

Rapport de stage – HT09

02/02/2026 – 17/07/2026

Développement d'outils moléculaires de suivi du méroplancton pour comprendre les dynamiques écologiques en contexte éolien offshore



Organisme d'accueil : Station biologique de Roscoff, UMR 7144
Tuteurs de stage : Fabrice Not & Thomas Pommier
Suiveur UTC : Sonia Rippa

Remerciements

Je tiens tout d'abord à remercier sincèrement mon maître de stage Fabrice Not pour m'avoir fait confiance, accueilli et supervisé. Merci d'avoir toujours su choisir les bons mots pour m'expliquer des notions compliquées avec tant de bienveillance (et de m'avoir permis d'assister à une magnifique pièce de théâtre).

Un grand merci aussi à Sarah Romic grâce à qui les manipulations n'ont (presque) plus aucun secret pour moi. Merci de m'avoir tout expliqué et montré avec patience et disponibilité. Je pense que je n'aurais pas autant apprécié la partie laboratoire si je n'avais pas été aussi bien encadrée !

Merci aussi à Thomas Pommier de m'avoir supervisé à distance tout au long de ce projet. Merci d'avoir su répondre à mes questions et aiguiller mes démarches expérimentales durant ces 6 mois.

Je remercie aussi mes co-encadrants, Yann Patry et Thierry Comtet, pour leurs précieuses expertises, ainsi que l'ensemble de l'équipe ECOMAP pour leur accueil chaleureux à la station biologique.

Enfin, merci aux stagiaires présent·e·s, et doctorant·e·s des différentes équipes, pour leur bonne humeur qui a tant participé à l'atmosphère conviviale de ce stage.

Finalement, je remercie évidemment Nicolas Salzmann pour son accompagnement et son investissement sans relâche tout au long de mon cursus Hutech.

Table des matières

Remerciements	1
Table des matières	2
Résumé technique	4
Glossaire	5
Liste des figures et des tableaux	6
Introduction	7
I - Présentation de l'organisme d'accueil	8
1. La Station Biologique de Roscoff.....	8
2. L'UMR 7144 et l'équipe ECOMAP	8
3. Setec énergie environnement	9
II - Contexte du sujet	10
1. La nécessité de décrire la biodiversité benthique	10
2. Le méroplancton.....	10
3. L'ADN environnemental et le metabarcoding pour l'étude de la biodiversité marine	11
A. L'ADN environnemental.....	11
B. Le metabarcoding.....	12
C. Les limites du metabarcoding pour l'étude du méroplancton.....	12
4. Problématique et objectifs du stage	13
III – Matériel et méthodes	14
1. Échantillons.....	14
2. Extraction de l'ADN des échantillons.....	14
A. Quantification et vérification de la qualité des ADN	15
3. Test des différentes Taq polymérase par PCR.....	15
4. Construction des bases de données	17
A. Bases de données de mitogénomes.....	17
B. Bases de données COI & 16S.....	17
C. Bases de données des espèces « non cibles ».....	18
5. Choix et conception des amorces <i>in Sillico</i>	18
6. Test des amorces par PCR	19
IV – Résultats	20
1. Extraction de l'ADN des échantillons.....	20
A. Comparaison des kits d'extraction.....	20
B. Comparaison des méthodes de conservation	22

2.	Test des différentes Taq par PCR.....	23
3.	Bases de données	24
4.	Choix et conception des amorces <i>in Sillico</i>	27
5.	Test des amorces par PCR	30
6.	Test de la spécificité des amorces	33
V – Discussion		35
1.	Améliorations et perspectives de l'étude.....	35
2.	Intérêt des marqueurs additionnels.....	36
3.	Pertinence du séquençage long-read sur de l'ADN environnemental	36
Conclusion		37
1.	Conclusion du stage.....	37
2.	Conclusion personnelle.....	37
Bibliographie		39
Annexes		43

Résumé technique

Cette étude s'inscrit dans une nécessité de décrire la biodiversité des organismes benthiques afin d'évaluer les impacts des pressions anthropiques croissantes sur les écosystèmes marins, notamment liées au développement de parcs éoliens offshore. Le méroplancton, regroupant les stades larvaires des espèces benthiques, constitue un indicateur précieux mais les méthodes morphologiques classiques ne sont pas adaptées pour son identification. Le metabarcoding d'ADN environnemental apparaît alors comme une alternative prometteuse, mais reste limité par deux contraintes majeures : la longueur réduite des fragments séquencés et le manque d'exhaustivité des bases de données de référence qui limitent la résolution taxinomique des études. L'objectif de mes travaux était donc de développer une approche metabarcoding « long-read », basée sur la technologie Oxford Nanopore Technologies (ONT), afin d'améliorer le suivi du méroplancton. Ce travail s'est articulé autour de deux axes : la construction de bases de données de références de mitogénomes complets et des marqueurs COI et 16S, et le développement d'un protocole expérimental permettant l'amplification de longs fragments d'ADN compatibles avec un séquençage « long-read ». La démarche a consisté à comparer plusieurs méthodes de conservation, d'extraction et d'amplification de l'ADN, puis à concevoir et tester *in silico* et par PCR des amorces ciblant la région comprise entre le COI et le 16S. Les résultats ont permis de sélectionner le tampon NucleoProtect et le kit MasterPure comme méthodes préférentielles de préservation et d'extraction, offrant la meilleure intégrité et quantité d'ADN, ainsi que la Taq Quantabio® pour l'amplification de longs fragments. Les tests PCR ont cependant révélé un manque de spécificité des amorces initialement sélectionnées, ce qui a conduit à la conception de nouvelles amorces plus spécifiques à chaque phylum ciblés.

This study addresses the need to characterize and monitor benthic biodiversity in order to assess the impacts of increasing human pressures on marine ecosystems, particularly those related to the development of offshore wind farms. Meroplankton, which includes the larval stages of benthic species, is a valuable indicator; however, traditional morphological methods are not well-suited for its identification. Environmental DNA metabarcoding thus emerges as a promising alternative but remains limited by two major constraints: the short length of the sequenced fragments and the lack of comprehensiveness in reference databases. The objective of this study was therefore to develop a long-read metabarcoding approach, based on Oxford Nanopore Technologies (ONT), to improve the monitoring of meroplankton. This work focused on two main areas: the construction of reference databases of complete mitogenomes and COI and 16S markers, and the development of an experimental protocol enabling the amplification of long DNA fragments compatible with long-read sequencing. The approach involved comparing several methods of preservation, extraction, and amplification, followed by the design and testing, both *in silico* and via PCR, of primers targeting the region between COI and 16S. The results led to the selection of NucleoProtect and the MasterPure kit as the preferred methods, offering the best DNA integrity and yield, as well as Quantabio® Taq for the amplification of long fragments. However, PCR tests revealed a lack of specificity in the initially selected primers, which led to the design of new primers that were more specific to each phylum.

Glossaire

ADNmt : AND mitochondrial

ANDr : AND ribosomique

BET : Bromure d'éthidium

BOLD : Barcode Of Life Data System

CMR : Cancérogène, mutagène et reprotoxique

CNRS : Centre National de la Recherche Scientifique

COI : Cytochrome oxydase I

COII : Cytochrome oxydase II

COIII : Cytochrome oxydase III

CYTB : Cytochrome B

DISEEM : Dispersal, Speciation, and Evolution of Marine Species

DMSO : Diméthylsulfoxyde

ECOMAP : *Ecology of marine plankton*

ECOPHY : *Ecophysiology of marine invertebrates*

EDYMAR : *Ecogeochemistry and dynamics of marine systems*

MAFFT : Multiple Alignment using Fast Fourier Transform

NCBI: National Center for Biotechnology Information

NGS : *Next-generation sequencing*

ONT : *Oxford Nanopore Technologies*

PB : Paire de base (de nucléotides)

PCR : *Polymerase Chain Reaction*

SBR : Station Biologique de Roscoff

SU : Sorbonne Université

TM : Température de fusion

UMR : Unité mixte de recherche

Liste des figures et des tableaux

Liste des figures :

- Photo de couverture : Source : <https://www.guepe.qc.ca/blogs/le-plancton-et-la-colonne-deau>
- Figure 1 : Organisation des 3 axes de Setec énergie environnement. Source : <https://www.setec.fr/>
- Figure 2 : Larve « zoé » de crabe. Source : <https://www.aquaportail.com/dictionnaire/definition/4245/zoe>
- Figure 3 : « Larve D » d'un bivalve appelée ainsi du fait de sa forme caractéristique. Source :
- Figure 4 : Schéma des différentes étapes du metabarcoding.
- Figure 5 & 6 : Photos de l'échantillonnage du 09/03/2026 à bord du Neomysis (bateau de la station biologique de Roscoff).
- Figure 7 : Capture d'écran de la page web de NCBI Nucleotide lors de la recherche des mitogénomes des polychètes.
- Figure 8 : Graphique du rapport A260/A280 en fonction de la taxonomie des échantillons analysés. Résultat Nanodrop. Comparaison du kit d'extraction MasterPure (en vert) et NucleoSpin Tissue (en bleu).
- Figure 9 : Graphique du rapport A260/A230 en fonction de la taxonomie des échantillons analysés. Résultat Nanodrop. Comparaison du kit d'extraction MasterPure (en vert) et NucleoSpin Tissue (en bleu).
- Figure 10 : Graphique de la quantité d'ADN (μg) en fonction des échantillons, résultats du kit Qubit DS HS 1X. Comparaison du kit d'extraction MasterPure (en vert) et NucleoSpin Tissue (en bleu).
- Figure 11 : Gel d'électrophorèse, comparaison du kit d'extraction MasterPure (en vert) et NucleoSpin Tissue (en bleu).
- Figure 12 : Gel d'électrophorèse, comparaison des méthodes d'extraction : éthanol 96% (en orange) et NucleoProtect (en violet).
- Figure 13 : Gel d'électrophorèse, comparaison de 3 Taq polymérase par PCR.
- Figure 14 : Graphique du nombre de séquences des bases de données de mitogénomes avant et après avoir été triées.
- Figure 15 : Graphique du nombre de familles apportées par les gènes extraits de mitogénomes (en jaune), ceux trouvées sur NCBI (en vert) et les manquants (en gris).
- Figure 16 : Graphique du nombre de familles apportées par les gènes extraits de mitogénomes (en jaune), ceux trouvées sur NCBI (en vert) et les manquants (en gris).
- Figure 17 : Graphique de la proportion de chacun des groupes présents dans la base de données des espèces non-cibles.
- Figure 18 : Amorces placées sur un mitogénome de polychète de l'espèce « laeonereis culveri », le trait bleu représente la région à amplifier.
- Figure 19 : Graphique des tailles des produits PCR théoriques des polychètes.
- Figure 20 : Graphique des tailles des produits PCR théoriques des échinodermes.
- Figure 21 : Graphique des tailles des produits PCR théoriques des crustacés.
- Figure 22 : Graphique des tailles des produits PCR théoriques des mollusques.
- Figure 23 : Exemple de la composition génétique d'un fragment d'ADN de polychète obtenu par PCR.
- Figure 24 : Exemple de la composition génétique d'un fragment d'ADN d'échinoderme obtenu par PCR.
- Figure 25 : Exemple de la composition génétique d'un fragment d'ADN de crustacé obtenu par PCR.
- Figure 26 : Exemple de la composition génétique d'un fragment d'ADN de mollusque obtenu par PCR.
- Figure 27 : Gel d'électrophorèse, résultat du test PCR des amorces.
- Figure 28 : Gel d'électrophorèse, résultat du test PCR sur les sets 2 et 5 avec la Taq Advantage.

Liste des tableaux :

- Tableau 1 : Amorces utilisées pour tester les 3 Taq polymérase.
- Tableau 2 : Liste des échantillons utilisés pour tester les amorces par PCR.
- Tableau 3 : Comparaison du prix et des longueurs d'amplification de 3 Taq polymérase.
- Tableau 4 : Amorces ciblant le COI et le 16S utilisées dans cette étude avec leur Tm, longueur, %GC et source.
- Tableau 5 : Couples d'amorces testés par PCR dans cette étude.
- Tableau 6 : Longueurs théoriques des fragments amplifiés par chaque couple d'amorce PCR.
- Tableau 7 : Nouvelles amorces plus spécifiques ciblant le COI des échinodermes et des mollusques.
- Tableau 8 : Pourcentage de compatibilité entre les amorces et les différents groupes taxonomiques (les 4 phylums et les espèces « non-cible »).

Introduction

Aujourd'hui, la mer est de plus en plus exploitée pour ses ressources et son énergie. Que ce soit par le transport maritime, la pêche ou encore l'introduction de nouvelles infrastructures comme les parcs éoliens offshore, le milieu marin est toujours plus touché par les activités humaines. Le suivi de la biodiversité marine s'avère donc nécessaire afin de comprendre et limiter les impacts que pourraient avoir de telles activités sur ce milieu. Développer de nouveaux outils d'inventaire et de suivi des communautés marines devient une nécessité de plus en plus pressante, tant les méthodes traditionnelles peinent à couvrir l'ampleur des changements en cours (Danovaro *et al.*, 2016).

L'étude du méroplancton occupe une place clé car il regroupe les stades larvaires de nombreux invertébrés benthiques, présents de manière saisonnière dans la colonne d'eau (plusieurs jours à plusieurs semaines), et constitue un indicateur sensible des changements environnementaux. Les modifications éventuelles des dynamiques écologiques impliquant les communautés benthiques telles que la colonisation des nouvelles structures artificielles ou encore le risque d'introduction d'espèces non indigènes, se traduisent toutes par des signaux observables dans la composition de ces communautés larvaires. Ainsi, le suivi du méroplancton constitue un outil précieux afin d'évaluer les effets des activités anthropiques comme les parcs éoliens offshore sur la biodiversité marine.

Plusieurs facteurs m'ont poussé à choisir ce stage : mon souhait de découvrir le monde de la recherche ainsi que mon attrait pour la biodiversité marine, mais surtout pour la préservation de nos écosystèmes qui sont chaque jour plus menacés par les activités anthropiques.

Une première partie de ce rapport est dédiée à une description de la Station Biologique de Roscoff et de l'équipe ECOMAP qui m'a accueillie. Puis, le contexte scientifique est abordé afin de présenter les enjeux et les objectifs de l'étude que j'ai menée. La troisième et la quatrième partie décrivent quant à elles les résultats obtenus ainsi que leur analyse.

I - Présentation de l'organisme d'accueil

1. La Station Biologique de Roscoff

Fondée en 1872, la Station Biologique de Roscoff (SBR), située dans le Finistère, est un centre internationalement reconnu pour la recherche et la formation en biologie et écologie marines. Elle est affiliée au Centre National de la Recherche Scientifique (CNRS) et à Sorbonne Université (SU). Ce laboratoire de pointe accueille près de 200 scientifiques ainsi qu'une trentaine de doctorants qui travaillent à étudier la biologie des organismes marins (microbes, macroalgues, microalgues, métazoaires), et le fonctionnement des écosystèmes océaniques pour mieux comprendre leur adaptation face au changement global.

La recherche à la Station Biologique de Roscoff est menée au sein de deux unités de recherche :

- L'UMR 8227 - Laboratoire de Biologie Intégrative des Modèles Marins
- L'UMR 7144 - Laboratoire "Adaptation et Diversité en milieu Marin"

J'ai effectué ce stage dans cette dernière unité.

Enfin, une Fédération de Recherche (FR2424) regroupe les services communs en support à la recherche (service accueil scientifique, logistique, finances, etc.) et en soutien à la recherche (service mer et plongée, observation, informatique, plateformes technologiques, etc.).

2. L'UMR 7144 et l'équipe ECOMAP

Les travaux au sein de l'UMR 7144 portent sur l'écologie descriptive, fonctionnelle et évolutive des organismes et écosystèmes marins et s'appuient sur des campagnes océanographiques pour explorer des zones éloignées, du milieu pélagique jusqu'aux grands fonds. Ce laboratoire se compose de 4 équipes de recherche qui orientent leurs questionnements vers des organismes répartis sur l'ensemble de l'arbre du vivant : des procaryotes photosynthétiques à tous les eucaryotes unicellulaires du plancton, en passant par les macro-algues et les divers invertébrés.

- L'équipe ECOPHY : Ecophysiology of marine invertebrates
- L'équipe DISEEM : Dispersal, Speciation, and Evolution of Marine Species
- L'équipe EDYMAR : Ecogeochemistry and dynamics of marine systems
- L'équipe ECOMAP : Ecology of marine plankton

J'ai effectué ce stage au sein de l'équipe ECOMAP qui représente la plus grande équipe de l'UMR 7144. Cette équipe, comme son nom l'indique, étudie la structure et la dynamique des communautés planctoniques. Les membres d'ECOMAP possèdent une expertise sur toutes les classes de tailles du plancton, y compris les virus, les procaryotes hétérotrophes et photosynthétiques, les protistes et autres microeucaryotes. De plus, ces organismes sont étudiés à tous les niveaux d'organisation : des molécules, aux cellules, aux communautés, jusqu'aux écosystèmes et à l'océan global.¹

¹ Site internet de la Station biologique de Roscoff : <https://www.sb-roscoff.fr/fr> (consulté le 03 juin 2026).

3. Setec énergie environnement

Ce stage est également réalisé en collaboration avec Setec énergie environnement, une filiale du groupe Setec. Setec est une société d'ingénierie française fondée en 1957, regroupant plus de 4000 collaborateurs en France et à l'étranger, dont le but est d'accompagner les acteurs publics et privés face aux enjeux de la transition écologique. Elle offre des prestations depuis les études amont et la conception jusqu'au suivi des réalisations, elle intervient également pour des missions d'expertise et de conseil. Setec énergie environnement est la filiale spécialisée dans l'énergie et l'environnement du groupe Setec. Setec énergie environnement, et sa marque Setec *in vivo*, est une société répartie sur sept implantations en France : Paris (siège social), Bordeaux, La Forêt Fouesnant, Lille, Lyon, Nantes et Toulon. Elle associe une approche systémique entre énergie et environnement pour proposer aux porteurs de projet privés et publics de faire face aux défis écologiques et énergétiques, à travers trois grands axes cités dans la figure 1.



Figure 1 : Organisation des 3 axes de Setec énergie environnement

Ce stage est plus étroitement en lien avec la marque Setec *in vivo* qui réunit des experts dédiés à l'environnement marin, spécialisés en biodiversité et en monitoring environnemental. Ils réalisent des études en particulier sur des projets d'énergies marines renouvelables, de pose de câbles sous-marins, d'aménagements portuaires et d'infrastructures maritimes.²

² Site internet de Setec : <https://www.setec.fr/> (consulté le 03 juin 2026).

II - Contexte du sujet

1. La nécessité de décrire la biodiversité benthique

La description de la biodiversité benthique constitue aujourd'hui un enjeu scientifique d'une grande importance. En effet, les fonds marins subissent toujours plus de pressions anthropiques dues à l'implantation de nouvelles infrastructures telles que les câbles sous-marins, les plateformes pétrolières et gazières, les exploitations minières ou encore les parcs éoliens offshore. Ce dernier exemple illustre bien cette urgence et de nombreux chercheurs expriment leurs inquiétudes quant aux impacts potentiels de ces parcs sur l'environnement marin, qu'ils soient positifs ou négatifs (Watson *et al.*, 2025). Les espèces benthiques sont les premières touchées par ces constructions, que ce soit pendant la phase de construction, d'exploitation ou de démantèlement. En effet, la pose des fondations, l'ancrage des câbles, ou encore le bruit émis constituent de grandes perturbations pour les habitants des fonds marins. Des effets indirects participent aussi à une recomposition totale des communautés benthiques, à savoir un changement du réseau trophique, un effet récif, ou encore l'introduction d'espèces non indigènes (Dannheim *et al.*, 2020). Ces exemples montrent que décrire la biodiversité benthique est aujourd'hui une condition indispensable à une gestion durable des activités en mer.

