

Metagenomics and environmental community analyses

Outline

- Introduction
 - Culture-independent analyses and why they matter
- Metagenomics
 - Challenges and applications
- Case study on mercury methylation

Dan Jones, daniel.s.jones@nmt.edu

Guest lecture for BIOL 435/535: Bioinformatics (Scarborough)

1

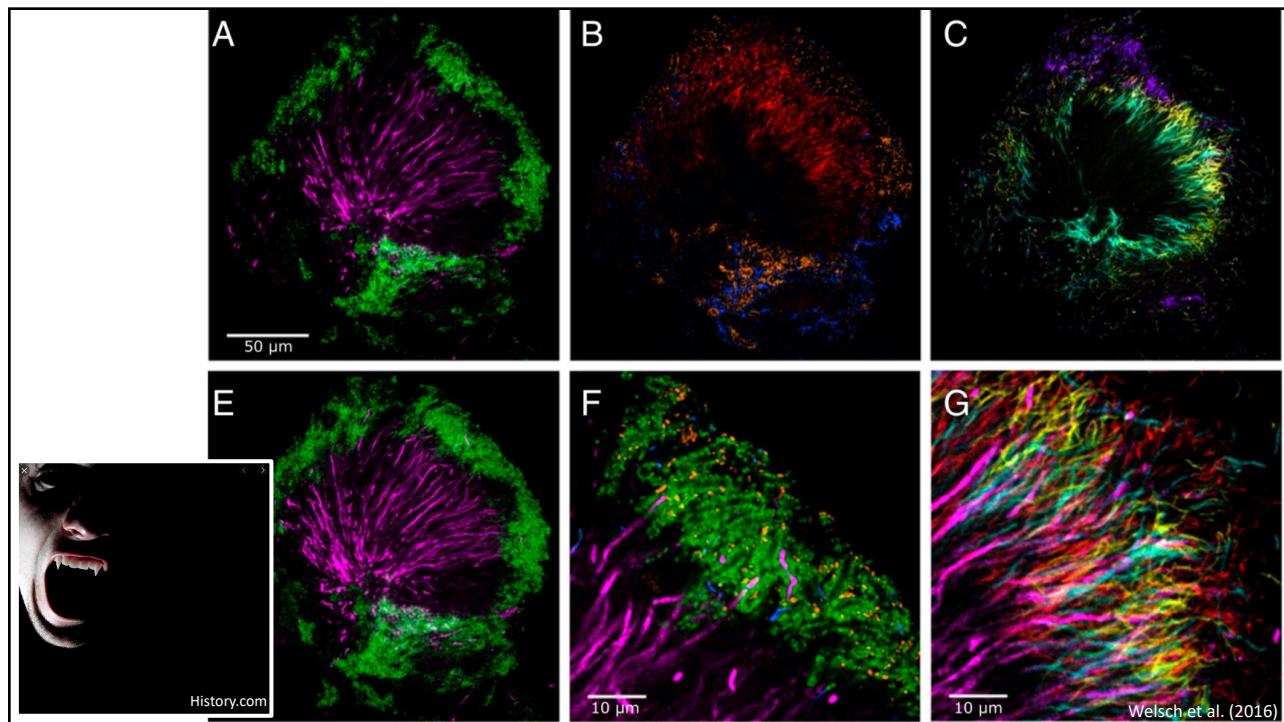
Environmental microbial communities are complex



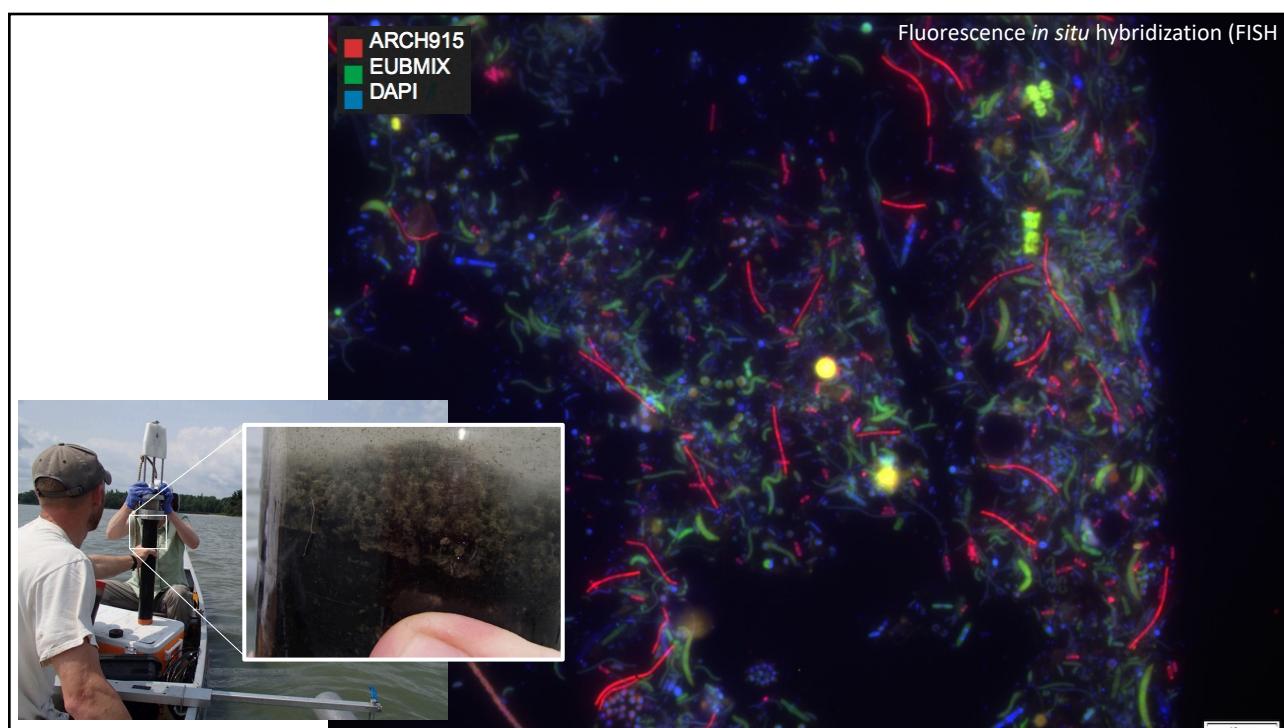
History.com



2



3

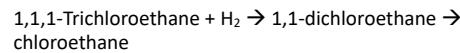
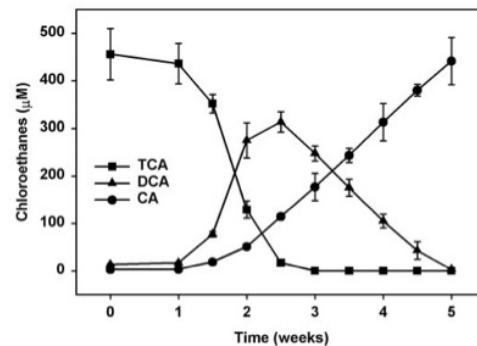


4

2

Traditionally, our primary source of information on microbial processes came from studies of microorganisms and microbial processes in culture

(And culture-based analyses remain exceedingly important)



Sun et al. (2002); Jones et al. (2016)

5

But, the most rapidly-growing laboratory strains are often only minor components of environmental microbial communities

- Organisms selected for in the lab are often not representative of the abundant or active communities in nature
 - Fast growing strains often win out in the lab
 - "Weeds"
 - Challenging to replicate the exact environmental conditions
- Enrichment bias or culture bias** can often be quite severe, and produce a biased view of microbial diversity

MEASUREMENT OF IN SITU ACTIVITIES OF NONPHOTOSYNTHETIC MICROORGANISMS IN AQUATIC AND TERRESTRIAL HABITATS

James T. Staley
Department of Microbiology and Immunology, University of Washington, Seattle, Washington 98195

