

Assemblage du génome de la seiche en intégrant les données issues de deux méthodes de séquençage complémentaires (reads courts NovaSeq Illumina vs reads longs Minion ONT)

Type de poste : Stage

Niveau d'étude minimal : Master2

Dates : janvier-juin

Durée du poste : 6 mois

Localisation :

Muséum national d'Histoire naturelle

UMR BOREA "Biologie des Organismes et Ecosystèmes Aquatiques"

43 rue Cuvier

750005 PARIS

France

Contacts : Yann Bassaglia (UMR 7208 BOREA) et Evelyne Duvernois-Berthet (UMR 7221 PhyMA)

Emails des contacts : yann.bassaglia@mnhn.fr / evelyne.duvernois@mnhn.fr

Candidature à envoyer avant le 30 décembre 2019.

Description :

Les Céphalopodes présentent des capacités cognitives dont les ressorts moléculaires et génétique commencent juste à être abordés. Cinq génomes de céphalopodes sont actuellement disponibles, dont trois d'octopodes (pieuvres), un de sépiole et un de calmar avec des différences dans la qualité des données à disposition [1-3]. Les génomes sont de grande taille, entre 2,8 pour *O. bimaculoides* [1] et 4,5 Gb estimé pour la seiche *S. officinalis* [4], et montrent de très nombreuses répétitions. Ainsi, l'assemblage de tel organisme, en plus de représenter un challenge bioinformatique, est une occasion de mieux comprendre l'évolution de ces métazoaires.

Jusqu'à présent, malgré leur diversité et leur intérêt écologique et économique, aucun génome de seiches n'est disponible pour la communauté. Une première analyse de données issues de séquençages illumina (HiSeq) lancée en collaboration avec notre équipe a produit une ébauche du génome de *Sepia officinalis* à faible couverture. De plus, notre équipe a produit plusieurs bibliothèques utilisant la technologie Oxford Nanopore (MinION) disponible au MNHN, et permettant la production de reads longs. Un second séquençage Illumina (NovaSeq) a aussi été récemment financé, et sera lancé début 2020.

Le stage proposé consiste tout d'abord en l'analyse et le nettoyage des reads MinION et Illumina obtenus. Par la suite, le/la stagiaire aura pour principale mission de produire un nouveau génome de *Sepia officinalis* par l'assemblage hybride utilisant les données issues des deux technologies de séquençage et de prédire/annoter les gènes. Notre objectif est d'obtenir un génome complet propre, fiable et robuste d'une part et également d'avoir une estimation des erreurs de séquençage lors de la lecture de longs fragments MinION.

Ce stage sera réalisé en lien étroit avec les deux co-encadrants (coordonnées ci-dessus) ainsi que d'autres participant.e.s du projet.

Références :

- [1] Albertin, C. B., Simakov, O., Mitros, T., Wang, Z. Y., Pungor, J. R., Edsinger-Gonzales, E., ... & Rokhsar, D. S. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature*, 524(7564), 220.
- [2] Kim, B. M., Kang, S., Ahn, D. H., Jung, S. H., Rhee, H., Yoo, J. S., ... & Cho, S. J. (2018). The genome of common long-arm octopus *Octopus minor*. *GigaScience*, 7(11), giy119.
- [3] Belcaid, M., Casaburi, G., McAnulty, S. J., Schmidbaur, H., Suria, A. M., Moriano-Gutierrez, S., ... & Collins, A. J. (2019). Symbiotic organs shaped by distinct modes of genome evolution in cephalopods. *Proceedings of the National Academy of Sciences*, 116(8), 3030-3035.

[4] Albertin, C. B., Bonnaud, L., Brown, C. T., Crookes-Goodson, W. J., da Fonseca, R. R., Di Cristo, C., ... & Koenig, K. M. (2012). Cephalopod genomics: a plan of strategies and organization. *Standards in genomic sciences*, 7(1), 175.