

# Data and analysis code used to experimentally evolve representatives of four phytoplankton functional types in co-culture with a heterotrophic bacterium under either present-day or predicted future pCO<sub>2</sub> conditions

**Website:** <https://www.bco-dmo.org/dataset/925872>

**Data Type:** experimental

**Version:** 1

**Version Date:** 2024-04-25

## Project

- » [Impacts of Evolution on the Response of Phytoplankton Populations to Rising CO<sub>2</sub>](#) (P-ExpEv)
- » [Collaborative Research: Ecology and Evolution of Microbial Interactions in a Changing Ocean](#) (LTPE)

## Program

- » [Science, Engineering and Education for Sustainability NSF-Wide Investment \(SEES\): Ocean Acidification \(formerly CRI-OA\)](#) (SEES-OA)

Contributors	Affiliation	Role
<a href="#">Morris, James Jeffrey</a>	University of Alabama at Birmingham (UA/Birmingham)	Principal Investigator
<a href="#">Entwistle, Elizabeth</a>	University of Alabama at Birmingham (UA/Birmingham)	Scientist
<a href="#">Lu, Zhiying</a>	University of Alabama at Birmingham (UA/Birmingham)	Scientist
<a href="#">Kuhl, Matthew</a>	University of Alabama at Birmingham (UA/Birmingham)	Student
<a href="#">Durrant, Alexander</a>	University of Alabama at Birmingham (UA/Birmingham)	Technician
<a href="#">Rauch, Shannon</a>	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

## Abstract

The CO<sub>2</sub> content of Earth's atmosphere is rapidly increasing due to human consumption of fossil fuels. Models based on short-term culture experiments predict that major changes will occur in marine phytoplankton communities in the future ocean, but these models rarely consider how the evolutionary potential of phytoplankton or interactions within marine microbial communities may influence these changes. Here we experimentally evolved representatives of four phytoplankton functional types (silicifiers, calcifiers, coastal cyanobacteria, and oligotrophic cyanobacteria) in co-culture with a heterotrophic bacterium, *Alteromonas*, under either present-day or predicted future pCO<sub>2</sub> conditions. The data and analysis code in this dataset show that the genomes of all four phytoplankton as well as *Alteromonas* evolved over the course of the experiment. Mutations in oxidative stress related genes (PTOX and thioredoxin reductase) were ubiquitous in evolved cultures of *Prochlorococcus*, suggesting adaptation in response to the well-studied deficiencies of this genus in terms of stress resistance in culture. With the exception of *Prochlorococcus*, most phytoplankton genomes appeared to experience mostly purifying selection, but *Alteromonas* genomes showed strong evidence of directional selection, particularly in co-culture with eukaryotic phytoplankton. Metabolic pathways were under intense selection for *Alteromonas*, and in particular adaptation to co-culture with eukaryotes appeared to select for a shift from growth on organic acids using an abbreviated TCA cycle to growth on more complex substrates using the complete TCA cycle. This work provides new insights on how phytoplankton will respond to anthropogenic change and on the evolutionary mechanisms governing the structure and function of marine microbial communities.

## Table of Contents

- [Coverage](#)
- [Dataset Description](#)
  - [Methods & Sampling](#)
  - [Data Processing Description](#)
- [Data Files](#)
- [Supplemental Files](#)
- [Related Publications](#)

- [Related Datasets](#)
  - [Parameters](#)
  - [Instruments](#)
  - [Project Information](#)
  - [Program Information](#)
  - [Funding](#)
- 

## Coverage

**Location:** Laboratories at the University of Alabama at Birmingham

**Temporal Extent:** 2013-08-01 - 2024-01-31

## Methods & Sampling

**Experimental evolution:** Cultures for experimental evolution were initiated with 12.3 milliliters (mL) of media, 0.2 mL of acid or base additions for carbonate chemistry manipulation, and 0.5 mL of a previous culture. All cultures were grown at 22 degrees Celsius (°C) under approximately 75 micromoles photons per square meter per second ( $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ ) in acid-washed conical-bottom glass tubes with airtight caps; with 13 mL of culture, almost no headspace existed in these tubes. All cultures except *Prochlorococcus* were grown on a rotating test tube wheel with illumination from the top and bottom; preliminary observations indicated that *Prochlorococcus* cells did not settle noticeably when grown in static test tube racks, whereas all other taxa did. Each phytoplankton clone was split into two culture lines, one maintained at 400 ppm pCO<sub>2</sub> and the other at 800 ppm pCO<sub>2</sub>. For all strains except *Synechocystis*, each clone was co-cultured with a single *Alteromonas* clone.