Allan Konopka
Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

THE GREAT PLATE COUNT ANOMALY

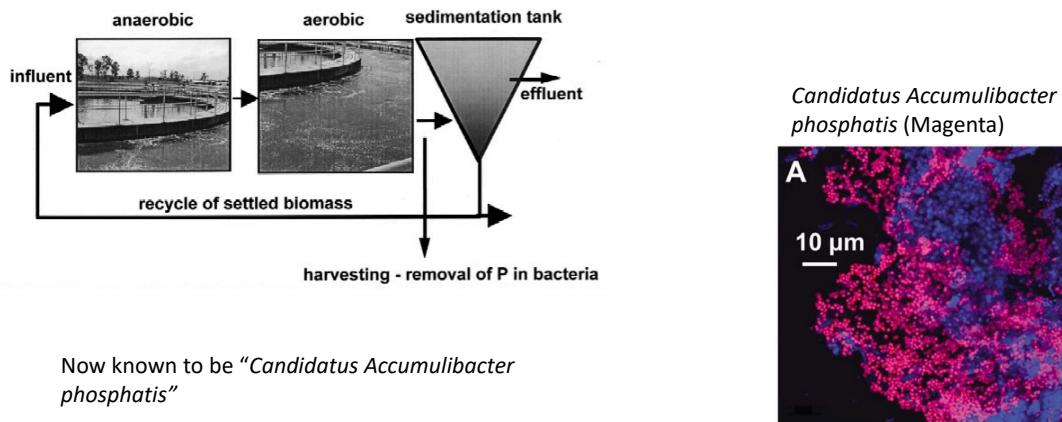
THE GREAT PLATE COUNT ANOMALY Perhaps the first report of the plate count anomaly was that of Kazunov (94), who noted a large discrepancy between the viable plate count and total direct microscopic count of bacteria from oligotrophic to mesotrophic aquatic habitats. He found higher numbers (by several orders of magnitude) by direct microscopic counting than by the plating procedure. He further noted that this same trend held in eutrophic canal water, but the differences between the two procedures were not nearly so great. Similar results were subsequently reported for marine habitats (49).

A contemporary illustration of the anomaly is provided in a seasonal distribution study of mesotrophic Lake Washington (32). Figure 1 shows the direct microscopic count using the acridine orange procedure (42), and Figure 2 shows the viable plate count for the same period of time. Viable plate counts

Staley and Konopka (1985) *Ann Rev Microbiol* 39

6

An example: The microorganism thought responsible for enhanced biological phosphorus removal (EBPR) was long thought to be *Acinetobacter* sp., based on enrichment culturing

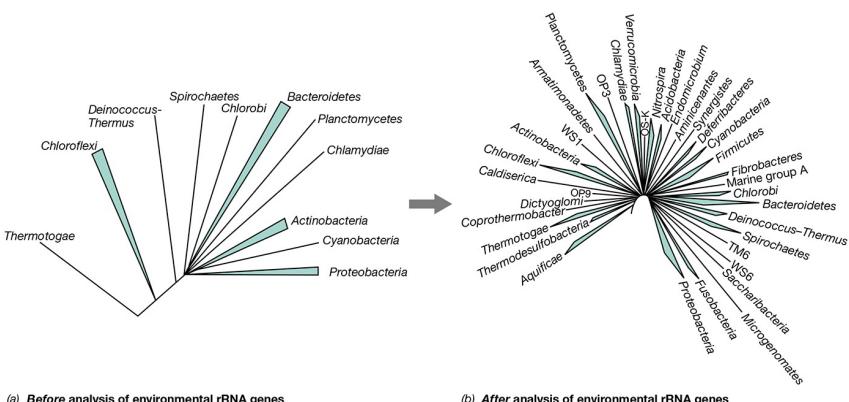


Has not yet been isolated—all of our knowledge of it comes from culture-independent analyses

Figures from Blackall et al. 2002 A Van Leeuwen 81

7

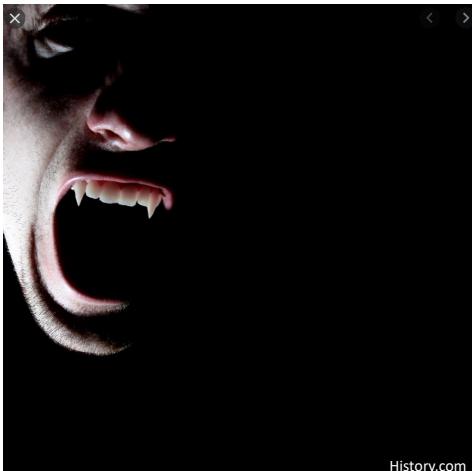
The domains bacteria and archaea contain enormous microbial diversity, much of which is only known from [environmental DNA sequencing](#)



Madigan et al. (2018)

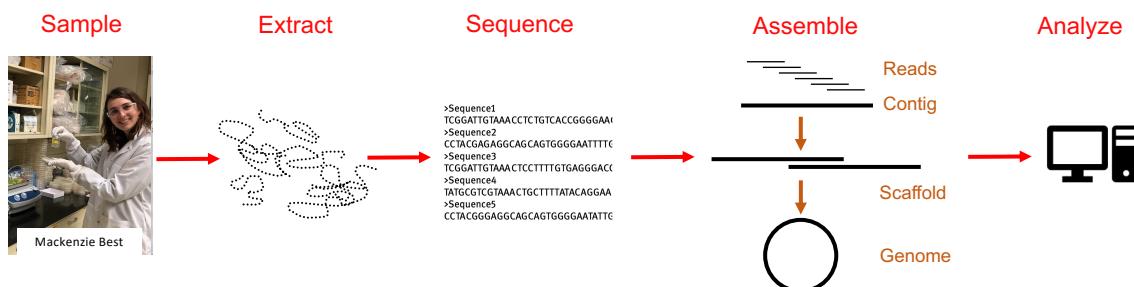
8

So how can we characterize the microbial community of an environmental sample?



