



HAL
open science

La surveillance de la biodiversité ichthyologique côtière par l'ADN environnemental. Evaluation de l'outil et de sa complémentarité aux comptages visuels en plongée.

Rapport final Septembre 2021

Anaïs Rey, Frédérique Viard, Erwan Corre, Anne Lizé, Laurent Levêque, Anna Le Joncour, Marine Delesalle, Alexandre Robert, Pierre Thiriet

► To cite this version:

Anaïs Rey, Frédérique Viard, Erwan Corre, Anne Lizé, Laurent Levêque, et al.. La surveillance de la biodiversité ichthyologique côtière par l'ADN environnemental. Evaluation de l'outil et de sa complémentarité aux comptages visuels en plongée. Rapport final Septembre 2021. PatriNat (OFB-CNRS-MNHN). 2021, 85 p. + 11 p. d'annexes. mnhn-04168495

HAL Id: mnhn-04168495

<https://mnhn.hal.science/mnhn-04168495>

Submitted on 21 Jul 2023

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

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

La surveillance de la biodiversité ichthyologique côtière par l'ADN environnemental : Evaluation de l'outil et de sa complémentarité aux comptages visuels en plongée

— Rapport final, septembre 2021 —



CNRS UPMC INSU
Station Biologique
Roscoff



MUSÉUM
NATIONAL D'HISTOIRE NATURELLE
Station Marine Dinard

Anaïs REY^{1,2,3}, Frédérique VIARD^{2,4}, Erwan CORRE³, Anne LIZE^{1,5}, Laurent LEVEQUE³, Anna LE JONCOUR¹, Marine DELESALLE¹, Alexandre ROBERT^{1,5} et Pierre THIRIET¹

¹ UMS PatriNat (OFB, CNRS, MNHN)

² UMR 7144 CNRS SU, Laboratoire AD2M, Station Biologique de Roscoff

³ FR 2424 CNRS SU, Station Biologique de Roscoff

⁴ ISEM, Université Montpellier, CNRS, IRD, Montpellier, France

⁵ MNHN-Station Marine de Dinard

Contacts : anaïs.rey@mnhn.fr ; frederique.viard@umontpellier.fr ; pierre.thiriet@mnhn.fr

Préambule :

Le présent document dresse un bilan final du projet « La surveillance des poissons côtiers par ADN environnemental ». Ce projet de développement méthodologique s'inscrit dans le cadre de la mise en œuvre de la Directive Cadre Stratégie pour le Milieu Marin (DCSMM). Ce projet court d'avril 2018 à avril 2021. Il est financé par le MTES au MNHN, tutelle de l'UMS PatriNat (titulaire des mandats RT et RS D1 PC côtiers de roche), *via* (1) une subvention spécifique (2018-2019) dans le cadre du programme 113 « paysages, eau, biodiversité » - action 07 « gestion des milieux et biodiversité » - sous-action 19 « espaces marins », et (2) la convention annuelle MTES-MNHN 2019, 2020 et 2021. Le financement du projet inclus également une part importante d'autofinancements des cinq structures collaborant pour le projet : UMS PatriNat (OFB, CNRS, MNHN), UMR 7144 CNRS SU « AD2M », FR 2424 CNRS SU, ISEM et MNHN-Station Marine de Dinard.

Citation du document :

Rey A., Viard F., Corre E., Lizé A., Lévêque L., Le Joncour A., Delesalle M., Robert A., et Thiriet P. (2021). *La surveillance de la biodiversité ichthyologique côtière par l'ADN environnemental. Evaluation de l'outil et de sa complémentarité aux comptages visuels en plongée. Rapport final Septembre 2021.* Rapport UMS PatriNat, 85 pages + annexes.

Crédit photo de la page de couverture :

Pierre Thiriet, UMS PatriNat

Sommaire

1	Résumé étendu	5
2	Introduction	5
2.1	Contexte : les programmes de surveillance DCSMM	14
2.2	Les comptages visuels en plongée, l'approche conventionnelle en milieux rocheux	17
2.3	L'analyse de l'ADN environnemental, une approche complémentaire ?	19
2.3.1	Qu'est-ce que l'ADN environnemental ?	19
2.3.2	La technique de metabarcoding sur ADNe comme outil de surveillance passive	20
2.3.3	La technique de droplet digital PCR sur ADNe comme outil de surveillance active	21
2.4	Hypothèses de travail concernant les complémentarités entre méthodes	22
2.5	Objectifs de l'étude	25
3	Méthodes	26
3.1	Collecte des données <i>in-situ</i> ADNe et UVC	26
3.2	Enrichissement de la base de référence moléculaire pour le metabarcoding	28
3.3	Acquisition des données moléculaires issues d'ADNe	29
3.3.1	Extraction ADNe	29
3.3.2	Séquençage haut-débit	29
3.3.3	Analyses bioinformatiques des données de metabarcoding	29
3.3.4	Analyses par ddPCR	31
3.4	Analyses statistiques de comparaison entre UVC et metabarcoding	32
3.5	Création d'un « workflow » d'analyses de différentes facettes de diversité à partir de données de metabarcoding	33
4	Résultats	34
4.1	Enrichissement des bases de séquences	34
4.2	Synthèse de l'analyse bioinformatique des données metabarcoding	35
4.2.1	Séquences conservées pour les analyses statistiques à l'issue des étapes de filtration bioinformatiques	35
4.2.2	Nombre de taxons et d'OTUs	35
4.2.3	Comparaison de la résolution taxinomique des deux marqueurs	37
4.3	La surveillance passive des communautés de poissons côtiers par metabarcoding	38
4.3.1	Comparaison des résultats des analyses par metabarcoding avec ou sans assignation taxinomique, et sélection du jeu de données pour la comparaison avec les données issues des comptages visuels	38
4.3.2	Analyse comparative des méthodes de détection	41
4.4	La surveillance active par ddPCR	54
4.4.1	Sélection des espèces cibles	54
4.4.2	Evaluation de la sensibilité de détection de la ddPCR par rapport au metabarcoding	55

4.4.3	Comparaison des patrons quantitatifs _____	57
4.4.4	Comparaison des patrons spatiaux-temporels des espèces cibles _____	59
4.4.5	Distinction des lignées Atlantique et Méditerranéenne du bar _____	60
4.5	Les premiers pas du « workflow » _____	61
5	Discussion _____	64
5.1	Bilans de l'étude pilote _____	64
5.1.1	Le metabarcoding sur ADNe a permis une analyse détaillée de la diversité ichtyologique de la baie de Morlaix _____	64
5.1.2	La ddPCR et le metabarcoding sont des outils complémentaires afin de surveiller la faune ichtyologique _____	68
5.1.3	Le travail de démocratisation de l'approche « ADNe » reste nécessaire _____	70
5.2	Recommandations pour la poursuite du développement de la surveillance ADNe _____	71
5.2.1	Poursuite du développement de la surveillance passive par metabarcoding : verrous et points clés _____	71
5.2.2	Poursuite du développement de la surveillance active par ddPCR : verrous et points clés _____	73
5.2.3	Poursuite du développement de l'estimation d'abondance à partir d'ADNe _____	74
5.2.4	Extension de l'utilisation de l'ADNe pour la surveillance de la biodiversité marine et d'autres composantes biocénétiques et d'autres politiques publiques _____	75
5.3	Pistes de développement d'une surveillance combinant ADNe et approches conventionnelles. _____	76
6	Conclusion _____	79
7	Bibliographie _____	80
8	Annexes _____	86

1 — Résumé étendu —

Le développement des programmes de surveillance des poissons côtiers pour la DCSMM

Afin de répondre aux enjeux de la Directive Cadre Stratégie pour le Milieu Marin (DCSMM) visant l'atteinte et/ou le maintien du Bon Etat Ecologique des écosystèmes marins d'Europe, des programmes de surveillance doivent être mis en place. Concernant les poissons des milieux côtiers démersaux de fonds rocheux, les dispositifs de surveillance qui sont en cours de développement le long des 3 façades d'Atlantique NE (réseau POCOROCH, Rey et al., 2021) sont basés sur les comptages visuels en plongée (UVC). Cette approche s'avère très efficace pour surveiller les poissons « necto-benthiques », *i.e.* faciles à détecter en plongée car nageant proche du fond. Les UVC permettent en effet d'évaluer les abondances, biomasses et spectres de tailles de ces espèces, et ainsi peuvent renseigner, pour ces espèces, le descripteur 1 (D1 : état des populations ; à l'exception de D1C1 : prises accessoires). Cependant, les UVC sont moins efficaces pour d'autres catégories d'espèces dont les fréquences d'observations en plongée sont trop faibles. Il s'agit principalement des espèces « crypto-benthiques », *i.e.* restant immobiles, cachées ou camouflées ; les espèces manifestant un comportement fuyant envers les plongeurs ; les espèces « benthopélagiques », *i.e.* fréquentant alternativement les milieux démersaux et pélagiques ; les espèces caractérisées par de faibles abondances entraînant donc une très faible probabilité d'occurrence à l'échelle des modestes surfaces échantillonnées en plongée. Il en résulte que les UVC sont peu adaptés à l'évaluation du D1 pour un certain nombre d'espèces, et de surcroît sont peu adaptés à l'évaluation de l'ensemble du peuplement via le Descripteur 4 (diversité, composition et structure des guildes trophiques, dont les poissons).

Dans le but d'échantillonner au mieux la diversité des poissons côtiers des fonds rocheux (pour *in fine* mieux évaluer l'état écologique de cette composante au travers du D1 comme du D4), l'étude de l'ADN environnemental (ADNe ; ADN extrait à partir d'un échantillon de l'environnement tel que de l'eau) apparaît très prometteuse en tant que méthode complémentaire aux UVC. En effet, de récentes études ont démontré l'intérêt d'utiliser l'ADNe pour surveiller les peuplements de poissons en milieux dulçaquicoles dans le cadre de la Directive Cadre sur l'Eau. C'est une méthode non-invasive et non destructrice, qui couplée à l'analyse par séquençage haut-débit, permet d'analyser facilement un grand nombre d'échantillons. L'ADNe permettrait donc, dans un plus grand nombre de site et/ou à plus haute fréquence, d'évaluer la diversité du peuplement (D4) et la présence, voire l'abondance, de certaines espèces cibles (D1). Cela reste néanmoins complémentaires aux UVC, car l'ADNe ne permet pas d'échantillonner certaines métriques indispensables à l'évaluation de certains critères et descripteurs (*e.g.* le spectre de taille pour le D1C3).

Ce projet vise donc à tester l'efficacité de l'utilisation de l'ADNe, et évaluer ses complémentarités par rapport aux méthodes de comptages visuels *in situ*, pour *in fine* déterminer la pertinence de

l'établissement d'un dispositif de surveillance utilisant l'ADNe, en complément des comptages visuels en plongée.

Le suivi des poissons côtiers de fonds rocheux en baie de Morlaix comme cas d'étude

La baie de Morlaix a été échantillonnée au Printemps et en Automne 2018, par UVC et ADNe. Sur les 8 sites échantillonnés en comptages visuels (dans le cadre du programme POCOROCH, voir Le Joncour et Thiriet 2020), 4 situés le long d'un gradient côte-large ont été sélectionnés pour réaliser les prélèvements d'eau pour les analyses ADNe (Figure 1). Afin d'évaluer l'impact de la stratégie d'échantillonnage des prélèvements d'eau sur les assemblages ichtyologiques observés, 3 types de prélèvement ont été réalisés : des prélèvements de 2L en point fixe au-dessus du fond (n = 24), des prélèvements de 2L en point fixe en surface (n = 24) et des prélèvements de 30L en surface en point mobile autour du site échantillonné (n = 16). Les échantillons d'ADNe ont été analysés par metabarcoding et par ddPCR. La première technique vise à décrire la composition des assemblages ichtyologiques. La seconde a pour objectif de détecter et quantifier l'abondance d'un nombre restreint d'espèces cibles.

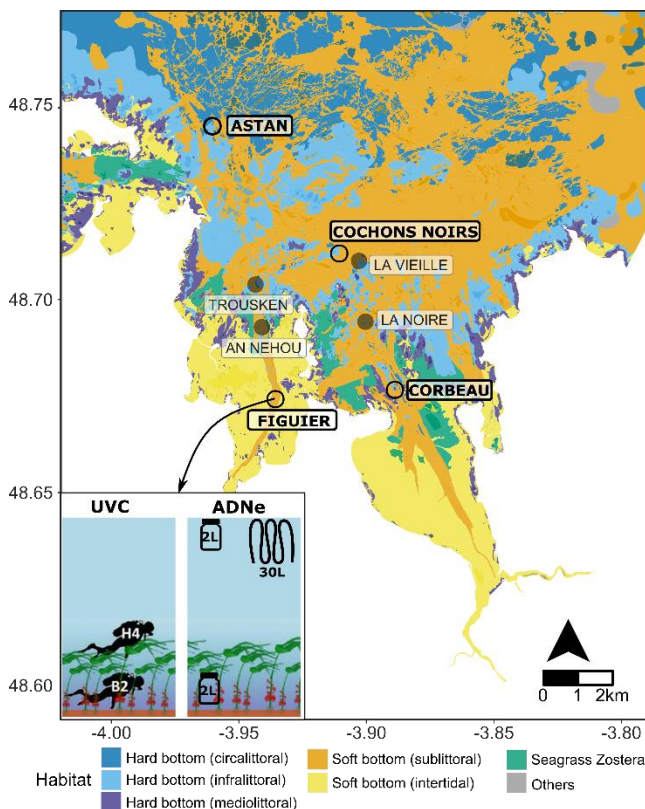


Figure 1. Carte d'échantillonnage de la baie de Morlaix indiquant la localisation des 8 sites échantillonnés en plongée (« UVC Baie ») dont 4 sites (en gras) ont également fait l'objet d'étude d'ADNe (« UVC 24h »). Les couleurs représentent les différents habitats benthiques de la baie et permettent d'évaluer la proximité du site d'habitat rocheux échantillonné avec les autres habitats. L'encadré représente le protocole d'échantillonnage simplifié des UVC et des 3 types de prélèvements d'eau pour l'ADNe.

Acquisition des données moléculaires issues du metabarcoding et de la ddPCR et création de la base de référence pour le metabarcoding

Pour le metabarcoding, les échantillons ont été amplifiés avec 2 couples d'amorces ciblant des régions courtes des marqueurs mitochondriaux 12S et 16S ; les amplicons ont ensuite été séquencés par séquençage haut-débit. Pour améliorer les assignations taxinomiques obtenues par metabarcoding, les bases de données publiques ont été enrichies grâce à la production de séquence de référence locales d'espèces fréquentant la baie de Morlaix. Une série d'étapes bioinformatiques a ensuite permis de 1) nettoyer le jeu de données, 2) assigner les séquences à un taxon grâce à la comparaison avec la base de référence créée en amont et 3) regrouper les séquences en OTU (Operational Taxonomic Unit) afin de

réaliser une analyse de diversité sans les biais engendrés par l'absence de séquences de référence pour certains taxons.

Pour la ddPCR, une première étape a consisté à sélectionner des espèces cibles en se basant sur une série de critères (faisabilité technique et intérêts des espèces). Des amorces et sondes spécifiques aux 6 espèces cibles sélectionnées ont été créés et validés (*i.e.* spécificité des sondes confirmée). Enfin, les analyses ddPCR des ADNe issues des prélèvements d'eau 2L (surface et fond) ont été effectués en routine pour chaque espèce cible.

Le metabarcoding identifie de nombreux taxons, grâce à une complémentarité entre marqueurs, cependant avec une plus grande efficacité du marqueur 12S

Sur l'ensemble des 64 échantillons, 77 taxons (dont 50 identifiés à l'espèce) et 29 taxons (dont 12 identifiés à l'espèce) ont pu être identifiés par metabarcoding, respectivement avec les marqueurs 12S et 16S. **La résolution des deux marqueurs est très différente** : alors que 51% des séquences sont assignées à l'espèce avec le marqueur 12S, seulement 27% le sont avec le marqueur 16S. Ce dernier permet néanmoins d'identifier certaines espèces non identifiées avec le marqueur 12S dont *Lepadogaster candolii*, *Mullus surmuletus*, *Trisopterus minutus* et *Callionymus lyra*. **Ceci montre l'intérêt de combiner différents marqueurs pour réaliser du metabarcoding. Néanmoins, si un choix devait être fait, le marqueur 12S serait à privilégier.**

S'affranchir de l'assignation taxinomique pour une meilleure comparaison des approches

La comparaison entre données issues du metabarcoding et données issues des comptages visuels a été réalisée en utilisant les listes et distributions d'OTUs issus du metabarcoding (et non pas à partir des seules séquences assignées à des taxons). En effet, les biais potentiels dus aux lacunes des bases de données (*i.e.*, absence de séquences de référence) sont absents en utilisant les OTUs. Par ailleurs, **la grande majorité (94% pour 12S et 92% pour 16S) des espèces identifiées est associée à un unique OTU**, permettant *in fine* d'identifier les espèces d'intérêt. De plus, **les analyses basées sur les assignations taxinomiques ou sur la diversité moléculaire (OTUs) montrent des résultats très proches en termes de diversité β** (*i.e.* les variations spatiales et temporelles observées sont similaires entre les deux approches). Il est à noter que trois espèces avec le 12S et une espèce avec le 16S sont néanmoins associées à deux OTUs. L'une d'entre elle est *Dicentrarchus labrax* pour laquelle ce résultat correspond à la présence de deux lignées évolutives bien connues impliquées dans la divergence entre le bar atlantique et le loup méditerranéen. Dans certains cas, l'utilisation de l'ADN pourrait permettre de suivre différentes lignées d'une même espèce, ce qui reste impossible avec les suivis visuels à moins de la présence d'un caractère visuel identifiable.

Accroître l'effort d'échantillonnage quelle que soit la stratégie d'échantillonnage d'eau ou la méthode (UVC ou ADNe)

L'objectif de cette présente étude était de comparer les méthodes d'observations et non pas de présenter une vision exhaustive de la communauté des poissons rocheux. Cependant quelle que soit la stratégie d'échantillonnage d'eau, les efforts d'échantillonnage devraient être accrus si l'on souhaite obtenir un inventaire plus proche de l'exhaustivité. Il est à noter par ailleurs que la diversité des OTUs est similaire quelle que soit la stratégie d'échantillonnage d'eau. Les patrons de diversité β sont aussi largement concordant entre ces types de prélèvements. Néanmoins, prélever de plus grand volume sur de plus grandes surfaces permet de détecter des taxons associés à des habitats autres que ceux ciblés initialement (habitats rocheux). **Ainsi, si l'objectif est d'identifier des taxons associés aux habitats rocheux dans des baies sans forte stratification verticale des masses d'eau (comme dans cette étude), la stratégie d'échantillonnage (volume, surface couverte ou profondeur d'échantillonnage) peut être choisie en fonction des moyens disponibles par les opérateurs. Mais de plus grands volumes et surfaces d'échantillonnage sont à privilégier pour identifier des taxons ayant des préférences d'habitats diversifiées, en particulier les poissons pélagiques.** Dans tous les cas, un accroissement du nombre de réplicats biologiques (nombre d'échantillons par site) est préconisé. Nos résultats obtenus en eaux côtières non stratifiées sur des fonds inférieurs à 30 m de profondeur ne peuvent donc pas être généralisés à des eaux plus profondes et/ou stratifiées. C'est un point important à considérer lors de l'élaboration d'une stratégie nationale d'échantillonnage.

Par comparaison aux UVC, le metabarcoding identifie un plus grand nombre de taxons, mais ne détecte pas certains taxons observés en plongée (i.e., faux-négatifs)

Sur l'ensemble des échantillons, 88 taxons (dont 52 espèces) ont été identifiés par metabarcoding et 38 par UVC. Le metabarcoding identifie 23 taxons observés en plongée (pendant les mêmes sessions d'échantillonnage que les prélèvements d'eau, i.e., UVC 24h, ou sur d'autres sessions de plongée dans cette même baie, i.e., UVC Baie, voir *Figure 2*), y compris cinq espèces rarement observées (moins de deux fois sur 34 sessions de plongée). Néanmoins, **15 taxons (dont 11 espèces) observés en plongée, y compris des espèces fréquemment observées telles que le lieu jaune (*Pollachius pollachius*) ou la vieille commune (*Labrus bergylta*), ne sont pas identifiées par metabarcoding.** Ces faux-négatifs sont dus à l'absence de référence dans les bases de données et/ou à un manque de résolution taxinomique des marqueurs utilisés. En effet, les bases de données (cumulant les données des bases de données publiques et celles obtenues lors de cette étude) sont encore incomplètes et manquent de résolution taxinomique. Sur les 212 espèces présentes ou potentiellement présentes en Manche, 31 espèces, soit 15 %, n'ont de séquence disponible pour aucun des 2 marqueurs 12S et 16S et 69 espèces, soit 32 % n'ont pas une résolution taxinomique à l'espèce pour au moins l'un des deux marqueurs. La résolution taxinomique est définie par le fait que toutes les séquences disponibles d'une espèce sont diagnostiques et permettent de les distinguer des autres espèces. Au final,

théoriquement seulement 112 des 212 espèces, soit 53 %, pourraient être détectées et identifiées si présentes dans les échantillons d'ADNe. Il est à noter que le marqueur 12S a un caractère diagnostique plus important que le 16S (98 espèces avec des séquences diagnostiques pour le 12S *versus* 28 pour le 16S). **Il est donc recommandé de poursuivre l'effort d'enrichissement des bases de données de références ainsi que de diversifier les marqueurs utilisés pour pallier les manques d'assignation ou de résolution taxinomique.**

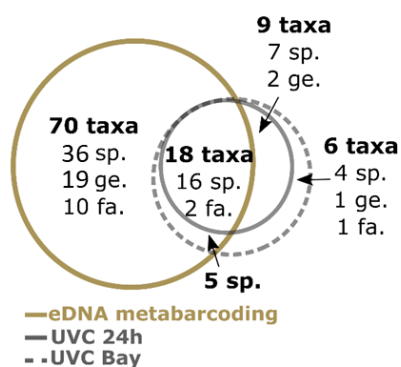


Figure 2. Diagramme de Venn comparant les taxons détectés par metabarcoding (cumul des résultats sur l'ensemble des types de prélèvements, marqueurs et saisons) et par les comptages visuels (cumulés sur UVC Baie et UVC 24h). Pour chaque aire du diagramme est noté le nombre total de taxons observés en gras et, en dessous, le nombre de taxons identifiés à la famille (fa.), au genre (ge.) et à l'espèce (sp.).

En revanche, le metabarcoding a permis de révéler 70 taxons (dont 36 espèces) non observés lors des comptages visuels, donc 32 espèces non strictement associées aux sites rocheux étudiés en plongée (22 espèces démersales d'habitats adjacents aux fonds rocheux, 4 espèces benthopélagiques et 6 espèces pélagiques). Par ailleurs, quatre espèces supplémentaires, montrant une affinité exclusive aux habitats rocheux, ont été identifiées par metabarcoding par rapport aux UVC. Cela s'explique par le fait que ces espèces sont des espèces crypto-benthiques difficilement observables en plongée. Ces résultats soulignent la capacité 1) à détecter des espèces quasiment impossibles à observer en plongée, car très petites, cachées, etc et 2) à fournir des indications de présence d'espèces non associées strictement aux habitats rocheux mais qui sont pourtant, soit bien présentes dans l'habitat rocheux échantillonné (passage fugace, migration, présence nocturne), soit présentes dans un habitat adjacent. L'ADNe apparait être une approche intéressante pour évaluer la diversité du peuplement à l'échelle du paysage côtier (regroupant l'ensemble des habitats), tout en pouvant distinguer des zones distantes de quelques kilomètres (cf section suivante).

Le metabarcoding permet une bonne résolution des patrons de diversité spatio-temporelle

Les analyses de structure des communautés observées par metabarcoding montrent des variations significatives des assemblages ichthyologiques en fonction de la saison et des sites, avec l'observation d'un gradient côte-large (Corbeau – Figuier *vs* Astan – Cochons Noirs). **Ces résultats suggèrent donc que le metabarcoding permet de distinguer des communautés à des résolutions spatiale (ici : au sein d'une même baie) et temporelle (ici : entre deux périodes d'échantillonnage séparées de 3 mois) relativement fines, et ce quel que soit le type de prélèvement.** Ces résultats illustrent la capacité potentielle du metabarcoding sur ADNe à évaluer des changements d'état

écologique (principalement la composition du peuplement D4C1), en prenant comme référence du Bon Etat Ecologique le temps (*e.g.* série temporelle) et/ou l'espace (*e.g.* zones protégées).

De la surveillance passive à la surveillance active par digital droplet PCR

Six espèces (représentant 7 taxons car les deux lignées de *Dicentrarchus labrax*, Atlantique et Méditerranéenne, ont été incluses) ont été sélectionnées pour tester l'hypothèse que la ddPCR permet une détection plus sensible des signaux ADNe que le metabarcoding pour des espèces cibles.

Par comparaison aux résultats obtenus par metabarcoding, **la ddPCR a détecté chaque espèce cible dans plus d'échantillons et sites que le metabarcoding** (Figure 3). Les échantillons où l'espèce cible a été détectée seulement avec la ddPCR ont en moyenne un nombre de copies d'ADN plus faible que dans les échantillons où l'espèce cible a été détectée simultanément par les 2 approches. Cela confirme que la ddPCR est plus sensible à la détection de signaux d'ADN faibles dans l'environnement que le metabarcoding.

Dans quelques rares cas (5% des échantillons), la ddPCR n'a pas détecté l'espèce cible alors qu'elle l'était par metabarcoding mais dans un seul ou deux des réplicats de PCR seulement. Or l'analyse par ddPCR n'a été réalisée que sur une seule prise d'ADN (au contraire du metabarcoding avec 12 prises, une pour chacun des réplicats). Ces non-détections sont donc très probablement liées à un effort d'analyse insuffisant (une seule analyse) fait avec la ddPCR, qu'il faudrait augmenter (ex. trois réplicats d'analyse par échantillon d'ADNe).

Un des atouts de la ddPCR, par rapport au metabarcoding, est de pouvoir estimer l'abondance relative d'une espèce (mesure quantitative entre échantillons traités de façon identique). De ce point de vue, il est intéressant de constater qu'une relation linéaire positive et significative entre le nombre de copies d'ADN obtenues par la ddPCR et le nombre de séquences obtenues en metabarcoding est observée pour le bar, le grand syngnathe (*S. acus*) et le centrolabre (*C. exoletus*). Dans le cas des espèces détectées plus rarement, il n'y a en revanche pas de concordance entre les distributions observées avec

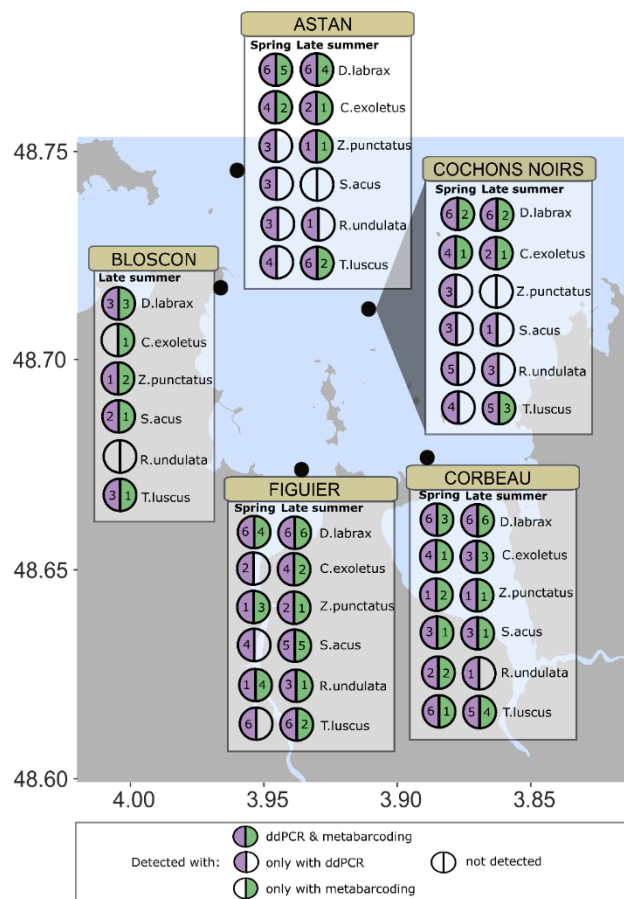


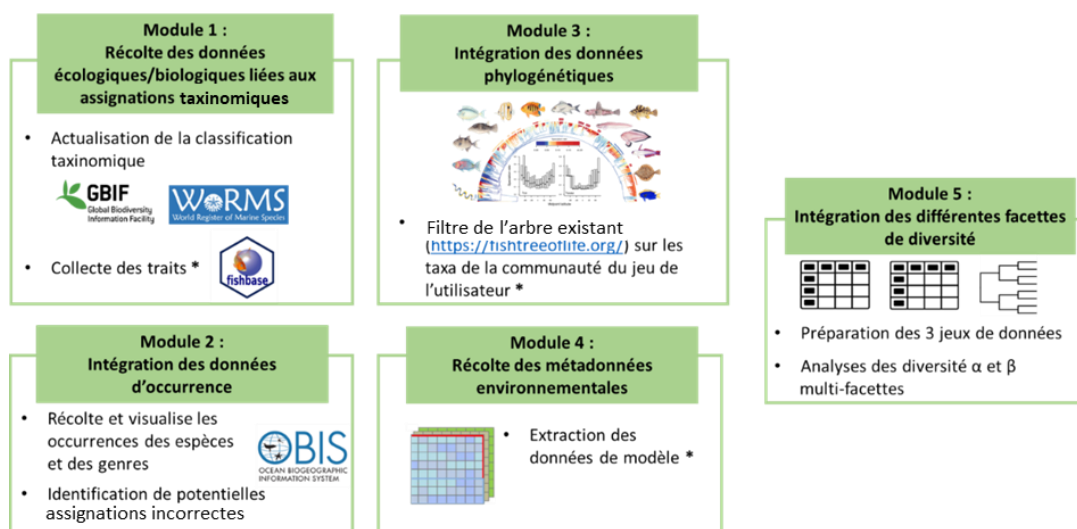
Figure 3. Evaluation de la détection des espèces cibles par les approches metabarcoding et ddPCR sur ADNe en baie de Morlaix. Le nombre d'échantillons dans lesquels l'espèce est détectée est inscrit sur la figure. Le nombre maximal d'échantillons utilisé pour le port de Blosccon par site à une saison est de trois (trois échantillons de surface) et six pour les quatre autres sites (trois échantillons de surface et trois de fond).

les deux approches. Ceci suggère que **le metabarcoding sans être une approche quantitative est à même de fournir des données semi-quantitatives pour les espèces qui ne sont pas trop rares.**

Un cas intéressant de notre étude concerne le bar puisque des sondes spécifiques des deux lignées Atlantiques et Méditerranéennes ont pu être dessinées. L'analyse par ddPCR montre une fréquence de la lignée Méditerranéenne (7%) proche de celle documentée dans d'autres études avec d'autres approches moléculaires (environ 5%), et peu variable entre sites. Le metabarcoding semble en revanche surestimer la présence de la lignée Méditerranéenne, peut-être du fait de biais d'amplification compétitive.

Un workflow d'analyse de diversité multi-facettes en cours de développement

Dans la perspective de proposer une stratégie de surveillance multi-sources de la diversité des poissons côtiers (ex. comptages UVC et ADNe), il nous a semblé pertinent de commencer à élaborer au sein de ce projet, un workflow d'analyse permettant de récolter les données de traits de vie et de phylogénie associées aux taxons détectés par une méthode d'observation et d'intégrer ces différentes facettes de diversité dans les analyses de communauté. La création de ce workflow, représenté en *Figure 4*, est en cours de développement. Une fois terminé, **il permettra de produire des analyses intégrant les diversités taxinomique, phylogénétique et fonctionnelle issues de communautés.**



* Dans le cas où l'utilisateur ne fournit pas la table de trait, l'arbre phylogénétique ou les métadonnées de l'échantillonnage

Figure 4. Représentation des modules en cours de création pour le workflow « Facettes de diversité » en cours de développement.

Conclusions : en complément des comptages UVC, l'ADNe est très pertinent pour la surveillance des poissons côtiers des habitats rocheux, et autres habitats côtiers

Cette étude pilote a fourni une première évaluation du potentiel de l'ADNe comme outil pour la surveillance des poissons côtiers des habitats rocheux. Le metabarcoding sur ADNe est un outil

complémentaire aux UVC très pertinent : il détecte un spectre plus large de diversité de l'ichtyofaune, tout en permettant de décrire des variations spatiales et temporelles à fine échelle.

Le metabarcoding sur ADNe apporte une vision intégrative des différents assemblages des poissons côtiers de roche comme illustré par la *Figure 5* : il identifie les poissons démersaux et associés à la roche mais aussi les assemblages en lien proche tels que les poissons benthopélagiques et pélagiques ou démersaux d'habitats voisins. Cet outil semble donc prometteur afin de d'évaluer la diversité d'assemblage ichtyologique pour les différents habitats côtiers, et ainsi contribuer au renseignement du descripteur 4 critère 1 de la DCSMM.

En termes de surveillance active (*i.e.*, ciblant certaines espèces d'intérêt), la ddPCR sur ADNe s'est révélée être plus sensible que le metabarcoding pour détecter des signaux d'ADNe faibles, illustrant donc son potentiel pour évaluer des espèces à occurrences rares (dont les amphihalins et elasmobranches à statut de conservation) ainsi que des espèces introduites.

Bien que des améliorations doivent être encore apportées (notamment concernant la complétude des bases de référence, le choix des marqueurs, ainsi que la stratégie d'échantillonnage), l'utilisation de l'ADNe en complément des autres méthodes d'observation permettra d'augmenter la couverture spatiale et temporelle des suivis de surveillance, favorisant une meilleure intégration des variations des communautés de poissons côtiers. En ce sens, basés sur les résultats de la présente étude, ainsi que sur l'analyse des dispositifs de suivis poissons existants en milieux côtiers et les travaux R&D en cours, nous proposons ici un concept de surveillance multi-source et multi-échelle. Notre proposition consiste à combiner plusieurs méthodes (ADNe, echosondage, comptages visuels, voire biophonie et/ou vidéo) complémentaires tant au niveau des catégories d'espèces surveillées, des métriques échantillonnées, et des couvertures et résolutions spatiales et temporelles (*Tableau 1*). Le déploiement de ces méthodes est à articuler à deux échelles spatiales :

- (1) L'observation par ADNe à l'échelle paysagère - permettant d'évaluer la composition des peuplements (metabarcoding) voire des indices d'abondance pour certaines espèces cibles (ddPCR) - vise à évaluer la tendance globale de l'état écologique à l'échelle de la façade, et dépister les zones à la dynamique particulière (pouvant faire l'objet d'études plus approfondies, *e.g.* un observatoire état-pression).

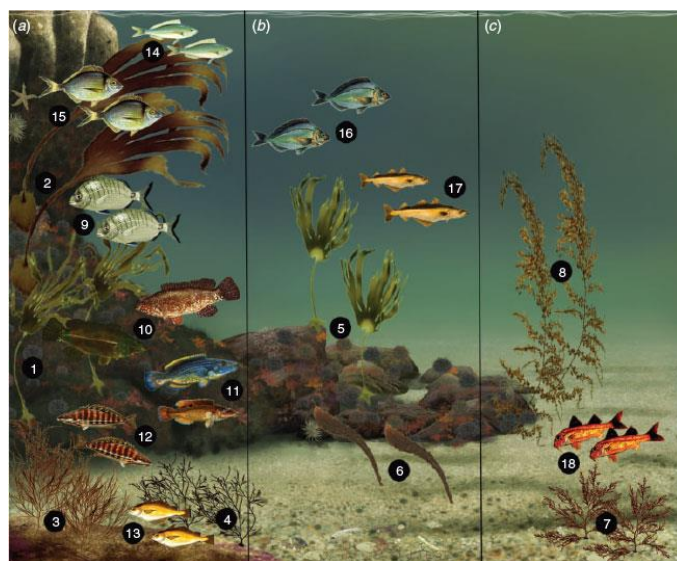


Figure 5. Illustration du paysage rocheux et de sa diversité issue de Pita et al. 2018. (a) représente la structuration verticale de l'habitat rocheux ; (b) représente la structuration horizontale de l'habitat rocheux et (c) représente l'habitat sableux avoisinant le récif

- (2) Les observatoires état-pression (*e.g.* à l'intérieur et autour d'une AMP, ou un parc éolien, à graduelle distance d'une zone urbanisée) combinent ADNe et comptages visuels (et autres méthodes selon l'habitat). Ces dernières méthodes dites conventionnelles, sont plus « lourdes » à mettre en œuvre mais apportent des métriques supplémentaires nécessaires à l'évaluation de certaines relations état-pression (et certains descripteurs et critères DCSMM).

Cette proposition de stratégie de surveillance des poissons côtiers, combinant approches conventionnelles propres à chaque habitat côtier, et approche inter-habitat telle que l'ADNe devra s'insérer dans une approche écosystémique plus globale, à discuter collégalement avec les autres RT/RS DCSMM en charge du suivi de composantes écosystémiques. Nous devons chercher une stratégie globale garantissant la cohérence des suivis et évaluations des poissons entre milieux, tant sur les aspects opérationnels qu'analytiques.

Tableau 1. Descripteurs et critères pouvant potentiellement être renseignés par les différentes méthodes d'observations proposées dans le concept multi-source/multi-échelle

	Echelle spatiale du dispositif	Milieux côtiers surveillés	Types d'espèces	Type d'évaluation	Descripteurs DCSMM potentiellement renseignés
Metabarcoding sur ADNe	« Screening de façade »	Tous les milieux (SP1 à 3)	Toutes les espèces	Tendances des diversités taxonomique, fonctionnelle et phylogénétiques (D4C1) Tendance de la diversité de groupes d'espèces fondés sur leur sensibilité aux pressions anthropiques Evolution des aires biogéographiques de distribution, par patrons d'occurrences (D1C4 et D2C2)	D1C4 D4C1 D2C2 + apport de compléments pour D1C2, D1C5, D4C2, D2C1, D2C3
ddPCR sur ADNe	« Screening de façade »	Tous les milieux (SP1 à 3)	Quelques espèces ciblées : e.g. espèces patrimoniales, amphihalins, élasmobranches à statut de conservation, ...	Tendance des abondances spécifiques (D1C2) Evolution des aires biogéographiques de distribution, par patrons d'abondances/biomasses (D1C4 et D2C2)	D1C2 D1C4 D2C2 + apport de compléments pour D1C5
Comptages visuels en plongée	Observatoires Etat-Pression	Milieux rocheux et habitats biogéniques (SP1)	Nectobenthiques	Indice de diversité d'abondance (D1C2) et évolution de la structure démographique (D1C3) Caractérisation de l'habitat pour établir les liens état-pressions (D1C5) Tendances des abondances des peuplements et structure en taille (D4C2 et D4C3)	D1C2 D1C3 D1C5 D4C2 D4C3 + apport de compléments pour D1C4 et D4C1
Echosondage	Observatoires Etat-Pression	Milieu pélagiques (SP3) Et espèces benthopélagiques en liens avec les différents habitats démersaux (SP1 et SP2)	Groupes d'écophases, basés sur la position dans la colonne d'eau, la taille individuelle et du comportement social/spatial (individus isolés vs agrégation vs bancs)	Tendances des abondances des peuplements et structure en taille (D4C2 et D4C3) Sous réserve de développement méthodologique : Indice de diversité d'abondance (D1C2) et évolution de la structure démographique (D1C3) Caractérisation de l'habitat pour établir les liens état-pressions (D1C5)	D4C2 et D4C3 Sous réserve de développement méthodo, D1C2, D1C3 et D1C5

2 Introduction

2.1 Contexte : les programmes de surveillance DCSMM

La Directive Cadre Stratégie pour le Milieu Marin (DCSMM) vise l'atteinte et/ou le maintien du Bon Etat Ecologique (BEE) des écosystèmes marins d'Europe. Pour sa mise en œuvre en France métropolitaine, les Plans d'Action pour le Milieu Marin (PAMM) prévoient cinq actions, dont le développement et la mise en œuvre de programmes de surveillance (4ème action adoptée dans chaque sous-région marine début 2015). La surveillance a pour finalité de collecter des données nécessaires à l'évaluation de l'état écologique du milieu marin et l'atteinte des objectifs environnementaux. Au total, 13 programmes de surveillance ont été retenus pour couvrir l'ensemble des composantes biocénotiques et des pressions pesant sur les écosystèmes marins, dont les Poissons et Céphalopodes (PC).

Les poissons osseux et cartilagineux représentent un groupe faunistique déterminant pour la structure et le fonctionnement des écosystèmes marins, notamment par leur diversité spécifique (plus de 1300 espèces sur l'Atlantique Nord-Est et la Méditerranée), la diversité des niches occupées par leurs différents stades ontogéniques (de la larve à l'adulte), leur abondance et leurs rôles trophiques (consommateurs primaires à prédateurs supérieurs). Par la diversité des niches écologiques qu'ils occupent, les poissons, à l'instar d'autres groupes faunistiques permettent aussi de caractériser le BEE dans tous les milieux marins. Les céphalopodes, dont la diversité spécifique est plus réduite dans les eaux françaises métropolitaines (une quarantaine d'espèces), ont été associés au groupe des poissons en raison de leur large mobilité et parce qu'ils occupent les mêmes habitats. Le suivi des céphalopodes n'est pas une priorité mais ceux-ci peuvent être échantillonnés en même temps que les poissons.

Le programme de surveillance D1-PC a pour fonction de collecter cycliquement les données nécessaires pour :

- Développer, réviser, compléter les indicateurs DCSMM du BEE des communautés de PC, relatifs aux descripteurs « Biodiversité », « Espèces non indigènes » et « Réseaux trophiques » (modifiés par la décision 2017/848 de l'UE) (Définition des Descripteurs en **Encadré 1**) ;
- Évaluer l'atteinte du BEE des communautés de PC, en renseignant les critères et indicateurs DCSMM du BEE ;
- Évaluer les impacts sur l'état écologique des pressions induites par les activités humaines ;
- Contribuer à l'évaluation de la réalisation des objectifs environnementaux définis dans les PAMM

Encadré 1 : Définition des descripteurs (D) et critères (C) du Bon Etat Ecologique qui doivent être renseignés grâce aux données des programmes de surveillance Poissons et Céphalopodes.

