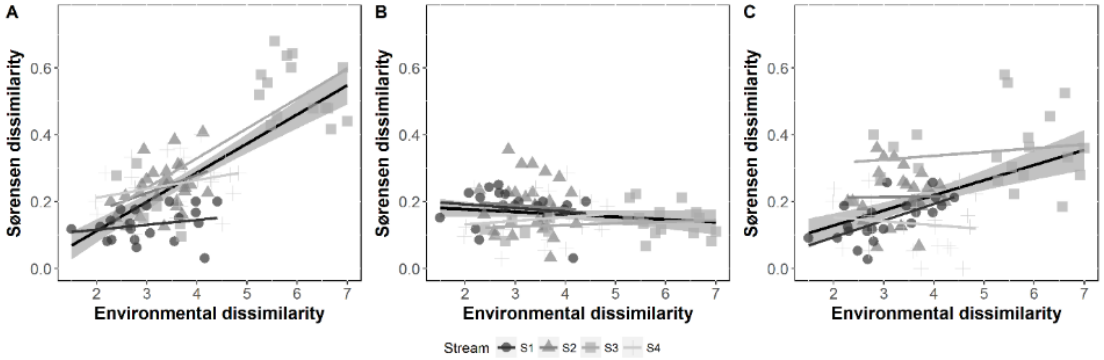


Figure V-5 Relationship among diatom (A), Chironomidae (B) and EPT (C) dissimilarities (Sørensen index; β_{SOR}) and the environmental dissimilarity among (black line) and within streams (color lines). Trend lines are modeled by using generalized least squares (GLS).

The relationships among the communities' beta diversity and the environmental dissimilarity caused by deforestation among-streams are represented in Figure V-5, while, within-stream relationships are summarized in Table V-2. We found that the diatom dissimilarity increased in line with environmental dissimilarity among-streams (Fig. V-5A, regression coefficient = 0.089, SE = 0.008, p-value < 0.001). Within-streams (Table V-2), the most pronounced change occurred at S3, and in a lesser degree in S2, which are the streams with a gradual loss of riparian coverage downstream of the reference reach. The Chironomidae communities did not significantly respond to the environmental changes driven by deforestation, neither in the within- nor among-stream comparisons (Fig. V-5B, regression coefficient = -0.008, SE = 0.005, p-value = 0.16). Finally, the dissimilarity of EPT communities increased in line with the environmental dissimilarity (Fig. V-5C, r = 0.045, SE = 0.009, p-value < 0.001).

Table V-2 Regression coefficients of the estimated relationships between diatom, Chironomidae and EPT dissimilarities (Sørensen index; β_{SOR}) and the environmental dissimilarity within streams.

Assemblage	Stream	Slope *(regression coefficient)*	SE	p-value	df
Diatoms	S1	0.017	0.012	0.17	25
	S2	0.067	0.025	0.01	25
	S3	**0.090**	**0.019**	**<0.001**	**25**
	S4	0.009	0.017	0.058	25
Chironomidae	S1	-0.011	0.014	0.47	25
	S2	-0.011	0.031	0.73	25
	S3	0.005	0.007	0.48	25
	S4	0.011	0.014	0.44	25
EPT	**S1**	**0.050**	**0.014**	**<0.001**	**25**
	S2	0.002	0.031	0.94	25
	S3	0.011	0.017	0.51	25
	S4	-0.012	0.018	0.52	25

*** In bold are represented those regressions with a p-value lesser than 0.01.**

Table V-3 summarizes the relationships between the environmental dissimilarity and the nestedness and turnover within each stream of the three communities (i.e. diatoms, Chironomidae and EPT). Nestedness and turnover of diatom communities increased among-streams. The higher dissimilarity value and the larger regression coefficient suggest that diatom dissimilarity was mainly driven by processes promoting nestedness. The turnover and nestedness components of Chironomidae communities significantly decreased with environmental changes only within the stream S4. Finally, in the EPT communities only the turnover component responded significantly to the environmental dissimilarity among-streams. Within-streams, the turnover and the nestedness components responded significantly at streams S1 and S3, respectively.

Table V-3 Regression coefficients of the estimated relationships between diatom, Chironomidae and EPT beta-diversity components (turnover and nestedness) and the environmental dissimilarity among and within streams.

Assemblage	Component		Stream	Slope *(regression coefficient)*	SE	p-value	df
Diatoms	Turnover (β_{SIM})	Among-streams	---	0.020	0.007	0.003	108
			S1	0. 013	0.011	0.27	25
			S2	0. 037	0.029	0.22	25
			S3	0. 032	0.015	0.05	25
			S4	0.008	0.016	0.61	25
	Nestedness (β_{NES})	Among-streams		0.069	0.009	<0.001	108
		Within-streams	S1	0. 004	0.009	0.68	25
			S2	0. 030	0.030	0.31	25
			S3	0. 059	0.028	0.05	25
			S4	0.002	0.010	0.86	25
Chironomidae	Turnover (β_{SIM})	Among-streams		0.0001	0.007	0.99	108
		Within-streams	S1	-0.011	0.018	0.56	25
			S2	-0.008	0.036	0.82	25
			S3	0.015	0.009	0.10	25
			S4	0.004	0.013	<0.001	25
	Nestedness (β_{NES})	Among-streams		-0.008	0.004	0.06	108
		Within-streams	S1	-0.0001	0.013	0.99	25
			S2	-0.002	0.013	0.87	25
			S3	-0.010	0.007	0.17	25
			S4	-0.031	0.012	0.01	25

Table V-3. *Continued*.

EPT	Turnover (βSIM)	Among-streams		0.046	0.009	<0.001	108
		Within-streams	S1	0.058	0.016	0.002	25
			S2	0.002	0.034	0.95	25
			S3	0.047	0.019	0.02	25
			S4	-0.015	0.009	0.11	25
	Nestedness (βNES)	Among-streams		-0.0007	0.006	0.90	108
		Within-streams	S1	-0.008	0.009	0.38	25
			S2	0.0001	0.020	0.99	25
			S3	-0.036	0.008	<0.001	25
			S4	0.004	0.018	0.83	25

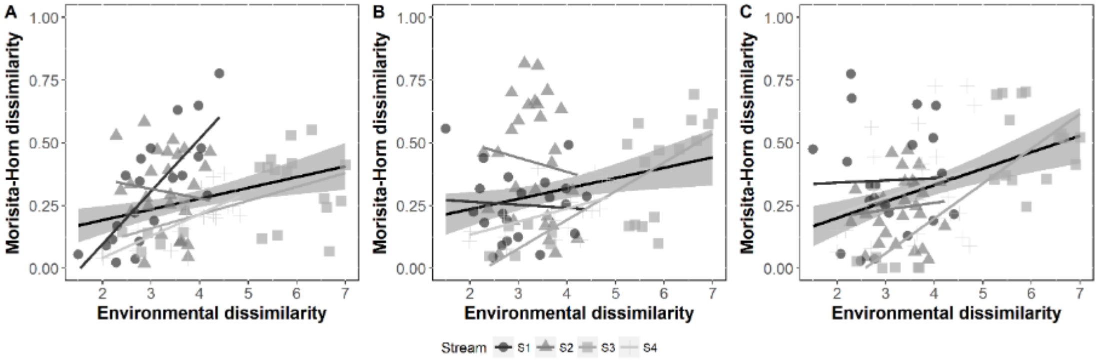


Figure V-6 Relationship among diatom (**A**), Chironomidae (**B**) and EPT (C) dissimilarities (Morisita-Horn index) and the environmental dissimilarity among (black line) and within streams (color lines). Trend lines are modeled by using generalized least squares (GLS).

The relationships between beta diversities and environmental dissimilarities became all significantly positive when considered the Morisita-Horn index (Figure V-6). Chironomidae communities exhibited the less pronounced changes in dissimilarity (Figure V-6B; r = 0.041, SE = 0.016, p-value = 0.013), followed by diatom (Figure V-6A; r= 0.043, SE = 0.013, p-value = 0.0019), and EPTs (Figure V-6C; r = 0.065, SE = 0.016, p-value = <0.001). Overall, the dissimilarity of the three communities increased with environmental dissimilarity within the S3 stream. We also observed a positive correlation within S4 stream for diatom and Chironomidae communities, and only within S1 stream for diatoms (Table V-4).

Table V-4 Regression coefficients of the estimated relationships between diatom, Chironomidae and EPT dissimilarities (Morisita-Horn index) and the environmental dissimilarity within streams.

Assemblage	**Stream**	**Slope** ***(regression coefficient)***	**SE**	**p-value**	**df**
Diatoms	**S1**	**0.210**	**0.040**	**<0.001**	**25**
	S2	-0.038	0.064	0.56	25
	S3	**0.055**	**0.021**	**0.01**	**25**
	S4	**0.087**	**0.015**	**<0.001**	**25**
Chironomidae	S1	-0.125	0.040	0.76	25
	S2	-0.056	0.094	0.56	25
	S3	**0.115**	**0.020**	**<0.001**	**25**
	S4	**0.052**	**0.020**	**0.01**	**25**
EPT	S1	0.009	0.063	0.89	25
	S2	0.028	0.049	0.57	25
	S3	**0.139**	**0.026**	**<0.001**	**25**
	S4	0.065	0.044	0.15	25

*** In bold are represented those regressions with a p-value lesser than 0.01**

DISCUSSION

The use of alpha-diversity and bioassessment-derived metrics relies on the assumption that biological communities track the environmental modifications caused by one or multiple stressors (Siqueira *et al.*, 2014b; Buss *et al.*, 2015). Here, we showed that three benthic communities (diatoms, non-biting midges and EPT) may not necessarily track the loss of forest cover occurring along high-Andean streams, even when the local environmental conditions changed as affected by this stressor. The distinct dispersal abilities of species from each community and the occurrence of upstream refugia in all streams provide enough evidence to hypothesize that dispersal may override the effects of deforestation, even when the forest loss is overwhelming (in the way it occurs in stream S3 or S4). Well-preserved upstream zones may act as source patches (refugia) of non-tolerant species that can persist in unsuitable conditions (“sink patches”) through dispersal-driven feedbacks.

The influence of dispersal-driven processes on diversity and distribution patterns has been already evidenced (reviewed in Tonkin et al. 2018). Briefly explained, at one end, lower dispersal ability can constrain species to track their preferred environmental conditions; and at the other, higher dispersal ability can override the constrains imposed by local environment conditions (Leibold *et al.*, 2004). So forth, our results support our hypothesis that dispersal ability can modulate the control exhorted by environmental filtering, and therefore the relationship between community (beta diversity) and environmental dissimilarities. However, it emerged in an unexpected way. We found that species with passive but higher dispersal ability (diatoms; Soininen 2007; Liu et al. 2013; Padial et al. 2014) were the best trackers of deforestation along the watercourse. Meanwhile, EPT and chironomid taxa, which a priori are active dispersers and can track for suitable conditions, were partially or not affected at all by deforestation.

Beta diversity patterns suggest different drivers underlying the responses of each community to deforestation. Environmental sorting can be the main driver of dissimilarity across communities in organisms with passive but higher dispersal ability, such as diatoms. Their beta-diversity tended to increase proportionally to the percentage of forest loss: the greater the loss, the greater the diatoms beta-diversity. This trend and the relatively higher importance of nestedness on their distribution suggest that deforestation is promoting the increase in species diversity via greater resource availability and habitat heterogeneity. In deforested reaches, the longer periods of light availability results in greater inputs of nutrients and diversity of microhabitats (i.e. filamentous algae) which may facilitate species coexistence by niche differentiation (Tilman *et al.*, 1982; Passy, 2008). In fact, we observed diatom guilds to expand from prevalent low-profile taxa with short stalks, or cells completely adhered at the substratum (i.e. *Achnanthidium*) in the undisturbed reaches, to others of higher profile with longer stalks or developing in mucilaginous tubes (such as *Gomphonema* and *Encyonema*) in the disturbed sections.

The insect communities did not track the loss of forest cover in the same manner that diatoms did. Their beta diversity within-streams was weak or unrelated to the environmental changes caused by deforestation. Taking into account that the two communities were made up by taxa with moderate to high dispersal abilities, downstream populations of non-tolerant species can be maintained by the flux of individuals from the upstream refugia patches (e.g. Downes et al. 2016). Depending on the community, the fluxes of individuals may override partially or completely the impact of deforestation at downstream reaches. Chironomidae communities, for instance, did not exhibit a change in their beta-diversity when used incidence data. Although non-biting midges are weak flyers, their larvae – especially from Orthocladiinae tribe – disperse commonly through the drift (Williams, 1989; Ríos-Touma *et al.*, 2011). Precisely, the Orthocladiinae is one of the dominant groups in Andean streams (Acosta & Prat, 2010; González-Trujillo & Donato-Rondon, 2016), and their constant flux from the upstream reaches could allow the persistence of non-adapted taxa under adverse conditions, and consequently, keeping lower dissimilarity within our study streams. This ‘mass effect’ process (Leibold *et al.*, 2004), can explain the null and weak changes in beta-diversity observed for incidence and abundance data, respectively. Mass effects maintain the occurrence of the major part of taxa along the river, but their abundances can diminish or increase according to their tolerance to the changes caused by deforestation. In addition, Chironomidae is a group composed by many species well-known as tolerant to environmental disturbance (e.g. Gresens et al. 2007), and consequently the stream impairment may affect less this assemblage than diatoms and EPT, driving a similar species composition among stream reaches.

The EPT communities exhibited a weaker relationship with environmental dissimilarity when used incidence data, but a stronger relationship when abundance data were used instead. It suggests that both the environmental conditions and the dispersal-ability are underlying the observed patterns. On one hand, the weaker relationship found when using incidence data provided some evidence that EPT communities are not dispersal-limited among- or within-streams. On the other hand, the strong relationship found by using the Morisita-Horn index indicates that, similar to chironomids, deforestation is probably not restricting taxa occurrence, but affecting their abundance according to their stress tolerance. The EPT communities are traditionally considered to be quite sensitive to environmental disturbances, because species usually track their preferred environmental conditions (Lenat, 1993). This may justify the stronger effects of deforestation on EPT dissimilarity than in Chironomidae among stream reaches.

However, the results of the within-stream relationships also suggest that deforestation effects on EPT communities could be context-dependent. We observed that the Sørensen dissimilarity increased according to the total amount of forest loss among-streams, but not within each stream (Fig. 4-5), and this suggests the existence of a threshold from which the composition of EPT taxa changes abruptly as affected by the environmental changes linked to deforestation. The absence of suitable places for laying eggs appears to be one of the major changes constraining EPT occurrence in deforested reaches. EPT taxa have been recorded in Andean, and worldwide streams, as active trackers of the best oviposition places, whose suitability is determined mainly by the substratum type, protruding and shading degree, and its associated water velocity (Hoffmann & Resh, 2003; Encalada & Peckarsky, 2006; Rios-Touma *et al.*, 2012; Lancaster & Downes, 2017). As those factors can be directly or indirectly affected by forest loss, the recruitment of EPT taxa can also be affected by the degree of deforestation. However, further understanding on oviposition patterns of these organisms is needed in order to disentangle the effect of deforestation on the recruitment process.

Implications for conservation

Our findings showed context-dependent responses to deforestation on high-Andean streams, where the type of community and the occurrence of upstream refugia can confound the assessment of this stressor. Along patchy corridors, as those of Andean streams, the existence of well-preserved upstream reaches can allow non-tolerant species to remain extant throughout dispersal-driven feedbacks. Therefore, not all communities might be adequate for assessing the impact of deforestation on high-Andean streams. Furthermore, considering both - the dispersal ability of bioindicators and the occurrence of refugia - is essential to achieve reliable conservation plans, given that they may mitigate negative local effects on biodiversity produced by human impacts such as deforestation (e.g. Bitušík et al. 2017).

The results reported here indicate three key aspects that need to be considered to enhance conservation and restoration efforts. Firstly, a deeper knowledge about the species traits (e.g. ability and mode dispersal) so our understanding of the effects of deforestation on biodiversity can be improved. Secondly, the conservation of upstream sections is a key aspect to consider for successful restoration or rehabilitation programs, as it has been previously suggested by Palmer et al. (1997). Thirdly, ecologists and conservationists should include beta-diversity measured through different indices incorporating species incidence and abundance, in addition to alpha-diversity metrics when studying deforestation effects on biodiversity.

VI. Historical legacies and contemporary processes shape beta diversity in Neotropical montane streams

ABSTRACT

Montane biodiversity is not homogeneously distributed in space. Here, we aimed to describe patterns of beta diversity in Neotropical montane streams and to define the roles that historical and contemporary factors likely play in determining them. We compared and modeled taxonomic-, functional- and phylogenetic- beta diversity. GLM models were fit to determine the major environmental, spatial and historical drivers of beta diversity and its turnover and nestedness in diatom and insect communities. pH, temperature and substratum mobility were included as environmental variables. We modeled the distances between streams in terms of altitude, latitude and longitude as spatial descriptors. Finally, we used a path-length matrix representing a hypothetical scenario of the basin's geological history as historical descriptor. We found that models explaining beta diversity of diatom and insect communities were fundamentally different regarding the number and type of significant predictors. However, we also found similar and significant effects of historical and environmental descriptors on the taxonomic and phylogenetic beta diversity patterns within each community. Functional beta diversity Patterns of diatoms and insects were significantly explained by temperature and pH gradients, respectively. The spatial structure did not influence any of the beta diversity facets. Overall, we conclude that tropical montane streams are ideal systems to bridge the gap between ecology and biogeography. We show that present-day patterns cannot be fully explained without incorporating the effect of past geological and climatic events on mountain landscapes. In the Neotropics, montane geology sets the stage for speciation and landscape formation, on which ecological (i.e. dispersal limitation) and environmental factors co-act to increase biodiversity.

INTRODUCTION

Montane biodiversity is not homogeneously distributed in space. One of the most well-known examples is the stepwise change in vegetation composition as elevation increases - described by Humboldt and Bonpland two centuries ago (von Humboldt & Bonpland, 1807). Such as vascular plants, most communities in montane landscapes exhibit a high degree of species turnover along elevation gradients (i.e. Wang *et al.*, 2012; Bishop *et al.*, 2015). While diverse lines of evidence indicate that this pattern is mainly shaped by elevation-driven environmental gradients (i.e. Bertuzzo *et al.*, 2016; Nottingham *et al.*, 2018), species turnover may also be partially or completely shaped by historical processes linked to past climatic (e.g. temperature oscillation) and geological (e.g. mountains uplift) events (Hazzi *et al.*, 2018). However, the study of this interplay between ecological and evolutionary processes has majorly focused in terrestrial communities (i.e. Anacker & Harrison, 2012), while is uncertain if it can be appropriate for freshwater communities. In these communities, contemporary constraints to dispersal and environmental conditions are usually invoked as major drivers of biodiversity patterns (Finn *et al.*, 2016; Encalada *et al.*, 2019).

Unlike terrestrial organisms, the dispersal of benthic stream organisms occurs within the dendritic-like structures of drainage basins (Brown & Swan, 2010). This means that the opportunity for them (including diatoms and some aquatic insects) to expand their distribution across two or more drainage basins depends on two major factors (Tonkin *et al.*, 2018): their dispersal traits (mode and ability), and the physical structure of stream networks. Multiple evidences show that the interplay of the two factors with environmental conditions drive species ranges size, and therefore species turnover patterns of benthic communities (i.e. Al-Shami *et al.*, 2013; Heino *et al.*, 2015; Datry *et al.*, 2017). However, further evidences (applied to European fishes) brought to light that contemporary distributions may also be the imprint of past historical events, such as the geohydrological process of river capture, or the confluence of river systems during low-sea-level periods (Dias *et al.*, 2014; Carvajal-Quintero *et al.*, 2019). The later has been described in Neotropical basins, where past geological and climatic events have shaped the contemporary structure of river networks and the distribution of biological communities (Hoorn *et al.*, 2010; Albert *et al.*, 2018; Bicudo *et al.*, 2019).

The assessment of beta diversity is essential for identifying key patterns and underlying mechanisms of species turnover at broad spatial extents (Qian *et al.*, 2005). Beta diversity not only quantifies the species turnover among local species assemblages, but may reveal the different forces operating in the assembly of communities at ecological and evolutionary timeframes (Devictor *et al.*, 2010). By incorporating functional and phylogenetic information, beta diversity metrics recognize a diverse array of life histories and different evolutionary histories (Cardoso *et al.*, 2014) within communities. This is especially useful for dissecting the role of historical legacies and contemporary processes on the assembly of biological communities. For example, while the functional dimension provides insight about the effect of contemporary processes, the phylogenetic dimension can reveal the signature of those processes at the evolutionary scale (Graham & Fine, 2008). Yet the influence of mountain ranges on all facets of beta diversity of benthic communities is yet to be explored.

In this study, we use beta diversity patterns to define the roles of historical and contemporary processes on benthic communities of diatoms and insects in tropical montane streams. These biological groups have disparate life histories, particularly regarding their dispersal mode and ability. Diatoms are sessile microorganisms whose small body size is usually related to a high dispersal capability (Padial *et al.*, 2014). Insects have different dispersal capabilities linked to their life histories and evolutionary background

(Shurin *et al.*, 2009; Saito *et al.*, 2015). While diatom communities may mainly respond to the contemporary environmental conditions rather than dispersal limitation factors (Passy, 2012), insect communities may respond to both environmental and dispersal constrains. In the later, there is some evidence that montane insect diversity is resulting from different allopatric events triggered by Andean uplifts (i.e. Chazot *et al.*, 2016), suggesting a greater effect of past history on insects dispersal limitation than on diatom communities.