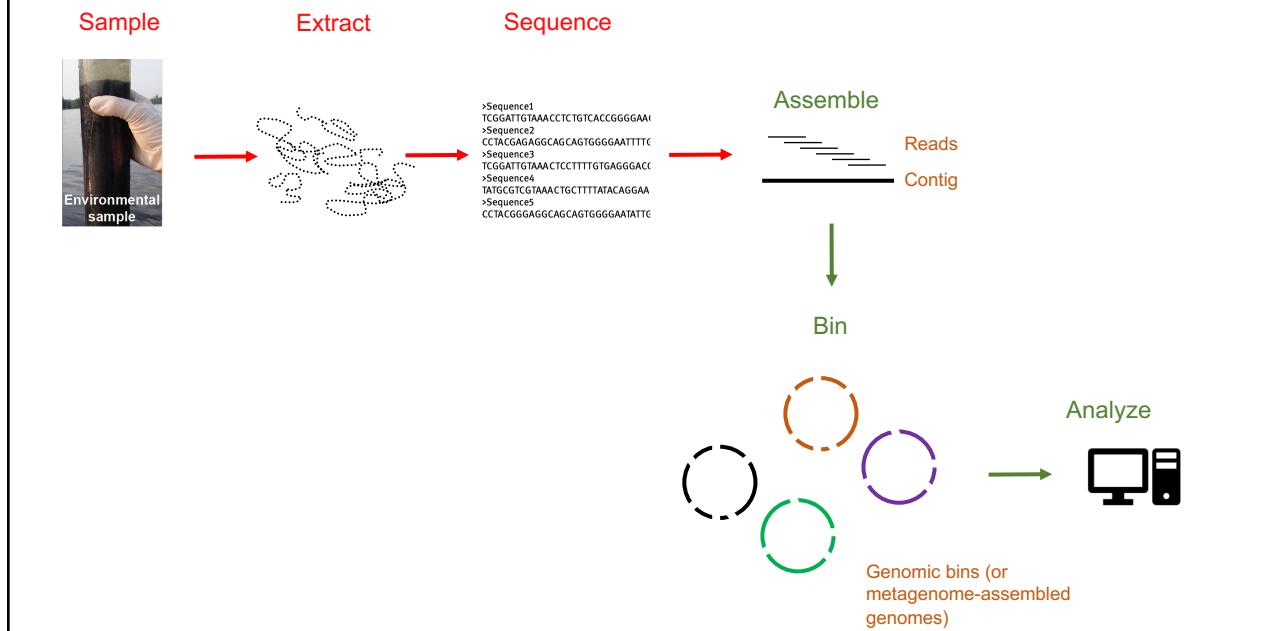
9

Brief workflow for isolate genome sequencing



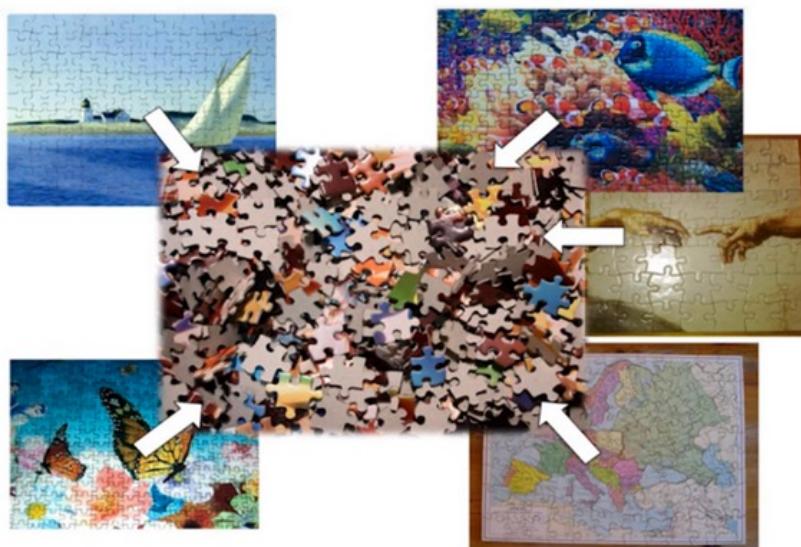
10

Metagenomics: Genomic DNA sequencing directly from an environmental sample



11

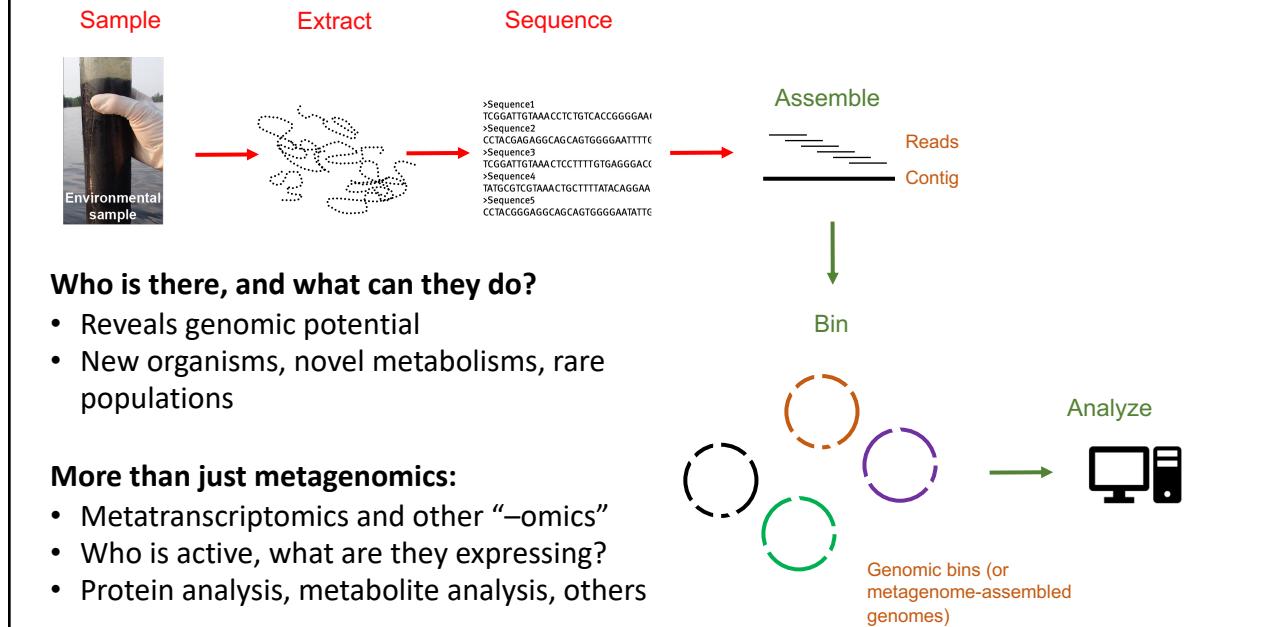
Analysis of metagenomic data is often described with a puzzle analogy



Batut and Hiltemann: <https://training.galaxyproject.org/training-material/topics/metagenomics/slides/introduction.html#1>

12

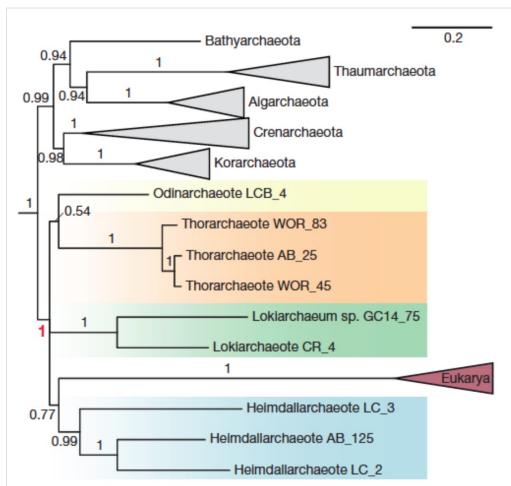
Metagenomics: Genomic DNA sequencing directly from an environmental sample



13

Example: microbial dark matter

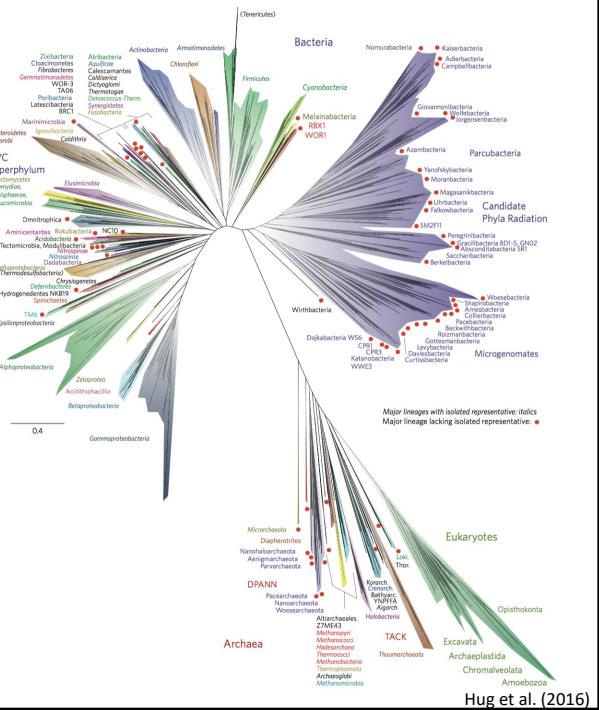
Asgard archaea illuminate the origin of eukaryotic cellular complexity

From Brett Baker's website, <http://sites.utexas.edu/baker-lab/>

14

The tree of life was recently dramatically expanded with the discovery of numerous new phyla (called the Candidate Phyla Radiation)

These organisms were previously “missed” due to biases in standard techniques for sequencing the 16S rRNA gene



15

And these exciting discoveries continue!

Article

Isolation of an archaeon at the prokaryote–eukaryote interface

<https://doi.org/10.1038/s41586-019-1916-6>

Received: 6 August 2019

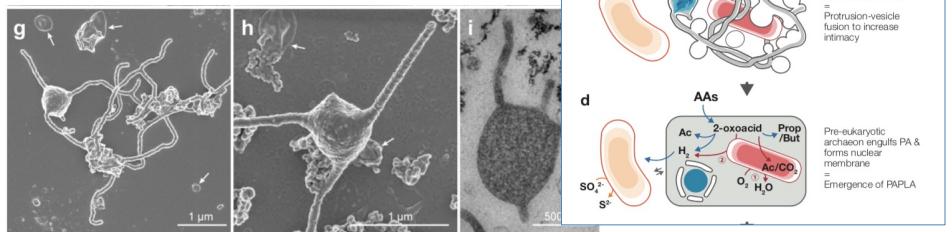
Accepted: 5 December 2019

Published online: 15 January 2020

Open access

Hiroyuki Imachi^{1,11*}, Masaru K. Nobu^{2,11*}, Nozomu Nakahara^{1,2,3}, Yuki Morono⁴, Miyuki Ogawara¹, Yoshihiro Takaki¹, Yoshinori Takano⁵, Katsuyuki Uematsu⁶, Tetsuro Ikuta⁷, Motoo Ito⁸, Yohei Matsui⁹, Masayuki Miyazaki¹, Kazuyoshi Murata⁹, Yumi Saito¹, Sanae Sakai¹, Chihong Song⁹, Eiji Tasumi¹¹, Yuko Yamanaka¹, Takashi Yamaguchi¹, Yoichi Kamagata², Hideyuki Tamaki¹ & Ken Takai^{1,10}

The origin of eukaryotes remains unclear^{1–4}. Current data suggest that eukaryotes may have emerged from an archaeal lineage known as ‘Asgard’ archaea^{5,6}. Despite the eukaryote-like genomic features that are found in these archaea, the evolutionary transition from archaea to eukaryotes remains elusive. Here we report the isolation of a new archaeon, *Candidatus Protarchaeum subterraneum*, from a marine sediment. This archaeon exhibits eukaryote-like features, such as membrane budding, vesicle transport and organelle-like structures.