Phytoplankton growth was measured every 48 hours using a Guava HT1 flow cytometer equipped with a 488 nanometer (nm) laser. Phytoplankton populations were identified by their clustering pattern on logarithmic plots of forward light scatter vs. 660 nm (chlorophyll) fluorescence. Cell densities within user-defined gates encircling the phytoplankton were calculated automatically by the Guava software. When phytoplankton cell densities crossed a cutoff value, cultures were diluted 26-fold (0.5 mL into a total volume of 13 mL) into fresh media. We targeted 108 transfers for each lineage, representing log<sub>2</sub>26 or 4.7 generations (although we did not achieve this goal for some lineages due to repeated crashes or contamination). Samples from each lineage were cryopreserved every 25 generations and again at the end of the experiment.

**Whole genome re-sequencing:** Post-evolution cultures were split into 5 replicate 13 mL culture tubes, grown to the cutoff transfer cell density, and then collected by gentle vacuum filtration on 0.2 micrometer ( $\mu\text{m}$ ) pore size polycarbonate filters, then flash frozen in liquid nitrogen and stored at -80°C. Genomic DNA was extracted from the filters using MoBio ProSoil kits, with the bead-beating step accomplished using a MP FastPrep-24 homogenizer. DNA was fragmented, ligated with Illumina adapters, and sequenced on an Illumina NextSeq500 device.

## Data Processing Description

FastQ files (NCBI BioSamples SAMN34542194 through SAMN34542251) were analyzed using breseq with default settings. *Synechocystis* PCC6803 cultures were assembled against the chromosomal reference genome (NCBI accession number NC\_000911) as well as its four plasmids (NC\_005229, NC\_005230, NC\_005231, and NC\_005232). *Prochlorococcus* MIT9312 and *Synechococcus* CC9311 were assembled against their RefSeq genomes (respectively, CP000111 and NC\_008319). The *T. oceanica* CCMP1005 and *E. huxleyi* CCMP371 reference genomes were still in draft form and were assembled in breseq with the contig flag activated; all contigs were downloaded as a single GenBank format file from NCBI BioProject PRJNA36595 (CCMP1005) or BioSample SRX112492 (CCMP371, also known as strain 12-1). The *T. oceanica* CCMP1005 sequence file contained the *T. oceanica* chloroplast sequence but *E. huxleyi* CCMP371 did not; we therefore used the chloroplast sequence of another *E. huxleyi* strain, CCMP373, also with downstream curation, to analyze mutational changes to those sequences. For both *T. oceanica* and *E. huxleyi*, breseq was set to not attempt to predict structural variants due to the much greater computational demand required to assemble these large genomes relative to the bacteria. *Alteromonas* genomes were assembled against the most recent EZ55 genome (40), including both the chromosomal (CABDXN010000001) and plasmid (CABDXN010000002) sequences in the same GenBank format file. For *Prochlorococcus* and *Synechococcus* lineages, EZ55 genomes were assembled in the same breseq run as the phytoplankton genome. However, because of the

computational demand of assembling *T. oceanica* and *E. huxleyi* genomes, the *Alteromonas* portion of these cultures was assembled separately to allow for structural variant prediction. All reference genomes are included with this archive in the format used for the analysis.