Historical processes left an imprint on the structure of Orinocan stream metacommunities, shaping different pools of species which partially matched with the distribution of ecoregions (González-Trujillo *et al.*, 2020). We here aimed to test for the influence of historical legacies in the present distribution of both diatom and insect stream metacommunities. To that purpose, we built a hypothetical scenario for evolutionary history of the Orinocan ecoregions (Figure 1, see the Methods section for further details) and included it as a factor to model taxonomic, functional and phylogenetic beta diversity patterns, among other contemporary factors. In the case the evolutionary history of benthic communities followed the stepwise-like model of other montane communities (i.e. von Humboldt & Bonpland, 1807; Wang *et al.*, 2012; Bishop *et al.*, 2015), we assumed that historical legacies would mostly explain the taxonomic and phylogenetic beta diversity patterns. In that case, the closer the age of divergence between two ecoregions, the most similar the species and lineage composition between communities. Otherwise, if contemporary processes were the major drivers, environmental conditions and geographical distances would be the major determinants of beta diversity patterns. This should be particularly true for functional beta diversity, since the environment sets a template onto which species are selected according to their traits (Townsend & Hildrew, 1994).

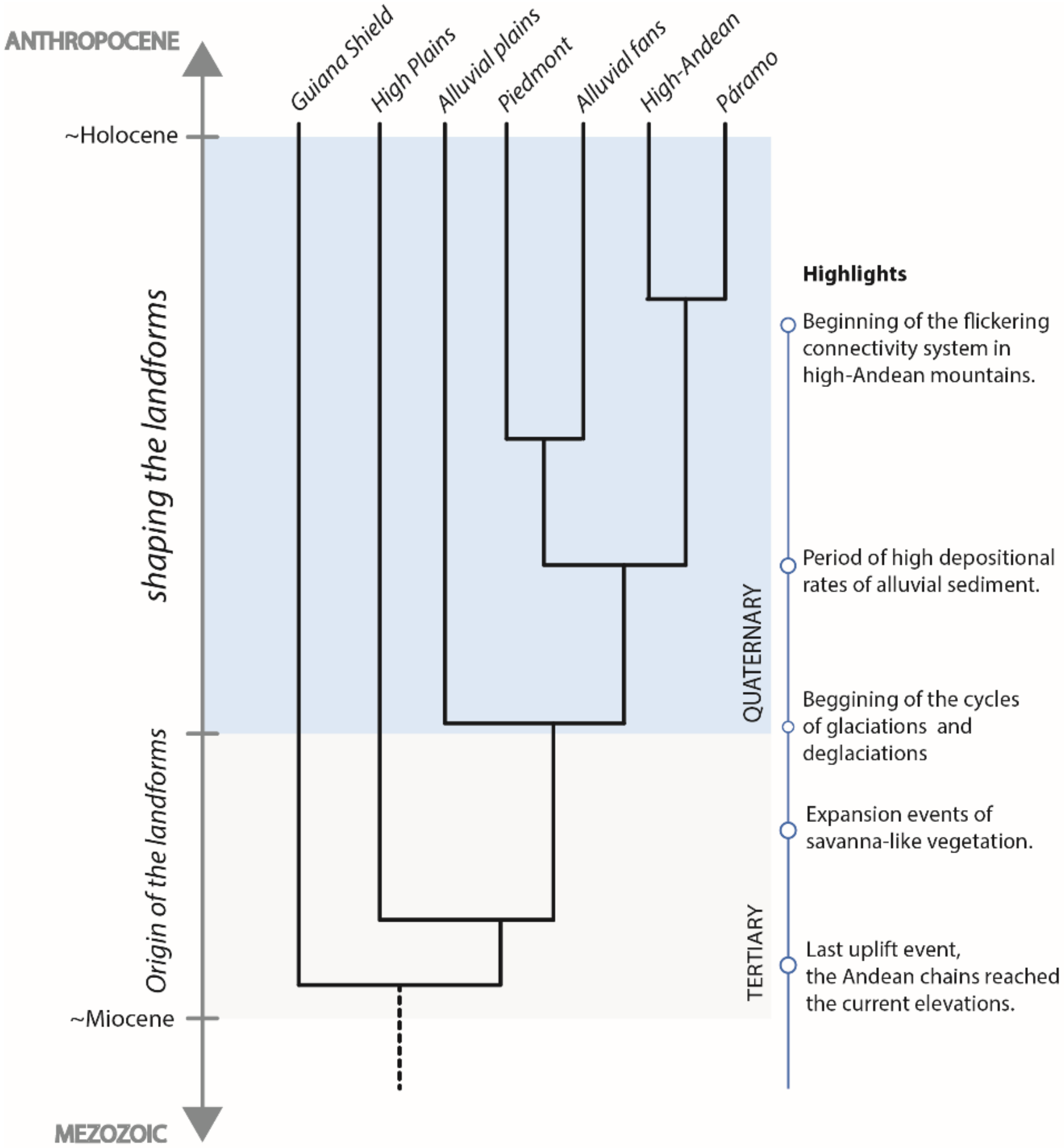


Figure VI-1 Cladogram that represents the path length matrix modeling the evolutionary history of ecoregions in the Orinoco basin. See chapter I for more details of its construction.

METHODS

Study area

This study was carried out in pristine or slightly modified streams within an area of about 40,000km^2 in the Colombian Orinoco basin. The Orinoco is the third largest basin in South America, and covers an area of about 990,000 km^2 in Venezuela and the eastern part of Colombia (Romero Ruíz *et al.*, 2004). Because the anthropogenic pressures received by these systems are still relatively small (though widening, see Sabater *et al.*, 2017), we expected that the effect of anthropogenic stressors is not confounding beta diversity patterns, as it can occur in more modified systems (i.e. Gutiérrez-Cánovas *et al.*, 2013).

The Orinoco basin hosts an intricate network of rivers and tributaries throughout three large relief forms in an altitudinal gradient from 0 to 3500m a.s.l.: the ancient massifs and shields, its recently raised ridges, and its tectonic depressions and sedimentary plains (Stallard, 1985). These reliefs were molded by the Andean orogenesis, and they are one of the major drivers of the in-stream habitat heterogeneity and water chemical characteristic, such as pH and conductivity (Stallard, 1985). In the case of the ancient massifs, for example, at least three weathering cycles (Goosen 1970) have occurred. Thus, rivers draining this relief have metamorphic rocks in their riverbeds and lower values of electric conductivity, and in general, lower concentrations of ions (the case of the Guiana Shield streams). On the other hand, the raised ridges shaped strong lateral influences which have driven increases in salinity at the basin lowlands (Gonzalez-Penagos *et al.*, 2014). However, these increases have not been uniform, since sedimentary plains are devoid of components associated with the meteoric water signature (Gonzalez-Penagos *et al.*, 2014).

Community sampling

In total, we sampled 26 (for diatoms) and 32 (for insects) stream segments during the dry season of 2016 (January-February). Details of the sampling sites can be found in Chapter I. Briefly, sites consisted of a segment including three consecutive riffles, representative of the stream's range of substrata types, flow velocities, channel widths and depths, and canopy cover. Sites were located in separated branches of different tributaries and scattered throughout the three large forms of Orinoco reliefs.

Diatom assemblages were collected in 8 cm^2 of surface, brush-scraped algal material from 30 boulders of similar size, in each of the stream sites (n=26). Algal material was pooled by riffle section (3 samples per stream segment) and subsequently preserved in a *Transeau* solution. Diatoms were identified at the finest level possible (usually morphospecies) using specialized monographs (Krammer & Lange-Bertalot, 1986, 1991; Metzeltin & Lange-Bertalot, 2007; Bellinger & Sigee, 2015). We assembled the 297 morphospecies into 76 taxonomic groups to alleviate taxonomic bias between samples since the identity of several taxa did not match the monograph descriptions (presumably new species).

Aquatic insects were sampled in stream segments using a multi-habitat sampling procedure, with 5 Surber (mesh size: 350 um; area: 0.09 m^2) samples in each of the riffle reaches (González-Trujillo *et al.*, 2020). Insects were sorted and identified to the genus level (including Diptera) in the laboratory, following Trivinho-Strixino & Strixino (1995), Domínguez & Fernández (2009) and González-Córdoba *et al.* (2015).

Pretreatment of community-level data

The considered grain sizes were a Surber sample for insects and a composed sample per riffle for diatoms. To create comparable datasets, we firstly pooled the abundances of each taxon per stream segment. Only taxa with counts greater than five individuals were kept in order to avoid the influence of very rare taxa. Two diatom morphospecies and 11 insect genera were then removed from datasets. Later, we converted the resultant data in matrices of presence-absence. The final data sets contained a total of 74 diatom morphospecies and 145 insect genera.

Functional aspects of the communities

The functional beta diversity of diatom communities was described using cell volume, growth form, attachment to substrata and motility (Table VI-I), which are documented as response traits to in-stream environmental conditions and dispersal-related constraints (i.e. Passy, 2007 and Lange, Townsend & Matthaei, 2016). Each trait was resolved as a categorical variable following microscope observations and the available literature (Lange *et al.*, 2016; www.algaebase.com).

In insects, functional beta diversity was described using seven traits: respiration mode, morphological adaptations to flow constraints, body size, feeding strategy, locomotion, aquatic stages and dispersal mode. As much as in diatoms, we focused on those traits accounting for taxon dispersal and responding to habitat heterogeneity, flow-related disturbance, and resource availability (i.e. Tomanova & Usseglio-Polatera, 2007; González-Trujillo, 2016). To describe each trait, we used field and laboratory observations as well as secondary information gathered from previous studies in the Neotropics (Tomanova & Usseglio-Polatera, 2007; González-Trujillo, 2016). In the field, we used diving masks to observe the type of locomotion. In the laboratory, we analyzed species mouthparts and gut contents to determine the feeding strategy. Each trait was resolved into different states and a fuzzing code procedure (Chevene *et al.*, 1994) was used to describe the affinity of each genus to each trait state. A score between 0 (no affinity) and 3 (high affinity) was assigned to each trait category. The affinity scores for each trait and genus were standardized so that their total sum per genus and trait was 1.

Table VI-1 Biological traits, and their respective states, used to estimate the functional diversity of diatom and insect communities.

Diatom community		
	Trait	**State**
A.	Cell volume	Nano (5 < 100 um^3), Micro (100 < 300 um^3), Meso (300 < 600 um^3), Large (600 < 1500 um^3), Very-Large (>1500 um^3)
B.	Growth form	Solitary, filamentous, colonial.
C.	Attachment to substrata	Postrate, Epyphite, non-attached, Pad-stalks, Mucilage tubes
D.	Motility	Drift, flag-glid, motile, non-motile
Insect community		
	Trait	**State**
A.	Respiration mode*	Spiracles – ‘spi’; Gas gills – ‘ggill’; Plastron and Spiracular gills – ‘pla:sgil’; Tegument – ‘teg’; Tracheal gills – ‘tgil’; blood-based respiration – ‘blo’
B.	Morphological adaptations to flow constraints	Case – ‘case’; Suckers – ‘Suc’; Silt glands – ‘silt’; abdominal claws – a.clw; tarsal claws – t.clw; no adaptation – ‘n.flow’
C.	Body size (At adult or last instars phases)	BS1 (<2.5mm); BS2 (2.5-5mm); BS3 (5-10mm); BS4 (10-20mm); BS5 (20-40mm); BS6 (>40mm)
D.	Feeding strategy	Collector-Gatherer – ‘CG’; Shredder – ‘SH’; Scrapper – ‘SC’; Filterer – ‘F’; Predator – ‘P’
E.	Locomotion	Surface swimmer – ‘SW’; Full water swimmer – ‘FW’; Crawler – ‘Cr’; Burrower/Interstitial – ‘Bu’; Temporary attached – ‘TA’; Permanently attached – ‘PA’
F.	Aquatic stages	Egg – ‘egg’; Larva/Nymph – ‘inmat’; pupa – ‘pup’; Adult – ‘adu’
G.	Dispersal mode	Aquatic passive – ‘AcP’; Aquatic active – ‘AcA’; Aerial passive – ‘AeP’; Aerial active – ‘AeA’

* Trait states are ordered according to their efficiency in oxygen uptake.

Phylogenetic aspects of the communities

We used the most inclusive time-calibrated diatom phylogeny to date (Nakov *et al.*, 2018), i.e. that reconstructed by maximum likelihood using 11 genes (18 S rRNA, 28 S rRNA, 16 S rRNA, *atp*B, *psa*A, *psa*B, *psb*A, *psb*C, *rbc*L, *cob* and *cox*I) and 38 calibration points. This phylogeny included almost all taxa of our diatom dataset (67 out of 74). The missing taxa were grafted to the most recent common ancestor of all members sharing the lowest taxonomic level available in the tree. This allowed each inserted taxon to be separated on the basis of its taxonomic name and then placed on an independent branch in the tree.

We used the supertree of Saito et al. (2015) as a basis for constructing our hypothetical phylogenetic tree for aquatic insects. This contains a large number of genera from our insect dataset (97 out of 145). Similar to the procedure for constructing the diatom supertree, the missing insect taxa were grafted to the most recent common ancestor of all members sharing the lowest taxonomic level available in the tree. To that purpose, we used information of age nodes compiled from several molecular phylogenies (Saito et al., 2015). When this information was not available, the topology was solved using taxonomy because it is a reasonable surrogate to phylogeny. In this regard, the Linnean classification system was used to establish relationships between sister groups. Further, we compiled the ages of 28 nodes and assigned the supertree branch lengths using the BLADJ algorithm in Phylocom. BLADJ spaces undated nodes evenly between dated nodes using an adjuster algorithm (Webb *et al.*, 2008). We used the Mesquite software (Maddison & Maddison, 2018) for editing the diatom and insect supertrees.

Beta diversity estimations

Taxonomic- functional- and phylogenetic-based beta diversities were estimated using Sørensen-based dissimilarity measures. Taxonomic beta diversity between stream communities was calculated using pairwise Sørensen dissimilarity (β_{SOR}). The beta diversity was further partitioned into the turnover (β_{SIM}; i.e. species replacement in the two communities) and the nestedness ($\beta_{SNE;}$ i.e. differences in species richness in the two communities) components (Baselga, 2010) because they represent different processes structuring dissimilarity among communities (Legendre, 2014).

Functional beta diversity (β_{FUNSOR}) was estimated by comparing the unshared functional space against the total functional space occupied by a pair of communities (Villéger *et al.*, 2013). Functional beta diversity was also further decomposed into turnover (β_{FUNSIM}) and nestedness (β_{FUNNES}) components. The functional space was estimated by calculating the Gower distances between each pair of genera in the community. After computing distances, a PCoA was carried out on the resultant distance matrix (Villéger *et al.*, 2008) and the first three axes were retained as informative descriptors of the functional space of each community (Maire *et al.*, 2015). In regard to diatom communities, the first three PCoA axes captured about 47% of the variation. This variation was split nearly evenly among the first (ca.19%), second (ca. 15%), and third (ca. 13%) axes. Given the loadings of the original traits, we interpreted the first three axes as describing the functional space of diatoms in terms of their successional stage and resistance to disturbances such as flow spates. The first three axes of the PCoA performed with the insect communities captured about 45% of the variation. This variation was mainly captured by the first axis (ca. 25%), and in a minor degree by the second (ca. 11%) and third (ca. 9%) axes. Given the loadings of the original traits, we interpreted the first three axes as describing the functional space of insects in terms of their specialization to occupy habitats with different water flow conditions, oxygen concentration and food availability (Supplementary Material).

Finally, phylogenetic beta diversity ($\beta_{PHYLOSOR}$) was calculated using the PhyloSor distance index. This was also decomposed into turnover ($\beta_{PHYLOSIM}$) and nestedness ($\beta_{PHYLOSNE}$) components. In the phylogenetic facet, the turnover component measures the 'true' lineage turnover, meanwhile the nestedness measures the differences linked to differences in Faith's Phylogenetic Diversity index (Leprieur *et al.*, 2012).

All the pairwise dissimilarity indices were calculated using R package 'betapart' (Baselga & Orme, 2012a).

Descriptors of beta diversity patterns

Environmental descriptors

Two respective sets of 4 and 5 environmental variables were used as determinants of the diatom and insect beta diversities. These variables were previously found as the major environmental explanatory variables of taxon occurrence and abundance in the Orinoco basin (Chapter III and IV). Those used as descriptors of the diatom beta diversity were water temperature, pH, reactive phosphorus (PRS) concentration and the number of days since the last spate events. The same variables except PRS were also used to model the beta diversity patterns of insect communities. In the insect models we additionally included substratum mobility (i.e. percentage of high mobility substrata in the streambed) and water oxygen saturation as variables.

Spatial descriptors

We calculated the distance between streams as a descriptor of the geographic distance. Distances were estimated in kilometers by using the longitudinal, latitudinal and altitudinal positions of each stream segments. This procedure was necessary because almost all sampled segments belong to independent micro-basins. Thus, the hydrological connection between two streams can pass by several stretches in other micro-basins, confounding the possible patterns on spatial structure. For instance, communities in two streams which are close to each other may be connected by aerial dispersal but not by aquatic dispersal.

Descriptors of basin's climatic and geological history

We built a path-length matrix (Legendre *et al.*, 1994) to account for the historical relationships between ecoregions. This matrix represents a hypothetical evolutionary scenario of the Orinoco basin, since it models the order of appearance of each ecoregion in the basin and the most plausible relationships among ecoregions regarding their origin (Figure IV-1). The matrix was constructed based on available geological, biological, historical and palynological evidences (van der Hammen, 1958, 1974; Goosen, 1971; Stallard, 1985; Flórez, 2003; Hubert & Renno, 2006; Hoorn *et al.*, 2010; Jaramillo *et al.*, 2010; Rangel-ch, 2016; Flantua & Hooghiemstra, 2018; Bicudo *et al.*, 2019; Flantua *et al.*, 2019). A detailed explanation of the Orinoco basin's evolutionary history and of the construction of the path length matrix is provided in the Chapter I.

Data analysis

The taxonomic, functional and phylogenetic pairwise dissimilarity was modelled against the geographic distance, the environmental dissimilarity matrices and the path-length matrix (which is modelling the evolutionary history of the basin) using a GLM approach with a Quasi-Binomial error distribution with a logit link function. As described by Dias *et al.* (2014), using this approach has three advantages: (1) it accounts for values found between 0 and 1, (2) it handles similarity values of 0 (minimal similarity) and 1 (maximal similarity) without previous transformation and (3) it accounts for the usually curvilinear shape of the distance–decay plots.

A problem linked to these models is that the cells of a similarity matrix are not statistically independent, making the “classical” tests invalid. To deal with this problem, we used a permutation method that consisted on permuting response variables (i.e. the rows and columns of benthic community matrices) while the original order of the exploratory variables (i.e. the environmental dissimilarity) was held constant. This method has been recommended over other permutation-based methods, and it has been recently used to test hypothesis in a similar scenario (Dias *et al.*, 2014). Under this method, the statistical significance was assessed using the frequency at which the null simulated coefficients were higher and/or lower than the observed coefficients.

RESULTS

The GLM results showed high values of pseudo-R^2 (ranging from 14 to 62%) and confirmed the importance of historical and contemporary factors in explaining the present-day patterns of beta diversity, particularly those taxonomic and phylogenetic (Tables IV-2 and IV-3). When significant, the effect (regression coefficient) of historical descriptors was not as strong as the effect of environmental descriptors, indicating that the sole effect of historical legacies or contemporary environmental conditions could not fully account for beta-diversity patterns.

Table VI-2 Effects of historical and contemporary factors on diatom communities' beta diversity. Beta diversity was estimated in terms of taxonomic, functional and phylogenetic dissimilarity. The standard errors of the slopes are indicated within parentheses. Historical descriptors refer to the path-length matrix representing basin's evolutionary history; geographic distance was log-transformed (km), while all environmental descriptors refer to Euclidean distance. PRS = reactive phosphorus concentrations.

