Title
A Collection of Algal Genomes from the JGI

Permalink
https://escholarship.org/uc/item/1gd533x8

Authors
Kuo, Alan
Grigoriev, Igor

Publication Date
2012-03-20
A Collection of Algal Genomes from JGI

Alan Kuo\textsuperscript{1*} (akuo@lbl.gov), Igor Grigoriev\textsuperscript{1}

\textsuperscript{1}DOE Joint Genome Institute, Lawrence Berkeley National Laboratory

March 2012

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.
ABSTRACT

Algae, defined as photosynthetic eukaryotes other than plants, constitute a major component of fundamental eukaryotic diversity. Acquisition of the ability to conduct oxygenic photosynthesis through endosymbiotic events has been a principal driver of eukaryotic evolution, and today algae continue to underpin aquatic food chains as primary producers. Algae play profound roles in the carbon cycle, can impose health and economic costs through toxic blooms, and are candidate sources for bio-fuels; all of these research areas are part of the mission of DOE’s Joint Genome Institute (JGI). A collection of algal projects ongoing at JGI contributes to each of these areas and illustrates analyses employed in their genome exploration.
A Collection of Algal Genomes from JGI

Algae and the JGI mission

Algae and the environment
50% of global photosynthesis
Base of food webs
Carbon fixation and cycling
Biomass accumulation, biogeochemistry

Algae and evolution
Very diverse, all over eukaryotic tree
Photosynthesis acquired many times by many clades
Endosymbiosis is a principal driver of evolution
Plastid loss: Osmocysta, evil plant pathogens

Algae and energy
Source of biofuels
Model systems for plants
Model systems for plant-fungal symbioses (lichens)

Resequencing of 13 other strains of Ehux from around the world.

~80 projects worldwide
10 genomes published by JGI (out of 12):
Thalassiosira (2004)
Ostreococcus, 2 sp. (2007)
Chlamydomonas (2007)
Phaeodactylum (2009)
Micromonas, 2 sp. (2008)
Volvox (2010)
Emiliania (2011)
Aureococcus (2011)
9 genomes released but not yet published:
Emiliania
Coccomyxa
Ostreococcus (3rd sp)
Cryptophyta
Chlorarachniophyta
Bigelowiella
Amphidinium
Guillardia
Pseudo-nitzschia

The dominant phytoplankter *Emiliania huxleyi*

Ehux is globally distributed, locally abundant, and well adapted to a wide range of environments.

The reference genome of Ehux, 1st sequenced Haptophyta, Hacrobia

Genome size (Mbp)

<table>
<thead>
<tr>
<th>Strain</th>
<th># scaffolds</th>
<th># genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ehux</td>
<td>168</td>
<td>39128</td>
</tr>
<tr>
<td>Bnatans</td>
<td>32341</td>
<td>122kbp</td>
</tr>
<tr>
<td>Gtheta</td>
<td>27065</td>
<td>9636</td>
</tr>
</tbody>
</table>

Resequencing of 13 other strains of Ehux from around the world.

Collect 13 strains.
Sequence to 2x-4x coverage.
Assemble into contigs.
Align to reference genome.
Many of the reference genes are absent from many of the 13 strains.

*Ehux* has core and variable genomes.

~15k core genes

~50% of genome is repetitive.
~50% of genes are in multigene families.
~75% of genes have introns.

Genomes of the 2 lichen symbionts.

Bnatans Gtheta each has 4 genomes.

Gene expression changes in the lichenoids.

Enrichment or depletion of Pfams in families clustered from *Asterochloris* and 10 other species of *Chlorella*.

Credits

These collaborators provided much data and many figures:

Ehux: Betsy Read, Xiaoyu Zhang
Bnatans=Gtheta: John Archibald, Bruce Curtis, Eunsoo Kim, Manuel Irimia
Lichen: Danielle Armalao, Olaf Mueller

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.