Imachi et al. (2020)

16

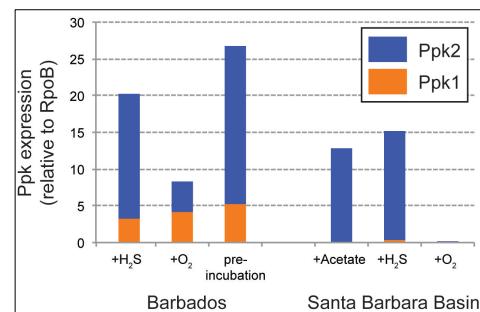
Metatranscriptomics is the analysis of gene expression from environmental microbial communities

DNA → mRNA → Protein

Metagenomics

Metatranscriptomics

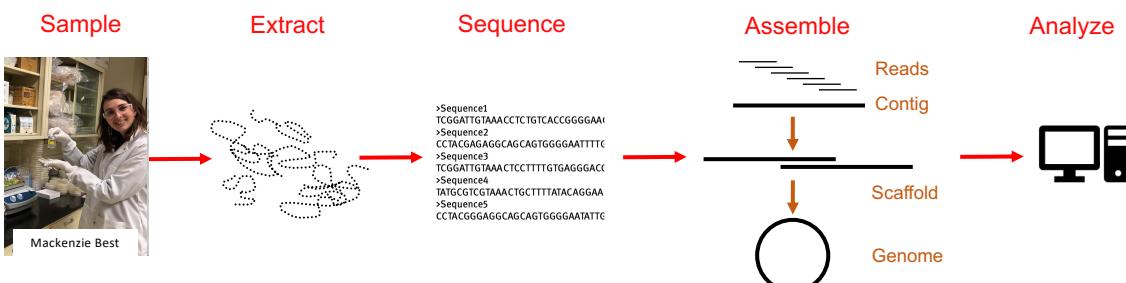
Right: Differential expression of polyphosphate kinase (*ppk*) in incubation experiments



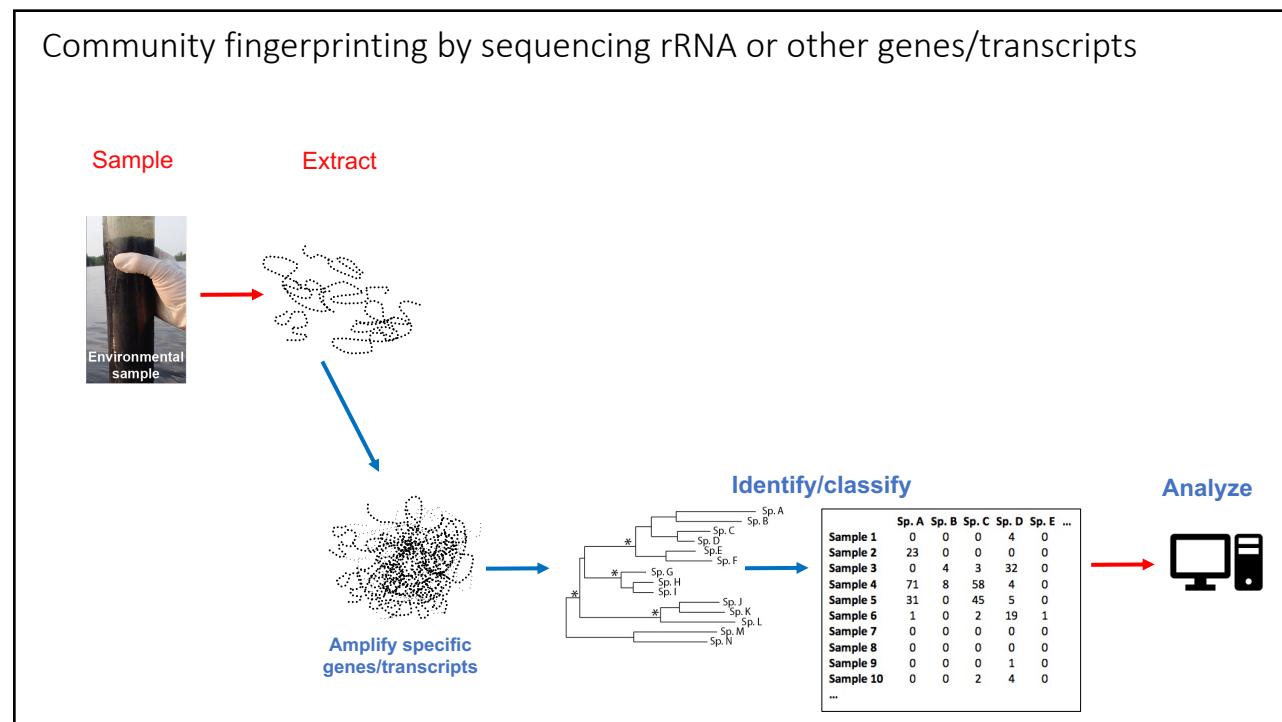
Jones et al. (2016)

17

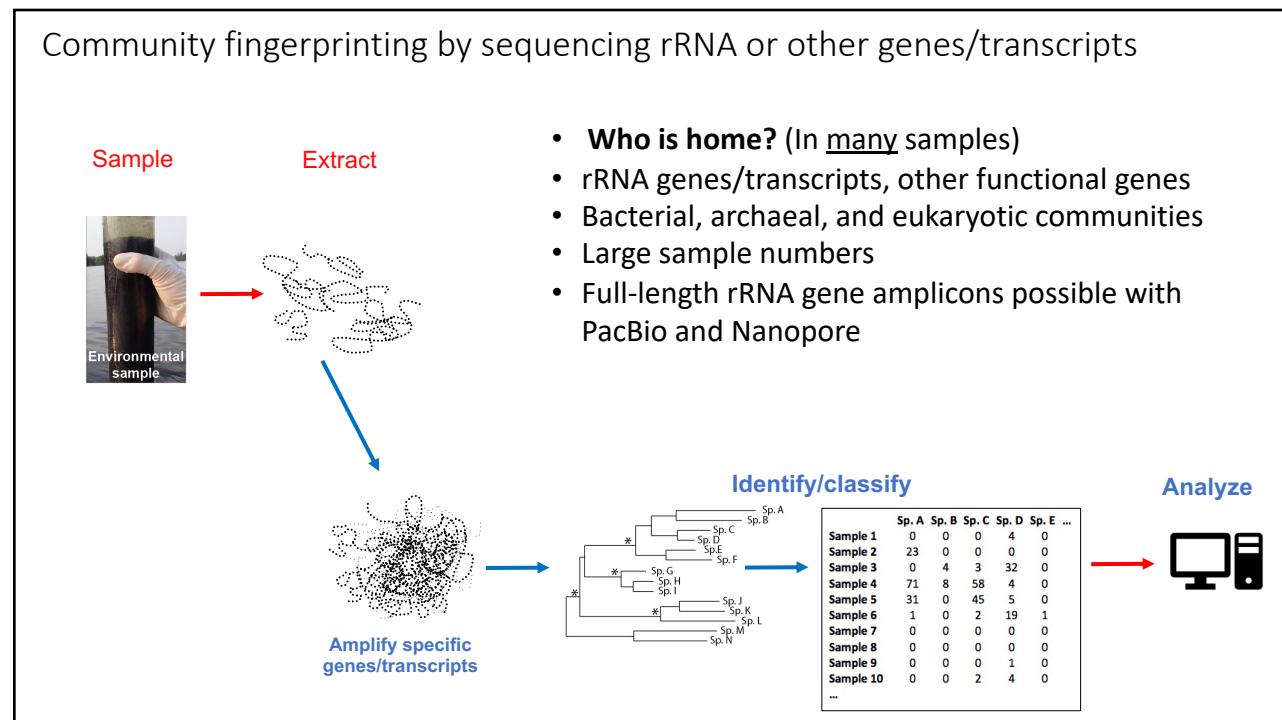
Community fingerprinting by sequencing rRNA or other genes/transcripts



