

# RÉSUMÉ ET ABTRACT

---

**Résumé:** Les données issues de l'ADN environnemental sont de plus en plus utilisées pour étudier la biodiversité et tester des hypothèses écologiques depuis longtemps étudiées. Cependant, la fiabilité des résultats obtenus par l'analyse d'ADN environnemental est remise en question car elle dépend de choix arbitraires faits tout au long du processus de pré-traitement bioinformatique. Nous avons donc cherché à quantifier l'impact de ces choix sur différentes questions écologiques et leurs mesures associées. Nous avons utilisé de l'ADN de sol obtenu à partir d'une méta-communauté végétale échantillonnée le long d'un gradient d'altitude dans les Alpes françaises. Nous avons testé l'effet de sept étapes de pré-traitement bioinformatique sur deux types de questions écologiques : 1) une analyse de partitionnement de diversité (diversités gamma, alpha et beta) et 2) une analyse spatiale (relation diversité-environnement) utilisant les nombres de collines. Nous avons démontré que les analyses écologiques étaient plus ou moins sensibles aux choix méthodologiques et que l'utilisation de la diversité de Shannon conduisait à des résultats plus fiables. De plus, nous avons démontré que le regroupement des unités taxonomiques opérationnelles (OTUs), l'élimination des erreurs de PCR et l'élimination des contaminations inter-échantillons ont une influence majeure sur les résultats des analyses écologiques, et doivent donc être pris en compte lors de la conception du processus de pré-traitement bioinformatique.

**Abstract:** Environmental DNA-based data is being increasingly used for measuring biodiversity patterns and testing long-standing hypotheses in ecology. However reliability of diversity estimates from environmental DNA data is questioned because it depends on arbitrary decisions made during the bioinformatic curation process, which are known to affect the results. Here, we aimed to test how sensitive are different ecological questions and related diversity measures to those decisions. We used soil DNA data obtained from a plant meta-community sampled across an elevation gradient in the French Alps. We tested the effect of seven bioinformatic curation steps on two types of ecological studies: 1) a diversity partitioning analysis (gamma, mean alpha and beta diversities) and 2) a spatial analysis (diversity-environment relationship), based on Hill numbers. We showed that ecological analyses had varying degrees of sensitivity to methodological choices and that the use of Shannon diversity led to more reliable results. Also, we demonstrated that operational taxonomic unit (OUT) clustering, removal of PCR errors and removal of cross-sample contaminations had major influence on ecological analyses, and must be taken in consideration when designing bioinformatic pipelines.