Any specific mutational call present in 100% of replicate lineages for a given organism was considered to have occurred prior to the initiation of the experiment (i.e., it was ancestral) and was removed from further consideration. Because of additional genomic complexity in the eukaryotic genomes of *T. oceanica* and *E. huxleyi*, we undertook additional efforts to curate these datasets. First, coverage for the eukaryote genomes was lower than for the bacterial genomes, and some loci had insufficient coverage to make predictions about mutations; all these loci were removed from breseq output files prior to further analysis. Also, because the GenBank files we used as reference sequences represented haploid sequences, we sought to discover potentially heterozygous loci in the ancestral genomes. To do this, we re-assembled the raw read files from the original genomic sequencing runs for both organisms against the GenBank reference sequences using breseq (SRX112492 for *E. huxleyi* and all sequencing runs within BioProject PRJNA36595 for *T. oceanica*). We identified several variants present in 100% of reads compared to the published GenBank sequences, perhaps suggesting differences between breseq's handling of data and the programs used to produce the published sequences; all mutations assigned by breseq to these loci were therefore removed from consideration as untrustworthy. All variants present in  $0% < n < 100%$  of sequences in the reference genome reassemblies were assumed to represent loci that were heterozygous in the ancestral population. Mutations mapped to these loci were only considered further if they became fixed in evolved lineages; if they remained at intermediate frequencies they were removed from subsequent analyses.

**Mutational analysis:** We used the application gdttools from the breseq package to convert breseq output genome difference files into long-format data files for subsequent analysis; these are provided in .tsv format with this archive. Custom python scripts were used to bin all mutations within a given gene in each lineage, producing mutational count tables of genes vs. lineages, also provided with this archive. For bacterial genomes, we produced count tables either with or without considering insertions, deletions, and intergenic mutations within 50 bp upstream of a gene's start codon (i.e., putative cis-acting promoter or other regulatory elements) in addition to mutations within a gene's coding region. Eukaryotic count tables only considered coding sequences; all intergenic and intron regions were removed from analysis. We used python scripts to analyze three metrics of directional selection: dN/dS (56), transition:transversion ratio (57), and nonsense mutation ratio (58). All three metrics exclusively used SNP data from coding regions of annotated genes. Effects of treatment groups on these metrics were analyzed using linear models in R.

We sought to test for convergent evolution of gene targets by determining which, if any, genes received more mutations than would be expected by chance under a completely random mutational model. This analysis was complicated by the fact that larger genes represented larger mutational targets than smaller ones, precluding the use of a simple Poisson model. Instead, we used a bootstrapped Monte Carlo procedure to produce randomized genomes with the same number of coding sequence mutations observed in our real dataset, only distributed randomly across the genome. Given the total coding genome size  $g$  as the sum of the lengths of all coding sequences in the genome, and the total number of observed mutations  $n$ , the probability  $p$  of any given base pair receiving a mutation is  $p = n/g$ , and the probability  $l$  of a gene of length  $l$  receiving a mutation is thus  $l = p \times l$ . For each lineage, we produced 100 randomized matrices where each gene in the genome was assigned several mutations drawn from a Poisson distribution with average rate of occurrence  $l$ . Each random matrix was compared to the real matrix of mutations by first converting each into an empirical cumulative distribution function (ECDF) and then applying either a Kolmogorov-Smirnov or dts test. The Kolmogorov-Smirnov (KS) test only considers the single value in the ECDF where the gap between the two samples is greatest, whereas the dts test considers the entire distribution. Because the greatest difference between our observed and Monte Carlo matrices was the fact that the real dataset had many more genes with large numbers of mutations than were ever observed in any simulation, the dts test was generally more sensitive in detecting significant differences. These procedures were conducted separately for nonsynonymous and synonymous mutations and revealed that the distribution of mutations was significantly different from random expectations.

Because genomes are known to include mutational hot-spots, we performed additional curation of multiply-mutated genes to find genes that were most likely targets of natural selection instead of accelerated mutational rates. We only considered genes as convergently evolved if they i) accumulated more mutations across replicate evolved lineages than any gene received in any randomized bootstrap trial or ii) were observed in at least half of all replicate evolved lineages. Additionally, reasoning that mutational hot spots should not be biased for or against silent mutations, the gene had to meet these criteria for the nonsynonymous mutation dataset but not also the synonymous mutation dataset. We applied these tests separately for each treatment group (e.g., pCO<sub>2</sub> treatments for phytoplankton genomes, or pCO<sub>2</sub> × partner for *Alteromonas* genomes). We further supported this curation step by performing a gene-by-gene linear model test in R to discover genes significantly more mutated under one treatment than another. Genes that met the curation criteria in one

treatment group only and that also demonstrated statistical enrichment in that group were considered optimal candidates for genes under natural selection for a given treatment.