Coefficients - Taxonomic beta diversity	**Diatoms**		
	Sørensen	Turnover	Nestedness
(intercept)	**-0.52 (0.13)**	**-1.19 (0.16)**	-1.70 (0.21)
Historical descriptors	**0.10 (0.02)**	**0.09 (0.03)**	0.04 (0.39)
Geographical distance	0.04 (0.06)	0.11 (0.08)	-0.13 (0.10)
Environmental descriptors			
Temperature	-0.001 (0.05)	-0.01 (0.06)	-0.02 (0.09)
pH	**0.21 (0.04)**	**0.30 (0.05)**	*-0.21 (0.07)*
# Days since the last spate	-0.03 (0.03)	-0.06 (0.04)	0.06 (0.06)
PRS	0.04 (0.03)	0.06 (0.03)	-0.06 (0.05)
Null deviance	29.71	41.36	26.61
Residual deviance	22.17	30.05	25.13
Pseudo-R2 V,adj (%) Zhang 2016	25.04	27.05	3.25
Pseudo-R2 (%)	25.38	27.35	5.56

Coefficients - Functional beta diversity	**Diatoms**		
	Sørensen	Turnover	Nestedness
(intercept)	-1.06 (0.17)	*-2.47 (0.23)*	-1.49 (0.23)
Historical descriptors	0.06 (0.03)	0.014 (0.03)	0.08 (0.04)
Geographical distance	0.001 (0.09)	0.24 (0.11)	-0.18 (0.11)
Environmental descriptors			
Temperature	0.05 (0.07)	*-0.24 (0.08)*	0.13 (0.10)
pH	-0.03 (0.05)	0.15 (0.06)	-0.19 (0.08)
# Days since the last spate	-0.003 (0.04)	-0.03 (0.05)	0.02 (0.06)
PRS	-0.07 (0.04)	- 0.004 (0.04)	-0.12 (0.06)

Table VI-2. *Continued.*

Null deviance	28.94	24.86	38.5
Residual deviance	28.05	23.69	36.75
Pseudo-R2 V,adj (%) Zhang 2016	1.01	2.99	2.25
Pseudo-R2 (%)	3.08	4.71	4.55

Coefficients - Phylogenetic beta diversity	**Diatoms**		
	Sørensen	Turnover	Nestedness
(intercept)	**-0.84 (0.12)**	**-1.69 (0.15)**	-1.68 (0.22)
Historical descriptors	**0.08 (0.02)**	0.04 (0.02)	0.09 (0.04)
Geographical distance	0.08 (0.06)	*0.21 (0.07)*	-0.20 (0.09)
Environmental descriptors			
Temperature	-0.03 (0.05)	-0.06 (0.06)	0.001 (0.09)
pH	0.10 (0.04)	**0.20 (0.04)**	*-0.21 (0.07)*
# Days since the last spate	-0.03 (0.03)	-0.06 (0.03)	0.04 (0.06)
PRS	-0.01 (0.28)	0.04 (0.03)	-0.12 (0.06)
Null deviance	20.58	26.45	29.59
Residual deviance	17.55	21.07	27.67
Pseudo-R2 V,adj (%) Zhang 2016	13.07	18.48	3.71
Pseudo-R2 (%)	13.47	19.43	6.49

italic a<0.05; **bold** a<0.01

Table VI-3 Effects of historical and contemporary factors on insect communities' beta diversity. Beta diversity was estimated in terms of taxonomic, functional and phylogenetic dissimilarity. The standard errors of the slopes are indicated within parentheses. Historical descriptors refer to the path-length matrix representing basin's evolutionary history; geographic distance was log-transformed (km), while all environmental descriptors refer to Euclidean distance.

Coefficients - Taxonomic beta diversity		**Insects**		
		Sørensen	Turnover	Nestedness
(intercept)		**-0.64 (0.08)**	**-1.32 (0.09)**	-1.79 (0.20)
	Historical descriptors	**0.13 (0.01)**	**0.16 (0.02)**	0.004 (0.03)
	Geographical distance	-0.08 (0.04)	-0.04 (0.05)	-0.09 (0.10)
Environmental descriptors	Temperature	**0.15 (0.03)**	**0.23 (0.04)**	-0.20 (0.09)
	pH	**0.21 (0.02)**	**0.11 (0.02)**	*0.24 (0.06)*
	# Days since the last spate	-0.01 (0.02)	*-0.04 (0.02)*	0.08 (0.05)
	Oxygen saturation	-0.03 (0.03)	- 0.003 (0.03)	-0.12 (0.08)
	Substratum mobility	**0.13 (0.02)**	**0.20 (0.02)**	-0.20 (0.06)
	Null deviance	36.74	51.23	44.53
	Residual deviance	13.93	19.29	38.08
	Pseudo-R2 V,adj (%) Zhang 2016	62.46	63.25	14.67
	Pseudo-R2 (%)	62.08	62.35	14.48

Coefficients - Functional beta diversity		**Insects**		
		Sørensen	Turnover	Nestedness
(intercept)		- 1.09 (0.17)	**-3.67 (0.28)**	-1.20 (0.21)
	Historical descriptors	0.02 (0.03)	0.13 (0.04)	-0.007 (0.04)
	Geographical distance	-0.06 (0.08)	0.20 (0.13)	-0.13 (0.10)
Environmental descriptors	Temperature	- 0.02 (0.07)	-0.26 (0.10)	0.08 (0.08)
	pH	**0.51 (0.05)**	-0.13 (0.07)	**0.64 (0.06)**
	# Days since the last spate	0.003 (0.04)	0.003 (0.05)	0.004 (0.05)
	Oxygen saturation	-0.14 (0.06)	0.15 (0.08)	-0.22 (0.07)
	Substratum mobility	-0.03 (0.05)	0.15 (0.06)	-0.10 (0.06)

Table VI-3. *Continued.*

Null deviance		82.06	36.11	108.56
Residual deviance		63.28	30.71	83.38
Pseudo-R2 V,adj (%) Zhang 2016		23.63	15.11	25.47
Pseudo-R2 (%)		22.89	14.95	23.19
Coefficients - Phylogenetic beta diversity		**Insects**		
		Sørensen	Turnover	Nestedness
(intercept)		**-0.99 (0.07)**	**-1.79 (0.09)**	-1.96 (0.20)
	Historical descriptors	**0.11 (0.01)**	**0.14 (0.02)**	0.01 (0.03)
	Geographical distance	-0.08 (0.03)	0.003 (0.04)	-0.14 (0.10)
Environmental descriptors	Temperature	**0.12 (0.03)**	**0.20 (0.03)**	-0.15 (0.08)
	pH	**0.20 (0.02)**	-0.01 (0.02)	**0.45 (0.06)**
	# Days since the last spate	0.01 (0.01)	-0.005 (0.02)	0.04 (0.05)
	Oxygen saturation	-0.07 (0.02)	-0.05 (0.03)	-0.09 (0.07)
	Substratum mobility	**0.09 (0.02)**	**0.18 (0.02)**	-0.17 (0.06)
Null deviance		22.65	29.38	44.66
Residual deviance		10.05	13.24	35.81
Pseudo-R2 V,adj (%) Zhang 2016		56.11	56.46	23
Pseudo-R2 (%)		55.63	54.94	19.82

italic a<0.05; **bold** a<0.01

As expected from the different life histories of diatoms and insects, the GLM results indicated dissimilar explanatory power (pseudo-R^2) and different explanatory factors driving each facet of beta diversity (Table VI-2 and VI-3). Overall, we found that GLM models performed better (higher pseudo-R^2 values) for insect than diatom communities. This was the case for the three metrics: taxonomic (62% versus 25%), functional (23% versus 3%) and phylogenetic (55% versus 14%). GLM models also indicated a partial congruence between diatom and insect beta-diversity patterns. We found that historical descriptors and pH were the only factors having a similar effect on the beta-diversity patterns of both communities. Particularly, historical descriptors were found to similarly explain the variability in taxonomic- (both Sørensen and turnover) and phylogenetic-based (only the Phylosor) beta-diversities. Meanwhile, pH was found to similarly explain the variability in taxonomical beta diversity (Sørensen and turnover indices).

Congruence among beta-diversity metrics within diatom communities

Functional-based patterns of diatom beta diversity were not concordant with the taxonomic- and phylogenetic-based patterns. While taxonomic and phylogenetic beta diversity of diatom communities were equally explained by the basin's history, this was not the case of the functional beta diversity (Table VI-2). When analyzed separately, we also found non-congruent patterns between the turnover and nestedness components of the three facets: basin´s history explained the taxonomic turnover, whereas water temperature explained the functional turnover (though with a low explanatory power; 5%). Overall, only the effect of pH was consistent on the taxonomic and phylogenetic facets: turnover increased with pH-based dissimilarity, while nestedness decreased (Figure VI-2 B-C, H-I).

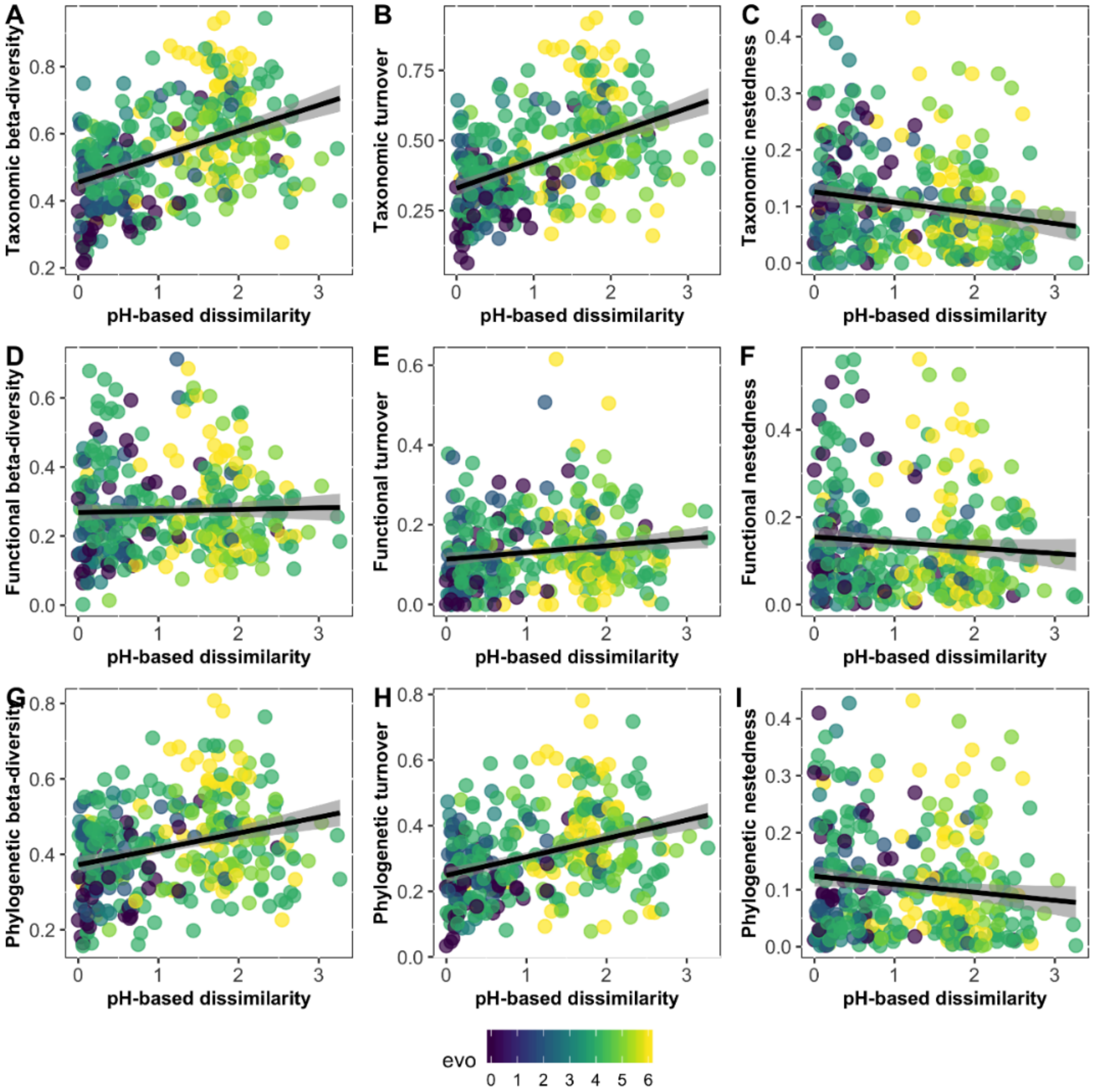


Figure VI-2 Relationship of beta diversities with environmental distance (pH-based) and historical distance (dot color) in diatom communities. (A-C) Taxonomic beta diversity and its turnover and nestedness components. (D-F) Functional beta diversity and its turnover and nestedness components. (G-I) Phylogenetic beta diversity and its turnover and nestedness components.

Congruence among beta-diversity metrics within insect communities

We observed a high concordance between the drivers of taxonomic and phylogenetic beta diversity (Table VI-3, Figure VI-3). However, the functional facet was not explained by the same factors. The taxonomic and phylogenetic facets were significantly explained by basin's history and a set of environmental descriptors (i.e. pH, temperature and substratum mobility), being stream water pH the factor with the strongest effect. After decomposing the three facets of beta diversity, our results indicated a similar effect of water temperature, substratum mobility and basin's history on the taxonomic and phylogenetic turnover (Table VI-3). Water pH was a good descriptor of the taxonomic turnover as well as of the nestedness component of the three facets (Figure 3 VI-E-F).

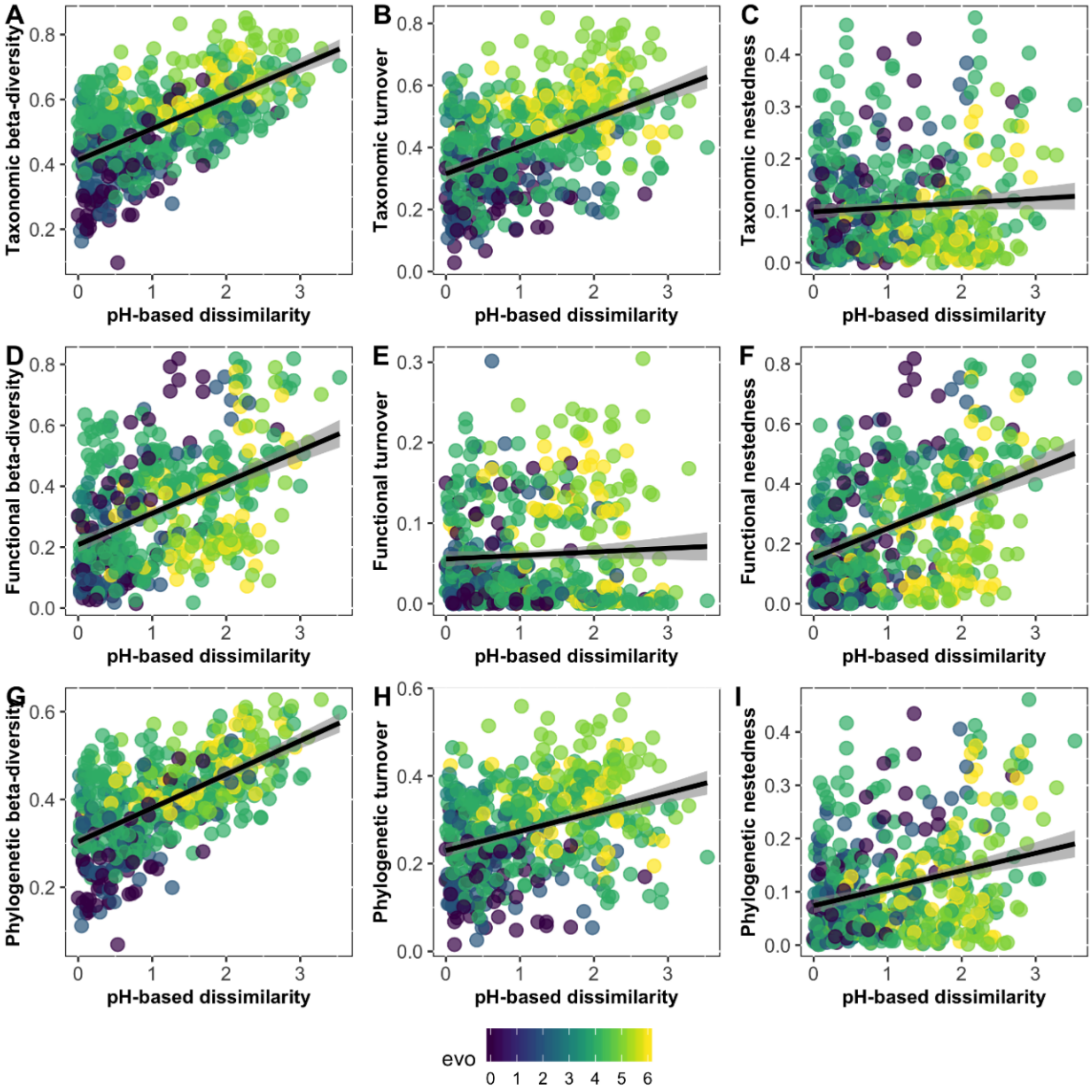


Figure VI-3 Relationship of beta diversities with environmental distance (pH-based) and historical distance (dot color) in insect communities. (A-C) Taxonomic beta diversity and its turnover and nestedness components. (D-F) Functional beta diversity and its turnover and nestedness components. (G-I) Phylogenetic beta diversity and its turnover and nestedness components.

DISCUSSION

Contemporary factors have been traditionally invoked to explain the turnover of species across tropical montane ecosystems (i.e. Anderson & Maldonado-Ocampo, 2011; Wang *et al.*, 2012; Liu *et al.*, 2018). However, under some scenarios, historical imprints may explain beta diversity patterns as much as present-day conditions. In this regard, we confirmed that beta diversity patterns of diatom and insect benthic communities could not be fully understood without considering the imprints of past geological and climatic events. This was the case mostly of taxonomic and phylogenetic beta diversity. Although GLM results pointed out to contemporary factors as the major drivers of the three facets of beta diversity, some of these factors (such as water pH) cannot be considered as totally independent from the historical background of the region. Our hypothetical reconstruction of the geological history of the Orinoco basin provides some insight to comprehend how ecological factors interacted with evolutionary forces to shape montane diversity in the Tropics.

Historical imprints on taxon distribution: the significance of the congruence between taxonomic and phylogenetic facets

Elevational patterns of species richness are associated to environmental controls, mostly associated to fluvial geomorphology (Bertuzzo et al. 2016). We observed a similar pattern in tropical montane systems, on which environmental controls are necessary to define beta diversity patterns. Besides, these controls were not sufficient to understand the beta diversity patterns in the Orinoco mountain range. We found that past historical events had a comparable effect to that of the environmental factors on the taxonomic and phylogenetical facets of beta diversity. Turnover occurs mostly between groups of lineages rather than following a random replacement of species (Leprieur *et al.*, 2012). This configures lineage groups that may have passed through similar processes within an evolutionary timeframe (e.g. species isolation or habitat expansion) and currently are part of the same species pool.

Overall, our explicit modeling of the historical background provides evidence supporting that past historical events influenced on control environmental conditions and geographical barriers, and also were essential on separating lineages (and taxa) in different ecoregions regardless of the long time available for dispersal (thousands or millions of years). These findings can be explained in the context of the evolutionary species pool hypothesis (SPH) (Pither & Aarssen, 2005), which states that the longer a set of environmental conditions persists in space and time, the greater the opportunity for the evolution of adapted species (Pither & Aarssen, 2005). Consequently, a higher number of species is expected to be shared (and beta diversity reduced) between regions that have had similar environmental conditions for a longer time. González-Trujillo et al. (2020) already brought to light the occurrence of several pools of taxa in the Orinoco basin, and we now prove that the taxonomic and phylogenetic turnover increase with weaker historical relationships among those pools (which partially match ecoregions' distribution).

The observed effect of pH on beta diversity of diatoms and insects in montane streams can also be seen as an expression of past historical events on contemporary patterns of beta diversity. Stream water pH is a derivative expression of the geochemical background of flowing waters throughout a basin. In this regard, pH values are highly dependent on the geological age of each ecoregion (Goosen, 1971; Stallard, 1985; Gonzalez-Penagos *et al.*, 2014). Lower pH values usually occur in the oldest ecoregions because of the high number of weathering cycles these have experienced (i.e. Guiana shield, Goosen 1971). Thus, it was not surprising to find that the lowest dissimilarity between communities occurred in ecoregions with the lower pH concentrations (the oldest), and the highest beta diversity occurred under more ‘benign’ (or more recent) pH concentrations. This finding supports Chase (2003, included in the SPH as mechanistic evidence) assumption that stable conditions in space and time favors multiple ways for communities to be assembled.

Our analysis also shows that historical legacies cannot be fully disentangled from environmental contemporary effects. Different number of contemporary factors were found to significantly account for beta diversity patterns, particularly in the case of insect communities. While this may be interpreted as evidence supporting that environmental sorting is majorly driven the assembly process, we posit that these results cannot be interpreted without considering the historical background of the basin. Stream water pH, temperature and substratum mobility cannot be considered as factors independent from historical legacies. In Neotropical streams, temperature is majorly driven by altitudinal ranges (Gill *et al.*, 2016) and the substratum mobility is highly dependent on the fluvial landform (Stallard, 1985). Therefore, in line with Perrigo *et al.* (2019), our findings emphasize that montane geology sets the stage for speciation and landscape formation, where different ecological and environmental factors co-act to increase biodiversity.

Insights on other drivers of the assembly process: the significance of the non-congruence of the functional beta diversity

Two reasons, not mutually exclusive, can support the non-congruence of functional and taxonomic/phylogenetic beta diversity patterns. The first one is that the included traits do not directly relate to the descriptors of our models. Bishop et al. (2015) observed a similar discrepancy in ant communities of Tropical mountains, asserting that the functional traits were probably unrelated to the contemporary or historical factors shaping taxonomic beta diversity patterns. As much as in this former study, we considered traits which mainly respond to characteristics of the local habitat (see Tomanova & Usseglio-Polatera, 2007; González-Trujillo, 2016), while others directly linked to taxon dispersal or temperature preferences could not be included due to the lack of primary data. It is likely that including other traits would help to disentangle the main drivers of functional diversity across different scales. In this regard, dispersal traits could help to explain the phylogenetic distance of Neotropical stream metacommunities (Saito *et al.*, 2015). Meanwhile, traits related to temperature tolerances could help to explain the patterns of richness distribution along mountain passes in the Andes (Gill *et al.*, 2016).

