ID: 737

Towards accurate species detection and identification: Calibrating metabarcoding methods based on multiplexing multiple markers

Guang K. Zhang[1], Frédéric J.J. Chain[1], Cathryn Abbott[2], Melania E. Cristescu[1]
[1]McGill University, Montréal, Québec, Canada [2]Fisheries and Oceans Canada, Pacific region

Abstract

Species-level identification in metabarcoding depends heavily on the choice of marker and primer pair, often with a trade-off between successful species amplification and taxonomic resolution. Thus, we present a versatile metabarcoding protocol for biomonitoring that involves the use of two barcode markers and multiple primer pairs per barcode in a single high-throughput run via sample multiplexing. With the use of all three COI primer pairs, we detected 61.5-82.8% of species across mock communities, while the use of the 18S primer pair resulted in the detection of 72.4-75.0% of species. The species detection level was significantly improved to 88.5-93.1% when both markers were used together. Furthermore, multiplexing did not have a negative impact on the proportion of reads assigned to each species. Overall, our metabarcoding approach utilizing two barcode markers and multiple primer pairs per barcode improved species detection rates over a single marker/primer pair by 13.5% to 34.7%, making it an attractive, cost effective method for biomonitoring natural zooplankton communities.

Method

- Primer testing: 13 COI primer pairs (COI-5P region) and one 18S primer pair (V4 region) tested on 104 species
- 3 COI primer pairs and one 18S primer pair for metabarcoding mock communities
- Assemblage of mock communities: total of 76 zooplankton species with various levels of genetic variation across 24 different mock communities
- Next-Generation Sequencing: Illumina MiSeq 2x300bp platform
- Evaluation: Read depth and/or species detection

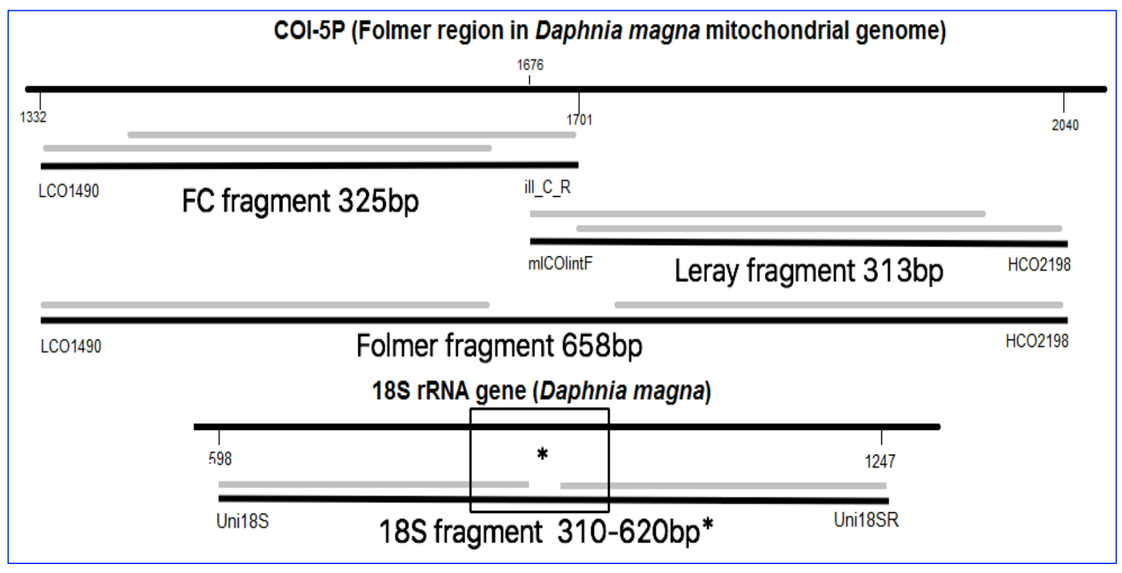

Fig. 1 The amplified fragments used for metabarcoding (FC: the 5' end of COI-5P gene; Leray: the 3' end of COI-5P gene; Folmer: whole COI-5P gene; 18S: V4 region). The primers are not included in the fragment lengths, and the gray lines refer to the forward and reverse reads from the paired-end 300bp Illumina MiSeq next-generation sequencing. *The 18S fragment sizes vary between species, resulting in some forward and reverse reads that do not overlap.

Results

The read depths differ by the lengths of amplicons and the species composition across mock communities.

