



HAL
open science

Dynamique des métacommunautés dans les têtes de bassin versant fragmentées : une perspective moléculaire

Maïlys Gauthier

► To cite this version:

Maïlys Gauthier. Dynamique des métacommunautés dans les têtes de bassin versant fragmentées : une perspective moléculaire. Sciences de la Terre. Université de Lyon, 2020. Français. NNT : 2020LYSE1171 . tel-03352894

HAL Id: tel-03352894

<https://theses.hal.science/tel-03352894v1>

Submitted on 23 Sep 2021

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.

N°d'ordre NNT :



THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 341
Evolution, Ecosystème, Microbiologie, Modélisation
(E2M2)

Spécialité de doctorat : Écologie
Discipline : Écologie des communautés

Soutenue publiquement le 10/12/2020, par :
Mailys Gauthier

**Dynamique des métacommunautés dans les têtes de
bassin versant fragmentées : une perspective
moléculaire**

Thèse dirigée par : Thibault Datry (Directeur de Recherche, HDR), INRAE, Lyon
Christophe Douady (Professeur des Universités, HDR), UCBL, Lyon

Devant le jury composé de :

Gibert, Patricia (Directrice de Recherche, HDR), CNRS, Lyon	Présidente
Bouchez, Agnès (Directrice de Recherche, HDR), INRAE, Thonon-les-Bains	Rapporteuse
Blanchet, Simon (Directeur de Recherche, HDR), CNRS, Saint-Girons	Rapporteur
Plenet, Sandrine (Maître de Conférences, HDR), UCBL, Lyon	Examinatrice
Poulet, Nicolas (Chargé de mission), OFB, Toulouse	Examineur

Datry, Thibault (Directeur de Recherche, HDR), INRAE, Lyon

Directeur de thèse

Université Claude Bernard – LYON 1

Administrateur provisoire de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Philippe CHEVALLIER
Vice-Président de la Commission de Recherche	M. Jean-François MORNEX
Directeur Général des Services	M. Pierre ROLLAND

COMPOSANTES SANTE

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud - Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directrice : Mme Christine VINCIGUERRA

COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

Département Génie Electrique et des Procédés (GEP)	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique (CPE Lyon)	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances (ISFA)	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Administrateur Provisoire : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE

Remerciements

Aucune thèse ne se ferait sans l'aide et le soutien des personnes autour de soi. Tout le monde le sait. Pour ma thèse, ce ne pourrait être plus vrai tant j'ai eu besoin et bénéficié de l'aide de nombreuses personnes pour les manips, les réflexions, le soutien mais aussi pour être tout simplement heureuse et optimiste au quotidien. J'ai eu la chance d'être entourée de personnes tout simplement incroyables, motivées (parfois plus que moi), qui m'ont soulevée dans les moments difficiles, fait rire, secouée, tout ça avec énormément de bienveillance, d'humour et de gentillesse. Bref, j'ai eu une chance extraordinaire de vous avoir tous et toutes rencontrés, j'aimerais vous remercier en racontant chaque anecdote que l'on a vécu ensemble et qui m'ont permis d'être heureuse ou de tenir dans les moments les plus durs. Mais je ne peux pas écrire 50 pages de remerciements. Ce que je vais écrire ici ne représente donc qu'un millième des remerciements que j'aimerais vous faire à toutes et tous alors multipliez-le par mille pour que ça représente vraiment à quel point je vous remercie. J'espère que vous le savez. Parce que je ne vous remercierai jamais assez.

Je remercie Patricia Gibert, Agnès Bouchez, Simon Blanchet, Sandrine Plenet et Nicolas Poulet d'avoir accepté de faire partie de mon jury de soutenance et d'avoir évalué mon travail, avec un remerciement particulier à Agnès et Simon qui ont accepté d'être mes rapporteurs.

Je remercie les membres de mon comité de suivi de thèse Nuria Bonada, Christophe Bouget, Ludovic Orlando et Sylvain Doledec, mon tuteur de thèse, qui m'ont aidé à lancer le projet dès les premiers mois et ont permis par leurs nombreux conseils, pertinents et bienveillants, et nos échanges enrichissants de faire de cette thèse ce qu'elle est aujourd'hui.

Je remercie bien sûr mes deux encadrants de thèse, Thibault et Christophe. Merci de m'avoir fait confiance pour mener à bien ce projet, merci d'avoir pris le temps de discuter de mes blocages existentiels qui m'empêchaient d'avancer, d'avoir été toujours disponibles quand j'en avais besoin et de m'avoir donné les mêmes arguments encore et encore parce que je n'étais plus convaincue de rien. Merci d'avoir pris le temps de me comprendre et de vous adapter à ma sensibilité et à mes humeurs en yo-yo qui sont difficiles à appréhender. Thibault, je te remercie pour ta disponibilité et ta réactivité qui je pense sont indétrônables sur la Terre entière. J'ai été impressionnée par la façon dont tu cogites à vitesse grand V et avec laquelle tu arrivais à envisager toutes les possibilités quand je me trouvais dans une impasse. Ton soutien et tes encouragements me démontrant par $a+b$ que ma thèse méritait d'exister m'ont vraiment permis de me motiver jusqu'au bout ! J'ai énormément appris en discutant avec toi, que ce soit sur les métacommunautés, les rouages de la recherche ou la politique dans les balkans en passant par l'apiculture et la construction d'une maison autonome. Travailler avec toi a été enrichissant sur énormément d'aspects, j'espère que ça a été le cas pour toi aussi. Christophe, merci d'avoir été présent malgré tes responsabilités. Tu as su être là dans les phases importantes et tu as porté un regard sur ma thèse toujours pertinent avec une concision, une exactitude et une précision que j'admire toute ma vie (étant l'exact opposé ahah). Merci aussi pour ta bienveillance et ton soutien, que ce soit par rapport à mon bien-être en tant que doctorante mais aussi en tant que personne.

Tristan, je pourrais écrire une liste avec des centaines de raisons pour lesquelles je dois te remercier en oubliant la moitié. Je ne pense pas avoir embêté quelqu'un autant que toi pendant toute cette thèse et je te

remercie de ta patience envers moi. Tu as agi comme un encadrant avec moi alors que ce n'était pas ton rôle et je ne pourrais jamais assez te remercier pour ça. Je n'aurais jamais pu avancer (commencer ?) sans ton aide en bio-informatique et ton suivi sur les manips de biologie moléculaire pour lesquelles je te sollicitais toutes les trente minutes dans les périodes de crise (souvent) ! Réfléchir avec toi sur mes manips en remplissant de schémas ton tableau blanc adoré pendant des heures a vraiment été une des parties les plus enthousiasmantes de ma thèse (sauf au bout de plusieurs mois d'échec ahah). Merci aussi pour ton soutien et ton oreille attentive quand j'étais au bord de la crise (de nerf, de panique, de rage), tu m'as souvent aidé à relativiser et à m'apaiser. Merci aussi pour les sauvetages chocolat dans les phases d'hypoglycémie.

Je remercie les équipes de mes deux laboratoires où j'ai pu échanger avec de nombreuses personnes. Merci déjà à toute l'équipe DYNAM de INRAE qui est vraiment une équipe soudée, conviviale et où il fait bon vivre. Je n'aurais pu rêver d'une meilleure équipe pour ces trois ans et demi. Je ne peux pas tous et toutes vous citer un·e par un·e, que ce soit les personnes de passage en post-doc ou stage pour quelques mois ou les membres permanents de l'équipe, mais cela compte pour tout le monde : les échanges du quotidien avec chacun·e, votre disponibilité quand j'avais besoin d'aide même pour un rien, la bonne humeur générale, les pauses café du matin et de l'après-midi, les apéros à 11h30 pétantes, l'humour souvent au rendez-vous m'ont permis de travailler dans des conditions idéales et m'ont énormément apporté personnellement et professionnellement. Merci aussi aux autres membres de ma nouvelle équipe EcoFlowS, ex-LHQ, que je connais moins mais avec qui j'ai bien rigolé et discuté en passant à l'improviste au coin café ! Je remercie particulièrement mes stagiaires, Mélanie, Claire, Quentin, Benjamin et Colin qui nous ont grandement aidé sur le terrain et sur le tri dans la bonne humeur et avec motivation. Maxence, tu as été un énorme soutien moral pour moi, j'aurai aimé que tu puisses plus faire partie de l'aventure mais j'ai été ravie de parler de biomol et d'invertébrés avec toi tout au long de ma thèse. Toujours disponible pour parler de tout et de rien, pour m'aider en stats (entre autres) alors que tu n'as pas assez de temps pour toi, tu es la gentillesse, l'écoute et la bienveillance réunies en une seule personne, tu es vraiment au top, merci pour tout. Gui, Bertrand, vous, vous avez fait partie de l'aventure pour le coup ! Sans vous, cette thèse n'existerait tout simplement pas. Un immense merci d'avoir mis les mains dans les ruisseaux trop froids pour moi, merci d'avoir prospecté, échantillonné, trié des centaines de milliers d'invertébrés pour cette thèse, d'avoir enduré le terrain à -5°C et à $+30^{\circ}\text{C}$ et parcouru des milliers de kilomètres (plus d'un aller-retour pôle Nord), d'avoir été super disponibles pour moi, j'ai eu de la chance de vous avoir à mes côtés. On a beaucoup râlé mais j'ai vraiment aimé travailler avec vous, passer mes journées et soirées sur le terrain avec vous, organiser nos pique-niques et rire avec vous. Le rituel du jeu des 1000 euros, les collections de photos idiotes, les centaines de blagues incomprises, les gentilles moqueries sur mon ectothermie, Jasper la mascotte, vélociraptor et le reste du matériel, ma capacité à vous orienter incroyablement bien sont un micro-échantillon des bons moments passés avec vous sur le terrain. La perte d'Olympus, appareil photo disparu sur l'avant-dernier site d'échantillonnage, restera gravé à jamais dans ma mémoire. Bertrand, sur la liste des personnes que j'ai bien embêtées durant ma thèse, tu es dans le top 5 ! Tu as aussi été un soutien incroyable et merci d'avoir été là au quotidien à écouter tous mes doutes, mes réflexions et de m'avoir aidé à démêler tout ça. Tu étais un peu mon deuxième cerveau et m'a aidé à réfléchir dès que j'avais un doute que ce soit pour le terrain, mes plans de manips ou ma thèse en général. Merci pour toutes ces discussions, ton aide précieuse dans le choix et le placement des sites, pour avoir cherché et contacté les personnes pour les infos sur les bassins, pour m'avoir nourri quand je mourrais de faim et tant d'autres choses. Hervé, merci pour toutes les mesures que tu as faites sur SIG souvent en urgence et pour nos

discussions sur l'écologie. Je ne débarquerai plus dans ton bureau à l'improviste en te suppliant de m'aider. Maria, Sophie, Nuria, Dani, merci pour votre aide quand j'étais bloquée dans mes questionnements divers et variés, dans mes analyses ou pour comprendre mes résultats. Merci Maria pour la relecture de l'introduction. Merci aussi à Edwige, Adeline, Marion et Vanessa qui étaient toujours disponibles pour répondre à mes questions administratives.

Merci à l'équipe e3s du LEHNA où la bonne humeur et les rires sont aussi au rendez-vous ! J'ai moins pu profiter du temps passé avec vous étant donné que vous êtes plus souvent à droite à gauche et que j'étais surtout au labo quand j'étais au LEHNA. Mais j'ai vraiment aimé le temps que j'ai passé avec vous à travers les échanges quotidiens ou les discussions sur mes manips qui étaient très enrichissantes. Merci d'ailleurs à Laurent pour les échanges sur tout le matériel du labo, merci à Michel et Michel pour nos discussions sur les Oligochètes, merci à Clémentine pour ton soutien et ton humour incroyable (et les apports nutritionnels en gâteaux et chocolat). Un immense merci infini pour celle sans qui je me serai noyée dans le désespoir pendant mes manips de biologie moléculaire : Lara. On a formé une équipe de choc toutes les deux et je ne le répéterai jamais assez : tu es la meilleure binôme de manip qui existe au monde ! Merci pour toute ton aide au début quand je ne connaissais pas encore tout, merci d'avoir appris avec moi comment fonctionnait la capture, d'avoir galéré à mes côtés quand ça ne marchait pas, de t'être enthousiasmée quand on pensait que ça allait marcher, d'avoir passé des heures à mes côtés sous la hotte, de m'avoir donné des conseils pour les manips et d'avoir planché sur tous les problèmes que l'on a rencontrés en m'envoyant parfois des messages tard le soir ou le week-end parce que tu avais eu un éclair de génie. Tout ça toujours dans la bonne humeur et l'enthousiasme et les discussions super chouettes. Merci du fond du cœur Lara. Merci aussi à Agnès N'guyen, notre experte en metabarcoding, qui nous a conseillées sur les manips, qui est toujours joviale et hyper disponible pour nous aider. J'espère qu'on continuera à se faire des restos toutes les trois même après ma thèse. A tous et toutes de mes deux laboratoires, je n'arrive pas à imaginer que je ne verrai plus vos têtes tous les jours (même si le confinement m'a un peu aidé à m'y préparer) et j'espère vraiment que l'on se recroisera souvent.

Grâce aux congrès et aux COST Action (SMIRES et DNAqua-Net), j'ai pu échanger avec de nombreuses personnes extrêmement douées qui m'ont aidé à aborder différemment ma thèse et à m'enrichir professionnellement, notamment au workshop au Monténégro. Je tiens à remercier particulièrement Vasco et Florian avec qui j'ai eu des discussions très intéressantes sur les protocoles de biologie moléculaire et qui m'ont énormément aidée sur cet aspect, mais aussi Agnès, Rachel, Romain et tant d'autres que j'oublie sûrement.

Je remercie l'ensemble des doctorant·es et post-doctorant·es de mes deux laboratoires qui pour la grande majorité sont devenu·es mes ami·es. Alexis et Sandrine, vous n'êtes pas doctorants mais je vous mets dans cette case-là. Une thèse, c'est des hauts et des bas et avoir des personnes autour de nous qui comprennent tout cela, qui l'ont vécu ou qui le vivent et qui peuvent vraiment nous soutenir dans les épreuves difficiles mais aussi dans les moments de joie, c'est extrêmement précieux. Merci à vous pour les pauses cafés et les discussions sérieuses ou complètement loufoques du côté de INRAE ; merci pour le rituel café-mots fléchés et les pauses sur la passerelle côté LEHNA. Que je vous ai rencontrés en début de thèse ou il y a peu, vous avez tous et toutes participé à apporter de la bienveillance, du soutien, de la joie de vivre, du fun et de la chouettitude au labo. Je ne peux pas vous citer un·e par un·e car vous êtes trop nombreux·ses. Mais je tiens à remercier particulièrement notre maman Laura qui était notre encyclopédie de thèse côté stats et administration, Julie qui travaillait sur la

même thématique et avec le même encadrant que moi et avec qui j'ai énormément échangé, réfléchi et râlé sur les métacommunautés. Merci à Maxime, mon co-bureau qui a dû supporter mes blablatages incessants et mes petites crises de folie. Traverser cette thèse avec toi (et nos plantes d'amour) a été un plaisir et un honneur et notre soutien mutuel et nos délires (Rosé le roseau, combat de boxe pour décompresser, communication par morse avec les bureaux voisins, j'en passe et des meilleures) tout au long de nos thèses m'ont aidée à tenir le coup bien des fois. Merci à Nathanaëlle, ma co-bureau de l'autre côté qui est une des personnes les plus bienveillantes que je connaisse et qui fait tout pour que tout le monde aille bien. Je n'étais pas souvent là mais j'ai adoré discuter et partager mon bureau avec toi.

Et il y a aussi toutes les personnes qui n'ont pas de lien direct avec ma thèse mais qui représentent les soutiens les plus importants de ma vie et qui m'ont plus d'une fois ramassée à la petite cuillère, secouée, relevée, faite relativiser et ont pris grand soin de moi. Un grand merci à tous·tes les copain·es qui sont présent·es depuis longtemps ou que j'ai rencontrés pendant ma thèse. Je vais sûrement oublier quelques personnes, j'en suis désolée mais vous savez à quel point je suis tête en l'air et que j'oublie des choses sur le moment. Les copines du Beaujolais, Marine, Manon, Marine, Brenda, Steph, Marlène, Cindy, vous êtes un rayon de soleil et de joie quand je vous vois, merci pour la simplicité de nos retrouvailles et les délires si faciles à avoir avec vous. Sara, Camille, Nathan, Aurélie, Manu, merci de faire encore partie de ma vie et d'être toujours vous-mêmes (c'est-à-dire absolument géniaux et géniales), d'avoir suivi mon parcours depuis le lycée et de me soutenir depuis que l'on se connaît. Je vous admire pour ce que vous avez fait depuis que l'on se connaît et ce que vous êtes devenu·es, vous êtes une source d'inspiration au quotidien pour moi. Merci Quentin et Marine d'arriver à consacrer un peu de votre précieux temps pour caler une rando avec météo foireuse assurée, d'être toujours heureux·se et de bonne humeur. Merci aux écologues de l'extrême loose, en particulier Camille, Chéimoute, Alizou, Vinc' et Valou, qui ont partagé le master avec moi et qui sont toujours là même si la vie nous a dispatchés dans le monde. Je remercie aussi toutes les personnes que j'ai rencontrées dans le milieu militant et qui m'ont soutenu et compris dans les moments les plus sombres où je ne croyais plus en rien. Marine, Luc, Pierre, Louise, Mimo, Greg (#lameute), Paillette, Mathilde, Maimai, Kev, Manue, Tobi, Arnaud, Arthur, Quentin, Terter, Cash, Amo, Bipbip et j'en oublie beaucoup, vous avez tous et toutes participé à rendre ma vie plus belle, nos discussions sur la vie et sur le monde m'ont complètement enrichie et je suis ravie de vous avoir rencontré·es. Luc, un remerciement spécial pour les illustrations magnifiques que tu as faites pour mon manuscrit et pour ta présence et ton soutien dans les dernières lignes droites (Pierre, merci aussi pour ça). J'ai commencé à écrire ma thèse en confinement et je remercie les copain·es qui ont supporté les semaines de rédaction et d'analyse intensives avec moi avec insomnies, cernes et état de folie pendant (et après) cette période : Clarisse, Héloïse, Charlotte, Mathieu et Juliette. J'ajoute à ces merveilleux·ses doctorant·es celles et ceux qui ont eu un rôle de soutien et de joie important pendant ma thèse et que je n'ai pas encore cité : Hugo, Valérian, Marine, Alba, Noëlle, Maxime G, Ivan, Mathieu, Paupau, Thomas, Samuel, Boris, Dom et Clara. Valérian, merci pour nos discussions jusqu'à 3 heures du matin sur ton balcon où l'on a refait le monde et discuté de nos questionnements sur la vie, merci pour ta bienveillance, les sessions escalade, jeux et j'en passe. Juliette, Héloïse, Clarisse et Antonin, merci tout simplement de faire partie de ma vie. Vous savez la place que vous avez dans mon cœur et je ne vous remercierai jamais assez de l'amour, du bonheur et du soutien que vous m'apportez au quotidien et de la vision que vous avez de la vie qui me fait toujours relativiser. Vous suivez chaque instant de ma vie à la minute près et m'aidez à prendre des décisions à chaque instant, merci d'être là et de m'avoir soutenu

jusqu'au bout. Juliette, merci de m'avoir montré que penser à soi, ce n'était pas si mal au final, et merci pour le recul incroyable que tu as sur toutes les situations et qui m'aide à voir la vie différemment (et bisous à Nonos). Héloïse, merci pour ta bonne humeur, ta motivation, ton humour, ton enthousiasme et ta prévenance infinie. Clarisse, ma colocataire que je considère comme ma petite sœur, je te remercie d'embellir mon quotidien par ta joie de vivre et tes rires aux éclats, tu es un vrai petit soleil. Antonin, merci pour tout ce que tu m'apportes, pour ton soutien infini, pour les heures passées au téléphone à parler de tout et de rien. Tu es la dose de calme, de sérénité et de positivisme dont j'ai besoin pour mon équilibre. Merci aussi de m'avoir autorisé à squatter votre coloc de rêve de nombreuses fois (petit clin d'œil à Gaëlle, Charles et Daniel). Merci à Juliette, Antonin et Clarisse pour la relecture de l'introduction. Merci Rémy d'avoir partagé un bout de ma vie, tu as été un soutien important de manière générale et au début de ma thèse et je te remercie pour ton aide, ton réconfort et d'avoir été là pour moi.

Et pour finir, je souhaite remercier du fond du cœur ma famille. Les cousin·es, même si on ne se voit pas beaucoup, vous voir pour faire des marathons jeux et mises à jour de vie est toujours un moment super chouette et agréable. Merci à mes tantes et oncles de partager ces moments avec nous, particulièrement à ma marraine Tatie qui suit ma vie de près avec bienveillance. Merci à mes mamies qui prennent des nouvelles de « mes études » et de l'« école », qui essaient toujours de comprendre ce qu'est une thèse et qui oublient que je ne travaille pas sur des grenouilles, vous me faites tellement rire. Merci à ma Moumoute qui est toute fière de moi et m'a secouée (un peu trop parfois ahah) dans les moments où je ne tenais pas le coup. Je sais que tu es capable de comprendre ma thèse alors j'espère que tu la liras. Merci à Alex et toi pour la fourniture d'huiles en tout genre pour nourrir la vegan que je suis. Merci Papoune, même si tu as essayé plus d'une fois de me sortir du « système », tu as accepté mes choix de vie et tu m'as aidé à surmonter les moments difficiles. Merci finalement à mes fratries, éloignées et proches. Elo, Jah, Yonnah et Cathy, vous serez toujours dans mon cœur même si des centaines de kilomètres nous séparent. Lorelei, Jorick et Kiran, c'est quand je suis avec vous que je me sens chez moi et c'est avec vous que j'oublie tous mes soucis pour rire et vivre l'instant présent. Je ne sais pas comment l'exprimer mieux. Même si on s'agace facilement avec nos caractères (un peu) forts, notre famille est soudée et je vous aime. Lolo, ma grande sœur, merci d'être à l'écoute quand j'en ai besoin, merci de prendre soin de moi, d'être enthousiaste et de suivre ma vie depuis si longtemps (et de t'être mis en mode Maïlys pour corriger mes fautes d'orthographe). Jorick et Kiran, je vous mets dans le même panier étant donné que Kiran suit la voie de son frère. Vous êtes les deux personnes qui me font le plus rire au monde, merci pour votre originalité, toutes vos blagues, vos imitations d'accents, pour vos jeux de mots vraiment nuls (heureusement que je suis bon public ahah). Vous avez le don de me changer les idées. A tous les trois, merci pour votre amour et votre bienveillance au quotidien. Merci d'avoir été là quand j'allais mal, de m'avoir fait rire, d'avoir pris soin de moi et de vous être inquiété·es pour moi dans ces moments difficiles. A ma famille proche, merci pour vos encouragements quand j'étais isolée dans un coin de la maison pendant que vous croquiez la vie à pleine dents et que vous passiez me dire « courage », « t'es la meilleure » !

Si vous êtes arrivé·e au bout, je vous remercie d'avoir lu tout ça. Je sais que c'est long mais je n'aurai pas terminé ma thèse sans toutes ces personnes et tout le soutien qu'elles m'ont offert. Remercier ces personnes qui m'ont aidé et qui comptent pour moi me tient particulièrement à cœur. Je pourrais d'ailleurs en écrire encore le double (celles et ceux qui me connaissent un peu ne devraient pas être étonnés de ma capacité à parler sans fin et à m'étaler). Bonne lecture si vous avez encore le courage de continuer !

Table des matières

Table des figures.....	1
Table des tableaux	3
Résumé.....	5
Abstract	7
Avant-propos	9
Chapitre 1 - Introduction	13
1 Pourquoi décrire et classifier le monde Vivant ?.....	15
2 Évolution des concepts en écologie des communautés.....	16
2.1 Niche écologique	17
2.2 Dispersion	18
2.3 Histoire évolutive.....	20
2.4 Stochasticité.....	22
2.5 Rassembler les approches.....	22
3 Écologie des communautés dans les rivières : vers une prise en compte de la dispersion.....	23
4 Spécificités des métacommunautés de têtes de bassin versant, milieux très dynamiques et fragmentés	30
5 Modèle invertébrés dans l'étude des métacommunautés : intérêts et limites.....	35
6 Affiner la compréhension des métacommunautés : une perspective moléculaire	38
7 Limites actuelles de l'identification moléculaire.....	39
8 Gestion et conservation des têtes de bassin versant : entre système ignoré et dynamique	42
9 Objectifs et structure de la thèse.....	45
10 Références	47
Chapitre 2 - Métacommunautés dans les têtes de bassin versant, le rôle de la dispersion et de la fragmentation	63
1 Matériel et méthodes - Chapitre 2	65
1.1 Choix des têtes de bassin versant : prospection et sélection.....	65
1.2 Echantillonnage	66

1.3	Caractérisation des états d'écoulement.....	70
1.4	Références	72
1.5	Annexes.....	73
2	Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1).....	75
2.1	Résumé étendu.....	75
2.2	Abstract.....	79
2.3	Introduction	79
2.4	Methods.....	82
2.5	Results.....	87
2.6	Discussion	93
2.7	Acknowledgements.....	98
2.8	Data accessibility	98
2.9	References	98
2.10	Supplemental Information	107
3	Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)	142
3.1	Résumé étendu.....	142
3.2	Abstract.....	145
3.3	Introduction	146
3.4	Material & methods	148
3.5	Results.....	153
3.6	Discussion	158
3.7	Acknowledgements.....	162
3.8	Data accessibility	162
3.9	References	162
3.10	Supplemental Information	167
Chapitre 3 - Apport de l'identification moléculaire pour l'étude des métacommunautés de têtes de bassin versant		175
1	Matériel et méthodes - Chapitre 3	177

1.1	Travail préliminaire à la conception de l'outil moléculaire : conception des sondes ARN	177
1.2	Protocoles de laboratoire : de l'extraction au séquençage.....	181
1.3	Enrichissement : amplification par PCR	183
1.4	Traitement bio-informatique des données issues du séquençage.....	188
1.5	Références	191
2	Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3).....	192
2.1	Résumé étendu.....	192
2.2	Abstract	195
2.3	Introduction	196
2.4	Materials & methods.....	200
2.5	Results.....	207
2.6	Discussion	213
2.7	Conclusion.....	218
2.8	Acknowledgements.....	219
2.9	Data accessibility	219
2.10	References	219
2.11	Supplemental Information	230
3	Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés - Étude exploratoire sur le bassin versant de l'Albarine (article 4)	244
3.1	Introduction	244
3.2	Matériel et méthodes.....	247
3.3	Résultats	250
3.4	Discussion	255
3.5	Références	259
3.6	Annexes.....	263
Chapitre 4 - Synthèse et Perspectives.....		269
1	Considération des résultats dans le cadre de la gestion et de la conservation des têtes de bassin versant.....	

2	Aller plus loin dans la description des processus structurant les métacommunautés des têtes de bassin versant	278
2.1	Mieux décrire les dynamiques de dispersion.....	278
2.2	Prise en compte de variables supplémentaires pour décrire la niche écologique et étude d'autres habitats	280
2.3	Intégration de la dynamique temporelle dans les analyses	281
2.4	Vers une meilleure description de la fragmentation	284
2.5	Discriminer les groupes d'invertébrés pour mieux appréhender les dynamiques.....	286
3	Amélioration de l'outil moléculaire pour décrire la biodiversité de têtes de bassin versant	287
3.1	Limites associées à l'ADN extrait à partir de l'éthanol	287
3.2	Perspectives concernant la capture par sondes ARN	290
3.3	Réflexion sur l'assignement des données issues de metabarcoding	293
4	Références.....	294

Table des figures

Chapitre 1 - Introduction

Figure 1.1 : Planches IV et V du <i>Systema Naturea</i> de Carl von Linné	15
Figure 1.2 : Extinction des espèces de vertébrés depuis 1500 et évolution de l'Indice Liste rouge depuis 1980	16
Figure 2.1 : Différents niveaux d'organisation pour l'étude de la biodiversité	17
Figure 2.2 : Les 4 paradigmes du concept de métacommunauté	18
Figure 2.3 : Échelles spatiales considérées dans l'étude des métacommunautés	19
Figure 2.4 : Structure des communautés en lien avec la phylogénie	21
Figure 2.5 : Théorie des communautés écologiques selon Vellend, 2010	22
Figure 2.6 : Différents concepts utilisés en écologie des communautés	23
Figure 3.1 : Différentes parties et structure dendritique d'un réseau hydrographique.....	26
Figure 3.2 : Cartes d'un réseau de référence du Service Géographique de l'Armée (gauche) et réseau réel prenant en compte les têtes de bassin versant	29
Figure 4.1 : Structure et dynamique d'une métacommunauté de tête de bassin versant contenant des tronçons intermittents et des structures anthropiques	30
Figure 4.2 : Différentes phases hydrologiques observées sur un tronçon intermittent	31
Figure 4.3 : Dynamique spatio-temporelle des habitats et de la connectivité dans un bassin	32
Figure 4.4 : Seuil en rivière	34
Figure 5.1 : Phylogénie des métazoaires	35
Figure 5.2 : Invertébrés aquatiques présentant des modes de dispersion variés.	36
Figure 7.1 : Schéma d'une PCR	40
Figure 8.1 : Carte de l'état écologique des cours d'eau en France en 2015 (DCE).....	43

Chapitre 2 - Métacommunautés dans les têtes de bassins versants, le rôle de la dispersion et de la fragmentation

Figure 1.1 : Localisation en France des 11 têtes de bassins versants étudiées	66
Figure 1.2 : Cartes détaillées et photos des 11 têtes de bassin versant étudiées	68
Figure 2.1 : Idealized dispersal dynamics in intermittent headwater stream networks and expected alpha- and beta-diversity patterns.	82
Figure 2.2 : The two study areas in France.....	85
Figure 2.3 : Alpha-diversity distribution for perennial and intermittent reaches for each FHW	89
Figure 2.4 : Coefficient of correlation of Mantel test for each community type and distance matrices	90

Figure 2.5 : Venn diagrams displaying unique and shared taxa proportion between intermittent and perennial reaches for each FHW.....	92
Figure 3.1 : Theoretical contribution of dispersal-based and niche-based processes in fragmented river systems across hydrological phases and expected beta-diversity patterns.....	147
Figure 3.2 : Location and map of the Thouaret at different hydrological phases.....	149
Figure 3.3 : Discharge time series for the year 2018 of the Thouaret.....	150
Figure 3.4 : Strength of the relationship between community dissimilarities and typical distances through the different sampling dates.....	155
Figure 3.5 : Relationships between community dissimilarities and the two fragmentation matrices.....	157

Chapitre 3 - Apport de l'identification moléculaire pour l'étude des métacommunautés de têtes de bassin versant

Figure 1.1 : Étapes effectuées pour établir la base de données COI.....	179
Figure 1.2 : Étapes suivies pour la conception des sondes.....	181
Figure 1.3 : Étapes pour l'obtention des librairies avec l'enrichissement par PCR.....	183
Figure 1.4 : Étapes pour l'obtention des librairies avec l'enrichissement par capture.....	185
Figure 2.1 : Design expérimental simplifié avec les deux types de communautés utilisés.....	193
Figure 2.2 : Overview of the experimental design with the two different mock communities.....	202
Figure 2.3 : Taxa recovery performance with bulk DNA for PCR and capture enrichment.....	210
Figure 2.4 : Biomass proportion and read proportion relationship for PCR and capture.....	211
Figure 2.5 : PCA based on dissimilarity in read proportions of each pair of the 10 species mock community samples following their DNA template and enrichment methods.....	213
Figure 3.1 : Pourcentage de taxa obtenus par identification moléculaire morphologique.....	251
Figure 3.2 : Carte de détection des taxa par les identifications moléculaire et morphologique à chaque site de prélèvement pour la sous-classe des Oligochètes.....	254
Figure 3.3 : Coefficient de corrélation des tests de Mantel pour chaque combinaison de groupe d'invertébrés sélectionné et de matrices de distance pour les deux méthodes d'identifications.....	255

Chapitre 4 - Synthèse et Perspectives

Figure 1.1 : Exemple du changement du nombre de sites d'évaluation d'une gestion centrée sur l'échelle locale cà une gestion intégrant l'échelle régionale.....	280
Figure 1.2 : Étapes pour intégrer le concept de métacommunauté dans l'évaluation du bon état des communautés dans les milieux très dynamiques par Cid et al (2020).....	277
Figure 2.1 : Plusieurs types de distances utilisées dans les études de métacommunautés.....	279
Figure 3.1 : Détection par l'identification moléculaire des familles identifiées en identification morphologique des données de l'article 4.....	288

Table des tableaux

Chapitre 1 - Introduction

Tableau 4.1 : Perturbations anthropiques des têtes de bassin versant et impact associé.	33
--	----

Chapitre 2 - Métacommunautés dans les têtes de bassins versants, le rôle de la dispersion et de la fragmentation

Tableau 1.1. Niveau d'identification des ordres d'invertébrés benthiques échantillonnés	70
Tableau 2.1. Main characteristics of the 10 studied headwater stream networks	83
Tableau 2.2. Correlation coefficient of the Mantel tests carried out to disentangle metacommunity processes in the 11 headwaters	91
Tableau 3.1. Environmental variables measured for each sampling date	153
Tableau 3.2. Descriptors of permanent and temporary fragmentation in the Thouaret network. ...	154
Tableau 3.3. r values from the Mantel tests based on Chao dissimilarity index and typical distances used in metacommunity studies in the Thouaret network.	156
Tableau 3.4. r values from the Mantel tests based on Chao dissimilarity index and fragmentation distances used in this study.	157

Chapitre 3 - Apport de l'identification moléculaire pour l'étude des métacommunautés de têtes de bassin versant

Tableau 1.1. Ordres utilisés pour la conception des sondes ARN	179
Tableau 1.2. Oligonucléotides utilisés pour le barcoding, l'amplification par PCR et la capture par sondes.....	187
Tableau 1.3. Critères de conservation des alignements lors de l'étape d'assignement	190
Tableau 2.1. Dry biomass in mg and number of individuals of freshwater invertebrate species in the 10 species mock communities.	201
Tableau 2.2. Assessment of PCR and capture enrichment specificity on bulk and ethanol DNA..	208
Tableau 2.3. Testing the link between read proportion and initial biomass proportion	211
Tableau 3.1. Taxa les plus abondants et fréquents dans les échantillons utilisés pour l'identification morphologique pour les sites pérennes, les sites intermittents et l'ensemble des sites.....	252
Tableau 3.2. Détails sur les groupes d'invertébrés étudiés.....	253

Résumé

Le concept de métacommunautés a permis d'intégrer explicitement le rôle de la dispersion dans la structure des communautés locales à de multiples échelles spatiales et temporelles. Dans les réseaux hydrographiques, des systèmes très structurés spatialement où les conditions hydrologiques et environnementales varient considérablement dans le temps et l'espace, la prise en compte de la dispersion a permis d'améliorer substantiellement nos connaissances sur l'organisation de la biodiversité aquatique, tout en fournissant des outils pour aider à sa conservation. Cependant, nos connaissances concernant la dynamique des métacommunautés dans les têtes de bassin versant, les parties les plus amont des réseaux hydrographiques, restent parcellaires. Pourtant, ces têtes de bassin versant représentent une majorité des cours d'eau et jouent un rôle clé dans le fonctionnement des réseaux hydrographiques de par leur biodiversité. Du fait de leur position amont qui les isolent au sein des réseaux hydrographiques et de la fragmentation forte que ces milieux subissent, naturelle par les assèchements et artificielle sous l'action de l'être humain, une approche métacommunauté semble la plus pertinente pour comprendre l'organisation spatio-temporelle de leur biodiversité.

L'objectif de cette thèse est d'améliorer la compréhension de la dynamique des métacommunautés de têtes de bassin versant (i) en déterminant les processus à l'œuvre dans dix têtes de bassins versants subissant des degrés différents de fragmentation, (ii) en caractérisant la dynamique temporelle de ces processus dans un bassin versant fortement fragmenté par des structures anthropiques et des assèchements et (iii) en développant une nouvelle méthode d'identification moléculaire pour affiner la différenciation des communautés d'invertébrés les unes des autres et ainsi mieux caractériser la dynamique des métacommunautés aquatiques.

Dans un premier temps, nos résultats ont montré que la diversité alpha était plus faible dans les tronçons intermittents que dans les tronçons pérennes et que les communautés de tronçons intermittents représentaient des sous-ensembles de celles des tronçons pérennes. Ceci confirme ce qui a été montré dans les parties plus à l'aval des réseaux hydrographiques. A l'échelle du réseau, les métacommunautés étaient structurées principalement par les distances spatiales, notamment celles associées à la fragmentation. L'ensemble de ces résultats suggèrent que le processus de dispersion est majoritairement à l'œuvre dans les têtes de bassin versant dynamiques et fragmentées, ceci à travers des événements d'extinction et de recolonisation des tronçons intermittents par des zones refuges pérennes et par la limitation de la dispersion due à la fragmentation et au dendritisme des réseaux. Ceci remet en question le point de vue actuel selon lequel la sélection de niche est le processus principal à l'œuvre dans les têtes de bassin versant du fait de leur isolement et des forts contrastes environnementaux qui y prévalent. Dans un deuxième temps, nous avons montré que les processus structurant les métacommunautés d'une tête de bassin versant dynamique et fragmentée variaient

considérablement dans le temps, faisant écho à de récentes études mettant l'accent sur l'importance d'intégrer la composante temporelle dans l'étude des métacommunautés des écosystèmes dynamiques. Ce résultat, couplé au fait que la dispersion est le processus dominant, reflète sans doute la réponse des communautés aux cycles hydrologiques caractérisant ces écosystèmes, impliquant des événements fréquents d'extinction et de recolonisation à partir de zones refuges. Dans ce contexte de forte fragmentation par les structures anthropiques, les réponses à l'assèchement étaient indécélables, suggérant que le rôle de la fragmentation dans la structure des métacommunautés dépend de la nature et du contexte de fragmentation. Enfin, nous avons développé une méthode d'identification moléculaire efficace et fiable pour obtenir des informations qualitatives, voire quantitatives, pour les études de métacommunautés par rapport aux méthodes d'identification moléculaire classiquement utilisées. L'application de la méthode développée sur des échantillons prélevés *in situ* a permis de montrer son efficacité pour mettre en évidence des patrons spatiaux de distribution chez des groupes d'invertébrés pour lesquels les analyses en identification morphologique n'étaient pas possibles dû à un niveau taxinomique grossier. Ce travail de thèse laisse entrevoir de nombreux axes d'approfondissements et d'amélioration que ce soit dans la compréhension de la structure des métacommunautés avec une meilleure description et quantification de la dispersion, de la fragmentation, le développement de nouveaux outils statistiques ou dans l'amélioration de la méthode moléculaire. Malgré tout, ce travail de thèse peut déjà servir à guider les programmes de gestion et de conservation des têtes de bassin versant en mettant en avant certaines priorités liées à la dispersion, à la variabilité temporelle et au contexte de fragmentation rencontrés dans ces écosystèmes. De plus, cette thèse propose de nombreuses pistes d'optimisation pour que la méthode d'identification moléculaire puisse devenir un outil d'identification de routine dans l'évaluation de l'état écologique des rivières.

Abstract

The metacommunity concept allowed to explicitly incorporate the role of dispersal in the structure of local communities at multiple spatial and temporal scales in ecological studies. In river systems, which are highly structured spatially and where hydrological and environmental conditions vary considerably in time and space, taking dispersal into account improved substantially our knowledge of aquatic biodiversity organisation while providing applied tools to improve river conservation. However, our knowledge of the metacommunity dynamics in the headwaters, the upstream parts of river networks, remains sparse. Yet, headwaters constitute the majority of river networks and play a key role in the functioning of river ecosystems thanks to their high biodiversity. Due to their upstream position, which isolates their communities within river networks, and the high fragmentation they undergo, by both natural drying and man-made structures, the metacommunity framework seems the most relevant to understand the spatio-temporal structure of their biodiversity.

The objective of this thesis is to improve our understanding of metacommunity dynamics in headwaters (i) by revealing the processes at work in ten headwaters undergoing different degrees of fragmentation, (ii) by describing the temporal dynamics of these processes in a headwater highly fragmented by man-made structures and drying, and (iii) by developing a novel molecular identification method to refine invertebrate community differentiation within the network and better characterize the dynamics of headwater aquatic metacommunities.

First, our results showed that alpha diversity was lower in intermittent reaches than perennial ones and that intermittent communities constituted subsets of those of perennial reaches. This confirms what was shown in more downstream parts of the river systems prone to drying. At the network scale, metacommunities were structured primarily by spatial distances, particularly those associated with fragmentation. All these results suggest that dispersal process is a major process in dynamic and fragmented headwaters, through extinction and recolonization events of intermittent reaches by perennial refuge zones and by dispersal limitation induced by fragmentation and dendritism of the networks. This reconsiders the current view that niche selection is the main process at work in headwaters due to their isolation and the strong environmental filtering prevailing there. Second, we showed that the processes structuring metacommunity in dynamic and fragmented headwater varied considerably through time, echoing recent studies which emphasizes the importance of integrating the temporal component in metacommunity studies of dynamic ecosystems. This result, coupled with the fact that dispersal is the dominant process, reflects the response of communities to the hydrological cycles characterizing these ecosystems, which involves frequent events of extinction and recolonization from refuges. In the headwater context of high fragmentation by man-made structures, responses to fragmentation by drying were undetectable, suggesting that the role of fragmentation in

metacommunity structure depends on the nature and context of the fragmentation. Finally, we developed an efficient and reliable molecular identification method to obtain qualitative, and even quantitative, information for metacommunity studies in comparison with molecular identification methods traditionally used. The use of the developed method on *in situ* collected samples demonstrated its effectiveness to reveal spatial distribution patterns in invertebrate groups for which morphological identification analyses were not possible due to too coarse taxonomic level. This thesis offers many areas for further research and improvement in the assessment of metacommunity structure with a need to better describe and quantify dispersal and fragmentation and to develop new statistical tools or in the improvement of molecular methods. Nevertheless, our findings can already be used to guide headwater management and conservation programs by highlighting some priorities, particularly related to the dispersal, temporal variability and fragmentation context encountered in these ecosystems. In addition, this thesis proposes many avenues for optimisations of the molecular identification method, so that it can become a routine identification tool in the assessment of river ecological status.

Avant-propos

Ma thèse a été financée par l'Office Français de la Biodiversité dans le cadre de l'Action n°3 « Fonctionnement des écosystèmes ».

Publications

Publiés

Gauthier, M., Konecny-Dupre, L., Nguyen, A., Elbrecht, V., Datry, T., Douady, C., & Lefébure, T. (2020). Enhancing DNA metabarcoding performance and applicability with bait capture enrichment and DNA from conservative ethanol. *Molecular Ecology Resources*, 20(1), 79-96.

Gauthier, M., Launay, B., Le Goff, G., Pella, H., Douady, C. J., & Datry, T. (2020). Fragmentation promotes the role of dispersal in determining ten intermittent headwater stream metacommunities. In Press. *Freshwater Biology*.

Soumis

Gauthier, M., Launay, B., Le Goff, G., Douady, C. J., & Datry, T. (in review). Drying does not alter metacommunity dynamics in a highly fragmented river network.

A finaliser

Gauthier, M., Datry, T., Douady, C., & Lefébure, T. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés - Étude exploratoire sur le bassin versant de l'Albarine.

Congrès et workshops

Gauthier, M., Lefébure, T., Forcellini, M., Douady, C., & Datry, T. (2017). Metacommunity organization in spatially constrained, dynamic systems: a molecular perspective. *Congrès Symposium for European Freshwater Sciences (SEFS) 10* – Olomouc, République Tchèque. Présentation orale.

Gauthier, M., Konecny-Dupre, L., Nguyen, A., Elbrecht, V., Datry, T., Douady, C., & Lefébure, T. (2018). Development of bait capture for metabarcoding purpose : a robust alternative to PCR enrichment? *Congrès Limnologia* – Coimbra, Portugal. Présentation orale.

Gauthier, M. (2018). Metacommunity structure in headwaters streams : a molecular approach. *Workshop Cost Action Joint Meeting (SMIRES & DNAqua-Net)* – Budva, Monténégro. Présentation orale.

Gauthier, M., Launay, B., Le Goff, G., Konecny-Dupre, L., Lefébure, T., Douady, C. J., & Datry, T. Metacommunity structure of benthic macroinvertebrates IRES with DNA metabarcoding and morphological identification (2019). *Congrès SEFS 11* – Zagreb, Croatie. Présentation orale. (Non présente mais résultats présentés par T. Datry).

Chapitre 1 - Introduction



Illustration : Carte de Cassini (XVIIIème siècle) montrant la tête de bassin versant du Vers (Lot).
Source : geoportail.

« On ne voit d'abord qu'une sorte de labyrinthe où dépressions et hauteurs alternent sans ordre : mais si l'on planait comme l'oiseau, ou si l'on se balançait dans la nacelle d'un ballon, on verrait que les limites du bassin s'arrondissent autour de toutes les sources du ruisseau comme un amphithéâtre et que tous les vallons ouverts dans la vaste rondeur s'inclinent en convergeant l'un vers l'autre et se réunissant en une vallée commune. La chaîne principale des hauteurs forme le bord le plus élevé du cirque ; deux autres côtés sont des chaînons latéraux qui s'abaissent graduellement en s'éloignant de la grande arête, et quelques collines basses se rapprochent pour fermer le cirque parallèlement aux montagnes ; mais elles laissent une issue, celle par laquelle échappe le ruisseau »

Jean Jacques Elisée Reclus dans « Histoire d'un ruisseau »

1 Pourquoi décrire et classer le monde Vivant ?

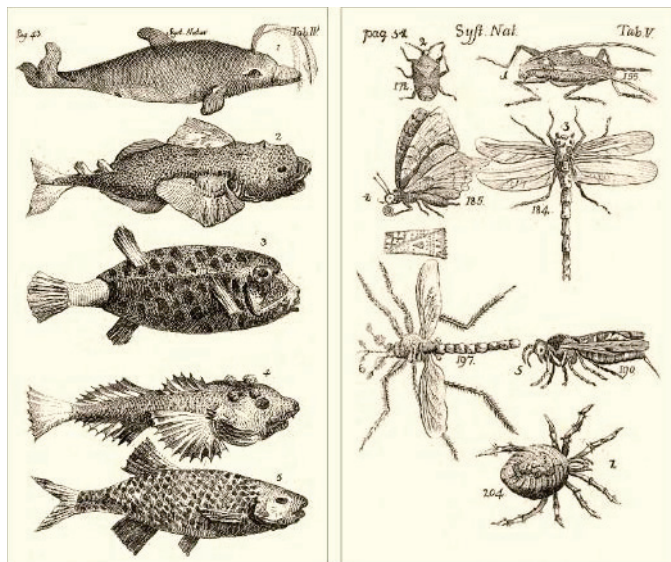


Figure 1.1 : Planches IV et V du *Systema Naturae* de Carl von Linné, premier essai de classification systématique

Depuis des millénaires, l'être humain décrit et classe le monde qui l'entoure. Aristote (384-322 av. J.C) classait les « êtres vivants » par l'observation de critères d'appartenance aux mondes minéral, végétal, animal et humain selon une complexification graduelle de la pierre à l'humain (Aristote, IVème siècle av. JC). Quelques siècles plus tard, Pline l'Ancien (23-79 ap. J.C) rédige en plusieurs volumes *Naturalis Historia* qui résume toutes les connaissances de l'Antiquité latine allant de la structure de l'Univers aux minéraux en passant par la médecine et l'anthropologie (Ier siècle). La classification moderne de la diversité du Vivant, ou biodiversité, est basée sur les travaux de Carl von Linné (1707-1778), qui, dans son œuvre *Systema*

Naturea, répertorie, décrit, nomme et classe l'essentiel des espèces vivantes connues à son époque selon un système binominal genre-espèce (**Erreur ! Référence non valide pour un signet.**, von Linné, 1735). L'espèce en est une notion clé et est restée comme l'unité élémentaire de classification, comme on peut le voir dans les programmes de conservation ([IUCN Red List](#)). Pourtant, la définition de la notion d'espèce est extrêmement floue, encore questionnée et subjective aujourd'hui. Ceci vient du fait qu'une espèce a longtemps été considérée du point de vue essentialiste, c'est-à-dire comme une entité naturellement délimitée et à catégoriser, sans prendre en compte qu'elle peut contenir de la variabilité et qu'elle est en constante évolution (Le Guyader, 2002), comme en témoignent certains phénomènes comme l'interfécondité entre espèces (Arntzen, 1986) ou la variation clinale dans des complexes d'espèces (*ring species*, Devitt et al., 2011). Des auteurs ont proposé d'autres nomenclatures prenant en compte le fait que les espèces évoluent en permanence (*PhyloCode*, LITU, Pleijel & Rouse, 2000) qui ont été adoptés en sciences de la biologie malgré que le terme « espèce » ait perduré pour définir les unités taxinomiques utilisées (Le Guyader, 2002). Au fil du temps et avec toujours comme unité de base l'espèce, la description du Vivant se complexifie avec l'émergence de nouvelles disciplines dont l'écologie scientifique qui cherche à étudier et décrire les règles d'assemblages des espèces entre elles et avec leur environnement à différents niveaux d'organisation, allant de l'espèce à l'écosystème, et à différentes échelles spatiales et temporelles. A l'instar des difficultés rencontrées avec le concept d'espèce, définir les règles d'assemblage dans des systèmes

complexes contenant des entités complexes et présentant de multiples facettes en constante évolution représente un challenge considérable (Lawton, 1999).

Alors, pourquoi l'être humain s'efforce-t-il à décrire et comprendre ces règles paraissant si complexes à définir ? Et desquelles on ne pourra que s'approcher sans les décrire parfaitement ? Certain-es y voient une exigence de l'esprit de quête de savoir et d'objectivité *per se*, une nécessité rationnelle ou une joie de la compréhension (Aristote, IVème siècle av. JC ; Paty, 2007), d'autres une terreur de l'inconnu (Joliot, 1947) ou un désir de puissance (Nietzsche, 1886). Quoiqu'il en soit, à l'ère de l'Anthropocène où l'être humain a un impact sans précédent sur la biosphère et le déclin des espèces (Figure 1.2), la conservation du Vivant est un enjeu de premier plan (Pimm et al., 2004 ; Rapport IPBES, 2019). Et il est nécessaire, pour préserver les espèces et leur écosystème, de décrire et comprendre leurs fonctionnements et rôles au sein de l'écosystème, leurs interactions avec d'autres espèces, de connaître leurs exigences environnementales, leurs réponses aux perturbations humaines, etc... afin de mettre en place les meilleurs outils de conservation et de gestion pour la biodiversité.

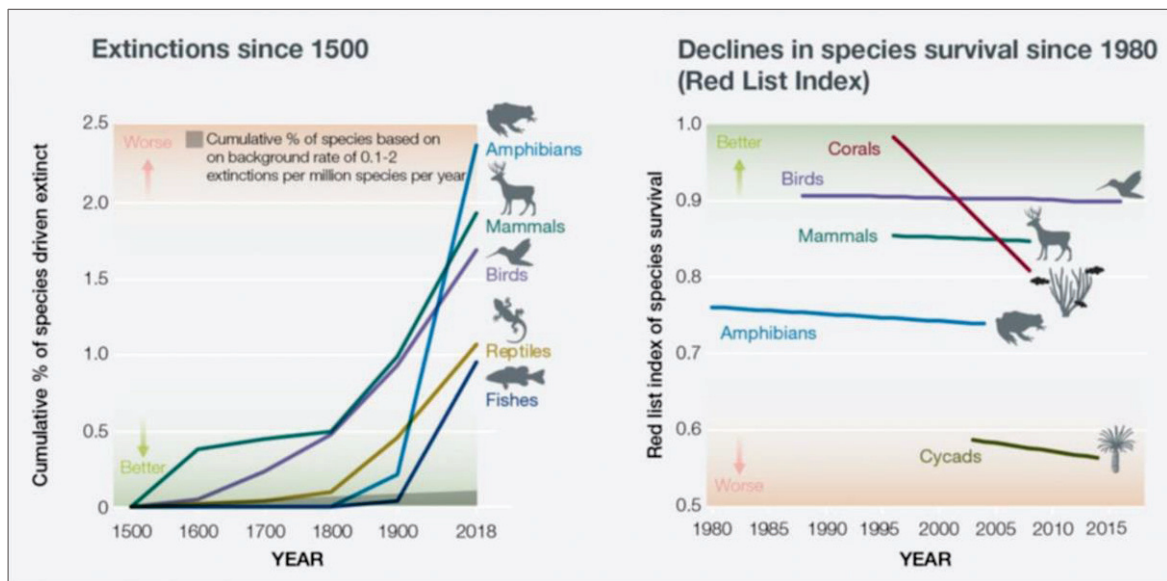


Figure 1.2 : Pourcentages cumulés d'extinction des espèces depuis 1500 pour les groupes de vertébrés (gauche) et évolution depuis 1980 de l'Indice Liste rouge représentant la survie des espèces pour certains groupes de la Liste rouge de l'UICN (droite). IPBES, rapport 2019.

2 Évolution des concepts en écologie des communautés

L'étude de la biodiversité peut se faire à plusieurs niveaux d'organisation : individu, population, communauté, écosystème, biosphère (Figure 2.1). Parmi elles, la communauté représente un groupe d'individus de plusieurs espèces différentes vivant dans le même espace géographique à un moment donné. L'écologie des communautés s'intéresse aux patrons spatio-temporels de distribution, d'abondance et d'interaction des espèces d'une même communauté et aux processus sous-jacents à ces



Figure 2.1 : Exemple des différents niveaux d'organisation pour l'étude de la biodiversité dans une mare. L'individu correspond à un individu d'une espèce donnée ; la population à l'ensemble des individus de cette même espèce ; la communauté à l'ensemble des espèces en interaction et l'écosystème, l'ensemble formé par la communauté en interaction avec l'environnement.

patrons (Lawton, 1999). Cette branche de l'écologie ne déroge pas à la règle en s'intéressant à des systèmes présentant de multiples facettes et pouvant être étudiés sous différents angles. Les communautés sont étudiées à travers différentes théories qui ont été mises en évidence et ont évolué au fur et à mesure du temps depuis le concept de niche écologique centrée sur l'échelle locale (Grinnell, 1917) jusqu'à aujourd'hui où différents processus et différentes échelles spatiales et temporelles sont abordés (Vellend, 2010).

2.1 Niche écologique

Le concept de niche écologique a été introduit en 1917 par Grinnell qui définit la niche comme tout facteur pouvant conditionner la présence d'une espèce à un endroit donné, qu'il soit abiotique (p.ex. température, pH, type de sols) ou biotique (p.ex. compétition, prédation, disponibilités des ressources). En 1927, Elton y ajoute une nouvelle dimension en mettant l'accent sur la place d'une espèce dans la chaîne trophique (herbivore, carnivore, etc...). En combinant ces deux visions de la niche, une espèce occupe dans une communauté un espace spécifique déterminé par (i) une niche spatiale, ensemble des facteurs biotiques et abiotiques favorables à la présence de l'espèce (niche *sensu* Grinnell) et (ii) une niche fonctionnelle, rôle de l'espèce dans l'écosystème (niche *sensu* Elton) qui, ensemble, forment la niche écologique, bien que les deux concepts ne soient pas totalement indépendants (Schoner, 1989 ; Whittaker, 1972). En 1957, Hutchinson met l'espèce au centre du concept en définissant sa niche

fondamentale correspondant à l'ensemble des valeurs des conditions environnementales biotiques et abiotiques favorables à l'établissement de l'espèce et sa niche réalisée qui correspond à une partie restreinte de la niche fondamentale réellement occupée par l'espèce dans un espace donné. La niche réalisée est dépendante de facteurs écologiques externes comme la compétition avec d'autres espèces (exclusion compétitive, Gause, 1935), la prédation ou la variabilité des conditions environnementales (Hutchinson, 1957). En résumé, quels que soient les auteurs, la niche écologique représente le lien entre chaque espèce et son habitat local via des processus liés aux facteurs abiotiques (filtre environnemental) ou biotiques (interactions entre espèces) et définit localement la distribution des espèces dans les communautés qui vont sélectionner les habitats correspondant à leur niche écologique.

2.2 Dispersion

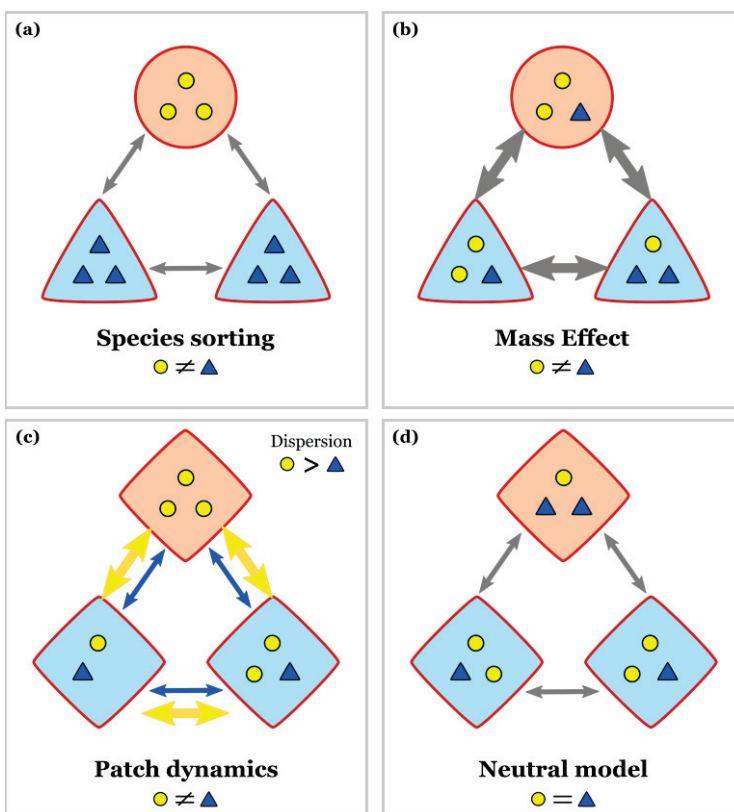


Figure 2.2 : Les 4 paradigmes du concept de métacommunauté, présentés par Leibold en 2004. Les formes à contour rouge représentent des communautés locales avec des environnements différents. Chaque forme au sein d'une communauté locale représente un individu d'une espèce donnée, chaque forme représentant une espèce différente. La concordance des formes entre espèces et communautés représentent la concordance espèce-habitat en lien avec sa niche. Les double-flèches représentent la dispersion entre les communautés avec l'épaisseur représentant l'intensité de dispersion.

La dispersion correspond aux mouvements des organismes entre localités dans le temps et dans l'espace (Bilton et al., 2001). Elle a d'abord été prise en considération seule à travers la théorie de la biogéographie insulaire (MacArthur & Levins, 1967) qui la première intègre la notion spatiale dans la distribution des espèces d'une communauté. Selon cette théorie, la richesse en espèces d'une île dépend de sa distance au continent, réservoir d'espèces qui immigrent vers l'île et s'y établissent avec succès en fonction du nombre d'espèces déjà présentes et de la surface de l'île. De cette vision purement spatiale et par l'influence des idées issues de l'étude des métapopulations (Hansson, 1991), la dispersion a été intégrée à différentes échelles et en interaction avec les processus associés à la niche écologique via le concept de métacommunauté (Leibold et al., 2004). Une métacommunauté est un ensemble de communautés locales reliées entre elle par la dispersion de multiples espèces en

interaction. Ce concept propose d'étudier pour la première fois la structure des communautés à deux échelles spatiales : locale où les processus associés à la niche écologique sont en jeu et régionale où la dispersion fait le lien entre les communautés locales et le pool régional d'espèces (Encadré 1). A l'origine, plusieurs paradigmes ont été proposés en fonction de la part des processus de sélection de niche et de dispersion dans la structure des métacommunautés : le *species sorting* (sélection des espèces), le *mass effect* (effet de masse), le *patch dynamic* (dynamique des patches) et le *neutral model* (modèle neutre) (Figure 2.2, Leibold et al., 2004). Dans le *species sorting*, les conditions environnementales sont différentes entre patches d'habitat et déterminent la composition des communautés au sein de chaque patch. Ce paradigme suppose une dispersion non limitante dans la métacommunauté pour permettre aux espèces de choisir leur habitat et les processus locaux y sont les processus dominants (Figure 2.2a). Le paradigme de *mass effect* est centré sur les dynamiques spatiales où la migration de patches « sources » (p.ex. de bonne qualité) vers des patches « puits » (p.ex. de mauvaise qualité) permet aux communautés de se maintenir dans des habitats non optimaux. L'effet local est amoindri par ces dynamiques et la dispersion y est le processus dominant (Figure 2.2b). Dans les paradigmes *patch dynamics* et *neutral model*, les habitats sont supposés identiques.

Encadré 1. Échelles étudiées et mesures de diversité associées dans l'étude des métacommunautés

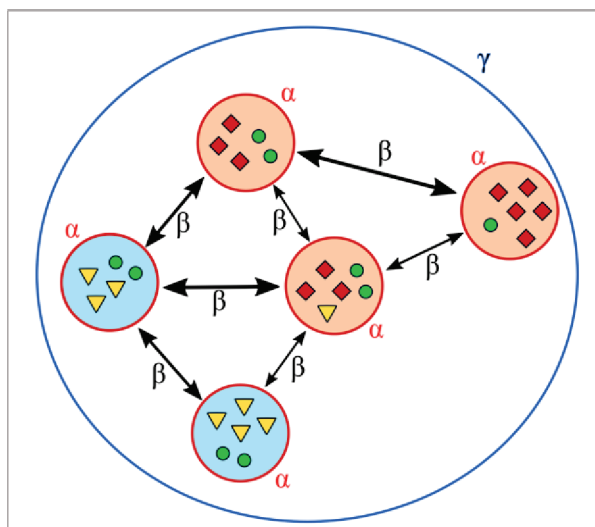


Figure 2.3 : Échelles spatiales considérées dans l'étude des métacommunautés. Les cercles rouges représentent des communautés locales présentant des conditions environnementales différentes (couleurs) et reliées entre elles par la dispersion (double flèche), le cercle bleu représente la métacommunauté. Chaque forme au sein d'une communauté locale représente un individu d'une espèce donnée, chaque forme représentant une espèce différente.

L'étude des métacommunautés s'intéresse à deux échelles spatiales : locale et régionale. Chaque communauté locale est un ensemble d'espèces vivant en interaction dans un habitat uniforme et un espace délimité (p.ex. une île dans un archipel). Les facteurs environnementaux et les interactions entre espèces vont impacter la composition et diversité en espèces de chaque communauté (Whittaker et al., 2001). On peut décrire chaque communauté par l'indice de diversité α qui représente la diversité spécifique de chaque communauté (p.ex. le nombre d'espèces). A l'échelle régionale, les communautés sont reliées entre elles par la dispersion. La métacommunauté représente l'ensemble des communautés locales (p.ex. un archipel d'îles). Elle est définie par sa diversité, mesurée par la diversité γ , et par le degré de différenciation entre communautés qui est mesurée par la diversité β (Marcon, 2015).

Dans le paradigme *patch dynamics*, la dynamique est dictée par les différences de dispersion des espèces couplée à des événements de colonisation et d'extinction locale, stochastiques ou déterministes (Figure 2.2c). Le paradigme *neutral model* suppose quant à lui des espèces identiques où la dynamique ne correspond qu'à des probabilités de gain et de perte d'espèces (Figure 2.2d) (Brown et al., 2017 ; Leibold et al., 2004). Ces paradigmes ont tout d'abord été considérés comme les quatre seuls types de métacommunautés existants, exclusifs les uns des autres, et les études cherchaient à connaître quel paradigme structurait telle ou telle métacommunauté (Cottenie, 2005 ; Logue et al., 2011). Mais dans la dernière décennie, des auteur·trices ont remis en question ces "big 4" en démontrant, qu'en plus de ne pas être exclusifs, ils n'étaient que des points de référence permettant d'illustrer le nouveau concept et le focus sur la seule délimitation des paradigmes entre eux représentait un frein pour l'avancée de l'écologie des communautés en tant que discipline (Brown et al., 2017 ; Winegardner et al., 2012). Elles et ils suggèrent de voir le concept de métacommunauté de manière linéaire et inclusive combinant processus locaux et régionaux plutôt que sous la forme d'une classification de quatre paradigmes bien délimités. Désormais, les métacommunautés sont principalement étudiées en mettant l'accent sur la part respective des processus dans la distribution des espèces comme la limitation à la dispersion (p.ex. Finn & Poff, 2011), le *mass effect* (p. ex. Richter-Boix et al., 2007), l'environnement local (p. ex. Heino et al., 2012) ou les interactions biotiques (p. ex. García-Girón et al., 2020), bien que l'approche « paradigmes » soit encore utilisée (Vieira et al., 2020). Ainsi, si la dispersion est limitée par les capacités de dispersion des espèces (Heino et al., 2017) ou par des facteurs extérieurs (p.ex. barrières géographiques, Blanchet et al., 2020), une espèce ne peut pas être présente dans toutes les communautés correspondant à sa niche écologique. A l'inverse, si la dispersion est forte, une dynamique source-puits peut se mettre en place comme dans le paradigme *mass effect*. Dans les deux cas, le rôle de la dispersion sera plus important que celui de l'habitat local dans la distribution des espèces. Si la dispersion est suffisante, une espèce peut être présente dans tous les habitats favorables à sa niche et la distribution des espèces sera déterminée par leurs niches écologiques, rejoignant en partie le paradigme de *species sorting*.

2.3 Histoire évolutive

L'intégration de l'histoire évolutive en écologie des communautés provient du constat que les écologues se basaient seulement sur l'étude des processus locaux agissant à une échelle temporelle fixe pour comprendre la distribution des espèces au sein des communautés. L'ensemble des processus agissant à des échelles spatiales et temporelles plus larges dans lesquelles les communautés sont incorporées n'étaient pas prises en compte (Ricklefs, 1987). Pourtant, deux habitats très similaires ne possèdent pas systématiquement le même cortège d'espèces car les patrons de distribution passés liés à des événements paléo-climatiques (p.ex. glaciation, régression marine) et/ou paléogéographiques (p.ex. dérive des continents) peuvent expliquer une part des patrons de distribution actuels (Graham &

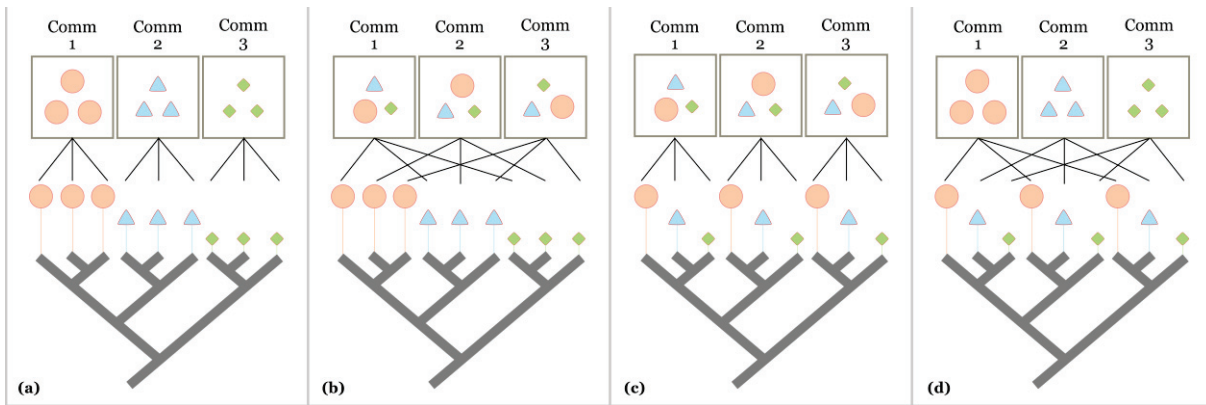


Figure 2.4 : Structure des communautés en lien avec la phylogénie. Les traits phénotypiques associés à la niche écologique (couleur et forme) peuvent être conservés au sein de l'évolution (a), (b) ou non (c), (d). Les espèces peuvent être regroupés phylogénétiquement dans les communautés (a), (c) ou non (b), (d). Chaque configuration permet d'inférer sur les processus ayant opérés au sein des communautés. (a) Conservation des traits et regroupement phylogénétique dans les communautés traduitent un fort filtre environnemental. (b) Conservation des traits et répartition phylogénétique dans les communautés traduitent une forte interaction entre espèces. (c) Changement phylogénétique des traits et regroupement phylogénétique dans les communautés traduitent une adaptation évolutive des traits vis-à-vis de l'occupation de niche. (d) Changement phylogénétique des traits et répartition phylogénétique dans les communautés traduitent une adaptation vis-à-vis de la niche et un fort filtre environnemental. Schéma reproduit à partir de Emerson & Gillespie, 2008.

Fine, 2008 ; Leibold et al., 2010). En 2002, Webb et al. introduisent la prise en compte de l'histoire évolutive des espèces en écologie des communautés via la phylogénie. Pour elles et eux, d'après le concept de niche écologique, les espèces interagissent au sein des communautés à travers la variabilité de leurs traits phénotypiques (traits biologiques, fonctionnels, comportementaux déterminant la niche d'une espèce), eux-mêmes basés sur leurs histoires évolutives. Ainsi, il s'agit de mettre en évidence si la distribution actuelle des espèces est corrélée à la phylogénie et si le patron observé permet d'inférer sur la part du filtre environnemental et des interactions biotiques dans la structure ancienne des communautés (Webb et al., 2002). Par exemple, si les traits phénotypiques sont phylogénétiquement conservés et que la distribution des espèces dans les communautés se fait par groupes phylogénétiques et donc phénotypiques, cela traduit un fort filtre environnemental (Figure 2.4a). A l'inverse, si la distribution des espèces dans les communautés ne se fait pas par groupes phylogénétiques, cela traduit une forte exclusion compétitive entre espèces proches (Figure 2.4b, Emerson & Gillespie, 2008). Mais cette approche reste centrée sur la niche et néglige les processus agissant à une échelle spatiale plus large, qui, s'ils sont supposés, ne sont jamais mesurés (phylogénie niche-centrée). La mesure de la diversité phylobeta a été proposée pour intégrer explicitement les larges échelles spatiales dans les analyses phylogénétiques de la structure des communautés. Cet indice mesure comment les distances phylogénétiques des communautés changent au sein d'un paysage pour intégrer les échelles locale et régionale directement dans la description des communautés. Ainsi, cette mesure permettrait d'avoir un accès simultané de la part des interactions biotiques, de l'environnement, de l'isolement géographique passée ou présente et des contraintes phylogénétiques dans la structure des communautés (Emerson et al., 2011).

2.4 Stochasticité

Les processus présentés jusqu'ici sont déterministes mais une part de hasard et d'aléatoire est aussi en jeu dans la distribution des espèces dans les communautés avec un nombre fini d'individus (MacArthur & Levins, 1967). A contre-pied du concept de niche écologique où la composition des communautés est déterminée par les différences entre espèces (de niche, de capacités de dispersion), Hubbell (2001) propose la théorie neutre dans laquelle les espèces sont équivalentes d'un point de vue démographique (c'est-à-dire même taux de natalité et de mortalité, nombre de descendants et de migrants) quel que soit l'habitat. Ainsi, la dynamique des communautés est stochastique et ne dépend pas de leurs compositions. Les patrons de distribution des espèces sont déterminés par des événements de spéciation, d'extinction et des événements d'immigration entre communautés dont l'intensité est dépendante de la taille de la communauté et de l'abondance de chaque espèce.

2.5 Rassembler les approches

Au final, l'écologie des communautés présente de multiples théories et concepts ayant chacun leurs processus propres qui se sont ajoutés les uns aux autres au cours du temps et qui ont tous leur pertinence. A ceci s'ajoute la difficulté à déterminer le lien entre processus et patrons observés car plusieurs patrons observés empiriquement peuvent résulter de différents processus (Lawton, 1999). L'étude de la structure des communautés représente ainsi une discipline complexe pour laquelle des règles d'assemblages générales semblent ne pas pouvoir être établies (Belovsky et al., 2004 ; Lawton, 1999). En 2010, Vellend tente pourtant de concilier toutes ces approches en proposant la « théorie des

communautés écologiques » dans laquelle 4 processus sont proposés : (i) la sélection qui est l'équivalent du concept de niche écologique avec pour facteurs importants la réponse des espèces à l'environnement abiotique, aux ressources, aux perturbations et aux interactions entre espèces ; (ii) la spéciation qui intègre le taux de spéciation et l'origine du pool des espèces (domaine de la biogéographie et de la macroévolution), qui, à travers son effet sur le pool régional, est un déterminant indirect dans la distribution des espèces à l'échelle locale ; (iii) la dispersion qui est le même processus que présenté précédemment et (iv) la dérive écologique qui correspond à la composante stochastique. Tous les processus ne sont pas exclusifs et interagissent entre eux à toutes les échelles spatiales et

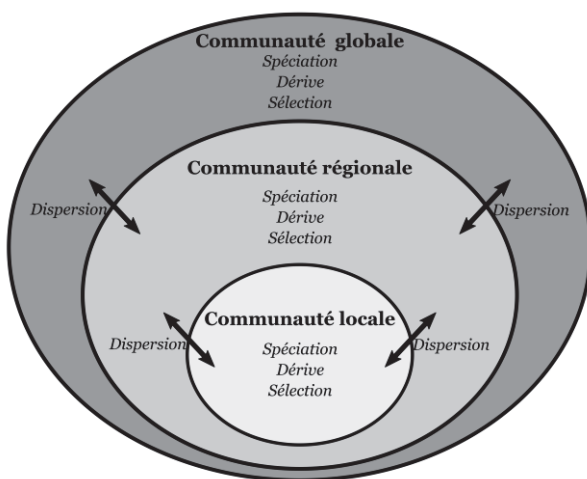


Figure 2.5 : Théorie des communautés écologiques selon Vellend, 2010. Les processus structurant les communautés interagissent entre eux à travers les différentes échelles spatiales. Reproduite à partir de Vellend, 2010.

temporelles (Figure 2.5). Les espèces sont ajoutées aux communautés via la spéciation et la dispersion, leurs abondances relatives sont déterminées par la sélection et la dérive ainsi que par la dispersion qui joue sur la dynamique des communautés. Ce cadre théorique permet d'inclure tous les patrons observés dans la nature mais reste très général, ce qui pourrait amener à toujours retrouver un cas unique où tout patron observé serait le résultat d'une interaction entre tous ces processus sans comprendre la manière avec laquelle ils interagissent ou quels processus sont dominants dans la distribution des communautés. De plus, il reste difficile de prendre en compte tous les processus simultanément dans une étude (Belovsky et al., 2004).

En résumé, la distribution des communautés peut être expliquée dans un cadre général, celui proposé par Vellend dans lequel peuvent s'inscrire tous les concepts présentés précédemment. Pour décrire plus finement les différentes facettes de cette distribution, l'écologie des communautés met à disposition des outils intellectuels riches et variés à utiliser selon l'échelle spatiale et temporelle que l'on considère et aux processus associés (Figure 2.6).

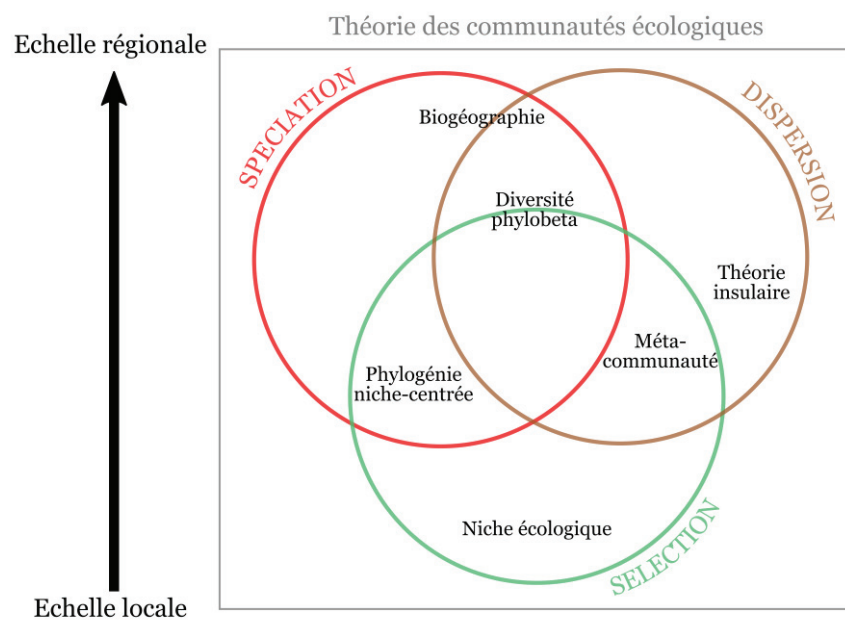


Figure 2.6 : Schéma présentant plusieurs concepts utilisés en écologie des communautés et présentés dans ce manuscrit, replacés en fonction des différents processus proposés par Vellend, 2010 et de l'échelle spatiale considérée. La dérive n'est pas présentée car elle agit à toutes les échelles spatiales et en interaction avec tous les processus (Vellend, 2010).

3 Écologie des communautés dans les rivières : vers une prise en compte de la dispersion

Les rivières sont essentielles pour l'être humain. Depuis des centaines d'années, il s'est installé près des rivières pour l'approvisionnement en eau, la production et l'irrigation des cultures, la

nourriture qu'elles fournissent, les voies de navigation et le commerce, la production d'énergie mais aussi pour leur valeur récréative, culturelle et esthétique de par la biodiversité unique qu'elles renferment (Böck et al., 2018 ; Chevallier, 2007). Les écosystèmes de rivières font partie, avec l'ensemble des eaux douces, des écosystèmes les plus menacés actuellement (Encadré 2) avec un déclin de 83% des abondances moyennes de 1970 à 2014 dans des populations de référence (Indice Planète Vivante) tandis qu'ils représentent moins de 1% de la surface du globe et hébergent plus de 10% des espèces animales décrites (WWF, 2018). Il est important de comprendre comment s'organise la biodiversité dans ce hotspot de biodiversité (Strayer & Dudgeon, 2010) afin de guider les politiques de gestion et de conservation à leur égard.

Les rivières sont organisées dans le paysage en réseau hydrographique, un ensemble hiérarchisé de chenaux assurant le drainage superficiel d'un bassin versant ou d'une région (Chevallier, 2007). Un réseau hydrographique est constitué en amont de chevelus de rivières qui forment les têtes de bassin versant qui se regroupent en aval dans des cours d'eau principaux plus larges qui se jettent dans la mer ou l'océan par un estuaire ou un delta, le cas échéant (Figure 3.1). Les réseaux hydrographiques présentent des particularités structurelles et fonctionnelles qui rendent les écosystèmes de rivières uniques au regard des autres écosystèmes et qui impactent la distribution des espèces dans leurs communautés : le dendritisme du réseau, le flux unidirectionnel du courant de l'amont vers l'aval et le dynamisme du régime hydrologique (Brown et al., 2011 ; Tonkin et al., 2017).

Un réseau dendritique est défini comme une structure hiérarchique en arborescence constitué d'un bras principal pourvu d'un nœud donnant deux branches et ainsi de suite. Les branches diminuent en

Encadré 2. Menaces pesant sur les écosystèmes de rivières

Les menaces principales de la biodiversité des rivières ont été regroupées en cinq catégories (Dudgeon et al., 2006) : (1) la surexploitation, qui affecte principalement les vertébrés (Arshad & Umar, 2020), (2) la pollution de l'eau par l'apport de nutriments ou de pesticides par l'agriculture ou les rejets domestiques (Wato, 2020), (3) les altérations du débit qui sont les conséquences de rectification des lits des rivières, pompages, mise en place de barrages et de retenues d'eau (Poff & Zimmerman, 2010), (4) la dégradation et la destruction des habitats comme par exemple l'exploitation du substrat des lits des rivières ou les modifications liées à la pollution de l'eau et la régulation des débits (Eby et al., 2005) et (5) l'invasion par des espèces exotiques (Devin et al., 2001). Ces menaces agissent de manière combinée et en interaction sur fond de changement climatique qui exacerbe leurs effets sur la biodiversité aquatique, par exemple en accélérant l'altération des cycles hydrologiques (augmentation des assèchements et des crues) ou celle des conditions environnementales des habitats (augmentation de la température). A ces menaces majeures s'ajoutent 12 menaces émergentes identifiées par Reid et al. en 2010. Ces nouvelles menaces sont une conséquence de l'expansion de la population humaine, du commerce libéral et de l'industrialisation associée dont l'augmentation du nombre des centrales hydroélectriques, la pollution aux micro-plastiques, la pollution sonore et lumineuse et les nouveaux

taille et largeur et augmentent en nombre hiérarchiquement de bas en haut (Figure 3.1, Campbell Grant et al., 2007). Contrairement aux systèmes en patches où la dispersion peut se faire dans toutes les directions, la dispersion est contrainte par la structure du réseau où les nœuds servent de points de transfert. Ainsi, la distance effective entre deux points au sein du réseau sera plus longue dans les parties amont du réseau que dans les parties aval (Figure 3.1a, Altermatt, 2013). A cela s'ajoute la topographie du réseau hydrographique qui induit un écoulement de l'eau unidirectionnel de l'amont vers l'aval (Figure 3.1) (Benda et al., 2004 ; Brown et al., 2011). Cet écoulement entraîne une dispersion des individus asymétrique et biaisée vers l'aval (Altermatt, 2013). Ces deux caractéristiques du réseau influencent les caractéristiques hydrauliques (p.ex. largeur de lit, débit, Malavoi & Bravard, 2010) et écologiques (p.ex. transfert des sédiments et des nutriments organiques et inorganiques, Vannote et al., 1980) de l'amont vers l'aval. Ainsi, les habitats disponibles pour les espèces et donc leur répartition dans le réseau hydrographique varient selon un gradient longitudinal (Figure 3.1b). De nombreux concepts ou typologies existent pour expliquer la distribution des espèces dans les écosystèmes de rivières, d'abord en se centrant sur le gradient amont-aval (Huet, 1949 ; Illies & Botosaneanu, 1963 ; Vannote et al., 1980 ; Verneaux, 1973) puis en intégrant les échanges avec la plaine alluviale (flood pulse concept, Junk et al., 1989), les berges et la zone riparienne (riverine productivity model, Thorp & Delong, 1994) ou les caractéristiques des vagues fluviales longitudinalement et latéralement (Humphries et al., 2014). La distribution des espèces est aussi influencée par le régime hydrologique qui est dynamique dans le temps. Le régime hydrologique est caractérisé par des variations de débit des rivières qui se répètent dans le temps (période de crues, d'étiage et/ou d'assèchement) et dans l'espace (sections du réseau exposés à différents degrés à ces phénomènes) (Lytle & Poff, 2004). Ces variations sont influencées à la fois par le climat, l'hydrologie et le topologie (Poff et al., 1997). Le dynamisme du régime hydrologique est une composante naturelle d'un réseau hydrographique qui participe à la structure physique de l'environnement (p.ex. transport de matériaux abiotiques et biotiques, morphologie du lit des rivières). Elle entraîne des variations des conditions environnementales localement et impacte aussi la dispersion des organismes. Par exemple, lors des crues, certains organismes sont emportés par le courant par dérive et colonisent des sites en aval (Waters, 1972). A l'inverse, lors de la période d'étiage, des assèchements de certains tronçons peuvent avoir lieu et altérer la connectivité au sein du réseau (Boulton et al., 2017 ; Datry et al., 2016b).

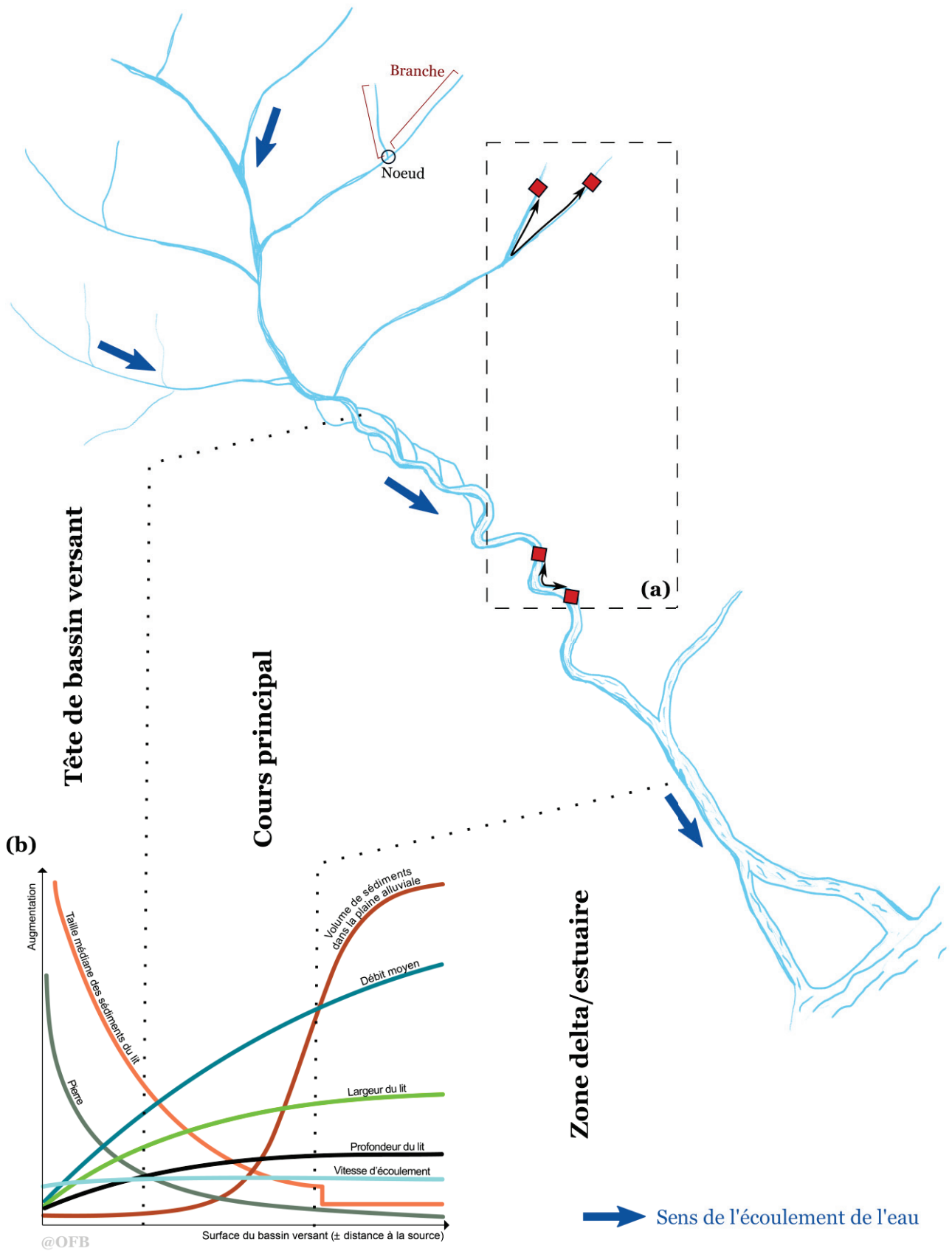


Figure 3.1 : Schéma montrant les différentes parties d'un réseau hydrographique ainsi que sa structure dendritique. (a) Comparaison de la distance à parcourir via le réseau entre deux communautés à même distance à vol d'oiseau suivant leur position dans le réseau. La distance effective est plus grande dans les têtes de bassin dû à la structure hiérarchisée du réseau. (b) Évolution de certaines variables physiques le long du gradient amont-aval.

Les écosystèmes de rivières sont donc des systèmes très structurés et contraints spatialement. Malgré les caractéristiques précédemment citées montrant que la dispersion est contrainte dans les réseaux hydrographiques, la recherche en écologie des communautés dans les rivières s'est focalisée pendant longtemps sur l'échelle locale et les processus de niche contrairement à d'autres disciplines telle que l'écologie des populations (Allendorf & Phelps, 1981 ; Stoeckel et al., 1997). L'avènement du concept de métacommunautés a proposé un cadre de travail pertinent pour l'étude de la structure des communautés dans les écosystèmes de rivières. Cela a entraîné une prise en compte exponentielle de la dispersion dans la structure des communautés de rivières dans les dernières années (Brown et al., 2011). De nombreuses théories et études se sont évidemment intéressées à la comparaison entre les différentes parties des réseaux en différenciant les parties amont, correspondant aux têtes de bassin versant, des parties aval, correspondant aux cours principaux (Brown & Swan, 2010). Dans ces deux parties, en plus des différences de conditions hydrauliques et environnementales, les contraintes à la dispersion y sont supposées différenciées. Ainsi, Brown et Swan proposent en 2010 un cadre conceptuel présentant les contributions relatives des processus structurant les métacommunautés des têtes de bassin versant et des cours principaux. En aval, la dispersion est le processus dominant dû à l'apport des organismes par les crues et le flux amont/aval associés à la structure du réseau. Selon le paradigme de *mass effect* (section 2.2), cela masquerait le rôle des processus locaux (associés à la niche écologique). A l'inverse, dans les têtes de bassins versants, les communautés sont isolées dû à leur position amont couplée à la dispersion biaisée vers l'aval dans le réseau et les échanges entre communautés faibles. Les processus dominant en amont sont donc plutôt associés à des processus locaux, notamment au fort filtre environnemental selon le paradigme de *species sorting* (section 2.2). Ce cadre théorique a été accepté et adopté largement par la plupart de la communauté scientifique (Brown et al., 2017 ; Brown et al., 2018 ; Carrara et al., 2012 ; Clarke et al., 2008 ; Heino et al., 2015b ; Schmera et al., 2018).

Pourtant, il a été récemment mis en évidence que de nombreux facteurs influençaient les processus liés à la niche écologique (filtre environnemental et interactions entre espèces) et à la dispersion dans les réseaux hydrographiques. L'échelle spatiale et la variation de l'environnement associée sont des facteurs importants agissant sur les processus de la dynamique des métacommunautés (Göthe et al., 2013 ; Heino et al., 2015a ; Viana & Chase, 2019). Par exemple, la limitation à la dispersion est supposée plus forte à mesure que l'échelle spatiale augmente car les individus n'ont pas la capacité d'évaluer exhaustivement les conditions environnementales de tous les habitats de la métacommunauté pour s'établir dans leur habitat optimal (Heino et al., 2015a). La topologie et l'hydrogéologie des réseaux ont aussi un impact sur les dynamiques de dispersion comme la localisation des assèchements sur un bassin versant. Par exemple, des assèchements en amont induisent des contraintes à la dispersion plus fortes que lorsqu'ils sont situés en aval (Crabot et al., 2020). Les hypothèses de la théorie de Brown & Swan ont été effectuées pour des réseaux sans

perturbations humaines. Or, les écosystèmes de rivières font partie des plus menacés et les altérations qu'ils subissent peuvent impacter fortement les processus structurant leurs métacommunautés comme la modification du mode d'occupation des sols et la pollution potentiellement associée qui modifient les caractéristiques de l'habitat (Jones et al., 2020 ; Turunen et al., 2018) ou la fragmentation qui limite la dispersion (Fuller et al., 2015 ; Poff & Hart, 2002). Pour finir, l'influence des processus peut varier selon les capacités et modes de dispersion des espèces (Cañedo-Argüelles et al., 2015 ; Heino et al., 2015a). Des espèces strictement aquatiques comme les poissons devraient être beaucoup plus contraintes spatialement que des espèces pouvant disperser par voie terrestre comme les stades adultes de certains insectes. Dans une étude sur la structure spatiale de deux invertébrés, un strictement aquatique, *Gamarrus fossarum*, et l'autre avec une phase adulte aérienne, *Baetis rhodani*, Alp et al. (2012) ont montré une limitation à la dispersion forte chez *G. fossarum* tandis que *B. rhodani* ne présentait aucune structure spatiale.

En résumé, les écosystèmes de rivières sont des systèmes complexes où la distribution des espèces dans les communautés est déterminée par de multiples facteurs et dépend du contexte de l'écosystème. Ainsi, prendre en considération le contexte par rapport à ces différents facteurs pour mettre en évidence des règles d'assemblage des communautés dans les écosystèmes de rivières, que ce soit en tête de bassin versant ou dans les cours principaux, semble être une nécessité.

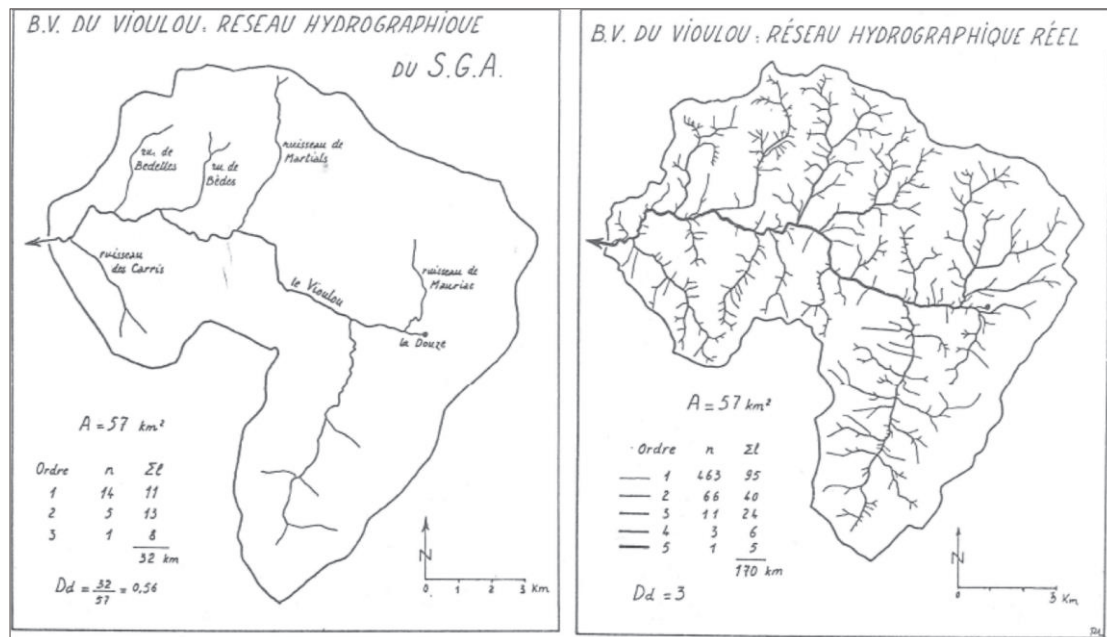
Encadré 3. Tête de bassin versant : définition floue et négligence

Figure 3.2 : Cartes montrant un réseau de référence du Service Géographique de l'Armée (gauche) et le réseau réel prenant en compte les têtes de bassin versant (droite). Lambert, 1996.

A ce jour, aucune définition claire des têtes de bassin versant n'existe. Rien ne définit réellement les limites des têtes de bassin versant et la seule certitude est qu'elles représentent la partie la plus amont des réseaux hydrographiques. Dans une synthèse bibliographique, Clarke et al. (2008) montrent que selon les auteurs, les têtes de bassin versant sont définies par un rang maximum de Strahler ou par la taille du bassin qui peut varier entre études d'un facteur de 500 (entre 12 et 6000 ha). Dans sa thèse, Lhéritier (2012) présente toutes les définitions que peut avoir une tête de bassin versant et en conclut que les limites d'une tête de bassin versant sont variables, d'un pays à l'autre, d'une région à l'autre, se basant sur des zonations, la superficie ou les rangs de Strahler. Cette absence de critères et de valeurs associées est sûrement due à la négligence générale à l'égard des petits cours d'eau, déjà remarqué en 1869 par Elisée-Reclus, en contraste à l'intérêt porté aux grands fleuves d'une importance économique et sociétale considérable. Comme on peut le voir sur la Figure 3.2, la prise en compte des rivières amont dans le référencement des réseaux hydrographiques était inexistante dans le passé, ce qui est toujours d'actualité (Benstead & Leigh, 2012). Heureusement, un regain d'intérêt pour les têtes de bassin versant est en cours par les services de l'État, les agences de l'eau, les associations et le domaine scientifique (Baattrup-Pedersen et al., 2018) attestant de l'importance enfin reconnue de ces systèmes pour l'intégrité des réseaux hydrographiques.

4 Spécificités des métacommunautés de têtes de bassin versant, milieux très dynamiques et fragmentés

A l'amont des réseaux hydrographiques se trouvent les « têtes de bassin versant » (Encadré 3). Les têtes de bassin versant constituent une part substantielle des réseaux hydrographiques avec, en Europe, 80 à 90% du linéaire total du réseau et 58% du continent représenté par leurs rivières et leurs bassins versants (Globevnik, 2007). De plus, elles jouent un rôle fondamental dans le bon fonctionnement des réseaux hydrographiques allant de l'apport de sédiments et de matière organique aux sections en aval à des zones de refuges ou de ponte pour certaines espèces (Meyer et al., 2007) et abritent une biodiversité unique (Clarke et al., 2008). Pourtant cette partie des réseaux a longtemps fait l'objet d'un désintérêt par les scientifiques et les politiques publiques (plus de détails dans l'encadré 3 et la section 8) et comprendre comment s'organise la biodiversité dans ces systèmes est primordial tant d'un point de vue fondamental, pour la compréhension de la dynamique des métacommunautés dans les écosystèmes de rivières, qu'appliqué, pour guider les pratiques de gestion ou de conservation.

Les têtes de bassin versant présentent plusieurs spécificités qui influent sur la niche écologique et la dispersion et sur leurs contributions respectives dans la distribution de leurs communautés. Bien que la niche écologique ait été présentée dans la section précédente comme l'unique processus à l'œuvre dans les têtes de bassin versant (Brown & Swan, 2010), ces spécificités pourraient rendre les processus basés sur la dispersion tout aussi primordiaux, voire prépondérants, dans ces systèmes. Tout d'abord, l'isolement des communautés de têtes de bassin versant dû à leur position en amont dans les réseaux mis en avant par Brown & Swan pour y justifier la prépondérance de la sélection de niche peut aussi

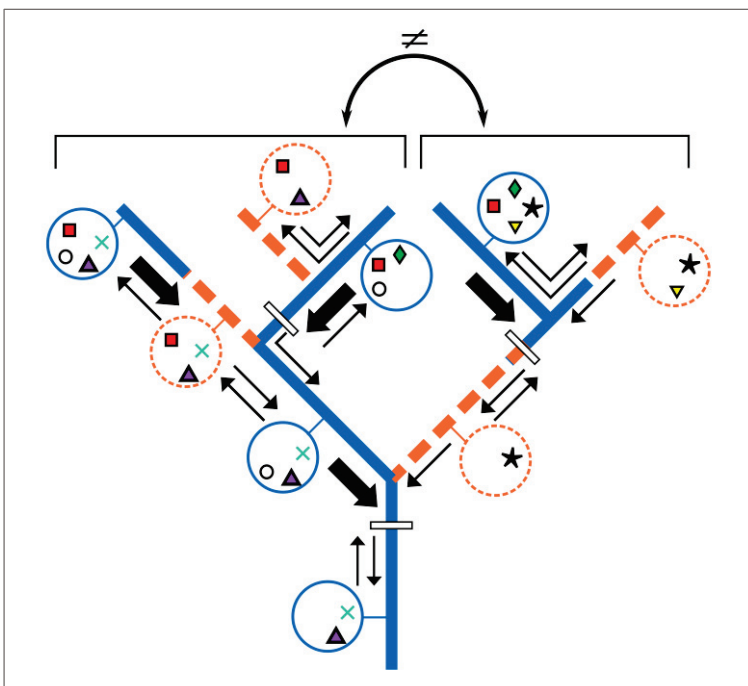


Figure 4.1 : Structure et dynamique d'une métacommunauté de tête de bassin versant contenant des tronçons intermittents (orange, pointillé) et des structures anthropiques fragmentant le réseau (rectangle blanc). L'intensité de dispersion est représentée par l'épaisseur des flèches. Les espèces sont représentées dans chaque communauté (cercle) par une forme de couleur différente. La limitation à la dispersion et le flux du courant amont-aval entraînant une dispersion biaisée vers l'aval freine les espèces de disperser dans tout le bassin, résultant en des communautés dans les deux parties du réseau avec des cortèges d'espèces différents. A cela s'ajoute les tronçons intermittents et les structures anthropiques qui représentent des obstacles à la dispersion. L'intermittence dans les tronçons intermittents induit une diversité alpha plus faible que dans les tronçons pérennes (bleu, plein) qui représentent des refuges et des sources de colonisateurs pour les tronçons intermittents. En conséquence, les communautés représentent un sous-ensemble de celles

4. Spécificités des métacommunautés de têtes de bassin versant, milieux très dynamiques et fragmentés

induire une limitation à la dispersion empêchant les espèces d'être présentes dans tous les habitats correspondant à leur niche écologique (Figure 4.1, Sarremejane et al., 2017b). Ensuite, l'approvisionnement en eau des rivières de têtes de bassin versant est intimement liée aux précipitations, aux réserves d'eau accumulées dans les nappes phréatiques ou les sols (Datry et al., 2017). Il en résulte un fort dynamisme de leur régime hydrologique avec des crues et des étiages sévères conduisant souvent, en association avec leur faible largeur de lit, à des assèchements partiels voire complets de leur réseau (Benstead & Leigh, 2012). A l'échelle locale, au niveau d'un tronçon intermittent, c'est-à-dire un tronçon où l'eau cesse de s'écouler ou dont le lit s'assèche entièrement à certaines périodes, trois phases hydrologiques se succèdent (Datry et al., 2014) : une phase lotique avec un écoulement de l'eau en surface (Figure 4.2a,b), une phase d'assèchement, lenticule avec un arrêt de l'écoulement et la présence de mouilles (Figure 4.2c) et une phase sèche où le lit est entièrement asséché (Figure 4.2d). De ce dynamisme résulte un changement constant et drastique à l'échelle locale des conditions environnementales faisant varier les habitats dans les tronçons intermittents et impactant la composition des communautés (Datry et al., 2016a). Les assèchements impactent la diversité en espèces qui est plus faible dans les tronçons intermittents que dans les tronçons pérennes (c'est-à-dire les tronçons qui présentent constamment un écoulement, Soria et al., 2017) et qui diminue d'autant plus que la durée de l'assèchement est longue (Crabot et al., 2020 ; Leigh & Datry, 2017). Ce déclin de la diversité spécifique est souvent accompagné du phénomène de *nestedness*, où les communautés avec une diversité faible sont des sous-ensembles de communautés plus riches taxinomiquement et où les tronçons subissant moins fortement les assèchements ou les tronçons pérennes servant de zone refuges pendant la perturbation (Figure 4.1 ; Datry et al., 2014). Ainsi, au niveau des tronçons intermittents, des événements d'extinction ou d'émigration vers et de recolonisation à partir de zones refuges ont lieu en lien avec les cycles d'assèchement et de remise en eau, ce qui rend crucial le rôle de la dispersion à cette échelle locale. Le filtre environnemental induit par les changements de conditions environnementales peut aussi résulter en la maintenance des espèces adaptées aux assèchements dans les tronçons intermittents en empêchant les espèces non



Figure 4.2 : Exemple des différentes phases hydrologiques observées sur un tronçon intermittent de la tête de bassin versant du Thouaret (Deux-Sèvres). Phases lotique à écoulement fort (a) et faible (b), phase lenticule ou d'assèchement (c) et phase sèche (d).

adaptées de s'y établir (Cañedo-Argüelles et al., 2015 ; Datry et al., 2017). Les assèchements vont aussi avoir des effets à l'échelle régionale sur la dispersion. En effet, les sections asséchées fragmentent le réseau de manière temporaire pendant la période d'étiage et limitent la dispersion des individus en diminuant la connectivité au sein du réseau (Cañedo-Argüelles et al., 2015). La connectivité hydrologique varie donc dans le temps au gré du dynamisme du régime hydrologique, de la configuration spatiale des phases hydrologiques et des habitats associés (Figure 4.3). On peut s'attendre à ce que la structure des métacommunautés soit dynamique dans le temps et que la part des processus structurant les métacommunautés dans ces systèmes dynamiques varient également au cours du temps (Datry et al., 2016a ; Sarremejane et al., 2017a).

Finalement, les têtes de bassin versant subissent des perturbations d'origine anthropique fortes (Tableau 4.1), liées en partie au désintérêt dont elles font l'objet (Lassaletta et al., 2010). Ces perturbations dérèglent le cycle hydrologique et le fonctionnement naturels de leurs rivières qui y sont particulièrement sensibles en comparaison aux rivières plus larges (Lowe & Likens, 2005). Du fait de la situation actuelle où les demandes en eau et les changements climatiques s'intensifient, ces perturbations ne feront qu'augmenter dans le futur (Dudgeon et al., 2006). Elles impactent le réseau soit en altérant son régime hydrologique naturel (Poff et al., 1997), soit en faisant office d'obstacles à l'écoulement (Fuller et al., 2015), soit les deux (Tableau 4.1). Une altération hydrologique est définie comme une perturbation d'origine anthropique du régime hydrologique naturel d'un cours d'eau pouvant présenter de multiples symptômes (Poff et al., 1997). Par exemple, l'imperméabilisation des sols peut augmenter l'intensité des crues tandis que la construction de barrages et de seuils induit une stabilisation du débit (Poff & Schmidt, 2016 ; Steward et al., 2012). La mise en place de retenues collinaires et les prélèvements entraînent une diminution du niveau d'eau moyen historique,

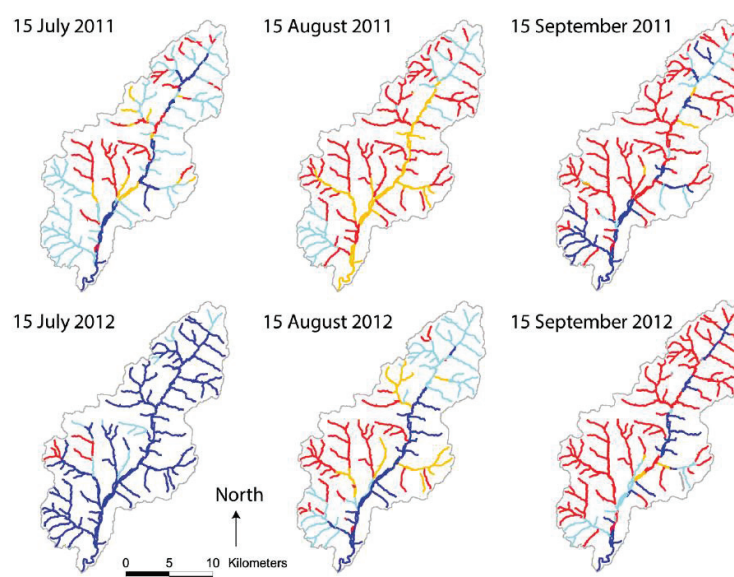


Figure 4.3 : Exemple des dynamiques spatio-temporelles des habitats et de la connectivité dans le bassin versant de la Tude, France. ; Bleu=Lotique ; Jaune=Lentique ; Rouge=Assèchement. Datry et al, 2016

4. Spécificités des métacommunautés de têtes de bassin versant, milieux très dynamiques et fragmentés

augmentant la probabilité des assèchements sur la section prélevée et son aval (Steward et al., 2012) tout comme le changement climatique peut augmenter la fréquence et la durée des assèchements (Acuña et al., 2014).

A l'échelle locale, la distribution des espèces est altérée par ces deux types de perturbations principalement en modifiant les conditions environnementales des habitats (Fuller et al., 2015 ; Poff & Zimmerman, 2010). Par exemple, la stabilisation du débit modifie la composition en espèces végétales des rivières, impactant les communautés de macro-invertébrés qui seront moins diverses et abondantes (Belmar et al., 2011 ; Mittal et al., 2015 ; Sheldon & Thoms, 2006). Les structures qui retiennent l'eau telles que les seuils et les barrages retiennent également les gros sédiments (Ward & Stanford, 1995). En conséquence, les secteurs en aval de ces constructions sont composés uniquement de sédiments fins et s'érodent, modifiant la géomorphologie et l'habitat et, en conséquence, le cortège d'espèces (Poff et al., 1997). De plus, ces constructions entraînent une modification des conditions hydrauliques, qui passent de lotiques à lenticues dans un certain périmètre autour d'elles, impactant également la composition des communautés dans ce périmètre (Figure 4.4, Fuller et al., 2015).

Tableau 4.1 : Ensemble des perturbations anthropiques observables dans les têtes de bassin versant et leur impact associé.

Perturbations	Obstacles à l'écoulement	Altérations hydrologiques
Barrages	X	X
Seuils	X	X
Buses et passage à gué	X	
Routes	X	
Pollution/Qualité de l'eau	X	
Introduction d'espèces invasives	X	
Prélèvements/Retenues		X
Modification de l'utilisation des sols		X
Réchauffement climatique		X



Figure 4.4 : Seuil en rivière qui altère hydrologiquement le réseau et représente un obstacle à l'écoulement et à la dispersion. Bassin de la Petite Barguelonne.

A l'échelle régionale, les perturbations d'origine anthropique augmentent la fragmentation longitudinale, la fragmentation le long du réseau hydrographique (Figure 4.1). La fragmentation longitudinale présente plusieurs caractéristiques (Fuller et al., 2015). Elle peut avoir des origines naturelles comme les cascades, ou anthropiques comme les seuils (Figure 4.4). Il en existe plusieurs types : des barrières physiques (p.ex. canyons, seuils), d'« habitat » qui sont des zones présentant des conditions physico-chimiques très différentes au sein du réseau (p.ex. zone lenticule ou

assèchements) ou biologique (p.ex. concentration de prédateurs ou de compétiteurs). Elle présente aussi une composante temporelle avec des fragmentations permanentes comme celles induites par les barrages ou temporaires illustrés par les assèchements (Lake, 2003 ; Taylor et al., 2018). Les perturbations anthropiques augmentant considérablement la fragmentation longitudinale sont les obstacles à l'écoulement qui agissent comme barrières physiques à la dispersion et freinent ou bloquent la dispersion des organismes (Tableau 4.1). Ces obstacles sont caractérisés dans les rivières par leur forme, leur perméabilité, la proportion d'individus capable de les traverser (passabilité), leur abondance et leur localisation (Fuller et al., 2015). Ils sont en général très nombreux et infranchissables (Gough et al., 2017). Ils augmentent donc significativement la fragmentation du réseau et ce de manière permanente, et induisent une diminution de la biodiversité à l'échelle locale en modifiant l'habitat disponible et en isolant les populations, notamment celles des espèces exclusivement aquatiques comme les poissons (Fahrig, 2003).