The second reason behind the non-congruent pattern of functional beta diversity lies in a high similarity between the functional spaces of benthic communities. It could be argued that the functional spaces may be similar irrespective of the geological history of the basin, being the functional dissimilarity more related to local environmental characteristics For example, functional dissimilarity among Neotropical invertebrate communities was the highest at the mesohabitat scale, driven by variations in water turbulence and benthic organic matter (Tomanova and Usseglio-Polatera, 2006). However, dispersal may override the strong control exerted by the features of the local environment at the stream or the micro-basin scales (González-Trujillo *et al.*, 2019), stressing the need for a better knowledge of dispersal-related traits to better understand the factors shaping the functional biogeography of these species.

Implications for conservation

Montane rivers in the tropics host high biodiversity and generate important benefits to human societies, yet still are inadequately studied (Encalada et al. 2019). Generating more primary data (Perrigo *et al.* 2019), and better integrating geological and climatic history into biological models of diversity and evolution are pressing needs. We contributed to both gaps by (i) synthetizing and testing a simple model representing an hypothetical scenario of the geological history of the basin; (ii) demonstrating that the understanding of contemporary patterns requires to assess for the effect of past geological and climatic events, in the manner already done in terrestrial communities (i.e. Bicudo *et al.*, 2019; Flantua *et al.*, 2019); and (iii) showing that tropical montane streams can be ideal systems to dissect the role that ecological and evolutionary forces may play in the community assembly process. Overall, we showed that knowing the historical background of a region is essential to gain better understanding of the mechanisms supporting community-level patterns.

.

VII. Concluding remarks and perspectives

Throughout this dissertation, I have focused on using a ***biogeographical oriented metacommunity framework*** to dissect the relative influences of *dispersal* (in ecological and evolutionary timeframes), *selection* (driven by abiotic factors) and *ecological drift* on the assembly process of benthic communities. By using a pattern-matching approach, I provide evidences supporting that the three processes are directly involved in the assembly of benthic communities. Notwithstanding, their role and influence in the assembly process is scale dependent.

I found evidence supporting *dispersal* as the major driver of taxon distribution; and thus, community composition at large scale. Particularly, the results indicate that one or more events of dispersal limitation (i.e. allopatric isolation) in an evolutionary timeframe shaped distinct pools of taxa in the basin. The extent of these pools partially matches the distribution of the ecoregions, suggesting that those events molding the riverscapes and the vegetation structure affect similarly the diversity and distribution of benthic species.

At the scale of each ecoregion, *dispersal, selection* and *drift* interact to constrain the structure and dynamics of communities and metacommunities among and within streams. Depending on the taxa belonging to each pool of species, the role of one of those processes may be greater respect to the others. For example, selection seems to have a greater effect on the composition of diatoms than on the composition of insects from high-Andean streams.

So forth, these findings have implications for the basic and applied research in the disciplines of metacommunity and freshwater ecology as well as of conservation and biogeography, which I detail below.

Implications for the ecology of stream (meta)communities

The major thesis of this study is that historical events have an important role in the assembly of metacommunities at the basin scale. Past climatic and geological shifts have shaped the contemporary distribution of benthic communities, particularly throughout limiting species dispersal within evolutionary timeframes.

In riverine ecosystems, the basin has been used as an appropriate scale to study the causes and consequences of biological diversity (Brown & Swan, 2010; Grönroos et al., 2013; Cañedo-Argüelles et al., 2015). However, the basin did not accurately reflect the assemblages ‘structure in highly heterogeneous basins such as the Orinoco basin.

I have shown that the occurrence of different regional pools can override the effects of other factors previously described as determinants of metacommunity structure (e.g. local environment or the spatial structure of tributaries). Similar to previous findings in algal communities (Benito et al. 2018), neotropical species distribution seemed to be significantly restricted to each ecoregion or biogeographic district. Hence, the communities that we studied do not constitute a metacommunity in an ecological timeframe or in the basin extent.

Conversely, the Orinoco basin can be viewed as a mosaic of functional metacommunity units, with the structure of each unit varying according to the taxa belonging to each species pool and the environmental context at each ecoregion. Within each ecoregion, *selection, dispersal* and *drift* processes are interacting to shape the structure of the metacommunity in the present time. In light of this, biomonitoring plans should be adapted to every ecoregion, or at least, match the extent of species pools.

Implications for the conservation of stream (meta)communities

A basic premise of conservation biology is that habitat units have unequal biological value; some are more important than others for maintaining total biodiversity and thus have a higher priority for protection. In regard to riverine ecosystems, the terminal segments of the riverine networks have been considered as the units with the highest biodiversity; and therefore, units of greater conservation value (Finn *et al.*, 2011b).

This thesis brought to light that before establishing a conservation unit in tropical basins, a first step should be determining the extent of the pool of species, and their relative influence on metacommunity structure. This agrees with the observation of Poiani et al., (2000), who reported that the units for conservation should be delineated according to the context of each basin or sub-basin. Our results indicate that such delineation should consider the type and extent of the different ecoregions in the basin, given that ecoregions harbor distinct pools of species because of historical contingencies. Therefore, they can be considered as functional metacommunity units for assessing biodiversity patterns and designing effective conservation actions.

Perspectives for future studies

The understanding of metacommunity dynamics in riverine ecosystems has been greatly enhanced by considering rivers as dendritic networks (Tonkin et al., 2018). However, few studies have addressed how the occurrence of several pools of species can modulate the control exerted by the riverine network in a basin (e.g. Heino et al., 2015b; Tonkin et al., 2015; Brasil et al., 2017). Considering the main findings of this dissertation, it will be relevant to test if historical imprints remain or not in a small- to medium-size riverine network draining.

The observed patterns also suggest that understanding the dispersal ability is key to uncovering the main drivers of the structure and dynamics of some assemblages of benthic communities (i.e. Chironomidae or Trichoptera). For example, chapter IV showed that considering both - the dispersal ability of bioindicators and the occurrence of refugia - is essential to achieve reliable conservation plans, given that they may mitigate negative local effects on biodiversity produced by human impacts such as deforestation. Therefore, characterizing dispersal traits – among other biological traits – at the genus or species level is a necessary step towards understanding the assembly process at mid-level scales.

It was someway unexpected to observe evidence of phylogenetic signal in our analyses of beta diversity. This finding indicates that the taxa belonging to each pool are phylogenetically related, also suggesting that they could have a similar origin and evolutionary history. Testing the feasibility of this suggestion seems to be a logical step towards understanding the causes of diversity in Neotropical streams. Studies on population genetics, phylogeography and molecular biogeography (i.e. Bonada *et al.*, 2009; Múrria *et al.*, 2015; Finn *et al.*, 2016) will help to deal with this research gap as well as establishing new brides between biogeography and metacommunity ecology. In this context, paleoecological reconstructions based on benthic species (e.g. diatoms or chironomids) will be also helpful to disentangle the interplay between ecological and evolutionary processes.

Conclusions

- I provide evidences supporting that *dispersal*, *selection* and *drift* are directly involved in the assembly of benthic communities. However, their role and influence in the assembly process is dependent on the type of community (diatoms or invertebrates) and on the scale considered.

- Dispersal was the major driver of taxon distribution; and thus, the factor determining the metacommunity structure at the scale of the Orinoco basin. Our findings indicate that one or more events of dispersal limitation (i.e. allopatric isolation) in an evolutionary timeframe shaped distinct pools of taxa in the basin.

- The regional species pools partially matched the distribution of the ecoregions, suggesting that events molding the riverscapes and the vegetation structure likely affected the diversity and distribution of benthic species in the basin.

- Orinoco freshwater communities do not constitute a single metacommunity within a given ecological timeframe. Conversely, the freshwater community in the Orinoco basin can be viewed as a mosaic of functional metacommunity units, with the structure of each unit varying according to the taxa belonging to each species pool and the environmental context in each of the ecoregions.

- Depending on the assemblages (diatoms vs invertebrates), the role of *dispersal*, *selection* or *drift* may become preferential respect to the others. In the high-Andean ecoregion, for instance, while *selection* had a greater effect on the composition of diatoms, *dispersal* had the highest role on the composition of insect assemblages. This means that Andean diatom metacommunities are structured by environmental sorting processes; and therefore, they are more prone to respond to human-driven impacts on streams and rivers. On the other hand, Andean insect metacommunities are structured by dispersal-driven processes, such as Mass effects or Patch dynamics. This means that local communities can persist in time under the effect of one or multiple stressors because of the fluxes of non-adapted species coming from 'source' communities.

References

Acosta, R. & Prat, N. (2010) Chironomid assemblages in high altitude streams of the Andean region of Peru. *Fundamental and Applied Limnology / Archiv für Hydrobiologie*, **177**, 57–79.

Al-Shami, S.A., Heino, J., Che Salmah, M.R., Abu Hassan, A., Suhaila, A.H. & Madrus, M.R. (2013) Drivers of beta diversity of macroinvertebrate communities in tropical forest streams. *Freshwater Biology*, **58**, 1126–1137.

Albert, J.S., Val, P. & Hoorn, C. (2018) The changing course of the Amazon River in the Neogene: center stage for Neotropical diversification. *Neotropical Ichthyology*, **16**.

Altermatt, F. (2013) Diversity in riverine metacommunities: A network perspective. *Aquatic Ecology*, **47**, 365–377.

Altermatt, F. & Fronhofer, E.A. (2018) Dispersal in dendritic networks: Ecological consequences on the spatial distribution of population densities. *Freshwater Biology*, **63**, 22–32.

Altermatt, F., Seymour, M. & Martinez, N. (2013) River network properties shape α-diversity and community similarity patterns of aquatic insect communities across major drainage basins. *Journal of Biogeography*, **40**, 2249–2260.

Anacker, B.L. & Harrison, S.P. (2012) Historical and ecological controls on phylogenetic diversity in Californian plant communities. *American Naturalist*, **180**, 257–269.

Andersen, P., Cranston, P.S. & Epler, J.H. (2013) *Chironomidae of the Holarctic Region: Keys and Diagnoses, Part 1: Larvae*, Entomological Society of Lund, Sweden.

Anderson, E.P. & Maldonado-Ocampo, J.A. (2011) Una Perspectiva Regional de la Diversidad y Conservación de Peces Andinos Tropicales. *Conservation Biology*, **25**, 30–39.

Anderson, M.J. (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, **62**, 245–253.

Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L., Sanders, N.J., Cornell, H. V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N.J.B., Stegen, J.C. & Swenson, N.G. (2011) Navigating the multiple meanings of β diversity: A roadmap for the practicing ecologist. *Ecology Letters*, **14**, 19–28.

Anderson, M.J., Ellingsen, K.E. & McArdle, B.H. (2006) Multivariate dispersion as a measure of beta diversity. *Ecology Letters*, **9**, 683–693.

Anderson, M.J. & Willis, T.J. (2003) Canonical analysis of principal coordinates: a useful method of contrained ordination for ecology. *Ecology*, **84**, 511–525.

Armenteras, D., Rodríguez, N., Retana, J. & Morales, M. (2011) Understanding deforestation in montane and lowland forests of the Colombian Andes. *Regional Environmental Change*, **11**, 693–705.

Barbour, M.T., Gerritsen, J., Snyder, B.D. & Stribling, J.B. (1999) *Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish*, US Environmental Protection Agency, Office of Water Washington, DC.

Barlow, J., Lennox, G.D., Ferreira, J., Berenguer, E., Lees, A.C., Nally, R. Mac, Thomson, J.R., Ferraz, S.F.D.B., Louzada, J., Oliveira, V.H.F., Parry, L., Ribeiro De Castro Solar, R., Vieira, I.C.G., Aragaõ, L.E.O.C., Begotti, R.A., Braga, R.F., Cardoso, T.M., Jr, R.C.D.O., Souza, C.M., Moura, N.G., Nunes, S.S., Siqueira, J.V., Pardini, R., Silveira, J.M., Vaz-De-Mello, F.Z., Veiga, R.C.S., Venturieri, A. & Gardner, T.A. (2016) Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature*, **535**, 144–147.

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

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

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

Bellinger, E.G. & Sigee, D.C. (2015) *Freshwater algae: identification and use as bioindicators*, John Wiley & Sons.

Benito, X., Fritz, S.C., Steinitz-Kannan, M., Vélez, M.I. & Mcglue, M.M. (2018) Lake regionalization and diatom metacommunity structuring in tropical South America. *Ecology and Evolution*, 7865–7878.

Bertuzzo, E., Carrara, F., Mari, L., Altermatt, F., Rodriguez-Iturbe, I. & Rinaldo, A. (2016) Geomorphic controls on elevational gradients of species richness. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 1737–1742.

Bicudo, T.C., Sacek, V., Almeida, R.P. De, Bates, J.M. & Ribas, C.C. (2019) Andean Tectonics and Mantle Dynamics as a Pervasive Influence on Amazonian Ecosystem. *Scientific Reports*, **9**, 1–11.

De Bie, T., De Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., Hampel, H., Denys, L., Vanhecke, L., Van der Gucht, K., Van Wichelen, J., Vyverman, W. & Declerck, S.A.J. (2012) Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters*, **15**, 740–747.

Bishop, T.R., Robertson, M.P., van Rensburg, B.J. & Parr, C.L. (2015) Contrasting species and functional beta diversity in montane ant assemblages. *Journal of Biogeography*, **42**, 1776–1786.

Bitušík, P., Svitok, M., Novikmec, M., Trnková, K. & Hamerlík, L. (2017) Biological recovery of acidified alpine lakes may be delayed by the dispersal limitation of aquatic insect adults. *Hydrobiologia*, **790**, 287–298.

Bohonak, A.J. & Jenkins, D.G. (2003) Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecology Letters*, **6**, 783–796.

Bojsen, B.H. & Jacobsen, D. (2003) Effects of deforestation on macroinvertebrate diversity and assemblage structure in Ecuadorian Amazon streams. *Archiv für Hydrobiologie*, **158**, 317–342.

Bonada, N., Múrria, C., Zamora-Muñoz, C., El Alami, M., Poquet, J.M., Puntí, T., Moreno, J.L., Bennas, N., Alba-Tercedor, J., Ribera, C. & Prat, N. (2009) Using community and population approaches to understand how contemporary and historical factors have shaped species distribution in river ecosystems. *Global Ecology and Biogeography*, **18**, 202–213.

Bonada, N., Zamora-Muñoz, C., Rieradevall, M. & Prat, N. (2005) Ecological and historical filters constraining spatial caddisfly distribution in Mediterranean rivers. *Freshwater Biology*, **50**, 781–797.

Booth, G.D., Niccolucci, M.J. & Schuster, E.G. (1994) *Identifying proxy sets in multiple linear regression: an aid to better coefficient interpretation*,.

Borcard, D., Gillet, F. & Legendre, P. (2018) *Numerical Ecology with R. Second Edition*, Springer.

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.

Bottin, M., Soininen, J., Ferrol, M. & Tison-Rosebery, J. (2014) Do spatial patterns of benthic diatom assemblages vary across regions and years? *Freshwater Science*, **33**, 402–416.

Brasil, L.S., Vieira, T.B., de Oliveira-Junior, J.M.B., Dias-Silva, K. & Juen, L. (2017) Elements of metacommunity structure in Amazonian Zygoptera among streams under different spatial scales and environmental conditions. *Ecology and Evolution*, **7**, 3190–3200.

Brown, B.L. & Swan, C.M. (2010) Dendritic network structure constrains metacommunity properties in riverine ecosystems. *Journal of Animal Ecology*, **79**, 571–580.

Brown, B.L., Swan, C.M., Auerbach, D.A., Campbell Grant, E.H., Hitt, N.P., Maloney, K.O. & Patrick, C. (2011) Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *Journal of the North American Benthological Society*, **30**, 310–327.

Brown, J.H. (1995) *Macroecology*, University of Chicago Press.

Burdon, F.J., Reyes, M., Alder, A.C., Joss, A., Ort, C., Räsänen, K., Jokela, J., Eggen, R.I.L. & Stamm, C. (2016) Environmental context and magnitude of disturbance influence trait-mediated community responses to wastewater in streams.

Ecology and Evolution, **6**, 3923–3939.

Buschke, F., Brendonck, L. & Vanschoenwinkel, B. (2014) Differences between regional and biogeographic species pools highlight the need for multi-scale theories in macroecology. *Frontiers of Biogeography*, **6**, fb_23415.

Buss, D.F., Carlisle, D.M., Chon, T.-S., Culp, J., Harding, J.S., Keizer-Vlek, H.E., Robinson, W.A., Strachan, S., Thirion, C. & Hughes, R.M. (2015) Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. *Environmental monitoring and assessment*, **187**, 4132.

Buss, D.F. & Vitorino, A.S. (2010) Rapid Bioassessment Protocols using benthic macroinvertebrates in Brazil: evaluation of taxonomic sufficiency. *Journal of the North American Benthological Society*, **29**, 562–571.

Buytaert, W., Célleri, R., De Bièvre, B., Cisneros, F., Wyseure, G., Deckers, J. & Hofstede, R. (2006) Human impact on the hydrology of the Andean páramos. *Earth-Science Reviews*, **79**, 53–72.

Cáceres, M. De & Legendre, P. (2009) Associations between species and groups of sites: indices and statistical inference. *Ecology*, **90**, 3566–3574.

Cadotte, M.W. & Tucker, C.M. (2017) Should Environmental Filtering be Abandoned? *Trends in Ecology & Evolution*, **32**, 429–437.

Campbell Grant, E.H., Lowe, W.H. & Fagan, W.F. (2007) Living in the branches: Population dynamics and ecological processes in dendritic networks. *Ecology Letters*, **10**, 165–175.

Cañedo-Argüelles, M., Boersma, K.S., Bogan, M.T., Olden, J.D., Phillipsen, I., Schriever, T.A. & Lytle, D.A. (2015) Dispersal strength determines meta-community structure in a dendritic riverine network. *Journal of Biogeography*, **42**, 778–790.

Cañedo-Argüelles, M., Grantham, T.E., Perrée, I., Rieradevall, M., Céspedes-Sánchez, R. & Prat, N. (2012) Response of stream invertebrates to short-term salinization: A mesocosm approach. *Environmental Pollution*, **166**, 144–151.

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

Carvajal-Quintero, J., Villalobos, F., Oberdorff, T., Grenouillet, G., Brosse, S., Hugueny, B., Jézéquel, C. & Tedesco, P.A. (2019) Drainage network position and historical connectivity explain global patterns in freshwater fishes' range size. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 13434–13439.

Chase, J.M. (2003) Community assembly: When should history matter? *Oecologia*, **136**, 489–498.

Chave, J. (2013) The problem of pattern and scale in ecology: What have we learned in 20 years? *Ecology Letters*, **16**, 4–16.

Chazot, N., Willmott, K.R., Condamine, F.L., De-Silva, D.L., Freitas, A.V.L., Lamas, G., Morlon, H., Giraldo, C.E., Jiggins, C.D., Joron, M., Mallet, J., Uribe, S. & Elias, M. (2016) Into the Andes: multiple independent colonizations drive montane diversity in the Neotropical clearwing butterflies Godyridina. *Molecular Ecology*, **25**, 5765–5784.

Chessel, D., Dufour, A.B. & Thioulouse, J. (2017) The ade4 package-I-one-table methods. R news. 2004; 4: 5–10.

Chevene, F., Dolédec, S. & Chessel, D. (1994) A fuzzy coding approach for the analysis of long-term ecological data. *Freshwater Biology*, **31**, 295–309.

Clarke, K.R. (1993) Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, **18**, 117–143.

Connor, E.F. & Simberloff, D. (1979) The Assembly of Species Communities: Chance or Competition? *Ecology*, **60**, 1132.

Cornell, H. V. & Harrison, S.P. (2014) What Are Species Pools and When Are They Important? *Annual Review of Ecology, Evolution, and Systematics*, **45**, 45–67.

Crist, T.O., Veech, J.A., Gering, J.C. & Summerville, K.S. (2003) Partitioning Species Diversity across Landscapes and Regions: A Hierarchical Analysis of a, b, and g Diversity. *The American Naturalist*, **162**, 734–743.

Dallas, T. (2014) metacom: an R package for the analysis of metacommunity structure. *Ecography*, **37**, 402–405.

Datry, T., Melo, A.S., Moya, N., Zubieta, J., De la Barra, E. & Oberdorff, T. (2016) Metacommunity patterns across three Neotropical catchments with varying environmental harshness. *Freshwater Biology*, **61**, 277–292.

Datry, T., Vander Vorste, R., Goïtia, E., Moya, N., Campero, M., Rodriguez, F., Zubieta, J. & Oberdorff, T. (2017) Context-dependent resistance of freshwater invertebrate communities to drying. *Ecology and Evolution*, **7**, 3201–3211.

Declerck, S.A.J., Coronel, J.S., Legendre, P. & Brendonck, L. (2011) Scale dependency of processes structuring metacommunities of cladocerans in temporary pools of High-Andes wetlands. *Ecography*, **34**, 296–305.