Genes remaining after this curation process were then analyzed for function. First, we binned genes into KEGG pathway groups using Over Representation Analysis (ORA) with the function enrichKEGG in clusterProfiler in R. The argument 'organism' was set to the KEGG organism codes 'pmi' and 'syn' for *Prochlorococcus* and *Synechocystis*, respectively, whereas it was set to 'ko' (KEGG Orthology) for *Synechococcus*, *T. oceanica*, *E. huxleyi* and *Alteromonas*. KO identifiers to individual genes were assigned for each organism using the KEGG automatic annotation server website by the bi-directional best hit (BBH) method. P-values correspond to the comparison between the gene ratio (i.e., the number of genes that match that gene set divided by the number of genes in the 'hit' database) and the background ratio (i.e., the number of genes in the gene set divided by the total number of unique genes in the genome database). Under specific contrast comparisons for *Alteromonas* considering the mixed effect of co-cultures, KO identifiers were examined directly using the KEGG database website.

Second, the degree of association between mutating genes in EZ55 was analyzed using the Species Pairwise Association Analysis (SPAA) algorithm in R, reasoning that presence vs. absence of mutations in a given gene in a culture was mathematically comparable to presence vs. absence of species in an ecosystem. The mutation data was transformed into a binary presence vs. absence matrix, with genes having at least 1 mutation in each lineage coded as 1 and those without mutations coded as 0. The logistic regression function was used to calculate the odds ratio for mutations in each gene as described in, using cutoffs of 0.5 and 0.9 for the minimum and maximum mutation frequencies, respectively. SPAA estimated Spearman's rank correlation coefficient for each gene pair, measuring the monotonic predictive relationship, i.e., the likelihood of a mutation in one gene predicting a mutation in the other. After manually examining the data, we further filtered the pairwise predictions to consider only the most positively and negatively correlated pairs (coefficients between 0.5 to 0.9 for positive relationships and -0.3 to -0.2 for negative relationships). The curated matrix was imported in Cytoscape 3.8.2 for network visualization. The network was analyzed as a directed graph to obtain the indegree and outdegree of each node, i.e., the number of other genes to which each gene is connected as a predictive target or source, respectively. The target gene of each set of genes was illustrated as an arrow pointed towards the target gene, and the gradient of the edge connecting genes was set to reflect either a positive or negative correlation.

Finally, we manually examined several genes of interest identified from these analyses using BLAST against the NCBI nr database to attempt to provide superior annotations for hypothetical proteins or other vaguely described gene products.

**Copy number analysis:** We used R scripts to extract coverage data from all breseq-assembled genomes. Estimated copy numbers for plasmids were initially obtained simply by dividing average plasmid coverage by average chromosome coverage. However, inspection of coverage maps for the *Alteromonas* plasmid revealed highly uneven coverage. Closer analysis revealed three plasmid regions with different patterns of coverage. One region was generally present at the same or greater coverage as the chromosome, one was often completely absent, and a third often existed at an intermediate coverage level. We reasoned that these differences may reflect homologous recombination events leading to the excision of parts of the plasmid and/or insertion of plasmid regions into the *Alteromonas* chromosome and therefore looked for homologous regions between the plasmid and chromosome sequences using a dot plot analysis via the D-GENIES web interface. This analysis revealed several points of homology, including one approximately 9kb sequence, thus suggesting two things. First, a reduced size plasmid likely evolved as a subpopulation in several lineages, with approximately 50kb deleted following an unknown but reproducible event. Second, an approximately 70kb section likely integrated into the chromosome of most lineages, and in several lineages became the only surviving portion of the plasmid, with no evidence remaining of subpopulations carrying the remaining 150kb of the plasmid. Because of these trends, we calculated average coverage of the plasmid at three different regions: 170kb to 180kb to estimate the "insertable" portion of the plasmid, 120kb to 130kb to estimate the full, free plasmids, and 40kb to 50kb to represent the possible reduced-size plasmid. We tested the effects of phytoplankton partner and pCO<sub>2</sub> treatment on *Alteromonas* plasmid copy number using pairwise Mann-Whitney tests in R with manually calculated Holm-Bonferroni corrections for multiple tests. We also compared the likelihood of plasmid loss (defined as less than 1% coverage of plasmid region 120kb to 130kb) in *Alteromonas* paired with cyanobacteria vs. eukaryotic phytoplankton using Fisher's exact test in R.