La durée de ce stage ne permettant pas de mener une étude sur toutes les espèces benthiques, nous nous intéresserons ici aux 4 phylums suivants : les mollusques, les crustacés, les échinodermes et les polychètes. Ces groupes d'invertébrés ont été choisis car ils dominent la majeure partie de la biomasse et de la diversité des fonds marins (Saeedi *et al.*, 2022).

Le méroplancton, regroupant les stades larvaires des espèces benthiques, semble être un bon indicateur de ces changements écologiques car sa composition est en lien direct avec celle de la biodiversité benthique.

2. Le méroplancton

Le plancton (du grec *planktos*, signifiant « errant ») désigne tous les organismes marins flottant passivement dans les milieux aquatiques et ne se déplaçant qu'au gré des courants, on y retrouve des procaryotes et des eucaryotes. On peut diviser le plancton en deux groupes : le phytoplancton et le zooplancton. Le phytoplancton englobe les organismes photo-autotrophes unicellulaires, comme les micro-algues. Celui-ci est à la base de tous les réseaux trophiques aquatiques (Reynolds, 2006). Le zooplancton regroupe des organismes unicellulaires ou pluricellulaires hétérotrophes, qui consomment de la matière organique déjà constituée. Le zooplancton peut être divisé en deux groupes selon que les organismes font partie intégrante du plancton ou qu'ils n'y séjournent que temporairement. L'holoplancton comprend les organismes qui passent toute leur vie au sein du plancton, comme les copépodes ou les cténaires. Le méroplancton désigne les organismes qui ne passent qu'une partie de leur cycle de vie (souvent au stade larvaire) au sein du plancton. On y retrouve par exemple les larves d'espèces benthiques sessiles comme les éponges et les anémones, mais aussi des espèces mobiles vivant au fond de l'eau telles que les mollusques et les crustacés (Dipper, 2022). Ces larves passent des durées variables dans le plancton et présentent de très grandes différences avec l'adulte qu'elles deviendront, comme on peut le voir sur les figures 2 & 3. Le méroplancton comprend également les œufs et larves d'espèces nectoniques (qui sont en capacité de nager) telles que les poissons ou les céphalopodes. Étant

donné que nous nous concentrons sur les espèces benthiques dans cette étude, ce ne sont pas celles qui vont nous intéresser.

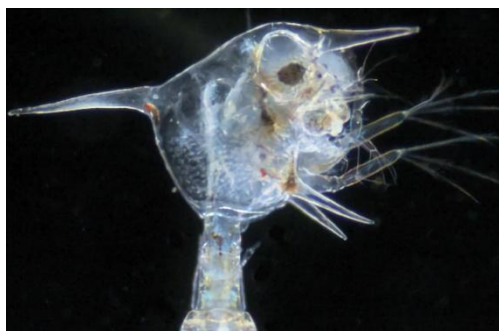


Figure 2 : Larve « zoé » de crabe

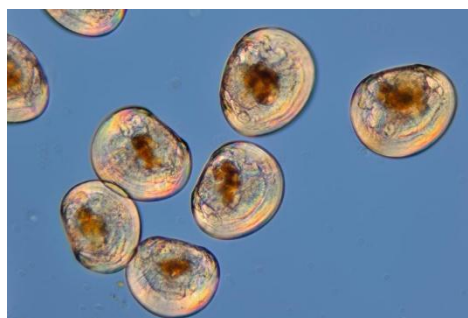


Figure 3 : « Larve D » d'un bivalve appelée ainsi du fait de sa forme caractéristique

Bien qu'il s'agisse d'un stade de vie important pour de nombreuses espèces benthiques clés, le méroplancton a longtemps été négligé. En effet, les études sur le méroplancton sont entravées par sa présence sporadique dans la colonne d'eau, ainsi que par l'absence de caractéristiques morphologiques distinctives permettant d'identifier les spécimens au-delà du niveau du phylum ou de la classe (Ershova *et al.*, 2019). Jusqu'à récemment, la seule technique permettant une identification des espèces du méroplancton était l'observation microscopique. Malheureusement, cette méthode n'est efficace que pour les stades larvaires qui se distinguent sur le plan morphologique. Bon nombre des stades larvaires les plus précoces, par exemple chez les bivalves (figure 3), présentent une apparence si similaire au début de leur développement qu'il est impossible de les identifier précisément, au niveau de l'espèce (Graland *et al.*, 2002). De plus, différencier les larves en fonction de leur forme, leur dimension et leur couleur nécessite des connaissances très spécialisées et l'intervention de nombreux experts. (Hendriks *et al.*, 2005). Face aux limites des études morphologiques, les méthodes moléculaires paraissent très prometteuses pour l'identification des communautés larvaires, quel que soit leur stade de développement. Le metabarcoding d'ADN environnemental par exemple, révèle une diversité benthique plus riche que ce que suggèrait les méthodes de prélèvement classiques (Martino *et al.*, 2025).

3. L'ADN environnemental et le metabarcoding pour l'étude de la biodiversité marine

A. L'ADN environnemental

L'ADN environnemental (ou eDNA) correspond à l'ADN extrait d'un échantillon environnemental dans le but d'obtenir les informations taxonomiques ou fonctionnelles les plus complètes possibles concernant l'écosystème étudié. Il peut s'agir de sol, de sédiments, d'eau, mais aussi de matériel issu de la filtration de ces milieux. L'ADN environnemental total contient à la fois de l'ADN intracellulaire et extracellulaire. L'ADN intracellulaire provient de cellules vivantes ou d'organismes multicellulaires présents dans l'échantillon tandis que l'ADN extracellulaire résulte des déchets et des fragments de cellules mortes libérés par ces organismes. Celui-ci peut être dégradé par des processus physiques, chimiques ou biologiques (Taberlet *et al.*, 2012).

B. Le metabarcoding

La biodiversité marine d'un échantillon environnemental peut être étudiée grâce à une technique de biologie moléculaire appelée « metabarcoding » (figure 4). Le metabarcoding permet la caractérisation génétique des individus présents dans un échantillon à partir d'un gène ou marqueur (barcode), par exemple une courte région d'ADN riche en information taxonomique.

Chez les eucaryotes, les marqueurs génétiques standards sont les marqueurs de l'ADN mitochondrial (ADNmt) et de l'ADN ribosomique (ADNr). Dans cette étude, il a été décidé de cibler des marqueurs de l'ADN mitochondrial qui permettent souvent d'identifier les taxons au niveau de l'espèce plus efficacement que les marqueurs ribosomiques (Atienza *et al.*, 2020) qui sont moins résolutifs. De plus, l'ADNmt bénéficie généralement d'un nombre élevé de copies au sein des cellules, ce qui améliore la sensibilité de détection dans les études d'ADN environnemental (Girard *et al.*, 2022).

Une fois le marqueur d'intérêt choisi, cette région est amplifiée par PCR puis séquencée. Il existe plusieurs types de séquençage différents pour faire du metabarcoding. Le séquençage à haut-débit ou « next-generation sequencing » (NGS) permet de produire rapidement des millions de séquences à partir d'un échantillon en un seul passage « run » à un coût par échantillon relativement faible. On retrouve par exemple la méthode Illumina, introduite en 2006, qui entraîne une explosion des études de diversité mais qui ne permet cependant de séquencer que des fragments courts (150 à 400 pb maximum). Face à ces enjeux, apparaissent les méthodes de séquençage de « troisième génération » qui permettent cette fois de faire du séquençage « long-read ». On retrouve par exemple la technologie utilisée dans cette étude nommée Oxford Nanopore Technologies (ONT). Cette dernière fonctionne grâce à un réseau de minuscules orifices, des nanopores, qui émettent un signal électrique à chaque fois qu'une molécule les traverse. Ces signaux sont ensuite décodés afin d'identifier quelle base vient de traverser le nanopore et ainsi déterminer la séquence d'ADN correspondante (Tyagy *et al.*, 2020).

Les résultats de ce séquençage sont ensuite comparés à une base de données de référence spécifique de ce barcode afin de faire le lien entre ces séquences et leur taxonomie. Cette base de données constitue un élément clé dans le processus de metabarcoding car il a été démontré que la fiabilité et la précision du classement taxonomique des données de séquençage issue d'échantillons environnementaux dépendent de manière déterminante de la qualité et de l'exhaustivité des bases de données de référence (Jeunen *et al.*, 2022). Une fois l'affiliation taxonomique faite, il ne reste plus qu'à étudier la composition des communautés qui nous intéressent (dans notre cas, les communautés benthiques méroplanctoniques).

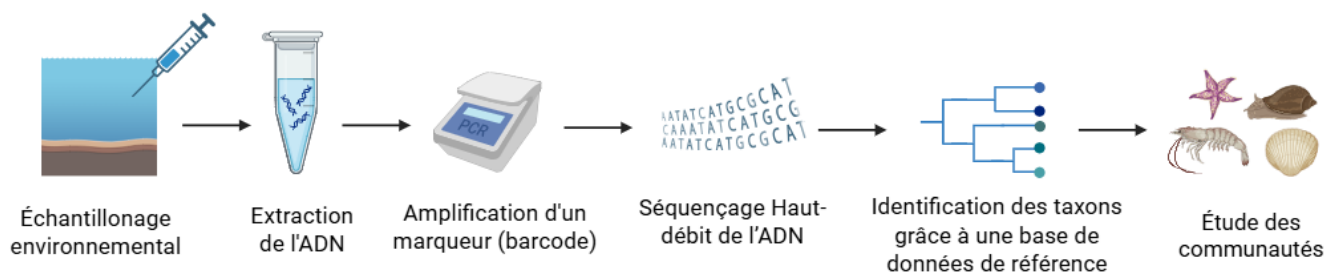


Figure 4 : Schéma des différentes étapes du metabarcoding

C. Les limites du metabarcoding pour l'étude du méroplancton

Si l'approche de metabarcoding est de plus en plus plébiscitée pour décrire les communautés présentes dans l'environnement, cette méthode est encore trop peu utilisée pour le méroplancton. Cette

approche moléculaire fait aujourd'hui face à 2 problèmes majeurs qui nuisent à son efficacité. En effet, la longueur réduite des fragments d'intérêts séquencés, ainsi que le manque d'exhaustivité des bases de données de références ne permettent souvent pas une affiliation taxonomique précise de ces communautés.

4. Problématique et objectifs du stage

Le méroplancton apparaît comme un indicateur pertinent pour la description des communautés benthiques. Les approches morphologiques n'étant pas efficaces pour discriminer les espèces qui le constituent, des méthodes moléculaires semblent être prometteuses. Cependant, ces approches viennent aussi avec leur lot de limites concernant le méroplancton. Pour faire face à ces enjeux, une approche de séquençage long-read semble comme prometteuse car elle permettrait d'obtenir des fragments longs d'ADN de génomes mitochondriaux. Cette approche impliquerait ainsi l'obtention simultanée de plusieurs marqueurs, permettant d'augmenter la résolution des assignations taxonomiques. Ces fragments incluraient idéalement des gènes permettant de distinguer les familles, mais aussi d'autres gènes plus spécifiques qui permettraient de distinguer les espèces voire les sous-espèces. L'objectif principal de ce stage est donc d'optimiser les approches de metabarcoding long-read à une partie encore peu explorée des communautés benthiques marines : celle du méroplancton. Cet objectif principal se découpe en deux sous-objectifs, répondant directement aux deux limites du metabarcoding appliqué au méroplancton.

- Le premier sous-objectif de cette étude va être de pallier l'exhaustivité des bases de données de référence en construisant des bases de données des génomes mitochondriaux des espèces benthiques. Cette base de données ne sera pas seulement utile pour l'affiliation taxonomique, mais permettra aussi le choix et la conception des amorces.
- Le deuxième sous-objectif de cette étude va être de pallier la longueur réduite des fragments d'intérêts séquencés en développant un protocole de séquençage long-read, grâce à la technologie « Oxford Nanopore Technologies » (ONT). À ce stade de l'étude, nous nous intéressons surtout aux étapes précédant celle du séquençage, à savoir : le processus d'extraction de l'ADN (quelle méthode de conservation, quel kit d'extraction,) et d'amplification de l'ADN (quelle Taq polymérase, quelles amorces) afin de les rendre compatibles avec ce séquençage long-read.

III – Matériel et méthodes

1. Échantillons

Pour cette étude, les prélèvements ont été réalisés à partir d'espèces correspondant aux quatre phylums qui nous intéressent lors de deux sorties en mer en baie de Morlaix, sur le site « Pierre Noire » (figures 5 & 6). Selon leur taille, des individus entiers ou des fragments de tissus ont été disséqués, puis deux réplicats ont été fixés respectivement dans 1 mL d'Ethanol absolu ou dans une solution de stabilisation Nucleoprotect (Macherey-Nagel®) afin de préserver les acides nucléiques. Tous les échantillons ont ensuite été stockés à -20°C. Afin de tester la spécificité des amorces, des fragments de tissus de poissons ont également été prélevés selon le même protocole. La liste des échantillons et les conditions de prélèvements sont présentées en annexe 1.



Figure 5 & 6 : Photos de l'échantillonnage du 09/03/2026 à bord du Neomysis (bateau de la station biologique de Roscoff)

2. Extraction de l'ADN des échantillons

Deux méthodes d'extraction d'ADN ont été testées sur les échantillons : le kit MasterPure (Lucigen™) et le kit NucleoSpin Tissue (Macherey-Nagel®).

Le kit MasterPure (Lucigen™), généralement utilisé pour les extractions d'ADN de matériel « trace » type cellule unique (Romac *et al.*, 2023) permet d'obtenir de longs fragments d'ADN. Les échantillons sont mis en suspension dans 300 µL d'un tampon de lyse en présence de 2 µL de protéinase K (20 mg/mL) et incubés à 65°C pendant 2 heures à 600 rpm. Le tampon de lyse désintègre les membranes permettant la libération de l'ADN contenu dans les cellules et la protéinase K élimine les protéines présentes. Le lysat obtenu est ensuite centrifugé en présence de 150 µL de tampon de précipitation MPC issu du kit afin de séparer les acides nucléiques des autres débris organiques, le surnageant est conservé. L'ADN est ensuite purifié successivement par précipitation v/v avec de l'isopropanol 100% et de l'éthanol 70%. Après élimination du surnageant, les pellets d'ADN obtenus sont séchés puis élués dans 30 µL de tampon d'éluion Tris-EDTA 1x et stockés à -20°C.

Le kit NucleoSpin Tissue (Macherey-Nagel®) est déjà couramment utilisé au laboratoire pour les extractions d'ADN de métazoaires. Les échantillons sont mis en suspension dans 360 µL d'un tampon

de lyse en présence de 25 µL de protéinase K (20 mg/mL) et incubés à 65°C pendant 2 heures. Une deuxième lyse est effectuée à l'aide d'une incubation à 70°C pendant 10 minutes. Le lysat obtenu est centrifugé afin d'éliminer les débris cellulaires. L'ADN est ensuite collecté sur colonne de silice avant d'être lavé. Une fois lavée, la membrane en silice est séchée puis l'ADN est élué dans 100 µL de tampon d'élution TE Tris-EDTA 1x et stocké à -20°C.

A. Quantification et vérification de la qualité des ADN

Les concentrations en ADN des extraits ont ensuite été mesurées grâce au kit Qubit ds HS 1X (Invitrogen™). Cette méthode utilise un colorant sélectif cible qui émet un signal fluorescent lorsqu'il est lié à l'ADN double brin. La quantité de fluorescence émise est directement proportionnelle à la concentration en ADN présente dans l'échantillon. La fluorescence émise est détectée à l'aide du fluorimètre Qubit™ 4 (Invitrogen™).

La qualité de l'extraction d'ADN a été mesurée grâce au spectrofluorimètre à microvolume Nanodrop (ThermoFisher®). Cette méthode informe sur la qualité des échantillons en mesurant :

- Le ratio A260/A280 (où A280 est la longueur d'onde d'absorbance des protéines et A260 est la longueur d'onde d'absorbance de l'ADN). La valeur du ratio A260/A280 doit être comprise entre 1,7 et 2 pour être considérée comme correcte. Cela renseigne ainsi sur d'éventuelles contaminations par des résidus protéiques. Si le ratio est compris entre 2 et 2,2, il témoigne de la présence d'ARN dans l'échantillon.
- Le ratio A260/A230 (où A230 est la longueur d'onde d'absorbance des résidus chimiques issus des procédés d'extraction et A260 est la longueur d'onde d'absorbance de l'ADN). La valeur du ratio A260/A230 doit être située au-dessus de 1 et tendre vers 2 pour être considérée comme correcte. Un rapport A260/A280 dépassant 2 n'est pas le signe d'un problème. Un rapport A260/A280 ne dépassant pas 1 témoigne généralement de la présence d'éthanol.

Enfin, une électrophorèse sur gel d'agarose a été menée afin de savoir le niveau de fragmentation des extraits d'ADN. Pour cela, un gel d'agarose à 1.2 % est préparé dans du TAE 0.5 % (Tris-Acétate EDTA) avant d'y ajouter du BET (Bromure d'éthidium). Ce dernier est un intercalant de l'ADN cancérigène, mutagène et reprotoxique (CMR) qui permet de visualiser les bandes d'ADN par les UV. Lors du dépôt, 5 µL d'ADN sont mélangés à 1 µL de bleu 6x, un tampon qui contient un agent de densité permettant de charger l'ADN et de le faire descendre au fond des puits. Il contient aussi un colorant ce qui permet de surveiller la migration et de voir l'échantillon lors du dépôt. En parallèle, le marqueur de taille Smartladder (Eurogentec™) est déposé, il permet de connaître la taille des bandes). Pour cette étude, la migration se fait à 140 V pendant 35 minutes dans du tampon TAE 0.5 %. À la fin de la migration, l'observation des résultats se fait à partir d'ImageQuant LAS 4000 (GE Healthcare).

3. Test des différentes Taq polymérase par PCR

La Taq polymérase est une variété d'ADN polymérase thermostable largement utilisée en biologie moléculaire pour réaliser des *Polymerase Chain Reaction* (PCR). Afin de pouvoir tester les amorces par PCR, il a d'abord fallu choisir une enzyme capable d'amplifier de longs fragments d'ADN. Trois différentes Taq ont donc été comparées par PCR : la Taq Phusion (Thermo Fisher®), la Taq repliQa HiFi ToughMix (Quantabio®), et enfin, la Taq Owi AptaFi Hot Start (Owibio®). Ces 3 enzymes ont été testées

avec un couple d'amorces validé (tableau 1) tiré de Chwalińska *et al.*, 2025, ciblant une portion d'environ 4000 pb entre les gènes de l'ARN ribosomique 18S et 28S de l'ensemble des eucaryotes. Les échantillons utilisés pour cette PCR sont des échantillons d'ADN environnemental de Méditerranée. Des témoins négatifs et positifs ont également été testés pour vérifier le bon déroulement de la réaction d'amplification et l'absence de contamination.

