March 1, 2023

Metagenome-assembled genomes(MAGs) Analysis for Assessing Extreme Storm (Florence) Impacts On Neuse River Estuary Plankton

Jorden Rabasco, Sangmi Jeong

Summary

To assess the impacts of Florence on microbial composition and metabolic shifts in the estuary before and after the landfall, we will generate metagenome-assembled genomes(MAGs) using metagenomic shotgun-sequencing data. To do this the raw reads from the shotgun sequencing will be run through QC, then assembly, metagenome binning and then bin taxonomic assignment. The binning will be done both within location and then when all reads are pooled together as if they were in one location, to allow comparison between the two approaches in the bins generated. To further analyze the microbial shifts due to the impacts of Florence, relative abundances will be generated on the raw reads prior to binning for each of the locations listed. This same process will be conducted on the binned results as well to facilitate further assessment. A markdown file will then be generated to assist the client with following and rescreating all analysis requested.

Background

Landfall of Hurricane Florence on September 14, 2018 near Wrightsville Beach, NC caused an extreme spike in river flow. Chemical conditions were also markedly altered post-Florence landfall (Paerl et al. 2020, 22026). The metagenomic shotgun-sequencing data of paired-end reads were obtained from 29 samples. The samples were collected at four different stations of the estuary (NRE30, 70, 100, and 180) at six post-strom times and a pre-storm time as a baseline.

Goal

1. Generate MAGs from timepoints and locations before and after the storm
2. QC bins and assess diversities of bins/taxa and assess diversity via relative abundances on unbinned data
3. Provide a pipeline for reproducing MAG generation

Methods

Bioinformatics tools for generating MAGs are introduced in (Yang et al. 2021, 6301-6314). Among the various methods, one or two methods for each stage of analysis were chosen based on Dr. Pearl's preferences (MetaBAT2, GTDB-Tk, and Kaiju, etc.) and other current literature on MAG analysis. The methods are listed in Table 1.

Disclaimer: for all new methodologies not laid out in previous literature, outcomes are not guaranteed

Table 1. The pipeline and computational tools for generating MAGs from metagenomic sequencing data

	Steps	Tools
Data	Raw reads (paired-end reads 2 x 150bp)	
Sequence QC	QC reads	fastQC
	Trim the reads	Trimmomatic
	Remove contaminant reads	bowtie2
Taxa reads	Taxonomic assignment of reads	Kaiju
Assembly	Co-assemble the filtered reads to contigs	MEGAHIT / metaSPAdes
	Assembly QC	MetaQUAST
Metagenome binning	Cluster contigs (Metagenome binning)	MetaBAT2
	Bin Quality/statistics (completeness, contamination)	CheckM
MAG	MAG taxonomic classification	GTDB-Tk
	MAG abundance	Kaiju
Analysis	Visualization of compositional diversities	R(ggplot)

Expected Results

We are expecting based on the literature that there will be significant microbial compositional shifts in the estuary associated with Florence landfall. The final result of this project is to investigate how diverse the compositions of bins/taxa in the samples are. This can be presented in stack bar charts as shown in Figure 1.
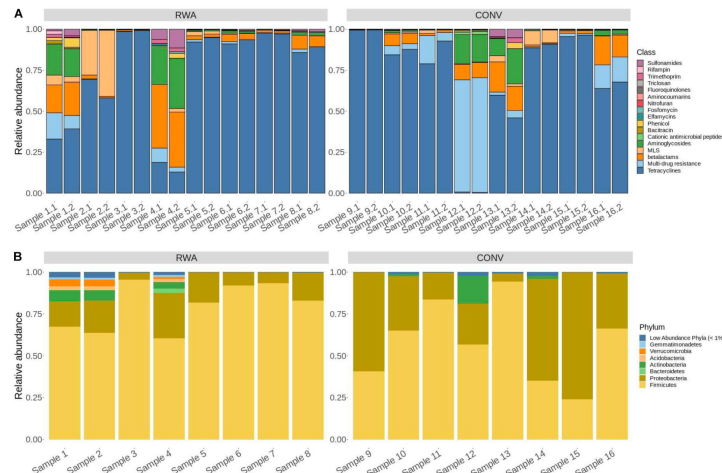
Disclaimer: for all new methodologies not laid out in previous literature, outcomes are not guaranteed

Figure 1. An example of compositional information

Deliverables

- Composition of bin diversities and composition of taxa diversities (class/order/genus/species level) presented in the stacked bars
- Composition of QC'd reads in terms of taxa present (stack bar charts)
- Bin quality : Completeness and contamination of bins
- MAG abundance table and taxa table of bins separated by location and then with locations pooled together.
- Markdown document on all analysis performed for easy replication

References

Bibliography

Paerl, Ryan W., Rebecca E. Venezia, Joel J. Sanchez, and Hans W. Paerl. 2020. "Picophytoplankton Dynamics in a Large Temperate Estuary and Impacts of Extreme Storm Events." *Scientific Reports* 10 (1) (Dec 16,): 22026. doi:10.1038/s41598-020-79157-6. https://www.ncbi.nlm.nih.gov/pubmed/33328574.

Yang, Chao, Debajyoti Chowdhury, Zhenmiao Zhang, William K. Cheung, Aiping Lu, Zhaoxiang Bian, and Lu Zhang. 2021. "A Review of Computational Tools for Generating Metagenome-Assembled Genomes from Metagenomic Sequencing Data." *Computational and Structural Biotechnology Journal* 19: 6301-6314. doi:10.1016/j.csbj.2021.11.028. https://dx.doi.org/10.1016/j.csbj.2021.11.028.

Zhou, Lei, Shihui Huang, Jiayi Gong, Peng Xu, and Xiande Huang. "500 metagenome-assembled microbial genomes from 30 subtropical estuaries in South China." *Scientific Data* 9, no. 1 (2022): 310. https://www.nature.com/articles/s41597-022-01433-z

Disclaimer: for all new methodologies not laid out in previous literature, outcomes are not guaranteed