

SFS Annual Meeting – May 20 – 24, 2018

<https://freshwater-science.org/annual-meeting>

Understanding Variation Associated With Bioinformatics Processing Of Environmental Genomic Data

Eric D. Stein¹, Joshua A. Steele¹, Anthony Chariton², Paul A. Greenfield², Luke Thompson³, John Martinson⁴, Mehrdad Hajibabei⁵, Donald J. Baird⁶, Erik Pilgrim⁴

1. Southern California Coastal Water Research Project, CA, USA.

2. Macquarie University, Sydney, Australia.

3. NOAA SWFSC La Jolla, CA

4. EPA ORD, Cincinnati, OH, USA.

5. University of Guelph, ON, Canada.

6. Environment and Climate Change Canada.

Abstract

Metabarcoding is increasingly used in biomonitoring worldwide. The lack of standardized bioinformatics approaches for environmental metabarcoding data compromises data quality. Understanding the sources of variability in the bioinformatics process will help practitioners make informed decisions about where standardization is important and account for uncertainty associated with the bioinformatics process. This study examined variation among a diverse group of bioinformatics workflows through intercalibration. By quantifying the relative contribution of variance at each step of the bioinformatics process, we documented how bioinformatics pipelines influence the ecological interpretation of environmental metabarcoding assessments. Five participating labs each received the same 18S rRNA sequence dataset to analyze through their pipelines. We compared output between labs at sequence counts, OTUs or sequence-types, community composition, taxonomic ID and ecological relationships. Initial analyses resulted in substantial interlaboratory variability in sequence numbers, OTUs and number of families. We reduced variability among labs by reconciling several key steps in the bioinformatic process. The intercalibration process allowed us to develop recommendations that will reduce the overall uncertainty associated with application of metabarcoding for environmental assessment while still allowing flexibility in the bioinformatic process.