Tableau 1 : Amorces utilisées pour tester les 3 Taq polymérase

Sens	Nom	Séquence 5'-3'	Longueur	Tm °C
Forward	SA_Min_V14	TTT-CTG-TTG-GTG-CTG-ATA-TTG-CAA-CCT-GGT-TGA-TCC-TGC-CAG-T	43	63,3
Reverse	D2C-R_Min_V14	ACT-TGC-CTG-TCG-CTC-TAT-CTT-CCC-TTG-GTC-CGT-GTT-TCA-AGA	42	64,3

Des protocoles spécifiques et des programmes PCR ont été développés pour chaque enzyme :

- Pour la Taq Phusion, la réaction PCR se fait dans un volume réactionnel de 25 µL contenant : le mix Phusion 1X, 0.35 µM de chaque amorce et 3% de Diméthylsulfoxyde (DMSO). Après avoir rajouté 10 ng d'ADN, le mélange réactionnel est placé dans un thermocycleur, appareil qui applique un programme d'amplification précis en faisant varier les cycles de température au cours du temps. Le programme PCR commence par une dénaturation initiale de l'ADN à 98°C pendant 30s, il se poursuit par 30 cycles de dénaturation à 98°C pendant 10s, d'hybridation des amorces pendant 30s 55°C et d'élongation à 72°C pendant 30s, le programme se termine par un cycle de dénaturation finale à 72°C pendant 10 minutes.
- Pour la Taq Quantabio®, la réaction PCR se fait dans un volume réactionnel de 25 µL contenant : le repliQa HiFi ToughMix 1X, 0.35 µM de chaque amorce, 5% de DMSO et 10 ng d'ADN. Le programme PCR commence par une dénaturation initiale de l'ADN à 98°C pendant 30s, il se poursuit par 30 cycles de dénaturation à 98°C pendant 10s, d'hybridation des amorces pendant 5s 55°C et d'élongation à 68°C pendant 1 minute.
- Pour la Taq Owibio®, la réaction PCR se fait dans un volume réactionnel de 25 µL contenant : le Owi AptaFi Hot Start Master Mix 1X, 0.35 µM de chaque amorce, 3% de DMSO et 10 ng d'ADN. Le programme PCR commence par une dénaturation initiale de l'ADN à 98°C pendant 1 minute, il se poursuit par 30 cycles de dénaturation à 98°C pendant 15s, d'hybridation des amorces pendant 15s 55°C et d'élongation à 72°C pendant 2 minutes.

Enfin, pour vérifier le bon déroulement de la réaction et estimer la taille des fragments amplifiés, les produits PCR ont été déposés sur un gel d'agarose 1, 2% selon le même protocole que celui détaillé pour tester la qualité des extraits d'ADN.

4. Construction des bases de données

A. Bases de données de mitogénomes

Il a été décidé de créer 4 bases de données de mitogénomes entiers, une pour chaque phylum. La construction des bases de données a été réalisée avec l'aide de Geneious Prime version 2026 1.1, un logiciel de biologie moléculaire et d'analyse de séquences. Les séquences de mitogénomes ont été téléchargées depuis NCBI Nucleotide (National Center for Biotechnology Information), une des plus grandes bases de données génomiques publiques. Ces séquences ont toutes été téléchargées en format GenBank depuis cette plateforme (Figure 7). Les bases de données ont été générées à l'aide de la recherche suivante, en faisant varier le terme « Polychaeta » pour « Mollusca », « Crustacean » ou encore « Echinodermata » pour les 3 autres phylums : « ("Polychaeta"[Organism] OR Polychaeta[All Fields]) AND mitochondrial[All Fields] AND complete[All Fields] AND genome[All Fields] AND mitochondrion[filter] ». Ces bases de données ont ensuite été triées sur Geneious Prime afin de supprimer les séquences des espèces non aquatiques et aux longueurs aberrantes (environ entre 10 kb et 20 kb selon les phylums). Afin de ne pas alourdir les bases de données et ainsi de rallonger la durée des opérations il a aussi été décidé de créer une autre base de données en ne gardant qu'un seul représentant de chaque genre, spécialement pour les alignements et les tests d'amorces. Le genre étant le rang taxonomique se situant entre la famille et l'espèce, il permet de garder une large couverture taxonomique tout en réduisant grandement le nombre de séquences à manier.

The screenshot shows the NCBI Nucleotide search interface. At the top, the NIH logo and 'National Library of Medicine' are visible. The search bar contains the query: `((\"Polychaeta\"[Organism] OR (\"Polychaeta\"[Organism] OR Polychaeta[All Fields])) AND mitochondrion`. Below the search bar, the results are displayed. On the left, there are filters for Species, Molecule types, Source databases, Sequence Type, Genomic compartments, and Sequence length. The main results area shows two items, both from *Decemunciger* sp. AB-2017, with details on their size and accession numbers. On the right, there are sections for 'Results by taxon' and 'Find related data'.

Figure 7 : Capture d'écran de la page web de NCBI Nucleotide lors de la recherche des mitogénomes des polychètes

B. Bases de données COI & 16S

À partir de ces bases de données, ainsi que d'une recherche bibliographique, les gènes du cytochrome-oxydase I (COI) ainsi que celui de L'ARN ribosomique 16S ont été choisis comme marqueurs. Ces deux régions sont intéressantes car elles sont déjà très documentées et encadrent un fragment long. De plus, les gènes présents dans ce fragment constituent d'autres marqueurs, améliorant l'assignation taxonomique et permettant une étude plus fine des populations. Ces deux régions ont été extraites des bases de données de mitogénomes grâce à la fonction « extract annotation » de Geneious

Prime, afin de créer des bases de données spécifiques de ces deux marqueurs. Afin d'avoir une représentation taxonomique la plus complète possible, ces bases de données ont ensuite été agrémentées de nouvelles séquences trouvées sur NCBI Nucleotide. Les recherches ont été menées au niveau taxonomique de la famille et ont été basées sur les inventaires locaux réalisés par la station biologique de Roscoff, permettant d'assurer que les espèces trouvées soient bien marines. Voici un exemple de requête pour trouver le COI d'un mollusque, de la famille des « margaritidae » : "margaritidae"[Organism:exp] AND ("cytochrome c oxidase subunit I "[Gene] OR COI[Gene] OR COI[All Fields] OR cox1[All Fields]) AND "Mollusca"[Organism]. Voici cette fois un exemple de requête pour trouver le 16S d'un crustacé, de la famille des « palaeocopida » : "palaeocopida"[Organism:exp] AND ("16S ribosomal RNA"[Gene] OR "16S rRNA"[All Fields] OR "16S"[All Fields]) AND "Crustacean"[Organism]. Ces opérations ont été menées pour chaque famille des 4 phylums. Une fois ces bases de données améliorées terminées, elles ont été alignées avec le programme MAFFT (Multiple Alignment using Fast Fourier Transform) afin de faire ressortir les régions conservées de l'ADN, où l'amorce viendra se fixer.

C. Bases de données des espèces « non cibles »

Enfin, il était nécessaire de créer une base de données de mitogénomes d'espèces non cibles afin de pouvoir tester la spécificité de nos amorces par la suite. Cette base de données a aussi été générée grâce à NCBI à l'aide de la requête suivante : "complete genome"[Title] AND "mitochondrion"[Title] AND "marine"[Habitat] AND RefSeq[Filter] NOT Mollusca[Organism] NOT Crustacea[Organism] NOT Polychaeta[Organism] NOT Echinodermata[Organism] AND (("10000"[SLEN] : "20000"[SLEN])). Le filtre « RefSeq » renvoie ici à la collection de séquences de référence de NCBI qui fournit un ensemble complet, intégré, non redondant et bien annoté de séquences. Ces séquences ont été directement triées entre 10 kb et 20 kb sur NCBI par la mention : (("10000"[SLEN] : "20000"[SLEN])).

5. Choix et conception des amorces *in Silico*

Afin d'amplifier la région du mitogénome qui se situe entre le COI et le 16S, il a été décidé de prendre une première amorce *forward* (5'-3') dans le COI et une deuxième amorce *reverse* (3'-5') dans le 16S. Une étape de recherche bibliographique a été réalisée dans le but de recenser de la manière la plus exhaustive possible l'ensemble des amorces déjà utilisées pour l'amplification du COI et du 16S des invertébrés. Ces recherches ont permis d'orienter la stratégie à adopter, à savoir : choisir une amorce *reverse* dans le 16S plutôt générale, qui cible tous les métazoaires et choisir une amorce *forward* dans le COI spécifique des invertébrés benthiques voire de chaque phylum. Cette stratégie paraissait être la plus logique étant donné qu'il existe déjà des amorces spécifiques des invertébrés pour le COI, contrairement au 16S qui est une région beaucoup moins documentée. Ces recherches ont permis de sélectionner plusieurs amorces, mais aussi de soulever la nécessité d'en modifier certaines et même d'en dessiner de nouvelles. La sélection et le design des amorces spécifiques *in silico* ont été réalisés en respectant plusieurs critères théoriques et techniques. Les amorces doivent :

- Être spécifique aux invertébrés benthiques ou à seulement un des 4 phylums, c'est-à-dire qu'elles doivent permettre l'amplification des mollusques, des crustacés, des échinodermes et des polychètes sans amplifier d'autres espèces marines ;
- Cibler les domaines du COI et du 16S de l'ADN mitochondrial ;

- Avoir une taille entre 18 et 25 nucléotides ;
- Avoir au moins un G ou un C en 3' mais pas plus de 3 d'affilée ;
- Avoir un %GC proche du %AT (~50%) ;
- Avoir une température de fusion (TM) se situant entre 57°C et 64°C ;
- Ne pas s'hybrider entre elles ;

Ces amorces ont été dessinées et testées à partir des alignements des bases de données du COI et du 16S développées dans cette étude. Elles ont aussi été testées sur la base de données des espèces non cibles, puis sur NCBI BLAST afin de s'assurer de leur spécificité. Les amorces choisies ont été synthétisées par Eurogentec. Enfin, ces amorces ont été positionnées sur les mitogénomes entiers afin de visualiser la taille et la composition théorique des fragments amplifiés grâce à la fonction « extract PCR product » de Geneious Prime.

6. Test des amorces par PCR

Les tests des amorces par PCR ont été réalisés sur des ADN extraits d'échantillons conservés dans le NucleoProtect, choisis de manière à représenter chacun des 4 phylums de l'étude, ainsi que 2 espèces de poissons afin de tester la spécificité des amorces (tableau 2). Les extraits d'ADN ont tous été réalisés avec le kit MasterPure et ont été dilués à 10ng/μL avant la PCR, à partir des résultats de la quantification Qubit™. Des témoins positifs et négatifs ont également été testés pour vérifier le bon déroulement de la réaction d'amplification et l'absence de contamination, respectivement.

Tableau 2 : Liste des échantillons utilisés pour tester les amorces par PCR

Code	Classe	Genre, espèce	Fixateur	Date d'échantillonnage	Kit d'extraction
mtPN7	Ophiuroidea	<i>Amphiura filiformis</i>	NucleoP	28/01/2026	MasterPure
mtPN14	Polycheta	<i>Glycera tridactyla</i>	NucleoP	28/01/2026	MasterPure
mtPN21	Malacostraca	<i>Ampelisca brevicornis</i>	NucleoP	28/01/2026	MasterPure
mtPN31	Gasteropode	<i>Nassarius reticulatus</i>	NucleoP	28/01/2026	MasterPure
mtPN57	Actinopterygii	<i>Callionymus lyra</i>	NucleoP	09/03/2026	MasterPure
mtPN58	Actinopterygii	<i>Mullus surmuletus</i>	NucleoP	09/03/2026	MasterPure

Les amorces ont été testées par PCR avec la Taq Quantabio®. La composition du volume réactionnel est similaire à celle décrite pour tester les Taq polymérase à l'exception qu'il a été ajouté 5μL d'ADN extrait au lieu de 2 μL. Le programme de PCR est aussi similaire à celui décrit pour le test des Taq polymérase, à l'exception que le temps d'élongation est passé de 1 à 5 minutes afin de permettre la synthèse de fragments encore plus longs.

Enfin, pour vérifier le bon déroulement de la réaction et estimer la taille des fragments amplifiés, les produits PCR ont été déposés sur un gel d'agarose 1,2% selon le même protocole que celui détaillé pour tester la qualité des extraits d'ADN.

IV – Résultats

1. Extraction de l'ADN des échantillons

A. Comparaison des kits d'extraction

Afin de développer le protocole d'extraction de l'ADN, 2 kits d'extraction ont été comparés : le MasterPure et le NucleoSpin Tissue. L'analyse de la qualité des extraits a été mesurée grâce au Nanodrop, l'analyse de la quantité d'ADN des extraits a été mesurée grâce au Qubit™, et enfin, l'intégrité des extraits a été mesurée par une électrophorèse sur gel d'agarose.

Les résultats du Nanodrop sont présentés en figure 8 et 9. Comme il est observable sur la figure 8, les rapports A260/A280 des extraits d'ADN de tous les échantillons se situent au-dessus de 1,7. Certains d'entre eux (PN1, PN17, PN18, PN19, PN26 et PN27) se situent légèrement au-dessus de 2.

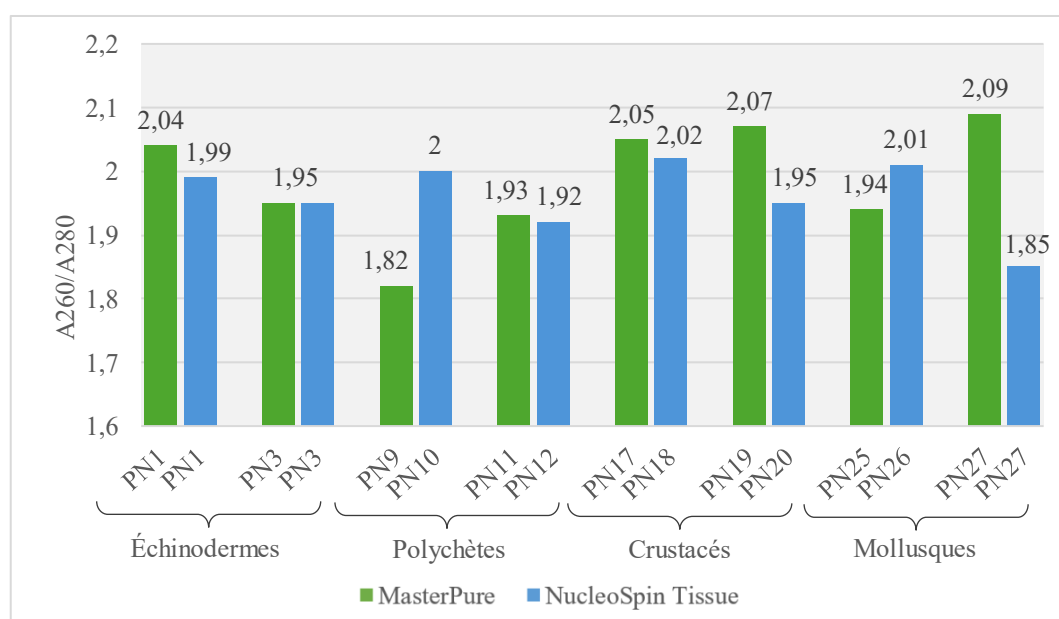


Figure 8 : Graphique du rapport A260/A280 en fonction de la taxonomie des échantillons analysés. Résultat Nanodrop. Comparaison du kit d'extraction MasterPure (en vert) et NucleoSpin Tissue (en bleu)

Comme il est observable sur la figure 9, les rapports A260/A230 des extraits d'ADN de tous les échantillons se situent tous largement au-dessus de 1 (le plus bas étant 1,7) et avoisinent le 2 voire le dépassent. Globalement, en considérant les deux ratios, les résultats Nanodrop témoignent que la qualité des extractions est correcte pour les 2 kits d'extraction.

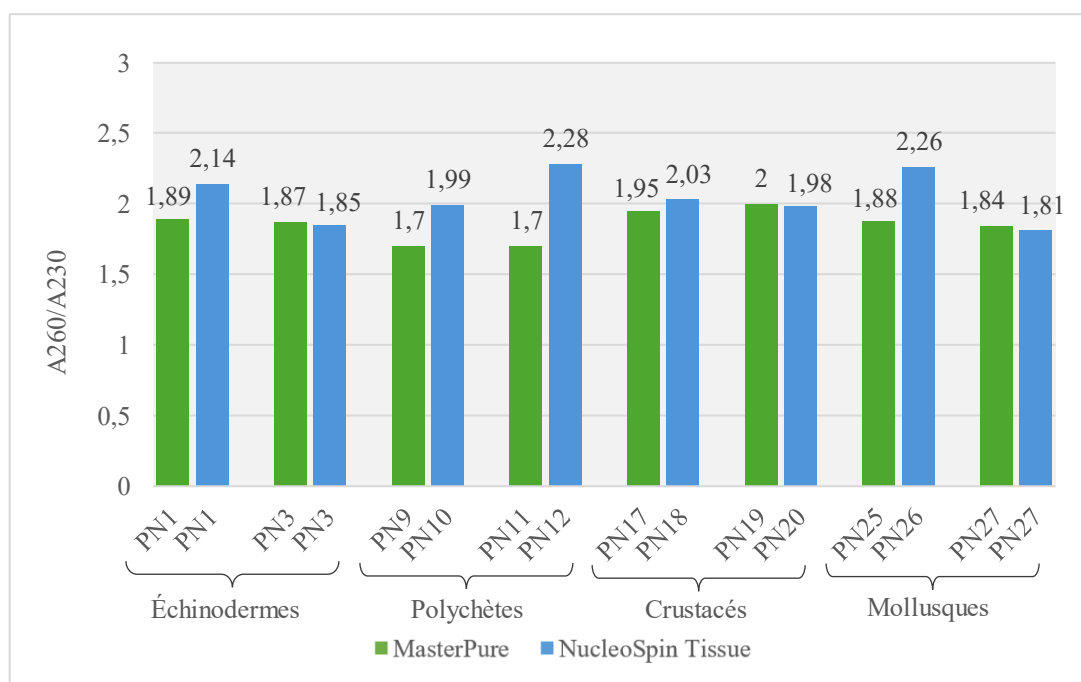


Figure 9 : Graphique du rapport A260/A230 en fonction de la taxonomie des échantillons analysés. Résultat Nanodrop. Comparaison du kit d'extraction MasterPure (en vert) et NucleoSpin Tissue (en bleu)

Les résultats de la quantification d'ADN mesurée grâce au kit Qubit ds HS 1X sont présentés en figure 10. On peut observer que la quantité d'ADN est très variable en fonction de la nature des échantillons et de la méthode d'extraction. Les extractions faites avec le kit MasterPure présentent une quantité d'ADN nettement supérieure aux extractions faites avec le kit NucleoSpin Tissue.

