LIÈGE université
InBioS

Function over origin: kleptoplasty as a driver of gene mosaicism in complex algae

Mick Van Vlierberghe, Denis Baurain

University of Liège

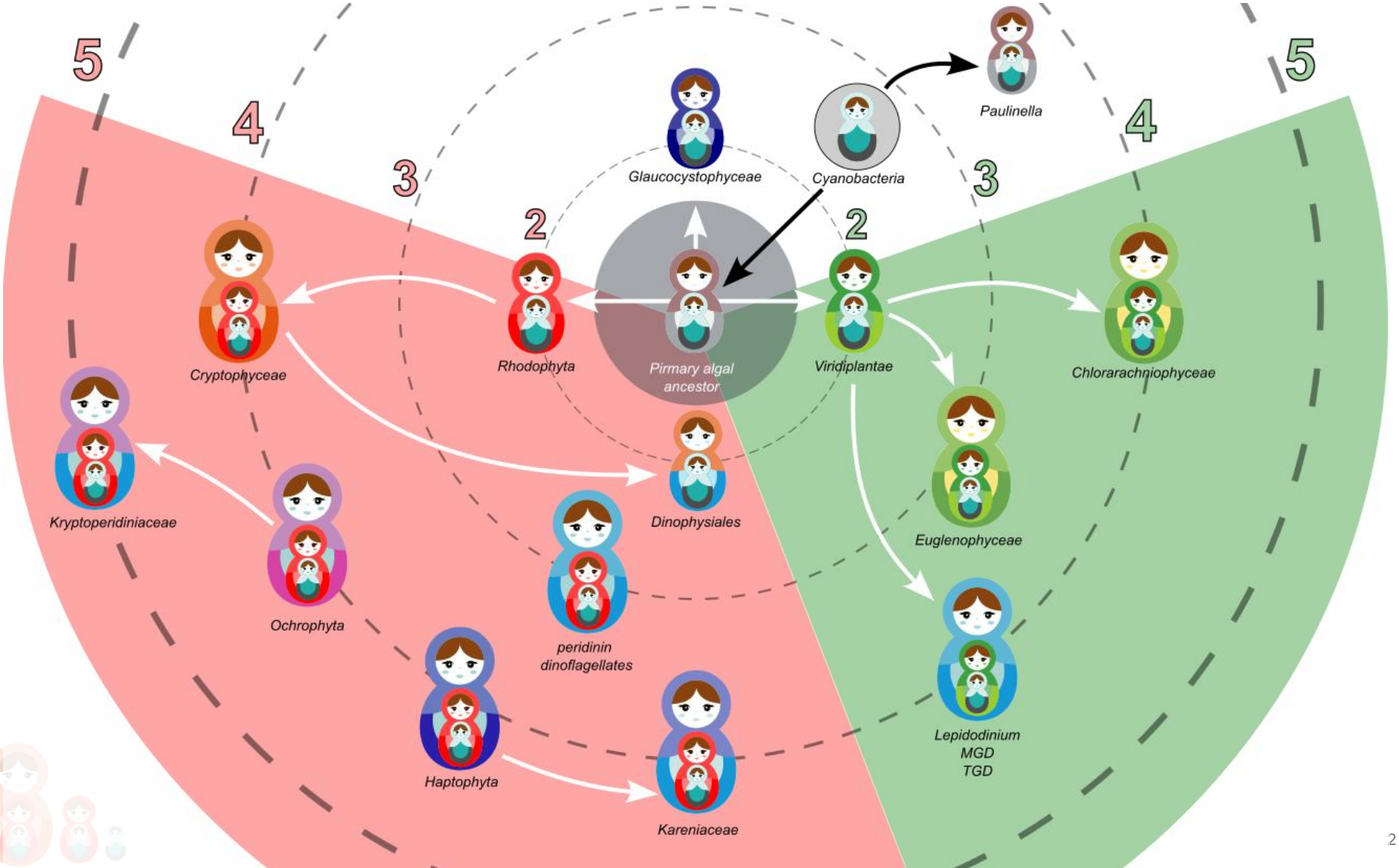
Eukaryotic Phylogenomics Unit



ISE 2026 Ostrava

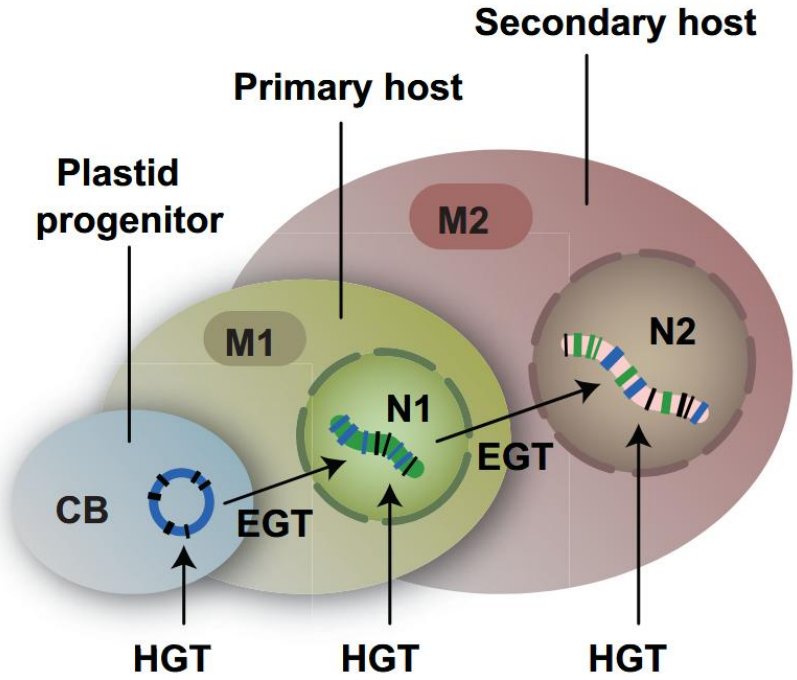
September 17th, 2026

Complex algal network



Observed nucpt# (nucleus-encoded plastid targeted) gene mosaicism

- Mosaicism is widespread
  - Nm-bearing algae (Curtis et al., 2012)
  - Euglenids (Maruyama et al., 2011)
  - Ochropyhytes (Dorrell et al., 2017)
  - Dinophyisales (Hongo et al., 2019)
  - *Rapaza viridis* (Karknowska et al., 2023)
  - …
- Gene transfer (EGT & HGT)
- Is there a mechanism favoring nucpt# mosaicism?
  - Endosymbiosis? Kleptoplasty? Other?
  - Quantify and compare mosaicism
  - Unified framework, single dataset



Elias and Archibald 2009

Quantitative phylogenomics analysis



PhD thesis *(publication in prep.)*

73 curated proteomes

Data note | Open access | Published: 17 April 2021

Broadly sampled orthologous groups of eukaryotic proteins for the phylogenetic study of plastid-bearing lineages

Mick Van Vlierberghe, Hervé Philippe & Denis Baurain

BMC Research Notes 14, Article number: 143 (2021) | Cite this article

1577 Accesses | 5 Citations | 2 Altmetric | Metrics

TAG-LOC

tagging OGs (nucpt#)

Data note | Open access | Published: 09 August 2021

Decontamination, pooling and dereplication of the 678 samples of the Marine Microbial Eukaryote Transcriptome Sequencing Project

Mick Van Vlierberghe, Arnaud Di Franco, Hervé Philippe & Denis Baurain

BMC Research Notes 14, Article number: 306 (2021) | Cite this article

1862 Accesses | 7 Citations | 2 Altmetric | Metrics

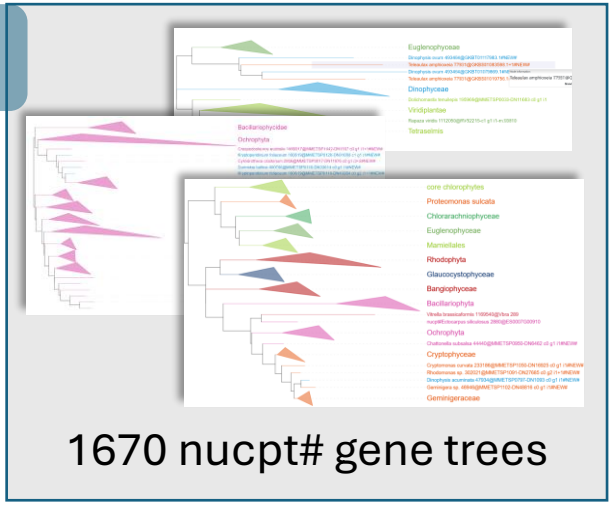
224 curated MMETSP transcriptomes

Forty-Two

dev. by Denis Baurain; in prep.



IQTREE
LG+C20+G



1670 nucpt# gene trees



Paraphyly-**A**ware
Gene **T**ransfer **I**nference

Quantitative phylogenomics analysis



PhD thesis *(publication in prep.)*

73 curated proteomes

Data note | Open access | Published: 17 April 2021

Broadly sampled orthologous groups of eukaryotic proteins for the phylogenetic study of plastid-bearing lineages

Mick Van Vlierberghe, Hervé Philippe & Denis Baurain

BMC Research Notes 14, Article number: 143 (2021) | Cite this article

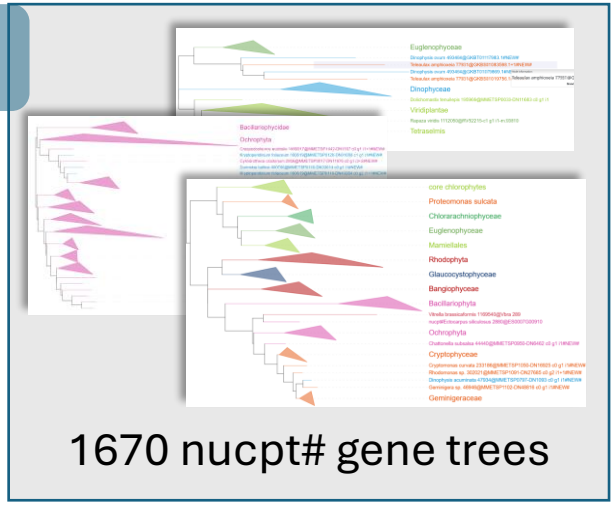
1577 Accesses | 5 Citations | 2 Altmetric | Metrics

TAG-LOC

tagging OGs (nucpt#)

Forty-Two

IQTREE
LG+C20+G



1670 nucpt# gene trees

JOURNAL ARTICLE

The P10K database: a data portal for the protist 10 000 genomes project

Xinxin Gao, Kai Chen, Jie Xiong, Dong Zou, Fangdian Yang, Yingke Ma, Chuanqi Jiang,

METdb: A GENOMIC REFERENCE DATABASE FOR MARINE SPECIES

Guita Niang, Mark Hoebeke, Arnaud Meng, Xi Liu, Maxim Scheremetjew, Rob Finn, Eric Pelletier[1], Erwan Corre[2]

PUBLISHED 06 JUN 2020 (https://doi.org/10.7490/f1000research.1118000.1)

New species?
Remake coming…

PAGTI

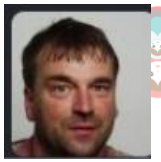
Paraphyly-Aware Gene Transfer Inference

PERL

Paraphyly-**A**ware
Gene **T**ransfer **I**nference

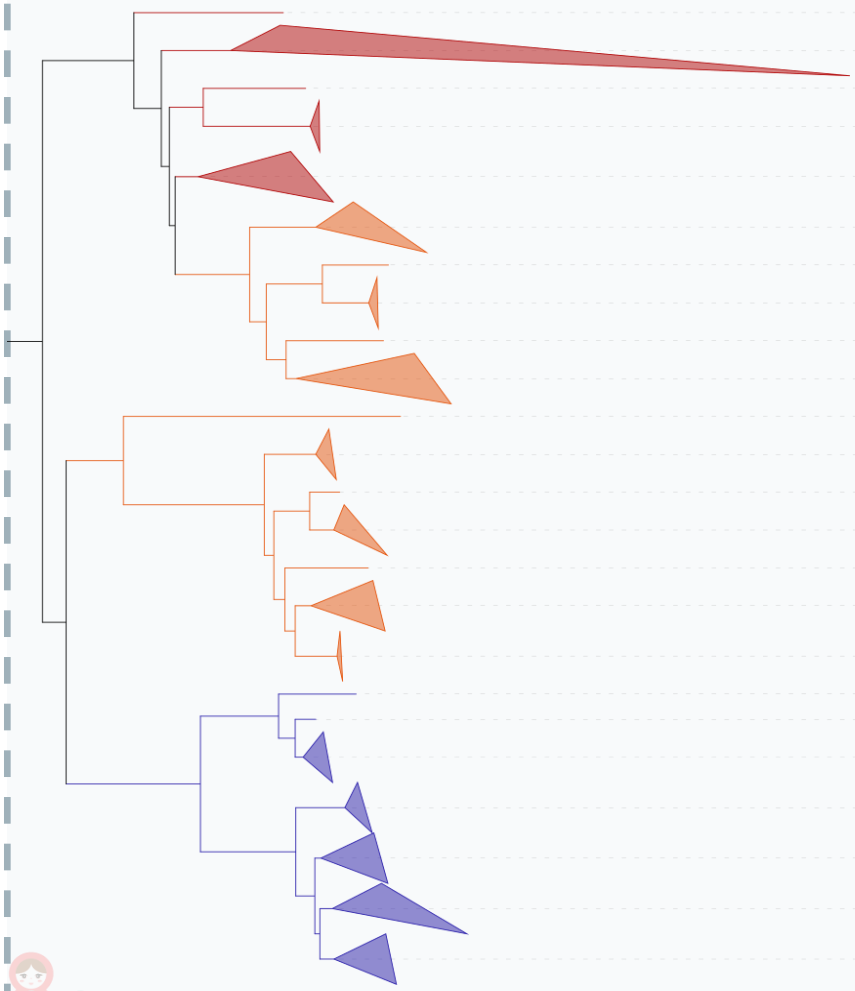
Keeling et al., 2014; Johnson et al. 2016

PAGTI Perl



manual validation
Henner Brinkmann, DSMZ

Paraphily-Aware Gene Transfer Inference



Rhodella violacea 38298@MMETSP0167-DN10244 c1 g1 i1#NEW#
Porphyridiaceae
Chondrus crispus 2769@XP 005717389.1
Bangiaceae
Bangiophyceae
Cryptomonas
Chroomonas mesostigmatica 195065@MMETSP0047-DN21743 c1 g1 i1#NEW#
Hemiselmis
Rhodomonas sp. 302021@MMETSP1091-DN28815 c16 g1 i1#NEW#
Geminigeraceae
Goniomonas avonlea 1255295@comp62657 c0 seq1 6 ORF28 816
Cryptomonas
Chroomonas mesostigmatica 195065@MMETSP0047-DN23805 c0 g1 i1#NEW#
Hemiselmis
Geminigera sp. 46946@MMETSP1102-DN38556 c0 g1 i1#NEW#
Pyrenomonadales
Geminigeraceae
Pavlovales sp. 483367@MMETSP0984-DN6682 c0 g1 i1#NEW#
Chrysoculter rhomboideus 259385@MMETSP1335-DN14908 c0 g1 i1#NEW#
Pavlovaceae
Phaeocystis
Prymnesiales
Prymnesiophyceae
Isochrysidales

Paraphyletic Relationship = Directional Transfer

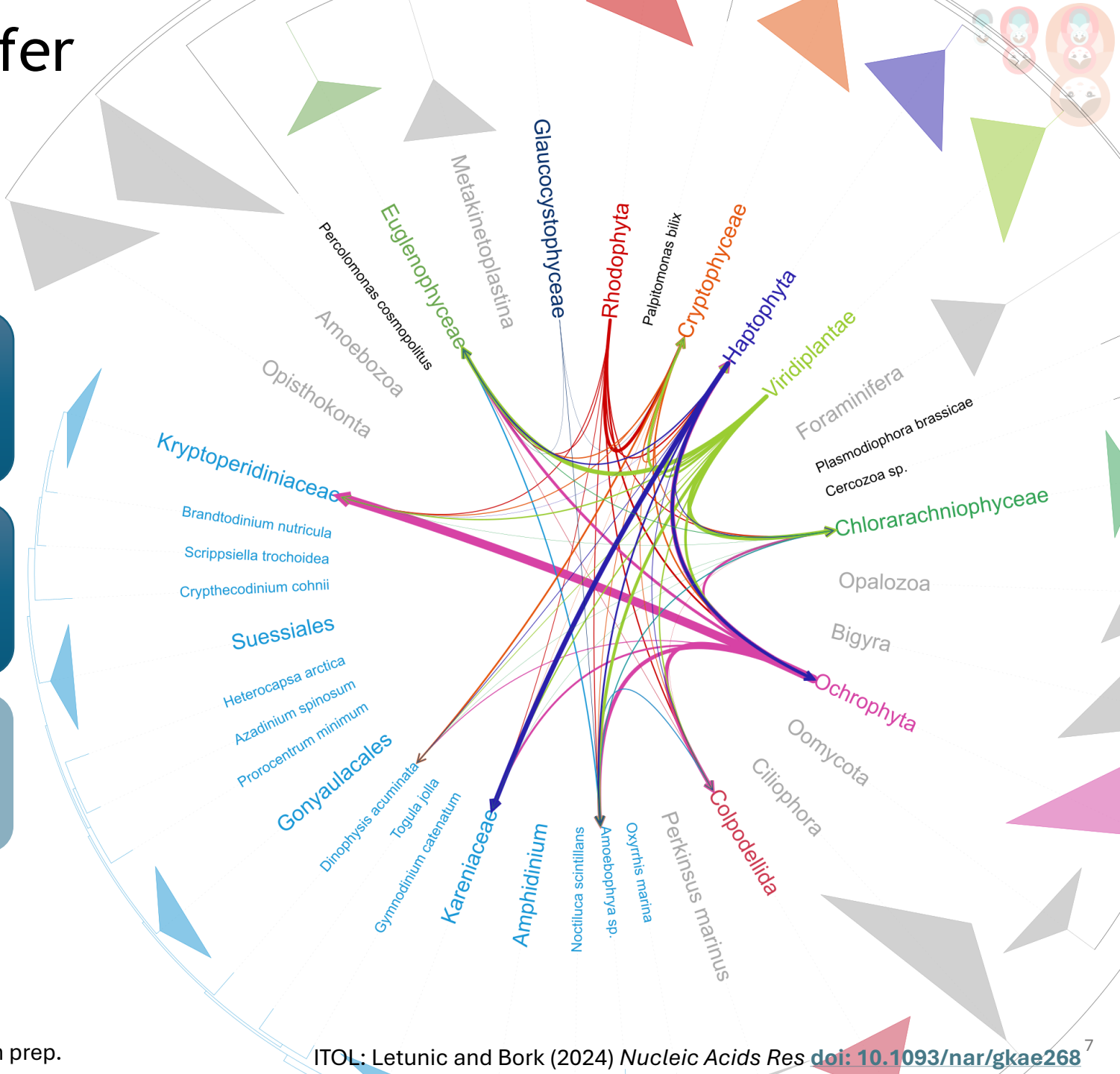
Sister Group Relationship ≠ Transfer

Gene transfer network

1670 nucpt# OGs
3353 transfers

42 primary algae
169 complex algae

78 ribosomal markers
293 species
12,224 sites
IQTREE LG+C20+G



This study; in prep.

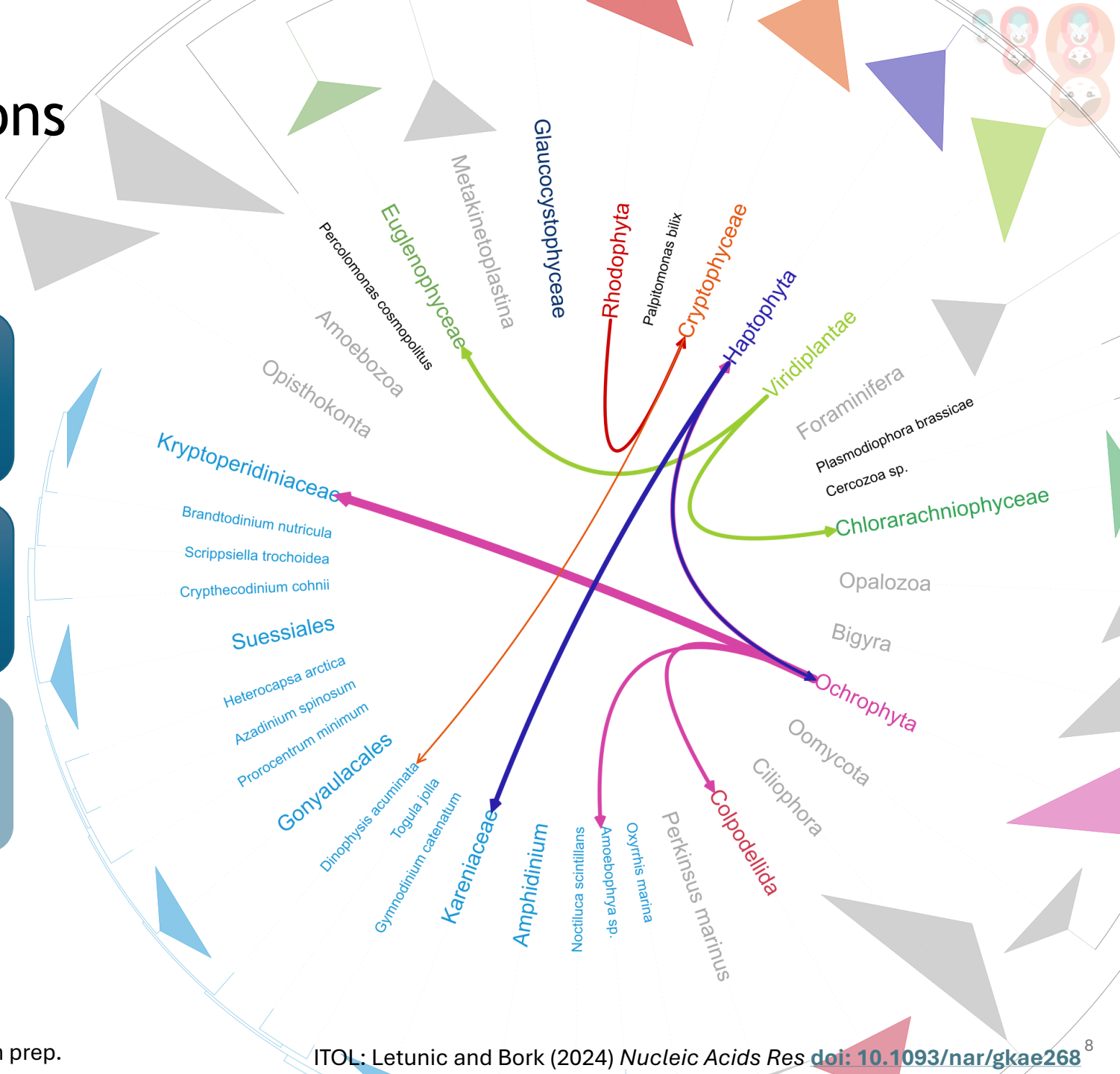
ITOL: Letunic and Bork (2024) *Nucleic Acids Res* doi: 10.1093/nar/gkae268

Major contributions

1670 nucpt# OGs
3353 transfers

45 primary algae
161 complex algae

78 ribosomal markers
293 species
12,224 sites
IQTREE LG+C20+G



This study; in prep.

