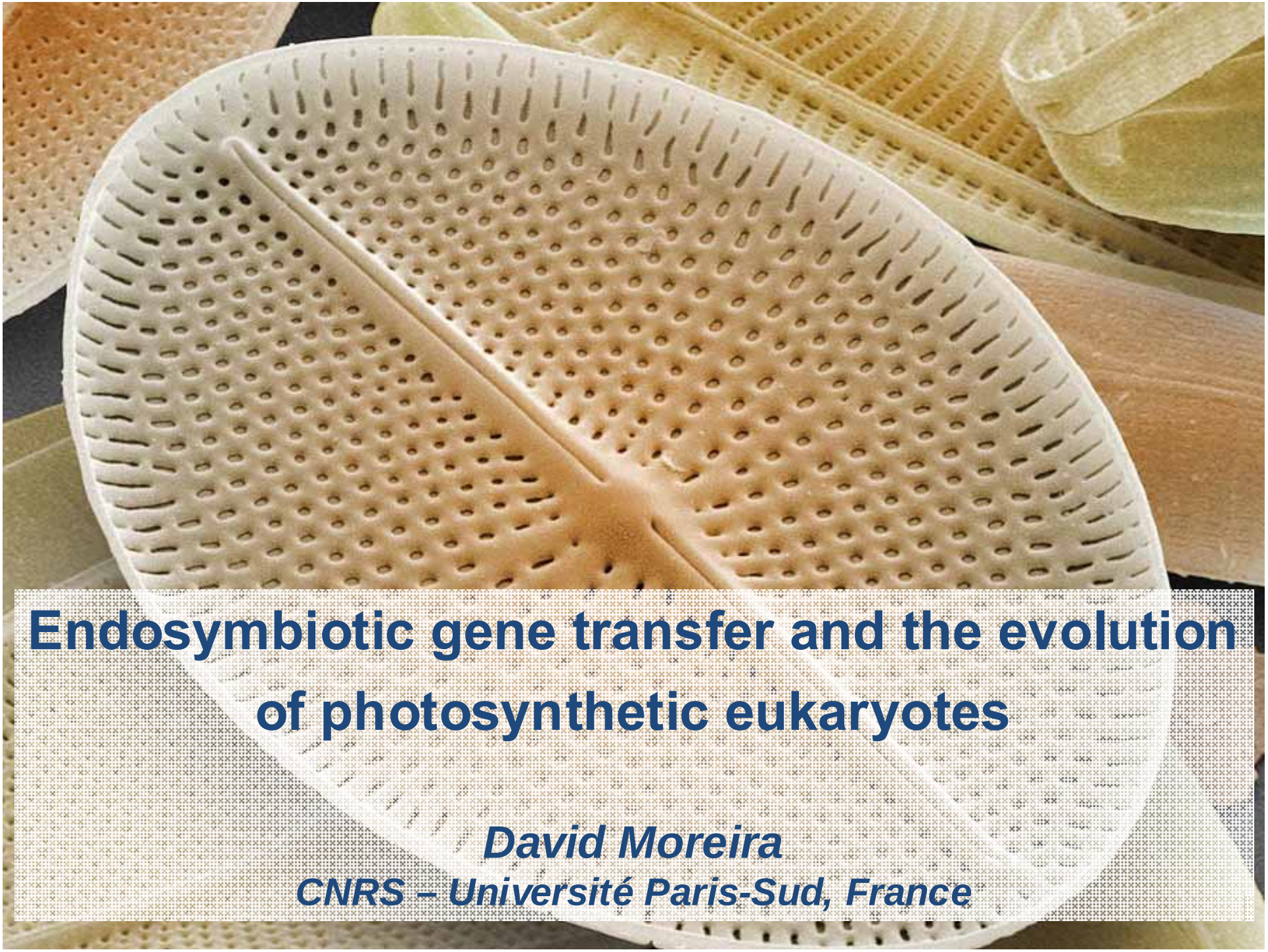
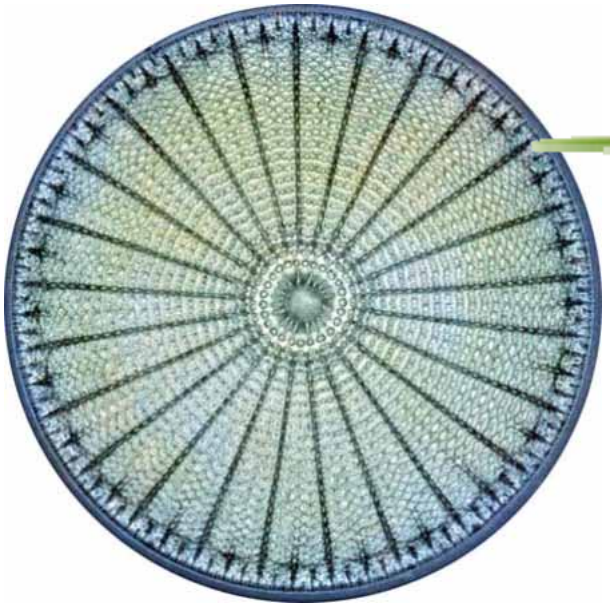
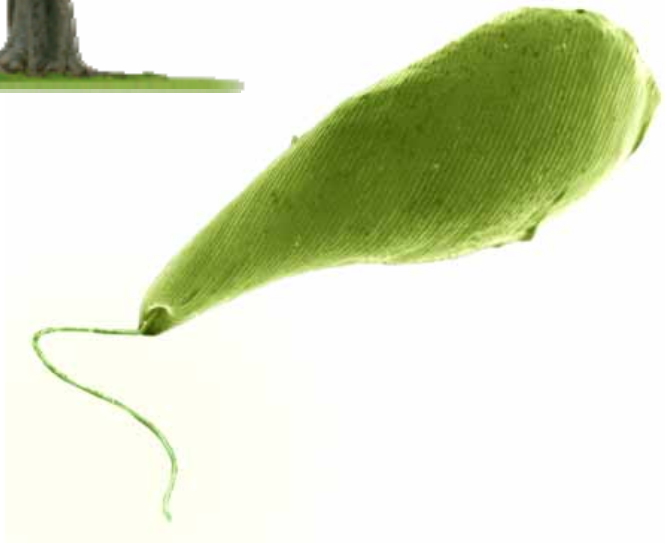


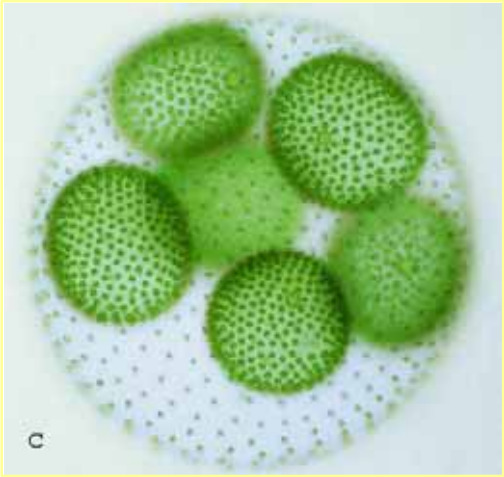
Endosymbiotic gene transfer and the evolution of photosynthetic eukaryotes

David Moreira
CNRS – Université Paris-Sud, France



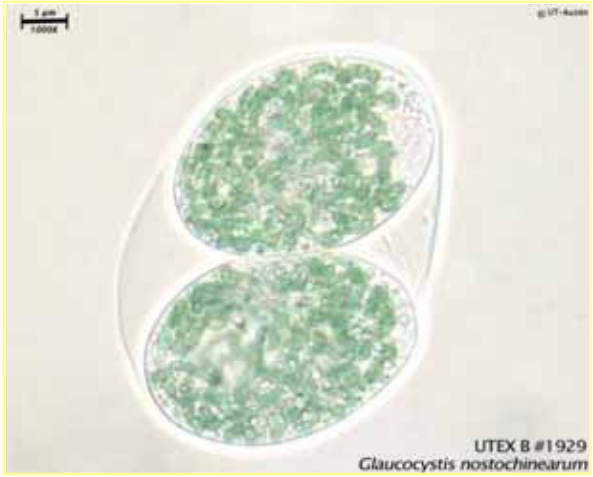


Two types of plastids: 2-membrane plastids



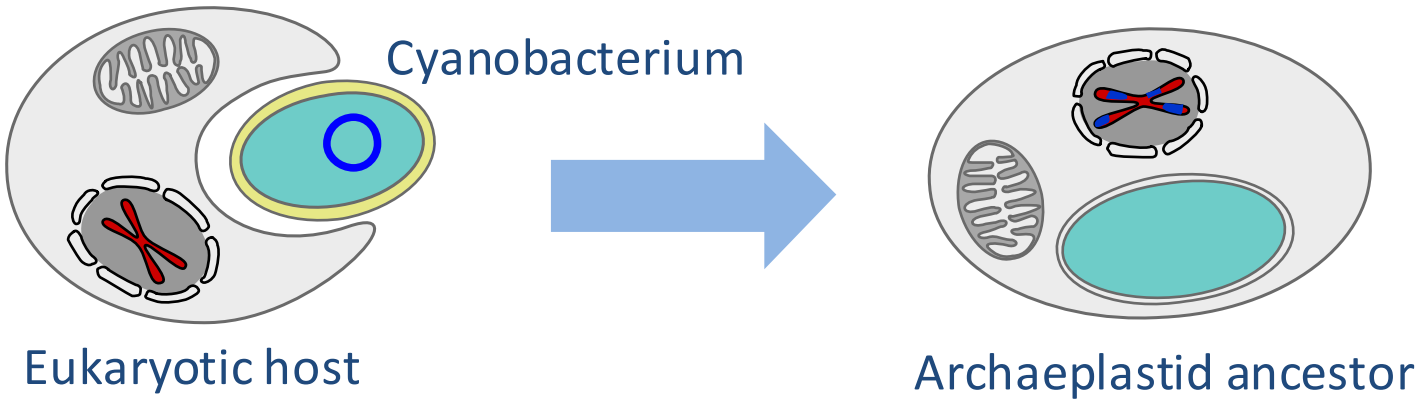
Viridiplantae

Rhodophyta

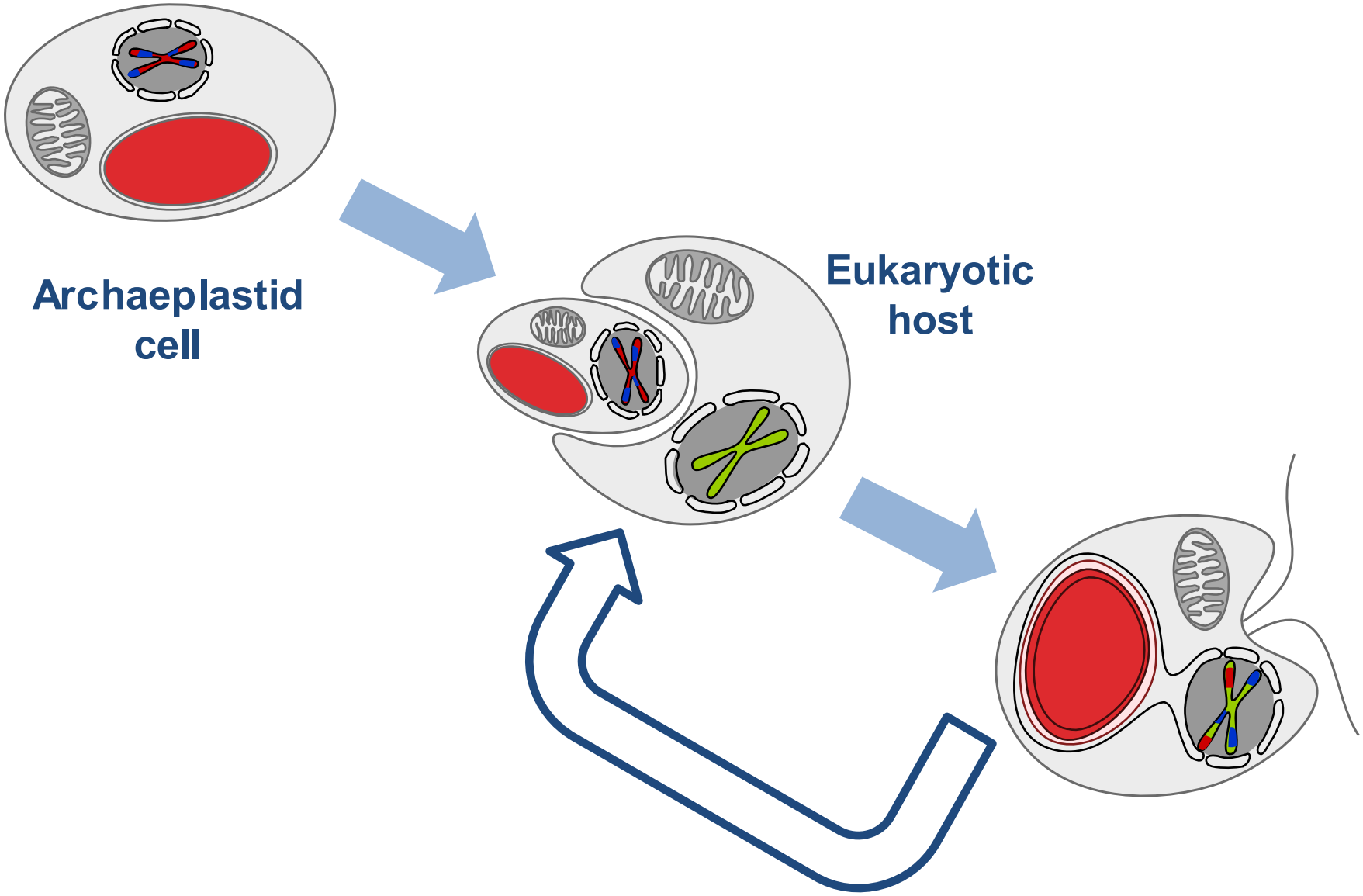


Glaucophyta

Archaeplastida (or Plantae)

