bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

A Timescale for the Radiation of Photosynthetic Eukaryotes

Eliane Evanovich[1,2], João Farias Guerreiro[1]

Filiation:

[1]Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas - Universidade Federal do Pará, Pará, Brasil.

[2]Universidade Aberta do Brasil, Ciências Biológicas, Instituto de Ciências Biológicas - Universidade Federal do Pará, Pará, Brasil.

Email address for correspondence: E-mail: lianevanovich@gmail.com

Co-author email address: joao.guerreiro53@gmail.com

Mailing address: Laboratório de Genética Humana e Médica - Universidade Federal do Pará, Rua Augusto Corrêa, 1 - Guamá, Belém - PA, 66075-110.

ORCID ID: https://orcid.org/0000-0001-7076-5621

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

Oxygenic photosynthesis is considered the most important evolutionary innovation in the history of Earth. It depends on two photosystems, responsible for the photolysis of water and the reduction of carbon dioxide. Oxygen and carbohydrates are released at the end of the reaction. Extraordinary, the oxygen released created the stratospheric ozone layer, and transformed the ocean chemistry, whereas the carbohydrates are the primary source of energy for complex cells. Several lines of evidence indicate the photosynthesis arose in the ancestors of cyanobacteria. It was spread over some eukaryotes by the acquisition of a free-living cyanobacterium, which evolved into photosynthetic plastid, the chloroplast. The timing of the chloroplast emergence is still controversy. Estimated ages range from 600 to 2100 million years ago (Mya) in accordance to previous studies. The aim of this study is to clarify several aspects of the origin and diversification of photosynthetic eukaryotes. For this purpose, we utilized a data set based on 27 protein-coding genes from genomes of cyanobacteria and photosynthetic eukaryotes, more genes than other papers that also utilized plastid genes, and performed the Bayesian analysis method to estimate the divergence times of the photosynthetic eukaryotes. Results showed photosynthetic eukaryotes emerged Late Mesoproterozoic about 1342 Mya. The Early Proterozoic oceans did not have adequate conditions for eukaryotes, because chemical elements such as zinc and molybdenum were at reduced concentrations, and they are essential to the formation of eukaryotic proteins.

Keywords: Chloroplast-encoded genes, fossil record, molecular dating, molecular timescale, photosynthetic eukaryotes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

1. Background

Oxygenic photosynthesis is probably the most important metabolic innovation of the Earth history. It produces, at the end of the process, water, oxygen and carbohydrates. Remarkable, free oxygen created the stratospheric ozone layer, and changed the ocean chemistry, while the carbohydrates are the main source of energy for complex cells [1, 2, 3, 4]. Most living beings, save for some exceptions, depend directly or indirectly of the oxygenated photosynthesis.

In the Archean Eon, the terrestrial atmosphere and the oceans were anoxic. This began to change about 2500 million years ago (Mya) after the Great Oxygenation Event (GOE). Later, four more oxygenation events happened in the primitive Earth. In the Proterozoic from 1850 to 850 Mya, the oceans became slightly less anoxic and sulfidic [5, 6, 7, 8]. These environmental changes appear to have triggered an increase in Proterozoic biodiversity, even from the complex eukaryotes carrying specialized structures [6, 8].

Several studies suggest that the oxygenic photosynthesis arose in the ancestors of Cyanobacteria (blue-green algae), and it was spread over some eukaryotes, through the acquisition of a free-living cyanobacterium, which evolved into photosynthetic plastid, the chloroplast [1, 4, 9]. This event, known as primary endosymbiosis arose in Archaeplastida, a monophyletic group formed by Rhodophyta, Viridiplantae and Glaucophyta [10, 11]. In contrast to Archaeplastida, CASH organisms (Cryptophyceae, Alveolata Stramenopiles, and Haptophyta) are paraphyletic, however, share a red plastid, which was obtained by secondary endosymbiosis with a red alga. CASH plastids (or complex red plastids) have the same protein transport apparatus, and present four membranes [13, 14, 15, 16, 17, 18, 19]. Conforming to the hypothesis of a unique origin of the red plastids, the loss of plastids in groups as some alveolates (apicomplexans and dinoflagellates), *Goniomonas* (Cryptophyceae), and Stramenopiles (Bicosoecida, Labyrinthulomycetes, Oomycetes, and Opalinidae) would be a derived condition [20, 21]. In accordance with the taxonomic review for

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Eukarya [22], Stramenopiles, Alveolata and Rhizaria form the SAR clade, whereas Cryptophyceae and Haptophyta do not have established phylogenetic positions within the Diaphoretickes (Domain which comprises Cryptophyceae, Haptophyta, Archaeplastida, SAR, and Euglenozoa). Dinoflagellate plastid *Durinskia* and *Kryptoperidinium* likely were retrieved from diatoms, while the Chlorarachniophyceae plastid derived from a Trebouxiophyceae alga by tertiary endosymbiosis [23, 24].