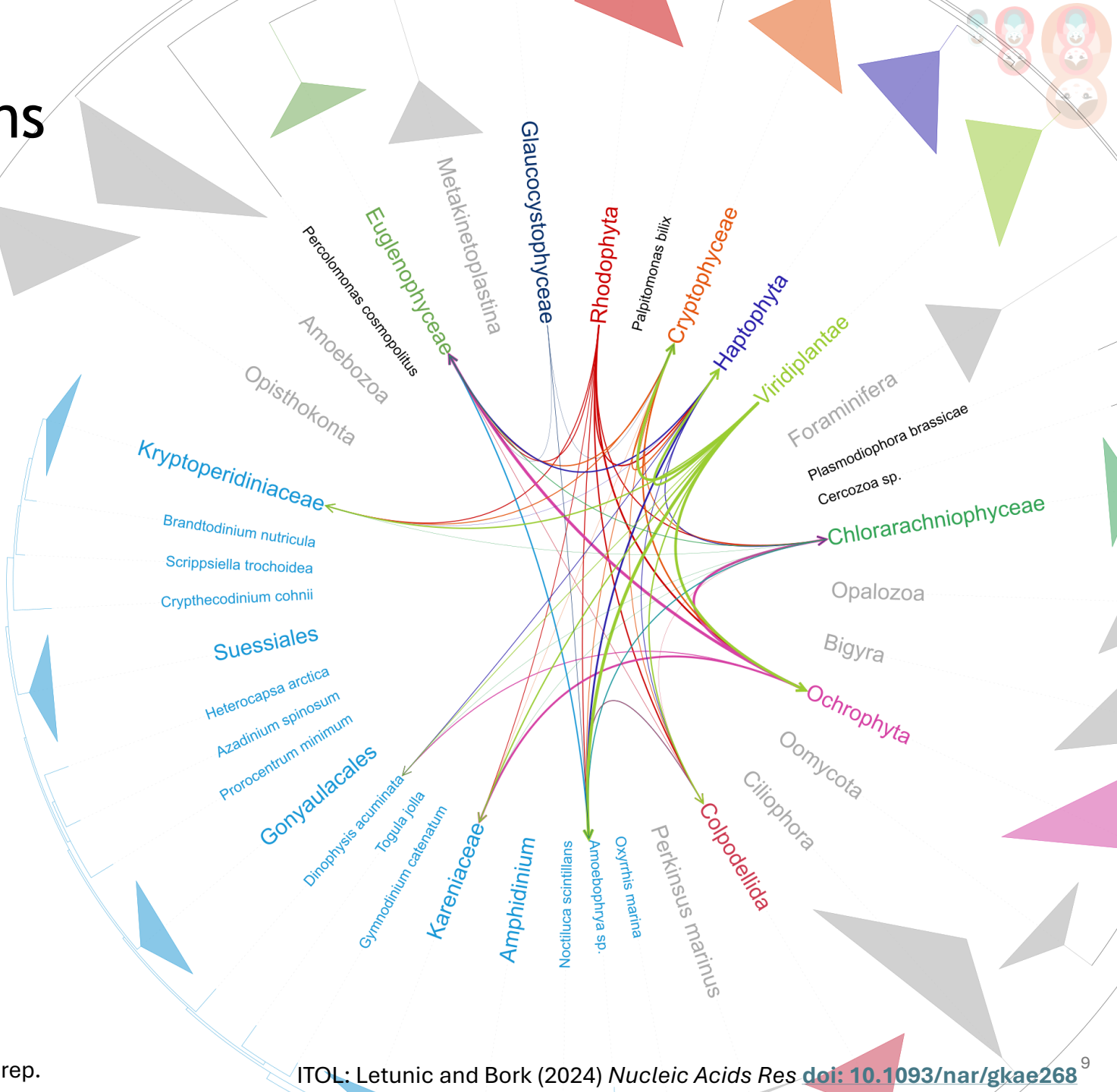
ITOL: Letunic and Bork (2024) *Nucleic Acids Res* doi: 10.1093/nar/gkae268

Minor contributions

1670 nucpt# OGs
3353 transfers

45 primary algae
161 complex algae

78 ribosomal markers
293 species
12,224 sites
IQTREE LG+C20+G



This study; in prep.

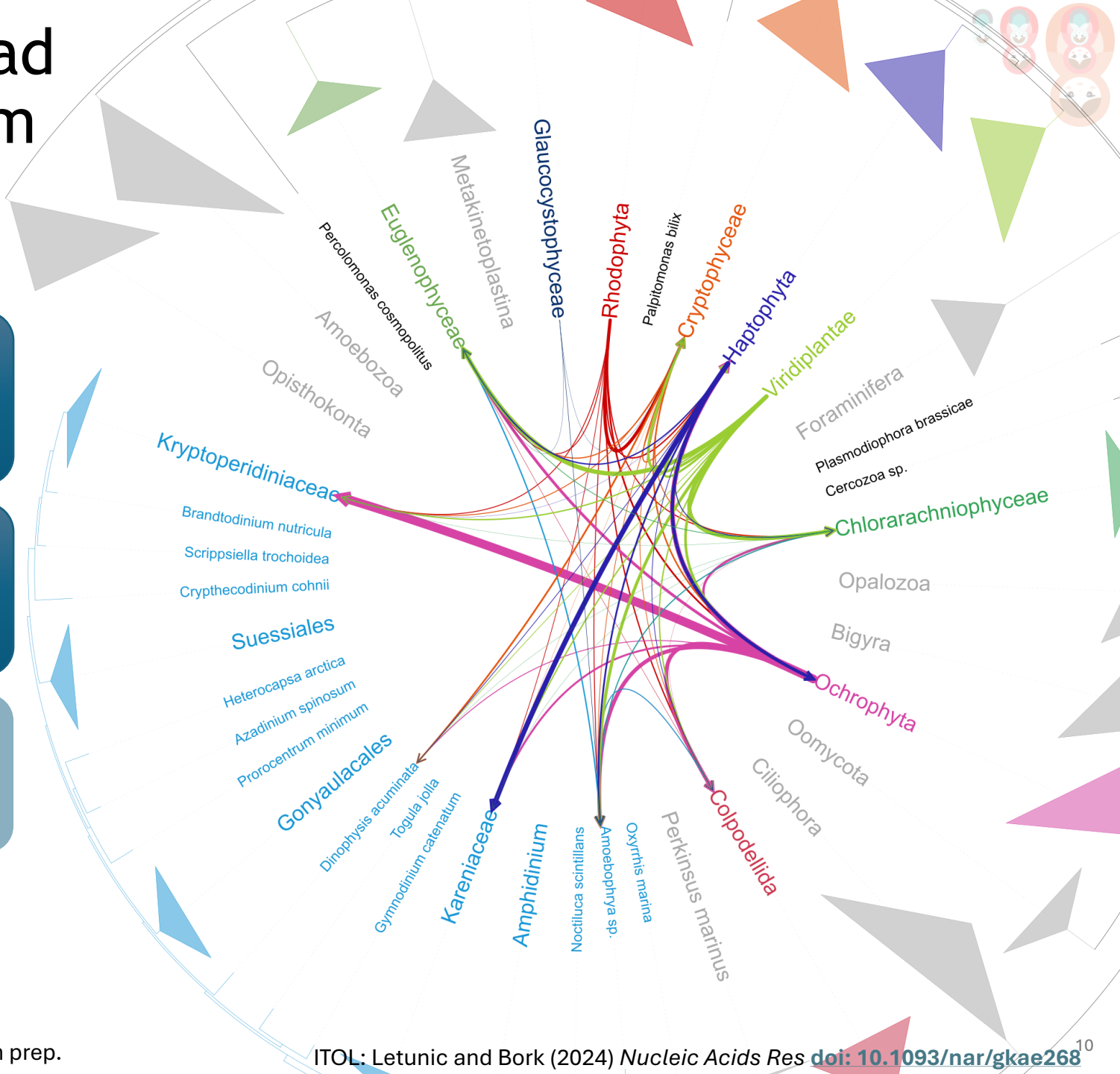
ITOL: Letunic and Bork (2024) *Nucleic Acids Res* doi: 10.1093/nar/gkae268

Widespread mosaicism

1670 nucpt# OGs
3353 transfers

45 primary algae
161 complex algae

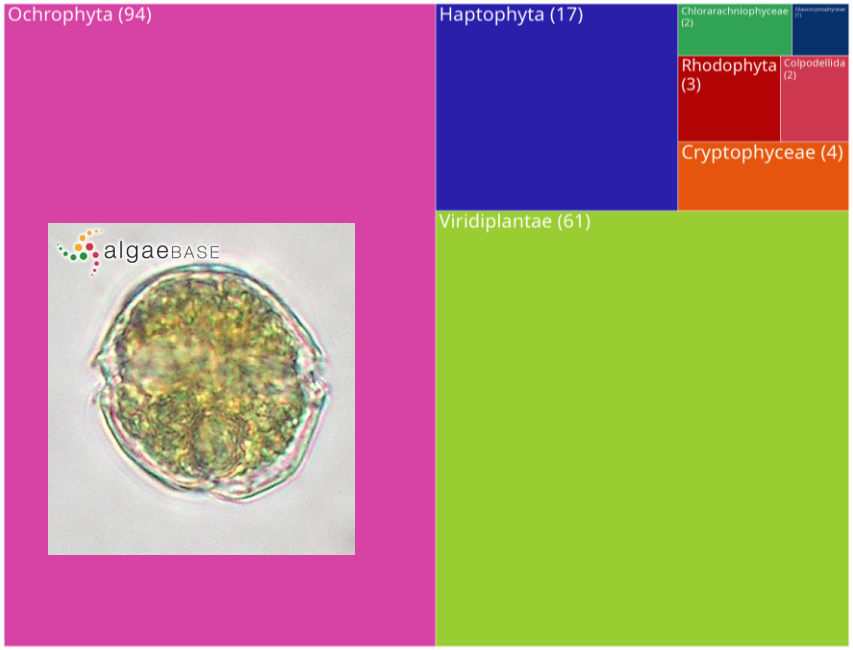
78 ribosomal markers
293 species
12,224 sites
IQTREE LG+C20+G



This study; in prep.

ITOL: Letunic and Bork (2024) *Nucleic Acids Res* doi: 10.1093/nar/gkae268

Quantify and compare mosaicism



- *Alexandrium tamarense* (peridinin dino)
- Gene transfer events = 182
- Shannon-Wiener's index (H') = 1.41



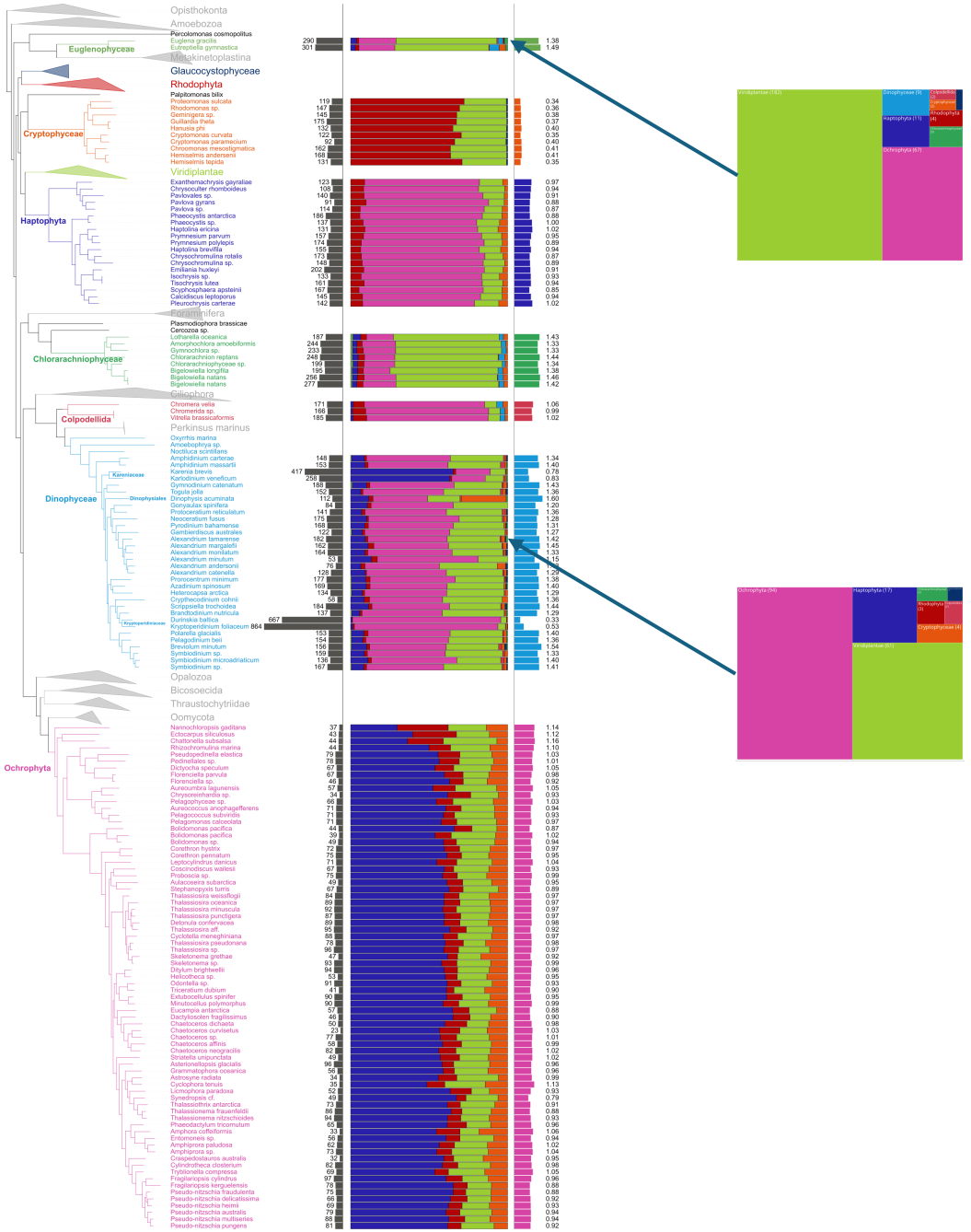
- *Euglena gracilis* (Euglenophyceae)
- Gene transfer events = 290
- Shannon-Wiener's index (H') = 1.38

This study; in prep.

Per species α-diversity

- Euglenozoa
- Haptophyta
- Rhodophyta
- Ochrophyta
- Viridiplantae
- Colpodellida
- Dinophyceae
- Cryptophyceae
- Glaucocystophyceae
- Chlorarachniophyceae

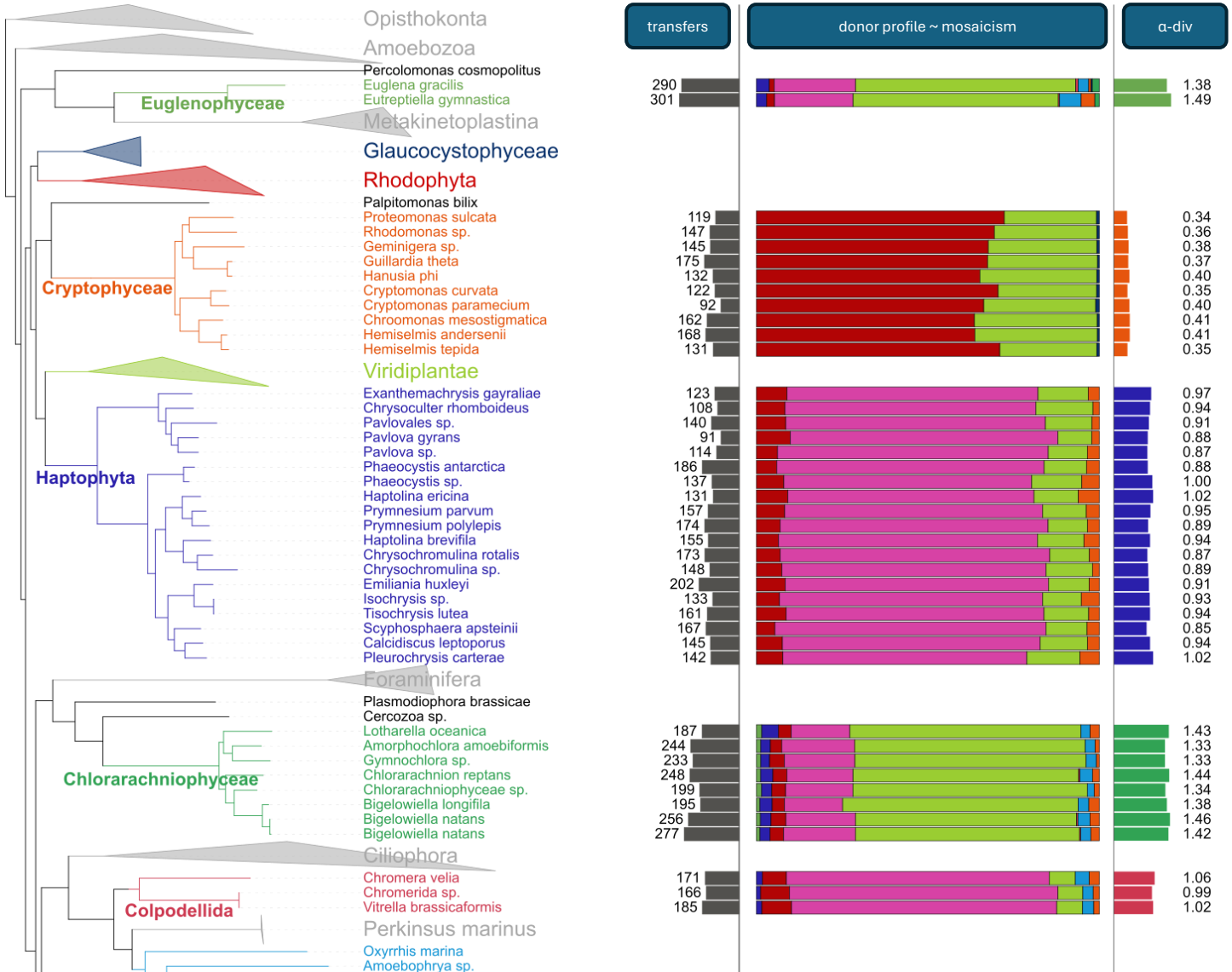
78 ribosomal markers
293 species
12,224 sites
IQTREE LG+C20+G