Examination of coverage maps further revealed the presence of elevated copy numbers for some chromosomal regions, suggesting the presence of duplications. We therefore investigated these more closely using R scripts, retrieving any regions where the average coverage was 5× the coverage (or 5× the standard deviation of coverage for low-coverage eukaryote genomes) across the chromosome. We excluded from further analysis all duplications in eukaryotic genomes where repetitive elements led to tens of thousands of

qualifying duplications, as well as duplications falling outside of coding sequences or cis-acting promoter regions. After this process, only one duplication remained of interest: a promoter region mutation in the apolipoprotein N-acyltransferase gene in *Prochlorococcus* MIT9312, clearly visible in coverage maps and present in most lineages with up to 120× duplication.

[ [table of contents](#) | [back to top](#) ]

## Data Files

File
<b>1_Genome_Assembly.zip</b> (ZIP Archive (ZIP), 156.42 MB) MD5:8101ce620743433bd8e6604e173ba2fb Contains the code used to assemble the raw Illumina sequencing reads against reference genomes and predict where mutations had occurred using breseq, as well as to analyze the frequency of mutations with different proportional abundances in the population. The breseq output was curated as described in these files and the final mutation databases are stored in the folder "Mutations". Code in this folder was used to produce Figures S7-S9. Refer to file "ReadMe.GenomeAssembly.txt" within this folder for more information.
<b>2_Mutation_Analysis.zip</b> (ZIP Archive (ZIP), 748.44 KB) MD5:f74096f0ccff4b68092f24fc3f389fe7 Contains the code used to do coverage analysis, including detection of large duplications and possible regions of recombination. Code from this folder was used to produce. It also contains the code used to do the primary analyses on the mutation data, including i) mutation type distribution, ii) statistical analysis of the impact of pCO <sub>2</sub> and partner treatments on mutations, iii) analysis of metrics of natural selection, and iv) statistical analysis of deviation of mutation counts from neutral predictions. Code from this folder was used to produce main text figure 3 as well as figures S1-S6, S10-S13, S16, S18, and S20-S21. Refer to file "ReadMe.mutations.txt" within this folder for more information.
<b>3_ORA_Analysis.zip</b> (ZIP Archive (ZIP), 263.05 KB) MD5:171dfa6b201b25c044079260964e32eb Contains the code used to the over-representation analysis of those genes identified during the analyses in Folder 3 as being mutated more often than expected by chance. Code from this folder was used to produce figures S14, S15, S17, and S19. Refer to file "ReadMe.ORA.txt" within this folder for more information.
<b>4_SPAA_Analysis.zip</b> (ZIP Archive (ZIP), 13.21 KB) MD5:70963b09972a5c518e0a77d6b3bababd Contains the code used to do the SPAA network analysis of those genes in <i>Alteromonas</i> EZ55 identified during the analyses in Folder 3 as being mutated more often than expected by chance. Code from this folder was used to produce main text figure 5. A cytoscape-formatted XML file is also included. Refer to file "ReadMe.SPAA.txt" within this folder for more information.
<b>Mutations.tar.gz</b> (ZIP Archive (ZIP), 377.14 MB) MD5:8fae612fc2455a63ab43e31ee4289463 Contains the final mutation databases.

[ [table of contents](#) | [back to top](#) ]

## Supplemental Files

File
<b>ReadMe.genomics.txt</b> (Plain Text, 2.17 KB) MD5:d918d2a304e32435d86665b7f9d4895d This document gives an overview of the code and data archives necessary to replicate the analysis of all genomic data presented in the manuscript "Marine phytoplankton and heterotrophic bacteria rapidly adapt to future pCO <sub>2</sub> conditions in experimental co-cultures" by Zhiying Lu et al.