Chloroplast DNA (cpDNA) has about one-twentieth the size of cyanobacterial genomes [25]. It is circular or linear and includes genes involved in transcriptional machinery (ribosomal RNA, transfer RNA, and RNA polymerase genes), photosynthesis, and biosynthesis of molecules as vitamins, amino acids, pigments, and fatty acids [25, 26]. The origin of the chloroplast is still a controversial issue. Previous studies assign an older aged between 1423 to 2100 Mya [27, 28, 29, 30, 31], while others have estimated a younger aged between 600 to 1250 Mya [32, 33, 34, 35]. Results found by Betts et al. [35] for the divergence time among primary endosymbiosis showed larger limits, between 1118 to 1774 Mya. Estimates of divergence times cited above were based on molecular data, except for the one proposed by Cavalier-Smith [33] which was based on fossil records and their geochemical impacts. In this paper, we performed relaxed Bayesian estimates based on 27 plastid genes to infer the possible date of origin and diversification of photosynthetic eukaryotes. Compared to dataset utilized by Yoon [28], we used more genes (21 more genes) and algae groups (Chlorarachniophyceae and Dinoflagellata), and found an age of 1342 Mya for the origin of photosynthetic eukaryotes.

2. Methods

2.1. Samples and Alignment

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Genome sequences utilized in this paper were available from the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov). Description and source of the samples are on additional file: Table S1. MUSCLE software [36] was utilized to perform sequence alignment of 27 protein-coding genes (additional file: Table S2). The concatenated sequence had 23240 nucleotides and was later submitted to Gblocks 0.91b [37], which selected the most conserved matrix data. Gblocks generated an alignment containing 11606 ungapped nucleotides, 49% of the original sequence.

2.2. Molecular Dating Analyses

GTR+I+G was the most appropriate model and was selected by Jmodeltest 2.1.10 [38]. RaxML 8 software [39] was applied to generate phylogenies. The best tree was utilized as an input tree on BEAST. Phylogenies obtained are in the additional file: Dataset S1. BEAST v2.6.0 [40] was performed to measure the divergence dates. The analysis was based on the following models: GTR+I+G, gamma-distributed heterogeneity of substitution rates among sites, uncorrelated log-normal relaxed clock (UCLN), and Yule. The computational performance was improved by BEAGLE library [41]. We run five independent Metropolis-coupled Markov chain Monte Carlo (MCMCMC) in BEAST. Each had a chain of 100 million generations and burn-in of 10 million. Tracer 1.7 [42] was utilized to monitor the analyzes and indicate when there was the convergence. At the end of the analysis, 100 trees were generated. TreeAnnotator [43] was applied to measure the consensus tree that was viewed with FigTree v1.4.4 [44]. The fossil calibration points are shown in Table 1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Table 1. Description of the fossils utilized as calibration points in this study.

Taxon	Fossil	Calibration point (Mya)	Type constraits	Reference
Cyanobacteria	Possible microfossils	3000	Maximum	45
Cyanobacteria	Biomarker evidence	1640	Minimum	46
Eukarya	*Grypania spiralis*	2100	Maximum	47
Eukarya	Possible protist microfossils	1500	Minimum	48
Rhodophyta: Bangiophyceae	*Bangiomorpha pubescens*	1047	Minimum	34, 49
Viridiplantae: Chlorophyta	*Proterocladus antiquus*	1000	Maximum	50
Stramenopiles:Xanthophyceae	*Palaeovaucheria*	1000	Maximum	51
Stramenopiles: Xanthophyceae	*Jacutianema*	750	Minimum	52
Rhodophyta: Florideophyceae	Oldest known Florideophyceae fossil	550	Minimum	48
Stramenopiles: Phaeophyceae	*Punctariopsis latifolia*	490	Minimum	53
Haptophyceae	First appearance of calcareous nannofossil	225	Minimum	54, 55
Viridiplantae: Zygnemophyceae	*Tetraporina horologia*	345	Maximum	56
Viridiplantae: Zygnemophyceae	*Paleoclosterium spiralis*	93	Minimum	57,58
Stramenopiles: Bacillariophyta	Oldest known diatom	185	Minimum	59
Stramenopiles: Bacillariophyta	Pennate diatoms	50	Minimum	60

