1. Final publishable summary report

Our efforts to isolate single-cell genomes from diverse environmental samples resulted in four unique archaeal SAGs (Single-cell Amplified Genomes). They consist of two members of MCG (Miscellaneous Crenarchaeotic Group) archaea now known as ‘Bathyarchaeota’, one member of SAGMEG (South African Gold Mine Euryarchaeal Group) archaea, and one member of the Korarchaeota. The two MCGs and the SAGMEG were isolated from a hot spring sediment sample from Yellowstone National Park in USA. The new Korarchaeum was isolated from a hot spring sample from Little Hot Creek in California, USA. The discovery and phylogenetic placement of the two MCG and the new Korarchaeal SAGs were reported in the journal Philosophical Transactions of the Royal Society B in 2015 [1]. This paper reviewed the recent technological advances and discoveries made in the field of microbial ecology by the use of both single-cell and metagenomic approaches. The paper also described the novel archaeal draft genomes obtained through single-cell genomics and metagenomics in our laboratory.

With the recent advances in metagenomic binning (i.e., separating genomic DNA fragments by organism from a community), we embarked on a few projects to extract community DNA from several environmental samples and sequenced their metagenomes, a conglomeration of DNA fragments originating from a multitude of organisms. One particular sample was obtained from a gravity core sample taken from within the vicinity of Loki’s Castle hydrothermal vent in the mid-Atlantic ocean near Svalbard. Using a form of binning method on the genomic fragments obtained from this particular sample, we managed to extract three near-complete draft genomes of members of archaea previously known as the DSAG (Deep-Sea Archaeal Group).

We discovered that the DSAG contained genes that have only been found in eukaryotes so far and the finding was unprecedented. These genes consist of actin, GTPases, and several other genes known to be involved in intracellular vesicle trafficking; all are hallmark features of eukaryotic cellular machinery. In addition, using conserved single-copy marker genes, phylogenomic reconstruction of these three archaea was carried out and the results showed that they are they closest archaeal (and microbial) relatives of eukaryotes. The results strongly alluded to the fact that they may in fact be so-called the ‘missing links’ in the origin of eukaryotes. The implications from this study were immense and the findings were published in the journal Nature in 2015 and the DSAG were given a name: Lokiarchaeota [2].

The Lokiarchaeota paper in Nature attracted huge media attention and several news media publicized the findings. Some of the high-profile presses were: the New York Times, National Geographic, BBC, and NPR. Based on Altmetrics, a system of scoring the publicity of articles based on how they have been shared or discussed online, the article ranked number one among all the articles published from Uppsala University and second among all the articles published from Sweden in 2015 (http://www.natureindex.com/country-outputs/Sweden). This has also resulted in collaborative projects with several researchers from other institutions for the follow-up studies. Currently a manuscript is being prepared to present more relatives of the Lokiarchaeota that may change the way we think about the origin of eukaryotes.

Another high-profile publication resulting from the project was the genomic reconstruction of the members of SAGMEG. One of the SAGMEGs isolated from the Yellowstone National Park sample sparked our interest to sequence the metagenome of the same sample (see **Figure 1**). The resulting metagenome reveal a highly diverse community consisting of both bacteria and archaea. Using DNA fragments assembled from the SAG as a cue, we identified genomic 'bins' from this metagenome and retrieved near-complete draft genomes of two SAGMEG archaea from it using an ESOM (Emergent Self Organizing Maps) based approach (see **Figure 2**). Two related SAGMEGs were also obtained from a marine sediment sample by our collaborator Dr. Brett Baker from University of Texas and we worked collaboratively to understand their phylogenetic placement and metabolism. They occupy a new place along the archaeal tree of life and are found to be a new class of archaea. They were subsequently renamed as “Hadesarchaea” and the findings were published in Nature Microbiology in 2016 [3]. The Hadesarchaea also seem to be metabolically versatile, having genes to metabolize carbon monoxide, a highly poisonous gas. The combined single-cell and metagenomics approach also highlighted the power of both techniques and how one can really harness the best of both worlds to come up with the best results.

Figure 1. Hot spring near Lower Culex Basin in Yellowstone National Park, USA.

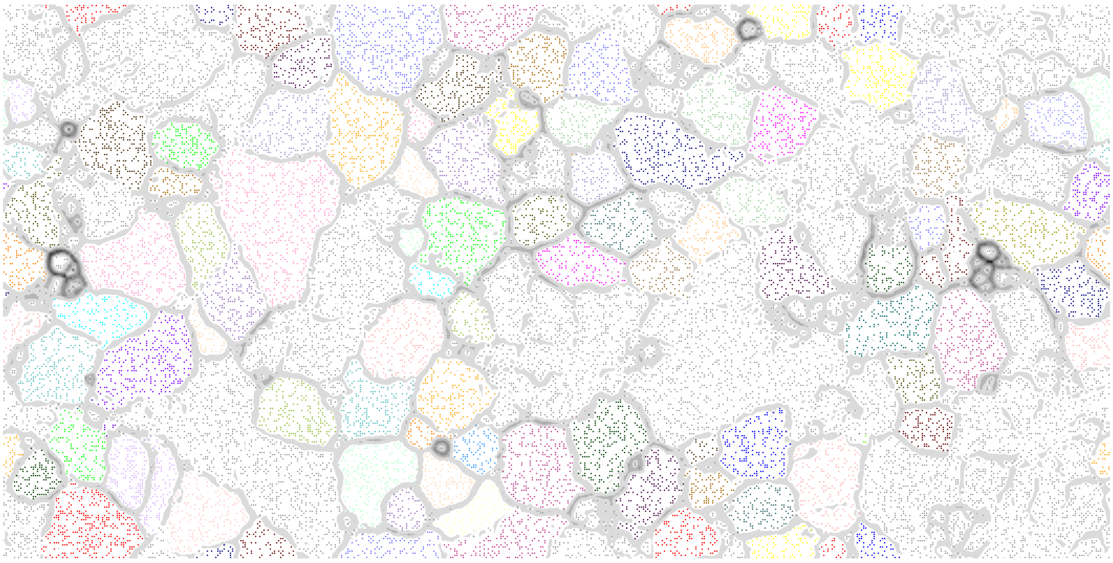

Figure 2. Emergent Self-Organizing Map (ESOM) of metagenomic sample from Yellowstone showing genome bins in different colors.

As part of our efforts to characterize uncultured microbial majority, we have currently obtained 16S rRNA amplicon sequences from about 180 environmental samples. They samples obtained were diverse in nature and came from all over the world. We are currently analyzing the microbial diversity of these samples an in time they will be processed for both single-cell genomics and metagenomics studies. The range of novel organisms discovered (both bacteria and archaea) will very likely grow as a direct result of these efforts.

All recent updates from the project are shown on the lab website at: http://www.ettemalab.org/

References

1. Saw JH, Spang A, Zaremba-Niedzwiedzka K, Juzokaite L, Dodsworth JA, et al. (2015) Exploring microbial dark matter to resolve the deep archaeal ancestry of eukaryotes. Philos Trans R Soc Lond B Biol Sci 370.
2. Spang A, Saw JH, Jorgensen SL, Zaremba-Niedzwiedzka K, Martijn J, et al. (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature 521: 173-179.
3. Baker BJ, Saw JH, Lind AE, Lazar CS, Hinrichs KU, et al. (2016) Genomic inference of the metabolism of cosmopolitan subsurface Archaea, Hadesarchaea. Nat Microbiol 1.