



HAL
open science

Patterns and processes of neglected biodiversity through the lens of environmental DNA

Lucie Zinger

► **To cite this version:**

Lucie Zinger. Patterns and processes of neglected biodiversity through the lens of environmental DNA. Life Sciences [q-bio]. Université de Toulouse III - Paul Sabatier, 2024. tel-04799149

HAL Id: tel-04799149

<https://hal.science/tel-04799149v1>

Submitted on 22 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Candidature pour Habilitation à diriger des recherches

**Patterns and processes of neglected biodiversity
through the lens of environmental DNA**

Lucie Zinger

Devant le jury composé de

Isabelle Domaizon (DR INRAE, rapporteuse)
Emma Rochelle-Newall (DR IRD, rapporteuse)
Philippe Vandenkoornhuyse (PR Université de Rennes, rapporteur)
Régis Cereghino (PR Université Toulouse III, examinateur)
Jérôme Chave (DR CNRS, parrain)

2024

Contents

Acknowledgements	5
Presentation of the candidate	7
1 Curriculum Vitae	9
1.1 General information	9
1.2 Education and diplomas	9
1.3 Employment	10
1.4 Teaching activities	10
1.5 Teaching responsibilities	11
1.6 Research themes and production	11
1.7 List of funded research projects	11
1.8 Other fellowships and distinctions	13
1.9 Research supervision	13
1.10 Other responsibilities / activities	18
2 Complete list of publications and oral communications	23
2.1 Articles published in peer-reviewed journals	23
2.2 Books, book chapters, dissertations	30
2.3 Oral presentations, seminars, invited lectures and posters.	31
Dissertation	35
3 List of Abbreviations	37
4 Introduction	39
4.1 A brief refresher of community ecology	39
4.2 A short history of microbial biogeography	41
4.3 1990-2010's: a series of molecular revolutions	44
5 A DNA-based toolbox for biodiversity research	47
5.1 General context	47
5.2 Molecular biases in eDNA metabarcoding	49
5.3 Bioinformatics for eDNA data	51
5.4 How to sample eDNA?	59
5.5 Ecological analyses of eDNA data	63
5.6 Conclusion	66
6 Patterns and processes of the neglected biodiversity	69
6.1 General context	69
6.2 Local spatio-temporal dynamics of soil microbial communities	71
6.3 A microbial perspective from landscapes to the global scale	75

6.4	The particular case of host islands as habitats	83
6.5	Towards an understanding of the multi-trophic neglected biodiversity	89
6.6	Conclusion	96
7	Perspectives of research	99
7.1	General context	99
7.2	Towards eDNA-based observatories of tropical forests?	101
7.3	Tropical above-belowground feedbacks in a changing world	107
	Bibliography	113

Acknowledgements

First, I would like to thank the three external experts Isabelle Domaizon, Emma Rochell-Newall, and Philippe Vandenkoornhuysen for reviewing this work. I also would like to thank Régis Céréghino for accepting to represent the University of Toulouse, and Jérôme Chave for being my sponsor. I'm deeply grateful to all of you for accepting to be part of this HDR jury.

Throughout my scientific career, I have had the privilege of crossing the path of several exceptional individuals who have played a significant role in shaping me into the researcher I am today. To them, I would like to express my deepest gratitude. First, to Roberto Geremia, my PhD thesis advisor, for his consistent belief in my abilities and continued support, even after the completion of my PhD thesis. Second, to Pierre Taberlet and Eric Coissac. Almost all my background in molecular ecology and bioinformatics comes from you, and your intuition, creativity and energy is a great source of inspiration. I fondly remember you as “les tontons” back then, but with the passage of time (and the appearance of grey hair on my head) we've grown into a fantastic professional partnership. Third, to Antje Boetius, for introducing me to the wonders of oceanography and the fascinating world of microbial life in the oceans and for being a role model as a woman leader in an academic world that, even today, is still challenging for us. Fourth, to Chris Bowler, for welcoming me into his team upon my recruitment at ENS, for his trust, for opening the doors to the Tara universe, and for the many opportunities I've gained through our collaboration. Last but not least, I'm also deeply grateful to Jérôme Chave. Jérôme has played an important role in my life. Professionally through his guidance and support, and by taking me on a journey under a more theoretical perspective and of discovery through the tropical forests. But also personally, as it is through Jérôme that I met Renato, the father of my son. Jérôme, I recall how much I was intimidated when I first met you at the Nouragues Research Station. Now, I view you with great fondness and friendship, not without admiration for your scientific excellence and dedication to our field.

This journey wouldn't have been the same without other key figures, which I would like also to thank for their trust and for the stimulating discussions we had / still have, both within the framework of our collaborations or on other topics, namely Julien Cote, Frédéric Boyer, Staffan Jacob, Aurélie Bonin, Heidy Schimann, Wilfried Thuiller, Guillaume Besnard, Silvia de Monte, Jérôme Orivel, Céline Leroy, Mélanie Roy, Angélique Gobet, Johan Pansu. There are many others. Thanks to all of you.

I would also like to express my special thanks to the members of the B2M team and especially to Amaia Iribar, who I see as a sister, to Anne-Sophie Benoiston, to Uxue Suescun, to Lucie Moreau, and to Yves Cuenot without whom there would be almost no eDNA studies and developments at CRBE. It's a real pleasure to work and collaborate with them and I'm deeply grateful for all the technical support they provide to us. I have also special thoughts to Ludo, Delphine, and Xtian who ensure the same technical support at the LECA.

This work wouldn't have been possible without the great PhD students and post-docs I had the chance to work with and from who I'm learning a lot both scientifically and humanely. Fédérico, Julian, Léa, Emma, Jérémy, Léonard, Pedro, thanks a lot for your work, your patience, your creativity and your enthusiasm. I also have a special thought for Caroline, the first PhD student I co-mentored. I'm very proud of what she accomplished. From pharmacology to microbial community ecology. What a journey!

The friendly atmosphere at former EDB (and now CRBE) would not have been the same without its members. I'm thinking in particular to Antoine, Robin, Lisa, Joël, Pierrick, Pablo, Thierry, Aurèle.

I have also a special and heartfelt thought for my fellow members of the pedagogic team at the ENS. Andréa, Morgane, Marie, Barbara, Helena, Renaud, Pierre, Jean-François, Samuel, and more lately Amaury. I admire your dedication. My time at the Biology Department hasn't been easy, but your welcome, your guidance and support, and your positivity have been invaluable. Thank you so much!

Finally, my thoughts go to my dearest friends, Yza and Greg, to my parents, to Renato, and my wonderful little Gaspar. Thank you for your unconditional love and support. I know I'm not always easy to deal with. I love you all.

Presentation of the candidate

Curriculum Vitae

Candidate: Lucie ZINGER
Graduate school: SEVAB

1.1 General information

Name: Lucie, Tiphaine ZINGER
Date and place of Birth: 31st December 1982, Paris XIV, France
Nationality: French
Family status: Maritale life, 1 child (born 2020)

Current function: Assistant professor from ENS Ulm - PSL University in secondment as CNRS researcher
Professional address: Centre de Recherche sur la Biodiversité et l'Environnement (CRBE)
UMR 5300 CNRS-UT3-IRD-INP
Université Toulouse III – Paul Sabatier
Bât. 4R1, 118 route de Narbonne
31062 Toulouse cedex 9, FRANCE
Phone/Whatsapp: +33 625 458 707
Emails: lucie@zinger.fr, lucie.zinger@univ-tlse3.fr

Author IDs: ResearcherID: D-2527-2010
ORCID: 0000-0002-3400-5825
Google Scholar: link
Website: luciezinger.wordpress.com
Twitter: @LucieZinger

1.2 Education and diplomas

2009 **PhD** in Ecology, Université Joseph Fourier¹, Grenoble, France.
2005 **Master** in Ecology, Université Joseph Fourier, Grenoble, France.
2005 **Maîtrise** in Ecology, Université Joseph Fourier, Grenoble, France.
2003 **Licence** in Biology, Université Joseph Fourier, Grenoble, France.
2002 **DEUG** in Biology, Université Joseph Fourier, Grenoble, France.
2000 **Scientific Baccalaureate**, Lycée les Eaux Claires, Grenoble, France.

¹Now Université Grenoble Alpes

1.3 Employment

- 2024 - **Secondement as CNRS researcher**,
Centre de Recherche sur la Biodiversité et l'Environnement (CRBE), Toulouse,
France.
- 2022 - **Correspondent researcher**
for Naturalis Biodiversity Center, Leiden, the Netherlands.
- 2021 **Senior researcher** (Tenure Track). Declined for personal reasons.
Naturalis Biodiversity Center, the Netherlands.
- 02/2020 - 05/2020 Maternity leave.
- 2016 - **Associate professor**,
Institut de Biologie de l'École Normale Supérieure (IBENS), Paris, France.
- 2013 - 2016 **Post-doctoral researcher**,
Laboratoire Évolution & Diversité Biologique (EDB), Toulouse, France.
Project: Community assembly of tropical soil organisms through eDNA.
Supervision: Jérôme Chave.
- 2011 - 2013 **Post-doctoral researcher**,
Laboratoire d'Écologie Alpine (LECA), Grenoble, France.
Project: Metatranscriptomics for soil microeukaryotes,
Supervision: Éric Coissac.
- 2009 - 2011 **Post-doctoral researcher**,
Max Planck Institute for Marine Microbiology, Bremen, Germany.
Project: Macroecological patterns of marine bacteria.
Supervision: Alban Ramette and Antje Boetius.
- 2005 - 2009 **PhD fellow**,
Laboratoire d'Écologie Alpine (LECA), Grenoble, France.
Project: Spatio-temporal dynamics of the alpine soil microflora,
Supervision: Roberto Geremia.

1.4 Teaching activities

- 2016 - 2022 **Lectures and practicals** (192h eqTD/y) for Licence 3 and Master students
in Ecology, Macroecology and Biogeography, Environmental Genomics, and
Biostatistics.
Referee for L3 and M1 internship reports and defences.
Tutoring of 10 ENS students.
- 2012 - Regular **invited lecturer** in thematic schools where I teach eDNA-based com-
munity ecology, and animate bioinformatics, and/or biostatistics tutorials.
- 2005 - 2008 **Practicals** (64h eqTD/y) in Ecology, Ecotoxicology, and Plant Biology for
Licence 1-3 students.

1.5 Teaching responsibilities

- 2021 - 2022 **Responsible of the Master 2 Ecology & Evolution path** of the Master IMaLiS (Interdisciplinary Master in Life Sciences) of ENS and PSL.
- 2018 - 2019 **Co-organisator** of metabarcoding thematic schools. [link](#).
- 2017 - 2021 **Co-organisator** of L3 and M1 students recruitments for the ENS Biology Department.
- 2016 - 2022 **(Co-)coordinator of modules:** Ecology & Evolution (Licence 3), Data Analysis (Master 1/2), and Biology of Ecological Systems (Master 1/2) from the Master IMaLiS.
- 2016 - 2022 **Responsible of L3 internship defence** in Ecology for the ENS Biology department.

1.6 Research themes and production

Themes microbial macroecology and community ecology, host-microbiote interactions, land use and climate change effects on soil/aquatic food webs, environmental DNA, bioinformatics.

Publications 70 peer-reviewed articles (15 as lead author, 6 as senior author, 13 as 2nd author). These include publications in high-impact journals like *Cell* (1), *Ecology Letters* (1), *Nature Ecology & Evolution* (2), *Trends in Ecology & Evolution* (2), *the ISME Journal* (1).
1 book, 3 book chapters.
ISI h-index=31, Times cited=3712 (17/06/2024).
Google scholar h-index=36, i10 index=61, Times cited=5655 (17/06/2024).
1 R package.
See complete list in section 2

1.7 List of funded research projects

- 2024 - 2027 **MONITOR** (PEPR FORESTT): Smart ecological forest monitoring systems (Task-lead with Flavien Chantreau [RNF], PIs Jérôme Chave [CRBE] and Camille Piponiot-Laroche [AMAP], 5M€).
- 2024 - 2027 **junior package TULIP** (LabEx TULIP): Above-belowground feedbacks in tropical forests in a changing world (PI, 300k€).
- 2023 - 2026 **BARCODRAIN** (OFB): Developing tools based on rain eDNA for the monitoring of biodiversity in tropical ecosystems (co-PI with Amaia Iribar [CRBE], 170k€). =>website.
- 2023 - 2027 **MARCO-BOLO** (Horizon EU 2022): Marine and coastal biodiversity and organismal long-term observation framework (WP lead², PI: Nicolas Pade [EM-BRC], 8m€). =>website.

- 2023 - 2026 **EdaFoBio** (Spanish ministry of science and innovation): Global Survey of Edaphic Arthropod Forest Biodiversity: understanding the eco-evolutionary patterns and processes of diversification and community assembly of soil forest arthropods (participant, PI: Carmelo Andujar, [IPNA-CSIC]). =>website.
- 2022 - 2025 **NeoFINEO** (CESAB/FAPESP/SinBiose) : Networks of Fungal Interactions in the Neotropics (participant, PI: Paulo R. Guimarães, [LET,USP] and Mélanie Roy [IFAECI], 200k€). =>website.
- 2022 - 2023 **DYNAMICS** (AnAEE France): eDNA degradation rate and processes across matrices in micro and macrocosms (participant, PI: Béatriz Decencièrre [CEREPP-Ecotron-IdF], 30k€).
- 2021 - 2026 **TransAlp** (ANR): Linking multi-trophic diversity and ecosystem functions in a warming world in alpine grasslands (WP co-lead, PI: Tamara MÜNKENMÜLLER [LECA], 487 k€). =>website.
- 2021 - 2024 **INTERACTIF** (LabEx CEBA): Deforestation and climate change effect on neotropical freshwater ecosystems (WP co-lead, PI: Céline Leroy [EcoFoG/AMAP], 100 k€). =>website.
- 2020 - 2024 **XPRIZE Rrainforest - Brazilian team**: Developing new technologies and strategies to study tropical forests (WP lead³, PI: Vinicius Castro-Souza [ESALQ]). =>website.
- 2020 - 2022 **HOLOBROM** (ANR): Maternal and environmental effects on the microbial fluxes from parent to their offspring (participant, PI: Céline Leroy [EcoFoG/AMAP], 50 k€). =>website.
- 2020 - 2023 **SOILInvaders** (Spanish ministry of science and innovation): Magnitude, drivers and risks of mesofauna introduction within soils: a toolbox for biosecurity and conservation of insular biodiversity (participant, PI: Paula Arribas [IPNA-CSIC]).
- 2019 - 2025 **ECOFEEED** (ERCCoG): Altered eco-evolutionary feedbacks in a future climate (WP co-lead, PI: Julien Cote [EDB], 2 m€). =>website.
- 2019 - 2021 **EPIFUN** (ANR): Epiphytism as a Fungal Ecosystem conquered by Vascular Plants (WP co-leader, PI: Florent Martos [ISYEB], 288 k€).
- 2017 - 2018 **VERTIGE** (LabEx CEBA): Phylogenetic and functional diversity of phyllosphere communities (participant, PI: Heidy Schimann [EcoFoG], 20 k€).
- 2016 - 2021 **GlobNets** (ANR): Global biogeography of soil ecological networks (WP leader, PI: Wilfried Thuiller [LECA], 751 k€). =>website.
- 2016 - 2019 **DIAMOND** (LabEx CEBA): Monitoring amazonian diversity (participant, PIs: Jérôme Murienne [EDB], and Jérôme Orivel [EcoFoG], 277 k€). =>website.
- 2016 - 2017 **GUMDROP** (LabEx CEBA): Community assembly of the gut microbiota of tropical predatory arthropods (co-PI with David Duneau [EDB], 20 k€).

- 2014 – 2016 **DEFANTS** (CNRS PEPS): Diversity and metabolic potential of tropical ant cuticular microbiota (participant, PI: C. Duplais, EcoFoG, 20 k€).
- 2005 – 2008 **PhD MRT fellowship** (French Ministry of Higher Education and Research, 45 k€).

1.8 Other fellowships and distinctions

- jan 2023 USP/CAPES PrInt visiting Professor scholarship, Piracicaba, Brasil (2,5k€).
2022 “Hors-Classe” status as Associate professor position at ENS.
2020 Leave for research or thematic reconversion (CRCT).
2019 - 2022 “Institut de Convergence” INCEPTION: Teaching fellowships (4k€ / year).
jun 2018 Martin & Temminck visiting scholarship, Naturalis, Netherlands (1,5 k€).
2005 – 2008 Teaching assistant fellowship, Université Joseph Fourier, France (7 k€).

1.9 Research supervision

I have participated to the supervision of **master (3) and PhD students (4), engineers (2) and early career researchers (5)**, with different levels of involvement.

1.9.1 Master students

Below are reported only internships of which duration was longer than 3 months.

Emma Fromm	Degree:	M2 - Master Biodiversité, Écologie, Évolution, Sorbonne Université.
	Defence year :	2021.
	Host lab:	Laboratoire Évolution et Diversité Biologique (EDB).
	Topic:	Response of the lizard’s gut microbiota to warming.
	Supervision rate:	40% with Julien Cote (60%).
Marion Lemoine	Degree:	M2 - Master Biodiversité, Écologie, Évolution, Université Toulouse III.
	Defence year :	2017.
	Host lab:	Laboratoire Évolution et Diversité Biologique (EDB).
	Topic:	Disentangling the determinants of the gut microbiota in tropical predatory arthropods.
	Supervision rate:	50% with David Duneau (50%).
Guillaume Lentendu	Degree:	M2 - Master Biodiversité, Écologie, Évolution, Université Toulouse III.
	Defence year :	2009.
	Host lab:	Laboratoire d’Écologie Alpine (LECA).
	Topic:	Development of a bioinformatic tools to estimate microbial diversity through 454 sequencing.

³While I wrote the WP proposal and am still involved in the project, I resigned from the WP leader position when it got funded due to personal reasons and new professional opportunities.

³Resigned from WP position after winning the semifinal in July 2023 due to new working constraints. Now participant.

Supervision rate: 10% with Éric Coissac (10%), Stéphanie Manel (40%), Christelle Melo de Lima (40%).

Published work: 1) Lentendu G.*, **Zinger L.***, Manel S., Coissac E., Choler P., Geremia R.A., Melo de Lima C. (2011) Assessment of soil fungal diversity in different alpine tundra habitats by means of pyrosequencing, *Fungal Diversity* 49(1):113:123. =>link.

* equal contributions

1.9.2 PhD candidates

Below I list only official supervision or mentoring, but I've been also involved in the mentoring (~5-10%) of other PhD students on matters related to eDNA and microbial ecology such as Staffan Jacob, Julien Roy, Guilhem Sommeria-Klein, or Irène Calderon-Sanou.

Caroline Birer	Host lab:	Laboratoire Écologie des Forêts de Guyane (EcoFoG).
	Graduate school:	ED Diversités, Santé et développement en Amazonie, Université de Guyane
	Starting year:	2014
	Defence year :	2017.
	Topic:	Bacterial microbiota of ant's cuticle in French Guiana : antibiotic activities and community ecology.
	Supervision rate:	non-official supervision (listed as co-mentor only). In practice 25% with Bruno Figadère (10%, HDR), C. Duplais (40%, HDR obtained during the thesis), and Niklas Tysklind (25%).
	Published work:	1) <u>Birer C.</u> , Moreau C.S., Tysklind N., Zinger L. *, Duplais C.* (2020) Determinants of bacterial communities and metabolites on the cuticle of two Amazonian ants species living in the same nest. <i>Molecular Ecology</i> 29(7):1372-1385. =>link. 2) <u>Birer C.</u> , Tysklind N., Zinger L. *, Duplais C.* (2017) Comparative analysis of DNA extraction methods to study the body surface microbiota of insects: a case study with ant cuticular bacteria. <i>Molecular Ecology Resources</i> 29(7):1372-1385. =>link.
	Current situation:	Assistant Professor at Université de Tours, Laboratoire de Biologie Cellulaire et Biochimie Végétale, EA 2106 - Biomolécules et Biotechnologies Végétales (BBV), UFR de Sciences Pharmaceutiques.
Léonard Dupont	Host lab:	Station d'Écologie Théorique et Expérimentale (SETE)
	Graduate school:	ED Sciences Écologiques, Vétérinaires, Agronomiques et Bioingenieries, Université Toulouse III.
	Starting year:	2021.
	Defence year :	scheduled for end of 2024.
	Topic:	Phenotypic plasticity in fluctuating environments : between microcosmes and natural communities.
	Supervision rate:	Officially 50% with Staffan Jacob (50%, HDR). In practice, I'm much less involved than Staffan in Leonard's thesis supervision. His thesis subject has evolved a little and relies less on my skills than what we planned.

Published work: 1) Dupont L., Thierry M., **Zinger L.**, Legrand D., Jacob S. (2023) Beyond reaction norms: the temporal dynamics of phenotypic plasticity, *Trends in Ecology and Evolution* 39(1):41-51. =>link.
 Current situation: Analyses and writing for Chapter II and III on-going.

Emma Fromm
 Host lab: Centre de Recherche sur la Biodiversité (CRBE).
 Graduate school: ED Sciences Écologiques, Vétérinaires, Agronomiques et Bioingenieries, Université Toulouse III.
 Starting year: 2021.
 Defence year : scheduled for 2025
 Topic: Response of the host-microbiota interactions to climate change and habitat fragmentation.
 Supervision rate: 50% with Julien Cote (50%, HDR)
 Published work: 1) Fromm E., **Zinger L.**, Pellerin F., Di Gesu L., Jacob S., Winandy L., Aguilée R., Parthuisot N., Iribar Pelozuelo A., Sescun U., White J., Bestion E., Cote J. (in press) Warming effects on lizard gut microbiota depends on habitat connectivity. *Proceedings of the Royal Society B* 291:20240220 =>link.
 Current situation: Analyses and writing for Chapter II and III on-going.

Jérémy Raynaud
 Host lab: Centre de Recherche sur la Biodiversité (CRBE).
 Graduate school: ED Sciences Écologiques, Vétérinaires, Agronomiques et Bioingenieries, Université Toulouse III.
 Starting year: 2024.
 Defence year : scheduled for 2027
 Topic: Plant-soil feedbacks in tropical forests in a changing world.
 Supervision rate: 100% (exemption for HDR obtained in Jan. 2024)
 Current situation: The thesis has started in April 2024. Analyses on already available datasets for Chapter I ongoing, Field sampling design and work in preparation for Chapters II and III.

* equal contributions

1.9.3 Engineers

I've been also co-supervising **2 engineers** in bioinformatics/bioanalyses

Anne-Sophie Benoiston
 Host Lab: Station d'Écologie Théorique et Expérimentale (SETE).
 Dates: 2021 - 2022.
 Supervision rate: 50% with Julien Cote (50%).
 Tasks: Building of a bioinformatic pipeline and analysis of the soil diversity data of the ECOFEED project. Field work and data analyses.
 Published work: In preparation.
 Current situation: IRD Engineer at the Centre de Recherche sur la Biodiversité et l'Environnement (CRBE).

Vincent Schilling
 Host Lab: Laboratoire Évolution et Diversité Biologique (EDB).
 Dates: 2014 - 2016.
 Supervision rate: 70% with Jérôme Chave (30%).

Tasks:	Development of a semi-automatic bioinformatic procedure to analyse eDNA metabarcoding data.
Published work:	Chalmandrier L., Pansu J., [...] Schilling V., Taberlet P., Zinger L. , Lavergne S., Thuiller W. (2019) Abiotic and biotic determinants of soil microbial communities -diversity across spatial and phylogenetic scales. <i>Ecography</i> 42:2144-2156. => link. Zinger L. , Taberlet P., Schimann H., Bonin A., Boyer F., [...] Schilling V., [...] Chave J. (2019) Body size determines soil community assembly in a tropical forest. <i>Molecular Ecology</i> 28:528-543. => link. Ohlmann M., Mazel F., Chalmandrier L., Bec S., Coissac E., Pansu J., Schilling V., Taberlet P., Zinger L. , Chave J., Thuiller W. (2018) Mapping the imprint of biotic interactions on β -diversity. <i>Ecology Letters</i> 21(11): 1660-1669. => link. Zinger L. , Chave J., Coissac E., Iribar A., Louisanna E., Manzi S., Schilling V., Schimann H., Sommeria-Klein G., Taberlet P. (2016) Extracellular DNA extraction is a fast, cheap and reliable approach for high-throughput soil biodiversity assessments from environmental DNA, <i>Soil Biology & Biochemistry</i> 96:16-19. => link.
Current situation:	Quality-Security-Environment assistant and Biotechnologies engineer at SIER Blagnac.

1.9.4 Post-doctoral researchers

Since 2016, I have also supervised (or is currently supervising) the work of **5 early career scientists**.

Léa Beaumelle	Host lab:	Centre de Recherche sur la Biodiversité (CRBE).
	Dates:	2022 -
	Supervision rate:	50% with Julien Cote (50%).
	Topic:	Response of soil multi-trophic communities to defaunation and warming.
	Published work:	In preparation.
Anne-Sophie Benoiston	Host lab:	Institut de Biologie de l'ENS (IBENS).
	Dates:	2019 - 2021.
	Supervision rate:	100%.
	Topic:	Response of soil communities to land use and development of bioinformatics tools for eDNA metabarcoding analyses. This contract was an ATER so it included a teaching part.
	Published work:	1) Zinger L. , Lionnet C., <u>Benoiston A.-S.</u> , Donald, J., Mercier, C., Boyer F. (2021) metabaR: an R package for the post bioinformatics evaluation and improvement of metabarcoding data quality, <i>Methods in Ecology and Evolution</i> 12(4):586-592. => link.
	Current situation:	IRD Engineer at the Centre de Recherche sur la Biodiversité et l'Environnement (CRBE).
Julian Donald	Host Lab:	Laboratoire Évolution et Diversité Biologique (EDB).
	Dates:	2019 - 2021.

Supervision rate: 30% with Jérôme Chave 40% and Jérôme Murienne 30%.
 Topic: Community assembly of the soil biota across tropical habitats with eDNA.
 Published work: 1) Donald J., Murienne J., Chave J., Iribar A., Manzi S., Tao S., Louisanna E., Roy M., Schimann H., Orivel J., Zinger L. (2021) Multi-taxa environmental DNA inventories reveal distinct taxonomic and functional diversity in urban tropical forest fragments. *Global Ecology and Conservation* 29: e01724. => link.
 2) **Zinger L.**, Lionnet C., Benoiston A.-S., Donald, J., Mercier, C., Boyer F. (2021) metabar: an R package for the post bioinformatics evaluation and improvement of metabarcoding data quality, *Methods in Ecology and Evolution* 12(4):586-592. => link.
 3) **Zinger L.**, Donald J., Brosse S., Iribar A., Gonzalez M.A., Leroy C., Murienne J., Orivel J., Schimann H., Taberlet P. Lopes C.M. (2020) Advances and prospects of eDNA in the Neotropical rainforests. *Advances in Ecological Research* 62:331-373. => link.
 Current situation: Senior advisor for the UK government conservation body.

Federico Ibarbalz Host lab: Institut de Biologie de l'ENS (IBENS).
 Dates: 2016 - 2019.
 Supervision rate: 50% with Chris Bowler (50 %).
 Topic: Macroecology of the sunlit plankton at the global scale.
 Published work: 1) Ibarbalz F.M., Henry N., Brandão M.C., Martini S., Busseni G., [...] Bowler C.*, **Zinger L.*** (2019) Global trends of marine plankton diversity across kingdoms of life. *Cell* 179(5):1084-1097.e21. => link.
 Royo-Llonch M., Sánchez P., Ruiz-González C., Salazar G., Pedrós-Alió C., [...] Ibarbalz F., **Zinger L.**, [...] Chris Bowler, Silvia G Acinas (2021) Compendium of 530 metagenome-assembled bacterial and archaeal genomes from the polar Arctic ocean. *Nature Microbiology* 6, 1561–1574. => link.
 2) Ibarbalz F., Henry N., Ardyna M., Lombard F., [...] **Zinger L.**, [...] Bowler C., Karp-Boss L. (2023) Pan-Arctic plankton community structure and its global connectivity. *Elementa* 11(1): 00060. => link.
 Current situation: Researcher at Centro de Investigaciones del Mar y la Atmosfera of the Buenos Aires University.

Pedro Junger Host lab: Institut de Biologie de l'ENS (IBENS).
 Dates: 2023 -
 Supervision rate: 20% with Chris Bowler (80 %).
 Topic: Macroecology of the sunlit plankton at the global scale and use of eDNA for its monitoring.
 Published work: In preparation.

* equal contributions

1.10 Other responsibilities / activities

1.10.1 Scientific committees

2022 - Member of the scientific committee of the CESAB (Centre de Synthèse et d'Analyse sur la Biodiversité).
 2022 - Member of the scientific committee of the B2M platform (Biologie Moléculaire et Microbiologie) of the CRBE.
 2021 - Member of the Tara Ocean consortium.

1.10.2 Editorial activities

2021 - **Recommender** for *Peer Community In Ecology*
 2021 - 2023 **Associated Editor** for *Journal of Biogeography*^A.
 2015 - **Associated Editor** for *Molecular Ecology* and *Molecular Ecology Resources*.
 30 manuscripts/y

1.10.3 Peer-review

Since 2009, I'm **referee for different peer-reviewed journals** including *Science*, *Science Advances*, *Nature Communication*, *Ecology Letters*, *ISME Journal*, *New Phytologist*, *Communication Biology*, *Ecography*, *Scientific Report*, *PLoS One*, *Frontiers in Microbiology*, *Environmental Microbiology Report*, *FEMS Microbiology*, and others.

1.10.4 Referee for funding agencies

Since 2015, I am invited as an **external expert to evaluate research proposals** for both national and international funding agencies/organisations.

At the **national level**, I have evaluated several research annual project proposals for the LabEx CEBA (Centre d'Étude de la Biodiversité Amazonienne), and the LabEx OSUG (Observatoire des Sciences de l'Univers de Grenoble). I have also assessed a research project for the CESAB. I have also evaluated one PhD exchange project for the French Embassy at São Paulo, Brazil (program TerEE).

At the **international level**, I have also evaluated several research projects: 1 for the FRQNT (Fonds de Recherche du Québec – Nature et Technologies), 1 for the SNSF (Swiss National Science Foundation), 1 for the NWO (Dutch Research Council), 1 for the COST Actions. I've been also invited to assess two ERC consolidator projects. I declined the first due to conflict of interest, and have to review the second for end of August 2024.

1.10.5 Jury staff recruitment

I have been member of **5 recruitment panels**: two engineers positions (one non-permanent position at CNRS, one permanent position at ENS), one research engineer permanent position (CNRS), and two assistant professor position (one at Université Grenoble Alpes and one at the University of Copenhagen).

1.10.6 Jury for PhD defences

Since 2021, I've been member of the PhD Defence Jury for **8 PhD students**, as “examinatrice”:

Didac Barroso Bergadà Host lab: UMR Agroécologie, Université Bourgogne Franche-Comté.
 Title: *Automatic learning of interaction networks from next-generation sequence data*.
 Supervision: David. A. Bohan.

^AI resigned in 2023 as 95% of the editorial board due to strong disagreements with Wiley's new OA policies.

- Defence year : 2022.
- Marine Biguet Host lab: Laboratoire Écosystèmes, Biodiversité, Évolution (ECOBIO), Université de Rennes I.
Title: *The grapevine as an Holobiont: validity of the concept, temporal dynamics, environmental drivers and application to the definition of the microbial terroir.*
Supervision: Philippe Vandenkoornhuyse, Véronique Chable, co-mentored with Sabrina Pernet.
Defence year : 2021
- Samuel Hamard Host lab: Laboratoire Écologie Fonctionnelle et Environnement (LEFE), Université Toulouse III.
Title: *Contribution of phototrophic microbes to peatland carbon uptake in a changing climate.*
Supervision: Régis Cereghino and Vincent Jassey.
Defence year : 2022.
- Coline Lenoir de Carlan Host lab: Plants and Ecosystems (PLECO), University of Antwerp.
Title: *Climate change direct and indirect effects on soil microbial communities in subarctic grasslands.*
Supervision: Erik Verburggen and Caroline de Tender.
Defence year : planned for the 9th of July 2024.
- Eduard Mas Carrio Host lab: Département d'Écologie et Évolution, Université de Lausanne.
Title: *Environmental DNA for management and ecosystem biomonitoring: Applications and methodological advances.*
Supervision: Luca Fumagalli.
Defence year : 2023.
- Laeticia Mathon Host lab: Centre d'Écologie Fonctionnelle et Évolutive (CEFE), Université de Montpellier.
Title: *Environmental DNA to describe large-scale fish distribution patterns and inform conservation.*
Supervision: Stéphanie Manel, co-mentoring by David Mouillot and Laurent Vigliola.
Defence year : 2023.
- Sylvain Moinard Host lab: Laboratoire d'Écologie Alpine (LECA), Université Grenoble Alpes.
Title: *Robust biodiversity estimators for metabarcoding.*
Supervision: Éric Coissac and Christelle Melo de Lima.
Defence year : 2023.
- Judith Schneider Host lab: Département d'Écologie et Évolution, Université de Lausanne.
Title: *Molecular assessment of wild vervet monkeys' diet to highlight social factors.*
Supervision: Luca Fumagalli and Erica van de Waal.
Defence year : 2023.

1.10.7 PhD advisory committees

Since 2013, I've participated to the advisory committees of **14 PhD students**:

- Irene Calderón-Sanou Host lab: Laboratoire d'Écologie Alpine (LECA), Université Grenoble Alpes.
Supervision: Wilfried Thuiller, Tamara Münkemüller.
Defence year: 2022.
- Isabel Cantera Host lab: Laboratoire Évolution & Diversité Biologique (EDB), Université Toulouse III.
Supervision: Sébastien Brosse.
Defence year: 2019.
- Lucie Cartairade Host lab: Centre de Recherches Insulaires et Observatoire de l'Environnement (CRIOBE), école Pratique des Hautes études.
Supervision: Serge Planes, Patrick Wincker.
Defence year: scheduled for 2025/2026.
- Loic Chalmandrier Host lab: Laboratoire d'Écologie Alpine (LECA), Université Grenoble Alpes.
Supervision: Wilfried Thuiller, Tamara Munkemuller, Sébastien Lavergne.
Defence year: 2015.
- Sacha Delmotte Host lab: Centre d'Écologie Fonctionnelle et Évolutive (CEFE), Université de Montpellier.
Supervision: Thibaud Decaëns, Alain Brauman, co-mentoring by Caroline Brunel, Antoine Versini.
Defence year: scheduled for 2025.
- Coralie Drake Host lab: Centre de Recherche sur la Biodiversité et l'Environnement (CRBE), Université Toulouse III.
Supervision: Joël White, Philipp Heeb.
Defence year: scheduled for 2025.
- Le Thi Huong Host lab: Laboratoire Écologie Microbienne (LEM), Université Claude Bernard.
Supervision : Thomas Pommier, Emma Rochelle-Newall.
Defence year: 2018.
- Maeva Iannelli Host lab: Laboratoire Eco&Sols, Université de Montpellier.
Supervision: Mikaël Hedde, co-mentoring by Claire Marsden and Jean Trap.
Defence year: withdrawal.
- Sébastien Kally Host lab: Laboratoire Évolution & Diversité Biologique (EDB), Université Toulouse III.
Supervision: Jérôme Muriene.
Defence year: 2019.
- Eduard Mas Carrio Host lab: Département d'Écologie et Évolution, Université de Lausanne.
Supervision: Luca Fumagalli.
Defence year: 2023.
- Johan Pansu Host lab: Laboratoire d'Écologie Alpine (LECA), Université Grenoble Alpes.
Supervision: Pierre Taberlet and Philippe Choler.
Defence year: 2014.
- Rémi Petrolli Host lab: Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle.
Supervision: Florent Martos and Marc-André Selosse.

Defence year: 2022.

Judith Schneider Host lab: Département d'Écologie et Évolution, Université de Lausanne.
Supervision: Luca Fumagalli and Erica van de Waal.
Defence year: 2023.

Valentin Verdon Host lab: Département d'Écologie et Évolution, Université de Lausanne.
Supervision: Antoine Guisan.
Defence year: scheduled for 2025.

1.10.8 Invited referee for Master theses

I've been also **external referee** of Master 2 internships in Ecology for the ENS Lyon, the Master Biodiversité, Écologie, Évolution (BEE) from Sorbonne Université, as well as for the Master écologie des Forêts Tropicales from Université de Guyane.

1.10.9 Outreach

2024-03-26 Discussion on eDNA on France Culture's radio program "CQFD". => [link](#).
2024-02-13 Interview for the "Compteur Biodiversité Outre-mer" on rain eDNA. => [link](#).
2020-12-23 Interview for France Culture "La Méthode Scientifique" the ocean's microbiome. => [link](#).
2019-11-15 Interview for France Inter on global warming and the ocean's plankton. => [link](#).
2019-11-15 Interview for Radio Canada on the latitudinal gradients of ocean's plankton. => [link](#).

Complete list of publications and oral communications

Publications written with students/early career scientists that I've supervised are **colored**. Equal contributions as lead author, second or senior author are indicated with a *.

2.1 Articles published in peer-reviewed journals

- a70. Litchman E., Villéger S., **Zinger L.**, Auguet J.-C., Thuiller W., Philippot L., Violle C. (in press) Refocusing the microbial rare biosphere concept through a functional lens. *Trends in Ecology and Evolution*.
Contribution: conceptualization, writing.
- a69. **Fromm E.**, **Zinger L.**, Pellerin F., Di Gesu L., Jacob S., Winandy L., Aguilée R., Parthuisot N., Iribar Pelozuelo A., Sescun U., White J., Bestion E., Cote J. (2024) Warming effects on lizard gut microbiota depends on habitat connectivity. *Proceedings of the Royal Society B*, 291:20240220, doi:10.1098/rspb.2024.0220.
Contribution: conceptualization, bioinformatics, writing.
- a68. Calderón-Sanou I., Ohlmann M., Münkemüller T., **Zinger L.**, Hedde M., Lionnet C., Martinez-Almoyna C., Saillard A., Renaud J., Le Guillarme N., Gielly L., Orchamp Consortium, Thuiller W. (2024) Mountain soil food webs shaped by the interplay between habitat and pedoclimatic conditions. *Soil Biology and Biochemistry*, 190:109282. doi:10.1016/j.soilbio.2023.109282.
Contribution: data analysis, manuscript revisions.
- a67. Fountain-Jones N., Giraud T., **Zinger L.**, Bik H., Creer S., Videvall E. (2024) Molecular ecology and microbiomes in the wild: methodological advances, common pitfalls and future directions. *Molecular Ecology*, 33(2):e17223. doi:10.1111/mec.17223.
Contribution: conceptualization, writing.
- a66. El Hourany R., Pierella Karlusich J.J., **Zinger L.**, Liosel H., Levy M., Bowler C. (2024) Linking satellites to genes with machine learning to estimate major phytoplankton groups from space. *Ocean Science*, 20:217–239. doi:10.5194/os-20-217-2024.
Contribution: conceptualization, writing.
- a65. **Dupont L.**, Thierry M., **Zinger L.**, Legrand D., and Jacob S. (2023) Beyond reaction norms: the temporal dynamics of phenotypic plasticity. *Trends in Ecology and Evolution*, 39(1):41-51. doi:10.1016/j.tree.2023.08.014.
Contribution: manuscript revisions | Media coverages: CNRS.
- a64. Schimann, H., Coste S., Fort, T., Vacher C, **Zinger L.** (2023) Determinants of the vertical distribution of the phyllosphere differ between microbial groups and the epi- and endosphere. *Phytobiomes Journal*,

- 7(3): 312-323. doi:10.1094/PBIOMES-02-23-0013-R.
Contribution: conceptualization, field work, analyses design, and writing.
- a63. [Ibarbalz F.](#), Henry N., Ardyna M., Lombard F., Jaillon O., Iudicone D., Malviya S., Lovejoy C., **Zinger L.**, Tara Oceans coordinators, Babin M., Boss E., Chaffron S., Wincker P., de Vargas C., Bowler C., Karp-Boss L. (2023) Pan-Arctic plankton community structure and its global connectivity. *Elementa*, 11(1):00060. doi:10.1525/elementa.2022.00060.
Contribution: conceptualization, writing.
- a62. Pierella Karlusich J.J., Pelletier E., **Zinger L.**, Lombard F., Zingone A., Colin S., Gasol J., Dorell R., Scalco E., Acinas S., Wincker P., de Vargas C., Bowler C. (2023) A robust approach to estimate relative phytoplankton cell abundance from metagenomes. *Molecular Ecology Resources*, 23(1):16-40. doi:10.1111/1755-0998.13592.
Contribution: conceptualization, writing.
- a61. Petrolli R., **Zinger L.**, Perez-Lamarque B., Collobert G., Griveau C., Pailler T., Selosse, M.A., Martos, F. (2022) Spatial turnover of fungi and partner choice shape mycorrhizal networks in epiphytic orchids. *Journal of Ecology*, 110(11):2568-2584. doi:10.1111/1365-2745.13986.
Contribution: conceptualization, manuscript revisions.
- a60. Arribas P., Andújar C., Bohmann K., deWaard J.R., Economo E., Elbrecht V., Geisen S., Goberna M., Krehenwinkel H., Novotny V., **Zinger L.**, Creedy T.J., Meramveliotakis E., Nogueras V., Overcast I., Morlon H., Papadopoulou A., Vogler A.P., Emerson B.C. (2022) Toward global integration of biodiversity data: a harmonised metabarcode data generation module for terrestrial arthropods. *Gigascience*, 11:1-12. doi:10.1093/gigascience/giac065.
Contribution: conceptualization, manuscript revisions.
- a59. Martinez-Almoyna C., Saillard A., **Zinger L.**, Lionnet C., Foulquier A., Gielly L., Piton G., The Orchamp Consortium, Münkemüller T., Thuiller W. (2022) Differential effects of soil food webs on decomposition in mountain ecosystems. *Soil Biology and Biochemistry*, 172:108771. doi:10.1016/j.soilbio.2022.108771.
Contribution: conceptualization, analyses, manuscript revisions.
- a58. Tedersoo, L., Bahram, M., **Zinger, L.**, Nilsson, H., Kennedy, P., Yang, T., Anslan, S., Mikryukov, V. (2022) Best practices in metabarcoding of fungi: from experimental design to results. *Molecular Ecology*, 31(10):2769-2795. doi:10.1111/mec.16460.
Contribution: manuscript revisions.
- a57. Calderón-Sanou I., **Zinger L.**, Hedde M., Martinez-Almoyna C., Saillard A., Renaud J., Gielly L., Khedim N., Lionnet C., Ohlmann M., Orchamp Consortium, Münkemüller T., Thuiller W. (2022) Energy and physiological tolerance explain multi-trophic soil diversity in temperate mountains. *Diversity and Distribution*, 28:2549-2564. doi:10.1111/ddi.13529.
Contribution: conceptualization, analyses, manuscript revisions.
- a56. Peguero G., Ferrín M., Grau O., Sardans J., Verbruggen E., Rojas I., Van Langenhove L., Murienne J., Iribar A., **Zinger L.**, Orivel J., Asensio D., Gargallo-Garriga A., Llusía J., Margalef O., Ogaya R., Richter A., Janssens I.A., Peñuelas J. (2022) Decay of similarity across tropical rainforest communities: integrating scale-dependency in the analysis of spatial and nutrient distances. *Ecology*, 103(2):e03599. doi:10.1002/ecy.3599.
Contribution: bioinformatics analyses, manuscript revisions.
- a55. Royo-Llonch M., Sánchez P., Ruiz-González C., Salazar G., Pedrós-Alió C., Labadie K., Paoli L., Tara Oceans Coordinators, [Ibarbalz F.](#), **Zinger L.**, Samuel Chaffron, Damien Eveillard, Eric Karsenti, Shinichi Sunagawa, Patrick Wincker, Lee Karp-Boss, Chris Bowler, Silvia G Acinas (2021) Ecogenomics of key prokaryotes in the arctic ocean. *Nature Microbiology*, 6:1561–1574. doi:10.1038/s41564-021-00979-9
Contribution: conceptualization, manuscript revisions
- a54. [Donald J.](#), Murienne J., Chave J., Iribar A., Manzi S., Tao S., Louisanna E., Roy M., Orivel J.,

- Schimann H., **Zinger L.** (2021) Multi-taxa environmental DNA inventories reveal distinct taxonomic and functional diversity in urban tropical forest fragments. *Global Ecology and Conservation* 29:e01724. doi: 10.1016/j.gecco.2021.e01724.
Contribution: co-supervision, writing.
- a53. Calderón-Sanou I., Münkemüller T., **Zinger L.**, Schimann H., Yoccoz N.G., Gielly L., Foulquier A., Hedde M., Olhmann M., Roy M., Si-Moussi S., Thuiller W. (2021) Cascading effects of moth outbreaks on subarctic soil food web. *Scientific Reports* 11:15054. doi:10.1038/s41598-021-94227-z.
Contribution: field work, molecular analyses, supervision of bioinformatics analyses, writing.
- a52. Rodríguez-Ezpeleta N., **Zinger L.**, Kinziger A., Bik H., Bonin A., Coissac E., Creer S., Dumbrell A., Emerson B., Taberlet P., West K., Narum, S. (2021) Biodiversity monitoring using environmental DNA. Editorial for *Molecular Ecology Resources* 21(5):1405-1409. doi:10.1111/1755-0998.13399.
Contribution: writing | Editorial note.
- a51. **Zinger, L.**, Lionnet C., Benoiston A.-S., Donald, J., Mercier, C., Boyer F. (2021) metabar: an R package for the post bioinformatics evaluation and improvement of metabarcoding data quality. *Methods in Ecology and Evolution* 12(4):586-592. doi: 10.1111/2041-210X.13552.
Contribution: co-supervision of the R package conceptualization, code writing, paper writing.
- a50. Marden E., Abbott R.J., Ortiz-Barrientos D., Bongaerts P., Bonin A., Bonneaud C., Buerkle C.A., Davison A., DeWoody A.J., Emerson B.C., Fountain-Jones N.M., Giraud T., Heuertz M., Hooper R., Hohenlohe P., Kinziger A.P., Moreau C.S., Nazareno A.G., Pelletier T.A., Renaut S., Riginos C., Rodríguez-Ezpeleta N., Rogers S.M., Russell J.A., Schoville S.D., Shi S., Smith M., Taberlet P., Videvall E., Waits L., Warschefsky E., Whibley A., Willoughby J., **Zinger L.**, Sibbett B., Narum S., Rieseberg L.H (2021) Sharing and Reporting Benefits from Biodiversity Research. *Molecular Ecology* 30(5):1103-1107. doi:10.1111/mec.15702.
Contribution: revision of the paper | Editorial note.
- a49. Arribas P., Andujar C., Bidartondo M., Bohmann K., Coissac E., Creer S., de Waard J., Elbrecht V., Ficetola F., Goberna M., Krehenwinkel H., Leese F., Novotny V., Ronquist F., Yu D.W., **Zinger L.**, Creedy T., Meramveliotakis E., Noguerales V., Overcast I., Morlon H., Vogler A.P., Papadopoulou A., Emerson B.C. (2021) Connecting high-throughput biodiversity inventories - opportunities for a site-based genomic framework for global integration and synthesis. *Molecular Ecology* 30(5):1120-1135. doi:10.1111/mec.15797.
Contribution: conceptualization, revision of the paper | Meeting review.
- a48. Birer C., Moreau C.S., Tysklind N., **Zinger L.***, Duplais C.* (2020) Determinants of bacterial communities and metabolites on the cuticle of two Amazonian ants species living in the same nest. *Molecular Ecology* 29(7):1372-1385. doi : 10.1111/mec.15400.
Contribution: conceptualization and co-supervision, bioinformatics analyses, writing.
- a47. Sommeria-Klein G., **Zinger L.**, Coissac C., Iribar A., Schimann H., Taberlet P., Chave J. (2020) Latent Dirichlet Allocation reveals spatial and taxonomic structure in a DNA-based census of soil biodiversity from a tropical forest. *Molecular Ecology Resources*. 20:371-386. doi: 10.1111/1755-0998.13109.
Contribution: conceptualization, field work, bioinformatics analyses, writing.
- a46. Calderón-Sanou I., Münkemüller T., Boyer F., **Zinger L.**, Thuiller W. (2020) From environmental DNA sequences to ecological conclusions: How strong is the influence of methodological choices? *Journal of Biogeography* 47:193-206. doi:10.1111/jbi.13681.
Contribution: conceptualization, supervision of bioinformatics analyses, writing.
- a45. Chalmandrier L., Pansu J., Münkemüller T., Boyer F., Génin A., Gielly L., Taberlet P., **Zinger L.**, Lavergne S., Thuiller W. (2019) Abiotic and biotic determinants of soil microbial communities β -diversity across spatial and phylogenetic scales. *Ecography* 42:2144-2156. doi:10.1111/ecog.04492.
Contribution: field work, bioinformatics analyses, writing.
- a44. Bestion E., Soriano-Redondo A., Cucherousset J., Fournet L., Jacob S., White J., **Zinger L.**, Di Gesu L., Teyssier A., Cote J. (2019) Altered trophic niche in warmer climates: an experimental appraisal.

Proceedings of the Royal Society B. 286: 20192227. doi:10.1098/rspb.2019.2227.

Contribution: supervision of molecular/bioinformatics analyses, revisions.

- a43. [Ibarbalz F.M.](#), Henry N., Brandao M.C., Martini S., Busseni G., Byrne H., Endo H., Gregory A.C., Rigonato J., Royo-Llonch M., Salazar G., Soviadan D., Zayed A., Labadie K., Ferland J., Marec C., Kandels S., Picheral M., Dimier C., Poulain J., Pisarev S., Carmichae M., Pesant S. Tara Oceans Coordinators, Babin M., Iudicone D., Jaillon O., Wincker P., Acinas S.G., Ogata H., Pelletier E., Stemmann L., Sullivan M.B., Sunagawa S., Bopp L., de Vargas C., Karp Boss L., Lombard F., Bowler C.* , **Zinger L.*** (2019) Global trends of marine plankton diversity across kingdoms of life. *Cell* 179(5):1084-1097.e21. doi:10.1016/j.cell.2019.10.008.
Contribution: conceptualization and supervision, statistical analyses, writing | Research highlights: Nature Methods | Media coverages: Phys.org - EurekAlert - Science Daily (English) - CNRS - CEA - France Inter - Radio Canada - Le Temps - GEO - France Culture - etc.. (French)
- a42. Roy J., Mazel F., Sosa-Hernandez M., Duenas J., Hempel S., **Zinger L.**, Rillig M.C. (2019) The relative importance of ecological drivers of arbuscular mycorrhizal fungal distribution varies with taxon phylogenetic resolution. *New Phytologist* 224(2):936-948. doi: 10.1111/nph.16080.
Contribution: supervision of bioinformatics/statistical analyses, writing.
- a41. Martinez-Almoyna C., Thuiller W., Chalmandrier L., Olhman M., Foulquier A., Clement J.C., **Zinger L.**, Münkemüller T. (2019) Multi-trophic β -diversity mediates the effect of environmental gradients on the turnover of multiple ecosystem functions. *Functional Ecology* 33(10):2053-2064. doi: 10.1111/1365-2435.13393.
Contribution: field work, bioinformatics analyses, writing.
- a40. **Zinger L.**, Bonin A., Alsos I.G., Bálint M., Bik H., Boyer F., Chariton A.A., Creer S., Coissac E., Deagles B.E., De Barba M., Dickie I.A., Dumbrell A., Ficetola G.F., Fierer N., Fumagalli L., Gilbert M.T.P., Jarman S., Jumpponen A., Kauserud H., Orlando L., Pansu J., Pawlowski J., Tedersoo L., Thomsen P.F., Willerslev E., Taberlet P. (2019) DNA metabarcoding - need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology* 28(8):1857-1862. doi:10.1111/mec.15060.
Contribution: conceptualization, writing | Editorial note.
- a39. Corse E., Tougaard C., Archambaud G., Bilong Bilong C.F., Duneau D., **Zinger L.**, Chappaz R., Xu C.R., Mègez E., Dubut V. (2019) One-locus-several-primers: a strategy to improve the taxonomic and haplotypic coverage in diet metabarcoding studies. *Ecology and Evolution* 9(8):4603-4620. doi:10.1002/ece3.5063.
Contribution: field work, revisions.
- a38. **Zinger L.**, Taberlet P., Schimann H., Bonin A., Boyer F., De Barba M., Gaucher P., Gielly L., Giguët-Covex C., Iribar A., Réjou-Méchain M., Rayé G., Rioux D., Schilling V., Thymen B., Viers J., Zouiten C., Thuiller W., Coissac E., Chave J. (2019) Body size determines soil community assembly in a tropical forest. *Molecular Ecology* 28:528-543. doi:10.1111/mec.14919.
Contribution: conceptualization, field work, molecular/bioinformatics/statistical analyses, writing | “From the Cover” article, image cover of the Journal’s issue | Research highlights: Molecular Ecology
- a37. Ohlmann M., Mazel F., Chalmandrier L., Bec S., Coissac E., Pansu J., Schilling V., Taberlet P., **Zinger L.**, Chave J., Thuiller W. (2018) Mapping the imprint of biotic interactions on β -diversity. *Ecology Letters* 21(11): 1660-1669. doi:10.1111/ele.13143.
Contribution: field work, bioinformatics analyses, writing.
- a36. Ser-Giacomi E., **Zinger L.***, Malviya S.* , De Vargas C., Karsenti E., Bowler C., De Monte S. (2018) Ubiquitous abundance distribution of non-dominant plankton across the world’s ocean. *Nature Ecology & Evolution* 2:1243-1249. doi:10.1038/s41559-018-0587-2.
Contribution: statistical analyses, writing | Image cover of the Journal’s issue | Research highlights: Nature Ecology & Evolution.
- a35. Jacob S., Sallé L., **Zinger L.**, Chaine A., Ducamp C., Boutault L., Russell A.F., Heeb P. (2018) Chemical regulation of body feather microbiota in a wild bird. *Molecular Ecology* 27(7):1727-1738.

