

2020-~~XXX~~-MISTRAL-UMR5023- Metabarcoding pour le Suivi des opérations de Restauration de l'ALbarine

Rapport final

Action n°~~XXX~~ du Programme 2010 au titre de l'accord cadre
Agence de l'Eau ZABR

RESPONSABLE SCIENTIFIQUE DU PROJET
Tristan Lefébure (LEHNA, UMR 5023, équipe E3S)

REDACTION DU RAPPORT
Maïlys Gauthier (LEHNA, UMR 5023, équipe E3S)

EQUIPES DE RECHERCHES ZABR CONCERNEES ET PERSONNELS IMPLIQUES
- LEHNA E3S : Lefébure Tristan, Konecny Lara, Gauthier Maïlys, Malard Florian, Douady Christophe
- INRAE EcoFlowsS : Datry Thibault, Forcellini Maxence, Launay Bertrand, Le Goff Guillaume

CONTACT SCIENTIFIQUE DE L'EQUIPE
- LEHNA E3S : Lefébure Tristan
- INRAE EcoFlowsS : Datry Thibault

AUTRES PARTENAIRES
• Recherche : Arnaud Foulquier du LECA (Laboratoire d'Écologie Alpine, UMR 5553)
• Institutionnel : Florent Pellizzaro puis Samuel Monnet du SR3A

THEME DE RATTACHEMENT ZABR
Flux forme habitat biocénoses

THEME DE RATTACHEMENT AGENCE DE L'EAU -QUESTIONS AGENCE DE L'EAU
Thème de rattachement Agence de l'Eau (accord cadre)
Thème : La protection, la restauration des milieux et les gains écologiques
Thématique : Restauration, fonctionnement physique
Inventaire des besoins en matière de connaissance opérationnelle pour la gestion des milieux
aquatiques :
3.1.1 – renforcer la panoplie de l'ingénierie écologique
2.1 – Compléter et optimiser la surveillance de l'état et la qualité des masses d'eau

SITE OU OBSERVATOIRE DE RATTACHEMENT ZABR
SARAM

REFERENT AGENCE DE L'EAU
Stéphane Stroffek

TABLE DES MATIERES

TABLE DES MATIERES	1
1 CONTEXTE GÉNÉRAL DU PROJET	2
2 OBJECTIFS DU PROJET.....	2
3 PROBLÉMATIQUES & PHASES DU PROJET.....	3
Problématique 1 : Extraction ADN	3
Problématique 2 : Enrichissement.....	3
Problématique 3 : Applicabilité et transférabilité au suivi des ORE	4
Phases du projet.....	4
Livrables	5
4 MÉTHODES & RÉSULTATS	6
PHASE 1 : État des lieux des études et protocoles de métabarcoding.....	6
PHASE 2 : Quel protocole d'extraction ADN pour les échantillons de Surber ?	6
Test in vitro sur des communautés artificielles – Méthodes	6
Test in vitro sur des communautés artificielles – Résultats.....	12
PHASE 3 : Applicabilité d'un protocole de métabarcoding au suivi ORE	20
Test in situ – Méthodes	20
Test in situ – Résultats.....	24
Test in situ – Conclusion.....	32
PHASE 4 : Transférabilité aux gestionnaires	38
5 ECHANGES ENTRE EQUIPES IMPLIQUEES & ECHANGES AVEC LES REFERENTS AGENCE DE L'EAU.....	39
6 REFERENCES BIBLIOGRAPHIQUES	40
7 ANNEXES	42
ANNEXE 1 – Tableau des tests post-hoc - Test in vitro sur des communautés artificielles	42
ANNEXE 2 – Protocole TRI ET DETERMINATION DU MACROZOOBENTHOS / MODIFIE DE NF T90-388 (NORME I2M2) ...	46
ANNEXE 3 – Protocole final pour le suivi des communautés de macrozoobenthos	51
ANNEXE 4 – Treemaps pour les Mollusca et Malacostraca	65
ANNEXE 5 – Comparaison des listes taxonomiques morphologiques et moléculaires à l'échelle du prélèvement	66

1 CONTEXTE GÉNÉRAL DU PROJET

Les opérations de restauration écologique (ORE) des cours d'eau se multiplient, notamment sur les nombreux affluents du Rhône tels que l'Albarine. L'efficacité de ces actions est aujourd'hui difficile à évaluer de par (1) la méconnaissance des états initiaux, (2) la forte variabilité associée à des contextes bruités par de multiples facteurs, et enfin (3) la dynamique temporelle des ORE. De nos jours, les ORE sont évaluées grâce à des mesures de biodiversité (p.ex. communautés de zoobenthos) qui dépendent de méthodes d'identification morphologique qui sont à la fois peu sensibles (Dolédéc et al 2015) et qui nécessitent un fort effort d'échantillonnage (Vaudor et al 2015) et sont donc onéreuses. De plus, ces mesures de biodiversité classiques apportent une vision incomplète de la biodiversité car elles ne prennent pas en compte des groupes ubiquistes difficilement identifiables morphologiquement (e.g. diptères, acariens, oligochètes) mais qui pourraient se révéler être de très bons indicateurs (Cañedo et al 2016). L'identification moléculaire de communauté entière (le métabarcoding) pourrait lever ces verrous en apportant de la précision taxonomique mais également en réduisant les temps et les coûts d'identification et des suivis, permettant ainsi des suivis en routine plus longs et plus fréquents, avant et après restauration (Ji et al 2013). De plus le métabarcoding peut s'appliquer sur les mêmes échantillons que les méthodes classiques (p.ex. Surber de l'I2M2) permettant à la fois de maintenir une continuité dans les protocoles d'échantillonnage mais également de s'affranchir des problèmes liés à l'utilisation des ADN environnementaux et leur transfert dans les hydrosystèmes. Néanmoins, il n'existe pas aujourd'hui de protocole de métabarcoding en cours d'eau standardisé et opérationnel pour des échantillons de Surber.

2 OBJECTIFS DU PROJET

Dans ce projet, nous avons proposé de tester sur le bassin de l'Albarine différents protocoles de métabarcoding et de les confronter aux approches morphologiques classiques afin d'évaluer la faisabilité et l'apport des différents protocoles moléculaires pour le suivi des communautés de macrozoobenthos. Plus précisément, nous avons souhaité (1) déterminer quelles méthodes moléculaires sont les plus efficaces pour décrire les communautés du macrozoobenthos de l'Albarine et (2) identifier un/des protocole(s) efficace(s), opérationnel(s) et facilement transférable(s) au suivi de l'Albarine, y compris dans un contexte ORE. L'objectif premier du projet était de mettre en place un outil efficace et opérationnel pour améliorer les estimations de biodiversité aquatique et permettre des suivis temporels à long terme des opérations de restauration, et, dans le cadre de l'Albarine, d'évaluer l'efficacité des ORE passées et en cours sur l'Albarine avec un grain beaucoup plus fin qu'avec les données I2M2 actuellement utilisées et qui sont malheureusement peu informatives dans le suivi ORE et dans un contexte d'intermittence d'écoulement.

3 PROBLÉMATIQUES & PHASES DU PROJET

Ce projet se focalise sur un type d'échantillon normalisé : les Surber tels qu'ils sont utilisés dans la norme I2M2. Ces échantillons sont fixés à l'alcool et contiennent une part minérale (sable, gravier) et une part organique – elle-même composée d'organismes et de matière organique particulaire (litière, spermaphytes, etc...). Le protocole de métabarcoding comporte plusieurs étapes : extraction ADN, enrichissement du marqueur d'identification moléculaire, séquençage et traitement bio-informatique (Liu et al 2019). Plusieurs protocoles peuvent être établis pour chacune de ses étapes. Ce projet se concentre initialement sur les deux premières étapes où les choix des protocoles peuvent sensiblement affecter la description des communautés de macrozoobenthos (Elbrecht et al 2017).

Problématique 1 : Extraction ADN

Plusieurs matrices sont envisageables pour l'extraction de l'ADN des échantillons Surber :

- **le « tri-pool-broyage » ou bulk.** Il correspond à la matrice ADN la plus utilisée et consiste à sortir les organismes de l'échantillon puis à les grouper, les broyer et extraire les ADN de ce broyat (p.ex. Elbrecht & Steinke 2019). Cette méthode a l'avantage d'être robuste car elle élimine les potentiels inhibiteurs des réactions en aval mais elle nécessite un travail conséquent de tri en amont qui augmente fortement les coûts et les risques de contamination ADN entre échantillons. De plus, elle implique la destruction des organismes échantillonnés.

- **le bulk in toto.** Contrairement au bulk, le principe ici est de broyer l'intégralité de l'échantillon (organismes, débris végétaux, sable, etc...) avant d'en extraire l'ADN (Majaneva et al 2018). Il peut être difficile à mettre en œuvre quand la part non-organique de l'échantillon est importante ou quand les inhibiteurs de réaction sont abondants (p.ex. acide humique). Bien que plus rapide, économe, diminuant les biais opérateur.ices et préservant des contaminations ADN inter-échantillons, ce protocole implique également la destruction des organismes.

- **l'éthanol de conservation.** Le principe est d'extraire l'ADN se trouvant en solution dans l'éthanol de conservation des échantillons. Cette méthode a l'avantage d'être rapide, économe et permet de conserver tous les organismes d'un échantillon pour d'éventuelles futures analyses (Gauthier et al 2020). Néanmoins, les ADN de l'éthanol sont généralement de moins bonne qualité, en faible concentration et des inhibiteurs de réactions peuvent aussi se trouver dans l'éthanol de conservation.

Problématique 2 : Enrichissement

Les extractions d'ADN (obtenues à partir de l'une des trois matrices précédentes) doivent ensuite être enrichies en marqueurs d'identification moléculaire (ici, le gène de la COI pour les animaux). Il existe à ce jour deux méthodes d'enrichissement :

- **la PCR (Polymerase Chain Reaction)**. C'est la méthode classique d'enrichissement en métabarcoding. Elle a l'avantage d'être simple, rapide et peu coûteuse, mais présente le défaut de montrer des biais d'amplification en faveur de certains taxons suivant le couple d'amorces utilisé (Elbrecht & Leese 2015). Ainsi certaines espèces seront moins bien, voire non, amplifiées. Par conséquent, elle peut être non exhaustive et ne permet pas d'estimer de manière fiable les abondances relatives des taxons.

- **la capture par sondes**. La capture utilise des sondes ARN qui s'hybrident aux fragments d'intérêt. Il est possible de synthétiser un grand nombre de sondes différentes et de grande taille (120 bases) permettant de capturer une grande diversité de fragments (et donc d'espèces). Ainsi, les biais spécifiques sont largement diminués et les abondances relatives sont mieux estimées malgré des optimisations nécessaires (Gauthier et al, 2020).

Problématique 3 : Applicabilité et transférabilité au suivi des ORE

Après la mise en place d'un protocole de métabarcoding en laboratoire, il est important de tester son efficacité sur des échantillons récoltés *in situ*. A l'heure actuelle, ceci est encore peu expérimenté.

Le bassin de l'Albarine est un bassin très suivi par l'équipe de recherche EcoFlowS (UR RiverLy, INRAE Lyon-Villeurbanne) : suivi depuis 2007, campagne I2M2 (2017-2018), projet AFB sur l'intermittence (2016-2020) et projet Horizon DRYvER (2021-...). Ces suivis permettent d'avoir une connaissance très fine de la diversité taxonomique présente sur ce bassin. De plus, il existe trois ORE en cours sur le bassin. L'Albarine est donc idéal pour tester différents protocoles de métabarcoding sur Surber et évaluer leur applicabilité pour le suivi d'ORE.

Phases du projet

Le projet s'est articulé en différentes phases :

La première phase a consisté en l'état des lieux des études et protocoles de métabarcoding pouvant être utilisés dans le cas des communautés du macrozoobenthos échantillonnées par Surber.

La deuxième phase a été méthodologique et a eu pour objectif de tester les différents protocoles d'extraction ADN élaborés d'après la littérature.

La troisième phase a consisté en l'application du protocole d'extraction ADN sélectionné en phase 2 sur des échantillons prélevés *in situ* sur l'Albarine. Différentes méthodes d'enrichissement sélectionnés d'après la littérature devaient être évaluées durant cette phase. Les résultats du métabarcoding ont été comparés à ceux de l'identification sur critères morphologique sous forme de rapport (Livrable 3).

La quatrième phase portait sur la transférabilité aux gestionnaires avec la production d'un rapport technique sur le protocole détaillé de métabarcoding utilisé dans le projet (Livrable 1) et une/des journée(s) de formation à l'intention des gestionnaires (Livrable 2).

Livrables

Livrable 1 : ce livrable correspond à l'ANNEXE 3 du présent rapport (disponible aussi sans en-tête dans le dossier DocSupp_Rapp_final joint). Il s'agit du protocole détaillé mis en place lors de ce projet, en français, comportant les préconisations pour le terrain à l'ensemble du traitement des échantillons en laboratoire avant séquençage.

Livrable 2 : ce livrable a été rendu sous deux formes : une demi-journée d'échanges et d'introduction aux méthodes moléculaires à la SR3A le 31/08/2021 et un café ZABR organisé par le GRAIE avec plus de 100 participants le 22/09/2023.

Livrable 3 : ce livrable correspond à la partie « PHASE 3 : Applicabilité d'un protocole de métabarcoding au suivi ORE » du présent rapport. Il s'agit de l'analyse comparative de l'approche moléculaire et de l'approche morphologique sur des échantillons prélevés *in situ* dans l'Albarine.

4 MÉTHODES & RÉSULTATS

PHASE 1 : État des lieux des études et protocoles de métabarcoding

Cette phase a permis de sélectionner plusieurs protocoles pouvant être utilisés sur les échantillons de Surber. La recherche portait sur l'ADN environnemental de l'eau et des sols, l'ADN *etoh*, l'ADN *bulk* et *bulk in toto* mais aussi sur les différentes méthodes d'enrichissement.

A la suite de cet état des lieux, il a été décidé de ne pas poursuivre le travail sur les méthodes d'enrichissement (problématique 2). En effet, la littérature s'intéresse peu à la capture par sondes malgré les avantages qu'elle peut apporter en comparaison à l'amplification par PCR. Ainsi, travailler sur la capture par sondes ARN demandait un investissement en temps et argent que le projet MISTRAL ne permettait pas pour optimiser cette méthodologie. De plus, étant donné qu'elle est coûteuse et peu appliquée, son utilisation n'était pas pertinente dans le cadre de ce projet, cherchant à démocratiser un outil moléculaire pour les gestionnaires.

PHASE 2 : Quel protocole d'extraction ADN pour les échantillons de Surber ?

Test *in vitro* sur des communautés artificielles – Méthodes

Un enjeu majeur concernant les échantillons de Surber est la présence de la matière « non désirée » dans les échantillons (sable, graviers, litière, etc...) qui peuvent représenter une part conséquente de l'échantillon. La matrice *bulk* consistant en la part « organismes » uniquement de l'échantillon a bien été étudiée et ne demande pas de nouvelles techniques. Ce qui n'est pas le cas pour la matrice *ethanol* et la matrice *bulk in toto*. A partir de la phase 1, des protocoles adaptés à la matrice *etoh* et à la matrice *bulk in toto* ont été élaborés. Les différents protocoles et leur coûts et temps associés sont disponibles dans le dossier DocSupp_Rapp_final joint.

Pour évaluer ces différents protocoles, des communautés artificielles ont été conçues contenant 10 espèces du macrozoobenthos et différents substrats présents dans les échantillons Surber (Tableau 1, Figure 1). Les espèces ont été choisies pour représenter un large panel taxonomique et des traits morphologiques divers et sont des espèces facilement identifiables à l'œil nu ou pour lesquelles aucune confusion avec des espèces proches n'était possible lors de la collecte (p.ex. seule espèce connue sur une station donnée d'après un suivi de plusieurs années) (Tableau 2, Figure 2).



Figure 1. Substrats ajoutés dans les communautés artificielles. (a) Algues, (b) Sable, (c) Graviers, (d) Bryophytes et (e) Litière + vase.

Tableau 1. Date et lieu de collecte pour les substrats ajoutés dans les communautés artificielles

Type	Date collecte	Lieu collecte
Algue	25/02/2021	Canaux LEHNA
Sable	25/02/2021	Rizan - Miribel
Graviers	25/02/2021	Rizan - Miribel
Bryophytes	25/02/2021	Albarine - Brénod
Vase + Litière	25/02/2021	Etang LEHNA

Z A B R

Zone Atelier Bassin du Rhône

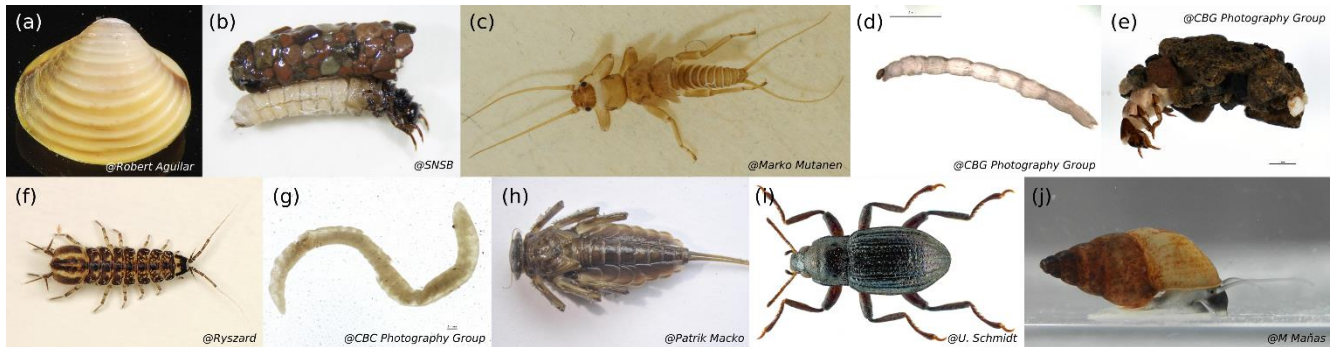






Figure 2. Espèces composant les communautés artificielles. (a) *Corbicula fulminea*, (b) *Drusus annulatus*, (c) *Brachyptera risi*, (d) *Chironomus riparius*, (e) *Glossossoma conforme*, (f) *Asellus aquaticus*, (g) *Tubifex tubifex*, (h) *Epeorus assimilis*, (i) *Riolus subviolaceus*, (j) *Potamopyrgus antipodarum*.

Tableau 2 : Espèces utilisées pour les communautés artificielles

Ordre	Espèces	Date collecte	Lieu collecte	Biomasse sèche de 10 organismes (mg)	Traits (taille - sclérisation)
Trichoptera	<i>Glossossoma conforme</i>	17/02/21	Fontaine noire	47.1	moyenne - tête sclérisée/ fourreau / forme en étui
Coleoptera	<i>Riolus subviolaceus</i>	17/02/21	Charabotte	0.9	petite - fortement sclérisée
Ephemeroptera	<i>Epeorus assimilis</i>	19/02/21	Valserine - Lelex	120.2	grande - sclérisée
Trichoptera	<i>Drusus annulatus</i>	19/02/21	Brénod	7.13	moyenne - tête sclérisée/ fourreau / forme en étui
Plecoptera	<i>Brachyptera risi</i>	19/02/21	Brénod	27.23	moyenne - sclérisée
Haplotaxida	<i>Tubifex tubifex</i>	18/02/21	Elevage - GREBIL SARL	3.5	petite - corps mou
Isopoda	<i>Asellus aquaticus</i>	19/02/21	Étang LEHNA	56.47	grande - sclérisée
Venerida	<i>Corbicula fulminea</i>	23/02/21	Pierre-Bénite	165.7	grande - mollusque à opercule / coquille / forme en étui
Littorinimorph a	<i>Potamopyrgus antipodarum</i>	23/02/21	Pierre-Bénite	NA	petite - mollusque à opercule / coquille / forme en étui
Diptera	<i>Chironomus riparium</i>	25/02/21	Elevage - INRAE (Ecotox)	1.7	petite/moyenne - tête sclérisée

Ces communautés permettent d'avoir de vrais réplicats et de pouvoir tester uniquement la qualité et l'efficacité des extractions. Les communautés ont été composées comme suit :

<p>1. les flacons ont été remplis avec 150 mL de liquide de conservation (TNES-Urea ou éthanol 96%)</p>		
<p>2. durant la semaine du 17/02/21 au 25/02/21, les individus ont été récupérés dans des cours d'eau (ou élevage) puis placés vivants dans les flacons (10 individus par espèce)</p>	<p>Vérification à la loupe de terrain</p>	<p>Communautés artificielles après l'ajout de toutes les espèces (25/02) pour le liquide de conservation éthanol 96 % (gauche) et TNES-Urea (droite)</p>
<p>3. le volume que représentait les organismes a été mesuré (~28 mL) et 1 volume de chaque substrat a été ajouté dans les communautés après tri des organismes visibles à l'œil nu des substrats (càd ~ 28 mL par substrat)</p>		
<p>4. le volume du liquide de conservation a été ajusté pour avoir un ratio de 2 : 1 entre le volume de liquide de conservation et le reste de l'échantillon (organisme + substrats)</p>		 <p>Communautés artificielles après l'ajout des substrats pour le liquide de conservation éthanol 96 % (gauche) et TNES-Urea (droite)</p>

Les communautés artificielles ont ensuite été laissées à incuber 2 semaines avant que leur ADN soit extrait. Différents protocoles ont été testés combinant des prétraitements et des extractions différents (Figure 3).

Les détails des protocoles et leur coûts et temps associés sont disponibles dans le dossier DocSupp_Rapp_final joint à ce rapport. Une communauté de composition macrozoobenthos identique mais sans substrat (c'est-à-dire *bulk*) a également été faite et a servi comme contrôle positif lors de l'évaluation des protocoles d'extraction. Des témoins négatifs (blancs de prétraitement et d'extraction) ont également été inclus dans l'expérience. La Figure 3 présente le plan expérimental, les méthodes utilisées pour l'évaluation des protocoles et leurs analyses statistiques associées.

L'évaluation de la performance des différents protocoles d'extraction a été faite à travers plusieurs mesures : (1) la quantité et la qualité de l'extraction via la mesure de la quantité d'ADN extrait (totale et rapportée au mg extrait permettant de mesurer le rendement) et via la mesure du ratio A280/260, A260/230 et [Qubit]/[Nanodrope] ; (2) le succès de PCR via l'amplification de plusieurs couples d'amorces (amplifiant différentes tailles de fragments, Tableau 3) et (3) la quantification spécifique via la mesure du nombre de copies 16S (total et rapporté au mg extrait permettant de mesurer le rendement) pour chaque espèce présente dans les échantillons (Tableau 4).

Pour chaque matrice ADN, des modèles linéaires (quantité d'ADN) ou des modèles mixtes linéaires (nombre de copies 16S) ont été utilisés pour déterminer s'il existait une différence statistiquement significative de la quantité d'ADN ou du nombre de copies 16S entre les prétraitements, les méthodes d'extraction ou les protocoles (prétraitements*méthodes d'extraction). Les modèles linéaires mixtes avec le protocole comme variable indépendante ont été construits pour chaque matrice ADN et pour chaque espèce (effet aléatoire sur la communauté). Les modèles linéaires mixtes avec le prétraitement ou la méthode d'extraction comme variable indépendante ont été construits pour toutes les espèces ensemble (effet aléatoire sur l'espèce). Les modèles significatifs ont été suivis de tests post hoc pour évaluer quels groupes différaient. Pour les analyses de prétraitement et d'extraction de l'ADN *bulk in toto*, des données ont été supprimées afin d'obtenir un plan expérimental complet (PMS pour le prétraitement, UREA et BUFFER pour la méthode d'extraction). De plus, les modèles de l'ADN *bulk in toto* ont uniquement comparé les protocoles de l'ADN *bulk in toto* à la référence *bulk*. Le lien entre biomasse des espèces dans l'échantillon et nombre de copies 16S a également été évalué. Pour l'ADN *etoh*, un modèle linéaire généralisé (famille binomiale) a été utilisé pour le tester. Pour l'ADN *bulk in toto*, un modèle linéaire mixte avec le protocole comme effet aléatoire sur l'intercept a été utilisé pour le tester. Enfin, la cohérence entre réplicats de protocole a été évaluée à l'aide du test de Levene.

La partie **Test in vitro sur des communautés artificielles – Résultats** ci-après présente les résultats. Un article scientifique pour valoriser cette expérience est en cours de rédaction.