En résumé, les métacommunautés de têtes de bassin versant semblent considérablement influencées par plusieurs facteurs liés à la structure du réseau, à la dynamique du régime hydrologique et aux perturbations anthropiques. Ces facteurs peuvent agir de concert, voire en interaction, sur les processus liés à la niche écologique et à la dispersion. Par exemple, une accentuation de la fragmentation longitudinale par les structures anthropiques entraînerait une augmentation de l'isolement des communautés résultant en un effet synergique où la dispersion des organismes est encore plus limitée dans les têtes de bassin versant perturbées. Ces facteurs impliquent des processus liés à la dispersion et il semble nécessaire de les prendre en compte afin d'évaluer si la dispersion joue un rôle dans la structuration des métacommunautés de têtes de bassin versant et de guider au mieux leur gestion et leur conservation.

5 Modèle invertébrés dans l'étude des métacommunautés : intérêts et limites

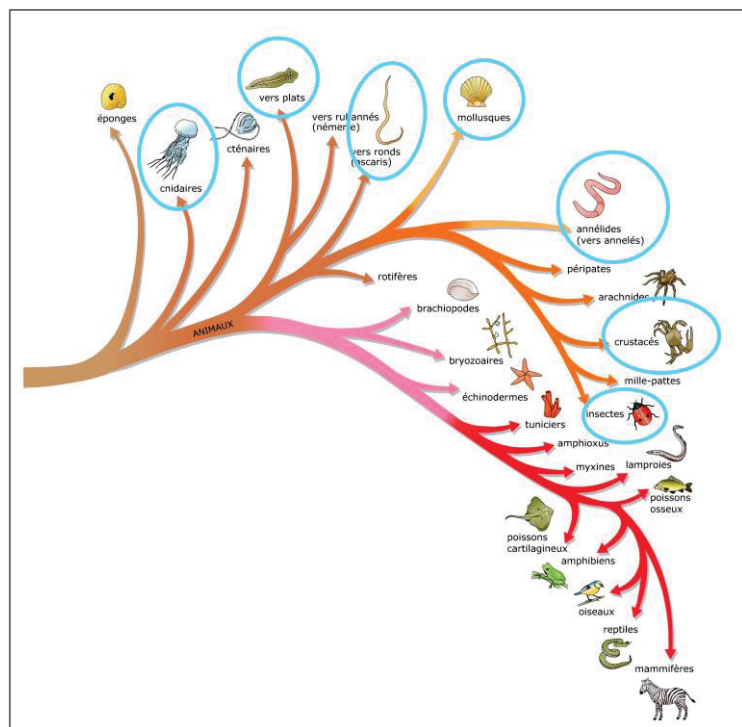


Figure 5.1 : Phylogénie des métazoaires montrant les groupes considérés comme faisant partie du groupe non valide des invertébrés dans le cadre des études de communautés de rivières (cercles bleu clair).

Les invertébrés sont regroupés sur la base d'une absence de caractère, l'absence de vertèbres. Or, ce taxon (c'est-à-dire unité taxinomique) n'est pas valide phylogénétiquement car non basé sur le partage d'un caractère commun de tous les descendants (Cavalier-Smith, 1998) et regroupe des clades d'embranchements très diverses au sein des métazoaires (Figure 5.1). De nombreuses études se focalisent pourtant sur ce groupe car il contient des organismes clés qui fournissent des services écosystémiques essentiels et variés tant dans le fonctionnement des écosystèmes (p.ex. décomposition de la matière organique) que dans la chaîne trophique (p.ex. source de nourriture importante) (Clarke et al., 2008 ; Wallace & Webster, 1996). Ils sont aussi utilisés comme bio-indicateurs de l'état des rivières du fait de leurs exigences particulières et spécifiques vis-à-vis des conditions

environnementales (Cairns & Pratt, 1993 ; Resh, 2008). Cet intérêt à leur égard et leur rôle essentiel dans les écosystèmes peut s'expliquer par la diversité en termes de traits biologiques, de préférence de niche écologique, de stratégies d'histoire de vie et d'interactions trophiques rencontrées dans ce groupe (Bonada et al., 2007 ; Resh, 2008 ; Sarremejane et al., 2020).

Cette diversité de traits en fait un bon modèle pour l'étude des métacommunautés. En effet, comme dit en section 3, les espèces sont influencées différemment par les processus structurant les communautés selon leurs modes et capacités de dispersion, dont les invertébrés présentent un panel immense (Figure 5.2). Par exemple, les crustacés sont strictement aquatiques et certains dispersent principalement dans l'eau tandis que les insectes possèdent, en général, un ou plusieurs stade(s) aquatique(s) suivi pour la grande majorité d'un stade adulte terrestre et ailé leur permettant de disperser par voie aérienne (Bilton et al., 2001). Certaines espèces dispersent de manière passive, c'est-à-dire qu'elles se déplacent via des facteurs extérieurs, c'est ce qu'on appelle la dérive, d'autres de manière active. Pour les invertébrés de rivières, la dérive se fait par le courant, notamment lors des

crues, par les animaux ou le vent (Bilton et al., 2001). A travers une catégorisation des modes de dispersion (Bohonak & Jenkins, 2003 ; Göthe et al., 2013) ou des métriques traduisant la capacité de dispersion permettant de quantifier la dispersion spatiale des espèces (Bonada et al., 2012 ; Li et al., 2018), il a été montré que les patrons de distribution des espèces dans les métacommautés variaient selon les groupes définis. Par exemple, chez les taxa dispersant activement par voie aérienne, la part de la structure des communautés expliquée par la sélection de l'habitat était plus forte que celle de la dispersion, supposant qu'elles sont capables de repérer les habitats avec les conditions environnementales adaptées à leur niche pour leur établissement sans être limitées par la distance (Gronroos et al., 2013). Li et al. (2016) observent que des espèces possédant de fortes capacités de dispersion sont les premières à coloniser des tronçons nouvellement restaurés et vierges de tout organisme, suivies l'année suivante par des espèces avec des capacités de dispersion plus faibles, montrant que les capacités de dispersion des espèces jouent un rôle important dans la dynamique spatiale et temporelle des métacommautés d'invertébrés de rivières.

Les invertébrés présentent aussi des stratégies de réponse différentes face aux perturbations. Dans le cas des assèchements, deux stratégies sont observées. Soit les espèces possèdent des traits de résistance aux assèchements sous forme de tolérance à la dessiccation, de diapause, d'enkystement ou même d'adaptations comportementales (Lytle & Poff, 2004 ; Vander Vorste et al., 2016), qui leur permettent de subsister dans les tronçons asséchés et d'émerger lors de la remise en eau ; soit les espèces ont des traits de résilience liés aux traits de dispersion qui leur permettent de revenir s'installer rapidement sur les tronçons asséchés et atteindre leur état avant perturbation (Chester & Robson, 2011 ; Fagan, 2002). Les espèces dispersant massivement, activement et sur une longue distance représentent les espèces les plus résilientes. La réponse aux perturbations d'origine anthropique, notamment la pollution des eaux, est aussi très variée avec des espèces tolérantes comme les

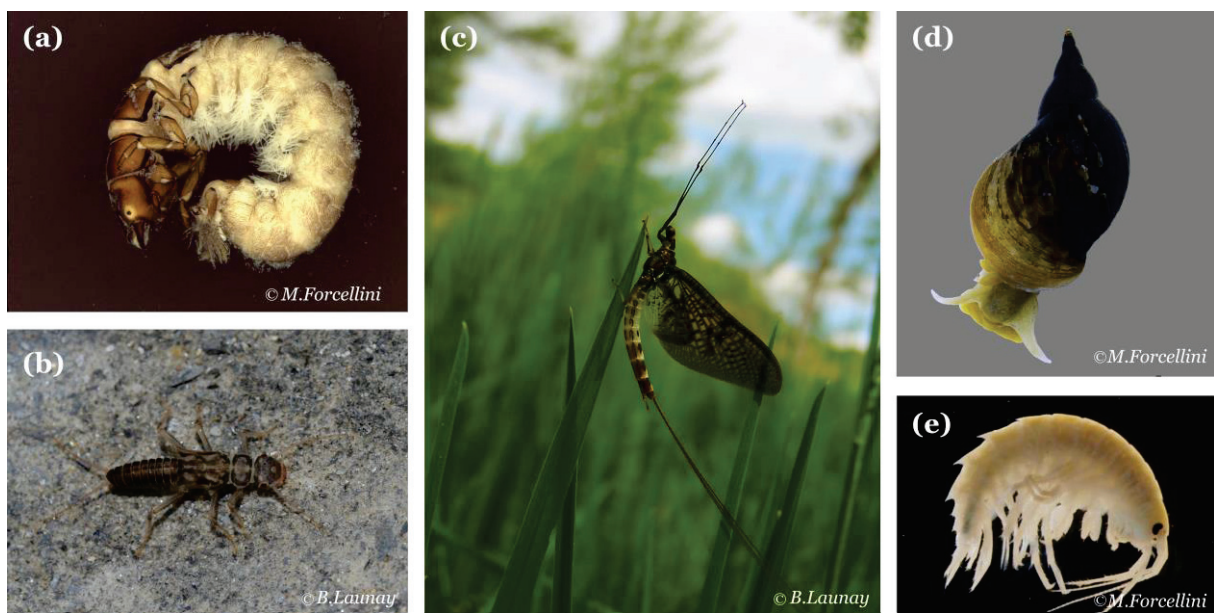


Figure 5.2 : Invertébrés aquatiques que l'on peut rencontrer dans les têtes de bassin versant et qui présentent des modes de dispersion variés. (a) *Hydropsyche incognita*, Trichoptera, (b) *Isoperla rivulorum*, Plecoptera, (c) *Ephemera danica*, Ephemeroptera, (d) *Lymnaea stagnalis*, Gastropoda et (e) *Gammarus roeseli*, Amphipoda. (a), (b) et (c) sont des insectes possédant un stade adulte ailé leur permettant de disperser par voie terrestre tandis que (d) et (e) sont strictement aquatiques.

Chironomidae (Lenat, 1983) et d'autres très sensibles comme certaines espèces d'éphémères, de plécoptères et de trichoptères (EPTs, Goodnight, 1973). Cette palette de traits biologiques et de réponses au dynamisme et à la variabilité environnementale rencontrée dans les métacommunautés de têtes de bassin versant permet d'étudier la diversité des processus en jeu au sein d'une même métacommunauté à travers différents prismes de lecture lorsque l'on considère les groupes séparément.

Cependant, étudier les métacommunautés d'invertébrés présente encore certaines limites. En effet, ce groupe est très diversifié et contient des dizaines de centaines d'espèces. Il est ainsi difficile de mesurer les caractéristiques, par exemple de dispersion ou de tolérance à la pollution, pour toutes ces espèces et, en général, un ou quelques représentant(s) est (sont) décrit(s) et utilisé(s) comme proxies pour représenter un groupe en particulier (p.ex. une famille, Mondy et al., 2012 ; Tachet et al., 2000). Les variations des traits biologiques peuvent pourtant être fortes entre espèces au sein d'une famille comme chez les Chironomidae, très diversifié écologiquement, où les espèces diffèrent dans leur tolérance à la qualité de l'eau (Pinder, 1986) et aux polluants (De Bisthoven et al., 2005). Cela peut même avoir lieu au niveau du genre, comme chez le genre *Deleatidium* (Ephemeridae) où les espèces répondent différemment aux facteurs de stress environnementaux (Macher et al., 2016). A cela s'ajoutent des limites associées à l'identification morphologique (c'est-à-dire à l'identification qui se base sur des critères morphologiques), qui ne sont pas exclusives aux invertébrés. De manière générale, l'identification morphologique ne permet pas de différencier les espèces cryptiques (p.ex. Gill et al., 2016) et du cryptisme a déjà été observé chez certains groupes taxinomiques chez les invertébrés (p.ex. sous-embanchement des Crustacés, Lefebure et al., 2006 ; au sein des ordres des EPTs, Gill et al., 2014) qui pourrait être présent à différents degrés dans tous les groupes. Certains groupes sont difficilement identifiables à l'espèce et demandent soit un investissement considérable en termes de temps et d'argent, soit l'intervention de spécialistes qui sont rares, souvent experts d'un groupe particulier et ne peuvent répondre à la demande des nombreuses études portant sur les eaux douces (p.ex. Sweeney et al., 2011). L'identification morphologique des invertébrés d'eaux douces souffre aussi d'un manque de connaissances et les clés de détermination sont souvent incomplètes, voire inexistantes, notamment chez les larves d'insectes (e.g. Limnephilidae) ou dans certaines zones géographiques (p.ex. hémisphère Sud). Finalement, beaucoup de spécimens sont trop immatures, trop petits ou endommagés lors du prélèvement pour être identifiés à l'espèce sans risque d'erreur (Elbrecht et al., 2017b ; Martin et al., 2016). Ainsi, lorsque l'on étudie les métacommunautés d'invertébrés de rivières, la résolution taxinomique est rarement celle du niveau spécifique pour tous les taxa (Datry et al., 2017 ; Martin et al., 2016 ; Sweeney et al., 2011), ce qui pourrait masquer certains patrons de distribution. Martin et al. (2016) mettent en évidence qu'une résolution taxinomique à l'espèce a permis de déceler des structures spatiales qui n'étaient pas observables à des niveaux taxinomiques grossiers (genre et famille) et que le rôle des processus associés à la niche écologique augmentait avec

l'affinement de la résolution taxinomique. Ainsi, pour certains groupes tels que les oligochètes (sous-classe) ou les diptères (ordre) qui ne sont pas identifiés à un niveau plus fin que la sous-classe ou la famille et sont souvent considérés comme ubiquistes dans une métacommunauté (Serra et al., 2016), une meilleure identification pourrait faire émerger des distributions d'espèces non homogènes dans la métacommunauté avec par exemple certaines espèces structurées par des processus de niche et d'autres par la limitation de la dispersion.

6 Affiner la compréhension des métacommunautés d'invertébrés de rivières : une perspective moléculaire

En 2003, Hebert et al. proposent une méthode d'identification moléculaire basée sur l'ADN permettant de résoudre les différents problèmes liés à l'identification morphologique : le barcoding. Ils proposent d'utiliser un fragment d'ADN mitochondrial de 648 paires de bases (pb) codant pour le gène cytochrome oxydase I (COI) du génome mitochondrial comme "code-barres ADN", ou barcode, afin de distinguer les espèces animales entre elles. Un barcode ADN est une séquence ADN, c'est-à-dire un enchaînement des nucléotides Adénosine, Cytosine, Guanine et Thymine (A, C, G et T, respectivement), qui a la particularité d'avoir une variabilité forte entre espèces et faible au sein d'une espèce permettant de distinguer les espèces les unes des autres selon un seuil de divergence et ainsi de les identifier sans ambiguïtés. Le fragment du COI respecte cette condition chez les espèces étudiées par Hebert et al. (2003). Le problème du barcoding dans le cadre d'études à large échelle et/ou avec un grand nombre d'individus, ce qui est courant en écologie et surtout dans l'étude des métacommunautés, est qu'il repose sur le séquençage Sanger classique qui ne permet l'identification que d'un spécimen par réaction de séquençage résultant en une identification coûteuse et chronophage (Taberlet et al., 2012).

L'arrivée dans la dernière décennie du séquençage haut débit grâce au développement des séquenceurs de nouvelle génération (NGS en anglais) a permis de surmonter ce problème en permettant l'identification de plusieurs espèces et échantillons en un événement de séquençage (Taberlet et al., 2012). Se basant sur le principe du barcoding, le metabarcoding effectue plusieurs réactions de séquençage en simultané où une réaction implique une séquence. Ce procédé produit des millions de lectures, *reads* en anglais, et un traitement bio-informatique des lectures permet, entre autres, de regrouper les séquences en OTUs (Operational Taxonomic Units), souvent considérés comme représentant une espèce. Grâce à cette méthode, il est possible de séquencer et d'identifier l'ADN d'une communauté contenant plusieurs organismes d'espèces différentes (p.ex. un échantillon de Surber dans les cours d'eau, Yu et al., 2012) ou même à partir d'un échantillon prélevé dans l'environnement (ADN environnemental, ADNe, provenant du sol, de l'eau, des fèces, Deiner et al., 2017). Ainsi, le metabarcoding serait un outil très utile dans les études d'évaluation de la biodiversité (Shokralla et al., 2012) et particulièrement dans celle des métacommunautés où de nombreux organismes et espèces et de nombreux échantillons sont prélevés. En plus de pouvoir traiter de

nombreux échantillons dans un laps de temps court, le métabarcoding permet d'avoir, tout comme le barcoding, une résolution taxinomique au niveau spécifique, sous forme d'OTUs. Avoir accès aux assemblages d'espèces dans les métacommunautés pour tous les groupes permettrait de préciser les patrons de distribution de la biodiversité et d'améliorer notre compréhension des processus sous-jacents (Li et al., 2019 ; Pauls et al., 2014).

7 Limites actuelles de l'identification moléculaire

Le métabarcoding a été adopté comme un outil d'identification de routine dans le monde des micro-organismes (bactéries, champignons, procaryotes) où l'identification sur critères morphologiques des organismes est souvent impossible (Abdelfattah et al., 2018) mais ne l'a pas été pour les communautés de macroorganismes dont font partie les invertébrés où l'identification morphologique est traditionnellement utilisée et est ancrée dans les pratiques, que ce soit dans la recherche ou dans les domaines appliqués. Plusieurs limitations méthodologiques empêchent la démocratisation de cet outil, qui sont rencontrées à toutes les étapes du traitement des échantillons (Alberdi et al., 2017 ; Blackman et al., 2019). Nous discuterons ici seulement des limitations qui représentent spécifiquement un frein pour l'adoption du métabarcoding pour l'étude des métacommunautés d'invertébrés.

La première limitation dans le cas des invertébrés consiste au tri systématique des organismes et à leur broyage avant l'extraction. A l'instar du tri effectué en identification morphologique, ce tri vise à séparer les individus du substrat (p.ex. feuilles, sable...) avant l'extraction de l'ADN pour éviter l'inhibition de la PCR (*Polymerase Chain Reaction*, Elbrecht et al., 2017b) et pour limiter la quantité de matériel à extraire (Zizka et al., 2019). Bien qu'elle soit nécessaire, particulièrement pour les échantillons ayant un faible ratio organismes/substrat, cette étape est longue et coûteuse. L'extraction de l'ADN directement à partir de l'agent fixateur dans lequel sont placés les organismes lors de l'échantillonnage, principalement l'éthanol, a été proposée comme alternative au tri des organismes (Carew et al., 2018 ; Erdozain et al., 2019 ; Hajibabaei et al., 2012 ; Martins et al., 2019 ; Zizka et al., 2019). En plus d'un gain de temps considérable lorsque l'éthanol est utilisé comme matrice d'extraction ADN, les organismes ne sont pas broyés pour l'extraction et peuvent être conservés pour des vérifications taxinomiques ou des analyses ultérieures (Martins et al., 2019). Actuellement, peu de recherches ont été menées sur cette alternative et d'autres recherches sont nécessaires pour vérifier son potentiel en tant que matrice ADN de routine dans les études de biologie. Une autre matrice ADN proposée comme alternative à la méthode standard est l'utilisation d'ADNe qui consiste (comme mentionné précédemment) à l'extraction du milieu où vivent les organismes, typiquement l'eau dans le cas des invertébrés aquatiques (Harper et al., 2018 ; Valentini et al., 2016). Cette alternative est actuellement en plein essor (Thomsen & Willerslev, 2015) et de nombreuses recherches portent sur

son optimisation pour de multiples applications dont l'étude des communautés (Klymus et al., 2017 ; Leese et al., 2020 ; Lim et al., 2016 ; Pont et al., 2018 ; Sakata et al., 2020).

Dans la plupart des études de metabarcoding, après extraction, l'ADN cible est enrichi dans l'échantillon par amplification PCR (Figure 7.1). Cette étape est souvent considérée comme l'étape la plus restrictive du metabarcoding car elle peut biaiser la détection des espèces et l'estimation de leurs abondances initiales (Elbrecht et al., 2017b ; Leese et al., 2018 ; Pinol et al., 2019). En effet, la PCR utilise une paire d'amorces, séquence oligonucléotidique, pour l'amplification de l'ADN. Les amorces sont conçues en se basant sur des séquences de référence du gène ciblé, l'objectif étant de trouver une région (site de fixation des amorces) avec une variabilité assez faible entre ces séquences pour que les amorces puissent s'y hybrider avec la même affinité. Cependant, les sites de fixation étant variables entre espèces, les amorces peuvent avoir plus d'affinités (p.ex. posséder plus de bases en commun) avec le site de fixation d'une espèce que celui d'une autre induisant ce qu'on appelle un biais taxinomique (Hajibabaei et al., 2020 ; Pinol et al., 2015). Ce biais taxinomique entraîne une amplification différentielle entre séquences des espèces avec certaines sur-amplifiées et d'autres peu amplifiées. Il a été démontré empiriquement pour les champignons (Bellemain et al., 2010), les microorganismes (Frank et al., 2008), les invertébrés (Pinol et al., 2015) et les vertébrés (Arif et al., 2011). D'autres facteurs influent sur la PCR, comme la teneur en bases GC, les inhibiteurs qui peuvent subsister après l'extraction d'ADN, les erreurs de la polymérase lors de l'élongation induisant la création de chimères ou de faux négatifs (Aird et al., 2011) ou le nombre de copies de gènes qui peut varier selon les espèces. Par exemple, chez les eucaryotes, le nombre de génomes mitochondriaux est différent selon les cellules, espèces ou tissus (Rooney et al., 2015). Mais le biais lié aux amorces est

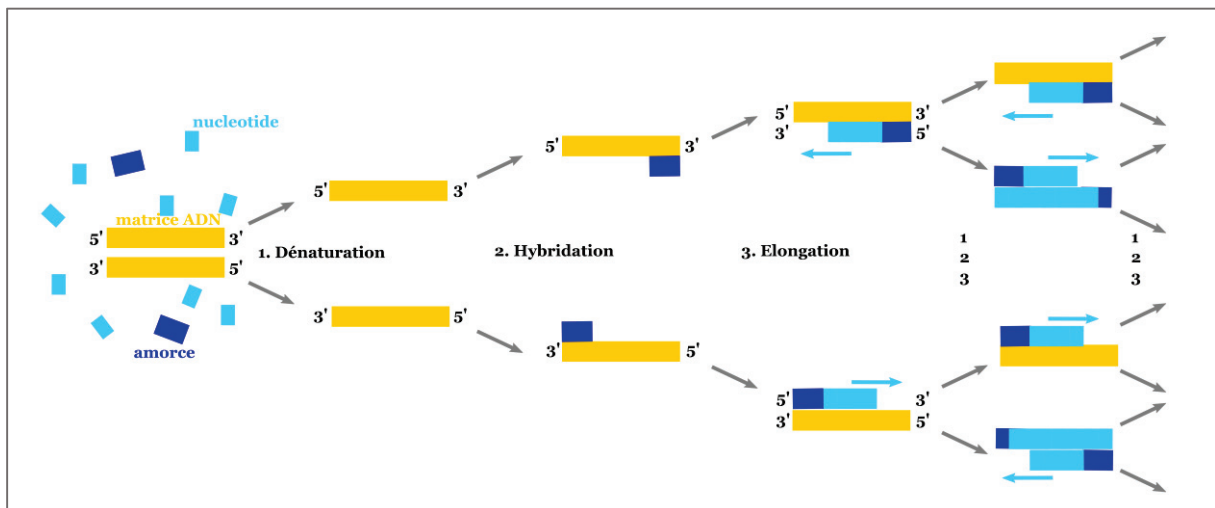


Figure 7.1 : Schéma d'une PCR. Cette méthode consiste en plusieurs réplifications successives d'une matrice ADN grâce à deux amorces (oligonucléotides) qui délimitent une séquence spécifique (p.ex. COI) à amplifier. Chaque cycle est constitué de 3 phases : la dénaturation permettant de séparer les deux brins de l'ADN, l'hybridation où les amorces vont se fixer sur les brins matrices et l'élongation où une enzyme appelée polymérase va synthétiser le brin complémentaire. De nombreux cycles ont lieu et l'échantillon sera enrichi avec des milliers de copies de la région d'intérêt.

toujours présenté comme la cause principale des erreurs en metabarcoding (p.ex. Elbrecht & Leese, 2015 ; Pinol et al., 2015 ; Pinto & Raskin, 2012). En conséquence, de nombreuses études se sont intéressées à la conception d'amorces pour améliorer la détection des espèces ou le rétablissement de l'abondance initiale. Par exemple avec l'utilisation de différentes paires d'amorces ou l'amplification de plusieurs marqueurs génétiques (Elbrecht & Leese, 2017 ; Elbrecht et al., 2016 ; Gibson et al., 2015 ; Leray et al., 2013). D'autres études ont porté sur le traitement des organismes avant leur extraction en prélevant la même quantité de tissu par spécimen (Beng et al., 2016) ou en regroupant les spécimens par groupe de taille (Elbrecht et al., 2017a). Malgré le fait que ces méthodes aient amélioré la détection des espèces, aucune d'entre elles n'a permis de résoudre le biais quantitatif (c'est-à-dire l'estimation fiable de l'abondance initiale des espèces, Pinol et al., 2018).

Pour éviter les biais induits par les amorces, il a été suggéré d'éviter l'étape de PCR et de nouvelles méthodes sans PCR ont émergé, provenant de la génomique et de la métagénomique pour être appliquées dans un contexte de metabarcoding (Porter & Hajibabaei, 2018). Une première proposition est l'utilisation du séquençage aléatoire de l'ADN (*shot gun sequencing*) sans cibler de marqueur ADN. Par exemple, Zhou et al. (2013) proposent d'enrichir les mitochondries dans l'échantillon puis de séquencer aléatoirement le génome mitochondrial. Ils trouvent des résultats fiables au niveau de la détection et de la représentativité de l'abondance mais le rendement est extrêmement faible (<1% des séquences appartiennent à de l'ADN mitochondrial). D'autres chercheurs proposent de travailler sur les mitogénomes entiers sans méthode d'enrichissement, méthode efficace pour détecter les espèces et estimer leur biomasse initiale, bien que la proportion de lectures mitochondriales récupérées soit également extrêmement faible dû à sa nature non spécifique (Bista et al., 2018 ; Gómez- Rodríguez et al., 2015 ; Tang et al., 2015 ; Zhou et al., 2013). Plus généralement, le séquençage du génome des organelles nécessite un effort de séquençage très important pour récupérer suffisamment de lectures et obtenir des résultats de bonne qualité (p.ex. pour détecter tous les taxa), ce qui le rend plus coûteux que le metabarcoding par PCR à l'heure actuelle (Bista et al., 2018). De plus, seule une petite partie du génome est utilisable pour l'assignement taxinomique étant donné que les bases de données de référence contiennent principalement des séquences de barcodes (COI pour la méiofaune et la macrofaune, Creer et al., 2016). Compte tenu des lacunes actuelles des bases de données de référence des génomes, leur séquençage représenterait un processus long et coûteux à l'image du projet "Barcode of Life" initié par Hebert en 2003 pour le barcoding qui n'a toujours pas référencé toutes les espèces connues à ce jour, même si des méthodes sont proposées pour en réduire le coût (Macher et al., 2017). Une autre alternative sans PCR est l'enrichissement par capture par sondes, où des séquences ciblées hybridées à des sondes sont récupérées par magnétisme (Horn, 2012). Contrairement au séquençage du mitogénome, la cible de l'enrichissement par capture est spécifique à un ensemble de gènes et/ou d'espèces, ce qui augmente la proportion de lectures ciblées récupérées. Les sondes sont des oligonucléotides à longs fragments (plus de 60 pb) qui sont conçus à partir de

séquences de référence. Pour une application donnée, des milliers des sondes différentes peuvent être conçues en fonction du nombre de gènes et/ou d'espèces ciblées et de la variabilité des séquences ciblées. Étant donné que l'hybridation se produit sur un long fragment et qu'une sonde est spécifique à un gène ou à une espèce, l'enrichissement de la capture devrait être plus robuste au biais taxinomique que les amorces de PCR. De plus, si des biais taxinomiques peuvent exister dans le cas où les bases de données ne soient pas exhaustives pour leur conception, les sondes pourraient capturer des espèces non référencées grâce à leur longue taille. Par exemple, une sonde conçue pour une espèce du genre *Danio* a permis de détecter d'autres espèces du genre *Danio* dans le bassin amazonien où l'espèce de référence était absente (Mariac et al., 2018). Les premières études ont montré que cette méthode était efficace pour capter l'ADN cible (Liu et al., 2016), détectait plus de taxa que la méthode PCR (Dowle et al., 2016 ; Shokralla et al., 2016) et donnait de meilleurs résultats d'abondance en diminuant le biais taxinomique si des corrections spécifiques aux espèces étaient appliquées (Liu et al., 2016).

Même si les limites associées au metabarcoding ne sont pas totalement résolues, de nombreux efforts sont consacrés à son optimisation. Par exemple, des consortiums de chercheur·es ont été mis en place, parfois à l'échelle d'un continent (Cost Action [DNAqua-Net](#)), pour échanger sur les pratiques utilisées, homogénéiser les méthodes et proposer des standards pour l'évaluation de la biodiversité et faire de l'identification moléculaire un outil d'identification standard fiable dans un futur proche.

8 Gestion et conservation des têtes de bassin versant : entre système ignoré et dynamique

En France, la protection des rivières est associée aux lois sur l'eau (Chevallier, 2007). La première loi sur l'eau date de 1964 à la suite de l'augmentation de la pollution dans les milieux naturels aquatiques, entre autres, et a permis la création des agences de l'eau qui allouaient une partie des redevances à la protection et la restauration des milieux naturels aquatiques. Dans les années qui suivent, d'une prise de conscience des problèmes environnementaux résultent des lois tels que la loi sur la protection de la Nature pour la sauvegarde des milieux naturels et la protection d'espèces menacées et la loi Pêche de 1984 qui définit les débits réservés, débits minimum à maintenir lors de la construction d'un ouvrage hydraulique. En 1992, une nouvelle loi sur l'eau voit le jour suite aux décisions et directives européennes sur l'eau qui vise la sauvegarde du patrimoine écologique et naturel que représente l'eau : protection de la qualité de l'eau, protection des écosystèmes aquatiques, protection et répartition des ressources en eau et création des Schémas Directeurs d'Aménagement et de Gestion des Eaux (SDAGE) pour s'occuper, entre autres, de l'orientation de la protection et du maintien des milieux naturels. Finalement, en 2000, un texte majeur, la Directive Cadre Européenne (DCE, EC, 2003) transposée en 2006 en France par la Loi sur l'Eau et les Milieux Aquatiques (LEMA), est proposée par l'Union Européenne. Ce texte demande aux États membres d'atteindre le « bon état écologique » de l'ensemble de leurs masses d'eau pour 2015. D'après la DCE, le bon état

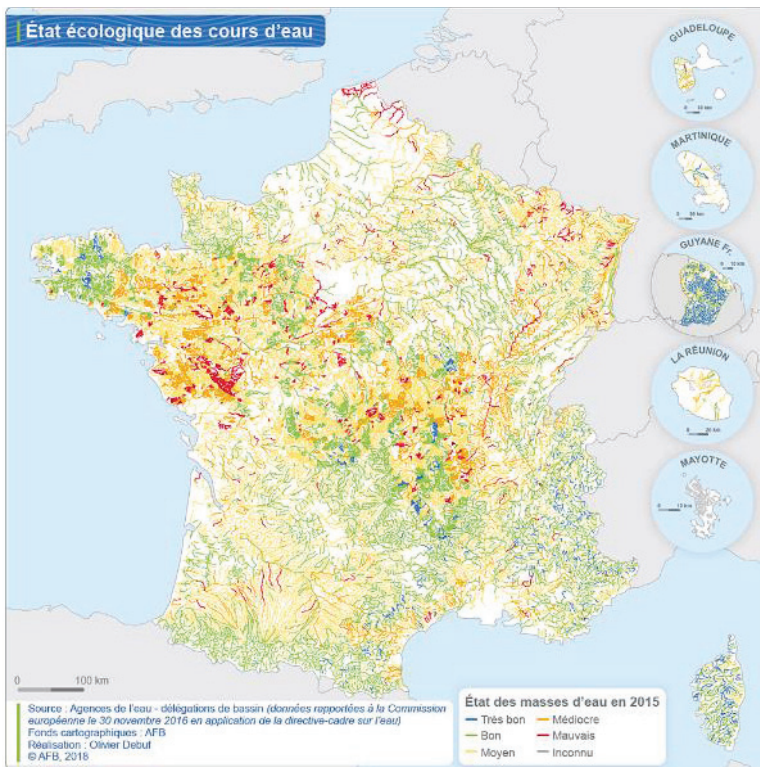


Figure 8.1 : Carte de l'état écologique des cours d'eau considérées dans le cadre de la DCE. En 2015, seulement 43,5 % des cours d'eau ont au moins un bon état écologique, loin des 100 % demandé par l'Union Européenne.

écologique correspond à l'appréciation de la structure et du fonctionnement des écosystèmes aquatiques associés aux eaux de surface et s'appuie sur des critères appelés « éléments de qualité » qui peuvent être de nature biologique (dont font partie les invertébrés qui sont largement utilisés comme éléments de qualité biologique, Bonada et al., 2007 ; Mondy et al., 2012), hydro-morphologique ou physico-chimique. Il caractérise un écart à des conditions dites de référence (conditions représentatives d'un cours d'eau pas ou très peu influencé par l'activité humaine). La situation de référence doit être représentative des conditions proches de la nature. L'atteinte du bon état écologique a été un échec par rapport aux objectifs fixés par la DCE. En effet, en 2015, l'ensemble des états

membres de l'Union Européenne sont loin des 100 % de leur masses d'eau atteignant un bon état écologique (Figure 8.1). La date dérogatoire de 2027 est remise en question en 2019 par Veronica Manfredi qui annonce à la Commission Européenne qu'« au vu du travail à fournir, 2027 reste un horizon très ambitieux » tandis que la question de réviser le texte se pose pour réévaluer les objectifs.

Dans l'ensemble de ces textes de lois, les têtes de bassin versant sont intégrées dans les hydrosystèmes sans statut de protection particuliers, ni distinction des autres milieux (Rocle, 2007). De plus, dans le cadre de la DCE, les rivières dont le bassin versant présente une superficie de moins de 10 km² ou d'une longueur < à 5 km ne sont pas définies comme des masses d'eau et peuvent être ignorées des politiques de gestion si elles ne sont pas incluses dans des masses d'eau plus grandes (p.ex arrêté ministériel du 4 mai 2017 [AGRG1632554A](#)). Pourtant, leur sensibilité aux perturbations anthropiques intenses qu'elles subissent (Rasmussen et al., 2013 ; Snell et al., 2014) couplé à leur abondance (p.ex. Lassaletta et al., 2010) et leur unicité (Clarke et al., 2008) devraient les rendre prioritaires dans les politiques de gestion et de conservation. De plus, leur position en amont les rend primordiales pour les rivières en aval qui en dépendent pour l'apport d'eau, de sédiments, de matières organiques, etc... Par exemple, de Lassaletta et al. (2010) ont montré que la pollution par les nitrates dans des têtes de bassin versant cascadaient dans les rivières en aval et que, si 76% des rivières de têtes de bassin versant composaient le réseau, seulement 22,6% (représentant 3,1% de leur linéaire

total) étaient considérées comme des masses d'eau dans le cadre de la DCE et toutes avaient des niveaux de pollution comparables (Lassaletta et al., 2010). Ainsi, la majorité des rivières de têtes de bassin versant n'était pas prise en compte et leur pollution était diffusée en aval, empêchant une gestion efficace du réseau de rivières. Si aucun effort de gestion n'est effectué en amont, les effets cascaderont jusqu'en aval, rendant inefficaces tous les efforts faits sur ces parties du réseau.

Le dynamisme fort observé dans les têtes de bassin versant présente aussi de nombreux problèmes vis-à-vis des demandes de la DCE. Tout d'abord, aucune recommandation n'est faite par rapport aux rivières intermittentes (Straka et al., 2019) qui ne sont pas reconnues comme une unité de gestion dans le cadre de la DCE. Dans le cas des indices biologiques se basant sur les invertébrés, les contrôles de surveillance sont effectués sur des tronçons de référence sans prendre en compte les hydrosystèmes dans leur ensemble. De plus, la qualité du milieu est évaluée à l'échelle locale, sans prendre en compte que les processus liés à la dispersion pourraient entrer en jeu dans la composition des communautés, selon le paradigme de *species sorting* (Stubbington et al., 2018). Dans le cas des hydrosystèmes en général mais plus particulièrement dans les hydrosystèmes dynamiques, l'évaluation devrait être faite à l'échelle de l'hydrosystème plutôt que localement sur des tronçons isolés car l'intégrité biologique y est définie globalement (Sarremejane et al., 2017b ; Straka et al., 2019). De plus, la fréquence des contrôles imposés par la DCE pour l'indice biologique invertébrés est de tous les 3 ans maximum ([Circulaire DCE](#), 2006). Le choix de la date de contrôle est donc primordial dans le cas de systèmes dynamiques et plus particulièrement dans les têtes de bassin versant (Callanan et al., 2018 ; Hastings et al., 2016), mais aucun critère objectif pour déterminer la date de contrôle n'a été déterminé par la DCE. A cela s'ajoute la comparaison à des « conditions de référence » qui sont établies à partir de rivières pérennes et ne sont pas pertinentes pour les rivières intermittentes qui présentent des compositions des communautés différentes de celles des rivières pérennes et qui de surcroît sont fluctuantes dans le temps et l'espace, complexifiant l'utilisation de telles conditions (Sánchez-Montoya et al., 2020 ; Stubbington et al., 2019).

Afin d'améliorer la conservation des têtes de bassin versant, il est nécessaire d'étudier la dynamique de la biodiversité au sein de leurs réseaux pour mettre en avant le besoin d'intégrer une vision globale plutôt que locale dans les politiques de gestion s'il s'avère que la dispersion est un processus majeur dans la structure de leurs métacommunautés. Pour ce faire, l'identification moléculaire pourrait fournir des outils fiables et rapides permettant d'évaluer la qualité de têtes de bassin versant à large échelle spatiale et temporelle, comme c'est déjà le cas pour d'autres éléments de qualité biologique (p.ex. diatomées, Valentin et al., 2019 ; Vilmi et al., 2016). De plus, bien que de nombreux protocoles de bio-évaluation, notamment en France avec l'IBGN ou son remplaçant l'I2M2 (Mondy et al., 2012), ne nécessitent qu'une identification à un niveau taxinomique élevé (famille et genre), il serait avantageux de mettre en évidence des patrons de distribution de la biodiversité au niveau spécifique, notamment dans les rivières dynamiques comme celles des têtes de bassin versant

où les espèces sont mal connues, pour conserver ces systèmes de la manière la plus adéquate (Stubbington et al., 2018).

9 Objectifs et structure de la thèse

La conservation des têtes de bassin versant représente un enjeu majeur à notre époque, tant pour maintenir l'intégrité des écosystèmes de rivières que pour préserver la biodiversité unique qu'elles renferment. Nos connaissances sur ces écosystèmes restent encore parcellaires et demandent à être approfondies pour guider au mieux les politiques de gestion et de conservation les concernant. Cette thèse tente de contribuer à cet effort en améliorant notre compréhension de la dynamique de la biodiversité aquatique dans les têtes de bassin versant. Pour cela, la thèse s'articule autour de **deux axes principaux** présentés sous la forme de deux parties dans le cadre de l'étude des métacommunautés des invertébrés benthiques de têtes de bassin versant : l'étude de la dynamique de leurs métacommunautés dans les têtes de bassin versant en contexte de fragmentation (articles 1 et 2) et l'évaluation de l'apport d'une identification moléculaire dans la compréhension des processus structurant leurs métacommunautés de tête de bassin versant (articles 3 et 4).

Les têtes de bassin versant représentent des milieux où la dispersion est contrainte par les caractéristiques physiques de leur réseau dendritique mais aussi par le caractère intermittent de certaines de leurs sections et par la présence de structures anthropiques qui peuvent limiter la dispersion des organismes en son sein. **Le premier article** cherche à mettre en évidence si des processus basés sur la dispersion sont à l'œuvre dans des têtes de bassin versant dynamiques et fragmentées. En étudiant la structure des métacommunautés des macro-invertébrés benthiques de 10 têtes de bassin versant présentant des degrés variés de fragmentation, nous avons cherché si des règles générales peuvent être établies pour ces milieux ou si les patrons observés dépendent du contexte de fragmentation. Les processus structurant les métacommunautés pouvant aussi varier au cours du temps en suivant les cycles dynamiques du régime hydrologique des têtes de bassin versant, leur description à un temps donné pourrait ne refléter que partiellement leur rôle dans l'assemblage des métacommunautés. **Le deuxième article** tente ainsi d'affiner la compréhension des métacommunautés des têtes de bassin avec un suivi spatio-temporel intense d'une tête de bassin versant sur 6 mois permettant d'étudier la variabilité temporelle des processus structurant les métacommunautés de macro-invertébrés benthiques et d'analyser plus précisément les influences respectives de la fragmentation permanente et temporaire aux différentes phases hydrologiques auxquelles elle est soumise.

L'identification morphologique ne permettant pas une identification au grain spécifique, certains patrons de distribution pourraient être cachés dans les métacommunautés, brouillant notre compréhension des processus sous-jacents à leur organisation. **Le troisième article** représente un axe

méthodologique dans lequel un outil moléculaire a été développé et évalué en laboratoire dans le but d'obtenir des informations qualitative et quantitative de bonne qualité pour les études de métacommunautés d'invertébrés. Les outils développés en laboratoire devant être testés pour une application réelle, **le quatrième article** teste cet outil dans le cadre d'une étude de métacommunauté d'invertébrés classiques sur des échantillons typiques collectés *in situ* afin de comparer la structure des métacommunautés de groupes d'invertébrés clé des têtes de bassin versant à travers le prisme de l'identification morphologique et celui de l'identification moléculaire.

10 Références

- Abdelfattah, A., Malacrino, A., Wisniewski, M., Cacciola, S.O., Schena, L., 2018. Metabarcoding: A powerful tool to investigate microbial communities and shape future plant protection strategies. *Biological Control* 120, 1-10.
- Acuña, V., Datry, T., Marshall, J., Barceló, D., Dahm, C.N., Ginebreda, A., McGregor, G., Sabater, S., Tockner, K., Palmer, M.A., 2014. Why Should We Care About Temporary Waterways? *Science* 343, 1080-1081.
- 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, R18.
- Alberdi, A., Aizpurua, O., Gilbert, M.T.P., Bohmann, K., Mahon, A., 2017. Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution* 9, 134-147.
- Allendorf, F.W., Phelps, S.R., 1981. Use of Allelic Frequencies to Describe Population Structure. *Canadian Journal of Fisheries and Aquatic Sciences* 38, 1507-1514.
- Alp, M., Keller, I., Westram, A.M., Robinson, C.T., 2012. How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology* 57, 969-981.
- Altermatt, F., 2013. Diversity in riverine metacommunities: a network perspective. *Aquatic Ecology* 47, 365-377.
- Arif, I.A., Khan, H.A., Al Sadoon, M., Shobrak, M., 2011. Limited efficiency of universal mini-barcode primers for DNA amplification from desert reptiles, birds and mammals. *Genet Mol Res* 10, 3559-3564.
- Aristote, IVème siècle av. JC. *Métaphysique A*.
- Arntzen, J.W., 1986. Aspects of introgressive hybridization in crested and marbled newts (*Amphibia: Caudata*). Preliminary results, In *Studies in Herpetology*. ed. Z. Rocek, pp. 155–158.
- Arshad, I., Umar, R., 2020. Status of urban hydrogeology research with emphasis on India. *Hydrogeology Journal* 28, 477-490.
- Baatrup-Pedersen, A., Larsen, S.E., Andersen, D.K., Jepsen, N., Nielsen, J., Rasmussen, J.J., 2018. Headwater streams in the EU Water Framework Directive: Evidence-based decision support to select streams for river basin management plans. *Sci Total Environ* 613-614, 1048-1054.
- Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., Kauserud, H., 2010. ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiol* 10, 189.
- Belmar, O., Velasco, J., Martinez-Capel, F., 2011. Hydrological classification of natural flow regimes to support environmental flow assessments in intensively regulated Mediterranean rivers, Segura River Basin (Spain). *Environ Manage* 47, 992-1004.

Belovsky, G.E., Botkin, D.B., Crowl, T.A., Cummins, K.W., Franklin, J.F., Hunter, M.L., Joern, A., Lindenmayer, D.B., MacMahon, J.A., Margules, C.R., Scott, J.M., 2004. Ten Suggestions to Strengthen the Science of Ecology. *Bioscience* 54, 345-345.

Benda, L.E.E., Poff, N.L., Miller, D., Dunne, T., Reeves, G., Pess, G., Pollock, M., 2004. The Network Dynamics Hypothesis: How Channel Networks Structure Riverine Habitats. *BioScience* 54.

Beng, K.C., Tomlinson, K.W., Shen, X.H., Surget-Groba, Y., Hughes, A.C., Corlett, R.T., Slik, J.W., 2016. The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Sci Rep* 6, 24965.

Benstead, J.P., Leigh, D.S., 2012. An expanded role for river networks. *Nature Geoscience* 5, 678-679.

Bilton, D.T., Freeland, J.R., Okamura, B., 2001. Dispersal in Freshwater Invertebrates. *Annual Review of Ecology and Systematics* 32, 159-181.

Bista, I., Carvalho, G.R., Tang, M., Walsh, K., Zhou, X., Hajibabaei, M., Shokralla, S., Seymour, M., Bradley, D., Liu, S., Christmas, M., Creer, S., 2018. Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. *Mol Ecol Resour.*

Blackman, R.C., Mächler, E., Altermatt, F., Arnold, A., Beja, P., Boets, P., Egeter, B., Elbrecht, V., Filipe, A.F., Jones, J.I., Macher, J., Majaneva, M., Martins, F.M.S., Múrria, C., Meissner, K., Pawlowski, J., Schmidt Yáñez, P.L., Zizka, V.M.A., Leese, F., Price, B., Deiner, K., 2019. Advancing the use of molecular methods for routine freshwater macroinvertebrate biomonitoring – the need for calibration experiments. *Metabarcoding and Metagenomics* 3, e34735.

Blanchet, S., Prunier, J.G., Paz- Vinas, I., Saint- Pé, K., Rey, O., Raffard, A., Mathieu- Bégné, E., Loot, G., Fourtune, L., Dubut, V., 2020. A river runs through it: The causes, consequences and management of intraspecific diversity in river networks. *Evolutionary Applications*, 1-19.

Böck, K., Polt, R., Schülting, L., 2018. Ecosystem Services in River Landscapes, In *Riverine Ecosystem Management: Science for Governing Towards a Sustainable Future*. eds S. Schmutz, J. Sendzimir, pp. 413-433. Springer International Publishing, Cham.

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

Bonada, N., Dolédec, S., Statzner, B., 2012. Spatial autocorrelation patterns of stream invertebrates: exogenous and endogenous factors. *Journal of Biogeography* 39, 56-68.

Bonada, N., Rieradevall, M., Prat, N., 2007. Macroinvertebrate community structure and biological traits related to flow permanence in a Mediterranean river network. *Hydrobiologia* 589, 91-106.

Boulton, A.J., Rolls, R.J., Jaeger, K.L., Datry, T., 2017. Chapter 2.3 - Hydrological Connectivity in Intermittent Rivers and Ephemeral Streams, In *Intermittent Rivers and Ephemeral Streams*. eds T. Datry, N. Bonada, A. Boulton, pp. 79-108. Academic Press.

Brown, B.L., Sokol, E.R., Skelton, J., Tornwall, B., 2017. Making sense of metacommunities: dispelling the mythology of a metacommunity typology. *Oecologia* 183, 643-652.

- Brown, B.L., Swan, C.M., 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. *J Anim Ecol* 79, 571-580.
- Brown, B.L., Swan, C.M., Auerbach, D.A., Campbell Grant, E.H., Hitt, N.P., Maloney, K.O., Patrick, C., 2011. Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *Journal of the North American Benthological Society* 30, 310-327.
- Brown, B.L., Wahl, C., Swan, C.M., 2018. Experimentally disentangling the influence of dispersal and habitat filtering on benthic invertebrate community structure. *Freshwater Biology* 63, 48-61.
- Cairns, J.J., Pratt, J.R., 1993. A History of Biological Monitoring Using Benthic Macroinvertebrates. In: Eds., *Freshwater Biomonitoring and Benthic Macroinvertebrates*, Chapman/Hall, New York, 10-27. , In *Freshwater Biomonitoring and Benthic Macroinvertebrates*. ed. D.M.a.R. Rosenberg, V.H., pp. 10-27, Chapman/Hall, New York,.
- Campbell Grant, E.H., Lowe, W.H., Fagan, W.F., 2007. Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecol Lett* 10, 165-175.
- Cañedo-Argüelles, M., Boersma, K.S., Bogan, M.T., Olden, J.D., Phillipsen, I., Schriever, T.A., Lytle, D.A., Ladle, R., 2015. Dispersal strength determines meta-community structure in a dendritic riverine network. *Journal of Biogeography* 42, 778-790.
- Carew, M.E., Coleman, R.A., Hoffmann, A.A., 2018. Can non-destructive DNA extraction of bulk invertebrate samples be used for metabarcoding? *PeerJ* 6, e4980.
- Carrara, F., Altermatt, F., Rodriguez-Iturbe, I., Rinaldo, A., 2012. Dendritic connectivity controls biodiversity patterns in experimental metacommunities. *Proc Natl Acad Sci U S A* 109, 5761-5766.
- Cavalier-Smith, T., 1998. A revised six-kingdom system of life. *Biological Reviews* 73, 203-266.
- Chester, E.T., Robson, B.J., 2011. Drought refuges, spatial scale and recolonisation by invertebrates in nonperennial streams. *Freshwater Biology* 56, 094-2104.
- Chevallier, H., 2007. *L'eau, un enjeu pour demain*
- Clarke, A., Mac Nally, R., Bond, N., Lake, P.S., 2008. Macroinvertebrate diversity in headwater streams: a review. *Freshwater Biology* 53, 1707-1721.
- Cottenie, K., 2005. Integrating environmental and spatial processes in ecological community dynamics. *Ecol Lett* 8, 1175-1182.
- Crabot, J., Heino, J., Launay, B., Datry, T., 2020. Drying determines the temporal dynamics of stream invertebrate structural and functional beta diversity. *Ecography* 43, 620-635.
- Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W.K., Potter, C., Bik, H.M., Freckleton, R., 2016. The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution* 7, 1008-1018.

Datry, T., Bonada, N., Boulton, A.J., 2017. Intermittent Rivers and Ephemeral Streams: Ecology and Management. Academic Press.

Datry, T., Bonada, N., Heino, J., 2016a. Towards understanding the organisation of metacommunities in highly dynamic ecological systems. *Oikos* 125, 149-159.

Datry, T., Larned, S.T., Tockner, K., 2014. Intermittent Rivers: A Challenge for Freshwater Ecology. *Bioscience* 64, 229-235.

Datry, T., Pella, H., Leigh, C., Bonada, N., Hugueny, B., 2016b. A landscape approach to advance intermittent river ecology. *Freshwater Biology* 61, 1200-1213.

De Bisthoven, L.J., Gerhardt, A., Soares, A.M.V.M., 2005. Chironomidae larvae as bioindicators of an acid mine drainage in Portugal. *Hydrobiologia* 532, 181-191.

Deiner, K., Bik, H.M., Machler, E., Seymour, M., Lacoursiere-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N., Pfrender, M.E., Bernatchez, L., 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Mol Ecol* 26, 5872-5895.

Devin, S., Beisel, J.N., Bachmann, V., Moreteau, J.C., 2001. *Dikerogammarus villosus* (Amphipoda : Gammaridae) : another invasive species newly established in the Moselle river and French hydrosystems. *Ann. Limnol. - Int. J. Lim.* 37, 21-27.

Devitt, T.J., Baird, S.J.E., Moritz, C., 2011. Asymmetric reproductive isolation between terminal forms of the salamander ring species *Ensatina eschscholtzii* revealed by fine-scale genetic analysis of a hybrid zone. *BMC Evolutionary Biology* 11, 245.

Dowle, E.J., Pochon, X., J, C.B., Shearer, K., Wood, S.A., 2016. Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: a case study using freshwater macroinvertebrates. *Mol Ecol Resour* 16, 1240-1254.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Lévêque, C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L.J., Sullivan, C.A., 2006. Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews of the Cambridge Philosophical Society* 81, 163-182.

Eby, L.A., Crowder, L.B., McClellan, C.M., Peterson, C.H., Powers, M.J., 2005. Habitat degradation from intermittent hypoxia: impacts on demersal fishes. *Marine Ecology Progress Series* 291, 249-261.

EC, 2003. Common Implementation Strategy for the Water Framework Directive. Guidance Document No. 2, Identification of Water Bodies

Brussels.

Elbrecht, V., Leese, F., 2015. Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass--Sequence Relationships with an Innovative Metabarcoding Protocol. *PLoS One* 10, e0130324.

- Elbrecht, V., Leese, F., 2017. Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental Science* 5.
- Elbrecht, V., Peinert, B., Leese, F., 2017a. Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecol Evol* 7, 6918-6926.
- Elbrecht, V., Taberlet, P., Dejean, T., Valentini, A., Usseglio-Polatera, P., Beisel, J.N., Coissac, E., Boyer, F., Leese, F., 2016. Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ* 4, e1966.
- Elbrecht, V., Vamos, E.E., Meissner, K., Aroviita, J., Leese, F., Yu, D., 2017b. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution* 8, 1265-1275.
- Élisée Reclus, J., 1869. *Histoire d'un ruisseau*, Pennti Éditions edn.
- Elton, C.S., 1927. *Animal Ecology*. Macmillan Co.
- Emerson, B.C., Cicconardi, F., Fanciulli, P.P., Shaw, P.J., 2011. Phylogeny, phylogeography, phylobetadiversity and the molecular analysis of biological communities. *Philos Trans R Soc Lond B Biol Sci* 366, 2391-2402.
- Emerson, B.C., Gillespie, R.G., 2008. Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution* 23, 619-630.
- Erdozain, M., Thompson, D.G., Porter, T.M., Kidd, K.A., Kreuzweiser, D.P., Sibley, P.K., Swystun, T., Chartrand, D., Hajibabaei, M., 2019. Metabarcoding of storage ethanol vs. conventional morphometric identification in relation to the use of stream macroinvertebrates as ecological indicators in forest management. *Ecological Indicators* 101, 173-184.
- Fagan, W.F., 2002. CONNECTIVITY, FRAGMENTATION, AND EXTINCTION RISK IN DENDRITIC METAPOPOPULATIONS. *Ecology* 83, 3243-3249.
- Fahrig, L., 2003. Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 34, 487-515.
- Finn, D.S., Poff, N.L., 2011. Examining spatial concordance of genetic and species diversity patterns to evaluate the role of dispersal limitation in structuring headwater metacommunities. *Journal of the North American Benthological Society* 30, 273-283.
- Frank, J.A., Reich, C.I., Sharma, S., Weisbaum, J.S., Wilson, B.A., Olsen, G.J., 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol* 74, 2461-2470.
- Fuller, M.R., Doyle, M.W., Strayer, D.L., 2015. Causes and consequences of habitat fragmentation in river networks. *Ann N Y Acad Sci* 1355, 31-51.
- García-Girón, J., Heino, J., García-Criado, F., Fernández-Aláez, C., Alahuhta, J., 2020. Biotic interactions hold the key to understanding metacommunity organisation. *Ecography*, 1-11.
- Gause, G.F., 1935. The Struggle for Existence. *Annals of the Entomological Society of America* 28, 59-59.

Gibson, J.F., Shokralla, S., Curry, C., Baird, D.J., Monk, W.A., King, I., Hajibabaei, M., 2015. Large-Scale Biomonitoring of Remote and Threatened Ecosystems via High-Throughput Sequencing. *PLoS One* 10, e0138432.

Gill, B.A., Harrington, R.A., Kondratieff, B.C., Zamudio, K.R., LeRoy Poff, N., Chris Funk, W., 2014. Morphological taxonomy, DNA barcoding, and species diversity in southern Rocky Mountain headwater streams. *Freshwater Science* 33, 288-301.

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

Globevnik, L., 2007. Briefing small water bodies. Report of the European Environment Agency, In No. EEA/ADS/06/001/Water.

Gómez- Rodríguez, C., Crampton- Platt, A., Timmermans, M.J.T.N., Baselga, A., Vogler, A.P., Gilbert, M., 2015. Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. *Methods in Ecology and Evolution* 6, 883-894.

Goodnight, C.J., 1973. The use of aquatic macroinvertebrates as indicators of stream pollution. *Trans Am Microsc Soc* 92, 1-13.

Göthe, E., Friberg, N., Kahlert, M., Temnerud, J., Sandin, L., 2013. Headwater biodiversity among different levels of stream habitat hierarchy. *Biodiversity and Conservation* 23, 63-80.

Gough, P., Fernández Garrido, P., Van Herk, J., 2017. Dam Removal. A viable solution for the future of our European rivers. *Dam Removal Europe*, 73.

Graham, C.H., Fine, P.V., 2008. Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecol Lett* 11, 1265-1277.

Grinnell, J., 1917. The Niche-Relationships of the California Thrasher. *The Auk* 34, 427-433.

Gronroos, M., Heino, J., Siqueira, T., Landeiro, V.L., Kotanen, J., Bini, L.M., 2013. Metacommunity structuring in stream networks: roles of dispersal mode, distance type, and regional environmental context. *Ecol Evol* 3, 4473-4487.

Hajibabaei, M., Porter, T.M., Wright, M., Rudar, J., 2020. COI metabarcoding primer choice affects richness and recovery of indicator taxa in freshwater systems.

Hajibabaei, M., Spall, J.L., Shokralla, S., van Konynenburg, S., 2012. Assessing biodiversity of a freshwater benthic macroinvertebrate community through nondestructive environmental barcoding of DNA from preservative ethanol.

Hansson, L., 1991. Dispersal and connectivity in metapopulations, In *Metapopulation Dynamics: Empirical and Theoretical Investigations*. eds M. Gilpin, I. Hanski, pp. 89-103. Academic Press.

Harper, L.R., Buxton, A.S., Rees, H.C., Bruce, K., Brys, R., Halfmaerten, D., Read, D.S., Watson, H.V., Sayer, C.D., Jones, E.P., Priestley, V., Mächler, E., Múrria, C., Garcés-Pastor, S., Medupin, C., Burgess, K., Benson, G., Boonham, N., Griffiths, R.A., Lawson Handley, L., Hänfling, B., 2018.

Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. *Hydrobiologia* 826, 25-41.

Hebert, P.D., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc Biol Sci* 270, 313-321.

Heino, J., Alahuhta, J., Ala-Hulkko, T., Antikainen, H., Bini, L.M., Bonada, N., Datry, T., Erős, T., Hjort, J., Kotavaara, O., Melo, A.S., Soininen, J., 2017. Integrating dispersal proxies in ecological and environmental research in the freshwater realm. *Environmental Reviews* 25, 334-349.

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

Heino, J., Melo, A.S., Siqueira, T., Soininen, J., Valanko, S., Bini, L.M., 2015a. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Freshwater Biology* 60, 845-869.

Heino, J., Nokela, T., Soininen, J., Tolkkinen, M., Virtanen, L., Virtanen, R., 2015b. Elements of metacommunity structure and community-environment relationships in stream organisms. *Freshwater Biology* 60, 973-988.

Horn, S., 2012. Target enrichment via DNA hybridization capture. *Methods Mol Biol* 840, 177-188.

Hubbell, S.P., 2001. *The Unified Neutral Theory of Biodiversity and Biogeography* Princeton University Press.

Huet, M., 1949. Aperçu des relations de la pente et des populations piscicoles des eaux courantes. *Hydrol. II*, 322-351.

Humphries, P., Keckeis, H., Finlayson, B., 2014. The River Wave Concept: Integrating River Ecosystem Models. *Bioscience* 64, 870-882.

Hutchinson, G.E., 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22, 415-427.

Illies, J., Botosaneanu, L., 1963. Problèmes et méthodes de la classification et de la zonation écologique des eaux courantes, considérée surtout du point de vue faunistique. *Mitt. int. Ver. Limnol.* 12, 1-57.

Joliot, F., 1947. In *Conférences de l'UNESCO*.

Jones, P.E., Consuegra, S., Börger, L., Jones, J., Garcia de Leaniz, C., 2020. Impacts of artificial barriers on the connectivity and dispersal of vascular macrophytes in rivers: A critical review. *Freshwater Biology* 65, 1165-1180.

Junk, W., Bayley, P., Sparks, R., 1989. The flood pulse concept in river-floodplain systems. *Canadian Journal of Fisheries and Aquatic Sciences Special Publication* 106, 110-127.

Klymus, K.E., Marshall, N.T., Stepien, C.A., 2017. Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PLoS One* 12, e0177643.

l'Ancien, P., 1er siècle. *Histoire Naturelle*.

Lake, P.S., 2003. Ecological effects of perturbation by drought in flowing waters. *Freshwater Biology* 48, 1161-1172.

Lassaletta, L., García-Gómez, H., Gimeno, B.S., Rovira, J.V., 2010. Headwater streams: Neglected ecosystems in the EU Water Framework Directive. Implications for nitrogen pollution control. *Environmental Science and Policy* 13, 423-433.

Lawton, J.H., 1999. Are There General Laws in Ecology? *Oikos* 84, 177-177.

Le Guyader, H., 2002. Doit-on abandonner le concept d'espèce ? *Courrier de l'environnement de l'INRA*, 51-64.

Leese, F., Bouchez, A., Abarenkov, K., Altermatt, F., Borja, Á., Bruce, K., Ekrem, T., Čiampor, F., Čiamporová-Zaťovičová, Z., Costa, F.O., Duarte, S., Elbrecht, V., Fontaneto, D., Franc, A., Geiger, M.F., Hering, D., Kahlert, M., Kalamujić Stroil, B., Kelly, M., Keskin, E., Liska, I., Mergen, P., Meissner, K., Pawlowski, J., Penev, L., Reyjol, Y., Rotter, A., Steinke, D., van der Wal, B., Vitecek, S., Zimmermann, J., Weigand, A.M., 2018. Why We Need Sustainable Networks Bridging Countries, Disciplines, Cultures and Generations for Aquatic Biomonitoring 2.0: A Perspective Derived From the DNAqua-Net COST Action, In *Next Generation Biomonitoring: Part 1*. pp. 63-99.

Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., Zizka, V.M.A., 2020. Improved freshwater macroinvertebrate detection from eDNA through minimized non-target amplification.

Lefebure, T., Douady, C.J., Gouy, M., Gibert, J., 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. *Mol Phylogenet Evol* 40, 435-447.

Leibold, M.A., Economo, E.P., Peres-Neto, P., 2010. Metacommunity phylogenetics: separating the roles of environmental filters and historical biogeography. *Ecol Lett* 13, 1290-1299.

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

Leigh, C., Datry, T., 2017. Drying as a primary hydrological determinant of biodiversity in river systems: a broad-scale analysis. *Ecography* 40, 487-499.

Lenat, D.R., 1983. Chironomid Taxa Richness: Natural Variation and Use in Pollution Assessment. *Freshwater Invertebrate Biology* 2, 192-198.

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T., Machida, R.J., 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology* 10, 34.

Lhéritier, N., 2012. Les têtes de bassin : de la cartographie aux échelles mondiale et française à la caractérisation des ruisseaux limousins, pp. 1 vol., (476 p.).

Li, F., Tonkin, J.D., Haase, P., 2018. Dispersal capacity and broad-scale landscape structure shape benthic invertebrate communities along stream networks. *Limnologia* 71, 68-74.

- Li, X., Huo, S., Zhang, J., Ma, C., Xiao, Z., Zhang, H., Xi, B., Xia, X., 2019. Metabarcoding reveals a more complex cyanobacterial community than morphological identification. *Ecological Indicators* 107.
- Lim, N.K., Tay, Y.C., Srivathsan, A., Tan, J.W., Kwik, J.T., Baloglu, B., Meier, R., Yeo, D.C., 2016. Next-generation freshwater bioassessment: eDNA metabarcoding with a conserved metazoan primer reveals species-rich and reservoir-specific communities. *R Soc Open Sci* 3, 160635.
- Liu, S., Wang, X., Xie, L., Tan, M., Li, Z., Su, X., Zhang, H., Misof, B., Kjer, K.M., Tang, M., Niehuis, O., Jiang, H., Zhou, X., 2016. Mitochondrial capture enriches mito-DNA 100 fold, enabling PCR-free mitogenomics biodiversity analysis. *Mol Ecol Resour* 16, 470-479.
- Logue, J.B., Mouquet, N., Peter, H., Hillebrand, H., Metacommunity Working, G., 2011. Empirical approaches to metacommunities: a review and comparison with theory. *Trends Ecol Evol* 26, 482-491.
- Lowe, W.H., Likens, G.E., 2005. Moving Headwater Streams to the Head of the Class. *BioScience* 55.
- Lytle, D.A., Poff, N.L., 2004. Adaptation to natural flow regimes. *Trends Ecol Evol* 19, 94-100.
- MacArthur, R.H., Levins, R., 1967. The limiting similarity, convergence and divergence of coexisting species. *American Naturalist* 101, 377-385.
- Macher, J.N., Salis, R.K., Blakemore, K.S., Tollrian, R., Matthaei, C.D., Leese, F., 2016. Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. *Ecological Indicators* 61, 159-169.
- Macher, J.N., Zizka, V.M.A., Weigand, A.M., Leese, F., Price, S., 2017. A simple centrifugation protocol for metagenomic studies increases mitochondrial DNA yield by two orders of magnitude. *Methods in Ecology and Evolution* 9, 1070-1074.
- Malavoi, J.R., Bravard, J.P., 2010. *Eléments d'hydromorphologie fluviale*. Onema.
- Marcon, E., 2015. *Mesures de la Biodiversité*. AgroParisTech, Kourou, France.
- Mariac, C., Vigouroux, Y., Duponchelle, F., Garcia-Davila, C., Nunez, J., Desmarais, E., Renno, J.F., 2018. Metabarcoding by capture using a single COI probe (MCSP) to identify and quantify fish species in ichthyoplankton swarms. *PLoS One* 13, e0202976.
- Martin, G.K., Adamowicz, S.J., Cottenie, K., 2016. Taxonomic resolution based on DNA barcoding affects environmental signal in metacommunity structure. *Freshwater Science* 35, 701-711.
- Martins, F.M.S., Galhardo, M., Filipe, A.F., Teixeira, A., Pinheiro, P., Pauperio, J., Alves, P.C., Beja, P., 2019. Have the cake and eat it: Optimizing nondestructive DNA metabarcoding of macroinvertebrate samples for freshwater biomonitoring. *Mol Ecol Resour* 19, 863-876.
- Meyer, J.L., Strayer, D.L., Wallace, J.B., Eggert, S.L., Helfman, G.S., Leonard, N.E., 2007. The Contribution of Headwater Streams to Biodiversity in River Networks1. *JAWRA Journal of the American Water Resources Association* 43, 86-103.
- Mittal, N., Bhawe, A.G., Mishra, A., Singh, R., 2015. Impact of Human Intervention and Climate Change on Natural Flow Regime. *Water Resources Management* 30, 685-699.

Mondy, C.P., Villeneuve, B., Archambault, V., Usseglio-Polatera, P., 2012. A new macroinvertebrate-based multimetric index (I 2M 2) to evaluate ecological quality of French wadeable streams fulfilling the WFD demands: A taxonomical and trait approach. *Ecological Indicators* 18, 452-467.

Nietzsche, F., 1886. Par delà le bien et le mal. Prélude d'une philosophie de l'avenir.

Paty, M., 2007. La religion cosmique d ' Einstein To cite this version : La religion cosmique d " Einstein.

Pauls, S.U., Alp, M., Bálint, M., Bernabò, P., Čiampor, F., Čiamporová-Zaťovičová, Z., Finn, D.S., Kohout, J., Leese, F., Lencioni, V., Paz-Vinas, I., Monaghan, M.T., 2014. Integrating molecular tools into freshwater ecology: developments and opportunities. *Freshwater Biology* 59, 1559-1576.

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M., Sexton, J.O., 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science* 344, 1246752.

Pinder, L.C.V., 1986. Biology of Freshwater Chironomidae. *Annual Review of Entomology* 31, 1-23.

Pinol, J., Mir, G., Gomez-Polo, P., Agusti, N., 2015. Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Mol Ecol Resour* 15, 819-830.

Pinol, J., Senar, M.A., Symondson, W.O.C., 2019. The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. *Mol Ecol* 28, 407-419.

Pinto, A.J., Raskin, L., 2012. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLoS One* 7, e43093.

Pleijel, F., Rouse, G.W., 2000. Least-inclusive taxonomic unit: a new taxonomic concept for biology. *Proceedings. Biological sciences* 267, 627-630.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegard, K.L., Richter, B.D., Sparks, R.E., Stromberg, J.C., 1997. The Natural Flow Regime. *Bioscience* 47, 769-784.

Poff, N.L., Hart, D.D., 2002. How Dams Vary and Why It Matters for the Emerging Science of Dam Removal: An ecological classification of dams is needed to characterize how the tremendous variation in the size, operational mode, age, and number of dams in a river basin influences the potential for restoring regulated rivers via dam removal. *Bioscience* 52, 659-668.

Poff, N.L., Schmidt, J.C., 2016. How dams can go with the flow. *Science* 353, 1099-1100.

Poff, N.L., Zimmerman, J.K.H., 2010. Ecological responses to altered flow regimes: A literature review to inform the science and management of environmental flows. *Freshwater Biology* 55, 194-205.

- Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., Roset, N., Schabuss, M., Zornig, H., Dejean, T., 2018. Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Sci Rep* 8, 10361.
- Porter, T.M., Hajibabaei, M., 2018. Scaling up: A guide to high-throughput genomic approaches for biodiversity analysis. *Mol Ecol* 27, 313-338.
- Rasmussen, J.J., McKnight, U.S., Loinaz, M.C., Thomsen, N.I., Olsson, M.E., Bjerg, P.L., Binning, P.J., Kronvang, B., 2013. A catchment scale evaluation of multiple stressor effects in headwater streams. *Science of the Total Environment* 442, 420-431.
- Reid, A.J., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T.J., Kidd, K.A., MacCormack, T.J., Olden, J.D., Ormerod, S.J., Smol, J.P., Taylor, W.W., Tockner, K., Vermaire, J.C., Dudgeon, D., Cooke, S.J., 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews* 94, 849-873.
- Resh, V.H., 2008. Which group is best? Attributes of different biological assemblages used in freshwater biomonitoring programs. *Environmental Monitoring and Assessment* 138, 131-138.
- Richter-Boix, A., Llorente, G.A., Montori, A., 2007. Structure and dynamics of an amphibian metacommunity in two regions. *J Anim Ecol* 76, 607-618.
- Ricklefs, R.E., 1987. Community Diversity: Relative Roles of Local and Regional Processes. *Science* 235, 167.
- Rocle, N., 2007. Point sur la réglementation relative aux zones humides et ruisseaux de tête de bassin in *Gestion des ruisseaux de tête de bassin et zones humides associées*, Saint-Brisson.
- Rooney, J.P., Ryde, I.T., Sanders, L.H., Howlett, E.H., Colton, M.D., Germ, K.E., Mayer, G.D., Greenamyre, J.T., Meyer, J.N., 2015. PCR based determination of mitochondrial DNA copy number in multiple species. *Methods Mol Biol* 1241, 23-38.
- Sakata, M.K., Yamamoto, S., Gotoh, R.O., Miya, M., Yamanaka, H., Minamoto, T., 2020. Sedimentary eDNA provides different information on timescale and fish species composition compared with aqueous eDNA. *Environmental DNA*.
- Sánchez-Montoya, M.M., Tockner, K., von Schiller, D., Miñano, J., Catarineu, C., Lencina, J.L., G. Barberá, G., Ruhi, A., 2020. Dynamics of ground-dwelling arthropod metacommunities in intermittent streams: The key role of dry riverbeds. *Biological Conservation* 241, 108328-108328.
- Sarremejane, R., Cañedo-Argüelles, M., Prat, N., Mykrä, H., Muotka, T., Bonada, N., 2017a. Do metacommunities vary through time? Intermittent rivers as model systems. *Journal of Biogeography* 44, 2752-2763.
- Sarremejane, R., Cid, N., Datry, T., Stubbington, R., Alp, M., Cañedo-Argüelles, M., Cordero-Rivera, A., Csabai, Z., Gutiérrez-Cánovas, C., Heino, J., Forcellini, M., Millán, A., Paillex, A., Pařil, P., Polářek, M., de Figueroa, J.M.T., Usseglio-Polatera, P., Zamora-Muñoz, C., Bonada, N., 2020. DISPERSE: A trait database to assess the dispersal potential of aquatic macroinvertebrates. In review.

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

Schmera, D., Árva, D., Boda, P., Bódis, E., Bolgovics, Á., Borics, G., Csercsa, A., Deák, C., Krasznai, E.Á., Lukács, B.A., Mauchart, P., Móra, A., Sály, P., Specziár, A., Süveges, K., Szivák, I., Takács, P., Tóth, M., Várbíró, G., Vojtkó, A.E., Erős, T., 2018. Does isolation influence the relative role of environmental and dispersal-related processes in stream networks? An empirical test of the network position hypothesis using multiple taxa. *Freshwater Biology* 63, 74-85.

Schoner, T.W., 1989. The Ecological Niche, In *Ecological Concepts : The Contribution of Ecology to an Understanding of the Natural World*. ed. J.M. Cherrett.

Serra, S.R.Q., Cobo, F., Graça, M.A.S., Dolédec, S., Feio, M.J., 2016. Synthesising the trait information of European Chironomidae (Insecta: Diptera): Towards a new database. *Ecological Indicators* 61, 282-292.

Sheldon, F., Thoms, M.C., 2006. Relationships between flow variability and macroinvertebrate assemblage composition: data from four Australian dryland rivers. *River Research and Applications* 22, 219-238.

Shokralla, S., F. Gibson, J., King, I., J. Baird, D., H. Janzen, D., Hallwachs, W., Hajibabaei, M., 2016. Environmental DNA Barcode Sequence Capture: Targeted, PCR-free Sequence Capture for Biodiversity Analysis from Bulk Environmental Samples.

Shokralla, S., Spall, J.L., Gibson, J.F., Hajibabaei, M., 2012. Next-generation sequencing technologies for environmental DNA research. *Mol Ecol* 21, 1794-1805.

Snell, M.A., Barker, P.A., Surridge, B.W.J., Large, A.R.G., Jonczyk, J., Benskin, C.M.W.H., Reaney, S., Perks, M.T., Owen, G.J., Cleasby, W., Deasy, C., Burke, S., Haygarth, P.M., 2014. High frequency variability of environmental drivers determining benthic community dynamics in headwater streams. *Environmental Sciences: Processes and Impacts* 16, 1629-1636.

Soria, M., Leigh, C., Datry, T., Bini, L.M., Bonada, N., 2017. Biodiversity in perennial and intermittent rivers: a meta-analysis. *Oikos* 126, 1078-1089.

Steward, A.L., von Schiller, D., Tockner, K., Marshall, J.C., Bunn, S.E., 2012. When the river runs dry: human and ecological values of dry riverbeds. *Frontiers in Ecology and the Environment* 10, 202-209.

Stoeckel, J.A., Schneider, D.W., Soeken, L.A., Blodgett, K.D., Sparks, R.E., 1997. Larval Dynamics of a Riverine Metapopulation: Implications for Zebra Mussel Recruitment, Dispersal, and Control in a Large-River System. *Journal of the North American Benthological Society* 16, 586-601.

Straka, M., Polášek, M., Syrovátka, V., Stubbington, R., Zahrádková, S., Němejcová, D., Šikulová, L., Řezníčková, P., Opatřilová, L., Datry, T., Pařil, P., 2019. Recognition of stream drying based on benthic macroinvertebrates: A new tool in Central Europe. *Ecological Indicators* 106.

- Strayer, D.L., Dudgeon, D., 2010. Freshwater biodiversity conservation: Recent progress and future challenges. *Journal of the North American Benthological Society* 29, 344-358.
- Stubbington, R., Chadd, R., Cid, N., Csabai, Z., Miliša, M., Morais, M., Munné, A., Pařil, P., Peřić, V., Tziortzis, I., Verdonshot, R.C.M., Datry, T., 2018. Biomonitoring of intermittent rivers and ephemeral streams in Europe: Current practice and priorities to enhance ecological status assessments. *Science of The Total Environment* 618, 1096-1113.
- Stubbington, R., Paillex, A., England, J., Barthès, A., Bouchez, A., Rimet, F., Sánchez-Montoya, M.M., Westwood, C.G., Datry, T., 2019. A comparison of biotic groups as dry-phase indicators of ecological quality in intermittent rivers and ephemeral streams. *Ecological Indicators* 97, 165-174.
- Sweeney, B.W., Battle, J.M., Jackson, J.K., Dapkey, T., 2011. Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *Journal of the North American Benthological Society* 30, 195-216.
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol Ecol* 21, 2045-2050.
- Tachet, H., Richoux, P., Bournaud, M., Usseglio-Polatera, P., 2000. Invertébrés d'eau douce: systématique, biologie, écologie.
- Tang, M., Hardman, C.J., Ji, Y., Meng, G., Liu, S., Tan, M., Yang, S., Moss, E.D., Wang, J., Yang, C., Bruce, C., Nevard, T., Potts, S.G., Zhou, X., Yu, D.W., 2015. High-throughput monitoring of wild bee diversity and abundance via mitogenomics. *Methods Ecol Evol* 6, 1034-1043.
- Taylor, A.T., Papeř, M., Long, J.M., 2018. Incorporating fragmentation and non-native species into distribution models to inform fluvial fish conservation. *Conservation Biology* 32, 171-182.
- Thomsen, P.F., Willerslev, E., 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation* 183, 4-18.
- Thorp, J., Delong, M., 1994. The riverine productivity model: An heuristic view of carbon sources and organic processing in large river ecosystems. *Oikos* 70, 305-308.
- Tonkin, J.D., Bogan, M.T., Bonada, N., Rios-Touma, B., Lytle, D.A., 2017. Seasonality and predictability shape temporal species diversity. *Ecology* 98, 1201-1216.
- Turunen, J., Louhi, P., Mykrä, H., Aroviita, J., Putkonen, E., Huusko, A., Muotka, T., 2018. Combined effects of local habitat, anthropogenic stress, and dispersal on stream ecosystems: a mesocosm experiment. *Ecol Appl* 28, 1606-1615.
- Unies, N., 2019. Plateforme intergouvernementale scientifique et politique sur la biodiversité et les services écosystémiques, pp. 1-15.
- Valentin, V., Frédéric, R., Isabelle, D., Olivier, M., Yorick, R., Agnès, B., 2019. Assessing pollution of aquatic environments with diatoms' DNA metabarcoding: experience and developments from France water framework directive networks. *Metabarcoding and Metagenomics* 3.
- Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P.F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G.H., Geniez,

P., Pont, D., Argillier, C., Baudoin, J.-M., Peroux, T., Crivelli, A.J., Olivier, A., Acqueberge, M., Le Brun, M., Møller, P.R., Willerslev, E., Dejean, T., 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Mol Ecol* 25, 929-942.

Vander Vorste, R., Corti, R., Sagouis, A., Datry, T., 2016. Invertebrate communities in gravel-bed, braided rivers are highly resilient to flow intermittence. *Freshwater Science* 35, 164-177.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R., Cushing, C.E., 1980. The River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences* 37, 130-137.

Vellend, M., 2010. Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology* 85, 183-206.

Verneaux, J., 1973. Cours d'eau de Franche-Comté (Massif du Jura). Recherches écologiques sur le réseau hydrographique du Doubs. Essai de biotypologie, In Thèse d'état. Besançon.

Viana, D.S., Chase, J.M., 2019. Spatial scale modulates the inference of metacommunity assembly processes. *Ecology* 100, e02576.

Vieira, T.B., da Silva, L.C.N., Casatti, L., Romero, R., Tejerina Garro, F.L., Uchôa de Aquino, P.D.P., Santos Pompeu, P.d., De Marco, P., 2020. PARTITIONING AMONG NICHE AND NEUTRAL EXPLANATIONS FOR METACOMMUNITY PATTERNS IN CERRADO STREAM FISH COMMUNITIES. *Biorix*.

Vilmi, A., Karjalainen, S.M., Hellsten, S., Heino, J., 2016. Bioassessment in a metacommunity context: Are diatom communities structured solely by species sorting? *Ecological Indicators* 62, 86-94.

von Linné, C., 1735. *Systema naturae per regna tria naturae : secundum classes, ordines, genera, species cum characteribus, differentiis, sinonimis, locis Tomus I (Regnum animale)*.

Wallace, J.B., Webster, J.R., 1996. The Role of Macroinvertebrates in Stream Ecosystem Function. *Annual Review of Entomology* 41, 115-139.

Ward, J.V., Stanford, J.A., 1995. Ecological connectivity in alluvial river ecosystems and its disruption by flow regulation. *Regulated Rivers: Research & Management* 11, 105-119.

Waters, T.F., 1972. The drift of stream insects. *Annu. Rev. Entomol.* 17, 253-272.

Wato, T., 2020. The Agricultural Water Pollution and Its Minimization Strategies – A Review. *Journal of Resources Development and Management* 64, 10-22.

Webb, C.O., Ackerly, D.D., McPeck, M.A., Donoghue, M.J., 2002. Phylogenies and Community Ecology. *Annual Review of Ecology and Systematics* 33, 475-505.

Whittaker, R.H., 1972. *Evolution and Measurement of Species Diversity* Published by : International Association for Plant Taxonomy (IAPT). *Taxon* 21, 213-251.

Whittaker, R.J., Willis, K.J., Field, R., 2001. Scale and species richness: Towards a general, hierarchical theory of species diversity. *Journal of Biogeography* 28, 453-470.

Winegardner, A.K., Jones, B.K., Ng, I.S., Siqueira, T., Cottenie, K., 2012. The terminology of metacommunity ecology. *Trends Ecol Evol* 27, 253-254.

- WWF, 2018. Living Planet Report - 2018: Aiming Higher, Gland, Switzerland.
- Yu, D.W., Ji, Y., Emerson, B.C., Wang, X., Ye, C., Yang, C., Ding, Z., 2012. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution* 3, 613-623.
- Zhou, X., Li, Y., Liu, S., Yang, Q., Su, X., Zhou, L., Tang, M., Fu, R., Li, J., Huang, Q., 2013. Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. *GigaScience* 2, 4.
- Zizka, V.M.A., Leese, F., Peinert, B., Geiger, M.F., 2019. DNA metabarcoding from sample fixative as a quick and voucher-preserving biodiversity assessment method (1). *Genome* 62, 122-136.

Chapitre 2

Métacommunautés dans les têtes de bassin versant, le rôle de la dispersion et de la fragmentation



Illustration : Ensemble de photographies présentant les outils et méthodes utilisés lors de l'échantillonnage effectué pendant cette thèse. Photos : G. Le Goff, M. Gauthier.

1 Matériel et méthodes - Chapitre 2

1.1 Choix des têtes de bassin versant : prospection et sélection

Dans le cadre de cette thèse, trois régions d'intérêt ont été identifiées (Auvergne-Rhône-Alpes, Occitanie, Nouvelle-Aquitaine) en lien avec l'Office Français de la Biodiversité (OFB) et certaines de ses Directions régionales (DiRs) : bassins d'Adour-Garonne, Rhône-Méditerranée & Corse et Loire-Bretagne. Pour l'article 1, dix têtes de bassin versant ont été sélectionnées dans les bassins d'Adour-Garonne, Rhône-Méditerranée & Corse et, pour l'article 2, une tête de bassin versant a été sélectionnée dans le bassin de Loire-Bretagne.

En 2016, 69 bassins versants potentiels ont été proposés par l'OFB, en lien avec les différentes DiRs concernées et après consultation de leurs services départementaux (Annexe A1). Plusieurs critères de présélections ont été établis afin de réduire le nombre de bassins versants à prospecter, à savoir la taille des bassins qui devait être inférieure à 200 km², l'ordre maximum des rangs de Strahler (Strahler, 1957) des cours d'eau inférieurs à 3 ainsi que la présence faible d'anthropisation hors altérations hydrologiques (physico-chimiquement et morphologiquement). Les 46 bassins versants restants après cette sélection (Annexe A1) ont été prospectés en janvier 2017 afin d'estimer *in situ* les degrés de dégradation morphologique, chimique et hydrologique et le nombre de cours d'eau potentiellement intermittents. Sur SIG, d'autres critères ont été ajoutés à ces informations : le nombre de structures anthropiques bloquant l'écoulement (Référentiel d'Obstacles à l'Écoulement, [ROE](#)), le nombre et le volume des captages d'eau (Banque nationale des prélèvements en eau, [BNPE](#)), le mode d'occupation des sols (CORINE Land Cover, [CORINE](#)) ainsi que la possibilité de placer 20 à 30 sites au sein d'un bassin versant.

Au final, 5 têtes de bassins versants ont été retenues dans le massif du Quercy (Ouest du Massif Central) et 5 dans le massif du Jura pour l'article 1 (Figure 1.1). Ces deux secteurs ont été retenus pour différentes raisons. Tout d'abord, ils présentent un contexte hydrogéologique et un contexte climatique similaires (cours d'eau issus de massifs calcaires karstifiés et situés en moyenne altitude en milieu tempéré) permettant de comparer les résultats de structure de métacommunautés des invertébrés benthiques dans des bassins bio-géographiquement différents (c'est-à-dire pour lesquels les pools régionaux d'espèces sont différents selon les écorégions sensu Illies, 1978, [écorégions sensu Illies](#)). Ensuite, il a été possible de sélectionner, pour chacune des régions, des bassins avec différents degrés d'altérations humaines en se basant sur les données ROE, BNPE, CORINE Land Cover et les observations *in situ*. Pour l'article 2, la tête de bassin versant du Thouaret dans les Deux-Sèvres a été retenue (Figure 1.1).

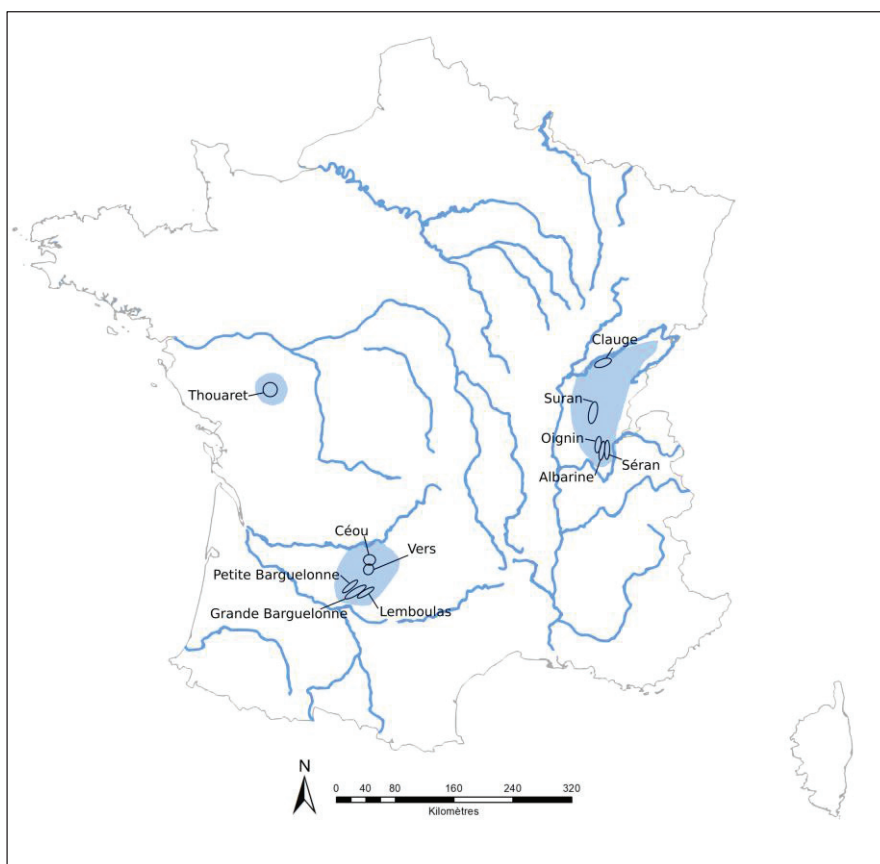


Figure 1.1 : Localisation en France des 11 têtes de bassins versants sélectionnées pour le projet de thèse.

1.2 Échantillonnage

La définition des sites d'échantillonnage a été faite dans un premier temps sur SIG mais des modifications a posteriori ont pu être apportées au regard des conditions observées *in situ* (p.ex., absence de radiers, hauteur d'eau trop élevée, accès au site impossible). Vingt à trente sites d'échantillonnage ont été placés par tête de bassin versant selon plusieurs règles : (i) les sites devaient être répartis uniformément sur tout le linéaire afin de représenter au mieux la tête de bassin versant, p.ex. en plaçant un site sur chaque tronçon hydrologique, (ii) un site ne pouvait être localisé à moins de 200 mètres d'une confluence (Altermatt, 2013) ou d'une structure anthropique comme les seuils (Fuller et al., 2015) afin de limiter les biais sur la composition des communautés liés à ces facteurs, (iii) au regard des connaissances de la tête de bassin versant, toutes les sections intermittentes devaient comporter un site (Figure 1.2). Pour l'article 2, le point (iii) a été différent du fait qu'en Poitou-Charentes, lors de la période d'étiage (généralement de juin à octobre), un suivi bimensuel des états d'écoulement des linéaires de cours d'eau est effectué par les Fédérations de pêche sur les bassins sensibles, dont fait partie le Thouaret ([Suivi Fédération de pêche](#)). Grâce aux cartes d'observations, il a été possible d'identifier les sections pérennes et intermittentes du Thouaret et les sites ont été placés afin d'avoir un site par section intermittente du bassin et un site par section pérenne en amont des

sections intermittentes (potentielle source de colonisateurs, refuges, etc...), en se basant sur les cartes de 2017 (Figure 1.2).

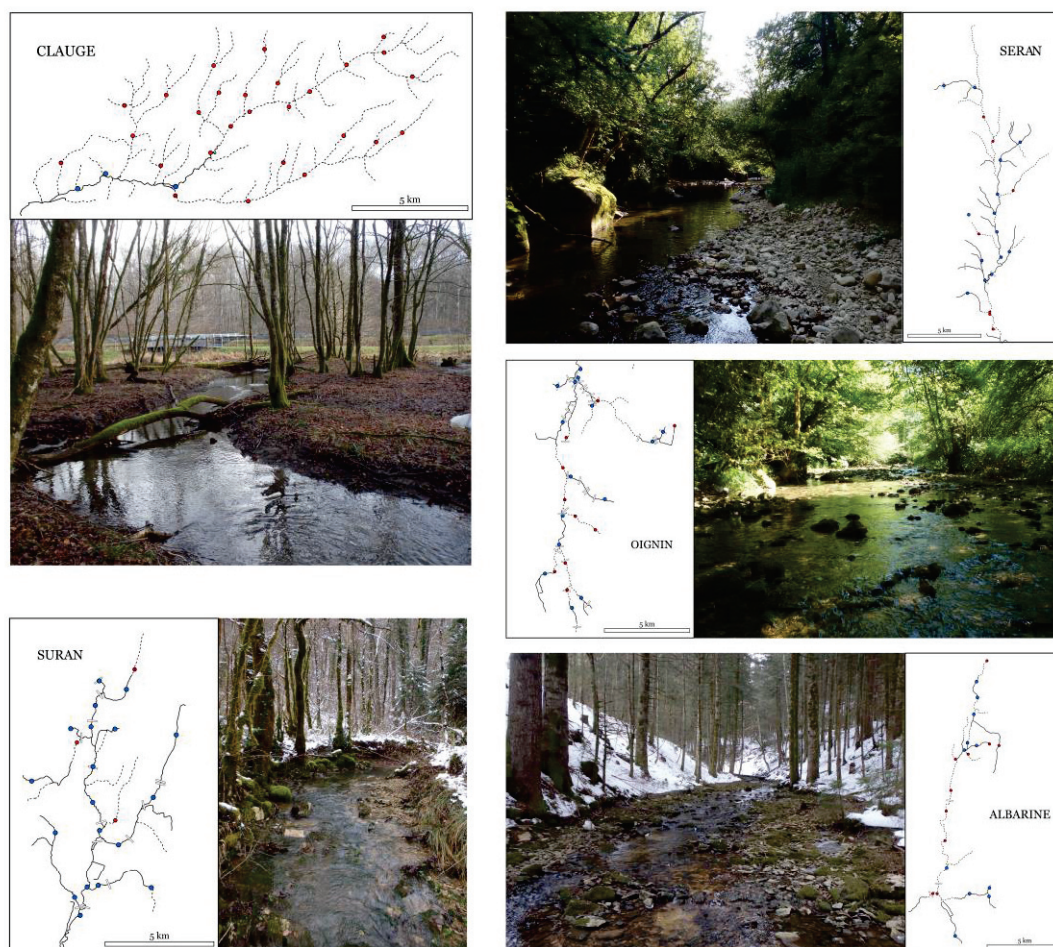


Figure 1.2 (1)

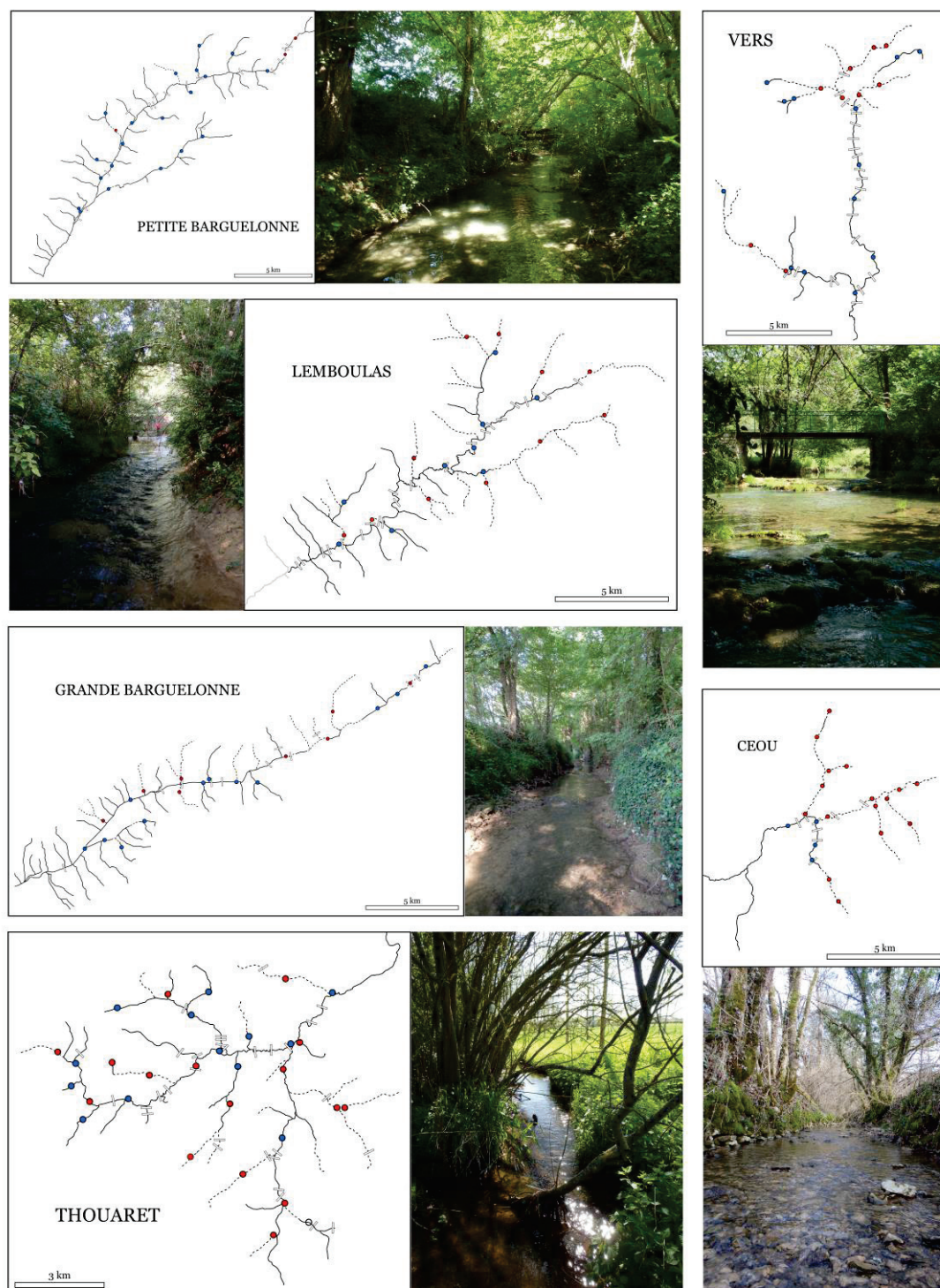


Figure 1.2 (2)

Figure 1.2 : Cartes détaillées des 11 têtes de bassin versant étudiées durant cette thèse montrant la localisation des sites d'échantillonnage, des obstacles à l'écoulement (Base de données ROE) et des sections asséchées et photos. Les cercles bleus et rouges correspondent aux sites de prélèvements pérennes et intermittents, respectivement. Les lignes pleines et en pointillé correspondent aux sections pérennes et intermittentes, respectivement. Les obstacles à l'écoulement sont représentés par des rectangles blancs sur le réseau.

Pour l'article 1, afin de caractériser au mieux la structure spatiale des communautés dans les têtes de bassin versant, deux périodes d'échantillonnage ont été prises en compte : avant étiage au printemps 2017 (juin 2017) et après étiage (après la remise en eau) à l'hiver 2017-2018 (décembre 2017-février 2018). Pour l'article 2, les campagnes d'échantillonnage ont eu lieu tous les mois de mars 2018 à décembre 2018 (sauf novembre 2018 dû à des crues trop importantes) pour un total de 9 campagnes d'échantillonnage. Cependant, seulement 6 campagnes ont été traitées dans cette thèse, englobant les 4 phases hydrologiques caractérisant les rivières intermittentes (phase « humide » où toutes les rivières du bassin sont en eau en mars, d'assèchement en avril et juin, sèche en septembre et octobre, de remise en eau en décembre). Pour les deux articles, l'échantillonnage s'est effectué selon le même protocole. Sur chaque station, deux prélèvements ont été effectués en tête de radier. Ce double échantillonnage tient au fait qu'un échantillon était destiné à l'identification morphologique, et l'autre à l'identification moléculaire (cf article 4). Pour chaque point de prélèvement, un unique prélèvement a été fait avec un filet Surber (0,1 m²) à maille de 500 µm. Ce design expérimental a permis d'avoir un effort d'échantillonnage spatial et temporel conséquent mais peut présenter une faiblesse quant à la robustesse des résultats. Nous avons testé le design expérimental en faisant des analyses sur des jeux de données externes possédant des répliques de radiers ou d'autres habitats qui ont permis de valider notre design (les détails se trouvent dans l'Appendix S2 de l'article 1). Le prélèvement a ensuite été nettoyé et tamisé puis le refus du tamis a été conditionné dans des sacs Whirl-pak® pour les échantillons destinés à l'identification morphologique et dans des flacons hermétiques pour ceux destinés à l'identification moléculaire. Ils ont enfin été fixés avec de l'alcool à 96°.

L'identification morphologique des échantillons s'est faite en laboratoire. Le contenu de chaque échantillon a été tamisé sur 2 mailles différentes, 2 mm et 500 µm, par une mise en suspension dans l'eau afin de ne pas endommager les individus. Les deux fractions ont été comptabilisées séparément. Les invertébrés ont ensuite été identifiés au niveau taxinomique le plus fin, dans les limites des connaissances taxinomiques et des possibilités techniques. Suivant les groupes, les niveaux sont donc variables (Tableau 1.1). Le tri et l'identification ont été réalisés par plusieurs opérateurs, dont deux employés permanents d'INRAE qui ont contrôlé systématiquement tous les individus (Bertrand LAUNAY et Guillaume LE GOFF). Les détails concernant l'identification moléculaire sont présentés dans le Matériel et méthodes - Chapitre 3).

A chaque site d'échantillonnage, 12 variables environnementales ont été mesurées. Au point de prélèvement des invertébrés, la hauteur d'eau (cm), la largeur mouillée (cm), la vitesse du courant (cm.s⁻¹ mesurée à 40% de la hauteur d'eau avec un anémomètre Miniwater20 (Schiltknecht, Gossau, Switzerland) et le type de substrat primaire et secondaire ont été mesurés. Le type de substrat était classé en 14 classes (surface uniforme dure, algue, sable, limon, vase, spermaphyte émergent, graviers, blocs, galets, substrat ligneux, chevelu racinaire, litière, spermaphyte immergé, bryophyte). L'altitude (m), la latitude et la longitude de chaque site ont été obtenues avec un GPS Garmin. Des mesures

physico-chimiques ont également été faites en amont des points de prélèvement pour éviter un biais dans les mesures : la conductivité ($\mu\text{S.cm}^{-1}$) et la température ont été mesurées avec un conductimètre HQ14d, le pH et le pourcentage d'oxygène dissous ont été mesurés avec une sonde multi-paramètres HQ40d (HACH, Loveland, Colorado, USA). Sur un tronçon de 20 fois la largeur mouillée, le pourcentage d'ombrage a été estimé visuellement et la pente (%) a été mesurée avec un clinomètre. Toutes les mesures ont été effectuées par la même personne, toutes dates et sites confondus pour éviter les biais liés à l'expérimentateur·trice.

Tableau 1.1. Liste des ordres d'invertébrés benthiques échantillonnés et leur niveau d'identification maximal, listé alphabétiquement par ordre au sein de groupes plus élevés (Ph, Phylum; Cl, Classe; sCl, Sous-classe). Le niveau défini ici ne concerne que les individus identifiables, c'est-à-dire ceux proches des derniers stades de développement et correctement conservés dans l'échantillon. Les individus immatures sont souvent identifiés au niveau taxinomique inférieur.

Groupe	Ordre	Niveau d'identification
Acari (sCl)	Hydrachnidia	Ordre
Bivalvia (sCl)		Espèce
Copepoda (sCl)		Ordre
Gastropoda (Cl)		Espèce
Hirudinea (sCl)		Genre
Hydrozoa (Cl)		Ordre
Insecta (Cl)	Coleoptera	Genre
	Diptera	Famille/ Genre
	Ephemeroptera	Espèce
	Hemiptera	Genre
	Hymenoptera	Espèce
	Megaloptera	Espèce
	Odonata	Espèce
	Plecoptera	Espèce
	Trichoptera	Genre/Espèce
Malacostraca (Cl)	Amphipoda	Genre
	Decapoda	Espèce
	Isopoda	Espèce
Nematoda (Ph)		Ordre
Oligochaeta (sCl)		Sous-classe
Ostracoda (Cl)		Ordre
Platyhelminthes (Ph)	Tricladida	Genre/ Espèce

1.3 Caractérisation des états d'écoulement

Afin de caractériser au mieux les états d'écoulement des sites échantillonnés, des enregistrements ont été obtenus grâce à la pose, sur quelques sites pour l'article 1 et sur tous les sites pour l'article 2, de capteurs d'état constitués d'un boîtier étanche, d'un enregistreur Hobo « Data logger state », d'un

câble électrique et d'une sonde (Onset Hobo®, Intermountain Environmental, Logan, UT). Le boîtier était placé en tête de radier, habitats dans lesquels l'eau cesse de couler en premier, le câble enterré et caché et la sonde fixée dans le lit du cours d'eau. Une mesure était enregistrée toutes les 2 heures.

Dans le cadre de l'article 1, une campagne d'observation a été effectuée en août 2017, période où les assèchements étaient supposés les plus forts, afin de caractériser les régimes hydrologiques des sites de prélèvements. Les états d'écoulement ont été attribués selon une des catégories suivantes : en écoulement lorsque le tronçon était en eau avec un écoulement de surface ; en rupture lorsque le tronçon comprenait des mouilles déconnectées sans écoulement d'eau en surface ; et sec lorsqu'il n'y avait pas d'eau visible au niveau du tronçon. Ensuite, un régime hydrologique a été assigné à chaque tronçon de chaque tête de bassin versant en couplant ces informations aux données des capteurs d'état d'écoulement : pérenne (état d'écoulement en écoulement) ou intermittent (états d'écoulement en rupture et sec).

Dans le cadre de l'article 2, la caractérisation des assèchements s'est effectuée en croisant trois sources d'informations. Premièrement, les observations visuelles faites par les Fédérations de pêche pendant la période d'étiage ([Suivi Fédération de pêche](#)). Ces observations ont été effectuées toutes les deux semaines par des membres bénévoles de l'association Carg'eau (Collectif Associatif Régional pour la Gestion de l'EAU) de juin à octobre pour le Thouaret en 2018 avec attribution des cours d'eau à l'un des états d'écoulement suivants : écoulement lorsque l'eau coule sur toute la largeur du chenal ; faible débit lorsque moins de 50% de la largeur du chenal coule ; aucun débit lorsque le chenal est composé de bassins déconnectés sans écoulement de surface ; et sec lorsque l'eau de surface est absente. Les observations ont été faites sur l'ensemble du réseau chaque fois que cela était possible, en fonction de l'accessibilité. Une description détaillée de la procédure se trouve dans Datry et al. (2016). Ensuite, l'état d'écoulement était évalué à chaque campagne d'échantillonnage qui s'effectuait dans un maximum de quatre jours après les observations des membres de l'association. Enfin, les capteurs placés sur chaque site ont permis de mesurer les durées d'assèchement durant la période d'échantillonnage, de confirmer les informations des membres de l'association et de caractériser le régime hydrologique de chaque site.

1.4 Références

Altermatt, F., 2013. Diversity in riverine metacommunities: a network perspective. *Aquatic Ecology* 47, 365-377.

Datry, T., Pella, H., Leigh, C., Bonada, N., Hugueny, B., 2016. A landscape approach to advance intermittent river ecology. *Freshwater Biology* 61, 1200-1213.

Fuller, M.R., Doyle, M.W., Strayer, D.L., 2015. Causes and consequences of habitat fragmentation in river networks. *Ann N Y Acad Sci* 1355, 31-51.

1.5 Annexes

Annexe 1. Liste des bassins versants proposés par l'OFB. Les bassins versants en gras ont été prospectés après présélection.

Bassins versants	Département	Région	Massif
Albarine	01	Auvergne-Rhône-Alpes	Jura
Séran	01	Auvergne-Rhône-Alpes	Jura
Oignin	01	Auvergne-Rhône-Alpes	Jura
Suran	01/39	Auvergne-Rhône-Alpes	Jura
Clauge	39	Bourgogne-Franche-Comté	Jura
Doüe	24	Nouvelle-Aquitaine	Massif-Central
Bournave	30	Occitanie	Massif-Central
Ribeyrette	30	Occitanie	Massif-Central
Cèze	30	Occitanie	Massif-Central
Boyne	34	Occitanie	Massif-Central
Dourbie	34	Occitanie	Massif-Central
Chassezac	48	Occitanie	Massif-Central
Nize	48	Occitanie	Massif-Central
Gupie	47	Nouvelle-Aquitaine	NA
Thoré	81	Occitanie	NA
Angle	81	Occitanie	NA
Pangaline	81	Occitanie	NA
Ro oriental	81	Occitanie	NA
Agros	81	Occitanie	NA
Ruisseau d'en Guibaud	81	Occitanie	NA
Chalaronne	01	Auvergne-Rhône-Alpes	Plaine
Moignans	01	Auvergne-Rhône-Alpes	Plaine
Glenne	01	Auvergne-Rhône-Alpes	Plaine
Salençon	01	Auvergne-Rhône-Alpes	Plaine
Nouère	16	Nouvelle-Aquitaine	Plaine
Ecly	16	Nouvelle-Aquitaine	Plaine
Maurry	16	Nouvelle-Aquitaine	Plaine
Bramerit	17	Nouvelle-Aquitaine	Plaine
Thouaret	79	Nouvelle-Aquitaine	Plaine
Sevre Niortaise	79	Nouvelle-Aquitaine	Plaine
Boutonne	79	Nouvelle-Aquitaine	Plaine
Herbasse	26	Auvergne-Rhône-Alpes	Pré-Alpes
Chalon	26	Auvergne-Rhône-Alpes	Pré-Alpes
Savasse	26	Auvergne-Rhône-Alpes	Pré-Alpes

Véore	26	Auvergne-Rhône-Alpes	Pré-Alpes
Ecoutay	26	Auvergne-Rhône-Alpes	Pré-Alpes
Gère	38	Auvergne-Rhône-Alpes	Pré-Alpes
Varèze	38	Auvergne-Rhône-Alpes	Pré-Alpes
Galaure	38	Auvergne-Rhône-Alpes	Pré-Alpes
Ourtau	64	Nouvelle-Aquitaine	Pyrénées
Espalunge	64	Nouvelle-Aquitaine	Pyrénées
Uhaitza	64	Nouvelle-Aquitaine	Pyrénées
Lez	09	Occitanie	Pyrénées
Rx de Plo	09	Occitanie	Pyrénées
Senesse	09	Occitanie	Pyrénées
Latou	09	Occitanie	Pyrénées
Rx de Monesple	09	Occitanie	Pyrénées
Estrique	09	Occitanie	Pyrénées
Rx de Roziès	09	Occitanie	Pyrénées
Rx de la Borie	09	Occitanie	Pyrénées
Rx St Nicolas	09	Occitanie	Pyrénées
Oule	09	Occitanie	Pyrénées
Clamoux	11	Occitanie	Pyrénées
Vallera	66	Occitanie	Pyrénées
Ample	66	Occitanie	Pyrénées
Maureillas	66	Occitanie	Pyrénées
Céou	46	Occitanie	Quercy
Bléou	46	Occitanie	Quercy
Vers	46	Occitanie	Quercy
Grande Barguelonne	46	Occitanie	Quercy
Petite Barguelonne	46	Occitanie	Quercy
Séoune	46	Occitanie	Quercy
Lautheronne	46	Occitanie	Quercy
Bartac	46	Occitanie	Quercy
Lemboulas	46	Occitanie	Quercy

2 Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Titre en anglais : Fragmentation promotes the role of dispersal in determining ten intermittent headwater stream metacommunities

Gauthier M.^{1,2}, Launay B.¹, Le Goff G.¹, Pella H.¹, Douady C. J.², Datry T.¹

¹ INRAE, UR-Riverly, Centre de Lyon-Villeurbanne, 5 rue de la Doua CS70077, 69626 VILLEURBANNE Cedex, France

² Univ Lyon, Université Claude Bernard Lyon 1, CNRS, ENTPE, UMR5023 LEHNA, F-69622, Villeurbanne, France

2.1 Résumé étendu

Depuis l'avènement du concept de métacommunauté, la dispersion est devenue un processus essentiel à prendre en compte dans l'étude de la structure des communautés (Heino et al., 2017 ; Leibold et al., 2004). Au sein d'une métacommunauté, la dispersion peut être élevée, suffisante ou limitée et interagit avec le processus de sélection de niche pour déterminer l'occupation des patches d'habitat par le pool régional d'espèces (Leibold et al, 2004). Dans les écosystèmes fragmentés par des facteurs naturels (p.ex. chutes d'eau, assèchements) ou anthropiques (p.ex. seuils en rivière), la dispersion peut être fortement limitée (p.ex., Alp et al, 2012). Les effets de la fragmentation sur la dispersion devraient être amplifiés dans les réseaux dendritiques où elle est déjà fortement contrainte spatialement (Campbell Grant et al, 2007 ; Tonkin et al, 2018b). Les têtes de bassin versant représentent un modèle particulièrement pertinent pour explorer l'influence de la fragmentation sur la dynamique des métacommunautés. En effet, leurs communautés sont les plus isolées au sein d'un réseau hydrographique de par leur position amont, leur réseau est fortement impacté par la construction de structures anthropiques (Lehner et al, 2011) et possède un régime hydrologique naturellement très dynamique (c'est-à-dire une amplitude du débit importante allant souvent jusqu'à l'assèchement, Benstead & Leigh, 2012). Que ce soit via la limitation à la dispersion due à leur isolement et à la fragmentation par les structures anthropiques et assèchements ou via la dynamique d'extinction et de recolonisation liée au dynamisme du régime hydrologique, les processus liés à la dispersion semblent jouer un rôle primordial dans les têtes de bassin versant.