18

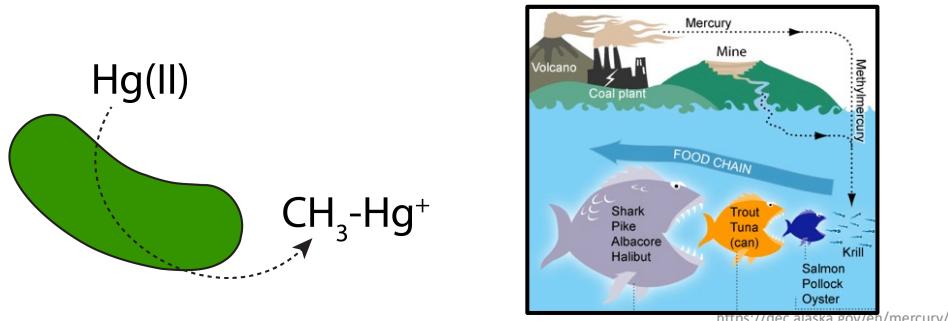


19

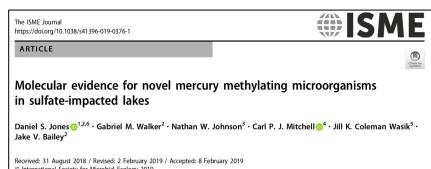


20

Case study on mercury methylating microbial communities

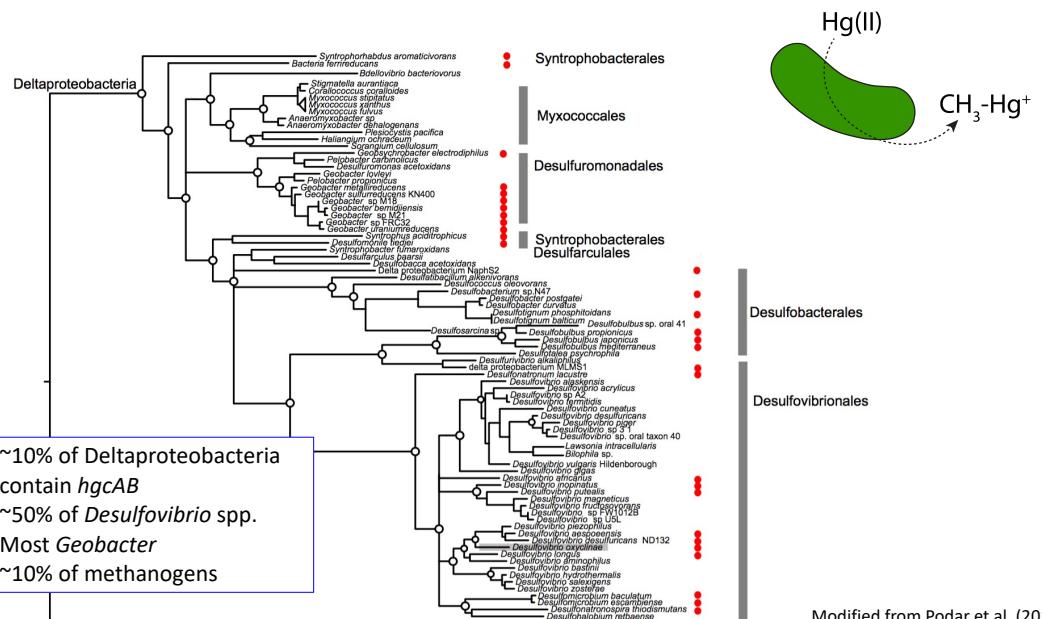


<https://dec.alaska.gov/en/mercury/>



24

Some (but only some) anaerobic microorganisms methylate mercury



25

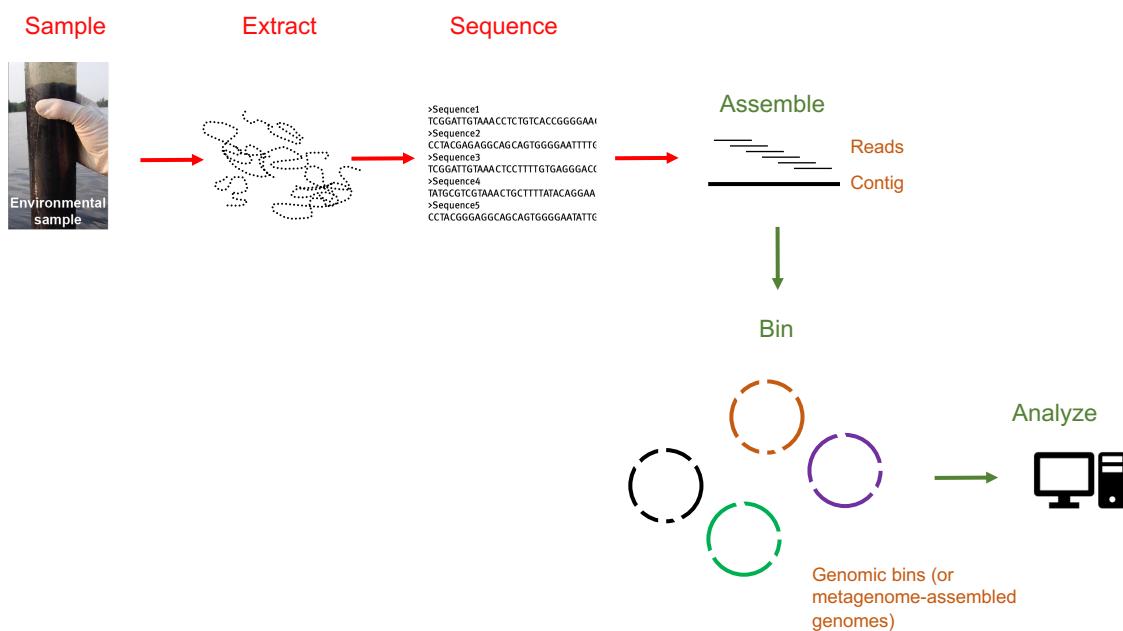
Lake Manganika is a hotspot for methylmercury production in Minnesota



Is that because of an unusual microbial community? Are methylating microbes especially abundant or diverse?

