

A Collection of Algal Genomes from JGI

Alan Kuo^{1*} (akuo@lbl.gov), Igor Grigoriev¹

¹DOE Joint Genome Institute, Lawrence Berkeley National Laboratory

March 2012

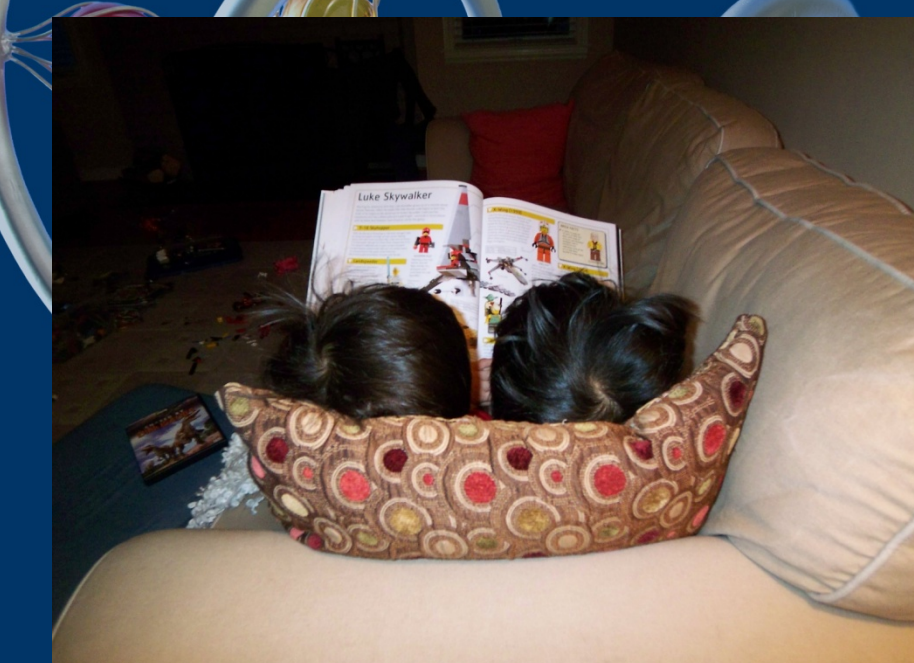
The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231

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.



Algae and the JGI mission

Algae and the environment

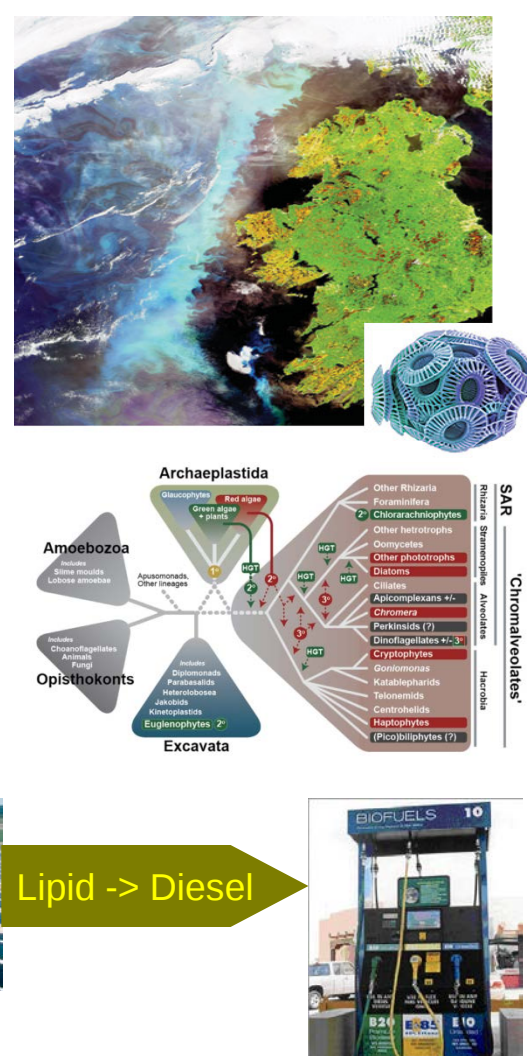
50% of global photosynthesis
Base of food webs
Carbon fixation and cycling
Biomineralization, biogeochemistry
Blooms can be toxic, blooms can be HUGE

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: Oomycota, evil plant pathogens

