



## **ABSTRACT**

*“How the life cycle impacts transcriptome and genome contents in the globally significant coccolithophorid *Emiliana huxleyi*”*

**Peter von Dassow**

**European Commission Marie Curie International Incoming Research**

**Fellow**

**Centre National de la Recherche Scientifique**

**Station Biologique de Roscoff**

**FRANCE**

*Emiliana huxleyi* has become a favorite model coccolithophorid as it forms massive blooms, is easily grown in the laboratory, and is now the subject of two whole genome sequencing efforts. *E. huxleyi* exhibits the general haplo-diplontic life cycle, alternating between haploid (1N) and diploid (2N) cells that are both capable of unlimited vegetative growth. *E. huxleyi* 2N cells are calcified and non-motile whereas 1N cells are non-calcified, covered by organic scales, and highly motile. 1N and 2N cells also differ in physiology and susceptibility to viruses. Our comparisons of isogenic 1N and 2N cells using normalized Sanger-sequenced EST libraries and non-normalized EST libraries sequenced by 454 showed great differences between the two transcriptomes. Only 27% of expressed genes were observed to be common between the two Sanger-sequenced libraries, and re-sampling statistics estimated that only 50% of expressed genes were shared. The 1N library was estimated to contain 16-19% fewer expressed genes than the 2N library. Lower transcriptome richness in haploid cells suggests they are intrinsically more streamlined whereas diploid cells may be intrinsically more prepared to exploit a diversity of rich environments. The major functional category of transcripts differentiating haploids included signal transduction and motility genes. Diploid-specific transcripts included  $\text{Ca}^{2+}$ ,  $\text{H}^+$ , and  $\text{HCO}_3^-$  pumps possibly important in calcification. Potential control factors differentiating the transcriptomes included haploid-specific Myb homologs and an unusual diploid-specific histone H4 homolog. BLAT comparisons of the 1N and 2N libraries generated in our study (from strain RCC1216/1217) with EST libraries generated from two other strains (CCMP1516 and CCMP371) showed that average sequence divergence among *E. huxleyi* strains is very low. In contrast, mapping of RCC1216/1217 Sanger and 454 ESTs onto the CCMP1516 genome sequences from JGI suggests a striking level of gene content variability between strains, with the sequenced 1516 genome selectively losing haploid-specific genes. Most surprisingly, the CCMP1516 genome appears to have lost potentially 50% of flagellar genes identified in the 1N transcriptome, genes which have been highly conserved and maintained over most of eukaryotic evolution. Gene loss and rearrangement in CCMP1516 is being confirmed by PCR and genomic microarray comparisons. This suggests that the *E. huxleyi* genome is highly flexible and undergoing rapid genome evolution, perhaps in culture and perhaps in nature. Broadly, analysis of transcriptome changes through life cycles in phytoplankton proves to be a potent strategy for unveiling likely candidate genes for interesting processes including biomineralization and transcriptional control, and also for revealing targets of on-going genomic evolution.