Ainsi, la composition des communautés dans les têtes de bassin versant seraient contrôlées par différents facteurs. D'une part, les communautés locales pourraient subir des événements récurrents

d'extinction et de recolonisation à partir de zones refuges en lien avec les cycles d'assèchement et de remise en eau. Par conséquent, la diversité alpha devrait être plus faible dans les tronçons intermittents que dans les tronçons pérennes, plus stables, et la diversité beta devrait être dominée par de la *nestedness* (lorsque certaines communautés sont des sous-ensembles de communautés plus riches taxinomiquement) plutôt que par du turnover (remplacement des taxa entre communautés) avec les communautés des tronçons intermittents représentant des sous-ensembles des tronçons pérennes. D'autre part, si la limitation de la dispersion est prépondérante dans les métacommunautés de têtes de bassin versant, la composition des communautés locales devrait être liée aux distances spatiales, notamment à celles exprimant l'intensité de fragmentation, ce qui, à notre connaissance, n'a pas encore été testé dans ces milieux. Ainsi, la diversité beta devrait être structurée par les distances spatiales plutôt que par les distances environnementales. Par ailleurs, la quantification du rôle de la dispersion dépend fortement du contexte dans lequel l'étude se place, qui combine échelle spatiale et variabilité environnementale associée, modes et capacités de dispersion intrinsèques aux espèces et variations de la topologie des réseaux hydrographiques (Crabot et al., 2019a ; Heino et al., 2015 ; Tonkin et al., 2016, 2018a). La prise en compte de plusieurs réseaux plutôt que d'un seul (Cañedo-Argüelles et al., 2015 ; Datry et al., 2014 ; Tonkin et al., 2016) lorsque l'on s'intéresse aux processus structurant les métacommunautés dans des systèmes où de nombreux facteurs sont en jeu est donc souhaitable afin de prendre en compte la potentielle variabilité des réponses.

Dans cette étude, nous avons quantifié les diversités alpha et beta ainsi que ses composantes dans des métacommunautés de macro-invertébrés benthiques de 10 têtes de bassin versant fragmentées par des assèchements naturels et des structures anthropiques, réparties dans deux régions de France (cf Matériel et méthodes - Chapitre 2). Ceci afin d'évaluer et répondre à l'hypothèse principale selon laquelle le rôle de la dispersion est prépondérant dans la structure des communautés de ces systèmes dynamiques et contraints spatialement.

Pour cela, plusieurs tronçons (20 à 30 selon les bassins) uniformément répartis dans chaque tête de bassin versant ont été sélectionnés. Les invertébrés ont été échantillonnés dans un radier par tronçon à deux dates, avant et après la période d'étiage de l'année 2017 (au printemps et à l'hiver), couplé à des mesures systématiques des conditions environnementales (détails dans le Matériel et méthodes - Chapitre 2), et identifiés au niveau taxinomique le plus fin possible. La différence de diversité alpha, correspondant à la richesse taxinomique, entre tronçons intermittents et pérennes a été évaluée à l'aide de modèles linéaires à effets mixtes. La part de *nestedness* et de *turnover* de la diversité beta dans chaque tête de bassin versant et la proportion de taxa partagés entre tronçons intermittents et pérennes ont aussi été évaluées. Enfin, les relations entre diversité beta et matrices de distances environnementales et spatiales (euclidiennes, réseaux et de fragmentation) ont été quantifiées et testées avec des tests de Mantel.

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Nos résultats montrent que la diversité alpha était plus faible dans les tronçons intermittents que dans les tronçons pérennes quel que soit la tête de bassin versant ou la période d'échantillonnage. Ce résultat est en adéquation avec la littérature émergente sur le fort filtre environnemental qu'exerce les assèchements sur les communautés d'invertébrés (Aspin et al., 2018 ; Crabot et al., 2019a ; Datry et al., 2014 ; Leigh & Datry, 2017) et montre aussi que, dans le contexte des têtes de bassin versant, ce filtre peut toujours être visible même après plusieurs mois de remise en eau. A ce résultat s'ajoute la part importante de *nestedness* observée dans les têtes de bassin versant étudiées en comparaison avec la littérature (Datry et al, 2016b ; Perez-Rocha et al, 2018 ; Soinen et al, 2018) et la comparaison des compositions des communautés intermittentes et pérennes suggérant que les communautés des tronçons intermittents étaient des sous-ensembles de celles des tronçons pérennes.

L'ensemble de ces résultats suggère que les tronçons pérennes ont le rôle de sources de colonisateurs pour les tronçons intermittents dans les têtes de bassin versant et que cette recolonisation par les tronçons pérennes est un processus lent qui pourrait être associé à l'isolement des communautés, au manque de zones hyporhéiques et à la disparition de mouilles ou de conditions humides pendant l'étiage dans les têtes de bassin versant (Finn et al, 2011 ; Piano et al, 2019). Bien que nécessitant des approfondissements notamment sur la résolution taxinomique qui, couplé à notre plan expérimental, pourrait sous-estimer la part du *turnover*, la prise en compte d'autres habitats (comme les mouilles) ou l'utilisation des marqueurs génétiques pour détecter les sources de colonisateurs, ces résultats suggèrent que les communautés de têtes de bassin versant sont particulièrement sensibles aux assèchements.

Nos résultats montrent également que la diversité beta était expliquée principalement par les distances spatiales et jamais par les distances environnementales, suggérant que la limitation de la dispersion joue un rôle dans la structure des métacommunautés des macro-invertébrés benthiques des têtes de bassin versant contrairement à la sélection de niche. Cette conclusion remet en cause des points de vue théoriques répandus selon lesquels, dû au dendritisme du réseau, les communautés des têtes de bassin versant seraient structurées par la sélection de niche et celles des cours principaux par la sélection de niche et la dispersion (Brown & Swan, 2010). Notre conclusion contradictoire par rapport à cette théorie et aux études associées peut s'expliquer par plusieurs facteurs : la plupart des études n'ont pas pris en compte la fragmentation et n'ont pas étudié des systèmes très dynamiques, sous estimant le rôle de la dispersion (p.ex., Heino et al., 2012; Tonkin et al., 2016) ; l'échelle spatiale des analyses était inadéquate pour l'étude des métacommunautés et comparait des communautés qui n'était pas connectées par la dispersion (Sarremejane et al., 2017a ; Brown & Swan, 2010) et les conclusions pouvant être dépendantes du contexte, celles tirées d'une seule métacommunauté sont susceptibles de ne pas être transférables à d'autres (McCluney et al., 2014). Notre étude montre aussi une variabilité dans les relations entre diversité beta et les distances spatiales entre dates d'échantillonnage, avec notamment des relations moins fortes après la remise en eau. Ces résultats

suggèrent une variation de l'importance des processus dans la structure des métacommunautés de macro-invertébrés benthiques de têtes de bassin versant dans le temps, voire une dominance de la dérive écologique (Hubbell, 2001) dans les premières étapes de successions écologiques après remise en eau. Finalement, les modes de dispersion des taxa ont aussi été un facteur mitigeant l'effet de la fragmentation avec les taxa ayant un stade de vie aérien montrant un lien plus fréquent et fort avec les distances spatiales que les espèces strictement aquatiques qui seraient structurées par la dérive écologique. Les explications seraient un rôle potentiel de barrières physiques (Chester et al, 2015) ou d'habitat (Anna et al, 2008 ; Alp et al, 2013) par les tronçons asséchés ou en rupture d'écoulement sur les espèces à dispersion aérienne faible ou modérée ou une limitation de la dispersion terrestre par la forêt dense dans les têtes de bassin versant (Petersen et al, 2004).

L'ensemble de ces résultats concordants entre les dix têtes de bassin versant sur la structure des communautés à l'échelle régionale indique que (i) la limitation de la dispersion joue un rôle prépondérant dans les têtes de bassin versant fragmentés et dynamiques, (ii) la part explicative des processus n'est pas la même suivant la phase hydrologique à laquelle on se place et (iii) varie suivant les modes et capacités de dispersion des espèces. Pour approfondir ces différents points abordés dans cet article, il est nécessaire (i) d'améliorer notre compréhension de la réponse des communautés à la fragmentation, notamment en caractérisant et en étudiant cette réponse face aux deux types de fragmentation (par les assèchements et par les structures anthropiques) séparément, (ii) d'étudier la structure des métacommunautés à une échelle temporelle plus fine (p.ex. mensuelle) pour mieux comprendre la dynamique des processus structurant les métacommunautés dans les systèmes dynamiques et (iii) d'affiner la description des modes et capacités de dispersion des espèces et les prendre en compte dans les analyses afin de décrire plus finement et précisément comment se structurent les métacommunautés.

Au final, l'ensemble des résultats de cette étude confirme le rôle prépondérant de la dispersion dans la structure des métacommunautés de systèmes contraints spatialement, fragmentés et dynamiques, à travers la dynamique d'extinction et de recolonisation des tronçons intermittents par des zones refuges ou par la limitation de la dispersion. Les réseaux de rivières en général, mais plus particulièrement les têtes de bassin versant, vont devenir dans un futur proche de plus en plus fragmentés dû aux changements globaux induisant des assèchements plus fréquents et le développement de structures anthropiques dans les rivières (Anderson et al., 2019 ; Lehner et al, 2011). Il est fort probable que ces systèmes stables structurés par la sélection de niche deviennent graduellement des systèmes limités par la dispersion. Ce changement de régime hydrologique de pérenne à intermittent pourrait représenter un point de bascule écologique critique (Smol & Douglas, 2007) où la composition des communautés serait modifiée de manière irréversible. Bien que ne pas franchir ce point de bascule serait la stratégie optimale pour préserver ces écosystèmes, il est nécessaire d'anticiper un tel changement en intensifiant les recherches tant fondamentales

qu'appliquées sur les systèmes déjà fragmentés et dynamiques comme ceux de cette étude pour préserver au mieux la biodiversité des rivières, leur intégrité fonctionnelle et les services écosystémiques qu'elles fournissent à la société (Datry et al., 2018 ; Tonkin et al., 2019).

2.2 Abstract

1. Dispersal, defined as the movement of individuals among local communities in a landscape, is a central regional determinant of metacommunity dynamics in ecosystems. Whereas both natural and anthropogenic ecosystem fragmentations can limit dispersal, previous attempts to measure such limitations have faced considerable context-dependency, due to a combination of spatial extent and associated environmental variability, the wide range of dispersal modes and abilities of organisms or variation in network topologies. Therefore, the role dispersal plays compared to local environmental filtering in explaining metacommunity dynamics remains unclear in fragmented dendritic ecosystems.

2. We quantified alpha and beta-diversity components of invertebrate metacommunities across ten fragmented headwater stream networks and tested the hypothesis that dispersal is the primary determinant of biodiversity organisation in these dynamic and spatially-constrained ecosystems.

3. Alpha-diversity was much lower in intermittent than perennial reaches, even long after rewetting, indicating an overwhelming effect of drying including a legacy effect on local communities.

4. Beta-diversity was never correlated with environmental distances but predominantly explained by spatial distances accounting for river network fragmentation. The nestedness proportion of beta-diversity was considerable and reflected compositional differences where communities from intermittent reaches were subsets of perennial ones.

5. Altogether, these results indicate dispersal as the primary process shaping metacommunity dynamics in these ten headwater stream networks, where local communities recurrently undergo extinction and recolonization events. This challenges previous conceptual views that local environment filtering is the main driver of headwater stream metacommunities.

6. As river networks become increasingly fragmented due to global change, our results suggest that some freshwater ecosystems currently driven by local environment filtering could gradually become dispersal-limited. In this perspective, shifts from perennial to intermittent flow regimes represent ecological thresholds that should not be crossed to avoid jeopardising river biodiversity, functional integrity and the ecosystem services they provide to society.

2.3 Introduction

In the recent years, community ecology has acknowledged that dispersal, defined as the movement of individuals between local communities scattered in a landscape, is a pivotal regional process

determining metacommunity dynamics in ecosystems (Brown & Swan, 2010; Heino et al., 2017; Leibold et al., 2004). In interaction with local environmental selection, dispersal can promote the coexistence of species on patchy habitat landscapes. For example, when the dispersal rate among communities is very high, it can contribute to species coexistence by favouring the colonisation of habitat patches after disturbances even when organisms are locally maladapted to their environment (i.e. mass-effect, Leibold et al., 2004). In contrast, when dispersal is limited, organisms are precluded to colonize their favourable habitats in the landscape and such dispersal limitation can cascade into biodiversity patterns with imprints onto alpha, beta and gamma-diversity patterns in space (e.g. Langenheder & Lindstrom, 2019) and time (e.g. Datry, Bonada, & Heino, 2016a; Jabot et al, 2020).

Dispersal can be dramatically limited in naturally and anthropogenically fragmented ecosystems, which are increasing due to global change and intensifying land-use changes. Natural landscape barriers, such as high relief in valleys, water falls or natural drying events in river networks can be permanent or temporary obstacles to organism movements within the landscape and thus influence the dynamics of metacommunities (e.g. Alp, Keller, Westram, & Robinson, 2012; Cañedo-Argüelles et al., 2015; Keleher & Rader, 2008). Dispersal limitation can also originate from the intrinsic dispersal abilities and modes of organisms (Sarremejane, Stubbington, Dunbar, Westwood, & England, 2019; Siqueira et al., 2019; Tonkin et al., 2018a). For example, two species with contrasted dispersal abilities can exhibit very different patterns within the same ecosystem, reflecting their relative ability to reach all localities within a given landscape (Rader, Unmack, Christensen, & Jiang, 2019). Similarly, the presence of weirs and culverts in river networks may not affect metacommunities of species possessing aerial stages such as ephemeropteran insects as strongly as it does for strictly aquatic species like the amphipod *Gammarus* spp. (Alp et al., 2012). The theoretical importance of dispersal in determining metacommunity dynamics is well recognized (e.g. Mouquet & Loreau, 2003; Schmera et al., 2018) and some recent empirical tests in perennial river networks confirm it (e.g. Downes & Lancaster, 2018; Downes, Lancaster, Glaister, Bovill, & Strecker, 2017). However, research is still scarce in non-perennial rivers fragmented by drying, possibly due to the inherent difficulties in assessing dispersal abilities for many organisms (Cañedo-Argüelles et al., 2015; Heino et al., 2017; Sarremejane et al., 2019) and to the limitations of existing statistical tools (Crabot, Clappe, Dray, Datry, & McCrea, 2019a; Jabot et al, 2020).

The effects of fragmentation, that is the disruption of physical connection within an ecosystem (Fuller, Doyle, & Strayer, 2015), on metacommunity dynamics should be exacerbated in dendritic ecosystems, such as hedge and river networks, where dispersal is already spatially constrained (Brown & Swan, 2010; Campbell Grant, Lowe, & Fagan, 2007; Finn & Poff, 2011; Tonkin, Heino, & Altermatt, 2018b). Accordingly, the headwaters of river networks are particularly relevant to explore how fragmentation can influence metacommunity dynamics (Figure 2.1a). This is because they are the most isolated habitat patches of river networks due to their upstream spatial position at the interface

between terrestrial and aquatic domains (Brown & Swan, 2010; Finn, Bogan, & Lytle, 2009; Tonkin, Heino, Sundermann, Haase, & Jähnig, 2016) and because most headwaters are naturally highly-dynamic due to recurrent dry-wet cycles (Benstead & Leigh, 2012; Crabot et al., 2019a; Datry et al., 2014; Lowe & Likens, 2005). Although the local effects of such cycles are well known today (e.g. Crabot, Heino, Launay, & Datry, 2019b; Datry et al., 2014; Leigh & Datry, 2017; Resh et al., 2012), their effects at larger scales caused by fragmentation of river networks is still poorly explored. Therefore, local communities of headwater streams could undergo recurrent extinction and recolonization events, making dispersal from nearby refuges the primary control of headwater stream metacommunity (Figure 2.1a). This should leave imprints on the different aspects of biodiversity, including alpha and beta-diversity and their components (Figure 2.1b-d). Moreover, if dispersal is the primary process shaping metacommunities in headwater streams, local communities should differ according to spatial distances that account for the fragmentation, which, to our knowledge, has not yet been tested. In addition, previous efforts to quantify the role of dispersal have faced a particular high context-dependency due to a combination of spatial extent and associated environmental variability, the wide range of dispersal modes and abilities of stream organisms as well as variations in network topology (e.g. Crabot et al., 2019a; Heino, Grönroos, Soininen, Virtanen, & Muotka, 2012; Tonkin et al., 2016). This clearly calls for considering multiple river networks when aiming to disentangle the factors shaping headwater metacommunities rather than focusing on a single one, as is usually done (e.g. Cañedo-Argüelles et al., 2015; Datry et al., 2014; Tonkin et al., 2016).

In this study, we investigated alpha- and beta-diversity components of benthic macroinvertebrate metacommunities across 10 headwater stream networks fragmented by natural drying events and small anthropogenic structures (e.g. weirs). We hypothesized dispersal to be the primary determinant of biodiversity organisation in these dynamic and spatially-constrained ecosystems. Accordingly, we predicted that (i) alpha-diversity would be lower in intermittent reaches than in perennial ones due to the legacy effect of drying on ecological successions (Figure 2.1b), (ii) beta-diversity would be primarily determined by physical distances, and particularly by those accounting for the network fragmentation, rather than by environmental distances (Figure 2.1c) and (iii) beta-diversity would be dominated by nestedness, local communities at intermittent reaches being subsets of those from perennial reaches (Figure 2.1d).

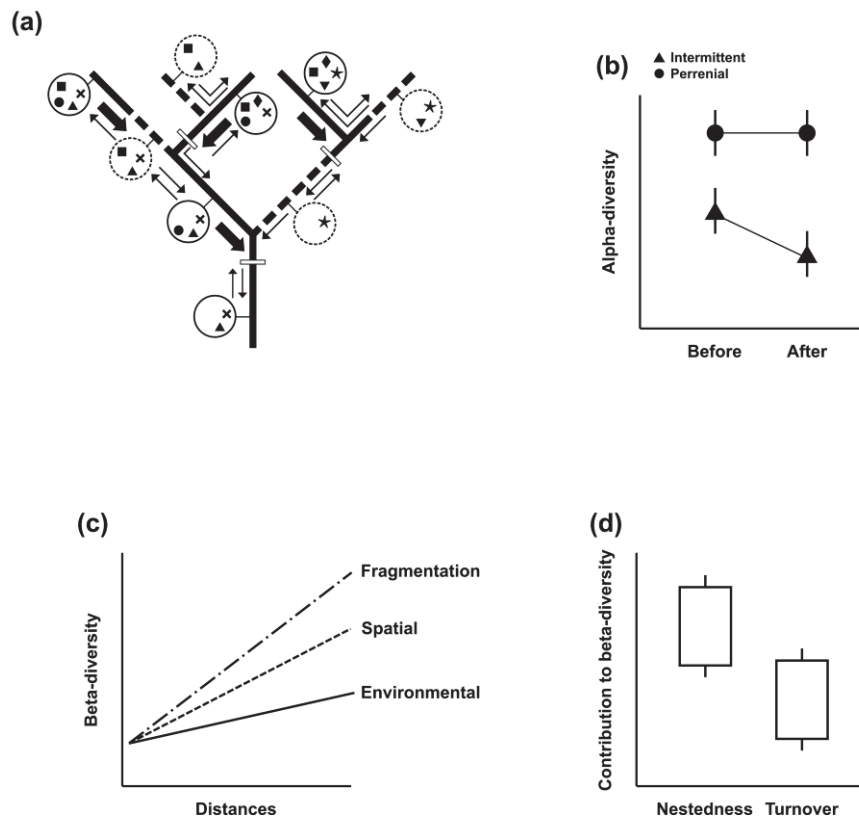


Figure 2.1 : (a) Idealized dispersal dynamics in intermittent headwater stream networks. Dispersal rates are represented by the thickness of the arrows. Communities of perennial reaches (filled line circles) represent sources of colonizers for communities of intermittent reaches (dashed line circles) with a stronger rate of dispersal downstream due to the direction of water flow from up- to downstream. Drying reaches (dashed lines) and anthropogenic structures (white rectangles) represent barriers to dispersal, at least temporarily for drying. Shapes in circles represent different species in each local community. Consequently, and if dispersal is the primary factor shaping metacommunities of headwater streams, one should expect (b) a lower alpha-diversity in intermittent compared to perennial reaches and (c) a beta-diversity structured mostly by spatial rather than environmental variables and notably those including network fragmentation, and (d) beta-diversity dominated by a nestedness component reflecting the colonisation from nearby refuges.

2.4 Methods

2.4.1 Study sites

Ten headwater stream networks fragmented by natural drying events (hereafter, FHWs) were selected in two biogeographically-contrasted areas (i.e. contrasted regional species pool according to ecoregions sensu Illies, 1978) of France presenting similar geological and climatic contexts (Tableau 2.1 Tableau 2.1. Main characteristics of the 10 studied headwater stream networks.). FHWs were selected according to the following rules: (i) a maximum catchment size of 200 km² to account for spatial dependence (Heino et al., 2012), (ii) a maximum Strahler order of three for the most downstream reach, (iii) the presence of at least four intermittent reaches and (iv) pristine physico-

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

chemical conditions . In addition, to account for context-dependency, FHWs were selected to cover different levels of fragmentation generated by drying and anthropogenic structures (Tableau 2.1). In each FHW, 20 to 30 study reaches (mean \pm SD.: 22 ± 3 , Tableau 2.1) were selected to be uniformly spread across the FHW yet as close as logistically practical to confluences and anthropogenic structures to control for the impact of these network features on community composition (Altermatt, Schreiber, & Holyoak, 2011; Fuller et al., 2015) (Figure 2.2). Detailed maps of each FHW are available in Supporting Information (Appendix S1).

Headwater stream network name	Headwater stream network code	Drainage area (km ²)	Altitudinal range (m)	Number of sampled reaches	Mountain range	Climate type	Anthropogenic structure number	Total linear length (m)	Total drying linear length (m)	Percentage of drying linear (%)
Clauge	CLAU	112	212-266	30	Jura	Temperate	0	57,574	121,397	46.8
Suran	SURA	132	326-437	22	Jura	Temperate	15	73,750	15,681	18.1
Oignin	OIGN	93	521-874	25	Jura	Temperate	45	76,160	21,879	16.3
Albarine	ALBA	99	747-985	21	Jura	Temperate	3	86,473	24,900	32.7
Séran	SERA	156	341-941	21	Jura	Temperate	6	114,709	33,027	28.8
Céou	CEOU	134	227-288	20	West of Massif central	Temperate	9	123,520	20,526	35.7
Vers	VERS	91	177-337	22	West of Massif central	Temperate	21	134,118	21,810	29.6
Petite Barguelonne	PBAR	120	111-232	23	West of Massif central	Temperate	19	148,764	8,719	7.1
Grande Barguelonne	GBAR	110	116-215	20	West of Massif central	Temperate	10	194,696	28,253	14.5
Lemboulas	LEMB	126	131-238	22	West of Massif central	Temperate	18	259,215	44,987	30.2

Tableau 2.1. Main characteristics of the 10 studied headwater stream networks.

2.4.2 Sampling design

All reaches were sampled before and after the summer period during which drying events occur. Pre-drying samples were collected in May-June 2017 and post-drying samples were collected between December 2017 and February 2018, following at least four weeks of continuous flow after rewetting had been observed. At each sampling reach and date, 12 environmental variables were measured. At the sampling point where invertebrates were sampled (see below), water depth (cm), wetted width

(cm), flow velocity [$\text{cm}\cdot\text{s}^{-1}$ measured at 40% of water depth with a Miniwater20 velocity meter (Schiltknecht, Gossau, Switzerland) and main and secondary substratum types were measured. Substratum type was classified into 14 classes (mud, silt, sand, gravel, rocks, slab, litter, root tufts, wood, algae, helophyte, spermaphyte and bryophyte). Elevation (m), latitude and longitude of each study site were obtained with a Garmin GPS. Conductivity ($\mu\text{S}\cdot\text{cm}^{-1}$) and temperature were measured with an HQ14d conductivity meter and pH and dissolved oxygen (%) were measured with an HQ40d multi parameter (HACH, Loveland, Colorado, USA). On a reach length defined as 20-fold the wetted width measured at the sampling point (reach width: 233 ± 180 cm ; reach length: 46 ± 36 m before the summer period ; reach width: 375 ± 484 cm ; reach length: 75 ± 96 m after), percent canopy cover was assessed visually by sight and slope (%) were measured with a clinometer. These variables are known to directly influence the composition of riffle invertebrate communities (Cañedo-Argüelles et al., 2015; Karaouzas, Theodoropoulos, Vourka, Gritzalis, & Skoulikidis, 2019).

Benthic invertebrates were sampled using a Surber sampler (sampled surface area 0.1 m^2 , mesh size $500 \mu\text{m}$). At each reach, one Surber sample was taken from a riffle habitat to promote between FHW comparisons rather than to inventory local biodiversity by sampling all habitat types; such approach has been widely used to generate flow-ecology relationships in intermittent rivers and is particularly suitable for cross-river network analyses (e.g. Arscott, Larned, Scarsbrook, & Lambert, 2010; Crabot et al., 2019a; Datry et al., 2014; Vander Vorste, Malard, & Datry, 2016). To test for the reliability of this sampling design to quantify metacommunity processes, preliminary analyses on two different datasets were carried out (Supporting Information, Appendix S2). Briefly, increasing the number of within-reach samples did not alter the inferences on metacommunity processes. Samples were preserved in 96% ethanol and organisms of each sample were counted and identified to the lowest practical taxonomic level using available identification keys (list of references for identification available on Zenodo, see Data Repository). Most insects, crustaceans, Tricladida and all molluscs were identified at the species level whenever possible (i.e. if organisms were mature enough). At a minimum, organisms of the insect order Diptera were identified at the genus level, those of insect orders Coleoptera, Hemiptera and a part of Trichoptera and Amphipoda were identified at the genus level and those of Ostracoda and Copepoda were identified at the order level. Achaeta and Tricladida were identified at the genus level. Nematoda, mites, Hydrozoa and Oligochaeta were identified at the order or sub-order level. A detailed table with identification taxonomic level for each group is presented in Supporting Information (Appendix S3). Lastly, each taxon was assigned to two dispersal categories according to Tachet et al. (2000): invertebrates that are strictly aquatic and can only disperse along the river network (90 taxa) and invertebrates with an aerial life stage (327 taxa) which can disperse by flying or crawling between reaches.

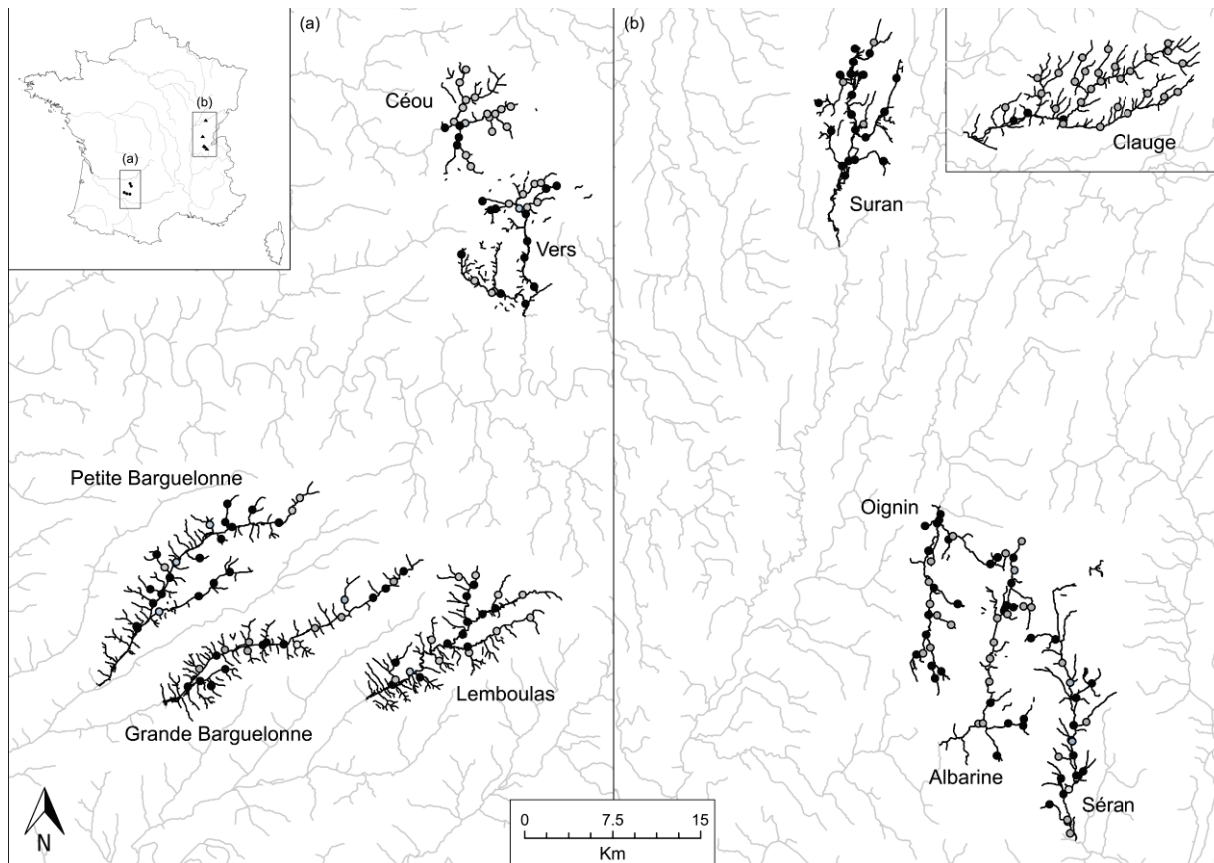


Figure 2.2 : The two study areas in France: (a) West of Massif Central and (b) Jura. The studied FHWs are represented by black lines. Grey and black circles correspond to the intermittent and perennial reaches sampled in this study, respectively. Detailed maps for each FHW are provided as Supporting Information (Appendix S1).

2.4.3 Data analysis

(a) Alpha-diversity in FHWs

To test our first hypothesis that alpha-diversity would be lower in intermittent compared to perennial reaches, taxonomic richness was estimated for each reach and was tested for differences between flow regimes (intermittent versus perennial) and sampling period (before and after the drying season). Four nested linear mixed-effect models with increasing levels of complexity were fitted. The fixed effects of the models were (i) none for the null model; (ii) flow regime for the second model; (iii) flow regime and sampling period for the third model without interaction between the two factors and (iv) flow regime and sampling period for the fourth model with interaction between the two factors. For each model, a random intercept was fitted by FHW to control for between-FHW variability and account for possible context-dependency. The statistical significance of the fixed effects was assessed using likelihood-ratio tests on the models comparing each model to the one with the closest upper complexity level (Bolker et al., 2009). Normality and homoscedasticity of residuals were visually

assessed for each model were verified by eye with plots of residuals against fitted values. One FHW (Clauge, Figure 2.2b) was omitted from model fitting as most reaches were already dry during the sampling in May-June.

(b) Beta-diversity patterns in FHWs

To test our second hypothesis that beta-diversity would be primarily determined by physical distances, and particularly when including fragmentation, rather than by environmental distances, we quantified and tested the relationships between beta-diversity and distance matrices.

Distance matrices. For each FHW, one environmental and three spatial distance matrices were computed. The environmental distance matrix (ENV) was calculated using Euclidian distances on the 12 environmental variables measured during sampling and the flowing regime of each reach (i.e. perennial or intermittent) and after variable normalization (mean = 0, S.D. = 1, Cañedo-Argüelles et al., 2015). Two spatial distance matrices related strictly to geographic distances among reaches in the network while the third one also integrating river network fragmentation by drying and anthropogenic structures. An overland matrix (EUC) was constructed from Euclidian distance between each pair of sampled reaches and was calculated from site GPS coordinates with `vegdist` function of `vegan` (Oksanen et al., 2019). A network matrix (NET) was constructed from the distances along the river channel network between all pairs of reaches which were calculated on the RHT river network (Pella et al, 2014) with QGIS modules (`postres`, `postgis` and `pgrouting`; QGIS v3.10). Both distances are commonly used in stream metacommunity research (e.g. Cañedo-Argüelles et al., 2015; Datry et al., 2016b; Sarremejane et al., 2017a). Last, the third distance matrix (FRAG) integrating the fragmentation between each pair of sampled reaches was generated. Fragmentation by drying was determined in the field by visually assessing the flow state of all reaches in each FHW during August 2017 (the driest period of the year in the studied areas) based on the Carthage network (<http://www.sandre.eaufrance.fr/atlas/srv/fre/catalog.search;jsessionid=j7hgqzjk61161ru6loglyweeg?mode=ModeMeta&uuid=3409c9c3-9836-43be-bac3-b110c82b3a25#/meta1>). Each reach was assigned one of the following flow states: flowing when water was flowing in the reach; no flow when the reach was composed of disconnected pools without surface flow; and dry when there was no surface water in the reach. This also allowed classifying each reach into the following category: perennial (when flowing) or intermittent (when non-flowing or dry). The presence of small anthropogenic structures was assessed across each FHW with the ROE database, an inventory of anthropogenic barriers at the national scale (<http://www.sandre.eaufrance.fr/atlas/srv/fre/catalog.search#/metadata/59057026-b40c-4cf9-9e3e-7296e0aa1a78>). These structures were principally weirs (64%) and culverts (14%) but were also represented by bridge inverters (8%), fords (2%) and small dams (6%). For each pair of sampled reaches, the fragmentation distance matrix was then constructed by summing up the number of dry patches and anthropogenic structures between sampled reaches using QGIS v3.10.

Beta-diversity measures. Matrices of beta-diversity between pairwise combinations of sampled reaches were constructed using three indices: (i) one based on presence-absence data (Jaccard dissimilarity index) (ii) two based on abundance data with a $\log(x + 1)$ transformation to better represent rare species (Bray-Curtis and Chao dissimilarity indexes) (Cañedo-Argüelles et al., 2015; Sarremejane, Mykrä, Bonada, Aroviita, & Muotka, 2017b). These matrices were calculated for each FHW with sampling periods considered individually. In each case, three beta-diversity matrices were generated: one with the total community, one with strictly aquatic taxa and one with taxa possessing an aerial stage; for a total of nine beta-diversity matrices for each FHW-sampling period combination.

The relationships between beta-diversity and distance matrices were quantified and tested using Mantel tests (Legendre & Legendre, 2012). For environmental distances, we used the MSR (Moran spectral randomization) procedure to correct for potential spatial autocorrelation that can alter the robustness of Mantel tests (Crabot et al., 2019a). Briefly, this method uses a spatially-constrained permutation procedure to account for the spatial structure of the environment and avoids spurious inflation of correlations found in Mantel tests (Crabot et al., 2019a), Appendix 5). Holm correction was then used to adjust p-values for multiple comparisons in Mantel tests (Heino & Tolonen, 2017; Holm, 1979).

(c) Nestedness component

To address our third prediction that beta-diversity would be dominated by nestedness, as a result of local communities at intermittent reaches being subsets of those from perennial reaches, total beta-diversity across sampling periods was partitioned into nestedness and turnover components using the `beta.div.comp` function in R (Podani-family decomposition, Legendre, 2014). The percentage of unique and shared taxa between intermittent and perennial reaches in each FHW was also assessed. Difference of the shared percentages between reach types was then tested using Wilcoxon tests.

All statistical analyses were conducted with R (Team, 2019) using the `lme4` package (mixed-effect models, Bates et al, 2015), `vegan` (distance matrices, Oksanen et al., 2019), `ade4` (mantel tests, S Dray & Dufour, 2007), `adespatial` (`beta.div.comp`, Dray et al, 2019) and `spdep` (MSR, S Dray et al., 2019) packages.

2.5 Results

2.5.1 Context-dependency among FHWs

The two studied areas were located in limestone ranges dominated by a temperate climate (Tableau 2.1). The catchment area of the FHWs ranged from 91 to 156 km², with an average of 117.3 ± 20.5 km². FHW spanned altitudes from 111 to 985 m a.s.l. and the total length of river channels ranged from 58.6 to 259.1 km (mean \pm S.D. = 126.9 ± 62) (Tableau 2.1). The drying period lasted ~ 4 months

on average, with intermittent reaches drying for one to six months. Drying reaches represented from 7.1 to 46.8% of the total river length across the FHWs ($26.0 \pm 11.8\%$, Tableau 2.1). Intermittent reaches sampled represented a consistent proportion of all sampled reaches for each FHWs ($13.5 \pm 1.5\%$). The number of anthropogenic structures varied greatly among FHWs (15 ± 13) with some FHW having none (e.g. Clauge) and some with up to 45 structures (Oignin) (Tableau 2.1).

2.5.2 *Invertebrate communities*

A total of 165 and 220 reaches were sampled before and after the drying season, respectively. In total, 106,954 individuals from 210 taxa were collected. On average, a sample contained 353 ± 1421 individuals (mean \pm S.D.) belonging to 18 ± 7 taxa. Abundances varied between sampling periods: before the drying period, reaches contained 625 ± 786 individuals whereas after the drying period, they contained 35 ± 14 individuals. Mean abundances for each taxon and FHW are shown in Appendix S4. Most taxa belonged to groups of invertebrate with an aerial stage ($66 \pm 4.1\%$). Invertebrate communities were dominated by Simuliidae (Diptera, relative abundance: 18.2%), Gammaridae (Amphipoda, 13.6%) and *Potamopyrgus antipodarum* (Tateidae, Gastropoda, 9.5%). Oligochaeta and Chironomidae (Diptera) were the most ubiquitous taxa across FHWs being found in 92% of all sampled reaches, followed by Orthocladinae (Diptera, 89%) and Simuliidae (Diptera, 71%).

2.5.3 *Alpha-diversity of perennial and intermittent reaches*

Regardless of the sampling period, alpha-diversity was lower at intermittent than perennial reaches (mean \pm S.D.: $\alpha_{INT} = 16.34 \pm 6.44$; $\alpha_{PER} = 21.97 \pm 6.33$, $\chi^2=49.37$, p-value < 0.001 , Figure 2.3). Alpha-diversity was lower after than before the drying period ($\alpha_{BEFORE} = 21.54 \pm 6.65$; $\alpha_{AFTER} = 18.42 \pm 6.85$, $\chi^2=14.90$, p-value < 0.001 , Figure 2.3). There was no interaction between flow regime and sampling period ($\chi^2=0$, p-value=1), in spite of the high variability observed among FHWs (Figure 2.3).

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

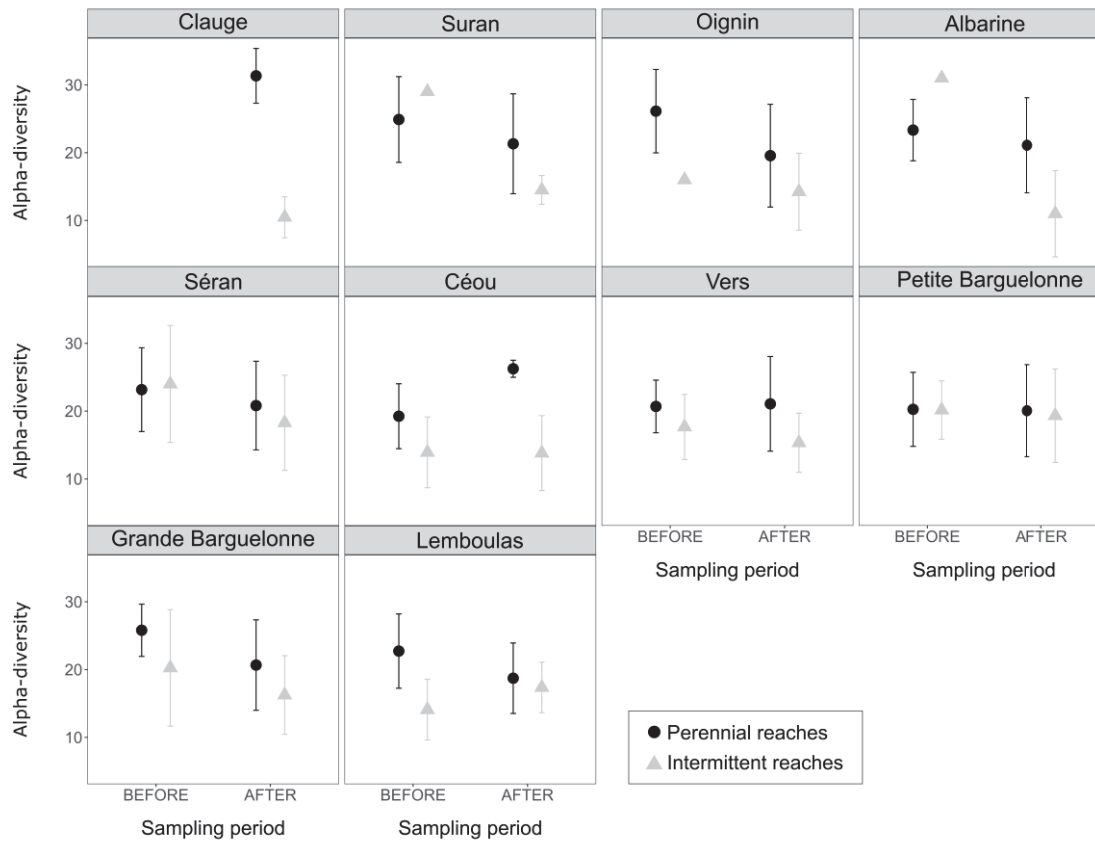


Figure 2.3 : Alpha-diversity distribution (mean \pm standard deviation) for perennial (black circle) and intermittent (grey triangle) reaches for each FHW and for both sampling period (before and after the drying season).

2.5.4 Beta-diversity of perennial and intermittent reaches

All three beta-diversity indices provided consistent results, hence only results based on Chao dissimilarity are shown here, while those for Jaccard and Bray-Curtis are available in Supporting Information (Appendix S5). Across all FHWs and sampling periods and in spite of the high among-FHWs variability, beta-diversity was not correlated to ENV and was systematically correlated to one or several spatial distance matrices, except for two FHWs (Figure 2.4, Tableau 2.2). FRAG was the strongest predictor of beta-diversity across FHWs with an average r of 0.41 ± 0.07 , including all sampling periods (Tableau 2.2). Globally, spatial distances were correlated more with beta-diversity before than after the drying period (Figure 2.4, Tableau 2.2). Lastly, the beta-diversity of organisms with aerial stages showed stronger relationships with physical distances than did beta-diversity of strictly aquatic organisms, regardless of sampling period, except for the FRAG matrix before the drying season (Figure 2.4, Tableau 2.2).

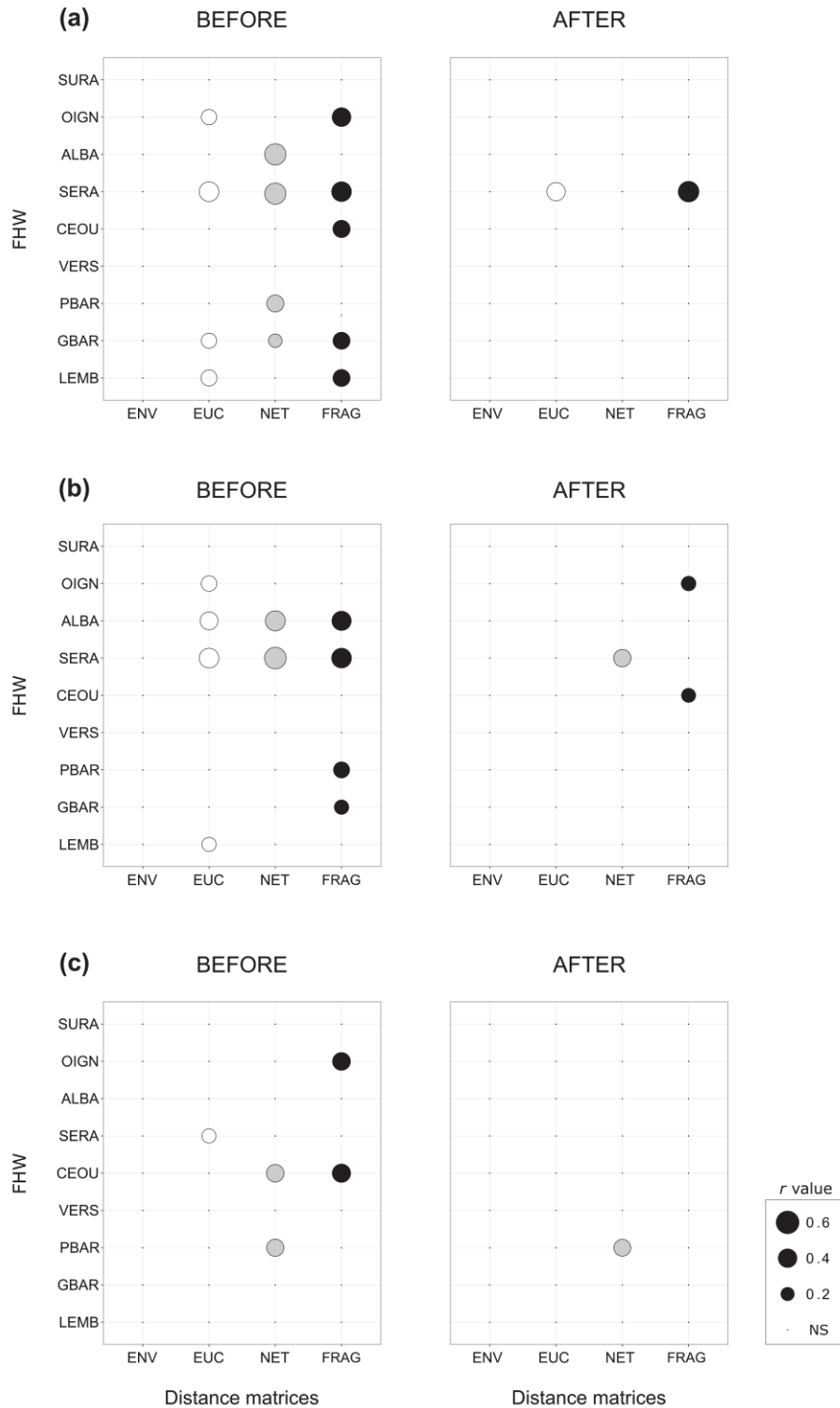


Figure 2.4 : Coefficient of correlation (r value represented by point size) of Mantel test results (see Table 2 for exact values of r) for each combination of community type (Chao dissimilarity index) and distance matrices for each FHW before and after the drying season, for the whole community (a), the community of taxa with aerial stages (b) and strictly aquatic taxa (c). Distance matrices are ENV, environmental distance; EUC, overland distance; NET, network distance; FRAG, distance based on number of dry patches and anthropogenic structures. Only r for which p -value is significant (≤ 0.05) are shown. Corresponding distance-decay relationships are available in Appendix S6.

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Tableau 2.2. Correlation coefficient of the Mantel tests based on Chao dissimilarity index carried out to disentangle metacommunity processes. With respect to beta-diversity: TOT, the whole community, AER, taxa with an aerial life stage, AQU, strictly aquatic taxa. With respect to distances: ENV, environmental; EUC, overland; NET, network; FRAG, number of dry patches and anthropogenic structures. Significant relationships are shown in bold (p-value ≤ 0.05 *). Holm correction was used to adjust p-values for multiple comparisons.

FHW	Distance matrices	Before drying season			After drying season		
		TOT	AER	AQU	TOT	AER	AQU
Clauge	ENV				0.00	-0.02	0.12
	EUC				0.26	0.27 *	0.14
	NET				-0.08	-0.03	-0.14
	FRAG				-0.22	-0.23	-0.11
Suran	ENV	0.07	0.12	-0.04	-0.11	-0.11	-0.05
	EUC	0.11	0.10	0.05	0.00	0.00	0.01
	NET	0.10	0.16	0.01	0.02	0.06	-0.04
	FRAG	0.10	0.16	0.01	0.06	0.05	0.04
Oignin	ENV	-0.16	-0.18	-0.08	0.01	0.01	0.02
	EUC	0.27 *	0.28 *	0.16	0.08	0.15	-0.06
	NET	0.29	0.29	0.28	0.14	0.21	-0.03
	FRAG	0.41 *	0.41	0.37 *	0.18	0.24 *	0.01
Albarine	ENV	-0.09	-0.21	0.05	0.32	0.29	0.24
	EUC	0.36	0.37 *	0.20	-0.01	0.06	-0.09
	NET	0.53 *	0.46 *	0.38	-0.10	-0.05	-0.14
	FRAG	0.45	0.43 *	0.27	0.01	0.05	-0.05
Séran	ENV	0.27	0.22	0.26	0.13	0.08	0.09
	EUC	0.45 *	0.45 *	0.23 *	0.39 *	0.36	0.16
	NET	0.52 *	0.55 *	0.21	0.31	0.35 *	0.03
	FRAG	0.45 *	0.45 *	0.22	0.50 *	0.48	0.22
Céou	ENV	0.20	0.23	0.02	-0.03	-0.02	-0.01
	EUC	0.32	0.34	0.13	-0.10	0.02	-0.18
	NET	0.26	0.11	0.36 *	0.21	0.22	0.07
	FRAG	0.34 *	0.23	0.39 *	0.20	0.23 *	0.04
Vers	ENV	-0.03	0.12	-0.15	0.00	-0.04	0.06
	EUC	0.13	0.12	0.10	0.15	0.13	0.07
	NET	0.05	-0.09	0.12	-0.03	-0.03	0.00
	FRAG	0.04	-0.05	0.09	-0.01	0.01	-0.03
Peite Barguelonne	ENV	0.18	0.22	0.11	0.16	0.04	0.16
	EUC	0.23	0.22	0.15	0.11	0.05	0.12
	NET	0.33 *	0.12	0.34 *	0.13	-0.16	0.33 *
	FRAG	0.31	0.30 *	0.20	0.22	0.13	0.21
Grande Barguelonne	ENV	-0.02	0.06	-0.07	0.15	0.21	-0.04
	EUC	0.26 *	0.22	0.17	0.25	0.28	0.05
	NET	0.21 *	0.19	0.13	0.16	0.21	-0.02
	FRAG	0.34 *	0.24 *	0.27	0.14	0.14	0.05
Lemboulas	ENV	0.13	0.11	0.11	-0.06	0.01	-0.09
	EUC	0.29 *	0.23 *	0.24	-0.01	-0.02	-0.01
	NET	0.17	0.09	0.20	0.03	-0.02	0.07
	FRAG	0.34 *	0.33	0.22	-0.01	0.01	-0.05

2.5.5 *Nestedness and composition differences among perennial and intermittent reaches*

Nestedness and turnover contributed similarly to beta-diversity in all studied FHW, in spite of some among-FHWs variability. On average, 45.01 ± 6.2 % of beta-diversity was due to nestedness (range: [32.7, 51.9]). Intermittent reaches shared more taxa with perennial reaches than perennial reaches shared with intermittent ones (Wilcoxon test, $W=44$, p -value < 0.01). On average, intermittent reaches shared 73.9 ± 10.2 % of taxa with perennial reaches, whereas perennial reaches shared only 53.9 ± 13.6 % of taxa with intermittent ones (Figure 2.5).

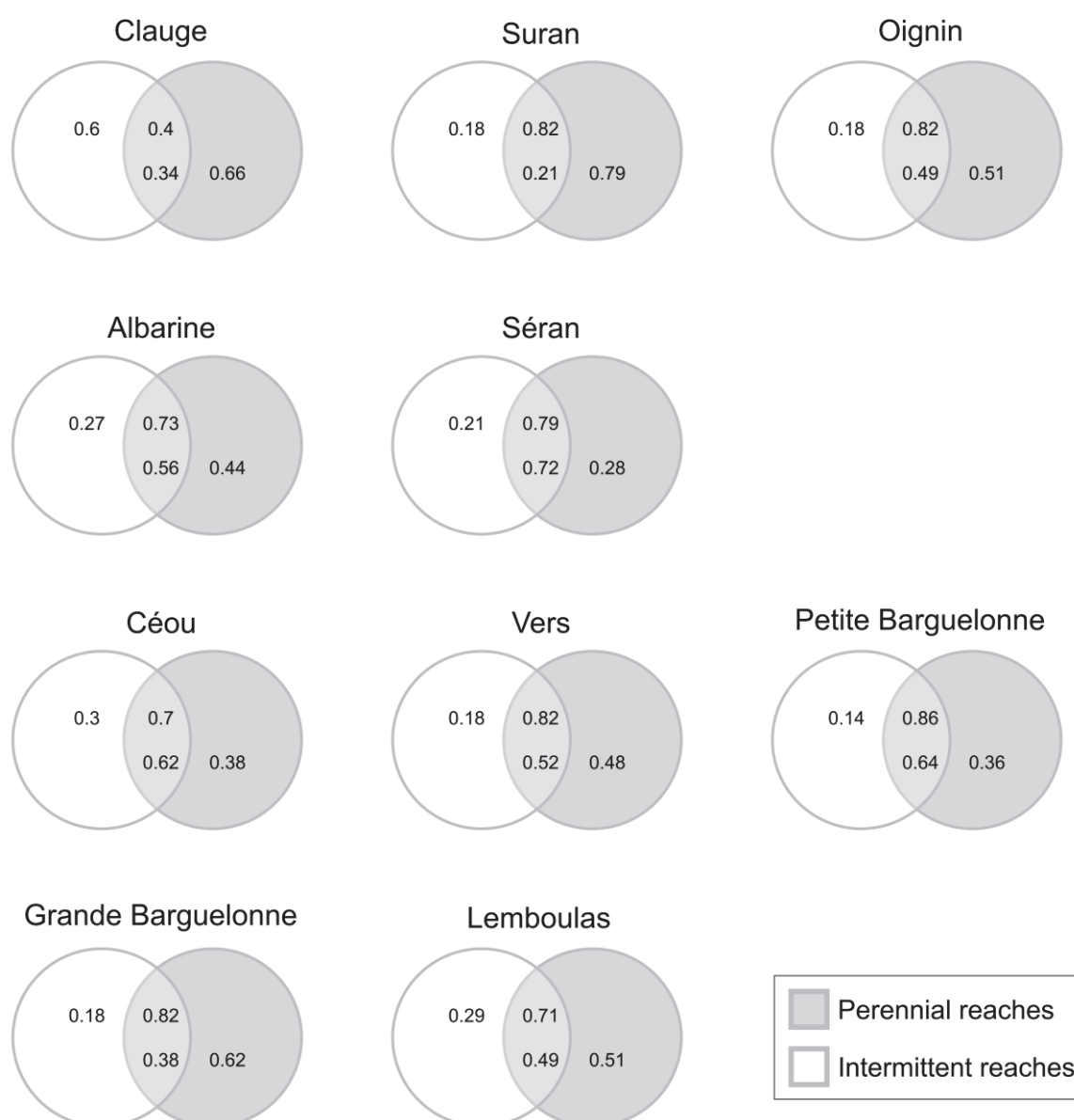


Figure 2.5 : Venn diagrams displaying unique and shared taxa proportion between intermittent (white) and perennial (grey) reaches for each FHW. The difference between shared taxa proportion was tested with a Wilcoxon test between the two reach types and was significant ($W=44$, p -value=0.01).

2.6 Discussion

In the past decade, dispersal has become a pivotal process to consider in community ecology and this particularly in fragmented ecosystems (Altermatt & Fronhofer, 2018; Datry et al., 2016a; Heino et al., 2017; Siqueira et al., 2019). However, dispersal is also one of the most difficult processes to study in nature even for a single individual or a single species, which is clearly exacerbated when considering dozens to hundreds of species in a metacommunity (Heino et al., 2017). This makes it difficult to systematically account for dispersal in metacommunity ecology (Heino et al., 2017; Jacobson & Peres-Neto, 2009; Jones et al., 2015). In general, ecologists are using either proxies for dispersal (e.g. using biological traits, Aspin et al., 2018; Crabot et al., 2019a) or statistical tools (e.g. distance decay relationships, variance partitioning) to unravel its roles in determining metacommunity dynamics (Heino et al., 2017). Recent experimental and observational studies have quantified dispersal directly by measuring drift and terrestrial adult dispersal distances and shown dispersal was a predominant processes structuring perennial river metacommunities (Brooks, Wolfenden, Downes, & Lancaster, 2018; Downes & Lancaster, 2018; Downes et al., 2017). However, whereas there is a growing interest in metacommunities from perennial river networks, little research has considered the effect of occasional fragmentation by drying and effects of small anthropogenic structures on their dynamics (but see Cañedo-Argüelles et al., 2015; Sarremejane et al., 2017b). Even rarer are studies focusing on intermittent headwater streams which lie at the interface between terrestrial and aquatic biomes, experience repeated dry and wet cycles and due to their position, are the most isolated habitats within river networks (Benstead & Leigh, 2012; Finn et al., 2009; Lowe & Likens, 2005). Our study across 10 intermittent headwater metacommunities revealed for the very first time that dispersal is the main determinant of metacommunity dynamics in these fragmented and dynamic ecosystems. As a result, and in spite of the high context-dependency reported here and elsewhere (e.g. Datry et al., 2016b; Heino et al., 2012; Tonkin et al., 2016), fragmentation and associated dispersal limitations led to predictable spatio-temporal patterns of alpha and beta-diversity. As river fragmentation is increasing due to global change, and notably in headwaters of river networks, our results suggest that some freshwater ecosystems currently driven by environmental selection could gradually shift towards dispersal-limited ecosystems. Such shift from perennial to intermittent flow regimes could represent ecological thresholds (sensus Smol & Douglas, 2007), as it is likely to alter irreversibly community composition or eliminate taxa from the metacommunity (Datry, Pella, Leigh, Bonada, & Hugueny, 2016c). Crossing such thresholds will have major basic and applied implications for understanding and managing freshwaters, which are among the most threatened ecosystems on Earth (Datry et al., 2018; Reid et al., 2019; Tonkin et al., 2019).

As predicted, alpha-diversity was lower at reaches prone to drying, regardless of sampling period or the FHW considered aligning with the growing literature on the effects of drying on riverine communities (Aspin et al., 2018; Crabot et al., 2019b; Datry et al., 2014; Leigh & Datry, 2017). While

it is clear today that the immediate effects of natural drying include a sharp decrease in taxonomic richness by filtering species adapted to drying (i.e. resistant species), notably for stream invertebrates (e.g. Aspin et al., 2018; Crabot et al., 2019a; Datry et al., 2014; Leigh & Datry, 2017) but also for algae (Robson & Matthews, 2004; Robson, Matthews, Lind, & Thomas, 2008), our result confirms that drying can have, in some cases, a legacy effect long after rewetting. Here, drying reaches were still exhibiting a lower taxonomic richness than perennial ones even after up to 6-9 months of flowing conditions. This long-term effect of drying most likely reflects the fact that, headwater streams being isolated in river networks, their recolonization by taxa originating from perennial refuges in the networks is a slow process. Such legacy effect was already reported for intermittent rivers and streams lacking refuges during the dry periods across Europe (e.g. Piano et al., 2019), the US (e.g. Bogan & Boersma, 2012) or Australasia (e.g. Arscott et al., 2010). In contrast, in intermittent rivers with abundant refuges during the dry phases, such as braided systems with underneath alluvial hyporheic zones (Vander Vorste et al., 2016), streams with disconnected pools during the dry period (Bonada, Rieradevall, & Prat, 2007) or streams that maintain a high sediment moisture due to high riparian shading (Pařil et al., 2019), community recovery is generally much quicker and occurs within 3-4 weeks. Headwater streams generally lack hyporheic zones, and their relatively small size can preclude the maintenance of pools or humid conditions in their dry riverbeds (Finn, Bonada, Múrria, & Hughes, 2011). Nevertheless, in the Mediterranean region, small pools located in headwaters can act as refuges: for example, in Victoria, Australia, stream temperatures can exceed 30°C, but groundwater-fed pools can maintain a temperature < 20°C (Davis, Pavlova, Thompson, & Sunnucks, 2013). However, in general, the spatial isolation of headwaters in river networks limit dispersal rates and make them highly sensitive to drying.

Beta-diversity was predominantly explained by physical distances that accounted for fragmentation and never by environmental distances. This suggests that dispersal is the major process shaping metacommunity dynamics in these 10 fragmented headwater stream networks, as we initially predicted, despite our inclusion of 12 relevant environmental descriptors leading to a finer description of invertebrates' niches than in many previous stream metacommunity studies (e.g. Brown & Swan, 2010; Cañedo-Argüelles et al., 2015; Datry et al., 2016b; Heino et al., 2012; Tonkin et al., 2016). While this result is consistent with recent studies indicating that dispersal might play a stronger role than initially thought in shaping metacommunities of dynamic freshwater ecosystems (Cañedo-Argüelles et al., 2015; Crabot et al., 2019a; Sarremejane et al., 2017a), it also challenges previous conceptual views of how metacommunity processes are spatially organised in river networks. Due to the dendritism of river networks, it has indeed been proposed that communities in isolated headwaters of river networks would be structured by local environmental filtering, while mainstem reaches would be structured by both environmental filtering and dispersal (Brown & Swan, 2010). Yet, our results clearly indicate a major role of dispersal in these 10 headwater networks, contradicting this conceptual

view. Several reasons could explain such discrepancy. First, most studies examining metacommunity dynamics in headwaters have not quantified fragmentation (e.g. Heino et al., 2012; Tonkin et al., 2016), hence possibly underestimating the role of dispersal. Second, in some cases, the spatial scale at which metacommunities were considered was unclear (e.g. Sarremejane et al., 2017a), although it is a crucial factor directly affecting which processes are emphasized in metacommunity structure (Heino et al., 2012), or the analyses were carried out on data from agency monitoring networks (e.g. Brown & Swan, 2010). Because such networks are generally not designed to study metacommunities, it could lead to conceptual shortcomings by considering local communities, which are physically not connected by dispersal. Last, context-dependency can be substantial in metacommunities and conclusions drawn from a single metacommunity are likely to be weak and non-transferable across systems (McCluney et al., 2014; Tonkin et al., 2016). From this perspective, and in spite of considerable among-network variability, we found congruent patterns across ten metacommunities, which strongly support our conclusion that dispersal play a key role in headwater streams and has clear imprints onto alpha- and beta-diversity components. To explore further the role of dispersal in fragmented headwater streams, further work is needed to quantify the sources of fragmentation (e.g. by drying or small human structures), as invertebrate communities could respond differently to them.

The imprints of dispersal were also detectable within beta-diversity components, as nestedness was remarkably high across all FHWs and sampling dates. The proportion of total beta-diversity being due to nestedness was far much higher than previously reported in various freshwater ecosystems (Soininen, Heino, & Wang, 2018). For example, nestedness contributed less than 20% to total beta-diversity of tropical fish and invertebrate communities in some Bolivian intermittent headwater streams (Datry et al., 2016b). Such contribution was even smaller than 10% in diatom and invertebrates metacommunities of perennial headwaters in the boreal region (Perez Rocha et al., 2019). The high nestedness proportion we found is coherent with the fact that local communities from intermittent reaches were subsets of those from perennial ones. Many taxa were shared between intermittent and perennial reaches, whereas only a few taxa were uniquely found in intermittent reaches. This also suggests that perennial reaches act as sources of colonists for intermittent reaches in these 10 metacommunities. However, our sampling design was more likely to sample the core taxa of the local communities. As taxonomic resolution was not systematically at the species level, these taxa were likely to be ubiquitous, potentially resulting in an underestimation of beta diversity and particularly of its turnover component. Further insights with a more detailed within-reach community description (e.g. multiple habitat sampling per reach) and a better identification level (e.g. species or even OTU level) would be needed. Further, the use of genetic markers to acknowledge the source of dispersers in intermittent headwater metacommunities would be welcomed in future studies to better understand dispersal dynamics between perennial and intermittent reaches as it has already been shown in previous studies on intermittent rivers (e.g. Zickovich & Bohonak, 2007).

Interestingly, significant relationships between beta-diversity and physical distances were more observed more often and were stronger before than after drying events. Yet, the reciprocal was not true: relationships between beta-diversity and environmental distances were not stronger or more frequent after drying events. This suggests that ecological drift (i.e. neutral metacommunity dynamics governed by demographic stochasticity, Vellend, 2010) could predominate in the early successional stages following stream channel rewetting. Conceptually, the lack of environmental or spatial patterns in metacommunity structure is expected to occur in highly-dynamic and disturbed ecosystems (Chase, 2007; Gilbert & Levine, 2017; Silva, Souza, Jardim, & Goto, 2015). Headwaters prone to drying most likely fall into this category and the very low number of significant relationships found after rewetting suggest that metacommunities are not driven by dispersal during the early successional stages, although dispersal may predominate in later stages. Mechanistically, some taxa can quickly recover from desiccation-resistant stages or by being tolerant to low sediment moisture (Pařil et al., 2019; Strachan, Chester, & Robson, 2015; Stubbington & Datry, 2013). These initial successional stages could be independent from either distances to sources of colonists or local environmental conditions. Such temporal changes in metacommunity processes in highly-dynamic ecosystems have been conceptualized (Datry et al., 2016a) and more and more studies are indicating shifts in metacommunity processes occur through time (e.g. Crabot et al., 2019a; Sarremejane et al., 2017a; Jabot et al., 2020). Although beyond the scope of the present study, the discrepancies in metacommunity dynamics before and after the drying period call for a better understanding of how assembly processes are alternating in time in highly-dynamic ecosystems.

The effects of fragmentation on metacommunity dynamics were partly mediated by the dispersal modes and abilities of taxa, as has been shown previously. Aquatic invertebrates show a particularly high variety of dispersal modes and capacities (Heino et al., 2017), which can mitigate the negative effects of network fragmentation. For example, metacommunities of taxa that can disperse overland during their adult stages, such as Ephemeroptera, Trichoptera, or Plecoptera are generally less affected by fragmentation due to drying or other causes, than taxa that are strictly aquatic (Alp et al., 2012; Cañedo-Argüelles et al., 2015; Crabot et al., 2019a). Surprisingly, we found globally a stronger spatial structure of metacommunities of taxa with aerial stages compared to those of strictly aquatic taxa. This finding is supported by the results from studies of invertebrates population genetics in streams that show that dry reaches act as barriers for weak or moderate aerial dispersers dispersing within dry channels, like some Ephemeroptera (Chester, Miller, Valenzuela, Wickson, & Robson, 2015). When not totally dry, lentic conditions prevailing in stream channels during dry phases are often associated with high temperature and low dissolved oxygen, (Anna, Yorgos, Konstantinos, & Maria, 2008; Boulton, 2003), which could potentially create “desert areas” for aerial dispersers with selective oviposition behavior. Alternatively, if pools are used for oviposition, poor water quality may decrease a negative effect by decreasing egg or larval survival, thus limiting settlement and spread of aerial

dispersers (Alp, Indermaur, & Robinson, 2013; Encalada & Peckarsky, 2007; Lancaster & Downes, 2018). In addition to enhance the understanding of underlying mechanisms in fragmented ecosystems (e.g. expected stronger species sorting in pools, Datry et al., 2016a), exploring metacommunity patterns from pools which can host strong dispersers (e.g. Anna et al., 2008), the hyporheic zone which can host weak dispersers (e.g. Stubbington, Sarremejane, & Datry, 2019) or integrate other taxonomic phyla sensitive to fragmentation and for whom further insights could be drawn (e.g. fish, diatoms, Chester et al., 2014; Ludlam & Magoulick, 2009) could help to understand how dispersal abilities can mediate the negative (Petersen, Masters, Hildrew, & Ormerod, 2004) effects of fragmentation. Another explanation could stem from limitations to overland dispersal in headwater streams, particularly when heavily-forested as in some FHWs studied here. There were some previous indications of such limitation in other systems in which taxa with aerial stages dispersed mostly along river networks than overland (Bogan & Boersma, 2012; Griffith, Barrows, & Perry, 1998; Petersen et al., 2004; Phillipsen & Lytle, 2013; Razeng et al., 2016) and, when dispersing overland, more easily in cleared catchments (Petersen et al., 2004). This further clearly supports recent calls (Heino et al., 2017; Sarremejane et al., 2019) for a better understanding of dispersal abilities when exploring freshwater metacommunity dynamics.

The fragmentation of river networks by drying and anthropogenic structures is increasing in space and time as a response to global change (Anderson et al., 2019; de Graaf, Gleeson, Rens van Beek, Sutanudjaja, & Bierkens, 2019; Döll & Schmied, 2012; Grill et al., 2019). Our results suggest that riverine metacommunities currently driven by environmental selection could gradually shift towards dispersal-limited ecosystems due to dispersal limitation. Such a shift would echo the rapid change of arctic pond ecosystems in the last two decades and the call to react rapidly to anticipate these shifts for their management (Smol, 2016; Smol & Douglas, 2007). The shifts from perennial to intermittent flow regimes represent an ecological threshold that should not be crossed to avoid jeopardising river biodiversity, functional integrity and the ecosystem services streams provide (Datry et al., 2018; Tonkin et al., 2019).

2.7 Acknowledgements

This work was supported by the French Agency for Biodiversity (AFB) through the program n° 14 entitled “Headwater biodiversity dynamics”. This manuscript benefited from the meaningful comments from Miguel Cañedo-Argüelles, Mathis Messenger and 2 anonymous reviewers and the editor. We thank the AFB members, in particular Claire Magand and Lionel Sainte-Olympe, the members of syndicate cc-Bugey Sud, cc-Haut-Bugey, SMISA and SR3a, and members of ONF for their help to select and expertise on studied headwater stream networks. We thank Claire Darras, Mélanie Milin, Quentin Schaming, Benjamin Gerfand, Colin Issartel for help with sample collection and processing. We thank Maxence Forcellini, Julien Barnasson and Jacques Mouthon for identification expertise on Trichoptera and Mollusks. We thank Julie Crabot, Nuria Cid-Puey and Maxence Forcellini for useful comments on data analysis. The authors declare no conflict of interest.

References used to identify organisms morphologically organisms are available from the Zenodo Digital Repository: <https://doi.org/10.5281/zenodo.3569563>

2.8 Data accessibility

FHW environmental tables are available from the Zenodo Digital Repository: <https://doi.org/10.5281/zenodo.3569563>. Invertebrate data are available from the authors upon reasonable request.

2.9 References

- Alp, M., Indermaur, L., & Robinson, C. T. (2013). Environmental constraints on oviposition of aquatic invertebrates with contrasting life cycles in two human-modified streams. *Freshwater Biology*, 58(9), 1932-1945. doi:10.1111/fwb.12181
- Alp, M., Keller, I., Westram, A. M., & Robinson, C. T. (2012). How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology*, 57(5), 969-981. doi:10.1111/j.1365-2427.2012.02758.x
- Altermatt, F., & Fronhofer, E. A. (2018). Dispersal in dendritic networks: Ecological consequences on the spatial distribution of population densities. *Freshwater Biology*, 63(1), 22-32. doi:10.1111/fwb.12951
- Altermatt, F., Schreiber, S., & Holyoak, M. (2011). Interactive effects of disturbance and dispersal directionality on species richness and composition in metacommunities. *Ecology*, 92(4), 859-870. doi:10.1890/10-1095.1

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

- Anderson, R. L., Anderson, C. A., Larson, J. H., Knights, B., Vallazza, J., Jenkins, S. E., & Lamer, J. T. (2019). Influence of a high-head dam as a dispersal barrier to fish community structure of the Upper Mississippi River. *River Research and Applications*. doi:10.1002/rra.3534
- Anna, A., Yorgos, C., Konstantinos, P., & Maria, L. (2008). Do intermittent and ephemeral Mediterranean rivers belong to the same river type? *Aquatic Ecology*, 43(2), 465-476. doi:10.1007/s10452-008-9176-9
- Arscott, D. B., Larned, S., Scarsbrook, M. R., & Lambert, P. (2010). Aquatic invertebrate community structure along an intermittence gradient: Selwyn River, New Zealand. *Journal of the North American Benthological Society*, 29(2), 530-545. doi:10.1899/08-124.1
- Aspin, T. W. H., Matthews, T. J., Khamis, K., Milner, A. M., Wang, Z., O'Callaghan, M. J., & Ledger, M. E. (2018). Drought intensification drives turnover of structure and function in stream invertebrate communities. *Ecography*, 41(12), 1992-2004. doi:10.1111/ecog.03711
- Benstead, J. P., & Leigh, D. S. (2012). An expanded role for river networks. *Nature Geoscience*, 5(10), 678-679. doi:10.1038/ngeo1593
- Bogan, M. T., & Boersma, K. S. (2012). Aerial dispersal of aquatic invertebrates along and away from arid-land streams. *Freshwater Science*, 31(4), 1131-1144. doi:10.1899/12-066.1
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H., & White, J. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol*, 24(3), 127-135. doi:10.1016/j.tree.2008.10.008
- Bonada, N., Rieradevall, M., & Prat, N. (2007). Macroinvertebrate community structure and biological traits related to flow permanence in a Mediterranean river network. *Hydrobiologia*, 589(1), 91-106. doi:10.1007/s10750-007-0723-5
- Boulton, A. J. (2003). Parallels and contrasts in the effects of drought on stream macroinvertebrate assemblages. *Freshwater Biology*, 48(7), 1173-1185. doi:10.1046/j.1365-2427.2003.01084.x
- Brooks, A. J., Wolfenden, B., Downes, B. J., & Lancaster, J. (2018). Barriers to dispersal: The effect of a weir on stream insect drift. *River Research and Applications*, 34(10), 1244-1253. doi:10.1002/rra.3377
- Brown, B. L., & Swan, C. M. (2010). Dendritic network structure constrains metacommunity properties in riverine ecosystems. *J Anim Ecol*, 79(3), 571-580. doi:10.1111/j.1365-2656.2010.01668.x

Campbell Grant, E. H., Lowe, W. H., & Fagan, W. F. (2007). Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecol Lett*, 10(2), 165-175. doi:10.1111/j.1461-0248.2006.01007.x

Cañedo-Argüelles, M., Boersma, K. S., Bogan, M. T., Olden, J. D., Phillipsen, I., Schriever, T. A., . . . Ladle, R. (2015). Dispersal strength determines meta-community structure in a dendritic riverine network. *Journal of Biogeography*, 42(4), 778-790. doi:10.1111/jbi.12457

Chase, J. M. (2007). Drought mediates the importance of stochastic community assembly. *Proceedings of the National Academy of Sciences*, 104, 17430-17434.

Chester, Matthews, T. G., Howson, T. J., Johnston, K., Mackie, J. K., Strachan, S. R., & Robson, B. J. (2014). Constraints upon the response of fish and crayfish to environmental flow releases in a regulated headwater stream network. *PLoS One*, 9(3), e91925. doi:10.1371/journal.pone.0091925

Chester, Miller, A. D., Valenzuela, I., Wickson, S. J., & Robson, B. J. (2015). Drought survival strategies, dispersal potential and persistence of invertebrate species in an intermittent stream landscape. *Freshwater Biology*, 60(10), 2066-2083. doi:10.1111/fwb.12630

Crabot, J., Clappe, S., Dray, S., Datry, T., & McCrea, R. (2019a). Testing the Mantel statistic with a spatially- constrained permutation procedure. *Methods in Ecology and Evolution*, 10(4), 532-540. doi:10.1111/2041-210x.13141

Crabot, J., Heino, J., Launay, B., & Datry, T. (2019b). Drying determines the temporal dynamics of stream invertebrate structural and functional beta diversity. *Ecography*, 1-16. doi:10.1111/ecog.04835

Datry, T., Bonada, N., & Heino, J. (2016a). Towards understanding the organisation of metacommunities in highly dynamic ecological systems. *Oikos*, 125(2), 149-159. doi:10.1111/oik.02922

Datry, T., Boulton, A. J., Bonada, N., Fritz, K., Leigh, C., Sauquet, E., . . . Dahm, C. N. (2018). Flow intermittence and ecosystem services in rivers of the Anthropocene. *J Appl Ecol*, 55(1), 353-364. doi:10.1111/1365-2664.12941

Datry, T., Larned, S. T., Fritz, K. M., Bogan, M. T., Wood, P. J., Meyer, E. I., & Santos, A. N. (2014). Broad-scale patterns of invertebrate richness and community composition in temporary rivers: effects of flow intermittence. *Ecography*, 37(1), 94-104. doi:10.1111/j.1600-0587.2013.00287.x

Datry, T., Melo, A. S., Moya, N., Zubietta, J., De la Barra, E., & Oberdorff, T. (2016b). Metacommunity patterns across three Neotropical catchments with varying environmental harshness. *Freshwater Biology*, 61(3), 277-292. doi:10.1111/fwb.12702

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Datry, T., Pella, H., Leigh, C., Bonada, N., & Hugueny, B. (2016c). A landscape approach to advance intermittent river ecology. *Freshwater Biology*, 61(8), 1200-1213. doi:10.1111/fwb.12645

Davis, J., Pavlova, A., Thompson, R., & Sunnucks, P. (2013). Evolutionary refugia and ecological refuges: key concepts for conserving Australian arid zone freshwater biodiversity under climate change. *Glob Chang Biol*, 19(7), 1970-1984. doi:10.1111/gcb.12203

de Graaf, I. E. M., Gleeson, T., Rens van Beek, L. P. H., Sutanudjaja, E. H., & Bierkens, M. F. P. (2019). Environmental flow limits to global groundwater pumping. *Nature*, 574(7776), 90-94. doi:10.1038/s41586-019-1594-4

Döll, P., & Schmied, H. M. (2012). How is the impact of climate change on river flow regimes related to the impact on mean annual runoff? A global-scale analysis. *Environmental Research Letters*, 7(1). doi:10.1088/1748-9326/7/1/014037

Downes, B. J., & Lancaster, J. (2018). Itinerant, nomad or invader? A field experiment sheds light on the characteristics of successful dispersers and colonists. *Freshwater Biology*, 63(11), 1394-1406. doi:10.1111/fwb.13167

Downes, B. J., Lancaster, J., Glaister, A., Bovill, W. D., & Strecker, A. (2017). A fresh approach reveals how dispersal shapes metacommunity structure in a human-altered landscape. *Journal of Applied Ecology*, 54(2), 588-598. doi:10.1111/1365-2664.12759

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., . . . Wagner, H. H. (2019). *adespatial: Multivariate Multiscale Spatial Analysis*. R package version 0.3-7.

Dray, S., & Dufour, A. (2007). The *ade4* Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software*, 22(4), 1-20. doi:doi: 10.18637/jss.v022.i04

Encalada, A. C., & Peckarsky, B. L. (2007). A comparative study of the costs of alternative mayfly oviposition behaviors. *Behavioral Ecology and Sociobiology*, 61(9), 1437-1448. doi:10.1007/s00265-007-0376-4

Finn, D. S., Bogan, M. T., & Lytle, D. A. (2009). Demographic stability metrics for conservation prioritization of isolated populations. *Conserv Biol*, 23(5), 1185-1194. doi:10.1111/j.1523-1739.2009.01226.x

Finn, D. S., Bonada, N., Múrria, C., & Hughes, J. M. (2011). Small but mighty: headwaters are vital to stream network biodiversity at two levels of organization. *Journal of the North American Benthological Society*, 30(4), 963-980. doi:10.1899/11-012.1

- Finn, D. S., & Poff, N. L. (2011). Examining spatial concordance of genetic and species diversity patterns to evaluate the role of dispersal limitation in structuring headwater metacommunities. *Journal of the North American Benthological Society*, 30(1), 273-283. doi:10.1899/10-035.1
- Fuller, M. R., Doyle, M. W., & Strayer, D. L. (2015). Causes and consequences of habitat fragmentation in river networks. *Ann N Y Acad Sci*, 1355, 31-51. doi:10.1111/nyas.12853
- Gilbert, B., & Levine, J. M. (2017). Ecological drift and the distribution of species diversity. *Proceedings of the Royal Society B: Biological Sciences*, 284(1855), 20170507. doi:doi:10.1098/rspb.2017.0507
- Griffith, M. B., Barrows, E. M., & Perry, S. A. (1998). Lateral Dispersal of Adult Aquatic Insects (Plecoptera, Trichoptera) following Emergence from Headwater Streams in Forested Appalachian Catchments. *Annals of the Entomological Society of America*, 91(2), 195-201. doi:10.1093/aesa/91.2.195
- Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., . . . Zarfl, C. (2019). Mapping the world's free-flowing rivers. *Nature*, 569(7755), 215-221. doi:10.1038/s41586-019-1111-9
- Heino, J., Alahuhta, J., Ala-Hulkko, T., Antikainen, H., Bini, L. M., Bonada, N., . . . Soininen, J. (2017). Integrating dispersal proxies in ecological and environmental research in the freshwater realm. *Environmental Reviews*, 25, 334-349. doi:10.1139/er-2016-0110#.WlyFLHIG2Uk
- Heino, J., Grönroos, M., Soininen, J., Virtanen, R., & Muotka, T. (2012). Context dependency and metacommunity structuring in boreal headwater streams. *Oikos*, 121(4), 537-544. doi:10.1111/j.1600-0706.2011.19715.x
- Heino, J., & Tolonen, K. T. (2017). Ecological drivers of multiple facets of beta diversity in a lentic macroinvertebrate metacommunity. *Limnology and Oceanography*, 62(6), 2431-2444. doi:10.1002/lno.10577
- Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2), 65-70.
- Jabot, F., Laroche, F., Massol, F., Arthaud, F., Crabot, J., Dubart, M., . . . Datry, T. (2020). Assessing metacommunity processes through signatures in spatiotemporal turnover of community composition. *Ecology Letters*.
- Jacobson, B., & Peres-Neto, P. R. (2009). Quantifying and disentangling dispersal in metacommunities: how close have we come? How far is there to go? *Landscape Ecology*, 25(4), 495-507. doi:10.1007/s10980-009-9442-9

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

- Jones, N. T., Germain, R. M., Grainger, T. N., Hall, A. M., Baldwin, L., Gilbert, B., & Jacquemyn, H. (2015). Dispersal mode mediates the effect of patch size and patch connectivity on metacommunity diversity. *Journal of Ecology*, 103(4), 935-944. doi:10.1111/1365-2745.12405
- Karaouzas, I., Theodoropoulos, C., Vourka, A., Gritzalis, K., & Skoulikidis, N. T. (2019). Stream invertebrate communities are primarily shaped by hydrological factors and ultimately fine-tuned by local habitat conditions. *Sci Total Environ*, 665, 290-299. doi:10.1016/j.scitotenv.2019.02.134
- Keleher, M. J., & Rader, R. B. (2008). Dispersal limitations and history explain community composition of metaphyton in desert springs of the Bonneville Basin, Utah: A multiscale analysis. *Limnol. Oceanogr.*, 54(4), 1604-1613.
- Lancaster, J., & Downes, B. J. (2018). Aquatic versus Terrestrial Insects: Real or Presumed Differences in Population Dynamics? *Insects*, 9(4). doi:10.3390/insects9040157
- Langenheder, S., & Lindstrom, E. S. (2019). Factors influencing aquatic and terrestrial bacterial community assembly. *Environ Microbiol Rep*, 11(3), 306-315. doi:10.1111/1758-2229.12731
- Legendre, P. (2014). Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography*, 23.
- Legendre, P., & Legendre, L. (2012). *Numerical ecology* (E. e. Elsevier Ed. 3 ed.). Amsterdam: English edn. Elsevier.
- Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., . . . Gonzalez, A. (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, 7(7), 601-613. doi:10.1111/j.1461-0248.2004.00608.x
- Leigh, C., & Datry, T. (2017). Drying as a primary hydrological determinant of biodiversity in river systems: a broad-scale analysis. *Ecography*, 40(4), 487-499. doi:10.1111/ecog.02230
- Lowe, W. H., & Likens, G. E. (2005). Moving Headwater Streams to the Head of the Class. *BioScience*, 55(3). doi:10.1641/0006-3568(2005)055[0196:Mhstth]2.0.Co;2
- Ludlam, J. P., & Magoulick, D. D. (2009). Spatial and temporal variation in the effects of fish and crayfish on benthic communities during stream drying. *Journal of the North American Benthological Society*, 28(2), 371-382. doi:10.1899/08-149.1
- McCluney, K. E., Poff, N. L., Palmer, M. A., Thorp, J. H., Poole, G. C., Williams, B. S., . . . Baron, J. S. (2014). Riverine macrosystems ecology: sensitivity, resistance, and resilience of whole river basins with human alterations. *Frontiers in Ecology and the Environment*, 12(1), 48-58. doi:10.1890/120367
- Mouquet, N., & Loreau, L. (2003). Community Patterns in Source- Sink Metacommunities. *The American Naturalist*, 162(5), 544-557. doi:10.1086/378857

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., . . . Wagner, H. (2019). *vegan: Community Ecology Package*. R package version 2.5-5.

Pařil, P., Polářek, M., Loskotová, B., Straka, M., Crabot, J., & Datry, T. (2019). An unexpected source of invertebrate community recovery in intermittent streams from a humid continental climate. *Freshwater Biology*, 64(11), 1971-1983. doi:10.1111/fwb.13386

Perez Rocha, M., Bini, L. M., Gronroos, M., Hjort, J., Lindholm, M., Karjalainen, S. M., . . . Heino, J. (2019). Correlates of different facets and components of beta diversity in stream organisms. *Oecologia*, 191(4), 919-929. doi:10.1007/s00442-019-04535-5

Petersen, I., Masters, Z., Hildrew, A. G., & Ormerod, S. J. (2004). Dispersal of adult aquatic insects in catchments of differing land use. *Journal of Applied Ecology*, 41(5), 934-950. doi:10.1111/j.0021-8901.2004.00942.x

Phillipsen, I. C., & Lytle, D. A. (2013). Aquatic insects in a sea of desert: population genetic structure is shaped by limited dispersal in a naturally fragmented landscape. *Ecography*, 36(6), 731-743. doi:10.1111/j.1600-0587.2012.00002.x

Piano, E., Doretto, A., Falasco, E., Fenoglio, S., Gruppuso, L., Nizzoli, D., . . . Bona, F. (2019). If Alpine streams run dry: the drought memory of benthic communities. *Aquatic Sciences*, 81(2). doi:10.1007/s00027-019-0629-0

Rader, R. B., Unmack, P. J., Christensen, W. F., & Jiang, X. (2019). Connectivity of two species with contrasting dispersal abilities: a test of the isolated tributary hypothesis. *Freshwater Science*, 38(1), 142-155. doi:10.1086/701671

Razeng, E., Morán-Ordóñez, A., Brim Box, J., Thompson, R., Davis, J., & Sunnucks, P. (2016). A potential role for overland dispersal in shaping aquatic invertebrate communities in arid regions. *Freshwater Biology*, 61(5), 745-757. doi:10.1111/fwb.12744

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., . . . Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biol Rev Camb Philos Soc*, 94(3), 849-873. doi:10.1111/brv.12480

Resh, V. H., Bêche, L. A., Lawrence, J. E., Mazor, R. D., McElravy, E. P., O'Dowd, A. P., . . . Carlson, S. M. (2012). Long-term population and community patterns of benthic macroinvertebrates and fishes in Northern California Mediterranean-climate streams. *Hydrobiologia*, 719(1), 93-118. doi:10.1007/s10750-012-1373-9

Robson, B. J., & Matthews, T. G. (2004). Drought refuges affect algal recolonization in intermittent streams. *River Research and Applications*, 20(7), 753-763. doi:10.1002/rra.789

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Robson, B. J., Matthews, T. G., Lind, P. R., & Thomas, N. A. (2008). Pathways for algal recolonization in seasonally-flowing streams. *Freshwater Biology*, 53(12), 2385-2401. doi:10.1111/j.1365-2427.2008.02061.x

Sarremejane, R., Cañedo-Argüelles, M., Prat, N., Mykrä, H., Muotka, T., & Bonada, N. (2017a). Do metacommunities vary through time? Intermittent rivers as model systems. *Journal of Biogeography*, 44(12), 2752-2763. doi:10.1111/jbi.13077

Sarremejane, R., Mykrä, H., Bonada, N., Aroviita, J., & Muotka, T. (2017b). Habitat connectivity and dispersal ability drive the assembly mechanisms of macroinvertebrate communities in river networks. *Freshwater Biology*, 62(6), 1073-1082. doi:10.1111/fwb.12926

Sarremejane, R., Stubbington, R., Dunbar, M. J., Westwood, C. G., & England, J. (2019). Biological indices to characterize community responses to drying in streams with contrasting flow permanence regimes. *Ecological Indicators*, 107. doi:10.1016/j.ecolind.2019.105620

Schmera, D., Árva, D., Boda, P., Bódis, E., Bolgovics, Á., Borics, G., . . . Erős, T. (2018). Does isolation influence the relative role of environmental and dispersal-related processes in stream networks? An empirical test of the network position hypothesis using multiple taxa. *Freshwater Biology*, 63(1), 74-85. doi:10.1111/fwb.12973

Silva, J. L. A., Souza, A. F., Jardim, J. G., & Goto, B. T. (2015). Community assembly in harsh environments: the prevalence of ecological drift in the heath vegetation of South America. *Ecosphere*, 6(7). doi:10.1890/es14-00548.1

Siqueira, T., Saito, V. S., Bini, L. M., Melo, A. S., Petsch, D. K., Landeiro, V. L., . . . Heino, J. (2019). Community size affects the signals of selection and ecological drift on biodiversity. *bioRxiv*.

Smol, J. P. (2016). Arctic and Sub-Arctic shallow lakes in a multiple-stressor world: a paleoecological perspective. *Hydrobiologia*, 778(1), 253-272. doi:10.1007/s10750-015-2543-3

Smol, J. P., & Douglas, M. S. V. (2007). Crossing the final ecological threshold in high Arctic ponds. *Proceedings of the National Academy of Sciences*, 104, 12395-12397.

Soininen, J., Heino, J., & Wang, J. (2018). A meta-analysis of nestedness and turnover components of beta diversity across organisms and ecosystems. *Global Ecology and Biogeography*, 27(1), 96-109. doi:10.1111/geb.12660

Strachan, S. R., Chester, E. T., & Robson, B. J. (2015). Freshwater Invertebrate Life History Strategies for Surviving Desiccation. *Springer Science Reviews*, 3(1), 57-75. doi:10.1007/s40362-015-0031-9

Stubbington, R., & Datry, T. (2013). The macroinvertebrate seedbank promotes community persistence in temporary rivers across climate zones. *Freshwater Biology*, 58(6), 1202-1220. doi:10.1111/fwb.12121

Stubbington, R., Sarremejane, R., & Datry, T. (2019). Alpha and beta diversity of connected benthic–subsurface invertebrate communities respond to drying in dynamic river ecosystems. *Ecography*, 42(12), 2060-2073. doi:10.1111/ecog.04592

Tachet, H., Richoux, P., Bournaud, M., & Usseglio-Polatera, P. (2000). *Invertébrés d'eau douce: systématique, biologie, écologie* (P. CNRS éditions Ed.).

Team, R. C. (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Tonkin, J. D., Altermatt, F., Finn, D. S., Heino, J., Olden, J. D., Pauls, S. U., & Lytle, D. A. (2018a). The role of dispersal in river network metacommunities: Patterns, processes, and pathways. *Freshwater Biology*, 63(1), 141-163. doi:10.1111/fwb.13037

Tonkin, J. D., Heino, J., & Altermatt, F. (2018b). Metacommunities in river networks: The importance of network structure and connectivity on patterns and processes. *Freshwater Biology*, 63(1), 1-5. doi:10.1111/fwb.13045

Tonkin, J. D., Heino, J., Sundermann, A., Haase, P., & Jähnig, S. C. (2016). Context dependency in biodiversity patterns of central German stream metacommunities. *Freshwater Biology*, 61(5), 607-620. doi:10.1111/fwb.12728

Tonkin, J. D., Leroy Poff, N., Bond, N. R., Horne, A., Merritt, D. M., Reynolds, L. V., . . . Lytle, D. A. (2019). Prepare river ecosystems for an uncertain future. *Nature*, 570, 301-303. doi:10.1038/d41586-019-01877-1

Vander Vorste, R., Malard, F., & Datry, T. (2016). Is drift the primary process promoting the resilience of river invertebrate communities? A manipulative field experiment in an intermittent alluvial river. *Freshwater Biology*, 61(8), 1276-1292. doi:10.1111/fwb.12658

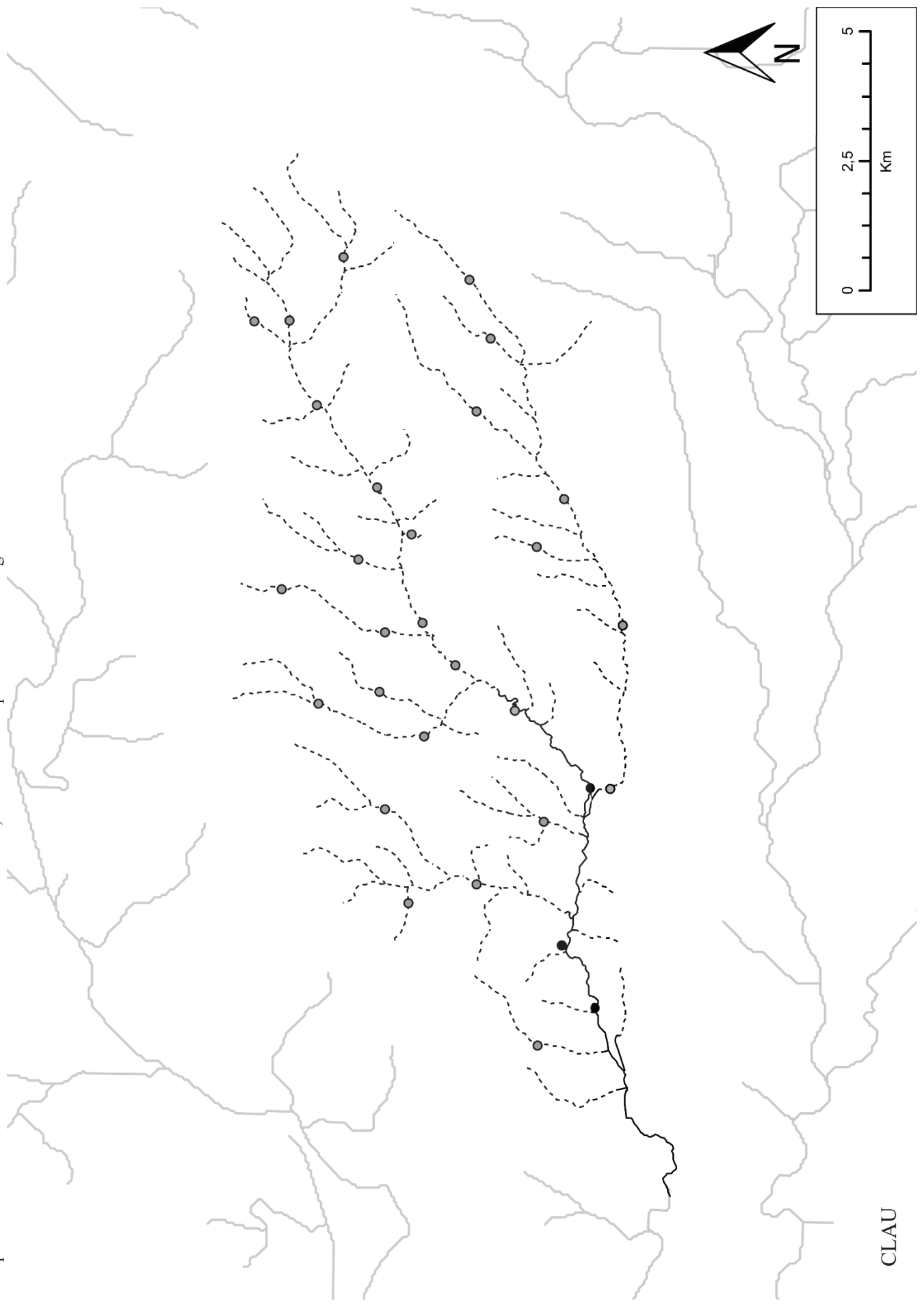
Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*, 85(2), 183-206. doi:10.1086/652373

Zickovich, J. M., & Bohonak, A. J. (2007). Dispersal ability and genetic structure in aquatic invertebrates: a comparative study in southern California streams and reservoirs. *Freshwater Biology*, 52(10), 1982-1996. doi:10.1111/j.1365-2427.2007.01822.x

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

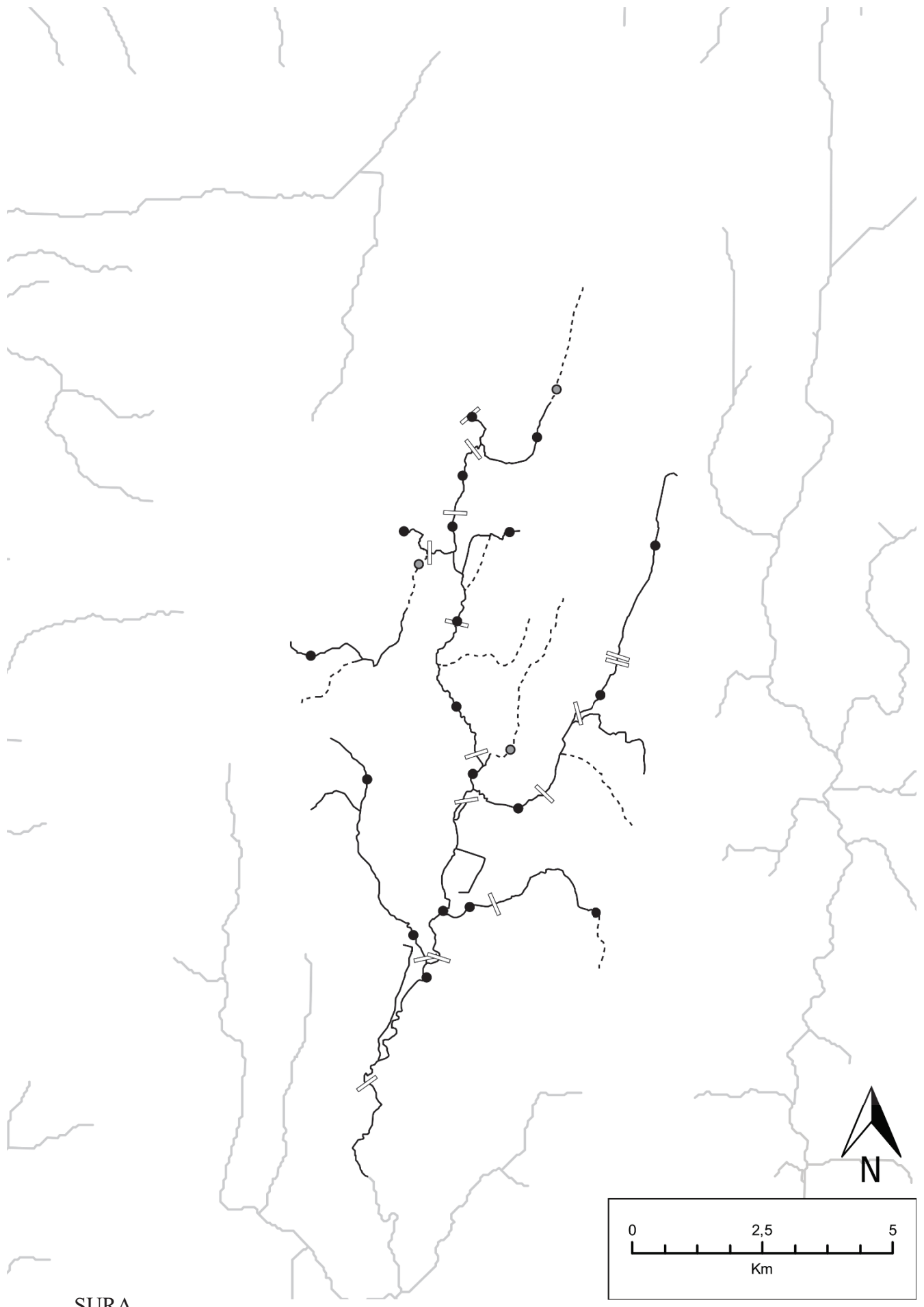
2.10 Supplemental Information

Appendix S1. Detailed maps for each FWH showing site locations, weir and drying locations. *Grey and black circles correspond to intermittent and perennial sampling sites, respectively. Dashed and filled lines correspond to intermittent and perennial reaches, respectively. Anthropogenic structures are represented by white rectangles on the network.*



CLAU

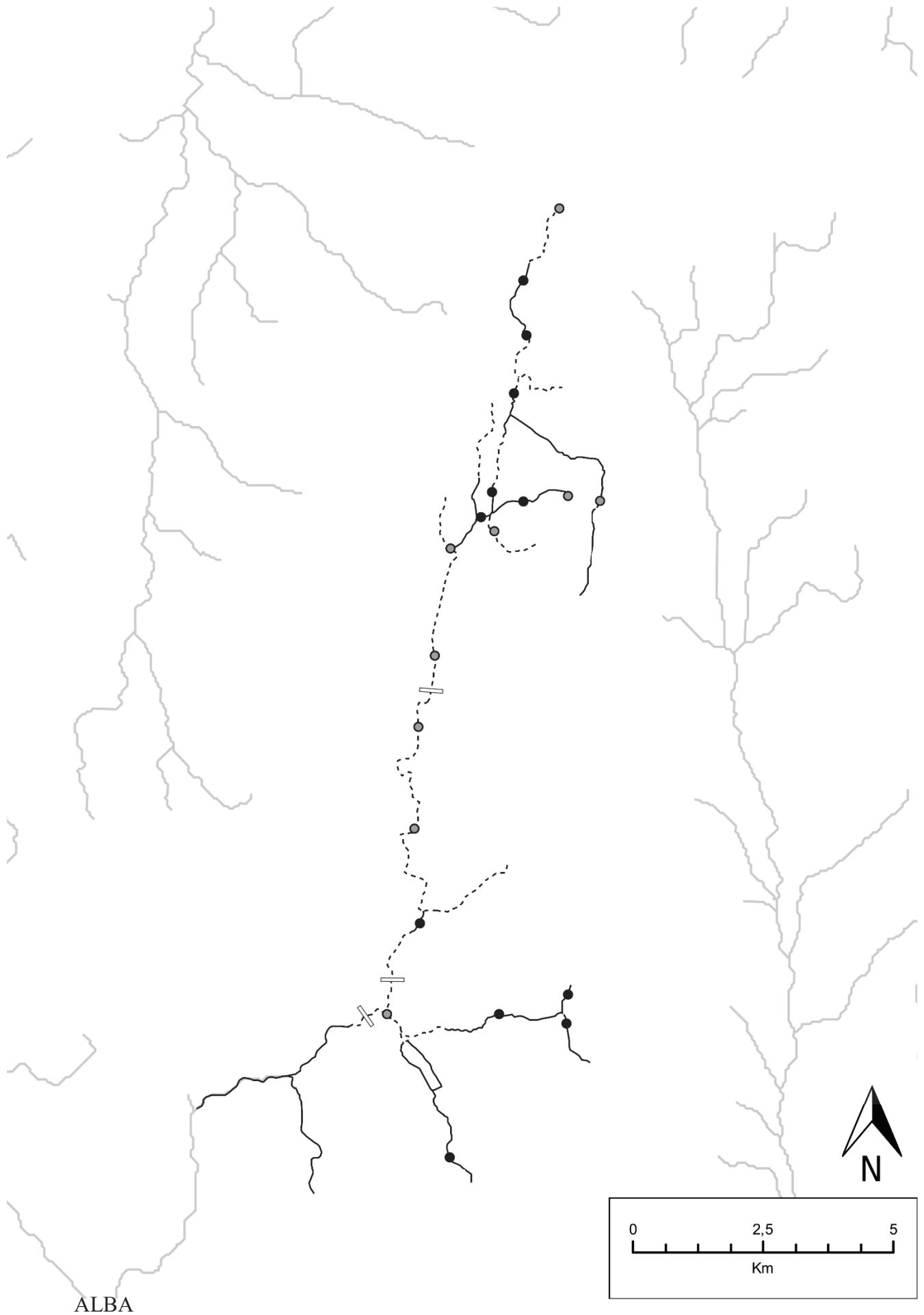
2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)



SURA

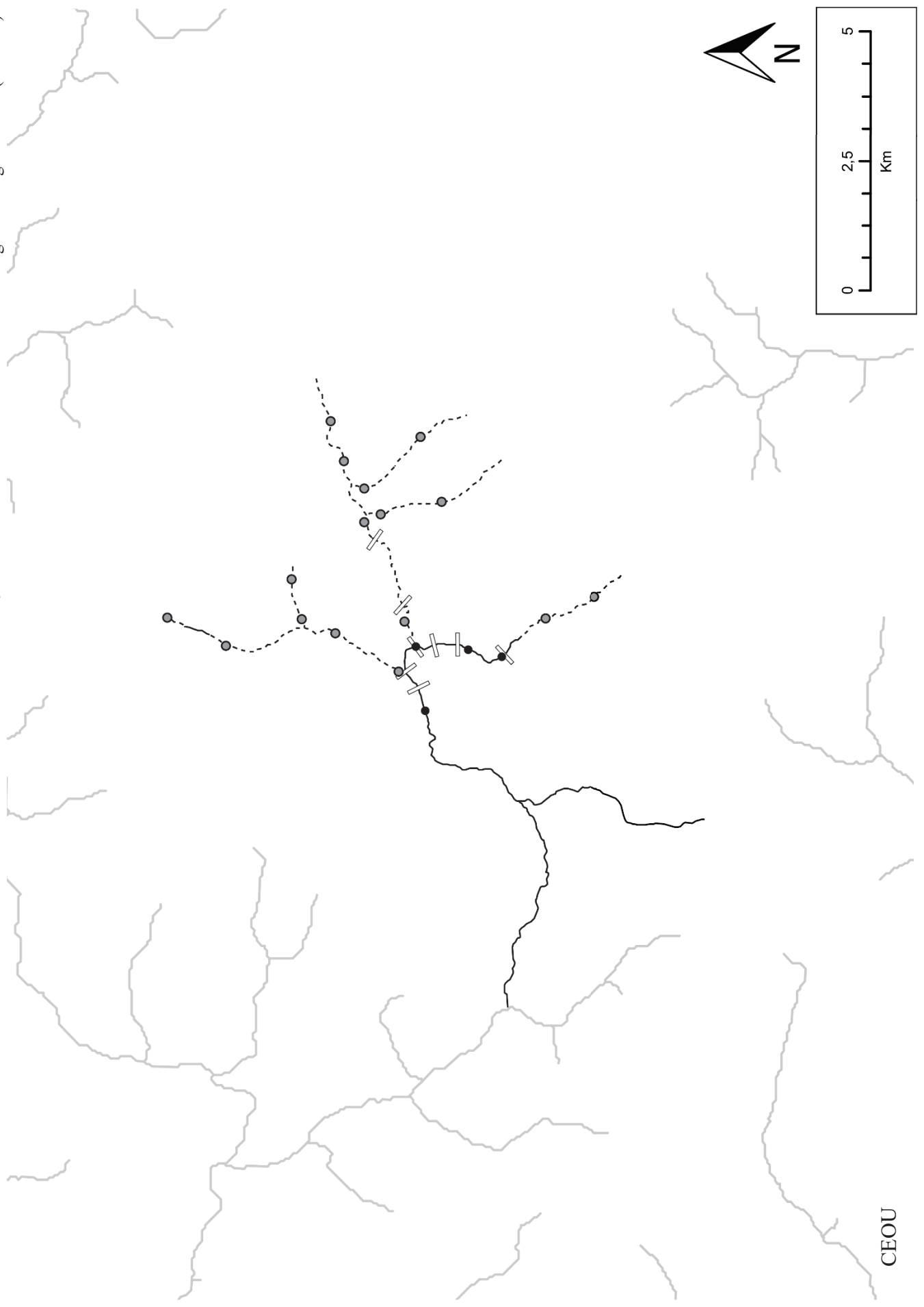


2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)





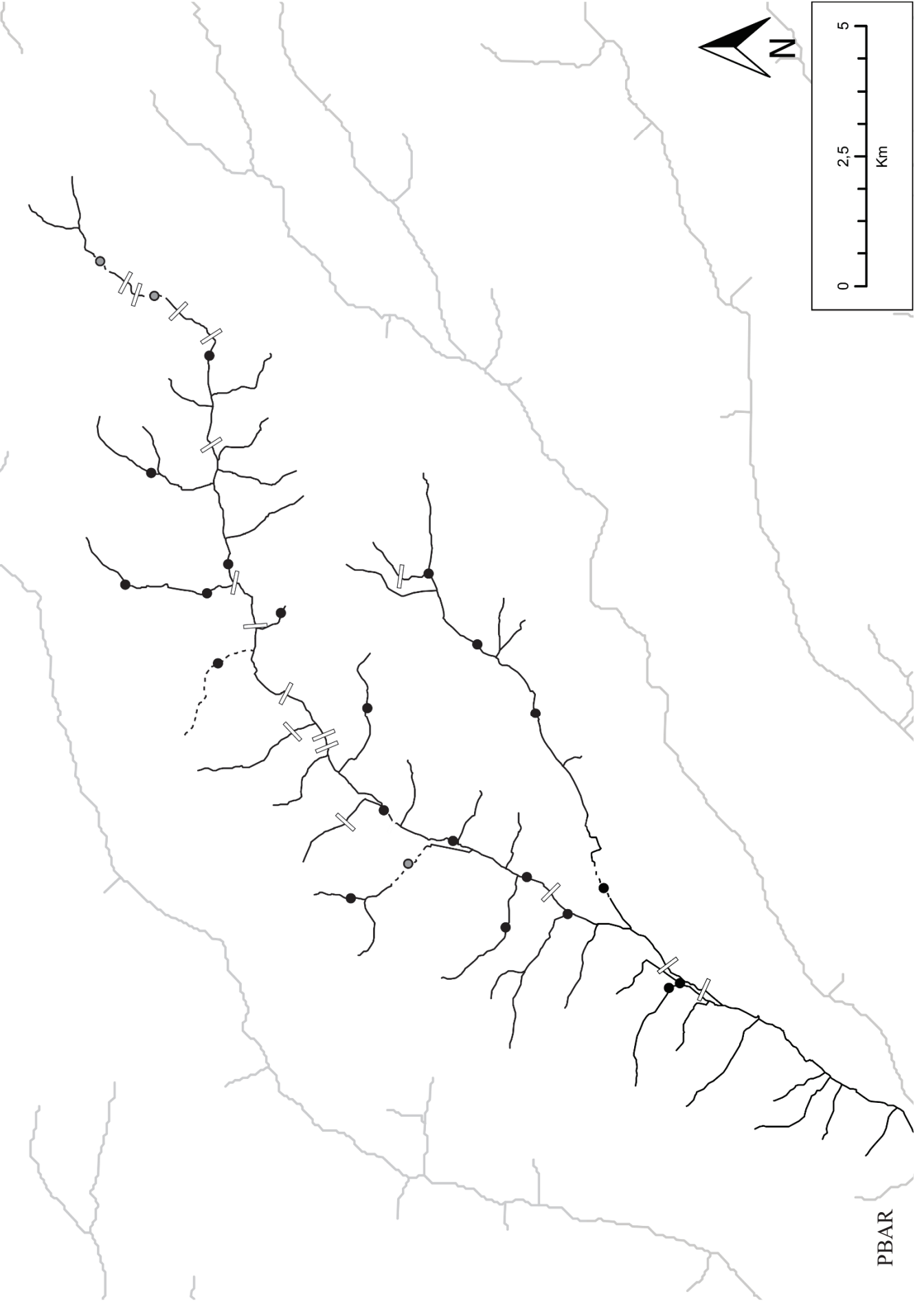
2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

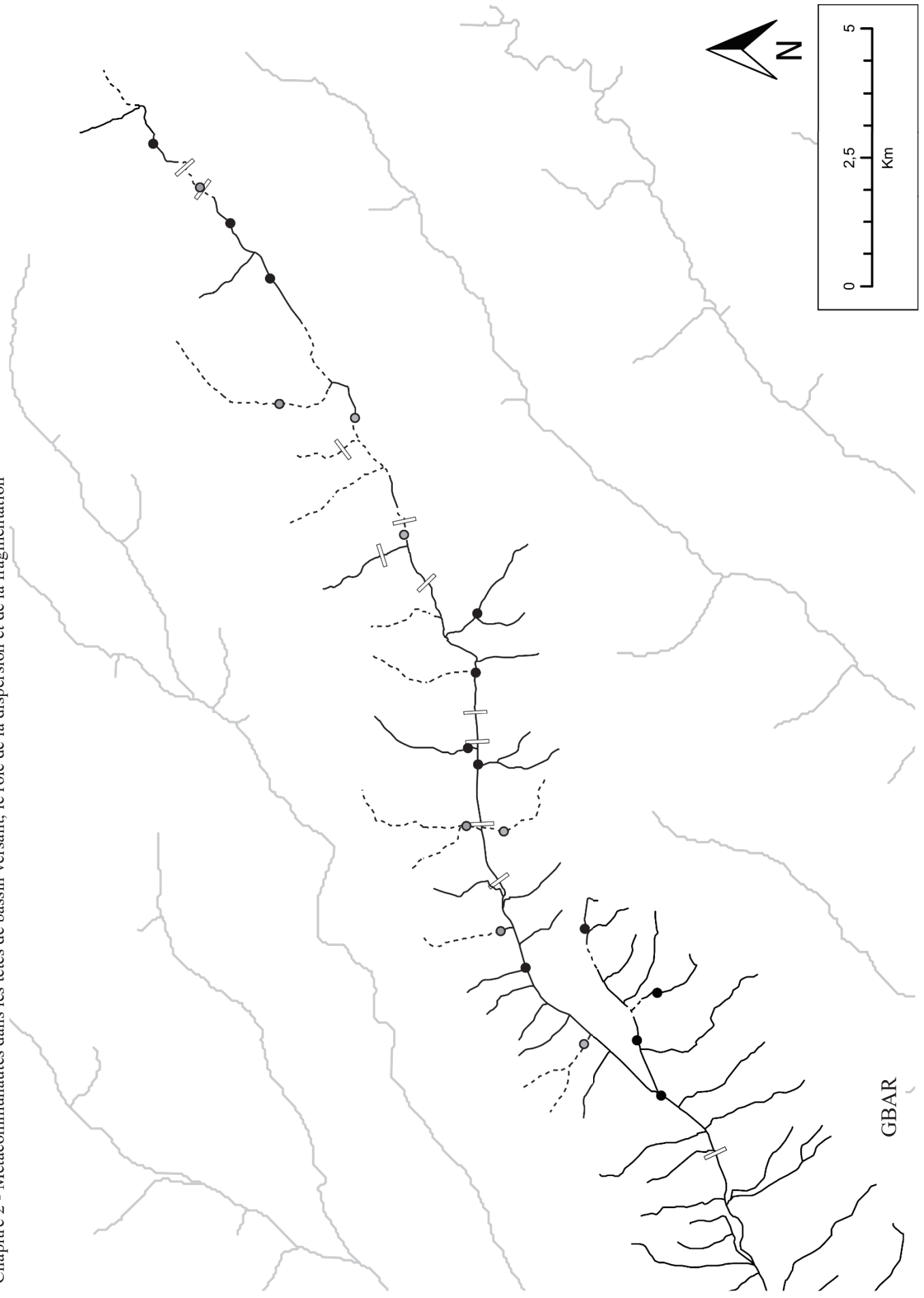


CEOU

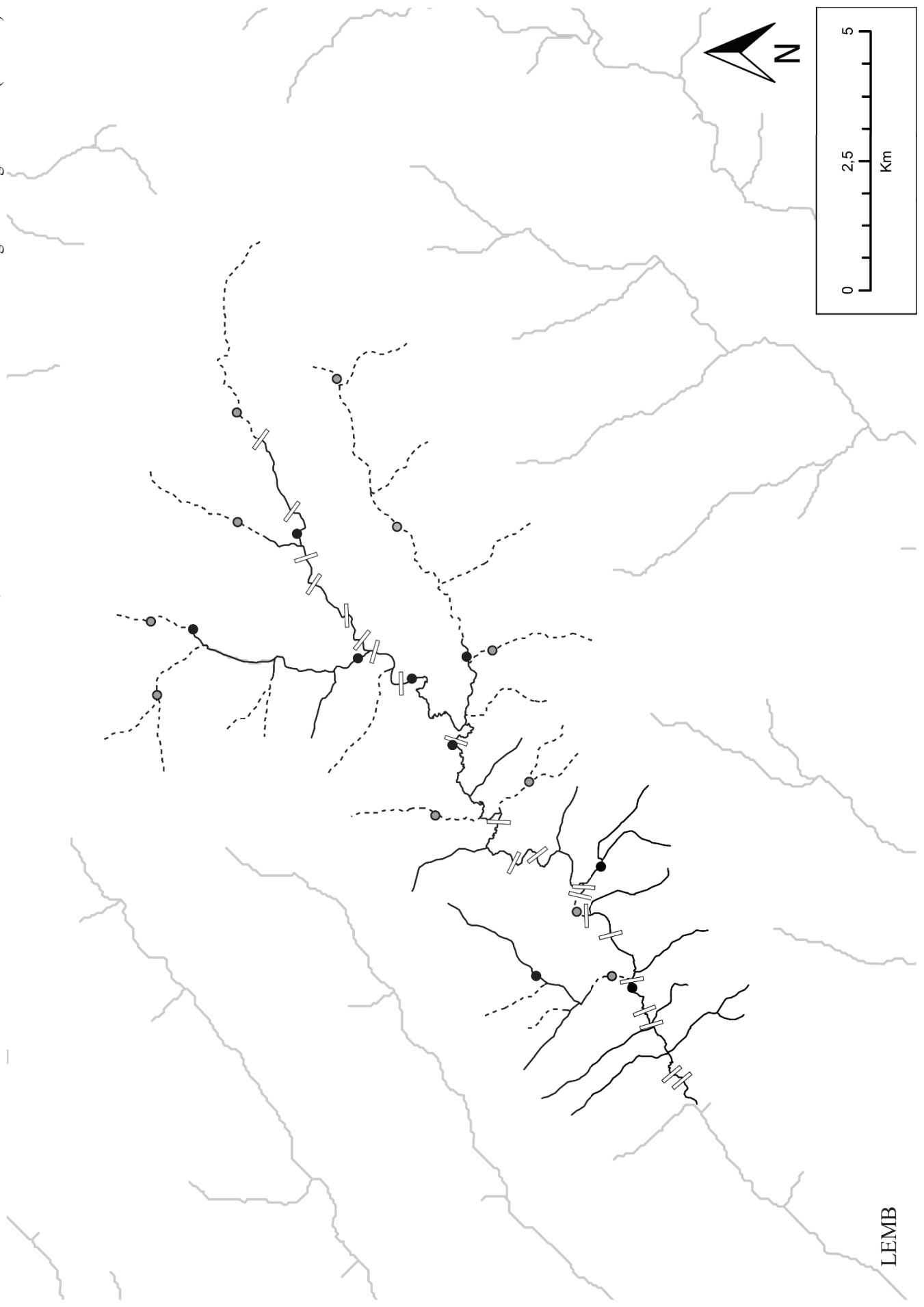


2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)





2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)



LEMB

Appendix S2. Supplementary analyses to test for the reliability of the study sampling design to quantify metacommunity processes

Table S2.1. Correlation coefficient r of the Mantel tests based on Chao dissimilarity index on a dataset composed of 8 river networks, each containing 6 sampled reaches. For more details, see Crabot et al, 2019. Results are shown for the two replicates and for their sum (a+b). With respect to distances: ENV, environmental; EUC, overland; Significant relationships are shown in bold (p -value ≤ 0.05 *). This indicates that inferences from either each individual or combined samples are identical, supporting our choice to put a stronger sampling effort to allow between-metacommunity comparisons than to produce a very detailed analyses of a single metacommunity, ignoring the context dependence warning from recent studies (Brown et al, 2011; Tonkin et al, 2016; Viana et al, 2008).