Z A B R

Zone Atelier Bassin du Rhône

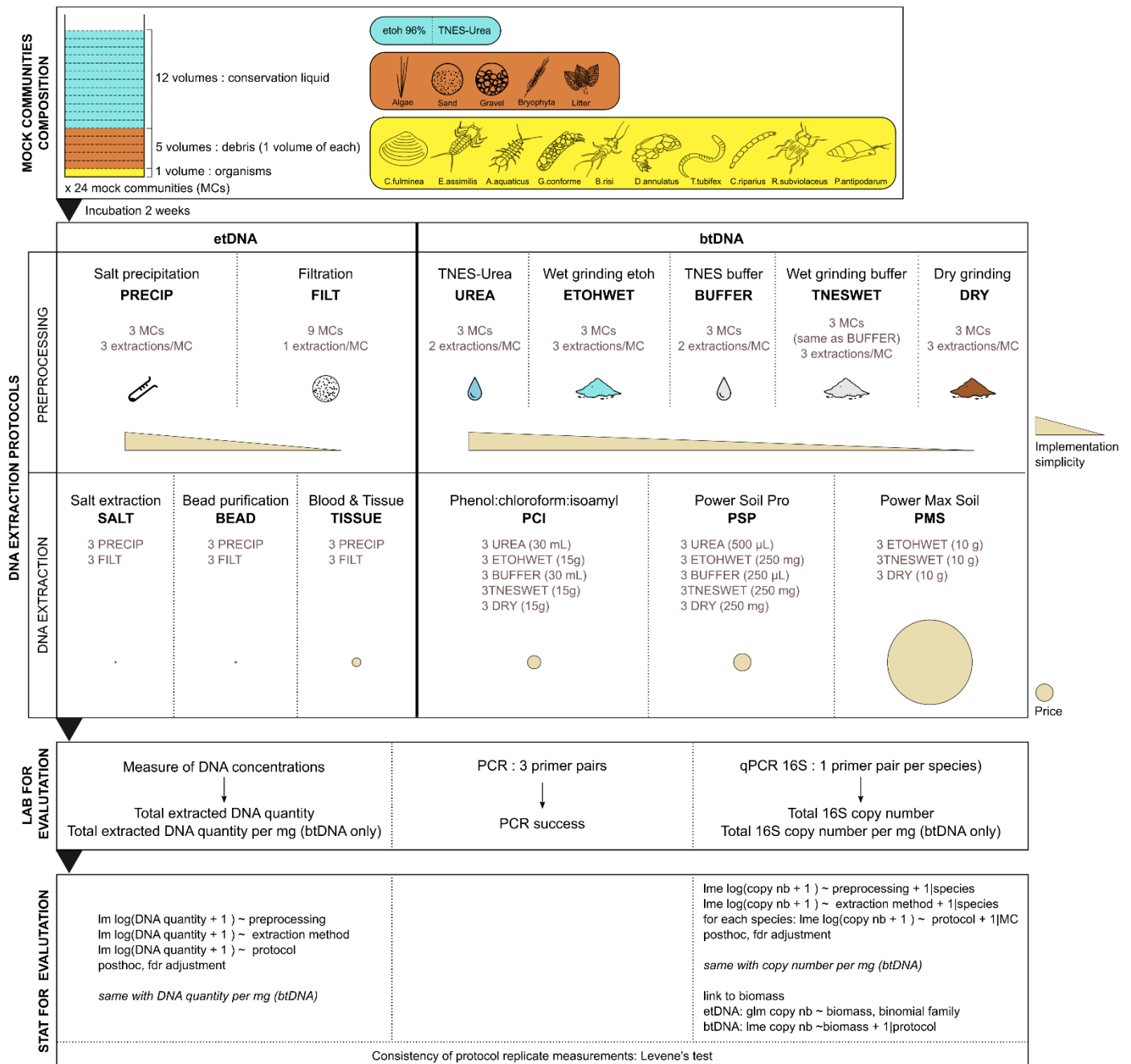


Figure 3. Résumé de l'expérience testant différents protocoles d'extraction ADN sur des communautés artificielles de la mise en place des communautés à l'analyse statistique. Chaque communauté est composée d'organismes, de substrats et de liquide de conservation (ethanol 96% ou TNES). Des blancs d'expérience et un contrôle positif (matrice Bulk) ont également été traités contenant du liquide de conservation et du liquide de conservation et des organismes, respectivement. La simplicité de mise en œuvre a été ajoutée pour le prétraitement uniquement, car c'est l'étape pour laquelle cette information est pertinente. Le prix a été ajouté à l'étape de la méthode d'extraction uniquement pour la même raison. Les détails de tous les protocoles ainsi que leur coûts et temps associés sont disponibles en anglais (extraction) et français (prétraitement) dans le dossier DocSupp_Rapp_final joint à ce rapport.

Test in vitro sur des communautés artificielles – Résultats

Quantité d'ADN extraite / Rendement d'extraction

Les ADN extraits ont été quantifiés avec un kit Qubit dsDNA et avec un Spectrophotomètre Nanodrop permettant aussi de mesurer les ratios A280/260, A260/230.

Dû à une concentration faible dans certains échantillons (<20 ng/μL) ou à la présence de contaminants résiduels dans l'éluat d'extraction (extraction PCI en particulier), les données issues du Nanodrop n'étaient exploitables que pour un nombre très faible d'échantillons et il a été décidé de ne pas exploiter ces données (concentration Nanodrop, ratios A280/260 et A260/230 et ratio [Qubit]/[Nanodrop]).

ADN etoh (ADNet)

Pour l'ADNet, aucune différence significative n'a été observé entre les traitements Filtration et Précipitation au sel (F=2.43, pvalue=0.14) ou les méthodes d'extraction (F= 2.22, pvalue=0.14) (Figure 4). Des différences significatives ont été observées entre 2 paires de protocoles : le protocole FILT*SALT significativement supérieur au protocole PRECIP*BEAD (t=3.79 ; pvalue=0.01) et au protocole PRECIP*SALT (t=4.25, pvalue=0.01). Le test de Levene n'a pas révélé de différences de variance significatives entre les protocoles (F=1,11, pvalue=0,4).

ADN bulk in toto (ADNbt)

Concernant les traitements, le broyage sec (DRY) et tampon (TNESWET) étaient significativement supérieurs aux trois autres traitements (UREA, ETOHWET and BUFFER, ANNEXE 1 Tableau A1).

Concernant les méthodes d'extractions, le kit PowerMaxSoil (PMS) a présenté des quantités extraites significativement plus élevée que le kit PowerSoilPro (PSP, t=-4.84, pvalue=0.0002) et que l'extraction Phenol-Chloro (PCI, t=-4.38, pvalue=0.0003).

La comparaison des différents protocoles au protocole de référence (ADN bulk) montre (i) que 3 protocoles étaient équivalents à la référence (TNESWET*PSP, DRY*PCI, DRY*PSP), (ii) que tous les protocoles PMS étaient supérieurs à la référence (figure 4) et (iii) que les protocoles restants étaient inférieurs (UREA*PCI, UREA*PSP, ETOHWET*PCI, ETOHWET*PSP, BUFFER*PCI, BUFFER*PSP) (Figure 4, ANNEXE 1 Tableau A2).

Rapporté au mg extrait, aucune différence significative entre les quantités d'ADN des différents traitements n'a été observée (ANNEXE 1 Tableau A1). Au niveau des méthodes d'extraction, les 2 kits d'extraction ne montraient plus de différences entre eux (t=-0.505, pvalue=0.62) mais l'extraction PCI restait inférieure aux deux kits (avec PMS : t=-4.5557, pvalue=0.0004 ; avec PSP : t=-4.051, pvalue=0.0007). Ainsi, rapporté au mg extrait, les deux kits d'extraction présentaient le même rendement et pouvaient être considérés comme équivalents. Finalement, les quantités ADN au mg extrait de tous les protocoles étaient significativement inférieurs à la référence (bulk en rouge, Figure 4, ANNEXE 1 Tableau A2).

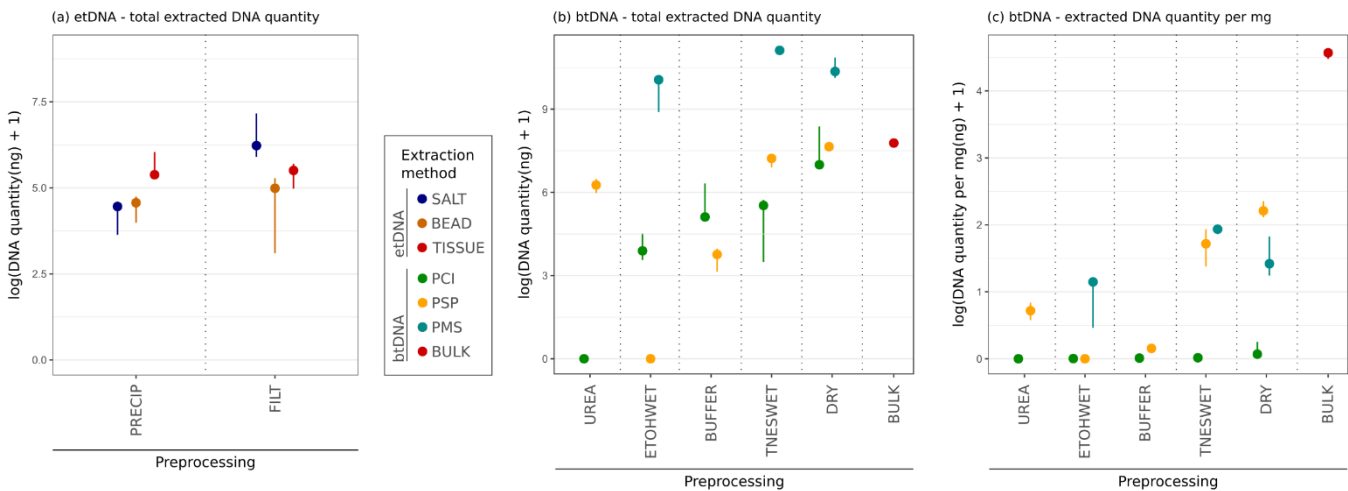


Figure 4. Log quantité minimale, médiane et maximale de l'ADN total extrait pour les différents protocoles utilisés pour l'ADNet (a), l'ADNbt (b) et rapportée par mg extrait pour l'ADNbt (c). (b) et (c) incluent également les mesures du protocole de référence bulk. Chaque protocole correspond à un prétraitement (x) associé à une méthode d'extraction (couleurs). La signification des abréviations des prétraitements et des méthodes d'extraction est disponible dans la Figure 3.

Inhibition des réactions en aval : PCR avec plusieurs couples d'amorces

Quatre couples d'amorces ciblant le gène COI couramment utilisés pour l'identification des invertébrés et amplifiant des fragments de tailles différentes (Tableau 3) ont été testés sur les échantillons. Les PCR ont été faites avec le kit Qiagen Multiplex PCR Plus Kit dans 25 μ L avec 0,50 μ M de chaque amorce et en suivant les protocoles d'amplification des auteurs.

Pour l'ADNet, tous les protocoles ont amplifié les trois amplicons ciblés, à l'exception des protocoles PRECIP*BEAD et FILT*SALT qui n'ont pas réussi à amplifier l'amplicon FwhF1/FwhR1 pour 1 des 3 réplicats. Pour l'ADNbt, 9 protocoles sur 13 ont amplifié avec succès tous leurs réplicats pour les trois amplicons. Les 4 protocoles restants étaient les suivants : UREA*PCI dont seulement 2/3 des réplicats ont amplifié BF2/BR2 et 1/3 a amplifié mlCOIintF/jgHCO2198 ; ETOHWET*PCI dont seulement 1/3 des réplicats ont amplifié BF2/BR2 ; ETOHWET*PSP qui n'a pas amplifié l'amplicon FwhF1/FwhR1, et BUFFER*PCI qui n'a pas amplifié FwhF1/FwhR1 pour 1 des 3 réplicats. Pour la référence bulk, les 3 réplicats ont amplifié avec succès les trois amplicons.

Tableau 3. Amorces COI utilisées pour les PCR.

Nom amorce Forward	Séquence oligonucléotidique (5' → 3')	Nom amorce Reverse	Séquence oligonucléotidique (5' → 3')	Longueur (paire de bases)	Littérature
BF2	GCHCCHGAYATRGCHTTYCC	BR2	TCDGGRTGNCCRAARAAYCA	421	Elbrecht & Leese, 2017
FwhF2	GGDACWGGWTGAACWGTWTAYCCHCC	EPTDr2n	CAAACAAATARDGGTATTCGDTY	171	Leese et al, 2021
FwhF1	YTCHACWAAYCAYAARGAYATYGG	FwhR1	ARTCARTTWCRAAHCHCC	142	Vamos et al, 2017
mIColintF	GGWACWGGWTGAACWGTWTAYCCYCC	HC02198	TAAACTTCAGGGTGACCAAAAAATCA	313	Leray et al, 2013

Quantification spécifique : mesure de l'ADN extrait des différentes espèces

Cette analyse avait pour objectif d'évaluer si des espèces sont plus ou moins détectées en fonction des protocoles d'extraction utilisés, de leurs traits morphologiques et de leur biomasse.

Un travail conséquent de plusieurs mois a été consacré à la conception d'amorces spécifiques à chaque espèce, c'est-à-dire la conception d'amorces qui amplifie uniquement une espèce de la communauté artificielle sans amplifier les autres. Plusieurs individus de chaque espèce ont d'abord été extraits puis une région du 16S a été amplifiée avec les amorces 16S rRNA universelles 16Sbr (5'-CCGGTCTGAACTCAGATCACGT-3') et 16Sar (5'-CGCCTGTTTATCAAAAACAT-3') et séquencées. Les séquences obtenues ont été nettoyées et alignées avec d'autres séquences 16S des espèces présentes dans les communautés artificielles si elles existaient (NCBI). Puis, pour chaque espèce, des zones conservées au sein de l'espèce et très variables avec les autres espèces ont été identifiées *in silico*. Ensuite, leur spécificité a été vérifiée et optimisée par des PCR jusqu'à obtenir des conditions permettant des amplifications spécifiques par qPCR. Une fois la méthodologie validée, elle était vérifiée en qPCR avant de passer les vrais échantillons. Au final, des amorces ont pu être conçues pour 7 espèces sur 10 (Tableau 4). Pour les 3 autres espèces, il n'a pas été possible de concevoir des amorces spécifiques. Des sondes hydrolyses-MGB de qPCR ont été conçues et ajoutées à la qPCR pour augmenter la spécificité mais, malgré de nombreux tests, nous ne sommes pas arrivés à des amplifications spécifiques satisfaisantes. Ces 3 espèces n'ont pas pu être prises en compte pour les analyses.

Pour chaque espèce avec un couple d'amorces spécifique conçu, les échantillons ont été amplifiés par qPCR dans 10µL de réaction avec 0.50µM de chaque amorce (SsoAdvanced Universal SYBR Green SuperMix Biorad). Le programme qPCR était le suivant : dénaturation initiale à 98°C, 3mn, suivi de 40 cycles de dénaturation à 98°C, 10 sec suivi de 30 sec à la température d'hybridation spécifique (Tableau 4), puis d'une melt curve de 65°C à 95°C. Les qPCR ont été faites en duplicats ou en triplicats lorsque les résultats des duplicats n'étaient pas cohérents

Z A B R

Zone Atelier Bassin du Rhône

Tableau 4. Amorces de qPCR utilisées pour la quantification spécifique des espèces des communautés artificielles. Pour chaque espèce, les amorces forward et reverse, la température spécifique d'hybridation et la quantité de biomasse qu'elles représentent dans chaque communauté artificielle.

Espèce	Amorce spécifique forward (5'-3')	Amorce spécifique reverse (5'-3')	Température d'hybridation/élongation (°C)	Quantité de biomasse représentée par les 10 individus dans les communautés artificielles (mg)
Corbicula fulminea	TAAAAGGTTGGGCCTGCCCGGTGAT	CTTTTTTTATTTTTTAAAACAGCTTTTTTAC	55	165,7
Epeorus assimilis	Non conçue	Non conçue	NA	120,2
Asellus aquaticus	TATTAATATTTAAAATAGAAAATTAGAAC	TTCAACATCGAGGTCATAAACTTTTCTGCC	55	56,47
Glossossoma conforme	Non conçue	Non conçue	NA	47,1
Brachyptera risi	CGAGGTATTGACTGTCTCAATATAAATGAG	AAAAAATTCTTCTAAATTTATAATAAATAC	60	27,23
Drusus annulatus	Non conçue	Non conçue	NA	7,13
Tubifex tubifex	AGAGCTTTACCATAAATCTTATCA	TTTTTAGTGATTAGTAAATCTTAAAGTTTT	55	3,5
Chironomus riparium	TTAAAAATAAATGAATTTATTATTTTAGTA	CTTTAAATCACCCCAATTAATAATTAACA	60	1,7
Riolus subviolaceus	TTTATATTATTTAAAATGTATATTTGTAGG	AAAAAAAATTGGATCAAAAAGTCAAATATA	60	0,9
Potamopyrgus antipodarum	CATGCAAAGCTTCCACCACAGTAATAAAGG	TAGGATATTCTAGTCCAACATCGAGGTCAC	60	NA

Z A B R

Zone Atelier Bassin du Rhône

ADN etoh (ADNet)

NB : Pour *Corbicula fulminea*, les qPCR montraient une amplification non spécifique hors espèces des communautés artificielles (amplicon de taille différente, supposé du 16S bactérien) et le nombre de copies est donc sur-estimé pour cette espèce. Cependant, il est sûr que l'espèce ait bien été détectée en qPCR.

Le nombre de copies moyen était significativement plus élevé pour le traitement FILT par rapport à la PRECIP ($\chi^2=6.22$, pvalue=0.01). Concernant les méthodes d'extraction, seul TISSUE a montré des valeurs de nombre de copies moyen significativement plus élevé que la méthode SALT (t=2.50, pvalue=0.04), les autres paires de méthodes étaient équivalentes.

Pour l'évaluation des protocoles, les tests statistiques n'ont pas pu être faits pour *T. tubifex*, *C. riparius*, *R. subviolaceus* et *P. antipodarum* étant donné le nombre trop important de 0. Pour les autres espèces, des résultats différents entre espèces ont été observés, semblant suggérer que la détection est espèce-dépendante. Malheureusement, il n'a pas été possible de tester si des traits morphologiques impactaient l'extraction de l'ADN des différentes espèces étant donné qu'on ne pouvait pas les différencier de l'effet biomasse. Il semble y avoir un effet seuil pour la détection des espèces corrélé à la quantité de biomasse représentée par les espèces dans les échantillons. Nous avons d'abord testé ce lien avec un modèle linéaire généralisé (passage en détection/non détection des données, famille binomiale) qui a été significatif ($\chi^2=67.7$, pvalue<0.0001). Les espèces en faible quantité (<5mg) ne sont pas ou très peu détectées. Les tests de Levene ont mis en évidence des différences de variance significatives entre les protocoles (F=4,56, pvalue=0,0008) et les tests posthoc ont montré des différences entre le protocole PRECIP*TISSUE et deux protocoles : FILT*BEAD (pvalue=0,001) et PRECIP*BEAD (pvalue<0,0001).

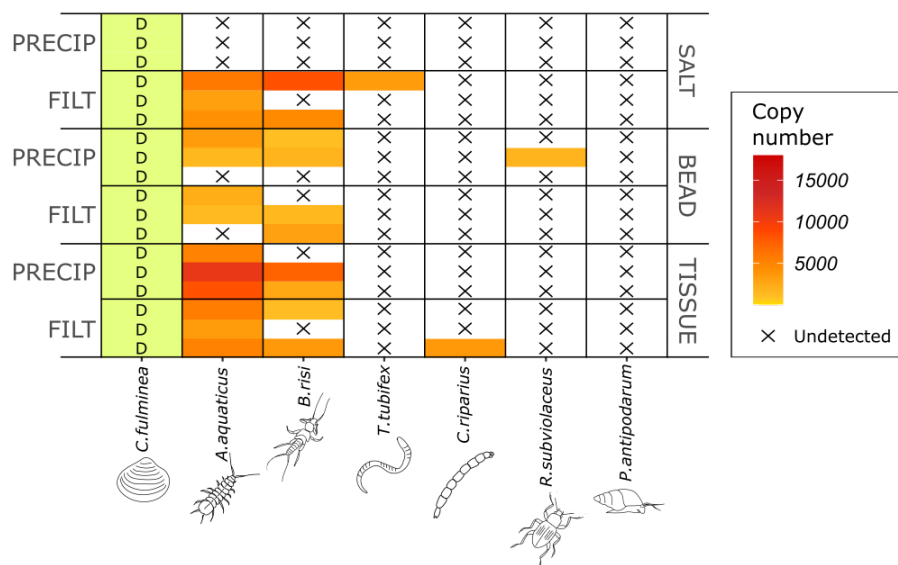


Figure 5. Nombre de copies 16S totales moyen par échantillon mesuré en qPCR pour chaque espèce des communautés artificielles pour les différents protocoles de la matrice ADNet. Les espèces sont classées de la plus abondante (à gauche) en termes de biomasse représentée dans la communauté artificielle à la moins abondante (à droite). Une cellule barrée indique un nombre de copies nul et un D dans une cellule verte indique que l'espèce a été détectée mais n'a pas pu être quantifiée en raison d'une amplification parasite lors de la qPCR.

ADN bulk in toto (ADNbt)

Tous les traitements ont montré des nombres de copies moyen significativement différents les uns des autres (ANNEXE 1 Tableau A3). L'ordre du plus performant (c'est-à-dire présentant le nombre de copies moyen le plus élevé en moyenne) au moins performant est le suivant : DRY > TNESWET > UREA > BUFFER > ETOHWET (Figure 6). Concernant les méthodes d'extraction, on retrouve également une différence significative entre chaque traitement, présenté ici selon l'ordre de la plus à la moins performante : PSM > PSP > PCI (PMS/PSP : $t=-12.33$, $pvalue<0.0001$; PMS/PCI : $t=-17.22$, $pvalue<0.0001$; PSP/PCI : $t=-4.88$, $pvalue<0.0001$).

Les comparaisons des protocoles à la référence (ADN bulk) ont montré des tendances similaires malgré des différences entre espèces (ANNEXE 1 Tableau A4). De manière générale, les protocoles utilisant le kit PowerMaxSoil étaient équivalents à la référence tout comme les protocoles utilisant le kit PowerSoilPro mais uniquement pour les traitements DRY et TNESWET. Tous les autres protocoles (protocoles utilisant le phenol-chloroform et le kit PSP et traitements BUFFER, ETOHWET et UREA) étaient moins performants que la référence. Comme pour la quantité d'ADN extrait, on peut se demander si ces résultats changeaient lorsque les données étaient passées au mg extrait (ANNEXE 1 Tableau A4). Ramené au mg extrait, tous les protocoles étaient significativement différents de la référence et ont donné de moins bons résultats pour toutes les espèces. Cependant les résultats observés pour les traitements et méthodes d'extraction restaient inchangés (Tableau A3 et A5). Les tests de Levene ont mis en évidence des différences de variance

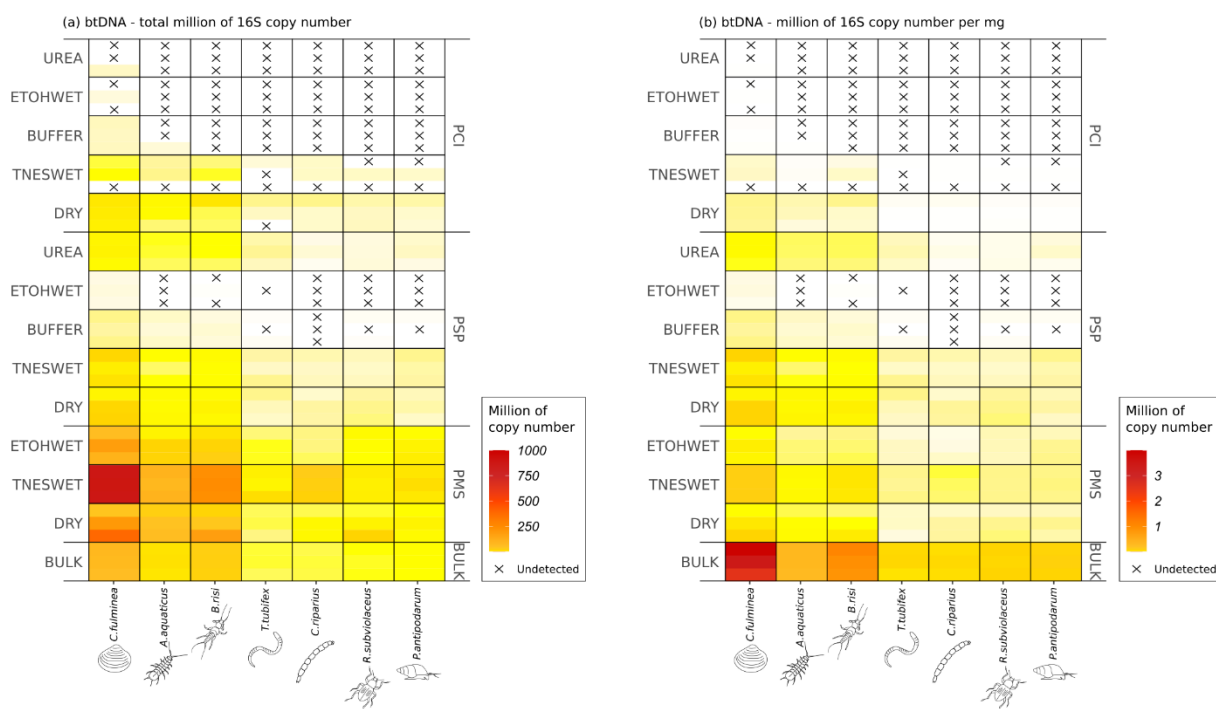


Figure 6. Nombre moyen de copies 16S par échantillon totales (a) et rapporté au mg extrait (b) mesuré en qPCR pour chaque espèce des communautés artificielles pour les différents protocoles de la matrice ADNbt. Les espèces sont classées de la plus abondante (à gauche) en termes de biomasse représentée dans la communauté artificielle à la moins abondante (à droite).

significatives entre les protocoles ($F=12,8$, $pvalue<0,0001$) et les tests posthoc ont montré que l'homogénéité de la variance de TNESWET*PMS différait de celle de tous les autres protocoles, à l'exception de UREA*PCI et ETOHWET*PCI.