Dedieu, N., Vigouroux, R., Cerdan, P. & Céréghino, R. (2015) Invertebrate communities delineate hydro-ecoregions and respond to anthropogenic disturbance in East-Amazonian streams. *Hydrobiologia*, **742**, 95–105.

Devictor, V., Mouillot, D., Meynard, C., Jiguet, F., Thuiller, W. & Mouquet, N. (2010) Spatial mismatch and congruence between taxonomic, phylogenetic and functional diversity: The need for integrative conservation strategies in a changing world. *Ecology Letters*, **13**, 1030–1040.

Diamond, J.M. (1975) The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation*, **7**, 129–146.

Dias, M.S., Oberdorff, T., Hugueny, B., Leprieur, F., Jézéquel, C., Cornu, J.F., Brosse, S., Grenouillet, G. & Tedesco, P.A. (2014) Global imprint of historical connectivity on freshwater fish biodiversity. *Ecology Letters*, **17**, 1130–1140.

Dollar, E.S.J., James, C.S., Rogers, K.H. & Thoms, M.C. (2007) A framework for interdisciplinary understanding of rivers as ecosystems. *Geomorphology*, **89**, 147–162.

Domínguez, E. (2006) *Ephemeroptera of South America*, PENSOFT Publishers.

Domínguez, E. & Fernández, H.R. (2009) *Macroinvertebrados bentónicos sudamericanos: sistemática y biología*, Fundación Miguel Lillo Tucumán.

Dong, X., Li, B., He, F., Gu, Y., Sun, M., Zhang, H., Tan, L., Xiao, W., Liu, S. & Cai, Q. (2016) Flow directionality, mountain barriers and functional traits determine diatom metacommunity structuring of high mountain streams. *Scientific Reports*, **6**, 24711.

Downes, B.J., Lancaster, J., Glaister, A. & Bovill, W.D. (2016) A fresh approach reveals how dispersal shapes metacommunity structure in a human-altered landscape. *Journal of Applied Ecology*.

Dray, S., Blanchet, G., Borcard, D., Clappe, S., Jombart, G.G.T., Larocque, G., Legendre, P., Madi, N. & Wagner, H.H. (2018) adespatial: Moran's eigenvector maps and related methods for the spatial multiscale analysis of ecological communities.

Elosegi, A. & Sabater, S. (2008) Conceptos y técnicas en ecología fluvial. *Verhandlungen der Internationalen Vereinigung für …*, **30**, 339–344.

Encalada, A.C., Flecker, A.S., Poff, N.L., Suárez, E., Herrera-R, G.A., Ríos-Touma, B., Jumani, S., Larson, E.I. & Anderson, E.P. (2019) A global perspective on tropical montane rivers. *Science*, **365**, 1124–1129.

Encalada, A.C. & Peckarsky, B.L. (2006) Selective oviposition of the mayfly Baetis bicaudatus. *Oecologia*, **148**, 526–537.

Etter, A., Mcalpine, C. & Possingham, H. (2008) Historical Patterns and Drivers of Landscape Change in Colombia Since 1500 : A Regionalized Spatial Approach Annals of the Association of American Geographers. *Annals of the Association of American Geographers*, **98**, 2–23.

Falasco, E., Bona, F., Monauni, C., Zeni, A. & Piano, E. (2019) Environmental and spatial factors drive diatom species distribution in Alpine streams: Implications for biomonitoring. *Ecological Indicators*, **106**, 105441.

Fattorini, S., Di Biase, L. & Chiarucci, A. (2019) Recognizing and interpreting vegetational belts: New wine in the old bottles of a von Humboldt's legacy. *Journal of Biogeography*, **46**, 1643–1651.

Fenchel, T. & Finlay, B.J. (2005) Bacteria and Island Biogeography. *Science*, **309**, 1997–1999.

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

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

Finn, D.S., Encalada, A.C. & Hampel, H. (2016) Genetic isolation among mountains but not between stream types in a tropical high-altitude mayfly. *Freshwater Biology*, **61**, 702–714.

Flantua, S.G.A. & Hooghiemstra, H. (2018) *Historical connectivity and mountain biodiversity*. *Mountains, Climate and Biodiversity* (ed. by C. Hoorn), A. Perrigo), and A. Antonelli), pp. 171–185. Wiley-Blackwell.

Flantua, S.G.A., O'Dea, A., Onstein, R.E., Giraldo, C. & Hooghiemstra, H. (2019) The flickering connectivity system of the north Andean páramos. *Journal of Biogeography*, **46**, 1808–1825.

Flórez, A. (2003) *Colombia: evolución de sus relieves y modelados*, 1st edn. Universidad Nacional de Colombia, Bogotá D.C.

Frissell, C.A., Liss, W.J., Warren, C.E. & Hurley, M.D. (1986) A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. *Environmental Management*, **10**, 199–214.

Gaston, K.J., Blackburn, T.M., Greenwood, J.J.D., Gregory, R.D., Quinn, R.M. & Lawton, J.H. (2000) Abundance–occupancy relationships. *Journal of Applied Ecology*, **37**, 39–59.

Gauch, H.G. (1982) *Multivariate analysis in community ecology*, Cambridge University Press.

Gill, B.A., Kondratieff, B.C., Casner, K.L., Encalada, A.C., Flecker, A.S., Gannon, D.G., Ghalambor, C.K., Guayasamin, J.M., Poff, N.L., Simmons, M.P., Thomas, S.A., Zamudio, K.R. & Funk, W.C. (2016) Cryptic species diversity reveals biogeographic support for the 'mountain passes are higher in the tropics' hypothesis. *Proceedings of the Royal Society B: Biological Sciences*, **283**, 20160553.

González-Córdoba, M., del Carmen Zúñiga, M. & Manzo, V. (2015) Riqueza genérica y distribución de Elmidae (Insecta: Coleoptera, Byrrhoidea) en el departamento del Valle del Cauca, Colombia. *Biota Colombiana*, **16**, 51–74.

Gonzalez-Penagos, F., Moretti, I., France-Lanord, C. & Guichet, X. (2014) Origins of formation waters in the Llanos foreland basin of Colombia: Geochemical variation and fluid flow history. *Geofluids*, **14**, 443–458.

González-Trujillo, J.D. (2016) Trait-based responses of caddisfly assemblages to the partial channelization of a High-Andean stream. *Hydrobiologia*, **766**, 381–392.

González-Trujillo, J.D. & Donato-Rondon, J.C. (2016) Changes in invertebrate assemblage structure as affected by the flow regulation of a páramo river. *Annales de Limnologie*, **52**.

González-Trujillo, J.D., Donato-Rondon, J.C., Muñoz, I. & Sabater, S. (2020) Historical processes constrain metacommunity structure by shaping different pools of invertebrate taxa within the Orinoco basin. *Diversity and Distributions*, **26**, 49–61.

González-Trujillo, J.D., Petsch, D.K., Córdoba-Ariza, G., Rincón-Palau, K., Donato-Rondon, J.C., Castro-Rebolledo, M.I. & Sabater, S. (2019) Upstream refugia and dispersal ability may override benthic-community responses to high-Andean streams deforestation. *Biodiversity and Conservation*, **28**, 1513–1531.

Goosen, D. (1971) *Physiography and soils of the Llanos Orientales, Colombia*, International Institute for Aerial Survey and Earth Sciences.

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

Gotelli, N.J. & Graves, G.R. (1996) *Null models in ecology*, Smithsonian Institution.

Graham, C.H. & Fine, P.V.A. (2008) Phylogenetic beta diversity: Linking ecological and evolutionary processes across space in time. *Ecology Letters*, **11**, 1265–1277.

Graham, S.E., Storey, R. & Smith, B. (2017) Dispersal distances of aquatic insects: upstream crawling by benthic EPT larvae and flight of adult Trichoptera along valley floors. *New Zealand Journal of Marine and Freshwater Research*, **51**, 146–164.

Gresens, S.E., Belt, K.T., Tang, J.A., Gwinn, D.C. & Banks, P.A. (2007) Temporal and spatial responses of Chironomidae (Diptera) and other benthic invertebrates to urban stormwater runoff. *Hydrobiologia*, **575**, 173–190.

Grönroos, M., Heino, J., Siqueira, T., Landeiro, V.L., Kotanen, J. & Bini, L.M. (2013) Metacommunity structuring in stream networks: Roles of dispersal mode, distance type, and regional environmental context. *Ecology and Evolution*, **3**, 4473–4487.

Gutiérrez-Cánovas, C., Millán, A., Velasco, J., Vaughan, I.P. & Ormerod, S.J. (2013) Contrasting effects of natural and anthropogenic stressors on beta diversity in river organisms. *Global Ecology and Biogeography*, **22**, 796–805.

van der Hammen, T. (1957) Climatic periodicity and evolution of South American Maestrichtian and tertiary floras (a study based on pollen analysis in Colombia). *Boletín Geológico*, **5**.

van der Hammen, T. (1958) Estratigrafía del Terciario y Maestrichtiano continentales y tectogénesis de los Andes Colombianos, Colombia. *Serv Geol. Nac. Bol.*

Geol., **6**, 67–128.

van der Hammen, T. (1974) The Pleistocene Changes of Vegetation and Climate in Tropical South America. *Journal of Biogeography*, **1**, 3.

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

Hawkes, H.A. (1998) Origin and development of the biological monitoring working party score system. *Water Research*, **32**, 964–968.

Hazzi, N.A., Moreno, J.S., Ortiz-Movliav, C. & Palacio, R.D. (2018) Biogeographic regions and events of isolation and diversification of the endemic biota of the tropical Andes. *Proceedings of the National Academy of Sciences*, 201803908.

Heino, J. (2013) The importance of metacommunity ecology for environmental assessment research in the freshwater realm. *Biological Reviews*, **88**, 166–178.

Heino, J., Grönroos, M., Ilmonen, J., Karhu, T., Niva, M. & Paasivirta, L. (2013) Environmental heterogeneity and β diversity of stream macroinvertebrate communities at intermediate spatial scales. *Freshwater Science*, **32**, 142–154.

Heino, J., Grönroos, M., Soininen, J., Virtanen, R. & Muotka, T. (2012) Context dependency and metacommunity structuring in boreal headwater streams. *Oikos*, **121**, 537–544.

Heino, J., Melo, A.S. & Bini, L.M. (2015a) Reconceptualising the beta diversity-environmental heterogeneity relationship in running water systems. *Freshwater Biology*, **60**, 223–235.

Heino, J., Muotka, T. & Paavola, R. (2003) Determinants of macroinvertebrate diversity in headwater streams: Regional and local influences. *Journal of Animal Ecology*, **72**, 425–434.

Heino, J., Soininen, J., Alahuhta, J., Lappalainen, J. & Virtanen, R. (2015b) A comparative analysis of metacommunity types in the freshwater realm. *Ecology and Evolution*, **5**, 1525–1537.

Heino, J., Soininen, J., Alahuhta, J., Lappalainen, J. & Virtanen, R. (2017) Metacommunity ecology meets biogeography: effects of geographical region, spatial dynamics and environmental filtering on community structure in aquatic organisms. *Oecologia*, **183**, 121–137.

Higgins, J. V, Bryer, M.T., Khoury, M.L. & Fitzhugh, T.W. (2005) A freshwater classification approach for biodiversity conservation planning. *Conservation Biology*, **19**, 432–445.

Hitt, N.P. & Angermeier, P.L. (2008) Evidence for fish dispersal from spatial analysis of stream network topology. *Journal of the North American Benthological Society*, **27**, 304–320.

Hoffmann, A. & Resh, V.H. (2003) Oviposition in three species of limnephiloid caddisflies (Trichoptera): Hierarchical influences on site selection. *Freshwater Biology*, **48**, 1064–1077.

Holyoak, M., Leibold, M.A. & Holt, R.D. (2005) *Metacommunities: spatial dynamics and ecological communities*, 1st edn. University of Chicago Press, Chicago.

Hoorn, C., Wesselingh, F.P., ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Särkinen, T. & Antonelli, A. (2010) Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. *Science*, **330**, 927–931.

Horn, H.S. (1966) Measurement of" overlap" in comparative ecological studies. *The American Naturalist*, **100**, 419–424.

Hubbell, S.P. (2006) Neutral theory and the evolution of ecological equivalence. *Ecology*, **87**, 1387–1398.

Hubert, N. & Renno, J.F. (2006) Historical biogeography of South American freshwater fishes. *Journal of Biogeography*, **33**, 1414–1436.

Hughes, C. & Eastwood, R. (2006) Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences*, **103**, 10334–10339.

von Humboldt, A. & Bonpland, A. (1807) *Essai sur la géographie des plantes*,.

Hutchinson, G.E. (1959) Homage to Santa Rosalia or why are there so many kinds of animals? *The American Naturalist*, **93**, 145–159.

Iñiguez-Armijos, C., Leiva, A., Frede, H.G., Hampel, H. & Breuer, L. (2014) Deforestation and benthic indicators: How much vegetation cover is needed to sustain healthy Andean streams? *PLoS ONE*, **9**.

Jackson, S.T. (2009) Alexander von Humboldt and the general physics of the Earth. *Science*, **324**, 596–597.

Jamoneau, A., Passy, S.I., Soininen, J., Leboucher, T. & Tison-Rosebery, J. (2018) *Beta diversity of diatom species and ecological guilds: Response to environmental and spatial mechanisms along the stream watercourse*. *Freshwater Biology*, pp. 62–73.

Jaramillo, C., Ochoa, D., Contreras, L., Pagani, M., Carvajal-Ortiz, H., Pratt, L.M., Krishnan, S., Cardona, A., Romero, M., Quiroz, L., Rodriguez, G., Rueda, M.J., de la Parra, F., Morón, S., Green, W., Bayona, G., Montes, C., Quintero, O., Ramirez, R., Mora, G., Schouten, S., Bermudez, H., Navarrete, R., Parra, F., Alvarán, M., Osorno, J., Crowley, J.L., Valencia, V. & Vervoort, J. (2010) Effects of rapid global warming at the Paleocene-Eocene boundary on neotropical vegetation. *Science (New York, N.Y.)*, **330**, 957–61.

Jenkins, D.G. & Ricklefs, R.E. (2011) Biogeography and ecology: two views of one world. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **366**, 2331–2335.

Jost, L. (2007) Partitioning Diversity into Independent Alpha and Beta Components. *Ecology*, **88**, 2427–2439.

Jost, L., Chao, A. & Chazdon, R.L. (2011) *Compositional similarity and beta diversity*. *Biological diversity: frontiers in measurement and assessment*, pp. 66–84. Oxford University Press Oxford.

Jost, L., Devries, P., Walla, T., Greeney, H., Chao, A. & Ricotta, C. (2010) Partitioning diversity for conservation analyses. *Diversity and Distributions*, **16**, 65–76.

Kärnä, O.M., Grönroos, M., Antikainen, H., Hjort, J., Ilmonen, J., Paasivirta, L. & Heino, J. (2015) Inferring the effects of potential dispersal routes on the metacommunity structure of stream insects: As the crow flies, as the fish swims or as the fox runs? *Journal of Animal Ecology*, **84**, 1342–1353.

Kimura, M. (1991) The neutral theory of molecular evolution: a review of recent evidence. *Idengaku zasshi*, **66**, 367–86.

Kindt, R. & Kindt, M.R. (2018) Package 'BiodiversityR.'

Kitto, J.A.J., Gray, D.P., Greig, H.S., Niyogi, D.K. & Harding, J.S. (2015) Meta-community theory and stream restoration: Evidence that spatial position constrains stream invertebrate communities in a mine impacted landscape. *Restoration Ecology*, **23**, 284–291.

Kondolf, G.M. & Piégay, H. (2005) *Tools in fluvial geomorphology*, (ed. by G.M. Kondolf) and H. Piégay) Wiley.

Kottek, M., Grieser, J., Beck, C., Rudolf, B. & Rubel, F. (2006) World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, **15**, 259–263.

Krammer, K. & Lange-Bertalot, H. (1991) *Bacillariophyceae. 2/3: Centrales, Fragilariaceae, Eunotiaceae*. *Süsswasserflora von Mitteleuropa*,.

Krammer, K. & Lange-Bertalot, H. (1986) Bacillariophyceae. *Susswasserflora von Mitteleuropa*.

Lake, P.S. (2000) Disturbance, patchiness, and diversity in streams. *Journal of the North American Benthological Society*, **19**, 573–592.

Lancaster, J. & Downes, B.J. (2017) Dispersal traits may reflect dispersal distances, but dispersers may not connect populations demographically. *Oecologia*, **184**, 171–182.

Lande, R. (1996) Statistics and Partitioning of Species Diversity, and Similarity among Multiple Communities. *Oikos*, **76**, 5.

Lange, K., Townsend, C.R. & Matthaei, C.D. (2016) A trait-based framework for stream algal communities. *Ecology and Evolution*, **6**, 23–36.

Lawton, J.H. (1999) Are There General Laws in Ecology? *Oikos*, **84**, 177–192.

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

Legendre, P. & Anderson, M.J. (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological monographs*, **69**, 1–24.

Legendre, P. & Gallagher, E.D. (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.

Legendre, P., Lapointe, F.-J. & Casgrain, P. (1994) MODELING BRAIN EVOLUTION FROM BEHAVIOR: A PERMUTATIONAL REGRESSION APPROACH. *Evolution*, **48**, 1487–1499.

Legendre, P. & Legendre, L. (2012) *Numerical Ecology*, Elsevier.

Leibold, M.A. & Chase, J.M. (2018) *Metacommunity ecology*, 1st edn. Princeton University Press, Princeton.

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

Leibold, M.A. & Mikkelson, G.M. (2002) Coherence, species turnover, and boundary clumping: elements of meta-community structure. *Oikos*, **97**, 237–250.

Leira, M. & Sabater, S. (2005) Diatom assemblages distribution in catalan rivers, NE Spain, in relation to chemical and physiographical factors. *Water Research*, **39**, 73–82.

Lenat, D.R. (1993) A biotic index for the southeastern United States: derivation and list of tolerance values, with criteria for assigning water-quality ratings. *Journal of the North American Benthological Society*, **12**, 279–290.

Lencioni, V., Marziali, L. & Rossaro, B. (2012) Chironomids as bioindicators of environmental quality in mountain springs. *Freshwater Science*, **31**, 525–541.

Lepori, F. & Hjerdt, N. (2006) Disturbance and Aquatic Biodiversity: Reconciling Contrasting Views. *BioScience*, **56**, 809.

Leprieur, F., Albouy, C., de Bortoli, J., Cowman, P.F., Bellwood, D.R. & Mouillot, D. (2012) Quantifying phylogenetic beta diversity: Distinguishing between “true” turnover of lineages and phylogenetic diversity gradients. *PLoS ONE*, **7**.

Lessard, J.P., Borregaard, M.K., Fordyce, J.A., Rahbek, C., Weiser, M.D., Dunn, R.R. & Sanders, N.J. (2012) Strong influence of regional species pools on continent-wide structuring of local communities. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 266–274.

Liu, C., Dudley, K.L., Xu, Z.H. & Economo, E.P. (2018) Mountain metacommunities:

climate and spatial connectivity shape ant diversity in a complex landscape. *Ecography*, **41**, 101–112.

Liu, J., Soininen, J., Han, B.P. & Declerck, S.A.J. (2013) Effects of connectivity, dispersal directionality and functional traits on the metacommunity structure of river benthic diatoms. *Journal of Biogeography*, **40**, 2238–2248.

Lorion, C.M. & Kennedy, B.P. (2009) Relationships between deforestation, riparian forest buffers and benthic macroinvertebrates in neotropical headwater streams. *Freshwater Biology*, **54**, 165–180.

MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*, Princeton university press.

Macneale, K.H., Peckarsky, B.L. & Likens, G.E. (2005) Stable isotopes identify dispersal patterns of stonefly populations living along stream corridors. *Freshwater Biology*, **50**, 1117–1130.

Maddison, W.P. & Maddison, D.R. (2018) Mesquite: a modular system for evolutionary analysis. Version 3.51.

Maire, E., Grenouillet, G., Brosse, S. & Villéger, S. (2015) How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. *Global Ecology and Biogeography*, **24**, 728–740.

Malmqvist, B. (2000) How does wing length relate to distribution patterns of stoneflies (Plecoptera) and mayflies (Ephemeroptera)? *Biological Conservation*, **93**, 271–276.

Marcilio-Silva, V., Zwiener, V.P. & Marques, M.C.M. (2017) Metacommunity structure, additive partitioning and environmental drivers of woody plants diversity in the Brazilian Atlantic Forest. *Diversity and Distributions*, **23**, 1110–1119.

Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S., Øvreås, L., Reysenbach, A.-L., Smith, V.H. & Staley, J.T. (2006) Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, **4**, 102.

McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution*, **21**, 178–185.

Merritt, R.W. & Cummins, K.W. (2008) *An introduction to the Aquatic insects of North America*, Kendall/Hunt Pub. Co.

Metzeltin, D. & Lange-Bertalot, H. (2007) *Tropical Diatoms of South America II. Special remarks on biogeography disjunction. Iconographia Diatomologica 18: 1-877, incl. 296 pl.,*.

Milošević, D., Simić, V., Stojković, M., Čerba, D., Mančev, D., Petrović, A. & Paunović, M. (2013) Spatio-temporal pattern of the Chironomidae community:

toward the use of non-biting midges in bioassessment programs. *Aquatic Ecology*, **47**, 37–55.