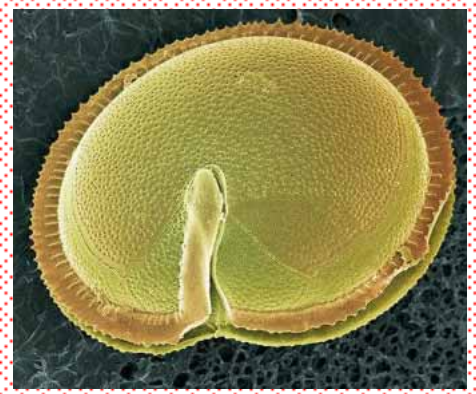


Two types of plastids : >2-membrane plastids

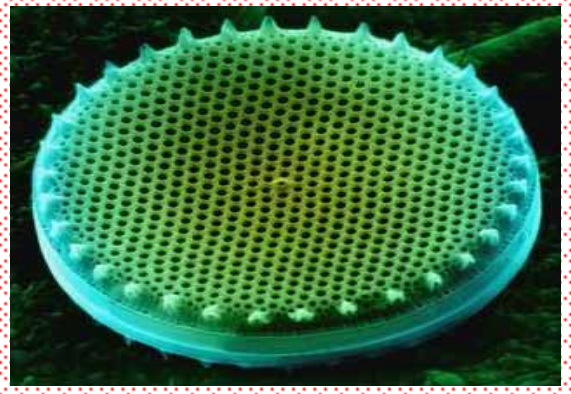


Two types of plastids : >2-membrane plastids

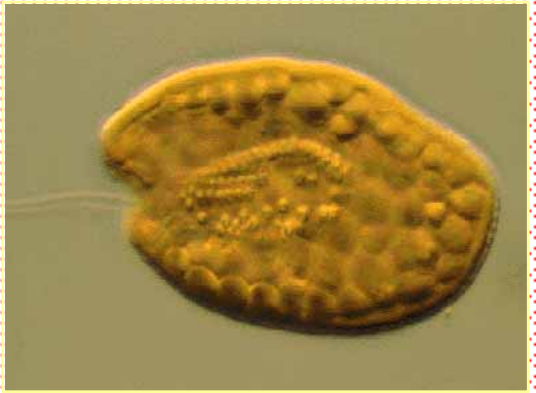
RED LINE



Alveolata

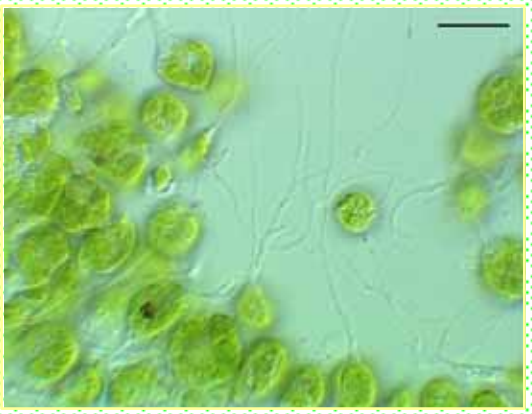


Heterokonta



Cryptophyta

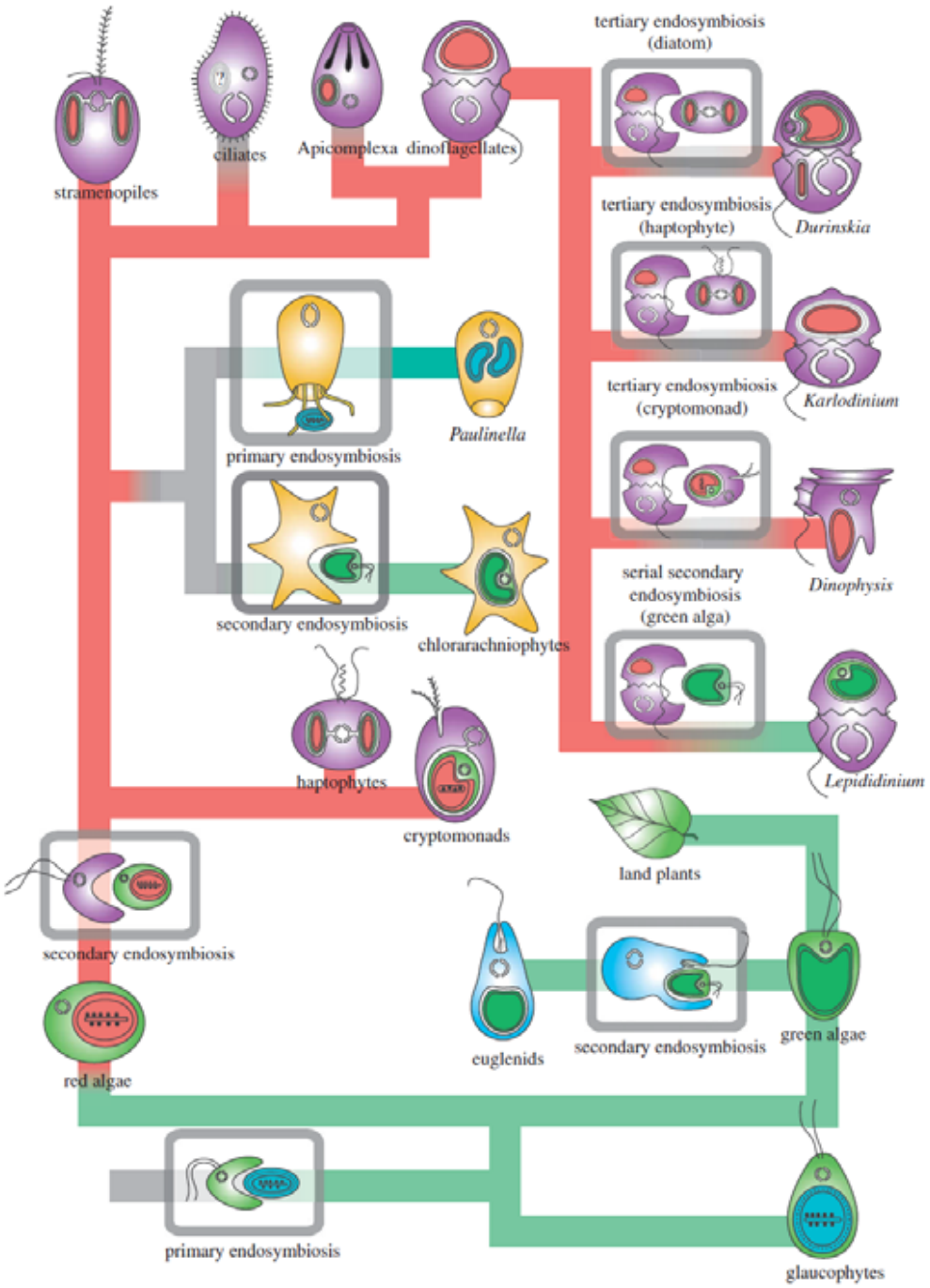
Haptophyta



Chlorarachniophyta

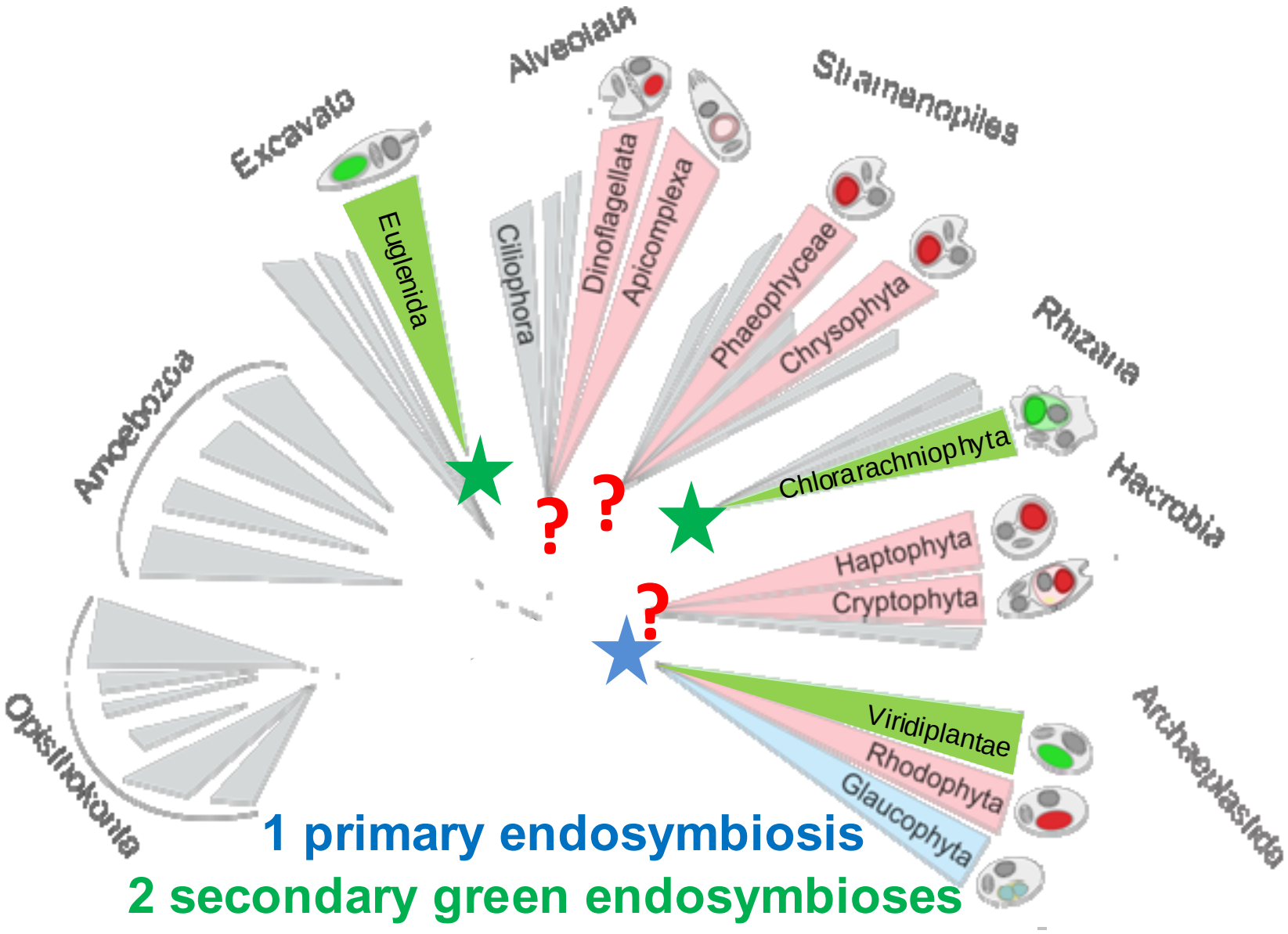
Euglenida

GREEN LINE



(Keeling, 2010)

