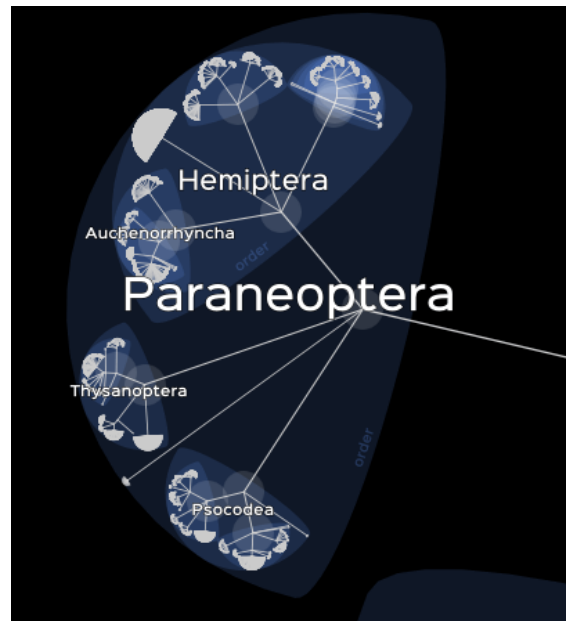


Création d'un package R et d'une App Shiny pour la visualisation de données biologiques avec Lifemap

Encadrement Damien de Vienne - damien.de-vienne@cnrs.fr
Aurélié Siberchicot - aurelie.siberchicot@univ-lyon1.fr

CONTEXTE

Lifemap (de Vienne 2016) est un outil web¹ permettant d'explorer de façon interactive et intuitive l'intégralité de l'arbre du vivant, c'est-à-dire les liens de parenté entre toutes les espèces connues à ce jour (environ 2 millions). Cet outil, développé au Laboratoire de Biométrie et Biologie Évolutive (LBBE), est utilisé dans le monde entier², majoritairement par le grand public qui y trouve un point d'accès vers des photographies et des articles issus de Wikipedia, et par les enseignants (du secondaire au supérieur) qui s'en servent comme support pédagogique pour aborder les notions d'évolution et de classification du vivant.



Les Paraneoptera (un ordre d'insectes) vus depuis l'outil Lifemap

Depuis 2022, le projet Lifemap se tourne vers le monde de la recherche. Plusieurs projets émergent qui visent à utiliser Lifemap non plus comme un simple explorateur du monde vivant, mais soit comme une véritable **plateforme** permettant d'accéder aux données génomiques massives disponibles dans de nombreuses bases de données (comme celle du NCBI), soit comme un **fond de carte** sur lequel représenter des informations de toutes natures relatives à des espèces d'intérêt. C'est sur ce second point que porte le présent projet.

PROJET

À la manière des SIG (Systèmes d'Information Géographique) où des données géoréférencées peuvent être représentées et analysées sur un fond de carte géographique, ce stage propose de **développer un package R et une appli Shiny** (Chang et al. 2022) permettant de **visualiser et d'analyser des données taxonomiquement référencées sur un fond de carte Lifemap**. La liste des données possiblement visualisables est infinie : données de conservation (statut IUCN), données génomiques (taille des génomes, taux de GC, nombre de chromosomes), données métagénomiques (présence/absence d'espèces dans un échantillon), etc. Quelles qu'elles soient, l'utilisateur apportant ces données pourra, grâce à ce nouveau package et cette application Shiny, les visualiser directement (et de façon interactive) sur le fond de carte **Lifemap**. Le package R (*LifemapR*) et

¹ <https://lifemap.univ-lyon1.fr>

² plus de 100000 téléchargement sur le PlayStore (au 1er octobre 2022)

l'application Shiny associée viendront ainsi étendre la collection d'outils déjà existants autour de Lifemap (le widget *lifemap-query*³ pour intégrer une vue personnalisée de Lifemap dans un site web et la page web *lifemap-ncbi*⁴ pour accéder aux informations du NCBI depuis Lifemap).

L'étudiant-e développera des fonctions en langage R pour (i) lire et stocker les données fournies par l'utilisateur, (ii) requêter des informations complémentaires nécessaires relatives aux espèces concernées (coordonnées sur la carte Lifemap, identifiants des noeuds ascendants, descendants, noms latins des espèces, etc.), (iii) inférer éventuellement des valeurs pour les noeuds internes de l'arbre, (iv) représenter les données sur le fond de carte Lifemap (couleurs, formes, intensité des marqueurs utilisés, couleur des branches, etc.), (v) exporter la carte générée et assurer son bon fonctionnement malgré les mises à jour des différentes bases de données associées. Il/elle publiera un package R fonctionnel et développera ensuite l'application Shiny faisant appel à ces fonctions et permettant à des utilisateurs de charger leurs propres données sur une interface web et de les visualiser directement sur Lifemap. Un article scientifique présentant ce nouvel outil pourra être produit en fin de stage.

PRÉREQUIS

L'étudiant-e devra avoir un intérêt pour la biologie et en particulier la biologie évolutive, connaître le langage R, avoir des bases dans les langages du web (html, javascript, css) et maîtriser git. Une connaissance de la librairie leaflet (<https://leafletjs.com/>) et de Shiny serait un plus.

RÉFÉRENCES

Chang W., Cheng J., Allaire J.J., Sievert C., Schloerke B., Xie Y., Allen J., McPherson J., Dipert A., Borges B. 2022. shiny: Web Application Framework for R. .
de Vienne D.M. 2016. Lifemap: Exploring the Entire Tree of Life. PLOS Biol. 14:e2001624.

³ <https://github.com/damiendevidienne/lifemap-query>

⁴ <https://lifemap-ncbi.univ-lyon1.fr>