Mittelbach, G.G. & Schemske, D.W. (2015) Ecological and evolutionary perspectives on community assembly. *Trends in Ecology and Evolution*, **30**, 241–247.

Mori, A.S., Isbell, F. & Seidl, R. (2018) β-Diversity, Community Assembly, and Ecosystem Functioning. *Trends in Ecology and Evolution*, **33**, 549–564.

Morton, E.R., McGrady, M.J., Newton, I., Rollie, C.J., Smith, G.D., Mearns, R. & Oli, M.K. (2018) Dispersal: a matter of scale. *Ecology*, **99**, 938–946.

Muellner-Riehl, A.N., Schnitzler, J., Kissling, W.D., Mosbrugger, V., Rijsdijk, K.F., Seijmonsbergen, A.C., Versteegh, H. & Favre, A. (2019) Origins of global mountain plant biodiversity: Testing the 'mountain-geobiodiversity hypothesis.' *Journal of Biogeography*, jbi.13715.

Múrria, C., Rugenski, A.T., Whiles, M.R. & Vogler, A.P. (2015) Long-term isolation and endemicity of Neotropical aquatic insects limit the community responses to recent amphibian decline. *Diversity and Distributions*, **21**, 938–949.

Myers, N., Fonseca, G.A.B., Mittermeier, R. a, Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Nakov, T., Beaulieu, J.M. & Alverson, A.J. (2018) Accelerated diversification is related to life history and locomotion in a hyperdiverse lineage of microbial eukaryotes (Diatoms, Bacillariophyta). *New Phytologist*, **219**, 462–473.

Neff, M.R. & Jackson, D.A. (2013) Regional-scale patterns in community concordance: testing the roles of historical biogeography versus contemporary abiotic controls in determining stream community composition. *Canadian Journal of Fisheries and Aquatic Sciences*, **70**, 1141–1150.

Nottingham, A.T., Fierer, N., Turner, B.L., Whitaker, J., Ostle, N.J., McNamara, N.P., Bardgett, R.D., Leff, J.W., Salinas, N., Silman, M., Kruuk, L. & Meir, P. (2018) Microbes follow Humboldt: temperature drives plant and soil microbial diversity patterns from the Amazon to the Andes. *Ecology*, **99**, 2455–2466.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2013) Package 'vegan.' *Community ecology package, version*, **2**.

Padial, A.A., Ceschin, F., Declerck, S.A.J., De Meester, L., Bonecker, C.C., Lansac-Tôha, F.A., Rodrigues, L., Rodrigues, L.C., Train, S., Velho, L.F.M. & Bini, L.M. (2014) Dispersal ability determines the role of environmental, spatial and temporal drivers of metacommunity structure. *PLoS ONE*, **9**, 1–8.

Palmer, M.A., Ambrose, R.F. & Poff, N.L. (1997) Ecological Theory and Community Restoration Ecology. *Restoration Ecology*, **5**, 291–300.

Passy, S.I. (2012) A hierarchical theory of macroecology. *Ecology Letters*, **15**, 923–934.

Passy, S.I. (2016) Abundance Inequality in Freshwater Communities Has an Ecological Origin. *The American Naturalist*, **187**, 502–516.

Passy, S.I. (2008) Continental diatom biodiversity in stream benthos declines as more nutrients become limiting. *Proceedings of the National Academy of Sciences*, **105**, 9663–9667.

Passy, S.I. (2007) Diatom ecological guilds display distinct and predictable behavior along nutrient and disturbance gradients in running waters. *Aquatic Botany*, **86**, 171–178.

Passy, S.I., Larson, C.A., Jamoneau, A., Budnick, W., Leboucher, T., Tison-rosebery, J. & Soininen, J. (2018) The biogeographical patterns of species richness and abundance distribution in stream diatoms are driven by climate and water chemistry. **192**.

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

Pero, E.J.I., Hankel, G.E., Molineri, C. & Domínguez, E. (2019) Correspondence between stream benthic macroinvertebrate assemblages and ecoregions in northwestern Argentina. *Freshwater Science*, **38**, 64–76.

Perrigo, A., Hoorn, C. & Antonelli, A. (2019) Why mountains matter for biodiversity. *Journal of Biogeography*, jbi.13731.

Petts, G.E. & Amoros, C. (1996) *Fluvial hydrosystems*, Chapman & Hall.

Pither, J. & Aarssen, L.W. (2005) The evolutionary species pool hypothesis and patterns of freshwater diatom diversity along a pH gradient. *Journal of Biogeography*, **32**, 503–513.

Poff, N.L. (1997) Landscape Filters and Species Traits: Towards Mechanistic Understanding and Prediction in Stream Ecology. *Journal of the North American Benthological Society*, **16**, 391–409.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E. & Stromberg, J.C. (1997) The Natural Flow Regime. *BioScience*, **47**, 769–784.

Poff, N.L., Olden, J.D., Vieira, N.K.M., Finn, D.S., Simmons, M.P. & Kondratieff, B.C. (2006) Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships. *Journal of the North American Benthological Society*, **25**, 730–755.

Poiani, K.A., Richter, B.D., Anderson, M.G. & Richter, H.E. (2000) Biodiversity Conservation at Multiple Scales: Functional Sites, Landscapes, and Networks. *BioScience*, **50**, 133.

Poulíčková, A., Veselá, J., Neustupa, J. & Škaloud, P. (2010) Pseudocryptic Diversity versus cosmopolitanism in diatoms: A case study on Navicula cryptocephala Kütz. (Bacillariophyceae) and morphologically similar taxa. *Protist*, **161**, 353–369.

Prat, N., González-Trujillo, J.D. & Ospina-Torres, R. (2014) Key to chironomid pupal exuviae (Diptera: Chironomidae) of tropical high Andean streams. *Revista de Biologia Tropical*, **62**.

Prat, N., Rieradevall, M., Acosta, R. & Villamarín, C. (2011) Guía para el reconocimiento de las larvas de Chironomidae (díptera) de los ríos altoandinos de Ecuador y Perú. *Grupo de Investigación FEM Departamento de Ecología, Universidad de Barcelona, España*, **78**.

Presley, S.J., Cisneros, L.M., Patterson, B.D. & Willig, M.R. (2012) Vertebrate metacommunity structure along an extensive elevational gradient in the tropics: A comparison of bats, rodents and birds. *Global Ecology and Biogeography*, **21**, 968–976.

Presley, S.J., Higgins, C.L. & Willig, M.R. (2010) A comprehensive framework for the evaluation of metacommunity structure. *Oikos*, **119**, 908–917.

Presley, S.J. & Willig, M.R. (2010) Bat metacommunity structure on Caribbean islands and the role of endemics. *Global Ecology and Biogeography*, **19**, 185–199.

Qian, H., Ricklefs, R.E. & White, P.S. (2005) Beta diversity of angiosperms in temperate floras of eastern Asia and eastern North America. *Ecology Letters*, **8**, 15–22.

R Core Team (2018) R: A Language and Environment for Statistical Computing.

Rangel-ch, J.O. (2016) *Ecosistemas de la Orinoquía de Colombia*. *Colombia Diversidad Biótica XIV*, pp. 807–847. Universidad Nacional de Colombia, Bogotá D.C.

Restrepo, J.J. & Toussaint, J.F. (1988) Terranes and continental accretion in the Colombian Andes. *Episodes*, **11**, 189–193.

Říčan, O. (2011) Historical biogeography of Neotropical freshwater fishes.

Richardson, J.S. & Béraud, S. (2014) Effects of riparian forest harvest on streams: A meta-analysis. *Journal of Applied Ecology*, **51**, 1712–1721.

Ricklefs, R. (1987) Community diversity: relative roles of local and regional processes. *Science*, **235**, 167–171.

Ricklefs, R.E. & Jenkins, D.G. (2011) Biogeography and ecology: towards the integration of two disciplines. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **366**, 2438–2448.

Rimet, F. & Bouchez, A. (2011) Use of diatom life-forms and ecological guilds to assess pesticide contamination in rivers: Lotic mesocosm approaches. *Ecological Indicators*, **11**, 489–499.

Rios-Touma, B., Encalada, A.C. & Prat, N. (2012) Oviposition of Aquatic Insects in a Tropical High Altitude Stream. *Environmental Entomology*, **41**, 1322–1331.

Ríos-Touma, B., Prat, N. & Encalada, A.C. (2011) Invertebrate drift and colonization processes in a tropical andean stream. *Aquatic Biology*, **14**, 233–246.

Rodríguez Eraso, N., Armenteras Pascual, D. & Alumbrerosa, R. (2012) Land use and land cover change in the Colombian Andes: dynamics and future scenarios. *Journal of Land Use Science*, **8**, 154–174.

Romero Ruíz, M.H., Galindo García, G., Otero García, J. & Armenteras Pascual, D. (2004) *Ecosistemas de la cuenca del Orinoco colombiano*, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, Colomobia.

Rull, V. (2008) Speciation timing and neotropical biodiversity: The Tertiary-Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology*, **17**, 2722–2729.

Sabater, S., González-Trujillo, J.D., Elosegi, A. & Donato Rondón, J.C. (2017) Colombian ecosystems at the crossroad after the new peace deal. *Biodiversity and Conservation*, **26**, 3505–3507.

Saito, V.S., Soininen, J., Fonseca-Gessner, A.A. & Siqueira, T. (2015) Dispersal traits drive the phylogenetic distance decay of similarity in Neotropical stream metacommunities. *Journal of Biogeography*, **42**, 2101–2111.

Sarremejane, R., Mykrä, H., Bonada, N., Aroviita, J. & Muotka, T. (2017) Habitat connectivity and dispersal ability drive the assembly mechanisms of macroinvertebrate communities in river networks. *Freshwater Biology*, **62**, 1073–1082.

Schumm, S.A. (1977) *The fluvial system*, Wiley.

Seckbach, J. & Oren, A. (2007) *Oxygenic Photosynthetic Microorganisms in Extreme Environments*. *Algae and Cyanobacteria in Extreme Environments* (ed. by J. Seckbach), pp. 3–25. Springer Netherlands, Dordrecht.

Sedell, J.R., Reeves, G.H., Hauer, F.R., Stanford, J.A. & Hawkins, C.P. (1990) Role of refugia in recovery from disturbances: Modern fragmented and disconnected river systems. *Environmental Management*, **14**, 711–724.

Shurin, J.B., Cottenie, K. & Hillebrand, H. (2009) Spatial autocorrelation and dispersal limitation in freshwater organisms. *Oecologia*, **159**, 151–159.

Siqueira, T., Durães, L. & Roque, F. de O. (2014a) *Predictive Modelling of Insect Metacommunities in Biomonitoring of Aquatic Networks*. *Ecological Modelling Applied to Entomology*, pp. 109–126. Springer International Publishing, Cham.

Siqueira, T., Durães, L.D. & de Oliveira Roque, F. (2014b) *Predictive Modelling of Insect Metacommunities in Biomonitoring of Aquatic Networks*. *Ecological Modelling Applied to Entomology*, pp. 109–126. Springer International Publishing, Cham.

Siqueira, T., Saito, V.S., Bini, L.M., Melo, A.S., Petsch, D.K., Landeiro, V.L., Tolonen, K.T., Jyrkänkallio-Mikkola, J., Soininen, J. & Heino, J. (2020) Community size can affect the signals of ecological drift and niche selection on biodiversity. *Ecology*, **n/a**, e03014.

Socolar, J.B., Gilroy, J.J., Kunin, W.E. & Edwards, D.P. (2016) How Should Beta-Diversity Inform Biodiversity Conservation? *Trends in Ecology and Evolution*, **31**, 67–80.

Soininen, J. (2007) Environmental and spatial control of freshwater diatoms—a review. *Diatom Research*, **22**, 473–490.

Soininen, J. (2012) Macroecology of unicellular organisms - patterns and processes. *Environmental Microbiology Reports*, **4**, 10–22.

Sørensen, T. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biol. Skr.*, **5**, 1–34.

Srivastava, D.S. (1999) Using Local-Regional Richness Plots to Test for Species Saturation- Pitfalls and Potentials.pdf. *Journal of Animal Ecology*, **68**, 1–16.

Stallard, R.F. (1985) *River chemistry, geology, geomorphology, and soils in the Amazon and Orinoco basins*. *The chemistry of weathering*, pp. 293–316. Springer.

Témez, J.R. (2003) Facetas del cálculo hidrometeorológico y estadístico de máximos caudales. *Revista de Obras Públicas*, 47–51.

Thompson, R. & Townsend, C. (2006) A truce with neutral theory: Local deterministic factors, species traits and dispersal limitation together determine patterns of diversity in stream invertebrates. *Journal of Animal Ecology*, **75**, 476–484.

Thorp, J.H., Thoms, M.C. & Delong, M.D. (2006) The riverine ecosystem synthesis: Biocomplexity in river networks across space and time. *River Research and Applications*, **22**, 123–147.

Thorp, J.H., Thoms, M.C. & Delong, M.D. (2010) *The riverine ecosystem synthesis: toward conceptual cohesiveness in river science*, Elsevier.

Tilman, D., Kilham, S.S. & Kilham, P. (1982) Phytoplankton Community Ecology: The Role of Limiting Nutrients. *Annual Review of Ecology and Systematics*, **13**, 349–372.

Tomanova, S. & Usseglio-Polatera, P. (2007) Patterns of benthic community traits in neotropical streams: relationship to mesoscale spatial variability. *Fundamental and Applied Limnology / Archiv für Hydrobiologie*, **170**, 243–255.

Tonkin, J.D., Altermatt, F., Finn, D.S., Heino, J., Olden, J.D., Pauls, S.U. & Lytle, D.A. (2018) The role of dispersal in river network metacommunities: Patterns, processes, and pathways. *Freshwater Biology*, **63**, 141–163.

Tonkin, J.D., Stoll, S., Sundermann, A. & Haase, P. (2014) Dispersal distance and the pool of taxa, but not barriers, determine the colonisation of restored river reaches by benthic invertebrates. *Freshwater Biology*, **59**, 1843–1855.

Tonkin, J.D., Sundermann, A., Jähnig, S.C. & Haase, P. (2015) Environmental controls on river assemblages at the regional scale: An application of the elements of metacommunity structure framework. *PLoS ONE*, **10**, e0135450.

Tornés, E., Mor, J.-R., Mandaric, L. & Sabater, S. (2018) Diatom responses to sewage inputs and hydrological alteration in Mediterranean streams. *Environmental Pollution*, **238**, 369–378.

Tornés, E. & Ruhí, A. (2013) Flow intermittency decreases nestedness and specialisation of diatom communities in Mediterranean rivers. *Freshwater Biology*, **58**, 2555–2566.

Townsend, C.R. & Hildrew, A.G. (1994) Species traits in relation to a habitat templet for river systems. *Freshwater Biology*, **31**, 265–275.

Trivinho-Strixino, S. & Strixino, G. (1995) *Larvas de Chironomidae (Diptera) do Estado de São Paulo: guia de identificação e diagnose dos gêneros*,.

Varanka, S. & Luoto, M. (2012) Environmental determinants of water quality in boreal rivers based on partitioning methods. *River Research and Applications*, **28**, 1034–1046.

Vellend, M. (2010) Conceptual synthesis in community ecology. *The Quarterly review of biology*, **85**, 183–206.

Verleyen, E., Vyverman, W., Sterken, M., Hodgson, D.A., De Wever, A., Juggins, S., Van De Vijver, B., Jones, V.J., Vanormelingen, P., Roberts, D., Flower, R., Kilroy, C., Souffreau, C. & Sabbe, K. (2009) The importance of dispersal related and local factors in shaping the taxonomic structure of diatom metacommunities. *Oikos*, **118**, 1239–1249.

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

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

Vilmi, A., Karjalainen, S.M. & Heino, J. (2017) Ecological uniqueness of stream and lake diatom communities shows different macroecological patterns. *Diversity and Distributions*, **23**, 1042–1053.

Violle, C., Reich, P.B., Pacala, S.W., Enquist, B.J. & Kattge, J. (2014) The emergence and promise of functional biogeography. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 13690–13696.

Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D.A., Mann, D.G., Juggins, S., Van de Vijver, B., Jones, V., Flower, R., Roberts, D., Chepurnov, V.A., Kilroy, C., Vanormelingen, P. & De Wever, A. (2007a) Historical processes constrain patterns in global diatom diversity. *Ecology*, **88**, 1924–31.

Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D.A., Mann, D.G., Juggins, S., Van De Vijver, B., Jones, V., Flower, R., Roberts, D., Chepurnov, V.A., Kilroy, C., Vanormelingen, P. & De Wever, A. (2007b) Historical processes constrain patterns in global diatom diversity. *Ecology*, **88**, 1924–1931.

Wang, J., Soininen, J., Zhang, Y., Wang, B., Yang, X. & Shen, J. (2012) Patterns of elevational beta diversity in micro- and macroorganisms. *Global Ecology and Biogeography*, **21**, 743–750.

Ward, J. V. (1998) *Riverine landscapes: Biodiversity patterns, disturbance regimes, and aquatic conservation*. *Biological Conservation*, pp. 269–278. Elsevier.

Webb, C.O., Ackerly, D.D. & Kembel, S.W. (2008) Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, **24**, 2098–2100.

Wentworth, C.K. (1922) A scale of grade and class terms for clastic sediments. *The journal of geology*, **30**, 377–392.

Whitmore, T.J. (1989) Florida diatom assemblages as indicators of trophic state and pH. *Limnology and Oceanography*, **34**, 882–895.

Whittaker, R.H. (1972) Evolution and Measurement of Species Diversity. *Taxon*, **21**, 213.

Wickham, H. (2016) *ggplot2: elegant graphics for data analysis*, Springer.

Wickham, H. (2010) ggplot2: elegant graphics for data analysis. *J Stat Softw*, **35**, 65–88.

Widder, S., Besemer, K., Singer, G.A., Ceola, S., Bertuzzo, E., Quince, C., Sloan, W.T., Rinaldo, A. & Battin, T.J. (2014) Fluvial network organization imprints on microbial co-occurrence networks. *Proceedings of the National Academy of Sciences*, **111**, 12799–12804.

Wiens, J.A. (2002) Riverine landscapes: Taking landscape ecology into the water. *Freshwater Biology*, **47**, 501–515.

Wiens, J.A. (1989) Spatial Scaling in Ecology. *Functional Ecology*, **3**, 385.

Williams, C.J. (1989) Downstream drift of the larvae of Chironomidae (Diptera) in the River Chew, SW England. *Hydrobiologia*, **183**, 59–72.

Willig, M.R., Presley, S.J., Bloch, C.P., Castro-Arellano, I., Cisneros, L.M., Higgins, C.L. & Klingbeil, B.T. (2011) Tropical metacommunities along elevational gradients: Effects of forest type and other environmental factors. *Oikos*, **120**, 1497–1508.

Winking, C., Lorenz, A.W., Sures, B. & Hering, D. (2014) Recolonisation patterns of benthic invertebrates: A field investigation of restored former sewage channels. *Freshwater Biology*, **59**, 1932–1944.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A. & Smith, G.M. (2009) Mixed effects models and extensions in ecology with R. Gail M, Krickeberg K, Samet JM, Tsiatis A, Wong W, editors. *New York, NY: Spring Science and Business Media*.

Supplementary Material

Supplementary Material – Chapter III.

List of invertebrate genera selected as indicator taxa for the IndVal Analysis.

