plancton, la nouvelle frontière

publié le 11.01.19 |par flora vincent

l'expédition tara océans, menée entre 2009 et 2013, avait notamment pour objectif de mieux caractériser le plancton. les études génétiques ont permis d'identifier les espèces présentes en différentes régions du globe, mais aussi des montrer que les océans renferment de nombreuses espèces inconnues. la détermination des gènes présents dans l'océan permet également de faire des déductions sur les caractéristiques, notamment écologiques, des espèces ainsi mises en évidence.

1. l'importance du plancton

l'océan représente le plus grand écosystème continu sur terre, et 98 % de sa biomasse est composée d'organismes qui sont invisibles à l'œil nu : les micro-organismes marins, dont nombre d'entre eux vivent comme « plancton ». le terme « plancton » vient du grec *planktos*, ou « errant » ; il désigne les organismes qui vivent dans la colonne d'eau et sont incapables de nager contre le courant. le plancton marin est ainsi composé de bactéries, de protistes (des eucaryotes unicellulaires), de virus, d'archées, mais aussi des stades larvaires d'organismes plus gros, comme les larves de poissons ou de crustacés.

alors que les membres du plancton de l'ordre du millimètre ont été étudiés depuis plus d'un siècle, ce n'est que depuis les années 1970 que l'abondance et la diversité des microbes et virus ont été mises au jour. il a été montré depuis qu'un litre d'eau de mer peut contenir jusqu'à 10^9 bactéries. dans les années 1990, les premières techniques d'évaluation de la diversité bactérienne marine indépendantes de mise en culture – notamment génétique – montrent que les bactéries sont diverses, et que la majorité des groupes marins étaient précédemment inconnus. simultanément, la découverte des virus dans l'océan, atteignant presque 10^{10} particules/l, rajoute une nouvelle couche de complexité dans notre compréhension de ce qui génère, et maintient la diversité microbienne. la diversité des protistes – eucaryotes unicellulaires – est révélée dans les années 2000, à travers l'usage de la génétique, et rend notre connaissance de la diversité planctonique assez récente.

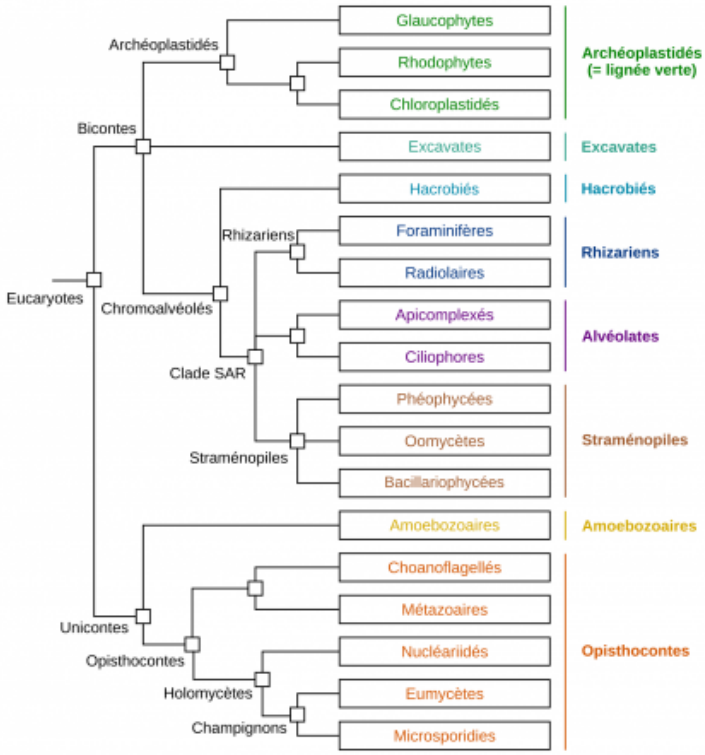

figure 1 - arbre phylogénétique des eucaryotes
les protistes forment un groupe polyphylétique correspondant aux eucaryotes unicellulaires. les taxons représentés sont soit formés uniquement d'unicellulaires, soit d'unicellulaires et de pluricellulaires (par exemple, chloroplastidés et eumycètes). seuls les métazoaires ne contiennent que des êtres vivants pluricellulaires. autrement dit, à l'exception du groupe des métazoaires, on trouve des protistes dans tous les groupes d'eucaryotes notés ici. d'après lecointre et le guyader, 2016.