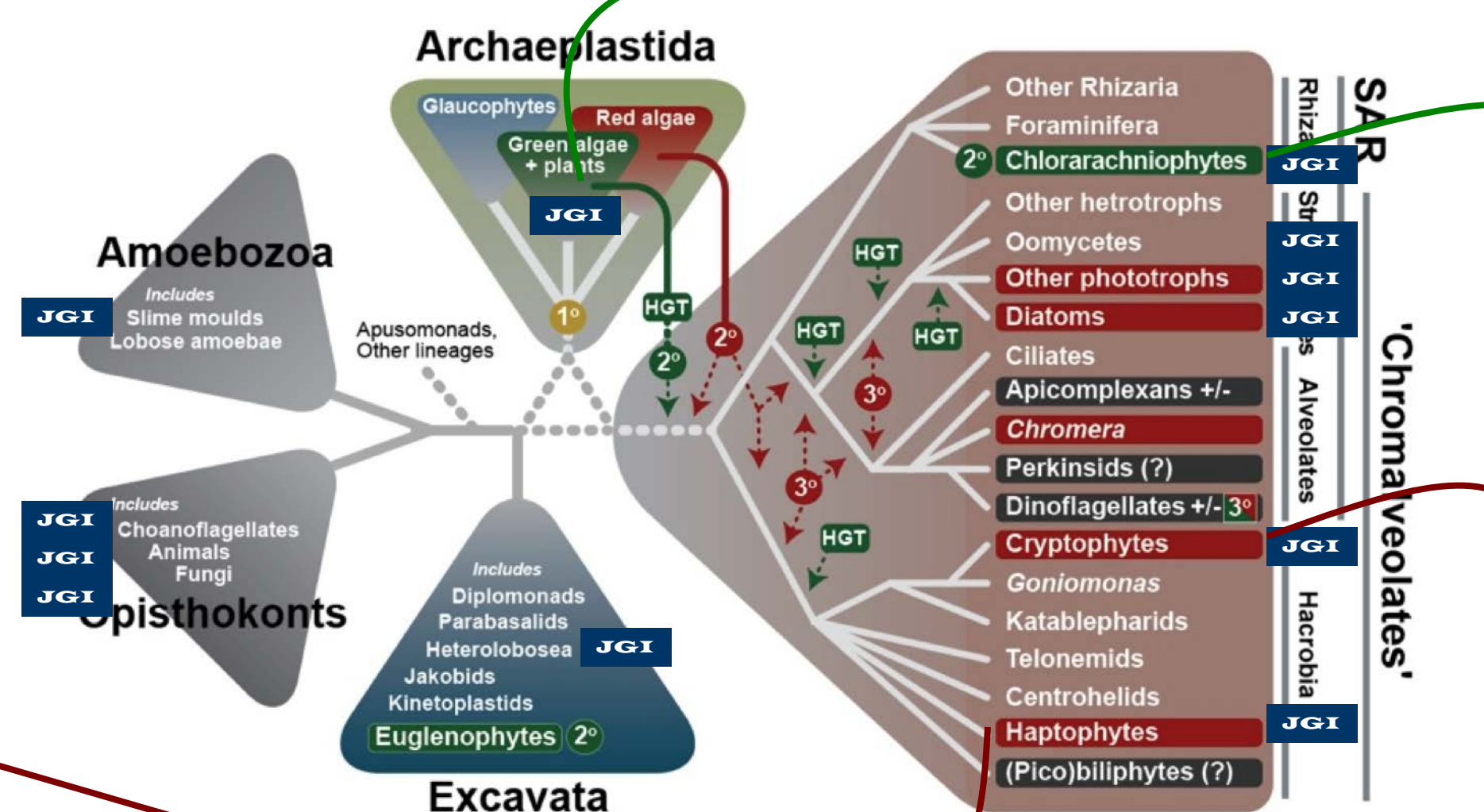
Algae and energy

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



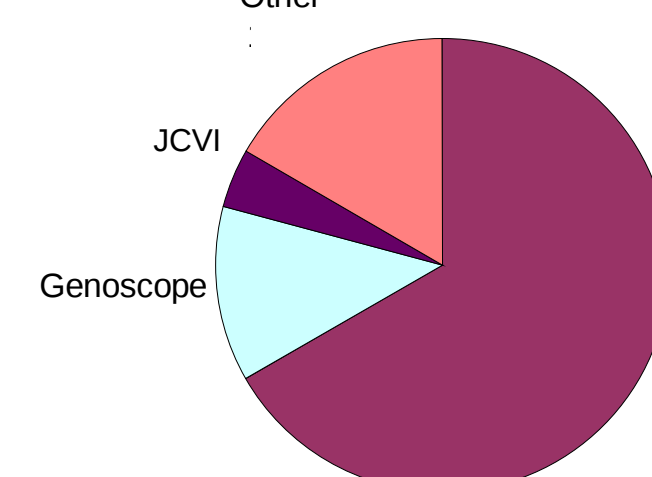
Abstract

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.



Algae sequenced by JGI

~80 projects worldwide



51 projects at JGI

10 genomes published by JGI (out of 12):

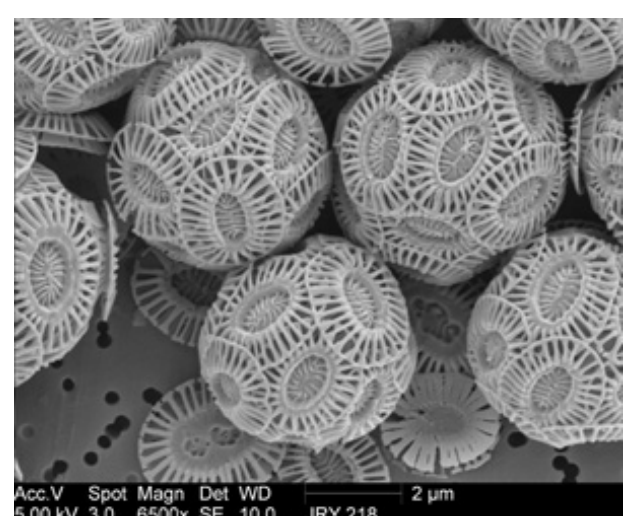
Thalassiosira (2004)
Ostreococcus, 2 sp. (2007)
Chlamydomonas (2007)
Phaeodactylum (2008)
Micromonas, 2 sp. (2009)
Volvox (2010)
Chlorella (2011)
Aureococcus (2011)

8 genomes released but not yet published:

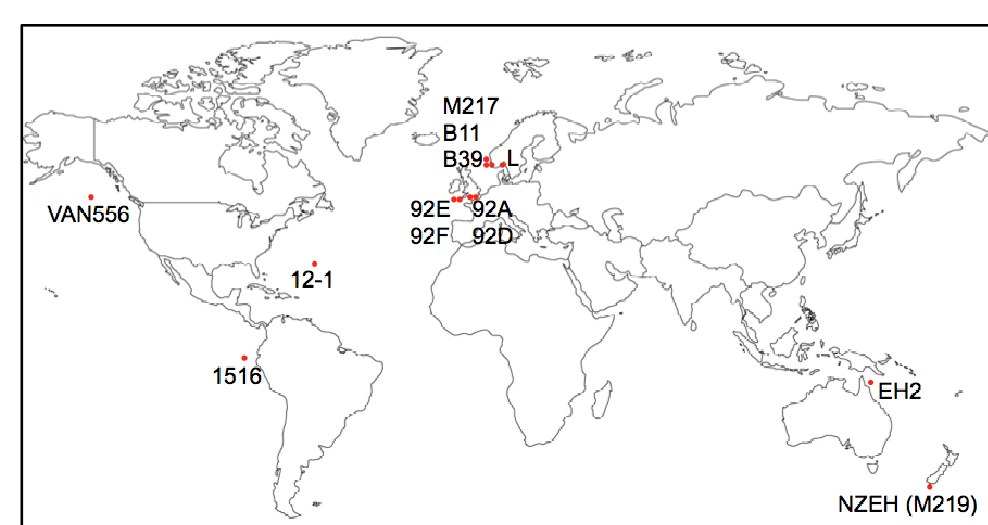
Emiliania
Coccomyxa
Ostreococcus (a 3rd sp)
Flagellariopsis
Bigeloviella
Asterochloris
Guillardia
Pseudo-nitzschia

JGI Algal Portal: <http://jgi.doe.gov/Algae/>

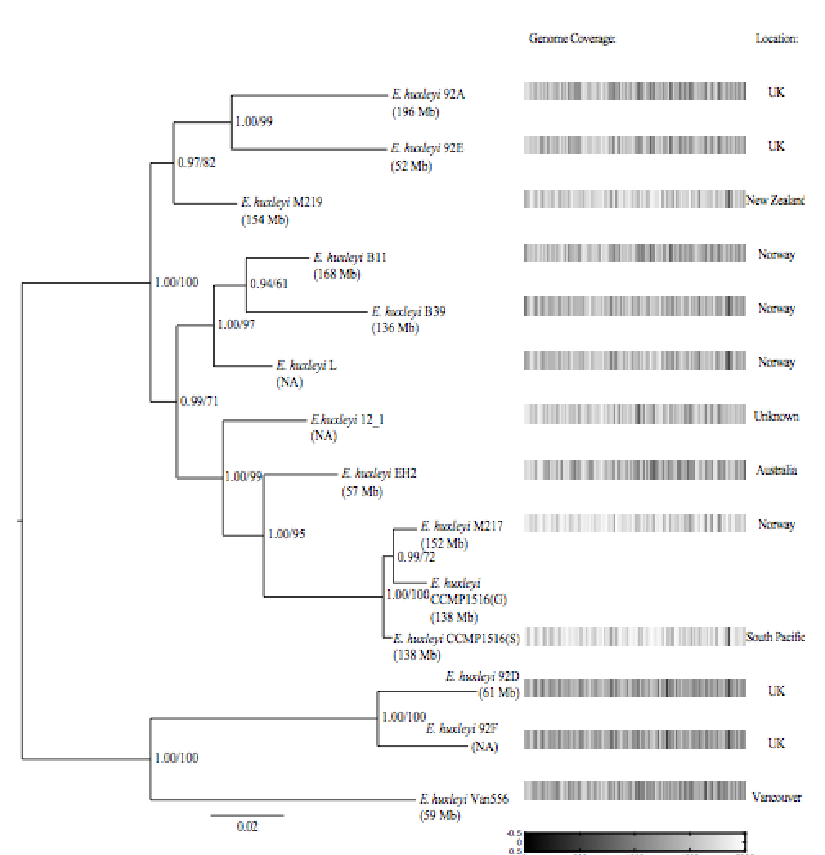
The dominant phytoplankter *Emiliania huxleyi*



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



Geography and phylogeny of 14 strains of *Ehux*.



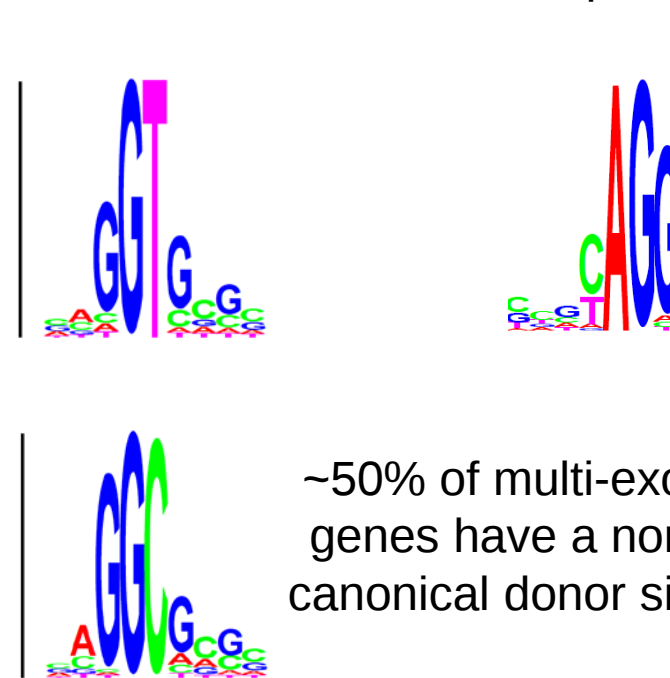
The reference genome of *Ehux*.

1st sequenced Haptophyta, Hacrobia

Genome size (Mbp)	168
# scaffolds	7809
# genes	39126
# genes (reduced)	33341

~50% of genome is repetitive.
~55% of genes are in multigene families
~75% of genes have introns.