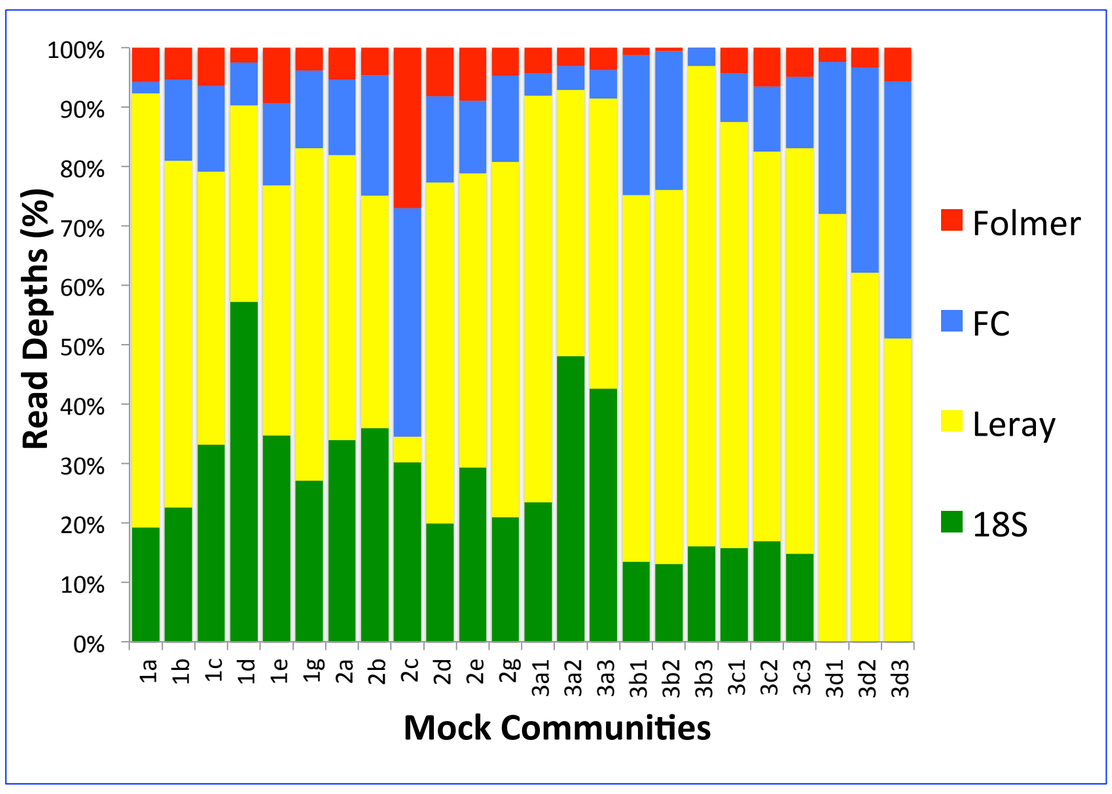

Fig.2 Read depths (%) of filtered reads of the 4 fragments. The libraries are presented as follows: 1a-1g: Single Individuals per Species (SIS); 2a-2g: Multiple Individuals per Species (MIS); 3a1-3d3: Populations of Single Species (PSS). Note the low abundant reads of 18S fragments in libraries 3d1-3d3.

The combination of 3 COI primer pairs and both COI and 18S markers increased the species recovery.

