Fueling Future with Algal Genomics

Igor V. Grigoriev
US DOE Joint Genome Institute, Walnut Creek, CA 94589

July 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

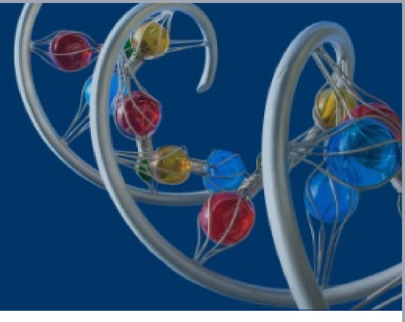
DISCLAIMER

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.

Fueling Future with Algal Genomics

Igor Grigoriev (ivgrigoriev@lbl.gov)
US DOE Joint Genome Institute, Walnut Creek, CA 94598

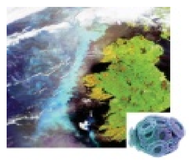
DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE



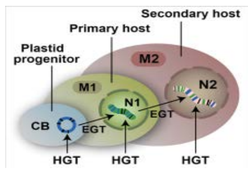
Algae for Energy and Environment

Energy
50% of global photosynthesis
Biomass for biofuels
Model systems for plants

Environment
Carbon fixation and cycling
Biomineralization, biogeochemistry
Blooms can be toxic, blooms can be HUGE



Evolution
Very diverse, all over eukaryotic tree
Photosynthesis acquired many times
Endosymbiosis is a principal driver of evolution

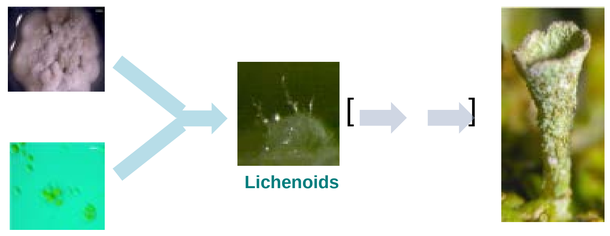


Abstract

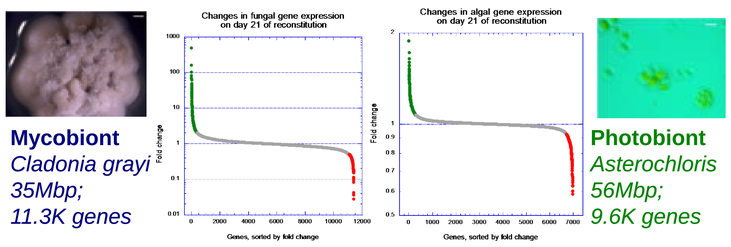
Algae constitute a major component of fundamental eukaryotic diversity, play profound roles in the carbon cycle, and are prominent candidates for biofuel production. The US Department of Energy Joint Genome Institute (JGI) is leading the world in algal genome sequencing (http://jgi.doe.gov/Algae) and contributes ¾ of the algal genome projects worldwide (GOLD database, 2012). The sequenced algal genomes offer catalogs of genes, networks, and pathways. The sequenced 'first of its kind' genomes of a haptophyte *E.huxleyii*, chlorarachniophyte *B.natans,* and cryptophyte *G.theta* fill the gaps in the eukaryotic tree of life and carry unique genes and pathways as well as 'molecular fossils' of secondary endosymbiosis. Natural adaptation to conditions critical for industrial production is encoded in algal genomes, for example, growth of *A.anophagefferens* at very high cell densities during the harmful algae blooms or a global distribution across diverse environments of *E.huxleyii,* able to live on sparse nutrients due to its expanded pan-genome. Communications and signaling pathways can be derived from simple symbiotic systems like lichens or complex marine algae metagenomes. Collectively these datasets derived from algal genomics contribute to building a comprehensive parts list essential for algal biofuel development.

Algal Symbiosis: lichen (fungus+alga)

Genomics of lichen symbiosis, and its *in vitro* reconstitution.



Differences in gene expression in the lichenoids.

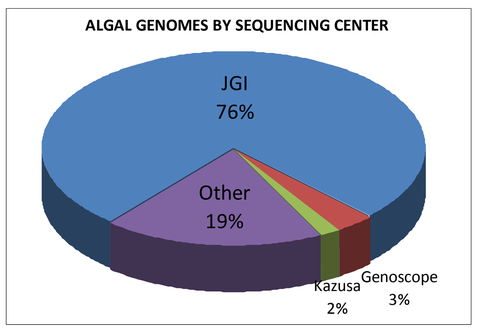


Mycobiont
Cladonia grayi
35Mbp;
11.3K genes

Photobiont
Asterochloris
56Mbp;
9.6K genes

PI: Danielle Armaleo, Duke University

Algal Genomes Sequenced by JGI



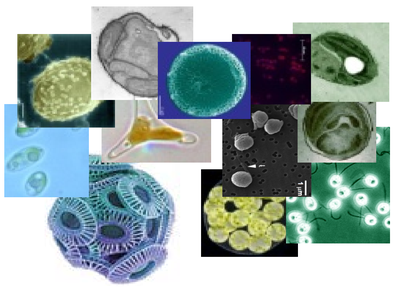
11 genomes published by JGI:

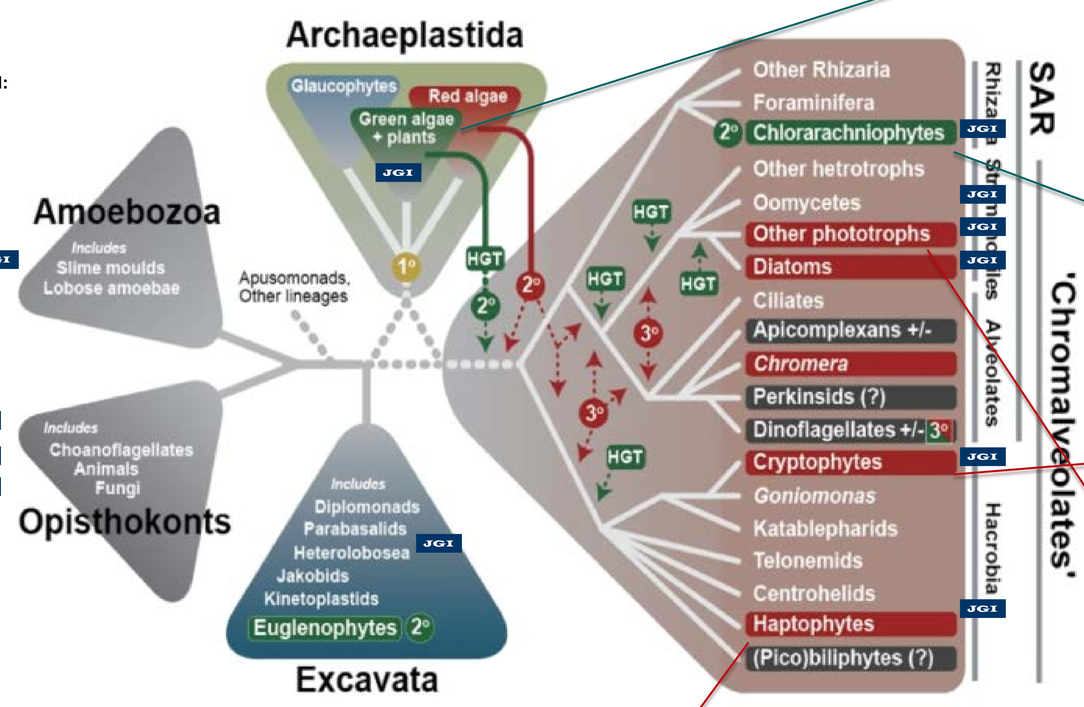
Thalassiosira (2004)
Ostreococcus, 2 sp. (2007)
Chlamydomonas (2007)
Phaeodactylum (2008)
Micromonas, 2 sp. (2009)
Volvox (2010)
Chlorella (2011) JGI
Aureococcus (2011)
Coccomyxa (2012)

7 more genomes released:

Emiliania
Ostreococcus (a 3rd sp) JGI
Fragilariopsis
Bigelowiella JGI
Asterochloris JGI
Guillardia
Pseudo-nitzschia

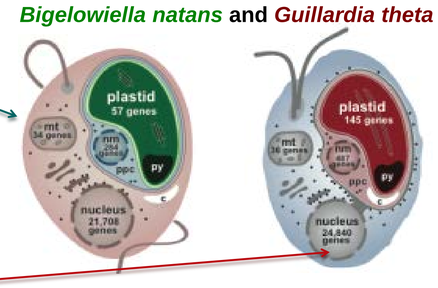
Algal projects worldwide





Snapshot of Algal Evolution: Nucleomorph-retaining algae

Bigelowiella natans* and *Guillardia theta

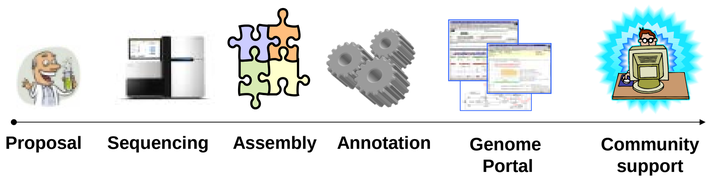


genome	size	# genes	size	# genes
nucleomorph	373 kbp	283	551 kbp	485
chloroplast	69 kbp	61	122 kbp	147
mitochondrion	37 kbp	22	N/A	N/A
nucleus	95 Mbp	21380	87 Mbp	24525