Descripteur 1, Biodiversité : La diversité biologique est conservée. La qualité des habitats et leur nombre, ainsi que la distribution et l'abondance des espèces sont adaptées aux conditions physiographiques, géographiques et climatiques existantes. Le descripteur 1 s'évalue selon 5 critères qui doivent être renseignés pour une liste d'espèces représentative¹.

- ❖ **D1C1 Taux de mortalité par captures accidentelles** : Le taux de mortalité par espèce dû aux captures accidentelles est inférieur au niveau susceptible de constituer une menace pour l'espèce, de sorte que la viabilité à long terme de celle-ci est assurée.
- ❖ **D1C2 Abondance des populations** : Les pressions anthropiques n'ont pas d'effets néfastes sur l'abondance des populations des espèces concernées, de sorte que la viabilité à long terme de ces populations est garantie.
- ❖ **D1C3 Caractéristiques démographiques des populations** : Les caractéristiques démographiques (par exemple structure par taille ou par âge, répartition par sexe, taux de fécondité, taux de survie,) des populations des espèces témoignent d'une population saine, qui n'est pas affectée par les pressions anthropiques.
- ❖ **D1C4 Distribution spatiale des populations** : L'aire de répartition des espèces et, le cas échéant, leur schéma de répartition dans ladite aire, est conforme aux conditions physiographiques, géographiques et climatiques dominantes.
- ❖ **D1C5 Habitat des espèces** : L'habitat des espèces offre l'étendue et les conditions nécessaires pour permettre à celles-ci d'accomplir les différentes étapes de leur cycle biologique.

Descripteur 2, Espèces non-indigènes : Les espèces non indigènes introduites par le biais des activités humaines sont à des niveaux qui ne perturbent pas les écosystèmes.

- ❖ **D2C1 Espèces non indigènes nouvellement introduites** : Le nombre d'espèces non indigènes nouvellement introduites dans le milieu naturel par le biais d'activités humaines, par période d'évaluation (six ans) est réduit au maximum et, à terme, tend vers zéro.
- ❖ **D2C2 Espèces non indigènes établies** : Abondance et répartition spatiale des espèces non indigènes établies, en particulier, les espèces envahissantes, qui contribuent de manière notable aux effets néfastes sur certains groupes d'espèces ou grands types d'habitats.
- ❖ **D2C3 Effets néfastes dus à la présence d'espèces non indigènes** : Proportion du groupe d'espèces ou étendue spatiale du grand type d'habitat subissant des altérations néfastes dues à la présence d'espèces non indigènes, en particulier des espèces non indigènes envahissantes.

Descripteur 4, Ecosystème et Réseaux trophiques :

- ❖ **D4C1 Diversité des espèces de la guildes trophique** : La diversité (composition des espèces et abondance relative de celles-ci) de la guildes trophique n'est pas affectée par les pressions anthropiques.
- ❖ **D4C2 Abondance dans les guildes trophiques** : L'équilibre de l'abondance totale entre les guildes trophiques n'est pas affecté par les pressions anthropiques.
- ❖ **D4C3 Distribution des tailles de guildes trophique** : La répartition par taille des individus au sein d'une guildes trophique n'est pas affectée par les pressions anthropiques.
- ❖ **D4C4 Productivité de guildes trophique** : La productivité de la guildes trophique n'est pas affectée par les pressions anthropiques.

¹ [Arrêté national du 09/09/2019 relatif à la définition du BEE des eaux marines et aux normes méthodologiques d'évaluation](#)

Pour chacune des quatre sous-régions marines (SRM) de France métropolitaine (**Figure 1**), le programme de surveillance PC 1^{er} cycle (2014–2020) est divisé en quatre sous-programmes (**Encadré 2**) organisés selon un triple gradient environnemental : distance à la côte, types de milieux (poissons vivant en lien avec le fond - démersaux - ou poissons vivant dans la colonne d'eau - pélagiques) et nature du substrat. Les responsabilités scientifiques du développement et de la coordination de ces quatre sous-programmes et des bio-indicateurs associés ont été réparties entre l'Ifremer et l'UMS PatriNat en fonction de leurs compétences respectives² (**Encadré 2**).

Encadré 2 : Structure et contenu du programme de surveillance D1PC 2nd cycle (2021-2026)			
Sous-programme par type de milieu	Dispositifs de surveillance retenus* pour le 2nd cycle 2021-2026	Méthode de suivi et stratégie d'échantillonnage	Responsables Thématique et Surveillance
SP1 PC benthodémersaux sur substrats durs et habitats biogéniques côtiers	Aucun dispositif retenu*	Développement en cours (pêche trémail dans les prés salés, comptage visuel en plongée en milieux rocheux, ADN pour l'ensemble des milieux)	UMS PatriNat
SP2 PC benthodémersaux sur substrats meubles côtiers	Campagnes d'observation halieutiques : nourriceries (NOUR), évaluation des stocks, OBSMER	Prélèvement au chalut, embarquement à bord des navires de pêche	Ifremer
SP3 PC pélagiques en milieux côtiers	Aucun dispositif retenu*	Développement en cours (suivi par échosondage)	UMS PatriNat
SP4 PC pélagiques et benthodémersaux sur le plateau continental et au large	Campagnes halieutiques (CGFS, EVHOE, IBTS, BFTAS, OBSMER, PELGAS, PELMED)	Campagnes d'observations halieutiques, campagnes aériennes de suivi du thon rouge, observation des captures de pêche	Ifremer

* « retenu » signifie ici que le dispositif est considéré comme non-opérationnel ou opérationnel

² Le co-pilotage scientifique des SP1, et SP3 a été assuré par le MNHN de 2010 à 2017. Suite à la création de l'AFB en 2017, l'équipe MNHN en charge de la thématique DCSMM Poissons-Céphalopodes a été rattachée à l'Unité Mixte de Service Patrimoine Naturel - UMS PatriNat.

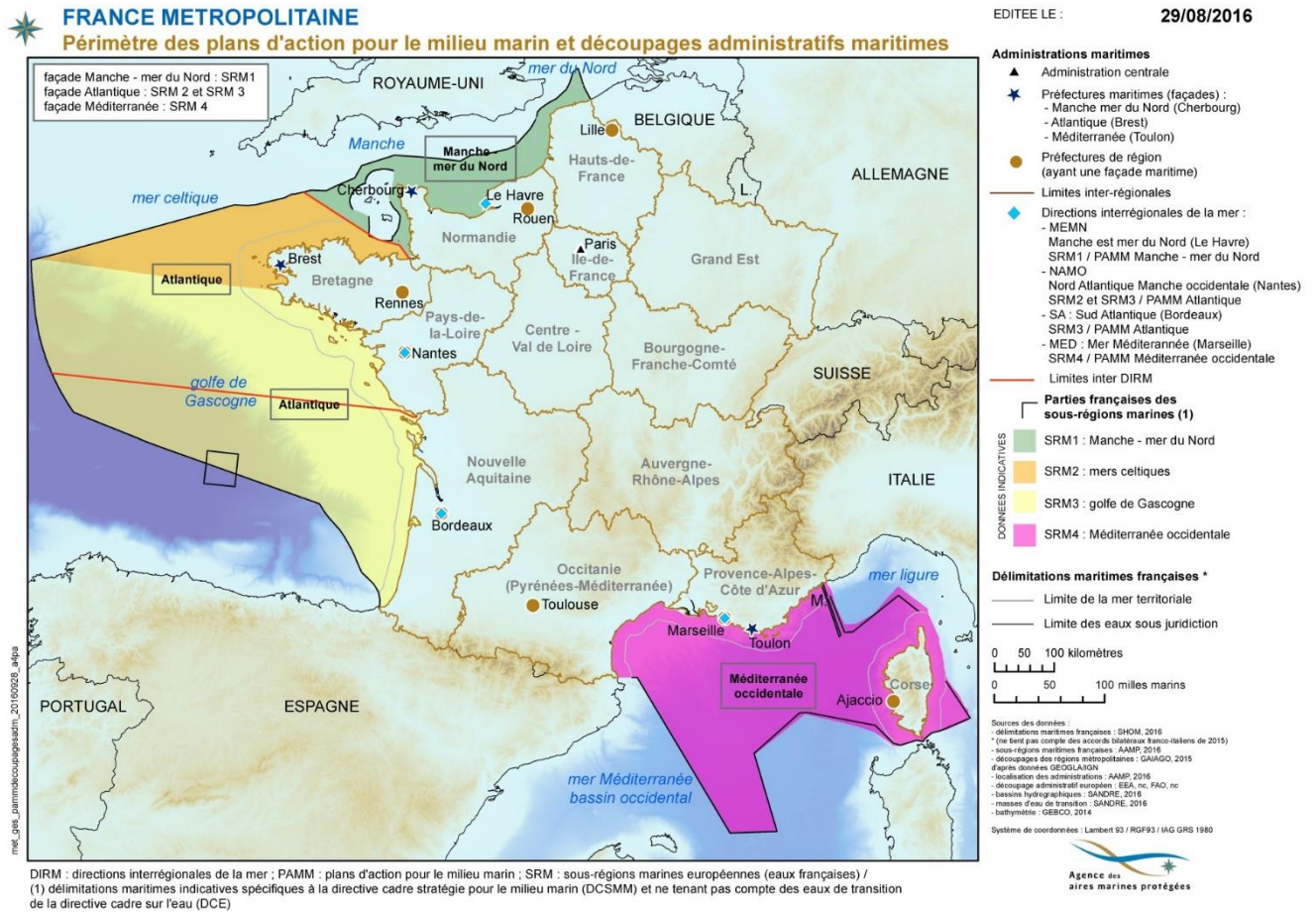


Figure 1. Les quatre sous-régions marines (SRM) de France métropolitaine

2.2 Les comptages visuels en plongée, l'approche conventionnelle en milieux rocheux

Le dispositif POCOROCH est en cours de développement pour la surveillance des poissons-céphalopodes dans milieux côtiers rocheux (SP1) des 3 façades d'Atlantique NE. Il repose sur des comptages visuels en plongées sous-marines, qui visent à estimer les abondances et tailles individuelles par espèce. De tels suivis existent également pour la façade Méditerranée Occidentale principalement dans le cadre des suivis d'Aires Marines Protégées. Ces derniers sont en cours d'intégration pour le 2nd cycle de surveillance DCSMM D1PC SP1 SRM MO.



L'approche par comptage visuels en plongée est dite conventionnelle car c'est la plus utilisée dans le monde pour les suivis de PC dans les milieux récifaux (tempérés et tropicaux) (Caldwell et al., 2016). Cette approche s'avère être très efficace pour dénombrer et mesurer la taille des poissons dits « necto-benthiques », *i.e.* ceux faciles à détecter en plongée car nageant proche du fond (catégories 3 et 4 de la Figure 2). Les observations visuelles en plongée sont toutefois moins efficaces pour les espèces

dites « benthopélagiques », *i.e.* celles fréquentant alternativement les milieux démersaux et pélagiques en manifestant une mobilité verticale (et horizontale) importante (catégorie 1 et 2 de la **Figure 2**), ainsi que pour les espèces dites « crypto-benthiques », *i.e.* celles restant immobiles et cachées (*e.g.* dans des crevasses, sous des tapis d'algues, les crampons de laminaires) ou camouflées (*e.g.* poissons plats) (catégories 5 et 6 de la **Figure 2**). Les abondances et diversités de ces espèces sont en effet quasi-systématiquement sous-évaluées par les comptages visuels à cause de la difficulté à détecter leur présence (Thiriet et al., 2016a). D'autres espèces peuvent également être difficiles à détecter : celles manifestant un comportement fuyant envers les plongeurs (souvent des espèces chassées) et celles dites « rares » ou du moins dont les faibles abondances entraînent une très faible probabilité d'occurrence à l'échelle des surfaces échantillonnées couramment en plongées (*i.e.* < 1 ha / plongée) (Prato et al., 2017).

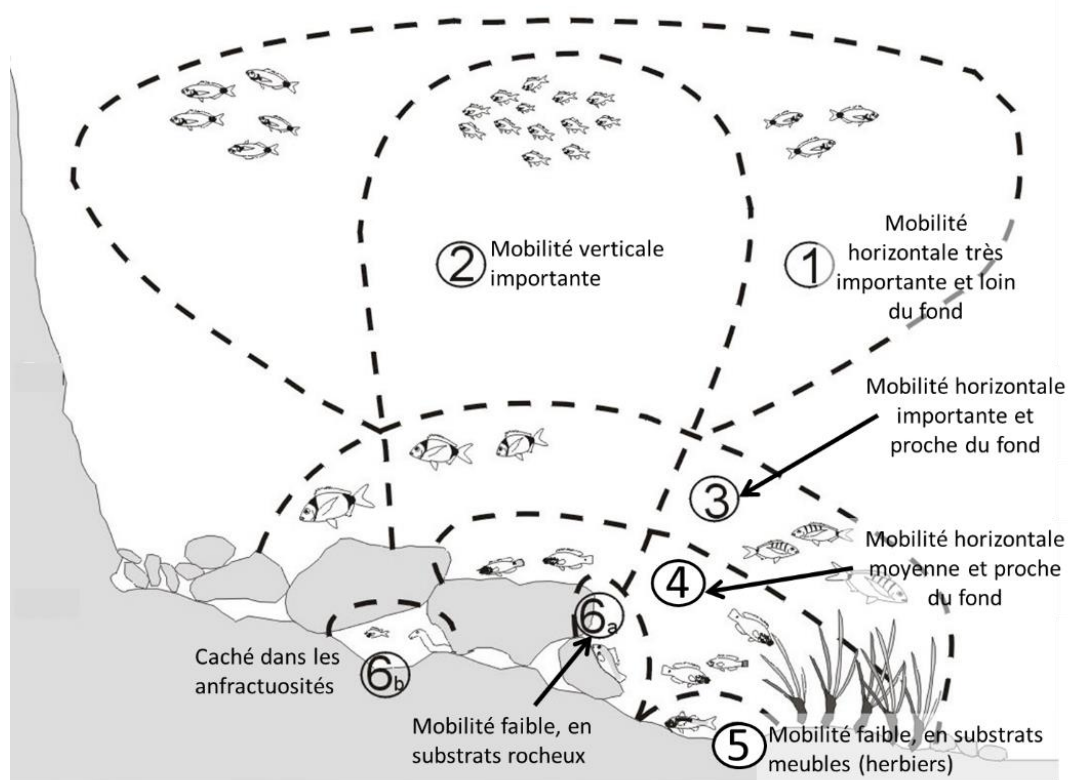


Figure 2. Les différentes catégories de poissons au regard de leur mobilité et de leur utilisation de l'habitat démersal. Schéma modifié d'après (Harmelin, 1987).

2.3 L'analyse de l'ADN environnemental, une approche complémentaire ?

2.3.1 Qu'est-ce que l'ADN environnemental ?

Bien qu'il y ait un débat dans la communauté scientifique sur la définition stricte de l'ADN environnemental (ADNe) (Lacoursière-Roussel and Deiner, 2019; Pawlowski et al., 2020), nous considérons ici la définition initialement proposée par Taberlet et al. (2012) qui définit l'ADNe comme « l'ADN qui peut être extrait d'échantillons environnementaux tels que l'eau, le sol ou l'air sans avoir isolé physiquement un groupe d'organismes ». Il rassemble donc un ensemble d'ADN extra- ou intracellulaires de nature variée. Par exemple, dans le cas des communautés ichthyologiques ciblées ici, la présence de mucus, écaille, fèces, gamètes, cellules etc. dans la colonne d'eau sont les sources principales d'ADNe. L'analyse de l'ADNe par séquençage haut-débit a une propriété intrinsèque lui donnant une valeur ajoutée sur les comptages individuels : la possibilité d'analyser un grand nombre d'échantillons (ex. nombre de sites, séries temporelles) avec un bon rapport coût/efficacité, ce que permettent rarement les protocoles en comptage visuels (nécessitant des conditions météorologiques et de turbidités particulières, un grand nombre de plongeurs etc.). L'utilisation d'ADNe est une méthode non invasive et non destructrice puisqu'elle ne nécessite pas de prélever des individus. L'analyse de l'ADN environnemental (ADNe ; ADN extrait à partir d'un échantillon de l'environnement) obtenu à partir d'un échantillon d'eau est ainsi depuis quelques années une piste prometteuse pour étudier les communautés de poissons.

L'analyse de l'ADNe peut être réalisée par plusieurs techniques moléculaires choisies en fonction des objectifs à atteindre. Les plus communes à l'heure actuelle sont le metabarcoding et la digital droplet PCR (ddPCR). Ces deux techniques metabarcoding et ddPCR - dont les étapes principales sont représentées dans la **Figure 3** - seront ici mises en œuvre.

L'approche par metabarcoding vise à analyser en une fois l'ensemble des ADN présents dans l'échantillon, et donc la communauté entière (concept de surveillance « passive »), ou a minima un large groupe taxinomique, comme ce sera le cas ici en ciblant plus particulièrement les poissons téléostéens. Les approches par ddPCR sont utilisées pour cibler une espèce en particulier (concept de surveillance « active »). Nous considérerons dans ce rapport que l'approche metabarcoding telle qu'employée dans cette étude est un outil de **surveillance passive** et que l'approche ddPCR est un outil de **surveillance active**, concepts communément employés dans la littérature « ADNe » (Blackman et al., 2020).

de biais d'amplification, Hatzenbuhler et al. 2017). Face aux limites de l'utilisation de COI, d'autres équipes ont travaillé à la définition de nouveaux marqueurs (par ex. gènes 12S et 16S) permettant une haute résolution taxinomique tout en limitant les biais d'amplification (Kelly et al., 2016; Miya et al., 2015). Les bases de référence sur ces marqueurs sont moins développées que pour COI mais des efforts importants sont faits par la communauté scientifique pour les enrichir, un effort auquel notre consortium participe. La complémentarité des approches par metabarcoding par rapport aux autres approches (par exemple visuelles) pour l'étude des assemblages ichthyologiques a été montrée dans le cadre d'étude de biodiversité de ces assemblages (Stat et al., 2019; Thomsen et al., 2012) ainsi que pour des applications en aquarium (Kelly et al., 2014) ou pour les pêcheries (Hansen et al., 2018) mais reste encore inexplorée en Manche notamment dans la détection des poissons des habitat côtiers rocheux. Le metabarcoding d'ADNe n'est toutefois pas exempt de limitations, notamment concernant les estimations d'abondance (Fraija-Fernandez et al., 2020; Lamb et al., 2019; Thomsen et al., 2016) et l'impossibilité d'accéder à la structure de taille (âge) des populations.

2.3.3 La technique de droplet digital PCR sur ADNe comme outil de surveillance active

La technique dite de droplet digital PCR (ddPCR) est une technique alternative au séquençage haut-débit pour analyser des échantillons d'ADNe. A l'inverse du metabarcoding, cette approche est spécifique et quantitative, c'est-à-dire vise à détecter la présence et estimer l'abondance relative d'une espèce donnée (l'abondance de l'espèce est comparée au sein d'une série d'ADNe obtenue dans les mêmes conditions d'extraction et d'analyse). Elle présente de nombreux avantages par rapport aux techniques de PCR quantitative, en particulier en s'affranchissant de la nécessité de développer un standard pour effectuer les dosages, et donc la quantification. La ddPCR a été en premier lieu utilisée pour estimer l'abondance de poissons d'eau douce (Capo et al., 2020; Doi et al., 2015). La méthode est décrite dans Hindson et al. (2011). En résumé, un petit volume d'ADNe, est incorporé à une émulsion huile et d'eau pour former des gouttelettes, chacune faisant l'objet d'une PCR avec des amorces et une sonde fluorescente, toutes trois spécifiques de l'espèce cible. La fluorescence est détectée dans chacune des gouttelettes, ce qui permet de quantifier le nombre de copie d'ADN correspondant à l'espèce cible sur l'ensemble des gouttelettes, et donc par volume de prise. Si cette méthode semble à la fois sensible et spécifique, elle a un inconvénient par rapport au metabarcoding : elle ne peut avec des coûts raisonnables n'être utilisée que sur un petit nombre d'espèces.

2.4 Hypothèses de travail concernant les complémentarités entre méthodes

D'après la littérature, il semblerait que UVC et ADNe soient complémentaires, tant au niveau des catégories d'espèces surveillées, des métriques échantillonnées, et des couvertures et résolutions spatiales et temporelles (**Tableau 1**).

Afin de tirer avantages des complémentarités entre UVC et ADNe pour l'évaluation des différents descripteurs et critères du D1, D2 et D4 relatifs aux poissons fréquentant les milieux côtiers (**Tableau 2**), il apparaît pertinent d'articuler le déploiement de ces approches à deux échelles spatiales:

(1) l'observation par ADNe - permettant d'évaluer la composition des peuplements (metabarcoding) voir des indices d'abondance et/ou de biomasse pour certaines espèces cibles (ddPCR) - vise à évaluer la tendance globale de l'état écologique à l'échelle de la façade, et dépister les zones à la dynamique particulière (pouvant faire l'objet d'études plus approfondies, e.g. un observatoire état-pression)

(2) les observatoires état-pression (e.g. dans et autour une AMP, ou un parc éolien, à graduelle distance d'une zone urbanisée) combinent ADNe et comptages visuels (et autres méthodes selon l'habitat). Ces dernières méthodes sont plus « lourdes » à mettre en œuvre mais apportent des métriques supplémentaires nécessaires à l'évaluation de certaines relations état-pression (et certains descripteurs et critères DCSMM). Les observatoires état-pression visent ainsi à :

- (a) évaluer finement les relations état-pression (dont résistance et résilience sur le long terme) grâce aux complémentarités des différentes méthodes
- (b) calibrer les indicateurs du Screening ADNe, et éclairer leur interprétation grâce à l'inter-calibration de l'ADNe « boîte noire » avec les méthodes d'observations directes.

Ce concept de surveillance multi-source et multi-échelle est un cadre de travail dans lequel s'insère le présent rapport. Le présent rapport est une première étape visant à vérifier/tester, pour les milieux côtiers rocheux de Manche Occidentales, les avantages/inconvénients présentés en **Tableau 1**, et développer/comparer certains protocoles ADNe. Des travaux ultérieurs à plus large échelle permettront de tester l'efficacité d'un déploiement à deux échelles pour l'évaluation des différents descripteurs et critères du D1, D2 et D4 relatifs aux poissons fréquentant les milieux côtiers (présentés en **Tableau 2**).

Tableau 1. Hypothèses de travail concernant les caractéristiques des méthodes de suivis comptages visuels en plongée (UVC) et ADN

Caractéristiques d'échantillonnage	UVC	ADNe
Exhaustivité des espèces/taxons observés	Esp. nectobenthique diurnes bien détectées Autres espèces mal détectées (e.g. cryptobenthiques, fuyantes, nocturnes, et rares)	L'ensembles des espèces, quelques soient leurs traits de vie et comportementaux, sont potentiellement bien détectés
Résolution taxinomique	Résolution élevée pour les espèces facilement identifiables (et détectables). Cependant les espèces de certaines familles ne peuvent être déterminée visuellement (e.g. Mugilidae) Impossibilité de travailler sur de la diversité intra-spécifique	Résolution élevée pour la grande majorité des espèces, dépendante toutefois de la disponibilité des séquences de référence et des marqueurs génétiques utilisés pour l'identification à l'espèce Possibilité pour quelques espèces de travailler sur la diversité intra-spécifique
Résolution et couverture spatio-temporelle	Grande résolution spatiale : possibilité de distinguer des micro-habitats au sein des récifs rocheux, et de bien distinguer les habitats adjacents même très proches Grande résolution temporelle : les espèces sont observées directement (le jour). Néanmoins, des espèces présentes et/ou actives la nuit ne peuvent être détectées.	Résolution spatiale intermédiaire : le pool d'espèces détecté peut-être présent dans l'habitat échantillonné et/ou dans les habitats adjacents (→ échantillonnage d'un paysage plus que d'un habitat). Des stations d'échantillonnage distantes de quelques kms peuvent néanmoins être discriminer. Résolution temporaire incertaine, la persistance de l'ADN dans l'eau serait potentiellement inférieure à 72h. Cela peut être avantageux pour les espèces nocturnes et/ou de passage
Coûts de réplification	Coûts élevés, indépendant du nombre de sites et/ou période d'échantillonnage	Coûts <i>a priori</i> plus faibles, avec cout unitaire pouvant diminuer avec le nombre d'échantillons
Estimation de l'abondance / biomasse des individus	Estimation directe et précise (pour les espèces nectobenthiques diurnes)	Estimations <i>a priori</i> peu précises par metabarcoding (mais couvrant l'ensemble des espèces) et potentiellement assez précise par ddPCR (pour quelques espèces cibles)
Estimation de la taille, sexe, morphologie et stade de vie	Estimations des tailles et stades de vie possibles (dépendant de la formation des plongeurs et du protocole mis en place), voir du sexe pour certaines espèces à dimorphisme sexuel	Impossibilité de fournir ce type de données avec la méthode

Tableau 2. Hypothèses de travail concernant les descripteurs et critères pouvant potentiellement être renseignés par les différentes méthodes d'observation

	Echelle spatiale du dispositif	Milieux côtiers surveillés	Types d'espèces	Type d'évaluation	Descripteurs DCSMM potentiellement renseignés
Metabarcoding sur ADNe	« Screening de façade »	Tous les milieux (SP1 à 3)	Toutes les espèces	Tendances des diversités taxonomique, fonctionnelle et phylogénétiques (D4C1) Tendance de la diversité de groupes d'espèces fondés sur leur sensibilité aux pressions anthropiques Evolution des aires biogéographiques de distribution, par patrons d'occurrences (D1C4 et D2C2)	D1C4 D4C1 D2C2 + apport de compléments pour D1C2, D1C5, D4C2, D2C1, D2C3
ddPCR sur ADNe	« Screening de façade »	Tous les milieux (SP1 à 3)	Quelques espèces ciblées : e.g. espèces patrimoniales, amphihalins, élaémobranches à statut de conservation, ...	Tendance des abondances spécifiques (D1C2) Evolution des aires biogéographiques de distribution, par patrons d'abondances/biomasses (D1C4 et D2C2)	D1C2 D1C4 D2C2 + apport de compléments pour D1C5
Comptages visuels en plongée	Observatoires Etat-Pression	Milieux rocheux et habitats biogéniques (SP1)	Nectobenthiques	Indice de diversité d'abondance (D1C2) et évolution de la structure démographique (D1C3) Caractérisation de l'habitat pour établir les liens état-pressions (D1C5) Tendances des abondances des peuplements et structure en taille (D4C2 et D4C3)	D1C2 D1C3 D1C5 D4C2 D4C3 + apport de compléments pour D1C4 et D4C1

2.5 Objectifs de l'étude

Afin de tester l'efficacité de l'utilisation de l'ADN environnemental, et évaluer ses complémentarités par rapport aux méthodes de comptage visuels *in situ* pour *in fine* déterminer la pertinence de l'établissement d'un dispositif de surveillance utilisant l'ADN environnemental (en complément aux approches conventionnelles), ce projet vise à :

- Contribuer à la production de références moléculaires en particulier pour les espèces de la zone d'étude, encore peu présentes dans les bases de données (*i.e.*, objectif d'enrichissement des bases de données publiques) ; la disponibilité de séquences de référence est en effet indispensable aux analyses par metabarcoding sur ADNe pour réaliser les assignations taxinomiques au niveau spécifique ;
- Comparer dans les mêmes sites aux mêmes dates la liste d'espèces obtenue en plongée par rapport à celle obtenue par metabarcoding, afin d'identifier le niveau de redondance et d'exclusivité des informations de présence/absence obtenues par ces deux méthodes ;
- Comparer les métriques de diversité de peuplement obtenues quand elles sont estimées sur la base des données issues du metabarcoding (*i.e.*, assignation taxinomique) par rapport à l'analyse de la seule diversité moléculaire obtenues par le séquençage haut-débit (*i.e.*, travail sans assignation taxinomique, en utilisant les OTUs – Operational Taxonomic Units) ; l'utilisation de ce second type de données permettrait de s'affranchir des lacunes (*i.e.*, absence de référence) des bases de données ;
- Tester la possibilité d'utiliser du Digital Droplet PCR pour estimer l'abondance de certaines espèces cibles à partir d'ADNe et ainsi offrir une alternative plus robuste pour le suivi des espèces cibles non ou mal détectées par comptages plongée (e.g. espèces crypto-benthiques, espèces rares) ;
- Evaluer la pertinence de la méthode metabarcoding pour révéler *a minima* des niveaux d'abondance (*i.e.* données semi-quantitatives), par comparaison avec les résultats de la ddPCR ;
- Comparer les performances de 3 types de prélèvement d'eau pour les suivis de l'ichtyofaune côtière des milieux rocheux par ADNe : (1) des prélèvements à un point fixe de 2L d'eau réalisés en plongée près du fond (*i.e.*, échantillonnage au sein du milieu d'intérêt), (2) des prélèvements à un point fixe de 2L d'eau réalisés en sub-surface depuis un bateau, permettant une facilité opérationnelle, et (3) des prélèvements de 30L d'eau réalisés en couvrant plusieurs centaines de mètres autour du site et réalisés en sub-surface depuis un bateau (plus grand volume et plus grande emprise spatiale qui pourrait faciliter la détection des espèces rares mais nécessitant un équipement spécifique à bord) ;
- Développer une analyse critique de différentes stratégies d'échantillonnage et d'acquisition de données de diversité spécifique selon ces différentes méthodes afin de proposer des protocoles opérationnels de surveillance ;
- Proposer une stratégie de workflow permettant l'intégration des données des différentes facettes de diversité (taxinomique, phylogénétique et fonctionnelle) ainsi que des données issues des bases de données de biodiversité afin d'accroître le potentiel des approches telles que le metabarcoding sur ADNe pour détecter la diversité des poissons côtiers.

3 — Méthodes —

3.1 Collecte des données *in-situ* ADNe et UVC

Les échantillonnages (comptages visuels en plongées et échantillons d'eau pour les analyses ADNe) ont été effectués au Printemps et en Automne 2018 dans la baie de Morlaix. Au sein de la baie, 9 sites ont été échantillonnés en plongée représentant l'échantillonnage complet de la baie de Morlaix en comptage visuel de l'année 2018 (Le Joncour and Thiriet, 2020) appelé « UVC Baie » dans ce rapport (**Figure 4**). Au sein de ces 9 sites, 4 ont été sélectionnés (« UVC 24h », sites en gras sur la **Figure 4**) pour être échantillonnés aussi dans le but d'obtenir de l'ADNe : 2 sites sont situés à l'extérieur de la baie (Astan et Cochons Noirs) et 2 sites sont situés en fond de baie (Figuier et Corbeau). Ces derniers sont donc potentiellement plus influencés par les apports des bassins versants et présentent une exposition à la houle moindre. Des prélèvements d'eau pour les analyses ADNe ont été effectués dans le port du Bloscon qui n'a pas été échantillonné avec les UVC. Ces échantillons d'eau ont donc été uniquement analysés dans le cadre de la comparaison ddPCR et metabarcoding.

Les données d'observation visuelle en plongée ont été obtenues grâce à deux types de protocoles : le transect métrique (30 m, environ 8 min) et le parcours libre (8 min, environ 30 m). Les complémentarités entre méthodes de comptages visuels ne faisant pas l'objet de la présente étude (Rey et al., 2021), l'ensemble des données issues de ces deux méthodes visuelles, conçues pour être interopérables, ont été rassemblées au sein d'un jeu de données unique. Tous les sites ont été échantillonnés à 2 dates par saison (chaque date séparée de 2 semaines), à l'exception du site Paradis échantillonné à une seule date au Printemps.

Pour les échantillons d'eau servant à l'extraction de l'ADNe, trois types de prélèvement ont été réalisés. Le premier, appelé « 2L fixed surface », est un triplicat d'échantillons de 2 L d'eau de mer prélevé à l'aide d'une outre, en un même point en surface (depuis le support de surface utilisé). Le second, appelé « 2L fixed bottom » est un triplicat d'échantillons de 2 L, prélevés également avec une outre au fond, *i.e.* à environ 50 cm au-dessus du substrat rocheux à une profondeur d'environ -18m cote marine, réalisé par un opérateur en scaphandre autonome immergé. La comparaison de ces deux types de prélèvements (surface versus fond) vise à évaluer l'effet de l'horizon d'échantillonnage sur la capacité de détection de différents types de poissons (ex : pélagique, démersal/benthique). Ces prélèvements ont été réalisés au sein de chacun des 4 sites et au cours de chacune des deux saisons. Dès le retour au laboratoire, ces échantillons d'eau ont été filtrés sur des sterivex 0.2µm qui ont été conservés à -20°C dans une solution tampon en vue de l'extraction ultérieure de l'ADN. Le troisième type de prélèvement appelé « 30L area surface » est une série de 4 échantillons de 30 L d'eau filtrée en continu, pendant un parcours de 30 min, à l'aide d'une pompe depuis un bateau naviguant à faible allure (3 nœuds) en faisant des va et viens au-dessus et dans le périmètre de chaque site échantillonné (surface parcourue de l'ordre de 200 x 200 m, **Figure 1** en annexe). La filtration réalisée à bord emploie des kits commerciaux permettant une filtration sur 0.2µm de grands volumes. Utiliser cette approche vise à tester l'effet

d'utiliser des échantillons intégrant à la fois un volume d'eau filtré beaucoup plus grand et couvrant une zone étendue autour du site. Ces prélèvements ont été réalisés dans les même 4 sites que les prélèvements de 2 L mais uniquement en Automne.

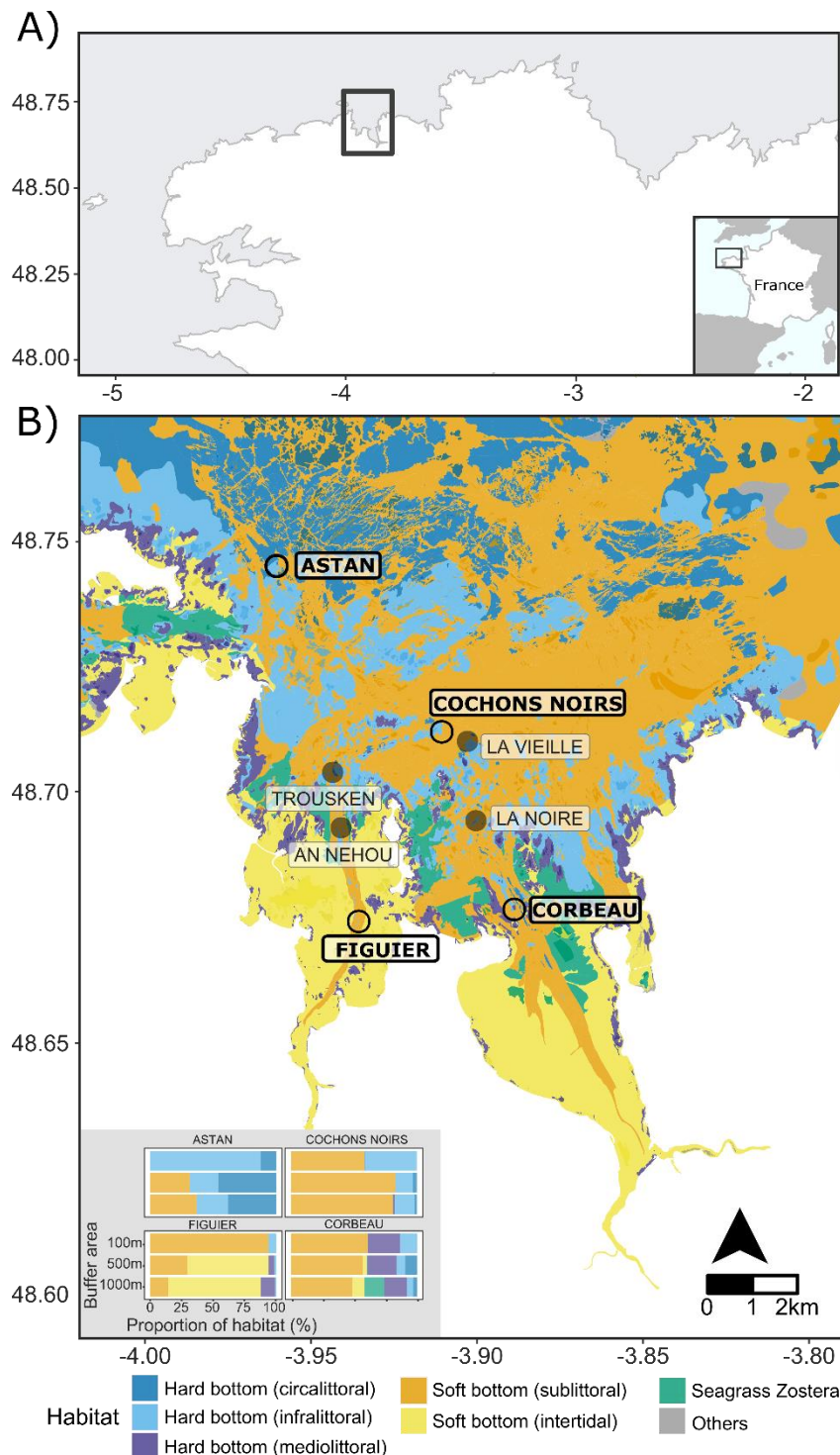


Figure 4. Carte d'échantillonnage de la baie de Morlaix indiquant la localisation des 8 sites échantillonnés en plongée (« UVC Baie ») dont 4 sites (en gras) ont également fait l'objet d'étude d'ADNe (« UVC 24h »). Les couleurs représentent les différents habitats benthiques de la baie et permettent d'évaluer la proximité du site d'habitat rocheux échantillonné avec les autres habitats (source : <https://www.data.gouv.fr/fr/datasets/carte-de-synthese-des-habitats-benthiques-du-site-natura-2000-fr5300015-baie-de-morlaix/>).

3.2 Enrichissement de la base de référence moléculaire pour le metabarcoding

Pour réaliser les assignations taxinomiques des ADN présents dans les échantillons d'eau, nous devons utiliser une base de référence (voir **Figure 3**). Cette base de référence doit contenir les espèces listées D1 (arrêté BEE 2019) mais aussi toutes les autres espèces potentiellement présentes afin de pouvoir évaluer les indices de diversité ichthyologique pouvant renseigner le descripteur 4. Pour cela, une liste de référence de l'ichtyofaune côtière potentiellement présente en Manche a été établie en combinant les espèces observées en comptages visuels des milieux rocheux réalisés en plongée en Manche dans le cadre du réseau POCOROCH (Le Joncour and Thiriet, 2020), les listes d'espèces issues du dire d'expert et les données récoltées dans la littérature incluant les milieux estuariens (Teichert et al., 2017a, 2017b). Au moment de la création de la liste de référence en début de projet, c'est-à-dire en 2019, la base de référence GAPeDNA (<https://shiny.cefe.cnrs.fr/GAPeDNA/>, Marques et al. 2021) présentant la disponibilité des séquences pour les poissons recensés dans chaque écorégion marine n'était pas encore publiée. Nous avons cependant décidé de continuer avec la liste de référence créée dans ce projet car nous avons constaté des différences de présence d'espèce entre la liste de référence du projet et les espèces recensées dans l'écorégion Mer Celtique de la base GAPeDNA. Par exemple, *Parablennius pilicornis*, qui est en extension d'aire de répartition en Atlantique Manche, est référencée dans la liste de référence du projet car elle a été observée en plongée en baie de Morlaix alors qu'elle n'est pas recensée dans l'écorégion Mer Celtique de la base GAPeDNA.

Cette liste de référence non exhaustive comprenant un total de 212 espèces nous a servi à évaluer la complétude des bases de données moléculaires pour notre étude. En effet, les bases de données moléculaires fournissant les références pour le metabarcoding sont lacunaires (pas de données moléculaires pour certaines espèces ou données sujettes à caution). Par ailleurs, les travaux menés par les équipes de Roscoff ont montré l'importance d'obtenir des références locales pour une assignation taxinomique plus robuste par metabarcoding (Couton et al., 2019). Nous avons donc cherché à améliorer (enrichir) ces bases de référence.

Des échantillons étaient disponibles pour certaines espèces à la Station Marine de Dinard. Pour d'autres espèces manquantes, plusieurs techniques ont été utilisées pour collecter 1-3 individus par espèce, notamment des pêches au chalut, à la senne de plage, à pied le long de l'estran, ainsi que des chasses en plongée et des visites de criées. Chaque individu a ensuite été disséqué dans les heures suivants sa collecte (dans la mesure du possible les poissons étaient maintenus en eau jusqu'à leur dissection). Deux morceaux de muscle et deux morceaux de nageoire ont été prélevés par individu et conservés dans de l'éthanol à 99% en vue de leur analyse génétique. L'ADN de deux échantillons de tissus (muscle et nageoire) de 81 espèces, soit un total de 160 échantillons, a ensuite été extrait en utilisant les kits NucleoSpin® Tissue, NucleoSpin® DNA Lipid Tissue ou NucleoSpin® 96 Tissue (Macherey-Nagel). La qualité de l'ADN extrait a été vérifiée sur des gels d'agarose à 0,8%. Une amplification par polymérisation en chaîne (PCR) de l'ADN a ensuite été réalisée en utilisant trois

marqueurs et des amorces généralistes ou ciblant certains groupes taxinomiques: 1) 18S, un gène nucléaire codant pour la petite sous-unité de l'ARN ribosomique, avec les amorces NSF4 et NSR58 ciblant l'ensemble des eucaryotes (Winnepenninckx et al., 1994) ; ce marqueur généraliste n'est pas utilisé dans nos études de metabarcoding mais nous sert de contrôle positif pour la PCR (test de qualité d'ADN), et la production de données de référence pourra être utile pour d'autres études ; 2) 12S, un marqueur mitochondrial, avec les amorces V05F_898 et Teleo-R spécifiques des poissons téléostéens (Thomsen et al., 2016) et 3) 16S, un marqueur mitochondrial, avec des amorces spécifiques des metazoaires (cible Euka03, Taberlet et al. (2018)). La réussite et qualité de l'amplification a été analysée par migration des produits de PCR sur des gels d'agarose à 1,6%. Les produits PCR répondant aux critères de qualité ont été envoyés pour séquençage Sanger à la compagnie Eurofins.