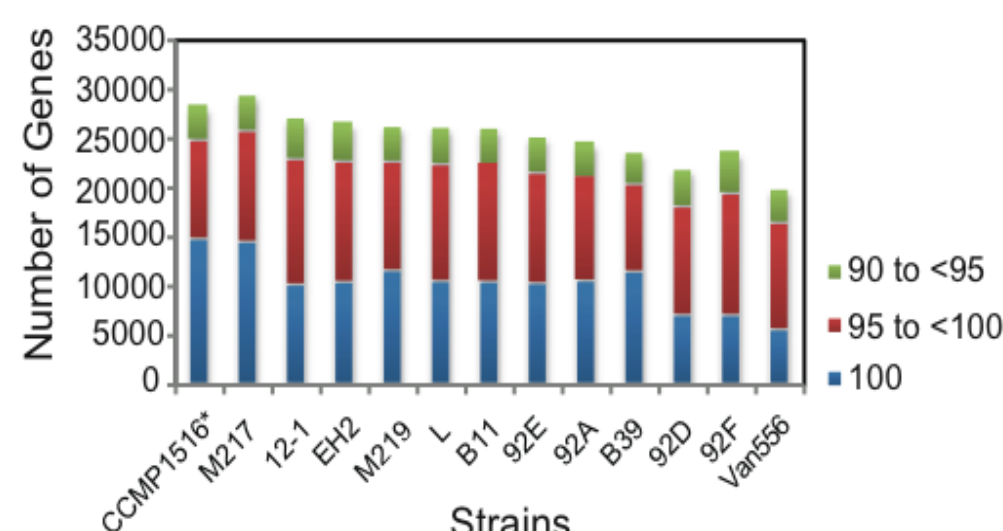
Donor Acceptor



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

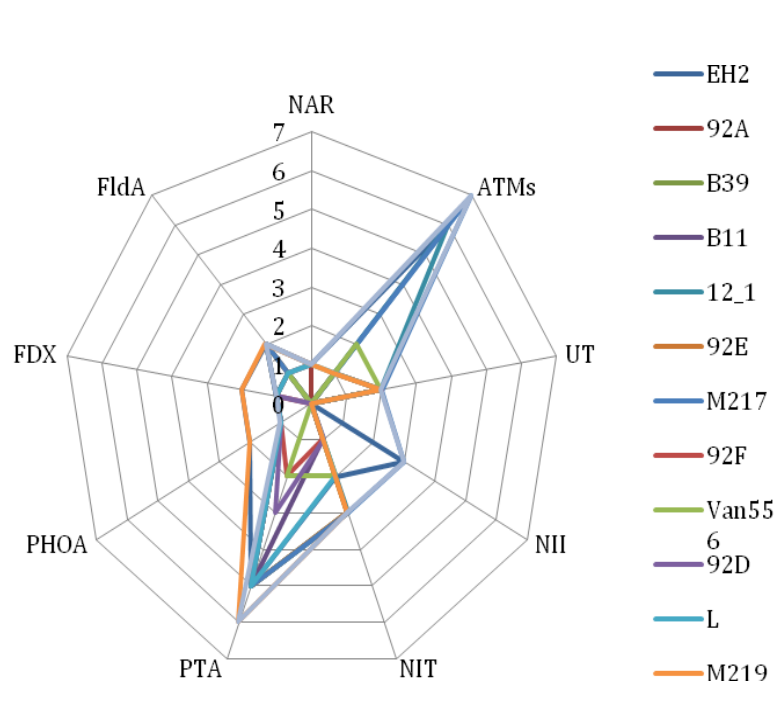
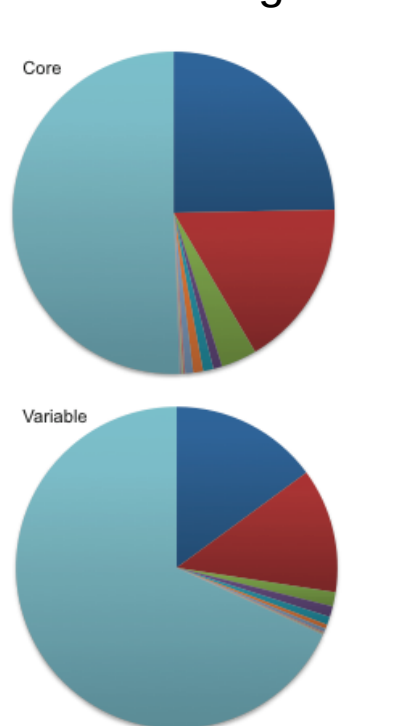
Collect 13 strains.
Sequence to 9-49x coverage.
Assemble into contigs.
Align to reference proteins.

Many of the reference genes are absent from many of the 13 strains.



Ehux has core and variable genomes.