Contrôles négatifs

Des contaminations d'ADN ont eu lieu au cours de l'expérience pour les deux types de contrôle négatif. En PCR, 14 des 19 contrôles négatifs ont amplifié au moins un amplicon, principalement BF2/BR2 et Fwh1/Fwh2. Les contrôles de prétraitement de TNESWET*PSP, BUFFER*PCI et UREA*PCI et ceux d'extraction de PSP et BEAD n'ont amplifié aucun amplicon. Deux contrôles négatifs ont été amplifiés pour les trois amplicons : bulk et un contrôle de contamination. En qPCR, les contrôles négatifs de l'ADNet ont détecté entre 1 et 3 espèces, à l'exception du contrôle d'extraction de SALT et BEAD, qui n'a détecté aucune espèce. Pour l'ADNbt, 6 des 13 contrôles négatifs n'ont détecté aucune espèce (tous les PCI sauf DRY, DRY*PSP and BUFFER*PSP), 6 ont détecté 1 espèce (PMS, tous les PSP sauf DRY, BUFFER and UREA et DRY*PCI) et 1 a détecté 3 espèces (PSP*UREA). *C. fulminea* était la principale espèce détectée (12 fois), suivie de *R. subviolaceus* (2 fois) et de *C. riparius* et *P. antipodarum* (1 fois). Le contrôle de contamination a détecté 3 espèces (les trois espèces ayant la biomasse représentée la plus élevée dans les communautés artificielles) ce qui signifie que les espèces ayant la biomasse représentée à l'échantillon la plus élevée sont les plus susceptibles de contaminer d'autres échantillons. Enfin, le contrôle négatif de l'extraction bulk a détecté les 7 espèces en qPCR.

Conclusion

ADN etoh

L'ensemble des résultats suggère qu'aucun protocole n'est adapté pour l'extraction et la détection du macrozoobenthos lorsqu'une espèce a une biomasse faible dans un échantillon. En effet, il semble exister un effet seuil au niveau de la biomasse ne permettant pas de détecter les espèces dans un échantillon (entre 4 et 25 mg, Tableau 2). La large gamme proposée pour le seuil est due à l'absence de mesures pour une espèce de biomasse intermédiaire (*Drusus annulatus*). Pour le reste des espèces, aucun protocole ne s'est distingué même si le traitement Filtration (FILT) semble avoir été plus performant (mais plus chronophage) que la Précipitation au sel (PRECIP). Ce traitement s'est montré plus fiable et devrait être plus exhaustif dans des communautés plus diversifiées puisque tout l'éthanol passe par le filtre (tant que l'espèce est en quantité suffisante). Concernant les méthodes d'extraction, le kit commercial (TISSUE) et la purification par billes (BEAD) étaient équivalentes et l'une ou l'autre peut être utilisée. La purification par billes est une méthode d'extraction manuelle et moins coûteuse que l'extraction avec kit, ce qui peut être intéressant pour des grands programmes de surveillance. Mais une méthode d'extraction commerciale sous forme de kit a l'avantage d'être plus standardisée et donc plus reproductible d'un.e exécutant.e à

l'autre. Un compromis doit être trouvé entre ces deux composantes en fonction des objectifs d'une étude donnée.

Etant donné que l'ADN *etoh* ne semble pas être une méthode intéressante pour établir une liste exhaustive des espèces dans un échantillon de type Surber, **son utilisation pour la phase 3 n'a pas été retenue.** L'*etoh* de conservation des échantillons de la phase 3 a tout de même été récupéré et mis de côté en vue de leur traitement potentiel.

ADN bulk in toto

L'ensemble des analyses met en avant le broyage sec (DRY) et le broyage avec tampon (TNESWET) comme les deux traitements les plus performants pour récupérer une grande quantité d'ADN et pour détecter toutes les espèces dans des échantillons représentatifs de Surber. Ces deux traitements sont également les plus exigeants en termes de temps de traitement et de mise en œuvre. Le traitement TNES-Urea malgré un faible nombre de copies mesurés pour chaque espèce a tout de même permis de détecter toutes les espèces avec un temps de préparation et de mise en œuvre moindre, ce qui diminue le potentiel de contamination croisée et de temps de préparation. Ce traitement pourrait être un bon compromis pour la détection des espèces. Ce traitement prometteur (première fois qu'il est utilisé dans le cadre de l'identification du macrozoobenthos) mérite des approfondissements (p.ex. plusieurs temps d'incubation) pour pouvoir mieux évaluer son potentiel. Au niveau des méthodes d'extraction, les kits d'extraction ont donné de meilleurs résultats que l'extraction au Phenol:chloroform :Isoamyl que ce soit pour la quantité d'ADN extraite ou pour le nombre de copies de qPCR. Le kit PowerMaxSoil a présenté des résultats plus élevés que le kit PowerSoilPro mais les résultats de ce kit étaient largement satisfaisants. Comme attendu, aucun protocole n'a surpassé le protocole de référence (ADN bulk). Mais de nombreux protocoles ont permis de détecter toutes les espèces présentes dans les communautés artificielles malgré des biomasses apportées dans les communautés très basses et sont donc applicables à des échantillons réels de Surber.

Pour la phase 3, **le protocole sélectionné a été celui couplant le traitement de broyage TNES avec le kit Qiagen DNEasy PowerSoilPro (TNESWET*PSP)**, qui correspond à un compromis entre performance du protocole, facilité de mise en œuvre et coût.

PHASE 3 : Applicabilité d'un protocole de métabarcoding au suivi ORE

Cette phase avait pour objectif de valider le protocole développé en laboratoire.

Test in situ – Méthodes

Protocole d'échantillonnage

Etant donné que la phase d'expérimentation du projet a duré un an, il n'a pas été possible de tester le protocole dans le cadre d'un suivi d'une opération de restauration écologique. Nous avons donc sélectionné sur l'Albarine plusieurs sites d'échantillonnage correspondant à 3 couples de sites amont et aval de station d'épuration sur les communes de Brénod (BREN), Bettant (BETA) et Ambérieu (AMBE) (Figure 7). Chaque station (couple amont-aval) possède un niveau d'intermittence différent (intermittence nulle, modérée, haute). Ce design nous a permis de vérifier si notre protocole permettait de détecter les différences de communautés entre sites présentant des caractéristiques différentes (niveau d'intermittence différent, amont-aval de station d'épuration pour lesquels une pollution avait été détecté dans un projet précédent soutenu par l'Agence de l'eau, Crabot et al, 2021) et pour lesquels nous attendions des compositions de communautés différentes. L'échantillonnage a eu lieu entre le 06/05/2021 et le 01/05/2021 lorsque les rivières étaient en eau.

Liste taxonomique morphologique

Les 72 échantillons (12 prélèvements*6 sites) ont été triés et identifiés par Colin Issartel au LEHNA selon un protocole strict de décontamination pour éviter les contaminations croisées (javel, port de gants et de blouse). Par ailleurs, un protocole de décontamination ADN a aussi été suivi lors de l'échantillonnage pour limiter les contaminations croisées (bain javel du matériel entre chaque prélèvement, cf ANNEXE 3 pour plus de détails). Cette identification a permis d'obtenir une liste taxinomique des taxa présents dans chaque échantillon (cf ANNEXE 2 pour le détail du protocole de tri et les niveaux de détermination par groupe, basés sur l'I2M2 en partie).

NB : avant le tri et la détermination des échantillons, l'éthanol de conservation a été prélevé et conservé à -20°C au cas où ces échantillons soient traités pour faire du métabarcoding.

Liste taxonomique moléculaire

Ces mêmes échantillons ont ensuite été utilisés pour l'extraction ADN. Le protocole utilisé du terrain au traitement des échantillons en laboratoire avant séquençage correspond à l'ANNEXE 3 de ce rapport (également dans le dossier DocSupp_Rapp_final joint). Rapidement, l'ADN de chaque échantillon a été extrait selon le protocole en ANNEXE 3 (broyage avec du tampon et extraction kit PowerSoil Pro de Qiagen). Puis il a été dosé et amplifié avec deux couples d'amorces : Fwh2/EPTDr2n (Leese et al, 2021 ; Tableau 3) pour permettre une comparaison avec les données obtenues par un autre projet fait à la même période et mlCOIintF/jgHCO2198 (Leray et al, 2013 ; Tableau 3) pour avoir des amplicons plus longs permettant

d'avoir des données de meilleure qualité pour l'identification ADN. Ce sont les données de ce dernier couple d'amorces qui ont été utilisés dans ce rapport. Les amplifications ont été effectuées avec le kit Qiagen Multiplex PCR plus Kit avec 0.36µM de chaque amorce pour le couple Fwh2/EPTDr2n et 0.5 µM de chaque amorce pour le couple mlCOIintF/jgHCO2198 en suivant les protocoles des auteur·ices (cf ANNEXE 3). Les PCR ont ensuite été envoyées pour séquençage à la plateforme GenoToul (Toulouse) en février 2022 et les données ont été fournies début avril 2022.

La plateforme fournit les lectures (les séquences issues du séquençage, aussi appelées « reads ») démultiplexées, c'est-à-dire séparées par échantillon selon un double-indexage, et nettoyés des adaptateurs Illumina, tags et séquences de P5 et P7. Après un contrôle qualité des lectures brutes avec fastqc (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) et multiqc (Ewels et al, 2016). Le traitement des échantillons a été fait avec le package dada2 sur R (Callahan et al, 2016) : 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. Le taux d'erreur a ensuite été inféré afin de vérifier qu'il diminuait bien avec le score de qualité. Puis, après une étape de déreplcation (identification de séquences uniques appelées ASVs, Amplicon Sequence Variants), dada2 a inféré la composition des échantillons en les corrigeant à partir du modèle d'erreur et les lectures R1 et R2 ont été fusionnées. 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 extérieure. Les ASVs ont finalement été regroupés en OTUs (Operational Taxonomic Units) avec l'outil de clustering swarm (Mahé et al, 2021).

Les séquences obtenues en sortie de dada2 (OTUs) peuvent être utilisées pour l'assignement. Suite à des évaluations non présentées ici, l'assignement a été fait en deux phases permettant d'améliorer sa qualité (càd de maximiser la résolution taxonomique des assignements tout en minimisant la proportion de faux assignements). Tout d'abord, les séquences ont été assignées avec RDP Classifier (Wang et al, 2007) qui classe les séquences en extrayant des caractéristiques de composition (ici les profils de fréquence des nucléotides) avant de construire un modèle qui relie ces profils à des groupes taxonomiques spécifiques. Les assignements à l'espèce de RDP Classifier ont été conservées pour construire le jeu de données finales. Toutes les séquences n'ayant pas été assignées à l'espèce par RDP Classifier ont été alignées sur la base de données de référence avec BLASTn (Altschul et al, 1990). Les résultats d'assignement (hits) ont ensuite été traités par BASTA, un outil permettant une annotation taxonomique basée sur le dernier ancêtre commun (*Last Common Ancestor*, LCA, Kahlke et al, 2019). Les assignements en sortie de BASTA ont été ajoutés au jeu de données final. La même base de données a été utilisée par les deux outils d'assignements : la base de

données CO1 Classifier (release v5.1.0, Porter & Hajibabaei, 2018). Les paramètres utilisés pour chaque outil sont regroupés dans le Tableau 5. Pour finir, le jeu de données obtenu a été nettoyé en se basant sur les contrôles négatifs de l'expérience (recherche de contamination) et en retirant les ASVs non assignés.

Comparaison des identifications morphologique et moléculaire

Dans un premier temps, nous avons examiné l'ensemble de la taxonomie (diversité Gamma) et les diversités Beta associés pour avoir une vue d'ensemble des deux approches. Puis, nous nous sommes placés au niveau du prélèvement puis du site pour comparer plus en détail les deux approches. Le détail des analyses et les résultats sont présentés dans la partie **Test in situ – Résultats** ci-après.

Tableau 5. Paramètres utilisés pour chaque outil d'assignement

Méthode d'assignement	Outil d'assignement	Paramètres utilisés
Composition de séquence	RDP Classifier	-C 0,8 (seuil de confiance d'assignement) Pas de paramètres sur blastn
Similarité de séquence	BLASTn & BASTA	BASTA: -n 25 (nombre maximum de hits à utiliser pour définir le LCA) -e 1e-50 (seuil pour la e-value) -m 1 (nombre minimum de hits qu'une séquence doit avoir pour se voir attribuer un LCA) -p 80 (% de hits du blastn utilisés pour définir le LCA) -i 80 (% d'identité minimale)

Z A B R

Zone Atelier Bassin du Rhône

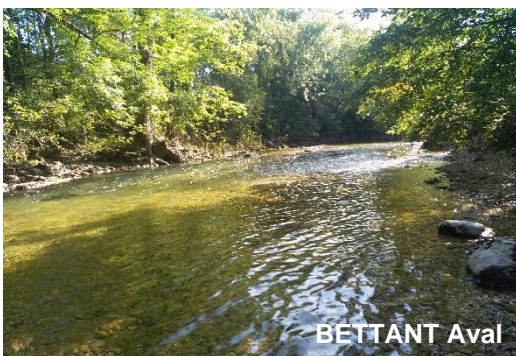
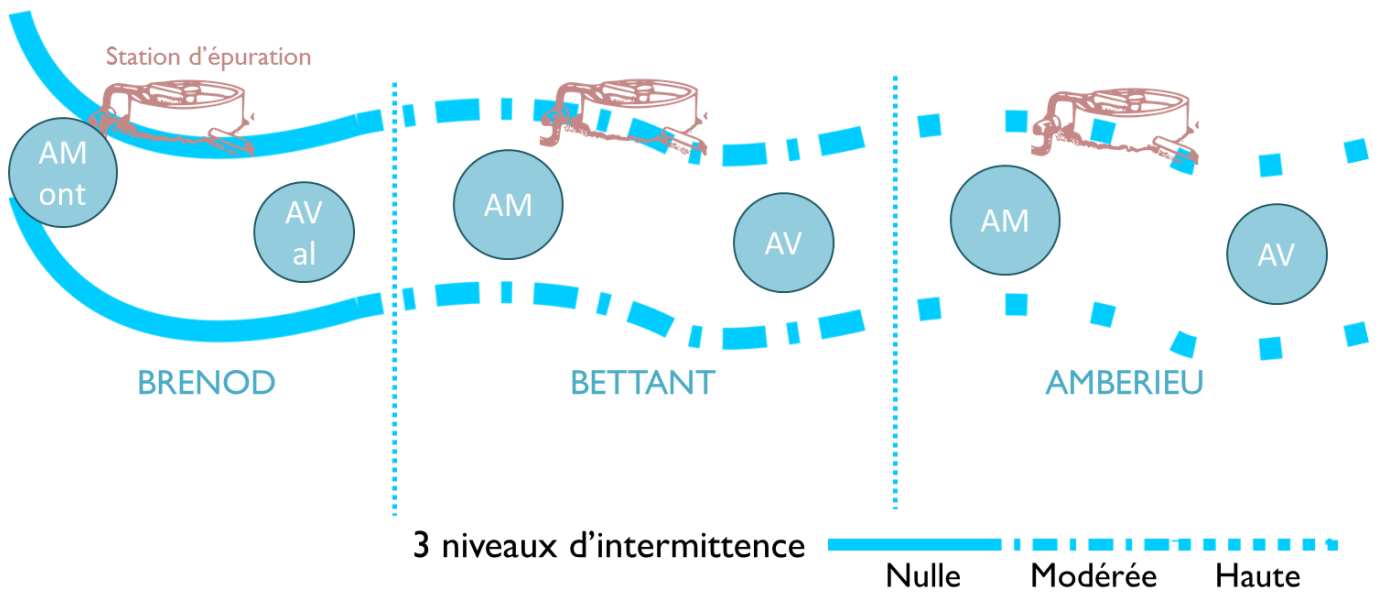


Figure 7. Représentation schématique et photographies des sites échantillonnés. Chaque site correspondait à l'AMont ou l'AV al d'une station d'épuration et correspondait à un niveau d'intermittence différent.

Test in situ – Résultats

Diversité Gamma

Au total, 936 identifications ont été effectuées en morphologie correspondant à 113 taxa différents assignés à des rangs allant du genre à la classe (Figure 8). En identification moléculaire, 15 289 OTUs ont été obtenus représentant des phylums de la quasi-majorité des règnes du vivant (*Animalia, Fungi, Plantae, Bacteria, Protista*). **L'identification moléculaire permet donc d'avoir accès à d'autres groupes taxonomiques que les « invertébrés ».**

Sur ces 15 289 OTUs, 4 059 correspondaient à des phylums ciblés dans le cadre du projet (phylums « invertébrés ») : *Arthropoda, Mollusca, Annelida, Nematelminthes, Platyhelminthes, Cnidaria, Rotifera, Porifera, Nemertea, Bryozoa*. Ces 4 059 OTUs représentaient près de 75% du nombre de lectures total du jeu de données. Un nettoyage bio-informatique permettant de normaliser (raréfaction à 40 960 lectures pour chaque échantillon, library *phyloseq*, McMurdie & Holmes, 2013) et de nettoyer le jeu de données sur les OTUs rares (pour chaque échantillon, les OTUs représentant moins de 0,01% du nombre de lecture total étaient retirés) a été appliqué pour aboutir à 2 306 OTUs. Cette opération a fait perdre 6 prélèvements pour lesquels le nombre de lectures total était trop bas (i.e. <40 960 lectures ; 1 prélèvement de AMBE-Amont, 4 de AMBE-Aval et 1 de BETA-Amont). Le jeu de données a ensuite été dégradé pour permettre une comparaison plus juste avec les identifications morphologiques. En effet, en identification morphologique, on identifie un organisme à un nom de taxon (à un rang taxonomique pouvant varier) et pas à un morphotype. Par exemple, si dans un groupe où les clés d'identification permettent des assignements à la famille uniquement, on observe des morphotypes correspondant potentiellement à des genres ou espèces différentes, l'identification sera tout de même donnée à la famille. L'identification moléculaire permet de dépasser ce problème avec les OTUs qui permettent d'avoir une différenciation à la même unité taxonomique même si ces unités ont des noms d'assignements à des taxa de rangs différents. Cela rend la comparaison entre morphologie et moléculaire déséquilibrée en faveur du moléculaire. Nous avons donc « écrasé » le jeu de données moléculaire : pour chaque échantillon, si plusieurs OTUs avaient le même assignement/nom de taxon, ils étaient combinés pour ne représenter plus qu'un taxon dans cet échantillon. Dans la suite de cette analyse, le terme taxon ne correspondra plus à un OTU pour le jeu de données moléculaire mais à un nom de taxon comme pour le jeu de données morphologique. Le nouveau jeu de données contenait 645 taxa différents. Le niveau d'assignement pour ces taxa finaux allait de l'espèce au domaine (Figure 8). **L'identification moléculaire a permis un gain de résolution taxonomique** avec plus de 80% des taxa possédant des assignements au moins au genre (65% des assignements à l'espèce et près de 20% au genre) en comparaison à l'identification morphologique avec moins de 50% des assignements au genre (Figure 8).

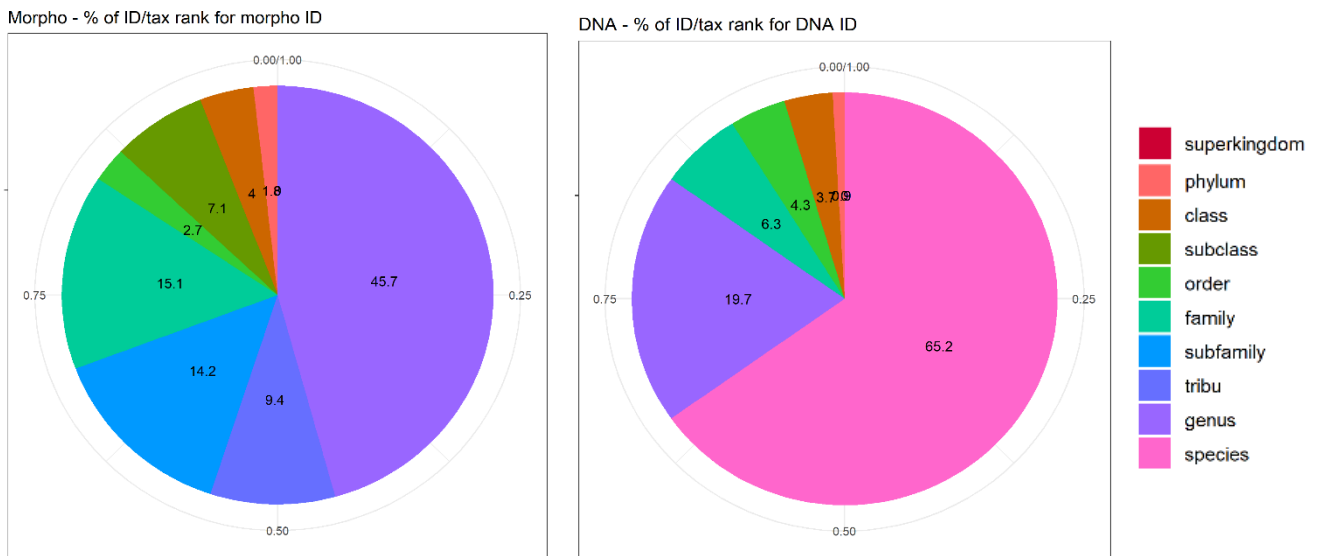


Figure 8. Pourcentage des rangs taxonomiques des assignations issues de l'identification morphologique (à gauche) et de l'identification moléculaire (à droite).

De plus, les proportions de taxa dans les différents phylums étaient équivalents entre les deux approches. On remarque que l'identification moléculaire a permis la détection de 4 phylums supplémentaires par rapport à la morphologie mais ne détectait pas les Nematelminthes (Figure 9).