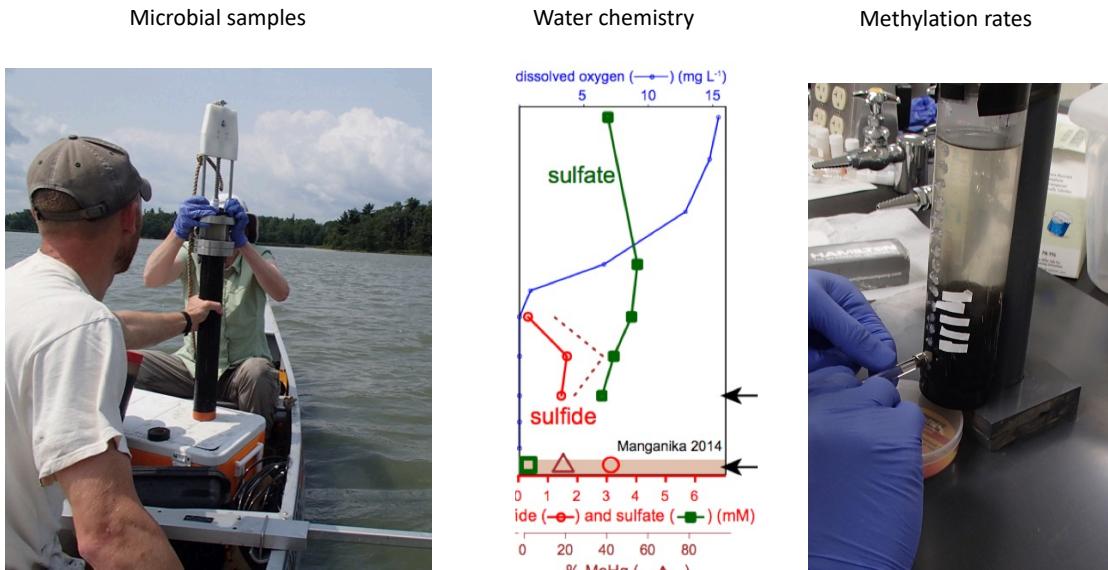
26

Metagenomic analysis of sediments and water from Lake Manganika



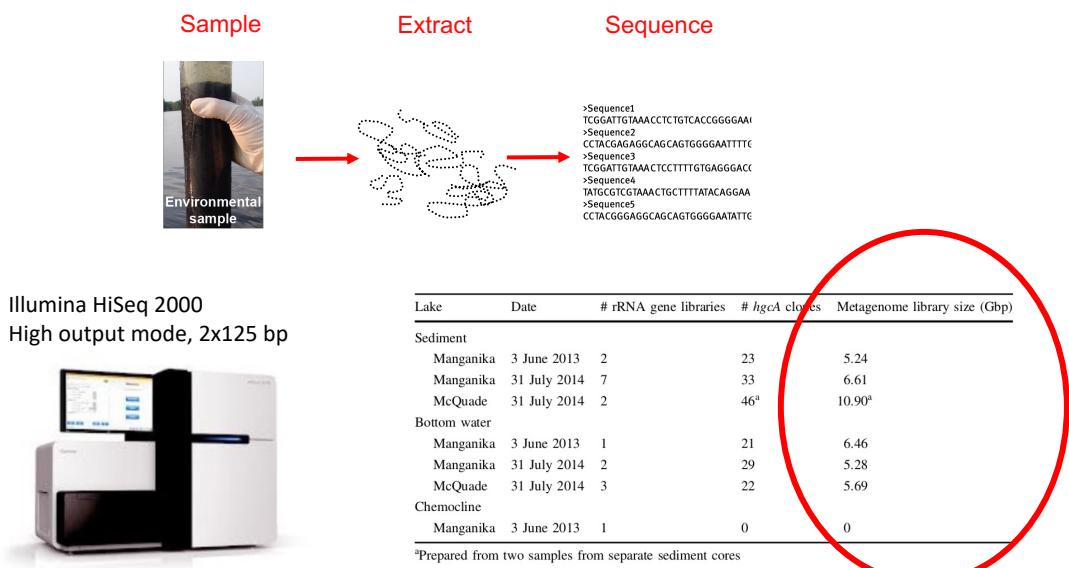
27

Step 1: Sample communities, collect associated geochemical and other data



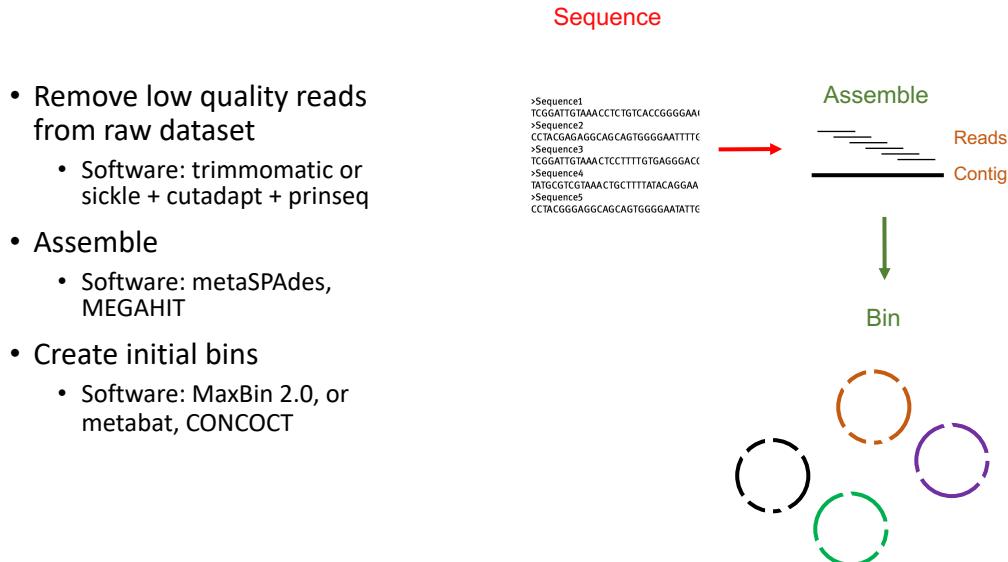
28

Step 2: Extract and sequence DNA



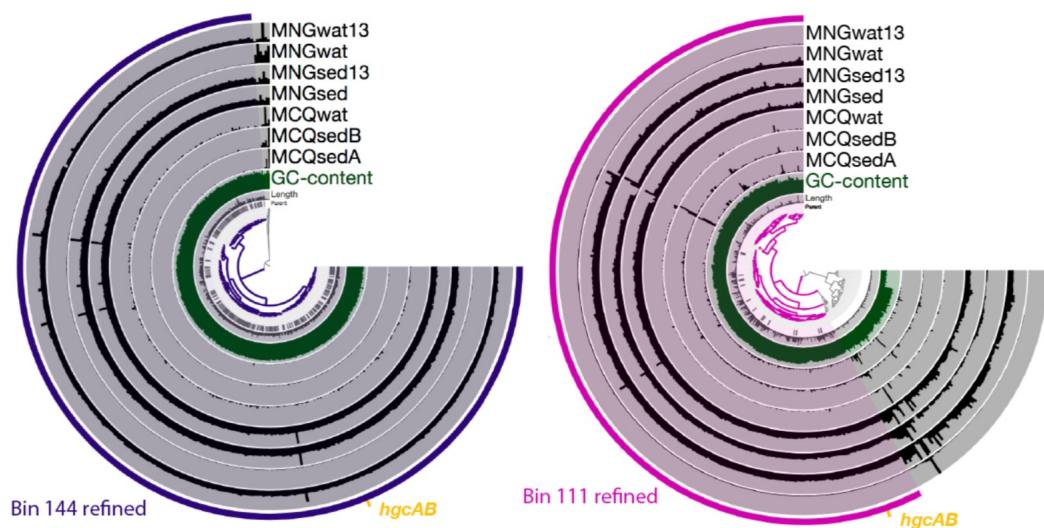
29

Step 3: Quality trim/filter raw reads, assemble, automated binning



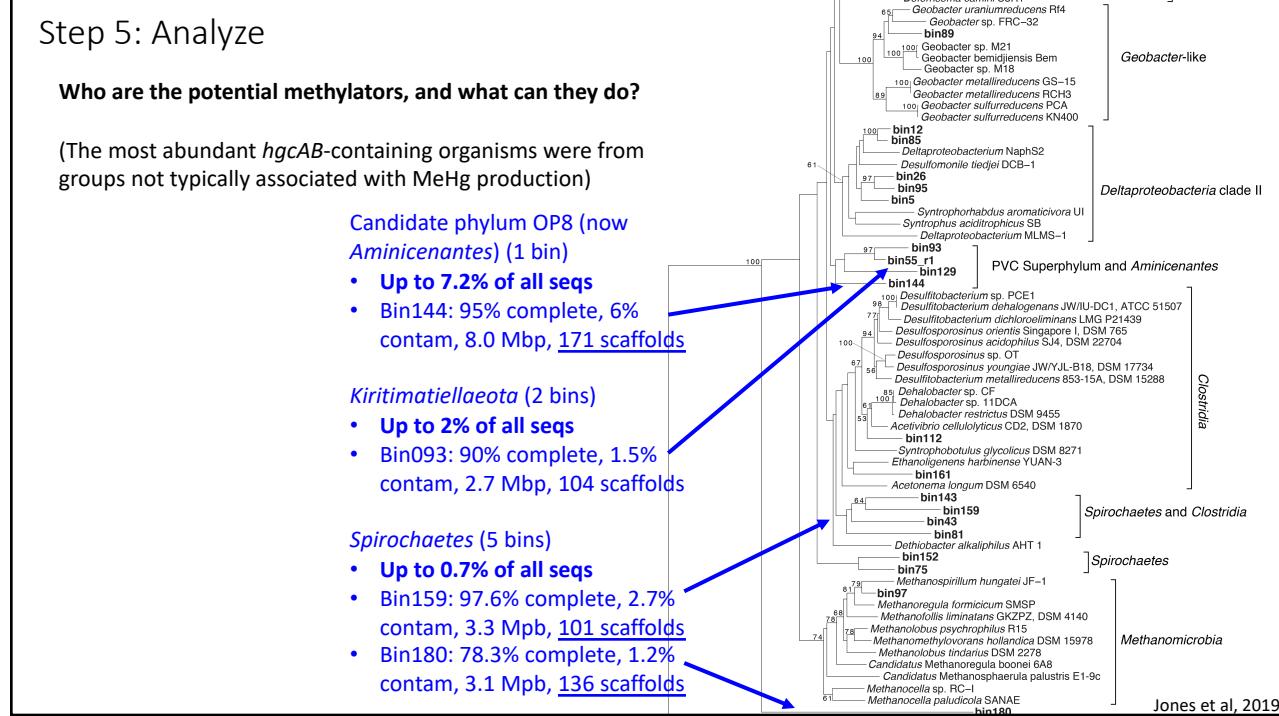
30

Step 4: Manually curate bins (with anvi'o)