Catchement name	Distance matrices	Samples		
		a	b	a + b
Aigue brun	ENV	-0.162	0.133	-0.029
	EUC	0.754	0.146	0.555
Petit buech	ENV	0.553	0.515	0.557
	EUC	0.434*	0.557*	0.610*
Cèze	ENV	0.044	0.291	0.241
	EUC	0.633*	0.848*	0.893*
Clauge	ENV	-0.032	0.081	0.081
	EUC	0.728	0.841	0.829
Lez	ENV	-0.539	-0.244	-0.265
	EUC	0.018	0.301	0.272
Roubion	ENV	0.208	0.175	-0.080
	EUC	0.499	0.222	0.539
Seguissous	ENV	0.500	0.538	0.541
	EUC	0.666	0.675	0.438
Toulourenc	ENV	0.122	0.285	0.255
	EUC	0.272	-0.005	0.072

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Table S2.2. Coefficient of correlation of the Mantel tests based on Chao dissimilarity index on a dataset composed of one river network (Suran) containing 7 sampled reaches. Twenty samples were sampled per reach. Results are shown for the four riffle replicates taken separately, their cumulative sum (a+b, a+b+c, a+b+c+d) and the total of all twenty samples (total20). With respect to distances: ENV, environmental; EUC, overland; FRAG, number of anthropogenic structures. Significant relationships are shown in bold (p-value ≤ 0.05 *). This, as A5, indicates that increasing the sampling effort at the reach scale does not alter the conclusions of our study.

Distance matrices	Samples							
	a	b	c	d	a+b	a+b+c	a+b+c+d	total20
ENV	0.297	-0.309	0.078	0.247	0.132	0.169	0.192	0.188
EUC	-0.304	0.322	-0.028	0.189	0.076	0.203	0.328	0.388
FRA	-0.306	0.303	-0.114	0.148	0.053	0.164	0.278	0.301

Appendix S3. Identification taxonomic level to which invertebrates orders were identified in the study, listed alphabetically by order within higher rank group (Ph, Phylum; Cl, class; sCl, Subclass).

Higher rank	Order	Identification level
Acari (sCl)	Hydrachnidia	Order
Bivalvia (sCl)		Species
Cladocera (sCl)		Order
Copepoda (sCl)		Order
Gastropoda(Cl)		Species
Achaeta (Cl)		Genus
Hydrozoa (Cl)		Order
	Coleoptera	Genus
	Diptera	Family/Genus
	Ephemeroptera	Species
	Hemiptera	Genus
Insecta (Cl)	Hymenoptera	Species
	Megaloptera	Species
	Odonata	Species
	Plecoptera	Species
	Trichoptera	Genus/Species
	Amphipoda	Genus
Malacostraca (Cl)	Decapoda	Species
	Isopoda	Species
Nematoda (Ph)		Order
Oligochaeta (sCl)		Subclass
Ostracoda (sCl)		Order
Platyhelminthes (Ph)	Tricladida	Genus/Species

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Appendix S4. Mean abundances for each taxon and FHW

Table S4.1. Mean value of abundance for each taxa and each FHW for the sampling period before the drying season. Higher rank group and mean specific richness are also provided.

				Intermittent headwater									
				ALBA	CEOU	GBAR	LEMB	OIGN	PBAR	SERA	SURA	VERS	
Achaeta	Achaeta	Achaeta	Achaeta	0	0	0	0.05	0	0	0	0	0	
Mollusca	Gastropoda	Acroloxidae	<i>Acroloxus lacustris</i>	0	0	0	0	0.29	0	0	0	0	
Insecta	Trichoptera	Leptoceridae	<i>Adicella reducta</i>	0	0	0	0.05	0	0	0	0	0	
Insecta	Odonata	Aeshnidae	Aeshnidae	0	0	0.05	0	0	0	0	0	0	
Insecta	Coleoptera	Dytiscidae	<i>Agabus sp</i>	0.1	0	0.11	0	0.06	0.09	0	0	0.05	
Insecta	Trichoptera	Glossosomatidae	Agapetinae	0	0	0	0.05	0	0	0	0	0	
Insecta	Trichoptera	Hydroptilidae	<i>Agraylea sp</i>	0	0	0	0	0	0	0	0.05	0	
Insecta	Hymenoptera	Icheumonidae	<i>Agriotypus armatus</i>	0	0	0.11	0	0	0.09	0	0	0.05	
Insecta	Ephemeroptera	Baetidae	<i>Alainites muticus</i>	1.7	0.31	0.84	0	5.82	0.13	2.21	0.05	1.37	
Insecta	Trichoptera	Limnephilidae	<i>Allogamus sp</i>	0.1	0	0	0	0	0	0.58	0	0	
Insecta	Plecoptera	Nemouridae	<i>Amphinemura sp</i>	0.4	0.25	0.05	0.18	0	0	0	0	2.32	
Insecta	Trichoptera	Limnephilidae	<i>Anobolia nervosa</i>	0.2	0	0	0	0	0	0	0	0	
Mollusca	Gastropoda	Planorbidae	<i>Ancylus fluviatilis</i>	6.6	35.44	1.68	38.09	8.24	5	0	4.35	20.11	
Mollusca	Gastropoda	Planorbidae	<i>Anisus spirorbis</i>	0	2.81	12	16	0	24.35	0	0	0	
Insecta	Diptera	Culicidae	<i>Anopheles sp</i>	0	0	0.05	0	0	0	0	0	0	
Insecta	Diptera	Anthomyiidae	Anthomyiidae	0	0	0	0	0.12	0	0	0	0	
Insecta	Diptera	Limoniidae	<i>Antocha vitripennis</i>	0	0	1.84	0	0.12	0.09	0.05	9.15	0	
Insecta	Hemiptera	Aphelocheiridae	<i>Aphelocheirus aestivalis</i>	0	0	0	0	0	0	0	0.05	0	
Malacostraca	Isopoda	Asellidae	<i>Asellus aquaticus</i>	0.7	0	0	0.5	0.06	0	0	0	0	
Insecta	Diptera	Athericidae	<i>Atherix sp</i>	5.2	1.06	0.16	0.09	1.76	0.22	0.79	0.15	1	
Insecta	Diptera	Athericidae	<i>Athericops crassipes</i>	0	0	0.74	0.09	0	0.22	0.05	0	0	
Insecta	Trichoptera	Leptoceridae	<i>Athripsodes sp</i>	0	0	1.05	0.18	0	0.74	0	0	0	
Insecta	Diptera	Ceratopogonidae	<i>Atrichopogon sp</i>	0	0	0.05	0	0	0	1.42	0	0	
Insecta	Ephemeroptera	Baetidae	<i>Baetis gr alpinus</i>	0	0	0	0	4.18	0	0	0	0	
Insecta	Ephemeroptera	Baetidae	<i>Baetis sp</i>	23.7	86.44	46.05	54.23	84.12	67.48	45.74	62.45	31.95	
Achaeta	Rhynchobdellida	Glossiphoniidae	<i>Batracobdella algira</i>	0	0	0	0	0.29	0	0	0	0	
Mollusca	Gastropoda	Amnicolidae	<i>Belgrandia sp</i>	0	0	0	0.14	0	0.3	0	0	0.68	
Insecta	Trichoptera	Beraeidae	<i>Beraea pullata</i>	0.1	0	0	0.05	0	0.04	0	0	0.05	
Insecta	Odonata	Aeshnidae	<i>Boyeria irene</i>	0	0	0	0.05	0	0	0	0	0	
Insecta	Trichoptera	Brachycentridae	<i>Brachycentrus sp</i>	0	0.06	0	0	0	0	0	0	0	
Mollusca	Gastropoda	Amnicolidae	<i>Bythinella sp</i>	0	0	0	0.14	0	0	0	0	11.58	
Mollusca	Gastropoda	Amnicolidae	<i>Bythiospeum sp</i>	0	0.12	0	0	0.06	0	0	0	0	
Insecta	Ephemeroptera	Caenidae	<i>Caenis horaria</i>	0	0	0	0	3.76	0	0	0	0	
Insecta	Ephemeroptera	Caenidae	<i>Caenis luctuosa</i>	0	0	0.05	0.09	0	0	0	0.1	0	
Insecta	Ephemeroptera	Caenidae	<i>Caenis macrura</i>	0	0	0.63	0.09	0	0	0	8.6	0	
Insecta	Ephemeroptera	Caenidae	<i>Caenis rivulorum</i>	0	0	0	0	0.65	0	0	0	0	
Insecta	Odonata	Calopterygidae	<i>Calopteryx sp</i>	0	0	0	0	0	0.04	0	0	0	
Insecta	Ephemeroptera	Baetidae	<i>Centroptilum luteolum</i>	0.8	0	0.32	0.09	0.29	0	0.05	0.15	0.11	
Insecta	Diptera	Ceratopogonidae	Ceratopogoninae	3.9	1.12	2.53	1.45	4.53	0.83	0.68	0.7	2.16	

Chapitre 2 - Métacommunautés dans les têtes de bassin versant, le rôle de la dispersion et de la fragmentation

Insecta	Trichoptera	Limnephilidae	<i>Chaetopteryx sp</i>	0.9	0	0.37	0	0.12	0.74	0	0.1	0.05
Insecta	Trichoptera	Hydropsychidae	<i>Cheumatopsyche lepida</i>	0	0	0	0	0	0	0	0.15	0
Insecta	Diptera	Chironomidae	<i>Chironomini</i>	0.4	0.19	1	1	1.24	1.57	1.42	2.55	0.89
Insecta	Diptera	Empididae	Clinocerinae	0	0	0	0	0.47	0.04	0.21	0.15	0.16
Insecta	Coleoptera	Dytiscidae	Colymbetinae	0	0	0.05	0	0	0.04	0	0	0
Insecta	Odonata	Cordulegastridae	<i>Cordulegaster sp</i>	0	0	0.89	0	0.12	0.13	0	0	0.11
Malacostraca	Cyclopoïda	Cyclopoïda	Cyclopoïda	0	0.06	0.11	0.32	0.18	0	0	1.15	0
Insecta	Diptera	Ceratopogonidae	Dasyheleinae	0	0	0.05	0	0.06	0	0	0	0.05
Insecta	Coleoptera	Dytiscidae	<i>Deronectes sp</i>	0	0	0	0.05	0	0	0	0	0
Insecta	Diptera	Limoniidae	<i>Dicranota sp</i>	0.7	0.94	1.37	1.18	0.12	0.87	7.26	0.65	0.21
Insecta	Plecoptera	Perlidae	<i>Dinocras cephalotes</i>	0	0	0	0	0.24	0	0	0	0
Insecta	Diptera	Dixidae	<i>Dixa sp</i>	0	0.06	0.32	0.64	0	0.22	0	0	0
Insecta	Diptera	Dolichopodidae	Dolichopodidae	0	0.06	0	0.05	0	0	0.05	0	0.05
Insecta	Trichoptera	Limnephilidae	<i>Drusus sp</i>	3.5	0	0	0	1.59	0	0.47	2.45	0
Insecta	Coleoptera	Dryopidae	<i>Dryops sp</i>	0	0.19	0.63	0.14	0	0	0	0	0
Platyhelminthes	Tricladida	Dugesidae	<i>Dugesia sp</i>	2.8	0	0	0	2.82	0.39	0.11	0.1	0
Platyhelminthes	Tricladida	Dugesidae	<i>Dugesia tigrina</i>	0	0	0.21	0	0	2.35	0	0	0
Insecta	Ephemeroptera	Heptageniidae	<i>Ecdyonurus sp</i>	4.6	0.75	2.32	0	0.65	0	2.37	0.8	0
Malacostraca	Amphipoda	Eulimnogammaridae	<i>Echinogammarus sp</i>	0	0	35.21	16.32	0	31.35	0	0	0
Insecta	Ephemeroptera	Heptageniidae	<i>Electrogena sp</i>	0	0	0	0	0	0	0.11	0	0
Insecta	Coleoptera	Elmidae	<i>Elmis sp</i>	30.8	1.94	16.16	4.09	14.94	12.39	2.84	25.5	15.16
Insecta	Coleoptera	Scirtidae	<i>Elodes sp</i>	0.3	0	0.11	0	0	0	0	0	0.74
Insecta	Ephemeroptera	Heptageniidae	<i>Epeorus assimilis</i>	0	0	0	0	0.12	0	0	0	0
Insecta	Ephemeroptera	Ephemeridae	<i>Ephemera danica</i>	0	0	0.74	0.41	0	0	0	0	0
Insecta	Ephemeroptera	Ephemeridae	<i>Ephemera sp</i>	0	0	0	0	0	0.22	0	0	0
Insecta	Diptera	Ephydriidae	<i>Ephydriidae</i>	0	0	0	0	0.06	0	0	0	0
Insecta	Diptera	Limoniidae	<i>Eriopterini</i>	0.3	0	0	0	0.18	0.17	0.05	0	0
Achaeta	Arynchobdellida	Erpobdellidae	<i>Erpobdella sp</i>	1.1	0.75	0.58	0.14	1.18	0	0.21	0.05	1
Insecta	Coleoptera	Elmidae	<i>Esolus sp</i>	4.4	4.31	23.58	2.64	2.71	23.26	23.53	53.1	16
Insecta	Diptera	Ceratopogonidae	Forcipomyinae	0.2	0.06	0	0.05	0	0	0	0	0
Mollusca	Gastropoda	Lymnaeidae	<i>Galba truncatula</i>	0.2	2.94	8.37	61.68	0.06	18.09	0.26	0.5	0.11
Malacostraca	Amphipoda	Gammaridae	Gammaridae	0	0	0	0.23	0	0	0	0	0
Malacostraca	Amphipoda	Gammaridae	<i>Gammarus sp</i>	460.7	15.25	68.68	20.86	120.88	43.3	234.21	772.65	164.26
Insecta	Hemiptera	Gerridae	<i>Gerris sp</i>	0	0	0	0	0	0	0.16	0.05	0
Achaeta	Rhynchobdellida	Glossiphoniidae	<i>Glossiphonia sp</i>	4.8	0	2.21	0.23	0.65	1	0	0	0.79
Insecta	Trichoptera	Glossosomatidae	Glossosomatidae	0	16.62	0.11	0	0.06	0	0	0	0
Insecta	Trichoptera	Goeridae	<i>Goera pilosa</i>	0	0	1.95	0	0	0	0	0	0
Insecta	Trichoptera	Goeridae	Goeridae	0	0	0	1.45	0	0	0	0	0.63
Insecta	Odonata	Gomphidae	Gomphidae	0	0	1	0.41	0	0	0	0	0
Nematoda	Gordea	Gordiidae	<i>Gordius sp</i>	0.1	0	0	0	0	0	0	0	0
Mollusca	Gastropoda	Planorbidae	<i>Gyraulus albus</i>	0.6	0	0	0	0.06	0	0	0.05	0
Mollusca	Gastropoda	Planorbidae	<i>Gyraulus sp</i>	0	0	0	0.36	0	0	0	0	0
Insecta	Ephemeroptera	Leptophlebiidae	<i>Habropleptoides sp</i>	13.2	0	0	0	2.24	0	0.68	1.15	0
Insecta	Ephemeroptera	Leptophlebiidae	<i>Habroplebia fusca</i>	10.3	0	0	0	1.12	0	0	5.65	0
Insecta	Ephemeroptera	Leptophlebiidae	<i>Habroplebia lauta</i>	10.4	5.44	0	0	3.12	4.3	5.26	3.1	9.37
Achaeta	Arynchobdellida	Haemopidae	<i>Haemopsis sanguinosa</i>	0	0	0	0	0.18	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Halesus sp</i>	0	0	0	0	0	0.04	0.05	0	0
Insecta	Coleoptera	Haliplidae	<i>Halipilus sp</i>	0	0.06	0.53	0.14	0.06	0	0	0	0

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Mollusca	Gastropoda	Amnicolidae	<i>Hauffenia sp</i>	0	0	0.26	0	0	0	0	0	0.37
Achaeta	Rhynchobdellida	Glossiphoniidae	<i>Helobdella stagnalis</i>	0.1	0	0	0	0.06	0.13	0	0	0
Insecta	Coleoptera	Helophoridae	<i>Helophorus sp</i>	0.1	0.31	0.21	0	0	0	0.05	0.1	0.05
Insecta	Diptera	Empididae	Hemerodromiinae	0	1.75	0.53	0.05	0.76	0.35	0.26	3.15	0.42
Insecta	Ephemeroptera	Heptageniidae	Heptageniidae	0	0	0	4	0	0.52	0	0	0
Insecta	Diptera	Limoniidae	<i>Hexatomini</i>	0	0	0.11	0	0	0	0	0	0
Mollusca	Gastropoda	Planorbidae	<i>Hippeutis complanatus</i>	0	0	0	0.05	0	0	0	0	0
Acari	Acari	Acari	<i>Hydracarina</i>	2.1	3.75	3.42	2.05	4.82	3.09	2.79	18.85	3.89
Insecta	Coleoptera	Hydraenidae	<i>Hydraena sp</i>	3.5	0.25	1.05	0.32	2.12	0.13	7.53	5.15	4.21
Insecta	Coleoptera	Hydrophilidae	<i>Hydrochara sp</i>	0	0	0	0	0	0.04	0	0	0
Insecta	Coleoptera	Hydrochidae	<i>Hydrochus sp</i>	0	0	0.32	0	0	0	0	0	0
Insecta	Coleoptera	Scirtidae	<i>Hydrocyphon sp</i>	0	0	0.16	0.05	4.94	0.65	0.47	3.7	2.63
Insecta	Hemiptera	Hydrometridae	<i>Hydrometra sp</i>	0	0	0	0	0	0	0.32	0	0
Insecta	Coleoptera	Hydrophilidae	Hydrophilinae	0	0	0	0	0	0	0.05	0.35	0
Insecta	Coleoptera	Dytiscidae	Hydroporinae	1.3	0.88	0.05	0	0	0	0	0	1.63
Insecta	Coleoptera	Dytiscidae	<i>Hydroporus sp</i>	0	0	0	0	0	0.04	0	0	0
Insecta	Trichoptera	Hydropsychidae	<i>Hydropsyche angustipennis</i>	0	0	0.89	0	0	0.35	0	0	0
Insecta	Trichoptera	Hydropsychidae	<i>Hydropsyche bulbifera</i>	0	0	0.63	0	0	0.43	0	0	0
Insecta	Trichoptera	Hydropsychidae	<i>Hydropsyche siltalai</i>	0	0	0.26	0	0.24	0.57	0	2.3	0
Insecta	Trichoptera	Hydropsychidae	<i>Hydropsyche sp</i>	0	0.12	0	0.05	4.06	0	0.11	17	0
Insecta	Trichoptera	Hydroptilidae	<i>Hydroptila sp</i>	0	0.44	1.42	0.05	0	0.13	0.11	3.7	0.05
Insecta	Trichoptera	Hydroptilidae	<i>Hydroptilidae</i>	0	0	0	0	0.29	0	0	0	0
Hydrozoa	Hydrozoa	Hydrozoa	Hydrozoa	0.2	0	0	1.23	0	0.26	0.58	0.1	0
Insecta	Trichoptera	Rhyacophilidae	<i>Hyporhyacophila sp</i>	0	0	0	0	0.24	0	0.95	0	0
Insecta	Plecoptera	Perlodidae	<i>Isoperla grammatica</i>	0	0	0	0	0.06	0	0	0	0
Insecta	Plecoptera	Perlodidae	<i>Isoperla sp</i>	0.1	6.19	2	1.23	0.06	0.35	2.21	0	3.05
Insecta	Coleoptera	Hydrophilidae	<i>Laccobius sp</i>	0	0	0	0	0	0.04	0	0	0
Insecta	Coleoptera	Dytiscidae	<i>Laccophilus sp</i>	0	0	0.05	0.05	0	0	0	0	0
Insecta	Trichoptera	Leptoceridae	Leptoceridae	0	0	0	0	0	0	0	1.1	0
Insecta	Ephemeroptera	Leptophlebiidae	Leptophlebiidae	0	0	8.21	10.36	0	0	0	0	0
Insecta	Plecoptera	Leuctridae	<i>Leuctra geniculata</i>	0	0.56	0.89	0.14	0	16.43	0	0.2	0
Insecta	Plecoptera	Leuctridae	<i>Leuctra sp</i>	4.9	0.94	1.37	0.05	7.94	2.78	7.58	5.95	0.26
Insecta	Trichoptera	Limnephilidae	Limnephilidae	0	0	0	0.09	0	0	0.05	0	0
Insecta	Trichoptera	Limnephilidae	<i>Limnephilus sp</i>	0	0.12	0.11	0	4.06	0	0	0	0
Insecta	Coleoptera	Elmidae	<i>Limnius sp</i>	3.4	1.19	4.95	0.73	2.59	3.96	2	24.7	1.79
Insecta	Diptera	Limoniidae	Limoniidae	0	0.06	0	0	0	0	0	0	0
Mollusca	Gastropoda	Lymnaeidae	Lymnaeidae	0	0	0	0.23	0	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Melampophylax mucoreus</i>	0	0	0	0	0	0	0	2	0
Insecta	Trichoptera	Limnephilidae	<i>Melampophylax sp</i>	3.9	0	0	0	3.06	0	0.79	0	0
Insecta	Hemiptera	Corixidae	<i>Micronecta sp</i>	0	0	0	0	0.18	0	0.11	0.2	0
Insecta	Trichoptera	Limnephilidae	<i>Micropterna sp</i>	0	0.25	0	0	0.12	0	0	0	0.47
Insecta	Trichoptera	Leptoceridae	<i>Mystacides sp</i>	0	0	0	0.05	0	0	0	0	0
Nematoda	Nematoda	Nematoda	Nematoda	1	2.06	0.42	0.68	1.71	0.43	1.05	0.5	1.05
Insecta	Plecoptera	Nemouridae	<i>Nemoura marginata</i>	0	0	0	0	0.06	0	0	0	0
Insecta	Plecoptera	Nemouridae	<i>Nemoura sp</i>	2.9	2.12	0.05	0	0	0	1.32	3.75	1.32
Malacostraca	Amphipoda	Niphargidae	<i>Niphargus sp</i>	0	0.25	0.05	3.27	0	0.17	0	0	0.16
Insecta	Trichoptera	Odontoceridae	<i>Odontocerum albicorne</i>	0.4	0	0	0.05	0.53	0	1.42	1.55	0
Oligochaeta	Oligochaeta	Oligochaeta	Oligochaeta	57.4	6.25	21.68	10.95	30.65	13	16.32	11.95	5.21

Chapitre 2 - Métacommunautés dans les têtes de bassin versant, le rôle de la dispersion et de la fragmentation

Insecta	Odonata	Gomphidae	<i>Onychogomphus sp</i>	0	0	0	0	0	1.13	0	0	0
Insecta	Coleoptera	Gyrinidae	<i>Orectochilus villosus</i>	0	0	0	0	0	0	0	0	0.05
Insecta	Coleoptera	Dytiscidae	<i>Oreodytes sp</i>	0	0	0	0	0.06	0	0	0	0
Insecta	Diptera	Chironomidae	Orthocladinae	9.2	29.56	11.42	2.64	27.47	1.96	23.37	12.9	16.84
Insecta	Neuroptera	Osmylidae	<i>Osmylus fulvicephalus</i>	0	0	0	0	0.06	0	0	0	0
Malacostraca	Ostracoda	Ostracoda	Ostracoda	1.3	1	2.05	1.09	0.35	3.17	1	2.3	0.47
Insecta	Coleoptera	Elmidae	<i>Oulimnius sp</i>	7.8	0.81	15.79	44.41	10.12	4.26	0.11	4.7	0
Insecta	Diptera	Limoniidae	<i>Pedicia sp</i>	0.4	0	0	0	0	0	0	0.1	0
Insecta	Plecoptera	Perlidae	<i>Perla sp</i>	0	0	0	0	0.06	0	0	0	0
Insecta	Plecoptera	Perlodidae	<i>Perlodes jurassicus</i>	0.4	0	0	0	0.65	0	0	0.65	0
Insecta	Trichoptera	Philopotamidae	Philopotamidae	0	0	0	0	0	0.26	0	0	0
Mollusca	Gastropoda	Physidae	<i>Physa acuta</i>	0	0	0	0.27	0	0.48	0	0	0
Achaeta	Rhynchobdellida	Glossiphoniidae	<i>Piscicola sp</i>	0	0	0	0	0	0	0	0.05	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium casertanum</i>	0	0	0	0	0	0	0	0.15	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium henslowanum</i>	0	0	0	0	0	0	0	3.75	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium hibernicum</i>	0	0	0	0	0	0.04	0	0	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium milium</i>	2	0	0	0	0	0.13	0	0	2.95
Mollusca	Venerida	Sphaeriidae	<i>Pisidium nitidum</i>	9.2	0	0.16	0	43.18	2.17	0	1.25	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium personatum</i>	3.8	0	3.21	2.18	0.12	2.7	2.21	0	0.37
Mollusca	Venerida	Sphaeriidae	<i>Pisidium subtruncatum</i>	0.4	0	3.89	0.45	0	0.74	0.05	0.95	1
Mollusca	Venerida	Sphaeriidae	<i>Pisidium tenuilineatum</i>	0	0	0	0	0	0	0	0.25	0
Insecta	Trichoptera	Polycentropodidae	<i>Plectrocnemia sp</i>	0.8	0.12	0	0	0.12	0	0	0.1	0
Platyhelminthes	Tricladida	Planariidae	<i>Polycelis felina</i>	0	0	0	0	0	0	0	0.5	0
Platyhelminthes	Tricladida	Planariidae	<i>Polycelis sp</i>	13	0	0.26	0.18	12.41	0.04	0.21	20.6	0
Insecta	Trichoptera	Polycentropodidae	<i>Polycentropodidae</i>	0	0	0.58	0.68	0	0	1.11	0	0.32
Insecta	Trichoptera	Limnephilidae	<i>Potamophylax sp</i>	0.1	0.06	0	0.05	0.71	0	0.21	0.3	0
Mollusca	Gastropoda	Amnicolidae	<i>Potamopyrgus antipodarum</i>	0	5.88	188.47	329.73	0	101	0.16	0	106.32
Malacostraca	Isopoda	Asellidae	<i>Proasellus meridianus</i>	0	0	1	0	0	1.35	0	0	0
Malacostraca	Isopoda	Asellidae	<i>Proasellus sp</i>	0	0	0	0.36	0	0	0	0	0
Insecta	Ephemeroptera	Baetidae	<i>Procloeon sp</i>	0.1	0.06	0	0	0	0	0	0	0
Insecta	Plecoptera	Nemouridae	<i>Protonemura sp</i>	1.1	0	0	0	7.53	0	0.68	0.05	0
Insecta	Diptera	Psychodidae	Psychodidae	0.2	0.12	0.16	0	0.18	0	0.84	0.05	0.05
Insecta	Trichoptera	Psychomyiidae	<i>Psychomyia pusilla</i>	0	0	0	0	0	0	0	0.2	0
Mollusca	Gastropoda	Lymnaeidae	<i>Radix labiata</i>	0	0	0	0	0	0	0	0	0.11
Mollusca	Gastropoda	Lymnaeidae	<i>Radix sp</i>	0.5	0.19	0.32	10.36	3.41	0.17	0.05	9.3	0
Insecta	Ephemeroptera	Heptageniidae	<i>Rhithrogena picteti</i>	0	0	0.32	0	0	0	0	0	0
Insecta	Ephemeroptera	Heptageniidae	<i>Rhithrogena sp</i>	4.4	0	0	0	7.76	0	8.58	8.35	2.68
Insecta	Trichoptera	Rhyacophilidae	<i>Rhyacophila sp</i>	0.3	1.88	0.89	0.95	3.65	0.48	5.79	3.05	2.11
Insecta	Coleoptera	Elmidae	<i>Riolus sp</i>	0.1	10.75	5.89	4.05	13.12	6.78	4.42	37.7	14.32
Insecta	Trichoptera	Sericostomatidae	<i>Sericostoma sp</i>	0.7	0	1.32	0	0.12	0.91	0	0	0.95
Insecta	Ephemeroptera	Ephemerellidae	<i>Serratella ignita</i>	0.1	62.06	15.47	18.68	24.59	25.65	3.79	27.3	2.05
Insecta	Megaloptera	Sialidae	<i>Sialis sp</i>	0.1	0	0.05	0.23	0	0	0.05	0	0
Insecta	Trichoptera	Goeridae	<i>Silo nigricornis</i>	0	0	0	0	0.59	0	0	0	0
Insecta	Trichoptera	Goeridae	<i>Silo sp</i>	0.2	0.19	0.05	0	0	0.91	0	0.2	0
Insecta	Diptera	Simuliidae	Simuliidae	7	34.5	4.37	26.64	13.29	2.83	17.84	2.95	1.26
Insecta	Ephemeroptera	Siphonuridae	<i>Siphonurus aestivalis</i>	0	0.06	0	0	0	0	0	0	0
Insecta	Ephemeroptera	Siphonuridae	<i>Siphonurus sp</i>	0	0	0	0	0	0	0	0	0.16
Mollusca	Venerida	Sphaeriidae	<i>Sphaerium corneum</i>	0	0	0	0.09	0	0	0	0	0

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Mollusca	Gastropoda	Lymnaeidae	<i>Stagnicola sp</i>	0	0	0.05	0	0	0	0	0	0
Insecta	Coleoptera	Elmidae	<i>Stenelmis sp</i>	0	0	0.11	0	0	0	0	0.2	0
Insecta	Trichoptera	Limnephilidae	<i>Stenophylax sp</i>	0	0	0.05	0	0	0	0	0	0
Insecta	Diptera	Stratiomyidae	Stratiomyidae	0.2	0.25	0.16	0.14	0.06	0.22	0.16	0.1	0.16
Insecta	Trichoptera	Glossosomatidae	<i>Synagapetus sp</i>	0	0	0	0	0	2.57	0.11	0.1	17.84
Insecta	Diptera	Tabanidae	Tabanidae	0.3	0	0.47	0	0.47	0.3	0	0.05	0.16
Insecta	Diptera	Chironomidae	Tanypodinae	7.2	0.44	1.63	0.95	9.41	2.26	7.74	9.45	1.37
Insecta	Diptera	Chironomidae	Tanytarsini	4.3	3.12	21.26	11.55	31.53	7	42.11	16.8	12.05
Insecta	Trichoptera	Psychomyiidae	<i>Tinodes sp</i>	0.5	0.25	3.53	0.5	1.59	2.22	4.42	0.85	3.32
Insecta	Diptera	Tipulidae	Tipulidae	0	0.06	0	0.18	0	0	0	0	0
Insecta	Ephemeroptera	Ephemerellidae	<i>Torleya major</i>	0	0	0	0	0	0	0	0.15	0
Platyhelminthes	Tricladida	Tricladida	Tricladida	0	0	0.05	0	0	0	0	0	0
Mollusca	Unionida	Margaritiferidae	<i>Unio crassus</i>	0	0	0	0	0	0	0	0.05	0
Mollusca	Gastropoda	Valvatidae	<i>Valvata cristata</i>	0.8	0	0	0	0	0	0	0	0
Mollusca	Gastropoda	Valvatidae	<i>Valvata piscinalis</i>	0	0	0	0	0	0	0	0.05	0
Insecta	Hemiptera	Veliidae	<i>Velia sp</i>	0	0	0	0	0.06	0	0	0	0
Insecta	Hemiptera	Veliidae	Veliidae	0	0.19	0	0	0	0	0.11	0.05	0
Insecta	Trichoptera	Philopotamidae	<i>Wormaldia sp</i>	0	0	0	0.05	0	0	0	0	0.11
Specific richness				24.1	15.25	23.32	18.41	25.53	20.7	23.32	25.1	19.74

Table S4.2. Mean value of abundance for each taxa and each FWH for the sampling period after the drying season. Higher rank group and mean specific richness are also provided.

				ALBA	CEOU	CLAU	GBAR	LEMB	OIGN	PBAR	SERA	SURA	VERS
Hirudinae	Hirudinae	Hirudinae	Hirudinae	0	0	0.03	0	0	0	0	0	0	0
Mollusca	Gastropoda	Acroloxidae	<i>Acroloxus lacustris</i>	0	0	0.07	0	0	0.04	0	0	0.05	0
Insecta	Trichoptera	Leptoceridae	<i>Adicella reducta</i>	0	0	0	0.05	0.05	0	0	0	0	0
Insecta	Coleoptera	Dysticidae	<i>Agabus sp</i>	0.1	0.3	0.1	0.21	0.41	0	0.09	0	0.05	0.14
Insecta	Trichoptera	Glossosomatidae	Agapetinae	0	0	0	0	0.09	0	0	0	0	0
Insecta	Trichoptera	Glossosomatidae	<i>Agapetus sp</i>	0	0	0.07	0	0	0	0	0	0	0.05
Insecta	Ephemeroptera	Baetidae	<i>Alainites muticus</i>	0.2	0	0	0.11	0	0.33	0.04	0.32	0	0.32
Insecta	Plecoptera	Nemouridae	<i>Amphinemura sp</i>	0.4	0.25	0.03	0.37	0.32	0.33	0.26	0	0	0.36
Mollusca	Gastropoda	Planorbidae	<i>Ancylus fluviatilis</i>	0.05	0.35	0.2	0.05	0.09	0.08	0.09	0	0.1	0.18
Mollusca	Gastropoda	Planorbidae	<i>Anisus spirorbis</i>	0	0.3	0	0.11	0.27	0	0.04	0	0	0.05
Insecta	Diptera	Anthomyiidae	Anthomyiidae	0	0	0	0	0	0	0.04	0	0	0
Insecta	Diptera	Limoniidae	<i>Antocha vitripennis</i>	0	0	0	0	0	0.12	0	0.05	0.33	0
Insecta	Hemiptera	Aphelocheiridae	<i>Aphelocheirus aestivalis</i>	0	0	0	0	0	0	0	0	0.05	0
Malacostraca	Isopoda	Asellidae	<i>Asellus aquaticus</i>	0.05	0	0	0	0.05	0	0	0	0.05	0
Insecta	Diptera	Athericidae	<i>Atherix sp</i>	0.05	0.05	0.03	0.11	0	0.21	0.17	0.05	0.1	0.18
Insecta	Diptera	Athericidae	<i>Athricops crassipes</i>	0	0	0	0.11	0	0	0.13	0	0.14	0
Insecta	Trichoptera	Leptoceridae	<i>Athripsodes sp</i>	0	0.05	0.07	0.11	0	0	0.04	0	0	0
Insecta	Ephemeroptera	Baetidae	<i>Baetis gr alpinus</i>	0	0	0	0	0	0.21	0	0	0	0
Insecta	Ephemeroptera	Baetidae	<i>Baetis sp</i>	0.3	1	0	0.79	0.91	0.54	0.74	0.47	0.67	0.95
Mollusca	Gastropoda	Planorbidae	<i>Bathymphalus contortus</i>	0	0	0.07	0	0	0	0	0	0	0
Mollusca	Gastropoda	Amnicolidae	<i>Belgrandia sp</i>	0.05	0	0	0	0.27	0.04	0.22	0	0	0.05
Insecta	Trichoptera	Beraeidae	<i>Beraea pullata</i>	0.05	0	0	0	0	0	0	0.05	0	0
Insecta	Plecoptera	Taeniopterygidae	<i>Brachyptera risi</i>	0.7	0	0	0	0	0.92	0	0	0	0
Insecta	Plecoptera	Taeniopterygidae	<i>Brachyptera sp</i>	0	0.45	0.67	0.26	0.36	0	0	0.53	0.14	0.27
Mollusca	Gastropoda	Amnicolidae	<i>Bythinella sp</i>	0	0	0	0	0	0	0	0	0	0.14
Insecta	Ephemeroptera	Caenidae	<i>Caenis luctuosa</i>	0	0	0.03	0.05	0.05	0	0.04	0	0.33	0
Insecta	Ephemeroptera	Caenidae	<i>Caenis macrura</i>	0	0.05	0	0.16	0	0.04	0.13	0.11	0.14	0
Insecta	Ephemeroptera	Caenidae	<i>Caenis rivulorum</i>	0	0	0	0	0	0.17	0	0	0	0
Insecta	Ephemeroptera	Caenidae	<i>Caenis sp</i>	0.1	0	0	0	0	0	0	0	0	0
Insecta	Plecoptera	Capniidae	<i>Capnioneura nemouroïdes</i>	0.1	0	0	0	0	0.12	0	0	0	0
Insecta	Ephemeroptera	Baetidae	<i>Centroptilum luteolum</i>	0.05	0	0	0	0	0.04	0	0	0.05	0.05
Insecta	Diptera	Ceratopogonidae	Ceratopogoninae	0.45	0.6	0.4	0.42	0.5	0.29	0.43	0.37	0.24	0.36
Insecta	Diptera	Chironomidae	<i>Chironomini</i>	0.25	0.05	0.17	0.16	0.14	0.04	0.22	0.26	0.33	0.27
Malacostraca	Cladocera	Cladocera	Cladocera	0.05	0	0.03	0	0.05	0	0.26	0	0	0
Insecta	Diptera	Empididae	Clinocerinae	0	0	0	0	0	0.04	0.04	0	0	0
Insecta	Coleoptera	Dysticidae	Colymbetinae	0.05	0.05	0	0	0	0	0.04	0.05	0	0.05
Mollusca	Venerida	Cyrenidae	<i>Corbicula fluminalis</i>	0	0	0	0.05	0	0	0	0	0	0
Insecta	Odonata	Cordulegastridae	<i>Cordulegaster sp</i>	0	0	0	0.21	0	0	0.04	0	0	0.18
Insecta	Odonata	Corduliidae	Corduliidae	0	0	0	0	0.05	0	0	0	0	0
Insecta	Odonata	Corduliidae/libellulidae	Corduliidae/libellulidae	0	0	0	0	0	0	0.04	0	0	0
Platyhelminthes	Tricladida	Planariidae	<i>Crenobia alpina</i>	0	0	0	0	0	0.04	0	0	0	0
Malacostraca	Cyclopoïda	Cyclopoïda	Cyclopoïda	0.05	0	0.03	0.21	0.23	0.04	0.3	0.26	0.1	0
Insecta	Diptera	Ceratopogonidae	Dasyheleinae	0	0	0	0	0	0	0.09	0	0	0

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Platyhelminthes	Tricladida	Dendrocoelidae	<i>Dendrocoelum lacteum</i>	0	0	0.07	0	0	0	0	0	0	0
Insecta	Diptera	Chironomidae	Diamesinae	0.35	0	0	0	0	0.42	0	0	0	0
Insecta	Diptera	Limoniidae	<i>Dicranota sp</i>	0.2	0.05	0	0.21	0	0.08	0.13	0.42	0.1	0.05
Insecta	Plecoptera	Perlidae	<i>Dinocras cephalotes</i>	0	0	0	0	0	0.04	0	0	0	0
Insecta	Diptera	Diptera	Diptera	0.05	0	0	0	0	0	0	0	0	0
Insecta	Diptera	Dixidae	<i>Dixa sp</i>	0	0	0	0	0	0	0	0	0	0.05
Insecta	Diptera	Dolichopodidae	Dolichopodidae	0.05	0	0	0.16	0	0	0.04	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Drusus sp</i>	0.2	0	0	0	0	0.08	0	0	0	0
Insecta	Coleoptera	Dryopidae	<i>Dryops sp</i>	0	0.1	0.13	0.32	0.14	0	0.09	0.11	0.05	0
Platyhelminthes	Tricladida	Dugesidae	<i>Dugesia sp</i>	0.1	0	0.07	0.05	0.05	0.12	0.09	0	0	0
Platyhelminthes	Tricladida	Dugesidae	<i>Dugesia tigrina</i>	0	0	0	0	0	0	0.17	0	0	0
Insecta	Ephemeroptera	Heptageniidae	<i>Ecdyonurus sp</i>	0	0	0	0.11	0	0.04	0	0.05	0.05	0
Malacostraca	Amphipoda	Eulimnogammaridae	<i>Echinogammarus sp</i>	0	0	0	0.16	0.14	0	0.17	0	0	0
Insecta	Ephemeroptera	Heptageniidae	<i>Electrogena lateralis</i>	0.15	0.15	0	0	0	0	0	0	0	0.05
Insecta	Ephemeroptera	Heptageniidae	<i>Electrogena sp</i>	0	0	0	0	0	0.04	0	0	0	0
Insecta	Coleoptera	Elmidae	<i>Elmis sp</i>	0.5	0.15	0.13	0.63	0.27	0.42	0.61	0.42	0.81	0.5
Insecta	Coleoptera	Scirtidae	<i>Elodes sp</i>	0	0.05	0	0.05	0	0	0.04	0.05	0.05	0.05
Insecta	Coleoptera	Hydrophilidae	<i>Enochrus sp</i>	0	0	0	0	0	0	0.04	0	0	0
Insecta	Ephemeroptera	Heptageniidae	<i>Epeorus assimilis</i>	0	0	0	0	0	0.17	0	0	0	0
Insecta	Ephemeroptera	Ephemeridae	<i>Ephemera danica</i>	0	0	0	0.16	0.09	0	0	0	0.29	0
Insecta	Ephemeroptera	Ephemeridae	<i>Ephemera sp</i>	0.05	0.05	0.1	0	0	0	0.17	0.11	0	0.05
Insecta	Diptera	Limoniidae	<i>Eriopterini</i>	0.1	0	0.1	0	0.09	0.04	0.13	0	0.05	0
Hirudinae	Arynhobdellida	Erpobdellidae	<i>Erpobdella sp</i>	0.25	0.5	0.2	0.16	0.18	0.17	0	0	0.19	0.41
Insecta	Coleoptera	Elmidae	<i>Esolus sp</i>	0.2	0.3	0.1	0.32	0.32	0.25	0.43	0.42	0.71	0.5
Mollusca	Gastropoda	Lymnaeidae	<i>Galba truncatula</i>	0.15	0.45	0.03	0.16	0.45	0.21	0.13	0.32	0.19	0.09
Malacostraca	Amphipoda	Gammaridae	<i>Gammarus sp</i>	0.65	0.25	0.13	0.26	0.05	0.75	0.48	0.79	0.95	0.64
Hirudinae	Rhynchobdellida	Glossiphoniidae	<i>Glossiphonia sp</i>	0.05	0	0	0.11	0.09	0.04	0.04	0.05	0	0.09
Hirudinae	Rhynchobdellida	Glossiphoniidae	<i>Glossiphonia verrucata</i>	0	0.05	0	0	0	0.04	0	0	0.05	0.05
Insecta	Trichoptera	Glossosomatidae	<i>Glossosoma sp</i>	0	0	0	0	0	0	0	0	0.05	0
Insecta	Trichoptera	Glossosomatidae	Glossosomatidae	0	0.35	0	0.16	0	0.08	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Glyphotaelius pellucidus</i>	0	0	0.03	0	0	0	0	0	0	0
Insecta	Trichoptera	Goeridae	<i>Goera pilosa</i>	0	0	0	0	0	0	0	0	0.1	0
Insecta	Trichoptera	Goeridae	Goeridae	0	0	0	0	0	0	0	0	0	0.09
Insecta	Odonata	Gomphidae	Gomphidae	0	0	0	0	0.05	0	0	0	0.05	0.05
Nematoda	Gordea	Gordiidae	<i>Gordius sp</i>	0	0	0	0	0	0	0	0.05	0	0.09
Mollusca	Gastropoda	Planorbidae	<i>Gyraulus albus</i>	0	0	0.03	0	0	0	0	0	0.14	0
Insecta	Ephemeroptera	Leptophlebiidae	<i>Habroleptoides confusa</i>	0	0.05	0.13	0	0	0	0.09	0	0	0
Insecta	Ephemeroptera	Leptophlebiidae	<i>Habroleptoides sp</i>	0.3	0	0	0	0	0.29	0	0.47	0.19	0.36
Insecta	Ephemeroptera	Leptophlebiidae	<i>Habrophlebia fusca</i>	0.1	0	0	0	0	0.12	0	0	0	0
Insecta	Ephemeroptera	Leptophlebiidae	<i>Habrophlebia lauta</i>	0.2	0.25	0.1	0	0	0.38	0.09	0.42	0.24	0.27
Hirudinae	Arynhobdellida	Haemopidae	<i>Haemopsis sanguinosa</i>	0	0	0	0	0	0	0	0	0.05	0
Insecta	Trichoptera	Limnephilidae	<i>Halesus digitatus</i>	0	0	0	0.05	0	0	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Halesus radiatus</i>	0.05	0	0	0.05	0	0	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Halesus sp</i>	0	0	0	0	0.05	0	0.09	0	0	0
Insecta	Coleoptera	Halipidae	<i>Halipus sp</i>	0.05	0	0	0	0.05	0	0	0.05	0	0
Mollusca	Gastropoda	Amnicolidae	<i>Hauffenia sp</i>	0	0.1	0	0.11	0.14	0.04	0.09	0	0	0.09
Hirudinae	Rhynchobdellida	Glossiphoniidae	<i>Helobdella stagnalis</i>	0	0	0	0	0	0	0	0.05	0	0
Insecta	Coleoptera	Helophoridae	<i>Helophorus sp</i>	0.05	0	0	0	0	0	0	0	0	0

Chapitre 2 - Métacommunautés dans les têtes de bassin versant, le rôle de la dispersion et de la fragmentation

Insecta	Diptera	Empididae	Hemerodromiinae	0	0.15	0.07	0.21	0.09	0.04	0.04	0.11	0.19	0.27
Insecta	Ephemeroptera	Heptageniidae	Heptageniidae	0	0	0	0	0.09	0	0.04	0	0	0
Insecta	Diptera	Limoniidae	<i>Hexatoma sp</i>	0.1	0	0	0	0	0	0.04	0	0	0
Insecta	Diptera	Limoniidae	<i>Hexatomini</i>	0	0	0.13	0	0	0	0	0	0	0.05
Mollusca	Gastropoda	Planorbidae	<i>Hippeutis complanatus</i>	0	0	0.03	0	0.05	0	0	0	0.05	0
Acari	Hydracarina	Hydracarina	Hydracarina	0.3	0.45	0.4	0.58	0.68	0.25	0.74	0.58	0.62	0.77
Insecta	Coleoptera	Hydraenidae	<i>Hydraena sp</i>	0.35	0.05	0.13	0.05	0	0.25	0	0.63	0.67	0.18
Insecta	Coleoptera	Hydrochidae	<i>Hydrochus sp</i>	0	0	0	0.05	0	0	0	0	0	0
Insecta	Coleoptera	Scirtidae	<i>Hydrocyphon sp</i>	0	0	0	0	0	0.04	0	0	0.1	0
Insecta	Trichoptera	Hydropsychidae	<i>Hydropsyche siltalai</i>	0	0	0	0.11	0.09	0.21	0.09	0	0	0
Insecta	Trichoptera	Hydropsychidae	<i>Hydropsyche sp</i>	0.05	0	0.03	0	0.09	0.33	0	0.11	0.29	0
Insecta	Trichoptera	Hydroptilidae	<i>Hydroptila sp</i>	0	0.05	0	0	0.05	0	0	0	0	0.05
Hydrozoa	Hydrozoa	Hydrozoa	Hydrozoa	0	0	0.03	0.05	0.05	0	0.26	0	0.05	0
Insecta	Trichoptera	Rhyacophilidae	<i>Hyporhyacophila sp</i>	0	0	0	0.05	0.05	0.29	0	0.05	0.05	0
Insecta	Trichoptera	Limnephilidae	<i>Ironoquia dubia</i>	0	0	0.03	0	0	0	0	0	0	0
Insecta	Plecoptera	Perlodidae	<i>Isoperla grammatica</i>	0	0	0	0	0	0.04	0	0	0	0
Insecta	Plecoptera	Perlodidae	<i>Isoperla rivulorum</i>	0	0	0	0	0	0.04	0	0	0	0
Insecta	Plecoptera	Perlodidae	<i>Isoperla sp</i>	0.2	0.95	0.47	0.63	0.86	0.04	0.52	0.42	0	0.45
Insecta	Trichoptera	Hydroptilidae	<i>Ithytrichia sp</i>	0	0	0	0	0.05	0	0	0	0.05	0
Insecta	Trichoptera	Leptoceridae	Leptoceridae	0	0	0	0	0	0	0	0	0.24	0
Insecta	Ephemeroptera	Leptophlebiidae	Leptophlebiidae	0	0	0	0.05	0.27	0	0	0	0	0
Insecta	Plecoptera	Leuctridae	<i>Leuctra geniculata</i>	0	0	0	0	0	0	0.04	0	0	0
Insecta	Plecoptera	Leuctridae	<i>Leuctra sp</i>	0.4	0	0	0	0	0.42	0.09	0.26	0	0
Insecta	Trichoptera	Limnephilidae	Limnephilidae	0	0	0.07	0	0.05	0	0	0.21	0.14	0
Insecta	Trichoptera	Limnephilidae	Limnephilinae	0	0	0.4	0.21	0.05	0.04	0	0.47	0.05	0
Insecta	Trichoptera	Limnephilidae	<i>Limnephilus sp</i>	0.1	0.1	0	0.05	0.05	0.04	0.04	0	0	0
Insecta	Coleoptera	Elmidae	<i>Limnius sp</i>	0.2	0.05	0.13	0.16	0.05	0.17	0.35	0.21	0.62	0.23
Insecta	Trichoptera	Psychomyiidae	<i>Lype reducta</i>	0	0	0.03	0	0	0	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Melampophylax sp</i>	0.05	0	0	0	0	0.04	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Mesophylax sp</i>	0	0	0	0.05	0.05	0	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Micropterna sequax</i>	0	0	0	0.11	0	0	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Micropterna sp</i>	0.05	0.6	0.1	0.05	0.09	0.04	0.09	0	0	0.55
Insecta	Trichoptera	Limnephilidae	<i>Micropterna testacea</i>	0.1	0	0	0	0.09	0	0	0	0	0
Nematoda	Nematoda	Nematoda	Nematoda	0	0.1	0.2	0.11	0.18	0.17	0.35	0.05	0.19	0.09
Insecta	Plecoptera	Nemouridae	<i>Nemoura cinerea</i>	0	0	0	0	0	0.12	0	0	0	0
Insecta	Plecoptera	Nemouridae	<i>Nemoura marginata</i>	0	0	0	0	0	0.21	0	0	0	0
Insecta	Plecoptera	Nemouridae	<i>Nemoura sp</i>	0.85	0.85	0.87	0.58	0.86	0	0.57	0.84	0.43	0.59
Malacostraca	Amphipoda	Niphargidae	<i>Niphargus sp</i>	0	0.2	0.23	0.05	0.14	0	0.09	0	0	0.09
Insecta	Coleoptera	Elmidae	<i>Normandia sp</i>	0	0	0	0	0	0	0	0	0.1	0
Insecta	Coleoptera	Scirtidae	<i>Odeles sp</i>	0	0	0	0	0	0	0.04	0	0	0.05
Insecta	Trichoptera	Odontoceridae	<i>Odontocerum albicorne</i>	0.05	0	0	0	0	0.04	0	0.21	0.14	0
Oligochaeta	Oligochaeta	Oligochaeta	Oligochaeta	0.8	0.85	0.97	0.95	0.91	0.88	1	1	0.9	0.95
Insecta	Trichoptera	Phryganeidae	<i>Oligostomis reticulata</i>	0	0	0.03	0	0	0	0	0	0	0
Insecta	Odonata	Gomphidae	<i>Onychogomphus sp</i>	0	0	0	0	0	0	0.22	0	0	0
Insecta	Coleoptera	Gyrinidae	<i>Orectochilus villosus</i>	0	0	0	0	0	0	0	0	0.19	0.05
Insecta	Diptera	Chironomidae	Orthocladinae	0.65	0.95	1	0.89	0.95	0.92	0.91	0.79	0.86	0.91
Insecta	Trichoptera	Hydroptilidae	<i>Orthotrichia sp</i>	0	0	0	0	0	0	0	0	0.1	0
Malacostraca	Ostracoda	Ostracoda	Ostracoda	0.25	0.7	0.1	0.68	0.55	0.08	0.52	0.47	0.19	0.18

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Insecta	Coleoptera	Elmidae	<i>Oulimnius sp</i>	0.2	0.15	0.07	0.42	0.64	0.21	0.52	0.16	0.57	0.09
Insecta	Coleoptera	Hydrophilidae	<i>Paracymus sp</i>	0	0	0.03	0	0	0	0	0	0	0
Insecta	Diptera	Limoniidae	<i>Pediciini</i>	0	0	0	0	0	0.04	0.04	0.05	0.05	0
Insecta	Plecoptera	Perlidae	<i>Perla sp</i>	0	0	0	0	0	0.04	0	0	0	0
Insecta	Plecoptera	Perlodidae	<i>Perlodes jurassicus</i>	0.1	0	0	0	0	0.17	0	0	0.05	0
Platyhelminthes	Tricladida	Planariidae	<i>Phagocata vitta</i>	0	0	0	0	0	0	0	0.05	0	0
Mollusca	Gastropoda	Physidae	<i>Physa acuta</i>	0	0	0	0.05	0.05	0	0.13	0	0	0
Insecta	Diptera	Limoniidae	<i>Pilaria sp</i>	0.05	0	0	0.11	0	0	0	0.11	0	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium casertanum</i>	0	0	0.23	0	0.05	0.08	0	0.05	0.05	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium henslowanum</i>	0	0	0	0	0	0	0	0	0.1	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium milium</i>	0	0	0.03	0	0	0	0.09	0	0	0.05
Mollusca	Venerida	Sphaeriidae	<i>Pisidium nitidum</i>	0.1	0	0.1	0.11	0.05	0.04	0.13	0	0.14	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium obtusale</i>	0	0	0	0	0	0	0	0	0	0.05
Mollusca	Venerida	Sphaeriidae	<i>Pisidium personatum</i>	0.4	0.05	0.43	0.32	0.41	0.21	0.35	0.37	0.33	0.18
Mollusca	Venerida	Sphaeriidae	<i>Pisidium subtruncatum</i>	0.1	0	0.1	0.21	0.09	0	0.26	0.05	0.24	0.09
Mollusca	Venerida	Sphaeriidae	<i>Pisidium tenuilineatum</i>	0	0	0	0	0	0	0	0	0.1	0
Mollusca	Gastropoda	Planorbidae	<i>Planorbarius corneus</i>	0	0	0	0	0	0	0.04	0	0	0
Mollusca	Gastropoda	Planorbidae	<i>Planorbis carinatus</i>	0	0	0.03	0	0	0	0	0	0	0
Insecta	Odonata	Platycnemididae	<i>Platycnemis sp</i>	0	0	0	0.05	0	0	0	0	0	0
Insecta	Trichoptera	Polycentropodidae	<i>Plectrocnemia sp</i>	0.25	0.15	0.03	0	0	0.17	0	0	0.1	0
Platyhelminthes	Tricladida	Planariidae	<i>Polycelis felina</i>	0	0	0	0	0	0.29	0	0	0	0
Platyhelminthes	Tricladida	Planariidae	<i>Polycelis sp</i>	0.05	0	0.03	0.05	0.05	0.12	0	0.11	0.43	0
Insecta	Trichoptera	Polycentropodidae	<i>Polycentropodidae</i>	0	0	0	0.16	0.09	0	0	0.21	0	0.09
Insecta	Coleoptera	Dryopidae	<i>Pomatinus substriatus</i>	0	0	0	0	0	0	0	0.05	0	0
Insecta	Trichoptera	Limnephilidae	<i>Potamophylax sp</i>	0	0	0.1	0	0	0	0	0.05	0.05	0.09
Mollusca	Gastropoda	Amnicolidae	<i>Potamopyrgus antipodarum</i>	0	0.1	0	0.37	0.32	0	0.7	0.16	0.1	0.36
Malacostraca	Isopoda	Asellidae	<i>Proasellus meridianus</i>	0	0.05	0.03	0.16	0	0	0.26	0	0	0.09
Malacostraca	Isopoda	Asellidae	<i>Proasellus sp</i>	0	0	0	0	0.23	0	0	0	0	0
Insecta	Plecoptera	Nemouridae	<i>Protonemura sp</i>	0.35	0	0	0	0	0.29	0	0	0	0
Insecta	Diptera	Psychodidae	Psychodidae	0.2	0.25	0.07	0.26	0.14	0.29	0.3	0.58	0.33	0.23
Insecta	Trichoptera	Psychomyiidae	<i>Psychomyia pusilla</i>	0	0	0	0	0	0	0	0	0.05	0.05
Mollusca	Gastropoda	Lymnaeidae	<i>Radix sp</i>	0.05	0.05	0.07	0	0	0.04	0	0.21	0.14	0
Insecta	Diptera	Rhagionidae	Rhagionidae	0	0	0	0.11	0.09	0	0.09	0	0	0
Insecta	Ephemeroptera	Heptageniidae	<i>Rhithrogena sp</i>	0.45	0	0	0	0	0.54	0	0.58	0.43	0.23
Insecta	Trichoptera	Rhyacophilidae	<i>Rhyacophila sp</i>	0.2	0.3	0	0.37	0.23	0.21	0.48	0.37	0.43	0.27
Insecta	Coleoptera	Elmidae	<i>Riolus sp</i>	0.1	0.25	0	0.26	0.23	0.25	0.39	0.21	0.76	0.45
Insecta	Diptera	Scatophagidae	Scatophagidae	0	0	0	0	0.05	0	0	0	0	0
Insecta	Diptera	Limoniidae	<i>Scleroprocta sp</i>	0	0	0	0	0	0.04	0	0	0	0
Insecta	Trichoptera	Sericostomatidae	<i>Sericostoma sp</i>	0.05	0	0.1	0.26	0.09	0.21	0.13	0.16	0.05	0.18
Insecta	Ephemeroptera	Ephemerellidae	<i>Serratella ignita</i>	0	0.3	0	0.37	0.5	0.08	0.7	0	0.33	0.14
Insecta	Megaloptera	Sialidae	<i>Sialis sp</i>	0.05	0	0	0	0	0	0	0.11	0.1	0
Insecta	Trichoptera	Goeridae	<i>Silo nigricornis</i>	0	0.05	0.07	0	0	0	0	0	0	0
Insecta	Trichoptera	Goeridae	<i>Silo sp</i>	0	0	0	0.16	0	0	0	0	0.1	0
Insecta	Diptera	Simuliidae	Simuliidae	0.7	0.9	0.97	0.58	0.64	0.88	0.52	0.47	0.71	0.59
Insecta	Ephemeroptera	Siphonuridae	<i>Siphonurus aestivalis</i>	0	0	0	0	0.05	0	0	0	0	0
Insecta	Ephemeroptera	Siphonuridae	<i>Siphonurus sp</i>	0	0	0.03	0.05	0	0	0	0	0	0.09
Insecta	Plecoptera	Chloroperlidae	<i>Siphonoperla sp</i>	0	0	0	0	0	0	0	0.05	0	0
Insecta	Plecoptera	Chloroperlidae	<i>Siphonoperla torrentium</i>	0.05	0	0	0	0	0	0	0	0	0

Chapitre 2 - Métacommunautés dans les têtes de bassin versant, le rôle de la dispersion et de la fragmentation

Mollusca	Venerida	Sphaeriidae	<i>Sphaerium corneum</i>	0	0	0.03	0	0	0	0	0	0	0
Insecta	Coleoptera	Elmidae	<i>Stenelmis sp</i>	0	0	0	0.05	0	0	0	0	0.05	0
Insecta	Trichoptera	Limnephilidae	<i>Stenophylax mitis</i>	0	0	0	0	0	0.04	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Stenophylax mucronatus</i>	0	0.05	0	0.05	0	0	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Stenophylax permistus</i>	0.05	0.2	0.1	0	0.14	0.04	0	0	0	0.18
Insecta	Trichoptera	Limnephilidae	<i>Stenophylax sp</i>	0	0	0.13	0	0.09	0.04	0.04	0	0	0
Insecta	Diptera	Stratiomyidae	Stratiomyidae	0	0.15	0	0.37	0.45	0.21	0.39	0.47	0.05	0.36
Insecta	Trichoptera	Glossosomatidae	<i>Synagapetus sp</i>	0.05	0	0	0	0	0	0.04	0	0.05	0.41
Insecta	Diptera	Tabanidae	Tabanidae	0.05	0.05	0.03	0.16	0.05	0	0.26	0	0.1	0.05
Insecta	Diptera	Chironomidae	Tanypodinae	0.2	0.1	0.07	0.32	0.14	0.29	0.3	0.58	0.38	0.41
Insecta	Diptera	Chironomidae	<i>Tanytarsini</i>	0.4	0.25	0.1	0.26	0.41	0.33	0.17	0.84	0.48	0.32
Insecta	Trichoptera	Psychomyiidae	<i>Tinodes sp</i>	0.1	0.05	0	0.16	0.09	0.12	0.26	0.42	0.19	0.32
Insecta	Trichoptera	Psychomyiidae	<i>Tinodes waeneri</i>	0.05	0	0	0	0	0	0	0	0	0
Insecta	Diptera	Tipulidae	Tipulidae	0.15	0.3	0.07	0.05	0.05	0.12	0.09	0.16	0.14	0.09
Insecta	Ephemeroptera	Ephemerellidae	<i>Torleya major</i>	0	0	0	0	0	0.04	0	0	0.24	0
Platyhelminthes	Tricladida	Tricladida	Tricladida	0	0.05	0	0.11	0.05	0	0	0	0	0.09
Porifera	Demospongiae	Spongillidae	<i>Trochospongilla horrida</i>	0	0	0	0	0	0.04	0	0	0	0
Mollusca	Gastropoda	Valvatidae	<i>Valvata sp</i>	0	0	0	0	0	0	0.04	0	0	0
Insecta	Trichoptera	Philopotamidae	<i>Wormaldia sp</i>	0	0.1	0	0.05	0	0.04	0	0	0	0.23
Insecta	Plecoptera	Capniidae	<i>Zwicknia bifrons</i>	0	0	0.07	0	0	0	0	0	0	0
Insecta	Plecoptera	Capniidae	<i>Zwicknia ledoarei</i>	0.25	0	0	0	0	0.12	0	0.58	0	0
Insecta	Plecoptera	Capniidae	<i>Zwicknia westermanni</i>	0	0	0.6	0.42	0.27	0	0.04	0	0	0.45
Specific richness				16.05	16.3	12.57	18.84	18.05	17.79	19.35	19.84	20.67	18.73

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

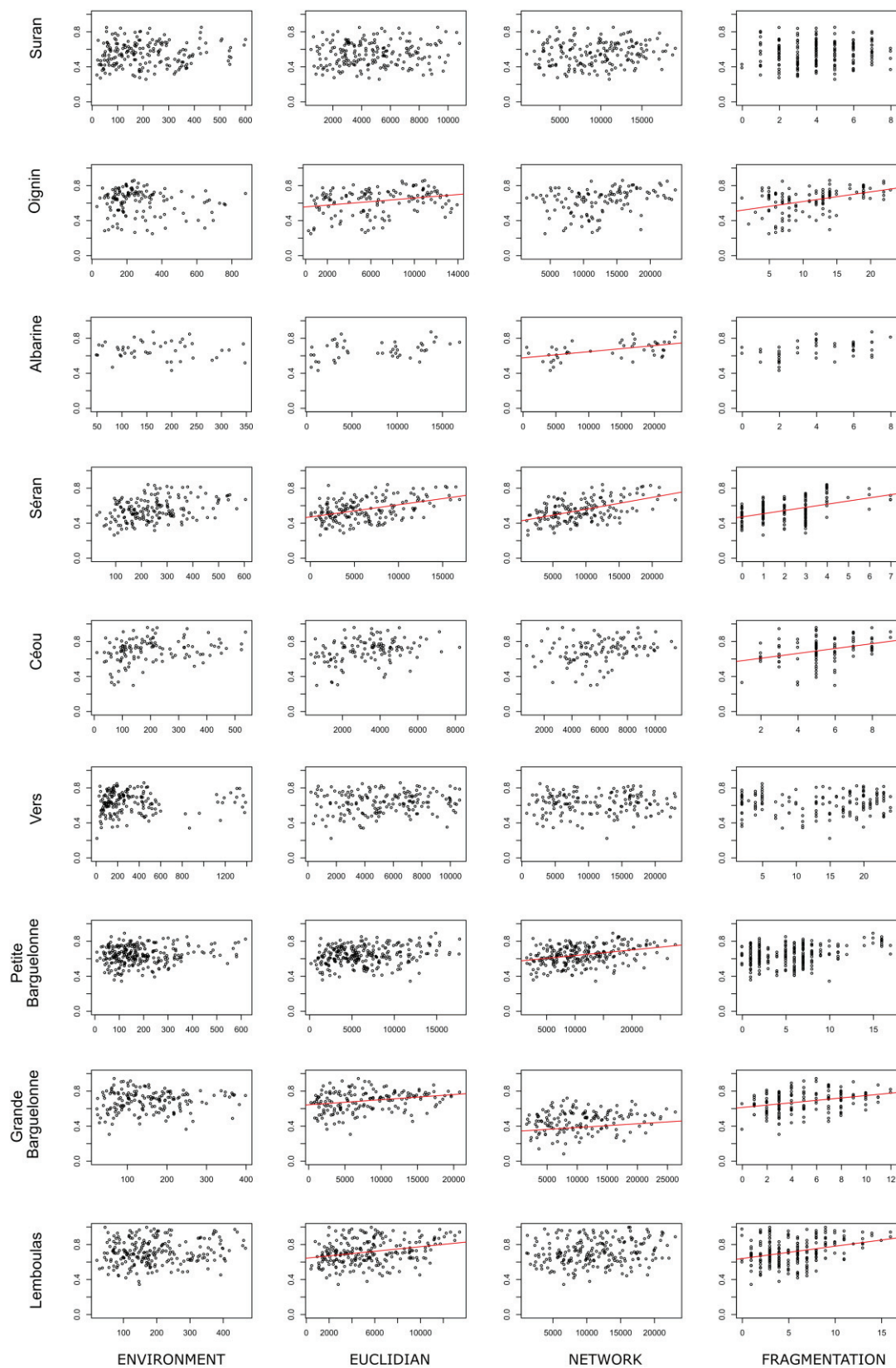
Appendix S5. Table of coefficient of correlation of the Mantel tests based on Jaccard and Bray-Curtis dissimilarity indices carried out to disentangle metacommunity processes. With respect to beta-diversity: TOT, the whole community, AER, taxa with an aerial life stage, AQU, strictly aquatic taxa. With respect to distances: ENV, environmental; EUC, overland; NET, network; FRAG, number of dry reaches and anthropogenic structures. Significant relationships are shown in bold (p-value ≤ 0.05 *). Holm correction was used to adjust p-values for multiple comparisons.

		Jaccard dissimilarity index					
FWH	Distances matrices	Before drying season			After drying season		
		TOT	AER	AQU	TOT	AER	AQU
Clauge	ENV				0.00	-0.03	0.11
	EUC				0.26 *	0.27 *	0.14
	NET				-0.08	-0.03	-0.14
	FRAG				-0.22	-0.23	-0.11
Suran	ENV	0.09	0.15	-0.04	-0.11	-0.12	-0.04
	EUC	0.10	0.10	0.02	0.01	0.00	0.01
	NET	0.09	0.17	-0.02	0.02	0.06	-0.04
	FRAG	0.09	0.15	-0.01	0.06	0.05	0.04
Oignin	ENV	-0.14	-0.19	0.01	0.00	0.00	0.02
	EUC	0.26 *	0.26 *	0.16	0.08	0.15 *	-0.06
	NET	0.28	0.22	0.28	0.14	0.21 *	-0.03
	FRAG	0.40 *	0.33	0.38 *	0.18 *	0.24 *	0.01
Albarine	ENV	-0.09	-0.19	0.05	0.31	0.01	0.24
	EUC	0.28	0.34	0.06	-0.01	0.27	-0.09
	NET	0.47 *	0.45 *	0.26	-0.11	0.06	-0.14
	FRAG	0.39	0.42	0.13	0.01	-0.02	-0.05
Séran	ENV	0.27	0.19	0.30	0.13	0.08	0.08
	EUC	0.43 *	0.44 *	0.22	0.39 *	0.36	0.16
	NET	0.52 *	0.54 *	0.24	0.31	0.35	0.03
	FRAG	0.45 *	0.45 *	0.22	0.50 *	0.48 *	0.22
Céou	ENV	0.17	0.20	-0.01	-0.01	-0.03	-0.01
	EUC	0.27 *	0.32 *	0.04	-0.10	0.02	-0.18
	NET	0.23	0.09	0.31	0.21	0.22	0.07
	FRAG	0.31 *	0.19	0.36 *	0.20	0.23	0.04
Vers	ENV	-0.05	0.08	-0.16	0.00	-0.03	0.06
	EUC	0.12	0.14	0.07	0.15	0.13	0.07
	NET	0.07	-0.04	0.12	-0.03	-0.03	0.00
	FRAG	0.06	-0.01	0.09	-0.01	0.01	-0.03
Petite Barguelonne	ENV	0.20	0.20	0.12	0.16	0.05	0.17
	EUC	0.21	0.21	0.13	0.11	0.04	0.12
	NET	0.31 *	0.12	0.32 *	0.13	-0.17	0.33 *
	FRAG	0.28	0.28 *	0.15	0.22	0.13	0.21
Grande Barguelonne	ENV	-0.03	0.04	-0.10	0.14	0.21	-0.03
	EUC	0.25	0.22	0.16	0.25 *	0.28	0.05
	NET	0.17	0.14	0.12	0.16	0.21	-0.02
	FRAG	0.30 *	0.21	0.26	0.14	0.14	0.05
Lemboulas	ENV	0.10	0.06	0.09	-0.05	0.01	-0.09
	EUC	0.27 *	0.21	0.24	-0.01	-0.02	-0.01
	NET	0.18	0.10	0.22	0.03	-0.02	0.07
	FRAG	0.33 *	0.33 *	0.19	-0.01	0.01	-0.05

		Bray-Curtis dissimilarity index					
FWH	Distance matrices	Before drying season			After drying season		
		TOT	AER	AQU	TOT	AER	AQU
Clauge	ENV				-0.01	-0.05	0.15
	EUC				0.27 *	0.28 *	0.15
	NET				-0.08	-0.03	-0.14
	FRAG				-0.24	-0.25	-0.11
Suran	ENV	0.05	0.12	-0.07	-0.11	-0.11	-0.07
	EUC	0.14	0.14	0.07	0.00	-0.02	0.00
	NET	0.09	0.15	0.03	0.02	0.07	-0.05
	FRAG	0.09	0.15	0.03	0.06	0.06	0.03
Oignin	ENV	-0.15	-0.14	-0.14	0.00	-0.01	0.03
	EUC	0.23 *	0.23 *	0.12	0.07	0.14	-0.06
	NET	0.27 *	0.23	0.17	0.13	0.20 *	-0.03
	FRAG	0.38 *	0.34	0.26	0.17	0.23	0.03
Albarine	ENV	-0.01	-0.14	0.08	0.36	0.31	0.26
	EUC	0.32	0.31	0.25	-0.02	0.04	-0.11
	NET	0.52	0.43 *	0.43 *	-0.11	-0.06	-0.15
	FRAG	0.48 *	0.41	0.37	0.01	0.04	-0.06
Séran	ENV	0.26	0.26	0.17	0.11	0.08	0.09
	EUC	0.50 *	0.48 *	0.34 *	0.39 *	0.37 *	0.17
	NET	0.57 *	0.57 *	0.36 *	0.33 *	0.37 *	0.03
	FRAG	0.50 *	0.47 *	0.39 *	0.50 *	0.47 *	0.23 *
Céou	ENV	0.22	0.22	0.01	-0.01	0.00	-0.01
	EUC	0.22	0.23	0.07	-0.11	-0.02	-0.19
	NET	0.17	0.05	0.32	0.19	0.19	0.06
	FRAG	0.31 *	0.19	0.38 *	0.17	0.20	0.03
Vers	ENV	-0.04	0.20	-0.21	0.03	-0.01	0.07
	EUC	0.13	0.14	0.08	0.14	0.14	0.04
	NET	0.08	-0.06	0.14	-0.05	-0.03	-0.04
	FRAG	0.09	-0.01	0.13	-0.02	0.01	-0.06
Petite Barguelonne	ENV	0.15	0.18	0.07	0.16	0.06	0.16
	EUC	0.22	0.22 *	0.13	0.11	0.04	0.13
	NET	0.36 *	0.19	0.34 *	0.13	-0.18	0.34 *
	FRAG	0.34	0.32	0.21	0.23	0.13	0.22
Grande Barguelonne	ENV	-0.05	-0.01	-0.05	0.16	0.25	0.00
	EUC	0.18	0.13	0.17	0.26 *	0.29	0.04
	NET	0.18	0.16	0.13	0.18	0.24 *	-0.01
	FRAG	0.30	0.23	0.26	0.14	0.15	0.05
Lemboulas	ENV	0.15	0.16	0.07	-0.05	0.01	-0.10
	EUC	0.29 *	0.25	0.22 *	-0.02	-0.01	-0.02
	NET	0.19	0.15	0.16	0.03	-0.01	0.08
	FRAG	0.35	0.38	0.22	-0.02	0.01	-0.04

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

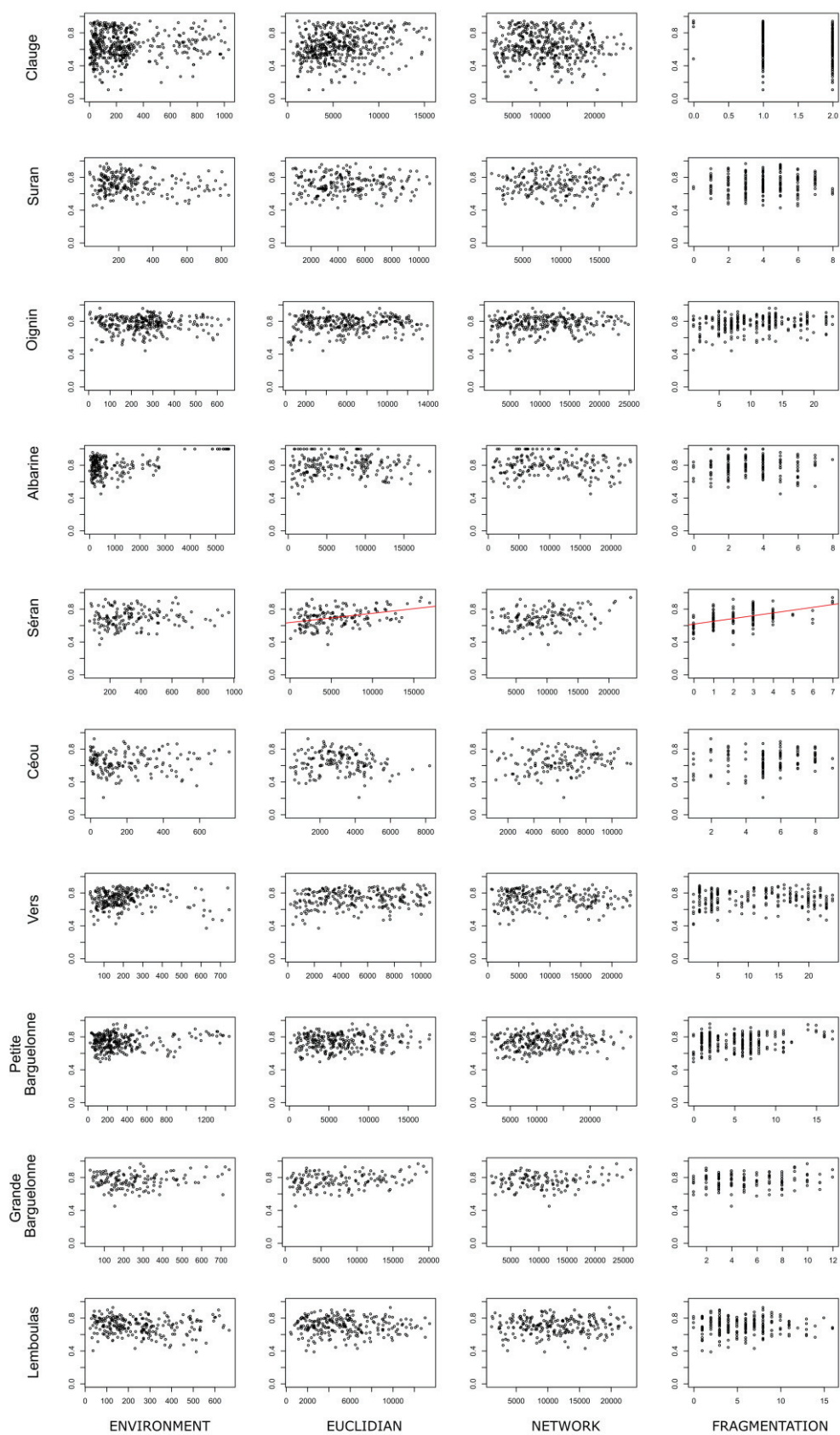
Appendix S6: Distance decay relationships for each combination of community type (Chao dissimilarity index) and distance matrices for each FHW before (a, c, e) and after (b, d, f) the drying season, for the whole community (a,b), the community of taxa with aerial stages (c,d) and strictly aquatic taxa (e,f). Distance matrices are ENV, environmental distance; EUC, overland distance; NET, network distance; FRAG, distance based on number of dry reaches and anthropogenic structures. Only significant relationships are shown (≤ 0.05). Corresponding exact r values are available in Tableau 2.2.



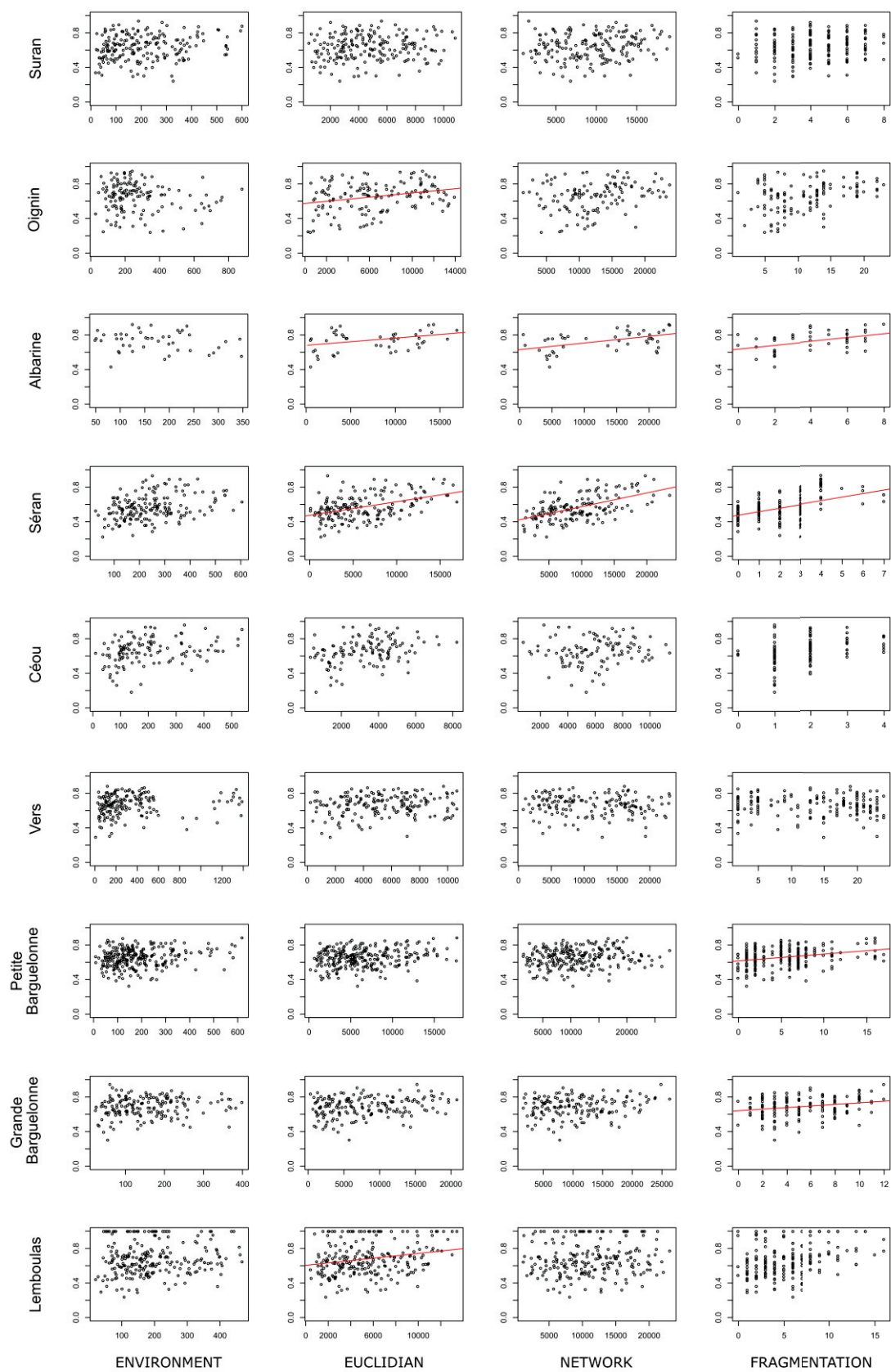
(a)

2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

(b)

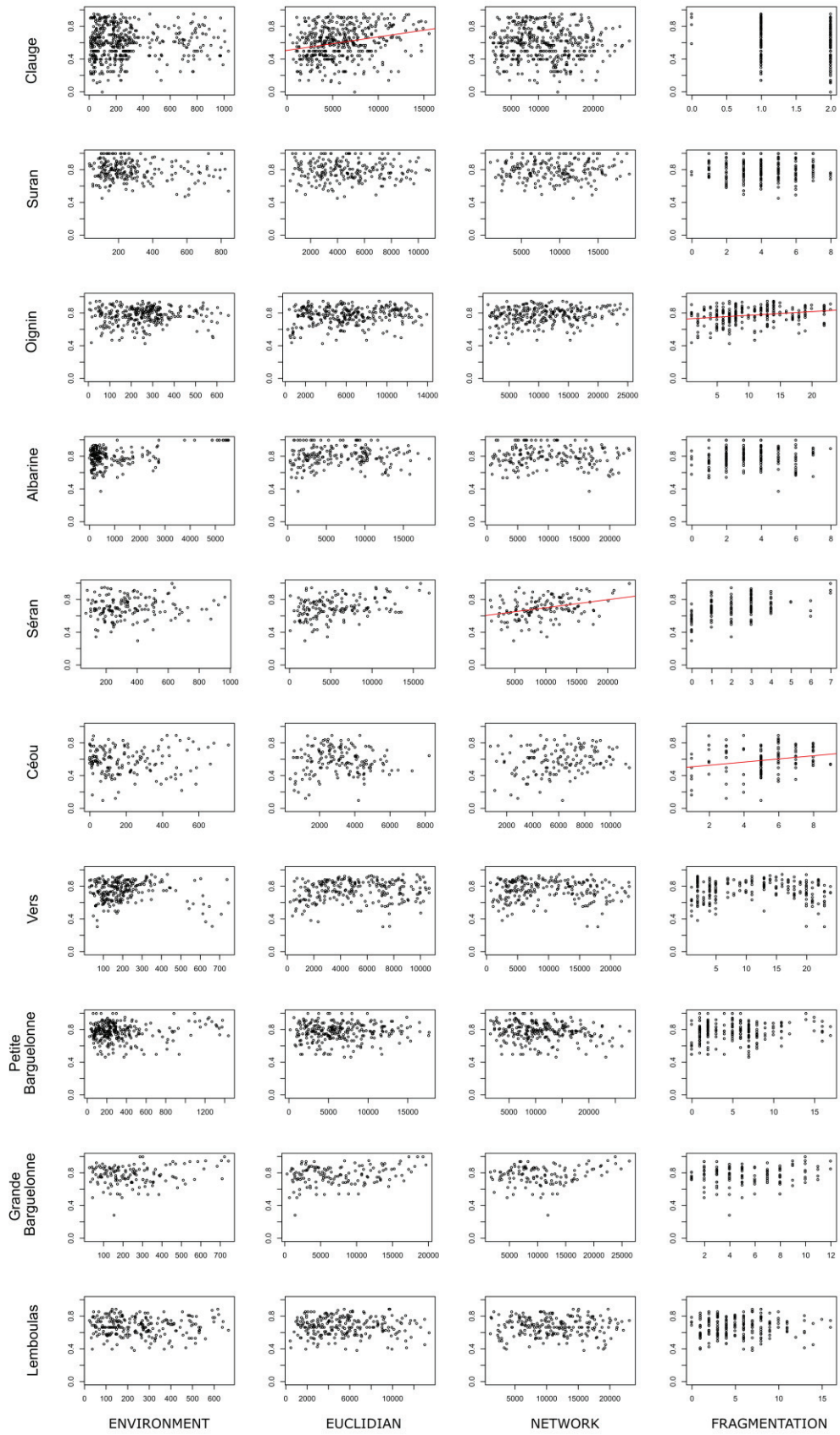


(c)

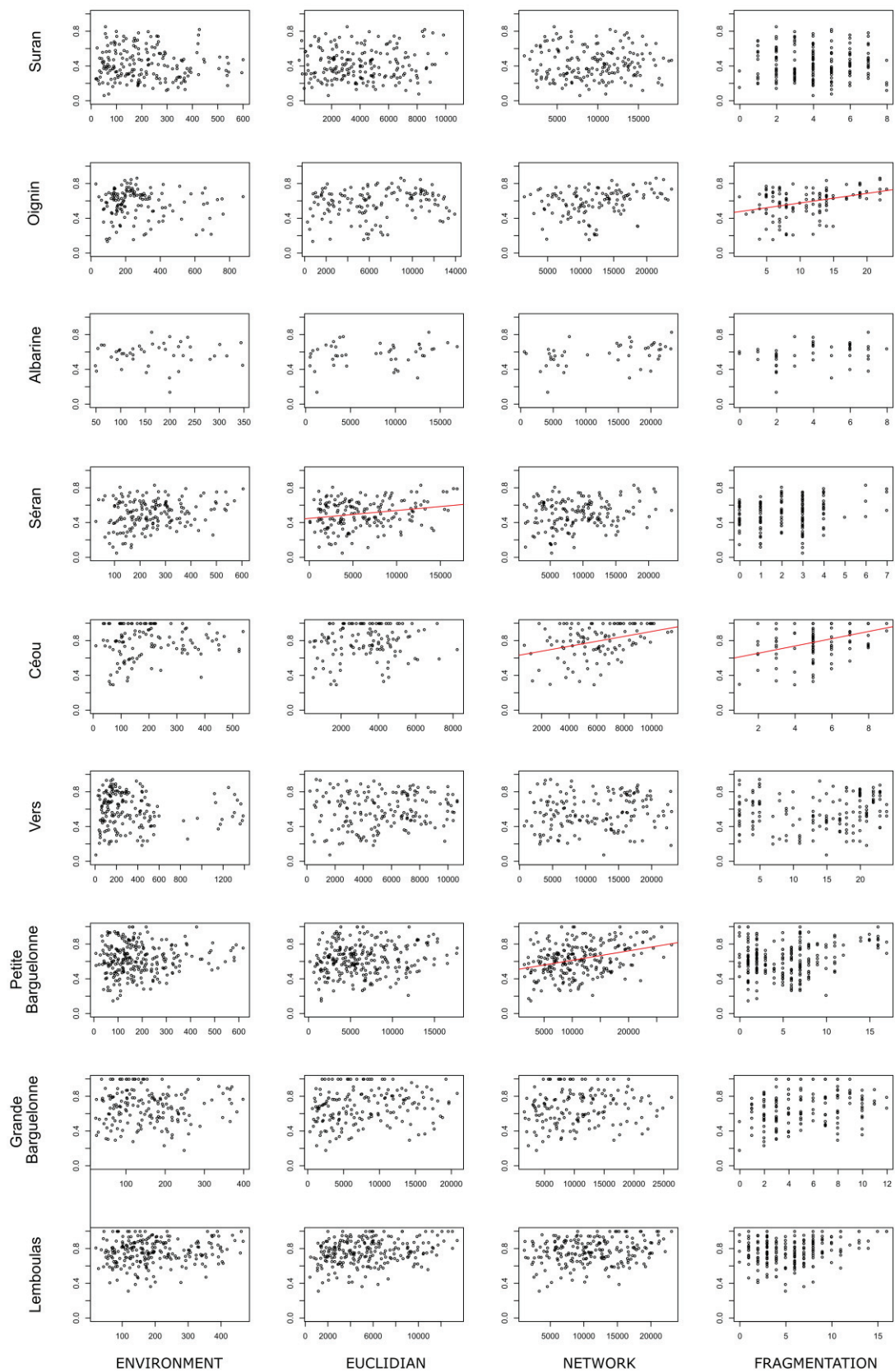


2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

(d)

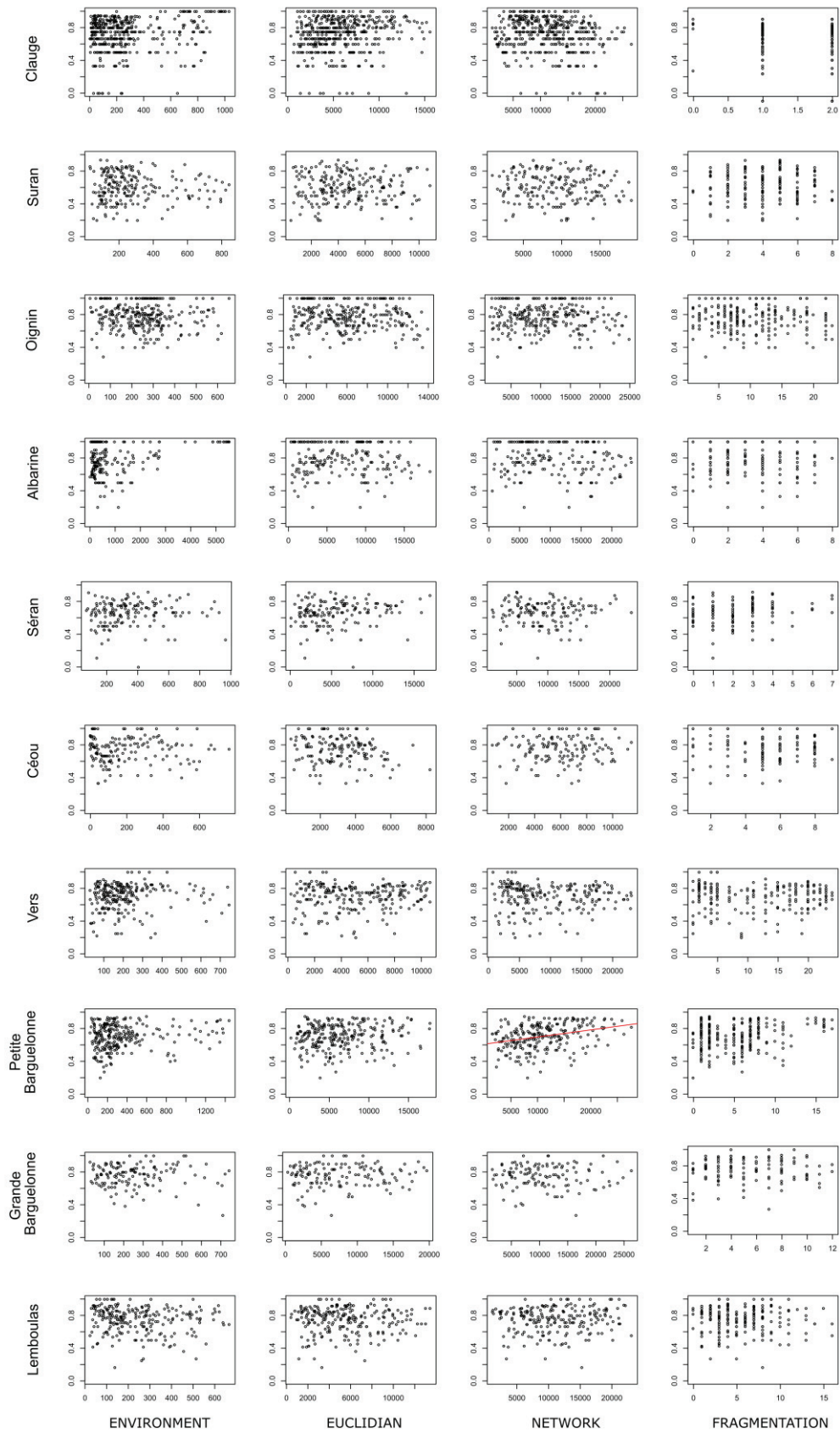


(e)

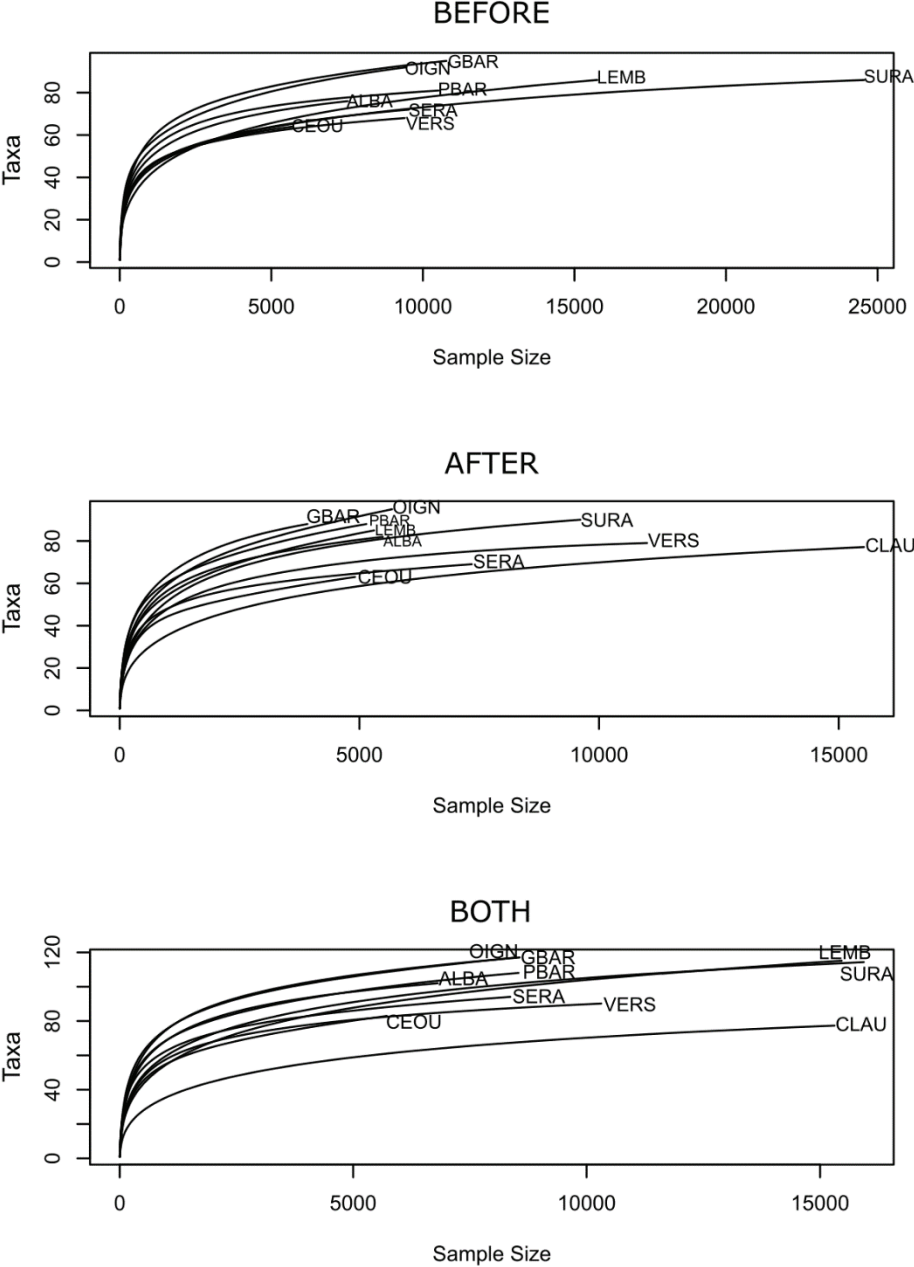


2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

(f)



Appendix S7. Rarefaction curves for each headwater stream network and for different sampling period (BEFORE and AFTER the drying period and the mean of the two sampling period, BOTH).



2. Étude de la structure des métacommunautés dans 10 têtes de bassin versant subissant différents degrés de fragmentation (article 1)

Literature Cited in Supplemental Information

Brown, B. L., Swan, C. M., Auerbach, D. A., Campbell Grant, E. H., Hitt, N. P., Maloney, K. O., & Patrick, C. (2011). Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *Journal of the North American Benthological Society*, 30(1), 310-327. doi:10.1899/10-129.1

Crabot, J., Heino, J., Launay, B., & Datry, T. (2019). Drying determines the temporal dynamics of stream invertebrate structural and functional beta diversity. *Ecography*, 1-16. doi:10.1111/ecog.04835

Tonkin, J. D., Heino, J., Sundermann, A., Haase, P., & Jähnig, S. C. (2016). Context dependency in biodiversity patterns of central German stream metacommunities. *Freshwater Biology*, 61(5), 607-620. doi:10.1111/fwb.12728

Viana, D. S., & Chase, J. M. (2019). Spatial scale modulates the inference of metacommunity assembly processes. *Ecology*, 100(2), e02576. doi:10.1002/ecy.2576

3 Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

Titre en anglais : Drying does not alter metacommunity dynamics in a highly fragmented river network

Gauthier M.^{1,2}, Le Goff G.¹, Launay B.¹, Douady C. J.², Datry T.¹

¹ INRAE, UR-Riverly, Centre de Lyon-Villeurbanne, 5 rue de la Doua CS70077, 69626 VILLEURBANNE Cedex, France

² Univ Lyon, Université Claude Bernard Lyon 1, CNRS, ENTPE, UMR5023 LEHNA, F-69622, Villeurbanne, France

Dans l'article précédent, les métacommunautés de macro-invertébrés benthiques de têtes de bassins ont été étudiées en privilégiant une approche multi-bassins permettant de prendre en compte différents contextes de fragmentation et de mettre en évidence des tendances dans leur structure ainsi que les processus associés. Bien que l'effort d'échantillonnage soit conséquent, nous avons tout de même pu aborder dans l'article 1 la question de la variation temporelle des processus avec deux dates d'échantillonnage et avons mis en évidence des variations dans les processus structurant les métacommunautés. L'article 2 se concentre sur cette variation temporelle avec le suivi de la métacommunauté de macro-invertébrés benthiques d'une tête de bassin versant sur plusieurs mois en englobant les différentes phases hydrologiques caractérisant les rivières intermittentes. De plus, cet article cherche à approfondir la compréhension du rôle de la fragmentation dans la structure des métacommunautés en analysant les réponses des communautés aux fragmentations par les assèchements et par les structures anthropiques séparément.

3.1 Résumé étendu

A l'ère de l'Anthropocène, les réseaux hydrographiques sont de plus en plus régulés par les actions de l'être humain de manière directe et indirecte (de Graaf et al, 2019, Reid et al, 2019) avec comme conséquence une intensification de la régulation des débits et de la fragmentation des réseaux à l'échelle globale (Grill et al., 2019). Ces altérations menacent la biodiversité aquatique qui présente l'un des taux les plus élevés de perte de diversité (Reid et al., 2019). Le développement de l'hydroélectricité (Zarlf et al, 2014) couplé à la diminution du débit lié au changement climatique (Reid et al, 2019) dans les réseaux hydrographiques a entraîné un développement rapide de structures

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

tels que des barrages, des seuils ou des retenues d'eau de plus ou moins grande taille (Lehner et al, 2011). Ces structures constituent des barrières fragmentant de manière permanente les réseaux et empêchent la dispersion des organismes au sein des réseaux ou entre eux (Fuller et al, 2015). En réponse à l'augmentation de construction de retenues d'eau et de prélèvements d'eau et au changement climatique, une augmentation de la fréquence et de la durée des assèchements qui fragmente temporairement les réseaux hydrographiques est également en cours (Datry et al, 2018).