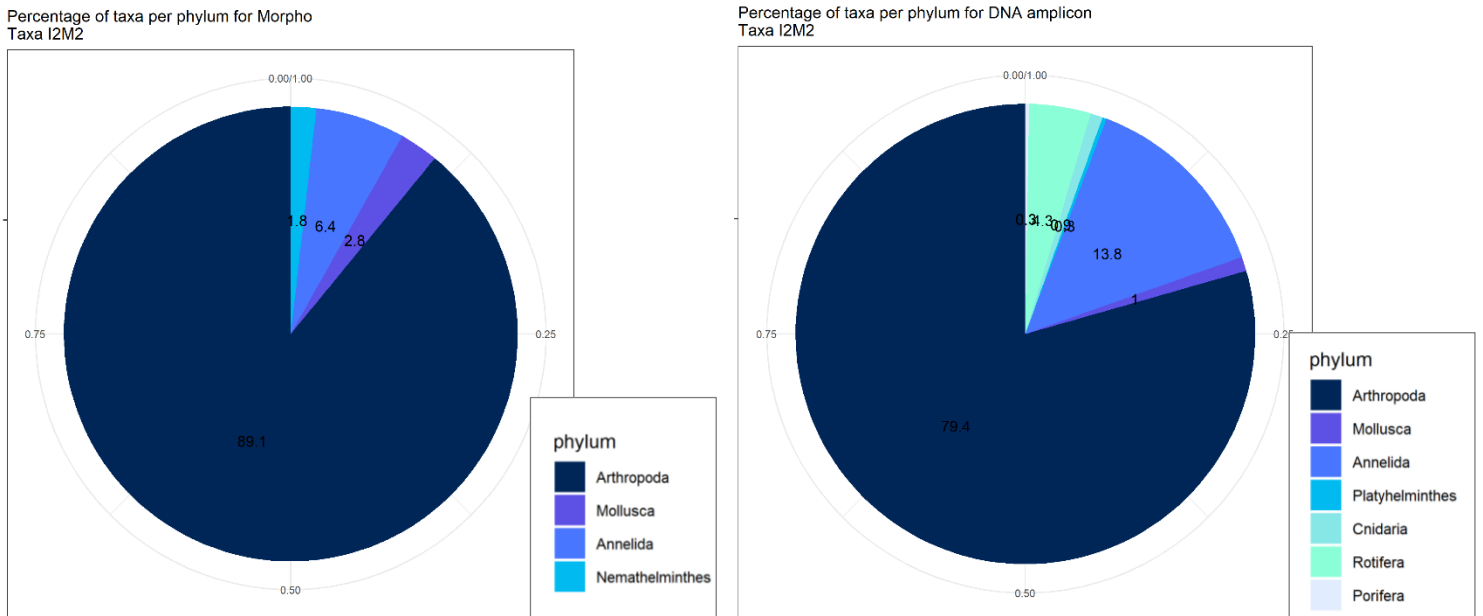


Figure 9. Pourcentage des taxa dans chaque phylum « invertébrés » pour l'identification morphologique (à gauche) et l'identification moléculaire (à droite).

Pour nous intéresser plus finement à la composition taxonomique dans des groupes d'intérêt dans le suivi du macrozoobenthos de rivières (i.e. *Insecta*, *Mollusca*, *Malacostraca*, *Annelida*), nous avons utilisé des graphiques treemaps. Ces graphiques représentent la composition hiérarchique d'une communauté en termes de diversité et d'abondance (nombre d'individus pour la morphologie, nombre de lectures pour le moléculaire) pour un groupe donné (*Insecta* et *Annelida* : Figure 10 ; *Mollusca* et *Malacostraca* : ANNEXE 4). Au niveau de la diversité au rang « taxon », on a observé des ordres de la classe *Insecta* pour qui la diversité était similaire entre les deux approches et qui étaient principalement des ordres avec des bonnes connaissances en taxonomie (p.ex. Ephemeroptera). Ce patron a également été retrouvé dans d'autres groupes (*Mollusca* et *Malacostraca*) pour qui la diversité est à priori faible sur les sites étudiés. Pour d'autres ordres ou classes, un gain taxonomique très important a été observé (p.ex. *Clitellata* du phylum *Annelida* ou Diptera de la classe *Insecta*). Ces groupes sont généralement décrits à des rangs taxonomiques grossiers morphologiquement, même si un gain taxonomique peut aussi être observé chez des groupes bien décrits morphologiquement. L'identification a donc permis une meilleure description de ces groupes. De plus, il ne faut pas oublier que le passage de OTU à « taxon » a dégradé le jeu de données moléculaire et que la diversité aurait été encore plus forte au grain « OTU ».

Au niveau des abondances, un signal « nombre d'individus » en morphologie versus un signal « biomasse » en moléculaire ont été observés (Bista et al, 2018 ; Gauthier et al, 2019). Ces deux signaux différents peuvent expliquer par exemple que les éphéméroptères (organismes gros) étaient majoritaires en moléculaire alors que les diptères (présents en général en grand nombre) l'étaient en morphologie (classe *Insecta*). Cependant, pour *Insecta* et *Annelida*, les ordres ou classes les plus représentés restaient les mêmes entre les deux approches (Figure 10). Ce n'était pas le cas pour les *Mollusca* et *Malacostraca* (ANNEXE 4), ce qui pourrait s'expliquer pas des biais d'amorces. **Avec l'approche moléculaire, on perd donc les données de comptage du nombre d'individus mais on conserve une information biomasse, qui peut être biaisée selon les groupes et amorces utilisées.**



Figure 10. Treemaps représentant la diversité de 3 niveaux hiérarchique au sein de la classe Insecta (haut) et du phylum Annelida (bas) pour l'identification morphologique (gauche) et moléculaire (droite). Le premier niveau hiérarchique (bordures très épaisses, nom en blanc) correspond aux ordres pour Insecta et aux classes pour Annelida, le deuxième (bordures épaisses, nom en noir) aux familles (Insecta) et aux ordres (Annelida) et le troisième aux taxa. Attention : les couleurs entre identifications morphologique et moléculaire ne correspondent pas forcément pour un groupe donné entre les deux identifications.

Diversité Beta

Les NMDS (Non-Metric Multi-dimensional Scaling) permettent de visualiser la dissimilarité (diversité Beta) entre paire de prélèvements dans un espace à faible dimension (Figure 11) : plus deux prélèvements sont éloignés dans l'espace, plus leurs communautés ont des compositions différentes. Les deux approches d'identification ont bien permis de différencier les différents sites (AMBE, BETA et BREN) entre eux (Figure 11, Tableau 6) et également, pour Ambérieu et Brénod, les communautés amont et aval de la station d'épuration (sauf Ambérieu avec l'approche moléculaire mais il y a un effet marginal, Tableau 6). Globalement, **l'identification moléculaire est cohérente avec l'identification morphologique et permet de discriminer des communautés de sites différents entre elles.**

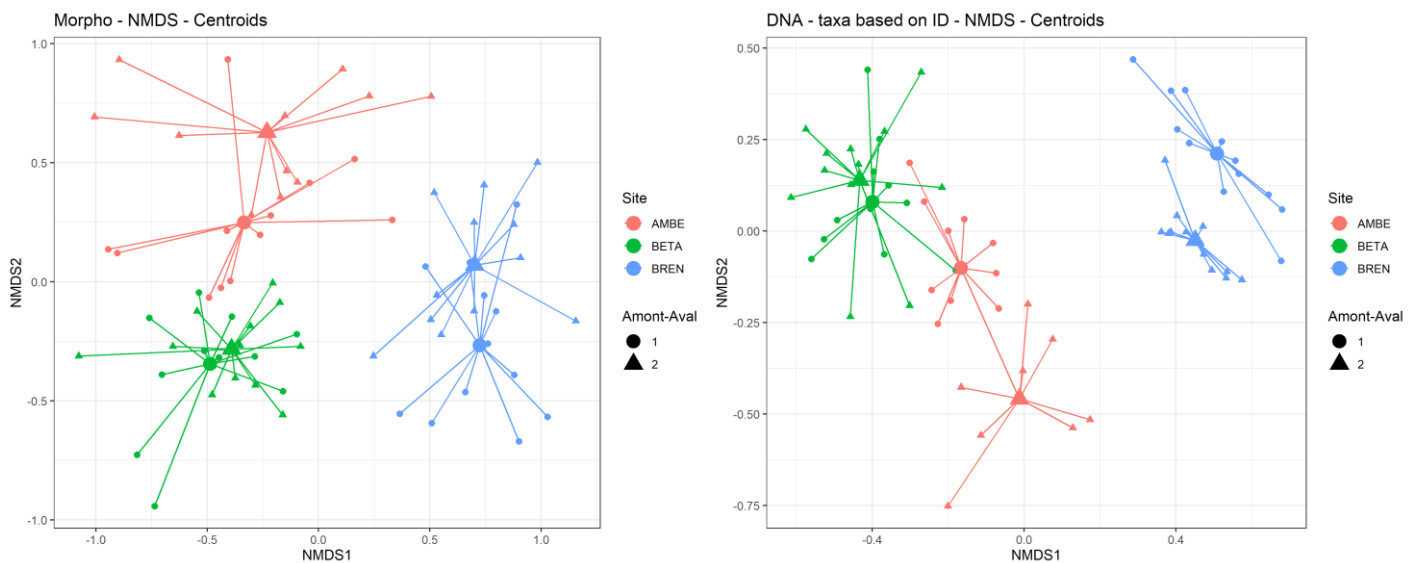


Figure 11. NMDS des communautés des prélèvements des trois couples de sites prélevés. Les couleurs correspondent aux stations et les formes aux sites Amont (rond/1) et Aval (triangle/2). Chaque point correspond à un prélèvement et sont reliés à leur centroïde (site*amont/aval).²

Tableau 6. Statistique F et p-values des PERMANOVA et des tests post-hocs associés testant l'effet Site et Amont-Aval sur les diversités beta des identifications morphologiques et moléculaires. Une pvalue<0.05 est considérée comme significative.

Variable/Comparaison post hoc	Identification morphologique		Identification moléculaire	
	F value	p-value	F value	p-value
Site	22,35	0,001	10,84	0,001
AMBE vs BETA	12,82	0,003	4,05	0,001
AMBE vs BREN	17,62	0,003	10,07	0,001
BETA vs BREN	31,83	0,003	16,96	0,001
Amont-Aval	2,20	0,03	2,23	0,008
AMBE : Amont-Aval	3,54	0,004	1,63	0,06 (marginal)
BETA : Amont-Aval	0,52	0,81	1,27	0,21
BREN : Amont-Aval	5,09	0,001	3,58	0,006

Comparaison à l'échelle du prélèvement

Après les analyses sur le jeu de données global, nous nous sommes intéressé.es à la comparaison des deux approches à l'échelle du prélèvement. Ainsi, pour chaque prélèvement, nous avons identifié les taxa en commun, ceux non détectés et ceux gagnés avec l'identification moléculaire au regard de la liste taxonomique morphologique du prélèvement (Figure 12, détail par prélèvement en ANNEXE 5).

Globalement, **l'identification moléculaire a permis de retrouver une majorité des taxa identifiés en morphologie** (entre 57-100%, $\bar{x} \pm sd = 87\% \pm 11$, Figure 12 et ANNEXE 5). La résolution taxonomique de ces taxa retrouvés en identification moléculaire étaient toujours soit au genre (44%), soit à l'espèce (56%). Les taxa non retrouvés par l'identification moléculaire représentait entre 0 et 41% d'un prélèvement ($\bar{x} \pm sd = 12\% \pm 10$). Les cas de taxa manqués dans l'ensemble des prélèvements correspondaient aux taxa suivants : 17 *Nemathelminthes*, 23 *Mollusca*, 17 *Annelida* (seulement des Achaeta), 216 *Arthropoda* dont 13 *Maxillopoda* (seulement Copepoda), 29 *Malacostraca* (principalement des Gammaridae), 151 *Insecta* (principalement Diptera et Coleoptera, quelques Hemiptera, Ephemeroptera, Plecoptera, Trichoptera (<10 cas)). Une partie de ces taxa correspondent à des taxa mal ciblés par les amorces (p.ex. les *Nemathelminthes*, Leray et al, 2003) ou à des taxa souvent difficiles à amplifier dû à leur diversité génétique élevée (p.ex. Gammaridae). Certains sites présentaient une part plus importante de taxa non retrouvés (AMBERIEU-Amont et l'Amont et Aval de BRENOD). Ces sites présentaient une diversité plus importante par rapport aux autres sites, ce qui expliquerait un taux de détection plus faible.

Des taxa supplémentaires ont été détectés par l'identification moléculaire au regard de la liste taxonomique morphologique (4-53 taxa par prélèvement, $\bar{x} \pm sd = 20 \pm 12$, Tableau 7). Ces taxa représentaient des taxa non-aquatiques, des taxa aquatiques non ou peu comptabilisés en identification morphologique (*Cnidaria*, *Platyhelminthes*, *Porifera*, *Rotifera*) et des taxa aquatiques comptabilisés en identification morphologique. On a obtenu avec l'identification moléculaire une diversité plus importante de ces taxa aquatiques que celle observée en morphologie. La présence de ces taxa pourrait avoir plusieurs origines. Il pourrait s'agir de taxa issus de contamination entre échantillons durant leur préparation pour le metabarcoding ou d'erreur d'assignement lors de l'étape de bio-informatique (il y a toujours un pourcentage de faux positifs durant cette étape). La piste de la contamination est à écarter car nous avons procédé à des analyses post-traitement bio-informatique et aucun patron de contamination n'a été mis en évidence. Des faux positifs d'assignements ne permettent pas non plus d'expliquer la quantité de taxa supplémentaires observée. Une piste intéressante est que ces taxa étaient vraiment présents dans l'échantillon sous des formes qui n'ont pas été détectées par l'œil humain : des individus trop petits, des œufs dans les substrats, de l'ADN présent dans les tubes digestifs, de l'ADN environnemental piégé dans le sédiment, les biofilms ou dans l'eau. Cette piste nous a amené à la question de savoir si l'approche moléculaire était plus intégrative que l'approche morphologique, c'est-à-dire si un prélèvement moléculaire représentait mieux la diversité d'un site entier que la morphologie.

Z A B R

Zone Atelier Bassin du Rhône

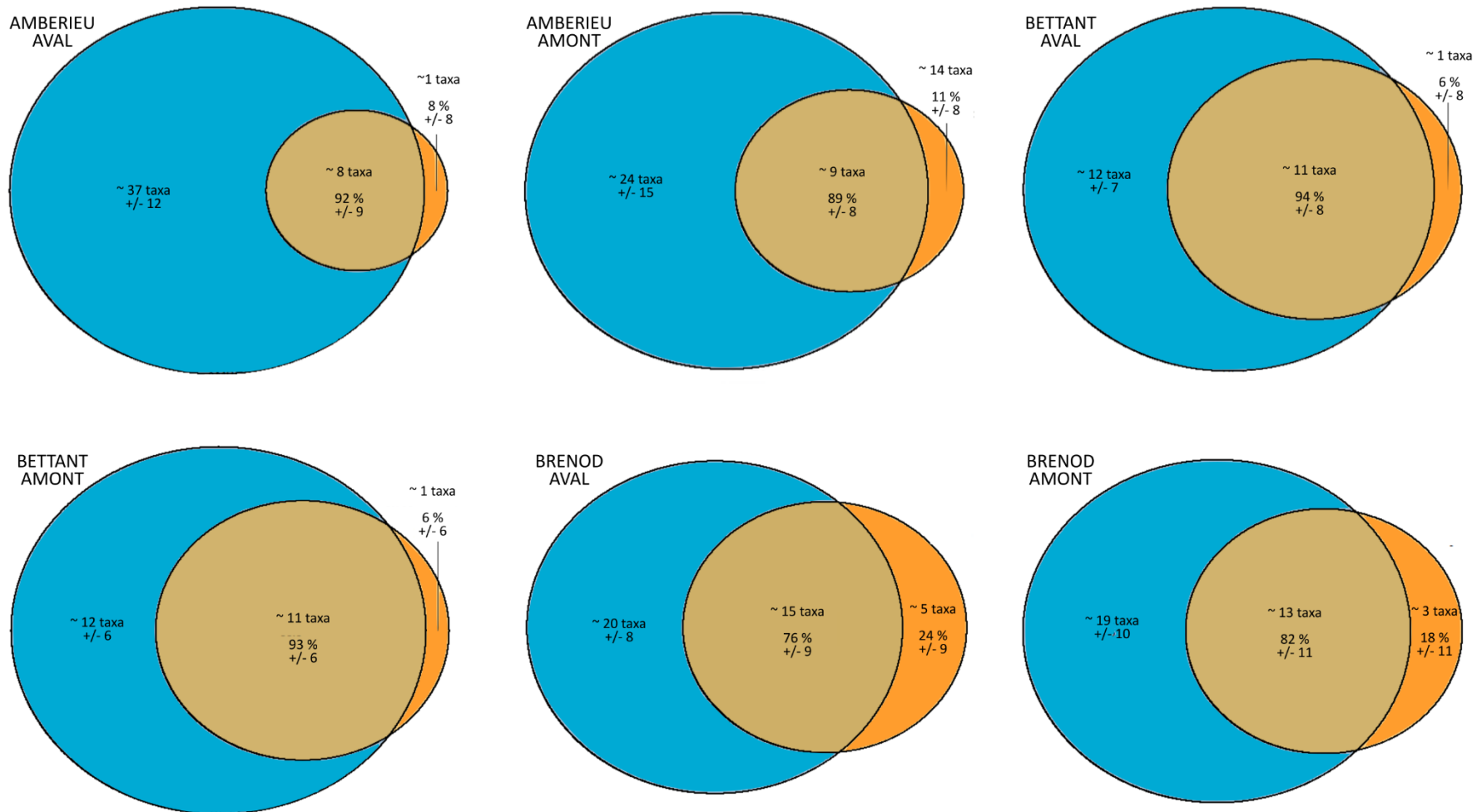


Figure 12. Diagramme de Venn présentant pour chaque site le nombre moyen de taxa partagés (au centre, marron) et uniques à l'identification morphologique (à droite, orange) et à l'identification moléculaire (à gauche, bleu). La part unique à l'identification moléculaire correspond donc à des taxa « gagnés » et celle unique à la morphologie des taxa « perdus » par l'identification moléculaire. Les pourcentages dans les parts orange et marron correspondent aux pourcentages de la liste taxonomique morphologique qu'elles représentent. Les valeurs +/- correspondent aux écarts-types des moyennes.

Tableau 7. Nombre de taxa supplémentaires obtenus avec l'identification moléculaire par rapport à la liste taxonomique morphologique à l'échelle du prélèvement et du site pour différents groupes taxonomiques et le taux de compensation d'un prélèvement par d'autres prélèvements du site.

	Taxa	Echelle prélèvement	Echelle site	Compensation d'un prélèvement par les autres prélèvements du site (%)
Phylum	Annelida	98	0	100
	Arthropoda	980	473	51,7
	Cnidaria	29	29	0
	Mollusca	30	5	83,3
	Platyhelminthes	10	10	0
	Porifera	9	9	0
	Rotifera	148	148	0
Classes au sein des Arthropoda	Arachnida	165	141	14,5
	Chilopoda	1	1	0
	Collembola	19	19	0
	Diplopoda	2	2	0
	Hexanauplia	6	6	0
	Insecta	730	267	63,4
	Malacostraca	11	6	45,5
	Ostracoda	16	1	93,8

Comparaison à l'échelle du site

Pour tester si l'identification moléculaire était plus intégrative que l'identification morphologique, nous avons dans un premier temps mis en commun les listes taxonomiques morphologiques par site pour obtenir une liste taxonomique du site. Ensuite, nous avons calculé pour chaque liste taxonomique moléculaire et morphologique de chaque prélèvement, le pourcentage de la liste « site » qu'elle représentait, c'est-à-dire le niveau de représentativité de la communauté du site décrite morphologiquement. Ces résultats sont illustrés dans la Figure 13. Pour un site donné, les identifications moléculaires ($\bar{x} \pm sd = 64,7\% \pm 8,3$) représentaient mieux sa communauté entière que les identifications morphologiques ($\bar{x} \pm sd = 33,5\% \pm 11,4$). De plus, à l'échelle du site, beaucoup de taxa supplémentaires ont été compensés en partie par d'autres prélèvements du site (Tableau 7). Cette compensation dépendait des groupes mais, par exemple, les *Annelida* ont été compensés à 100% par d'autres prélèvements, les insectes à plus de 60%. Par ailleurs, il ne faut pas oublier que certains groupes, par exemple les *Arachnida*, contiennent des taxa terrestres qui ne pouvaient donc pas être compensés par d'autres prélèvements, ne faisant pas partie de la liste taxonomique « site » et, pour ce groupe, l'ensemble des Trombidiformes (unique ordre des *Arachnida* dans les rivières) ont été compensés. Finalement, nous avons effectué une analyse qui a montré que les listes taxonomiques « site » n'étaient pas exhaustives et il est possible que ces taxa supplémentaires étaient en réalité des taxa non décrits en morphologie à l'échelle de l'ensemble des prélèvements mais bien présents sur le site. Ces résultats nous ont permis de conclure que **l'identification moléculaire représente un outil plus intégratif du site que l'identification morphologique.**

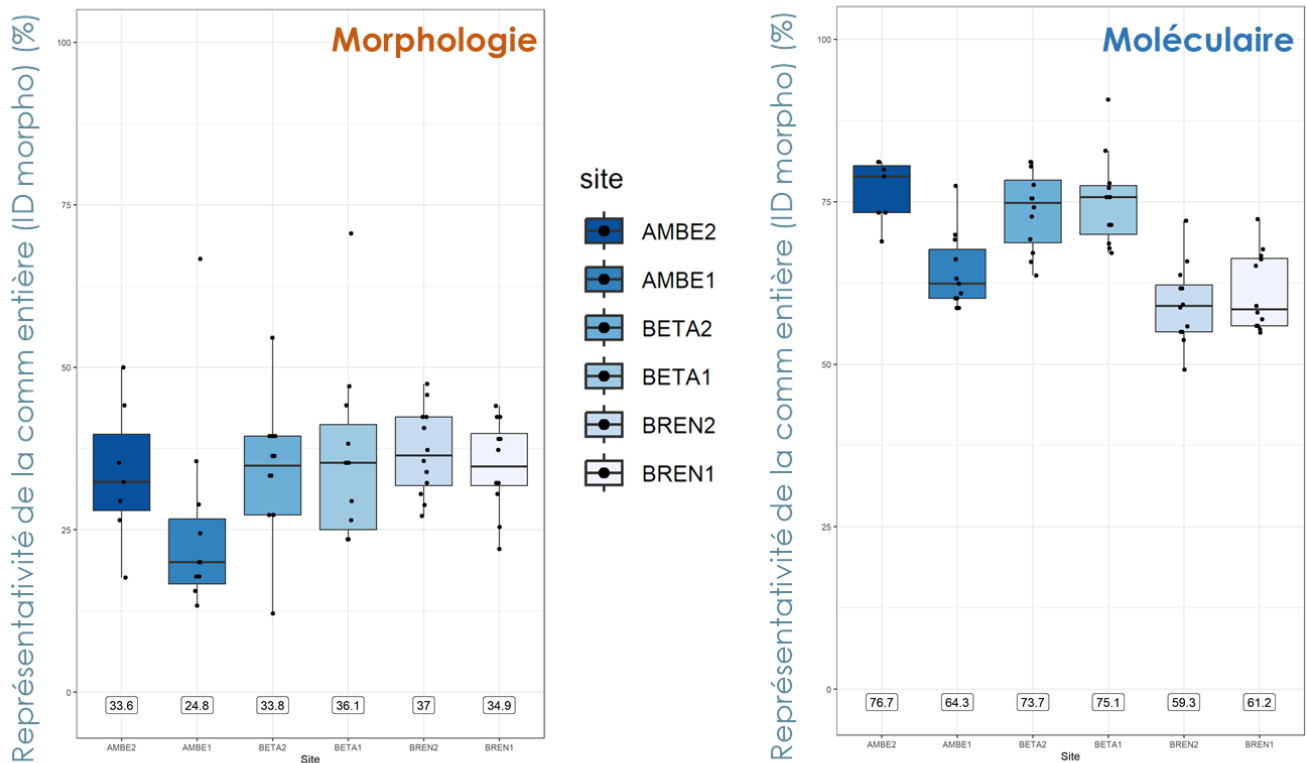


Figure 13. Distribution du pourcentage de représentativité de la communauté entière décrite en morphologie par les prélèvements pour les différents sites (abscisse) et pour les deux types d'identification (morphologique à gauche et moléculaire à droite). Les encadrés représentent la moyenne pour un site donné. Abréviations : AMBE : Ambérieu ; BETA : Bettant ; BREN : Brénod ; 1 : site amont ; 2 : site aval.

Test in situ – Conclusion

Cette étude comparative des identifications morphologique et moléculaire avait pour objectif d'évaluer les gains et les désavantages associés à l'identification moléculaire pour le suivi des communautés de macrozoobenthos.