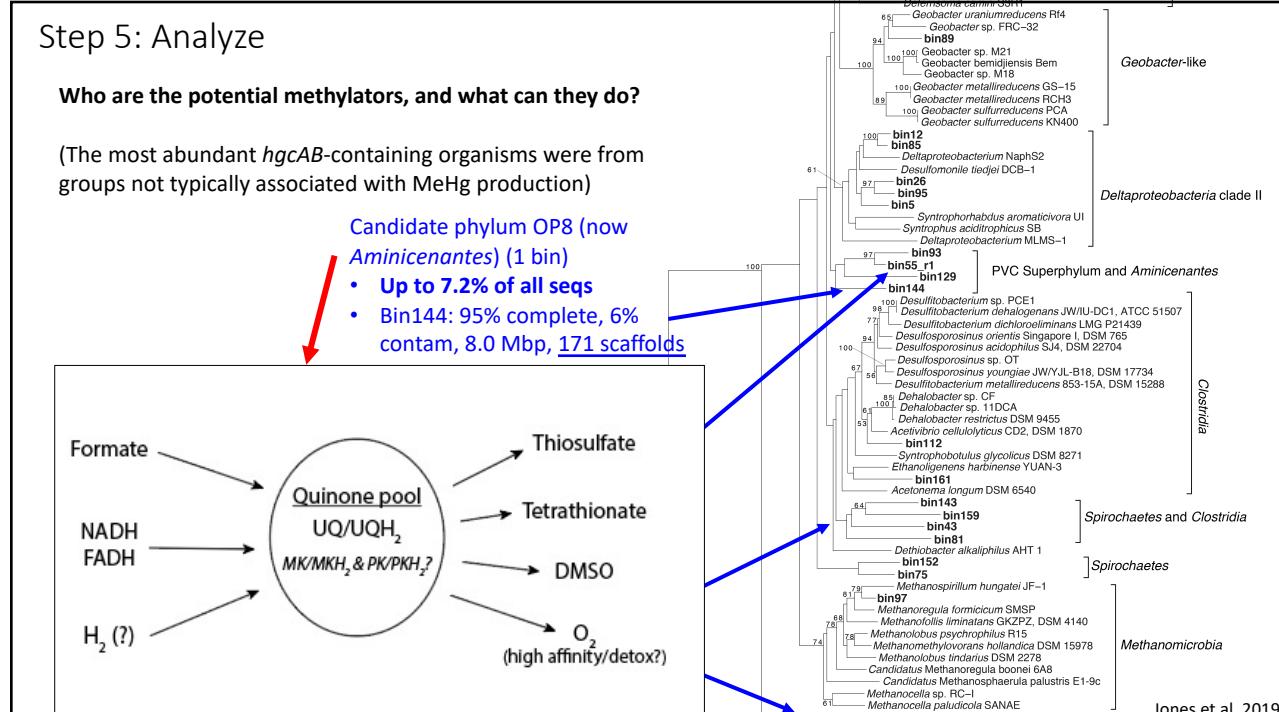


Jones et al. (2019)

31



32

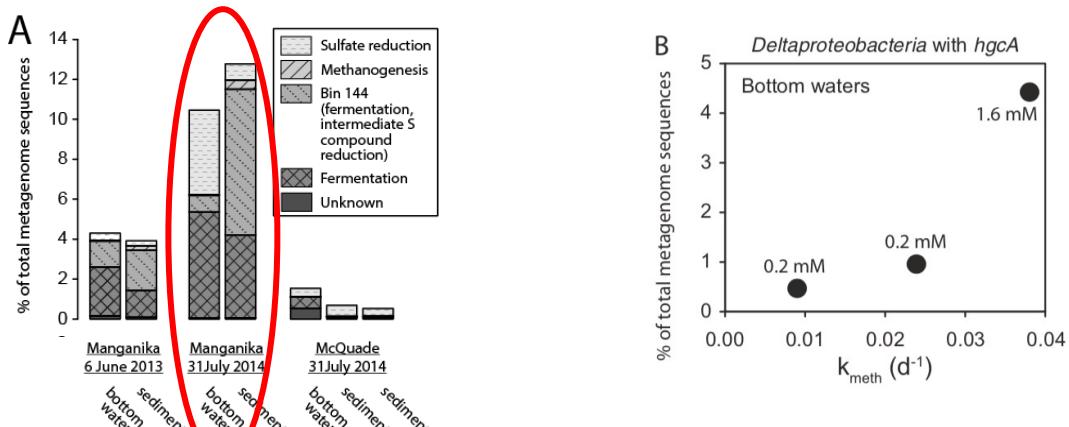


33

Step 5: More analysis

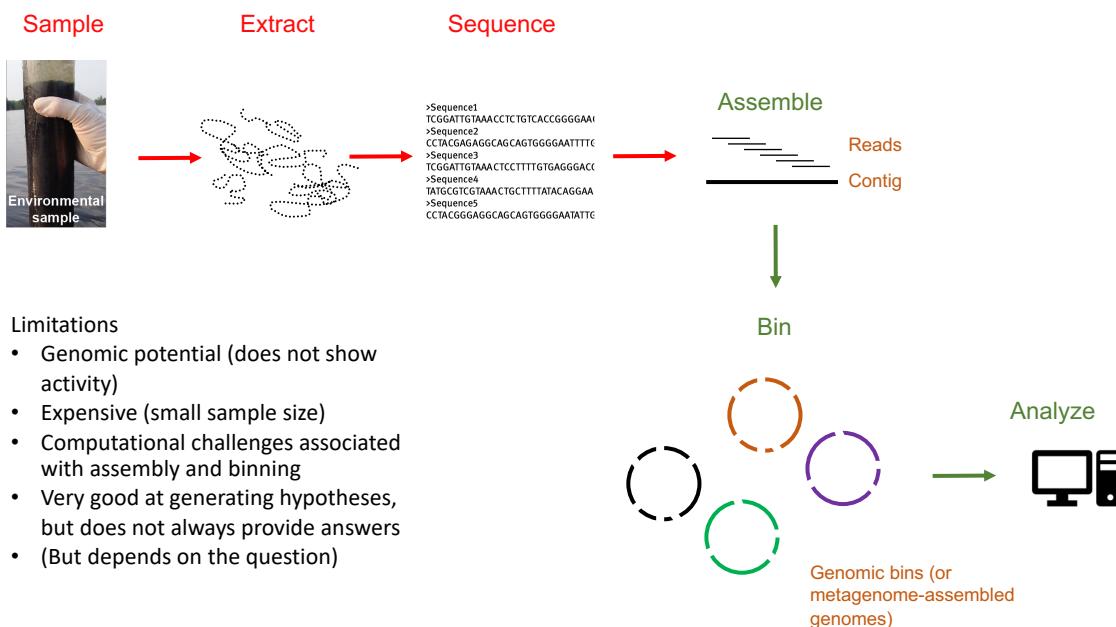
The methylating community at Manganika is especially abundant and diverse, both phylogenetically and physiologically

Sulfate reducer abundance increases with potential methylation rates (k_{meth}) in the waters



34

Metagenomics is powerful but comes with important limitations



35

Step 6: MOAR SAMPLING!!