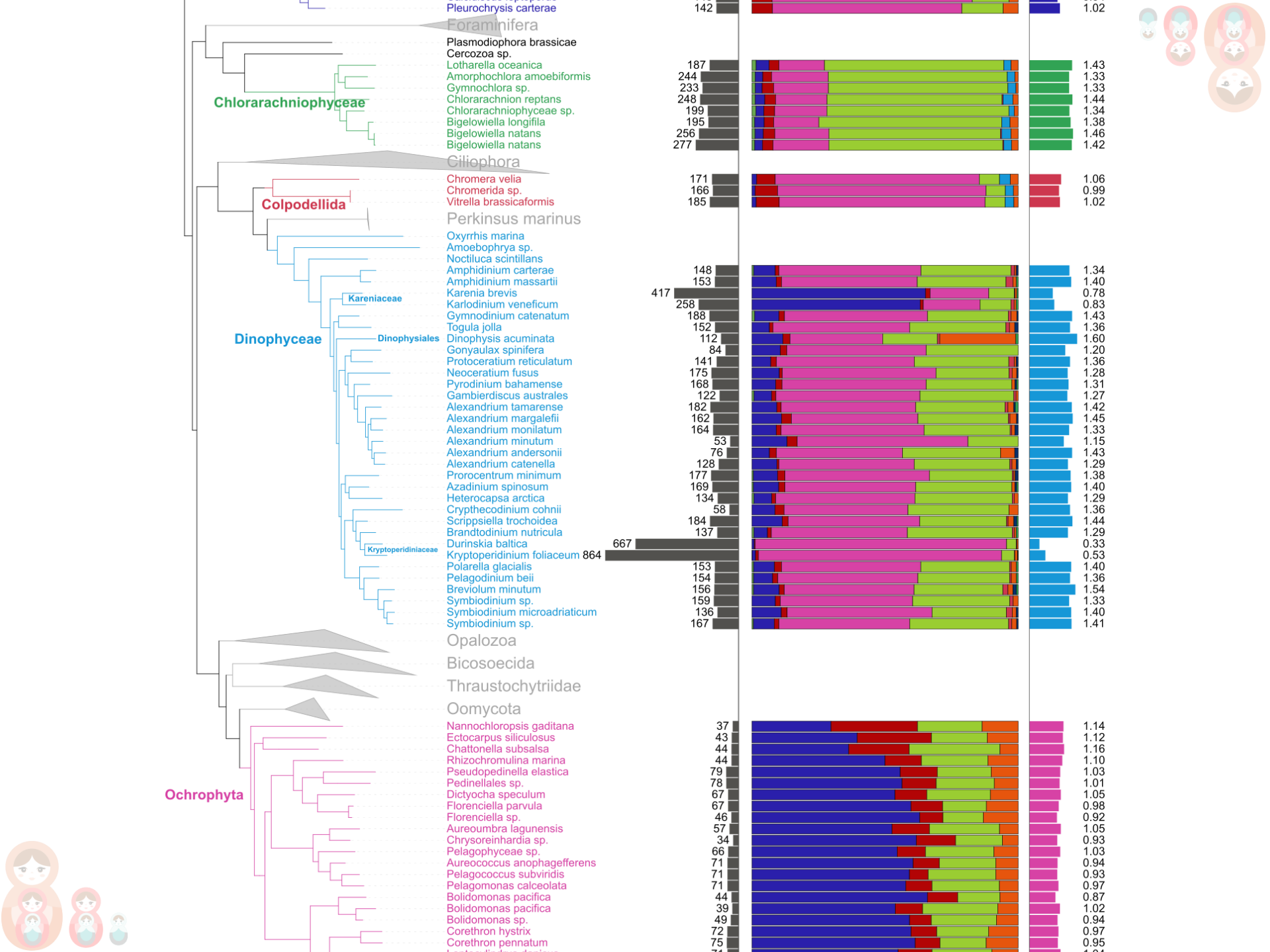


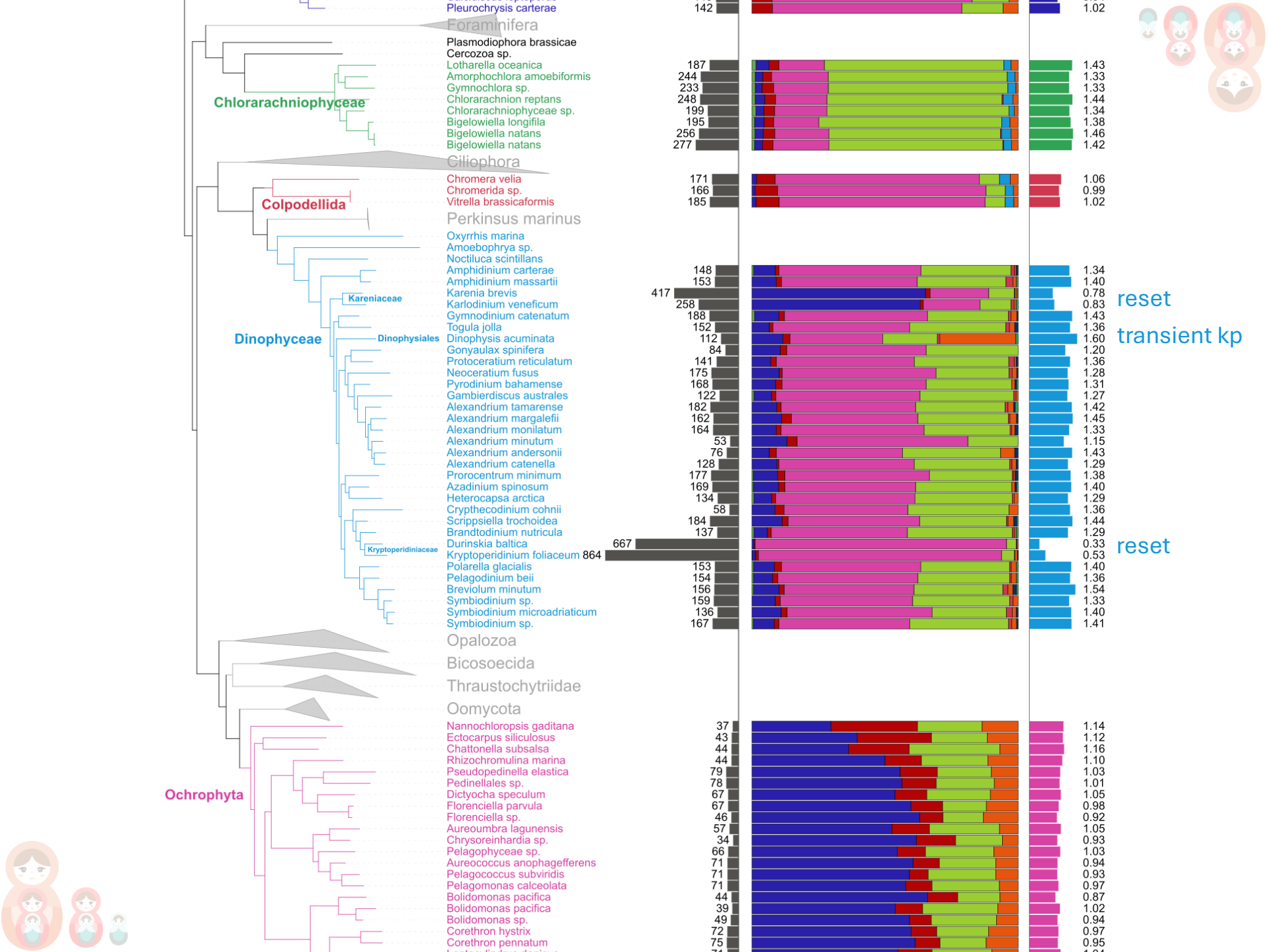
This study; in prep.

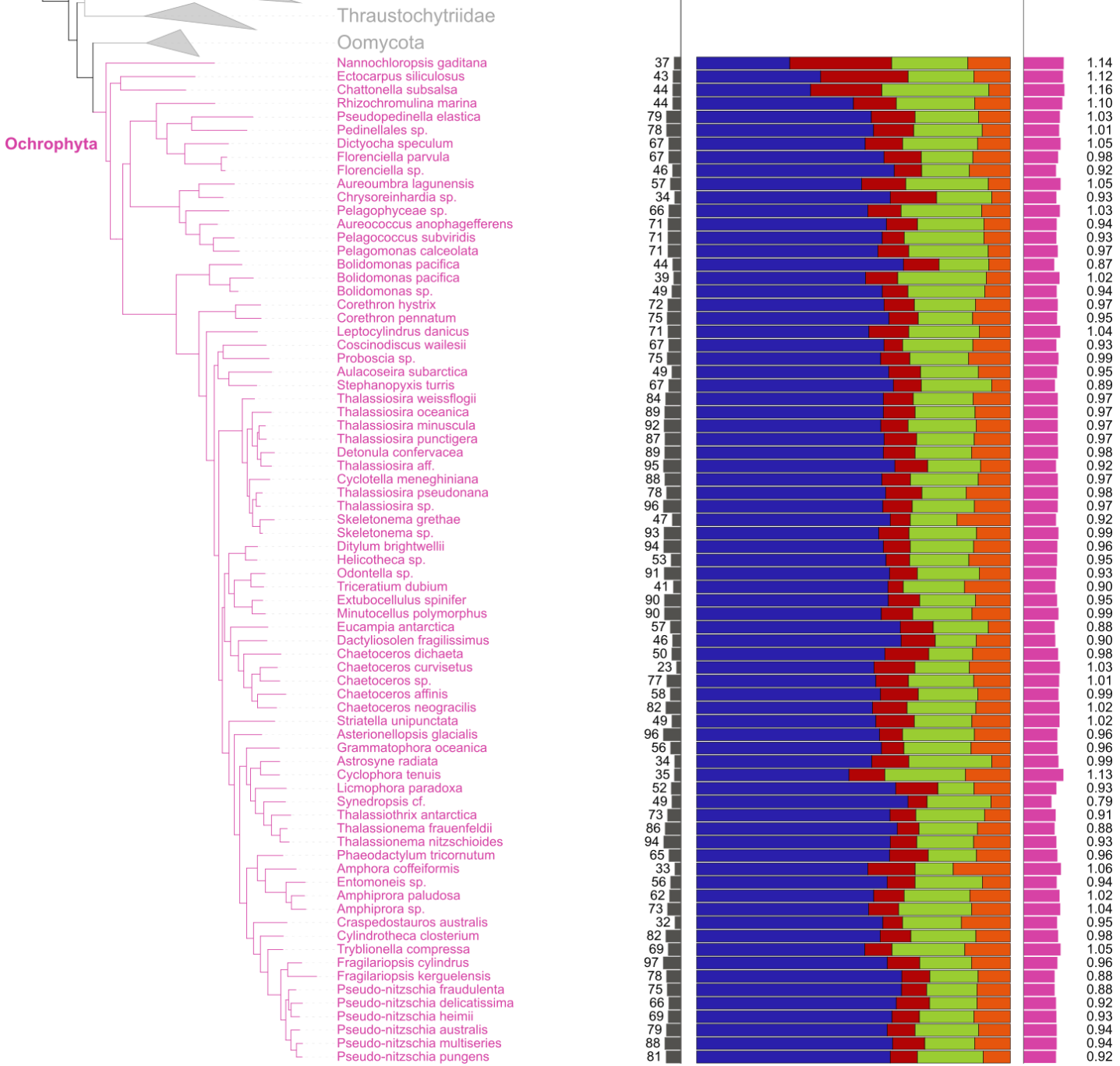
ITOL: Letunic and Bork (2024) *Nucleic Acids Res* doi: 10.1093/nar/gkae268

Consitent intra-lineage donor profiles




Pleurochrysis carterae
142
1.02
Foraminifera
Plasmodiophora brassicae
Cercozoa sp.
Lotharella oceanica
Amorphochlora amoebiformis
Gymnochlora sp.
Chlorarachnion reptans
Chlorarachniophyceae sp.
Bigelowiella longifila
Bigelowiella natans
Bigelowiella natans
Chlorarachniophyceae
187
244
233
248
199
195
256
277
1.43
1.33
1.33
1.44
1.34
1.38
1.46
1.42
Ciliophora
Chromera velia
Chromerida sp.
Vitrella brassicaformis
Colpodellida
171
166
185
1.06
0.99
1.02
Perkinsus marinus
Oxyrrhis marina
Amoebophrya sp.
Noctiluca scintillans
Amphidinium carterae
Amphidinium massartii
Karenia brevis
Karlodinium veneficum
Gymnodinium catenatum
Togula jolla
Dinophysis acuminata
Gonyaulax spinifera
Protoceratium reticulatum
Neoceratium fusus
Pyrodinium bahamense
Gambierdiscus australes
Alexandrium tamarense
Alexandrium margalefii
Alexandrium monilatum
Alexandrium minutum
Alexandrium andersonii
Alexandrium catenella
Prorocentrum minimum
Azadinium spinosum
Heterocapsa arctica
Crypthecodinium cohnii
Scrippsiella trochoidea
Brandtodinium nutricula
Durinskia baltica
Kryptoperidinium foliaceum
Polarella glacialis
Pelagodinium beii
Breviolum minutum
Symbiodinium sp.
Symbiodinium microadriaticum
Symbiodinium sp.
Kareniaceae
Dinophysiales
Kryptoperidiniaceae
Dinophyceae
148
153
417
258
188
152
112
84
141
175
168
122
182
162
164
53
76
128
177
169
134
58
184
137
667
864
153
154
156
159
136
167
1.34
1.40
0.78
0.83
1.43
1.36
1.60
1.20
1.36
1.28
1.31
1.27
1.42
1.45
1.33
1.15
1.43
1.29
1.38
1.40
1.29
1.36
1.44
1.29
0.33
0.53
1.40
1.36
1.54
1.33
1.40
1.41
Opalozoa
Bicosoecida
Thraustochytriidae
Oomycota
Nannochloropsis gaditana
Ectocarpus siliculosus
Chattonella subsalsa
Rhizochromulina marina
Pseudopedinella elastica
Pedinellales sp.
Dictyocha speculum
Florenciella parvula
Florenciella sp.
Aureoumbra lagunensis
Chrysoreinhardia sp.
Pelagophyceae sp.
Aureococcus anophagefferens
Pelagococcus subviridis
Pelagomonas calceolata
Bolidomonas pacifica
Bolidomonas pacifica
Bolidomonas sp.
Corethron hystrix
Corethron pennatum
Ochrophyta
37
43
44
44
79
78
67
67
46
57
34
66
71
71
71
44
39
49
72
75
1.14
1.12
1.16
1.10
1.03
1.01
1.05
0.98
0.92
1.05
0.93
1.03
0.94
0.93
0.97
0.87
1.02
0.94
0.97
0.95


Pleurochrysis carterae
Foraminifera
Plasmodiophora brassicae
Cercozoa sp.
Lotharella oceanica
Amorphochlora amoebiformis
Gymnochlora sp.
Chlorarachnion reptans
Chlorarachniophyceae sp.
Bigelowiella longifila
Bigelowiella natans
Bigelowiella natans
Chlorarachniophyceae
Ciliophora
Chromera velia
Chromerida sp.
Vitrella brassicaformis
Colpodellida
Perkinsus marinus
Oxyrrhis marina
Amoebophrya sp.
Noctiluca scintillans
Amphidinium carterae
Amphidinium massartii
Karenia brevis
Karlodinium veneficum
Gymnodinium catenatum
Togula jolla
Dinophysis acuminata
Gonyaulax spinifera
Protoceratium reticulatum
Neoceratium fusus
Pyrodinium bahamense
Gambierdiscus australes
Alexandrium tamarense
Alexandrium margalefii
Alexandrium monilatum
Alexandrium minutum
Alexandrium andersonii
Alexandrium catenella
Prorocentrum minimum
Azadinium spinosum
Heterocapsa arctica
Crypthecodinium cohnii
Scrippsiella trochoidea
Brandtodinium nutricula
Durinskia baltica
Kryptoperidinium foliaceum
Polarella glacialis
Pelagodinium beii
Breviolum minutum
Symbiodinium sp.
Symbiodinium microadriaticum
Symbiodinium sp.
Kareniaceae
Dinophysiales
Kryptoperidiniaceae
Dinophyceae
Opalozoa
Bicosoecida
Thraustochytriidae
Oomycota
Nannochloropsis gaditana
Ectocarpus siliculosus
Chattonella subsalsa
Rhizochromulina marina
Pseudopedinella elastica
Pedinellales sp.
Dictyocha speculum
Florenciella parvula
Florenciella sp.
Aureoumbra lagunensis
Chrysoreinhardia sp.
Pelagophyceae sp.
Aureococcus anophagefferens
Pelagococcus subviridis
Pelagomonas calceolata
Bolidomonas pacifica
Bolidomonas pacifica
Bolidomonas sp.
Corethron hystrix
Corethron pennatum
Ochrophyta
142 1.02
187 1.43
244 1.33
233 1.33
248 1.44
199 1.34
195 1.38
256 1.46
277 1.42
171 1.06
166 0.99
185 1.02
148 1.34
153 1.40
417 0.78
258 0.83
188 1.43
152 1.36
112 1.60
84 1.20
141 1.36
175 1.28
168 1.31
122 1.27
182 1.42
162 1.45
164 1.33
53 1.15
76 1.43
128 1.29
177 1.38
169 1.40
134 1.29
58 1.36
184 1.44
137 1.29
667 0.33
864 0.53
153 1.40
154 1.36
156 1.54
159 1.33
136 1.40
167 1.41
37 1.14
43 1.12
44 1.16
44 1.10
79 1.03
78 1.01
67 1.05
67 0.98
46 0.92
57 1.05
34 0.93
66 1.03
71 0.94
71 0.93
71 0.97
44 0.87
39 1.02
49 0.94
72 0.97
75 0.95
reset
transient kp
reset



This study; in prep.

Per **species** α-diversity values

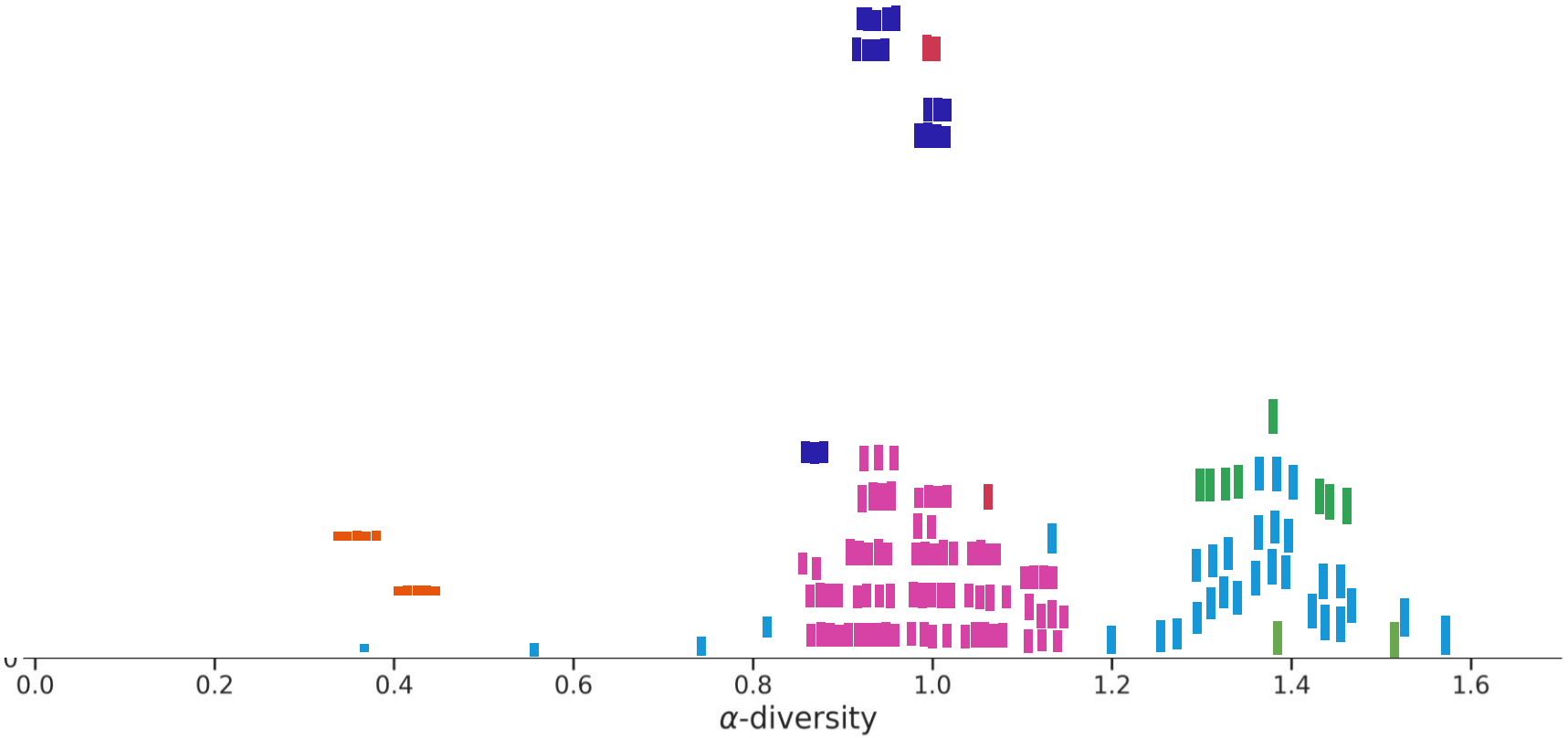


Per **lineage** α-diversity values

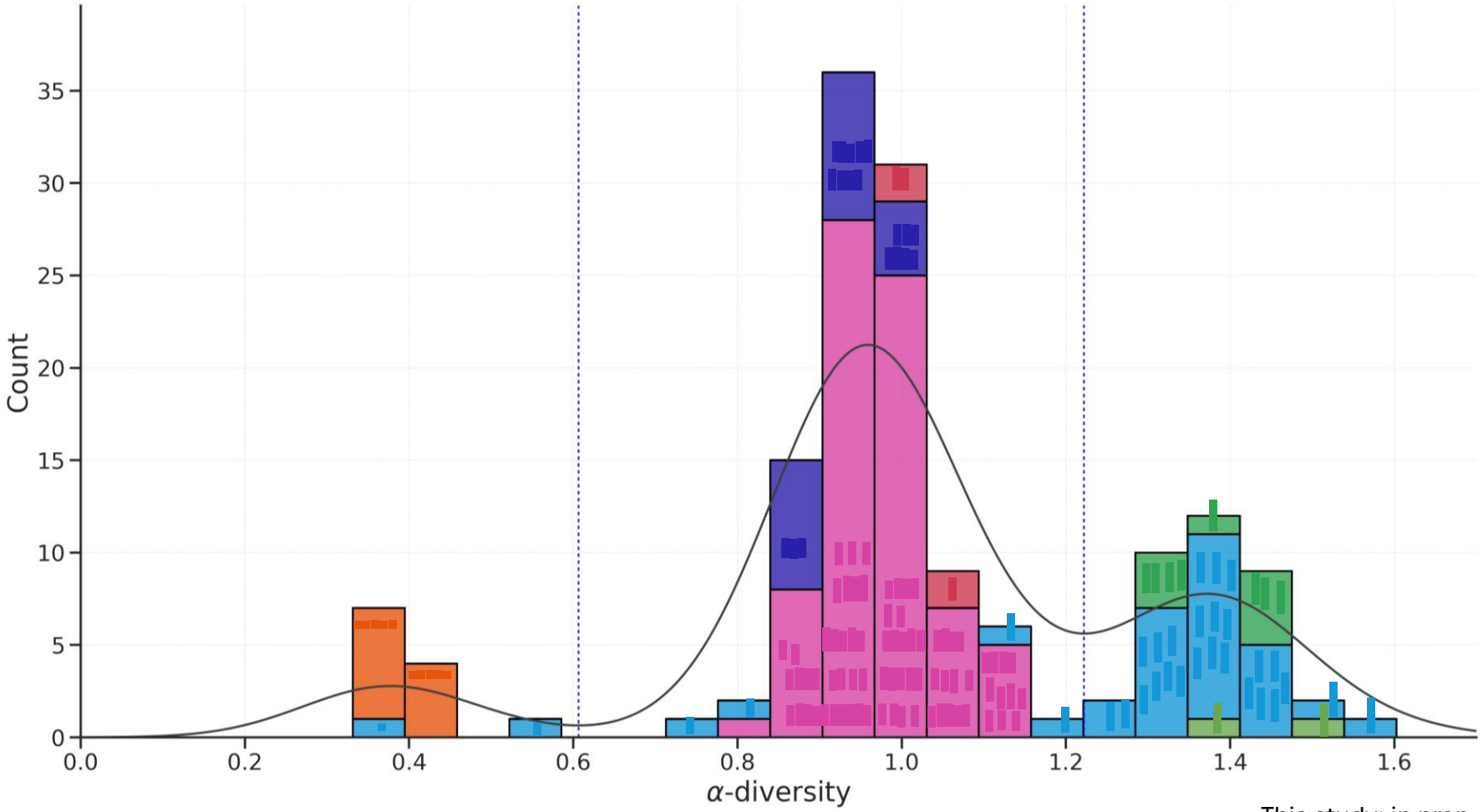


This study; in prep.

