Metagenomic Analysis of Polyphosphate-Accumulating Organisms During Upset Events

Ho Wa Chu[1,2], Jessica A. Deaver[1,2], Douglas F. Call[1,2]
[1]Department of Civil, Construction, and Environmental Engineering, NC State University, Raleigh, NC 27607
[2]STEPS Center, NC State University, Raleigh, NC 27607

Introduction

Polyphosphate-accumulating organisms (PAOs) are crucial to the nutrient removal processes within a wastewater treatment system.

- Capable of hyper-accumulating phosphorus up to 15% of their dry weight

However, observed instability events have disrupted PAO activity, reducing their phosphorus removal efficiency.

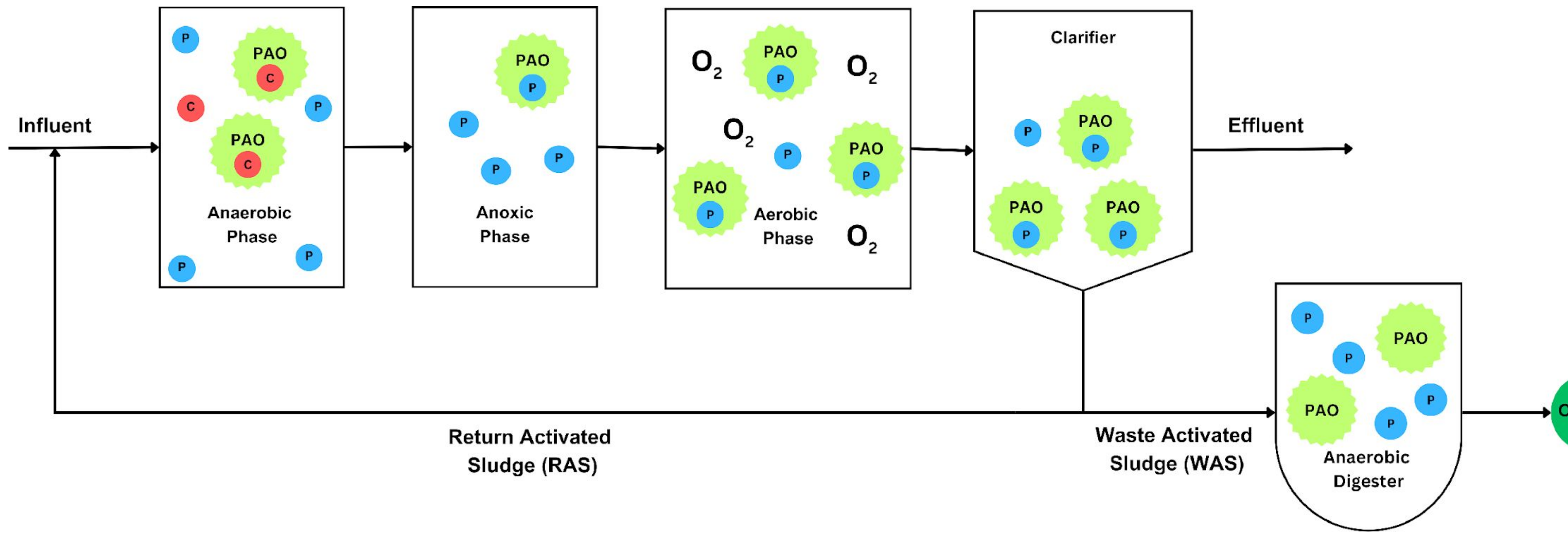
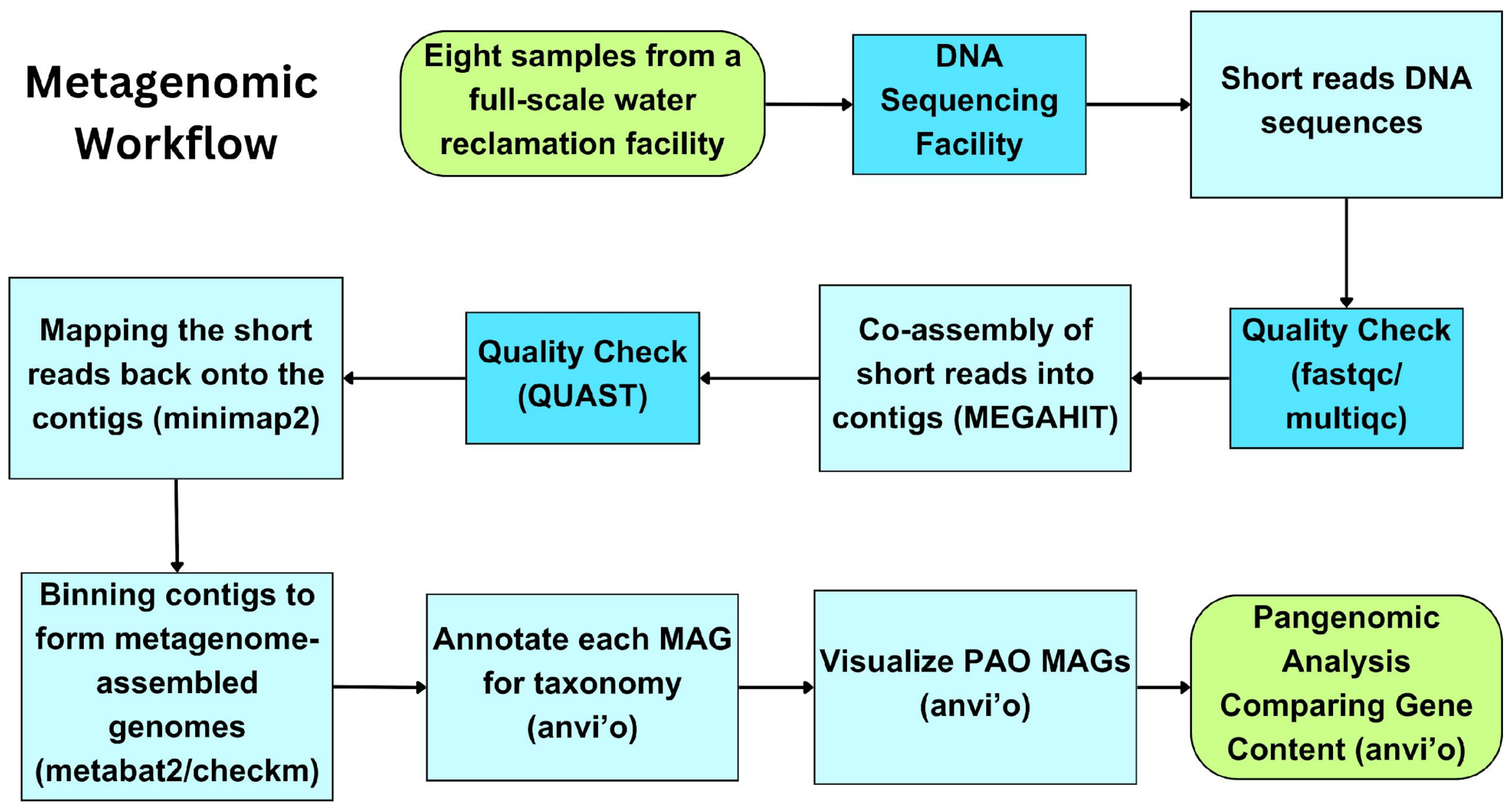