auteur(s)/autrice(s) : pascal combemorel licence : cc-by

l'exploration de la distribution globale des communautés et diversités de microbes marins devient quantitative grâce au séquençage à haut débit disponible au milieu des années 2000, et ouvre la voie pour des campagnes d'échantillonnage à grande échelle spatiale. en effet, la distribution du plancton dépend fortement de facteurs abiotiques, comme la lumière, les nutriments, la turbulence, la température, la salinité ou le ph, et de facteurs biotiques, comme la présence d'autres organismes tels des prédateurs ou des symbiontes. même si l'abondance locale du plancton varie de façon horizontale, verticale et saisonnière, les organismes planctoniques sont présents partout dans les océans.

l'importance du plancton à l'échelle planétaire est multiple. il est à la base des chaînes alimentaires et représente 50% de la production annuelle de dioxygène sur terre (field, 1998). le métabolisme du plancton joue un rôle majeur dans les grands cycles biogéochimiques du carbone, de l'oxygène, de l'azote, du phosphore et du soufre.

2. l'expédition tara océans

l'expédition tara océans (2009-2013) a eu pour objectif de découvrir la grande variété d'organismes planctoniques (des virus aux larves de poissons) de l'océan de surface (entre 0 et 200 m) et mésopélagique (200 à 1000 m) à l'échelle planétaire. au total, 40 000 échantillons d'eau de mer et de plancton ont été prélevés dans 210 stations réparties dans 20 provinces biogéographiques. de nombreuses questions ont animé cette expédition : quelle est la vraie nature de la diversité planctonique dans nos océans ? quels sont les organismes qui portent les fonctions les plus importantes ? quel est l'effet des paramètres environnementaux et des interactions biotiques sur l'écosystème océanique ?

tara expéditions

le site tara expéditions propose de nombreux outils pédagogiques : kits de données, activités et expériences à réaliser en classe. vous pourrez notamment y retrouver un ensemble de données sur l'analyse génétique des échantillons de plancton.

figure 2 - la goélette tara pendant l'expédition tara océans

auteur(s)/autrice(s) : tara expéditions licence : reproduit avec autorisation

afin de répondre à ces questions, tara a regroupé plus de 100 scientifiques à travers le monde et procédé à un échantillonnage strictement identique pendant plus de trois ans sur la goélette de 36 mètres.

le programme d'échantillonnage standard a été conçu pour étudier une grande variété d'écosystèmes marins : remontées d'eau (*upwellings*), points chauds de biodiversité, zones de bas ph ou pauvres en dioxygène… un total de 210 stations a été défini sur lesquelles une caractérisation environnementale plus précise a été conduite, afin de contextualiser les prélèvements morphologiques et génétiques du plancton. pour chaque station, les prélèvements d'eau ont été filtrés puis soumis à différentes analyses génétiques :

- **séquençage de l'adnr 18s (*metabarcoding*) ;**
- **métatranscriptomique : détermination de l'ensemble des transcriptomes présents, c'est-à-dire de l'ensemble des arn produits par les différents organismes ;**
- **métagénomique : caractérisation de l'ensemble des génomes.**

des échantillons d'eau préservés grâce à du paraformaldéhyde sont utilisés pour la microscopie à haute résolution, tandis que d'autres sont fixés dans de l'éthanol ou du lugol pour permettre d'examiner les populations ultérieurement, pour des projets auxquels nous n'aurions pas encore pensé.

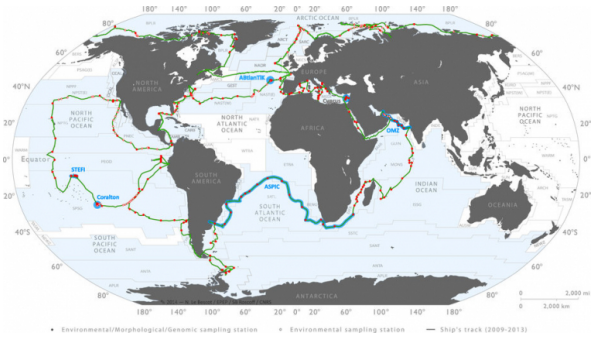

figure 3 - trajet de la goélette tara

auteur(s)/autrice(s) : tara expéditions licence : reproduit avec autorisation

3. tara et la génétique

le séquençage adn à haut débit ouvre la porte de l'exploration de la diversité génétique microbienne d'échantillons environnementaux, à la fois qualitative et quantitative. elle permet de comprendre qui est là, qui fait quoi, et quel est le répertoire de gènes présent dans l'océan.

