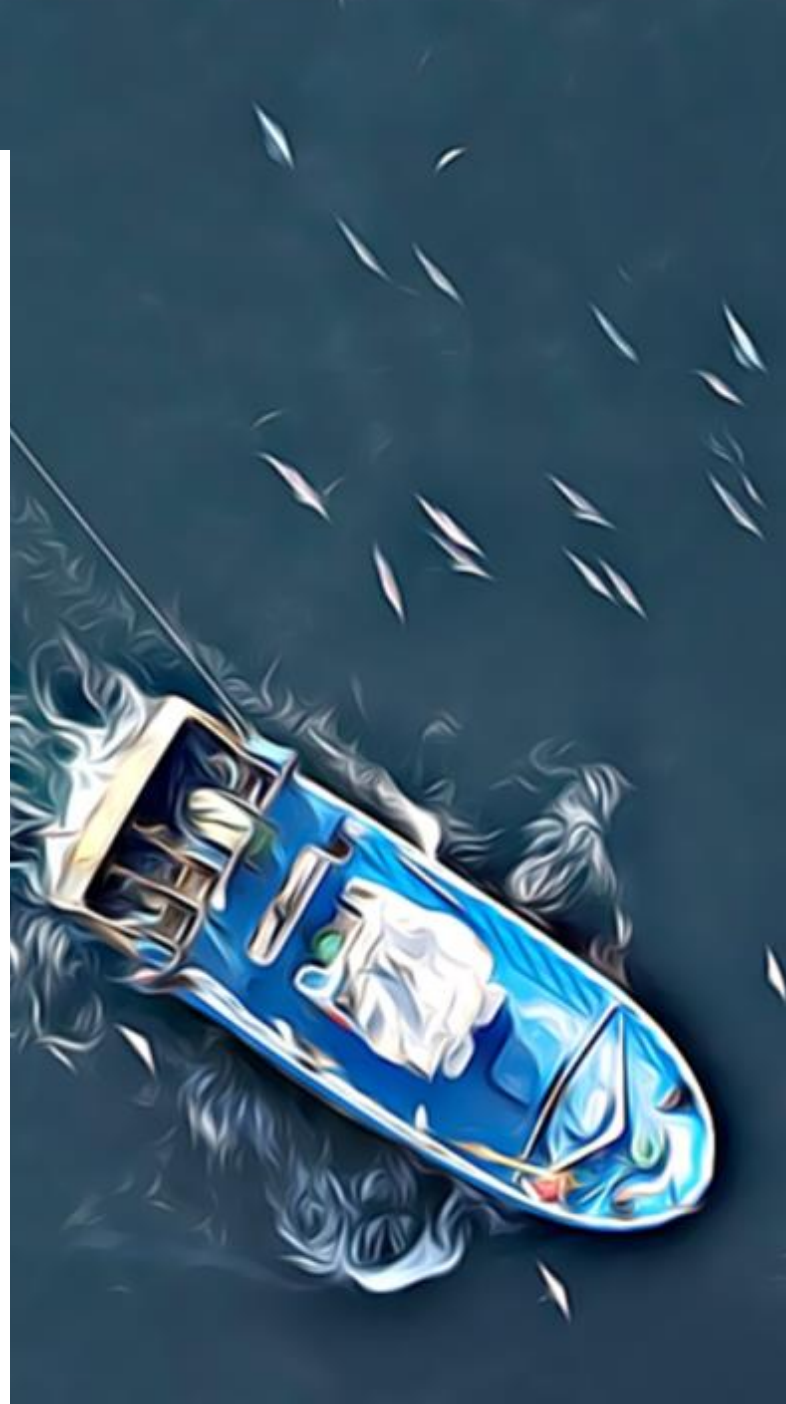




DELphinus MOuvements GESTion

Septembre 2024

Jeux de séquences obtenues à
partir d'ADNe





Durée du projet : 3 ans

Date de lancement : 01/03/2022

Date de fin : 30/06/2025

Coordinateurs de projet : Clara Ulrich, Pierre Petitgas, Jérôme Spitz, Marion PILLET.