3.3 Acquisition des données moléculaires issues d'ADNe

3.3.1 Extraction ADNe

Un total de 67 échantillons d'ADNe a été obtenu : 51 ADNe issus des triplicats de 2 L de fond ou de surface, obtenus par l'équipe de l'UMR7144 impliquée dans le projet (protocole mis au point par cette équipe), et 16 issus des échantillons de 30 L, obtenus par la société développeuse de ces kits (SPYGEN).

3.3.2 Séquençage haut-débit

L'amplification, la construction des bibliothèques et les séquençages des 67 ADNe ont été réalisés par la société SPYGEN. Chaque extrait d'ADNe a été amplifié avec 2 marqueurs mitochondriaux, 12S et 16S respectivement avec les amorces développées par Valentini et al. (2016) pour détecter en particulier les poissons téléostéens et Kelly et al. (2016) ciblant plus généralement les métazoaires, et particulièrement efficaces sur les chordés. A noter que les fragments obtenus sont plus courts (adaptés au séquençage haut-débit) que ceux obtenus pour le barcoding moléculaire (cf. 3.2) mais sont contenus dans les fragments produits pour le barcoding. Les 67 ADNe ont été amplifiés 12 fois (=12 répliquats de PCR) afin de maximiser la détection des espèces dont le signal ADN est plus faible et de minimiser les biais stochastiques de PCR. Les 12 répliquats de PCR ont été séquençés séparément afin de pouvoir observer si besoin la variabilité entre répliquats de PCR.

3.3.3 Analyses bioinformatiques des données de metabarcoding

Les approches exploitant les séquençage haut-débit issus de l'ADNe, telles que le metabarcoding, reposent sur des analyses bioinformatiques visant à nettoyer le jeu de données afin de garder seulement les séquences de qualité parmi les millions de séquences obtenues puis identifier des OTUs (« Operational Taxonomic Units ») pour ensuite réaliser des assignations taxinomiques (mise en relation d'un OTU avec une espèce, par l'exploitation de bases de données de référence), en amont des analyses statistiques et calculs d'indices. Dans la version préliminaire du rapport (Rey et al., 2019), un premier jeu d'analyse fournie par la société SPYGEN suivant leur pipeline d'analyse, qui intègre

notamment des outils de la suite OBITOOLS (Boyer et al., 2016) avait été utilisé. Cependant, 3 facteurs limitaient l'exploitabilité de ce jeu de données : 1) les assignations taxinomiques étaient restreintes à l'utilisation des bases de données publiques et donc n'intégraient pas les nouvelles références moléculaires obtenues ici (cf. 3.2) ; 2) la procédure détaillée bio-informatique était manquante, ce qui empêchait la reproductibilité des analyses ainsi que la compréhension de la prise en compte de contaminations diverses (laboratoire, séquençage, ...), étape importante dans les analyses de metabarcoding pour s'assurer de la qualité des résultats obtenus ; et 3) aucune analyse de diversité des OTU n'était possible car seule une liste de taxons issue de l'analyse des 67 échantillons était fournie. Ces limitations nous ont donc conduit à ré-analyser les données brutes ; ces nouvelles analyses sont brièvement décrites ici. Le **Tableau 1** en annexe détaille les programmes et paramètres utilisés dans chacune des étapes bio-informatiques. La première étape bioinformatique qui consiste à rassembler ensemble les séquences venant d'un même échantillon et d'un même marqueur (12S ou 16S) a été faite avec le programme Cutadapt (Martin, 2011). Ensuite une série d'étapes de nettoyage permettant d'enlever les séquences considérées comme des erreurs (provenant des étapes de PCR, des constructions de librairie et des séquençages) a été réalisée avec les packages R DADA2 (Callahan et al., 2016), LULU (Frøslev et al., 2017) ainsi qu'avec des scripts R que nous avons créés pour ce projet. Les séquences ont ensuite été assignées taxinomiquement en les comparant avec une base de référence « maison » qui a été créée pour ce projet. Pour chacun des deux marqueurs 12S et 16S, cette base de référence a été créée en reprenant toutes les séquences de la base de références publique NCBI des chordés et des métazoaires et en y ajoutant les séquences de référence obtenues dans ce projet (cf. 3.2). L'assignation taxinomique a été réalisée en utilisant 3 méthodes différentes d'assignations, couramment utilisées en metabarcoding, BLAST (Altschul et al., 1990), RDP Classifier (Wang et al., 2007) et ECOTAG (Boyer et al., 2016). L'assignation consensus entre les 3 méthodes a été choisie comme assignation taxinomique finale. Les séquences identifiées dans les contrôles négatifs et les séquences assignées à des taxons non ciblés dans le projet ont été enlevées du jeu de données. Finalement les séquences uniques (ASV ; Amplicons Sequence Variants) ont été regroupées en OTU avec le programme SWARM (Mahé et al., 2015) et les OTUs singletons (trouvés dans un seul réplicat de PCR d'un seul échantillon) ont été enlevés. L'assignation taxinomique de la séquence représentative des séquences regroupées dans un OTU est choisie comme assignation taxinomique de l'OTU. Le groupement des séquences en OTUs permet de faire l'analyse de la diversité moléculaire obtenue par le séquençage haut-débit et donc de s'affranchir des lacunes (*i.e.*, absence de référence) des bases de données qui limite les analyses faites sur les assignations taxinomiques. La partie des résultats 4.3.1 compare ces approches (OTUs *vs* assignations taxinomiques) présentées en **Figure 5** et évalue l'impact du choix de ces approches sur les analyses de diversité. Concernant les approches basées sur les assignations taxinomiques, pour éviter les redondances dans le jeu de données, 2 méthodes ont été considérées :

- une méthode conservatrice (sous-estimation potentielle de la diversité) : Quand une espèce a été identifiée, toutes les séquences assignées aux rangs taxinomiques supérieurs de cette espèce ne sont pas considérées comme des espèces nouvelles et sont supprimées du jeu de données.
- une méthode non conservatrice (surestimation potentielle de la diversité) : Tout ce qui n'est pas identifié à l'espèce est un nouveau taxon pour lequel nous n'avons pas la séquence de référence dans la base

Après les étapes de nettoyage bio-informatique, on a le jeu de données suivant :			Approche basée sur les assignations taxinomiques : méthode conservatrice		Approche basée sur les assignations taxinomiques : méthode non conservatrice		Approche basée sur la diversité moléculaire : OTU		
ASV	Assignment de l'unité taxinomique	Abondance dans Echantillon 1	Assignment de l'unité taxinomique	Abondance dans Echantillon 1	Assignment de l'unité taxinomique	Abondance dans Echantillon 1	OTU	Assignment de l'unité taxinomique	Abondance dans Echantillon 1
ASV1	Labrus bergylta	150	Labrus bergylta	350	Labrus bergylta	350	OTU1	Labrus bergylta	150
ASV2	Labrus bergylta	200					OTU2	Labrus bergylta	200
ASV3	Labrus viridis	15	Labrus viridis	15	Labrus viridis	15	OTU3	Labrus viridis	15
ASV4	Labrus	20	ASV supprimé du jeu	0	Labrus_other	42	OTU4	Labrus	42
ASV5	Labrus	22	ASV supprimé du jeu	0					
ASV6	Labridae	50	ASV supprimé du jeu	0	Labridae_other	55	OTU5	Labridae	50
ASV7	Labridae	5	ASV supprimé du jeu	0					
ASV8	Pollachius	55	Pollachius pollachius	55	Pollachius pollachius	55	OTU7	Pollachius pollachius	55
ASV9	Scomber	15	Scomber	15	Scomber	15	OTU8	Scomber	15
ASV10	Actinopterygii	1000	ASV supprimé du jeu	0	Actinopterygii_other	1800	OTU9	Actinopterygii	1000
ASV11	Actinopterygii	800	ASV supprimé du jeu	0			OTU10	Actinopterygii	800

Figure 5. Explication des jeux de données utilisés pour faire les analyses basées sur les assignations taxinomiques et sur la diversité moléculaire (OTU).

3.3.4 Analyses par ddPCR

Notre premier travail a été de sélectionner des espèces candidates. Cette sélection s'est faite en fonction de critères à remplir pour pouvoir réaliser de la ddPCR : **(1)** des espèces d'intérêt par rapport aux objectifs de ce projet (par ex., une espèce vue abondamment en UVC, une espèce abondante en metabarcoding mais pas en UVC, etc.) **(2)** la disponibilité de plusieurs séquences pour la région 12S étudiée par metabarcoding, sur l'espèce cible (pour déterminer son niveau de polymorphisme et la fiabilité des sondes), **(3)** la disponibilité de ce même type de séquence sur plusieurs espèces proches (pour être certain du diagnostic spécifique), **(4)** la disponibilité d'ADN des espèces cibles et proches afin de valider expérimentalement la spécificité des sondes, **(5)** des échantillons d'ADNe chez lesquels la cible est absente ou présente par metabarcoding afin de pouvoir réaliser des comparaisons entre les deux approches. Suite à cette sélection d'espèce, des analyses bioinformatiques ont été réalisées par la société IAGE, en collaboration avec l'équipe projet, pour définir des zones candidates pour dessiner les amorces et les sondes. Ce prestataire a ensuite dessiné les amorces et sondes puis a réalisé des tests de validation, dans le but de procéder par la suite à des analyses en routine sur les 51 ADNe obtenus par l'équipe UMR7144.

3.4 Analyses statistiques de comparaison entre UVC et metabarcoding

Les données issues du metabarcoding réalisé à partir de l'ADNe extrait des différents types de prélèvement d'eau ont été comparées aux données issues des comptages visuels en plongées, afin de déterminer leur complémentarité et caractère opérationnel (*i.e.* liste des objectifs pouvant être atteint par l'une et/ou l'autre des deux approches) pour de futurs suivis.

Les données moléculaires ont été analysées pour comparer la diversité et la structure des assemblages ichtyologiques entre les différentes catégories de sites (saison, gradient côte-large), de types de données (visuelles *in situ* et issues de l'analyse de l'ADNe), et d'échantillons (facteurs profondeur, volume filtré, point fixe *vs* zone étendue). A l'issue de ces analyses, des stratégies de surveillance combinant analyses d'ADNe et comptages visuels pourraient être proposées, si cela s'avère pertinent. Les données issues des 2 méthodes d'observation ont été transformées en présence/absence afin de comparer la diversité taxinomique observée par le metabarcoding et les UVC.

Afin de mieux évaluer les spécificités des différentes stratégies d'échantillonnage et méthodes, les analyses comparatives ont aussi pris en compte la préférence d'habitat des taxons observés. Cette préférence d'habitat a été obtenue par dire d'expert et 5 catégories ont été créées :

- **Pélagique** (pelagic) : Espèce qui vit et se nourrit dans la colonne d'eau, loin du fond
- **Démersale – spécialiste rocheux** (demersal-rocky substrate) : Espèce qui vit et se nourrit sur ou proche du fond, avec une affinité exclusive au substrat rocheux
- **Démersale – non spécialiste rocheux / autres substrats** (demersal-other substrates) : Espèce qui vit et se nourrit sur ou proche du fond sans affinité exclusive au substrat rocheux. Cela inclut des espèces ayant des affinités aux substrats sableux et/ou herbiers ou bien des espèces démersales sans affinité à un substrat particulier.
- **Benthopélagique** (benthopelagic) : Espèce qui se déplace régulièrement/quotidiennement entre les zones démersale et pélagique.
- **Sans préférence d'habitat** (no associated habitat) : Taxon non identifié à l'espèce pour lequel il est impossible de définir une préférence d'habitat

3.5 Création d'un « workflow » d'analyses de différentes facettes de diversité à partir de données de metabarcoding

L'idée principale du workflow est de faciliter l'intégration des données caractérisant différentes facettes de la diversité des communautés de poissons : diversité taxinomique, diversité fonctionnelle et diversité phylogénétique, à partir des données issues de metabarcoding (**Figure 6**).

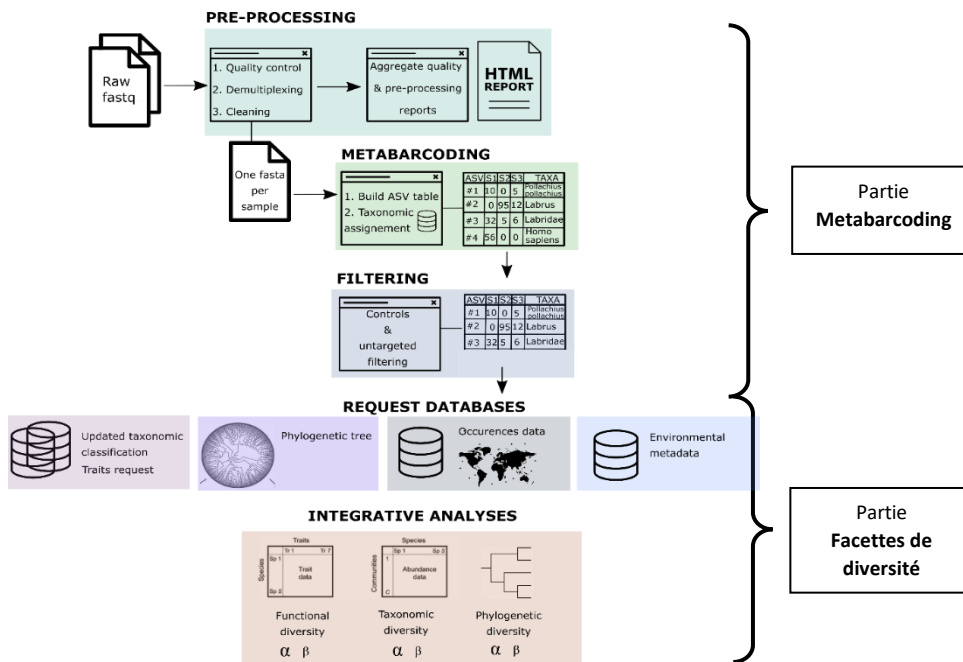
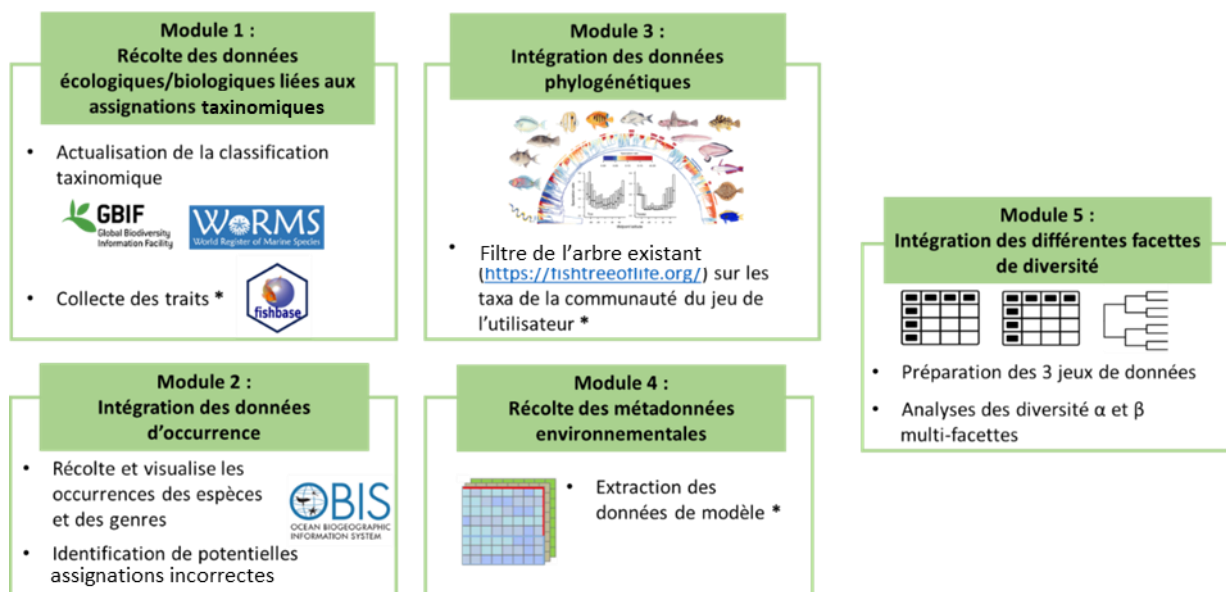


Figure 6. Description schématique du workflow intégrant le traitement de données metabarcoding jusqu'à l'intégration des analyses de diversité taxinomique, fonctionnelle et phylogénétique.

Pour ce projet, nous avons choisi de nous concentrer sur la partie la plus novatrice, *i.e.* la partie concernant l'intégration des différentes facettes de diversité. Nous avons donc travaillé sur la création avec le langage R de 5 modules représentés dans la **Figure 7**.



* Dans le cas où l'utilisateur ne fournit pas la table de trait, l'arbre phylogénétique ou les métadonnées de l'échantillonnage

Figure 7. Représentation des modules en cours de création pour la partie « Facettes de diversité » du workflow en cours.

4 Résultats

4.1 Enrichissement des bases de séquences

Sur les 212 espèces présentes ou potentiellement présentes en Manche, 176 espèces, soit 83 %, avaient au moins une séquence disponible pour l'un des marqueurs 12S ou 16S (dont 126, soit 59.4 %, et 165 soit 77.8 %, respectivement pour le marqueur 12S et 16S) au moment de la recherche dans les bases publiques pour la création de la base de séquences, c'est-à-dire en mai 2020. Pour améliorer la complétude des bases de données, notamment en références issues de spécimens collectés dans l'aire d'étude, un effort particulier a été mis en œuvre pour obtenir des références moléculaires pour les espèces fréquemment rencontrées en baie de Morlaix ou absentes de ces bases de séquences. Au total, 81 espèces ont été collectées. Les extractions et amplifications réalisées pour plusieurs individus par espèce ont été satisfaisantes pour 77 individus avec l'amorce 18S, 74 avec l'amorce 12S et 159 avec l'amorce 16S. Pour un individu par espèce, les couples ADN/amorce permettant théoriquement le séquençage ont ensuite été sélectionnés (53 ADN pour 18S, 48 pour 12S et 81 pour 16S) pour être séquencés par la société Eurofins. La qualité des séquences n'était pas optimale mais un total de 40 et 15 séquences pour 12S et 16S respectivement ont pu être utilisées pour compléter la base de séquences, permettant donc l'ajout de séquence locale pour 48 espèces (voir **Tableau 2 en Annexe**). Au moment de la rédaction du rapport (Mai 2021), 12 espèces pour le marqueur 12S et 1 espèce pour le 16S pour lesquelles nous avons obtenus une séquence locale n'ont toujours pas de séquence disponible dans les bases publiques (voir **Tableau 2 en Annexe**). Les séquences obtenues dans cette étude seront rendues publiques en 2021.

Après l'ajout des séquences locales, 181 espèces, soit une complétude de la base finale de 85 %, avaient au moins une séquence disponible pour l'un des marqueurs 12S ou 16S (dont 141, soit 66.5 %, et 166 soit 78.3 %, respectivement pour le marqueur 12S et 16S). Il est important de noter que malgré la complétude importante de la base, les séquences disponibles ne sont pas forcément des séquences dites diagnostiques, c'est-à-dire permettant l'identification de l'espèce par rapport aux autres espèces de la base. En effet pour 112 espèces toutes leurs séquences sont diagnostiques pour au moins l'un des marqueurs (dont 98 et 28 espèces respectivement pour le marqueur 12S et 16S), ce qui fait que théoriquement seulement 53 % des espèces de la base peuvent être détectées (si présentes dans les échantillons), les autres n'ayant soit pas de séquence disponible (15 %), soit pas de séquence

diagnostique (32 %) (**Figure 8**).

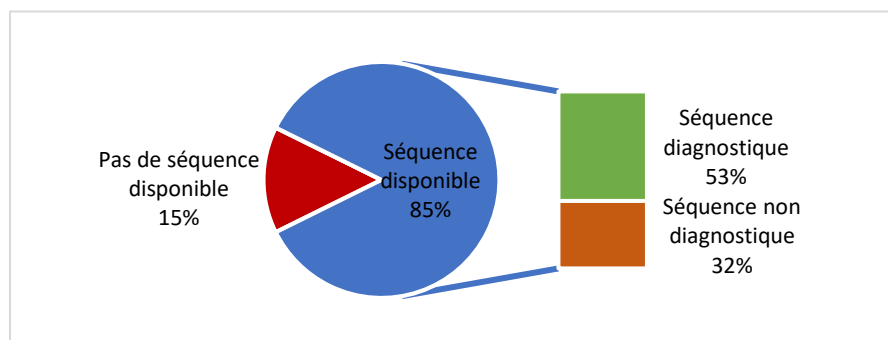


Figure 8. Synthèse de la disponibilité des séquences et de leur caractère diagnostique des 212 espèces présentes dans la base finale recensant les espèces présentes ou potentiellement présentes dans l'aire d'étude.

4.2 Synthèse de l'analyse bioinformatique des données metabarcoding

4.2.1 Séquences conservées pour les analyses statistiques à l'issue des étapes de filtration bioinformatiques

A l'issue du séquençage haut-débit, un total de 76 446 331 et 175 683 371 séquences ont été obtenues respectivement pour 12S et 16S. Chacune des étapes de filtration/nettoyage, visant notamment à éliminer les erreurs de séquençage et à cibler les taxons d'intérêts, a ensuite conduit à une diminution du nombre de séquences conservées : le nombre de séquences gardées à chaque étape est présenté dans le **tableau 3 en annexe**. L'étape qui a le plus fortement réduit le nombre de séquences, et ce pour les 2 marqueurs, est la filtration sur les taxons cibles, à savoir les actinoptérygiens et chondrichthyens. Au final, un total de 18 867 477 et 33 213 663, respectivement pour 12S et 16S ont été conservés pour les regroupements en OTUs et assignations taxinomiques.

4.2.2 Nombre de taxons et d'OTUs

Sur l'ensemble du jeu de données, les séquences uniques, conservées à l'issue des étapes de nettoyage et filtration, ont été assignées taxinomiquement à 85 (dont 53 espèces) et 33 (dont 15 espèces) taxons présents dans au moins 1 réplicat de PCR d'un échantillon biologique, respectivement pour 12S et 16S. Ces taxons sont représentés par 112 et 38 OTUs respectivement pour 12S et 16S (**Tableau 3** et **Figure 9**). A l'issue du regroupement en OTU, il y a un total de 50 et 12 espèces respectivement pour 12S et 16S. Il est à noter que la plupart des espèces sont chacune représentées par 1 seul OTU à part pour quatre espèces qui sont représentées par 2 OTUs dont un OTU est largement plus abondant et présent sur l'ensemble des échantillons (**Figure 10**). Dans le cas du bar (*D.labrax*), qui est la seule espèce où l'OTU non dominant représente plus de 10% du nombre de séquences assignées à cette espèce, les deux OTUs représentent 2 lignées bien connues chez cette espèce, les lignées Atlantique et Méditerranéenne (voir partie 4.4.5). Concernant les 3 autres espèces, la similarité par rapport à la séquence de référence de l'espèce dans la base de données est plus faible pour l'OTU non dominant que pour l'OTU dominant (par ex. par la méthode d'assignation par BLAST, la similarité à la séquence de référence présente dans la base de données est de 100% pour l'OTU dominant et de 95 à 98% pour l'OTU minoritaire). Pour ces 3 espèces, il est possible que l'OTU minoritaire soit un artefact dû à des erreurs de séquençage non supprimées par les étapes bioinformatiques de « nettoyage », ou révèle une lignée cryptique inconnue et légèrement divergente de la lignée majoritaire (à l'instar du bar).

Par ailleurs suite au regroupement en OTU, 8 (dont 3 espèces) et 4 taxons (dont 3 espèces) ont été perdus respectivement pour 12S et 16S car leurs séquences ont été groupées avec des séquences associées à d'autres taxons dans un unique OTU (**Tableau 4**).

Tableau 3. Synthèse des taxons avant et après groupement des séquences en OTU

	12S	16S
Nombre de taxon (espèces) avant regroupement des séquences uniques en OTU	85 (53)	33 (15)
Nombre d'OTU	112	38
Nombre de taxon (espèces) après regroupement en OTU	77 (50)	29 (12)
Espèces représentées par plusieurs OTU (nb d'OTU)	<ul style="list-style-type: none"> ○ <i>Scylliorhinus canicula</i> (2) ○ <i>Sardina pilchardus</i> (2) ○ <i>Dicentrarchus labrax</i> (2) 	<ul style="list-style-type: none"> ○ <i>Callionymus lyra</i> (2)

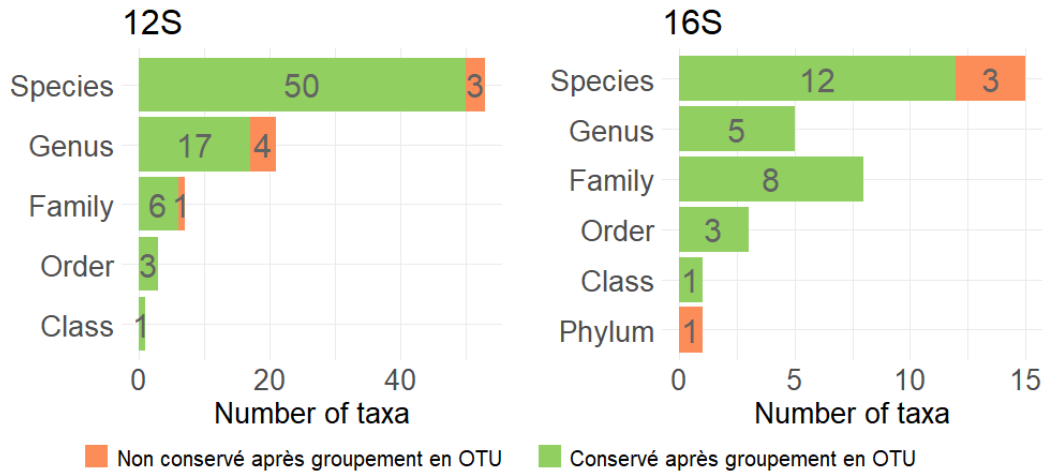


Figure 9. Comparaison du nombre de taxons par rang taxinomique avant et après groupement des séquences en OTU – Genus renseigné pour les assignations aux niveau genre et niveau spp.

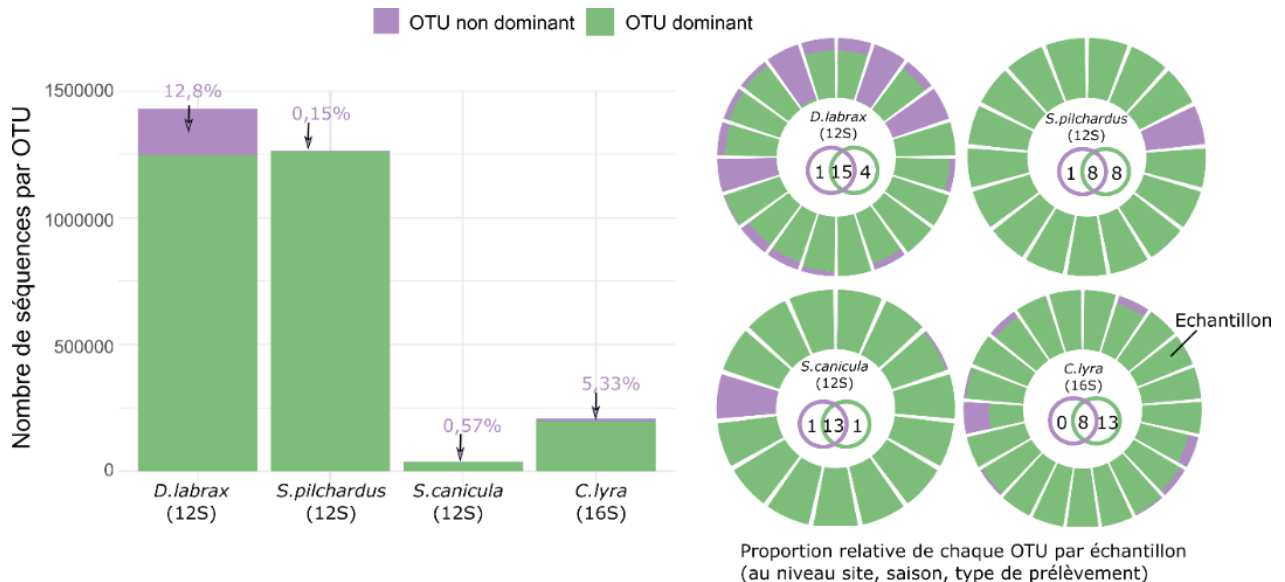


Figure 10. Distribution des OTUs pour les espèces représentées par 2 OTUs. L'histogramme de gauche indique l'abondance relative des deux OTUs sur le jeu de données du marqueur correspondant. Les représentations circulaires à droite indiquent leur proportion relative pour chaque « unité d'observation » (site, saison et type d'échantillon). Les diagrammes de Venn au centre des cercles indiquent pour chaque espèce le nombre d'échantillons pour lesquels les 2 OTUs sont détectés ou pour lesquels seulement l'un des 2 OTUs est détecté.

Tableau 4. Détail des taxons non conservés après regroupement des séquences en OTU

Marqueur	Taxon non conservé après regroupement en OTU	Rang	Correspondance taxonomique de l'OTU dans lequel le taxon a été finalement regroupé après l'étape de regroupement en OTU
12S	<i>Istiophorus albicans</i>	Espèce	Engraulis
	<i>Engraulis encrasicolus</i>	Espèce	Engraulis
	<i>Trisopterus minutus</i>	Espèce	<i>Trisopterus luscus</i>
	Lepidotrigla	Genre	Chelidonichthys
	<i>Oncorhynchus spp.</i>	Genre *car séquence assignée à <i>O. clarkii</i> , espèce absente de la région étudiée	Oncorhynchus
	Symphodus	Genre	<i>Symphodus melops</i>
	<i>Conger spp.</i>	Genre *car séquence assignée à <i>C. oceanicus</i> , espèce absente de la région étudiée	Conger
16S	Gadidae	Famille	Gadiformes
	<i>Lipophrys pholis</i>	Espèce	Gadiformes
	<i>Trisopterus luscus</i>	Espèce	Gadiformes
	<i>Solea senegalensis</i>	Espèce	Soleidae
	Chordata	Phylum	Actinopterygii

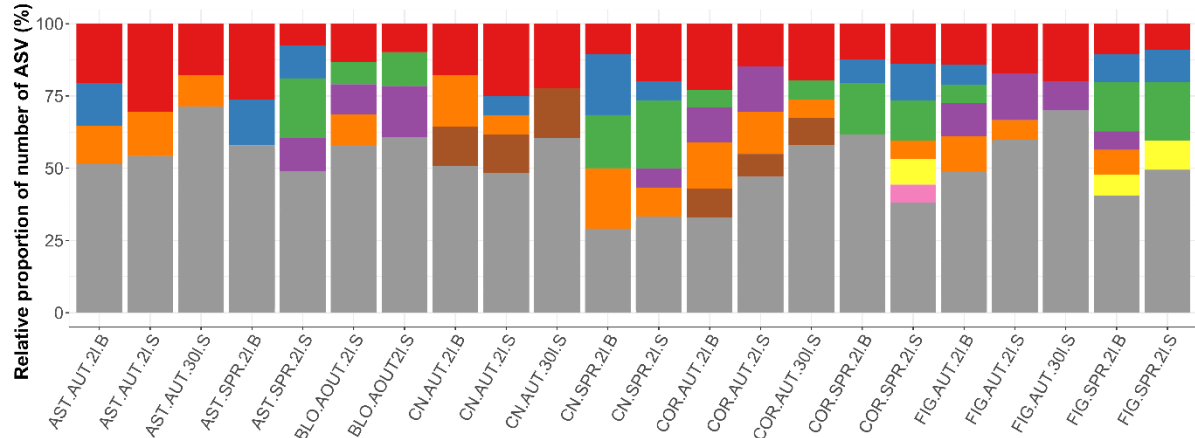
4.2.3 Comparaison de la résolution taxinomique des deux marqueurs

Une différence importante entre les 2 marqueurs est observée : dans le cas de 12S, 73% des ASVs sont assignés au genre ou à l'espèce alors que ce nombre n'est que de 30% pour le marqueur 16S (**Tableau 5**). Cette différence est aussi observée par échantillon biologique avec, pour le marqueur 16S, entre 30 et 60 % des ASVs assignés seulement à la classe (Actinopterygii) (**Figure 11**). Cependant, certaines espèces dont *Lepadogaster candolii*, *Mullus surmuletus*, *Trisopterus minutus* et *Callionymus lyra*, vues lors des comptages visuels, ont été identifiées grâce au seul marqueur 16S, ce qui illustre l'intérêt de combiner différents marqueurs pour réaliser du metabarcoding.

Tableau 5. Résumé global des assignations taxinomiques des ASVs pour chaque marqueur

Rang taxonomique	% du nombre d'ASV assignés à un des rangs taxinomiques	
	12S	16S
Classe	16.8 %	38.1%
Ordre	1.40 %	6.35 %
Famille	8.68 %	15.9 %
Genre	21.6 %	3.17 %
Espèce	51.5 %	27.0 %

12S



16S

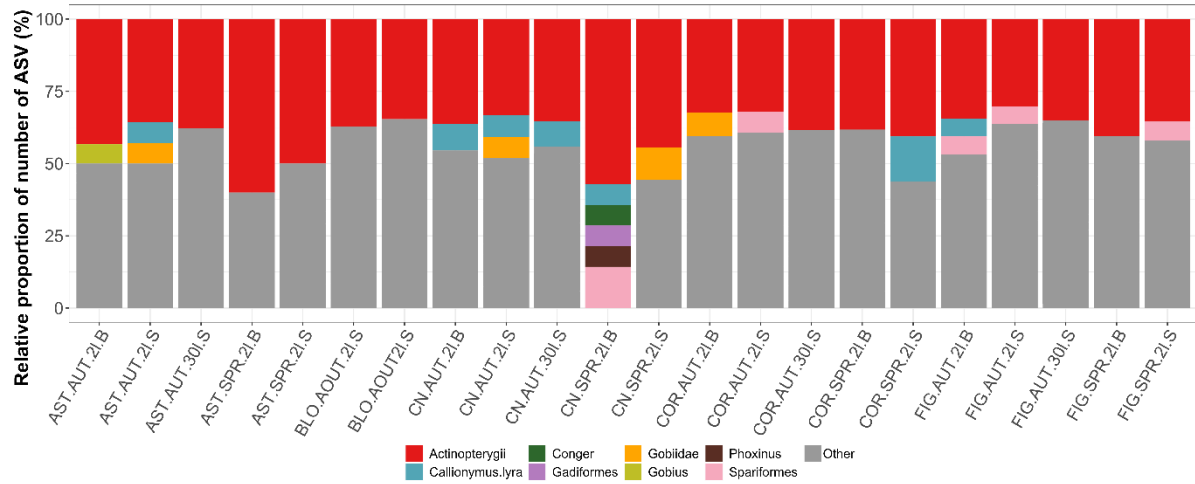


Figure 11. Proportion relative du nombre d'ASVs assignés à un taxon par échantillon biologique, pour les marqueurs 12S et 16S. Les taxons avec une proportion relative en nombre d'ASVs de plus de 6% sont représentés et les autres taxons sont indiqués en gris. La nomenclature des échantillon est « Site.Saison.Type de prélèvement » avec AST pour Astan, CN pour Cochons noirs, COR pour Corbeau, FIG pour Figuier, AUT pour Automne, SPR pour Printemps, 2L.B pour les échantillons 2L au fond, 2L.S, pour les échantillons 2L à la surface, 30L.S pour les échantillons 30L à la surface.

4.3 La surveillance passive des communautés de poissons côtiers par metabarcoding

4.3.1 Comparaison des résultats des analyses par metabarcoding avec ou sans assignation taxinomique, et sélection du jeu de données pour la comparaison avec les données issues des comptages visuels

Trois jeux de données issus du metabarcoding ont été obtenus par les approches suivantes (**Figure 5**): 1) une assignation taxinomique avec une méthode conservatrice pour éviter les redondances entre rang d'assignation ; 2) une assignation taxinomique avec une méthode non conservatrice pour éviter de réduire la diversité et 3) une analyse sur la base des OTUs sans assignation taxinomique. Concernant les deux premières approches, en termes de richesse, les différences majeures observées concernent les séquences assignées au rang de la classe (*i.e.* à Actinopterygii) (**Figure 12**). La classe des Actinopterygii, représente une unité pour le jeu de données basé sur les assignations avec

la méthode non conservatrice, mais disparaît du jeu de données avec la méthode conservatrice. Or cette classe est représentée par 30 et 9 OTUs dans les jeux de données moléculaire, respectivement pour 12S et 16S.

Pour le marqueur 12S, les patrons de diversité α (richesse et richesse pondérée par l'abondance des unités taxinomiques) et de diversité β (structure ou turnover des communautés) sont très peu influencés par le jeu de données utilisé. Concernant la diversité α , les courbes de diversité des nombres de Hill observées pour chaque jeu de données ont la même tendance et ce sont les unités taxinomiques les moins abondantes qui impactent le plus la différence entre les 3 jeux (**Figure 13**). En effet, il y a plus de différence de diversité entre les 3 jeux de données à $q=0$ où c'est seulement la richesse qui est prise en compte en comparaison à $q=2$, équivalent de l'inverse de Simpson où les unités taxinomiques les plus abondantes ont le plus de poids dans l'estimation de la diversité α . Concernant la diversité β (structure ou turnover), les coefficients RV comparant la concordance entre les structures des communautés des 3 jeux de données indiquent une forte similitude (coefficients RV supérieurs à 0.9, **Tableau 6**). Le peu d'impact observé par le choix d'un des jeux de données dans le cas du marqueur 12S s'explique par le fait que les unités taxinomiques qui expliquent ces patrons de diversité α et β sont majoritairement au rang d'espèce, qui sont aussi associées aux séquences les plus nombreuses et abondantes (**Figure 12**).

Pour le marqueur 16S, les patrons de diversité α et de diversité β sont plus impactés par le choix du jeu de données utilisé. En effet, pour la diversité α , les courbes de diversité n'ont pas la même tendance (**Figure 13**). Par exemple, dans le cas du jeu basé sur les assignations avec une méthode conservatrice, la courbe diminue moins fortement que pour les 2 autres jeux indiquant que les unités taxinomiques de ce jeu ont relativement un même poids. Concernant les patrons de diversité β , les coefficients RV sont plus faibles que pour le marqueur 12S, principalement pour les jeux de données prenant en compte l'abondance ($0.6 < \text{coefficient RV des jeux de données en abondance} < 0.7$, **Tableau 6**), ce qui indique aussi que les structures des communautés sont plus impactées par le choix du jeu de données. Cela s'explique par le fait que contrairement au marqueur 12S, les OTUs les plus abondants sont ceux assignés à la classe des Actinopterygii (**Figure 12**) et ce sont ces derniers qui sont le plus impactés par les approches utilisées sur les jeux de données (voir explication ci-dessus).

Compte-tenu des résultats issus de ces comparaisons entre jeux de données, il a été fait le choix de comparer les résultats obtenus par les comptages visuels avec les résultats obtenus par metabarcoding en considérant le jeu de données d'OTUs (*i.e.*, analyses sans assignation taxinomique). En effet :

- La majorité des espèces sont représentées par un unique OTU unique (cf. **Tableau 3**). La correspondance avec l'espèce peut donc être *in fine* facilement réalisée.

- Les approches basées sur les assignations ou sur la diversité moléculaire (OTUs) présentent des patrons de diversité β très proches (cf. **Tableau 6**).
- Les biais potentiels dus aux lacunes des bases de données (*i.e.*, absence de séquences de référence) sont absents avec l'utilisation du jeu de données de diversité des OTUs.

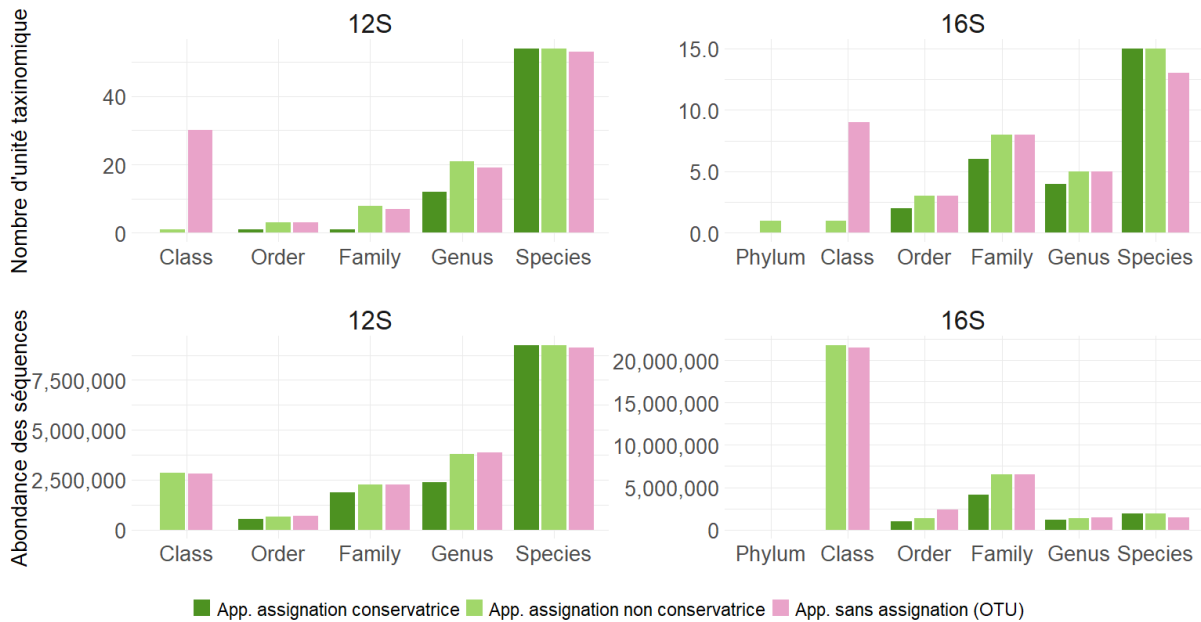


Figure 12. Résumé des assignations taxinomiques pour les 2 marqueurs des jeux de données obtenus à partir des 3 approches testées. Les figures du haut représentent le nombre d'unités taxinomiques assignées à un des rangs taxonomiques pour chaque approche et les figures du bas leur abondance.