3.1. qui est là ?

3.1.1. barcoding

la phylogénie des micro-organismes a longtemps été fondée sur des caractères morphologiques et biochimiques. récemment, les marqueurs moléculaires (barcodes) ont été utilisés pour reconstruire l'histoire évolutive des organismes vivants, basé sur l'idée – simplifiée – que plus deux organismes sont distants évolutivement, plus la différence entre leur séquence génomique sera grande. *stricto sensu*, un code-barre adn (barcode) est une courte séquence (typiquement 100 à 400 paires de bases) correspondant à une portion standard du génome (par exemple adn ribosomal 18s), qui peut être utilisée pour identifier les espèces, comme un code-barre utilisé au supermarché qui permet de faire la relation entre le produit (la séquence) et son prix (l'identification de l'espèce). cette séquence est choisie sur la base de critères très précis (valentini *et al.*, 2009) : la variabilité *intra*-espèce doit être faible (cette séquence doit être quasi identique chez tous les organismes de la même espèce) et en même temps la variabilité *inter*-espèces doit être grande (qu'on puisse différencier deux espèces différentes sur la base de leur séquence).

les marqueurs moléculaires adn les plus utilisés pour la reconstruction phylogénétique sont les gènes codants les arnr des sous-unités ribosomales, les adnr. chez la majorité des eucaryotes, l'arnr 18s est présent dans la petite sous-unité ribosomale, tandis que la grande sous-unité contient trois molécules d'arnr (5s, 5,8s, 28s chez les mammifères et 25s chez les plantes). les gènes codant les arnr sont souvent regroupés en batterie (*cluster*), et séparés par des espaceurs internes transcrits (*internal transcribed spacers*, its1 et its2) et un espaceur intergénique (igs, figure 4).

figure 4 - structure d'un adnr eucaryote
organisation typique d'une batterie (*cluster*) d'adnr chez les eucaryotes. les gènes 18s, 5,8s et 28s codent les arn ribosomaux ; its1 et its2 sont les espaceurs internes transcrits (*internal transcribed spacers*); les espaceurs intergéniques (igs : *intergenic spacers*) séparent les nombreuses copies de ces gènes.

auteur(s)/autrice(s) : zagoskin et al., 2014 licence : cc-by

3.1.2. metabarcoding

l'impact des code-barres comme outil moléculaire va bien au-delà d'une phylogénie à plus grande résolution des espèces connues : conservation, découvertes d'espèces, et écologie des communautés en bénéficient (kress *et al.*, 2015). grâce à l'arrivée du séquençage à haut débit, le barcoding est devenu un outil répandu pour l'écologie des communautés eucaryotes (et procaryotes !) à travers une méthode connue sous le nom de metabarcoding, « meta » exprimant l'idée que le barcoding est réalisé non pas sur une seule espèce mais sur un ensemble d'espèces provenant d'échantillons environnementaux.

plus formellement, « le metabarcoding adn fait référence à l'identification automatisée de plusieurs espèces provenant d'un même échantillon contenant les organismes entiers, ou d'un échantillon environnemental contenant de l'adn dégradé (du sol, de l'eau, fèces, etc.) » (taberlet *et al.*, 2012). un bon code-barre pour une étude metabarcoding (comme l'adnr 18s) doit (i) correspondre à une portion de gène quasi identique au sein des individus de la même espèce, mais différent entre espèces (ii) être utilisable pour toutes les espèces considérées dans l'étude, (iii) permettre l'assignation taxonomique et (iv) ce à différents niveaux taxonomiques (valentini *et al.*, 2009). après extraction d'adn de l'échantillon, une pcr avec des amorces dites « universelles » est réalisée, qui permet d'amplifier la portion choisie (par exemple, l'adnr 18s) de *tous* les organismes présents.