- doi:10.1111/mec.14551.
Contribution: conceptualization, molecular/bioinformatics analyses, revisions.
- a34. [Birer C.](#), Tyskland N., **Zinger L.***, Duplais C.* (2017) Comparative analysis of DNA extraction methods to study the body surface microbiota of insects: a case study with ant cuticular bacteria. *Molecular Ecology Resources* 17(6):e34-e45. doi:10.1111/1755-0998.12688.
Contribution: conceptualization, co-supervision, bioinformatics analyses, writing.
- a33. Bestion E., Jacob S., **Zinger L.**, Di Gesu L., Richard M., White J., Cote J. (2017) Climate warming reduces gut microbiota diversity in a vertebrate ectotherm. *Nature Ecology & Evolution* 1(6): 0161. doi:10.1038/s41559-017-0161.
Contribution: supervision of molecular/bioinformatics analyses, writing | Research highlights: Science, Nature Climate Change | Media coverages: The Scientist - Daily Mail - International Business Times - Popular Science - Nature Ecology & Evolution blog - Phys.org - Earth.com - EurekAlert - UPI (English) - CNRS - 20 minutes - Ouest France - La Depeche - RTS 1ere (French)
- a32. Soininen E.M., **Zinger L.**, Gielly L., Yoccoz N.G., Henden J.A., Ims R.A. (2017) Not only mosses: lemming winter diets as described by DNA metabarcoding. *Polar Biology* 40(10):2097-2103. doi:10.1007/s00300-017-2114-3.
Contribution: bioinformatics analyses, revisions.
- a31. Kocher A., de Thoisy B., Catzeflis F., Huguin M., Valiere S., **Zinger L.**, Murienne J. (2017) Evaluation of short mitochondrial metabarcodes for the identification of Amazonian mammals. *Methods in Ecology and Evolution* 8(10):1276-1283. doi:10.1111/2041-210X.12729.
Contribution: supervision of bioinformatics analyses, revisions..
- a30. Kocher A., Gantier J-C., Gaborit P., **Zinger L.**, Holota H., Valière S., Dusfour I., Girod R., Banuls A-L., Murienne J. (2017) Vector soup: high-throughput identification of neotropical phlebotomine sand flies using metabarcoding. *Molecular Ecology Resources* 17(2):172-182. doi:10.1111/1755-0998.12556.
Contribution: supervision of bioinformatics analyses, revisions.
- a29. Sommeria-Klein G., **Zinger L.**, Taberlet P., Coissac E., Chave J. (2016) Inferring neutral biodiversity parameters using environmental DNA data sets. *Scientific Reports* 6:35644. doi:10.1038/srep35644.
Contribution: bioinformatics analyses, writing.
- a28. Pornon A., Escaravage N., Burrus M., Holota H., Khimoun A., Mariette J., Pellizzari C., Iribar A., Etienne R., Taberlet P., Vidal M., Winterton P., **Zinger L.**, Andalo C. (2016) Using metabarcoding to reveal and quantify plant-pollinator interactions. *Scientific Reports* 6:27282. doi:10.1038/srep27282.
Contribution: conceptualization molecular analyses, bioinformatics analyses, writing.
- a27. **Zinger L.**, Philippe H. (2016) Coalescing molecular evolution and DNA barcoding. *Molecular Ecology* 25(9):1908-1910. doi:10.1111/mec.13639. Contribution: conceptualization, writing | Invited perspective paper.
- a26. **Zinger L.**, Chave J., Coissac E., Iribar A., Louisanna E., Manzi S., Schilling V., Schimann H., Sommeria-Klein G., Taberlet P. (2016) Extracellular DNA extraction is a fast, cheap and reliable approach for high-throughput soil biodiversity assessments from environmental DNA, *Soil Biology & Biochemistry* 96:16-19. doi:10.1016/j.soilbio.2016.01.008.
Contribution: conceptualization, field work, bioinformatics/statistical analyses, writing.
- a25. Bienhold C., **Zinger L.**, Boetius A., Ramette A. (2016) Diversity and Biogeography of Bathyal and Abyssal Seafloor Bacteria, *PLoS One* 11(1):e0148016. doi:10.1371/journal.pone.0148016.
Contribution: conceptualization, statistical analyses, revisions.
- a24. Jusselme M.D., Saccone P., **Zinger L.**, Faure M., Le Roux X., Guillaumaud N., Bernard L., Clement J-C., Poly F. (2016) Variations in snow depth modify N-related soil microbial abundances and functioning during winter in subalpine grassland, *Soil Biology & Biochemistry* 92:27-37. doi:.
Contribution: molecular/statistical analyses, revisions.
- a23. Geremia R., Puşcaş M., **Zinger L.**, Bonneville J-M., Choler P. (2016) Contrasting microbial biogeographical patterns between anthropogenic subalpine grasslands and natural alpine grasslands, *New*

- Phytologist* 209(3):1196-1207. doi:10.1111/nph.13690.
Contribution: molecular/statistical analyses, revisions | Media coverage: CNRS.
- a22. Papadopoulou A., Taberlet P., **Zinger L.** (2015) Metagenome skimming for phylogenetic community ecology: a new era in biodiversity research, *Molecular Ecology* 24(14): 3515-3517. doi:10.1111/mec.13263.
Contribution: conceptualization, writing | Invited perspective article.
- a21. Pansu J., Giguët-Covex C., Ficotola G., Gielly L., Boyer F., **Zinger L.**, Fabien A., Poulenard J., Taberlet P., Choler P. (2015) Reconstructing long-term human impacts on plant communities: an ecological approach based on lake sediment DNA, *Molecular Ecology* 24(7):1485-1498. doi:10.1111/mec.13136.
Contribution: supervision of bioinformatics analyses, revisions.
- a20. Pansu J., de Danieli S., Puissant J., Gonzalez J-M., Gielly L., Cordonnier T., **Zinger L.**, Brun J-J., Choler P., Taberlet P., Cécillon L. (2015) Landscape-scale distribution patterns of earthworms inferred from soil DNA, *Soil Biology & Biochemistry* 83:100-105. doi:10.1016/j.soilbio.2015.01.004.
Contribution: supervision of bioinformatics analyses, revisions.
- a19. **Zinger L.**, Boetius A., Ramette A. (2014) Bacterial taxa-area and distance-decay relationships in marine environments, *Molecular Ecology* 23(4):954-964. doi:10.1111/mec.12640.
Contribution: conceptualization, statistical analyses, writing.
- a18. Mouhamadou B., Puissant J., Personeni E., Desclos-Theveniau M., Kastl E.M., Schloter M., **Zinger L.**, Roy J., Geremia R.A., Lavorel S. (2013) Effects of two grass species on the composition of soil fungal communities, *Biology and Fertility of Soils* 49:1131-1139. doi:10.1007/s00374-013-0810-x.
Contribution: bioinformatics analyses, revisions.
- a17. Soininen E.M., **Zinger L.**, Gielly L., Bellemain E., Bråthen K.A., Brochmann C., Epp L.S., Gussarova G., Hassel K., Henden J.A., Killengren S.T., Rämä T., Stenøien H.S., Yoccoz N.G., Ims R.A. Shedding new light on the diet of Norwegian lemmings: DNA metabarcoding of stomach content, *Polar Biology* (2013) 36(7):1069-1076. doi:10.1007/s00300-013-1328-2.
Contribution: bioinformatics analyses, writing.
- a16. Roy J., Albert C.H., Choler P., Clement J-C., Ibanez S., Lavergne S., Saccone P., **Zinger L.**, Geremia R.A. (2013) Microbes on the cliff: alpine cushion plants structure bacterial and fungal communities, *Frontiers in Microbiology* 4:00064. doi:10.3389/fmicb.2013.00064.
Contribution: conceptualization, supervision of statistical analyses, revisions.
- a15. Amend A.S., Oliver T.A., Amaral-Zettler L.A., Boetius A., Fuhrman J.A., Horner-Devine M.C., Huse S.M., Mark Welch D.B., Martiny A.C., Ramette A., **Zinger L.**, Sogin S., Martiny J.B.H. (2012) Macroecological patterns of marine bacteria on a global scale, *Journal of Biogeography* 40(4):800-811. doi:10.1111/jbi.12034.
Contribution: statistical analyses, revisions.
- a14. Bouasria A., Mustapha T., de Bello F., Geremia R., **Zinger L.**, Lemperiere G., Choler P. (2012) Changes in root-associated microbial communities are determined by species-specific plant growth responses to stress and disturbance, *European Journal of Soil Biology* 52:59-66. doi:10.1016/j.ejsobi.2012.06.003.
Contribution: molecular analyses, revisions.
- a13. Shahnavaz B., **Zinger L.**, Lavergne S., Choler P., Geremia R.A. (2012) Phylogenetic clustering reveals selective events driving the turnover of bacterial community in alpine tundra soils. *Arctic, Antarctic and Alpine Research* 44(2):232-238. doi:10.1657/1938-4246-44.2.232.
Contribution: conceptualization, molecular/statistical analyses, writing.
- a12. **Zinger L.***, Gobet A.*, Pommier T. (2012) Two decades of describing the unseen majority of aquatic microbial diversity. *Molecular Ecology* 21(8):1878-1896. doi:10.1111/j.1365-294X.2011.05362.x.
Contribution: conceptualization, statistical analyses, writing | Invited Review.

Complete list of publications and oral communications

- a11. **Zinger L.**, Amaral-Zettler L.A., Fuhrman J.A., Horner-Devine M.C., Huse S.M., Mark Welch D.B., Martiny J.B.H., Sogin M., Boetius A., Ramette A. (2011) Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PLoS One* 6(9):e24570. doi:10.1371/journal.pone.0024570.
Contribution: conceptualization, database formatting and management, statistical analyses, writing.
- a10. **Lentendu G.***, **Zinger L.***, Manel S., Coissac E., Choler P., Geremia R.A., Melo de Lima C. (2011) Assessment of soil fungal diversity in different alpine tundra habitats by means of pyrosequencing, *Fungal Diversity* 49(1):113:123. doi:10.1007/s13225-011-0101-5.
Contribution: conceptualization, field work, molecular analyses, writing.
- a9. **Zinger L.**, Lejon D.P.H., Bouasria A., Baptist F., Aubert S., Geremia R.A., Choler P. (2011) Contrasting patterns of crenarchaeal, bacterial and fungal soil communities in an alpine landscape, *PLoS One* 6(5):e19950. doi:10.1371/journal.pone.0019950.
Contribution: conceptualization, field work, molecular/statistical analyses, writing.
- a8. **Zinger L.**, Coissac E., Choler P., Geremia R.A. (2009) Microbial Communities Assessed by Graph Partitioning Approach: Study of Soil Fungi in Two Alpine Meadows. *Applied and Environmental Microbiology* 75(18):5863-5870. doi:10.1128/AEM.00748-09.
Contribution: conceptualization, field work, bioinformatics/statistical analyses, writing.
- a7. **Zinger L.**, Shahnavaz B., Baptist F., Geremia R.A., Choler P. (2009) Microbial diversity in alpine tundra soils correlates with snow cover dynamics. *ISME Journal* 3:850-859. doi:10.1038/ismej.2009.20.
Contribution: conceptualization, field work, molecular/bioinformatics/statistical analyses, writing.
- a6. Baptist F., **Zinger L.**, Clement J-C., Gallet C., Guillemin R., Martins J.M., Sage L., Shahnavaz B., Choler P., Geremia R.A. (2008) Tannin impacts on microbial diversity and the functioning of alpine soils: a multidisciplinary approach. *Environmental Microbiology* 10:799-809. doi:10.1111/j.1462-2920.2007.01504.x.
Contribution: field work, molecular/statistical analyses, writing.
- a5. Gury J., **Zinger L.**, Gielly L., Taberlet P., and Geremia R.A. (2008) Exonuclease activity of proof-reading DNA polymerases is at the origin of artifacts in molecular profiling studies. *Electrophoresis* 29:2437-2444. doi:10.1002/elps.200700667.
Contribution: statistical analyses, writing/revision.
- a4. **Zinger L.**, Gury J., Alibeu O., Rioux D., Gielly L., Sage L., Pompanon F., Geremia R.A. (2008) CE-SSCP and CE-FLA, simple and high-throughput alternatives for fungal diversity studies. *Journal of Microbiological Methods* 72:42-53. doi:10.1016/j.mimet.2007.10.005.
Contribution: conceptualization, molecular/statistical analyses, writing.
- a3. **Zinger L.**, Gury J., Giraud F., Krivobok S., Gielly L., Taberlet P., and Geremia R.A. (2007) Improvements of polymerase chain reaction and capillary electrophoresis single-strand conformation polymorphism methods in microbial ecology: Toward a high-throughput method for microbial diversity studies in soil. *Microbial Ecology* 54:203-216. doi:10.1144/SP279.15.
Contribution: study design, molecular/statistical analyses, writing.
- a2. Chalmin E., d'Orlye F., **Zinger L.**, Charlet L., Geremia R.A., Oriol G., Menu M., Baffier D., Reiche I. (2007) Biotic versus abiotic calcite formation on prehistoric cave paintings: the Arcy-sur-Cure 'Grande Grotte' (Yonne, France) case. *Geological Society, London, Special Publications* 279:185-197. doi:10.1144/SP279.15.
Contribution: molecular analyses.
- a1. Gallet C., Ibanez S., **Zinger L.**, Taravel F.R., Trierweiler M., Jeacomine I., and Despres L. (2007) Plant chemical defense induced by a seed-eating pollinator mutualist. *Journal Of Chemical Ecology* 33:2078-2089. doi:10.1007/s10886-007-9362-6.
Contribution: field work, biochemical analyses, statistical analyses.

2.2 Books, book chapters, dissertations

- b6. **Zinger L.**, [Donald J.](#), Brosse S., Iribar A., Gonzalez M.A., Leroy C., Murienne J., Orivel J., Schimann H., Taberlet P. Lopes C.M. (2020) Advances and prospects of eDNA in the Neotropical rainforests. In *Advances in Ecological Research* (Vol. 62, pp. 331-373). Academic Press. doi:10.1016/bs.aecr.2020.01.001
Contribution: conceptualization, coordination and writing | Invited review.
- b5. Taberlet P., Bonin A., **Zinger L.**, Coissac E. (2018) Environmental DNA for biodiversity research and monitoring, 1st edition. Oxford University Press, Oxford. pp. 1-268. Online INSB: 9780198767220. doi:10.1093/oso/9780198767220.001.0001.
Contribution: conceptualization, statistical analyses, writing (6 chapters), revisions (all chapters) | Research highlights: Trends in Ecology & Evolution, Molecular Ecology, The Quarterly Review of Biology, IMA Fungus.
- b4. Murienne J., Chave J., Elias M., Fontaine C., Nicolas V., Pugnolle F., Réjou-Méchain M., Taberlet P., **Zinger L.** (2014) Genèse et évaluation de la biodiversité. In Forget P.M. (Ed.) *Prospective écologie tropicale : une prospective de l'Institut Ecologie et Environnement du CNRS*, CNRS-INEE, Paris. pp. 29-44.
Contribution: analyses, writing.
- b3. Fraissinet-Tachet L., Marmeisse R., **Zinger L.**, Luis P. (2014) Metatranscriptomics of soil eukaryotic communities. In Martin F. (Ed.) *The Ecological Genomics of Fungi*, John Wiley & Sons, Chichester. pp. 305-323. doi:10.1007/978-1-4939-3369-3_16.
Contribution: writing.
- b2. Geremia R.A., **Zinger L.** (2013) Molecular fingerprinting of fungal communities in soil. In Gupta, V.K., Tuohy, M., Ayyachamy, M., Turner, K.M., O'Donovan, A (Eds.) *Laboratory protocols in fungal biology: current methods in fungal biology*, Springer, New York, NY. pp. 349-356. doi:10.1007/978-1-4614-2356-0_31.
Contribution: writing.
- b1. **Zinger L.** (2009) Spatial and temporal variations of microbial communities in alpine tundra soils. PhD thesis (french, chapters in english), Université Joseph Fourier. pp. 1-258. Available on Archive ouverte HAL.
Contribution: conceptualization, writing.

2.2.1 Manuscripts submitted / in (advanced) preparation

- S3. Lombard F., Guidi L., Brandao M. C., Coelho L. P., Dolan J. R., Gasol J., Grondin P.L., Henry N., Ibarbalz F., Jalabert L., Loreau M., Martini S., Mériquet Z., Picheral M., Pierella Karlusich J.J., Rainer P., Romagnan J.-B., **Zinger L.**, Tara Oceans Coordinators, Stemmann L., Acinas S., Karp-Boss L., Boss E., Sullivan M.B., de Vargas C., Bowler C., Karsenti E., Gorsky G. Ubiquity of inverted 'gelatinous' ecosystem pyramids in the global ocean. Encouraged to resubmit to *Science*. preprint:=> link bioRxiv.
Contribution: conceptualization and revision of the paper.
- S2. Violle C., Mahaut L., Mouillot D., Diaz S., Mouquet N., Gaüzère P., Enquist B., Maire A., Grenié M., O'Connor L.M.J., Kraft N.J.B., Auber A., Cadotte M.W., Cardou F., Delalandre L., Judah A.B., Klausmeier C.A., Litchman E., Loiseau N., Maitner B.S., McGill B., McLean M., Munoz F., Ostling A., Pimiento C., Toussaint A., Villéger S., **Zinger L.**, Thuiller W. Trait-based pathways to protect nature and its contributions to people. Submitted to *Science Advances*.
Contribution: discussions and revision of the paper.
- S1. Pierella Karlusich J.J., Cosnier K., **Zinger L.**, Henry N., Nef C., Bernard G., Scalco E., Tara Oceans Coordinators, Vieira F.R.J., Delage E., Chaffron S., Ovchinnikov S., Zingone A., Bowler C. Patterns and drivers of diatom diversity and abundance in the global ocean. Submitted to *Nature*

Communications.

Contribution: conceptualization, writing.

- P3. Martinez-Almoyna C., Saillar A., **Zinger L.**, Lionnet C., Arnoldi C., Foulquier A., Gielly L., Piton G., Münkemüller T., Thuiller W. Differential effects of soil trophic networks on microbial decomposition activity in mountain ecosystems. preprint: => link authorea.

Contribution: discussions and revision of the paper.

- P2. [Beaumelle L.](#), Cote J., Bestion E., Benoiston A.-S., Di Gesu L., Iribar A., Mathieu J., Pellerin F., Sescun U., **Zinger L.** Trophic structure influences warming effects on above-belowground multitrophic communities. In preparation for *Nature Climate Change*.

Contribution: conceptualization, supervision of molecular and bioinformatics analyses, writing.

- P1. Taberlet P., Bonin A., Boyer F., Zinger L., Coissac E. Environmental DNA for biodiversity research and monitoring, 2nd edition. Oxford University Press, Oxford.

Contribution: conceptualization, statistical analyses, writing (9 chapters), revisions (all chapters)

2.3 Oral presentations, seminars, invited lectures and posters.

- 2024 **Invited seminar**, Colloque ID-ADN 2024 - L'ADN pour identifier et suivre la biodiversité, Paris, France.

L'ADN de la pluie pour décrire la biodiversité des forêts tropicales. Scheduled for Nov 2024.

- 2023 **Invited seminar**, French Academy of Agriculture Colloquium :L'ADN (ADNe) et l'ARN (ARNe) environnementaux: marqueurs du vivant dans les écosystèmes, Paris, France.

Potentiel et limites de l'ADN environnemental et du metabarcoding ADN pour des applications en écologie.

- 2023 **Invited seminars**, Cycle of talks on eDNA, organised within the context of my CAPES/USP PrInt visiting Professor scholarship, Piracicaba, Brazil.

From molecular diversity patterns to ecological processes: how far are they in the plant-soil system?.

- 2023 **Invited seminars**, Cycle of talks on eDNA, organised within the context of my CAPES/USP PrInt visiting Professor scholarship, Piracicaba, Brazil.

The XPRIZE Rainforest challenge the eDNA solution of the Brazilian Team.

- 2023 **Invited seminars**, Cycle of talks on eDNA, organised within the context of my CAPES/USP PrInt visiting Professor scholarship, Piracicaba, Brazil.

Environmental DNA to study biodiversity in space and time.

- 2022 **Invited online lecture**, Workshop Taller de cocreación - lineamientos para la evaluación de impactos por derrame con base técnicas de genómica, Bogota, Colombia.

Designing a metabarcoding experiment: what one should consider?.

- 2021 **Invited online seminar**, ATBC 2021, Cartagena, Colombia.

the value of eDNA for tropical diversity assessments.

- 2021 **Invited lecture**, Quantitative ecological genomics in the Tara Ocean, Online QLife winter school.

eDNA and community ecology.

- 2020 **Invited seminar**, iBioGen Next-Generation Biodiversity Monitoring 2nd Symposium, Online.

Sampling soil biodiversity with eDNA.

- 2019 **Invited seminar**, iBioGen Next-Generation Biodiversity Monitoring 1st Symposium, Nicosia, Cyprus.

DNA metabarcoding and trans-kingdom biodiversity assessments.

- 2019 **Invited lectures**, the 9th DNA metabarcoding spring school, Villa de Leyva, Colombia.
DNA metabarcoding and community ecology.
- 2019 **Invited seminar**, International symposium of metbarcoding, Bogotá, Colombia.
Sequencing soils from tropical forests: what can be learned?.
- 2019 **Invited keynote seminar**, Ecology and Evolutionary Biology Symposium, Ankara, Turkey.
Microbial macroecology.
- 2018 **Keynote seminar**, the 8th DNA metabarcoding spring school. Nouragues Reserve, French Guiana.
Assessing soil tropical diversity with eDNA.
- 2018 **Invited lecture**, IFREMER Metabarcoding courses 2018, Roscoff, France.
Environmental DNA to study biodiversity in space and time.
- 2018 **Invited seminar**, Smithsonian Tropical Research Institute, Panama city, Panama.
From diversity patterns to ecosystem processes? a microbial perspective.
- 2018 **Invited seminar**, Institute of Ecology and Environmental Sciences (iEES), Paris, France.
Environmental DNA to investigate biodiversity in space and time.
- 2017 **Invited seminar**, Centre d'Ecologie et des Sciences de la Conservation (CESCO), Paris, France.
Environmental DNA to investigate biodiversity in space and time.
- 2016 **Invited lectures**, the 6th DNA metabarcoding spring school, Kunming, China.
DNA metabarcoding and community ecology.
- 2016 **Poster** ATBC 2016, Montpellier, France.
Soil community assembly varies across body sizes in a tropical forest.
- 2015 **Oral presentation** 3^{eme} Colloque de Génomique Environnementale. Montpellier, France.
Advances and Prospects of environmental DNA in French Guiana.
- 2015 **Invited keynote seminar** CEBA 4th annual meeting. Cayenne, French Guiana.
Advances and Prospects in Metabarcoding.
- 2015 **Invited lectures**, the 5th DNA metabarcoding spring school, Bialowieza, Poland.
DNA metabarcoding and community ecology.
- 2015 **Invited keynote seminar**, short DNA metabarcoding school, Cluj, Romania.
Exploring soil biodiversity via DNA metabarcoding.
- 2014 **Invited lecture**, the 3rd DNA metabarcoding spring school, Tromsø, Norway.
Exploring soil biodiversity via DNA metabarcoding.
- 2014 **Invited keynote seminar** CEBA 3rd Annual Meeting. Toulouse, France.
Soil Biodiversity: a molecular perspective.
- 2014 **Invited seminar**, Institute for Applied Ecology, Canberra, Australia.
From molecular diversity patterns to processes: how far can we conciliate community ecology and DNA metabarcoding?.
- 2013 **Invited lecture** CEBA 1st thematic school. Kourou, French Guiana.
Microbial Biogeography: Lesson learnt and new frontiers.
- 2013 **Invited lecture** the 2nd DNA metabarcoding spring school, Heraklion, Greece.
DNA metabarcoding in community ecology: lesson learnt from microbial ecology.
- 2013 **Oral presentation** 2nd Colloque de Génomique Environnementale. Rennes, France.
Mapping tropical biodiversity via DNA metabarcoding with a focus on plant and soil fungal communities.
- 2013 **Oral presentation** Microbiotoul 5th edition. Toulouse, France.
Fungal diversity and distribution in relation to plant cover in a tropical rainforest
- 2012 **Invited seminar** Laboratoire Evolution et Diversité Biologique, Toulouse, France.
Macroecological patterns of microbes in the sea and the soil.
- 2012 **Invited lecture** the 1st DNA metabarcoding spring school, Pinsot, France.
Bacterial and Fungal diversity estimation.
- 2011 **Oral presentation** Integrative ecological genomics, Jacques Monod Conferences. Roscoff, France.

Complete list of publications and oral communications

- The taxa-area and distance decay relationships of marine Bacteria: methodological aspects and ecosystem effects.*
- 2011 **Poster** 1st Colloque Génomique Environnementale, Lyon, France.
EUMETASOL: Eukaryotic Metatranscriptomes of soils in a gradient of constraints.
- 2010 **Oral presentation** ISME 13th Conferences. Seattle, USA.
Global-scale analysis of water column and sediment bacterial communities indicates contrasting beta-diversity patterns.
- 2010 **Poster** ISME 13th Conferences. Seattle, USA.
Contrasting patterns of changes of crenarchaeal, bacterial and fungal soil communities in an alpine landscape.
- 2008 **Poster** 7^{eme} Colloque National FRB, Les ressources génétiques à l'heure des génomes. Strasbourg, France.
Dynamique saisonnière des communautés microbiennes alpines: influence de l'enneigement.
- 2008 **Poster** Mountain soils under a changing climate and land-use. Birmensdorf, Switzerland.
Seasonal variations of microbial communities in two alpine soils: a molecular approach.
- 2007 **Poster** Environmental genomics: from individual genomes to genomes of complex communities, Jacques Monod Conferences. Roscoff, France.
Microbial diversity and Carbon cycle in alpine ecosystems.
- 2007 **Poster** VII^{eme} Congrès National de la SFM. Nantes, France.
Diversité de la microflore alpine: une variation spatio-temporelle.

Dissertation

List of Abbreviations

ASV	Amplicon sequence variant
CE-SSCP	Capillary-Electrophoresis Single Strand Conformation Polymorphism
COI	cytochrome c oxidase I
dNTP	deoxyribonucleotide triphosphate. N can be A (adenine), T (thymine), C (cytosine), and G (guanine).
dPCR	digital PCR.
ddPCR	droplet digital PCR.
DDR	Distance-decay of similarity
DNA	Deoxyribonucleic acid
eDNA	environmental DNA
HTS	High-Throughput Sequencing
ITS	Internal Transcribed Spacer
matK	Maturase K
MOTU	Molecular Operational Taxonomic Unit
PCR	Polymerase Chain Reaction
rbcL	ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) large subunit
RNA	Ribonucleic acid
rRNA	ribosomal RNA
TAR	Taxa-area relationship

Introduction

Microorganisms and small organisms in general, typically invertebrates (hereafter referred to as “neglected diversity”), constitute an important part of the biosphere biomass (Bar-On et al., 2018) and dominate in terms of abundance (Whitman et al., 1998) and diversity (Louca et al., 2019; Mora et al., 2011). They underpin the functioning of both marine (Azam et al., 1983; Falkowski et al., 2008; Guidi et al., 2016) and terrestrial ecosystems (Bardgett and van der Putten, 2014; Lavelle et al., 2006; van der Heijden et al., 2008). Their influence extends far beyond these environments, as they can have direct impacts on human (e.g. disease-causing pathogens, beneficial gut microbiota) and environmental health (e.g. climate change mitigation, plant pathogens/pests regulation). Until recently, studying these organisms was highly challenging from a methodological point of view, limiting our ability to anticipate how this biological compartment and its associated functions would respond to global changes. The emergence of high throughput sequencing (HTS) and the field of environmental DNA (eDNA) - which occurred more or less at the time I started my PhD in 2005 - now enables the description of different facets of the neglected diversity, i.e. taxonomic, phylogenetic, and functional, but also raises important methodological and conceptual questions.

My research activities fall within this general historical and scientific context and address questions that are at the interface between community ecology, macroecology, microbial ecology, molecular biology, and bioinformatics. I will hence briefly reintroduce some of these different fields below. Over the last decade, my main goals have been: (i) to develop methodological frameworks based on environmental genomics to (ii) better understand the different ecological processes shaping microbial and invertebrates communities and how these respond to global changes, in particular climate warming. Throughout this dissertation, I will mostly focus on free-living or host-associated microbial communities, but will at the end extend these questions to the neglected, multi-trophic diversity in general, which I started to study more recently. This dissertation will thus discuss two main axes of my research: (i) the development of an eDNA-based toolbox for biodiversity research at molecular, analytical, and conceptual levels (section 5), and (ii) investigating the patterns and processes of free-living or host-associated microbiota, and multi-trophic communities (section 6). Throughout this dissertation, I will present results obtained in different biological systems, ranging from soils, marine waters and sediments, to various macro-organisms hosts, i.e. lizards, birds, arthropods, or trees. Following this synthesis, I will propose several perspective of research (section 7) with the following broad objectives: (i) to advance the eDNA-based toolbox for a more holistic, robust, and quantitative assessment of multi-trophic communities, to (ii) investigate how tropical forests above-belowground communities and their associated functions respond to different climatic/biogeographic contexts, land use types and restoration strategies.

4.1 A brief refresher of community ecology

One central and long-standing quest in Ecology is to identify and quantify the processes that shape ecological assemblages in space and time. About a century of documenting patterns of α , β , and γ diversity (Rosenzweig, 1995, see definitions in Figure 4.1) has led to the emergence of a plethora of concepts and theoretical frameworks that were, until recently, considered a “mess” (Lawton, 1999). This was not only due to the intrinsic complexity of the many rules governing ecological assemblages. The mess also resulted from the in-

dependent emergence of many, often overlapping, concepts and theories in different fields of Ecology. Various attempts of synthesis of this vast body of literature have been made to resolve this complexity (e.g. Leibold et al., 2004; Ricklefs, 1987). Mark Vellend’s conceptual synthesis (Vellend, 2010) made considerable advances by reducing all previously identified processes to make them analogous to the “big four” in population genetics (i.e. *selection*, *drift*, *mutation*, and *gene flow*) at the level of biological organisation of communities. More specifically, Vellend proposes that a local assemblage of species is a subset of the regional species pool influenced by four, non-exclusive, fundamental processes: *selection*, *drift*, *speciation* and *dispersal* (Figure 4.1).

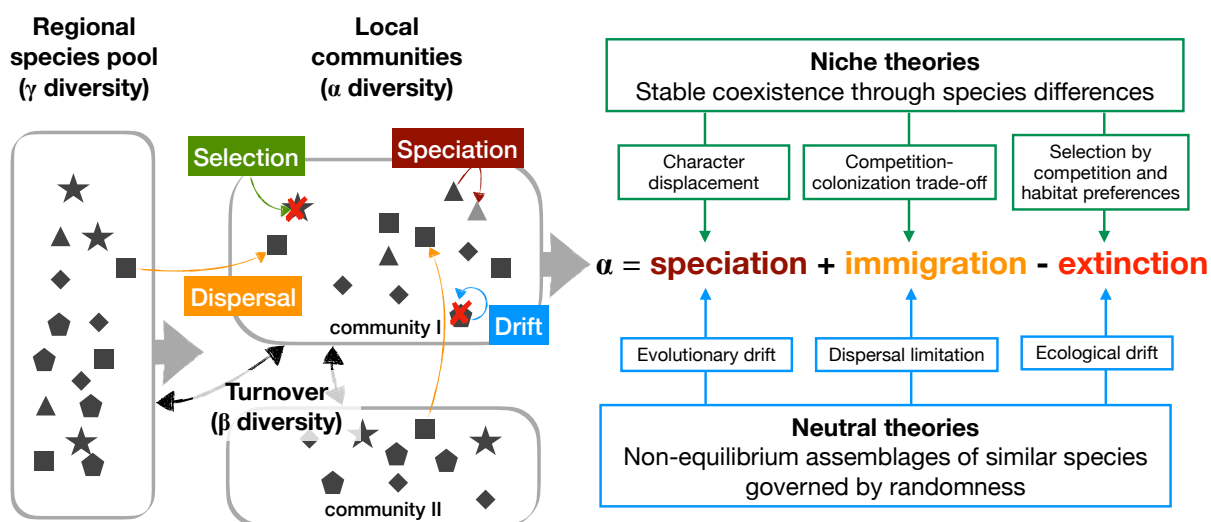


Figure 4.1: Overview of the processes shaping biological assemblages and associated main theories. Biological assemblages can be partitioned into different diversity components, i.e. α and γ diversities, which refer to local vs. regional diversities respectively while β diversity refers to the differences between local communities or between local communities and the regional species pool (Whittaker, 1972). Each diversity component depends on the species present locally and their abundances, which are influenced by four fundamental factors according to Vellend (2010), depicted here in full coloured boxes. The right side of the figure gives a simplified overview of the main families of theories explaining each process, whether they are related to the niche (green) or to neutrality (blue) concepts.

Although often associated with evolutionary dynamics, *selection* here relates to the concept of the niche (Chase and Leibold, 2003; Hutchinson, 1957), which originally designated the space different species occupy in their environment through their various relationships to resources, predators, and habitat. This concept was already implicit in Charles Darwin’s “struggle for existence” (Darwin, 1859) and has largely evolved since then (Chase and Leibold, 2003; Pocheville, 2015). *Selection* hence refers to the process by which a set of environmental factors such as abiotic conditions (e.g. temperature, nutrient availability) or biotic interactions (e.g. predation, symbiosis, or competition) influence the population growth rate of a given species and can ultimately lead to local extinction (i.e. extirpation). Therefore, these factors *select* a set of species able to thrive and co-exist locally, either in a constant, density-dependant, or spatiotemporally variable way (Chesson, 2000; Vellend, 2010), while the other species present in the system go extinct locally through abiotic or biotic filtering.

This deterministic view is often opposed to the idea that community dynamics can result from neutral processes, i.e. can occur irrespective of species ecological differences. Neutrality is a more recent concept in the history of community ecology and has been popularized by Stephen P. Hubbell’s unified neutral theory of biodiversity and biogeography (Hubbell, 2001). This theory builds on other neutral models (reviewed in Chave, 2004; Pocheville, 2015), including Robert MacArthur & Edward O. Wilson’s famous theory of island biogeography (MacArthur and Wilson, 1967). In a neutral community, species populations undergo birth, reproduction, or death that all occur at random irrespective of species identity. Under this assumption of

ecological equivalence, these three demographic processes randomly sample individuals. In a community with a finite number of individuals, this re-sampling causes fluctuations in species abundances through time where most species ultimately *drift* to extinction. Hubbell’s neutral theory has been proven successful in making realistic predictions on species diversity or abundance distributions and is often used as a null hypothesis to test the niche theory (reviewed in Chave, 2004).

Both *drift* and *selection* lead to similar outcomes, the dominance or extinction of species. Hence, these two processes fail to explain why many species can co-exist locally. *Dispersal*, on the other hand, can contribute to species coexistence. Dispersal corresponds to the movement of organisms across space from source to recipient communities, subsequently replacing local communities in their regional context (Ricklefs, 1987). *Dispersal* can counteract local extinctions caused by *selection* or *drift* (reviewed in Chase and Myers, 2011; Vellend, 2010). For example, from a niche perspective, a species that would have been extirpated through competition can still co-exist regionally with its competitor if it has a greater ability to disperse in sites where its competitor is absent (i.e. Tilman, 1994). Likewise, species with high dispersal rates have a higher probability to colonize and persist in areas that are yet less favourable (often referred to as “mass effect”). From a neutral perspective, the simplest version of the theory of island biogeography explains the species richness of an island or site as an equilibrium between colonization and random extinction processes (MacArthur and Wilson, 1967). In both perspectives, local extinctions increase with increasing *dispersal* limitation.

Finally, *speciation*, i.e. the appearance of a new species, also contributes to the build-up of local communities. Although more related to fields like biogeography or macroevolution, *speciation* must be considered when explaining local community composition in a larger-scale context, e.g. when comparing different geographic regions or steep environmental gradients. As for *dispersal*, *speciation* can result from either niche-related or neutral dynamics. From a neutral perspective, random mutations in two sub-populations can lead to the appearance of new species when the gene flow is disrupted (e.g. mountain range, distances), a process also known as “evolutionary drift” or “allopatric speciation”. *Speciation* can also result from niche-related processes, which brings us back to Darwin’s “struggle for existence” and his “principle of divergence” (Darwin, 1859): slight niche differences within a species should be gradually exacerbated through time to reduce competitive pressure up to the creation of new species with different niche requirements (discussed in Mayr, 1992). In both cases, *speciation* increases local species diversity. This process remains, however, seldom considered in community ecology because the focus is on the local species pools, where the speciation rate is expected to be negligible. As this process falls outside the direct scope of my investigations, I will limit its discussion in my work synthesis and elaborate a bit more in the research perspectives section.

Whether communities are governed by neutral vs. niche-based processes has been so far assessed mainly using either inductive or deductive approaches, or both, but making such inferences remains to date challenging (reviewed in Chase and Myers, 2011; Vellend et al., 2014). Inductive approaches in community ecology primarily involve statistical modelling and consist in explaining community diversity or composition with descriptors either related to *selection*, i.e. abiotic/biotic conditions or reflecting neutral/random processes, e.g. geographic proximity as a proxy of dispersal limitation and hence, of *drift*, area of the patch as proxy of sampling effects, etc. (Cottenie, 2005; Legendre, 2008; Tuomisto et al., 2003). The advantage of this approach is that the relative contribution of each process can be somewhat estimated. Its main limit - amongst others - is that unexplained variance can always be due to unmeasured parameters (Gilbert and Bennett, 2010; Smith and Lundholm, 2010; Vellend et al., 2014). By contrast, deductive approaches aim at testing a theory, here that the community results from niche-based, deterministic processes. Falsifying this theory is done by comparing certain features from observed communities (e.g. species abundance distributions, co-occurrence, or β diversity patterns) against modelled features of stochastic or neutral communities. Two types of models are used in this respect: null models, which are build by randomising certain properties of the community data (Gotelli and Graves, 1996), and neutral models such as Hubbell’s neutral theory (Hubbell, 2001).

4.2 A short history of microbial biogeography

Community ecology and biogeography applied to the microscopic world (i.e. viruses, archaea, bacteria, protists, and fungi) have gained momentum in the two last decades owing to several methodological break-

throughs that I will describe below (see section 4.3). These advances have not only enabled descriptions of patterns of microbial diversity at unprecedented scales; they have also revealed how important microbes are for nutrient cycling and ecosystem functioning (Falkowski et al., 2008; van der Heijden et al., 2008), as well as the nutrition and health of larger organisms (Hacquard et al., 2015; Mueller and Sachs, 2015), with the latter topic having been largely developed and popularized after the launch Human Microbiome Project (The Human Microbiome Project Consortium, 2012). However, initial attempts to understand the factors shaping the composition and distribution of microbial assemblages date back to the very beginning of bacteriology in the 19th century, and the concepts and ideas developed at that time were surprisingly modern.

Naturalist and microscopist Christian Gottfried Ehrenberg, a contemporary of Darwin, pioneered microbial biogeography by studying the geographical distribution of *Tetracyclus* Ralfs, a genus of diatoms (Figure 4.2a). Based on the morphology of strains present in samples coming from various parts of the world, his observations led him to suggest that members of this group were less divergent between America and China than between North and Southern America. He explained this particular distribution by the presence of mountain ranges, which may constitute greater barriers to the dispersal of these organisms than oceans (reviewed in Williams and Huxley, 1998). His “infusion” experiments from environmental samples further suggested that the community composition of microbial cultures was a product of chance. These findings and ideas strikingly echo the modern neutral theories and evolutionary *drift* (reviewed in De Wit and Bouvier, 2006; O’Malley, 2007).

On the other hand, Darwin’s theory of evolution, which also relied on observations of microbial life forms, was received with indifference or even scepticism by the microbiological community, which is probably why so few exchanges occurred between microbiology and ecology until recently. In the late 19th century, the field was indeed more committed to experimental approaches than empirical, *in situ* ones. In addition, Darwin’s theory of evolution was - incorrectly - perceived as resulting from spontaneous generation, which had just been falsified by Louis Pasteur (reviewed in O’Malley, 2009). It is only in the early 20th century that niche-based, deterministic explanations for microbial community assembly re-emerged, with the experiments of Martinus Willem Beijerinck (Beijerinck, 1913) and later of Lourens Baas Becking (Baas-Becking, 1934). Their culture enrichment experiments produced similar sets of species when providing the same environmental conditions, leading to Baas Becking’s famous principle of “*everything is everywhere, but, the environment selects*”, pre-dating the later emergence of niche theory, where *selection* determines microbial assemblages (De Wit and Bouvier, 2006; O’Malley, 2007).

One common point in the early neutral or niche-based views of microbial assemblages is the fundamental role given to *dispersal*. Ehrenberg’s explained differences between communities to be directly linked to their connectivity, thus implicitly invoking dispersal limitation. Likewise, the first part of Baas Becking’s principle “*everything is everywhere*” makes the explicit hypothesis of a microbial cosmopolitanism - implicitly through unlimited dispersal. The assumption of large, if not unlimited dispersal in microbiology idea finds its roots in Augustin P. de Candolle’s work, for whom dispersal was an explanation for the absence of endemism in “low organisation” organisms, typically fungi (reviewed in O’Malley, 2007). The idea that dispersal is allometric was later further developed to explain why the distribution of microbial species differs so much from that of macroorganisms (Finlay, 2002). Indeed, microorganisms have large population sizes, fast population growth rates, and shorter generation times. These properties not only contribute to higher dispersal but also make microbes more likely to successfully colonize new locations (Figure 4.2b, reviewed in Barberán et al., 2015). The expected high dispersal capacity of microorganisms should buffer local extinctions caused by either *drift* or *selection*, reduce speciation rates, and ultimately lead to the absence of biogeographical patterns for microbes (Finlay, 2002).

The advent of novel molecular techniques stimulated a renewed interest in the search for microbial biogeographical patterns in the mid-2000’s - more or less when I started my PhD - by enlarging the scale at which microbial diversity was studied, not only spatially or temporally, but also by giving access to unculturable organisms. These are usually thought to represent > 99% of the whole microbial diversity (a phenomenon often referred to as “the Great Plate Count Anomaly”, Staley and Konopka, 1985, but see ¹). Emerging stud-

¹A not so recent study (Tanaka et al., 2014) has shown that this anomaly is - at least partly - due to inappropriate protocols for culture media preparation: autoclaving phosphate together with agar to prepare solid growth media seem to generate bacterial growth inhibitory compounds.

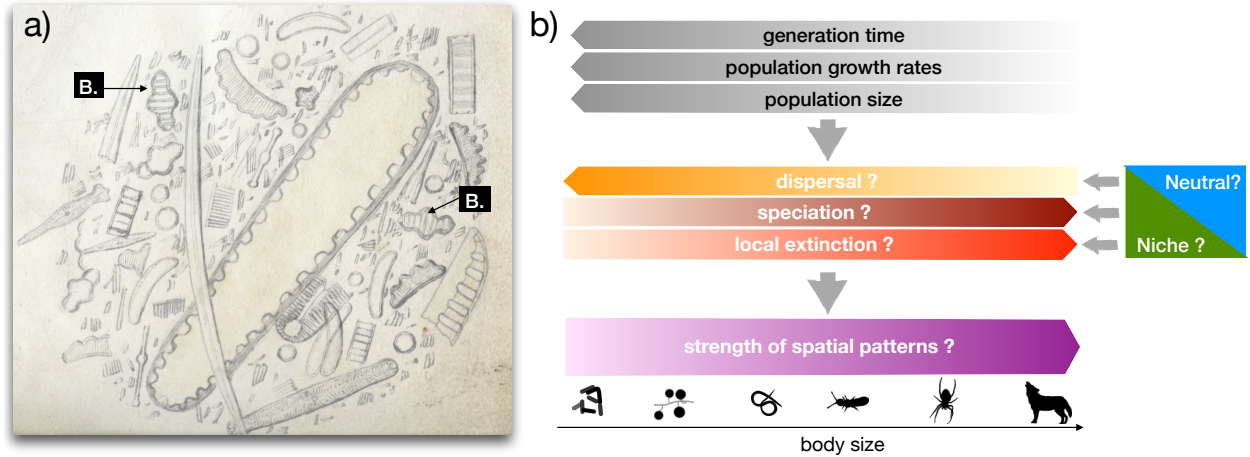


Figure 4.2: Microbial community assembly (a) An aquatic microbial community. Detail of a drawing of naturalist Ehrenberg, depicting various diatom species including members of the *Biblarium* genus (now *Tetracyclus* Ralfs), denoted by “B.”, of which Ehrenberg studied the geographic distribution. Image from the Ehrenberg Collection Database freely available on the Natural History Museum of Berlin website. (b) Hypothetical relationships between body size, intrinsic properties of organisms, and community assembly processes. The darker the colour of the boxes, the more important the property/process. The relationships are not necessarily linear.

ies relying on molecular approaches seriously questioned the hypothesis of a microbial cosmopolitanism by revealing macroecological patterns in microorganisms that were hitherto challenging to observe. For example, a study revealed the existence of a species-area relationship for bacteria (Bell et al., 2005; Horner-Devine et al., 2004), i.e. the fact that the number of species (or rather taxa or MOTUs when talking about microbes²) increases when increasing the area of observation. The same year, another study showed that soil ascomycete fungi do exhibit a distance-decay of similarity (Green et al., 2004), i.e. the fact that communities tend to be more different with increasing geographic distances. On the other hand, these studies, amongst others, all reported that microorganisms do indeed exhibit weaker spatial patterns as compared to their “macro-bial” counterpart (reviewed Prosser et al., 2007; Soininen, 2011) as well as existence of several cosmopolitan MOTUs (Pommier et al., 2005), therefore supporting the idea of higher dispersal rates in microbes.

While the hypothesis of an absence of biogeographical pattern in microbes has been quickly abandoned following the above-mentioned observations (Martiny et al., 2006), the microbial cosmopolitanism hypothesis (or that of the microbial endemism) remains to date not testable in practice. Indeed, appropriately sampling microbial diversity remains a challenging endeavour even with HTS-based molecular methods, as they are subjected to several biases (section 5) and often lack taxonomic resolution (discussed in section 5.5 and in Achtman and Wagner, 2008; Hanson et al., 2012; Rosselló-Mora and Amann, 2001; Stackebrandt and Goebel, 1994). Several cases of microbial endemism have been reported and attributed to dispersal limitation and allopatric speciation (Pommier et al., 2005; Whitaker et al., 2003), but again, the sampling coverage of these studies is largely insufficient to validate endemism. Whether microbes are cosmopolitan or not, the question of whether community assembly processes are rather niche-based or neutral (Figure 4.1) remains up to now a central question in microbial biogeography, as it is the case for macroorganisms. I will develop this in section 6 of this dissertation.

²I won’t cover the matter of the species concept for microbes here, for which I refer the reader to other readings (e.g. Achtman and Wagner, 2008; Rosselló-Mora and Amann, 2001). The MOTU concept, widely used as a proxy of species in microbial ecology for convenience, classifies microorganisms into operational taxonomic groups exhibiting a certain level of genetic similarity, in general on variable regions the 16S rRNA gene (see section 4.3). The minimum similarity required to belong to the same bacterial species is usually 97% of similarity on the full 16S rRNA gene, which is the similarity threshold below which two 16S rRNA gene sequences always belong to two different bacterial species as defined by DNA-DNA hybridization (Stackebrandt and Goebel, 1994).

4.3 1990-2010's: a series of molecular revolutions

The last three decades have seen the emergence of several major technological breakthroughs that considerably changed the way we now inventory biodiversity (see Figure 4.3 for an overview). Microbiologists have pioneered such innovations in several ways. One of the most important contributions is perhaps the identification of genomic regions bearing a taxonomic or phylogenetic signal, typically the 16S rRNA gene (Fox et al., 1977; Woese and Fox, 1977). This genomic region has rapidly become the gold-standard region for classifying members of the Archaea and Bacteria domains. It was not until the mid 2000s that such standard genomic regions, now referred to as “barcodes” (Hebert et al., 2003), were defined for macro-organisms; namely the 18S rRNA gene for protists (Pawlowski et al., 2012), the ITS³, located in the ribosomal genes cluster, for fungi (Schoch et al., 2012), the mitochondrial *COI*⁴ for metazoa (Hebert et al., 2003), or several chloroplastic genes (e.g. *rbcL*⁵ and *matK*⁶) for plants (CBOL Plant Working Group et al., 2009). Barcoding regions typically exhibit a nucleotidic sequence that is variable enough amongst species/taxa to differentiate them, but are flanked by sequences that are conserved enough across a given lineage to enable the DNA region amplification or capture with specific primers (Figure 4.3). This DNA-based identification system is expected to help overcome the Linnean shortfall (i.e. the lack of formal description for most species). Indeed, DNA provides hundreds of objective classification criteria (i.e. suites of nucleotides) compared to the tens morphological or behavioural, sometimes subjective, criteria used with more traditional approaches. But we’ll see later that barcode-based approaches also come with their load of limitations (section 5).

Other major developments enabled microbiologists, and later on plant or animal ecologists, to use eDNA to characterise biodiversity. For example, the first procedures enabling the isolation of DNA molecules from environmental matrices such as sediments (Ogram et al., 1987) or soils (Zhou et al., 1996) were developed to study microbial diversity, thereby leading the emergence of what is now referred to as “environmental DNA” based techniques (Pawlowski et al., 2020; Taberlet, Coissac, Hajibabaei, and Rieseberg, 2012). Note that here and throughout this dissertation, I will use the term “eDNA” *sensu* Pawlowski et al. (2020), i.e. to designate any DNA present in environmental samples, whether it belongs to micro- or macroorganisms and whether it is intra- or extracellular. Likewise, the invention of PCR amplification (Saiki et al., 1988) has been rapidly used to amplify barcode genes from environmental samples to describe seawater (Giovannoni et al., 1990) or soil microbial communities (Torsvik et al., 1990). The first transposition of this approach to macroorganisms emerged only two decades later, to detect frog presence from water samples (Ficetola et al., 2008). This approach, now referred to as DNA metabarcoding (Taberlet, Coissac, Pompanon, et al., 2012), typically consists in isolating gold-standard barcodes from a metagenome through PCR amplification or hybridization capture enrichment, and sorting/identifying these barcodes with fingerprinting techniques, cloning and/or Sanger/HTS sequencing (reviewed in Metzker, 2009), or microarray methods (reviewed in Kirk et al., 2004, Figure 4.3). I will cover DNA metabarcoding more extensively in section 5. In parallel, direct sequencing of pieces of eDNA extracted from soils or seawater has also been rapidly enabled describing not only the taxonomic diversity of microbial communities, but also their metabolic functions (Handelsman et al., 1998; Venter et al., 2004). In both eDNA metabarcoding and metagenomics, the sequencing step acts like a sampling process of the DNA molecules contained in a given environmental sample. Increasing the number of sequences (i.e. sequencing depth) to describe that sample inherently improves the coverage of the diversity of amplicons or genomic fragments present in the sample.

³Internal Transcribed Spacer.

⁴cytochrome c oxidase I gene.

⁵ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) large subunit.

⁶Maturase K.

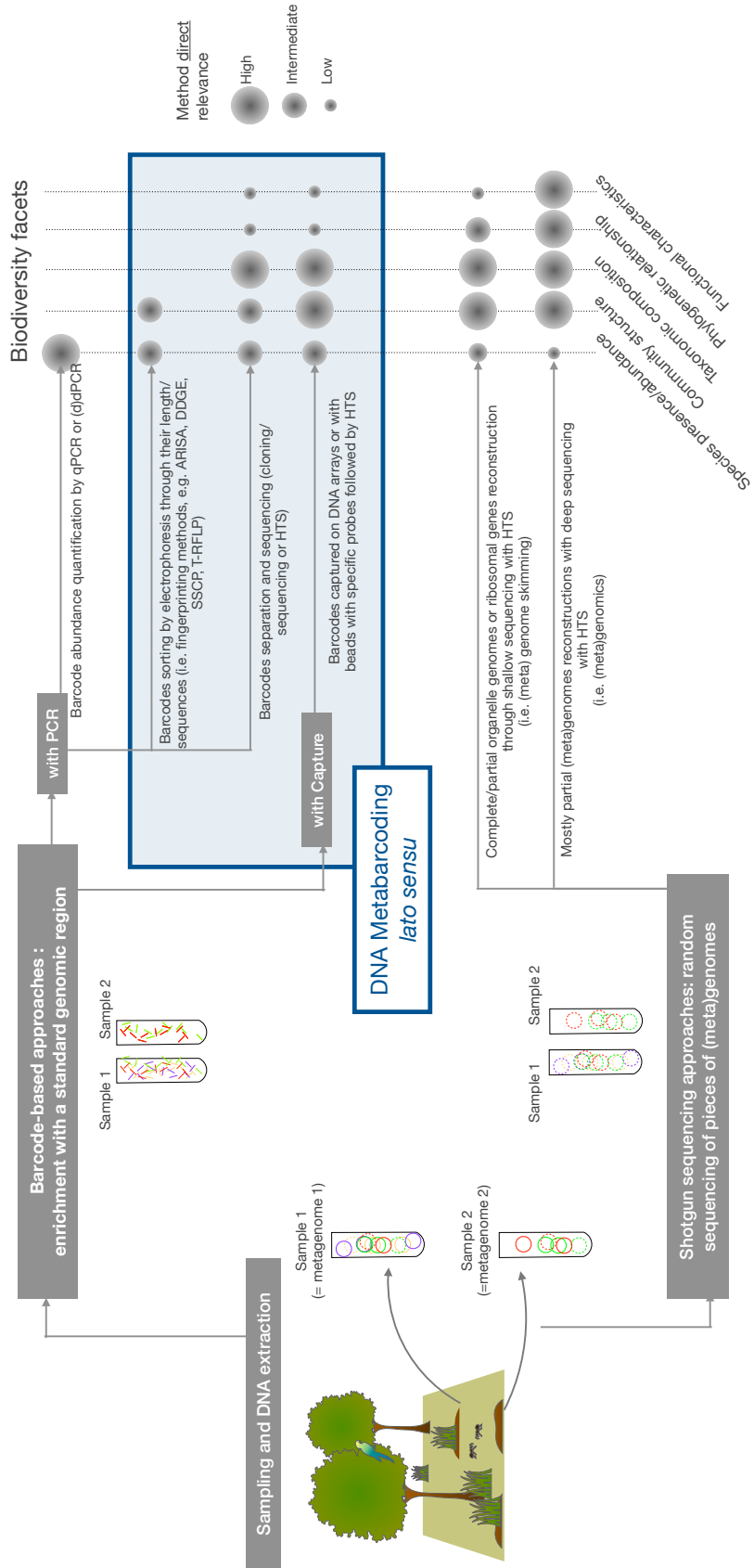


Figure 4-3: Overview of the main molecular methods for studying biodiversity. The right side of the figure classifies the different molecular methods, while their relevance for exploring each biodiversity facet is depicted by the dot sizes in the dot plot on the left (from no circle: inappropriate; to large circle: highly appropriate). All methods require a first step of sampling and DNA extraction. The resulting DNA extract, which contains a metagenome, is then processed with two broad families of approaches, either by amplifying by PCR/capturing a barcode, or through shotgun sequencing. DNA metabarcoding techniques - included in the blue box - correspond to the former category, and basically consist in classifying the amplified/enriched barcodes based on their properties (e.g. length/3-D conformation when using fingerprinting methods, and/or nucleotide sequence when using sequencing). Although outdated, I have included fingerprinting methods here because a non-negligible part of my thesis relied on one of them, the Capillary-Electrophoresis Single Strand Conformation Polymorphism (CE-SSCP). The second category - shotgun sequencing - is here subdivided into two approaches depending on the sequencing effort. Metagenome skimming (Papadopoulou et al., 2015) is based on a shallow sequencing of the metagenome. It thus samples the most repeated regions of the genome, typically organelle genomes and the ribosomal genes cluster, that are phylogenetically informative. Genes involved in less generic functions to cell functioning are in comparison much less repeated, and detecting them hence requires deeper sequencing effort.

All these approaches differ in their ability to describe the different facets of biodiversity, i.e. taxonomic, phylogenetic, and functional diversity (Figure 4.3). The more DNA regions per genome sequenced, as it is the case for metagenomics, the better one will describe the different facets of biodiversity. But it comes at a price: sequencing more regions per genome inherently reduces one's capacity to describe all genomes in a given sample. By contrast, amplifying/capturing barcodes fails to provide information on the functions or metabolic capacities of the species present in the sampled environment unless one has *a priori* knowledge on these taxa (e.g. using species functional databases). In any case, focusing the sequencing on the information solely bearing a taxonomic signal inherently increases our capacity to describe the whole community. On the other hand, the phylogenetic information yielded by one barcoding region is often poorly informative as compared to the one retrieved with metagenome skimming. The latter technique indeed enables the parallel sequencing of multiple barcoding loci, hence providing a better description of the evolutionary history of the species present in the environmental sample. All these methods - and especially DNA metabarcoding - are now being increasingly used in research and beyond with the ever-increasing throughput and ever-decreasing costs of HTS technologies. However, working with eDNA comes also with several methodological and conceptual challenges, which I will now discuss in section 5.

A DNA-based toolbox for biodiversity research

5.1 General context

The last decade has witnessed a revolution in the way we describe biodiversity owing to the rise of eDNA and HTS (see Figure 4.3 for an overview of available methods). This has rendered possible the study a full range of taxa and habitats, from soil microbes (e.g. Thompson et al., 2017) to large marine fish (e.g. Boussarie et al., 2018), and from contemporary to tens of thousands-year-old biological communities (e.g. Willerslev et al., 2014). The breadth of potential applications of eDNA is immense. It encompasses the study of diversity patterns of full biological communities, including the neglected biodiversity component, of species trophic (e.g. De Barba et al., 2014) or mutualistic interactions (e.g. Pornon et al., 2016), as well as more routine biomonitoring applications to assess ecosystem health (Pawłowski et al., 2018). DNA metabarcoding, in particular, is a tool of choice for such analyses because of its cost-effectiveness and - at first sight - easiness of implementation (see a detailed explanation of the DNA metabarcoding workflow in Figure 5.1). At a time when the monitoring of biodiversity and ecosystems becomes increasingly critical, the approach hence stands out as a powerful and promising tool.

However, eDNA-based techniques also come with several hurdles at the molecular, bioinformatics, analytic, and conceptual levels that I will discuss in this section. This discussion will generally be illustrated with applications for the study of the neglected biodiversity in the soil environment. Though I wouldn't consider myself an expert in each of the disciplinary fields involved eDNA-based approaches, I have examined closely several of the steps of the eDNA metabarcoding workflow. This was mainly motivated by the necessity for me to really understand the characteristics of the data I handle, but also because I had sometimes to adapt the technique to certain type of questions/systems. I have also been fortunate to start working with eDNA at a pivotal time - the field was exploding with innovations - and with exceptional mentorships and collaborative research environments. This has allowed me to learn about and contribute to various aspects of the eDNA toolbox during my scientific career and gave me, I believe, a relatively good vision of the promises and limitations of the whole eDNA metabarcoding workflow.

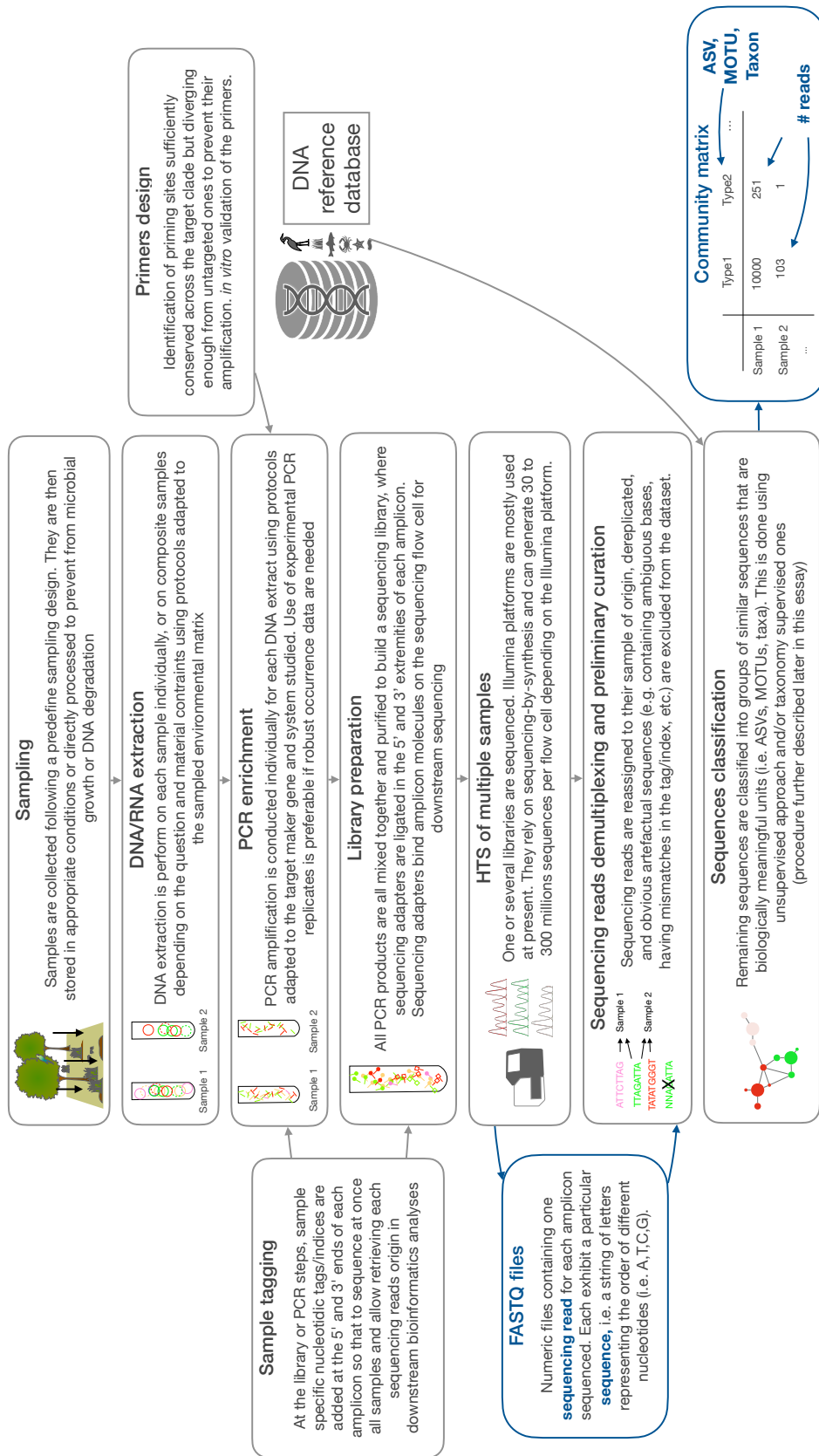


Figure 5.1: A typical DNA metabarcoding workflow. In blue boxes are indicated the output of a particular step as well as some definitions about their content. See section 5.3 for more details on the sequence classification step.