Globalement, en plus **d'avoir accès à une plus grande diversité de groupes taxonomiques** (p.ex *Fungi* ou *Plantae*), l'identification moléculaire a permis de **nettement améliorer la résolution taxonomique dans les phylums ciblés par cette étude**. Ce gain taxonomique était moins important chez des groupes bien décrits et connus morphologiquement ou chez des groupes pour qui la diversité est a priori faible dans les communautés étudiées (quelques espèces ou genres). Pour ces derniers, le gain apporté par le moléculaire se ferait à l'échelle de l'OTU plutôt qu'à celle du nom d'assignement. Pour des groupes mal décrits, ce gain taxonomique est extrêmement important. Le gain taxonomique aurait été encore plus important pour tous les groupes considérés si nous avions travaillé au niveau de l'OTU pour l'identification moléculaire.

Au niveau de la différenciation des communautés, l'identification moléculaire a pu **discriminer des stations de prélèvement et des sites amont-aval** pour lesquels les communautés étaient supposées différentes, en cohérence avec l'identification morphologique. Elle possède donc un signal local assez fort pour discriminer des communautés, permettant à priori des suivis temporels de communautés et des études comparatives entre sites.

L'identification moléculaire a également **détecté la majorité des taxa décrits morphologiquement** (toujours plus de 57% avec une moyenne à 86%). Certains taxa non détectés correspondaient à des taxa ayant des affinités faibles avec les amorces. Si ces taxa représentent un intérêt fort pour une étude, d'autres amorces avec de meilleures affinités avec eux pourraient être utilisées. Le taux de taxa non détectés était aussi plus élevé dans des communautés plus diverses (au regard de la liste taxonomique morphologique), laissant supposer que plus une communauté est riche, plus le taux de détection diminue. Mais cela demande à être vérifié car nous n'avons pas assez de prélèvements pour l'affirmer et ce taux devrait dépendre aussi de la composition des communautés.

De nombreux taxa supplémentaires ont été détectés avec l'identification moléculaire. On ne peut pas distinguer, ni évaluer les parts des causes diverses de la présence de ces taxa : espèces présentes sous des formes invisibles à l'œil humain, erreur lors du traitement bio-informatique, contamination, ADN venant d'autres sites. Nous pensons qu'une majorité de ces taxa représente un signal local et non un bruit de fond. D'une part, nous avons pu montrer que l'identification moléculaire était **plus intégrative** de la communauté d'un site et qu'une majorité de ces taxa étaient compensés par d'autres échantillons de la communauté du site. D'autre part, la liste taxonomique morphologique du site n'était pas exhaustive et ne correspondait donc pas à une référence absolue : il est donc possible que ces taxa étaient réellement présents localement sur le site. Si ces taxa sont toujours appréhender avec méfiance, il est possible d'ajouter des étapes ou filtres pour rendre l'analyse plus stricte afin d'être sûr d'avoir un signal local. En laboratoire, les échantillons peuvent être extraits en triplicats et les OTUs présents uniquement dans les 3 réplicats conservés pour établir la liste taxonomique. Cette méthode permet de retirer les taxa issus de contamination et rares dans les échantillons. Il est également possible de mettre un filtre sur l'abondance plus strict (p.ex. à 0,1%) lors du traitement bio-informatique afin d'« assurer » un signal local étant donné que les taxa les plus abondants (nombre de lectures) sont majoritairement ceux qui ont le plus contribué à l'ADN de l'échantillon, c'est-à-dire les organismes avec des tissus. Il faut bien considérer que ces étapes et filtres vont diminuer les faux positifs potentiels mais peuvent également faire diminuer le taux de vrais positifs, en mettant un filtre trop strict sur les données (notamment les taxa rares). Concernant les taxa non aquatiques, il est facile de les retirer en mettant en place des listes de famille, genre, espèce de taxa non ciblés qui serait appliqué au jeu de données. Finalement, le risque d'identifier des taxa provenant d'autres sites via de l'ADN environnemental est très faible. Le risque de présence d'ADN environnemental existe mais nous avons travaillé ici sur l'ADN d'organismes et non sur les traces de leur ADN dans l'environnement

et leur ADN est donc majoritaire dans nos échantillons. De plus, nous avons ciblé un fragment d'ADN « long » par rapport à ce qui est utilisé en ADN environnemental et il y a peu de chances que l'ADN environnemental, qui est dégradé (petits fragments), puisse être amplifié par nos amorces. Pour finir, même si certains ADNe avaient été amplifiés, cela aurait été dans des faibles proportions et notre traitement bio-informatique aurait supprimé les OTUs concernés. Pour conclure, il n'est pas possible de connaître la part de signal local et de bruit de fond dans ces taxa supplémentaires. Nous pensons tout de même que le signal local est majoritaire, mais il est possible de mitiger le potentiel bruit de fond en étant plus strict sur le protocole d'analyse. L'approche utilisée devra dépendre de la question posée par une étude donnée.

Un désavantage de l'identification moléculaire est la **perte de données de comptage** qui peuvent être importants pour les gestionnaires. L'identification moléculaire ne pourra jamais combler ce désavantage, à moins qu'une correction par espèce soit appliquée impliquant le nombre de copie du gène d'intérêt par organisme et par espèce et l'affinité aux amorces utilisées, ce qui n'est pas faisable dans le cadre d'étude avec des centaines d'espèces. Néanmoins, l'identification morphologique a permis de retrouver les ordres d'abondances pour certains groupes. Et une corrélation entre abondance et nombre de lectures a déjà été mis en évidence dans la littérature pour différentes espèces (Bista et al, 2018). L'identification moléculaire permet donc pour une espèce donnée, voire même un ensemble d'espèces du même groupe, de faire des études comparatives (càd entre sites ou entre dates) car les biais d'amorces seront identiques et, à priori, d'avoir accès aux taxa les plus abondants en abondance relative, sauf s'il y a des biais d'amorces très forts (affinité très forte ou très faible). Comme dit précédemment, si des taxa d'intérêt pour lesquels l'information d'abondance est nécessaire ont des affinités faibles avec les amorces, il est possible d'utiliser d'autres amorces. L'accès au nombre d'individus est inatteignable avec l'identification moléculaire mais une information d'abondance relative, voire semi-quantitative est tout de même possible.

Pour finir, l'identification moléculaire permettrait de **gagner énormément de temps** (facteur 2 à 4 selon les échantillons) et de **diminuer le coût d'une étude**. A titre d'exemple, nous avons estimé le temps, le coût et les niveaux d'expertise nécessaires à cette étude dans le tableau 8 après l'étape d'échantillonnage qui est commune aux deux approches. Pour l'identification morphologique, le nombre d'échantillons traités en un jour pouvant être très variable suivant le milieu échantillonné (par exemple, un échantillon rempli d'algues est très long à trier, contrairement à un échantillon de rivière avec uniquement des graviers et sable), nous avons fait des estimations sur une large fourchette de temps passé sur un échantillon. Globalement, l'identification moléculaire permet un gain de temps et de coût par rapport à l'identification morphologique, même si celle-ci se fait à une bonne cadence de tri et de détermination. De plus, comme l'approche moléculaire est plus intégrative, il est possible de réduire le nombre d'échantillons par site et ainsi d'augmenter le nombre de sites étudiés pour le même coût. Elle permettrait donc d'augmenter l'effort d'échantillonnage spatialement et temporellement d'un suivi.

Tableau 8. Temps, coût et personnel requis pour les différentes étapes de l'identification morphologique et moléculaire. Les prix indiqués sont en Hors Taxes. Les salaires ont été pris à titre indicatif en se référant aux salaires des employé.es CNRS en prenant la moyenne entre la rémunération brute minimale et maximale en début et fin de carrière respectivement (<https://carrieres.cnrs.fr/remuneration-des-fonctionnaires/>) : salaire brut journalier Technicien.ne de Recherche = ~ 132€ ; Ingénieur.e de recherche : ~ 170€.

Identification morphologique (72 échantillons)			
Etape	Investissement minimum (hors appareils)	Consommables (prix/éch)	Temps et personnel requis (salaire brut total)
Tri et détermination	Boîtes de pétri : 51€ Pincettes : 40€ Tamis 500 µm : 120€ Tamis 2 mm : 76€	Ethanol 96% : 1€	Temps par échantillon peut varier entre 1/jour à 5/jour suivant les milieux échantillonnés Technicien.ne avec des connaissances en identification niveau genre pour plusieurs groupes taxonomiques 14-30-72 jours : 1850-3960-9515€
Entrée et traitement des données			Technicien.ne : 6 jours : 792€
Total	287€	72€	Entre 20, 36 et 78 jours : 2597-4752-10 307€

Identification moléculaire (95 échantillons)			
Préparation des échantillons avant séquençage	<i>Prétraitement</i> Billes de broyage, maille et élastique : 34€	<i>Prétraitement</i> TNES : 2,5€ <i>Extraction</i> : 8€ <i>PCR</i> : 3,8€ <i>Dosage</i> : 0,6€ <i>Purification (facultatif)</i> : 4,7€ <i>Séquençage</i> : 36,8€	Technicien.ne de biologie moléculaire 10 jours : 1320€
Traitement des données			Ingénieur.e de biologie moléculaire ou bioinformatique 1 jour (voire quelques heures si bien automatisé) si pipeline en place : 170€ 6 jours si pipeline à faire : 1020€
Total	34€	52-56,7€	Entre 11 et 17 jours : 1490-2340€

Pour conclure, l'ensemble de nos résultats suggère que l'identification moléculaire est une méthode **intégrative** qui permet **un suivi et une discrimination** des communautés à **large échelle et sur du long-terme** avec une **résolution taxinomique fine**. Elle s'accompagne d'une **perte des données de comptage** et d'un **bruit de fond qui peut être mitigé**. Bien qu'il ait été développé sur des prélèvements de filet Surber, le protocole développé dans ce projet est facilement transférable à tout type d'échantillonnage des « invertébrés » benthiques aboutissant à un échantillon contenant organismes et substrats.

Continuité du projet

Cet outil, testé dans le cadre d'un projet de démonstration, n'a pas pu être testé sur un suivi ORE. Ces premiers résultats indiquent qu'il est prêt à être testé et même utilisé sur un suivi. L'outil proposé correspond à un outil d'indice de biodiversité et non de qualité biologique/bio-indication (cf ANR AquaDNA Meyer et al, 2020) : il est à destination de l'évaluation de l'évolution d'une ou plusieurs communautés dans le temps et dans l'espace suite à une action sur celles-ci ou non ou de comparaison de communautés. Il est aussi possible, à partir du jeu de données établi, de dégrader l'information pour aboutir à des listes taxonomiques contenant uniquement les informations nécessaires aux indices de bio-indication, basés sur l'occurrence uniquement (informations intéressantes à ce sujet dans le document de Cruickshanks et al, 2022, partie « Using DNA for reporting against CBD & other policy frameworks »).

Une phase de validation dans le cadre du suivi pour cet outil serait d'être intégré au suivi scientifique minimal du réseau de site de démonstration avec une collaboration entre gestionnaires et scientifiques. Les approches moléculaire et morphologique pourraient y être couplées et cela permettrait un suivi concret sur du long terme et dans des types de cours d'eau différents (p.ex. lônes, fleuves) pour évaluer la réelle capacité de la méthode à mettre en évidence des variations des communautés dans le temps et dans le cadre d'une ORE. En parallèle, il serait intéressant d'essayer à partir des données récoltées de construire un indicateur de variation de la biodiversité pour donner un sens aux variations observées, aux actions menées. La mise en place d'un tel indicateur semble compliquée car il rejoint un questionnement qui a été mis en avant plusieurs fois lors de nos échanges avec les gestionnaires.

Ce questionnement portait sur l'interprétation biologique et, en prolongement, fonctionnelle des données issues de l'identification moléculaire : quel est l'intérêt de décrire mieux et plus de diversité spécifique si on ne connaît pas leur fonction dans les écosystèmes ? Dans le cadre d'un suivi de biodiversité cherchant à évaluer des variations dans le temps ou dans l'espace, l'information taxinomique est suffisante ; dans le cadre de la compréhension du fonctionnement des écosystèmes et l'estimation du niveau de restauration ou perturbation, l'information taxinomique est insuffisante. C'est une question qui sort de notre expertise et qui doit être prise en main par les expert.es écologues/taxonomistes/systématicien.nes qui sont capables de calculer et interpréter des traits en fonction des habitats. Malheureusement, les spécialistes

sont rares et ceci de plus en plus car peu soutenus financièrement et non remplacés avec le temps, amenant à de faible effectif pour faire avancer les connaissances taxonomiques et les exigences écologiques associées. Une solution est de réinvestir des ressources dans la description des espèces et de leur habitat pour allier identification moléculaire fine et interprétations biologiques et fonctionnelles des données, et potentiellement à un indicateur fonctionnel de variation de la biodiversité.

Mise en œuvre du protocole (ANNEXE 3)

Au niveau de l'échantillonnage, la mise en œuvre du protocole est assez simple. Tout d'abord, comme l'identification moléculaire est plus intégrative, un nombre moins important d'échantillons peut être effectué. Néanmoins, nous n'avons pas pu estimer de nombre minimum car nous n'avons prélevé que 6 sites. Si le protocole est utilisé à terme dans des suivis scientifiques minimum, il sera possible d'estimer ce nombre minimum, potentiellement dans des types de cours d'eau différents, et de le réduire au fil du temps. Au niveau du prélèvement, l'échantillonnage reste un échantillonnage classique avec la nécessité de mettre en place une bonne pratique évitant les contaminations entre échantillons (cf ANNEXE 5). L'échantillonnage pourra ainsi toujours être effectué par les parties prenantes des projets et est même applicable dans le contexte de sciences participatives (cf projet STREAM au Canada).

Les étapes de préparation des échantillons en laboratoire peuvent être effectuées dans le cadre de collaboration entre le milieu académique et les gestionnaires. Il peut également s'agir d'une demande spécifique des gestionnaires à des bureaux d'études possédant les infrastructures nécessaires (bureaux d'études faisant de la biologie moléculaire) ou des plateformes de biologie moléculaire universitaires avec le protocole à suivre. Ce protocole n'est pas un protocole fastidieux pour les personnes formées à la biologie moléculaire.

L'étape de traitement bio-informatique nécessite une personne formée en traitement des données issues de metabarcoding. Elle peut être acquise en interne si des compétences en bio-informatique sont déjà présentes chez les gestionnaires par des formations ou en autodidactie. Le traitement bio-informatique utilisé pour ce projet peut être diffusé sur demande. Il peut également s'agir d'une collaboration ou d'un financement avec le milieu académique ou d'une externalisation en bureau d'étude.

PHASE 4 : Transférabilité aux gestionnaires

Une première réunion a eu lieu le 04/05/2021 avec l'ensemble des partenaires (Florent Pellizzaro, SR3A ; Anne Clémens, GRAIE ; Laurent Simon, LEHNA/ZABR ; Stéphane Stoffek, Agence de l'eau ; Arnaud Foulquier, LECA ; Thibault Datry, INRAE ; Tristan Lefébure, LEHNA ; Mailys Gauthier, LEHNA). Cette réunion a permis de présenter le projet et les premiers résultats et de discuter autour du projet en général. Nous avons également discuté sur les attentes des gestionnaires afin de préparer au mieux la phase 4. A été mis en évidence le fait :

- que les gestionnaires n'étaient pas intéressé-es par la partie méthodologique des identifications ADN (c'est-à-dire la phase 1 et 2) mais par les possibilités que pourraient leur apporter les outils moléculaires de manière générale étant donné que de manière générale, ils et elles sous-traitent les mesures.

- que les besoins des gestionnaires portaient sur une continuité de la méthode actuelle (travailler sur les Surber permet cette continuité), sur une méthode unique, stable et partagée, sur une méthode répondant à une question écologique donnée, sur un suivi à long terme et à large échelle.

- que les gestionnaires aimeraient découvrir les méthodes de biologie moléculaire et comprendre leurs différences, suivre le train des évolutions et avoir une nuance quant à leur performance.

- que plus d'échanges entre gestionnaires et scientifiques étaient nécessaires pour identifier ce que les scientifiques peuvent leur apporter.

- que le suivi des ORE n'était pas la priorité pour elles et eux et qu'il y avait un besoin de redéfinir la manière dont on gère les bassins versants (ceci est en dehors de la portée du projet MISTRAL).

Lors de la réunion, il a donc été décidé que le Livrable 2 du projet serait plutôt un échange avec les gestionnaires qui pourraient se pérenniser à long terme pour suivre les évolutions des outils moléculaires. Un premier temps d'échanges a été décidé avec la SR3A pour avoir une présentation couplée entre gestionnaires et scientifiques.

Cet échange a eu lieu le 31/08/2021. Etaient présent-es les personnels de la SR3A (dont Florent Pellizzaro), Tristan Lefébure et Mailys Gauthier (LEHNA). FP et ses collègues ont présenté le SR3A ainsi que leurs différentes missions, suivi d'une présentation par MG des méthodes ADN et d'une discussion générale pour répondre à la question « Les méthodes moléculaires peuvent-elles répondre à certaines attentes/besoins des gestionnaires ? ». Cette réunion a permis de mettre en évidence sur quels projets du SR3A l'ADN pourrait être un outil utile (de manière générale, pas pour les aspects réglementaires qui demandent un outil agréé mais pour des suivis de biodiversité).

Le 28/08/2023, une réunion de préparation d'un café ZABR a eu lieu avec Anne Clémens (GRAIE), Samuel Monnet (SR3A), Stéphane Stroffek (Agence de l'eau), Tristan Lefébure et Mailys Gauthier (LEHNA). Cette réunion a permis de discuter en amont du café ZABR et en détail des résultats du projet et de discuter de la mise en œuvre du protocole sur des projets actuels de l'agence de l'eau et de la SR3A. Le café ZABR qui a eu lieu le 22/09/2023 a permis de diffuser les résultats à l'échelle nationale à un large public (universitaires,

collectivités, bureaux d'étude, agence de l'eau, gestionnaires des cours d'eau). La présentation et les échanges sont disponibles sur la chaîne Youtube du GRAIE, permettant de continuer la diffusion des résultats. Les échanges et demandes issus de ce café ZABR ont été intégrés dans la conclusion du livrable 2/PHASE 3 : Applicabilité d'un protocole de métabarcoding au suivi ORE.

Finalement, une fiche outil ZABR sera peut-être rédigée par le GRAIE.

5 ECHANGES ENTRE EQUIPES IMPLIQUEES & ECHANGES AVEC LES REFERENTS AGENCE DE L'EAU

Le choix et la collecte des espèces pour les communautés artificielles ont impliqué des membres de l'équipe EcoFlowS (connaissance des espèces et sites de prélèvement) et e3s (plan expérimental) au printemps 2021.

Un échange a eu lieu le 04/05/2021 avec tous-tes les partenaires du projet pour présenter le projet et le faire évoluer dans un sens pertinent aux regards des attentes des gestionnaires et de l'agence de l'eau. Une réunion à la SR3A a eu lieu le 31/08/2021.

Une réunion a eu lieu le 28/08/2023 afin de présenter les résultats finaux du projet avec l'agence de l'eau (Stéphane Stroffek) et la SR3A (Samuel Monnet) et de préparer le café ZABR qui a eu lieu le 22/09/2023.

6 REFERENCES BIBLIOGRAPHIQUES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410.
- Bista, I, Carvalho, GR, Tang, M, et al.. (2018). Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. *Mol Ecol Resour.* 18: 1020–1034. <https://doi.org/10.1111/1755-0998.12888>
- Cañedo-Argüelles M., Bogan M.T., Lytle D.A., & Prat N. (2016) Are Chironomidae (Diptera) good indicators of water scarcity? Dryland streams as a case study. *Ecological Indicators*, 71, 155-162.
- Callahan B. J., McMurdie P. J., Rosen M. J., Han A. W., Johnson A. J. A., & Holmes S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583.
- Crabot, Julie & Dolédec, Sylvain & Forcellini, Maxence & Datry, Thibault. (2021). Efficiency of invertebrate-based bioassessment for evaluating the ecological status of streams along a gradient of flow intermittence. *Ecological Indicators*. 133. 10.1016/j.ecolind.2021.108440.
- Cruickshanks K., Paupério Castro J., Tayleur C., Bakker J., Craig H., Harper L. R., Wort E., Flintham H., Victor-Matei A., Devenish C., Bruce K., Egeter B. (2022). Phase 1 Report - Developing Habitat Scale DNA Monitoring in Support of Post 2020 Biodiversity Reporting Requirements. NMP/001/20. *NatureMetrics*.
- Dolédec, S., Forcellini, M., Olivier, J. M., & Roset, N. (2015). Effects of large river restoration on currently used bioindicators and alternative metrics. *Freshwater Biology*, 60(6), 1221-1236.
- 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(7), e0130324.
- Elbrecht, V., & Leese, F. (2017). Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental Science*, 5(April), 1–11.
- 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. <https://doi.org/10.1111/2041-210X.12789>
- Elbrecht, V., & Steinke, D. (2019). Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biology*, 64(2), 380-387.
- Ewels P., Magnusson M., Lundin V. and Käller M.. (2016)MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*
- Gauthier, M., Konecny-Dupré, L., Nguyen, A., Elbrecht, V., Datry, T., Douady, C. J., & Lefébure, T. (2019). Enhancing DNA metabarcoding performance and applicability with bait capture enrichment and DNA from conservative ethanol. *Molecular Ecology Resources*, <https://doi.org/10.1111/1755-0998.13088>.

- Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., ... & Larsen, T. H. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology letters*, 16(10), 1245-1257.
- Kahlke, T, Ralph, PJ. (2019). BASTA – Taxonomic classification of sequences and sequence bins using last common ancestor estimations. *Methods Ecol Evol*. 2019; 10: 100– 103.
<https://doi.org/10.1111/2041-210X.13095>
- Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., & Zizka, V. M. A. (2020). Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification. *Environmental DNA*. <https://doi.org/10.1002/edn3.177>
- 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(1), 1–14. <https://doi.org/10.1186/1742-9994-10-34>
- Liu, M., Clarke, L.J., Baker, S.C., Jordan, G.J. & BurrIDGE, C.P. (2020). A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology*, 45, 373-385.
- Mahé F, Czech L, Stamatakis A, Quince C, de Vargas C, Dunthorn M, Rognes T. (2021) Swarm v3: towards tera-scale amplicon clustering. *Bioinformatics* doi: 10.1093/bioinformatics/btab493
- Majaneva M., Diserud O.H., Eagle S.H.C., Hajibabaei M. & Ekrem T. (2018). Choice of DNA extraction method affects DNA metabarcoding of unsorted invertebrate bulk samples. *Metabarcoding and Metagenomics*, 2(e26664). <https://doi.org/10.3897/mbmg.2.26664>
- Martin, M. (2014). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*, 17, 10. doi:10.14806/ej.17.1.200
- McMurdie P.J. and Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8(4):e61217.
- Meyer, A, Boyer, F, Valentini, A, et al. Morphological vs. DNA metabarcoding approaches for the evaluation of stream ecological status with benthic invertebrates: Testing different combinations of markers and strategies of data filtering. (2021) *Mol Ecol.*; 30: 3203–3220. doi.org/10.1111/mec.15723
- Porter, T.M., & Hajibabaei, M. (2018) Automated high throughput animal CO1 metabarcode classification. *Scientific Reports*, 8, 4226.
- Vamos, E. E., Elbrecht, V., & Leese, F. (2017). Short coi markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics*, 1, 1–20.
<https://doi.org/10.3897/mbmg.1.14625>
- Vaudor, L., Lamouroux, N., Olivier, J. M., & Forcellini, M. (2015). How sampling influences the statistical power to detect changes in abundance: an application to river restoration. *Freshwater Biology*, 60(6), 1192-1207.
- Wang, Q, G. M. Garrity, J. M. Tiedje, and J. R. Cole. (2007). Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol*. 73(16):5261-7.

7 ANNEXES

ANNEXE 1 – Tableau des tests post-hoc - Test in vitro sur des communautés artificielles

Tableau A1. Statistique *t* et significativités des comparaisons deux à deux des traitements de la matrice ADN *bulk in toto* des quantités d'ADN extraites et rapportées au mg extrait. Une *p*value<0.05 est considérée comme significative.