Comme présenté en Introduction et montré dans l'article 1, la fragmentation a un impact sur la dispersion des organismes au sein des réseaux hydrographiques et peut altérer les dynamiques des métacommunautés. Pourtant, à ce jour, les effets individuels de la fragmentation permanente par les structures anthropiques et de la fragmentation temporaire par les assèchements sur les métacommunautés aquatiques n'ont pas encore été explorés. Dans un contexte de fragmentation permanente intense, l'effet de la fragmentation temporaire sur la structure des communautés devrait être limité, la dispersion des organismes étant déjà considérablement limitée par les structures anthropiques. En revanche, dans un contexte de fragmentation permanente faible ou modérée, la fragmentation temporaire devrait induire une limitation de la dispersion forte étant donné que les sections asséchées peuvent s'étendre sur de longues distances (plusieurs kilomètres). C'est pourquoi il est fondamental d'évaluer les effets des assèchements en interaction avec ceux des structures anthropiques. En effet, cela pourrait amener à des implications importantes concernant les priorités de gestion et de conservation à mettre en œuvre pour la sauvegarde de la biodiversité des réseaux hydrographiques dynamiques et fragmentés dont font partie les têtes de bassin versant. De plus, la variabilité de la fragmentation temporaire et plus généralement du débit et des cycles d'assèchement et de remise en eau pose la question de la variation dans le temps des processus structurant les métacommunautés des systèmes dynamiques sujets aux assèchements. Pendant la phase d'assèchement, les conditions environnementales changent drastiquement dans les rivières avec un glissement de conditions lotique à lentique, rendant à priori le processus de sélection de niche dominant dans la structure des métacommunautés. En revanche, pendant et après la remise en eau, le processus de dispersion devrait prédominer avec une recolonisation passive ou active des sections asséchées.

Dans cette étude, nous avons exploré la variabilité temporelle des processus structurant les métacommunautés de macro-invertébrés benthiques dans une tête de bassin versant et évalué les influences respectives de la fragmentation permanente et temporaire. Pour cela, nous avons couplé un échantillonnage spatial (30 sites) et temporel (6 dates représentant 4 phases hydrologiques) intense dans une tête de bassin versant présentant une fragmentation forte par les structures anthropiques et sujette aux assèchements. Les invertébrés ont été échantillonnés dans un radier par site à chaque date, couplé à des mesures systématiques des conditions environnementales, et identifiés au niveau taxinomique le plus fin possible (détails dans le Matériel et méthodes - Chapitre 2). Les relations entre

diversité beta et matrices de distances environnementales et spatiales (euclidiens, réseaux, de fragmentation permanente par les structures anthropiques, de fragmentation temporaire par les assèchements) ont été quantifiées et testées avec des tests de Mantel.

La tête de bassin versant était fragmentée par 32 structures anthropiques (principalement des barrages et des seuils) et plus de 40% de linéaire était asséché au maximum de la période d'été. Nos résultats mettent en évidence une variation de la part des processus et de leur intensité dans la structure des métacommunautés à travers les différentes phases hydrologiques. Cela fait écho à de récentes études mettant l'accent sur l'importance d'intégrer la composante temporelle dans l'étude des métacommunautés de réseaux hydrographiques dynamiques (Datry et al, 2016 ; Jabot et al, 2020 ; Sarremejane et al, 2017). Comme attendu, le rôle du processus de dispersion était prédominant dans ce réseau hydrographique fragmenté et dynamique où des événements d'assèchement et de remise en eau se succèdent et où la dispersion est limitée (article 1). Cependant, aucune empreinte du processus de sélection de niche n'a été mis en évidence même pendant la phase d'assèchement. Malgré des conditions environnementales variables entre sites et dates d'échantillonnage, leurs valeurs ne semblaient pas dépasser des limites seuils pouvant engendrer un fort filtre environnemental (Quinn et al, 1994), ce qui pourrait expliquer ce résultat non attendu. Les habitats des mouilles qui pourraient connaître des changements de conditions environnementales plus prononcés lors des assèchements n'ont pas été échantillonnés dans cette étude et leur inclusion pourrait permettre de mettre en évidence des processus basés sur la sélection de niche.

Nos résultats montrent également que pendant la période d'été et au-delà, la diversité beta était expliquée par les distances de fragmentation permanente et non par celles de fragmentation temporaire, rejoignant plusieurs études sur la limitation de la dispersion par les structures anthropiques chez des communautés et populations d'invertébrés au sein des réseaux hydrographiques (Blanchet et al, 2010 ; Katano et al, 2006). Comme énoncé précédemment, l'absence de limitation de la dispersion par les sections asséchées pourraient s'expliquer par une limitation suffisamment forte par les structures anthropiques, rendant celle induite par les assèchements soit inexistante, soit indétectable. Les effets des assèchements sur la dispersion des espèces au sein d'un réseau ont pourtant déjà été montrés dans une précédente étude où la fragmentation par les structures anthropiques étaient faibles (Cañedo-Argüelles et al. 2015), mettant à nouveau en avant l'importance du contexte dans l'étude des métacommunautés. De plus, comme pour l'article 1, malgré des résultats divergents, des réponses différentes ont été observées selon les modes de dispersion des taxa. Les taxa strictement aquatiques ont été les seuls impactés par la limitation à la dispersion, que ce soit par les distances euclidiennes ou celles associées à la fragmentation permanente en contraste avec les taxa ayant un stade de vie aérien qui peuvent coloniser les sections intermittentes en volant sans être impactés par les structures anthropiques. En complément d'une meilleure description des modes et capacités de dispersion des espèces, l'estimation de la dispersion effective (c'est-à-dire d'une dispersion suivie d'une reproduction

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

réussie) en comparant la structure génétique des métapopulations permettrait de mieux comprendre la dynamique des métacommunautés dans les systèmes fragmentés (Coulon, 2010).

L'ensemble des résultats de cette étude indique que les processus structurant les métacommunautés de macro-invertébrés benthiques dans de réseaux hydrographiques dynamiques varient dans le temps et suggèrent qu'il est nécessaire de reconnaître et d'intégrer cette composante temporelle dans l'étude de leur biodiversité et leur conservation. Cette intégration nécessitant un effort d'échantillonnage et d'identification conséquent et potentiellement difficile à appliquer, de nouveaux outils d'identification moléculaire tel que le metabarcoding pourrait offrir la possibilité de surmonter ce problème (p.ex. Bush et al, 2020). Autrement, la définition de phases hydrologiques clés sur lesquelles se baser pour évaluer l'état écologique des rivières et améliorer les pratiques de gestion pourrait aussi être pertinent. Une première étape cruciale serait de mieux comprendre la variabilité des processus structurant les métacommunautés pour identifier ces phases hydrologiques. Pour ce faire, notre étude souligne, en accord avec l'article 1, la nécessité de prendre en compte la complexité pouvant être rencontrée dans les réseaux hydrographiques, notamment celle des capacités de dispersion des espèces ainsi que celle concernant la nature et les effets de la fragmentation permanente et temporaire sur les métacommunautés. En effet, leurs rôles dans la structure des métacommunautés semblent dépendre du contexte de fragmentation du réseau et il est essentiel d'évaluer ces impacts dans des contextes où les intensités de fragmentation permanente et temporaire varient afin d'explorer les synergies et les antagonismes entre les deux types de fragmentation. Il est urgent de centrer les recherches sur la quantification de ces effets à l'ère de l'Anthropocène où les réseaux de rivières subissent une fragmentation accrue par les assèchements et le développement de structures anthropiques pour l'irrigation et les besoins humains (Reid et al, 2019) afin d'établir des liens entre la biodiversité des rivières et la fragmentation et développer des outils efficaces et pertinents pour la gestion et la conservation des réseaux hydrographiques.

3.2 Abstract

In the Anthropocene, river networks are globally threatened by human activities, with dramatic consequences for aquatic biodiversity. Permanent and temporary fragmentation by man-made structures and drying events, respectively, are both increasing in rivers worldwide. Although these two fragmentation types can limit dispersal, their individual and interacting effects at both local (site-level) and regional (network-level) scales in shaping metacommunity dynamics has not been explored yet. Here, we examined whether processes structuring aquatic invertebrate metacommunities vary through time in response to flow variability in a river network affected by drying and severe fragmentation by man-made structures. We also compared the relative influences of permanent and temporary fragmentation. We hypothesized permanent fragmentation to be the primary determinant of metacommunity dynamics. We conducted an intensive sampling effort (30 sampling sites × 6 dates)

and measured 12 local environmental variables and 4 spatial distances to assess the relative importance of niche- and dispersal-based processes on benthic invertebrate metacommunities across hydrological phases. Spatial distances considered here were overland, network and two fragmentation distances integrating the permanent fragmentation by man-made structures and temporary fragmentation by drying. We used Mantel tests to identify relationships between community dissimilarities and environmental and spatial distance matrices. We identified significant temporal variability in metacommunity processes, with a predominant role of dispersal and no effect of niche-based processes. Metacommunities were shaped primarily by permanent fragmentation, whereas we detected no effect of fragmentation by network-scale drying. Dispersal-based metacommunity processes varied over time, following the wet-dry cycles that characterize dynamic river networks. Our results suggest the importance of key hydrological phases that should be incorporated into conservation perspectives. In addition, we highlight the need to quantify context dependency in metacommunity studies, to optimize biodiversity conservation strategies in river networks ecosystems.

3.3 Introduction

Ecosystem fragmentation, the division of habitats into smaller and isolated fragments separated by a matrix of human-transformed land cover, is a major threat to biodiversity in the Anthropocene era (Haddad et al. 2015, Crooks et al. 2017). As well as decreasing habitat availability, fragmentation alter the movements of organisms among local habitat patches within ecosystems, i.e. dispersal (Leibold et al. 2004, Brown et al. 2011, Heino et al. 2017). Dispersal interacts with local environmental selection to play a major role in promoting species coexistence at regional levels (Poff 1997, Leibold et al. 2004, Brown et al. 2010, Sarremejane et al. 2020).

Globally, human activities including land use change, dams and extraction of surface and ground water are altering river flow regimes by modifying runoff patterns through catchments (de Graaf et al. 2019, Reid et al. 2019). In recent decades, the speed of these changes is unprecedented, accelerating the fragmentation and intensity of flow regulation of river networks in most parts of the world (Grill et al. 2019). These alterations threaten freshwater biodiversity, which is experiencing higher extinction rates than in other ecosystems (Reid et al. 2019).

Increasing hydropower development in many countries has fragmented river networks through the construction of large and small structures including dams, weirs and culverts (Zarfl et al. 2014). Similarly, the increasing frequency of droughts in the past decades is motivating the development of small water retention structures, often located within a stream. These structures include small to very small impoundments with an impounded area $<0.1 \text{ km}^2$ and a volume $<0.2 \text{ hm}^3$ (Lehner et al. 2011). In spite of the low volume of stored water, they represent 99.5% of the total number of reservoirs worldwide and are the most common anthropogenic fragmentation structures in river networks

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

(Lehner et al. 2011). Together with large dams, they constitute barriers that permanently fragment rivers thus preventing dispersal of organisms within or across networks (Fuller et al. 2015).

In response to water abstraction and climate change, natural drying events that temporarily fragment river networks are also becoming increasingly common (Datry et al. 2018, de Graaf et al. 2019). Drying is defined here as the cessation of flow or the complete disappearance of surface water in a given reach (Ruhi et al. 2017). An estimated half of the total length of river network is prone to drying, and this proportion is increasing in regions experiencing drying trends and over-abstraction (Döll et al. 2012). Whereas the local effects of drying on aquatic biodiversity are fairly well-understood (e.g. Datry et al. 2017), the effects of fragmentation by drying on aquatic biodiversity remain understudied at the network scale for biotic groups including aquatic invertebrates (but see Cañedo-Argüelles et al. 2015, Crabot et al. 2020, Gauthier et al. 2020, Sarremejane et al. 2020).

For metacommunities, defined as sets of local communities connected by dispersal (Leibold et al. 2004), fragmentation alters the dispersal of organisms among localities (Fagan 2002), which can alter metacommunity dynamics (Cañedo-Argüelles et al. 2015). Using the biogeographical framework of distance decay relationships (Nekola et al. 1999), many studies have identified local, niche-based processes as the main determinants of freshwater metacommunity dynamics (Heino et al. 2015). In contrast, metacommunities in fragmented freshwater systems can be structured by dispersal-based processes, as an imprint of fragmentation (Thrush et al. 2008, Brown et al. 2011, Sarremejane et al. 2017, Gauthier et al. 2020). However, whether this imprint is stable or varies over time, notably according to flow variability and wet-dry cycles, remains unknown (Ruhi et al. 2017). During drying periods, local environmental conditions can change dramatically (e.g. increased temperature, biotic interaction changes, decreased dissolved oxygen), which could favour niche-based processes. In contrast, during rewetting phases, dispersal could become the major determinant of metacommunity structure due to active and passive recolonization (Sarremejane et al. 2017). Characterizing temporal

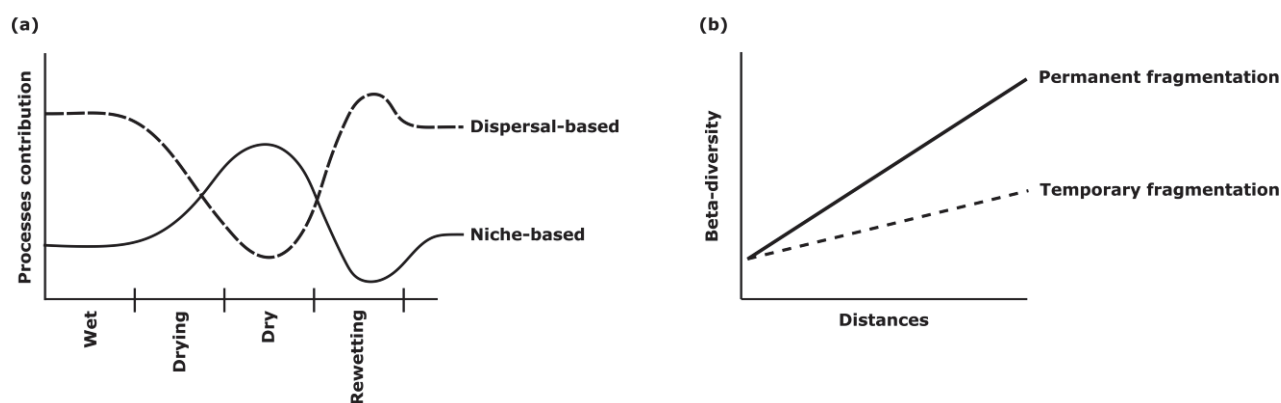


Figure 3.1 : Theoretical contribution of dispersal-based and niche-based processes on metacommunity assembly in fragmented river systems across hydrological phases (a) and expected beta-diversity relationships with permanent (i.e. artificial or natural barriers) and temporary (i.e. drying events) in such systems (b).

variability in metacommunity processes in river networks prone to drying will promote more efficient monitoring and management of their biodiversity.

How permanent and temporary fragmentations of river networks jointly shape aquatic metacommunity patterns has not been explored, leaving open questions and important implications for biodiversity conservation. When man-made structures permanently limit aquatic organisms' dispersal, the effect of drying on metacommunity dynamics could be reduced as it represents only temporary fragmentation. This is because during flowing periods, the movement of organisms is already constrained by man-made structures and, thus, temporary dispersal limitation could be undetectable. In contrast, where drying extends over long distances (e.g. several kilometres), dry sections could represent stronger dispersal barriers than man-made structures, particularly for strictly aquatic organisms. Therefore, contextualizing the effects of drying in interaction with man-made structures on riverine biodiversity is fundamental in the Anthropocene, as it could lead to contrasted conservation strategies.

Here, we explored the temporal dynamics of aquatic invertebrate metacommunities in a river network prone to drying and severe fragmentation by man-made structures through the biogeographical distance decay lens. We first hypothesized that processes structuring metacommunity would change within a hydrological year (Figure 3.1a). Accordingly, we predicted niche-based process to be the predominant determinant of metacommunity dynamics during drying periods, as shifts from lotic to lentic environmental conditions develop. Conversely, we predicted dispersal-based process to be the predominant upon flow resumption when organisms recolonize previously dry habitats from perennial refuges in the network. Our second hypothesis was that permanent fragmentation by man-made structures would be the primary factor determining metacommunity dynamics in this network (Figure 3.1b). Accordingly, we predicted stronger relationships between community dissimilarity and physical distances representing permanent fragmentation compared with those including drying events. To test these predictions, we coupled an intensive sampling campaign in both time and space to a fine-scale description of permanent and temporary fragmentation patterns in the Thouaret river network, western France.

3.4 Material & methods

3.4.1 Study area

The Thouaret river network is a sub-tributary of the Loire river in the Deux-Sèvres department in western France (Figure 3.2a). Its hydrological regime reflects the region's oceanic climate with rainy and cold winters (mean \pm SD cumulative 1958–2019 rainfall 233 ± 87 mm and temperature $5.16 \pm 1.37^\circ\text{C}$; Vidal et al. 2010) and hot and dry summers (mean cumulative rainfall for the period 1958–2019: 135 mm (± 59 SD); mean temperature for the period 1958–2019: 18.07°C (± 1.25 SD)). The impermeable granitic bedrock of the region results in flashy flood events and a lack of baseflow during

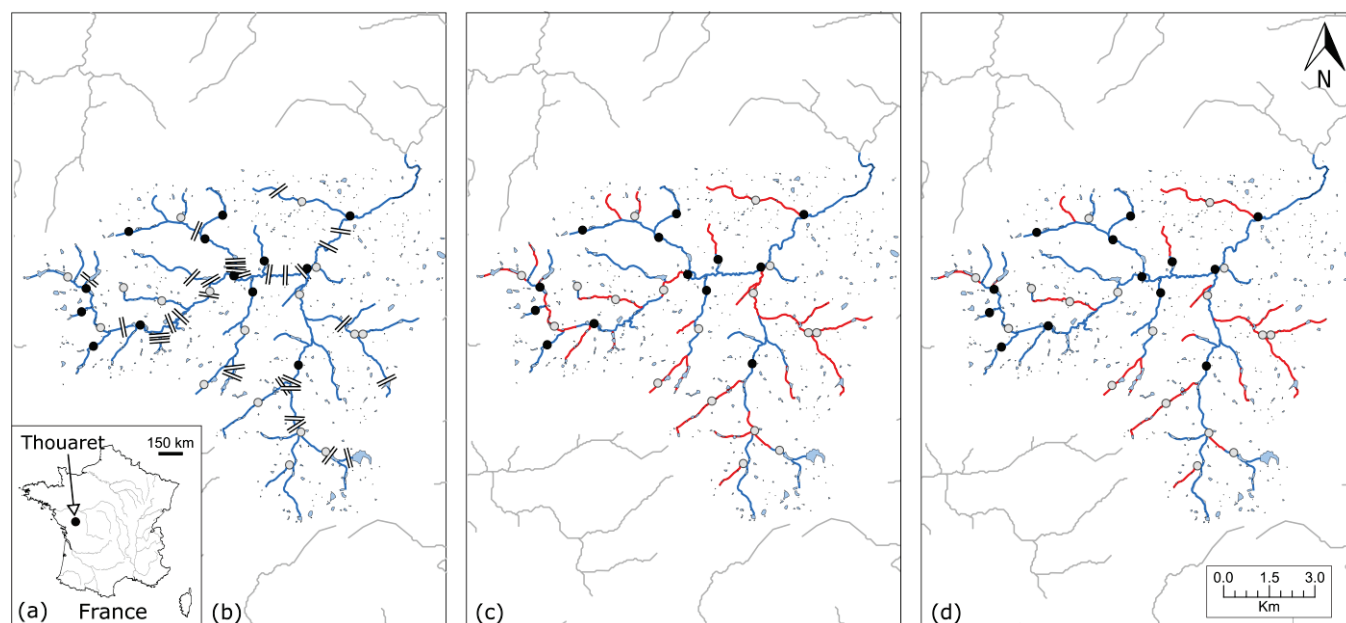


Figure 3.2 : The study river network, the Thouaret, is situated in Western France (a) at different hydrological phases (b,c,d) with site locations, man-made structures and drying patches locations. (b) refers to hydrological phases for which the entire catchment is flowing [i.e. wet (March, April), drying (June) and rewetting (December)], (c) and (d) to the dry phase [(c) September and (d) October]. Grey and black circles correspond to intermittent and perennial sampling sites, respectively. Red and blue lines correspond to dry and flowing reaches, respectively. Small water retention structures are also shown in light blue. Man-made structures are only presented in (b) by black double lines on the network for sake of clarity.

summer. Hydrological variability is thus considerable, ranging between flooding and severe summer low flows and associated drying events (Datry et al. 2016). The 96.0-km² Thouaret headwater network drains an area dominated by agriculture including pasture (35.4%), arable land (28.3%), complex cultivation patterns (23.9%), forest (10%) and urban areas (2.4%). Thousands of small water retention structures have been constructed for agriculture and public water supply in the past decades (exact installation dates are unknown), and some small lentic water bodies along and within the river network are used for recreation and aquaculture activities. These structures permanently fragment the network (Figure 3.2b). In addition, dry reaches temporarily fragment the network during low-flow periods between August and October (Figure 3.3).

3.4.2 Quantification of drying

Drying events were quantified by cross-referencing three different information sources. First, visual observations were made every two weeks by volunteer members of Carg'eau association (Collectif Associatif Régional pour la Gestion de l'EAU; www.eau-poitou-charentes.org) from June to October for the Thouaret in 2018. Observations were classified as one of four states: flowing, when water flowed across the whole channel width; low flow, when <50% of the reach width was flowing; no flow, when the reach was dry apart from disconnected pools; and dry, when surface water was absent. Observations were made along the entire network whenever possible, depending on accessibility.

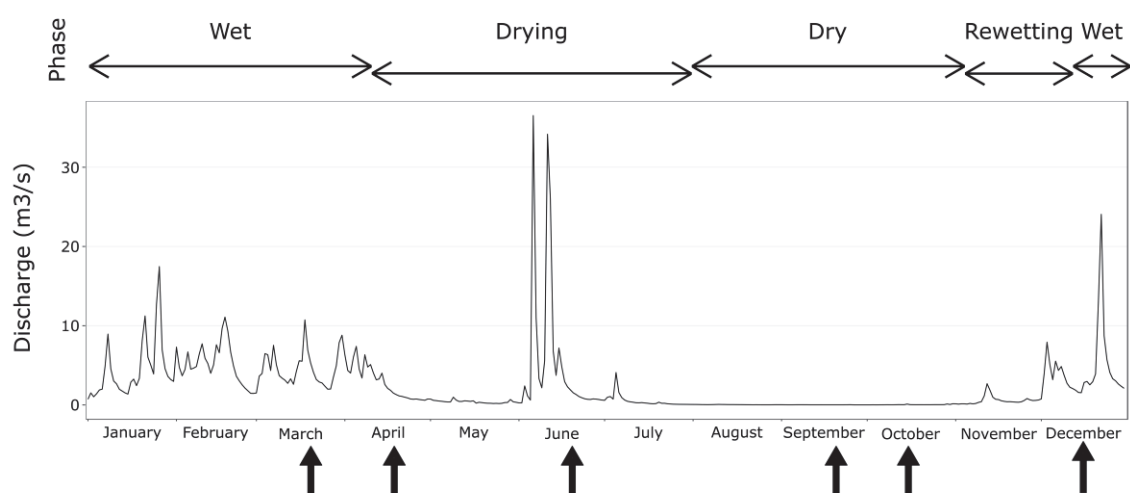


Figure 3.3 : Discharge time series for the year 2018 of a gauging station situated 55 km downstream of the studied Thouaret headwater. Hydrological phases are shown. Sampling dates are represented by vertical black arrows on x-axis. Data were downloaded from the Hydro database (www.hydro.eaufrance.fr): gauging station L8213010.

Datry et al. (2016) provide a detailed description of the method. Second, we visually assessed the same instream states at each sampling site and along each reach visited during the sampling campaigns, within < 4 days of the volunteers' observations. Finally, water presence loggers (Onset Hobo®, Intermountain Environmental, Logan, UT) were installed on the bed at each sampling site in riffle heads, from which dry first. The presence or absence of surface water was recorded every two hours from March 2018 to December 2018.

3.4.3 Sampling design

(a) Site selection

Thirty sampling sites were selected to encompass spatial variability within the Thouaret network, including sites in 13 perennial and 17 intermittent reaches (Figure 3.2b,c,d), and to control for the effects of confluences and man-made structures on community composition. From March 2018 to December 2018, each site was sampled 6 times, if flowing, to characterize different hydrological phases. One sampling date was during the wet phase in March, two during the drying phase in April and June, two during the dry phase in September and October, and one in December during the rewetting phase (Figure 3.3).

(b) Characterization of environmental variables

At each sampling site and date, 12 environmental variables were measured. Water depth (cm), wetted width (cm), flow velocity (cm.s⁻¹, measured at 0.4× water depth with a Miniwater20 velocity meter;

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

Schiltknecht, Gossau, Switzerland), main and secondary substrate types were measured at each invertebrate sampling point (see below). Substrate type was classified into 14 classes (mud, silt, sand, gravel, pebbles, rocks, slab, litter, root tufts, wood, algae, helophyte, spermaphyte and bryophyte). At the sampling site, elevation (m), latitude and longitude were obtained with a Garmin GPS. Specific conductivity ($\mu\text{S}\cdot\text{cm}^{-1}$) and temperature were measured using an HQ14d meter, and pH and dissolved oxygen percentage were measured using an HQ40d multi meter (HACH, Loveland, Colorado, USA). Along an upstream distance defined as 20-fold the wetted width of the sampling point, percent canopy cover was assessed visually and channel slope quantified with a clinometer. These environmental variables allow defining fairly well the habitat of riffle-associated invertebrate communities (Karaouzas et al. 2019).

(c) Benthic invertebrate sampling

Organisms were sampled from riffle habitat using a Surber sampler (area 0.1 m², mesh size 500 μm) to explore between-date differences in metacommunity patterns (Arscott et al. 2010, Datry et al. 2014, Ruhi et al. 2017, Crabot et al. 2020). This sampling design was validated from a metacommunity perspective through preliminary analysis of two other datasets (see Appendix A2 in Supporting Information). We focused on riffles, which are the first instream habitats to dry. Samples were preserved in 96% ethanol and organisms counted and identified to the lowest practical taxonomic level using available identification keys (list of references for identification available on Zenodo, <https://doi.org/10.5281/zenodo.3569563>). Most insects, crustaceans, triclads and all mollusks were identified to species, except for immature specimens lacking identification features. At a minimum, Diptera were identified to genus, sub-family or tribe depending on family; Coleoptera, Hemiptera, Trichoptera and Crustacea (Amphipoda) were identified to genus, and Ostracoda and Copepoda were identified to order. Achaeta and Tricladida were identified to genus, Nematoda, Acarina, Hydrozoa and Oligochaeta to order or sub-order. Each taxon was assigned to one of two dispersal categories according to Tachet et al. (2000): strictly aquatic invertebrates that can only disperse in water (57 taxa), and those with an aerial life stage (110 taxa) which can also disperse by flying or crawling.

3.4.4 Data analysis

(a) Description of permanent and temporary fragmentation types

To characterize permanent fragmentation by man-made structures, we calculated density (number of structures/km) and mean structure height. To characterize temporary fragmentation, the number of dry reaches, the mean and total length of dry reaches, the percentage of the catchment were calculated as in Datry et al. 2016). Logged water presence data were used to calculate the dry duration at each site, defined as the time in seconds without flowing water (i.e. no-flow and dry states).

All measurements related to man-made structures were based on a national ROE database (<http://www.sandre.eaufrance.fr/atlas/srv/fire/catalog.search#/metadata/59057026-b40c-4cf9-9e3e-7296e0aa1a78>). Dry patch delimitations were determined using flow state observations of volunteers validated by our visual assessments and logged water presence data and were integrated in Carthage network (<https://geo.data.gouv.fr/fr/datasets/54917fd94fb1cd2fcb6f5d3295dbf33acaf1847e>). All network measurements were then calculated with QGIS v3.10 modules postres, postgis and pgrouting.

(b) Hypothesis 1: Temporal variation of metacommunity processes within a hydrological year

To assess the metacommunity patterns across hydrological phases, Mantel tests (Legendre and Legendre 2012) were used to analyse the relationships between community beta diversity and environmental and spatial distances between pairwise combinations of sites on each sampling date. Beta-diversity matrices based on the Chao dissimilarity index were created to measure dissimilarity after $\log(x+1)$ transformation of abundance data to better represent rare taxa (Cañedo-Argüelles et al. 2015). Beta-diversity matrices were calculated for each sampling date, for the whole community and assemblages of taxa in the two dispersal categories (i.e. strictly aquatic taxa and taxa with an aerial life stage). Environmental distance matrices were constructed using Euclidean distances based on the 12 environmental variables and the dry duration, after variable standardization ($\text{mean} \pm 1 \text{ SD} = 0$) with the `vegdist` function in `vegan` (Oksanen et al. 2019). Two spatial distance matrices commonly used in stream metacommunity research were assessed (e.g. Sarremejane et al. 2017): an overland matrix based on Euclidean distances, constructed from site GPS coordinates with `vegdist`, and a network matrix based on instream distances within the river network constructed using the theoretical hydrographic network described by Pella et al. (2012). For environmental distances, the Moran spectral randomization procedure was used to account for spatial patterns of the environment and to avoid inflation of correlations in Mantel tests (Crabot et al. 2019). The Holm correction was used to adjust p-values for multiple comparisons in Mantel tests (Heino et al. 2017).

(d) Hypothesis 2: Effects of permanent and temporary fragmentation on metacommunity structure

To assess the importance of both permanent and temporary fragmentation on metacommunity structure, Mantel tests were used to analyse relationships between matrices of community dissimilarities and two fragmentation distances. First, the correlation between two fragmentation matrices was explored to determine whether to use classic or partial (i.e. correcting for the influence of the other matrix) Mantel tests. The permanent fragmentation matrix was calculated by summing the height of man-made structures (ROE database) between pairwise combinations of sites using QGIS v3.10. Height can impact the ability of organisms to pass a structure (Fuller et al. 2015). Although this

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

approach did not distinguish the effect of one large dam from that of several smaller structures, cumulative height can be more relevant than number of structures when studying network-scale dispersal (Perkin et al. 2015). The temporary fragmentation matrix was calculated using QGIS v3.10 by summing the length of dry patches observed in September between pairwise combinations of sites. Dry patches in September were used because they represent the maximum observed fragmentation level encountered in the Thouaret during the study. Mantel tests were thus performed for matrices including the final three sampling dates to be ecologically meaningful.

3.5 Results

3.5.1 Environmental variability

We observed considerable spatial variability in environmental variables on each sampling date. The network comprised sites experiencing a wide range of hydraulic conditions, including small and shallow to large and deep reaches (e.g. June wetted widths of 4.69–419.5 cm and depths of 0.61–25.0 cm), as well as standing to fast-flowing conditions (e.g. June flow velocity in: 0–0.25 m.s⁻¹). Physicochemical variables were also spatially variable, except for pH, and canopy cover encompassed both unshaded and forested reaches (Tableau 3.1). As is typical in riffle habitats, the main substrates were pebbles and gravel, across all sites and dates. The secondary substrate type was mainly composed of gravel, followed by sand and pebble (Tableau 3.1).

Tableau 3.1. Environmental variables measured for each sampling date. Mean (\pm SD) are shown, with the range for each variable [*Minimum – Maximum*]. For the main and secondary substrata, percentage for each substratum type is provided.

Sampling dates	March	April	June	September	October	December
Hydrologic phase	Wet	Drying	Drying	Dry	Dry	Rewetting
Altitude (m)	163.03 (19.23) [126 - 194]					
Slope (%)	1.27 (0.64) [0.5 - 3]					
Depth (cm)	22.64 (10.72) [5 - 46]	16.9 (9.8) [4 - 45]	14.2 (6.12) [0.61 - 25]	6.03 (2.92) [2 - 12]	7.02 (3.18) [2 - 14]	16.43 (8.68) [3 - 44]
Wetted width (cm)	230.77 (163.63) [51.3 - 750]	179.72 (117.44) [45.4 - 429.1]	162.77 (111.54) [4.69 - 419.5]	116.88 (113.09) [30 - 457.3]	118.16 (104.47) [30.1 - 424.1]	191.17 (128.87) [30.1 - 532.3]
Flow velocity (m.s ⁻¹)	0.44 (0.19) [0.21 - 0.8]	0.4 (0.26) [0.1 - 1.22]	0.43 (0.19) [0.11 - 0.72]	0.1 (0.08) [0 - 0.25]	0.19 (0.15) [0.03 - 0.66]	0.56 (0.31) [0.02 - 1.15]
Temperature (°C)	9.89 (1.15) [7.1 - 11.3]	13.71 (2.5) [9 - 17.7]	18.88 (1.22) [16.3 - 20.7]	17.04 (0.85) [15.8 - 19]	14.18 (1.18) [12.4 - 16.6]	7.25 (0.7) [6 - 8.2]

Conductance (uS.cm ⁻¹)	225.89 (55.55) [96.3 - 366]	230.55 (41.5) [116.5 - 338]	245.19 (46.56) [177.6 - 373]	333.21 (129.16) [183.2 - 685]	346.95 (96.38) [192.8 - 537]	271.82 (76.71) [32.5 - 476]
Dissolved oxygen (%)	94.2 (6.32) [83.5 - 108.8]	87.87 (18.41) [10.26 - 102.9]	82.77 (7.74) [64.8 - 97.2]	43.51 (28.89) [0.6 - 89.1]	53.72 (22.45) [1.5 - 87.4]	88.28 (6.88) [62.3 - 96.6]
pH	7.42 (0.31) [6.81 - 8.18]	7.47 (0.22) [6.91 - 7.76]	7.5 (0.45) [6.86 - 8.5]	7.42 (0.17) [7.09 - 7.71]	7.17 (0.26) [6.39 - 7.61]	7.29 (0.31) [6.67 - 8.01]
Canopy cover (%)	[0-25 - 25-50]	[0-25 - 25-50]	[0-25 - 75-100]	[25-50 - 75-100]	[0-25 - 75-100]	[0-25 - 75-100]
Main substratum type (%)	Pebbles 53.6	Pebbles 53.3	Pebbles 53.3	Pebbles 46.7	Pebbles 40.9	Pebbles 46.7
	Gravels 28.6	Gravels 33.3	Gravels 30	Gravels 26.7	Gravels 31.8	Gravels 23.3
	Sand 10.7	Sand 3.3	Sand 6.7	Mud 13.3	Mud 9.1	Wood 13.3
	Slab 3.6	Mud 3.3	Mud 6.7	Litter 6.7	Sand 9.1	Helophytes 6.7
	Mud 3.6	Helophytes 3.3	Helophytes 3.3	Sand 6.7	Silt 4.5	Slab 6.7
	Litter 3.3			Litter 4.5	Sand 3.3	

3.5.2 Permanent and temporary fragmentation of the Thouaret network

We recorded 32 man-made structures corresponding to a density of 0.32 man-made structures per linear kilometre (Tableau 3.2). Dams were the most abundant structures (n = 15), supporting water retention for agriculture and aquaculture purpose (60%) and recreation activities (40%). Other man-made structures comprised weirs (n = 12), bridge inverts (n = 3) and culverts (n = 2). Man-made structure heights ranged from 0.2 to 4.5 m (mean \pm SD, 1.69 \pm 1.34 m). Temporary fragmentation by dry reaches was observed on both dry phase sampling dates, in September and October (dry phase, Figure 3.2c,d). There were 20 dry reaches in September and 14 in October, with a mean length of 20.57 and 20.34 km, respectively. Dry reaches represented 46.09% and 32.63% of the network in September and October, respectively. Drying duration ranged from 5 hours to 3 months (mean duration: 4.2 weeks).

Tableau 3.2. Descriptors of permanent and temporary fragmentation in the Thouaret network. For the temporary fragmentation, values for both dry phase sampling dates are provided.

Fragmentation type	Variables	Value	
Permanent by man-made structures	Number of structures	32	
	Density (number of structures per linear km)	0.37	
	Height of structures in meter (mean)	1.69 (\pm 1.34 SD)	
Temporary by dry reaches	Variables	September	October
	Number of dry patches	20	14
	Average length of dry patch (km)	20.57	20.34
	Total length of dry reaches (km)	40.69	28.80
	Percentage of the catchment (%)	46.09	32.63

3.5.3 Invertebrate communities

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

A total of 74,337 individuals from 167 taxa were collected, with a mean \pm SD of $15,723 \pm 3,188$ individuals per date (Appendix A1) and of 1226 ± 1146 individuals from 19 ± 6 taxa per sample. On average, $58 \pm 2.8\%$ of taxa had an aerial stage and $42 \pm 2.8\%$ were strictly aquatic. Communities were dominated by Oligochaeta (relative abundance: 32.26 %), Tanytarsini (Diptera, 9.05%) and Gammaridae (Amphipoda, 8.84%). Oligochaeta was also the most ubiquitous taxon, occurring at 92% of all sites, followed by Orthoclaadiinae (Diptera, 84.4%) and Tanytarsini (Diptera, 72.5%). Taxa richness was 21.29 ± 5.78 taxa per sample across all sites and dates (Appendix A1) and were comparable at intermittent and perennial sites (Student test, $t=-0.006$, $p=0.99$). The highest average taxa richness was encountered in September (26.87), during the dry period, for which flowing reaches can act as refuges (Bonada et al. 2017) and density of species is high due to habitat area decrease (Dewson et al. 2007).

3.5.4 Hypothesis 1: Temporal variation of metacommunity processes

Our first hypothesis was partly supported: the processes determining aquatic invertebrate metacommunities varied during the hydrological year. Overland distances were strongly associated with community dissimilarities during the dry phase (mean $r = 0.24$), whereas associations were moderate during the rewetting (mean $r = 0.13$) and wet phases ($r = 0.12$), and no relationships were found during the drying phase. The relationships between community dissimilarities and overland distances were significant for strictly aquatic taxa but not for taxa with an aerial life stage (Figure 3.4, Tableau 3.3). Community dissimilarities were not associated with environmental or network distances on any date, indicating that niche-based processes were not important determinants of metacommunity dynamics in this network based on the 12 measured variables (Figure 3.4, Tableau 3.3).

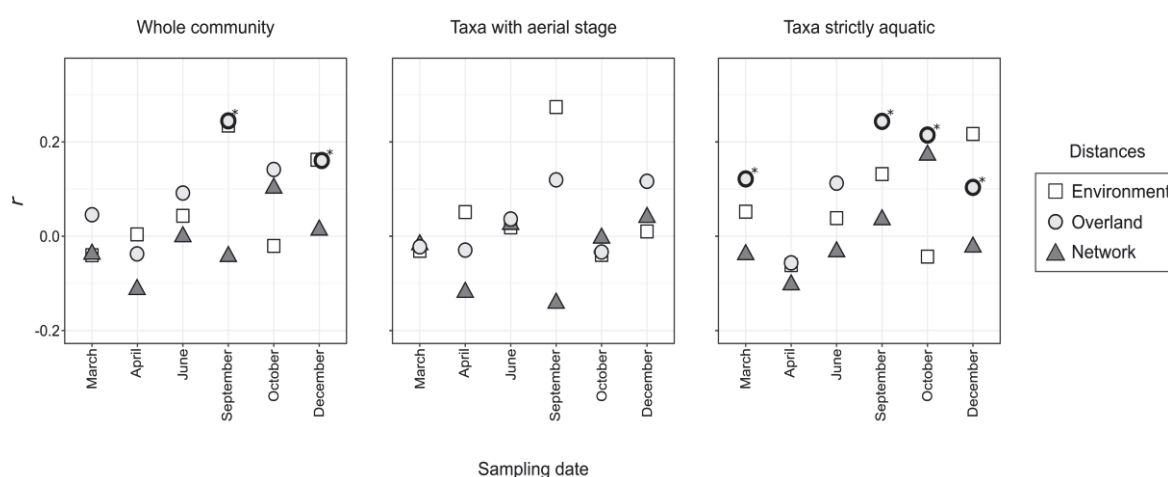


Figure 3.4 : Strength of the relationship (i.e. r values) between community dissimilarities and typical distances (environmental, overland and network distances) used in metacommunity studies through the different sampling dates for the whole community (left), the community of taxa with aerial life stage (middle) and strictly aquatic taxa (right). Significant relationships are shown with thicker outlines and a star. r values of all Mantel tests are available in Tableau 3.3.

Tableau 3.3. *r* values from the Mantel tests based on Chao dissimilarity index and typical distances used in metacommunity studies. With respect to distances: Envi, environmental; Over, overland; Netw, network. Significant relationships are shown in bold (*p*-value $\leq 0.05^*$)

Sampling date	March			April			June			September			October			December		
Hydrologic phase	Wet			Drying			Drying			Dry			Dry			Rewetting		
Distances	Envi	Over	Netw	Envi	Over	Netw	Envi	Over	Netw	Envi	Over	Netw	Envi	Over	Netw	Envi	Over	Netw
Whole community	-0.040	0.048	-0.038	0.004	-0.035	-0.113	0.043	0.094	-0.001	0.235	0.247*	-0.043	-0.021	0.144	0.102	0.162	0.163*	0.013
Taxa with aerial life stage	-0.031	-0.020	-0.018	0.051	-0.027	-0.119	0.019	0.039	0.025	0.274	0.122	-0.142	-0.040	-0.031	-0.004	0.010	0.119	0.040
Taxa strictly aquatic	0.052	0.124*	-0.039	-0.061	-0.054	-0.103	0.038	0.115	-0.033	0.132	0.246*	0.035	-0.043	0.217*	0.172	0.217	0.106*	-0.023

3.5.5 Hypothesis 2: Effect of permanent and temporary fragmentation on metacommunity dynamics

No significant correlation was found between permanent and temporary fragmentation matrices (Mantel test, $r = 0.10$, $p = 0.21$) and classic Mantel tests were used to explore the influence of fragmentation type on community dissimilarities. During the dry phase, permanent fragmentation by man-made structures explained more variance the metacommunity structure than the temporary fragmentation, supporting our second hypothesis. Despite low correlation coefficients, dissimilarities for the whole community and strictly aquatic taxa were significantly positively correlated with the permanent fragmentation matrix (mean $r = 0.20$, Figure 3.5, Tableau 4). In contrast, taxa with an aerial life stage had no significant relationships with permanent fragmentation matrices. Community dissimilarities were not correlated with temporary fragmentation matrices. Positive significant relationships were observed between community dissimilarities and permanent fragmentation before the dry phase (March, June) for strictly aquatic taxa, suggesting a long-term effect of man-made structures (Tableau 3.4).

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

Tableau 3.4. r values from the Mantel tests based on Chao dissimilarity index and fragmentation distances used in this study. With respect to distances: Perm. permanent fragmentation by man-made structures. Temp. maximal temporary fragmentation by dry reaches observed in 2018 (September). Significant relationships are shown in bold (p -value $\leq 0.05^*$).

Sampling date	March		April		June		September		October		December	
Hydrologic phase	Wet		Drying		Drying		Dry		Dry		Rewetting	
Distances	Perm	Temp	Perm	Temp	Perm	Temp	Perm	Temp	Perm	Temp	Perm	Temp
Whole community	0.103	-0.180	0.091	-0.267	0.130	-0.101	0.181	-0.146	0.209*	-0.024	0.177*	-0.129
Taxa with aerial life stage	0.049	-0.156	0.073	-0.250	0.071	-0.063	0.003	-0.138	0.062	0.006	0.130	-0.042
Taxa strictly aquatic	0.137*	-0.132	0.045	-0.209	0.166*	-0.121	0.207	-0.069	0.265*	-0.015	0.123	-0.135

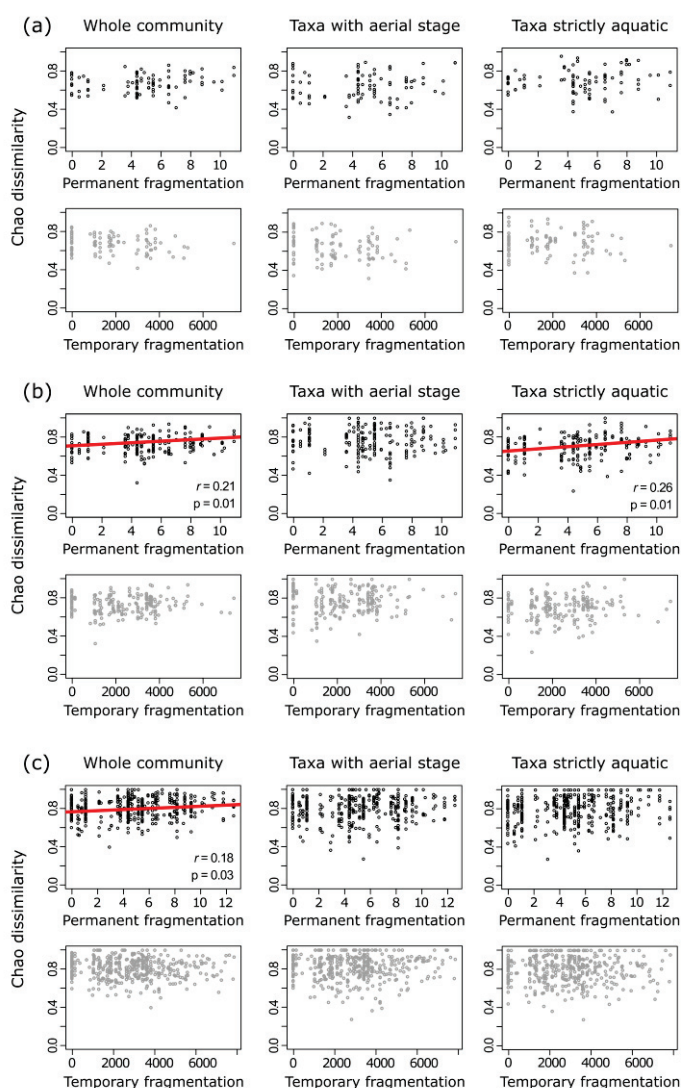


Figure 3.5 : Relationships between community dissimilarities (Chao dissimilarity index) and the two fragmentation matrices for (A) September, (B) October and (C) December sampling dates and for each community type. Permanent and temporary fragmentation matrices are shown in black and grey circles, respectively. Correspondence of sampling date and hydrological phases: September and October: Dry; December: Rewetting. Only significant relationships are shown (≤ 0.05) with corresponding exact r and p -values indicated. r values of all Mantel tests are available in Tableau 3.4.

3.6 Discussion

River network fragmentation is one of the most important anthropogenic alterations threatening freshwater biodiversity (Reid et al. 2019). Fragmentation alters the dispersal of organisms within river networks, and thus has cascading effects on metacommunity dynamics (Gauthier et al. 2020). However, to date, most studies have focused on the effects of permanent fragmentation by man-made structures, whereas the consequences of temporary fragmentation of river networks by drying have only recently been explored (Sarremejane et al. 2017, Gauthier et al. 2020). These recent studies suggest that the temporary fragmentation of river networks leads to shifts in metacommunity processes over time. By repeatedly quantifying the associations between metacommunity and fragmentation patterns in a river network affected by both permanent and temporary fragmentation, our results confirmed that metacommunity processes can greatly vary over time, following wet-dry cycles and associated changes in local environmental conditions and temporary fragmentation patterns at the network scale. We found that fragmentation by drying had little effect on metacommunity dynamics within a network subject to considerable permanent fragmentation by man-made structures, indicating that contextualization of the effects of fragmentation by drying is critical to protect and conserve biodiversity of river networks in the Anthropocene.

As hypothesized, processes structuring the invertebrate metacommunity varied through time across hydrological phases. This finding echoes recent conceptual (Datry et al. 2016), empirical (Sarremejane et al. 2017) and methodological (Jabot et al. 2020) calls to incorporate temporal variability into the metacommunity framework. These studies have also suggested that the relative importance of local and regional processes structuring metacommunities in dynamic river networks could experience dramatic shifts over time. Our metacommunity study, which to our knowledge is the first to describe the physical and biological model of a river network across multiple hydrological phases at such a fine spatiotemporal scale, confirmed that the relative roles of metacommunity processes change over time. We suggest that this conclusion may apply to other dynamic freshwater, terrestrial and marine ecosystems (e.g. Gerisch 2014, Dell et al. 2019). Although we acknowledge that the effort expended during our field campaigns and quantification of biodiversity through sample processing and invertebrate identification were substantial, novel approaches such as metabarcoding represent opportunities improve the efficiency of research into metacommunity ecology (Bush et al. 2020).

Although metacommunity processes varied over time, our conceptual model predicting shifts between local, niche-based and regional, dispersal-based processes was partially validated. First, environmental conditions defined by the 12 locally measured variables were not important in determining metacommunity structures in any hydrological phase. Nevertheless, we expected them to be influential during drying, when temperature, nutrient levels and interspecific competition typically increase and when dissolved oxygen can dramatically decrease (Datry et al. 2016). In the Thouaret network,

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

although environmental variables varied significantly in space and time, they did not vary considerably between dates and may not have reached ecologically relevant thresholds. For example, during the drying phase, recorded temperatures did not exceed 20.7°C at any site, far below the thermal threshold of most aquatic invertebrates (Quinn et al. 1994). We did not sample pool habitats, which could experience greater change in physicochemical conditions during drying; their inclusion may have indicated a stronger role of niche-based processes. However, the limited influence of niche-based processes on riffle communities in this fragmented river network suggests that metacommunities in intermittent river networks are predominantly controlled by dispersal. Many taxa have adaptations promoting resistance to drying, including the production of desiccation-tolerant dormant life stages (e.g. Stubbington and Datry 2013), which may partly explain the absence of relationships with local environmental conditions in this network. Such strategies are conceptualized as enabling ‘temporal dispersal’ in metacommunities (e.g. Datry et al. 2016x, Wisnoski et al. 2019).

Dispersal was the main process structuring invertebrate metacommunities in the Thouaret network, independently of hydrological phases. On four sampling dates, community dissimilarities were exclusively explained by physical, overland distances. However, they only explained a small proportion of total variation, which may reflect ecological drift (Silva et al. 2015), unmeasured environmental factors (e.g. pollutants, Jacobsen 1998) or fragmentation (see below). These results still confirmed recent studies proposing that metacommunities in dynamic river ecosystems are primarily structured by dispersal (Sarremejane et al. 2017, Gauthier et al. 2020). In intermittent river networks, aquatic communities experience locally wet-dry cycles which involve repeated extinction and recolonization events (Datry et al. 2016). At the river network scale, recurrent recolonization through active and passive dispersal from nearby refuges is a major influence on the organization of biodiversity in space and time (Crabot et al. 2020).

As hypothesized, the permanent fragmentation of the Thouaret network by man-made structures was the primary factor determining metacommunity dynamics, whereas temporary fragmentation by drying had little effect. We found more, and stronger, distance decay relationships between community dissimilarities and physical distances accounting for permanent fragmentation compared to distances including the spatial extent of drying. This suggests that small man-made structures strongly alter invertebrate species dispersal within our river network, as reported for other river metacommunities (Katano et al. 2006, Gauthier et al. 2020) and metapopulations (Blanchet et al. 2010, Taylor et al. 2018). The relationships, although significant, were not very strong, which could be because the high level of fragmentation lead to high isolation, thus increasing ecological drift and species turnover (Merritt et al. 2010).

Furthermore, the effects of river fragmentation are generally more evident for fish than invertebrates (Van Looy et al. 2014). This partly reflects the highly diverse dispersal modes and capacities of

aquatic invertebrates (Heino et al. 2017), which lead to contrasting responses to fragmentation. For example, we observed much stronger fragmentation effects for strictly aquatic invertebrate taxa than for those that have an aerial life stage. The latter, such as mayflies, caddisflies and non-biting midges, can colonize previously dry reaches by flying without being severely limited by man-made structures. In contrast, the dispersal of crustaceans, mollusks and others strictly aquatic taxa was limited, as also observed for fish. In addition to ongoing calls and efforts to better quantify the dispersal routes and capacities of aquatic invertebrates (Downes et al. 2017, Heino et al. 2017), assessing effective dispersal (i.e. dispersal followed by successful reproduction, ensuring gene flow) by measuring genetic variation among local populations represent the next step to understanding of metacommunity dynamics in fragmented systems and provide guidance for conservation plans (Coulon et al. 2010).

Reflecting the extensive permanent fragmentation of the Thouaret network by man-made structures, the effects of fragmentation by drying were not detected, highlighting the effects of drying on riverine biodiversity as context dependent. Previous studies reporting strong effects of fragmentation by drying on metacommunity dynamics were in systems subject to little fragmentation by man-made structures (e.g. Cañedo-Argüelles et al. 2015). In such cases, dry reaches limited aquatic invertebrate dispersal during dry phases, which cascaded onto metacommunity dynamics. In the Thouaret network, permanent fragmentation had a considerable effect on invertebrate dispersal to the extent that the effects of fragmentation by drying were either null or undetectable at the network scale. In contrast, local-scale observed differences between perennial and intermittent sites reflect the key role of drying in structuring aquatic communities (e.g. Leigh and Datry 2017, Ruhi et al. 2017).

Our study – which, to our knowledge, is the first to document the effects of both permanent and temporary fragmentation of riverine metacommunities – highlights the need to account for context-dependency in metacommunity ecology, in particular for conservation purposes. For example, we observed no additive effects of permanent and temporary fragmentation in any hydrological phases or for any community type (results not shown), contrasting with other fragmentation contexts in headwater streams of France (Gauthier et al, 2020). Our results also indicate the need to consider the complexity of the effects of drying in river ecology, particularly when exploring network-scale patterns and processes. Further research should explore the synergies and antagonisms between permanent and temporary fragmentation of river networks, which is critical for refining the management and conservation of the rivers of the Anthropocene.

Implications for the management and conservation of river networks

Our results add to recent calls to recognize the temporal variability of metacommunity processes in biodiversity research (Datry et al. 2016, Sarremejane et al. 2017, Jabot et al. 2020). That such processes can vary over short time scales in river networks, and most likely in other dynamic ecosystems (e.g. Fernández et al. 2014, Wolkovich et al. 2014), has multiple implications for

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

ecosystem management and conservation planning. For example, certain hydrological phases could be particularly relevant than others to assess the ecological status of monitoring sites, which is currently mostly based on niche-based processes (Cid et al. 2020). Identifying such key hydrological phases, as a parallel for keystone species or habitats (Kotliar 2000), and during which increased understanding of critical metacommunity processes is crucial to improve the success of management practices. Typically, conservation actions would also need to promote hydrological connectivity during dispersal-limited phases. In cases where fragmentation is mostly due to drying, limiting surface and groundwater abstraction during summer could ensure sufficient connectivity within drying river networks (de Graaf et al. 2019). Last, but not least, many conservation tools and software (e.g. Marxan, Zonation) are static and then are poorly suited to highly dynamic ecosystems such as most river networks (Hermoso et al. 2018).

In the Anthropocene, river networks are prone to increasing fragmentation by drying events and the development of small water retention structures for irrigation and human needs (Reid et al. 2019). While previous studies showed little or no effect of fragmentation by man-made structures on invertebrate communities compared to fish communities (Van Looy et al. 2014, Hwan et al. 2016), our metacommunity approach indicates that small water retention structures can have severe effects on biodiversity patterns. More importantly, the cumulative effect of such structures on riverine biodiversity have not yet been investigated, but could be substantial (Linares et al. 2019). As small water retention structures are increasingly promoted as an adaptation strategy to climate change (Habets et al. 2013), we argue that they might considerably alter metacommunity dynamics, resulting in riverine biodiversity loss. Further, determining tipping points of fragmentation levels on metacommunity dynamics would allow prioritization of management efforts, notably with respect to restoration. For example, identifying structure characteristics (e.g. height, passability, Fuller et al. 2015) and their associated effects on metacommunities, as well as determining what fragmentation level is acceptable without altering dispersal among local communities could be done by gathering similar datasets as the one presented here. In turn, generating such fragmentation-ecology relationships could reconcile water uses and biodiversity, as well as optimizing the conservation of river networks of the Anthropocene.

3.7 Acknowledgements

This work was supported by the French Agency for Biodiversity (AFB) through the program n° 14 entitled “Headwater biodiversity dynamics”. We thank the AFB members, in particular Claire Magand and Claire Bramard, for their help to select and expertise on Poitou-Charente rivers. We thank the Nouvelle-Aquitaine Regional Agency for Biodiversity, particularly Arnaud Vaudelet, and Hervé Pella and Jean-Philippe Vidal for their help on GIS and fishing federation data. We thank Julie Crabot, Christel Blot and Benjamin Gerfand for help with sample collection and processing. We thank Maxence Forcellini, Julien Barnasson and Jacques Mouthon for identification expertise on Trichoptera and Mollusks. We thank Julie Crabot and Nuria Cid-Puey for useful comments on hypothesis and data analysis. This manuscript benefited from the meaningful comments from Rachel Stubbington, Romain Sarremejane, Jani Heino and 2 anonymous reviewers and the associate editor.

3.8 Data accessibility

Mean value of abundance for each taxon and each sampling date (Appendix A1), supplementary analyses to test for the reliability of the study sampling design (Appendix A2) and outputs of non-parametric permutational multivariate analysis of variances testing differences in community composition between intermittent and perennial sites (Appendix A3) are available online.

3.9 References

- Arcscott, D. B., S. Larned, M. R. Scarsbrook, and P. Lambert. 2010. Aquatic invertebrate community structure along an intermittence gradient: Selwyn River, New Zealand. *Journal of the North American Benthological Society* 29:530-545.
- Blanchet, S., O. Rey, R. Etienne, S. Lek, and G. Loot. 2010. Species-specific responses to landscape fragmentation: Implications for management strategies. *Evolutionary Applications* 3:291-304.
- Boulton, A. J. 2003. Parallels and contrasts in the effects of drought on stream macroinvertebrate assemblages. *Freshwater Biology* 48:1173-1185.
- Bush, A., W. A. Monk, Z. G. Compson, D. L. Peters, T. M. Porter, S. Shokralla, M. T. G. Wright, M. Hajibabaei, and D. J. Baird. 2020. DNA metabarcoding reveals metacommunity dynamics in a threatened boreal wetland wilderness. *Proc Natl Acad Sci U S A*.
- Cañedo-Argüelles, M., K. S. Boersma, M. T. Bogan, J. D. Olden, I. Phillipsen, T. A. Schriever, D. A. Lytle, and R. Ladle. 2015. Dispersal strength determines meta-community structure in a dendritic riverine network. *Journal of Biogeography* 42:778-790.

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

Coulon, A., J. W. Fitzpatrick, R. Bowman, and I. J. Lovette. 2010. Effects of Habitat Fragmentation on Effective Dispersal of Florida Scrub-Jays. *Conservation Biology* 24:1080-1088.

Crabot, J., J. Heino, B. Launay, and T. Datry. 2019. Drying determines the temporal dynamics of stream invertebrate structural and functional beta diversity. *Ecography*:1-16.

Crabot, J., S. Clappe, S. Dray, T. Datry, and R. McCrea. 2019. Testing the Mantel statistic with a spatially- constrained permutation procedure. *Methods in Ecology and Evolution* 10:532-540.

Datry, T., N. Bonada, and J. Heino. 2016. Towards understanding the organisation of metacommunities in highly dynamic ecological systems. *Oikos* 125:149-159.

Datry, T., N. Bonada, and A. J. Boulton. 2017. *Intermittent Rivers and Ephemeral Streams: Ecology and Management*. Academic Press.

Datry, T., H. Pella, C. Leigh, N. Bonada, and B. Hugueny. 2016. A landscape approach to advance intermittent river ecology. *Freshwater Biology* 61:1200-1213.

Datry, T., S. T. Larned, K. M. Fritz, M. T. Bogan, P. J. Wood, E. I. Meyer, and A. N. Santos. 2014. Broad-scale patterns of invertebrate richness and community composition in temporary rivers: effects of flow intermittence. *Ecography* 37:94-104.

Datry, T., A. J. Boulton, N. Bonada, K. Fritz, C. Leigh, E. Sauquet, K. Tockner, B. Hugueny, and C. N. Dahm. 2018. Flow intermittence and ecosystem services in rivers of the Anthropocene. *J Appl Ecol* 55:353-364.

de Graaf, I. E. M., T. Gleeson, L. P. H. Rens van Beek, E. H. Sutanudjaja, and M. F. P. Bierkens. 2019. Environmental flow limits to global groundwater pumping. *Nature* 574:90-94.

Dell, J. E., D. M. Salcido, W. Lumpkin, L. A. Richards, S. M. Pokswinski, E. L. Loudermilk, J. J. O'Brien, and L. A. Dyer. 2019. Interaction Diversity Maintains Resiliency in a Frequently Disturbed Ecosystem. *Frontiers in Ecology and Evolution* 7.

Döll, P., and H. M. Schmied. 2012. How is the impact of climate change on river flow regimes related to the impact on mean annual runoff? A global-scale analysis. *Environmental Research Letters* 7.

Downes, B. J., J. Lancaster, A. Glaister, and W. D. Bovill. 2017. A fresh approach reveals how dispersal shapes metacommunity structure in a human-altered landscape. *Journal of Applied Ecology* 54:588-598.

Fernández, M. E., J. E. Gyenge, S. Varela, and M. de Urquiza. 2014. Effects of the time of drought occurrence within the growing season on growth and survival of *Pinus ponderosa* seedlings. *Trees* 28:745-756.

Fuller, M. R., M. W. Doyle, and D. L. Strayer. 2015. Causes and consequences of habitat fragmentation in river networks. *Ann N Y Acad Sci* 1355:31-51.

Gauthier, M., B. Launay, G. Le Goff, H. Pella, C. J. Douady, and T. Datry. 2020. Fragmentation promotes the role of dispersal in determining 10 intermittent headwater stream metacommunities. In review.

Gerisch, M. 2014. Non-random patterns of functional redundancy revealed in ground beetle communities facing an extreme flood event. *Functional Ecology* 28:1504-1512.

Grill, G., B. Lehner, M. Thieme, B. Geenen, D. Tickner, F. Antonelli, S. Babu, P. Borrelli, L. Cheng, H. Crochetiere, H. Ehalt Macedo, R. Filgueiras, M. Goichot, J. Higgins, Z. Hogan, B. Lip, M. E. McClain, J. Meng, M. Mulligan, C. Nilsson, J. D. Olden, J. J. Opperman, P. Petry, C. Reidy Liermann, L. Saenz, S. Salinas-Rodriguez, P. Schelle, R. J. P. Schmitt, J. Snider, F. Tan, K. Tockner, P. H. Valdujo, A. van Soesbergen, and C. Zarfl. 2019. Mapping the world's free-flowing rivers. *Nature* 569:215-221.

Habets, F., J. Boé, M. Déqué, A. Ducharne, S. Gascoïn, A. Hachour, E. Martin, C. Pagé, E. Sauquet, L. Terray, D. Thiéry, L. Oudin, and P. Viennot. 2013. Impact of climate change on the hydrogeology of two basins in northern France. *Climatic Change* 121:771-785.

Heino, J., and K. T. Tolonen. 2017. Ecological drivers of multiple facets of beta diversity in a lentic macroinvertebrate metacommunity. *Limnology and Oceanography* 62:2431-2444.

Heino, J., A. S. Melo, T. Siqueira, J. Soininen, S. Valanko, and L. M. Bini. 2015. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Freshwater Biology* 60:845-869.

Heino, J., J. Alahuhta, T. Ala-Hulkko, H. Antikainen, L. M. Bini, N. Bonada, T. Datry, T. Erős, J. Hjort, O. Kotavaara, A. S. Melo, and J. Soininen. 2017. Integrating dispersal proxies in ecological and environmental research in the freshwater realm. *Environmental Reviews* 25:334-349.

Hermoso, V., A. Morán-Ordóñez, and L. Brotons. 2018. Assessing the role of Natura 2000 at maintaining dynamic landscapes in Europe over the last two decades: implications for conservation. *Landscape Ecology* 33:1447-1460.

Hwan, J. L., and S. M. Carlson. 2016. Fragmentation of an Intermittent Stream During Seasonal Drought: Intra-annual and Interannual Patterns and Biological Consequences. *River Research and Applications* 32:856-870.

Jabot, F., F. Laroche, F. Massol, F. Arthaud, J. Crabot, M. Dubart, S. Blanchet, F. Munoz, P. David, and T. Datry. In Press. Assessing metacommunity processes through signatures in spatiotemporal turnover of community composition. *Ecology letters*.

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

Karaouzas, I., C. Theodoropoulos, A. Vourka, K. Gritzalis, and N. T. Skoulikidis. 2019. Stream invertebrate communities are primarily shaped by hydrological factors and ultimately fine-tuned by local habitat conditions. *Sci Total Environ* 665:290-299.

Katano, O., T. Nakamura, S. Abe, S. Yamamoto, and Y. Baba. 2006. Comparison of fish communities between above- and below-dam sections of small streams; barrier effect to diadromous fishes. *Journal of Fish Biology* 68:767-782.

Kotliar, N. B. 2000. Application of the new keystone-species concept to prairie dogs: How well does it work? *Conservation Biology* 14:1715-1721.

Legendre, P., and L. Legendre. 2012. *Numerical ecology*. 3 edition. English edn. Elsevier, Amsterdam.

Lehner, B., C. R. Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, J. Magome, C. Nilsson, J. C. Robertson, R. Rödel, N. Sindorf, and D. Wisser. 2011. High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment* 9:494-502.

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

Leigh, C., and T. Datry. 2017. Drying as a primary hydrological determinant of biodiversity in river systems: a broad-scale analysis. *Ecography* 40:487-499.

Linares, M. S., W. Assis, R. R. de Castro Solar, R. P. Leitão, R. M. Hughes, and M. Callisto. 2019. Small hydropower dam alters the taxonomic composition of benthic macroinvertebrate assemblages in a neotropical river. *River Research and Applications* 35:725-735.

Malavoi, J. R., and J. P. Bravard. 2010. *Eléments d'hydromorphologie fluviale*. Onema.

Nekola, J. C., and P. S. White. 1999. The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* 26:867-878.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlenn, Minchin P.R., R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner. 2019. *vegan: Community Ecology Package*. R package version 2.5-5.

Pella, H., J. Lejot, N. Lamouroux, and T. Snelder. 2012. Le réseau hydrographique théorique (RHT) français et ses attributs environnementaux. *Géomorphologie : relief, processus, environnement* 18:317-336.

- Perkin, J. S., K. B. Gido, A. R. Cooper, T. F. Turner, M. J. Osborne, E. R. Johnson, K. B. Mayes, and C. Nilsson. 2015. Fragmentation and dewatering transform Great Plains stream fish communities. *Ecological Monographs* 85:73-92.
- Quinn, J. M., G. L. Steele, C. W. Hickey, and M. L. Vickers. 1994. Upper thermal tolerances of twelve New Zealand stream invertebrate species. *New Zealand Journal of Marine and Freshwater Research* 28:391-397.
- Reid, A. J., A. K. Carlson, I. F. Creed, E. J. Eliason, P. A. Gell, P. T. J. Johnson, K. A. Kidd, T. J. MacCormack, J. D. Olden, S. J. Ormerod, J. P. Smol, W. W. Taylor, K. Tockner, J. C. Vermaire, D. Dudgeon, and S. J. Cooke. 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews* 94:849-873.
- Ruhi, A., T. Datry, and J. L. Sabo. 2017. Interpreting beta-diversity components over time to conserve metacommunities in highly dynamic ecosystems. *Conserv Biol* 31:1459-1468.
- Sarremejane, R., M. Cañedo-Argüelles, N. Prat, H. Mykrä, T. Muotka, and N. Bonada. 2017. Do metacommunities vary through time? Intermittent rivers as model systems. *Journal of Biogeography* 44:2752-2763.
- Tachet, H., P. Richoux, M. Bournaud, and P. Usseglio-Polatera. 2000. *Invertébrés d'eau douce: systématique, biologie, écologie*.
- Taylor, A. T., M. Papeş, and J. M. Long. 2018. Incorporating fragmentation and non-native species into distribution models to inform fluvial fish conservation. *Conservation Biology* 32:171-182.
- Van Looy, K., T. Tormos, and Y. Souchon. 2014. Disentangling dam impacts in river networks. *Ecological Indicators* 37:10-20.
- Vidal, J. P., E. Martin, L. Franchistéguy, M. Baillon, and J. M. Soubeyroux. 2010. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology* 30:1627-1644.
- Wolkovich, E. M., B. I. Cook, K. K. McLauchlan, and T. J. Davies. 2014. Temporal ecology in the Anthropocene. *Ecology Letters* 17:1365-1379.
- Zarfl, C., A. E. Lumsdon, J. Berlekamp, L. Tydecks, and K. Tockner. 2014. A global boom in hydropower dam construction. *Aquatic Sciences* 77:161-170.

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

3.10 Supplemental Information

Appendix S1. Mean value of abundance for each taxa and each sampling dates for the Thouaret headwater. Higher rank groups (family, order and class when possible) and mean specific richness are also provided.

Class	Order	Family	Taxa	Sampling dates					
				March	April	June	September	October	December
Arachnida	Hydracarina	Hydracarina	Hydracarina	0.11	0.177	0.148	0.105	0.099	0.082
Clitellata	Oligochaeta	Oligochaeta	Oligochaeta	32.468	29.797	28.886	27.78	38.645	38.751
Hirudinae	Arynchobdellida	Erpobdellidae	<i>Erpobdella</i>	0.268	0.064	0.201	1.237	0.553	0.405
Hirudinae	Rhynchobdellida	Glossiphoniidae	<i>Glossiphonia</i>	0	0.007	0.091	0.055	0.092	0.03
Hirudinae	Rhynchobdellida	Glossiphoniidae	<i>Glossiphonia verrucata</i>	0.118	0	0	0	0	0
Hirudinae	Rhynchobdellida	Glossiphoniidae	<i>Helobdella stagnalis</i>	0	0.007	0.01	0.022	0.007	0.045
Hydrozoa	Hydrozoa	Hydrozoa	Hydrozoa	0.3	5.297	0.329	0.066	0.026	0.247
Insecta	Coleoptera	Curculionidae	Curculionidae	0	0	0	0	0.007	0
Insecta	Coleoptera	Dryopidae	<i>Dryops</i>	0.103	0.057	0.062	0.006	0.026	0.142
Insecta	Coleoptera	Dytiscidae	<i>Agabus</i>	0.024	0.021	0	0	0	0.03
Insecta	Coleoptera	Dytiscidae	<i>Ilybius</i>	0	0	0	0.006	0	0.007
Insecta	Coleoptera	Dytiscidae	<i>Laccornis oblongus</i>	0	0	0.005	0	0	0
Insecta	Coleoptera	Dytiscidae	<i>Platambus</i>	0	0	0	0.028	0	0
Insecta	Coleoptera	Elmidae	<i>Elmis</i>	0.426	0.353	0.521	3.617	1.937	1.566
Insecta	Coleoptera	Elmidae	<i>Esolus</i>	0	0.021	0.005	0.138	0.481	0.052
Insecta	Coleoptera	Elmidae	<i>Limnius</i>	0	0	0	0.006	0.013	0
Insecta	Coleoptera	Elmidae	<i>Oulimnius</i>	0.11	0.064	0.177	1.568	0.988	0.352
Insecta	Coleoptera	Gyrinidae	<i>Gyrinus</i>	0	0	0	0.006	0	0
Insecta	Coleoptera	Helophoridae	<i>Helophorus</i>	0	0	0.005	0.006	0.007	0
Insecta	Coleoptera	Hydraenidae	<i>Hydraena</i>	0	0.014	0.01	0.326	0.303	0.067
Insecta	Coleoptera	Hydraenidae	<i>Limnebius</i>	0	0	0	0.006	0	0
Insecta	Coleoptera	Hydraenidae	<i>Ochthebius</i>	0	0	0	0.006	0	0
Insecta	Coleoptera	Hydrophilidae	<i>Anacaena</i>	0	0	0	0.006	0	0
Insecta	Coleoptera	Hydrophilidae	<i>Laccobius</i>	0	0.007	0	0	0	0
Insecta	Coleoptera	Scirtidae	<i>Elodes</i>	0.071	0.021	0.014	0.271	0.191	0.06
Insecta	Diptera	Anthomyidae	Anthomyidae	0	0	0.01	0.028	0.059	0
Insecta	Diptera	Athericidae	<i>Athricops crassipes</i>	0.016	0	0	0	0.02	0.022
Insecta	Diptera	Ceratopogonidae	<i>Atrichopogon</i>	0	0	0	0.006	0	0
Insecta	Diptera	Ceratopogonidae	Ceratopogoninae	1.136	1.787	0.191	0.453	0.395	3.432
Insecta	Diptera	Chaoboridae	<i>Chaoborus</i>	0	0	0	0	0.152	0
Insecta	Diptera	Chironomidae	<i>Chironomini</i>	0.308	1.243	1.16	1.165	1.436	0.644
Insecta	Diptera	Chironomidae	Orthocladinae	8.267	13.08	6.208	2.772	3.024	3.237

Chapitre 2 - Métacommunautés dans les têtes de bassin versant, le rôle de la dispersion et de la fragmentation

Insecta	Diptera	Chironomidae	Tanypodinae	0.11	0.438	0.726	0.723	0.613	0.217
Insecta	Diptera	Chironomidae	<i>Tanytarsini</i>	3.857	6.519	14.889	16.554	5.633	1.199
Insecta	Diptera	Culicidae	Culicidae	0	0	0.014	0.006	0.02	0
Insecta	Diptera	Diptera	Diptera	0	0	0	0	0	0.007
Insecta	Diptera	Dixidae	<i>Dixa</i>	0.016	0.014	0.033	0.392	0.152	0.015
Insecta	Diptera	Dixidae	<i>Dixella</i>	0	0	0.005	0	0.007	0.007
Insecta	Diptera	Dolichopodidae	Dolichopodidae	0	0.021	0	0	0.007	0
Insecta	Diptera	Empididae	Clinocerinae	0	0	0.005	0.011	0.013	0
Insecta	Diptera	Empididae	Hemerodromiinae	0.008	0.021	0.033	0.006	0	0
Insecta	Diptera	Limoniidae	<i>Dicranota</i>	0.008	0.127	0.372	0.133	0.158	0.172
Insecta	Diptera	Limoniidae	<i>Eloephila</i>	0.032	0.014	0	0.017	0.04	0.022
Insecta	Diptera	Limoniidae	<i>Eriopterini</i>	0.016	0.014	0.01	0	0.02	0.015
Insecta	Diptera	Limoniidae	<i>Hexatoma</i>	0	0.014	0.005	0.006	0.053	0.007
Insecta	Diptera	Limoniidae	<i>Hexatomini</i>	0.063	0	0.005	0	0.013	0.022
Insecta	Diptera	Limoniidae	Limoniidae	0.024	0.007	0	0	0	0
Insecta	Diptera	Limoniidae	<i>Pedicia</i>	0	0	0.014	0.039	0.013	0.03
Insecta	Diptera	Limoniidae	<i>Pilaria</i>	0.126	0.127	0.01	0.105	0.27	0.555
Insecta	Diptera	Psychodidae	Psychodidae	0.142	0.127	0.11	0.039	0.105	0.285
Insecta	Diptera	Ptychopteridae	<i>Ptychoptera</i>	0.189	0.007	0	0.033	0.171	0.052
Insecta	Diptera	Rhagionidae	Rhagionidae	0.039	0.021	0.01	0	0.007	0.067
Insecta	Diptera	Simuliidae	Simuliidae	15.713	2.444	5.692	1.242	0.217	8.161
Insecta	Diptera	Tabanidae	Tabanidae	0.047	0.071	0.019	0.232	0.329	0.547
Insecta	Diptera	Tipulidae	Tipulidae	0.016	0.007	0.048	0.033	0.013	0.037
Insecta	Ephemeroptera	Baetidae	<i>Baetis</i>	0.876	2.401	1.681	1.314	0.81	0.487
Insecta	Ephemeroptera	Caenidae	<i>Caenis horaria</i>	0.008	0	0	0	0	0
Insecta	Ephemeroptera	Caenidae	<i>Caenis luctuosa</i>	0	0	0.005	0	0.013	0.015
Insecta	Ephemeroptera	Ephemerellidae	<i>Serratella ignita</i>	0.355	4.909	0.606	0.011	0	0
Insecta	Ephemeroptera	Ephemeridae	<i>Ephemera</i>	0	0	0	0	0.007	0
Insecta	Ephemeroptera	Ephemeridae	<i>Ephemera danica</i>	0.008	0	0	0	0	0
Insecta	Ephemeroptera	Heptageniidae	<i>Ecdyonurus dispar</i>	0	0	0.005	0	0	0
Insecta	Ephemeroptera	Heptageniidae	<i>Ecdyonurus torrentis</i>	0	0	0.005	0	0	0
Insecta	Ephemeroptera	Heptageniidae	Heptageniidae	0.008	0.042	0	0	0	0.007
Insecta	Ephemeroptera	Leptophlebiidae	<i>Habrophlebia lauta</i>	0.426	0.636	0.029	0	0	0
Insecta	Ephemeroptera	Leptophlebiidae	Leptophlebiidae	0	0	0	0	0	0.037
Insecta	Ephemeroptera	Siphonuridae	<i>Siphonurus</i>	0	0	0	0	0	0.12
Insecta	Hemiptera	Corixidae	Corixidae	0	0	0.005	0	0.046	0
Insecta	Hemiptera	Corixidae	<i>Micronecta</i>	0.039	0	0	0	0	0.007
Insecta	Hemiptera	Corixidae	<i>Sigara</i>	0	0	0	0.006	0	0
Insecta	Hemiptera	Nepidae	<i>Nepa cinerea</i>	0	0	0	0.011	0.013	0
Insecta	Hemiptera	Veliidae	Veliidae	0	0.007	0	0	0	0

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

Insecta	Megaloptera	Sialidae	<i>Sialis</i>	0	0	0.01	0	0	0
Insecta	Megaloptera	Sialidae	<i>Sialis lutaria</i>	0	0	0	0.011	0	0
Insecta	Neuroptera	Sisyridae	<i>Sisyra</i>	0	0	0	0.006	0	0
Insecta	Odonata	Calopterygidae	<i>Calopteryx</i>	0	0	0	0.022	0	0
Insecta	Odonata	Coenagrionidae	<i>Coenagrionidae</i>	0	0	0	0.006	0	0
Insecta	Odonata	Cordulegastridae	<i>Cordulegaster</i>	0	0	0	0.011	0.046	0.037
Insecta	Odonata	Cordulegastridae	<i>Cordulegaster boltonii</i>	0.008	0.014	0.048	0	0	0
Insecta	Odonata	Corduliidae	Corduliidae	0.008	0	0	0	0	0
Insecta	Odonata	Gomphidae	Gomphidae	0.008	0	0	0	0	0
Insecta	Odonata	Lestidae	Lestidae	0	0.007	0	0	0	0
Insecta	Plecoptera	Capniidae	<i>Zwicknia</i>	0.047	0	0	0	0	0.382
Insecta	Plecoptera	Nemouridae	<i>Amphinemura</i>	2.004	2.225	0	0	0	0
Insecta	Plecoptera	Nemouridae	<i>Amphinemura standfussi</i>	0	0	0.005	0	0	0
Insecta	Plecoptera	Nemouridae	<i>Nemoura</i>	1.483	0.833	0.014	0	0.013	3.305
Insecta	Plecoptera	Perlodidae	<i>Isoperla</i>	0.379	0.579	0	0	0	0.465
Insecta	Plecoptera	Taeniopterygidae	<i>Brachyptera</i>	0.513	0	0	0	0	1.821
Insecta	Plecoptera	Taeniopterygidae	<i>Brachyptera risi</i>	0	0.042	0	0	0	0
Insecta	Trichoptera	Beraeidae	<i>Beraea maurus</i>	0	0.021	0.01	0	0	0
Insecta	Trichoptera	Beraeidae	<i>Beraea pullata</i>	0	0.007	0	0	0	0.03
Insecta	Trichoptera	Beraeidae	<i>Beraeodes minutus</i>	0	0	0	0.011	0.007	0
Insecta	Trichoptera	Glossosomatidae	<i>Agapetus</i>	0	0.021	0.005	0	0.079	0
Insecta	Trichoptera	Glossosomatidae	<i>Glossosomatidae</i>	0	0	0	0.066	0	0
Insecta	Trichoptera	Goeridae	<i>Goera pilosa</i>	0	0	0	0.006	0.013	0
Insecta	Trichoptera	Goeridae	Goeridae	0	0	0.153	0	0	0
Insecta	Trichoptera	Goeridae	<i>Lithax obscurus</i>	0	0.007	0	0.017	0.105	0.03
Insecta	Trichoptera	Goeridae	<i>Silo nigricornis</i>	0.008	0.014	0	0.022	0.013	0
Insecta	Trichoptera	Goeridae	<i>Silo pallipes</i>	0	0	0	0	0.007	0
Insecta	Trichoptera	Hydropsychidae	<i>Diplectrona felix</i>	0.016	0.007	0	0	0.013	0.015
Insecta	Trichoptera	Hydropsychidae	<i>Hydropsyche</i>	0.158	0.092	7.779	6.637	4.809	0
Insecta	Trichoptera	Hydropsychidae	<i>Hydropsyche angustipennis</i>	0	0	0	0	0	0.749
Insecta	Trichoptera	Hydropsychidae	<i>Hydropsyche siltalai</i>	0.024	0.021	0.019	0.006	0.04	0.015
Insecta	Trichoptera	Leptoceridae	<i>Athripsodes</i>	0.032	0.155	0.01	0.271	0.501	0.195
Insecta	Trichoptera	Leptoceridae	<i>Ceraclea albimacula</i>	0	0.007	0	0	0	0
Insecta	Trichoptera	Leptoceridae	<i>Mystacides</i>	0	0	0	0.011	0	0
Insecta	Trichoptera	Limnephilidae	<i>Anabolia</i>	0.008	0	0	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Chaetopteryx</i>	0	0.035	0.033	0.083	0	0
Insecta	Trichoptera	Limnephilidae	<i>Drusus</i>	0	0	0.014	0	0	0
Insecta	Trichoptera	Limnephilidae	<i>Glyphotaelius pellucidus</i>	0	0	0	0	0	0.037
Insecta	Trichoptera	Limnephilidae	<i>Halesus radiatus</i>	0.032	0	0	0	0	0
Insecta	Trichoptera	Limnephilidae	Limnephilinae	0.016	0	0.024	0	0	0.39

Chapitre 2 - Métacommunautés dans les têtes de bassin versant, le rôle de la dispersion et de la fragmentation

Insecta	Trichoptera	Limnephilidae	<i>Limnephilus</i>	0.15	0.579	0.019	0	0.013	0.15
Insecta	Trichoptera	Limnephilidae	<i>Micropterna lateralis</i>	0	0.014	0	0	0	0.03
Insecta	Trichoptera	Polycentropodidae	<i>Plectrocnemia</i>	0.008	0.007	0	0	0.007	0
Insecta	Trichoptera	Polycentropodidae	<i>Polycentropus</i>	0	0	0	0	0.033	0
Insecta	Trichoptera	Psychomyiidae	<i>Lype</i>	0	0	0	0	0	0.007
Insecta	Trichoptera	Psychomyiidae	<i>Tinodes</i>	0	0	0.01	0	0	0.007
Insecta	Trichoptera	Psychomyiidae	<i>Tinodes waeneri</i>	0	0	0.005	0	0	0.022
Insecta	Trichoptera	Rhyacophilidae	<i>Rhyacophila</i>	0	0.007	0.019	0	0.013	0.007
Insecta	Trichoptera	Sericostomatidae	<i>Notidobia ciliaris</i>	0	0	0	0.006	0	0.015
Insecta	Trichoptera	Sericostomatidae	<i>Sericostomatidae</i>	0	0	0.005	0	0	0
Malacostraca	Amphipoda	Eulimnogammaridae	<i>Echinogammarus</i>	0.915	1.109	2.058	2.485	5.356	0.937
Malacostraca	Amphipoda	Gammaridae	Gammaridae	0.032	0.184	0	0	0	0
Malacostraca	Amphipoda	Gammaridae	<i>Gammarus</i>	4.346	5.198	10.047	11.541	11.569	8.311
Malacostraca	Amphipoda	Niphargidae	<i>Niphargus</i>	0	0.049	0	0	0	0.007
Malacostraca	Calanoïda	Calanoïda	<i>Calanoïda</i>	0	0.021	0	0	0.007	0.127
Malacostraca	Cladocera	Cladocera	Cladocera	15.934	0.537	0.21	9.961	0.58	2.158
Malacostraca	Cyclopoïda	Cyclopoïda	Cyclopoïda	0.237	0.247	0.033	0.409	1.074	1.162
Malacostraca	Decapoda	Astacidae	Astacidae	0	0	0	0	0.007	0
Malacostraca	Decapoda	Cambaridae	Cambaridae	0	0	0	0	0	0.007
Malacostraca	Decapoda	Cambaridae	<i>Orconectes</i>	0	0	0.005	0	0	0
Malacostraca	Harpacticoïda	Harpacticoïda	Harpacticoïda	0	0.021	0	0	0	0.06
Malacostraca	Isopoda	Asellidae	Asellidae	0.142	0.021	0	0	0	0
Malacostraca	Isopoda	Asellidae	<i>Asellus aquaticus</i>	0.458	2.684	1.165	1.778	1.344	0.465
Malacostraca	Isopoda	Asellidae	<i>Proasellus</i>	0.174	1.123	4.833	0	0	0
Malacostraca	Isopoda	Asellidae	<i>Proasellus meridianus</i>	0	0	0	0.574	1.067	1.296
Malacostraca	Ostracoda	Ostracoda	Ostracoda	0.426	0.699	0.038	0.121	0.257	0.435
Mollusca	Gastropoda	Acroloxidae	<i>Acroloxus lacustris</i>	0	0	0	0.011	0.02	0.007
Mollusca	Gastropoda	Amnicolidae	<i>Potamopyrgus antipodarum</i>	2.09	3.955	0.019	0.21	1.331	1.409
Mollusca	Gastropoda	Lymnaeidae	<i>Galba truncatula</i>	0	0	0.005	0	0.007	0.022
Mollusca	Gastropoda	Lymnaeidae	<i>Radix</i>	0.016	0.014	0.014	0.011	0.026	0.007
Mollusca	Gastropoda	Physidae	<i>Physa acuta</i>	0	0.007	0.014	0.022	0.053	0.075
Mollusca	Gastropoda	Planorbidae	<i>Ancylus fluviatilis</i>	0.323	0.134	4.446	0.442	0.422	0.742
Mollusca	Gastropoda	Planorbidae	<i>Anisus spirorbis</i>	0	0.007	0.029	0.006	0	0.022
Mollusca	Gastropoda	Planorbidae	<i>Gyraulus albus</i>	0	0	0	0.028	0.033	0.097
Mollusca	Gastropoda	Planorbidae	<i>Hippeutis complanatus</i>	0	0.007	0.005	0.006	0.033	0.165
Mollusca	Gastropoda	Planorbidae	<i>Menetus dilatatus</i>	0	0	0.005	0	0.04	0
Mollusca	Gastropoda	Planorbidae	Planorbidae	0	0	0.005	0	0	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium</i>	0	0	0	0.453	0	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium casertanum</i>	0.079	0.212	0.287	0	0.026	0.382

3. Variabilité temporelle des processus et influences des fragmentations permanente et temporaire dans une métacommunauté de tête de bassin versant (article 2)

Mollusca	Venerida	Sphaeriidae	<i>Pisidium henslowanum</i>	0	0	0.019	0.072	0.013	0.03
Mollusca	Venerida	Sphaeriidae	<i>Pisidium hibernicum</i>	0	0	0	0.006	0	0
Mollusca	Venerida	Sphaeriidae	<i>Pisidium milium</i>	0.024	0.268	0.024	0.072	0.283	0.562
Mollusca	Venerida	Sphaeriidae	<i>Pisidium nitidum</i>	1.586	4.188	1.084	0.878	3.795	0.45
Mollusca	Venerida	Sphaeriidae	<i>Pisidium personatum</i>	0.284	0.558	1.156	0.155	1.515	3.837
Mollusca	Venerida	Sphaeriidae	<i>Pisidium subtruncatum</i>	1.743	2.444	2.025	0.9	3.551	4.744
Mollusca	Venerida	Sphaeriidae	<i>Pisidium supinum</i>	0	0.007	0	0	0	0
Mollusca	Venerida	Sphaeriidae	<i>Sphaerium corneum</i>	0.032	0.12	0.134	0.094	1.555	0.12
Nematoda	Nematoda	Nematoda	Nematoda	0.213	0.777	0.31	0.138	0.369	0.442
Nematomorpha	Gordea	Gordiidae	<i>Gordius</i>	0	0	0	0	0.007	0.007
Nemertea	Hoplonemertea	Tetrastemmatidae	<i>Prostoma</i>	0	0	0	0	0.007	0.007
Platyhelminthes	Tricladida	Dendrocoelidae	<i>Dendrocoelum lacteum</i>	0	0	0.005	0.011	0.026	0
Platyhelminthes	Tricladida	Dugesiiidae	<i>Dugesia</i>	0	0	0	0	0.013	0.015
Platyhelminthes	Tricladida	Dugesiiidae	<i>Dugesia tigrina</i>	0	0	0.005	0.326	0	0
Platyhelminthes	Tricladida	Planariidae	<i>Polycelis</i>	0.189	0	1.461	1.513	2.701	2.908
Platyhelminthes	Tricladida	Tricladida	Tricladida	0.008	0.706	0.014	0	0	0
Specific richness				18.07	20.33	19.93	26.87	25.10	20.97

Appendix 2. Non-parametric permutational multivariate analysis of variances (PERMANOVA) testing differences in community composition between intermittent and perennial sites at each sampling date. Chao index was used after $\log(x+1)$ transformation of abundance to calculate community dissimilarities. Significant relationships are shown in bold (p-value ≤ 0.05 *).

Sampling date	Hydrological phase	F	p-value
March	Wet	2.807	0.001
April	Drying	2.614	0.006
June	Drying	2.589	0.005
September	Dry	0.303	0.990
October	Dry	0.635	0.853
December	Rewetting	2.234	0.008

Chapitre 3

Apport de l'identification moléculaire pour l'étude des métacommunautés de têtes de bassin versant

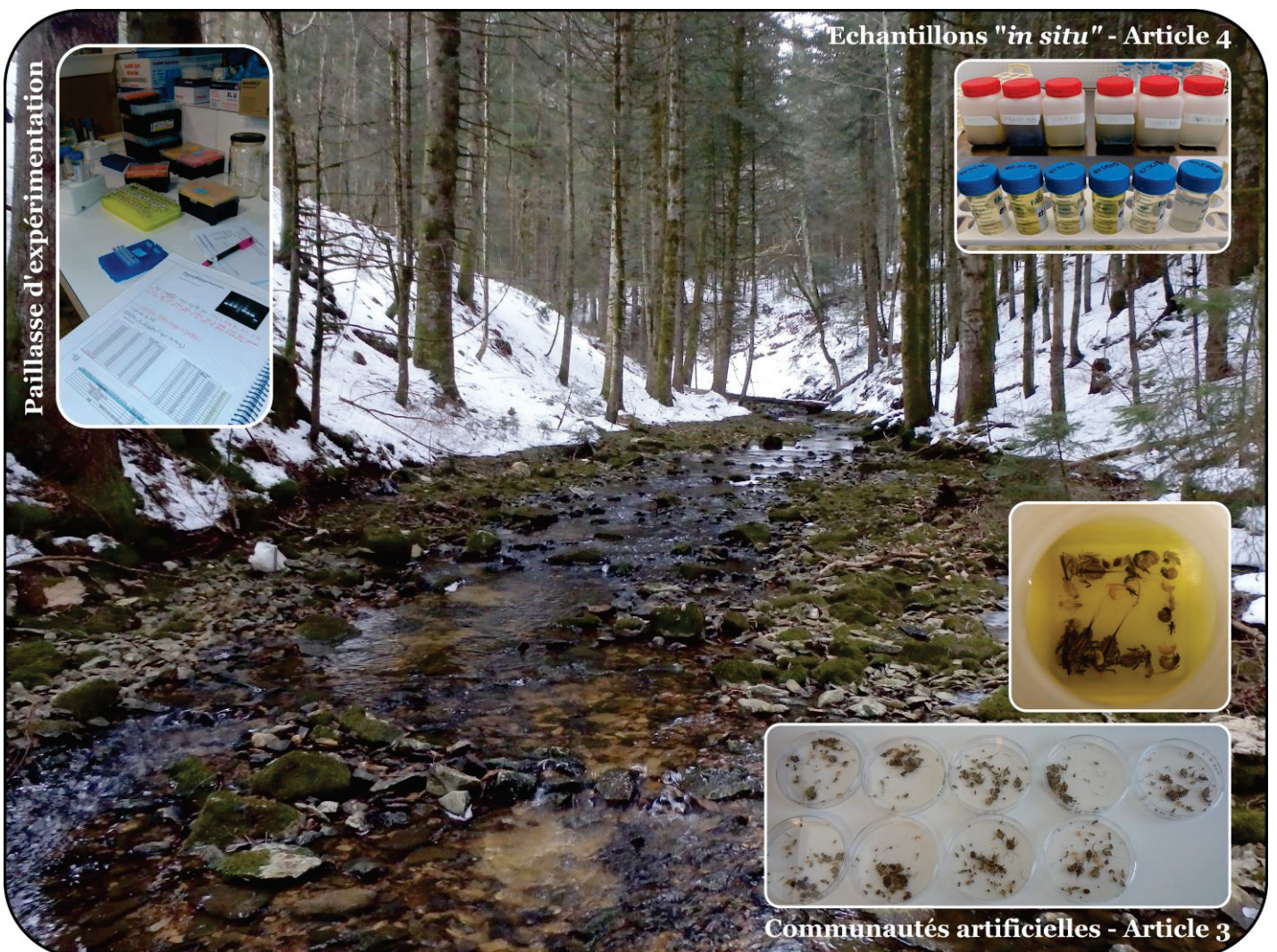


Illustration : Ensemble de photographies présentant un site de prélèvement sur l'Albarine, étudié dans le Chapitre 3, et quelques échantillons utilisés. Photos : L. Konecny-Dupré, M. Gauthier.

1 Matériel et méthodes - Chapitre 3

Ce chapitre porte sur l'évaluation de l'apport d'une identification moléculaire dans la compréhension des processus structurant les métacommunautés de macro-invertébrés benthiques dans les têtes de bassin versant. Pour ce faire, plusieurs outils d'identification moléculaire ont été développés et/ou utilisés et comparés entre eux en laboratoire (Article 3) pour ensuite être appliqués sur des échantillons collectés *in situ* (Article 4). Ce matériel et méthodes détaille les différentes étapes de développement des sondes ARN pour la méthode d'enrichissement par capture (Article 3) ainsi que les différents protocoles utilisés dans les articles 3 et 4 de l'extraction au traitement bioinformatique pour les différentes matrices ADN (ADN bulk et ADNet) et méthodes d'enrichissement (PCR et capture par sondes ARN).

1.1 Travail préliminaire à la conception de l'outil moléculaire : conception des sondes ARN

Le développement de l'outil moléculaire utilisant la capture par sondes ARN a demandé un travail préliminaire concernant la conception des sondes. En effet, la capture par sondes ARN est un outil utilisé en génomique et génétique pour séquencer plusieurs loci (Hodges et al., 2007 ; McCormack et al., 2013) ou des SNPs (polymorphisme nucléotidique, Yang et al., 2009) d'une espèce mais n'est pas utilisé en écologie des communautés où l'objectif est de séquencer un locus chez plusieurs espèces dans un but d'identification. A l'inverse des amorces utilisées en PCR en routine pour amplifier un locus chez plusieurs espèces qui ont été développées pour de nombreux groupes, gènes (Creer et al., 2016) et matrices ADN (p.ex. l'ADN dégradé, Vamos et al., 2017), il n'existe pas de set de sondes disponible pour le metabarcoding.

Dans le cadre de cette thèse, il a donc été nécessaire de développer un set de sondes pour les macro-invertébrés présents en France métropolitaine dans les têtes de bassin versant. Pour cela, il a fallu élaborer une base de données de séquences de la région Folmer du COI utilisée comme barcode chez les métazoaires (658 paires de bases, Hebert et al., 2003) pour les espèces d'intérêt qui a ensuite servi pour la conception des sondes ARN.

1.1.1 Élaboration d'une base de données COI

Pour élaborer la base de données contenant les COI des espèces d'intérêt, une base de données interne à l'équipe EcoFlowS de INRAE actualisée régulièrement et contenant les taxa de macro-invertébrés connus en France métropolitaine en se basant sur les connaissances actuelles a été utilisée. En novembre 2016, la base de données contenait 3 245 espèces comprises dans 22 ordres de macro-invertébrés (Figure 1.1). De cette base de données ont été exclues les familles de diptères n'appartenant pas à la famille des Chironomidae car la liste exhaustive des espèces de France pour ces

familles est inconnue dû notamment au manque de connaissances à leur égard et il n'était pas envisageable de travailler avec l'ensemble des espèces mondiales de cet ordre extrêmement diversifié (Wiegmann & Yeates, 2007). Un tri des familles a été effectué à partir de la base de données en enlevant celles marginales dans les cours d'eau de têtes de bassin versant sur dires d'experts (p.ex., Corixidae), retirant 122 espèces et 6 ordres. Sur les ordres restants, un tri des espèces sur ces critères a aussi été fait et les séquences du COI des 2 448 espèces restantes ont été téléchargées à partir des bases de données publiques GenBank et BOLD en décembre 2016 via un package du logiciel R permettant de télécharger les données d'un gène spécifique (ici COI) à partir d'une liste d'espèces ou de taxa (PrimerMiner v.0.18, Elbrecht & Leese, 2017a). Pour chaque ordre, il a été regardé si suffisamment de données génétiques (50% des espèces avec une séquence de COI) étaient disponibles dans les bases de données et les 4 ordres pour lesquels ce n'était pas le cas ont été retirés (Acariformes, Enchytraeida, Haplotaxida, Lumbriculida). Pour les 12 ordres restants, les espèces ont été classées en deux catégories après avoir retiré les espèces absentes des zones d'études des articles 1 et 2 : celles possédant une séquence COI dans les bases de données et celles n'en possédant pas. Pour cette dernière catégorie représentant 261 espèces, la région Folmer du COI (Folmer et al., 1994) a été séquencée après extraction au Chelex (Bio-Rad) et séquençage Sanger (détails dans la section 1.2.2) de l'espèce lorsqu'un organisme était disponible dans les collections de l'équipe EcoFlowS. Quarante-deux espèces ont ainsi pu être séquencées (numéros d'accès sur GenBank : MK584300:MK584515). Au final, une base de données de séquences de COI appartenant à la région Folmer contenant des barcodes pour 1 525 espèces sur les 1 689 d'intérêt (plus de 90%) et représentant 12 ordres de macro-invertébrés benthiques des têtes de bassins versants a été élaborée (Tableau 1.1).

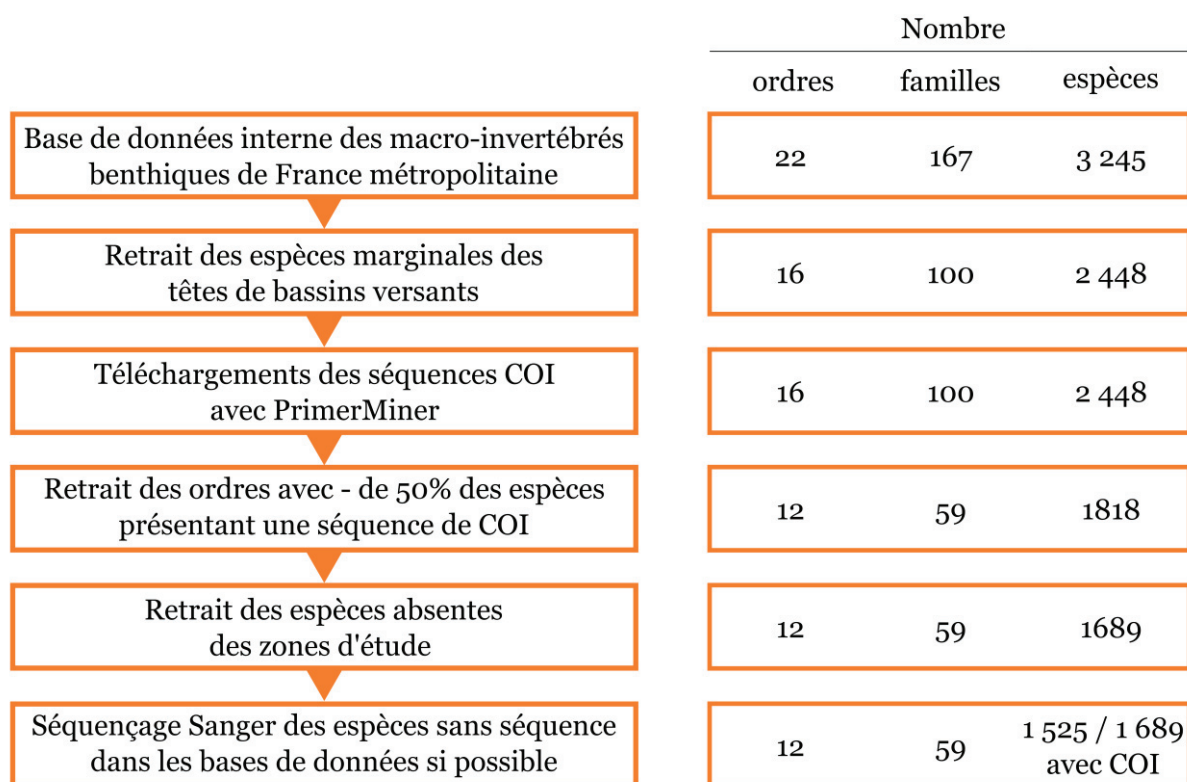


Figure 1.1 : Différentes étapes effectuées pour établir la base de données COI pour la conception des sondes ARN

Tableau 1.1. Liste des ordres utilisés pour la conception des sondes ARN avec le nombre d'espèces d'intérêt total, avec des barcodes disponibles dans les bases de données et le nombre de sondes générées pour chaque ordre.

Ordres	Nombre d'espèces d'intérêt	Nombre d'espèces d'intérêt avec barcodes	Nombre de sondes générées
Amphipoda	19	14	374
Arhynchobdellida	8	6	94
Coleoptera	416	346	2176
Diptera	670	670	6785
Ephemeroptera	92	74	1121
Gastropoda	35	31	501
Isopoda	5	4	312
Odonata	30	23	192
Plecoptera	118	99	786
Rhynchobdellida	15	11	156
Trichoptera	266	232	2442
Venerida	15	15	99

1.1.2 Conception des sondes

La conception des sondes a été faite grâce au package BaitFisher, un outil optimisant la conception de sondes pour enrichir plusieurs loci d'un génome (Mayer et al., 2016). Ce package a été conçu pour concevoir des sondes d'hybridation à partir d'alignements de séquences multiples en évitant la redondance dans la construction des sondes, c'est-à-dire en concevant moins de sondes dans les régions très conservées d'un génome et plus de sondes dans les régions variables. Dans le cadre de cette thèse, nous avons appliqué cet outil pour enrichir un même locus chez plusieurs espèces différentes.

Pour chaque ordre de macro-invertébrés, un alignement protéique à une séquence de COI de référence (*Drosophila yakubai* ; No d'accès NCBI : NC_001322.1) a été fait. Les séquences strictement identiques (les Ns et les gaps aux extrémités des séquences ne sont pas considérés comme des différences) ont ensuite été fusionnées afin d'éviter la redondance au sein du groupement de séquences avec un script perl (https://github.com/TristanLefebure/collapse_to_uniq_seq). L'alignement COI obtenu a ensuite été traité par le package BaitFisher. Dans un premier temps, 10 fenêtres de lecture de 120 bases ont été définies dans l'alignement avec un chevauchement entre fenêtres de 60 bases (Figure 1.2). A cette étape, plusieurs positions de départ sont testées (de la longueur de la sonde, ici 120 premières bases) afin de trouver la meilleure pour découper l'alignement. Cette position est définie comme celle comportant les fenêtres pour lesquelles au moins une séquence ne comporte pas de gap ou d'ambiguïtés nucléotidiques. L'alignement est ensuite découpé en plusieurs fenêtres selon la meilleure position de départ et, pour chaque fenêtre, les séquences sont regroupées selon un seuil de dissimilarité maximal, fixé ici à 5%. Ainsi, pour chaque fenêtre, un nombre différent de groupes peut être obtenu. Pour chaque groupe de séquences, une séquence nucléotidique artificielle a été inférée. Cette séquence présente la plus petite distance maximale à toutes les séquences présentes dans le groupe et constitue la sonde (Figure 1.2). Au final, 15 038 sondes ont été générées à partir de 100 367 séquences (Tableau 1.1). Les sondes ont ensuite été envoyées à Arbor Biosciences (USA) qui ont synthétisé les sondes d'hybridation ARN. Le set de sondes est disponible sur Zenodo ([Sondes Zenodo](#)).

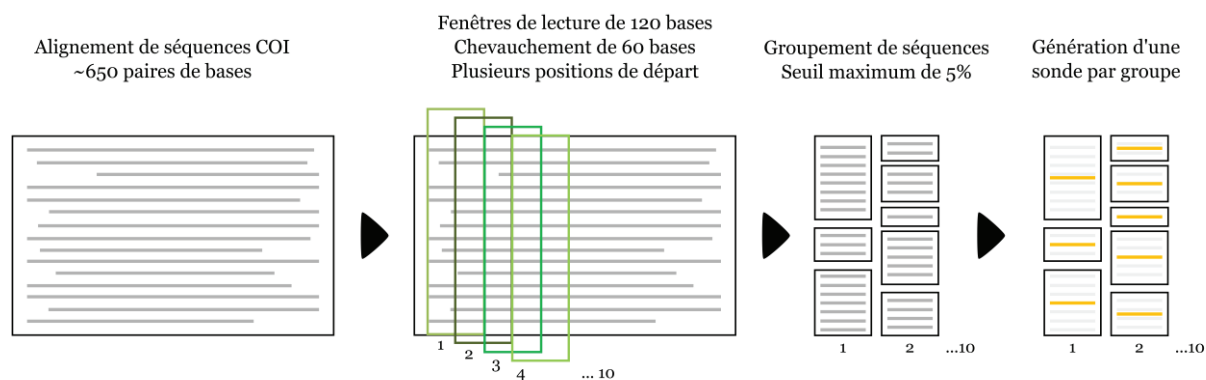


Figure 1.2 : Étapes suivies pour la conception des sondes par le logiciel BaitFisher (Mayer et al, 2016)

1.2 Protocoles de laboratoire : de l'extraction au séquençage

Le metabarcoding repose sur plusieurs grandes étapes qui sont présentées dans les sections suivantes. En résumé, après le prélèvement des échantillons qui sont stockés à froid en attendant d'être traités, l'ADN total est extrait à partir de chaque échantillon. Ensuite, le barcode est enrichi dans l'échantillon, classiquement par amplification par PCR à l'aide d'amorces spécifiques, mais d'autres méthodes d'enrichissement peuvent être utilisées (cf Introduction, section 7). Après enrichissement, les séquences des barcodes enrichies sont séquencées, c'est-à-dire que leur enchaînement de nucléotides est déterminé puis les données brutes de séquençage sont nettoyées bio-informatiquement et chaque séquence conservée sera confrontée à une base de données de référence pour se voir assigner un nom d'espèce ou de taxon (étapes détaillées dans la section 1.4).

1.2.1 Extraction de l'ADN total

Pour cette thèse, deux types de matrice ADN ont été utilisées : l'homogénat des organismes (ADN bulk) et l'éthanol de conservation (ADNet), impliquant deux protocoles d'extraction différents. L'ADN bulk a été utilisé seulement dans l'article 3 sur des communautés artificielles (des communautés composées en laboratoire qui ne contenaient que des organismes et de l'éthanol, sans matière non désirée comme des feuilles, du sable, etc...) et l'ADNet a été utilisé dans les articles 3 et 4 sur des communautés artificielles et des communautés échantillonnées *in situ* (des communautés échantillonnées dans des rivières donc contenant de la matière non désirée).

(a) *ADN bulk* : les organismes ont été sortis de l'échantillon et placés dans des boîtes de pétri stériles dans lesquels ils ont été laissés pendant une nuit pour que l'éthanol s'évapore. Les organismes secs étaient ensuite broyés dans un broyeur à billes MM200 (Retsch) pendant 4 minutes à 30 Hz. Le broyat a ensuite été transféré dans un tube de 2 mL. Les capsules de broyage étaient systématiquement nettoyées entre chaque broyage avec un trempage pendant 10 secondes dans de la javel suivie d'un rinçage à l'éthanol 96%. Le broyat était ensuite extrait à l'aide du kit d'extraction FastDNA® Spin Kit

for Soil (MP Biomedicals) en suivant les instructions du fabricant. L'ADN a été élué dans 60 μ L de tampon d'éluion du kit chauffé à 70°C pour améliorer le rendement d'extraction.

(b) *ADNet* : pour chaque échantillon, 48 mL d'éthanol de conservation a été prélevé à l'aide de pipettes stériles après homogénéisation de l'échantillon et placé dans un tube Falcon® 50 mL. Ensuite, 1500 μ L d'acétate de sodium 3M et 10 μ L de glycoène étaient ajoutés afin de faciliter la précipitation de l'ADN. L'échantillon était ensuite homogénéisé et placé pendant au moins 72 heures à -80°C. Après 72 h, une centrifugation était faite à 14 000g et 6°C pendant 30 minutes. Le surnageant (fraction liquide en surface) était ensuite retiré, suivi d'une nouvelle centrifugation pendant 15 mn (14 000 g, 6°C). La majorité du surnageant restant était retiré et les échantillons étaient placés sous une hotte à flux horizontaux jusqu'à évaporation totale de l'éthanol. Le résidu sec était alors extrait et dissous dans la solution tampon du kit d'extraction (volume dépendant du kit utilisé) puis transféré dans un tube de 1,5 mL. Deux kits d'extraction différents ont été utilisés entre les deux articles car, dans l'article 3, l'éthanol des communautés artificielles était relativement pur et les risques de présence d'inhibiteurs de PCR ou d'autres méthodes étaient faibles. Pour ces échantillons, le kit NucleoSpin Tissue® kit (Macherey- Nagel GmbH) a été utilisé, en suivant les instructions du fabricant, avec une durée de lyse de 1h et l'emploi de RNase et une éluion dans 46 μ L de tampon d'éluion chauffé à 70°C. Dans l'article 4, les communautés étaient prélevées avec le substrat des rivières et les échantillons contenaient des feuilles, du sable et d'autres composés, notamment des acides humiques connus comme des inhibiteurs de PCR. Ainsi, il a été choisi d'extraire l'éthanol de ces échantillons avec le kit FastDNA® Spin Kit for Soil (MP Biomedicals) qui élimine les potentiels inhibiteurs lors de l'extraction avec une éluion dans 60 μ L de tampon d'éluion après une incubation de 5 minutes à 55°C pour optimiser le rendement d'extraction.