3. Results and Discussion

3.1. Phylogenetic Relationships

Topology found in this paper was similar to that obtained by Janouškovec et al. [18], which also utilized chloroplast genes. Phylogenetic trees based on chloroplast genes have often failed to recover the Archaeplastida monophyly because they reflect more recent phylogenetic relationships containing algae arising from secondary and tertiary endosymbiosis. Figure 1 shows the time scale for the origin and diversification of the photosynthetic eukaryotes. Phylogenetic tree with the minimum and maximum limits is shown in the additional file: Figure S1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

3.2. Primary Plastid Endosymbiosis

Origin of the chloroplast was about 1342 Mya (95% highest posterior density or HPD: 1155-1848 Mya) according to our results, and it is closer to limits estimated by Betts et al. [35], and Yoon et al. [28]. Gibson et al. [34] rejected *Bangiomorpha* as representative of Bangiales-Rhodophyta. Our results do not reject the phylogenetic position of these fossils and other Proterozoic fossils such as *Proterocladus antiquus*, *Palaeovaucheria*, and *Jacutianema*.

It is noteworthy that the low oxygen levels in the Paleoproterozoic (2500 to 1600 Mya ago) were incompatible with the bioavailability of essential micronutrients for the enzymatic activity of eukaryotic cell [5, 6]. In addition to having low levels of micronutrients, the Paleoproterozoic oceans also had high sulfate concentrations [61, 62, 63]. The biogeochemical stasis previously experienced by the planet was finally broken in Mesoproterozoic [64], and abrupt innovations appeared on Earth for the first time during the Neoproterozoic (1000 to 541 Mya ago) as ice sheets, alteration in carbon and oxygen levels, and increase in primary productivity [64]. The emergence and diversification of eukaryotic algae may probably have had to play a decisive role in these transformations.

3.3. Glaucophyceae

In Archaeplastida, Glaucophyceae was the first lineage to emerge about 1322 Mya (95% HPD: 1103-1824 Mya), shortly after the emergence of photosynthetic eukaryotes. These algae retain primitive characters, similar to an ancestral cyanobacterium, such as a vestigial peptidoglycan wall, non-stacked thylakoid, and carboxysome-like bodies, therefore it is considered a living fossil [65]. It has little diversity, restricted to four genera (Cyanophora, Cyanoptyche, Glaucocystis, and Gloeochaete), which contrasts with the vast diversity found in Rhodophyta and Viridiplantae.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

3.4. Viridiplantae

The divergence between Streptophyta and Chlorophyta was estimated at 1108 to 1401 Mya (mean 1244 Mya), whereas the emergence of Chlorophyta was between 984 to 1203 Mya (mean 1093 Mya). This knowledge is consistent with a possible Chlorophyta fossil called *Proterocladus antiquus* dated 1000 Mya [50]. Radiation of Viridiplantae may have influenced the fluctuations in the carbon cycle and consequently would have helped to trigger the initiation of Snowball Earth [64].

3.5. Rhodophyta

The age estimated for the Rhodophyta radiation was 1127 Mya (95% HPD: 1061-1242 Mya), which is consistent with the age of 1047 Myr[34]. In according to Cavalier-Smith [64], *Bangiomorpha pubescens* would be a bacterial complex, whereas Berney and Pawlowski [33] and Gibson et al. [34] suggest that *Bangiomorpha* might be a stem photosynthetic eukaryotic extinct. The large size, sexual reproduction, multicellularity, and radial cell division are the traits that supported *Bangiomorpha* as a member of the Rhodophyceae [49].

3.6. Secondary Plastid Endosymbiosis

Secondary endosymbiosis at least three types of plastids, found in Euglenozoa, Chlorarachniophyceae, and CASH organisms [66, 67]. Results show the engulfment of Pyramimonadales algae by Euglenozoa ancestor occurred about 366 Mya in the Upper Devonian. Oldest known fossil of Euglenozoa dates from Lower Cretaceous [68, 69]. Chlorarachniophyte plastids were originated from a Trebouxiophyceae alga [24] during the Carboniferous (344 Mya).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Our estimates indicate the red plastid emerged about 1156 Mya (95% HPD: 1070-1313) during the Neoproterozoic. Stramenopiles arose between 1044-1182 Mya (mean 1017 Mya). The date is within the range estimated by Berney and Pawlowski [33]. And Coccolithophores (Haptophyta) would have appeared about 538 to 570 Mya (mean 554 Mya). These algae have a fundamental role in the production and precipitation of calcium carbonate (CaCO3) and besides influence the light scattering and cloud cover from the Jurassic [70].