Paire de traitements comparée	Quantité d'ADN extraite		Quantité d'ADN rapportée au mg extrait	
	t-ratio	p-value	t-ratio	p-value
ETOHWET/BUFFER	-2.197	0.062	-0.212	0.833
ETOHWET/TNESWET	-3.467	0.006	-2.128	0.122
ETOHWET/DRY	-4.759	0.001	-2.966	0.054
ETOHWET/UREA	-0.966	0.343	-0.893	0.512
BUFFER/TNESWET	-1.271	0.256	-1.915	0.134
BUFFER/DRY	-2.563	0.039	-2.754	0.054
BUFFER/UREA	1.231	0.256	-0.681	0.559
TNESWET/DRY	-1.292	0.256	-0.838	0.512
TNESWET/UREA	2.501	0.039	1.235	0.381
DRY/UREA	3.793	0.004	2.073	0.122

Abréviations : DRY=Broyage sec ; TNESWET=Broyage avec tampon – TNES ; BUFFER=Tampon TNES (récupéré avant le broyage) ; ETOHWET=Broyage avec etoh ; UREA=Tampon TNES-Urea

Tableau A2. Statistique *t* et significativités des comparaisons des protocoles de la matrice ADN *bulk in toto* au protocole de référence (ADN *bulk*) des quantités d'ADN extraites et rapportées au mg extrait. Une *p*value<0.05 est considérée comme significative.

Protocole comparé	Quantité d'ADN extraite		Quantité d'ADN rapportée au mg extrait	
	t-ratio	p-value	t-ratio	p-value
ETOHWETPCI	-8.738	<0.0001	-33.075	<0.0001
ETOHWETPMS	4.456	0.0002	-26.256	<0.0001
ETOHWETPSP	-17.989	<0.0001	-33.102	<0.0001
BUFFERPCI	-5.277	<0.0001	-32.891	<0.0001
BUFFERPSP	-9.590	<0.0001	-32.062	<0.0001
DRYPCI	-0.755	0.4946	-32.164	<0.0001
DRYPMS	6.229	<0.0001	-22.223	<0.0001
DRYPSP	-0.276	0.7843	-16.918	<0.0001
TNESWETPCI	-6.604	<0.0001	-33.011	<0.0001
TNESWETPMS	7.674	<0.0001	-19.036	<0.0001
TNESWETPSP	-1.402	0.2030	-20.902	<0.0001
UREAPCI	-17.989	<0.0001	-33.102	<0.0001
UREAPSP	-3.522	0.0019	-27.929	<0.0001

Abréviations : Traitement : DRY=Broyage sec ; TNESWET=Broyage avec tampon – TNES ; BUFFER=Tampon TNES (récupéré avant le broyage) ; ETOHWET=Broyage avec etoh ; UREA=Tampon TNES-Urea ; Méthode d'extraction : PCI=Phenol :Chloroform ; PMS=PowerMaxSoil ; PSP=PowerSoilPro.

Tableau A3. Statistique *t* et significativités des comparaisons deux à deux des traitements de la matrice ADN *bulk in toto* des nombres de copies observés moyens et rapportés au mg extrait. Une *p*value<0.05 est considérée comme significative.

Paire de traitements comparée	Nombre de copies moyen		Nombre de copie moyen rapporté au mg extrait	
	t-ratio	p-value	t-ratio	p-value
ETOHWET/BUFFER	-4.798	<.0001	-2.929	0.004
ETOHWET/TNESWET	-13.106	<.0001	-9.970	<.0001
ETOHWET/DRY	-17.902	<.0001	-12.870	<.0001
ETOHWET/UREA	-7.903	<.0001	-6.483	<.0001
BUFFER/TNESWET	-8.308	<.0001	-7.041	<.0001
BUFFER/DRY	-13.104	<.0001	-9.941	<.0001
BUFFER/UREA	-3.104	0.002	-3.554	0.0006
TNESWET/DRY	-4.796	<.0001	-2.900	0.004
TNESWET/UREA	5.204	<.0001	3.487	0.0007
DRY/UREA	9.999	<.0001	6.388	<.0001

Abréviations : DRY=Broyage sec ; TNESWET=Broyage avec tampon – TNES ; BUFFER=Tampon TNES (récupéré avant le broyage) ; ETOHWET=Broyage avec etoh ; UREA=Tampon TNES-Urea

Z A B R

Zone Atelier Bassin du Rhône

Tableau A4. Statistique *t* et significativités des comparaisons des protocoles de la matrice ADN *bulk in toto* au protocole de référence (ADN *bulk*) des nombres de copies observés et rapportés au mg extrait. Une *p*value<0.05 est considérée comme significative.

Protocole comparé	Nombres de copies observés						
	Aa	Br	Cf	Cr	Pa	Rs	Tt
ETOHWETPCI	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ETOHWETPMS	0.889	0.784	0.857	0.411	0.591	0.843	0.938
ETOHWETPSP	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BUFFERPCI	0.000	0.000	0.001	0.000	0.000	0.000	0.000
BUFFERPSP	0.001	0.000	0.006	0.000	0.000	0.000	0.000
DRYBnT	0.000	0.000	0.000	0.000	0.000	0.000	0.000
DRYPCI	0.298	0.107	0.247	0.008	0.007	0.027	0.003
DRYPMS	0.325	0.672	0.748	0.493	0.767	0.424	0.802
DRYPSP	0.547	0.226	0.310	0.013	0.023	0.184	0.166
TNESWETPCI	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TNESWETPMS	0.153	0.160	0.271	0.005	0.290	0.281	0.371
TNESWETPSP	0.245	0.129	0.307	0.009	0.073	0.018	0.203
UREAPCI	0.000	0.000	0.000	0.000	0.000	0.000	0.000
UREAPSP	0.233	0.064	0.154	0.000	0.004	0.001	0.209
Protocole comparé	Nombres de copies observés rapportés au mg extrait						
	Aa	Br	Cf	Cr	Pa	Rs	Tt
ETOHWETPCI	0	0	0	0	0	0	0
ETOHWETPMS	0	0	0	0	0	0	0
ETOHWETPSP	0	0	0	0	0	0	0
BUFFERPCI	0	0	0	0	0	0	0
BUFFERPSP	0	0	0	0	0	0	0
DRYBnT	0	0	0	0	0	0	0
DRYPCI	0	0	0	0	0	0	0
DRYPMS	0	0	0	0	0	0	0
DRYPSP	0	0	0	0	0	0	0
TNESWETPCI	0	0	0	0	0	0	0
TNESWETPMS	0	0	0	0	0	0	0
TNESWETPSP	0	0	0	0	0	0	0
UREAPCI	0	0	0	0	0	0	0
UREAPSP	0	0	0	0	0	0	0

Abréviations des espèces : Cf=Corbicula fulminea ; Aa=Asellus aquaticus ; Br=Brachyptera risi ; Tt=Tubifex tubifex ; Cr= Chironomus riparium ; Rs=Riolus subviolaceus ; Pa=Potamopyrgus antipodarum

Tableau A5. Statistique *t* et significativités des comparaisons deux à deux des méthodes d'extraction de la matrice ADN *bulk in toto* des nombres de copies moyens rapportés au mg extrait. Une *pvalue*<0.05 est considérée comme significative.

	<i>t</i> -ratio	<i>pvalue</i>
PCI/PSP	-12.754	<0.0001
PCI/PMS	-20.364	<0.0001
PSP/PMS	-7.610	<0.0001

ANNEXE 2 – Protocole TRI ET DETERMINATION DU MACROZOOBENTHOS / MODIFIE DE NF T90-388 (NORME I2M2)

1. Principe

L'analyse consiste à :

(a) préparer l'échantillon : si besoin, laver le substrat pour éliminer les conservateurs éventuels et les éléments fins, organiques ou minéraux qui pourraient gêner le tri et retirer les éléments grossiers (graviers, feuilles...). Le cas échéant, constituer des fractions (séparation sur colonne de tamis, par élutriation) pour faciliter le tri ultérieur.

(b) trier et déterminer les taxons à un niveau au moins égal à celui demandé par le protocole choisi, pour établir une liste faunistique.

2. Matériel

(a) Matériel de préparation de l'échantillon

Gants

Tamis de vide de maille 500 µm,

Tamis de maille 100 µm pour récupérer les éléments <500 µm lors du lavage

Tamis 2 mm pour séparation en 2 fractions si besoin

Grand récipient pour élutriation permettant de réaliser un mouvement suffisant sans perdre une partie de l'échantillon

(b) Matériel de tri

Gants

Pipettes de tri (facultatif et à usage unique (1/échantillon))

Pissette d'eau et d'alcool

Matériel optique au grossissement x2 minimum pour le tri du refus des plus petits tamis

2 jeux de :

Pincés ultra fines, pincés souples

Récipient de tri à fond clair + récipients (par exemple boîte de Pétri)

(c) Matériel de détermination

Loupe binoculaire avec grossissement minimum x45 (détermination niveau A) et un appareil optique avec grossissement x80 (détermination niveau B)

Eclairage suffisamment puissant

Documents pour la détermination taxinomique

(d) Matériel de décontamination

Gants

Sopalin

Bassine

Solution décontaminante (10% javel, 1% NaOH, 1% détergent en poudre) diluée 3 fois

1% NaOH (200g per 20L)

1% détergent (200g per 20L)

Eau déminéralisée

3. Étapes d'analyse des échantillons

Les paires de gants doivent être changées entre chaque échantillon ou lorsqu'une contamination des gants est supposée.

(a) Préparation de l'échantillon

- Tout le matériel utilisé doit être décontaminé avant utilisation.

- Lavage pour éliminer les conservateurs éventuels et les sédiments fins qui pourraient gêner le tri ultérieur.

Il est réalisé sur une colonne de tamis avec un tamis 100 µm et 500 µm. Le lavage doit être doux pour ne pas dégrader les organismes. Il est conseillé de le réaliser en plaçant la colonne de tamis dans une cuvette.

- Retrait des éléments plus grossiers dans le tamis 500 µm (pierres, branchages...) après contrôle visuel de l'absence d'organismes.

- Le refus de tamis 100 µm et les éléments grossiers doivent être replacés dans le flacon d'échantillonnage.

- Constitution de fractions (facultatif) : cette étape permet d'augmenter l'efficacité du tri. Les techniques suivantes peuvent être utilisées selon la nature des substrats à traiter :

- Passage sur une colonne de tamis 500 µm et 2 mm pour séparer l'échantillon

- Elutriation pour séparer la fraction surnageante et le refus d'élutriation

- « Nettoyage » des végétaux (séparer feuilles et tiges en les agitant) (bryophytes, hydrophytes, héliophytes) en pleine eau et séparation du surnageant contenant la majorité des organismes et le refus

(b) Reconnaissance des taxons

Tous les organismes présents dans l'échantillon doivent être :

- déterminés dans les limites de tailles (0,5 mm), de grossissements des appareils optiques fixés dans le présent document.

- inscrits sur la(les) liste(s) faunistique(s) du point de prélèvement au niveau de détermination requis (niveau A, souvent la famille, ou niveau B, souvent le genre). Ce niveau peut être exceptionnellement moins précis uniquement pour les nymphes et pour les organismes trop jeunes ou abîmés qui ne peuvent être déterminés avec certitude au niveau requis.

- le tableau en fin d'annexe fixe les limites de détermination et le nombre d'organismes à déterminer par groupe.

- pour les taxons non déterminés avec certitude, ajouter l'information en commentaire.

(c) Tri, différenciation, dénombrement et extraction des individus

- **Tous les organismes doivent être extraits, comptés et identifiés au niveau MISTRAM puis remis dans le flacon d'échantillonnage avec de l'étoh.**

- **Tout le matériel utilisé doit être décontaminé avant utilisation.**

- La totalité de l'échantillon doit être observée, même en cas de fractionnement. Une fraction peut elle-même être divisée en plusieurs unités de tri.

- Pour permettre une bonne visibilité et assurer une distinction efficace des particules minérales, organiques et des organismes, il convient de placer chaque unité de tri en quantité limitée dans un récipient, d'adapter le niveau d'eau dans ce récipient et d'utiliser une source lumineuse suffisante.

- Si l'unité de tri à examiner provient d'un refus de tamis 500 µm ou plus, les taxons sont tous visibles à l'œil nu et l'usage d'un grossissement n'est pas nécessaire. Si elle provient d'un refus de tamis de maille inférieure, le tri doit être finalisé à l'aide d'un matériel optique grossissant au minimum 2 fois.

- Parasites : certains taxons parasites sont à examiner sur (groupe 18-Branchiobdellida sur les carapaces d'écrevisses) ou à l'intérieur d'un hôte (groupe 10-Hymenoptera dans les fourreaux des trichoptères et ou de diptères ; groupe 24-Nemathelminthe dans de nombreux taxons). Un parasite visible à l'intérieur d'un hôte non disséqué est comptabilisé, ainsi que l'hôte. Concernant les hyménoptères parasites de trichoptères, seuls les individus à l'intérieur des fourreaux sont pris en compte. L'hôte n'est pas compté s'il n'est plus visible.

- Les exuvies, les fourreaux et les coquilles vides sans partie organiques, les statoblastes de Bryozaires et les gemmules de Spongiaires ne sont pas pris en compte mais leur présence peut être signalée en commentaires.

- **Le comptage doit être exhaustif.**

- **Une liste faunistique doit être établie en indiquant les taxons et les effectifs trouvés par échantillon.**

(d) Décontamination

- Pulvériser l'ensemble du matériel réutilisable avec la solution décontaminante et le placer 5 mn dans une bassine propre.

- Rincer abondamment à l'eau déminéralisée et laisser sécher ou essuyer avec du sopalin.

Z A B R

Zone Atelier Bassin du Rhône

Limite de détermination et nombre d'individus à déterminer par groupe

Groupe	Niveau de détermination A	Niveau de détermination B	Niveau de détermination MISTRAL
1-PLECOPTERA	Famille	Genre	Genre
2-TRICHOPTERA	Famille	Genre sauf sous-famille pour les Glossosomatidae et les Limnephilidae	Genre sauf sous-famille pour les Limnephilidae
3-Ephemeroptera	Famille	Genre	Genre
4-Heteroptera	Famille	Famille sauf sous-famille pour les Corixidae	Famille sauf sous-famille pour les Corixidae
5-Coleoptera	Famille	Genre sauf sous-famille pour les Dysticidae et Hydrophilidae	Genre
6-Diptera	Famille	Famille	Tribu
7-Odonata	Famille	Genre	Genre
8-Megaloptera	Famille	Genre	Genre
9-Planipenne	Famille	Genre	Genre
10-Hymenoptera	HYMENOPTERA	HYMENOPTERA	HYMENOPTERA
11-Lepidoptera	Crambidae	<i>Paraponyx</i>	<i>Paraponyx</i>
12-Crustacea-Malacostraca	Famille	Genre sauf famille pour les Asellidae	Genre sauf famille pour les Asellidae (essayer d'aller le plus loin possible)
13-Crustacea-Branchiopoda-Phyllozoa	Famille	Famille	Famille
14-Crustacea (autres)	Branchiura Cladocères Copépodes Ostracodes	Branchiura Cladocères Copépodes Ostracodes	Branchiura Cladocères Copépodes Ostracodes
15-Hydracarina	HYDRACARINA	HYDRACARINA	HYDRACARINA
16-Bivalvia	Famille	Genre	Genre
17-Gastropoda	Famille	Genre sauf famille pour les Planorbidae	Genre sauf famille pour les Planorbidae
18-Branchiobdellida	BRANCHIOBDELLIDA	BRANCHIOBDELLIDA	BRANCHIOBDELLIDA
19-Achaeta-Hirudinea	Famille	Famille	Genre
20-Oligochaeta	OLIGOCHAETA	OLIGOCHAETA	OLIGOCHAETA
21-Polychaeta	Famille	Genre	Genre
22-Turbellaria-Tricladida	Famille	Famille	Famille
23-Nemertea	Famille	Genre	Genre
24-Nemaltheiminta	NEMALTHELMINTHA	NEMALTHELMINTHA	NEMALTHELMINTHA
25-Hydrozoa	HYDROZOA	HYDROZOA	HYDROZOA

Z A B R

Zone Atelier Bassin du Rhône

26-PORIFERA-SPONGILLIDAE	SPONGILLIDAE	SPONGILLIDAE	SPONGILLIDAE
27-BRYOZOA	BRYOZOA	BRYOZOA	BRYOZOA
28-ROTIFERA	ROTIFERA formant des colonies	ROTIFERA formant des colonies	ROTIFERA formant des colonies

ANNEXE 3 – Protocole final pour le suivi des communautés de macrozoobenthos**ECHANTILLONNAGE****Matériel :**

Matériel à décontaminer avant l'échantillonnage	Autre matériel
<p><u>Protocole décontamination javel/détergent + autoclave</u></p> <ul style="list-style-type: none"> - flacons de prélèvements autoclavables (possible d'acheter des flacons neufs stériles qui n'auront pas besoin d'être décontaminés) - tamis 500 µm en inox <p><u>Protocole de décontamination javel</u></p> <ul style="list-style-type: none"> - filet Surber (cadre 1/20 de m², 500µm de vide de maille) ou autre filet de prélèvement - seaux de prélèvement - entonnoir - glacière pour ranger les échantillons 	<ul style="list-style-type: none"> - waders - gants en nitrile ou latex - ethanol 96% + pissette ethanol - jerrycan d'eau déminéralisée - seau rempli de javel 2,5% - seau « déchets liquides » - blocs froids pour la glacière - produit décontaminant (2,5% javel, 0,25% NaOH, 0,25% détergent) dans un pulvérisateur à gâchette. <p><i>Il est possible de faire une solution mère de solution décontaminante (10% javel, 1% NaOH et 1% détergent) que l'on dilue avec de l'eau déminéralisée 2 fois pour avoir la concentration finale souhaitée.</i></p>

Avant l'échantillonnage, il est nécessaire de décontaminer le matériel « sensible », c'est-à-dire qui va rentrer en contact avec l'échantillon, s'il n'est pas neuf ou stérile.

Protocole décontamination javel + autoclave

1. Nettoyer le matériel avec un détergent.
2. Si le matériel n'est pas en inox, le placer dans un bain de javel 5% pendant 10 mn. Le rincer abondamment à l'eau déminéralisée et le laisser sécher.
3. Passer le matériel à l'autoclave.
Attention, certains matériels ont des parties non autoclavables (p.ex. joint sur les tamis) et doivent être retirés pour l'autoclave.
4. Placer le matériel sous des UVs pendant 15 mn.

Protocole décontamination javel

Ce protocole est fait pour le gros matériel ou le matériel qui ne passe pas à l'autoclave

1. Nettoyer le matériel avec un détergent.
2. Dans une bassine ou évier, pulvériser sur le matériel du produit décontaminant et laisser agir 5 mn
3. Rincer abondamment à l'eau déminéralisée et laisser sécher.

Echantillonnage

Le protocole ci-après est issu des protocoles d'échantillonnage Surber mais tout autre type de prélèvement consistant à récupérer des organismes et leur substrat est adapté.

Si plusieurs prélèvements sont faits sur un même tronçon, il est nécessaire de réaliser les prélèvements de l'aval vers l'amont pour éviter la gêne occasionnée par le trouble éventuel de l'eau, de récolter des invertébrés en dérive et d'endommager des habitats non prélevés.

Tout l'échantillonnage doit se faire en portant des gants qui doivent être changés entre chaque échantillon.

1. Placer le filet Surber sur la zone à échantillonner face au courant et agiter, frotter et/ou gratter l'habitat dans le but de décrocher les individus fixés. Une partie du substrat peut être ramener dans le filet à la main.

Veiller à ce que le courant entraîne les éléments récoltés (macroinvertébrés et, le cas échéant, le substrat) dans le filet. Pour les courants lents, l'entraînement dans le filet peut être réalisé par le mouvement de la main.

2. Verser le contenu du filet dans un seau de prélèvement et nettoyer les éléments volumineux.
3. Récupérer les organismes (et une partie du substrat) dans un tamis 500 µm par élutriation. Bien vérifier qu'il ne reste pas d'organismes résistants à l'élutriation dans l'échantillon.
4. A l'aide d'une pissette contenant de l'éthanol 96%, vider le contenant du tamis dans un flacon de prélèvement.
5. Remplir le flacon avec de l'éthanol 96% à l'aide de l'entonnoir pour qu'il y ait au minimum un volume d'échantillon pour un volume d'éthanol.
6. Placer le flacon de prélèvement dans la glacière contenant des blocs froids.

Entre chaque échantillon : décontamination du matériel

1. Après les avoir rincés dans la rivière, secouer le Surber et le tamis dans le seau « javel » et les laisser moins d'une minute (pour ne pas trop les abîmer car la javel érode l'acier inoxydable qui finira par rouiller).
L'entonnoir peut aussi suivre ce protocole sans être rincé dans la rivière s'il a touché l'échantillon.
2. Changer de gants.
3. Les rincer à l'eau déminéralisée au-dessus du seau « déchets liquides ».
4. Un nouvel échantillon peut être fait.

En retour d'échantillonnage, les échantillons doivent être préservés à 4°C.

TRAITEMENT PRE-EXTRACTION : BROYAGE TAMPON

Matériel :

Matériel à autoclaver	Autre matériel
<p>JOUR 1</p> <ul style="list-style-type: none"> - bécher - mailles nylon (A/éch)/tamis (1/éch) avec mailles 600 µm - élastiques (si filet, 1/filet) - TNES (100 mM Tris-HCl, 100 mM Edta, 1,5 M NaCl) <p>JOUR 2</p> <ul style="list-style-type: none"> - bécher - billes acier inoxydable 10 mm (1/éch) - billes acier inoxydable 5 mm (500g/éch) - pinces (1/éch) - spatules en acier (1/éch) ou spatules stériles (ne pas autoclaver) 	<ul style="list-style-type: none"> - gants en nitrile ou latex - papier de protection de paille (facultatif) - produit décontaminant (2,5% javel, 0,25% NaOH, 0,25% détergent) dans un pulvérisateur à gâchette. <p><i>Il est possible de faire une solution mère de solution décontaminante (10% javel, 1% NaOH et 1% détergent) que l'on dilue avec de l'eau déminéralisée 2 fois pour avoir la concentration finale souhaitée</i></p>

Ce protocole se déroule sur 2 jours. L'idéal est de manipuler sous hotte aspirante pour limiter la contamination.

Autoclaver le matériel avant utilisation : mailles nylon ou tamis, élastiques, billes, pinces.

JOUR 1

Passer tout le matériel sous UVs 15 mn avant de commencer (bécher, élastiques, mailles Jour 1/tamis, benchtop, TNES)

Penser à faire un contrôle négatif consistant en un flacon de prélèvement contenant 100 mL d'etoh et qui suit le même protocole.

1. Vider l'etoh du flacon à l'aide d'une maille nylon 600 µm maintenue avec un élastique/tamis dans un bécher.

Lors du versement d'etoh ou de TNES, faire attention à ce que la maille soit bien tendue et bien tirée au niveau de l'élastique pour éviter les fuites

Bien vérifier qu'il ne reste pas de matière (en particulier des organismes) sur la maille. S'il en reste, les remettre dans le flacon (p.ex. en tapotant le dessus de la maille)

2. Ajouter dans le flacon du tampon TNES (100 mM Tris-HCl, 100 mM Edta, 1,5 M NaCl) pour avoir un ratio 2 :1 (TNES : échantillon) et homogénéiser (12 A/R).

Pour le blanc : 100 mL de TNES

3. Laisser les échantillons dans le TNES toute la nuit à 4°C.

Traitement des déchets et matériels fin du jour 1

- Vider les béciers d'éthanol dans un bidon qui sera traité comme un déchet 14 06 03/UN 1993.
- Mettre les élastiques et mailles/tamis sales de côté pour être nettoyé à la fin du Jour 2.

JOUR 2

Ce protocole utilise un GenoGrinder (SPEX® SamplePrep, USA) qui est un broyeur utilisant des auxiliaires de broyage (billes) pour broyer. Les échantillons sont placés sur un support et le broyage se fait par de rapides mouvements verticaux permettant aux billes d'impacter les échantillons. D'autres machines peuvent être utilisées. L'objectif est qu'à la fin du broyage, l'échantillon soit homogénéisé.

Passer tout le matériel sous UVs 15 mn avant de commencer (bécher, élastiques, mailles Jour 2, pinces, spatules, billes, benchtop)

1. Récupérer les flacons du jour 1 et les placer dans de la glace.
2. Vider le TNES du flacon à l'aide d'une maille nylon 600 µm maintenue avec un élastique dans un bécher.
Lors du versement d'etoh ou de TNES, faire attention à ce que la maille soit bien tendue et bien tirée au niveau de l'élastique pour éviter les fuites
Bien vérifier qu'il ne reste pas de matière (en particulier des organismes) sur la maille. S'il en reste, les remettre dans le flacon (p.ex. en tapotant le dessus de la maille)
3. Ajouter 1 bille en acier inoxydable 10 mm et 100g de billes en acier inoxydable 5 mm dans le flacon.
4. Nettoyer la paillasse et remplacer le benchtop par un nouveau.
5. Broyer à l'aide d'un GenoGrinder 2010 comme suit : 3 cycles de (5*2 mn, 30 sec de pause, 1500 rpm), 10 mn pause entre chaque cycle (pour que l'échantillon ne chauffe pas trop).
Faire un cycle et voir l'état de l'échantillon avant d'en relancer un, stopper quand l'échantillon est broyé au maximum
6. A nouveau sous la hotte, passer directement à la pesée pour l'extraction (cf fiche Extraction, étape 1).
A cette étape, on peut également enlever les billes de l'échantillon.
7. Passer directement à l'extraction ou conserver à 4°C ou à -20 °C jusqu'à l'extraction.