[ [table of contents](#) | [back to top](#) ]

---

## Related Publications

Lu, Z., Entwistle, E., Kuhl, M. D., Durrant, A. R., Filho, M. M. B., Goswami, A., & Morris, J. J. (2024). Marine phytoplankton and heterotrophic bacteria rapidly adapt to future pCO<sub>2</sub> conditions in experimental co-cultures. <https://doi.org/10.1101/2024.02.07.579367>  
*Results*

[ [table of contents](#) | [back to top](#) ]

---

## Related Datasets

### IsRelatedTo

Morris, J. J., Entwistle, E., Lu, Z. (2024) **Data and code from an examination of growth rates of cyanobacteria co-cultured with a heterotrophic bacterium, *Alteromonas*, under either present-day or predicted future pCO<sub>2</sub> conditions.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2024-04-24 doi:10.26008/1912/bco-dmo.925841.1 [[view at BCO-DMO](#)]

[ [table of contents](#) | [back to top](#) ]

---

## Parameters

*Parameters for this dataset have not yet been identified*

[ [table of contents](#) | [back to top](#) ]

---

## Instruments

<b>Dataset-specific Instrument Name</b>	NextSeq500 (Illumina)
<b>Generic Instrument Name</b>	Automated DNA Sequencer
<b>Generic Instrument Description</b>	General term for a laboratory instrument used for deciphering the order of bases in a strand of DNA. Sanger sequencers detect fluorescence from different dyes that are used to identify the A, C, G, and T extension reactions. Contemporary or Pyrosequencer methods are based on detecting the activity of DNA polymerase (a DNA synthesizing enzyme) with another chemoluminescent enzyme. Essentially, the method allows sequencing of a single strand of DNA by synthesizing the complementary strand along it, one base pair at a time, and detecting which base was actually added at each step.

<b>Dataset-specific Instrument Name</b>	Guava HT1 flow cytometer (Luminex)
<b>Generic Instrument Name</b>	Flow Cytometer
<b>Generic Instrument Description</b>	Flow cytometers (FC or FCM) are automated instruments that quantitate properties of single cells, one cell at a time. They can measure cell size, cell granularity, the amounts of cell components such as total DNA, newly synthesized DNA, gene expression as the amount messenger RNA for a particular gene, amounts of specific surface receptors, amounts of intracellular proteins, or transient signalling events in living cells. (from: <a href="http://www.bio.umass.edu/micro/immunology/facs542/facswhat.htm">http://www.bio.umass.edu/micro/immunology/facs542/facswhat.htm</a> )

<b>Dataset-specific Instrument Name</b>	NextSeq500 (Illumina)
<b>Generic Instrument Name</b>	Homogenizer
<b>Generic Instrument Description</b>	A homogenizer is a piece of laboratory equipment used for the homogenization of various types of material, such as tissue, plant, food, soil, and many others.

[ [table of contents](#) | [back to top](#) ]

## Project Information

### Impacts of Evolution on the Response of Phytoplankton Populations to Rising CO<sub>2</sub> (P-ExpEv)

**Coverage:** Experiment housed in laboratories at Michigan State University

Note: This project is also affiliated with the [NSF BEACON Center for the Study of Evolution in Action](#).

#### *Project Description from NSF Award:*

Human activities are driving up atmospheric carbon dioxide concentrations at an unprecedented rate, perturbing the ocean's carbonate buffering system, lowering oceanic pH, and changing the concentration and composition of dissolved inorganic carbon. Recent studies have shown that this ocean acidification has many short-term effects on phytoplankton, including changes in carbon fixation among others. These physiological changes could have profound effects on phytoplankton metabolism and community structure, with concomitant effects on Earth's carbon cycle and, hence, global climate. However, extrapolation of present understanding to the field are complicated by the possibility that natural populations might evolve in response to their changing environments, leading to different outcomes than those predicted from short-term studies. Indeed, evolution experiments demonstrate that microbes are often able to rapidly adapt to changes in the environment, and that beneficial mutations are capable of sweeping large populations on time scales relevant to predictions of environmental dynamics in the coming decades. This project addresses two major areas of uncertainty for phytoplankton populations with the following questions:

- 1) What adaptive mutations to elevated CO<sub>2</sub> are easily accessible to extant species, how often do they arise, and how large are their effects on fitness?
- 2) How will physical and ecological interactions affect the expansion of those mutations into standing populations?