3.7. Tertiary Plastid Endosymbiosis

Some Peridiniales plastids were obtained from red algae, however, none of them is represented in this paper. Our analysis was restricted to *Durinskia* and *Kryptoperidinium*. Their plastids were acquired from the engulfment of a diatom between 64.5 to 194 234 Mya (mean 109 Mya). The most reliable fossil record of these algae appear since Triassic, had an increase in Cenozoic Era, and is currently in decline [71].

4. Conclusion

The biogeochemical stasis of the Earth lasted about 3 billion years, and it was broken during the Late Mesoproterozoic. Earlier the oceans were inhospitable, in other words, highly sulfidic and poorly oxygenated. Sulfides are toxic to eukaryotic algae, and the low levels of oxygen and micronutrients inhibit the formation of proteins in these organisms. Our results indicate that photosynthetic eukaryotes arise around 1048 Mya, at the end of Mesoproterozoic, and diversified rapidly. The radiation of photosynthetic eukaryotes helped progressively change the chemical composition of the oceans, increased the primary productivity and favoured the emergence of new life forms. Early eukaryotes may have had a wide variety of organisms, such as *Bangiomorpha* and

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Paleovaucheria. Many of them have disappeared or have a yet unknown phylogenetic affinity, similar to what happened too many representatives of the Ediacaran biota.

Abbreviattions

$CaCO_3$ - calcium carbonate

cpDNA - Chloroplast DNA

GOE - Great oxygenation event

HPD – 95% Highest posterior density

LECA - last eukaryotic common ancestor

MCMCMC - Metropolis-coupled Markov chain Monte Carlo

MRCA - most recent common ancestor

Mya - million years ago

NCBI - National Center for Biotechnology Information

PRONE - plant-specific Rop nucleotide exchanger

Re-Os - Rhenium-Osmium

rRNA - ribosomal RNA

tRNA - transfer RNA

Data Availability

Supporting data are enclosed as additional files.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Additional Points

This manuscript is part of the Ph.D. thesis of Eliane Evanovich from Programa de Pós-graduação em Genética e Biologia Molecular, Universidade Federal do Pará. The authors wish to express their thanks for the financial support of Pró-Reitoria de Pesquisa e Pós-Graduação (PROPESP).

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' contributions

The authors read and approved the final manuscript.

References

1. N. Rassmussen, I. R. Fletcher, J. J. Brocks et al., "Reassessing the first appeareance of eukaryotes and cyanobacteria", *Nature*, vol. 455, n. 7216, pp. 1101–1104, 2008.

2. Margulis, Lynn and Chapman, Michael J. *Kingdoms and Domains: An Illustrated Guide to the Phyla of Life on Earth,* Philadelphia: Academic Press, 2010.

3. S. A. Kulasooriya, "Cyanobacteria: Pioneers of Planet Earth", *Ceylon Journal of Science*, vol. 40, n. 2, pp. 71-88, 2011.

4. M. F. Hohmann-Marriott and R. E. Blankenship, "Evolution of photosynthesis", *Annual Review of Plant Biology*, vol. 62, pp. 515-548, 2011.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

5. H. D. Holland, "The oxygenation of the atmosphere and oceans", *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 361, n. 1470, pp. 903-15, 2006.

6. N. J. Planavsky, L. J. Tarhan, E. J. Bellefroid et al., "Late Proterozoic transitions in climate, oxygen, and tectonics, and the rise of complex life", *The Paleontological Society Papers*, vol. 21, pp. 2015.

7. M. A. Saito, D. M Sigman and F. M. M. Morel, "The bioinorganic chemistry of the ancient ocean: the co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean–Proterozoic boundary?", *Inorganica Chimica Acta*, vol. 356, pp. 308–318, 2003.

8. P. A. Cohen PA and F. A. Macdonald, "The Proterozoic Record of Eukaryotes", *Paleobiology*, vol. 41, n. 4, pp. 610–632, 2015.

9. W. Martin and K. V. Kowallik, "Annonated English translation of Mereschkowski 1905 paper", *European Journal of Phycology*, vol. 34, n. 3, pp. 287–295, 1999.

