

# Finding eukaryotic needles in the metagenomic haystack

Abigail L. Lind  
Katherine S. Pollard

---

## Video Byte

**Keywords:** fungi, protists, algae, nematodes, arthropods, choanoflagellates, microbial eukaryotes, helminths, whole metagenome sequencing, metagenomics, EukDetect, gene database, marker genes, Microbiome

**DOI:** <https://doi.org/10.21203/rs.3.rs-318782/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Natural microbial systems are all around us, from nutrient-rich soil to churning seawater to our own bodies. These complex systems include bacteria, archaea, viruses, and microbial eukaryotes. One of the best methods available for analyzing these systems is shotgun sequencing, which generates vast quantities of DNA sequence data. However, current data annotation methods don't include a dedicated way to find eukaryotic sequences. Now, researchers have introduced a bioinformatics method called EukDetect. EukDetect uses a database of over 500,000 universal marker genes from 241 conserved gene families across thousands of eukaryotes. Broad taxonomic coverage and accurate identification of low-abundance and high-similarity sequences were possible with EukDetect, and bacterial contamination was no obstacle to identifying eukaryotic species. EukDetect highlights information that could be missed or obscured in standard shotgun sequence analysis. Future improvement of EukDetect will depend on new high-quality genome data from diverse eukaryotic species. EukDetect will illuminate the contributions of microbial eukaryotes to many different types of microbiomes as producers, predators, decomposers, and symbiotes.