Overcoming limitations to environmental DNA studies: A coastal temperate reference sequence database for multiple chloroplast gene regions generated in a single assay.

Nicole Foster¹, Kor-jent Dijk², Ed Biffin¹, Jennifer Young³, Vicki Thomson², Bronwyn Gillanders², Alice Jones¹, and Michelle Waycott²

¹The University of Adelaide

²University of Adelaide

³Flinders University

September 24, 2021

Abstract

A proliferation in environmental DNA (eDNA) research has increased the reliance on reference sequence databases to assign unknown DNA sequences to known taxa. Without comprehensive reference databases, DNA extracted from environmental samples cannot be correctly assigned to taxa, limiting the use of this genetic information to identify organisms in unknown sample mixtures. For animals, standard metabarcoding practices involve amplification of the mitochondrial Cytochrome-c oxidase subunit 1 (CO1) region, which is a universally amplifiable region across majority of animal taxa. This region, however, does not work well as a DNA barcode for plants and fungi, and there is no similar universal single barcode locus that has the same species resolution. Therefore, generating reference sequences has been more difficult and several loci have been suggested to be used in parallel to get to species identification. For this reason, we developed a multi-gene targeted capture approach to generate reference DNA sequences for plant taxa across 20 target chloroplast gene regions in a single assay. We successfully compiled a reference database for 93 temperate coastal plants including seagrasses, mangroves, and saltmarshes/samphire's. We demonstrate the importance of a comprehensive reference database to prevent species going undetected in eDNA studies. We also investigate how using multiple chloroplast gene regions impacts the ability to discriminate between taxa.

Hosted file

Foster et al. manuscript.docx available at <https://authorea.com/users/435785/articles/538567-overcoming-limitations-to-environmental-dna-studies-a-coastal-temperate-reference-sequence-database-for-multiple-chloroplast-gene-regions-generated-in-a-single-assay>