This study will address these questions by coupling experimental evolution with computational modeling of ocean biogeochemical cycles. First, cultured unicellular phytoplankton, representative of major functional groups (e.g. cyanobacteria, diatoms, coccolithophores), will be evolved under simulated year 2100 CO<sub>2</sub> concentrations. From these experiments, estimates will be made of a) the rate of beneficial mutations, b) the magnitude of fitness gains conferred by these mutations, and c) secondary phenotypes (i.e., trade-offs) associated with these mutations, assayed using both physiological and genetic approaches. Second, an existing numerical model of the global ocean system will be modified to a) simulate the effects of changing

atmospheric CO<sub>2</sub> concentrations on ocean chemistry, and b) allow the introduction of CO<sub>2</sub>-specific adaptive mutants into the extant populations of virtual phytoplankton. The model will be used to explore the ecological and biogeochemical impacts of beneficial mutations in realistic environmental situations (e.g. resource availability, predation, etc.). Initially, the model will be applied to idealized sensitivity studies; then, as experimental results become available, the implications of the specific beneficial mutations observed in our experiments will be explored.

This interdisciplinary study will provide novel, transformative understanding of the extent to which evolutionary processes influence phytoplankton diversity, physiological ecology, and carbon cycling in the near-future ocean. One of many important outcomes will be the development and testing of nearly-neutral genetic markers useful for competition studies in major phytoplankton functional groups, which has applications well beyond the current proposal.

## **Collaborative Research: Ecology and Evolution of Microbial Interactions in a Changing Ocean (LTPE)**

**Coverage:** Lab work: Birmingham, Alabama and New York, New York. Field Work: Bermuda Atlantic Time Series.

### *NSF Award Abstract:*

Carbon dioxide released from fossil fuels is causing the ocean to become more acidic. Much attention has been given to how this will affect shelled animals like corals, but acidification also affects the algae that form the base of the ocean food chain. It is possible that future algal communities will look very different than they do today, with potentially negative consequences for fisheries, recreation, and climate. Alternatively, it is possible that these algae will be able to adapt rapidly enough to avoid the worst of it. This study looks at algae adapting to acidification in real time in the lab, focusing on "marketplace" interactions between the algae and the bacteria they live alongside. The researchers also go to sea to learn whether adaptations from the lab experiments are beneficial under real-world conditions. Ultimately, this project is helping scientists better understand how the ocean's most important and most overlooked organisms will respond to the changes humans are causing in their habitat. The researchers also use their scientific work to create fun educational opportunities from grade school to college, including agar art classes where students learn about microbial ecology by "painting" with freshly-isolated ocean bacteria.

The effect of ocean acidification on calcifying organisms has been well-studied, but less is known about how changing pH will affect phytoplankton. Previous work showed that the mutualistic interaction between the globally abundant cyanobacterium *Prochlorococcus* and its "helper" bacterium *Alteromonas* broke down under projected future CO<sub>2</sub> conditions, leading to a strong decrease in the fitness of *Prochlorococcus*. It is possible that such interspecies interactions between microbes are important for many ecological processes, but a lack of understanding of how these interactions evolve makes it difficult to predict how important they are. This project is using laboratory evolution experiments to discover how evolution shapes the interactions between bacteria and algae like *Prochlorococcus*, and how these co-evolutionary dynamics might influence the biogeochemical processes that shape Earth's climate. Four research cruises to the Bermuda Atlantic Time Series are also planned to study how natural algal/bacterial communities respond to acidification, and whether evolved microbes from laboratory experiments have a competitive advantage in complex, natural communities exposed to elevated CO<sub>2</sub>. The ultimate goal of this project is to gain a mechanistic understanding of microbial interactions that can be used to inform models of Earth's oceans and biological feedbacks on global climate.