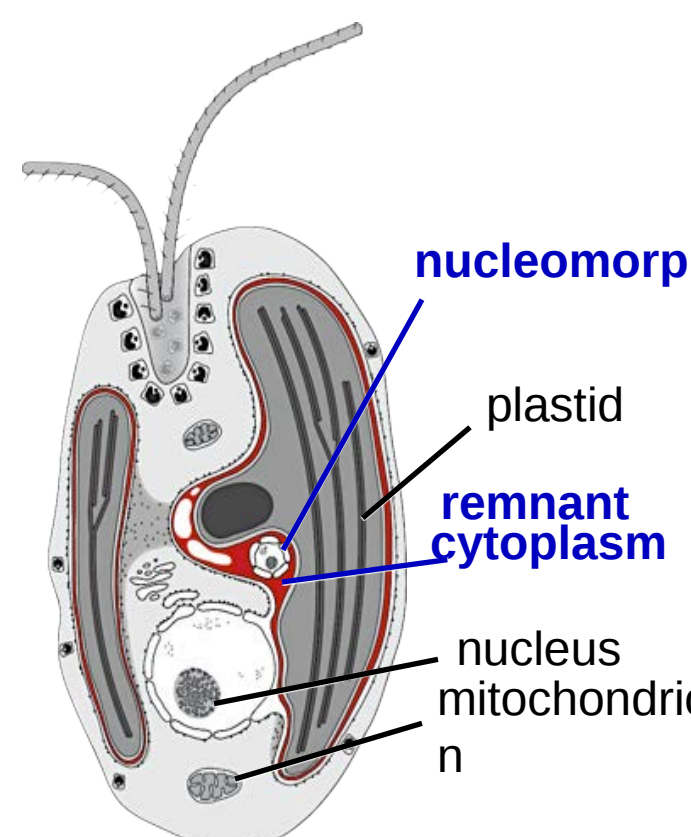
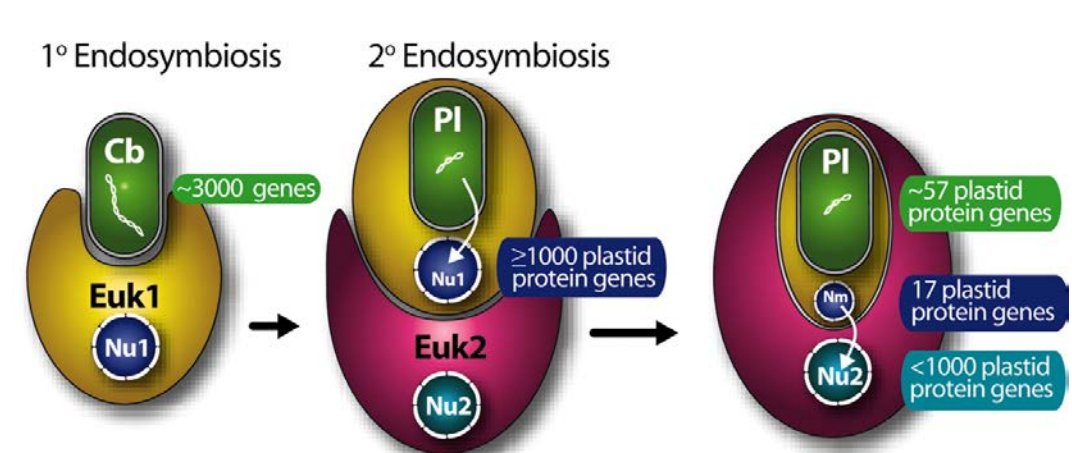
~18k core genes



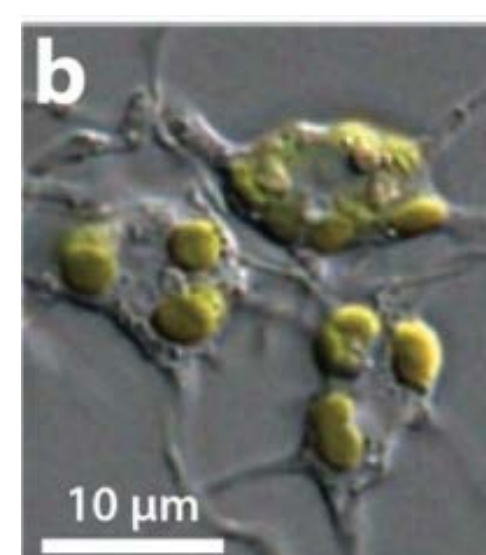
~10k variable genes

The nucleomorph-retaining algae *Bigeloviella natans* and *Guillardia theta*

Nucleomorph is the residual nucleus of an algal 2° endosymbiont



Bnatans and *Gtheta* each has 4 genomes.



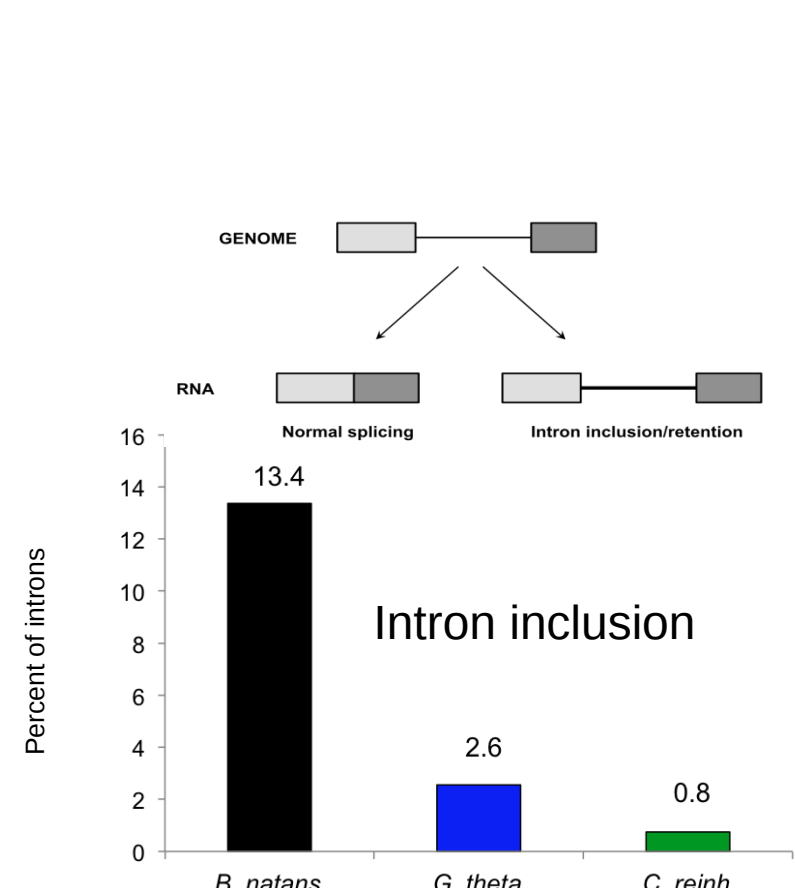
Bnatans
1st sequenced
Chlorarachniophyta,
Rhizaria.