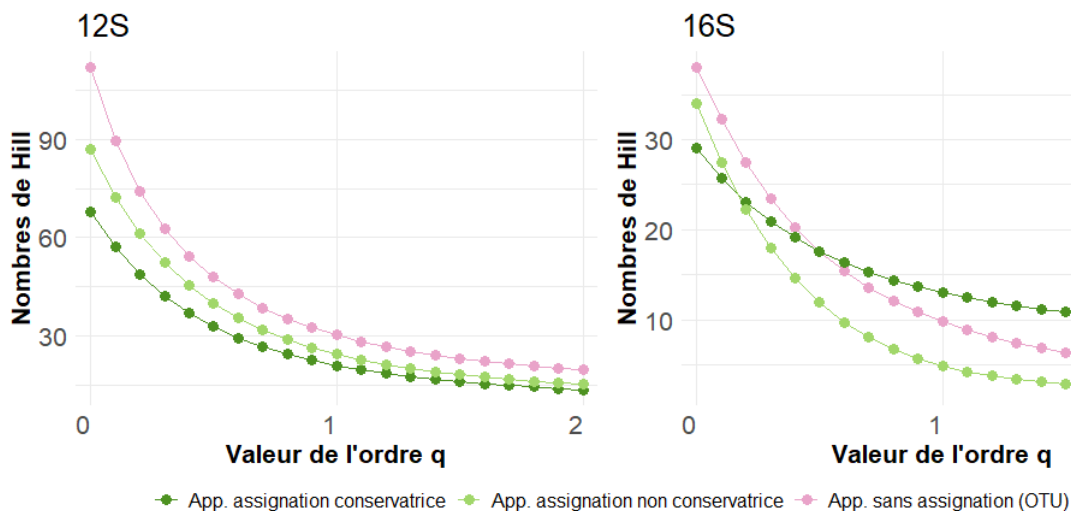


Figure 13. Indices de diversité α basés sur les différents ordres des nombres de Hill pour les matrices de communautés provenant des 3 approches pour les 2 marqueurs. Plus l'ordre est élevé, plus de poids est donné aux unités taxinomiques les plus abondantes. L'ordre $q=0$ représente le nombre d'unité taxinomiques (*i.e.* richesse), $q=1$ représente l'indice de Shannon et $q=2$ représente l'inverse de l'indice de Simpson.

Tableau 6. Coefficients RV des analyses de co-inertie en fonction des comparaisons de matrices de communautés provenant des 3 approches pour les 2 marqueurs. Les analyses de co-inertie sont faites sur les données de présence-absence et d'abondances à partir d'Analyses en Composantes Principales (ACP). Pour cela, les matrices de communautés ont été transformées avec la méthode Hellinger (Legendre and Legendre, 2012). Le coefficient RV est un coefficient de corrélation (qui varie entre 0 et 1) entre deux tables X et Y (ici correspondant aux jeux de données comparés). Plus ce coefficient est proche de 1, plus la corrélation entre les jeux de données est forte. Les valeurs au-dessus de la diagonale grise représentent les coefficients RV des comparaisons des jeux de données en abondance et celles en dessous de la diagonale grise représentent les coefficients RV des comparaisons des jeux de données en présence-absence.

12S	App. assignation conservatrice	App. assignation non conservatrice	App. sans assignation (OTU)
App. assignation conservatrice		0.96	0.95
App. assignation non conservatrice	0.96		0.98
App. sans assignation (OTU)	0.92	0.95	
16S	App. assignation conservatrice	App. assignation non conservatrice	App. sans assignation (OTU)
App. assignation conservatrice		0.72	0.61
App. assignation non conservatrice	0.91		0.73
App. sans assignation (OTU)	0.81	0.87	

4.3.2 Analyse comparative des méthodes de détection

4.3.2.1 Evaluation de l'effort d'échantillonnage

Aucune des stratégies d'échantillonnage d'eau pour l'ADNe ne permet d'atteindre un plateau pour la détection des OTUs avec les jeux de données issus du marqueur 12S (**Figure 14 A** et **Figure 15 A**). Concernant ceux du marqueurs 16S, le plateau de détection des OTUs est observé pour les échantillonnages « 30L area surface » en Automne et « 2L fixed bottom » aux deux saisons. Cette différence entre marqueur peut s'expliquer par la faible diversité détectée avec le 16S en comparaison au 12S. De plus, les intervalles de confiance de chaque courbe se superposent ce qui indique qu'il n'y a pas de type de prélèvement d'eau accumulant plus d'OTU qu'un autre. Par ailleurs, les différents types de prélèvement d'eau ont détecté la même diversité d'OTU. Enfin, les OTUs non partagés par les 3 types de prélèvement d'eau (**Figure 14 C** et **Figure 15 B**) représentent moins de 2 % des séquences. On remarquera que les prélèvements « 2L fixed bottom » sont ceux associés au moindre nombre d'OTUs et qui possèdent le moins d'OTUs uniques. Cependant la différence avec les autres types de prélèvement d'eau reste très faible. Ainsi, les trois stratégies d'échantillonnage d'eau sont donc similaires en termes de diversité d'OTUs représentée.

Concernant les méthodes comptage visuel, le jeu 24h étant un « subset » du jeu Baie, le but n'est pas de les comparer mais d'observer que dans les deux cas aucun plateau n'est atteint (**Figure 14**) suggérant que tout comme pour les types de prélèvement d'eau pour l'ADNe, un effort supplémentaire d'échantillonnage serait nécessaire si l'on souhaite obtenir un inventaire plus proche de l'exhaustivité.

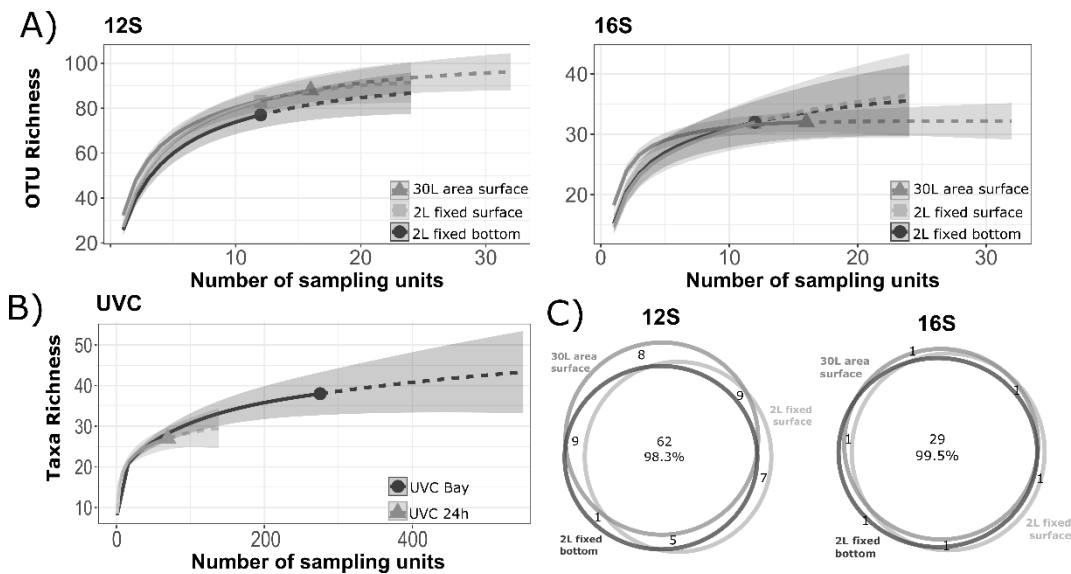


Figure 14. Les panneaux A) et B) représentent les courbes de raréfaction montrant le lien entre le nombre d'OTUs (OTU richness) pour chaque type de prélèvement d'eau pour l'ADNe en Automne (A) ou le nombre de taxons (taxa richness) pour les comptages visuels (B) et le nombre de sites échantillonnés. Un intervalle de confiance de 95% autour de chaque courbe est représenté ainsi que l'extrapolation (en pointillés) de la richesse si le nombre d'échantillon était doublé. Le panneau C) représente le nombre d'OTUs partagés entre les types de prélèvement d'eau pour l'ADNe en Automne. Le pourcentage représente le nombre de séquences partagées par les 3 types de prélèvement d'eau en fonction du nombre total de séquences.

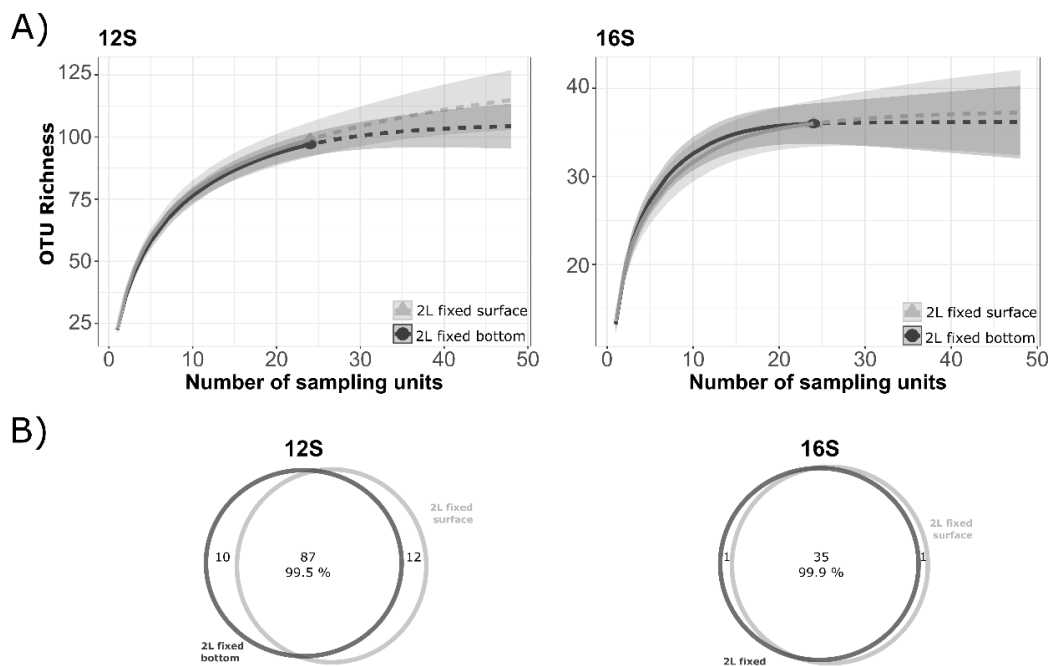


Figure 15. A) Courbes de raréfaction montrant le lien entre la richesse d'OTU pour les types de prélèvement d'eau 2L aux deux saisons d'échantillonnage et le nombre de sites échantillonnés. Un intervalle de confiance de 95% autour de chaque courbe est représenté ainsi que l'extrapolation en pointillée de la richesse d'OTU si le nombre d'échantillon était doublé. B)

Nombre d'OTUs partagés entre les types de prélèvement d'eau 2L aux deux saisons d'échantillonnage. Le pourcentage représente le nombre de séquences partagées par les 2 types de prélèvement d'eau en fonction du nombre total.

Lorsque l'on regarde à une échelle plus locale (*i.e.* au sein de chaque site), l'accumulation de la richesse en OTUs par catégorie de préférence d'habitat des taxons correspondant à ces OTUs, chaque type de prélèvement détecte une richesse relativement similaire pour les taxons associés aux habitats rocheux et benthopélagiques (**Figure 16**). Cependant, dans le cas des taxons associés aux habitats démersaux et non associés au substrat rocheux, les prélèvements « 30L area surface » détectent une richesse plus importante que les deux types de prélèvements « 2L », et ce principalement dans les sites proches de la côte (Corbeau et Figuier) qui sont caractérisés par une plus grande diversité d'habitats côtiers (notamment présence d'herbiers à *Zoostera noltii*, et proximité d'estuaires). Il en est de même pour les taxons pélagiques mais principalement dans les sites plus au large (Astan et Cochons Noirs) propices aux poissons pélagiques. Ainsi les différences entre les méthodes 2L et la méthode « 30L area surface » résideraient notamment dans le fait que le prélèvement « 30L area surface » réalisé sur de plus grandes surfaces couvrent plusieurs habitats. La méthode « 30L area surface » apparaît donc intéressante dans l'objectif d'évaluer la diversité ichthyologique côtière à l'échelle du paysage (et non des différents habitats le composant). Les méthodes 2L, par leur échantillonnage ponctuel, seraient potentiellement plus adaptées au suivi de l'ichtyofaune à des échelles spatiales plus fines, pour éventuellement mieux discriminer entre habitats les patrons de diversité ichthyologiques.

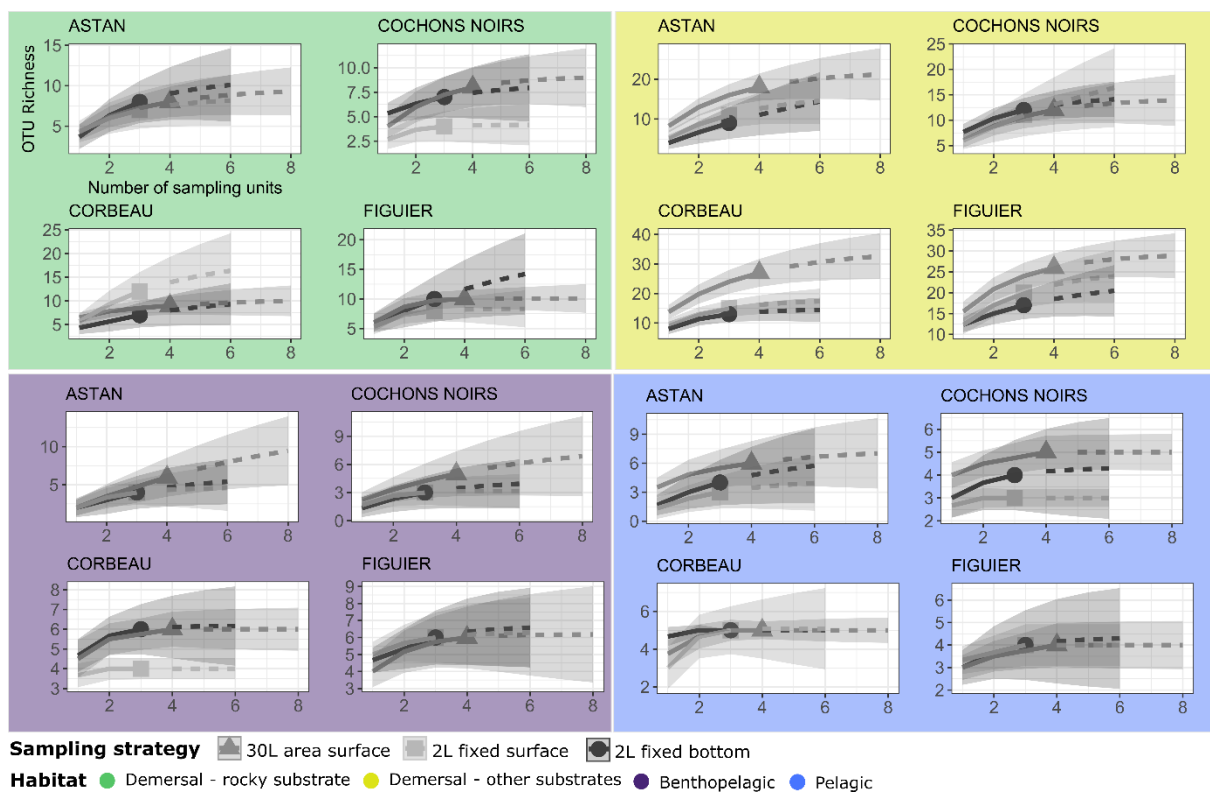


Figure 16. Courbes de raréfaction montrant le lien entre la richesse d'OTU pour chaque type de prélèvement d'eau en Automne par site et par catégorie de préférence d'habitats des taxons assignés aux OTUs. Un intervalle de confiance de 95% autour de chaque courbe est représenté ainsi que l'extrapolation en pointillée de la richesse d'OTU ou taxinomique si le nombre d'échantillon était doublé.

4.3.2.2 Comparaison des listes de taxons identifiés

Au total 18 taxons, appartenant aux actinoptérygiens et aux élasmobranches observés en plongée dans les mêmes sites et à la même session d'échantillonnage que pour l'ADNe (« UVC 24h »), ont été identifiés par metabarcoding sur ADNe (**Figure 17 A**). De plus, 5 espèces (*Dicentrarchus labrax*, *Symphodus bailloni*, *Torpedo marmorata*, *Lophius piscatorius* et *Lepadogaster candolii*) qui ont été observées dans d'autres lieux ou à d'autres sessions de plongée (« UVC Baie ») que ceux concordant à la période d'échantillonnage pour l'ADNe ont été également détectées par metabarcoding sur ADNe. Ces 5 espèces ont été observées moins de 2 fois sur les 34 observations en plongée (site-saison-session) combinant les 2 méthodes de comptage visuel. Ce résultat indique que le metabarcoding sur ADNe a permis la détection d'espèce à occurrence rare en comptage visuel (**Tableau 7**). Néanmoins, des faux-négatifs (taxons observés en plongée mais non identifiés par metabarcoding) sont présents dans les données de metabarcoding (**Figure 17 A, B** et **Tableau 7**). Les 7 espèces observées seulement dans les comptages visuels du jeu « UVC 24h » et non détectées avec le metabarcoding ne possèdent pas de séquences de référence couvrant la région ciblée par les amorces 12S/16S utilisées ou permettant de discriminer uniquement cette espèce dans la base de données de référence (**Tableau 7**). Par exemple, pour les espèces fréquemment observées en comptage telles que *Pollachius pollachius* et *Parablennius gattorugine*, le travail effectué pour compléter les bases de référence a permis l'obtention de séquences 16S mais qui n'ont pas permis de discriminer ces espèces d'autres espèces proches. Concernant *Conger conger* et *Labrus bergylta*, des OTU ont été assignés à *Conger* et *Labrus* spp. avec le metabarcoding. Néanmoins, ces OTUs correspondent très probablement aux espèces identifiées dans les comptages visuels. En effet, il n'est recensé qu'une seule espèce du genre *Conger* dans la zone étudiée. Pour *Labrus bergylta*, des séquences ont été assignées à *Labrus* spp. car il y a 100% de similarité entre les séquences 12S de *L. bergylta*, *L. merula* et *L. viridis* mais les deux dernières sont absentes de la zone étudiée. Ces résultats sont intéressants car ils indiquent que, pour ces 7 espèces observées en UVC mais non identifiées par metabarcoding, des améliorations (compléter la base de référence ou choisir un autre marqueur qui permettrait l'assignation à l'espèce) pourraient permettre leur détection à l'avenir. Par exemple, dans le cas de *Pollachius pollachius*, l'utilisation du couple d'amorces de Miya et al. (2015) ciblant une région du 12S mitochondrial aurait probablement permis sa détection (**Tableau 4 en annexe**).

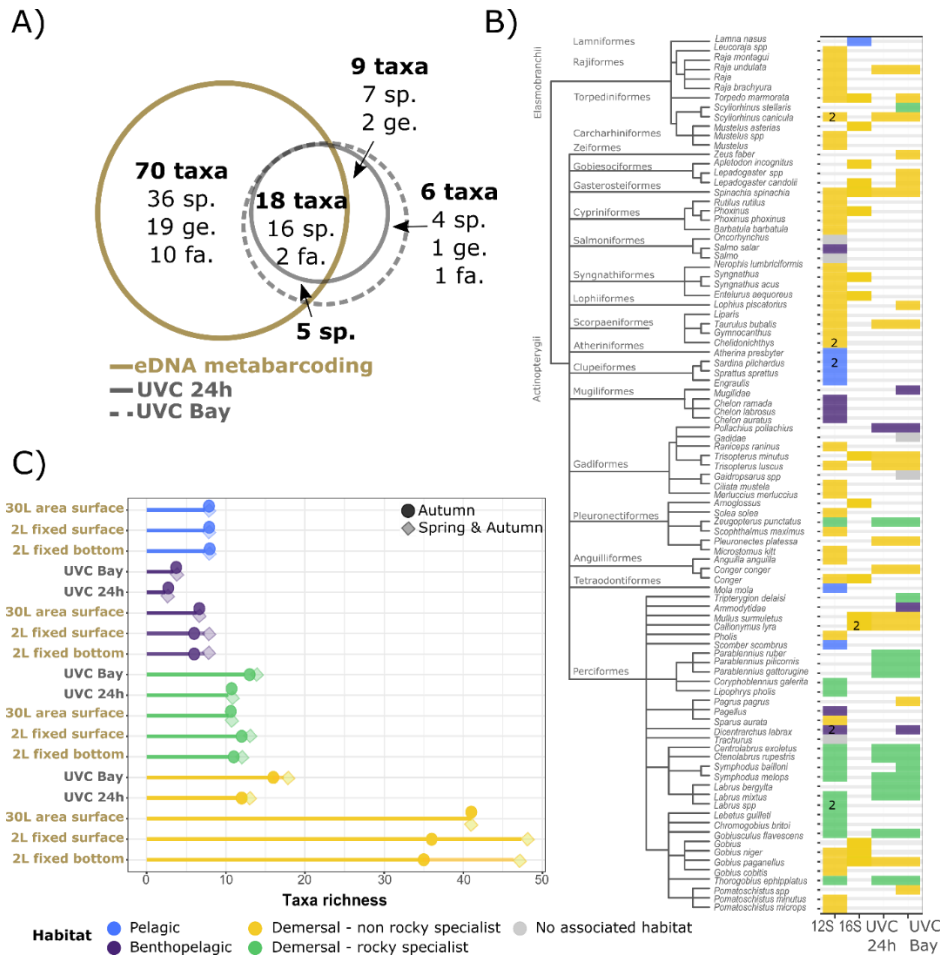


Figure 17. Synthèse des taxons identifiés par les différentes approches et types de prélèvements A) Diagramme de Venn comparant les taxons détectés par metabarcoding (cumul des résultats sur l'ensemble des types de prélèvements, marqueurs et saisons) et par les comptages visuels (cumulés sur UVC Baie et UVC 24h). Pour chaque aire du diagramme est noté le nombre total de taxons observés en gras et, en dessous, le nombre de taxons identifiés à la famille, au genre et à l'espèce. B) Cladogramme taxinomique montrant les taxons identifiés au genre ou à l'espèce par metabarcoding avec le marqueur 12S, par metabarcoding avec le marqueur 16S, par comptage visuel Baie et par comptage visuel 24h. Chaque couleur représente la préférence d'habitat du taxon quand elle est connue. Le nombre d'OTU par taxon et par marqueur est indiqué s'il dépasse la valeur de 1. C) Nombre de taxons par préférence d'habitat en fonction des stratégies d'échantillonnage. Les ronds foncés représentent le jeu de données à l'Automne uniquement et les losanges clairs le jeu de données aux deux saisons.

Tableau 7. Disponibilité de séquences de référence pour chacun des taxons (au genre ou à l'espèce) identifiés par comptages visuels. Trois statuts sont décrits dans les colonnes des marqueurs 12S et 16S : NON (rouge) : pas de séquence de référence disponible, OUI (vert) : séquence diagnostique (i.e. permettant l'identification de l'espèce) et ND (orange) : séquence disponible mais non diagnostique. Dans le cas où les séquences de référence n'étaient pas disponibles au moment de la création de la base de données en Mai 2020, nous avons vérifié si c'était toujours le cas au moment de la rédaction du rapport (Mai 2021). Il n'y a pas eu d'ajout de séquence dans la base publique NCBI depuis Mai 2020 pour les espèces sans séquences à l'exception de *Mullus surmuletus* où un génome mitochondrial complet a été rajouté, rendant donc disponible une séquence de référence pour les marqueurs 12S et 16S.

Taxon vu en comptage visuel	Nombre d'échantillon dans UVC Baie (total 34)	Méthodes d'observation			Disponibilité de séquences de référence	
		UVC Baie	UVC 24h	ADNe	12S	16S
<i>Scyliorhinus stellaris</i>	1	OUI	NON	NON	NON	NON
<i>Lepadogaster candolii</i>	1	OUI	NON	OUI	NON	OUI

<i>Lophius piscatorius</i>	1	OUI	NON	OUI	OUI	ND
<i>Symphodus bailloni</i>	1	OUI	NON	OUI	OUI	ND
<i>Parablennius ruber</i>	1	OUI	OUI	NON	NON	ND
<i>Pleuronectes platessa</i>	1	OUI	OUI	NON	ND	ND
<i>Pagrus pagrus</i>	2	OUI	NON	NON	ND	ND
<i>Dicentrarchus labrax</i>	2	OUI	NON	OUI	OUI	ND
<i>Torpedo marmorata</i>	2	OUI	NON	OUI	OUI	OUI
<i>Gaidropsarus spp</i>	3	OUI	NON	NON	NON	ND
<i>Tripterygion delaisi</i>	3	OUI	NON	NON	OUI	ND
<i>Zeus faber</i>	3	OUI	NON	NON	OUI	ND
<i>Raja undulata</i>	3	OUI	OUI	OUI	OUI	ND
<i>Spinachia spinachia</i>	4	OUI	OUI	OUI	OUI	OUI
<i>Taurulus bubalis</i>	4	OUI	OUI	OUI	OUI	ND
<i>Zeugopterus punctatus</i>	5	OUI	OUI	OUI	OUI	ND
<i>Scyliorhinus canicula</i>	12	OUI	OUI	OUI	OUI	ND
<i>Mullus surmuletus</i>	14	OUI	OUI	OUI	NON	OUI
<i>Lepadogaster spp</i>	15	OUI	OUI	NON	OUI (L. <i>lepadogaster</i>)	ND (L. <i>lepadogaster</i>)
<i>Parablennius pilicornis</i>	15	OUI	OUI	NON	NON	ND
<i>Conger conger</i>	19	OUI	OUI	NON	NON	ND
<i>Callionymus lyra</i>	19	OUI	OUI	OUI	NON	OUI
<i>Gobius paganellus</i>	23	OUI	OUI	OUI	OUI	OUI
<i>Pomatoschistus spp</i>	26	OUI	OUI	OUI (P. <i>minutus</i> et <i>P. microps</i>)	OUI	ND
<i>Gobiusculus flavescens</i>	28	OUI	OUI	OUI	OUI	ND
<i>Trisopterus minutus</i>	28	OUI	OUI	OUI	OUI	OUI
<i>Trisopterus luscus</i>	30	OUI	OUI	OUI	OUI	OUI
<i>Labrus mixtus</i>	31	OUI	OUI	OUI	OUI	ND
<i>Parablennius gattorugine</i>	32	OUI	OUI	NON	NON	ND
<i>Symphodus melops</i>	32	OUI	OUI	OUI	OUI	ND
<i>Labrus bergylta</i>	33	OUI	OUI	NON	ND	ND
<i>Pollachius pollachius</i>	33	OUI	OUI	NON	ND	ND
<i>Centrolabrus exoletus</i>	33	OUI	OUI	OUI	OUI	ND
<i>Ctenolabrus rupestris</i>	33	OUI	OUI	OUI	OUI	ND
<i>Thorogobius ephippiatus</i>	33	OUI	OUI	OUI	OUI	NON

Le metabarcoding n'a pas identifié toutes les espèces vues lors des comptages visuels mais, comme attendu, il a révélé un plus grand nombre de taxons (**Figure 17 A**), en particulier avec le marqueur 12S en comparaison du marqueur 16S, comme cela a déjà été évoqué en Section 4.2.3 (**Figure 17 B**).

Sur les 70 taxons uniquement détectés par le metabarcoding, 55 (dont 32 des 36 espèces) proviennent d'habitats adjacents aux sites rocheux échantillonnés, et ce quel que soit le type de prélèvement d'eau utilisé (**Figure 17 B et C**). Le metabarcoding a notamment permis la détection d'un grand nombre de taxons démersaux non exclusivement associés au milieu rocheux ainsi que de taxons vivant dans la colonne d'eau.

Concernant les 18 espèces (nb : ici on considère que *Labrus spp.* est *L. bergylta*) identifiées par metabarcoding et qui ont une affinité exclusive aux habitat rocheux, 9 sont communément observées par metabarcoding et UVC, 4 (*Lipophrys pholis*, *Chromogobius britoi*, *Coryphoblennius galerita*, *Lebetus guilleti*) sont observées uniquement par metabarcoding et 5 uniquement par UVC (les 3 espèces de *Parablennius* observées dans les données « UVC Baie » et « UVC 24h », *Scyllorhinus stellaris* et *Tripterygion delaisi* observées uniquement dans les données « UVC Baie ») (**Figure 18**). Il est intéressant de noter que les 4 espèces uniquement observées par metabarcoding sont des espèces qui peuvent être considérées comme crypto-benthiques, comme par exemple le gobie nain (*Lebetus guilleti*) qui ne mesure pas plus de 2.5cm.

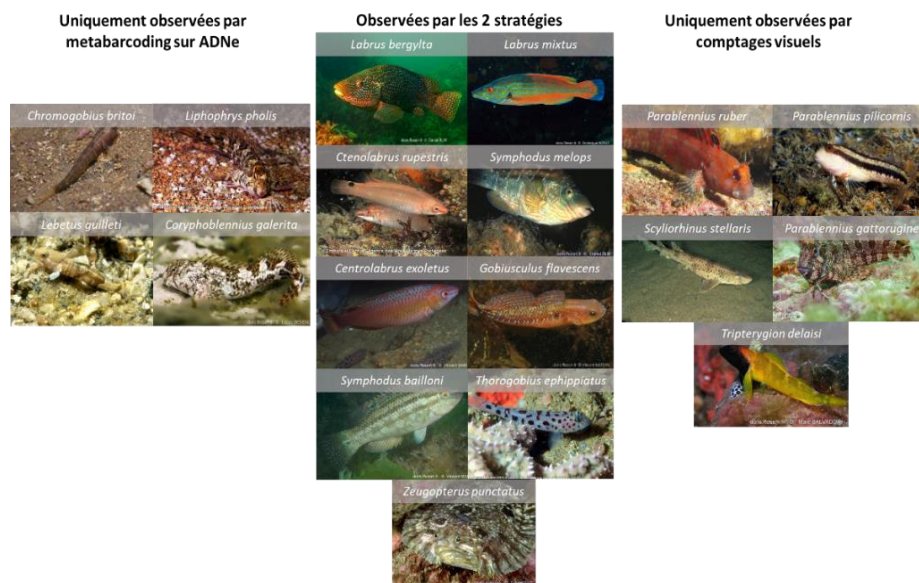


Figure 18. Espèces à affinité exclusive au substrat rocheux observées par metabarcoding et/ou UVC

4.3.2.3 Comparaison des occurrences (intra- et inter-sites) des espèces

De manière générale, les deux méthodes d'observation (UVC *vs* metabarcoding) semblent détecter de façon concordante les espèces fréquemment ou peu observées (**Figure 19**). Il existe cependant des discordances entre les deux approches concernant les espèces *Callionymus lyra* (plus fréquemment observée en ADNe qu'en UVC) et *Thorogobius ephippiatus* (plus fréquemment observée en UVC qu'en ADNe). *Callionymus lyra* est une espèce discrète vivant sur fonds sableux (source : DORIS) ce qui pourrait expliquer sa faible détection avec les comptages visuels réalisés en milieu rocheux. Les détections par ADNe concerneraient alors avant tout des individus présents dans les fonds sableux adjacents aux milieux rocheux suivis en UVC. Concernant *Thorogobius ephippiatus* qui

est une espèce associée au milieu rocheux, facilement identifiable et régulièrement observée en UVC (voir photo **Figure 18**), diverses raisons peuvent expliquer sa mauvaise détection par metabarcoding : source d'ADNe dans l'environnement faible, compétition au sein des PCRs, etc.

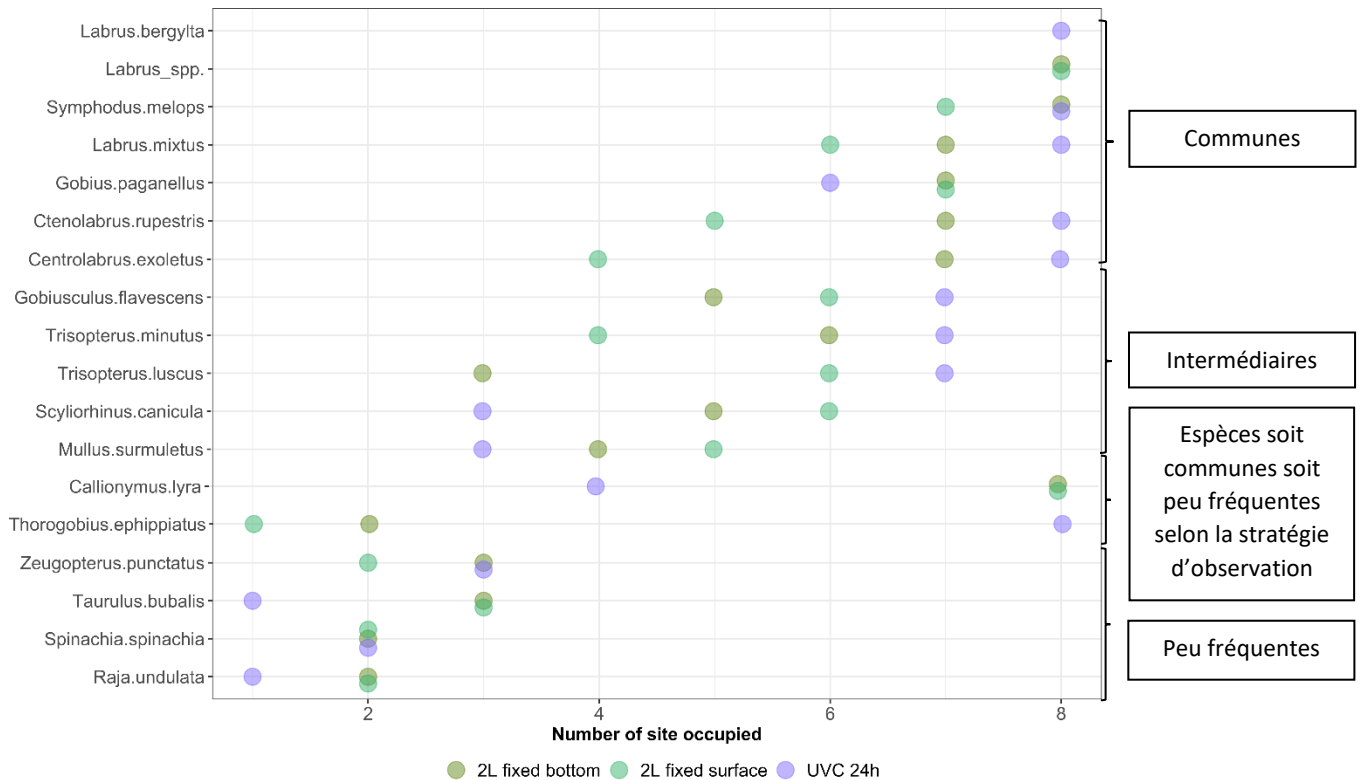


Figure 19. Nombre de sites où les espèces observées à la fois par metabarcoding (avec « ADNe 2L ») et par UVC (UVC 24h) sont détectées par chaque méthode.

Concernant les fréquences d'occurrence des taxons selon les types de prélèvement réalisés pour le metabarcoding, et comme il a pu être observé dans les figures précédentes, il semble que les 3 types de prélèvements d'eau permettent de révéler une diversité α similaire à l'échelle du jeu de données entier. Aucun de ces types de prélèvements ne semble avoir sur- ou sous-estimé la détection de taxons, et ce quelle que soit la préférence d'habitat des taxons. Ce patron est confirmé lorsque l'on compare la fréquence d'occurrence des OTUs au sein des échantillons (réplicats et sites) entre chaque type de prélèvement (**Figure 20** et **Figure 21**). Quelle que soit la préférence d'habitat du taxon assigné, les OTUs ont des fréquences d'occurrence similaires entre type de prélèvement pris deux à deux.

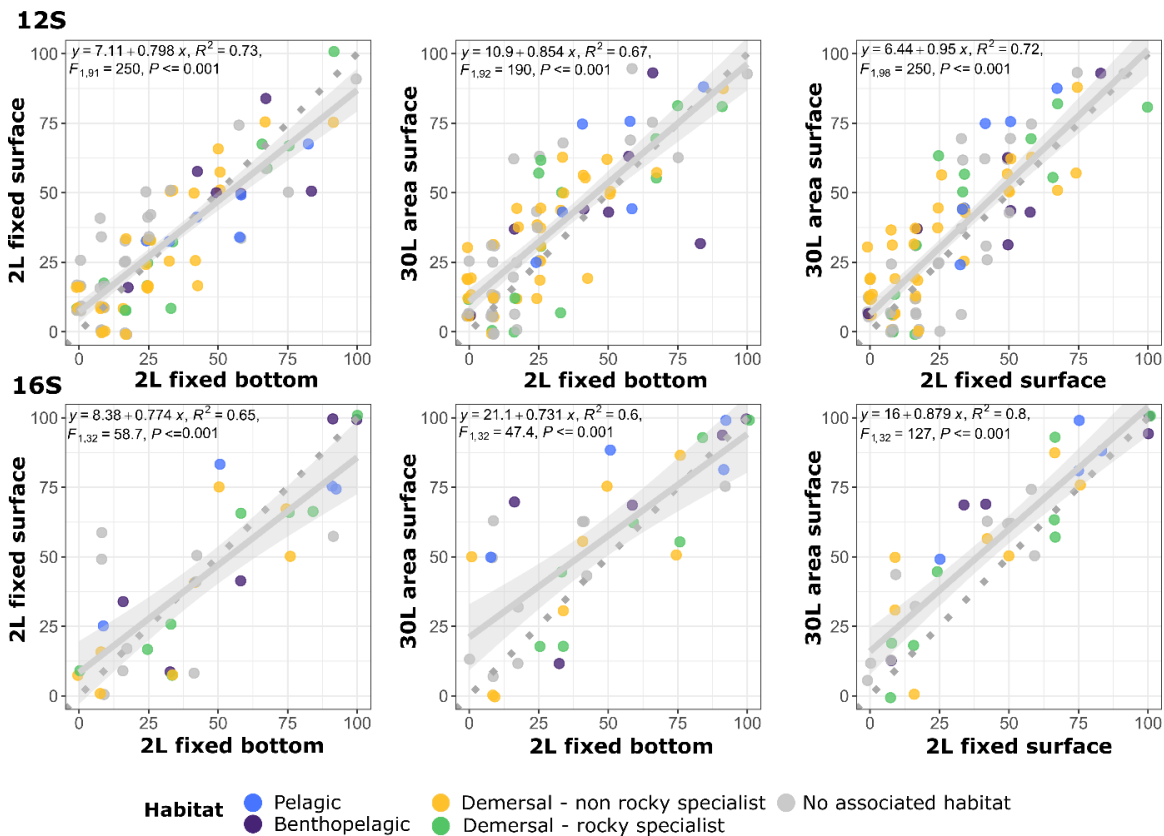


Figure 20. Relation entre les fréquences d'occurrence des OTUs entre deux types de prélèvement d'eau pour la saison Automne. Les OTUs sont colorés en fonction de la préférence d'habitat du taxon assigné à l'OTU. La ligne solide représente la droite de régression ($y \sim x$) avec son intervalle de confiance à 95%. La ligne pointillée représente la droite 1:1 le long de laquelle la fréquence de l'OTU serait égale avec les deux types de prélèvements comparés.

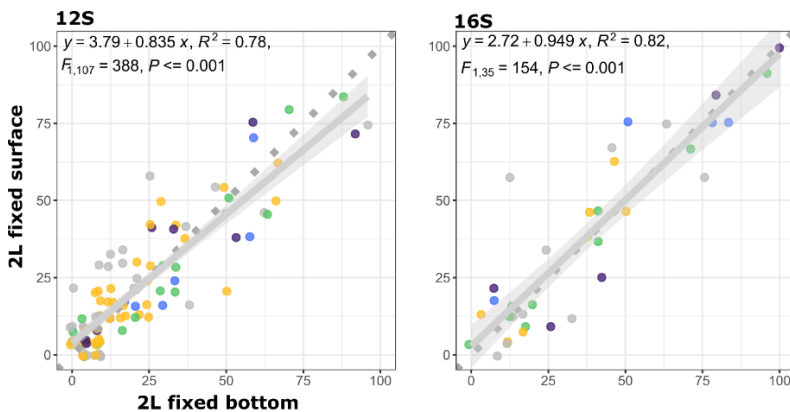


Figure 21. Relation entre les fréquences d'occurrence des OTUs entre les deux types de prélèvement (2L-fond et 2L-surface) aux deux saisons d'échantillonnage. Les OTUs sont colorés en fonction de la préférence d'habitat du taxon assigné à l'OTU. La ligne solide représente la droite de régression ($y \sim x$) avec son intervalle de confiance à 95%. La ligne pointillée représente la droite 1:1 le long de laquelle la fréquence de l'OTU serait parfaitement égale avec les deux types de prélèvements.

4.3.2.4 Comparaison des patrons de diversité β

Impact de la nature du jeu de données sur les comparaisons des patrons de diversité β

Les coefficients RV des analyses de co-inertie nous permettent d'évaluer à quel point les patrons de diversité β concordent selon l'origine des jeux de données de présence-absence (**Tableau 8**). Ces coefficients sont élevés pour les comparaisons des jeux au sein de la méthode d'observation ADNe ($0.83 < RV < 0.90$) indiquant alors que les patrons des communautés sont relativement proches entre 1) les 2 marqueurs utilisés, 2) entre les types de prélèvements. Néanmoins entre les méthodes d'observations, les coefficients RV sont modérés ($0.59 < RV < 0.68$) soulignant alors que les patrons de diversité β entre UVC et metabarcoding au niveau des sites concordent moyennement.