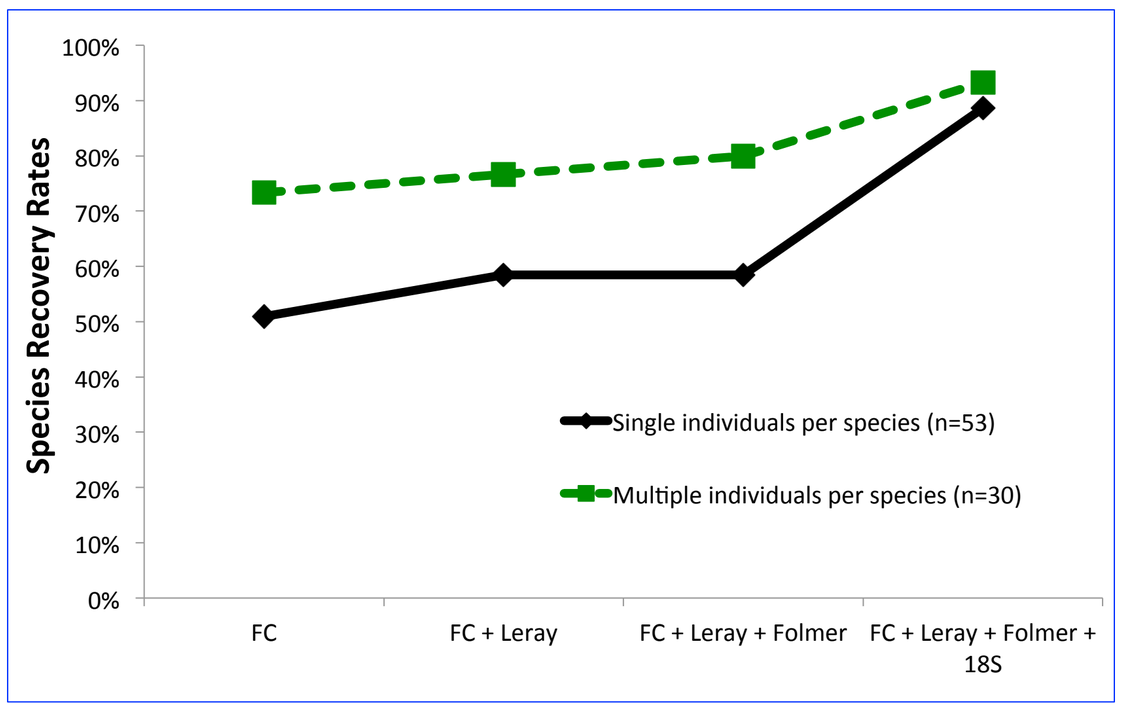

Fig. 3 Cumulative species recovery percentages after including different fragments in two mock communities that are made up of either a single individual per species or multiple individuals per species.

Discussion

Multiple primer pairs:

- The mitochondrial COI marker technically challenging for amplification of broad taxonomic groups due to the lack of conserved priming sites (Deagle *et al.* 2014)
- Multiple primer pairs per marker suggested and shown to improve amplification success (Clarke *et al.* 2014)
- Using 3 COI primer pairs covering different regions of the COI-5P gene improved species detection rates by 3.3-16.7%.

Marker choice:

- Choice of marker greatly affects species estimates in the metabarcoding studies (Bucklin *et al.* 2016)
- COI marker estimated more species than morphospecies, whereas 18S marker underestimated species richness (Tang *et al.* 2012)
- Difficulty in amplifying COI gene (Young *et al.* 2016)
- Problems in assigning reads to species (Brown *et al.* 2015)
- Combining 18S and COI markers improved by 11.3-30.2%

Conclusion

Our results suggest that a multiplexed metabarcoding approach that uses multiple markers and primer pairs can overcome amplification biases, improve taxonomic resolution across groups of zooplankton, and ultimately achieve more accurate biodiversity estimates. The calibrated approach proposed in this study of mock communities is cost effective, and useful for biomonitoring zooplankton in natural communities.

References

Bucklin A, Lindeque PK, Rodriguez-Ezpeleta N, Albaina A, Lehtiniemi M (2016) Metabarcoding of marine zooplankton: prospects, progress and pitfalls. *Journal of Plankton Research*, **38**, 393-400.

Clarke LJ, Soubrier J, Weyrich LS *et al.* (2014) Environmental metabarcodes for insects: *in silico* PCR reveals potential for taxonomic bias. *Molecular Ecology Resources*, **14**, 1160-1170.

Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biological Letter*, **10**, 20140562.

Tang CQ, Leasi F, Obertegger U, Kieneke A, Barraclough TG, Fontaneto D (2012). The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. *PNAS*, **109**, 16208-16212.

Young RG, Abbott C, Therriault T, Adamowicz SJ (2016) Barcode-based species delimitation in the marine realm: a test using Hexanauplia (Multicrustacea: Thecostraca and Copepoda). *Genome*, 10.1139/gen-2015-0209.

Funding Sources/Acknowledgements

We thank R. Young and S. Adamowicz for generating the reference sequences for many of the zooplankton species included in this study. We also thank E. Brown for help with preparing libraries. Furthermore, we would like to thank T. Crease, E. Brown and J. Flynn for helpful advice. Many of the zooplankton samples used in this study were collected by the sampling team of the Canadian Aquatic Invasive Species Network (CAISN) and Fisheries and Oceans Canada (DFO). This research was supported by the NSERC CAISN to MEC and CA and the NSERC Discovery Grant to MEC.