1.2.2 Base de données COI de référence

Dans l'article 3, une base de données de COI pour les 10 espèces utilisées (cf. article 3 pour les détails des espèces) a été construite pour l'étape d'identification des lectures issues du séquençage (cf. section Traitement bio-informatique des données issues du séquençage 1.4 pour les détails). L'ADN d'un organisme de chaque espèce a été extrait selon le protocole de l'ADN bulk de la section 1.2.1. La région Folmer du COI (Folmer et al., 1994) a ensuite été amplifiée à partir des ADN extraits. Les amplifications ont été effectuées dans une réaction de 25 μ L contenant une solution tampon 1X contenant du 3 MgCl₂ 3 mM (Eurobio, France), 0,05 U/ μ l de polymérase EurobioTaq DNA polymerase (Eurobio), 0,8 mM de chaque dNTP (Eurogentec), 0,1 mg/mL de BSA (New England BioLabs) et 0,4 μ M de chaque amorce (LCO1490/HCO2198, Tableau 1.2) et 0,5 μ L d'ADN. L'amplification consistait en une dénaturation initiale de 3 min à 95°C suivi par 40 cycles composés d'une étape de dénaturation à 95°C de 20 sec, d'une d'hybridation à 51°C de 30 sec et d'une d'élongation à 72°C de 45 sec suivi d'une extension finale à 72°C de 5 minutes. Les amplicons

produits ont ensuite été séquencés après purification par séquençage Sanger à partir de l'amorce forward et de l'amorce reverse pour obtenir un fragment le plus long possible et de bonne qualité par la société Biofidal. Les séquences obtenues ont été vérifiées et nettoyées manuellement avec le logiciel Finch TV 1.4.0. Le séquençage n'ayant pas permis d'obtenir une séquence de bonne qualité pour une espèce, elle a été téléchargée depuis NCBI pour compléter la base de données COI (numéro d'accès sur GenBank : HE651395.1).

1.3 Enrichissement : amplification par PCR

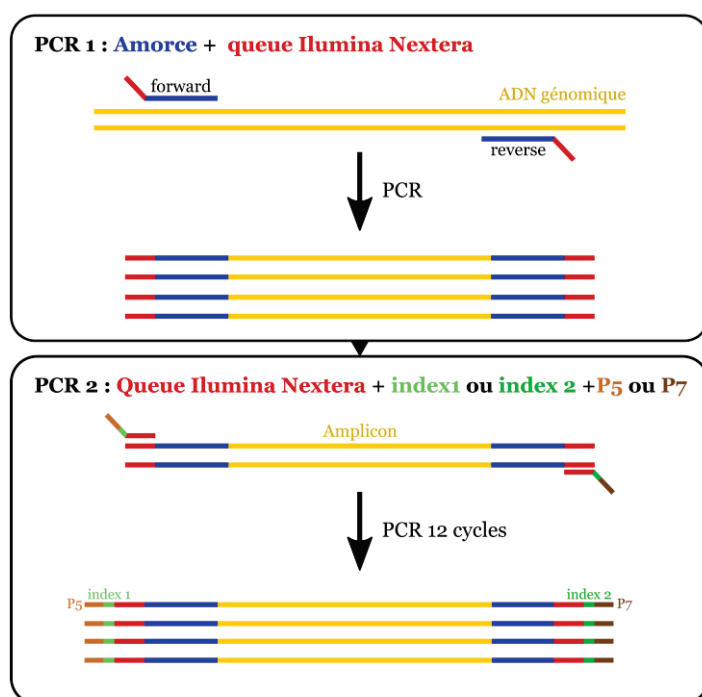


Figure 1.3 : Schéma illustrant les deux étapes de l'enrichissement par PCR pour obtenir des bibliothèques pour le séquençage. La PCR 1 permet d'amplifier la région d'intérêt (le barcode) et la PCR 2 permet d'ajouter les tags et les extrémités P5/P7 qui sont nécessaires pour le séquençage.

Le séquençage haut-débit nécessite une préparation des échantillons en amont spécifique que l'on appelle la préparation des bibliothèques. Une bibliothèque de séquences Illumina correspond à un ensemble de fragments de taille homogène et qui portent toutes à leurs extrémités un ensemble de séquences contenant entre autres (i) une étiquette, ou tag, spécifique (des codes uniques, ici composés de 8 bases) permettant de connaître l'origine de chaque fragment lors du séquençage pour le démultiplexage et (ii) des adaptateurs permettant de s'hybrider sur les flowcell (cellules de séquençage).

Pour l'enrichissement par PCR, une double PCR a été effectuée avec une première PCR (PCR 1) permettant d'amplifier la région d'intérêt avec des amorces lui étant spécifiques et une deuxième PCR (PCR 2) permettant de préparer les échantillons pour le multiplexage et le séquençage (Figure 1.3).

La PCR1 a été effectuée avec deux paires d'amorces adaptées pour les invertébrés d'eau douce ciblant des loci au sein de la région Folmer du COI : BFR2 amplifiant un amplicon de 421 pb (Elbrecht & Leese, 2017b) et Fwh1 amplifiant un amplicon de 178 pb (Vamos et al., 2017). Pour permettre la double PCR, des queues Illumina Nextera® ont été ajoutées aux séquences des deux couples d'amorces pour préparer la PCR 2 (Figure 1.3, Tableau 1.2). La PCR 2 permet le double-indexage de chaque échantillon et leur préparation pour le séquençage avec l'ajout d'oligonucléotides (P5 ou P7) spécifiques à la technologie utilisée pour l'indexage et le séquençage, ici Illumina

Nextera® : c'est à cette étape que l'on obtient des librairies. Le double-indexage permet d'avoir des combinaisons de tags unique pour chaque échantillon et ainsi diminuer les fausses attributions lors du démultiplexage (Kircher et al., 2012).

La PCR 1 a été faite en triplicatas dans un volume total de 25 µL avec une solution tampon 1X contenant du 3 MgCl₂ 3mM, 400 µM de chaque dNTP et la polymérase (QIAGEN® Multiplex PCR kit, Qiagen,Allemagne), 0.5 µM de chaque amorce et 10 ng d'ADN pour l'ADN bulk ou 5 µL d'ADN pour l'ADNet. L'amplification consistait en une dénaturation initiale de 5 min à 95°C suivi par 25 (ADN bulk) ou 35 (ADNet) cycles composés d'une étape de dénaturation à 95°C de 30 sec, d'une d'hybridation à 50°C (ADN bulk) ou 52°C (ADNet) de 30 sec et d'une d'élongation à 72°C de 2 min suivie d'une extension finale à 72°C de 10 minutes. Les produits PCR ont été regroupés par triplicatas et purifiés avec des billes de purification Agencourt AMPure XP (Beckman Coulter) pour éliminer les amorces restantes et les fragments non cibles. Les produits PCR ont été quantifiés avec le kit QuantiFluor® dsDNA System (Promega). Dix nanogrammes de chaque produit PCR ont été utilisés pour une seconde PCR.

La PCR2 a été faite dans un volume total de 25 µL avec une solution tampon 1X (BIOAmp® Blend Mix ; Biofidal), 2.5 nM de MgCl₂, 200 µM de chaque dNTP, 0.25 µM de d'amorce Illumina (Tableau 1.2) et 0.02 U/µL de HOT BioAmp Taq (Biofidal). L'amplification consistait en une dénaturation initiale de 5 min à 96°C suivi par 12 cycles composés d'une étape de dénaturation à 95°C de 20 sec, d'une d'hybridation à 54°C de 30 sec et d'une d'élongation à 72°C de 1 min suivie d'une extension finale à 72°C de 10 minutes. Les librairies obtenues ont ensuite été purifiées avec des billes de purification Agencourt AMPure XP (Beckman Coulter), quantifiées avec le kit QuantiFluor® dsDNA System (Promega) puis regroupées à la même concentration (2 nM) pour le séquençage (étape de multiplexage). Les librairies d'amplicons de PCR ont été séquencées en utilisant le kit de séquençage 2*250 paired-end V3 MiSeq par l'entreprise Biofidal.

1.3.1 Enrichissement : capture par sondes ARN

La préparation des bibliothèques pour la capture par sondes ARN est plus complexe que celle faite lors de l'amplification par PCR. Tout d'abord, l'ADN génomique est fragmenté à la taille souhaitée pour le séquençage (ici par sonication par ultrasons en ciblant des fragments d'environ 600 pb). Les bibliothèques sont ensuite préparées (Figure 1.3). Les premières étapes assurent que les fragments sont prêts pour la ligation : une étape de « end repair » qui assure que les extrémités des fragments sont franches (c'est-à-dire pas de dépassement de bases sur un des deux brins de l'ADN), une étape avec un ajout d'un phosphate à l'extrémité 5' (5' phosphorylation) et d'un A à l'extrémité 3' (dA-tailing). Les fragments ainsi obtenus possèdent une extrémité 5' phosphorylée et une extrémité 3' avec un A supplémentaire. L'étape de ligation peut alors avoir lieu et consiste en l'ajout d'adaptateurs, qui peuvent contenir les tags ou non. Pour cette thèse, les adaptateurs contenaient uniquement les queues Illumina Nextera®. L'adaptateur est lié aux fragments par une enzyme ligase. Enfin, une amplification pour ajouter les

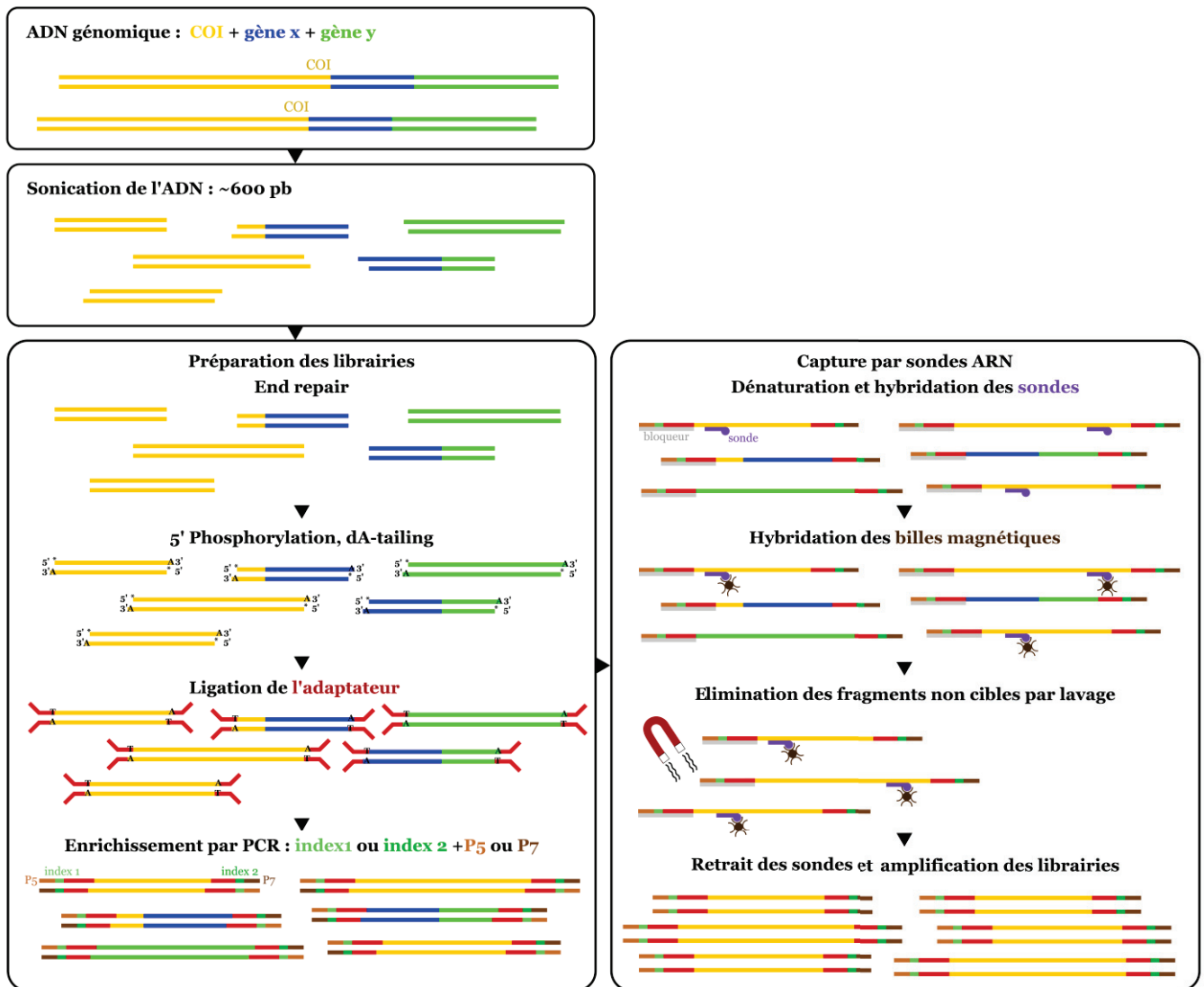


Figure 1.4 : Schéma illustrant les différentes étapes de l'enrichissement par capture par sondes ARN pour obtenir des bibliothèques pour le séquençage.

tags et les oligonucléotides P5 et P7 est effectuée et les librairies sont prêtes. L'enrichissement par capture par sondes ARN est faite à partir de ces librairies. Pour préparer l'hybridation des sondes, un des adaptateurs Illumina Nextera® entier est bloqué par des oligonucléotides spécifiques après dénaturation des fragments pour éviter des hybridations aléatoires sur les régions des adaptateurs. Les sondes sont ensuite ajoutées et s'hybrident avec leur séquence complémentaire sur les fragments d'intérêt. Ensuite, les complexes fragment/sonde sont liés à des billes magnétiques et les séquences non magnétisées vont être éliminées par des séries de lavage. Enfin, les librairies restantes sont séparées des sondes par dénaturation (95°C) qui sont retirées du milieu par magnétisme. Les librairies restantes dans le milieu, dont celles qui correspondent aux fragments d'intérêt (ici, fragments de COI de la région Folmer), sont amplifiées en préparation du séquençage.

La quantité d'ADN génomique de départ idéale est de 1 µg, quantité qui a pu être obtenue pour le bulk DNA mais pas pour l'ADNet (entre 14 et 57 ng). La sonication mécanique par ultrasons visant un découpage en fragments d'environ 600 pb a été faite avec un Qsonica Q800R (Qsonica). La préparation des librairies a été faite avec le kit de préparation de librairies NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England BioLabs) en suivant les instructions du fabricant avec une sélection de la taille des librairies entre 500 et 700 nucléotides en utilisant les billes de purification AMPure XP purification beads (Beckman Coulter). L'enrichissement par capture a été effectué avec le kit myBaits® Custom (Arbor Biosciences) en suivant les instructions du fabricant. Les librairies ont été regroupées par 2 ou 3 (par matrice ADN et longueur de librairies similaires) avant la capture pour diminuer le nombre de réactions de capture. Au départ, 100 ng et 230 ng de librairies ont été utilisées pour l'ADN bulk et l'ADNet en accord avec les instructions du fabricant concernant l'ADN dégradé. Comme dit dans la section 1.1, la capture par sondes n'est en général pas utilisée pour enrichir le même gène chez plusieurs espèces mais plusieurs gènes chez une espèce. Cette particularité qui rend notre set de sondes très concentré pour un seul gène nous a permis de diluer 10 fois les sondes sans avoir à priori d'impact sur les étapes ultérieures. L'hybridation a duré 24h pour les deux matrices ADN mais à une température de 55°C au lieu de 65°C pour l'ADNet en suivant les instructions du fabricant concernant l'ADN dégradé. L'amplification finale a été faite dans un volume total de 50 µL avec la polymérase KAPA HiFi DNA Polymerase (Kapa Biosystems) et les amorces P5 and P7 d'Illumina (Tableau 1.2). L'amplification consistait en une dénaturation initiale de 2 min à 98°C suivi par 21 cycles composés d'une étape de dénaturation à 98°C de 20 sec, d'une d'hybridation à 60°C de 30 sec et d'une d'élongation à 72°C de 1 min suivie d'une extension finale à 72°C de 5 minutes. Les librairies obtenues ont ensuite été purifiées avec des billes de purification Agencourt AMPure XP (Beckman Coulter), quantifiées par qPCR avec le kit KAPA qPCR (KAPA Library Quant Kit ; Kapa Biosystems) et regroupées à la même concentration pour le séquençage. Les librairies de capture ont été séquencées en utilisant le kit de séquençage 2*300 paired-end V3 MiSeq par l'entreprise Biofidal.

Tableau 1.2. Oligonucléotides utilisés pour le barcoding, l'amplification par PCR et la capture par sondes. Les amorces modifiées correspondent aux queues Illumina ajoutées aux amorces de la littérature pour la PCR 1. Les tags utilisés pour le double indexage sont des combinaisons uniques de 8 bases et sont représentées ici par des *XXXXXXXX*.

Barcoding		
Amorce	Séquence oligonucléotidique (5' → 3')	Littérature
LCO1490	GGTCAACAAATCATAAAGATATTGG	Folmer et al, 1994
HCO2198	TAAACTTCAGGGTGACCAAAAAATCA	Folmer et al, 1994
Enrichissement par PCR (PCR 1)		
	Séquence oligonucléotidique (5' → 3')	Littérature
Amorce BF2 modifiée avec queue Illumina Nextera®	TCGTCGGCAGCGTCAGATGTGTATAAGAG ACAGGCHCCHGAYATRGCHTTYCC	Elbrecht et al, 2017
Amorce BR2 modifiée avec queue Illumina Nextera®	GTCTCGTGGGCTCGGAGATGTGTATAAGA GACAGTCDGGRTGNCCRAARAAYCA	Elbrecht et al, 2017
Amorce fwhF1 modifiée avec queue Illumina Nextera®	TCGTCGGCAGCGTCAGATGTGTATAAGAG ACAGYTCHACWAAYCAYAARGAYATYGG	Vamos et al, 2017
Amorce fwhR1 modifiée avec queue Illumina Nextera®	GTCTCGTGGGCTCGGAGATGTGTATAAGA GACAGARTCARTTWCCRAAHCHCC	Vamos et al, 2017
Enrichissement par capture		
	Séquence oligonucléotidique (5' → 3')	Littérature
Queue Illumina Nextera® P5	TCGTCGGCAGCGTCAGATGTGTATAAGAG ACAG*T	
Queue Illumina Nextera® P7	[Phos]CTGTCTCTTATACACATCTCCGAGCC CACGAGAC	
P5	AATGATACGGCGACCACCGA	
P7	CAAGCAGAAGACGGCATACGA	
Indexage (PCR (PCR 2) et capture)		
	Séquence oligonucléotidique (5' → 3')	Littérature
Adaptateur Illumina P5 + tag	AATGATACGGCGACCACCGAACACXXXXX XXX TCGTCGGCAGCGTC	
Adaptateur Illumina P7 + tag	CAAGCAGAAGACGGCATACGAGATXXXXX XXXGTCTCGTGGGCTCGG	

1.4 Traitement bio-informatique des données issues du séquençage

L'entreprise BioFidal fournit les lectures (les séquences issues du séquençage, aussi appelées « reads ») démultiplexées, c'est-à-dire séparées par échantillon selon le double-indexage, et avec les adaptateurs Illumina, les tags et les séquences de P5 et P7 déjà retirées des lectures. Le protocole de traitement des lectures, ou pipeline, a été différent pour les lectures issues des enrichissements par PCR et capture. Les deux pipelines seront donc présentés indépendamment et sont disponibles sur Github (<https://github.com/mailysgauthier/bioinf-cap-PCR>). Par souci de cohérence, le même logiciel a été utilisé pour les deux méthodes d'enrichissement : Vsearch 2.8.4 développé par Rognes et al. (2016). Les matrices ADN (ADN bulk et ADNet) et les types de communautés (cf. article 3) ont été traitées avec les mêmes pipelines mais indépendamment.

1.4.1 Pipeline des données de séquençage obtenues avec l'enrichissement par PCR

Les lectures forward et reverse appartenant au même fragment ont été fusionnées (étape de « merge ») avec un minimum de recouvrement de 10 nucléotides entre les deux lectures et une différence maximum sur la zone de recouvrement de 5 nucléotides. Les lectures pour lesquelles cette fusion n'était pas possible ont été retirées. Ensuite les séquences des amorces ont été retirées avec cutadapt 1.9.1 (Martin, 2014) et les lectures ont été filtrées selon leur qualité (étape de « quality filtering ») avec comme critères : une valeur ee (expected error, erreur attendue) maximale de 1, une longueur de séquence minimale de 200 nucléotides et aucun N (nucléotide inconnu) autorisé. Les lectures ont été dérépliquées (les lectures strictement identiques ne sont représentées que par une seule séquence avec conservation du nombre de lectures lui étant affilié) et celles représentées par une unique séquence (singletons) ont été retirées. Les chimères, des séquences artificielles produites lors de la PCR lorsque qu'une séquence en cours d'élongation se détache de sa matrice ADN pour aller sur une autre et termine son élongation sur celle-ci, ont ensuite été *de novo* détectées, c'est-à-dire en utilisant les lectures générées par le séquençage et non une base de données de références extérieures pour détecter les chimères. Après retrait des chimères, l'étape de « clustering » a été faite, dans laquelle les lectures sont regroupées selon un seuil de similarité à 97%. Chaque cluster correspond à un OTU (Unité Taxinomique Opérationnelle, Operational Taxonomic Unit en anglais). Chaque OTU se voit attribuer une séquence qui correspond à celle de la première lecture utilisée pour composer le cluster et qui sera utilisée pour l'assignement. Dans l'article 4, étant donné que l'on ne travaillait pas avec des communautés dont la composition était connue, une étape de filtre des abondances des OTUs a été ajoutée pour retirer ceux pouvant être générés par des erreurs de PCR et de séquençage. L'abondance de chaque OTU pour un échantillon donné a été mesurée en divisant le nombre de lectures correspondant à l'OTU dans l'échantillon par le nombre total de lectures de l'échantillon. Les OTUs avec une abondance inférieure à 0.01% dans tous les échantillons ont été retirés. L'étape d'assignation des séquences à une espèce en utilisant des bases de données de séquences a ensuite été effectuée.

Suivant les articles et les objectifs, les bases de données et critères d'attribution étaient différents et sont résumés dans le Tableau 1.3.

1.4.2 Pipeline des données de séquençage obtenues avec l'enrichissement par capture

Les lectures forward et reverse appartenant à la même séquence ont été fusionnées avec un minimum de recouvrement de 10 nucléotides entre les deux lectures et une différence maximum sur la zone de recouvrement de 5 nucléotides. Étant donné que la sonication peut amener à des fragments de tailles différentes, et notamment à des fragments plus long que 600 nucléotides (longueur maximale du séquençage), les lectures qui n'ont pas fusionnées ont été conservées pour la suite du pipeline. Les lectures ont ensuite été filtrées selon leur qualité avec comme critères : une valeur ee (expected error, erreur attendue) maximale de 1, une longueur de séquence minimale de 150 nucléotides et 50 Ns autorisés. Le set de sondes contenait des sondes ciblant d'autres gènes (16S, NAD1, NAD4, NAD5, CYTB et ATP6) pour d'autres projets (cf Chapitre 4 - 3.2.2) et il était nécessaire de retirer les lectures leur correspondant pour la suite des analyses. Elles ont été identifiées en les assignant à une base de données interne contenant les gènes en question (Tableau 1.3). Les lectures identifiées comme du non-COI ont été retirées et les lectures restantes ont été assignées à une espèce en utilisant des bases de données de séquences. Étant donné que les séquences non fusionnées pouvaient toutes les deux passer les contrôles qualité et être assignées, le meilleur assignement était conservé (assignement avec le pourcentage d'identité la plus haute). Suivant les articles et les objectifs, les bases de données et critères d'attribution étaient différents et sont résumés dans le Tableau 1.3.

Tableau 1.3. Critères de conservation des alignements lors de l'étape d'assignement selon les articles et les objectifs de l'assignement. Les détails concernant les bases de données et leur construction sont présentés dans les articles correspondants.

article	Données issues de	Objectifs	Base de données et outil d'alignement utilisé	Critères
3	PCR	Assignement aux espèces des communautés artificielles	nucléotidique, région Folmer du COI (<i>section 4.2.2 ; Elbrecht & Leese, 2017</i>) outil d'alignement : blastn	e-value < 1 ^{E-10} identité > 97% ADN bulk : alignment length > 200 ADNet : alignment length > 90
3	PCR et Capture	Détection du COI d'autres espèces dans les communautés artificielles	protéique, COI complet (<i>8 espèces d'invertébrés de NCBI</i>) outil d'alignement : blastx (diamond 0.9.22)	option « more sensitive » e-value < 1 ^{E-10}
3	PCR et Capture	Estimation du taux de contamination	protéique, protéomes entiers (<i>100 de Ensembl, Ensembl Metazoa, 837 de Microbial Genome Database for Comparative Analysis</i>) outil d'alignement : blastx (diamond 0.9.22)	option « more sensitive » e-value < 1 ^{E-10}
3	Capture	Retrait des gènes non-cibles (16S, NAD1, NAD4, NAD5, CYTB et ATP6)	nucléotidique, séquences complètes de chaque gène sauf 16S (<i>8 espèces d'invertébrés de NCBI</i>) (<i>16S : 10 espèces séquencées en 16S, 52 taxa de NCBI</i>) outil d'alignement : blastn-short	e-value < 1 ^{E-10}
3	Capture	Assignement aux espèces des communautés artificielles	nucléotidique, région Folmer du COI (<i>section 4.2.2 ; Elbrecht & Leese, 2017</i>) outil d'alignement : blastn	e-value < 1 ^{E-10} identité > 97% alignment length > 250
4	PCR	Assignement à des espèces	nucléotidique, région Folmer du COI (<i>BOLD system, base de données publique</i>) outil d'alignement : BOLD system	critères inhérents à BOLD (Ratnasingham & Hebert, 2007) + vérification manuelle sur les arbres

1.5 Références

- Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W.K., Potter, C., Bik, H.M., Freckleton, R., 2016. The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution* 7, 1008-1018.
- Elbrecht, V., Leese, F., 2017a. PrimerMiner: an R package for development and in silico validation of DNA metabarcoding primers. *Methods in Ecology and Evolution* 8, 622-626.
- Elbrecht, V., Leese, F., 2017b. Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental Science* 5.
- Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3, 294-299.
- Hebert, P.D., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc Biol Sci* 270, 313-321.
- Hodges, E., Xuan, Z., Balija, V., Kramer, M., Molla, M.N., Smith, S.W., Middle, C.M., Rodesch, M.J., Albert, T.J., Hannon, G.J., McCombie, W.R., 2007. Genome-wide in situ exon capture for selective resequencing. *Nature Genetics* 39, 1522.
- Kircher, M., Sawyer, S., Meyer, M., 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* 40, e3.
- Mayer, C., Sann, M., Donath, A., Meixner, M., Podsiadlowski, L., Peters, R.S., Petersen, M., Meusemann, K., Liere, K., Wagele, J.W., Misof, B., Bleidorn, C., Ohl, M., Niehuis, O., 2016. BaitFisher: A Software Package for Multispecies Target DNA Enrichment Probe Design. *Mol Biol Evol* 33, 1875-1886.
- McCormack, J.E., Harvey, M.G., Faircloth, B.C., Crawford, N.G., Glenn, T.C., Brumfield, R.T., 2013. A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. *PLoS One* 8, e54848.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahe, F., 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584.
- Vamos, E., Elbrecht, V., Leese, F., 2017. Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics* 1.
- Wiegmann, B.M., Yeates, D.K., 2007. *Diptera. True Flies*.
- Yang, H., Ding, Y., Hutchins, L.N., Szatkiewicz, J., Bell, T.A., Paigen, B.J., Graber, J.H., de Villena, F.P.-M., Churchill, G.A., 2009. A customized and versatile high-density genotyping array for the mouse. *Nature Methods* 6, 663.

2 Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Titre en anglais : Enhancing DNA metabarcoding performance and applicability with bait capture enrichment and DNA from conservative ethanol

Gauthier M.^{1,2}, Konecny-Dupré L.¹, Nguyen A.³, Elbrecht V.⁴, Datry T.², Douady C.J.¹, Lefébure T.¹

¹ Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5023, ENTPE, Laboratoire d'Ecologie des Hydrosystèmes Naturels et Anthropisés, F-69622 Villeurbanne, France

² IRSTEA, UR-RiverLy, Centre de Lyon-Villeurbanne, Villeurbanne Cedex, France

³ Biofidal, Vaulx-en-Velin, France

⁴ Centre for Biodiversity Genomics, University of Guelph, 50 Stone Road East, Guelph, Ontario, N1G 2W1, Canada

DOI : 10.1111/1755-0998.13088

2.1 Résumé étendu

Une identification fiable et précise est fondamentale dans le domaine de l'écologie, que ce soit en science fondamentale pour inférer les processus structurant la biodiversité à différentes échelles spatiales et temporelles ou en science appliquée pour gérer et conserver les systèmes impactés par l'humain (Bortolus, 2008 ; Prié et al., 2016). Comme présenté en Introduction, l'identification des organismes sur critères morphologiques est parfois associée à une identification à résolution taxinomique grossière ou incorrecte (Creer et al, 2016). L'identification basée sur l'ADN a ainsi été proposée comme une alternative pour surmonter les limites liées à l'identification morphologique avec le développement du barcoding (Hebert et al, 2003) puis du metabarcoding (Pompanon et al, 2011). Dans le contexte de l'étude des communautés, le metabarcoding a été adopté comme un outil d'identification de routine dans le monde des micro-organismes (Abdelfattah et al, 2018) alors qu'il peine à se faire une place dans le monde des macroorganismes où l'identification morphologique a toujours été utilisée. A cette tradition de l'identification morphologique s'ajoutent des limites méthodologiques qui n'ont pas encore été résolues vis-à-vis du metabarcoding et qui limitent l'adoption systématique de l'identification moléculaire pour l'étude des macroorganismes (Elbrecht et al, 2017). En particulier, deux obstacles majeurs doivent être pris en considération. Premièrement, l'extraction de l'ADN à partir de l'homogénat des organismes (ADN

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

bulk), qui demande le tri et la destruction des organismes, étapes chronophages et coûteuses qui ne permettent pas la conservation des organismes pour des analyses ultérieures. Deuxièmement, l'enrichissement par PCR des barcodes ADN qui induit des erreurs dans l'estimation de la diversité des espèces et des biomasses relatives.

Dans cette étude, nous nous sommes intéressés à ces deux obstacles (i) en évaluant la capacité de l'enrichissement par capture par sondes ARN à détecter et estimer les biomasses relatives des espèces par rapport à l'enrichissement par PCR et (ii) en testant si l'ADN extrait de l'éthanol de conservation (ADNet) de communautés d'invertébrés pouvait être utilisé comme une alternative à l'ADN bulk en évaluant la détection et l'estimation des biomasses relatives des espèces par l'ADNet avec les deux méthodes d'enrichissement.

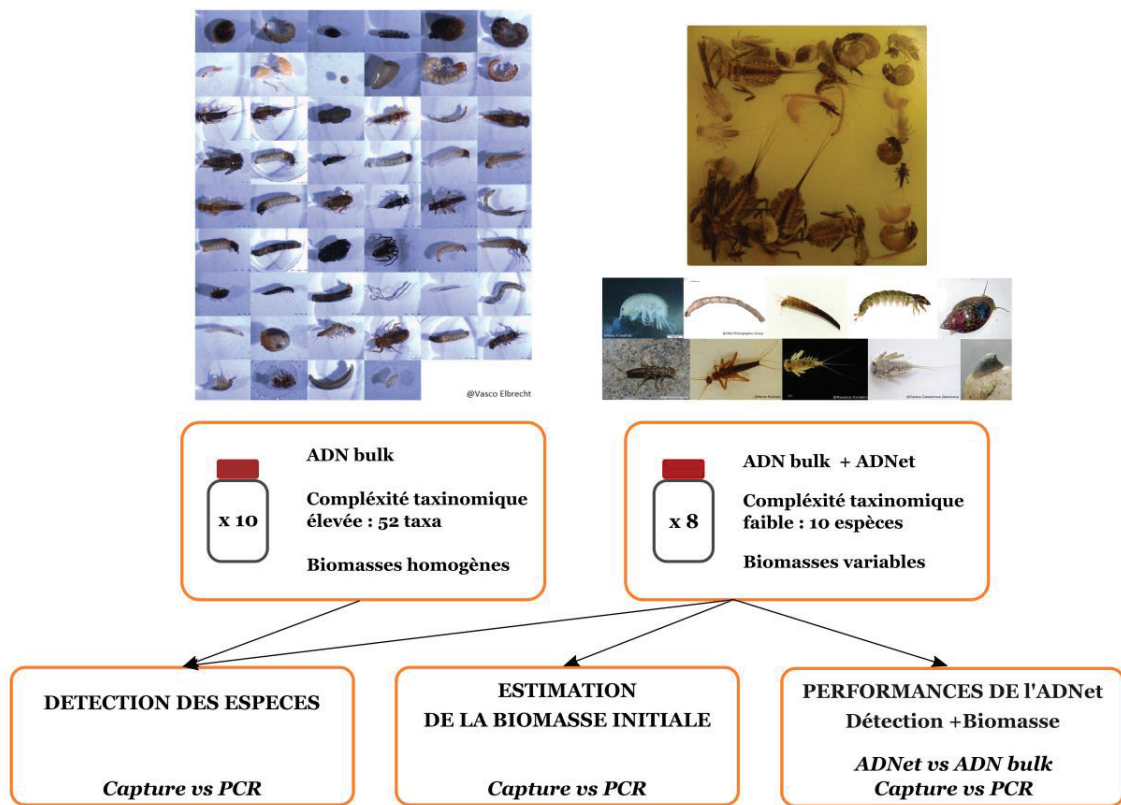


Figure 2.1 : Schéma simplifié du design expérimental avec les deux types de communautés utilisés, en regard des objectifs de l'article.

Pour cela, nous avons utilisé deux types de communautés artificielles avec des complexités taxinomiques différentes (Figure 2.1). Le premier type de communautés contenait 10 espèces d'invertébrés (complexité taxinomique faible) avec des biomasses sèches variables. Huit échantillons ont été constitués dont l'ADN a été extrait à partir de l'éthanol de conservation (ADNet) et de l'homogénéat par broyage des organismes séchés (ADN bulk). Le deuxième type de communautés était composé de 52 taxa (complexité taxinomique élevée) avec une biomasse homogène entre les taxa et comportait 10 échantillons. Ces échantillons ont été constitués par Elbrecht & Leese en 2015 et de l'ADN bulk uniquement en a été extrait, à partir des organismes séchés. Les ADN extraits ont été enrichis par capture par sondes ARN (set de 15 038 sondes, cf Matériel et méthodes - Chapitre 3) et par PCR avec le couple d'amorces BFR2 (amplicons de 421 paires de base (pb) ; Elbrecht & Leese, 2017). Les ADNet ont aussi été amplifiés avec un couple d'amorces ciblant un fragment plus court du COI de 178 pb (Fwh1, Vamos et al, 2017). La capture par sonde ARN n'étant pas utilisée classiquement pour le metabarcoding, le rendement et la spécificité de capture du COI ont été évalués. Les comparaisons des deux méthodes d'enrichissement et des deux matrices ADN (ADN bulk et ADNet) vis-à-vis de la détection et de l'estimation de la biomasse initiale des espèces ont été faites en comparant la sensibilité de détection des espèces et les relations entre la biomasse initiale et le nombre de lectures issues du séquençage mesurée par des modèles logistiques à effets mixtes.

Nos résultats montrent que la capture par sondes ARN détecte plus d'espèces que l'enrichissement par PCR, quelles que soient la matrice ADN ou la complexité taxinomique de l'échantillon. Pour les deux méthodes d'enrichissement, l'absence de détection d'une espèce était liée à la distance à l'oligonucléotide, que ce soit l'amorce pour la PCR ou la sonde pour la capture, en accord avec de précédentes études (Liu et al, 2015 ; Portik et al, 2016) et l'hypothèse selon laquelle cette distance est une source majeure de biais pour la détection des espèces (Elbrecht et al, 2017). Cependant, la capture par sondes est plus robuste à cette distance car (i) elle est moins sensible aux différences entre la séquence et l'oligonucléotide (ou mismatches) avec une divergence maximale de 15 % pour la capture contre 5% pour la PCR pour permettre la détection d'une espèce, (ii) la probabilité de rencontrer des mismatches est réduite du fait de la conception de centaines de sondes différentes pour chaque région de 120 bases et (iii) plusieurs régions de 120 bases peuvent être ciblées (5 dans cet article) augmentant la probabilité qu'une sonde fonctionnelle soit conçue pour une espèce donnée, même si elle n'était pas présente dans la base de données lors de la conception des sondes.

La capture par sondes a aussi donné de meilleures estimations des biomasses initiales des espèces que l'enrichissement par PCR. En effet, les relations entre les biomasses initiales et le nombre de lectures étaient plus proches d'une relation linéaire attendue 1:1 pour l'enrichissement par capture. Cependant, un

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

effet espèce a aussi été mis en évidence indiquant que les relations entre les biomasses initiales et le nombre des lectures sont différentes entre espèces, impliquant une difficulté à estimer précisément les biomasses initiales sans connaître pour chaque espèce la valeur de la relation. Cette variation pourrait être expliquée par les biais liés aux distances entre sondes et séquences déjà évoqués mais peuvent aussi être inhérentes au gène d'intérêt, qui est mitochondrial. En effet, la quantité d'ADN mitochondrial dépend du nombre de génomes mitochondriaux par cellule qui peut varier entre espèces, entre stades de développement et même entre tissus (Rooney et al, 2015). Ainsi, reconstituer les biomasses initiales de toute une communauté sans connaître les corrections spécifiques à appliquer aux estimations semble très compliqué avec de l'ADN mitochondrial. Cependant, nous avons pu reconstituer les rangs de biomasses des espèces au sein des échantillons, ce qui pourrait être assez informatif dans certains contextes comme la bio-évaluation (Elbrecht et al, 2017).

Concernant l'ADN extrait à partir de l'éthanol, nos résultats montrent qu'il pourrait être utilisé pour remplacer l'ADN bulk dans des études où la détection des espèces seulement est nécessaire. En effet, la détection des espèces était équivalente entre l'ADNet et l'ADN bulk quelle que soit la méthode d'enrichissement utilisée. Cependant, l'ADNet n'a pas permis d'estimer les biomasses initiales des espèces, ce qui peut s'expliquer par le relargage différentiel d'ADN entre les espèces quand elles sont placées dans l'éthanol. L'ADNet est donc une méthode rapide et non destructive pour décrire qualitativement un échantillon mais ne semble pas être utilisable pour le décrire quantitativement.

Bien que l'enrichissement par capture par sondes ARN donne une représentation plus fiable et plus précise de la présence des espèces et des biomasses relatives par rapport à la PCR et soit déjà un outil efficace pour le metabarcoding, des optimisations sont possibles. Ces optimisations concernent (i) l'efficacité de capture qui, malgré un rendement fort (au moins 214 fois plus de COI capturé avec la capture par rapport à un séquençage aléatoire), montre une spécificité variable qui semble dépendre de la complexité taxinomique et de l'échantillon, de la capture possible de NUMTs et du pipeline bio-informatique choisi ; (ii) la conception des sondes en réduisant la distance entre les sondes et les séquences d'intérêt par rapport au seuil de 15% ; (iii) le protocole de laboratoire avec un travail nécessaire sur certains facteurs clés comme la température d'hybridation ou le regroupement des échantillons lors de l'hybridation des sondes et (iv) le traitement bio-informatique des lectures qui nécessiterait un pipeline dédié aux lectures de metabarcoding enrichies par capture.

2.2 Abstract

Metabarcoding is often presented as an alternative identification tool to compensate for coarse taxonomic resolution and misidentification encountered with traditional morphological approaches.

However, metabarcoding comes with two major impediments which slow down its adoption. First, the picking and destruction of organisms for DNA extraction are time and cost consuming and do not allow organism conservation for further evaluations. Second, current metabarcoding protocols include a PCR enrichment step which induces errors in the estimation of species diversity and relative biomasses. In this study, we first evaluated the capacity of capture enrichment to replace PCR enrichment using controlled freshwater macrozoobenthos mock communities. Then, we tested if DNA extracted from the fixative ethanol (etDNA) of the same mock communities can be used as an alternative to DNA extracted from pools of whole organisms (bulk DNA). We show that capture enrichment provides more reliable and accurate representation of species occurrences and relative biomasses in comparison with PCR enrichment for bulk DNA. While etDNA does not permit to estimate relative biomasses, etDNA and bulk DNA provide equivalent species detection rates. Thanks to its robustness to mismatches, capture enrichment is already an efficient alternative to PCR enrichment for metabarcoding and, if coupled to etDNA, is a time-saver option in studies where presence information only is sufficient.

2.3 Introduction

Reliable and accurate taxa identification is fundamental in biological sciences. Poor taxonomic identification can lead to cascades of error affecting our knowledge and understanding not only in theoretical and fundamental biology, but also in applied fields leading to poor management decisions (Bortolus, 2008). It distorts our ability to infer processes in ecology and evolution, to manage and conserve human-impacted systems and to carry out human health and resource programs (Bortolus, 2008; Leys, Keller, Räsänen, Gattolliat, & Robinson, 2016; Prié, Puillandre, & Bouchet, 2013). Poor taxonomic identification occurs when identification to species level is not possible (coarse taxonomic resolution) or when identification is incorrect (misidentification). For example, Martin, Adamowicz, and Cottenie (2016) investigated macrozoobenthos community distribution in freshwater streams at different taxonomic level (family, genus and species) and found a spatial structure only when the species identification level was reached. In Vietnam, misidentification in *Anopheles* species lead to mismanagement in control program of the Malaria disease as the main target species was a non-vector one (Van Bortel et al., 2001). This kind of difficulties leading to poor taxonomic identification and thus to misinterpretation and erroneous conclusions is often associated to morphological based taxonomy (Baird & Hajibabaei, 2012; Creer et al., 2016). Cryptic species, limited expertise, damaged or juvenile specimens and even cost and time constraints in the case of applied research are all problems encountered in morphological identification leading to misidentification or coarse taxonomic resolution (Bringloe, Cottenie, Martin, & Adamowicz, 2016; Hajibabaei, Baird, Fahner, Beiko, & Golding, 2016; Ji et al., 2013; Sweeney et al., 2011).

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

In the past decade, DNA-based identification has been proposed as an alternative to morphological approaches. DNA barcoding uses the sequence of a genetic marker of one specimen, hereafter called a DNA barcode, usually of an organelle genome for eukaryotes (e.g. mitochondria for animals and chloroplast for plants, see Creer et al. (2016) for overview) and assign it to a species name within a reference database (Hebert, Ratnasingham, & de Waard, 2006). DNA barcoding supposedly addresses limitations in morphological identification by accurately discriminating species regardless of their morphology (Sweeney et al., 2011), development stages (Hubert, Delrieu-Trottin, Irisson, Meyer, & Planes, 2010) or sex (Forshaw, 2010). Recent advances in high throughput technologies enabled the emergence of DNA metabarcoding, the barcoding of pool of specimens in a single reaction which permit to work on whole community at once (Pompanon et al, 2011). Metabarcoding has already been adopted as a routine identification tool in microbial community ecology (Abdelfattah, Malacrinò, Wisniewski, Cacciola, & Schena, 2018). In macro-organism community studies, the last ten years have seen the development of invertebrate derived DNA (iDNA, (Calvignac-Spencer et al., 2013; Schnell et al., 2012)) and the analysis of DNA associated with environmental matrices such as soil or water (eDNA, (Harper et al., 2018; Thomsen & Willerslev, 2015)). Metabarcoding was instrumental in the emergence of these new approaches for which, by definition, there is no non-molecular alternative. Conversely, in conventional macro-organism community studies where the traditional identification using morphological criteria has been used for many years, metabarcoding is still currently marginally used as a routine identification tool. Yet, metabarcoding could enhance the capacity to characterize these communities through iDNA, eDNA or through the extraction of DNA from a homogenate of whole organisms collected together (bulk DNA; (Deiner et al., 2017)). A major limitation to the democratization of metabarcoding is the absence of standardized metabarcoding protocol that has been established and validated by the scientific community (Leese et al., 2018). In particular, methodological roadblocks are encountered throughout each step of sample processing (i.e. DNA extraction, enrichment of the targeted DNA barcode, sequencing, bioinformatic treatment and taxonomic assignment). Hereafter, we focused on two major metabarcoding roadblocks: specimen picking before extraction and DNA barcode enrichment.

Specimen picking is a major concern for bulk DNA samples and is particularly critical when the ratio of targeted organism over substrate is low (e.g. invertebrates sampled in streambed). It aims to separate individuals from substrate (e.g. leaves, sand...) prior to DNA extraction to avoid PCR inhibition (Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017) and to limit the quantity of material to be processed during DNA extraction. This step is time and cost consuming. Direct extraction from fixative agent, usually ethanol, has been proposed as a time-saver alternative to specimen picking (Hajibabaei, Spall, Shokralla, & van Konynenburg, 2012; Zizka, Leese, Peinert, & Geiger, 2018). When fixative agent is used as a DNA template, organisms are not destroyed and are conserved for further taxonomic work or

downstream analyses (Leese et al., 2016). Similar techniques have already been employed by the ancient DNA community where the destruction of samples such as museum specimens is to be avoided (Paijmans, Fickel, Courtiol, Hofreiter, & Forster, 2016; van der Valk, Lona Durazo, Dalen, & Guschanski, 2017). Yet, little research on bulk DNA has been conducted on this alternative (but see Hajibabaei et al. (2012); Zizka et al. (2018)) and a rigorous comparison with traditional specimen picking is warranted before it can be used as a standard template of DNA in community studies.

PCR enrichment bias is often considered as the most problematic roadblock in metabarcoding because it may alter species detection and relative abundance recovery of species (Elbrecht & Leese, 2017; Leese et al., 2018; Piñol, Senar, & Symondson, 2018). Prior to sequencing, DNA barcodes are first amplified by PCR using primers that may not have the same number of mismatches with the targeted sequences across taxa (Piñol, Mir, Gomez-Polo, & Agustí, 2015). During a PCR, if DNA barcode sequences from two different species are in equimolar concentration, but that the first species presents less mismatches with the primers, this species' sequence will be amplified preferentially. Consequently, amplification efficiency is expected to be non-equal across taxa, leading from under-amplification to no amplification of some taxa in the worst case scenario. PCR bias was demonstrated for fungi (Bellemain et al., 2010), bacteria (Frank et al., 2008), invertebrates (Piñol et al., 2015) and vertebrates (Arif, Khan, Al Sadoon, & Shobrak, 2011). Other factors affect PCR like GC content (Aird et al., 2011) or inhibitors that can remain after DNA extraction but primer bias is commonly presented as the major cause of biases in metabarcoding (e.g. (Elbrecht & Leese, 2015; Piñol et al., 2015; Pinto & Raskin, 2012)). In consequence, a lot of work focused on primer design to decrease PCR biases with, for instance, the use of several primer pairs, degenerated primers or amplification of several DNA barcodes (Drummond et al., 2015; Elbrecht & Leese, 2017; Elbrecht et al., 2016; Gibson et al., 2015; Jusino et al., 2019; Leray & Knowlton, 2017; Zhang, Chain, Abbott, & Cristescu, 2018). These efforts increased the species detection rate but the quantitative bias was not solved completely (Piñol et al., 2018).

Avoiding PCR enrichment will, by definition, solve the PCR bias issue (Porter & Hajibabaei, 2018). Low (Linard, Crampton-Platt, Gillett, Vogler, & Timmermans, 2015) or high (Porter & Hajibabaei, 2018) coverage metagenome sequencing (i.e. sequencing a community DNA without any enrichment) can be used to assemble entire organelle genomes. This approach provides an efficient way to recover species richness and taxa relative biomass, although the proportion of organelle reads is extremely low making metagenome sequencing much more expensive than PCR metabarcoding (Bista et al., 2018; Gómez-Rodríguez, Crampton-Platt, Timmermans, Baselga, & Vogler, 2015; Zhou et al., 2013). Furthermore, only a small part of an organelle genome is usable for taxonomic assignment as reference databases mostly contain DNA barcode sequences (e.g. COI for metazoan, 16S for bacteria, ITS for fungi, (Creer et

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

al., 2016)). Although methods are being developed to reduce organelle genome sequencing cost (Macher, Zizka, Weigand, & Leese, 2017), the construction of exhaustive organelle genome reference databases will be a long-term and expensive process. Another PCR-free alternative is capture enrichment where targeted sequences hybridize to baits and are retrieved by magnetism (Dowle, Pochon, C. Banks, Shearer, & Wood, 2016). Contrary to metagenome sequencing, capture enrichment increases the proportion of targeted reads reducing the sequencing cost (Jones & Good, 2016). Baits are long oligonucleotides (more than 60 bp) which are designed from reference sequences. Capture enrichment has already been extensively used in genomics and genetics where it is used to retrieve thousands of loci (e.g. exons capture, (Hodges et al., 2007)) and SNPs arrays (Yang et al., 2009) of one species in a single reaction prior to sequencing. It is also commonly used for ancient DNA to both increase the amount of endogenous DNA and allow loci enrichment where DNA fragmentation tends to complicate PCR reactions (Avila-Arcos et al., 2011; Horn, 2012). Capture enrichment has also recently been used in phylogenetics where the goal is to target thousands of loci in many related species (McCormack et al., 2013; Phuong & Mahardika, 2018). Conversely, capture enrichment is not commonly used in community ecology (but see Dowle, Pochon, J, Shearer, & Wood, (2016); Shokralla et al., (2016)) which opens new challenges as in this new context the diversity of organism is extremely high but the number of targeted loci is rather small. Capture enrichment should be more robust to identify taxa in a community than PCR enrichment because (i) thousands of different baits can be designed and (ii) when sequences are unknown or species are polymorphic, few mismatches between the baits and the targeted sequences should not bias DNA enrichment as they would do with PCR primers (Li, Hofreiter, Straube, Corrigan, & Naylor, 2013; Pajmans et al., 2016; Portik, Smith, & Bi, 2016). For example, one bait designed for a species of the genus *Danio* permitted to detect others *Danio* species in the Amazon basin where the reference species was absent (Mariat et al., 2018). In a single study, capture enrichment was compared to PCR enrichment and was found to detect more taxa but was unable to estimate relative abundances (Dowle et al., 2016). Liu et al. (2016) and Wilcox et al. (2018) have shown that relative abundances can be recovered with capture enrichment if species-specific corrections were to be applied to take into account variation in the number of mitochondria copy number among species (Liu et al., 2016; Wilcox et al., 2018). Such corrections are unfortunately not suitable to the complexity of field samples. Dowle et al. (2016) study was based on natural communities only described at a coarse taxonomic level (family to genus level) and with indirect biomass measurements. Thus, the capacity of capture enrichment to describe both community diversity, and relative biomass without species-specific biomass corrections, requires further testing with controlled communities.

In this study, we first investigated the capacity of cytochrome oxidase subunit I (COI) gene capture enrichment to detect taxa and retrieve initial biomass without species specific correction. Second, we

evaluated if DNA extracted from ethanol (etDNA) can be used as an alternative to organism picking and homogenization (bulk DNA) by assessing species detection and initial biomass recovery of this alternative template DNA with PCR and capture enrichment. Tests were carried out using two types of freshwater mock communities (MC): (i) low diversity MC (10 species) with variable dry biomass across taxa and (ii) high diversity MC (52 taxa) with homogeneous biomass across taxa from (Elbrecht & Leese, 2015).

2.4 Materials & methods

2.4.1 Mock community design

Freshwater macrozoobenthos specimens for the 10 species mock communities (MC) were sampled from various streams in east of France during May 2017, except for two species, *Gammarus fossarum* and *Chironomus riparius*, which came from Irstea livestock, France (ECOTOX team, RiverLy, France). The 10 species were chosen to represent a wide taxonomic range and because they can be easily identified to the species level by the naked eye (Tableau 2.1 Tableau 2.1. Dry biomass in mg and number of individuals (*italic*) of freshwater invertebrate species in the 10 species mock communities.). Two hundred milliliters of ethanol 96% (EtOH) were dispense in bottles and individuals were placed alive in the ethanol. Eight samples with different relative biomass for the 10 species were constructed (Figure 2.2, Tableau 2.1). Samples were then stored for 6 months at 4°C until DNA extraction. DNA was extracted from the whole organisms (bulk DNA) and from the preservative EtOH (etDNA).

The 52 taxa mock communities were designed by Elbrecht and Leese (2015). In a nutshell, for ten different MC, a roughly similar dry biomass was collected from 52 taxa identified to the lowest taxonomic level based on morphology, and the homogenized tissue pool was then extracted for each sample using a salt extraction protocol (Figure 2.2, see Elbrecht & Leese (2015) for details). Because the biomass was fairly homogeneous among taxa in these MC, we only used these MC to compare the performance of capture and PCR enrichment in predicting species occurrences.

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Tableau 2.1. Dry biomass in mg and number of individuals (*italic*) of freshwater invertebrate species in the 10 species mock communities.

Species	<i>Mock community</i>							
	1	2	3	4	5	6	7	8
<i>Chironomus riparius</i> Meigen, 1804	1.84 (5)	1.83 (5)	2.58 (5)	2.42 (5)	3.66 (10)	4.62 (11)	3.89 (10)	4.44 (10)
<i>Epeorus assimilis</i> Eaton, 1885	102.32 (5)	55.53 (6)	77.89 (5)	96.47 (5)	62.94 (4)	110.73 (4)	114.80 (4)	74.23 (4)
<i>Heptagenia sulphurea</i> (O.F. Müller, 1776)	15.73 (5)	20.79 (5)	21.08 (5)	18.02 (5)	38.31 (10)	28.19 (10)	32.58 (10)	38.37 (10)
<i>Isoperla rivulorum</i> (Pictet, 1841)	37.97 (5)	31.88 (5)	36.28 (5)	30.89 (5)	35.11 (6)	33.39 (6)	36.22 (6)	39.47 (6)
<i>Nemurella picteti</i> Klapálek, 1900	2.83 (5)	2.49 (5)	2.34 (5)	3.60 (5)	0.50 (1)	0.97 (1)	0.77 (1)	0.56 (1)
<i>Hydropsyche siltalai</i> Doehler, 1963	68.42 (5)	53.15 (5)	53.63 (5)	53.32 (5)	80.36 (8)	66.27 (8)	86.71 (8)	79.86 (8)
<i>Athripsodes aterrimus</i> (Stephens, 1836)	6.10 (6)	6.68 (5)	7.87 (5)	6.23 (5)	9.10 (6)	7.96 (6)	5.39 (6)	7.01 (6)
<i>Physella acuta</i> (Draparnaud, 1805)	23.30 (5)	24.85 (5)	23.03 (5)	31.30 (5)	16.81 (4)	15.22 (4)	13.25 (4)	20.68 (4)
<i>Ancylus fluviatilis</i> O.F. Müller, 1774	13.42 (5)	18.16 (5)	18.40 (5)	15.42 (5)	20.79 (8)	20.47 (8)	23.21 (8)	15.98 (8)
<i>Gammarus fossarum</i> Koch, 1836	23.07 (5)	20.28 (5)	16.89 (5)	16.72 (5)	4.4 (1)	4.02 (1)	5.17 (1)	3.76 (1)

2.4.1 DNA extraction of the 10 species mock communities

Bulk DNA. For each MC, individuals were picked up and sorted by species in petri dishes. They were let to dry overnight and dry biomass for each species was weighted. Mollusc shells and Trichoptera cases were removed prior to the weighing. Individuals were pooled together and the entire community was grounded with a bead mill MM200 (Retsch) during 4 min at 30 Hz. The whole homogenate was extracted (mean \pm SD: 245.8 \pm 28.6 mg) with FastDNA® Spin Kit for Soil (MP Biomedicals, USA) following manufacturer's protocol. The extracted DNA was purified with Agencourt AMPure XP purification beads (Beckman Coulter, USA) to further remove solvents. A negative extraction control was also made, following the same extraction protocol but without starting material.

etDNA. For each mock community, roughly 50 mL of preservative ethanol was collected. Glycogen and sodium acetate were added to precipitate DNA and samples were placed at -80°C for at least 72 hours. After centrifugation, ethanol was removed and after total ethanol evaporation, the dried residue was dissolved in the buffer solution of NucleoSpin Tissue® kit (Macherey-Nagel GmbH, Germany) and DNA was extracted following manufacturer's instructions. As the amount of DNA was limiting, extractions were repeated twice: a first extraction to compare PCR and capture enrichment methods

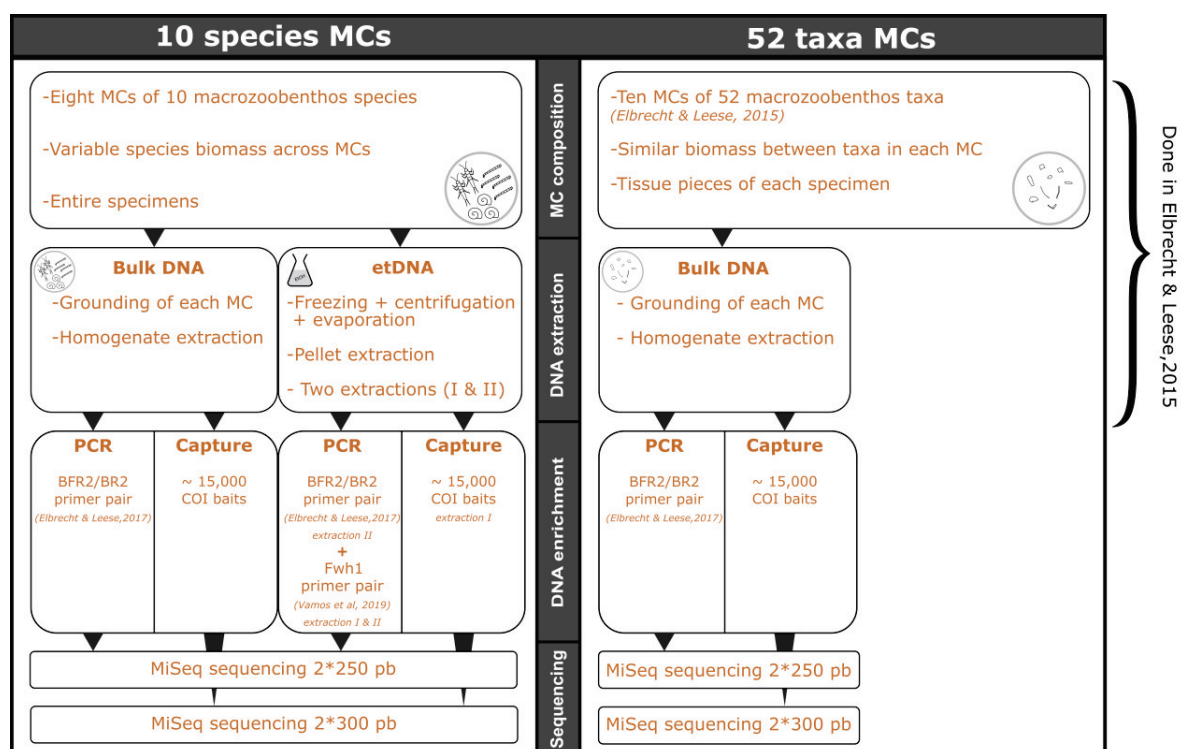


Figure 2.2 : Overview of the experimental design of the study. Two different mock communities (MCs) were used: 10 species MCs and 52 taxa MCs. Ten species MCs were used to assess initial biomass recovery and etDNA performances with PCR and capture enrichment. Both MC were used to assess taxa detection. etDNA: ethanol DNA.

(extraction I); a second to compare two different primer pairs performance for this supposedly degraded DNA template (extraction II). Four negative extraction controls were also made during extraction II by filling tubes with 96% ethanol which were then extracted following the same protocol.

2.4.2 COI mock community reference databases

The 10 species reference database was built by sequencing the COI Folmer region of each species of the 10 species mock communities (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994). One specimen per species was extracted following the same protocol as for bulk DNA. PCR reactions were performed in a total volume of 25 μ L with 1X of PCR standard buffer (including 3 mM MgCl₂, Eurobio, France), 0.05 U/ μ L of EurobioTaq DNA polymerase (Eurobio, France), 0.8 mM of each dNTP (Eurogentec, Belgium), 0.1 mg/mL of BSA (New England BioLabs, USA), 0.4 μ M of each primer (LCO1490/HCO2198, Table S1) and 0.5 μ L of template DNA. The amplification consisted in an initial denaturation at 95°C for 3 min, followed by 40 cycles of denaturation at 95°C for 20 s, annealing at 51°C for 30 s and extension at 72°C for 45 s, with a final extension at 72°C for 5 min. Purified template DNA was sequenced on both strands with the PCR primers using standard Sanger sequencing (Biofidal, France). All sequences were manually

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

checked for errors and cleaned up with Finch TV software 1.4.0. This protocol did not work for one species (*Heptagenia sulphurea*) and its sequence was downloaded from NCBI to complete the 10 species reference database (GenBank Accession Number HE651395.1).

The 52 taxa reference database was already available for the 52 taxa MC (Elbrecht & Leese, 2017). It includes the haplotypes of all organisms used in their experiment, leading to a reference database of 212 COI sequences. For one taxon (Nematoda), Elbrecht and Leese (2017) could not get any amplification. Since this taxon is missing in Elbrecht and Leese (2017) haplotype reference database, it was not taken into account for downstream analysis.

2.4.3 Amplicon sequencing and analysis

Library preparation and sequencing. For the 10 species and 52 taxa MC bulk DNA and etDNA (extraction II), a 421 bp fragment within the COI Folmer region was amplified with the BF2/BR2 primer set with Illumina Nextera tails which is well suited for freshwater macrozoobenthos (Elbrecht & Leese, 2017), Table S1). For etDNA extractions I and II, a shorter region of 178 bp (Fwh1 primer set with Illumina Nextera tails, (Vamos, Elbrecht, & Leese, 2017), Table S1) was also targeted. PCR reactions were performed in triplicates in a total volume of 25 μ L with 1X of PCR buffer (including 3mM MgCl₂ and 400 μ M each dNTP, QIAGEN® Multiplex PCR kit, Qiagen, Germany), 0.5 μ M of each primer and 10 ng of template DNA for bulk DNA or 5 μ L of template DNA for etDNA. The amplification consisted in an initial denaturation at 95°C for 5 min, followed by 25 (bulk DNA) or 35 (etDNA) cycles of denaturation at 95°C for 30 s, annealing at 50°C (bulk DNA) or 52°C (etDNA) for 30 s and extension at 72°C for 2 min, with a final extension at 72°C for 10 min. Triplicate PCR products were pooled and purified with Agencourt AMPure XP purification beads (Beckman Coulter, USA) and quantified using QuantiFluor® dsDNA System (Promega, USA). Ten ng of each purified PCR products were then used in a second PCR to dual index each sample with a unique tag combination (Table S3 for tag combinations details) and to add Illumina adapters. PCR reactions were performed in a total volume of 25 μ L with 1X of PCR buffer (BIOAmp® Blend Mix, Biofidal, France), 2.5 nM of MgCL₂, 200 μ M of each dNTP, 0.25 μ M of Illumina primer (Table S1) and 0.02 U/ μ L of HOT BioAmp Taq (Biofidal, France). PCR products were purified with Agencourt AMPure XP purification beads (Beckman Coulter, USA), quantified using QuantiFluor® dsDNA System (Promega, USA) and pooled at the same concentration (2 nM) for sequencing. The PCR amplicon libraries were sequenced using a 2*250 paired-end V3 MiSeq sequencing kit (Biofidal, France). Negative controls were included to every PCR during the library preparation.

Bioinformatic analysis. Reads were delivered demultiplexed and adapter trimmed. The reads of the 52 taxa mock communities, bulk DNA and etDNA of the 10 species mock communities were processed

independently with the same bioinformatic pipeline. First, forward and reverse reads were merged with Vsearch 2.8.4 with a minimum overlapping of 10 nucleotides and a maximum difference of the overlapping region of 5 nucleotides (Rognes, Quince, Nichols, Flouri, & Mahé, 2016). Then, the primer regions were removed from the merged reads with cutadapt 1.9.1 (M. Martin, 2014) and the reads were quality filtered (maximum expected error (ee) of 1, minimum length of 200, no Ns allowed) and dereplicated with Vsearch. Sequences observed less than twice (< 2 reads) were removed. Then, chimeras were de novo detected and removed with Vsearch. Finally, sequences were clustered with a similarity cutoff of 97% identity into MOTUs (Molecular Operational Taxonomic Units). MOTUs were assigned to species using the 10 species reference database or the 52 taxa reference database with the blastn algorithm (Camacho et al., 2009). Only alignments with an e-value under 1E-10, a query cover over 200 bases for bulk DNA and over 90 bases for etDNA and an identity over 97% were conserved as good alignments for further analysis. MOTUs were also compared to a complete COI protein reference database made from 8 taxa (Asellus aquaticus ADA69754.1, Daphnia pulex AAD33231.1, Dinocras cephalotes AGZ03516.1, Gammarus fossarum YP_009379680.1, Physella acuta YP_008994230.1, Radix balthica HQ330989.1, Sericostoma personatum AJR19241.1, Thremma gallicum AJR19254.1) present in one or both MCs using diamond 0.9.22 (blastx, more sensitive option, e-value threshold of 1E-10 (Buchfink, Xie, & Huson, 2014)). This allowed us to detect COI sequences even if they did not belong to the species used in the MC. We also estimated the rate of contamination by comparing the quality filtered reads to a protein database containing the proteomes of 100 eukaryotic species from Ensembl (Zerbino et al., 2017) and Ensembl Metazoa (Kersey et al., 2017) as well as the proteomes of 837 prokaryotic species retrieved from the Microbial Genome Database for Comparative Analysis (Uchiyama, Mihara, Nishide, & Chiba, 2015) selecting one species per genus. Sequences were assigned to coarse taxonomic groups (archaea, eubacteria, fungi, plant, protist, protostomia— here Arthropoda, Annelida, Brachiopoda, Mollusca, Nematoda and Platyhelminthes – and other metazoa – corresponding to Deuterostomia) using the best diamond hit (blastx, more sensitive option, e-value threshold of 1E-10).

2.4.4 Capture sequencing and analysis

Bait design. While one of the primary goals of the present study is to assess capture versus PCR enrichment efficiency using mock communities, we chose to develop a larger set of baits that could be used in future assessment of French freshwater macrozoobenthos diversity. Up to now and aside from non-Chironomidae dipteran, 3,245 macrozoobenthos species belonging to 22 orders have been identified in small and medium streams of France (Aukema & Rieger, 2013; Coppa, 2019; D'Hondt & Ben Ahmed, 2009; Dusoulier, 2008; Gargominy et al, 2011; Grand & Boudot, 2007; Henry & Magniez, 1983; Le Doaré & Vinçon, 2019; Pattée & Gourbault, 1981; Piscart & Bollache, 2012; Queney, 2011; Serra et al,

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

2015; Souty-Grosset et al, 2006; Thomas, 2019; Tillier, 2019; Vallenduuk, 2004). One hundred twenty two species of six orders can be considered as marginal in streams and were thus discarded. All the available COI sequences were downloaded for the 3,123 remaining species from GenBank and BOLD with PrimerMiner 0.18 in December 2016 (Elbrecht & Leese, 2016). Four orders had less than half of their known taxa with available COI barcode and were discarded. Within the 12 remaining orders, the species without COI sequences were barcoded whenever possible, i.e. when tissues could be obtained. Organisms were extracted using Chelex (BioRad, USA) and the PCR and sequencing conditions were identical to those of section 3 of the methods. Sequences were deposited on GenBank (97 species, GenBank Accession Numbers MK584300:MK584515). At the end, 1,525 species out of 1,689 known species had a DNA barcode available, representing more than 90% of the targeted species (Table S2). For bait development, sequences were processed by order. First, COI sequences were aligned with blastx 2.7.1 to a reference *Drosophila yakubai* COI sequence (Accession Number: NC_001322.1) and identical sequences were collapsed using a perl script (http://github.com/TristanLefebure/collapse_to_uniq_seq). Then, 120 bp baits were constructed in silico using BaitFisher with a tiling of 60 bp and a cluster threshold of 5%, leading to a total of 15,038 baits generated from 100,367 unique sequences (Mayer et al., 2016). The COI in silico baits were then sent for RNA bait synthesis to Arbor Biosciences (USA).

Library preparation, hybridization and sequencing. Starting quantity of DNA was 1 µg for bulk DNA (10 and 52 taxa MC) and between 14 and 57 ng for etDNA (extraction I). DNA of each sample was sheared into approximately 600 bp nucleotide fragments by ultrasound sonication with a Qsonica Q800R (Qsonica, USA). Library preparation was conducted using NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England BioLabs, USA) following manufacturer's instructions. Briefly, after sheared DNA end repair, the 5' ends were phosphorylated and the 3' ends were A-tailed. Then, Illumina Nextera tails (Table S1) were ligated to the DNA fragments followed by a clean-up and a size selection of 500-700 nucleotides long fragments with Agencourt AMPure XP purification beads (Beckman Coulter, USA). Finally, DNA fragments were amplified to dual index the libraries (Table S3 for tag combinations details) and to add Illumina adapters (Table S1). COI capture enrichment was conducted using myBaits® Custom kit following manufacturer's instructions (Arbor Biosciences, USA). One hundred ng of library DNA was used for capture enrichment for bulk DNA and 230 ng for etDNA following manufacturer's instructions for degraded DNA. Baits were diluted 10 times and hybridization lasted 24h for both DNA templates with the exception that hybridization for etDNA was done at 55°C instead of 65°C following manufacturer's instructions for degraded DNA. The final library amplification was performed in a total volume of 50 µL per reaction with the KAPA HiFi DNA Polymerase (Kapa Biosystems, USA) with P5 and P7 Illumina primers (Table S1) using the following conditions: initial denaturation at 98°C for 2 min, followed by 21 cycles of denaturation at 98°C for 20 s, annealing at 60°C for 30 s and extension at 72°C for 1 min, with a

final extension at 72°C for 5 min. Capture library concentrations were determined by qPCR with a KAPA qPCR kit (KAPA Library Quant Kit, Kapa Biosystems, USA) and pooled at the same concentration for sequencing. The capture libraries were sequenced using a 2*300 paired-end V3 MiSeq sequencing kit (Biofidal, France).

Bioinformatic analysis. Reads were delivered demultiplexed and adapter trimmed. As for PCR analysis, the reads of the 52 taxa MC, bulk DNA and etDNA of the 10 species MC were processed independently with the same bioinformatic pipeline. First, forward and reverse reads were merged with Vsearch with a minimum overlapping of 10 nucleotides and a maximum difference of the overlapping region of 5 nucleotides (Rognes et al., 2016). Because shearing could lead to fragment longer than 600 nucleotides, merged and non-merged sequences were conserved for downstream bioinformatic steps. Reads were then quality filtered (maximum ee of 1, minimum length of 150, 50 Ns allowed) with Vsearch. Because the bait set contained baits for other genes for other projects (i.e. 16S, NAD1, NAD4, NAD5, CYTB, and ATP6) and because their presence in the bait set can alter downstream results, the reads corresponding to these genes were removed from the quality filtered reads in all downstream analysis. They were recovered with a blastn on a reference database (blastn-short, e-value threshold of 1E-10) containing the complete sequence of NAD1, NAD4, NAD5, CYTB and ATP6 genes of 8 species (*Asellus aquaticus* ADA69754.1, *Daphnia pulex* AAD33231.1, *Dinocras cephalotes* AGZ03516.1, *Gammarus fossarum* YP_009379680.1, *Physella acuta* YP_008994230.1, *Radix balthica* HQ330989.1, *Sericostoma personatum* AJR19241.1, *Thremma gallicum* AJR19254.1), the 16S of the 10 species MC (Table S1) and the 16S corresponding to the 52 taxa MC (downloaded from NCBI). The remaining reads were assigned to species using the 10 species reference database or the 52 taxa reference database using BLAST algorithm (Camacho et al., 2009). Only alignments with an e-value below 1E-10, a query cover over 250 and an identity over 97 were used for taxonomic assignment. As for amplicon, filtered reads were also compared to a complete COI protein reference database to estimate the total number of COI reads and to a protein database to estimate the rate and origin of contaminations.

2.4.5 Capture efficacy

Capture efficacy (sensus (Cha & Thilly, 1993)) was evaluated by measuring the percentage of targeted reads for each capture library (i.e. capture specificity) and the X-fold enrichment (i.e. capture efficiency) as in Maggia et al. (2017). The percentage of targeted reads is the ratio of the number of reads assigned with diamond on the COI protein database or with blastn to a 16S, NAD1, NAD4, NAD5, CYTB and ATP6 nucleotide reference database to the number of quality filtered reads. To estimate the X-fold enrichment, four samples from the 10 species MC were sequenced without any enrichment. The X-fold

enrichment was calculated using the ratio of the percentage of targeted reads from the capture library to the percentage of targeted reads from the enrichment-free library.

2.4.6 *Species detection and initial biomass recovery*

To compensate for sequencing effort as well as total biomass variation among samples, read count assigned to species and biomass (mg) were transformed in proportion of reads (i.e. ratio of the number of COI reads assigned to the species to the total number of COI reads assigned to the COI mock community reference database) and proportion of biomass (i.e. ratio of the biomass of the species to the total biomass of the MC), respectively.

Species detection. A species was considered present in a sample when represented by at least one read for the capture sequencing pipeline and one MOTU for the amplicon sequencing pipeline. We calculated the detection sensitivity (S) which measures the number of detection success to the total number of trials ($S = \frac{\text{total number of species detected}}{\text{number of samples} * \text{number of taxa}}$).

Biomass recovery. For the bulk DNA and etDNA of the 10 species MC, logistic models were built to investigate the relationship between read proportion and biomass proportion. For each DNA template and enrichment method, we compared sets of mixed effects models (i.e. no fixed effect and biomass as fixed effect with no random effect, random intercept, random intercept and slope by species or samples) using Aikake's Information Criterion (AIC, (Burnham & Anderson, 2002)). We also summarized to what extent each method was able to predict species biomass by comparing the observed read proportions to the expected read proportions where read proportion perfectly predict biomass proportion (i.e. a $y=x$ relationship). To this aim, we calculated the mean absolute error (MAE) for each method as follow: $\frac{1}{n} \sum |observed - expected|$ (Willmott & Matsuura, 2005). The lower the MAE is, the closer the observed read proportions are to the expected read proportions.

All statistical analyses were conducted with R (R Core team, 2018). All logistic models relating biomass proportion to read proportion were fitted using lme4 package (glmer functions, with a Logit link and a Binomial family, (Bates, Mächler, Bolker, & Walker, 2014)).

2.5 *Results*

2.5.1 *Sequencing results*

For bulk DNA, amplicon sequencing produced 60,337 to 365,331 raw reads and capture sequencing 165,089 to 857,040 raw reads per sample. For etDNA, amplicon sequencing produced 209,425 to 365,331

raw reads for Fwh1 primer pair of extraction I, 219,124 to 314,441 raw reads for Fwh1 primer pair of extraction II, 172,036 to 280,065 raw reads for BF2/BR2 primer pair of extraction II and capture sequencing 217,240 to 911,484 raw reads (Tableau 2.2). In both approaches, the capture sequencing effort was higher but a higher proportion of reads was discarded through the bioinformatic pipeline, particularly for etDNA (Tableau 2.2).

Tableau 2.2. Assessment of PCR and capture enrichment specificity on bulk and ethanol DNA. "% of targeted reads": for capture enrichment only, percentage of reads that align to the COI, 16S, NAD1, NAD4, NAD5, CYTB, and ATP6 reference databases, "% of COI reads": percentage of reads that align to the COI protein reference database, "% of COI reads assigned to a MC species": percentage of the number of reads that were successfully assigned to a species using the COI nucleotide reference databases to the COI assigned reads and "% of non-protostomian reads": percentage of reads that align to non-protostomian groups. Indicated values are mean number and standard deviation. For sample values, see Table S4 (PCR enrichment) and Table S5 (Capture enrichment). 10 sps: 10 species MC; 52 taxa:52 taxa MC; Ø means that the measure cannot be calculated.

bulk DNA	<i>Mock community</i>	<i>Raw reads</i>	<i>Quality filtered reads</i>	<i>% targeted reads</i>	<i>% COI reads</i>	<i>% of COI reads assigned to a MC species</i>	<i>% of non-protostomian reads (contaminants)</i>
<i>PCR enrichment</i>	10 sps	170,026 ± 40,272	114,701 ± 31,160	Ø	49.54 ± 4.83	96.01 ± 3.24	2.77 ± 1.30
	52 taxa	135,168 ± 40,871	91,935 ± 32,720	Ø	42.33 ± 4.61	96.55 ± 3.80	2.37 ± 1.36
<i>Capture enrichment</i>	10 sps	493,967 ± 288,240	243,483 ± 86,526	43.34 ± 24.52	25.16 ± 18.81	61.17 ± 5.55	10.73 ± 3.78
	52 taxa	553,825 ± 165,868	264,694 ± 75,864	80.21 ± 5.09	69.30 ± 6.51	63.27 ± 1.96	5.31 ± 1.70
<i>Enrichment-free</i>	10 sps	921,863 ± 68,162	574,443 ± 43,547	1.19 ± 0.31	0.023 ± 0.028	26.14 ± 9.21	31.66 ± 1.96
etDNA	<i>Mock community</i>	<i>Raw reads</i>	<i>Quality filtered reads</i>	<i>% targeted reads</i>	<i>% COI reads</i>	<i>% of COI reads assigned to a MC species</i>	<i>% of non-protostomian reads (contaminants)</i>
<i>PCR enrichment - Fwh1 extraction I</i>	10 sps	283,644 ± 52,944	233,856 ± 43,256	Ø	62.97 ± 4.47	46.36 ± 11.47	18.61 ± 12.78
<i>PCR enrichment - Fwh1 extraction II</i>	10 sps	205,229 ± 81,671	169,772 ± 69,282	Ø	66.34 ± 5.88	47.72 ± 14.65	10.13 ± 4.12
<i>PCR enrichment - BF2/BR2 extraction II</i>	10 sps	203,406 ± 43,493	123,148 ± 47,936	Ø	48.97 ± 9.81	60.95 ± 18.88	28.21 ± 19.29
<i>Capture Enrichment – extraction I</i>	10 sps	517,202 ± 234,942	309,637 ± 147,304	27.18 ± 8.39	19.09 ± 5.82	33.71 ± 10.96	56.81 ± 10.17

2.5.1 Capture efficacy

The percentage of targeted reads after capture enrichment (i.e. capture specificity) ranged from 16.24 to 86.66% (mean \pm se: 63.82% \pm 24.83) for bulk DNA and from 19.52 to 40.79% (mean \pm se: 27.18 \pm 8.39) for etDNA (Tableau 2.2, Table S5). For bulk DNA, the 52 taxa MC showed higher and more homogeneous capture specificity (mean \pm se: 80.21 \pm 5.09) in comparison with bulk DNA of the 10 species MC (mean 43.34 \pm 24.52). The average X-fold enrichment (i.e. capture efficiency) was 974 (range: 214-2470) meaning that on average 974-fold more targeted reads were sequenced with capture enrichment than without any enrichment.

The percentage of COI reads that were assigned to a MC species was very heterogeneous among DNA template and enrichment methods. Amplicons on bulk DNA gave the best results (mean: 96.45% COI assignment), followed by capture on bulk (62.33 %), amplicon on etDNA (46.36 %) and finally capture on etDNA (33.71 %). Reads that do not match to the reference MC COI database can have multiple origins including contamination from other organisms. The majority of reads were assigned to protostomians for all but one experiment – capture enrichment on etDNA – where half of the reads belonged to eubacteria and the other half to protostomians (Tableau 2.2, Figure S2).

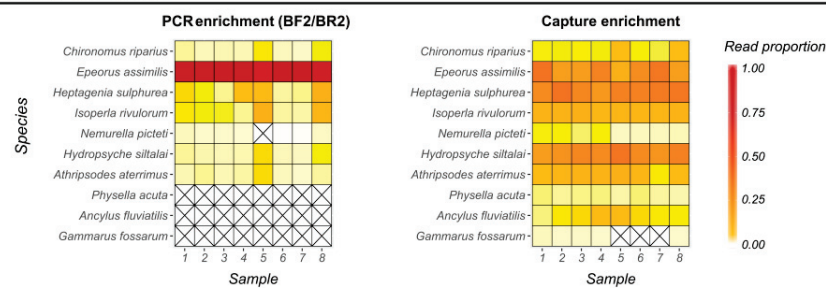
2.5.2 PCR and capture enrichment for bulk DNA

Species detection. Capture enrichment detected systematically more species ($S=0.96$) than PCR ($S=0.68$) among the 10 species MC (Figure 2.3). In PCR libraries, three species belonging to Gastropoda (*Ancylus fluviatilis* and *Physella acuta*) and Amphipoda (*Gammarus fossarum*) were never detected. With capture enrichment, every species was detected in 5 out of 8 samples (Figure 2.3). In the 52 taxa MC, we found the same pattern, with capture enrichment ($S=0.96$) detecting more species than PCR enrichment ($S=0.93$) (Figure 2.3). Each enrichment method eventually failed to detect a small but different set of taxa.

Biomass recovery. Whatever the enrichment method, following the AIC, the best mixed model to predict initial species biomass using species read proportion was the model combining biomass as a fixed effect and species as a random effect on the intercept and slope (Tableau 2.3). This suggests that the mean read proportion between species is different independently of their biomass and that the relationship between biomass proportion and read proportion also differs among species (Figure 2.4). Using this type of mixed models, a significant relationship between biomass and read proportion was found for both enrichment methods (PCR: p-value=0.024; capture: p-value=0.001, Figure 2.4). The mean absolute errors between observed and expected read proportion if biomass proportions were to be perfectly translated into

read proportions (i.e. $y=x$ relationship) were higher for PCR (MAE=0.11) than for capture (MAE=0.056). Under a scenario where there is no relationship between biomass and read proportions and where each species contribute to the same read proportion independently of its biomass (i.e. 1/10 read proportion), the MAE would be of 0.07, again highlighting that the PCR enrichment step wiped out most of the biomass signal. While absolute biomass variations may be lost, we also tested if biomass ranks could be recovered using Spearman's rank correlation coefficient. Again PCR enrichment performed poorly compared to capture enrichment (average Spearman rho 0.53 and 0.67 for PCR and capture enrichment, respectively).

10 species mock communities



52 taxa mock communities

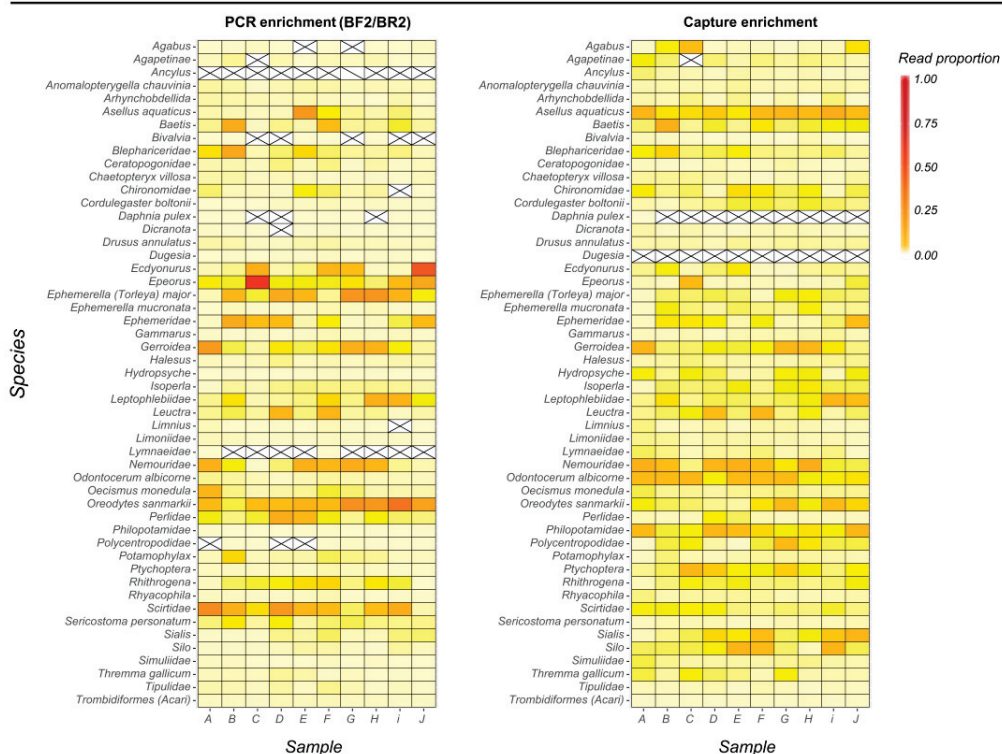


Figure 2.3 : Taxa recovery performance assessed using read proportion for PCR (BF2/BR2 primer pair, left) and capture (right) enrichment and for bulk DNA of the 10 species mock communities (MCs) (top) and bulk DNA of the 52 taxa MC (bottom). Read proportions are shown for each taxa (in rows) and mock community (in column). A crossed out cell indicates no assigned read. For detailed read counts, see Tables S6 (10 species MCs) and S7 (52 taxa MCs).

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Tableau 2.3. Testing the link between read proportion (dependant variable) and initial biomass proportion. Only relative ΔAIC to the best model are shown (i.e. $AIC_{\text{model}} - AIC_{\text{best model}}$). Models with and without fixed effect (biomass proportion) and two random effects (species and sample) were built. For each random effect, three models were built: no random effect, random intercept and random intercept and slope. Best models correspond to the models with relative ΔAIC of 0. \emptyset means that the model could not be tested.

bulk DNA	Fixed effect	Random effect	Random effect		
			None	Intercept	Intercept + slope
PCR enrichment	1	Species	2,012,800	31,216	\emptyset
		Sample		2,012,802	
	biomass proportion	Species	537,148	6,674	0
		Sample		496,237	382,656
Capture enrichment	1	Species	320,984	9,223	\emptyset
		Sample		320,986	
	biomass proportion	Species	185,777	5,154	0
		Sample		185,322	182,666
etDNA	Fixed effect	Random effect	None	Intercept	Intercept + slope
PCR enrichment	1	Species	2,602,013	121,099	\emptyset
		Sample		2,602,015	
	biomass proportion	Species	1,897,689	97,590	0
		Sample		1,894,974	1,748,461
Capture enrichment	1	Species	195,670	43,136	\emptyset
		Sample		195,672	
	biomass proportion	Species	166,761	26,717	0
		Sample		166,760	102,699

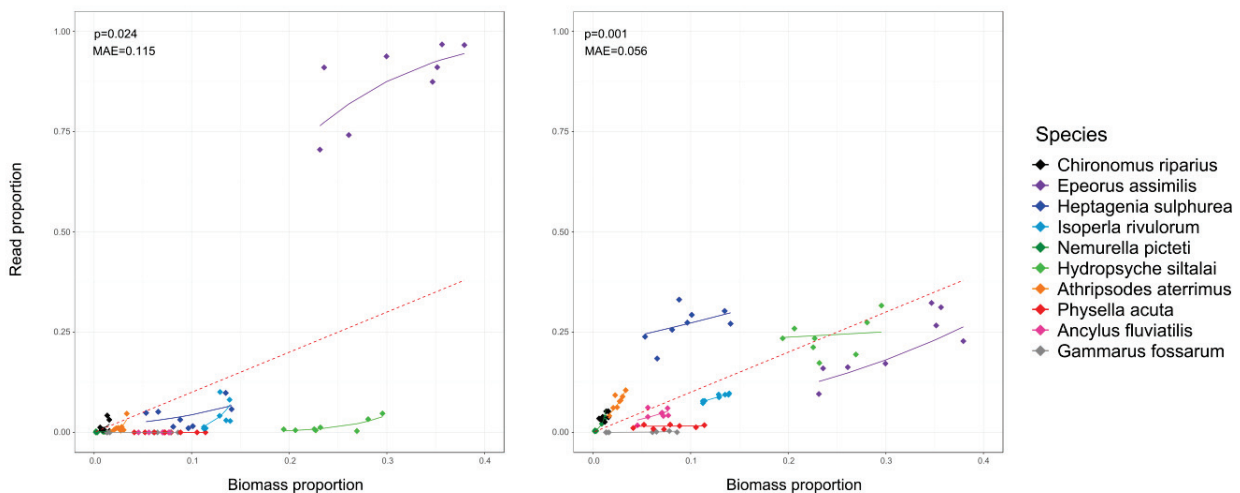


Figure 2.4 : Relationship between biomass proportion and read proportion for PCR (left) and capture (right) enrichment method. Logistic mixed model predictions with random intercept and slope by species are shown (solid line). The dashed red line corresponds to the expected linear relationship where one unit of read proportion equals to one unit of biomass proportion.

2.5.3 etDNA performance

As for bulk DNA, capture enrichment detected more species ($S=0.97$) than PCR ($S=0.85$) (Figure 2.5, extraction I) on the etDNA template. Concerning PCR, the primer pair Fwh1 performs better ($S=0.86$) than the primer pair BF2/BR2 ($S=0.7$) on etDNA (Figure 2.5, extraction II). This was expected as the long BF2/BR2 fragment may be difficult to amplify on degraded DNA. Nevertheless the two primers pairs also failed to detect different species probably indicating a primer bias rather than a degradation problem. To confirm this, we tested if the sample composition was mostly driven by DNA template and enrichment method using a Principal Coordinate Analysis (PCoA). Samples clustered by enrichment method and PCR primers but not by DNA template (Figure 2.6). Therefore, in this experiment, the DNA template had little to no impact compared to the enrichment method and primers. Concerning initial biomass recovery, the same models as for bulk DNA were selected for etDNA using the AIC (Tableau 2.3). However, contrary to bulk DNA, for both enrichment methods, no significant relationship between biomass proportion and read proportion was found (PCR: $p\text{-value}=0.37$; capture: $p\text{-value}=0.58$). The difficulty to infer biomass from read proportion using etDNA was also supported by high MAE values (0.102 and 0.104 for PCR and capture enrichment, respectively) and low rank correlations (average Spearman $\rho=0.40$ and 0.28 for PCR and capture enrichment, respectively).

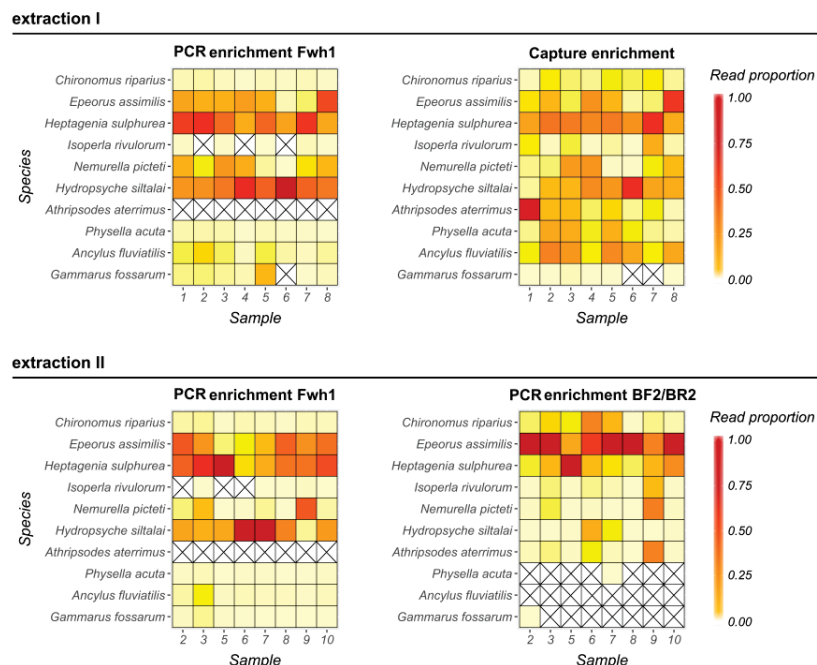


Figure 2.5 : Taxa recovery performance assessed using read proportion for etDNA of the 10 species mock communities. Differences between PCR (top left) and capture (top right) enrichment were assessed using a first ethanol extraction (extraction I). Differences between Fwh1 (bottom left) and BF2/BR2 (bottom right) primer pairs were assessed using a second ethanol extraction (extraction II). Sequence abundances are shown for each taxa (in rows) and mock community (in column). A crossed out cell indicates no assigned read. For detailed read counts, see

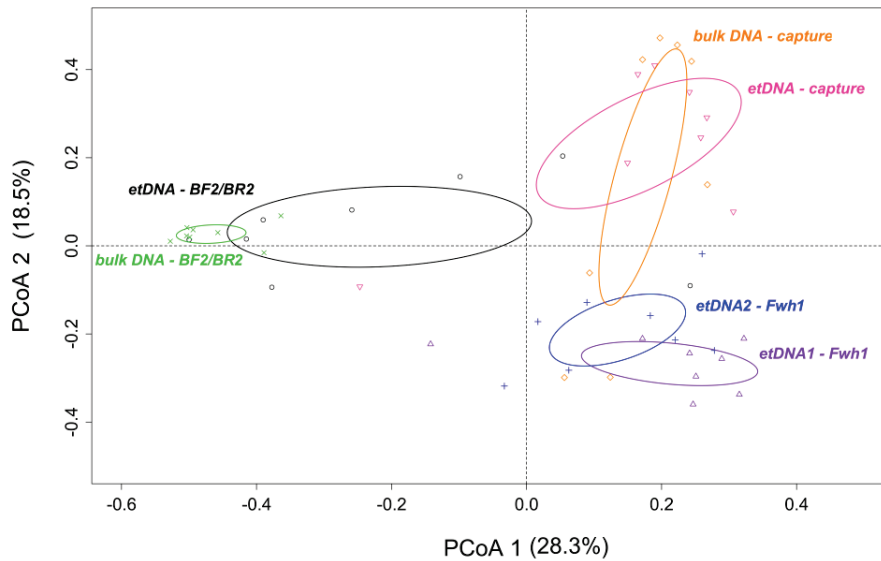


Figure 2.6 : Principal Coordinate Analysis based on dissimilarity (Jaccard distances) in read proportions of each pair of the 10 species mock community samples. Samples are classified following their DNA template (etDNA and bulk DNA) and enrichment methods (BF2/BR2 amplicon, Fwh1 amplicon and capture). etDNA1 and etDNA2 correspond to the first and second extraction of DNA from ethanol, respectively. Percentage of variance explained by each axis is shown in bracket. Ellipses are drawn with a confidence limit of 0.95. Samples are grouped by enrichment method rather than by DNA template.

2.6 Discussion

2.6.1 Capture versus PCR enrichment for metabarcoding

Capture enrichment consistently led to higher species detection rates compared to PCR enrichment whatever the DNA template or sample taxonomic complexity. Using capture in the 10 species MC, only one species was missed in one third of the samples where it was at low abundance (dry biomass < 1 mg). In the 52 taxa MC, capture enrichment also globally performed better than PCR enrichment in terms of species detection (20 against 35 failed detections, respectively). Both methods also missed different sets of taxa. In PCR enrichment, the commonly missed taxa systematically presented 1 or 2 mismatches with the primer pair (Figure S1) reinforcing the view that primer mismatches are the main cause for species non-detection (Elbrecht et al., 2017). Regarding capture enrichment, 8 taxa had no bait designed for them (i.e. Ceratopogonidae, Blephariceridae, Dicranota, Simuliidae, Tipulidae, Trombidiformes, Dugesia and *Daphnia pulex*) but only two taxa (*Daphnia pulex* and *Dugesia*) were almost systematically missed. These two taxa have COI haplotypes that are at a divergence of at least 15% to any bait whereas the 6 detected taxa have haplotypes that are closer to baits originally designed for other taxa (Figure S1). The divergence to the oligonucleotide has already been described as an important parameter for species detection (Liu et al, 2015; Portik et al, 2016; Vallender et al, 2011; van der Valk et al, 2017). In particular, Liu et al (2015)

using a pool of 49 taxa found a lower enrichment efficiency when the divergence to the bait was higher than 20%. Portik et al (2016), this time working on a multi-loci capture data set, found a negative linear relationship between the probability of sequencing a locus and the divergence to the bait. This translated into a sharp decline from 60% to 20% chance of sequencing when the divergence increased from 10% to 20%. Altogether, for both enrichment methods, the divergence to the oligonucleotide (i.e. primer or bait) appears to be a determinant factor for species detection. Nevertheless, capture enrichment is much more robust to this divergence issue as it combines three characteristics that PCR enrichment misses. First, capture enrichment is less sensitive to mismatches (18 mismatches or 15% divergence) than PCR (2 mismatches or 5% divergence). Second, because thousands of different baits are designed for a given 120-base region, the probability to encounter mismatches is reduced with capture. Finally, several 120-base regions (5 in this study leading to a total cover of 600 bases) are targeted, increasing the probability to design at least one functional bait for a given species even if this species was not in the database used to design the baits. Alternatively species detection can be improved in traditional PCR-based metabarcoding by targeting multiple DNA fragment (e.g. (Drummond et al., 2015; Zhang, Chain, Abbott, & Cristescu, 2018)). This is also likely to be true but easier to implement with capture enrichment where baits for different loci can be used simultaneously (e.g. Liu et al, 2016). The ability to detect more robustly known and unknown biodiversity using a single enrichment step is a pivotal step towards a better understanding of ecosystem function and structure.

Capture enrichment also yields read proportions that are better predictors of species relative biomasses than PCR enrichment for bulk DNA. Both enrichment methods presented a positive relationship between relative biomass proportion and read proportion but this relationship was closer to a linear $y=x$ relationship for capture enrichment. That said, a random species effect still remained in the capture enrichment model. Its origin could be attributed to a bait bias similar to the primer bias encountered in PCR enrichment but it can also be explained by other factors. Indeed, using a mitochondrial marker to reconstruct an entire dry biomass community is eventually doomed to fail. First, the amount of mitochondrial DNA belonging to a species in a sample will vary as a function of the average number of mitochondrial genomes per cell which is likely to vary extensively among species, life stages and tissues (Rooney et al, 2015). Wilcox et al. (2018) found that total initial species DNA abundances could be recovered using capture enrichment if a correction was applied to normalize for variation in the number of mitochondrial genome per cell. Second, dry biomass proportions are a very rough estimator of cell numbers. Thus, variability in cell number per biomass and in mitochondria per cell combines to blur the biomass/read proportion relationship. A similar mock community experiment but with precise knowledge of the number of mitochondrial genome per species in a sample is needed to better test for the existence of bait biases. One can argue that even if capture enrichment yields good estimates of each species

mitochondrial genome proportion in a sample, this estimate will still be a poor biomass or abundance proxy. Interestingly, here we found that read proportions from gene capture enrichment already provides biomass rank estimates which in many ecological contexts, for instance in bioassessment (Elbrecht et al., 2017), would already be informative and may be sufficient.

2.6.2 Ethanol DNA as a fast alternative to bulk DNA for species detection

In agreement with Martins et al. (2019), our results demonstrate the potential of etDNA to replace bulk DNA for macrozoobenthos samples if only species occurrence data is required. Species detection was not determined by DNA template but by the enrichment method and the primer pair used. Hence, species detection with etDNA was equivalent to species detection with bulk DNA. Nonetheless, no relationship between read proportion and biomass proportion was highlighted for etDNA. For biomass recovery to work with etDNA, each species should release a quantity of DNA in the ethanol proportionally to its biomass. However, the release of DNA in the ethanol might differ among development stages or species. When put alive in ethanol, some individuals regurgitate and release a lot of their DNA (Anderson et al., 2013). For example, in the class Gastropoda, one sub-class has an operculum which isolates them from the environment. When put in ethanol, these gastropods close their shell and the quantity of DNA released will be lower than for other gastropods without operculum. Also, after ethanol fixation, depending on the presence of a shell or a thick exoskeleton, the amount of released DNA during the soaking period may also be extremely variable across taxa. In conclusion, while DNA from ethanol offers a fast and non-destructive way to identify the species in a sample, differences in the way species release their DNA in ethanol seem to prevent the use of ethanol for quantitative estimates. Compared to bulk DNA, DNA extracted from preservative ethanol is more fragmented and much less concentrated. In addition to environmental contaminants such as the bacterial contaminants observed in this study, this template may also be more susceptible to reagent contamination and cross-contamination. The adoption of this alternative template may therefore need stricter sampling, sample handling and DNA extraction protocols than the one regularly used in community ecology laboratories.

2.6.3 Optimising capture enrichment for metabarcoding

Despite capture enrichment already delivers better results than PCR enrichment in terms of species detection or biomass prediction, the efficiency and specificity of this emerging approach can still be optimized for metabarcoding purposes. Although capture enrichment was efficient to enrich the COI sequences of the macrozoobenthos present in our samples with an enrichment of at least 214-fold compared to non-enriched libraries, the percentage of targeted reads was highly heterogeneous and low

for some samples. The specificity of the capture enrichment was particularly low when the diversity of the sample was low (43% of targeted reads for the 10 species MC compared to 80% for the 52 taxa MC). How the taxonomic diversity of a sample influences the efficiency of capture enrichment remains unexplained and warrants further experiments where the impact of a gradually increasing sample diversity is tested while controlling for other factors such as the extraction protocol or the overall phylogenetic diversity.

The specificity of capture enrichment estimated by the number of COI reads that matched a taxa included in the MC was systematically lower than with PCR enrichment (about 40% of the COI reads did not match a MC taxa). When looking at the coarse taxonomic composition of the COI reads sequenced after capture enrichment, we found that for bulk DNA, most of the reads (> 95%) were assigned to protostomians ruling out the possibility that bacterial or plant COI contaminations alone are reducing capture specificity. DNA from other protostomians that interact with the taxa included in the MC (e.g. preys or parasites) may contaminate the samples and reduce capture specificity. Nevertheless, we would expect the same to happen with PCR enrichment which is not the case. Alternatively, Li, Schroeder, Ko, and Stoneking (2012) found that the robustness of capture to mismatches is also a drawback when targeting mitochondrial genes: in addition to capture the targeted mitochondrial gene, it also captures nuclear copies of this gene (NUMTs). The number of NUMT loci per genome is highly variable across taxa but can easily reach hundreds to thousands of loci (Hazkani-Covo, Zeller, & Martin, 2010). In some species there might therefore be more NUMTs in a cell than there are mitochondrial genomes. Given that baits can capture divergent loci and that there is many different baits, NUMTs may represent a significant burden in mitochondrial capture enrichment. Finally, capture enrichment also increase the sequencing coverage of the flanking regions of the targeted sequences (Portik et al, 2016). With the pipeline used in this study, the reads containing more than 41% of a flanking region will not be classified as a read coming from a targeted region. The lower specificity of capture enrichment observed in our samples, but not observed with PCR enrichment, might therefore be explained by a combined effect of all these factors.

For capture enrichment on etDNA, we estimated that almost half of the reads belonged to eubacteria and the other half to protostomians. This is likely to be linked to the regurgitation of gastric microbiota and the release of epidermic microbiota by some species. By this process, bacterial DNA may actually dominate the ethanol DNA. In this experiment, a lower hybridization temperature was used for etDNA (55°C) than for bulk DNA (65°C) in the hope of capturing more fragmented DNA. Retrospectively, this lower temperature is likely to have decreased the specificity of the capture enrichment and increased the representation of bacterial DNA. Indeed, the specificity of capture enrichment is higher for degraded DNA and samples with a lot of contaminants when a higher hybridization temperature is used (Paijmans

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

et al., 2016). Noteworthy, 95% of the bacterial reads did not share any similarity with the mitochondrial COI, again reinforcing the idea that the large amount of bacterial reads in etDNA samples is probably the result of the combination of non-stringent hybridization conditions and bacterial DNA dominance in etDNA.

As found in this study and others (e.g. Liu et al., 2016; Mariac et al., 2018; Portik et al., 2015), the divergence between the targeted DNA and the baits is a critical factor in capture enrichment. As such, to recover all taxa in a community, the bait design must be optimized to reduce this divergence. The priority is thus to design baits using a reference database that is exhaustive. While reference DNA barcodes are available for most species for some groups, this is far from being the case for many groups (e.g. (Sonet et al., 2013)). Knowing the divergence threshold of 15% (this study) over which species are not captured and detected, alternative strategies could be deployed to barcode key missing taxa in the reference databases. Indeed, a very limited set of baits is sufficient to represent a whole family or a genus if the intra-group genetic diversity is lower than the divergence threshold (Mariac et al., 2018). This diversity is heterogeneous between taxa group, for example the maximum intra-family COI divergence in Gammaridea (Amphipoda) is 30% but is only 17% in Chloroperlidae (Plecoptera). Therefore, prior investigations are needed to establish the diversity of each group and the number of taxa that have to be described before a robust set of baits can be designed. Another alternative would be to design baits from a set of representative COI sequences and mutate them according to a given divergence to obtain a set of baits that can hybridize to most sequences. Such baits would permit to capture non-barcoded or even new species but could also reduce capture specificity by capturing untargeted sequences such as bacterial genes or NUMTs.

In addition to the bait design strategy, other capture enrichment parameters may need further optimization for a metabarcoding application. The hybridization temperature is known to be a major factor controlling capture efficacy (Li et al., 2013). However, Paijmans et al. (2016) compared different hybridization temperatures but unexpectedly did not find any interaction between hybridization temperature and distance to the bait. But Paijmans et al. (2016) used a pool of 21 feline species, with all the probes designed using a single reference species. Therefore their results probably need to be tested on a data-set that is more similar to a metabarcoding data-set, with more divergent taxa and probes designed from a large database, before the interaction between hybridization temperature and distance to the bait can be completely dismissed. The number of pooled libraries in a single hybridization reaction is another potentially important parameter. While pooling many samples in a single capture reaction significantly cuts the hybridization budget, Portik et al. (2016) showed that it significantly reduces the complexity of the sequencing libraries. In a metabarcoding context, this may result in a lower species discovery rate in

disfavor of the rare species. In the same vein, the number of capture rounds is a substantial parameter increasing capture efficacy but also the hybridization budget (Li et al., 2013; Mariac et al., 2018; Templeton et al., 2013; van der Valk et al., 2017). With a second round of capture, Mariac et al. (2018) increased capture specificity from 0.57% to 70.5% and capture efficiency by 40 times but with the penalty of doubling hybridization costs. After sequencing capture enriched metabarcoding libraries, another challenge is the bioinformatic treatment of the reads. Indeed, processing capture data targeting one or few DNA barcodes for many species is considerably different than assembling multiple targeted genes for a single organism. To our knowledge, while there is several assembly pipelines that were developed for the latter case (co-assembly of different loci from a single species, (e.g. (Allio et al., 2019))), there is no available bioinformatic pipelines for the former. Assembling metabarcoding capture reads is challenging as it requires the independent assembly of orthologous sequences belonging to different species which may be more or less divergent and more or less polymorphic. An alternative to assembly is to directly map the capture reads on a reference database as done in this study. While functional, this solution is sub-optimal because capture reads do not start and end at the same position as in amplicon sequencing. This complicates the read clustering and taxonomic assignment, with many capture reads probably lost as they only partially map to the reference library. While the sequencing of the flanking regions of targeted region would probably be highly valuable for taxonomic assignment, as long as there is no dedicated bioinformatic solution, it paradoxically complicates the processing and analysis of capture reads. Finally, a comparison of these mock community results with field samples in which the diversity is much higher and where the template is also much more complex is warranted to confirm the potential of capture enrichment in community ecology.

2.7 Conclusion

Capture enrichment is a promising alternative to PCR enrichment for metabarcoding. Its main advantage is to provide better species detection thanks to its robustness to mismatches. At this point, while performing much better than PCR enrichment, absolute biomass reconstruction is not applicable without species mitogenome copy number correction and is therefore not tractable for community studies where hundreds of species can be encountered. Yet, biomass rank ordination appears to be robust and could be used for ecological purpose. Albeit more bacterial contamination, the use of etDNA coupled to capture enrichment presents interesting compromises. It saves a lot of time by ignoring the organism picking step and permits to save organisms for further analyses. However, etDNA should be used for studies where quantitative information is not required. Finally, these promising results call for additional testing of the bait design strategy and hybridization reaction parameters, as well as testing capture enrichment on complex field samples.

2.8 Acknowledgements

We thank Maxence Forcellini, Bertrand Launay and Guillaume Le Goff for morphological macrozoobenthos identification and fieldwork assistance. We thank Clémentine François and Florian Leese for advices and assistance in bioinformatic analysis and bait design. We thank the three anonymous reviewers for their helpful and relevant comments on the manuscript. This study was supported by the French Biodiversity Agency (AFB) “Grant 26: headwater biodiversity dynamics: a molecular perspective”, and the CNRS Mission pour les Initiatives Transverses et Interdisciplinaires (project XLIFE CAPTAS).

2.9 Data accessibility

Additional reference sequences developed for bait design are available on Genbank (Accession Numbers MK584300:MK584515). COI and 16S sequences barcoded for the 10 species MC assignment are available on GenBank (Accession Numbers MK584516:MK584524 for COI and MK584525:MK584534 for 16S). The bait set designed for this study is available on Zenodo (Zenodo (<https://doi.org/10.5281/zenodo.2581410>)). FASTQ files from PCR and capture enrichment are available on NCBI (Accession Number SRR8745499:SRR8745570). Scripts for bioinformatic analysis are available on GitHub (<https://github.com/mailysgauthier/bioinf-cap-PCR>).

2.10 References

- Abdelfattah, A., Malacrinò, A., Wisniewski, M., Cacciola, S. O., & Schena, L. (2018). Metabarcoding: A powerful tool to investigate microbial communities and shape future plant protection strategies. *Biological Control*, 120, 1-10. doi:10.1016/j.biocontrol.2017.07.009
- Aird, D., Ross, M. G., Chen, W. S., Danielsson, M., Fennell, T., Russ, C., . . . Gnirke, A. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biology*, 12, R18. doi:10.1186/gb-2011-12-2-r18
- Allio, R., Schomaker-Bastos, A., Romiguier, J., Prosdocimi, F., Nabholz, B., & Delsuc, F. (2019). doi:10.1101/685412
- Anderson, J. T., Zilli, F. L., Montalto, L., Marchese, M. R., McKinney, M., & Park, Y.-L. (2013). Sampling and Processing Aquatic and Terrestrial Invertebrates in Wetlands. In *Wetland Techniques* (Vol. 1, pp. 143-195).

Arif, I. A., Khan, H. A., Al Sadoon, M., & Shobrak, M. (2011). Limited efficiency of universal mini-barcode primers for DNA amplification from desert reptiles, birds and mammals. *Genetics and Molecular Research*, 10, 3559-3564. doi:10.4238/2011.October.31.3

Aukema, B. & Rieger C. (1995) Catalogue of the Heteroptera of the Palaearctic Region. Vol. 1 : Enicocephalomorpha, Dipsocoromorpha, Nepomorpha, Gerromorpha and Leptopodomorpha. Netherlands Entomological Society, Wageningen, The Netherlands. 222 p.

Aukema, B., Rieger, C. & Rabitsch, W. (2013) Catalogue of the Heteroptera of the Palaearctic Region. Vol. 6 : Supplement. Netherlands Entomological Society, Wageningen, The Netherlands. 629 p.