Diagram of Enhanced Biological Phosphorus Removal (EBPR) process.

Objective

The purpose of this project is to recover individual PAO genomes from metagenome samples using computational software.

Methods



Metagenomic Workflow

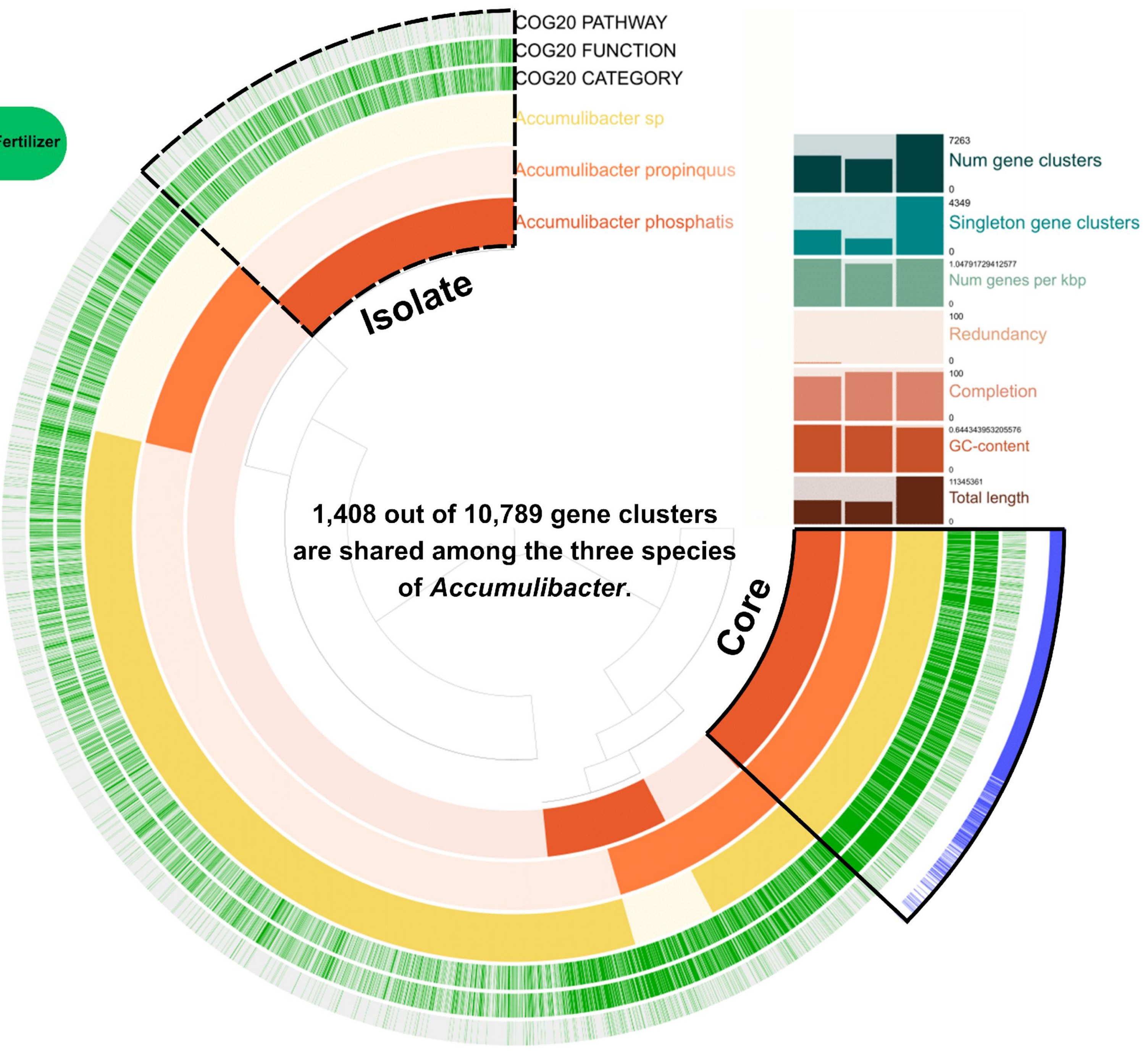
1. **Eight samples from a full-scale water reclamation facility**
2. **DNA Sequencing Facility**
3. **Short reads DNA sequences**
4. **Quality Check (fastqc/multiqc)**
5. **Co-assembly of short reads into contigs (MEGAHIT)**
6. **Quality Check (QUAST)**
7. **Mapping the short reads back onto the contigs (minimap2)**
8. **Binning contigs to form metagenome-assembled genomes (metabat2/checkm)**
9. **Annotate each MAG for taxonomy (anvi'o)**
10. **Visualize PAO MAGs (anvi'o)**
11. **Pangenomic Analysis Comparing Gene Content (anvi'o)**

Results

Table 1: Taxonomic classification of 4 metagenome-assembled genomes (MAGs) of PAOs based on single copy genes using Anvi'o.

Bins	Percent Completion	Percent Redundancy	Class	Order	Family	Genus	Species
bin_262	97.18	5.63	Actinomycetia	Actinomycetales	Dermatophilaceae	JADJVJ01	JADJVJ01 sp016716625
bin_169	91.55	0.00	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Accumulibacter	
bin_370	91.55	0.00	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Accumulibacter	Accumulibacter propinquus
bin_371	83.10	2.82	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Accumulibacter	Accumulibacter sp017302455

Out of 454 bins, we found 8 that are PAOs, 4 of which had completion > 80%



1,408 out of 10,789 gene clusters are shared among the three species of *Accumulibacter*.

Figure 1: Pangenome of three *Accumulibacter* species with core genes shared among all samples highlighted in blue.