10. S. M. Adl SM, A. G. B, Simpson, M. A. Farmer et al., "The new higher level classification of eukaryotes with emphasis on the taxonomy of protists", *Journal of Eukaryotic Microbiology*, vol. 52, n. 5, pp. 399–451, 2005.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

11. T. Cavalier-Smith, "The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa", *International Journal of Systematic and Evolutionary Microbiology*, vol. 52, n. 2, pp. 297–354, 2002.

12. J. de Vries and J. M. Archibald, "Plastid genomes", *Current Biology*, vol. 28, pp. R336-R337, 2018.

13. F. Burki, K. Shalchian-Tabrizi, J. Pawlowski, “Phylogenomics reveals a new ‘megagroup’ including most photosynthetic eukaryotes,” *Biology Letters*, vol. 4, pp. 366–369, 2008.

14. M. S. Sommer, S. B. Gould, P. Lehmann, A. Gruber, J. M. Przyborski, U. G. Maier, "Der1-mediated pre-protein import into the periplastid compartment of chromalveolates?” , *Molecular Biology and Evolution*, vol. 24, pp. 918-928, 2007.

15. K. Bolte, L. Bullmann, F. Hempel, A. Bozarth, S. Zauner, U. G. Maie, "Protein targeting into secondary plastids", *Journal of Eukaryotic Microbiology*, vol. 56, pp. 9-15, 2009.

16. M. H Smith, H. L. Ploegh, J. S Weissman, "Road to ruin: targeting proteins for degradation in the endoplasmic reticulum", *Science*, vol. 334, pp. 1086-1090, 2009.

17. Janouškovec, A. Horák, M. Oborník, J. Lukeš and P. J. Keeling, "A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, n. 24, pp. 10949-10954, 2010.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

18. D. Baurain, H. Brinkmann, J. Petersen, N. Rodríguez-Ezpeleta, et al., “Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles”, *Molecular Biology and Evolution, vol.* 27, pp. 1698–709, 2010.

19. J. Petersen, A. K. Ludewig AK, V. Michael, B. Bunk, et al., "*Chromera velia*, endosymbioses and the rhodoplex hypothesis–plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages)", *Genome Biology and Evolution*, vol. 6:666-684, 2014.

20. P. J. Keeling, "Chromalveolates and the evolution of plastids by secondary endosymbiosis", *Journal of Eukaryotic Microbiology*, vol. 56, n. 1, pp. 1-8, 2009.

21. N. Arisue and T. Hashimoto, "Phylogeny and evolution of apicoplasts and apicomplexan parasites", *Parasitology International*, vol. 64, n. 3, pp. 254-259, 2015.

22. S. M. Adl, D. Bass, C. E. Lane, J. Lukeš *et al*., "Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes", *Journal of Eukaryotic Microbiology*, vol. 66, n. 1, pp. 4-119, 2019.

23. B. Imanian, J. F. Pombert, P.J. Keeling, “The complete plastid genomes of the two 'dinotoms' *Durinskia baltica* and *Kryptoperidinium foliaceum*”, *PLoS One*, vol. 5, n. 5, p. e10711, 2010.

24. M. B. Rogers, P. R. Gilson, V. Su, G. I. McFadden, P. J. Keeling, “The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts”, *Molecular Biology and Evolution*, vol. 24, n. 1, pp. 54-62, 2007.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

25. B. Green, "Chloroplast genomes of photosynthetic eukaryotes", *The Plant Journal*, vol. 66, n. 1, pp. 34-44, 2011.

26. Y. Jiao and H. Guo, "Chloroplast genome evolution in land plants", *Advances in Botanical Research: Genomes of Herbaceous Land Plants*, Vol. 69, n. 3-5, p. 226, 2014.

27. S. B. Hedges, J. E. Blair, M. L. Venturi and J. L. Shoe, "A molecular timescale of eukaryote evolution and the rise of complex multicellular life", *BMC Evolutionary Biology*, vol. 4, n. 4, pp. 1-9, 2004.

28. H. S. Yoon, J. D. Hackett, C. Ciniglia et al., "A molecular timeline for the origin of photosynthetic eukaryotes", *Molecular Biology and Evolution*, vol. 21, n. 5, pp. 809-818, 2004.

29. E. J. Douzery, E. A. Snell, E. Bapteste, "The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils?" *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, n. 45, pp. 15386-15391, 2004.