Gtheta
1st sequenced
Cryptophyta

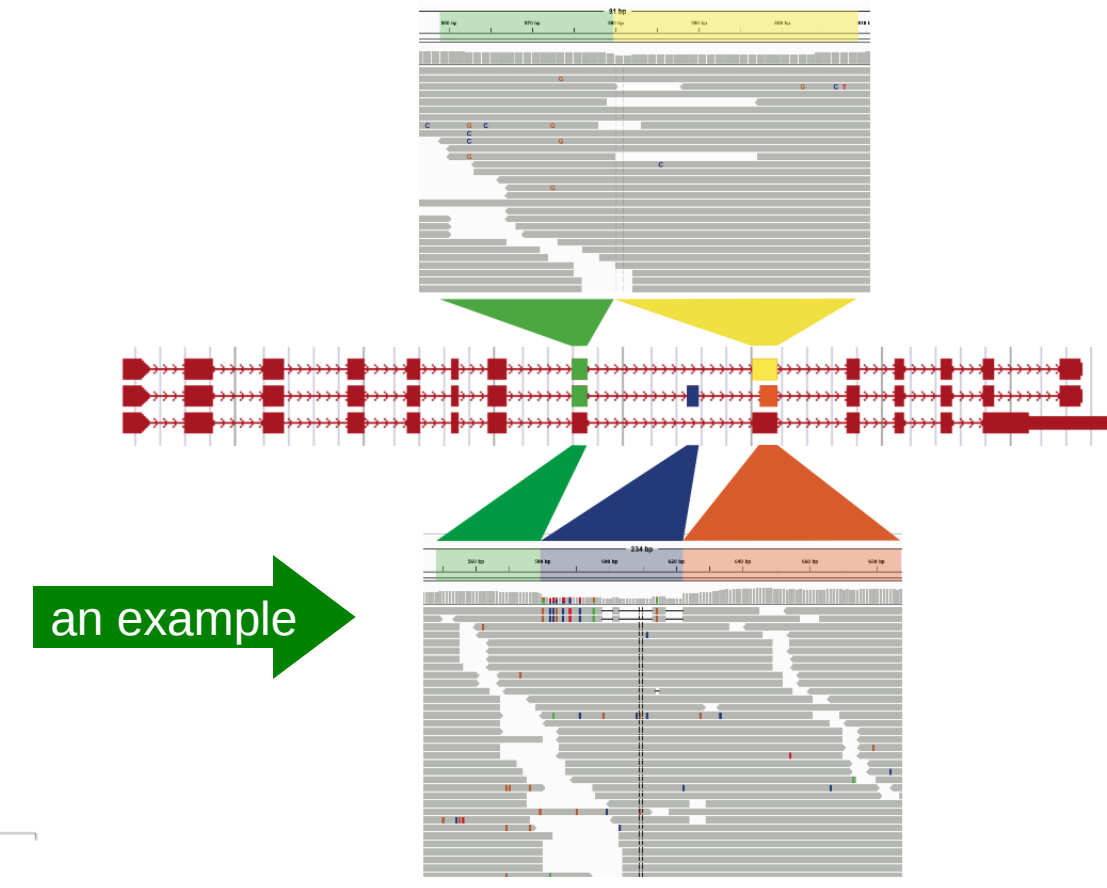
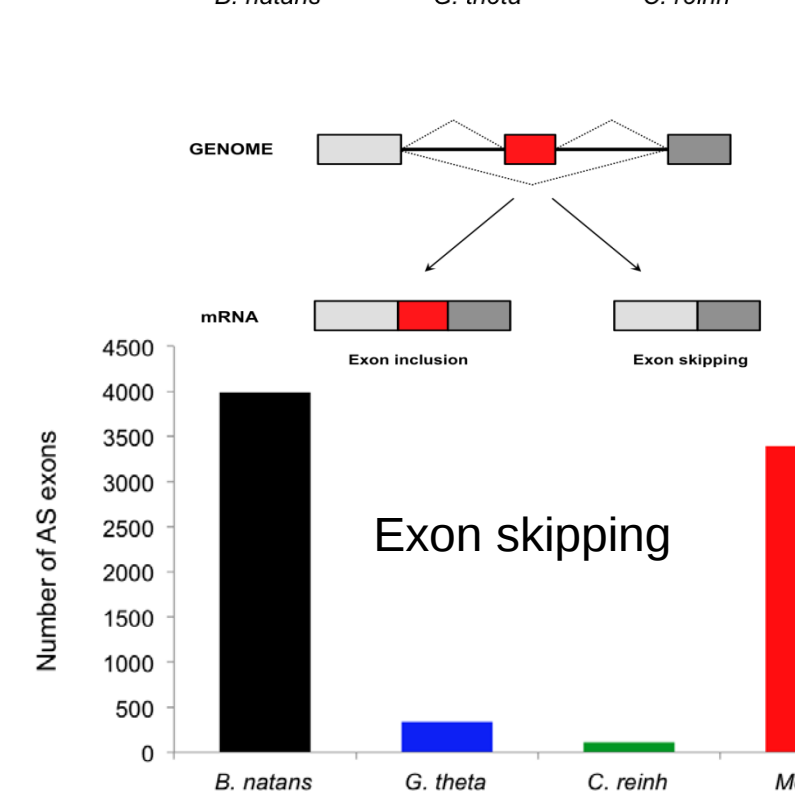
genome	size	# chromo.	# genes	size	# chromo.	# genes
nucleomorph	373 kbp	3	283	551 kbp	3	485
chloroplast	69 kbp	1	61	122 kbp	1	147
mitochondrion	37 kbp	1	22	? kbp	1	?
nucleus	95 Mbp	< 302	21380	87 Mbp	< 670	24525

Bnatans has unexpectedly large amounts of alternative splicing (AS)