5.2 Molecular biases in eDNA metabarcoding

Every technological innovation, while offering new possibilities, also comes with its own set of new biases. This phenomenon is well-illustrated by the eDNA field: it heavily relies on molecular biology, where techniques are both in constant evolution and subjected to a variety of biases (see Table 5.1). In the 2000s (c.a. ten years after the first DNA metabarcoding study), environmental microbiologists already identified a range of biases inherent to DNA extraction and PCR, typically taxon-dependant sampling biases or generation of artefactual DNA molecules. These biases were shown to collectively distort taxa abundances and overestimate bacterial diversity. The use of HTS further amplified this issue by making artefacts more detectable through an increased sampling effort of PCR products, but also by introducing new sets of biases that are HTS-platform dependent, hence requiring continuous vigilance and adaptation¹. Depending on the DNA metabarcoding experimental design and target system, the amount of artefactual signal in raw DNA metabarcoding datasets ranges between ~10 to 30% of the total number of *sequencing reads*, and between ~30 to 95% of the total number of *sequences* (Taberlet et al., 2018, see Figure 5.1 for the difference between the two terms, and Figure 5.3 for an example of the extent of the problem). Such an amount of experimental noise emphasizes the critical need to develop effective protocols at different steps of the DNA metabarcoding workflow to ensure data and associated inferences reliability.

Bias/error	Definition	References
Biased DNA extractions	Certain microbial clades have cell walls that are recalcitrant to specific DNA extraction protocols, hence resulting in strong discrepancies in community composition depending on what extraction protocol is used.	Martin-Laurent et al., 2001
Preferential PCR amplification	Certain clades can be preferentially amplified over others, either because the priming sites of certain clades exhibit mismatches with the primers used, or because of the presence of secondary structures outside the region targeted that renders the primer site inaccessible for elongation.	Hansen et al., 1998; Polz and Cavanaugh, 1998.
PCR errors	Artefactual DNA fragments can be generated during the PCR, either because DNA polymerase errors can occur at the elongation step, with a frequency that depends on the type of DNA polymerase used, or by the annealing and elongation of a partial or full-length sequence with another one, resulting in chimeric or heteroduplex DNA molecules.	Acinas et al., 2005; Koczynski et al., 1994; Thompson et al., 2002
Sequencing errors	Artefactual sequences can be created when the sequencer misreads the nucleotide emitting the signal, typically in homopolymer regions, or when the wrong nucleotide is incorporated in the case of sequencing by synthesis. This error rate depends on the sequencing platform.	Kunin et al., 2010; Quince et al., 2009
Preferential sequencing	Sequences with strongly unbalanced GC content tend to be under amplified during the library preparation step and hence less sequenced	Aird et al., 2011; Dohm et al., 2008
Tag/index jumps	Several molecular processes at the library preparation step generate chimeric fragments at the priming site of the amplicons. The products of such chimeric fragments exhibit erroneous tags/indices combinations while conserving a valid barcode sequence. Such sequences can hence be assigned to a wrong sample, resulting in an apparent sample cross-contamination	Carlsen et al., 2012; Schnell et al., 2015

Table 5.1: Overview of the molecular biases impacting biodiversity observation in eDNA data. This list is not exhaustive and does not consider contaminations.

¹One will remember the initial excitement surrounding the vastness of the deep sea rare bacterial biosphere revealed by the pyrosequencing technology (Sogin et al., 2006), which ultimately turned out to be largely composed of molecular artefacts (Kunin et al., 2010)

Part of my research activities has hence so far consisted in optimizing/validating molecular biology protocols to maximize the resolution and reproducibility of diversity patterns retrieved from eDNA metabarcoding following the most recent technical advances. For example, while I have worked on the optimization of the CE-SSCP fingerprinting technique to study fungal communities (Zinger et al., 2008), its limitations, together with the promises of HTS led me to rapidly switch to HTS during my thesis. In addition, the molecular protocols also need to be tailored to the system studied. For example, for [Caroline Birer](#)'s PhD work, which I co-supervised, we wanted to study specifically the ant's cuticular microbiota, which is expected to be involved in host defence against pathogens (Boucias et al., 2018). But targeting specifically ant cuticular microbiota without detecting the gut one is quite a challenge: it requires desorbing minute amounts of bacterial cells from the ant cuticle while ensuring ant individual integrity and avoiding faecal evacuation. We hence developed a DNA extraction procedure able to handle such sensitive samples (Birer et al., 2017). To keep things interesting, I will only dive into some examples of developments I've worked on for the DNA extraction and PCR amplification steps.

5.2.1 eDNA extraction

The topic of DNA extraction methods is a big one, especially when it comes to the soil matrix. The soil matrix is indeed mainly constituted of humic substances, which are complex chemically and contain notorious inhibitors of the PCR reaction (Rådström et al., 2004). Soils also exhibit a large range of chemical and texture properties, which will have an impact on DNA extraction success (Philippot et al., 2012). Like many environmental microbiologists, I have also tested the performance of different DNA extraction protocols and commercially available kits designed for soil samples, to make more robust the study of soil bacterial communities during my PhD (Zinger et al., 2007). However, these protocols, particularly commercial kits, typically require a starting material of only 250 mg, which is insufficient to appropriately sample microbial communities (Ranjard et al., 2003).

Accurately capturing soil biodiversity, i.e. including also soil fauna, necessitates processing samples several orders of magnitude larger (c.a. hundreds of grams, Andersen et al., 2012; Arribas et al., 2016). When I began working on soil multitrophic communities, Taberlet et al. (2012) had already proposed an approach compatible with large soil volumes and targeting extracellular DNA. This method is also suitable for large-scale applications because it is relatively simple and cost-effective to implement. Briefly, large amounts of soil are vigorously mixed with a phosphate-saturated buffer. The phosphate competes with- and desorbs extracellular DNA molecules, also containing phosphates, that are bound to organo-mineral particles. A portion of the resulting homogenised suspension is then processed with a commercial kit. While targeting explicitly extracellular DNA is necessary when studying macroorganisms, as they are in most cases not present in the environmental sample, this approach has received criticisms when applied to microbial communities. Extracellular DNA has been indeed suggested to distort estimates of contemporaneous soil microbial diversity, as it would correspond to past (or "relic") communities (Carini et al., 2016, but see section 5.4 for a discussion on this assumption).

We've thus assessed how this approach compared with a commercial kit extracting total DNA from large volumes of soil (ten grams). We did so by comparing the diversity trends obtained for bacteria, whole eukaryotes, and plant communities (Zinger et al., 2016). We found very similar diversity patterns across approaches, although differences existed in the relative abundances of specific clades. Note though that such differences in clade abundance are always observed when comparing any DNA extraction protocol (e.g. Tedersoo et al., 2010; Terrat et al., 2012). Soil communities obtained with Taberlet et al. (2012) protocol also tended to yield more rare MOTUs, suggesting potential undersampling issues. Yet, it remains a very good trade-off between cost, labour, and accuracy for high-throughput DNA metabarcoding studies of soil biodiversity. Ultimately, this approach allowed us (and is still allowing us) to process hundreds of kilograms of soil samples across study systems and projects (some examples of study using this protocols will be presented in section 6.5).

5.2.2 PCR amplification

The question of the type of DNA polymerase used for PCR amplification, albeit seldom assessed, seems to have been of lesser interest to the scientific community, to the point that the enzyme’s type and its identity are rarely explicitly reported in research articles. However, it has long been known that DNA polymerases have an error rate that is strongly variable between enzymes, ranging from about c.a. 10^{-4} mismatch/bp for low-fidelity polymerases (typically *Taq* types) to 10^{-6} to 10^{-7} for high-fidelity ones (Kunkel and Bebenek, 2000)². High-fidelity polymerases, also called “proofreading” polymerases, differ from regular *Taq* DNA polymerases by (i) the fact that they do not add a final A tail in the 3’ end of the synthesised strand, and (ii) by their additional 3’→5’ exonucleasic activity, which removes mispaired bases by cutting out the last base incorporated in the synthesised strand. Thus, the proof-reading activity makes proof-reading enzymes more faithful in DNA replication, but also more demanding in reagents (i.e. require higher e.g. Mg²⁺ or dNTP concentrations), which inherently increases the costs of the reaction.

But the exonucleasic activity of proofreading polymerases has more insidious effects. Indeed, by generating fingerprinting profiles of amplicons obtained from pure bacterial cultures and amplified with different types of polymerases, we have demonstrated during my PhD that proofreading enzymes overestimate microbial diversity (Gury et al., 2008). We further showed that the exonucleasic activity of proof-reading enzymes generates trimmed DNA fragments of the same initial DNA molecule and that this feature tends to be exacerbated when PCR reagents are limiting. Such truncated DNA fragments can further anneal to wrong templates and continue to extend during the next PCR cycles, thus generating more chimeric fragments than do *Taq* DNA polymerases (Ahn et al., 2012). Hence, despite apparent desirable properties and wide use in the fields of eDNA and environmental microbiology, proofreading polymerases may not be the best choice for DNA metabarcoding analyses, chimeras being more difficult to identify than polymerase errors (section 5.3). This observation, alongside the fact DNA polymerases have different preferences for sequences with specific GC content (Nichols et al., 2018) and provide different diversity profiles (Zinger et al., 2007), emphasizes the need for explicit reporting of the DNA polymerase used in research articles making use of DNA metabarcoding.

5.3 Bioinformatics for eDNA data

Bioinformatic processing of eDNA data usually involves a preliminary step of data curation. This step generally consists in excluding obvious artefacts from the datasets (Figure 5.1), such as sequencing reads having mismatches in primers or tags/indices. Some researchers also trim the obtained sequencing reads by removing portions of the sequencing reads exhibiting low phred quality scores³. However, I’m not sure this approach is appropriate, as it generates truncated sequences that are difficult to deal with both bioinformatically and conceptually when it comes to sequence classification. It is also of common practice to exclude singleton sequences, as they correspond to errors in most cases and remain challenging to interpret anyway from an ecological perspective.

The second phase of the analysis lies in the classification of a list of DNA sequences that include both genuine and artefactual variation (Figure 5.2a) into biologically meaningful classes. In most cases, this classification is done with a limited to absent *a priori* knowledge of the samples taxonomic composition. This is probably the biggest challenge in bioinformatics applied to eDNA metabarcoding data. There are many ways this classification can be done, which are summarized in Figure 5.2b-d, with all approaches having their *pros* and *cons*. I have more specifically worked on unsupervised classification methods, as well as on sequence classification supervised by experimental controls.

²Estimates are provided for the same number of amplicon doublings.

³A mesure of the probability the sequencer has correctly read the nucleotide.

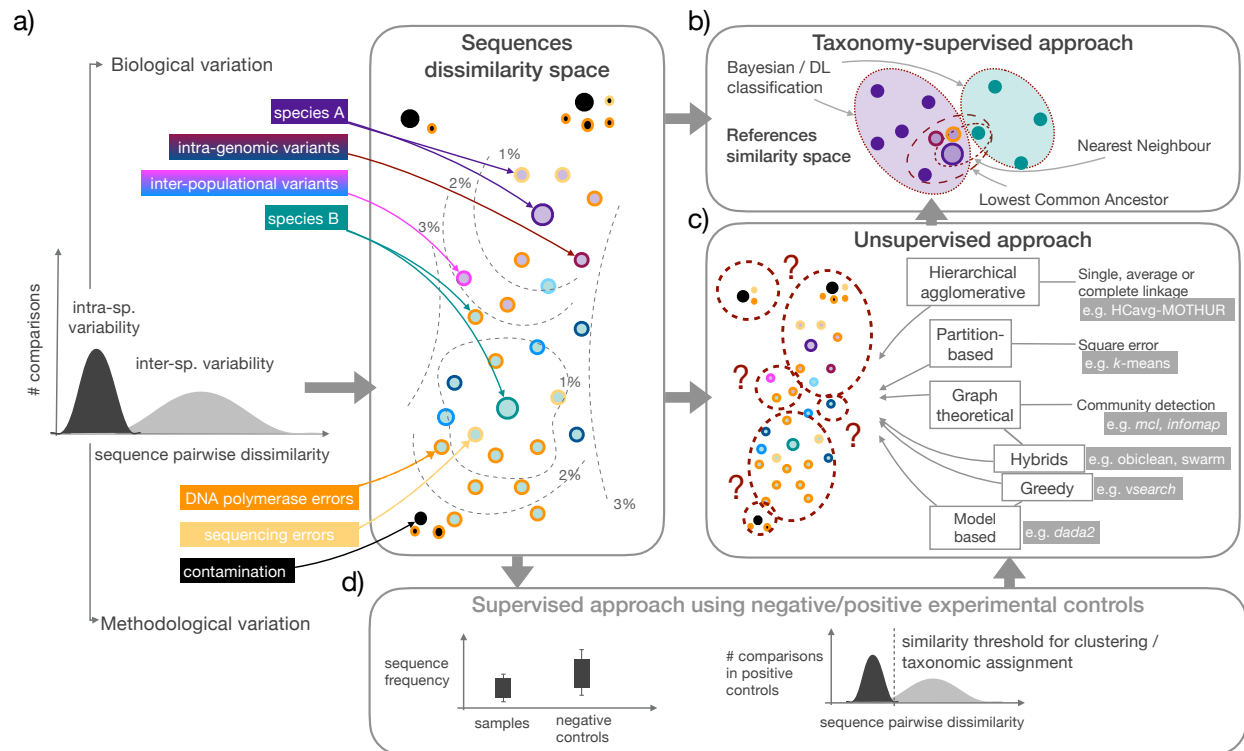


Figure 5.2: The problem of sequence classification in eDNA metabarcoding datasets.

(a) Each species detected in a DNA metabarcoding dataset exhibits a certain level of sequence variability. This variability can be biological, i.e. caused by intragenomic variation as for bacterial or fungal rRNA genes (Acinas et al., 2004; Simon and Weiß, 2008) or interpopulational variation as for the *COI* gene (Craft et al., 2010). It can also be caused by molecular artefacts (Table 5.1) or contaminations. Considering this variability in a sequence dissimilarity space, where dashed greyed isoclines represent sequence dissimilarity, illustrates the difficulty to distinguish/classify both biological and artefactual variants. (b) One way to do so is to assign each sequence a taxon name by comparison against a reference database (e.g. Genbank). This assignation can be the closest DNA reference sequence to the query, or to the Lowest Common Ancestor of a set of DNA references exhibiting a certain level of similarity with the query sequence (e.g. Huson et al., 2007), or by employing naive Bayesian classifiers (e.g. Wang et al., 2007) or more recently deep neural networks (e.g. Mock et al., 2022). However, this approach assumes that the reference database is errorless and complete, which is not the case for most organisms and DNA markers. (c) Alternatively, an unsupervised classification can be used. This is a vast topic. Many algorithms have been implemented to form groups of objects based on their (dis)similarities and can be easily transposed to our sequence classification problem. I will not detail them here as we did it already extensively in Taberlet et al. (2018). Briefly, there are several broad families: (i) agglomerative hierarchical clustering methods, (ii) machine learning partitioning methods, such as the *k-means* (Hartigan and Wong, 1979), (iii) community detection algorithms such as *mc1* (markov cluster algorithm, Dongen, 2000) or *infomap* (Rosvall and Bergstrom, 2008), (iv) graph and model-based approaches, where the graph is built based on a particular process, here PCR errors, such as *obiclean* (Boyer et al., 2016), or *swarm* (Mahé et al., 2014), (v) greedy (i.e. heuristic) versions of the latter, such as *vsearch* (Rognes et al., 2016), and finally (vi) denoising approaches that are based on a probabilistic error models, such as *dada2* (Callahan et al., 2016), who produce the so called “Amplicon Sequence Variants” (ASVs). Note that both (b) and (c) rely on sequence dissimilarities and hence require aligning sequences, which can be done through different algorithms (e.g. Needleman-Wunsch alignment, Needleman and Wunsch, 1970 or BLAST, Altschul et al., 1990) that have different performances. (b) and (c) also rely on measures of sequence dissimilarity, which can be done with different indices (e.g. raw or normalized editing distances, phylogenetic distances, etc.) that can also influence the sequence classification outcomes (Barley and Thomson, 2016; Zinger and Philippe, 2016). Note also here that I do not consider *k-mer*-based approaches. (d) Finally, another type of supervised classification relies on the signal contained in negative and positive controls to classify sequences (or MOTUs) as genuine or artefactual. Examples of strategies to do so are described in the main text.

5.3.1 Clustering methods for DNA sequences

During my thesis, available sequence classification tools consisted in applying hierarchical clustering on phylogenies built from multiple alignments of barcode sequences (Schloss and Handelsman, 2005). However, multiple alignments are not suitable for barcodes that are highly polymorphic in composition and length (e.g. the ITS, a marker I use to study fungal communities) or for datasets with a large phylogenetic breadth. More pragmatically, multiple alignment algorithms at that time were also too computationally expensive for large datasets. In collaboration with Eric Coissac, we built a high-performance MOTU definition procedure based on the Needleman-Wunsch pairwise sequence alignment algorithm Needleman and Wunsch, 1970 and a community detection method, `mc1` (markov cluster algorithm, Dongen, 2000) that we used successfully to describe soil alpine microbial communities (Zinger, Coissac, et al., 2009), with diversity patterns that were in line with those obtained with fingerprinting methods. We further upscaled the approach to analyse the landscape-scale distribution of soil alpine fungal communities, our first HTS dataset (generated through 454 pyrosequencing, Lentendu et al., 2011).

Later on, during my postdoc at EDB with Jérôme Chave, we needed an effective sequence classification method able to deal with large trans-kingdom datasets obtained from tropical soil samples and generated by Illumina sequencing. I have hence benchmarked several clustering methods encompassing the different clustering families shown in Figure 5.2, using a DNA metabarcoding dataset obtained from a plant mock community. This analysis has never been published because it was part of a manuscript that we abandoned. I hence take the opportunity of this HDR to make this analysis available, and because I think the dataset used for this benchmark illustrates very well various aspects of the eDNA metabarcoding data curation problem. This dataset corresponded to sequencing data obtained from eight PCR replicates conducted on a mock community containing a mixture of genomic DNA from 12 alpine plants (Figure 5.3a). It contained 1,062,395 reads and 6128 unique sequences, which is 500 times more than the 12 plant species initially expected (Figure 5.3a). This illustrates the high proportion of PCR/sequencing errors in this experiment. But by experience, it is a typical feature of DNA metabarcoding data⁴. Note here that many of the artefactual sequences were present in the 8 PCR replicates, demonstrating that sequence repeatability alone isn't sufficient to confirm a sequence's authenticity.

The benchmark included an average linkage clustering method, as implemented in `MOTHUR` (hereafter `HCavg-MOTHUR`), which belongs to the family of ascendant hierarchical clustering and has been widely used in microbial community studies before the rise of “denoising methods” (Figure 5.2). I also tested two methods belonging to the family of “greedy” clustering algorithms, namely `UPARSE` (Edgar, 2013) and `sumac1ust` (Mercier et al., 2013). Both algorithms consider the most abundant sequence as a seed for a first cluster. They then assign the other sequences to that cluster if they are similar enough to the cluster seed above a certain threshold, chosen beforehand. If not, the sequence is used as a seed for a new cluster. The same process is repeated until all sequences are classified. Third, I tested two algorithms that consider that errorless sequences should be more abundant than 1-error ones, and so forth (`obiclean`, Boyer et al., 2016 and `swarm`, Mahé et al., 2014), mimicking an accumulation of errors during the PCR amplification process. The two algorithms differ in their implementation but basically consist in detecting groups of sequences – or connected components – differing by 1 mismatch. Finally, I included in the benchmark two community detection algorithms based on information theory (`mc1`, Dongen, 2000, and `infomap`, Rosvall and Bergstrom, 2008). These algorithms detect communities in a graph as in Figure 5.3a by using random walks. This approach considers that if a random walker tends to be trapped in subregions of the graph, then the nodes are highly connected, and this subregion should be a “community”, i.e. a cluster.

Figure 5.3b-c demonstrates the very poor performance of `HCavg-MOTHUR` and `UPARSE`. Both methods produced respectively 40 and 15 times more MOTUs than expected (480 ± 70 and 180 ± 28 MOTUs) and displayed a low clustering quality ($NMI < 0.8$, see Figure 5.3c for an explanation of this index). Note that the low performance of these two algorithms can be partly due to their implementation, which relies on a percentage

⁴Such assessments are seldom made in the many microbiology studies studying sequence error rates. But often, the samples used are environmental samples, for which one has no expectations on the ground truth. Sometimes, mock communities are used but microbial mock communities are not always convincing, as it is challenging to really control what bacterial sequences are introduced in the mock community prior to sequencing due to a much higher contamination probability, if only to consider reagent contaminants (Salter et al., 2014)

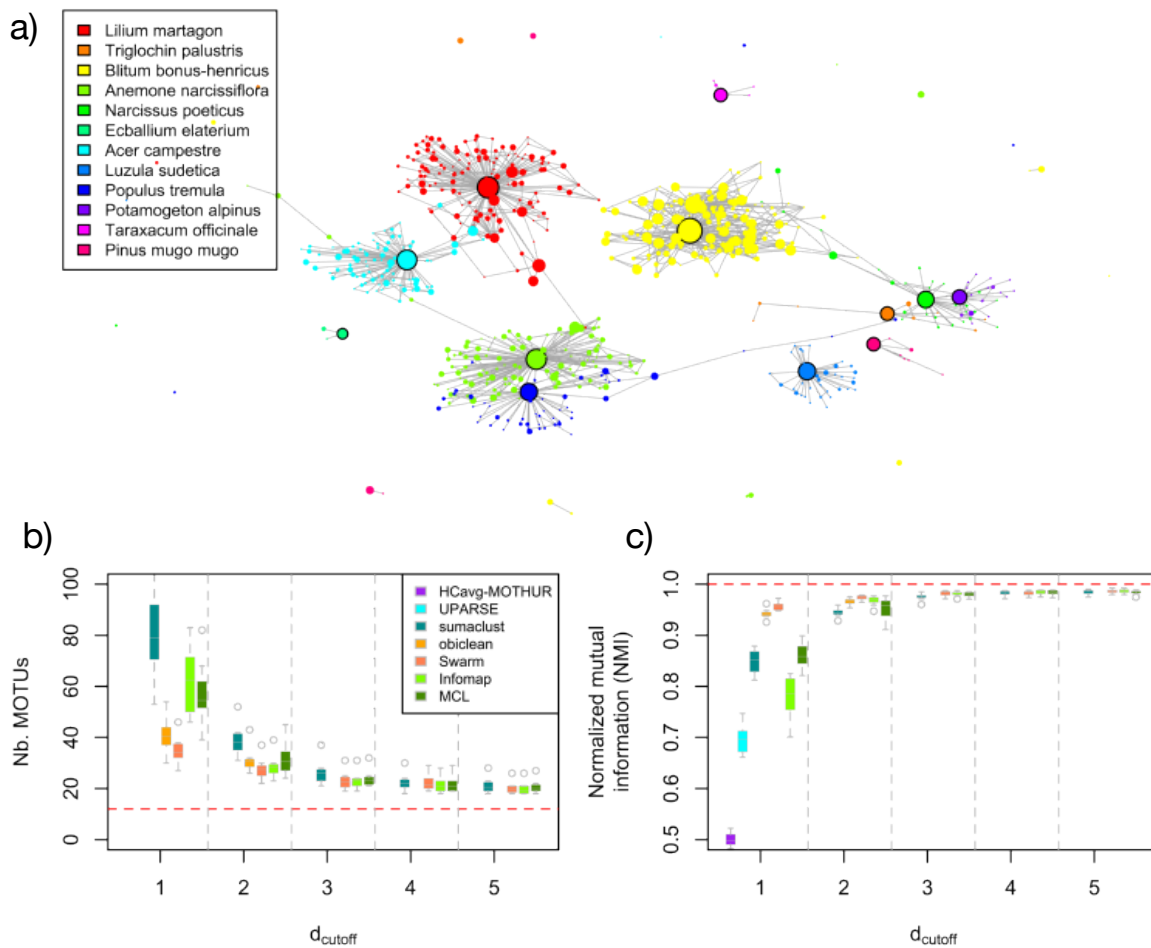


Figure 5.3: Benchmark of clustering methods.

(a) List of the 12 plant species included in the mock community, as well as their corresponding sequences dissimilarity space. The DNA metabarcoding data was generated using the *trnL* P6 loop marker (Taberlet et al., 2007) and similar protocols as in De Barba et al. (2014). This graph exemplifies the level of noise contained in these data: here, each node corresponds to one sequence, and edges represent a pairwise dissimilarity of 1 mismatch. Node sizes are proportional to each sequence log₁₀-transformed abundance, and black-framed nodes represent genuine sequences. Nodes are colored according to their taxonomic assignments, hence representing the “ground truth”. (b) and (c) Illustrate the clustering methods performances on eight PCR replicates of the 12 plant mock community: (b) Number of MOTUs found for different clustering method clustering thresholds (d_{cutoff}). HCavg-MOTHUR and UPARSE values are not represented in the plot due to a much higher number of MOTUs. Obiclean-based clustering was not done at $d_{cutoff} > 2$ due to computational constraints. The red dotted line indicates the expected number of MOTUs. (c) Normalized Mutual Information (NMI) scores of the methods tested at different clustering thresholds. Sequence taxonomic assignments were used as the “ground truth” to compute NMI scores. NMI measures the degree of similarity between two partitions. NMI=1 when the two clustering method outcomes are identical, as depicted by the red dotted line.

of dissimilarity and was adapted for longer markers. A percentage of dissimilarity may not be adapted to short markers or markers with high polymorphism in length as it is the case here⁵. This suggests that using absolute number of mismatches to measure the distance between pairs of unique sequences may be more appropriate in certain cases. For the other algorithms, clustering qualities varied but tended to stabilize at a clustering threshold of 3 mismatches. Because `obiclean` and `swarm` are explicitly implemented to detect PCR errors, the main source of variants in the studied dataset (and most likely in all other ones), they were the closest from the real grouping, both in terms of number and composition, even at a clustering threshold at 1 mismatch. Yet, they predicted nearly 3 times more MOTUs than the real figure at this threshold. The performance of all five algorithms converged for a clustering threshold of 3 mismatches, suggesting that the clustering is very robust at a certain similarity threshold irrespective of the method chosen. For later analyses I conducted for this specific project and some others, we made a trade-off between computational requirements and result quality and opted for `infomap`, which produced partitions with the highest NMI scores and has been further shown to be very robust on graphs of varying in size and properties (Lancichinetti and Fortunato, 2009).

5.3.2 Some thoughts on denoising methods

At the time when I initially conducted this benchmark, the so-called “denoising approaches” such as `dada2` (Callahan et al., 2016, 5.3c) were not published yet. I’ll discuss here shortly - and perhaps a bit provocatively - these methods and in particular the algorithm `dada2`. Its development undoubtedly marked a turning point in the quest for solving the problem of DNA metabarcoding sequence classification. Indeed, it proposed an elegant approach relying on a probabilistic model explicitly depicting the process of error generation during Illumina sequencing. Briefly, in this model, errors are expected to occur independently within a sequencing read and independently between sequencing reads. Based on this model, the algorithm then classifies sequences as genuine or artificially deriving from another one and excludes the latter⁶. Without questioning the validity of `dada2`’s error model, one may question its main assumption, i.e. that most errors do occur during the sequencing phase. To my knowledge, no study has yet explicitly quantified at what step most errors do occur. On the other hand, the model developed by Sylvain Moinard (Moinard, 2023), which explicitly reproduces a PCR error process, i.e. an accumulation of mutations globally and within sequences when the number of PCR cycles increases, also performs well in classifying artefactual from genuine signal. This, together with the fact that several empirical studies have shown that the error rate in DNA metabarcoding data tends to increase with the number of PCR cycles (Jönsson et al., 2023; Patin et al., 2013; Sze and Schloss, 2019), question `dada2` model’s main assumption.

Denoising methods have further been often presented as disruptive, or even opposed to more traditional clustering approaches (Callahan et al., 2017). However, from my understanding, they eventually do not differ so much: they include a sequence alignment procedure, estimate sequence similarity with classical approaches, and include a step of partitioning (or, broadly speaking, of clustering). The difference hence lies in that the partitioning is based on the sequencing error model. Finally, denoising methods have been presented as almost infallible, and as a way to facilitate cross-studies comparisons, as their products are sequences (“Exact Sequence Variants” or “Amplicon Sequence Variants”) when MOTU are described by unlabelled entities that are dataset-dependent Callahan et al., 2017. Apart from the fact that a MOTU can be quite robustly represented by its dominant sequence (5.3a), I think that this ASV vs. MOTU dichotomy somehow distorts the initial meaning of what a MOTU is. The MOTU concept has clear roots in taxonomy and phylogeny sciences: it’s a biological unit regrouping individuals that exhibit a certain level of sequence similarity for a given DNA marker. It ignores methodological/artefactual sequence variability and does not explicitly rely on a particular partitioning method. In that context, an ASV is not more than a particular case of MOTU at 100% of sequence identity of genuine biological sequences. Then, the sequence similarity threshold required to appropriately define a biological entity and to study a given phenomenon is obviously marker/taxon dependent (Bonin et al., 2023), or even question dependent. That `dada2` still performs well in general (Callahan et al., 2017; Caruso et al., 2019; Pauvert et al., 2019) mirrors the convergence of

⁵This observation may well apply for the widely used plants or fungal ITS1 or ITS2 markers

⁶Note that I do not master fully the intricacies of the `dada2` error model, its parametrization, or implementation, but I had many discussions with experts on the topic, including Eric Coissac, Frédéric Boyer, and Sylvain Moinard (whose part of the thesis examine specifically this algorithm).

algorithms performance observed in Figure (5.3b-c), as well as the fact that the diversity patterns obtained with denoising vs. clustering approaches do correlate very well in general (Glassman and Martiny, 2018). This suggests that the sequence dissimilarity space of DNA metabarcoding data might be sufficiently discontinuous to be partitioned relatively consistently, whatever the approach used (few exceptions excepted).

I won't further develop here other claimed advantages of the denoising approaches that I find questionable (e.g. their better appropriateness to assess intraspecific variation when most markers we generally use do not reach this level of resolution, or the assertion that an ASV is dataset independent, when it relies on a partitioning step that is inherently dataset dependent). At present, I lack more material of discussion and fully covering the denoising methods topic would require a dedicated article (which we actually started to write with Eric Coissac). But while I do believe that current denoising approaches can be valid to curate the data from a particular type of noise, i.e. sequencing errors, their superiority over traditional clustering approaches, remains, in my opinion, to be demonstrated. In addition to the points raised above, several analyses now suggest that denoising methods tend to oversplit the data either by revealing some intragenomic variants, or because they still contains errors (e.g. Estensmo et al., 2021; Johnson et al., 2019; Nearing et al., 2018; Schloss, 2021). To conclude, my overall feeling is that there has been quite a hype in the way these methods have been presented to the scientific community, which has created more confusion and dogmatism than the method itself has really improved the sequence classification problem. I've witnessed this dogmatism myself on many occasions as an author or through my editorial activities, where I've seen certain editors or referees take strong, often aggressive, and often scientifically unsupported positions on the superiority of denoising tools to invalidate results based on more traditional clustering approaches. I really hope this will change in the future.

5.3.3 Using experimental controls to detect artefactual signal

Back to our benchmark of clustering methods, Figure 5.3b suggests that even for high clustering thresholds, the MOTU richness remains 3 times higher than in reality. Comparing the most abundant sequence of these additional MOTUs against Genbank showed that part of them corresponded to contaminants, i.e. plant species often amplified in the lab or recurrent reagent contaminants (e.g. hemp or potatoes), as well as potential chimeras, these latter exhibiting similarity scores against Genbank DNA references below 80% sequence identity. Regarding the problem of chimeras detection (Table 5.1), there are, to my knowledge, no satisfying tools allowing us to detect them correctly, as available tools generate substantial amounts of false positives and/or negatives (e.g. Bjørnsgaard Aas et al., 2017; Pauvert et al., 2019). They further perform poorly on sequences below 300 bp (Wright et al., 2012), which is the typical length of DNA metabarcoding sequences (c.a. 50 to 300 bp). Regarding the problem of reagent contaminants, the "tag-jumps" bias (Table 5.1, Figure 5.4) renders it impossible to curate the dataset solely based on a sequence/MOTU presence in negative controls.

Between 2015 and 2020, Frédéric Boyer and I separately developed a range of procedures (written in the R language) to deal with these problems, amongst other things. All these procedures have the same philosophy: they rely on the behaviour of the information contained in experimental controls as compared to PCR products of environmental samples, as well as on technical replicates (Figure 5.3d). For example, I've developed a function able to automatically detect reagent contaminants: any MOTU with a frequency (over the entire dataset) being maximal in one negative control or being on average maximal across all negative controls must be a contaminant. Note that albeit this function performs relatively well, it does not work well when a negative control has been contaminated with a DNA extract from the experiment, or when the experimental design involves samples with very low PCR success, in which case they will behave similarly to negative controls.

Frédéric and I identified a gap in the existing literature regarding such tools. We therefore decided to encapsulate our different procedures in a single R package, `metabar` (Zinger et al., 2021, see also the GitHub page and companion descriptive website), together with Clément Lionnet, an informatic engineer supervised by Frédéric, as well as [Anne-Sophie Benoiston](#) and [Julian Donald](#), who were postdocs I was mentoring at that time. To deal with the intrinsic multidimensionality of DNA metabarcoding data, we developed a single object containing the MOTU by samples abundance community matrix, descriptions relative to each MOTU, as well as metadata at both the PCR level and sample level. We also added several functions

allowing import/export of objects, functions allowing further manipulation of the sequences and dataset, and functions allowing better visualization of some features of the DNA metabarcoding datasets in their experimental context (e.g. Figure 5.4). This tool is now routinely used in our group and others, and we plan further improvements and integration with other tools developed by Eric Coissac in the mid-term.

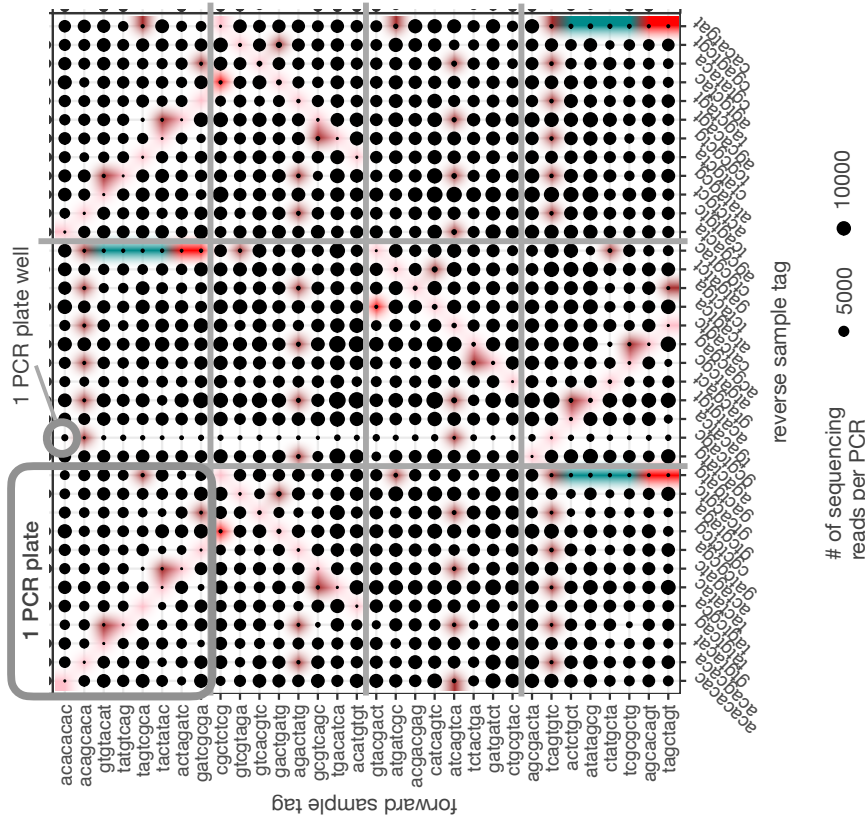


Figure 5.4: A typical DNA metabarcoding experimental design and output. The figure shows a typical experimental design, consisting of 12 PCR plates, where both samples and different types of experimental controls are included. Nucleotide tags represented at each row and column of this design correspond to the tags added in the 5' end of the forward or reverse primer. One well therefore has a unique tag combination across the full dataset and allows us to identify each independent PCR in downstream analyses. As the reader might see, all negative controls do contain at least a few sequencing reads, even sequencing controls (i.e. empty PCR wells, or in other words, unused tag combinations in our design). This is a typical feature of DNA metabarcoding data and this is caused by contaminations and tagjumps. Note that this dataset has been generated with a library procedure supposed to strongly reduce tag-jumps in the data (i.e. the METAFAST protocol developed by the Fastaris sequencing company). For this specific example, one can also notice that all PCR reactions conducted with the reverse primer including the tag “gggtcagc” exhibit low amounts of reads. It turns out that this particular tag forms a secondary structure with the reverse primer used here, hence considerably lowering the efficiency of the PCR. This figure is an example of output from the `ggpcrtag` function of the `metabar` R package.

Control type:

Sequencing blank controls:
 Consist in detecting at the bioinformatics level all sample tag/indices combinations that were not used in the DNA metabarcoding experiment. Physically in this design, the PCR plate wells were left empty.
 → Detection and estimation of “tag-jumps” rate.

PCR blank controls:
 Sterile water used to prepare the PCR reaction mix is used in place of the DNA template for PCR amplification.
 → Detection of “tag-jumps” and identification of contaminants introduced at the PCR stage

DNA extraction blank controls:
 The whole DNA extraction procedure is conducted without starting material and result in a DNA extract supposed to contain only sterile elution buffer.
 → Detection of “tag-jumps” and identification of contaminants introduced at the PCR and DNA extraction steps.

Positive controls:
 PCR of a given set of species for which DNA sequence and copy number of the target gene are known.
 → Detection of “tag-jumps”, of contaminants introduced at the PCR and DNA extraction steps, estimation of error rate and of amplification biases.

5.4 How to sample eDNA?

The question of sampling is essential in ecology, including when dealing with eDNA. Surprisingly, it has comparatively received much less attention than molecular and bioinformatic approaches discussed above (Blackman et al., 2024; Dickie et al., 2018). I guess it is due to (i) our collective tendency to focus on techno-scientific advances and solutions at the expense of more fundamental questions⁷, but also (ii) because this problem is not a trivial one. Indeed, the spatio-temporal distribution of eDNA results from a multitude of biological and abiotic processes (Figure 5.5). This multiplicity causes large uncertainties on the spatio-temporal window covered by a single punctual eDNA-based observation, which has direct implications for the inferences that can be made from eDNA metabarcoding data. While I've not published yet explicitly on these issues, I'm regularly confronted with this problem when analysing my data and have conducted several analyses in this regard. More specifically, I will discuss briefly some thoughts about two aspects of eDNA sampling: (i) the problem of sampling itself, focusing more particularly on the soil environment, and (ii) the actual meaning of eDNA populations (i.e. total, intracellular, or extracellular) sampled.

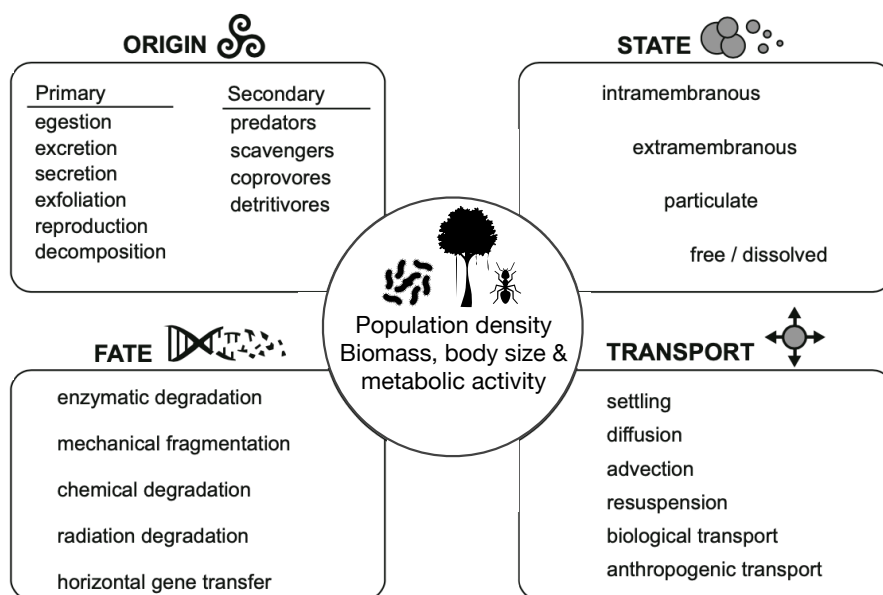


Figure 5.5: Determinants of eDNA abundance and distribution in space and time. Figure modified from Barnes and Turner, 2016. I added population size, organism biomass and metabolic activity as central determinants of the initial number of eDNA molecules (both intra- or extra-cellular) of a given species in the environment. Typically, most samples have an eDNA largely dominated by the DNA of microbial populations present in the sample. Likewise, dense populations or large macroorganisms expel greater amounts of DNA in their environment (reviewed in Taberlet et al., 2018). In addition, the number of ribosomal genes or organelle correlates with cell size for both uni- and multicellular organisms, because larger cells have usually higher demand of energy (Marguerat and Bähler, 2012; Okie et al., 2016). As such, a recent study has shown that allometric equations can improve the relationship between eDNA abundance and organisms abundance (Yates et al., 2021). The number of copy of rRNA genes have also been reported to depend on organisms k/r strategy, and hence of their metabolic rates (Roller et al., 2016). Then, the eDNA spatiotemporal distribution depends on the transport of both extra- and intracellular DNA, often through water movement, as well as extracellular DNA persistence in the environment, which can last from several hours to thousands of years depending on e.g. temperature, UV radiation, and microbial activity Barnes and Turner, 2016; Nagler, Insam, et al., 2018.

⁷I'm aware that part of my work is a good example of this behaviour. It actually raise a number of questions related to techno-solutionist thinkings, that are discussed in Shen et al. 2023.

5.4.1 Accounting for the variation of eDNA in space and time

Soils are probably the most complex and dynamic environmental matrix in terms of physical, chemical, and biological properties. In particular, the soil biota is expected to exhibit a steep heterogeneity horizontally, vertically, and through time (reviewed in Bardgett and van der Putten, 2014; Berg, 2013; Ettema and Wardle, 2002; Nemergut et al., 2013). This heterogeneity is visible through eDNA: all soil eDNA datasets I have analysed in the past exhibit a signature that is very punctual spatially and temporally for both micro- and macro-organisms. For example, considering individual soil cores as a basic sampling unit, bacterial, eukaryotic and plant communities retrieved with eDNA metabarcoding can be taxonomically just as different over short distances (10 meters) as they are over much larger distances (10 kilometres), hence resulting in almost flat distance-decay curves (Figure 5.6). However, distance-decay patterns do emerge when aggregating soil community data at coarser sampling grains. This is consistent with previous studies, reporting the better predictability of soil microbes composition when increasing the scale of observation (Averill et al., 2021; Barberán et al., 2015). The underlying explanation usually proposed is a scale-dependency of community assembly processes, with neutral processes being more important at smaller scales than niche-based ones (Chase, 2014, see also section 6).

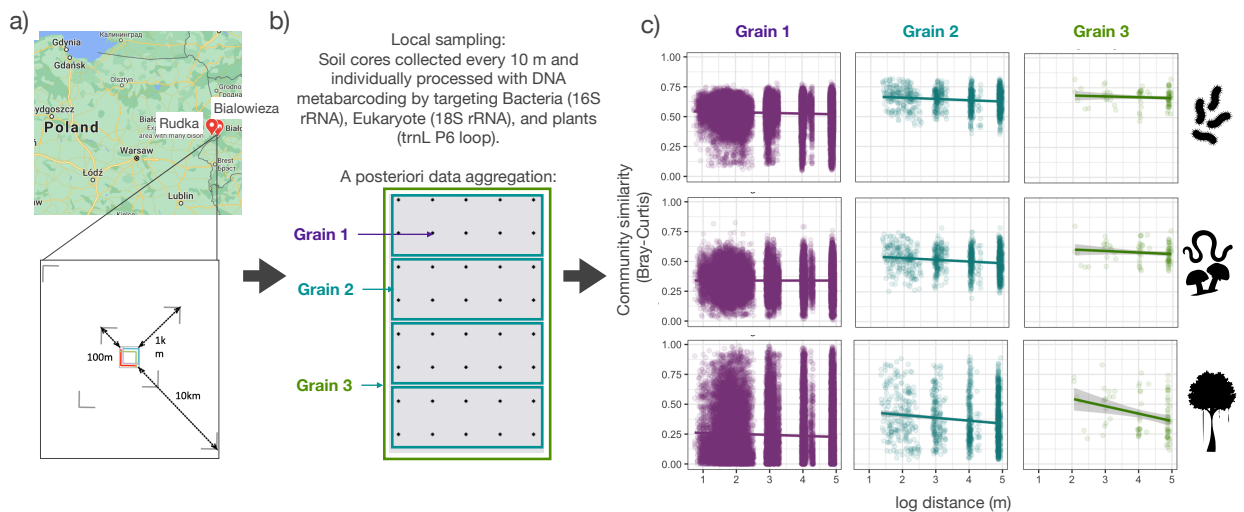


Figure 5.6: Spatial patterns of the soil biota retrieved with eDNA. The example patterns shown here result from a sampling campaign conducted in Bialowieza (Poland) that I designed and conducted together with Pierre Taberlet, Frédéric Boyer, Eric Coissac, and Heidy Schimann to assess the spatial patterns of soil biodiversity in forest ecosystems. (a) The map shows the location of two studied forest fragments: Rudka (managed) and Bialowieza (old growth) forests. Below is shown the distribution of different plots (full grids vs “light grids” plots having a “L” form) in each site. (b) Sampling design in each local plot (shown here only for the full grid design), as well as indications on DNA metabarcoding processing and post-data aggregation (i.e. sum of MOTUs abundances) at different spatial grains. (c) Distance-decay of similarity (Bray-Curtis index) for bacteria, eukaryote and plant communities at different grains of data aggregation. Unpublished results and data analysed with [Anne-Sophie Benoiston](#).

Neutral processes are probably not the only explanation behind this result. Severe undersampling of communities having species with highly clumped spatial distributions, as is the case in the soil environment (Berg, 2013; Ettema and Wardle, 2002), can also explain these results (Beck et al., 2013). Also, there is a recurrent scale mismatch between the set of environmental drivers we usually consider, which are “more human-scale

compatible”, and the actual variation we are measuring with eDNA at the scale of a soil core, which may be better explained by fine-scale biotic interactions or micro-heterogeneity of abiotic parameters that cannot be easily measured. To reduce fine-scale heterogeneity while allowing studying ecological systems at a scale that remains compatible with traditional approaches (i.e. ~ 0.01 and 1 ha in grassland and forests respectively), we now use alternative grid-sampling schemes, with fewer sampling points per sampling unit, but where each sampling point is made of a composite sample of several soil cores collected in the sampling point vicinity. The approach is more cost effective, as the number of samples to process is lower, and may be more appropriate for sampling multi-trophic communities given the body-size dependence of distance-decay patterns, even at small spatial scales (Figure 5.6). However, we still need to test this experimentally and assess the design efficiency to capture plot-scale diversity for multiple taxa.

The above discusses horizontal spatial variation, but soil diversity is also highly compartmentalized vertically through steep gradients in abiotic and biotic conditions (i.e. organic matter quality, rhizospheric influence, etc., Bardgett and van der Putten, 2014; Berg, 2013; Ettema and Wardle, 2002). Accordingly, we have shown that soil bacterial and eukaryotic communities are compositionally much more different when comparing vertical layers (i.e. litter vs. topsoil compartments) within a single habitat than when comparing communities from distinct tropical forest habitats (i.e. white-sand *vs.* *terra firme* forests). Litter communities indeed tend to be enriched in DNA belonging to copiotrophic bacteria (e.g. Actinomycetota, Alphaproteobacteria), ants, and nematodes, when topsoil communities are rather enriched in DNA corresponding to oligotrophic bacteria (e.g. Verrucumicrobiota, Acidobacteriota), termites, and earthworms (Zinger et al., 2020). The distinct features we observed across the soil profile strongly suggest steep variations in carbon cycling between soil horizons. This is of course not surprising. But because most eDNA studies of soil biodiversity only analyse the topsoil matrix, this result raise a critical question: do we currently correctly sample this compartment? how can we better incorporate the soil vertical dimension, especially the litter layer, when linking eDNA-derived soil biodiversity data to ecosystem functions like ecosystem productivity and C storage? Does the diversity of the two layers can be appropriately assessed on composite samples? These are questions we plan to ask as part of [Jérémy Raynaud's](#) PhD thesis.

The problem of scaling and compartmentalization of soil biodiversity holds true along the temporal axis (Bardgett and van der Putten, 2014), and this applies to soil biodiversity's eDNA as well. First across seasons, where ecosystems do experience strong changes in plant communities, soil organic matter quality, temperatures/moisture, soil communities as observed through eDNA do exhibit steep seasonal dynamics in various environments, ranging from alpine grasslands, where microbial communities do experience full successional dynamics over the year (Schmidt et al., 2007; Zinger, Shahnava, et al., 2009), to tropical forests, where both microbial and fauna communities of the soil environment largely differ between the dry and the wet season when plant ones - mainly composed of trees - do change (Zinger et al., 2020).

Beyond seasonal dynamics, we are now at a stage where eDNA-based time series spanning multiple years are becoming available. While designed to minimize interannual fluctuations, e.g. by sampling at the peak of the growing season, eDNA-based time series data often exhibit surprising interannual variability that are sometimes even stronger than habitats or treatments effects (see Figure 5.7), as we observed for example for the ECOFEED project, led by Julien Cote. Similar observations were made for the TransAlp project, led by Tamara Münkemüller, where soils from different altitudes were cross-transplanted to study the effect of warming on alpine communities, and for which we also observe important interannual variations. Inter-annual technical effects are unlikely, as, in the case of the ECOFEED project, all samples were stored at -80°C and then all processed for DNA metabarcoding analyses at the same time. I'll discuss the underlying biological reasons for such dynamics later (section 6). In any case, the seasonal or inter-annual dynamics observed in soil eDNA metabarcoding data pose a challenge for generalizations of soil diversity patterns from single-point-in-time observations and highlight the need for improved approaches in both sampling design and/or modelling for future studies.

To sum up, with eDNA methods, we most likely face a mismatch between how an ecological phenomenon is measured and the actual scale at which this phenomenon occurs. This mismatch is not unique to eDNA-based approaches, it is a recognized issue in ecology today (Estes et al., 2018). Development of biodiversity measurements through multiple approaches, both *in situ* and remote can bridge this gap. While the high throughput of eDNA methods holds promise in that sense, unlocking its full potential still requires addressing

critical unknowns about the spatial and temporal window and type of information captured by eDNA across spatial and temporal scales.

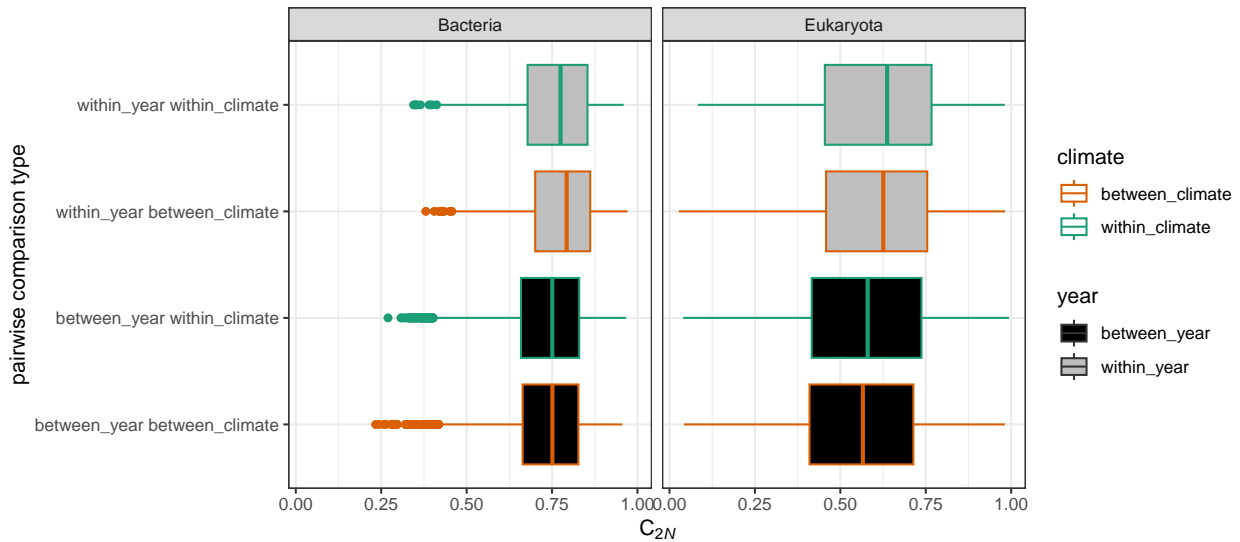


Figure 5.7: Soil community similarity across treatments and years. The example dataset results from a 7-year-long experiment manipulating climatic conditions to study the response of soil communities to warming. It is part of the ECOFEED project, led by Julien Cote in which I’m strongly involved (see section 6.5). Here are shown the pairwise similarity values, C_{2N} , at $q=2$ in the Hill number framework (corresponding to the Morisita-Horn index, Chao et al., 2014) within and between experimental treatments and sampling years. For each condition, the level of replication is $n = 12$ plots of 100m^2 , where 3 to 5 soil samples were collected, and subjected to DNA metabarcoding targeting the 16S rRNA gene for Bacteria, and the 18S rRNA gene for eukaryotes. MOTU abundances retrieved from each soil core replicate were aggregated at the plot level as discussed above. A PERMANOVA of corresponding dissimilarity patterns including climatic treatments and sampling years as explanatory variables suggests that only sampling years have an effect on bacterial dissimilarity patterns ($R^2 = 0.08$, $p=0.001$), and that the sampling year effect is equivalent to that of the climatic treatment for eukaryotes ($R^2 \sim 0.021$ for both factors, $p<0.02$). Unpublished results and data.

5.4.2 Total, intracellular or extracellular eDNA?

The temporal dynamics of the eDNA signal discussed above were observed on data obtained with Taberlet et al. 2012 DNA extraction protocol, which is biased towards extracellular DNA. That we observe dynamics on this specific DNA population is somewhat in contradiction with the long persistence of DNA observed for some species in e.g. alpine soils (Yoccoz et al., 2012), as well as the common belief in microbial ecology that persistence of extracellular DNA (i.e. “relic” DNA) obscures contemporaneous signal (Carini et al., 2016). Indeed, the death of microbial cells or of organisms in general is not synonymous of the extinction of their population of origin. Active microbial populations can release DNA through a continuous process of cell turnover, viral predation, quorum-sensing signals, and biofilm construction (Vorkapic et al., 2016; Williamson et al., 2017). Consequently, extracellular DNA could actually be indicative of active and abundant populations in soils, as recently suggested (Nagler, Podmirseg, et al., 2018).

On the other hand, targeting specifically intracellular DNA when studying microbial communities, as it has been often recommended, may actually enrich the data with dormant cells, which can constitute up to 80% of soil bacterial diversity (Lennon and Jones, 2011), hence potentially biasing the results towards inactive members of the community. The debate on what DNA population between intra vs. extracellular DNA needs to be sampled to appropriately describe biological community also arose in the eDNA community recently, where some authors argued that we should differentiate intra- vs extra-organismal DNA when interpreting eDNA metabarcoding data because extra-organismal DNA can correspond to organisms/populations that

are spatially or temporally distant from the sampling point (Rodriguez-Ezpeleta et al., 2021). I do not believe that making this distinction is very helpful. First because the reasons I invoke above apply to macro-organisms as well: that the DNA of a given species is in an extracellular form does not always mean that its population of origin is distant geographically nor that it has gone extinct. Second, because in practice, it remains difficult and costly to separate and study in parallel these two DNA populations (Pawlowski et al., 2021). A selection by DNA fragment size (e.g. filtration with different mesh size as done in the *Tara* Ocean project (Karsenti et al., 2011) or amplification of markers of different length, as proposed recently for diet analyses (Kreherwinkel et al., 2017)) may distinguish at minimum expenses fresh vs. older eDNA. However, this would require calibration, as, to my knowledge, these approaches have not been tested yet for soil biodiversity assessments.