Site web : <https://delmoges.recherche.univ-lr.fr>

Livrable

WP concerné : WP1

Responsables du WP : Tiphaine Chouvelon (ULR/CNRS), Amélia Viricel (UBO)

Livrable L.1.1.4

Date de production : 30 septembre 2024

Titre : Jeux de séquences obtenues à partir d'ADNe

Auteurs : Sylvie Lapègue (Ifremer), Tamás Malkócs (UBO), Florence Cornette (Ifremer), Paula Méndez-Fernandez (ULR), Jérôme Spitz (ULR), Aurélie Bonin (Argaly), Céline Reisser (Ifremer), Sophie Arnaud (Ifremer), Cyril Noël (Ifremer), Amelia Viricel-Pante (UBO)

Résumé

Depuis les années 1990, la France connaît régulièrement des épisodes de mortalités importantes de dauphins, qui entraînent des pics d'échouages sur le littoral Atlantique en hiver. Depuis 2016, les échouages de petits cétacés dans le golfe de Gascogne présentant des traces de capture, atteignent des niveaux inédits. Si les données scientifiques actuelles permettent d'évaluer globalement le risque induit par ces captures accidentelles pour la conservation de la population de dauphins communs, elles sont toutefois trop lacunaires pour comprendre les déterminants écosystémiques et halieutiques à l'origine de ces captures. En concertation avec l'Office français de la biodiversité, les professionnels de la pêche et l'Etat, l'Université la Rochelle-CNRS et l'Institut français de recherche pour l'exploitation de la mer (Ifremer) ont construit le projet Delmoges (*Delphinus Mouvements Gestion*). Il vise, dans un premier temps, à combler ces lacunes en allant chercher des nouvelles données sur les habitats des dauphins, sur leurs interactions trophiques dans l'écosystème et leurs interactions techniques avec les engins de pêche. Ensuite, le projet propose d'intégrer les connaissances sur l'ensemble du socio-écosystème pour envisager une diversité de scénarios de diminution des captures accidentelles incluant des solutions technologiques et, enfin, d'en évaluer les conséquences biologiques et socio-économiques.

Ce livrable fournit l'accès aux jeux de données ADNe obtenus dans le cadre du travail du WP1 sur l'évaluation de la structure des populations de dauphins communs et sur l'identification des espèces de poissons présentes à proximité des groupes de dauphins observés lors des campagnes en mer Delgost I et II. Le livrable décrit le contenu de trois jeux de données qui sont stockés dans la base DATAREF de l'Ifremer et seront déposés sur une base de données spécifique aux séquences d'ADN : la base de données ENA (European Nucleotide Archive).

Dissémination

Type de livrable : Données de séquençage

Public : Oui en 2025

Lieux de stockage : DATAREF (Ifremer) et base de données ENA (European Nucleotide Archive ; <https://www.ebi.ac.uk/ena/browser/home>)

Consortium scientifique



La Rochelle Université
23 avenue Albert Einstein
BP 33060
17031 La Rochelle

<https://www.univ-larochelle.fr/>



Centre national de la recherche scientifique (CNRS)
3, rue Michel-Ange
75794 Paris cedex 16

<https://www.cnrs.fr/fr>



Institut Français pour l'Exploitation de la Mer (Ifremer)
1625 route de Sainte-Anne - CS 10070
29280 Plouzané

www.ifremer.fr/



Université
de Bretagne
Occidentale

Université de Bretagne Occidentale (UBO)
3 rue des Archives
CS93837
29238 Brest cedex 3

<https://nouveau.univ-brest.fr/>



Comité National des Pêches Maritimes et des Elevages Marins (CNP MEM)
134 avenue de Malakoff
75116 Paris

<https://www.comite-peches.fr/>

Table des matières

1	Contexte	5
1.1	Contexte environnemental et scientifique.....	5
1.2	Rôle du livrable	5
1.3	Acronymes et abréviations	5
2	Échantillonnage de l'ADN environnemental.....	5
2.1	Périodes et zones d'échantillonnage.....	5
2.2	Protocole d'échantillonnage	6
3	Séquençage	6
3.1	Choix des marqueurs	6
3.2	Protocole de séquençage	7
3.2.1	Extraction d'ADN	7
3.2.2	Amplification, purification et séquençage	7
3.2.3	Contrôles qualité	8
3.2.4	Demultiplexage bioinformatique	8
4	Bilan des séquences obtenues	8
4.1	Nombre de séquences	8
4.2	Stockage et mise à disposition	9
5	Bibliographie	10

1 Contexte

1.1 CONTEXTE ENVIRONNEMENTAL ET SCIENTIFIQUE

Dans le cadre du WP1, la tâche 1.1 a pour objectif de mieux caractériser la (les) population(s) de dauphins impactée(s) par le phénomène des captures dont la structuration côte-large. Une des approches consiste en l'analyse de métabarcodes sur la base de prélèvements d'eau réalisés en parallèle des biopsies de dauphins, afin de définir des stratégies d'échantillonnage et de prélèvement d'eau non invasives pour délimiter et étudier les populations en présence.