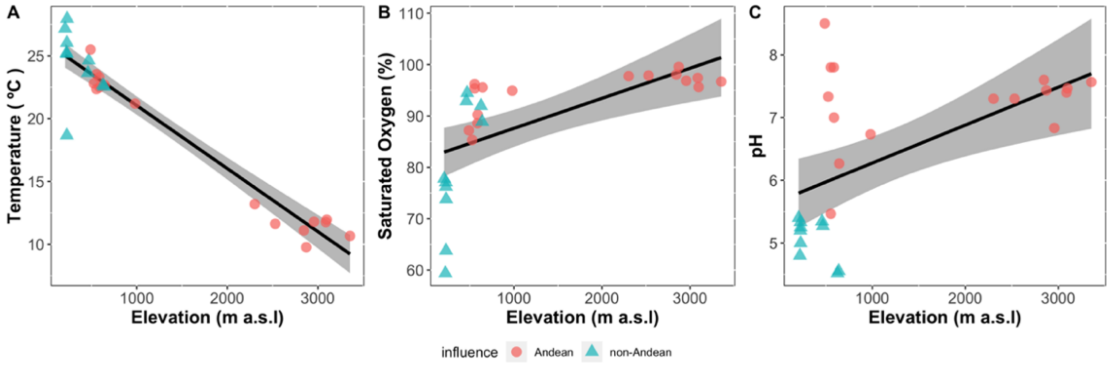
```
 Multilevel pattern analysis
 ---------------------------

 Association function: IndVal.g
 Significance level (alpha): 0.05

 Total number of species: 133
 Selected number of species: 114 
 Number of species associated to 1 group: 67 
 Number of species associated to 2 groups: 28 
 Number of species associated to 3 groups: 7 
 Number of species associated to 4 groups: 9 
 Number of species associated to 5 groups: 3 

 List of species associated to each combination: 

 Group Alluvial_fans  #sps.  6 
                       A      B  stat p.value   
Thraulodes        0.9419 0.5455 0.717   0.002 **
Hexanchorus       0.7740 0.4545 0.593   0.009 **
Neotrichia        0.9353 0.3636 0.583   0.008 **
near.Endotribelos 1.0000 0.2727 0.522   0.016 * 
Varipes           0.9603 0.2727 0.512   0.033 * 
Corydalus         0.7102 0.3636 0.508   0.046 * 

 Group Guiana_shield  #sps.  17 
                      A      B  stat p.value    
Americabaetis    1.0000 1.0000 1.000   0.001 ***
Cloeodes         0.9987 1.0000 0.999   0.001 ***
Petrophila       0.8783 1.0000 0.937   0.001 ***
Tholymis         0.9865 0.8889 0.936   0.001 ***
Tricorythopsis   0.9609 0.8889 0.924   0.001 ***
Thiennemannimyia 0.9557 0.8889 0.922   0.001 ***
Microcylloepus   0.8733 0.7778 0.824   0.001 ***
Oxyethira        1.0000 0.6667 0.816   0.001 ***
Mayatrichia      0.9933 0.6667 0.814   0.001 ***
Haeterina        1.0000 0.5556 0.745   0.001 ***
Xenelmis         0.9333 0.5556 0.720   0.001 ***
Labrundinia      1.0000 0.4444 0.667   0.002 ** 
Scaphydra        1.0000 0.4444 0.667   0.002 ** 
Orthocladinae.spA 0.9920 0.3333 0.575   0.004 ** 
Argia            0.9474 0.3333 0.562   0.007 ** 
Pelocoris        0.9415 0.3333 0.560   0.015 *  
Miroculis        1.0000 0.2222 0.471   0.032 *  

 Group high-Andean  #sps.  8 
                  A      B  stat p.value    
Metrichia    0.9707 1.0000 0.985   0.001 ***
Gigantodax   0.9219 0.9286 0.925   0.001 ***
Nectopsyche  1.0000 0.7857 0.886   0.001 ***
Onychelmis   0.9699 0.6429 0.790   0.001 ***
Elodes       1.0000 0.5714 0.756   0.001 ***
Podonomopsis 1.0000 0.5714 0.756   0.001 ***
Hyalella     1.0000 0.5000 0.707   0.001 ***
Microcara    1.0000 0.3571 0.598   0.006 ** 

 Group highPlains  #sps.  23 
```

```
                       A      B  stat p.value    
Ablabesmyia       1.0000 1.0000 1.000   0.001 ***
Hexacylloepus     1.0000 1.0000 1.000   0.001 ***
Chironomus        0.9678 1.0000 0.984   0.001 ***
Polycentropus     0.9373 1.0000 0.968   0.001 ***
Stenochironomus   0.8793 1.0000 0.938   0.001 ***
Stegoelmis        1.0000 0.8333 0.913   0.001 ***
Oecetis           0.9582 0.8333 0.894   0.001 ***
Phylloicus        0.9561 0.8333 0.893   0.001 ***
Lauterborniella   0.8844 0.8333 0.858   0.001 ***
Aeschnosoma       1.0000 0.6667 0.816   0.001 ***
Endotribelos      0.9643 0.6667 0.802   0.001 ***
Fissimentum       0.9070 0.6667 0.778   0.001 ***
Djalmabatista     0.7472 0.6667 0.706   0.002 ** 
Huleechius        0.8736 0.5000 0.661   0.004 ** 
Caenis            0.6522 0.6667 0.659   0.001 ***
Macrobrachium     0.6429 0.6667 0.655   0.001 ***
Disersus          1.0000 0.3333 0.577   0.006 ** 
Hintoelmis.cf     1.0000 0.3333 0.577   0.006 ** 
Paratanytarsus.cf 1.0000 0.3333 0.577   0.008 ** 
Pelomus.cf        1.0000 0.3333 0.577   0.008 ** 
Procladius        1.0000 0.3333 0.577   0.008 ** 
Tenagobia         1.0000 0.3333 0.577   0.006 ** 
Nilotanypus       0.7903 0.3333 0.513   0.022 *  

 Group Paramo  #sps.  8 
                      A      B  stat p.value    
Oliveiriella     0.9356 1.0000 0.967   0.001 ***
Neoplasta        0.7729 1.0000 0.879   0.001 ***
Ochrotrichia     1.0000 0.6667 0.816   0.001 ***
Pristina         1.0000 0.6667 0.816   0.001 ***
Parakiefferiella 0.8145 0.7778 0.796   0.001 ***
Limonia          0.9725 0.5556 0.735   0.001 ***
Limaya           0.8305 0.5556 0.679   0.002 ** 
near.Lymnophyes  1.0000 0.3333 0.577   0.008 ** 

 Group Piedmont  #sps.  5 
                    A      B  stat p.value    
Ceratotrichia  0.9828 0.6842 0.820   0.001 ***
Maurina        0.9124 0.5263 0.693   0.003 ** 
Blepharopus    0.9147 0.4737 0.658   0.006 ** 
Guajirolus     0.8579 0.4211 0.601   0.026 *  
Cerasmotrichia 1.0000 0.3158 0.562   0.006 ** 

 Group Alluvial_fans+high-Andean  #sps.  1 
              A    B stat p.value   
Alotanypus 1.00 0.36  0.6   0.005 **

 Group Alluvial_fans+Piedmont  #sps.  3 
                     A      B  stat p.value    
Camelobaetidius 0.9994 0.7000 0.836   0.001 ***
Paracloeodes    1.0000 0.5333 0.730   0.002 ** 
Meridialis      1.0000 0.3333 0.577   0.014 *  

 Group Guiana_shield+high-Andean  #sps.  4 
                 A      B  stat p.value    
Paraponyx   0.9932 0.8261 0.906   0.001 ***
Triaenodes  1.0000 0.5217 0.722   0.002 ** 
Psephenus   1.0000 0.3913 0.626   0.005 ** 
Haplohyphes 0.9211 0.3913 0.600   0.005 ** 

 Group Guiana_shield+highPlains  #sps.  4 
                   A      B  stat p.value    
Beardius      1.0000 0.6000 0.775   0.002 ** 
Macropelopia  0.9532 0.6000 0.756   0.001 ***
Phaenopsectra 1.0000 0.4000 0.632   0.002 ** 
```