Traitement des déchets et matériel fin du jour 2

- Vider le bécher de tampon
- Mettre les élastiques et mailles sales/tamis de côté
- Récupérer les billes (avec l'aide d'une pince si besoin) dans les échantillons. *Cette étape peut se faire pendant la mise en tubes d'extraction ou après.*
- Laver les élastiques en les passant à l'eau déminéralisée
- Décontaminer le matériel (bécher, billes, pinces, spatules, mailles/tamis) : asperger de produit et laisser agir une 10aine de minutes, puis rincer à l'eau déminéralisée

EXTRACTION KIT QIAGEN DNEASY POWERSOILPRO (modifié à partir du protocole Qiagen)Matériel :

- kit Qiagen DNeasy PowerSoil Pro (Ref Qiagen : 47014)
- spatules stériles ou autoclavées et passées aux UVs
- centrifugeuse
- broyeur/homogénéiseur
- vortex
- pipettes P1000 et P200 et cônes à filtre associés
- gants en nitrile ou latex

Penser à faire un contrôle négatif (blanc d'extraction) qui suit le même protocole sans ajout de matériel à l'étape 1b.

1. a. Faire un spin du tube PowerBead Pro pour s'assurer que les billes se sont déposées au fond.
- b. Ajouter jusqu'à 250 mg d'homogénat dans le tube Power Bead Pro à l'aide d'une spatule propre. (250 µL de tampon TNES pour le contrôle négatif du traitement pré-extraction)
- c. Ajouter 800 µl de la solution CD1.

Une fois l'échantillon chargé dans le tube PowerBead Pro, l'étape suivante est une procédure d'homogénéisation et de lyse. Le tube PowerBead Pro contient un tampon qui va aider à disperser les particules de sol, commencer à dissoudre les acides humiques et protéger les acides nucléiques de la dégradation.

2. Homogénéiser soigneusement les échantillons à l'aide du GenoGrinder : 10 mn, 1750 rpm.

Un autre appareil peut être utilisé pour homogénéiser les échantillons : vortex avec un adaptateur pour tubes 2 mL à vitesse maximale, 10 mn ; PowerLyzer 24 Homogenizer, 2(2000 rpm, 30 s) avec une pause de 30 s ; broyeur horizontal, 2*(25Hz, 5 mn) en inversant les tubes de côté entre les deux broyages*

L'agitation est essentielle pour une homogénéisation et une lyse cellulaire complètes. Les cellules sont lysées par une combinaison d'agents chimiques de l'étape 1 et d'une agitation mécanique introduite à cette étape. L'agitation aléatoire des billes en présence d'agents perturbateurs entraînera la collision des billes avec les cellules et l'ouverture de ces dernières.

3. Centrifuger le tube PowerBead Pro à 15 000 g pendant 1 min.
4. Transférer le surnageant dans un tube Microcentrifuge 2 mL (fourni).
5. Ajouter 200 µL de la solution CD2 et vortexer pendant 5 s.

La solution CD2 contient un réactif capable de précipiter les matières organiques et organiques non liées à l'ADN, y compris les substances humiques, les débris cellulaires et les protéines. Il est important d'éliminer les matières organiques et inorganiques contaminantes qui peuvent réduire la pureté de l'ADN et la qualité de l'échantillon et d'entraver les applications d'ADN en aval.

6. Centrifuger à 15 000 g pendant 1 min. En évitant le culot, transférer jusqu'à 700 µL de surnageant dans un tube Microcentrifuge 2 mL.

À ce stade, le culot contient des matières organiques et inorganiques autres que l'ADN, notamment des acides humiques, des débris cellulaires et des protéines. Pour obtenir les meilleurs rendements et la meilleure qualité d'ADN, éviter de transférer une partie du culot.

7. Ajouter 600 μ L de la solution CD3 et vortexer pendant 5 s.
La solution CD3 est une solution saline très concentrée. Étant donné que l'ADN se lie étroitement à la silice à des concentrations élevées de sel, la solution CD3 ajustera la concentration de sel de la solution d'ADN pour permettre la fixation de l'ADN, mais pas des matières organiques et inorganiques autres que l'ADN qui peuvent encore être présentes à de faibles concentrations sur la colonne MB Spin Column.
8. Transférer 650 μ L du lysat sur une colonne MB Spin Column et centrifuger à 15 000 g pendant 1 min.
L'ADN est sélectivement lié à la membrane de silice de la colonne MB Spin Column en présence d'une solution saline élevée. Les contaminants passent à travers la membrane, ne laissant que l'ADN lié à la membrane.
9. Jeter l'éluat de centrifugation et répéter l'étape 8 pour s'assurer que tout le lysat est passé dans la colonne MB Spin Column.
10. Placer soigneusement la colonne MB Spin Column dans un tube Collection tube de 2 ml (fourni). Éviter d'éclabousser l'éluat de centrifugation sur la colonne MB Spin Column.
11. Ajouter 500 μ L de la solution EA dans la colonne MB. Centrifuger à 15 000 g pendant 1 min.
La solution EA est un tampon de lavage qui élimine les protéines et autres contaminants non aqueux de la membrane de la colonne.
12. Jeter l'éluat de centrifugation et replacer la colonne dans le même tube.
13. Ajouter 500 μ L de la solution C5 à la colonne. Centrifuger à 15 000 g pendant 1 minute.
La solution C5 est une solution de lavage à base d'éthanol utilisée pour nettoyer davantage l'ADN lié à la membrane. Cette solution de lavage élimine les sels résiduels, l'acide humique et d'autres contaminants tout en permettant à l'ADN de rester lié à la membrane de silice.
14. Jeter l'éluat de centrifugation.
15. Centrifuger jusqu'à 16 000 g pendant 2 minutes. Placer soigneusement la colonne dans un tube d'élution de 1,5 ml (fourni).
Cette centrifugation permet d'éliminer les résidus de la solution C5. Il est essentiel d'éliminer toute trace de solution C5 car l'éthanol qu'elle contient peut interférer avec les applications d'ADN en aval, telles que la PCR, les digestions par restriction et l'électrophorèse sur gel.
16. Ajoutez 70 μ L de la solution C6 au centre de la membrane.
Lorsque la solution C6 traverse la membrane de silice, l'ADN qui était lié en présence de sel élevé est sélectivement libéré par la solution C6 (Tris 10 mM) qui ne contient pas de sel.
17. Centrifuger à 15 000 g pendant 1 min. Jeter la colonne MB Spin Column. L'ADN est maintenant prêt pour les applications en aval.

L'ADN peut être conservé à 4°C s'il est utilisé rapidement ou conservé à -20°C pour une conservation à plus long terme.

EXTRACTION KIT QIAGEN DNEASY POWERSOILPRO (modifié à partir du protocole Qiagen) English version

Remember to make a negative control consisting of a tube filled with 250 µL of TNES used for the samples and which follows the same protocol.

1. a. Spin the PowerBead Pro Tube briefly to ensure that the beads have settled at the bottom.
b. Add up to 250 mg of matrice in the Power Bead Pro Tube with a clean spatula.
(250 µL of TNES buffer for negative control)
c. Add 800 µL of solution CD1.
2. Homogenize samples thoroughly using the GenoGrinder: 10 mn, 1750 rpm.
Another device can be used to homogenise the samples: vortex with an adaptor for 2mL tubes at maximum speed, 10 min; PowerLyzer 24 Homogenizer, 2(2000 rpm, 30 s) with a 30 s pause; horizontal grinder, 2*(25Hz, 5 min) by inverting the tubes between the two grindings.*
3. Centrifuge the PowerBead Pro Tube at 15,000 x g for 1 min.
4. Transfer the supernatant to a clean 2 ml Microcentrifuge Tube (provided).
5. Add 200 µl of Solution CD2 and vortex for 5 s.
6. Centrifuge at 15,000 x g for 1 min. Avoiding the pellet, transfer up to 700 µl of supernatant to a clean 2 ml Microcentrifuge Tube (provided).
7. Add 600 µl of Solution CD3 and vortex for 5 s.
8. Load 650 µl of the lysate onto an MB Spin Column and centrifuge at 15,000 x g for 1 min.
9. Discard the flow-through and repeat step 8 to ensure that all of the lysate has passed through the MB Spin Column.
10. Discard the flow-through. Avoid splashing any flow-through onto the MB Spin Column.
11. Add 500 µl of Solution EA to the MB Spin Column. Centrifuge at 15,000 x g for 1 min.
12. Discard the flow-through and place the MB Spin Column back into the same 2 ml Collection Tube.
13. Add 500 µl of Solution C5 to the MB Spin Column. Centrifuge at 15,000 x g for 1 min.
14. Discard the flow-through.
15. Centrifuge at up to 16,000 x g for 2 min. Carefully place the MB Spin Column into a new 1.5 ml Elution Tube (provided).
16. Add 70 µl of solution C6 to the center of the white filter membrane.
17. Centrifuge at 15,000 x g for 1 min. Discard the MB Spin Column. The DNA is now ready for downstream applications.

DNA can be stored at 4°C for rapid use or at -20°C for longer term storage.

DOSAGE ADN - Qubit dsDNA High Sensitivity (ref Q32851, Invitrogen)

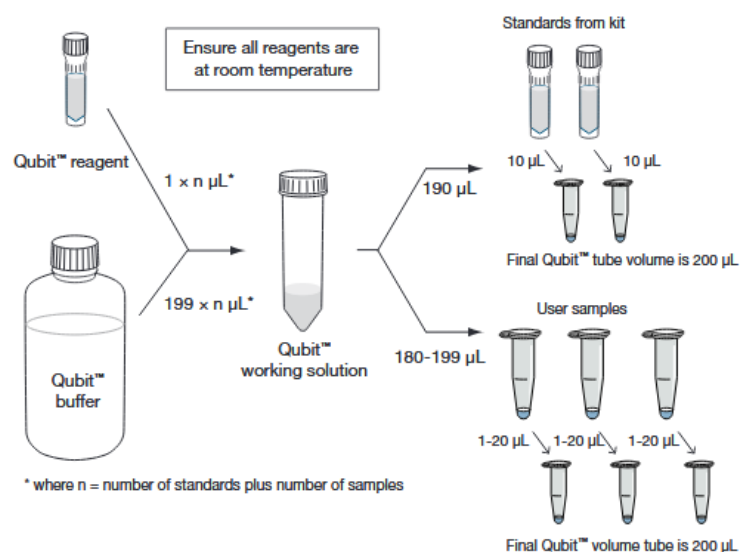
Matériel :

- kit Qubit dsDNA High Sensitivity et tubes pour Qubit (ref Q32856)
- pipettes P10, P200 et P1000 et cônes à filtre associés
- tube Falcon™
- vortex
- fluorimètre Qubit
- gants en nitrile ou latex

1. Préparez deux tubes pour les standards et un tube à essai pour chaque échantillon.
2. Préparer la *working solution* (tampon + réactif) en diluant le réactif Qubit 1:200 dans le tampon Qubit : préparer 200 µl de solution chaque standard et échantillon.
Le fluorimètre Qubit 5 calcule automatiquement le volume de tampon et le volume.
3. Dans les tubes standards, mettre 10 µL de standard et 190 µL de solution de travail. Pour les échantillons, mettre 2 µL d'ADN et 198 µL de solution de travail.
4. Vortexer tous les tubes pendant 2 à 3 secondes.
5. Incuber les tubes pendant 2 minutes à température ambiante. Les tubes sont prêts pour la mesure dans le fluorimètre Qubit.

Note : Si vous mesurez beaucoup d'échantillons en même temps, faites des séries de 24 échantillons par exemple pour une mesure précise car la fluorescence diminue avec le temps.

Remarque : si la mesure de l'échantillon est Out of Range (trop élevée), diluez l'échantillon par dix et recommencez la mesure.

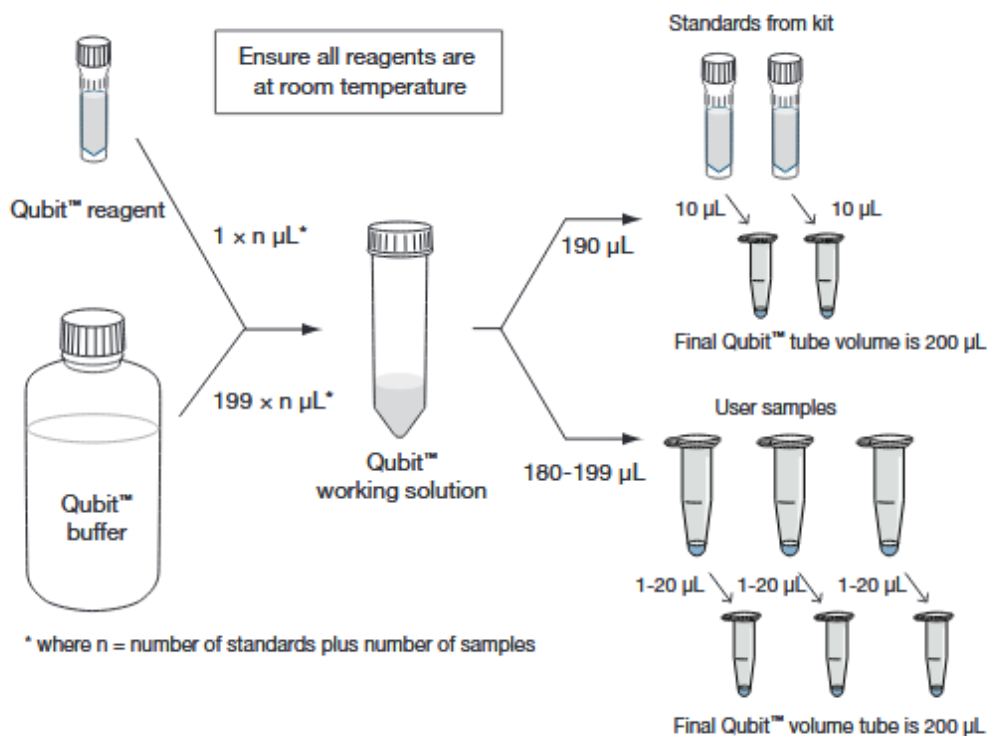


DOSAGE ADN - Qubit dsDNA High Sensitivity (ref Q32851, Invitrogen)
English version

1. Set up two assay tubes for the standards and one assay tube for each sample.
2. Prepare the working solution (buffer + reagent) by diluting the Qubit reagent 1:200 in Qubit buffer: prepare 200 μL of working solution for each standard and sample.
The Qubit 5 Fluorometer calculate the volume of buffer and volume automatically.
3. In the standard tubes, put 10 μL of standard and 190 μL of working solution. For the samples, 2 μL of DNA and 198 μL of working solution.
4. Vortex all tubes for 2-3 seconds.
5. Incubate the tubes for 2 mn at room temperature. Tubes are ready for measurement in the Qubit Fluorometer.

Note: If you measure a lot of samples at the same time, make series of e.g. 24 samples for precise measurement because fluorescence decrease over time.

Note: If sample measurement is Out of Range (Too High), dilute sample by ten and measure again.



PREPARATION AU SEQUENCAGE ADN

Matériel

Préparation des échantillons

- eau ultra-pure pour PCR
- tubes eppendorf 1,5 mL
- pipettes (dépend de la dilution, P10 à P200) et cônes à filtre associés
- gants en nitrile ou latex

PCR 1 : Qiagen Multiplex PCR Plus Kit (Reference : 206152)

- Qiagen Multiplex PCR Plus Kit
- eau ultra-pure pour PCR
- tubes eppendorf 1,5 mL
- amorces diluées à 10 μ M
- plaque PCR ou barrettes
- pipettes P10, P200, P1000 et cônes à filtre associés
- thermocycleur
- centrifugeuse de paillasse
- gants en nitrile ou latex

Gel d'agarose

- poudre agarose
- GelRed™ (ref 41003-T, Biotum)
- TAE 1X
- bleu de charge
- marqueur de taille électrophorèse
- électrophorèse
- pipettes P10 et cônes associés
- gants en nitrile ou latex

Purification des amplicons PCR – SPRI Beads (ref B23317, Beckman Coulter)

- SPRI Beads (ref B23317, Beckman Coulter)
- barrettes
- ethanol pure
- Tris 10 mM, TE ou eau ultrapure pour l'élution
- support magnétique pour barrettes
- vortex
- pipettes P10, P200 et cônes à filtre associés
- gants en nitrile ou latex

Préparation des échantillons

Si les échantillons ont une concentration $> 5 \text{ ng}/\mu\text{L}$, les diluer à $5 \text{ ng}/\mu\text{L}$ dans le l'eau ultra-pure pour PCR avant la PCR.

Si les échantillons ont une concentration $< 5 \text{ ng}/\mu\text{L}$, les diluer pour qu'il y ait 10 ng d'ADN dans $10 \mu\text{L}$ d'échantillon.

S'il n'est pas possible de faire cette dilution, $10 \mu\text{L}$ de l'échantillon pur seront mis pour la PCR.

PCR 1 : Qiagen Multiplex PCR Plus Kit (Reference : 206152)

Le mix PCR doit être fait en premier, dans en salle blanche (= où aucun ADN ne rentre). On ne peut pas aller en salle blanche si on a déjà manipulé de l'ADN dans la journée. Idéalement, il faut que le mix PCR soit la première chose que l'on fasse de la journée.

Penser à ajouter des blancs de PCR (eau ultrapure à la place de l'ADN) pour chaque mix.

Le couple d'amorces utilisé est issu de Leray et al, 2003 (<https://doi.org/10.1186/1742-9994-10-34>)

Il s'agit d'un couple d'amorces amplifiant un amplicon de 313 bp du COI : mICOInt/jgHCO2198.

Pour le metabarcoding, il est nécessaire d'ajouter une « queue » aux amorces utilisées en PCR1 pour la PCR de séquençage (PCR2). Cette queue peut varier selon les prestataires de séquençage et il est nécessaire d'échanger avec ces derniers avant de commander les amorces de PCR1. Ici, un exemple du couple d'amorce avec des queues adapté à la technologie Illumina :

Nom amorce	Amorces + ½ queue Illumina (sens 5' – 3')
mICOIntF_illadptr	CTTTCCTACACGACGCTCTCCGATCTGGWACWGGWTGAACWGTWTAYCCYC C
jgHCO2198_illadptr	GGAGTTCAGACGTGTGCTCTCCGATCTTAIACYTCIGGRTGICCRARAAYCA

1. Préparer le mix PCR

Selon ce tableau si [ADN] dilué $5 \text{ ng}/\mu\text{L}$:

réactifs PCR	35,0µl vol final		
	µl pour 1 éch	Concentration finale	µL pour n éch
Eau ultra-pure pour PCR	12		n*12
2X Multiplex PCR Master Mix	17,50	1,00X	n*17,50
10µM primer mICOIntF_illadp	1,75	0,5µM	n*1,75
10µM primer jgHCO2198_illadptr	1,75	0,5µM	n*1,75

2. Répartir le mix dans les tubes ou plaque PCR

3. Ajouter l'ADN ($5 \text{ ng}/\mu\text{L}$)	2
---	---

Selon ce tableau si [ADN] < 5 ng/μL :

35,0μl vol final			
réactifs PCR	μl pour 1 éch	Concentration finale	μL pour n éch
Eau ultra-pure pour PCR	4		n*4
2X Multiplex PCR Master Mix	17,50	1,00X	n*17,50
10μM primer mICOIntF_illadptt	1,75	0,5μM	n*1,75
10μM primer jgHC02198_illadptr	1,75	0,5μM	n*1,75
2. Répartir le mix dans les tubes PCR			
3. Ajouter l'ADN		10	

En général, on utilise un fichier Excel pour faire les calculs des volumes pour les mix automatiquement.

4. Lancer la PCR selon le protocole suivant :

Dénaturation initiale	95°C	5min	
Dénaturation	95°C	10sec	
Hybridation	Touch down 62-46°C (-1°C/cycle)		x 16 cycles
Elongation	72°C	1min	
Dénaturation	95°C	10sec	
Hybridation	46°C	30sec	x 25 cycles
Elongation	72°C	1min	
Elongation finale	68°C	10mn	

Gel d'agarose

Faire un gel d'agarose permet de visualiser les tailles des amplicons PCR. Il est possible de faire le gel sur seulement quelques échantillons représentant bien l'ensemble des échantillons considérés par le projet.

Plusieurs protocoles de gel d'agarose existent. Ici est présenté un protocole utilisant du GelRed™ (ref 41003-T, Biotum).

1. Faire un gel d'agarose à 3% avec GelRed incorporé (4 μL de GelRed 10 000 X dans 100 mL de gel).
2. Une fois le gel fait et durci, charger les puits avec :
 - 5 μL de PCR et 1 μL de bleu de charge
 - 1,5 μL de marqueur de taille
3. Lancer l'électrophorèse pour 35 minutes puis visualiser sous UVs.

Si les amplicons ADN sont bien de la taille attendu (= longueur de l'amplicon + longueur des amorces avec demi-queue Illumina), les échantillons sont prêts à être envoyé en séquençage. Certaines entreprises/platformes demandent à ce que les échantillons soient purifiés avant la PCR2, d'autres non. Nous avons donc mis ci-dessous le protocole de purification.

Purification des amplicons PCR – SPRI Beads (ref B23317, Beckman Coulter)

Ce protocole utilise des billes pour purifier (c'est-à-dire enlever les amorces et dimères d'amorces) de l'échantillon. Tout autre protocole de purification post-PCR peut être utilisé à la place.

1. a. (optionnel) Si vous avez fait des triplicats de PCR, regrouper les triplicats de PCR de chaque échantillon à purifier dans un seul puits.
Si la PCR a été faite en barrette, vous pouvez mettre deux puits en un. Si la PCR a été faite en plaque, vous devez mettre les triplicats en barrettes.
b. Remettre en suspension les billes SPRI. Ajouter un volume de SPRIselect pour obtenir un rapport de 0,8x avec l'échantillon.

2. Mélanger le volume total de réaction en pipettant 10 fois et incuber à température ambiante pendant 1 minute.

OU

Vortexer pendant 1 minute à une vitesse appropriée jusqu'à obtention d'un mélange homogène.

Une homogénéisation insuffisante de l'échantillon et de SPRIselect conduira à des résultats de sélection de taille incohérents. Veiller à bien mélanger.

3. Placer les barrettes sur un support magnétique approprié et laisser les billes SPRI décanter sur l'aimant.

Les temps de décantation varient ; un volume d'échantillon initial plus élevé, un rapport SPRIselect plus élevé ou des aimants plus faibles nécessiteront un temps de décantation plus long.

4. Retirer et jeter le surnageant clair.

Il faut veiller à ne pas aspirer plus qu'une infime quantité de billes au cours de cette étape, car l'ADN est lié aux billes. Une perte importante de billes entraînera une diminution du rendement.

5. En laissant la barrette dans le support magnétique, ajouter 180 µL d'éthanol à 85 % (préparer de l'éthanol frais pour des résultats optimaux) et incuber à température ambiante pendant 30 secondes. Retirer et jeter le surnageant d'éthanol.

Veillez à éliminer tout l'éthanol du fond du puits, car il s'agit d'un inhibiteur connu de PCR. Un temps de séchage de quelques minutes à température ambiante peut être effectué pour s'assurer que toutes les traces d'éthanol sont éliminées, mais il faut veiller à ne pas trop sécher l'anneau de billes (diminution du rendement).

Il faut veiller à ne pas aspirer plus qu'une infime quantité de billes au cours de cette étape, car l'ADN est lié aux billes. Une perte importante de billes entraînera une diminution du rendement.

6. a. Retirer la barrette du support magnétique et ajouter 25 µL de Tris 10 mM (pH 8).
b. Mélanger en pipettant 10 fois pour remettre les billes en suspension et incuber à température ambiante pendant 1 minute.
c. Placer la barrette sur le support magnétique et laisser les billes SPRI se décanter sur l'aimant.
7. Transférer l'éluat (échantillon purifié) dans une plaque ou barrette de PCR pour la PCR2.