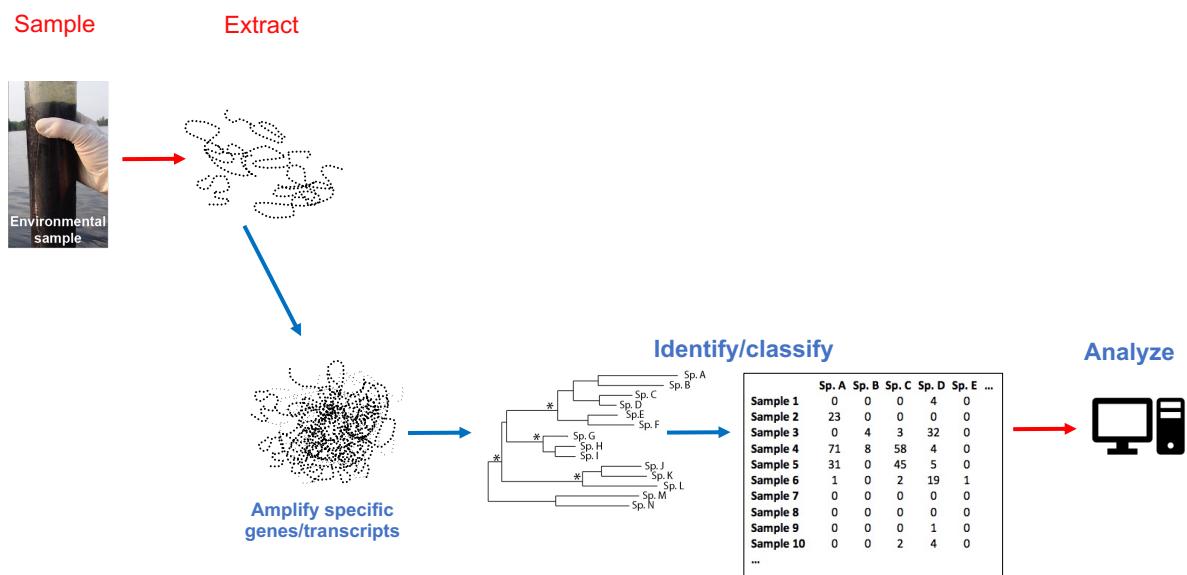
- Sulfate loading experiment
- 5 sulfate amendment levels (7, 50, 100, 150, and 300 mg/L), with 6 replicates each



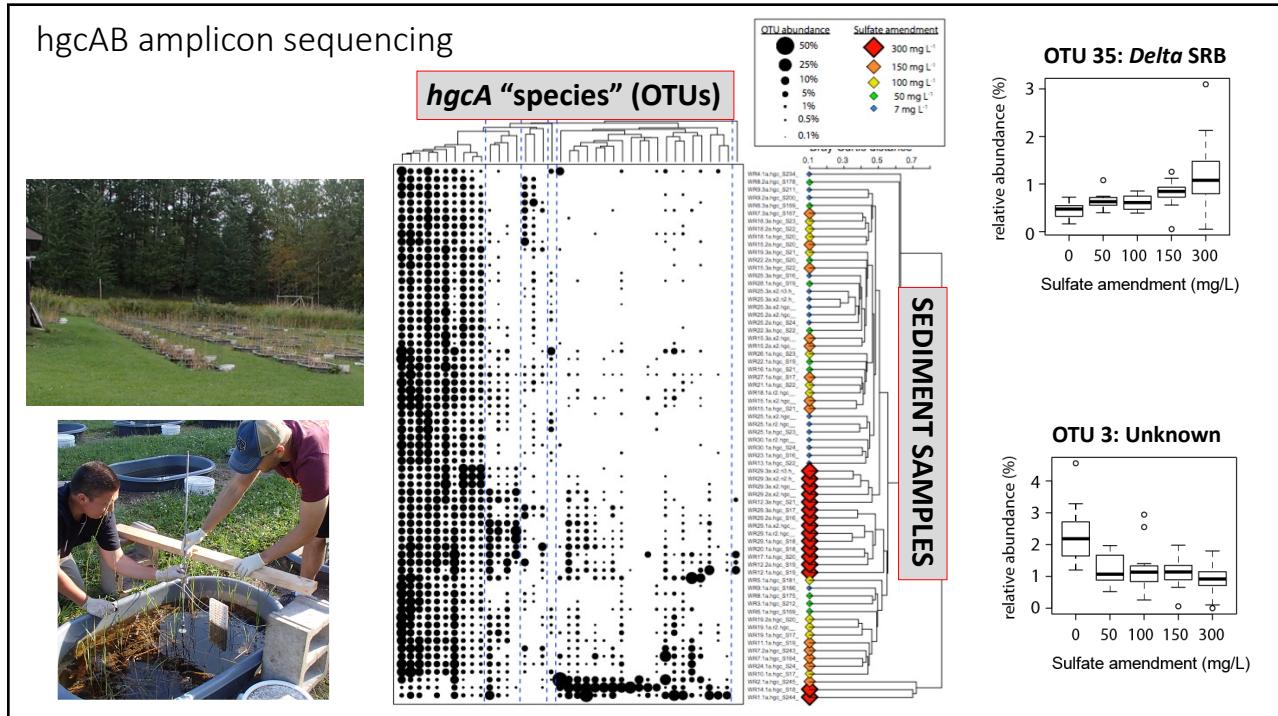
Jones et al. (2020)

36

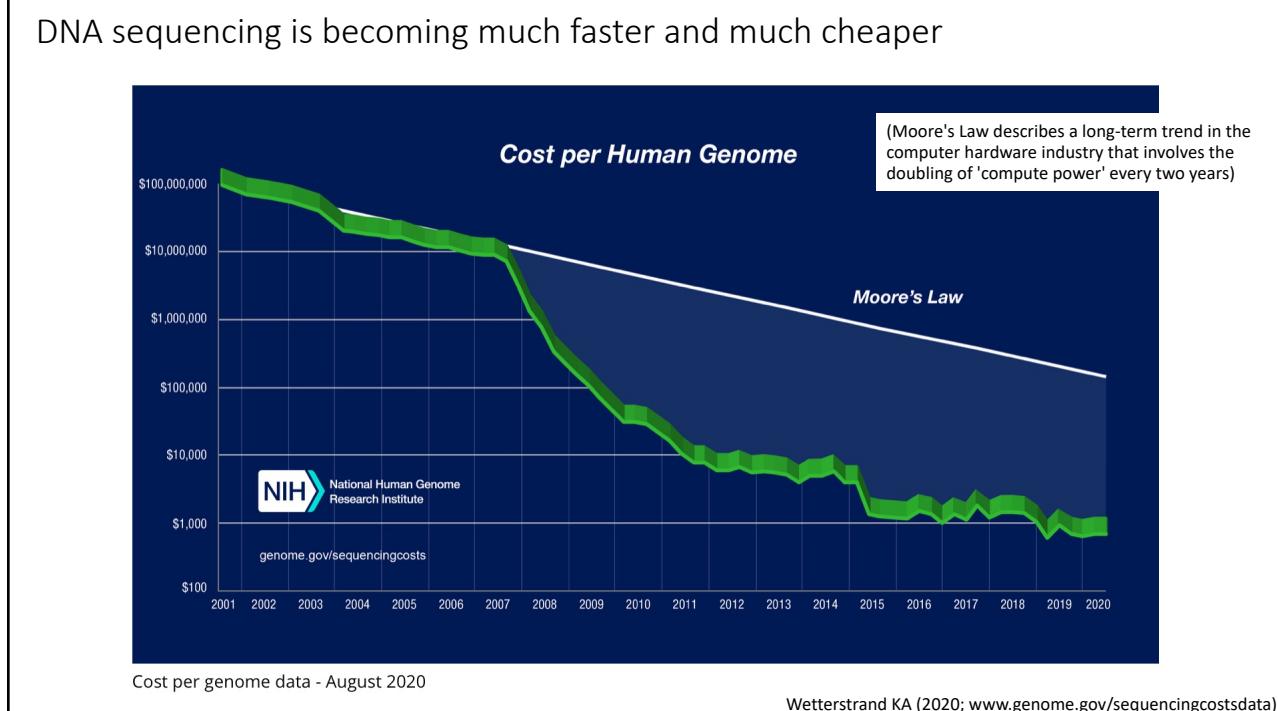
In a subsequent study, we used *hgcAB* amplicon sequencing to assess the relationship between methylator communities and geochemical variables over a much larger number of samples



37



38



39