A vast diversity of photosynthetic eukaryotes



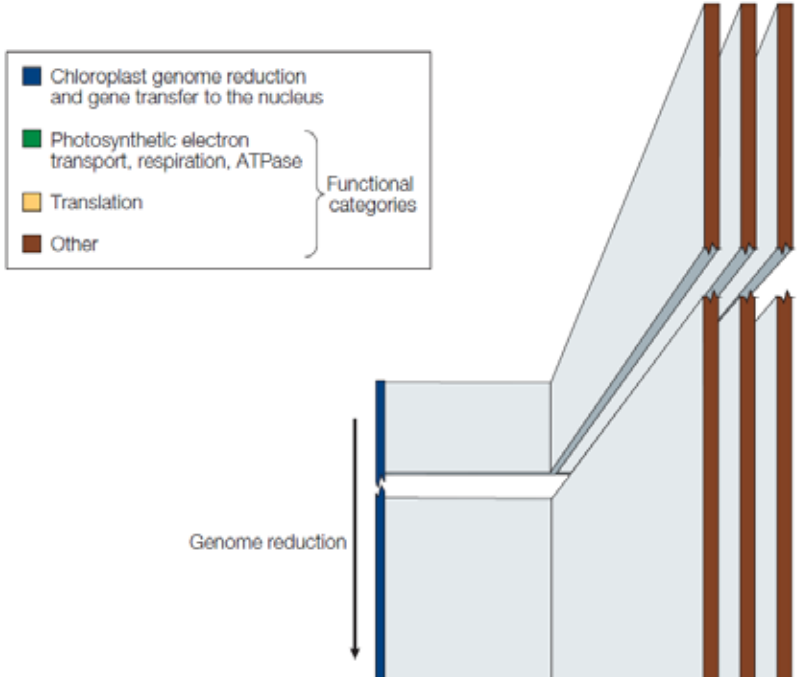
1 primary endosymbiosis

2 secondary green endosymbioses

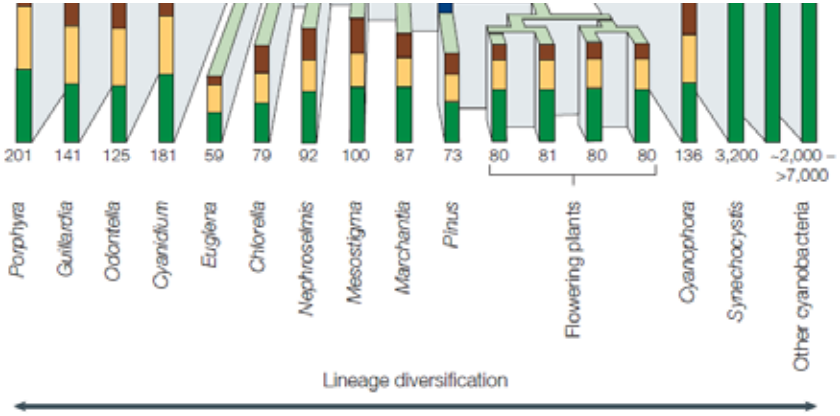
X secondary red endosymbioses

Endosymbiotic gene transfer: EGT

Genome	Length [kbp]	Number of protein-coding genes
Algae		
cp *Porphyra purpurea*	191	200
cp *Cyanidium caldarium*	165	197
cp *Guillardia theta*	122	148
cp *Cyanophora paradoxa*	136	136
cp *Odontella sinensis*	120	124
cp *Euglena gracilis*	143	58
Land plants		
cp *Marchantia polymorpha*	121	84
cp *Chlorella vulgaris*	151	78
cp *Nicotiana tabacum*	156	76
cp *Oryza sativa*	134	76
cp *Zea mays*	140	76
cp *Pinus thunbergii*	120	69
Non-phosynthetic plastids		
cp *Toxoplasma gondii*	35	26
cp *Eimeria tenella*	35	28
cp *Epifagus* virginiana	70	21
Cyanobacteria		
Synechocystis sp.	3573	3168
Prochlorococcus marinus	1660	1884
Nostoc PCC 7120	6413	5368
Nostoc punctiforme	~9000	~7400



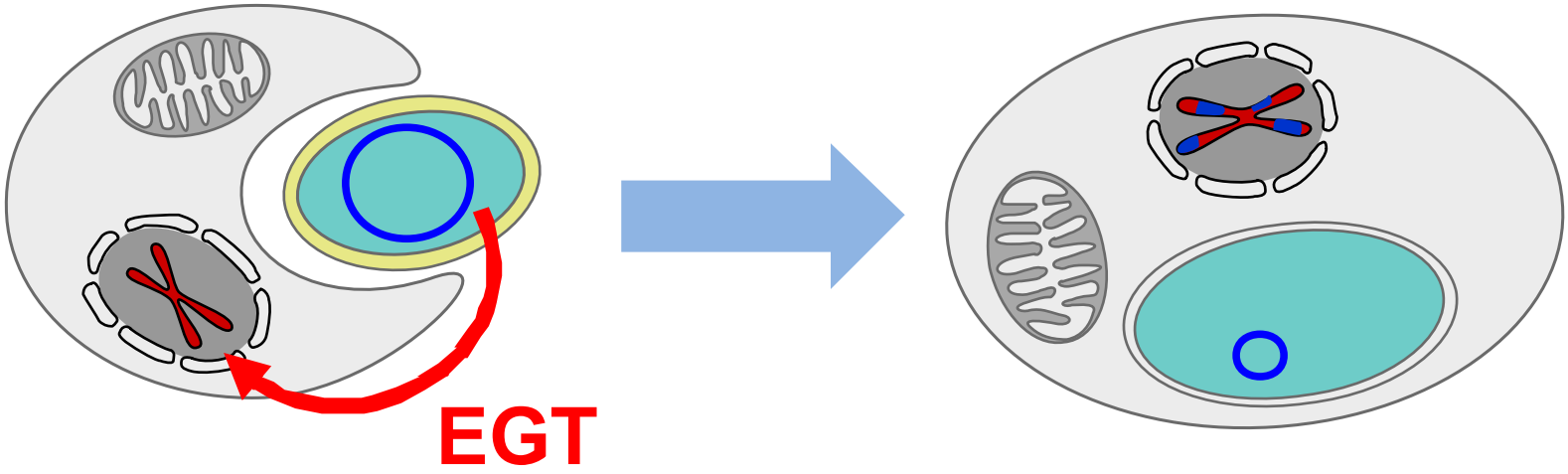
However, plastids need many more proteins to function



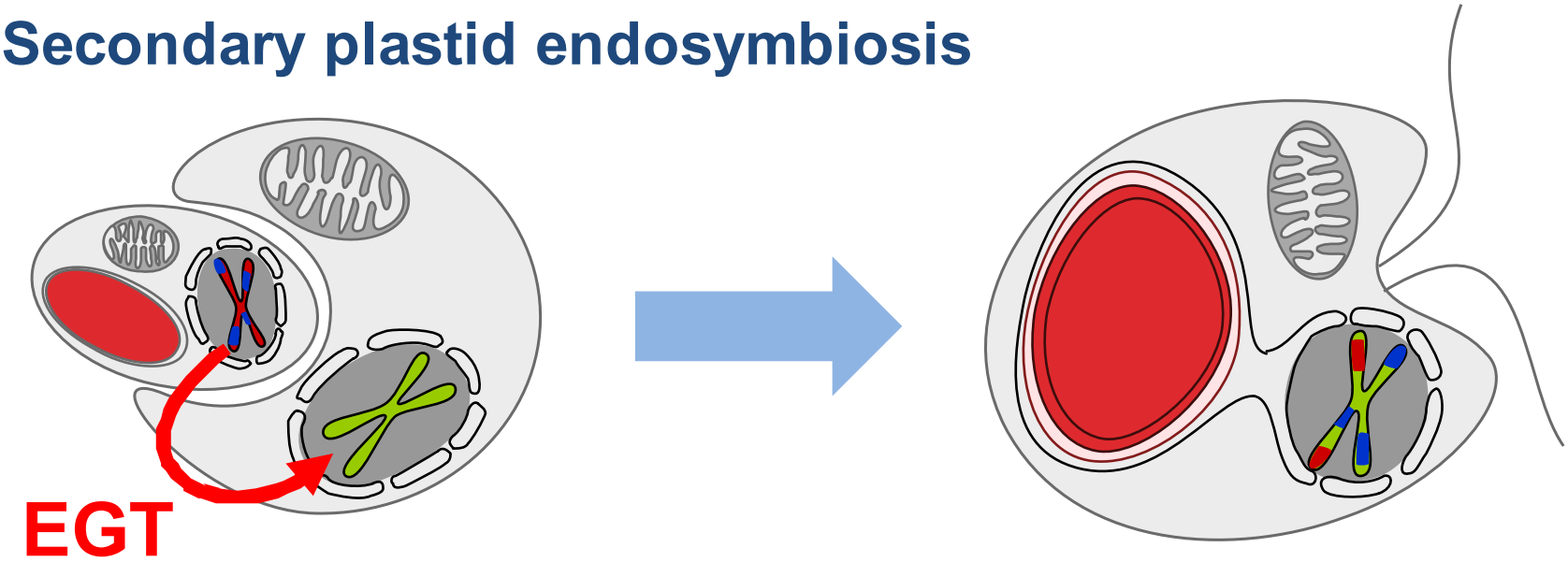
(Timmis et al. 2004)

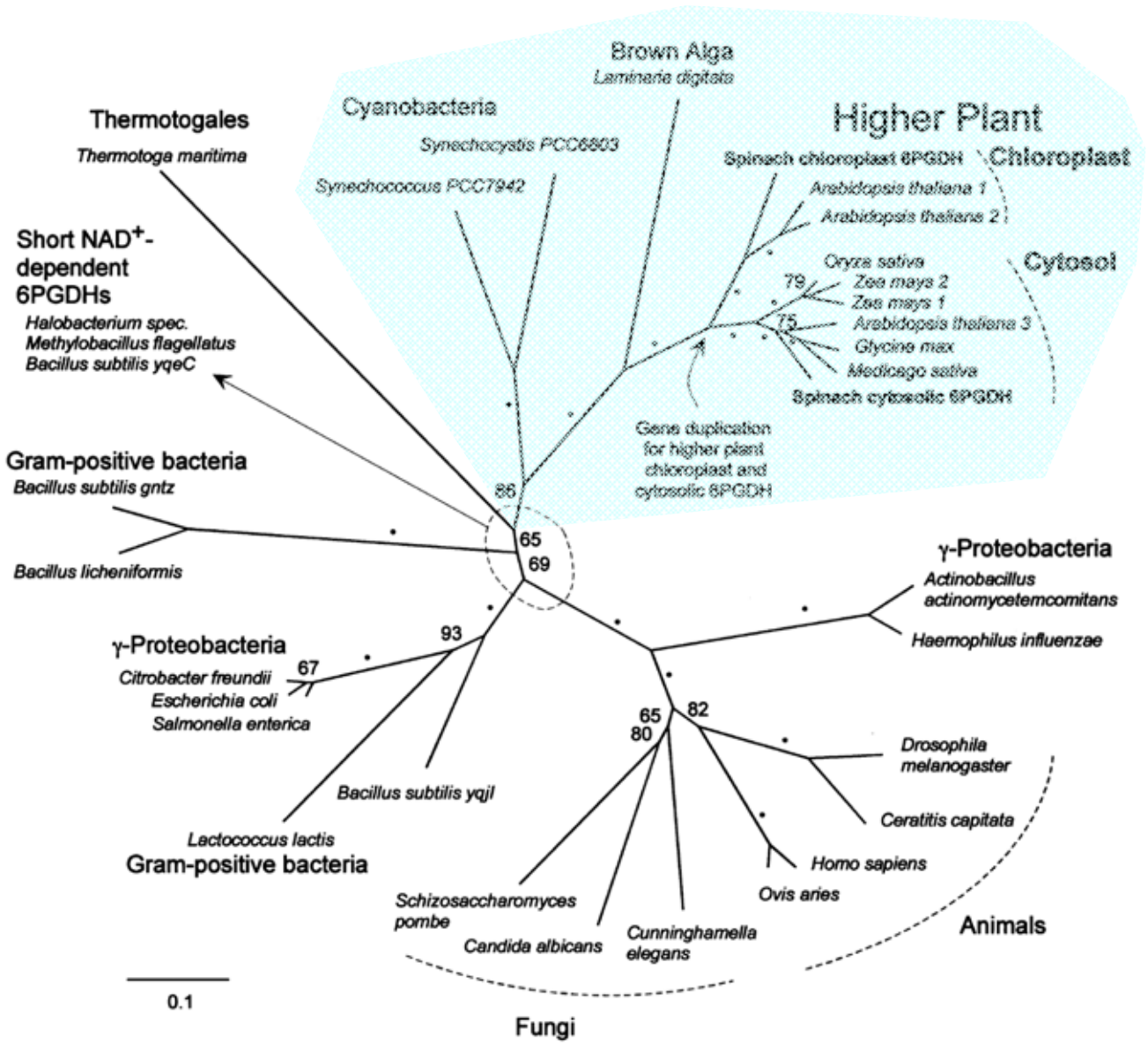
Endosymbiotic Gene Transfer (EGT)

Primary plastid endosymbiosis



Secondary plastid endosymbiosis





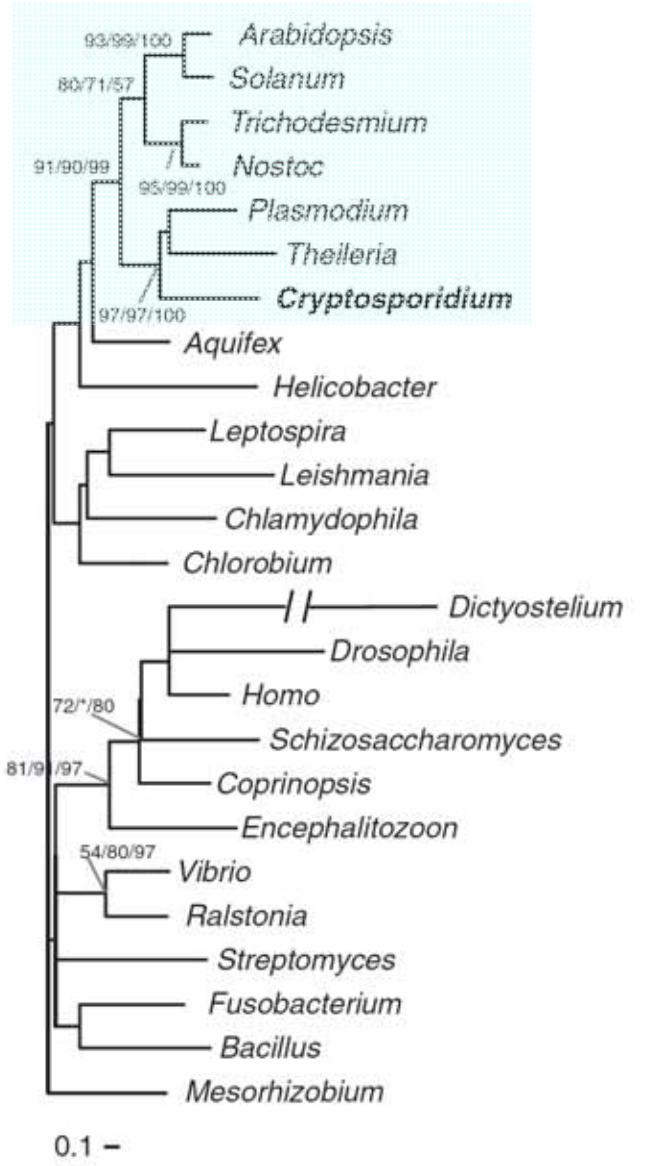
6-phosphogluconate dehydrogenase (Krepinsky et al. 2001)