30. L. W. Parfrey, D. J. Lahr, A. H. Knoll and L. A. Katz, "Estimating the timing of early eukaryotic diversification with multigene molecular clocks", *Proceedings of the National Academy of Sciences*, vol. 108, n. 33, pp. 13624-13629, 2011.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

31. P. Sánchez-Baracaldo, J. A. Raven, D. Pisani and A. H. Knoll, “Early photosynthetic eukaryotes inhabited low-salinity habitats, *Proceedings of the National Academy of Sciences*, vol. 114, n.37, pp. E7737-E7745, 2017.

32. T. Cavalier-Smith, "Deep phylogeny, ancestral groups and the four ages of life", *Philosophical Transactions of the Royal Society*, vol. 365, n. 1537, pp. 111-132, 2010.

33. C. Berney and J. Pawlowski, "A molecular time-scale for eukaryote evolution recalibrated with the continuous microfossil record", *Philosophical Transactions of the Royal Society*, vol. 27, n. 1596, pp. 31867-31872, 2006.

34. T. M. Gibson, M. Timothy, P. M. Shih et al., "Precise age of *Bangiomorpha pubescens* dates the origin of eukaryotic photosynthesis", *Geology*, vol. 46, n. 2, pp. 135–138, 2017.

35. H. C. Betts, M. N. Puttick, J. W. Clark, et al., "Integrated genomic and fossil evidence illuminates life’s early evolution and eukaryote origin", *Nature Ecology and Evolution*, vol. 2, n. 10, pp. 1556-1562, 2018.

36. R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput", *Nucleic Acids Research*, vol. 32, n. 5, pp. 1792-1797, 2004.

37. G. Talavera and J. Castresana, "Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments", *Systematic Biology*, vol. 56, n. 4, pp. 564-577, 2007.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

38. D. Darriba, G. L. Taboada, R. Doallo, D. Posada, "jModelTest 2: more models, new heuristics and parallel computing", *Nature Methods*, vol. 9, n. 8, pp. 772, 2012.

39. A. Stamatakis, "RAxML Version 8: A tool for Phylogenetic Analysis and post-analysis of large phylogenies", *Bioinformatics*, vol. 30, n. 9, pp. 1312–1313, 2014.

40. R. Bouckaert, T. G. Vaughan, J. Barido-Sottani et al., "BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis", *PLoS computational biology*, vol. 15, n. 4, article e1006650, 2019.

42. M. A. Suchard and A. Rambaut, "Many-core algorithms for statistical phylogenetics", *Bioinformatics*, vol. 25, n. 11, pp. 1370–1376, 2009.

43. A. Rambaut, A. J. Drummond, D. Xie, G. Baele and M. A. Suchard, "Posterior summarisation in Bayesian phylogenetics using Tracer 1.7", *Systematic Biology*, vol. 67, n. 5, pp. 901-904. 2018.

44. A. Rambaut, *FigTree v1.4.1. Institute of Evolutionary Biology*, University of Edinburgh, Edinburgh. http://tree.bio.ed.ac.uk/software/figtree/

45. S. A. Crowe, N. L. Døssing, N.J. Beukes, M. Bau, S.J. Kruger, R. Frei, D.E. Canfield, “Atmospheric oxygenation three billion years ago”, *Nature,* vol. 501: 535–538, 2013.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

46. J. J. Brocks, G. D. Love, R. E. Summons et al., "Biomarker Evidence for Green and Purple Sulfur Bacteria in a stratified Paleoproterozoic Sea", *Nature*, vol. 437, n. 7060, pp. 866-870, 2005.

47. T. M. Han and B. Runnegar, "Megascopic eucaryotic algae from the 2.1-billion-year-old Negaunee Iron Formation, Michigan" *Science*, vol. 257, n. 5067, pp. 232-235, 1992.

48. E. J. Javaux, A. H. Knoll and M. R. Walter, "Morphology and ecological complexity in early eukaryotic ecosystems", *Nature*, vol. 412, n. 6842, pp. 66-69, 2001.

49. J. Butterfield, "Bangiomorpha pubescens n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes", *Paleobiology*, vol. 26, n. 3, pp. 386–404, 2000.

50. Q. Tang, K. Pang, X. Yuan, S. Xiao, "A one-billion-year-old multicellular chlorophyte", *Nature Ecology and Evolution*, https://doi.org/10.1038/s41559-020-1122-9, 2020.