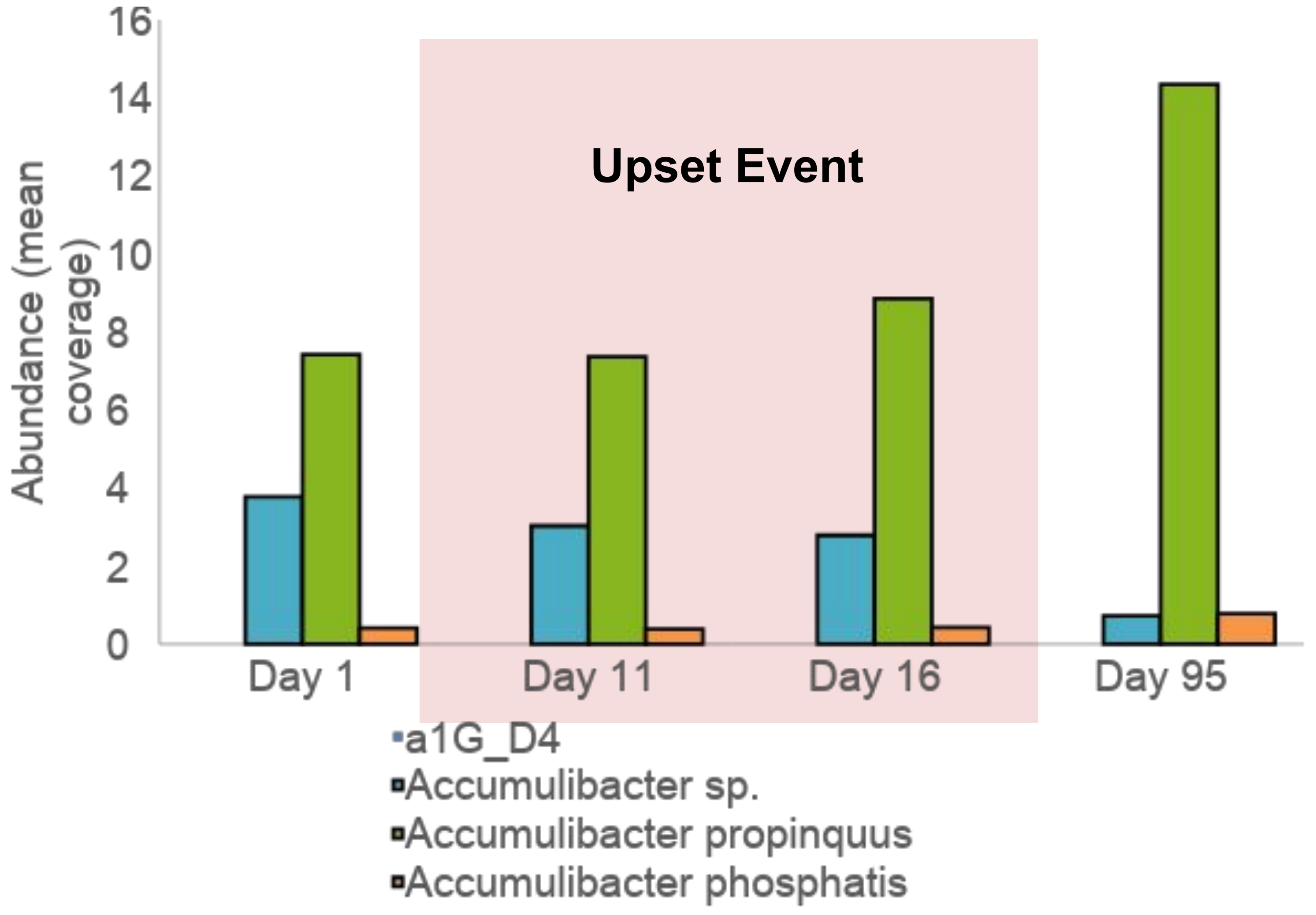


Figure 2: Abundance of three *Accumulibacter* samples from train 2 over four days.

References

Crocetti, G. R., et al. (2000). Identification of Polyphosphate-Accumulating Organisms and Design of 16S rRNA-Directed Probes for Their Detection and Quantitation. Applied and Environmental Microbiology, 66(3), 1175–1182.

Eren et al. (2020). Community-led, integrated, reproducible multi-omics with anvi'o. Nature Microbiology, 6(1), 3–6.

Lee, M. (2019). Happy Belly Bioinformatics: An open-source resource dedicated to helping biologists utilize bioinformatics. Journal of Open Source Education, 2(19), 53

Discussion / Future Works

1. High-quality MAGs from this study will serve as vital references for RNA transcripts.
2. Determine which metabolic pathways are being activated based on gene expression before, during, and after an upset event.
3. Metapangenomics compares gene content across different environments. Helps discover environmental core and accessory genes (ECG/EAG).

This material is based upon work supported by the National Science Foundation CBET-2019435.