Endosymbiotic Gene Transfer (EGT)

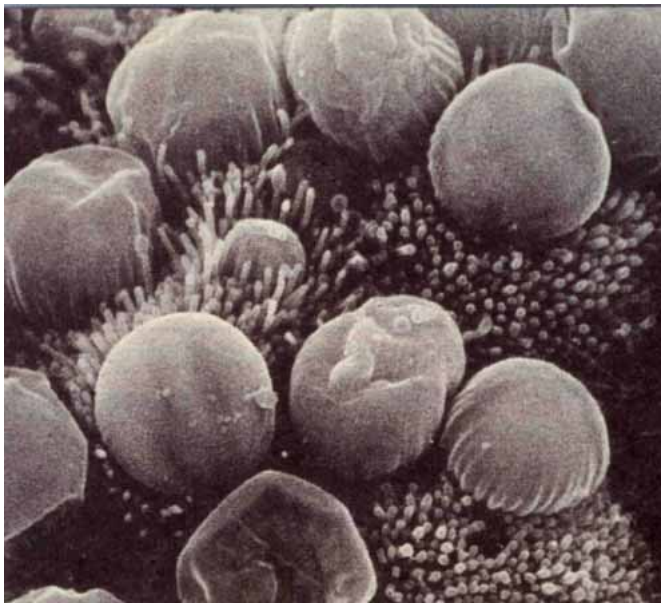
18% of *Arabidopsis* proteins (~4500) have been acquired from the plastid

Martin et al. 2002

EGT and detection of cryptic endosymbioses

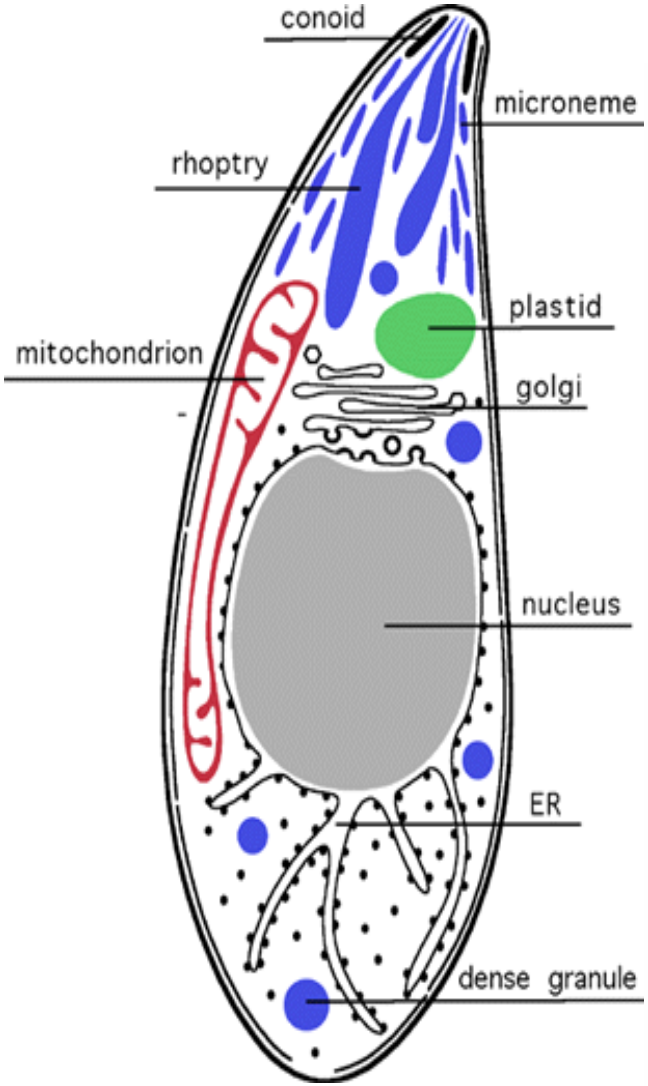


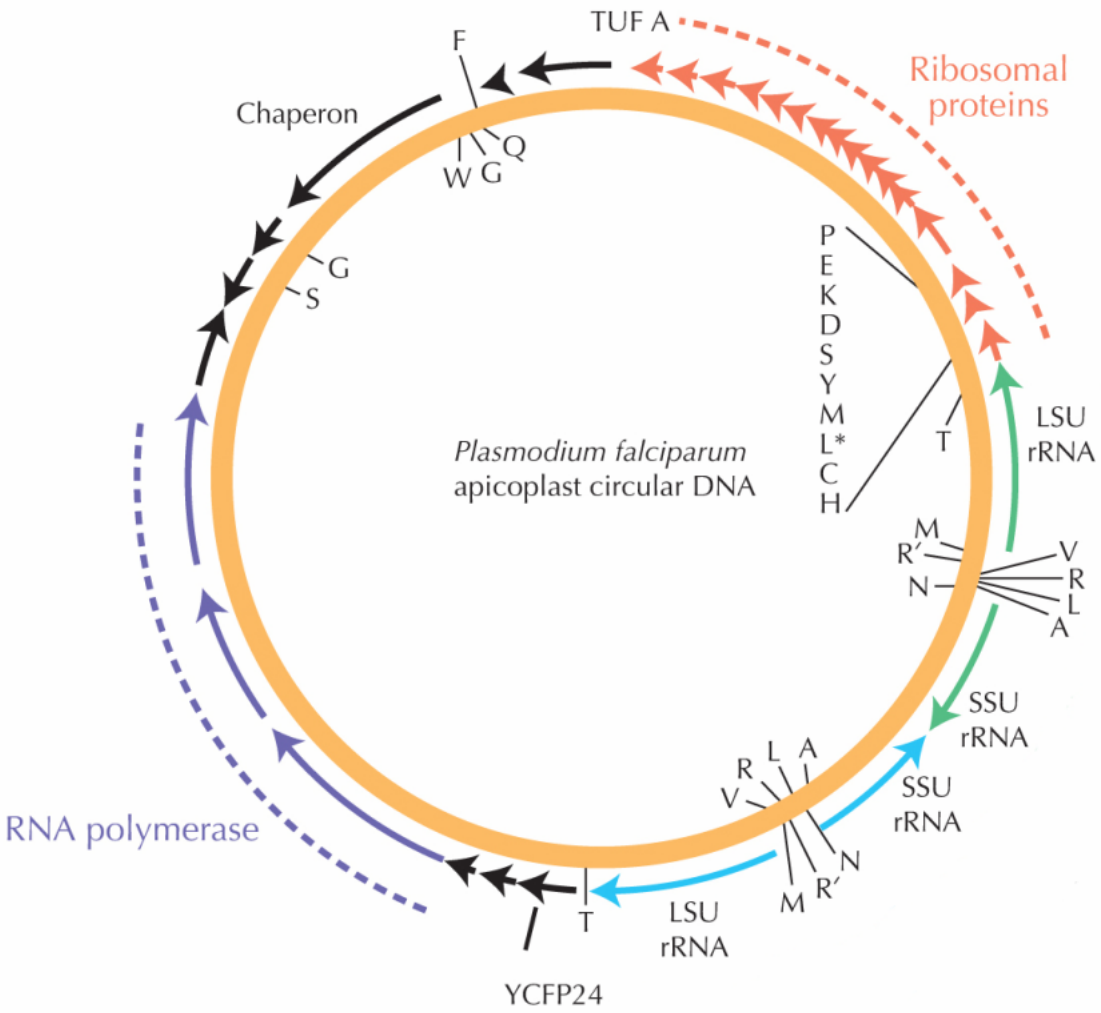
Leucine aminopeptidase (Huang et al. 2004)



Cryptosporidium parvum

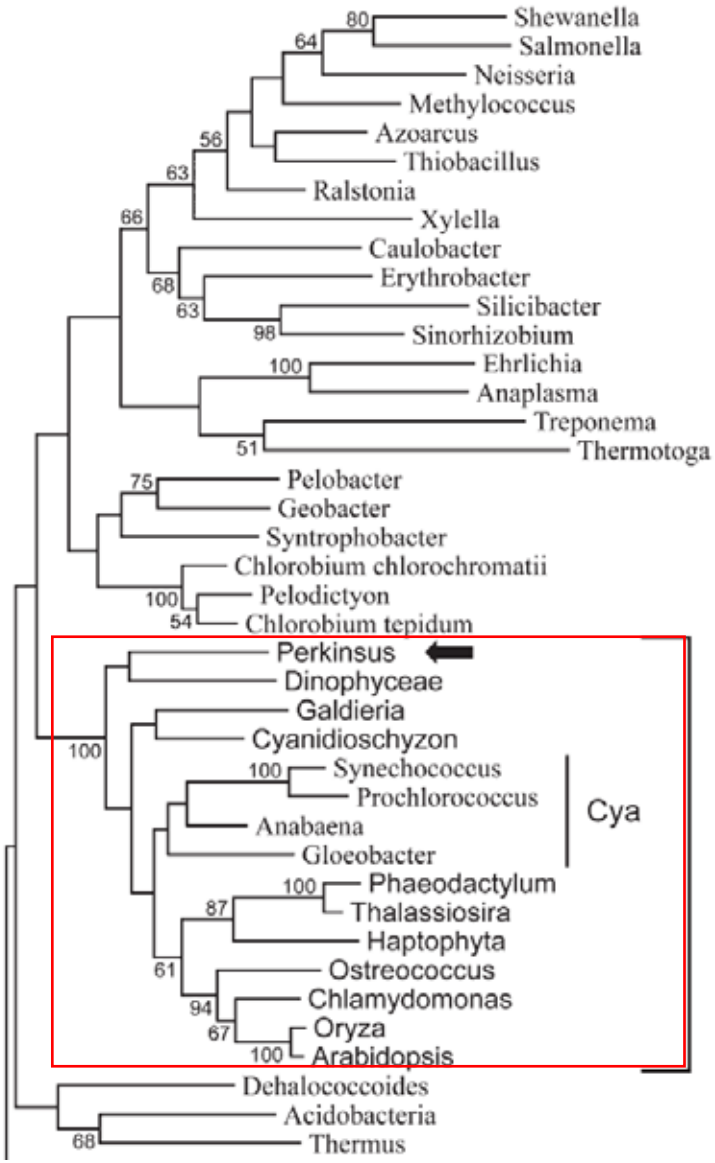
EGT and detection of cryptic endosymbioses





EGT and detection of cryptic endosymbioses

Perkinsus marinus



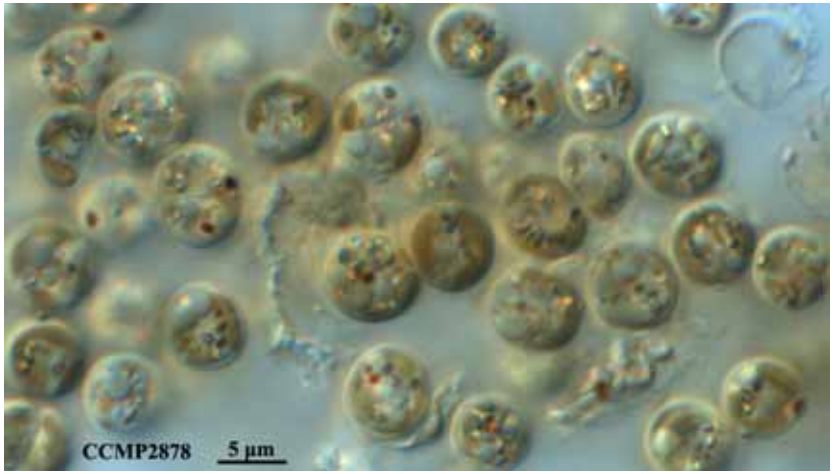
Matsuzaki *et al.* - *MBE* 2008

Unexpected cryptic endosymbioses?