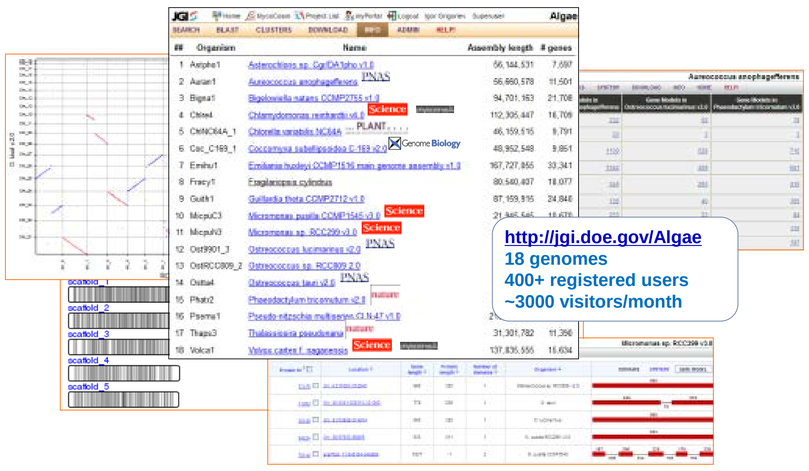
Four genomes in each algal

PI: John Archibald, Dalhousie University

JGI Genome Project Pipeline



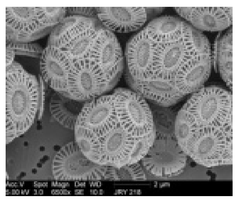
JGI Algal Genome Portal (http://jgi.doe.gov/Algae/): Comparative genomics tools



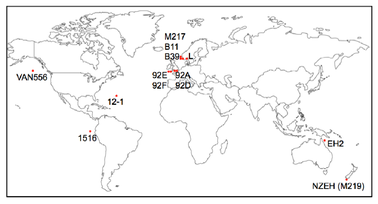
Propose your genome projects for the next JGI Community Sequencing Program Call (Spring 2013) at http://www.jgi.doe.gov/CSP/

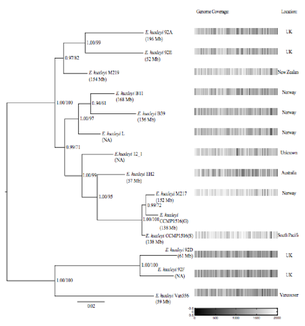
The dominant phytoplankter *Emiliania huxleyi*

Haptophyte
E.huxleyi
168Mbp;
33.3K genes



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

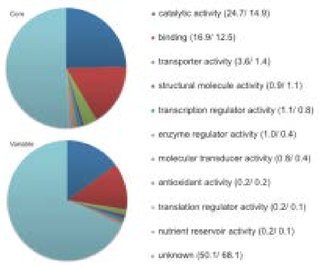


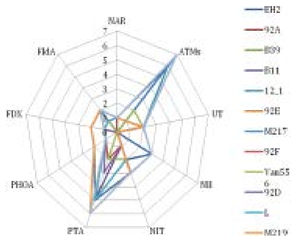


Geography and phylogeny of 14 strains of Ehux.

Ehux has core and variable genomes.

~18k core genes



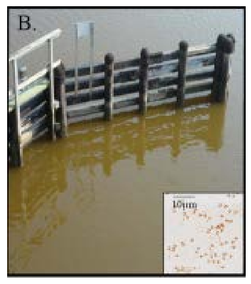


~ 10k variable genes

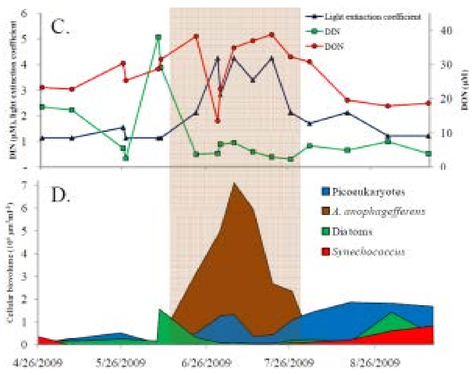
PI: Betsy Read, Cal State University

The Harmful Algal Bloom Explained

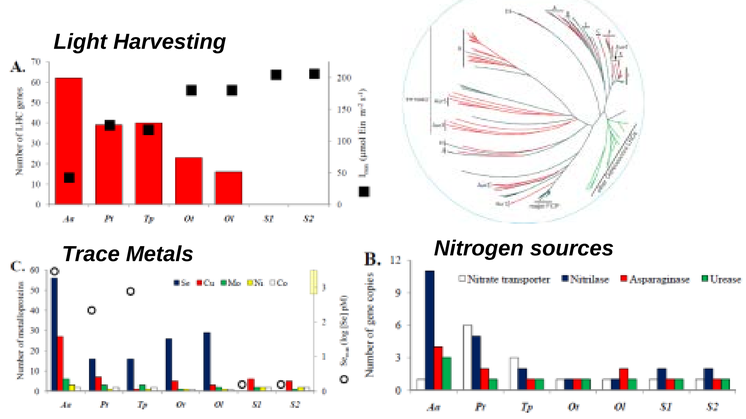
Pelagophyte
Aureococcus anophagefferens
57 Mbp;
11.5K genes



Estuary niche:
low light, high organics, trace metals



Comparative genomics: gene family expansions to fit the niche



PI: Chris Gobler, Stony-Brooks (PNAS 2011)

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.