1.2 ROLE DU LIVRABLE

Ce livrable vise à mettre à disposition les séquences obtenues à partir d'ADNe afin de pouvoir analyser la diversité génétique obtenue à partir de cette matrice, que ce soit la diversité génétique inter- mais aussi intra-spécifique des dauphins présents au moment de l'échantillonnage, ou la diversité des poissons présents à proximité de groupes de dauphins. Il est composé de trois jeux de données distincts.

1.3 ACRONYMES ET ABREVIATIONS

ADN	Acide DesoxyriboNucléique
ADNe	ADN obtenu à partir d'une source environnementale
ARN	Acide RiboNucléique
Métabarcode	Analyse d'un fragment de gène permettant l'identification des espèces présentes
ENA	European Nucleotide Archive

2 Échantillonnage de l'ADN environnemental

2.1 PERIODES ET ZONES D'ÉCHANTILLONNAGE

Les échantillonnages d'eau de mer ont été réalisés au cours de deux campagnes menées dans le golfe de Gascogne dans le cadre du projet Delmoges : la campagne Delgost 1 (2 - 17 juin 2022) et la campagne Delgost II (31 mai - 17 juin 2023). L'un des objectifs principaux de ces campagnes était l'obtention de biopsies (peau et lard) de dauphins dont l'origine océanique est avérée et de collecter de l'eau de mer à proximité des groupes de dauphins en parallèle pour répondre à l'enjeu de développement d'outils d'analyses génétiques non-invasifs. L'eau a été filtrée pour permettre la collecte d'ADN environnemental. Un deuxième objectif était d'identifier les espèces de poissons présentes à proximité des groupes de dauphins observés grâce aux mêmes prélèvements d'ADN environnemental.

La zone ciblée pour la collecte de biopsies et d'eau de mer était la zone océanique du golfe de Gascogne dans la ZEE (Zone Économique Exclusive) française, principalement au-delà du talus continental. La mise en œuvre et les résultats de ces deux campagnes sont détaillés dans les rapports correspondant aux livrables L1.1.1 (Delgost I) et L1.1.1b (Delgost II). Les prélèvements ont été réalisés à partir de catamarans de la marque Nautitech. Pendant Delgost I, des prélèvements d'ADN environnemental ont été réalisés sur 7 stations différentes, dont 6 en présence et une en absence de détections visuelles de cétacés. Lors de la deuxième campagne (Delgost II), les prélèvements d'eau de mer ont été effectués sur 12 stations différentes. Les prélèvements ont été réalisés en présence de cétacés faisant suite à des détections visuelles et la mise en œuvre d'une collecte de biopsies. Dans la plupart des cas, la station était réalisée une fois les autres travaux terminés avec un retour et une mise en dérive du bateau sur un secteur où les dauphins avaient été observés. La présence immédiate de dauphins à proximité du bateau lors de la filtration n'était alors pas systématique.

2.2 PROTOCOLE D'ÉCHANTILLONNAGE

La filtration d'eau de mer a été réalisée à partir d'un système développé par Argaly (<https://www.argaly.com/capsules-de-filtration-waterra>) comportant une pompe péristaltique associée à un débitmètre. Les capsules de filtration utilisées possèdent un filtre de polyethersulphone d'une surface de 600 centimètres carrés et d'un diamètre de pores de 0,45 microns (Waterra ©). Le dispositif de filtration développé par Argaly permet une filtration en amont du système de pompage, ce qui limite les risques de contamination mais également les déchets plastiques générés en permettant de ne pas avoir à changer tous les tuyaux à chaque station. Trois à quatre répliques par station ont été échantillonnées en filtrant 60 litres par échantillon (environ une heure de filtration). Les répliques ont été réalisés simultanément à chaque station en avant du bateau à l'aide d'une perche de 12 pieds et d'un système de fixation permettant la mise en œuvre simultanée jusqu'à quatre capsules. À la fin de chaque filtration, l'eau de mer à l'intérieur des capsules a été vidée et les capsules ont été remplies avec du tampon de conservation. Les capsules ont été énergiquement agitées à l'aide d'un agitateur mécanique. Le tampon de conservation a été ensuite immédiatement récupéré dans un tube à centrifugation et stocké à bord à température ambiante, puis à 4°C au laboratoire jusqu'à l'analyse des échantillons. Nous avons suivi un protocole strict pour éviter toute contamination lors des prélèvements et de la préparation des échantillons (usage de gants uniques, matériel stérile, désinfection des espaces de travail...).

