

Modélisation de la contribution des alpha-cyanobactéries aux cycles biogéochimiques dans les océans (CyaNEMO)

Proposition de thèse entre l'équipe ComBi (LS2N, Université de Nantes) et l'équipe ECOMAP (UMR7144 AD2M, Station Biologique de Roscoff)

Description du projet

Résumé

Ce projet de thèse vise à intégrer, dans le modèle climatique Nemo-PISCES, la diversité métabolique des alpha-cyanobactéries, les organismes photosynthétiques les plus abondants du milieu marin, afin d'améliorer la représentation des processus biotiques océaniques. Il repose sur la reconstruction de modèles métaboliques à l'échelle du génome des cyanobactéries (CyanoGEMs), la définition et la validation des niches métaboliques à partir de données expérimentales, ainsi que leur intégration dans des simulations biogéochimiques. En combinant des données génomiques et des approches écogénomiques, ce travail contribuera à proposer une nouvelle modélisation numérique fondée sur la programmation par contraintes pour mieux comprendre le rôle écologique des alpha-cyanobactéries et affiner les prédictions climatiques. **Une demi-bourse de thèse est d'ores et déjà acquise pour ce sujet dans le cadre du projet ANR TaxCy et le complément fait l'objet de demandes en cours auprès de différentes sources de financement.**

Enjeux du projet

Malgré leur intérêt majeur pour les décideurs, les modèles climatiques constituent des défis de taille pour les scientifiques en raison de leur complexité inhérente. Ces modèles intègrent une composante océanique, essentielle à leur fonctionnement, mais qui leur confère des incertitudes considérables. En effet, si des modèles du système Terre tels que Nemo-PISCES¹ démontrent leur efficacité dans la prévision de variables abiotiques comme la température et les niveaux de CO₂, ils sont nettement moins performants dans l'évaluation des facteurs biotiques, en particulier la dynamique des nutriments². Il est largement reconnu que ces imprécisions résultent d'une prise en compte insuffisante de la biodiversité des organismes marins.

Dans ce contexte, les avancées récentes en écogénomique, un champ allant du métabarcodage à la métagénomique, permettent une caractérisation plus fine de la biodiversité, en particulier au sein des communautés microbiennes³. La disponibilité de nombreuses données génomiques de souches et d'organismes incultivés permet désormais d'appréhender la diversité génétique et fonctionnelle au sein de différentes lignées d'organismes. Cela est particulièrement pertinent pour les alpha-cyanobactéries qui constituent un composant clé des écosystèmes marins ayant une influence significative sur les scénarios climatiques⁴. En effet, de par leur abondance et leur ubiquité *in situ*, ces microorganismes photosynthétiques jouent un rôle crucial dans les grands cycles biogéochimiques, notamment celui du carbone, ainsi que dans la dynamique des réseaux trophiques marins. Elles sont à la base des modèles océans globaux où elles sont actuellement modélisées par une équation différentielle ordinaire (ODE) ultra-paramétrée⁵. Cependant, cette représentation ne prend pas en

¹ Aumont, O., Ethé, C., Tagliabue, A., Bopp, L. & Gehlen, M. PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geoscientific Model Development* **8**, 2465–2513 (2015).

² A. Tagliabue, 'Oceans are hugely complex': modelling marine microbes is key to climate forecasts. *Nature* **623**, 250–252 (2023).

³ Sunagawa, S. *et al.* Tara Oceans: towards global ocean ecosystems biology. *Nature Reviews Microbiology* **18**, 428–445 (2020).

⁴ Guidi, L. *et al.* Plankton networks driving carbon export in the oligotrophic ocean. *Nature* **532**, 465–470 (2016).

⁵ Fennel, K. *et al.* Ocean biogeochemical modelling. *Nat Rev Methods Primers* **2**, 76 (2022).

compte leur vaste diversité, caractérisée par l'existence d'écotypes adaptés à des niches écologiques distinctes. Ainsi, on retrouve des écotypes adaptés aux eaux chaudes ou froides, mais aussi aux eaux pauvres en azote, en phosphore et/ou en fer, ou encore aux faibles ou forts éclairements⁷. Cette adaptabilité repose sur une forte diversité génomique⁶ qui leur confère une flexibilité métabolique et la capacité à exploiter diverses ressources en orchestrant des voies biochimiques complexes pour leur croissance et leur survie. Cette adaptabilité est bien définie par les concepts de niche écologique explorés dans des recherches antérieures⁷.