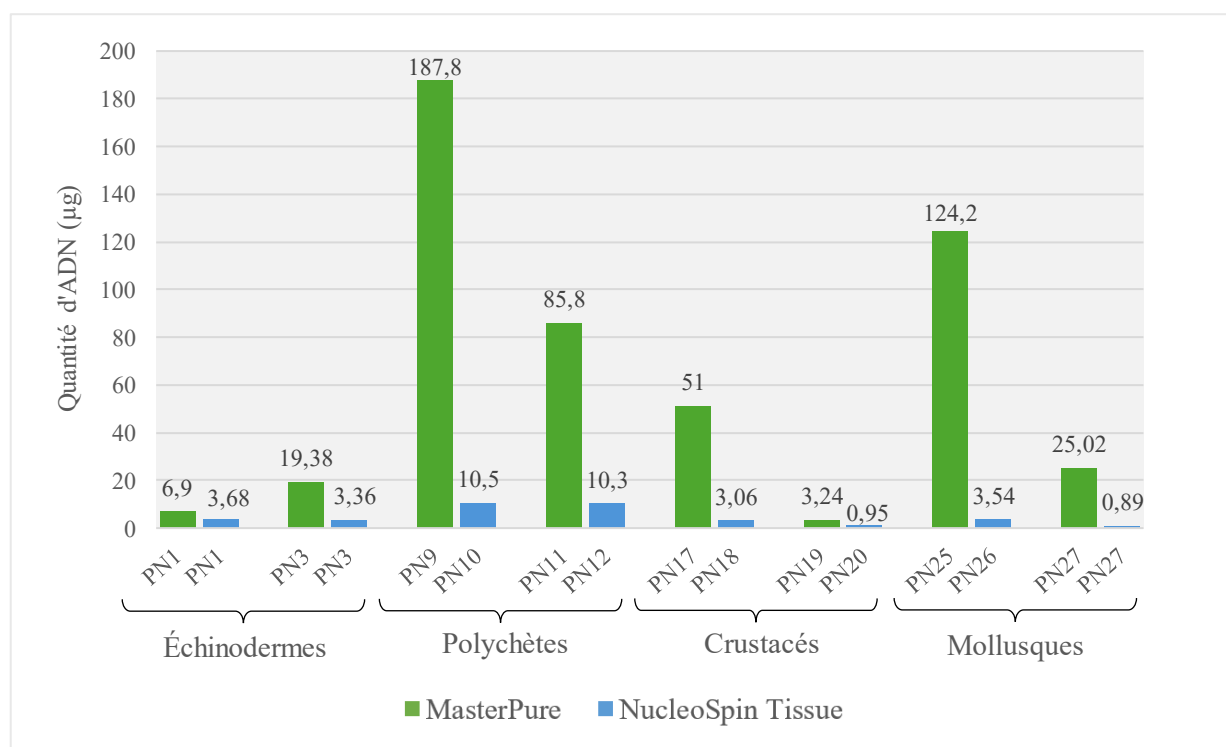


Figure 10 : Graphique de la quantité d'ADN (µg) en fonction des échantillons, résultats du kit Qubit DS HS 1X. Comparaison du kit d'extraction MasterPure (en vert) et NucleoSpin Tissue (en bleu)

Enfin, les résultats de l'électrophorèse sur gel d'agarose sont présentés en figure 11. On peut observer que les fragments d'ADN se situent pratiquement tous au-dessus de 10kb, ce qui est prometteur pour pouvoir faire du séquençage long-read. De plus, les bandes des extractions faites avec le kit MasterPure font moins de smear (une tache floue et étirée verticalement qui traduit la présence de nombreux fragments de tailles différentes au lieu de ne former qu'une seule bande précise) que celles faites avec le kit NucleoSpin Tissue, l'intégrité des fragments d'ADN semble donc être meilleure.

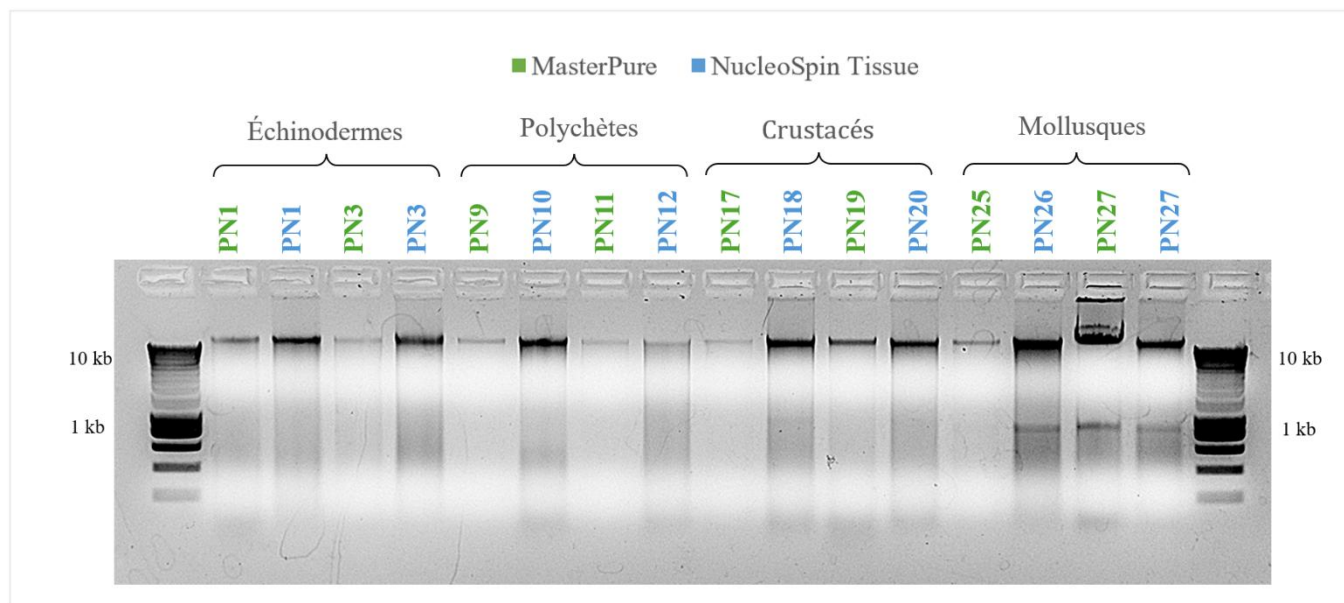


Figure 11 : Gel d'électrophorèse, comparaison du kit d'extraction MasterPure (en vert) et NucleoSpin Tissue (en bleu)

Les résultats Qubit™ et le gel d'électrophorèse permettent de constater que les extractions faites avec le kit MasterPure ont une quantité et une qualité d'ADN supérieures aux extractions faites avec le kit NucleoSpin Tissue. Il a été décidé que le kit d'extraction de l'ADN préférable serait le kit d'extraction MasterPure.

B. Comparaison des méthodes de conservation

Une fois la méthode d'extraction validée, deux méthodes de conservation ont été comparées : l'éthanol 96% et le NucleoProtect. Afin de pouvoir comparer ces deux fixateurs, il a été choisi des répliques ou tripliquas d'espèces de polychètes, échinodermes et crustacés. Tous les ADN ont été extraits avec le kit MasterPure puis la qualité des extraits a été mesurée grâce à une électrophorèse sur gel d'agarose (figure 11). On peut observer que les extraits d'ADN d'échantillons conservés dans le NucleoProtect sont beaucoup plus visibles que ceux des échantillons conservés dans l'éthanol 96% malgré qu'ils soient tous à la même concentration (10 ng/μL). À partir de ces résultats, il a été décidé que la méthode de conservation préférable serait le NucleoProtect.

Les expériences qui ont suivi ont donc toutes été réalisées à partir d'extraits d'ADN d'échantillons conservés dans le **NucleoProtect** et extraits avec le kit **MasterPure**.

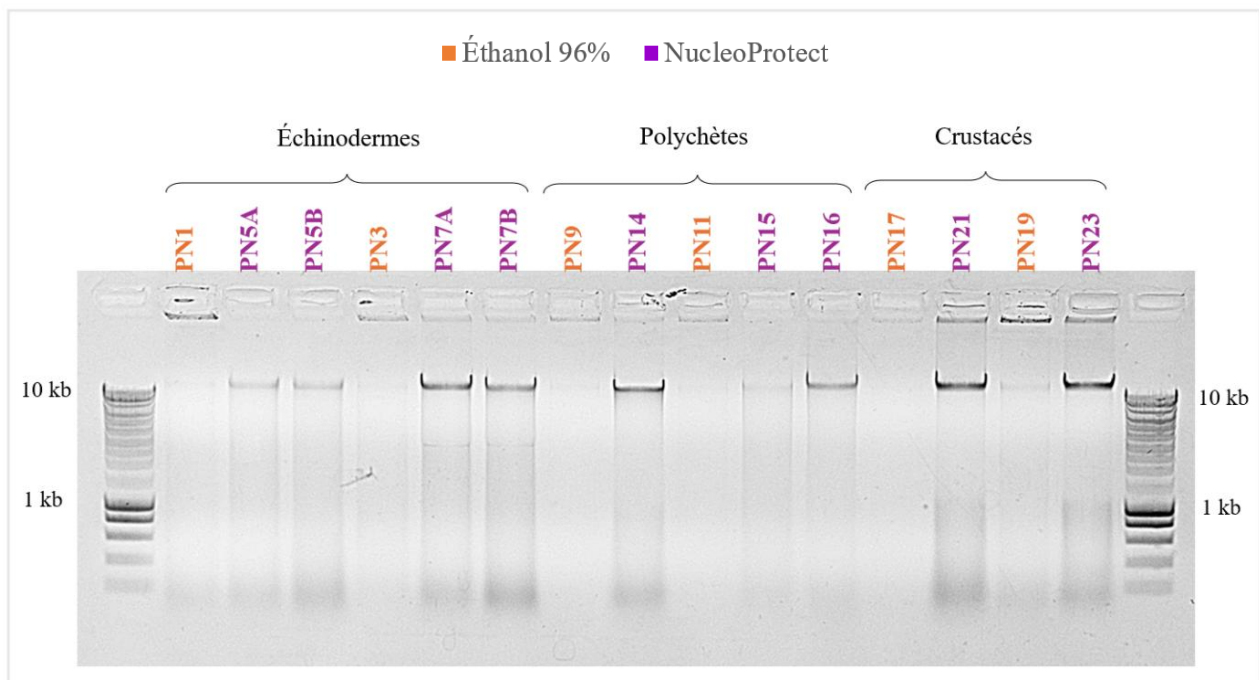


Figure 11 : Gel d'électrophorèse, comparaison des méthodes d'extraction : éthanol 96% (en orange) et NucleoProtect (en violet)

2. Test des différentes Taq par PCR

Après validation des méthodes de préservation et d'extraction de l'ADN, il a fallu sélectionner une Taq polymérase capable d'amplifier de longs fragments d'ADN par PCR. Pour cela, nous avons comparé 3 Taq polymérase : la Taq Phusion, la Taq repliQa HiFi ToughMix (Quantabio®), et enfin, la Taq Owi AptaFi Hot Start (Owibio®). Ces 3 enzymes ont été testées avec un couple d'amorces identiques ciblant une portion d'environ 4000 pb entre les gènes de l'ARN ribosomique 18S et 28S. La qualité des amplifications a été mesurée grâce à une électrophorèse sur gel d'agarose (figure 13).

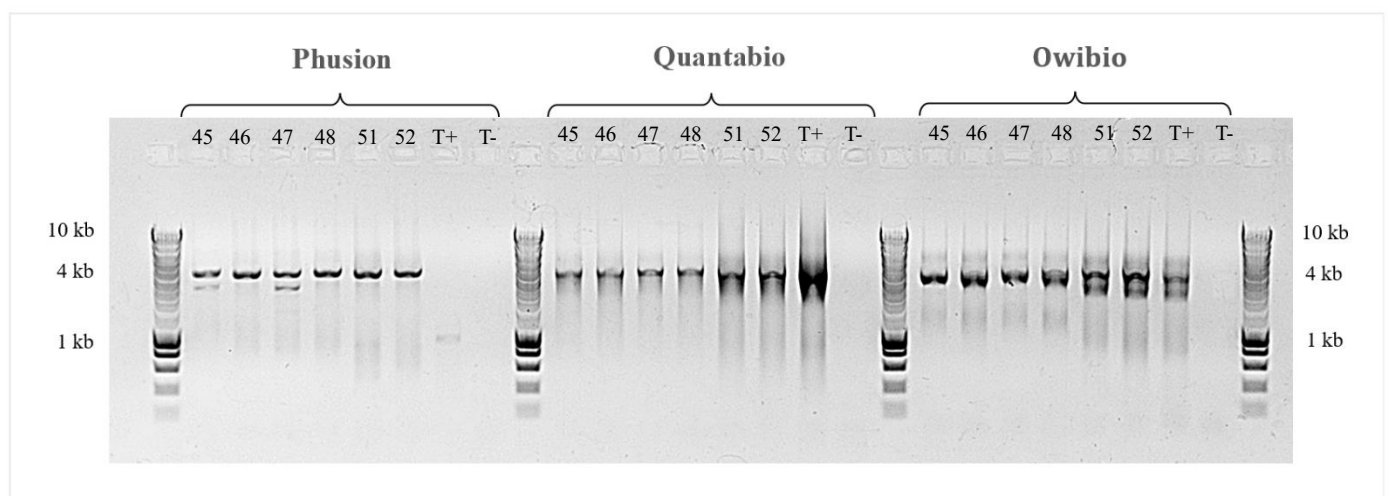


Figure 13 : Gel d'électrophorèse, comparaison de 3 Taq polymérase par PCR

On peut observer une absence d'amplification pour tous les témoins négatifs (T-), et une amplification à la bonne taille (4 kb) pour les témoins positifs (T+) des Taq Quantabio® et Owibio®, le témoin positif de la Taq Phusion se situant plutôt aux alentours des 1 kb. On peut aussi observer des amplifications à la bonne taille pour tous les échantillons avec les 3 Taq. Il est aussi observable que les bandes de la Taq Phusion sont moins étalées que celles de la Taq Quantabio® et Owibio®. Étant donné que les trois enzymes donnent de bons résultats de PCR, il a été décidé de les discriminer sur d'autres critères : le prix et leur longueur maximale d'amplification selon le site vendeur (tableau 3).

Tableau 3 : Comparaison du prix et des longueurs d'amplification de 3 Taq polymérases

Supplier	Vendeur	Nom	Nombre de runs	Prix €/run	Moyenne	Volume de solution	Longueur maximale des amplicons
Phusion	ThermoFisher	Phusion High-Fidelity PCR Master Mix avec HF buffer	100	2,29	2,048	2 × 1.25 mL	7,5 kb
			500	1,796		10 × 1.25 mL	
		Phusion High-Fidelity PCR Master Mix avec GC buffer	100	2,31		2 × 1.25 mL	
			500	1,796		10 × 1.25 mL	
Quantabio	VWR Avantor	RepliQa HiFi ToughMix	25	2,208	1,958	1 x 0.625 mL	24 kb
			100	1,95		1 x 2.50 mL	
			500	1,716		1 x 12.50 mL	
Owibio	Ozyme	Owi Aptafi Hot Start Master Mix	100	1,54	1,649	2 x 1.25 mL	17kb
			500	1,308		10 x 1.25 mL	
			20	2,1		1 x 0.5 mL	

Dans ce tableau, nous pouvons observer que la Taq Phusion est la plus chère (2,048€/run) et que sa longueur d'amplification est la plus courte. Nous pouvons aussi observer que la Taq Owibio® est la moins chère (1,649€/run). Enfin, la Taq Quantabio® a la longueur d'amplification la plus longue. À partir des résultats du gel d'électrophorèse et du tableau de comparaison des prix et de la longueur d'amplification, il a été décidé de faire la suite des expériences avec la **Taq Quantabio®**.

3. Bases de données

Une fois la méthode d'extraction et la Taq polymérase validées, il a fallu construire les bases de données afin de choisir et dessiner les amorces spécifiques des différents groupes ciblés (polychètes, échinodermes, crustacés et mollusques) pour permettre des amplifications de longs fragments par PCR. Les mitogénomes téléchargés depuis NCBI Nucleotide ont été triés pour supprimer les séquences des espèces non aquatiques, celles aux longueurs aberrantes et pour ne garder qu'un représentant par genre.

Les résultats de ce tri sont présentés sur la figure 14. On peut observer qu'il y a un nombre de mitogénomes bien plus élevé de crustacés et de mollusques que d'échinodermes et de polychètes. Ceci est expliqué par la différence du nombre d'espèces de ces phylums. En effet, on compte environ 85 000 espèces de mollusques et 60 000 espèces de crustacés contre environ 10 000 espèces de polychètes et d'échinodermes. Les longueurs des mitogénomes se situent à peu près toutes entre 13 kb et 20 kb (à part pour les mollusques où certaines séquences vont jusqu'à 22 kb).

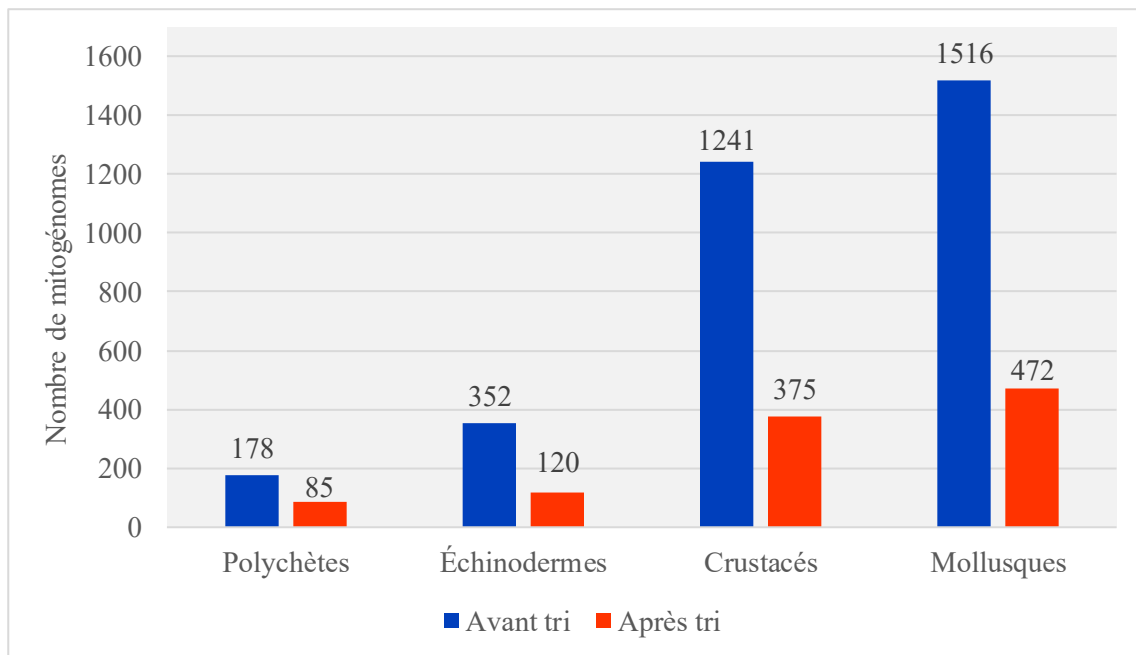


Figure 14 : Graphique du nombre de séquences des bases de données de mitogénomes avant et après avoir été triées

Une fois ces premières bases de données de mitogénomes entières constituées, il a été décidé de créer des bases de données spécifiques à nos deux marqueurs : le COI et le 16S. Si ces nouvelles bases de données ont été créées après extraction de ces gènes depuis les mitogénomes entières, elles ont été améliorées avec des séquences plus courtes des deux gènes cibles trouvées sur NCBI Nucleotide afin d'optimiser la couverture taxonomique lors de la conception des amorces et d'augmenter la représentativité des différentes lignées au sein du groupe étudié.