To sum up, both intra- and extracellular fractions can encompass inactive populations, corresponding to either dormant cells or propagules or necromass. In the case of extracellular DNA, it is reasonable to expect that the relative abundance of extracellular DNA from a given species is inversely proportional to the species spatio-temporal distance from the sampling point, because in dynamics systems such as stream or ocean surface waters, as well as in surface soils, the decay rate of eDNA is relatively fast (i.e. from about few days to several weeks, Barnes and Turner, 2016; Nagler, Insam, et al., 2018), which limits its transport. Consequently, I would rather tend to analyse and interpret extracellular DNA of soil samples as a signal that is mostly contemporaneous and local in most ecosystems/environments.

5.5 Ecological analyses of eDNA data

Compared to traditional surveys, eDNA metabarcoding data have specific intrinsic properties which may need to be accounted for in downstream analyses. Albeit very similar to traditional species abundance data, eDNA metabarcoding data are expressed in fundamental units (i.e. sequencing reads for each MOTUs/Taxa, 5.1) that largely differ from traditional observations. We've also seen above that several molecular biases (Table 5.1) can distort species abundance distribution and artificially inflate the observed diversity. The implications of such differences on subsequent ecological inferences remain for now not well identified/quantified but may be non-negligible, in particular when the inferences rely on theoretical frameworks and models that explicitly refers to species and individual-based abundances (e.g. niche or neutral models, species abundance distributions). Accordingly, we have shown that adding different levels of noise in both abundance and species richness in species abundance distribution can bias the estimation of certain parameters of Hubbell's neutral theory model when the noise is above a certain level; a level that is reached in mock communities (Sommeria-Klein et al., 2016). These considerations challenge the validity/relevance of several diversity estimates or ecological inferences that have been made directly from traditional ecology approaches or theoretical models (e.g. Locey and Lennon, 2016; Shoemaker et al., 2017; Sloan et al., 2006). Here, I will discuss what approaches may allow us to better analyse and interpret eDNA metabarcoding data.

5.5.1 Uncertainties on the taxonomic/phylogenetic resolution

The first issue lies in that an ensemble of ASVs or MOTUs cannot really be considered as a proxy of an ensemble of species. First, the intra and interspecific DNA marker variability (and hence evolutionary rates) do vary across lineages (e.g. Bonin et al., 2023; Nilsson et al., 2008). eDNA data hence includes taxonomic entities of varying taxonomic levels. Second, it is unlikely that we reach the species or even genus level for most markers used for eDNA metabarcoding applications. MOTUs have been historically often defined at a sequence similarity threshold of 97% to 99% across marker genes and taxonomic groups (e.g. Stackebrandt and Goebel, 1994 for the bacterial 16S rRNA gene, Hebert et al., 2003 for metazoan's COI marker, Schoch et al., 2012 for ribosomal markers in fungi, and Kocher, De Thoisy, et al., 2017 for the 12S mt rRNA gene of mammals), based on assessments made on full-length genes. However, in most case, these thresholds only correspond to the level of sequence similarity below which one is sure to observe different species. In other word, many species actually exhibit the same barcode sequence. In addition, eDNA metabarcoding markers are about one order of magnitude shorter than full barcoding genes, which inevitably reduces data taxonomic resolution. Using long-reads sequencing platforms (e.g. Oxford Nanopore Technologies or PacBio sequencers) will certainly alleviate this issue (Tedessoo et al., 2018), but they remain relatively costly and/or

unstable. They are also appropriate only for particular applications/questions that target intracellular DNA: extracellular DNA can be indeed highly degraded/fragmented, which prevents PCR amplification of long fragments (Taberlet et al., 2018).

When the target DNA marker lacks resolution, it could be possible to refine the taxonomic inference *a posteriori*, if one knows what species are locally present and if these species are not too close phylogenetically. For example, if the presence of a plant family is identified by metabarcoding⁸ and there is only one locally known representative species of this family, then we can reasonably conclude that the detected signal comes from this species. However, local checklists are often not available. Soon, we will most likely be able to solve this problem for some taxa, by using species distribution models coupled with occurrence databases, such as the Global Biodiversity Information Facility (GBIF, www.gbif.org) to classify sequences in a taxonomy- and geography- supervised manner, as it is done for example for plants in the citizen science platform Pl@ntNet (Botella et al., 2019, plantnet.org). However, achieving this is likely out of reach for the neglected biodiversity. This highlights the need to estimate and account for the taxonomic resolution of the chosen DNA marker at the result interpretation stage (Bonin et al., 2023).

By avoiding the definition of classes, phylogenetic-based diversity or community similarity indices can, to a certain extent, better account for the actual nature of eDNA data. Indeed, they allow dealing with DNA marker's heterogeneous taxonomic resolution and can downweight artefactual MOTUs, of which DNA sequence is close from genuine ones. Phylogenetic-based approach can further greatly improve our understanding of various ecological patterns, as they result from processes that occur across multiple phylogenetic scales (Diniz-Filho et al., 2024; Graham et al., 2018). However, the short and hypervariable DNA markers typically used for DNA metabarcoding often exhibit high genetic saturation. As such, they are selected to maximize distinction between taxa, but not necessarily to accurately represent their respective evolutionary histories. Consequently *de novo* phylogenetic inferences, i.e. phylogenies built solely based on eDNA sequence data, often result in poor phylogenies and poor associated phylogenetic diversity estimates (e.g. Dupuis et al., 2012 for eukaryote markers, Hassler et al., 2022 for the bacterial 16S rRNA gene).

Phylogenetic placement methods represent an interesting alternative, provided that a robust backbone phylogenetic tree is available (e.g. `pplacer`, Matsen et al., 2010, `EPA-ng`, Barbera et al., 2019, or `EPIK`, Romashchenko et al., 2023). The emergence of large-scale initiatives of genome sequencing such as the European Reference Genome Atlas (ERGA, www.erga-biodiversity.eu) will certainly accelerate the production of such backbone, multi-locus phylogenies. However, they're for now available only for few emblematic groups (e.g. angiosperms, vertebrates), and remain lacking for the neglected biodiversity. Soil biodiversity phylogenetic structure can be retrieved with metagenomics or metagenome skimming approaches (Figure 4.3, Andújar et al., 2015; Papadopoulou et al., 2015), but these impose other challenges related to sampling, costs, and computational approaches, that I will not develop here.

There are more pragmatic ways to better account for the phylogenetic/taxonomic proximity of observed MOTUs and study processes that may occur at different phylogenetic scales that I've used in the past. For example, one can use phylogenetic diversity metrics either based on pairwise sequence similarity rather than on phylogenetic distances *per se*, or by stitching MOTUs on taxonomic trees or, when available, on backbone phylogenies, as we did to study the co-variation between alpine soil microbial and plant communities (Chalmandrier et al., 2019). Alternatively, one can analyse diversity patterns at different taxonomic resolutions (Roy et al., 2019; Zinger, Amaral-Zettler, et al., 2011). Simple heuristic approaches have been already shown to perform better and faster than more elegant/elaborated ones for taxonomic assignments (e.g. Hleap et al., 2021). An interesting complementary analysis would be to empirically assess whether the phylogenetic distances and structure inferred with the aforementioned heuristic approach are consistent with the genuine phylogenetic structure of communities and whether they outperform or not those obtained through *de novo* phylogenetic trees inference from eDNA metabarcoding data.

⁸Typically the `*trn*L` or `*rbc*L` marker genes do not exhibit sequence variation within several plant families Taberlet et al., 2007

5.5.2 Sequencing reads abundances and diversity measurements

The other uncertainty of eDNA data relates to the meaning of sequencing read counts. We've seen above that molecular biases can distort taxa abundances. Another matter is that of eDNA data compositionality, which is linked to the standardization of PCR products before sequencing or to the fixed sequencing capacity of HTS platforms. This problem is now often put forward in environmental microbiology (Gloor et al., 2017) because data compositionality imposes several constraints in terms of data analysis. At first glance, I think that this particularity may be a problem for specific ecological questions that require absolute quantification of taxa (e.g. identify indicator species) than for questions that are more related to community structure or diversity. It might not be so much an issue if one analyse non-standardized PCR products sequenced on a single HTS run, as me and my collaborators do in our experimental designs. But I won't dig further into this matter, as for now, I have limited hindsight/reflection on the issue. In any case, it will certainly be resolved by making eDNA metabarcoding data more quantitative (Lloréns-Rico et al., 2021), e.g. by quantifying/estimating the number of copies of the target gene using e.g. (d)dPCR, or by using spiking approaches (i.e. introduce in the PCR reaction a set of DNA molecules of known composition and number). I will develop this a bit in section 7.

More fundamentally and still leaving molecular biases aside, the number of target copy genes per genome or individual varies greatly across clades (e.g. Kembel et al., 2012 for bacteria, Gong and Marchetti, 2019 for protists, Lofgren et al., 2019 for fungi, etc.) and, in multicellular organisms, across tissue types (e.g. root vs. leaves will obviously not contain the same amount of chloroplastic genes). As a consequence, it has been often advocated to correct relative abundance retrieved in eDNA metabarcoding data, either with gene copy number databases associated with predictive tools based on the principle that such traits are phylogenetically conserved (Gao and Wu, 2023; Kembel et al., 2012), or with allometric equations (Yates et al., 2021). While this endeavour is necessary when relating eDNA metabarcoding data to individual or population-based processes, as in *lodka-volterra* or neutral models (Hubbell, 2001; van den Berg et al., 2022; Woodcock et al., 2007), the approach remains difficult to implement because this trait is not known for many clades. One way to circumvent this would be to use instead single-copy genes as DNA markers, as we proposed recently to study jointly bacterial and eukaryotic phytoplankton communities (Pierella Karlusich et al., 2023). Alternatively, observed relative abundance in eDNA data may be interpreted as proxies of biomass/activity, but in the latter case, this interpretation should be accounted for in the analyses or models used, and their associated ecological inferences, as species abundance distributions do not necessarily have the same properties than species biomass distributions, nor they are driven by the same processes (Matthews and Whittaker, 2015; Morlon et al., 2009).

The uncertainties on the meaning of sequencing read count abundances and of their associated bias introduced throughout the eDNA metabarcoding workflow have often led researchers to prefer presence-absence indices over abundance-based ones to analyse diversity patterns (Fonseca, 2018; Jurburg et al., 2021; Lamb et al., 2019). But on the other hand, presence-absence data remain highly sensitive to molecular artefacts that often remain in the data even after the most conservative bioinformatics procedures, hence questioning the validity of using MOTUs richness, or richness estimators based on singletons such as Chao1, as well as presence-absence indices as descriptors of α , β or γ diversities (Figure 4.1). Remaining artefacts, by co-occurring with the genuine MOTUs (Frøslev et al., 2017) can further inflate the structuration or correlation with environmental parameters when using presence-absence data compared to abundance-based ones, as we observed in tropical soil bacterial and eukaryotic communities (e.g. Sommeria-Klein et al., 2020). To alleviate this issue, several developments have been made recently to correct eDNA metabarcoding data, either through analysis of co-occurrence patterns (Frøslev et al., 2017), or by correcting presence-absence data from false negatives with multi-species occupancy models (Ficetola et al., 2015). Some of them further include autoregressive terms to facilitate the detection of both false negative and positives (Chen and Ficetola, 2019, reviewed in Burian et al., 2021). While extremely promising, these approaches are still in their infancy, and their efficiency in correcting the biases introduced by the eDNA workflow remains to be assessed across systems and taxa. They also often require additional, high credibility (and often non-eDNA-based) data that may not be available for many groups, typically the neglected biodiversity.

Alternatively, one may consider that abundance data are valid, or, at least as imperfect as presence/absence ones. Indeed, as discussed above, albeit abundance data may be distorted by DNA extraction or PCR

biases, several studies suggest that they provide relatively reliable information on species dominance in terms of biomass and activity. From a more conservative viewpoint, read counts data can also be seen as a probability of MOTUs genuineness, artefacts being in general much rarer than genuine sequences. Such data can be analysed with the Hill number framework (see Equation (5.1) for its initial development, Hill, 1973), which constitutes an interesting perspective not only because it unifies mathematically well-known diversity measures through a unique parameter q (i.e. richness at $q = 0$, the exponential of Shannon entropy at $q \sim 1$ and the inverse of Simpson at $q = 2$), but also because with this q parameter, one can tune the weight given to rare taxa. In other words, this framework allows adjusting the level of confidence one may have in rare MOTUs ($q = 0$ being the highest confidence level).

$${}^qD = \left(\sum_{i=1}^S p_i^q \right)^{1/(1-q)} \quad (5.1)$$

Using this framework, we compared plant diversity patterns derived from traditional vs. eDNA-based observations along an altitudinal gradient. We have shown that robust inferences on plant α diversity and associated patterns with eDNA are possible at $q = 1$ (Calderón-Sanou et al., 2020). On the other hand, we found that β diversity patterns are more robust. This is consistent with earlier observations I made on the β patterns of marine bacteria, retrieved with abundance data, presence/absence data, or abundance data from which rarer species are removed: all these are often very similar (Zinger, Amaral-Zettler, et al., 2011; Zinger et al., 2014). The Hill number framework also now presents the advantage of encompassing phylogenetic, functional and network diversity indices (reviewed in Chao et al., 2014, Ohlmann et al., 2019). At $q > 0$, both α and β diversity estimates are also often independent of sequencing depth (Calderón-Sanou et al., 2020; Ibarbalz et al., 2019; Modin et al., 2020). Consequently, Hill numbers at $q > 0$ are now recommended as a more robust way of measuring diversity for both microbial (Haegeman et al., 2013; Lloréns-Rico et al., 2021; Modin et al., 2020) or more generally eDNA data (Alberdi and Gilbert, 2019).

Downstream possible analyses often aim at quantifying community assembly processes, building occurrence networks, capturing spatial/temporal dynamics, etc. They often rely on statistical approaches traditionally used in ecology (e.g. multivariate or distance-based analyses, null models, etc., reviewed in Dray et al., 2012; Ulrich and Gotelli, 2010), mechanistic models (Sloan et al., 2006; van den Berg et al., 2022), and, more recently on new tools based on machine learning algorithms or joint species distribution modelling (reviewed in Hartig et al., 2024; Trego et al., 2022). But given the different limitations discussed above, a conservative approach would be to analyse eDNA metabarcoding data within the limits of what it can tell us, that is to say, essentially on the distribution and α and β diversity of the most abundant taxa, using analytic approaches that make no or few assumptions on the input data and/or that are poorly sensitive to remaining errors and biases. On the other hand, some mechanistic models appear to make very good predictions of microbial taxa abundances, community diversity, and community structure across environments, as observed through eDNA metabarcoding (e.g. Grilli, 2020). I clearly lack the background (and time to gain more of it during these last years) to discuss these aspects further. Certainly, we still need to assess qualitatively and quantitatively the extent to which the different biases and issue I've mentioned earlier would compromise the ecological inference made from these new approaches and models (Hartig et al., 2024; van den Berg et al., 2022) and/or identify what are the critical steps where biases have an impact. I am very curious to dive into these aspects more in the future and see how the field will evolve.

5.6 Conclusion

I've discuss above about the uncertainties we have on, and the biases introduced by, the eDNA metabarcoding approach throughout the entire workflow. All the aforementioned considerations are now synthesized in several reviews, certain of which I have contributed to - or even led (Arribas et al., 2021, 2022; Fountain-Jones et al., 2024; Taberlet et al., 2018; Tedersoo et al., 2022; Zinger, Bonin, et al., 2019; Zinger et al., 2012, 2020). My understanding of the current literature is that we still have a limited understanding on the eDNA spatio-temporal dynamics, of the error/bias generation process, and we lack robust and generalized quantification of how each of these issues can compromise downstream analyses and ecological inferences (though I suspect

that it is the way we collect samples that will be the most determinant). Unfortunately, carrying out this quantification, despite its importance, remains a thankless task that is difficult to fund and value in relation to the effort it requires. In the absence of such knowledge, providing definitive recommendations for eDNA metabarcoding data analysis remain challenging. However, I do believe that we can still find our way through these complexities by adopting a flexible and context-specific approach, by carefully evaluating available tools and approaches at each step of the eDNA metabarcoding workflow (Figure 5.8), while continuing to actively contribute to the development and validation of best practices. By embracing these principles, I believe that we will be able to use more effectively and gain more valuable insights from eDNA-based measures.

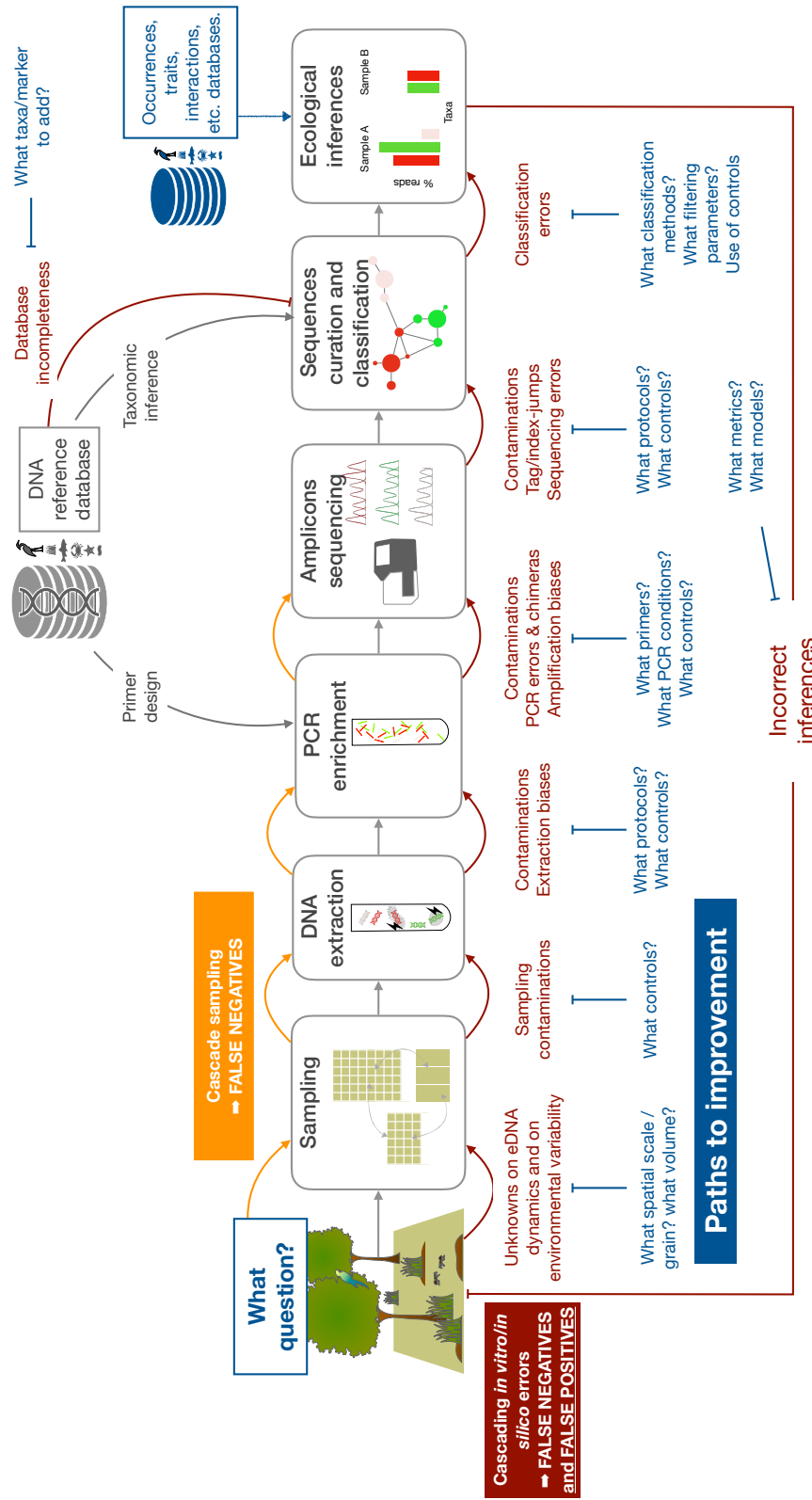


Figure 5.8: Steps, problems, and paths to improvements for a typical DNA metabarcoding experiment. DNA metabarcoding relies on a particular material, i.e. DNA molecules, and is prone to many errors producing false negatives and positives, which can jointly lead to spurious ecological inferences. Minimizing these issues lies in (i) having some *a priori* knowledge of the studied site and organisms targeted, built through e.g. visual census, specimen collection, or building of a local reference database as well as some expectations on local spatio-temporal eDNA dynamics. (ii) incorporating systematically field, extraction, PCR, and tagging-system negative and positive controls throughout the whole workflow, as they will help in identifying potential contaminants and rate of molecular/bioinformatics artefacts. (iii) The use of diversity metrics and models that are relevant for eDNA metabarcoding particularities.

Patterns and processes of the neglected biodiversity

6.1 General context

Community ecology and conservation biology have traditionally placed a strong emphasis on two iconic taxonomic groups: vertebrates and green plants. While past methodological limitations have undeniably skewed our focus towards easily observable organisms, this bias still hinders our understanding of biological systems by neglecting a significant portion of biodiversity. First in terms of global biomass. While green plants are by far major contributors (80%, Bar-On et al., 2018), vertebrates, on the other hand, only represent 0.15% of total biomass. The unbalance is even more visible when considering global biodiversity. Indeed, green plants and vertebrates make up only 20% of described species (Díaz et al., 2019). But this figure is misleading. Indeed, it is only opposed to the number of accepted species of arthropods and fungi, which is itself largely underestimated (Hyde et al., 2020; Stork, 2018). Furthermore, it completely overlooks the vast majority of the tree of life that is essentially made up of protists, archaea, and, above all, bacteria (Burki et al., 2020; Hug et al., 2016). These understudied taxonomic groups collectively represent what I refer to as “neglected biodiversity”, and constitute the largest reservoir of biological diversity on Earth (Figure 6.1).

There is now unequivocal evidence that members of the neglected biodiversity are pivotal for the functioning of ecosystems across both aquatic and terrestrial domains. In the soil environment, both soil microorganisms and fauna can have important impacts on soil erosion, soil aeration, C and N cycles, as well as plant productivity through their role in organic matter shredding and mineralization, etc. (Bardgett and van der Putten, 2014; Lavelle et al., 2006; van der Heijden et al., 2008). Likewise, in the global ocean, plankton communities sustain major biogeochemical cycles (Azam et al., 1983; Falkowski et al., 2008; Worden et al., 2015), being responsible for c.a. 50% of the Earth’s primary production, and constituting the basis of the whole marine food chain. Plankton communities also fuel the global ocean biological pump, a mechanism of C sequestration mediated by the plankton food web that transfers carbon from the surface ocean to the deep ocean (Field et al., 1998; Guidi et al., 2016). In both marine and terrestrial domains, part of the neglected biodiversity is further involved in intimate interactions with plants or animals (e.g. mycorrhizal associations, gut microbiota) with direct consequences on their health (Bardgett and van der Putten, 2014; Trevelline et al., 2019; van der Heijden et al., 2008). There is hence now an increasing awareness of the necessity to better document the neglected biodiversity, not only to better understand and anticipate ecosystem functioning and service delivery in the face of global change (e.g. Cavicchioli et al., 2019; Winding et al., 2020), but also from a conservation perspective (Caldwell et al., 2024; Hochkirch et al., 2021; McQuatters-Gollop et al., 2017; Trevelline et al., 2019).

In the past, I had the opportunity to study neglected biodiversity patterns in different ecosystems from both terrestrial and marine domains. I started first studying microbial communities in alpine tundra soils during my PhD, initially to link them to C fluxes. At that time, we were only starting to explore community ecology related questions in the microbial world, thanks to molecular, and in particular HTS techniques (Green et al., 2004; Horner-Devine et al., 2004; Martiny et al., 2006; Prosser et al., 2007). My thesis hence took a turn

more oriented towards the understanding of the processes involved in the spatial distribution of soil microbial diversity. At the end of my PhD, I was keen to continue addressing such questions at larger spatial scales and to gain more experience in macroecology. I had the chance to find a postdoc in that area, relying on one of the largest eDNA metabarcoding datasets of bacterial diversity available at that time, built by the International Census of Marine Microbes (ICOMM). Back in France, I switched to the study of soil multi-trophic diversity in tropical forests during my last postdoc, to eventually enlarge my scope of research to host-microbiota, marine, and soil multi-taxa systems when I got my position of assistant professor.

This trajectory, marked by transitions between diverse models and spatial scales, may seem like an inventory “à la Prévost” and rather opportunistic¹. Yet, the questions addressed remain the same: (i) do communities of the neglected biodiversity exhibit spatial (and temporal in a lesser extent) patterns? (ii) What are the determinants and underlying processes that can explain these patterns? At first glance, a comparative analysis of contrasting models and organisms surely has the potential to unlock a deeper comprehension of the fundamental mechanisms that shape community dynamics, here of the neglected biodiversity. But I have to acknowledge that it was challenging to find a satisfying and cohesive narrative from this diversity of systems without simply listing off a series of unconnected works. Eventually, it’s the notion of scales that I’ve found to best unify these studies. This resonates with Levin’s seminal paper on the problem of pattern and scale (Levin, 1992), although I clearly don’t presume to match Levin’s depth and theoretical perspective on the topic here.

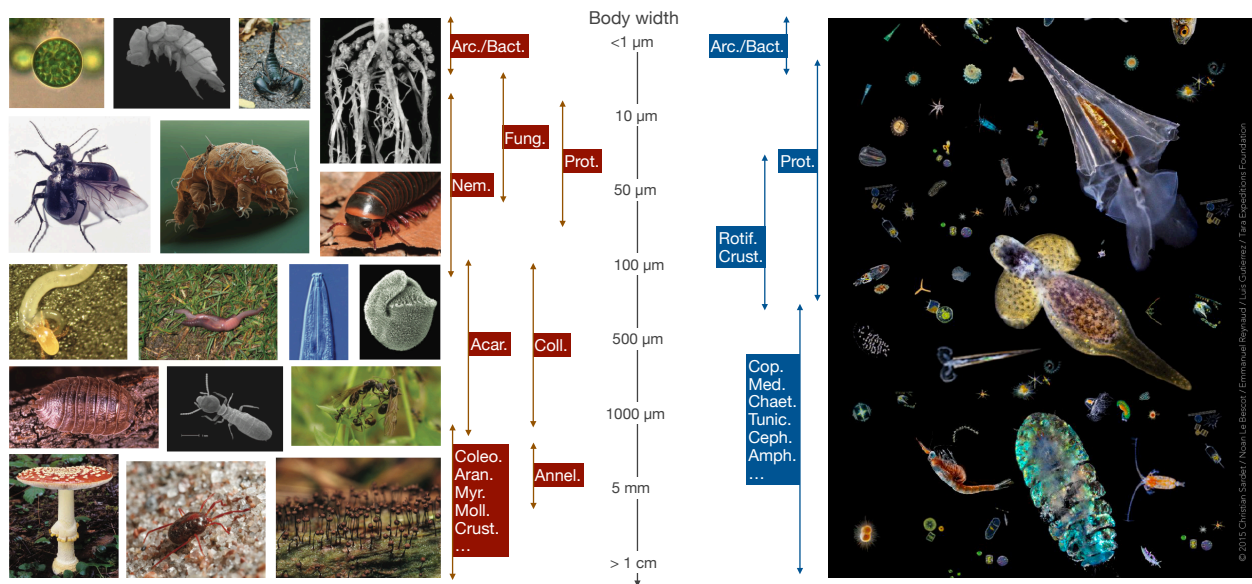


Figure 6.1: Overview of the neglected biodiversity in soil (left) and marine (pelagic, right) environments. Neglected biodiversity includes organisms that are distributed across the Tree of Life and of which body size encompasses several orders of magnitude. Despite their evolutive and ecological importance, I’ve voluntarily excluded viruses from this overview as I’ve considered them only once in my work, and because of their particular status regarding the Tree of Life. Clades abbreviations: Acar.: Acari (mites), Amph.: Amphipoda, Annel.: Annelida (earthworms), Aran.: Aranea (spiders), Arc.: Archaea, Bact.: Bacteria, Ceph.: Cephalopoda, Chaet.: Chaetognaths (arrow worms), Coleo.: Coleoptera (beetles), Coll.: Collembola (spring-tails), Cop.: Copepoda, Crust.: Crustaceans, Fung.: Fungi, Med.: Medusae (jellyfish), Moll.: Mollusca, Myr.: Myriapoda (millipedes), Nem.: Nematoda, Prot.: Protists, Rotif.: Rotifera, Tunic.: Tunicata. Body size distributions are here simplified from Decaëns, 2010 for soil biodiversity, and derived from Karsenti et al., 2011) for marine planktonic biodiversity. Soil biodiversity pictures are from Montanarella et al., 2010. Marine plankton picture credits: © Christian Sardet, Noan Le Bescot, Emmanuel Reynaud, Luis Guierrez, and Tara Expedition Foundation.

¹It is obviously partially opportunistic, the realities of the academic job market necessitating adaptability in a field with increasingly limited permanent positions and funding instruments for project-based post-doctoral fellowships.

It is well-known that patterns of α and β diversity and their underlying drivers do vary across spatial and temporal scales (Averill et al., 2021; Chase et al., 2018; Shinohara et al., 2023, see also Figure 5.6). For example, different global change drivers do not have the same effects on species richness across spatial scales (Chase et al., 2018). To understand how biological communities do assemble, it is therefore necessary to recognise the importance of scales. This is because the key processes of community assembly operate at distinct spatial scales. *Selection* (e.g. biotic interactions) and *drift* (i.e. demographic processes) are primarily local phenomena, while *dispersal* and *speciation* act across boarder spatial scales (Ricklefs, 1987; Vellend, 2010). Furthermore, the spatial or temporal scale at which organisms experience/perceive their environment determines populations or communities patterns across space or time. There are several intrinsic properties of living organisms that modulate this perception, amongst which lies the dispersal capacity and dormancy (Levin, 1992), two key life-history traits often invoked to explain why the spatial scaling and community structure of microbes differ so much from that of macroorganisms (Barberán et al., 2014, section 4.2). However, patterns can also be constrained by the external physical complexity of the area under study, whether it is in terms of heterogeneity/fluctuation in resources and other abiotic parameters, or in terms of connectivity. Finally, the pattern (and perceived underlying processes) also depends on the scale at which we, as external observers, do study a given phenomenon. This interplay between organism perception, external constraints, and scale of our observations underscores the importance of adopting a multi-faceted approach to better understand communities assembly and dynamics. Below, I will hence discuss what are the community assembly processes, in particular the relative contribution of niche-based vs. neutral ones, from small- to large spatial scales for environmental microbial communities, host-microbiota systems, and finally mutli-trophic communities.

6.2 Local spatio-temporal dynamics of soil microbial communities

6.2.1 Processes underlying soil microbial fine-scale spatial distribution

My past work shows that at the local scale (~0.05-12 ha), microbial community spatial turnover and assembly processes already depend on the intrinsic spatial heterogeneity of environmental conditions in the ecosystem. For example the alpine tundra I studied during my PhD exhibit steep topographical and sun exposure gradients over very short distances (i.e. ~ 10 meters). These meso-topographical variations induce gradients in snow-cover dynamics, with upper slope sites being deprived of snow cover almost throughout the year, while lower slope or north exposed sites can keep their snow cover up to late of May. As snow cover dynamics determine the length of the plant growing season, early and late snowmelt sites exhibit also contrasted plant communities, whose composition is dominated by slow-growing vs. fast-growing plant species with contrasted leaf decomposability (Choler, 2005). These differences drive soil properties, resulting in contrasted organic matter content and soil pH between early and late snowmelt locations (Zinger, Shahnava, et al., 2009), but also markedly different bacterial and fungal communities, both in terms of taxonomic and phylogenetic diversity and composition (Shahnava et al., 2012; Zinger, Coissac, et al., 2009; Zinger, Shahnava, et al., 2009).

From my past work, it remains difficult to isolate whether these observed differences resulted from the identity of plant species, soil characteristics *per se*, or plant-related soil parameters, the bacterial and fungal taxa that were over-represented in one condition or the other also had different affinity to recalcitrant organic matter. For example, we have shown experimentally on soil cores from the above-studied locations that these differences can be partly explained by the tannic composition of the litter from the dominant plant species (Baptist et al., 2008). On the other hand, subalpine grasslands harbouring contrasted snow cover dynamics and yet similar plant communities exhibited different soil abiotic conditions and depletion in nitrifying and denitrifying bacteria abundance and activity in sites with thicker winter snowpacks (Jusselme et al., 2016). Finally, while we did not find strong relationships of β diversity patterns between plant communities and soil microbes at small spatial scales (Chalmandrier et al., 2019), we have also shown in several common garden experiments that the functional characteristics of subalpine plant species are also important determinants of soil bacterial and fungal community composition (Bouasria et al., 2012; Mouhamadou et al., 2013). These findings hence suggest that *selection* is a prominent process in the community assembly of microbial communities in mountain soils, mainly driven by an interplay between plant functional characteristics and mediated

effects on soils, as well as of soil properties per se.

However, niche-based processes are not so marked in tropical rainforests at comparable spatial scale (12 ha, Petit plateau plot, French Guiana) for various microbial taxa (bacteria, archaea, fungi, and protists; Zinger, Taberlet, et al., 2019). In these environments and at this spatial scale, local topographical gradients, and more broadly abiotic gradients, tend to be much weaker than those found in the alpine systems discussed above. Accordingly, environmental factors seem to play a less prominent role in shaping the spatial distribution of tree communities. Instead, neutral demographic processes appear to exert a stronger influence (Chave, 2008). It's actually using tree inventories from plots of ~10-50 ha of spatial extent that Hubbell conceptualized and parametrized its first version of the neutral theory model (Hubbell, 2001).

As for trees, that the environmental conditions are more homogeneous does not necessarily mean that microbial communities are spatially homogeneously distributed at this spatial scale (see Figure 6.2a for a selection of abundant taxa). However, only part of this spatial variation can be explained by niche-based processes (Sommeria-Klein et al., 2020; Zinger, Taberlet, et al., 2019), i.e. related to topographical variations, which co-varies strongly with soil pH, widely known as a good predictor of soil microbial communities (Delgado-Baquerizo et al., 2018; Fierer and Jackson, 2006), but also with soil moisture or soil metals content, the latter being notoriously toxic for microbes (Lemire et al., 2013). Likewise, microbial α diversity or β diversity² correlated only very poorly with plant ones within the Petit Plateau plot (Zinger 2019 body). Accordingly, we found weak significant effects of the identity of the local dominant plant (identified through eDNA metabarcoding) on microbial β diversity, and a few significant plant-fungi MOTUs co-occurrences (6.2b). This observation aligns with another study conducted in a similar environment and at a comparable spatial scale (Barberán et al., 2015).

However, it contradicts the predictions of the Janzen-Connell (Connell, 1978; Janzen, 1970), or reverse Janzen-Connell hypotheses (Zahra et al., 2021). These hypotheses propose that host-specific interactions either reduce (in the case of pathogens) or increase (in the case of mutualists) the fitness of tree offspring close to parent tree individuals. Under these hypotheses, a general expectation is that the β diversity of soil microbes should be related to plant ones, as found in other studies (e.g. Peay et al., 2013). I do not exclude that the molecular marker used here lacks taxonomic resolution, which may blur plant-microbial-specific associations. But there are several alternative explanations. First, due to functional redundancy: several microbial mutualists/pathogens taxa may share the same host, even though their own interaction is host-specific. Second, at least for fungal communities, it is possible that common fungal networks, i.e. networks of fungal hyphae that connect the roots of multiple plants individuals and species, are more common in these environments. Another explanation could be that my observations are made at the scale of the microbial community when Janzen-Connell effects (or reverted ones) may arise from a small fraction of soil microbial communities. Finally, the spatial extent considered here may be too small to observe Janzen-Connell effects, the recruitment distance from tree adults exceeding 50 meters for many tree species (Terborgh, 2020).

Besides environmental drivers, we also found that the spatial patterning of microbial communities in the Petit plateau plot may partly arise from low-distance dispersal processes. Indeed, both α and β diversities exhibited spatial autocorrelation that was independent of the spatial structuring of environmental parameters (Zinger, Taberlet, et al., 2019). However, all the explanatory variables, either environmental or spatial ones, explained at best 10% of the variance of soil microbial community composition, consistent with another observation in similar forests of French Guiana (Peguero et al., 2022). To further verify that microbial communities' spatial structuring did not result from unmeasured environmental factors in the Petit plateau, we simulated null communities and compared the resulting β diversity patterns against the observed ones using part of the Stegen et al. (2013) approach. In this framework, when observed pairwise β diversities are higher than null ones, it should reflect *heterogeneous selection* or *drift with dispersal limitation*. When they are lower than null expectations, they should reflect *homogeneous selection* or *homogenizing dispersal*. If they are equal to those expected by chance, they should reflect *ecological drift without dispersal limitation*. We found observed pairwise β diversities to not deviate from null expectations. Consistent with this result, microbial MOTUs had overall quite large distribution ranges across the whole plot (> 200 m on average). This suggests that even though there are some effects of *selection* and *dispersal limitation* on soil microbial communities of

²respectively defined as the exponential of the Shannon index, and pairwise Bray-curtis dissimilarities.

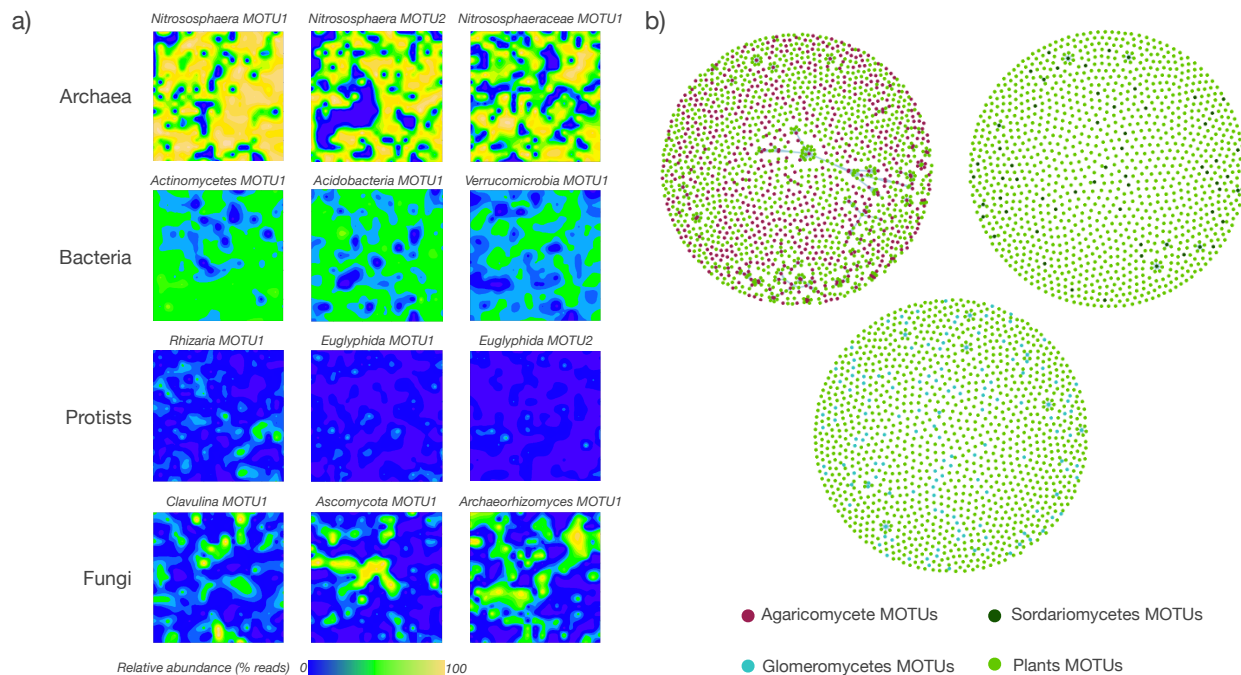


Figure 6.2: Patterns of microbial communities in a 1 ha plot of tropical rainforest. (a) Distribution of few dominant archaeal, bacterial, protist, and fungal MOTUs. Color scales are only comparable within each kingdom. Data were obtained through eDNA metabarcoding of soil cores collected every 5 meters. We used kingdom-specific primers except for protists, where a universal marker targeting all eukaryotes was used (18S rRNA gene). Maps were obtained with a kriging approach based on an exponential variogram model. (b) Co-occurrence patterns between plant and fungal MOTUs in the 1 ha plot. Plants MOTUs, which mainly corresponded to tree taxa, were retrieved for each soil core independently through eDNA metabarcoding using the *trnL*-P6-loop marker. Plant-fungi MOTUs associations were inferred following Faust et al. 2012, with significant associations being those having a FDR-corrected p-value > 0.05. Co-occurrence associations with plants were assessed separately for 3 families of fungi known to interact strongly with trees i.e. Agaricomycetes, a clade that includes saprotrophs, pathogens, or ectomycorrhizal fungi; Sordariomycetes, which also encompasses saprotrophs, pathogens and dark septate fungi; and Glomeromycetes, who correspond to arbuscular mycorrhizal fungi. The 1 ha plot that was analysed here is a subplot from the Petit plateau plot studied in Zinger et al. (2019), the latter have been sampled every 10 meters.

tropical forests, they remain more prominently driven by *ecological drift without dispersal limitation* at this spatial scale.

6.2.2 Dynamics of soil microbial communities

On the temporal axis, I have observed marked temporal dynamics in both alpine and tropical ecosystems, that seem primarily driven by *selection*. In the tropical forest plot discussed above, we have found in a pilot study that soil microbial communities exhibit marked seasonal dynamics between the dry and the wet season (Taberlet et al., 2018; Zinger et al., 2020). The observed changes in microbial community composition reflect the changes in soil moisture between the two seasons. For example, during the wet season, a time at which soils can be waterlogged in tropical forests, we found higher relative abundances of *Deltaproteobacteria*, a group mainly composed of anaerobes (Waite et al., 2020). On the other hand, during the dry season, when tropical rainforests of French Guiana can experience drought (c.a. 21 consecutive dry days on average), the relative abundance of *Actinobacteria* is higher. Members of this phylum have been repeatedly observed to resist drought and enhance plant water and nutrient uptake in water deficit stress (Ebrahimi-Zarandi et al., 2023). During the dry season, we also observe an increase in the relative abundance of *Basidiomycota* fungi, which is consistent with earlier studies reporting the increase in abundance of certain members of this clade after drought treatments (Treseder et al., 2018).

In the alpine ecosystems studied during my PhD, microbial communities from early vs. late snowmelt locations differ throughout the year. First, early snowmelt locations exhibit seasonal dynamics in microbial communities from spring to fall. This contrasts with late snowmelt locations where such variations are less pronounced during the growing season. These differences likely stem from the growth strategies of the dominant plants at each location. Dominant plants of early snowmelt sites produce more recalcitrant litter, which may lead to a more pronounced seasonal pattern in the amount and quality of plant carbon inputs to the soil. Second, we observed drastic changes in microbial communities in both locations in wintertime. In winter, the drop in temperatures comes along a change in soil humidity, nutrient availability and pH regardless of snow cover dynamics (Zinger, Shahnava, et al., 2009). In addition, at that time of the year, soil microbes experience reduced competition for N and other nutrients from plants (reviewed in Schmidt et al., 2007). Accordingly, winter conditions seem to recruit the same bacterial and fungal groups in early and late snowmelt sites across years (Zinger, Shahnava, et al., 2009), with bacterial communities further exhibiting phylogenetic clustering at that time of the year (Shahnava et al., 2012). Winter conditions hence likely act as a strong *selective* event, potentially recruiting psychrophilic microbes able to degrade plant litter, a more complex organic material, that has accumulated in the system at the end of fall (Schmidt et al., 2007).

Soil microbial communities also exhibit inter-annual variations. I first observed this in alpine tundra soils: while winter conditions recruit similar communities across our 2-year sampling, communities sampled during the growing season tend to exhibit more inter-annual differences than seasonal ones. This is consistent with the observations I made later on for the ECOFEED and TransAlp projects (cf. section 5.4.1 and Figure 5.7). Inter-annual fluctuation of climatic conditions across years could explain this pattern. Indeed, for the two latter projects, we identified independently the annual number of freezing days as a relatively good predictor of the observed inter-annual variabilities (articles in preparation). Such inter-annual climatic fluctuations may have direct effects on soil microbial community dynamics, e.g. by recruiting psychrophilic groups at different times of the year, but can also reflect differences in growing season length and/or of the timing of plant phenology. The quantity and quality of plant-mediated C supplies or N dynamics in the soil is known to strongly vary across plant phenology, and this variation likely has an impact on soil microbial community dynamics (Schmidt et al., 2007).

“Priority effects”, i.e. the effect the identity of early colonizing species has on the community dynamics, could be another explanation for the inter-annual dynamics observed in alpine soil microbial communities (Nemergut et al., 2013; Stroud et al., 2024). Indeed, winter microbial populations of alpine tundra soils undergo a drastic crash during snowmelt (Schmidt et al., 2007). This causes a massive release of dissolved N and C, that may recruit different, yet functionally redundant fast-growing microbial taxa from one year to another. Either through niche preemption or through facilitation of later-colonizing taxa, the identity of the early colonizer may determine the successional trajectory of soil microbial communities throughout the rest of the year, with possible consequences on the ecosystem functioning (Fanin et al., 2021; Stroud et al., 2024).

6.3 A microbial perspective from landscapes to the global scale

6.3.1 Landscape ecology and biogeography of soil microbes

My work also shows that *selection* has more importance in the assembly of soil microbial communities as the scale of observation increases through the emergence of steeper environmental gradients. At the landscape scale (c.a. 1-100 km), alpine tundra are characterized by an important turnover of plant communities distributed along meso-topographical gradients (Choler, 2005). During my PhD, we analysed the distribution of alpine tundra soil archaeal, fungal and bacterial communities in eleven types of alpine habitats, dominated by different plant species, distributed across a high-elevation watershed, ranging from 1900 to 2800 meters of altitude and exhibiting steep sun exposure gradient (Zinger, Lejon, et al., 2011, Figure 6.3a). We did so to test whether the landscape-scale structuring observed for plant communities also exists in soil microbial communities, or whether their spatial distribution was driven by other environmental conditions or by dispersal limitation. Using a distance-based redundancy analysis and variation partitioning approach (summarized in Legendre, 2008), we found no pure spatial structuring, hence rejecting the hypothesis of dispersal limitation at this scale for the three microbial domains. On the other hand, plant community composition was an important predictor of archaeal, fungal and bacterial β diversity. However, the influence of plant community composition relative to abiotic predictors differed across microbial domains. It was the best predictor of fungal β diversity. In contrast, bacterial β diversity was better explained by confounding effects of plant community composition and abiotic conditions (typically annual radiations, which drive plant community composition, and soil pH which is influenced by plant functional characteristics as well). Finally, archaeal β diversity was better explained by soil pH. These differences may lie in the different degree of association between plants and microbial taxa belonging to each of these three domains.

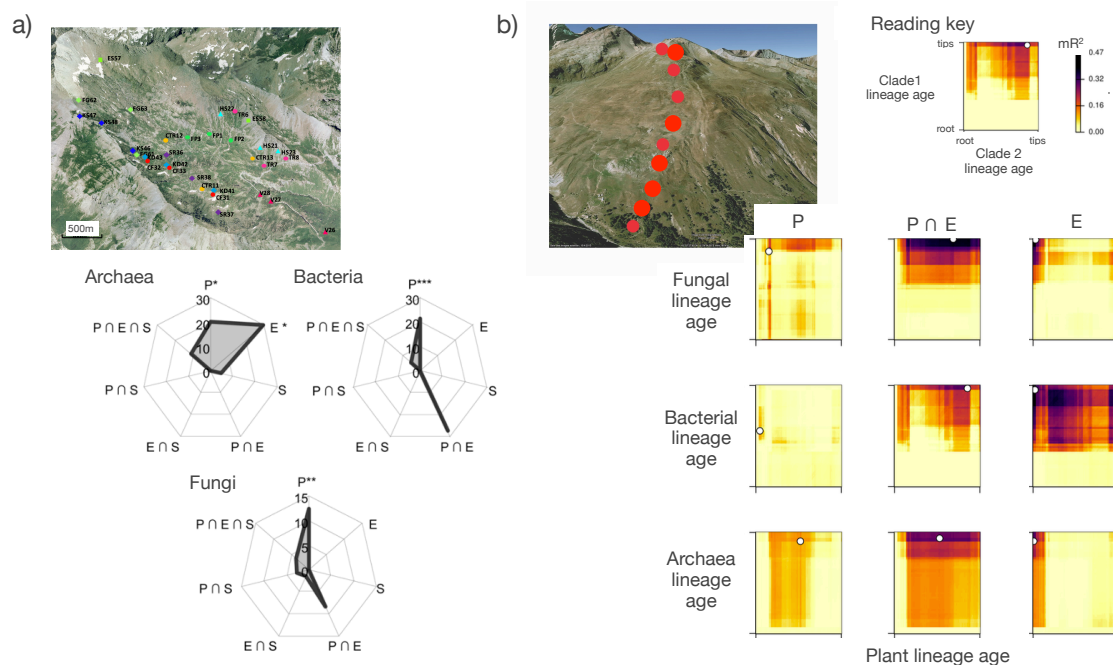


Figure 6.3: Drivers of soil microbial β diversity in alpine landscapes. Both study sites are located near the Col du Galibier. P: pure effect of plant community composition, E: pure effect of abiotic conditions, S: pure effect of geographic distances. The symbol represent shared of the effect between the different predictors. (a) in the Combe Roche Noire watershed (Zinger, Lejon, et al., 2011), where we assessed soil microbial β diversity with CE-SSCP fingerprinting methods (Figure 4.3). (b) in the commune of Valloire (Chalmandrier et al., 2019), where soil microbial β diversity was analyzed through eDNA metabarcoding.

In another study conducted in alpine ecosystems along an elevation gradient ranging from ca. 1800 to 2700

meters of altitude (Chalmandrier et al., 2019, Figure 6.3b), we observed similar differences of co-variation of β diversity patterns between plant and the three microbial domains. However, we went further in this study by quantifying whether the co-variation of soil microbial β diversities with plant ones depended on the phylogenetic grain at which plant or microbial communities are observed. To modulate the phylogenetic grain of observation, we mapped the taxonomic names of plant or microbial MOTUs, as retrieved from soil samples using eDNA metabarcoding, on backbone phylogenies and truncated the resulting trees at different depth. We found that plant β diversity best explained fungal and bacterial ones when plant communities were considered at an intermediate phylogenetic grain corresponding roughly to the family or order level. This suggests that plant–microbe associations result from relatively old plant lineages, most likely exhibiting each particular ecophysiological traits. For example, the elevation gradient studied here exhibit a shift from Poaceae to Cyperaceae dominated grasslands, two families that are colonized by different types of mycorrhizal fungi (i.e. arbuscular mycorrhizal fungi vs. dark septate/ectomycorrhizal fungi respectively, Gardes and Dahlberg, 1996). Likewise, root and leaf N content is rather low within Asterales and high within Fabales (Valverde-Barrantes et al., 2017), the latter group being well known to form associations of N-fixing bacteria and to occur more at lower altitudes. Hence, these results constitute nice examples of the below-ground consequences of above-ground evolutionary history over large temporal scales.

In tropical ecosystems, landscape-scale variations of soil microbial communities are better explained by environmental variables than locally. For example, in French Guiana forests, we found that soil bacterial, protists, and fungal communities do exhibit distance-decay patterns (Figure 6.4). These are better explained by soil environmental parameters (e.g. soil Fe or P content, soil texture). More specifically, microbial β diversities correlation with geographic distances is actually due to the spatial structuring of soil abiotic conditions, except for fungi. This may suggest that soil fungi have dispersal limitation in tropical forests. However, fungi are expected to have good dispersal abilities in general. Plant-fungi associations may be an alternative explanation of the observed correlation of fungal β diversity to both environmental dissimilarities and geographical distances. Indeed, in tropical forests, tree β diversity has been reported to correlate well with geographical distance, suggesting that these communities are driven by dispersal-limited neutral processes (Condit et al., 2002). It is hence possible that the dependency of fungal β diversity to geographic distances may actually reflect niche-based processes, through their trophic interactions with trees. Further analyses comparing more explicitly microbial and plant β diversities will allow further testing this hypothesis.

The landscape scales and above are also relevant to assess land use effects on soil microbial communities. For example in many tropical countries, urban expansion imposes major threats to forest biodiversity (McDonald et al., 2020). Remaining forest fragments within urban landscapes can yet maintain, if not too pressured, native organisms and ecosystem functions. However, the extent to which they differ from undisturbed forests from a microbiological point of view remains poorly known. In a study led by [Julian Donald](#) conducted in French Guiana, we asked whether tropical urban forest fragments of the city of Cayenne harbour levels of microbial taxonomic and functional diversity that are comparable with those of nearby continuous forest (Donald et al., 2021). While α diversity of fungi was similar at the plot scale, γ diversity, i.e. at the scale of the landscape was much lower in urban forests. This could result from dispersal limitation, higher drift, as well as higher abiotic stresses, which may lead to higher species extirpation risks in urban forests fragments. Bacterial communities exhibited contrasted trends, with both higher α and γ diversities in urban forests, which may result from increased soil chemical heterogeneity and availability of labile nutrients and other organic compounds released by different human activities (Decina et al., 2017). We also found that microbial communities composition differed between land uses, with urban forests further exhibiting greater turnover between fragments potentially caused by ecological drift and limited dispersal. On the other hand, urban and undisturbed forests exhibited microbial communities with overall similar functional characteristics such as the ratios of abundances of copiotrophs vs. oligotrophs or of arbuscular vs. ecto- mycorrhizal fungi. This suggests that although urban forest fragments do not seem to conserve the native soil microbial communities found in old-growth forests, they yet allow conserving basic soil functional properties. Note that in this study, we obtained similar conclusions for other components of the neglected diversity (i.e. soil nematodes, earthworms and arthropods) as well as plants, as retrieved with eDNA metabarcoding.

Back to alpine soils, I've been involved at the end of my PhD in a project aiming at studying how land-use changes impact on alpine soil microbial communities across the Alpo-Carpathian ranges (Geremia et al., 2016). Pasture is a very widespread practice across the Alps and Carpathian mountains. This leads to

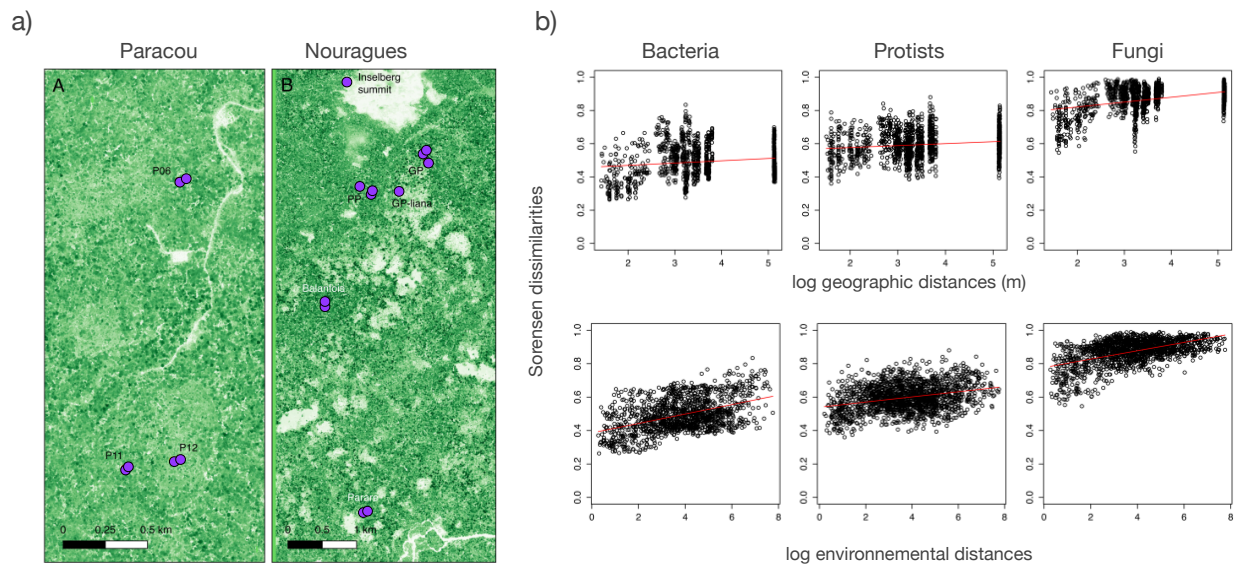


Figure 6.4: Relationship between soil microbial β diversity, geographic and environmental distances in tropical forests. (a) Data were obtained through eDNA metabarcoding of soil cores collected in 19 plots of 1ha at Paracou and the Nourague natural reserve. The markers used corresponded to the 16S rRNA gene for bacteria, the 18S rRNA gene for protists, and the ITS1 for fungi. (b) Environmental dissimilarities correspond to Euclidean distances between samples obtained from a PCA of edaphic parameters. The red line corresponds to the linear regression. All correlations are significant (Mantel tests $\text{pearson}'r > 0.1$, $p < 0.05$), but partialling out environmental distances result in non-significant relationships between soil bacterial and protist β diversity and geographic distances. Figure and associated analyses were made by Guilhem Sommeria-Klein as part of his thesis, unpublished data.

a large and interconnected distribution of grasslands dominated by the subalpine poaceae *Nardus stricta*. In contrast, natural alpine grasslands dominated by *Carex curvula* have a highly fragmented distribution and a long-term separation between populations from the Carpathian and the Alps (Puşcaş et al., 2008). Despite the large scale studied here, we found a signature of biogeographic provinces as defined through plants only for MOTUs statistically strongly associated to *Carex curvula*. This suggests either that even at the scale of the entire Alpo-Carpathian ranges, most microorganisms have unlimited dispersal, or that we lack phylogenetic resolution to observe this. On the other hand, bacterial and fungal communities strongly differed between land uses, consistent with the covariation between floristic and soil microbial community composition discussed above and also reported at the global scale (Prober et al., 2015). Soil microbial communities in *Carex curvula* grasslands were heterogeneous and harbored taxa associated with biotrophic lifestyles, suggesting tighter interactions with the dominant plant. *Nardus stricta* grasslands exhibited lower γ diversity, in agreement with the larger interconnectivity of these grasslands. This comparative analysis suggests a potential homogenization of microbial communities in *Nardus stricta* dominated grassland through the dispersal of microbes by human activities.