51. K. N. Woods, A.H Knoll, and T. German, “Xanthophyte algae from the Mesoproterozoic/Neoproterozoic transition: confirmation and evolutionary implications”, *Geological Society of America Bulletin*, 30, 232, 1998.

52. N. J. Butterfield, “A vaucheriacean alga from the middle Neoproterozoic of Spitsbergen: implications for the evolution of Proterozoic eukaryotes and the Cambrian explosion”, *Paleobiology,* 30, 231–252, 2004.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

53. Xu , Z. L. New discoveries of phaeophycean fossils in the Early Cambrian, Haikou, Kunming, Yunnan, Southwest China" *Acta Botanica Sinica,* vol. 43 : 1072 – 1076, 2001.

54. D. Janofske, "Calcareous nannofossils of the Alpine Upper Triassic", *Nannoplankton Research*, vol. 1, pp. 87–109, 1992.

55. Bown, Paul. Calcareous nannofossil biostratigraphy. Cambridge: Kluver Acad. Publish, 1998.

56. Mullins GL, Servais T (2008) The diversity of the Carboniferous phytoplankton. Rev Palaeobot Palynol 149:29–49.

57. H. Dörfelt, and U. Schäfer, *Palaeozygnema spiralis* , ein Vertreter der Conjugatophyceae in mesozoischem Bernstein aus Bayern . Hoppea, *Denkschriften der Regensburgischen Botanischen Gesellschaft,* 61 : 785–793, 2000.

58. A. R. Schmidt, H. von Eynatten, M. Wagreich, M. Wagreich, "The Mesozoic amber of Schliersee (southern Germany) is Cretaceous in age", *Cretaceous Research*, vol. 22,423-428.

59 Barron John A. Diatoms, Fossil Prokaryotes and Protists, Boston: Blackwell Scientific Publications, 1993.

60. Strelnikova, Nina. Evolution of the diatoms during the Cretaceous and Paleogene periods. Königstein: Költz Scientific Books, 1990.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

61. A. D. Anbar, A. H. Knoll, "Proterozoic Ocean Chemistry and Evolution: A Bioinorganic Bridge?", *Science*, vol. 297, n. 5584, pp. 1137-1142, 2002.

62. D. T. Johnston, F. Wolfe-Simon, A. Pearson and A. H. Knoll, "Anoxygenic photosynthesis modulated Proterozoic oxygen and sustained Earth's middle age", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, n. 40, pp. 16925-16929, 2009.

63. C. L. A. Dupont, R. E. Butcher, P.E. Valas et al., "History of biological metal utilization inferred through phylogenomic analysis of protein structures", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, n. 23, pp. 10567-10572, 2010.

64. E. Tziperman, I. Halevy, D. T. Johnston, A. H. Knoll, and D. P. Schrag, "Biologically induced initiation of Neoproterozoic snowball-Earth events", *Proceedings of the National academy of Sciences*, vol. 108, n. 38, pp. 15091–15096, 2011.

65. T. Cavalier-Smith, "Principles of protein and lipid targeting in secondary symbiogenesis: Euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree", Journal of Eukaryotic Microbiology, vol. 46, n. 4, pp. 347–66, 1999.

66. M. Turmel, M. C. Gagnon, C. J. O'Kelly, C. Otis and C. Lemieux, "The chloroplast genomes of the green algae *Pyramimonas, Monomastix,* and *Pycnococcus* shed new light on the evolutionary history of prasinophytes and the origin of the secondary chloroplasts of euglenids", *Molecular Biology and Evolution*, vol. 26, n. 3, pp. 631–648, 2009.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.18.047969; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

67. C. Ascaso, J. Wierzchos, M. Speranza, et al., "Fossil protists and fungi in amber and rock substrates", *Micropaleontology*, vol. 51, n.1, pp. 59–72, 2005.

68. T. Tyrrell and J. R. Young, “Coccolithophores”. Encyclopedia of Ocean Sciences, pp. 606-614, 2009.

69. Falkowski, Paul, G. II, Knoll, Andrew H, Evolution of Primary Producers in the Sea. Burlington, MA: Elsevier Academic Press, 2007.

70. R. A. Fensome, J. F. Saldarriaga and M. F. J. Taylor, "Dinoflagellate phylogeny revisited: reconciling morphological and molecular based phylogenies", *Grana*, vol. 38, n. 2-3, pp. 66-80, 1999.

71. R. A. MacRae, R. A. Fensome and G. L. Williams, "Fossil dinoflagellate diversity, originations, and extinctions and their significance", *Canadian Journal of Botany*, vol, 74 n. 11, pp. 1687-1694, 1996.