3 Séquençage

3.1 CHOIX DES MARQUEURS

En ce qui concerne les cétacés, deux marqueurs mitochondriaux ont été ciblés : i) une portion (465 pb) du gène codant pour le cytochrome b (cyt-b) et ii) une portion de la D-loop (527 pb). Des amorces spécifiques aux cétacés sont disponibles (tableau ci-dessous) ce qui permet de limiter ou d'éviter l'amplification d'ADN issu d'organismes d'autres groupes taxonomiques. Le

choix de ces marqueurs s'est également basé sur la variabilité génétique qu'ils fournissent : ces deux marqueurs permettent d'assigner chaque séquence à l'espèce (variabilité interspécifique) et sont suffisamment variables pour mesurer la diversité génétique intraspécifique.

De plus, une portion courte du gène codant pour l'ARN ribosomique 16S mitochondrial (tableau ci-dessous), spécifique des poissons, a été choisie pour essayer de caractériser les espèces de poissons présentes à proximité des dauphins étudiés.

MARQUEUR	AMORCES	REFERENCES
cyt-b	- L14724 : 5'-TGACTTGAARAACCAAYCGTTG-3' - H15149 : 5'-CAGAATGATATTTGTCCTCA-3'	- Palumbi <i>et al.</i> 1991 - Kocher <i>et al.</i> 1989
D-loop	- L15824 : 5'-CCTCACTCCTCCCTAAGACT-3' - H16498 : 5'-CCTGAAGTAAGAACCAGATG-3'	- Rosel <i>et al.</i> 1999 - Rosel <i>et al.</i> 1994
16S	- 16S fish-specific F: 5'-GGTCGCCCAACCRAAG-3' - 16S fish-specific R: 5'- CGAGAAGACCCTWTGGAGCTTIAG-3'	- Shaw <i>et al.</i> 2016 - Shaw <i>et al.</i> 2016

3.2 PROTOCOLE DE SEQUENÇAGE

3.2.1 Extraction d'ADN

L'ADN des filtrations d'eau a été extrait dans un laboratoire dédié à la manipulation d'échantillons aquatiques par la société Argaly. Les tubes 50 mL ont d'abord été centrifugés 15 min à 15 000 g. L'ADN des culots a ensuite été extrait selon le protocole commercial NucleoSpin eDNA water (Macherey Nagel). Des contrôles d'extraction ont été ajoutés afin d'examiner les potentielles contaminations lors de cette étape.

3.2.2 Amplification, purification et séquençage

Un test qPCR a été effectué sur plusieurs extraits d'ADN issus de filtrations et sélectionnés aléatoirement, afin de déterminer la dilution optimale des ADN et le nombre de cycles PCR à effectuer avec chaque couple d'amorces. Chaque ADN a ensuite été amplifié en 8 répliques PCR avec le couple d'amorces du fragment visé. Chaque réplique PCR a été identifiée de manière

unique par une combinaison de deux tags de huit bases accolées en 5' à chaque amorce de PCR. Ces tags servent à assigner les séquences au réplica correspondant pendant l'analyse bioinformatique. Après amplification, tous les échantillons ont été poolés et purifiés avec le kit de purification MinElute (Qiagen GmbH). La construction de la librairie ainsi que le séquençage ont ensuite été effectués par la société FASTERIS (Genève, Suisse). La librairie de séquençage a été préparée suivant le protocole Metafast destiné à limiter les artéfacts de séquençage (<https://www.fasteris.com/dna/?q=content/metafast-protocol-amplicon-metagenomic-analysis>). La librairie de séquençage a ensuite été séquencée dans un run Illumina MiSeq avec des lectures paired-end de 2x300 pb pour les fragments de la D-loop et du Cytochrome B, et 2x150 pb pour le fragment 16S.