6.3.2 Marine microbial macroecology

The field of marine microbiology has deep roots in the pioneering spirit of late 19th-century naturalist and large-scale oceanographic expeditions. Initiatives like the British *HSM Challenger* (1872-1876) and the French *Talisman* and *Travailleur* (1880-1883) expeditions paved the way for this new discipline, fueled by the revolutionary discoveries of Louis Pasteur in microbiology at that time (Adler and Dücker, 2018). The progress made in molecular techniques and the increasing awareness of the critical role of marine microbial lifeforms for the environment over the following century led to several extensive expeditions across multiple ocean basins dedicated to the study of the diversity of microbial life in the global ocean such as *Global Ocean Sampling* (GOS) expedition (2003), the *Malaspina expedition* (2010) or *Tara Oceans expeditions* (2009-on going). These initiatives pre-date their terrestrial counterpart (e.g. Earth microbiome project, soil BON, etc.), explaining why global-scale patterns of microbial communities have been explored earlier in the ocean.

During my postdoc with Alban Ramette and Antje Boetius at the Max Planck Institute for Marine Microbiology, I had the chance to work on one of the largest and most unique eDNA metabarcoding datasets of bacterial diversity available at that time, built by the International Census of Marine Microbes (ICOMM). This dataset indeed gathered hundreds of samples collected in various marine habitats and from the ocean surface to the deep-sea floor (Figure 6.5). We explored how bacterial communities differed between the water column and sediments and whether their β diversity patterns were driven by the same processes (Zinger, Amaral-Zettler, et al., 2011). We found strong differences in community composition that were mainly driven by heterotrophic bacteria exhibiting different terminal electron acceptors, from oxygen in the water column to sulfur or methane in sediments, consistent with the oxygen availability gradient from the sea surface to sediments. We further found that community assembly processes differed, with niche-related factors (i.e. habitat, ocean surface production) having higher relative importance in the structuring of water column bacterial communities when spatial variables better explained bacterial β diversity patterns in sediments. These observations are in line with the contrasted characteristics of these two marine realms: seawater is characterised by strong physical mixing, nutrient state heterogeneity, and harbour diluted and drifting bacteria, more likely to follow the Baas-Becking principle of “everything is everywhere, but, the environment selects”. In contrast, the sea floor is more stable and harbour bacteria with a sessile lifestyle, which should be hence more dispersal-limited. Yet, ocean surface primary production equally contributed to the β diversity patterns of bacterial communities in both realms, reflecting a certain dependency of the ocean seafloor to ocean surface primary productivity (Azam et al., 1983; Jørgensen and Boetius, 2007).

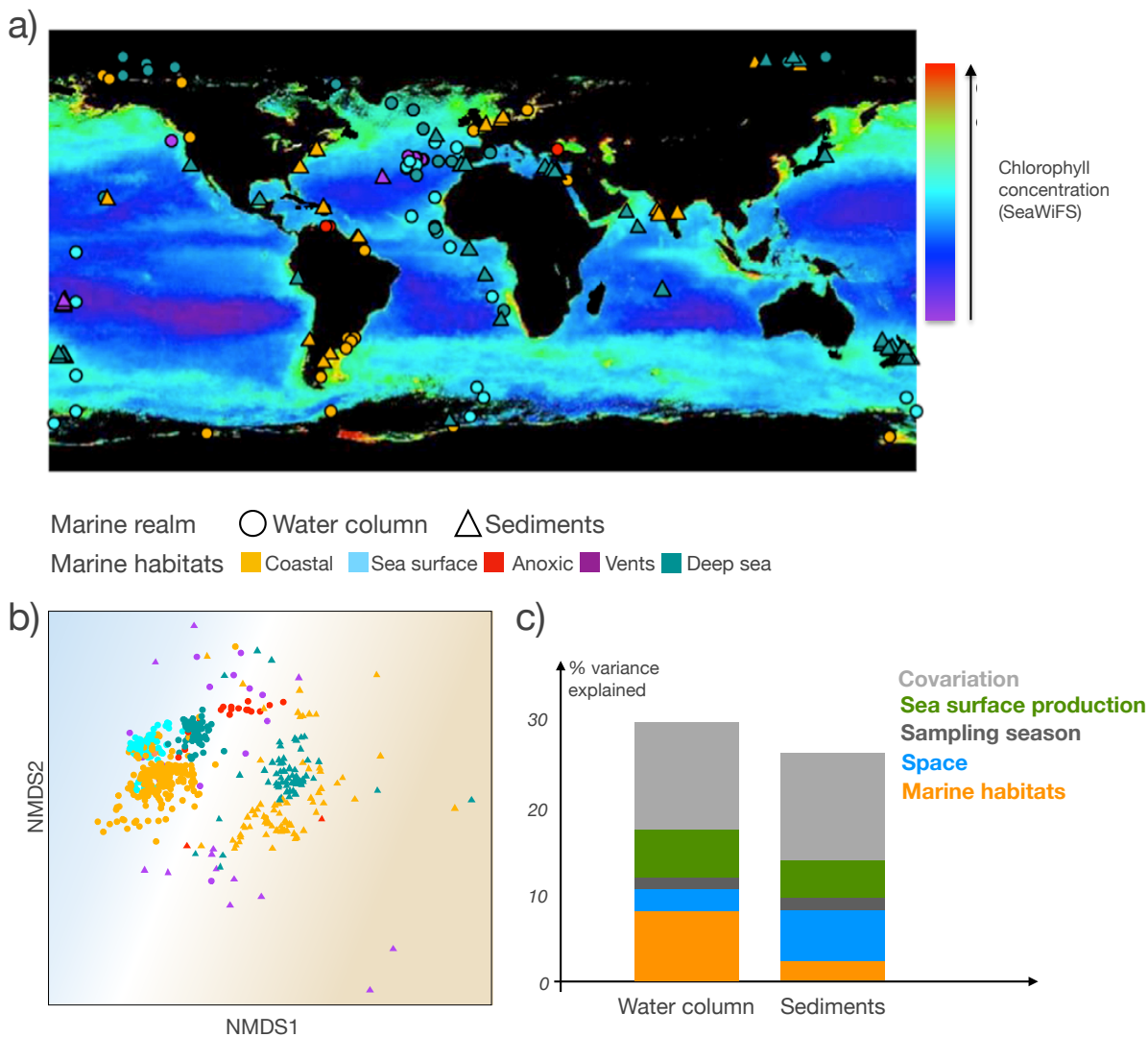


Figure 6.5: Distribution and drivers of bacterial communities in the global ocean. (a) Distribution of ICOMM samples. Data were obtained through eDNA metabarcoding of different samples using the same methodology, the 16S rRNA gene marker, and 454 pyrosequencing. (b) NMDS ordination of the dissimilarity in bacterial community composition. (c) Partitioning of the variation of bacterial composition according to different drivers. From Zinger, Amaral-Zettler, et al., 2011.

We also used the ICOMM dataset to explore if macroecological patterns widely observed in the “macrobial” world were observable in the microbial one, namely the taxa-area relationship (TAR) and the distance-decay relationship (DDR, [Bienhold et al., 2016](#); [Zinger et al., 2014](#), Figure 6.6), which are both indicative of the spatial community turnover rate. We also wanted to assess how different marine habitat ranked in terms of community turnover rate, with the expectation that coastal areas and sediments should exhibit steeper TAR/DDR, due to their higher environmental heterogeneity and because dispersal limitation is likely to be more important in these environments. We also wanted to test TAR and DDR’s slopes estimates robustness to sampling issues, as well as their congruence across habitats because at the time of the study, it was common to derive the slope of the TAR, z , from DDR’s slope β values ([Green et al., 2004](#); [Harte et al., 1999](#)). We found TAR and DDR patterns to be very robust to under-sequencing and quite so to undersampling. As expected, TAR and DDR slopes were steeper in sediments and in coastal habitats. We also found that although TAR and DDR slopes obtained for each marine habitat rank overall similarly. However, we also found that deriving TAR’s slope z from DDR’s slope leads to an important underestimation of microbial community turnover. This result hence provided an empirical support for questioning this approach that was quite popular at that time, hence providing empirical support of earlier theoretical works (e.g. [Woodcock et al., 2006](#)).

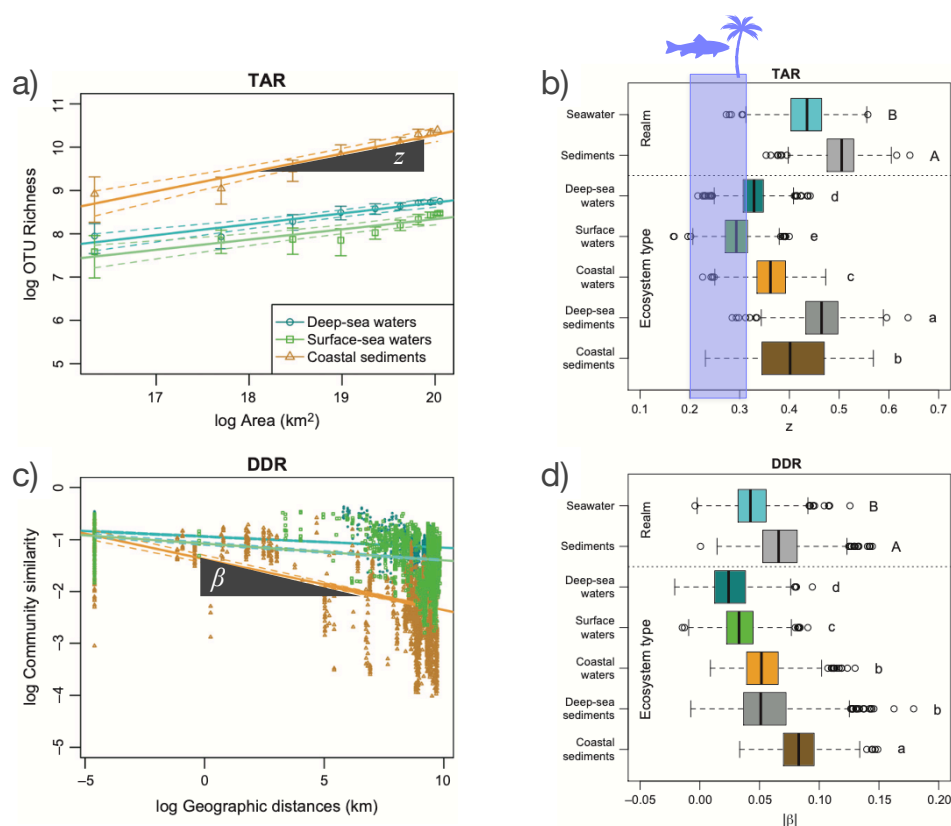


Figure 6.6: Taxa-area and distance-decay relationships of bacterial communities in the global ocean. (a) Taxa-area relationship and (b) its slope z across marine habitats, as assessed following the Arrhenius’ model. (c) Distance decay of similarity and (d) its slopes across marine habitats, as assessed with log-transformed values. Slope values were obtained by randomly resampling 40 samples and 5000 sequences per sample 1000 times in the initial community tables. Upper/lower case letters indicate significant differences (Mann–Whitney tests, Holm-corrected $P < 0.05$) between realms/ecosystem types. From [Zinger et al., 2014](#). Violet polygons indicate the TAR slope values observed in macro-organisms ([Drakare et al., 2006](#)). Equivalent comparisons for DDR are more difficult to draw, as the DDR’s models and dissimilarity metrics used are more variable across studies.

With the ICOMM dataset in hand, it was very tempting to investigate the well-known latitudinal gradient of diversity, i.e. the decrease of diversity from the tropics towards the poles, for marine bacteria. More specifically, we were interested to know whether this macroecological pattern was consistent across the different marine habitats (Figure 6.7a). We were very excited about the results, which confirmed the existence of a bacterial latitudinal gradients in surface waters (Fuhrman et al., 2008), but also indicated the absence of the pattern in the deep sea (i.e. below 200 meters of water depth). Unfortunately, we lacked at that time the contextual parameters necessary to further explain the observed trends, and were concerned by the sampling in sediments and in the deep sea because, which was relatively shallow and biased towards Zealandia. I was therefore very eager to re-explore the latitudinal gradient of diversity nearly a decade later, on the *Tara* Ocean dataset, when I joined Chris Bowler team at IBENS. Following a part of my application proposal, we thus initiated a work, led by our post-doc [Federico Ibarbalz](#), on this topic. The results were quite consistent with my earlier observations (Figure 6.7b): while the latitudinal gradient of diversity is visible in surface waters and pervasive across microbial domains (viruses excluded), it vanishes in the deep sea.

Many hypotheses have been proposed to explain the latitudinal gradient of diversity (reviewed in Dowle et al., 2013) involving energy and area as two main determinants, and proposing different ways these determinants would modulate immigration, speciation, and extinction rates across latitudes. We tested only some of them. Temperature is often thought to be one of the major drivers of this pattern, either through selection of physiologically tolerant taxa, as temperature can impose abiotic constraints on species distribution range, with fewer species tolerating cold conditions or metabolic processes. Indeed, the metabolic theory (reviewed in Brown et al., 2004) posits that higher temperatures, which correspond to a kinetic energy, increase the rate of metabolic reactions, shorten generation times, speed up ecological or physiological processes, ultimately leading to higher mutation and speciation rates. We also tested the “productivity/resources hypothesis”, which proposes that higher resource availability and/or primary production, as it is the case in tropical terrestrial areas, promotes species coexistence by limiting local extinction. Finally, the “environmental stability hypothesis” asserts that unstable environments as it is the case outside the tropics require particular and costly physiological adaptations and would preclude speciation.

The diversity patterns of marine microbial plankton in surface water were very well explained by sea surface temperatures, and to a lesser extent by nutrient availability and of environmental variability (Ibarbalz et al., 2019), consistent with previous work on bacteria (fuhrman2008latitudinal) and larger organisms (Tittensor et al., 2010). We further found that the abundance of phototrophic bacteria (mainly cyanobacteria) declined at cooler higher latitudes, whereas that of phototrophic eukaryotes (mainly diatoms) increased, which may reflect contrasting thermal niches, diatoms generally having larger thermal niche breadths and lower minimal thermal growth than cyanobacteria (Chen, 2015). This may suggest that the latitudinal gradient of diversity partly results from thermal physiological constraints. These constraints are either contemporaneous, or result from past climatic shifts that caused important extinctions. This last mechanism is one of those that is often invoked to explain the existence of Rapoport patterns, i.e. the fact that the latitudinal range of species is on average wider at higher latitudes than at lower ones. It may also hold true for marine bacteria, as we did observe a Rapoport pattern for sea surface bacterial communities (Amend et al., 2013). However, we cannot exclude that temperature has also kinetic effects on e.g. population density, interactions, speciation (Brown, 2014), as suggested by other studies on foraminifera (Allen et al., 2006) and diatoms (Lewitus et al., 2018) and further analyses would be necessary to disentangle these two broad processes.

In contrast to their surface counterpart, deep-sea environments are isolated from sunlight and steep temperature gradients. This particular feature is likely the main underlying reason for the absence of latitudinal gradients of diversity in the deep sea. In addition, ocean carbon export in the deep sea is usually higher at high latitudes (Henson et al., 2012). Higher C supply in high latitudes in the deep sea may hence compensate for the reduction of diversity caused by thermal physiological constraints, in line with the productivity/resources hypothesis. A less plausible explanation lies in the migration of surface species to deep waters through passive vertical fluxes may also contribute to compensate temperature effects. Finally, both sediments and water layers below the photic zone harbour organisms that are primarily heterotrophic or chemolithoautotrophic, and thus highly depend on the availability of organic or inorganic compounds (Danovaro et al., 2016; Woolley et al., 2016). This supports the idea that different channels of energy sustain life across water layers, switching from solar energy or kinetic effects of temperature in epipelagic waters to chemical energy (i.e., carbon- or mineral-based) in the deep sea (Woolley et al., 2016).

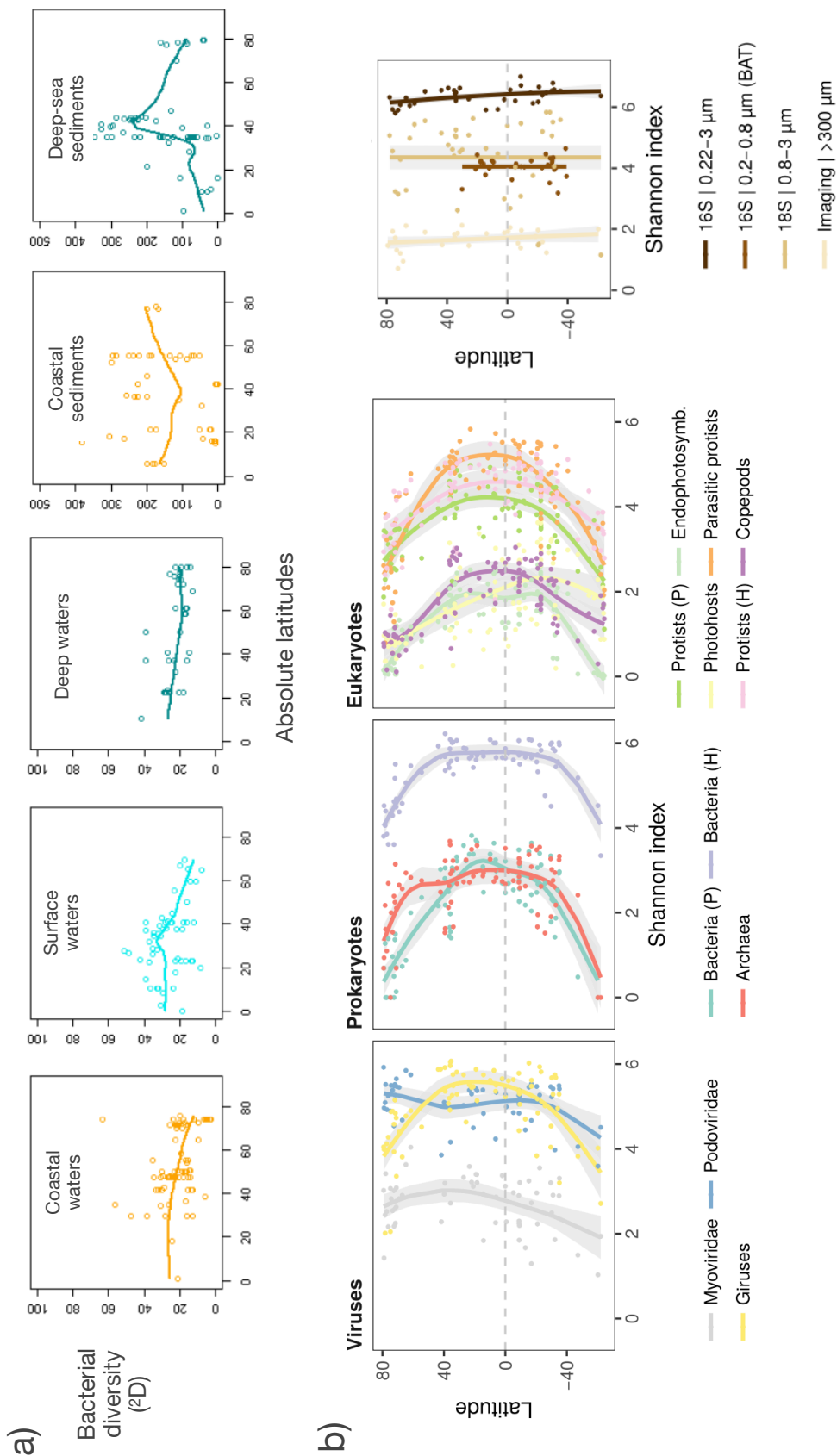


Figure 6.7: Microbial latitudinal gradients in the global ocean. (a) Using the ICOMM dataset, across marine habitats. 2D corresponds to the Hill diversity at $q=2$. Unpublished data. (b) Using the *Tara Ocean* data for surface waters and deep sea waters, from Ibarbalz et al. (2019).

6.4 The particular case of host islands as habitats

The launch of the Human Microbiome Project in 2007 has spurred a renewed wave of interest in the field of symbiosis research, aiming at deepening our understanding of the intimate relationships between macroorganisms and their associated microorganisms (i.e. their microbiota), or, in other words, of the *Holobiont* entity. New technological advances have identified the existence of *phylosymbiosis* patterns, i.e. the fact that close host lineages share more similar bacterial communities (Kohl, 2020) than distant ones. They also have enabled asking more complex questions on how the combined genomes of a holobiont (i.e. the hologenome) collectively interact and/or co-evolve (reviewed in Simon et al., 2019), leading ultimately to global projects such as the Earth Hologenome Initiative, launched in May 2024. Finally, the idea of considering these entities at larger scales has brought the emergence of the concepts of meta-holobionts, similar to that of meta-communities, corresponding to interconnected clonal plants or social insects (Vannier et al., 2019). The holobiont/hologenome concepts as stable selectable entities have been largely debated (reviewed in Simon et al., 2019), but as my research hasn't really addressed this topic nor the effect the host microbiota can have on the host fitness, I prefer to not discuss this particular point in this dissertation. Here, I will deliberately derive from the holobiont concept and will consider microbial communities as the central object of observation, living in host islands, themselves included in larger, more continuous environments, both layers exerting their own influence locally and on the overall microbiota "meta-community" dynamics.

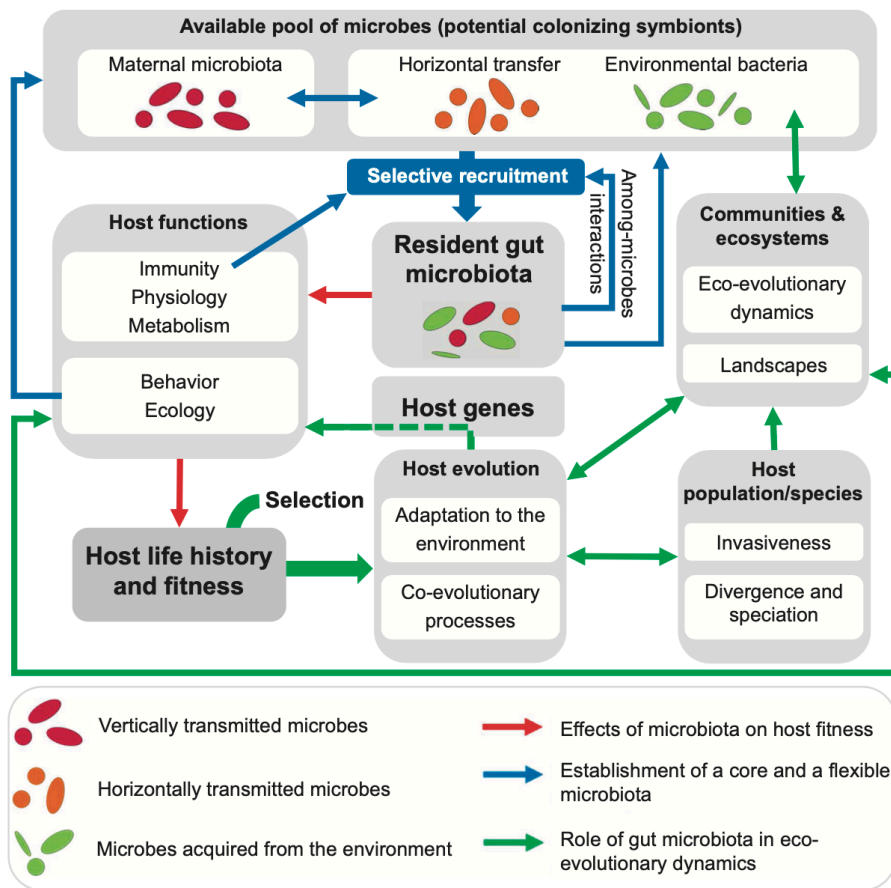


Figure 6.8: Assembly processes of the resident host-associated microbiota. Figure from Make et al. (2017). This diagram is initially intended to describe the community assembly process of the gut microbiota, but can be actually applied to any body part, including to some extent to intracellular microbiota for which both maternal and horizontal transmission has been shown (Chrostek et al., 2017).

Shapira (2016) proposed a framework aiming at better understanding the composition of the host-associated

microbiota, by categorizing microbes based on the tightness of their association with the host. He defined three main classes: (i) the *core*, maternally transmitted, microbiota, (ii) the *flexible* microbiota, rather adapted to host conditions and acquired through horizontal transmission and exchanged with the environment, and (iii) the *environmental* microbiota. As for environmental microbial communities, these fractions are expected to be shaped by the fundamental processes of *selection*, *dispersal*, *speciation*, and *drift*, with the hosts acting as an additional source of - , and filtering layer on - , the local environmental microbial pool (Figure 6.8, Kohl, 2020; Macke et al., 2017). Note though that the flexible and environmental microbiota are not necessarily selectively recruited by the host, even if they can have beneficial or detrimental effects on the host's fitness, and hence contribute too to phylosymbiosis patterns (Groussin et al., 2017; Macke et al., 2017; Mazel et al., 2018). However, the relative contribution of these different fractions on the whole community, and hence of associated community assembly processes, remain unclear. They likely vary across host species, body parts or environments, which may explain why the phylosymbiosis pattern is not consistently found in nature (reviewed in Mazel et al., 2018).

During the past decade, I have tackled related questions in both plants and animals and in both external and internal body parts. The plant and animal microbiota that are involved in nutrition acquisition, i.e. in the root and the gut microbiota respectively, exhibit almost no compositional overlap (Hacquard et al., 2015). The reasons for such large differences remain to be identified, but have so far been suggested to arise from major differences in the way the initial colonization of the host is done, i.e. an prominence of horizontal acquisition from the environment for plants vs. a prominence of vertical acquisition for animals. The large differences in microbiota composition between plants and animals can also result from host habitat specificities, plant root and animal gut environments largely differing in e.g. oxygen availability, organic carbon availability, pH, etc. Such large differences likely influence the relative contribution of host *selection* in the two systems as well. In addition, the filtering effect of the host is expected to be more pronounced in internal body parts such as the gut or root tissues, compared to external ones like the skin or the rhizosphere, the latter having more physical connectivity with the surrounding environment. Accordingly, the phylosymbiosis signal have been found to be stronger in internal body compartment, likely resulting from a stronger filtering effect of the host (Mazel et al., 2018). The collective results of my different studies related to this topic tend to illustrate these expectations, as I will discuss (but not formally quantify) below.

6.4.1 Plant-associated microbiota

The effect of plant community composition on the soil microbiota reported earlier in this manuscript can arise from plant host recruitment of its associated microbiota from the local bacterial taxa pool, either through plant-microbes specific mutualistic or antagonistic interactions or through a modification of soil abiotic conditions through plant particular litter or root exudate or architecture characteristics in the plant vicinity. This question can be partially answered by comparing rhizospheric microbial communities of cushion or tussock plants to those in the surrounding bare soils. In sparsely vegetated mountainous areas, such plants indeed act as ecosystem engineers as most of the soil organic C below these plants directly comes from their own litter and root activity. We conducted such an analysis for the cushion plant *Silene acaulis*, which grows on bedrocks that are either calcareous (with a soil pH rather neutral) or siliceous (i.e rather acidic soils, Roy et al., 2013). While bare soil microbial communities largely differ between calcareous and siliceous bedrocks due to strong discrepancies in soil pH, we found that the plant presence tends to recruit similar bacterial and fungal communities in both bedrocks by buffering the soil pH and increasing local nutrient availability.

A subsequent study where I was not involved further showed that not only did the fungal communities from *Silene acaulis* rhizosphere differ from the surrounding bare soil, but also that these communities differ between two distinct genotypes of the host, one occurring only on calcareous bedrock and another occurring both on siliceous and calcareous bedrocks. These differences of microbiota were mainly driven by biotrophic and pathogenic fungal MOTUs (Roy et al., 2018). These observations are consistent with those made in agricultural plants (Trivedi et al., 2020), but also in the *Carex curvula* rhizospheric microbiota (section 6.3.1, Geremia et al., 2016), for which both bacterial and fungal MOTUs significantly associated with the plant species differed across biogeographic regions. Altogether, these results suggest that indeed, the plant host exerts strong filtering indirectly, i.e. by modifying their environment and recruiting free-living decomposers, but also likely more directly through plant–fungi specific interactions. Further work on these alpine plant

models will enable determining whether these specific associations result from co-evolution *per se* or from ecological filtering driven by particular host traits that would diverge between populations and would selectively recruit adapted microbes to these traits that are yet locally different.

This figure might not hold when looking at the fine-scale distribution of the microbiota fraction corresponding to plant-associated, mycelia-forming fungi. This is the case for example for the fungal communities associated with tropical epiphytic orchids. In such constrained systems, we found that tree barks are colonized by mycorrhizal fungi that are shared among several co-occurring epiphytic orchid species. This suggests the existence of epiphytic common mycorrhizal networks among orchids of a single tree (Petrolli et al., 2022), fungal structures that are suggested to play a crucial role in facilitating the exchange of nutrients and other beneficial compounds amongst different plant individuals and promote plant species coexistence (Bever et al., 2010). I could not observe this in the alpine tundra ecosystems mentioned above because the designs were not adapted for that. In the case of the *Silene acaulis* and *Carex curvula* studies, the rhizosphere of only one species was collected. In the others, soils rather than root samples, were collected, precluding assessing whether some particular MOTUs are detected across spatially close, yet different plant species. In both cases, the spatial sampling coverage was too shallow at the local scale. Yet, I suspect common mycorrhizal networks, or more broadly common fungal networks to be non-negligible drivers of plant communities distribution and diversity in the rather harsh conditions that characterize alpine environments.

More locally than in alpine environments, we assessed the bacterial and fungal endo- and epi- phyllosphere of tropical trees (Schimann et al., 2023) along a vertical, micro-climatic gradient; the temperature, air moisture and light conditions being markedly different from the top of tree crowns to the understory in tropical forests. In such a setting, we expected the endophyllosphere to be more responsive than the epiphyllosphere to the tree host identity and of their respective leaf traits, and less so to the microclimatic gradient, which should better predict the β diversity of the epiphyllosphere. These expectations were met for bacterial communities. By contrast, we observed a great deal of variability of the fungal phyllosphere between tree individuals, and only a very small part of this variation could be explained by the different parameters we considered (i.e. spatial structures related to tree location, tree host species identity and leaf chemistry, and microclimatic gradient). Yet, null models analysis revealed that fungal β diversity patterns largely deviated from null expectations, suggesting (i) the action of *selection* processes from either unmeasured parameters, related e.g. to plant individual physiology or volatile organic compounds, or resulting from priority effects, or (ii) dispersal limitation at a scale that we could not detect with multivariate analyses with our sampling design (Trivedi et al., 2020). In addition, fungal epi- and endo-phyllosphere behaved similarly, contrasting with the expectations that the host filtering is stronger in internal compartments. This may be due to the spatial distribution of fungal mycelium on the leaf, as it can colonize both the epi- and endosphere (e.g. through stomates) as well as to the ability of foliar fungi to switch their trophic niche from green to decaying leaves (Vacher et al., 2016).

6.4.2 Animal-associated microbiota

In animals, I had the opportunity to work on both invertebrates and vertebrates. The thesis of [Caroline Birer](#), who was a pharmacist by training, initially intended to use ant microbial symbionts as models for drug discovery. Invertebrates, in particular ants, have a variety of defensive mechanisms against entomopathogens, and there is growing evidence that one of them lies in the bacteria that colonize their cuticle. These indeed can protect their host from natural fungal or other enemies by producing antimicrobial molecules (Kaltenpoth, 2009), but also by occupying a niche space that may preclude the cuticle colonization by other, potentially harmful, microbes. Due to delays in her chemical experiments, we reoriented part of Caroline's thesis objectives towards the study of community assembly processes of the ant cuticular microbiota (Birer, 2017). In particular, we asked what was the relative contribution of the host *selection* effects relative to horizontal acquisition from the environmental bacterial species pool. To this end, we studied the cuticular microbiota of two co-existing, yet phylogenetically distant ant species living in ant garden nests systems, and compared them against the nest environmental microbiota (Figure 6.9a-b).

We found the nest identity to be the best predictor of the composition of microbial communities on both ant cuticular and nest material (Figure 6.9c), but also of the ant metabolic profiles, as measured by Liquid chromatography – mass spectrometry (Birer et al., 2020). Albeit we lack the data to confirm this, we

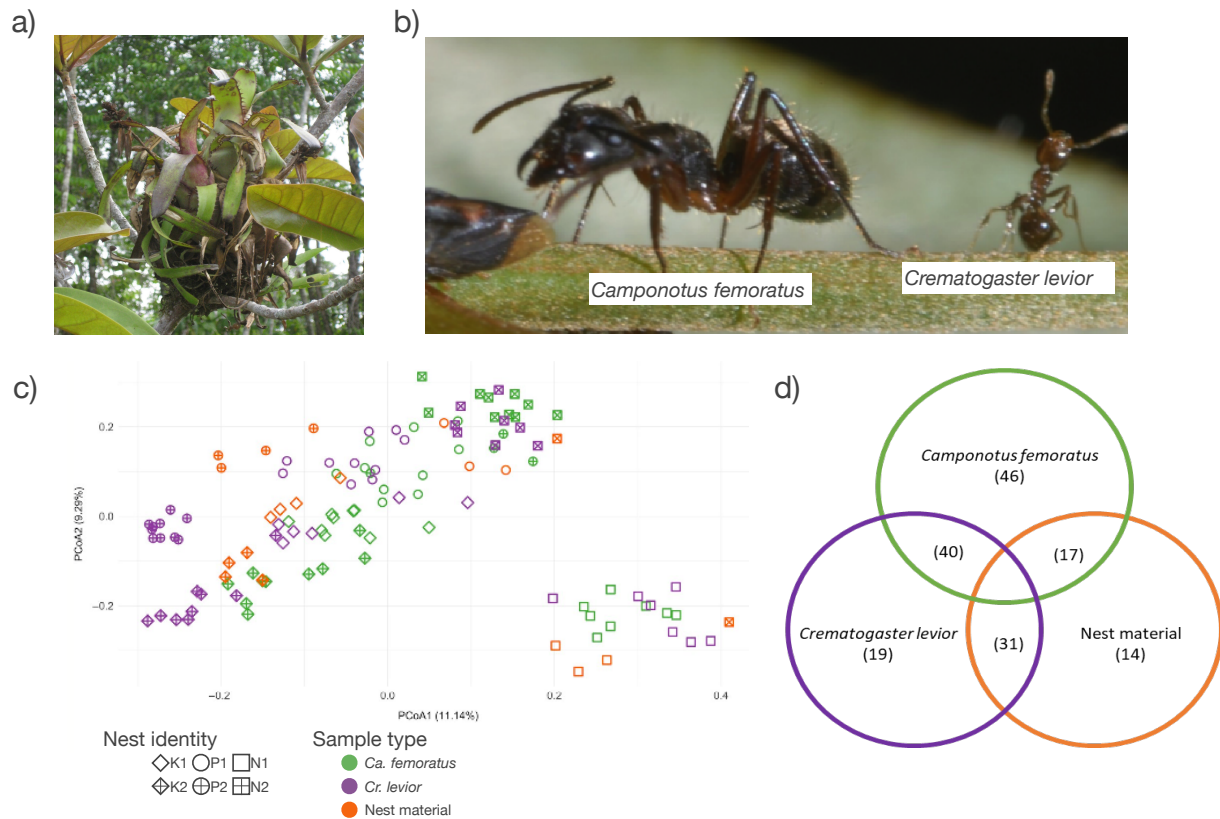


Figure 6.9: Drivers of the cuticular microbiota in ant garden nests from French Guiana. (a) Photography of an ant garden, photo credits Caroline Birer. (b) Picture of the two studied ant species, photo credits Alain Dejean. (c) Ordination representing the similarities of ant cuticular microbial communities across nests and host species vs. nest material. K, P, N correspond to different sampling locations in French Guiana, namely Montagne de Kaw (K), Petit Saut (P), and the Nouragues research field station (N). Two independent nests were sampled in each location. (d) Number of MOTUs statistically associated with either one of the host species or the nest material, as well as shared fractions across all nests sampled, as identified through an Indic species analysis (De Cáceres et al., 2010). (c-d) are from Birer et al. (2020).

suspect the nest specificity of the signal to be linked to the identity of the epiphytic plants making up the nest. Nevertheless, few MOTUs were consistently associated with one of the ant species and are likely acquired through maternal or social interaction with con-specifics (Figure 6.9d). These included MOTUs assigned to the *Streptomyces* genus, known for producing antibacterial metabolites used in many symbiotic interactions (Kaltenpoth, 2009), as well as to *Wolbachia*. Members of this genus are notorious arthropod bacterial endosymbionts acquired vertically. They play a crucial role in arthropods reproduction and evolution and are further highly prevalent in neotropical arthropods (Werren et al., 1995). This study hence exemplifies that even if the host can selectively recruit particular bacterial taxa (of maternal, conspecific, or environmental origin), most of the body surface microbiota seem prominently acquired passively from the environment.

The important contribution of the environmental bacterial species pool to the body surface microbiota as observed above does not always hold true. For example, in two populations of the great tit, we found that the feather microbiota is less diversified and distinct from that found in the nests (Jacob et al., 2018). This suggests a strong selection by the host. This selection can be simply due to the nature of the feather habitat itself, which is depleted in organic matter. This may preclude the colonization of taxa usually associated with plant litter material, such as *Actinobacteria*. Accordingly, we found this clade to be largely under-represented in the feather microbiota. Alternatively, these differences might be due to antimicrobial molecules secreted by the host. For example, uropygial oily secretions might act as a physical barrier by limiting the adsorption of many bacteria on the feather surface. However, we do not know whether this physical effect would be general or might exclude particular taxa, as we observed here. We also identified particular chemicals secreted by the uropygial gland of which abundance correlated negatively with feather bacterial richness and positively with the difference between nest vs. feather diversity. While the apparent negative impact of these chemicals seems to be pervasive across bacterial phyla, it remains difficult to conclude about their broad spectrum antimicrobial activity *per se*, as we did not quantify the overall bacterial load in both systems to verify this hypothesis. It appears clear that the host control was higher in that case compared to the ant cuticular microbiota. Additional work would be required to better understand the differences between these two cases.

We are currently asking similar questions to that of above but this time for animal-associated microbiota in internal body parts for the thesis of [Emma Fromm](#) which I co-supervise together with Julien Cote. The exception here is that we are not trying to determine the effect of the host itself, but rather that of the different conditions in which the host evolves and responds to, whether these effects are direct, or mediated by the host. More specifically, Emma's thesis aims to determine the impact of both climate change and habitat fragmentation on the gut microbiota of the common lizard, *Zootoca vivipara*, the extent to which this can affect the host's fitness, as well as to quantify the ecological and evolutionary processes involved in the responses of both the host and its microbiota to warming and habitat fragmentation. This project thus goes a step further than what is done currently in the holobiont research, by framing it in a global change context. The project builds on a ten-year-long experiment conducted in the metatron (Legrand et al., 2012, => [link](#)) a semi-natural setup, composed of mesocosms where both the climate (i.e. present vs. +2°C) and connectivity (isolated vs. connected mesocosms) are manipulated (Figure 6.10a-b).

Using this experimental setup, Julien Cote, Elvire Bestion, and colleagues have already shown that warming has dramatic impacts on lizards' vital characteristics and functions (e.g. body temperature, nutrition), and life history traits (e.g. growth rate, survival, reproduction, dispersal propensity, Bestion, Clobert, and Cote, 2015; Bestion, Teyssier, et al., 2015 amongst other references). We have also already shown in the past that warmer climates can induce in the short-term a decline in lizard's gut microbiota diversity that is pervasive across bacterial phyla (Bestion et al., 2017). We further found that these changes were concomitant with a shift in lizard diet, with lizard individuals collectively consuming higher proportions of predatory arthropods, and individually exhibiting narrower dietary breadths (Bestion et al., 2019).

The first step of Emma's thesis was to explore how the warming effects observed in the short term (<1 year) evolved over longer time scales (3 years of experimental treatments) and further understand the interplay between warming and habitat fragmentation (Fromm et al., 2024). Emma's work confirmed that the negative effect on the gut microbiota persisted, and even tend to increases after 3 years of climatic treatments in isolated mesocosms (Figure 6.10c), in particular through a decrease in *Firmicutes* abundance and diversity. *Firmicutes* are important members of the gut microbiota that have been already shown to respond negatively to warming (Sepulveda and Moeller, 2020). However, the inter-individual variability of the gut microbiota

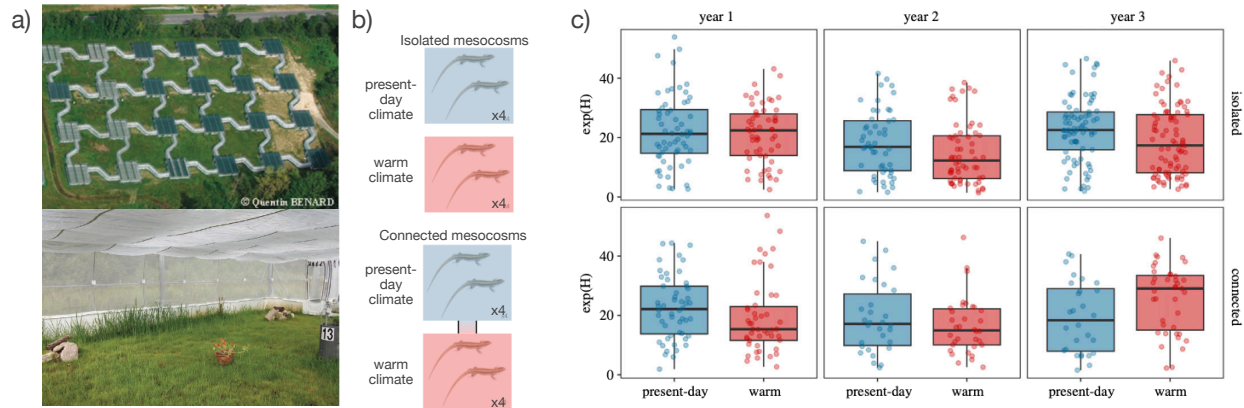


Figure 6.10: Impact of climate warming and habitat fragmentation on lizard gut microbiota. (a) Aerial picture of part of the metatron setup and photography of the inside of a mesocosm. Images from Legrand et al., 2012. (b) Experimental design. Figure from Emma Fromm. (c) Lizards gut diversity, expressed as the exponential of Shannon index ($\exp(H)$) across climatic and connectivity treatments over time. Figure from Fromm et al. (2024). Statistics are described in the paper.

β diversity remains very high, and appears to be prominently (though not fully) stochastically assembled using a null model approach. Accordingly, we could not identify particular members of the community that would be significantly over-represented in one climatic condition or the other. With no clear response of the lizard gut microbiota composition to the climatic treatment, I suspect the reduction in diversity we observe to result from a decrease of the overall microbial load (i.e. community size) in these communities in warmer conditions. Unfortunately, we did not measure this important parameter.

Most importantly, we also found that enabling connectivity between mesocosms subjected to different climates cancels or even reverses the negative effects of warming on lizard gut microbiota α diversity (Figure 6.10c). In other words, *dispersal* balanced the *selective* effect of warming. We knew already that lizard life-history traits responses to warming are buffered in connected habitats through access to more thermal habitats and food resources (Pellerin et al., 2022). From a microbial viewpoint, dispersal generally likely exposes hosts to a greater diversity of environmental and/or gut bacterial species (reviewed in Miller et al., 2018). Here, we lacked of statistical power to identify what mechanisms are involved in this response, but access to more thermal niches will likely allow lizards to avoid at least temporary warmer temperatures and mitigate their negative effects on the gut bacterial diversity. Likewise, climate-induced changes in lizard prey communities or in the vegetation can increase the diversity of the bacterial pool in which dispersing lizard individuals navigate. Finally, host dispersal can also increase the number of sexual and social partners experienced by hosts. To sum up, these results hence shows that preserving habitat connectivity in real world ecosystem can not only protect the host from climate change effects, but also help preserving the diversity of its gut microbiota. A question mark remain on the functional link between lizards gut microbiota and their own condition.

The next steps of Emma's thesis will be to assess the long-term ecological and evolutionary effects of warming (seven years) on the lizard gut microbiota. Diet is a possible ecological driver of the gut microbiota in two ways: (i) by providing the gut microbiota with nutrients that will depend on what the host eats and (ii) by introducing new microbial cells harboured by the preys. We will disentangle these two processes by confronting data of (i) the abundance and composition of available prey species in the mesocosms across climatic treatment (see section 6.5), (ii) the isotopic signature of both lizard and three broad trophic classes of invertebrates found in the mesocosms (detritivores, herbivores, and predators), and (iii) the total microbiota of the invertebrates present in the mesocosms (i.e. potential preys). A preliminary analysis of the latter suggests a large variation of prey microbiota composition across prey species, many of them having no MOTUs at all in common, and that these differences tend to be independent of the prey trophic guild (Figure 6.11a). Instead, the prey microbiota tend to be more consistent between individuals belonging to

the same families (Figure 6.11b). This strongly suggests the existence of a phyllosymbiosis signal in this particular case, where we considered the host’s internal compartment (individuals surface was sterilised). In addition, Emma found that for invertebrates of the same families, their microbiota seemed even more similar when the comparison involved individuals from the same mesocosms (data not shown), hence suggesting some effects of the local bacterial environmental pool. This specific analysis of lizard’s prey microbiota will be the subject of a dedicated article (and a chapter in Emma’s thesis). The long-term dynamics of the lizard gut microbiota together with other information related to lizard life-history traits and evolution will allow us to assess potential evolutionary effects.

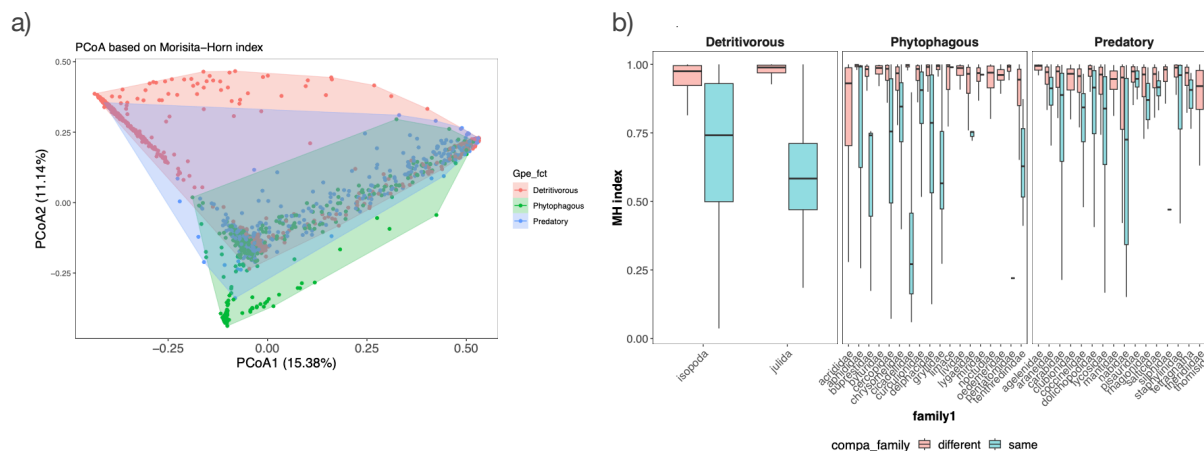


Figure 6.11: Dissimilarity patterns of the total microbiota of lizard’s prey invertebrates in the ECOFEED experiment. (a) Ordination representing the dissimilarity patterns of invertebrate microbiota collected in the metatron mesocosms over experimental years, with dissimilarity being here measured with the Morisita-Horn index. (b) Distribution of the Morisita-Horn dissimilarity index between and within invertebrate families. The analysis comprises in total about 3000 individuals corresponding to about thirty invertebrate families. Figures from Emma Fromm, work in progress.

6.5 Towards an understanding of the multi-trophic neglected biodiversity

Up to recently, holistic descriptions of biological communities, encompassing different kingdoms of life or trophic levels, relied meta-analyses of, when available, fragmentary and often non-standard resources focused on single-taxon communities (e.g. Cottenie, 2005; Drakare et al., 2006; Hillebrand, 2004; Soininen et al., 2007). The transposition of molecular tools from microbial to non-microbial life in the 2010’s has opened new avenues of research by rendering possible the sampling of multi-taxa biological communities with a single approach. These advances offer the possibility to re-examine community assembly processes and major macroecological patterns across multiple taxa, including neglected ones, in the light of their differences in trophic or dispersal mode, general lifestyle, body size, etc, as well as of their inter-taxa interactions, hence enabling to reveal more fundamental principles underlying biodiversity distribution (Shade et al., 2018). More importantly, a multi-trophic perspective can allow better connecting biological communities to ecosystem functioning and stability (reviewed in Eisenhauer et al., 2019). For about seven years now, I’ve explored such multi-trophic systems, mostly in the soil environment through the ANR METABAR (PI Pierre Taberlet) and the ANR GlobNets (PI Wilfried Thuiller), but also in the ocean through the Tara Ocean initiative (current coord: Chris Bowler). Using mostly eDNA data in these studies, the approach I commonly use is to infer a trait/trophic level from the MOTUs taxonomy using a variety of articles or databases gathering such information (e.g. GLOBI, Poelen et al., 2014; FungalTrait, Pölme et al., 2020, etc.).

6.5.1 Spatial patterns of diversity across trophic/taxonomic groups and scales.

In the Petit Plateau plot of tropical forest mentioned earlier (section 6.2), we did not only characterize soil microbial communities but also that of the soil fauna (Zinger, Taberlet, et al., 2019). This work constituted one of the first study of its kind, describing the fine-scale distribution patterns of the whole soil biota (viruses excepted). Contrary to the few trans-kingdom works available so far, who generally analysed separately e.g. bacteria vs. eukaryote communities, my initial expectation was that not all groups would be influenced by the same assembly processes. First, soil organisms can exhibit very various trophic strategies, from chemolithoautotrophy to predation or parasitism. Also, microbial groups, i.e. bacteria, protists and fungi should exhibit different patterns from macro-organisms due to higher dispersal and speciation rates (Barberán et al., 2014; Hanson et al., 2012; Martiny et al., 2006). Beyond this binary comparison, the soil biota exhibit a large range of body sizes. According to the metabolic theory of biodiversity, body size is a major life-history trait that determines organisms' metabolic rate, population size, and diversity (Brown et al., 2004). Body size also correlates with dispersal capacity and dispersal distances: small-bodied organisms have larger population size and have thus more chances to disperse (Finlay, 2002; Martiny et al., 2006), larger active dispersers are expected to disperse over longer distances (Jenkins et al., 2007), while passive dispersers are likely better dispersed through water/air/animal movements when they are small (De Bie et al., 2012; Finlay, 2002). More generally, body size determines the scale at which an organism perceives its environment, which can translate into a scale- or body-size dependency of observed community assembly processes.

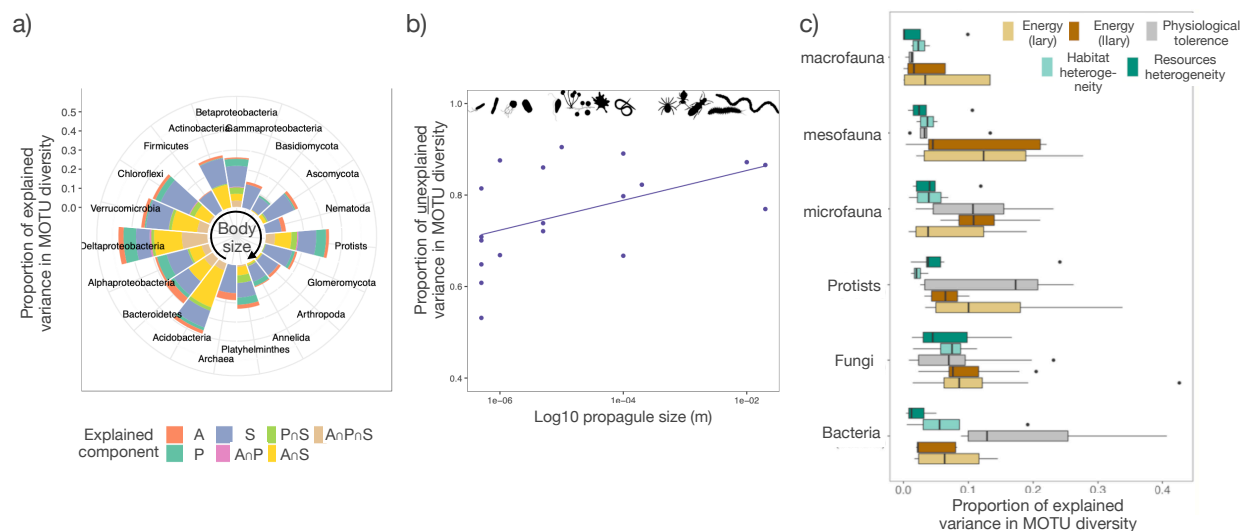


Figure 6.12: α diversity patterns are better explained by environmental factors in small-bodied organisms. (a) Variance in α diversity explained by abiotic factors (A), plant dominant species identity (P) and spatial variables (S) as well as their covariation across clades in the Petit plateau plot of tropical forest studied in Zinger et al. (2019) and (b) linear relationship between remaining unexplained variance and clade average propagule size. (c) Predictors contribution to the variation of α diversity patterns in temperate mountainous soils across organisms of different body sizes. Each hypotheses (colors) include different parameters. Energy (primary): NDVI, Solar radiation. Energy (secondary): organic matter content, C/N. Physiological tolerance: freezing degree days and soil pH. Habitat heterogeneity (clay and bulk density). Resource heterogeneity: prey diversity and plant functional diversity. Figure from Calderon-Sanou et al. (2022).

To test the above hypothesis, I analysed separately major soil phyla whose body sizes are relatively well conserved (in terms of orders of magnitude). I was initially expecting more neutral spatial structuring (i.e. spatial autocorrelation of community composition that is not caused by a spatially structured environment) for larger-bodied organisms, as reported earlier (Cottenie, 2005; De Bie et al., 2012). Within their distribution range, organisms are indeed often reported to exhibit clumped distribution, which can be caused

either by habitat heterogeneity, but also by dispersal limitation as found for tree subpopulations in tropical forests (Hardy and Sonké, 2004). Such clumped neutral distribution patterns have been also reported for the soil mesofauna, who usually displays spatial aggregation below 10 m (Bahram et al., 2016; Berg, 2013). In the Petit plateau plot, the spatial structuring of soil micro- and macro-organisms was overall weak, and most MOTUs had overall large distribution ranges (> 200 m) locally, suggesting unlimited dispersal. This does not exclude the existence of spatial aggregation patterns, but these were unlikely to be observed with the sampling grain used here, which was above 10 m. On the other hand, we found a linear positive relationship between clade average body size and the level of β diversity in general but also of the level of stochasticity in community assembly processes. A similar trend was observed for α diversity patterns (Figure 6.12a-b). This result brings us back to the idea of scale-dependence of community assembly processes: for an observation area of a fixed spatial extent, the organism body size will modulate the scale at which this spatial extent is perceived. In other words, a 12 ha plot may represent a larger area with more environmental heterogeneity from a microbial point of view than from a macrobial one.

According to the above findings and as explained above, we observed no or very low co-variation of β diversity patterns between the studied taxonomic groups in the Petit plateau plot. However, we could have expected a certain level of co-variation, at least between adjacent trophic levels or with plant communities, as a result of species-specific or trait-related biotic interactions (Wardle, 2006). We've actually found that these expectations are theoretically supported using a stochastic model of multi-trophic community assembly (Ohlmann et al., 2018)³. Again, as I explained above, the environmental variation was very low in the Petit plateau plot, and trees there are most likely to be neutrally assembled as well. I expect co-variation of β diversity patterns between trophic levels to emerge when there is sufficient environmental variation in the studied area, allowing a less random distribution of plant diversity. This is what we found in the alpine elevation gradient I presented earlier (section 6.2, Chalmandrier et al., 2019), for which we generated multi-trophic eDNA-based data as well (Ohlmann et al., 2018). For this specific dataset, we found more co-variation of β diversity patterns between trophic groups than with pedoclimatic variables, and with fungal symbionts and saprobes appearing as central in the structuring of β diversity patterns across trophic groups (Ohlmann et al., 2018), most likely through their influence on plant diversity, plant biomass production and plant biomass degradation.

We conducted a similar analysis at a regional scale in mountainous ecosystems. This time, we analysed the spatial patterns of α diversity for groups of organisms belonging to different trophic levels within which we differentiated taxa by their body size (Calderón-Sanou et al., 2022). Similarly to the study I described earlier testing the drivers of the latitudinal gradient of diversity in the ocean (Ibarbalz et al., 2019), we tested several hypotheses that could explain α diversity spatial distribution. First, we tested effects energy availability, as more energy (thermic, solar or in terms of organic matter) would lead to higher diversity through increasing of speciation rates and/or of population size, the latter limiting local extinction. Second, we tested effects of physiological tolerance, which as I explained above (section 6.3.2), but this time by considering frost days and soil pH as potential stressors. Finally, we tested the effect of soil structural microheterogeneity or of resources heterogeneity (in terms of prey diversity or plant functional diversity), as both would lead to higher species coexistence.