comment construire des amorces universelles, c'est-à-dire capables d'amplifier l'adn de n'importe quel organisme présent dans un échantillon ? l'idée est de partir des séquences d'adnr 18s déjà connues et de comparer ces séquences (on parle d'alignement) entre divers organismes. cette démarche a permis d'identifier une région appelée v9, longue de 130 paires de bases dont les extrémités sont très conservées chez tous les organismes connus. les extrémités de cette séquence peuvent donc servir de point d'ancrage pour les amorces lors de la pcr. par construction, il est possible que ces amorces ne détectent pas certains micro-organismes absents des bases de données de références, qui auraient une séquence v9 relativement différente de celle des autres eucaryotes. c'est ainsi que des recherches ont pu montrer que des amorces, que l'on pensait universelles, ne l'étaient pas en réalité.

le produit pcr final est donc un mélange de toutes les séquences v9 des organismes présents. chacune des copies est ensuite séquencée : on obtient l'ensemble de la diversité des v9 de notre échantillon. les séquences présentant un degré de similarité de 97 % ou plus sont regroupées en unités taxonomiques opérationnelles (uto, ou otu en anglais pour *operational taxonomic unit*) - (blaxter *et al.*, 2005). il faut ensuite faire correspondre à chaque uto une espèce. cette assignation taxonomique se fait en comparant les uto de l'échantillon avec des bases de données de référence d'organismes en culture. l'idée est de voir de quelle séquence connue l'uto se rapproche le plus, et dans quelle proportion. le degré de similarité permettra de donner avec un certain degré de certitude un nom d'espèce, ou de genre, ou de famille. seul hic, plus de 40 % des uto déterminées dans le cadre de tara océans sont totalement nouvelles et ne correspondent à aucune espèce connue. on sait simplement qu'il s'agit d'eucaryotes, mais il est quasiment impossible de les placer dans un arbre phylogénétique. excitant !

il faut noter que ces technologies de metabarcoding évoluent aussi vite que baisse le prix du séquençage. d'autres marqueurs plus longs, tels le v4 de 380 paires de bases, auparavant trop chers, sont maintenant accessibles. leur plus grande taille augmente la résolution de la phylogénie. aujourd'hui, et principalement pour le domaine procaryote, d'autres approches s'appuient sur les séquençages de génomes entiers, ou métagénomique. au lieu d'utiliser des amorces spécifiques, l'ensemble des génomes dans l'échantillon environnemental est séquencé. la première étape, après extraction de

l'adn, consiste à fragmenter tous les adn présents dans l'échantillon en morceaux très courts puis à les séquencer ; on parle de *shotgun sequencing*. par la suite, les fragments séquencés sont assemblés bioinformatiquement à partir des régions chevauchantes, afin de reconstruire les génomes d'origines. il est possible d'identifier des portions clefs (les code-barres) *a posteriori*, appelés « $_{mi}$tags » (logares *et al.*, 2014), sur la base de leur ressemblance avec d'autres code-barres de la même famille. on identifie ainsi une portion du génome séquencé qui ressemble à l'adnr 16s (pour les procaryotes) ou 18s (pour les eucaryotes). la démarche est donc différente de celle du metabarcoding où ces régions étaient ciblées *a priori* par la pcr. l'approche $_{mi}$tags permet ainsi d'identifier les code-barres au sein des génomes, mais aussi les gènes présents.

3.2. qui fait quoi et qui peut faire quoi ?

au-delà de faire un inventaire des espèces présentes, l'échantillonnage de tara a aussi eu recours aux techniques de métatranscriptomique. elles permettent de rendre compte des gènes exprimés dans l'échantillon en séquençant les arn messagers (arnm). pour étudier spécifiquement les transcrits eucaryotes, il est possible de sélectionner uniquement les arnm polyadénylés (dotés en 3' d'une queue faite d'une succession d'adénosine, absente chez les procaryotes). ces arnm sont ensuite rétrotranscrits et les adn complémentaires obtenus sont séquencés et comparés à des bases de données de références pour essayer à la fois de donner une annotation fonctionnelle (de quels gènes ces transcrits proviennent-ils ? à quoi peuvent-ils servir ?) et taxonomique (de quels organismes ces transcrits proviennent-ils ?). c'est le « qui fait quoi », détaillé dans (carradec *et al.*, 2018).