Tableau 8. Coefficients RV des analyses de co-inertie comparant la similarité des patrons de diversité β entre les jeux de données 12S et 16S, entre les types de prélèvements 2L ADNe et entre les méthodes d'observation. Les analyses de co-inertie sont faites à partir d'Analyses en Composantes Principales (ACP) des données de présence-absence transformées Hellinger et moyennées au niveau du site ($n = 8$; 4 sites à 2 saisons). Afin de prendre en compte les variations spatiales et temporelles, le jeu de données de metabarcoding issu des prélèvements « 30L area surface » n'a pas été pris en compte car l'échantillonnage a eu lieu en Automne uniquement. Le coefficient RV est un coefficient de corrélation (qui varie entre 0 et 1) entre deux tables X et Y (ici correspondant aux jeux de données comparés). Plus ce coefficient est proche de 1, plus la corrélation entre les jeux de données est forte.

Comparaison effectuée avec l'analyse de co-inertie	Metabarcoding				Coefficient RV	
	UVC	12S		16S		
		2L-surface	2L-bottom	2L-surface		2L-bottom
Marqueurs : 12S vs 16S		x		x	0.89	
			x		x	0.89
Types de prélèvement : 2L-surface vs bottom		x	x		0.90	
				x	x	0.83
Méthodes d'observation : UVC vs Metabarcoding	x	x			0.66	
	x		x		0.68	
	x			x	0.59	
	x			x	0.61	

Evaluation de la résolution spatiale et temporelle

Pour les données de metabarcoding, traitées en présence-absence d'OTUs, l'ordination non contrainte met en avant avec les 2 premiers axes (21.2% de la variation totale) l'importance de la saison et des sites, avec l'observation d'un gradient côte-large (Corbeau – Figuier vs Astan – Cochons Noirs) dans les patrons de communautés des assemblages ichthyologiques (**Figure 22 B**). Les analyses de PERMANOVA présentant des valeurs de test significatives confirment ce patron (**Annexe - tableau 5**). Ces résultats suggèrent donc que le metabarcoding permet de distinguer des communautés à des résolutions spatiale et temporelle (ici : saisons) relativement fines. Il faut aussi noter que le type de prélèvement a une influence minimale sur les principaux profils observés, même si dans le cas des sites plus au large (Astan – Cochons Noirs), les structures révélées avec les prélèvements de 30L semblent

plus marquées que celles observées avec les prélèvements de 2L. En accord avec l'ACP et les analyses de co-inertie (**Tableau 8**), les analyses de PERMANOVA indiquent que l'impact du type de prélèvement sur les patrons de communautés reste faible en comparaison à la saisonnalité et aux sites (**Annexe - tableau 5**). Les principaux profils identifiés en UVC, également en présence-absence de taxons, sont en revanche un peu différents de ceux identifiés avec les données de metabarcoding ce qui explique la concordance modérée entre les 2 méthodes d'observations observée avec les analyses de co-inertie (**Tableau 8**). En effet, la différence en UVC est surtout observée entre les sites (**Figure 22 A**), cependant les analyses de PERMANOVA confirment comme avec les données de metabarcoding, l'importance des effets site et saisons sur les communautés (**Annexe - tableau 5**). Lorsque nous regardons les patrons de communauté en commun identifiés par la co-inertie entre les 2 méthodes d'observation, le signal spatial gradient côte-large est mis en évidence (**Figure 23**). Sur le premier axe de la co-inertie représentant 57% de la variation totale, Figuiers (site à la côte) est le plus éloigné d'Astan et Cochons Noirs (sites au large) et Corbeau (site à la côte) présente une position intermédiaire. De plus, les taxons partagés entre UVC et metabarcoding montrent la plupart du temps des variations similaires.

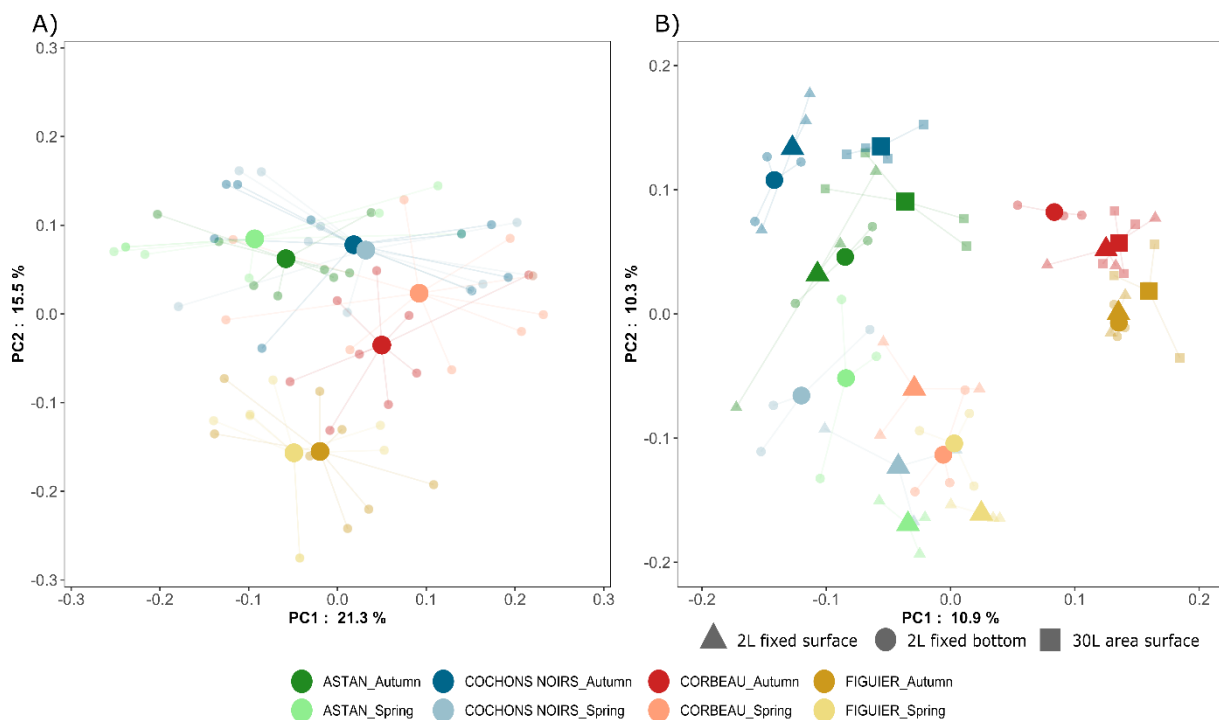


Figure 22. Analyses en Composantes Principales des données (A) de présence-absence transformées Hellinger pour la matrice de communauté basée sur les taxons observés en UVC 24h, et (B) pour la matrice de communauté basée sur les OTUs observés en metabarcoding avec le marqueur 12S. Les 2 premiers axes représentent 36.8% et 21.2% de la variation totale de la communauté respectivement pour le jeu UVC et le jeu metabarcoding. Les ACP sont représentées en scaling 1 afin de préserver les distances entre les sites. Les grands symboles représentent les positions des centroïdes des transects pour le jeu UVC et des répliquats pour le jeu metabarcoding au niveau du site échantillonné, et pour chaque type de prélèvement pour le jeu metabarcoding seulement.

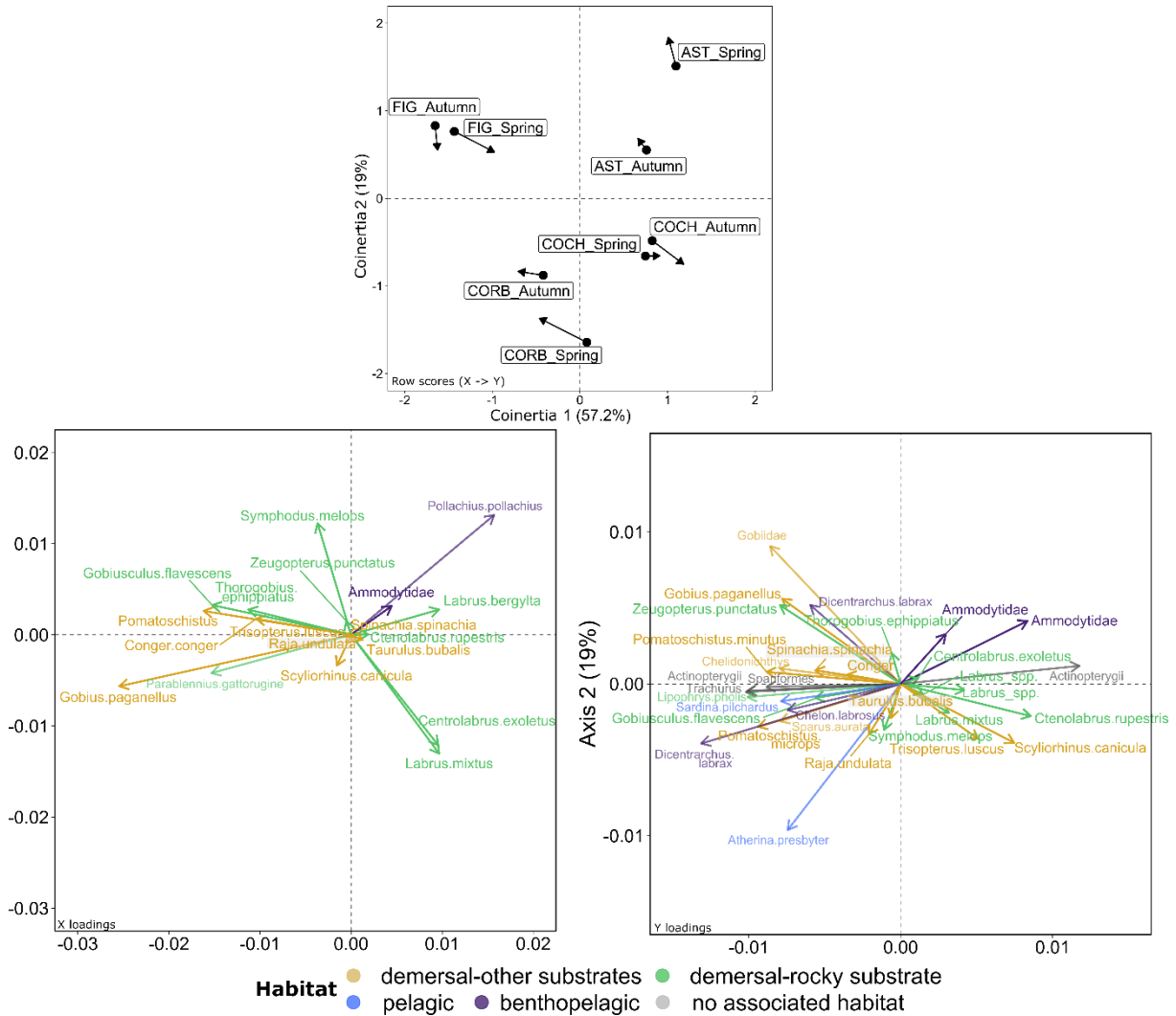


Figure 23. Représentation graphique de l'analyse de co-inertie présentée en Tableau 8. Coefficients RV des analyses de co-inertie comparant la similarité des patrons de diversité β entre les jeux de données 12S et 16S, entre les types de prélèvements 2L ADNe et entre les méthodes d'observation. Les analyses de co-inertie sont faites à partir d'Analyses en Composantes Principales (ACP) des données de présence-absence transformées Hellinger et moyennées au niveau du site ($n = 8$; 4 sites à 2 saisons). Afin de prendre en compte les variations spatiales et temporelles, le jeu de données de metabarcoding issu des prélèvements « 30L area surface » n'a pas été pris en compte car l'échantillonnage a eu lieu en Automne uniquement. Le coefficient RV est un coefficient de corrélation (qui varie entre 0 et 1) entre deux tables X et Y (ici correspondant aux jeux de données comparés). Plus ce coefficient est proche de 1, plus la corrélation entre les jeux de données est forte. comparant les patrons de diversité β entre les 2 méthodes d'observation. Le metabarcoding est représenté par le jeu de données 12S « 2L fixed bottom » et les UVC par le jeu de données 24h. Les 2 premiers axes sont montrés et représentent 76,2% de la variance totale expliquée par la co-inertie dont le coefficient RV est de 0.66. Le plot du dessus indique les positions des sites sur les axes de la co-inertie dont le symbole rond représente le jeu UVC et la tête de flèche le jeu metabarcoding. Plus la flèche est courte, plus la correspondance entre les 2 projections concorde. Les abréviations des sites sont les suivantes : AST pour ASTAN, COCH pour COCHONS NOIRS, CORB pour CORBEAU, FIG pour FIGUIER. Les 2 plot du dessous indiquent la contribution des groupes de variables (Taxons des UVC sur la gauche et OTUs sur la droite). Les vecteurs pointant dans la même direction sont corrélés et plus les vecteurs sont longs, plus ils contribuent à la structure. Les taxons (à gauche) et les OTU (à droite) sont colorés en fonction de leur habitat associé. Seuls les taxons partagés par les 2 méthodes d'observation (couleur foncée et grande taille) et les 20% des taxons qui avaient la longueur de vecteur la plus haute (couleur claire, petite taille) sont représentés.

Contributions des taxons à la diversité β

Concernant la communauté observée avec le metabarcoding, la contribution des OTUs à la diversité β en fonction de leur préférence d'habitat est similaire entre les 2 types de prélèvement 2L (**Figure 24**). Les OTUs assignés à des taxons associés à l'habitat pélagique présentent les plus hauts scores de SCBD (« Species Contribution to β diversity ») dont *Atherina presbyter* associés aux OTUs ayant les scores les plus élevés. Les OTUs assignés à des taxons démersaux non spécialistes du substrat rocheux et les OTUs dont les taxons n'ont pas d'habitat associé ont les plus faibles scores médians de SCBD, suggérant que la plupart de ces OTUs sont plus rarement observés. Néanmoins certains de ces OTUs ont contribué fortement à la diversité β tels que la petite roussette *Scyliorhinus canicula*, espèce démersale non spécialiste du substrat rocheux ou bien l'OTU assigné à *Oncorhynchus* (sans habitat associé). Lorsque l'on compare les 2 méthodes d'observation (UVC vs metabarcoding), les scores de SCBD en fonction des préférences d'habitat des taxons diffèrent (**Figure 24**). Par exemple, les taxons démersaux spécialistes du substrat rocheux contribuent le plus à la diversité β de la communauté observée par les UVC, ce qui n'est pas le cas avec le metabarcoding. Cependant, certaines espèces contribuent fortement à la diversité β des communautés observées en metabarcoding et en UVC comme par exemple *Gobius paganellus*, *Trisopterus luscus* et *Symphodus melops*.

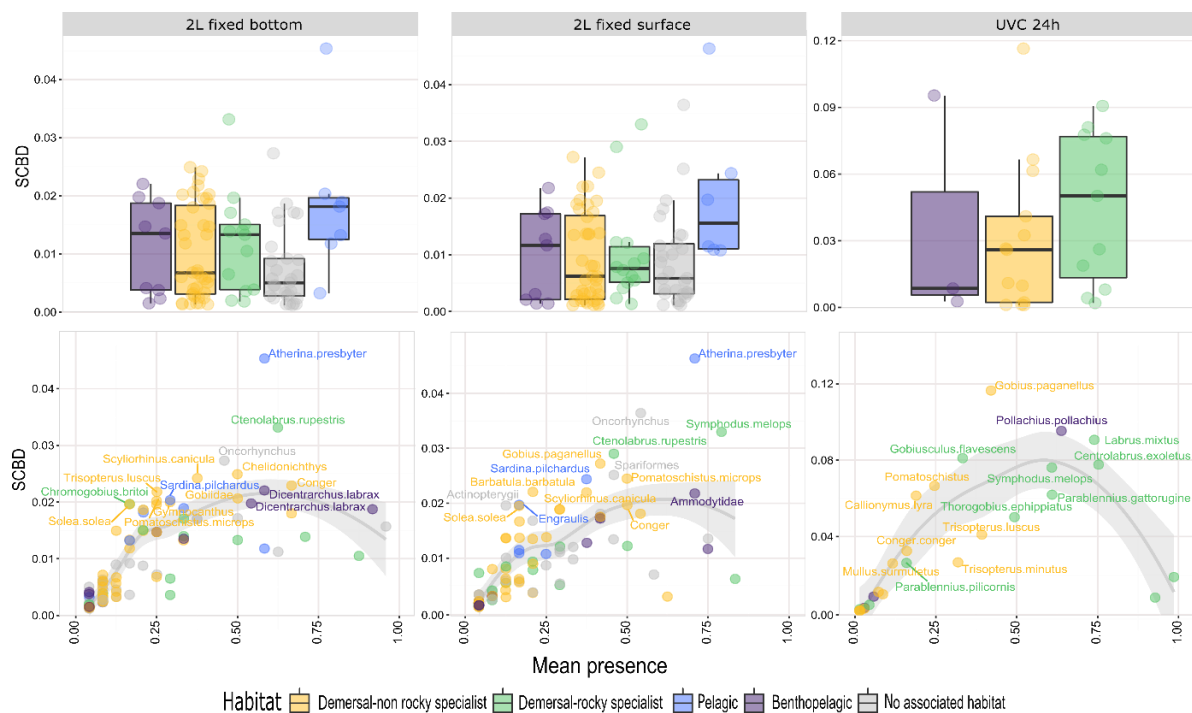


Figure 24. Analyse « Species Contribution to β diversity » effectuée sur les matrices de communautés de présence-absence transformées Hellinger et moyennées au niveau du site ($n = 8$; 4 sites à 2 saisons) pour les 2 types de prélèvements 2L avec le marqueur 12S et pour le jeu de données UVC 24h. (A) Boîte à moustache des valeurs de SCBD pour chaque OTU (jeu de données metabarcoding) et taxons (jeu de données UVC) en fonction de la préférence d'habitat du taxon. (B) Présence moyenne de chaque OTU (pour les données de metabarcoding) ou taxon (pour les données UVC 24h) par rapport au nombre total de répliqués (pour les données de metabarcoding) ou transects (pour les données UVC 24h) en fonction de sa valeur de SCBD. Les 15 OTUs (pour les données de metabarcoding) ou taxon (pour les données UVC 24h) avec les plus fortes valeurs de SCBD sont indiqués sur la figure. Afin de prendre en compte les variations spatiales et temporelles, le jeu de données ADNe « 30L area surface » n'a pas été pris en compte car l'échantillonnage a eu lieu en Automne uniquement.

4.4 La surveillance active par ddPCR

4.4.1 Sélection des espèces cibles

La première étape était de déterminer si nous avions des espèces candidates pour une étude par ddPCR. Pour cela nous avons élaboré une grille d'évaluation basée sur 5 critères (cf. 3.3.4). Le croisement des critères 1 à 5 a limité les possibilités de sélection à un petit nombre d'espèces (<12). Des alignements ont ensuite été réalisés pour déterminer la pertinence de l'approche, c'est-à-dire la présence de sites diagnostiques de l'espèce cible par rapport à des espèces proches, tel qu'illustré avec le cas de l'espèce cible *Syngnathus acus* dans la **Figure 25**.

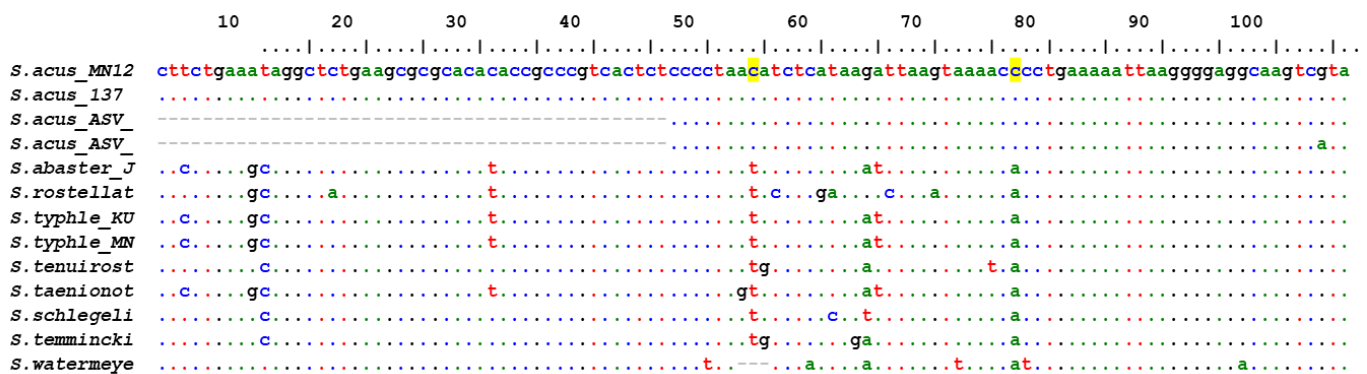


Figure 25. Alignement des séquences 12S sur la portion analysée en metabarcoding obtenu pour la cible *Syngnathus acus* et d'autres espèces du même genre. Deux sites (surlignés en jaune dans la première séquence de l'espèce cible) sont spécifiques de cette espèce, et utilisés pour définir des sondes spécifiques.

A l'issue de ce travail un total de 6 espèces candidates (représentant 7 taxons car les deux lignées, Atlantiques et Méditerranéennes, de *Dicentrarchus labrax* ont été incluses) ont été sélectionnées pour une première phase de design et validation des amorces et sondes. Des amorces et sondes ont pu être définies et validées *in silico* puis de façon empirique pour chacune des 6 espèces, comme indiqué dans le **Tableau 9**. Trois multiplexes permettent d'analyser les résultats des hybridations des 7 sondes. Les ADNe (n=51) obtenus à partir des prélèvements de 2L-fond et 2L-surface ont été analysés par la société IAGE.

Il faut noter que les espèces étudiées par ddPCR nous permettent de répondre à plusieurs objectifs : **(1)** évaluer la possibilité d'obtenir des données quantitatives d'abondance par rapport au metabarcoding dans le cas d'espèces abondantes ou d'abondance variable telles que vues en UVC (ex. centrolabre, tacaud), **(2)** déterminer l'apport de prélèvements d'ADNe par rapport aux approches UVC, pour une estimation des abondances d'espèces cryptiques/cachées (ex. targeur), furtives (ex. bar) ou présentes dans des habitats adjacents (ex. grand syngnathe), et **(3)** analyser le potentiel d'utilisation en routine de ddPCR pour des espèces d'intérêt car objets de mesures de gestion/conservation ou patrimoniales (ex. bar, raie brunette).

Tableau 9. Espèces cibles et sondes validées pour l'analyse par ddPCR. Les critères de choix utilisés sont rappelés. « ADN cible » indique la disponibilité et le nombre d'ADN de la cible disponibles (au laboratoire de l'équipe projet) pour des tests de validation. « ADN esp. proche » liste les espèces proches de la cible pour lesquelles des ADN sont également disponibles. « Bioinfo » correspond au critère de disponibilité de séquences pour la cible et des espèces proches, ainsi que la présence de sites spécifiques de la cible (cf. **Figure 25**). « Mult » indique le multiplex associant les sondes indiquées dans la colonne « Sonde ». « Metabar. » indique la disponibilité d'échantillons ADN analysés par metabarcoding qui ont détecté ou pas l'espèce cible. « Surv. » définit l'intérêt de l'espèce dans un contexte de surveillance.

Espèce	ADN cible	ADN Esp. proches	Bio-info	Mult.	Sonde	Meta bar.	Surv.
Centrolabrus exoletus (centrolabre)	OUI (1)	<i>Symphodus bailloni</i> <i>Symphodus melops</i>	OUI	Duplex 1	HEX-0108	OUI	Espèce abondante vue en UVC
Dicentrarchus labrax (bar)	OUI (2)	Aucune autre.	OUI	Triplex	HEX-0109 (atl) FAM-0110 (Med)	OUI	Espèce abondante mais pas ou très peu vue en UVC
Raja undulata (raie brunette)	OUI (2)	<i>Raja clavata</i> R. <i>microocoellata</i> R. <i>montagui</i>	OUI	Duplex 1	FAM-0111	OUI	Statut conservation (en danger) – problématique gestion/conservation
Syngnathus acus (grand syngnathe)	OUI (2)	<i>S. rostellatus</i> , <i>S. typhile</i>	OUI	Duplex 2	HEX-0112	OUI	Espèce d'habitats adjacents (herbiers) rarement vue en UVC rocheux
Trisopterus luscus (tacaud)	OUI (1)	<i>T. minutus</i>	OUI	Duplex 2	FAM-0113	OUI	Espèce abondante vue en UVC
Zeugopterus punctatus (targeur)	OUI (1)	<i>Lepidorhombus whiffiagonis</i> <i>Scophtalmus maximus</i> <i>S. rhombus</i>	OUI	Triplex	HEX-0114/FAM-0115	OUI	Espèce a priori abondante mais cryptique donc rarement vue en UVC

4.4.2 Evaluation de la sensibilité de détection de la ddPCR par rapport au metabarcoding

La ddPCR a détecté chaque espèce cible dans plus d'échantillons que le metabarcoding (**Tableau 10-A**). Les échantillons où l'espèce cible a été détectée seulement avec la ddPCR ont en moyenne un nombre de copies d'ADN plus faible que dans les échantillons où l'espèce cible a été détectée simultanément par les 2 approches, à l'exception du tacaud (*T. luscus*) (**Figure 26**).

Dans le cas du tacaud qui est une espèce fréquente de la baie de Morlaix, nous émettons l'hypothèse que le metabarcoding sous-estime sa détection et pourrait présenter une vision biaisée de son abondance. Cela expliquerait i) pourquoi cette espèce est détectée très souvent avec la ddPCR (89% des échantillons, **Tableau 10-A**) et très peu avec le metabarcoding (26%) et ii) pourquoi la ddPCR ne semble pas plus sensible à détecter des signaux faibles d'ADN comme cela est observé pour les autres cibles. Plusieurs explications pourraient causer cette sous-estimation dont des biais de PCR, une limitation dans la résolution taxinomique ou des étapes bioinformatiques trop conservatrices.

Concernant la dernière explication, cela ne semble pas être le cas pour les étapes bioinformatiques effectuées après les assignations des ASVs. La **figure 2 en annexe** montre qu'avec le jeu de données metabarcoding n'ayant pas subi les dernières étapes de nettoyage, bien que le nombre d'échantillons où *T.luscus* a été détecté soit plus important (17 en plus), aucune séquence de l'espèce n'a été détectée dans les échantillons où le nombre de copies en ddPCR était le plus fort. Cela indique donc que les étapes de nettoyage finales ne sont pas responsables pour la non concordance entre ddPCR et metabarcoding.

Pour 4 des 6 cibles, il y a 16 échantillons (5% au total) où la ddPCR n'a pas détecté les cibles alors qu'elles ont pourtant été vues avec le metabarcoding (**Tableau 10-B**). Sur ces 16 échantillons, 3 (2 pour *Raja undulata* et 1 pour *Zeugopterus punctatus*) sont en-dessous du seuil de confiance que nous avons fixé pour éviter les faux-positifs en ddPCR (c'est-à-dire 2 copies) mais pour lesquels un signal ADN est observé en ddPCR (nombre de copies : 1.19, 1.42, 1.63) ; pour les 13 autres, 0 copies ont été obtenues en ddPCR. Sur ces 16 cas, la cible a été détecté dans maximum 1 (81 % des cas) ou 2 réplicats de PCR sur les 12 effectués par échantillon biologique (rappel du protocole metabarcoding **3.3.2**). Cela indique une faible détection avec le metabarcoding pour ces cas précis même après 12 prises d'ADN en comparaison à la ddPCR pour laquelle une seule prise a été effectuée.

Tableau 10-A. Synthèse des détections obtenues par ddPCR et metabarcoding sur les 51 échantillons biologiques

Nombre d'échantillons (%)	<i>Dicentrarchus labrax</i>	<i>Centrolabrus exoletus</i>	<i>Zeugopterus punctatus</i>	<i>Syngnathus acus</i>	<i>Raja undulata</i>	<i>Trisopterus luscus</i>
Détection ddPCR	51 (100%)	25 (49%)	13 (26%)	24 (47%)	19 (37%)	45 (89%)
Détection metabarcoding	35 (69%)	12 (24%)	10 (20%)	8 (16%)	7 (14%)	13 (26%)

Tableau 10-B. Détail sur la concordance entre ddPCR et metabarcoding sur les 51 échantillons biologiques

Nombre d'échantillons (%)	<i>Dicentrarchus labrax</i>	<i>Centrolabrus exoletus</i>	<i>Zeugopterus punctatus</i>	<i>Syngnathus acus</i>	<i>Raja undulata</i>	<i>Trisopterus luscus</i>
Détection simultanée avec les 2 approches	35 (69%)	10 (20%)	3 (6%)	8 (16%)	2 (4%)	11 (22%)
Détection seulement avec la ddPCR	16 (31%)	15 (29%)	10 (20%)	16 (31%)	17 (33%)	34 (67%)
Détection seulement avec le metabarcoding	0 (0%)	2 (4%)	7 (14%)	0 (0%)	5 (10%)	2 (4%)
Non détection avec les 2 approches	0 (0%)	24 (47%)	31 (61%)	27 (53%)	27 (53%)	4 (8%)

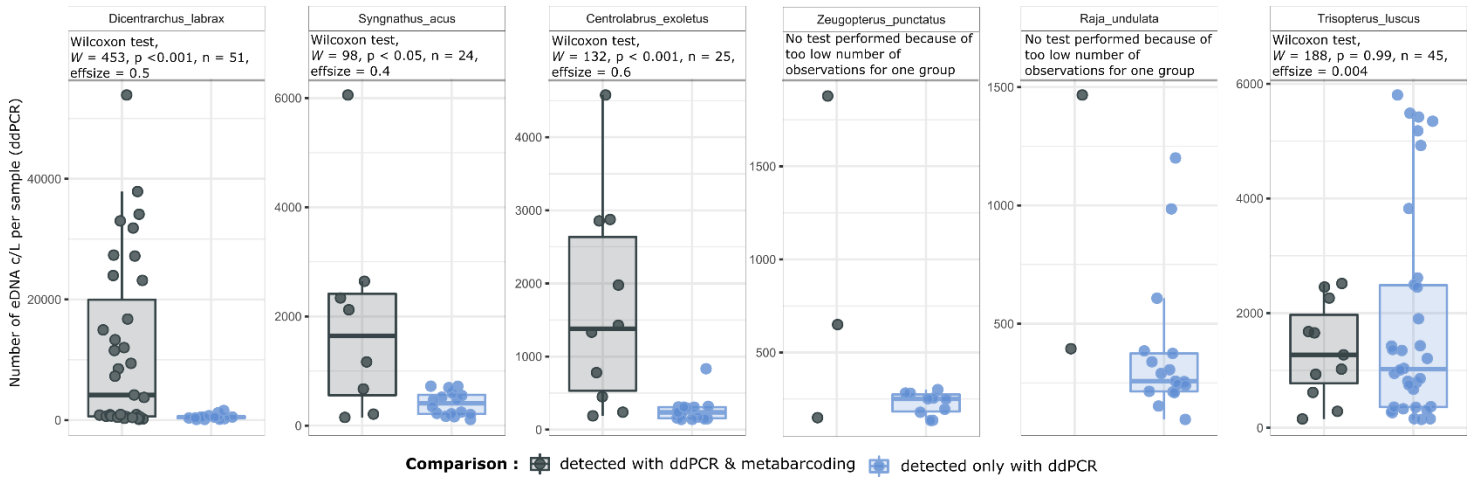


Figure 26. Comparaison du nombre de copies observé en ddPCR entre les échantillons où la cible est retrouvée simultanément avec la ddPCR et le metabarcoding (gris) et les échantillons où la cible est détectée seulement avec la ddPCR (bleu).

4.4.3 Comparaison des patrons quantitatifs

Sur l'ensemble du jeu de données, les 2 approches décrivent une même tendance dans les patrons d'abondance du signal ADN entre espèces à l'exception du tacaud (**Figure 27**). Ce dernier représente 14% du nombre de copies sur les 6 espèces en ddPCR alors qu'il ne représente que 2.5% du nombre de séquences en metabarcoding (en lien avec l'hypothèse de sous-estimation du metabarcoding pour cette espèce exposée ci-dessus). Concernant les autres espèces cibles, on observe avec les 2 approches que le bar (*D.labrax*) est l'espèce qui présente le plus grand nombre de copies d'ADN dans les échantillons analysés alors que la raie brunette (*R.undulata*) et le targeur (*Z.punctatus*) sont celles avec les nombres de copies les plus faibles (**Figure 27 B**).

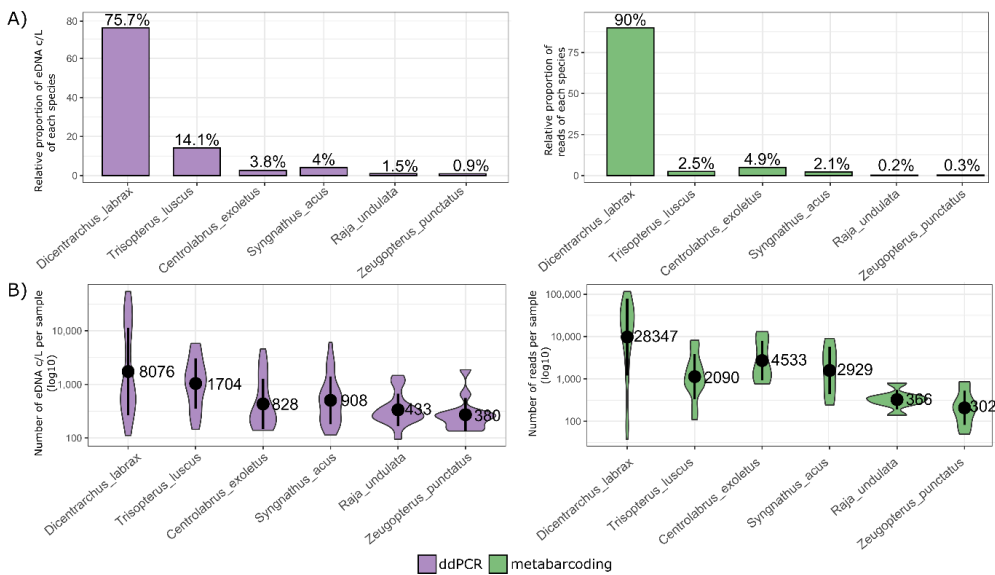


Figure 27. Résumé des patrons quantitatifs sur l'ensemble du jeu de données. (A) Proportion relative du nombre de copies (ddPCR) et de reads (metabarcoding) pour chaque espèce cible sur l'ensemble du jeu de données. (B) Violin plot représentant la distribution du nombre de copies (ddPCR) et de séquences (metabarcoding) par échantillon pour chaque espèce cible. Le point noir représente le nombre moyen.

Une relation linéaire positive et significative entre le nombre de copies obtenu en ddPCR et le nombre de séquences obtenu en metabarcoding est observée pour le bar, le grand syngnathe (*S.acus*) et le centrolabre (*C.exoletus*) (**Figure 28**). Pour ces espèces cibles, la corrélation entre les rangs des échantillons est haute pour le bar (0.7) et modérée pour le syngnathe et le centrolabre (0.5). Dans le cas des espèces qui sont détectées en plus faible quantité (**Figure 27**), il n'y a pas de concordance entre les patrons retrouvés par les deux approches (**Figure 28**).

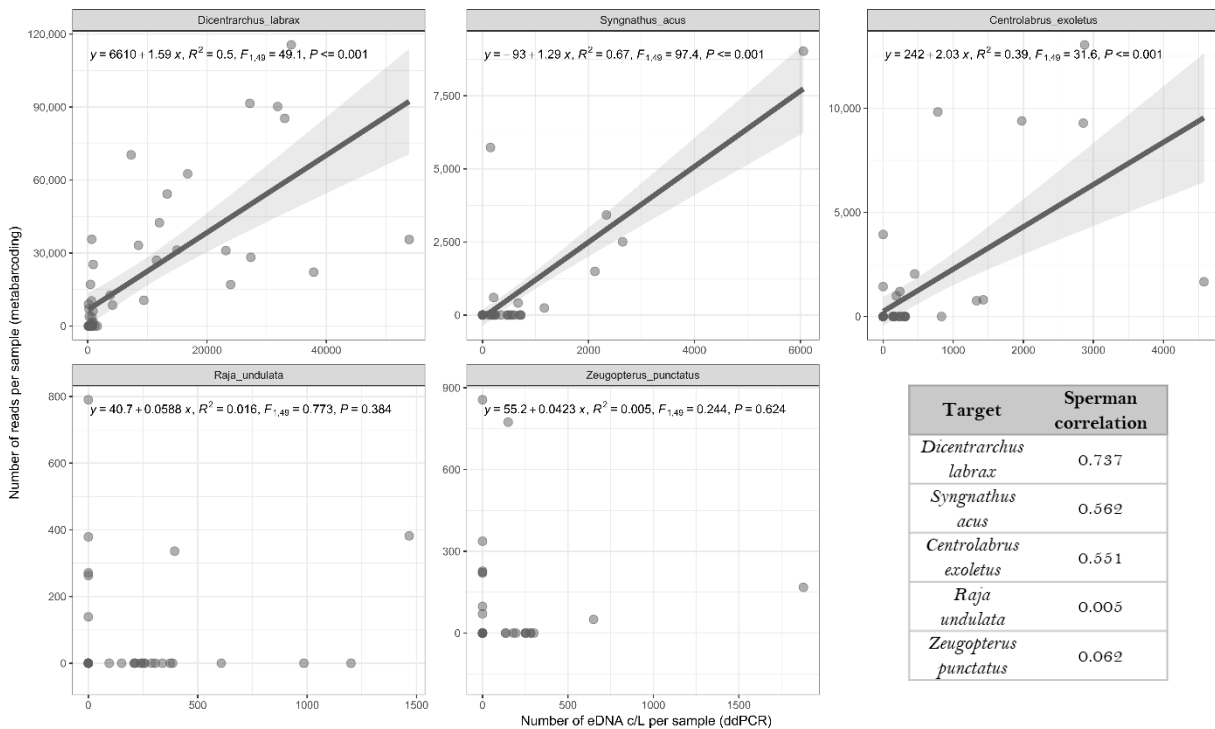


Figure 28. Relation linéaire pour chaque cible entre le nombre de ADNe c/L obtenu en ddPCR et le nombre de séquences par échantillon obtenu en metabarcoding pour les 51 échantillons. Le tableau indique les résultats de la corrélation de Spearman basée sur les rangs entre les patrons quantitatifs observés en ddPCR et metabarcoding pour les 51 échantillons. NB : Le tacaud n'est pas considéré ici puisque le metabarcoding semble sous-estimer la détection de cette espèce.

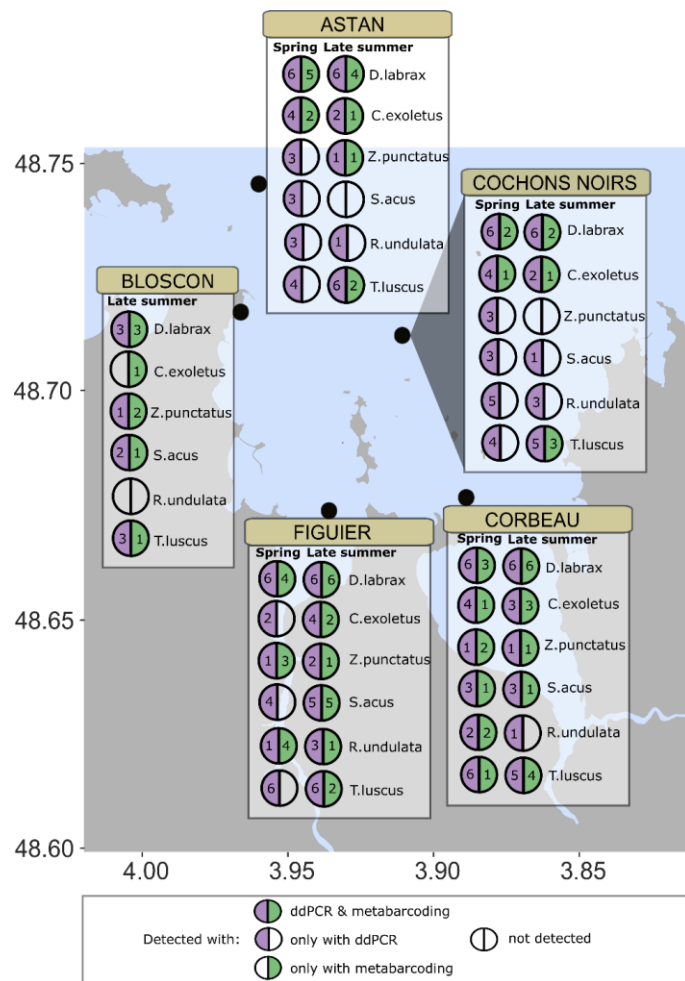
Pour le tacaud et le centrolabre, espèces fréquemment observées en UVC en baie de Morlaix, l'analyse de corrélation de Spearman ($-0.3 > \text{corrélation UVC vs ADNe} > 0.1$) à l'échelle des sites ($n=8$) ainsi que l'observation des patrons spatio-temporels n'indique *a priori* pas de concordance entre les patrons d'abondance vus en UVC et signaux ADNe estimés à partir de la ddPCR et du metabarcoding (voir détails **Figure 3** en annexe). Cette analyse reste préliminaire pour conclure sur la concordance entre les patrons quantitatifs observés par les différentes méthodes d'observation et nécessite un travail plus approfondi afin de comprendre les différences observées.

Tout comme il a été observé à l'échelle de la communauté avec les analyses metabarcoding (cf **4.3.2**), les signaux quantitatifs d'ADNe pour ces 6 espèces cibles ne sont pas différents en fonction du type de prélèvement des échantillons ADNe (testés ici avec 2L fond vs surface, voir **Figure 4** en annexe).

