



Université Claude Bernard



DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **14 décembre 2021**

Nom de famille et prénom de l'auteur : **Monsieur TRICOU Théo**

Titre de la thèse : « *Détecter et exploiter les flux de gènes dans une biodiversité majoritairement inconnue* »

Résumé



En première approximation, toutes les espèces sont éteintes. Et celles qui ne le sont pas nous sont pour la plupart inconnues.

Durant les 4 milliards d'années d'évolution ayant engendré cette immense biodiversité, des organismes se sont transmis du matériel génétique, soit verticalement, de manière généalogique, soit horizontalement, par des transferts entre espèces distinctes. Cette deuxième composante est maintenant reconnue comme une force évolutive majeure ayant remodelé les génomes tout au long de l'évolution.

L'observation conjointe d'une biodiversité largement inconnue et de l'existence de flux génomiques horizontaux implique que certains gènes, présents dans les espèces observées aujourd'hui, sont vraisemblablement apparus et ont évolué pendant un certain temps dans des organismes aujourd'hui éteints ou encore inconnus. Il est donc légitime de se demander si l'étude des flux de gènes, menée habituellement sans prendre en compte ces lignées fantômes, ne peut pas conduire à des conclusions erronées.

Au cours de ma thèse, j'ai développé et utilisé des approches *in silico* pour explorer cette question. J'ai tout d'abord collaboré au développement d'un outil bioinformatique, Zombi, permettant de simuler les composantes verticales et horizontales de l'évolution des génomes le long des branches d'un arbre d'espèces tout en considérant des lignées fantômes. J'ai ensuite utilisé cet outil pour explorer trois aspects de l'influence des lignées fantômes sur certains résultats en évolution moléculaire. Premièrement j'ai examiné les erreurs commises, si on néglige les espèces fantômes, lors de l'interprétation de la statistique-D pour détecter les introgressions. J'ai montré que les interprétations erronées, qui étaient considérées comme des exceptions, étaient en réalité probablement la règle. Deuxièmement, j'ai ré-analysé trois études qui ont utilisé les longueurs de branches pour l'analyse des flux génomiques horizontaux et j'ai montré que leurs conclusions s'inversaient lorsque l'on considérait que des lignées fantômes pouvaient être impliquées. Enfin, j'ai fourni une preuve du concept que les flux de gènes provenant de lignées fantômes pouvaient être utilisés pour révéler l'existence et la nature de certaines lignées fantômes. La détection des flux de gènes pourrait ainsi remplacer les fossiles lorsqu'ils sont indisponibles, comme chez les micro-organismes.

L'apport de cette thèse est à la fois de montrer l'importance de prendre en compte la biodiversité fantôme dans l'étude des flux de gènes, de fournir des outils pour cette prise en compte, et donc d'offrir un nouveau cadre de travail pour la recherche future sur les flux de gènes dans tous les domaines du vivant.