Per lineage α-diversity values

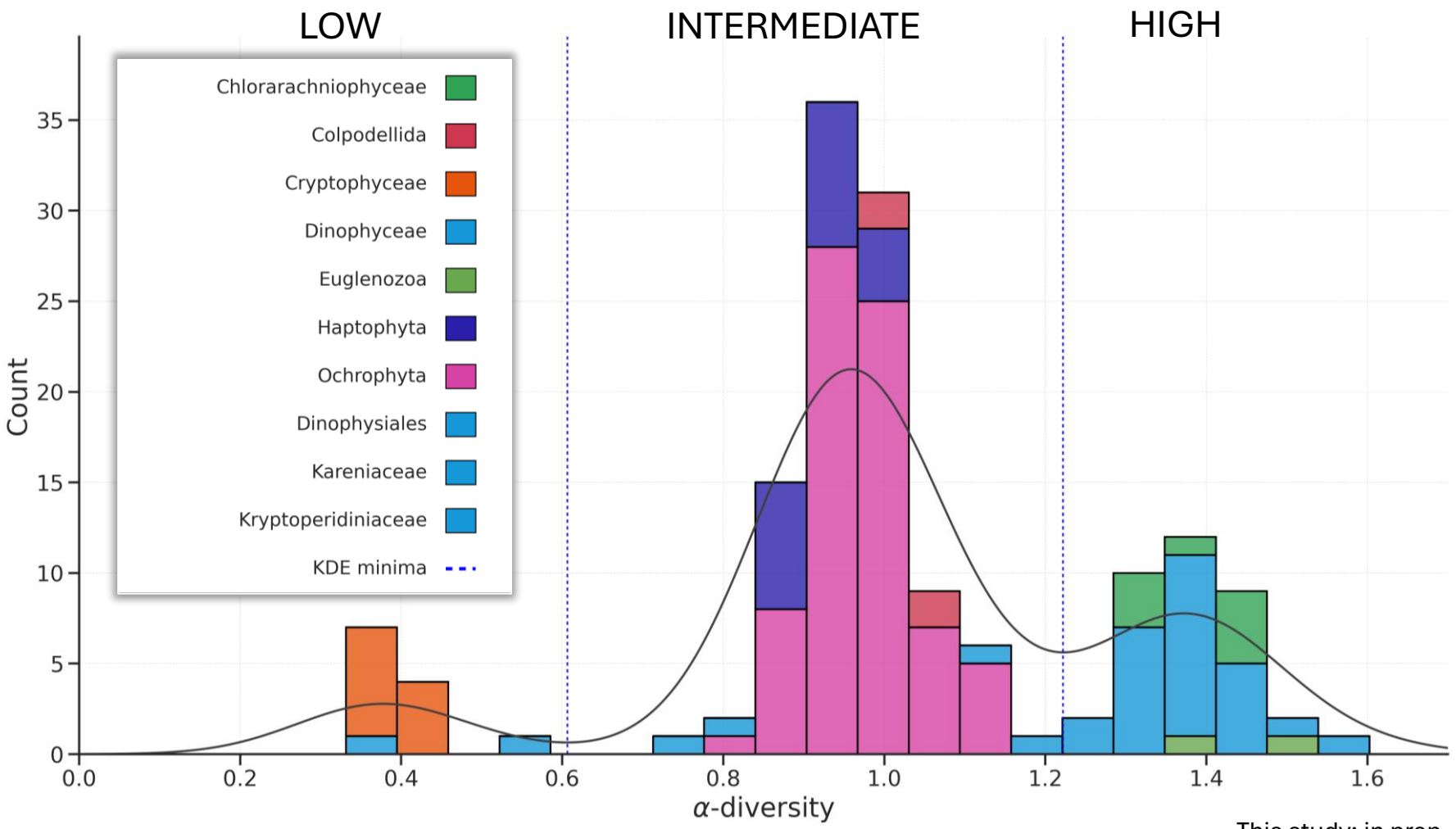


Per lineage α-diversity values



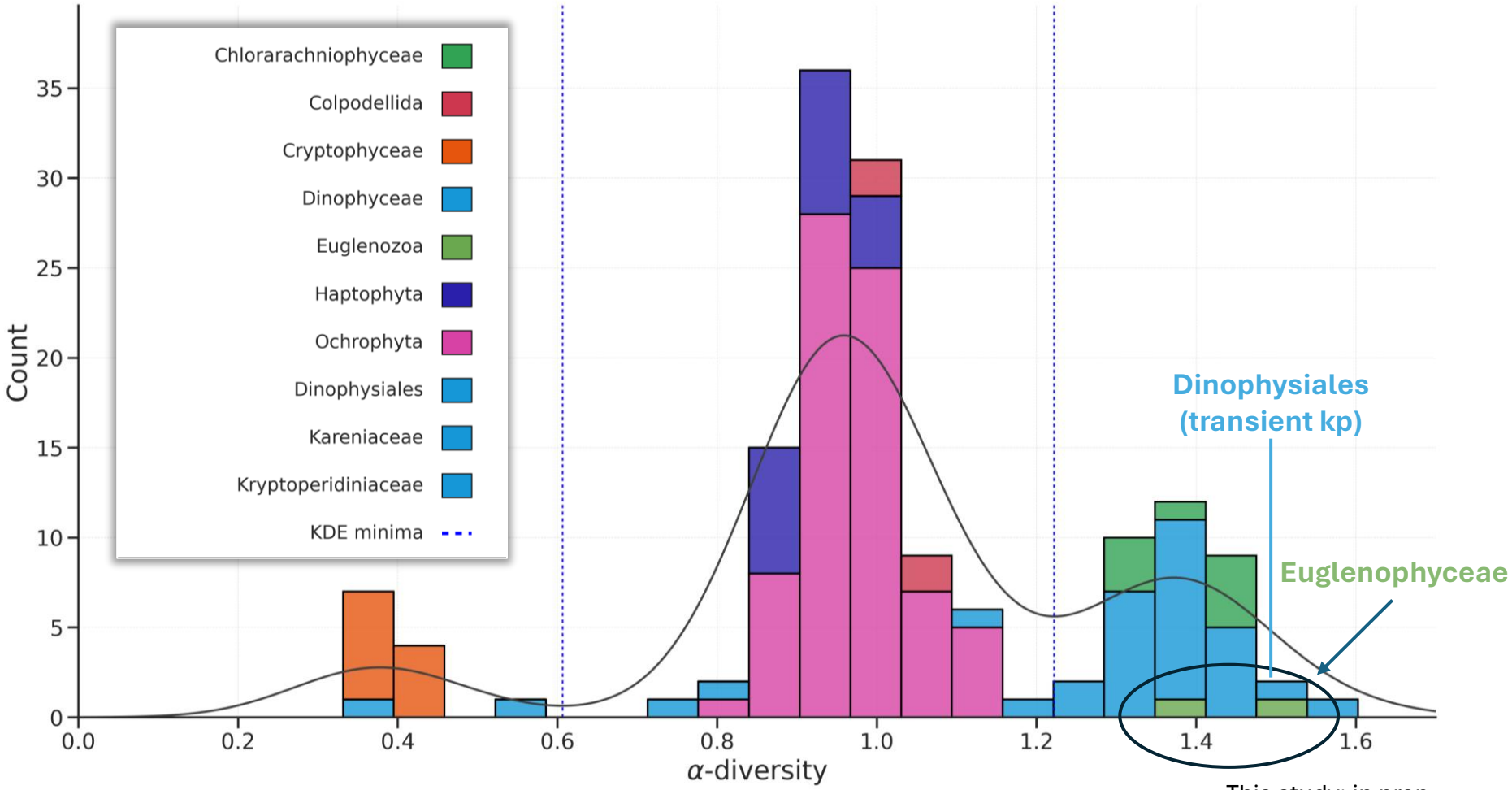
This study; in prep.

3 α-diversity regimes



This study; in prep.

3 α-diversity regimes



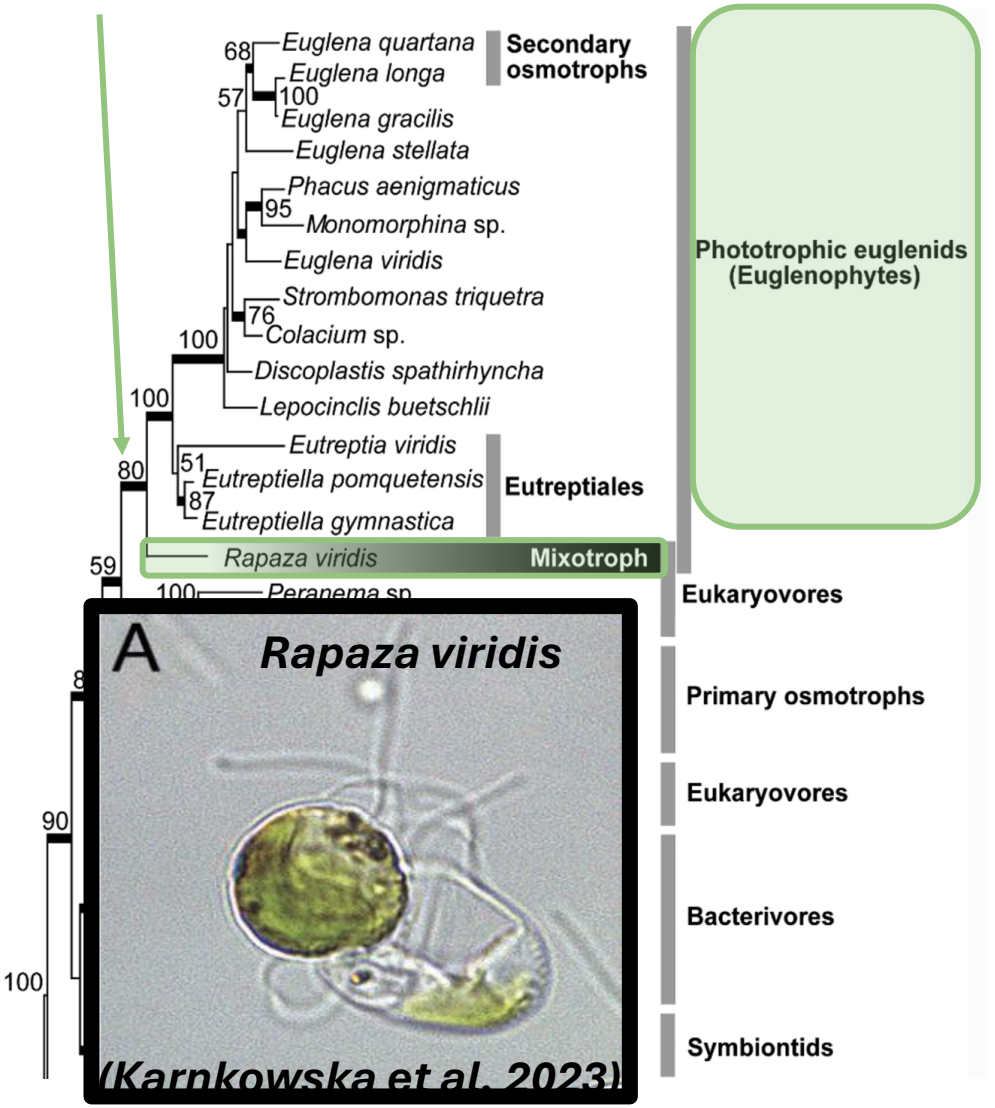
This study; in prep.

Euglenophyceae: a kleptoplasty-driven acquisition

- **Phylogenetic position of *Rapaza viridis***
  - *Yamaguchi et al. 2012*
  - *Rackevei et al. 2023*
- **(Shared) nucpt# repertoire**
  - Karnkowska et al. 2023
  - Maruyama et al. 2023
  - Kashiyama et al. 2026
- **Mosaicism**
  - Maruyama et al. 2011
  - Novák Vanclová et al. 2020
  - This study; in prep



Kleptoplastic ancestor

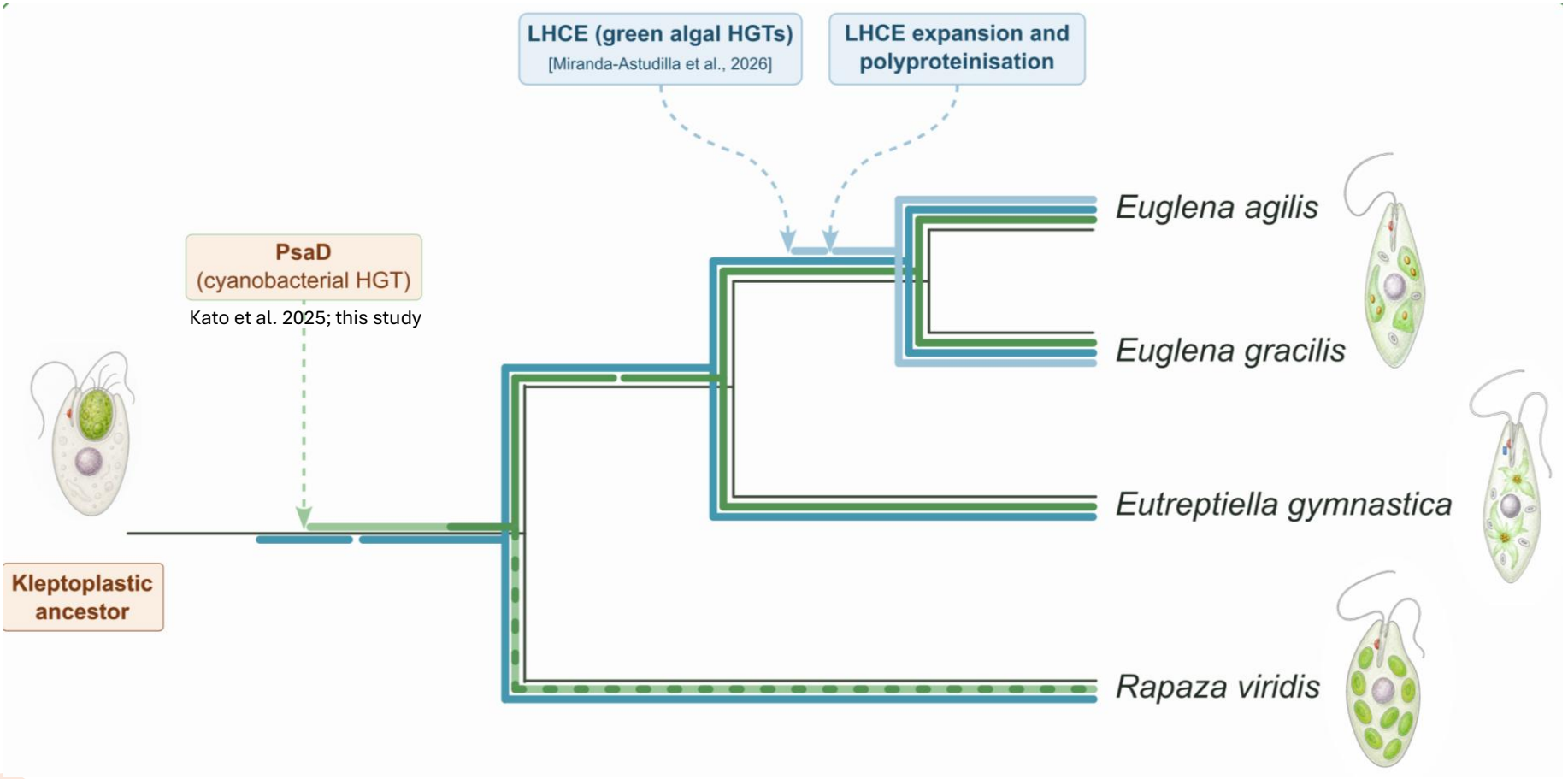


Yamaguchi et al., 2012

Mosaicism at work – PSI complex - PsaD

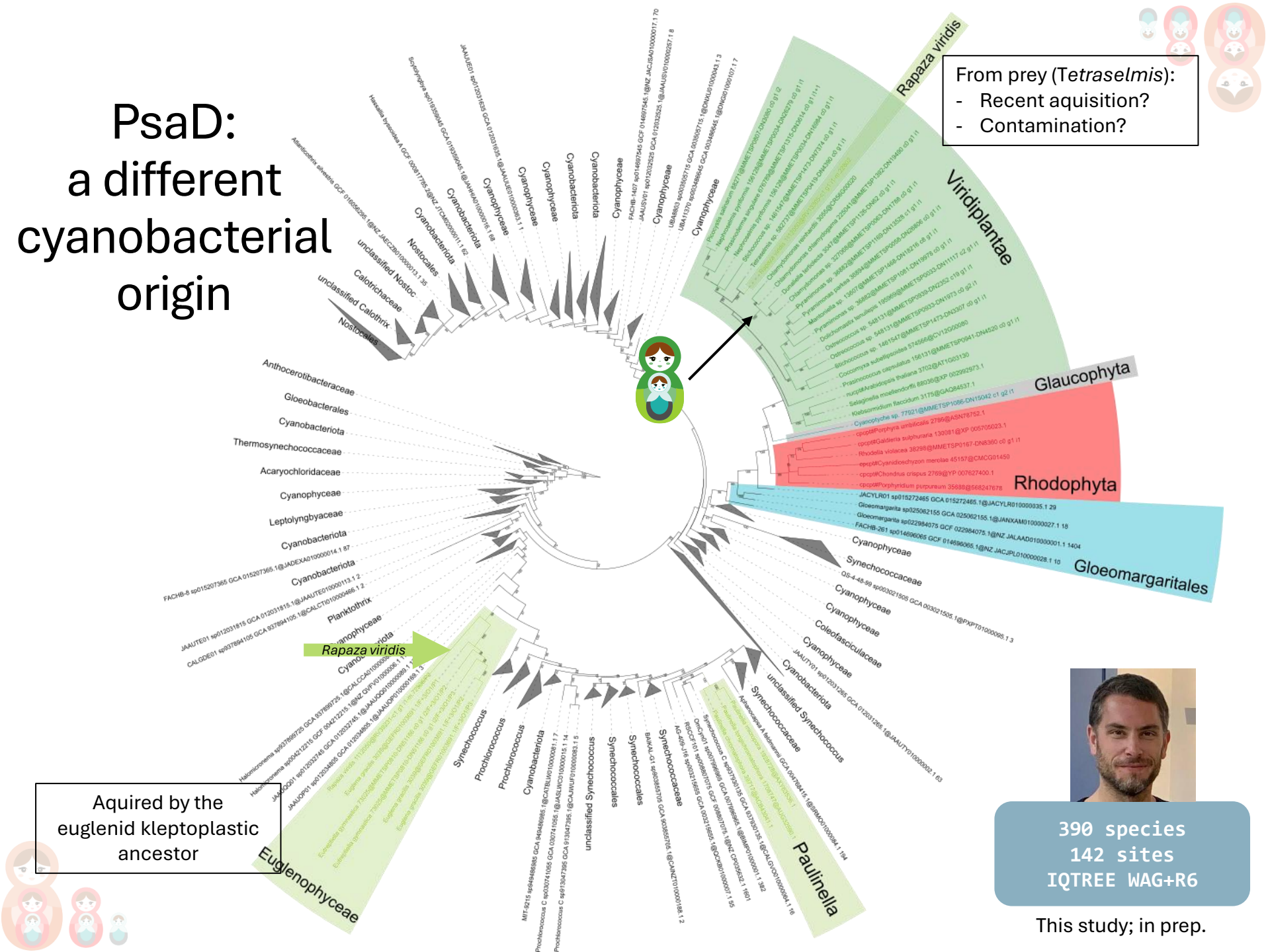


Pierre Cardol, ULiege



This study; in prep.