[ [table of contents](#) | [back to top](#) ]

---

## **Program Information**

**Science, Engineering and Education for Sustainability NSF-Wide Investment (SEES): Ocean Acidification (formerly CRI-OA) (SEES-OA)**

**Website:** [https://www.nsf.gov/funding/pgm\\_summ.jsp?pims\\_id=503477](https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503477)

**Coverage:** global

NSF Climate Research Investment (CRI) activities that were initiated in 2010 are now included under Science, Engineering and Education for Sustainability NSF-Wide Investment (SEES). SEES is a portfolio of activities that highlights NSF's unique role in helping society address the challenge(s) of achieving sustainability. Detailed information about the SEES program is available from NSF ([https://www.nsf.gov/funding/pgm\\_summ.jsp?pims\\_id=504707](https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504707)).

In recognition of the need for basic research concerning the nature, extent and impact of ocean acidification on oceanic environments in the past, present and future, the goal of the SEES: OA program is to understand (a) the chemistry and physical chemistry of ocean acidification; (b) how ocean acidification interacts with processes at the organismal level; and (c) how the earth system history informs our understanding of the effects of ocean acidification on the present day and future ocean.

**Solicitations issued under this program:**

[NSF 10-530](#), FY 2010-FY2011

[NSF 12-500](#), FY 2012

[NSF 12-600](#), FY 2013

[NSF 13-586](#), FY 2014

NSF 13-586 was the final solicitation that will be released for this program.

**PI Meetings:**

[1st U.S. Ocean Acidification PI Meeting](#) (March 22-24, 2011, Woods Hole, MA)

[2nd U.S. Ocean Acidification PI Meeting](#) (Sept. 18-20, 2013, Washington, DC)

3rd U.S. Ocean Acidification PI Meeting (June 9-11, 2015, Woods Hole, MA - Tentative)

**NSF media releases for the Ocean Acidification Program:**

[Press Release 10-186 NSF Awards Grants to Study Effects of Ocean Acidification](#)

[Discovery Blue Mussels "Hang On" Along Rocky Shores: For How Long?](#)

[Discovery nsf.gov - National Science Foundation \(NSF\) Discoveries - Trouble in Paradise: Ocean Acidification This Way Comes - US National Science Foundation \(NSF\)](#)

[Press Release 12-179 nsf.gov - National Science Foundation \(NSF\) News - Ocean Acidification: Finding New Answers Through National Science Foundation Research Grants - US National Science Foundation \(NSF\)](#)

[Press Release 13-102 World Oceans Month Brings Mixed News for Oysters](#)

[Press Release 13-108 nsf.gov - National Science Foundation \(NSF\) News - Natural Underwater Springs Show How Coral Reefs Respond to Ocean Acidification - US National Science Foundation \(NSF\)](#)

[Press Release 13-148 Ocean acidification: Making new discoveries through National Science Foundation research grants](#)

[Press Release 13-148 - Video nsf.gov - News - Video - NSF Ocean Sciences Division Director David Conover answers questions about ocean acidification. - US National Science Foundation \(NSF\)](#)

[Press Release 14-010 nsf.gov - National Science Foundation \(NSF\) News - Palau's coral reefs surprisingly resistant to ocean acidification - US National Science Foundation \(NSF\)](#)

[Press Release 14-116 nsf.gov - National Science Foundation \(NSF\) News - Ocean Acidification: NSF awards \\$11.4 million in new grants to study effects on marine ecosystems - US National Science Foundation \(NSF\)](#)

[ [table of contents](#) | [back to top](#) ]

---

## Funding

<b>Funding Source</b>	<b>Award</b>
<a href="#">NSF Division of Ocean Sciences (NSF OCE)</a>	<a href="#">OCE-1316101</a>
<a href="#">NSF Division of Ocean Sciences (NSF OCE)</a>	<a href="#">OCE-1540158</a>
<a href="#">NSF Division of Ocean Sciences (NSF OCE)</a>	<a href="#">OCE-1851085</a>

[ [table of contents](#) | [back to top](#) ]