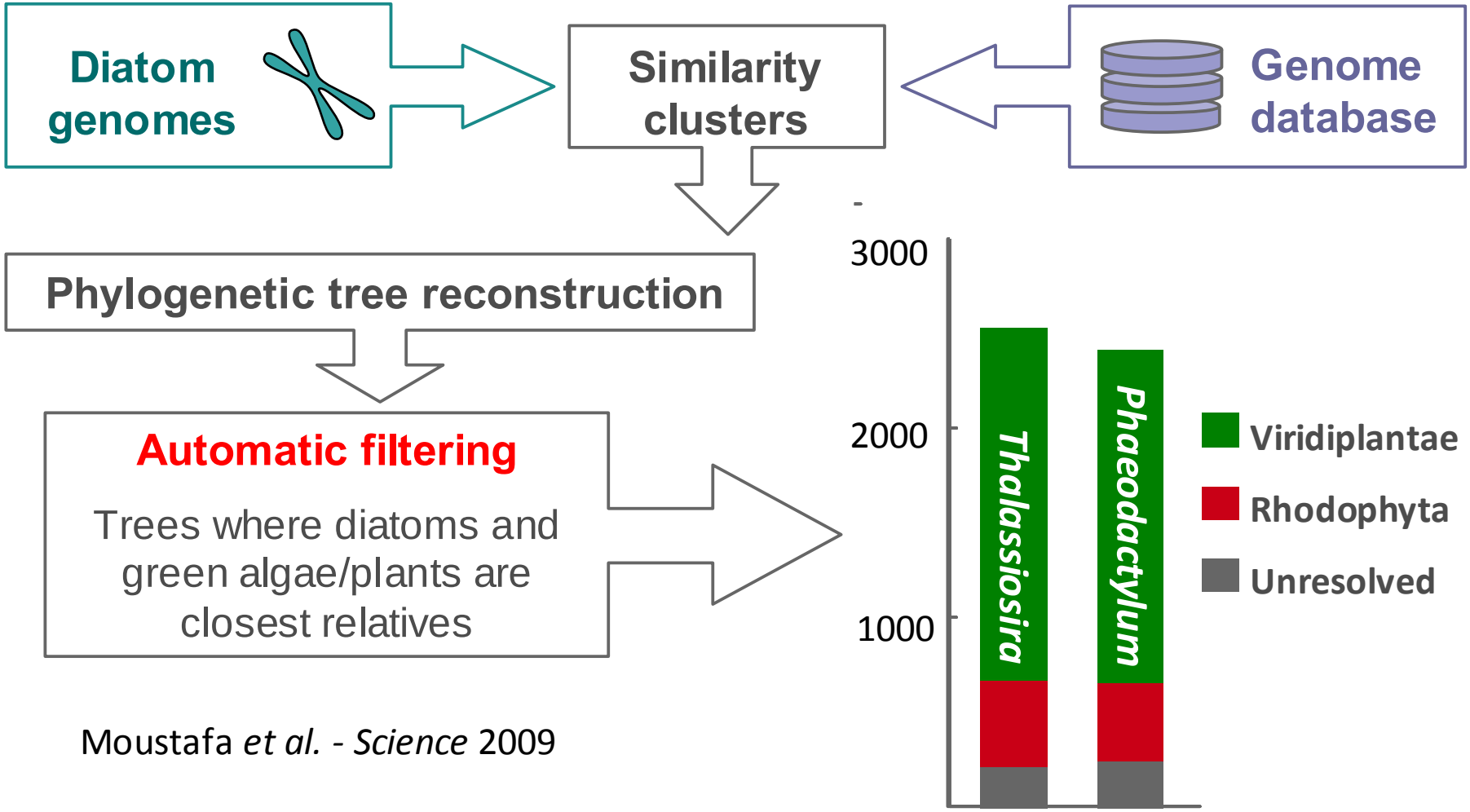
Based on genomic surveys:

Moustafa *et al.*
Genomic footprints of a cryptic plastid endosymbiosis in diatoms.
Science, 2009, 324, 1724-1726

Woehle *et al*.
Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and Apicomplexa-related *Chromera velia*
Genome Biol Evol, 2011, 3:1220–1230



Green genes in diatoms

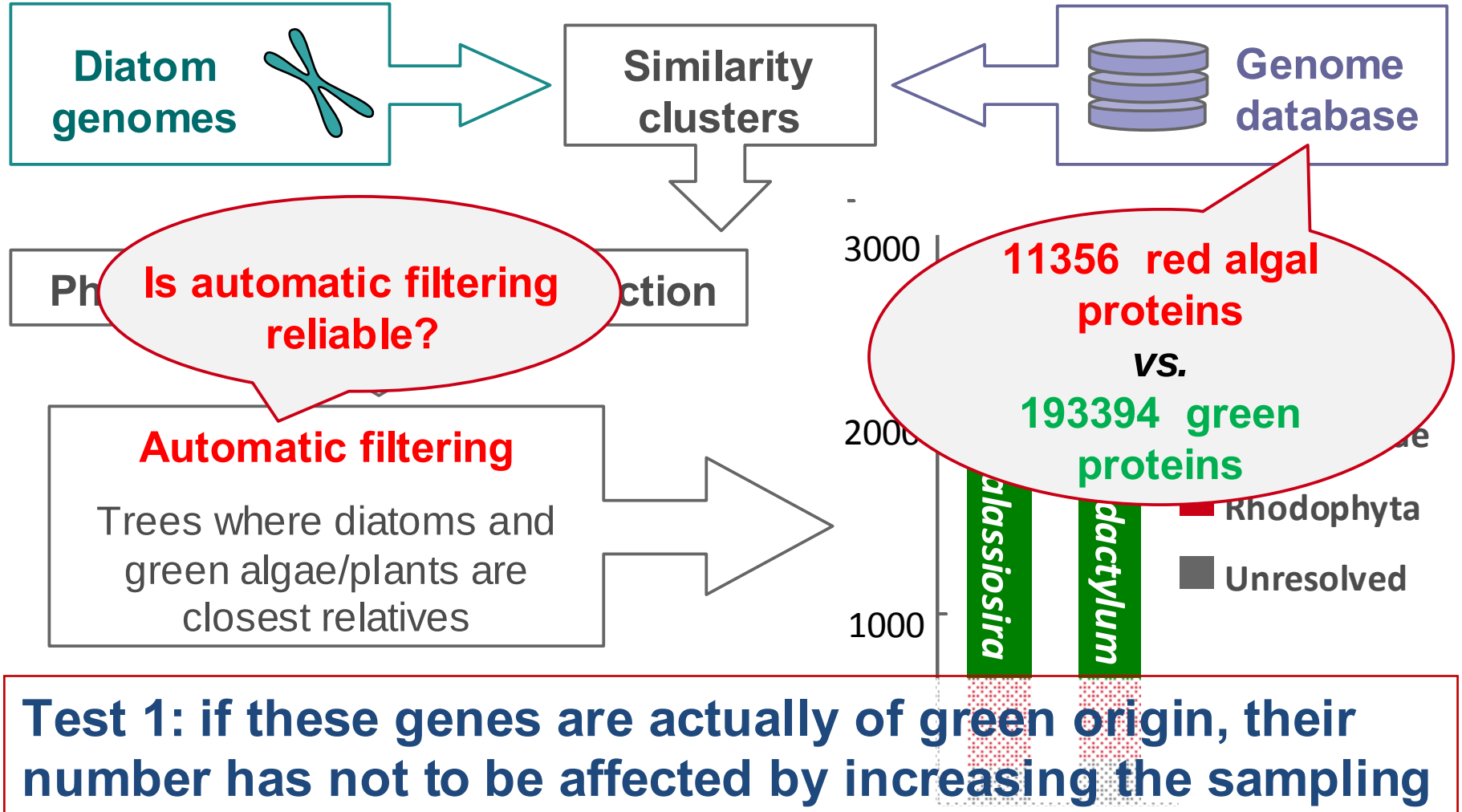


Moustafa *et al.* - *Science* 2009

~3550 putative EGTs of green origin!

Traces of a cryptic green algal endosymbiosis?

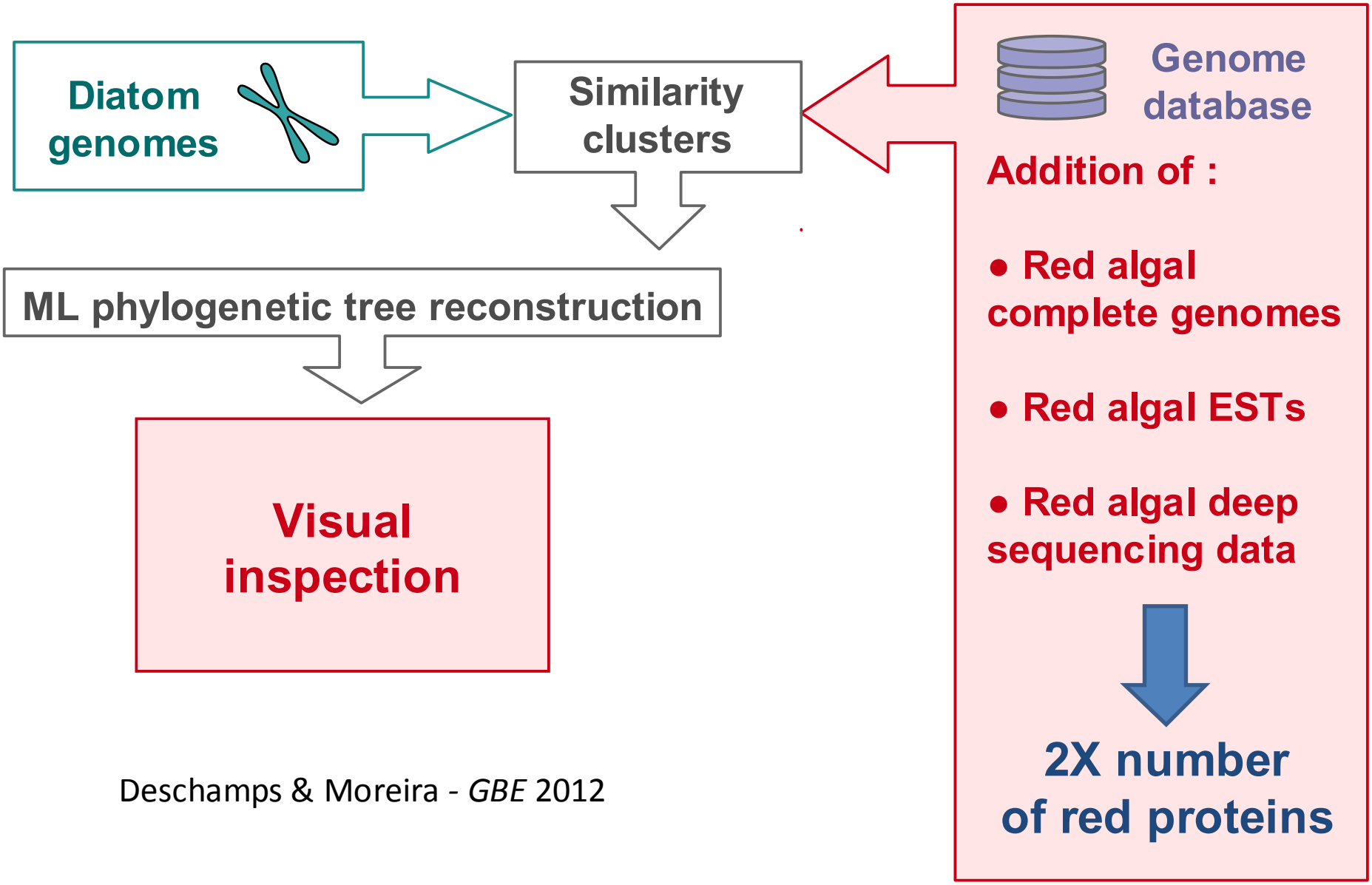
Green genes in diatoms



Test 1: if these genes are actually of green origin, their number has not to be affected by increasing the sampling of red algal sequences

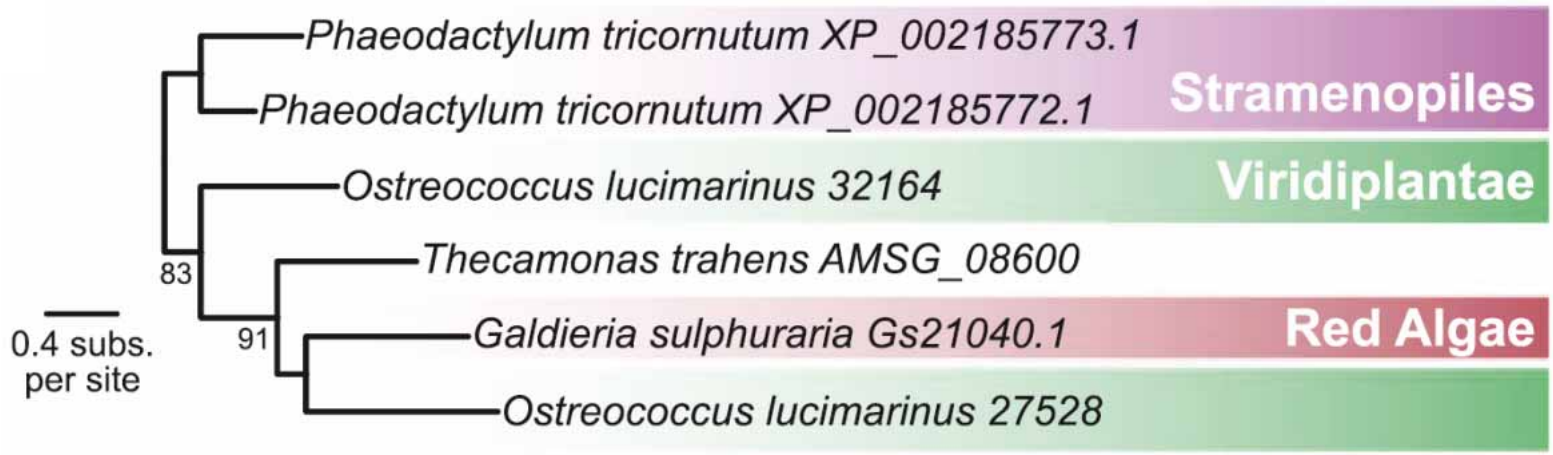
Test 2: replace automatic filtering by visual inspection