PsaD: a different cyanobacterial origin

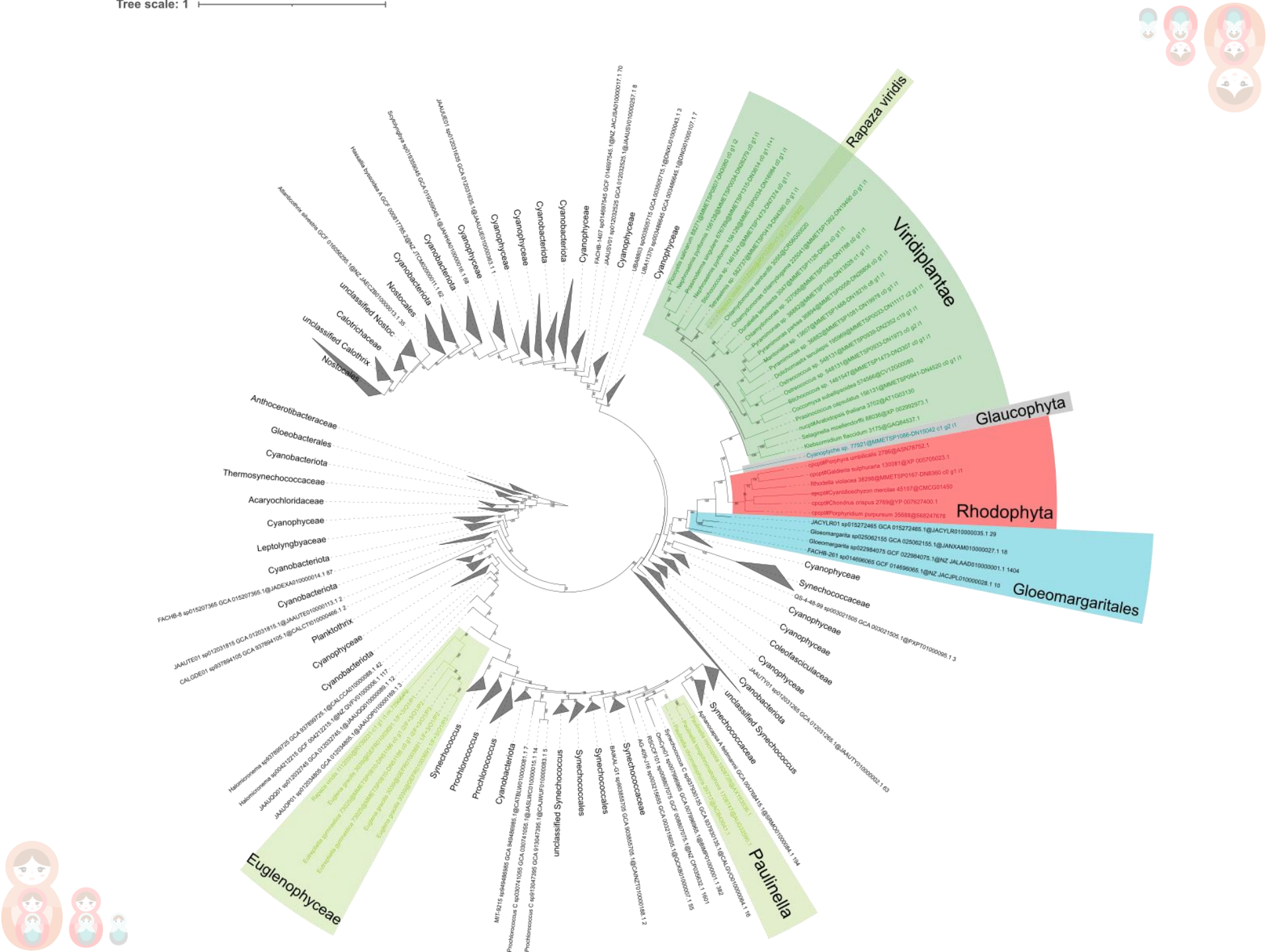


From prey (*Tetraselmis*):
- Recent aquisition?
- Contamination?

Aquired by the euglenid kleptoplastic ancestor

390 species
142 sites
IQTREE WAG+R6

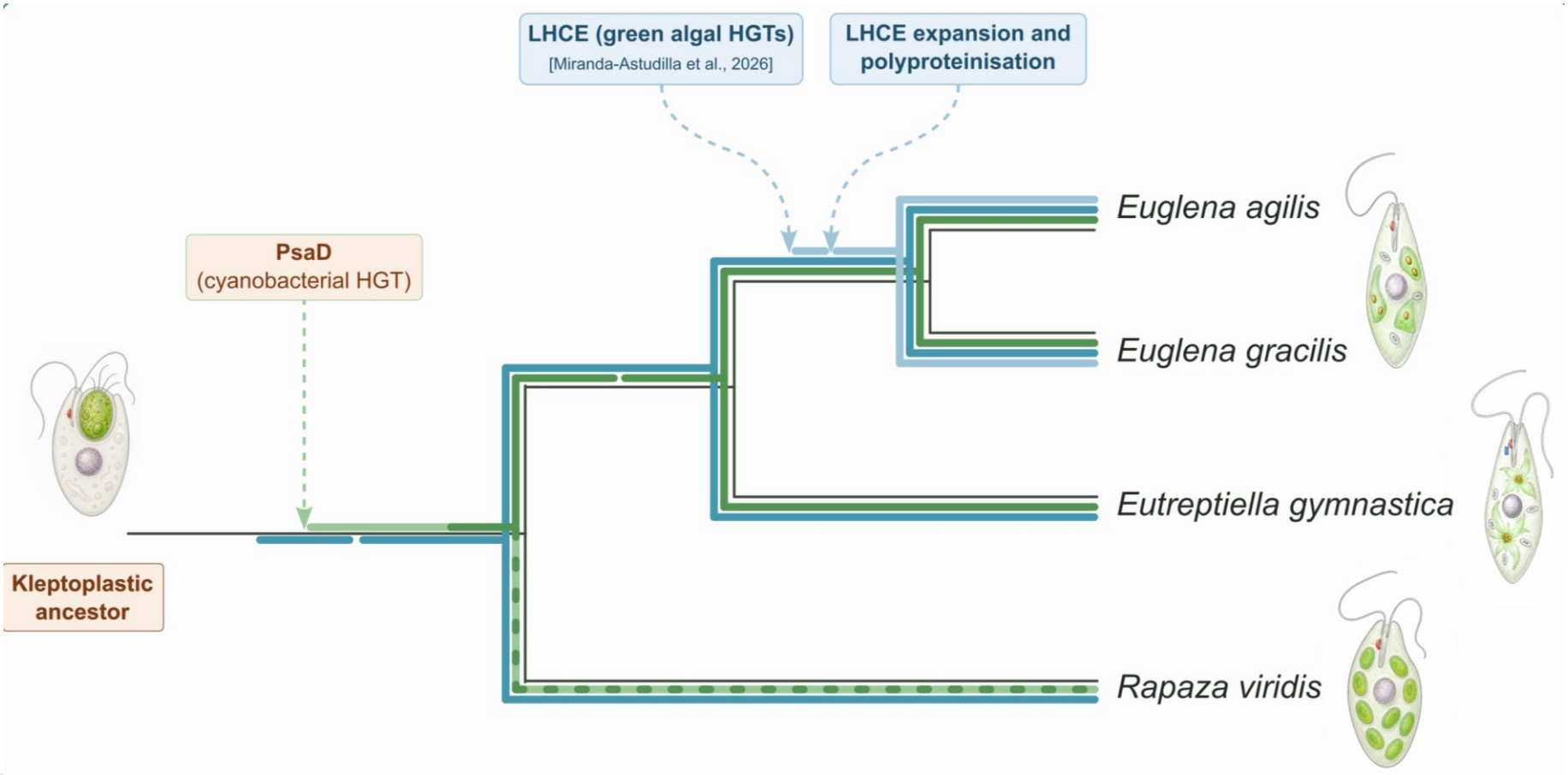
This study; in prep.


Tree scale: 1
Rapaza viridis
Viridiplantae
Glaucophyta
Rhodophyta
Gloeomargaritales
Paulinella
Euglenophyceae
Cyanophyceae
Cyanobacteriota
Synechococcus
Prochlorococcus
Synechococcales
Synechococcaceae
unclassified Synechococcus
Coleofasciculaceae
Synechococcaceae
Nostocales
unclassified Nostoc
Calotrichaceae
unclassified Calothrix
Anthocerotibacteraceae
Gloeobacterales
Thermosynechococcaceae
Acaryochloridaceae
Leptolyngbyaceae
Planktothrix

Mosaicism at work – PSI complex - LHCE



Pierre Cardol, ULiege

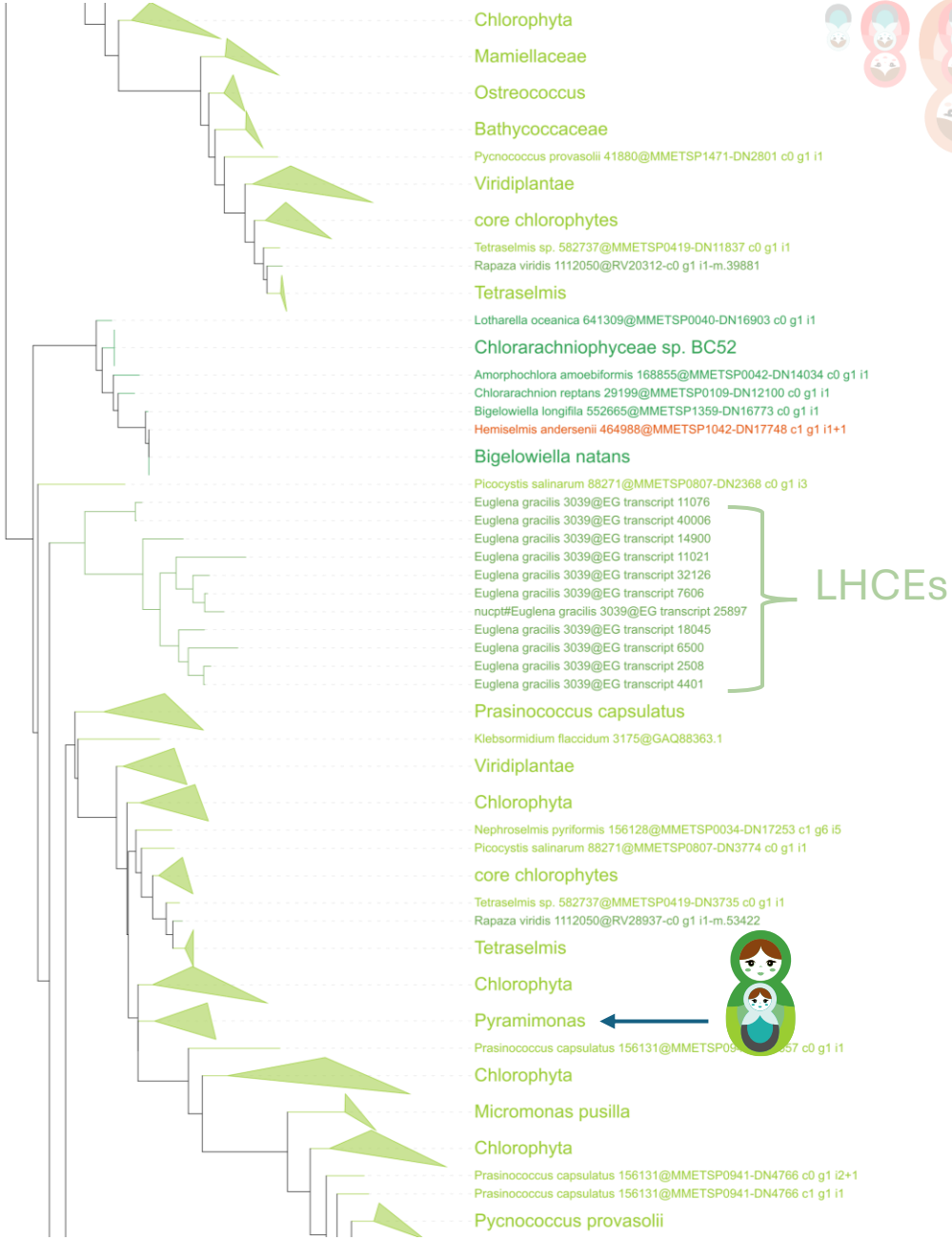


This study; in prep.

OG0000092 - LHCs

LHCE: a different green algal origin

LHC superfamily gene tree
OG0000092
IQTREE LG+C20+G

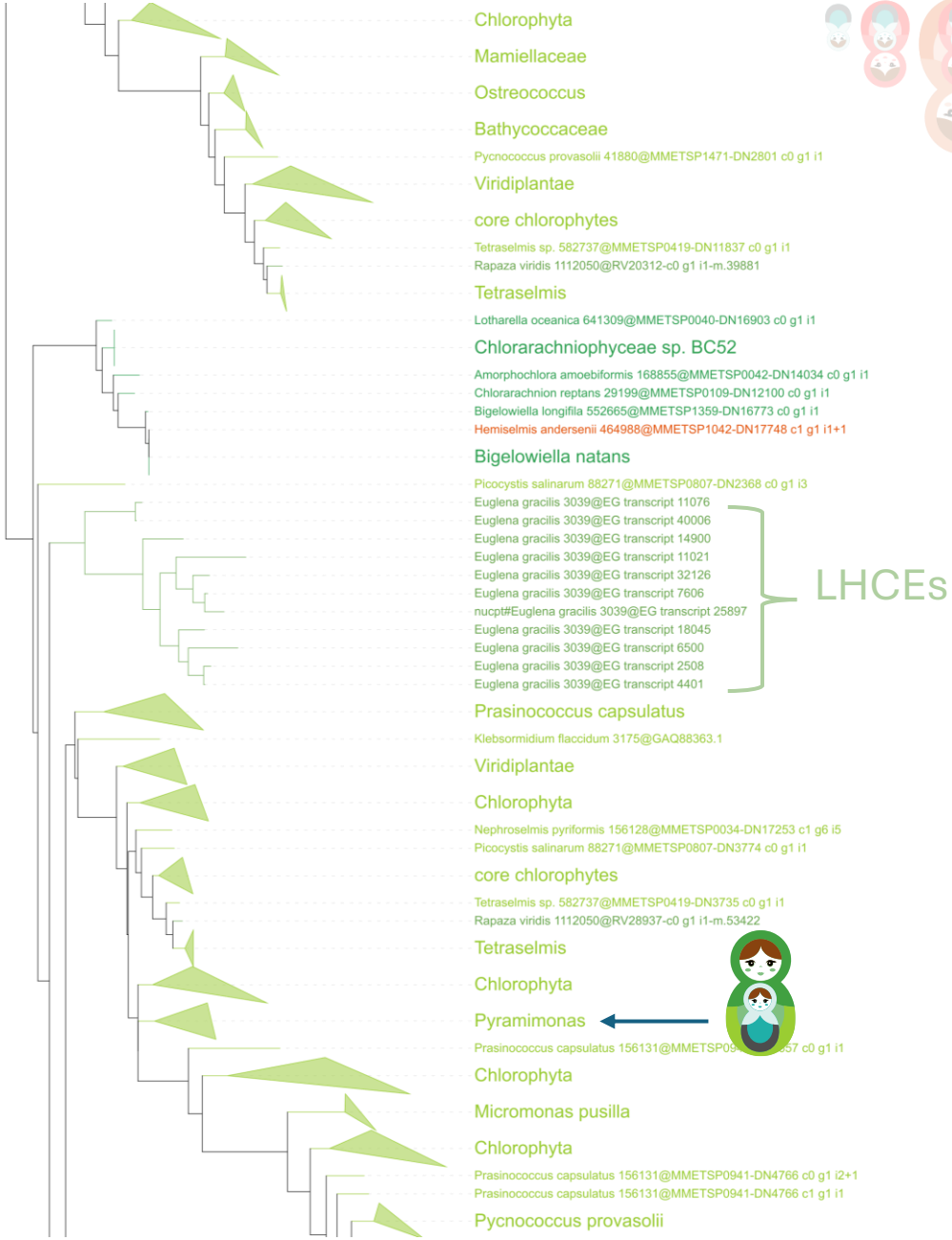


This study; in prep.

OG0000092 - LHCs

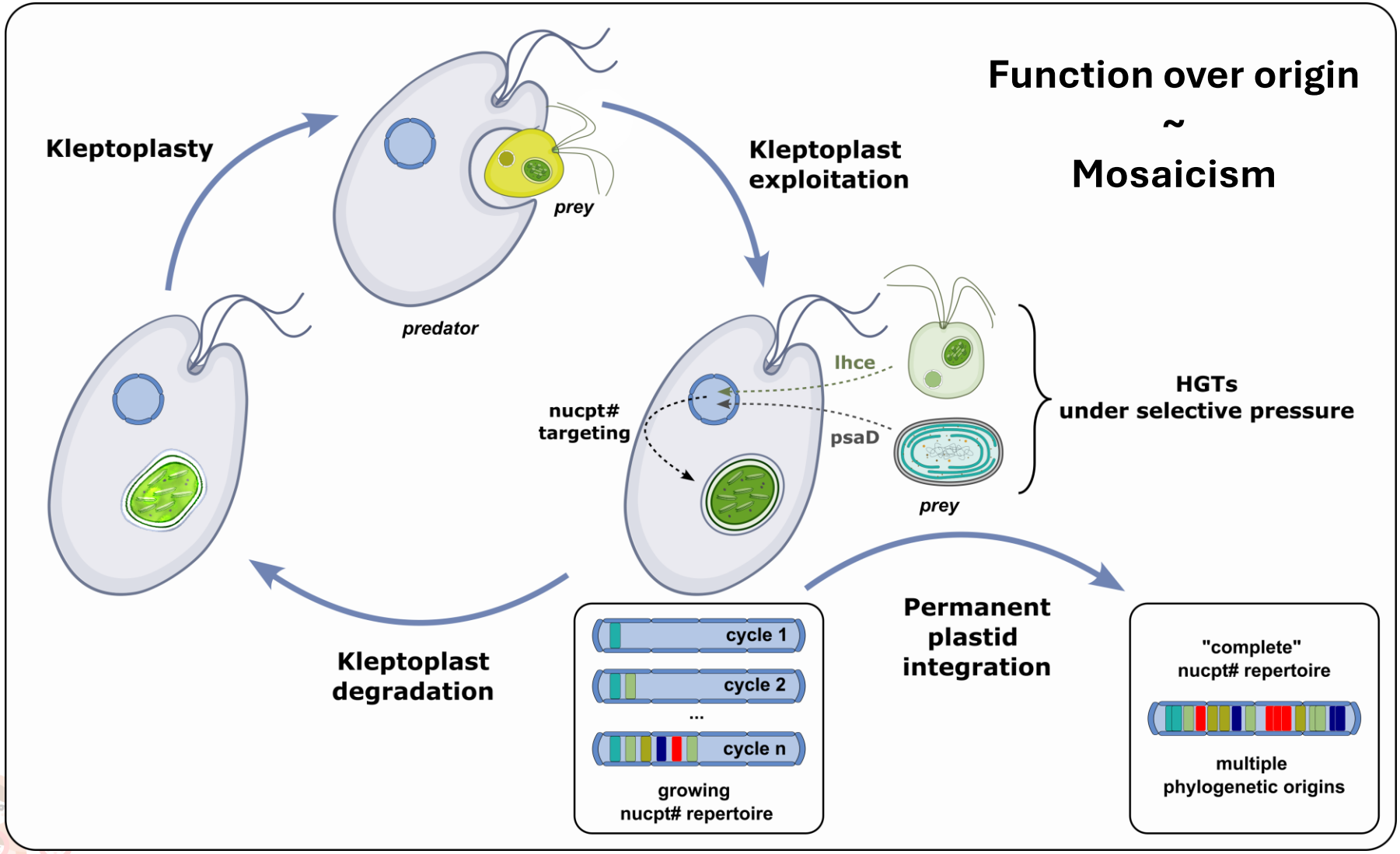
LHCE: a different green algal origin

red-type diadinoxanthin-diatoxanthin photoprotective cycle (Kato et al., 2025)



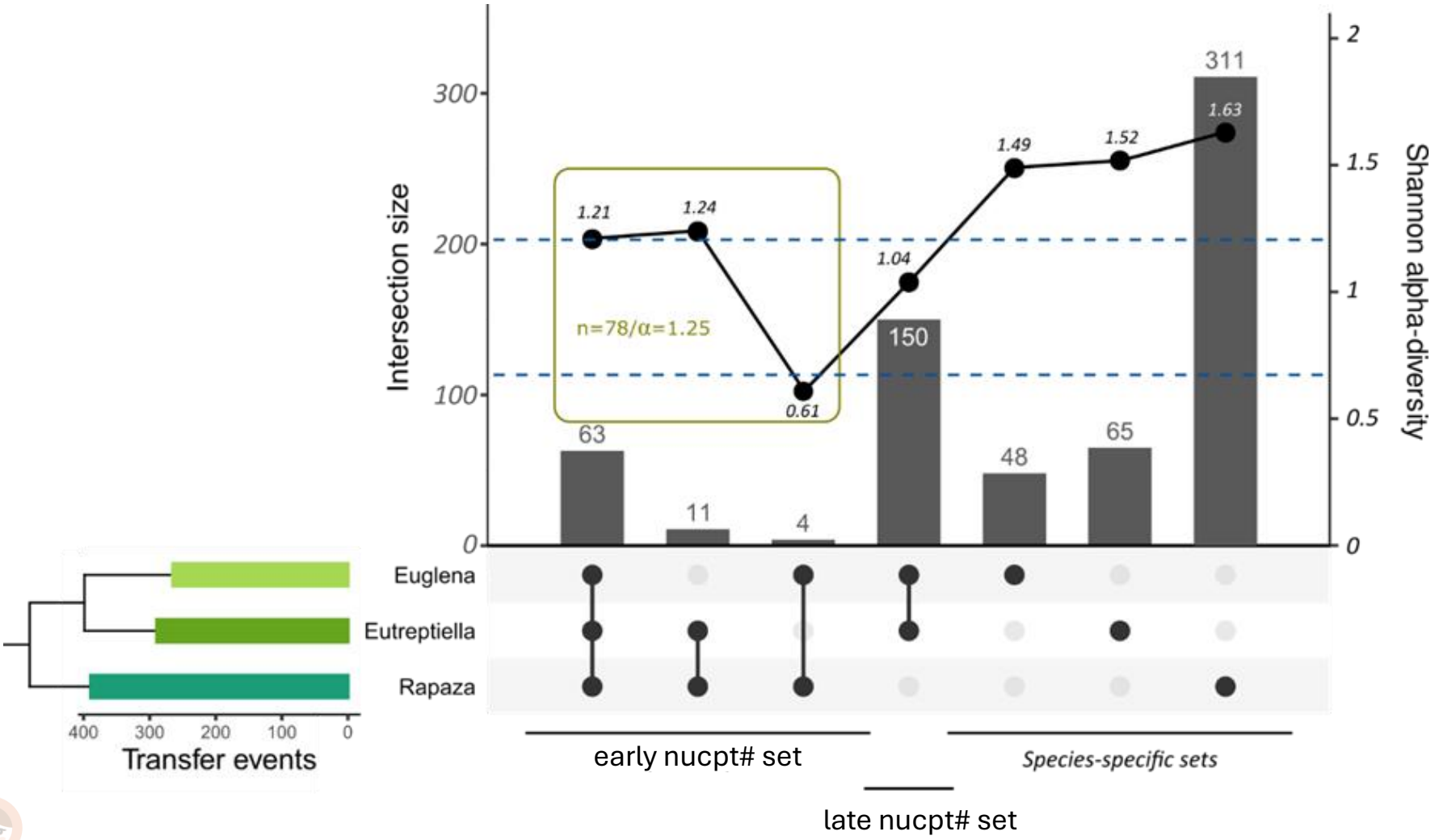
This study; in prep.