Les figures 15 et 16 montrent la couverture taxonomique apportée par ces nouvelles séquences au niveau des familles puis des ordres, respectivement. Ces graphiques ne représentent pas la proportion de séquences provenant de chaque source, mais bien la proportion de familles et d'ordres représentés par l'ensemble de ces séquences. On peut voir sur la figure 15 qu'une grande part de la diversité des familles des bases de données a été apportée par les séquences trouvées sur NCBI Nucleotide, allant jusqu'à 50% pour les COI des Polychètes. On peut aussi observer que la base de données la moins complète au niveau des familles est celle des 16S des échinodermes (73% de familles manquantes). Ce manque de représentativité est cependant moins présent au niveau des ordres dont la plupart sont déjà présents dans les séquences extraites des mitogénomes. Il est aussi important de noter que les séquences téléchargées depuis NCBI Nucleotides sont moins longues que celles directement extraites des mitogénomes. En effet, pour ces dernières, les longueurs des gènes du COI se situent entre 1500 et 1600 pb et celles du 16S sont entre 1300 et 1400pb, tandis que les longueurs des séquences téléchargées depuis NCBI Nucléotides se situent entre 600 et 700 pb pour le COI et entre 500 et 600 pour le 16S. Ceci est expliqué par le fait que les séquences ajoutées sur NCBI Nucleotide sont souvent le résultat de PCR ne ciblant qu'une partie du gène qui n'est alors pas complet, contrairement à ceux extraits des mitogénomes. Une fois améliorées, ces bases de données spécifiques au COI et au 16S ont été alignées.

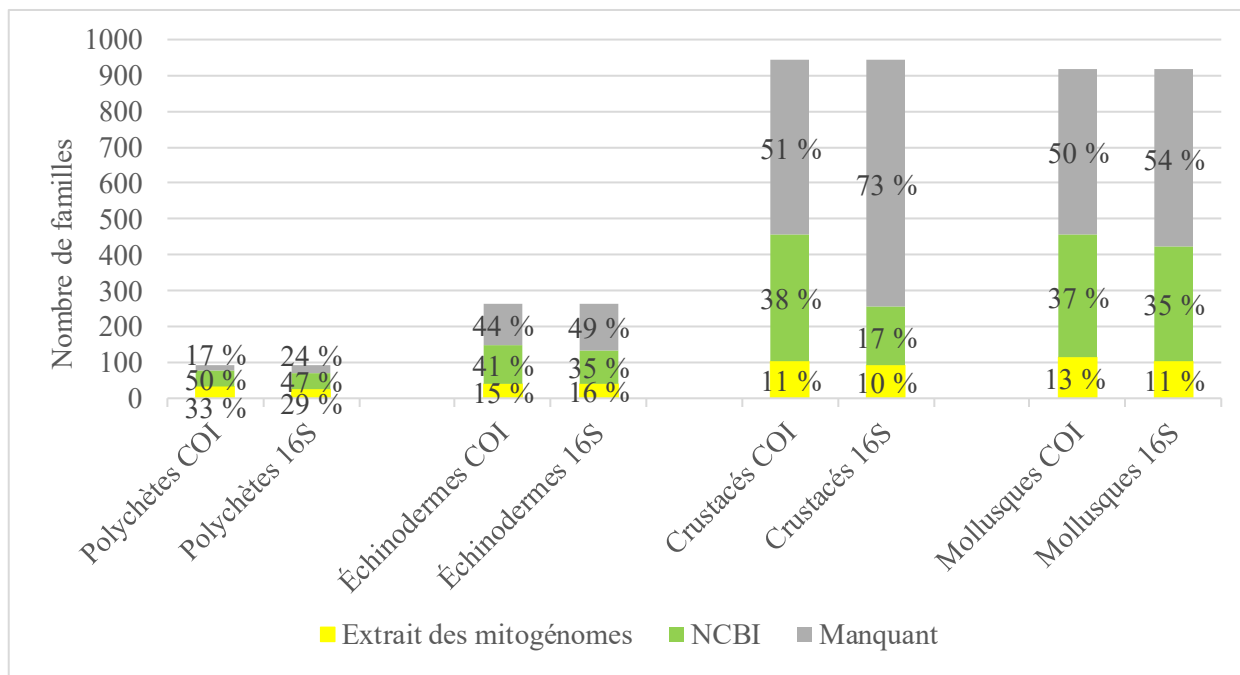


Figure 15 : Graphique du nombre de familles apportées par les gènes extraits de mitogénomes (en jaune), ceux trouvées sur NCBI (en vert) et les manquants (en gris)

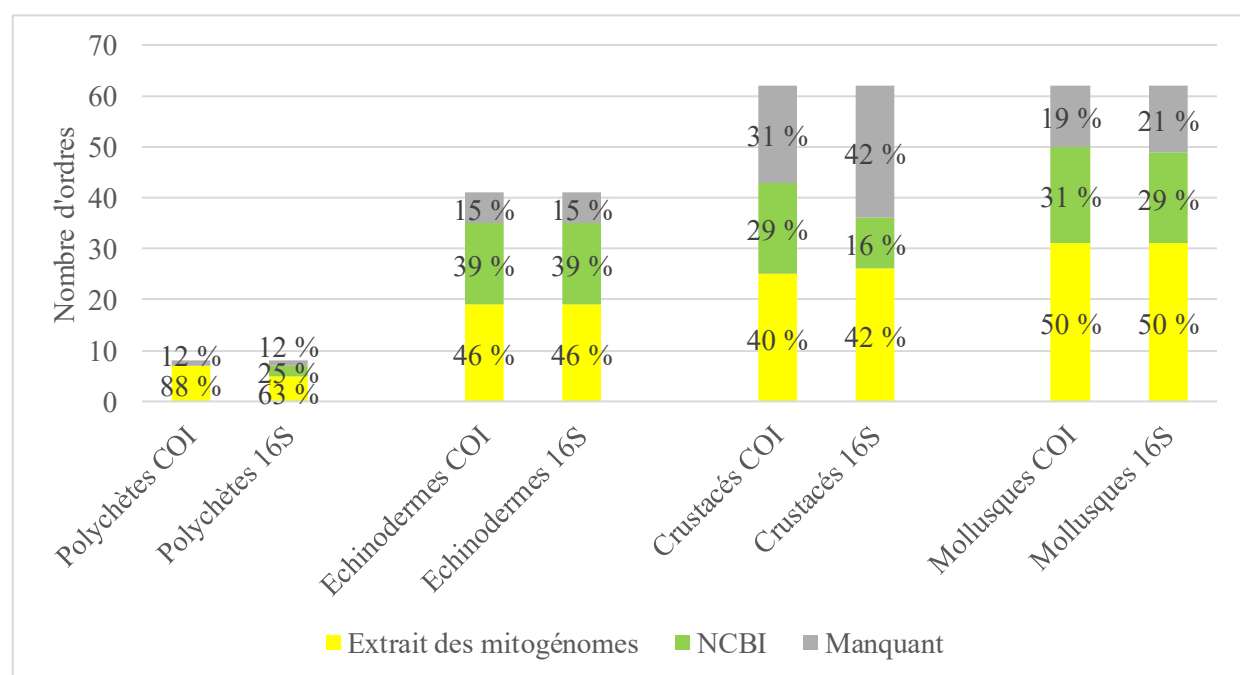


Figure 16 : Graphique du nombre de familles apportées par les gènes extraits de mitogénomes (en jaune), ceux trouvées sur NCBI (en vert) et les manquants (en gris)

Enfin, une base de données d'espèces non cibles a été conçue afin de pouvoir tester la spécificité des amorces. Le but de celle-ci est de représenter le maximum d'espèces pouvant être présentes dans le même environnement que les mollusques, crustacés, échinodermes et polychètes afin de s'assurer que si l'ADN de ces espèces non cible est échantillonné, il ne pourra pas être amplifié par les amorces choisies. La figure 17 présente la diversité de cette base de données, constituée de 1039 séquences.

Même si celle-ci est majoritairement constituée de poissons, on y retrouve aussi des protistes, des mammifères, des reptiles etc.

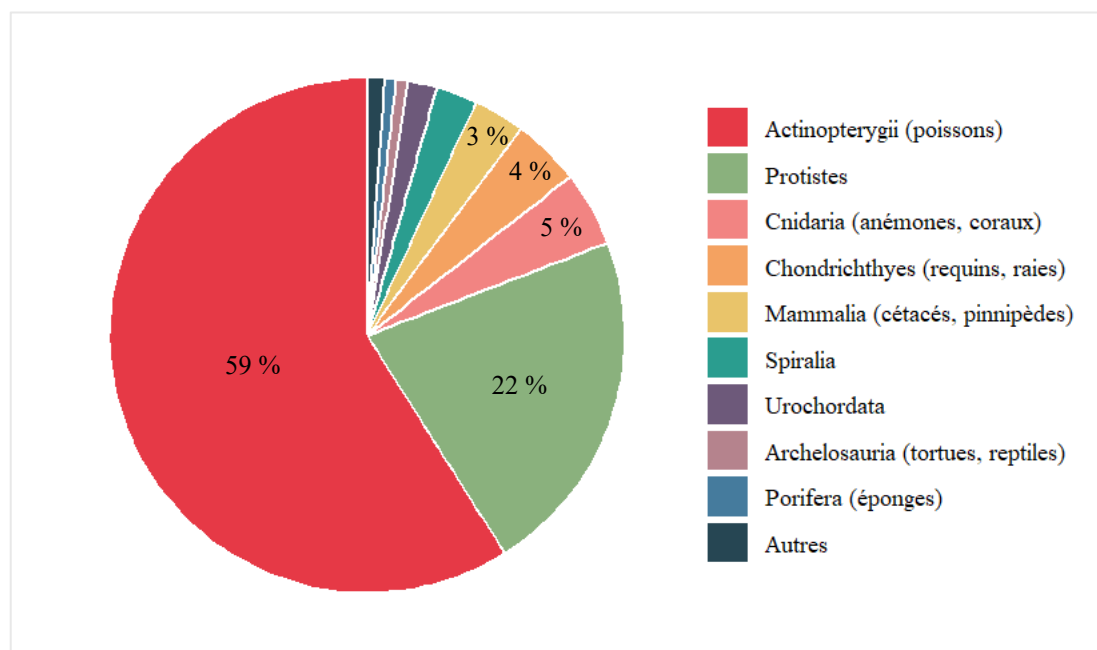


Figure 17 : Graphique de la proportion de chacun des groupes présents dans la base de données des espèces non-cibles

4. Choix et conception des amorces *in Sillico*

À partir des différents alignements de séquences COI et 16S, il a été possible de choisir et dessiner les amorces. Le tableau 4 présente les amorces retenues : 2 amorces *reverse* ciblant le 16S et 3 amorces *forward* ciblant le COI afin d'amplifier toute la partie du mitogénome se situant entre ces deux gènes. La deuxième amorce du 16S (16S-Metazoa_RevMod) a été modifiée à partir de celle de Kelly *et al.* afin de cibler davantage d'espèces de nos 4 phylums.

Tableau 4 : Amorces ciblant le COI et le 16S utilisées dans cette étude avec leur Tm, longueur, %GC et source

Nom	Région cible	Groupe cible	Sens	5'-3'séquence	Tm°C	Longueur (pb)	%GC	Source
16S-Metazoa_Rev	16S	Métazoaires	R	CCG-GTC-TGA-ACT-CAG-ATC-AYG-T	59,8-62	22	52,4	Kelly <i>et al.</i> , 2016
16S-Metazoa_RevMod	16S	Métazoaires	R	CG-KTC-TRA-ACT-CAW-CTC-AYG-T	51,5-58,6	20	47,1	This study
LCO1490	COI	Invertébrés	F	GGT-CAA-CAA-ATC-ATA-AAG-ATA-TTG-G	56,4	25	32,0	Folmer <i>et al.</i> , 1994
jgLCO1490	COI	Invertébrés	F	TI-TCI-ACI-AAY-CAY-AAR-GAY-ATT-GG	51,8-65,4	25	33,3	Geller <i>et al.</i> , 2013
mLCOintF	COI	Invertébrés	F	GGW-ACW-GGW-TGA-ACW-GTW-TAY-CCY-CC	60,5-64,3	26	50	Leray <i>et al.</i> , 2013

Une fois ces amorces sélectionnées, elles ont été placées sur les mitogénomes de notre base de données afin de visualiser le fragment théorique à amplifier comme il est observable sur la figure 18. On peut y voir un mitogénome de 14 918 pb d'une espèce de polychète où sont placées les amorces. Le trait reliant les amorces étant la région théoriquement amplifiée par PCR.

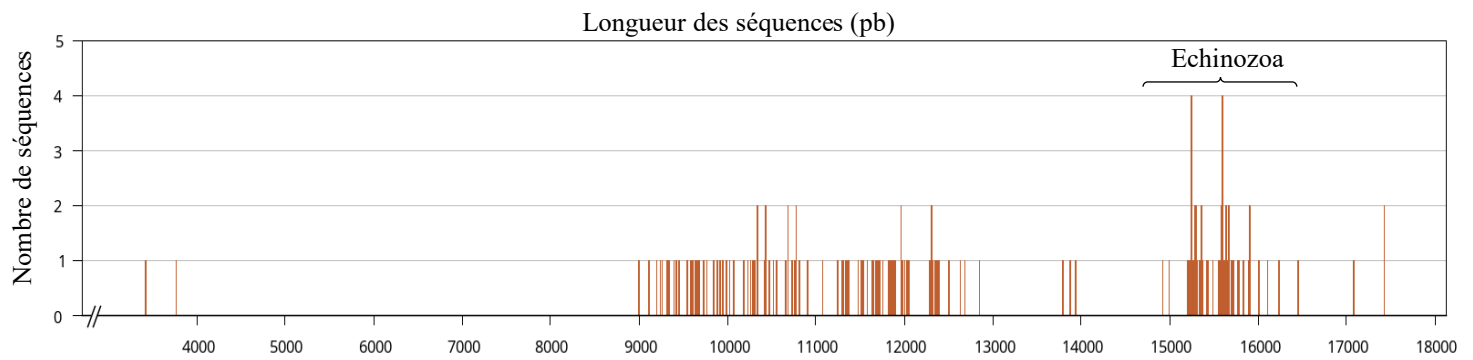


Figure 20 : Graphique des tailles des produits PCR théoriques des échinodermes

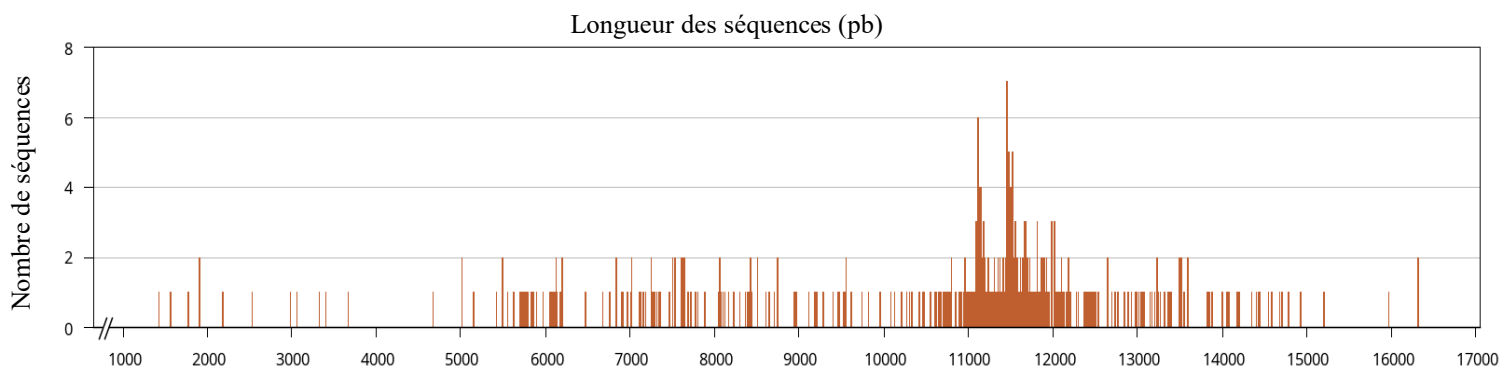


Figure 21 : Graphique des tailles des produits PCR théoriques des crustacés

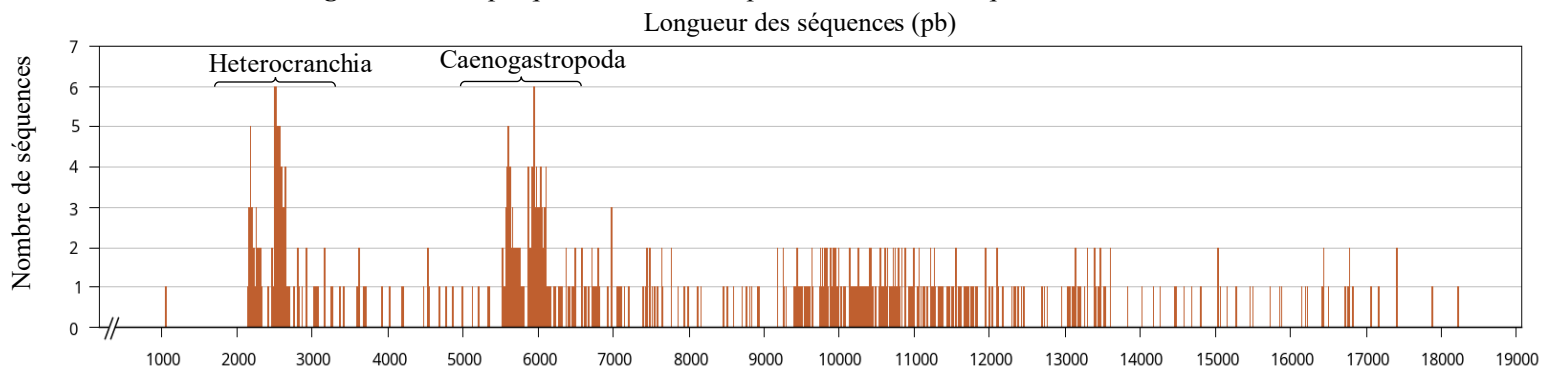


Figure 22 : Graphique des tailles des produits PCR théoriques des mollusques

Les figures 23 à 26 montrent des exemples de fragments amplifiés par PCR pour chacun des phylums. On peut observer que tous les fragments commencent logiquement par le COI et se terminent par le 16S. La composition génétique n'est pas exactement la même d'un phylum à un autre mais certains gènes reviennent souvent comme le 12S, le cytochrome b (noté COB ou CYTB sur les figures) ou encore les cytochrome-oxydase II et cytochrome-oxydase III (notés COX2 et COX3 sur les figures).