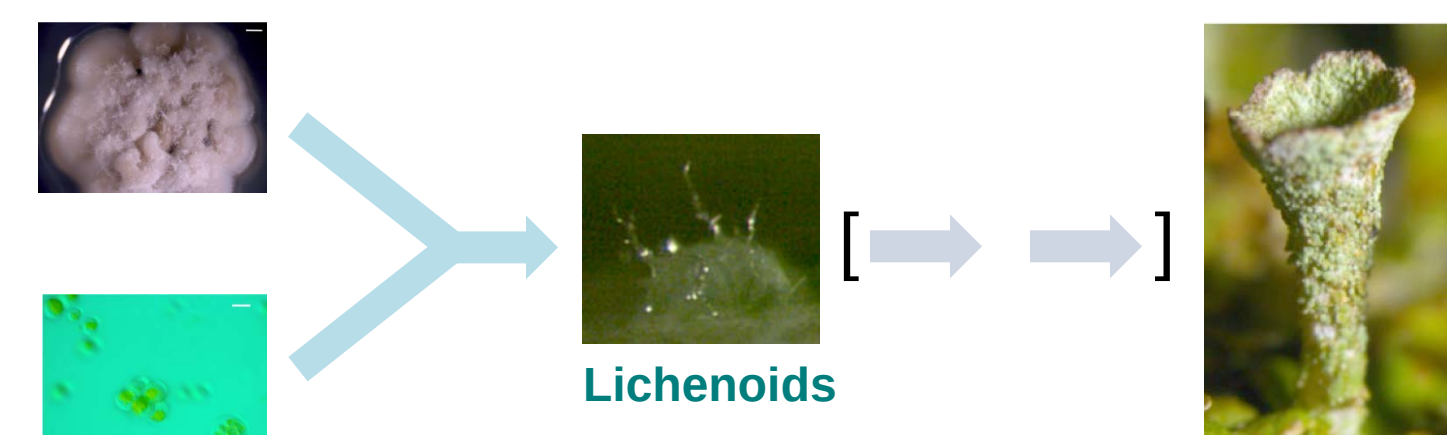
	<i>Bnatans</i>	<i>Gtheta</i>
# genes	21380	24525
# transcripts	27065	25438
Avg. # exons / transcript	9.4	7.1

Lots of introns.
15% of genes are AS.
33% of transcripts are AS.

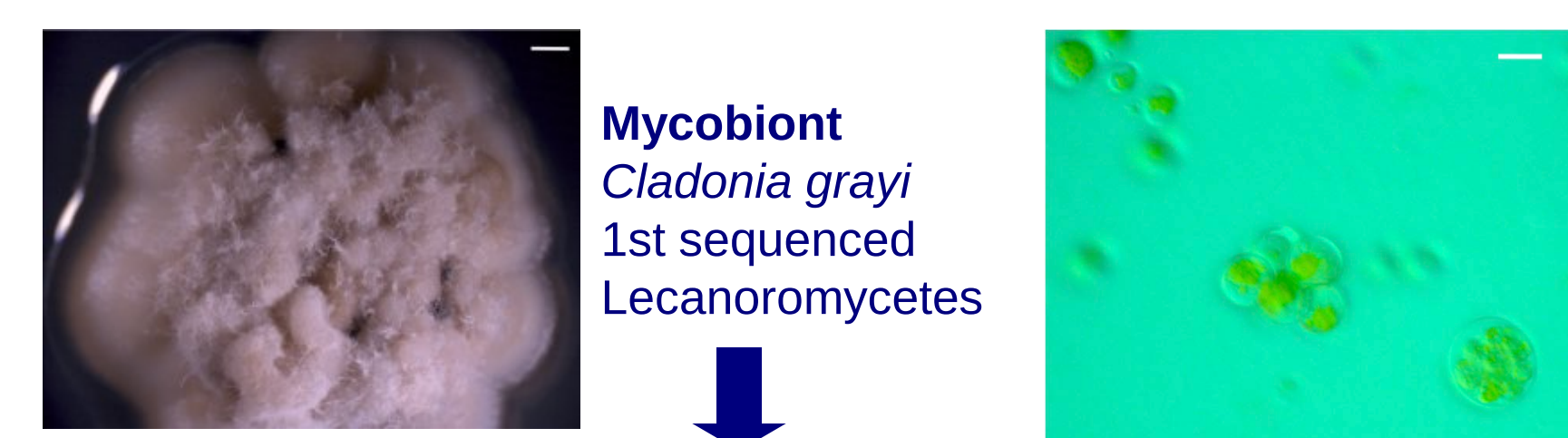


The lichen photobiont *Asterochloris* sp.

The lichen symbiosis, and its *in vitro* reconstitution.