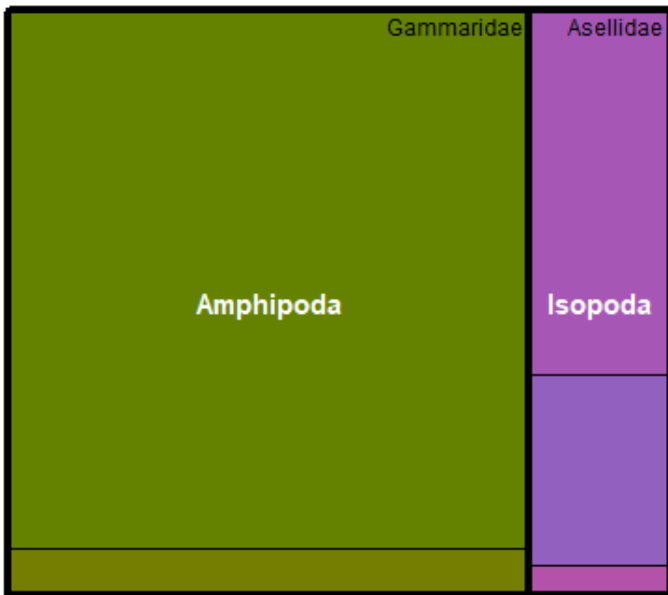
Purification des amplicons PCR – SPRI Beads (ref B23317, Beckman Coulter)**English version**

1. a. (optional) If you made PCR triplicates, pool PCR triplicates of each sample to purify in one well.
If the PCR was done in strips, you can put two wells in one. If the PCR was done in plate, you have to put the triplicates in strips.
b. Thoroughly resuspend the SPRI beads. Add a volume of SPRIselect to have a 0.8x ratio to the sample.
2. Mix the total reaction volume by pipetting 10 times and incubate at RT for 1 minute
OR
Vortex for 1 minute at an appropriate speed until homogenous
Insufficient mixing of sample and SPRIselect will lead to inconsistent size selection results. Make sure to mix well.
3. Place the reaction vessel on an appropriate magnetic stand and allow the SPRI beads to settle to the magnet.
Settle times will vary; a higher initial sample volume, higher SPRIselect ratio or weaker magnets will require a longer settle time.
4. Remove and discard the clear supernatant.
Care should be taken not to aspirate more than a trace amount of beads during this step, as the desired library is associated with the beads. Significant bead loss will result in reduced yield.
5. With the reaction vessel still on the magnet, add 180 μL of 85% ethanol (prepare fresh etoh for optimal results) and incubate at RT for 30 seconds. Remove and discard the ethanol supernatant.
Be sure to remove all of the ethanol from the bottom of the well as it is a known PCR inhibitor. A dry time of few minutes at RT can be done to ensure all traces of etoh are removed but take care not to over dry the bead ring (decrease elution efficiency).
Care should be taken not to aspirate more than a trace amount of beads during this step, as the desired library is associated with the beads. Significant bead loss will result in reduced yield.
6. a. Remove the reaction vessel from the magnet and add 25 μL of 10 mM Tris (pH 8).
b. Mix the total elution volume by pipetting 10 times to resuspend the beads and incubate at RT for 1 minute.
c. Place the reaction vessel on an appropriate magnetic stand and allow the SPRI beads to settle to the magnet.
7. Transfer the eluate (size selected sample) to an appropriate storage vessel (PCR plate or strips).

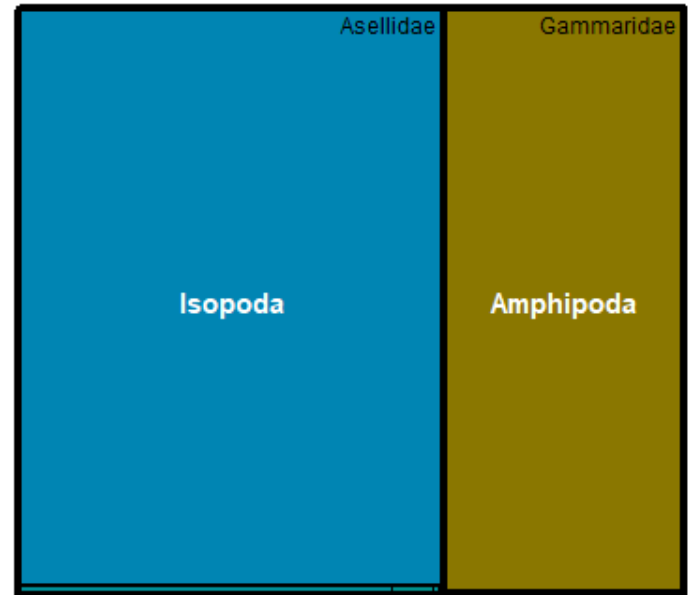
ANNEXE 4 - Treemaps pour les Mollusca et Malacostraca

Treemaps représentant la diversité de 3 niveaux hiérarchique au sein de la classe *Malacostraca* (haut) et du phylum *Mollusca* (bas) pour l'identification morphologique (gauche) et moléculaire (droite). Le premier niveau hiérarchique (bordures très épaisse, noms en blanc) correspond aux ordres pour *Malacostraca* et aux classes pour *Mollusca*, le deuxième (bordures épaisses, noms en noir) aux familles et le troisième aux taxa. Attention : les couleurs entre identifications morphologique et moléculaires ne correspondent pas forcément pour un groupe donné.

IDENTIFICATION MORPHOLOGIQUE

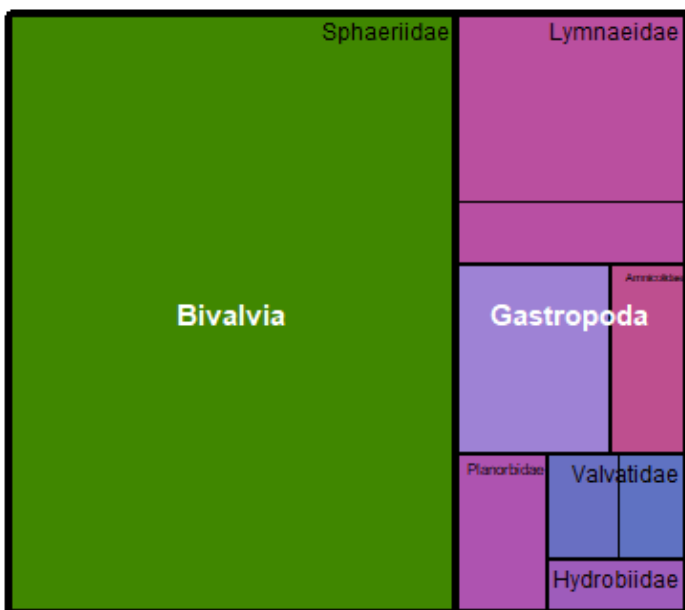


IDENTIFICATION MOLECULAIRE

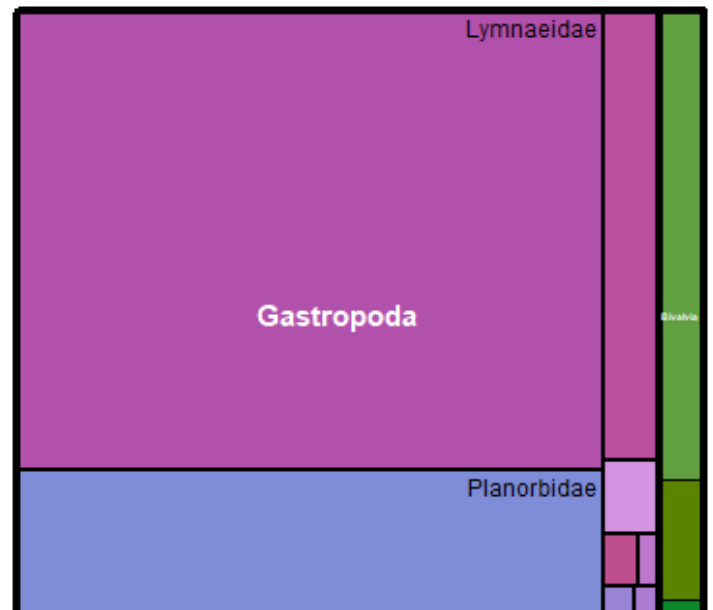


MALACOSTRACA

IDENTIFICATION MORPHOLOGIQUE



IDENTIFICATION MOLECULAIRE



MOLLUSCA

Z A B R

Zone Atelier Bassin du Rhône

ANNEXE 5 – Comparaison des listes taxonomiques morphologiques et moléculaires à l'échelle du prélèvement

Comparaison des listes taxonomiques morphologique (morpho, à gauche) et moléculaire (DNA, à droite) pour chaque prélèvement des différents sites. Pour chaque prélèvement, le nom de taxon avec ses rangs supérieurs abrégés se trouve en ordonnée. Pour chaque taxon, le rang taxonomique de l'assignement en identification morphologique (à gauche) et en identification moléculaire (à droite) est renseigné par des couleurs. Une croix correspond à une non-détection par l'identification moléculaire.

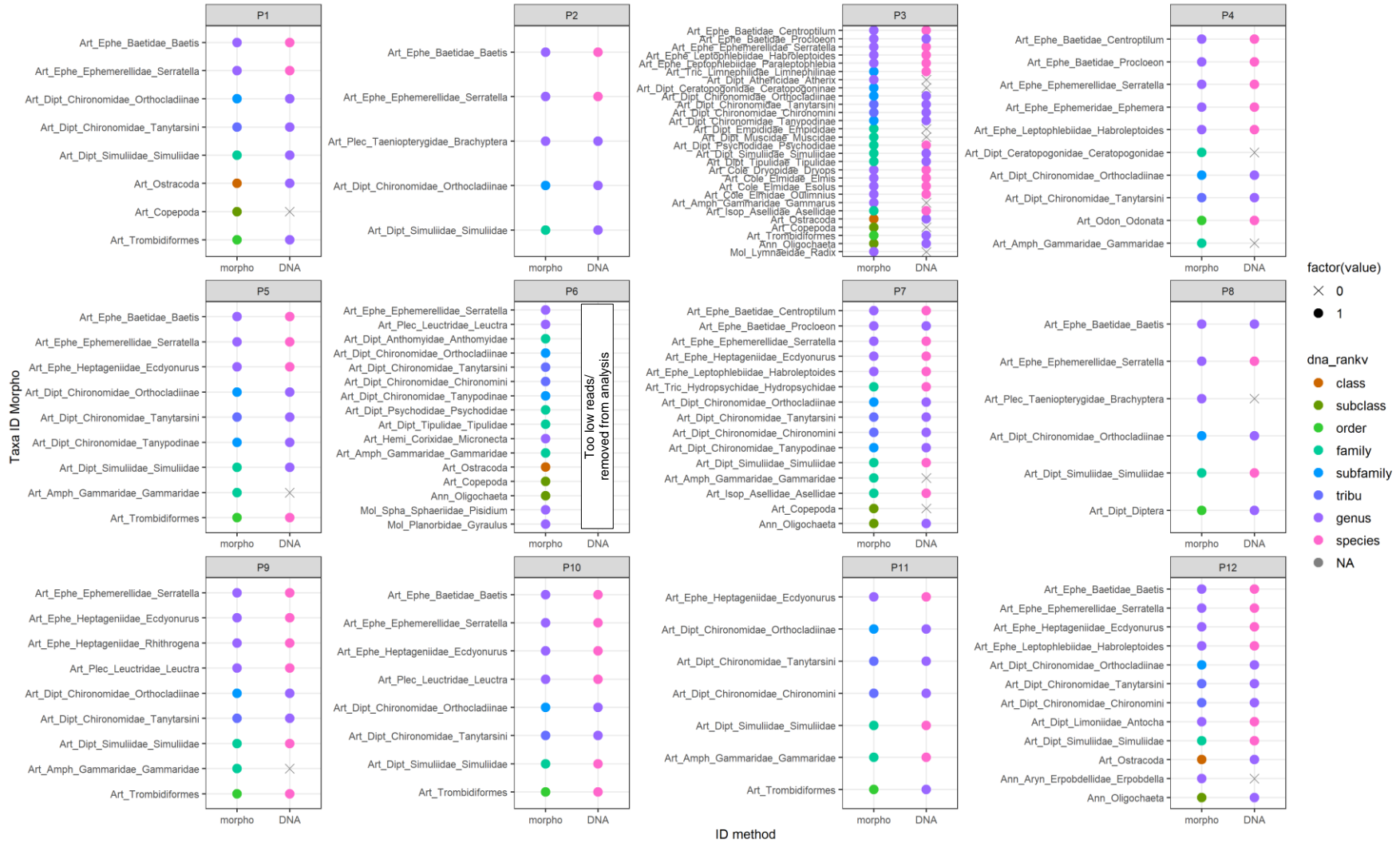
AMBERIEU-AVAL



Z A B R

Zone Atelier Bassin du Rhône

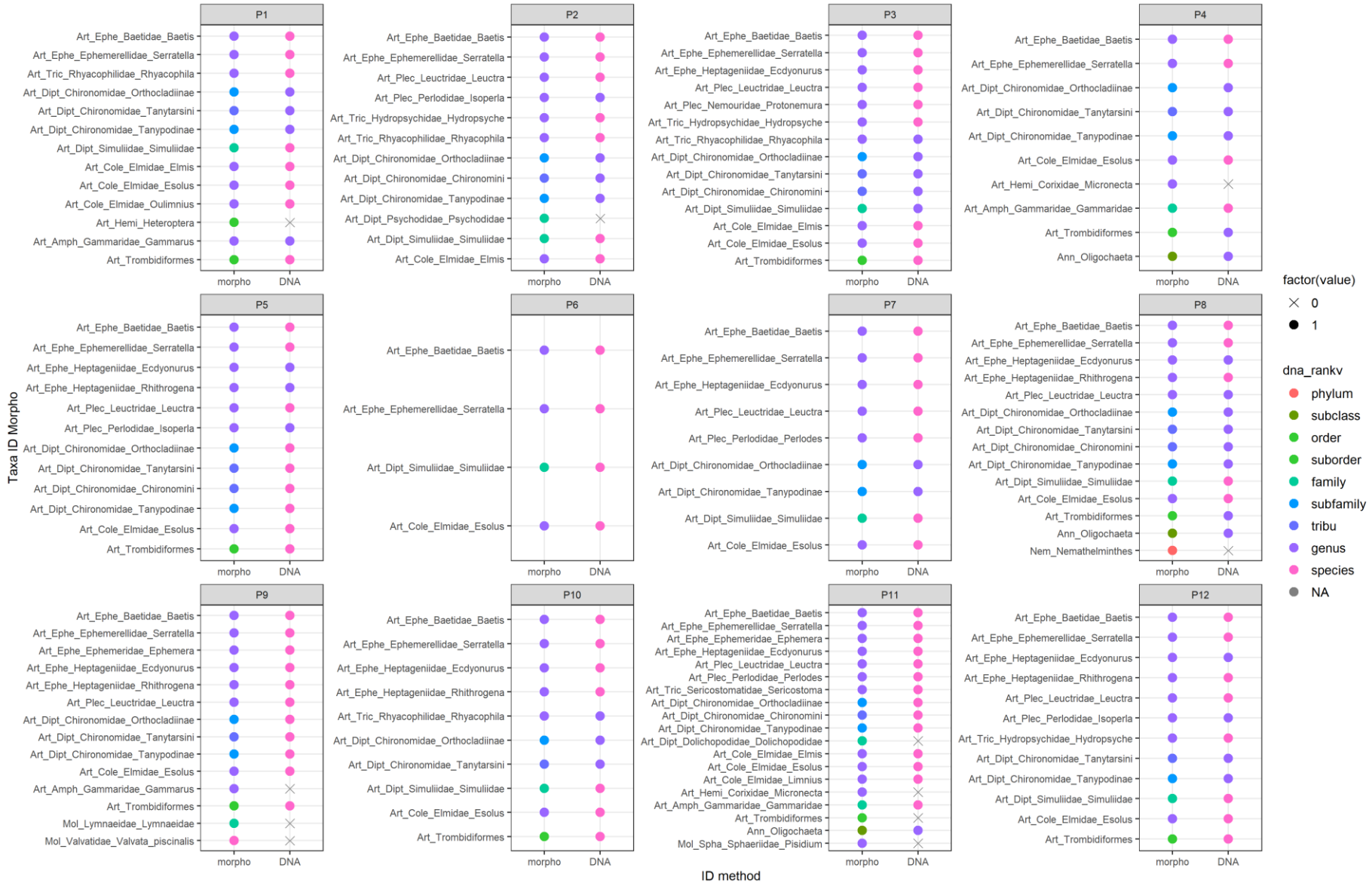
AMBERIEU-AMONT



Z A B R

Zone Atelier Bassin du Rhône

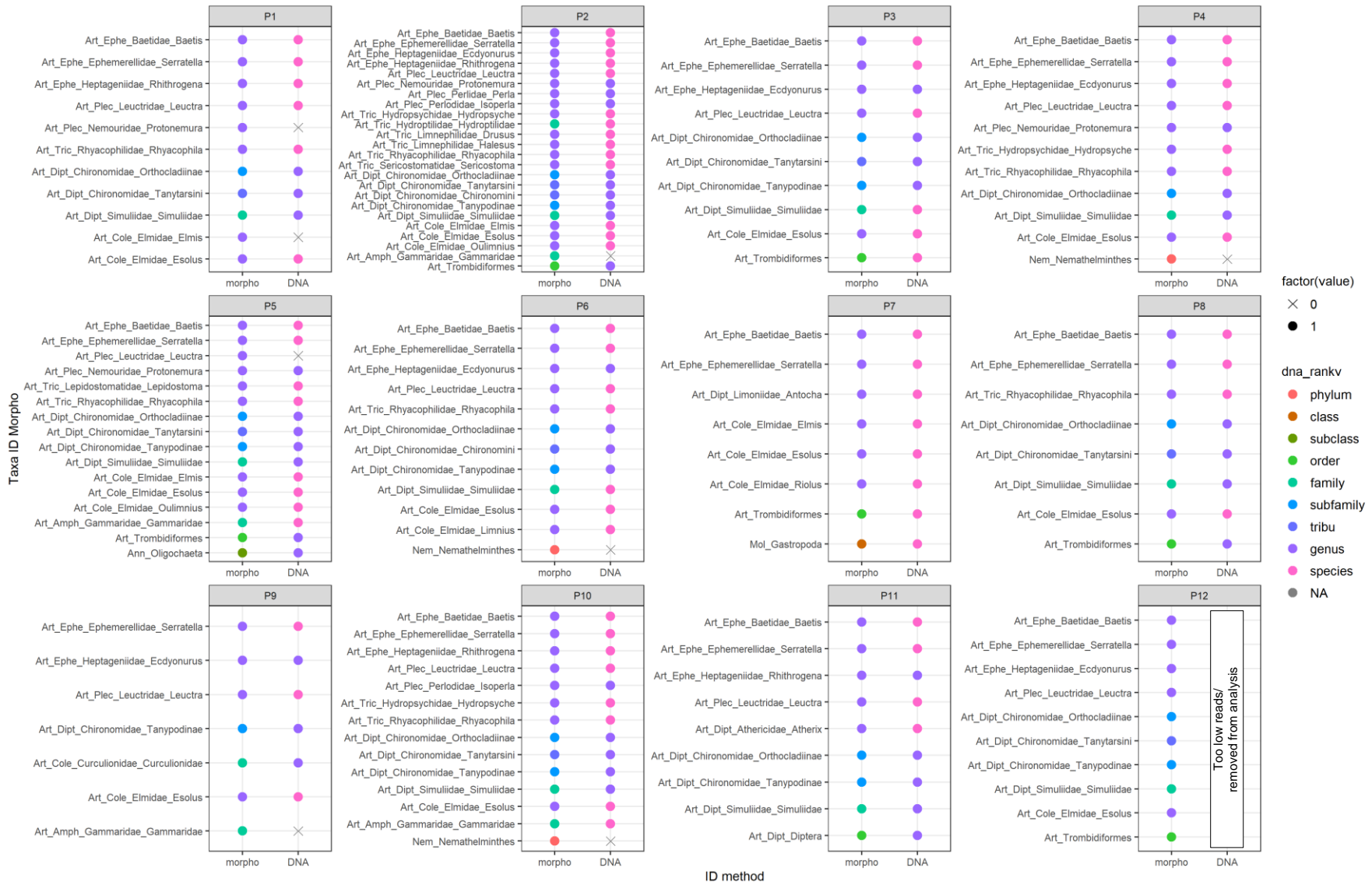
BETTANT - AVAL



Z A B R

Zone Atelier Bassin du Rhône

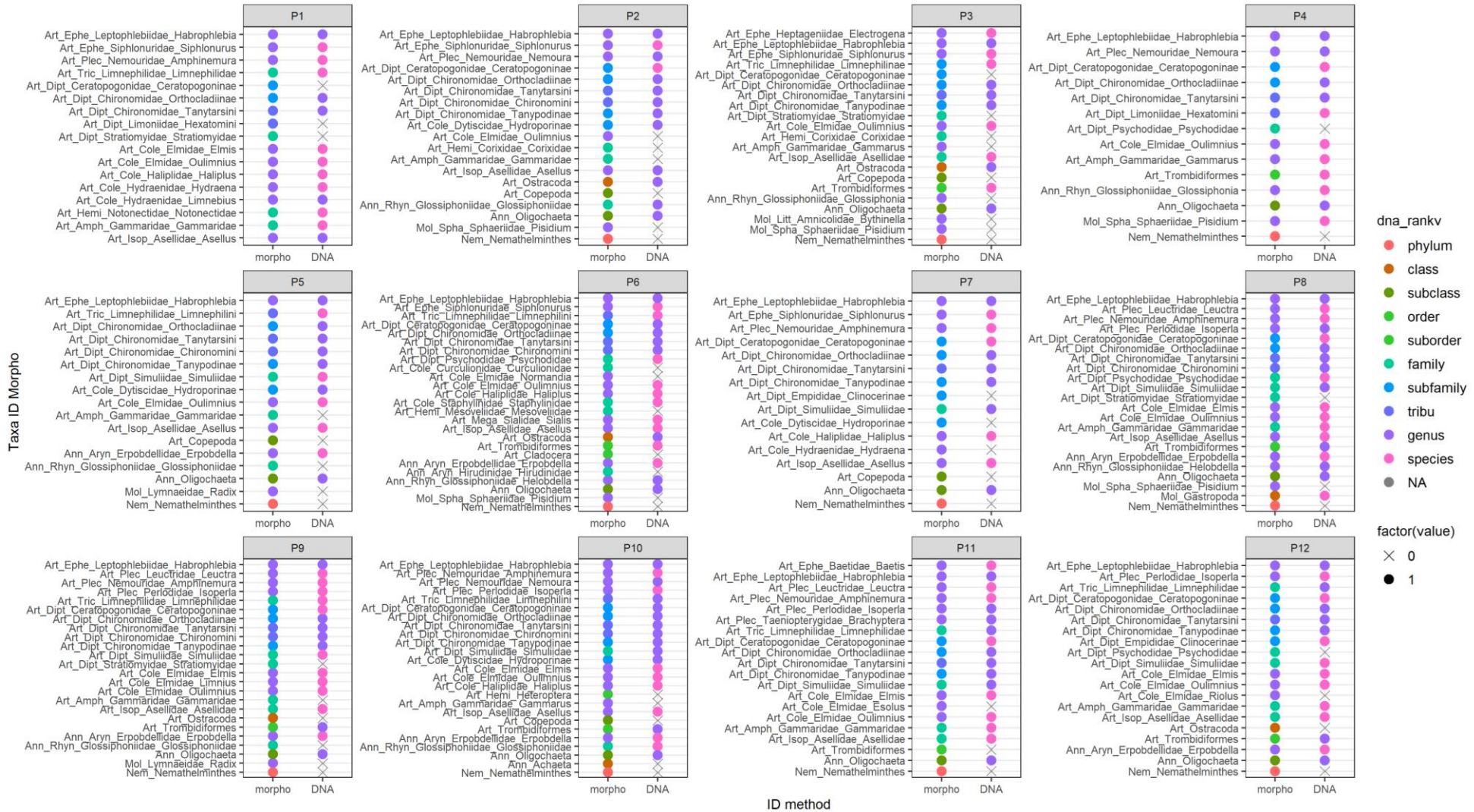
BETTANT - AMONT



Z A B R

Zone Atelier Bassin du Rhône

BRENOD - AVAL



Z A B R

Zone Atelier Bassin du Rhône

BRENOD - AMONT

Taxa ID Morpho