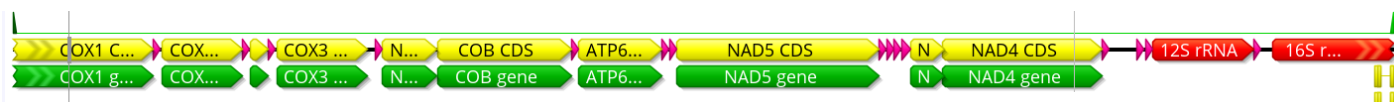


Figure 23 : Exemple de la composition génétique d'un fragment d'ADN de polychète obtenu par PCR



Figure 24 : Exemple de la composition génétique d'un fragment d'ADN d'échinoderme obtenu par PCR

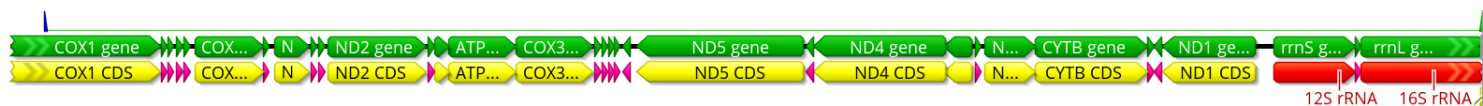


Figure 25 : Exemple de la composition génétique d'un fragment d'ADN de crustacé obtenu par PCR

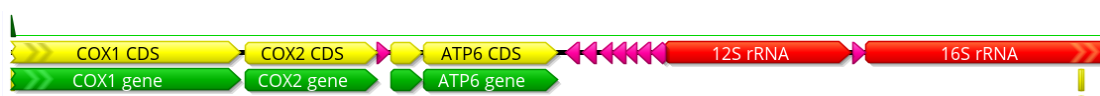


Figure 26 : Exemple de la composition génétique d'un fragment d'ADN de mollusque obtenu par PCR

5. Test des amorces par PCR

Une fois les amorces sélectionnées *In Silico*, elles ont été testées par PCR sur nos extraits d'ADN à partir d'échantillons choisis (tableau 2). Le tableau 5 présente les couples d'amorces retenues pour ces tests : les 6 premiers couples sont les amorces à tester et les couples 7 et 8 sont des couples contrôles. Le couple 7 cible une petite partie du 16S (~160 pb) et le couple 8 est le même couple d'amorce utilisé pour tester les Taq polymérases, il cible une portion d'environ 4000 pb entre le 18S et le 28S. Cela permet d'avoir deux contrôles différents : un pour les fragments courts et un pour les fragments longs.

Tableau 5 : Couples d'amorces testés par PCR dans cette étude

Couple	Nom des amorces	Cible	Sens	Séquence 5'-3'
1	LCO1490	COI	F	GGTCAACAAATCATAAAGATATTGG
	16S-Metazoa_Rev	16S	R	CCGGTCTGAACTCAGATCAYGT
2	jgLCO1490	COI	F	TITCIACIAAYCAYAARGAYATTGG
	16S-Metazoa_Rev	16S	R	CCGGTCTGAACTCAGATCAYGT
3	mICOIntF	COI	F	GGWACWGGWTGAACWGTWTAYCCYCC
	16S-Metazoa_Rev	16S	R	CCGGTCTGAACTCAGATCAYGT
4	LCO1490	COI	F	GGTCAACAAATCATAAAGATATTGG
	16S-Metazoa_RevMod	16S	R	CGKTCTRAACTCAWCTCAYGT
5	jgLCO1490	COI	F	TITCIACIAAYCAYAARGAYATTGG
	16S-Metazoa_RevMod	16S	R	CGKTCTRAACTCAWCTCAYGT
6	mICOIntF	COI	F	GGWACWGGWTGAACWGTWTAYCCYCC
	16S-Metazoa_RevMod	16S	R	CGKTCTRAACTCAWCTCAYGT
7	16S-Metazoa_fwd	16S	F	AGTTACYYTAGGGATAACAGCG
	16S-Metazoa_Rev	16S	R	CCGGTCTGAACTCAGATCAYGT
8	SA_Min_V14	18S-28S	F	TTTCTGTTGGTGCTGATATTGCAACCTGGTTGATCCTGCCAGT
	D2C-R_Min_V14	18S-28S	R	ACTTGCCTGTCGCTCTATCTTCCCTTGGTCCGTGTTCAAGA

Une étude préalable aux PCR a été menée afin d'estimer la longueur des amplicons pour chaque échantillon (tableau 6). Lorsqu'il n'était pas possible de retrouver exactement le mitogénome correspondant à l'espèce dans la base de données, la longueur du produit PCR a été estimée à partir des mitogénomes qui s'en rapprochaient le plus taxonomiquement. Ceci explique qu'il n'y ait pas de longueur exacte pour l'espèce de crustacé. Les cases du tableau barrées signifient qu'il ne devrait pas y avoir d'amplification, soit parce que ce ne sont pas les espèces cibles des amorces (pour les poissons) soit parce que les tests de compatibilité des amorces *In Silico* étaient négatifs.

Tableau 6 : Longueurs théoriques des fragments amplifiés par chaque couple d'amorce PCR

SET#	Nom des amorces	PN7 (Échinoderme)	PN14 (Polychète)	PN21 (Crustacé)	PN31 (Mollusque)	PN57 (poisson)	PN58 (poisson)
1	LCO1490			7 kb - 10,7 kb			
	16S-Metazoa_Rev						
2	16S-Metazoa_Rev	10 kb	11,7 kb	7 kb - 10,7 kb	5,9 kb		
	16S-Metazoa_Rev						
3	mICOLintF	9,5 kb		6,6 kb - 10,4 kb	5,6 kb		
	16S-Metazoa_Rev						
4	LCO1490			7 kb - 10,7 kb			
	16S-Metazoa_RevMod						
5	16S-Metazoa_RevMod	10 kb	11,7 kb	7 kb - 10,7 kb	5,9 kb		
	16S-Metazoa_RevMod						
6	mICOLintF	9,5 kb		6,6 kb - 10,4 kb	5,6 kb		
	16S-Metazoa_RevMod						
7	16S-Metazoa_fwd	160 pb	160 pb	160 pb	160 pb	160 pb	160 pb
	16S-Metazoa_Rev						
8	SA_Min_V14	4 kb	4 kb	4 kb	4 kb	4 kb	4 kb
	D2C-R_Min_V14						

Les résultats de l'électrophorèse sur gel d'agarose sont présentés en figure 27. On peut observer que les contrôles négatifs ont bien fonctionné, en revanche, le contrôle positif n'est pas adapté car il ne présente jamais d'amplification à part pour le set 8 (contrôle). Le set contrôle fragments longs (set 8) a très bien fonctionné, en revanche, le set contrôle fragments courts ne présente des amplifications que pour les 2 poissons. Les échantillons d'invertébrés ne présentent que rarement des bandes aux longueurs attendues, seulement pour le set 3 où l'on peut observer des bandes aux bonnes longueurs et le set 6 où on peut distinguer une légère amplification du crustacé (PN21). De plus, les sets 2 et 5 ne présentent quasiment aucune amplification, ce qui pourrait provenir d'une incompatibilité entre la Taq Quantabio® et le primer « jgLCO1490 » qui est composé d'inosines. Enfin, cette PCR soulève surtout le problème de la non-spécificité des amorces. En effet, même si les amorces *forward* du COI ne sont censées cibler que les invertébrés, ce n'est en réalité pas le cas car les 2 poissons (PN57 et PN58) ont été amplifiés par les sets 1, 2, 3 et 6.

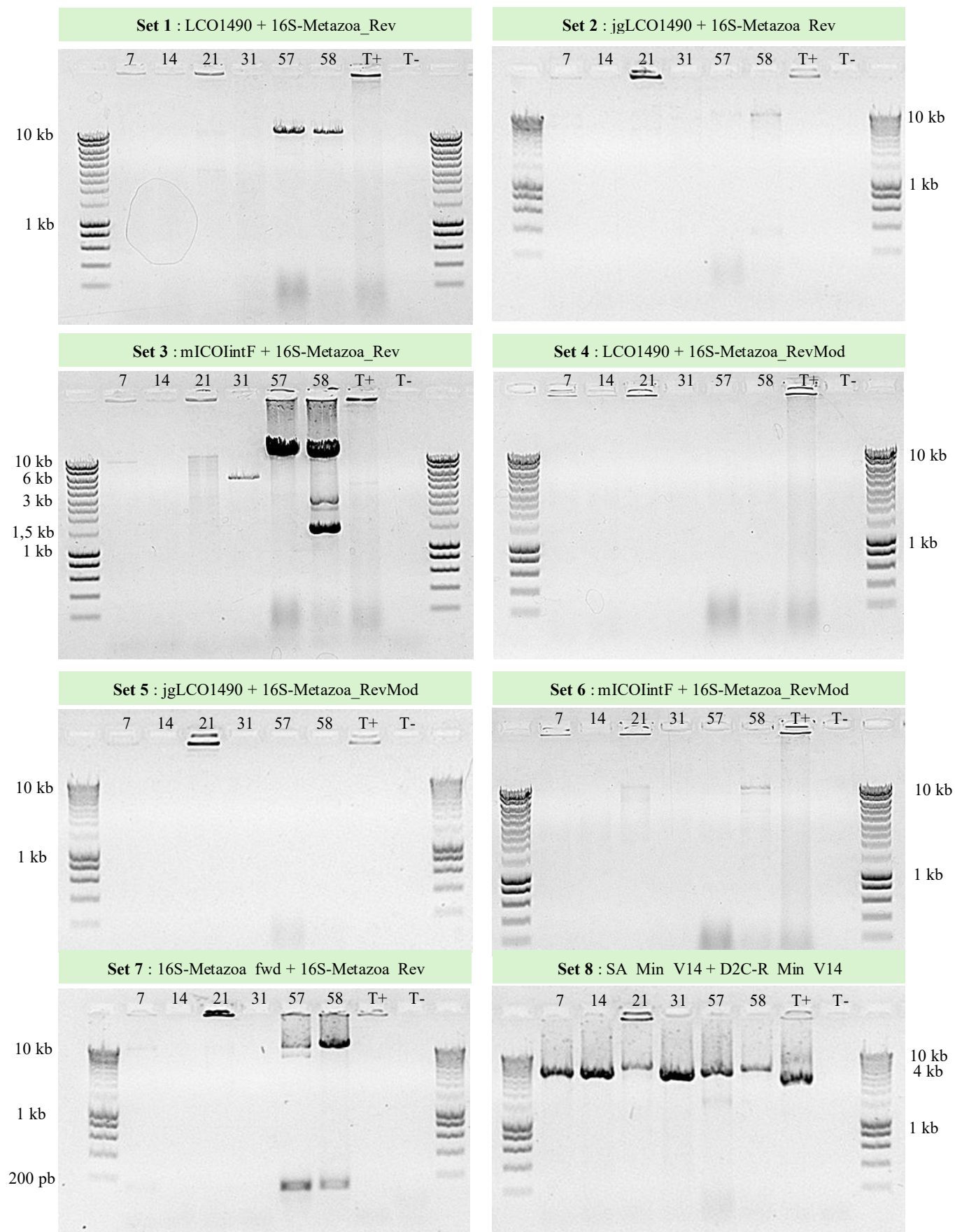


Figure 27 : Gel d'électrophorèse, résultat du test PCR des amorces

Afin de tester si les mauvais résultats d'amplification des sets 2 et 5 provenaient bien de la Taq Quantabio®, une nouvelle PCR a été menée seulement sur ces 2 sets mais cette fois avec la Taq Advantage® (Takara Bio), décrite dans la littérature comme fonctionnant avec les inosines (figure 28). Cette enzyme a été utilisée en suivant le protocole décrit dans Leray, 2013. On peut observer que cette fois, le mix contrôle fragments courts a fonctionné mais pas le mix contrôle fragments longs. Les sets 2 et 5 ne présentent toujours pas d'amplification. Bien que la Taq Advantage® (Takara Bio) soit compatible avec les inosines, elle semble plutôt faite pour l'amplification de fragments courts, ce qui pourrait expliquer l'absence d'amplification.

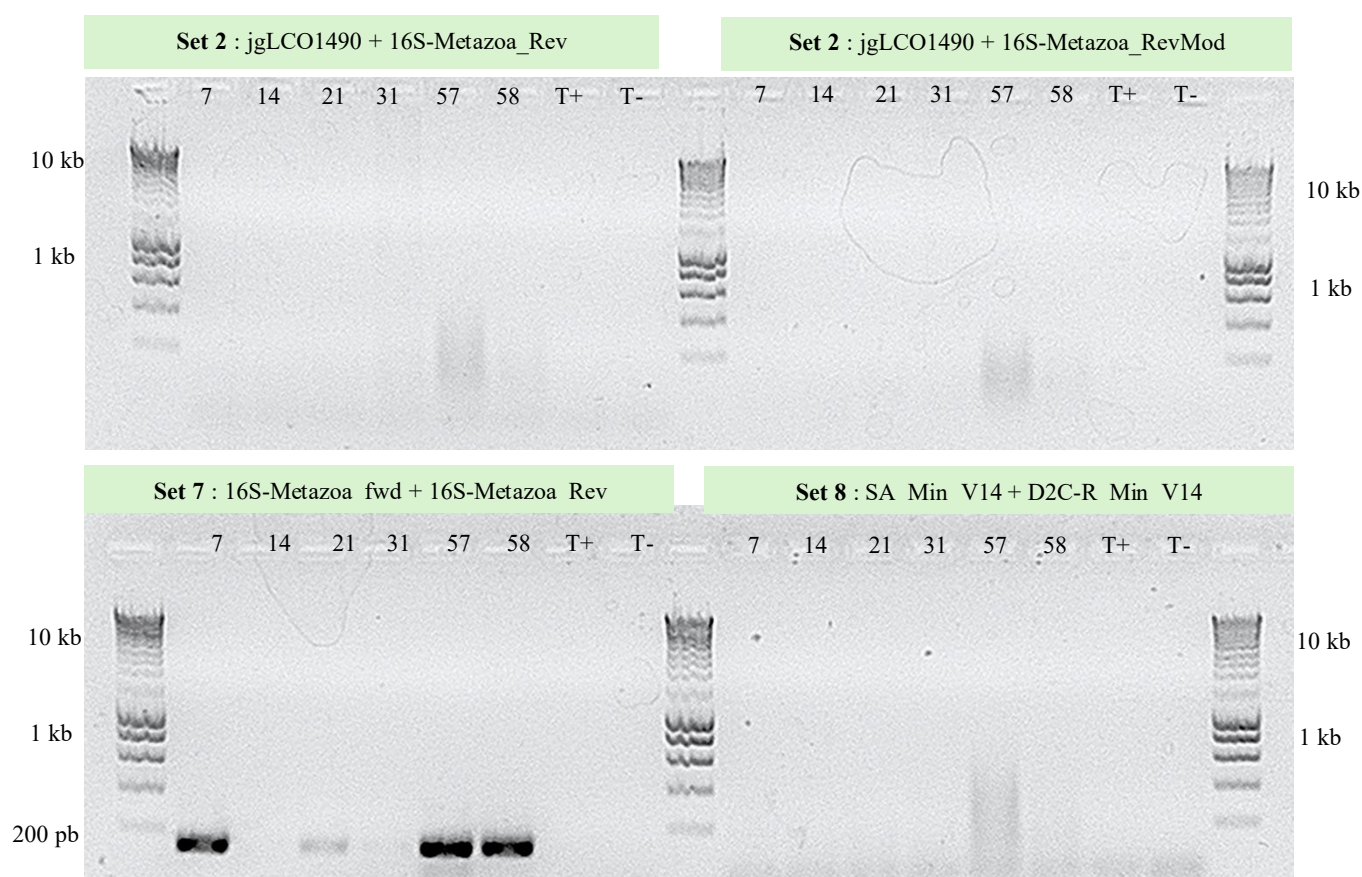


Figure 28 : Gel d'électrophorèse, résultat du test PCR sur les sets 2 et 5 avec la Taq Advantage

6. Test de la spécificité des amorces

Les PCR ayant relevé le manque de spécificité des amorces, il a été décidé de choisir et de dessiner de nouvelles amorces plus spécifiques, qui ne cibleraient cette fois pas tous les invertébrés mais chacun des 4 phylums choisis plus spécifiquement. À titre d'exemple, le tableau 7 présente les nouvelles amorces retenues pour les mollusques et les échinodermes. On y retrouve 2 amorces dessinées pour les échinodermes et 4 pour les mollusques. La deuxième amorce des échinodermes « COIeF1_long_F » a été dessinée en allongeant celle de Arndt *et al.*, 1996. Les amorces « Moll_COI_945F » et « Moll_COI_757F » n'ont pas été dessinées à partir d'amorces déjà existantes.

Tableau 7 : Nouvelles amorces plus spécifiques ciblant le COI des échinodermes et des mollusques

Nom	Région cible	Groupe cible	Sens	5'-3'sequence	Tm°C	Longueur (pb)	%GC	Source
COIeF1	COI	Échinoderme	F	ATAATGATAGGAGGRT TTGG	49,2 - 51,6	20	36,8	Arndt <i>et al.</i> , 1996
COIeF1_long F	COI	Échinoderme	F	CCAATAATGATAGGAG GRTTTGG	53,6- 59,5	23	42,9	This study
Moll_COI_253 F	COI	Mollusque	F	GGAGTAGGAAGTGGTT GGAC	57,2	20	55	Shi <i>et al.</i> , 2024
Moll_COI_945 F	COI	Mollusque	F	TACTACTATTATTAATA TACGATGAC	50,2	26	23	This study
GHalf	COI	Mollusque	F	TCAACAAATCATAAAG ATATYGG	51-52,7	23	27,3	Saunders, 2008
Moll_COI_757 F	COI	Mollusque	F	TAGAAAGAGGAGCTGG AACT	55,2	20	45	This study

Toutes les amorces ont ensuite été testées sur les bases de données de mitogénomes des 4 phylums et des espèces non cible. Les amorces auraient aussi pu être testées sur les bases de données du COI et du 16S plutôt que sur les mitogénomes afin d'avoir une meilleure représentation taxonomique. Cependant, les amorces n'étaient parfois pas compatibles avec les séquences du COI et du 16S du fait qu'elles n'étaient pas complètes, ce qui faussait les résultats.

Le tableau 8 recense toutes les amorces, ainsi que leur pourcentage de compatibilité pour chaque groupe. On peut observer que les deux amorces 16S sont très compatibles avec tous les groupes, ce qui était attendu étant donné qu'elles sont dessinées pour cibler les métazoaires. Il est aussi notable que l'amorce « 16S_Metazoa_rev » est moins compatible avec le groupe non-cible (74,4 %) que la « 16S_Metazoa_revMod » (94,13 %). Une amorce COI idéale serait une amorce très compatible avec les 4 phylums - ou un des 4 phylums - mais peu compatible avec le groupe « non-cible ». Ce tableau permet de confirmer que les 3 amorces COI testées par PCR (« LCO1490 », « jgLCO1490 » et « mICOLintF ») ne sont pas spécifiques aux invertébrés. En effet, la première n'est compatible avec aucun des groupes et les deux dernières sont compatibles avec tous les groupes. Les amorces ciblant les échinodermes semblent être prometteuses car elles sont très compatibles avec les échinodermes et peu avec les « non cibles ». Concernant les amorces dessinées pour les mollusques, seule l'amorce « GHalf » semble être intéressante, les 3 autres n'étant compatibles avec aucun des groupes. Cette amorce est d'autant plus prometteuse qu'elle est aussi compatible avec les polychètes et les crustacés.