en parallèle, le séquençage métagénomique rend compte de l'ensemble des gènes non pas exprimés, mais présents. ce type de séquençage informe sur le potentiel génétique de la communauté, c'est le « qui peut faire quoi », mais ne le fait pas forcément. cette approche au potentiel colossal se confronte à de nombreuses limitations techniques. à l'époque de tara océans, la taille des fragments d'adn pouvant être séquencés à haut débit dépassait à peine les 500 paires de base. si l'on considère par exemple le génome d'une diatomée (phytoplancton abondant), qui fait environ 30 millions de paires de bases, cela signifie qu'il faut séquencer au minimum 60 000 fragments d'adn pour reconstituer son génome complet. en pratique, les analyses ne sont jamais réalisées sur un seul individu, notamment parce qu'il faut disposer de fragments d'adn qui se chevauchent pour pouvoir reconstituer l'ordre des séquences. en réalité, le séquençage est donc réalisé à partir d'un mélange de fragments d'adn provenant d'individus différents. il n'est donc pas possible de se limiter au séquençage de 60 000 fragments au hasard, car sinon certaines parties du génome de notre diatomée auront été séquencées plusieurs fois, mais d'autres aucunes et le génome reconstitué comportera donc des « trous ». pour éviter ce problème, on séquence l'équivalent non pas d'un seul génome (60 000 fragments) mais, par exemple, de vingt génomes (donc 1 200 000 fragments !). à l'époque de tara océans, séquencer les génomes de plusieurs centaines ou milliers d'espèces présentes dans un échantillon d'eau semblait une tâche démesurée. les défis en termes d'assemblage et d'annotation sont colossaux, mais des progrès sont réalisés chaque année.

4. grands principes d'échantillonnage

pour chacune des 210 stations de l'expédition tara océans, l'échantillonnage s'est fait à deux profondeurs. la première est la couche d'eau de surface (sur), définie comme la couche de trois à sept mètres sous la surface. la seconde est la couche de « *deep chlorophyll maximum* » (dcm) qui correspond à la zone d'abondance maximale du plancton photosynthétique, déterminée grâce à la mesure de la concentration en chlorophylle par fluorimétrie.

pour chaque station, l'échantillonnage s'est fait sur plusieurs fractions de taille. le plancton prélevé pendant tara océans recouvre six ordres de grandeur en termes de taille qui correspondent aux virus, virus géants (girus), procaryotes (bactéries et archées), eucaryotes unicellulaires (protistes), et eucaryotes pluricellulaires (comme les copépodes). les cellules des protistes mesurent entre 0,8 et 2000 microns. des filets ont été utilisés afin de créer plusieurs fractions de taille : 0,8-5 microns, 5-20 microns, 20-180 microns, 180-2000 microns.

pour chaque station, une analyse morphologique a été réalisée pour différentes classes d'organismes. d'une part, des systèmes de reconnaissance automatisés (le flowcam et le zooscan) permettent des mesures quantitatives de la biodiversité des organismes allant de 20 microns à quelques centimètres. d'autre part, la microscopie confocale 3d et la microscopie électronique à transmission permettent des analyses ultrastructurales détaillées des petits protistes.

le projet tara océans a utilisé de nombreuses nouvelles technologies et outils d'analyses pour établir le premier effort de collecte de données à l'échelle planétaire qui couple biogéographie, écologie, génétique et morphologie, regroupant une communauté internationale de scientifiques issus de disciplines bien différentes : écologistes marins, microbiologistes, océanographes, statisticiens, biogéochimistes, informaticiens, ingénieurs. cinq articles fondateurs ont été publiés dans le journal *science* en mai 2015, et plus de 75 papiers scientifiques ont été publiés depuis le début de l'aventure.

