

# Sujet de thèse Unité Evo-Eco-Paléo – campagne 2021

**Université :** Université de Lille, Faculté des Sciences et Technologies

**Ecole doctorale:** ED104 Sciences de la matière, du rayonnement et de l'environnement (SMRE)

**Domaine scientifique, Spécialité :** Biologie de l'environnement, des organismes, des populations, écologie

**Titre de la thèse :** Inférer conjointement l'histoire démographique des espèces et les cibles génomiques de la sélection naturelle en développant une approche basée sur les réseaux de neurones convolutifs

**Direction de thèse :** Xavier VEKEMANS et Camille ROUX

**Email de contact:** camille.roux@univ-lille.fr

**Laboratoire(s) de Rattachement :** Unité Evolution, Ecologie et Paléontologie, UMR CNRS 8198

**Programme(s) de Rattachement :** ANR AI\_PhD@Lille

**Financement acquis ou demandé :** Allocation isite ULille (demandé)

## SUJET DE THESE

Comprendre la dynamique d'émergence de nouvelles espèces requiert une bonne description de leur histoire démographique et de la sélection naturelle agissant dans les génomes en réponse aux changements environnementaux. À ce jour, ces deux aspects sont étudiés séparément malgré leurs interactions : si les méthodes statistiques pour inférer l'histoire démographique sont biaisées par les effets de la sélection naturelle, les méthodes pour détecter la sélection sont quant à elles impactées par la démographie. Nous proposons i) de développer une nouvelle approche basée sur les réseaux de neurones convolutifs (CNNs) pour inférer conjointement la démographie et les effets de la sélection naturelle à partir de génomes séquencés, ii) de l'appliquer à des données génomiques disponibles et en cours d'obtention chez diverses familles de plantes et iii) d'intégrer cette nouvelle approche dans notre plateforme web de science collaborative en génomique des populations développée pour les Random Forests.

**Compétences recherchées :** Formation en biologie évolutive, bio-analyse de larges jeux de données de séquences.

**Mots clefs :** Biodiversité ; Génomique ; Réseaux de neurones convolutifs ; Science collaborative

**Publication du laboratoire sur le sujet :** C. Fraisse, I. Popovic, C. Mazoyer, B. Spataro, S. Delmotte, J. Romiguier, E. Loire, A. Simon, N. Galtier, L. Duret, N. Bierne, X. Vekemans, C. Roux, 2021. DILS : Demographic Inferences with Linked Selection by using ABC. *Molecular Ecology Resources*, in press. (<https://doi.org/10.1111/1755-0998.13323>)

## PHD SUBJECT

**Joint inferences of the demographic history of species and genomic targets of natural selection by developing a new approach based on convolutional neural networks**

Understanding the evolution of new species requires a better reconstruction of the past demography and detection of the impacts let by natural selection in their genomes, including environmental changes. To date, these two main goals have been investigated separately despite their interactions: demographic inference is affected by selection throughout genomes, while genome scans to detect selection are affected by demography. Here, we propose (i) to develop a novel approach based on convolutional neural networks (CNNs) to jointly infer historical demography and the effects of natural selection from sequenced genomes, (ii) to apply it to available and in-progress genomic data in various plant families, and (iii) to integrate this new approach into our web-based platform for collaborative science in population genomics already using Random Forests.

**Requested skills:** Training in evolutionary biology, modelisation, analysis of sequenced dataset.

**Key words :** Biodiversity; Genomics; Convolutional neural networks ; Collaborative science

**Publication of the team on related subject :** C. Fraisse, I. Popovic, C. Mazoyer, B. Spataro, S. Delmotte, J. Romiguier, E. Loire, A. Simon, N. Galtier, L. Duret, N. Bierne, X. Vekemans, C. Roux, 2021. DILS : Demographic Inferences with Linked Selection by using ABC. *Molecular Ecology Resources*, in press. (<https://doi.org/10.1111/1755-0998.13323>)