4.4.4 Comparaison des patrons spatiaux-temporels des espèces cibles

La ddPCR détecte plus souvent les espèces cibles au sein des sites par rapport au metabarcoding (**Figure 29**). Le nombre moyen d'espèces détectées par site à une saison donnée est de 5.5 avec la ddPCR et de 4 avec le metabarcoding et le nombre moyen d'occurrences totales des espèces présentes par site à une saison donnée est de 19.1 avec la ddPCR et de 9.6 avec le metabarcoding. Le bar est l'unique espèce détectée dans tous les sites avec les 2 approches et est présent dans 100% et en moyenne dans 70% des échantillons d'un site à une saison donnée pour la ddPCR et le metabarcoding respectivement (**Figure 29**). Avec la ddPCR on note que le tacaud est aussi relativement fréquent et est présent en moyenne dans 89 % des échantillons d'un site à une saison donnée (**Figure 29**).

Figure 29. Evaluation de la détection des espèces cibles par les approches metabarcoding et ddPCR sur ADNe au niveau de la baie de Morlaix. Le nombre d'échantillon dans lesquels l'espèce est détectée est inscrit sur la figure. Le nombre maximal d'échantillon pour le port de Bloiscon par site à une saison est de 3 (3 échantillons de surface) et 6 pour les 4 autres sites (3 échantillons de surface et 3 de fond).



Les patrons semi-quantitatifs spatiaux-temporels entre la ddPCR et le metabarcoding pour les espèces (bar, grand syngnathe et centrolabre) où la relation entre nombre de copies en ddPCR et nombre de séquences en metabarcoding était linéaire et significative (cf. **Figure 28**) concordent (**Figure 30**). En effet, pour ces 3 espèces, les sites à l'intérieur de la baie (Corbeau et Figuier) en Septembre sont ceux pour lesquels on observe des signaux quantitatifs d'ADNe plus forts que dans les autres sites (**Figure 30**).

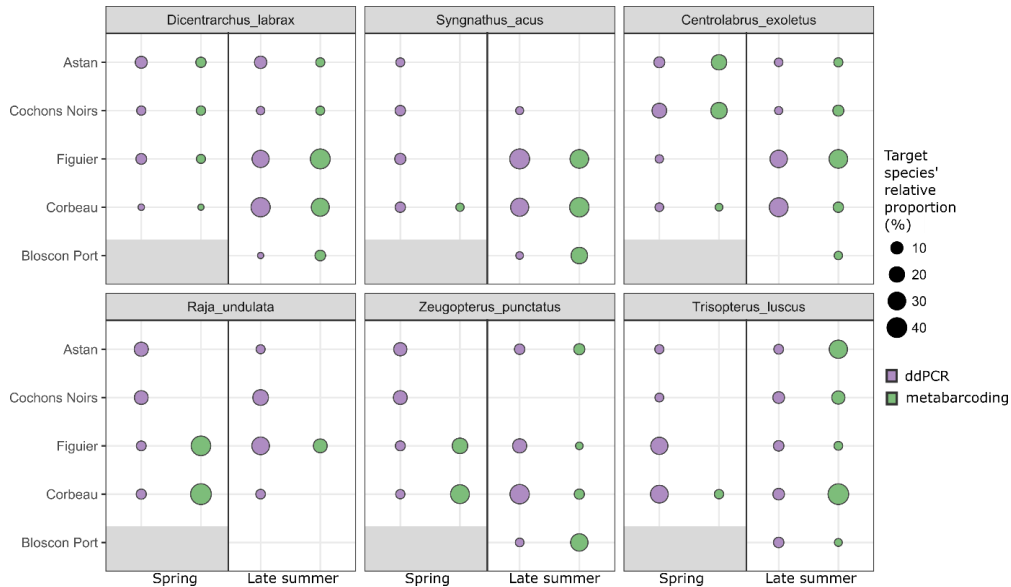


Figure 30. Proportion relative de l'espèce dans chaque site à une saison pour chaque approche (ddPCR et metabarcoding) par rapport à l'abondance totale de l'espèce. L'absence de point signifie une valeur d'abondance de 0.

4.4.5 Distinction des lignées Atlantique et Méditerranéenne du bar

Les 2 approches ADNe ont permis de distinguer les 2 lignées de bar, illustrant leurs potentiels pour détecter la diversité à des échelles infra-spécifiques. Néanmoins, la ddPCR a observé une fréquence de la lignée Méditerranée (7.7%) proche de celle documentée dans d'autres études (5%) et qui reste peu variable entre sites (proportion de la lignée Méditerranée entre 3 et 9.7%) quel que soit le nombre de copies d'ADN (Figure 31). Le metabarcoding semble surestimer la lignée Méditerranée (13.9%) par rapport à la lignée Atlantique (86.1%). De plus, les ratios des 2 lignées varient plus fortement entre sites (proportion de la lignée Méditerranée entre 0 et 100%, Figure 31). Il est possible que ces résultats inattendus soient dus à des compétitions entre ADN lors de l'amplification pour le metabarcoding, rappelant la nature qualitative plutôt que quantitative du metabarcoding.

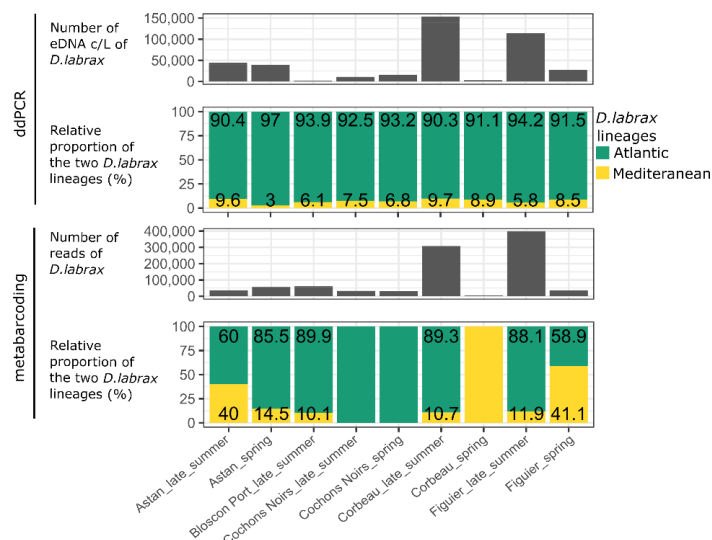


Figure 31. Détail des ratios des 2 lignées du bar (*D. labrax*) au niveau d'un site à une saison donnée observées en ddPCR et metabarcoding

4.5 Les premiers pas du « workflow »

Les premières étapes de réflexion et de création des premiers scripts ont été effectuées durant ce projet. La création du workflow complet nécessite plus de temps et pourrait être continuée à partir de 2022 dans le cadre du projet PoCoRocH afin de proposer une analyse des facettes de diversité possible à partir de jeux de données de communautés issues de source diverses (metabarcoding, comptages visuels, ...). Le but de ce workflow ainsi que des analyses bioinformatiques faites en amont (cf. 3.3.3) est de faciliter et de faire gagner du temps dans les projets à venir utilisant le metabarcoding sur ADNe. Les scripts associés seront rendus accessibles à la communauté scientifique sur github une fois terminés. Cependant un premier travail a été fait afin de réfléchir aux différents modules (cf. 3.5 et Figure 7) et commencer à créer les fonctions R associées. L'état de ce travail est détaillé dans le Tableau 11 et quelques résultats des fonctions sont présentés dans les Figure 32 et Figure 33.

Tableau 11. Résumé de l'état du workflow. Les packages R indiqués sont uniquement ceux permettant la requête d'une base de données ou la création d'indices de diversité ou de manipulation de données de diversité

Modules	Fonctions	Package R principal	Description	Statut
Module 1 : Réculte des données écologiques/ biologiques liées aux assignations taxonomiques	get.taxo.class	taxize; worrms	1. Corrige de potentielles fautes dans le nom de l'assignation 2. Récupère la classification acceptée de WORMS 3. Si taxon absent de WORMS, chercher la classification de GBIF	Créée. Liens à faire avec TaxRef en plus de GBIF
	fishbase.trait	rfishbase	1. Récupère les données de traits présentes dans fishbase pour les espèces 2. Propose un protocole pour assigner une valeur de trait aux taxons assignés au genre ou à la famille	Créée mais améliorations nécessaires pour simplifier le choix des traits pour l'utilisateur. Liens à faire avec la BDC Traits de l'INPN
	iucn.status	rredlist	Récupère le statut IUCN des espèces. Récupérer d'autres statuts via la BDC Statut de l'INPN	A créer
Module 2 : Intégration des données d'occurrence	id.redflag.obis	robis	Identifie les espèces du jeu de l'utilisateur absentes de la base OBIS ou bien de la zone géographique ciblée par l'utilisateur	Créée
	occ.species.obis plot.occ.species.obis	robis	Récupère les données d'occurrence de la base OBIS pour une espèce en particulier et représenter ces occurrences avec celles observées dans le jeu de l'utilisateur	Créées mais améliorations nécessaires pour la représentation graphique de plusieurs espèces
	occ.genus.obis plot.occ.genus.obis	robis	Récupère et « plote » les données d'occurrence des taxons uniquement identifiés au genre dans le	Créée

			jeu de l'utilisateur afin d'identifier quelles espèces pourraient potentiellement être présentes dans la zone ciblée par l'utilisateur	
Module 3 : Intégration des données phylogénétiques	get.fishtree	fishtree	Récupère l'arbre phylogénétique de fishtreeoflife	Créée
Module 4 : Récolte des métadonnées environnementales	na	na	na	Mis de côté pour l'instant dû à la complexité d'accessibilité des données de modèles
Module 5 : Intégration des différentes facettes de diversité	traits.comm.pruning tree.comm.pruning tree.traits.comm.pruning	picante	Fonctions permettant de créer le sous-ensemble de taxons utilisés pour les analyses de diversité (<i>i.e.</i> taxons de la communauté du jeu de l'utilisateur qui sont présents dans la base de trait ainsi que dans l'arbre phylogénétique)	Créées
	vd.3div.plot vd.3div.overlap	pas de package spécifique	Fonction permettant d'identifier les taxons de la communauté initiale gardés pour les analyses de diversité et de voir les taxons dont les informations sur les traits ou la phylogénie sont manquantes	Améliorations prévues ultérieurement : - analyse de co-inertie entre le jeu initial et le jeu utilisé pour les analyses de diversité - nombres de séquences gardées pour les analyses de diversité
	agg.replicates filter.samples select.level tree.phylo.comm.sel	pas de package spécifique	Diverses fonctions permettant à l'utilisateur de filtrer le jeu de données pour les analyses de diversité	Créés mais améliorations nécessaires
	hill.div.alpha	hillr; adiv	Analyse de diversité α avec les nombres de Hill	A créer
	hill.div.beta	hillr; adiv	Analyse de diversité β avec les nombres de Hill	A créer

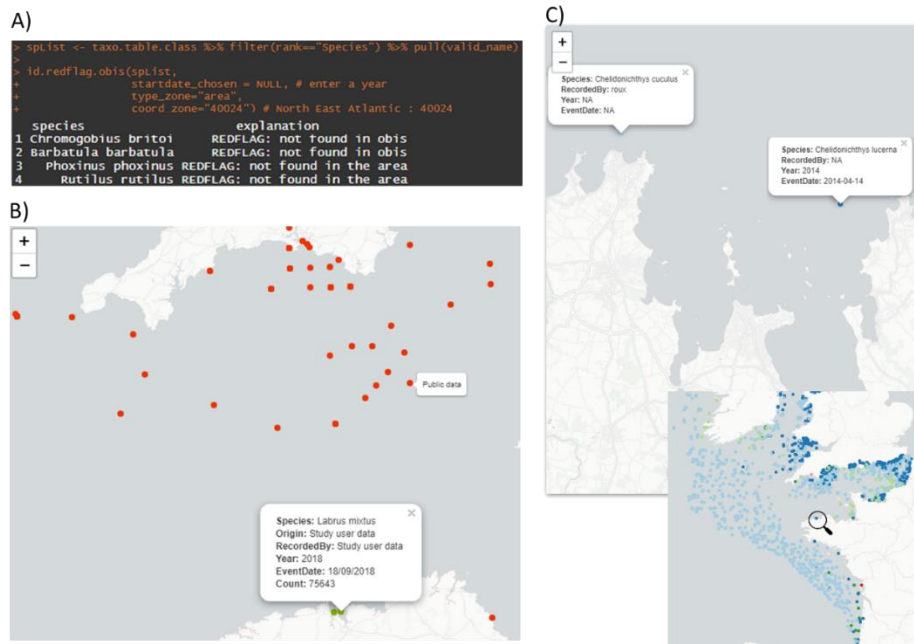


Figure 32. Exemples de sortie des fonctions créées dans le module 2. La fonction présentée en A) a permis d'identifier dans le jeu utilisateur ces 4 espèces dont B.barbatula, P.phoxinus et R.rutilus sont des espèces d'eau douce d'où leur absence dans la base de données marine OBIS. La figure B) représente un zoom sur la carte intégrant les données d'OBIS (en rouge) et les données de l'utilisateur (en vert) pour Labrus mixtus. La figure C) représente les occurrences d'OBIS des espèces du taxon assigné seulement au genre dans le jeu utilisateur pour le genre Chelidonichthys.

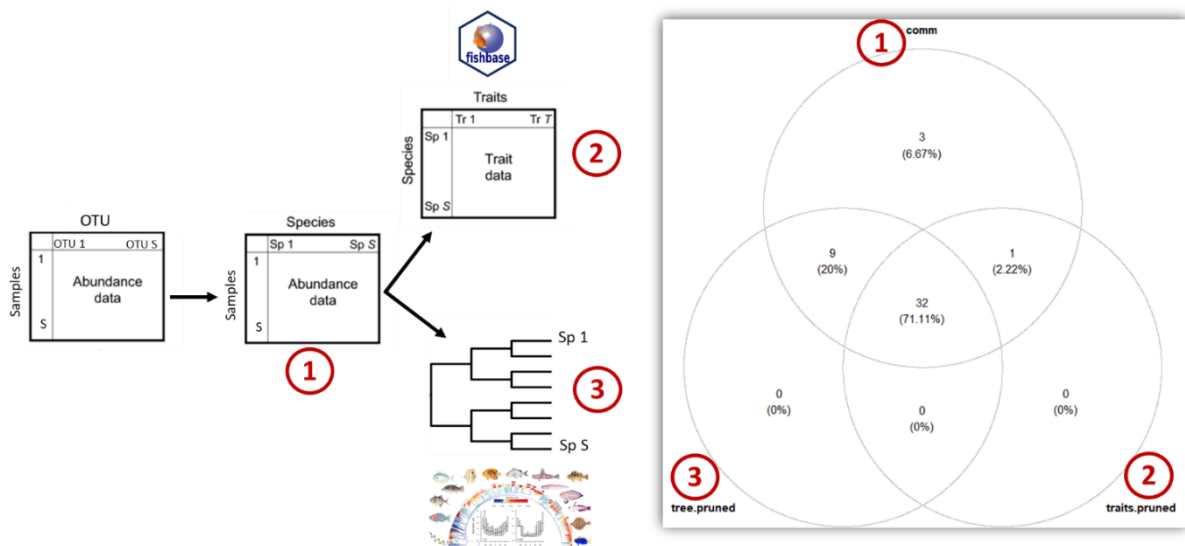


Figure 33. Représentation des étapes de préparation des jeux de diversité taxinomique, fonctionnelle et phylogénétique ainsi que du résultat de la fonction vd.3div.plot du module 5

5 Discussion

5.1 Bilans de l'étude pilote

5.1.1 Le metabarcoding sur ADNe a permis une analyse détaillée de la diversité ichtyologique de la baie de Morlaix

- L'assignation taxinomique est l'élément clé du metabarcoding pour la détection d'espèces : le choix des marqueurs et amorces, la création de la base de référence et la vérification des assignations taxinomiques en sont des étapes indispensables.

Nos résultats sont en accord avec les nombreuses études ayant utilisées le metabarcoding sur ADNe à des fins de surveillance des poissons, soulignant en particulier que cette technique détecte plus d'espèces que les méthodes conventionnelles, mais qu'elle n'est pas en mesure de détecter systématiquement toutes les espèces observées avec les méthodes conventionnelles.

Les principales limites du metabarcoding sur ADNe proviennent du manque de résolution taxinomique et/ou de l'absence de données de référence (Aglieri et al., 2020; DiBattista et al., 2017; Eble et al., 2020; Juhel et al., 2020; Nester et al., 2020). Par exemple, dans le cas des poissons côtiers de la baie de Morlaix ciblés ici, 7 des 23 espèces observées avec UVC (jeux de données 24h) n'ont pas été identifiées par aucun des 2 marqueurs.

Une analyse au cas par cas permet potentiellement d'enrichir la liste des espèces obtenue par metabarcoding. Ainsi, pour 3 espèces proches (*Labrus bergylta*, *L. viridis*, *L. merula*) les séquences de référence sont identiques à 100% ne permettant pas leur détection individuelle. Néanmoins, il est raisonnable de pouvoir assigner les séquences assignées à ces différentes espèces à *Labrus bergylta*, observée avec les UVC, du fait de l'absence des 2 autres espèces dans la région d'étude (Iglesias, 2020). Ce type de choix doit néanmoins être fait avec une grande prudence, et en complément d'autres observations ou connaissances. En effet, il est possible que de nouvelles espèces aient été introduites dans la zone d'étude. De même, bien que le metabarcoding sur ADNe soit une méthode efficace pour surveiller les changements d'aire de distribution, toute nouvelle observation doit être prise avec prudence. Dans notre étude, cela est illustré par *Chromogobius britoi*, espèce crypto-benthique associée au substrat rocheux, détectée uniquement avec le metabarcoding sur ADNe, et jamais observée plus au nord que les eaux galiciennes (Bañón and Martínez, 2019). *C. britoi* est la seule espèce connue du genre à être présente dans l'Atlantique du Nord-Est (les autres étant observées seulement en Méditerranée), ce qui pourrait rendre probable le fait qu'elle soit détectée par metabarcoding. Cependant, une seule séquence 12S est disponible dans NCBI pour l'ensemble du genre, ce qui empêche d'évaluer le pouvoir de discrimination taxinomique du 12S et donc de statuer réellement sur la fiabilité de cette assignation.

Concernant les problèmes de résolution taxinomique, dans le cas du marqueur 12S, le plus efficace des deux marqueurs utilisés, nous avons travaillé avec les amorces Teleo (Valentini et al.,

2016). Ces amorces sont très intéressantes car elles permettent d'identifier à la fois des actinoptérygiens et des élasmobranches (Collins et al., 2019). Cependant, leur résolution taxinomique peut être limitée. L'utilisation des amorces Mifish (Miya et al., 2015) ciblant le même marqueur 12S aurait potentiellement permis d'augmenter la détection d'espèces observées en UVC, notamment concernant le lieu jaune (*Pollachius pollachius*), le pagre commun (*Pagrus pagrus*) et le mullet rouget-barbet de roche (*Mullus surmuletus*) (**Tableau 4** en annexe). Le cas du lieu jaune qui est une espèce importante des milieux rocheux (retrouvé dans 33 des 34 échantillons du jeu UVC Baie) illustre à quel point le choix des amorces est crucial. C'est pourquoi nous recommandons une approche multi-marqueur ainsi qu'une évaluation des marqueurs et amorces en amont des futurs projets de surveillance en fonction des communautés visées afin d'améliorer le potentiel d'utilisation du metabarcoding.

o Le metabarcoding sur ADNe apporte une vision intégrée de la communauté des poissons côtiers.

En plus des taxons spécifiques aux habitats rocheux, le metabarcoding sur ADNe a permis la détection de taxons pélagiques et benthopélagiques vivant dans la colonne d'eau, ainsi que d'autres taxons démersaux des habitats avoisinant la roche (sédiment meuble, herbier, ...). Ces taxons ne sont pas (ou rarement) signalés par UVC. Concernant les taxons inféodés aux habitats voisins, les UVC ne les détectent pas puisqu'ils sont réalisés strictement en milieux rocheux. Concernant les poissons benthopélagiques - très mobiles dans la colonne d'eau et dans les différents habitats démersaux (et souvent timides/fuyants envers les plongeurs), les UVC sous-estiment les abondances, certainement pour plusieurs raisons combinées : les UVC ne sont pas réalisés au moment où les espèces passent dans l'habitat rocheux, et/ou les UVC sont réalisés sur des surfaces trop petites au regard de la « patchiness » de la distribution spatiale de ces espèces souvent agrégatives, et/ou les UVC ne détectent pas des individus pourtant présent car ils fuient la présence des plongeurs. Le bar (*Dicentrarchus labrax*) en est un exemple remarquable d'espèces benthopélagique sous-échantillonnée en UVC. Cette espèce a été en effet fréquemment détectée dans les données de metabarcoding (51 échantillons sur un total de 64) alors qu'elle très rarement observée par UVC (3 transects sur 271 du jeu de données Baie). Cet exemple illustre ainsi la valeur du metabarcoding sur ADNe pour augmenter la détectabilité des espèces mobiles et timides, vivant dans ou à proximité d'habitats rocheux, mais souvent sous-estimées avec les UVC (Thanopoulou et al., 2018). Nos résultats sont donc en accord avec ceux des études précédentes, comparant le metabarcoding sur ADN aux méthodes conventionnelles de surveillance des poissons côtiers (données d'observation et de capture de la pêche), qui montraient que le metabarcoding sur ADN détecte des diversités taxinomiques, phylogénétiques et fonctionnelles plus grandes et que les assemblages de poissons observés par cette approche étaient les moins redondants (Aglieri et al., 2020; Polanco Fernández et al., 2021). Ces études ont également mis en évidence que les taxons mobiles, en banc et pélagiques étaient plus facilement détectés avec le metabarcoding sur ADNe, comme également observé dans notre étude avec le prêtre (*Atherina*

presbyter), espèce côtière pélagique qui a le plus largement contribué à la diversité β calculée dans notre étude.

Concernant les poissons crypto-benthiques tels que les Gobiidae et les Blenniidae vivants sur/dans le substrat rocheux, le metabarcoding sur ADNe a détecté des espèces non observées avec les UVC telles que *Lipophrys pholis*, *Coryphoblennius galerita* et *Lebetus guilleti*. L'utilité du metabarcoding sur ADNe pour dévoiler les poissons crypto-benthiques a souvent été notée (Aglieri et al., 2020; Bessey et al., 2020; Juhel et al., 2020; Polanco Fernández et al., 2021). Cette technique pourrait donc contribuer à mieux évaluer la répartition et l'occurrence de ces poissons souvent négligés alors qu'ils sont diversifiés et fréquents et ont une forte influence sur le fonctionnement de l'écosystème des habitats rocheux et environnants (Goatley and Brandl, 2017; Thiriet et al., 2016b). Les caractéristiques écologiques et physiques de ces poissons crypto-benthiques pourraient expliquer l'absence d'observation avec UVC pour le petit (<à 2 cm) et bien camouflé *Lebetus guilleti* ou pour les poissons vivants dans les crevasses avec des comportements de « homing » tels que *Coryphoblennius galerita* et *Lipophrys pholis*. Cependant, le transport d'ADN des habitats adjacents semble être une explication alternative pertinente concernant leur détection avec le metabarcoding sur ADNe. En effet, *Lebetus guilleti* vit principalement sur les bancs de maërl, tandis que *Coryphoblennius galerita* et *Lipophrys pholis* vivent principalement en milieux intertidaux rocheux ; or ces habitats n'ont pas été échantillonnés en UVC. Cela suggère à nouveau le caractère intégrateur à l'échelle du paysage du metabarcoding sur ADNe. Cette idée est d'ailleurs renforcée par le résultat suivant : *L. pholis* a été détecté dans 23 échantillons, dont 20 provenant de Corbeau et Figuier, les sites à l'intérieur de la baie avec la plus forte proportion d'habitats rocheux médiolittoraux adjacents. A noter cependant que *C. galerita*, trouvée que dans 4 échantillons, ne présente pas ce patron paysagé. Il se pourrait que le comportement (la mobilité) propre à chaque espèce puisse expliquer ces différences, en affectant la dissémination de leur ADN dans la colonne d'eau. *L. pholis* est en effet connu pour effectuer à marée haute de grands déplacements en dehors de son refuge (utilisé à marée basse), tandis que *C. galerita* reste à proximité de son refuge à marée basse (Faria and Almada, 2006).

Concernant les espèces démersales spécialistes du substrat rocheux et cibles des UVC, les 2 méthodes d'observations UVC et ADNe ont révélé une richesse similaire au niveau de la baie. Dans le cas des Labridae, une famille fréquemment observée dans les assemblages de poissons rocheux de l'Atlantique Nord-Est (Jackson et al., 2006; Pita et al., 2018), le metabarcoding sur ADNe a détecté cinq (*Labrus mixtus*, *Ctenolabrus rupestris*, *Centrolabrus exoletus*, *Symphodus melops*, *S. bailloni*) des 6 espèces observées avec UVC. La sixième espèce, *Labrus bergylta*, n'a pas été identifiée du fait d'absence de séquence de référence diagnostique, mais est certainement l'espèce assignée à *Labrus* sp. (voir ci-dessus). De plus, les cinq espèces ont été observées dans la même gamme de fréquence d'occurrence, à l'exception de *Symphodus bailloni*. Cette dernière espèce n'a en effet été observée qu'une seule fois avec les UVC (jeux de données Baie) alors qu'elle était fréquemment observée avec le metabarcoding sur

ADNe (19 échantillons). Plusieurs études, dont deux en Manche (Dunn and Brown, 2003; Göktürk et al., 2012; Maran and Chanet, 2011), ont mis en garde contre la sous-détection potentielle de cette espèce par les UVC en raison d'un manque de connaissances sur son habitat et d'une possible confusion avec *S. melops*. Ils ont conclu que *S. bailloni* était l'espèce du genre *Symphodus* la plus abondante dans leur zone d'étude et commune dans les habitats estuariens, y compris sur plusieurs types de fonds (rocheux, boueux, sableux et herbiers), présents en baie de Morlaix. Cet exemple particulier met en évidence le potentiel du metabarcoding sur ADNe pour enrichir les connaissances sur la distribution des espèces, supposées peu fréquentes selon les méthodes d'observations conventionnelles, alors qu'elles sont potentiellement communes (Stoeckle et al., 2020b).

- L'ADN environnemental a une forte résolution temporelle et a une résolution spatiale qui a permis de différencier les communautés des sites le long d'un gradient côte-large.

Notre étude s'intègre dans une littérature grandissante illustrant que même dans des environnements côtiers marins très dynamiques, le metabarcoding sur ADNe reste sensible pour capturer la diversité locale, et peut donc être utilisé pour caractériser les assemblages de poissons à une résolution spatiale fine (Collins et al., 2018; Jeunen et al., 2019; Oka et al., 2021; Port et al., 2016; Valdivia-Carrillo et al., 2021). L'ADNe a en effet permis de distinguer la diversité ichtyologique de sites distants que de quelques kilomètres. Le transport d'ADN sur de petites distances depuis les habitats voisins (par ex., les fonds sableux démersaux ou les zones intertidales avoisinant le site rocheux échantillonné) ne peut cependant pas être exclu (Valdivia-Carrillo et al., 2021).

Des variations saisonnières ont également été observées pour certaines espèces, indiquant la fiabilité du metabarcoding sur ADNe pour décrire les changements temporels dans la distribution des espèces. Par exemple, le prêtre (*Atherina presbyter*) était principalement associé à des échantillons prélevés en juin, ce qui peut s'expliquer par son activité de frai qui a lieu cette saison (Henderson et al., 1984; Moreno et al., 2005). De plus, le mulot porc (*Chelon ramada*), espèce amphihaline, n'a été trouvée qu'en septembre, or cette espèce est connue pour se reproduire dans les zones côtières à l'automne avant de retourner dans la rivière le reste de l'année (comm. pers.). Une corrélation positive entre la concentration d'ADNe et le nombre de poissons pendant les périodes de frai a été récemment mise en évidence avec les approches de surveillance active comme la ddPCR (Thalinger et al., 2019) et la qPCR (Tsuji and Shibata, 2021). Notre étude nous a aussi permis d'identifier des zones potentielles de nourriceries en baie de Morlaix pour le bar (*D. labrax*). En effet, le metabarcoding et la ddPCR ont conjointement observé que la quantité d'ADN estimée était plus forte dans les sites à l'intérieur de la baie en Septembre, ce qui est cohérent avec la littérature décrivant l'arrivée des juvéniles dans les estuaires et zones côtières proches des estuaires de l'Atlantique Nord-Est entre la fin du printemps et l'automne (par ex. les premiers juvéniles capturés en baie de Douarnenez dans le cadre du suivi conduit par le Parc Naturel Marin d'Iroise le sont fin juillet-début août, <https://archimer.ifremer.fr/doc/00379/48987/49408.pdf>). Ces résultats encourageants suggèrent

que l'ADNe pourrait être approprié pour détecter les mouvements saisonniers et les événements de frai d'espèces côtières et estuariennes (Sigsgaard et al., 2017; Stoeckle et al., 2017).

- La stratégie d'échantillonnage de l'eau, utilisée pour obtenir l'ADN environnemental, n'a pas fortement influencé la détection des communautés des poissons côtiers dans la baie de Morlaix, une baie caractérisée par une absence de stratification verticale de la colonne d'eau.

La distance au fond marin (surface *vs* fond) ainsi que la couverture spatiale et le volume plus élevé associé (2L stationnaire *vs* 30L en mouvement autour du site rocheux) ont eu un impact très faible sur la richesse taxonomique, la détectabilité des OTU, l'accumulation de la diversité des OTU et les patrons de diversité β globaux (dissimilarité des échantillons et contribution des OTU à la diversité β). La seule exception concernait les taxons démersaux non associés aux habitats rocheux, pour lesquels les échantillons d'eau de 30 L ont détecté plus de richesse dans les sites à l'intérieur de la baie, qui ont la plus forte proportion de ce type d'habitat à proximité du site d'étude rocheux (**Figure 4**). Bien qu'il n'y ait pas de lignes directrices ou de recommandations communes concernant les volumes d'eau à filtrer, 1 à 2 L est souvent choisi comme compromis entre la quantité et la faisabilité de la filtration (Shu et al., 2020). Concernant la distance au fond marin, d'autres études comparant des échantillons d'eau de surface et d'eau de fond dans les zones côtières, avaient montré des différences en termes de diversité des poissons. Ces études suggéraient que le transport vertical des ADN était limité en raison des thermoclines ou des haloclines empêchant le mélange vertical des couches d'eau (Andruszkiewicz et al., 2017; Jeunen et al., 2020; Sigsgaard et al., 2020; Yamamoto et al., 2017, 2016). Stoeckle et al. (2020a) n'ont cependant pas observé de telles différences pour les taxons habitant les fonds sableux démersaux. Notre site d'étude, la baie de Morlaix, est caractérisé par de fortes marées sans thermocline saisonnière se traduisant par une colonne d'eau en permanence bien mélangée (Wafar et al., 1983). Cette caractéristique de l'ouest de la Manche pourrait expliquer que nous n'ayons pas trouvé de différences entre les prélèvements réalisés en surface ou près du fond. Nos résultats obtenus en eaux côtières non stratifiées sur des fonds inférieurs à 30 m de profondeur ne peuvent donc pas être généralisés à des eaux plus profondes et/ou stratifiées. C'est un point important à considérer lors de l'élaboration de la stratégie nationale d'échantillonnage.

5.1.2 La ddPCR et le metabarcoding sont des outils complémentaires afin de surveiller la faune ichtyologique

- La surveillance active par ddPCR est plus sensible que la surveillance passive par metabarcoding pour détecter des espèces cibles et permet une quantification de l'ADNe présent dans l'environnement.

Notre étude a permis de développer un outil de surveillance active pour 6 espèces de poissons côtiers ainsi que pour les lignées atlantique et méditerranéenne du bar. Les tests en laboratoire effectués par le prestataire sur les espèces proches ont montré la spécificité des sondes définies pour chaque espèce cible, validant donc leur utilisation comme outil de détection spécifique par rapport à

ces espèces proches. Le bilan de notre évaluation de la ddPCR par rapport au metabarcoding, pour chaque espèce cible, est présenté dans le **Tableau 12**. La conclusion majeure est que la ddPCR a permis d'affiner l'observation de chacune des cibles dans la baie de Morlaix par rapport au metabarcoding en détectant des signaux faibles d'ADNe non observés en metabarcoding (à l'exception du tacaud, voir ci-dessous), et ce indépendamment de la fréquence de détection de la cible.

Tableau 12. Bilan de la ddPCR par rapport au metabarcoding. La capacité de détection est indiquée en pourcentage d'échantillons dans lesquels l'espèce a été détectée. La concordance dans la détection et dans la concordance des patrons spatio-temporels est une appréciation qualitative basée sur les corrélations d'observations par site et période.

Cible	Capacité de détection de la ddPCR > à celle du metabarcoding	Concordance dans la détection	Concordance dans l'estimation de l'abondance du signal ADNe	Concordance des patrons spatio-temporels semi-quantitatifs	Bilan
Bar (<i>Dicentrarchus labrax</i>)	OUI (100 vs 69 %)	OUI	OUI (Abondante dans les 2 approches)	OUI	<ul style="list-style-type: none"> o Espèces communes dans la baie de Morlaix et détectées par analyse de l'ADNe : o Le signal ADNe est fort dans le milieu o Les patrons observés par les 2 approches concordent o La ddPCR est plus sensible que le metabarcoding
Grand syngnathe (<i>Syngnathus acus</i>)	OUI (47 vs 16 %)	OUI	OUI - (Intermédiaire dans les 2 approches)	OUI	
Centrolabre (<i>Centrolabrus exoletus</i>)	OUI (49 vs 24 %)	OUI	OUI (Intermédiaire dans les 2 approches)	OUI	
Tacaud (<i>Trisopterus luscus</i>)	OUI (89 vs 26 %)	NON	NON – (Abondante en ddPCR, faible en metabarcoding)	NON	<ul style="list-style-type: none"> o Espèce commune dans la baie de Morlaix, détectée par ddPCR mais sous-estimée avec le metabarcoding. La ddPCR est plus sensible que le metabarcoding
Raie brunette (<i>Raja undulata</i>)	OUI (37 vs 14 %)	Difficile à estimer	OUI (Faible dans les 2 approches)	NON	<ul style="list-style-type: none"> o Ces espèces ont un signal ADNe faible dans le milieu. o Les patrons observés par les 2 approches ne concordent pas. o La ddPCR est plus sensible que le metabarcoding
Targeur (<i>Zeugopterus punctatus</i>)	OUI (26 vs 20 %)	Difficile à estimer	OUI (Faible dans les 2 approches)	NON	

La ddPCR, permet une quantification des molécules d'ADN présentes dans les échantillons. Elle a permis de mettre en évidence la pertinence du metabarcoding pour estimer des niveaux relatifs d'abondance (*i.e.* données semi-quantitative) pour les espèces fréquemment détectées par les 2 approches (bar, centrolabre et grand syngnathe). Néanmoins, pour le tacaud, espèce fréquemment détectée en ddPCR, le metabarcoding a systématiquement sous-estimé sa présence, suggérant l'existence de biais lors des analyses par metabarcoding pour cette espèce. Concernant les espèces faiblement observées par les 2 approches (targeur et raie brunette), il n'y a pas eu de concordance entre les niveaux d'abondance. Pour ces espèces, il semblerait que le signal ADNe détecté soit trop faible

pour permettre une détection « fiable » avec le nombre de points d'analyse effectué (un seul par échantillon pour la ddPCR) ; l'augmentation du nombre de prises (=réplicats) d'ADNe à analyser par ddPCR est donc ici recommandé. Globalement, la ddPCR apparaît comme une approche très pertinente lorsqu'une évaluation détaillée est nécessaire pour des espèces cibles préalablement identifiées (par ex. suivis d'espèces non-indigènes, d'espèces à occurrences rares telles que les amphihalins et les élasmobranches à status de conservation, ...). A contrario, le metabarcoding permet une surveillance à l'échelle de la communauté et peut offrir une estimation d'abondance semi-quantitative pour les espèces communes, mais il convient d'être vigilant car le metabarcoding est sujet à des biais pour certaines espèces.

5.1.3 Le travail de démocratisation de l'approche « ADNe » reste nécessaire

- Développer des workflows accessibles pour démocratiser l'approche « ADNe » et amplifier la valorisation des données qui en sont issues

Nous avons pu observer durant cette étude que les analyses bioinformatiques, biostatistiques et de visualisation de données liées à l'ADNe sont en train de se démocratiser très rapidement. Par exemple, i) FROGS (<http://frogs.toulouse.inra.fr/>) disponible sur l'instance Galaxy, et dont la section Galaxy-E est développée par le Pôle National de Données de Biodiversité pour faciliter l'exploitation des données de biodiversité, ou ii) l'application shiny eDNA (<https://seak.shinyapps.io/eDNA/>) permettant de modéliser les faux positifs et faux négatifs obtenus en qPCR et d'estimer de manière plus robuste la présence de l'espèce cible permettent à des utilisateurs sans compétence en bioinformatique d'analyser leur données. L'IFREMER a aussi développé le pipeline SAMBA (<https://github.com/ifremer-bioinformatics/samba>), un workflow d'analyse metabarcoding qui offre très bonne reproductibilité dans les analyses bien que nécessitant des compétences en bioinformatique. Bien que les workflows clés en main soient nécessaires pour ainsi faciliter l'analyse des données et donc populariser l'utilisation du metabarcoding, il est important de « mettre en garde » sur les points suivants :

- (1) en fonction du workflow, des compétences en bioinformatique pourront être nécessaires,
- (2) la compréhension des paramètres clés de chaque étape bioinformatique reste indispensable pour cerner leurs impacts sur le jeu de données final,
- (3) en fonction du volume de données (*i.e.* séquences brutes), le workflow pourrait rencontrer des limites de fonctionnement et/ou l'utilisation d'un cluster de calcul serait nécessaire,
- (4) en fonction des stratégies de préparation de bibliothèques et de séquençage, les premières étapes de bioinformatique enlevant les amorces et adaptateurs de séquençage ne seront pas forcément adaptables à un workflow
- (5) il est important de vérifier la pérennité du workflow car tous ne sont pas maintenus et mis à jour,
- (6) il est important de s'assurer de la reproductibilité des analyses effectuées par le workflow.

Finalement, un domaine encore peu exploré mais cependant indispensable afin d'exploiter au mieux les données de communautés basées sur n'importe quelle méthode d'observation concerne l'intégration des données de diversité taxinomique, fonctionnelle et phylogénétique comme cela a été proposé par Aglieri et al (2020). Cette approche a été le sujet central du workflow commencé dans ce projet. Il semble donc pertinent de poursuivre cette action dans les prochains projets de R&D.

5.2 Recommandations pour la poursuite du développement de la surveillance ADN_e

5.2.1 Poursuite du développement de la surveillance passive par metabarcoding : verrous et points clés

Pour une future surveillance de routine des poissons côtiers des sous-régions marines :

- Adapter et choisir les marqueurs et amorces optimales
- Enrichir les bases de références

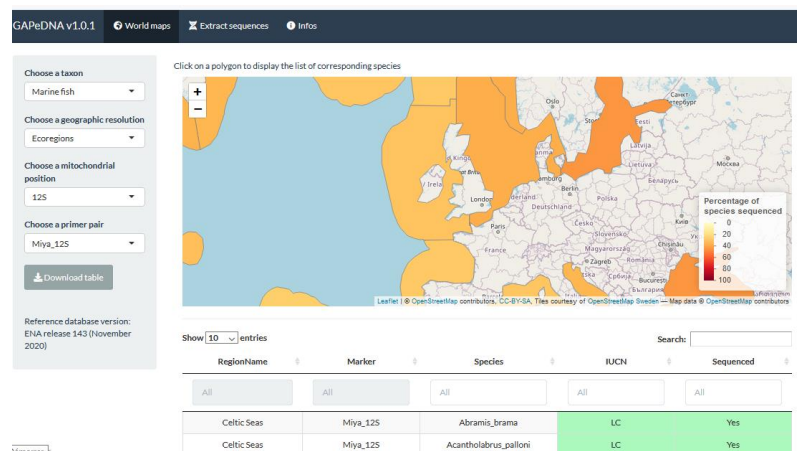
Notre étude pilote confirme les enjeux cruciaux liés au choix des marqueurs et amorces et à la complétude des bases de séquence de référence pour fournir la liste d'espèce la plus fiable et exhaustive possible. La communauté scientifique de manière générale (Collins et al., 2019; Marques et al., 2021; Wang et al., 2021) s'accorde sur le fait que le marqueur 12S semble être le marqueur de choix pour la détection des communautés de poisson en comparaison avec le COI, 16S ou cyt b. Notre étude le confirme.

Le marqueur COI, aussi bien avec des amorces ciblant les metazoaires que les poissons, manque de spécificité et a tendance à amplifier les ADN d'organismes procaryotes et eucaryotes non ciblés présents dans les échantillons (par ex. phytoplancton)(Collins et al., 2019). Pour le cas des marqueurs basés sur la région 16S, notre étude a permis d'observer que la majorité des séquences de référence des espèces était 100% identique avec des séquences d'autres espèces, à quelques exceptions près, favorisant donc son utilisation en complément d'autres marqueurs. Concernant 12S, ce marqueur est encore faiblement représenté dans les bases de références génétiques, en comparaison de COI ou 16S, ce qui limite son potentiel pour de la surveillance de routine. Il est donc nécessaire de mettre en place un plan d'action basé sur un réseau d'acteurs œuvrant pour l'élaboration et la mise à jour d'une base de référence (ciblant en priorité le marqueur 12S) des poissons côtiers des eaux métropolitaines. Pour cela, plusieurs actions seraient envisageables :

- (1) Définir une liste (non-exhaustive et « ré-actualisable ») d'espèces présentes dans les SRM des eaux côtières
- (2) Récupérer les séquences de bases publiques (ex. NCBI, ENA), les vérifier et les mettre à jour régulièrement ainsi que mutualiser les séquences non publiques (avec accord préalable)
- (3) Obtenir les séquences de référence pour les espèces manquantes (et des populations locales) en travaillant avec les groupes de recherche ADN_e poissons français, les taxonomistes et les muséums qui ont des collections de tissus, voir même des extraits d'ADN et des séquences de référence déjà disponibles (Iglesias, 2020).