Re-analyzing green genes in diatoms

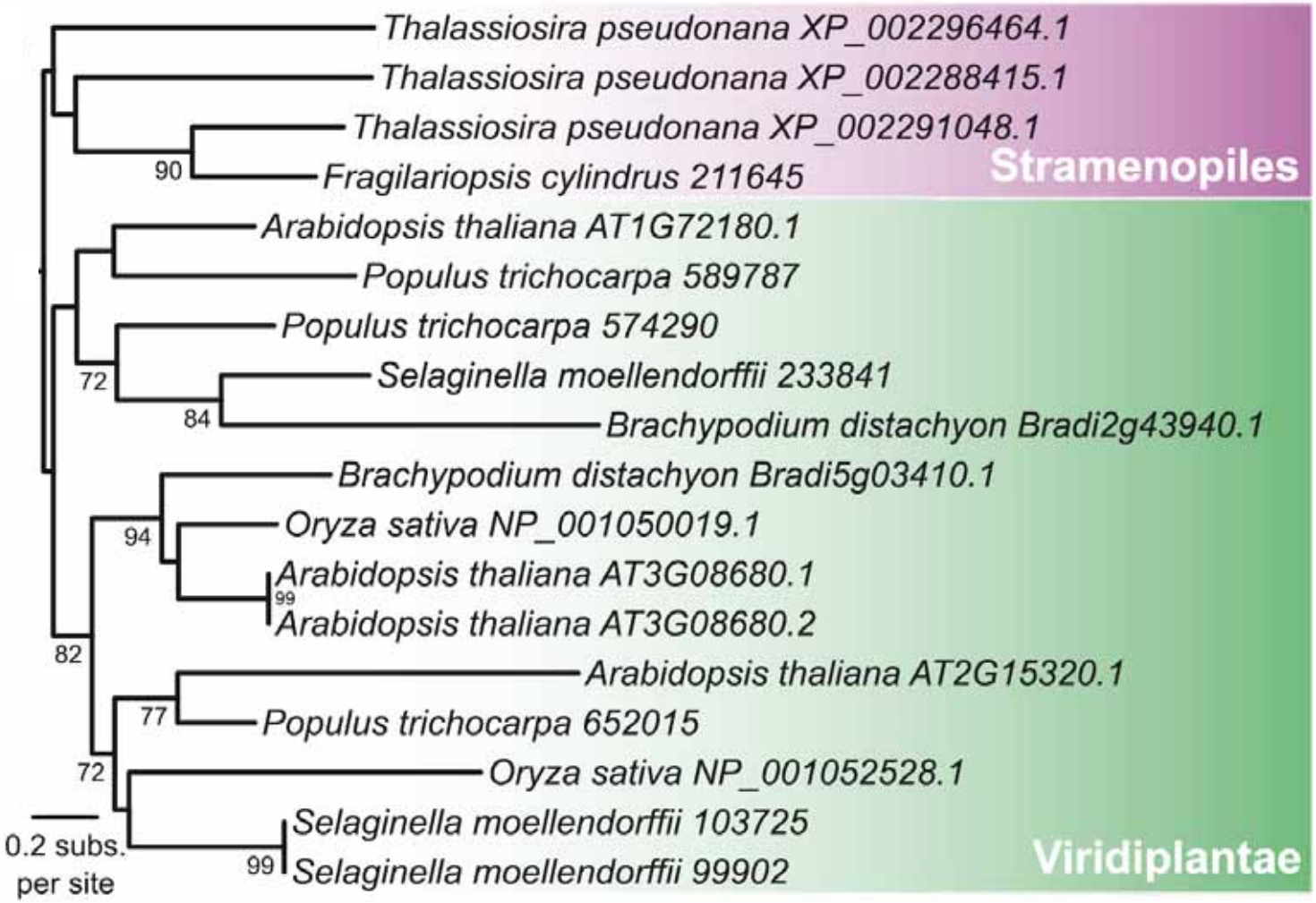


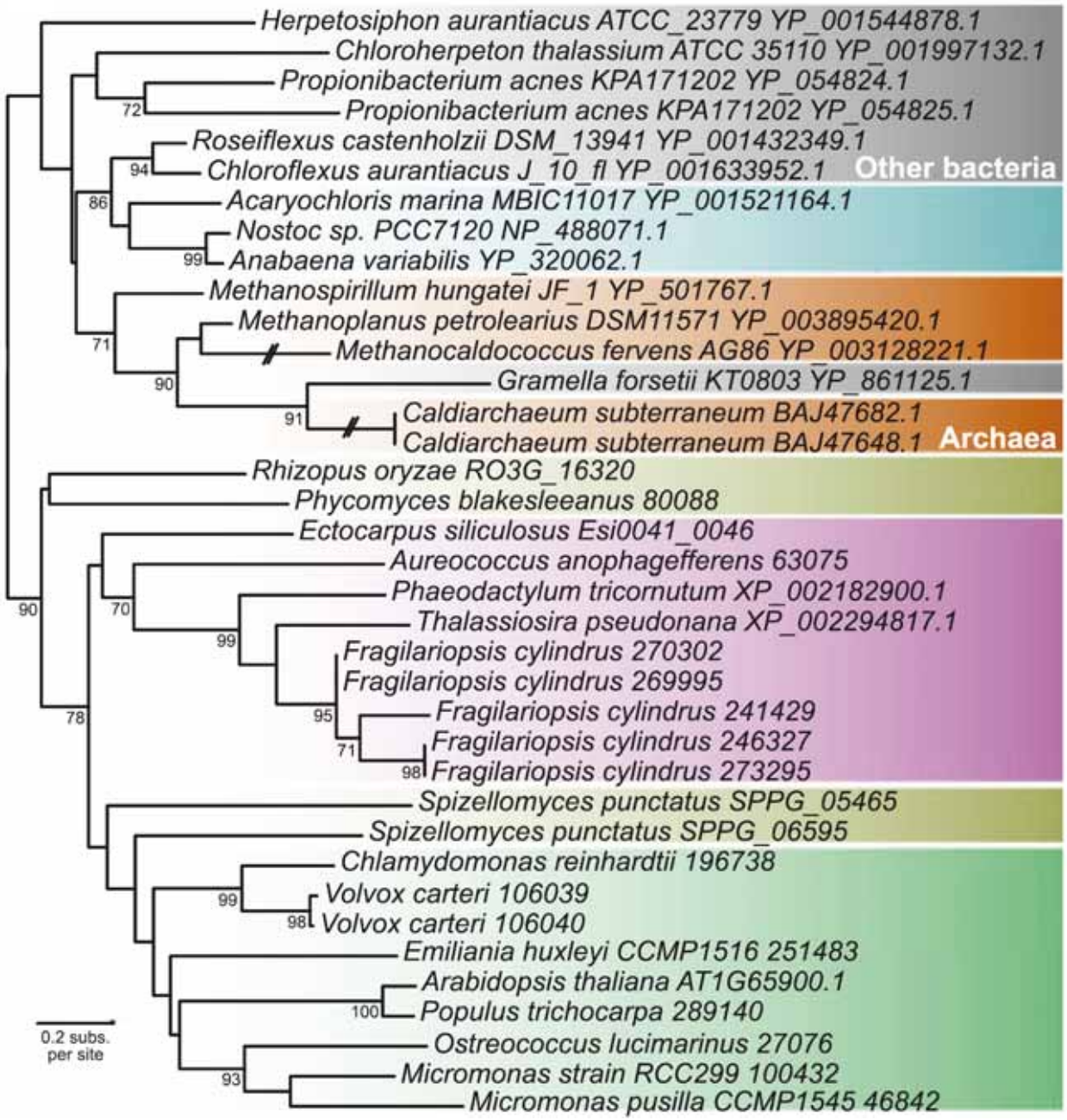
Deschamps & Moreira - *GBE* 2012

« Poor sampling » case

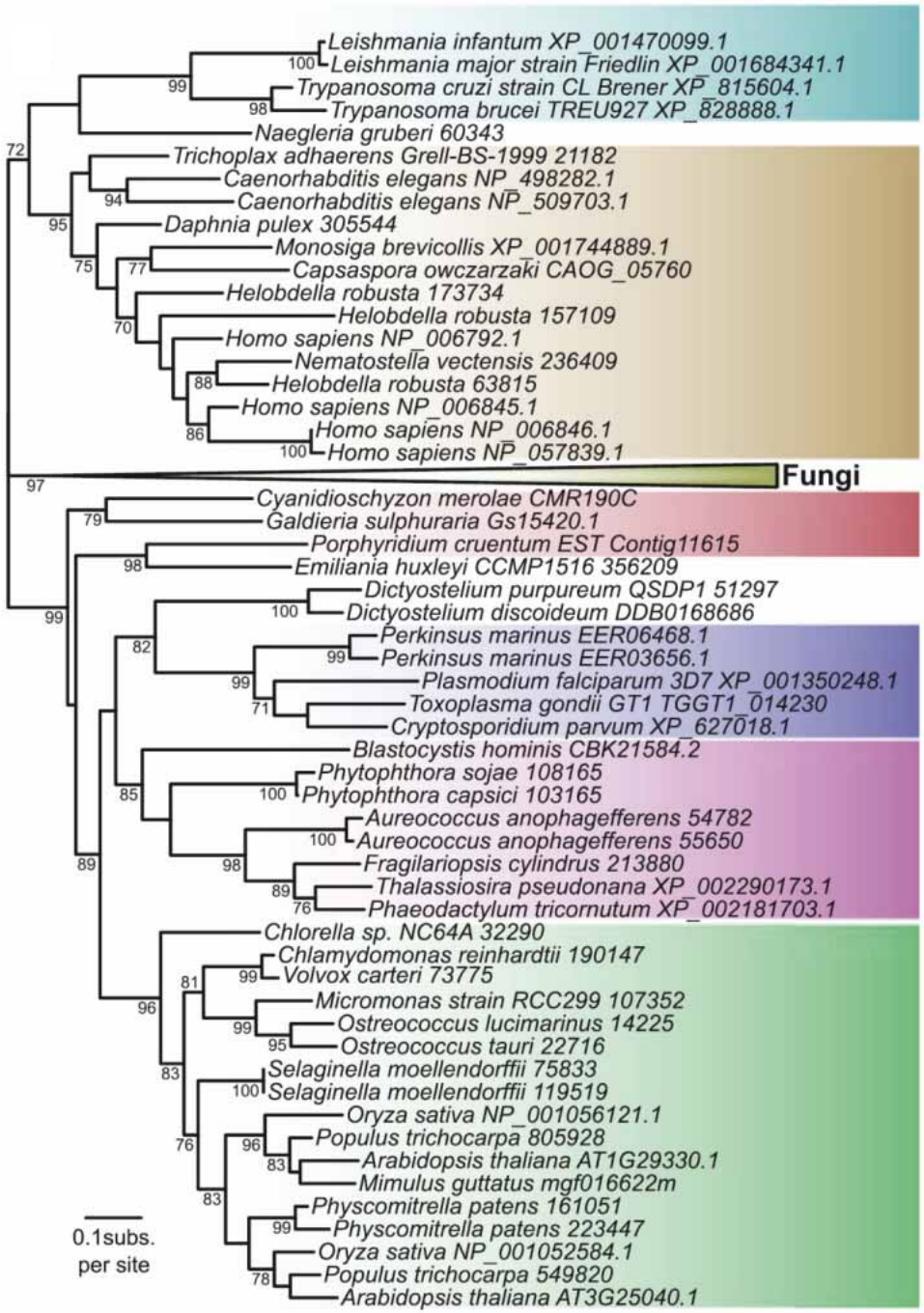


« Only Viridiplantae + diatoms » case



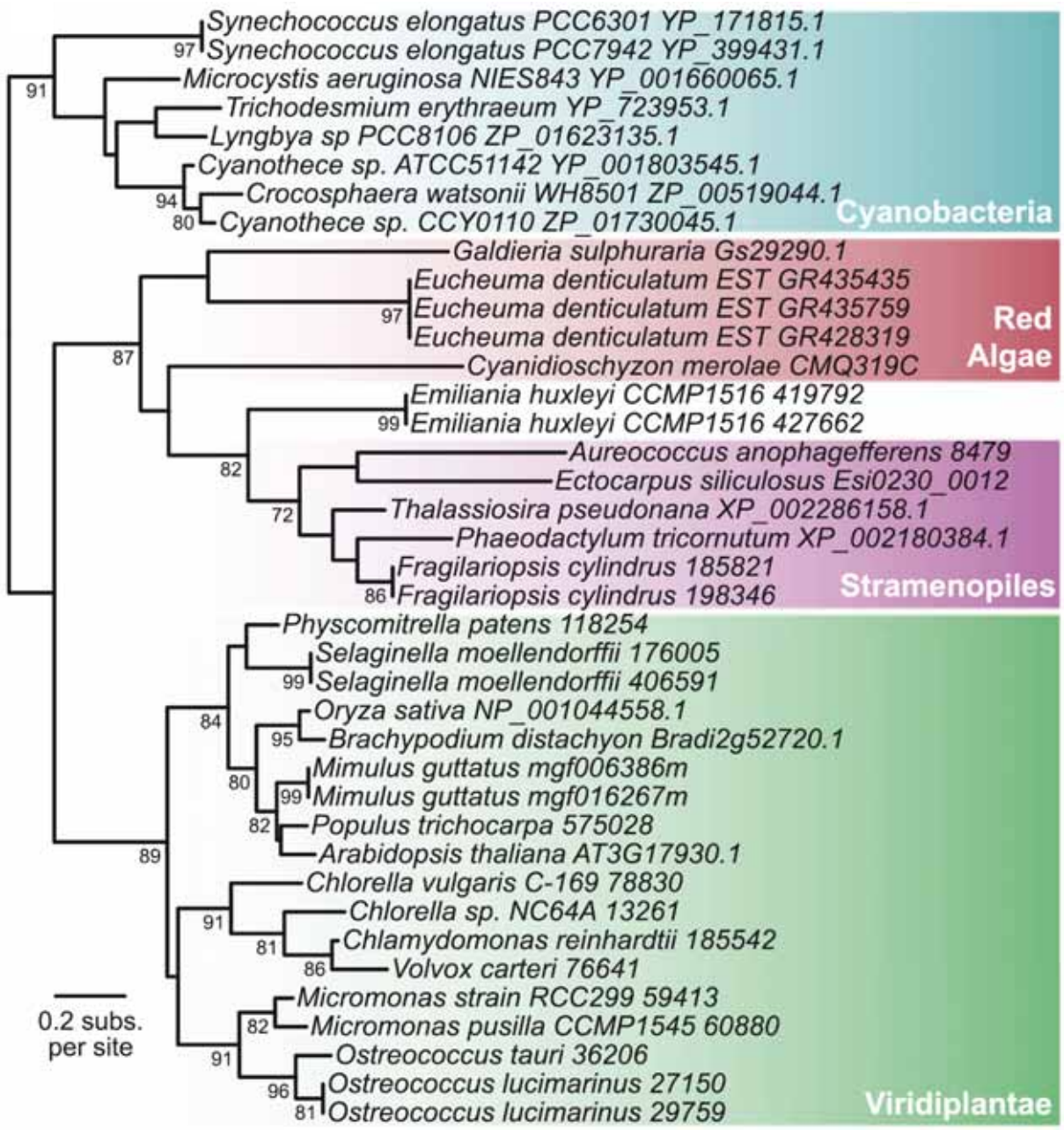


« Unresolved » case



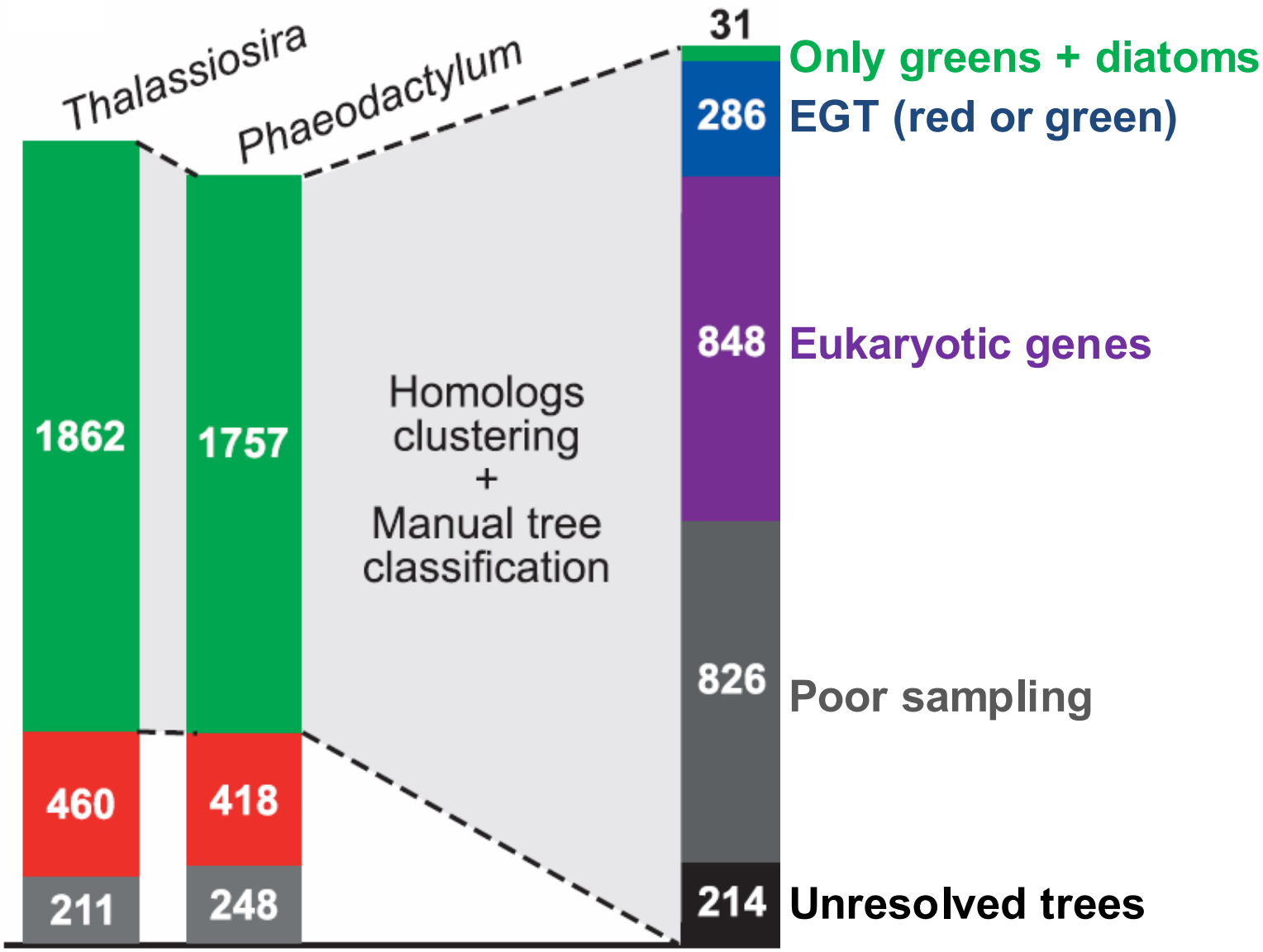
« Eukaryotic gene » case

(no evidence of EGT)

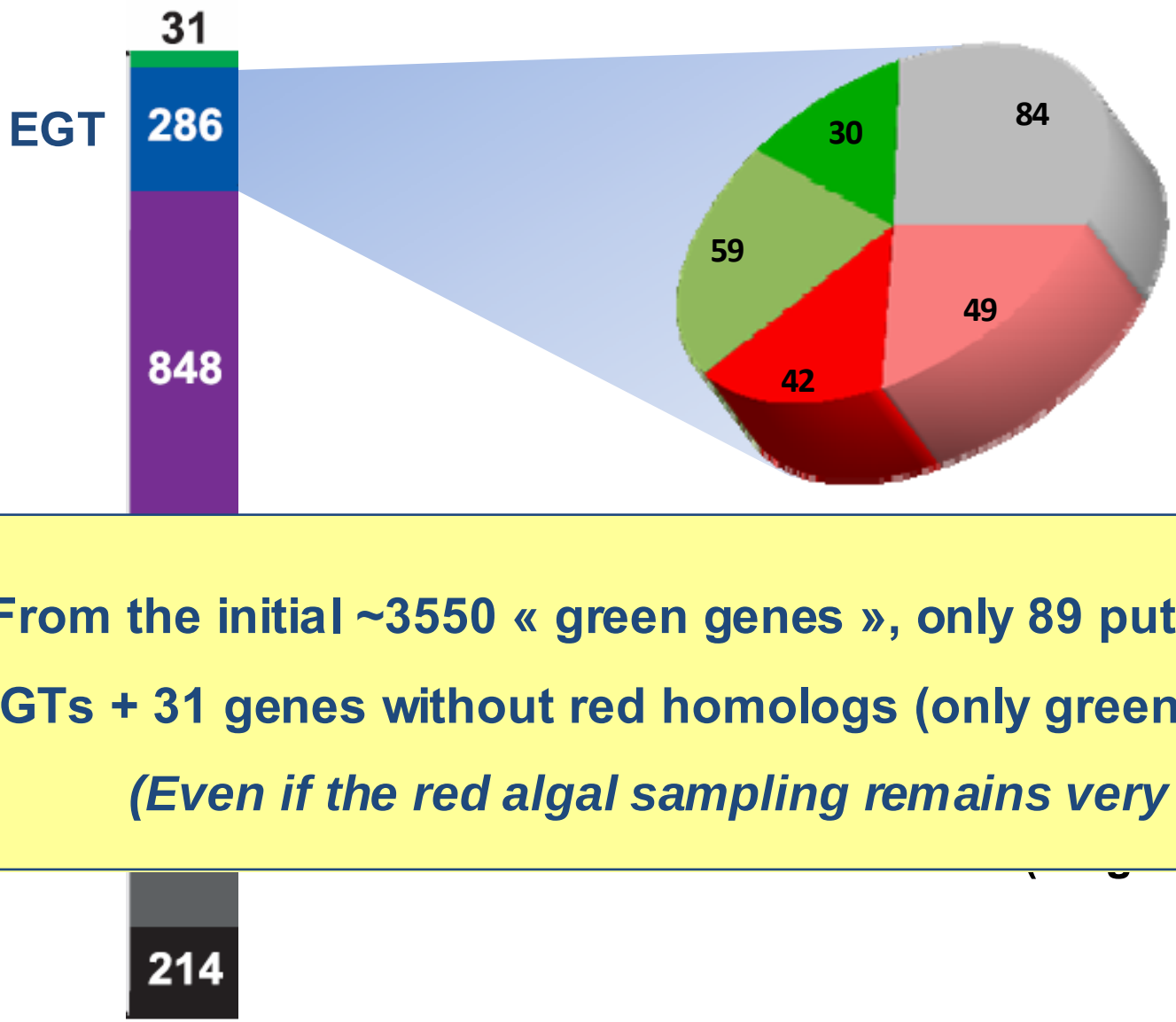


« EGT gene »
(red case)

Re-analyzing green genes in diatoms: summary



Re-analyzing green genes in diatoms: summary