Although we observed a great deal of variability in the best predictors of the α diversity of the focal soil trophic group considered, we found that factors related to either energy or physiological tolerance hypotheses were often amongst the most important ones (Figure 6.12c). Note though as at smaller scales, the variance in α diversity patterns explained by all tested factors was again greater for small-bodied organisms (i.e. bacteria, protists, and microfauna), bringing back on the discussion of body size effect on the scale of observation. The fact that the α diversity of preys/plants did not emerge as important predictors of that of higher trophic levels somehow contradicts what I discuss above for β diversity (Ohlmann et al., 2018). On the other hand, it is consistent with previous meta-analyses (Prober et al., 2015; Wardle, 2006). Scale-dependency of the processes shaping soil α diversity has been suggested as a possible explanation, as biotic interactions are

³The theoretical model used here is inspired by the Trophic Theory of Island Biogeography (Gravel et al., 2011), which assumes that a given species will colonize a local community only if its preys are present, and will go extinct if its preys do so too. So in reality, this model is not totally stochastic in the sense that predator-prey interactions are species-specific, although the level of diet breadth of each species can be modulated. We actually tested this, but I will not go further on this aspect, as this specific part of the work was developed by other colleagues.

expected to be more influential at the local spatial scale, while larger scale patterns may rather result from dispersal limitation, habitat heterogeneity and/or plant-community level differences rather than individual plant species effects (Ettema and Wardle, 2002; Wardle, 2006).

In the global ocean, [Federico Ibarbalz](#) and I also compared α diversity patterns across various marine planktonic groups exhibiting differences in trophic strategies and very broad body sizes (Figure 6.7b, Ibarbalz et al., 2019). We found that the form of the latitudinal gradient of diversity seemed consistent across groups exhibiting similar broad trophic modes, i.e. phototrophic versus heterotrophic/chemotrophic organisms. This is in line with the hypothesis that contrasting strategies in energy acquisition and processing should result in different forms of the latitudinal gradient of diversity (Hillebrand, 2004). However, we could not relate body size classes neither to particular forms of the latitudinal gradient of diversity nor to the way α diversity patterns were explained by different environmental predictors, in contrast with earlier expectations (Hillebrand, 2004) and what I discuss above. I do not exclude this contradicting result to be genuine, but at present, I fail to find a rational, biological explanation. On the other hand, the absence of relationship with body size may result from our artificial grouping of marine planktonic groups, which was more based on taxonomy and trophic mode than on body size *per se*. Some of the groups we considered indeed exhibited large within-group ranges of body size/mass that overlap the between-group variation (Hatton et al., 2021). This later explanation is more likely, as a later study on the same dataset where I was not involved has shown body-size dependency of both α and β diversity patterns and underlying processes (Sommeria-Klein et al., 2021), with small-bodied organisms being more influenced by environmental factors while larger-bodied organisms being more structured by ocean basins, hence suggesting dispersal limitation.

6.5.2 Linking multitrophic communities, ecosystem functions/services and global-change?

A key benefit of understanding how biological communities assemble is the ability to better predict their responses to global change and the potential impacts on ecosystems in order to develop more effective conservation / ecosystem management strategies. I will illustrate with a selection of studies how analysing communities from a multi-trophic perspective allows us to get better insights into communities and ecosystems responses to climate change's direct or indirect effects on natural ecosystems. These studies are the products of projects I've been strongly involved, and include the ANR Globnets (PI Wilfried Thuiller), to which I've importantly contributed in terms of project writing, field work, and analysis; the *Tara* Ocean project, for a study I co-supervised with Chris Bowler, and the ERC ECOFEED led by Julien Cote, where I've been also strongly involved since the project writing. Note that I will not discuss the matter of scales for this specific section, as I lack material to do so compared to the previous sections.

A first example of global change effect is the increase of invasions resulting from climate change. In northern Fennoscandia, subarctic birch forests have experienced moth outbreaks of unprecedented scale and severity in recent decades due to climate change. By attacking the foliage of the dominant plants, this invasion has led to a persistent change in vegetation structure, from birch forests associated with understory dwarf shrubs, to grass-dominated systems associated with high tree mortality (Figure 6.13a). We investigated how this invasion impacted soil biodiversity by reconstructing broad food webs from eDNA metabarcoding data obtained from soils from undamaged and defoliated birch forests (Calderón-Sanou et al., 2020). Our main hypothesis was that changes in vegetation and basal resources, driven by moth invasion, would have bottom-up effects, cascading up to higher trophic compartments. Accordingly, we found that defoliated forests exhibited drastic changes in soil food web overall structure, with a decrease in saprobes and ectomycorrhiza fungi relative abundance, an increase in bacterivores, omnivores, and predators and a general increase in soil food web diversity and complexity. The observed changes are consistent with Wardle et al. (2004) framework, where energy use is expected to differ when shifting from an unfertile to a fertile system. Unfertile systems typically rely on fungal decomposition and are energetically more conservative, while fertile ones exhibit faster energy flow across trophic levels through bacterial organic matter decomposition and mineralization. In the same line, we found similar transitions in soil food web structure and composition when comparing mountainous forests versus grasslands habitats (Calderón-Sanou et al., 2024), which can give some glimpse into how soil multi-trophic diversity and potential associated energetic fluxes would change under a scenario of forest conversion to pasture or, at the opposite, of an upward shift of the treeline in high mountains with

global warming.

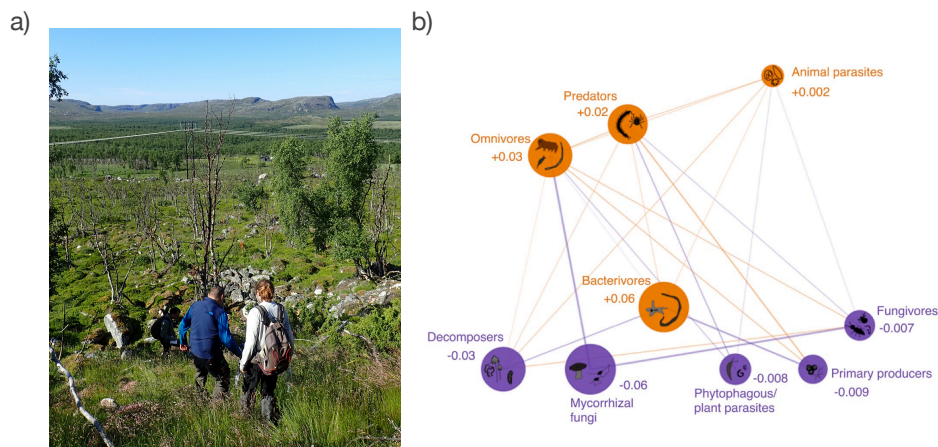


Figure 6.13: Impact of moth outbreaks across trophic levels in Northern Fennoscandia birch forests. (a) Fieldwork in an impacted birch forest in the first plan with Ludovic Gielly. A non-impacted forest is visible in the background. Photo credit Heidi Schimann. (b) Differences in structure of broad soil food web, with orange nodes representing trophic groups that increased in relative abundance, and purple ones groups that decreased in relative abundance from undamaged to defoliated forests. Nodes are distributed vertically according to their trophic level. The size of the nodes is proportional to the value of change in relative abundances and are indicated with numbers. Links width is proportional to the probability of interaction between two classes given the links between their respective trophic groups and the relative abundances of these groups. Figure from Calderon-Sanou et al. (2022).

Still related to climate change, our multi-trophic analysis of planktonic communities in the global ocean, led by [Federico](#), not only revealed their macroecological patterns and their underlying drivers, but also provided insights into how the α diversity of different planktonic groups would respond to climate change and what would be the potential impacts of these changes (Ibarbalz et al., 2019). Indeed, the statistical models built to identify the determinants of planktonic α diversity had very good explanatory power for several marine planktonic groups with only two predictors: sea surface temperatures and chlorophyll *a* concentration, a proxy of primary production. Owing to earth system coupled models (here CMIP5⁴, Bopp et al., 2013), these two parameters can be simulated globally at a relatively fine grain resolution (100km x 100km) in both recent/present and future climate under different scenarios. We hence used sea surface temperature and chlorophyll *a* concentration simulated by CMIP5 at the beginning and end of the 21st century under a scenario of severe climate warming (RCP 8.5) as an input of our planktonic α diversity models. That way, we were able to map current and future α diversity patterns across marine planktonic groups and identify where the largest diversity changes are to be expected (Figure 6.14a).

We identified high-latitude oceanic regions, in particular the Arctic Ocean, to be those that will experience the most drastic changes, hence supporting earlier reports on the tropicalisation of temperate diversity in several planktonic and fish taxa. In particular, high-latitude oceanic regions should experience strong increase in both *Cyanobacteria* and *Copepoda* diversity. The consequences of these changes remain to be determined but are likely to have cascading effects. Indeed, in warmer polar oceans, bacterial phytoplankton may outcompete local, larger-bodied diatoms (Chen, 2015) and result in a reduced local productivity and energy flow over the marine food web (Ullah et al., 2018). More generally, warmer sea surface temperatures are expected to recruit smaller-bodied organisms, according to the Bergmann rule (Sommer et al., 2017) or the metabolic theory (Brose et al., 2012). Although we did not perform a formal analysis, the *Tara* Ocean data tend to exhibit such a Bergmann pattern: the relative abundance of both diatoms and copepods are higher in samples

⁴Stands for Coupled Model Intercomparison Project (CMIP) Phase 5. CMIP is an international climate modelling project, designed to provide scientists a database of coupled global model simulations result from different climate models

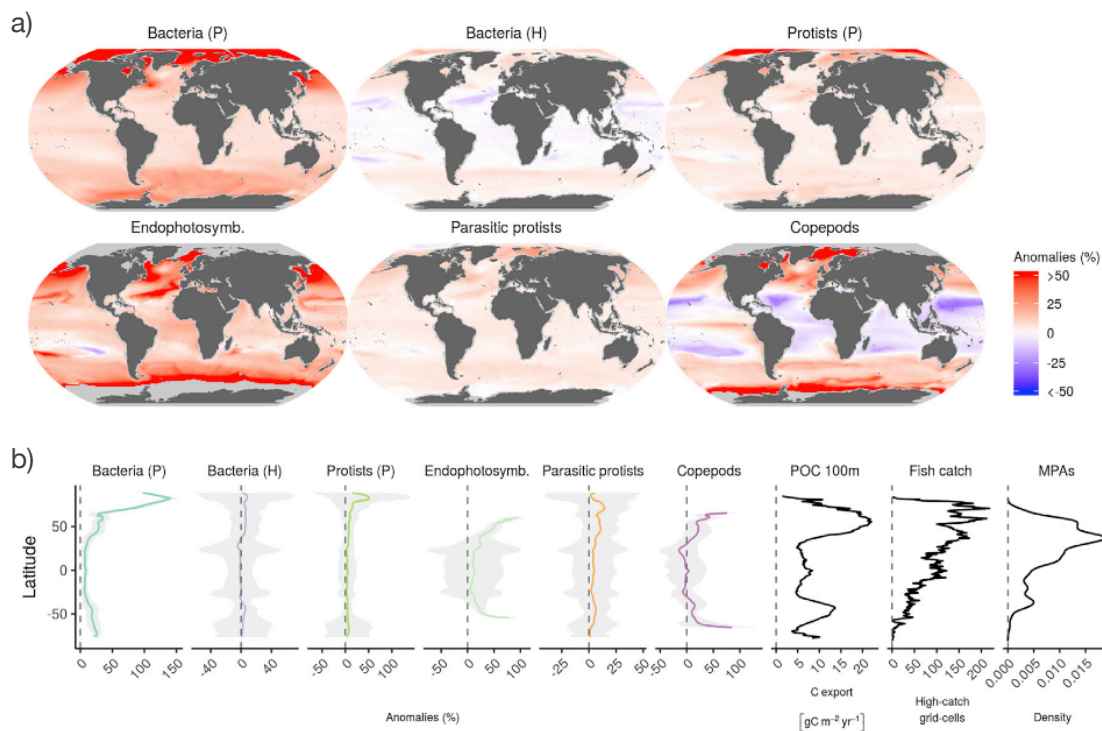


Figure 6.14: Predicted changes in marine plankton diversity in the surface ocean by the end of the 21st century. (a) Maps of predicted changes in marine plankton group diversity between the beginning and end of the 21st century, represented only for marine planktonic groups for which diversity models explained more than >60% of deviance. Copepods, photosynthetic protists, parasitic protists, and endophotosymbiont diversities were measured through eDNA metabarcoding (18S rRNA gene). Diversity of bacteria was assessed with metagenomics. (b) Average values from (a) across latitudes and their uncertainties. The last three panels represent latitude averages for particulate organic carbon (POC) export at 100 m, the number of grid cells with a high marine fisheries catch (>200 kg km⁻² year⁻¹), and marine protected area (MPA) latitude kernel density plots. Figure from Ibarbalz et al. 2019.

corresponding to large size plankton (water filtration between 180-2000 μm) in polar oceans, and higher in samples corresponding to medium size plankton (water filtration between 20-180 μm) in other, non-polar oceans (data not shown). A general reduction of body size may thus further reduce the energy transfer to higher trophic levels (Beaugrand et al., 2008) in these particular areas.

In addition, we found that the areas where the most important changes are expected do overlap with areas that currently exhibit higher C export to the deep ocean, higher fisheries catch, and higher density of marine protected areas. In other words, these correspond to oceanic regions of very high ecological and socio-economic value at present, hence raising questions about the fate of these ecosystem services in the near future as well. Of course, the projection exercise we did here has a number of limitations that we largely discuss and acknowledge in the paper: we indeed used a correlative approach, and we did not consider either adaptation/evolutionary processes or the matter of ocean acidification. Yet, this work remains in my view a necessary preliminary step to draw new hypotheses about the fate of multi-trophic planktonic diversity in the global ocean and of associated potential functional and socio-economic consequences.

Last but not least, I would like to mention a more recent work in preparation I co-supervise together with Julien Cote, as part of his ERC ECOFEED. The work is led by [Léa Beaumelle](#), postdoc on the project, and aims to understand the impact of climate change on above- and belowground biodiversity. Due to the complexity of climate change impacts and the diverse direct and indirect pathways, including biotic interactions, through which it can affect both the above and the belowground, the current state of the art on the topic suggests that climate-induced response of the plant-soil system is highly context-dependent (Pugnaire et al., 2019). Building on other studies focusing on aboveground or aquatic systems (e.g. O’Gorman et al., 2023; Rudolf and Roman, 2018), we hypothesized that the context-dependency of climate change impacts on above-belowground systems are, at least partly, modulated by the initial structure of their associated food webs. While trophic network structure and complexity can vary naturally, it is now clear that current human activities are leading to trophic downgrading, i.e. the loss of top predators, with cascading consequences on the whole ecosystem (Estes et al., 2011). Hence, studying the interactive effect of climate and trophic downgrading seemed us as highly relevant objective.

The metatron experimental setting, which we used in the ECOFEED project (Figure 6.10a), includes semi-natural communities composed of complex above and belowground multi-trophic communities subjected to climatic treatments (Figure 6.15a-b). In addition to the experiment I described earlier, we further included mesocosms where the apex predator, *Zootoca vivipara*, is absent. In each mesocosm, we characterized aboveground diversity using traditional approaches, and belowground diversity with eDNA metabarcoding on soil samples. In agreement with our expectation and experimental objectives, we observed a trophic cascade in α diversities across trophic levels in response to apex predator removal in present climate, with an increase of aboveground invertebrates diversity resulting from a release of predation pressure on these communities, and lower diversity of plants and soil fungi as a result of an increasing grazing pressure on plants (Figure 6.15c).

Now studying the effect of the climatic treatment on these two types of food webs, we found contrasted non-significant trends depending on the climatic treatment, more marked for plant communities in absence of apex predator (Figure 6.15d), of which diversity significantly increased. The apparent absence of clear response to the warming treatment actually result from mixed, sometimes inverted responses of certain trophic guilds within the clades analysed in Figure 6.15d. Dissecting these clades into broad trophic guilds, more significant differences did emerge. For example, we found lower omnivore and herbivore α diversity in warmer climates without the top predator, and an reverted trend when the apex predator was present. Likewise, several plant functional groups had opposed α diversity responses to warming depending on the apex predator presence. Belowground, the changes in trends observed in presence or absence of lizards suggest again differences in energy use by the soil food web (Wardle et al., 2004), with higher diversity of fast-growing bacteria, phytoparasites and arbuscular mycorrhizal fungi and soil macrofauna predators in warmed mesocosms when the apex predator is present. In contrast, in the absence of the apex predator, warming increased the diversity of all plant functional groups except graminoids, and so did ectomycorrhizal and saprobic fungi (data not shown, but I will likely do so for the defence).

The observed differences in α diversity patterns across trophic guilds echo again the contrast of ecosystem functioning between fertile vs unfertile depicted by Wardle et al. (2004). This suggests that when the apex predator is present, warming would lead to more fertile productive systems than in present climate, while

when the apex predator is extirpated, warming should make the system less fertile. Léa is currently finishing additional analyses on several key ecosystem functions (i.e. decomposition, productivity, and soil microbial respiration) to confirm (or not) this hypothesis. Note however than here, contrary to what we observed in the global ocean and to theoretical expectations (Brose et al., 2012), we did not see significant changes in food web size structure. At this stage, it is difficult to say if it's a genuine result, a problem in our MOTU trophic/size classification approach, a lack of marker resolution, or the fact that we are considering trophic guilds diversity rather than abundances (note though that some of the results described above are conserved when considering relative abundances instead of diversity).

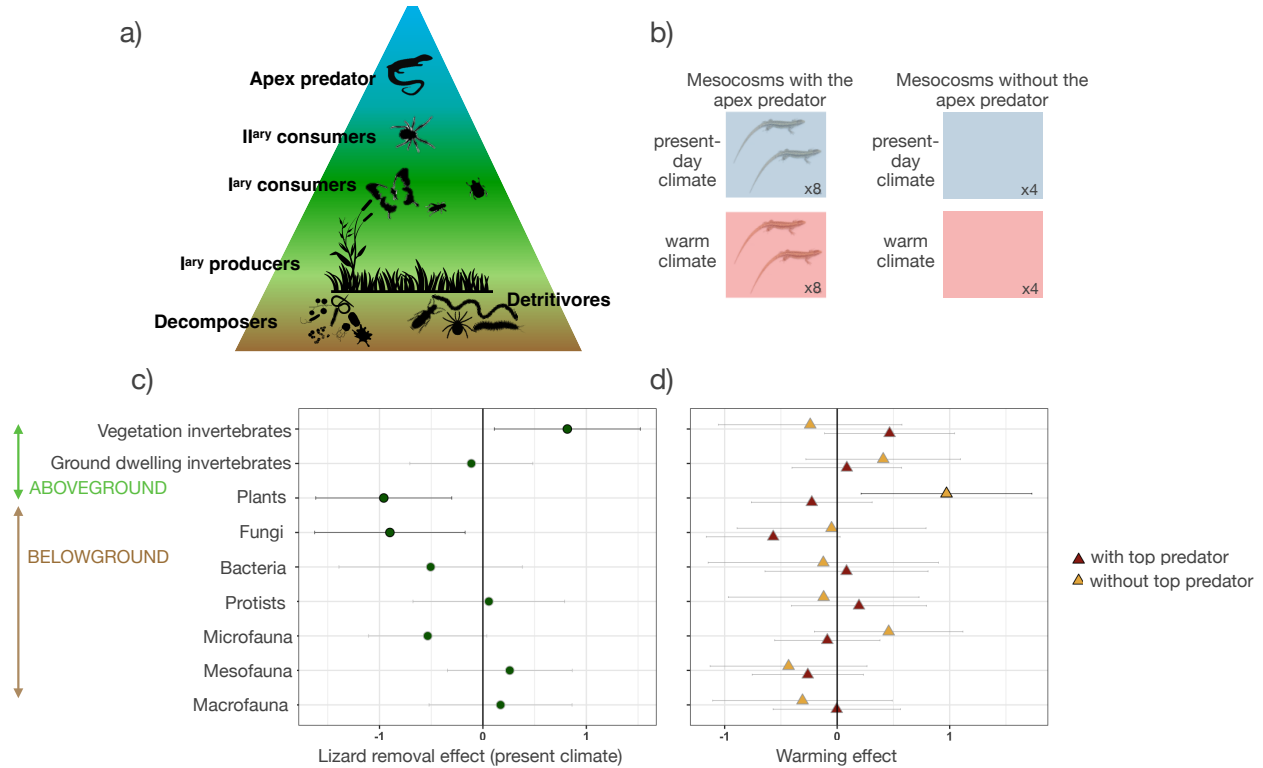


Figure 6.15: Above and belowground diversity responses to trophic downgrading and warming. (a) Overview of the above-belowground food web in the metatron (b) Experimental design depicting warming and apex predator removal treatments. (c) Effect of the apex predator removal on above-belowground diversity across broad taxonomic clades in present climate. Positive values indicate higher diversity when lizards are not present while negative values indicate a lower diversity. (d) Response of the same diversity components to warming in presence (brown triangles) or absence (yellow triangles) of the apex predator. Positive and negative effects indicate respectively a higher, or lower diversity in warmer climate as compared to present climate. In (c) and (d), the effects correspond to standardized coefficients from linear mixed effect models, with zero values indicating no treatment effects. Error bars depict 95% confidence intervals, and black points and lines indicate statistically clear contrasts. (c) and (d) figures and associated analyses by Léa Beaumelle, article in preparation.

6.6 Conclusion

A renewed emphasis has emerged in recent years on the importance of better considering the scale dependence of ecological/evolutionary patterns and processes to better understand biological communities dynamics, their link with ecosystem functioning, and to identify solutions to better manage biodiversity and ecosystems in a changing world (e.g. Chase, 2014; Chase et al., 2018; Estes et al., 2018; Leibold et al., 2022; Lu and Jetz, 2023; Shinohara et al., 2023; Simmons et al., 2021; Viana and Chase, 2019). Reexamining my past work under this

multi-scale perspective gives insights into the context-dependence of results observed in both my research and the broader literature regarding the neglected biodiversity (Figure 6.16). It resolves the niche-neutral debate for environmental microbial communities, consistent with earlier observations for plant communities (Chase, 2014). In addition, considering differences in organism body size explicitly further allows integrating different domains of life in this framework. Finally, it emphasizes the importance of the external physical properties of the system studied, which may influence both environmental heterogeneity and system spatial connectivity (Shinohara et al., 2023). However, this framework intentionally omits temporal scales, for which I have little elements of discussion. It also leaves host-microbiota and multi-trophic systems out of the picture, because at present and with the time constraints I have to send this HDR, I simply fail to replace them in a convincing way. However, I suspect them to relate to an additional scale axis invoked by Levin (Levin, 1992), that is the organizational scales. More formal analyses incorporating additional data (as well as deeper reading of the theoretical literature on the topic on my side) are required to further explore these questions. In this context, eDNA-based methods, despite their limitations (section 5) can capture biodiversity at multiple spatio/temporal scales and with an extremely large taxonomic breadth, and thus undoubtedly hold great promise for addressing the fascinating question of scales in ecology. This is a long-term task that I intend to revisit later, but in the medium term, I will focus on other another thematic described in the following section.

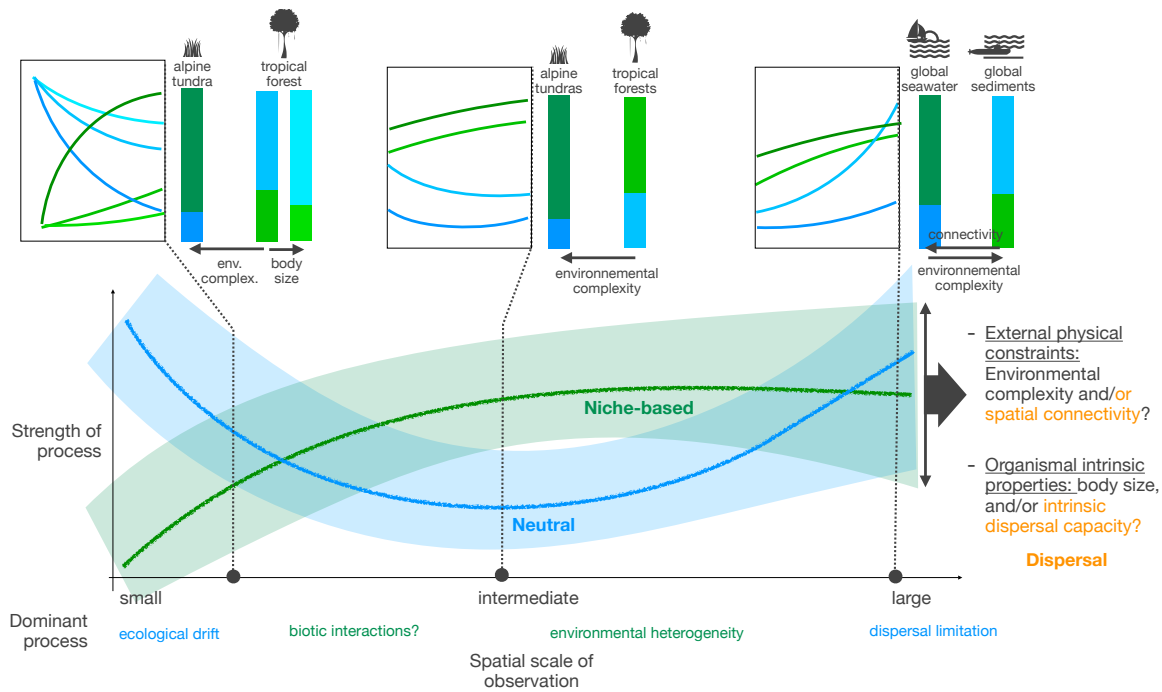


Figure 6.16: Scale dependency of community assembly processes of the neglected biodiversity.

In this scheme, I represented how niche vs. neutral process are generally expected to contribute to community assembly across spatial scales of observation, with illustrations from examples of studies discussed in this manuscript. This comparative analysis allows showing how community assembly processes scaling is itself dependent on external physical properties of the environment, e.g. environmental complexity/heterogeneity, spatial connectivity, as well as intrinsic properties of the organism studied, e.g. body size, which correlates positively with metabolic rate and the scale at which the organisms themselves perceive their environment, and negatively with passive dispersal probability. At small spatial extents, dispersal is not limited. If environmental complexity is high, as in alpine tundra, environmental selection predominates, but if it's low, as in the Petit plateau plot, neutral demographic processes predominate. At intermediate spatial extents, niche-based processes tend to predominate, and will be stronger if environmental heterogeneity is high, and lower if spatial connectivity is high. We also may expect body size to modulate the strength of niche-based processes, as the dispersal probability should decrease with body size at such scales (in case of passive dispersal). This relates to the initial assumptions of the absence or lower spatial patterning of microbial communities as compared to macrobial ones (section 4.2). Finally, at large spatial extents, environmental variation is at its highest, and dispersal probability is reduced, leading to a prevalence of neutral process and a non negligible contribution of niche-based ones. However, in systems that are subjected to strong physical mixing as it is the case in the pelagic realm of the global ocean.

Perspectives of research

7.1 General context

Part of the work presented so far in this dissertation has rather focused on patterns of diversity and community assembly processes. However, I now feel the need to go beyond, as it is already perhaps visible in the last section of my former work synthesis (section 6.5.2). As I mentioned above, through eDNA-based approaches, our ability to examine full, multi-trophic communities from the local to the global scales now offers great opportunities to assess empirically the ecological and evolutionary feedbacks that could occur across spatio-temporal scales and trophic levels, as well as their functional impacts in a changing world.

Ecosystems are indeed threatened by multiple stressors that can each alter directly or indirectly abiotic conditions, biodiversity, as well as the functioning of ecosystems and their contribution to people (e.g. food provisioning, water quality, C storage; FAO, 2019; IPBES, 2019). They can also have cumulative effects across biological compartments (Jaureguiberry et al., 2022; Rillig et al., 2023). Anticipating and mitigating the effect of these stressors is thus a pressing challenge for research and institutional and non-governmental organisations. As such, these last years have seen the emergence of territorial management approaches, political instruments, and global initiatives inspired by the functioning and eco-evolutionary dynamics occurring in natural ecosystems to restore, adapt them, or at least mitigate global change impacts (Moore and Schindler, 2022; Moreno-Mateos et al., 2020).

While there is now compelling evidence that human-induced biodiversity loss have a negative impacts on the provisioning and stability of multiple ecosystem services (Cardinale et al., 2012; Mori et al., 2018; Naeem et al., 2012), the magnitude and pathways of these impacts remain uncertain. This is partly because impact studies have long been focused on single taxonomic groups. As I mention earlier in this dissertation, a more holistic view of ecological communities is needed to better understand the links between biodiversity and ecosystem functions (reviewed in Eisenhauer et al., 2019, Soliveres et al., 2016). In that context, studying above-belowground feedbacks, i.e. the reciprocal effects between the aboveground biota, in particular plants, and the belowground neglected biodiversity, is particularly relevant. Above-belowground feedbacks are indeed increasingly acknowledged to be pivotal in regulating climate, water cycle, above-ground productivity and diversity, and other biogeochemical cycles and appear to be a way forward to mitigating global change impacts (De Long et al., 2019; Eisenhauer et al., 2019; Kardol and Wardle, 2010; Pugnaire et al., 2019; van der Putten et al., 2016).

The interactions between the above- and belowground is also an interesting - yet complex - model to study ecological and evolutionary dynamics. First, both *in natura* observations (section 6.2, 6.4.1) and experimental work (Bailey et al., 2006; Baxendale et al., 2014; Cortois et al., 2016) suggest that plant identity, genetics, and functional traits have an important influence on the above and belowground biota and their abiotic context (e.g. the soil moisture, its chemistry, parasite density, etc.). Likewise, the biotic and abiotic status of the plant-soil system is also influenced by aboveground consumers (van der Putten et al., 2016), as we demonstrate empirically with the ERC ECOFEED (6.5.2), and obviously by the soil biota itself (Bardgett and van der Putten, 2014; Lavelle et al., 2006; van der Heijden et al., 2008). In addition, both ecological and evolutionary processes can occur on relatively short time scales, as many soil organisms have a short

generation time (Bardgett and van der Putten, 2014; Van der Putten et al., 2013; van der Putten et al., 2016). This feature is particularly true for microorganisms, who are also able of horizontal transfers and for which ecological and evolutionary dynamics actually occur simultaneously (Angulo et al., 2022; Martiny et al., 2023).

These different characteristics promote niche construction, i.e. the way an organisms alters its local environment, with either positive or negative consequences for plants at small spatial scales. Janzen-Connell direct/reverse effects (Connell, 1978; Janzen, 1970; Zahra et al., 2021) with aboveground or soil-borne enemies or mutualists are a good example of above-belowground feedbacks that may involve both ecological and evolutionary dynamics (Van der Putten et al., 2013). Another example is the “home-field advantage”, i.e. the idea that a plant associates with specific decomposer communities in the soil and in its phyllosphere in such a way that it accelerates its own litter decomposition (Fanin et al., 2021). At larger spatio-temporal scales, eco-evolutionary dynamics likely occur during range expansion, as suggested for the plant-soil system (Van Nuland et al., 2024), as well as in the co-evolution of traits or lineages (Cornelissen et al., 2023; Van Nuland et al., 2016). These larger scale dynamics can translate into a spatial co-variation of diversity between trophic guilds (Ohlmann et al., 2018), with ecosystem properties (e.g. aboveground biomass, C storage), or in phylosymbiosis patterns (Mazel et al., 2018).

Understanding these processes is particularly needed for tropical forest ecosystems. Indeed, tropical forests, and here I’ll restrict my discussion to those of the Americas, represent an enormous reservoir of biodiversity and resources for local communities and humanity (FAO, 2019; IPBES, 2018). But these ecosystems are increasingly threatened by human activities (e.g. deforestation, gold/oil mining pollution) and the increase of the mean and extremes if both temperatures and drought events, (IPBES, 2018; Lapola et al., 2023). In addition, above-belowground feedbacks remain challenging to investigate in these ecosystems, which are characterised by their remoteness and where the Linnean shortfall is high. Available data on above- and, most importantly, belowground tropical diversity remain scarce and highly fragmentary (Decaëns, 2010; Guerra et al., 2020), leading to large uncertainties on the tropical forests’ above and belowground collective responses to global changes.

Meta-analyses suggest that unsustainable land use and climate change have already caused a decline of plant, mammal and bird taxonomic and functional diversity (Lapola et al., 2023), and could potentially lead to critical ecosystem transitions (e.g. Boulton et al., 2022; Sales et al., 2020). Afforestation and forest restoration are expected to mitigate global change impacts on tropical forests biomass and C storage (Koch and Kaplan, 2022) and may facilitate the recovery of species richness in the mid term, but their composition, and hence initial functions may take centuries to recover (Jakovac et al., 2021; Rozendaal et al., 2019). The response of the soil biota to climate and land use change, albeit translating into compositional turnover, appears to be strongly context dependent and variable across trophic groups (De Long et al., 2019; Le Provost et al., 2021; Pugnaire et al., 2019), although recent reports suggest shifts in energy fluxes at the scale of the whole food web (Potapov et al., 2024; Zhou et al., 2022).

My next objectives are thus to study above-belowground feedbacks in tropical ecosystems in South America and their potential repercussions on ecosystem functioning in a changing world. I intend to develop 2 complementary axes of research combining methodological developments, and *in natura* and experimental studies. More specifically, these axes are the following: (i) improving current eDNA-based tools to better describe and monitor both above- and belowground biodiversity and functions in tropical systems, (ii) studying how soil/plant diversities and several ecosystem functions co-vary across land uses and climatic conditions and dissecting above-belowground feedbacks more experimentally or through longer-term monitoring. For now, I will focus on plant-soil feedbacks, but I expect some work developed in the methodological axis to rapidly produce tools providing more holistic assessments of above and belowground communities in the short- to mid-term. Some of the aspects of these research perspectives have started very recently through the funding of several projects of which I’m PI, co-PI or partner, the most important ones being the junior package TULIP, which I earned in end 2023 and which funds the PhD of [Jérémy Raynaud](#), BARCODRAIN, of which I’m co-PI with Amaia Iribar, and the XPRIZE rainforest competition, to which I participate as part of the Brazilian Team, led by Vinicius Castro-Souza.

7.2 Towards eDNA-based observatories of tropical forests?

This axis of research aims to pursue some of the methodological aspects I described earlier (section 5). The use of eDNA-based tools is now becoming more accessible, HTS capacities are still evolving, therefore opening new possibilities to develop large-scale and long-term monitoring of biodiversity (Thomsen et al., 2024). Such eDNA-based observatories are even more relevant for tropical ecosystems compared to other ecosystems, as in these environments, traditional monitoring (i) involves high costs of labor and field related expenses due to the remoteness and high diversity of these habitats (ii) rely on a dwindling taxonomic expertise, and (iii) remain often incomplete as a large part of the biodiversity remains unknown to science or poorly described, which is especially true for the neglected biodiversity component. However, there are still several challenges that still need to be met before fully exploiting the potential of eDNA in general and for particular case of above- and belowground biodiversity of tropical forest ecosystems.

7.2.1 eDNA-based description from the belowground to the top of the canopy

An important challenge in describing terrestrial tropical biodiversity with eDNA lies in that we are dealing with ecosystems that are intrinsically highly heterogeneous both spatially and vertically (above vs. belowground, vertically stratified vegetation) and where eDNA does not diffuse as in aquatic ecosystems, making the eDNA signal very punctual (Zinger et al., 2020). This is problematic to appropriately sample the soil environment. While I describe earlier several progresses we made in the past on the sampling of soil eDNA, I also identified several caveats related to the spatial and temporal sampling grain (section 5.4), that I intend to investigate more formally in collaboration with Irene Calderon-Sanou with different eDNA metabarcoding datasets that we already obtained from tropical environments through various projects (e.g. METABAR, PI Pierre Taberlet; Globnets, PI Wilfried Thuillier; DIAMONDS, PIs Jérôme Orivel and Jérôme Muriene; BUG and BING, PIs Jérôme Orivel and Heidy Schimann, etc.).

But the most challenging biota to sample currently through eDNA remains the aboveground community. Current eDNA-based approaches to effectively sample vertebrates or plant communities either rely on a destructive approach, i.e. the sampling of bulks of arthropods feeding on these organisms (e.g. Carvalho et al., 2022; Kocher, Gantier, et al., 2017), or on the sampling of local water bodies, which is not invasive, but only possible in presence of a water body. Soil eDNA, on the other hand, contains no or very minute amounts of DNA belonging to plant epiphytes or other aboveground organisms (Zinger et al., 2020). Finally, while eDNA contained in the air has been recently proposed as a promising alternative to detect vertebrates and plant diversity (e.g. Lynggaard et al., 2022), current proofs of concepts were conducted in systems of low diversity and high organism density, which is far from the reality of tropical forests. Accordingly, Pierre Taberlet and colleagues made pilot experiments with air eDNA in the Nouragues Natural Reserve in French Guiana, and while they succeeded in amplifying microbes and some plants, their attempts to detect vertebrates from this matrix, even close to bird nests, were unsuccessful (personal communication).

Rainwashes from forest canopy or that run-off at ground level represent an interesting alternative to the above matrices. Indeed, rainwashes transports and concentrates eDNA originating from tissue renewal, excretions or faeces of organisms living in the canopy or in the understory. Accordingly, both experimental and *in situ* studies have demonstrated that it was possible to detect canopy arthropod species directly from rainwash collected below the canopy (Macher et al., 2023; Valentin et al., 2021). This matrix would thus allow giving access not only to ground animals and plants if sampled at the ground level, but also to the canopy biota, which is assumed to shelter the largest portion of plant and animal diversity in tropical forests (Nakamura et al., 2017). I co-lead together with Amaia Iribar a project called BARCODRAIN, funded by the Office Français de la Biodiversité (OFB), aiming to test the relevance of this environmental matrix and develop associated efficient - and if possible low-cost - protocols of sampling for biomonitoring applications. At present, we only have preliminary results for plant communities of a pilot experiment (see Figure 7.1a-b for the experiment objective and design). We found that the eDNA signal in this matrix can persist up to about a month, which means that we can leave the eDNA collectors for a certain amount of time on the field. This strategy would allow us to reduce the punctuality of the signal recovered, a problem I mention in section 5.4, as well as the need for field personnel to be mobilized too frequently to collect the eDNA collectors (Figure 7.1c-d). We are also able to distinguish quite well different forest types (Figure 7.1a-e). We are currently

producing the data for vertebrates and invertebrate communities for the same experiment and for a second one aiming to identify the sampling effort (i.e. number of collection points) needed to estimate appropriately the plot-scale diversity, and plan to upscale the approach and test it in other ecosystems next year (savannas, white sand forests, *terra firme* forests, and lowland forests).

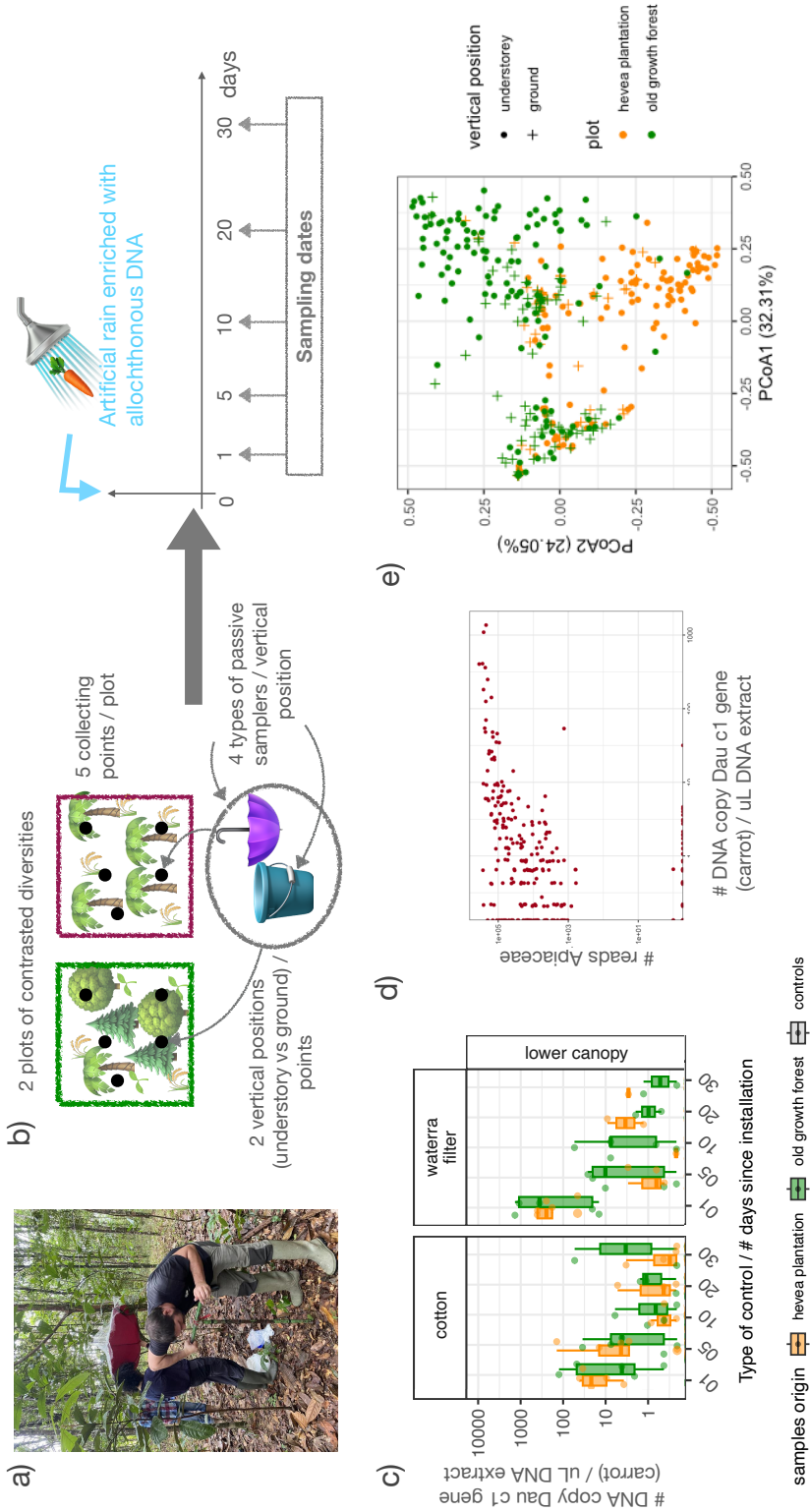


Figure 7.1: Canopy and ground rainwash eDNA to sample aboveground tropical forest biodiversity. (a) Setting up the rainwash canopy and ground collectors at the Paracou station, French Guiana, by Amaia Iribar and Frédéric Petitclerc. Photo credits: Céline Leroy (b) Design of the first BARCODRAIN experiment aiming to assess the persistence and relevance of rainwash eDNA signal in two tropical forests of contrasted diversity. We test passive sampling methods of rainwash eDNA (e.g. water filtration by gravity with water filters, adsorption on different types of substrates such as cotton or ceramic). We pulverised the juice of an exogenous plant at the start of the experiment, here the carrot (*Apiaceae*, a family absent in the studied area), to measure more objectively the decay of rainwash eDNA. (c) Decay of canopy rainwashes eDNA over 30 days in cotton and water filters collectors as assessed through dPCR of *Dau c1*, a single-copy gene often used for detecting carrot allergens. (d) Relationship between the number of copies of *Dau c1* and the number of reads corresponding to *Apiaceae* retrieved through eDNA metabarcoding (portion of the *rbcL* gene). PCR products were not standardized prior library preparation and sequencing to obtain a semi-quantitative signal between samples. Pearson's $r = 0.48$, $p < 0.01$. (e) Compositional dissimilarity of plant communities retrieved with eDNA metabarcoding across forest types and vertical positions, represented in a Principal Coordinates Analysis (PCoA). All collector types are included. The cloud of points located on the left likely corresponds to dysfunctional PCRs and mainly includes samples obtained from two specific types of collectors of poor performances. The composition of the two forest types and retrieved with different types collectors is significantly distinct (PERMANOVA $R^2 = 0.7$ and 0.4 respectively, $p < 0.001$). Unpublished results, manuscript in preparation.

7.2.2 Beyond taxonomic diversity

While eDNA metabarcoding is quite powerful in describing taxonomic diversity, its use in a multi-trophic context presents an important limit: it currently fails to provide at least a semi-quantitative view on the abundance, biomass or activity across trophic levels (see section 5.5.2), especially when several DNA markers are used in parallel. Yet, this information can be useful for identifying a coarse functioning of above-belowground systems (e.g. the fertile vs. unfertile systems described by Wardle et al. 2004). To a certain extent, it may also help, together with additional information on metabolism, to assess more quantitatively the energy fluxes and dynamics of food webs (Barnes et al., 2018). I intend to develop an approach relying on DNA spike-ins introduced at the PCR amplification level to enable absolute quantification of the number of copies of each DNA marker, as developed for host-associated microbiota 16S quantification (Barlow et al., 2020; Lloréns-Rico et al., 2021), but this time in a framework using multiple universal and more clade-specific primer pairs amplified simultaneously in a single PCR assay (Ficetola and Taberlet, 2023). Such PCR multiplexes can be now done with minimum modification on the primers pool with commercial kits containing stabilizing components that facilitate primer binding to their targets even in a non-optimal melting temperature. They can therefore allow us to quantify the number of DNA molecules from various markers within a single PCR reaction. The challenge will thus lie in finding the appropriate balance of the different primer pairs and DNA spike-ins concentrations to maximize the coverage and reliability of the quantification for all target clades while keeping the experiment as standard as possible. If successful, this development will hence provide a more quantitative perspective of multi-trophic communities, even though interpreting the number of copies of a DNA marker will have to be done with caution (section 5.5.2) and cross-validated with alternative measures of biomass or activity (e.g. cell counts, lamina baits, soil respiration, etc.). [Jérémy Raynaud](#) will participate to these developments as part of his PhD.

Another limitation of eDNA metabarcoding to study above-belowground feedbacks is that they lack functional information. In the work I've conducted or to which I've participated so far, we use the MOTU taxonomy to assign them a function (typically a trophic guild) or trait (typically body size) based on expertise, literature or on databases that are fragmented across clades and types of ecosystems. Acknowledging limitations such as intra-specific variation and the fact that some organisms can switch their trophic mode throughout their lifespan, this approach remains valuable and has proven effective in better describing and understanding the assembly processes of multi-trophic communities (e.g. section 6.5.2). I have compiled a more homogeneous database during my maternity leave in 2020, which we used in different studies of the GlobNets or ECOFEED project, but this database still needs more work and be encapsulated in a R package. I am also starting to connect with experts on tropical soil fauna and microflora (e.g. Jérôme Mathieu, Miguel Cooper, and other members of their project SoilFaunaService (Brown et al., 2024), Luca Mendes and Tsai Siu Mui from the Brazilian Center for Nuclear Energy in Agriculture, and all the members of the NEFINEO project led by Mélanie Roy) in order to gain more description on tropical taxa.

Additionally to eDNA metabarcoding approaches, I intend to start working more with metagenomics (Figure 4.3)¹. More specifically, I plan to couple both Illumina and long-read Oxford Nanopore Sequencing to take advantage respectively from their low error rate vs. the possibility of sequencing long fragments (see also more on Oxford Nanopore Sequencing in 7.2.3). That way, we may be able to retrieve ecological and short-to-long-term evolutionary information for multiple taxa (Andújar et al., 2015; Sigsgaard et al., 2020; Tedersoo et al., 2021) and thus better identify eco-evolutionary dynamics from eDNA data.

Using these metagenomes, I'm also considering several approaches based on the functional profiling of samples from metagenomes (e.g. placement in the microbial economics spectrum, Piton et al., 2023, metabolic mapping and reverse ecology, i.e. the inference of microbial interactions from the community-scale metabolism Diener et al., 2020; Levy and Borenstein, 2012). I'm also considering to build a database that allows to trace specific traits or functions that could be involved in above-belowground feedbacks or stress responses

¹I may change of mind later, but for now, I'm not planning to use metatranscriptomics. We encountered many issues of RNA extraction success and protocol standardization during my postdoc with Eric Coissac at LECA. In addition, the short half-life of RNA renders its signal very punctual. While this property is desired for short-term observations or for better understanding particular mechanisms of response to change, it makes it less relevant for larger spatio-temporal and organisational scales of observation, as I plan to do here. Finally, RNA extraction is more difficult to implement outside a dedicated lab, which might be a problem when working on tropical soils. This is not the case for the extraction of DNA, which we were able to perform with minimum equipment in hotel rooms or in garages (see our field protocol in Taberlet et al., 2018).

to land use change or climate from metagenomes. For example, microbial genes involved in organic matter degradation and nutrient mineralization (e.g. carbohydrate-active enzymes) participate to niche construction, by promoting access of plant and belowground organisms to resources, and hence influence their productivity. For plants, several microbial genes can increase their tolerance to climate or pathogens such as those involved in biofilm building, secretion of plant hormones or defence proteins, and that should hence be involved in eco-evolutionary feedbacks between plants and their microbiota (Angulo et al., 2022). Other genes involved in greenhouse gas emissions (e.g. genes involved in N₂O or CH₄ production/regulation, Hallin et al., 2018; Knief, 2019), in stress response (e.g. dormancy/sporulation, exopolysaccharides production), or more broadly of life history strategies (e.g. virulence, genome size, Piton et al., 2023). Note here that this specific part of the work will be more microbial focused due to the dominance of bacterial and fungal DNA in soils. However, I do not exclude exploring possibilities to retrieve further genomics information from other organisms in the obtained metagenomes and, if necessary with enrichment methods in the longer term (e.g. sampling of soil fauna bulks, fragment size selection, single-cell sequencing).

Note though that gathering and formatting both trophic, traits and genes information is a laborious and long-term task that will certainly benefit from both artificial Intelligence tools such as those developed by Nicolas Leguillarme on soil biodiversity (Le Guillarme and Thuiller, 2022; Le Guillarme et al., 2023)). This field is beyond my expertise, but its potential for eDNA applications is immense. While I will probably not make development in that area myself, I'm eager to collaborate with researchers in this field and follow the field progress in the future.

7.2.3 Upscaling the approach.

Another challenge for building eDNA-based observatories in tropical forests lies in a difficulty of access to sampling sites, as well as the more limited availability of local molecular/sequencing facilities in the Global South in general. These limitations require developing robotic instruments to facilitate the sampling of environmental matrices in the most remote/unreachable areas, as well as field-compatible molecular protocols, which are part of the development we do in the BARCODRAIN project and the XPRIZE rainforest competition² (Figure 7.2a). For the XPRIZE competition, Brazilian Team's robotic subgroup, led by Marco H. Terra, is currently developing terrestrial robots and drones able to conduct water filtration, or collection of tree leaves and litter/soil material. Within the eDNA subgroup, we are developing field-compatible molecular protocols for DNA barcoding and eDNA metabarcoding. These protocols inevitably rely on MinION sequencers from Oxford Nanopore Technologies (Figure 7.2b), which are the only portable and real-time sequencing devices available at the moment. The use of this particular sequencing technologies comes with its load of challenges: the error pattern strongly differ from what is obtained with Illumina sequencing, as the nanopore sequencing error rate remains quite high (Figure 7.2c). Therefore, we had to develop both molecular and bioinformatics procedures to minimize this noise. At present, I cannot give more details on the robots or on the molecular protocols developed, but will be able to do so after the competition final, which will take place at the end of July 2024 close to Manaus, Brazil. Our BARCODRAIN project explores some of these advancements as well. Additionally, with BARCODRAIN, we are investigating the potential application of isothermal amplification methods (reviewed in Oliveira et al., 2021), which may be more easy to conduct at the field (no need of thermocycling machine) and are less sensitive to DNA template impurity than the PCR. While we currently explore these methods for species detection, I also envisage exploring their feasibility/performance for eDNA metabarcoding applications in the future.

While robotic technologies hold promises for collecting samples in highly remote, contaminated, or otherwise inaccessible areas such as most tropical forest fragments, their high cost may limit their broad application for the building eDNA observatories in tropical ecosystems. To achieve widespread and sustainable monitoring of these environments, a more effective strategy would involve local researchers and communities leading the observatory development process. However, there are large inequities of access to the eDNA technology

²Briefly, the competition itself consists in providing a list of the species and of their "insights" (e.g conservation value, traditional use, etc..) of an unknown 100 ha plot of tropical forest in less than 72 hours. No humans are allowed in the plot, and the sampling devices are allowed to be present in the plot only for the first 24 hours. The team that provides the largest list of species and insights wins. The general frame and rules of the competition make little sense to be honest (not to mention that they evolve constantly and not necessarily in the sense of improvement in my opinion..), but the competition itself has the merit to stimulate technological developments to improve biodiversity monitoring in tropical forests.

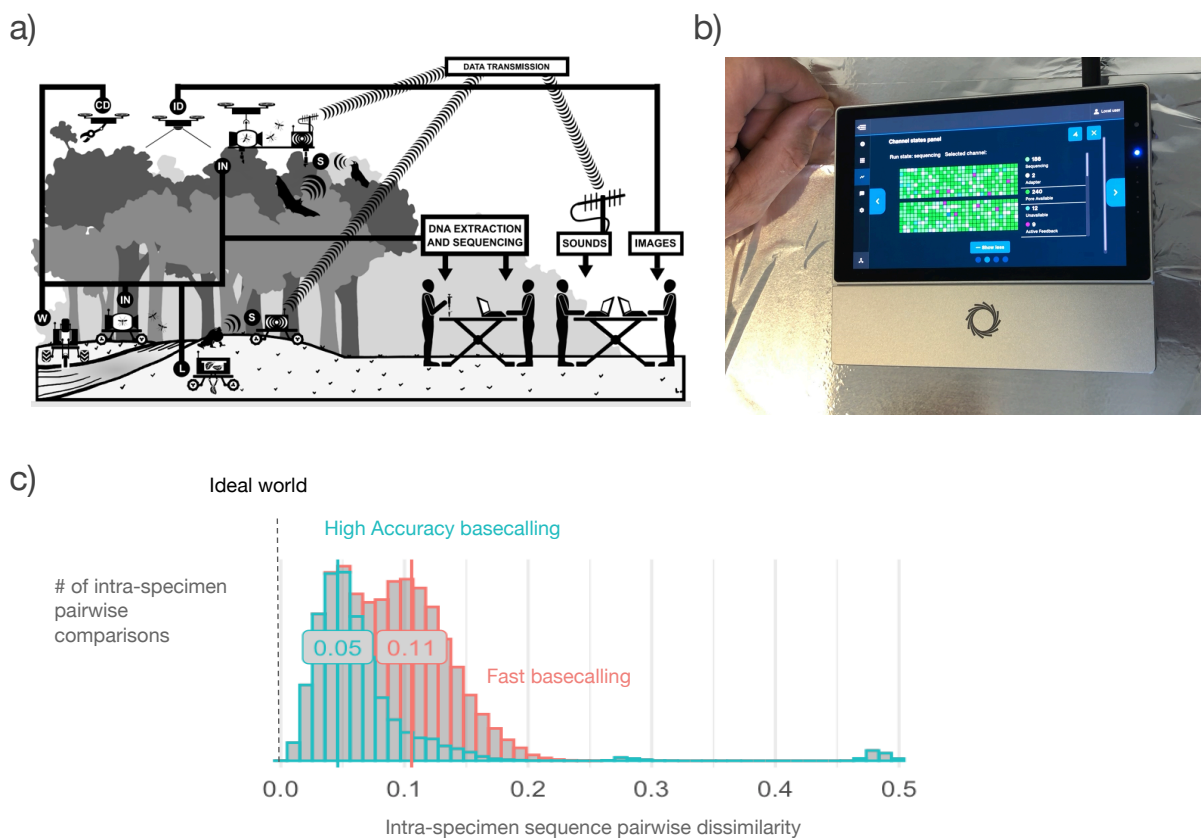


Figure 7.2: Towards real-time and large scale eDNA-based biodiversity assessment?. (a) Biodiversity census strategy of the Brazilian Team for the XPRIZE rainforest competition, which combines robotics, remote sensing, biostatistics, eDNA, as well as taxonomy and social sciences. Non-eDNA approaches are not further discussed here as I'm not directly involved in their development and because it is not directly relevant to this perspective. Modified from the Brazilian Team website. (b) Picture of a MinION MK1C device performing the sequencing of several specimens and environmental samples collected on the ESALQ Campus of Piracicaba, Brazil. Photo credit Lucie Zinger. (c) Distribution of pairwise sequences dissimilarities obtained from a single plant specimen through MinION sequencing. The marker used is the full-length *rbcL* gene. The two colours indicate by what basecalling algorithm the DNA sequences were inferred, i.e. fast (low accuracy) or high accuracy. In both case, we obtained very little to no repeated sequences.

between the global north and the global south (Shen et al., 2023). Beyond the challenges of funding and researcher training, tropical countries face the additional barrier that the molecular consumables are much more expensive due to higher taxes and transportation costs, less furnishers, hence less competition, which inherently raises the prices. For example, a MinION flowcell costs about 1000 euros in France. In Brazil, it costs 3000 euros (i.e. roughly 16 000 reais, which is about two month of salary for an experienced post-doc). Perhaps the development of international funding instruments specifically designed to offset the local costs associated with research collaborations may foster a true spirit of co-leadership between institutions from both worlds. Such an approach would not only alleviate the financial burden on institutions in tropical countries, but also promote a more equitable distribution of research partnerships. However, I must acknowledge that my limited experience and expertise of this sensitive topic preclude me to from discuss further its complexities.