Kleptoplastic ratchet shopping bag model

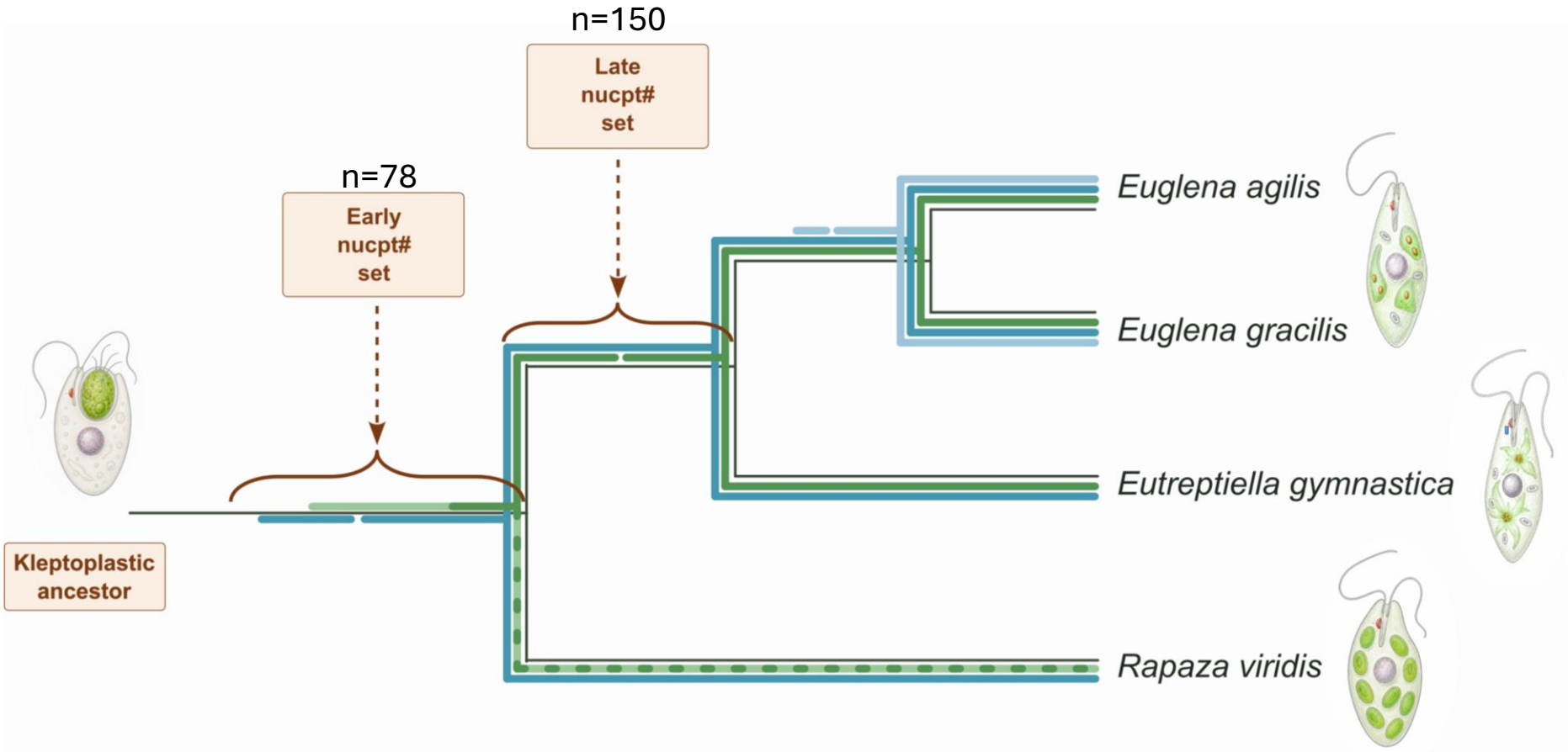


This study; in prep.

Euglenophyceae ~ gene sets

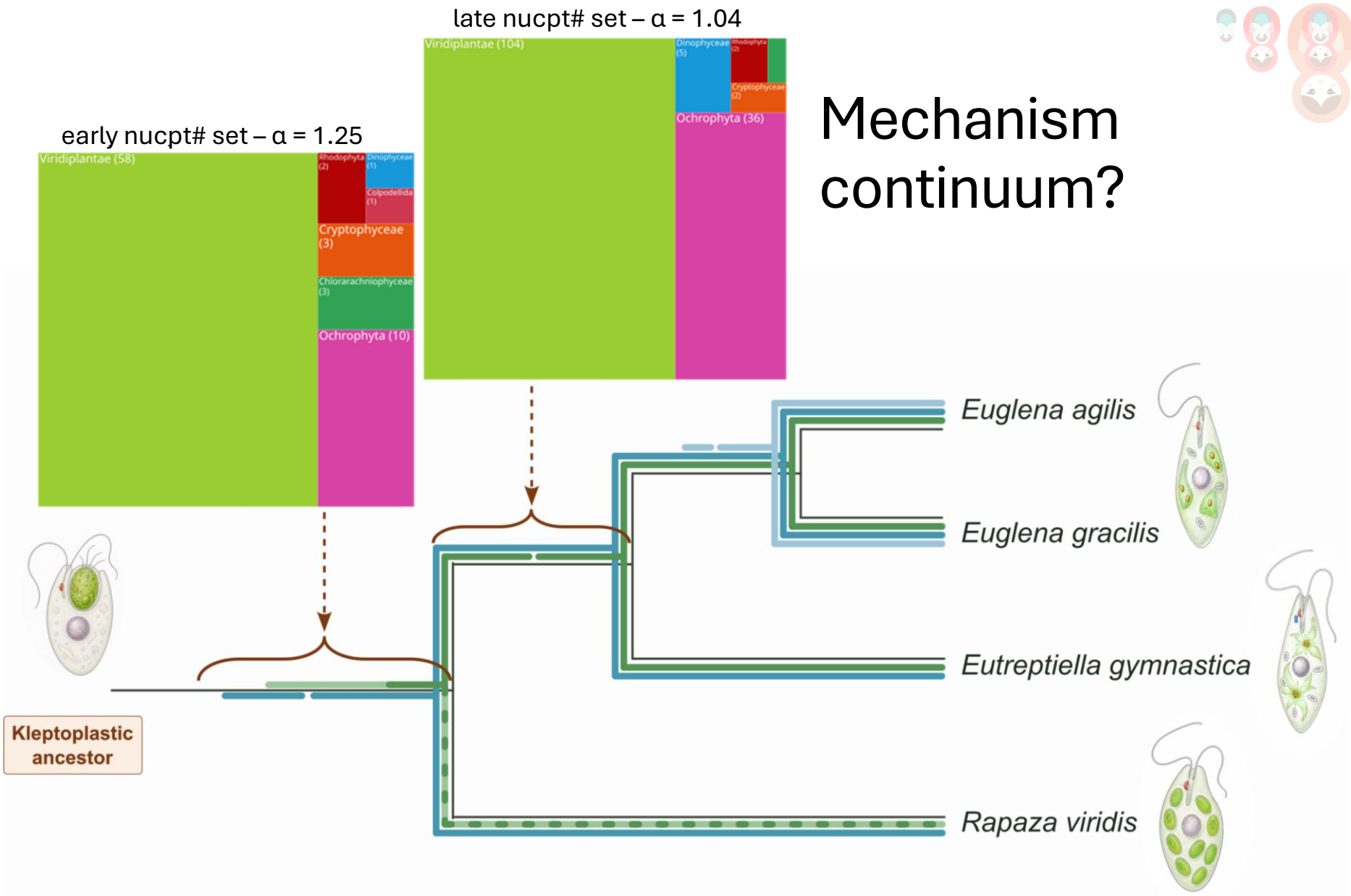


Integration continuum?



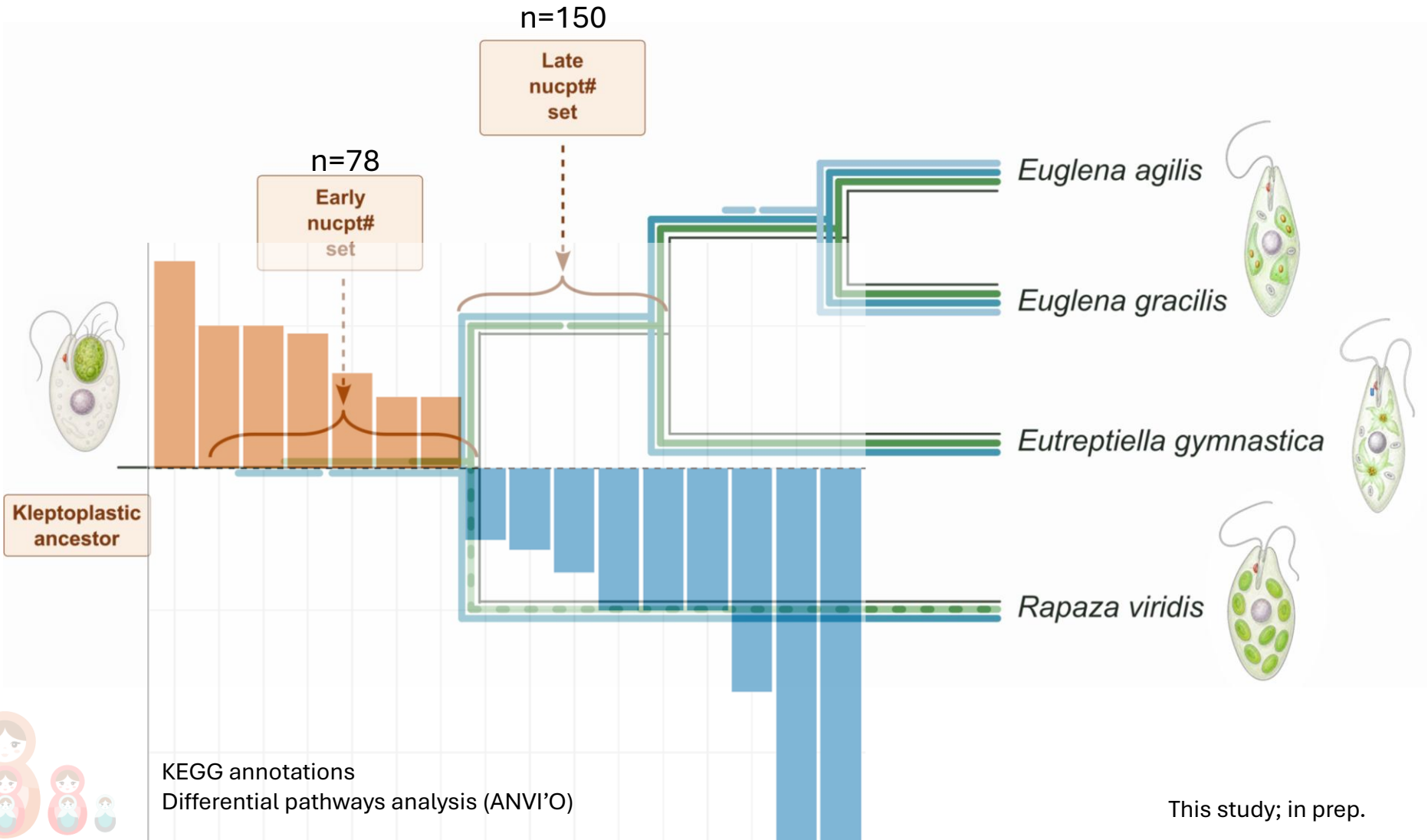
This study; in prep.

Mechanism continuum?



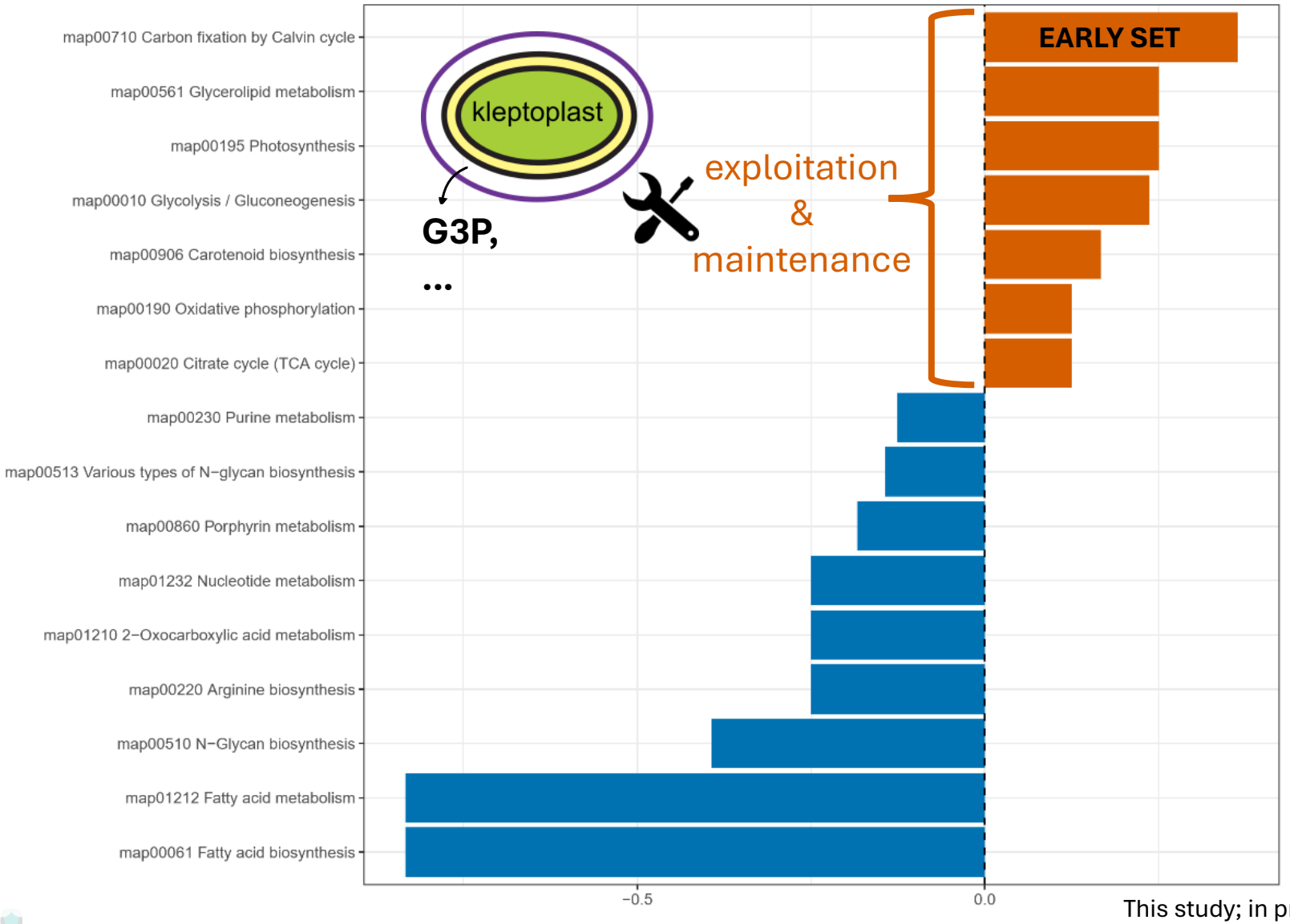
This study; in prep.

Functionnal partitioning?

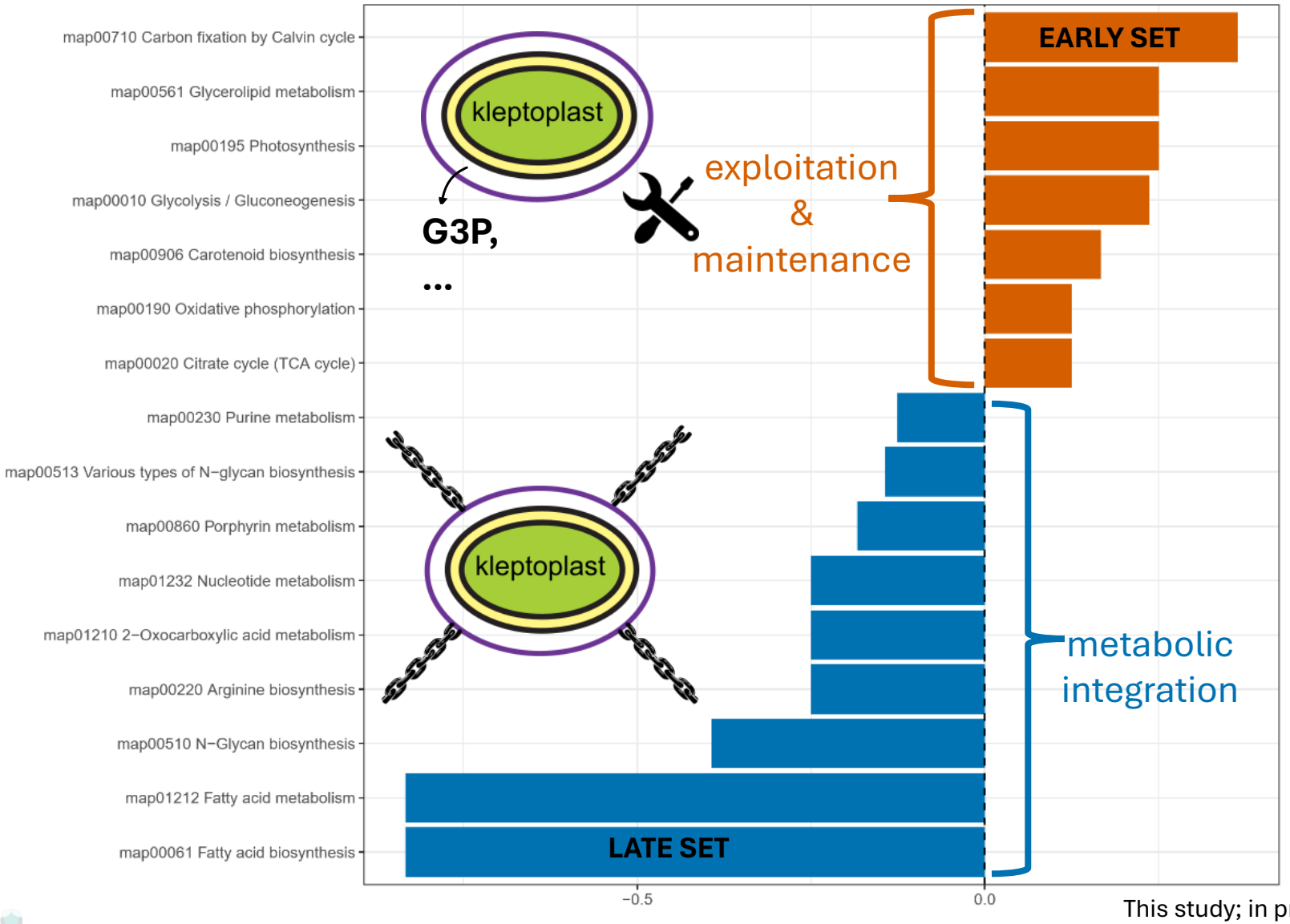


KEGG annotations
Differential pathways analysis (ANVI'O)

This study; in prep.