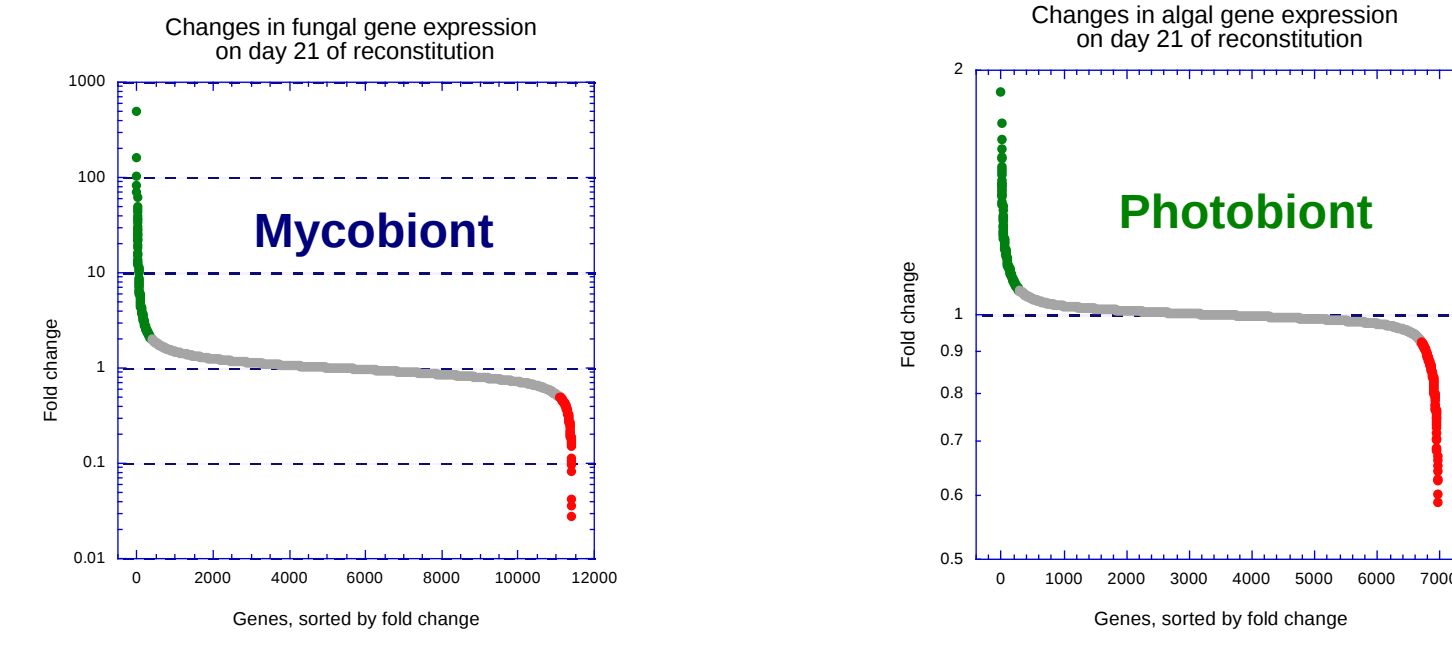
Genomes of the 2 lichen symbionts.



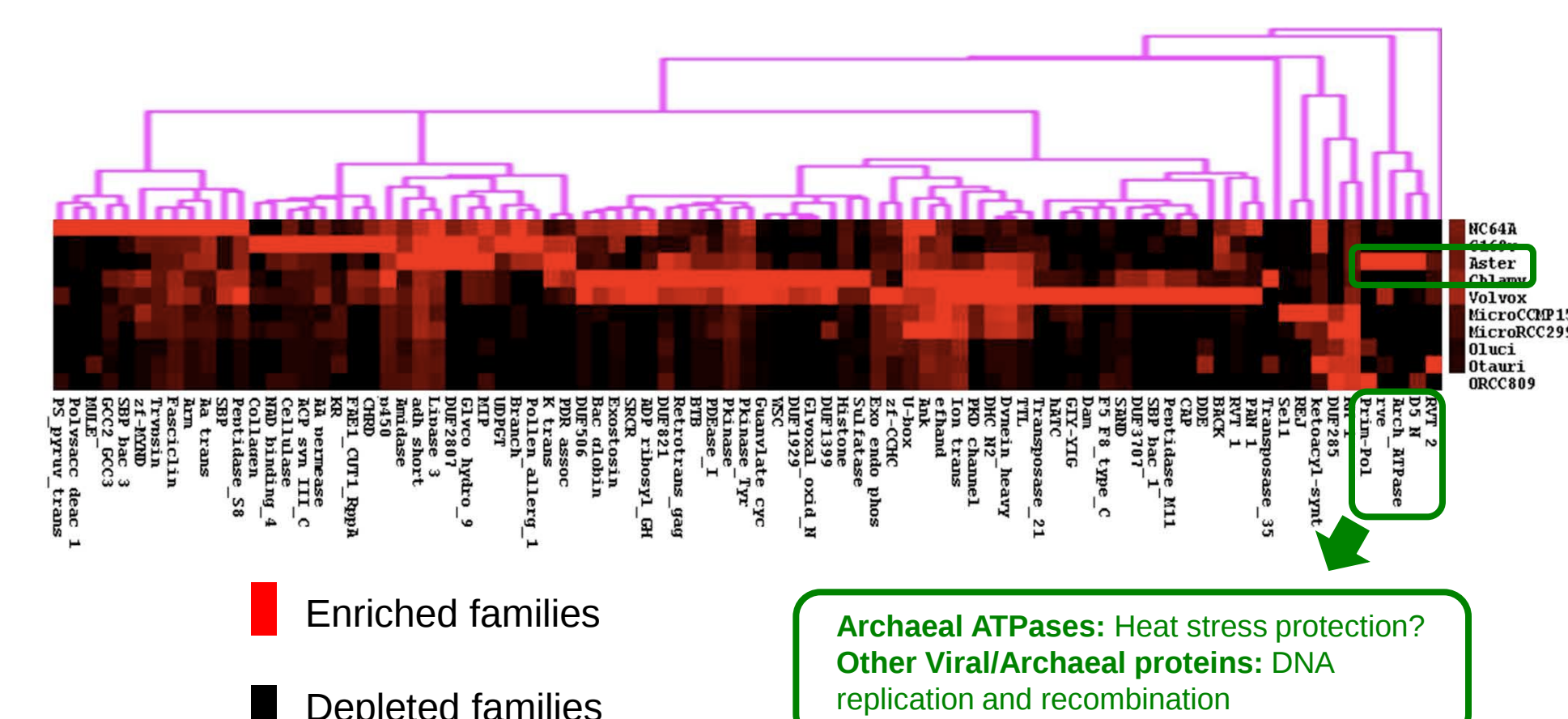
Genome size (Mbp)	35	56
# scaffolds	414	151
# genes	11274	9636
# transcripts	11389	10025
Avg. # exons / transcript	3.0	9.0

Photobiont
Asterochloris
3rd sequenced
Trebouxiphyceae,
after *Chlorella* and
Coccomyxa

Gene expression changes in the lichenoids.



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



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 Armaleo, Olaf Mueller