- (4) Produire une interface facile d'utilisation permettant d'avoir accès à la base de référence mais aussi d'aller chercher des informations spécifiques sur une espèce en particulier (par ex. si la séquence de référence discrimine jusqu'à l'espèce). Dans ce sens, l'application GapeDNA (
- (5) **Figure 34**) permet d'avoir une première évaluation à l'échelle des SRM du nombre d'espèces ayant une séquence de référence ainsi que de leur identité (Marques et al., 2021). Pour aller plus loin, il serait pertinent de rajouter dans ces interfaces, une manière d'évaluer le pouvoir discriminant jusqu'à l'espèce des séquences de références.
- (6) Choisir le(s) couple(s) d'amorces qui permettront de détecter de manière la plus exhaustive possible les espèces de la SRM/zone d'étude ciblée.
- (7) Evaluer l'efficacité relative des amorces choisies pour amplifier l'ADN des espèces de la communauté ciblée (ex. pour mieux comprendre la sous-estimation du tacaud) à partir d'échantillons de tissus de différentes espèces (Couton et al., 2019).

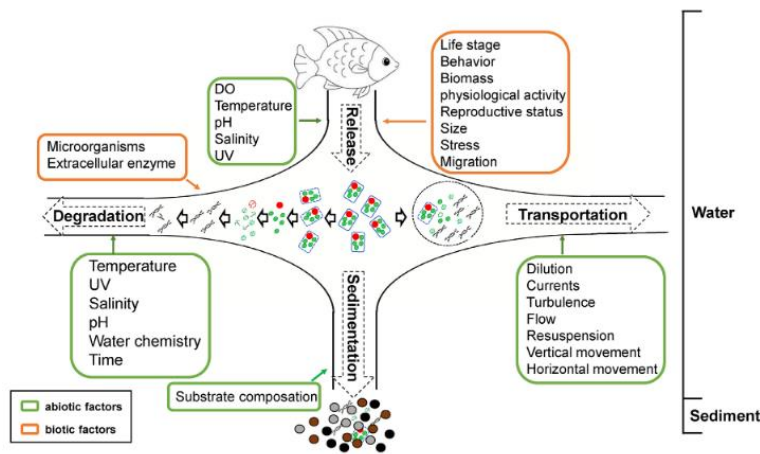
Figure 34. Impression d'écran du site <https://shiny.cefe.cnrs.fr/GAPeDNA/> développé par Marques et al. 2021



- Optimiser les protocoles d'échantillonnage en fonction des caractéristiques de l'environnement échantillonné et des taxons ciblés

De manière générale, le type et la complexité des habitats, les caractéristiques environnementales de la zone d'étude (par ex. saisonnalité, marée, stratification verticale, température) et les caractéristiques écologiques des espèces (par ex. période de frai, comportement) des communautés de poissons ciblées peuvent jouer un rôle important sur la présence et la persistance de l'ADN dans la colonne d'eau. Par conséquent, ces facteurs, entre autres (**Figure 35**), doivent être pris en compte au moment de décider du protocole d'échantillonnage de l'eau (par ex. volume, reproductibilité, couverture spatiale, profondeur). Dans notre cas d'étude ciblant des poissons provenant d'habitats subtidiaux rocheux dans une colonne d'eau bien mélangée, même un petit volume (3 fois 2 L) collecté à la surface, qui est la méthode la plus pratique (sans plongée), semble satisfaisant. Cependant, ce résultat n'est pas généralisable au reste des eaux métropolitaines car la masse d'eau de notre zone d'étude n'est pas stratifiée verticalement. Par exemple, le test fond *vs* surface devrait être reproduit dans des zones où il y a une stratification verticale de la colonne d'eau (*e.g.* Golfe de Gascogne) afin de mieux évaluer l'impact de la stratification sur le signal ADN présent dans l'eau pour la communauté des poissons côtiers mais principalement pour les espèces démersales, cachées dans le

substrat, peu mobiles, peu occurrente, de petite taille, ... De plus, pour une surveillance intégrée à l'échelle paysagère de la biodiversité des poissons (c.-à-d. dans tous les habitats et sans distinction des habitats), un plus grand volume collecté sur une vaste zone (plusieurs milles nautiques) pourrait être préféré. Dans l'idée d'intégrer l'ADNe comme approche transversale permettant de détecter les communautés de poissons côtiers, un travail de R&D dans des zones côtières aux caractéristiques contrastées représentant les milieux les plus fréquents dans les eaux métropolitaines permettrait de



définir le ou les type(s) de prélèvement à mettre en place en routine.

Figure 35. Schéma représentant les associations connues entre les facteurs abiotiques et biotiques ayant une influence sur la persistance et présence de l'ADNe dans l'environnement - Repris de Wang et al. (2021)

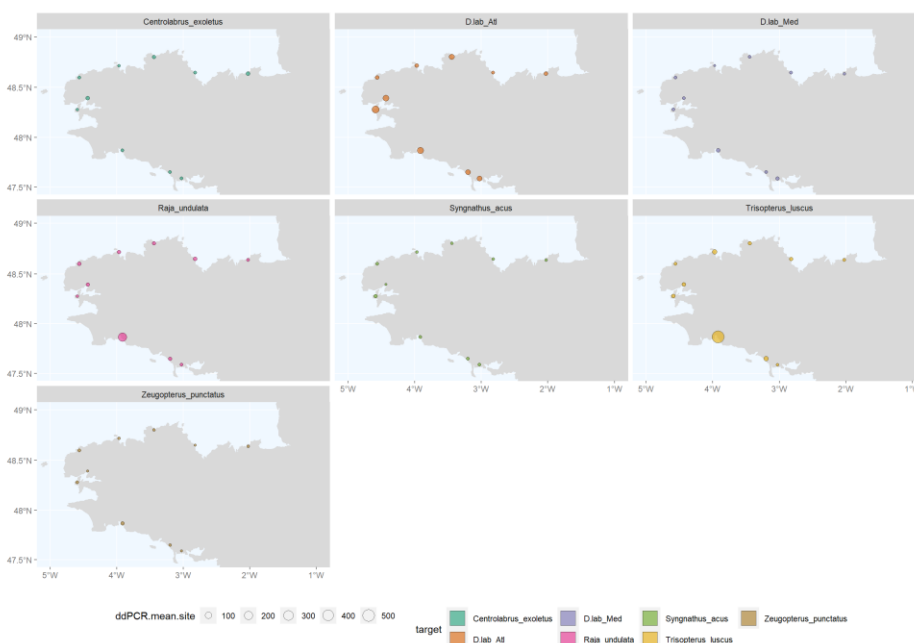
5.2.2 Poursuite du développement de la surveillance active par ddPCR : verrous et points clés

La ddPCR offre une alternative plus robuste que le metabarcoding pour la détection et l'estimation des patrons d'abondance. Bien que cette approche ne puisse pas être développée pour évaluer tous les poissons côtiers, sa future application pour des espèces clés (statut de conservation, espèce commerciale, ...) pourrait permettre d'obtenir des données sur un grand nombre d'échantillons

et donc *a fortiori* une couverture spatiale et temporelle plus large qu'avec des approches conventionnelles

(Figure 36).

Figure 36. Exemple de couverture spatiale envisageable pour les 6 espèces cibles permettant d'identifier des changements d'abondance d'ADNe. Les données présentées sont issues d'un autre projet (AQUANIS 2.0, coord. F.Viard)



concernant la faune des habitats portuaires.

Cependant, afin de proposer une surveillance fiable, une série d'étapes de validation serait à ajouter dans les futurs développements basés sur les recommandations de la littérature (Thalinger et al., 2021; Xia et al., 2021) (**Figure 37**), incluant notamment des gammes de dilution de l'ADN de chaque cible pour déterminer la sensibilité et la limite de détection ainsi qu'une meilleure couverture en terme de stratégie d'échantillonnage. Notre étude est au niveau 3 considéré comme le niveau intermédiaire où il est encore impossible de conclure si la cible est réellement absente lorsqu'elle n'est pas détectée (du fait de l'absence de réplication des amplifications par ddPCR pour les espèces avec des signaux d'ADNe faibles).

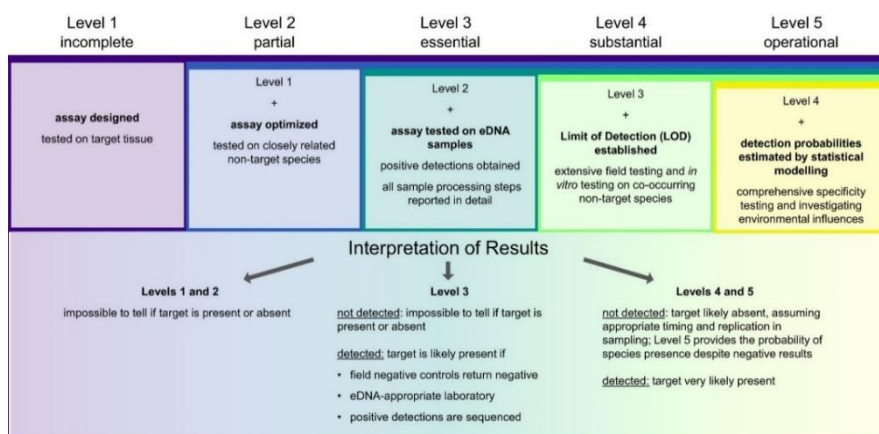


Figure 37. Synthèse des 5 niveaux de validation allant d'un niveau incomplet à opérationnel proposé par Thalinger et collaborateurs (2021)

Finalement, compte-tenu des très bons résultats obtenus pour la distinction des lignées de bars, une poursuite intéressante serait d'exploiter le potentiel de la ddPCR sur ADNe comme outil de surveillance identifiant des possibles « échappements » des bars européens issus de l'aquaculture en suivant l'évolution du ratio lignée Méditerranée/Atlantique dans les eaux environnantes des sites d'élevage puisque c'est la lignée Méditerranée qui est communément élevée.

5.2.3 Poursuite du développement de l'estimation d'abondance à partir d'ADNe

Afin de devenir un outil de surveillance (active & passive), un des défis majeurs de l'ADNe est d'évaluer son potentiel quantitatif. Une série de facteurs biotiques et abiotiques jouent un rôle dans la distribution du signal d'ADNe dans l'environnement (**Figure 35**) et nécessite des études spécifiques aux poissons côtiers afin de mieux évaluer leur rôle et de proposer des stratégies d'échantillonnage les prenant en compte. Néanmoins, la synthèse bibliographique effectuée par Rourke et al. (2021) a montré que parmi les études publiées, plus de 90% d'entre elles avaient rapportées des corrélations positives entre signaux quantitatifs d'ADNe (concentration d'ADNe pour qPCR et ddPCR et nombre relatif de séquences pour metabarcoding) et abondance ou biomasse des espèces de poissons. De plus, plusieurs études ont montré l'application concrète de l'ADNe utilisé en approche active ou passive pour fournir des données essentielles à l'évaluation des espèces et des écosystèmes. Par exemple, Li et al. (2019) ont illustré l'intérêt du metabarcoding sur ADNe pour estimer l'abondance relative des espèces de poissons d'eau douce en utilisant l'échelle DAFOR (Dominant-Abundant-Frequent-Occasional-Rare) ; ce qui

permettrait une intégration dans les approches de la Directive Cadre sur l'Eau où des classes d'abondance sont couramment utilisées. Les mêmes auteurs concluent qu'aussi bien en termes de détection d'espèce, d'estimation d'abondance relative et de caractérisation écologique des communautés, le metabarcoding sur ADNe a clairement un potentiel comme outil d'évaluation des communautés ichtyologiques des lacs pour la DCE. Riaz et al. (2020) montrent l'intérêt de la qPCR sur ADNe (approche active) afin d'évaluer la réintroduction d'espèces de poissons rares et d'évaluer si l'espèce atteindra le bon état écologique selon les critères de la DCE. Concernant l'approche quantitative par metabarcoding, Stoeckle et al. (2020a), en comparant les patrons d'abondance sur 1 an entre les captures au chalut de fond et le metabarcoding sur ADNe pour les poissons démersaux des fonds meubles, ont aussi mis en avant la concordance des patrons d'abondance saisonniers entre les 2 méthodes pour 70% des espèces. Bien qu'il reste un travail d'optimisation afin d'affiner les estimations d'abondances/biomasses à partir du signal ADNe quantifié par différentes techniques (Rourke et al., 2021; Wang et al., 2021), il semble concevable d'obtenir des données relatives et semi-quantitatives, et donc pertinent de continuer ces développements dans un but d'intégration future dans des réseaux de surveillance.

5.2.4 Extension de l'utilisation de l'ADNe pour la surveillance de la biodiversité marine à d'autres composantes biocénétiques et d'autres politiques publiques

Un webinaire est organisé fin 2021 (**Tableau 6 en annexe** pour les détails des présentations prévues) dans le but d'obtenir un retour d'expériences sur les projets ADNe en milieu aquatique afin de permettre aux gestionnaires et responsables thématiques et de surveillance DCSMM de mieux comprendre comment l'ADNe pourrait s'intégrer dans des réseaux de surveillance des autres espèces et habitats. Il y sera prévu de prendre en exemple le cas des poissons puisque plusieurs groupes de recherche français utilisent l'ADNe et aussi pour bénéficier des retours d'expériences de l'intégration de l'ADNe dans la surveillance des milieux dulçaquicoles pour la DCE. Durant ce webinaire, les réflexions et interrogations suivantes seront discutées :

- Quelles sont les différentes politiques publiques, et quelles composantes biocénétiques marines, pourraient tirer avantage d'une surveillance par ADNe ?
- Besoin de développer de nouveaux indicateurs du bon état écologique (BEE) et/ou calibrer les indicateurs construits sur des méthodes conventionnelles
- Est-il possible du point de vue méthodologique et scientifique d'élaborer un dispositif de surveillance ADNe mutualisé, permettant de suivre plusieurs composantes biocénétiques en parallèle et/ou répondre aux besoins/enjeux de plusieurs politiques publiques ? Quelles étapes peuvent-être mutualisées (*i.e.* collecte de données, laboratoire, bio-informatique, bio-statistiques & bio-indication) ?
- Besoin en infrastructures et plateformes (de l'extraction d'ADN à l'analyse bio-informatique)

5.3 Pistes de développement d'une surveillance combinant ADNe et approches conventionnelles

Basés sur les résultats de la présente étude, ainsi que sur l'analyse des dispositifs de suivis poissons existants en milieux côtiers et les travaux R&D en cours, nous proposons ici un concept de surveillance multi-source et multi-échelle. Notre proposition consiste à combiner plusieurs méthodes (ADNe, echosondage, comptages visuels, voire biophonie et/ou vidéo) complémentaires tant au niveau des catégories d'espèces surveillées, des métriques échantillonnées, et des couvertures et résolutions spatiales et temporelles (**Figure 38**).

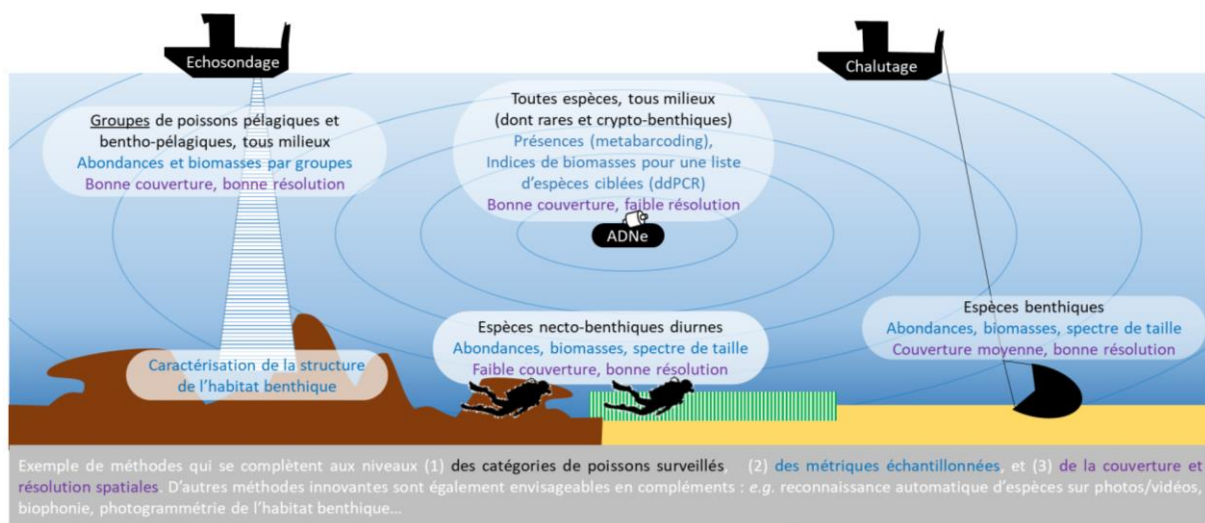


Figure 38. Représentation du concept de surveillance multi-sources et multi-échelles

Le déploiement de ces méthodes est à articuler à deux échelles spatiales (**Figure 39**) :

(1) l'observation par ADNe à l'échelle paysagère - permettant d'évaluer la composition des peuplements (metabarcoding) voire des indices d'abondance pour certaines espèces cibles (ddPCR) - vise à évaluer la tendance globale de l'état écologique à l'échelle de la façade, et dépister les zones à la dynamique particulière (pouvant faire l'objet d'études plus approfondies, e.g. un observatoire état-pression)

(2) les observatoires état-pression (e.g. dans et autour une AMP, ou un parc éolien, à graduelle distance d'une zone urbanisée) combinent ADNe et méthodes complémentaires (echosondage, comptages visuels, chalut benthique, voire biophonie et/ou vidéo). Ces dernières méthodes dites conventionnelles, sont plus « lourdes » à mettre en œuvre mais apportent des métriques supplémentaires nécessaires à l'évaluation de certaines relations état-pression (et certains descripteurs et critères DCSMM). Les observatoires état-pression visent ainsi à :

- (a) évaluer finement les relations état-pression (dont résistance et résilience sur le long terme) grâce aux complémentarités des différentes méthodes

(b) calibrer les indicateurs de l'observation par ADNe, et éclairer leur interprétation grâce à l'inter-calibration de l'ADNe avec les méthodes conventionnelles.

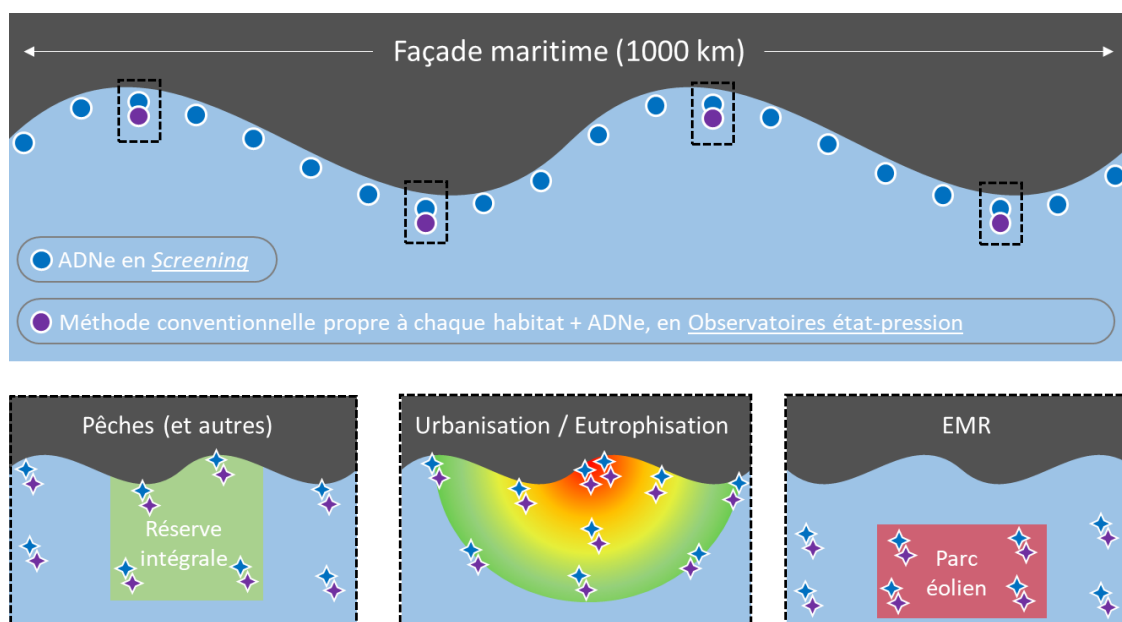


Figure 39. Représentation du concept de déploiement à large échelle de la surveillance basée sur l'utilisation conjointe de l'ADNe et des méthodes conventionnelles.

Les descripteurs et critères potentiellement renseignés par cette approche sont présentés en **Tableau 13**. Cette proposition de stratégie de surveillance des poissons côtiers, combinant approches conventionnelles propres à chaque habitat côtier, et approche inter-habitat telle que l'ADNe (voire autre telle que biophonie) devra s'insérer dans une approche écosystémique plus globale, à discuter collégialement avec les autres RT/RS DCSMM en charge du suivi de composantes écosystémiques. Nous devons chercher une stratégie globale garantissant la cohérence des suivis et évaluations des poissons entre milieux, tant sur les aspects opérationnels qu'analytiques. L'observation par ADNe, comme proposée dans le concept de surveillance multi-source et multi-échelle, a déjà été mise en place en Méditerranée depuis 2020 (voir projet PISCIS https://medtrix.fr/portfolio_page/piscis/). Il sera donc pertinent dans les étapes à suivre de R&D d'entamer un travail collaboratif afin de mettre en cohérence les différentes stratégies d'échantillonnage ADNe dans les différentes façades. De même, nous devons considérer les suivis (réalisés/envisagés) des forçages environnementaux naturels et anthropiques (*e.g.* habitats pélagiques et benthiques, pressions de pêche, etc.), afin de mieux interpréter les évolutions de la structure des peuplements de poissons, et pouvoir évaluer les relations états-pressions.

Tableau 13. Descripteurs et critères pouvant potentiellement être renseignés par les différentes méthodes d'observations

	Echelle spatiale du dispositif	Milieux côtiers surveillés	Types d'espèces	Type d'évaluation	Descripteurs DCSMM potentiellement renseignés
Metabarcoding sur ADNe	« Screening de façade »	Tous les milieux (SP1 à 3)	Toutes les espèces	Tendances des diversités taxonomique, fonctionnelle et phylogénétiques (D4C1) Tendance de la diversité de groupes d'espèces fondés sur leur sensibilité aux pressions anthropiques Evolution des aires biogéographiques de distribution, par patrons d'occurrences (D1C4 et D2C2)	D1C4 D4C1 D2C2 + apport de compléments pour D1C2, D1C5, D4C2, D2C1, D2C3
ddPCR sur ADNe	« Screening de façade »	Tous les milieux (SP1 à 3)	Quelques espèces ciblées : e.g. espèces patrimoniales, amphihalins, élasmobranches à statut de conservation, ...	Tendance des abondances spécifiques (D1C2) Evolution des aires biogéographiques de distribution, par patrons d'abondances/biomasses (D1C4 et D2C2)	D1C2 D1C4 D2C2 + apport de compléments pour D1C5
Comptages visuels en plongée	Observatoires Etat-Pression	Milieux rocheux et habitats biogéniques (SP1)	Nectobenthiques	Indice de diversité d'abondance (D1C2) et évolution de la structure démographique (D1C3) Caractérisation de l'habitat pour établir les liens état-pressions (D1C5) Tendances des abondances des peuplements et structure en taille (D4C2 et D4C3)	D1C2 D1C3 D1C5 D4C2 D4C3 + apport de compléments pour D1C4 et D4C1
Echosondage	Observatoires Etat-Pression	Milieu pélagiques (SP3) Et espèces benthopélagiques en liens avec les différents habitats démersaux (SP1 et SP2)	Groupes d'écophases, basés sur la position dans la colonne d'eau, la taille individuelle et du comportement social/spatial (individus isolés vs agrégation vs bancs)	Tendances des abondances des peuplements et structure en taille (D4C2 et D4C3) Sous réserve de développement méthodologique : Indice de diversité d'abondance (D1C2) et évolution de la structure démographique (D1C3) Caractérisation de l'habitat pour établir les liens état-pressions (D1C5)	D4C2 et D4C3 Sous réserve de développement méthodo, D1C2, D1C3 et D1C5

6 Conclusion

Cette étude pilote a fourni une première évaluation du potentiel de l'ADNe comme outil dans la surveillance des poissons côtiers des habitats rocheux. La surveillance passive avec metabarcoding sur ADNe a détecté la diversité des communautés des substrats rocheux ainsi que celle d'habitats avoisinants, tout en permettant de retrouver des signaux spatiaux et temporels à fine échelle. Cette combinaison de plus grande détectabilité et de résolution spatio-temporelle relativement fine représente une valeur ajoutée par rapport aux méthodes d'observation conventionnelles non conçues pour capturer un large spectre de biodiversité. En termes de surveillance active, la ddPCR sur ADNe s'est révélée être plus sensible que le metabarcoding pour détecter des signaux d'ADNe faibles, illustrant donc son intérêt pour évaluer des espèces où la distribution est peu connue ou en cours d'évolution, ou bien pour la surveillance d'espèces introduites. Bien que des défis d'optimisation restent à relever, l'utilisation de l'ADNe permettra d'augmenter la couverture spatiale et temporelle des suivis de surveillance favorisant donc une meilleure intégration des variations des communautés de poissons côtiers.

Nous en concluons que l'ADNe est approprié pour le suivi de la diversité ichthyologique d'habitats spécifiques, mais que la liste d'espèces détectées contient (1) des espèces dont la présence est la plus probable dans l'habitat, (2) d'autres dont la présence est la plus probable dans des habitats adjacents, (3) d'autres encore dont la présence est potentiellement dans tous les habitats (*e.g.* espèces mobiles, pélagiques et benthopélagiques). Cette interprétation probabiliste du cortège détecté au sein d'un habitat est à faire *a posteriori* par l'analyse croisée de la composition et distribution des habitats adjacents (analyse paysagère par cartographie) et des préférences d'habitat des différentes espèces d'après la littérature (*e.g.* traits de vie dans Fishbase ou TaxRef). De manière alternative ou complémentaire, l'ADNe peut également être utilisée comme une méthode intégrative à l'échelle paysagère de la diversité ichthyologique des différents habitats présents dans la zone échantillonnée. Une méthode de prélèvement d'eau le long de radiales couvrant explicitement les différents habitats composant une zone pourrait être avantageuse pour cet objectif. Les stratégies d'échantillonnage de l'ADNe à l'échelle habitat et/ou paysage seront discutées ultérieurement dans le cadre du développement de la stratégie multi-source multi-échelle, notamment au travers de l'atelier national « Développement d'un réseau de surveillance & Observation des poissons côtiers rocheux » de Décembre 2021 et des futurs projets R&D.

7 Bibliographie

- Aglieri, G., Baillie, C., Mariani, S., Cattano, C., Calò, A., Turco, G., Spatafora, D., Franco, A.D., Lorenzo, M.D., Guidetti, P., Milazzo, M., 2020. Environmental DNA effectively captures functional diversity of coastal fish communities. *Molecular Ecology* n/a. <https://doi.org/10.1111/mec.15661>
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Andruszkiewicz, E.A., Starks, H.A., Chavez, F.P., Sassoubre, L.M., Block, B.A., Boehm, A.B., 2017. Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. *PLOS ONE* 12, e0176343. <https://doi.org/10.1371/journal.pone.0176343>
- Bañón, R., Martínez, L.A., 2019. New northernmost distributional limit of *Chromogobius britoi* (Gobiidae) in the eastern Atlantic. <https://doi.org/10.26028/CYBIUM/2019-433-010>
- Bessey, C., Jarman, S.N., Berry, O., Olsen, Y.S., Bunce, M., Simpson, T., Power, M., McLaughlin, J., Edgar, G.J., Keesing, J., 2020. Maximizing fish detection with eDNA metabarcoding. *Environmental DNA* 2, 493–504. <https://doi.org/10.1002/edn3.74>
- Bhattacharya, M., Sharma, A.R., Patra, B.C., Sharma, G., Seo, E.-M., Nam, J.-S., Chakraborty, C., Lee, S.-S., 2016. DNA barcoding to fishes: current status and future directions. *Mitochondrial DNA A DNA Mapp Seq Anal* 27, 2744–2752. <https://doi.org/10.3109/19401736.2015.1046175>
- Blackman, R.C., Ling, K.K.S., Harper, L.R., Shum, P., Hänfling, B., Lawson-Handley, L., 2020. Targeted and passive environmental DNA approaches outperform established methods for detection of quagga mussels, *Dreissena rostriformis bugensis* in flowing water. *Ecology and Evolution* 10, 13248–13259. <https://doi.org/10.1002/ece3.6921>
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., Coissac, E., 2016. obitools: a unix-inspired software package for DNA metabarcoding. *Mol Ecol Resour* 16, 176–182. <https://doi.org/10.1111/1755-0998.12428>
- Caldwell, Z.R., Zgliczynski, B.J., Williams, G.J., Sandin, S.A., 2016. Reef Fish Survey Techniques: Assessing the Potential for Standardizing Methodologies. *PLOS ONE* 11, e0153066. <https://doi.org/10.1371/journal.pone.0153066>
- Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* 13, 581–583. <https://doi.org/10.1038/nmeth.3869>
- Capo, E., Spong, G., Koizumi, S., Puts, I., Olajos, F., Königsson, H., Karlsson, J., Byström, P., 2020. Droplet digital PCR applied to environmental DNA, a promising method to estimate fish population abundance from humic-rich aquatic ecosystems. *Environmental DNA* n/a. <https://doi.org/10.1002/edn3.115>
- Collins, R.A., Bakker, J., Wangensteen, O.S., Soto, A.Z., Corrigan, L., Sims, D.W., Genner, M.J., Mariani, S., 2019. Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution* 10, 1985–2001. <https://doi.org/10.1111/2041-210X.13276>
- Collins, R.A., Wangensteen, O.S., O’Gorman, E.J., Mariani, S., Sims, D.W., Genner, M.J., 2018. Persistence of environmental DNA in marine systems. *Communications Biology* 1, 1–11. <https://doi.org/10.1038/s42003-018-0192-6>
- Couton, M., Comtet, T., Le Cam, S., Corre, E., Viard, F., 2019. Metabarcoding on planktonic larval stages: an efficient approach for detecting and investigating life cycle dynamics of benthic aliens. *Management of Biological Invasions* 10, 657–689.
- DiBattista, J.D., Coker, D.J., Sinclair-Taylor, T.H., Stat, M., Berumen, M.L., Bunce, M., 2017. Assessing the utility of eDNA as a tool to survey reef-fish communities in the Red Sea. *Coral Reefs* 36, 1245–1252. <https://doi.org/10.1007/s00338-017-1618-1>

- Doi, H., Uchii, K., Takahara, T., Matsubashi, S., Yamanaka, H., Minamoto, T., 2015. Use of Droplet Digital PCR for Estimation of Fish Abundance and Biomass in Environmental DNA Surveys. *PLoS One* 10. <https://doi.org/10.1371/journal.pone.0122763>
- Dunn, M.R., Brown, M.J., 2003. The occurrence of *Symphodus bailloni* on the south coast of England. *Marine Biological Association of the United Kingdom. Journal of the Marine Biological Association of the United Kingdom* 83, 875.
- Eble, J.A., Daly-Engel, T.S., DiBattista, J.D., Koziol, A., Gaither, M.R., 2020. Chapter Two - Marine environmental DNA: Approaches, applications, and opportunities, in: Sheppard, C. (Ed.), *Advances in Marine Biology*. Academic Press, pp. 141–169. <https://doi.org/10.1016/bs.amb.2020.01.001>
- Faria, C., Almada, V., 2006. Patterns of spatial distribution and behaviour of fish on a rocky intertidal platform at high-tide. *Marine Ecology Progress Series* 316, 155–164. <https://doi.org/10.3354/meps316155>
- Ficetola, G.F., Miaud, C., Pompanon, F., Taberlet, P., 2008. Species detection using environmental DNA from water samples. *Biol. Lett.* 4, 423–425. <https://doi.org/10.1098/rsbl.2008.0118>
- Frajia-Fernandez, N., Bouquieaux, M., Rey, A., Mendibil, I., Cotano, U., Irigoien, X., Santos, M., Rodriguez-Ezpeleta, N., 2020. Marine water environmental DNA metabarcoding provides a comprehensive fish diversity assessment and reveals spatial patterns in a large oceanic area. *Ecology and Evolution* 10. <https://doi.org/10.1002/ece3.6482>
- Frøslev, T.G., Kjøller, R., Bruun, H.H., Ejrnæs, R., Brunbjerg, A.K., Pietroni, C., Hansen, A.J., 2017. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications* 8, 1188. <https://doi.org/10.1038/s41467-017-01312-x>
- Goatley, C.H.R., Brandl, S.J., 2017. Cryptobenthic reef fishes. *Current Biology* 27, R452–R454. <https://doi.org/10.1016/j.cub.2017.03.051>
- Göktürk, D., Karakulak, F.S., Ünsal, N., Kahraman, A.E., 2012. A New Record for Occurrence of *Symphodus bailloni* (Osteichthyes: Perciformes: Labridae) in the Western Black Sea Coast of Turkey. *ScientificWorldJournal* 2012. <https://doi.org/10.1100/2012/615318>
- Hansen, B.K., Bekkevold, D., Clausen, L.W., Nielsen, E.E., 2018. The sceptical optimist: challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries* 19, 751–768. <https://doi.org/10.1111/faf.12286>
- Harmelin, J.-G., 1987. Structure et variabilité de l'ichtyofaune d'une zone rocheuse protégée en Méditerranée (Parc national de Port-Cros, France). *Marine Ecology* 8, 263–284. <https://doi.org/10.1111/j.1439-0485.1987.tb00188.x>
- Hatzenbuehler, C., Kelly, J.R., Martinson, J., Okum, S., Pilgrim, E., 2017. Sensitivity and accuracy of high-throughput metabarcoding methods for early detection of invasive fish species. *Sci Rep* 7. <https://doi.org/10.1038/srep46393>
- Henderson, P.A., Turnpenny, A.W.H., Bamber, R.N., 1984. Long-term stability of a sand smelt (*Atherina presbyter* Cuvier) population subject to power station cropping. *Journal of Applied Ecology* 1–10.
- Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., Bright, I.J., Lucero, M.Y., Hiddessen, A.L., Legler, T.C., Kitano, T.K., Hodel, M.R., Petersen, J.F., Wyatt, P.W., Steenblock, E.R., Shah, P.H., Bousse, L.J., Troup, C.B., Mellen, J.C., Wittmann, D.K., Erndt, N.G., Cauley, T.H., Koehler, R.T., So, A.P., Dube, S., Rose, K.A., Montesclaros, L., Wang, S., Stumbo, D.P., Hodges, S.P., Romine, S., Milanovich, F.P., White, H.E., Regan, J.F., Karlin-Neumann, G.A., Hindson, C.M., Saxonov, S., Colston, B.W., 2011. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Anal Chem* 83, 8604–8610. <https://doi.org/10.1021/ac202028g>
- Iglesias, S., 2020. *Piscibus Marinis - Guide des poisons marins, Europe et eaux adjacentes (Une classification naturelle basée sur des spécimens de collection, des barcodes ADN et des photos standardisées)*, (No. Version provisoire 11).

- Jackson, E.L., Attrill, M.J., Rowden, A.A., Jones, M.B., 2006. Seagrass complexity hierarchies: Influence on fish groups around the coast of Jersey (English Channel). *Journal of Experimental Marine Biology and Ecology*, A Tribute to Richard M. Warwick 330, 38–54. <https://doi.org/10.1016/j.jembe.2005.12.016>
- Jeunen, G.-J., Knapp, M., Spencer, H.G., Lamare, M.D., Taylor, H.R., Stat, M., Bunce, M., Gemmell, N.J., 2019. Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. *Mol Ecol Resour* 19, 426–438. <https://doi.org/10.1111/1755-0998.12982>
- Jeunen, G.-J., Lamare, M.D., Knapp, M., Spencer, H.G., Taylor, H.R., Stat, M., Bunce, M., Gemmell, N.J., 2020. Water stratification in the marine biome restricts vertical environmental DNA (eDNA) signal dispersal. *Environmental DNA* 2, 99–111. <https://doi.org/10.1002/edn3.49>
- Juhel, J.-B., Utama, R.S., Marques, V., Vimono, I.B., Sugeha, H.Y., Kadarusman, null, Pouyau, L., Dejean, T., Mouillot, D., Hocdé, R., 2020. Accumulation curves of environmental DNA sequences predict coastal fish diversity in the coral triangle. *Proceedings of the Royal Society B: Biological Sciences* 287, 20200248. <https://doi.org/10.1098/rspb.2020.0248>
- Kelly, R.P., O'Donnell, J.L., Lowell, N.C., Shelton, A.O., Samhour, J.F., Hennessey, S.M., Feist, B.E., Williams, G.D., 2016. Genetic signatures of ecological diversity along an urbanization gradient. *PeerJ* 4, e2444. <https://doi.org/10.7717/peerj.2444>
- Kelly, R.P., Port, J.A., Yamahara, K.M., Crowder, L.B., 2014. Using Environmental DNA to Census Marine Fishes in a Large Mesocosm. *PLOS ONE* 9, e86175. <https://doi.org/10.1371/journal.pone.0086175>
- Lacoursière-Roussel, A., Deiner, K., 2019. Environmental DNA is not the tool by itself. *Journal of Fish Biology* n/a. <https://doi.org/10.1111/jfb.14177>
- Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G., Taylor, M.I., 2019. How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology* 28, 420–430. <https://doi.org/10.1111/mec.14920>
- Le Joncour, A., Thiriet, P., 2020. Surveillance des poissons côtiers rocheux de Manche Occidentale par comptages visuels en plongée sous-marine : structure et contenu des bases de données générées par les suivis POCOROCH 2016-2019. Rapport UMS PatriNat.
- Legendre, P., Legendre, L., 2012. *Numerical ecology*. Elsevier.
- Li, J., Hatton-Ellis, T.W., Handley, L.-J.L., Kimbell, H.S., Benucci, M., Peirson, G., Hänfling, B., 2019. Ground-truthing of a fish-based environmental DNA metabarcoding method for assessing the quality of lakes. *Journal of Applied Ecology* 56, 1232–1244. <https://doi.org/10.1111/1365-2664.13352>
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., Dunthorn, M., 2015. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* 3. <https://doi.org/10.7717/peerj.1420>
- Maran, V., Chanet, B., 2011. Observations du crénilabre de Baillon *Symphodus bailloni* (Valenciennes, 1839), Labridae, dans et autour de la rade de Brest (Nord Bretagne, France). *Bulletin de la Société des sciences naturelles de l'Ouest de la France* (1983) 33, 57–60.
- Marques, V., Milhau, T., Albouy, C., Dejean, T., Manel, S., Mouillot, D., Juhel, J.-B., 2021. GAPeDNA: Assessing and mapping global species gaps in genetic databases for eDNA metabarcoding. *Diversity and Distributions* n/a. <https://doi.org/10.1111/ddi.13142>
- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. <https://doi.org/10.14806/ej.17.1.200>
- Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J.Y., Sato, K., Minamoto, T., Yamamoto, S., Yamanaka, H., Araki, H., others, 2015. MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society open science* 2, 150088.
- Moreno, T., Castro, J.J., Socorro, J., 2005. Reproductive biology of the sand smelt *Atherina presbyter* Cuvier, 1829 (Pisces: Atherinidae) in the central-east Atlantic. *Fisheries Research* 72, 121–131. <https://doi.org/10.1016/j.fishres.2004.06.016>

- Nester, G.M., Brauwer, M.D., Koziol, A., West, K.M., DiBattista, J.D., White, N.E., Power, M., Heydenrych, M.J., Harvey, E., Bunce, M., 2020. Development and evaluation of fish eDNA metabarcoding assays facilitate the detection of cryptic seahorse taxa (family: Syngnathidae). *Environmental DNA* 2, 614–626. <https://doi.org/10.1002/edn3.93>
- Oka, S., Doi, H., Miyamoto, K., Hanahara, N., Sado, T., Miya, M., 2021. Environmental DNA metabarcoding for biodiversity monitoring of a highly diverse tropical fish community in a coral reef lagoon: Estimation of species richness and detection of habitat segregation. *Environmental DNA* 3, 55–69. <https://doi.org/10.1002/edn3.132>
- Pawlowski, J., Apothéloz-Perret-Gentil, L., Altermatt, F., 2020. Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. *Molecular Ecology* 29, 4258–4264. <https://doi.org/10.1111/mec.15643>
- Pita, P., Fernández-Márquez, D., Freire, J., 2018. Spatiotemporal variation in the structure of reef fish and macroalgal assemblages in a north-east Atlantic kelp forest ecosystem: implications for the management of temperate rocky reefs. *Mar. Freshwater Res.* 69, 525–541. <https://doi.org/10.1071/MF17193>
- Polanco Fernández, A., Marques, V., Fopp, F., Juhel, J.-B., Borrero-Pérez, G.H., Cheutin, M.-C., Dejean, T., González Corredor, J.D., Acosta-Chaparro, A., Hocdé, R., 2021. Comparing environmental DNA metabarcoding and underwater visual census to monitor tropical reef fishes. *Environmental DNA*.
- Port, J.A., O'Donnell, J.L., Romero-Maraccini, O.C., Leary, P.R., Litvin, S.Y., Nickols, K.J., Yamahara, K.M., Kelly, R.P., 2016. Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Mol. Ecol.* 25, 527–541. <https://doi.org/10.1111/mec.13481>
- Prato, G., Thiriet, P., Franco, A.D., Francour, P., 2017. Enhancing fish Underwater Visual Census to move forward assessment of fish assemblages: An application in three Mediterranean Marine Protected Areas. *PLOS ONE* 12, e0178511. <https://doi.org/10.1371/journal.pone.0178511>
- Rey, A., Gauthier, O., Delesalle, M., Le Joncour, A., Ternon, Q., Thiriet, P., 2021. Développement méthodologique d'un dispositif de surveillance et d'observation des poissons dans les milieux côtiers rocheux de la façade Atlantique-Manche-Mer du Nord, par comptages visuels en plongée sous-marine. Phase 1 : validation des protocoles et de la stratégie d'échantillonnage intra-site. Rapport UMS PatriNat.
- Rey, A., Viard, F., Corre, E., Lizé, A., Lévêque, L., Robert, A., Feunteun, E., Thiriet, P., 2019. La surveillance de la biodiversité ichtyologique côtière par l'ADN environnemental : Evaluation de l'outil et de sa complémentarité aux comptages visuels en plongée. Rapport intermédiaire, juillet 2019. Rapport UMS PatriNat.
- Riaz, M., Kuemmerlen, M., Wittwer, C., Cocchiararo, B., Khaliq, I., Pfenninger, M., Nowak, C., 2020. Combining environmental DNA and species distribution modeling to evaluate reintroduction success of a freshwater fish. *Ecological Applications* 30, e02034. <https://doi.org/10.1002/eap.2034>
- Rourke, M., Fowler, A., Hughes, J., Broadhurst, M., Dibattista, J., Fielder, D., Wilkes Walburn, J., Furlan, E., 2021. Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. *Environmental DNA*. <https://doi.org/10.1002/edn3.185>
- Shu, L., Ludwig, A., Peng, Z., 2020. Standards for Methods Utilizing Environmental DNA for Detection of Fish Species. *Genes (Basel)* 11. <https://doi.org/10.3390/genes11030296>
- Sigsgaard, E.E., Nielsen, I.B., Carl, H., Krag, M.A., Knudsen, S.W., Xing, Y., Holm-Hansen, T.H., Møller, P.R., Thomsen, P.F., 2017. Seawater environmental DNA reflects seasonality of a coastal fish community. *Mar Biol* 164, 128. <https://doi.org/10.1007/s00227-017-3147-4>
- Sigsgaard, E.E., Torquato, F., Frøslev, T.G., Moore, A.B.M., Sørensen, J.M., Range, P., Ben-Hamadou, R., Bach, S.S., Møller, P.R., Thomsen, P.F., 2020. Using vertebrate environmental DNA from