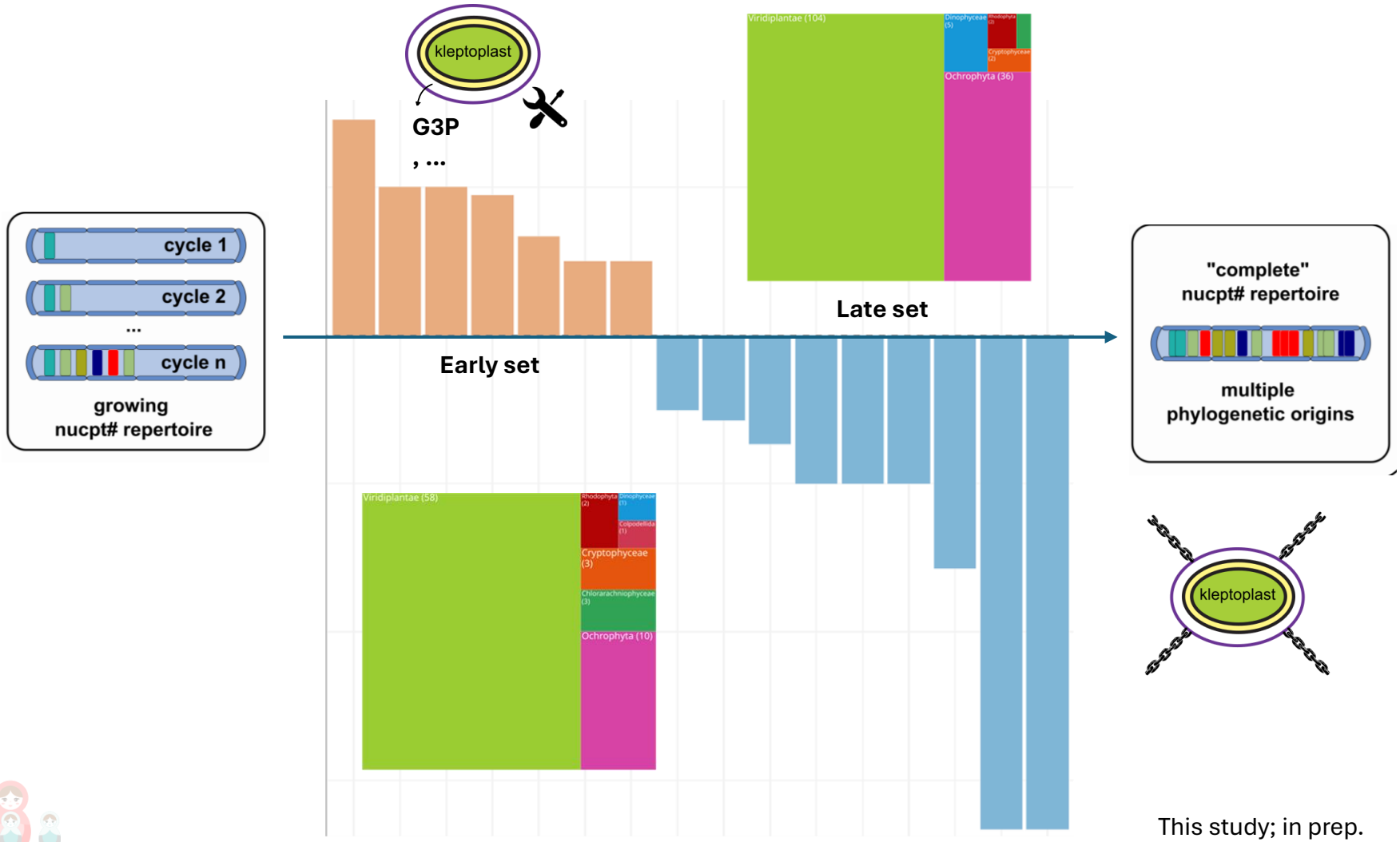


This study; in prep.



This study; in prep.

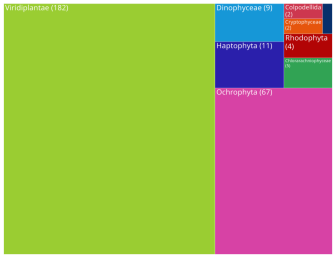
Integration continuum

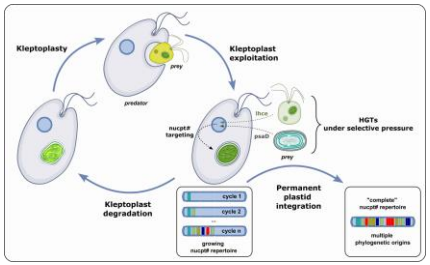


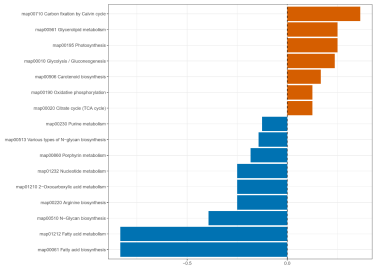
This study; in prep.

Take home message

- Kleptoplasty explains well mosaicism
  - Strong selective pressure exerted by the kleptoplast
  - Favors function over origin
- Kleptoplast integration continuum
  - From photosynthesis to broader metabolic functions
- Widespread mosaicism
  - More kleptoplasty-mediated aquisition?







Acknowledgements

- Henner Brinkmann, DSMZ
- Pierre Cardol, ULiege





Eukaryotic Phylogenomics Unit





DB MVV

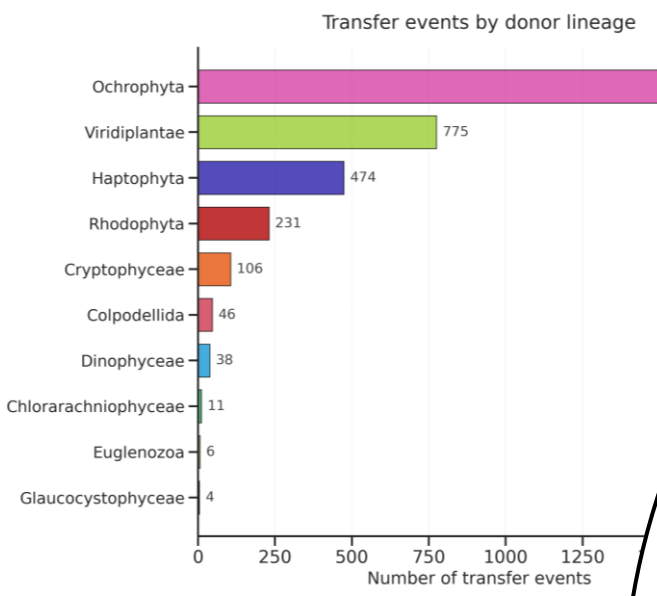
Me

Thank you!

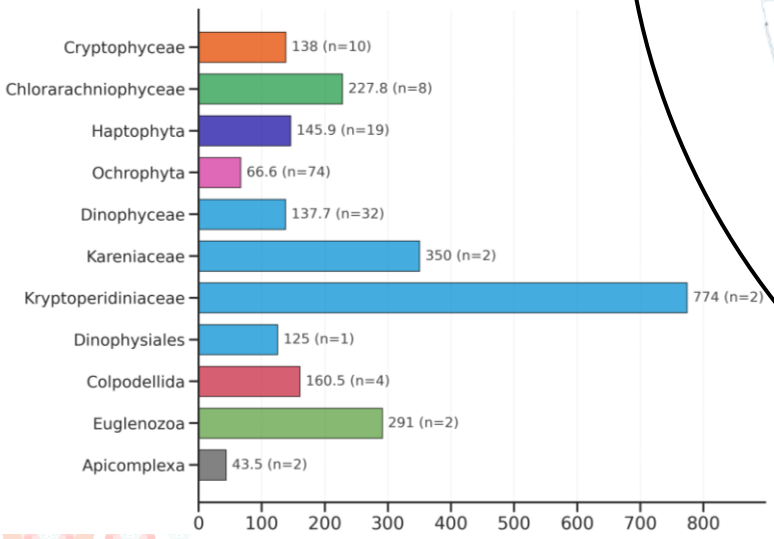
LIÈGE université

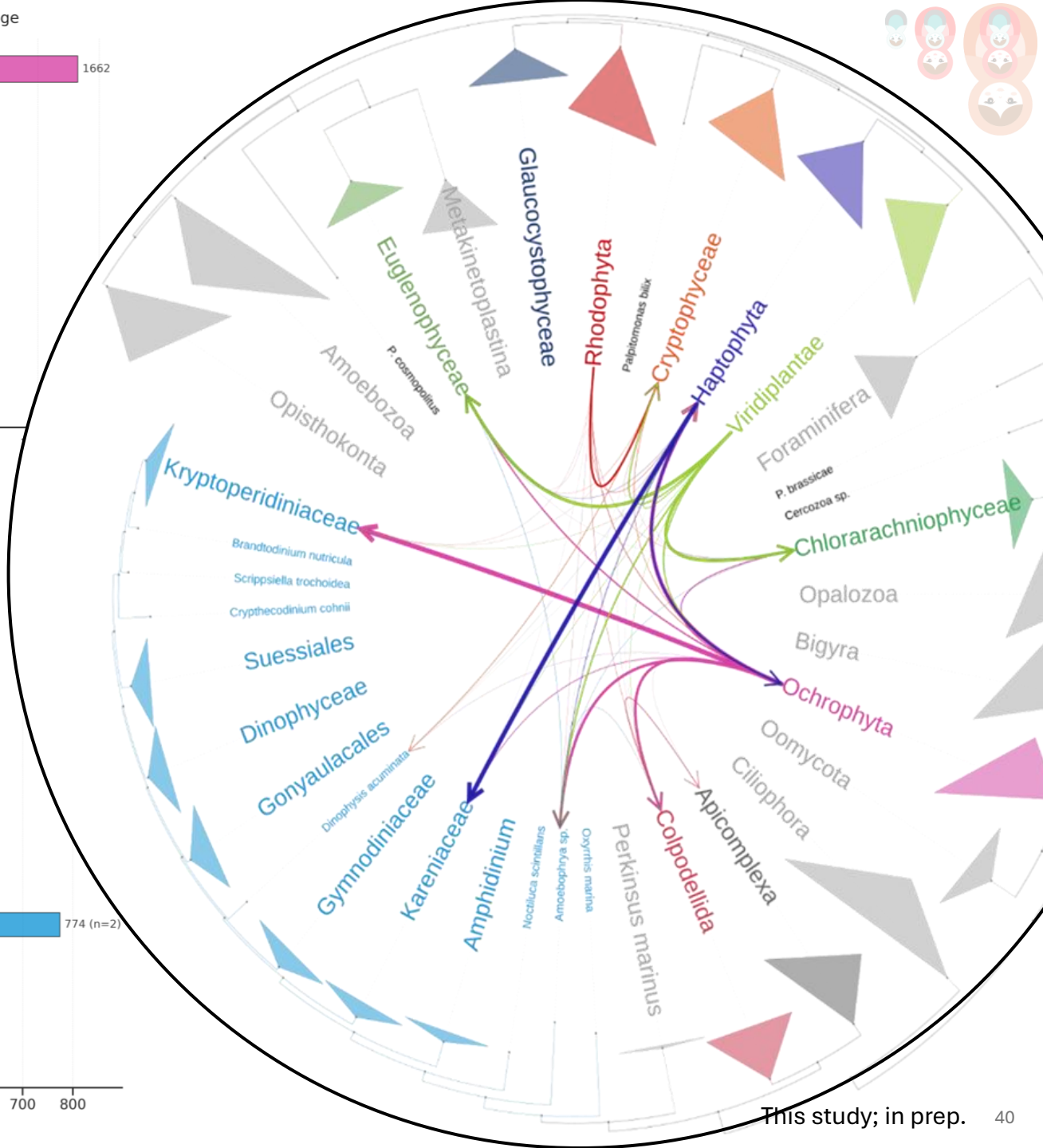
fnrs LA LIBERTÉ DE CHERCHER

Images ©2019 Google, Data SIO, NOAA, U.S. Navy, NGA, GEBCO, Landsat / Copernicus, Données cartographiques ©2019



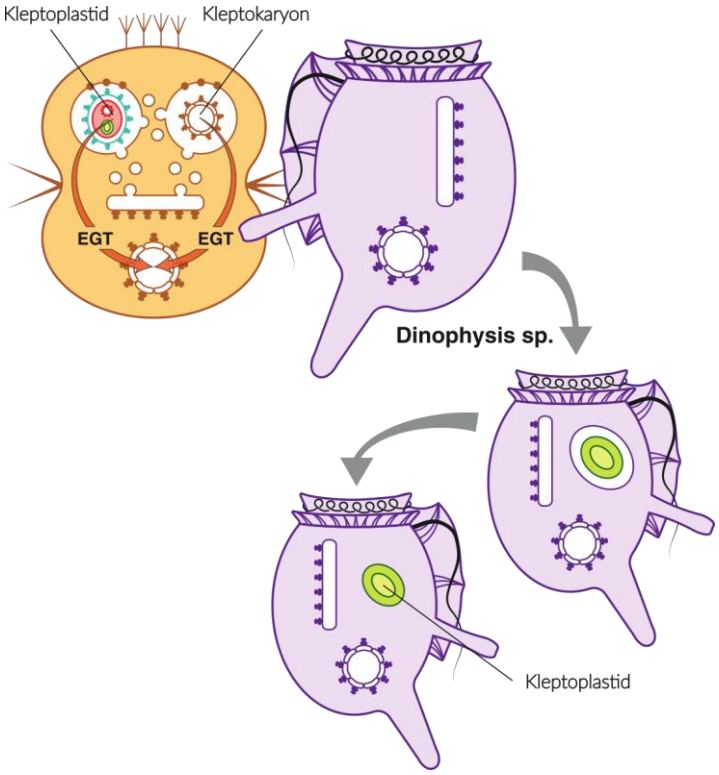
3353 nucpt# gene transfer network





This study; in prep.

Kleptoplasty by myzocytosis



Dinophysis acuminata_47934 — α = 1.576



SCIENTIFIC REPORTS

OPEN

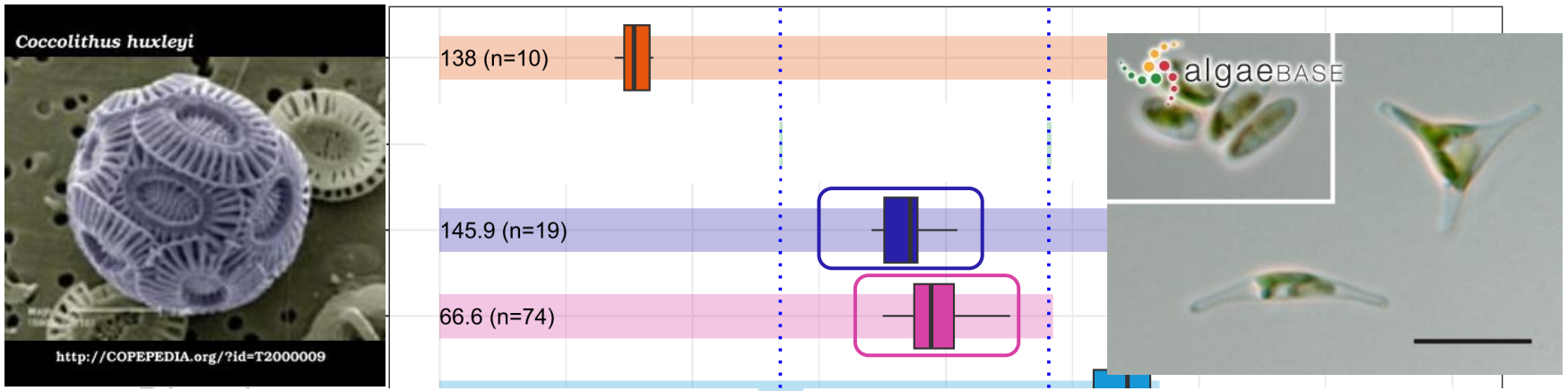
Genes functioned in kleptoplastids of *Dinophysis* are derived from haptophytes rather than from cryptophytes

Yuki Hongo[1], Akinori Yabuki[2], Katsunori Fujikura[2] & Satoshi Nagai[1]

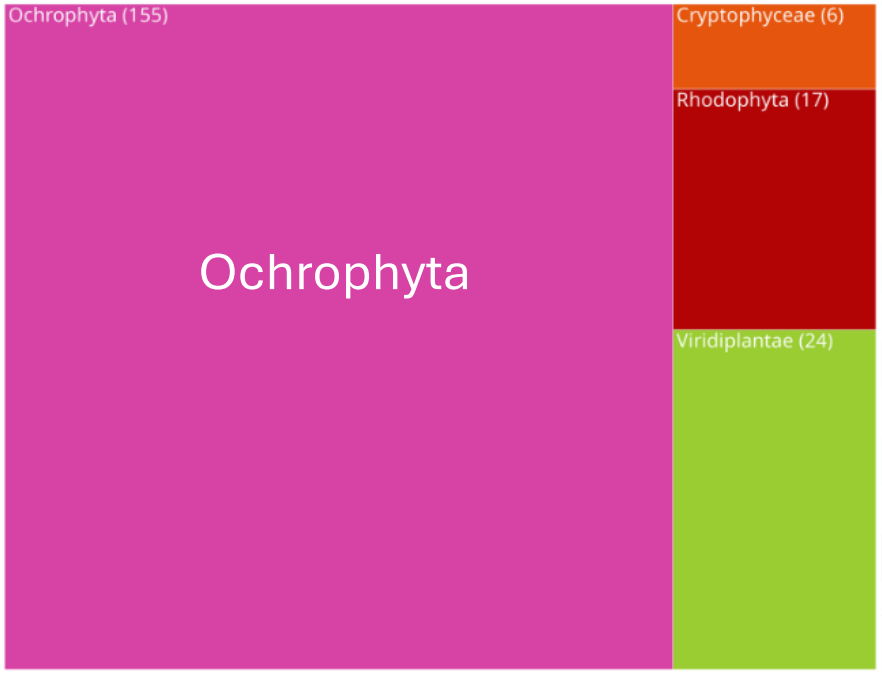
Received: 30 October 2018
Accepted: 5 June 2019
Published online: 21 June 2019

This study; in prep.