Finally, and independently from the developments proposed above, the increasing availability of eDNA data, whether they are uni- or multi-taxa but also their particularities compared to more traditional techniques raise the question of how to incorporate eDNA-based biodiversity products into more formal and interdisciplinary frameworks such as the Essential Biodiversity Variables (Pereira et al., 2013) from the Group on Earth Observations Biodiversity (GEO-BON), or into national, european or international environmental policy frameworks. For example, the European Parliament has adopted in March 2024 a directive on soil monitoring and resilience ([link](#)), which will certainly align with other EU environmental frameworks such as the European Water Framework Directive or the European Marine Strategy Framework Directive. eDNA-based methods have been clearly identified as pivotal for the implementation of these directives, as evidenced by several recent Horizon Europe calls for projects on these themes for aquatic environments.

I participate to one of such project, MARCO-BOLO (PI Nicolas Pade), which aims to strengthen marine biodiversity observations in support of decision making for environmental management. The project gather multiple approaches, from remote sensing and eDNA sciences. I initially led MARCO-BOLO's eDNA working package, but due to several personal and professional changes since mid-2022, I now follow its development in a more supporting role. More specifically, I co-supervise with Chris Bowler the postdoctoral work of [Pedro Junger](#). The main objectives of this work is to explore whether we can link different types of eDNA-based indicators (taxon-related or not, use of different diversity indices) with existing biotic indicators. More broadly, we aim to define a framework for producing essential biodiversity variables for marine plankton with eDNA metabarcoding, metagenomics, metatranscriptomics. We are also starting to investigate how we could link these data with satellite products, such as Ocean color data (El Hourany et al., 2024). In the longer run, I would like to explore these aspects for soil biodiversity, to a certain extent an European context, but more importantly in tropical ecosystems. But this endeavour is possible only with sufficient data availability. We have data that were already produced in French Guiana and Colombia, and I anticipate additional soil collection within the framework of starting and future projects that we could develop in forest plot networks coordinated the ONF and CIRAD in French Guiana, or in Brazil by P. Brancalion and R. Ribeiro Rodrigues.

7.3 Tropical above-belowground feedbacks in a changing world

7.3.1 Above/belowground diversity across land uses and climates

Despite the inconsistencies and fragmented nature of existing work on land use and climate change effects on above-belowground systems, especially in tropical forests, we can draw broad general expectations from recent advances on soil biodiversity. For example, Anton Potapov has recently made considerable progresses in better describing soil food webs functioning by incorporating in a single framework resource-based and size-based energy channelling (2022, 2021, Figure 7.3a). Although the framework presents some limits, e.g. the disconnection between the microbial component and plants, it remains valuable by proposing different types of broad soil food web structures that we can assess to a certain extent with eDNA metabarcoding data (e.g. section 6.5.2), even without the developments I discuss above (section 7.2.2). Other progresses made on fungal or bacterial economics spectrum (Figure 7.3b, Camenzind et al., 2024; Piton et al., 2023) can also provide indications on the pace of nutrient cycling, metabolic capacities, and stress level of soil microbial communities, mostly through metagenomics, but also to a certain extent through eDNA metabarcoding as

certain of microbial life-history traits are conserved at high taxonomic ranks (Martiny et al., 2015)³.

Based in these frameworks, in a scenario of forest conversion to agro-systems, we may thus expected a bottom up effect through the change of vegetation and level of disturbance, with impacts on the soil food web. More specifically, by planting species with fast-growing strategies and increasing the soil disturbance magnitude and frequency as it is the case in agro-systems, we may expect the soil food web to exhibit lower complexity, higher abundance of arbuscular mycorrhizal fungi, faster and rather bacterial-based energy channels dominated by fast-growing taxa, as well as higher belowground biomass (Figure 7.3b, Jakovac et al., 2021; Neyret et al., 2024; Potapov et al., 2024). Climate change effects on the other hand are expected to be less clear as I already discuss above, but we may hypothesize changes in food web size structure (Brose et al., 2012), or, as I describe in section (6.5.2), a dependency of climate effects to the initial food web structure, which is itself likely modulated by the change of land use. Forest restoration, whether it is passive (i.e. forest recolonization of abandoned lands) or active (i.e. tree planting, above-ground animals reintroduction, or even soil inoculations) might reverse land use effects or mitigate those of climate change, but further studies are needed to determine the extent of the resilience of soil communities with different restoration strategies, as well as the pathways by which it takes (or does not take) place (Jakovac et al., 2021; Robinson et al., 2024).

Finally, in both cases, whether these responses would depend on biogeographical characteristics is unknown, despite those might exhibit significant ecological and evolutionary differences in terms of taxonomic and, in some extent functional composition of the soil biota due to current or past geographic or climatic constraints. A striking example of such long term effects is the case of North America soils, of which food webs were reconstituted without earthworms at the end of the last ice age, after glaciers retreated. These particular earthworm-free soil food webs are now threaten since earthworms reintroduction by the arrival of Europeans. Indeed, earthworms are highly invasive in these systems, and have been shown to cause a decline of above-ground insect diversity through below-to-aboveground trophic cascades, i.e. by competing with belowground detritivores and omnivores that are consumed by aboveground insect predators (Jochum et al., 2022).

In a context of limited expectations, especially in tropical soils, descriptive analysis of *in natura* large-scale patterns of diversity remains valuable, even though they cannot allow themselves directly observing above-belowground feedbacks (De Long et al., 2019). Indeed, this observational approach still allows us to test broad hypotheses such as those discussed above while simultaneously refining them for more targeted observational and experimental studies. With the thesis of [Jérémy Raynaud](#), I plan a two step approach. First, we will ask what is the relative importance of land use change vs. biogeographic/climatic/geomorphologic characteristics in driving the soil biota's taxonomic and functional composition and diversity. To this end, we will rely on an existing, unpublished eDNA metabarcoding dataset describing soil multi-trophic diversity from tropical dry forests of Colombia (PI Maily Gonzalez and Jérôme Chave). This dataset encompasses samples from eleven sites from two different ecoregions exhibiting contrasted rainfall regimes and tree community composition (Dryflor et al., 2016), and where, at each site, pairs of plots of pasture vs forest ecosystems were sampled.

Preliminary analyses conducted by [Julian Donald](#) suggest that the taxonomic α and β diversity seems rather site specific, but locally, forest conversion to pasture tend to reduce α diversity. On the other hand, the trophic guild composition clearly responded to land use (data not shown). However, these results may evolve with further data filtration and updated taxonomic and functional assignments, a work currently done by Jérémy. Besides land use change effects, he will also test more explicitly the effect of the biogeography (present and past climatic context, plant communities). I anticipate this analysis will give us some insights into (i) whether land use change has a constant, or variable impact on soil multi-trophic diversity across bioclimatic/biogeographic regions and, in the latter case, what are the regions that are the most affected. Note also that these results can be seen and interpreted in the other sense using a space-for-time substitution approach, and allow us to make some hypothesis on whether climate change have more impact on pasture or forest ecosystems.

Next, we plan to go a step further with Jérémy's thesis by considering very different bioclimatic/biogeographic conditions but increasing the breadth of land use types considered. More specifically, I envision assessing how the plant-soil system varies across 5 land use types. These will be more specifically defined depending

³There is also a substantial number of works and theories on multi-trophic biodiversity link with ecosystem functioning and stability (Barnes et al., 2018; Eisenhauer et al., 2019; Srednick, 2024) that I still need to read more thoroughly and digest. This together with the time constraints I have for writing this HDR prevent me to further develop the topic here.

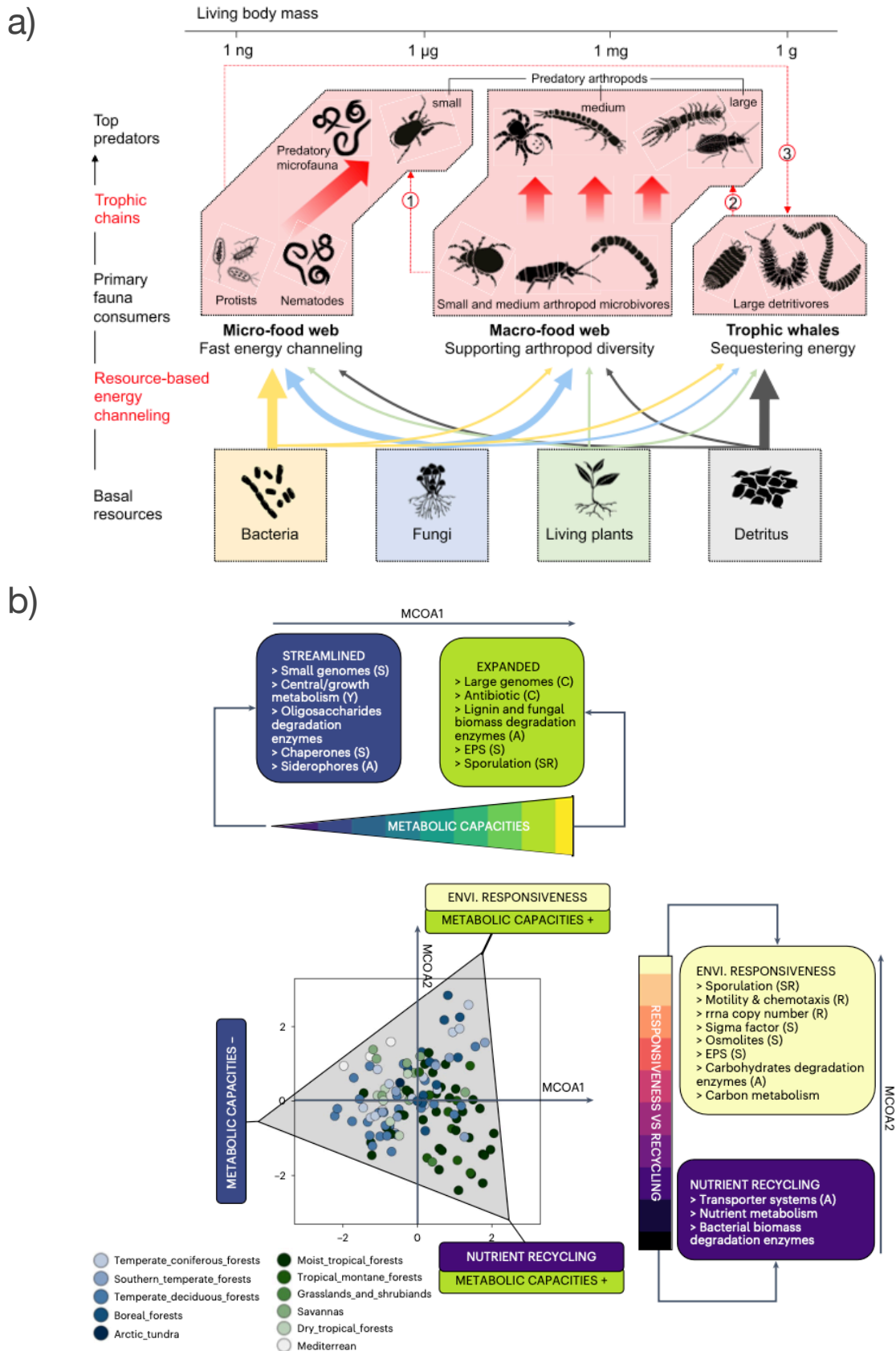


Figure 7.3: Hypotheses on soil food web structure and microbial functions. (a) Conceptual scheme of soil faunal food webs, figure from Potapov et al. (2021). (b) Life-history traits associations across soil bacterial communities, inferred from soil metagenomes. Figure from Piton et al. (2023).

on local conditions and human activities but will correspond broadly to: (i) pristine old-growth forests as a reference, (ii) intensive annual crop systems (e.g. sugar cane) or pastures and (iii) eucalyptus plantations and two mitigation management strategies, i.e. (iv) active forest restoration and (v) passive restoration (i.e. secondary forests). These modalities will be studied in two biomes with different drought frequency and intensity and biogeographic histories, i.e. the Amazonian vs. Atlantic forests (in French Guiana and the São Paulo state respectively, through collaborations with colleagues from the EcoFoG lab in Kourou, i.e. Irene Calderon Sanou, Jérôme Orivel, Céline Leroy, Raphaël Marichard, and from the University of São Paulo, i.e. Ricardo Ribeiro Rodrigues, Pedro Brancalion, Renato Ferreira de Lima, as well as the re.green company). We will then (i) characterize the whole soil food web, including plants, with a more quantitative eDNA metabarcoding approach as I indicate above (section 7.2.2). This will allow us to calculate, for each locality studied, several food web topological properties (e.g. vertical or horizontal α and β diversities, trophic guild composition, food web topological complexity, etc.) across land use types. We will also measure/retrieve from satellite products or models more contextual parameters such as soil abiotic conditions (e.g. soil pH, C/N), climate, and ecosystem functions (e.g. aboveground biomass). Finally, we plan to generate 1 metagenome per plot.

Using the produced eDNA metabarcoding data, the plan would be to assess the overall effect of land use change on soil biodiversity and soil food web structure? More specifically we will (i) test if there is a shift from macro or fungal-dominated food webs that sequester energy in forests to micro/bacteria dominated food webs that are fast energy channels in agro-systems (Figure 7.3), and, by further analysing the co-dependencies of different trophic guilds (e.g. as in Ohlmann et al., 2018, or using structural equation or species joint distribution models) identify potential trophic cascades induced by land use or climatic conditions. We will also (ii) connect soil multi-trophic properties to ecosystem level functions. Finally, we will (iii) assess the extent to which the different restoration strategies are able to reverse the observed land use change effects on the soil food web and broad ecosystem functions and whether the magnitude of recovery is equivalent in different bioclimatic / biogeographic settings. Under a space-for-time substitution perspective, that would give some hints on the potential impact of climate change on our capacities to restore the soil biodiversity of tropical forest. Besides the above mentioned collaborators, I'm also in discussion with Jérôme Mathieu to start a joint work for this specific part of the project, as he plans to analyse soil fauna response to land use change with more traditional approaches.

In parallel, soil metagenomes will be analysed so as to unravel broad microbial metabolic pathways and broad responses to the tested factors, using an approach similar to Piton et al. (Piton et al., 2023) or the functional database I mention above (section 7.2.2). This will allow consolidating observations from on eDNA metabarcoding-based food-webs, measure potential differences in several broad potential soil functions (e.g. greenhouse gas production, N cycle, etc.) but also to identify potential above-belowground feedbacks through the presence/abundance of microbial candidate genes involved in plant-soil interactions or stress responses. I also plan to analyse metagenomes phylogenetic signal, as these will also provide insights into potential evolutionary processes in response to land use or climatic conditions, as shown in earlier studies (reviewed in Martiny et al., 2023). Finally, based on metabolic pathway identification, I plan to infer microbial interactions from the community-scale metabolism (Diener et al., 2020; Levy and Borenstein, 2012), which, coupled with more phylogenetic analyses, may provide very interesting insights into potential eco-evolutionary processes within the microbial component. Note though that this part of the proposal is more risky as at present, I'm not familiar with these approaches, and that developing them will require a certain time of adaptation on my side. At present, it is not clear for me whether these modelling approaches, developed currently for gut microbiota, will also perform well or be computationally feasible on soil metagenomes, which are far more complex. Finally, I also do not exclude that we may fail to find close relative of many environmental DNA sequences in current databases due to the large sampling deficit of soil microbes genomes or metagenomes in the tropics. Some of the metagenomic data may be used for the thesis of Jérémy, but this will depend on his progresses.

Upscaling some of these analyses at larger spatial scales is now envisageable in the mid-term through the growing availability of soil eDNA multitrophic metabarcoding data in tropical forests. Besides those already mentioned above, there are several projects using the same sampling designs or DNA metabarcoding strategies in which I've been involved in French Guiana such as DIAMOND (PIs Jérôme Orivel and Jérôme Muriene), BING/BUG (PIs Jérôme Orivel and Heidy Schimann), in Australia (PIs Pierre Taberlet and Mauricio

Rossetto), but also developed by other researchers with which I collaborate or starting connections (e.g. in Mexico with Roberto Garibay Orijel, in Brazil with Markus Gastauer and Luca Mendes. Building such a database could be done with CESAB or iDIV synthesis projects. Using such data, it will be then possible to refine the observations from Jérémy's thesis and further assess more quantitatively the drivers of multi-trophic communities across land use types and biogeographic regions and along more continuous gradients of climates and land use intensity. Using such database, in link with above-belowground eco and/or evolutionary feedbacks, and building on my previous experience, I envision addressing the following questions: (i) what are the large-scale determinants of soil α and β diversity in tropical forests across soil clades and depending on models explanatory power, can we map them? (similarly to what we did for ocean's plankton, Ibarbalz et al., 2019), (ii) do the differences in patterns of these emergent properties depend on taxa body size or trophic modes?, (iii) do tree β diversity co-varies with that of soil clades and does the level of plant-soil clade β diversity co-variation depends on the soil-clade trophic proximity to plants (Ohlmann et al., 2018)? and (iv) Considering particular plant families with large spatial distributions and detected in soil samples: are there phylosymbiosis patterns between trees and their associated soil biota and if yes, do soil clade β diversity or subsets of taxa co-variation with trees phylogenies depends on trophic guild proximity to primary producers?. These later analyses may provide indirectly some insights into potential «home field advantage» for decomposer organisms, or Janzen-Connell direct or reverse effects for pathogens, herbivores, or mutualistic symbionts (e.g. mycorrhiza).

In the longer term, I intend to initiate more experimental work to gain more mechanistic understanding of plant-soil feedbacks⁴ in the light of forest restoration. For example, there is increasing evidence that plant communities restoration is favoured with soil *inocula* from pristine conditions (reviewed in Robinson et al., 2024). Likewise, it has been recently shown that fungal communities *inocula* adapted to warmer and drought conditions tend to promote tree tolerance to drought (Allsup et al., 2023). Following this ideas and building on both the methodological developments and observations made from the studies mentioned above, I intend to develop several experiments that should enable progressing our basic understanding of plant-soil eco-evolutionary feedbacks, as well as in the identification of nature-based solutions that may mitigate the impacts of land use or climate changes.

First, I would like to test the impact of plant genetic variation on soil abiotic and biotic conditions using common garden experiments (i.e. evo- to ecological feedbacks). I plan to grow tree saplings of varying genetic backgrounds, i.e. sourced from different populations (or close species), different land uses and climates, under identical environmental conditions, i.e. on sterile soils and on standard mock soil multitrophic communities of varying complexity. This will allow us to assess plant genetics and associated functional characteristics direct effects on soil the soil abiotic context and to test whether the effects on soil communities depend on their complexity. In addition, it will also allow to test whether soil communities of different complexities have an effect on plant performances (biomass, leaf and root traits, mortality). I imagine a similar experiment with the same pool of plants, but by inoculating them with soil *inocula* originating from the same or from different sites, land uses and climates from that of their plant. Such experiments will allow us to evaluate whether plant performance (biomass, leaf and root traits, mortality) is maximised with the soil communities from which they originate from or, if not what are the soil biota settings that maximise them (i.e. ecological and potentially evolutionary feedbacks). For these two types of experiments, I also envision to add several additional modalities by excluding part(s) of the food web (e.g. with bactericides, fungicides, or insecticides), in order to further understand how depletion of certain trophic guilds disrupt the plant-soil feedbacks that could be identified with undisturbed food webs.

Finally, I would like to assess the extent to which the acclimatisation of a plant to a soil *inoculum* from an environmental context that differs from its original one enhances the plant performance/tolerance to this new environmental context. Building on the above common garden experiments, we can conduct transplant experiments. Such experiment would consist in planting the different saplings, acclimated or not to non-native soil biotic conditions produced above, in both native and non non-native conditions. An improvement of tree saplings performances in non-native conditions would reflect a positive ecological effect on plant adaptation to these new conditions. Further capturing both ecological and evolutionary dynamics on the longer-term

⁴I leave here deliberately the above-ground component aside because the above-ground component, plant excepted, imposes more experimental constraints.

is a challenging endeavour for old growth tropical forests, where tree generation times exceed the decades on average, but could be reasonably implemented in either annual crop systems (e.g. sugarcane). There are many opportunities to develop such monitoring with collaborators from the University of São Paulo, in particular on the Luiz de Queiroz College of Agriculture campus. Alternatively, these processes will likely be identifiable through the monitoring of acclimated trees and the associated soil biota from the experimental approaches I mention above, and or, in parallel the monitoring of the dynamics of plant performances and of above ground and belowground taxonomic, trophic, and phylogenetic diversity in restoration plots of different ages.

Bibliography

- Achtman, M., & Wagner, M. (2008). Microbial diversity and the genetic nature of microbial species. *Nature Reviews Microbiology*, 6(6), 431–440.
- Acinas, S. G., Marcelino, L. A., Klepac-Ceraj, V., & Polz, M. F. (2004). Divergence and redundancy of 16s rRNA sequences in genomes with multiple rrn operons. *Journal of Bacteriology*, 186(9), 2629–2635.
- Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., & Polz, M. F. (2005). PCR-induced sequence artifacts and bias: Insights from comparison of two 16s rRNA clone libraries constructed from the same sample. *Applied and Environmental Microbiology*, 71(12), 8966–8969.
- Adler, A., & Dücker, E. (2018). When pasteurian science went to sea: The birth of marine microbiology. *Journal of the History of Biology*, 51, 107–133.
- Ahn, J.-H., Kim, B.-Y., Song, J., & Weon, H.-Y. (2012). Effects of PCR cycle number and DNA polymerase type on the 16s rRNA gene pyrosequencing analysis of bacterial communities. *Journal of Microbiology*, 50, 1071–1074.
- Aird, D., Ross, M. G., Chen, W.-S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D. B., Nusbaum, C., & Gnirke, A. (2011). Analyzing and minimizing PCR amplification bias in illumina sequencing libraries. *Genome Biology*, 12, 1–14.
- Alberdi, A., & Gilbert, M. T. P. (2019). A guide to the application of hill numbers to DNA-based diversity analyses. *Molecular Ecology Resources*, 19(4), 804–817.
- Allen, A. P., Gillooly, J. F., Savage, V. M., & Brown, J. H. (2006). Kinetic effects of temperature on rates of genetic divergence and speciation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(24), 9130–9135.
- Allsup, C. M., George, I., & Lankau, R. A. (2023). Shifting microbial communities can enhance tree tolerance to changing climates. *Science*, 380(6647), 835–840.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410.
- Amend, A. S., Oliver, T. A., Amaral-Zettler, L. A., Boetius, A., Fuhrman, J. A., Horner-Devine, M. C., Huse, S. M., Welch, D. B. M., Martiny, A. C., Ramette, A., et al. (2013). Macroecological patterns of marine bacteria on a global scale. *Journal of Biogeography*, 40(4), 800–811.
- Andersen, K., Bird, K. L., Rasmussen, M., Haile, J., Breuning-Madsen, H., Kjaer, K. H., Orlando, L., Gilbert, M. T. P., & Willerslev, E. (2012). Meta-barcoding of ‘dirt’DNA from soil reflects vertebrate biodiversity. *Molecular Ecology*, 21(8), 1966–1979.
- Andújar, C., Arribas, P., Ruzicka, F., Crampton-Platt, A., Timmermans, M. J., & Vogler, A. P. (2015). Phylogenetic community ecology of soil biodiversity using mitochondrial metagenomics. *Molecular Ecology*, 24(14), 3603–3617.
- Angulo, V., Beriot, N., Garcia-Hernandez, E., Li, E., Masteling, R., & Lau, J. A. (2022). Plant–microbe eco-evolutionary dynamics in a changing world. *New Phytologist*, 234(6), 1919–1928.
- Arribas, P., Andújar, C., Bidartondo, M. I., Bohmann, K., Coissac, E., Creer, S., Dewaard, J. R., Elbrecht, V., Ficetola, G. F., Goberna, M., et al. (2021). Connecting high-throughput biodiversity inventories: Opportunities for a site-based genomic framework for global integration and synthesis. *Molecular Ecology*, 30(5), 1120–1135.
- Arribas, P., Andújar, C., Bohmann, K., DeWaard, J. R., Economo, E. P., Elbrecht, V., Geisen, S., Goberna, M., Krehenwinkel, H., Novotny, V., et al. (2022). Toward global integration of biodiversity big data: A harmonized metabarcode data generation module for terrestrial arthropods. *GigaScience*, 11, giac065.

- Arribas, P., Andújar, C., Hopkins, K., Shepherd, M., & Vogler, A. P. (2016). Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. *Methods in Ecology and Evolution*, 7(9), 1071–1081.
- Averill, C., Werbin, Z. R., Atherton, K. F., Bhatnagar, J. M., & Dietze, M. C. (2021). Soil microbiome predictability increases with spatial and taxonomic scale. *Nature Ecology & Evolution*, 5(6), 747–756.
- Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., & Thingstad, F. (1983). The ecological role of water-column microbes in the sea. *Marine Ecology Progress Series*, 10(3), 257–263.
- Baas-Becking, L. (1934). *Geobiologie of inleiding tot de milieukunde*. WP Van Stockum & Zoon NV.
- Bahram, M., Kohout, P., Anslan, S., Harend, H., Abarenkov, K., & Tedersoo, L. (2016). Stochastic distribution of small soil eukaryotes resulting from high dispersal and drift in a local environment. *The ISME journal*, 10(4), 885–896.
- Bailey, J. K., Wooley, S. C., Lindroth, R. L., & Whitham, T. G. (2006). Importance of species interactions to community heritability: A genetic basis to trophic-level interactions. *Ecology Letters*, 9(1), 78–85.
- Baptist, F., Zinger, L., Clement, J., Gallet, C., Guillemain, R., Martins, J., Sage, L., Shahnavaz, B., Choler, P., & Geremia, R. (2008). Tannin impacts on microbial diversity and the functioning of alpine soils: A multidisciplinary approach. *Environmental Microbiology*, 10(3), 799–809.
- Barbera, P., Kozlov, A. M., Czech, L., Morel, B., Darriba, D., Flouri, T., & Stamatakis, A. (2019). Epa-ng: Massively parallel evolutionary placement of genetic sequences. *Systematic Biology*, 68(2), 365–369.
- Barberán, A., Casamayor, E. O., & Fierer, N. (2014). The microbial contribution to macroecology. *Frontiers in Microbiology*, 5, 89056.
- Barberán, A., McGuire, K. L., Wolf, J. A., Jones, F. A., Wright, S. J., Turner, B. L., Essene, A., Hubbell, S. P., Faircloth, B. C., & Fierer, N. (2015). Relating belowground microbial composition to the taxonomic, phylogenetic, and functional trait distributions of trees in a tropical forest. *Ecology Letters*, 18(12), 1397–1405.
- Bardgett, R., & van der Putten, W. (2014). Belowground biodiversity and ecosystem functioning. *Nature*, 515(7528), 505–511.
- Barley, A. J., & Thomson, R. C. (2016). Assessing the performance of DNA barcoding using posterior predictive simulations. *Molecular Ecology*, 25(9), 1944–1957.
- Barlow, J. T., Bogatyrev, S. R., & Ismagilov, R. F. (2020). A quantitative sequencing framework for absolute abundance measurements of mucosal and luminal microbial communities. *Nature Communications*, 11(1), 2590.
- Barnes, A. D., Jochum, M., Lefcheck, J. S., Eisenhauer, N., Scherber, C., O’Connor, M. I., de Ruiter, P., & Brose, U. (2018). Energy flux: The link between multitrophic biodiversity and ecosystem functioning. *Trends in Ecology & Evolution*, 33(3), 186–197.
- Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, 17(1), 1–17.
- Bar-On, Y., Phillips, R., & Milo, R. (2018). The biomass distribution on earth. *Proceedings of the National Academy of Sciences of the United States of America*, 115(25), 6506–6511.
- Baxendale, C., Orwin, K. H., Poly, F., Pommier, T., & Bardgett, R. D. (2014). Are plant–soil feedback responses explained by plant traits? *New Phytologist*, 204(2), 408–423.
- Beaugrand, G., Edwards, M., Brander, K., Luczak, C., & Ibanez, F. (2008). Causes and projections of abrupt climate-driven ecosystem shifts in the north atlantic. *Ecology Letters*, 11(11), 1157–1168.
- Beck, J., Holloway, J. D., & Schwanghart, W. (2013). Undersampling and the measurement of beta diversity. *Methods in Ecology and Evolution*, 4(4), 370–382.
- Beijerinck, M. (1913). *De infusies en de ontdekking der bakterien*. Johannes Müller.
- Bell, T., Ager, D., Song, J.-I., Newman, J., Thompson, I., Lilley, A., & van der Gast, C. (2005). Larger islands house more bacterial taxa. *Science*, 308(5730), 1884–1884.
- Berg, M. P. (2013). Patterns of biodiversity at fine and small spatial scales. In D. H. Wall, R. D. Bardgett, V. Behan-Pelletier, J. E. Herrick, T. H. Jones, K. Ritz, J. Six, D. R. Strong, & W. H. van der Putten (Eds.), *Soil ecology and ecosystem services* (pp. 136–152). Oxford University Press Oxford, UK.
- Bestion, E., Clobert, J., & Cote, J. (2015). Dispersal response to climate change: Scaling down to intraspecific variation. *Ecology Letters*, 18(11), 1226–1233.

BIBLIOGRAPHY

- Bestion, E., Jacob, S., Zinger, L., Di Gesu, L., Richard, M., White, J., & Cote, J. (2017). Climate warming reduces gut microbiota diversity in a vertebrate ectotherm. *Nature Ecology & Evolution*, *1*(6), 0161.
- Bestion, E., Soriano-Redondo, A., Cucherousset, J., Jacob, S., White, J., Zinger, L., Fourtune, L., Di Gesu, L., Teyssier, A., & Cote, J. (2019). Altered trophic interactions in warming climates: Consequences for predator diet breadth and fitness. *Proceedings of the Royal Society B*, *286*(1914), 20192227.
- Bestion, E., Teyssier, A., Richard, M., Clobert, J., & Cote, J. (2015). Live fast, die young: Experimental evidence of population extinction risk due to climate change. *PLoS Biology*, *13*(10), e1002281.
- Bever, J. D., Dickie, I. A., Facelli, E., Facelli, J. M., Klironomos, J., Moora, M., Rillig, M. C., Stock, W. D., Tibbett, M., & Zobel, M. (2010). Rooting theories of plant community ecology in microbial interactions. *Trends in Ecology & Evolution*, *25*(8), 468–478.
- Bienhold, C., Zinger, L., Boetius, A., & Ramette, A. (2016). Diversity and biogeography of bathyal and abyssal seafloor bacteria. *PLoS One*, *11*(1), e0148016.
- Birer, C. (2017, April). *Le microbiote bactérien cuticulaire des fourmis de Guyane : pouvoir antibiotique et écologie des communautés* (Publication No. 2017YANE0003) [Theses]. Université de Guyane.
- Birer, C., Moreau, C. S., Tysklind, N., Zinger, L., & Duplais, C. (2020). Disentangling the assembly mechanisms of ant cuticular bacterial communities of two amazonian ant species sharing a common arboreal nest. *Molecular Ecology*, *29*(7), 1372–1385.
- Birer, C., Tysklind, N., Zinger, L., & Duplais, C. (2017). Comparative analysis of DNA extraction methods to study the body surface microbiota of insects: A case study with ant cuticular bacteria. *Molecular Ecology Resources*, *17*(6), e34–e45.
- Bjørnsgaard Aas, A., Davey, M. L., & Kauserud, H. (2017). Its all right mama: Investigating the formation of chimeric sequences in the its 2 region by DNA metabarcoding analyses of fungal mock communities of different complexities. *Molecular Ecology Resources*, *17*(4), 730–741.
- Blackman, R., Couton, M., Keck, F., Kirschner, D., Carraro, L., Cereghetti, E., Perrelet, K., Bossart, R., Brantschen, J., Zhang, Y., et al. (2024). Environmental DNA: The next chapter. *Molecular Ecology*, e17355.
- Bonin, A., Guerrieri, A., & Ficetola, G. F. (2023). Optimal sequence similarity thresholds for clustering of molecular operational taxonomic units in DNA metabarcoding studies. *Molecular Ecology Resources*, *23*(2), 368–381.
- Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Seferian, R., et al. (2013). Multiple stressors of ocean ecosystems in the 21st century: Projections with cmip5 models. *Biogeosciences*, *10*(10), 6225–6245.
- Botella, C., Joly, A., Bonnet, P., Munoz, F., & Monestiez, P. (2019). A new multi-species method to correct for sampling bias in presence only niche models, and its application to pl@ ntnet citizen science data in france. *ESA 2019-Annual Meeting of the Ecological Society of America*.
- Bouasria, A., Mustafa, T., De Bello, F., Zinger, L., Lemperiere, G., Geremia, R. A., & Choler, P. (2012). Changes in root-associated microbial communities are determined by species-specific plant growth responses to stress and disturbance. *European Journal of Soil Biology*, *52*, 59–66.
- Boucias, D. G., Zhou, Y., Huang, S., & Keyhani, N. O. (2018). Microbiota in insect fungal pathology. *Applied Microbiology and Biotechnology*, *102*, 5873–5888.
- Boulton, C. A., Lenton, T. M., & Boers, N. (2022). Pronounced loss of amazon rainforest resilience since the early 2000s. *Nature Climate Change*, *12*(3), 271–278.
- Boussarie, G., Bakker, J., Wangenstein, O. S., Mariani, S., Bonnin, L., Juhel, J.-B., Kiszka, J. J., Kulbicki, M., Manel, S., Robbins, W. D., et al. (2018). Environmental DNA illuminates the dark diversity of sharks. *Science Advances*, *4*(5), eaap9661.
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). Obitoools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*(1), 176–182.
- Brose, U., Dunne, J. A., Montoya, J. M., Petchey, O. L., Schneider, F. D., & Jacob, U. (2012). Climate change in size-structured ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1605), 2903–2912.
- Brown, G. G., Demetrio, W. C., Gabriac, Q., Pasini, A., Korasaki, V., Oliveira, L. J., Dos Santos, J. C., Torres, E., Galerani, P. R., Gazziero, D. L., et al. (2024). Soil macrofauna communities in brazilian land-use systems. *Biodiversity Data Journal*, *12*.
- Brown, J. H. (2014). Why are there so many species in the tropics? *Journal of Biogeography*, *41*(1), 8–22.

- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, *85*(7), 1771–1789.
- Burian, A., Mauvisseau, Q., Bulling, M., Domisch, S., Qian, S., & Sweet, M. (2021). Improving the reliability of eDNA data interpretation. *Molecular Ecology Resources*, *21*(5), 1422–1433.
- Burki, F., Roger, A. J., Brown, M. W., & Simpson, A. G. (2020). The new tree of eukaryotes. *Trends in Ecology & Evolution*, *35*(1), 43–55.
- Calderón-Sanou, I., Münkemüller, T., Boyer, F., Zinger, L., & Thuiller, W. (2020). From environmental DNA sequences to ecological conclusions: How strong is the influence of methodological choices? *Journal of Biogeography*, *47*(1), 193–206.
- Calderón-Sanou, I., Ohlmann, M., Münkemüller, T., Zinger, L., Hedde, M., Lionnet, C., Martinez-Almoyna, C., Saillard, A., Renaud, J., Le Guillarme, N., et al. (2024). Mountain soil multitrophic networks shaped by the interplay between habitat and pedoclimatic conditions. *Soil Biology and Biochemistry*, *190*, 109282.
- Calderón-Sanou, I., Zinger, L., Hedde, M., Martinez-Almoyna, C., Saillard, A., Renaud, J., Gielly, L., Khedim, N., Lionnet, C., Ohlmann, M., et al. (2022). Energy and physiological tolerance explain multi-trophic soil diversity in temperate mountains. *Diversity and Distributions*, *28*(12), 2549–2564.
- Caldwell, I. R., Hobbs, J.-P. A., Bowen, B. W., Cowman, P. F., DiBattista, J. D., Whitney, J. L., Ahti, P. A., Belderok, R., Canfield, S., Coleman, R. R., et al. (2024). Global trends and biases in biodiversity conservation research. *Cell Reports Sustainability*.
- Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, *11*(12), 2639–2643.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). Dada2: High-resolution sample inference from illumina amplicon data. *Nature Methods*, *13*(7), 581–583.
- Camenzind, T., Aguilar-Trigueros, C. A., Hempel, S., Lehmann, A., Bielcik, M., Andrade-Linares, D. R., Bergmann, J., dela Cruz, J., Gawronski, J., Golubeva, P., et al. (2024). Towards establishing a fungal economics spectrum in soil saprobic fungi. *Nature Communications*, *15*(1), 3321.
- Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., Wardle, D. A., et al. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59–67.
- Carini, P., Marsden, P. J., Leff, J. W., Morgan, E. E., Strickland, M. S., & Fierer, N. (2016). Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nature Microbiology*, *2*(3), 1–6.
- Carlsen, T., Aas, A. B., Lindner, D., Vrålstad, T., Schumacher, T., & Kauserud, H. (2012). Don't make a mista (g) ke: Is tag switching an overlooked source of error in amplicon pyrosequencing studies? *Fungal Ecology*, *5*(6), 747–749.
- Caruso, V., Song, X., Asquith, M., & Karstens, L. (2019). Performance of microbiome sequence inference methods in environments with varying biomass. *MSystems*, *4*(1), e00163–18.
- Carvalho, C. S., De Oliveira, M. E., Rodriguez-Castro, K. G., Saranholi, B. H., & Galetti Jr, P. M. (2022). Efficiency of eDNA and iDNA in assessing vertebrate diversity and its abundance. *Molecular Ecology Resources*, *22*(4), 1262–1273.
- Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., Behrenfeld, M. J., Boetius, A., Boyd, P. W., Classen, A. T., et al. (2019). Scientists' warning to humanity: Microorganisms and climate change. *Nature Reviews Microbiology*, *17*(9), 569–586.
- CBOL Plant Working Group, Hollingsworth, P., Forrest, L., Spouge, J., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M., Cowan, R., Erickson, D., Fazekas, A., Graham, S., James, K., Kim, K.-J., Kress, W., Schneider, H., van AlphenStahl, J., Barrett, S., van den Berg, C., ... Little, D. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(31), 12794–12797.
- Chalmandrier, L., Pansu, J., Zinger, L., Boyer, F., Coissac, E., Génin, A., Gielly, L., Lavergne, S., Legay, N., Schilling, V., et al. (2019). Environmental and biotic drivers of soil microbial β -diversity across spatial and phylogenetic scales. *Ecography*, *42*(12), 2144–2156.
- Chao, A., Chiu, C.-H., & Jost, L. (2014). Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers. *Annual Review of Ecology, Evolution, and Systematics*, *45*, 297–324.

BIBLIOGRAPHY

- Chase, J., & Leibold, M. (2003). *Ecological niches: Linking classical and contemporary approaches*. University of Chicago Press.
- Chase, J., & Myers, J. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1576), 2351–2363.
- Chase, J. M. (2014). Spatial scale resolves the niche versus neutral theory debate. *Journal of vegetation science*, 25(2), 319–322.
- Chase, J. M., McGill, B. J., McGlinn, D. J., May, F., Blowes, S. A., Xiao, X., Knight, T. M., Purschke, O., & Gotelli, N. J. (2018). Embracing scale-dependence to achieve a deeper understanding of biodiversity and its change across communities. *Ecology Letters*, 21(11), 1737–1751.
- Chave, J. (2004). Neutral theory and community ecology. *Ecology Letters*, 7(3), 241–253.
- Chave, J. (2008). Spatial variation in tree species composition across tropical forests: Pattern and process. In W. Carson & S. A. Schnitzer (Eds.), *Tropical forest community ecology* (pp. 11–30). Wiley-Blackwell Chichester, UK.
- Chen, B. (2015). Patterns of thermal limits of phytoplankton. *Journal of Plankton Research*, 37(2), 285–292.
- Chen, W., & Ficetola, G. F. (2019). Conditionally autoregressive models improve occupancy analyses of autocorrelated data: An example with environmental DNA. *Molecular Ecology Resources*, 19(1), 163–175.
- Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, 31(1), 343–366.
- Choler, P. (2005). Consistent shifts in alpine plant traits along a mesotopographical gradient. *Arctic, Antarctic, and Alpine Research*, 37(4), 444–453.
- Chrostek, E., Pelz-Stelinski, K., Hurst, G. D., & Hughes, G. L. (2017). Horizontal transmission of intracellular insect symbionts via plants. *Frontiers in Microbiology*, 8, 301416.
- Condit, R., Pitman, N., Leigh Jr, E. G., Chave, J., Terborgh, J., Foster, R. B., Núñez, P., Aguilar, S., Valencia, R., Villa, G., et al. (2002). Beta-diversity in tropical forest trees. *Science*, 295(5555), 666–669.
- Connell, J. H. (1978). Diversity in tropical rain forests and coral reefs: High diversity of trees and corals is maintained only in a nonequilibrium state. *Science*, 199(4335), 1302–1310.
- Cornelissen, J. H. C., Cornwell, W. K., Freschet, G. T., Weedon, J. T., Berg, M. P., & Zanne, A. E. (2023). Coevolutionary legacies for plant decomposition. *Trends in Ecology & Evolution*, 38(1), 44–54.
- Cortois, R., Schröder-Georgi, T., Weigelt, A., van der Putten, W. H., & De Deyn, G. B. (2016). Plant–soil feedbacks: Role of plant functional group and plant traits. *Journal of Ecology*, 104(6), 1608–1617.
- Cottenie, K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, 8(11), 1175–1182.
- Craft, K. J., Pauls, S. U., Darrow, K., Miller, S. E., Hebert, P. D., Helgen, L. E., Novotny, V., & Weiblen, G. D. (2010). Population genetics of ecological communities with DNA barcodes: An example from new guinea lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America*, 107(11), 5041–5046.
- Danovaro, R., Molari, M., Corinaldesi, C., & Dell’Anno, A. (2016). Macroecological drivers of archaea and bacteria in benthic deep-sea ecosystems. *Science Advances*, 2(4), e1500961.
- Darwin, C. (1859). *On the origin of species*. London: John Murray.
- De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources*, 14(2), 306–323.
- De Bie, T., De Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., Hampel, H., Denys, L., Vanhecke, L., Van der Gucht, K., et al. (2012). Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters*, 15(7), 740–747.
- De Cáceres, M., Legendre, P., & Moretti, M. (2010). Improving indicator species analysis by combining groups of sites. *Oikos*, 119(10), 1674–1684.
- De Long, J. R., Fry, E. L., Veen, G., & Kardol, P. (2019). Why are plant–soil feedbacks so unpredictable, and what to do about it? *Functional Ecology*, 33(1), 118–128.
- De Wit, R., & Bouvier, T. (2006). ‘everything is everywhere, but, the environment selects’; what did baas becking and bejerinck really say? *Environmental Microbiology*, 8(4), 755–758.

- Decaëns, T. (2010). Macroecological patterns in soil communities. *Global Ecology and Biogeography*, 19(3), 287–302.
- Decina, S. M., Templer, P. H., Hutyla, L. R., Gately, C. K., & Rao, P. (2017). Variability, drivers, and effects of atmospheric nitrogen inputs across an urban area: Emerging patterns among human activities, the atmosphere, and soils. *Science of the Total Environment*, 609, 1524–1534.
- Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., Maestre, F. T., Singh, B. K., & Fierer, N. (2018). A global atlas of the dominant bacteria found in soil. *Science*, 359(6373), 320–325.
- Díaz, S., Settele, J., Brondízio, E. S., Ngo, H. T., Agard, J., Arneth, A., Balvanera, P., Brauman, K. A., Butchart, S. H., Chan, K. M., et al. (2019). Pervasive human-driven decline of life on earth points to the need for transformative change. *Science*, 366(6471), eaax3100.
- Dickie, I., Boyer, S., Buckley, H., Duncan, R., Gardner, P., Hogg, I., Holdaway, R., Lear, G., Makiola, A., Morales, S., Powell, J., & Weaver, L. (2018). Towards robust and repeatable sampling methods in eDNA-based studies. *Molecular Ecology Resources*, 18(5), 940–952.
- Diener, C., Gibbons, S. M., & Resendis-Antonio, O. (2020). Micom: Metagenome-scale modeling to infer metabolic interactions in the gut microbiota. *MSystems*, 5(1), 10–1128.
- Diniz-Filho, J. A. F., Bini, L. M., Targueta, C. P., de Campos Telles, M. P., Jardim, L., Machado, K. B., Nabout, J. C., Nunes, R., Vieira, L. C. G., & Soares, T. N. (2024). Environmental DNA and biodiversity patterns: A call for a community phylogenetics approach. *Perspectives in Ecology and Conservation*.
- Dohm, J. C., Lottaz, C., Borodina, T., & Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research*, 36(16), e105.
- Donald, J., Murienne, J., Chave, J., Iribar, A., Louisanna, E., Manzi, S., Roy, M., Tao, S., Orivel, J., Schimann, H., et al. (2021). Multi-taxa environmental DNA inventories reveal distinct taxonomic and functional diversity in urban tropical forest fragments. *Global Ecology and Conservation*, 29, e01724.
- Dongen, S. (2000). *A cluster algorithm for graphs*. CWI (Centre for Mathematics; Computer Science).
- Dowle, E., Morgan-Richards, M., & Trewick, S. (2013). Molecular evolution and the latitudinal biodiversity gradient. *Heredity*, 110(6), 501–510.
- Drakare, S., Lennon, J. J., & Hillebrand, H. (2006). The imprint of the geographical, evolutionary and ecological context on species–area relationships. *Ecology Letters*, 9(2), 215–227.
- Dray, S., Péliissier, R., Couteron, P., Fortin, M.-J., Legendre, P., Peres-Neto, P. R., Bellier, E., Bivand, R., Blanchet, F. G., De Cáceres, M., et al. (2012). Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs*, 82(3), 257–275.
- Dryflor, Banda-R, K., Delgado-Salinas, A., Dexter, K. G., Linares-Palomino, R., Oliveira-Filho, A., Prado, D., Pullan, M., Quintana, C., Riina, R., et al. (2016). Plant diversity patterns in neotropical dry forests and their conservation implications. *Science*, 353(6306), 1383–1387.
- Dupuis, J. R., Roe, A. D., & Sperling, F. A. (2012). Multi-locus species delimitation in closely related animals and fungi: One marker is not enough. *Molecular Ecology*, 21(18), 4422–4436.
- Ebrahimi-Zarandi, M., Etesami, H., & Glick, B. R. (2023). Fostering plant resilience to drought with actinobacteria: Unveiling perennial allies in drought stress tolerance. *Plant Stress*, 100242.
- Edgar, R. C. (2013). Uparse: Highly accurate otu sequences from microbial amplicon reads. *Nature Methods*, 10(10), 996–998.
- Eisenhauer, N., Schielzeth, H., Barnes, A. D., Barry, K. E., Bonn, A., Brose, U., Bruehlheide, H., Buchmann, N., Buscot, F., Ebeling, A., et al. (2019). A multitrophic perspective on biodiversity–ecosystem functioning research. In *Advances in ecological research* (pp. 1–54, Vol. 61). Elsevier.
- El Hourany, R., Pierella Karlusich, J., Zinger, L., Loisel, H., Levy, M., & Bowler, C. (2024). Linking satellites to genes with machine learning to estimate phytoplankton community structure from space. *Ocean Science*, 20(1), 217–239.
- Estensmo, E. L. F., Maurice, S., Morgado, L., Martin-Sanchez, P. M., Skrede, I., & Kauserud, H. (2021). The influence of intraspecific sequence variation during DNA metabarcoding: A case study of eleven fungal species. *Molecular Ecology Resources*, 21(4), 1141–1148.
- Estes, J. A., Terborgh, J., Brashares, J. S., Power, M. E., Berger, J., Bond, W. J., Carpenter, S. R., Essington, T. E., Holt, R. D., Jackson, J. B., et al. (2011). Trophic downgrading of planet earth. *Science*, 333(6040), 301–306.

BIBLIOGRAPHY

- Estes, L., Elsen, P. R., Treuer, T., Ahmed, L., Caylor, K., Chang, J., Choi, J. J., & Ellis, E. C. (2018). The spatial and temporal domains of modern ecology. *Nature Ecology & Evolution*, *2*(5), 819–826.
- Ettema, C. H., & Wardle, D. A. (2002). Spatial soil ecology. *Trends in Ecology & Evolution*, *17*(4), 177–183.
- Falkowski, P., Fenchel, T., & Delong, E. (2008). The microbial engines that drive earth's biogeochemical cycles. *Science*, *320*(5879), 1034–1039.
- Fanin, N., Lin, D., Freschet, G. T., Keiser, A. D., Augusto, L., Wardle, D. A., & Veen, G. (2021). Home-field advantage of litter decomposition: From the phyllosphere to the soil. *New Phytologist*, *231*(4), 1353–1358.
- FAO. (2019). *The state of the world's biodiversity for food and agriculture*. (J. Bélanger & D. Pilling, Eds.).
- Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., & Huttenhower, C. (2012). Microbial co-occurrence relationships in the human microbiome. *PLoS Computational Biology*, *8*(7), e1002606.
- Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguët-Covex, C., De Barba, M., Gielly, L., Lopes, C. M., Boyer, F., Pompanon, F., et al. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, *15*(3), 543–556.
- Ficetola, G. F., & Taberlet, P. (2023). Towards exhaustive community ecology via DNA metabarcoding. *Molecular Ecology*, *32*(23), 6320–6329.
- Ficetola, G., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, *4*(4), 423–425.
- Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science*, *281*(5374), 237–240.
- Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 626–631.
- Finlay, B. (2002). Global dispersal of free-living microbial eukaryote species. *Science*, *296*(5570), 1061–1063.
- Fonseca, V. G. (2018). Pitfalls in relative abundance estimation using eDNA metabarcoding. *Molecular Ecology Resources*, *18*, 923–926.
- Fountain-Jones, N. M., Giraud, T., Zinger, L., Bik, H., Creer, S., & Videvall, E. (2024). Molecular ecology of microbiomes in the wild: Common pitfalls, methodological advances and future directions. *Molecular Ecology*, *33*(2), e17223.
- Fox, G., Pechman, K., & Woese, C. (1977). Comparative cataloging of 16s ribosomal ribonucleic acid: Molecular approach to procaryotic systematics. *International Journal of Systematic and Evolutionary Microbiology*, *27*(1), 44–57.
- Fromm, E., Zinger, L., Pellerin, F., Di Gesu, L., Jacob, S., Winandy, L., Aguilée, R., Parthuisot, N., Iribar, A., White, J., et al. (2024). Warming effects on lizard gut microbiome depend on habitat connectivity. *Proceedings of the Royal Society B*, *291*(2021), 20240220.
- Frøslev, T. G., Kjølner, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., & Hansen, A. J. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications*, *8*(1), 1188.
- Fuhrman, J. A., Steele, J. A., Hewson, I., Schwalbach, M. S., Brown, M. V., Green, J. L., & Brown, J. H. (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(22), 7774–7778.
- Gao, Y., & Wu, M. (2023). Accounting for 16s rRNA copy number prediction uncertainty and its implications in bacterial diversity analyses. *ISME communications*, *3*(1), 59.
- Gardes, M., & Dahlberg, A. (1996). Mycorrhizal diversity in arctic and alpine tundra: An open question. *New Phytologist*, *133*(1), 147–157.
- Geremia, R. A., Puşcaş, M., Zinger, L., Bonneville, J.-M., & Choler, P. (2016). Contrasting microbial biogeographical patterns between anthropogenic subalpine grasslands and natural alpine grasslands. *New Phytologist*, *209*(3), 1196–1207.
- Gilbert, B., & Bennett, J. (2010). Partitioning variation in ecological communities: Do the numbers add up? *Journal of Applied Ecology*, *47*(5), 1071–1082.
- Giovannoni, S., Britschgi, T., Moyer, C., & Field, K. (1990). Genetic diversity in sargasso sea bacterioplankton. *Nature*, *345*(6270), 60–63.
- Glassman, S. I., & Martiny, J. B. (2018). Broad-scale ecological patterns are robust to use of exact sequence variants versus operational taxonomic units. *MSphere*, *3*(4), 10–1128.

- Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017). Microbiome datasets are compositional: And this is not optional. *Frontiers in Microbiology*, *8*, 294209.
- Gong, W., & Marchetti, A. (2019). Estimation of 18s gene copy number in marine eukaryotic plankton using a next-generation sequencing approach. *Frontiers in Marine Science*, *6*, 219.
- Gotelli, N., & Graves, G. (1996). *Null models in ecology*. Smithsonian Institution Press.
- Graham, C. H., Storch, D., & Machac, A. (2018). Phylogenetic scale in ecology and evolution. *Global Ecology and Biogeography*, *27*(2), 175–187.
- Gravel, D., Massol, F., Canard, E., Mouillot, D., & Mouquet, N. (2011). Trophic theory of island biogeography. *Ecology Letters*, *14*(10), 1010–1016.
- Green, J., Holmes, A., Westoby, M., Oliver, I., Briscoe, D., Dangerfield, M., Gillings, M., & Beattie, A. (2004). Spatial scaling of microbial eukaryote diversity. *Nature*, *432*(7018), 747–750.
- Grilli, J. (2020). Macroecological laws describe variation and diversity in microbial communities. *Nature Communications*, *11*(1), 4743.
- Groussin, M., Mazel, F., Sanders, J. G., Smillie, C. S., Lavergne, S., Thuiller, W., & Alm, E. J. (2017). Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. *Nature Communications*, *8*(1), 14319.
- Guerra, C. A., Heintz-Buschart, A., Sikorski, J., Chatzinotas, A., Guerrero-Ramírez, N., Cesarz, S., Beaumelle, L., Rillig, M. C., Maestre, F. T., Delgado-Baquerizo, M., et al. (2020). Blind spots in global soil biodiversity and ecosystem function research. *Nature Communications*, *11*(1), 3870.
- Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., Darzi, Y., Audic, S., Berline, L., Brum, J., Coelho, L., Espinoza, J., Malviya, S., Sunagawa, S., Dimier, C., Kandels-Lewis, S., Picheral, M., Poulain, J., Searson, S., ... Gorsky, G. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature*, *532*(7600), 465–470.
- Gury, J., Zinger, L., Gielly, L., Taberlet, P., & Geremia, R. A. (2008). Exonuclease activity of proofreading DNA polymerases is at the origin of artifacts in molecular profiling studies. *Electrophoresis*, *29*(11), 2437–2444.
- Hacquard, S., Garrido-Oter, R., González, A., Spaepen, S., Ackermann, G., Lebeis, S., McHardy, A., Dangl, J., Knight, R., Ley, R., & Schulze-Lefert, P. (2015). Microbiota and host nutrition across plant and animal kingdoms. *Cell Host & Microbe*, *17*(5), 603–616.
- Haegeman, B., Hamelin, J., Moriarty, J., Neal, P., Dushoff, J., & Weitz, J. S. (2013). Robust estimation of microbial diversity in theory and in practice. *The ISME journal*, *7*(6), 1092–1101.
- Hallin, S., Philippot, L., Löffler, F. E., Sanford, R. A., & Jones, C. M. (2018). Genomics and ecology of novel n2o-reducing microorganisms. *Trends in Microbiology*, *26*(1), 43–55.
- Handelsman, J., Rondon, M., Brady, S., Clardy, J., & Goodman, R. (1998). Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chemistry & Biology*, *5*(10), R245–R249.
- Hansen, M. C., Tolker-Nielsen, T., Givskov, M., & Molin, S. (1998). Biased 16s rDNA PCR amplification caused by interference from DNA flanking the template region. *FEMS Microbiology Ecology*, *26*(2), 141–149.
- Hanson, C., Fuhrman, J., Horner-Devine, M., & Martiny, J. (2012). Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology*, *10*(7), 497–506.
- Hardy, O. J., & Sonké, B. (2004). Spatial pattern analysis of tree species distribution in a tropical rain forest of cameroon: Assessing the role of limited dispersal and niche differentiation. *Forest Ecology and Management*, *197*(1-3), 191–202.
- Harte, J., Kinzig, A., & Green, J. (1999). Self-similarity in the distribution and abundance of species. *Science*, *284*(5412), 334–336.
- Hartig, F., Abrego, N., Bush, A., Chase, J. M., Guillera-Aroita, G., Leibold, M. A., Ovaskainen, O., Pellissier, L., Pichler, M., Poggiato, G., et al. (2024). Novel community data in ecology-properties and prospects. *Trends in Ecology & Evolution*.
- Hartigan, J. A., & Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *28*(1), 100–108.
- Hassler, H. B., Probert, B., Moore, C., Lawson, E., Jackson, R. W., Russell, B. T., & Richards, V. P. (2022). Phylogenies of the 16s rRNA gene and its hypervariable regions lack concordance with core genome phylogenies. *Microbiome*, *10*(1), 104.