Avila-Arcos, M. C., Cappellini, E., Romero-Navarro, J. A., Wales, N., Moreno-Mayar, J. V., Rasmussen, M., . . . Gilbert, M. T. (2011). Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. *Sci Rep*, 1, 74. doi:10.1038/srep00074

Baird, D. J., & Hajibabaei, M. (2012). Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, 21, 2039-2044. doi:10.1111/j.1365-294X.2012.05519.x

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. doi:10.18637/jss.v067.i01

Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., & Kausrud, H. (2010). ITS as an environmental DNA barcode for fungi: An in silico approach reveals potential PCR biases. *BMC Microbiology*, 10, 1-9. doi:10.1186/1471-2180-10-189

Bista, I., Walsh, K., Zhou, X., Seymour, M., Christmas, M., Bradley, D., . . . Creer, S. (2018). Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. *Molecular Ecology Resources*, 18, 1020-1034. doi:10.1111/1755-0998.12888

Bortolus, A. (2008). Error Cascades in the Biological Sciences: The Unwanted Consequences of Using Bad Taxonomy in Ecology. *AMBIO: A Journal of the Human Environment*, 37, 114-118. doi:10.1579/0044-7447(2008)37[114:ecitbs]2.0.co;2

Bringloe, T. T., Cottenie, K., Martin, G. K., & Adamowicz, S. J. (2016). The importance of taxonomic resolution for additive beta diversity as revealed through DNA barcoding. *Genome*, 1140, 1130-1140.

Buchfink, B., Xie, C., & Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12, 59-60. doi:10.1038/nmeth.3176

Burnham, K. P., & Anderson, D. R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed). *Ecological Modelling*, 172, 488. doi:10.1016/j.ecolmodel.2003.11.004

Calvignac-Spencer, S., Merkel, K., Kutzner, N., Kuhl, H., Boesch, C., Kappeler, P. M., . . . Leendertz, F. H. (2013). Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Mol Ecol*, 22(4), 915-924. doi:10.1111/mec.12183

Camacho, C., Madden, T. L., Ma, N., Coulouris, G., Avagyan, V., Bealer, K., & Papadopoulos, J. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, 10, 421. doi:10.1186/1471-2105-10-421

Cha, R. S., & Thilly, W. G. (1993). Specificity, efficiency, and fidelity of PCR. *Genome Res.*, 3, 18-29.

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, K., . . . Bik, H. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution*, 56, 68-74. doi:10.1111/2041-210X.12574

Deiner, K., Bik, H. M., Machler, E., Seymour, M., Lacoursiere-Roussel, A., Altermatt, F., . . . Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Mol Ecol*, 26(21), 5872-5895. doi:10.1111/mec.14350

de Moraes-Barros, N., Silva, J. A. B., & Morgante, J. S. (2011). Morphology, molecular phylogeny, and taxonomic inconsistencies in the study of *Bradypus* sloths (Pilosa: Bradypodidae) *Journal of Mammalogy*, 92, 86-100. doi:10.1644/10-mamm-a-086.1

D'Hondt, J. L., & BEN A., R. (2009). Catalogue et clés tabulaires de détermination des Hirudinées d'eau douce de la faune Française. *Bulletin de la Société zoologique de France*, 134, 263-298.

Dowle, E. J., Pochon, X., C. Banks, J., Shearer, K., & Wood, S. A. (2016). Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: a case study using freshwater macroinvertebrates. *Molecular Ecology Resources*, 16, 1240-1254. doi:10.1111/1755-0998.12488

Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C. M., . . . Nelson, N. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. *GigaScience*, 4. doi:10.1186/s13742-015-0086-1

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE*, 10, 1-16. doi:10.1371/journal.pone.0130324

Elbrecht, V., & Leese, F. (2016). Development and validation of DNA metabarcoding COI primers for aquatic invertebrates using the R package "PrimerMiner". *PeerJ*, 1-23. doi:10.7287/peerj.preprints.2044v1

Elbrecht, V., & Leese, F. (2017). Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental Science*, 5, 1-11. doi:10.3389/fenvs.2017.00011

Elbrecht, V., Taberlet, P., Dejean, T., Valentini, A., Usseglio-Polatera, P., Beisel, J.-N., . . . Leese, F. (2016). Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ*, 4, e1966. doi:10.7717/peerj.1966

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, 8, 1265-1275. doi:10.1111/2041-210X.12789

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5), 294-299.

Forshaw, J. M. (2010). *Parrots of the World* (P. U. Press Ed. Vol. 70).

Frank, J. A., Wilson, B. A., Weisbaum, J. S., Sharma, S., Reich, C. I., & Olsen, G. J. (2008). Critical Evaluation of Two Primers Commonly Used for Amplification of Bacterial 16S rRNA Genes. *Applied and Environmental Microbiology*, 74, 2461-2470. doi:10.1128/aem.02272-07

Forster, D. W., Bull, J. K., Lenz, D., Autenrieth, M., Paijmans, J. L. A., Kraus, R. H. S., . . . Fickel, J. (2018). Targeted resequencing of coding DNA sequences for SNP discovery in nonmodel species. *Mol Ecol Resour*, 18(6), 1356-1373. doi:10.1111/1755-0998.12924

Gargominy, O., Prié, V., Bichain, J. M., Cucherat, X., & Fontaine, B. (2011). Liste de référence annotée des mollusques continentaux de France. *MalaCo*, 7, 307-382.

Gibson, J. F., Shokralla, S., Curry, C., Baird, D. J., Monk, W. A., King, I., & Hajibabaei, M. (2015). Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. *PLoS ONE*, 10, e0138432. doi:10.1371/journal.pone.0138432

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Gómez-Rodríguez, C., Crampton-Platt, A., Timmermans, M. J. T. N., Baselga, A., & Vogler, A. P. (2015). Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. *Methods in Ecology and Evolution*, 6, 883-894. doi:10.1111/2041-210X.12376

Grand, D., & Boudot, J. P. (2007). *Les libellules de France, Belgique et Luxembourg*. Biotope

Hajibabaei, M., Baird, D. J., Fahner, N. A., Beiko, R., & Golding, G. B. (2016). A new way to contemplate Darwin's tangled bank: how DNA barcodes are reconnecting biodiversity science and biomonitoring. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 371, 20150330. doi:10.1098/rstb.2015.0330

Hajibabaei, M., Spall, J. L., Shokralla, S., & van Konynenburg, S. (2012). Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. *BMC Ecology*, 12, 28. doi:10.1186/1472-6785-12-28

Harper, L. R., Buxton, A. S., Rees, H. C., Bruce, K., Brys, R., Halfmaerten, D., . . . Hänfling, B. (2018). Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. *Hydrobiologia*, 826(1), 25-41. doi:10.1007/s10750-018-3750-5

Hazkani-Covo, E., Zeller, R. M., & Martin, W. (2010). Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet*, 6(2), e1000834. doi:10.1371/journal.pgen.1000834

Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. (2006). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270, 96-99. doi:10.1098/rsbl.2003.0025

Henry, J. P., & Magniez, G. (1983). *Introduction pratique à la systématique des organismes des eaux continentales françaises-4. Crustacés Isopodes (principalement Asellotes)*. Publications de la Société Linnéenne de Lyon. 52., 319-357.

Hodges, E., Xuan, Z., Balija, V., Kramer, M., Molla, M. N., Smith, S. W., . . . McCombie, W. R. (2007). Genome-wide in situ exon capture for selective resequencing. *Nature Genetics*, 39, 1522. doi:10.1038/ng.2007.42

Horn, S. (2012). Target enrichment via DNA hybridization capture. *Methods Mol Biol*, 840, 177-188. doi:10.1007/978-1-61779-516-9_21

Hubert, N., Delrieu-Trottin, E., Irisson, J. O., Meyer, C., & Planes, S. (2010). Identifying coral reef fish larvae through DNA barcoding: A test case with the families Acanthuridae and Holocentridae. *Molecular Phylogenetics and Evolution*, 55, 1195-1203. doi:10.1016/j.ympev.2010.02.023

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., . . . Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, 16, 1245-1257. doi:10.1111/ele.12162

Jones, M. R., & Good, J. M. (2016). Targeted capture in evolutionary and ecological genomics. *Mol Ecol*, 25(1), 185-202. doi:10.1111/mec.13304

Jusino, M. A., Banik, M. T., Palmer, J. M., Wray, A. K., Xiao, L., Pelton, E., . . . Lindner, D. L. (2019). An improved method for utilizing high-throughput amplicon sequencing to determine the diets of insectivorous animals. *Molecular Ecology Resources*, 19, 176-190. doi:10.1111/1755-0998.12951

Kersey, P. J., Stein, J., Zadissia, A., Yates, A., Paulini, M., Urban, M., . . . Bolser, D. M. (2017). Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species. *Nucleic Acids Research*, 46, D802-D808. doi:10.1093/nar/gkx1011

Leese, F., Altermatt, F., Bouchez, A., Ekrem, T., Hering, D., Meissner, K., . . . Zimmermann, J. (2016). DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. *Research Ideas and Outcomes*, 2, e11321. doi:10.3897/rio.2.e11321

Leese, F., Bouchez, A., Abarenkov, K., Altermatt, F., Borja, Á., Bruce, K., . . . Weigand, A. M. (2018). Why We Need Sustainable Networks Bridging Countries, Disciplines, Cultures and Generations for Aquatic Biomonitoring 2.0: A Perspective Derived From the DNAqua-Net COST Action. *Advances in Ecological Research*, 58, 63-99. doi:10.1016/bs.aecr.2018.01.001

Leray, M., & Knowlton, N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ*, 5, e3006. doi:10.7717/peerj.3006

Leys, M., Keller, I., Räsänen, K., Gattolliat, J. L., & Robinson, C. T. (2016). Distribution and population genetic variation of cryptic species of the Alpine mayfly *Baetis alpinus* (Ephemeroptera: Baetidae) in the Central Alps. *BMC Evolutionary Biology*, 16, 1-15. doi:10.1186/s12862-016-0643-y

Li, M., Schroeder, R., Ko, A., & Stoneking, M. (2012). Fidelity of capture-enrichment for mtDNA genome sequencing: Influence of NUMTs. *Nucleic Acids Research*, 40. doi:10.1093/nar/gks499

Li, C., Hofreiter, M., Straube, N., Corrigan, S., & Naylor, G. J. (2013). Capturing protein-coding genes across highly divergent species. *Biotechniques*, 54(6), 321-326. doi:10.2144/000114039

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Linard, B., Crampton-Platt, A., Gillett, C. P. D. T., Vogler, A. P., & Timmermans, M. J. T. N. (2015). Metagenome Skimming of Insect Specimen Pools: Potential for Comparative Genomics. *Genome Biology and Evolution*, 7, 1474-1489. doi:10.1093/gbe/evv086

Liu, S., Wang, X., Xie, L., Tan, M., Li, Z., Su, X., . . . Zhou, X. (2016). Mitochondrial capture enriches mito-DNA 100 fold, enabling PCR-free mitogenomics biodiversity analysis. *Molecular Ecology Resources*, 16, 470-479. doi:10.1111/1755-0998.12472

Macher, J. N., Zizka, V. M. A., Weigand, A. M., & Leese, F. (2017). A simple centrifugation protocol for metagenomic studies increases mitochondrial DNA yield by two orders of magnitude. *Methods in Ecology and Evolution*, 9(4), 1070-1074. doi:10.1111/2041-210X.12937

Maggia, M. E., Vigouroux, Y., Renno, J. F., Duponchelle, F., Desmarais, E., & Nunez, J. (2017). DNA Metabarcoding of Amazonian Ichthyoplankton Swarms. 1-14. doi:10.1371/journal.pone.0170009

Mariac, C., Vigouroux, Y., Duponchelle, F., García-Dávila, C., Nunez, J., Desmarais, E., & Renno, J. F. (2018). Metabarcoding by capture using a single COI probe (MCSP) to identify and quantify fish species in ichthyoplankton swarms. *PLoS ONE*, 13, 1-15. doi:10.1371/journal.pone.0202976

Martin, G. K., Adamowicz, S. J., & Cottenie, K. (2016). Taxonomic resolution based on DNA barcoding affects environmental signal in metacommunity structure. *Freshwater Science*, 35, 701-711. doi:10.1086/686260.

Martin, M. (2014). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*, 17, 10. doi:10.14806/ej.17.1.200

Martins, F. M. S., Galhardo, M., Filipe, A. F., Teixeira, A., Pinheiro, P., Pauperio, J., . . . Beja, P. (2019). Have the cake and eat it: Optimizing nondestructive DNA metabarcoding of macroinvertebrate samples for freshwater biomonitoring. *Mol Ecol Resour*, 19(4), 863-876. doi:10.1111/1755-0998.13012

Mayer, C., Sann, M., Donath, A., Meixner, M., Podsiadlowski, L., Peters, R. S., . . . Niehuis, O. (2016). BaitFisher: A Software Package for Multispecies Target DNA Enrichment Probe Design. *Molecular biology and evolution*, 33, 1875-1886. doi:10.1093/molbev/msw056

McCormack, J. E., Harvey, M. G., Faircloth, B. C., Crawford, N. G., Glenn, T. C., & Brumfield, R. T. (2013). A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. *PLoS One*, 8(1), e54848. doi:10.1371/journal.pone.0054848

Oksanen, J., Blanchet G.F., Friendly M., Kindt R., Legendre P., McGlenn D., Minchin P.R., O'Hara R. B., Simpson G. L., Solymos P., Henry M., Stevens H., Szoecs E. and Wagner H. (2019). *vegan*: Community Ecology Package. R package version 2.5-5. <https://CRAN.R-project.org/package=vegan>

Pajmans, J. L., Fickel, J., Courtiol, A., Hofreiter, M., & Forster, D. W. (2016). Impact of enrichment conditions on cross-species capture of fresh and degraded DNA. *Mol Ecol Resour*, 16(1), 42-55. doi:10.1111/1755-0998.12420

Pattée, E., & Goubault, N. (1981). Introduction pratique à la systématique des organismes des eaux continentales françaises. 1 Turbellariés Tricladés Paludicoles (planaires d'eau douce). *Publications de la Société Linnéenne de Lyon*, 50, 279-304.

Phuong, M. A., & Mahardika, G. N. (2018). Targeted Sequencing of Venom Genes from Cone Snail Genomes Improves Understanding of Conotoxin Molecular Evolution. *Mol Biol Evol*, 35(5), 1210-1224. doi:10.1093/molbev/msy034

Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, 15, 819-830. doi:10.1111/1755-0998.12355

Piñol, J., Senar, M. A., & Symondson, W. O. C. (2018). The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. *Molecular Ecology*. doi:10.1111/mec.14776

Pinto, A. J., & Raskin, L. (2012). PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLoS ONE*, 7. doi:10.1371/journal.pone.0043093

Piscart, C., & Bollache, L. (2012). Crustacés amphipodes de surface: gammares d'eau douce. *Association française de limnologie*. p122.

Porter, T. M., & Hajibabaei, M. (2018). Scaling up: A guide to high-throughput genomic approaches for biodiversity analysis. *Molecular Ecology*, 27, 313-338. doi:10.1111/mec.14478

Prié, V., Puillandre, N., & Bouchet, P. (2013). Bad taxonomy can kill: molecular reevaluation of *Unio mancus* Lamarck, 1819 (Bivalvia: Unionidae) and its accepted subspecies. *Knowledge and Management of Aquatic Ecosystems*, 11. doi:10.1051/kmae/2013071

Queney, P. (2004). Liste taxinomique des Coléoptères 'aquatiques' de la faune de France (avec leur répartition sommaire). *Le Coléoptériste*, 7, 3-27

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

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

Rognes, T., Quince, C., Nichols, B., Flouri, T., & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, 4, e2584. doi:10.7717/peerj.2584

Schnell, I. B., Thomsen, P. F., Wilkinson, N., Rasmussen, M., Jensen, L. R., Willerslev, E., . . . Gilbert, M. T. (2012). Screening mammal biodiversity using DNA from leeches. *Curr Biol*, 22(8), R262-263. doi:10.1016/j.cub.2012.02.058

Shokralla, S., F. Gibson, J., King, I., J. Baird, D., H. Janzen, D., Hallwachs, W., & Hajibabaei, M. (2016). Environmental DNA Barcode Sequence Capture: Targeted, PCR-free Sequence Capture for Biodiversity Analysis from Bulk Environmental Samples. doi:10.1101/087437

Sonet, G., Jordaens, K., Braet, Y., Bourguignon, L., Dupont, E., Backeljau, T., . . . Desmyter, S. (2013). Utility of GenBank and the Barcode of Life Data Systems (BOLD) for the identification of forensically important Diptera from Belgium and France. *ZooKeys*, 365, 307-328. doi:10.3897/zookeys.365.6027

Souty-Grosset, C., Holdich, D., Noel, P., Reynolds, J. D., & Haffner, P. (2006). Atlas of crayfish in Europe. *Muséum national d'Histoire naturelle*. p188

Stein, E. D., White, B. P., Mazor, R. D., Miller, P. E., & Pilgrim, E. M. (2013). Evaluating Ethanol-based Sample Preservation to Facilitate Use of DNA Barcoding in Routine Freshwater Biomonitoring Programs Using Benthic Macroinvertebrates. *PLoS ONE*, 8, 1-7. doi:10.1371/journal.pone.0051273

Sweeney, B. W., Battle, J. M., Jackson, J. K., & Dapkey, T. (2011). Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *Journal of the North American Benthological Society*, 30, 195-216. doi:10.1899/10-016.1

Tachet, H., Richoux, P., Bournard, M., & Usseglio-Polatera, P. (2010). *Invertébrés d'eau douce: systématique, biologie, écologie*. Paris: CNRS éditions.

Templeton, J. E. L., P.M., B., Llamas, B., J., S., Haak, W., Cooper, A., & Austin, J. J. (2013). DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investigative Genetics*, 5(26).

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, 183, 4-18. doi:10.1016/j.biocon.2014.11.019

Uchiyama, I., Mihara, M., Nishide, H., & Chiba, H. (2015). MGD update 2015: Microbial genome database for flexible ortholog analysis utilizing a diverse set of genomic data. *Nucleic Acids Research*, 43, D270-D276. doi:10.1093/nar/gku1152

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., . . . Dejean, T. (2015). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, n/a-n/a. doi:10.1111/mec.13428

Vallenduuk, H. J., & Cuppen, M. J. (2004). The aquatic living caterpillars (Lepidoptera: Pyraloidea: Crambidae) of Central Europe. A key to the larvae and autecology. *Lauterbornia*, 49, 1-17.

Vamos, E., Elbrecht, V., & Leese, F. (2017). Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics*, 1, e14625. doi:10.3897/mbmg.1.14625

Van Bortel, W., Harbach, R. E., Trung, H. D., Roelants, P., Backeljau, T., & Coosemans, M. (2001). Confirmation of *Anopheles varuna* in Vietnam, previously misidentified and mistargeted as the malaria vector *Anopheles minimus*. *American Journal of Tropical Medicine and Hygiene*, 65, 729-732. doi:10.4269/ajtmh.2001.65.729

van der Valk, T., Lona Durazo, F., Dalen, L., & Guschanski, K. (2017). Whole mitochondrial genome capture from faecal samples and museum-preserved specimens. *Mol Ecol Resour*, 17(6), e111-e121. doi:10.1111/1755-0998.12699

Wilcox, T. M., Piggott, M. P., Young, M. K., McKelvey, K. S., Schwartz, M. K., & Zarn, K. E. (2018). Capture enrichment of aquatic environmental DNA: A first proof of concept. *Molecular Ecology Resources*, 18, 1392-1401. doi:10.1111/1755-0998.12928

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, 79-82. doi:10.3354/cr00799

Yang, H., Ding, Y., Hutchins, L. N., Szatkiewicz, J., Bell, T. A., Paigen, B. J., . . . Churchill, G. A. (2009). A customized and versatile high-density genotyping array for the mouse. *Nature Methods*, 6, 663. doi:10.1038/nmeth.1359

Zerbino, D. R., Flicek, P., Juettemann, T., Zadissa, A., Lavidas, I., Achuthan, P., . . . Loveland, J. E. (2017). Ensembl 2018. *Nucleic Acids Research*, 46, D754-D761. doi:10.1093/nar/gkx1098

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Zhang, G. K., Chain, F. J. J., Abbott, C. L., & Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. *Evolutionary Applications*, 11, 1901-1914. doi:10.1111/eva.12694

Zhou, X., Li, Y., Liu, S., Yang, Q., Su, X., Zhou, L., . . . Huang, Q. (2013). Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. *GigaScience*, 2, 4. doi:10.1186/2047-217X-2-4

Zizka, V. M. A., Leese, F., Peinert, B., & Geiger, M. F. (2018). DNA metabarcoding from sample fixative as a quick and voucher-preserving biodiversity assessment method. *Genome*, 1-41. doi:10.1139/gen-2018-0048

2.11 Supplemental Information

Table S1 : Oligonucleotides used for barcoding, PCR and capture enrichment. The modified primers for amplicon enrichment correspond to Illumina tails plus the primer from the literature for the first PCR. The tags used for dual indexing correspond to a unique combination of eight bases that are represented here as XXXXXXXX (Table S3).

Barcoding		
Primer name	Oligonucleotide sequence (5' → 3')	Literature
LCO1490	GGTCAACAAATCATAAAGATATTGG	Folmer et al, 1994
HCO2198	TAACTTCAGGGTGACCAAAAAATCA	Folmer et al, 1994
16SarDr	CGCCTGTTTAAACAAAAACAT	Palumbi et al, 1991
16Sbr	CCGGTCTGAACTCAGATCACGT	Palumbi et al, 1996
PCR enrichment (first PCR)		
Primer name	Oligonucleotide sequence (5' → 3')	Literature
Modified primer BF2 with Illumina Nextera tail	TCGTCCGACGCGTCAGATGTGTATAAAGAGAC AGGCHCCHGAYATRGCHTTYCC	Elbrecht et al, 2017
Modified primer BR2 with Illumina Nextera tail	GTCTCGTGGGCTCGGAGATGTGTATAAAGAGA CAGTCDGGRTGNCCRAARAAYCA	Elbrecht et al, 2017
Modified primer fwhF1 with Illumina Nextera tail	TCGTCCGACGCGTCAGATGTGTATAAAGAGAC AGYTCHACWAAYCAYAARGAYATYGG	Vamos et al, 2017
Modified primer fwhR1 with Illumina Nextera tail	GTCTCGTGGGCTCGGAGATGTGTATAAAGAGA CAGARTCARTTWCCRAAHCCHCC	Vamos et al, 2017
Capture enrichment		
	Oligonucleotide sequence (5' → 3')	Literature
Illumina Nextera tail P5	TCGTCCGACGCGTCAGATGTGTATAAAGAGAC AG*T	
Illumina Nextera tail P7	[Phos]CTGTCTCTTATACACATCTCCGAGCCCA CGAGAC	
P5	AATGATACGGCGACCACCGA	
P7	CAAGCAGAAGACGGCATACGA	
Indexing (both PCR (second PCR) and capture enrichment)		
	Oligonucleotide sequence (5' → 3')	Literature
Illumina adapter P5 + tag	AATGATACGGCGACCACCGAACACXXXXXXX TCGTCCGACGCGTC	
Illumina adapter P7 + tag	CAAGCAGAAGACGGCATACGAGATXXXXXXX GTCTCGTGGGCTCGG	

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Table S2 : Retained order for bait design after selection based on sequence availability and marginality in streams upon a French species list. For each order, the numbers of species with and without a barcode available for bait design are provided as well as the percentage of species with a barcode.

Order	Number of species with a barcode	Number of species with no barcode	Percentage of species with a barcode
Amphipoda	14	5	73.68
Arhynchobdellida	6	2	75.00
Coleoptera	346	70	83.17
Diptera (Chironomidae)	670	0	100.00
Ephemeroptera	74	18	80.43
Gastropoda	31	4	88.57
Isopoda	4	1	80.00
Odonata	23	7	76.67
Plecoptera	99	19	83.90
Rhynchobdellida	11	4	73.33
Trichoptera	232	34	87.22
Venerida	15	0	100.00

Table S3 : Tag combination of each sample for all enrichment methods (BF2/BR2 PCR, fwh1 PCR, capture and no enrichment) and DNA templates (bulk and ethanol DNA). etDNA1 and etDNA2 correspond to the first and second extraction of DNA from ethanol, respectively.

PCR enrichment				
		Tag i5	Tag i7	
Bulk DNA	1	CTATTAAG	AGGCAGAA	
	2	TTCTAGCT	TAAGGCGA	
	3	GAGCCTTA	CGTACTAG	
	4	CCTAGAGT	TAAGGCGA	
	5	GCGTAAGA	TAAGGCGA	
	6	CTATTAAG	TAAGGCGA	
	7	AAGGCTAT	TAAGGCGA	
	8	AAGGCTAT	CGTACTAG	
	A	TCGACTAG	CGTACTAG	
	BF2/BR2	B	TTCTAGCT	CGTACTAG
		C	CCTAGAGT	CGTACTAG
		D	GCGTAAGA	CGTACTAG
		E	CTATTAAG	CGTACTAG
		F	TTATGCGA	CGTACTAG
G		TCGACTAG	AGGCAGAA	
H		TTCTAGCT	AGGCAGAA	
i		CCTAGAGT	AGGCAGAA	
J		GCGTAAGA	AGGCAGAA	
etDNA1		1	CTATTAAG	TCCTGAGC
	2	TCGACTAG	TCCTGAGC	
	3	GCGTAAGA	TCCTGAGC	
	4	CCTAGAGT	TCCTGAGC	
	fwh1	5	AAGGCTAT	TCCTGAGC
		6	TTCTAGCT	TCCTGAGC
		7	GAGCCTTA	TCCTGAGC
		8	TTATGCGA	TCCTGAGC
etDNA2	2	CCTAGAGT	TAGGCATG	
	3	TCGACTAG	TAGGCATG	
	5	TTCTAGCT	TAGGCATG	
	6	AAGGCTAT	TAGGCATG	
	fwh1	7	GAGCCTTA	TAGGCATG
		8	TTATGCGA	TAGGCATG
		9	GCGTAAGA	TAGGCATG
		10	CTATTAAG	TAGGCATG
	etDNA2	2	CCTAGAGT	CGAGGCTG
		3	TCGACTAG	CGAGGCTG
5		TTCTAGCT	CGAGGCTG	
6		TCGACTAG	AAGAGGCA	
BF2/BR2		7	TTCTAGCT	AAGAGGCA
		8	CCTAGAGT	AAGAGGCA
		9	GCGTAAGA	CGAGGCTG
		10	CTATTAAG	CGAGGCTG

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Capture enrichment			
		Tag i5	Tag i7
Bulk DNA	1	TCGACTAG	GTAGAGGA
	2	TTCTAGCT	GTAGAGGA
	3	GCGTAAGA	AAGAGGCA
	4	CTATTAAG	AAGAGGCA
	5	CCTAGAGT	GTAGAGGA
	6	GCGTAAGA	GTAGAGGA
	7	AAGGCTAT	AAGAGGCA
	8	GAGCCTTA	AAGAGGCA
	A	TTATGCGA	GTAGAGGA
	B	TCGACTAG	AAGAGGCA
	C	TTATGCGA	AAGAGGCA
	D	CCTAGAGT	CGAGGCTG
	E	TTCTAGCT	AAGAGGCA
	F	CCTAGAGT	AAGAGGCA
G	GCGTAAGA	CGAGGCTG	
H	CTATTAAG	CGAGGCTG	
i	AAGGCTAT	CGAGGCTG	
J	GAGCCTTA	CGAGGCTG	
etDNA	1	CTATTAAG	TAGGCATG
	2	TCGACTAG	TAGGCATG
	3	GCGTAAGA	TAGGCATG
	4	CCTAGAGT	TAGGCATG
	5	AAGGCTAT	TAGGCATG
	6	GAGCCTTA	TAGGCATG
	7	GCGTAAGA	CTCTCTAC
	8	TCGACTAG	CTCTCTAC
Enrichment-free	2	TTCTAGCT	CGAGGCTG
	4	TTATGCGA	CTCTCTAC
	5	GAGCCTTA	CTCTCTAC
	8	AAGGCTAT	CTCTCTAC

Table S4 : Assessment of COI PCR enrichment specificity on two types of DNA templates. "% of COI reads" : percentage of reads that align to the COI protein reference database, "% of COI reads assigned to a MC species": percentage of the number of reads that were successfully assigned to a species using the COI nucleotide reference databases to the COI assigned reads and "% of non-protostomian reads": percentage of reads that align to non-protostomian groups. Indicated values are mean number and standard deviation. For sample values, see Table S4 (PCR enrichment) and Table S5 (Capture enrichment). 10 sps: 10 species MC; 52 taxa:52 taxa MC.

Bulk DNA							
Samples	MC type	Raw	Quality filtered	Dereplicated	% reads assigned to COI	% of COI reads assigned to a MC species	% of reads assigned to non-protostomian
1	10 sps	232,590	154,023	112,841	45.60	97.62	2.89
2	10 sps	146,043	101,579	75,380	47.71	98.35	1.97
3	10 sps	128,848	68,221	48,296	56.33	98.88	0.89
4	10 sps	153,984	109,616	83,361	42.65	97.44	3.00
5	10 sps	146,348	95,113	71,087	45.33	94.38	3.01
6	10 sps	170,792	124,516	88,660	53.66	98.37	1.77
7	10 sps	149,057	102,314	72,570	53.08	93.28	5.27
8	10 sps	232,553	162,224	111,594	51.50	89.74	3.42
A	52 taxa	125,332	87,765	63,465	45.29	93.03	5.83
B	52 taxa	110,528	71,808	52,415	40.71	97.15	2.56
C	52 taxa	132,420	93,280	66,136	51.60	94.30	1.18
D	52 taxa	60,337	31,948	24,348	36.20	104.42	2.00
E	52 taxa	106,747	64,778	46,054	46.14	98.17	1.18
F	52 taxa	216,804	156,491	115,405	42.73	95.28	2.70
G	52 taxa	158,951	109,633	82,051	40.15	97.97	2.49
H	52 taxa	134,076	92,950	69,692	39.06	98.85	2.76
i	52 taxa	157,139	109,428	79,209	43.91	95.99	1.63
J	52 taxa	149,350	101,269	76,552	37.49	90.32	1.38
etDNA - extraction I - Fwh1							
Samples	MC type	Raw	Quality filtered	Dereplicated	% reads assigned to COI	% of COI reads assigned to a MC species	% of reads assigned to non-protostomian
1	10 sps	350,270	285,167	179,267	61.64	34.66	19.90
2	10 sps	249,675	205,449	131,510	63.49	32.57	46.08
3	10 sps	261,439	216,397	142,800	58.02	41.80	18.87
4	10 sps	209,425	171,501	115,639	61.99	48.67	5.12
5	10 sps	300,533	247,413	143,985	67.89	50.74	19.44

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

6	10 sps	365,331	301,504	165,707	69.96	66.47	15.05
7	10 sps	251,022	207,394	127,199	66.66	56.36	19.65
8	10 sps	281,454	236,022	134,735	69.36	39.65	4.80

etDNA - extraction II - Fwh1

Samples	MC type	Raw	Quality filtered	Dereplicated	% reads assigned to COI	% of COI reads assigned to a MC species	% of reads assigned to non-protostomian
2	10 sps	219,124	183,826	117,572	63.53	36.50	12.00
3	10 sps	221,143	174,165	116,192	56.36	27.78	8.35
5	10 sps	199,922	159,514	100,204	63.45	66.78	8.16
6	10 sps	314,441	269,902	142,307	75.55	61.46	7.78
7	10 sps	210,351	173,972	103,731	68.44	63.97	9.08
8	10 sps	220,764	185,927	108,669	70.16	47.81	2.14
9	10 sps	233,181	197,180	120,291	63.32	37.88	10.44
10	10 sps	229,080	191,159	112,987	69.98	39.61	6.12

etDNA - extraction II - BFR2/BR2

Samples	MC type	Raw	Quality filtered	Dereplicated	% reads assigned to COI	% of COI reads assigned to a MC species	% of reads assigned to non-protostomian
2	10 sps	172,036	89,048	63,687	49.06	85.89	14.47
3	10 sps	259,551	126,222	92,409	46.57	56.43	43.17
5	10 sps	205,397	124,323	89,377	46.37	41.03	26.02
6	10 sps	280,065	179,815	126,041	48.82	28.16	54.87
7	10 sps	155,288	91,748	66,017	47.92	46.02	50.38
8	10 sps	196,966	149,779	95,228	62.46	98.68	1.26
9	10 sps	153,339	101,732	74,365	39.58	52.86	13.14
10	10 sps	204,611	122,519	86,720	50.99	78.54	22.43

Table S5 : Assessment of COI capture enrichment specificity on two types of DNA templates. "% of targeted reads": percentage of reads that align to the COI protein reference database and 16S, NAD1, NAD4, NAD5, CYTB, and ATP6 nucleotide reference databases, "% of COI reads": percentage of reads that align to the COI protein reference database, "% of COI reads assigned to a MC species": percentage of the number of reads that were successfully assigned to a species using the COI nucleotide reference databases to the COI assigned reads and "% of non-protostomian reads": percentage of reads that align to non-protostomian groups. Indicated values are mean number and standard deviation. For sample values, see Table S4 (PCR enrichment) and Table S5 (Capture enrichment). 10 sps : 10 species MC ; 52 taxa :52 taxa MC.

Bulk DNA

Samples	MC type	Raw	Quality filtered	% of targeted reads	% reads assigned to COI	% of COI reads assigned to a MC species	% of reads assigned to non-protostomian
1	10 sps	456,417	289,201	17.33	6.66	61.73	14.52
2	10 sps	165,089	75,345	51.77	27.26	63.25	8.80
3	10 sps	857,040	339,778	70.20	45.44	61.55	8.94
4	10 sps	383,568	236,817	17.37	6.52	63.01	15.67
5	10 sps	459,650	234,785	37.26	17.94	62.92	10.56
6	10 sps	820,343	326,158	73.47	52.75	55.36	7.21
7	10 sps	381,006	173,037	63.06	38.78	60.60	6.52
8	10 sps	428,625	272,747	16.24	5.89	60.94	14.83
A	52 taxa	404,000	206,457	81.45	71.99	62.46	8.71
B	52 taxa	424,222	201,446	83.52	73.55	64.99	5.53
C	52 taxa	521,296	264,357	70.61	56.95	60.05	4.05
D	52 taxa	365,939	163,938	82.55	71.24	61.09	3.75
E	52 taxa	501,838	234,229	80.69	69.19	62.59	6.09
F	52 taxa	437,073	209,903	80.72	70.28	63.79	5.74
G	52 taxa	787,464	378,592	77.52	66.10	66.62	4.29
H	52 taxa	643,607	298,314	85.19	75.35	64.29	6.34
i	52 taxa	597,073	303,811	73.20	60.48	64.56	5.44
J	52 taxa	855,743	385,902	86.66	77.90	62.46	2.69

etDNA

Samples	MC type	Raw	Quality filtered	% of targeted reads	% reads assigned to COI	% of COI reads assigned to a MC species	% of reads assigned to non-protostomian
1	10 sps	575584	348292	28.20	21.93	46.65	49.14
2	10 sps	268611	150820	21.92	16.31	17.08	60.72
3	10 sps	217240	130783	19.52	12.50	35.76	69.49
4	10 sps	451088	257570	23.65	15.21	36.38	61.68
5	10 sps	772633	506973	24.44	18.51	22.57	60.24
6	10 sps	440802	262404	19.72	13.78	31.04	59.43
7	10 sps	500176	289496	39.19	28.05	30.81	35.94
8	10 sps	911484	530759	40.79	26.42	49.39	49.81

Table S6 : Number of reads recovered for each species in each sample for PCR and capture enrichment and for bulk and ethanol DNA on the 10 species mock communities. Reads were assigned with the blastn algorithm (Camacho et al, 2009). Only alignments with an e-value under 1.10^{-10} , a query cover over 200 bases for bulk DNA and over 90 bases for etDNA and an identity over 97% were conserved.

Samples	PCR enrichment									
	<i>Chironomus riparius</i>	<i>Epeorus assimilis</i>	<i>Heptagenia sulphurea</i>	<i>Isoptera rivulorum</i>	<i>Nemurella picteti</i>	<i>Hydropsyche siltalai</i>	<i>Athripsodes aterrimus</i>	<i>Physella acuta</i>	<i>Ancyclus fluviatilis</i>	<i>Gammarus fossarum</i>
1	840	59,940	3,333	2,838	301	851	467	0	0	0
2	250	43,366	1,509	1,431	67	385	652	0	0	0
3	261	35,628	539	1,078	60	200	231	0	0	0
4	277	41,473	2,340	609	42	344	464	0	0	0
5	1,716	28,691	2,354	4,098	2	1,922	1,905	0	0	0
6	246	63,481	690	689	4	326	296	0	0	0
7	104	49,004	787	485	9	179	89	0	0	0
8	2,357	55,593	7,386	6,114	202	2,476	845	0	0	0
1	335	8587	30182	121	6464	12287	2	436	1533	979
2	370	24030	24030	2	1233	9153	0	355	2124	807
3	163	5584	18747	51	9760	15954	0	390	1177	664
4	24	9054	5919	0	5555	30708	0	43	136	299
5	290	9056	32034	7	961	34095	0	435	2163	6185
6	289	766	21367	0	37	116177	0	254	1305	2
7	328	1640	41916	330	3467	29846	0	96	225	59
8	104	30227	8591	33	4687	20229	0	32	757	253
2	410	17950	16782	0	790	6121	0	88	382	107
3	343	5365	15949	13	1463	2718	0	133	921	368
5	88	1280	56779	0	15	9145	0	18	176	87
6	724	4124	5790	4	128	113707	0	172	610	52
7	307	4970	8221	7	980	61324	0	121	233	4
8	85	24863	20503	25	22	16647	0	14	197	7
9	242	10507	15707	126	20005	581	2	39	72	8
10	208	18051	24550	7	509	9348	0	89	185	37
2	814	35327	1022	55	97	16	190	0	0	3
3	1648	27180	2276	614	857	21	576	0	0	0
5	745	3228	19523	49	19	28	62	0	0	0
6	6489	12788	1454	377	54	2815	750	0	0	0
7	1425	17016	901	128	9	650	104	2	0	0
8	109	90009	2083	51	5	47	6	0	0	0
9	335	5883	1841	1352	6056	2	5818	0	0	0
10	373	37236	11167	56	120	42	71	0	0	0

Chapitre 3 - Apport de l'identification moléculaire pour l'étude des métacommunautés de têtes de bassin versant

Samples	Capture enrichment									
	<i>Chironomus riparius</i>	<i>Epeorus assimilis</i>	<i>Heptagenia sulphurea</i>	<i>Isoperla rivulorum</i>	<i>Nemurella picteti</i>	<i>Hydropsyche siltalai</i>	<i>Athripsodes aterrimus</i>	<i>Physella acuta</i>	<i>Ancyclus fluvialis</i>	<i>Gammarus fossarum</i>
1	411	3838	2837	1122	411	2058	729	231	210	41
2	408	2074	4300	1222	473	2756	1035	165	551	6
3	3674	16319	24326	9227	2043	24608	8501	1532	4668	141
4	308	2595	1793	710	362	2283	903	175	603	3
5	1393	2538	7179	2315	70	8379	2788	236	1597	0
6	3560	21675	26072	7480	340	22341	7396	1829	4560	0
7	1059	12690	11915	3179	144	7900	1670	446	1665	0
8	516	1593	2966	920	49	2685	613	74	380	0
1	168	1552	5488	1390	557	412	24330	539	1185	4
2	166	363	1348	13	83	244	269	566	1147	2
3	121	144	1759	152	1026	481	384	678	1078	26
4	203	2926	3545	28	2918	3377	318	405	524	8
5	640	1743	6211	224	38	3919	1012	2200	5185	14
6	268	109	2327	6	18	6505	155	359	1477	0
7	998	505	14337	2845	791	3491	739	363	956	0
8	858	37178	8248	565	4540	7602	373	538	9359	5
2	0	0	0	1	1	1	4	0	0	0
4	0	1	1	1	2	2	3	0	0	0
5	1	19	13	2	6	49	5	3	21	0
8	0	1	5	0	2	7	4	1	0	0

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Table S7 : Number of reads recovered for each taxon in each sample for PCR and capture enrichment for bulk DNA on the 52 taxa mock communities. Reads were assigned with the blastn algorithm (Camacho et al, 2009). Only alignments with an e-value under 1.10^{-10} , a query cover over 200 bases for bulk DNA and an identity over 97% were conserved.

PCR enrichment										
Taxa	A	B	C	D	E	F	G	H	i	J
<i>Agabus</i>	8	34	109	4	0	33	2	4	5	95
<i>Agapetinae</i>	123	375	0	87	20	323	133	215	25	41
<i>Ancylus</i>	0	0	0	0	0	2	0	0	0	0
<i>Anomalopterygella chauvinia</i>	392	253	66	71	272	558	190	201	215	155
<i>Arhynchobdellida</i>	345	227	260	50	210	258	90	10	352	9
<i>Asellus aquaticus</i>	306	178	140	49	5681	2430	618	263	915	168
<i>Baetis</i>	554	3175	108	188	218	3474	526	494	1261	349
<i>Bivalvia</i>	12	4	0	0	2	10	0	4	0	0
<i>Blephariceridae</i>	1671	3670	310	250	1402	1348	543	366	782	154
<i>Ceratopogonidae</i>	284	210	176	209	165	1036	55	310	60	100
<i>Chaetopteryx villosa</i>	334	245	127	56	68	27	76	48	30	130
<i>Chironomidae</i>	763	135	172	32	1032	1402	304	380	0	8
<i>Cordulegaster boltonii</i>	29	13	32	27	81	295	81	89	62	44
<i>Daphnia pulex</i>	9	7	0	0	3	11	7	0	3	4
<i>Dicranota</i>	66	9	22	0	3	63	13	25	9	6
<i>Drusus annulatus</i>	340	227	112	32	216	611	209	156	129	105
<i>Dugesia</i>	25	45	28	14	40	160	30	51	45	19
<i>Ecdyonurus</i>	222	438	4848	101	485	5346	2448	96	214	14696
<i>Epeorus</i>	1444	788	24613	361	932	1813	1892	731	2782	4384
<i>Ephemera (Torleya) major</i>	207	1970	1214	1453	2169	843	8644	7155	4177	1173
<i>Ephemera mucronata</i>	16	137	68	15	68	451	105	194	60	30
<i>Ephemeraeidae</i>	86	3407	2364	798	221	2466	106	134	1137	2054
<i>Gammarus</i>	50	15	40	8	46	353	110	54	145	24
<i>Gerroidea</i>	6617	699	444	513	707	2746	4354	3336	1723	361
<i>Halesus</i>	135	64	198	190	87	411	136	182	335	223
<i>Hydropsyche</i>	6	28	15	5	38	177	41	42	31	26
<i>Isoperla</i>	48	332	225	184	522	1038	592	554	582	254
<i>Leptophlebiidae</i>	544	1225	274	150	513	3105	693	3602	4281	1006
<i>Leuctra</i>	531	691	534	988	521	5155	471	513	126	315
<i>Limnius</i>	174	7	19	17	21	308	50	23	2	5
<i>Limoniidae</i>	166	109	20	19	29	166	75	50	99	13
<i>Lymnaeidae</i>	7	0	0	0	0	6	0	0	2	0
<i>Nemouridae</i>	3909	941	41	240	2933	4677	4818	3027	647	246
<i>Odontocerum albicorne</i>	543	302	252	23	170	677	348	164	231	135
<i>Oecismus monedula</i>	2882	454	262	111	375	1700	856	268	336	258
<i>Oreodytes sanmarkii</i>	2354	727	2363	681	2696	4021	9409	6883	15257	5078
<i>Perlidae</i>	1077	478	1219	1342	2022	1851	616	1238	1007	605

Chapitre 3 - Apport de l'identification moléculaire pour l'étude des métacommunautés de têtes de bassin versant

<i>Philopotamidae</i>	49	16	9	6	30	124	19	46	24	34
<i>Polycentropodidae</i>	0	13	12	0	0	48	79	30	23	17
<i>Potamophylax</i>	186	1386	223	109	282	1472	694	470	150	108
<i>Ptychoptera</i>	119	101	146	54	83	526	196	107	58	70
<i>Rhithrogena</i>	136	641	1224	427	1347	3184	729	1506	1214	6
<i>Rhyacophila</i>	97	71	44	16	76	374	31	15	62	15
<i>Scirtidae</i>	8718	2882	2055	2212	2244	3879	955	1877	5143	260
<i>Sericostoma personatum</i>	695	1163	509	428	297	1314	1028	598	718	542
<i>Sialis</i>	44	41	71	184	295	1419	185	158	868	707
<i>Silo</i>	57	25	72	41	279	547	96	2	473	111
<i>Simuliidae</i>	80	68	29	2	15	177	4	72	23	23
<i>Thremma gallicum</i>	144	136	132	152	253	65	125	15	54	11
<i>Tipulidae</i>	271	166	114	121	61	901	248	41	175	56
<i>Trombidiformes (Acari)</i>	104	73	71	56	114	327	93	93	70	5

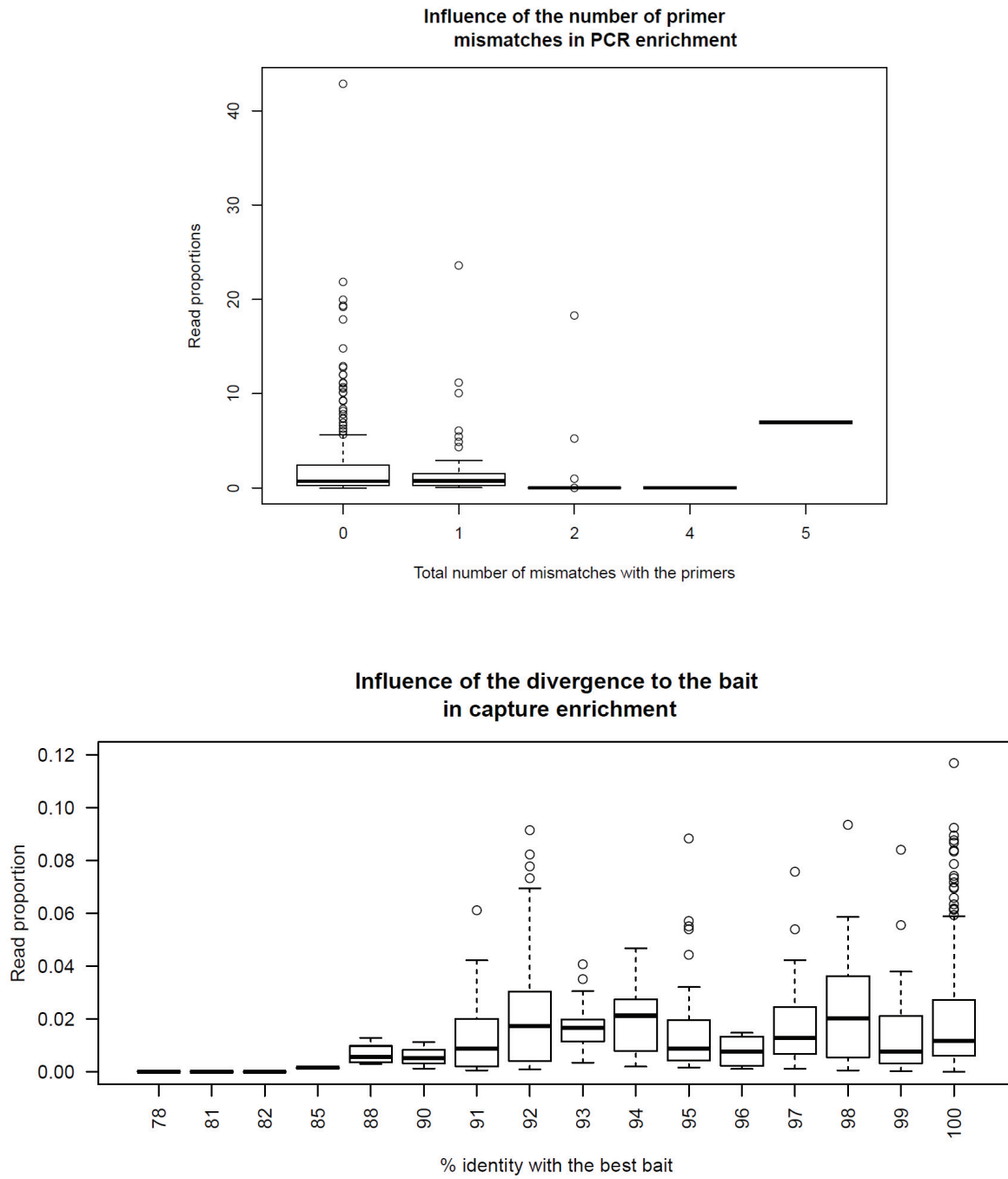
Capture enrichment

Taxa	A	B	C	D	E	F	G	H	i	J
<i>Agabus</i>	151	3341	5449	125	433	368	249	717	511	8472
<i>Agapetinae</i>	2809	1926	0	1375	763	1585	1699	2617	1532	1311
<i>Ancylus</i>	1840	814	585	174	856	620	799	951	245	445
<i>Anomalopterygella chauvinia</i>	671	758	565	506	1131	521	896	1179	1025	1330
<i>Arhynchobdellida</i>	826	619	1561	243	1451	190	1268	115	2185	242
<i>Asellus aquaticus</i>	7267	4036	4456	3549	3995	5734	13119	8695	12719	10836
<i>Baetis</i>	1937	9326	1301	1914	1439	3944	3336	4001	4143	6070
<i>Bivalvia</i>	531	242	179	51	86	109	23	1013	167	50
<i>Blephariceridae</i>	3487	4579	1926	1710	3497	1585	2367	2478	2620	2106
<i>Ceratopogonidae</i>	355	89	154	451	303	308	1427	486	250	713
<i>Chaetopteryx villosa</i>	1056	1137	1338	611	495	82	880	787	442	2142
<i>Chironomidae</i>	3694	1526	2201	320	4547	3841	3746	5078	528	4796
<i>Cordulegaster boltonii</i>	735	273	1158	1140	2620	2484	3403	4104	2598	3120
<i>Daphnia pulex</i>	5	1	4	1	1	1	0	0	0	0
<i>Dicranota</i>	815	223	728	215	185	520	1066	757	406	486
<i>Drusus annulatus</i>	738	861	944	328	1311	776	1873	1735	1181	2424
<i>Dugesia</i>	0	0	0	0	0	0	4	0	0	0
<i>Ecdyonurus</i>	1361	3236	601	1445	4083	345	432	1933	1954	2274
<i>Epeorus</i>	574	290	5879	265	510	455	521	571	732	5046
<i>Ephemerella (Torleya) major</i>	219	1958	2092	1725	2007	353	4885	5415	2720	3739
<i>Ephemerella mucronata</i>	236	2703	1846	734	2254	1593	2409	5252	1356	2599
<i>Ephemeridae</i>	274	3859	4127	2556	800	2818	848	1019	2532	11004
<i>Gammarus</i>	328	227	624	132	632	621	1075	1091	1357	937
<i>Gerroidea</i>	7833	1859	2109	2159	2597	3472	11773	11183	4865	3251
<i>Halesus</i>	748	991	1831	854	914	726	1236	1781	1662	2807
<i>Hydropsyche</i>	3170	1247	3008	1698	734	937	5741	4677	608	6350
<i>Isoperla</i>	371	2113	1874	1565	3084	1463	4778	6121	2852	4591
<i>Leptophlebiidae</i>	1459	4139	1522	1713	2264	1781	3848	4616	10116	10286

2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

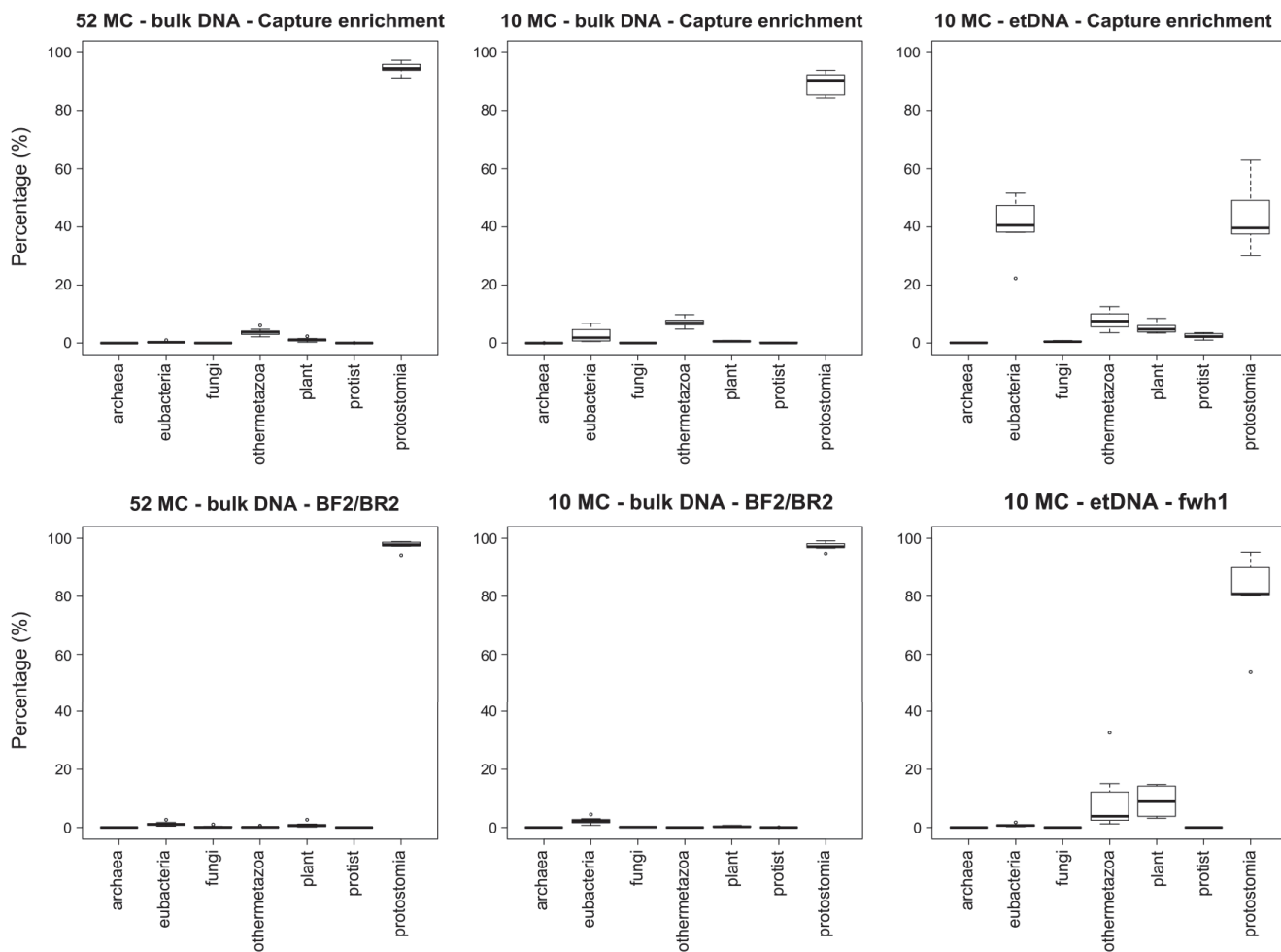
<i>Leuctra</i>	1398	2119	2990	3742	2087	6721	2728	5210	1040	4361
<i>Limnius</i>	1576	277	414	747	262	888	1474	899	273	823
<i>Limoniidae</i>	1422	1191	506	704	877	731	1223	846	793	775
<i>Lymnaeidae</i>	2112	1544	405	850	1155	1672	984	1476	1933	415
<i>Nemouridae</i>	9675	5020	1655	5883	10129	6112	6591	16068	3186	4872
<i>Odontocerum albicorne</i>	5699	5470	7248	2492	5915	4751	9563	4745	4733	8410
<i>Oecismus monedula</i>	2314	1286	1347	688	1357	1453	2880	1704	1646	2800
<i>Oreodytes sanmarkii</i>	3101	2153	2108	1000	978	4052	10687	4615	7555	9667
<i>Perlidae</i>	614	348	456	2709	2257	627	636	861	437	734
<i>Philopotamidae</i>	6080	2416	3397	5190	5376	4442	4982	6736	3949	13881
<i>Polycentropodidae</i>	143	2366	3220	707	93	2404	10300	6222	2970	4956
<i>Potamophylax</i>	209	1707	667	280	722	1139	1669	1672	1021	1132
<i>Ptychoptera</i>	1314	2139	6760	3532	3631	2781	8030	4719	2230	7487
<i>Rhithrogena</i>	388	1702	3360	1794	2877	1676	3265	2289	2425	5984
<i>Rhyacophila</i>	1580	578	1055	730	1139	1101	841	733	806	1617
<i>Scirtidae</i>	3237	2971	3747	2398	1740	1441	3282	2652	3215	3315
<i>Sericostoma personatum</i>	363	651	579	473	253	336	944	761	582	1363
<i>Sialis</i>	339	1301	2331	3421	4012	4720	3742	4589	5986	15964
<i>Silo</i>	2120	2137	2448	2387	8151	5953	3442	668	9485	4623
<i>Simuliidae</i>	2339	1335	937	689	985	800	2048	2244	1063	1362
<i>Thremma gallicum</i>	2384	964	3499	1599	1677	527	5659	1107	848	1272
<i>Tipulidae</i>	464	360	460	461	150	607	702	197	482	456
<i>Trombidiformes (Acari)</i>	143	131	178	100	183	163	272	363	168	380

Figure S1 : Robustness of PCR (top) and capture (bottom) enrichment to mismatches between the targeted species COI and the best pair of degenerated primers or the best bait.



2. Développement d'un outil moléculaire pour l'étude des métacommunautés (article 3)

Figure S2 : Percentage of reads for different coarse taxonomic group (archaea, eubacteria, fungi, plant, protist, protostomia (here Arthropoda, Annelida, Brachiopoda, Mollusca, Nematoda and Platyhelminthes) and other metazoa (corresponding to deuterostomia)) for each mock community type (10 sps = 10 species MC; 52 taxa=52 taxa MC), template DNA (bulk DNA and etDNA) and enrichment method (capture enrichment and PCR enrichment (BF2/BR2 primer pair for bulk DNA and fwh1 primer pair for etDNA)). The assignment was obtained by comparing the quality filtered reads to a protein database designed using diamond (blastx, more sensitive option, e-value threshold of



1.10⁻¹⁰) (Buchfink et al, 2015).

3 Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés - Étude exploratoire sur le bassin versant de l'Albarine (article 4)

Après avoir développé dans l'article 3 une méthode sur des communautés composées en laboratoire pour optimiser l'identification par metabarcoding des invertébrés benthiques de têtes de bassin versant (c'est-à-dire obtenir des données de bonne qualité pour l'étude de métacommunautés), une application sur des échantillons récoltés *in situ* (c'est-à-dire des communautés échantillonnées dans des rivières) pour la valider semble pertinente et nécessaire.

Dans cet article, nous avons donc appliqué une des méthodes développées dans l'article 3, à savoir l'amplification par PCR à partir de l'ADN extrait de l'éthanol de conservation des invertébrés (etDNA), sur des échantillons prélevés lors de la campagne d'hiver 2017-2018 du bassin versant de l'Albarine (cf. Matériel et méthodes - Chapitre 2). Comme vu en Introduction et dans l'article 3, l'identification moléculaire permettrait d'identifier les organismes composant une communauté à un niveau taxinomique fin, celui de l'espèce. L'objectif de cet article est de comparer et d'évaluer les complémentarités des identifications morphologique et moléculaire pour l'étude des métacommunautés de macro-invertébrés benthiques avec un focus sur les groupes peu connus, abondants et fréquents des têtes de bassin versant.

3.1 Introduction

Détecter et mettre en évidence les processus environnementaux, spatiaux et évolutifs qui conditionnent la structure des communautés est un des objectifs principaux en écologie des communautés (Vellend, 2010). L'approfondissement des connaissances sur les patrons de distribution permet d'appuyer les concepts théoriques et de comprendre de plus en plus finement ces processus. Un axe d'amélioration majeur par rapport à cet approfondissement porte sur l'utilisation systématique d'une résolution taxinomique à l'espèce plutôt qu'à des niveaux taxinomiques plus élevés. En effet, traditionnellement, l'identification des organismes se base sur des critères morphologiques et est associée à plusieurs limites (Hebert et al., 2003). La perte des compétences en taxinomie et le manque d'experts, les lacunes dans les clés de détermination, la difficulté à identifier les stades immatures ou les organismes endommagés, couplé à l'augmentation importante des coûts et du temps que l'identification à l'espèce engendre, amène à une identification à des niveaux taxinomiques élevés chez certains groupes (p.ex., classe ou famille) plutôt qu'à l'espèce (Heino & Soininen, 2007). Les études actuelles font l'hypothèse de la suffisance taxinomique (« taxonomic sufficiency » en anglais) mais sans l'avoir testée préalablement dans le contexte particulier d'une étude, ce qui pourrait affecter notre capacité à caractériser et comprendre la structure des communautés (Godoy et al., 2019) et

3. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés - Étude exploratoire sur le bassin versant de l'Albarine (article 4)

masquer des différences entre communautés que ce soit pour estimer des liens entre les communautés et les conditions environnementales (Heino, 2014), les facteurs spatiaux (Godoy et al., 2019) ou les facteurs historiques (Losos, 2008).

L'identification moléculaire par barcoding a été proposée et utilisée pour résoudre les problèmes liés à l'identification morphologique (Hebert et al., 2003). D'une part, elle permet de surmonter les problèmes inhérents à l'identification morphologique. D'autre part, elle permet d'identifier des OTUs (Operational Taxonomic Units ou Unités taxinomiques opérationnelles) utilisés comme proxy d'espèce, mais pouvant même être plus performants en identifiant des structures génétiques au sein d'une même espèce (Elbrecht et al., 2018). Le barcoding a déjà permis de mettre en évidence des compositions et structures de communauté qui n'étaient pas visibles avec la seule identification morphologique. Par exemple, Sweeney et al. (2011) ont montré que l'identification moléculaire augmentait significativement les indices de richesse, notamment chez les groupes mal connus comme les Chironomidae présents en forte abondance dans les cours d'eau pollués. Le barcoding a aussi permis de différencier des communautés de sites altérés et non altérés dans des rivières (Stein et al., 2014) ou de mettre en évidence des patrons de distribution spatiaux chez des groupes identifiés morphologiquement à des niveaux taxinomiques grossiers (Martin et al., 2016). Le barcoding possède cependant une limite quant à l'étude des communautés : il ne permet l'identification que d'un spécimen par événement de séquençage (c'est-à-dire, traitement et identification d'un organisme à la fois). Le metabarcoding a alors émergé et a permis de surmonter ce problème en permettant d'identifier les espèces de communautés de plusieurs échantillons en un événement de séquençage (Pompanon et al., 2011), malgré des challenges méthodologiques toujours présents (Elbrecht et al., 2017, article 3). Le metabarcoding a déjà montré son efficacité par rapport à l'identification morphologique dans le cadre des effets des conditions environnementales sur les communautés (Andujar et al., 2018) mais ses apports dans le cadre des métacommunautés - l'étude des effets des processus liés à l'environnement et à la dispersion sur la structure des communautés - restent à être montrés.

Les rivières forment un réseau dendritique interconnecté où la dispersion est déterminante dans la structure des communautés (Benda et al., 2004 ; Tonkin et al., 2018). Bien que la génétique des populations ait permis de mettre en évidence des structures spatiales chez certaines espèces (Alp et al., 2012 ; Prunier et al., 2018), cette approche est limitée lorsque l'on s'intéresse à des groupes très diversifiés avec une taxinomie complexe comme les invertébrés. L'étude des métacommunautés au grain spécifique représente alors un cadre de travail privilégié qui pourrait mettre en évidence des tendances générales dans la distribution des communautés ou de groupes spécifiques en affinant la description des liens entre communautés, environnements et distances spatiales. Ceci est d'autant plus pertinent dans les réseaux de rivières présentant des tronçons intermittents où les groupes abondants sont souvent mal connus et à priori ubiquistes comme les Chironomidae ou la sous-classe des

Oligochaeta (Stubbington et al., 2018 ; Stubbington et al., 2009). Pour ces taxa chez qui la résolution taxinomique est faible, on ne sait pas si (i) un faible nombre d'espèces est représenté qui sont effectivement abondantes et ubiquistes dans le bassin versant et possèdent des fortes capacités de dispersion ou si (ii) un grand nombre d'espèces proches sont représentées et montrent des structures indécélables à une résolution taxinomique grossière, par exemple liées à l'environnement. Actuellement, de par leur abondance forte et leur omniprésence dans les communautés, ces taxa « brouillent » les indices de diversité. Pourtant, ces groupes pourraient présenter une diversité spécifique et fonctionnelle forte révélant des patrons de distribution invisible avec la seule identification morphologique. Par exemple, la famille des Chironomidae est représentée par des centaines d'espèces et présente une large gamme de réponses aux assèchements (différentes stratégies de résistance et résilience, Cañedo-Argüelles et al., 2015). Étant souvent identifié à la famille ou à la sous-famille, il est supposé qu'une meilleure résolution taxinomique grâce à l'identification moléculaire permettrait une meilleure caractérisation de la réponse des espèces à la variabilité environnementale dans les rivières intermittentes et ainsi permette de mieux décrire l'état écologique de ces systèmes (Stubbington et al., 2018). Une meilleure caractérisation de leurs communautés permettrait également de mieux différencier les communautés les unes des autres et de mettre en évidence des patrons de distribution en lien avec des conditions environnementales particulières qui traduiraient par exemple une adaptation locale des espèces à un fort filtre environnemental. Elle permettrait aussi une description plus réaliste des processus de dispersion ayant lieu au sein des métacommunautés. Par exemple, au sein des Chironomidae, Cañedo-Argüelles et al. (2015) ont mis en évidence que les assèchements représentaient un filtre environnemental fort en empêchant les espèces non adaptées aux assèchements de coloniser et de s'installer dans les cours d'eau intermittents, mettant en évidence une différence de distribution en fonction des types de cours d'eau. De plus, certaines espèces non adaptées aux assèchements seraient résilientes et coloniseraient les cours d'eau intermittents à partir de zones refuges, ce qui devrait laisser une empreinte dans les patrons de diversité alpha et beta.

Dans cet article, l'objectif a été de comparer les identifications morphologique et moléculaire pour l'étude de métacommunautés d'invertébrés de rivières chez des groupes abondants et à priori ubiquistes où l'identification morphologique à l'espèce n'est pas atteinte. L'identification morphologique ne permet pas une résolution taxinomique fine chez certains groupes et regroupe ainsi plusieurs espèces, occultant pour ces groupes des patrons de distribution spatiale ou liée à l'environnement si leurs préférences environnementales divergent. Par conséquent, nous faisons l'hypothèse que l'identification moléculaire va permettre de mettre en évidence de nouveaux patrons de distribution chez ces groupes grâce à une résolution taxinomique moléculaire à l'OTU induisant une meilleure discrimination de leurs communautés entre elles. Pour cela, nous avons comparé les résultats d'analyse de métacommunautés pour les deux approches, morphologique et moléculaire, sur

3. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés - Étude exploratoire sur le bassin versant de l'Albarine (article 4)

des groupes ciblés de communautés d'invertébrés de la tête de bassin de l'Albarine en France étudiée dans l'article 1.

3.2 *Matériel et méthodes*

Les échantillons associés à l'identification morphologique ont déjà fait l'objet d'une étude (article 1, Gauthier et al, 2020, In Press) mais les étapes de traitement principaux les concernant sont répétés dans cet article.

3.2.1 *Échantillonnage*

Le bassin versant de l'Albarine est un sous bassin de l'Ain, lui-même sous bassin du Rhône, situé dans le département de l'Ain dans l'Est de la France (Figure 1.1). Le placement des sites d'échantillonnage sur la tête de bassin versant de l'Albarine, représentant 99 km², a été faite selon plusieurs critères : les sites devaient être répartis de manière homogène dans le bassin versant, un site ne pouvait être localisé à moins de 200 mètres d'une confluence (Altermatt, 2013) ou d'une structure anthropique comme des seuils (Fuller et al., 2015) afin de limiter les biais liés à ces facteurs sur la composition des communautés et, au regard des connaissances du bassin versant, toutes les sections intermittentes devaient comporter un site d'échantillonnage. Vingt-et-un sites d'échantillonnage ont été définis et les échantillons ont été prélevés en février 2018 dans 20 sites (un des sites étant asséché, Figure 3.2). Sur chaque site, deux prélèvements ont été effectués en tête de radier : un destiné à l'identification morphologique, l'autre à l'identification moléculaire. Les prélèvements ont été réalisés à l'aide d'un filet Surber (0,1 m²) à maille de 500 µm. Ils ont ensuite été nettoyés et tamisés puis le refus du tamis a été conditionné dans des sacs Whirl-pak® pour les échantillons destinés à l'identification morphologique et dans des flacons hermétiques pour ceux destinés à l'identification moléculaire, et, dans les deux cas, fixés avec de l'alcool à 96°.

A chaque site d'échantillonnage, 12 variables environnementales ont été mesurées. Aux deux points de prélèvement des invertébrés, la hauteur d'eau (cm), la largeur mouillée (cm), la vitesse du courant (cm.s⁻¹ mesurée à 40% de la hauteur d'eau avec un anémomètre Miniwater20 (Schiltknecht, Gossau, Switzerland)) et le type de substrat primaire et secondaire ont été mesurés. Le type de substrat était classé en 14 classes (surface uniforme dure, algue, sable, limon, vase, spermaphyte émergent, graviers, blocs, galets, substrat ligneux, chevelu racinaire, litière, spermaphyte immergé, bryophyte). L'altitude (m), la latitude et la longitude de chaque site ont été obtenus avec un GPS Garmin. La conductivité (µS.cm⁻¹) et la température (°C) ont été mesurées avec un conductimètre HQ14d. Le pH et le pourcentage d'oxygène dissous ont été mesurés avec une sonde multi-paramètres HQ40d (HACH, Loveland, Colorado, USA) en amont des points de prélèvements pour éviter un biais dans les mesures. Sur un tronçon de 20 fois la largeur mouillée moyenne, le pourcentage d'ombrage a été estimé visuellement et la pente (%) a été mesurée avec un clinomètre.

3.2.2 Identification des invertébrés

(a) Identification morphologique

L'identification morphologique des invertébrés s'est effectuée en laboratoire. Les organismes ont été triés, comptés et identifiés au niveau taxinomique le plus fin, dans les limites des connaissances taxinomiques et des possibilités techniques (cf. liste des références, <https://doi.org/10.5281/zenodo.3569563>). Suivant les groupes, les niveaux d'identification ont donc été variables (cf Tableau 1.1). La plupart des insectes, crustacés, Tricladida et tous les mollusques ont été identifiés à l'espèce lorsque cela était possible (i.e. si les organismes étaient suffisamment matures). Au minimum, les organismes de l'ordre des Diptera ont été identifiés au niveau du genre, ceux des ordres Coleoptera, Hemiptera et une partie des Trichoptera et Amphipoda au niveau du genre et ceux des Ostracoda et Copepoda au niveau de l'ordre. Les Achaeta et Tricladida ont été identifiés au niveau du genre. Les Nematoda, Hydracaria, Hydrozoaires et Oligochaeta ont été identifiés au niveau de la classe ou de la sous-classe.

(b) Identification moléculaire

Pour l'identification moléculaire, l'ADN extrait de l'éthanol de conservation des organismes (ADNet) a été utilisé. En effet, les conclusions de l'article 3 nous ont amenées à opter pour cette matrice rapide et non destructive qui apporte les mêmes informations qualitatives (c'est-à-dire de détection) que l'ADN bulk, seules informations actuellement utilisables avec le metabarcoding. Pour chaque échantillon, 48 mL d'éthanol de conservation ont été collectés. Du glycogène et de l'acétate de sodium ont été ajoutés pour précipiter l'ADN. Les échantillons ont été placés à -80°C pendant 72 heures. Après centrifugation, l'éthanol a été retiré et, après évaporation totale de l'éthanol, le résidu sec a été dissous dans la solution tampon du kit FastDNA® Spin Kit for Soil (MP Biomedicals), kit permettant d'éliminer les potentiels inhibiteurs de PCR (comme l'acide humique) supposés présents dans l'éthanol de conservation de communautés prélevés *in situ*. L'ADN a été extrait en suivant les instructions du fabricant. Un blanc d'extraction contrôle a également été effectué par remplissage d'un échantillon d'éthanol 96% qui a ensuite été extrait selon le même protocole.

Les échantillons ont été amplifiés par PCR avec le couple d'amorces Fwh1 auxquelles ont été ajoutés des adaptateurs Illumina Nextera® (Vamos et al., 2017). Ce couple d'amorces amplifiant un fragment de 178 pb dans la région Folmer du COI a été conçu pour détecter les invertébrés dans des matrices ADN dégradées tel que l'ADNet. L'enrichissement par PCR a été préféré à celui par capture par sondes ARN car certains taxa potentiellement cibles, à savoir les Oligochaeta et les autres familles de Diptera, n'étaient pas représentés dans le set de sondes conçu lors de cette thèse. La PCR a été faite en triplicatas dans un volume total de 25 µL avec une solution tampon 1X contenant du MgCl₂ 3 mM, 400 µM de chaque dNTP et la polymérase (QIAGEN® Multiplex PCR kit, Qiagen, Allemagne), 0.5

3. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés - Étude exploratoire sur le bassin versant de l'Albarine (article 4)

μM de chaque amorce et 10 ng d'ADN pour l'ADN bulk ou 5 μL d'ADN pour l'ADNet. L'amplification consistait en une dénaturation initiale de 5 min à 95°C suivie par 35 cycles composés d'une étape de dénaturation à 95°C de 30 sec, d'une d'hybridation à 52°C de 30 sec et d'une d'élongation à 72°C de 2 min suivie d'une extension finale à 72°C de 10 minutes. Les produits PCR ont été regroupés par triplicatas et purifiés avec des billes de purification Agencourt AMPure XP (Beckman Coulter). Les produits PCR ont été quantifiés avec le kit QuantiFluor® dsDNA System (Promega). Dix ng de chaque produit PCR ont été utilisés pour une seconde PCR qui a été faite dans un volume total de 25 μL avec une solution tampon 1X (BIOAmp® Blend Mix ; Biofidal), 2.5 nM de MgCl_2 , 200 μM de chaque dNTP, 0.25 μM de d'amorce Illumina (Tableau 1.2) et 0.02 U/ μL de HOT BioAmp Taq (Biofidal). L'amplification consistait en une dénaturation initiale de 5 min à 96°C suivie par 12 cycles composés d'une étape de dénaturation à 95°C de 20 sec, d'une d'hybridation à 54°C de 30 sec et d'une d'élongation à 72°C de 1 min suivie d'une extension finale à 72°C de 10 minutes. Les librairies obtenues ont ensuite été purifiées avec des billes de purification Agencourt AMPure XP (Beckman Coulter), quantifiées avec le kit QuantiFluor® dsDNA System (Promega) puis regroupées à la même concentration pour le séquençage. Un blanc de PCR contrôle a également été fait en remplaçant l'ADN par de l'eau ultra pure à la première PCR. Les librairies ont été séquencées en utilisant le kit de séquençage 2*250 paired-end V3 MiSeq par l'entreprise Biofidal.

Les lectures ont été livrées démultiplexées et sans les adaptateurs Illumina Nextera®. Les lectures forward et reverse ont été fusionnées avec un minimum de recouvrement de 10 nucléotides et une différence maximale sur la zone de recouvrement de 5 nucléotides avec Vsearch 2.8.4 (Rognes et al., 2016). Ensuite les séquences des amorces ont été retirées avec cutadapt 1.9.1 (Martin, 2014) et les lectures ont été filtrées avec Vsearch (valeur ee maximale de 1, longueur minimale de 200 nucléotides et aucun N autorisé). Les séquences ont été dérépliquées et les singletons retirés. Les séquences chimères ont ensuite été *de novo* détectées et retirées avec Vsearch. Ensuite les séquences ont été regroupées en cluster selon un seuil de similarité à 97% avec Vsearch, chaque cluster étant représenté par un OTU. Enfin, les OTUs avec une abondance inférieure à 0.01% dans tous les échantillons ont été retirés car ces OTUs peuvent être le produit d'erreurs de PCR ou de séquençage. Les OTUs ont été assignés à des espèces en utilisant la base de données publique BOLD et l'algorithme associé (Ratnasingham & Hebert, 2007). Comme conseillé par les auteurs, une vérification systématique de l'assignement a été effectuée par la visualisation des arbres générés par Neighbour-Joining lors de l'assignement et les identifications ont été corrigées en fonction de ces arbres.

3.2.3 Analyse de métacommunautés

Les groupes les plus abondants (avec les nombres d'organismes les plus hauts tous sites confondus) et les groupes les plus fréquents (détectés dans les plus grands nombres de sites) ont été déterminés pour tous les sites échantillonnés et en prenant les sites intermittents et les sites pérennes séparément à

partir de données issues de l'identification morphologique, étant donné que l'identification moléculaire avec l'ADNet ne permet pas de quantifier les abondances de manière fiable (article 3). Un groupe était considéré abondant lorsque son abondance dans le jeu de données était supérieure à 10%. Un groupe était considéré fréquent lorsqu'il était présent dans plus de 70% des sites d'échantillonnage. Pour définir les groupes d'intérêt parmi ces groupes pour les analyses, deux critères ont été utilisés : (i) la résolution taxinomique du groupe est supérieure à l'espèce, (ii) les groupes sont présents dans plus de 5 sites pour les deux méthodes d'identification pour permettre une analyse de métacommunautés.

Sur ces groupes, les relations entre la diversité beta (c'est-à-dire les dissimilarités de composition et d'abondance des communautés entre paire de sites) et les distances environnementales (c'est-à-dire les dissimilarités environnementales entre paires de sites) et spatiales ont été analysées par des tests de Mantel. Pour chaque groupe, deux matrices de diversité beta avec l'indice de Jaccard basé sur des données de présence-absence ont été construites, une pour chaque méthode d'identification. La matrice des distances environnementales a été construite en utilisant les distances euclidiennes sur les 12 variables environnementales mesurées pendant l'échantillonnage et le régime hydrologique de chaque site (pérenne ou intermittent) et après normalisation des variables (moyenne = 0, SD = 1, Cañedo-Argüelles et al., 2015). Trois matrices spatiales ont été utilisées : une matrice de distances euclidiennes construite à partir des coordonnées GPS des sites deux à deux avec la fonction *vegdist* du package *vegan* (Oksanen et al., 2019) ; une matrice de distances réseau construite à partir des distances le long du réseau entre chaque paire de sites calculées sur le réseau RHT (Pella et al., 2012) avec QGIS 2.16 v3.10 (*postres*, *postgis* et *pgrouting*) ; une matrice de distances intégrant la fragmentation entre chaque paire de sites échantillonnés calculée à partir de la somme du nombre de tronçons asséchés et du nombre de structures anthropiques entre chaque paire de sites. Les détails pour la construction de cette matrice sont disponibles dans l'article 1. Les relations entre la diversité beta et les quatre matrices de distance ont été quantifiées et testées à l'aide de tests de Mantel (Legendre & Legendre, 2012). Pour les distances environnementales, la procédure MSR (randomisation spectrale de Moran) a été utilisée pour corriger l'autocorrélation spatiale potentielle qui peut altérer la robustesse des tests de Mantel (Crabot et al., 2019).

3.3 Résultats

3.3.1 Identification moléculaire et identification morphologique

Le séquençage a produit un total de 4 862 173 de lectures brutes (moyenne \pm SD : 226 094 \pm 71 120) dont 4 525 642 lectures conservées après l'étape filtrant les séquences de bonne qualité (207 107 \pm 68 600). A la suite du traitement bio-informatique, 517 OTUs ont été définis. Une affiliation taxinomique a pu être attribuée à 250 OTUs dont 246 avec une affiliation à un taxon invertébré. Ces 246 OTUs représentaient 48 taxa uniques (plusieurs OTUs peuvent avoir la même

3. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés - Étude exploratoire sur le bassin versant de l'Albarine (article 4)

affiliation) et trois classes (Figure 3.1). Une majorité des OTUs appartenait aux ordres de la classe des insectes (presque 75%), puis aux ordres de la classe des clitellates. Concernant les taxa n'ayant pas d'affiliation par l'intermédiaire de BOLD, les séquences ont été comparées à la base de données NCBI en utilisant l'outil d'alignement blastn et correspondaient à des bactéries (14,6%), des eukaryotes (81,6% dont 10,8% à des arthropodes, 2,2% à des annélides, 7,1% à des diatomées, 10% à des algues vertes, 62% à des eukaryotes non déterminés et les pourcentages restants à des protozoaires, de algues rouges, des cnidaires ou des Chordés) et des séquences non identifiées (3,7%).

Au total, 5 487 individus ont été collectés sur les 20 sites échantillonnés. En moyenne, un site contenait 274 ± 400 individus. Un total de 91 taxa ont été identifiés morphologiquement à différentes résolutions taxinomiques (de l'espèce à la classe) avec en moyenne 16 ± 8 taxa par site. Six classes d'invertébrés étaient représentées avec une majorité de taxa appartenant aux ordres de la classe des insectes (Figure 3.1).

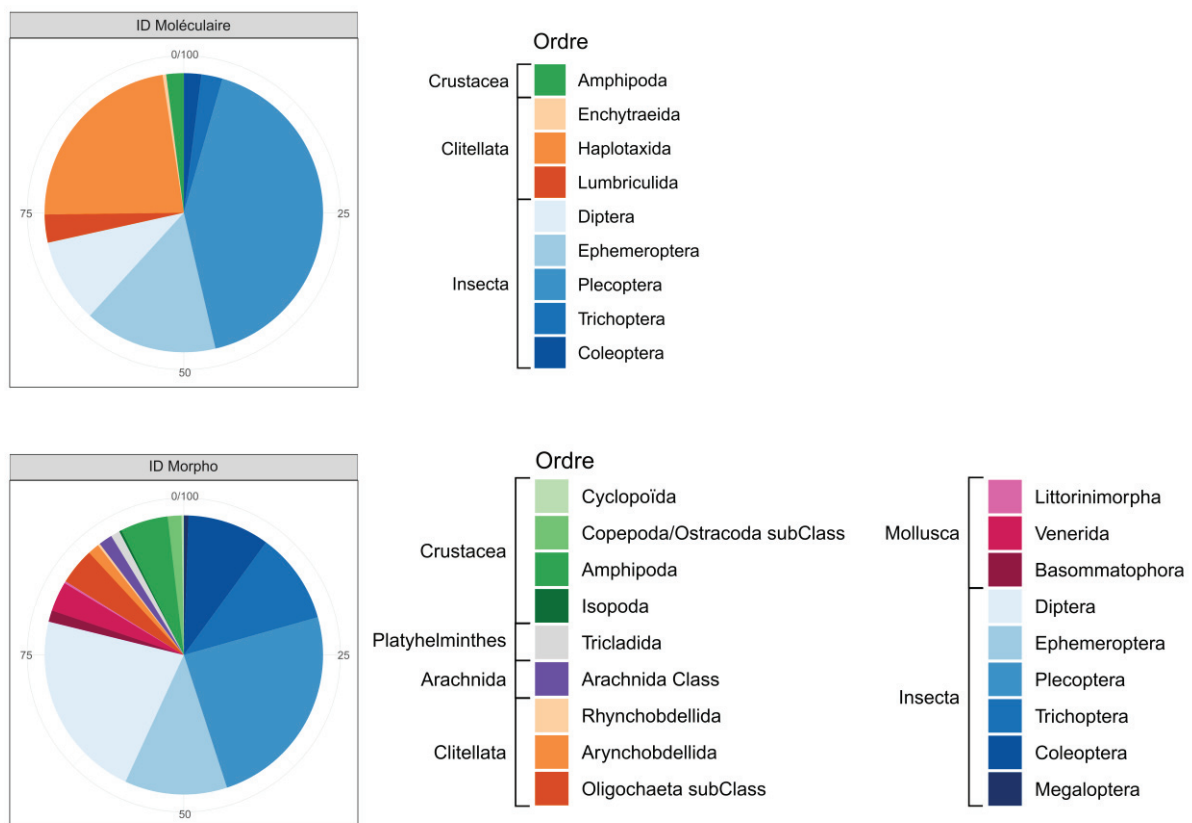


Figure 3.1 : Pourcentage d'OTUs obtenus par identification moléculaire (haut) et de taxa obtenus par identification morphologique (bas) pour différents ordres d'invertébrés (résolution taxinomique précisée si non à l'ordre).

3.3.2 Groupes d'invertébrés abondants et fréquents : sélection des groupes pour l'analyse

En prenant en compte tous les sites, les communautés étaient dominées en terme d'abondance par cinq taxa et de fréquence pas cinq taxa (Tableau 3.1). Quatre des 5 taxa les plus abondants étaient aussi les plus fréquents. Lorsque l'on s'intéresse aux sites pérennes et intermittents séparément, on remarque des différences entre les deux types de sites, que ce soit pour les taxa abondants ou fréquents, même si certains taxa sont représentés dans les deux types de sites. Les taxa abondants et/ou fréquents dans les sites de tronçons intermittents sont tous des taxa connus pour leur caractère résistant (Nemouridae, Taeniopterygidae, Chironomidae, Oligochaeta) et/ou résilient (Gammaridae, Simuliidae, Chironomidae) face aux assèchements des rivières (Khoo, 1968 ; Stubbington et al., 2018 ; Stubbington et al., 2009).

Tableau 3.1. Taxa les plus abondants et fréquents dans les échantillons utilisés pour l'identification morphologique pour les sites pérennes, les sites intermittents et l'ensemble des sites. Les taxa en italique sont ceux sélectionnés pour les analyses. Excepté pour les Oligochaeta qui représentent une sous-classe de Clitellata, les taxa présentés sont des familles d'invertébrés.

	Taxa avec une abondance > 10%		Taxa présents dans plus de 70% sites		
	Taxa	Abondance (%)	Taxa	Fréquence (% de sites)	
Sites pérennes			<i>Oligochaeta</i>	100	
			Elmidae	90	
			<i>Gammaridae</i>	90	
			<i>Nemouridae</i>	90	
			<i>Simuliidae</i>	90	
		<i>Chironomidae</i>	21,21	Leptophlebiidae	80
		<i>Simuliidae</i>	17,44	<i>Chironomidae</i>	70
		Elmidae	13,27	Heptageniidae	70
Sites intermittents		<i>Nemouridae</i>	12,83	Taeniopterygidae	70
		<i>Oligochaeta</i>	20,80		
		<i>Gammaridae</i>	20,09	<i>Nemouridae</i>	90
		<i>Simuliidae</i>	12,41	<i>Chironomidae</i>	70
Tous les sites		<i>Chironomidae</i>	11,33	Taeniopterygidae	70
		<i>Chironomidae</i>	18,70	<i>Nemouridae</i>	90
		<i>Simuliidae</i>	16,17	<i>Oligochaeta</i>	80
		<i>Nemouridae</i>	11,66	<i>Chironomidae</i>	70
		<i>Oligochaeta</i>	11,52	<i>Simuliidae</i>	70
	Elmidae	10,01	Taeniopterygidae	70	

3. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés -
Étude exploratoire sur le bassin versant de l'Albarine (article 4)

Parmi tous les groupes définis comme abondants et/ou fréquents dans les sites échantillonnés, nous avons sélectionné 5 taxa suivant les critères définis en matériel et méthodes. Le Tableau 3.2 résume les informations concernant ces 5 taxa. On remarque que les groupes ont été détectés dans plus de sites en identification morphologique (sauf pour les Oligochaeta) mais que le nombre d'OTUs détectés est systématiquement plus élevé que le nombre de taxa (Tableau 3.2). Au niveau des sites, la détection d'un groupe par les deux méthodes d'identification était parfois cohérente (les deux méthodes détectent ou ne détectent pas le groupe) et parfois une des deux méthodes détectait le groupe tandis que l'autre non (Figure 3.2). Néanmoins, l'accès aux OTUs nous permet d'avoir une vision plus diversifiée de la variabilité de distribution au sein de la tête de bassin.

Tableau 3.2. Détails pour les groupes d'invertébrés sélectionnés pour les analyses de cette étude. Sont présentés la résolution taxinomique atteinte en identification morphologique, le nombre d'OTUs (pour l'identification moléculaire) et de taxa (pour l'identification morphologique) ainsi que le nombre de sites où le groupe a été détecté.

Groupe sélectionné	Résolution taxinomique	ID Moléculaire		ID Morpho	
		Nb OTUs	Nb de sites	Nb taxa	Nb de sites
Gammaridae	Genre	5	8	1	13
Simuliidae	Famille	9	7	1	14
Chironomidae	Sous-famille Tribu	14	10	5	14
Nemouridae	Genre	57	15	3	18
Oligochaeta	Sous-classe	65	19	1	16

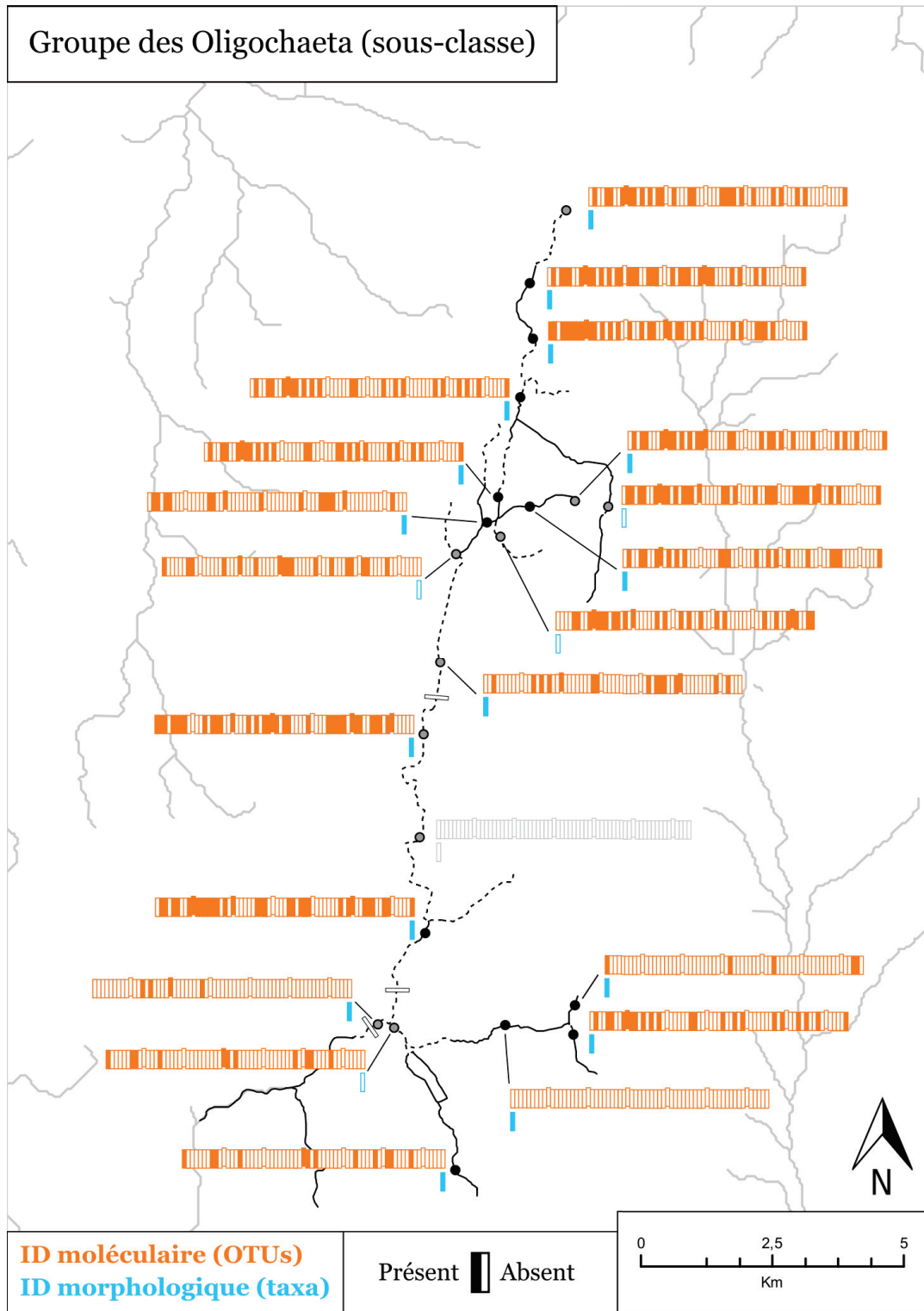


Figure 3.2 : Détection des OTUs de l'identification moléculaire (en orange) et des taxa de l'identification morphologique (en bleu) pour chaque site de prélèvement pour la sous-classe des Oligochètes. Une barre représente un OTU ou taxon. Lorsqu'une barre est pleine, l'OTU/taxon a été détecté tandis qu'une barre vide signifie qu'il ne l'a pas été. Tous les 10 OTUs/taxa, la barre est plus haute afin de mieux repérer leur placement. Le site avec les barres grisées est un site qui n'a pas pu être échantillonné. Les cercles gris et noirs correspondent aux sites de prélèvement intermittent et pérenne, respectivement. Les lignes pointillées et pleines correspondent aux sections intermittentes et pérennes de la tête de bassin, respectivement. Les structures anthropiques sont représentées par des rectangles blancs sur le réseau. Les cartes de distribution pour les autres groupes sélectionnés sont présentées en Annexe A1 de cet article.

3. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés - Étude exploratoire sur le bassin versant de l'Albarine (article 4)

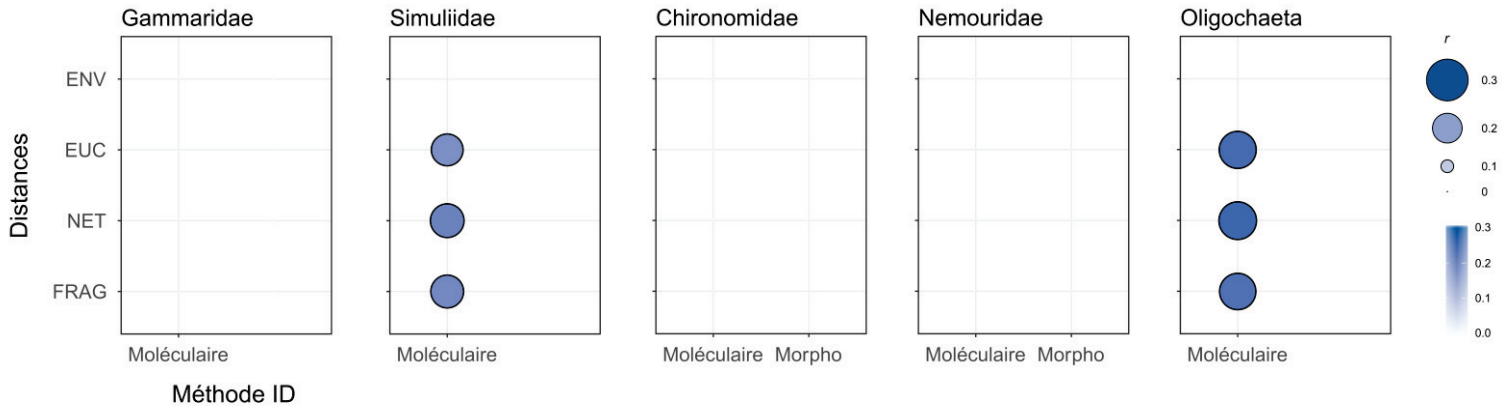


Figure 3.3 : Coefficient de corrélation (valeur de r représentée par la taille et la couleur du point) des résultats du test de Mantel pour chaque combinaison de groupe d'invertébrés sélectionné (indice de dissimilarité de Jaccard basé sur l'ensemble du groupe) et de matrices de distance pour les deux méthodes d'identifications. Seul les relations pour lesquelles la p-value est significative ($\leq 0,05$) sont présentées. ENV, distance environnementale ; EUC, distance euclidienne ; NET, distance réseau ; FRAG, distance basée sur le nombre de tronçons asséchés et de structures anthropiques entre paires de sites. Les valeurs exactes des r pour chaque test sont disponibles en Annexe 2.

3.3.3 Structure de la diversité beta

En identification morphologique, seulement deux groupes ont pu être utilisés pour effectuer des tests de Mantel (Chironomidae et Nemouridae) car ils étaient représentés par plusieurs taxa dans la tête de bassin versant. Pour ces deux groupes, aucune corrélation entre la diversité beta (distance de Jaccard) et les distances explicatives n'a pu être mise en évidence dans la métacommunauté en identification morphologique ou en identification moléculaire (Figure 3.3). Cependant, pour les 3 groupes restants pour lesquels l'analyse n'est pas envisageable via l'identification morphologique, des corrélations entre diversité beta et les 3 distances spatiales ont pu être mises en évidence chez deux d'entre eux, le groupe des Simuliidae et celui des Oligochaeta (Figure 3.3).

3.4 Discussion

Apport du moléculaire dans la structure des métacommunautés

Dans cette étude exploratoire, nous nous sommes intéressé·es à l'apport d'une identification moléculaire dans l'étude d'une métacommunauté de macro-invertébrés benthiques de tête de bassin versant comportant des tronçons intermittents, milieux dans lesquels les invertébrés abondants et ubiquistes représentent des groupes mal connus et par conséquent souvent identifiés à une résolution taxinomique très grossière en identification morphologique (famille ou sous-classe, Stubbington et al., 2018 ; Stubbington et al., 2019, article 1). Nos résultats ont démontré le potentiel de l'identification moléculaire pour mettre en évidence des patrons de distribution de la diversité beta chez des groupes qu'il n'aurait pas été possible d'étudier en identification morphologique. En effet, dans cette étude, les Oligochètes et les Simuliidae n'étaient représentés que par un unique taxon, empêchant ainsi une

quelconque analyse de diversité beta. Pour les deux groupes, la diversité beta était expliquée par les trois distances spatiales et non par les distances environnementales, suggérant un rôle clé de la dispersion dans la structure de leurs communautés. Les Oligochètes représentent des taxa généralistes (c'est-à-dire qu'ils supportent une large gamme de conditions environnementales) et présentent la particularité de survivre en milieu semi-aquatique, voire terrestre, leur permettant de disperser par voie terrestre entre les tronçons d'un bassin et de survivre dans les tronçons asséchés (Steward et al., 2017). Ils présentent donc un fort potentiel de colonisation du réseau entier. Cependant, ce potentiel ne peut être entièrement réalisé car ce groupe possède également une capacité de dispersion faible (Behrend et al., 2012 ; Palmer et al., 1996) qui serait d'autant plus limitée dans les têtes de bassin versant (Fagan, 2002) et leur dispersion est fortement impactée par les structures anthropiques tels que les seuils (Suriani et al., 2007). Les résultats obtenus mis en lien avec les caractéristiques du groupe suggèrent que les communautés d'Oligochètes sont structurées dans ce bassin par un mécanisme de limitation à la dispersion induite par leur capacité de dispersion faible couplée à la fragmentation du réseau et à son dendritisme. Les Simuliidae, quant à elles, représentent une famille de Diptères complexe avec de nombreuses espèces peu décrites et de nombreux traits biologiques et stratégies différentes. On sait que la dispersion aérienne passive est dominante pour l'adulte tandis que la dérive à l'état larvaire est prépondérante, les rendant résilientes aux assèchements avec une recolonisation forte des sites remis en eau par dérive de l'amont (Sarremejane et al., 2020). Cependant, ayant un mode de nutrition filtreur, les organismes filtrent les particules de l'eau pour se nourrir et ont besoin de courant pour se nourrir (Tachet et al., 2000). Elles ne survivraient donc pas en cas d'assèchement. Tout comme pour les Oligochètes, nos résultats montrent que la dispersion est un facteur important dans la structure des communautés de Simuliidae dans le bassin avec comme mécanismes une recolonisation des sites asséchés par dérive (avec un potentiel *mass effect*) associée à une limitation de la dispersion due aux assèchements qui y sont très présents et au dendritisme du réseau. Cette limitation empêche une homogénéisation des OTUs à l'ensemble de la tête de bassin versant.

Contrairement aux Oligochètes et Simuliidae, aucune relation entre diversité beta et distances environnementales ou spatiales n'a été mise en évidence pour les groupes des Gammaridae, Chironomidae et Nemouridae. Pour ces deux derniers groupes, l'identification moléculaire et l'identification morphologique ont amené aux mêmes conclusions. Si l'identification moléculaire ne permet pas d'affiner notre compréhension de la structure des communautés pour ces groupes, il est possible que la résolution taxinomique soit suffisante en identification morphologique (sous-famille/tribu et genre pour les Chironomidae et les Nemouridae, respectivement) pour apporter une information fiable et inférer les mécanismes structurant les communautés. Ici, aucun patron de distribution de la diversité beta n'a pu être mis en évidence pour ces groupes, suggérant que la dérive écologique (c'est-à-dire une structure régie par la stochasticité démographique, Hubbell, 2001) serait prédominante dans la structure des communautés de ces groupes à cette phase hydrologique (après la

3. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés - Étude exploratoire sur le bassin versant de l'Albarine (article 4)

remise en eau). Par exemple, pour les Chironomidae, la dispersion passive et la dérive des larves sont deux formes de résilience prédominantes dans le groupe comme pour les Simuliidae (Sarremejane et al., 2020). Par contre, beaucoup d'espèces de Chironomidae sont tolérantes et résistantes aux assèchements et certaines sont généralistes (Carew et al., 2007). Ainsi, la dérive et la dispersion aérienne passives ont pu permettre aux individus de la famille des Chironomidae de s'installer dans de nombreux sites et de coloniser l'ensemble du réseau. Alternativement, il est possible que les variables environnementales mesurées représentent mal celles déterminantes pour la présence d'un groupe à un site donné, malgré la prise en compte de 12 variables environnementales pertinentes pour la description des niches des invertébrés (Brown & Swan, 2010 ; Cañedo-Argüelles et al., 2015 ; Datry et al., 2016). P.ex., pour les Chironomidae, bien que certains paramètres déterminants pour la présence de certaines espèces aient été mesurés (% d'oxygène dissous, conductivité), d'autres n'ont pas été mesurés, notamment celles liées aux perturbations anthropiques telles que la pollution chimique et organique, l'utilisation des sols ou la modification du lit du cours d'eau (Jacobsen, 1998 ; Lencioni et al., 2012 ; Puntí et al., 2009).

Aller plus loin dans la compréhension des processus structurant les métacommunautés : axes d'amélioration

Cette étude a permis de montrer que l'identification moléculaire avait le potentiel de mettre en évidence des patrons de distribution de diversité beta chez des groupes mal connus mais qui sont cependant abondants et ubiquistes dans les rivières. Notre étude reste cependant exploratoire et appelle à d'autres études pour tester et améliorer cette approche. Tout d'abord, il serait intéressant de pouvoir, comme dans la première partie de cette thèse, tester la méthode sur plusieurs têtes de bassin versant (article 1) et à plusieurs périodes d'échantillonnage (article 2) afin de voir si les mécanismes structurant ces groupes d'invertébrés varient entre bassins versants et dans le temps.

Avant de tester cette approche à large échelle spatiale ou temporelle, des améliorations méthodologiques sur l'identification moléculaire sont à mon avis nécessaires. Comme dit en Introduction, l'identification moléculaire par metabarcoding présente encore des limites. Dans cette étude, certains groupes n'ont pas été détectés en identification moléculaire sur des sites où ils étaient présents car confirmés par l'identification morphologique (Figure 3.2, Annexe A1). Une faiblesse de notre étude est que les échantillons ayant servi pour l'identification morphologique et l'identification moléculaire étaient différents, ne permettant pas de faire une comparaison stricte des deux méthodes d'identification et de déterminer l'origine des différences observées par rapport à la détection des taxa. Ainsi, si un groupe n'a pas été détecté en identification moléculaire sur un site, il est impossible de savoir si cela est dû à une composition différente entre les deux communautés échantillonnées ou à des biais inhérents à l'identification moléculaire (p.ex., biais de PCR, profondeur de séquençage, matrice ADN). Il serait ainsi intéressant d'étudier la structure des communautés avec des identifications

morphologiques et moléculaires à partir d'un même échantillon, ce qui permettrait (i) de tester finement la complémentarité des deux approches sur les groupes abondants mais aussi sur les communautés entières et (ii) d'évaluer la fiabilité de l'identification moléculaire pour étudier les métacommunautés d'invertébrés de rivières. Pour ce faire, l'idéal serait d'avoir un échantillon pour un site donné où l'éthanol serait prélevé pour l'extraction d'ADNet et les organismes seraient identifiés morphologiquement. Cela n'a pas pu être fait dans cette étude car, à l'origine, les échantillons pour l'identification moléculaire auraient dû être utilisés pour extraire de l'ADN bulk et de l'ADNet afin d'avoir deux niveaux de comparaisons avec l'identification morphologique. L'ADN bulk n'a pu être extrait pour des raisons de développement méthodologique conséquent (p.ex. gérer l'extraction d'ADN à partir d'échantillons récoltés *in situ* représentant un volume conséquent et contenant potentiellement beaucoup d'inhibiteurs) qui n'était pas envisageable pour des raisons de calendrier.

Conclusion

Grâce à l'identification moléculaire, une nouvelle façon d'aborder l'étude de la biodiversité a été proposée avec la détermination d'OTUs correspondants à un grain taxinomique équivalent à celui de l'espèce, voire plus fin (Elbrecht et al., 2018). Dans les études de métacommunautés où l'identification morphologique est la norme, cette approche permettra une meilleure résolution taxinomique pour affiner la description des patrons de distribution et des mécanismes sous-jacents à cette distribution comme l'a montrée notre étude et d'autres (Martin et al., 2016). Bien que l'étude des métacommunautés avec une identification moléculaire aux OTUs demande encore à être testée avec plusieurs axes d'améliorations et d'ajustements, c'est une méthode qui semble prometteuse pour l'étude des métacommunautés. En effet, l'identification moléculaire pourra (i) permettre de décrire les patrons de distribution plus finement chez des groupes où l'identification morphologique est un frein et (ii) en permettant une meilleure discrimination des groupes entre eux, permettre d'étudier les groupes d'invertébrés indépendamment les uns des autres et ainsi de mettre en évidence l'ensemble des mécanismes complexes structurant les métacommunautés plus finement et plus précisément que lorsque les communautés entières sont prises en compte.

3.5 Références

- Alp, M., Keller, I., Westram, A.M., Robinson, C.T., 2012. How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology* 57, 969-981.
- Altermatt, F., 2013. Diversity in riverine metacommunities: a network perspective. *Aquatic Ecology* 47, 365-377.
- Andujar, C., Arribas, P., Gray, C., Bruce, C., Woodward, G., Yu, D.W., Vogler, A.P., 2018. Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Mol Ecol* 27, 146-166.
- Behrend, R., Takeda, A., Gomes, L., Fernandes, S., 2012. Using oligochaeta assemblages as an indicator of environmental changes. *Brazilian Journal of Biology* 72, 873-844.
- Benda, L.E.E., Poff, N.L., Miller, D., Dunne, T., Reeves, G., Pess, G., Pollock, M., 2004. The Network Dynamics Hypothesis: How Channel Networks Structure Riverine Habitats. *BioScience* 54.
- Brown, B.L., Swan, C.M., 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. *J Anim Ecol* 79, 571-580.
- Cañedo-Argüelles, M., Boersma, K.S., Bogan, M.T., Olden, J.D., Phillipsen, I., Schriever, T.A., Lytle, D.A., Ladle, R., 2015. Dispersal strength determines meta-community structure in a dendritic riverine network. *Journal of Biogeography* 42, 778-790.
- Carew, M.E., Pettigrove, V., Cox, R.L., Hoffmann, A.A., 2007. The response of Chironomidae to sediment pollution and other environmental characteristics in urban wetlands. *Freshwater Biology* 52, 2444-2462.
- Crabot, J., Clappe, S., Dray, S., Datry, T., McCrea, R., 2019. Testing the Mantel statistic with a spatially- constrained permutation procedure. *Methods in Ecology and Evolution* 10, 532-540.
- Datry, T., Melo, A.S., Moya, N., Zubieta, J., De la Barra, E., Oberdorff, T., 2016. Metacommunity patterns across three Neotropical catchments with varying environmental harshness. *Freshwater Biology* 61, 277-292.
- Elbrecht, V., Vamos, E.E., Meissner, K., Aroviita, J., Leese, F., Yu, D., 2017. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution* 8, 1265-1275.
- Elbrecht, V., Vamos, E.E., Steinke, D., Leese, F., 2018. Estimating intraspecific genetic diversity from community DNA metabarcoding data. *PeerJ* 6, e4644.
- Fagan, W.F., 2002. CONNECTIVITY, FRAGMENTATION, AND EXTINCTION RISK IN DENDRITIC METAPOPOPULATIONS. *Ecology* 83, 3243-3249.
- Fuller, M.R., Doyle, M.W., Strayer, D.L., 2015. Causes and consequences of habitat fragmentation in river networks. *Ann N Y Acad Sci* 1355, 31-51.

Godoy, B.S., Faria, A.P.J., Juen, L., Sara, L., Oliveira, L.G., 2019. Taxonomic sufficiency and effects of environmental and spatial drivers on aquatic insect community. *Ecological Indicators* 107.

Hebert, P.D., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc Biol Sci* 270, 313-321.

Heino, J., 2014. Taxonomic surrogacy, numerical resolution and responses of stream macroinvertebrate communities to ecological gradients: Are the inferences transferable among regions? *Ecological Indicators* 36, 186-194.

Heino, J., Soininen, J., 2007. Are higher taxa adequate surrogates for species-level assemblage patterns and species richness in stream organisms? *Biological Conservation* 137, 78-89.

Hubbell, S.P., 2001. *The Unified Neutral Theory of Biodiversity and Biogeography* Princeton University Press.

Jacobsen, D., 1998. The effect of organic pollution on the macroinvertebrate fauna of Ecuadorian highland streams. *Fundamental and Applied Limnology* 143, 179-195.

Khoo, S.G., 1968. Experimental studies on diapause in stoneflies. II. Eggs of *Diura bicaudata* (L.). *Proceedings of the Royal Entomological Society of London. Series A, General Entomology* 43, 49-56.

Legendre, P., Legendre, L., 2012. *Numerical ecology*, 3 edn. English edn. Elsevier, Amsterdam.

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

Losos, J.B., 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol Lett* 11, 995-1003.

Martin, G.K., Adamowicz, S.J., Cottenie, K., 2016. Taxonomic resolution based on DNA barcoding affects environmental signal in metacommunity structure. *Freshwater Science* 35, 701-711.

Martin, M., 2014. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 7, 10-12.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin P.R, O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2019. *vegan: Community Ecology Package*. R package version 2.5-5.

Palmer, M., A., Allan, D., Butman, C.A., 1996. Dispersal as a regional process affecting the local dynamics of marine and stream benthic invertebrates. *TREE* 11, 322-325.