- seawater in biomonitoring of marine habitats. *Conservation Biology* 34, 697–710.
<https://doi.org/10.1111/cobi.13437>
- Stat, M., John, J., DiBattista, J.D., Newman, S.J., Bunce, M., Harvey, E.S., 2019. Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conserv. Biol.* 33, 196–205. <https://doi.org/10.1111/cobi.13183>
- Stoeckle, M.Y., Adolf, J., Charlop-Powers, Z., Dunton, K.J., Hinks, G., VanMorter, S.M., 2020a. Trawl and eDNA assessment of marine fish diversity, seasonality, and relative abundance in coastal New Jersey, USA. *ICES Journal of Marine Science*.
- Stoeckle, M.Y., Das Mishu, M., Charlop-Powers, Z., 2020b. Improved Environmental DNA Reference Library Detects Overlooked Marine Fishes in New Jersey, United States. *Front. Mar. Sci.* 7. <https://doi.org/10.3389/fmars.2020.00226>
- Stoeckle, M.Y., Soboleva, L., Charlop-Powers, Z., 2017. Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. *PLOS ONE* 12, e0175186. <https://doi.org/10.1371/journal.pone.0175186>
- Taberlet, P., Bonin, A., Zinger, L., Coissac, E., 2018. *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford University Press.
- Taberlet, P., Coissac, E., Hajibabaei, M., Rieseberg, L.H., 2012. Environmental DNA. *Molecular Ecology* 21, 1789–1793. <https://doi.org/10.1111/j.1365-294X.2012.05542.x>
- Teichert, N., Lepage, M., Sagouis, A., Borja, A., Chust, G., Ferreira, M.T., Pasquaud, S., Schinegger, R., Segurado, P., Argillier, C., 2017a. Functional redundancy and sensitivity of fish assemblages in European rivers, lakes and estuarine ecosystems. *Scientific Reports* 7, 17611. <https://doi.org/10.1038/s41598-017-17975-x>
- Teichert, N., Pasquaud, S., Borja, A., Chust, G., Uriarte, A., Lepage, M., 2017b. Living under stressful conditions: Fish life history strategies across environmental gradients in estuaries. *Estuarine, Coastal and Shelf Science* 188, 18–26. <https://doi.org/10.1016/j.ecss.2017.02.006>
- Thalinger, B., Deiner, K., Harper, L.R., Rees, H.C., Blackman, R.C., Sint, D., Traugott, M., Goldberg, C.S., Bruce, K., 2021. A validation scale to determine the readiness of environmental DNA assays for routine species monitoring. *Environmental DNA* n/a. <https://doi.org/10.1002/edn3.189>
- Thalinger, B., Wolf, E., Traugott, M., Wanzenböck, J., 2019. Monitoring spawning migrations of potamodromous fish species via eDNA. *Scientific Reports* 9, 15388. <https://doi.org/10.1038/s41598-019-51398-0>
- Thanopoulou, Z., Sini, M., Vatikiotis, K., Katsoupis, C., Dimitrakopoulos, P.G., Katsanevakis, S., 2018. How many fish? Comparison of two underwater visual sampling methods for monitoring fish communities. *PeerJ* 6. <https://doi.org/10.7717/peerj.5066>
- Thiriet, P.D., Franco, A.D., Cheminée, A., Guidetti, P., Bianchimani, O., Basthard-Bogain, S., Cottalorda, J.-M., Arceo, H., Moranta, J., Lejeune, P., Francour, P., Mangialajo, L., 2016a. Abundance and Diversity of Crypto- and Necto-Benthic Coastal Fish Are Higher in Marine Forests than in Structurally Less Complex Macroalgal Assemblages. *PLOS ONE* 11, e0164121. <https://doi.org/10.1371/journal.pone.0164121>
- Thiriet, P.D., Franco, A.D., Cheminée, A., Guidetti, P., Bianchimani, O., Basthard-Bogain, S., Cottalorda, J.-M., Arceo, H., Moranta, J., Lejeune, P., Francour, P., Mangialajo, L., 2016b. Abundance and Diversity of Crypto- and Necto-Benthic Coastal Fish Are Higher in Marine Forests than in Structurally Less Complex Macroalgal Assemblages. *PLOS ONE* 11, e0164121. <https://doi.org/10.1371/journal.pone.0164121>
- Thomsen, P.F., Kielgast, J., Iversen, L.L., Møller, P.R., Rasmussen, M., Willerslev, E., 2012. Detection of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples. *PLOS ONE* 7, e41732. <https://doi.org/10.1371/journal.pone.0041732>
- Thomsen, P.F., Møller, P.R., Sigsgaard, E.E., Knudsen, S.W., Jørgensen, O.A., Willerslev, E., 2016. Environmental DNA from Seawater Samples Correlate with Trawl Catches of Subarctic, Deepwater Fishes. *PLOS ONE* 11, e0165252. <https://doi.org/10.1371/journal.pone.0165252>

- Tsuji, S., Shibata, N., 2021. Identifying spawning events in fish by observing a spike in environmental DNA concentration after spawning. *Environmental DNA* 3, 190–199.
<https://doi.org/10.1002/edn3.153>
- Valdivia-Carrillo, T., Rocha-Olivares, A., Reyes-Bonilla, H., Domínguez-Contreras, J.F., Munguia-Vega, A., 2021. Integrating eDNA metabarcoding and simultaneous underwater visual surveys to describe complex fish communities in a marine biodiversity hotspot. *Molecular Ecology Resources* n/a. <https://doi.org/10.1111/1755-0998.13375>
- Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P.F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G.H., Geniez, P., Pont, D., Argillier, C., Baudoin, J.-M., Peroux, T., Crivelli, A.J., Olivier, A., Acqueberge, M., Le Brun, M., Møller, P.R., Willerslev, E., Dejean, T., 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Mol. Ecol.* 25, 929–942.
<https://doi.org/10.1111/mec.13428>
- Wafar, M.V.M., Le Corre, P., Birrien, J.L., 1983. Nutrients and primary production in permanently well-mixed temperate coastal waters. *Estuarine, Coastal and Shelf Science* 17, 431–446.
[https://doi.org/10.1016/0272-7714\(83\)90128-2](https://doi.org/10.1016/0272-7714(83)90128-2)
- Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* 73, 5261–5267.
<https://doi.org/10.1128/AEM.00062-07>
- Wang, S., Yan, Z., Hänfling, B., Zheng, X., Wang, P., Fan, J., Li, J., 2021. Methodology of fish eDNA and its applications in ecology and environment. *Science of The Total Environment* 755, 142622.
<https://doi.org/10.1016/j.scitotenv.2020.142622>
- Ward, R.D., 2012. FISH-BOL, A Case Study for DNA Barcodes, in: *DNA Barcodes*. Humana Press, Totowa, NJ, pp. 423–439. https://doi.org/10.1007/978-1-61779-591-6_21
- Winnepenninckx, B., Backeljau, T., De Wachter, R., 1994. Small ribosomal subunit RNA and the phylogeny of the Mollusca. *NAUTILUS-GREENVILLE* 98–110.
- Xia, Z., Zhan, A., Johansson, M.L., DeRoy, E., Haffner, G.D., Maclsaac, H.J., 2021. Screening marker sensitivity: Optimizing eDNA-based rare species detection. *Diversity and Distributions* n/a.
<https://doi.org/10.1111/ddi.13262>
- Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., Minamoto, T., Miya, M., 2017. Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports* 7, 40368. <https://doi.org/10.1038/srep40368>
- Yamamoto, S., Minami, K., Fukaya, K., Takahashi, K., Sawada, H., Murakami, H., Tsuji, S., Hashizume, H., Kubonaga, S., Horiuchi, T., Hongo, M., Nishida, J., Okugawa, Y., Fujiwara, A., Fukuda, M., Hidaka, S., Suzuki, K.W., Miya, M., Araki, H., Yamanaka, H., Maruyama, A., Miyashita, K., Masuda, R., Minamoto, T., Kondoh, M., 2016. Environmental DNA as a ‘Snapshot’ of Fish Distribution: A Case Study of Japanese Jack Mackerel in Maizuru Bay, Sea of Japan. *PLOS ONE* 11, e0149786. <https://doi.org/10.1371/journal.pone.0149786>

8 Annexes

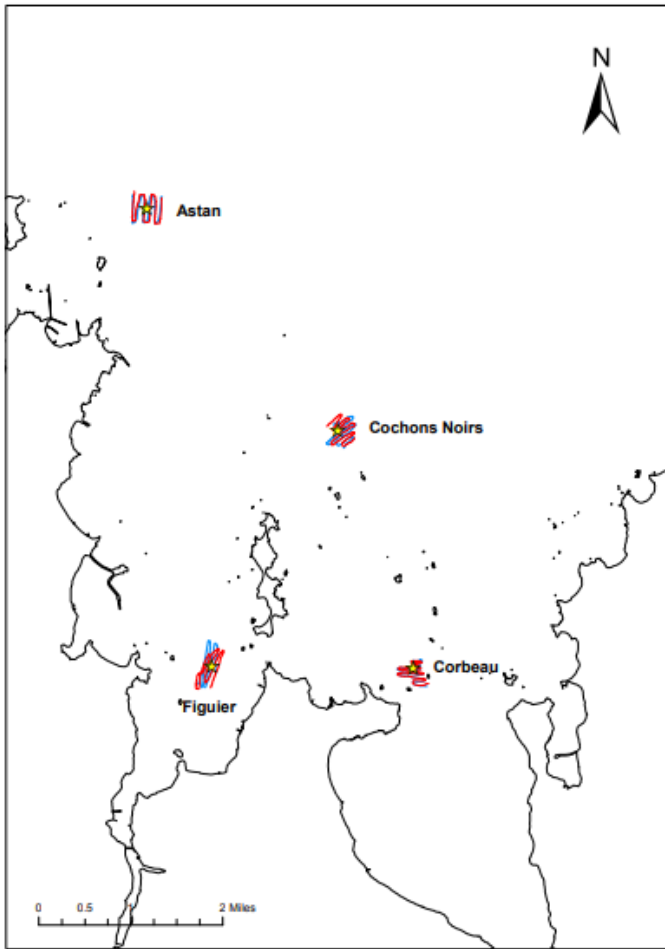


Figure 1. Traits d'échantillonnage parcourus pendant 30 minutes à chaque site pour le type de prélèvement d'eau ADNe « 30L area surface »

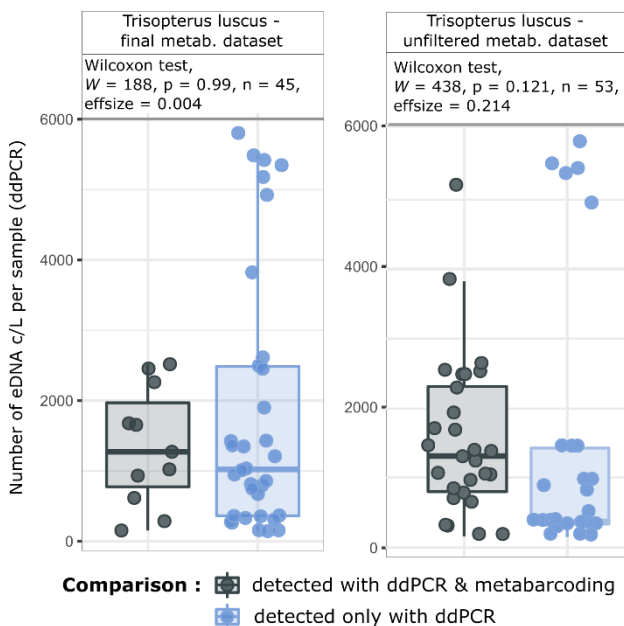


Figure 2. Comparaison du nombre de copies estimé en ddPCR entre les échantillons où la cible est retrouvée avec la ddPCR et le metabarcoding (gris) et les échantillons où la cible est détectée seulement avec la ddPCR (bleu) pour le tacaud (*T.luscus*) en fonction des jeux de données metabarcoding : jeu de données final utilisé dans l'étude (à gauche) et jeu de données après assignations taxinomiques et qui n'a pas encore subi les étapes finales de nettoyage (à droite).

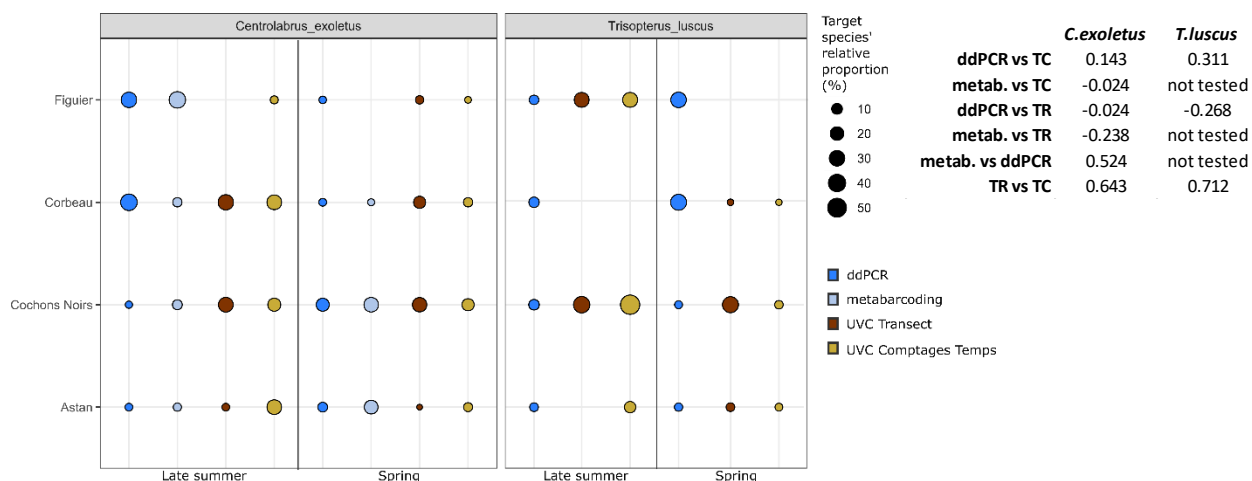


Figure 3. Proportion relative du centrolabre (*C.exoletus*) et du tacaud (*T.luscus*) dans chaque site à une saison pour chaque approche (ddPCR, metabarcoding, UVC Transect, UVC Comptage Temps) par rapport à l'abondance totale de l'espèce estimée par chaque approche. L'abondance pour le jeu UVC Transect représente le nombre d'individus par m², pour le jeu UVC Comptage Temps, le nombre d'individus par classe, pour le metabarcoding le nombre de séquence et pour le jeu ddPCR le nombre de copie. L'absence de point signifie une valeur d'abondance de 0. Le tableau à droite indique les résultats de la corrélation de Spearman basée sur les rangs entre les patrons quantitatifs observés en ddPCR, metabarcoding, UVC Transect et UVC Comptage Temps au niveau des 8 sites. NB : Le tacaud n'est pas considéré qu'avec la ddPCR pour les approches sur ADN puisque le metabarcoding semble sous-estimer la détection de cette espèce.

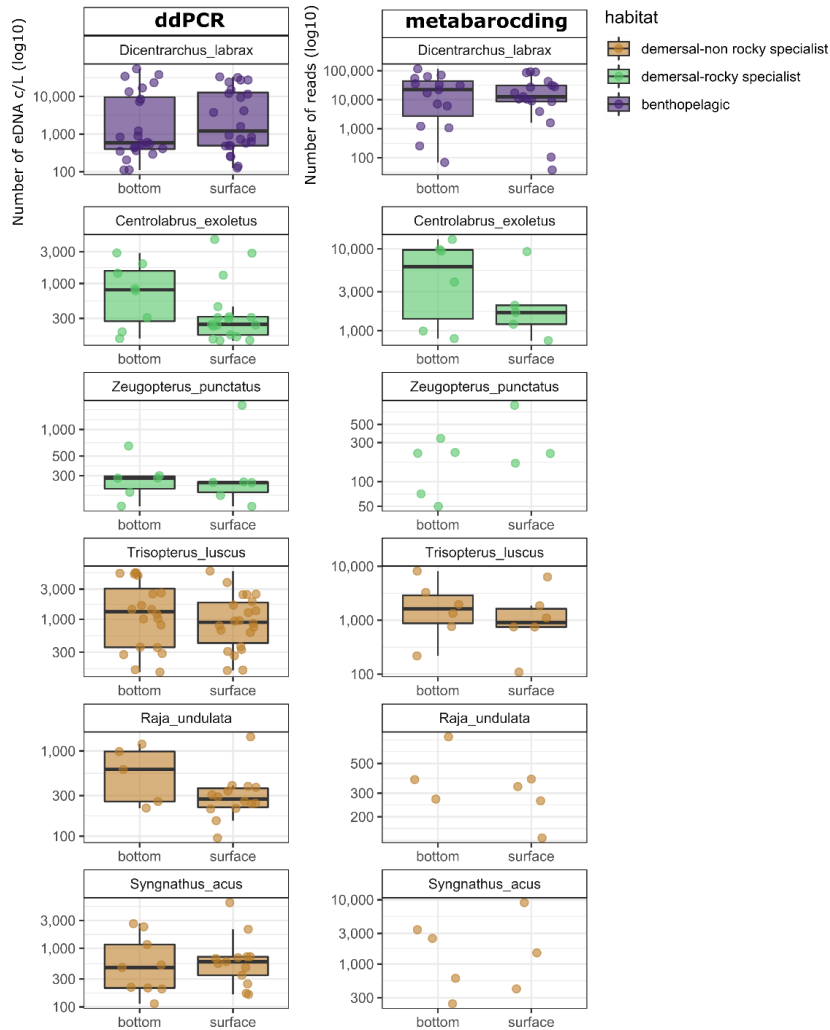


Figure 4. Distribution des quantité d'ADNe obtenues en ddPCR et en metabarcoding pour les échantillons pris en surface et au-dessus du fond pour les espèces cibles. Les boîtes à moustache sont ajoutées seulement si le nombre d'échantillons où la cible est détectée est supérieur à 5 pour les 2 types de prélèvement.

- Les comparaisons de moyenne entre les 2 types de prélèvement ADNe pour chaque cible et approche effectuées avec le test de Wilcoxon sont non-significatives ($p > 0.05$)
- Il n'y a pas d'impact du type de prélèvement sur la quantité d'ADNe des espèces cibles qu'importe leur préférence d'habitat, ce qui reste cohérent avec l'hypothèse de l'absence de stratification des eaux en baie de Morlaix.

Tableau 1. Résumé des étapes bio-informatiques des données de metabarcoding

Bioinformatic step	Software (version)	Command	Parameters added or for which the default parameter was changed
Control quality of raw sequences	FASTQC (0.11.8)	fastqc	default parameters
Demultiplexing on tag	Cutadapt (2.3)	cutadapt	--no-indels; --discard-untrimmed ; --pair-adapters; -overlap 8 ; -minimum-length 107
Demultiplexing on TELEO primers	Cutadapt (2.8)	cutadapt	-a tel_f=^ACACCGCCCGTCACTCT...CATGGTAAGTGTACCGGAAG; -A tel_f=^CTTCCGGTACTTACCATG...AGAGTGACGGGCGGTGT; -a tel_r=^CTTCCGGTACTTACCATG...AGAGTGACGGGCGGTGT; -A tel_r=^ACACCGCCCGTCACTCT...CATGGTAAGTGTACCGGAAG; -m 45; -M 96; --discard-untrimmed; --pair-adapters --no-indels
Demultiplexing on METAZ primers	Cutadapt (2.8)	cutadapt	-g met_f=^AGTTACYYTAGGGATAACAGCG; -G met_f=^CCGGTCTGAACTCAGATCAYGT; -g met_r=^CCGGTCTGAACTCAGATCAYGT; -G met_r=^AGTTACYYTAGGGATAACAGCG; --discard-untrimmed; --pair-adapters; --no-indels
Reverse Complement to have in each sample only one orientation of reads	Obitools (1.2.11)	obicomplement	default parameters
Quality filtering of reads	DADA2 (1.14.0)	filterAndTrim	maxEE=1; truncQ=2; maxN=0
Learn forward and reverse error rates	DADA2 (1.14.0)	learnErrors	nbases=1e8, randomize=random
Dereplicating of reads	DADA2 (1.14.0)	derepFatsq	default parameters
Removal of sequencing errors	DADA2 (1.14.0)	dada	default parameters
Merging forward and reverse reads	DADA2 (1.14.0)	mergePairs	default parameters
Chimera removal	DADA2 (1.14.0)	removeBimera Denovo	default parameters
Fragment size filtering on TELEO fragment	custom R script	na	minimum length was 50 bp and maximum length was 96 bp
Fragment size filtering on METAZ fragment	custom R script	na	minimum length was 110 bp and maximum length was 120 bp
Taxonomic assignment with RDP Classifier	DADA2 (1.14.0)	assignTaxonomy	default parameters
Taxonomic assignment with BLAST	Blast (2.7.1)	blastn	-max_target_seqs 100

Taxonomic assignment with ECOTAG	Obitools (1.2.11)	ecotag	default parameters
Index jumping filtering (for HISEQ libraries)	custom R script	na	na
Normalization between libraries (for 12S libraries only)	custom R script	na	na
Filtering on targeted taxa (Actinopteri and Chondrichthyes)	custom R script	na	na
LULU curation of ASVs on TELEO fragment	LULU (0.1.0)	lulu	minimum_ratio_type = "min"; minimum_ratio = 1; minimum_match = 96; minimum_relative_cooccurrence = 0.95
LULU curation of ASVs on METAZ fragment	LULU (0.1.0)	lulu	minimum_ratio_type = "min"; minimum_ratio = 1; minimum_match = 98; minimum_relative_cooccurrence = 0.95
Clustering of ASVs into OTUs	SWARM (3.0.0)	swarm	-d 1; -f
Removal of OTUs found only once in the dataset	custom R script	na	na

Tableau 2. Espèces (n=48) pour lesquelles une séquence de référence a été obtenue lors de ce projet avec les marqueurs 12S et/ou 16S. Un astérisque est ajouté quand l'espèce n'avait pas de séquence déjà publiée dans les bases publiques au moment de la recherche en Mai 2020 ainsi qu'au moment de la rédaction du rapport en Mai 2021. Les séquences obtenues lors de ce projet seront rajoutées dans les bases de références publiques courant 2021.

Espèce	Marqueur
<i>Arnoglossus imperialis</i>	12S*
<i>Atherina presbyter</i>	12S*
<i>Chelidonichthys lucerna</i>	12S*
<i>Coryphoblennius galerita</i>	12S*
<i>Hyperoplus immaculatus</i>	12S*
<i>Lepidorhombus whiffiagonis</i>	12S*
<i>Lipophrys pholis</i>	12S*
<i>Nerophis lumbriciformis</i>	12S
<i>Scophthalmus rhombus</i>	12S*
<i>Buglossidium luteum</i>	12S* & 16S
<i>Labrus bergylta</i>	12S & 16S
<i>Lepadogaster lepadogaster</i>	12S* & 16S
<i>Symphodus bailloni</i>	12S* & 16S
<i>Taurulus bubalis</i>	12S* & 16S
<i>Zeugopterus punctatus</i>	12S & 16S
<i>Conger conger</i>	16S*
<i>Callionymus lyra</i>	16S
<i>Lepadogaster candolii</i>	16S
<i>Mullus surmuletus</i>	16S
<i>Parablennius gattorugine</i>	16S
<i>Pleuronectes platessa</i>	16S
<i>Spondyliosoma cantharus</i>	16S
<i>Symphodus melops</i>	16S
<i>Alosa fallax</i>	12S
<i>Anguilla anguilla</i>	12S
<i>Aphia minuta</i>	12S
<i>Centrolabrus exoletus</i>	12S
<i>Dicentrarchus labrax</i>	12S
<i>Engraulis encrasicolus</i>	12S
<i>Entelurus aequoreus</i>	12S
<i>Gobius cobitis</i>	12S
<i>Gobius niger</i>	12S
<i>Hippocampus hippocampus</i>	12S
<i>Leucoraja fullonica</i>	12S
<i>Leucoraja naevus</i>	12S
<i>Merluccius merluccius</i>	12S
<i>Nerophis ophidion</i>	12S
<i>Pholis gunnellus</i>	12S
<i>Raja brachyura</i>	12S
<i>Raja clavata</i>	12S

<i>Raja microocellata</i>	12S
<i>Raja montagui</i>	12S
<i>Raja undulata</i>	12S
<i>Syngnathus acus</i>	12S
<i>Syngnathus rostellatus</i>	12S
<i>Syngnathus typhle</i>	12S
<i>Tripterygion delaisi</i>	12S
<i>Ciliata mustela</i>	12S & 16S

Tableau 3. Nombre de séquences conservées après chaque étape bioinformatique de nettoyage et filtration des données

Step	TOTAL 12S	TOTAL 16S
Raw	370,258,632	
Demultiplexed	76,446,331	175,683,371
DADA2:filtered	76,251,344	171,567,834
DADA2:denoisedR	76,199,443	171,373,853
DADA2:denoisedF	76,202,688	171,350,043
DADA2:merged	75,519,787	167,225,163
DADA2:no_chimera	74,758,065	164,874,170
DADA2:expected_lenght_barcode	74,662,954	162,461,856
Index jumping filtering (for HISEQ librairies only)	57,720,261	161,995,523
Normalization between librairies (for 12S librairies only)	30,853,377	161,995,523
Target taxa (Actinopteri and Chondrichthyes)	18,867,477	33,213,663
LULU curation	18,867,477	33,213,663
Sequences clustered into OTUs with SWARM	18,867,477	33,213,663
Sequences after removal of OTUs found only once	18,782,045	33,197,741

Tableau 4. Filtre du **Tableau 7 du rapport** sur les taxons non détectés avec les amorces 12S Teleo utilisées dans cette étude. La colonne 12S – Amorces Mifish a été ajoutée au tableau et présente la disponibilité des séquences de référence pour un autre couple d’amorce fréquemment utilisé pour les analyses metabarcoding sur ADNe ciblant les poissons (Miya et al., 2015). Disponibilité de séquences de référence pour chacun des taxons (au genre ou à l’espèce) identifiés par comptages visuels. Trois statuts sont décrits dans les colonnes des marqueurs 12S et 16S : NON (rouge) : pas de séquence de référence disponible, OUI (vert) : séquence diagnostique (i.e. permettant l’identification de l’espèce) et ND (orange) : séquence disponible mais non diagnostique. Dans le cas où les séquences de référence n’étaient pas disponibles au moment de la création de la base de données en Mai 2020, nous avons vérifié si c’était toujours le cas au moment de la rédaction du rapport (Mai 2021). Il n’y a pas eu d’ajout de séquence dans la base publique NCBI depuis Mai 2020 pour les espèces sans séquences à l’exception de *Mullus surmuletus* où un génome mitochondrial complet a été rajouté, rendant donc disponible une séquence de référence pour les marqueurs 12S et 16S.

Taxon vu en comptage visuel	Nombre d'échantillon dans UVC Baie (total 34)	Méthodes d'observation			Disponibilité de séquences de référence		
		UVC Baie	UVC 24h	ADNe	16S	12S – Amorces Teleo (utilisées ici)	12S – Amorces Mifish (non utilisées)
<i>Pleuronectes platessa</i>	1	OUI	OUI	NON	ND	ND	ND avec <i>P. quadrituberculatus</i> mais différentiable à 1base près des esp. <i>Platichthys stellatus</i> , <i>P. flesus</i> et <i>Myzopsetta ferruginea</i> contrairement aux amorces à Teleo
<i>Pagrus pagrus</i>	2	OUI	NON	NON	ND	ND	OUI
<i>Labrus bergylta</i>	33	OUI	OUI	NON	ND	ND	ND entre <i>L. bergylta</i> et <i>L. merula</i>
<i>Pollachius pollachius</i>	33	OUI	OUI	NON	ND	ND	OUI dont <i>P. virens</i> à 1 base près
<i>Scyliorhinus stellaris</i>	1	OUI	NON	NON	NON	NON	NON
<i>Parablennius ruber</i>	1	OUI	OUI	NON	ND	NON	NON
<i>Parablennius pilicornis</i>	15	OUI	OUI	NON	ND	NON	NON
<i>Conger conger</i>	19	OUI	OUI	NON	ND	NON	NON
<i>Parablennius gattorugine</i>	32	OUI	OUI	NON	ND	NON	NON
<i>Lepadogaster candolii</i>	1	OUI	NON	OUI	OUI	NON	NON
<i>Mullus surmuletus</i>	14	OUI	OUI	OUI	OUI	NON	OUI
<i>Callionymus lyra</i>	19	OUI	OUI	OUI	OUI	NON	NON
<i>Gaidropsarus spp</i>	3	OUI	NON	NON	ND	NON	ND
<i>Lepadogaster spp</i>	15	OUI	OUI	NON	ND (<i>L. lepadogaster</i>)	OUI (<i>L. lepadogaster</i>)	NON
<i>Tripterygion delaisi</i>	3	OUI	NON	NON	ND	OUI	OUI
<i>Zeus faber</i>	3	OUI	NON	NON	ND	OUI	OUI

Tableau 5. Analyses de PERMANOVA des jeux de données en présence/absence des 2 méthodes d'observation évaluant l'impact des facteurs type de prélèvement ADN (pour metabarcoding seulement), site et saison (pour UVC et metabarcoding) sur les patrons des communautés. Les noms des facteurs sont les suivants : AST pour ASTAN, CN pour COCHONS NOIRS, COR pour CORBEAU, FIG pour FIGUIER, SPR pour SPRING, AUT pour AUTUMN, 2L.S pour 2L fixed surface, 2L.F pour 2L fixed bottom pour 30L.S for 30L area surface. Df : degrés de liberté, SS : sommes des carrés, Pseudo-F : valeur du test, P(perms) : p-valeur associée au test, perms : nombre de permutations. Les tests pour lesquels l'homogénéité des variances intra-groupes n'était pas respectée (p-valeur < 0.05 au test betadisper) sont indiqués par un astérisque (*) dans la colonne Pseudo-F. ** indique que le nombre de permutation était trop faible pour effectuer les « pairwise » test quand l'interaction entre facteurs était significative à la p-valeur = 0.05.

Dataset	Type of PERMANOVA	Output of the test																									
2L fixed water sample dataset (125)	Main test	Source	df	SS	Pseudo-F	P(perms)	perms																				
		locality	3	3,7293	2,7363	0,0001	9809																				
		seasonality	1	2,2511	4,9552	0,0001	9883																				
		type of water sample	1	0,69172	1,5226	0,0135	9856																				
		locality x seasonality	3	3,1188	2,2883	0,0001	9806																				
		locality x type of water sample	3	1,897	1,3919	0,002	9786																				
		seasonality x type of water sample	1	0,4923	1,0837	0,3256	9864																				
		locality x seasonality x type of water sample	3	1,617	1,1864*	0,067	9780																				
		Residuals	32	14,538																							
		Total	47	28,335																							
		2L fixed water sample dataset (125)	Pairwise test locality x seasonality	Within level 'AUT' of factor 'seasonality'						Within level 'SPR' of factor 'seasonality'						Within level 'COR' of factor 'locality'						Within level 'FIG' of factor 'locality'					
				Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms				
AST, CN	1,3431			0,0087	8844	AST, CN	1,3544	0,0041	8848	AST, CN	1,3544	0,0041	8848	AST, CN	1,3544	0,0041	8848	AST, CN	1,3544	0,0041	8848						
AST, COR	1,5049			0,0021	8848	AST, COR	1,2	0,0672	8872	AST, COR	1,2	0,0672	8872	AST, COR	1,2	0,0672	8872	AST, COR	1,2	0,0672	8872						
AST, FIG	1,7679			0,002	8828	AST, FIG	1,3508	0,0053	8864	AST, FIG	1,3508	0,0053	8864	AST, FIG	1,3508	0,0053	8864	AST, FIG	1,3508	0,0053	8864						
CN, COR	2,0654			0,0013	8865	CN, COR	1,4714	0,0061	8866	CN, COR	1,4714	0,0061	8866	CN, COR	1,4714	0,0061	8866	CN, COR	1,4714	0,0061	8866						
Pairwise test locality x type of water sample	Within level 'AST' of factor 'locality'						Within level 'CN' of factor 'locality'						Within level 'COR' of factor 'locality'						Within level 'FIG' of factor 'locality'								
	Groups		t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms						
	AUT, SPR		1,2672	0,0306	8853	AUT, SPR	2,0006	0,0023	8820	AUT, SPR	1,7768	0,0028	8890	AUT, SPR	1,9474	0,0022	8900	AUT, SPR	1,9474	0,0022	8900						
	Within level '2L.B' of factor 'type of water sample'						Within level '2L.S' of factor 'type of water sample'																				
	Groups		t	P(perms)	perms	Groups	t	P(perms)	perms																		
	AST, CN		1,3054	0,0293	8849	AST, CN	1,201	0,0445	8854																		
2L fixed water sample dataset (165)	Main test	Source	df	SS	Pseudo-F	P(perms)	perms																				
		locality	3	2,3781	2,4419	0,0001	9836																				
		seasonality	1	1,8671	5,7517	0,0001	9912																				
		type of water sample	1	0,61889	1,9065	0,0098	9914																				
		locality x seasonality	3	2,2692	2,3301	0,0002	9849																				
		locality x type of water sample	3	1,1585	1,1895	0,1442	9834																				
		seasonality x type of water sample	1	0,36831	1,1346	0,299	9892																				
		locality x seasonality x type of water sample	3	1,204	1,2363*	0,1019	9847																				
		Residuals	32	10,388																							
		Total	47	20,252																							
		All types of water sample, AUT only (125)	Main test	Source	df	SS	Pseudo-F	P(perms)	perms																		
				locality	3	5,0614	4,05	0,0001	9834																		
type of water sample	2			1,0281	1,234	0,0673	9796																				
locality x type of water sample	6			2,7361	1,0955*	0,1251	9650																				
Residuals	28			11,664																							
Total	39			20,454																							
All types of water sample, AUT only (165)	Main test			Source	df	SS	Pseudo-F	P(perms)	perms																		
				locality	3	2,4458	2,8654	0,0001	9849																		
				type of water sample	2	1,073	1,8855	0,0013	9885																		
				locality x type of water sample**	6	2,3596	1,3822	0,0028	9777																		
				Residuals	28	7,9666																					
				Total	39	13,835																					
		UVC 24h	Main test	Source	df	SS	Pseudo-F	P(perms)	perms																		
				locality	3	5,1701	5,2916	0,0001	9875																		
				seasonality	1	0,7171	2,2019	0,0146	9927																		
				locality x seasonality	3	1,4308	1,4644*	0,0471	9868																		
				Residuals	61	19,866																					
				Total	68	27,151																					
UVC 24h	Pairwise test locality x seasonality			Within level 'AUT' of factor 'seasonality'						Within level 'SPR' of factor 'seasonality'						Within level 'COR' of factor 'locality'						Within level 'FIG' of factor 'locality'					
				Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms				
				AST, COCH	0,98404	0,4301	8148	AST, COCH	1,367	0,0763	8088	AST, COCH	1,367	0,0763	8088	AST, COCH	1,367	0,0763	8088	AST, COCH	1,367	0,0763	8088				
				AST, CORB	1,6482	0,0002	8143	AST, CORB	2,0464	0,0036	7417	AST, CORB	2,0464	0,0036	7417	AST, CORB	2,0464	0,0036	7417	AST, CORB	2,0464	0,0036	7417				
				AST, FIG	2,1731	0,0003	5029	AST, FIG	2,3743	0,0003	3103	AST, FIG	2,3743	0,0003	3103	AST, FIG	2,3743	0,0003	3103	AST, FIG	2,3743	0,0003	3103				
				COCH, CORB	1,4955	0,0297	8092	COCH, CORB	0,91687	0,5772	9154	COCH, CORB	0,91687	0,5772	9154	COCH, CORB	0,91687	0,5772	9154	COCH, CORB	0,91687	0,5772	9154				
	Pairwise test locality x type of water sample	Within level 'AST' of factor 'locality'						Within level 'CN' of factor 'locality'						Within level 'COR' of factor 'locality'						Within level 'FIG' of factor 'locality'							
		Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms	Groups	t	P(perms)	perms						
		AUT, SPR	1,2363	0,1415	4011	AUT, SPR	0,79726	0,7796	9279	AUT, SPR	1,1903	0,1301	8118	AUT, SPR	1,9531	0,0008	4023	AUT, SPR	1,9531	0,0008	4023						
		Within level '2L.B' of factor 'type of water sample'						Within level '2L.S' of factor 'type of water sample'																			
		Groups	t	P(perms)	perms	Groups	t	P(perms)	perms																		
		AST, CN	1,3054	0,0293	8849	AST, CN	1,201	0,0445	8854																		

Tableau 6. Détails des présentations prévues dans le cadre du webinaire ADN

Intervenant	Horaire	Milieu	Composante	Intervention
Accueil et introduction :				
Pierre Thiriet & Anaïs Rey UMS PatriNat	9h30 – 9h40	-	-	Objectifs du webinaire Présentation du déroulé de la journée
Session 1 : Développer des réseaux de surveillance basés sur l'ADNe				
Agnès Bouchez INRAE	9h40 – 10h00	Aquatique	-	Définitions et grands concepts autour de l'ADNe et des méthodes génétiques appliquées à l'ADNe ; RETEX DNAqua-Net
Arnaud Collin VigiLife	10h00 – 10h15	Aquatique et terrestre	-	Présentation de l'observatoire VigiLife
Raffaele Siano IFREMER	10h15 – 10h30	Marin	Communauté microbienne	ROME : un réseau d'observation des communautés microbiennes des écosystèmes côtiers basé sur l'ADNe
Frédérique Viard CNRS/ISEM	10h30 – 10h45	Marin	Espèces non-indigènes	Surveillance active/passive ; Intérêt du suivi de la diversité intra-spécifique
Collégial	10h45 – 11h00	Discussion et synthèse de la session 1		
Pause-Café	11h00 – 11h15			
Session 2 : Avancées concernant les habitats benthiques et pélagiques				
Frédéric Rimet INRAE	11h15 – 11h30	Eau douce	Microalgues benthiques	Bioindication diatomées et ADN environnemental, où en sommes-nous ? Développements et exemples
Alice Valentini SPYGEN	11h30 – 11h45	Eau douce	Mollusques benthiques	RETEX travaux menés sur les mollusques en eau douce
Albin Meyer INRAE	11h45 – 12h00	Eau douce	Habitats benthiques	Comparaison des approches morphologiques et par metabarcoding dans l'évaluation de la diversité taxonomique et fonctionnelle des macro-invertébrés benthiques en rivière
Benjamin Alric SBR	12h00 – 12h15	Marin	Habitats pélagiques	Le développement d'indicateurs dans le cadre de la DCSMM à travers le projet INDIGENE
Collégial	12h15 – 12h30	Discussion et synthèse de la session 2		

Pause déjeuner	12h30 – 14h00			
Session 3 : Avancées concernant les espèces mobiles				
Isabelle Domaizon INRAE	14h00 – 14h15	Eau douce	Poissons lacustres	L'ADNe pour la détection et suivi d'espèces piscicoles
Verena Trenkel IFREMER	14h15 – 14h30	Marin	Stocks halieutiques	Application des méthodes ADNe pour la surveillance des stocks halieutiques
Naiara Rodriguez Ezpeleta - AZTI	14h30 – 14h45	Marin	Poissons pélagiques	Le potentiel de l'ADN environnemental pour surveiller les poissons des estuaires aux profondeurs de l'océan
Jean-Luc Jung UBO/ISYEB	14h45 – 15h00	Marin	Mammifères marins	Détection de mammifères marins par analyse d'ADNe dans des filtrats d'eau de mer
Collégial	15h00 – 15h15	Discussion et synthèse de la session 3		
Pause-Café	15h15 – 15h30			
Session 4 : Perspectives concernant la surveillance des réseaux trophiques				
Erwan Quéméré INRAE	15h30 – 15h45	Marin	Communautés côtières	L'ADNe trophique pour l'étude de la structure et du fonctionnement des communautés associées aux champs de laminaires – projet LAMINET
Collégial	15h45-17h00	Synthèse des sessions, Discussion générale, Conclusions & Perspectives		