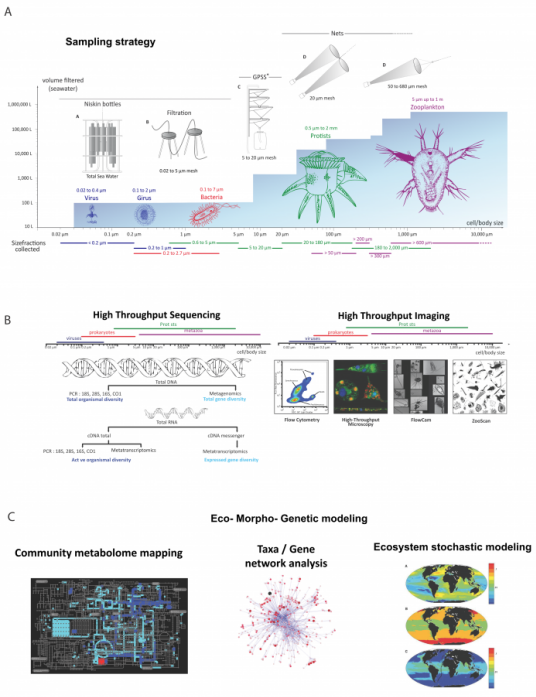

figure 5 - méthodes d'échantillonnage et d'analyse dans tara océans
(a) méthode pour échantillonner des organismes par fractions de taille. le fond bleu représente le volume d'eau filtré afin d'obtenir suffisamment de biomasse pour l'analyse. (b) méthode de traitement des échantillons : les images à droite proviennent de tara océans. (c) séquençage à haut débit et imagerie quantitative apportent des données sur l'évolution, le métabolisme et les interactions entre les organismes, et permettent de reconstruire les métabolomes de la communauté, des réseaux de gène et d'organismes, ainsi que des modèles de distribution spatiale d'espèces.

auteur(s)/autrice(s) : karsenti et al., 2011 licence : cc-by

5. les premiers résultats en génétique de tara océans

les éléments de cette partie sont issus d'un communiqué de presse de tara océans. chaque sous-partie correspond au compte-rendu d'un article de recherche, dont la référence est donnée.

5.1. la découverte de plus de 100 000 types de protistes dans le plancton mondial enrichit notre compréhension des écosystèmes marins

après trois ans de navigation et d'étude des zones baignées de lumière des océans planétaires, les chercheurs ont dévoilé une diversité insoupçonnée chez les organismes unicellulaires eucaryotes aussi appelés protistes. le séquençage de près d'un milliard de codes-barres génétiques a montré que les protistes sont largement plus diversifiés que les bactéries ou les animaux, et que la plupart d'entre eux appartiennent à des groupes peu connus de parasites, de symbiontes, et de prédateurs en tout genre. ces résultats changent radicalement notre vision de la diversité biologique et fonctionnelle du plancton mondial, écosystème-clef pour le fonctionnement de notre biosphère.

dans cette étude, les chercheurs ont déchiffré et analysé près d'un milliard de séquences d'adn ribosomique, issues de 334 sites d'échantillonnage, qui correspondent à des marqueurs de la biodiversité des eucaryotes, des plus petits organismes unicellulaires (<1 micron) aux animaux planctoniques de quelques millimètres.

la grande quantité de code-barres génétiques générés a tout d'abord permis de caractériser la quasi-totalité des espèces eucaryotes du plancton de la zone photique analysée. 150 000 types génétiques de plancton eucaryote ont été mis en évidence, ce qui représente une diversité insoupçonnée par rapport aux 11 000 espèces décrites jusqu'à présent. la grande majorité des types génétiques répertoriés n'a pas de référent proche dans les bases de données génétiques actuelles, démontrant que ces organismes sont pour la plupart non-répertoriés et incultivables. un tiers de la diversité génétique n'a même pu être associé à aucune des grandes lignées eucaryotes reconnues aujourd'hui.

parmi les types génétiques pouvant être classés dans l'arbre de la vie eucaryote, la plupart correspondent à des organismes unicellulaires (protistes), avec une diversité phénoménale de parasites, d'espèces symbiotiques, et de prédateurs en tout genre. les organismes photosynthétiques sont, quant à eux, bien moins diversifiés, plus petits, et représenteraient une biomasse largement plus faible.

référence : *eukaryotic plankton diversity in the sunlit ocean*, c. de vargas, s. audic, n. henry *et al.*, science, 22 mai 2015

5.2. structure et fonction du microbiome océanique

les virus et les microorganismes ([de taille inférieure à] 3 micromètres) dominent numériquement les écosystèmes planctoniques marins avec 10 à 100 milliards de cellules dans chaque millilitre d'eau de mer. leur rôle majeur dans les grands processus biogéochimiques est bien connu et il est donc important de cataloguer leur diversité biologique et fonctionnelle et comprendre comment ils sont affectés par leur environnement. ces questions ont été abordées pour la première fois à l'échelle planétaire grâce à la métagénomique, c'est-à-dire au séquençage massif du matériel

génétique issu de communautés entières de microorganismes. les communautés de microorganismes de tailles variées et vivant à différentes profondeurs ont été prélevés dans l'ensemble des océans planétaires et des relevés de plusieurs paramètres physicochimiques de leur environnement ont été effectués en parallèle.