Pella, H., Lejot, J., Lamouroux, N., Snelder, T., 2012. Le réseau hydrographique théorique (RHT) français et ses attributs environnementaux. *Géomorphologie : relief, processus, environnement* 18, 317-336.

Pompanon, F., Coissac, É., Taberlet, P., 2011. Metabarcoding, a new way of analysing biodiversity.

Prunier, J.G., Dubut, V., Loot, G., Tudesque, L., Blanchet, S., 2018. The relative contribution of river network structure and anthropogenic stressors to spatial patterns of genetic diversity in two freshwater fishes: A multiple-stressors approach. *Freshwater Biology* 63, 6-21.

3. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés -
Étude exploratoire sur le bassin versant de l'Albarine (article 4)

Puntí, T., Rieradevall, M., Prat, N., 2009. Environmental factors, spatial variation, and specific requirements of Chironomidae in Mediterranean reference streams. *Journal of the North American Benthological Society* 28, 247-265.

Ratnasingham, S., Hebert, P.D.N., 2007. BOLD : The Barcode of Life Data System. *Molecular Ecology Notes* 7, 355-364.

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahe, F., 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584.

Sarremejane, R., Cid, N., Datry, T., Stubbington, R., Alp, M., Cañedo-Argüelles, M., Cordero-Rivera, A., Csabai, Z., Gutiérrez-Cánovas, C., Heino, J., Forcellini, M., Millán, A., Paillex, A., Pařil, P., Polářek, M., de Figueroa, J.M.T., Usseglio-Polatera, P., Zamora-Muñoz, C., Bonada, N., 2020. DISPERSE: A trait database to assess the dispersal potential of aquatic macroinvertebrates. In review.

Stein, E.D., White, B.P., Mazor, R.D., Jackson, J.K., Battle, J.M., Miller, P.E., Pilgrim, E.M., Sweeney, B.W., 2014. Does DNA barcoding improve performance of traditional stream bioassessment metrics? *Freshwater Science* 33, 302-311.

Steward, A.L., Langhans, S.D., Corti, R., Datry, T., 2017. Chapter 4.4 - The Biota of Intermittent Rivers and Ephemeral Streams: Terrestrial AND Semiaquatic Invertebrates, In *Intermittent Rivers and Ephemeral Streams*. eds T. Datry, N. Bonada, A. Boulton, pp. 245-271. Academic Press.

Stubbington, R., Chadd, R., Cid, N., Csabai, Z., Miliřa, M., Morais, M., Munné, A., Pařil, P., Peřić, V., Tziortzis, I., Verdonschot, R.C.M., Datry, T., 2018. Biomonitoring of intermittent rivers and ephemeral streams in Europe: Current practice and priorities to enhance ecological status assessments. *Science of The Total Environment* 618, 1096-1113.

Stubbington, R., Greenwood, A.M., Wood, P.J., Armitage, P.D., Gunn, J., Robertson, A.L., 2009. The response of perennial and temporary headwater stream invertebrate communities to hydrological extremes. *Hydrobiologia* 630, 299-312.

Stubbington, R., Paillex, A., England, J., Barthès, A., Bouchez, A., Rimet, F., Sánchez-Montoya, M.M., Westwood, C.G., Datry, T., 2019. A comparison of biotic groups as dry-phase indicators of ecological quality in intermittent rivers and ephemeral streams. *Ecological Indicators* 97, 165-174.

Suriani, A.L., FRANÇA , R.S., PAMPLIN , P.A.Z., MARCHESE, M., LUCCA, J.V., ROCHA , O., 2007. Species richness and distribution of oligochaetes in six reservoirs on Middle and Low Tietê River (SP, Brazil). 19 4.

Sweeney, B.W., Battle, J.M., Jackson, J.K., Dapkey, T., 2011. Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *Journal of the North American Benthological Society* 30, 195-216.

Tachet, H., Richoux, P., Bournaud, M., Usseglio-Polatera, P., 2000. Invertébrés d'eau douce: systématique, biologie, écologie.

Tonkin, J.D., Heino, J., Altermatt, F., 2018. Metacommunities in river networks: The importance of network structure and connectivity on patterns and processes. *Freshwater Biology* 63, 1-5.

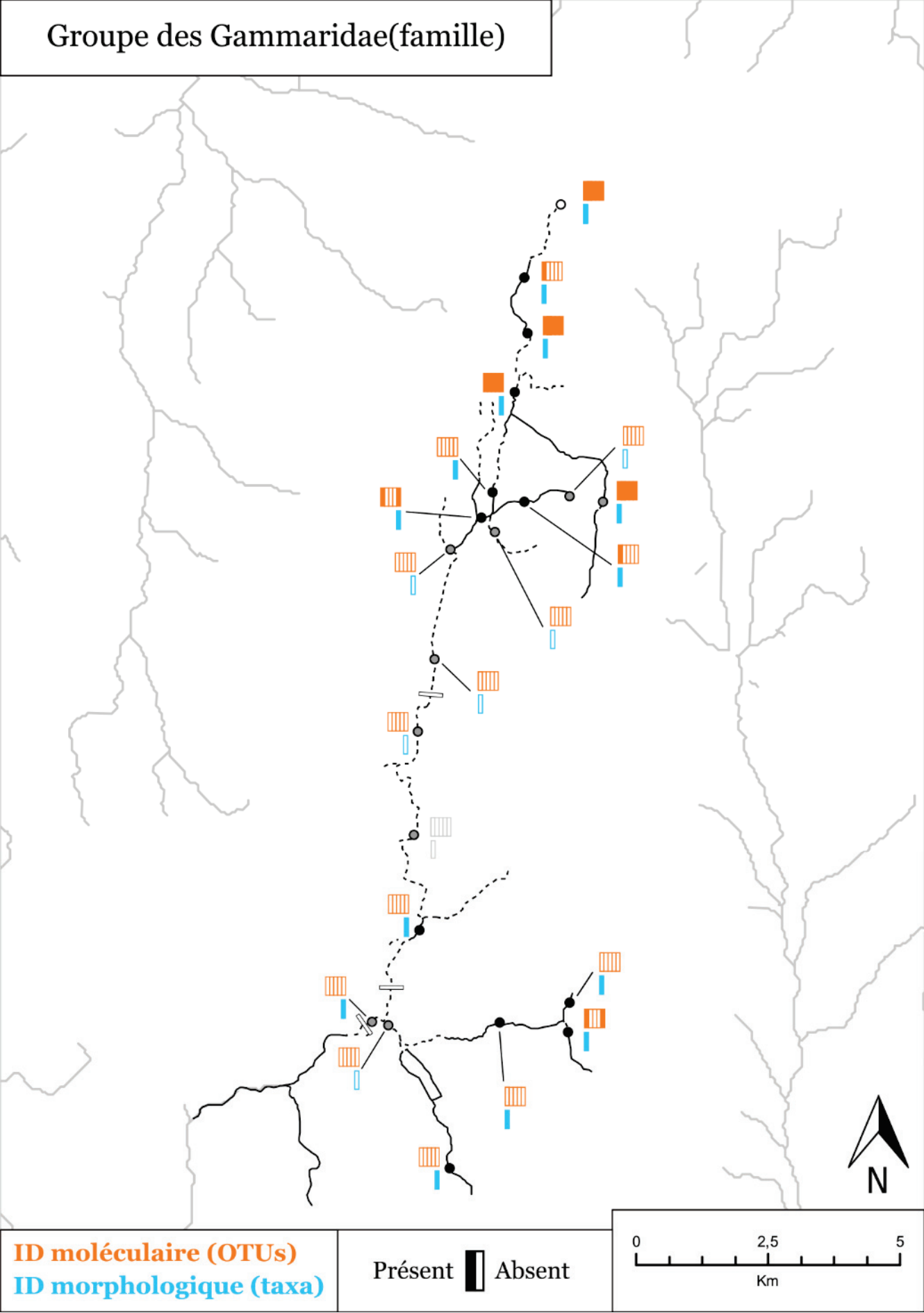
Vamos, E., Elbrecht, V., Leese, F., 2017. Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics* 1.

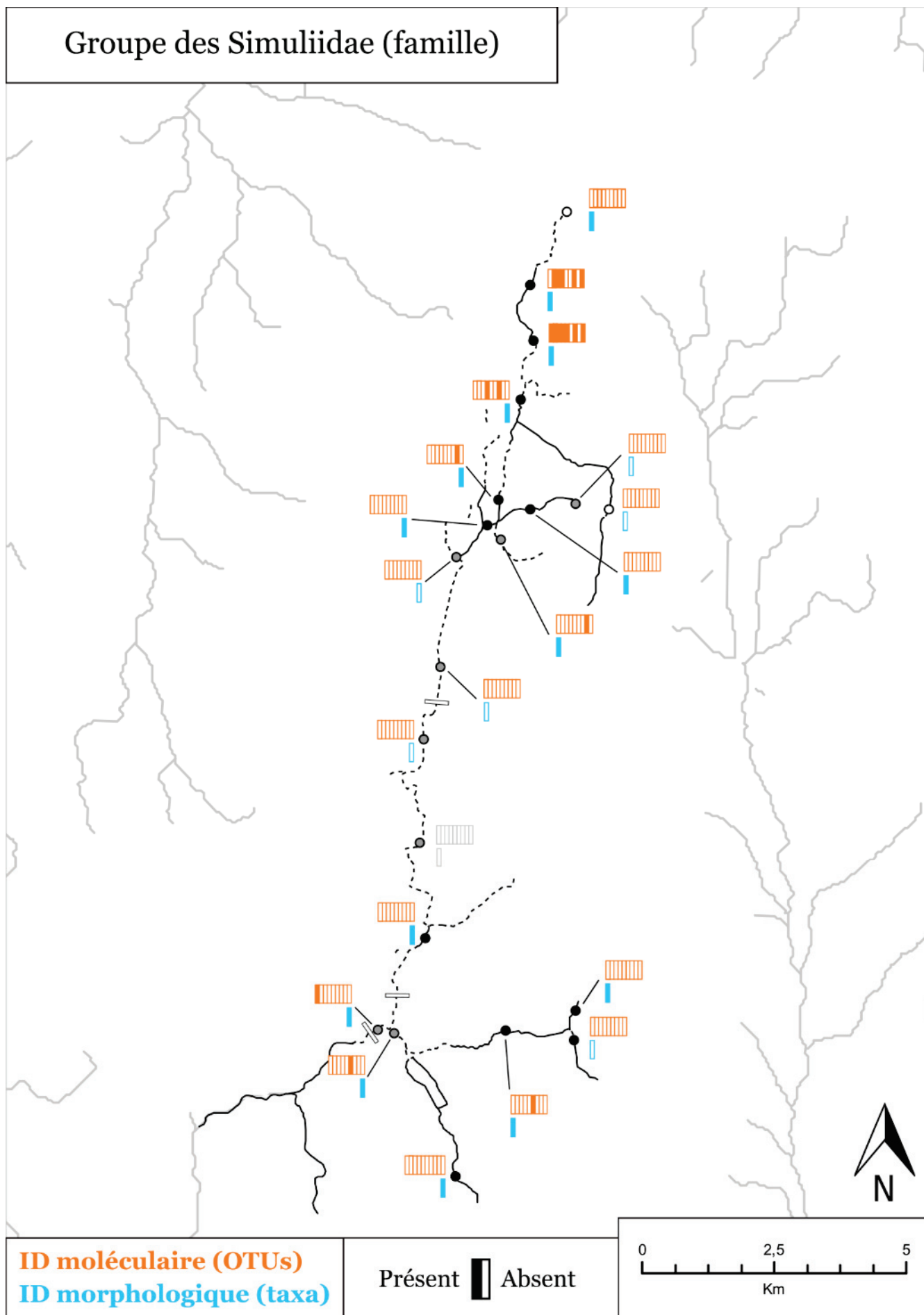
Vellend, M., 2010. Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology* 85, 183-206.

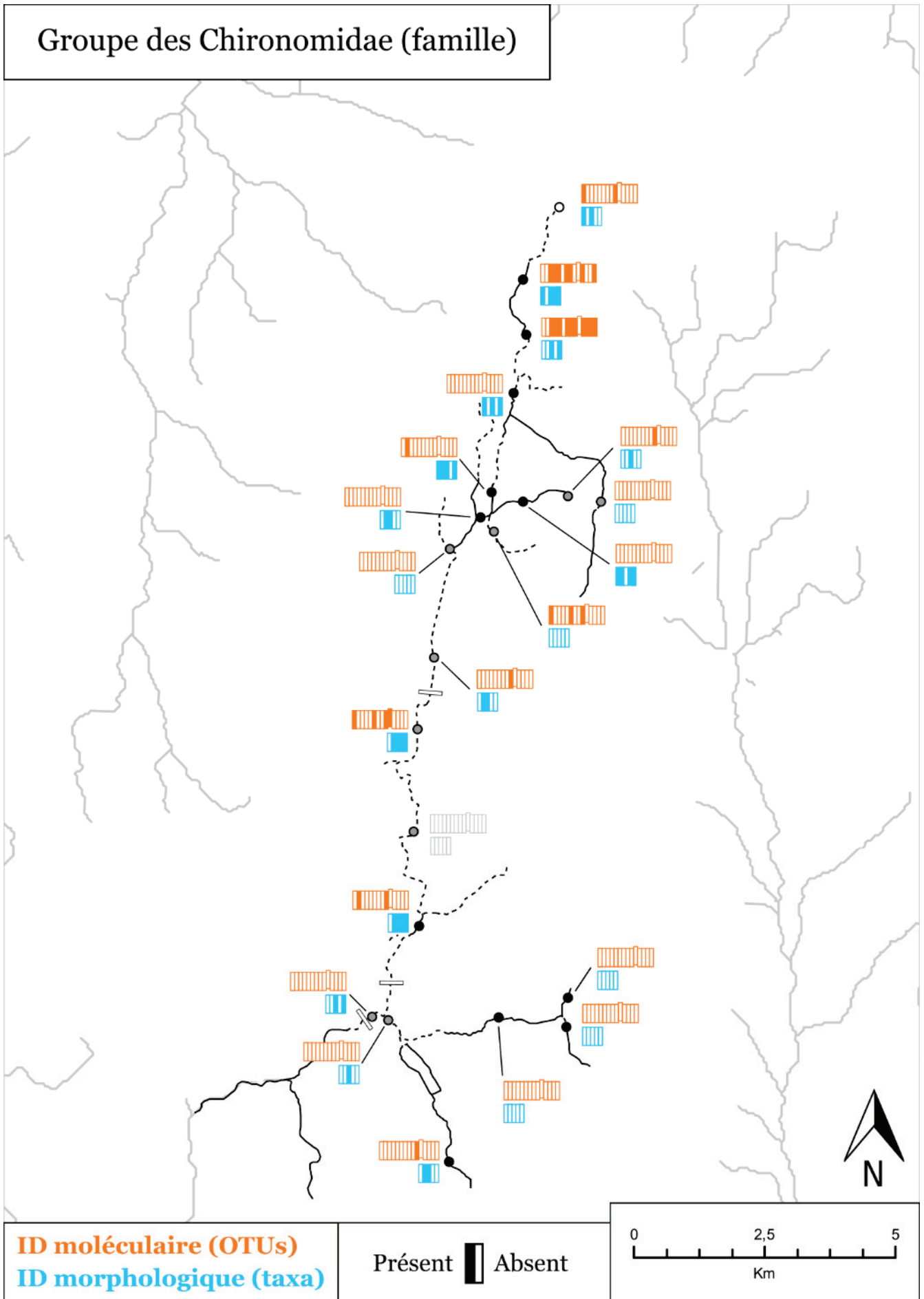
3. Apport de l'identification moléculaire dans la compréhension de la structure des métacommunautés -
Étude exploratoire sur le bassin versant de l'Albarine (article 4)

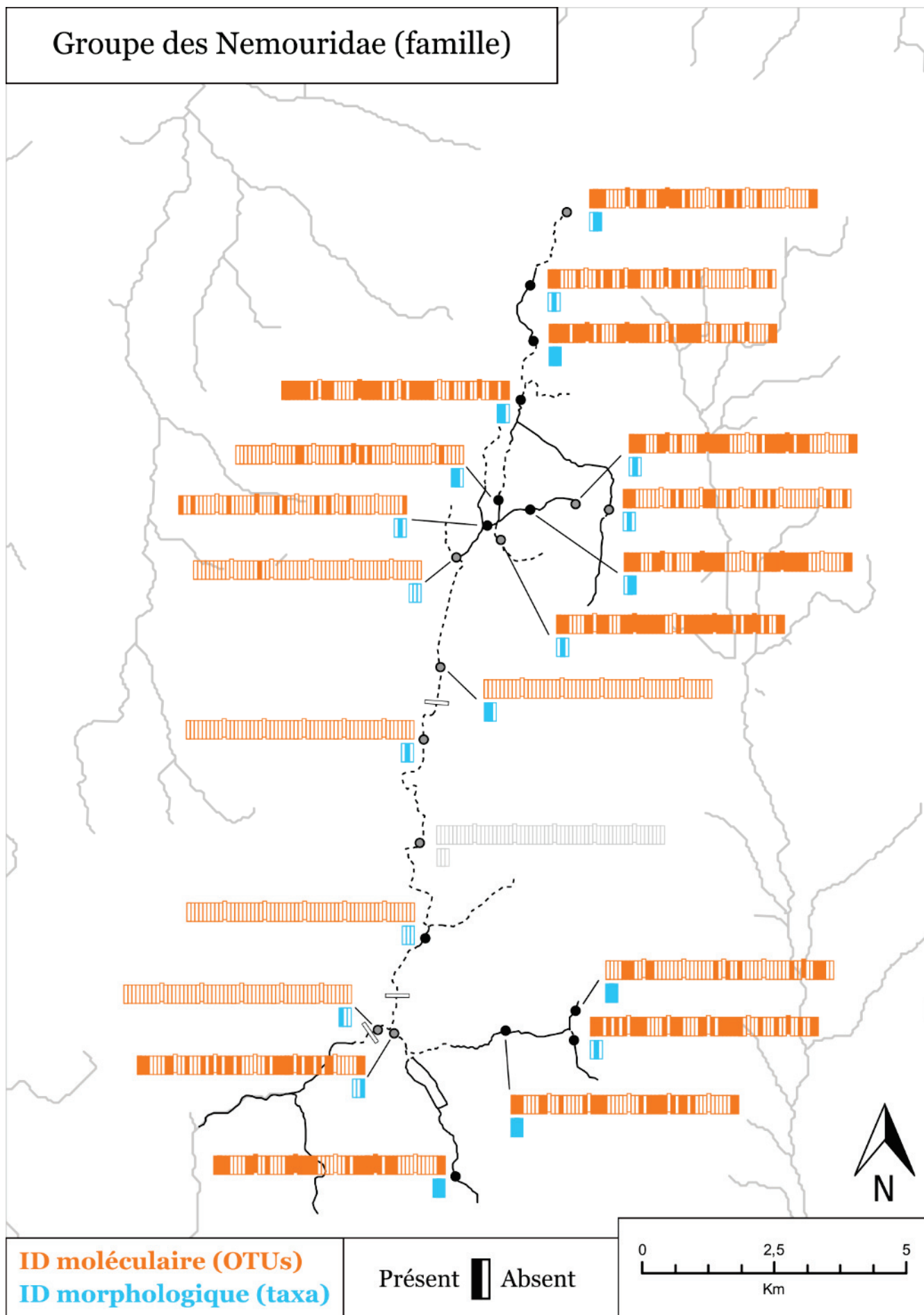
3.6 Annexes

Annexe A1 : Détection des OTUs de l'identification moléculaire (en orange) et des taxa de l'identification morphologique (en bleu) pour chaque site de prélèvement pour les différents groupes d'invertébrés étudiés lors de cet article. Une barre représente un OTU ou taxon. Lorsqu'une barre est pleine, l'OTU/taxon a été détecté tandis qu'une barre vide signifie qu'il ne l'a pas été. Le site avec les barres grisées est un site qui n'a pas pu être échantillonné. Les cercles gris et noirs correspondent aux sites de prélèvement intermittent et pérenne, respectivement. Les lignes pointillées et pleines correspondent aux sections intermittentes et pérennes de la tête de bassin, respectivement. Les structures anthropiques sont représentées par des rectangles blancs sur le réseau.









Annexe A2. Tableau des valeurs des coefficients de corrélation des tests de Mantel basé sur l'indice de dissimilarité de Jaccard effectués pour déterminer les processus structurant les métacommunautés. ENV, environnemental ; EUC, terrestre ; NET, réseau ; FRAG, nombre de sections asséchées et de structures anthropiques. Les relations significatives sont indiquées en gras (valeur $p \leq 0,05$ *).

Identification moléculaire						
Matrice de distances	Gammaridae	Simuliidae	Chironomidae	Nemouridae	Oligochaeta	
ENV	0	0	0	0	0	0
EUC	0	0,173*	0	0	0	0,237*
NET	0	0,195*	0	0	0	0,245*
FRAG	0	0,185*	0	0	0	0,228*
Identification morphologique						
Matrice de distances	Gammaridae	Simuliidae	Chironomidae	Nemouridae	Oligochaeta	
ENV	NA	NA	0	0	0	NA
EUC	NA	NA	0	0	0	NA
NET	NA	NA	0	0	0	NA
FRAG	NA	NA	0	0	0	NA

Chapitre 4 - Synthèse et Perspectives



Illustration : Panel d'invertébrés benthiques de rivière

« Je viens juste de comparer entre elles des définitions de l'espèce [...], il est vraiment comique de voir à quel point peuvent être diverses les idées qu'ont en tête les naturalistes lorsqu'ils parlent de l'espèce : chez certains, la ressemblance est tout, et la descendance de parents communs compte pour peu de choses ; chez d'autres, la ressemblance ne compte pratiquement pour rien, et la création est l'idée dominante ; pour d'autres encore, la descendance est la notion-clé ; chez certains, la stérilité est un test infallible, tandis que chez d'autres, cela ne vaut pas un sou. **Tout cela vient, je suppose, de ce que l'on essaie de définir l'indéfinissable** »

Darwin

Le concept de métacommunauté a permis de prendre en compte la dispersion de manière explicite dans l'étude de la structure des communautés. Il a proposé un cadre de travail pertinent pour les écosystèmes de rivières, réseau dendritique où les communautés sont interconnectées et où la structure des habitats et la dispersion sont spatialement contraintes par le courant de l'eau d'amont en aval et le dynamisme de son hydrologie. A ce jour, des lacunes subsistent dans la compréhension de la distribution de la biodiversité des têtes de bassin versant qui sont relativement négligées, notamment par la Directive Cadre sur l'Eau (DCE) de l'Union Européenne. Pourtant, les têtes de bassin versant présentent une biodiversité unique, sont extrêmement sensibles aux perturbations anthropiques et ont un rôle essentiel pour l'intégrité de tout l'écosystème que représente un réseau hydrographique. Bien que la considération de l'ensemble des rivières de têtes de bassin versant semble difficile du fait d'un compromis entre faisabilité et coût économique et logistique (Oenema et al., 2009), atteindre un bon état écologique pour les rivières en aval semble voué à l'échec sans les considérer. Baattrup et al. (2018) ont ainsi proposé un outil de prise de décision permettant de définir au Danemark les rivières de têtes de bassin versant pouvant répondre aux objectifs de bonne qualité actuellement fixés par la DCE afin de les inclure dès aujourd'hui dans les politiques de gestion. Bien qu'elle permette de préserver des rivières qui seraient ignorées sinon, cette méthode n'est qu'une étape transitoire en attendant de pouvoir clairement définir les besoins des rivières de têtes de bassin versant et de les gérer convenablement.

S'intéresser à la description des dynamiques en place dans les têtes de bassin versant est donc nécessaire pour améliorer notre compréhension des processus en jeu dans la structure de leurs communautés et pour guider les mesures de gestion et de conservation de ces milieux. Cette thèse souhaite participer à l'effort grandissant en faveur de leur conservation dont le programme LIFE « Ruisseaux de têtes de bassin et faune patrimoniale associée » a été l'initiateur en France (LIFE Ruisseaux, 2004-2009). Ce travail visait à améliorer la compréhension de l'organisation de la biodiversité dans les têtes de bassin versant grâce au cadre de travail qu'offre le concept de métacommunauté. Dans une première partie, il s'est intéressé à décrire la structure spatiale et la dynamique temporelle des métacommunautés de macro-invertébrés benthiques en y intégrant la fragmentation liée aux assèchements et aux structures anthropiques. Les résultats ont montré dans ces systèmes (i) que les processus basés sur la dispersion étaient majoritairement à l'œuvre dans les dynamiques de métacommunautés (article 1 et 2), à travers les événements d'extinction et de recolonisation des tronçons intermittents à partir des tronçons pérennes qui servent de refuges et par la limitation de la dispersion induite par la fragmentation et le dendritisme inhérent à la structure de ces réseaux, (ii) que les assèchements jouaient un rôle primordial dans la structure des communautés à l'échelle locale à travers un filtre environnemental fort et que les communautés des têtes de bassin versant y étaient très sensibles (article 1), (iii) que la contribution respective de ces processus dépendaient de la phase hydrologique et variaient dans le temps (article 2) et (iv) que les réponses des

communautés étaient aussi dépendantes du contexte et de la nature de la fragmentation ainsi que des modes de dispersion des espèces (articles 1 et 2).

Dans une deuxième partie, un outil moléculaire a été développé puis testé avec pour objectif de mieux décrire la biodiversité d'invertébrés benthiques des têtes de bassin versant et, en conséquence, leurs patrons de distribution utilisés dans les analyses de métacommunautés. Les résultats ont montré (i) que la méthode de capture par sondes ARN pour l'enrichissement de l'ADN était plus fiable et robuste que la méthode par amplification traditionnellement utilisée et permettait de mieux décrire les communautés qualitativement et quantitativement (article 3), (ii) que l'ADN extrait à partir de l'éthanol de conservation était aussi efficace que l'ADN extrait à partir de l'homogénat d'organismes dans la détection des taxa (article 3), (iii) que l'ADN extrait à partir de l'éthanol pouvait être utilisé sur des échantillons représentatifs des études *in situ* de métacommunautés de macro-invertébrés benthiques (article 4) et (iv) qu'il permettait de mettre en évidence des patrons de distribution indécélables en identification morphologique chez certains groupes invertébrés mal connus grâce à l'identification de la biodiversité à l'OTU (article 4).

1 Considération des résultats dans le cadre de la gestion et de la conservation des têtes de bassin versant

Un résultat majeur de cette thèse réside dans la mise en évidence de la primauté des processus liés à la dispersion dans la dynamique des communautés de têtes de bassin versant et de leur variation dans le temps. En accord avec la DCE, la gestion et la conservation des écosystèmes de rivières se base sur l'évaluation de leur état écologique via la comparaison de l'état d'« éléments de qualité » à des états de référence. Pour les éléments de qualité biologique, l'évaluation se concentre sur les réponses des communautés aux conditions environnementales et ce à l'échelle locale uniquement. En plus de la nécessité d'intégrer les têtes de bassin versant, la DCE profiterait de la prise en compte de la dispersion qui agit à l'échelle régionale en considérant la connectivité au sein du réseau et les échanges associés ainsi que la variabilité hydraulique et environnementale associée au dynamisme du régime hydrologique (Cid et al., 2020).

Les problématiques associées au dynamisme du régime hydrologique ont déjà été abordées dans le cadre de la conservation des rivières intermittentes (Datry et al., 2016 ; Stubbington et al., 2018). Par exemple, un effort a été investi dans le développement de nouveaux indices prenant en compte la variabilité des habitats dûe aux assèchements (Drought Effect of Habitat Loss on Invertebrates, (Chadd et al., 2017) ou aux régimes hydrologiques (MIRAGE par CORDIS, 2015 ; Stubbington et al., 2018) afin d'avoir une bio-évaluation adaptée à la définition d'un état de bonne qualité dans ces milieux dynamiques. Cependant, ces indices restent centrés sur l'échelle locale et la niche écologique sans considération des systèmes à l'échelle régionale. Il a pourtant été montré que les assemblages des

1. Considération des résultats dans le cadre de la gestion et de la conservation des têtes de bassin versant

communautés d'invertébrés dans ces systèmes dynamiques étaient dépendants de la connectivité et de la structure du réseau et que des dynamiques « source-puits » y étaient à l'œuvre (Sarremejane et al., 2017). Un appel à leur prise en compte a été fait récemment à cet égard (Stubbington et al., 2018).

Dans de tels systèmes, il est nécessaire de passer d'une gestion à l'échelle locale, insuffisante et biaisée (p.ex si le site évalué correspond à un site intermittent), à une gestion intégrant l'échelle régionale. L'évaluation de l'état écologique d'une rivière évoluerait alors en l'évaluation de l'intégrité écologique d'un système de rivières (Figure 1.1). Cette nouvelle évaluation consisterait en la description de la connectivité biologique entre zones refuges (p.ex. tronçons pérennes) et zones sensibles (p.ex. tronçons intermittents) et en sa caractérisation (valeur, type de fragmentation, etc). Cela permettrait à terme de mettre en place des actions de protection et/ou de restauration dans les têtes de bassin versant : de protection en limitant la construction de seuils, de retenues collinaires, de pompages ou en freinant la destruction des zones refuges par les usagers (négligence à l'égard des rivières situés en amont, cf Introduction, section 4) ; de restauration en augmentant la connectivité au sein des systèmes, par exemple en détruisant les seuils qui compromettent l'intégrité du système comme le propose à l'échelle européenne le Dam Removal Project (Gough et al, 2018) et à l'échelle française le Plan d'Action de la Restauration de la Continuité Ecologique des cours d'eau (Bellec & Lefebvre, 2012).

La mise en œuvre de cette évaluation pourrait se faire à travers le cadre de l'étude des métacommunautés, sous réserve que les patrons mis en évidence à un instant t reflète la structure des communautés à une échelle temporelle courte (cf section 2.3 de ce chapitre pour des réflexions sur cette question). Idéalement, un outil basé sur l'étude des métacommunautés permettrait d'évaluer l'état d'intégrité écologique d'un système de rivières donné, d'identifier les causes d'un mauvais état d'intégrité (p.ex. assèchements, seuils) et de proposer des solutions pour atteindre un bon état, dit de référence. Pour obtenir un tel outil, plusieurs points sont à considérer. Premièrement, l'identification de systèmes de référence est nécessaire. Ces systèmes doivent être proches de conditions naturelles (c'est-à-dire non fragmentées par l'humain) et doivent représenter la variabilité de régime hydrologique

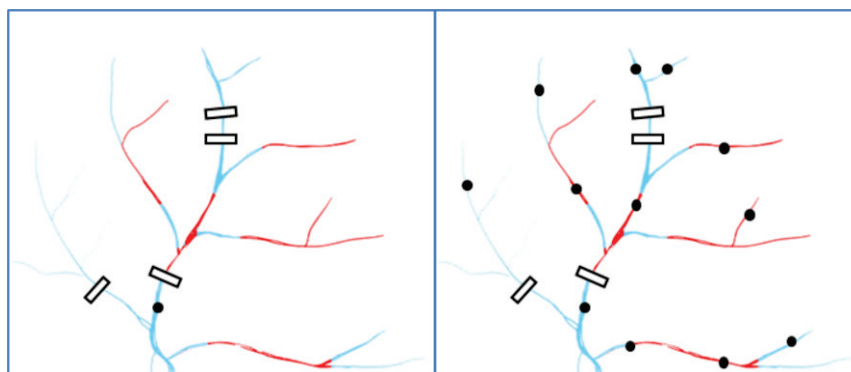


Figure 1.1 : Exemple du changement du nombre de sites d'évaluation d'une gestion centrée sur l'échelle locale comme actuellement (gauche) à une gestion intégrant l'échelle régionale (droite) permettant d'évaluer l'intégrité écologique d'une tête de bassin versant . Les points noirs représentent les sites d'évaluation, les barres des obstacles à la dispersion, les sections bleues et rouges les tronçons pérennes et intermittents, respectivement.

pouvant être rencontré dans l'ensemble des têtes de bassin versant (Stubbington et al, 2018). Deuxièmement, une échelle spatiale de gestion doit être déterminée. Faut-il établir une évaluation sur quelques km²? A l'échelle d'une région? Doit-on intégrer les cours principaux, potentiels sources de colonisateurs dans l'évaluation? Une bonne échelle spatiale de gestion est spécifique aux régions (Viana & Chase, 2019) ainsi qu'aux groupes biologiques étudiés (Padial et al, 2014). Dans un premier temps, il serait judicieux d'avoir une approche intégrative avec plusieurs échelles spatiales afin de définir la ou les meilleure(s) échelle(s) spatiale(s) de gestion dans le cadre des invertébrés et des têtes de bassin versant de France. Troisièmement, la variation temporelle doit être prise en considération, par exemple en identifiant des phases clés dans ces systèmes dynamiques, comme la période de fragmentation maximale pendant l'été et/ou la phase pré-assèchement où les sections intermittentes sont à un niveau de "résilience" maximale par rapport à la date de dernière remise en eau.

Récemment, Cid et al (2020) ont proposé un modèle prédictif comme outil pour évaluer l'état des milieux très dynamiques dans une approche métacommunauté qui permettrait de prendre en compte ces différents points (Figure 1.2). L'idée proposée par les auteur-trices est de coupler les données environnementales et les métriques du régime hydrologique locales aux distances régionales et aux traits biologiques pour développer des modèles prédictifs qui détermineraient les communautés attendues sur des systèmes de rivières de référence selon plusieurs critères. Ceci pouvant être fait à plusieurs échelles spatiales et pour différentes phases hydrologiques. Ces critères peuvent prendre en compte le contexte de fragmentation, d'altérations anthropiques et/ou de capacités de dispersion, permettant d'établir des classes le long d'un gradient de qualité. Des modèles seraient ensuite construits pour les systèmes de rivières à évaluer et comparés aux modèles de référence avec une analyse de régression (Figure 1.2). Les données produites durant cette thèse vont être utilisées pour tester les modèles en collaboration avec Nuria Cid pour une première analyse afin de tester leur efficacité et potentiel. Utiliser un tel modèle demande cependant de définir des caractéristiques pertinentes pour les têtes de bassin versant d'un point de vue local, régional mais aussi temporel. Ainsi, il est nécessaire d'approfondir nos connaissances sur ces milieux.

Le changement d'échelle proposé ici implique un effort d'échantillonnage conséquent qui semble difficile à mettre en œuvre en terme de coût et de temps avec une identification morphologique. L'outil moléculaire présenté en deuxième partie de cette thèse permettrait cet effort d'échantillonnage tout en affinant la description des communautés grâce à une résolution taxinomique aux MOTUs. De plus, l'identification moléculaire couplée aux MOTUs a le potentiel de mettre en évidence des patrons de distribution plus représentatifs de la réalité et donc d'en inférer les processus sous-jacents de manière plus fiable et juste (Article 4). Il serait aussi possible de mettre en évidence des MOTUs spécifiques à certains types de tronçons, comme les tronçons très isolés ou intermittents, que l'identification morphologique ne permet pas de discerner. Ceci permettrait de mettre en évidence des taxa adaptés localement à des conditions particulières au sein des têtes de bassin versant. Cependant,

1. Considération des résultats dans le cadre de la gestion et de la conservation des têtes de bassin versant

les résultats de cette thèse ont montré que l'identification moléculaire ne permettait pas de détecter exhaustivement l'ensemble des taxa présents dans les communautés d'invertébrés et laissent penser que l'identification moléculaire n'est pas encore à même de remplacer l'identification morphologique et nécessite encore des optimisations.

La suite de ce chapitre présente les axes d'approfondissements qui ressortent de cette thèse pour améliorer la compréhension de l'organisation de la biodiversité dans les têtes de bassin versant à travers l'amélioration de la description des processus écologiques et l'optimisation de l'outil moléculaire.

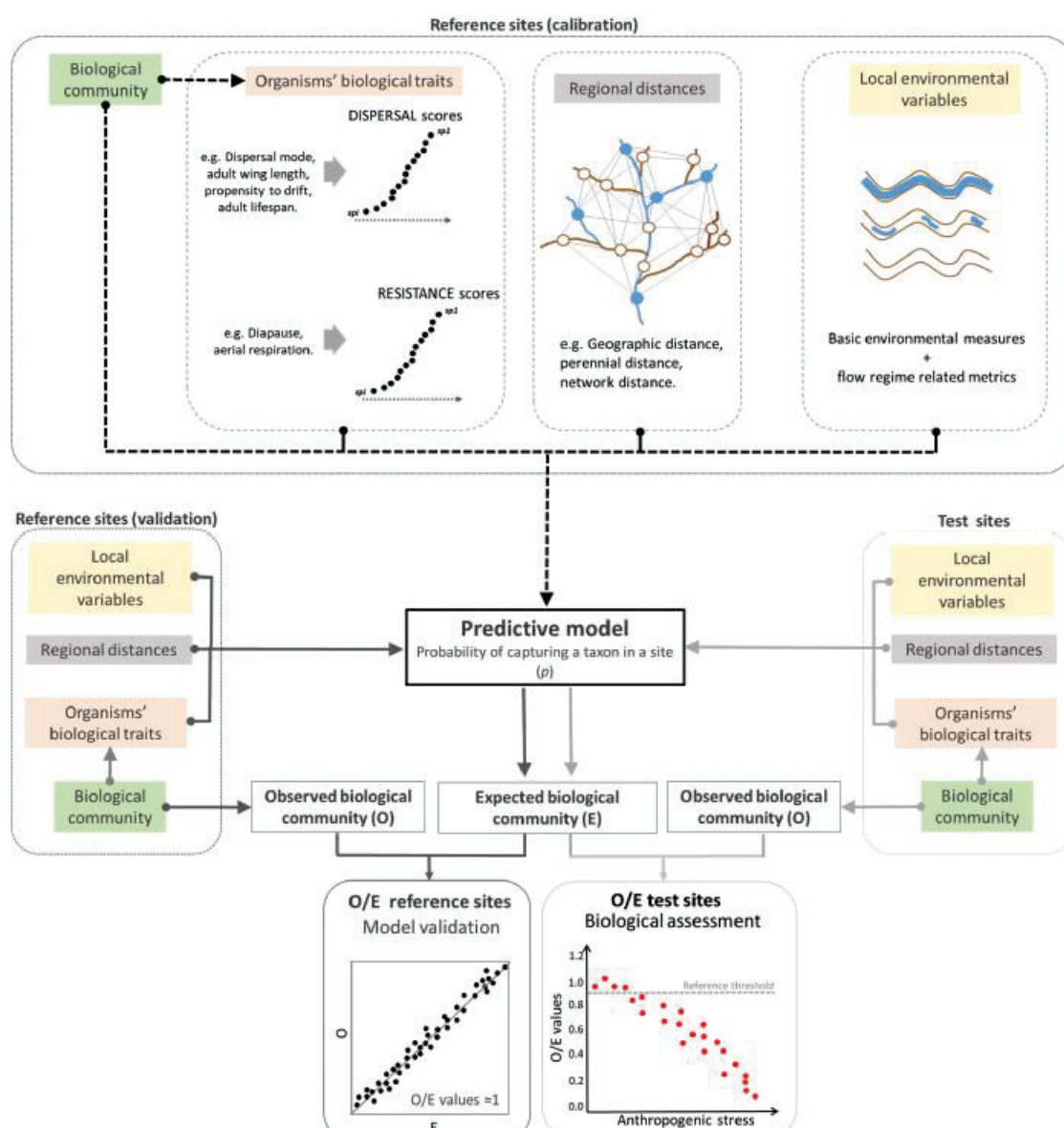


Figure 1.2 : Différentes étapes pour intégrer le concept de métacommunauté dans l'évaluation du bon état des communautés dans les milieux très dynamiques par Cid et al (2020). Les lignes noires en pointillé représentent les étapes pour construire le modèle prédictif, les lignes grises foncées (gauche) représentent le modèle de validation et les lignes gris clair (droite) l'évaluation des rivières. Proviens de Cid et al (2020).

2 Aller plus loin dans la description des processus structurant les métacommunautés des têtes de bassin versant

2.1 Mieux décrire les dynamiques de dispersion

La dispersion est un processus difficile à prendre en compte dans les métacommunautés où de nombreuses espèces sont étudiées. L'étude des métacommunautés s'appuie sur des proxys qui reflètent les effets de la dispersion sur la structure des communautés (Heino et al., 2015), comme les modes de dispersion des taxa ou les distances spatiales utilisées dans cette thèse. Diversifier ou affiner la description de ces proxys permettrait d'améliorer la compréhension des processus à l'œuvre dans les métacommunautés.

Dans les articles 1 et 2, nous avons vu que les taxa ayant un mode de dispersion strictement aquatique et ceux avec un stade de vie aérien ne présentaient pas les mêmes réponses à la fragmentation et aux autres distances spatiales. Améliorer la description de la distribution des espèces en discriminant les traits de dispersion en plus de catégories serait une première amélioration. Par exemple, en prenant en compte la nature active ou passive de la dispersion, en plus du mode de dispersion aquatique ou aérien. Des auteur·trices ont établi une métrique à partir de ces quatre modes de dispersion où un groupe taxinomique à stade adulte aérien et dispersion active tel que les Odonates ont une valeur de métrique forte (c'est-à-dire un potentiel de dispersion fort), contrairement aux bivalves, espèces aquatiques dispersant par voie passive, dont les capacités de dispersion sont faibles (Bonada et al., 2012 ; Li et al., 2018). De tels indices permettent d'interpréter de manière linéaire les réponses des espèces selon leur capacité de dispersion à des gradients longitudinaux ou des échelles spatiales (Li et al., 2018). D'autres données relatives à la dispersion des espèces peuvent aussi être intégrées pour former des groupes les plus informatifs possibles. Par exemple, la base de données DISPERSE (Sarremejane et al., 2020) a répertorié 9 traits biologiques pouvant être utilisé comme proxys de la dispersion. Un indice basé sur une combinaison des traits de cette base de données permettrait de définir les capacités de dispersion d'une manière très réaliste et demande à être testé. Pour finir, ajouter à cette dispersion spatiale la dimension temporelle via les traits de résistance aux perturbations (p.ex. diapause, enkystement, Lytle & Poff, 2004 ; Vander Vorste et al., 2016) ajouterait un niveau de compréhension dans la distribution des espèces (Bohonak & Jenkins, 2003).

Une deuxième amélioration qui permettrait d'aller encore plus loin mais qui demande un plus grand investissement en termes de coût et de temps serait de quantifier les phénomènes de dispersion au sein du réseau. Cela concerne la quantification de la dérive, la dispersion des adultes et les voies de dispersion qu'ils empruntent (p.ex. le long du réseau ou entre les rivières), la quantification de la dispersion efficace qui permet de prendre en compte l'établissement réel des espèces sur un lieu donné et l'identification des habitats sources de colonisateurs. Plusieurs méthodes ont déjà été utilisées pour

2. Aller plus loin dans la description des processus structurant les métacommautés des têtes de bassin versant

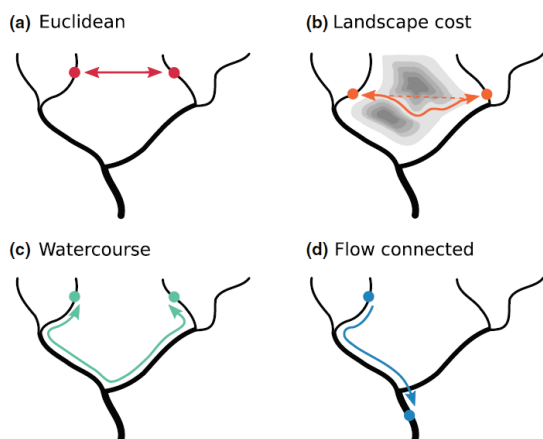


Figure 2.1 : Plusieurs types de distances utilisées dans les études de métacommautés. Distance (a) euclidienne, (b) liée à la conformation du paysage, (c) le long du réseau et (d) liée à l'écoulement de l'eau. Tonkin et al., 2017.

mesurer ces phénomènes comme l'utilisation d'isotopes stables pour marquer les individus et identifier leur distance de dispersion (Macneale et al., 2005) ou leur dérive (Hershey et al., 2010), la capture des individus qui dérivent avec des filets (p.ex. Shearer et al., 2002) ou qui dispersent le long des rivières ou à travers la forêt (p.ex. Petersen et al., 2004). Les méthodes génétiques utilisées pour les métapopulations s'avèreraient très utiles pour mettre en évidence ces phénomènes (Alp et al., 2012 ; Coulon et al., 2010 ; Phillipsen & Lytle, 2013). Bien sûr, il n'est pas possible de déterminer les structures génétiques pour des

centaines d'espèces. Une stratégie serait de déterminer des espèces ou taxa qui représenteraient des points de repère clés dans le panel des modes et capacités de dispersion rencontrés chez les invertébrés (p.ex. sur un gradient de capacité de dispersion ou dans chaque catégorie de mode de dispersion) et de s'intéresser aux phénomènes de dispersion chez ces espèces, comme tente de le faire le projet ALBACOM (www.graie.org/zabr/zabrdoc/AccordCadre-fiches/2017/2017-55-ALBACOM-projet.pdf). Cette stratégie permettrait d'avoir une vision assez réaliste de l'ensemble des processus à l'œuvre dans une métacommunauté. Alternativement, la méthode développée par Elbrecht et al. (2018) utilise des données issues de metabarcoding pour y détecter des variations intra-spécifiques chez plusieurs espèces en simultané et pourrait mettre en évidence certaines structures génétiques. Bien que cette méthode ne soit pas optimale et ait été testée sur une large échelle spatiale (partie Sud de la Finlande, plus de 50 000 km²), elle demande à être testée à différentes échelles spatiales et pourrait représenter un pas vers une quantification précise de la dispersion pour une large gamme d'espèces dans les rivières.

Une troisième et dernière amélioration serait d'intégrer dans les analyses des distances spatiales plus appropriées (Tonkin et al., 2018). En effet, cette thèse a utilisé plusieurs types de distances spatiales comme proxies de dispersion, la distance euclidienne et celle le long du réseau (celles en lien avec la fragmentation ne représente pas un proxy de dispersion mais un facteur la limitant, articles 1, 2 et 4). Mais d'autres distances spatiales existent comme celles utilisées par Cañedo-Argüelles (2015) ou présentées par Tonkin et al. (2017, Figure 2.1). Typiquement, ces distances permettent de représenter de façon plus réaliste certaines voies de dispersion. Par exemple, Cañedo-Argüelles et al. (2015) ont pris en compte la distance topographique qui prend en compte des voies de dispersion concaves qui sont supposées favorables pour la dispersion des adultes dispersant par voie terrestre ou aérienne. D'autres auteurs ont intégré la dispersion biaisée vers l'aval qui permet d'intégrer la contrainte que

l'écoulement de l'eau impose aux organismes (Padial et al., 2014). La distance aux sources pérennes et la position de ces dernières sont aussi importantes dans la réponse des espèces à l'assèchement (Davey & Kelly, 2007). De manière générale, ces distances ont montré leur intérêt pour améliorer la description de la distribution des espèces par rapport aux distance euclidiennes et permettent donc de mettre en évidence les voies de dispersion empruntées par les espèces (Tonkin et al., 2018). Par ailleurs, elles peuvent être utilisées dans la modélisation se basant sur la théorie des graphes où un réseau de rivières est représenté comme un ensemble de nœuds et de liens pondérés par les capacités de dispersion des espèces ou la distance et résistance du paysage (Heino et al., 2017). Dans le cas des têtes de bassin versant, les résultats de cette thèse suggèrent que la dispersion a lieu préférentiellement le long du réseau et que la dispersion à travers le paysage est limitée (article 1). Ainsi, il serait intéressant de tester la prise en compte du courant amont/aval, notamment pour les taxa strictement aquatique et dispersant par dérive. De plus, sur les bassins étudiés lors de cette thèse, les distances aux sources pérennes aval et amont ont été calculées et comparer la diversité et la composition des communautés en lien avec ces distances permettrait de démontrer la dépendance des communautés aux sources pérennes.

2.2 Prise en compte de variables supplémentaires pour décrire la niche écologique et étude d'autres habitats

Cette thèse s'est énormément centrée sur les dynamiques de dispersion au sein des métacommunautés de macro-invertébrés benthiques, notamment dûes aux assèchements, mais n'a pas négligé la description des habitats en incluant 12 variables environnementales pertinentes pour décrire l'habitat lors des prélèvements (articles 1, 2 et 4 ; Karaouzas et al., 2019). Il est pourtant certain que des facteurs influençant la présence des espèces d'invertébrés à un site donné n'aient pas été considérés, comme par exemple le débit ou la teneur en polluants (Jacobsen, 1998 ; Poff et al., 1997). Il a été montré chez les Chironomidae que la distribution des espèces était en grande partie déterminée par le degré de pollution des sédiments (Carew et al., 2007). Le compromis fait entre le nombre d'échantillons et le temps passé sur site pour atteindre l'échelle spatiale et temporelle des données de cette thèse ne nous a pas permis de mesurer ces facteurs pour chaque site de prélèvements. Pourtant, pouvoir inclure ces informations dans les mesures d'habitat va devenir crucial dans un futur proche où les têtes de bassin versant vont être de plus en plus être impactées par les perturbations anthropiques (Reid et al., 2019). Des méthodes pour estimer rapidement la pollution par certains composés ou pour établir la qualité d'un tronçon sont en développement pour augmenter la prise en compte des pollutions dans les mesures de l'habitat local (Bernalte et al., 2020 ; Effendi, 2016). Un autre facteur toujours supposé mais jamais quantifié dans les mesures d'habitat correspond aux interactions biotiques au sein des invertébrés mais égalementen interaction avec d'autres compartiments biologiques. Il est pourtant clair que les espèces interagissent entre elles via la compétition, la

prédation ou la facilitation (Ohlmann, 2019). Récemment, il a été mis en évidence par Garcia-Giron et al. (2020) que les interactions biotiques participaient considérablement au paradigme de *species sorting* dans la structure des métacommunautés et ne devaient pas être négligées dans les analyses de composition des communautés et les politiques de gestion.

Une faiblesse de notre étude liée à l'échelle spatio-temporelle fine à laquelle nous avons travaillé est le manque de répliques et d'habitats représentés dans le jeu de données. Comme mentionné dans les articles 1 et 2, nous avons pu mettre en évidence grâce à deux jeux de données que les inférences de nos analyses étaient robustes par rapport au nombre de répliques. Cependant, il est possible que notre plan expérimental ne favorisait pas la détection d'espèces rares et que couplé à l'identification grossière, il surestimait la part de *nestedness* dans les communautés. Évaluer la sensibilité des parts relatives des composantes de la diversité beta (*nestedness* et *turnover*) au nombre de répliques et à la résolution taxinomique est nécessaire pour explorer l'impact de ces deux facteurs sur leur estimation et attester de la véracité de nos résultats. Dans le futur, l'utilisation de l'identification moléculaire pourra régler ces deux limites d'une part en permettant le prélèvement de plus d'échantillons et d'autre part en améliorant la résolution taxinomique. Finalement, l'étude des métacommunautés de têtes de bassin versant gagnerait à prendre en compte d'autres habitats que les radiers ou de prendre en compte leur variabilité sur un tronçon avec d'autres méthodes d'échantillonnage (p.ex. kick sampling, Stubbington et al., 2009). Par exemple, les habitats des mouilles pourraient connaître des changements des conditions environnementales plus prononcés lors des assèchements. En effet, en phase d'assèchement, les conditions environnementales vont changer et devenir un facteur de sélection aussi bien dans les radiers que dans les mouilles. Néanmoins, dans les mouilles, ce changement et cette sélection se feront sur une plus longue durée que dans les radiers où l'on observe une perte de l'habitat plutôt qu'un changement des conditions environnementales. Il est donc possible que les processus liés à la niche y soient plus importants que dans les radiers. Lors des échantillonnages de cette thèse, il a été présenté que deux prélèvements ont été effectués : un dans les têtes de radiers pour l'identification morphologique, un autre dans les têtes de radiers pour l'identification moléculaire. En réalité, un troisième prélèvement a été effectué dans une mouille à chaque site. L'identification n'a pas encore été effectuée mais permettra d'intégrer un autre habitat dans les analyses et un nouveau niveau de compréhension des processus dans les têtes de bassin versant étudiées.

2.3 Intégration de la dynamique temporelle dans les analyses

Intégrer la dynamique temporelle dans l'étude des têtes de bassin versant implique plusieurs niveaux de réflexion : (i) quelle résolution temporelle est nécessaire pour capturer au mieux la variation des processus dans les métacommunautés ? (ii) comment améliorer la prise en compte de la variabilité du régime hydrologique et des conditions environnementales associées (p.ex. Snell et al., 2014) à l'échelle locale ? et (iii) comment prendre en compte les assèchements à l'échelle régionale ?

Dans les articles 1 et 2 et ailleurs (Cid et al., 2020 ; Datry et al., 2016 ; Tonkin et al., 2017), l'importance d'augmenter la résolution temporelle des prélèvements a déjà été mise en avant comme une nécessité pour comprendre les processus à l'œuvre et évaluer l'état écologique des rivières dans les systèmes très dynamiques. De tout évidence, la meilleure résolution temporelle serait la plus fine possible pour déterminer précisément comment les processus opèrent à différents instants de chaque phase hydrologique et aux patrons d'assèchements associés dans les têtes de bassin versant. Mais cette résolution demanderait un effort d'échantillonnage et d'identification important qui ne permettrait pas d'être utilisé à large échelle ou d'être pérennisé dans les politiques de gestion et de conservation. Dans l'article 2, nous mettons en avant l'intérêt d'identifier des phases hydrologiques clés pour évaluer l'état écologique des rivières. Pour cela, il faudrait pouvoir suivre idéalement plusieurs têtes de bassin versant à une résolution temporelle fine (15 jours, Crabot et al., 2020) et sur plusieurs années afin de déterminer ces phases clés. Dans le cadre de l'article 2, les données de 6 mois de prélèvements représentant les différentes phases hydrologiques ont été analysées. Cependant, les rivières du Thouaret ont été échantillonnées 9 fois sur la période de mars 2018 à décembre 2018. Intégrer le reste des données dans les analyses, qui à ce jour n'est pas encore trié et identifié, représente un premier pas vers l'identification de ces phases hydrologiques clés. L'utilisation de l'outil moléculaire pour augmenter le nombre de prélèvements en facilitant le temps de traitement et d'identification des échantillons pourrait aussi aider à identifier ces phases hydrologiques clés.

Dans cette thèse, la variabilité du régime hydrologique a été prise en compte soit qualitativement (nature du régime hydrologique des tronçons, article 1), soit quantitativement (durée d'assèchement sur la période d'échantillonnage des tronçons, article 2). Dans les deux cas, ces informations ont été incluses dans les matrices des mesures environnementales et analysées de manière répétée à chaque date d'échantillonnage. Pour étudier la dynamique temporelle du régime hydrologique et des conditions environnementales associées, il est aussi possible de mesurer une diversité beta temporelle qui correspond pour un site donné aux dissimilarités entre dates (Crabot et al., 2020 ; Jabot et al., 2020). Bien que cette mesure permette de mettre en évidence des liens entre diversité spécifique et durée et fréquence d'assèchements à l'échelle locale (Crabot et al., 2020), elle ne permet pas d'intégrer l'échelle spatiale et donc le processus de dispersion dans les analyses. Une autre approche consiste en l'analyse multivariée qui quantifie dans un premier temps les variations temporelles d'assemblage des invertébrés au niveau des sites puis, dans un deuxième temps, analyse si une variation spatiale de ces assemblages entre dates est visible (Doledec et al., 2017). Cette méthode est très utile pour explorer les différences d'assemblages dans le temps et l'espace entre type de communautés (p.ex. entre tronçons intermittents et pérennes) mais n'intègre toujours pas les distances spatiales et donc les échanges et le rôle de la dispersion dans les analyses. A ce jour et à ma connaissance, l'intégration concrète de la dynamique temporelle du régime hydrologique et des conditions environnementales associées pour l'étude de la structure des métacommunautés n'a été faite que par Jabot et al. (2020). Elles et ils

proposent un modèle prenant en compte à la fois les diversités beta temporelle, spatiale et ses composantes et l'influence directe et indirecte de l'environnement, de la dispersion et de la dérive écologique. Bien que cette méthode donne seulement une vision globale de la structure des métacommunautés, elle a le mérite de mettre en évidence en une unique analyse les points à approfondir dans la structure de telle ou telle métacommunauté et permettrait rapidement de voir par exemple si des métacommunautés supposées similaires sont structurées globalement par les mêmes facteurs ou si des dépendances à certains contextes sont à l'œuvre.

L'intégration précise de la dynamique temporelle, que ce soit à l'échelle locale ou à l'échelle régionale, demande une quantification des assèchements. A l'échelle locale, les informations pertinentes sont la durée et la fréquence des assèchements (Crabot et al., 2020 ; Datry et al., 2016). A l'échelle régionale, en plus des deux informations précédentes, la configuration des assèchements serait également à prendre en compte (Crabot et al., 2020). Dans l'article 2, il a été possible d'avoir ces informations grâce à la pose de capteurs d'état d'écoulement et à la science participative de l'association Carg'eau mais cela concernait 30 sites et un bassin versant. Pour l'article 1, il n'était pas envisageable de poser des capteurs sur les 220 sites étudiés et d'étudier les patrons d'assèchements pour les 10 têtes de bassin versant. A l'origine, nous avons placés des iButton® (maxim integrated, USA), sondes de température avec un enregistrement toutes les 2 heures, sur la moitié des sites de prélèvements. L'objectif était de déterminer le régime hydrologique à partir des données de température grâce à l'amplitude de température ou les températures maximales journalières (Marek Plosak, modèles non publiés). Ensuite, en collaboration avec l'équipe Hydrologie et Hydraulique de INRAE, il s'agissait de croiser ces informations avec les données pluviométriques, de température, de topographie et de débit si disponibles (base de données [ONDE](#)) pour extrapoler sur tout un bassin versant les patrons d'assèchements entre les deux dates de prélèvement. Malheureusement, les modèles n'étaient pas satisfaisants pour nos données et ces dernières n'ont pas été exploitées. L'extrapolation sur un bassin versant de la dynamique des assèchements serait pourtant extrêmement précieuse et bénéfique pour les études à large échelle spatio-temporelle comme c'est le cas pour les têtes de bassin versant et plus de recherche dans cette voie est souhaitable. De plus, dans de telles modélisations pourraient être intégrées les prélèvements en eau ou d'autres facteurs d'altération du débit amenant à des assèchements qui pourraient déterminer l'origine naturelle ou anthropique des assèchements qui impactent sur la réponse des espèces aux assèchements (Crabot et al., 2020 ; Gallart et al., 2017).

Pour finir et comme vu en début de section, la prise en compte de la dynamique temporelle amène à des réflexions sur les outils statistiques à utiliser. Un autre axe de réflexion porte selon moi sur la fenêtre d'action des processus mis en évidence qui représente une « boîte noire » dans l'étude de la structure des communautés. Les tests de Mantel qui ont servis tout au long de cette thèse sont-ils pertinents pour étudier les processus agissant à court terme ? Pendant cette thèse, beaucoup de

questionnements à ce sujet ont émergés. Que reflètent les processus mis en évidence à une date t : les processus agissant sur les dernières semaines ou les dernières années ? En conséquence, est-ce que la variation observée dans la contribution des processus sur une période courte reflète les dynamiques récentes ou est liée à des variations de la composition des espèces tout au long de l'année ou à leur phénologie ? Ou alors serait-ce une accumulation des processus agissant à plusieurs échelles temporelles qui font varier la contribution des processus aux différentes dates ? Auquel cas, comment différencier ces différents processus sans faire d'inférences ? L'étude des métacommunautés permet de mettre en évidence les grandes tendances qui ont lieu au sein des systèmes et une discussion sur les possibilités qu'offre ce cadre de travail et les outils utilisés doit avoir lieu pour orienter la mise en place d'outils pertinents pour étudier la dynamique des processus dans les métacommunautés.

2.4 Vers une meilleure description de la fragmentation

Cette thèse a permis de montrer que la fragmentation longitudinale joue un rôle important dans la structure des métacommunautés de macro-invertébrés benthiques dans les têtes de bassin versant et que le contexte de fragmentation semble être un facteur important dans cette structure (articles 1, 2 et 4). De nombreux axes d'améliorations sont à envisager pour mieux prendre en compte la fragmentation dans la structure des métacommunautés et augmenter le pouvoir explicatif des relations observées entre fragmentation et biodiversité.

Concernant la fragmentation par les assèchements, nous avons pu intégrer dans l'article 2 la longueur d'assèchement maximale dans le calcul de la matrice de distance. Intégrer la durée de l'assèchement dans les analyses serait aussi intéressant, quoique difficile à mettre en place. En effet, pour une section asséchée donnée, la durée d'assèchement peut être plus ou moins longue selon la partie de la section, par exemple les extrémités devraient s'assécher plus tard et se remettre en eau plus tôt que le centre de la section. A l'instar de la longueur maximale d'assèchement utilisé dans l'article 2, il serait possible d'imaginer une métrique qui reflèterait la contrainte à la dispersion que représente la durée d'assèchement. Par exemple, une métrique de contrainte prenant en compte la longueur et la durée d'assèchement moyenne comme cela est fait pour les structures anthropiques pourrait être mesurée (Cote et al., 2008). Ainsi chaque section intermittente aurait une valeur de « contrainte à la dispersion » permettant d'intégrer la durée et la longueur des assèchements dans le calcul des distances entre paire de sites. De plus, il serait intéressant de savoir si les assèchements sont d'origine naturelle ou anthropique (section précédente) et de les inclure dans les analyses afin d'étudier des fragmentations par les assèchements « anciennes et naturelles » à des fragmentations récentes et induites par l'humain. Intuitivement, une fragmentation récente devrait avoir un impact fort sur la structure des communautés localement car le pool d'espèces régional ne possède pas la capacité d'adaptation à ces perturbations nouvelles.

2. Aller plus loin dans la description des processus structurant les métacommunautés des têtes de bassin versant

Concernant la fragmentation par les structures anthropiques, des efforts doivent être investis dans la caractérisation et le référencement des structures. En effet, il serait pertinent d'améliorer les distances de fragmentation par les structures anthropiques en déterminant la capacité de franchissement de l'obstacle par les espèces avec une valeur dans le sens aval/amont et une valeur dans le sens amont/aval (Cote et al., 2008). Kemp & O'Hanley (2010) ont utilisé par exemple dans leur mesure de fragmentation la hauteur de la structure, la capacité de nage et la taille des espèces ainsi que la direction du courant dans leur mesure de connectivité pour des poissons. Dans l'article 2, nous avons pris en compte les hauteurs des structures renseignées par la base de données ROE dans les mesures des distances de fragmentation permanente. Cependant, ces informations n'étaient disponibles que pour ce bassin versant sur les onze étudiés de cette thèse. Il est donc important de mieux caractériser les structures pouvant faire obstacle à la dispersion pour estimer des indices de connectivité reflétant au mieux la réalité de dispersion (Rodeles et al., 2020). En France, l'Indice de Continuité Écologique a été récemment mis en place et a pour but de cartographier les obstacles à la dispersion avec une description précise des structures et un risque de blocage de la dispersion (Burgun et al., 2015). Cependant, ce calcul de risque est défini uniquement pour les poissons et non pour les invertébrés, pour lesquels définir ce risque représente un challenge du fait de la diversité des espèces et de leurs modes et capacités de dispersion. Une solution serait de définir des catégories de capacité de franchissement des structures selon des dires d'experts sur les structures et sur les espèces (p.ex. considérer les modes de dispersion aquatique et aérien séparément) comme l'ont fait Tonkin et al. (2014) avec un système de codage flou résultant en 5 catégories de barrière infranchissable à inexistante.

Pour les analyses des différents articles, seules les fragmentations par les structures anthropiques et les assèchements ont été considérées. Comme présenté en Introduction (section 4), la fragmentation longitudinale présente d'autres composantes freinant la dispersion comme des obstacles physiques naturels (p.ex. cascades, Kjærandsen, 2005 ; Tonkin et al., 2014) ou des obstacles biologiques ou chimiques (Fuller et al., 2015). L'intégration de ces fragmentations dans les matrices de fragmentation demande une expertise et une connaissance fortes des têtes de bassin versant étudiées mais permettrait d'estimer la fragmentation au sein du réseau de manière plus représentative.

Finalement, comme suggéré dans l'article 2, les effets de la fragmentation temporaire par les assèchements et de la fragmentation permanente par les structures anthropiques sur la dispersion des espèces serait spécifique à la tête de bassin versant étudié. Dans le cas du Thouaret, la limitation à la dispersion induite par la fragmentation permanente était suffisante et, dans ce contexte, les effets de la fragmentation temporaire étaient nuls ou indécélables. Nos résultats ouvrent la voie vers l'étude des synergies et des antagonismes entre ces deux types de fragmentation et appelle à intensifier la recherche sur ce sujet avec l'étude de plusieurs têtes de bassin versant présentant des intensités de fragmentation permanente et temporaire variables. D'autant plus qu'avec l'augmentation des

assèchements par le réchauffement climatique et les prélèvements (Jaeger et al., 2014), les interactions entre les deux types de fragmentation peuvent évoluer avec notamment une intensification des sections asséchées. Dans un contexte où l'eau sera plus rare, la transition pour certaines rivières d'un écoulement pérenne à un écoulement intermittent va impacter la composition des communautés à l'échelle locale, entraînant un impact sur la distribution des espèces et sur leur dispersion. Par exemple, dans un contexte fragmenté, une recherche de zone refuges plus importante induite par l'augmentation des assèchements pourrait être freinée par la fragmentation permanente. A cela s'ajoutent des effets synergiques avec d'autres perturbations anthropiques (Kalogianni et al., 2017 ; Radinger et al., 2017 ; Radinger et al., 2016). Par exemple, dans une étude portant sur les effets individuels et combinés des assèchements et de la pollution, il a été montré que la distribution des poissons dans des rivières intermittentes polluées et non polluées étaient très divergentes. Dans les rivières où pollution et stress hydrique étaient à l'œuvre, l'effet combiné des deux facteurs amenaient à des conditions hypoxiques induisant un filtre environnemental très fort. Comment la distribution des espèces sera impactée et comment les espèces s'adapteront à une combinaison d'effets de perturbations anthropiques qui ne font qu'augmenter et qui peuvent, en plus de se cumuler, agir en interaction et induire de nouvelles contraintes reste à être exploré.

2.5 Discriminer les groupes d'invertébrés pour mieux appréhender les dynamiques

L'étude des métacommunautés a pour but de décrire les tendances au sein de l'ensemble d'une communauté ou au sein de certaines catégories comme celles liées aux modes de dispersion utilisées pour cette thèse. Pourtant, dans l'article 4, il a été mis en évidence que les patrons de distribution étaient différents selon le groupe d'invertébrés étudié (familles d'insectes ou ordres) même s'il possédait des caractéristiques de dispersion communes avec les autres groupes, par exemple les Chironomidae et les Simuliidae qui font partie du même ordre d'insecte (Diptera). De manière générale, toutes les améliorations proposées dans cette discussion pour affiner la compréhension de la structure des communautés dans les têtes de bassin versant dépendent des caractéristiques des espèces, que ce soit à travers leurs préférences environnementales, leurs traits de dispersion ou leurs traits de résistance ou de résilience. Par exemple, la capacité de franchissement d'un obstacle à la dispersion sera très variable d'un ordre à l'autre, voire d'une espèce à l'autre (Kemp & O'Hanley, 2010). Les interactions biotiques et les préférences environnementales sont également propres aux espèces ou genres (Cadotte, 2006 ; Mondy et al., 2012).

Bien que demandant un développement méthodologique et des évaluations supplémentaires, l'identification moléculaire a le potentiel de différencier les taxa selon leur variabilité génétique et permet d'envisager, à l'interface entre écologie des communautés et des populations, l'étude des groupes d'invertébrés (p.ex. famille ou classe) indépendamment les uns des autres pour affiner la compréhension des processus structurant les têtes de bassin versant (article 4).

3 Amélioration de l'outil moléculaire pour décrire la biodiversité des têtes de bassin versant

Les articles 3 et 4 avaient pour ambition d'améliorer la description des communautés d'invertébrés des têtes de bassin versant grâce à un outil moléculaire, participant ainsi à l'effort grandissant et commun pour faire du metabarcoding un outil fiable et robuste pour la bio-évaluation des invertébrés au sein de l'Union Européenne (Blackman et al., 2019).

Cette section présente les développements possibles pour l'outil moléculaire liés aux thématiques abordées pendant la thèse d'un point de vue méthodologique mais aussi dans une perspective d'application de l'outil en condition naturelle pour la description la plus adéquate possible des communautés d'invertébrés pour des études fondamentales ou appliquées dans les têtes de bassin versant. Cependant, de l'extraction au traitement des données, d'autres sources de biais existent et ne doivent pas être négligées car elles empêchent l'adoption de la méthode comme outil d'identification standard pour la bio-évaluation. De plus, je rejoins l'appel de Blackman et al. (2019) sur la nécessité de mettre en place des protocoles standards pour les politiques de gestion, au moins à l'échelle européenne où les objectifs sont communs à tous les pays. En effet, du prélèvement au traitement bio-informatique, chacun-e a son propre protocole qui peut influencer les détections des espèces et les estimations de métriques (Brandon-Mong et al., 2015 ; Yang et al., 2013) et rend impossible l'acceptation du metabarcoding comme outil d'identification pour la bio-évaluation des rivières dans le cadre de la DCE.

3.1 Limites associées à l'ADN extrait à partir de l'éthanol

Le potentiel de l'ADN extrait à partir de l'éthanol de conservation (ADNet) a été démontré dans cette thèse pour l'obtention de données de présence/absence sur des communautés artificielles (article 3) et dans le cadre de l'étude des communautés prélevées *in situ* (article 4) en substitut à l'ADN bulk. Cependant, des limites méthodologiques ont été mises en évidence lors de la préparation des échantillons de l'article 4 et méritent d'être explorées, principalement pour l'application de la méthode en condition réelle.

L'ADNet est connu pour être de moins bonne qualité et en faible concentration par rapport à l'ADN bulk (Blackman et al., 2019). Dans l'article 3, contre toute attente, nous avons pourtant pu amplifier un long fragment du COI de 421 pb (couple d'amorces BFR2). En testant ce couple d'amorces pour les échantillons prélevés *in situ* de l'article 4, le nombre d'OTUs affiliés aux invertébrés après séquençage était très réduit (~30) et la profondeur de séquençage était très faible. Les données du couple d'amorces BFR2 n'ont pas été considérées comme fiables ou informatives pour les analyses et n'ont pas été exploitées. Ainsi, l'ADNet des échantillons de l'article 3 a été extrait dans des

conditions privilégiées où l'éthanol est resté très concentré et où l'ADN était mieux conservé, ce qui n'a pas lieu fréquemment en milieu réel (Barnes & Turner, 2015).

Deuxièmement, des différences de détection ont été observées entre l'identification morphologique et l'identification morphologique au niveau des familles. La Figure 3.1 présente les différentes familles identifiées morphologiquement et leur détection par l'ADNet et le couple d'amorce Fwh1.

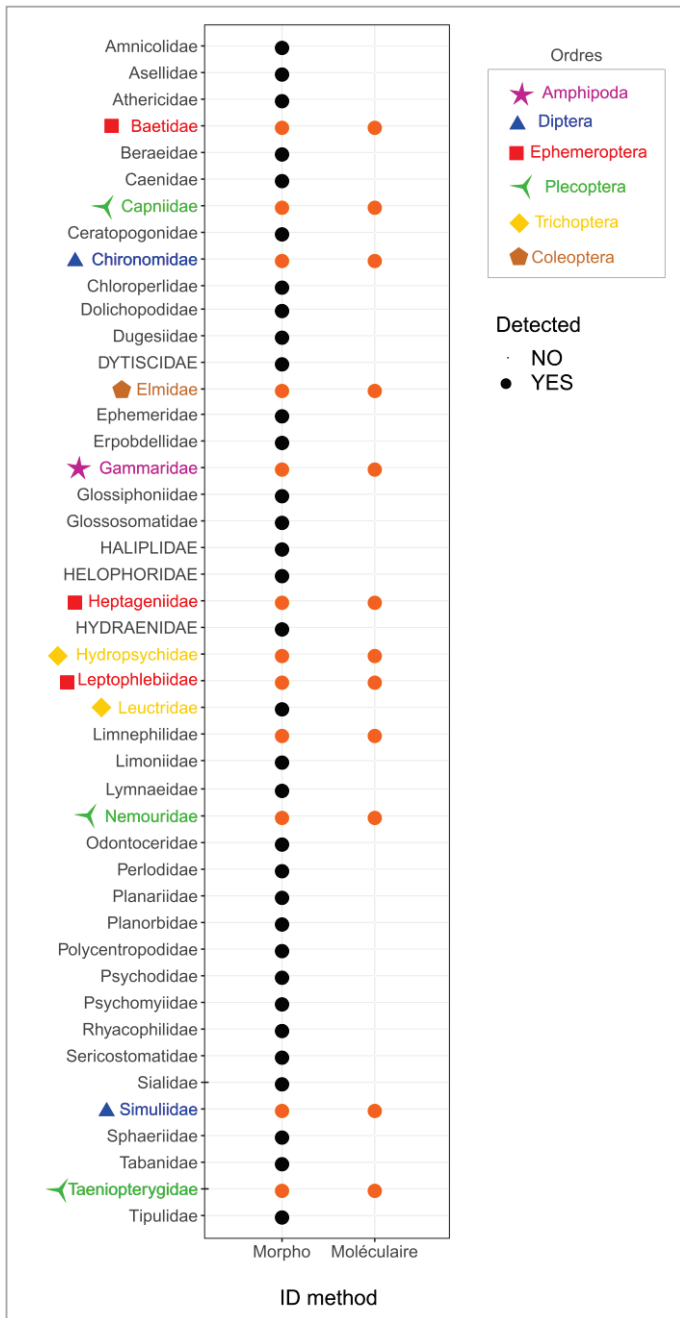


Figure 3.1 : Détection par l'identification moléculaire (droite) des familles identifiées en identification morphologique (gauche) des données de l'article 4. Une couleur orange indique que la famille a été détectée par l'identification moléculaire. Les ordres correspondant des familles détectées sont représentés par une forme et une couleur différentes

La Figure 3.1 présente les différentes familles identifiées morphologiquement et leur détection par l'ADNet et le couple d'amorce Fwh1. L'identification moléculaire détecte très peu de familles présentes dans l'échantillon d'identification morphologiquement (27%). Une limite de cette analyse déjà évoquée dans l'article 4 est que l'éthanol ayant servi pour l'identification moléculaire ne provenait pas des échantillons identifiés morphologiquement. Ainsi, il n'est pas possible de dissocier si l'absence de détection provient de l'absence dans les échantillons moléculaires des familles non détectées ou à des biais liés à l'ADNet et à la PCR. Il semble tout de même peu probable que des différences aussi fortes soient observées à l'échelle de la diversité gamma sur 20 double-échantillonnage du même habitat. Parmi les familles non détectées par l'identification moléculaire, neuf étaient présentes avec des abondances relativement élevées (entre 15 et 117 individus) et il est probable que ces familles aient été présentes dans les échantillons utilisés pour l'identification moléculaire. Les biais associés à l'identification moléculaire semblent donc à privilégier. Etant donné que le couple d'amorces Fwh1 permet d'amplifier des taxa dans la plupart des familles non détectées ici (Vamos et al., 2017), la matrice éthanol semble être la principale source de la non-détection des familles. En contraste avec les communautés artificielles de l'article 3, les échantillons étaient ici prélevés avec le substrat des organismes dont les composés diminuent la concentration en éthanol accélérant

3. Amélioration de l'outil moléculaire pour décrire la biodiversité des têtes de bassin versant

la dégradation de l'ADN qui y est en suspension et qui peuvent libérer des inhibiteurs de PCR. A ce jour et à ma connaissance, sur les cinq études portant sur l'ADNet (Carew et al., 2018 ; Erdozain et al., 2019 ; Hajibabaei et al., 2012 ; Martins et al., 2019 ; Zizka et al., 2020), seulement deux n'utilisent pas des communautés artificielles ou une méthode ne prélevant pas le substrat (Hajibabaei et al., 2012 ; Martins et al., 2019). Ces deux études montrent des taux de détection élevés par rapport à nos échantillons (>70%). La méthode d'extraction peut affecter la concentration d'inhibiteurs de PCR, le rendement de l'ADN et la détection des espèces (Deiner et al., 2015 ; Fredricks et al., 2005) et il est possible que la méthode d'extraction mise en place durant cette thèse ne soit pas optimale lorsqu'elle est appliquée à des échantillons prélevés *in situ* et mérite d'être optimisée. Il serait possible de quantifier le biais de détection en identifiant morphologiquement les échantillons dont provient l'éthanol extrait. Cela permettrait d'identifier si des facteurs spécifiques (obstruction de leur opercule des mollusques, épaisseur de l'exosquelette, taille de l'organisme) joue un rôle dans la détection des espèces par l'ADNet. En résumé, ces résultats mettent en avant l'importance de tester la fiabilité des outils développés dans des conditions de laboratoire sur des échantillons caractéristiques des échantillons en condition réelle.

L'utilisation de l'ADN bulk *in toto*, c'est-à-dire le broyat de l'intégralité de l'échantillon (organismes et substrat, Nichols et al., 2019), est prometteuse dans le cadre de l'étude de la structure des communautés d'invertébrés. Cette méthode représenterait un outil à la fois rapide car elle ne demande pas de tri des organismes, tout en permettant d'obtenir des fragments plus longs pour l'identification et de diminuer les risques de contamination liés à l'ADN bulk et à l'ADNet (moins d'étapes de manipulation). Cependant, son utilisation présente plusieurs challenges pour mettre en place un protocole pour son extraction qui concernent le ratio organismes/substrat dans l'échantillon, le volume que représente un échantillon du fait de la matière organique et inorganique non ciblée, la quantité d'inhibiteurs potentiellement haute, la contamination entre échantillons suivant le système de broyage. Les échantillons moléculaires de l'article 4 pourraient être utilisés pour une extraction *in toto* et permettraient de comparer le potentiel de cette matrice d'extraction à l'ADNet et à l'identification morphologique. Un projet du nom de MISTRAL va débiter au LEHNA en 2020 avec comme objectif de déterminer les meilleures méthodes moléculaires pour décrire les communautés d'invertébrés du bassin versant de l'Albarine dans un cadre de bio-évaluation et s'intéressera notamment à optimiser et poursuivre la mise en place de protocoles d'extraction d'ADN amorcées lors de cette thèse.

Pour de futures recherches de développement, un protocole idéal serait d'extraire l'éthanol de conservation après une période d'incubation à froid, de trier et identifier les organismes morphologiquement, puis de les broyer avec le substrat pour l'extraction bulk *in toto*.

3.2 Perspectives concernant la capture par sondes ARN

La capture par sondes ARN a prouvé son efficacité dans le cadre de l'étude des communautés d'invertébrés présentant des biomasses variables ou une complexité taxinomique élevée et supplantait l'amplification par PCR dans des conditions contrôlées en laboratoire (article 3). Cependant, il nécessite comme l'ADNet d'être testé sur des communautés plus riches et complexes prélevées *in situ* et des approfondissements méthodologiques allant des paramètres d'hybridation à la conception même des sondes.

3.2.1 Application de la méthode sur des échantillons prélevés in situ

Une prochaine étape pour la capture par sondes ARN est son application sur des échantillons prélevés en conditions réelles afin de tester son potentiel dans la description de communautés entières d'invertébrés. Bien que dans l'article 3, la capture par sondes ARN ait fonctionné pour préparer des bibliothèques de séquençage Illumina Nextera® sur de l'ADNet, des difficultés ont été rencontrées lors de son application sur des échantillons prélevés *in situ*. La préparation des bibliothèques, étape relativement coûteuse, n'aboutissait pas pour les deux tiers des échantillons et n'était pas répétable pour un même échantillon. La source de ce problème n'a pas été trouvée et nous avons dû mettre de côté le couple ADNet/capture pour le projet au profit d'une approche plus rapide et fiable que représente l'amplification par PCR. La capture par sondes ARN demande donc encore à être testée sur des échantillons prélevés *in situ* et sur d'autres matrices d'extraction d'ADN (ADN bulk, ADN bulk *in toto*). Il est possible qu'un couple ADN bulk *in toto*/capture soit un outil de choix pour des évaluations rapides, précises et fiables de la biodiversité des têtes de bassin versant.

3.2.2 Développement méthodologique nécessaire

D'un point de vue purement méthodologique, beaucoup d'optimisations vis-à-vis de cette méthode ont été abordées dans l'article 3 et vont être présentées succinctement ici dans le but de mettre en évidence les défis méthodologiques associés à la capture par sondes.

Tout d'abord, une réflexion doit être portée sur la conception des sondes. Dans cette thèse, cette dernière a été faite à partir du COI étant donné qu'il représente le barcode privilégié pour la plupart des groupes d'invertébrés (Evans & Paulay, 2012). Le COI est un gène mitochondrial et par sa nature pose un problème concernant l'estimation de la biomasse des espèces dans un échantillon. En effet, comme vu dans l'article 3 et dans d'autres études (Liu et al., 2016 ; Wilcox et al., 2018), un effet espèce a été mis en évidence dans les relations entre biomasse initiale et nombre de lectures. Cet effet peut être principalement attribué au gène d'intérêt, le COI, en raison de la variabilité du nombre de copies de génome mitochondrial entre tissus, stades de développement et espèces (Rooney et al., 2015) et aux potentiels NUMTs que les sondes peuvent capturer (Hazkani-Covo et al., 2010). Ainsi,

pour avoir une estimation de la biomasse initiale faudrait-il s'intéresser à d'autres marqueurs non mitochondriaux. Cependant, dans le cas des invertébrés, cela impliquerait une perte de l'information taxinomique, ce qui n'est pas profitable dans le cadre de l'étude des métacommunautés ou de la bio-évaluation. Les métacommunautés d'invertébrés peuvent être étudiées avec des indices de présence/absence et arriver aux mêmes conclusions qu'avec des métriques basées sur les données d'abondance (article 1). Concernant la bio-évaluation, de plus en plus d'études montrent le potentiel des données de présence/absence issues du metabarcoding pour évaluer l'état des rivières en comparaison avec les méthodes classiques (Elbrecht et al., 2017 ; Pereira-da-Conceicao et al., 2019 ; Zizka et al., 2020). Si l'objectif de la méthode est uniquement de détecter l'ensemble des espèces constituant les communautés, alors la capture par sondes ARN développée sur du COI est appropriée. De plus, même sans un besoin de données d'abondance, la capture par sondes ARN reste plus performante que l'amplification par PCR dans la détection des espèces, moyennant cependant un coût peut-être trop élevé dans le cadre des politiques de gestion et de conservation (article 3).

Les optimisations méthodologiques porteraient alors sur l'augmentation de la spécificité de capture, c'est-à-dire le pourcentage de lectures d'intérêt par rapport au pool de lectures séquencées. Des essais sont à effectuer sur la température d'hybridation (Li et al., 2013 ; Paijmans et al., 2016), le regroupement des librairies pour la capture, la multiplication des événements de capture (Mariat et al., 2018 ; Portik et al., 2016) et la complexité taxinomique de l'échantillon. Un dernier facteur particulièrement important sur lequel jouer est la divergence entre sondes et séquence d'intérêt. Dans l'article 3, rejoignant les résultats d'autres études (Liu et al., 2016 ; Portik et al., 2016 ; Vallender, 2011), la divergence maximale entre la sonde et le gène d'intérêt devait être de 15% pour pouvoir capturer un fragment. Plus cette distance est diminuée, plus l'hybridation des sondes sera spécifique et moins de séquences parasites seront capturées. Le risque est de diminuer le potentiel de capture de séquences d'intérêt non référencées dans la base de référence pour la conception des sondes. Un compromis entre ces deux paramètres doit être effectué de manière à optimiser la spécificité de capture.

Un autre aspect porte sur la transférabilité d'un set de sondes conçu à d'autres compartiments biologiques ou à d'autres régions. En effet, le set de sondes conçu pour ce projet de thèse visait les invertébrés benthiques des têtes de bassin versant de régions spécifiques en France métropolitaine. Le temps de conception des sondes, allant de la mise en place de la base de données à la conception *in silico*, a été relativement longue (plusieurs mois) et il est légitime de se demander l'intérêt d'un outil si spécifique. Cette étape de conception de sondes pourra se faire plus rapidement et simplement maintenant qu'un pipeline dédié au gène COI existe et à mesure du développement de la méthode. A terme, on peut envisager qu'il sera seulement nécessaire de préparer une base de données avec les taxa cibles mais il est aussi envisageable de concevoir un set de sondes très large contenant des centaines de milliers de sondes pour capturer un ensemble très large d'un compartiment biologique donné (e.g.

invertébrés benthiques). Du fait de la plasticité des sondes évoquée précédemment, un tel set de sondes pourrait permettre de capturer l'ADN d'espèces non présentes dans la base de données initiale.

Lors de la conception des sondes, certaines ont été conçues pour d'autres gènes mitochondriaux. Deux objectifs étaient visés. Le premier était de calculer si un biais de capture existait lorsqu'un grand nombre de sondes était conçu pour la même séquence. Pour cela, des sondes de 16S ont été conçues pour 5 espèces présentes dans les communautés artificielles de 10 espèces en découpant la séquence 16S séquencée au laboratoire en 6 sondes de 120 pb. Pour deux d'entre elles, 10 fois plus de sondes ont été commandées (60 sondes au lieu de 6). Le deuxième objectif était d'étudier si le pourcentage de divergence des sondes aux séquences d'intérêt impactait leur capture. Pour cela, cinq gènes (NAD1, NAD4, NAD5, CYTB et ATP6) de huit espèces proches des espèces des communautés artificielles de Elbrecht & Leese (2017) ont été mutés *in silico* selon un gradient de pourcentage de divergence (10, 15, 20, 25 et 30%). Les données pour ces deux objectifs ont été produites en même temps que celles de l'article 3 et n'ont pas encore été analysées. Connaître plus précisément les liens entre divergence aux sondes et abondances est primordial pour estimer la robustesse de la méthode de capture par sondes ARN et d'autres études à ce sujet doivent être mises en place. De plus, une telle analyse permettrait de démêler les contributions du biais de capture et de divergence aux biais « espèce » dans les biais d'abondance observés dans l'article 3. Pour finir, il serait par la suite intéressant d'étudier le comportement de divergence sur des échantillons tels que ceux de l'article 4 car il est possible que les propriétés de ces échantillons fassent varier la capacité ou la probabilité d'hybridation (p.ex. inhibiteurs, molécules chargées positivement) comme pour l'ADNet.

3.2.3 Conception d'une pipeline bio-informatique spécifique

Afin de pouvoir comparer les résultats issus d'enrichissement par PCR et ceux issus d'enrichissement par capture, un pipeline classique de metabarcoding par PCR a été utilisé (cf Matériel et méthodes - Chapitre 3). Il a fallu adapter ce pipeline conçu pour des amplicons (des séquences qui commencent et se terminent toutes à la même position) aux données issues de capture qui n'ont aucun pipeline dédié. Bien que fonctionnel, ce pipeline est loin d'être optimal. En effet, lors de la sonication, des fragments de tailles variables sont produits et ces fragments représentent des zones variables de la région du COI incluant les régions flanquantes (les séquences qui encadrent) de la zone d'intérêt (ici, la région Folmer du COI qui a servi pour la conception des sondes). Les séquences possédant plus de 41% d'une région flanquante ne sont pas considérées comme des séquences appartenant au COI avec le pipeline utilisé. Ceci induit une baisse de spécificité mais aussi une perte d'informations. Pourtant, avoir accès à plus d'informations génétiques pourrait permettre d'apporter un signal phylogénétique aux données de metabarcoding, ouvrant des portes vers la diversité phyllobeta ou vers une meilleure discrimination de la diversité génétique, notamment des haplotypes qui pourraient être utilisés pour mettre en évidence des patrons de distributions plus fins

chez certaines espèces. Les méthodes d'assemblages utilisées en méta-génomique et transcriptomique permettent de reconstruire une séquence originale, longue, en assemblant des fragments d'ADN. Cependant, elles ne sont pas non plus adaptées pour le couple capture par sondes ARN/gène unique car elles sont développées pour assembler des gènes différents (donc variables au niveau de leur séquence nucléotidique) d'un organisme et non une séquence (de facto similaire) chez de nombreuses espèces. Par exemple, des tests effectués sur les données des communautés artificielles avec l'assembleur Trinity (Haas et al., 2013) n'ont pas été fructueux avec notamment la perte totale de certaines espèces. Les séquences COI pouvant être similaires sur un fragment donné entre espèces, il est très probable que l'assembleur ne puisse pas les discriminer et combinent des lectures d'espèces différentes. Ainsi, il n'existe pas à ce jour de pipeline bio-informatique disponible pour les lectures de metabarcoding enrichies par capture et il est nécessaire de mettre en place un pipeline dédié.

3.3 Réflexion sur l'assignement des données issues de metabarcoding

L'étape d'assignement est cruciale dans les études de metabarcoding car c'est l'étape qui fait le lien entre lectures et taxinomie. De nombreux outils et de nombreuses bases de données existent dont les plus connus pour les métazoaires et le barcode universel sont NCBI et BOLD Systems. Une variété d'outils est à disposition se basant sur des arbres phylogénétiques (BOLD, Ratnasingham & Hebert, 2007), des mesures de similarités (blastn, Altschul et al., 1990), des inférences bayésiennes (SAP, Munch et al., 2008) ou des réseaux de neurones (Shadab et al, 2020). Peu d'attention a été portée sur la mise en évidence des similarités et différences entre ces différents outils par rapport à la qualité d'assignement. L'article 3 n'a pas présenté de problème au niveau de l'assignement étant donné que les haplotypes attendus étaient connus grâce aux bases de données spécifiques aux organismes des communautés artificielles à disposition. Pour l'article 4, le choix de l'assignement a suscité des questionnements sur l'outil à utiliser, sur le comportement des algorithmes et des sorties par rapport à la faible longueur des séquences (178 pb), sur les bases de données à utiliser ou sur les critères ou les méthodes à déterminer pour définir la résolution taxinomique d'un taxon quand l'assignement à l'espèce n'était pas possible. Des choix aussi cruciaux sont souvent laissés à l'appréciation de l'utilisateur·trice du logiciel alors qu'ils sont déterminants pour toutes les analyses de diversité, de communautés ou de bio-évaluation subséquentes. Il est nécessaire de mettre en place des guides accessibles pour les utilisateur·trices concernant ces assignations et les choix à faire en fonction de la nature des données, notamment si l'outil moléculaire a pour ambition d'être généralisé et utilisé par des personnes novices dans le domaine de la biologie moléculaire.

4 Références

Alp, M., Keller, I., Westram, A.M., Robinson, C.T., 2012. How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology* 57, 969-981.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J Mol Biol* 215, 403-410.

Baatrup-Pedersen, A., Larsen, S.E., Andersen, D.K., Jepsen, N., Nielsen, J., Rasmussen, J.J., 2018. Headwater streams in the EU Water Framework Directive: Evidence-based decision support to select streams for river basin management plans. *Sci Total Environ* 613-614, 1048-1054.

Barnes, M.A., Turner, C.R., 2015. The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics* 17, 1-17.

Bellec, P., Lefebvre, E., 2012. Plan d'actions pour la restauration de la continuité écologique des cours d'eau (Parce)-Diagnostic de mise en oeuvre. Rapport n° 008036-0.

Bernalte, E., Arévalo, S., Pérez-Taborda, J., Wenk, J., Estrela, P., Avila, A., Di Lorenzo, M., 2020. Rapid and on-site simultaneous electrochemical detection of copper, lead and mercury in the Amazon river. *Sensors and Actuators, B: Chemical* 307, 127620-127620.

Blackman, R.C., Mächler, E., Altermatt, F., Arnold, A., Beja, P., Boets, P., Egeter, B., Elbrecht, V., Filipe, A.F., Jones, J.I., Macher, J., Majaneva, M., Martins, F.M.S., Múrria, C., Meissner, K., Pawlowski, J., Schmidt Yáñez, P.L., Zizka, V.M.A., Leese, F., Price, B., Deiner, K., 2019. Advancing the use of molecular methods for routine freshwater macroinvertebrate biomonitoring – the need for calibration experiments. *Metabarcoding and Metagenomics* 3, e34735.

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

Bonada, N., Dolédec, S., Statzner, B., 2012. Spatial autocorrelation patterns of stream invertebrates: exogenous and endogenous factors. *Journal of Biogeography* 39, 56-68.

Brandon-Mong, G.J., Gan, H.M., Sing, K.W., Lee, P.S., Lim, P.E., Wilson, J.J., 2015. DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. *Bull Entomol Res* 105, 717-727.

Burgun, V., Chanseau, M., Kreutseberger, K., Marty, V., Pénil, C., Tual, M., Voegtli, B., 2015. ICE nformations sur la continuité écologique Protocole de terrain pour l'acquisition des données, ed. ONEMA, p. 88.

Cadotte, M.W., 2006. Metacommunity influences on community richness at multiple spatial scales: A microcosm experiment. *Ecology* 87, 1008-1016.

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

- Carew, M.E., Coleman, R.A., Hoffmann, A.A., 2018. Can non-destructive DNA extraction of bulk invertebrate samples be used for metabarcoding? *PeerJ* 6, e4980.
- Carew, M.E., Pettigrove, V., Cox, R.L., Hoffmann, A.A., 2007. The response of Chironomidae to sediment pollution and other environmental characteristics in urban wetlands. *Freshwater Biology* 52, 2444-2462.
- Chadd, R.P., England, J.A., Constable, D., Dunbar, M.J., Extence, C.A., Leeming, D.J., Murray-Bligh, J.A., Wood, P.J., 2017. An index to track the ecological effects of drought development and recovery on riverine invertebrate communities. *Ecological Indicators* 82, 344-356.
- Cid, N., Bonada, N., Heino, J., Cañedo-Argüelles, M., Crabot, J., Sarremejane, R., Soininen, J., Stubbington, R., Datry, T., 2020. A Metacommunity Approach to Improve Biological Assessments in Highly Dynamic Freshwater Ecosystems. *Bioscience* 70, 427-438.
- CORDIS, 2015. MIRAGE Project ID: 211732 – Mediterranean Intermittent River Management.
- Cote, D., Kehler, D.G., Bourne, C., Wiersma, Y.F., 2008. A new measure of longitudinal connectivity for stream networks. *Landscape Ecology* 24, 101-113.
- Coulon, A., Fitzpatrick, J.W., Bowman, R., Lovette, I.J., 2010. Effects of Habitat Fragmentation on Effective Dispersal of Florida Scrub-Jays. *Conservation Biology* 24, 1080-1088.
- Crabot, J., Heino, J., Launay, B., Datry, T., 2020. Drying determines the temporal dynamics of stream invertebrate structural and functional beta diversity. *Ecography* 43, 620-635.
- Datry, T., Fritz, K., Leigh, C., 2016. Challenges, developments and perspectives in intermittent river ecology. *Freshwater Biology* 61, 1171-1180.
- Davey, A.J.H., Kelly, D.J., 2007. Fish community responses to drying disturbances in an intermittent stream: A landscape perspective. *Freshwater Biology* 52, 1719-1733.
- Deiner, K., Walser, J.-C., Mächler, E., Altermatt, F., 2015. Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biological Conservation* 183, 53-63.
- Doledec, S., Tilbian, J., Bonada, N., 2017. Temporal variability in taxonomic and trait compositions of invertebrate assemblages in two climatic regions with contrasting flow regimes. *Sci Total Environ* 599-600, 1912-1921.
- Effendi, H., 2016. River Water Quality Preliminary Rapid Assessment Using Pollution Index. *Procedia Environmental Sciences* 33, 562-567.
- Elbrecht, V., Vamos, E.E., Meissner, K., Aroviita, J., Leese, F., Yu, D., 2017. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution* 8, 1265-1275.
- Elbrecht, V., Vamos, E.E., Steinke, D., Leese, F., 2018. Estimating intraspecific genetic diversity from community DNA metabarcoding data. *PeerJ* 6, e4644.
- Erdozain, M., Thompson, D.G., Porter, T.M., Kidd, K.A., Kreutzweiser, D.P., Sibley, P.K., Swystun, T., Chartrand, D., Hajibabaei, M., 2019. Metabarcoding of storage ethanol vs. conventional

morphometric identification in relation to the use of stream macroinvertebrates as ecological indicators in forest management. *Ecological Indicators* 101, 173-184.

Evans, N., Paulay, G., 2012. DNA Barcoding Methods for Invertebrates, In *DNA Barcodes: Methods and Protocols*. eds W.J. Kress, D.L. Erickson, pp. 47-77. Humana Press, Totowa, NJ.

Fredricks, D.N., Smith, C., Meier, A., 2005. Comparison of six DNA extraction methods for recovery of fungal DNA as assessed by quantitative PCR. *J Clin Microbiol* 43, 5122-5128.

Fuller, M.R., Doyle, M.W., Strayer, D.L., 2015. Causes and consequences of habitat fragmentation in river networks. *Ann N Y Acad Sci* 1355, 31-51.

Gallart, F., Cid, N., Latron, J., Llorens, P., Bonada, N., Jeuffroy, J., Jiménez-Argudo, S.-M., Vega, R.-M., Solà, C., Soria, M., Bardina, M., Hernández-Casahuga, A.-J., Fidalgo, A., Estrela, T., Munné, A., Prat, N., 2017. TREHS: An open-access software tool for investigating and evaluating temporary river regimes as a first step for their ecological status assessment. *Science of the Total Environment* 607-608, 519-540.

García-Girón, J., Heino, J., García-Criado, F., Fernández-Aláez, C., Alahuhta, J., 2020. Biotic interactions hold the key to understanding metacommunity organisation. *Ecography*, 1-11.

Gough, P., Fernández Garrido, P., Van Herk, J., 2017. Dam Removal. A viable solution for the future of our European rivers. *Dam Removal Europe*, 73.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols* 8, 1494-1512.

Hajibabaei, M., Spall, J.L., Shokralla, S., van Konynenburg, S., 2012. Assessing biodiversity of a freshwater benthic macroinvertebrate community through nondestructive environmental barcoding of DNA from preservative ethanol.

Hazkani-Covo, E., Zeller, R.M., Martin, W., 2010. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet* 6, e1000834.

Heino, J., Alahuhta, J., Ala-Hulkko, T., Antikainen, H., Bini, L.M., Bonada, N., Datry, T., Erős, T., Hjort, J., Kotavaara, O., Melo, A.S., Soininen, J., 2017. Integrating dispersal proxies in ecological and environmental research in the freshwater realm. *Environmental Reviews* 25, 334-349.

Heino, J., Melo, A.S., Siqueira, T., Soininen, J., Valanko, S., Bini, L.M., 2015. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Freshwater Biology* 60, 845-869.

Hershey, A.E., Pastor, J., Peterson, B.J., Kling, G.W., 2010. Stable Isotopes Resolve the Drift Paradox for *Baetis* Mayflies in an Arctic River *Stable Ecological Society of America* 74, 2315-2325.

- Jabot, F., Laroche, F., Massol, F., Arthaud, F., Crabot, J., Dubart, M., Blanchet, S., Munoz, F., David, P., Datry, T., 2020. Assessing metacommunity processes through signatures in spatiotemporal turnover of community composition. *Ecology Letters*.
- Jacobsen, D., 1998. The effect of organic pollution on the macroinvertebrate fauna of Ecuadorian highland streams. *Fundamental and Applied Limnology* 143, 179-195.
- Jaeger, K.L., Olden, J.D., Pelland, N.A., 2014. Climate change poised to threaten hydrologic connectivity and endemic fishes in dryland streams. *Proceedings of the National Academy of Sciences of the United States of America* 111, 13894-13899.
- Kalogianni, E., Vourka, A., Karaouzas, I., Vardakas, L., Laschou, S., Skoulikidis, N.T., 2017. Combined effects of water stress and pollution on macroinvertebrate and fish assemblages in a Mediterranean intermittent river. *Science of the Total Environment* 603-604, 639-650.
- Karaouzas, I., Theodoropoulos, C., Vourka, A., Gritzalis, K., Skoulikidis, N.T., 2019. Stream invertebrate communities are primarily shaped by hydrological factors and ultimately fine-tuned by local habitat conditions. *Sci Total Environ* 665, 290-299.
- Kemp, P.S., O'Hanley, J.R., 2010. Procedures for evaluating and prioritising the removal of fish passage barriers: a synthesis. *Fisheries Management and Ecology* 17, 297-322.
- Kjærandsen, J., 2005. Species assemblages and community structure of adult caddisflies along a headwater stream in southeastern Ghana (Insecta: Trichoptera). *Biodiversity and Conservation* 14, 1-43.
- Li, C., Hofreiter, M., Straube, N., Corrigan, S., Naylor, G.J., 2013. Capturing protein-coding genes across highly divergent species. *Biotechniques* 54, 321-326.
- Li, F., Tonkin, J.D., Haase, P., 2018. Dispersal capacity and broad-scale landscape structure shape benthic invertebrate communities along stream networks. *Limnologia* 71, 68-74.
- Liu, S., Wang, X., Xie, L., Tan, M., Li, Z., Su, X., Zhang, H., Misof, B., Kjer, K.M., Tang, M., Niehuis, O., Jiang, H., Zhou, X., 2016. Mitochondrial capture enriches mito-DNA 100 fold, enabling PCR-free mitogenomics biodiversity analysis. *Mol Ecol Resour* 16, 470-479.
- Lytle, D.A., Poff, N.L., 2004. Adaptation to natural flow regimes. *Trends Ecol Evol* 19, 94-100.
- Macneale, K.H., Peckarsky, B.L., Likens, G.E., 2005. Stable isotopes identify dispersal patterns of stonefly populations living along stream corridors. *Freshwater Biology* 50, 1117-1130.
- Mariac, C., Vigouroux, Y., Duponchelle, F., Garcia-Davila, C., Nunez, J., Desmarais, E., Renno, J.F., 2018. Metabarcoding by capture using a single COI probe (MCSP) to identify and quantify fish species in ichthyoplankton swarms. *PLoS One* 13, e0202976.
- Martins, F.M.S., Galhardo, M., Filipe, A.F., Teixeira, A., Pinheiro, P., Pauperio, J., Alves, P.C., Beja, P., 2019. Have the cake and eat it: Optimizing nondestructive DNA metabarcoding of macroinvertebrate samples for freshwater biomonitoring. *Mol Ecol Resour* 19, 863-876.
- Mondy, C.P., Villeneuve, B., Archaimbault, V., Usseglio-Polatera, P., 2012. A new macroinvertebrate-based multimetric index (I 2M 2) to evaluate ecological quality of French wadeable

streams fulfilling the WFD demands: A taxonomical and trait approach. *Ecological Indicators* 18, 452-467.

Munch, K., Boomsma, W., Huelsenbeck, J.P., Willerslev, E., Nielsen, R., 2008. Statistical assignment of DNA sequences using Bayesian phylogenetics. *Syst Biol* 57, 750-757.

Nichols, S.J., Kefford, B.J., Campbell, C.D., Bylemans, J., Chandler, E., Bray, J.P., Shackleton, M., Robinson, K.L., Carew, M.E., Furlan, E.M., 2019. Towards routine DNA metabarcoding of macroinvertebrates using bulk samples for freshwater bioassessment: Effects of debris and storage conditions on the recovery of target taxa. *Freshwater Biology* 65, 607-620.

Oenema, O., Witzke, H.P., Klimont, Z., Lesschen, J.P., Velthof, G.L., 2009. Integrated assessment of promising measures to decrease nitrogen losses from agriculture in EU-27. *Agriculture, Ecosystems & Environment* 133, 280-288.

Ohlmann, M., 2019. Vers une théorie spatiale des réseaux d'interaction en écologie : méthodes, concepts et applications. Université Grenoble Alpes (ComUE).

Padial, A.A., Ceschin, F., Declerck, S.A., De Meester, L., Bonecker, C.C., Lansac-Toha, F.A., Rodrigues, L., Rodrigues, L.C., Train, S., Velho, L.F., Bini, L.M., 2014. Dispersal ability determines the role of environmental, spatial and temporal drivers of metacommunity structure. *PLoS One* 9, e111227.

Paijmans, J.L., Fickel, J., Courtiol, A., Hofreiter, M., Forster, D.W., 2016. Impact of enrichment conditions on cross-species capture of fresh and degraded DNA. *Mol Ecol Resour* 16, 42-55.

Pereira-da-Conceicao, L., Elbrecht, V., Hall, A., Briscoe, A., Barber-James, H., Price, B., 2019. Metabarcoding unsorted kick-samples facilitates macroinvertebrate-based biomonitoring with increased taxonomic resolution, while outperforming environmental DNA.

Petersen, I., Masters, Z., Hildrew, A.G., Ormerod, S.J., 2004. Dispersal of adult aquatic insects in catchments of differing land use. *Journal of Applied Ecology* 41, 934-950.

Phillipsen, I.C., Lytle, D.A., 2013. Aquatic insects in a sea of desert: population genetic structure is shaped by limited dispersal in a naturally fragmented landscape. *Ecography* 36, 731-743.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegard, K.L., Richter, B.D., Sparks, R.E., Stromberg, J.C., 1997. The Natural Flow Regime. *Bioscience* 47, 769-784.

Portik, D.M., Smith, L.L., Bi, K., 2016. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). *Mol Ecol Resour* 16, 1069-1083.

Radinger, J., Essl, F., Holker, F., Horky, P., Slavik, O., Wolter, C., 2017. The future distribution of river fish: The complex interplay of climate and land use changes, species dispersal and movement barriers. *Glob Chang Biol* 23, 4970-4986.

Radinger, J., Hölker, F., Horký, P., Slavík, O., Dendoncker, N., Wolter, C., 2016. Synergistic and antagonistic interactions of future land use and climate change on river fish assemblages. *Global Change Biology* 22, 1505-1522.

- Ratnasingham, S., Hebert, P.D.N., 2007. BOLD : The Barcode of Life Data System. *Molecular Ecology Notes* 7, 355-364.
- Reid, A.J., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T.J., Kidd, K.A., MacCormack, T.J., Olden, J.D., Ormerod, S.J., Smol, J.P., Taylor, W.W., Tockner, K., Vermaire, J.C., Dudgeon, D., Cooke, S.J., 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews* 94, 849-873.
- Rodeles, A.A., Galicia, D., Miranda, R., 2020. Barriers to longitudinal river connectivity: review of impacts, study methods and management for Iberian fish conservation. *Limnetica* 39, 601-619.
- Rooney, J.P., Ryde, I.T., Sanders, L.H., Howlett, E.H., Colton, M.D., Germ, K.E., Mayer, G.D., Greenamyre, J.T., Meyer, J.N., 2015. PCR based determination of mitochondrial DNA copy number in multiple species. *Methods Mol Biol* 1241, 23-38.
- RUISSEAUX, L., 2004-2009. Résumé illustré du programme LIFE Ruisseaux de Têtes de Bassins et Faune Patrimoniales Associées - 2004-2009, p. 36.
- Sarremejane, R., Cid, N., Datry, T., Stubbington, R., Alp, M., Cañedo-Argüelles, M., Cordero-Rivera, A., Csabai, Z., Gutiérrez-Cánovas, C., Heino, J., Forcellini, M., Millán, A., Paillex, A., Pařil, P., Polářek, M., de Figueroa, J.M.T., Usseglio-Polatera, P., Zamora-Muñoz, C., Bonada, N., 2020. DISPERSE: A trait database to assess the dispersal potential of aquatic macroinvertebrates. In review.
- Sarremejane, R., Mykrä, H., Bonada, N., Aroviita, J., Muotka, T., 2017. Habitat connectivity and dispersal ability drive the assembly mechanisms of macroinvertebrate communities in river networks. *Freshwater Biology* 62, 1073-1082.
- Shearer, K.A., Hayes, J.W., Stark, J.D., 2002. Temporal and spatial quantification of aquatic invertebrate drift in the Maruia River, South Island, New Zealand. *New Zealand Journal of Marine and Freshwater Research* 36, 529-536.
- Snell, M.A., Barker, P.A., Surridge, B.W.J., Large, A.R.G., Jonczyk, J., Benskin, C.M.W.H., Reaney, S., Perks, M.T., Owen, G.J., Cleasby, W., Deasy, C., Burke, S., Haygarth, P.M., 2014. High frequency variability of environmental drivers determining benthic community dynamics in headwater streams. *Environmental Sciences: Processes and Impacts* 16, 1629-1636.
- Stubbington, R., Chadd, R., Cid, N., Csabai, Z., Miliřa, M., Morais, M., Munné, A., Pařil, P., Peřić, V., Tziortzis, I., Verdonschot, R.C.M., Datry, T., 2018. Biomonitoring of intermittent rivers and ephemeral streams in Europe: Current practice and priorities to enhance ecological status assessments. *Science of The Total Environment* 618, 1096-1113.
- Stubbington, R., Greenwood, A.M., Wood, P.J., Armitage, P.D., Gunn, J., Robertson, A.L., 2009. The response of perennial and temporary headwater stream invertebrate communities to hydrological extremes. *Hydrobiologia* 630, 299-312.
- Tonkin, J.D., Altermatt, F., Finn, D.S., Heino, J., Olden, J.D., Pauls, S.U., Lytle, D.A., 2018. The role of dispersal in river network metacommunities: Patterns, processes, and pathways. *Freshwater Biology* 63, 141-163.

Tonkin, J.D., Bogan, M.T., Bonada, N., Rios-Touma, B., Lytle, D.A., 2017. Seasonality and predictability shape temporal species diversity. *Ecology* 98, 1201-1216.

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

Vallender, E.J., 2011. Expanding whole exome resequencing into non-human primates. *Genome Biology* 12, R87.

Vamos, E., Elbrecht, V., Leese, F., 2017. Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics* 1.

Vander Vorste, R., Corti, R., Sagouis, A., Datry, T., 2016. Invertebrate communities in gravel-bed, braided rivers are highly resilient to flow intermittence. *Freshwater Science* 35, 164-177.

Viana, D. S., & Chase, J. M. (2019). Spatial scale modulates the inference of metacommunity assembly processes. *Ecology*, 100(2), e02576. doi:10.1002/ecy.2576

Wilcox, T.M., Zarn, K.E., Piggott, M.P., Young, M.K., McKelvey, K.S., Schwartz, M.K., 2018. Capture enrichment of aquatic environmental DNA: A first proof of concept. *Mol Ecol Resour* 18, 1392-1401.

Yang, C., Ji, Y., Wang, X., Yang, C., Yu, D.W., 2013. Testing three pipelines for 18S rDNA-based metabarcoding of soil faunal diversity. *Sci China Life Sci* 56, 73-81.

Zizka, V.M.A., Geiger, M.F., Leese, F., 2020. DNA metabarcoding of stream invertebrates reveals spatio-temporal variation but consistent status class assessments in a natural and urban river. *Ecological Indicators* 115.