From the initial ~3550 « green genes », only 89 putative green EGTs + 31 genes without red homologs (only greens + diatoms)

(Even if the red algal sampling remains very poor)

A similar example: *Chromera velia*

GBE

Red and Problematic Green Phylogenetic Signals among Thousands of Nuclear Genes from the Photosynthetic and Apicomplexa-Related *Chromera velia*

Christian Woehle, Tal Dagan, William F. Martin, and Sven B. Gould*

513 EGTs: 263 red genes vs. 250 green genes

But…

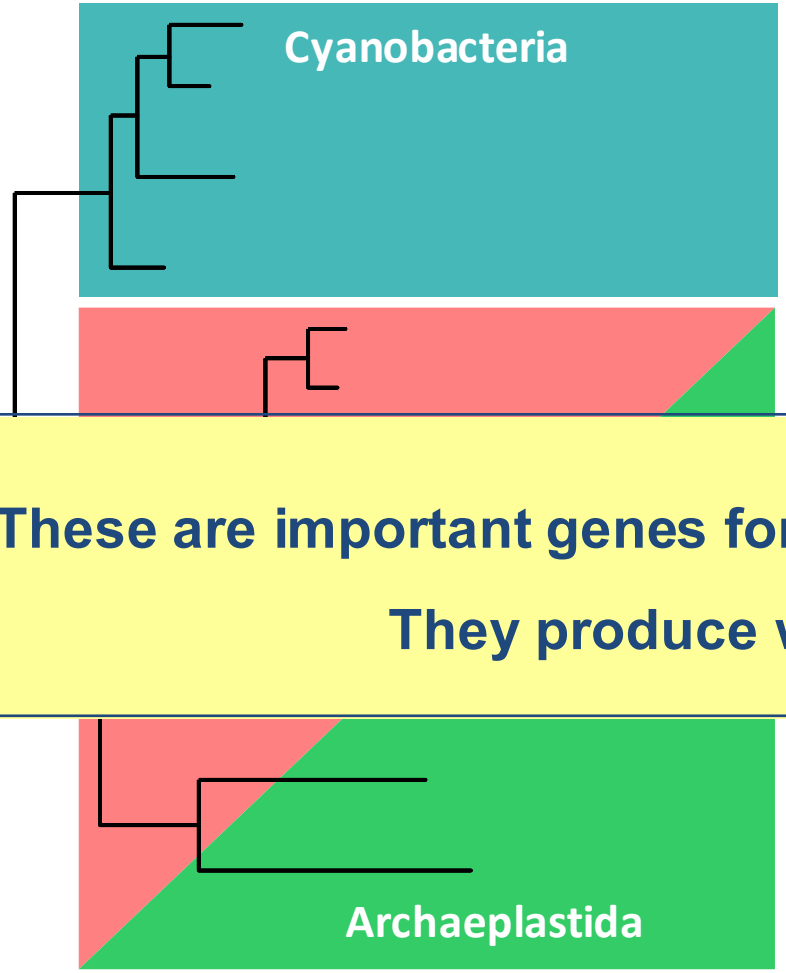
GBE

Re-evaluating the Green versus Red Signal in Eukaryotes with Secondary Plastid of Red Algal Origin