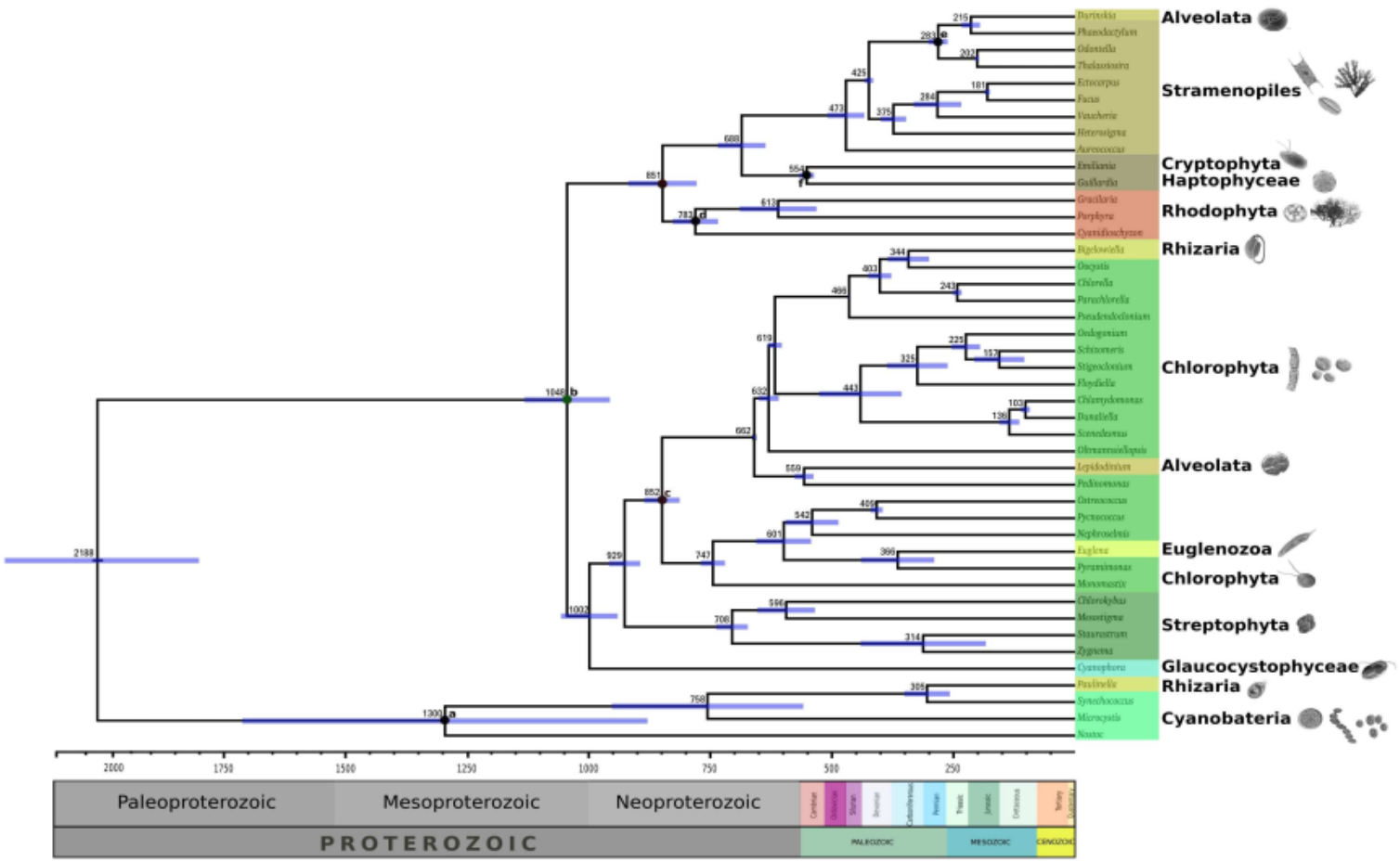

Figure 1: A time scale for the origin and diversification of the photosynthetic eukaryotes obtained by BEAST 2 based on 27 chloroplast-encoded genes. The estimated mean divergence times are shown at the nodes of the tree, and the blue bars at nodes represent the 95% confidence interval. Calibration nodes are indicated by the circles. Dark circles correspond to nodes with two calibration points, and light circles are nodes with one calibration point. Above of the tree is presenting the geological time.