Le projet CyaNEMO vise à proposer une définition formelle des niches écologiques des alpha-cyanobactéries à partir des capacités métaboliques de chaque écotype, dérivées des données génomiques. L'intérêt de cette définition réside dans l'intégration ultérieure des niches au modèle du système Terre Nemo-PISCES pour améliorer la prédiction climatique, mais aussi dans une meilleure appréciation du rôle des alpha-cyanobactéries dans les cycles biogéochimiques.

Pré-requis

Les travaux envisagés reposent sur deux piliers pré-établis. Le premier est que l'équipe ECOMAP, co-dirigée par Laurence Garczarek à la Station Biologique de Roscoff, maintient la **plus grande ressource génomique mondiale d'alpha-cyanobactéries** via la base de données [Cyanorak](#), dans laquelle les gènes issus de nombreux génomes représentatifs des principaux écotypes (97 génomes dans la base publique et plus de 250 dans la base privée) sont rassemblés en groupes d'orthologues afin de faciliter leur curation experte. L'équipe ECOMAP possède par ailleurs une solide expertise en écologie, physiologie et évolution de ces microorganismes. Le deuxième pilier est méthodologique. Il repose sur les travaux d'une thèse soutenue dans le groupe de Damien Eveillard au LS2N à Nantes (thèse CNRS 80 Prime), qui propose une formalisation de la niche métabolique par la **réolution d'un problème linéaire mixte (MILP) via des solveurs à base de contraintes**⁸. Le problème consiste à prendre en compte un réseau métabolique à l'échelle du génome (c-à-d l'ensemble des réactions biochimiques d'un individu qui permet, une fois simulé, d'estimer les flux métaboliques portés par chaque réaction). La niche métabolique ainsi formalisée tient compte de la composition en gènes des organismes modélisés. L'espace de solution de cette niche peut être ensuite utilisé pour simuler les taux de croissance des organismes et ainsi être intégré dans le modèle du système Terre Nemo-PISCES, permettant la simulation de l'organisme, sans ODE, et une estimation des capacités d'adaptation des organismes⁹. Des résultats préliminaires ont été obtenus sur une souche d'alpha-cyanobactérie (*Prochlorococcus MED4*) et demandent à être généralisés à l'ensemble des génomes disponibles dans Cyanorak via la reconstruction de modèles métaboliques correspondants et l'estimation de leurs niches métaboliques⁹.

Planning prévisionnel du doctorat

Première année : La thèse commencera par la reconstruction de modèles métaboliques spécifiques aux souches d'alpha-cyanobactéries (CyanoGEMs) à l'échelle des génomes. Cette étape repose sur l'adaptation de la méthode CARVEME afin d'automatiser cette reconstruction. Le processus implique la résolution d'un problème MILP permettant de simplifier un modèle métabolique universel, afin de conserver uniquement les réactions correspondant aux gènes d'une souche tout en assurant la

⁶ Doré, H. et al. Differential global distribution of marine picocyanobacteria gene clusters reveals distinct niche-related adaptive strategies. *ISME J* doi:[10.1038/s41396-023-01386-0](https://doi.org/10.1038/s41396-023-01386-0) (2023)

⁷ Farrant, G.K. et al. Delineating ecologically significant taxonomic units from global patterns of marine picocyanobacteria. *PNAS* 113(24):E3365-74 doi:[10.1073/pnas.1524865113](https://doi.org/10.1073/pnas.1524865113) (2016)

⁸ Régimbeau, A. et al. Contribution of genome-scale metabolic modelling to niche theory. *Ecology Letters* 25, 1352–1364 (2022).

⁹ Régimbeau, A. et al. Unveiling the link between phytoplankton molecular physiology and biogeochemical cycling via genome-scale modeling. *Sci. Adv.* 11, (2025).

fonctionnalité du modèle (c-à-d une croissance en état quasi-stationnaire respectant les principes de la loi d'action de masse). Cette approche, bien qu'optimisée pour les bactéries hétérotrophes^{10 11} et les eucaryotes autotrophes¹², sera adaptée ici aux cyanobactéries.

Deuxième année : L'objectif principal sera d'identifier les niches métaboliques propres à chaque CyanoGEM. Les niches seront validées en s'appuyant sur des données expérimentales recueillies à Roscoff, notamment les quelques 350 transcriptomes réalisés sous différentes conditions environnementales. Ces données seront comparées aux flux métaboliques estimés pour maintenir la niche, ainsi qu'aux taux de croissance mesurés. Par ailleurs, les hypervolumes des niches métaboliques seront analysés afin d'évaluer leurs chevauchements et d'estimer la distribution théorique des souches. Les résultats seront comparés à deux axes complémentaires : i) l'identification des écotypes⁷ et de leur diversité fonctionnelle⁶, et ii) l'estimation des niches générées par des outils existants, tels que les cartes de diffusion, qui utilisent la présence ou l'absence de gènes pour prédire les niches¹³.

