

## WHOLE GENOME SHOTGUN PHYLOGENOMICS IN THE THALASSIOSIRALES

Wade R. Roberts<sup>1</sup> and Andrew J. Alverson<sup>1</sup>

<sup>1</sup>Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701

Understanding how species evolve and adapt to new and changing environments have been long-standing goals in evolutionary biology. The Thalassiosirales is a diverse diatom lineage containing multiple transitions between freshwater and marine environments. We plan to use this lineage to investigate the genomic mechanisms involved in these transitions. Recently, it has become feasible and inexpensive to sequence the entire genomes of numerous species in parallel using shotgun sequencing. Using both short read (Illumina) and long read (Oxford Nanopore) technologies, we will sequence, assemble, and analyze low coverage genomes of > 50 Thalassiosirales species that span both freshwater and marine environments. These data will be used to explore phylogenetic relationships and how genome size, genome complexity, and gene family evolution relate to the adaptations to these different environments. We have currently sequenced and assembled draft genomes of *Cyclotella cryptica*, *Cyclostephanos tholiformis*, and *Thalassiosira profunda* and estimated their genome sizes to be 163 Mb, 33 Mb, and 45 Mb, respectively. Long read sequencing has greatly improved the reference genome of *C. cryptica*, bringing the number of contigs down from 116k to 839 while increasing the N50 from 12 kb to 596 kb. These preliminary results demonstrate the power of long read technologies to sequence and assembly quality, low coverage genomes across entire lineages.