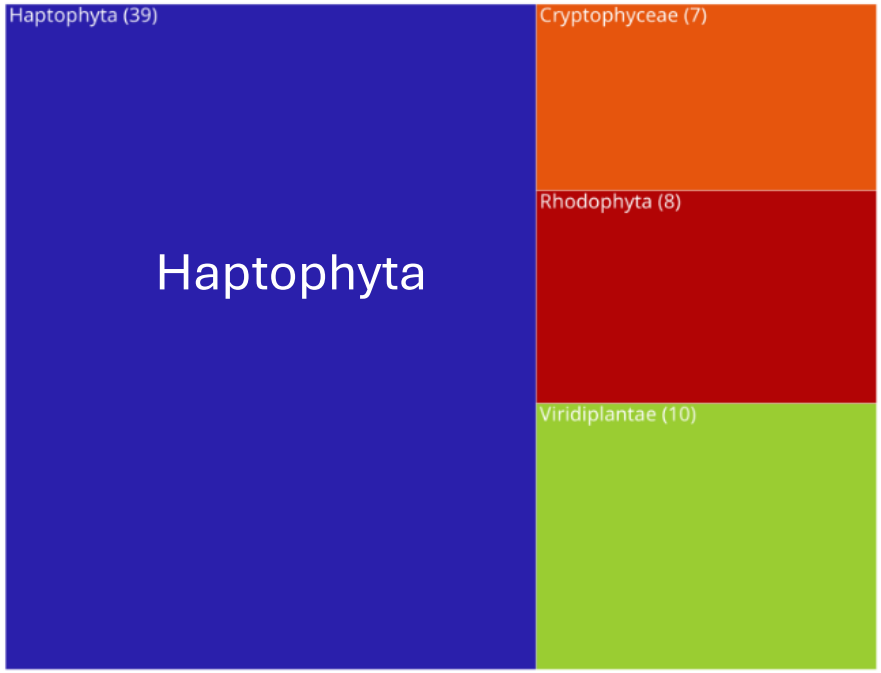
Unknown phylogenetic origin of Haptophyta and Ochrophyta



A **Emiliania huxleyi_280463 — α = 0.913**

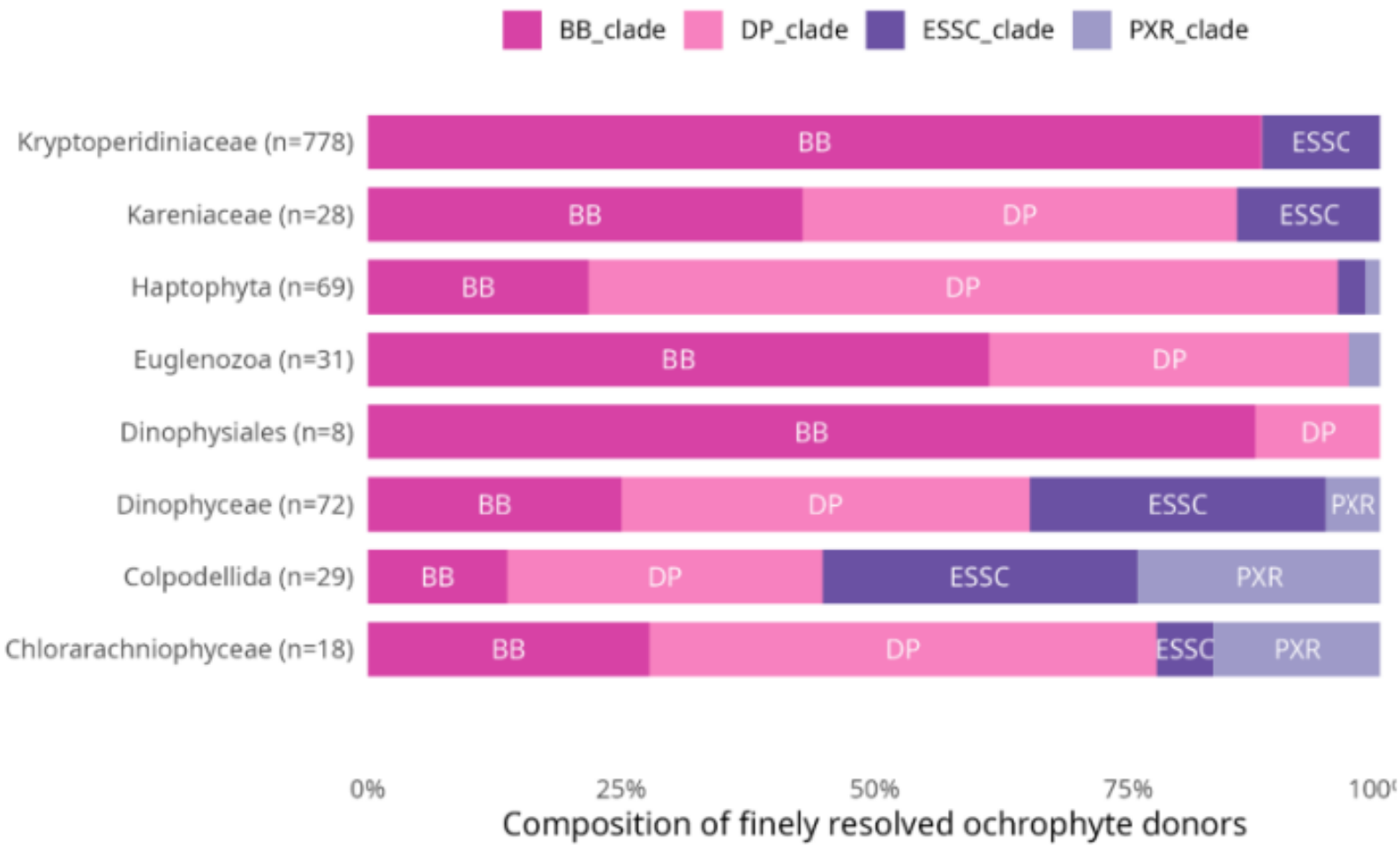


B **Phaeodactylum tricornutum_2850 — α = 0.966**

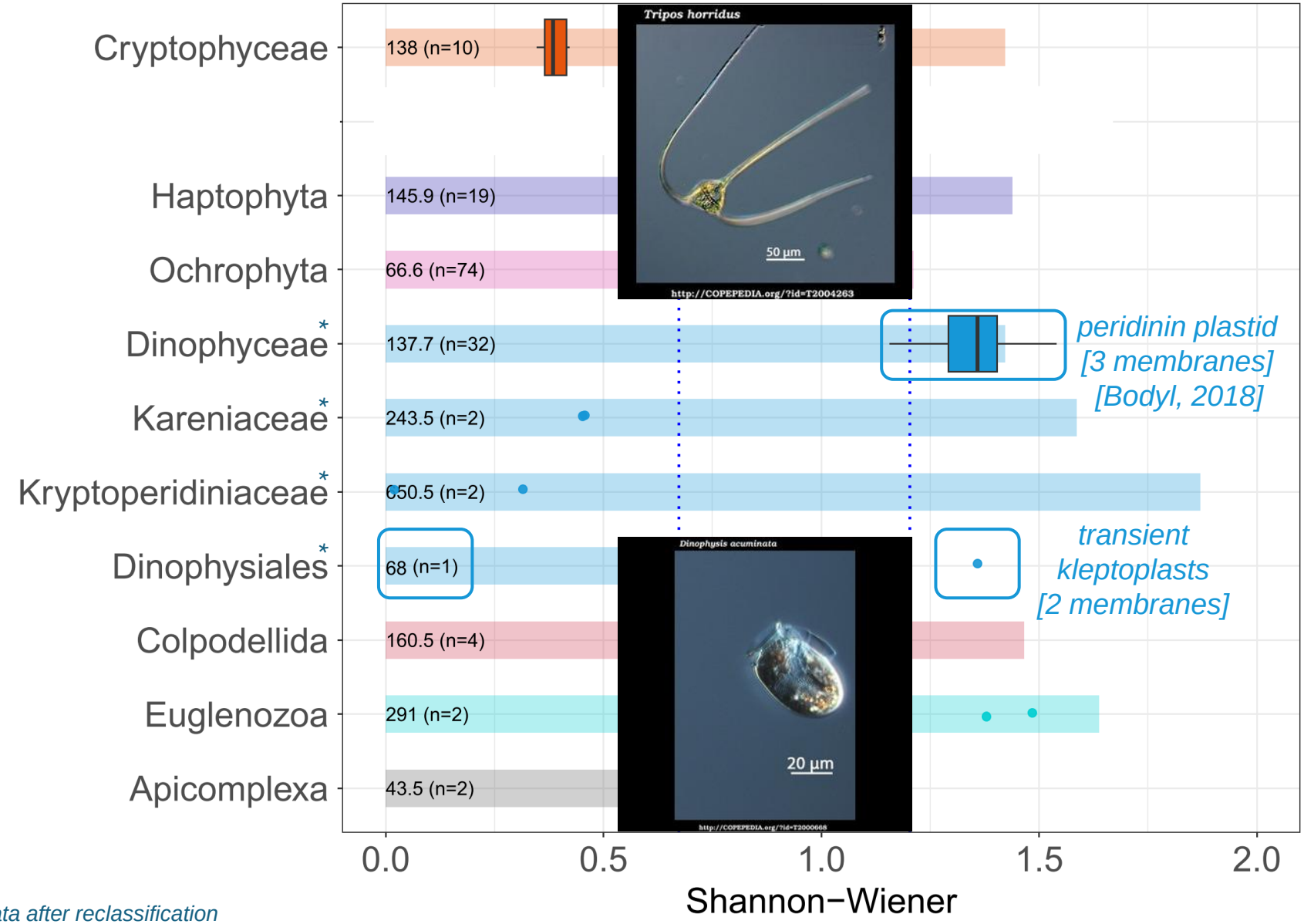


Shannon−Wiener

*Data after reclassification


BB_clade
DP_clade
ESSC_clade
PXR_clade
Kryptoperidiniaceae (n=778)
BB
ESSC
Kareniaceae (n=28)
BB
DP
ESSC
Haptophyta (n=69)
BB
DP
Euglenozoa (n=31)
BB
DP
Dinophysiales (n=8)
BB
DP
Dinophyceae (n=72)
BB
DP
ESSC
PXR
Colpodellida (n=29)
BB
DP
ESSC
PXR
Chlorarachniophyceae (n=18)
BB
DP
ESSC
PXR
0%
25%
50%
75%
Composition of finely resolved ochrophyte donors

Kleptoplasty-mediated acquisition in Dinophyceae?



*Data after reclassification

Plastid replacement in Dinophyceae

Waller & Kořený (2017)

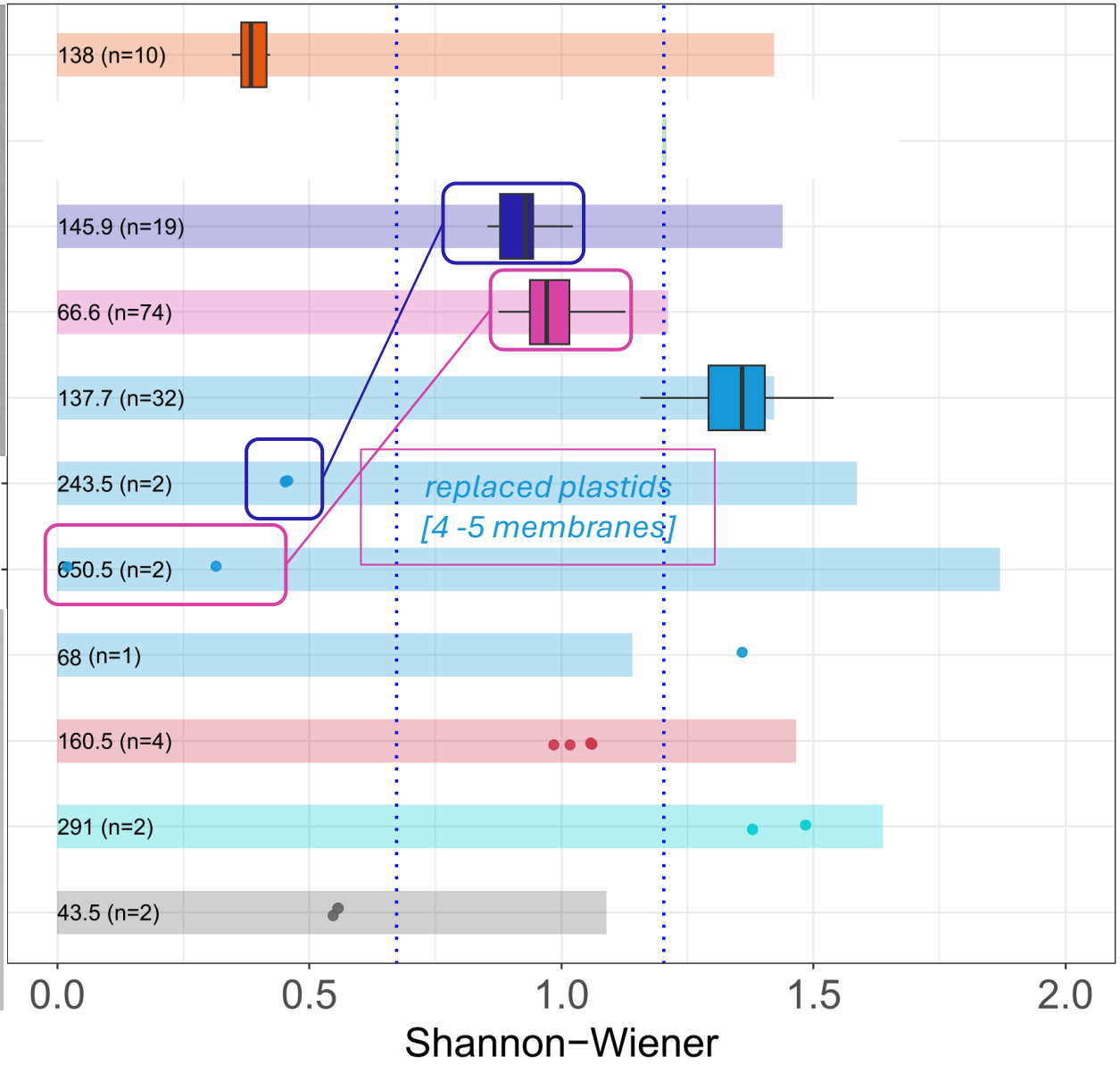


Kareniaceae*

Kryptoperidiniaceae*

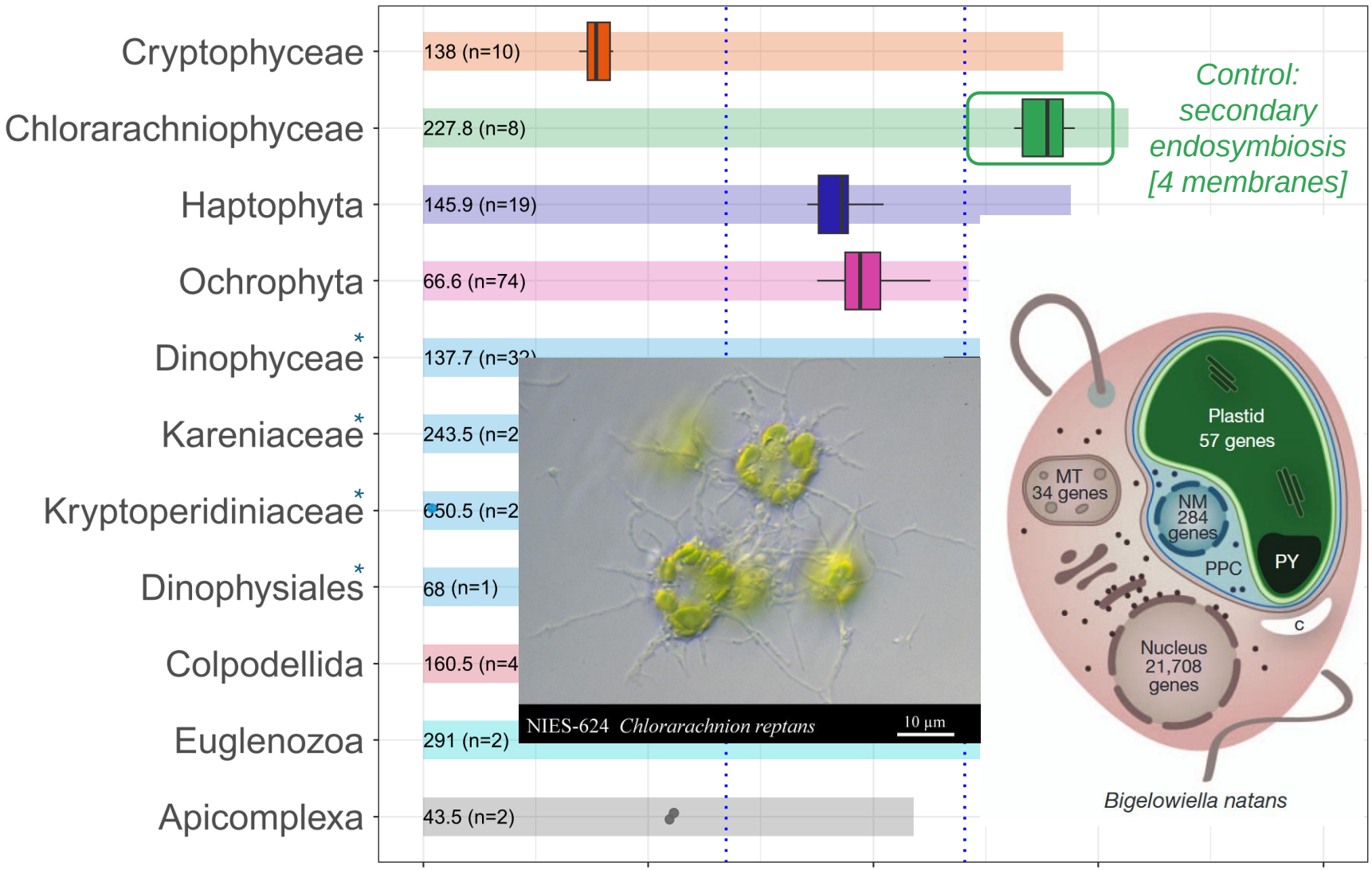


Waller & Kořený (2017)



*Data after reclassification

Prediction controls: endosymbiosis



*Data after reclassification

Membrane topologies

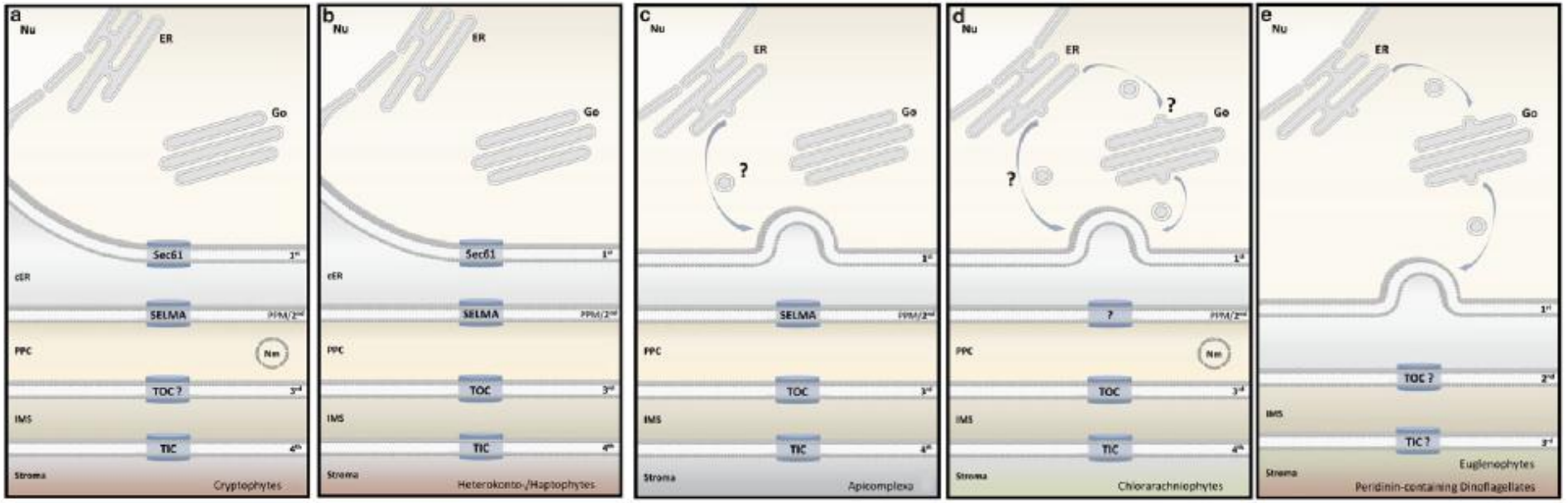
Cryptophyta
- 4 membranes
- nucleomorph

Haptophyta
Ochrophyta
- 4 membranes

Apicomplexa
Chromerida
- 4 membranes

Chlorarachniophyceae
- 4 membranes
- nucleomorph

Euglena
Dinophyceae
- 3 membranes



Guillardia theta

Vaucheria litorea
Emiliania huxleyi

Plasmodium falciparum

Bigelowiella natans

Euglena gracilis
Symbiodinium minutum

Grosche et al., 2014

	plastid protein import					ultrastructure				
	pre-sequence	vesicles	TIC	TOC	SELMA	#/cell	shape	thylakoids	pyrenoid	eyespot
Chloroplastida	+ (A)	-	✓	✓	-	1/2/+			-/✓	
Rhodophyta	+ (F)	-	✓	✓	-	1/+			(✓)	-
Cryptophyta	+ AXAF	-	red.	red.	✓	1/2			✓	
Haptophyta	+ F	-	red.	red.	✓	1/2			✓	
Ochrophyta	+ ASAFAP	PPC	red.	red.	✓	1/2/+			✓	
Chromerida	+ (ASA)F	ER	red.	red.	✓	1			✓	-
dinotoms	-	-	red.	red.	✓	+			✓	
fucoxanthin dinos	GRR	-	red.	red.	✓	+			✓	
peridinin dinos	+ F F	ER Golgi	red.	red.	-	1/+			(✓)	
Lepidodinium	+	PPC	red.	red.	-	1			✓	-
Chlorarachniophyta	+	ER	red.	red.	-	1/+			✓	-
Euglenophyta	+	ER Golgi	red.	- ?	?	1/+			✓	

	plastid genome					transcript editing	
	shape	size (kbp)	genes	% primary	introns	rubisco	subst.	poly-U
Chloroplastida	◯	100-200	90-130		+	(L)S	(✓)	-
Rhodophyta	◯	150-200	200-250		+	LS	-	-
Cryptophyta	◯	120-136	~180		+	LS	-	-
Haptophyta	◯	~100	~140		-	LS	-	-
Ochrophyta	◯	90-170	130-200		-	LS	-	-
Chromerida	∼	80-120	~110		-	type II	-	✓
dinotoms	◯	~120	~160		-	LS	-	-
fucoxanthin dinos	◯	~140	~100		+	LS	✓	✓
peridinin dinos	◦◦	~3/circle	~15		-	type II	✓	✓
MGD	◯	102	96		+	(L)S	✓	-
TGD	◯	71	94		+	(L)S	✓	-
Lepidodinium	◯	66	90		+	(L)S	✓	-
Chlorarachniophyta	◯	~70	~95		+	(L)S	-	-
Euglenophyta	◯	60-140	~90		+++	(L)S	-	-

	pigments						xantophyll cycle		pathways		
	chlorophylls	phycobilins	phytochr.	fucox.	vaucheriax.	peridinin	viola/zea	diadino/diato dino	ALA	haem	IPP
Chloroplastida	*a b*	-	✓	-	-	-	✓	-	C5		
Rhodophyta	*a*	✓	✓	(✓)	-	-	(✓)	-	C5		
Cryptophyta	*a* c_2	✓	✓	-	-	-	-	-	C5		
Haptophyta	*a* c_{123}	-	-	✓	-	-	-	✓	C5		
Ochrophyta	*a* c_{12}	-	✓	✓	✓	-	(✓)	✓	C5		
Chromerida	*a*	-	✓	✓	✓	-	✓	-	C4		
dinotoms	*a* c_{12}	-	✓?	✓	-	-	?	✓?	2×C5	?	
fucoxanthin dinos	*a* c_{123}	-	✓?	✓	-	-	-	✓	C5		
peridinin dinos	*a* c_2	-	-	-	-	✓	-	✓	C5	*	
Lepidodinium	*a b*	-	?	-	-	-	?	-	C5	?	
Chlorarachniophyta	*a b*	-	-	-	-	-	✓	-	C5 C4		
Euglenophyta	*a b*	-	-	-	-	-	✓	-	C5 C4		