

Optimization of Genome Engineering in the Model Pennate Diatom *Phaeodactylum tricorutum* using CRISPR-CAS9

Mark Moosburner^{1,2}, Andrew E. Allen^{1,2*} (aallen@jvvi.org)

¹Scripps Institution of Oceanography, UC San Diego, La Jolla, CA; ²Microbial and Environmental Genomics, J. Craig Venter Institute, La Jolla, CA

Genetic engineering can change an organism into a biological factory for producing industrially valuable products. CRISPR-CAS9, known as molecular scissors, alters an organisms' genome by a CAS9-based ribonuclease complex that is guided to a gene target by a programmable "guide" RNA called an sgRNA. CAS9 engineering of marine diatoms have recently been demonstrated and could potentially enhance the synthesis of naturally occurring high-valued metabolites such as fatty acids, pigments, and lipids for biodiesel production. In the Allen laboratory, we've been optimizing CRISPR-CAS9 genome engineering for multiple applications in the model marine diatom *Phaeodactylum tricorutum*. Primarily, CRISPR-CAS9 has been developed to be introduced to *Phaeodactylum* by bacterial-mediated conjugation DNA transformation. Targeting the *urease* gene resulted in 55% targeting efficiency compared to 17% targeting efficiency observed when using biolistics-based CRISPR-CAS9 delivery. However, a growth assay screen for *urease*, where mutants die when grown with urea, was crucial for screening transformed colonies prior to evaluating the targeted genomic loci. Nevertheless, a majority of *Phaeodactylum* genes do not have an associated phenotype to screen for. Also, conjugation here resulted in a loss of the CAS9 coding sequence in 90% of transformed *Phaeodactylum* which severely impacts the rate of producing false positive cell lines that do not contain a functional CAS9 enzyme. Recent efforts towards improving the bacterial-mediated conjugation methodology for introducing the CRISPR-CAS9 system have been made to boost CAS9 targeting efficiency. To ensure high-CAS9 expression in all conjugated *Phaeodactylum*, a cell line was produced harboring a CAS9 expression cassette in the genome. In order to continually track CAS9 expression, a fused eYFP was expressed with CAS9. Flow cytometry was used to quantify CAS9-eYFP expression for multiple cell lines. Two commonly used promoters, FcpB and pNR, were used to constitutively express CAS9 and induce CAS9 expression, respectively. A CAS9 cell line was chosen for high eYFP abundance within the culture and comparable growth to a wild-type *Phaeodactylum*. Two out of three lines could grow comparably to wild-type and expression of CAS9 widely varied between all picked lines, which indicates that CAS9 delivery via biolistics can impact CAS9 expression and cell physiology in unpredictable ways. Using a CAS9-eYFP *Phaeodactylum* cell line, sgRNAs will be introduced on a replicating episome. An episome has been constructed conferring resistance to the antibiotic *nourseothricin* and containing an sgRNA expression cassette. The episome has also been designed to allow the expression of multiple sgRNAs by harnessing a hierarchical Golden-Gate cloning scheme. The combination of a CAS9 cell line and bacterial conjugation delivery of sgRNAs will be used for the targeting of individual gene and also multiplexed targeting of multiple genes simultaneously. Lastly, the sgRNA episome has also been optimized for the introduction of a library of sgRNAs. Genome-wide mutagenesis of all coding gene or a "cherry-picked" set of genes will be performed using the sgRNA episome and a CAS9 *Phaeodactylum* cell line. Targeted mutagenesis of a massive number of genes in parallel will greatly accelerated

the identification and characterization of genes involved in the production of high-valued metabolites in *Phaeodactylum tricornutum*.

Funding Statement

Research supported by United States Department of Energy Genomics Science program grants DE-SC0008593 and DE-SC0018344.