Tableau 8 : Pourcentage de compatibilité entre les amorces et les différents groupes taxonomiques

Amorces	Groupe cible	Pourcentage de compatibilité				
		Polychètes	Échinodermes	Crustacés	Mollusques	Non cible
16S Metazoa rev	Métazoaires	95,29 %	99,17 %	95,72 %	98,09 %	74,40 %
16S_Metazoa_revMod	Métazoaires	100,00 %	90,83 %	97,59 %	97,24 %	94,13 %
LCO1490	Invertébrés	15,29 %	0,00 %	21,66 %	28,03 %	11,36 %
jgLCO1490	Invertébrés	95,29 %	96,67 %	97,86 %	91,93 %	97,31 %
mICOLintF	Invertébrés	92,94 %	86,67 %	95,19 %	83,65 %	84,89 %
COIeF1	Échinoderme	21,18 %	93,33 %	42,51 %	43,52 %	26,37 %
COIeF1_long	Échinoderme	58,82 %	88,33 %	14,71 %	22,08 %	22,33 %
Moll_COI_253_F	Mollusque	14,12 %	20,83 %	52,30 %	16,14 %	10,20 %
Moll_COI_945F	Mollusque	3,53 %	0,00 %	0,27 %	24,63 %	0,48 %
GHalf	Mollusque	77,65 %	22,50 %	82,09 %	80,25 %	22,91 %
Moll_COI_757F	Mollusque	1,18 %	51,67 %	21,66 %	18,90 %	2,02 %

V – Discussion

1. Améliorations et perspectives de l'étude

Afin d'amplifier la partie du mitogénome entre le COI et le 16S, plusieurs amorces ont été testées. Comme le gel d'électrophorèse (figure 27) et le tableau 8 l'ont exposé, les trois premières amorces ciblant le COI se sont révélées non spécifiques aux invertébrés. Ces amorces ne sont donc pas idéales pour faire de l'analyse d'ADN environnemental car elles ciblent aussi bien les 4 phylums qui nous intéressent que les autres espèces présentes dans le milieu aquatique. En revanche les nouvelles amorces dessinées pour le COI des échinodermes et des mollusques paraissent bien plus spécifiques et s'avèrent être de meilleures options. En particulier, les deux amorces ciblant les échinodermes peuvent être des choix envisageables. L'amorce « COIeF1 » paraît être davantage compatible avec les échinodermes que l'amorce « COIeF1_long », mais elle est aussi davantage compatible avec des espèces non cibles (ex : crustacés et mollusques). Une suite intéressante pour aller au-delà des tests *in silico* serait de tester ces deux amorces par PCR afin de constater laquelle est la plus adaptée. L'amorce GHalf semble aussi être une option intéressante, non seulement pour les mollusques mais aussi pour les polychètes et les crustacés. Encore une fois, un test par PCR sur des extraits d'ADN de différents spécimens permettrait de valider cette hypothèse de dessin d'amorces. Concernant les amorces ciblant le 16S, il semblerait que l'amorce non modifiée « 16S-Metazoa_rev » soit finalement plus adaptée que l'amorce « 16S-Metazoa_revMod » étant donné qu'elle n'engendre quasiment aucune perte de couverture taxonomique (voire en fait gagner pour les échinodermes) et qu'elle est moins compatible avec les espèces « non-cibles ». Il est tout de même important de noter que ces tests de compatibilité ont été faits sur des bases de données loin de représenter la diversité des 4 phylums. Il faudrait sûrement approfondir ces résultats en enrichissant les bases de données de mitogénomes. Enfin, pour compléter cette étude il faudrait séquencer ces résultats PCR à l'aide de la méthode de séquençage ONT afin de s'assurer que les protocoles d'extraction et d'amplification sont bien conciliables avec un séquençage long-read. Cela permettrait aussi de s'assurer de la nature des amplicons et ainsi de valider définitivement les amorces.

Les bases de données pourraient elles aussi nécessiter certaines améliorations à plusieurs niveaux. Premièrement, elles pourraient être complétées par des séquences trouvées sur d'autres bases de données que NCBI Nucleotide telles que MIDORI2 (Leray *et al.*, 2022) ou encore BOLD (Barcode Of Life Data System). On pourrait aussi y ajouter des métadonnées propres à chaque espèce pouvant indiquer son habitat, sa provenance, à quelle saison de l'année elle est censée être présente etc. Ces bases de données enrichies seraient ainsi capables de prévoir quelles espèces sont supposées être échantillonnées quand, où et en quelle quantité, afin de pouvoir suivre la composition des communautés méroplanctoniques. En ce sens, le choix de s'appuyer sur les inventaires créés par la Station Biologique de Roscoff, c'est-à-dire une base de données locale, est questionnable. En effet, si l'on veut à terme déployer ces outils moléculaires dans des eaux éloignées géographiquement, les bases de données créées durant cette étude seront peut-être moins pertinentes. Enfin, ces bases de données pourront être complétées par le séquençage de mitogénomes d'espèces pas encore disponibles sur les bases de données publiques une fois que le protocole expérimental sera finalisé. Cela pourrait être très intéressant étant donné que les chiffres révèlent une rareté générale des données de séquences dans les bases de données publiques. Par exemple, NCBI et BOLD contiennent une couverture totale de seulement 15,13 % de la biodiversité reconnue sur Terre pour le COI alors même qu'il constitue le gène mitochondrial le plus documenté (Kvist, 2013).

2. Intérêt des marqueurs additionnels

Comme il est observable sur les figures 23 à 26, le séquençage long-read peut désormais capturer une portion du génome plus large et donc inclure d'autres gènes plutôt qu'un seul « barcode » cible comme c'est le plus souvent le cas actuellement. Dans le cas de cette étude, on peut voir que les fragments amplifiés pourraient contenir le 12S, les sous-unités II et III du cytochrome oxydase, le CYTB mais aussi les sous-unités NADH déshydrogénase ((ND4, ND5 etc.) et les gènes ATP synthase (ATP6, ATP8 etc.) L'intérêt de ces marqueurs additionnels réside principalement dans leur pouvoir de résolution taxinomique complémentaire au COI et au 16S. Le COII présente par exemple un taux de réussite d'identification supérieur au COI chez certains groupes comme les oomycètes (Choi *et al.*, 2015). Mais il s'est aussi montré plus performant que le COI chez plusieurs groupes d'invertébrés comme les termites avec un taux d'identification correcte compris entre 91,8 % et 100 % (Monteiro *et al.*, 2025). Le CYTB reste cependant davantage intéressant pour les vertébrés tels que les poissons (Yang *et al.*, 2024). Les gènes NAD, quant à eux, présentent un intérêt particulier car ils possèdent une variabilité élevée. Les gènes ND1 et ND2 en particulier évoluent deux à trois fois plus rapidement que le COI (Jacobsen *et al.*, 2016). Cela leur confère un fort potentiel de résolution pour distinguer des populations ou des espèces très proches. À l'inverse, le gène ATP8 figure parmi les régions les plus lentement évolutives du mitogénome, ce qui limite son pouvoir discriminant à l'échelle spécifique mais le rend potentiellement utile pour des comparaisons à plus large échelle taxinomique (Jacobsen *et al.*, 2016). Le séquençage long-read présente ici un avantage décisif puisqu'il permet de récupérer l'ensemble de ces gènes à taux d'évolution et couverture taxinomique hétérogènes en une seule lecture, offrant une information génétique bien plus robuste qu'un marqueur unique. Le principal désavantage de ces marqueurs additionnels reste cependant la disponibilité restreinte des bases de données de référence, ce qui peut limiter la fiabilité de l'assignation taxinomique pour de nombreux taxons marins peu étudiés.

3. Pertinence du séquençage long-read sur de l'ADN environnemental

Enfin, l'approche séquençage long-read doit être nuancée par le fait que l'ADN environnemental est soumis à une fragmentation rapide dans le milieu naturel, ce qui limite la récupération de longs fragments. Les analyses faites sur de l'ADN environnemental se concentrent donc plus souvent sur des fragments courts allant de 100 à 600 pb environ (Jo *et al.*, 2017). Malgré cela, plusieurs études ont montré qu'il était possible d'amplifier de l'eDNA marin sur des longueurs largement supérieures à celles habituellement ciblées. Une étude menée sur le requin-tigre en Australie occidentale a confirmé que des fragments dépassant largement les tailles conventionnelles (636, 840 et 1518 pb) restaient amplifiables à partir d'eDNA marin, bien qu'avec une diminution marquée du nombre de copies au-delà de 1518 pb (West *et al.*, 2015). D'autres travaux ont démontré qu'il était possible d'amplifier des fragments mitochondriaux de 16 à 17 kb à partir d'ADN environnemental prélevé en mésocosme et en cours d'eau naturel, à l'aide d'une PCR longue portée et d'un séquençage Illumina (Deiner *et al.*, 2017). Plus récemment et grâce à un séquençage long-read ONT, une équipe a également réussi à amplifier des fragments mitochondriaux de 8 à 9,5 kb à partir d'échantillons d'eDNA de mésocosme et de terrain (Matthews *et al.*, 2026). L'ensemble de ces études prouve bien qu'il est possible d'effectuer du séquençage long-read sur des échantillons d'ADN environnemental. Cependant, cela reste plus fastidieux que de mener des études sur des fragments courts et il y a généralement davantage de précautions à prendre lors des manipulations.

Conclusion

1. Conclusion du stage

L'objectif principal de ce stage était d'optimiser le suivi du méroplancton en se basant sur une approche de metabarcoding long-read. Deux sous-objectifs se découpaient à partir de ce premier but : le premier était de construire de nouvelles bases de données de référence et le second de développer le protocole expérimental permettant de faire du séquençage long-read ONT.

L'étude effectuée durant ce stage a permis de construire de nouvelles bases de données de référence spécifiques pour 4 phylums des communautés benthiques. Il a été décidé de construire des bases de données pour les mitogénomes complets de ces espèces, allant de 85 séquences pour les polychètes à 472 séquences pour les mollusques, puis des bases de données spécifiques aux deux marqueurs sélectionnés : le 16S et le COI. Ces bases de données ont été utiles pour le choix et la conception des amorces et le seront encore davantage pour l'affiliation taxonomique. Une base de données regroupant des espèces « non cibles » regroupant 1039 séquences a aussi été créée afin de pouvoir tester la spécificité des amorces.

Cette étude a aussi permis de développer une partie du protocole expérimental permettant de faire du séquençage long-read. Concernant la conservation des échantillons, il a été décidé qu'il était préférable de les immerger dans du Nucleoprotect. Concernant la partie extraction de l'ADN, il a été décidé de sélectionner le kit MasterPure. Enfin, concernant la partie de l'amplification de l'ADN, il a été décidé de faire les PCR avec la Taq polymérase Quantabio®. Toutefois, les premières amorces testées par PCR ne s'étant pas révélées assez spécifiques, cette étude reste insuffisante pour conclure sur un outil de suivi du méroplancton finalisé. D'autres études sont nécessaires afin d'achever le protocole expérimental mais aussi d'améliorer les bases de données.

Malgré les études qu'il reste à fournir afin de mener cette étude à bout, ce stage a contribué à améliorer les approches pour l'étude du méroplancton. Cette étude pourra être utile dans un contexte académique mais elle pourra aussi l'être dans un contexte appliqué, par exemple pour étudier les impacts du parc éoliens offshore en construction au large de Roscoff³.

2. Conclusion personnelle

Ces 6 mois à la Station Biologique de Roscoff ont été une expérience très enrichissante à de nombreux égards. Du point de vue académique, réaliser une expérience pour tester des hypothèses préalablement formulées à partir de la littérature a été très formateur. J'ai adoré pouvoir mettre en application et surtout donner sens aux connaissances que j'ai acquises durant mes études (notamment sur la partie expérimentale). J'ai aussi découvert des logiciels et des plateformes informatiques auxquels je n'avais jamais été confronté avant comme Geneious Prime ou encore NCBI. Je suis aussi très reconnaissante du projet dont j'ai été chargée et des responsabilités qui m'ont été confiées, qui m'ont également permis d'affirmer ma pensée scientifique et d'être force de proposition. Finalement, les réunions auxquelles j'ai participé et les échanges que j'ai pu avoir avec les étudiants et professionnels du milieu m'ont permis d'apprendre à communiquer efficacement à l'oral, en français mais aussi en anglais. Ces échanges m'ont aussi permis de découvrir plus en détail de nombreux aspects de la biologie

³ Article Ouest France : [Parc éolien en baie de Morlaix : le collectif Nemo demande l'abandon du projet](#) (consulté le 16/08/2026).

et ainsi de mieux définir les domaines mais aussi des approches qui m'animent. Le seul petit point de frustration que je garde de ce stage est que je n'ai pas pu voir une application concrète de mon étude étant donné qu'il était à la genèse d'un sujet plus long que 6 mois.

Bibliographie

- Andruszkiewicz Allan, Elizabeth, Megan R. Shaffer, Ryan P. Kelly, et Kim Parsons. 2025. « Optimizing Target-to-Total DNA Ratio in eDNA Studies: Effects of Sampling, Preservation, and Extraction Methods on Single-Species Detection ». *PeerJ* 13: e20127. doi:10.7717/peerj.20127.
- Arndt, Allan, Carlos Marquez, Philip Lambert, et Michael J. Smith. 1996. « Molecular Phylogeny of Eastern Pacific Sea Cucumbers (Echinodermata: Holothuroidea) Based on Mitochondrial DNA Sequence ». *Molecular Phylogenetics and Evolution* 6(3): 425-37. doi:10.1006/mpev.1996.0091.
- Atienza, Sara, Magdalena Guardiola, Kim Præbel, Adrià Antich, Xavier Turon, et Owen Simon Wangenstein. 2020. « DNA Metabarcoding of Deep-Sea Sediment Communities Using COI: Community Assessment, Spatio-Temporal Patterns and Comparison with 18S rDNA ». *Diversity* 12(4): 123. doi:10.3390/d12040123.
- Baeza, J. Antonio. 2020. « Yes, We Can Use It: A Formal Test on the Accuracy of Low-Pass Nanopore Long-Read Sequencing for Mitophylogenomics and Barcoding Research Using the Caribbean Spiny Lobster *Panulirus Argus* ». *BMC Genomics* 21(1): 882. doi:10.1186/s12864-020-07292-5.
- Cahyani, Ni Kadek Dita, Aji Wahyu Anggoro, Rian Prasetya, Yuliana Fitri Syamsuni, Muhammad Danie Al Malik, Eka Maya Kurniasih, Nining Nursalim, et al. 2026. « eDNA Metabarcoding as a Promising Conservation Tool to Monitor Fish Diversity in Indonesia Marine Protected Areas ». *Environmental DNA* 8(1): e70242. doi:10.1002/edn3.70242.
- Choi, Young-Joon, Gordon Beakes, Sally Glockling, Julia Kruse, Bora Nam, Lisa Nigrelli, Sebastian Ploch, et al. 2015. « Towards a Universal Barcode of Oomycetes – a Comparison of the Cox1 and Cox2 Loci ». *Molecular Ecology Resources* 15(6): 1275-88. doi:10.1111/1755-0998.12398.
- Chwalińska, Małgorzata, Michał Karlicki, Sarah Romic, Fabrice Not, et Anna Karnkowska. 2025. « From Short to Long Reads: Enhanced Protist Diversity Profiling via Nanopore Metabarcoding ». *Metabarcoding and Metagenomics* 9: e163750. doi:10.3897/mbmg.9.163750.
- Côté, Geneviève. 2015. « Évaluation de la biodiversité marine sur substrat rocheux à l'aide du codebarre génétique et influence de deux types de substrats sur le recrutement des espèces endémiques et de l'espèce *Allochthona intestinalis* ». masters. Université du Québec à Rimouski. <https://semaphore.uqar.ca/id/eprint/1093/> (19 février 2026).
- Dannheim, Jennifer, Lena Bergström, Silvana N R Birchenough, Radosław Brzana, Arjen R Boon, Joop W P Coolen, Jean-Claude Dauvin, et al. 2020. « Benthic effects of offshore renewables: identification of knowledge gaps and urgently needed research ». *ICES Journal of Marine Science* 77(3): 1092-1108. doi:10.1093/icesjms/fsz018.
- Danovaro, Roberto, Emanuela Fanelli, Jacopo Aguzzi, David Billett, Laura Carugati, Cinzia Corinaldesi, Antonio Dell'Anno, et al. 2020. « Ecological Variables for Developing a Global Deep-Ocean Monitoring and Conservation Strategy ». *Nature Ecology & Evolution* 4(2): 181-92. doi:10.1038/s41559-019-1091-z.
- Deiner, Kristy, Mark A. Renshaw, Yiyuan Li, Brett P. Olds, David M. Lodge, et Michael E. Pfrender. 2017. « Long-range PCR Allows Sequencing of Mitochondrial Genomes from Environmental DNA » éd. Douglas Yu. *Methods in Ecology and Evolution* 8(12): 1888-98. doi:10.1111/2041-210X.12836.
- Dipper, Frances. 2022. « Chapter 4 - Open water lifestyles: marine plankton ». In *Elements of Marine Ecology* (Fifth Edition), Butterworth-Heinemann, 193-228. doi:10.1016/B978-0-08-102826-1.00005-3.
- Domingo-Bretón, Ricardo, Federico Moroni, Socorro Toxqui-Rodríguez, Álvaro Belenguer, M. Carla Piazzon, Jaime Pérez-Sánchez, et Fernando Naya-Català. 2024. « Moving Beyond Oxford Nanopore Standard Procedures: New Insights from Water and Multiple Fish Microbiomes ». *International Journal of Molecular Sciences* 25(23): 12603. doi:10.3390/ijms252312603.

- Doorenspleet, Karlijn, Lara Jansen, Saskia Oosterbroek, Pauline Kamermans, Oscar Bos, Erik Wurz, Albertinka Murk, et Reindert Nijland. 2025. « The Long and the Short of It: Nanopore-Based eDNA Metabarcoding of Marine Vertebrates Works; Sensitivity and Species-Level Assignment Depend on Amplicon Lengths ». *Molecular Ecology Resources* 25(4): e14079. doi:10.1111/1755-0998.14079.
- Egeter, Bastian, Joana Veríssimo, Manuel Lopes-Lima, Cátia Chaves, Joana Pinto, Nicoletta Riccardi, Pedro Beja, et Nuno A. Fonseca. 2022. « Speeding up the Detection of Invasive Bivalve Species Using Environmental DNA: A Nanopore and Illumina Sequencing Comparison ». *Molecular Ecology Resources* 22(6): 2232-47. doi:10.1111/1755-0998.13610.
- Ershova, E. A., R. Descoteaux, O. S. Wangensteen, K. Iken, R. R. Hopcroft, C. Smoot, J. M. Grebmeier, et B. A. Bluhm. 2019. « Diversity and Distribution of Meroplanktonic Larvae in the Pacific Arctic and Connectivity With Adult Benthic Invertebrate Communities ». *Frontiers in Marine Science* 6. doi:10.3389/fmars.2019.00490.
- Folmer, M Black, W Hoeh, R Lutz, et R Vrijenhoek. 1994. « DNA Primers for Amplification of Mitochondrial Cytochrome c Oxidase Subunit I from Diverse Metazoan Invertebrates ».
- Footitt, R. G., D. T. Lowery, H. E. L. Maw, M. J. Smirle, et G. Lushai. 2009. « Identification, Distribution, and Molecular Characterization of the Apple Aphids *Aphis Pomi* and *Aphis Spiraeicola* (Hemiptera: Aphididae: Aphidinae) ». *The Canadian Entomologist* 141(5): 478-95. doi:10.4039/n09-037.
- Furuta, Yoshikazu, Manami Kakita, et Hidenori Tanaka. 2025. « MitoCOMON: Whole Mitochondrial DNA Sequencing by Primer Design and Long Overlapping Amplicon Assembly ». *BMC Genomics* 26(1): 787. doi:10.1186/s12864-025-12010-0.
- Garland, Elizabeth D., et Cheryl Ann Zimmer. 2002. « Techniques for the Identification of Bivalve Larvae ». *Marine Ecology Progress Series* 225: 299-310. doi:10.3354/meps225299.
- Geller, J., C. Meyer, M. Parker, et H. Hawk. 2013. « Redesign of PCR Primers for Mitochondrial Cytochrome c Oxidase Subunit I for Marine Invertebrates and Application in All-taxa Biotic Surveys ». *Molecular Ecology Resources* 13(5): 851-61. doi:10.1111/1755-0998.12138.
- Girard, Elsa B., Anouk Langerak, Jamaluddin Jompa, Owen S. Wangensteen, Jan-Niklas Macher, et Willem Renema. 2022. « Mitochondrial Cytochrome Oxidase Subunit 1: A Promising Molecular Marker for Species Identification in Foraminifera ». *Frontiers in Marine Science* 9. doi:10.3389/fmars.2022.809659.
- González-Miguéns, Rubén, Àlex Gálvez-Morante, Margarita Skamnelou, Meritxell Antó, Elena Casacuberta, Daniel J Richter, Enrique Lara, et al. 2025. « A Novel Taxonomic Database for Eukaryotic Mitochondrial Cytochrome Oxidase Subunit I Gene (eKOI), with a Focus on Protists Diversity ». *Database* 2025: baaf057. doi:10.1093/database/baaf057.
- Haro, Ronie, Nikita Walunjkar, Soham Jorapur, et Claudio H. Slamovits. 2024. « Long-Read DNA Sequencing Reveals the Organization of the Mitochondrial Genome in the Early-Branching Dinoflagellate *Oxyrrhis Marina*. » *Protist* 175(6): 126071. doi:10.1016/j.protis.2024.126071.
- Hebert, Paul D. N., Alina Cywinska, Shelley L. Ball, et Jeremy R. deWaard. 2003. « Biological Identifications through DNA Barcodes ». *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270(1512): 313-21. doi:10.1098/rspb.2002.2218.
- Hendriks, Iris E., Luca A. van Duren, et Peter M. J. Herman. 2005. « Image analysis techniques: A tool for the identification of bivalve larvae? » *Journal of Sea Research* 54(2): 151-62. doi:10.1016/j.seares.2005.03.001.
- Highfield, James M., Damien Eloire, David V. P. Conway, Penelope K. Lindeque, Martin J. Attrill, et Paul J. Somerfield. 2010. « Seasonal dynamics of meroplankton assemblages at station L4 ». *Journal of Plankton Research* 32(5): 681-91. doi:10.1093/plankt/fbp139.
- Ikehara, K., et I. Y. Abdurakhmonov. 2025. *DNA Sequencing - History, Present and Future: History, Present and Future*. IntechOpen. <https://books.google.fr/books?id=11NQEQAQBAJ>.