Troisième année : La dernière phase de la thèse consistera à intégrer les CyanoGEMs au modèle biogéochimique Nemo-PISCES⁹. Cette étape vise à étudier la contribution de chaque génotype d'alpha-cyanobactérie aux cycles biogéochimiques et à mieux comprendre son rôle dans la dynamique des écosystèmes marins.

Lieu de la thèse et encadrement

La thèse sera jalonnée de réunions distancielles hebdomadaires avec les collaborateurs de Nantes, ainsi que de mobilités récurrentes entre les laboratoires de Nantes et de Roscoff. En ce qui concerne l'encadrement de la thèse, Laurence Garczarek (50 % de la codirection de la thèse) est directrice de recherches du CNRS et co-dirige l'équipe ECOMAP de la station biologique de Roscoff. Elle est spécialiste des cyanobactéries et des approches génomiques pour l'étude du plancton (7598 citations, h-index 42 sur Google Scholar) et a encadré 11 thèses (dont 1 en co-encadrement en cours) et 12 post-docs (dont 2 en cours). Damien Eveillard (50 % de la codirection de la thèse) est professeur en informatique et responsable de l'équipe COMBI. Il est spécialisé dans la modélisation des systèmes vivants par la programmation par contraintes (4787 citations, h-index 30 sur Google Scholar) et a encadré 13 thèses (dont 2 co-encadrements en cours) et 7 post-doctorants (dont 3 en cours). Tous deux sont membres du consortium Tara Océans. L'environnement de thèse (ordinateur portable, frais de publication et de colloques, etc.) et la moitié du salaire du doctorant seront financés sur le projet ANR TaxCy. L'autre moitié fait l'objet de plusieurs demandes en cours et déterminera l'école doctorale de rattachement : Sorbonne Université ou Université de Nantes. Le doctorant bénéficiera des moyens de calcul et de stockage des plateformes de bioinformatique nantaise (BiRD) et roscovite (ABIMS, IFB).

Profil Recherché

- Formation : Master 2 ou diplôme d'ingénieur en bio-informatique, biologie des systèmes, océanographie biogéochimique, écologie numérique, mathématiques appliquées ou domaine équivalent.
- Compétences techniques appréciées :
 - Forte aptitude pour la programmation (Python indispensable ; R, Bash, Julia un plus).

¹⁰ Machado, D., Andrejev, S., Tramontano, M. & Patil, K. R. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Research* **46**, 7542–7553 (2018).

¹¹ Lieven, C. et al. MEMOTE for standardized genome-scale metabolic model testing. *Nature Biotechnology* **38**, 272–276 (2020).

¹² Burel, M., Régimbeau, A., Chaffron, S., Eveillard, D. & Pelletier, E. *PhotoEukStein: Towards an Omics-Based Definition of Unicellular Eukaryote Phototrophs Functional Traits via Metabolic Modelling*. <http://biorkiv.org/lookup/doi/10.1101/2023.05.22.541783> (2023) doi:10.1101/2023.05.22.541783

¹³ Fahimipour, A. K. & Gross, T. Mapping the bacterial metabolic niche space. *Nature Communications* **11**, 415–418 (2020).

- Expérience ou fort intérêt pour la modélisation (métabolique, écologique ou biogéochimique).
- Connaissances en analyse de données biologiques (génomique, transcriptomique).
- Connaissances de base en écologie microbienne ou biogéochimie marine.
- Qualités personnelles :
 - Curiosité scientifique et motivation pour la recherche interdisciplinaire.
 - Autonomie, rigueur et capacité à mener un projet sur 3 ans.
 - Excellentes capacités de communication écrite et orale (en anglais)
 - Esprit d'équipe et aptitude à travailler dans un environnement collaboratif à distance (Roscoff/Nantes).

Comment Postuler

Les candidatures doivent être envoyées par email aux responsables du projet (voir ci-dessous) et inclure :

- Un CV détaillé.
- Une lettre de motivation expliquant votre intérêt pour le projet et l'adéquation de votre profil.
- Les coordonnées de 2 à 3 référents académiques.
- Les relevés de notes de Master ou cursus ingénieur.

Contacts pour l'encadrement et informations :

- Laurence Garczarek, Équipe ECOMAP, Station Biologique de Roscoff. (Responsable données biologiques/écologie) laurence.garczarek@sb-roscff.fr
- Damien Eveillard, Équipe ComBi, LS2N, Nantes. (Responsable modélisation/optimisation) damien.eveillard@univ-nantes.fr

Date limite de candidature : 15 mars 2026