la quantité d'adn séquencé correspond à environ deux millions de génomes bactériens, ou à peu près deux mille génomes humains. le catalogue de gènes issu de ces travaux comprend 40 millions de gènes de virus, de procaryotes et de pico-eucaryotes marins, nouveaux pour la plupart (> 80 %). il permet de cartographier la diversité fonctionnelle de ces microorganismes océaniques et constitue une ressource fondamentale pour de nombreuses études scientifiques à venir.

une première analyse a par exemple permis de dégager quels sont les paramètres environnementaux qui influencent la formation des communautés microbiennes dans la zone baignée de lumière des océans et d'identifier la température comme un des facteurs les plus importants. ces résultats impliquent que le réchauffement climatique pourrait avoir un fort impact sur ces communautés microbiennes, invisibles à l'œil nu, dont l'activité photosynthétique est à la base des chaînes alimentaires marines.

par ailleurs, une comparaison entre les familles de gènes qui sont au cœur du fonctionnement des communautés microbiennes océaniques et celles présentes dans le système digestif humain a montré que plus de la moitié sont partagées, indiquant des principes communs de la vie microbienne dans ces deux écosystèmes très distincts.

ce nouveau catalogue de gènes microbiens marins planétaires représente un outil indispensable pour mieux comprendre la biodiversité des microbes du plancton et de leurs fonctions, et ce plus particulièrement dans le contexte du changement climatique.

référence : *structure and function of the global ocean microbiome*, s. sunagawa, l.p. coelho, s. chaffron, *et al.*, science, 22 mai 2015.

5.3. cartographie des interactions planctoniques : le rôle majeur des parasites

si la biodiversité des communautés microscopiques qui sont à la base de la vie marine commence à être bien répertoriée, la structure et la dynamique de ces communautés sont encore mal connues. les microorganismes planctoniques interagissent de différentes façons (compétition, collaboration, prédation, symbiose, parasitisme) et forment de gigantesques chaînes alimentaires qui affectent des processus majeurs tels que la séquestration du carbone et la photosynthèse. cependant la plupart de ces interactions étaient jusqu'à présent largement inconnues.

dans cette étude, les chercheurs se sont intéressés à la cartographie des réseaux d'interactions entre espèces de plancton dans la zone photique, et à la façon dont la structure et la composition des communautés planctoniques sont façonnées par des facteurs biotiques (interactions entre espèces) et abiotiques (conditions environnementales et disponibilité en nutriments). ils ont développé de nouveaux modèles permettant de prédire les interactions au sein des communautés planctoniques. grâce à l'analyse des échantillons prélevés lors de l'expédition avec des techniques de microscopie de pointe, ils ont ensuite confirmé que les interactions prédites par les modèles étaient bien réelles dans les communautés observées.

les analyses dites de « réseaux de co-abondances » montrent que les associations au sein du plancton ne sont pas distribuées de façon aléatoire et que les facteurs abiotiques ont des effets plus limités que prévu sur la structure des communautés. ils soulignent le rôle des interactions biotiques (relation de type *top-down*, où les ressources sont régulées par les consommateurs) dans la zone supérieure de l'océan et notamment du parasitisme. les interactions parasitiques sont en effet les plus abondantes dans le réseau généré par ordinateur et ont également été observées de

façon répétée dans les échantillons. la fréquence élevée du parasitisme chez les microorganismes océaniques constitue l'observation la plus importante de cette étude et semble indiquer que les parasites jouent un rôle majeur et largement sous-estimé dans l'écologie du plancton marin.

cette première cartographie globale des interactions planctoniques constitue la matière première qui permettra de savoir comment les symbiontes, pathogènes, prédateurs et parasites interagissent avec leurs organismes cibles, d'élucider le fonctionnement des chaînes alimentaires et la circulation des flux de nutriments et d'énergie, pour à terme mieux comprendre et prédire la dynamique des écosystèmes océaniques.

référence : ***determinants of community structure in the global plankton interactome*** , g. lima-mendez, k. faust, n. henry ***et al.***, science, 22 mai 2015.