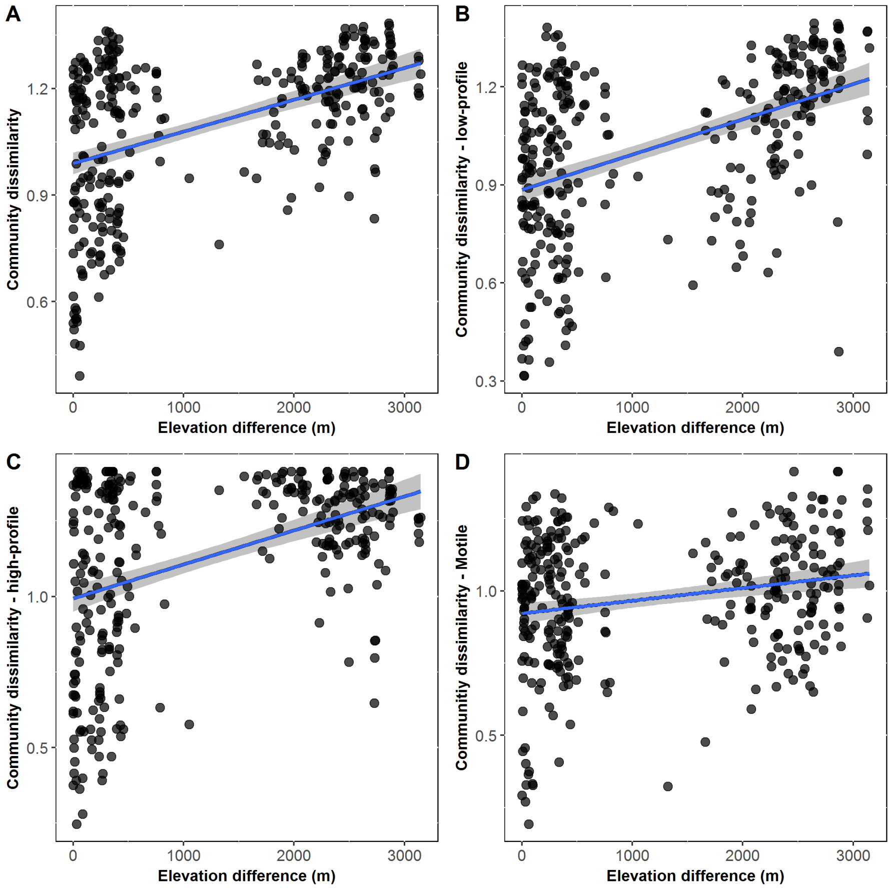
```
Polyplectropus 1.0000 0.2000 0.447   0.048 *

 Group Guiana_shield+Paramo  #sps.  1 
                A      B stat p.value   
Hydroptila 0.9584 0.5556 0.73   0.002 **

 Group high-Andean+highPlains  #sps.  1 
                 A      B  stat p.value    
Nanocladius 0.8837 0.6500 0.758   0.001 ***

 Group high-Andean+Paramo  #sps.  13 
                      A      B  stat p.value    
Morfo.X          1.0000 0.8696 0.933   0.001 ***
Andesiops        0.9288 0.9130 0.921   0.001 ***
Podonomus        0.9743 0.8696 0.920   0.001 ***
Contulma         1.0000 0.8261 0.909   0.001 ***
Xiphocentron.cf  0.9941 0.8261 0.906   0.001 ***
Parametriocnemus 0.8107 1.0000 0.900   0.001 ***
Genus.1          0.9693 0.8261 0.895   0.001 ***
Atopsyche        0.9068 0.8261 0.866   0.001 ***
Austrolimnius    0.8145 0.9130 0.862   0.001 ***
Blepharicera     0.9959 0.7391 0.858   0.001 ***
Atrichopogon     0.9011 0.6087 0.741   0.002 ** 
Metriocnemus     1.0000 0.3913 0.626   0.004 ** 
Paraheptagyia    1.0000 0.2609 0.511   0.020 *  

 Group high-Andean+Piedmont  #sps.  1 
                   A      B  stat p.value  
Tricorythodes 1.0000 0.2727 0.522   0.038 *

 Group Alluvial_fans+Guiana_shield+highPlains  #sps.  1 
               A      B  stat p.value    
Macrelmis 0.8254 0.7308 0.777   0.001 ***

 Group Alluvial_fans+high-Andean+Paramo  #sps.  1 
              A      B  stat p.value    
Neoelmis 0.9137 0.8235 0.867   0.001 ***

 Group Alluvial_fans+high-Andean+Piedmont  #sps.  1 
                 A      B  stat p.value    
Grumichella 0.9912 0.6591 0.808   0.001 ***

 Group Guiana_shield+high-Andean+highPlains  #sps.  4 
                  A      B  stat p.value    
Tanytarsus   0.9204 0.9310 0.926   0.001 ***
Helicopsyche 0.9497 0.8276 0.887   0.001 ***
Riethia      0.8705 0.8621 0.866   0.001 ***
Heterelmis   0.8641 0.6897 0.772   0.002 ** 

 Group Alluvial_fans+Guiana_shield+high-Andean+highPlains  #sps.  1 
                  A      B  stat p.value   
Anacroneuria 0.9274 0.6750 0.791   0.003 **

 Group Alluvial_fans+Guiana_shield+high-Andean+Paramo  #sps.  4 
                 A      B  stat p.value    
Onconeura   0.9471 1.0000 0.973   0.001 ***
Simulium    0.9768 0.9535 0.965   0.001 ***
Corynoneura 0.9527 0.9767 0.965   0.012 *  
Pentaneura  0.9796 0.7442 0.854   0.001 ***

 Group Alluvial_fans+Guiana_shield+high-Andean+Piedmont  #sps.  1 
               A      B  stat p.value   
Smicridea 0.9307 0.8491 0.889   0.003 **

 Group Alluvial_fans+high-Andean+Paramo+Piedmont  #sps.  2 
                     A      B  stat p.value
```

```
Baetodes        0.9999 0.9245 0.961   0.001 ***
Thiennemaniella 0.9998 0.8868 0.942   0.001 ***

 Group Guiana_shield+high-Andean+highPlains+Paramo  #sps.  1
              A      B  stat p.value
Bezzia.cf 0.9551 0.7895 0.868   0.001 ***

 Group Alluvial_fans+Guiana_shield+high-Andean+highPlains+Piedmont
#sps.  1
                 A      B  stat p.value
Polypedilum 1.0000 0.9492 0.974   0.001 ***

 Group Alluvial_fans+Guiana_shield+high-Andean+Paramo+Piedmont  #sps
.  2
                    A      B  stat p.value
Cricotopus     0.9998 1.0000 1.000   0.001 ***
Rheotanytarsus 0.9921 0.9516 0.972   0.001 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

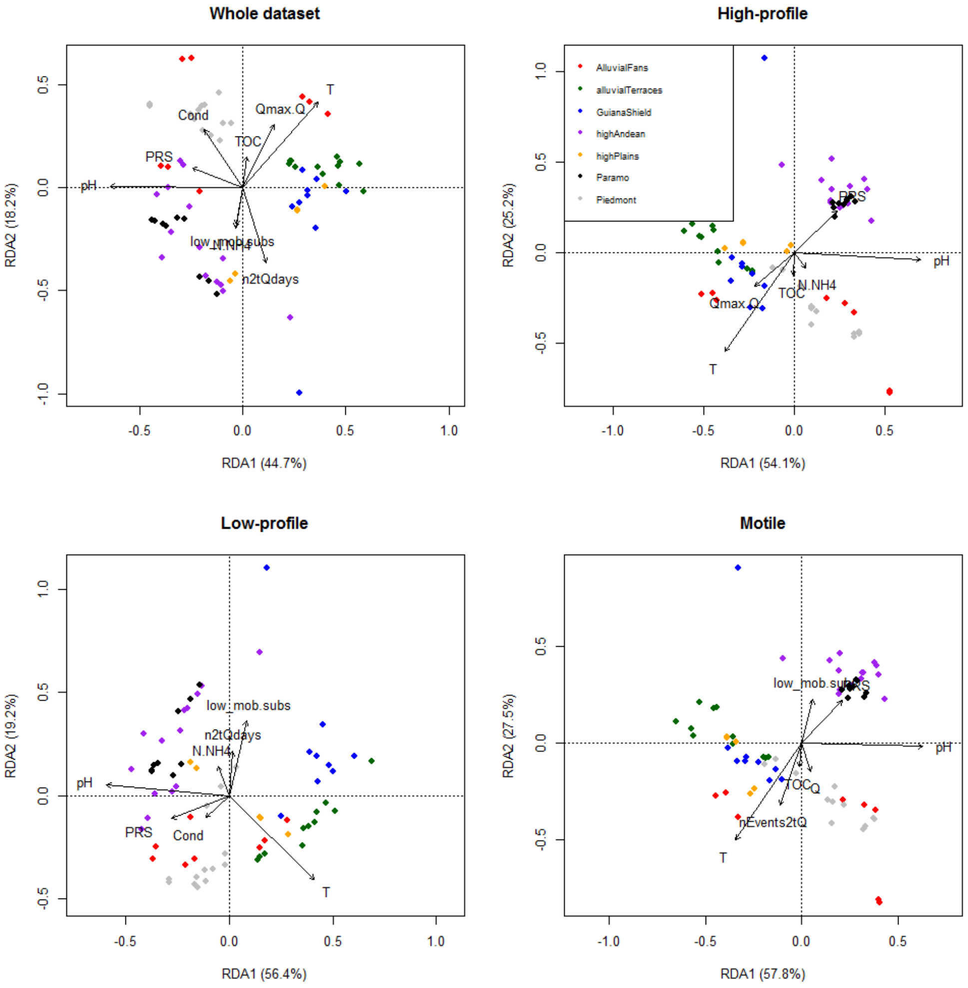
relationships between elevation and water chemistry variables. Only the variables with the highest correlation coefficient are showed here.



Relationships between the dissimilarity (ß-diversity) of diatom communities and elevation difference, estimated using the whole community dataset (A), and the datasets including low-profile (B), high-profile (C) and motile (D) diatoms.

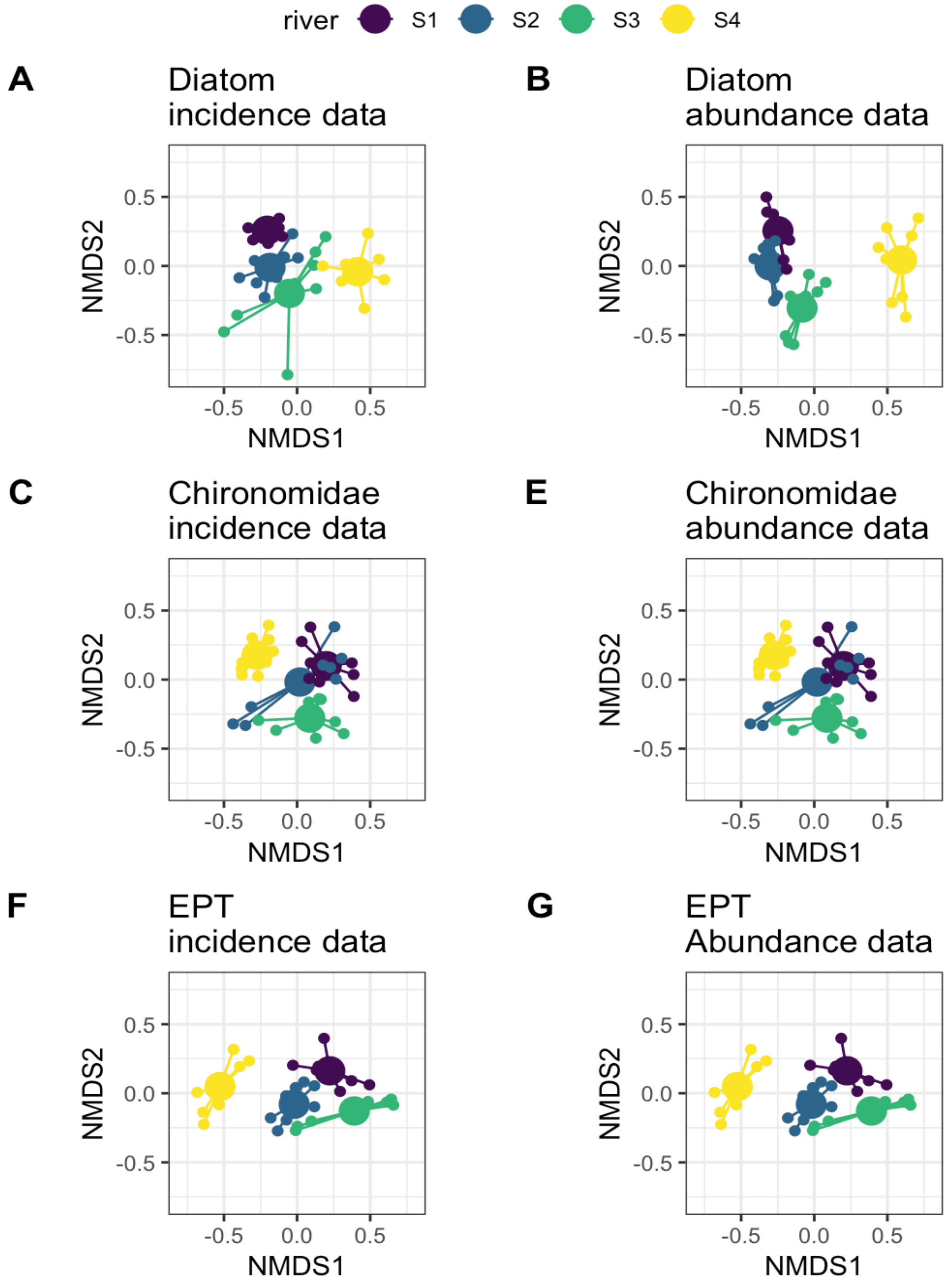


RDA triplots for the first and second significant axes (P<0.001) of each model, scaling 2. Community data was Hellinger-transformed. The angle between vectors represents the correlation degree of the environmental variables with the axes, and between themselves.



Supplementary Material – Chapter V

Non-metrical multidimensional scaling (nMDS) showing the dissimilarity between the samples (points) taken in the four study streams. The greater the distance between points, the greater is the dissimilarity between samples in terms of species composition. Big circles represent the centroid of each stream. Incidence and abundance data were Hellinger-transformed.



Results of ANOSIM tests. Hellinger transformed data. 999 permutations.

Dataset	ANOSIM R statistic	p-vale
Diatom – incidence data	0.5939	0.001
Diatom – abundance data	0.7162	0.001
Chironomidae – incidence data	0.3927	0.001
Chironomidae – abundance data	0.4906	0.001
EPT – incidence data	0.7738	0.001
EPT – abundance data	0.5914	0.001

List of diatom taxa at the studied reaches. Black squares indicate the occurrence of a given taxa.

		S1			S2			S3			S4	
Genus	**Ref**	**Imp 1**	**Imp 2**	**Ref**	**Imp 1**	**Imp 2**	**Ref**	**Imp 1**	**Imp 2**	**Ref**	**Imp 1**	**Imp 2**
Achnanthidium sp	■	■	■	■	■	■	■	■	■	■	■	■
Cocconeis placentula	■	■	■	■	■	■	0	■	■	■	■	■
Craticula cf	■	■	0	0	0	■	■	0	0	0	■	■
Cymbella sp	0	■	■	■	■	0	0	■	■	■	■	■
Diatoma sp	0	0	0	0	0	0	0	0	0	0	■	0
Encyonema sp	■	■	■	■	■	■	0	■	■	■	■	■
Ephithemia sp	0	0	0	0	0	0	0	0	0	0	0	0
Eunotia sp	0	■	0	0	0	■	0	■	■	■	■	■
Fragilaria sp	■	■	■	■	■	■	■	■	■	■	■	■
Frustulia sp	■	0	0	■	■	■	0	■	■	■	■	■
Gomphoneis clevei	■	■	■	■	■	■	■	■	■	■	■	0
Gomphonema sp	■	■	■	0	■	■	0	0	0	0	0	0
Hannaea arcus	0	0	0	0	■	■	0	■	■	■	■	0
Melosira varians	■	■	■	0	■	■	0	0	■	0	0	0
Meridion circulare	0	0	0	0	0	0	0	0	0	0	■	0
Navicula sp	■	■	■	■	■	■	0	■	■	■	■	■
Neidium sp	■	0	0	0	0	0	0	0	■	■	0	■
Nitzchia sp	■	■	■	0	0	■	0	■	■	■	0	■
Pinnularia sp	■	■	■	0	■	0	0	■	■	0	0	■
Planothidium lanceolatum	■	■	■	■	■	■	■	■	■	0	0	0
Psammothidium subatomoides	■	■	■	■	■	■	■	■	■	■	■	■
Reimeria sinuata	■	■	■	0	■	■	0	0	0	0	0	0
Rhoicosphaenia abreviata	■	■	■	0	■	■	0	■	■	0	■	0
Stauroneis obtusa	■	■	■	0	■	■	0	■	■	■	■	■
Surirella sp	■	■	■	0	0	0	0	0	■	0	0	0
Tabellaria sp	0	0	0	0	0	0	0	■	■	■	0	■

List of chironomid taxa occurring at the study streams. Black squares indicate the occurrence of a given taxa.

	S1			S2			S3			S4		
Genus	**Ref**	**Imp1**	**Imp 2**	**Ref**	**Imp 1**	**Imp 2**	**Ref**	**Imp1**	**Imp 2**	**Ref**	**Imp 1**	**Imp2**
Alotanypus	0	0	0	0	0	0	■	■	■	■	■	■
Cardiocladius	■	■	■	0	■	0	0	■	■	■	0	■
Corynoneura	■	■	■	■	■	■	■	■	■	■	■	■
Cricotopus	■	■	■	■	■	■	■	■	■	■	■	■
Genus 1	■	■	■	■	■	■	■	■	■	■	■	■
Gymnometriocnemus	0	0	0	■	0	0	0	0	0	0	0	0
Harnischia	0	0	0	0	0	0	0	0	0	0	■	0
Hudsonimyia	■	■	■	■	■	■	0	0	0	0	0	0
Lauterborniella	0	0	0	0	■	0	0	0	0	0	0	0
Lopescladius	■	■	0	0	■	0	■	0	■	0	0	■
Metriocnemus	0	0	0	■	0	0	■	0	■	0	■	0
Morfo X	■	■	■	■	■	■	■	■	■	■	■	■
Nanocladius	0	■	■	■	0	■	■	■	■	■	■	■
Nylotanypus.cf	■	0	0	0	0	0	0	■	■	0	■	■
Oliveiriella	■	■	0	■	0	0	0	0	0	0	0	0
Onconeura	■	■	■	■	■	■	■	■	■	■	■	■
Paraheptagyia	0	0	0	0	0	0	■	■	■	0	0	0
Parakiefferiella	■	■	■	■	■	■	0	■	0	0	0	0
Parametriocnemus	■	■	■	■	■	■	■	■	■	■	■	■
Pentaneura	■	■	■	■	■	■	■	■	■	■	■	■
Podonomopsis	■	■	■	■	0	0	■	0	0	0	■	0
Podonomus	■		■	■	■	■	■	■	■	■	■	■
Polypedilum	■	■	■	■	■	■	■	■	■	0	■	■
Rheotanytarsus	■	■	■	■	■	■	■	■	■	■	■	■
Riethia	■	■	■	■	■	■	■	0	■	■	■	0
Stenochironomus	0	0	■	■	0	0	0	■	0	0	■	0
Tanytarsus	■	■	■	■	■	■	■	■	■	■	■	■
Thiennemaniella	■	■	■	■	■	■	■	■	■	■	■	■

List of EPT taxa occurring at the study streams. Black squares indicate the occurrence of a given taxa.

		S1			S2			S3			S4		
Order	**Genus**	**Ref**	**imp1**	**Imp 2**	**Ref**	**imp 1**	**Imp 2**	**Ref**	**imp1**	**Imp2**	**Ref**	**imp 1**	**Imp 2**
Ephemeroptera	*Andesiops*	■	■	■	■	■	■	■	■	■	■	0	0
	Baetodes	■	■	■	■	■	■	■	■	■	0	0	0
	Farrodes	■	■	■	■	0	0	0	■	■	0	0	0
	Haplohyphes	■	■	■	■	■	■	0	0	■	0	0	0
	Leptohyphes	0	■	■	■	■	■	0	■	■	0	0	0
	Leptohyphodes	■	0	0	0	0	0	0	0	■	0	0	0
	Thraulodes	0	0	■	0	0	0	0	0	0	0	0	0
	Tricorythodes	■	■	■	■	■	■	0	0	0	0	0	0
	Tricorythopsis	0	0	0	0	■	■	0	0	0	■	■	■
Plecoptera	*Anacroneuria*	■	■	■	0	■	■	■	0	■	0	0	0
Trichoptera	*Atopsyche*	0	■	■	■	■	■	■	■	■	■	■	■
	Ceratotrichia	0	■	■	0	0	0	0	■	0	0	0	0
	Contulma	0	■	■	■	■	■	■	0	■	■	■	■
	Culoptila	0	■	■	0	0	■	0	0	0	0	0	0
	Grumichella	■	■	0	■	■	■	■	0	0	0	0	0
	Helicopsyche	■	■	■	■	■	■	■	■	■	0	0	0
	Hydroptila	0	■	■	■	0	0	0	0	0	■	0	0
	Leptonema	0	■	■	0	0	0	0	0	0	■	■	■
	Marilia	0	0	0	0	0	0	0	■	■	■	■	■
	Metrichia	0	■	■	■	■	■	■	■	■	0	0	0
	Nectopsyche	0	■	■	■	■	■	■	■	■	■	0	0
	Phylloicus	0	0	0	■	0	■	0	0	0	■	■	■
	Smicridea	■	■	■	■	■	■	■	■	■	0	0	0
	Triaenodes	■	0	0	0	0	0	0	0	0	■	■	■
	Triplectides	0	0	0	■	0	■	0	0	■	0	0	0
	Xiphocentron.cf	0	■	0	■	■	■	■	0	0	0	0	0

Supplementary Material – Chapter VI

Estimation of the functional spaces of diatom and insect communities

The functional space was estimated by calculating the Gower distances between each pair of genera in the community. Gower's distance allows different types of variables to be mixed while giving them equal weights (Pavoine *et al.*, 2009). After computing distances, a PCoA was carried out on the resultant distance matrix (Villéger *et al.*, 2008). The species coordinates in the first three axes were retained as the values of three synthetic functional traits describing diatom and insect functional diversity. The ecological meaning of these axes was interpreted based on the loadings of trait values.

Functional space of diatom communities

The first three PCoA axes captured about 47% of the variation present in diatom traits. This variation was split nearly evenly among the first (ca.19%), second (ca. 15%), and third (ca. 13%) axes. Given the loadings of the original traits, we interpret the first three axes as describing the variability of diatoms in terms of their successional stage and resistance to disturbances such as flow spates.

The first axis can be interpreted as a descriptor of species successional stage (Fig.SM1 and 2): pioneers (solitary and prostrate species of 'nano' sizes) had negative scores on axis 1, whereas late-succession species (weakly-attached or non-attached species of large size and those that conform filamentous or colonies) had positive scores on axis 1. In the center of this axis were mid-succession species (medium size species with structures to capture light more efficiently such as pad-stalks).

The second axis can be interpreted as a descriptor of species preference for habitats with different water flow velocity (Fig. SM1). Species with traits indicating that they occupy streams with low water flow velocity (non-attached diatoms with very large body size) had positive scores on axis

2, whereas species with traits indicating that they occupy streams with fast water flow velocity (different attachments to substratum and life forms) had negative scores on axis 2. Finally, we interpret the third axis as representing life-form complexity, sorting species from solitary to colonial life forms (Fig. SM2).

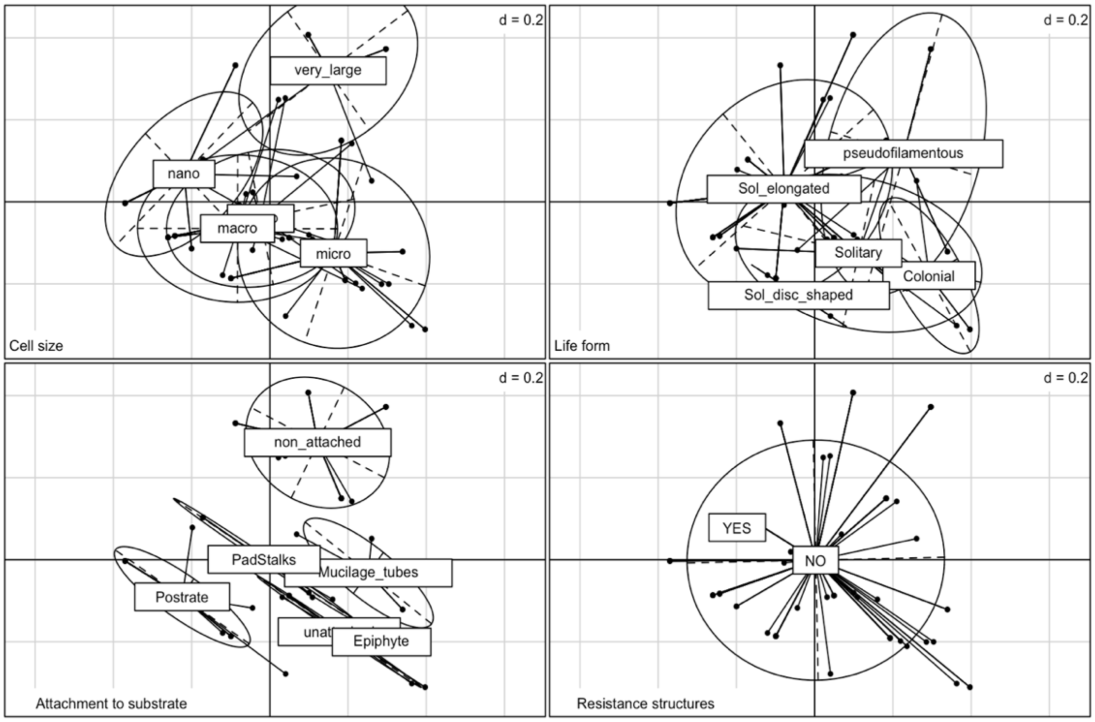


Figure SM-VI-1 Biplot displaying the structure of diatom functional space on the first two principal coordinate axes. Each data point is a morphospecies. To improve visualization, the categories of each traits are represented individually (Trait names are in the bottom-left corner). The loadings of each original trait are visualized as boxes. The circle around each box represents the area in which the 95% of observations (species) can be found.

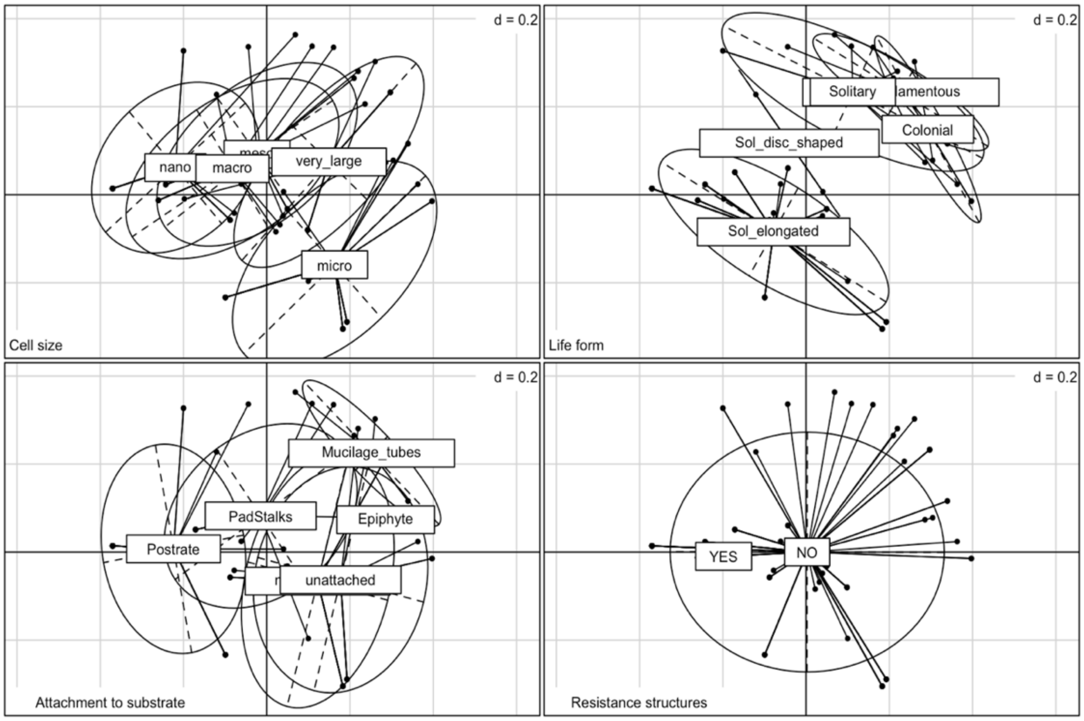


Figure SM-VI-2 Biplot displaying the structure of diatom functional space on the first and third principal coordinate axes. Each data point is a morphospecies. To improve visualization, the categories of each traits are represented individually (Trait names are in the bottom-left corner). The loadings of each original trait are visualized as boxes. The circle around each box represents the area in which the 95% of observations (species) can be found.

Functional space of insect communities

The first three PCoA axes captured about 45% of the variation present in insect traits. This variation was mainly captured by the first axis (ca.25%), and in a minor degree by the second (ca. 11%) and third (ca. 9%) axes. Given the loadings of the original traits, we interpret the first three axes as describing the variability of insect in terms of their specialization to occupy habitats with different water flow conditions, oxygen concentration and food availability.

The first axis can be interpreted as a gradient in predatory and non-benthic habitat specialization (Fig. SM2.3): large species with traits describing predatory specialists (swimmers with low efficiency of oxygen uptake, suggesting they use the water column or fast-flowing habitats) had negative scores on axis 1, whereas species with more generalized traits (different respiration systems, body sizes and feeding requirements) had positive scores on axis 1.

The second axis can be interpreted as a specialization gradient in terms of oxygen uptake and adaptations to flow constraints (Fig. SM2.3). Species with traits indicating no adaptation to flow constraints and low efficiency in oxygen uptake (spiracles and gas gills) had positive scores on axis 2, whereas species with a greater number of adaptations to flow constraints (suckers or silt glands) and adaptations indicating a high efficiency in oxygen uptake (tracheal gills) had negative scores on axis 2.

Finally, we interpret the third axis can be interpreted as a gradient in body size and in specialization towards filter-feeding behavior (Fig. SM2.3). Insect body size increased from negative to positive scores on axis 3. On the other hand, species with traits describing a filter-feeding behavior had negative scores on axis 3.

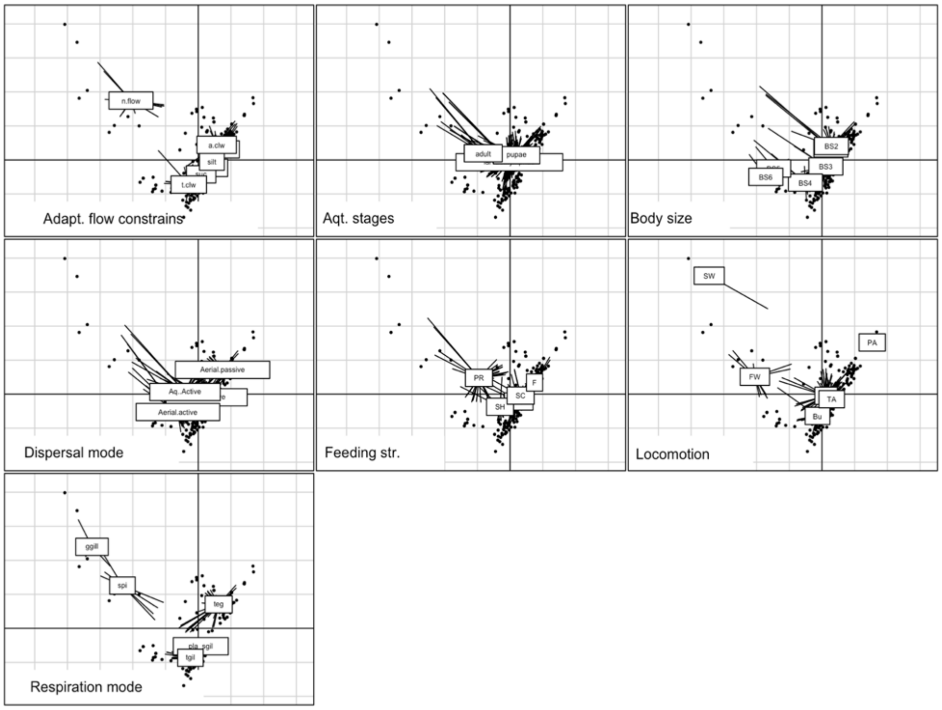


Figure SM-VI-3 Biplot displaying the structure of diatom functional space on the first two principal coordinate axes. Each data point is a genus. To improve visualization, the categories of each traits are represented individually (Trait names are in the bottom-left corner). The loadings of each original trait are visualized as boxes.

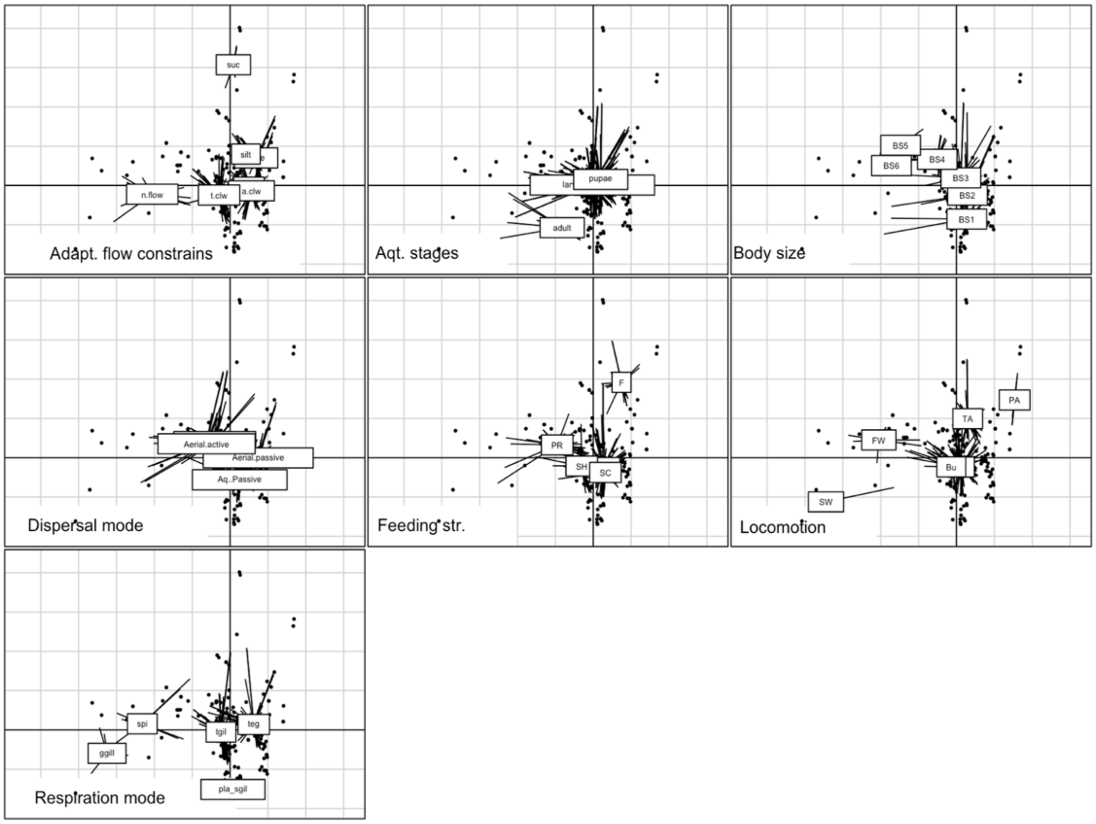


Figure SM-VI-4. Biplot displaying the structure of diatom functional space on the first and third principal coordinate axes. Each data point is a genus. To improve visualization, the categories of each traits are represented individually (Trait names are in the bottom-left corner). The loadings of each original trait are visualized as boxes.