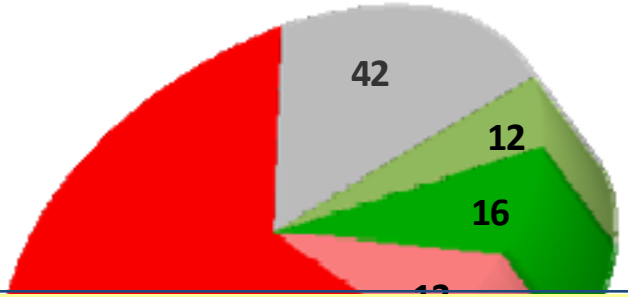
Fabien Burki[1,†], Pavel Flegontov[2,†], Miroslav Oborník[2,3,4], Jaromír Cihlář[2], Arnab Pain[5], Julius Lukeš[2,3], and Patrick J. Keeling[1,*]

Only 51 EGTs: 23 red genes vs. 9 green genes + 19 ambiguous

Genes with a complete EGT history



209 in diatoms



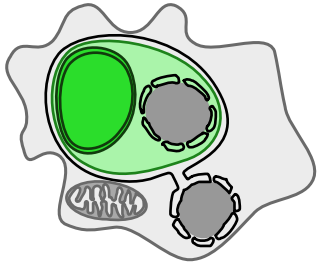
These are important genes for plastid function, well conserved.

They produce well resolved trees.

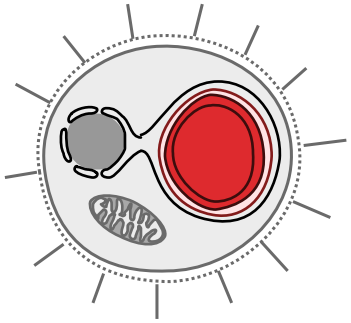
- Unresolved
- Green
- Putative green (no red in tree)
- Red
- Putative red (no green in tree)

From cyanobacteria to diatoms through the archaeplastida

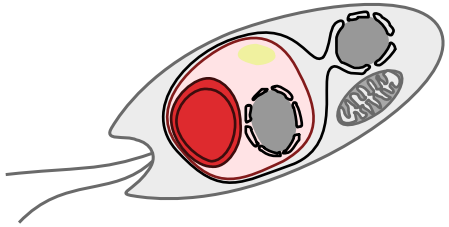
What about other photosynthetic eukaryotes?



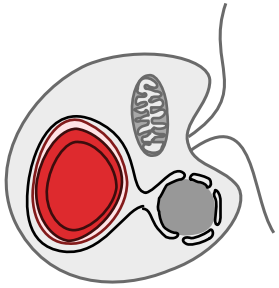
Chlorarachniophytes



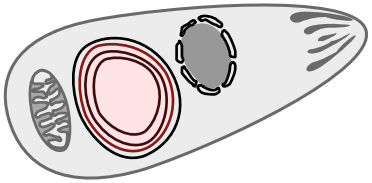
Haptophytes



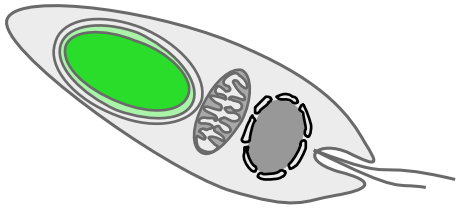
Cryptophytes



Stramenopiles

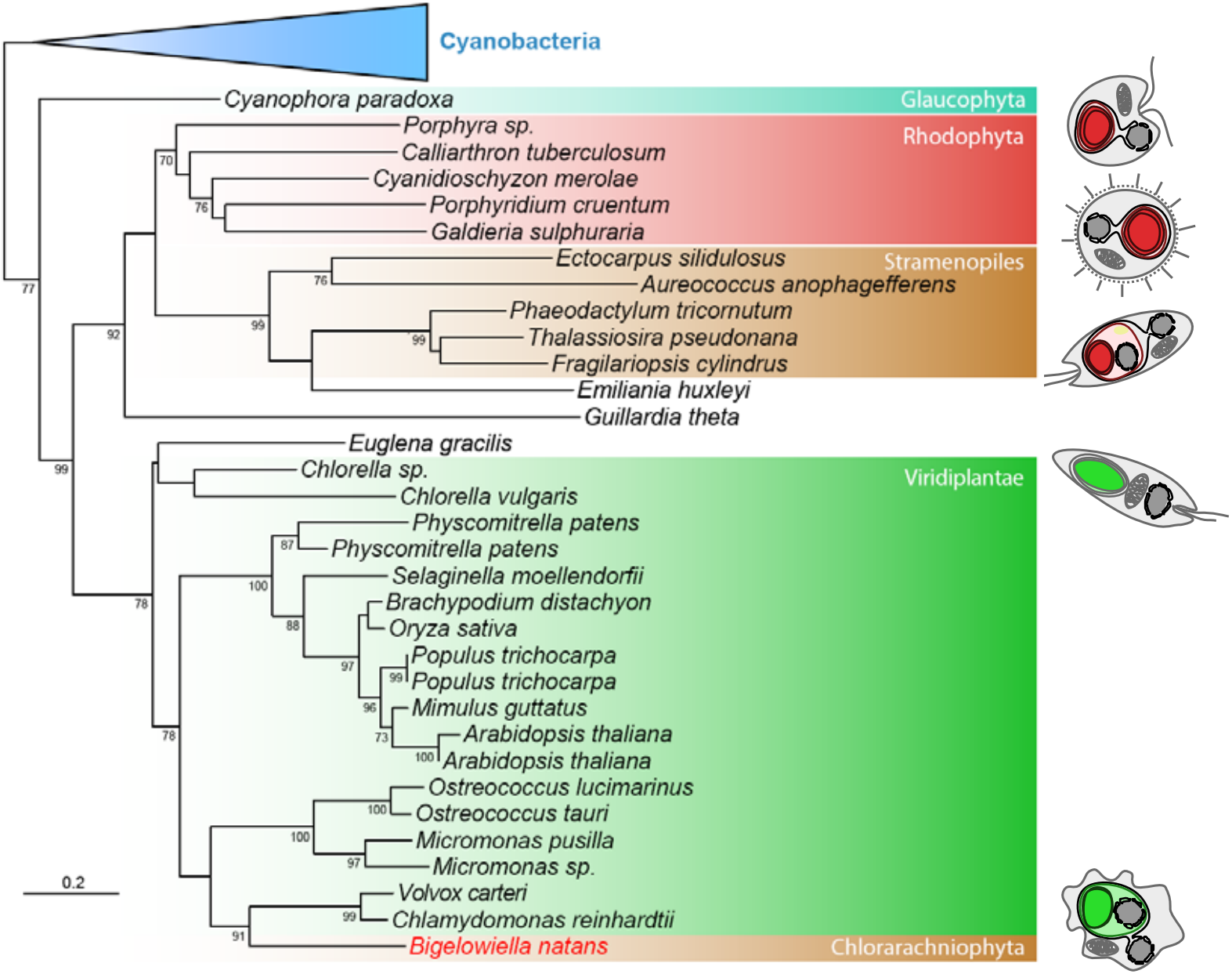


Apicomplexa

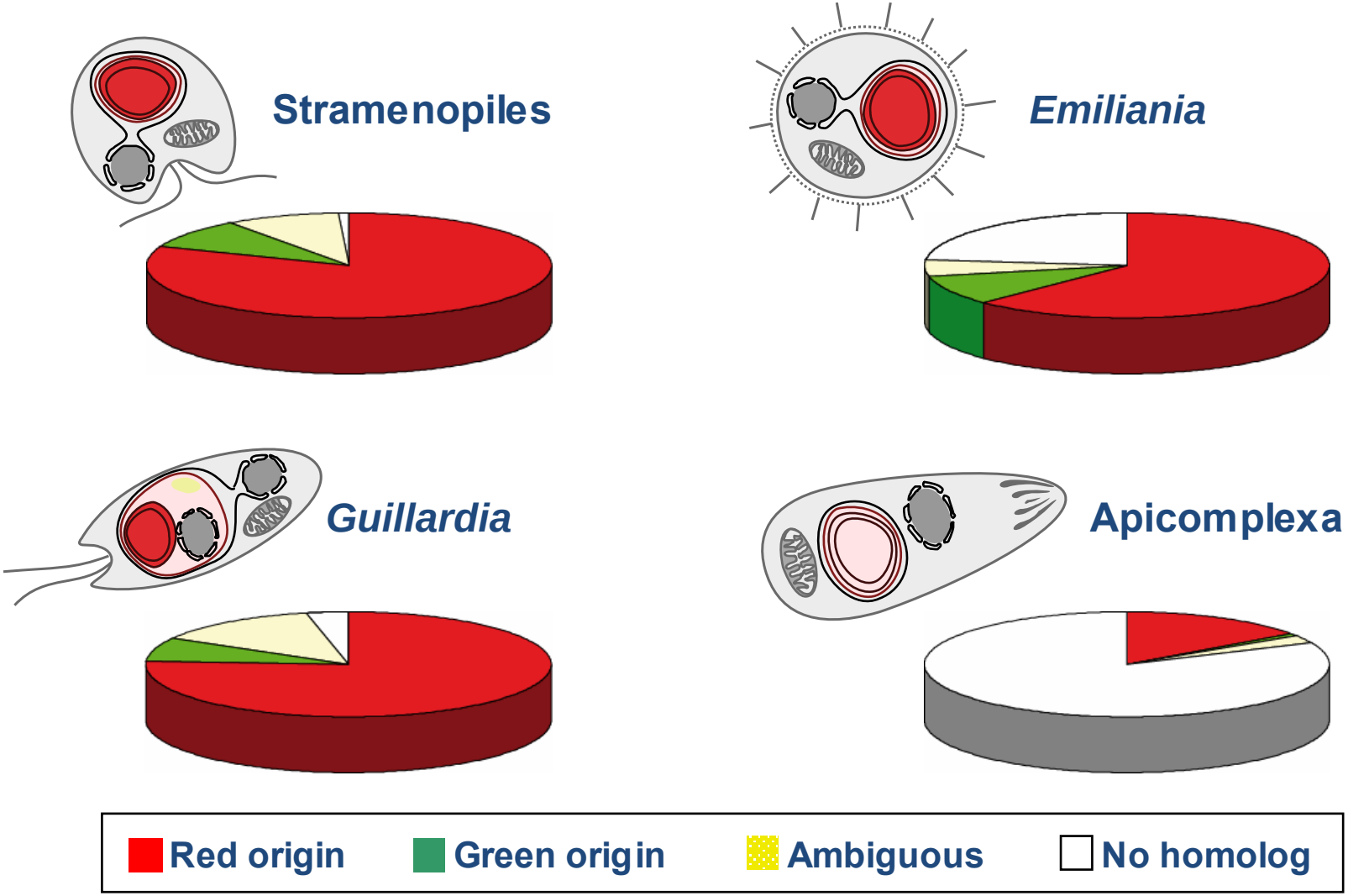


Euglenida

Analysis of 120 widespread genes with complete EGT history


Cyanobacteria
Glaucophyta
Cyanophora paradoxa
Rhodophyta
Porphyra sp.
Calliarthron tuberculosum
Cyanidioschyzon merolae
Porphyridium cruentum
Galdieria sulphuraria
Stramenopiles
Ectocarpus silidulosus
Aureococcus anophagefferens
Phaeodactylum tricornutum
Thalassiosira pseudonana
Fragilariopsis cylindrus
Emiliania huxleyi
Guillardia theta
Viridiplantae
Euglena gracilis
Chlorella sp.
Chlorella vulgaris
Physcomitrella patens
Physcomitrella patens
Selaginella moellendorfii
Brachypodium distachyon
Oryza sativa
Populus trichocarpa
Populus trichocarpa
Mimulus guttatus
Arabidopsis thaliana
Arabidopsis thaliana
Ostreococcus lucimarinus
Ostreococcus tauri
Micromonas pusilla
Micromonas sp.
Volvox carteri
Chlamydomonas reinhardtii
Chlorarachniophyta
Bigelowiella natans
0.2

Analysis of 120 widespread genes with complete EGT history



Small proportion (<10%) of « green genes »: Basal level of « phylogenetic noise » (HGT, hidden paralogies, etc.)

Conclusions

- EGTs are difficult to identify. Automatic analyses may detect many false positives.

- Several genes with a complete EGT history (cyanobacteria ➡ primary photosynthetic eukaryotes ➡ secondary photosynthetic eukaryotes) appear to be robust markers.