- Jacobsen, Magnus W., Rute R. da Fonseca, Louis Bernatchez, et Michael M. Hansen. 2016. « Comparative analysis of complete mitochondrial genomes suggests that relaxed purifying selection is driving high nonsynonymous evolutionary rate of the NADH2 gene in whitefish (*Coregonus* spp.) ». *Molecular Phylogenetics and Evolution* 95: 161-70. doi:10.1016/j.ympev.2015.11.008.
- Jeunen, Gert-Jan, Eddy Dowle, Jonika Edgecombe, Ulla von Ammon, Neil J. Gemmell, et Hugh Cross. 2023. « Crabs—A Software Program to Generate Curated Reference Databases for Metabarcoding Sequencing Data ». *Molecular Ecology Resources* 23(3): 725-38. doi:10.1111/1755-0998.13741.
- Jo, Toshiaki, Hiroaki Murakami, Reiji Masuda, Masayuki K. Sakata, Satoshi Yamamoto, et Toshifumi Minamoto. 2017. « Rapid Degradation of Longer DNA Fragments Enables the Improved Estimation of Distribution and Biomass Using Environmental DNA ». *Molecular Ecology Resources* 17(6): e25-33. doi:10.1111/1755-0998.12685.
- Kelly, Ryan P., James L. O'Donnell, Natalie C. Lowell, Andrew O. Shelton, Jameal F. Samhour, Shannon M. Hennessey, Blake E. Feist, et Gregory D. Williams. 2016. « Genetic Signatures of Ecological Diversity along an Urbanization Gradient ». *PeerJ* 4: e2444. doi:10.7717/peerj.2444.
- Krehenwinkel, Henrik, Aaron Pomerantz, James B Henderson, Susan R Kennedy, Jun Ying Lim, Varun Swamy, Juan Diego Shoobridge, et al. 2019. « Nanopore Sequencing of Long Ribosomal DNA Amplicons Enables Portable and Simple Biodiversity Assessments with High Phylogenetic Resolution across Broad Taxonomic Scale ». *GigaScience* 8(5): giz006. doi:10.1093/gigascience/giz006.
- Kvist, Sebastian. 2013. « Barcoding in the dark?: A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge ». *Molecular Phylogenetics and Evolution* 69(1): 39-45. doi:10.1016/j.ympev.2013.05.012.
- Layton, Kara K. S., Erin A. Corstorphine, et Paul D. N. Hebert. 2016. « Exploring Canadian Echinoderm Diversity through DNA Barcodes ». *PLOS ONE* 11(11): e0166118. doi:10.1371/journal.pone.0166118.
- Leduc, Noémie. 2019. « Évaluation de la biodiversité des invertébrés marins dans les ports commerciaux de l'Arctique grâce à l'ADN environnemental ». <https://corpus.ulaval.ca/entities/publication/cffa38d1-94e1-498a-ae5-61a7fb8814c5> (19 février 2026).
- Leray, Matthieu, Nancy Knowlton, et Ryuji J. Machida. 2022. « MIDORI2: A Collection of Quality Controlled, Preformatted, and Regularly Updated Reference Databases for Taxonomic Assignment of Eukaryotic Mitochondrial Sequences ». *Environmental DNA* 4(4): 894-907. doi:10.1002/edn3.303.
- Leray, Matthieu, Joy Y. Yang, Christopher P. Meyer, Suzanne C. Mills, Natalia Agudelo, Vincent Ranwez, Joel T. Boehm, et Ryuji J. Machida. 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): 34. doi:10.1186/1742-9994-10-34.
- Maggini, Sara, Magnus Wulff Jacobsen, Paulina Urban, Brian Klitgaard Hansen, Jos Kielgast, Dorte Bekkevold, Ernesto Jardim, et al. 2024. « Nanopore Environmental DNA Sequencing of Catch Water for Estimating Species Composition in Demersal Bottom Trawl Fisheries ». *Environmental DNA* 6(3): e555. doi:10.1002/edn3.555.
- Marín, Alan, Takafumi Fujimoto, et Katsutoshi Arai. 2015. « The Variable 5' End of the 16S rRNA Gene as a Novel Barcoding Tool for Scallops (*Bivalvia*, *Pectinidae*) ». *Fisheries Science* 81(1): 73-81. doi:10.1007/s12562-014-0819-6.
- Martino, Francesco, Gabriele Cananzi, Irene Gregori, Tianshi Li, Anna Benvenuto, Elisa Boscari, Elisa Camatti, et al. 2025. « Linking Water to the Bottom: eDNA Study of Benthic Invertebrates and Invasive Species in the Venice Lagoon ». *Environmental DNA* 7(2): e70093. doi:10.1002/edn3.70093.
- Meyer, Christopher. 2003. « Molecular systematics of cowries (*Gastropoda*: *Cypraeidae*) and diversification patterns in the tropics ». *Biological Journal of the Linnean Society* 79(3): 401-59. doi:10.1046/j.1095-8312.2003.00197.x.

Mizuno, Hinano, Hidenori Tanaka, Yoshikazu Furuta, et Nobuhiko Muramoto. 2026. « Non-invasive Reconstruction of Complete Mitochondrial Genomes from Aquatic Environmental DNA Using PCR-free Long-read Sequencing ». *Methods in Ecology and Evolution* 17(1): 99-110. doi:10.1111/2041-210x.70209.

Monteiro, Sara Rikeley Paulino, Antonio Carvalho, Renan Rodrigues Ferreira, Rozzanna Esther C. R. Figueirêdo, Alexandre Vasconcellos, et Ricardo Koroiva. 2025. « Efficiency of the Cytochrome c Oxidase Subunit II Gene for the Delimitation of Termite Species (Blattodea: Isoptera) in the State of Paraíba, Northeastern Brazil ». *PLOS ONE* 20(9): e0328685. doi:10.1371/journal.pone.0328685.

Patin, N. V., et K. D. Goodwin. 2022. « Long-Read Sequencing Improves Recovery of Picoeukaryotic Genomes and Zooplankton Marker Genes from Marine Metagenomes » éd. Holly Bik. *mSystems* 7(6): e00595-22. doi:10.1128/msystems.00595-22.

Ren, Wenting, Peng Zhou, et Dongsheng Zhang. 2025. « Assessing the efficiency of COI primers in eDNA metabarcoding for marine metazoan biodiversity ». *Marine Environmental Research* 211: 107450. doi:10.1016/j.marenvres.2025.107450.

Reynolds. 2006. « The Ecology of Phytoplankton (1re éd.). Cambridge University Press. »

Romac Sarah. 2023. Nucleic acids extraction from single cell using MasterPure Complete DNA purification (Epicenter) V.3. doi: <https://dx.doi.org/10.17504/protocols.io.x54v9mk3qg3e/v3>

Saeedi, Hanieh, Dan Warren, et Angelika Brandt. 2022. « The Environmental Drivers of Benthic Fauna Diversity and Community Composition ». *Frontiers in Marine Science* 9. doi:10.3389/fmars.2022.804019.

Saunders, Gary W. 2008. « A DNA Barcode Examination of the Red Algal Family Dumontiaceae in Canadian Waters Reveals Substantial Cryptic Species Diversity » *Botany* 86(7): 773-89. doi:10.1139/B08-001.

Shi, Xiaojing, Yihui Jiang, Ling Cao, et Cong Zeng. 2024. « Development of Environmental DNA Metabarcoding Primers for Marine Mollusks and Comparison with Published Primers ». *BMC Ecology and Evolution* 24(1): 73. doi:10.1186/s12862-024-02265-8.

Srivathsan, Amrita, Leshon Lee, Kazutaka Katoh, Emily Hartop, Sujatha Narayanan Kutty, Johnathan Wong, Darren Yeo, et Rudolf Meier. 2021. « ONTbarcoder and MinION Barcodes Aid Biodiversity Discovery and Identification by Everyone, for Everyone ». *BMC Biology* 19(1): 217. doi:10.1186/s12915-021-01141-x.

Stephanie A. Matthews, Olivia M. Scott, Yin Cheong Aden Ip, Elizabeth A. Allan, Ryan P. Kelly. 2026. « Long-Range PCR Amplification and Nanopore Sequencing of 8-10 Kb Mitochondrial Fragments from Environmental DNA ». doi: 10.64898/2026.01.13.699327

Tyagi, P., et M. Bhide. 2020. « History of DNA Sequencing ». *Folia Veterinaria* 64(2): 66-73. doi:10.2478/fv-2020-0019.

Wangensteen, Owen S., Creu Palacín, Magdalena Guardiola, et Xavier Turon. 2018. « DNA Metabarcoding of Littoral Hard-Bottom Communities: High Diversity and Database Gaps Revealed by Two Molecular Markers ». *PeerJ* 6: e4705. doi:10.7717/peerj.4705.

Watson, Stephen C. L., Claire L. Szostek, Andrew Edwards-Jones, Betheney Wills, Gordon J. Watson, et Nicola J. Beaumont. 2025. « Assessing, Monitoring and Mitigating the Effects of Offshore Wind Farms on Biodiversity ». *Nature Reviews Biodiversity* 1(9): 581-96. doi:10.1038/s44358-025-00074-5.

West, Katrina M., et Bruce Deagle. 2025. « How Fragmented Is eDNA? A Case Study on Shark DNA in Tropical Reef Seawater ». *Environmental DNA* 7(4): e70165. doi:10.1002/edn3.70165.

Yang, Tiezhu, Chenxi Tan, Liangjie Zhao, Zhiguo Hu, Chaoqun Su, Fan Li, Yuanye Ma, et al. 2024. « The Complete Mitochondrial Genome of the *Luciocyprinus Langsoni* (Cypriniformes: Cyprinidae): Characterization, Phylogeny, and Genetic Diversity Analysis ». *Genes* 15(12): 1621. doi:10.3390/genes15121621.

Annexes

Annexe 1 : Liste des échantillons utilisés dans cette étude, avec leur code, affiliation taxonomique, fixateur, nature et masse du tissu prélevé et date d'échantillonnage.

Code	Classe	Genre & espèce	Tissu prélevé	Masse / volume	Fixateur	Date d'échantillonnage	Commentaires
mtPN1	Ophiuroidea	Ophiura albida	1 bras	<30 mg	EtOH96	28/1/2026	
mtPN2	Ophiuroidea	Ophiura albida	1 bras	<30 mg	EtOH96	28/1/2026	
mtPN3	Ophiuroidea	Amphiura filiformis	1 bras	<30 mg	EtOH96	28/1/2026	
mtPN4	Ophiuroidea	Amphiura filiformis	1 bras	<30 mg	EtOH96	28/1/2026	
mtPN5	Ophiuroidea	Ophiura albida	1 bras	<30 mg	NucleoP	28/1/2026	
mtPN6	Ophiuroidea	Ophiura albida	1 bras	<30 mg	NucleoP	28/1/2026	
mtPN7	Ophiuroidea	Amphiura filiformis	1 bras	<30 mg	NucleoP	28/1/2026	
mtPN8	Ophiuroidea	Amphiura filiformis	1 bras	<30 mg	NucleoP	28/1/2026	
mtPN9	Polycheta	Glycera tridactyla	individu entier	<30 mg	EtOH96	28/1/2026	Élimination du tube digestif
mtPN10	Polycheta	Glycera tridactyla	individu entier	<30 mg	EtOH96	28/1/2026	Élimination du tube digestif
mtPN11	Polycheta	Aponuphis bilineata	individu entier	<30 mg	EtOH96	28/1/2026	Élimination du tube digestif
mtPN12	Polycheta	Aponuphis bilineata	individu entier	<30 mg	EtOH96	28/1/2026	Élimination du tube digestif
mtPN13	Polycheta	Glycera tridactyla	individu entier	<30 mg	NucleoP	28/1/2026	Élimination du tube digestif
mtPN14	Polycheta	Glycera tridactyla	individu entier	<30 mg	NucleoP	28/1/2026	Élimination du tube digestif
mtPN15	Polycheta	Aponuphis bilineata	individu entier	<30 mg	NucleoP	28/1/2026	Élimination du tube digestif
mtPN16	Polycheta	Aponuphis bilineata	individu entier	<30 mg	NucleoP	28/1/2026	Élimination du tube digestif
mtPN17	Malacostraca	Ampelisca brevicornis	individu entier	<30 mg	EtOH96	28/1/2026	
mtPN18	Malacostraca	Ampelisca brevicornis	individu entier	<30 mg	EtOH96	28/1/2026	
mtPN19	Malacostraca	Urothoe pulchella	individu entier	<30 mg	EtOH96	28/1/2026	
mtPN20	Malacostraca	Urothoe pulchella	individu entier	<30 mg	EtOH96	28/1/2026	
mtPN21	Malacostraca	Ampelisca brevicornis	individu entier	<30 mg	NucleoP	28/1/2026	
mtPN22	Malacostraca	Ampelisca brevicornis	individu entier	<30 mg	NucleoP	28/1/2026	
mtPN23	Malacostraca	Urothoe pulchella	individu entier	<30 mg	NucleoP	28/1/2026	
mtPN24	Malacostraca	Urothoe pulchella	individu entier	<30 mg	NucleoP	28/1/2026	
mtPN25	Bivalve	Abra alba	individu entier	~200 mg	EtOH96	28/1/2026	
mtPN26	Bivalve	Abra alba	individu entier	~200 mg	EtOH96	28/1/2026	
mtPN27	Gasteropode	Nassarius reticulatus	fragment du pied	~200 mg	EtOH96	28/1/2026	
mtPN28	Gasteropode	Nassarius reticulatus	fragment du pied	~200 mg	EtOH96	28/1/2026	
mtPN29	Bivalve	Abra alba	individu entier	~200 mg	NucleoP	28/1/2026	
mtPN30	Bivalve	Abra alba	individu entier	~200 mg	NucleoP	28/1/2026	
mtPN31	Gasteropode	Nassarius reticulatus	fragment du pied	~200 mg	NucleoP	28/1/2026	
mtPN32	Gasteropode	Nassarius reticulatus	fragment du pied	~200 mg	NucleoP	28/1/2026	
mtPN33	Bivalve	Abra alba	individu entier	~200 mg	NucleoP	28/1/2026	

mtPN34	Polycheta	Glycera tridactyla	individu entier	<30 mg	NucleoP	28/1/2026
mtPN35	Malacostraca	Ampelisca brevicornis	individu entier	<30 mg	NucleoP	28/1/2026
mtPN36	Malacostraca	Urothoe pulchella	individu entier	<30 mg	NucleoP	28/1/2026
mtPN37	Ophiuroidea	Ophiura albida	1 bras	<30 mg	NucleoP	28/1/2026
mtPN38	Ophiuroidea	Amphiura filiformis	1 bras	<30 mg	NucleoP	28/1/2026
mtPN39	Gasteropode	Nassarius reticulatus	fragment du pied	50 mg	NucleoP	28/1/2026
mtPN40	Malacostraca	Corystes cassivelaunus	pince	50 mg	EtOH96	9/3/2026
mtPN41	Actinopterygii	Trisopterus luscus	muscle	50 mg	EtOH96	9/3/2026
mtPN42	Actinopterygii	Arnoglossus sp.	muscle	50 mg	EtOH96	9/3/2026
mtPN43	Malacostraca	Crangon crangon	muscle	<30 mg	EtOH96	9/3/2026
mtPN44	Gasteropode	Philine quadripartita	parapode	<30 mg	EtOH96	9/3/2026
mtPN45	Cephalopode	Loligo sp.	manteau	<30 mg	EtOH96	9/3/2026
mtPN46	Malacostraca	Maja squinado	pince	50 mg	EtOH96	9/3/2026
mtPN47	Actinopterygii	Callionymus lyra	muscle	50 mg	EtOH96	9/3/2026
mtPN48	Actinopterygii	Mullus surmuletus	muscle	50 mg	EtOH96	9/3/2026
mtPN49	Actinopterygii	Buglossidium sp.	muscle	50 mg	EtOH96	9/3/2026
mtPN50	Ctenophora	Nuda sp.	individu entier	<30 mg	EtOH96	9/3/2026
mtPN51	Malacostraca	Macropodia	pince	<30 mg	EtOH96	9/3/2026
mtPN53	Malacostraca	Corystes cassivelaunus	pince	~200 mg	NucleoP	9/3/2026
mtPN54	Actinopterygii	Trisopterus luscus	muscle	50 mg	NucleoP	9/3/2026
mtPN55	Actinopterygii	Arnoglossus sp.	muscle	50 mg	NucleoP	9/3/2026
mtPN56	Actinopterygii	Buglossidium sp.	muscle	50 mg	NucleoP	9/3/2026
mtPN57	Actinopterygii	Callionymus lyra	muscle	50 mg	NucleoP	9/3/2026
mtPN58	Actinopterygii	Mullus surmuletus	muscle	50 mg	NucleoP	9/3/2026
mtPN59	Malacostraca	Crangon crangon	muscle	<30 mg	NucleoP	9/3/2026
mtPN60	Gasteropode	Philine quadripartita	parapode	<30 mg	NucleoP	9/3/2026
mtPN61	Cephalopode	Loligo sp.	manteau	<30 mg	NucleoP	9/3/2026
mtPN62	Ctenophora	Nuda sp.	individu entier	<30 mg	NucleoP	9/3/2026
mtPN63	Malacostraca	Macropodia	pince	<30 mg	NucleoP	9/3/2026
mtPN64	Malacostraca	Maja squinado	pince	50 mg	NucleoP	9/3/2026