

AffiliationExplorer: une application shiny pour résoudre les conflits de taxonomie

Mahendra Mariadassou¹

Cédric Midoux²

Olivier Rué³

Résumé (max 300 mots)

Les études de métabarcoding en écologie microbienne génèrent des données d'abondance de taxons, généralement représentés par la séquence d'un ASV (*Amplicon Sequence Variant*), où certains taxons peuvent être associés à plusieurs affiliations taxonomiques, parfois conflictuelles. Ces ambiguïtés (ou multi-affiliations) [1], fréquentes lors de l'analyse de séquences courtes et/ou de bases de référence incomplètes, compliquent l'interprétation écologique et nuisent à la reproductibilité des résultats. Selon l'origine des échantillons (alimentaire, fécale, etc.), une curation manuelle permet parfois de lever ces ambiguïtés. De la même manière, certaines incohérences sont facilement corrigeables de manière automatique. Le package `{affiliationExplorer}` propose une interface utilisateur dédiée à l'exploration et à la résolution de ces conflits, tout en s'intégrant aux workflows d'analyse de données de type BIOM (*Biological Observation Matrix*).

Développé en Shiny avec le framework `{golem}`, ce package offre une interface graphique intuitive permettant aux utilisateurs de ① charger un fichier BIOM, un fichier de multi-affiliations et, optionnellement un ou plusieurs dictionnaires séquence—taxonomie expertisés (au format FASTA), ② visualiser et trier les taxons multi-affiliés par abondance, ③ résoudre automatiquement les conflits grâce aux dictionnaires fournis pour les ASV qui y sont référencés, ④ procéder à une curation manuelle en sélectionnant l'affiliation la plus pertinente parmi celles proposées ou en la spécifiant manuellement, et ⑤ exporter un fichier BIOM corrigé ainsi qu'un dictionnaire séquence—taxonomie des curation effectuées, réutilisable lors de sessions de curation ultérieures. Le fichier BIOM produit suit les standards de `biom-format.org` et est directement compatible avec les workflows usuels de métabarcoding, tandis que les fichiers FASTA exportés peuvent être réutilisés tels quels dans l'interface.

Mots-clefs (3 à 5) : Biostatistique - Package - Shiny - Taxonomie - Métabarcoding

Bibliographie

- [1] M. Bernard, O. Rué, M. Mariadassou, et G. Pascal, « FROGS: a powerful tool to analyse the diversity of fungi with special management of internal transcribed spacers », *Briefings in Bioinformatics*, vol. 22, n° 6, août 2021, doi: 10.1093/bib/bbab318.

¹INRAE, mahendra.mariadassou@inrae.fr

²INRAE, cedric.midoux@inrae.fr

³INRAE, olivier.rue@inrae.fr