BIBLIOGRAPHY

- Hatton, I. A., Heneghan, R. F., Bar-On, Y. M., & Galbraith, E. D. (2021). The global ocean size spectrum from bacteria to whales. *Science Advances*, *7*(46), eabh3732.
- Hebert, P., Cywinska, A., Ball, S., & de Waard, J. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, *270*(1512), 313–321.
- Henson, S. A., Sanders, R., & Madsen, E. (2012). Global patterns in efficiency of particulate organic carbon export and transfer to the deep ocean. *Global Biogeochemical Cycles*, *26*(1).
- Hill, M. O. (1973). Diversity and evenness: A unifying notation and its consequences. *Ecology*, *54*(2), 427–432.
- Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *The American Naturalist*, *163*(2), 192–211.
- Hleap, J. S., Littlefair, J. E., Steinke, D., Hebert, P. D., & Cristescu, M. E. (2021). Assessment of current taxonomic assignment strategies for metabarcoding eukaryotes. *Molecular Ecology Resources*, *21*(7), 2190–2203.
- Hochkirch, A., Samways, M. J., Gerlach, J., Böhm, M., Williams, P., Cardoso, P., Cumberlidge, N., Stephenson, P. J., Seddon, M. B., Clausnitzer, V., et al. (2021). A strategy for the next decade to address data deficiency in neglected biodiversity. *Conservation Biology*, *35*(2), 502–509.
- Horner-Devine, M., Lage, M., Hughes, J., & Bohannon, B. (2004). A taxa-area relationship for bacteria. *Nature*, *432*(7018), 750–753.
- Hubbell, S. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton University Press.
- Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., Butterfield, C. N., HERNSDORF, A. W., Amano, Y., Ise, K., et al. (2016). A new view of the tree of life. *Nature Microbiology*, *1*(5), 1–6.
- Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). Megan analysis of metagenomic data. *Genome Research*, *17*(3), 377–386.
- Hutchinson, G. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, *22*, 415–427.
- Hyde, K. D., Jeewon, R., Chen, Y.-J., Bhunjun, C. S., Calabon, M. S., Jiang, H.-B., Lin, C.-G., Norphanphoun, C., Sysouphanthong, P., Pem, D., et al. (2020). The numbers of fungi: Is the descriptive curve flattening? *Fungal Diversity*, *103*, 219–271.
- Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., Coelho, L. P., Endo, H., Gasol, J. M., Gregory, A. C., et al. (2019). Global trends in marine plankton diversity across kingdoms of life. *Cell*, *179*(5), 1084–1097.
- IPBES. (2018, April). *The IPBES regional assessment report on biodiversity and ecosystem services for the Americas*. (J. Rice, C. S. Seixas, M. E. Zaccagnini, M. Bedoya-Gaitán, & V. Natalia, Eds.).
- IPBES. (2019, April). *Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* (E. Brondizio, S. Diaz, J. Settele, & H. T. Ngo, Eds.).
- Jacob, S., Sallé, L., Zinger, L., Chaine, A. S., Ducamp, C., Boutault, L., Russell, A. F., & Heeb, P. (2018). Chemical regulation of body feather microbiota in a wild bird. *Molecular Ecology*, *27*(7), 1727–1738.
- Jakovac, C. C., Junqueira, A. B., Crouzeilles, R., Peña-Claros, M., Mesquita, R. C., & Bongers, F. (2021). The role of land-use history in driving successional pathways and its implications for the restoration of tropical forests. *Biological Reviews*, *96*(4), 1114–1134.
- Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, *104*(940), 501–528.
- Jaureguiberry, P., Titeux, N., Wiemers, M., Bowler, D. E., Coscieme, L., Golden, A. S., Guerra, C. A., Jacob, U., Takahashi, Y., Settele, J., et al. (2022). The direct drivers of recent global anthropogenic biodiversity loss. *Science Advances*, *8*(45), eabm9982.
- Jenkins, D. G., Brescacin, C. R., Duxbury, C. V., Elliott, J. A., Evans, J. A., Grablow, K. R., Hillegass, M., Lyon, B. N., Metzger, G. A., Olandese, M. L., et al. (2007). Does size matter for dispersal distance? *Global Ecology and Biogeography*, *16*(4), 415–425.
- Jochum, M., Thouvenot, L., Ferlian, O., Zeiss, R., Klarner, B., Pruschitzki, U., Johnson, E. A., & Eisenhauer, N. (2022). Aboveground impacts of a belowground invader: How invasive earthworms alter aboveground arthropod communities in a northern north american forest. *Biological Letters*, *18*(3), 20210636.

- Johnson, J. S., Spakowicz, D. J., Hong, B.-Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., et al. (2019). Evaluation of 16s rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, *10*(1), 5029.
- Jønsson, K. A., Thomassen, E. E., Iova, B., Sam, K., & Thomsen, P. F. (2023). Using environmental DNA to investigate avian interactions with flowering plants. *Environmental DNA*, *5*(3), 462–475.
- Jørgensen, B. B., & Boetius, A. (2007). Feast and famine—microbial life in the deep-sea bed. *Nature Reviews Microbiology*, *5*(10), 770–781.
- Jurburg, S. D., Keil, P., Singh, B. K., & Chase, J. M. (2021). All together now: Limitations and recommendations for the simultaneous analysis of all eukaryotic soil sequences. *Molecular Ecology Resources*, *21*(6), 1759–1771.
- Jusselme, M.-D., Saccone, P., Zinger, L., Faure, M., Le Roux, X., Guillaumaud, N., Bernard, L., Clement, J.-C., & Poly, F. (2016). Variations in snow depth modify n-related soil microbial abundances and functioning during winter in subalpine grassland. *Soil Biology and Biochemistry*, *92*, 27–37.
- Kaltenpoth, M. (2009). Actinobacteria as mutualists: General healthcare for insects? *Trends in microbiology*, *17*(12), 529–535.
- Kardol, P., & Wardle, D. A. (2010). How understanding aboveground–belowground linkages can assist restoration ecology. *Trends in Ecology & Evolution*, *25*(11), 670–679.
- Karsenti, E., Acinas, S. G., Bork, P., Bowler, C., De Vargas, C., Raes, J., Sullivan, M., Arendt, D., Benzoni, F., Claverie, J.-M., et al. (2011). A holistic approach to marine eco-systems biology. *PLoS Biology*, *9*(10), e1001177.
- Kembel, S. W., Wu, M., Eisen, J. A., & Green, J. L. (2012). Incorporating 16s gene copy number information improves estimates of microbial diversity and abundance. *PLoS Computational Biology*, *8*(10), e1002743.
- Kirk, J., Beaudette, L., Hart, M., Moutoglis, P., Klironomos, J., Lee, H., & Trevors, J. (2004). Methods of studying soil microbial diversity. *Journal of Microbiological Methods*, *58*(2), 169–188.
- Knief, C. (2019). Diversity of methane-cycling microorganisms in soils and their relation to oxygen. *Current Issues in Molecular Biology*, *33*(1), 23–56.
- Koch, A., & Kaplan, J. O. (2022). Tropical forest restoration under future climate change. *Nature Climate Change*, *12*(3), 279–283.
- Kocher, A., De Thoisy, B., Catzeflis, F., Huguin, M., Valière, S., Zinger, L., Bañuls, A.-L., & Murienne, J. (2017). Evaluation of short mitochondrial metabarcodes for the identification of amazonian mammals. *Methods in Ecology and Evolution*, *8*(10), 1276–1283.
- Kocher, A., Gantier, J.-C., Gaborit, P., Zinger, L., Holota, H., Valiere, S., Dusfour, I., Girod, R., Bañuls, A.-L., & Murienne, J. (2017). Vector soup: High-throughput identification of neotropical phlebotomine sand flies using metabarcoding. *Molecular Ecology Resources*, *17*(2), 172–182.
- Kohl, K. D. (2020). Ecological and evolutionary mechanisms underlying patterns of phyllosymbiosis in host-associated microbial communities. *Philosophical Transactions of the Royal Society B*, *375*(1798), 20190251.
- Kopczynski, E. D., Bateson, M. M., & Ward, D. M. (1994). Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms. *Applied and Environmental Microbiology*, *60*(2), 746–748.
- Krehenwinkel, H., Kennedy, S., Pekár, S., & Gillespie, R. G. (2017). A cost-efficient and simple protocol to enrich prey DNA from extractions of predatory arthropods for large-scale gut content analysis by illumina sequencing. *Methods in Ecology and Evolution*, *8*(1), 126–134.
- Kunin, V., Engelbrekton, A., Ochman, H., & Hugenholtz, P. (2010). Wrinkles in the rare biosphere: Pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology*, *12*(1), 118–123.
- Kunkel, T. A., & Bebenek, K. (2000). DNA replication fidelity. *Annual Review of Biochemistry*, *69*(1), 497–529.
- Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology*, *28*(2), 420–430.
- Lancichinetti, A., & Fortunato, S. (2009). Community detection algorithms: A comparative analysis. *Physical review E*, *80*(5), 056117.

BIBLIOGRAPHY

- Lapola, D. M., Pinho, P., Barlow, J., Aragão, L. E., Berenguer, E., Carmenta, R., Liddy, H. M., Seixas, H., Silva, C. V., Silva-Junior, C. H., et al. (2023). The drivers and impacts of amazon forest degradation. *Science*, 379(6630), eabp8622.
- Lavelle, P., Decaens, T., Aubert, M., Barot, S., Blouin, M., Bureau, F., Margerie, P., Mora, P., & Rossi, J.-P. (2006). Soil invertebrates and ecosystem services. *European Journal of Soil Biology*, 42, Supplement 1(0), S3–S15.
- Lawton, J. (1999). Are there general laws in ecology? *Oikos*, 84(2), 177–192.
- Le Guillarme, N., Hedde, M., Potapov, A. M., Martínez-Muñoz, C. A., Berg, M. P., Briones, M. J., Calderón-Sanou, I., Degrune, F., Hohberg, K., Martínez-Almoyna, C., et al. (2023). The soil food web ontology: Aligning trophic groups, processes, resources, and dietary traits to support food-web research. *Ecological Informatics*, 78, 102360.
- Le Guillarme, N., & Thuiller, W. (2022). Taxonerd: Deep neural models for the recognition of taxonomic entities in the ecological and evolutionary literature. *Methods in Ecology and Evolution*, 13(3), 625–641.
- Le Provost, G., Thiele, J., Westphal, C., Penone, C., Allan, E., Neyret, M., Van Der Plas, F., Ayasse, M., Bardgett, R. D., Birkhofer, K., et al. (2021). Contrasting responses of above-and belowground diversity to multiple components of land-use intensity. *Nature Communications*, 12(1), 3918.
- Legendre, P. (2008). Studying beta diversity: Ecological variation partitioning by multiple regression and canonical analysis. *Journal of Plant Ecology*, 1(1), 3–8.
- Legrand, D., Guillaume, O., Baguette, M., Cote, J., Trochet, A., Calvez, O., Zajitschek, S., Zajitschek, F., Lecomte, J., Bénard, Q., et al. (2012). The metatron: An experimental system to study dispersal and metaecosystems for terrestrial organisms. *Nature Methods*, 9(8), 828–833.
- Leibold, M., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J., Hoopes, M., Holt, R., Shurin, J., Law, R., & Tilman, D. (2004). The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, 7(7), 601–613.
- Leibold, M. A., Govaert, L., Loeuille, N., De Meester, L., & Urban, M. C. (2022). Evolution and community assembly across spatial scales. *Annual Review of Ecology, Evolution, and Systematics*, 53, 299–326.
- Lemire, J. A., Harrison, J. J., & Turner, R. J. (2013). Antimicrobial activity of metals: Mechanisms, molecular targets and applications. *Nature Reviews Microbiology*, 11(6), 371–384.
- Lennon, J. T., & Jones, S. E. (2011). Microbial seed banks: The ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, 9(2), 119–130.
- Lentendu, G., Zinger, L., Manel, S., Coissac, E., Choler, P., Geremia, R. A., & Melodelima, C. (2011). Assessment of soil fungal diversity in different alpine tundra habitats by means of pyrosequencing. *Fungal Diversity*, 49, 113–123.
- Levin, S. A. (1992). The problem of pattern and scale in ecology: The robert h. macarthur award lecture. *Ecology*, 73(6), 1943–1967.
- Levy, R., & Borenstein, E. (2012). Reverse ecology: From systems to environments and back. *Advances in Experimental Medicine and Biology*, 751, 329–345.
- Lewitus, E., Bittner, L., Malviya, S., Bowler, C., & Morlon, H. (2018). Clade-specific diversification dynamics of marine diatoms since the jurassic. *Nature Ecology & Evolution*, 2(11), 1715–1723.
- Lloréns-Rico, V., Vieira-Silva, S., Gonçalves, P. J., Falony, G., & Raes, J. (2021). Benchmarking microbiome transformations favors experimental quantitative approaches to address compositionality and sampling depth biases. *Nature Communications*, 12(1), 3562.
- Locey, K. J., & Lennon, J. T. (2016). Scaling laws predict global microbial diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 113(21), 5970–5975.
- Lofgren, L. A., Uehling, J. K., Branco, S., Bruns, T. D., Martin, F., & Kennedy, P. G. (2019). Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and ecological lifestyles. *Molecular Ecology*, 28(4), 721–730.
- Louca, S., Mazel, F., Doebeli, M., & Parfrey, L. (2019). A census-based estimate of earth’s bacterial and archaeal diversity. *PLoS Biology*, 17(2), e3000106.
- Lu, M., & Jetz, W. (2023). Scale-sensitivity in the measurement and interpretation of environmental niches. *Trends in Ecology & Evolution*, 38(6), 554–567.

- Lynggaard, C., Bertelsen, M. F., Jensen, C. V., Johnson, M. S., Frøslev, T. G., Olsen, M. T., & Bohmann, K. (2022). Airborne environmental DNA for terrestrial vertebrate community monitoring. *Current Biology*, *32*(3), 701–707.
- MacArthur, R., & Wilson, E. (1967). *The theory of island biogeography*. Princeton University Press.
- Macher, T.-H., Schütz, R., Hörren, T., Beermann, A. J., & Leese, F. (2023). It's raining species: Rainwash eDNA metabarcoding as a minimally invasive method to assess tree canopy invertebrate diversity. *Environmental DNA*, *5*(1), 3–11.
- Macke, E., Tasiemski, A., Massol, F., Callens, M., & Decaestecker, E. (2017). Life history and eco-evolutionary dynamics in light of the gut microbiota. *Oikos*, *126*(4), 508–531.
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2014). Swarm: Robust and fast clustering method for amplicon-based studies. *PeerJ*, *2*, e593.
- Marguerat, S., & Bähler, J. (2012). Coordinating genome expression with cell size. *Trends in Genetics*, *28*(11), 560–565.
- Martin-Laurent, F., Philippot, L., Hallet, S., Chaussod, R., Germon, J., Soulas, G., & Catroux, G. (2001). DNA extraction from soils: Old bias for new microbial diversity analysis methods. *Applied and Environmental Microbiology*, *67*(5), 2354–2359.
- Martiny, J., Bohannan, B., Brown, J., Colwell, R., Fuhrman, J., Green, J., Horner-Devine, M., Kane, M., Krumins, J., Kuske, C., Morin, P., Naeem, S., Ovreås, L., Reysenbach, A.-L., Smith, V., & Staley, J. (2006). Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology*, *4*(2), 102–112.
- Martiny, J. B., Jones, S. E., Lennon, J. T., & Martiny, A. C. (2015). Microbiomes in light of traits: A phylogenetic perspective. *Science*, *350*(6261), aac9323.
- Martiny, J. B., Martiny, A. C., Brodie, E., Chase, A. B., Rodríguez-Verdugo, A., Treseder, K. K., & Allison, S. D. (2023). Investigating the eco-evolutionary response of microbiomes to environmental change. *Ecology Letters*, *26*, S81–S90.
- Matsen, F. A., Kodner, R. B., & Armbrust, E. V. (2010). Pplacer: Linear time maximum-likelihood and bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics*, *11*, 1–16.
- Matthews, T. J., & Whittaker, R. J. (2015). On the species abundance distribution in applied ecology and biodiversity management. *Journal of Applied Ecology*, *52*(2), 443–454.
- Mayr, E. (1992). Darwin's principle of divergence. *Journal of the History of Biology*, *25*(3), 343–359.
- Mazel, F., Davis, K. M., Loudon, A., Kwong, W. K., Groussin, M., & Parfrey, L. W. (2018). Is host filtering the main driver of phyllosymbiosis across the tree of life? *Msystems*, *3*(5), 10–1128.
- McDonald, R. I., Mansur, A. V., Ascensão, F., Colbert, M., Crossman, K., Elmquist, T., Gonzalez, A., Güneralp, B., Haase, D., Hamann, M., et al. (2020). Research gaps in knowledge of the impact of urban growth on biodiversity. *Nature Sustainability*, *3*(1), 16–24.
- McQuatters-Gollop, A., Johns, D. G., Bresnan, E., Skinner, J., Rombouts, I., Stern, R., Aubert, A., Johansen, M., Bedford, J., & Knights, A. (2017). From microscope to management: The critical value of plankton taxonomy to marine policy and biodiversity conservation. *Marine Policy*, *83*, 1–10.
- Mercier, C., Boyer, F., Bonin, A., & Coissac, E. (2013). Sumatra and sumacrust: Fast and exact comparison and clustering of sequences. *Programs and Abstracts of the SeqBio 2013 workshop. Abstract*, 27–29.
- Metzker, M. (2009). Sequencing technologies — the next generation. *Nature Reviews Genetics*, *11*(1), 31–46.
- Miller, E. T., Svanbäck, R., & Bohannan, B. J. (2018). Microbiomes as metacommunities: Understanding host-associated microbes through metacommunity ecology. *Trends in Ecology & Evolution*, *33*(12), 926–935.
- Mock, F., Kretschmer, F., Kriese, A., Böcker, S., & Marz, M. (2022). Taxonomic classification of DNA sequences beyond sequence similarity using deep neural networks. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(35), e2122636119.
- Modin, O., Liébana, R., Saheb-Alam, S., Wilén, B.-M., Suarez, C., Hermansson, M., & Persson, F. (2020). Hill-based dissimilarity indices and null models for analysis of microbial community assembly. *Microbiome*, *8*, 1–16.
- Moinard, S. (2023). *Estimateurs de biodiversité robustes pour le métabarcoding* [Doctoral dissertation] [s247636].

BIBLIOGRAPHY

- Montanarella, L., MARMO, L., MIKO, L., Ritz, K., Peres, G., Römbke, J., & Van der Putten, W. (2010). European atlas of soil biodiversity. *European Commission, Publications Office of the European Union, Luxembourg*.
- Moore, J. W., & Schindler, D. E. (2022). Getting ahead of climate change for ecological adaptation and resilience. *Science*, *376*(6600), 1421–1426.
- Mora, C., Tittensor, D., Adl, S., Simpson, A., & Worm, B. (2011). How many species are there on earth and in the ocean? *PLoS Biology*, *9*(8), e1001127.
- Moreno-Mateos, D., Alberdi, A., Morriën, E., van der Putten, W. H., Rodríguez-Uña, A., & Montoya, D. (2020). The long-term restoration of ecosystem complexity. *Nature Ecology & Evolution*, *4*(5), 676–685.
- Mori, A. S., Isbell, F., & Seidl, R. (2018). β -diversity, community assembly, and ecosystem functioning. *Trends in Ecology & Evolution*, *33*(7), 549–564.
- Morlon, H., White, E. P., Etienne, R. S., Green, J. L., Ostling, A., Alonso, D., Enquist, B. J., He, F., Hurlbert, A., Magurran, A. E., et al. (2009). Taking species abundance distributions beyond individuals. *Ecology Letters*, *12*(6), 488–501.
- Mouhamadou, B., Puissant, J., Personeni, E., Desclos-Theveniau, M., Kastl, E., Schloter, M., Zinger, L., Roy, J., Geremia, R., & Lavorel, S. (2013). Effects of two grass species on the composition of soil fungal communities. *Biology and Fertility of Soils*, *49*, 1131–1139.
- Mueller, U., & Sachs, J. (2015). Engineering microbiomes to improve plant and animal health. *Trends in Microbiology*, *23*(10), 606–617.
- Naeem, S., Duffy, J. E., & Zavaleta, E. (2012). The functions of biological diversity in an age of extinction. *Science*, *336*(6087), 1401–1406.
- Nagler, M., Insam, H., Pietramellara, G., & Ascher-Jenull, J. (2018). Extracellular DNA in natural environments: Features, relevance and applications. *Applied Microbiology and Biotechnology*, *102*, 6343–6356.
- Nagler, M., Podmirseg, S. M., Griffith, G. W., Insam, H., & Ascher-Jenull, J. (2018). The use of extracellular DNA as a proxy for specific microbial activity. *Applied Microbiology and Biotechnology*, *102*, 2885–2898.
- Nakamura, A., Kitching, R. L., Cao, M., Creedy, T. J., Fayle, T. M., Freiberg, M., Hewitt, C., Itioka, T., Koh, L. P., Ma, K., et al. (2017). Forests and their canopies: Achievements and horizons in canopy science. *Trends in Ecology & Evolution*, *32*(6), 438–451.
- Nearing, J. T., Douglas, G. M., Comeau, A. M., & Langille, M. G. (2018). Denoising the denoisers: An independent evaluation of microbiome sequence error-correction approaches. *PeerJ*, *6*, e5364.
- Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, *48*(3), 443–453.
- Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P., et al. (2013). Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews*, *77*(3), 342–356.
- Neyret, M., Le Provost, G., Boesing, A. L., Schneider, F. D., Baulechner, D., Bergmann, J., de Vries, F. T., Fiore-Donno, A. M., Geisen, S., Goldmann, K., et al. (2024). A slow-fast trait continuum at the whole community level in relation to land-use intensification. *Nature Communications*, *15*(1), 1251.
- Nichols, R. V., Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., Green, R. E., & Shapiro, B. (2018). Minimizing polymerase biases in metabarcoding. *Molecular Ecology Resources*, *18*(5), 927–939.
- Nilsson, R. H., Kristiansson, E., Ryberg, M., Hallenberg, N., & Larsson, K.-H. (2008). Intraspecific its variability in the kingdom fungi as expressed in the international sequence databases and its implications for molecular species identification. *Evolutionary Bioinformatics*, *4*, EBO–S653.
- O’Gorman, E. J., Zhao, L., Kordas, R. L., Dudgeon, S., & Woodward, G. (2023). Warming indirectly simplifies food webs through effects on apex predators. *Nature Ecology & Evolution*, *7*(12), 1983–1992.
- Ogram, A., Sayler, G., & Barkay, T. (1987). The extraction and purification of microbial DNA from sediments. *Journal of Microbiological Methods*, *7*(2-3), 57–66.
- Ohlmann, M., Mazel, F., Chalmandrier, L., Bec, S., Coissac, E., Gielly, L., Pansu, J., Schilling, V., Taberlet, P., Zinger, L., et al. (2018). Mapping the imprint of biotic interactions on β -diversity. *Ecology Letters*, *21*(11), 1660–1669.

- Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O’connor, L., & Thuiller, W. (2019). Diversity indices for ecological networks: A unifying framework using hill numbers. *Ecology Letters*, *22*(4), 737–747.
- Okie, J. G., Smith, V. H., & Martin-Cereceda, M. (2016). Major evolutionary transitions of life, metabolic scaling and the number and size of mitochondria and chloroplasts. *Proceedings of the Royal Society B*, *283*(1831), 20160611.
- Oliveira, B. B., Veigas, B., & Baptista, P. V. (2021). Isothermal amplification of nucleic acids: The race for the next “gold standard”. *Frontiers in Sensors*, *2*, 752600.
- O’Malley, M. (2007). The nineteenth century roots of ‘everything is everywhere’. *Nature Reviews Microbiology*, *5*(8), 647–651.
- O’Malley, M. (2009). What did darwin say about microbes, and how did microbiology respond? *Trends in Microbiology*, *17*(8), 341–347.
- Papadopoulou, A., Taberlet, P., & Zinger, L. (2015). Metagenome skimming for phylogenetic community ecology: A new era in biodiversity research. *Molecular Ecology*, *24*(14), 3515–3517.
- Patin, N. V., Kunin, V., Lidström, U., & Ashby, M. N. (2013). Effects of otu clustering and PCR artifacts on microbial diversity estimates. *Microbial Ecology*, *65*, 709–719.
- Pauvert, C., Buée, M., Laval, V., Edel-Hermann, V., Fauchery, L., Gautier, A., Lesur, I., Vallance, J., & Vacher, C. (2019). Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal Ecology*, *41*, 23–33.
- Pawlowski, J., Apothéoz-Perret-Gentil, L., & Altermatt, F. (2020). Environmental DNA: What’s behind the term? clarifying the terminology and recommendations for its future use in biomonitoring. *Molecular Ecology*, *29*, 4258–4264.
- Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S., Cepicka, I., Decelle, J., Dunthorn, M., Fiore-Donno, A., Gile, G., Holzmann, M., Jahn, R., Jirků, M., Keeling, P., Kostka, M., Kudryavtsev, A., Lara, E., ... de Vargas, C. (2012). Cbol protist working group: Barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS Biology*, *10*(11), e1001419.
- Pawlowski, J., Apothéoz-Perret-Gentil, L., & Altermatt, F. (2021). Environmental versus extra-organismal DNA. *Molecular Ecology*, *30*(19), 4606.
- Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéoz-Perret-Gentil, L., Beja, P., Boggero, A., Borja, A., Bouchez, A., Cordier, T., Domaizon, I., et al. (2018). The future of biotic indices in the ecogenomic era: Integrating (e) DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total Environment*, *637*, 1295–1310.
- Peay, K. G., Baraloto, C., & Fine, P. V. (2013). Strong coupling of plant and fungal community structure across western amazonian rainforests. *The ISME journal*, *7*(9), 1852–1861.
- Peguero, G., Ferrín, M., Sardans, J., Verbruggen, E., Ramírez-Rojas, I., Van Langenhove, L., Verryckt, L. T., Murienne, J., Iribar, A., Zinger, L., et al. (2022). Decay of similarity across tropical forest communities: Integrating spatial distance with soil nutrients. *Ecology*, *103*(2), e03599.
- Pellerin, F., Bestion, E., Winandy, L., Di Gesu, L., Richard, M., Aguilée, R., & Cote, J. (2022). Connectivity among thermal habitats buffers the effects of warm climate on life-history traits and population dynamics. *Journal of Animal Ecology*, *91*(11), 2301–2313.
- Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H., Cardoso, A., et al. (2013). Essential biodiversity variables. *Science*, *339*(6117), 277–278.
- Petrolli, R., Zinger, L., Perez-Lamarque, B., Collobert, G., Griveau, C., Pailler, T., Selosse, M.-A., & Martos, F. (2022). Spatial turnover of fungi and partner choice shape mycorrhizal networks in epiphytic orchids. *Journal of Ecology*, *110*(11), 2568–2584.
- Philippot, L., Ritz, K., Pandard, P., Hallin, S., & Martin-Laurent, F. (2012). Standardisation of methods in soil microbiology: Progress and challenges. *FEMS Microbiology Ecology*, *82*(1), 1–10.
- Pierella Karlusich, J. J., Pelletier, E., Zinger, L., Lombard, F., Zingone, A., Colin, S., Gasol, J. M., Dorrell, R. G., Henry, N., Scalco, E., et al. (2023). A robust approach to estimate relative phytoplankton cell abundances from metagenomes. *Molecular Ecology Resources*, *23*(1), 16–40.
- Piton, G., Allison, S. D., Bahram, M., Hildebrand, F., Martiny, J. B., Treseder, K. K., & Martiny, A. C. (2023). Life history strategies of soil bacterial communities across global terrestrial biomes. *Nature Microbiology*, *8*(11), 2093–2102.

BIBLIOGRAPHY

- Pocheville, A. (2015). The ecological niche: History and recent controversies. In *Handbook of evolutionary thinking in the sciences* (pp. 547–586). Springer.
- Poelen, J. H., Simons, J. D., & Mungall, C. J. (2014). Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets. *Ecological informatics*, *24*, 148–159.
- Pölme, S., Abarenkov, K., Henrik Nilsson, R., Lindahl, B. D., Clemmensen, K. E., Kauserud, H., Nguyen, N., Kjølter, R., Bates, S. T., Baldrian, P., et al. (2020). Fungaltraits: A user-friendly traits database of fungi and fungus-like stramenopiles. *Fungal Diversity*, *105*, 1–16.
- Polz, M. F., & Cavanaugh, C. M. (1998). Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*, *64*(10), 3724–3730.
- Pommier, T., Pinhassi, J., & Hagström, Å. (2005). Biogeographic analysis of ribosomal RNA clusters from marine bacterioplankton. *Aquatic Microbial Ecology*, *41*(1), 79–89.
- Pornon, A., Escaravage, N., Burrus, M., Holota, H., Khimoun, A., Mariette, J., Pellizzari, C., Iribar, A., Etienne, R., Taberlet, P., et al. (2016). Using metabarcoding to reveal and quantify plant-pollinator interactions. *Scientific Reports*, *6*(1), 27282.
- Potapov, A. M. (2022). Multifunctionality of belowground food webs: Resource, size and spatial energy channels. *Biological Reviews*, *97*(4), 1691–1711.
- Potapov, A. M., Drescher, J., Darras, K., Wenzel, A., Janotta, N., Nazarreta, R., Kasmiatun, Laurent, V., Mawan, A., Utari, E. H., et al. (2024). Rainforest transformation reallocates energy from green to brown food webs. *Nature*, 1–7.
- Potapov, A. M., Rozanova, O. L., Semenina, E. E., Leonov, V. D., Belyakova, O. I., Bogatyreva, V. Y., Degtyarev, M. I., Esaulov, A. S., Korotkevich, A. Y., Kudrin, A. A., et al. (2021). Size compartmentalization of energy channeling in terrestrial belowground food webs. *Ecology*, *102*(8), e03421.
- Prober, S. M., Leff, J. W., Bates, S. T., Borer, E. T., Firn, J., Harpole, W. S., Lind, E. M., Seabloom, E. W., Adler, P. B., Bakker, J. D., et al. (2015). Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecology Letters*, *18*(1), 85–95.
- Prosser, J., Bohannan, B., Curtis, T., Ellis, R., Firestone, M., Freckleton, R., Green, J., Green, L., Killham, K., Lennon, J., Osborn, A. M., Solan, M., van der Gast, C., & Young, J. (2007). The role of ecological theory in microbial ecology. *Nature Reviews Microbiology*, *5*(5), 384–392.
- Pugnaire, F. I., Morillo, J. A., Peñuelas, J., Reich, P. B., Bardgett, R. D., Gaxiola, A., Wardle, D. A., & Van Der Putten, W. H. (2019). Climate change effects on plant-soil feedbacks and consequences for biodiversity and functioning of terrestrial ecosystems. *Science Advances*, *5*(11), eaaz1834.
- Puşçaş, M., Choler, P., Tribsch, A., Gielly, L., Rioux, D., Gaudeul, M., & Taberlet, P. (2008). Post-glacial history of the dominant alpine sedge *Carex curvula* in the European alpine system inferred from nuclear and chloroplast markers. *Molecular Ecology*, *17*(10), 2417–2429.
- Quince, C., Lanzén, A., Curtis, T. P., Davenport, R. J., Hall, N., Head, I. M., Read, L. F., & Sloan, W. T. (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. *Nature Methods*, *6*(9), 639–641.
- Rådström, P., Knutsson, R., Wolffs, P., Lövenklev, M., & Löfström, C. (2004). Pre-PCR processing: Strategies to generate PCR-compatible samples. *Molecular Biotechnology*, *26*, 133–146.
- Ranjard, L., Lejon, D. P., Mougel, C., Schehrer, L., Merdinoglu, D., & Chaussod, R. (2003). Sampling strategy in molecular microbial ecology: Influence of soil sample size on DNA fingerprinting analysis of fungal and bacterial communities. *Environmental Microbiology*, *5*(11), 1111–1120.
- Ricklefs, R. (1987). Community diversity: Relative roles of local and regional processes. *Science*, *235*(4785), 167–171.
- Rillig, M. C., Van der Heijden, M. G., Berdugo, M., Liu, Y.-R., Riedo, J., Sanz-Lazaro, C., Moreno-Jiménez, E., Romero, F., Tedersoo, L., & Delgado-Baquerizo, M. (2023). Increasing the number of stressors reduces soil ecosystem services worldwide. *Nature Climate Change*, *13*(5), 478–483.
- Robinson, J. M., Liddicoat, C., Muñoz-Rojas, M., & Breed, M. F. (2024). Restoring soil biodiversity. *Current Biology*, *34*(9), R393–R398.
- Rodriguez-Ezpeleta, N., Morissette, O., Bean, C. W., Manu, S., Banerjee, P., Lacoursière-Roussel, A., Beng, K. C., Alter, S. E., Roger, F., Holman, L. E., et al. (2021). Trade-offs between reducing complex terminology and producing accurate interpretations from environmental DNA: Comment on “environmental DNA: What’s behind the term?” by pawlowski et al.,(2020). *Molecular Ecology*, *30*(19), 4601–4605.

- Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). Vsearch: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584.
- Roller, B. R., Stoddard, S. F., & Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nature Microbiology*, *1*(11), 1–7.
- Romashchenko, N., Linard, B., Pardi, F., & Rivals, E. (2023). Epik: Precise and scalable evolutionary placement with informative k-mers. *Bioinformatics*, *39*(12), btad692.
- Rosenzweig, M. (1995). *Species diversity in space and time*. Cambridge University Press.
- Rosselló-Mora, R., & Amann, R. (2001). The species concept for prokaryotes. *FEMS Microbiology Reviews*, *25*(1), 39–67.
- Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(4), 1118–1123.
- Roy, J., Albert, C. H., Choler, P., Clément, J.-C., Ibanez, S., Saccone, P., Zinger, L., & Geremia, R. A. (2013). Microbes on the cliff: Alpine cushion plants structure bacterial and fungal communities. *Frontiers in Microbiology*, *4*, 40260.
- Roy, J., Bonneville, J.-M., Saccone, P., Ibanez, S., Albert, C. H., Boleda, M., Gueguen, M., Ohlmann, M., Rioux, D., Clément, J.-C., et al. (2018). Differences in the fungal communities nursed by two genetic groups of the alpine cushion plant, *silene acaulis*. *Ecology and Evolution*, *8*(23), 11568–11581.
- Roy, J., Mazel, F., Sosa-Hernández, M. A., Dueñas, J. F., Hempel, S., Zinger, L., & Rillig, M. C. (2019). The relative importance of ecological drivers of arbuscular mycorrhizal fungal distribution varies with taxon phylogenetic resolution. *New Phytologist*, *224*(2), 936–948.
- Rozendaal, D. M., Bongers, F., Aide, T. M., Alvarez-Dávila, E., Ascarrunz, N., Balvanera, P., Becknell, J. M., Bentos, T. V., Brancalion, P. H., Cabral, G. A., et al. (2019). Biodiversity recovery of neotropical secondary forests. *Science Advances*, *5*(3), eaau3114.
- Rudolf, V. H., & Roman, A. (2018). Trophic structure alters consequences of environmental warming. *Oikos*, *127*(11), 1646–1656.
- Saiki, R. K., Gelfand, D., Stoffel, S., Scharf, S., Higuchi, R., Horn, G., Mullis, K., & Erlich, H. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, *239*(4839), 487–491.
- Sales, L. P., Galetti, M., & Pires, M. M. (2020). Climate and land-use change will lead to a faunal “savanization” on tropical rainforests. *Global Change Biology*, *26*(12), 7036–7044.
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J., & Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, *12*, 1–12.
- Schimann, H., Vacher, C., Coste, S., Louisanna, E., Fort, T., & Zinger, L. (2023). Determinants of the vertical distribution of the phyllosphere differ between microbial groups and the epi-and endosphere in a tropical forest. *Phytobiomes Journal*, *7*(3), 312–323.
- Schloss, P. D. (2021). Amplicon sequence variants artificially split bacterial genomes into separate clusters. *Msphere*, *6*(4), 10–1128.
- Schloss, P. D., & Handelsman, J. (2005). Introducing dotur, a computer program for defining operational taxonomic units and estimating species richness. *Applied and Environmental Microbiology*, *71*(3), 1501–1506.
- Schmidt, S., Costello, E., Nemergut, D., Cleveland, C. C., Reed, S., Weintraub, M., Meyer, A., & Martin, A. (2007). Biogeochemical consequences of rapid microbial turnover and seasonal succession in soil. *Ecology*, *88*(6), 1379–1385.
- Schnell, I. B., Bohmann, K., & Gilbert, M. T. P. (2015). Tag jumps illuminated—reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*, *15*(6), 1289–1303.
- Schoch, C., Seifert, K., Huhndorf, S., Robert, V., Spouge, J., Levesque, C., Chen, W., & the Fungal Barcoding Consortium. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for fungi. *Proceedings of the National Academy of Science of the United States of America*, *109*(16), 6241–6246.
- Sepulveda, J., & Moeller, A. H. (2020). The effects of temperature on animal gut microbiomes. *Frontiers in Microbiology*, *11*, 512773.

BIBLIOGRAPHY

- Shade, A., Dunn, R. R., Blowes, S. A., Keil, P., Bohannan, B. J., Herrmann, M., Küsel, K., Lennon, J. T., Sanders, N. J., Storch, D., et al. (2018). Macroecology to unite all life, large and small. *Trends in Ecology & Evolution*, *33*(10), 731–744.
- Shahnavaz, B., Zinger, L., Lavergne, S., Choler, P., & Geremia, R. A. (2012). Phylogenetic clustering reveals selective events driving the turnover of bacterial community in alpine tundra soils. *Arctic, Antarctic, and Alpine Research*, *44*(2), 232–238.
- Shapira, M. (2016). Gut microbiotas and host evolution: Scaling up symbiosis. *Trends in Ecology & Evolution*, *31*(7), 539–549.
- Shen, E. W., Vandenberg, J. M., & Moore, A. (2023). Sensing inequity: Technological solutionism, biodiversity conservation, and environmental DNA. *BioSocieties*, 1–25.
- Shinohara, N., Nakadai, R., Suzuki, Y., & Terui, A. (2023). Spatiotemporal dimensions of community assembly. *Population Ecology*, *65*(1), 5–16.
- Shoemaker, W. R., Locey, K. J., & Lennon, J. T. (2017). A macroecological theory of microbial biodiversity. *Nature Ecology & Evolution*, *1*(5), 0107.
- Sigsgaard, E. E., Jensen, M. R., Winkelmann, I. E., Møller, P. R., Hansen, M. M., & Thomsen, P. F. (2020). Population-level inferences from environmental DNA—current status and future perspectives. *Evolutionary Applications*, *13*(2), 245–262.
- Simmons, B. I., Blyth, P. S., Blanchard, J. L., Clegg, T., Delmas, E., Garnier, A., Griffiths, C. A., Jacob, U., Pennekamp, F., Petchey, O. L., et al. (2021). Refocusing multiple stressor research around the targets and scales of ecological impacts. *Nature Ecology & Evolution*, *5*(11), 1478–1489.
- Simon, J.-C., Marchesi, J. R., Mougél, C., & Selosse, M.-A. (2019). Host-microbiota interactions: From holobiont theory to analysis. *Microbiome*, *7*, 1–5.
- Simon, U. K., & Weiß, M. (2008). Intragenomic variation of fungal ribosomal genes is higher than previously thought. *Molecular Biology and Evolution*, *25*(11), 2251–2254.
- Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., & Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental microbiology*, *8*(4), 732–740.
- Smith, T., & Lundholm, J. (2010). Variation partitioning as a tool to distinguish between niche and neutral processes. *Ecography*, *33*(4), 648–655.
- Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., Arrieta, J. M., & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(32), 12115–12120.
- Soininen, J. (2011). Macroecology of unicellular organisms - patterns and processes. *Environmental Microbiology Reports*, *4*(1), 10–22.
- Soininen, J., McDonald, R., & Hillebrand, H. (2007). The distance decay of similarity in ecological communities. *Ecography*, *30*(1), 3–12.
- Soliveres, S., Van Der Plas, F., Manning, P., Prati, D., Gossner, M. M., Renner, S. C., Alt, F., Arndt, H., Baumgartner, V., Binkenstein, J., et al. (2016). Biodiversity at multiple trophic levels is needed for ecosystem multifunctionality. *Nature*, *536*(7617), 456–459.
- Sommer, U., Peter, K. H., Genitsaris, S., & Moustaka-Gouni, M. (2017). Do marine phytoplankton follow Bergmann’s rule sensu lato? *Biological Reviews*, *92*(2), 1011–1026.
- Sommeria-Klein, G., Watteaux, R., Ibarbalz, F. M., Pierella Karlusich, J. J., Iudicone, D., Bowler, C., & Morlon, H. (2021). Global drivers of eukaryotic plankton biogeography in the sunlit ocean. *Science*, *374*(6567), 594–599.
- Sommeria-Klein, G., Zinger, L., Coissac, E., Iribar, A., Schimann, H., Taberlet, P., & Chave, J. (2020). Latent dirichlet allocation reveals spatial and taxonomic structure in a DNA-based census of soil biodiversity from a tropical forest. *Molecular Ecology Resources*, *20*(2), 371–386.
- Sommeria-Klein, G., Zinger, L., Taberlet, P., Coissac, E., & Chave, J. (2016). Inferring neutral biodiversity parameters using environmental DNA data sets. *Scientific Reports*, *6*(1), 35644.
- Srednick, S., G. and Swearer. (2024). Understanding diversity–synchrony–stability relationships in multi-trophic communities. *Nature Ecology & Evolution*, *in press*.
- Stackebrandt, E., & Goebel, B. (1994). Taxonomic note: A place for DNA-DNA reassociation and 16s rRNA sequence analysis in the present species definition in bacteriology. *International journal of systematic and evolutionary microbiology*, *44*(4), 846–849.

- Staley, J., & Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology*, *39*(1), 321–346.
- Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., Rockhold, M. L., & Konopka, A. (2013). Quantifying community assembly processes and identifying features that impose them. *The ISME journal*, *7*(11), 2069–2079.
- Stork, N. E. (2018). How many species of insects and other terrestrial arthropods are there on earth? *Annual Review of Entomology*, *63*, 31–45.
- Stroud, J., Delory, B., Barnes, E., Chase, J., De Meester, L., Dieskau, J., Grainger, T., Halliday, F., Kardol, P., Knight, T., et al. (2024). Priority effects transcend scales and disciplines in biology. *Trends in Ecology & Evolution*.
- Sze, M. A., & Schloss, P. D. (2019). The impact of DNA polymerase and number of rounds of amplification in PCR on 16s rRNA gene sequence data. *Msphere*, *4*(3), 10–1128.
- Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. (2012). Environmental DNA. *Molecular Ecology*, *21*(8), 1789–1793.
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–2050.
- Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA: For biodiversity research and monitoring*. Oxford University Press.
- Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermet, T., Corthier, G., Brochmann, C., & Willerslev, E. (2007). Power and limitations of the chloroplast trn l (uaa) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*(3), e14–e14.
- Taberlet, P., Prud'homme, S. M., Campione, E., Roy, J., Miquel, C., Shehzad, W., Gielly, L., Rioux, D., Choler, P., Clément, J.-C., et al. (2012). Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Molecular Ecology*, *21*(8), 1816–1820.
- Tanaka, T., Kawasaki, K., Daimon, S., Kitagawa, W., Yamamoto, K., Tamaki, H., Tanaka, M., Nakatsu, C., & Kamagata, Y. (2014). A hidden pitfall in the preparation of agar media undermines microorganism cultivability. *Applied and Environmental Microbiology*, *80*(24), 7659–7666.
- Tedersoo, L., Albertsen, M., Anslan, S., & Callahan, B. (2021). Perspectives and benefits of high-throughput long-read sequencing in microbial ecology. *Applied and Environmental Microbiology*, *87*(17), e00626–21.
- Tedersoo, L., Bahram, M., Zinger, L., Nilsson, R. H., Kennedy, P. G., Yang, T., Anslan, S., & Mikryukov, V. (2022). Best practices in metabarcoding of fungi: From experimental design to results. *Molecular Ecology*, *31*(10), 2769–2795.
- Tedersoo, L., Nilsson, R. H., Abarenkov, K., Jairus, T., Sadam, A., Saar, I., Bahram, M., Bechem, E., Chuyong, G., & Kõljalg, U. (2010). 454 pyrosequencing and sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist*, *188*(1), 291–301.
- Tedersoo, L., Tooming-Klunderud, A., & Anslan, S. (2018). Pacbio metabarcoding of fungi and other eukaryotes: Errors, biases and perspectives. *New Phytologist*, *217*(3), 1370–1385.
- Terborgh, J. (2020). At 50, Janzen–Connell has come of age. *BioScience*, *70*(12), 1082–1092.
- Terrat, S., Christen, R., Dequiedt, S., Lelièvre, M., Nowak, V., Regnier, T., Bachar, D., Plassart, P., Wincker, P., Jolivet, C., et al. (2012). Molecular biomass and metataxogenomic assessment of soil microbial communities as influenced by soil DNA extraction procedure. *Microbial Biotechnology*, *5*(1), 135–141.
- The Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, *486*(7402), 207–214.
- Thompson, J. R., Marcelino, L. A., & Polz, M. F. (2002). Heteroduplexes in mixed-template amplifications: Formation, consequence and elimination by ‘reconditioning PCR’. *Nucleic Acids Research*, *30*(9), 2083–2088.
- Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., et al. (2017). A communal catalogue reveals earth’s multiscale microbial diversity. *Nature*, *551*(7681), 457–463.
- Thomsen, P. F., Jensen, M. R., & Sigsgaard, E. E. (2024). A vision for global eDNA-based monitoring in a changing world. *Cell*.
- Tilman, D. (1994). Competition and biodiversity in spatially structured habitats. *Ecology*, *75*(1), 2–16.

BIBLIOGRAPHY

- Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., & Worm, B. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, *466*(7310), 1098–1101.
- Torsvik, V., Goksøyr, J., & Daae, F. (1990). High diversity in DNA of soil bacteria. *Applied and Environmental Microbiology*, *56*(3), 782–787.
- Trego, A., Keating, C., Nzeteu, C., Graham, A., O’Flaherty, V., & Ijaz, U. Z. (2022). Beyond basic diversity estimates—analytical tools for mechanistic interpretations of amplicon sequencing data. *Microorganisms*, *10*(10), 1961.
- Treseder, K. K., Berlemont, R., Allison, S. D., & Martiny, A. C. (2018). Drought increases the frequencies of fungal functional genes related to carbon and nitrogen acquisition. *PloS one*, *13*(11), e0206441.
- Trevelline, B. K., Fontaine, S. S., Hartup, B. K., & Kohl, K. D. (2019). Conservation biology needs a microbial renaissance: A call for the consideration of host-associated microbiota in wildlife management practices. *Proceedings of the Royal Society B*, *286*(1895), 20182448.
- Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., & Singh, B. K. (2020). Plant–microbiome interactions: From community assembly to plant health. *Nature Reviews Microbiology*, *18*(11), 607–621.
- Tuomisto, H., Ruokolainen, K., & Yli-Halla, M. (2003). Dispersal, environment, and floristic variation of western amazonian forests. *Science*, *299*(5604), 241–244.
- Ullah, H., Nagelkerken, I., Goldenberg, S. U., & Fordham, D. A. (2018). Climate change could drive marine food web collapse through altered trophic flows and cyanobacterial proliferation. *PLoS Biology*, *16*(1), e2003446.
- Ulrich, W., & Gotelli, N. J. (2010). Null model analysis of species associations using abundance data. *Ecology*, *91*(11), 3384–3397.
- Vacher, C., Hampe, A., Porté, A. J., Sauer, U., Compant, S., & Morris, C. E. (2016). The phyllosphere: Microbial jungle at the plant–climate interface. *Annual Review of Ecology, Evolution, and Systematics*, *47*, 1–24.
- Valentin, R. E., Kyle, K. E., Allen, M. C., Welbourne, D. J., & Lockwood, J. L. (2021). The state, transport, and fate of aboveground terrestrial arthropod eDNA. *Environmental DNA*, *3*(6), 1081–1092.
- Valverde-Barrantes, O. J., Freschet, G. T., Roumet, C., & Blackwood, C. B. (2017). A worldview of root traits: The influence of ancestry, growth form, climate and mycorrhizal association on the functional trait variation of fine-root tissues in seed plants. *New Phytologist*, *215*(4), 1562–1573.
- van den Berg, N. I., Machado, D., Santos, S., Rocha, I., Chacón, J., Harcombe, W., Mitri, S., & Patil, K. R. (2022). Ecological modelling approaches for predicting emergent properties in microbial communities. *Nature Ecology & Evolution*, *6*(7), 855–865.
- Van der Putten, W. H., Bardgett, R. D., Bever, J. D., Bezemer, T. M., Casper, B. B., Fukami, T., Kardol, P., Klironomos, J. N., Kulmatiski, A., Schweitzer, J. A., et al. (2013). Plant–soil feedbacks: The past, the present and future challenges. *Journal of Ecology*, *101*(2), 265–276.
- van der Heijden, M., Bardgett, R., & van Straalen, N. (2008). The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters*, *11*(3), 296–310.
- van der Putten, W. H., Bradford, M. A., Pernilla Brinkman, E., van de Voorde, T. F., & Veen, G. (2016). Where, when and how plant–soil feedback matters in a changing world. *Functional Ecology*, *30*(7), 1109–1121.
- Van Nuland, M. E., Qin, C., Pellitier, P. T., Zhu, K., & Peay, K. G. (2024). Climate mismatches with ectomycorrhizal fungi contribute to migration lag in north american tree range shifts. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(23), e230881121.
- Van Nuland, M. E., Wooliver, R. C., Pfennigwerth, A. A., Read, Q. D., Ware, I. M., Mueller, L., Fordyce, J. A., Schweitzer, J. A., & Bailey, J. K. (2016). Plant–soil feedbacks. *Functional Ecology*, *30*(7), 1032–1042.
- Vannier, N., Mony, C., Bittebiere, A.-K., Theis, K. R., Rosenberg, E., & Vandenkoornhuyse, P. (2019). Clonal plants as meta-holobionts. *MSystems*, *4*(2), 10–1128.
- Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, *85*(2), 183–206.
- Vellend, M., Srivastava, D., Anderson, K., Brown, C., Jankowski, J., Kleyhans, E., Kraft, N., Letaw, A., Macdonald, A., Maclean, J., Myers-Smith, I., Norris, A., & Xu, X. (2014). Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos*, *123*(12), 1420–1430.

- Venter, J., Remington, K., Heidelberg, J., Halpern, A., Rusch, D., Eisen, J., Wu, D., Paulsen, I., Nelson, K., Nelson, W., Fouts, D., Levy, S., Knap, A., Lomas, M., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., ... Smith, H. (2004). Environmental genome shotgun sequencing of the sargasso sea. *Science*, *304*(5667), 66–74.
- Viana, D. S., & Chase, J. M. (2019). Spatial scale modulates the inference of metacommunity assembly processes. *Ecology*, *100*(2), e02576.
- Vorkapic, D., Pressler, K., & Schild, S. (2016). Multifaceted roles of extracellular DNA in bacterial physiology. *Current Genetics*, *62*, 71–79.
- Waite, D. W., Chuvochina, M., Pelikan, C., Parks, D. H., Yilmaz, P., Wagner, M., Loy, A., Naganuma, T., Nakai, R., Whitman, W. B., et al. (2020). Proposal to reclassify the proteobacterial classes deltaproteobacteria and oligoflexia, and the phylum thermodesulfobacteria into four phyla reflecting major functional capabilities. *International Journal of Systematic and Evolutionary Microbiology*, *70*(11), 5972–6016.
- Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, *73*(16), 5261–5267.
- Wardle, D. A. (2006). The influence of biotic interactions on soil biodiversity. *Ecology Letters*, *9*(7), 870–886.
- Wardle, D. A., Bardgett, R. D., Klironomos, J. N., Setälä, H., Van Der Putten, W. H., & Wall, D. H. (2004). Ecological linkages between aboveground and belowground biota. *Science*, *304*(5677), 1629–1633.
- Werren, J. H., Windsor, D., & Guo, L. R. (1995). Distribution of wolbachia among neotropical arthropods. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *262*(1364), 197–204.
- Whitaker, R., Grogan, D., & Taylor, J. (2003). Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science*, *301*(5635), 976–978.
- Whitman, W., Coleman, D., & Wiebe, W. (1998). Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(12), 6578–6583.
- Whittaker, R. (1972). Evolution and measurement of species diversity. *Taxon*, *21*(2-3), 213–251.
- Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergård, M., Gussarova, G., Haile, J., et al. (2014). Fifty thousand years of arctic vegetation and megafaunal diet. *Nature*, *506*(7486), 47–51.
- Williams, D., & Huxley, R. (1998). Christian gottfried ehrenberg (1795-1876): The man and his legacy. an introduction. *Linnean*, *1*, 1–13.
- Williamson, K. E., Fuhrmann, J. J., Wommack, K. E., & Radosevich, M. (2017). Viruses in soil ecosystems: An unknown quantity within an unexplored territory. *Annual Review of Virology*, *4*, 201–219.
- Winding, A., Singh, B. K., Bach, E., Brown, G., Zhang, J., Cooper, M., Dion, P., Mele, P., Eisenhauer, N., Pena-Neira, S., et al. (2020). State of knowledge of soil biodiversity: Status, challenges, and potentialities.
- Woese, C., & Fox, G. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(11), 5088–5090.
- Woodcock, S., Curtis, T. P., Head, I. M., Lunn, M., & Sloan, W. T. (2006). Taxa–area relationships for microbes: The unsampled and the unseen. *Ecology Letters*, *9*(7), 805–812.
- Woodcock, S., Van Der Gast, C. J., Bell, T., Lunn, M., Curtis, T. P., Head, I. M., & Sloan, W. T. (2007). Neutral assembly of bacterial communities. *FEMS microbiology ecology*, *62*(2), 171–180.
- Woolley, S. N., Tittensor, D. P., Dunstan, P. K., Guillera-Arroita, G., Lahoz-Monfort, J. J., Wintle, B. A., Worm, B., & O’Hara, T. D. (2016). Deep-sea diversity patterns are shaped by energy availability. *Nature*, *533*(7603), 393–396.
- Worden, A. Z., Follows, M. J., Giovannoni, S. J., Wilken, S., Zimmerman, A. E., & Keeling, P. J. (2015). Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science*, *347*(6223), 1257594.
- Wright, E. S., Yilmaz, L. S., & Noguera, D. R. (2012). Decipher, a search-based approach to chimera identification for 16s rRNA sequences. *Applied and Environmental Microbiology*, *78*(3), 717–725.
- Yates, M. C., Glaser, D., Post, J. R., Cristescu, M. E., Fraser, D. J., & Derry, A. M. (2021). The relationship between eDNA particle concentration and organism abundance in nature is strengthened by allometric scaling. *Molecular Ecology*, *30*(13), 3068–3082.

BIBLIOGRAPHY

- Yoccoz, N. G., Bråthen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., von Stedingk, H., Brysting, A., Coissac, E., Pompanon, F., et al. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, *21*(15), 3647–3655.
- Zahra, S., Novotny, V., & Fayle, T. M. (2021). Do reverse Janzen-Connell effects reduce species diversity? *Trends in Ecology & Evolution*, *36*(5), 387–390.
- Zhou, J., Bruns, M., & Tiedje, J. (1996). DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology*, *62*(2), 316–22.
- Zhou, Z., Krashevskaya, V., Widyastuti, R., Scheu, S., & Potapov, A. (2022). Tropical land use alters functional diversity of soil food webs and leads to monopolization of the detrital energy channel. *Elife*, *11*, e75428.
- Zinger, L., Coissac, E., Choler, P., & Geremia, R. (2009). Assessment of microbial communities by graph partitioning in a study of soil fungi in two alpine meadows. *Applied and Environmental Microbiology*, *75*(18), 5863–5870.
- Zinger, L., Amaral-Zettler, L. A., Fuhrman, J. A., Horner-Devine, M. C., Huse, S. M., Welch, D. B. M., Martiny, J. B., Sogin, M., Boetius, A., & Ramette, A. (2011). Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PLoS one*, *6*(9), e24570.
- Zinger, L., Boetius, A., & Ramette, A. (2014). Bacterial taxa–area and distance–decay relationships in marine environments. *Molecular Ecology*, *23*(4), 954–964.
- Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., Chariton, A. A., Creer, S., Coissac, E., Deagle, B. E., et al. (2019). DNA metabarcoding—need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, *28*(8), 1857–1862.
- Zinger, L., Chave, J., Coissac, E., Iribar, A., Louisanna, E., Manzi, S., Schilling, V., Schimann, H., Sommeria-Klein, G., & Taberlet, P. (2016). Extracellular DNA extraction is a fast, cheap and reliable alternative for multi-taxa surveys based on soil DNA. *Soil Biology and Biochemistry*, *96*, 16–19.
- Zinger, L., Donald, J., Brosse, S., Gonzalez, M. A., Iribar, A., Leroy, C., Muriene, J., Orivel, J., Schimann, H., Taberlet, P., et al. (2020). Advances and prospects of environmental DNA in neotropical rainforests. *Advances in Ecological Research*, *62*, 331–373.
- Zinger, L., Gobet, A., & Pommier, T. (2012). Two decades of describing the unseen majority of aquatic microbial diversity. *Molecular Ecology*, *21*(8), 1878–1896.
- Zinger, L., Gury, J., Alibeu, O., Rioux, D., Gielly, L., Sage, L., Pompanon, F., & Geremia, R. A. (2008). Ce-sscp and ce-fla, simple and high-throughput alternatives for fungal diversity studies. *Journal of Microbiological Methods*, *72*(1), 42–53.
- Zinger, L., Gury, J., Giraud, F., Krivobok, S., Gielly, L., Taberlet, P., & Geremia, R. A. (2007). Improvements of polymerase chain reaction and capillary electrophoresis single-strand conformation polymorphism methods in microbial ecology: Toward a high-throughput method for microbial diversity studies in soil. *Microbial Ecology*, *54*, 203–216.
- Zinger, L., Lejon, D. P., Baptist, F., Bouasria, A., Aubert, S., Geremia, R. A., & Choler, P. (2011). Contrasting diversity patterns of crenarchaeal, bacterial and fungal soil communities in an alpine landscape. *PLoS One*, *6*(5), e19950.
- Zinger, L., Lionnet, C., Benoiston, A.-S., Donald, J., Mercier, C., & Boyer, F. (2021). Metabar: An R package for the evaluation and improvement of DNA metabarcoding data quality. *Methods in Ecology and Evolution*, *12*(4), 586–592.
- Zinger, L., & Philippe, H. (2016). Coalescing molecular evolution and DNA barcoding. *Molecular Ecology*, *25*(9), 1908–1910.
- Zinger, L., Shahnavaz, B., Baptist, F., Geremia, R. A., & Choler, P. (2009). Microbial diversity in alpine tundra soils correlates with snow cover dynamics. *The ISME journal*, *3*(7), 850–859.
- Zinger, L., Taberlet, P., Schimann, H., Bonin, A., Boyer, F., De Barba, M., Gaucher, P., Gielly, L., Giguet-Covex, C., Iribar, A., et al. (2019). Body size determines soil community assembly in a tropical forest. *Molecular Ecology*, *28*(3), 528–543.