

DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **17 mai 2024**

Nom de famille et prénom de l'auteur : **Madame NGUYEN TRUNG Alexia**

Titre de la thèse : « *Odysée d'*Helicobacter pylori* : Itinéraires, Rencontres, Contacts* »



Résumé

Les génomes présentent une diversité remarquable, variant en taille, composition en base, expression, nombre de gènes *etc.* Comprendre l'origine de ces changements captive. Trois forces évolutives sont à l'œuvre: la mutation, *i.e.* la création de nouveaux allèles; la sélection, *i.e.* l'impact des allèles sur la fitness; la dérive génétique aléatoire, *i.e.* l'échantillonnage aléatoire des allèles au fil des générations. Si, au cours de l'évolution, des mutations bénéfiques (*i.e.* qui ont contribué à l'adaptation des espèces à leur environnement) se sont fixées dans les génomes, la sélection naturelle ne peut à elle seule expliquer toute la diversité observée au niveau moléculaire.

La théorie neutraliste de l'évolution nous apprend que les mutations ayant des effets négligeables sur la fitness d'un organisme jouent également un rôle important dans l'évolution des génomes. Plus précisément, la théorie quasi-neutraliste suggère que l'efficacité à purger (ou fixer) des mutations légèrement délétères (ou avantageuses), dépend de la capacité de la sélection à dominer les effets aléatoires de la dérive. À pression sélective égale, la capacité d'un génome à atteindre l'optimal est ainsi limitée par l'intensité de la dérive à laquelle il est soumis. Cette hypothèse, connue sous le nom de "barrière de la dérive", prédit que les espèces à forte dérive, et donc à faible taille efficace (N_e), ont un génome moins bien optimisé que celles qui présentent une N_e plus grande.

Au cours de ma thèse, j'ai étudié l'impact de la dérive génétique sur l'architecture des génomes chez les animaux. Pour ce faire, j'ai collecté de nombreux génomes et transcriptomes, provenant de 1,506 espèces et 15,935 échantillons RNA-seq, afin d'analyser la diversité transcriptomique et la composition des séquences codantes. Ces analyses sont présentées dans une base de données bio-informatique, incluant des estimateurs de N_e .

Ainsi, j'ai étudié l'influence de la dérive sur la diversité transcriptomique, à travers la quantification des variants d'épissage. Nos résultats ont démontré que la dérive limite la capacité de la sélection à optimiser la machinerie

d'épissage dans les génomes. Ce qui provoque la production de nombreux transcrits erronés chez les espèces à faible N_e , comme l'Homme.

Je me suis intéressé à un autre trait de la composition en base des génomes, et en particulier l'usage des codons synonymes chez les métazoaires. Nos résultats ont révélé que chez les espèces où la sélection traductionnelle (*i.e.* favorisant l'utilisation de codons optimisant le processus de traduction) est détectée, le coefficient de sélection à l'échelle populationnelle est faible. Cela suggère que la sélection traductionnelle ne peut être efficace que chez les espèces à grande N_e , justifiant sa rareté chez les métazoaires. Néanmoins, certaines espèces ayant une grande N_e ne montrent pas de signaux de sélection traductionnelle, suggérant que l'avantage sélectif à optimiser le processus de traduction varie.

Finalement, cette thèse illustre l'impact de la dérive sur l'architecture des génomes et fournit un cadre conceptuel intéressant, ainsi qu'une collection de données réutilisables, pour examiner ce qui est ou n'est pas soumis à la sélection dans nos génomes.