6. conclusion

à ce jour (7 décembre 2018), 68 publications émanent directement de l'analyse des données de tara océans. qu'elles s'intéressent à la diversité des virus, des bactéries, du phytoplancton, au lien entre cet écosystème planctonique et l'export de carbone ou encore à l'évolution de ces micro-organismes dans nos océans, l'avalanche de résultats démontre que ces données sont une mine d'or. les études représentent aussi bien une démarche scientifique traditionnelle, avec formulation d'hypothèses et validation expérimentale, que de nouvelles approches dites « ***data driven*** ». l'idée de celles-ci est de laisser les données parler par elles-mêmes pour nous indiquer des pistes de recherches auxquelles nous n'aurions pas forcément pensé, tant la nature est plus créative que le cerveau humain.

s'engager dans une telle aventure, avec une manne de données si colossale et complexe, suppose de repenser les façons de faire de la recherche scientifique au xxi^{e} siècle. d'une part, au sein du consortium, cela impose d'être collaboratif et interdisciplinaire. la force de ce jeu de données réside dans sa capacité à coupler génétique, morphologie, biogéochimie et physique de l'océan à grande échelle spatiale. il impose aux scientifiques d'oublier un instant leur jargon respectif et d'établir un langage commun pour exprimer simplement leurs idées et aboutir à des questionnements qui gomment les frontières disciplinaires. d'autre part, il serait illusoire de penser un instant qu'un groupe de 100 scientifiques suffirait à exploiter ces données dans leur intégralité. ainsi, ces ressources sont rendues accessibles, facilement utilisables et compréhensibles, afin que chaque scientifique où qu'il soit puisse les analyser librement. dans une époque où la donnée est plus que jamais monétisée et la science de plus en plus appliquée, tara océans nage à contre-courant en donnant librement accès, au nom de l'amélioration de nos connaissances, à la diversité du plus grand écosystème de la planète : l'océan.

7. références

- blaxter, m., mann, j., chapman, t., thomas, f., whitton, c., floyd, r., and eyualem, a. (2005) defining operational taxonomic units using dna barcode data. ***philos. trans. r. soc.*** 360: 1935–1943.
- carradec, q., pelletier, e., da silva, c., alberti, a., seeleuthner, y., blanc-mathieu, r., et al. (2018) a global ocean atlas of eukaryotic genes. ***nat. commun.*** 9: 1038.
- field, c.b. (1998) primary production of the biosphere: integrating terrestrial and oceanic components. ***science (80-.).*** 281: 237–240.
- karsenti, e., acinas, s.g., bork, p., bowler, c., de vargas, c., raes, j., et al. (2011) a holistic approach to marine eco-systems biology. ***plos biol.*** 9: e1001177.

- kress, w.j., erickson, d.l., uriarte, m., and garcı, c. (2015) dna barcodes for ecology , evolution , and conservation. 30.
- logares, r., sunagawa, s., salazar, g., cornejo-castillo, f.m., ferrera, i., sarmento, h., et al. (2014) using mitags to explore microbial communities. *env. microbiol* 2659–2671.
- taberlet, p., coissac, e., pompagnon, f., brochmann, c., and willerslev, e. (2012) towards next-generation biodiversity assessment using dna metabarcoding. *mol. ecol.* 21: 2045–2050.
- valentini, a., pompanon, f., and taberlet, p. (2009) dna barcoding for ecologists. *trends ecol. evol.* 24: 110–17.
- zagoskin, m., lazareva, v., grishanin, a., and mukha, d. (2014) phylogenetic information content of copepoda ribosomal dna repeat units: its1 and its2 impact. *biomed res int.*

CRÉDITS

AUTEUR(S)/AUTRICE(S)

flora vincent

spécialiste en microbiologie marine. après un doctorat à l’école normale supérieure où elle a travaillé sur les données de l’expédition tara océans, elle est aujourd’hui en post-doctorat au weizmann institute of science en israël.

MISE EN LIGNE

pascal combemorel

agrégé de svt, il est le responsable éditorial du site planet-vie depuis septembre 2016.

LICENCE DU TEXTE DE L'ARTICLE

creative commons - attribution - pas d'utilisation commerciale