3.2.3 Contrôles qualité

Différents contrôles à chaque étape du protocole permettent de détecter les éventuelles contaminations pour une meilleure interprétation des résultats. Dans le cadre de cette étude ont été réalisés : deux contrôles négatifs d'extraction, deux contrôles négatifs PCR, deux contrôles positifs et quarante-huit contrôles bioinformatiques (combinaisons de tags n'existant pas dans l'expérience et permettant de contrôler les niveaux de tag-jumps). Le succès des amplifications et des purifications a été vérifié sur gel d'agarose à 2% (E-Gel Power Snap, Invitrogen®).

3.2.4 Demultiplexage bioinformatique

Les données brutes de séquençage ont été demultiplexées à l'aide du programme obitagpccr des OBITools4 (<https://git.metabarcoding.org/obitools/obitools4/obitools4>), qui attribue chaque lecture au réplica PCR dont elle provient. Plus précisément, ce programme aligne d'abord chaque paire de lectures R1 (sens) et R2 (antisens), puis l'affecte à un réplica PCR sur la base de la combinaison des tags identifiés en 5' des amorces PCR. Les données ayant été séquencées par pcr_ligation, les fichiers R1 et R2 contenaient à la fois des données 'sens' et 'antisens'. Une étape supplémentaire a donc été nécessaire afin d'identifier et répartir les lectures 'sens' (R1) et 'antisens' (R2) dans les fichiers correspondants. Cette étape a été réalisée à l'aide de l'outil cutadapt (v4.1).

4 Bilan des séquences obtenues

4.1 NOMBRE DE SEQUENCES

Le nombre de séquences obtenues pour les différents échantillonnages est indiqué dans le tableau ci-dessous :

ANNEE	2022			2023
MARQUEUR	16S	cyt-b	D-loop	D-loop
NOMBRE DE SEQUENCES	5 691 875	7 281 038	3 404 926	12 621 927
DOI	https://doi.org/10.12770/7c2b753e-ba6d-46a9-ba8d-56e5a075970d	https://doi.org/10.12770/0ebdc7e0-d0bc-486d-8e49-6a33b45ec8bd	https://doi.org/10.12770/9087ff34-cbf1-417a-bf50-7dbb76a2d08d	

4.2 STOCKAGE ET MISE A DISPOSITION

Les séquences ont été stockées sur les disques sécurisés DATAREF de Ifremer et seront rendues publiques dans l'ENA à la fin du projet. Chacun des trois jeux de données, regroupés par marqueur, a obtenu un DOI :

Lapègue Sylvie, Malkocs Tamas, Cornette Florence, Méndez-Fernandez Paula, Spitz Jérôme, Bonin Aurélie, Reisser Céline, Arnaud-Haond Sophie, Noël Cyril, Viricel-Pante Amelia (2025). Cetacean NGS sequences from eDNA samples (cyt-b). IFREMER <https://doi.org/10.12770/0ebdc7e0-d0bc-486d-8e49-6a33b45ec8bd>

Lapègue Sylvie, Malkocs Tamas, Cornette Florence, Méndez-Fernandez Paula, Spitz Jérôme, Bonin Aurélie, Reisser Céline, Arnaud-Haond Sophie, Noël Cyril, Viricel-Pante Amelia (2025). Cetacean NGS sequences from eDNA samples (D-loop). IFREMER <https://doi.org/10.12770/9087ff34-cbf1-417a-bf50-7dbb76a2d08d>

Lapègue Sylvie, Malkocs Tamas, Cornette Florence, Méndez-Fernandez Paula, Spitz Jérôme, Bonin Aurélie, Reisser Céline, Arnaud-Haond Sophie, Noël Cyril, Viricel-Pante Amelia (2024). Fish NGS sequences associated to cetaceans from eDNA samples. IFREMER <https://doi.org/10.12770/7c2b753e-ba6d-46a9-ba8d-56e5a075970d>

5 Bibliographie

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc Natn Acad Sci USA* 86:6196-6200

Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski G (1991) The simples fool's guide to PCR. University of Hawaii, Honolulu

Rosel PE, Dizon AE Heyning JE (1994) Genetic analysis of sympatric morphotypes of common dolphins (genus *Delphinus*). *Mar Biol* 119:159-167

Rosel PE, France SC Wang JY, Kocher TD (1999) Genetic structure of harbour porpoise *Phocoena phocoena* populations in the northwest Atlantic based on mitochondrial and nuclear markers. *Mol Ecol* 8:S41–S54.

Shaw JLA, ClarkeLJ, Wedderburn SD, Barnes TC, Weyrich LS, Cooper A (2016) Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biol Conserv* 197:131-138.